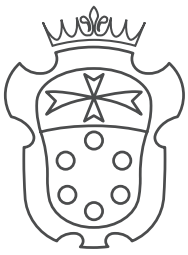

SCUOLA
NORMALE
SUPERIORE

Classe di Scienze

Corso di perfezionamento in
Fisica

XXXV ciclo

Black Holes through the lenses of Effective Field Theory

Settore Scientifico Disciplinare **FIS/02**

Candidato
dr. Francesco Serra

Relatore

Prof. Enrico Trincherini

Supervisione interna

Prof. Guilherme L. Pimentel

Anno accademico 2022–2023

To my family, cousins included

ABSTRACT

The detection of gravitational waves emitted by black hole binaries opens a window to test the theory of General Relativity and to probe the dynamics of gravity in regimes otherwise inaccessible. To fully exploit this opportunity, it is necessary to understand which are the consistent and detectable deviations from General Relativity. To single out these possible deviations, one must also be able to recognize effects due to environmental perturbations of the binary system, as these may lead to departures within General Relativity from the waveform of an ideal isolated binary. In this thesis, we approach some aspects of both these theoretical challenges from the point of view of effective field theory. First, we study how to narrow down the space of theories that can describe detectable deviations from General Relativity based on consistency with the fundamental principles of our description of nature. We consider the simple example of General Relativity modified by the presence of a shift-symmetric scalar field coupled only to gravity. In this context, we first show that only a specific scalar-graviton interaction can lead to black holes different from what General Relativity predicts while being consistently included in an effective field theory description. Then we study how causality, unitarity and locality constrain this interaction. We show that if this interaction is strong enough to leave an imprint detectable with the next gravitational wave interferometers, then causality would require new degrees of freedom to appear at very low energies. In the second part of this work, we consider one example of environmental perturbation that can affect black hole mergers: a distant third body orbiting the black hole binary. To study efficiently such a system, we derive a worldline effective action describing the relativistic effects due to the third body on timescales much longer than the orbital periods. Using techniques from non-relativistic General Relativity, we obtain a description of the two orbits as two interacting particles endowed with multipole moments. We carry out these computations up to quadrupole order in the three-body interaction, including the leading relativistic corrections. This approach allows to study novel long timescale effects that can enhance the rate of orbital flips of the inner binary.

PUBLICATIONS

Paolo Creminelli, Nicolas Loayza, Francesco Serra, Enrico Trincherini, Leonardo G. Trombetta. *Hairy Black-holes in Shift-symmetric Theories* . JHEP, 2020, 08, 045

Adrien Kuntz, Francesco Serra, Enrico Trincherini. *Effective two-body approach to the hierarchical three-body problem: Quadrupole to 1PN* . Phys. Rev. D, 2023, 107, 044011

Francesco Serra, Javi Serra, Enrico Trincherini, Leonardo G. Trombetta. *Causality constraints on black holes beyond GR* . JHEP, 2022, 08, 157

Adrien Kuntz, Francesco Serra, Enrico Trincherini. *Effective two-body approach to the hierarchical three-body problem: Quadrupole to 1PN* . Phys. Rev. D, 2023, 107, 044011

Carl Beadle, Giulia Isabella, Davide Perrone, Sergio Ricossa, Francesco Riva, Francesco Serra. *Dispersion relations without forward limits* . In preparation

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

INTRODUCTION

The formulation of Einstein's theory of General Relativity (GR) in the last century marked a paradigm shift in our understanding of gravity, introducing dynamic degrees of freedom to describe the gravitational interaction. This description of gravity revolutionized our comprehension of nature, establishing the dynamic role of spacetime and making possible to draw important analogies between gravity and other areas of particle physics. In this thesis we explore how these analogies can be leveraged to study the phenomenology of dynamical gravity in GR and beyond.

Among the several features predicted within the dynamical paradigm brought by GR, one of the most unique ones is the propagation of energy and momentum in the form of gravitational waves. Gravitational waves are dynamic disturbances in the geometry of spacetime, inducing repeated contractions and expansions of distances between different points. These are generally emitted when a source of gravitational field is accelerated.

Through immense scientific endeavors spanning decades, today the scientific community has developed the ability to detect gravitational waves. This by precision interferometers which measure microscopic variations in the path lengths traveled by light beams. Due to the weak nature of gravitational interactions, current detection capabilities are limited to capturing gravitational waves originating from exceptional events, such as the collision of stellar mass objects at least as compact as neutron stars.
Such collisions usually take place between objects that form a gravitationally bound binary system. In these cases the energy loss in gravitational waves and possibly internal friction determines a progressive shrinking of the binary orbit until the collision is unavoidable and the objects merge together. The more the objects are compact, the more the orbit will shrink before the collision, determining larger accelerations and a greater emission of gravitational waves. In particular, if the bodies are as compact as neutron stars, then the merger becomes a relativistic event that fully display the non-linear nature of gravitational interaction at length-scales comparable with the size of the components of the binary, typically tens of kilometers. As of today, the LIGO-Virgo-Kagra collaboration has detected 90 such events in a radius of 135 Mpc [1–3].

These outstanding results follow a rich history of experimental probes of the phenomenology predicted by GR in the past century. Notable examples include the cosmological evolution [4, 5], gravitational time delay [6, 7], gravitational lensing of light rays [8] and various effects in planetary as well as lunar and stellar motion [9–12]. These observations allowed to test GR in special configurations in which the sources of gravitational field

are highly symmetric or evolve negligibly due to the gravitational fields being weak.

Even when compared to these milestones, the detection of gravitational waves represents an unprecedented opportunity to explore gravity.
Indeed, at the most fundamental level gravitational waves allow to study directly the degrees of freedom that characterize gravity, encoded in the graviton. To date, the most striking result in this direction was the deduction of a bound on the speed at which gravitational waves propagate, thanks to the simultaneous detection of gravitational and electromagnetic signatures of a merger between two neutron stars, the GW170817 - GRB 170817A event [13, 14].
In addition to probing these fundamental aspects, detecting gravitational waves emitted by mergers of compact objects allows to test GR in a regime in which the dynamics of the gravitational sources is affected non-perturbatively by gravitational interactions. With no evident symmetry simplifying the dynamics, this setup probes the full complexity of Einstein's equations.
Similarly, gravitational waves produced by binary mergers give the unique opportunity of probing the nature of the most compact stellar-sized objects in the universe, making possible to study how matter behaves under extreme conditions.

Besides requiring sophisticated technology, detecting gravitational waves demands producing templates of the possible waveforms in order to single out the signal against the multiple sources of noise. This makes extremely important to study the dynamics of mergers across the binary phase-space. In this regard the complexity displayed by mergers of compact objects represents both an opportunity and a challenge, as it allows to probe a new regime of gravity but poses an obstacle to predicting the outcome of these extreme events. To handle this difficulty, various analytical techniques have been developed to study the early and late phases of the merger, the so called inspiral and ringdown phases, see e.g. [15, 16]. Even so, all known approaches bring little help in describing the instants surrounding the collision itself, when the components of the binary are very close to each other and reach relativistic speed. Currently, this phase of the merger can only be studied at the cost of running heavy numerical simulations [17–20].

Despite this technical challenge, it is still possible to draw some important and sharp predictions about the physics that governs the entire merger by studying analytically the inspiral and ringdown phases. Qualitatively, the inspiral phase describes the orbital motion of the bodies and the shrinking of the orbit, which can be characterized in terms of the orbital parameters as well as the deformability of the bodies and their interactions. Studying the dependence of the waveform on these details and comparing it with data, one can infer quantities like the mass and angular momentum of the two components of the binary. The ringdown phase of the merger, in which the final product of the collision reaches its stationary state, can be described efficiently in terms of perturbations of the final object itself. In other words, the gravitational radiation emitted only a short time after the collision can be described purely in terms of the stable, non-dynamical state that is eventually reached.

When it comes to compact objects in astrophysics, GR's prediction are impressively sharp. Infact, one of the most popularized achievements of this theory is to precisely describe the exterior of a black hole, that is a body which is so compact to have an escape velocity as high as the speed of light [21]. The boundary at which the escape velocity reaches the speed of light is called the black hole event horizon. Not only GR describes these objects precisely, but predicts their uniqueness up to three parameters: the mass, the angular momentum, and the electric charge of the black hole. This statement of uniqueness is the content of the so-called no-hair theorems[1] [22–28]. Besides describing these objects as solutions, GR predicts their formation as a result of gravitational collapse [29, 30]. This makes black holes, together with neutron stars, the natural candidates for stellar mass compact objects responsible for gravitational waves emitted during mergers.

Classically no excitation or particle traveling subluminally or at the speed of light can escape from the event horizon. For this reason, gravitational waves propagating away from the location of a black hole must respect boundary conditions that are infalling at the horizon. This translates into the fact that the perturbations of the gravitational field of a black hole that travel away from the horizon are exponentially damped oscillations with a discrete frequency spectrum, the so-called quasi-normal modes [31]. This distinguishing feature of black holes opens the possibility of performing a spectroscopic analysis of the ringdown radiation, in principle allowing for a direct test of the nature of compact objects [32].

The current precision of gravitational wave detection allows to measure the real and imaginary part of the longest-lived quasi-normal mode frequency with a precision of few ten per cents [33]. This allow to compare total mass and angular momentum of the final black hole with what is predicted numerically given the inspiral signal. The higher accuracy expected in the next generation of gravitational wave detectors, e.g. LISA [34–36], Einstein Telescope [37, 38], Cosmic Explorer [39], makes black hole spectroscopy a concrete opportunity to test the nature of compact objects in GR.

While the ringdown radiation is suppressed exponentially as time passes and can only be detected for fractions of a second, the radiation emitted during the inspiral phase is detectable for a much longer time. Currently, the gravitational waves are detected starting tens of seconds before the collision. However in the future, it will become possible to trigger a detection several months before the collision [40]. On one hand, this feature of the inspiral radiation allows to extract a large amount of information about the system, possibly allowing to probe gravitational dynamics with extreme precision. On the other hand, over the long duration of this phase feeble perturbations can pile up and bring non-negligible changes in the emitted waveform. For this reason, the inspiral phase can be affected more strongly by environmental effects. This makes the inspiral a less direct way to probe GR with respect to the ringdown. At the same time however it gives access to information about the surroundings of the binary, which can in principle allow to study interesting physical phenomena that may take place e.g. in the accretion disks of black holes.

1 Likely named after a pictorial comparison between the would-be extra features of black holes and unconventional hairdos of people

Therefore, analyzing the inspiral phase makes both interesting and necessary to parametrize systematically how the environment surrounding the binary affects its dynamics, be it gas, other astrophysical bodies, or other sources of perturbations.

Beyond the next generation gravitational wave interferometers such as LISA, the Einstein Telescope and Cosmic Explorer, which improve on the designs of the current detectors, other ideas are being explored. One example is given by pulsar timing arrays (PTA) [41–43], which allow to study the correlation in the signals received by an ensemble of pulsars to deduce whether a gravitational wave passed through the array, introducing time delays between the various signals. Given their size spanning galactic scales, these networks are sensitive to gravitational waves with exceptionally low frequencies, such as those that may be produced by mergers of supermassive black holes as heavy as 10^8 solar masses [44, 45]. More recent proposals with novel setups involve also microwave cavities [46, 47] as well as precision astrometry [48]. Besides the detection of gravitational waves, an other revolutionary development in the observation of gravitational phenomena is the imaging of accretion disks surrounding supermassive compact objects, by the Event Horizon Telescope [49, 50]. Such objects are found at the center of galaxies, and today have been photographed in the Milky Way as well as the M87 galaxy. This source of data allows to probe the strongly non-linear regime of gravity through the geodesics that are close to the innermost stable circular orbit (ISCO) [49].

These impressive progresses usher in a new era of exploration of the dynamics of gravity, allowing us to probe GR in regimes never accessed before. This makes possible not only to better understand the phenomenology of GR, but also to test whether nature deviates from GR predictions. In other words, gravitational wave detection gives the remarkable opportunity to test physics beyond GR (BGR) in the full non-linear regime at the scales that characterize binary mergers, that is the kilometer.
From the theoretical point of view, this possibility is particularly interesting as it might offer new perspectives on open questions in cosmology [51–55], such as modeling Dark Energy, Dark Matter or the early-time cosmological dynamics.

This opportunity comes with challenges. From the point of view of model-building, BGR physics that might be tested in the future must satisfy important constraints. From the outset, a model of BGR physics should not only predict deviations from GR but also be consistent with GR predictions that have been tested so far, within the current precision of the experiments. Much less trivially, the theories that one considers should be compatible with our theoretical understanding of how interactions work at the length-scales that characterize the new physics.
Moreover, from the point of view of fitting data, testing new physics makes all the more crucial understanding and identifying environmental effects as well as finite size effects (e.g. tidal deformability of neutron stars), since effects beyond GR might be similar to these departures from the ideal picture of an isolated black hole binary in GR. The vast size of parameter space of the possible perturbations and finite size effects make necessary to develop analytic tools to describe and classify these contributions.

Among the various theoretical tools and ideas developed in the context of high energy particle physics, effective field theory (EFT) stands out as the most appropriate one to address these challenges. Generally speaking, an EFT is a way to describe a given physical system in a chosen range of energies by singling out the degrees of freedom and interactions that are relevant for the dynamics. This is achieved by exploiting the separation of scales displayed by most physical systems, deriving power counting rules to compute e.g. the corrections due to microscopic, high energy (UV) dynamics on macroscopic, low energy (IR) observables. In this way, EFTs offer a highly versatile framework to understand and encode the structure governing the dynamics of a physical system across different scales.
For instance, an EFT can be very helpful in model building as it allows to parametrize efficiently models having similar degrees of freedom and symmetries at low energies, but possibly very different underlying microscopic UV descriptions. This bottom-up approach has been followed for instance in studying condensed matter systems, [56–59], as well as theories beyond the Standard Model of particle physics, see e.g. [60–62] and more recent works [63–67]. In the case of BGR physics, the bottom-up EFT approach simplifies the task of understanding whether a given interaction can fit with our current knowledge of gravitational phenomenology and whether it can be tested in future gravitational wave detections. Moreover, as we will discuss at length, this approach makes it simple to contrast a given model of BGR physics with the knowledge of how different interactions should be interconnected according to perturbation theory.
Similarly, given a system with a known microscopic dynamics, the EFT approach can be fruitfully used to derive a simplified description of its macroscopic, low-energy dynamics with no need of including the microscopic, high-energy degrees of freedom. Examples of this kind range from large-scale structure in the universe [68] to aspects of the Standard Model of particle physics, e.g. the weak force [69]. This top-down approach can also help simplifying the description of binaries, as proposed in the so-called Non Relativistic General Relativity (NRGR) method [70, 71]. Thanks to its versatility, this approach can be used to include finite size effects and is suitable to describe the effects due to environmental perturbations piling up over the inspiral phase.
With this in mind, in the following we will leverage on the insights gained through EFT in the context of particle physics and employ EFT methods to study compact objects and their dynamics in GR and beyond.

In particular, in the first part of this thesis we will use the bottom-up EFT approach to explore the space of theories that modify GR in a way that might be testable in the coming future. This approach allows to study the possible BGR theories in a way that is as much agnostic as reasonably possible with respect to the UV processes that govern physics on microscopic scales. From this point of view, we will analyze which of the models describing interesting gravitational dynamics are compatible with the rest of our knowledge and experience of the physical laws.
In Chapters 2 and 3, we will use EFT to study the predictivity of some models that could be phenomenologically interesting as well as their implications in terms of microscopic physics and causality.

In the second part of this thesis, we will take the direction of using the top-down approach to disentangle environmental effects from the dynamics of the merger. To this purpose, we will employ the EFT approach to describe binary systems as simple composite particles, using the NRGR formalism as well as an averaging technique [72]. In the specific, we will derive a simplified description of the evolution of a relativistic binary system perturbed by a distant third body. In this case the EFT point of view allows to exploit the separation of scales characterizing the system, encoded in the distances between the bodies, to simplify the dynamics. We will develop in Chapter 4 the description of the first non-trivial effects, while we will refine the construction to include subtler effects in Chapter 5.

Two additional sections at the beginning of the two parts of this thesis are devoted to a more detailed introduction of the two main topics. In the following we will use natural units in which both the reduced Planck constant and the speed of light are set to unity, $\hbar = c = 1$, and we choose the metric signature to be mostly positive: $(-,+,+,+)$.

Part I

RESTRICTING THEORIES BEYOND GR

PHYSICS BEYOND GR FROM THE EFT POINT OF VIEW

As mentioned in Chapter 1, the detection of gravitational waves emitted during binary mergers makes possible to probe the fully non-linear strong-field dynamics of gravity at lengthscales of order of kilometers. In this way, it is possible to test to which extent the dynamics of gravity is described by GR as opposed to theories that depart from GR predictions in the context of astrophysical binary mergers. Therefore, in practice one wants to parametrize theories beyond GR that are compatible with all current observations but predict deviations that would become testable in future experiments.
The bottom-up EFT approach represents a very efficient way to explore the parameter space of different models, allowing to keep track of which interactions affect more markedly the dynamics. In addition to this, EFTs are the natural way to frame our understanding of GR. Indeed, in analogy with the Standard Model, GR should be understood as a description of gravity only valid at low energies[2], determined by the symmetry under the Lorentz group of Special Relativity and by the presence of a massless spin-two particle, the graviton [74–76].

Technically, the physics described by GR in empty, asymptotically flat space-time can be encoded in the so-called Einstein-Hilbert action:

$$S_{EH} = \int d^4x \sqrt{-g} \frac{{M_{\rm Pl}}^2}{2} R \,, \tag{1.1}$$

where g is the determinant of the metric $g_{\mu\nu}$, R is the Ricci scalar obtained by the same metric and $M_{\rm Pl}$ is the Planck mass, related to the Newton's constant by ${M_{\rm Pl}}^2 = 1/8\pi G$. In this action, two degrees of freedom, the gravitons, are packaged into the metric $g_{\mu\nu}$. This geometric interpretation of gravity can be seen both as a result of describing a massless spin 2 field, the graviton, in a relativistic theory [74, 75] and as a result of the principle of general covariance. This principle states that different frames of reference and coordinate systems lead to equivalent descriptions of nature and translates in a symmetry under diffeomorphisms of the space-time manifold. For a thorough exploration of the geometric interpretation of GR, see e.g. [77]. In the presence of matter, GR describes gravitational interaction through minimal coupling, that is replacing partial derivatives ∂ with covariant (Levi-Civita) derivatives ∇ in the matter action and adding a factor $\sqrt{-g}$ to its integration measure. This ensures invariance of the theory under diffeomorphisms and implies the universality of gravitational attraction on point-like objects, i.e. the weak equivalence principle of GR.

In this setup one can go beyond GR by writing an effective action that contains the Einstein-Hilbert contribution of GR plus additional operators. This strategy allows to control how much GR is being deformed, making easier to understand which theories are compatible with our experimental knowledge of gravity. The operators appearing are fixed by the symmetries of the low energy descriptions and by the low energy

2 Due to its non-renormalizability, GR has long been known as a microscopically inconsistent description of gravity and space-time [73].

degrees of freedom. Once the symmetries and degrees of freedom are fixed, the theory will be specified by the dimensionful coefficients that multiply each of the allowed operators in the Lagrangian.

The most straightforward possibility is to consider operators that are invariant under the same symmetry group as GR, that is diffeomorphisms, without adding any other field besides the metric. This leads to an EFT in which the operators are geometric quantities built by fully contracting the tensor indices of powers of the Riemann tensor $R_{\mu\nu\rho\sigma}$ and covariant derivatives. Such theories have been studied extensively, see e.g. [78, 79].
In $D = 4$ space-time dimensions, all of the non-trivial operators of this kind lead to equations of motion of order higher than second. The most natural expectation, according to the EFT point of view, is that such theories are only valid at energies low enough to make the higher-derivative contributions small. This is for instance the case of chiral perturbation theory [62]. At higher energies a UV completion of the EFT including new degrees of freedom will instead take over and modify the dynamics. Indeed, the larger set of initial conditions needed to determine the evolution of the system when higher derivatives cannot be neglected would indicate the presence of extra degrees of freedom if the theory were extrapolated to high energy. A result known as the Ostrogradsky theorem, see e.g. [80, 81], states that generically in such cases the dynamics would develop instabilities, such as the runaway behavior in the Abraham-Lorentz radiation reaction force [82, 83]. This means that the theory loses consistency when higher derivatives are not perturbatively small.

The next-to-minimal way to parametrize theories beyond GR is to explicitly introduce one degree of freedom coupled to gravity. The simplest case is adding a scalar field ϕ, which is done in the so-called scalar-tensor EFTs beyond GR. The extra degree of freedom in these cases does not need to be ghost-like as in the case of Riemmannian EFTs, instead it is a physical low-energy degree of freedom that can be consistently described by the EFT.
Despite being a very simple deformation of GR, adding a scalar field to the gravitational sector makes possible to describe a considerably rich phenomenology beyond GR. This has received significant attention in the literature, see e.g. [84] and references therein. In the following we will mainly focus on a specific kind of scalar-tensor theories, in which the interactions enjoy a symmetry under constant shifts of the scalar field: $\phi \to \phi + c$, c being an arbitrary constant. The shift symmetry makes natural for the scalar field to be massless, allowing to envision it playing an important role in the dynamics on cosmological and astrophysical scales.

This possibility makes interesting to understand whether binary mergers can become a probe of these dynamical components of cosmological dynamics for instance if the presence of the scalar field alters the physics of compact objects [85]. In the last decade, this question has been addressed in the context of black hole physics, trying to determine whether a shift-symmetric scalar field can modify black holes and produce hair, by acquiring a non-trivial profile. Remarkably, a quite general no-hair theorem was proven in this context [86], leading to the understanding that only a few exceptional operators could lead to sizable, static non trivial solutions for ϕ in shift-symmetric scalar tensor theories [87, 88]. See [89] for other

types of hair in black hole geometries.

Despite being a remarkable insight, understanding which operators can endow black holes of features that are not predicted by GR does not make clear which theories are viable for describing testable effects in binary mergers. Here we tackle this question and explore the features of EFTs that include a sizable hair-inducing interaction, by taking into account two important points. First, the fact that the coefficients of the various operators in an EFT are not arbitrary, rather they are interconnected by the perturbative EFT structure. For instance this makes necessary to include other operators with non-suppressed coefficients besides those one wants to focus on, which might substantially alter the signatures expected in binary mergers. Second, the fact that any choice of values for the coefficients of an EFT determines the energy scale at which the EFT description breaks down, as well as the underlying microscopic dynamics that takes place at higher energies, the so-called UV completion. This fact can be especially relevant when the EFT breaks down at scales that can be accessed with experiments. These two aspects help us understanding the physical meaning of certain technical issues that can arise in the theories of interest. In particular, we will see that for the exceptional models that lead to shift-symmetric black hole hair, either certain quantities become divergent or certain excitations travel superluminally. Giving a physical interpretation of these issues will allow us to gain insight on whether models that predict shift-symmetric black hole hair are likely to be observed in future experiments. In stronger words, the EFT point of view will make possible to narrow down the space of theories that produce testable deviations from GR in binary mergers.

More in detail, as we will show in Chapter 2, in the case of non-spinning black holes only one of the hair-inducing operators leads to a theory free of divergent local, scalar quantities in regions of space-time surrounding a hairy black hole. These scalar quantities would generically appear in the EFT as operators, with a non-vanishing coefficient. The presence of such divergent operators, neglected when computing the background solution for the hairy black hole, would mean that the theory is not predictive and the solution cannot be trusted.
This analysis based on predictivity of the theories rules out all of the hair-inducing interactions besides one, the so-called scalar-Gauss-Bonnet (sGB) operator:

$$\alpha M_{\rm Pl}\, \phi\, \mathcal{R}^2_{GB}\,, \tag{1.2}$$

where α is a length square and $\mathcal{R}^2_{GB} = R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma} - 4R^{\mu\nu}R_{\mu\nu} + R^2$ is the Gauss-Bonnet invariant, a topological density quadratic in the Riemann tensor. Since this invariant is a total derivative, the sGB operator respects the shift symmetry of ϕ. As we will see, the topological origin of this operator is also the reason why its presence does not induce the divergence of local, scalar quantities, making it possible to consistently include this operator in an EFT. Moreover, despite containing contributions with four derivatives and three fields, this operator contributes to the equations of motion with terms that have at most second derivatives [90, 91].

With this result, one knows that any EFT containing the sGB operator will describe black holes with scalar hair. For this reason, black hole phenomenology in such models has been the focus of considerable efforts

[92–97].
As one might expect, the importance of the effects introduced by the sGB operator is controlled by the size of α. In particular, if α were as large as km^2, then the emission of scalar waves in addition to gravitational waves would determine an excessive energy loss during the inspiral phase of a black hole merger compared to what has been observed[3] [98]. This upper bound makes interesting to investigate values of α that are close to the detectability threshold, in order to characterize the signatures of the sGB operator that might be tested with the next generation of gravitational wave detectors.

With this perspective, in Chapter 3 we will explore the features of shift-symmetric theories that include the sGB operator with a sizable coefficient, with special attention to the cases in which its coefficient α is close to km^2. While in GR minimal coupling ensures that any probe traveling in a gravitational field accumulates a universal Shapiro time delay [6], in a theory in which non-minimal gravitational interactions are included, such as in the case of sGB, one can expect the Shapiro time delay to get modified and to depend on the details of the probe, for instance its spin. In particular, studying these questions in the context of EFTs including the sGB operator, we will diagnose the presence of superluminal excitations around black holes. This by finding that a time advance becomes unavoidable for specific states, much like in the case of a non-minimal three-graviton interaction [99]. In more precise terms, our result states that certain superpositions of scalar and graviton excitations travel faster than the speed at which light travels in flat space-time, when they propagate below a certain impact parameter with respect to a black hole. This happens when the impact parameter is of order of or lower than $\sqrt{\alpha}$.
This result, however interpreted, means that theories describing the sGB interaction do not fit painlessly in the usual relativistic picture of causal dynamics. A simple way to reconcile this finding with the standard understanding of causality is to attribute the prediction of superluminal propagation to a technical fault of our approach. Namely, assuming that the computations carried out in the EFT fail when the impact parameter becomes too small, leading to unphysical predictions. This will be the case when new degrees of freedom appear at length-scales lower than $\sqrt{\alpha}$, altering the dynamics. In other words, the natural expectation is that a UV completion of the EFT will appear at those scales, modifying the faulty predictions of the EFT and restoring causality. Strikingly, this resolution of the tension with causality implies that the EFT description breaks down at energies much lower than what one would expect based on NDA, since all of the higher derivative contributions are still well suppressed when the impact parameter is lowered to the value giving superluminality.
The conclusion that new degrees of freedom should appear at low energies, above a scale $\Lambda_{UV} \sim 1/\sqrt{\alpha}$, can be read as a bound in the EFT, stating that in a theory free of superluminal propagation, for a fixed cutoff, the coefficient of the sGB operator will have to be bounded by the cutoff: $\alpha \lesssim 1/\Lambda_{UV}^2$.
As we will argue, if α were close to values that would make it detectable,

3 Another upper bound on α comes from requiring the predictivity of the theory, i.e. absence of singularities, close to the black hole horizon [92]. As we will discuss in Chapter 3, if one poses that the EFT only describes black holes that are of astrophysical (or larger) mass, then the observational bound is the most restrictive. Moreover this bound becomes weaker when other operators are included in the EFT [85].

then this UV completion would call for an infinite tower of very light spinning particles, with masses as low as 10^{-10}eV. Having no experience of such particles altering gravity at scales below the kilometer, we conclude that the most likely explanation is that the sGB interaction, if present, has a much smaller coefficient than what would become testable in the foreseeable future.

In addition to analyzing the shift-symmetric scalar-tensor cases, we will consider the analogous class of models describing a shift-symmetric parity-breaking coupling to gravity. In this context, the dynamical Chern-Simons (CS) operator, of the form $\phi\, R_{\mu\nu\rho\sigma}\tilde{R}^{\mu\nu\rho\sigma}$, with $\tilde{R}$ the dual of the Riemann tensor, plays the same role of the sGB operator [100]. The main differences between the two operators is that the dynamical CS operator leads to third order equations of motion and produces hair only for spinning black holes. We will show that the same superluminality issues found in the sGB case arise for this operator.

Moving beyond the shift-symmetric theories, we will briefly consider the case of another operator that breaks minimal coupling, $\phi^2\,\mathcal{R}^2_{GB}$, which can be considered as a proxy for interactions giving rise to scalarized black holes [101–103]. Despite not finding an equally strong result as in the previous cases, we will point out various facts suggestive of a similar conclusion, i.e. that tensions with causality may force the EFT to break down at energies much lower than what would be inferred by NDA.

This analysis shows how theoretical considerations can help exploring the space of theories that might be of phenomenological interest. A complementary approach to diagnosing superluminal propagation is that of studying the S-matrix and its analytic structure. The analytic structure of the S-matrix has been actively explored in the last sixty years [104–106], see for a pedagogical review [107], with the last twenty years seeing many interesting applications to the study of EFTs [108–123]. In the context of two-to-two S-matrix elements, the control over the analytic structure of the scattering amplitude with respect to the Mandelstam invariants, combined with unitarity, locality and crossing symmetry, allows to apply the Cauchy residue theorem and derive constraints over the coefficients that characterize the low energy EFT description of the system. These constraints take the form of sum rules expressing the low energy EFT coefficients in terms of integrals over the real axis of e.g. the $s-$Mandelstam complex plane. In many cases these integrals are positive, leading to positivity bounds for the EFT coefficients, in complete analogy with what is found in Kramers-Kronig relations, see e.g. [124, 125]. Owing to this analogy, these sum rules are usually dubbed dispersion relations.
Dispersion relations have been shown to give similar constraints to those derived by requiring the absence of superluminality [108]. In quantum mechanics, this can be understood by the fact that causality implies the analyticity properties that are used to derive sum rules, precisely as in the Kramers-Kronig case. However, in quantum field theory an equally sound argument connecting causality to analyticity of S-matrix elements is still lacking. For this reason, the interplay between the two approaches is still not fully understood and remains an active field of research [126].
For instance, diagnosing superluminality can produce constraints for operators that depend on more than four fields, while no dispersion

relation has been derived for higher point amplitudes yet. Conversely, in the cases we study, we will see that our superluminality bounds are parametrically close to those recently found with dispersion relations [127, 128], which are however sharper due to the optimization procedure that can be implemented when deriving dispersion relations[4].

More importantly, the approach of studying dispersion relations for two-to-two scattering amplitudes can be employed to efficiently derive bounds on the relative size of different operators in the EFT, see e.g. [116, 119]. As we discuss in Chapter 3, deriving such bounds in the presence of gravitational interactions is more complex than in other cases. In technical terms, the long range Newtonian potential makes the two-to-two scattering amplitude divergent when the scattering angle is zero, i.e. in the forward limit. This prevents deriving the usual positivity bounds, and implies that a certain negativity is allowed for the coefficients of the EFT [118, 129]. Furthermore, the presence of loop effects makes important to introduce a new approach to derive dispersion relations for gravitational scattering amplitudes, which we will describe schematically [130]. Through the sGB example, this analysis illustrates how the predictions of a simplified model with a minimal set of interactions are affected by the operators that will appear in a generic EFT.

4 An analogous optimization strategy is not known for superluminality bounds, and it would require scanning efficiently different backgrounds while computing the propagator of the available excitations.

2

BLACK HOLE HAIR IN SHIFT-SYMMETRIC THEORIES

Do black holes have hair? Fifty years have passed since this question was first formulated but it still fuels new ideas. One of the reasons behind its longevity is that both the theoretical and the experimental context surrounding it have changed dramatically in the last half a century. For instance, while the original emphasis was on characterizing the possible existence of additional parameters—in addition to the black hole mass, charge and angular momentum—that can be seen from far away, after the beginning of the era of gravitational wave astronomy, a more promising perspective has come out. Indeed, the presence of a non-trivial background at length scales of order of the light ring can modify the quasi normal mode spectrum and leave a detectable imprint in the black hole ringdown, which can therefore serve as a window on the dynamics of the gravitational sector. At least from this point of view, today there is no reason to prefer long over short hair.

In the present Chapter, we will readdress once again the aforementioned question, focusing on a somehow specific though significant situation of scalar hair in shift-symmetric theories. Indeed if a scalar field is, with the exception of a cosmological constant, the most plausible ingredient to be added to General Relativity to explain e.g. the accelerated expansion of the Universe, the presence of a shift symmetry represents the minimal choice to guarantee that such a field will be almost massless and hence relevant on cosmological scales.

Within this specific setting a clear answer, for spherically symmetric and time-independent solutions going to a constant asymptotically, was obtained in [86]. In this proof that black holes have no shift-symmetric, scalar hair, a crucial role is played by the covariantly conserved current J^μ associated with the shift symmetry. The invariance under shifts $\phi \to \phi + c$ actually implies that the scalar equation of motion can be written as the conservation of this current, which depends on the field only through its derivatives. Because of the assumed spherical symmetry and static nature of the scalar and metric backgrounds, the only non-vanishing component can be J^r. As we will explain in detail at the beginning of the next section, for a black hole solution to be consistent with the paradigm of EFT, at the horizon every physical and local scalar quantity, and $J_\mu J^\mu$ in particular, must be finite. This implies that J^r has to vanish at the horizon, since g_{rr} diverges on that surface. Using now the conservation of the current, Ref. [86] argues that $J^r \equiv 0$ everywhere. At this point, one can argue that if the dependence on ϕ in the Lagrangian starts quadratically then the current will be proportional to ϕ':

$$J^r = \phi' F[\phi', g, g'] \quad \text{with } F \text{ a regular function}\,, \tag{2.1}$$

where the absence of ϕ'' and higher derivatives of the field is ensured if the theory has second-order equations of motion. The presence of a kinetic

term for the scalar field translates in the fact that F approaches a non-zero constant as $\phi' \to 0$. These two facts, along with the condition $J^r \equiv 0$, imply that if $F[\phi', g, g']$ is a regular function around $\phi' = 0$, then it must be $\phi' \equiv 0$ and therefore the hair vanishes.

Soon after the appearance of this theorem, however, it was realized [131] that such a simple and compelling argument admits a subtle exception. Consider for instance the theory:

$$S = \int d^4x\sqrt{-g}\left(\frac{M_{Pl}^2}{2}R - \frac{1}{2}(\partial\phi)^2 + \alpha M_{\rm Pl}\phi\mathcal{R}^2_{GB}\right) . \tag{2.2}$$

The Gauss-Bonnet invariant, $\mathcal{R}^2_{GB} \equiv R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma} - 4R^{\mu\nu}R_{\mu\nu} + R^2$, is a total derivative and thus its coupling with the scalar preserves the shift symmetry $\phi \to \phi + c$. This term gives a ϕ-independent contribution to the scalar equation of motion (and therefore to the current J^r), invalidating the assumption in (2.1). It acts as a source in the $J^r = 0$ equation, that does no longer allow for the trivial solution with a vanishing ϕ'. While the presence of a linear scalar Gauss-Bonnet (sGB) coupling is indeed a sufficient condition to guarantee that black holes have hair, the actual solution found in [131] seems puzzling. In this case not only J^r contains a ϕ-independent term, which is enough to circumvent the conclusion of the theorem, but also the norm of the current diverges at the horizon, as pointed out in [88].

A natural concern at this point is whether such a divergence, despite the regularity of the stress-energy tensor and of the resulting geometry at the horizon, is enough to conclude that solutions sourced by the Gauss-Bonnet coupling are not physical and therefore that the no-hair result in shift-symmetric theories is robust. A more optimistic perspective could instead be that the sGB example is just the first manifestation of a whole class of theories with hairy black holes, sourced by Lagrangian operators that give rise to ϕ-independent contributions to J^μ, among which there can be solutions with finite J^2.

A reason to consider the second possibility is the following. While the sGB coupling manifestly contains terms with higher derivatives on the metric, it gives rise to second-order equations of motion. This means that it has to belong to the large family of shift-symmetric scalar-tensor Lagrangian with this property, the so-called Horndeski theories [132] or generalized galileons [133]. Such an equivalence was pointed out in [91] and it will be discussed in App. A.2. From this point of view, the peculiarity of the sGB operator grows dim and in fact Ref. [88] finds several other operators of the Horndeski-type that give contributions to J^r that depend only on the metric in spherically symmetric and static backgrounds. The authors then conclude that there exist examples of black holes with hair and a finite norm of the current at the horizon. Despite this result, in a subsequent paper [134] it is claimed that all the hairy black hole solutions of this kind cannot be smoothly connected to Minkowski space-time, leaving those generated by the Gauss-Bonnet linear coupling as the only possibility.

Given the somehow unsettled status of the original question in the literature, in this chapter we will clarify if hairy black holes in shift-symmetric theories are One, No One or One Hundred Thousand, quoting Pirandello. The first step will be to show, in Section 2.1.1 that the Gauss-Bonnet current is not covariant under diffeomorphism and that, as a result, J^2 is not a scalar quantity. Its divergence at the horizon is therefore non-physical. The existence of an equivalent description of the sGB coupling in terms of a Horndeski operator, however, implies that there is a different form of the

current which is instead covariant and still divergent. In spite of that, as we will discuss in Section 2.1.2, in this case the vector J^μ and its norm contain powers of $(\partial\phi)^2$ at the denominator. This non-locality for $X \to 0$ does not affect in any way the dynamics, but deprives the divergence of J^2 of any physical meaning. The conclusion is that the presence of Gauss-Bonnet actually represents a well-defined exception to the no-hair theorem. Notice that only black holes feature "long" hair in these theories, i.e. solutions $\phi \propto 1/r$, while compact objects without horizon like neutron stars do not [135].

After having discussed the sGB models, in Section 2.2 we move on and examine the whole class of Horndeski shift-symmetric Lagrangians to identify if a similar behavior is present in other cases as well. While, as already noticed in [88], for static and spherically symmetric solutions there are several exceptional operators that contribute to J^μ with a regular and scalar-independent term, as soon as the background solution is slightly deformed, every operator of this type manifests its non-local nature and becomes divergent in the limit of Minkowski spacetime. To further assess the robustness of the no-hair result, in Section 2.3 we then extend the analysis to Lagrangians that still propagate 3 degrees of freedom (the graviton plus a scalar) but nonetheless have equations of motion with higher-order derivatives, the so called degenerate higher-order scalar-tensor (DHOST) theories. These include Horndeski and Beyond Horndeski as particular cases.

In Section 2.3.3 we briefly study the case of the most general shift-symmetric EFT. In this class a prototypical example, which shares many similarities with sGB, is given by $\phi R\tilde{R}$. Finally, conclusions are drawn in Section 2.4.

2.1 THE SCALAR GAUSS-BONNET OPERATOR

2.1.1 *The Gauss-Bonnet current*

We want to understand whether the divergence of J^2 at the horizon is a pathology of the sGB hairy solutions or not. Why should the divergence of a scalar quantity $\mathcal{O}$ be worrisome, even when the stress-energy tensor and the geometry are regular at the horizon? One reason is that a scalar quantity can be added to the Lagrangian of the system with an arbitrary coefficient in front $\mathcal{L} \supset \lambda\mathcal{O}$. In doing so the black hole solution will change and if $\mathcal{O}$ diverges at the horizon, this will happen no matter how small λ is. The solution cannot be trusted since it is extremely "unstable" if one modifies the theory. The situation is already pathological in classical physics, but it is even more so when we consider Quantum Mechanics, since loop corrections will induce $\lambda \neq 0$ even if we start with $\lambda = 0$. Another related way to see the pathology is that in general a particle will be coupled to the scalar $\mathcal{O}$. This means that one gets an effect on the dynamics of the particle (and on its stress-energy tensor) that diverges at the horizon. This suggests that the solution is unstable when matter is included in the picture.

What we said holds for a general operator $\mathcal{O}$, but J^2 turns out to be quite special. Indeed, the full current contains a part J^μ_{GB} associated with the Gauss-Bonnet term in the action, $\phi\mathcal{R}^2_{GB}$. This current *is* not covariant under diffeomorphisms and therefore any scalar built with it is not invariant under diffs. Therefore one cannot add to the Lagrangian λJ^2 (or write a coupling with a particle) and the issue above does not arise. Simply stated, J^2 is not a scalar quantity: its value, and thus its divergence, depends on the coordinates we choose. The divergence of J^2 is immaterial, like the diver-

gence of a component of the metric or of a Christoffel symbol. The current J^μ_{GB} satisfies $\nabla_\mu J^\mu_{GB} = \mathcal{R}^2_{GB}$, but the form of the current is ambiguous and there is no privileged expression, even when a coordinate system is chosen [90].

This statement is analogous to what happens in a (non-abelian) gauge theory for the term $TrF_{\mu\nu}\tilde{F}^{\mu\nu}$. This object is notoriously a total derivative $TrF_{\mu\nu}\tilde{F}^{\mu\nu} = \partial_\mu G^\mu$, with $G_\mu = \epsilon_{\mu\nu\lambda\sigma} TrA_\nu(F_{\lambda\sigma} - \frac{2}{3}A_\lambda A_\sigma)$. Similarly to our case, the current G^μ is not gauge-invariant and G^2 is not a gauge-invariant scalar that can be added to the Lagrangian.

Let us make some examples of the forms the current J^μ_{GB} can take for the Schwarzschild metric. In the presence of Killing vectors, there is a simple way to write J^μ_{GB} in the coordinates in which an isometry simply acts as a shift of one coordinate [90]. Suppose this coordinate direction has label W, then

$$J^\mu_{GB} = 2P^{W\mu\nu}{}_\rho \Gamma^\rho_{\nu W} \,, \tag{2.3}$$

where $P^{\mu\nu\rho\sigma} = \partial\mathcal{R}^2_{GB}/\partial R_{\mu\nu\rho\sigma}$ and Γ is the Christoffel symbol. This expression holds only in coordinates where the translation in W is an isometry. For the case of Schwarzschild, one can use this expression in the standard coordinates (t, r, θ, φ) either using $W = t$ or $W = \varphi$. In the first case one gets a current that points only in the radial direction (we temporarily suppress the subscript GB)

$$J^\mu_{(t)} = \left(0\,, -\frac{4r_s^2}{r^5}\,, 0\,, 0\right) \,, \qquad {J_{(t)}}^2 = \frac{16r_s^4}{r^9(r-r_s)} \,. \tag{2.4}$$

Where r_s is the Schwarzschild radius. This is exactly the current discussed in the Introduction and indeed ${J_{(t)}}^2$ diverges at the horizon, $r = r_s$. On the other hand, with the choice $W = \varphi$ one gets

$$\begin{aligned} J^\mu_{(\varphi)} &= \left(0\,, \frac{4r_s(r-r_s)}{r^5}\,, -\frac{8r_s\cot(\theta)}{r^5}\,, 0\right) \,, \\ {J_{(\varphi)}}^2 &= \frac{16r_s^2\left(-r_s + 4r\cot^2(\theta) + r\right)}{r^9} \,. \end{aligned} \tag{2.5}$$

The divergence of both these currents gives the Gauss-Bonnet invariant: $\nabla_\mu J^\mu_{(t)} = \nabla_\mu J^\mu_{(\varphi)} = \mathcal{R}^2_{GB} = 12r_s^2/r^6$. However ${J_{(\varphi)}}^2$ is finite at the horizon, while it diverges on the azimuthal axis. This example makes clear that J^2_{GB} is not a diff invariant quantity.

A general expression of the Gauss-Bonnet current, which does not assume isometries of the metric, can be given in terms of the spin connection. In the Appendix A.1 we compute this current in the Schwarzschild spacetime and show that it is not covariant by writing it in different coordinate systems, in particular in Kruskal-Szekeres coordinates where the metric is regular at the horizon.

2.1.2 *Horndeski form of sGB*

Since the sGB operator is such that the equations of motion are of second order and that there is symmetry under constant shift of the scalar field, it must be possible to express it in terms of the so-called shift-symmetric Horndeski Lagrangian. Indeed, the latter describes the most general shift-

symmetric scalar-tensor theory with second-order equations, and is given by the sum of the following terms:

$$\begin{aligned}
\mathcal{L}_2^H &= G_2(X)\,,\\
\mathcal{L}_3^H &= G_3(X)[\Pi]\,,\\
\mathcal{L}_4^H &= G_4(X)R - 2G_{4,X}(X)\left([\Pi]^2-[\Pi^2]\right),\\
\mathcal{L}_5^H &= G_5(X)G_{\mu\nu}\Pi^{\mu\nu} + \frac{1}{3}G_{5,X}(X)\left([\Pi]^3-3[\Pi][\Pi^2]+2[\Pi^3]\right),
\end{aligned} \tag{2.6}$$

where $X \equiv g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$, $\Pi_{\mu\nu} \equiv \nabla_\mu\nabla_\nu\phi$, $G_{\mu\nu}$ is the Einstein tensor and square brackets indicate the trace of an expression, e.g. $[\Pi] = \Box\phi$.

It has been pointed out in Ref. [91] that the choice $G_5 = \log(X)$ gives indeed the same equation of motion as the linear sGB operator (without any field redefinition). In Appendix A.2 we give some details about the proof of this equivalence. The benefit of this alternative way of writing the sGB operator is that now the Noether current J^μ_{H5} associated with the shift-symmetry is covariant. For this reason, contrarily to the previous case, the norm of this current $(J_{H5})^2$ is a true scalar and its divergence looks now problematic. For $G_5 \sim \log|X|$ one has

$$\begin{aligned}
(J_{H5})^2 = \frac{4}{X^4}\Bigg\{&-\frac{X}{18}\Big([\Pi]^3-3[\Pi][\Pi^2]+2[\Pi^3]\Big)^2 \\
&+\partial\phi\cdot\Bigg[\Pi^6-2[\Pi]\Pi^5+[\Pi]^2\Pi^4-\frac{\Pi^3}{3}\Big([\Pi]^3-3[\Pi][\Pi^2]+2[\Pi^3]\Big)\\
&\quad+\frac{\Pi^2}{12}\Big([\Pi]^4-6[\Pi]^2[\Pi^2]+8[\Pi][\Pi^3]-3[\Pi^2]^2\Big)\\
&\quad+\frac{\Pi}{6}\Big([\Pi^2]-[\Pi]^2\Big)\Big([\Pi]^3-3[\Pi][\Pi^2]+2[\Pi^3]\Big)\Bigg]\cdot\partial\phi\Bigg\}\\
&+\mathcal{O}(R_{\mu\nu\rho\sigma})\,.
\end{aligned} \tag{2.7}$$

Here we only wrote explicitly the terms that survive in flat space: the complete expression contains terms up to quadratic order in the curvature, indicated by $\mathcal{O}(R_{\mu\nu\rho\sigma})$. It is easy to see that $(J_{H5})^2$ is a non-local operator, with powers of X at denominator. As such it cannot be added to the action if one is interested in solutions for which $X \to 0$ somewhere. Therefore, its divergence is immaterial, in the same way one is not worried about $1/X$ going to infinity for a solution where the scalar is a constant. The above quantity is generally ill-defined as $X \to 0$. As discussed in Ref. [88], $(J_{H5})^2$ diverges on the horizon of a hairy black hole. However this does not invalidate the solution, since the operator is non-local.

Since we now understand that the operator $(J_{H5})^2$ is non-local and cannot be added to the Lagrangian, one may worry about the theory we started with, featuring $G_5 = \log(X)$. The appearance of powers of X in denominators suggests that the theory is pathological in the limit $X \to 0$. However, this cannot be the case, since the theory is equivalent to the original sGB. Indeed, the non-locality of $G_5 = \log(X)$ is only apparent as we are going to explicitly show in Section 2.2 and Appendix A.2.

2.1.3 *Boundedness of local scalar quantities*

Having established that the divergence of a nonlocal quantity does not invalidate the sGB solution, one may ask whether there can be instead a local

scalar quantity that diverges. One can then fully trust the solution only if no local scalar operators blow up (outside the physical singularities). Here we will verify that this is indeed the case for the sGB solution. We will see that requiring the boundedness of scalar quantities forces a condition on the scalar field ϕ, i.e. that all its radial partial derivatives $\partial_r^n \phi$ have to be bounded everywhere, and in particular at the black hole horizon.

While this result is evident far away from the black hole, it becomes less obvious at the horizon, where some coordinate systems display a non-physical singularity. In order to overcome this complication, one can choose coordinates in which all the geometrical quantities are smooth at the horizon. This can be achieved for instance by means of Kruskal-Szekeres-like coordinates or through locally inertial coordinates, where the metric is set to Minkowski in a specific point (e.g. $g_{\mu\nu}(p) = \eta_{\mu\nu}$ in a point p at the horizon). Choosing these last coordinates, the Christoffel symbols will vanish at the chosen point but will have non zero derivatives: these will describe (up to a Lorentz boost) the finite tidal forces experienced by a free-falling observer that is crossing the horizon.

Being interested in spherically symmetric and static solutions, it is enough to compute scalar quantities in a single point of the horizon. Moreover, since the geometry of a black hole is non-singular at the horizon, one can simply consider quantities that depend on the scalar field, for instance having the form $(\nabla^n \phi)^2$. For the same reason, in these quantities the terms displaying the most severe divergence when derivatives of ϕ are not well behaved will be those involving only partial derivatives[1].

Writing the metric in Schwarzschild-like coordinates as

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + \rho^2(r)(d\theta^2 + \sin^2\theta\, d\varphi^2)\,, \tag{2.8}$$

with $f = 0$ for $r = r_s$, we can define a locally inertial frame in a point p using coordinates $(\hat{t}, \hat{r}, \hat{\theta}, \hat{\varphi})$ having origin in p and such that in p:

$$d\hat{t} = \sqrt{f}dt\ ,\ d\hat{r} = \frac{1}{\sqrt{f}}dr\ ,\ d\hat{\theta} = \rho\, d\theta\ ,\ d\hat{\varphi} = \rho\, \sin\theta\, d\varphi\ . \tag{2.9}$$

The Jacobian of this transformation will be diagonal in p. For this reason we understand that in the leading term of $(\nabla^n \phi)^2$ in p only $\hat{r}$ partial derivatives will appear, each corresponding to a weighted r partial derivative: $\partial_{\hat{r}} = \sqrt{f}\partial_r$. In conclusion in the chosen point we have

$$(\nabla^n \phi)^2 \sim (\partial_{\hat{r}}^n \phi)^2 + \cdots \sim f^n (\partial_r^n \phi)^2 + f^{n-k-p}\partial_r^{n-p}\phi\partial_r^{n-k}\phi + \dots\ , \tag{2.10}$$

where the second term on rhs (with $k + p < n - 1$) indicates schematically a series of contributions having the same magnitude of the first one and the dots indicate smaller terms where derivatives hit the Christoffel symbols.

Writing $\phi \sim f^\gamma$ when $r \sim r_s$, it becomes clear that if γ is not a positive integer (or zero) there will be large enough values of n such that the terms in Eq. (2.10) will diverge at the horizon[2], making the whole scalar $(\nabla^n \phi)^2$ diverge as $f^{2\gamma - n}$. For this reason we see that all the scalar quantities built using the metric and the scalar field will be bounded when computed on

1 Even though divergent boost factors might make the derivatives of the Christoffel symbols diverge, these would still be subleading with respect to partial derivatives of the scalar field, which would get the same boost-enhancement.

2 For some special non-integer values of γ there will be a single integer n_γ for which the leading contributions in (2.10) add up to zero, but this does not change our conclusion, since infinitely many other scalars will diverge.

a sGB hairy background if the hair has $\partial_r\phi$ and its higher radial derivatives bounded at the horizon. This condition is satisfied by the perturbative solution of [92].

2.2 ADDITIONAL HAIR IN HORNDESKI?

In the Horndeski form, the sGB theory violates the assumptions of the no-hair theorem of [86] since G_5 in non-analytic for $X \to 0$. Indeed the current does not start linearly in ϕ' and Eq. (2.1) does not hold. A natural question is therefore whether one can find additional hairy solutions (under the same symmetry assumptions stated in the Introduction) when the other Horndeski functions are non-analytic for $X \to 0$. Examples of such theories have already been considered in [88], where some particular cases were studied in which the radial component of the current contains a ϕ'-independent term (i.e. a term in $F[\phi', g]$ proportional to $1/\phi'$), in a static and spherically symmetric setting. In this section (see also [134]) we study this possibility and we conclude that all these additional examples are pathological. Here we stick to theories with second-order equations of motion, while more general cases will be discussed in the next section.

Writing the spherically symmetric, static metric as

$$ds^2 = -h(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega^2 \,, \tag{2.11}$$

the radial component of the shift-symmetry Noether current for generic Horndeski Lagrangian (2.6) takes the form [88]:

$$\begin{aligned} J^r_H \;=\; & 2f\phi' G_{2X} + f\frac{rh'+4h}{rh} X G_{3X} - 4f\phi' \frac{fh - h + rfh'}{r^2 h} G_{4X} \\ & -8f^2\phi' \frac{h+rh'}{r^2 h} X G_{4XX} - fh' \frac{1-3f}{r^2 h} X G_{5X} + 2\frac{h' f^2}{r^2 h} X^2 G_{5XX} \,. \end{aligned} \tag{2.12}$$

Therefore, we see that it is possible to have contributions independent of ϕ' when the functions G_i behave at small X as

$$G_2(X) \sim \sqrt{|X|}\,, \;\; G_3(X) \sim \log|X|\,, \;\; G_4(X) \sim \sqrt{|X|}\,, \;\; G_5(X) \sim \log|X|\,, \tag{2.13}$$

where the last choice gives the sGB operator. Keep in mind that the function G_4 will always include a leading constant term that drops out of the current in Eq. (2.12) and corresponds to the Einstein-Hilbert part of the Lagrangian. The non-analytic behavior for $X \to 0$ is worrisome when one wants to study the Lorentz-invariant vacuum $X = 0$ or approaching it as it happens going far away from a localized black-hole solution. In the following we are going to show that these theories are indeed pathological, with the only exception of sGB.

2.2.1 *Troubles with a Lorentz-invariant solution*

We want to study Lorentz-invariant solutions of the theories with the non-analytic behaviors of Eq. (2.13). We are going to show that, with the exception of sGB, these solutions are pathological since the equations of motion are not continuous in this limit, i.e. the result depends on how the flat, Lorentz invariant solution is approached. Let us take the metric to be

Minkowski from the beginning $g = \eta$ (this defines a particular direction in which we approach the solutions we are interested in).

The equations of motion for generic Horndeski functions G_i read:

$$\begin{aligned}
\nabla_\mu J^\mu_{H2}\Big|_{g=\eta} &= 2G_{2,XX}\,\Pi^{\mu\nu}\,\partial_\mu\phi\,\partial_\nu\phi + G_{2,X}[\Pi]\,, \qquad (2.14)\\
\nabla_\mu J^\mu_{H3}\Big|_{g=\eta} &= 4G_{3,XX}\left[\Pi^{\mu\nu}\,[\Pi] - (\Pi^2)^{\mu\nu}\right]\partial_\mu\phi\,\partial_\nu\phi + 2G_{3,X}\left([\Pi]^2 - [\Pi^2]\right)\,,\\
\nabla_\mu J^\mu_{H4}\Big|_{g=\eta} &= 8G_{4,XXX}\left[(\Pi^3)^{\mu\nu} - (\Pi^2)^{\mu\nu}\,[\Pi] + \frac{1}{2}\Pi^{\mu\nu}\left([\Pi]^2 - [\Pi^2]\right)\right]\partial_\mu\phi\,\partial_\nu\phi\\
&\quad + 2G_{4,XX}\left([\Pi]^3 - 3[\Pi^2][\Pi] + 2[\Pi^3]\right)\,,\\
\nabla_\mu J^\mu_{H5}\Big|_{g=\eta} &= 4G_{5,XXX}\Big[(\Pi^4)^{\mu\nu} - (\Pi^3)^{\mu\nu}\,[\Pi] + \frac{1}{2}(\Pi^2)^{\mu\nu}\left([\Pi]^2 - [\Pi^2]\right)\\
&\quad - \frac{1}{6}\Pi^{\mu\nu}\left([\Pi]^3 - 3[\Pi][\Pi^2] + 2[\Pi^3]\right)\Big]\partial_\mu\phi\,\partial_\nu\phi\\
&\quad - \frac{1}{3}G_{5,XX}\Big([\Pi]^4 - 6[\Pi^2][\Pi]^2 + 3[\Pi^2]^2 + 8[\Pi^3][\Pi] - 6[\Pi^4]\Big)\,.
\end{aligned}$$

When the functions G_i, $i = 2,3,4,5$ behave as in Eq. (2.13), the above equations take the form

$$\nabla_\mu J^\mu_{Hi}\Big|_{g=\eta} \sim \frac{1}{X^{(i+1)/2}}\left(A_i^{\mu\nu}\,\partial_\mu\phi\,\partial_\nu\phi + [A_i]\,c_i X\right)\,, \qquad (2.15)$$

where $A_i^{\mu\nu}$ are tensors built out of the $(i-1)$-th power of $\Pi^{\mu\nu}$, and c_i are numerical coefficients that can be easily determined by inspection:

$$c_i = -\frac{1}{i-1}\,. \qquad (2.16)$$

Notice that all of the above equations of motion (in flat-space) are finite for time-independent and spherically symmetric backgrounds, as it can be confirmed by taking the divergence of (2.12) when the G_i's are given by (2.13) and then taking the Minkowski limit $f, h \to 1$. This is why no apparent problem arises when looking for hairy black-hole solutions.

However, if a Lorentz invariant solution were to exist in d dimensions, then one would have $A_i^{\mu\nu} = \eta^{\mu\nu}[A_i]/d$ (this can be seen as an additional assumption about the direction in which the limit is approached), and therefore the equation would simplify to

$$\nabla_\mu J^\mu_{Hi}\Big|_{g=\eta} \propto \left(\frac{1}{d} + c_i\right)\frac{[A_i]}{X^{(i-1)/2}} = P_i(d)\left(\frac{1}{d} + c_i\right)\left(\frac{[\Pi]}{X^{1/2}}\right)^{i-1}\,. \qquad (2.17)$$

In the last expression we used $\Pi^{\mu\nu} = \eta^{\mu\nu}[\Pi]/d$, with the prefactor given by $P_i(d) = \prod_{p=0}^{(i-2)}(d-p)$. Since $[\Pi]/X^{1/2} \sim \partial^2\phi/\partial\phi$, one has the same number of fields at numerator and denominator, so that the Lorentz-invariant limit is ambiguous. Consider for instance $\phi = A\,x_\mu x^\mu + b_\mu x^\mu + c$, for which $\Pi_{\mu\nu} = 2A\,\eta_{\mu\nu}$. The trivial Lorentz-invariant and translationally invariant configuration $\phi = const$ is reached when $A \to 0$ and $b_\mu \to 0$. Expressions like $[\Pi]/X^{1/2}$ depend on the order of these limits.

However, for each operator there is a critical dimension for which Eq. (2.17) identically vanishes, namely

$$d_i = (i-1)\,. \qquad (2.18)$$

This means that except for a single value of i, all of the other cases[3] are incompatible with a Lorentz invariant solution. This analysis is enough to conclude that in $d = 4$ all of the cases in (2.13) are not compatible with a Lorentz-invariant solution with the exception of sGB (in Appendix A.2 we are going to also study the $d = 2$ case with $G_3 \sim \log|X|$, which corresponds to a coupling $\phi^{(2)}R$).

2.2.2 *Troubles with perturbations*

A similar situation arises when considering arbitrary perturbations around an $X = 0$ background. For simplicity we will consider a Lorentz-invariant one. Indeed, consider a scalar quantity built with the scalar fields's first and second derivatives, $\mathcal{O}(\partial\phi, \Pi)$. Expanding in linear perturbations, $\phi = \phi_0 + \delta\phi$, it takes the form

$$\delta\mathcal{O} = B^\mu \, \partial_\mu \delta\phi + C^{\mu\nu} \, \partial_\mu \partial_\nu \delta\phi \, , \tag{2.19}$$

where B^μ and $C^{\mu\nu}$ depend on background quantities only, and for a Lorentz invariant background will satisfy

$$B^\mu = 0 \qquad ; \qquad C^{\mu\nu} \propto \eta^{\mu\nu}. \tag{2.20}$$

Therefore, it is enough to only track the perturbations with two derivatives acting on $\delta\phi$. For example, linear perturbations of the equations of motion (2.14) are

$$\delta\left(\nabla_\mu J^\mu_{Hi}\Big|_{g=\eta}\right) = \mathcal{Z}_i \, \Box\delta\phi \, , \tag{2.21}$$

with

$$\mathcal{Z}_i \propto \frac{P_{i+1}(d)}{\sqrt{X}} \left(\frac{[\Pi]}{X^{1/2}}\right)^{i-2} . \tag{2.22}$$

Again, we observe a problem in the limit $\phi \to const$ for the cases in (2.13) which is now even worse than for the background equations (2.17), since here there is an extra power of the field's first derivatives in the denominator. Also, similarly to what happened for the background equations discussed above, in $d = 4$ dimensions we see that the choice $G_5(X) \sim \log|X|$, i.e. sGB, is safe because the prefactor $P_6(4)$ vanishes (in a similar way in $d = 2$ we have an analogous result for the cubic Horndeski $P_4(2) = 0$). Of course $G_5(X) \sim \log|X|$ would continue to avoid problems, even going to higher order in perturbations and on more general backgrounds. Indeed, as we discussed, this case does not feature any true non-locality being equivalent to the sGB theory (see Appendix A.2).

It is important to point out that, besides the cases (2.13), many other choices of the G_i can produce hairy solutions, as long as J^r contains terms less than linear in ϕ', so that Eq. (2.1) does not hold. One such example is $G_3(X) \sim X^{1/4}$, which produces a term proportional to $\sqrt{\phi'}$. Even if in this case $\phi' = 0$ solves $J^r = 0$, from the explicit expression of J^r, Eq. (2.12) one finds also a non-zero solution:

$$\phi' \propto \frac{f^{1/2}(rh' + 4h)^2}{r^2 h^2} \quad \sim \frac{1}{r^2} \quad \text{as} \;\; r \to \infty \, . \tag{2.23}$$

3 Here we mean those that are not automatically trivial. As it is well known [136], for a given dimension d, the Galileon-like structures present in Horndeski theories with $i > d - 2$ are indeed trivial. In our setup this can be seen in Eq. (2.17) from the fact that $P_{(d+2)}(d) = 0$.

However, the same analysis carried out above for the cases (2.13) shows that also this case is pathological. The analogue of Eq. (2.22) now reads

$$\tilde{\mathcal{Z}}_3 = \frac{4}{d}(d-1)[\Pi]\left(\frac{2}{d}XG_{3XX} + G_{3X}\right) \sim \frac{[\Pi]}{X^{3/4}} \tag{2.24}$$

and again the Lorentz-invariant limit is not well-defined. These pathologies will arise for any non-analytic function at a certain order in perturbations. For instance even an apparently innocuous term $X^{n+1/2}$ will get corrections of the form $\sim X^{n-k+1/2}(\delta X)^k$ when we consider deformations $X \mapsto X + \delta X$ of the background solution. These terms will diverge as soon as $k > n$, making impossible to compute corrections whenever $X = 0$, both on the Lorentz invariant vacuum and on hairy solutions.

In conclusion, by dropping the assumption of Eq. (2.1) one gets healthy hairy solutions only in the case of sGB. The physical validity of the hairy black-hole solutions in theories of the form of Eq. (2.13) was studied in [134], reaching a similar result. However, the arguments of [134] are not completely conclusive in our view. The authors point out that if one sets $\phi = const$, with a spherically symmetric and static metric and takes the limit of Minkowski spacetime, J^r goes to zero only in the sGB case. This can be easily checked in the explicit expression of J^r of Eq. (2.12). However, in the case $G_4(X) \sim \sqrt{|X|}$ one gets $J^r \propto r^{-2}$ and this does not contribute to the equation of motion $\nabla_\mu J^\mu = 0$. (Notice that a static solution is effectively 3-dimensional, so that, following the argument of Eq. (2.18), it is not surprising that the $G_4 \sim \sqrt{|X|}$ case is healthy for a static solution.) Actually, as we discussed at length in the previous sections, $J^r = 0$ is not a necessary requirement when J^2 is a non-local operator, as it is the case for all the choices in Eq. (2.13), including sGB. As our analysis shows, one needs to go beyond static solutions to pinpoint the pathology. This also allows to exclude cases like $G_3(X) \sim X^{1/4}$ discussed above, which were not covered by the arguments of [134] since the current vanishes once $\phi' = 0$ is taken.

2.3 THEORIES WITH HIGHER-ORDER EQUATIONS OF MOTION

So far we focused on shift-symmetric theories with second-order equations of motion (Horndeski). However, the requirement that the field equations are of second order, which ensures there are no ghost degrees of freedom, can be relaxed. Indeed, even scalar-tensor theories leading to higher-order equations of motion can, in some cases, propagate only gravity plus a single extra scalar degree of freedom. For instance this happens when the following (shift-symmetric) Beyond Horndeski Lagrangian [137] is added to the Horndeski one (2.6):

$$\begin{aligned}\mathcal{L}_4^{BH} &= -F_4(X)\epsilon^{\mu\nu\rho}{}_\sigma\epsilon^{\mu'\nu'\rho'\sigma}\partial_\mu\phi\,\partial_{\mu'}\phi\,\Pi_{\nu\nu'}\Pi_{\rho\rho'}\,,\\ \mathcal{L}_5^{BH} &= -F_5(X)\epsilon^{\mu\nu\rho\sigma}\epsilon^{\mu'\nu'\rho'\sigma'}\partial_\mu\phi\,\partial_{\mu'}\phi\,\Pi_{\nu\nu'}\Pi_{\rho\rho'}\Pi_{\sigma\sigma'}\,,\end{aligned} \tag{2.25}$$

provided this degeneracy condition is satisfied [138]:

$$XG_{5X}F_4 = 3F_5\left[G_4 - 2XG_{4X}\right]\,. \tag{2.26}$$

There is an even larger set of such theories known as DHOST [139, 140], which includes both Horndeski and Beyond Horndeski as special cases. In the following we are going to extend the study of black-hole hair to this more general setup, always with the same symmetry assumptions made in the Introduction. Notice that the application of the no-hair theorem is now

not obvious, since now one expects the radial component of the current to also depend on $\phi''(r)$, violating Eq. (2.1).

2.3.1 *No-hair theorem for DHOST*

Let us start with the class of DHOST theories that can be obtained via invertible conformal and disformal transformations that depend on the scalar field:

$$\bar{g}_{\mu\nu} = \Omega(X)\, g_{\mu\nu} + \Gamma(X)\, \partial_\mu\phi\, \partial_\nu\phi \,. \tag{2.27}$$

The dependence of Ω and Γ on X only (and not on ϕ) ensures that the shift-symmetry is preserved. (Notice that the scalar field is not changed in the transformation.) The kinetic term transforms as

$$\bar{X} = \frac{X}{\Omega + X\Gamma} \,. \tag{2.28}$$

This relation with the Horndeski theories is a way to understand why these DHOST theories must propagate only gravity plus a single extra scalar degree of freedom. In particular, from Quartic and Quintic Horndeski one generates [141, 142]

$$\bar{\mathcal{L}}_4^H[\bar{G}_4] = \mathcal{L}_4^H[G_4] + \mathcal{L}_4^{BH}[F_4] + \sum_i \alpha_i L_i^{(2)}, \tag{2.29}$$

$$\bar{\mathcal{L}}_5^H[\bar{G}_5] = \mathcal{L}_5^H[G_5] + \mathcal{L}_5^{BH}[F_5] + \sum_j b_j L_j^{(3)}, \tag{2.30}$$

where the α_i's and b_j's are functions which parametrize the part of the DHOST Lagrangian which is neither Horndeski nor Beyond Horndeski[4], and $L_i^{(2)}$ and $L_i^{(3)}$ are terms quadratic and cubic in second derivatives of the scalar field respectively.

Let us consider first a theory with only Horndeski and Beyond Horndeski: it is generated by a purely disformal transformation, i.e. $\Omega(X) = 1$ and $\Gamma(X) \neq 0$. Since the equations of motion are of higher order, one would expect J^r to contain more derivatives with respect to the form of Eq. (2.1). However, this does not happen, as a consequence of the high degree of symmetry, and ϕ'' does not appear in J^r [88]:

$$J^r_{BH} = 4f^2\phi' \frac{h + rh'}{r^2 h} X(2F_4 + XF_{4X}) + 3\frac{f^2 h'}{rh} X^2(5F_5 + 2XF_{5X}) \,. \tag{2.31}$$

Therefore the no-hair theorem applies without any changes. We will discuss below new exceptions in the same vein of Eq. (2.13).

More generally, turning on the conformal part of the transformation, i.e. $\Omega_{,X} \neq 0$, allows to span this full DHOST class. In this case the current will contain higher derivatives of the scalar field: these arise from derivatives of the metric, once one uses the transformation of Eq. (2.27). Therefore extra derivatives come from the derivatives of $\Omega(X)$ and $\Gamma(X)\partial_\mu\phi\partial_\nu\phi$. Assuming that the functions Ω and Γ are regular for $X \to 0$ (Γ must start as a constant and Ω as a non-zero constant) extra derivatives of ϕ will always

4 These functions are not all independent, but satisfy relations in order to ensure the degeneracy conditions analogous to (2.26).

appear alongside extra powers of ϕ. Therefore the current, instead of being of the form of Eq. (2.1) is of the form[5]

$$J^r = \phi' F[\phi', \phi'', g'] \,. \tag{2.32}$$

Now we are in the position of extending the theorem to this case. Since in any EFT derivatives must be bounded, in the limit $\phi' \to 0$ we also have $\phi'' \to 0$. In this limit the function F must go to a constant as in the original case, since the new terms in the current are at least quadratic in ϕ. Therefore the logic of [86] still applies: since $J^r = 0$ (with the caveat of Gauss-Bonnet that we discussed at length) and $\phi' = 0$ asymptotically, it must remain so everywhere because for small values of the field the current is simply proportional to ϕ' so that this cannot move away from zero. In conclusion, the no-hair theorem is extended to DHOST theories which are connected to a healthy Horndeski theory (as defined in the previous Section) by means of a transformation with $\Omega(X)$ and $\Gamma(X)$ regular around $X = 0$.

2.3.2 *The fate of sGB*

Another way to see that the theorem still holds is to look at how black-hole solutions are transformed. Since the scalar field is not changed by the transformation, hair can neither be generated nor removed (grown nor cut) by these transformations. Moreover, the asymptotics of the solutions are preserved and their symmetries as well. Indeed, far away from the black hole the transformation (2.27) becomes trivial ($\partial_\mu \phi \to 0$),

$$\bar{g}_{\mu\nu} = \Omega(0)\, g_{\mu\nu} \qquad (r \to \infty), \tag{2.33}$$

where of course $\Omega(0) > 0$. This is a constant overall rescaling of the metric: spacetime is still asymptotically flat. Therefore the only DHOST theories with hair are the ones obtained via (2.27) starting from a sGB Horndeski theory, since this is the only Horndeski theory with hair. (Here we are not considering the possibility that a black-hole solution is mapped into a solution with a naked singularity, as discussed in [143].)

The new terms generated by such transformation from both Quartic and Quintic Horndeski, Eqs. (2.29) and (2.30), are given by

$$G_{4,X} = \bar{G}_{4,\bar{X}} \sqrt{\Omega} (\Omega + X\Gamma)^{1/2} \,, \tag{2.34}$$

$$F_4 = -\bar{G}_4 \frac{(\Gamma\Omega_X + \Omega\Gamma_X)}{\sqrt{\Omega}(\Omega + X\Gamma)^{1/2}} + 2\bar{G}_{4,\bar{X}} \frac{\sqrt{\Omega}(X\Gamma_X - \Omega_X)}{(\Omega + X\Gamma)^{3/2}} \,, \tag{2.35}$$

$$\alpha_5 = -\bar{G}_4 \frac{2\Omega_X(\Gamma\Omega_X + 2\Omega\Gamma_X)}{\Omega^{3/2}(\Omega + X\Gamma)^{1/2}} + 4\bar{G}_{4,\bar{X}} \frac{\Omega_X(2X\Gamma_X - \Omega_X)}{\sqrt{\Omega}(\Omega + X\Gamma)^{3/2}} \,, \tag{2.36}$$

for the Quartic part, while for the Quintic part

$$G_{5,X} = \bar{G}_{5,\bar{X}} \frac{\sqrt{\Omega}\,[\Omega - X(\Omega_X + X\Gamma_X)]}{(\Omega + X\Gamma)^{5/2}} \,, \tag{2.37}$$

$$F_5 = -2\bar{G}_{5,\bar{X}} \frac{\sqrt{\Omega}(\Omega_X + X\Gamma_X)}{3(\Omega + X\Gamma)^{5/2}} \,, \tag{2.38}$$

$$b_4 = \bar{G}_{5,\bar{X}} \frac{\sqrt{\Omega}\,\Omega_X}{3(\Omega + X\Gamma)^{5/2}} \,. \tag{2.39}$$

5 There are no terms with three or more derivatives of ϕ in the current, because they would give terms with four or more derivatives in the equations of motions. However the transformation (2.27) adds at most one derivative: starting with second-order equations, one ends up with at most three derivatives.

Due to the degeneracy conditions (see Refs. [141, 142]), the remaining α_i and b_j are determined by the ones shown, and therefore contain no new information. Starting with a Horndeski theory with hair, i.e. with $\bar{G}_5 = \log(\bar{X})$ (sGB) one wants to know whether it is possible to end up in a DHOST theory without the sGB term (and with all functions regular for $X \to 0$). From Eq. (2.37) it would seem that there is a possible choice of Ω and Γ in the transformation such that G_5 is regular in $X = 0$, namely

$$[\Omega - X(\Omega_{,X} + X\Gamma_{,X})] \to 0\,, \tag{2.40}$$

at least linearly in X. However, as discussed in Ref. [141], when the above combination vanishes the transformation admits a null eigenvector, i.e. it is not invertible and thus pathological.

Ref. [88] studied Beyond-Horndeski theories which could be exceptions to the no-hair theorem, along the lines of (2.13). These exceptions involve special choices of the Beyond Horndeski functions,

$$F_4(X) \sim |X|^{-3/2}\,, \quad F_5(X) \sim |X|^{-2}\,. \tag{2.41}$$

The transformation laws (2.35) and (2.38) show however that these are not reachable with regular transformations, neither starting from regular Horndeski functions, nor allowing for sGB. Indeed, in the latter case one would need to allow for $\Gamma \sim X^{-1}$ in order to generate $F_5(X) = |X|^{-2}$ from $\bar{G}_5 = \log(\bar{X})$. It is straightforward to check that, although such transformation is safe in a static and spherically symmetric background, it is ill defined for a general configuration.

We conclude then that it is not possible to remove the sGB operator with a regular and invertible transformation of the form (2.27). Therefore, the DHOST theories that we studied can be separated in two (invariant) subclasses, those with the sGB operator and therefore with hairy black holes and those without. In other words, for a given DHOST theory connected to Horndeski, in order to determine whether it can support healthy hairy black hole solutions or not, one only needs to check if Eqs. (2.34) to (2.39) can be satisfied with $\bar{G}_5(\bar{X}) \sim \log(\bar{X})$ for small $\bar{X}$.

2.3.3 *Other DHOST theories and beyond*

Besides the theories discussed in the previous sections, other DHOST classes can be defined imposing different degeneracy conditions on the higher-derivative operators added to the Lagrangian [141, 142]. This procedure outlines various DHOST classes featuring operators either quadratic or cubic in second derivatives of the scalar field. As discussed in [141, 142], further requirements might be imposed in order to select the theories which can be interpreted as a modification of General Relativity through the presence of an additional scalar degree of freedom.

In particular only some DHOST theories admit a ghost-free decoupling limit of the metric in flat spacetime. In addition to this, if one wishes to include operators from a cubic DHOST class, the degeneracy conditions required by these must be compatible with those of the quadratic DHOST theories which are necessary in order to include an Einstein-Hilbert term in the Lagrangian. As shown in [142], these two requirements narrow down the interesting classes to only two possibilities[6]. One of these is precisely

6 Classes ^{2}N-I + ^{3}M-I and ^{2}N-I + ^{3}N-I, as defined in Ref. [142].

the class studied in the previous sections, generated by conformal plus disformal invertible transformations of Horndeski theories[7]. The other class involves more complicated constraints and cannot be characterized as easily. In Appendix A.3 we show that although this class accommodates both quadratic and cubic DHOST, it contains only theories that do not allow for an Einstein-Hilbert term and are therefore unsuitable to describe a modification of General Relativity.

One can consider an even more general situation. Imposing either second order or degenerate equations of motion is motivated if at least one higher derivative (HD) operator becomes large on the solutions one is interested in. On the other hand, if HD operators can always be treated perturbatively, as it typically happens in more conventional EFTs, then such a requirement is no longer necessary and arbitrary HD operators can be considered. (Notice that this possibility is not that different from the case of sGB discussed so far: even if the sGB gives second-order equations of motion, these equations may be pathological, featuring ghost or gradient instabilities, when the sGB is as important as the scalar kinetic term [96].)

Interestingly the theorem of [86] can be extended to this very generic setting, as long as one considers energy scales below that at which the ghost degrees of freedom appear, i.e. in the regime of validity of the EFT. In a spherically symmetric and static spacetime the current will take the form:

$$J^r = \phi' F_1 + \phi'' F_2 + ... + \phi^{(n)} F_n \,, \tag{2.42}$$

where the functions F_i are assumed to be regular as ϕ' and its derivatives approach zero. Sufficiently far away, and within the regime of validity of the EFT, the leading term will be the first one in Eq. (2.42), so that following Ref. [86], $J^r = 0$ implies $\phi = const$.

One can also find exceptions to this extension of the theorem, similarly to the case of sGB, where the current contains ϕ-independent contributions. Among the various possible operators of this kind, the simplest example is given by $\phi R\tilde{R}$, i.e. a linear coupling between the scalar field and the Chern-Simons topological density (see for example [144])

$$\int d^4x \sqrt{-g}\, R\tilde{R} = \int d^4x \sqrt{-g} \nabla_\mu K^\mu \tag{2.43}$$

$$K^\mu = 2\frac{\epsilon^{\mu\alpha\beta\gamma}}{\sqrt{-g}}\Gamma^\tau_{\alpha\sigma}\left(\frac{1}{2}\partial_\beta\Gamma^\sigma_{\gamma\tau} + \frac{1}{3}\Gamma^\tau_{\beta\sigma}\Gamma^\sigma_{\gamma\tau}\right) \tag{2.44}$$

where $R\tilde{R} = R^{\mu\nu\rho}{}_\sigma \tilde{R}^\sigma{}_{\rho\mu\nu}$ and $\tilde{R}^\sigma{}_{\rho\mu\nu} := \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}R^{\alpha\beta\sigma}{}_\rho$. The current K^μ vanishes in any static spacetime and does not transform covariantly. Similarly to the sGB case, this current will forcibly source scalar hair around any (non-static) black holes. One might also consider operators with higher derivatives, for instance $\phi\,\Box(R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma})$. As remarked, for the theory to be consistent the generated hair must be small. Nonetheless, the presence of this kind of operators might be tested through future detections of gravitational waves.

2.4 SUMMARY AND DISCUSSION

In this chapter we have shown that asymptotically flat black holes in shift-symmetric scalar-tensor theories with no ghost degrees of freedom can have nontrivial scalar hair only in the presence of the operator $\phi\,\mathcal{R}^2_{GB}$ (sGB). Further assumptions include time-independence and spherical symmetry. We

7 Non-invertible transformations will land either outside this class, or in a theory involving non-regular functions.

have laid out this fact by building from the no-hair theorem of Ref. [86], which is directly applicable only to Horndeski theories. We have shown that this theorem allows a single pathology-free exception, by first addressing some concerns about the sGB solution and the infinite norm of its associated current at the black hole horizon. The fact that this object is either non diffeomorphism invariant or non-local devoids this divergence of physical meaning. Instead, any local scalar quantities were shown to be finite. In contrast, all of the other exceptions to the no-hair theorem within the realm of shift-symmetric Horndeski theories turn out to feature pathologies, such as the lack of a Lorentz invariant solution in flat space.

It is worth mentioning that although we have focused on static black holes and static hair, the no-hair theorem of [86] has been recently extended to black holes with possibly large spin [145], confirming the expectation that the results of our study outline intrinsic features of black holes in scalar tensor theories rather than accidents due to the static nature of the solutions.

Stepping away from theories with second-order equations of motion, we extended the applicability of the no-hair theorem to a larger class of shift-symmetric scalar-tensor theories, which nevertheless propagate no extra degrees of freedom (the so called DHOST). Among these, we focused on those which can recover General Relativity when $X = 0$, therefore selecting the class which also contains Horndeski and it is in fact generated from it by X-dependent invertible conformal plus disformal transformations of the metric. Leveraging this fact, we were able to show that no new operator that produces hair apart from sGB can arise in this larger class of theories, since hair cannot be generated nor removed by such transformations. Therefore, sGB remains the only consistent interaction sourcing shift-symmetric scalar hair.

It is in the context of shift-symmetric theories in which it was ultimately possible to give a sharp answer to the question of black hole hair. This is a compelling scenario since an approximately massless scalar field can be important throughout a large range of scales, from the cosmological to the astrophysical. One such interesting situation is when the effect of black hole hair on the production of gravitational waves in black hole mergers could help in unveiling the dynamics of the dark energy field. This scenario was put forward in [85], where in spite of there being only a single possible source of hair, i.e. sGB, the phenomenology is sensitive to the other operators present in the Lagrangian, allowing for a rich array of observational signatures.

We now turn to exploring the features of theories displaying the sGB interaction. In the next chapter, in particular, we will question how such theories fit in with the relativistic causal structure as well as the interplay between sGB and the various operators that generically will appear in a shift-symmetric scalar-tensor EFT.

3

CAUSALITY CONSTRAINTS ON BLACK HOLES BEYOND GR

As discussed in the previous chapters, the possibility to study gravity in the strong field regime for the first time has motivated a surge of interest in field theories that allow for black hole solutions different from the ones predicted by GR.

In the absence of a compelling guiding principle, the intrinsic complexity of the merger process has encouraged the study of simple models where deviations from GR could be order one. This is the case of scalar-tensor theories featuring the lowest-dimensional non-minimal couplings of a scalar field to gravity, capable of sourcing detectable scalar hair around black holes: a massless (shift-symmetric) scalar coupled to the Gauss-Bonnet (GB) invariant [92, 131] or to the Chern-Simons (CS) term, a.k.a. Pontryagin invariant, [146],

$$S = \int d^4x\sqrt{-g}\left(\frac{{M_{\rm Pl}}^2}{2}R - \frac{1}{2}(\nabla_\mu\phi)^2 + M_{\rm Pl}\alpha\phi\mathcal{R}^2_{\rm GB} + M_{\rm Pl}\tilde{\alpha}\phi R_{\mu\nu\rho\sigma}\tilde{R}^{\mu\nu\rho\sigma}\right), \tag{3.1}$$

where $\mathcal{R}^2_{\rm GB} \equiv R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma} - 4R^{\mu\nu}R_{\mu\nu} + R^2$ and $\tilde{R}^{\mu\nu\rho\sigma} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta}R_{\alpha\beta}{}^{\rho\sigma}$. The coefficients α and $\tilde{\alpha}$ are the length-scales (squared) parametrizing the strength of the non-minimal scalar couplings. In Chapter 2 we have seen how in shift-symmetric scalar theories a no-hair theorem [86] basically selects these two interactions as the only ones leading to black hole hair.

While very interesting from the phenomenological point of view, it is crucial to understand how much one can learn about the fundamental properties of gravity via the study of these models in the context of gravitational wave observations. To answer this important question, the very first step is to assess the consistency of such extensions of GR with what we already know: that GR provides a good description of gravitational interactions down to μm scales [147] and, at the most basic level, that the principles of unitarity, locality, and causality hold at such scales.

To this purpose we will first diagnose whether these theories are compatible with the picture of causal dynamics, in which excitations propagate inside the light cone. Due to the presence of superluminal excitations, we will come to the conclusion that the cutoff of the EFT in Eq. (3.1) is bounded from above as $\Lambda \lesssim 1/|\hat{\alpha}|^{1/2}$, where $\hat{\alpha} = \alpha + i\tilde{\alpha}$. Since the effects beyond GR, associated with the scalar hair of black holes, are observable when $|\hat{\alpha}|/r_s^2 = O(1)$, where r_s is the Schwarzschild radius, we find that for phenomenological applications, i.e. for black holes of astrophysical size, $\Lambda \lesssim 1/\text{ km}$. Therefore, the observability of black holes with scalar hair comes at the high price of a very limited regime of validity of these models. In fact, we will argue that the observation of $O(|\hat{\alpha}|/\text{ km}^2)$ non-standard effects due to the scalar hair of astrophysical black holes is likely at odds with standard gravity at distances

shorter than $|\hat{\alpha}|^{1/2}$, or, from a more dramatic perspective, it would point to the violation of fundamental principles below that scale.

Our causality bound is a generalization of the well-known fact that effective field theories exhibiting non-minimal 3-graviton or 2-photon plus 1-graviton interactions, if extrapolated beyond their regime of validity, display time advances when in a gravitational background, in conflict with causality [99, 148]. In Sec. 3.1 we show that the scalar-graviton mixing induced by the non-minimal couplings in Eq. (3.1) leads as well to a macroscopic violation of causality unless $\Lambda \lesssim 1/|\hat{\alpha}|^{1/2}$, in which case the time advance is never observable within the EFT regime of validity. As we discuss in Sec. 3.1.2, this result does not depend on the presence of other operators in the EFT. Based on this bound as well as those found in [99], along with the theoretical constraints on gravitational EFTs recently derived using dispersion relations [121, 127], we will extract generic lessons on the power counting of gravitational EFT operators, of relevance for gravitational wave science.

While our bound renders the EFT in Eq. (3.1) at the verge of its regime of validity for the physical systems of interest, there is a small range of scales where it could remain interesting. What are the effects one can expect from such a low cutoff? In Sec. 3.2.1 we investigate this question by means of dispersion relations, which connect observables at low energies, i.e. EFT coefficients, with the high-energy dynamics that lies behind them, on the basis of the unitarity, locality and causality of the scattering amplitudes. Due to the weakness of the non-minimal gravitational interactions compared to GR, as enforced by causality, we find that our one-loop positivity conditions do not subsist. Drawing from an upcoming work [130], in Sec. 3.2.2 we outline a strategy that allows to derive sharp but non-positive bounds from dispersion relations in the presence of loops, contrarily to what might be feared. Despite lacking one-loop positivity bounds, given that setting $|\hat{\alpha}| \sim 1/\Lambda^2$ fixes the power counting of the EFT, we are able to identify in Sec. 3.2.3 the leading higher-dimensional operators that should generically (yet not generally) become large when the BGR effects are maximized within the EFT regime.

In Sec. 3.3 we explore the phenomenological consequences of the additional EFT operators. The main generic lesson we extract is that it would be of great significance to extend the black hole solutions and numerical studies of their merger, obtained so far in the literature for the scalar-GB and dynamical-CS gravity theories ($\tilde{\alpha} = 0$ and $\alpha = 0$ respectively), to include these operators. This conclusion holds insofar there exist a UV completion in which gravity remains well described by GR at scales lower than $|\hat{\alpha}|^{1/2} \sim$ km, an important caveat that we chose to be agnostic about and leave for future investigation. We present our outlook and conclusions in Sec. 3.5.

In Sec. 3.4 we discuss how our arguments could be extended to place theoretical constraints on the idea of spontaneous scalarization around black holes [101, 102].

3.1 TIME ADVANCE BOUNDS

In this section we compute the time delay that the two graviton polarizations and the massless scalar experience when scattering against a very heavy (classical) gravitational source in the eikonal regime, following [99, 149] to include the effects of the non-minimal couplings in Eq. (3.1). These interactions lead to a non-diagonal transition amplitude between graviton and

scalar, such that one of the propagation eigenmodes travels faster than what is allowed by the causal structure of the asymptotic spacetime, thus violating asymptotic causality [150]. This is analogous to the case of gravitons and photons discussed in [99, 149], where non-minimal gravitational 3-point interactions, encoded by the operators $R_{\mu\nu\rho\sigma}R^{\rho\sigma}{}_{\alpha\beta}R^{\alpha\beta\mu\nu}$ and $F_{\mu\nu}F_{\rho\sigma}R^{\mu\nu\rho\sigma}$, give rise to a mixing between the two graviton or the two photon helicities, respectively, and which results in a net macroscopic time advance for one of the propagating eigenmodes.[1] Since this happens for scattering at sufficiently small impact parameters, avoiding causality violation sets an upper bound on the cutoff of the EFT, Λ, where dynamics that is not captured by the EFT must become relevant. For recent works discussing the notion of causality in the gravitational context, we point the reader to e.g. [152–157].

Let us then consider the scattering of graviton and scalar with an spectator of mass m, within the so-called eikonal limit, $s \gg t$, where s is the center of mass energy of the collision and $t = -|\vec{q}|^2 \equiv q^2$, where $\vec{q}$ is the exchanged momentum. We take the spectator to be very heavy and nearly at rest, acting as a gravitational source against which the massless probe particle, of energy ω, scatters. In such a kinematic configuration, $m \gg \omega \gg q$, the leading contribution to the gravitational amplitude for $r_s\omega > 1$, with $r_s = m/(4\pi {M_{\rm Pl}}^2)$ the Schwarzschild radius of the target, is obtained after summing over ladder diagrams from single graviton exchange, see Fig. 3.1. The S-matrix takes an exponential form, $S = e^{i\delta(\omega,b)}$, where

$$\delta(\omega,\vec{b}) = \frac{1}{4m\omega}\int \frac{d^{D-2}q}{(2\pi)^{D-2}} e^{i\vec{q}\cdot\vec{b}}\mathcal{M}(\omega,\vec{q})\,, \tag{3.2}$$

is the eikonal phase shift and $\vec{b}$ the impact parameter [158, 159]. As we show below, the phase shift is in general a matrix in helicity space, from which, after diagonalization, one can extract the classical time delay for the propagation eigenmodes simply as $\Delta t = \partial_\omega \delta$.[2]

Let us briefly go over the time delay for a probe particle minimally coupled to gravity, that is the Shapiro time delay. The tree-level amplitude is helicity-preserving and universal,

$$\mathcal{M}^{\rm GR}_{\rm tree} \simeq \frac{1}{{M_{\rm Pl}}^2}\frac{(2m\omega)^2}{q^2}\,. \tag{3.3}$$

We can compute the associated phase shift by performing the integral Eq. (3.2) in $D-2$ dimensions, where $D = 4-2\epsilon$ is used as a regularization,

$$\delta^{\rm GR} = \frac{m\omega}{4\pi {M_{\rm Pl}}^2}\Gamma\left(\frac{D-4}{2}\right)\frac{1}{b^{(D-4)/2}} = 2\omega r_s\left(-\frac{1}{2\epsilon} - \frac{\gamma_E}{2} - \log b\right) + O(\epsilon)\,, \tag{3.4}$$

where $b \equiv |\vec{b}|$. Subtracting the time delay measured at a reference impact parameter $b_0 \gg b$, we obtain the result,

$$\Delta t_{\rm GR} = 2r_s \log(b_0/b)\,. \tag{3.5}$$

This is the Shapiro time delay for a signal traveling at an impact parameter b from a source with Schwarzschild radius r_s, as measured by an observer at an impact parameter $b_0 \gg b$.

1 Similar ideas have been considered for quadratic gravity in [151].

2 One could consider as well, as done in [99], the sub-planckian scattering against a coherent state of a large number $N \gg 4\pi {M_{\rm Pl}}^2/s$ of relativistic particles, a.k.a. shock waves.

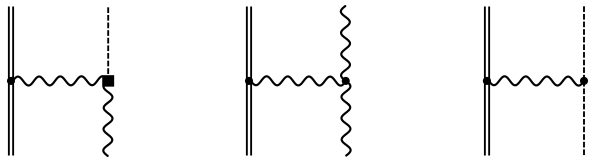

Figure 3.1: Leading tree-level diagrams for the eikonal scattering of graviton and scalar against a heavy target. Wiggly lines represent gravitons, dashed lines the massless scalar, and double lines the massive source. The square vertex corresponds to the ϕRR helicity-changing interaction.

Within GR, the leading corrections to the phase shift are of order r_s/b, associated to amplitude terms in momentum space of order q/ω, arising from the eikonal expansion as well as non-linear gravitational interactions [160]. Note that when these corrections become large, that is when $r_s \sim b$, the deflection angle of the probe, $\theta = -\omega^{-1}\partial_b \delta$, is no longer small. Let us point out as well that as long as $r_s\omega > 1$, the Shapiro time delay is larger than the quantum-mechanical uncertainty associated with the probe wave, i.e. $\Delta t_{\rm GR} > 1/\omega$.

3.1.1 *Non-minimal scalar-tensor trilinear interactions*

The (pseudo)scalar-graviton 3-point interactions associated with the $\phi\mathcal{R}^2_{\rm GB}$ and $\phi R\tilde{R}$ operators in Eq. (3.1) give rise to an eikonal phase shift that is not diagonal with respect to the helicity of the probe particle. This, along with the energy dependence of the interaction, results in time advances at energies where the EFT is still weakly coupled.

In order to compute the phase shift, we consider 4-point scattering amplitudes associated with tree-level graviton exchange between a scalar or graviton and a heavy spectator, which we take to be a scalar, S, without loss of generality. The corresponding Feynman diagrams are shown in Fig. 3.1.

Using spinor-helicity variables and taking all the particles (with complex momenta) as incoming, the relevant 3-point amplitudes read as follows:

$$\mathcal{M}^{\rm GB/CS}_{1_\phi 2_{h^{++}} 3_{h^{++}}} = \frac{2\hat{\alpha}}{M_{\rm Pl}}[23]^4\,, \quad \mathcal{M}^{\rm GB/CS}_{1_\phi 2_{h^{--}} 3_{h^{--}}} = \frac{2\hat{\alpha}^*}{M_{\rm Pl}}\langle 23\rangle^4\,, \tag{3.6}$$

where recall that $\hat{\alpha} = \alpha + i\tilde{\alpha}$ and that $\hat{\alpha} = 0$ and $\alpha = 0$ correspond to the scalar Gauss-Bonnet and dynamical Chern-Simons gravity theories, respectively. In the regime $m \gg \omega \gg q$, the relevant 4-point scattering amplitudes are given by

$$\begin{aligned}
\mathcal{M}^{\rm GB/CS}_{1_S 2_S 3_{h^{++}} 4_\phi} &= \mathcal{M}^{\rm GB/CS}_{1_S 2_S 3_\phi 4_{h^{++}}} \simeq -\frac{2\hat{\alpha}}{{M_{\rm Pl}}^2}\frac{(q_1+iq_2)^2}{q^2}(2m\omega)^2\,, \\
\mathcal{M}^{\rm GB/CS}_{1_S 2_S 3_{h^{--}} 4_\phi} &= \mathcal{M}^{\rm GB/CS}_{1_S 2_S 3_\phi 4_{h^{--}}} \simeq -\frac{2\hat{\alpha}^*}{{M_{\rm Pl}}^2}\frac{(q_1-iq_2)^2}{q^2}(2m\omega)^2\,,
\end{aligned} \tag{3.7}$$

where q_1, q_2 are the components of the exchanged momentum $\vec{q}$. Defining $b_\pm = (b_1 \pm i b_2)/2$, we have $\vec{b}\cdot\vec{q} = b_+(q_1 - iq_2) + b_-(q_1 + iq_2)$, and as before the eikonal phase shift matrix is obtained by taking the impact-parameter transform of the amplitudes,

$$\begin{aligned}
\delta^{\rm GB/CS}_{1_S 2_S 3_{h^{++}} 4_\phi} &= \delta^{\rm GB/CS}_{1_S 2_S 3_\phi 4_{h^{++}}} = -2\omega r_s\,\frac{\hat{\alpha}}{b_-^2}\,, \\
\delta^{\rm GB/CS}_{1_S 2_S 3_{h^{--}} 4_\phi} &= \delta^{\rm GB/CS}_{1_S 2_S 3_\phi 4_{h^{--}}} = -2\omega r_s\,\frac{\hat{\alpha}^*}{b_+^2}\,.
\end{aligned} \tag{3.8}$$

These helicity-changing contributions add up to the helicity-preserving ones from minimal coupling, to yield the phase shift matrix

$$\delta^{\text{GR+GB/CS}} \simeq 2\omega r_s \begin{pmatrix} D & 0 & A \\ 0 & D & A^* \\ A^* & A & D \end{pmatrix}, \tag{3.9}$$

with rows (h^{++}, h^{--}, ϕ) and

$$D = -\frac{1}{2\epsilon} - \frac{\gamma_E}{2} - \log b\,, \quad A = -\frac{\hat{\alpha}}{b_-^2}\,. \tag{3.10}$$

After diagonalizing, we find the eigenvalues

$$\begin{aligned} \delta_0 &= 2\omega r_s \left(-\frac{1}{2\epsilon} - \frac{\gamma_E}{2} - \log b\right), \\ \delta_\pm &= 2\omega r_s \left(-\frac{1}{2\epsilon} - \frac{\gamma_E}{2} - \log b \pm \sqrt{2}\frac{|\hat{\alpha}|}{b^2}\right), \end{aligned} \tag{3.11}$$

where the first corresponds to a pure graviton state, while the other two to a scalar-graviton mixed state. The time delay that the latter propagating eigenmodes acquire are

$$\Delta t_\pm = 2r_s \left(\log\frac{b_0}{b} \pm \sqrt{2}\frac{|\hat{\alpha}|}{b^2}\right). \tag{3.12}$$

At small enough impact parameters, Δt_- becomes negative, that is a time advance, signaling a potential violation of causality. Phrased in another way, for a given impact parameter there is a time advance if the GB/CS coefficient is large enough, $|\hat{\alpha}| \gtrsim b^2 \log(b_0/b)$. To avoid acausality at low energies, the EFT computation must therefore break down at distances such that this condition cannot be satisfied.[3] This implies the GB/CS coupling is parametrically bounded by the cutoff of the EFT as

$$|\hat{\alpha}| \lesssim \frac{\log(b_0\Lambda)}{\Lambda^2}\,. \tag{3.13}$$

Several comments are in order. For the violation of causality to potentially be resolvable and thus become problematic, the time advance should be larger than the quantum uncertainty of the wave-packet, $|\Delta t_-| > 1/\omega$. For impact parameters where the BGR contribution is assumed to dominate, this condition reads[4]

$$|\Delta t_-| \sim r_s \frac{|\hat{\alpha}|}{b^2} > \frac{1}{\omega}\,, \tag{3.14}$$

which for impact parameters down to the minimum cutoff length implied by Eq. (3.13), i.e. $b \sim |\hat{\alpha}|^{1/2}$ (neglecting the log), just requires $r_s\omega > 1$. Equivalently, Eq. (3.14) defines an impact parameter below which the would-be time advance is resolvable, $b_r = (r_s\omega|\hat{\alpha}|)^{1/2}$. This is larger than the minimum cutoff length within the EFT regime of validity, i.e. $b_r > |\hat{\alpha}|^{1/2}$, as long as $r_s\omega > 1$. Therefore, even if potentially resolvable, as long as Eq. (3.13) is

3 That is, at a distance $1/\Lambda > b_*$, where the largest impact parameter at which a time advance is found, b_*, is given by $|\hat{\alpha}| \sim b_*^2 \log(b_0/b_*)$. Note that the dynamics needed to restore causality on time-scales of order b_* should only involve momentum transfers $q_* \sim 1/b_*$, meaning that new physics must appear at scales $1/b_*$ or smaller, a priori regardless of ω and r_s.

4 This condition can also be interpreted as the one for which the beyond-GR contribution to the time delay is resolvable on its own.

satisfied there is never an actual time advance. Alternatively, one could also argue that the time advance is actually not resolvable at $b \sim |\hat{\alpha}|^{1/2}$ because the UV completion precludes $r_s\omega > 1$, which in practice means a cutoff such that $\omega \lesssim \Lambda \lesssim 1/r_s$. In this case, the condition $|\Delta t_-| < 1/\omega$ for $\omega \sim \Lambda$ and $b \sim r_s \sim 1/\Lambda$ implies $\Lambda \lesssim |\hat{\alpha}|^{1/2}$, just like in Eq. (3.13) up to $O(1)$ factors, which in any case we are oblivious about.

The bound Eq. (3.13) depends on the logarithm of an unspecified scale $b_0 \gg b$, because of the infrared (IR) divergent nature of gravity in four dimensions. While the identification of IR-finite scattering observables in gravity remains an open and interesting problem (see e.g. [161] for a recent attempt in the context of causality constraints), we do not regard this divergence as a serious drawback that invalidates our bounds. In fact, note that even if we consider an IR scale of the order of the size of the observable universe, we merely find $\log(\Lambda/H_0) \sim 50$.

Finally, let us point out that $\Lambda \lesssim |\hat{\alpha}|^{-1/2}$ implies that the EFT description must break down at energies much lower than the strong coupling scale associated with the GB/CS interactions. Indeed, the scale where the trilinear scalar-tensor coupling becomes large, indicated by e.g. the 4-graviton amplitude mediated by the scalar becoming strong, $\mathcal{M} \sim (\hat{\alpha}/M_{\rm Pl})^2 E^6 \sim 1$, is

$$\Lambda_\alpha = \left(\frac{M_{\rm Pl}}{|\hat{\alpha}|}\right)^{1/3}, \tag{3.15}$$

much larger than the actual cutoff of the EFT, unless $|\hat{\alpha}| \sim 1/{M_{\rm Pl}}^2$.

3.1.2 *Robustness of superluminality*

In order to understand whether the presence of superluminal excitations is a feature common to all models including a sizable sGB/CS interaction, we have to consider the possible effects due to other operators in the EFT as well as the backreaction of the background. A key observation to answer this question is that the only diagrams that exponentiate are those involving at least one graviton exchange. To understand this, it is enough to notice that $p_3 \cdot p_4 \sim q^2$. Moreover the contributions to the phaseshift that alter the time advance are those at least linear in ω. This implies that the only diagrams that can exponentiate to alter the time advance in the eikonal limit are those in which the indices of two powers of p_3 and p_4 are contracted with the indices of the graviton propagator, so that they can contribute as $(m\omega)^2$ to the amplitude.

This for instance means that both scalar-scalar-graviton vertices as well as the backreaction due to the scalar hair on the probe are negligible, being suppressed by $(q/\omega)^2$.

Similarly, diagrams involving one graviton exchange and more irrelevant graviton-graviton-scalar interactions are suppressed by higher powers of q^2, since the additional derivatives cannot be contracted with further powers of $p_{1/2}$. Diagrams with more than one graviton exchange will instead contribute at higher order in the Post-Minkowskian (PM) expansion, being suppressed by higher powers of r_s/b.

With these considerations, we conclude that the superluminal propagation is a common feature of any EFT including the most relevant shift-symmetric scalar-graviton-graviton interactions.

Even more, we can briefly consider the case of a spinning black hole with spin parameter a (a length). In this case, the leading corrections to the scattering amplitude will be suppressed by factors of a/b [162]. For a black hole

it will always hold $a < r_s$, meaning that the superluminality persists for spinning black holes. Interestingly, in the case of naked singularities, the spin parameter might be large enough to compete with or even dwarf the sGB/CS contribution, so that the same conclusion might not hold.

3.1.3 *Causality bounds on power counting*

In this section we reinterpret the causality constraints in terms of bounds on the power counting of gravitational EFTs. With this aim, let us consider the generic form of a scalar theory coupled to gravity, in which the heavy degrees of freedom, of mass Λ or higher (i.e. Λ is the EFT cutoff), have been integrated out

$$\mathcal{L} = \tfrac{1}{2}\hat{M}_{Pl}^2 R + \frac{\Lambda^4}{g^2} L^{(0)}\left(\frac{\nabla_\mu}{\Lambda}, \frac{\zeta R_{\mu\nu\rho\sigma}}{\Lambda^2}, \frac{g\phi}{\Lambda}\right) + \dots . \tag{3.16}$$

In the spirit of naive dimensional analysis (NDA) each covariant derivative ∇_μ is weighted by $1/\Lambda$, and each (scalar) field ϕ by g/Λ. The coupling g parametrizes the strength with which the heavy states couple to the light degrees of freedom, with $g \sim 4\pi$ the usual non-perturbative coupling limit. Note that instead of considering the Riemann tensor, $R_{\mu\nu\rho\sigma} \sim \partial_\mu\partial_\nu h_{\rho\sigma}$, simply as a two-derivative object thus weighted by $1/\Lambda^2$, we introduce a dimensionless parameter ζ to allow for the possibility that gravitational interactions beyond GR's minimal coupling are enhanced w.r.t. standard NDA.[5] We will elaborate on such a generalized power counting below. Each ϕ interaction comes with a decay constant f, identified with (or defined as)

$$f = \frac{\Lambda}{g}\,. \tag{3.17}$$

At this point we can already distinguish the two interesting scenarios, for which it is enough to consider the standard power counting $\zeta = 1$ and to realize that the EH action receives a contribution from both terms in Eq. (3.16). When $\hat{M}_{Pl}^2 \gg f^2$, the EH action is dominated by the first term and the effective Planck scale is $M_{\rm Pl} \sim \hat{M}_{Pl}$. Gravity is external to the ultraviolet (UV) dynamics giving rise to $L^{(0)}$, a.k.a. "elementary". Instead, when $\hat{M}_{Pl}^2 \ll f^2$, we have $M_{\rm Pl} \sim f$ and the heavy dynamics constitutes a bona fide UV completion of gravity. Phrasing it in terms of the coupling g, the minimum coupling $g \sim \Lambda/M_{\rm Pl}$ corresponds to the "composite" limit of gravity. This is the case of string theory (or more generally, potential tree-level UV completions with infinitely many higher-spins particles, see e.g. [99, 163]), where $\Lambda \sim M_s$ the string scale, as well as of loop-level completions based on a large number of species, $N \sim (4\pi M_{\rm Pl}/\Lambda)^2$, where $g \sim 4\pi/\sqrt{N}$ [164, 165]. Note that in this limit one finds the largest coefficients for gravitational EFT operators with none or a single matter field, since they scale as $1/g^2$ or $1/g$ respectively. The fact that $g \gtrsim \Lambda/M_{\rm Pl}$ is reminiscent of the weak gravity conjecture [166].

In terms of scattering amplitudes, the two scenarios are distinguished by the maximal size of e.g. 2-to-2 graviton processes within the EFT regime of validity, i.e. $E \lesssim \Lambda$. From minimal coupling we have $\mathcal{M}^{\rm GR} \sim (E/M_{\rm Pl})^2 \lesssim (\Lambda/M_{\rm Pl})^2$. Instead, an effective operator like $R^3_{\mu\nu\rho\sigma}$ leads to an amplitude

5 As usual we work with a dimensionless graviton field, whose interactions are eventually weighted by $1/M_{\rm Pl}$ once its kinetic term is canonically normalized, following the normalization of the Einstein-Hilbert (EH) term.

$\mathcal{M}^{\rm BGR}_{(\zeta=1)} \sim E^6/(g^2\Lambda^2 {M_{\rm Pl}}^4) \lesssim f^2\Lambda^2/{M_{\rm Pl}}^4$, smaller than GR except in the limit $M_{\rm Pl} \sim f$, in which case the two amplitudes are of the same size at the cutoff. The same analysis can be reproduced if instead of amplitudes one considers other (classical) gravitational observables and distances, rather than energies, within EFT control, i.e. $r \gtrsim 1/\Lambda$.

For the generalized power counting $\zeta > 1$, the discussion is very much analogous, except for the important difference that now the BGR effects can become larger than the GR prediction for energies well described by the EFT. The composite case corresponds to $\hat{M}_{Pl}^2 \ll \zeta f^2$, for which we have $M_{\rm Pl} \sim \sqrt{\zeta} f$. Therefore, $\zeta \ll (M_{\rm Pl}/f)^2$ corresponds to the case where gravity is external to the UV dynamics. Elementary or composite, we find that non-standard gravitational interactions, in the form of $R^3_{\mu\nu\rho\sigma}$, give rise to enhanced 4-graviton amplitudes

$$\mathcal{M}^{\rm BGR}_{1_h 2_h 3_h 4_h} \sim \zeta^3 \frac{E^6 f^2}{\Lambda^4 {M_{\rm Pl}}^4} \lesssim \zeta^3 \frac{\Lambda^2 f^2}{{M_{\rm Pl}}^4} \lesssim \frac{\Lambda^2 {M_{\rm Pl}}^2}{f^4} = \left(\frac{g M_{\rm Pl}}{f}\right)^2 , \tag{3.18}$$

where the first inequality follows from $E \lesssim \Lambda$ and the second from $\zeta \lesssim (M_{\rm Pl}/f)^2$. Note that for $\zeta \gtrsim (M_{\rm Pl}/f)^{2/3}$ the amplitude is larger than in GR, and it becomes non-perturbatively strong, i.e. $\mathcal{M} \sim (4\pi)^2$, for EFT cutoffs well below the maximal gravity cutoff given by $4\pi M_{\rm Pl}$. As we discuss in the following, it is precisely this possibility that causality constraints forbid.[6]

Let us start by recalling that each specific UV theory within the class of theories described in the IR by Eq. (3.16) comes with $O(1)$ factors not captured by the power counting. Even more importantly, the presence of symmetries can enforce some EFT operators to have vanishing coefficients, for instance if ϕ is Nambu-Goldstone boson with a shift symmetry $\phi \to \phi + c$ (as the scalar field that concerns us in this work), any potential term for ϕ vanishes. However, beyond the well-known selection rules from symmetries, there are further requirements that an EFT must satisfy if it is to be consistent with the fundamental principles of unitarity, locality, and causality (and if it is to arise from UV dynamics that abides by such principles). Indeed, it was found in [99] that causality, in the form of absence of a (resolvable) time advance, leads to a constraint on the size of corrections to the cubic graviton coupling, arising from an operator $\alpha_3 {M_{\rm Pl}}^2 R^3_{\mu\nu\rho\sigma}$, given by $\alpha_3 \lesssim 1/\Lambda^4$. In terms of the power counting Eq. (3.16), $\alpha_3 \sim \zeta^3/(g M_{\rm Pl}\Lambda)^2$, such a bound implies $\zeta \lesssim (M_{\rm Pl}/f)^{2/3}$, precisely such that the BGR effects never get to dominate over GR, see below Eq. (3.18). This conclusion seems to be generic. The similar bound we have derived in Sec. 3.1.1 on the GB/CS non-minimal coupling of gravitons to a scalar, $|\hat{\alpha}| \lesssim 1/\Lambda^2$, when interpreted in terms of our power counting, $|\hat{\alpha}| \sim \zeta^2/(g M_{\rm Pl}\Lambda)$, implies $\zeta \lesssim (M_{\rm Pl}/f)^{1/2}$. Once again, this forbids the 4-graviton amplitude mediated by the scalar from getting larger than in GR if restricted to energies within the EFT, $E \lesssim \Lambda$,

$$\mathcal{M}^{\rm GB/CS}_{1_{h^{++}} 2_{h^{++}} 3_{h^{--}} 4_{h^{--}}} \sim |\hat{\alpha}|^2 \frac{E^6}{{M_{\rm Pl}}^2} \lesssim \frac{\Lambda^2}{{M_{\rm Pl}}^2} \,. \tag{3.19}$$

6 It is perhaps instructive to compare with EFTs for spin-1 (abelian or non-abelian) gauge fields,

$$\mathcal{L} = \tfrac{1}{4\hat{e}^2} F_{\mu\nu}^2 + \frac{\Lambda^4}{g^2} L^{(0)}\left(\frac{D_\mu}{\Lambda}, \frac{\zeta F_{\mu\nu}}{\Lambda^2}, \frac{g\phi}{\Lambda}\right) ,$$

with the elementary and composite limits given respectively by $\hat{e} \ll g/\zeta$ (and effective gauge coupling $e \sim \hat{e}$) and $\hat{e} \gg g/\zeta$ ($e \sim g/\zeta$). As discussed in [167], in the strongly coupled gauge field scenario $\zeta \sim g/e \gtrsim 1$ one finds 4-point amplitudes (from e.g. $F^4_{\mu\nu}$ operators) $\mathcal{M} \sim g^2 (E/\Lambda)^4$, which can be larger than the amplitude from minimal coupling, $\mathcal{M} \sim e^2$, for energies within the EFT.

It is illuminating to realize that in the case of a standard power counting $\zeta = 1$, these causality constraints robustly imply that $g \gtrsim \Lambda/M_{\rm Pl}$ (or equivalently $f \lesssim M_{\rm Pl}$), as we expected from the simple NDA considerations on the elementary vs composite nature of gravity. In turn, if one is interested in genuine UV completions of gravity, i.e. $g \sim \Lambda/M_{\rm Pl}$, these bounds imply that $\zeta \lesssim 1$ and therefore that the EFTs in which non-minimal interactions are enhanced beyond standard NDA have no gravitational completions consistent with fundamental principles.

3.1.4 *Bounds from dispersion relations*

This conclusion is reinforced by recent progress on the derivation of theoretical constraints on gravitational EFTs that go beyond causality violation in classical observables and therefore beyond corrections to cubic gravitational interactions [116, 118, 121, 127, 129, 161, 168–174]. Such bounds are instead obtained via dispersion relations [175], which connect the coefficients of the EFT operators to the dynamics of their UV completions. These UV/IR relations, which we will review in some detail in Sec. 3.2.1, are very powerful because of their generality, relying only on the basic assumptions of unitarity, locality and causality (encoded as the analyticity, crossing symmetry and boundedness of the scattering amplitudes).[7] Of particular relevance for the physics of black holes are the results of [121], which derived lower bound on $\alpha_4 {M_{\rm Pl}}^2 R^4_{\mu\nu\rho\sigma}$ given by $\alpha_4 \gtrsim \alpha_3^2 \Lambda^2$ (recall $\alpha_3 {M_{\rm Pl}}^2 R^3_{\mu\nu\rho\sigma}$), and of [127], which derived the upper bound $\alpha_4 \lesssim 1/\Lambda^6$. Both constraints restrict the BGR contribution to gravitational observables to be smaller than the prediction of GR. In fact, we should stress that if similar bounds were to be derived on non-standard higher-point amplitudes (with $n \geqslant 5$ gravitons) from $R^n_{\mu\nu\rho\sigma}$ operators, we would be led to the conclusion that the power counting in Eq. (3.16) with $\zeta > 1$ is inconsistent altogether, i.e. regardless of f and not only for $f \sim M_{\rm Pl}$. While this seems like a plausible expectation, a robust derivation of theoretical constraints on higher-point amplitudes remains an open problem at the time of writing this work (see e.g. [176] for recent progress in this direction). If indeed $\zeta > 1$ is forbidden by fundamental principles, we would come to the sensible conclusion that in a gravitational EFT the largest effects for a fixed cutoff are found when $f \sim M_{\rm Pl}$, therefore when gravitational interactions should dramatically change above Λ. We will provide further insight into this fact in Sec. 3.5.

Before concluding this section, let us make some additional comments on the implications of the causality bounds on the phenomenology of black holes beyond GR. The constraint $\alpha_4 \lesssim 1/\Lambda^6$ [127] places modifications of GR due to quartic terms in the curvature on the same footing as those due to cubic terms. This means that there is no strong reason to discard the effects of $R^3_{\mu\nu\rho\sigma}$ operators, leading in the derivative expansion, while keeping those of $R^4_{\mu\nu\rho\sigma}$ [177–179].

The constraint we have obtained in Sec. 3.1.1 on BGR scalar-tensor cubic couplings have also been recently derived in [127] via dispersion relations. In this regard, it is important to point out that even though our bound is robust up to $O(1)$ factors, contrary to the more precise (yet still IR divergent) one from dispersion relations, we believe our derivation is very valuable because it comes from a simple physical setup in which causality violation

7 The link between dispersion relations and causality, expressed as the absence of superluminal propagation, was pointed out in [108], and its connection with time delay has been recently discussed in [116, 127].

is a classical, macroscopic effect. This makes possible to avoid relying on a priori stronger assumptions on the analyticity and polynomial boundedness of scattering amplitudes associated with causality. Moreover extracting the classical phase-shift simplifies the analysis of light (massless) loops in the EFT, which can be problematic in the context of dispersion relations, forcing in some cases to neglect these loops all-together, see e.g. [118, 121]. We will review these issues and present a strategy to include light loops in Sec. 3.2.2.

NDA expectations are also confirmed by dispersion relations involving operators with extra derivatives acting on the curvature, for instance of the form

$$\alpha_5 {M_{\rm Pl}}^2 R^2_{\mu\nu\rho\sigma}(\nabla_\eta R_{\mu\nu\rho\sigma})^2, \quad \alpha_6 {M_{\rm Pl}}^2(\nabla_\eta R_{\mu\nu\rho\sigma})^4, \tag{3.20}$$

which contribute to 4-graviton amplitudes as $\mathcal{M} \sim \alpha_J E^{2J}$ [121, 127, 174]. In particular, there are an infinite number of linear constraints on the EFT coefficients that take the form of two-sided bounds such as

$$-\alpha_4 \leqslant \alpha_5 \Lambda^2 \leqslant \alpha_4 \quad \text{and} \quad 0 \leqslant \alpha_6 \Lambda^4 \leqslant \alpha_4 \,, \tag{3.21}$$

and similar bounds for higher J, respectively odd or even.

Their interpretation in terms of a power counting is clear, subleading operators in the derivative expansion ∇/Λ cannot be enhanced over the leading ones.

In our discussion we have focused on cubic and quartic operators built out of the Riemann tensor, with no mention of terms quadratic in the curvature. This is because $R^2_{\mu\nu\rho\sigma}$ operators do not contribute to graviton scattering amplitudes, given that the GB term $\mathcal{R}^2_{GB}$ is a topological invariant (in $D = 4$) and because field redefinitions can be performed to eliminate any EFT operator built out of R and $R_{\mu\nu}$ in favor of matter terms (T and $T_{\mu\nu}$), therefore giving rise to amplitudes involving ϕ fields. For this reason, one might find it more convenient (although not necessary) to use a basis of EFT operators directly link to scattering amplitudes, such as the one systematically constructed in [180]. In this respect, note that the relevant object giving rise to processes with gravitons on-shell is the Weyl tensor, $C_{\mu\nu\rho\sigma} = R_{\mu\nu\rho\sigma} - (g_{\mu[\rho}R_{\sigma]\nu} - g_{\nu[\rho}R_{\sigma]\mu}) + \frac{1}{3}g_{\mu[\rho}g_{\sigma]\nu}R$. This means that the relevant parts of the GB/CS operators in Eq. (3.1) are $M_{\rm Pl}\alpha\,\phi C_{\mu\nu\rho\sigma}C^{\mu\nu\rho\sigma}$ and $M_{\rm Pl}\tilde{\alpha}\,\phi C_{\mu\nu\rho\sigma}\tilde{C}^{\mu\nu\rho\sigma}$. These are also the terms behind the scalar hair of black holes, since black holes are Ricci-flat gravitational solutions ($R, R_{\mu\nu} = 0$) at zeroth order in α, $\tilde{\alpha}$.

Finally, causality bounds on pure scalar operators are also relevant for the physics of hairy black holes, in particular

$$\tfrac{1}{4}c_2(\nabla_\mu\phi)^4 \,, \tag{3.22}$$

i.e. the leading operator in the derivative expansion. Several recent works on dispersion relations that incorporate gravity have argued that $c_2 \gtrsim -1/(\Lambda^2 {M_{\rm Pl}}^2)$ (and likewise for the equivalent operator $F^4_{\mu\nu}$ in a theory of photons) [118, 161, 168, 170, 171, 181], bound that becomes a standard positivity constraint, $c_2 > 0$ [108], when gravity decouples, $M_{\rm Pl} \to \infty$. In particular, [118] has shown via dispersion relations at finite impact parameter that this is indeed the case up to a $\log(b_0\Lambda)$, as in Eq. (3.13), when light loop corrections in the EFT are neglected. Furthermore, the upper bound $c_2 \lesssim (4\pi)^2/\Lambda^4$ has been derived using similar techniques [117, 182]. These constraints on c_2 can be easily understood in terms of the power counting in Eq. (3.16). Since $c_2 \sim g^2/\Lambda^4$, the upper and lower bounds correspond, respectively, to the maximum coupling in the spirit of NDA, $g \lesssim 4\pi$, and

to the minimum coupling in a gravitational theory, $g \gtrsim \Lambda/M_{\rm Pl}$ (recall that power counting estimates are insensitive to the sign of the operators' coefficient).

3.2 SIGNS OF UV COMPLETION

In the previous section we have argued that gravitational EFTs where black holes have scalar hair, Eq. (3.1), must have a cutoff $\Lambda \lesssim |\hat{\alpha}|^{-1/2}$. In terms of the power counting Eq. (3.16) with $\zeta = 1$, the maximum cutoff of these theories corresponds to the minimum NDA coupling $g \sim \Lambda/M_{\rm Pl}$, while EFTs with a larger coupling, or equivalently $f < M_{\rm Pl}$, must have a lower cutoff for the same value of $|\hat{\alpha}|$.

In this section we try to infer from an EFT point of view what additional low-energy effects are associated with a generic UV completion at the scale Λ, in particular one that is unitary, local and casual. We focus on the leading corrections in the derivative and field expansion [180], restricted to CP even operators

$$\Delta S = \int d^4x\sqrt{-g}\left[{M_{\rm Pl}}^2\left(\alpha_3\mathcal{I}+\alpha_4\mathcal{C}^2+\alpha_4'\tilde{\mathcal{C}}^2\right)+\frac{c_2}{4}(\nabla_\mu\phi)^4+\frac{d_1}{2}\mathcal{C}(\nabla_\mu\phi)^2\right], \tag{3.23}$$

where $\mathcal{C} = R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$, $\tilde{\mathcal{C}} = R_{\mu\nu\rho\sigma}\tilde{R}^{\mu\nu\rho\sigma}$ and $\mathcal{I} = R_{\mu\nu}{}^{\rho\sigma}R^{\mu\nu\alpha\beta}R_{\alpha\beta\rho\sigma}$. Note that the last operator is equivalent, by the leading-order scalar equation of motion, to the cubic Galileon term $(\nabla\phi)^2\Box\phi$. We first investigate the constraints on the coefficients above that arise from dispersion relations at one loop, which take the form of lower bounds that depend on $|\hat{\alpha}|$.[8] Precisely because of the upper bound $|\hat{\alpha}| \lesssim 1/\Lambda^2$, we find that such dispersion relations are in fact dominated by standard gravitational contributions, rendering the constraints on the operators in Eq. (3.23) inapplicable and phenomenologically irrelevant. We therefore leave aside general bounds and turn to generic expectations based on power counting. We show how typical UV completions of the scalar-GB or dynamical-CS theories likely give rise to higher-curvature terms of the same parametric size, i.e. $\alpha\Lambda^2 \sim \alpha_3\Lambda^4 \sim \alpha_4\Lambda^6$, if these arise at the same loop order.

3.2.1 *Beyond positivity constraints*

There exists an extensive literature on dispersion relations, in particular on non-gravitational theories, with many new results and applications found in recent years, see e.g. [108, 110–116, 119]. Dispersion relations are typically constructed by evaluating a 2-to-2 scattering amplitude $\mathcal{M}(s,t)$ over a closed circular contour in the complex s-plane,[9]

$$\Sigma_n(s,t) = \frac{1}{2\pi i}\oint_{\Gamma_s} ds' \frac{\mathcal{M}(s',t)}{\left(s'+\frac{t}{2}\right)^{n+1}}\,. \tag{3.24}$$

8 There are no constraints of this form from dispersion relations at tree level. This is in contrast with the lower bound $\alpha_4 \gtrsim \alpha_3^2\Lambda^2$ [121].

9 In a slight abuse of notation, we denote the amplitude as a function of the $s = -(p_1+p_2)^2$ and $t = -(p_1+p_3)^2$ Mandelstam variables as $\mathcal{M}$, like in Sec. 3.1 where instead was a function of ω and $\vec{q}$. Besides, we work with all momenta incoming and recall $u = -(p_1+p_4)^2 = -s-t+4m^2$, where m is now the mass of the scattered states, that we will eventually take to zero.

Let us start with the scattering of the GB/CS scalar ϕ at low energies, $s \ll \Lambda^2$, neglecting for the time being GR's minimal gravitational coupling. The 4-scalar EFT interaction in Eq. (3.23) leads to an amplitude,

$$\mathcal{M}^{\Delta S}_{1_\phi 2_\phi 3_\phi 4_\phi}(s,t) = \frac{c_2}{2}(s^2+t^2+u^2)\,, \tag{3.25}$$

which grows like s^2 for fixed t. Therefore, considering a small contour Γ_0 around $s' = -t/2$, the integral of the twice-subtracted amplitude, i.e. $n = 2$ in Eq. (3.24), yields $\Sigma_2(0,t) = c_2$. At this point, unitarity and analyticity in the upper half of the complex plane allow one to deform the contour away from the origin (for $0 \leqslant t \leqslant 4m^2$) in a controlled way. As discussed in the introduction to this part of the work, the analogy with quantum mechanics strongly suggests that causality is responsible for $\mathcal{M}(s,t)$ being analytic everywhere except in the real axis. On the real axis, as implied by unitarity, one finds singularities in the form of simple poles and branch cuts. The former correspond to particles exchanged at tree level going on shell, which in the case at hand belong only to the UV completion, either in the s-channel at $s \geqslant \Lambda^2$ or in the u-channel $s \leqslant -\Lambda^2 - t$. The branch cuts are associated with logarithms arising from loops and correspond to multi-particle production. Besides the loops of heavy states at and above the cutoff, there is an s-channel branch cut starting at $s = 4m^2$ from (one) loop diagrams of the IR degrees of freedom, and its $s \leftrightarrow u$ crossing symmetric counterpart. Because of real analyticity, $\mathcal{M}^*(s,t) = \mathcal{M}(s^*,t)$, these discontinuities are proportional to $Im\mathcal{M}(s,t)$, which is positive for elastic scattering around $t = 0$. In particular, for the zeroth-order term in an expansion around the forward limit, the optical theorem fixes $Im\mathcal{M}(s,0) = s\sqrt{1-4m^2/s}\,\sigma_T(s)$, where σ_T is the total cross section for $1\,2 \to$ *everything*. In addition, unitarity and analyticity in theories with a mass gap imply that amplitudes are polynomially bounded as $\mathcal{M}(s,t)/s^2 \to 0$ for $|s| \to \infty$ as a result of the Froissart-Jin-Martin bound [105, 106]. Even though here we are interested in theories with a massless graviton, it has been argued from different perspectives that a growth smaller than s^2 holds as well with dynamical gravity, see e.g. [99, 116, 118, 183, 184]. Note then that when this is the case, the integral Eq. (3.24) over a contour Γ_∞ at $|s| \to \infty$ vanishes for $n \geqslant 2$, i.e. $\Sigma_{n\geqslant 2}(\infty,t) = 0$. A dispersion relation is then finally derived by using Cauchy's theorem to deform the original contour Γ_0 to Γ_∞, leaving Σ_n as an integral over the aforementioned singularities. In the forward limit $t = 0$ of 4-scalar scattering, one then finds for $n = 2$:

$$c_2 = \sum_X \frac{2}{\pi}\int_0^\infty \frac{ds}{s^2}\sigma_{1_\phi 2_\phi \to X}(s) > 0\,. \quad \text{(w/o GR minimal coupling)} \tag{3.26}$$

This positivity constraint can be improved by noticing that, while the cross sections for production of the heavy states associated with the UV completion are by construction not computable within the EFT, those for production of the low-energy states are, as long as we restrict them to energies $s \leqslant \Lambda^2$ [109, 112, 185–187]. Since we are neglecting the minimal coupling of gravitons, the process with the largest cross section in the scalar EFT Eq. (3.1) is the production of a pair of gravitons via the GB/CS coupling. The corresponding amplitude is

$$\mathcal{M}^{\text{GB/CS}}_{1_\phi 2_\phi 3_{h^{++}} 4_{h^{--}}} = \left(\frac{2|\hat{\alpha}|}{M_{\text{Pl}}}\right)^2 \langle 4|1|3]^4 \left(\frac{1}{t}+\frac{1}{u}\right)\,. \tag{3.27}$$

Explicitly including this contribution in the twice-subtracted dispersion relation Eq. (3.26), we arrive at

$$c_2 > \frac{2}{\pi}\int_0^{\Lambda^2}\frac{ds}{s^2}\sigma^{\rm GB/CS}_{\phi\phi\to h^{--}h^{++}} = \frac{1}{60\pi^2}\left(\frac{|\hat{\alpha}|\Lambda^2}{M_{\rm Pl}}\right)^4 . \text{(w/o GR minimal coupling)} \tag{3.28}$$

If one could ignore GR's contributions to the dispersion relation, as we have done this far, such a beyond-positivity bound would imply that the GB/CS scalar-tensor theories in Eq. (3.1) are inconsistent with unitarity and causality unless they are supplemented with the $(\nabla\phi)^4$ operator. In particular, note that the larger the regime of validity of the EFT, i.e. the larger the cutoff Λ, the larger its coefficient c_2 would have to be.[10] However, neglecting GR's interactions would require in practice the existence of a consistent decoupling limit in which $M_{\rm Pl}\to\infty$ yet the lower bound on c_2 remains non-zero. Expressing Eq. (3.28) in terms of the strong coupling scale Eq. (3.15), $c_2 \gtrsim \frac{1}{16\pi^2}(\Lambda/\Lambda_\alpha)^8\Lambda_\alpha^{-4}$, this would require keeping Λ_α as well as Λ fixed. However, precisely because of the causality bound we derived in Sec. 3.1, $\Lambda \lesssim 1/|\hat{\alpha}|^{1/2}$ for $\log(b_0\Lambda)\sim 1$ (or equivalently $\Lambda\lesssim(\Lambda_\alpha^3/M_{\rm Pl})^{1/2}$), such a limit is not possible: if $M_{\rm Pl}\to\infty$, then either $\Lambda_\alpha\to\infty$ or $\Lambda\to 0$, rendering the EFT invalid. In fact, even if one saturates the upper bound on $|\hat{\alpha}|$, the beyond-positivity contribution to c_2 in Eq. (3.28) is only as large as a quantum correction in GR at one loop, i.e. $c_2\gtrsim\frac{1}{16\pi^2}{M_{\rm Pl}}^{-4}$.

We can explicitly check that one cannot ignore GR's minimal coupling if the upper bound $\Lambda\lesssim 1/|\hat{\alpha}|^{1/2}$ holds by retaking the steps above keeping $t\neq 0$ and with the low-energy contour now enclosing the graviton pole, at $s=0$ and $s=-t$ ($u=0$), of the 4-scalar amplitude in GR,

$$\mathcal{M}^{\rm GR}_{1_\phi 2_\phi 3_\phi 4_\phi} = -\frac{1}{2{M_{\rm Pl}}^2}\left(\frac{t^2+u^2}{s}+\frac{s^2+u^2}{t}+\frac{t^2+s^2}{u}\right) . \tag{3.29}$$

Note that the forward limit is ill-defined because of the graviton t-channel exchange. The $n=2$ dispersion relation then reads

$$-\frac{1}{{M_{\rm Pl}}^2 t}+c_2+\beta_2^{(t)}\log\frac{t}{t_0}+O(t) = \frac{2}{\pi}\int_0^\infty ds\frac{Im\mathcal{M}(s',t)}{\left(s+\frac{t}{2}\right)^3} . \tag{3.30}$$

We have included the one-loop UV divergence of the s^2 term of the 4-scalar amplitude arising from t-channel cuts, with β-function given by $\beta_2^{(t)} = +(13/160\pi^2){M_{\rm Pl}}^{-4}$. This is of the same loop order as the r.h.s. of Eq. (3.28). Indeed, as discussed in [119], the beyond-positivity contributions to the dispersion relation are equivalent to including the running of the coefficients of the EFT in the forward limit, associated with the UV divergences from s- and u-channel cuts. These cuts and the corresponding gravitational β-functions can be easily computed following the on-shell amplitude techniques presented in [188]. The $O(t)$ term in Eq. (3.30) encodes the subleading terms in the forward limit, arising from e.g. higher-order EFT operators in the derivative expansion. For instance, the first such correction comes from the Galileon-like term $(\nabla\phi)^2(\nabla\nabla\phi)^2$, which gives rise to an stu term in the amplitude. Most importantly, as advanced at the end of Sec. 3.1.3, the

10 In Sec. 3.3 we discuss the effects of the leading additional operators in Eq. (3.23) on black holes with scalar hair. From that analysis one can arrive at the conclusion that for astrophysical black holes where $r_s\sim|\hat{\alpha}|^{1/2}\sim$ km, the effects of $(\nabla\phi)^4$ with c_2 fixed by Eq. (3.28) would become $O(1)$, thus as important as the GB/CS term, for $\Lambda\gtrsim\mu{\rm m}^{-1}$, precisely of the same order as the smallest scales where gravity has been experimentally tested [147] and at least up to which one would want any BGR theory to hold.

$1/t$ term in Eq. (3.30) precludes setting a positive lower bound on c_2 as the one in Eq. (3.28), unless the beyond-positivity contributions are larger than $-(M_{\rm Pl}{}^2 t)^{-1} \gtrsim (M_{\rm Pl}\Lambda)^{-2}$ [118]. As we discussed above, this is not the case because of the causality bound $|\hat{\alpha}| \lesssim 1/\Lambda^2$.

While it might naively seem from the discussion above that the main obstruction for the derivation of meaningful beyond-positivity bounds in gravitational EFTs is the t-channel graviton pole, the real reason for their ineffectiveness is the fact that BGR amplitudes larger than in GR are not consistent with causality. To show this, let us consider a dispersion relation for the 4-graviton amplitude with two positive and two negative helicities. The amplitude in GR plus the leading BGR correction in the energy expansion from Eq. (3.23) is given by

$$\begin{aligned} \mathcal{M}^{\rm GR+\Delta S}_{1_{h^{++}}2_{h^{--}}3_{h^{--}}4_{h^{++}}}(s,t) &= \frac{\langle 23\rangle^4[14]^4}{M_{\rm Pl}{}^2} f(s,t)\,, \\ f(s,t) &= \frac{1}{stu} + 8(\alpha_4+\alpha_4')\,. \end{aligned} \tag{3.31}$$

Similarly to the scalar case, one can construct dispersion relations from the contour integral (see e.g. [116, 121, 127, 129] for more details)

$$\frac{1}{2\pi i}\oint_{\Gamma_s} ds' \frac{f(s',t)}{\left(s'+\frac{t}{2}\right)^{n+1}}\,. \tag{3.32}$$

In particular, for $n=0$ one arrives at

$$\begin{aligned} \frac{8(\alpha_4+\alpha_4')}{M_{\rm Pl}{}^2} + \frac{\gamma_4}{\Lambda^2 t}\log\frac{-t}{\mu^2} + O(t) &> \frac{2}{\pi}\int_0^{\Lambda^2}\frac{ds}{s^4}\sigma^{\rm GB/CS}_{h^{++}h^{--}\to\phi\phi,\,h^{--}h^{++}} \\ &\sim \frac{1}{16\pi^2}\left(\frac{|\hat{\alpha}|\Lambda}{M_{\rm Pl}}\right)^4. \end{aligned} \tag{3.33}$$

On the r.h.s. we have explicitly included the beyond-positivity contribution from a scalar as well as a graviton loop via the GB/CS coupling, computed in a dispersive way from the corresponding cross sections. The corresponding amplitudes are given by Eq. (3.27) and by

$$\mathcal{M}^{\rm GB/CS}_{1_{h^{++}}2_{h^{--}}3_{h^{++}}4_{h^{--}}} = -\left(\frac{2|\hat{\alpha}|}{M_{\rm Pl}}\right)^2\frac{\langle 24\rangle^4[13]^4}{t}\,, \tag{3.34}$$

which proceeds via scalar exchange. While there is no contribution from tree-level graviton exchange in Eq. (3.33), further forward limit singularities are generated at one loop in GR, with $\gamma_4 \sim +(1/16\pi^2)M_{\rm Pl}{}^{-4}$ [121]. Therefore, since the time-delay constraint $|\hat{\alpha}| \lesssim 1/\Lambda^2$ sets an upper bound on the r.h.s. $\lesssim \frac{1}{16\pi^2}(\Lambda M_{\rm Pl})^{-4}$, the beyond-positivity contribution is no larger than the one of GR, rendering the former immaterial to bound the quartic curvature operators.

In summary, because causality demands that gravitational amplitudes within the EFT domain are dominated by GR, loop corrections from BGR interactions never lead to robust lower bounds on the coefficients of the EFT.

3.2.2 *Improved dispersion relations at loop level*

Besides leading to the conclusion highlighted above, that GR loops dominate over the BGR ones as a result of causality, Eq. (3.33) highlights the

problematic role played by light loops in dispersion relations. Indeed, in order to neglect the $O(t)$ terms, one would like to restrict to a regime in which $t \ll \Lambda^2$. This however would enhance dramatically the GR's loop contribution on the l.h.s. of Eq. (3.33), making it difficult to find a sharp bound from below on the allowed negativity for $\alpha_4 + \alpha_4'$. Here we briefly outline a new method to deal with this issue, which is the subject of an upcoming work [130].

An approach valid at tree level to derive sharp constraints from dispersion relations without taking the $t \to 0$ limit, was proposed in [118] in the context of amplitudes affected by the graviton $t-$channel pole. This approach consists in combining different dispersion relations to obtain an improved dispersion relation, in which almost all the $O(t)$ terms have been subtracted, leaving a known polynomial in t with a finite number of terms. This allows to have control over the $O(t)$ terms and extract meaningful bounds on the EFT coefficients even when $t \sim \Lambda^2$.[11] Considering for instance $\Sigma_2(s,t)$, the idea of [118] is to subtract higher Σ_ns and their derivatives evaluated at $t = 0$ to obtain an integral, $\tilde{\Sigma}_2(s,t)$ whose low energy residue $\tilde{a}_2$ has only a finite number of terms of $O(t)$. Calling $a_n(t)$ the IR residue of $\Sigma_n(s,t)$, i.e. the $n-$th arc, one has:

$$\begin{aligned} \tilde{a}_2(t) &= a_2(t) - \sum_{n=3}^{\infty} \left(n t^{n-2} a_n(0) - t^{n-1} \partial_t a_n(0) \right) \\ &\equiv \frac{1}{-{M_{\rm Pl}}^2 t} + c_2 - c_3 t \, . \end{aligned} \tag{3.35}$$

Using each of the dispersion relations Σ_n it is possible to express the IR residues as a resummed UV integral, which has a certain allowed negativity. In this setup then, neglecting loop contributions, one can then derive sharp bounds on the EFT coefficients.

As exemplified by Eq. (3.33) however, GR loops lead to non-analytic contributions in the a_n at $t = 0$, e.g. in the form of $\log(t)/t$. Such loop contributions make actually impossible to define Eq. (3.35), since the infinite terms subtracted are divergent.

To circumvent this problem, as we present in [130], one can modify the approach used to define $\tilde{a}_2$ in such a way to never use dispersion relations evaluated at $t = 0$ and still retain only a polynomial of finite order in t in the arc. This modification makes much harder to find the right combination of higher a_n and derivatives to be subtracted, forcing one to use up to the nth derivative of each a_n, with coefficients to be determined algorithmically. Overcoming these technical complications, we find the well-defined improved arcs, $a_n^{imp}(s,t)$. For instance, for $n = 2$ we obtain:

$$\begin{aligned} a_2^{imp}(s,t) &= a_2(s,t) - \sum_{n=3}^{\infty} \sum_{m=0}^{n} c_{n,m} t^{n-2+m} \, \partial_t^m a_n(s,t) \\ &\equiv \frac{1}{-{M_{\rm Pl}}^2 t} + c_2 - c_3 t + logs \, , \end{aligned} \tag{3.36}$$

where *logs* indicate terms depending logarithmically on t and at least quadratically in the EFT coefficients. Quite non-trivially, it is possible to determine the generating function of the coefficients c_{nm}, i.e. a function $p(x,d)$ such that $p(x,d) = \sum_{n=3}^{\infty} \sum_{m=0}^{n} c_{n,m} d^m x^n$. Here we content ourselves in mentioning that this function, similarly to the generating functions for higher improved arcs, has a branch cut in x, which implies that the expression for the

11 To derive the actual constraints however this approach requires smearing the amplitude in impact parameter space, so that no constraint comes from a sharp value of t [118].

improved arc cannot be calculated straightforwardly if t is too close to Λ^2. This approach allows to compute the UV integrals through the dispersion relations Σ_n, finding new sharp and well defined bounds on the IR coefficients at finite t, having control on both loop contributions and $O(t)$ terms simultaneously. These bounds differ only parametrically from the bounds found in [118] using e.g. $\tilde{a}_2$ and neglecting light loops, as both of the approaches are parametrically close to non-improved bounds evaluated at small but finite t. Beyond the conceptual improvement, the difference between the two kind of bounds is relevant for characterizing the features of the boundary of the regions allowed by the constraints.

As can be understood from Eq. (3.33), these results do not change the fact that causality makes BGR loops subleading with respect to the GR ones, invalidating the loop positivity bounds.

3.2.3 *Power counting expectations*

The results of Sec. 3.2.1 can be understood in terms of NDA, when the power counting rules in Eq. (3.16) are extended to include the possibility that EFT operators can be generated at one-loop order,

$$\mathcal{L} = \tfrac{1}{2}\hat{M}_{Pl}^2 R + \frac{\Lambda^4}{g^2}\left[L^{(0)}\left(\frac{\nabla_\mu}{\Lambda}, \frac{R_{\mu\nu\rho\sigma}}{\Lambda^2}, \frac{g\phi}{\Lambda}\right) + \frac{g^2}{(4\pi)^2} L^{(1)}\left(\frac{\nabla_\mu}{\Lambda}, \frac{R_{\mu\nu\rho\sigma}}{\Lambda^2}, \frac{g\phi}{\Lambda}\right) + \cdots\right]. \tag{3.37}$$

This is because beyond-positivity contributions correspond to loop corrections within the EFT [116, 119, 188]. The one-loop NDA estimate for the $(\nabla\phi)^4$ operator is $(c_2)^{(1)} \sim \frac{1}{16\pi^2} g^4 \Lambda^{-4}$, which for $g \sim \Lambda/M_{\rm Pl}$ matches the maximal value of the r.h.s. of Eq. (3.28), i.e. for $|\hat{\alpha}| \sim 1/\Lambda^2$. Likewise, we can estimate the beyond-positivity contributions to quartic curvatures operators from $L^{(1)}$ in Eq. (3.37), $(\alpha_4)^{(1)}/{M_{\rm Pl}}^2 \sim \frac{1}{16\pi^2}(\Lambda M_{\rm Pl})^{-4}$, which coincides with the r.h.s. of Eq. (3.33) for the maximum value of the GB/CS coupling.

This discussion brings us to the important realization that, from the EFT standpoint, for UV completions where both the scalar-GB/CS term in Eq. (3.1) and the operators in Eq. (3.23) are generated at the same (tree-level) order, one should expect much larger coefficients for the latter than what discussed above. To see this, let us simply fix the power counting from the GB/CS term, assuming the maximal regime of validity of the EFT, $|\hat{\alpha}| \sim 1/\Lambda^2$ (a requirement, rather than a choice, if one interested in phenomenological applications, see Sec. 3.3). From $L^{(0)}$ in Eq. (3.37), this sets $g \sim 1/|\hat{\alpha}|\Lambda M_{\rm Pl} \sim \Lambda/M_{\rm Pl}$, corresponding to a bona-fide UV completion of gravity, as discussed in Sec. 3.1.3. Then, generic EFTs will feature

$$\alpha(\tilde{\alpha}) \sim \frac{1}{\Lambda^2}\,,\ \alpha_3 \sim \frac{1}{\Lambda^4}\,,\ \alpha_4, \alpha_4' \sim \frac{1}{\Lambda^6}\,,\ c_2 \sim \frac{1}{\Lambda^2 {M_{\rm Pl}}^2}\,,\ d_1 \sim \frac{1}{\Lambda^4}\,, \tag{3.38}$$

for the coefficients of the operators in Eq. (3.23).

In the next section we investigate the phenomenological consequences of these estimates for the physics of black holes with scalar hair.

3.3 PHENOMENOLOGICAL IMPLICATIONS

In this section we discuss the main implications on the phenomenology of black holes of the upper bound on the GB/CS coupling $|\hat{\alpha}| \lesssim 1/\Lambda^2$,

along with the implications associated with the additional EFT corrections that are expected from saturating such a bound, Eq. (3.38). Our focus is on astrophysical black holes, in particular those detectable by LIGO-Virgo, which have sizes of a few solar masses, corresponding to Schwarzschild radii $r_s \gtrsim 10\,\mathrm{km}$.

We will discuss separately the scalar-GB and dynamical-CS gravity theories, reviewing in each case their imprints on the physics of black holes as well as the current experimental bounds on the couplings α and $\tilde{\alpha}$, respectively.

3.3.1 *Black holes in scalar-GB gravity*

From a perturbative point of view, one can argue that the metric of a hairy black hole still displays a horizon, characterized by a linear zero of the metric, much like in the Schwarzschild and Kerr cases. The effects due to the BGR dynamics rapidly vanish far away from the horizon ($r = r_s$), following the fall-off of the GB invariant, which sources both the scalar hair and the deviations from GR in the metric.

More in detail, in the static and spherically symmetric case, we have the metric

$$ds^2 = -h(r)dt^2 + f(r)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)\,, \tag{3.39}$$

with

$$h(r), f(r) \sim \left(1 - \frac{r_s}{r}\right) \quad \text{for} \quad r \sim r_s\,. \tag{3.40}$$

At leading order in the dimensionless expansion parameter α/r^2, the GB invariant is the one of the Schwarzschild solution,

$$\mathcal{R}^2_{\rm GB} \equiv R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma} - 4R^{\mu\nu}R_{\mu\nu} + R^2 \sim \frac{r_s^2}{r^6}\,. \tag{3.41}$$

The scalar equation of motion reads

$$\Box\phi = M_{\rm Pl}\alpha\mathcal{R}^2_{\rm GB} \sim \frac{{M_{\rm Pl}}^2}{\Lambda_\alpha^3}\frac{r_s^2}{r^6}\,, \tag{3.42}$$

where in the last step we have traded the GB coupling α for the strong coupling scale Λ_α, given in Eq. (3.15) ($\tilde{\alpha} = 0$). The scalar field profile is then completely determined by requiring that invariant quantities built out of it do not diverge for $r \geq r_s$ [189]. At asymptotically large distances, $r \to \infty$, the solution behaves as $\phi(r) \sim 1/r$. In addition, let us note that the largest value of the scalar radial derivative is estimated as,

$$\phi' \lesssim \frac{{M_{\rm Pl}}^2}{(\Lambda_\alpha r_s)^3}\,. \tag{3.43}$$

A first bound on the GB coupling comes from the requirement of the existence of real solutions for the scalar profile. In the simple EFT Eq. (3.1), this condition requires that $\alpha^2 < r_s^4/192$ [92]. The constraint is however dependent on additional EFT corrections from Eq. (3.23) [85].

In order to estimate the impact of the scalar-GB operator on the background geometry, we can compute its ratio to GR, that is to ${M_{\rm Pl}}^2\mathcal{R}$ where $\mathcal{R} = \sqrt{R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}}$ is the typical curvature. Evaluating both terms on the background given by the Schwarzschild metric, $\mathcal{R} \sim r_s/r^3$, and the scalar solution from Eq. (3.42), one finds [85]

$$\varepsilon_0(r) = \frac{M_{\rm Pl}\alpha\phi\mathcal{R}^2_{\rm GB}}{{M_{\rm Pl}}^2\mathcal{R}} \sim \left(\frac{\alpha}{r^2}\right)^2\,. \tag{3.44}$$

Turning to perturbations, let us start by noting that the theory Eq. (3.1) has no scalar self-interactions. Therefore, there is simply no possible screening effect associated to classical non-linearities. On the other hand, there is no direct coupling between the scalar field and matter, therefore no screening mechanism is required to have agreement with fifth-force constraints (if the theory is valid at the scales of those experiments).

Instead, the scalar-GB term gives rise to a kinetic mixing between scalar and graviton (the same leading to the causality bound of Sec. 3.1.1), schematically of the form

$$M_{\rm Pl}\alpha\phi\mathcal{R}^2_{\rm GB} \supset \varepsilon_{mix}(r)\,\partial\phi\partial h\,, \tag{3.45}$$

where we are taking the fluctuations to be canonically normalized. This effect will be important when the mixing

$$\varepsilon_{mix}(r) \sim \frac{\alpha r_s}{r^3}\,, \tag{3.46}$$

becomes of order one. Note that the two estimators of the BGR effects are related to each other, namely

$$\varepsilon_0 \sim \left(\frac{r}{r_s}\varepsilon_{mix}\right)^2\,. \tag{3.47}$$

In this scenario, a sizable deviation of the quasi-normal mode (QNM) spectrum from the GR prediction is expected, strongly affecting the waveform during the ringdown phase of a merger.

Finally, let us turn to the phenomenology of a binary system of hairy black holes, each sourcing its own scalar profile as discussed before. The dynamical nature of the system implies that, just as it happens with gravitational waves, there will also be scalar wave emission. However, the latter is now dipolar instead of quadrupolar, therefore being much less suppressed than the former. This opens a new channel of power loss during the merger, which accelerates the rate of change in the orbital period. The effect accumulates during the inspiral phase, potentially producing an observable dephasing between the measured waveform and the one predicted by GR. To date, the absence of any observed effect of this type constitutes the most stringent experimental bound on the size of the scalar-GB coupling, $\alpha \lesssim (1.2\,\text{km})^2$ [98]. In terms of the strong coupling scale, this translates into $\Lambda_\alpha \gtrsim 10^{12}\,\text{km}^{-1}$. Let us add that such a bound strictly applies only to scalar-GB gravity with all other EFT corrections, in particular those in Eq. (3.23), neglected or irrelevantly small.

Considering a typical LIGO/Virgo black hole, with $r_s \sim 10\,\text{km}$, the above bound implies that the kinetic mixing is constrained to be $\varepsilon_{mix} \lesssim 10^{-2}$. Furthermore, according to Eq. (3.47) the effect on the background geometry is significantly suppressed, $\varepsilon_0 \lesssim 10^{-4}$. This conclusion justifies neglecting deviations from Schwarzschild background as we assumed initially.

3.3.2 *EFT implications on scalar-GB black holes*

The fact that the BGR effects on hairy black holes are relatively small, below the 10 % level, could have been anticipated from the requirement that the scalar-GB theory should be able to properly describe the black holes of interest. Indeed, as derived in Sec. 3.1.1, causality sets an upper bound on the GB coupling and therefore on the size of the observable corrections relative

to GR, which we denote generically with $\epsilon(r)$ – one instance being ε_{mix} in Eq. (3.46). The BGR effects are largest near the horizon,

$$\epsilon(r_s) \sim \frac{\alpha}{r_s^2} \lesssim (\Lambda r_s)^{-2}\,, \tag{3.48}$$

where the inequality follows from causality, Eq. (3.13) (with $\log(b_0\Lambda) \sim 1$ and $\tilde{\alpha} = 0$). Usefulness of the EFT requires a hierarchy between the cutoff and the relevant scales of the system. A sensible demand on the EFT is therefore that the black hole falls within the EFT regime of validity at least down to its Schwarzschild radius.[12] Therefore, BGR corrections can never become large, i.e. $\epsilon(r_s) \ll 1$. Furthermore, taking the current experimental upper bound on α as benchmark, the causality bound implies a very low maximal cutoff,

$$\Lambda \lesssim (1\,\mathrm{km})^{-1}\,, \tag{3.49}$$

certainly much smaller than the strong coupling scale $\Lambda_\alpha \gtrsim (10^{-12}\,\mathrm{km})^{-1}$.

Let us discuss now the additional BGR effects that could be expected from the UV completion in the form of higher-dimensional operators with coefficients fixed to Eq. (3.38), where let us recall that such NDA estimates correspond to the the maximal cutoff $\Lambda \sim \alpha^{-1/2} \sim (1\,\mathrm{km})^{-1}$.

The operators in Eq. (3.23) give rise to modifications of the geometry. We can estimate such modifications as in Eq. (3.44) for the scalar-GB term, which we recall scales as $\varepsilon_0(r) \sim (\alpha/r^2)^2 \sim (\Lambda r)^{-4}$. Similarly, we find

$$\begin{aligned} \frac{\alpha_3\mathcal{I}}{\mathcal{R}} &\sim \frac{r_s^2}{r^2}(\Lambda r)^{-4}\,, \quad \frac{\alpha_4\mathcal{C}^2}{\mathcal{R}} \sim \frac{r_s^3}{r^3}(\Lambda r)^{-6}\,, \\ \frac{c_2(\nabla\phi)^4}{{M_{\rm Pl}}^2\mathcal{R}} &\sim \frac{r^5}{r_s^5}(\Lambda r)^{-10}\,, \quad \frac{d_1\mathcal{C}(\nabla\phi)^2}{{M_{\rm Pl}}^2\mathcal{R}} \sim \frac{r}{r_s}(\Lambda r)^{-8}\,. \end{aligned} \tag{3.50}$$

While deviations introduced by the scalar-GB term are the largest, operators cubic in the Riemann tensor can become as important near the horizon. The deviations introduced by the rest of operators are subleading, being higher order in $(\Lambda r_s)^{-1}$, as expected from the derivative expansion and Eq. (3.43). In addition, the operators built out of the scalar give rise to modifications of the scalar field profile, which we can estimate as

$$\frac{c_2(\nabla\phi)^4}{(\nabla\phi)^2} \sim \frac{r^2}{r_s^2}(\Lambda r)^{-6}\,, \quad \frac{d_1\mathcal{C}(\nabla\phi)^2}{(\nabla\phi)^2} \sim \frac{r_s^2}{r^2}(\Lambda r)^{-4}\,. \tag{3.51}$$

Given the upper bound Eq. (3.49), we conclude that the operators in Eq. (3.23) should induce corrections on the metric and scalar of up to 0.01 % near the horizon of black holes with $r_s \sim 10\,\mathrm{km}$.

There are many other interesting signatures associated with the operators that we expect to be present in the scalar-GB EFT. The phenomenology of cubic and quartic curvature operators have been discussed in [190–194] and [177–179, 194], respectively. Besides modifications of the Schwarzschild (and Kerr) geometries, these include deviations from GR at the leading order in QNMs, quadrupole moments, non-vanishing Love numbers, and corrections to the gravitational-wave signals at relatively high post-Newtonian (PN) order. Since such corrections start at order $(\Lambda r)^{-4} \lesssim 10^{-4}$, they will not be easy to probe with the sensitivity of current experiments.

12 As a matter of fact, one would like a gravitational EFT to be valid at least up to the scales where gravity has been experimentally tested, that is a cutoff $\Lambda \gtrsim \mu\mathrm{m}^{-1}$ [147]. One is forced to give up on such a requirement if the scalar-GB theory is to be phenomenologically interesting for astrophysical black holes (see however the discussion in Sec. 3.5).

Operators involving the scalar field have received less attention in the literature. The impact of the cubic Galileon operator, which we have rewritten in Eq. (3.23) as $d_1 R^2_{\mu\nu\rho\sigma}(\nabla_\eta\phi)^2$, has been discussed in [85]. Along with the operator $c_2(\nabla_\mu\phi)^4$, these EFT terms introduce modifications e.g. in the scalar and gravitational wave spectrum, as well as in the QNMs, predicted by the pure scalar-GB theory Eq. (3.1), although a priori subleading due to the suppression by higher powers of $(\Lambda r)^{-2} \ll 1$. In particular, note that potential screening effects are not likely to be significant.

Let us recall once again that these conclusions appear to be a robust consequence of causality. Nevertheless, quantitative predictions for the gravitational observables, which typically require performing costly numerical simulations, are still interesting, if only to experimentally test the fundamental principles behind these expectations.

3.3.3 *Black holes in dynamical-CS gravity*

The dynamical-CS term in Eq. (3.1) (with $\alpha = 0$) it is usually studied in the context of a pseudoscalar field coupled to gravity, leading to a phenomenology of black holes similar to that discussed in the previous section, although with a few important differences. First, the Pontryagin invariant, $R_{\mu\nu\rho\sigma}\tilde{R}^{\mu\nu\rho\sigma}$, vanishes in the Schwarzschild geometry, while it is non-zero for the Kerr geometry. Therefore, one needs to consider rotating black holes in order to have a non vanishing scalar hair. To simplify the analysis, we follow [146] and treat the spin parameter of the black hole, a/r_s, perturbatively. We also work in an expansion in $\tilde{\alpha}/r_s^2$ since, similar to discussion for scalar-GB black holes in Sec. 3.3.2, this is a consequence of causality, $\tilde{\alpha} \lesssim 1/\Lambda^2$, along with the requirement that the black holes of interest fall within the regime of validity of the EFT, i.e. $(\Lambda r_s)^{-2} \ll 1$.

At leading order, the equation of motion for the pseudo-scalar ϕ is given by

$$\Box\phi = M_{\rm Pl}\tilde{\alpha} R_{\mu\nu\rho\sigma}\tilde{R}^{\mu\nu\rho\sigma} \sim \frac{{M_{\rm Pl}}^2}{\Lambda_{\tilde{\alpha}}^3}\frac{r_s^2}{r^6}\frac{a}{r}\cos\theta\,, \tag{3.52}$$

where in the last step we have traded the CS coupling $\tilde{\alpha}$ for the strong coupling scale $\Lambda_{\tilde{\alpha}}$, given in Eq. (3.15) ($\alpha = 0$ and changed subscript to avoid confusion). Subleading terms in the spin parameter scale as $(a/r_s)^3$. From the solution to this equation (see e.g. [146]), we can compute the scalar radial derivative, which roughly satisfies

$$\phi' \lesssim \frac{{M_{\rm Pl}}^2}{(\Lambda_{\tilde{\alpha}} r_s)^3}\frac{a}{r_s}\,. \tag{3.53}$$

Similarly to scalar-GB case, there is an approximate relation between corrections to the Kerr geometry, ε_0, and the kinetic mixing between pseudo-scalar and graviton, ε_{mix}, given by

$$\varepsilon_0(r) \sim \left(\frac{a}{r_s}\varepsilon_{mix}(r)\right)^2\,. \tag{3.54}$$

A second difference w.r.t. the scalar-GB case is the nature of the most stringent experimental bounds on the CS coupling. Since the scalar background sourced around an isolated (spinning) black hole is dipolar, the emission of scalar waves from a black hole binary system starts from a quadrupole moment. Therefore, energy loss via scalar emission is further suppressed in

the PN expansion compared to the scalar-GB case, such that no constraint can be derived on $\tilde{\alpha}$ given current sensitivities [195].

The strongest bound to date on the pseudo-scalar-CS coupling comes from independent measurements of the tidal deformability and of the moment of inertia in neutron stars [196]. The comparison between these measurements and the values predicted in dynamical-CS gravity yields the bound $\tilde{\alpha} \lesssim (8\,\mathrm{km})^2$. Note this is weaker than the most stringent bound on the scalar-GB coupling, $\alpha \lesssim (1.2\,\mathrm{km})^2$ by one order of magnitude. Nevertheless, if the pseudo-scalar-CS EFT is to be able to describe black holes with sizes down to $r_s \sim 10\,\mathrm{km}$ (recall that the smallest black holes display the largest BGR effects, which in any case cannot become $O(1)$) with at least 10 % accuracy, it seems wise to consider a different benchmark for the coupling $\tilde{\alpha}$. Maximal testability compatible with causality then suggests to take $\tilde{\alpha} \sim 1/\Lambda^2 \sim (3\,\mathrm{km})^2$.

In this case, the kinetic mixing between pseudo-scalar and graviton can induce stronger deviations in the QNM spectrum w.r.t. to GR [197–199], of order $\varepsilon_{mix} \lesssim (\tilde{\alpha}/r_s^2) \sim 10^{-1}$ for black holes with $r_s \sim 10\,\mathrm{km}$.

The discussion of the implications of the additional operators in Eq. (3.23), present in generic UV completions of pseudo-scalar-CS theory, largely parallels that of Sec. 3.3.2 and we do not repeat it here. Nevertheless, we wish to point out that to date much less work has been devoted to the study of these EFT effects for rotating black holes, see e.g. [192, 193].

Before closing the section, let us point out one last difference between the scalar-GB and dynamical-CS theories. While the scalar-GB operator leads to equations of motion of at most second order in (time) derivatives, the dynamical-CS operator gives rise to higher-derivative terms. These in principle could spoil the quantum stability of the theory. Considering perturbations of rotating black holes with pseudo-scalar-CS hair, for instance during the inspiral phase of a merger,[13] higher derivatives will become important at a mass scale $M_g^{-1} \sim \tilde{\alpha}\dot{\phi}_0/M_{\rm Pl}$, being $\dot{\phi}_0 \sim \omega\phi_0 \lesssim \phi_0/r_s$ the time derivative of the axionic field evaluated on the background. We can estimate this scale, at least in some appropriate regime, using the solution of Eq. (3.52), finding

$$M_g \sim \frac{1}{a}\left(\frac{r_s^2}{\tilde{\alpha}}\right)^2 . \tag{3.55}$$

Since causality requires $\Lambda \lesssim \tilde{\alpha}^{-1/2}$, we find that the ghost's mass is above the EFT cutoff.

3.4 GAUSS-BONNET SCALARIZATION

For scalar-tensor theories with no scalar shift-symmetry, one can consider non-linear couplings between ϕ and the GB invariant. Let us focus on the leading Z_2-symmetric ($\phi \to -\phi$) term in a field and derivative expansion,

$$S_{\rm GBization} = \int d^4x\sqrt{-g}\left(\frac{M_{\rm Pl}{}^2}{2}R - \frac{1}{2}(\nabla_\mu\phi)^2 + \lambda\phi^2\mathcal{R}^2_{\rm GB}\right) , \tag{3.56}$$

where λ has dimensions of a length square. The sign of λ determines whether $\phi = 0$ is a stable solution or not around a gravitational source like a (static or spinning) black hole [103]. Spontaneous scalarization, i.e. a

13 This system allows us to consider non-vanishing time derivatives of the scalar background, which can lead to ghosts instabilities.

non-trivial scalar profile, generically develops for $|\lambda|/r_s^2 \gtrsim 1$, with $\lambda > 0$ provided the GB invariant is positive.[14]

From simple little group (helicity) selection rules, one can derive the 4-point interaction of two scalars and two gravitons associated with the non-minimal coupling to GB,

$$\mathcal{M}^{\text{GBization}}_{1_\phi 2_\phi 3_{h^{++}} 4_{h^{++}}} = \frac{4\lambda}{{M_{\rm Pl}}^2}[34]^4\,, \quad \mathcal{M}^{\text{GBization}}_{1_\phi 2_\phi 3_{h^{--}} 4_{h^{--}}} = \frac{4\lambda}{{M_{\rm Pl}}^2}\langle 34\rangle^4\,. \tag{3.57}$$

Note that this is associated with an inelastic $\phi h^{++} \to \phi h^{--}$ scattering amplitude, that vanishes in the forward limit $t \to 0$,

$$\mathcal{M}^{\text{GBization}}_{1_\phi 2_{h^{++}} \to 3_\phi 4_{h^{--}}} = \frac{4\lambda}{{M_{\rm Pl}}^2} t^2 e^{4i\theta}\,, \tag{3.58}$$

where θ is just a phase (for physical momenta $[ij]^* = \langle ij\rangle = \sqrt{s_{ij}}e^{i\theta}$, with $s_{13} = t$).

While a dispersion relation for λ from the 2-scalar-2-graviton amplitude can be derived along the lines of Sec. 3.2.1, the inelasticity of the amplitude preclude the derivation of a positivity bound $\lambda > 0$. One can actually come up with a simple (yet partial) tree-level UV completion that shows that the sign of λ is not fixed. This involves an additional massive scalar field Φ,

$$\begin{aligned} S^{\rm UV}_{\rm GBization} = \int d^4x\sqrt{-g}\Big(&\tfrac{1}{2}{M_{\rm Pl}}^2 R - \tfrac{1}{2}(\nabla_\mu\phi)^2 - \tfrac{1}{2}(\nabla_\mu\Phi)^2 - \tfrac{1}{2}m_\Phi^2\Phi^2 \\ &+ M_{\rm Pl}\alpha_\Phi\Phi\mathcal{R}^2_{\rm GB} + g_\Phi\Phi\phi^2\Big)\,. \end{aligned} \tag{3.59}$$

One can see that upon integrating out the massive scalar, one obtains

$$\lambda = \frac{M_{\rm Pl}\alpha_\Phi g_\Phi}{m_\Phi^2}\,, \tag{3.60}$$

with no definite sign, since e.g. α_Φ can consistently be positive or negative. In addition, from this example one can infer that the size of λ is likely to be theoretically bounded. Indeed, since Eq. (3.59) is itself an effective action, from the generic EFT perspective presented in Sec. 3.1.3, the trilinear coupling is of order $g_\Phi \sim g\Lambda$ while $\alpha_\phi \sim 1/(gM_{\rm Pl}\Lambda)$, where Λ is the cutoff and g a coupling. The mass of the heavy scalar can be at most of the order of the cutoff, i.e. $m_\Phi \sim \Lambda$, and should itself be identified with the cutoff of Eq. (3.56). We are then led to the conjecture that the quadratic scalar-GB coupling should not be much larger than $\lambda \sim 1/\Lambda^2$ given the causality bound on α_Φ. Following similar arguments, one can start with the effective interaction $\lambda\phi^2\mathcal{R}^2_{\rm GB}$ and give the scalar a vacuum expectation value, which within the EFT can be at most $\langle\phi\rangle \sim \Lambda/g \lesssim M_{\rm Pl}$. This gives rise to the scalar-GB term in Eq. (3.1) with $M_{\rm Pl}\alpha \sim \lambda\langle\phi\rangle$, which for $\lambda \lesssim 1/\Lambda^2$ is consistent with the causality bound we derived in Sec. 3.1.1. Note that, at the end of the day, these arguments are just refined versions of the statement that, from the gravitational power counting discussed in Sec. 3.1.3, we expect $\lambda \sim \zeta^2/\Lambda^2 \lesssim 1/\Lambda^2$ for $\zeta \lesssim 1$. However, in this case one cannot reinterpret such expectation as a consequence of the requirement that the BGR amplitude, Eq. (3.57), should not become larger than GR's within the EFT, since the latter vanishes at tree level (and at one loop) [188].

14 This is the case e.g. on a static black hole background or away from the horizon on a rotating black hole background.

These expectations were recently confirmed in the context of dispersion relations valid at tree level [128], using a similar approach to that of [127]. Including loops, one can construct a dispersion relation for the 4-scalar amplitude along the lines of Sec. 3.2.1. In this case, the beyond positivity bound corresponding to Eq. (3.56) is associated with a cross section for $\phi\phi \to h^{\pm\pm}h^{\pm\pm}$ which, similar to Eq. (3.28), leads to $c_2 \gtrsim \frac{1}{16\pi^2}\lambda^2(\Lambda/M_{\rm Pl})^4$. If indeed $\lambda \lesssim 1/\Lambda^2$ regardless of the UV completion as long as this is unitary and causal, then the lower bound on c_2 is in fact inapplicable due to the graviton pole.

Finally, from the phenomenological point of view, scalarization of black holes turns out to be a dubious phenomenon, given that $\lambda/r_s^2 \lesssim (\Lambda r_s)^{-2} \lesssim 1$ if the EFT is to describe the black holes down to their horizon. However, let us recall that, differently from the cases discussed in the main text, for the theory of GB-scalarization we have not found solid evidence that causality forces the BGR effects to be subleading.

3.5 SUMMARY

In this chapter, we used causality arguments to derive phenomenologically interesting bounds on theories that, at low energies, comprise the graviton and a shift-symmetric scalar field. The presence of a coupling between the scalar and the Gauss-Bonnet or the Chern-Simons operators leads to black hole hair. If the couplings α and $\tilde{\alpha}$, in the notation of Eq. (3.1), are large enough, $|\hat{\alpha}| \sim r_s^2$, hair can be measured in astrophysical black holes. The first consequence that we derived imposing the absence of classical superluminality is that the cutoff of the EFT cannot be parametrically larger than $|\hat{\alpha}|^{-1/2}$, i.e. $1/$ km for solar mass BHs. This has implications for the structure of all the higher-dimensional operators in the theory. One could attempt to draw robust lower bounds on their coefficients using positivity constraints obtained via dispersion relations. However, the weakness of non-minimal gravitational interactions compared to GR, enforced by causality, implies that lower bounds from one-loop dispersion relations are phenomenologically irrelevant. In regard to this conclusion, we outlined how dispersion relations can be improved to accommodate gravitational effects both at tree level and at loop level.

On the other hand, for such a low cutoff $\Lambda \sim 1/$ km, we used general power counting arguments to show that if both the scalar-GB/CS term and other operators are generated at the same (tree-level) order, the latter will also give sizable contributions in black hole dynamics.

The result that the UV cutoff of a phenomenologically interesting and causal EFT describing gravity must be lower than km^{-1} should not be considered only as a source of potentially large corrections to the effective description of astrophysical black holes. Instead, we should demand that this result is reconciled with our experimental knowledge of gravity. As a matter of fact, to date gravity has been probed in table-top experiments down to the scales of tens of microns in length, showing good agreement with Newtonian theory [147]. In light of this, one should at least prefer, if not demand, that GR is extended using a theory valid down to the μm, in such a way to describe the same observations as GR does.

In the scenarios studied above, being the cutoff at a much lower scale than $\mu\mathrm{m}^{-1}$, one needs to trust that the UV completion that gives rise either to the scalar-GB or to the dynamical-CS EFTs, does indeed reproduce the Newtonian potential at microscopic lengths. Similarly to what was argued

in [177, 179] for quartic curvature terms, this might be the case of a "soft" UV completion that resolves the irrelevant operators in the EFT in such a way that interactions stop growing with the energy.

In addition to this requirement, we need the UV completion to contribute to the time delay in such a way that causality is preserved, in particular given the negative contributions (i.e. the time advance) from the low-energy operators.

This requirement is a substantial obstacle for both the scalar-GB and dynamical-CS EFTs.[15] Indeed, since the sizes of both the scalar-GB and of the axion-CS operators are chosen in such a way to follow the tree-level NDA, it appears difficult for loop-level effects in the UV physics to restore causality, unless the number of species scales as $(4\pi/g)^2 \sim (4\pi M_{\rm Pl}/\Lambda)^2$ [164, 165]. On the other hand, as it was argued in [99, 163], in order for causality to be restored by tree-level exchanges in the UV, one must introduce an infinite tower of higher-spin particles having a mass of order $M \sim \text{km}^{-1} \sim 10^{-10}\,\text{eV}$. In both these cases, the description of gravity below the km would be very different from what we know.

In light of this, from a conservative point of view it seems that both scalar-GB and pseudo-scalar-CS interactions cannot lead to testable modifications of GR. A less conservative attitude might be trying to understand whether a UV completion of these models exists that restores causality without clashing with our knowledge of gravity at small distances. In any case, the detection of black hole hair would be revolutionary, telling us there is something fundamentally unexpected and so far unknown about gravitational dynamics.

15 As a matter of fact, consistency with unitarity and causality is an obstacle as well for any BGR deformation of phenomenological relevance above the μm, in particular if it involves higher-order terms in the curvature [127].

Part II

BINARIES PERTURBED BY A DISTANT THIRD BODY

ENVIRONMENTAL EFFECTS: BINARIES PERTURBED BY A DISTANT THIRD BODY

As discussed in the Introduction, the study of binary mergers makes both desirable and necessary to parametrize efficiently the various possible environmental and finite size effects that can considerably change the binary inspiral over the course of many orbital periods.

Even in the ideal case of an isolated black hole binary, the study of the inspiral dynamics has greatly benefited from various ideas developed in the context of particle physics. The multitude of different methods that have been proposed in the last decades can be organized based on which parameters are treated perturbatively.
One possibility is to treat the gravitational interaction perturbatively in the so-called Post-Minkowskian expansion, in which $\frac{Gm}{r} \sim \frac{r_s}{r} \ll 1$ (r being the separation between the bodies). In this limit it is possible to employ various techniques of scattering amplitudes and quantum field theory to extract fully relativistic, classical information about unbounded binary systems. Notable examples are the binding potential [200–204], the eikonal phase shift [205–207], effects on the trajectory of the bodies [208] as well as the expectation value of various observables [209]. Once a quantity has been determined for an unbound system, one can try to obtain its bound system counterpart by analytic continuation [210, 211]. In the same PM expansion, the Bethe-Salpeter equation can be used to study bound systems directly [212].
A different possibility is to restrict to bound systems from the get-go, and use knowledge of the virial theorem in conjunction with the non-relativistic approximation to treat perturbatively the typical squared velocity of the bodies $v^2 \sim \frac{Gm}{r} \ll 1$, in what is called the Post-Newtonian expansion. This approximation is employed in the Non-Relativistic General Relativity (NRGR) approach [70, 213].
A third possible expansion parameter is given by the ratio of the masses of the two bodies. This is called the self-force (SF) expansion [214–216], and allows to obtain results that are non perturbative in the Newton's constant G. These different strategies to gain analytical control over the dynamics of the inspiral can be combined in different ways [217–220].

These analytical approaches have been used to study effects due to gravitational wave emission, i.e. the non conservative contribution to the binary dynamics, see e.g. [221], as well as effects due to the spin and finite size of the bodies, see e.g. [222–225]. The understanding of these effects in binary dynamics remains currently a very active area of research, see e.g. [226] for a summary of recent progresses.

As the precision of gravitational wave detection as well as our analytical understanding of binaries progress, there is both interest and opportunity to describe environmental effects by adapting existing analytical techniques. In particular, one can exploit a distinctive feature of bound binary dynamics, which is the natural separation of scales between high frequency modes on time-scales of the orbital period and the nearly adiabatic evolution of

the orbit due to departures from the Newtonian dynamics. This separation of scales is already present for an isolated binary in GR, where relativistic corrections both at the conservative and non-conservative level have a sizable effects, namely the periastron precession and the shrinking of the orbit, over time-scales much longer than the orbital period. The same holds for effects due to the finite size and spinning of the bodies, see e.g. [213]. This feature of binary dynamics makes particularly convenient to employ EFT techniques to describe a perturbed binary, as these make possible to obtain a simplified description of the system over long timescales. In particular, describing relativistic effects through the PN expansion allows to analyze the interplay and relative importance of the various relativistic effects and possible perturbations over certain time-scales.

Motivated by these opportunities and by the perspectives on future observations, here we take on the task analyzing the case of a specific kind of perturbation: the presence of a distant third body that orbits the binary.
Owing to the hierarchy in the typical distances between the bodies, these systems are referred to as hierarchical triples. Besides being a convenient choice of perturbation that can serve as a proxy for more generic scenarios, effects related to hierarchical triples are very common in nature [227]. Examples include satellites and asteroids in the Solar system [228, 229], triple stars [230–233], exoplanets [234–237] as well as triple systems including black holes and neutron stars [238–243]. In addition to these examples, future detections of relativistic triples are expected with LISA [244, 245].
As we expect, the large distance between the third body and the other two results in a nearly Keplerian dynamics with two independent orbits that evolve adiabatically. One representing the perturbed binary and the other representing e.g. the orbit of the third body. We will dub these the inner and outer orbit. This setup makes possible to describe the effects due to the third body in a multipole expansion controlled by the ratio $\varepsilon = \frac{a}{a_3} \ll 1$ of the semi-major axes of the inner and outer orbit[16]. With this perturbative approach, three body effects have been studied to understand e.g. the merger rate of compact objects [231, 240, 246–250] , the transits of exoplanets [251–253] or the evolution of triple star systems [254–257]. The perturbative description, valid as long as the third body remains far from the other two, makes possible to outline interesting mechanisms that occur on timescales much longer than the periods of the two orbits.
One example is given at the level of Newtonian interaction by the so-called Kozai-Lidov (KL) mechanism [258], consisting in large oscillations of the eccentricity and inclination of the inner orbit on timescales

$$T_{\rm KL} \sim \frac{T_3}{T} T_3 \sim \frac{1}{\varepsilon^{3/2}} T_3 \,, \tag{3.61}$$

T (T_3) being the period of the inner (outer) orbit. Due to this effect, for instance, hierarchical systems of black holes can feature dramatically reduced merger timescales [248]. Since these systems are quite common in dense stellar environments [259, 260], the Kozai-Lidov mechanism is especially relevant to gravitational wave astronomy. Over longer time-scales, Newtonian three-body interactions can induce other effects, such as orbital flips [261–263].

Moving past Newtonian interactions, the relation between two-body GR effects and KL oscillations has been studied by numerous authors,

16 As a convention, we will indicate with the subscript $_3$ quantities that describe the outer orbit, while we will use no subscripts for quantities describing the inner orbit.

e.g. [246, 264–266]. Generically, the inner binary precession tends to suppress the eccentricity oscillations if the PN precession time-scale (T/v^2) is much shorter than the KL one [263]. In other parts of the phase space, however, PN corrections combined with the three-body effects can result in exciting the inner orbit eccentricity [249]. Beyond these results much less is understood about the corrections brought by the genuine three-body relativistic effects. These terms are essential to understand the interplay between relativistic and three-body effects and obtain the correct time-evolution of the system in some parts of the parameter space [267]. Due to the fact that such effects appear suppressed by both ε and powers of v, they are usually called cross terms [268]. In the following we will address the question of describing relativistic three-body effects from the point of view of EFT. Leveraging on the separation of scales displayed by hierarchical triples, we will use EFT and NRGR to derive a simplified, effective description for the long time-scale dynamics including the first relativistic-quadrupole corrections. Our approach will make a clear connection between the multipole expansion of the three body effects and the effective long-timescale description of both the inner and outer orbits as composite particles. Importantly, this connection makes possible to use symmetry arguments to find the most convenient variables in which to encode the dynamics. In practice, this description is derived by computing averages of the multipole moments of the orbits over their periods, capturing backreaction as well as deviations from perfect adiabaticity of the orbital evolution. In Chapter 4 we will review the approach of NRGR and carry out in the simplest terms our construction, deriving an effective Lagrangian that describes the system over long timescales including $O(v^2\varepsilon^{3/2})$ contributions. In Chapter 5, we will refine our construction to derive relativistic quadrupole effects and obtain an effective Lagrangian including $O(v^2\epsilon^{5/2})$ contributions. Many of the technical details of this construction will be presented in the appendices B.

4

AN EFT DESCRIPTION OF HIERARCHICAL THREE BODY SYSTEMS

The hierarchical three body problem has been for centuries a central topic in celestial mechanics, see e.g. [269] for a comprehensive discussion of the Newtonian case. A conventional treatment for tackling it proceeds by expanding the Hamiltonian in the ratio of semimajor axes, ε, to then average it on both orbital timescales to obtain a set of long-timescale evolution equations of orbital quantities, a procedure known as double averaging. The set of equations thus obtained is commonly referred to as the Lagrange planetary equations. For instance, the quadrupole term of this averaged expansion gives rise to the KL oscillations discussed above.

Another valid approach is to further assume a hierarchy of masses $m_3 \gg m_1, m_2$ so that the system can be studied with black hole perturbation theory [270–278].

Most previous studies on the fully relativistic hierarchical three-body problem use a combination of the post-Newtonian formalism [15] and the quadrupole expansion at the level of either the equations of motion [267, 279–283] or the Hamiltonian [249, 284]. However, their studies do not take into account the so-called cross terms computed in [281–283]. The computations needed to identify these contributions can be quite cumbersome, making difficult to develop a physical intuition of the role of the various terms.

To tackle this issue and develop a satisfactory method to understand cross-terms, instead, we begin exploring three-body systems following a new approach, based on a number of powerful EFT techniques that have been developed for the relativistic two-body problem in recent years.

In particular, we employ the techniques of NRGR [70, 213] building on the strategy proposed to describe the spin of the constituents of a binary [71, 223, 285], and on the resulting spin-induced PN contributions that have been computed [71, 222, 286–289]. In the EFT language, the gravitational multipole expansion is implemented at the level of the action [290], using symmetries to restrict the form of the allowed terms [291]. The multipole expansion derived in this way can then be employed to compute the gravitational dissipative dynamics in the GR two-body problem [221].

In the following, we will apply similar ideas to the hierarchical three-body problem. As mentioned, the dynamics of (mildly) relativistic hierarchical triples is characterized by two small dimensionless ratios of scales, v and ε. For this reason, power counting rules can be derived following the EFT approach of NRGR to estimate easily the sizes of different contributions, thus predicting to what order in perturbation they have to be computed, for a given experimental accuracy. Moreover, the NRGR setup makes symmetries manifest at the level of the Lagrangian, providing guidance on the form that the various contributions will take once the appropriate variables

are selected. As we will see, this considerably simplifies the form of the cross-terms compared to the existing literature and it allows to gather some physical intuition about the effects of relativistic multipole corrections to the dynamics.

The very nature of the EFT framework requires to first identify the hierarchy of well-separated length scales involved in a system and remove (integrate out) each of them, one at a time, starting form the shortest.

Thus, we will first focus on the inner binary and integrate out the gravitational field in the presence of an external perturbation, which will be ultimately generated by the third body. The resulting theory will match onto a composite particle, endowed with spin and multipole moments, coupled to gravity. Such a treatment will be valid away from resonances [1], and as long as the ratio of semimajor axes ϵ remains small at all times. This procedure means replacing a three-body problem with a simpler two-body one, where one of the two point-particles is the inner binary.

Here we describe briefly the distinct steps of our approach and outline the structure of the chapter.

1. We will start from a system of three worldlines minimally coupled to gravity, where we have already integrated out the modes whose wavelengths are comparable with the size of the bodies. From this starting point, we will integrate out the off-shell modes that contribute to the gravitational potential of the inner binary, having momenta $k^\mu \sim (v/a, 1/a)$. Thus we will obtain an action describing the gravitational interaction of the two inner bodies in the presence of an external gravitational field. This will be done in Section 4.1.

2. Then we will first expand the Lagrangian in multipoles and, after that, since we are interested in long-time scale evolution, we integrate out the point-particle orbital modes with frequencies $\omega > v/a$. In practice this will be done by averaging over the period of the inner orbit. Doing so, we will obtain the action of a composite particle, whose spin is simply the orbital angular momentum of the inner binary, coupled to an external gravitational field. This step will be carried out in Section 4.2. Although the final result may seem straightforward from an EFT perspective (gauge invariance fixes all the terms in the action to dipolar order without any free parameter, so that the matching might seem superfluous), the computation will allow us to find the exact relation between the center-of-mass choice and the so-called "spin supplementary condition" (SSC), which is a particular gauge choice for the spin tensor.

3. Similarly to the first step, we will then consider the two worldlines, one for the third body and the other for the composite spinning particle representing the inner binary. We introduce the "effective two-body" EFT and show explicitly its power-counting rules in both expansion parameters v and ε. Integrating out the off-shell modes with momenta $k^\mu \sim (V/a_3, 1/a_3)$, we will obtain an action describing the gravitational interaction between the inner binary and the third body.

4. Finally, we will integrate out the remaining point-particle orbital modes with frequencies $\omega > V/a_3$, doing an average over the period

1 If the perturbation was in resonance with the modes of the inner orbital motion, then it would be much more difficult to integrate out these modes and to describe the inner binary as an effective point particle. The same obstacle is encountered in the double averaging procedure, see for instance Appendix A2 of Ref. [262].

of the outer orbit. In this way we will get to a Lagrangian representing the dynamics of the 3-body system as an interaction between the composite particle representing the inner binary and the outer body. These last two steps will be carried out in Section 4.3.

Besides these points, in Section 4.1 we will also comment on the relativistic definition of the center of mass and introduce the osculating orbital elements that describe the perturbed motion of the binary. We elaborate on the relation between the spin kinetic term and the Lagrange planetary equations in Appendix B.1, while in Appendix B.2 we provide details about the specific spin supplementary condition used in this article.

In the present chapter, we will carry out our computations up to dipole and 1PN order. Already at this stage, we will highlight a number of conceptual clarifications arising from the EFT treatment. However, several interesting new terms also arise at quadrupolar order, related to the corrections to the adiabatic approximation [281, 283]. The analysis of these terms will require subtler choices of variables as well as a refined procedure to integrate out the orbital modes, i.e. keeping their backreaction into account. We defer this analysis to Chapter 5.

Given the numerous different symbols appearing in this part of the work, we provide here a dictionary of our notation:

- $\boldsymbol{y}_1$, $\boldsymbol{v}_1$, $\boldsymbol{y}_2$, $\boldsymbol{v}_2$: positions and velocities of the two constituents of the inner orbit, of masses m_1 and m_2;
- $\boldsymbol{y}_3$, $\boldsymbol{v}_3$: position and velocity of the third body, of mass m_3;
- $\boldsymbol{Y}_{\rm CM}$, $\boldsymbol{V}_{\rm CM}$: position and velocity of the center-of-mass of the inner binary, defined in Section 4.1.2
- $\boldsymbol{r} = \boldsymbol{y}_1 - \boldsymbol{y}_2$, $r = |\mathbf{r}|$, $\mathbf{n} = \mathbf{r}/r$ $\boldsymbol{v} = \boldsymbol{v}_1 - \boldsymbol{v}_2$, $\boldsymbol{R} = \boldsymbol{Y}_{\rm CM} - y_3$, $R = |\mathbf{R}|$, $\mathbf{N} = \mathbf{R}/R$, $\boldsymbol{V} = \boldsymbol{V}_{\rm CM} - v_3$;
- $m = m_1 + m_2$ is the mass of the inner binary, $M = m_1 + m_2 + m_3$ is the total mass of the system, $\mu = m_1 m_2/m$ is the reduced mass of the inner and $\nu = \mu/m$ its symmetric mass ratio. Similarly, $\mu_3 = m_3 m/M$ and $\nu_3 = \mu_3/M$ are the reduced mass and symmetric mass ratio of the outer;
- a [a_3]: semimajor axis of the inner [outer] orbit;
- e [e_3]: eccentricity of the inner [outer] orbit;
- $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\gamma}}$ [$\hat{\boldsymbol{\alpha}}_3$, $\hat{\boldsymbol{\beta}}_3$, $\hat{\boldsymbol{\gamma}}_3$]: orthonormal basis of vectors characterizing the inner [outer] orbit, aligned respectively along the semimajor axis (pointing towards the pericenter), the semiminor axis, and the angular momentum;
- Ω, ω, ι: angles characterizing the orientation of the inner orbit (the "orbital elements"), defined by $\hat{\boldsymbol{\alpha}} = R_z(\Omega)R_x(\iota)R_z(\omega)\hat{\mathbf{u}}_x$ where the R_{x_i}'s are rotation matrices along the given axis x_i;
- u [η]: mean [eccentric] anomaly of the inner orbit;
- L, G, H: conjugate momenta to u, ω and Ω respectively, defined in Eq.(B.10);
- $\boldsymbol{J} = \mu\sqrt{G_N m a(1-e^2)}\hat{\boldsymbol{\gamma}}$ [$\boldsymbol{J}_3 = \mu_3\sqrt{G_N M a_3(1-e_3^2)}\hat{\boldsymbol{\gamma}}_3$]: angular momentum vector of the inner [outer] orbit;

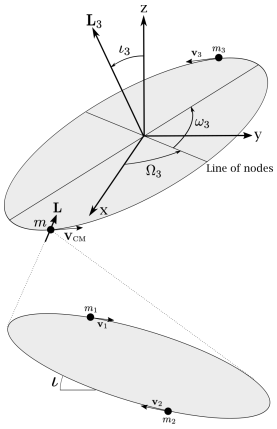

Figure 4.1: Illustration of the "effective two-body" description and of osculating elements. The inner binary is replaced with a point-particle whose spin and multipole moments are related to the osculating elements of the inner orbit.

- $\mathcal{E} = m - G_N m\mu/(2a)$: Total (mass and Newtonian) energy of the inner binary;

4.1 A BINARY SYSTEM IN AN EXTERNAL FIELD

In this Section, we will obtain the effective Lagrangian at the first post-Newtonian order for the inner two-body system using the background field method, which amounts to integrate out the metric fluctuations in the presence of an arbitrary external field. Up to dipole order, we will then explicitly match this Lagrangian to the one of a spinning point-particle coupled to gravity. This spin coupling induces the dominant non-trivial post-Newtonian evolution of the inner binary parameters in the hierarchical three-body problem.

Before integrating out the gravitational field, let us introduce a convenient notation. We will write the Lagrangian of the binary as

$$\mathcal{L} = \frac{1}{2}\mu v^2 + \frac{G_N \mu m}{r} + \mathcal{L}_1 \equiv \mathcal{L}_0 + \mathcal{L}_1 \,, \tag{4.1}$$

where $\mu = m_1 m_2/(m_1+m_2)$ is the reduced mass, $\mathbf{r} = \mathbf{y}_1 - \mathbf{y}_2$, $\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2$ and $\mathcal{L}_1 \equiv \mu\mathcal{R}$ is called the *perturbing function*. For instance, considering only the Newtonian-order perturbation due to the additional Newtonian potential Φ from the third body, the perturbing function reads

$$\mathcal{L}_1 = \frac{1}{2} m V_{\rm CM}^2 - m_1 \Phi(t, \mathbf{y}_1) - m_2 \Phi(t, \mathbf{y}_2) \,, \tag{4.2}$$

where $\mathbf{V}_{\rm CM}$ is the (Newtonian) center-of-mass velocity. The aim of this Section is to compute the 1PN terms in the perturbing function.

4.1.1 *The Lagrangian up to 1PN order*

In order to make the computations as simple as possible, we will use the Kaluza-Klein decomposition space+time of the metric presented in [292, 293], since in the NR regime the time dimension can be considered as compact in comparison to the spatial dimensions. The full metric is decomposed

into a scalar ϕ, a spatial vector A_i and a spatial metric γ_{ij} in the following way:

$$\mathrm{d}s^2 = -e^{2\phi}\left(\mathrm{d}t - A_i \mathrm{d}x^i\right)^2 + e^{-2\phi}\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j \,. \tag{4.3}$$

We take the field action to be the standard Einstein-Hilbert term with a harmonic gauge-fixing term [70],

$$S = \frac{{M_{\rm Pl}}^2}{2}\int \mathrm{d}^4x\sqrt{-g}\,R - \frac{{M_{\rm Pl}}^2}{4}\int \mathrm{d}^4x\sqrt{-g}\,g_{\mu\nu}\Gamma^\mu\Gamma^\nu \,, \tag{4.4}$$

where Γ^μ is the harmonic gauge condition,

$$\Gamma^\mu = \Gamma^\mu_{\nu\rho}g^{\nu\rho} \,. \tag{4.5}$$

In the non-relativistic limit and in the conservative sector of the dynamics, temporal derivatives are treated as an interaction term. Up to 1PN order we will only need the part of the action defining the ϕ and A_i propagators, so that the action simplifies to

$$S = \frac{{M_{\rm Pl}}^2}{2}\int \mathrm{d}^4x\left[2(\partial_\mu\phi)^2 - \frac{1}{2}(\partial_i A_j)^2\right] \,. \tag{4.6}$$

Consequently, the propagators of the (Fourier-space) fields are given in the non-relativistic regime by

$$\langle T\phi(\mathbf{k},t_1)\phi(\mathbf{q},t_2)\rangle = -\frac{i}{2\mathbf{k}^2{M_{\rm Pl}}^2}\delta^3(\mathbf{k}+\mathbf{q})\delta(t_1-t_2) \,, \tag{4.7}$$

$$\langle TA_i(\mathbf{k},t_1)A_j(\mathbf{q},t_2)\rangle = \frac{2i}{\mathbf{k}^2{M_{\rm Pl}}^2}\delta_{ij}\delta^3(\mathbf{k}+\mathbf{q})\delta(t_1-t_2) \,, \tag{4.8}$$

and there is an additional scalar temporal vertex whose expression is $-{M_{\rm Pl}}^2\int \mathrm{d}^4x\,\dot\phi^2$.

To the Einstein-Hilbert term we add two point-particles $A = 1,2$ whose action is

$$\begin{aligned} S_{\rm pp,A} &= -m_A\int \mathrm{d}t\sqrt{-g_{\mu\nu}v_A^\mu v_A^\nu} \\ &= -m_A\int \mathrm{d}t\; e^{\phi}\sqrt{(1-\mathbf{A}\cdot\mathbf{v}_A)^2 - e^{-4\phi}v_A^2} \,, \end{aligned} \tag{4.9}$$

where $v_A^\mu = (1,\mathbf{v}_A)$ is the coordinate velocity of the point-particle. We have set $\gamma_{ij} = \delta_{ij}$ since the fluctuations of γ_{ij} contribute only starting from 2PN order [292]. We expand the point-particle action for weak-field values. At 1PN order, the only vertices contributing are:

$$\begin{aligned} S_{\rm pp,A} = -m_A\int \mathrm{d}t\Bigg(&1 - \frac{v_A^2}{2} - \frac{v_A^4}{8} - \mathbf{A}\cdot\mathbf{v}_A \\ &+ \phi\left(1+\frac{3}{2}v_A^2\right) + \frac{\phi^2}{2}\Bigg) \,. \end{aligned} \tag{4.10}$$

We now use the background field method by splitting the fields according to $\phi = \bar\phi + \tilde\phi$, $A_i = \bar A_i + \tilde A_i$. The tilde quantities correspond to an external arbitrary field (later on, we will relate this field to the one generated by the third point-particle), while we integrate out the barred quantities corresponding to gravitons exchanges between the two bodies. The part of the Lagrangian which does not depend on $\tilde\phi$ and $\tilde A_i$ is the so-called EIH Lagrangian [294]. Since it has already been computed in this framework by

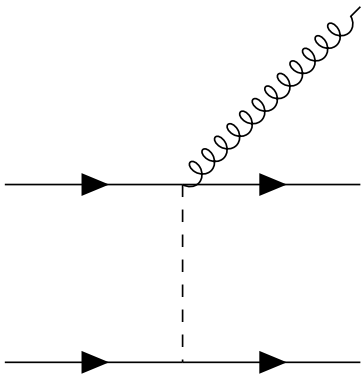

Figure 4.2: Feynman diagram contributing to the emission of one scalar, at order v^2.

several references [70, 292, 293], we will simply give its expression without explicitly computing the relevant Feynman diagrams:

$$\begin{aligned}\mathcal{L}_{\rm EIH} = &\frac{1}{2}m_1 v_1^2 + \frac{1}{2}m_2 v_2^2 + \frac{G_N m_1 m_2}{r} + \frac{1}{8}m_1 v_1^4 + \frac{1}{8}m_2 v_2^4 \\ &+ \frac{G_N m_1 m_2}{2r}\left[3v_1^2 + 3v_2^2 - 7\mathbf{v}_1\cdot\mathbf{v}_2 - \mathbf{v}_1\cdot\mathbf{n}\,\mathbf{v}_2\cdot\mathbf{n} - \frac{G_N m}{r}\right],\end{aligned} \tag{4.11}$$

where $\mathbf{r} = \mathbf{y}_1 - \mathbf{y}_2$, $r = |\mathbf{r}|$ and $\mathbf{n} = \mathbf{r}/r$.

Next, including background fields, we can compute the perturbing function $\mathcal{L}_1$ defined in Eq. (4.1), integrating out $\bar{\phi}$ and $\bar{A}_i$. At 1PN order the result is given by:

$$\begin{aligned}\mathcal{L}_1 \equiv \mathcal{L} - \mathcal{L}_0 = \mathcal{L}_{\rm EIH} - \mathcal{L}_0 &- m_1\tilde{\phi}(\mathbf{y}_1)\left(1 + \frac{3}{2}v_1^2\right) \\ &- \frac{m_1}{2}\tilde{\phi}(\mathbf{y}_1)^2 + m_1\tilde{\mathbf{A}}(\mathbf{y}_1)\cdot\mathbf{v}_1 + \frac{G_N m_1 m_2}{r}\tilde{\phi}(\mathbf{y}_1) \\ &+ (1 \leftrightarrow 2)\,,\end{aligned} \tag{4.12}$$

where $\mathcal{L}_0$ was introduced in Eq. (4.1), and the last term comes from the Feynman diagram with one external $\tilde{\phi}$ and one internal $\bar{\phi}$, represented in Figure 4.2.

4.1.2 *Center-of-mass coordinates*

Given the full 1PN two-body Lagrangian in Eq. (4.12), there remains to expand the two point-particle positions relatively to their common center-of-mass (CM). It is a well-known fact that there is no universal CM definition in General Relativity [295]. For example, the ambiguities in the choice of the CM are related to the so-called "spin supplementary condition" for spinning point-particles, which is a gauge choice for the spin degree of freedom [223, 296]. We provide a discussion about the spin of our system and its relation to the center of mass in Appendix B.2. In our case, we will adopt the standard post-Newtonian definition of the CM, i.e at 1PN order:

$$\begin{aligned} E\mathbf{Y}_{\rm CM} &= E_1\mathbf{y}_1 + E_2\mathbf{y}_2\,, \\ E_A &= m_A + \frac{1}{2}m_A v_A^2 - \frac{G_N m_1 m_2}{2r}\,, \\ E &= E_1 + E_2\,. \end{aligned} \tag{4.13}$$

Conversely, one can express the coordinates $\mathbf{y}_A$ using the relative separation $\mathbf{r}$ and the CM position $\mathbf{X}_{\rm CM}$:

$$\mathbf{y}_1 = \mathbf{Y}_{\rm CM} + (X_2 + \delta)\mathbf{r}\,, \quad \mathbf{y}_2 = \mathbf{Y}_{\rm CM} + (-X_1 + \delta)\mathbf{r}\,, \tag{4.14}$$

where we have defined

$$X_A = \frac{m_A}{m}\,, \quad m = m_1 + m_2\,, \quad \mu = \frac{m_1 m_2}{m}\,, \quad \nu = \frac{\mu}{m} \tag{4.15}$$

and to 1PN order we have:

$$\delta = -\nu \mathbf{V}_{\rm CM}\cdot\mathbf{v} + \nu(X_1 - X_2)\left(\frac{v^2}{2} - \frac{G_N m}{2r}\right). \tag{4.16}$$

In the absence of any external field, the CM follows a straight line in the post-Newtonian coordinates. However, in the hierarchical three-body problem the binary CM will not follow such a trajectory even at the Newtonian level.

We now expand the Lagrangian (4.12) in multipoles, e.g.

$$\begin{aligned}\tilde{\phi}(\mathbf{y}_1) = \tilde{\phi} &+ (y_1 - Y_{\rm CM})^i \partial_i \tilde{\phi} \\ &+ \frac{1}{2}(y_1 - Y_{\rm CM})^i (y_1 - Y_{\rm CM})^j \partial_i \partial_j \tilde{\phi} + \dots\,,\end{aligned} \tag{4.17}$$

where the field is now evaluated at the CM position $\mathbf{Y}_{\rm CM}$. The monopole corresponds to the term involving no derivatives of the fields, the dipole to the term involving first derivatives of the fields and so on.

4.1.3 *Osculating orbital elements*

Before expanding the Lagrangian (4.12) into multipoles and perform a matching with an effective point-particle action, we must eliminate an unwanted degree of freedom from the full theory. Indeed, we want to describe the evolution of the binary over a secular timescale, i.e. a time much longer than the period of the binary itself. In order to do so we average all quantities over the quick periodic motion of the binary, which can be approximated with an ellipse. Indeed, if the motion was purely Newtonian, the trajectory would be described by five constants of motion (six, if we count the initial time), which are nicely packaged in a set of geometrical elements, the so called osculating orbital elements. These are respectively the semimajor axis of the ellipse, the unit vector along the angular momentum and the Runge-Lenz vector:

$$\begin{aligned} a &= -\frac{G_N m}{2}\left(\frac{v^2}{2} - \frac{G_N m}{r}\right)^{-1}\,, \\ \hat{\boldsymbol{\gamma}} &= \frac{\mathbf{r}\times\mathbf{v}}{\sqrt{G_N m a(1-e^2)}}\,, \\ \mathbf{e} &= \frac{1}{G_N m}\mathbf{v}\times(\mathbf{r}\times\mathbf{v}) - \frac{\mathbf{r}}{r}\,. \end{aligned} \tag{4.18}$$

There are two angles in the unit vector $\hat{\boldsymbol{\gamma}}$; furthermore $\mathbf{e}$ is orthogonal to $\hat{\boldsymbol{\gamma}}$ (it points towards the perihelion) and its norm is equal to the eccentricity e. Conversely, the position and velocity vectors can be written as

$$\begin{aligned} \mathbf{r} &= a\left((\cos\eta - e)\,\hat{\boldsymbol{\alpha}} + \sqrt{1-e^2}\sin\eta\,\hat{\boldsymbol{\beta}}\right)\,, \\ \mathbf{v} &= \sqrt{\frac{G_N m}{a}}\frac{1}{1 - e\cos\eta}\left(-\sin\eta\,\hat{\boldsymbol{\alpha}} + \sqrt{1-e^2}\cos\eta\,\hat{\boldsymbol{\beta}}\right)\,, \end{aligned} \tag{4.19}$$

where $\hat{\boldsymbol{\alpha}} = \mathbf{e}/e$, $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\gamma}} \times \hat{\boldsymbol{\alpha}}$ and η is the eccentric anomaly, defined by

$$u = \sqrt{\frac{G_N m}{a^3}} t + \phi = \eta - e \sin \eta \,, \tag{4.20}$$

where ϕ is an arbitrary initial phase, and u is called the mean anomaly.

Now, if the motion is slightly perturbed by post-Newtonian or quadrupolar corrections these constant elements will generically vary slowly with time (compared to the orbital frequency). Thus, in this generic case, we define the osculating elements as the (time-dependent) values of a, e, ϕ, $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\gamma}}$ such that the instantaneous position and velocity of the binary is given by the formulae (4.19). This physically corresponds to drawing at each point the ellipse defined by the instantaneous position and velocity of the binary. We have mapped the six components of $\mathbf{r}$, $\mathbf{v}$ into six elements a, e, ϕ, $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\gamma}}$.

The equations of motion for the binary system can then be translated in a set of first-order equations on the osculating elements, called the Lagrange planetary equations (LPE). For completeness, we recall them in Appendix B.1. For our present purposes, though, it will be sufficient to state the result of Eq.(B.12), i.e. that the orbit-averaged LPE are completely equivalent to a spin kinetic term in the Lagrangian in flat spacetime:

$$\frac{1}{2}\mu v^2 + \frac{G_N \mu m}{r} \to \boldsymbol{J} \cdot \boldsymbol{\Omega} \,, \tag{4.21}$$

where $\mathbf{J}$ is the total angular momentum of the binary and Ω is an angular velocity defined by

$$\boldsymbol{J} = \mu \sqrt{G_N m a (1 - e^2)} \boldsymbol{\gamma} \,, \quad \boldsymbol{\Omega} = \hat{\boldsymbol{\alpha}} \times \hat{\dot{\boldsymbol{\alpha}}} \,. \tag{4.22}$$

Finally, we will average all quantities in the Lagrangian over one period of the binary, using the formula

$$\langle A \rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{d}\eta (1 - e \cos \eta) A(\eta) \,. \tag{4.23}$$

valid to lowest order for any quantity A (we recall that η is the eccentric anomaly defined in Eq. (4.20)). Thus, we will have removed from the Lagrangian the high-energy degree of freedom contained in the mean anomaly. As a consequence of the LPE (B.4), the semimajor axis a will be conserved. This can be understood as deriving from the fact that the conserved conjugate momentum associated to the mean anomaly depends only on a. As a side remark, notice that Eq. (4.23) is valid only if we assume that the binary exactly follows an ellipse. As explained in App. B.1, there will be higher-order corrections to this formula, which however are not needed for our present purposes.

4.2 MULTIPOLE EXPANSION

4.2.1 *The internal Lagrangian*

To begin with, let us deal with the very first term in the Lagrangian (4.12), namely the EIH Lagrangian. This term does not contain any coupling to the gravitational field. As explained before, it still contains the short-distance degree of freedom from the Kepler trajectory of the binary system. In order to remove it and keep only the long-distance degrees of freedom which can

be excited by the external field (in other words, the osculating elements), we should average the Lagrangian over the inner binary timescale, splitting the variables between the center-of-mass and the relative variables.

A priori, one should be careful about the fact that in the Newtonian kinetic energy one should use the relativistic center-of-mass definition in Eq.(B.71). However, the meaning of the supplementary 1PN term will be better understood in terms of spin coupling, so we defer its calculation to a later Subsection. Thus, in this Subsection we stick to the Newtonian definition of the center-of-mass. Carrying out the heavy but straightforward computations, we find by using Eq (4.23):

$$\langle \mathcal{L}_{\rm EIH} - \mathcal{L}_0 \rangle = \frac{1}{2} m V_{\rm CM}^2 + \frac{1}{8} m V_{\rm CM}^4 + 3\mu \frac{G_N^2 m^2}{a^2\sqrt{1-e^2}} - \frac{G_N \mu m}{4a} V_{\rm CM}^2 \,, \tag{4.24}$$

where we have dropped an unimportant constant term in the average (depending on the semimajor axis a only, which is constant in the adiabatic approximation). Each term in Eq. (4.24) lends itself to a very simple interpretation. The two first terms are just the usual relativistic expansion of the center-of-mass velocity $-m\sqrt{1-V_{\rm CM}^2}$. The third term is the average of the EIH Lagrangian of a binary system in isolation: used in the LPE equation (B.8), it gives rise to the celebrated perihelion precession formula. We call such a term the "internal" Lagrangian $\mathcal{L}_{\rm internal}$:

$$\mathcal{L}_{\rm internal} = 3\mu \frac{G_N^2 m^2}{a^2\sqrt{1-e^2}} \,. \tag{4.25}$$

Finally, the meaning of the last term in Eq.(4.24) will become clearer in the next Subsection.

4.2.2 *Monopole*

Starting from Eq.(4.12) we can collect all the terms coupling the binary system to the monopole of the external gravitational field:

$$\mathcal{L}_{\rm monopole} = -m\tilde{\phi}\left(1 + \frac{3}{2}V_{\rm CM}^2 + \frac{3}{2}\nu v^2 - \frac{2G_N\mu}{r}\right) - \frac{m}{2}\tilde{\phi}^2 + m\tilde{\mathbf{A}}\cdot\mathbf{V}_{\rm CM} \,, \tag{4.26}$$

where $\tilde{\phi}$ and $\tilde{\mathbf{A}}$ are evaluated at the CM position $\mathbf{Y}_{\rm CM}$. Note that in the term multiplying $\tilde{\phi}$ in the above equation, we have used the Newtonian version of the CM (i.e., we have set $\delta = 0$ in (B.72)) since the terms involving δ are of higher post-Newtonian order. Averaging over the binary orbital timescale, we find

$$\left\langle \mathcal{L}_{\rm monopole} \right\rangle = -m\tilde{\phi}\left(1 + \frac{3}{2}V_{\rm CM}^2 - \frac{G_N\mu}{2a}\right) - \frac{m}{2}\tilde{\phi}^2 + m\tilde{\mathbf{A}}\cdot\mathbf{V}_{\rm CM} \,, \tag{4.27}$$

Let us now gather this monopole coupling together with the average of the EIH Lagrangian (4.24) computed in the last Subsection. To 1PN order, we find

$$\left\langle \mathcal{L}_{\rm monopole+EIH} \right\rangle = \mathcal{L}_{\rm internal} - m\sqrt{-\tilde{g}_{\mu\nu}V_{\rm CM}^\mu V_{\rm CM}^\nu} + \frac{G_N\mu m}{2a}\left(\tilde{\phi} - \frac{V_{\rm CM}^2}{2}\right) \,. \tag{4.28}$$

To 1PN order, the last term can be exactly accounted for by replacing the mass m of the binary system (which is now treated as an effective point-particle) with its total energy in the worldline Lagrangian:

$$\left\langle \mathcal{L}_{\rm monopole+EIH} \right\rangle = \mathcal{L}_{\rm internal} - \mathcal{E}\sqrt{-\tilde{g}_{\mu\nu}V_{\rm CM}^\mu V_{\rm CM}^\nu} \,, \tag{4.29}$$

where $\mathcal{E}$ is defined as

$$\mathcal{E} = m - \frac{G_N m\mu}{2a} \,. \tag{4.30}$$

Thus, the binary moves in the external field with a total mass equal to its binding energy, as could have been anticipated from an EFT perspective [285]. However, our computation highlights the fact that one should also include the internal Lagrangian in the effective action so that the binary PN precession effects are taken into account.

4.2.3 *Dipole*

Expanding the Lagrangian (4.12) to dipole order (i.e, to first derivatives in the external fields) by taking into account the relativistic CM definition (B.71), we find at 1PN:

$$\begin{aligned}\mathcal{L}_{\rm dipole} = \mu r^i \partial_i \tilde{\phi} \Big[&- 2\mathbf{V}_{\rm CM}\cdot\mathbf{v} \\ &+ (X_1 - X_2)\left(v^2 - \frac{G_N m}{2r}\right)\Big] + \mu r^i v^j \partial_i \tilde{A}_j \,,\end{aligned} \tag{4.31}$$

where $X_A = m_A/m$. As before, one should average this Lagrangian over the inner binary timescale. We find that the term proportional to the difference of masses averages out, leaving us with an averaged Lagrangian

$$\left\langle \mathcal{L}_{\rm dipole} \right\rangle = \frac{\mu}{2}\sqrt{G_N m a(1-e^2)}\epsilon_{ijk}\hat{\gamma}^k \left(2V^i_{\rm CM}\partial^j\tilde{\phi} + \partial^i \tilde{A}^j\right) \,. \tag{4.32}$$

From this expression one easily recognizes the coupling of a spinning point-particle to gravity given in e.g. [223, 285]. In our case, the spin tensor $J_{\mu\nu}$ depends on the total orbital angular momentum of the binary system $\boldsymbol{J} = \mu\sqrt{G_N m a(1-e^2)}\hat{\boldsymbol{\gamma}}$ through the relations

$$J_{ij} = \epsilon_{ijk}J^k \,, \quad J_{0i} = 0 \,. \tag{4.33}$$

The second condition is called a spin supplementary condition, removing the unwanted degrees of freedom from the full spin tensor $J_{\mu\nu}$. As mentioned before, this gauge condition is related to the choice of a center-of-mass of the binary system; our particular CM choice in Eq. (B.71) has selected the spin supplementary condition $J_{0i} = 0$, which has already been discussed e.g. in [223, 297]. We further elaborate on this in Appendix B.2. Furthermore, note that in Eq. (4.33) the spin tensor has been projected to a locally flat frame through $J^{ab} = e^a_\mu e^b_\nu J^{\mu\nu}$, where we have introduced the worldline tetrads defined over all spacetime by $\tilde{g}_{\mu\nu}e^\mu_a e^\nu_b = \eta_{ab}$. As a side remark, note that on top of the spin supplementary condition, the components of the spin vector are not all independent degrees of freedom following the remark below Eq. (B.13). This reflects the fact that the spin of the inner binary contains two degrees of freedom once the orbital timescale has been integrated out, instead of the three degrees of freedom contained in the Euler angles of a generic spin.

To 1PN order, the spin coupling (4.32) can be written in a compact form using the Ricci rotation coefficients:

$$\left\langle \mathcal{L}_{\rm dipole} \right\rangle = \frac{1}{2}J_{ab}\omega^{ab}_\mu V^\mu_{\rm CM} \,, \quad \omega^{ab}_\mu = e^{a\nu}D_\mu e^b_\nu \,. \tag{4.34}$$

This formula gives back our previous equation (4.32) when expanded for weak-field values [285, 287]. We may be tempted to add to this spin coupling the kinetic term for the spin in Eq. (4.21) to obtain the minimal gravitational spin coupling which has been discussed at length in the NRGR formalism [71, 223]:

$$\mathcal{L}_{\rm spin} = \frac{1}{2} J_{\mu\nu}\Omega^{\mu\nu} \,. \tag{4.35}$$

In this equation the total angular velocity $\Omega^{\mu\nu}$ includes both the Ricci rotation coefficients from Eq. (4.34) and the locally flat angular velocity from Eq. (4.21). It is defined through

$$\Omega^{\mu\nu} = e^{\mu}_{a} e^{\nu}_{b} \left(\Omega^{ab}_{\rm flat} + V^{\alpha}_{\rm CM} \omega^{ab}_{\alpha} \right) \,, \tag{4.36}$$

Here $\Omega^{ab}_{\rm flat}$ is related to the tensor $\Omega^{ij} = \epsilon_{ijk}\Omega^{k}$ by a relation that we discuss in Appendix B.2, and the rotation vector $\mathbf{\Omega} = \hat{\boldsymbol{\alpha}} \times \dot{\hat{\boldsymbol{\alpha}}}$ has been defined in Eq (5.2). However, there is a small piece that is still missing to obtain the full Eq. (4.35), related to the choice of the center-of-mass. As we show in Appendix B.2, in the spin gauge we are using ($J^{0i} = 0$), there should be a supplementary spin kinetic term related to Thomas precession, which is 1PN order higher than the kinetic term (4.21):

$$\frac{1}{2} J_{\mu\nu}\Omega^{\mu\nu} \supset \mathbf{J} \cdot \mathbf{\Omega} + \frac{1}{2} J_{ij} A^{i}_{\rm CM} V^{j}_{\rm CM} \,, \tag{4.37}$$

where $\mathbf{A}_{\rm CM}$ is the acceleration of the center-of-mass. Such a term is related to the PN corrections to the center-of-mass position (and speed) which we ignored in Section 4.2.1. Indeed, using the full CM definition (B.71) in the Newtonian part of the EIH Lagrangian (4.11) gives a supplementary 1PN order term,

$$\mathcal{L}_{\rm Thomas} = m \mathbf{V}_{\rm CM} \cdot \frac{\mathrm{d}}{\mathrm{d}t} \left(\delta \mathbf{r} \right) \,, \tag{4.38}$$

where δ has been defined in Eq. (B.73). At first sight, we may be tempted to discard such a term when averaging out the internal binary timescale. However, one should not forget to take also the time derivative acting on $\mathbf{V}_{\rm CM}$ in δ, giving rise to

$$\left\langle \mathcal{L}_{\rm Thomas} \right\rangle = -\mu \left\langle r^{i} v^{j} \right\rangle V^{i}_{\rm CM} A^{j}_{\rm CM} = \frac{1}{2} J_{ij} A^{i}_{\rm CM} V^{j}_{\rm CM} \,, \tag{4.39}$$

which is exactly the additional spin kinetic term shown in Eq. (4.37).

4.2.4 *Quadrupole*

From the EFT point of view, at 1PN quadrupolar order the couplings to gravity are contained in two non-minimal worldline operators [291]:

$$\mathcal{O}_1 = \frac{1}{2} \int \mathrm{d}\tau E_{ij} Q^{ij}_{E} \,, \quad \mathcal{O}_2 = -\frac{4}{3} \int \mathrm{d}\tau B_{ij} Q^{ij}_{B} \,, \tag{4.40}$$

where Q^{ij}_{E} and Q^{ij}_{B} are the electric-type and magnetic-type quadrupole moments of the source, coupled to the corresponding parts of the Weyl tensor $C_{\mu\nu\alpha\beta}$:

$$\begin{aligned} E_{\mu\nu} &= C_{\mu\nu\alpha\beta} V^{\alpha}_{\rm CM} V^{\beta}_{\rm CM} \,, \\ B_{\mu\nu} &= \frac{1}{2} \epsilon_{\mu\alpha\beta\sigma} C^{\alpha\beta}{}_{\nu\rho} V^{\sigma}_{\rm CM} V^{\rho}_{\rm CM} \,. \end{aligned} \tag{4.41}$$

Furthermore,in Eq. (4.40) the tensors have been projected to the locally flat frame defined below Eq. (4.33).

We could proceed as before and carry out the integration procedure to obtain the quadrupole moment of the effective point-particle. However, at this order the procedure is somewhat more involved than one could naively expect. The first complication comes from the corrections to the time averages introduced in Eq. (4.23). Indeed, post-Newtonian corrections to the period of the system will matter when taking the average of the Newtonian quadrupole moment, combining to produce a quadrupolar 1PN term. In the same way, the Newtonian quadrupolar corrections to the motion of the inner binary should be taken into account in the average of the EIH Lagrangian.

The second complication comes from the corrections to the adiabatic approximation mentioned in the introduction. Indeed, in our analysis we are assuming that all the variables of the inner binary vary on long timescales (except of course the mean anomaly). This neglects short-timescale oscillations, which can ultimately have an effect on long-wavelength modes [298, 299]. It turns out that at lowest order this effect produces cross-terms of 1PN quadrupolar order [281, 283] (no such corrections appear at lower multipole orders). While noting in passing that these kind of corrections have a very transparent meaning in the EFT language (they are high-energy corrections to an effective low-energy action), we will defer their complete calculation to Chapter 5.

4.3 INTEGRATING OUT THE OUTER BINARY TIMESCALE

Now that we have replaced the binary system with an effective point-particle, we can integrate out the external fields $\tilde{\phi}$, $\tilde{\mathbf{A}}$ in the presence of a third point-particle of mass m_3. For simplicity, in the following we will assume this mass to be of the same order of the mass of the inner binary: $m_3 \sim m$. We will first derive the Feynman rules of the effective point-particle; then, in a second step, we will integrate out the outer binary timescale and comment on the different terms obtained in the expansion of the Lagrangian. For the 1PN precision we aim to, it will be sufficient to set the total Newtonian center-of-mass of the three-body system to the origin of coordinates (it will accelerate only at 2PN order [300]). Thus, we will have the expressions

$$\mathbf{Y}_{\rm CM} = X_3 \mathbf{R}\,, \quad \mathbf{y}_3 = -X_{\rm CM}\mathbf{R}\,, \tag{4.42}$$

where we recall that $\mathbf{Y}_{\rm CM}$ is the position of the center-of-mass of the inner binary, and we have defined $\mathbf{R} = \mathbf{Y}_{\rm CM} - \mathbf{y}_3$, $R = |\mathbf{R}|$, $\mathbf{N} = \mathbf{R}/R$, $M = m_1 + m_2 + m_3$, $X_3 = m_3/M$ and $X_{\rm CM} = m/M$. The averages over the outer binary timescale are then taken in the same way than in the preceding Section.

4.3.1 *Power-counting rules*

Let us recap what we have learned so far and set up power-counting rules for the vertex coupling the binary system (now treated as an effective point-particle) to gravity. Up to dipole order, the Lagrangian of the binary system can be written as

$$\mathcal{L} = \mathcal{L}_{\rm internal} - \mathcal{E}\sqrt{-\tilde{g}_{\mu\nu} V^\mu_{\rm CM} V^\nu_{\rm CM}} + \frac{1}{2} J_{\mu\nu}\Omega^{\mu\nu}\,. \tag{4.43}$$

Note that this Lagrangian has not yet been averaged over the period of the outer orbit T_3, and can therefore describe the secular dynamics on timescales

shorter than T_3. With such a simple Lagrangian, one can assign the standard power-counting rules of NRGR which have been described in e.g [70, 213, 301], considering the motion of the effective point-particle and the third mass (the outer orbit) for which one has $V_{\rm CM}^2 \sim v_3^2 \sim GM/a_3$ where a_3 is the semimajor axis of the outer orbit. Thus, spatial derivatives are treated as $\partial_i \sim a_3^{-1}$. Time intervals scale as $t \sim a_3/V_{\rm CM}$ and the metric perturbations scale as $\tilde{\phi} \sim \tilde{A}_i \sim V_{\rm CM}^{1/2}(M_{\rm Pl}a_3)^{-1}$. As usual in NRGR, the lowest-order Lagrangian scales as the orbital angular momentum of the outer orbit $J_3 \sim MV_{\rm CM}a_3$ which is treated non-perturbatively, higher-order corrections coming with higher powers of $V_{\rm CM}$.

However, one difference with respect to the standard NRGR power-counting rules is evidently the presence of two expansions, the first one in v and the second one in $\varepsilon \equiv a/a_3$. *A priori*, we could also have an expansion in the post-Newtonian parameter of the outer orbit $V_{\rm CM}$. However, not all these parameters are independent. We choose to write all the post-Newtonian corrections as an expansion in the velocity of the inner binary v, converting the center-of-mass velocity by means of the relation $V_{\rm CM} \sim v\varepsilon^{1/2}$, which holds when $m \sim m_3$. In Table 4.1 we give the power-counting rules of the monopole and dipole vertex which we computed in Sections 4.2.2 and 4.2.3. The effect of post-Newtonian corrections on the dynamics of the system is highly non-trivial, as it can lead to suppression as well as enhancement of the Kozai-Lidov oscillations depending on the part of parameter space explored [249]; we expect that our power-counting scheme will help in discriminating between the different behaviors observed.

Notice that the scaling of the spin is somewhat different than the one usually presented in NRGR [71, 287]. Indeed, when taking compact objects as point-particles the spin is given as an order-of-magnitude by

$$J \sim m r_s v_{\rm rot} < m r_s \,, \tag{4.44}$$

where $v_{\rm rot}$ is the rotation velocity of the object and $r_s \sim G_N m$ its size. As a consequence, the ratio of the spin coupling presented in Eq. (4.32) to the Newtonian gravitational coupling is of v^3 (1.5PN) order. However, in our case the spin order-of-magnitude is given by $J \sim \mu\sqrt{G_N m a}$ so that the ratio of (4.32) to the Newtonian coupling is

$$\frac{JV_{\rm CM}\partial\phi}{m\phi} \sim v^2\varepsilon^{3/2} \,. \tag{4.45}$$

Thus, the inner binary angular momentum coupling is formally of 1PN order, although it is suppressed by the small ratio $\varepsilon^{3/2}$. This power-counting is different from the one of the Lense-Thirring precession caused by the intrinsic spin of the objects, which has been studied in [267, 279] and enters at 1.5 PN.

4.3.2 *Monopole*

Let us begin by integrating out the vertex contained in the monopole operators of the effective binary system, i.e in the square root appearing in Eq. (4.43). The final effective action, including orders of $J_3 v^2 \varepsilon$, is given by:

$$\mathcal{L}_{\leq v^2\varepsilon} = \mathcal{L}_{\rm internal} + \tilde{\mathcal{L}}_{\rm EIH}^{\rm CM,3} \,, \tag{4.46}$$

where $\mathcal{L}_{\rm EIH}^{\rm CM,3}$ is the EIH Lagrangian of the system composed by the CM (of mass $\mathcal{E}$, defined in Eq. (4.29)) and the third particle.

Operator	Rule
$\frac{1}{2}mV_{\rm CM}^2$	J_3
$-m\tilde{\phi}$	$J_3^{1/2}$
$m\tilde{\mathbf{A}}\cdot\mathbf{V}_{\rm CM}$	$J_3^{1/2}v\varepsilon^{1/2}$
$\frac{1}{8}mV_{\rm CM}^4$	$J_3v^2\varepsilon$
$-\frac{3}{2}m\tilde{\phi}V_{\rm CM}^2$	$J_3^{1/2}v^2\varepsilon$
$-\frac{1}{2}m\tilde{\phi}^2$	$v^2\varepsilon$
$\frac{G_N\mu m}{2a}\tilde{\phi}$	$J_3^{1/2}v^2$
$-\frac{G_N\mu m}{4a}V_{\rm CM}^2$	J_3v^2
$J_{ij}V_{\rm CM}^i\partial^j\tilde{\phi}$	$J_3^{1/2}v^2\varepsilon^{3/2}$
$\frac{1}{2}J_{ij}\partial^i\tilde{A}^j$	$J_3^{1/2}v\varepsilon$
$\frac{1}{2}J_{ij}A_{\rm CM}^iV_{\rm CM}^j$	$J_3v^2\varepsilon^{3/2}$

Table 4.1: Power-counting rules for the vertices obtained by expanding the effective point-particle action (4.43) up to 1PN order, with $J_3 = (G_NM^3a_3)^{1/2}$, $v^2 = Gm/a$ and $\varepsilon = a/a_3$. For convenience, the integral over time is not displayed, although it should be included to obtain a dimensionless rule.

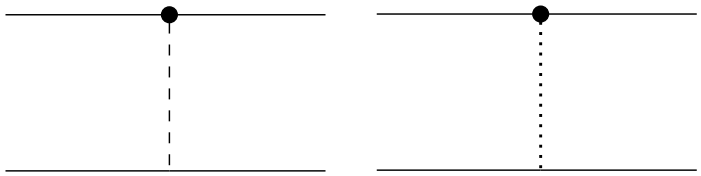

Figure 4.3: Feynman diagrams contributing to the lowest-order spin-orbit coupling, at order $J_3 v^2 \varepsilon^{3/2}$. The dot represent the insertion of a spin coupling from Eq. (4.32). The dotted line represents propagation of a scalar ϕ, while the dashed line stands for the propagation of a vector $\mathbf{A}$.

The Lagrangian in Eq. (4.46) involves a non-trivial coupling between the variables of the inner and outer binaries, given by

$$\mathcal{L}_{v^2} = -\frac{G_N \mu m}{2a}\left(\frac{V_{\rm CM}^2}{2} + \frac{G_N m_3}{R}\right) . \tag{4.47}$$

This contribution is of order v^2 with respect to the standard Newtonian term $\mathcal{L}_0 \sim G_N M/a_3$. We average this term over one orbit of the outer binary, which gives

$$\mathcal{L}_{v^2} = -\frac{G_N^2 M^2 \mu \nu_3}{2a a_3}\left(1 + \frac{X_3}{2}\right) , \tag{4.48}$$

where $M = m_3 + m$, $X_3 = m_3/M$ and $\nu_3 = m m_3/M$. This new monopole coupling has no effect on the dynamics. Indeed it depends only on the semi-major axes a and a_3. Consequently, in the Lagrange planetary equations this term will only enter in the equation for the mean anomaly (B.7), which is surely irrelevant in the adiabatic approximation. Therefore, at the level of the monopole, the resulting motion is the one of two ellipses precessing because of standard two-body GR effects.

In fact, one can be quite generic about the monopole terms. Indeed, the only planetary elements upon which the monopole terms could depend are the semimajor axes a, a_3 and the eccentricities e, e_3 (they do not involve angles). In the LPE the derivatives with respect to these elements enter only in the equations for the mean anomaly (B.7) and the perihelion angle (B.8). Thus, the only effect that monopole terms can have is to make the ellipses precess.

4.3.3 *Dipole*

In order to integrate out modes contributing to the potential at dipole order, we have to compute the diagrams related to spin-orbit coupling. These are shown in Figure 4.3. Using the Lagrangian averaged over the inner orbit Eq. (4.32), we find:

$$\mathcal{L}_{\rm spin-orbit} = \frac{1}{2} J_{ij} \frac{G_N m_3}{R^3} R^i \left(4 v_3^j - 2 V_{\rm CM}^j\right) . \tag{4.49}$$

At this order of approximation however, we should also take into account the Thomas precession term of Eq. (4.39). This gives a contribution of the same size of the terms in Eq.(4.49). We can replace the center-of-mass acceleration in Eq. (4.39) using the equation of motion, since the difference

between the two terms would contribute at a higher PN order (this is usually called the "double zero trick" [302, 303]). Thus, at order $J_3 v^2 \varepsilon^{3/2}$ the full Lagrangian is given by

$$\begin{aligned}\mathcal{L}_{v^2\varepsilon^{3/2}} &= \frac{1}{2} J_{ij} \frac{G_N m_3}{R^3} R^i \left(4 v_3^j - 3 V_{\rm CM}^j\right) \\ &= -\frac{1}{2} J_{ij} \frac{G_N m_3 (4m + 3m_3)}{M R^3} R^i V^j \,,\end{aligned} \tag{4.50}$$

which recovers the result already known in the NRGR approach [287]. Carrying out the average over the outer binary timescale in a way very similar to the previous Section, we find

$$\langle \mathcal{L}_{v^2\varepsilon^{3/2}} \rangle = -\frac{4m + 3m_3}{2m} \frac{G_N}{a_3^3 (1 - e_3^2)^{3/2}} \, \mathbf{J} \cdot \mathbf{J}_3 \,, \tag{4.51}$$

where $\mathbf{J}_3$ is the angular momentum vector of the outer orbit, $\mathbf{J}_3 = \mu_3 (G_N M a_3 (1 - e_3^2))^{1/2} \hat{\boldsymbol{\gamma}}_3$ (here $\hat{\boldsymbol{\gamma}}_3$ is the unit vector along the outer orbit angular momentum, and $\mu_3 = m m_3 / M$). Thus, this term is indeed a coupling between the angular momentum vectors of the two orbits.

From this expression one can obtain a precession equation for the inner orbit angular momentum. Indeed, varying the kinetic term for the spin with respect to the canonical variables $\hat{\boldsymbol{\alpha}}$ and $\boldsymbol{W} = \boldsymbol{J} \times \hat{\boldsymbol{\alpha}}$, one obtains the equations of motion

$$\frac{\mathrm{d}\boldsymbol{W}}{\mathrm{d}t} = -\Omega_{\rm prec} \boldsymbol{W} \times \boldsymbol{J}_3 \,, \tag{4.52}$$

$$\frac{\mathrm{d}\hat{\boldsymbol{\alpha}}}{\mathrm{d}t} = \Omega_{\rm prec} \boldsymbol{J}_3 \times \hat{\boldsymbol{\alpha}} \,, \tag{4.53}$$

where the precession frequency is equal to

$$\Omega_{\rm prec} = \frac{4m + 3m_3}{2m} \frac{G_N}{a_3^3 (1 - e_3^2)^{3/2}} \,. \tag{4.54}$$

From these two equations, and using the Jacobi identity for the cross product, one obtains the precession equation

$$\frac{\mathrm{d}\boldsymbol{J}}{\mathrm{d}t} = \Omega_{\rm prec} \boldsymbol{J}_3 \times \boldsymbol{J} \,, \tag{4.55}$$

which is in complete accordance with earlier results on the hierarchical three-body problem [267, 283]. Notice that conservation of the total angular momentum requires that $\boldsymbol{J}_3$ satisfies an analogous equation,

$$\frac{\mathrm{d}\boldsymbol{J}_3}{\mathrm{d}t} = \Omega_{\rm prec} \boldsymbol{J} \times \boldsymbol{J}_3 \,. \tag{4.56}$$

In particular, it was shown that this angular momentum precession may play an important role for stellar-mass binary mergers near a supermassive BH [267]. Quadrupolar terms would lead to further precession effects, of order $J_3 v^2 \varepsilon^2$ in the Lagrangian. We leave the computation and the astrophysical implications of such terms to further work.

4.4 SUMMARY

The NRGR approach to the two-body problem was designed to deal with extended compact objects. In this article, we have extended NRGR to the

setting of a hierarchical three-body problem. In the approximation that the inner orbit is much smaller in amplitude than the outer one, the inner binary system can be replaced by an effective point-particle endowed with multipole moments, which we explicitly computed up to dipole order. This is very natural from the EFT perspective and provides a new specific example of how an extended (and not so compact) system can be accounted for by means of a point-particle operator.

Our procedure consists in integrating out the short timescales associated with the period of the two hierarchical orbits. One notable result of this approach is to make explicit the link between the Lagrange planetary equations, describing the long-time evolution of the inner binary Keplerian parameters, and the kinetic term for a spin in the EFT language. We have also clarified the relation between the post-Newtonian definition of the center-of-mass and the spin supplementary condition for the angular momentum of the inner binary. The computation of quadrupolar PN terms including the corrections to the adiabatic approximation will be the subject of Chapter 5.

Our study moves towards a more systematic characterization of the dynamics of a binary system perturbed by a distant third body.

5

QUADRUPOLE RELATIVISTIC CORRECTIONS FROM A DISTANT THIRD BODY

In Chapter 4 we followed insights from NRGR and introduced a new EFT approach to the relativistic, hierarchical three-body problem to derive systematically a description of its dynamics on long time-scales. Thanks to the double perturbative expansion in v and ε made possible by this setup, the dynamics of a hierarchical three-body system can be matched to a simpler two-body interaction, in which the inner binary is described as a single point-particle endowed with multipole moments. This was achieved by performing an averaging procedure at the level of the Lagrangian. Most noticeably, the EFT approach exploits symmetries that are manifest in the effective Lagrangian, restricting the form of the allowed interaction terms. Moreover, working with a single functional rather than with several equations of motion makes it simpler to setup a systematic study of the three body system.

In Chapter 4 we presented the EFT setup and derived the effective Lagrangian describing the system on long time-scales up to 1PN dipole order, i.e. up to order $v^2\varepsilon^{3/2}$ beyond the leading Newtonian interaction. Instead, here we extend this computation up to 1PN quadrupolar order, i.e. $v^2\varepsilon^{5/2}$ beyond leading order. At this order, computations are substantially more complex with respect to what we have seen in the previous chapter. A first source of complexity is due to the averaging procedure. At lower orders the averaging can be performed in the so-called adiabatic approximation, i.e. neglecting variations of slowly evolving variables during the average over the period of both orbits. Instead, when accounting for terms of mixed quadrupolar and 1PN order, deviations from adiabaticity must be taken into account. In addition to this, backreaction from quickly oscillating terms that are suppressed in amplitude will also affect the averaging, contrarily to what happened at lower orders. We address these complications by following the method of near-identity transformations [72], which allows to consistently implement the averaging procedure to any order of accuracy. While several authors already studied quadrupolar couplings at 1PN order [249, 267, 279–284], we are aware of only three which took into account these deviations from the adiabatic approximation [268, 281–283]. However, we believe that we give in this work the first complete expressions of quadrupolar 1PN terms. Indeed, only the particular case of a circular outer orbit is considered in [268, 281, 282] neglecting some PN interactions that we describe in this work. On the other hand, the derivation in [283] reports a puzzling result that we will mention in Section 5.3.

Another source of complexity lies in finding the suitable dynamical variables to efficiently package relativistic corrections in our results. While the idea at the core of our approach of identifying the inner binary to a spinning point-particle with multipole moments is very intuitive, providing a definite relation between the variables describing the inner binary and the pa-

rameters of the effective point-particle is subtle in practice. For example, in Chapter 4 we showed how the choice of a Spin Supplementary Condition (a gauge condition on the spin tensor of the effective point-particle [223, 296]) is related to the center-of-mass choice of the inner binary. In the present computations, two new similar subtleties arise. The first one concerns the definition of osculating elements describing both inner and outer orbits. In our previous work, we followed the usual convention and used the osculating *orbital* elements defined as the parameters of the ellipse instantaneously tangent to the trajectory (described with positions and velocities). However, at 1PN quadrupolar order we find that it becomes more convenient to use osculating *contact* elements, which are defined through momenta rather than velocities [304]. Since PN corrections induce a non-trivial relation between momentum and velocity, these two sets of osculating elements will differ in general. One remarkable conclusion of the present analysis is that, while the slowly evolving part of the contact semimajor axes are conserved throughout the evolution of the system (as is common in long-timescale evolution of triple systems [227]), their orbital counterpart features small variations over long time-scales, which offers a new point of view on earlier findings of [268, 281–283].

The second subtle point in the matching between inner binary and point-particle is that the quantities describing the inner binary system are inherently defined in the rest frame of its center of mass, which is accelerating because of the presence of the third body. This entails non-trivial relations between the absolute positions of the inner binary components, defined in the rest frame of the three-body center-of-mass, and the relative quantities defined in the binary rest frame, as we show in Section 5.2.2. As far as we know, this point went so far unnoticed in the relativistic three-body literature. While this step just amounts to a redefinition of the osculating elements of the inner binary, it proves to be crucial in order to perform correctly the matching procedure described in Section 5.2.4.

Let us now describe in more detail the organization of this chapter. In Section 5.2, we will explain how to refine the procedure of integrating out the fast dynamics in order to fully capture relativistic quadrupole effects. In order to keep the discussion as simple as possible, we have deferred the computation of beyond-adiabatic corrections to Appendix B.5 and use only the final result of this appendix in the main text. In Section 5.3 instead we integrate out the gravitational field due to the third body. Finally, in Section 5.4 we provide a numerical solution implementing the new relativistic interaction derived in the present work and we show how it influences the long time-scale dynamics in the case of a particular three-body system. The rest of the Appendices are devoted to: a presentation of the averaging procedure that we employ (Appendix B.3); a discussion of the conservation of the semimajor axis (Appendix B.4); the derivation of the expressions connecting the absolute coordinates of the two inner bodies in the three body rest frame to their relative coordinates in the inner binary rest frame (Appendix B.7); an independent computation of the so-called quadrupole-squared terms of [298] (Appendix B.8).

As mentioned, in contrast with Chapter 4, we will work out the results in terms of the contact elements, the variables in which the PN Hamiltonian has the simplest expression. With respect to the usual Newtonian orbital elements, contact elements include PN corrections. We review their precise definition in Section 5.1 and examine their difference from

orbital elements in Appendix B.9. For ease of notation, we will denote them by the same symbols we used in Chapter 4 for the orbital elements. Orbital elements will instead be denoted with tilded symbols from now on.

With these improvements with respect to the approach followed in Chapter 4, we will be able to obtain a total Lagrangian:

$$\begin{aligned}\mathcal{L} =& \boldsymbol{J}\cdot\boldsymbol{\Omega} + \boldsymbol{J}_3\cdot\boldsymbol{\Omega}_3 + 3\mu\frac{G_N^2 m^2}{a^2\sqrt{1-e^2}} + 3\mu_3\frac{G_N^2 M^2}{a_3^2\sqrt{1-e_3^2}} \\ &- \frac{4m+3m_3}{2m}\Omega_{\rm prec}\boldsymbol{J}\cdot\boldsymbol{J}_3 + \left\langle \mathcal{L}_{\rm quad}^{\leq v^2\varepsilon^{5/2}} \right\rangle \,.\end{aligned} \tag{5.1}$$

In this equation, $\mathbf{J}$ ($\mathbf{J}_3$) and $\boldsymbol{\Omega}$ ($\boldsymbol{\Omega}_3$) are the angular momentum and rotation vectors of the inner (resp. outer) orbits, defined by

$$\boldsymbol{J} = \mu\sqrt{G_N m a(1-e^2)}\boldsymbol{\gamma}\,, \quad \boldsymbol{\Omega} = \boldsymbol{\alpha}\times\dot{\boldsymbol{\alpha}}\,, \tag{5.2}$$

with analogous formulas for the outer orbit. Let us comment on each component of Eq (5.1). The first two terms are the spin kinetic terms of the two orbits. Once a variational principle is applied, they will give rise to first-order evolution equations for the planetary elements of the two orbits, the Lagrange Planetary Equations (LPE), see Appendix B.1. Note that, as mentioned before, the spin $\mathbf{J}_3$ of the outer orbit is defined by using the 1PN energy $\mathcal{E}$ of the binary system as its effective mass, so that the spin kinetic term of the outer orbit secretly hides post-Newtonian terms (the other term where $\mathbf{J}_3$ appears in (5.1) is already of 1PN order, so that in this term the difference between using m or $\mathcal{E}$ is of 2PN order). The two next terms correspond to the well-known 1PN potentials inducing perihelion precession of the two orbits at 1PN order [305]. We refer to them as internal Lagrangians since they would be present even without any interaction between the two orbits. The third term is the coupling between the angular momenta of the two orbits at dipole order $v^2\varepsilon^{3/2}$ that we computed in the previous chapter. It involves a precession frequency given by Eq. (4.54). Finally, the term $\mathcal{L}_{\rm quad}^{\leq v^2\varepsilon^{5/2}}$ encodes the contributions from quadrupole-suppressed 0PN and 1PN interactions to the long time-scale dynamics, which we will compute in the next sections and which is the main result of the present chapter.

5.1 CONTACT ELEMENTS

As mentioned above, we find convenient to express all of the results of this chapter in terms of contact elements, which differ from the usual Newtonian orbital elements by PN corrections. Such differences will only become relevant if one considers high enough terms in the perturbative analysis of the problem, which happens to be the case for this chapter. The key difference is that contact elements are defined in terms of the canonical momenta of the system, rather than in terms of the velocities (see e.g. [304]). This allows to keep track of some PN corrections in a compact way, since the conjugate momenta will receive corrections as soon as the interaction Lagrangian depends on the velocities.

Let us now define the contact elements. In the rest frame of the inner binary center of mass, we can write the Lagrangian of the inner binary system as

$$\mathcal{L} = \frac{1}{2}\mu v^2 + \frac{G_N m\mu}{r} + \mu\mathcal{R}\,, \tag{5.3}$$

where $\mathcal{R}$ encodes any term in the Lagrangian beyond Newton's expression for the inner binary system. The conjugate momentum to the coordinate $\boldsymbol{r}$ will be:

$$\boldsymbol{p} = \mu \boldsymbol{v} + \mu \partial \mathcal{R} / \partial \boldsymbol{v} \,. \tag{5.4}$$

Then, rather than using Eq. (4.18), we will define the contact elements as follows:

$$\begin{aligned} a &= -\frac{G_N m}{2}\left(\frac{p^2}{2\mu^2} - \frac{G_N m}{r}\right)^{-1} , \\ e\boldsymbol{\alpha} &= \frac{1}{G_N m \mu^2}\mathbf{p}\times(\mathbf{r}\times\mathbf{p}) - \frac{\mathbf{r}}{r} , \\ \gamma &= \frac{\mathbf{r}\times\mathbf{p}}{\sqrt{G_N m \mu^2 a(1-e^2)}} , \end{aligned} \tag{5.5}$$

where the unit vector γ contains two angles and it is orthogonal to $\boldsymbol{\alpha}$. With these definitions, in analogy with the Kepler problem, we have:

$$\begin{aligned} \mathbf{r} &= a\left((\cos\eta - e)\,\boldsymbol{\alpha} + \sqrt{1-e^2}\sin\eta\,\boldsymbol{\beta}\right) , \\ \mathbf{p} &= \sqrt{\frac{G_N m}{a}}\frac{\mu}{1-e\cos\eta}\left(-\sin\eta\,\boldsymbol{\alpha} + \sqrt{1-e^2}\cos\eta\,\boldsymbol{\beta}\right) . \end{aligned} \tag{5.6}$$

where $\boldsymbol{\beta} = \gamma \times \boldsymbol{\alpha}$ and η is the (contact) eccentric anomaly, defined at a time t by

$$u \equiv nt + \phi = \sqrt{\frac{G_N m}{a^3}}t + \phi = \eta - e\sin\eta \,, \tag{5.7}$$

with ϕ the initial phase, and u the contact equivalent of the mean anomaly. In the following, it will be useful to have an explicit relation between the contact elements $(a, e, u, \omega, \Omega, \iota)$ and the orbital elements $(\tilde{a}, \tilde{e}, \tilde{u}, \tilde{\omega}, \tilde{\Omega}, \tilde{\iota})$. We derive such a relation in Appendix B.9, see Eqs. (B.110)-(B.113). An important outcome of Appendix B.9 which we will use in the main text is that the difference between orbital and contact elements is small in the sense that at *any time*, they differ by a 1PN quantity (i.e. this difference cannot grow on long timescales).

5.2 THE POINT-PARTICLE EFT TO QUADRUPOLAR ORDER

We now derive an effective Lagrangian describing the inner binary coupled to an external gravitational field on time-scales much longer than its orbital period. We start by describing the procedure to integrate out the fast (orbital) modes, the so-called averaging, which we present in detail in Appendix B.3. We then discuss in which reference frame to define the contact elements of the inner binary, highlighting the rest frame of the inner binary's center of mass as the most suitable choice. Having done that, we present the results of the averaging procedure, which is carried out in some detail in Appendix B.5. We close the section by matching the result to a world-line action for the inner binary.

5.2.1 *Integrating out fast modes*

In field theoretic terms, we wish to derive the effective action for the slow modes (the long time-scale dynamics) by integrating out the fast modes (the

orbital dynamics). In the classical limit, this corresponds to substituting the solutions to the equations of motion for the fast modes in terms of the slow modes in the Lagrangian. On top of that, we can average this effective Lagrangian, so as to remove any quickly oscillating terms, which only carry information about the already solved short time-scale dynamics. In our specific case, fast and slow modes are packed together in our variables, the osculating (orbital or contact) elements. Therefore, we need a way to split the dynamics over short and long time-scales. Intuitively, a method to achieve this is considering the average and the average-free part (with respect to the orbital time-scale) of the original equations of motion. The splitting allows to solve for the fast modes in terms of the slow ones, by solving the average-free equations. This method is broadly referred to as averaging, see for instance [72, 306].

Despite its intuitiveness, averaging presents a few subtleties which must be clarified in order to set up a systematic and consistent procedure. For instance, if the period of the fast oscillations of our system changes slowly, we have to understand how to account for variations of the period in our averages. Even more, we need to account for the small changes of the slow variables over the period of a quick oscillation. Up to which order can we compute averages while keeping fixed the slow variables?

While it is easy to estimate the size of these possible corrections, the task of choosing a set-up that makes transparent how to deal with these issues is much less straightforward. For instance, one can consider whether to average with respect to time or with respect to a dynamical variable (e.g. a time-dependent angle). We can promptly see that these two choices will lead to different quantities. For instance, suppose the dynamics has periodicity with respect to an angle u. If we call $T[\ell]$ the period as a function of the slow variables ℓ, the average of a quantity $A(\ell, u)$ can be defined in two different ways:

$$\langle A\rangle_u = \int_{u_0}^{u_0+2\pi} \frac{du}{2\pi} A(\ell, u) \quad , \tag{5.8}$$

$$\langle A\rangle_t = \frac{1}{T[\ell(t)]} \int_t^{t+T[\ell(t)]} dt' A(\ell, u(t')) = \frac{1}{T[\ell(t)]} \int_{u(t)}^{u(t)+2\pi} du \frac{dt}{du} A(\ell, u) \, . \tag{5.9}$$

Even when in the first integral $u_0 = u(t)$, these two quantities will be different as long as dt/du depends on u.

Intuition suggests that the difference between the choices that one can take at the level of the setup might be akin to the difference between the various choices of renormalization scheme. While it seems reasonable that different averaging methods and choices can be followed consistently to describe the same dynamics, some of the choices that we have to make might depend dramatically on the nature of the system itself. An extreme example is the case of the so-called *crude averaging*, see e.g. [306], which shows that in some cases there can be problems in the convergence of the averaged solutions unless one chooses to perform the average after having rewritten the equations of motion in a certain canonical form.

In light of these possible complications, following closely [72] we set up our averaging procedure by means of the so called near-identity transformations for an angle-periodic system written in its canonical form. This method consists in performing a change of variables such that the system of equations becomes independent on the angle. We review this construction in Appendix B.3.

The choice of this method has a few relevant consequences. First, it implies that the averages (both at the level of the equations and at the level of the Lagrangian), must be taken with respect to the slowly evolving part of the mean anomaly. For practical reasons, these averages will then be expressed as integrals over the eccentric anomaly (or rather, the corresponding contact element). This makes clear that the possibly changing value of the time-period of the orbit does not lead to any corrections to the averages. Moreover, as we show in Appendix B.3, the construction of near identity transformations is such that the averages are performed, to any order, keeping the slow variables fixed in the integrand. This might seem counterintuitive, but stems from the fact that the averages appear as a by-product of a certain transformation of our variables, rather than as a direct coarse graining of the dynamics.

Having defined the averaging procedure, our practical task is to apply it to the relativistic, hierarchical three-body problem, up to quadrupolar order. In doing so, we only have difficulties due to the somewhat large number of variables (the contact elements) and to the presence of two small parameters that characterize the perturbation theory: the ratio of semimajor axes and the typical velocity of the bodies. On top of this, as we have already remarked, we will have to perform two averaging procedures.

5.2.2 *Center-of-mass and relative coordinates in boosted frame*

In order to carry out the multipole expansion, we express the two inner bodies' coordinates $\boldsymbol{y}_1$, $\boldsymbol{y}_2$ as functions of the position of the inner binary center of mass and of the relative distance between the two inner bodies, $\boldsymbol{Y}_{\rm CM}$ and $\boldsymbol{r}$. Then, introducing the contact elements, we can describe the inner binary as a spinning particle endowed with multipole moments, coupled to gravity.

When this is done to 1PN order, in general one needs to account for the difference between the rest frame of the three-body center of mass and the rest frame of the center of mass of the inner binary, as illustrated in Figure B.1. In fact, one might express the Lagrangian in terms of contact elements that are defined in either of the two reference frames, by means of Eq. (5.6). These two frames are connected by a boost plus a translation, a transformation that is non-trivial starting at 1PN order. This entails a difference between the two sets of contact elements that one can define. Crucially, while it is natural to express the effective action in the rest frame of the three-body center of mass, we find that the matching procedure is considerably simplified when we express the Lagrangian in terms of the *intrinsic* contact elements of the inner binary, that is, those defined in the rest frame of the center of mass of the inner binary. This is due to the fact that an appropriate reference frame is needed to disentangle the gravitational field from the multipole moments of an object, as discussed in [223]. For this reason, we will express the terms that appear in the effective Lagrangian, a functional evaluated in the global three-body rest frame R, in terms of the coordinates of the inner binary rest frame R′, so as to obtain a functional depending on the *intrinsic* contact elements of the inner binary. As studied in Chapter 4, terms in the effective Lagrangian of order up to $\varepsilon^{3/2}$ are not affected by this difference in reference frame. However, when computing quadrupole order contributions, we will see that it becomes important to account for such a difference.

While deferring the explicit computations in Appendix B.7, let us just show here the final relation between the absolute coordinates $(\boldsymbol{y}_1, \boldsymbol{y}_2)$ and the relative ones, $\boldsymbol{Y}_{\rm CM}, \boldsymbol{V}_{\rm CM}$ (center-of-mass position and velocity of the inner binary to 1PN order) and $\boldsymbol{r}', \boldsymbol{p}'$ (relative distance and momentum in the inner binary rest frame to 1PN order):

$$\begin{aligned}
\boldsymbol{y}_1 &= \mathbf{Y}_{\rm CM} + (X_2+\delta)\mathbf{r}' + X_2(\boldsymbol{V}_{\rm CM}\cdot\boldsymbol{r}')\left[(X_1-X_2)\frac{\boldsymbol{p}'}{\mu} - \frac{\boldsymbol{V}_{\rm CM}}{2}\right], \\
\boldsymbol{y}_2 &= \mathbf{Y}_{\rm CM} + (-X_1+\delta)\mathbf{r}' - X_1(\boldsymbol{V}_{\rm CM}\cdot\boldsymbol{r}')\left[(X_1-X_2)\frac{\boldsymbol{p}'}{\mu} - \frac{\boldsymbol{V}_{\rm CM}}{2}\right],
\end{aligned} \tag{5.10}$$

where δ is a 1PN quantity defined by

$$\delta = -\frac{1}{m}\mathbf{V}_{\rm CM}\cdot\mathbf{p}' + \nu(X_1-X_2)\left(\frac{p'^2}{2\mu} - \frac{G_N m}{2r'}\right). \tag{5.11}$$

In particular, note that this result implies the following relation between $\boldsymbol{r} = \boldsymbol{y}_1 - \boldsymbol{y}_2$ and $\boldsymbol{r}'$:

$$\boldsymbol{r} = \boldsymbol{r}' - (\boldsymbol{V}_{\rm CM}\cdot\boldsymbol{r}')\left[\frac{\boldsymbol{V}_{\rm CM}}{2} + (X_2-X_1)\frac{\boldsymbol{p}'}{\mu}\right]. \tag{5.12}$$

The final step is just to express $\boldsymbol{r}'$, $\boldsymbol{p}'$ in terms of the contact elements as in Eq. (5.6). To avoid clutter, in the rest of the article we will suppress the primed label on $\boldsymbol{r}'$, $\boldsymbol{p}'$.

5.2.3 *Averaging the Lagrangian*

In this Section, we start the computation of the quadrupolar Lagrangian by carrying out the averaging of the fast orbital modes of the inner binary. Specifically, we expand the Lagrangian (4.12) in multipoles around the center-of-mass of the inner binary, using the formulas given in Eq. (5.10). When carrying out computations up to order $v^2\varepsilon^{5/2}$, it is crucial to include corrections to the leading order averaging procedure besides the quadrupolar order of the multipole expansion. These are due to deviation from perfect adiabaticity, i.e. small changes of slowly evolving quantities over the course of the orbital period, and to backreaction of the quickly oscillating terms on the long time-scale dynamics. In order to simplify the presentation, we leave the detailed analysis of these corrections to App. B.5. Here we only remark that such corrections come in the form of the so-called cross terms: in this case either PN corrections to 0PN quadrupole terms or quadrupolar corrections to 1PN terms of the Lagrangian.

To give a separate treatment of these different contributions, we split the final averaged Lagrangian for the inner binary coupled to an external gravitational field, $\mathcal{L}^{\leq v^2\varepsilon^{5/2}}_{\rm quad,12}$, in two terms: the quadrupole term coming from the multipole expansion, computed to 1PN order, and the cross-terms induced by corrections to the leading order averaging procedure, computed in App. B.5:

$$\mathcal{L}^{\leq v^2\varepsilon^{5/2}}_{\rm quad,12} = \langle \mathcal{L}^{(0)}_{\rm quad}\rangle + \langle \mathcal{L}_S\rangle\,. \tag{5.13}$$

From this Lagrangian, integrating out the potential gravitons exchanged with the third body and averaging over the outer orbit, in the next sections we will obtain the long-timescale Lagrangian $\mathcal{L}^{\leq v^2\varepsilon^{5/2}}_{\rm quad}$.

Starting with the quadrupole term of the multipole expansion, to order $v^2\varepsilon^{5/2}$ included, we obtain the expression

$$\begin{aligned}\mathcal{L}^{(0)}_{\rm quad} = -\frac{1}{2}\mu r^i r^j \partial_i\partial_j\tilde{\phi}\Big[1 - \frac{G_N m}{r}(1-2\nu) + (X_1 - X_2)\boldsymbol{V}_{\rm CM}\cdot\boldsymbol{r}\big(r^i v^j + r^j v^i\big) \\ + \frac{3}{2}\big(2(X_2-X_1)\mathbf{V}_{\rm CM}\cdot\mathbf{v} + (1-3\nu)v^2\big)\Big] \\ + \frac{1}{2}\mu r^i r^j \partial_i\partial_j\tilde{A}_k\big[V^k_{\rm CM} + (X_2-X_1)v^k\big]\,,\end{aligned} \tag{5.14}$$

where all the relevant quantities used in this equation have been defined in Section 5.2.2 and $\boldsymbol{r}$, $\boldsymbol{p}$ are the relative position and momentum vectors in the instantaneous rest frame of the inner binary center of mass.[1] We have made several simplifications in order to get to Eq. (5.14). First, we have dropped terms nonlinear in $\tilde{\phi}$ as well as $V^2_{\rm CM}$ corrections as they would be of order $v^2\varepsilon^3$ according to the power-counting rules of our theory, once the external gravitational field is integrated out. These terms would be relevant when considering the octupolar order in the center-of-mass expansion, which is beyond the scope of our present analysis. Second, although at this point the coupling $\partial_i\partial_j\tilde{A}_k V^k_{\rm CM}$ is of $\varepsilon^{5/2}v$ order, it cannot contribute to the final $\varepsilon^{5/2}v^2$ Lagrangian when $\tilde{A}_k$ is been integrated out. This is because the lowest-order coupling of $\tilde{A}_k$ to m_3 involves also v_3, so that it brings another $\varepsilon^{1/2}$ factor once we integrate out $\tilde{A}_k$, due to the scaling $v_3 \sim \varepsilon^{1/2}v$. Therefore we can ignore the term $\partial_i\partial_j\tilde{A}_k V^k_{\rm CM}$ in the present analysis. Finally, note that the use of intrinsic relative position vectors, defined in the inner binary instantaneous rest frame as explained in Section 5.2.2, manifests itself in the last factor that multiplies $\partial_i\partial_i\tilde{\phi}$.

Following the method of near-identity transformations, presented in Appendix B.3, we can implement the averaging procedure and eliminate the short time-scale dynamics from the Lagrangian. We leave to Appendix B.5 the corrections due to breaking of adiabaticity and backreaction of short modes, $\langle\mathcal{L}_S\rangle$ in Eq. (5.13), while here we average Eq. 5.14 over one orbit of the inner binary, using the definition of contact elements given in Eq. (5.5). We obtain:

$$\begin{aligned}\langle\mathcal{L}^{(0)}_{\rm quad}\rangle = -\frac{\mu a^2}{4}\Big[(1+4e^2)\alpha^i\alpha^j + (1-e^2)\beta^i\beta^j \\ + \frac{G_N m}{2a}\big((1-5\nu+4e^2(2\nu-1))\alpha^i\alpha^j + (1-5\nu)(1-e^2)\beta^i\beta^j\big)\Big]\partial_i\partial_j\tilde{\phi} \\ - \frac{1}{2}(X_2-X_1)ae\,\epsilon_{lik}J^l\alpha^j\big(\partial_i\partial_j\tilde{A}_k - 4V^k_{\rm CM}\partial_i\partial_j\tilde{\phi}\big)\,,\end{aligned} \tag{5.15}$$

where we recall that $\boldsymbol{J}$ is the angular momentum of the binary defined in Eq. (5.2).

At this point we can compute the whole contribution $\mathcal{L}^{\le v^2\varepsilon^{5/2}}_{\rm quad,12}$ by adding the result from our analysis in Appendix B.5, i.e. Eq. (B.67), to the contribution of Eq. (5.15):

$$\mathcal{L}^{\le v^2\varepsilon^{5/2}}_{\rm quad,12} = -\frac{1}{2}Q^{ij}_E\partial_i\partial_j\tilde{\phi} - \frac{1}{2}(X_2-X_1)ae\,\epsilon_{lik}J^l\alpha^j\big(\partial_i\partial_j\tilde{A}_k - 4V^k_{\rm CM}\partial_i\partial_j\tilde{\phi}\big)\,. \tag{5.16}$$

1 In this expression, we can substitute $\boldsymbol{v} = \boldsymbol{p}/\mu$.

Here the traceless "electric-type" quadrupole moment is given to 1PN order by

$$Q_E^{ij} = \frac{\mu a^2}{2}\left(f_\alpha(e)\alpha^i\alpha^j + f_\beta(e)\beta^i\beta^j - \frac{f_\alpha(e)+f_\beta(e)}{3}\delta^{ij}\right), \tag{5.17}$$

and the two functions of the eccentricity read

$$f_\alpha(e) = 1 + 4e^2 - \frac{G_N m}{2a\left(1-e^2+\sqrt{1-e^2}\right)}\Big[17 + 13\sqrt{1-e^2} + 5\nu\left(1+\sqrt{1-e^2}\right) + e^2\left(56 + 15\sqrt{1-e^2} - \nu(13 + 8\sqrt{1-e^2})\right) + 4e^4(3+2\nu)\Big], \tag{5.18}$$

$$f_\beta(e) = 1 - e^2 - \frac{G_N m}{2a\left(1-e^2+\sqrt{1-e^2}\right)}\Big[13 + 17\sqrt{1-e^2} + 5\nu\left(1+\sqrt{1-e^2}\right) + e^2\left(31 + 18\sqrt{1-e^2} - 5\nu(2+\sqrt{1-e^2})\right) + e^4(5\nu - 9)\Big]. \tag{5.19}$$

The first term of the Lagrangian (5.16) contains both Newtonian ε^2 and PN $v^2\varepsilon^2$ scalings, while its second term proportional to the difference of masses contains only the PN $v^2\varepsilon^{5/2}$ scaling. Note that, in order to remove the trace from the quadrupole moment, we have made use of the equation of motion $\partial_i\partial^i\tilde{\phi} = 0$.[2] Note also that the osculating elements used in these equations are the *intrinsic contact* elements, defined in the rest frame of center of mass of the inner binary, and not Newtonian orbital elements. The explicit difference between these two sets of osculating elements is detailed in Appendix B.9, see Eq. (B.116). As explained there, this difference is small in the sense that at any time, the contact elements differ from the osculating elements by a 1PN quantity. Despite the small difference, Eq. (B.116) implies that while the contact element a is constant by virtue of our averaging procedure, see Appendix B.4, the respective orbital element features post-Newtonian variations. This confirms the findings of [268, 281], who showed for the first time that cross-terms induce variations in the (orbital) semi-major axis a. However, our Lagrangian formalism allows us to assert that these variations always stay small over time and cannot accumulate over a long timescale to induce large variations in a, while this was left as an open question in previous works [268, 281].

5.2.4 *Matching*

We will now express the quadrupolar coupling (5.16) in a gauge-invariant way. As discussed in Sec. 4.2.4, the quadrupolar part of the effective action can be written in terms of two interactions, the electric-type and magnetic-type quadrupole terms:

$$\mathcal{S}_{\text{quad}} = -\frac{1}{2}\int \mathrm{d}\tau E_{ij}Q_E^{ij} + \frac{2}{3}\int \mathrm{d}\tau B_{ij}Q_B^{ij}, \tag{5.20}$$

with the definitions of Sec. 4.2.4.

Being only interested in terms of order $v^2\varepsilon^{5/2}$, we can expand $E_{\mu\nu}$ and $B_{\mu\nu}$ to linear order in the gravitational fields $\tilde{\phi}$ and $\tilde{A}_i$. Furthermore, at this order we can also ignore the interactions involving $\dot{\tilde{A}}_i$ as well as both V_{CM}

2 Even though we will eventually be interested in off-shell potential gravitons exchanged between the inner binary and the third body, using this on-shell condition does not alter the effective action, similarly to what is discussed in [307], since it amounts to neglecting a contact term.

and $\tilde{A}_i$ together. Finally, we will also make use of the equations of motion for the external fields which read $\partial_i\partial^i\tilde{\phi} = \partial_i\partial^i\tilde{A}_j = 0$,[3] and of the gauge condition on $\tilde{A}^i$ which is $\partial_i\tilde{A}^i = -4\dot{\tilde{\phi}}$ [292, 293]. Thus, the tensors are given by

$$E_{ij} = \partial_i\partial_j\tilde{\phi}\,, \tag{5.21}$$

$$B_{ij} = \frac{1}{2}\epsilon_{mn(i}\partial_{j)}\left[\partial^m\tilde{A}^n + 4V^m_{\rm CM}\partial^n\tilde{\phi}\right]\,. \tag{5.22}$$

The resulting magnetic-type quadrupole reproduces the coupling to $\tilde{A}$ and $\tilde{\phi}$ proportional to the difference of masses $X_1 - X_2$ that we found previously in Eq. (5.16) [4]:

$$Q^{ij}_B = \frac{3}{2}(X_1 - X_2)aeJ^{(i}\alpha^{j)}\,, \tag{5.23}$$

This is consistent with the standard definition of the magnetic-type quadrupole

$$Q^{ij}_B = \epsilon^{mn(i}\sum_A m_A x^{j)}_A x^m_A v^n_A\,, \tag{5.24}$$

see e.g. [213], once averaged over one orbit of the inner binary system, thus giving a strong check of the validity of our decomposition. Note that the use of the relative variables defined in the rest-frame of the inner binary, as explained in Section 5.2.2, is crucial to obtain the correct magnetic-type interaction. Had we used relative variables naively defined in the three-body rest frame (as it is the case in other studies [268, 281–283]), we would not have achieved this separation between a quadrupole moment intrinsic to the inner binary coupled to the external gravitational field. Instead, we would have obtained a quadrupole moment which depends on the center-of-mass velocity of the inner binary. This undesirable feature has been avoided thanks to our choice of variables.

To recap our results, we obtain the following action for the binary system treated as a point-particle up to quadrupole order:

$$S = \int \mathrm{d}\tilde{\tau}\left[-\mathcal{E}_{\rm 1PN} + \frac{1}{2}J_{\mu\nu}\Omega^{\mu\nu} - \frac{1}{2}E_{ij}Q^{ij}_E + \frac{2}{3}B_{ij}Q^{ij}_B\right]\,, \tag{5.25}$$

where Q^{ij}_B and Q^{ij}_E have been defined in Eqs. (5.23) and (5.17), $d\tilde{\tau} = dt\sqrt{-\tilde{g}_{\mu\nu}V^\mu_{\rm CM}V^\nu_{\rm CM}}$ and the first two terms represent the Lagrangian of the inner binary to dipole order, with $\mathcal{E}_{\rm 1PN}$ its energy to 1PN order, $J_{\mu\nu}$ its spin tensor and $\Omega^{\mu\nu}$ its angular velocity rotation vector. We can now move on to integrate out the external fields $\tilde{\phi}$, $\tilde{A}_i$ in the presence of a third point-particle m_3.

5.3 DOUBLE-AVERAGED LAGRANGIAN UP TO ORDER $v^2\varepsilon^{5/2}$

In the previous sections we have derived the Lagrangian $\mathcal{L}^{\leq v^2\varepsilon^{5/2}}_{\rm quad,12}$ describing the dynamics of the inner binary coupled to an external gravitational field

3 As mentioned in Footnote 2, even if we will eventually be interested in off-shell gravitons we still can use the equations of motion, which amounts to neglecting a contact term.

4 In order to get the coupling in Eq. (5.16) from the two equations (5.22) and (5.23), one has to make use of the gauge condition on $\tilde{A}_i$ and $\tilde{\phi}$ and the fact that

$$\frac{\mathrm{d}\tilde{\phi}}{\mathrm{d}t} = \dot{\tilde{\phi}} + V^i_{\rm CM}\partial_i\tilde{\phi}$$

so that this total derivative can be ignored from the vertex.

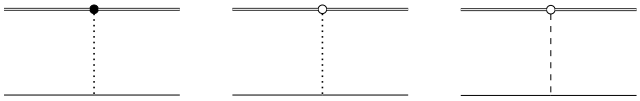

Figure 5.1: Feynman diagrams corresponding to integrating out the outer fields $\tilde{\phi}$, $\tilde{\boldsymbol{A}}$ and contributing to the quadrupolar Lagrangian (5.27) up to order $v^2\varepsilon^{5/2}$. The double line represents the inner binary system, while the single line stands for the third particle m_3. The black [white] dot represent the insertion of an "electric-type" ["magnetic-type"] quadrupolar coupling from Eq. (5.25). The dotted line represents the exchange of a scalar $\tilde{\phi}$, while the dashed line stands for the exchange of a vector $\tilde{\mathbf{A}}$.

Operator	Rule
$-\frac{1}{2}Q_E^{ij}\partial_i\partial^j\tilde{\phi}$	$J_3^{1/2}v^2\varepsilon^2$
$\frac{1}{3}Q_B^{ij}\epsilon_{mn(i}\partial_{j)}\partial^m\tilde{A}^n$	$J_3^{1/2}v\varepsilon^2$
$\frac{4}{3}Q_B^{ij}\epsilon_{mn(i}\partial_{j)}V_{\rm CM}^m\partial^n\tilde{\phi}$	$J_3^{1/2}v^2\varepsilon^{5/2}$

Table 5.1: Power-counting rules for the quadrupolar vertices contained in the effective point-particle action (5.25) up to 1PN order, with $J_3 = (G_N M^3 a_3)^{1/2}$, $v^2 = G_N m/a$ and $\varepsilon = a/a_3$. The rules are obtained using the scaling presented in Section 4.3. For convenience, the integral over time is not displayed, although it should be included to obtain a dimensionless rule. Furthermore, in the main text we will ignore the J_3 factors when discussing the scaling of an operator, since in the end the Lagrangian will always be proportional to J_3 (terms not proportional to the angular momentum represent true quantum loops whose contribution to the dynamics is completely negligible in the NRGR formalism [70])

over time-scales longer than the inner orbital period, up to order $v^2\varepsilon^{5/2}$. We now take into account the presence of the third body. This implies that we need to integrate out potential gravitons exchanged between the inner binary and the third body, and in turn adds a new time-scale to the problem: the orbital period of the outer orbit. Being interested in the evolution on much longer time-scales, we will repeat the averaging procedure carried out for the inner binary, starting from quantities that are already averaged with respect to the period of the inner binary.

The action of the third point-particle is given by:

$$\mathcal{L}_3 = -m_3\sqrt{-\tilde{g}_{\mu\nu}v_3^\mu v_3^\nu}\,, \tag{5.26}$$

and we integrate out the gravitational field $\tilde{g}_{\mu\nu}$. Since the corresponding Lagrangian up to dipole order $v^2\varepsilon^{3/2}$ has already been described in the previous chapter, we will concentrate on the quadrupolar contributions[5]. From Eq. (5.25), we see that three new vertex appear at quadrupolar order, whose power-counting rules are summarized in Table 5.1. Integrating out the external fields $\tilde{\phi}$, $\tilde{\boldsymbol{A}}$ as shown in the Feynman diagrams of Figure 5.1, we find

$$\mathcal{L}_{\rm quad}^{\leq v^2\varepsilon^{5/2}} = \frac{G_N m_3}{2R^3}\left(3N^iN^j-\delta^{ij}\right)\left[Q_{E,ij}+4(X_2-X_1)ae\epsilon_{lki}J^lV^k\alpha_j\right]\,, \tag{5.27}$$

where we have moved to the center-of-mass frame of the triple system by setting $\boldsymbol{V}_{\rm CM} = X_3\boldsymbol{V}$, $\boldsymbol{v}_3 = -X_{\rm CM}\boldsymbol{V}$, and we recall that Q_E^{ij} has been defined in Eq. (5.17). This Lagrangian contains terms with three different scalings: Newtonian quadrupolar (ε^2) in the Newtonian part of Q_E^{ij}, 1PN quadrupolar ($v^2\varepsilon^2$) in the 1PN part of Q_E^{ij}, and magnetic-type quadrupolar ($v^2\varepsilon^{5/2}$) in the second term proportional to the difference of masses.

We can now implement the second step of averaging and eliminate the dynamics on time-scales shorter than the period of the outer orbit by means of new near-identity transformations. We thus obtain the doubly averaged Lagrangian:

$$\begin{aligned}\left\langle\mathcal{L}_{\rm quad}^{\leq v^2\varepsilon^{5/2}}\right\rangle &= \frac{3G_Nm_3}{4a_3^3(1-e_3^2)^{3/2}}Q_E^{ij}\left(\alpha_3^i\alpha_3^j+\beta_3^i\beta_3^j\right) \\ &+\frac{3(X_2-X_1)}{X_{\rm CM}}\Omega_{\rm prec}\frac{a}{a_3}\frac{ee_3}{1-e_3^2}\Big((\boldsymbol{J}\times\boldsymbol{\alpha})\cdot(\boldsymbol{J}_3\times\boldsymbol{\alpha}_3)-2(\boldsymbol{J}\cdot\boldsymbol{J}_3)(\boldsymbol{\alpha}\cdot\boldsymbol{\alpha}_3)\Big)\,,\end{aligned} \tag{5.28}$$

where $\Omega_{\rm prec}$ is the precession frequency already defined in Eq. (4.54). This completes our derivation of the double-averaged Lagrangian up to order $v^2\varepsilon^{5/2}$. Note that in doing this last average, we did not need to compute corrections due to deviation from adiabaticity and backreaction of short modes, since these effects, similarly to what discussed in Appendix B.5, would contribute only to quadrupole times v_3^2 order which means starting from octupole order, since $v_3 \sim e^{1/2}\,v$.

Starting from this result, the LPE for the long time-scale evolution of both the inner and outer orbits are obtained by taking the relevant derivatives of the Lagrangian as shown in Eqs. (B.4)-(B.9). Note that the angular dependence of the quadrupolar $v^2\varepsilon^2$ terms is quite similar to the Newtonian quadrupole at ε^2, so that we expect that these $v^2\varepsilon^2$ terms will not induce a

5 Being interested in the dynamics of contact elements, we do not keep track of the possible coupling of the three body system itself to an external gravitational field.

qualitatively different behavior in the long-term evolution of the system. On the other hand, the angular structure of the $v^2\varepsilon^{5/2}$ terms is more involved and somewhat similar to the Newtonian octupole, which means that similarly to the octupole such terms can give rise to new behaviors at long times (see e.g. [227, 262] concerning the influence of octupole terms in the Kozai-Lidov problem). We will describe these new behaviors in the next Section, where we will numerically solve the LPE for the long time-scale dynamics obtained from the quadrupolar Lagrangian up to order $v^2\varepsilon^{5/2}$.

Before moving on, let us compare our result to the ones already present in the literature. As we stated in the introduction, we are aware of only four works which tackled the task of computing 1PN quadrupolar terms including the effects due to deviations from the adiabatic approximation and to backreaction of quickly oscillating modes. In three of them, a circular outer orbit is assumed [268, 281, 282]. As shown by Eq. (5.28), this causes to neglect the presence of magnetic-type quadrupolar terms of order $v^2\varepsilon^{5/2}$, which can lead to new interesting behaviors in the long-term evolution of the system as we show in Section 5.4. The explicit comparison between the LPE obtained in these works and our result is complicated by the fact that we do not use the exact same averaging procedure, see the comments at the end of Appendix B.9. On the other hand, the work in [283] reports a puzzling result. Indeed, it describes the effect of a so-called libration cross-term in the equations of motion which, once translated to the Lagrangian point of view, would scale as $v^2\varepsilon^{1/2}$ within our power-counting rules. Such a term is absent from our derivation and we believe that it should not be present [6].

5.4 NUMERICAL SOLUTION TO THE LPE

In this Section, we will numerically integrate the equations of motion stemming from the averaged Lagrangian (5.1) for both inner and outer orbits given some initial conditions. A systematic exploration of the (huge) parameter space, as has been done in e.g. [249] with lower-order perturbations in the Lagrangian, is beyond the scope of our work. Instead, we content ourselves with showing that the quadrupolar terms derived in this chapter can have some non-trivial consequences on the long-term evolution of relativistic three-body hierarchical systems.

Varying the total averaged Lagrangian (5.1) over planetary elements as described in Appendix B.1, we obtain the so-called Lagrange Planetary

6 Note that before averaging over the inner orbit, the Lagrangian indeed contains terms scaling as $v^2\varepsilon^{1/2}$ which are also dependent on the center-of-mass definition. However, one can check that even with the Newtonian definition of the center-of-mass used in [283], the average of these $v^2\varepsilon^{1/2}$ terms over the inner orbit still vanishes.

Equations (LPE) which dictate the evolution of orbital elements over long timescales. For the inner orbit, they are given by

$$\dot{a} = 0 \,, \tag{5.29}$$

$$\dot{e} = -\sqrt{\frac{1-e^2}{G_N m a e^2}}\frac{\partial \mathcal{R}}{\partial \omega} \,, \tag{5.30}$$

$$\dot{\imath} = -\frac{1}{\sqrt{G_N m a (1-e^2)}\sin\imath}\frac{\partial \mathcal{R}}{\partial \Omega} + \frac{\cos\imath}{\sqrt{G_N m a (1-e^2)}\sin\imath}\frac{\partial \mathcal{R}}{\partial \omega} \,, \tag{5.31}$$

$$\dot{\omega} = \sqrt{\frac{1-e^2}{G_N m a e^2}}\frac{\partial \mathcal{R}}{\partial e} - \frac{\cos\imath}{\sqrt{G_N m a (1-e^2)}\sin\imath}\frac{\partial \mathcal{R}}{\partial \imath} \,, \tag{5.32}$$

$$\dot{\Omega} = \frac{1}{\sqrt{G_N m a (1-e^2)}\sin\imath}\frac{\partial \mathcal{R}}{\partial \imath} \,, \tag{5.33}$$

where $\mathcal{R} = \mathcal{L}_1/\mu$ is the so-called *perturbing function*, with $\mathcal{L}_1$ containing all terms beyond the first two kinetic parts in the total Lagrangian (5.1). The LPE for the outer orbit are obtained by replacing all inner quantities with outer ones. However, one has to be careful to replace the mass m with the sum $m_3 + \mathcal{E}$, with $\mathcal{E} = m - G_N m/2a$, as was emphasized below Eq. (5.1).

Until now, we have not specified any orientation for our reference frame centered on the total center-of-mass of the three-body system. A straightforward application of Noether's theorem to the averaged Lagrangian (5.1) gives that the total momentum, $\boldsymbol{J} + \boldsymbol{J}_3$, is conserved. We can thus follow the conventional elimination of the nodes procedure, see e.g. [227], to choose the orientation of axis with the z axis parallel to $\boldsymbol{J} + \boldsymbol{J}_3$, in which the following relations between planetary elements hold true:

$$\Omega_3 = \Omega + \pi \,, \quad \|\boldsymbol{J}\|\sin\imath = \|\boldsymbol{J}_3\|\sin\imath_3 \,. \tag{5.34}$$

This allows us to eliminate two variables from the eight dynamical variables that we are solving for, by e.g. expressing all quantities only in terms of Ω and $\imath_{\rm tot} = \imath + \imath_3$. As a relevant remark, let us highlight two subtleties concerning the elimination of nodes. First, as discussed in [262], one can not directly use the relations (5.34) at the level of the Lagrangian or the Hamiltonian, because this would mean using a consequence of the equations of motion (conservation of angular momentum) in the Lagrangian itself, which can lead to wrong results. Indeed, many studies (including the original one of Kozai [308]) concluded that the z-components of both angular momentum are conserved *independently*, which is incorrect as shown in [262]. Second, note that the conserved angular momentum is defined with the contact elements, which differ from the osculating elements at 1PN order as we highlight in Appendix B.9. In other words, the angular momentum defined with osculating elements does feature variations at 1PN order, while the one defined with contact elements is a constant. Thus, one cannot eliminate the nodes by following the conventional procedure if one uses the osculating elements instead of the contact ones.[7] As an independent check of this procedure, our numerical simulation confirms that the projection of the angular momentum on the z-axis, $\|\boldsymbol{J}\|\cos\imath + \|\boldsymbol{J}_3\|\cos\imath_3$, is conserved through the evolution.

We now discuss our numerical solution for a particular set of parameters. We choose the inner binary to be composed of two black holes with total

7 The fact that $\dot{\Omega} \neq \dot{\Omega}_3$ for osculating elements was already noticed in Ref. [283]. In comparison, our discussion adds that the elimination of nodes can be carried out consistently at the level of contact elements.

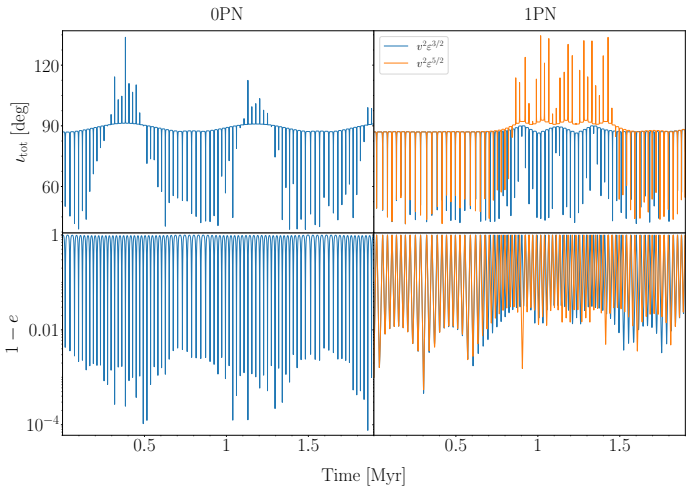

Figure 5.2: *Impact of the quadrupole-1PN terms on the evolution of a three-body system.* We solve the LPE (5.29) with the following parameters for a three-body system: $m = 50M_\odot$, $\nu = 0.15$, $a = 5$AU, $m_3 = 50M_\odot$, $a_3 = 350$AU. The initial conditions are $e = 0.001$, $e_3 = 0.7$, $\iota_{\rm tot} = 87°$, $\omega = 240°$, $\omega_3 = 0°$, $\Omega = 0°$. In the left plots we show the evolution of the total inclination $\iota_{\rm tot} = \iota + \iota_3$ and the eccentricity e without taking into account PN perturbations (i.e., setting $v^2 = 0$ in our power-counting rules), but including Newtonian quadrupolar and octupolar terms as described in e.g. [227]. The system features orbital flips and eccentricities as high as $1 - e \sim 10^{-4}$, which are typical of the octupolar Kozai-Lidov mechanism [227, 231, 240, 262]. In the right panels, we show the evolution of the same quantities taking into account higher terms in the PN expansion on top of the Newtonian quadrupolar and octupolar terms. Namely, we include PN terms up to $v^2\varepsilon^{3/2}$ order (in blue) and up to $v^2\varepsilon^{5/2}$ order (in orange). The lowest-order PN perturbations (which are well-known and have been studied in the context of three-body dynamics e.g. in [249, 263, 267, 279, 280, 309]) generically quench the orbital flips and eccentricities. However, the new $v^2\varepsilon^{5/2}$ terms computed in this article re-trigger the orbital flips.

mass $m = 20M_\odot$ and mass ratio $\nu = 0.15$, while the third body has a mass $m_3 = 50M_\odot$. The inner semimajor axis is $a = 5$AU, and the outer one is $a_3 = 350$AU. Such values are typical for black holes in dense nuclear clusters [231, 239, 310, 311]. The initial conditions of the system are described in Figure 5.2, where we plot the total inclination $\iota_{\rm tot} = \iota + \iota_3$ and the eccentricity e as functions of time. Without PN perturbations, the system undergoes flips of inclination and extreme eccentricities due to the octupole effects beyond the Kozai-Lidov mechanism. The presence of lower-order PN terms (up to $v^2\varepsilon^{3/2}$ i.e. dipole order) quenches the maximal eccentricity as well as the orbital flips. This behavior is well-known in the literature [249, 263, 309]. However, we find that adding the new 1PN quadrupolar terms that we computed in this article can re-trigger the flips in inclination. This is shown in Figure 5.2. Whether or not this ultimately influences the mechanisms leading to binary mergers in nuclear clusters is left as an interesting question for future work.

As a final comment, note that we have not studied the linear stability of the system, which would require diagonalizing the Hessian matrix of the inner and outer contact elements. Even without knowing the eigenvalues of the Hessian matrix, we can appreciate how no unstable direction was hit in our simulation, since amplitude and phases of the oscillations do not grow exponentially. Moreover, we take the persistence of quasi-periodic oscillations as a proxy for the absence of any resonant behavior (see discussion in Appendix B.4). Therefore, the accuracy of the simulation reported is only limited by the growth of higher order terms that we have neglected. These will possibly become of order one after a time which is parametrically larger than the interval explored in the simulation.

5.5 SUMMARY

In this chapter we have considered the effects of a distant third body orbiting a tightly bound binary system, forming a hierarchical triple. We have derived an effective action that includes relativistic effects up to 1PN order as well as multipole effects up to quadrupole order.

Our EFT approach allows to make use of symmetries to constrain the form of the interactions and makes clear contact with the underlying field theoretic description of gravity. In particular, starting from an action of three worldlines minimally coupled to gravity, we integrated out the off-shell gravitons that give rise to the binding potential and matched the theory to a system of interacting composite particles. We did this by first integrating potential gravitons binding the inner binary and matching the action of the inner binary to that of a composite particle coupled to an external gravitational field. Then, integrating out potential gravitons with wavelengths comparable to the size of the outer orbit, we obtained an action describing two interacting particles, the inner binary and the third body. This method allows to build towards a systematic understanding of the long time-scale dynamics away from resonances.

In practice, deriving the effective action presents a few challenges. For instance, it is important to understand which are the most suitable variables to describe the long time-scale dynamics and the relativistic corrections. In this study we have found that the contact elements defined in the center of mass frame of the inner binary allow to encode in a compact way various PN corrections. Moreover their slowly evolving parts are the quantities that carry only the relevant information to describe the system on long time-

scales. Using near-identity transformations, we were able to unambiguously integrate out the effect of fast orbital modes on the dynamics, outlining a procedure that can be used at any order in perturbation theory. With this approach, we managed to gain insight on the interplay between multipole expansion and relativistic effects. Our main result is that quadrupole-PN cross terms can retrigger orbital flips in spite of PN effects that appear at lower order in the multipole expansion.

Part III

CONCLUSION

6

CONCLUSIONS

In this thesis we have employed EFTs to tackle relevant questions for black hole physics and gravitational wave detection, illustrating both the predictive and descriptive power of the EFT framework.
In the first part of the thesis we considered models of physics beyond GR which have a low energy degree of freedom coupled non minimally to gravity, in the specific a shift symmetric (pseudo)scalar field.
Our considerations led us to understand that non-minimal coupling to invariants quadratic in the Riemann tensor is in tension with the expectations regarding the causal structure of a relativistic EFT. A UV completion can in principle remove this tension by canceling superluminal propagation. This solution still makes difficult to envision these interaction taking place with a strength such to affect astrophysical black holes. Indeed invoking the UV completion would imply the existence of a tower of spinning particles as light as $\mathrm{km}^{-1} \sim 10^{-10}\mathrm{eV}$, impacting severely the Newtonian potential measured at scales of few μm in table top experiments.

In this regard, it might be interesting to understand whether such a UV completion can exists that would not radically alter the Newtonian potential at shorter length-scales.
On a more fundamental level, as far as the first part of this work goes, we can outline a few points that deserve further study. One is the question of whether large non-minimal coupling to gravity leads in general to tensions with causality. This for instance seems to be the case for the operator ϕ^2-GB, but it would be useful to gather other examples also for different fields coupled to gravity. Another is the relation between superluminality bounds and results from dispersion relations, which is still not fully understood. In this perspective, studying constraints on higher point operators from both points of view seems an important and ambitious goal. In the same direction, it would also be useful to implement an optimization procedure for superluminality bounds, e.g. building on the attempts of [126]. Finally, we have omitted any discussion of the possible implications of superluminality. Exploring this topic further seems to be a fundamental direction of research which might lead to new insights about the interplay between quantum fields and Cauchy horizons. Moreover, from a more mathematical point of view, it would be interesting to understand if in the presence of gravity the finite-frequency phase velocity can be shown to be larger than the front velocity [312].

In the second part of this thesis, we have focused on studying environmental perturbations of a binary, in the specific case in which it is perturbed by a distant third body. We have developed an EFT to describe the long time-scale dynamics of such systems, by mapping the binary to

a composite particle endowed with multipole moments. We derived the effective Lagrangian for this system including the first order relativistic corrections as well as three-body effects up to the quadrupole order. The EFT approach made transparent how different terms combine together due to symmetries in the IR description. With our analysis, we were able to investigate regimes in which the relativistic-quadrupole cross terms can lead to orbital flips. However, this analysis should only be considered a starting point.

In fact, we can outline a few points that would deserve further study. First of all, despite carrying out non-trivial computations that probed to some depth the perturbative structure of the theory, we are still unable to provide a complete understanding of the NDA power counting scheme for the long time-scale effects. In particular, although we can determine by symmetry which operators will describe the long time-scale dynamics of the hierarchical triple, we have not yet outlined a way to predict the relations between their coefficients. Determining these relations by means of NDA would allow to explore very efficiently the interplay between different interactions. A similar analysis might also make simpler to detect the resonant behaviors that can be triggered by the various interactions. These usually become manifest e.g. through small denominators (the so-called small divisors [72]) in the Fourier transform of quantities computed perturbatively. Also in this case, NDA should make possible to sidestep part of the computations and predict by power counting the rate at which a resonance is hit, depending on the coefficients of the operators in the EFT.
Since the coefficients of the long time-scale operators are specified by the orbital parameters and masses of the three-body system, answering these questions would allow to efficiently study the features of the system throughout its phase-space. Besides studying the operator mixing and the meta-stability of the hierarchical configuration, it might be possible to further leverage insights from particle physics and compute e.g. the spectrum of eccentricity-inclination oscillations or the rate of orbital flips across the three-body phase-space.

From the phenomenological point of view, it would be interesting to quantify how these effects alter existing priors on the population of binaries that will be observed through gravitational wave experiments. Even more, our work could be used to better understand which three-body configurations can lead to detectable modifications of gravitational waveforms due to the long time-scale dynamics, and to provide a waveform template incorporating three-body effects. For instance, it has recently been proposed that Kozai-Lidov oscillations may be observable in waveforms [313, 314]; taking into account the relativistic three-body effects which we computed in this work may be crucial for the parameter estimation of these kind of systems.
Moving beyond the case of a distant third body orbiting an ideal binary, it would be relevant to extend the perturbed-NRGR approach we presented to describe more generic environmental effects. For instance, one could consider the NRGR description of a binary and add a generic stress energy tensor describing e.g. a cloud or halo surrounding the binary as well as distant sources. Continuing in this direction, the ultimate goal would be to piece together the puzzle of different possible effects (clouds, distant

sources, deformability of the bodies, etc.) and produce a clear picture of the features that could be displayed in GR binary mergers.

Remarkably, studying how finite size and environmental effects impact binary mergers might one day allow to shed light, or better gravity, on astrophysical phenomena in extreme conditions.

Part IV

APPENDICES

A

RESULTS FOR BLACK HOLE HAIR

A.1 THE MANY GAUSS-BONNET CURRENTS

As we have discussed in Section 2, the Gauss-Bonnet invariant is the divergence of a current which is not itself a tensorial object. For this reason taking its square does not give a quantity which is invariant under diffeomorphisms. In the main text we discussed the special expression that this current takes in the presence of a Killing vector aligned with a coordinate. In a generic case the current can be written is terms of the spin connection. Its expression does not give a covariant vector, in the same way the Christoffel symbols are not rank-3 tensors.

Using Greek and Latin letters for curved and flat indexes respectively, the vierbeins $e^a_\mu(x)$ will be defined through the following relation: $g_{\mu\nu} = e^a_\mu \eta_{ab} e^b_\nu$. The spin connection will be $\omega^a{}_b = \omega^a_{\mu b} dx^\mu = e^a_\nu \nabla_\mu e^\nu_b dx^\mu$ and the curvature form will be $R^{ab} = d\omega^{ab} + \omega^a{}_c \wedge \omega^{cb} = e^a_\mu e^b_\nu R^{\mu\nu}{}_{\rho\sigma} dx^\rho \wedge dx^\sigma$, where as usual flat indexes are lifted and lowered with the flat metric η_{ab}. Using these definitions we can express the Gauss-Bonnet invariant as a total derivative:

$$\begin{aligned}\int d^4x\sqrt{-g}\mathcal{R}^2_{GB} &= \int R^{ab}R^{cd}\epsilon_{abcd} = \int d\left(\epsilon_{abcd}\omega^{ab}\left(R^{cd} - \frac{1}{3}\omega^c{}_e\omega^{ed}\right)\right) \\ &= -\int d^4x\sqrt{-g}\nabla_\mu\left(\epsilon^{\mu\nu\rho\sigma}\epsilon_{\alpha\tau}{}^{\beta\lambda}\omega^\alpha_{\nu\beta}\left(\frac{1}{2}R^\tau{}_{\lambda\rho\sigma} - \frac{1}{3}\omega^\tau_{\rho\gamma}\omega^\gamma_{\sigma\lambda}\right)\right) ,\end{aligned} \tag{A.1}$$

where $\omega^\alpha_{\nu\beta} = \omega^a_{\nu b} e^\alpha_a e^b_\beta$ and $\epsilon_{\mu\nu\rho\sigma} = e^a_\mu e^b_\nu e^c_\rho e^d_\sigma \epsilon_{abcd}$, with ϵ_{abcd} the Levi-Civita symbol[1].

Evaluating the expression (A.1) in the Schwarzschild coordinates (with the natural induced vierbein) gives:

$$J^\mu_{(Schw)} = \left(0, \frac{2r_s(r-2r_s)}{r^5}, -\frac{4r_s\cot(\theta)}{r^5}, 0\right) , \qquad \nabla_\mu J^\mu_{(Schw)} = \frac{12r_s^2}{r^6} . \tag{A.2}$$

This current has a non-zero θ component and its square diverges both at the horizon and at the poles:

$$J^2_{(Schw)} = \frac{4r_s^2}{r^9}\left(4r\cot^2(\theta) - \frac{(r-2r_s)^2}{r_s - r}\right) . \tag{A.3}$$

This current is actually a linear combination of the currents defined in Section 2.1.1: $J^\mu_{(Schw)} = \frac{1}{2}J^\mu_{(t)} + \frac{1}{2}J^\mu_{(\varphi)}$ (and this of course implies it has the right divergence).

1 Notice that in the analogous expression given in [92] the term $-\frac{1}{3}\omega^\tau_{\rho\gamma}\omega^\gamma_{\sigma\lambda}$ was accidentally omitted.

The expression (A.1) holds in any coordinate system without the need of a Killing vector. In the case of Schwarzschild space-time, in Kruskal-Szekeres coordinates (T, R, θ, φ), see for instance [77], we have:

$$\begin{aligned} J^{\mu}_{(KS)} &= \left(\frac{T\left(2r_s^2 + r_s r + r^2\right)}{r_s^2 r^4}, \frac{R\left(2r_s^2 + r_s r + r^2\right)}{r_s^2 r^4}, -\frac{4r_s \cot(\theta)}{r^5}, 0 \right) , \\ \nabla_\mu J^{\mu}_{(KS)} &= \frac{12 r_s^2}{r^6} , \end{aligned} \tag{A.4}$$

where r must be understood as a function of the new coordinates T and R. The relation $J^R_{(KS)}/J^T_{(KS)} = R/T$ implies that this current has no time component once we transform it to Schwarzschild coordinates. The radial component reads:

$$\begin{aligned} J^r_{(KS\to Schw)} &= \frac{\partial r}{\partial T} J^T_{(KS)} + \frac{\partial r}{\partial R} J^R_{(KS)} \\ &= 2\frac{(r - r_s)}{r_s r^5}\left(2r_s^2 + r_s r + r^2\right) = J^r_{(Schw)} + \frac{2}{r_s r^2} . \end{aligned} \tag{A.5}$$

Therefore, transforming back to Schwarzschild coordinates we obtain the current (A.2) plus a divergenceless term. The square of the Kruskal-Szekeres current is divergent only at the poles:

$$J^2_{(KS)} = \frac{4}{r_s^2 r^9}\left[4r_s^4 r \cot^2(\theta) + (r - r_s)\left(2r_s^2 + r_s r + r^2\right)^2\right] . \tag{A.6}$$

This does not coincide with Eq. (A.3), as expected since J^2 is not a scalar.

Notice that it is possible to take an arbitrary linear combination of the various currents obtained above, and build another one with the proper divergence. For example we can combine the currents of Eqs. (2.5) and (A.5):

$$J^{\mu}_{(finite)} = -J^{\mu}_{(\varphi)} + 2J^{\mu}_{(KS\to Schw)} = \left(0, -\frac{4r_s^2}{r^5}\left(1 - \frac{r^3}{r_s^3}\right), 0, 0\right) . \tag{A.7}$$

This current gives the correct divergence and has a finite norm everywhere for $r > 0$:

$$J^2_{(finite)} = \frac{16}{r^7}(r - r_s)\left(\frac{r}{r_s} + \frac{r_s}{r} + 1\right)^2 . \tag{A.8}$$

A.2 EQUIVALENCE BETWEEN SGB AND QUINTIC HORNDESKI WITH $G_5 = \log(X)$

In this appendix we want to check explicitly the equivalence between the sGB operator and a shift-symmetric Quintic Horndeski with $G_5 = \log(X)$ [91].

As a warm up, we can first look at the analogous case of shift-symmetric Cubic Horndeski with $G_3 = \log(X)$ in $d = 2$ dimensions and the operator $\phi\,{}^{(2)}R$, where ${}^{(2)}R$ is the two-dimensional Ricci scalar (see also Ref. [315]). The scalar current for a generic $G_3(X)$ has the following form

$$J^{\mu}_{H3} = G_{3X}\left([\Pi] g^{\mu\nu} - \Pi^{\mu\nu}\right)\partial_\nu \phi . \tag{A.9}$$

The equation of motion then reads

$$\begin{aligned} \nabla_\mu J^{\mu}_{H3} \;=\; & G_{3X}\left([\Pi]^2 - [\Pi^2]\right) + 2G_{3XX}\,\partial_\alpha\phi\partial_\beta\phi\left([\Pi]\Pi^{\alpha\beta} - \Pi^{\alpha\mu}\Pi_\mu{}^{\beta}\right) \\ & + G_{3X}\, g_{\alpha\beta}\,\partial_\mu\phi\,\nabla^{[\mu}\nabla^{\alpha]}\partial^{\beta}\phi . \end{aligned} \tag{A.10}$$

Notice that the terms with three covariant derivatives acting on ϕ arrange in an antisymmetric way, leaving behind only a term proportional to the Riemann tensor, but no third derivatives of the field, as expected from a Horndeski Lagrangian. The above equation of motion in its current form obscures the fact that there is a choice of the function $G_3(X)$ that renders the equation ϕ-independent (in $d = 2$). In order to make this manifest, it is useful to consider the Cayley-Hamilton theorem, which states that any square matrix satisfies its own characteristic equation. In this case, consider the matrix of second derivatives of the field in a given basis, $\Pi^{\mu}{}_{\nu}$, a $d \times d$ matrix, the following local identity holds in $d = 2$:

$$(\Pi^2)^{\mu}{}_{\nu} - [\Pi]\Pi^{\mu}{}_{\nu} - \frac{1}{2}\delta^{\mu}{}_{\nu}\left([\Pi^2] - [\Pi]^2\right) = 0 \quad (d = 2)\,. \tag{A.11}$$

Then it is straightforward to rewrite the equation of motion (A.10) as follows

$$\nabla_\mu J^{\mu}_{H3}\Big|_{d=2} = (G_{3X} + X\,G_{3XX})\left([\Pi]^2 - [\Pi^2]\right) - G_{3X}\,{}^{(2)}R^{\mu\nu}\partial_\mu\phi\,\partial_\nu\phi\,. \tag{A.12}$$

Finally, using that in $d = 2$ the Ricci tensor is just ${}^{(2)}R_{\mu\nu} = {}^{(2)}R\,g_{\mu\nu}/2$ and the choice $G_3 = \log(X)$ we obtain

$$\nabla_\mu J^{\mu}_{H3}\Big|_{d=2} = -\frac{1}{2}{}^{(2)}R\,, \tag{A.13}$$

which is the expected result.

Now let us turn to our case of interest. In what follows we are going to be more schematic, however the story is conceptually similar, but the calculations considerably more cumbersome due to the sheer amount of terms involved. A generic shift-symmetric Quintic Horndeski in $d = 4$ dimensions will have a scalar current with two types of terms:

$$J_{H5} \sim G_{5X}\mathcal{R}(\nabla\nabla\phi)\partial\phi + G_{5XX}(\nabla\nabla\phi)^3\partial\phi\,, \tag{A.14}$$

where $\mathcal{R}$ stands generically for the curvature. There various terms of each kind have several different contractions among the tensors, which nevertheless enjoy a particular structure due to the theory being Horndeski. The equation of motion, in turn, will schematically have the following seven types of terms,

$$\begin{aligned}\nabla J_{H5} \;\sim\; & G_{5X}\left[\nabla\mathcal{R}(\nabla\nabla\phi)\partial\phi + \mathcal{R}([\nabla,\nabla]\partial\phi)\partial\phi + \mathcal{R}(\nabla\nabla\phi)^2\right] \\ & + G_{5XX}\left[\mathcal{R}(\nabla\nabla\phi)^2(\partial\phi)^2 + (\nabla\nabla\phi)^4 + (\nabla\nabla\phi)^2([\nabla,\nabla]\partial\phi)\partial\phi\right] \\ & + G_{5XXX}\,(\nabla\nabla\phi)^4(\partial\phi)^2\,,\end{aligned} \tag{A.15}$$

where the terms were arranged according to the number of X-derivatives acting on G_5. Once again notice that, since the theory is Horndeski, the equations of motion must be of second order. Indeed, the terms with $\nabla\mathcal{R}$ cancel identically by the differential Bianchi identities, while those with third derivatives acting on the scalar always appear antisymmetrically. Some of these terms are in fact the ones giving rise to terms quadratic in the curvature. We also emphasize that, much like in the cubic case, the terms contain various possible contractions. For example, in the last term of (A.15) the two factors of $\partial\phi$ are not contracted to each other, and thus are not forming the combination X.

We rearrange once again the types of terms in the equation of motion, now in increasing powers of the curvature, obtaining, schematically,

$$\begin{aligned}\nabla J_{H5} \sim & \left[G_{5XX}(\nabla\nabla\phi)^4 + G_{5XXX}(\nabla\nabla\phi)^4(\partial\phi)^2\right] \\ & + \mathcal{R}\left[G_{5X}(\nabla\nabla\phi)^2 + G_{5XX}(\nabla\nabla\phi)^2(\partial\phi)^2\right] \\ & + G_{5X}\mathcal{R}^2(\partial\phi)^2 \,. \end{aligned} \tag{A.16}$$

At this point, if one specializes to $d = 4$ one can simplify the way indices are contracted so that, similarly to the Cubic example above, all the terms that have a $\partial_\alpha\phi\partial_\beta\phi$ become proportional to X. For the first line we make use of the Cayley-Hamilton theorem in $d = 4$. For the second line instead, it is useful to first decompose the Riemann tensor

$$R_{\mu\nu\rho\sigma} = C_{\mu\nu\rho\sigma} + E_{\mu\nu\rho\sigma} + S_{\mu\nu\rho\sigma} \,, \tag{A.17}$$

where $C_{\mu\nu\rho\sigma}$ is the Weyl tensor, and

$$E_{\mu\nu\rho\sigma} = \frac{1}{d-2}\left[g_{\mu\rho}S_{\nu\sigma} - g_{\mu\sigma}S_{\nu\rho} + g_{\nu\sigma}S_{\mu\rho} - g_{\nu\rho}S_{\mu\sigma}\right] \,, \tag{A.18}$$

$$S_{\mu\nu\rho\sigma} = \frac{R}{d(d-1)}\left[g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho}\right] \,, \tag{A.19}$$

and $S_{\mu\nu} = R_{\mu\nu} - \frac{R}{d}g_{\mu\nu}$ is the traceless part of the Ricci tensor. The pieces involving the Ricci tensor quickly combine to be proportional to a metric. For the pieces involving the Weyl tensor, a bit more work is necessary to show this, but it ultimately follows by exploiting the fact it is a fully traceless tensor. Once the X is factorized, the resulting expression combines with the terms with one less X-derivative.

Finally, let us be more explicit with the part quadratic in the curvature (third line of (A.16)),

$$\begin{aligned}(\nabla_\mu J^\mu_{H5})^{(2)} = -G_{5X}\,\partial_\alpha\phi\,\partial_\beta\phi\Bigg[& R_{\mu\nu}R^{\mu\alpha\nu\beta} - \frac{1}{2}RR^{\alpha\beta} \\ & + R^\alpha{}_\mu R^{\mu\beta} - \frac{1}{2}R_{\mu\nu\rho}{}^\alpha R^{\mu\nu\rho\beta}\Bigg] \,. \end{aligned} \tag{A.20}$$

Using the decomposition (A.17), the above expression can be brought to the form

$$(\nabla_\mu J^\mu_{H5})^{(2)} = -G_{5X}\,\partial_\alpha\phi\,\partial_\beta\phi\left[\left(\frac{1}{4}S_{\mu\nu}S^{\mu\nu} - \frac{R^2}{48}\right)g^{\alpha\beta} - \frac{1}{2}C_{\mu\nu\rho}{}^\alpha C^{\mu\nu\rho\beta}\right] \,, \tag{A.21}$$

where again the nontrivial part is the one involving the Weyl tensor. In this case, it is necessary to further decompose it into its electric and magnetic parts, defined as

$$E_{\mu\nu} = C_{\mu\alpha\nu\beta}U^\alpha U^\beta \,, \qquad B_{\mu\nu} = \tilde{C}_{\mu\alpha\nu\beta}U^\alpha U^\beta \,, \tag{A.22}$$

where U^μ is any timelike unit vector defining a local frame, and $\tilde{C}_{\mu\alpha\nu\beta}$ is the dual of the Weyl tensor, $\tilde{C}^\sigma{}_{\rho\mu\nu} := \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}C^{\alpha\beta\sigma}{}_\rho$. Here, $E_{\mu\nu}$ and $B_{\mu\nu}$ are symmetric, traceless and transverse to U^μ. An explicit expression for $C_{\mu\alpha\nu\beta}$

in terms of them can be found in Ref. [316]. With these tools, it can be shown that

$$C_{\mu\nu\rho}{}^{\alpha}C^{\mu\nu\rho\beta} = 2(d-4)E^{\alpha\mu}E_{\mu}{}^{\beta} + 2(E_{\mu\nu}E^{\mu\nu} - B_{\mu\nu}B^{\mu\nu})g^{\alpha\beta}\,. \tag{A.23}$$

Therefore, in $d = 4$, this contribution is indeed proportional to the metric. Notice that, although $E_{\mu\nu}$ and $B_{\mu\nu}$ are frame dependent, the combination on the second term above is in fact invariant. With this, we can finally write

$$(\nabla_\mu J^\mu_{H5})^{(2)} = \frac{1}{8}XG_{5X}\left[-2S_{\mu\nu}S^{\mu\nu} + \frac{R^2}{6} + 8(E_{\mu\nu}E^{\mu\nu} - B_{\mu\nu}B^{\mu\nu})\right], \tag{A.24}$$

the quantity in brackets being no other than the Gauss-Bonnet invariant $\mathcal{R}^2_{GB}$.

Putting everything together, the equation of motion can be written in the following form

$$\begin{aligned}\nabla_\mu J^\mu_{H5}\Big|_{d=4} = \\ \tfrac{XG_{5X}}{8}\mathcal{R}^2_{GB} - 2\left(G_{5X} + XG_{5XX}\right)\Big[\tfrac{1}{3}\left([\Pi]^2 - [\Pi^2]\right)R \\ - \left([\Pi]\Pi_{\mu\nu} - \Pi^2_{\mu\nu}\right)R^{\mu\nu} + \Pi_{\mu\rho}\Pi_{\nu\sigma}C^{\mu\nu\rho\sigma}\Big] \\ -\tfrac{2}{3}\left(2G_{5XX} + XG_{5XXX}\right)\Big([\Pi]^4 - 6[\Pi^2][\Pi]^2 + 3[\Pi^2]^2 + 8[\Pi^3][\Pi] - 6[\Pi^4]\Big)\,.\end{aligned} \tag{A.25}$$

We emphasize again that we crucially rely on being in $d = 4$ dimensions in order to express the equation in this form. The unique choice $G_5 = \log|X|$ makes the whole ϕ-dependence go away, leaving only

$$\nabla_\mu J^\mu_{H5} = \frac{1}{8}\mathcal{R}^2_{GB}\,. \tag{A.26}$$

A.3 REQUIREMENTS ON DHOST THEORIES

As remarked in Section 2.3, the requirements of a ghost-free decoupling limit around flat spacetime and of the presence of the Einstein-Hilbert term define two different classes: one is generated by Horndeski Lagrangians by a conformal plus disformal transformation, while the other would appear more difficult to explore. Here we will show that despite admitting the presence of both quadratic and cubic DHOST operators, this second class never admits a standard Einstein-Hilbert term, making it impossible to recover General Relativity in the limit in which $X \to 0$.

The proof makes use of the result of [141], i.e. that every quadratic DHOST theory admitting a healthy decoupling limit is connected to the quartic Horndeski Lagrangians via an invertible conformal plus disformal transformation of the form (2.27). Thus one can start by examining the condition of compatibility of the cubic part with the quadratic one, when this last is chosen to be the quartic Horndeski[2]:

$$0 = -4\frac{G^2_{4X}}{G_4} + 4\frac{G_{4X}}{X} - \frac{G_4}{X^2} = -G_4\left(2\frac{G_{4X}}{G_4} - \frac{1}{X}\right)^2\,. \tag{A.27}$$

2 The generic compatibility conditions can be found in Section 4 of [142], conditions (1) and (3) in the second table. These conditions become degenerate when the quadratic DHOST part is simply a quartic Horndeski theory.

This means that either $G_4 \equiv 0$ or $G_4 \propto \sqrt{X}$. Both these solutions correspond to theories which contain no Einstein-Hilbert term.

If we do not restrict to quartic Horndeski, the quadratic-cubic compatibility conditions become more involved. However, knowing that all the quadratic DHOST theories that we are scanning can be obtained by a conformal plus disformal transformation of a quartic Horndeski theory, we can simply inspect how the transformation (2.27) will change a function G_4 that solves Eq. (A.27):

$$\bar{G}_4\sqrt{\Omega}(\Omega + X\Gamma)^{1/2} = G_4 \propto \sqrt{X}\,. \tag{A.28}$$

This means that as long as we require $\Omega(0) = 1$ and $\Gamma(X)$ to be smooth in $X = 0$, the function $\bar{G}_4$ will not contain a constant term, therefore making impossible to retrieve General Relativity when $X = 0$.

B

RESULTS FOR BINARIES AND NRGR

B.1 LAGRANGE PLANETARY EQUATIONS AND LEADING ORDER AVERAGING

This Appendix introduces a set of equations initially introduced by Lagrange. To begin with, note that the time-dependence of the osculating elements defined by Eq. (4.19) cannot be arbitrary. We must impose a gauge-fixing condition such that the velocity is indeed given by Eq. (4.19). We denote such a condition by

$$\mathbf{C} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{v} \,, \tag{B.1}$$

where the expression for the vector v was given in Eq. (4.19). Thus, there is a relation between time derivatives of the osculating elements. This gauge-fixing condition removes three degrees of freedom (equivalently, six variables in phase space) from the six degrees of freedom contained in the six osculating elements (equivalently, twelve variables in phase space).

Now, we could write a Lagrangian for the osculating elements by implementing this constraint with a Lagrange multiplier $\boldsymbol{\lambda}$, so that

$$\mathcal{L} = \frac{1}{2}\mu v^2 + \frac{G\mu m}{r} + \boldsymbol{\lambda}\cdot\left(\mathbf{C}-\mathbf{v}\right) + \mathcal{L}_1 \,, \tag{B.2}$$

where $\mathcal{L}_1$ has been defined in Eq (4.1). From there one could deduce the Lagrange planetary equations (LPE) which relate time derivatives of the osculating elements to the perturbing function $\mathcal{L}_1 \equiv \mu\mathcal{R}$. However, it is much easier to derive them in a Hamiltonian formalism, see e.g. [317] to which we refer the reader interested in the details of the derivation.

The LPE are traditionally expressed using the following angles: ι is the inclination, ω the argument of periapsis, and Ω the longitude of the ascending node. In term of these, the unit vectors $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\gamma}}$ are expressed as

$$\begin{aligned} \hat{\boldsymbol{\alpha}} &= R_z(\Omega)R_x(\iota)R_z(\omega)\hat{\mathbf{u}}_x \,, \\ \hat{\boldsymbol{\gamma}} &= R_z(\Omega)R_x(\iota)R_z(\omega)\hat{\mathbf{u}}_z \,, \end{aligned} \tag{B.3}$$

where $\hat{\mathbf{u}}_x$, $\hat{\mathbf{u}}_y$, $\hat{\mathbf{u}}_z$ are the Cartesian basis vectors. Using these angles, the LPE are given by [317]

$$\dot{a} = \sqrt{\frac{4a}{G_N m}}\frac{\partial \mathcal{R}}{\partial u}\,, \tag{B.4}$$

$$\dot{e} = -\sqrt{\frac{1-e^2}{G_N m a e^2}}\frac{\partial \mathcal{R}}{\partial \omega} + \frac{1-e^2}{\sqrt{G_N m a e}}\frac{\partial \mathcal{R}}{\partial u}\,, \tag{B.5}$$

$$\dot{\imath} = -\frac{1}{\sqrt{G_N m a(1-e^2)}\sin\imath}\frac{\partial \mathcal{R}}{\partial \Omega} + \frac{\cos\imath}{\sqrt{G_N m a(1-e^2)}\sin\imath}\frac{\partial \mathcal{R}}{\partial \omega}\,, \tag{B.6}$$

$$\dot{u} = \sqrt{\frac{G_N m}{a^3}} - \sqrt{\frac{4a}{G_N m}}\frac{\partial \mathcal{R}}{\partial a} - \frac{1-e^2}{\sqrt{G_N m a e}}\frac{\partial \mathcal{R}}{\partial e}\,, \tag{B.7}$$

$$\dot{\omega} = \sqrt{\frac{1-e^2}{G_N m a e^2}}\frac{\partial \mathcal{R}}{\partial e} - \frac{\cos\imath}{\sqrt{G_N m a(1-e^2)}\sin\imath}\frac{\partial \mathcal{R}}{\partial \imath}\,, \tag{B.8}$$

$$\dot{\Omega} = \frac{1}{\sqrt{G_N m a(1-e^2)}\sin\imath}\frac{\partial \mathcal{R}}{\partial \imath}\,. \tag{B.9}$$

It can be checked that the LPE can be derived from the following fist-order Lagrangian:

$$\mathcal{L} = \mu\left[\frac{G_N m}{2a} + L\dot{u} + G\dot{\omega} + H\dot{\Omega}\right] + \mathcal{L}_1\,, \tag{B.10}$$

The conjugate momenta are given by

$$L = \sqrt{G_N m a}\,,\; G = L\sqrt{1-e^2}\,,\; H = G\cos\imath\,. \tag{B.11}$$

This Lagrangian is exact, however, it is not manifestly invariant under a rotation of the basis vectors; such a manifest invariance can be recovered by noticing that the angular part can be rewritten as

$$\mu\left[G\dot{\omega} + H\dot{\Omega}\right] = \mu G\,\hat{\boldsymbol{\beta}}\cdot\dot{\hat{\boldsymbol{\alpha}}} = \boldsymbol{J}\cdot\boldsymbol{\Omega}\,, \tag{B.12}$$

where $\mathbf{J}$ is the total angular momentum of the binary and Ω is an angular velocity defined by

$$\boldsymbol{J} = \mu\sqrt{G_N m a(1-e^2)}\,\hat{\boldsymbol{\gamma}}\,,\quad \boldsymbol{\Omega} = \hat{\boldsymbol{\alpha}}\times\dot{\hat{\boldsymbol{\alpha}}}\,. \tag{B.13}$$

Thus, the angular kinetic term can be identified with a spin coupling in flat space (note that our sign convention for the metric is different from the one used in e.g Refs [223, 285], which explains the sign difference of the kinetic term). However, note that not all the components of the spin vector are independent, since the Lagrangian shown in (B.12) displays only two degrees of freedom (corresponding to four equations in phase space once a variational principle is applied). Indeed, notice that if one wants to vary the Lagrangian with respect to $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ in order to keep a manifest rotational invariance, one should also impose that these vectors should be unitary and orthogonal in order to preserve the right number of degrees of freedom.

In terms of these canonical variables we can derive the Lagrange Planetary Equations (B.4) by expressing the time-derivatives of all osculating orbital[1] elements in terms of derivatives of the Hamiltonian [304]. They can be obtained as the equations of motion stemming from the Lagrangian (B.52).

1 A similar analysis leads to the same conclusions in the case of the contact elements, i.e. the variables used in Chapter 5.

The derivatives of $\mathcal{R}$ with respect to canonical momenta can be transformed in derivatives with respect to planetary elements by inverting the relations (B.53):

$$a = \frac{L^2}{Gm}\,, \quad e = \sqrt{1-\frac{G^2}{L^2}}\,, \quad \cos\iota = \frac{H}{G}\,, \tag{B.14}$$

Thus, one has

$$\frac{\partial\mathcal{R}}{\partial L} = 2\sqrt{\frac{a}{G_N m}}\frac{\partial\mathcal{R}}{\partial a} + \frac{1-e^2}{e\sqrt{G_N m a}}\frac{\partial\mathcal{R}}{\partial e}\,, \tag{B.15}$$

$$\frac{\partial\mathcal{R}}{\partial G} = -\frac{\sqrt{1-e^2}}{e\sqrt{G_N m a}}\frac{\partial\mathcal{R}}{\partial e} + \frac{\cos\iota}{\sqrt{G_N m a(1-e^2)}\sin\iota}\frac{\partial\mathcal{R}}{\partial\iota}\,, \tag{B.16}$$

$$\frac{\partial\mathcal{R}}{\partial H} = -\frac{1}{\sqrt{G_N m a(1-e^2)}\sin\iota}\frac{\partial\mathcal{R}}{\partial\iota}\,. \tag{B.17}$$

In these equations, the derivatives with respect to the osculating elements are taken holding all other elements fixed. However, we should be careful because the eccentric anomaly η depends on both the eccentricity and the semimajor axis through the equation $\eta - e\sin\eta = \sqrt{G_N m/a^3}\,t + \sigma$. Consequently,

$$\left.\frac{\partial\eta}{\partial e}\right|_{\sigma,a\text{ fixed}} = \frac{\sin\eta}{1-e\cos\eta}\,, \quad \left.\frac{\partial\eta}{\partial a}\right|_{\sigma,e\text{ fixed}} = -\frac{3nt}{2a(1-e\cos\eta)}\,. \tag{B.18}$$

Thus we obtain:

$$\begin{aligned}\frac{\partial\mathcal{R}}{\partial e} &= \left.\frac{\partial\mathcal{R}}{\partial e}\right|_{\mathrm{J}\text{ fixed}} + \frac{\sin\eta}{1-e\cos\eta}\frac{\partial\mathcal{R}}{\partial\eta}\,,\\ \frac{\partial\mathcal{R}}{\partial a} &= \left.\frac{\partial\mathcal{R}}{\partial a}\right|_{\mathrm{J}\text{ fixed}} - \frac{3nt}{2a(1-e\cos\eta)}\frac{\partial\mathcal{R}}{\partial\eta}\,,\end{aligned} \tag{B.19}$$

leading to Eq. (B.4) given above.

Finally, to obtain a long time-scale description of the system, one needs to average out, the quasi-periodic orbital motion of the binary system. This corresponds to eliminating the short time-scale degree of freedom contained in the mean anomaly u. As a consequence, since the perturbing function does not depend on u any more, one expects the semimajor axis a to be constant through time from Eq. (B.4). We will clarify to what extent this is the case in App. B.4. This fact implies that the two-body Lagrangian shown in Eq.(B.10) is indeed equivalent to a spin kinetic term, since the term $G_N m/2a + L\dot{u}$ becomes an irrelevant constant.

After this elimination, the binary system is described by four dynamical quantities (the eccentricity e and the three Euler angles defined above) which vary over a timescale much greater than the period of the binary. At leading order, we can use the following formula to define the average of a given quantity A:

$$\begin{aligned}\langle A\rangle &= \frac{1}{T}\int_0^T \mathrm{d}t\; A(t) = \frac{1}{T}\int_0^{2\pi}\frac{\mathrm{d}t}{\mathrm{d}\eta}\mathrm{d}\eta\; A(\eta)\,,\\ T &= \int_0^{2\pi}\frac{\mathrm{d}t}{\mathrm{d}\eta}\mathrm{d}\eta\,.\end{aligned} \tag{B.20}$$

Using Eq. (4.20), one has

$$\frac{\mathrm{d}t}{\mathrm{d}\eta} = \frac{1-e\cos\eta}{\dot{u}+\dot{e}\sin\eta}\,. \tag{B.21}$$

At lowest order in the perturbing function $\mathcal{L}_1$, one has $\dot{u} + \dot{e}\sin\eta = \sqrt{G_N m/a^3}$ and $T = 2\pi\sqrt{a^3/(G_N m)}$, so that the mean value becomes

$$\langle A \rangle = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\eta (1 - e\cos\eta) A(\eta)\,. \tag{B.22}$$

Going beyond this naive averaging technique to capture further corrections due to either backreaction of fast modes or deviations from adiabaticity, we will need a more careful analysis of the dynamical variables in play, which we carry out in App. B.3. As we will see in detail in Chapter 5, these corrections contribute at the quadrupolar 1PN level.

B.2 SPIN KINETIC TERM AND GAUGE FIXING OF ROTATIONAL VARIABLES

In this appendix we provide some details of the computation of the spin kinetic term (4.35) as a function of the intrinsic angular momentum of the inner binary. The computations are analogous to those carried out in [223], with the difference that we specialize to the *no mass dipole* gauge in which the time components of the spin tensor are set to zero. This choice will make simple to connect the spin tensor to the orbital angular momentum.

First of all, it is useful to introduce a worldline tetrad $e^\mu_A(\sigma)$ defined only on the worldline $y^\mu(\sigma)$ (σ being the affine parameter of the curve) which represents a choice of axes in the rest-frame of the body and satisfies: $g_{\mu\nu}(y(\sigma))e^\mu_A(\sigma)e^\nu_B(\sigma) = \eta_{AB}$. This tetrad can be used to define the angular velocity of the body: $\Omega^{\mu\nu} = e^\nu_A(De^{\mu A}/D\sigma)$, whose conjugate is the spin tensor $J_{\mu\nu} = 2\partial\mathcal{L}/\partial\Omega^{\mu\nu}$. Both these tensors contain gauge degrees of freedom, since only the spatial orientation of the worldline tetrad has a physical meaning. In fact we can choose arbitrarily its time-like direction, encoded in $e^\mu_{[0]}$. This gauge choice corresponds to a redundant boost transformation of the worldline tetrad (in order to avoid ambiguities between the different set of indices, we are using square brackets to distinguish the flat indices of the worldline tetrad from the others).

The gauge fixing of $e^\mu_{[0]}$ must be supplemented with a gauge fixing of the conjugate variables in $J_{\mu\nu}$, the so-called Spin Supplementary Condition (SSC). Starting from a covariant gauge choice in which $e^\mu_{[0]} = p^\mu/\sqrt{-p^2}$ and the spin tensor satisfies the covariant SSC $J_{\mu\nu}p^\nu = 0$, the action of a boost will change the worldline tetrad and the angular velocity tensor. The changes produced by this transformation in the Lagrangian can be interpreted by means of a redefinition of the spin tensor and a consequent change of the SSC. We will therefore use this boost degree of freedom to first pick the *no mass dipole* SSC for the spin tensor and then to fix the canonical gauge for the angular velocity vector. In this way we will get an expression dependent only on the intrinsic angular momentum of the binary.

As computed in [223], the transformation of the spin kinetic term of the Lagrangian under a boost of the worldline tetrad (starting from the covariant gauge and SSC) is the following:

$$\frac{1}{2}J_{\mu\nu}\Omega^{\mu\nu} = \frac{1}{2}\hat{J}_{\mu\nu}\hat{\Omega}^{\mu\nu} + \frac{p^\lambda}{-p^2}\hat{J}_{\mu\lambda}\frac{Dp^\mu}{D\sigma}\,, \tag{B.23}$$

where we have used hatted symbols to label boosted variables and in particular we have defined the boosted spin tensor to be $J_{\mu\nu} = \hat{J}_{\mu\nu} - \delta z_\mu p_\nu + \delta z_\nu p_\mu$, with $\delta z_\mu = \hat{J}_{\mu\rho}p^\rho/(-p^2)$. We can interpret

this change of the spin tensor as due to a shift of the center of the body rotation, that is the point where the worldline intersects the body. In the case of the *no mass dipole* gauge, in which the spin tensor is purely spatial, this shift corresponds to setting the center of the worldline on the relativistic center of mass, as shown in the main text. The second term in Eq.(B.23) will instead contribute to the Thomas precession, which we can understand as due to a gravitational torque associated to the finite size of rotating objects in GR.

Before specifying the boost needed to get to the desired SSC, it is useful to disentangle the gravitational field from the spinning degrees of freedom. We can do so by introducing the gravitational tetrad field $g_{\mu\nu}(x)\tilde{e}^{\mu}_{a}(x)\tilde{e}^{\nu}_{b}(x) = \eta^{ab}$, which is defined on the whole space-time. This tetrad can be related to the worldline tetrad by means of a Lorentz transformation: $\tilde{e}^{\mu}_{a}(y(\sigma)) = \Lambda^{A}_{a}(\sigma)e^{\mu}_{A}(\sigma)$, being $\Lambda^{A}_{a}(\sigma)$ a Lorentz matrix dependent on the affine parameter of the worldline. As for the worldline tetrad flat indices, when needed we will use round brackets to distinguish the flat indices of the tetrad field from the others.

In this notation, once the gauge of the tetrad field is fixed, we can fix the time-like vector of the worldline tetrad by choosing the boosted zero components of the Lorentz matrices: $\hat{\Lambda}^{[0]}_{a}$. Moreover, introducing the tetrad field will make possible to write all the objects in the right hand side of Eq.(B.23) in terms of their counterparts with flat indices. Such quantities correspond to those computed in terms of the intrinsic angular momentum of the binary, as they are independent on the external gravitational field. In particular we have:

$$\frac{1}{2}\hat{J}_{\mu\nu}\hat{\Omega}^{\mu\nu} = \frac{1}{2}\hat{J}_{ab}\hat{\Omega}^{ab}_{flat} + \frac{1}{2}\hat{J}_{ab}\omega^{ab}_{\mu}u^{\mu}\,, \tag{B.24}$$

where $\omega^{ab}_{\mu} = \tilde{e}^{a}_{\nu}\nabla_{\mu}\tilde{e}^{b\nu}$ is the spin connection of the tetrad field, $u^{\mu} = dy^{\mu}/d\sigma$ is the worldline speed and we have defined $\hat{\Omega}^{ab}_{flat} = \hat{\Lambda}^{b}_{A}d\hat{\Lambda}^{aA}/d\sigma$.

At this point we can fix the gauge boost of the worldline tetrad. In order to set the time components of the spin tensor to zero, $\hat{J}_{a(0)} = 0$, we need to choose a boost such that $\sqrt{p^2}\hat{\Lambda}_{[0]a} = 2p_0\delta_{0a} - p_a$ (this can be understood by inspecting the generic expression for $\hat{J}_{\mu\nu}$, as discussed in [223]). Doing so, we obtain the following:

$$\begin{aligned}\frac{1}{2}J_{\mu\nu}\Omega^{\mu\nu} =& \frac{1}{2}\hat{J}_{(i)(j)}\hat{\Omega}^{(i)(j)}_{flat} + \frac{1}{2}\hat{J}_{(i)(j)}\omega^{(i)(j)}_{\mu}u^{\mu} \\ &+ \frac{p^{(j)}}{-p^2}\hat{J}_{(i)(j)}\tilde{e}^{(i)}_{\mu}\frac{Dp^{\mu}}{D\sigma}\,,\end{aligned} \tag{B.25}$$

This gauge choice makes possible to unpack $\hat{\Omega}^{(i)(j)}_{flat}$ and express $\hat{\Lambda}^{a}_{[0]}d\hat{\Lambda}^{b[0]}/d\sigma$ in terms of the momentum of the worldline, leaving to compute only on the spatial part of the Lorentz matrices. However, these spatial leftovers won't be $SO(3)$ matrices, since they need to satisfy the condition $\hat{\Lambda}^{a}_{A}\eta^{AB}\hat{\Lambda}^{b}_{B} = \eta^{ab}$ and will carry a dependence on the worldline momentum, due to the gauge condition on $\hat{\Lambda}^{a}_{[0]}$.

In order to obtain an angular velocity tensor defined in terms of rotation matrices and to remove its dependence on the worldline momentum, we can take a further boost of the worldline tetrad. This time however, we will not use a redefinition of the spin tensor to absorb the new terms appearing

in the Lagrangian after the transformation. Rather, we will retain the spin tensor satisfying the *no mass dipole* SSC and we will keep track of the new terms explicitly.

In order to make $\hat{\Lambda}^a_A$ an $SO(3)$ matrix, we need to choose a gauge in which $\hat{\Lambda}^a_{[0]} = \delta^a_0$. Therefore we implement a boost of the worldline tetrad that sends the time-like unit vector $(2p_0\delta^a_0 - p^a)/\sqrt{-p^2}$ to δ^a_0 . This transformation will change only the first term in Eq.(B.25) as follows:

$$\frac{1}{2}\hat{J}_{(i)(j)}\hat{\Omega}^{(i)(j)}_{flat} = \frac{1}{2}\hat{J}_{(i)(j)}\Omega^{(i)(j)}_{SO(3)} + \frac{1}{2}\hat{J}_{(i)(j)}u^{(i)}\frac{du^{(j)}}{d\sigma}\,, \tag{B.26}$$

where now $\Omega^{(i)(j)}_{SO(3)}$ is build out of rotation matrices and we have used $p^a = mu^a/\sqrt{-u^2}$, with $-u^2 = 1$ at leading order in the PN expansion.

Having fixed the gauge for both angular velocity and spin tensor, we can carry out the explicit computation of the last two terms in Eq.(B.25). In order to do so, we pick the tetrad field in such a way to have $\tilde{e}^{(i)}_0 = 0$. Then, at 1PN order we obtain:

$$\begin{aligned}\frac{1}{2}\hat{J}_{ab}\omega^{ab}_{\mu}u^{\mu} &= \frac{1}{2}\hat{J}_{(i)(j)}(4u^{(i)}\partial^{(j)}\tilde{\phi} + \partial^{(i)}\tilde{A}^{(j)})\,, \\ \frac{p^{(j)}}{-p^2}\hat{J}_{(i)(j)}\tilde{e}^{(i)}_{\mu}\frac{Dp^{\mu}}{D\sigma} &= \hat{J}_{(i)(j)}u^{(j)}\left(\frac{du^{(i)}}{d\sigma} + \partial^{(i)}\tilde{\phi}\right)\,.\end{aligned} \tag{B.27}$$

Plugging these results into Eq.(B.23), and identifying the worldline with the trajectory of the center of mass, $u^{(i)} = V^{(i)}_{CM}$ we finally get:

$$\begin{aligned}\frac{1}{2}J_{\mu\nu}\Omega^{\mu\nu} =& \frac{1}{2}\hat{J}_{(i)(j)}\Omega^{(i)(j)}_{SO(3)} + \frac{1}{2}\hat{J}_{(i)(j)}A^{(i)}_{CM}V^{(j)}_{CM} \\ &+ \frac{1}{2}\hat{J}_{(i)(j)}(2V^{(i)}_{CM}\partial^{(j)}\tilde{\phi} + \partial^{(i)}\tilde{A}^{(j)})\,.\end{aligned} \tag{B.28}$$

Then, with a mild abuse of notation, we can drop the index brackets and the hats so as to match the expressions used (for simplicity) in the main text: $\hat{J}_{(i)(j)} \mapsto J_{ij}$, $\Omega^{(i)(j)}_{SO(3)} \mapsto \Omega^{ij}$. We stress however that these are different from the $(\mu,\nu) = (i,j)$ components of $J_{\mu\nu}$ and $\Omega_{\mu\nu}$, which depend on the external gravitational field.

Using this notation and the definitions $J_{ij} = \epsilon_{ijk}J^k$, $\Omega_{ij} = \epsilon_{ijk}\Omega^k$, we can rewrite Eq.(B.28) as:

$$\begin{aligned}\frac{1}{2}J_{\mu\nu}\Omega^{\mu\nu} &= \boldsymbol{J}\cdot\boldsymbol{\Omega} + \frac{1}{2}J_{ij}A^i_{CM}V^j_{CM} \\ &\quad + \frac{1}{2}J_{ij}(2V^i_{CM}\partial^j\tilde{\phi} + \partial^i\tilde{A}^j)\,,\end{aligned} \tag{B.29}$$

which is the equation used in the main text.

REFINED AVERAGING AND CROSS TERMS

B.3 AVERAGING THROUGH NEAR-IDENTITY TRANSFORMATIONS

We here present the averaging procedure that we adopt in our computations of Chapter 5, the so-called averaging by near-identity transformations [72].

We do so in the context of a toy model that closely resembles the three-body problem, an angle-periodic system in its canonical form. While in the three body problem there are two angle variables (the mean anomalies of inner and outer orbits) and several slowly evolving variables, we consider a system with one angle variable, u, and one slowly evolving variable x:

$$\begin{aligned} \frac{d}{dt}x &= \mathrm{e}\, g_1(x,u) + \mathrm{e}^2 g_2(x,u) + \dots \,, \\ \frac{d}{dt}u &= H(x) + \mathrm{e}\, h_1(x,u) + \mathrm{e}^2 h_2(x,u) + \dots \,, \end{aligned} \tag{B.30}$$

where $\mathrm{e} \ll 1$, the dots indicate terms with $i > 2$, and the functions g_i , h_i are periodic in u with period 2π. Physically, we can think of e as a ratio of time-scales, since the slow variable x has excursions of order 1 over times that are $1/\mathrm{e}$ longer than those over which u changes its value by an order 1 factor. We implement the splitting between fast and slow dynamics using a change of variables:

$$\begin{aligned} u(t) &= u_L(t) + \mathrm{e}\, u_S(x_L, u_L) \,, \\ x(t) &= x_L(t) + \mathrm{e}\, x_S(x_L, u_L) \,, \end{aligned} \tag{B.31}$$

where u_S, x_S encode the short time-scale dynamics and are chosen to be periodic in u_L with period 2π. Eq. (B.31) is called a near-identity transformation, since for $\mathrm{e} \to 0$ it reduces to an identity. Physically, this ansatz encodes the fact that fast oscillations will have suppressed amplitude. The strategy that we adopt is to fix the functions u_S , x_S in such a way to cancel quickly oscillating terms in the equations (B.30) order by order in e, up to a desired accuracy. Then we truncate the equations, neglecting the higher order corrections that still contain oscillating terms. This will leave a system of equations for u_L , x_L that only depends on the long time-scale, up to corrections that are of arbitrarily high order (which are neglected after the truncation). Plugging the transformation in Eq. (B.30) we have, to second order:

$$\begin{aligned} \dot{x}_L + \mathrm{e}(\dot{x}_L\partial_{x_L} + \dot{u}_L\partial_{u_L})x_S =& \,\mathrm{e}\, g_1(x_L,u_L) + \mathrm{e}^2(x_S\partial_{x_L} + u_S\partial_{u_L})g_1(x_L,u_L) \\ &+ \mathrm{e}^2 g_2(x_L,u_L) + \mathcal{O}(\mathrm{e}^3) \,, \\ \dot{u}_L + \mathrm{e}(\dot{x}_L\partial_{x_L} + \dot{u}_L\partial_{u_L})u_S =& H(x_L) + \mathrm{e}\, H'(x_L)x_S + \frac{\mathrm{e}^2}{2}H''(x_L)x_S^2 \\ &+ \mathrm{e}\, h_1(x_L,u_L) + \mathrm{e}^2(x_S\partial_{x_L} + u_S\partial_{u_L})h_1(x_L,u_L) \\ &+ \mathrm{e}^2 h_2(x_L,u_L) + \mathcal{O}(\mathrm{e}^3) \,, \end{aligned} \tag{B.32}$$

Now, order by order, we can determine x_S , u_S in such a way that the derivatives $\dot{x}_L$, $\dot{u}_L$ only depend on x_L, up to terms of order e^3:

$$\begin{aligned} \dot{x}_L =& \,\mathrm{e}\, G_1(x_L) + \mathrm{e}^2 G_2(x_L) + \mathcal{O}(\mathrm{e}^3) \,, \\ \dot{u}_L =& H(x_L) + \mathrm{e}\, H_1(x_L) + \mathrm{e}^2 H_2(x_L) + \mathcal{O}(\mathrm{e}^3) \,. \end{aligned} \tag{B.33}$$

In order to obtain this, we write $x_S = x_S^{(0)} + \mathrm{e}\, x_S^{(1)} + \dots$ and $u_S = u_S^{(0)} + \mathrm{e}\, u_S^{(1)} + \dots$. Then, comparing Eq. (B.33) with Eq. (B.32) we find that at first order it must be:

$$\begin{aligned} \partial_{u_L} x_S^{(0)} =& \frac{1}{H(x_L)}\Big(g_1(x_L,u_L) - G_1(x_L)\Big) \,, \\ \partial_{u_L} u_S^{(0)} =& \frac{1}{H(x_L)}\Big(h_1(x_L,u_L) + H'(x_L)x_S^{(0)}(x_L,u_L) - H_1(x_L)\Big) \,. \end{aligned} \tag{B.34}$$

Integrating the first equation at fixed x_L, we find:

$$x_S^{(0)}(x_L, u_L) = \int_0^{u_L} \frac{ds}{H(x_L)} \Big(g_1(x_L, s) - G_1(x_L) \Big) + C_0(x_L) \,. \tag{B.35}$$

From this, since by assumption $x_S^{(0)}$ is periodic in u_L with period 2π, we see that it must be:

$$\int_0^{2\pi} \frac{ds}{H(x_L)} \Big(g_1(x_L, s) - G_1(x_L) \Big) = 0 \,, \tag{B.36}$$

which in turn implies that G_1 must be chosen to be the average of g_1:

$$G_1 = \langle g_1 \rangle_{u_L} = \int_0^{2\pi} \frac{ds}{2\pi} g_1(x_L, s) \,. \tag{B.37}$$

We can further set $C_0(x_L)$ in such a way that $x_S^{(0)}$ has zero average. Schematically we will write:

$$x_S^{(0)}(x_L, u_L) = \frac{1}{H(x_L)} \mathrm{Af}_{u_L} \Big(\int^{u_L} ds \mathrm{Af}_s(g_1(x_L, s)) \Big), \tag{B.38}$$

where the symbol Af_x indicates taking the average-free part of the argument with respect to the variable x.

Note that in this approach, the average is defined without ambiguities through an integral over u_L with x_L fixed, as a by-product of requiring x_S to be periodic. This procedure can be thought of as an average over the short time-scale characterizing the evolution of u_L, while keeping fixed the variable x_L. Therefore, we are not truly performing an average over time, but a procedure that is very similar and which ensures all the same an arbitrarily precise approximation.

Turning to the equation for $u_S^{(0)}$, we find in the same way that $H_1(x_L) = \langle h_1 \rangle_{u_L}$, since the quantity $H' x_S^{(0)}$ is average-free (thanks to our choice of C_0). This means that if h_1 is average-free, then $\dot{u}_L$ will receive corrections starting at order e^2. We will also choose the constant of integration for $u_S^{(0)}$ in such a way to make it average-free. Thus it will be:

$$u_S^{(0)}(x_L, u_L) = \frac{1}{H(x_L)} \mathrm{Af}_{u_L} \Big(\int^{u_L} ds \Big[\mathrm{Af}_s(h_1(x_L, s)) + H'(x_L) x_S^{(0)}(x_L, s) \Big] \Big) \,. \tag{B.39}$$

This first order truncation of the approximation, when applied to the Newtonian hierarchical three-body problem, gives the Kozai-Lidov long time-scale dynamics. Turning to second order, we have:

$$\begin{aligned} \partial_{u_L} x_S^{(1)} &= \frac{1}{H(x_L)} \Big(g_2 + (x_S^{(0)} \partial_{x_L} + u_S^{(0)} \partial_{u_L}) g_1 - H_1 \partial_{u_L} x_S^{(0)} - G_2(x_L) \Big) \,, \\ \partial_{u_L} u_S^{(1)} &= \frac{1}{H(x_L)} \Big(h_2 + H' x_S^{(1)} + \frac{H''}{2} (x_S^{(0)})^2 + (x_S^{(0)} \partial_{x_L} + u_S^{(0)} \partial_{u_L}) h_1 \\ &\qquad - H_1 \partial_{u_L} u_S^{(0)} - H_2(x_L) \Big) \,, \end{aligned} \tag{B.40}$$

where $g_{1,2}$, $h_{1,2}$ are evaluated in (x_L, u_L). Again, requiring the near-identity transformation to be periodic, we find that G_2 and H_2 will be averages of the other terms in the right hand side. Consequently, $x_S^{(1)}$ and $u_S^{(1)}$ will be integrals of average-free expressions, and we will be able to fix the constants

of integration so as to make $x_S^{(1)}$, $u_S^{(1)}$ average-free as well. In particular, we find:

$$G_2 = \int_0^{2\pi} \frac{ds}{2\pi} \Big(g_2(x_L, s) + (x_S^{(0)} \partial_{x_L} + u_S^{(0)} \partial_{u_L}) g_1(x_L, s) \Big) \,, \tag{B.41}$$
$$H_2 = \int_0^{2\pi} \frac{ds}{2\pi} \Big(h_2(x_L, s) + (x_S^{(0)} \partial_{x_L} + u_S^{(0)} \partial_{u_L}) h_1(x_L, s) + \frac{H''}{2} (x_S^{(0)})^2 \Big) \,,$$

where we have dropped the terms $H_1 \partial_{u_L} x_S^{(0)}$, $H_1 \partial_{u_L} u_S^{(0)}$ and $H' x_S^{(1)}$, which are average free. These functions can then be plugged in the equations for x_L and u_L, and the slow dynamics can be determined up to terms of order e^3. Similarly, one can derive the averaging procedure to any order in ϵ simply by fixing higher orders of the near-identity transformation, through the functions x_S and u_S.

The toy model just discussed describes well the procedure we would adopt if we expressed our quantities in terms of the mean anomaly. However, we can only give a closed form expression of the perturbing function in terms of the eccentric anomaly. If we wished to use the variable u as a proxy for the eccentric anomaly, then we would have to allow a dependence on u of the function H driving the leading order evolution of u. It would be then much less straightforward to understand how to define the near identity transformation, since we would have to decide whether we want to retain a u_L dependence in the functions H_i driving the evolution of u_L. Moreover, the equations for u_S would require a more cumbersome integration.

For this reason, we find useful to exploit the averaging procedure derived above, in terms of the mean anomaly, and to simply change integration variables in a consistent way to the eccentric anomaly. If we call u the mean anomaly and η the eccentric anomaly, then we know that Kepler's equation holds at all times:

$$u = \eta - e \sin \eta \,, \tag{B.42}$$

where e is the eccentricity, which we can regard as a component of what in general will be the x vector. This relation ensures that any function periodic in u is also be periodic in η. Given Kepler's equation, we can perform the near identity transformation and find a relation between u_L and the eccentric anomaly. Suppose the eccentricity is transformed as:

$$e = e_L(x_L) + \mathrm{e}\, e_S(x_L, u_L) \,, \tag{B.43}$$

then we see that it must hold:

$$u_L + \mathrm{e}\, u_S(x_L, u_L) = \eta - (e_L(x_L) + \mathrm{e}\, e_S(x_L, u_L)) \sin \eta \,. \tag{B.44}$$

Although the relation between η and u_L is very involved, we can avoid complications by performing a near-identity transformation on η as well:

$$\eta = \eta_L + \mathrm{e}\, \eta_S(x_L, u_L) \,. \tag{B.45}$$

Then, it is possible to fix η_S so as to retain, to all orders, the relation:

$$u_L = \eta_L - e_L \sin \eta_L \,. \tag{B.46}$$

This choice makes possible to express the integrands evaluated in u_L as simple functions of η_L. Moreover, it makes clear that at all orders the change of variables will be given by:

$$du_L = d\eta_L (1 - e_L \cos \eta_L) \,. \tag{B.47}$$

Note that crucially, as a result of integrating at fixed x_L, the Jacobian does not have the denominator factor $\dot{u} + \dot{e} \sin \eta$ that would appear in $dt/d\eta$. This is to say that the averaging procedure that we have presented provides a coarse graining of the system in the time coordinate while removing the need of actually performing an integral over time.

Before concluding, let us mention other methods to eliminate fast variables besides the method of near-identity transformations. One alternative is the so-called multiple scale analysis, for instance discussed in [318, 319]. The idea behind this method is to introduce fictitious variables corresponding to long time-scales, and to determine the dependence on these long time-scales by imposing the cancellation of terms that display a secular growth, i.e. terms growing linearly with time which would break the perturbative expansion early on. As far as we know, the method of near identity transformations is to be preferred over the multiple scale analysis if one is interest in estimating the range of validity in time of the approximate solution [72]. Besides the multiple scale analysis, we quote the method of Von Zeipel transformations, [249, 269], which works at the level of the Hamiltonian implementing canonical transformations that eliminate the dependence on short time-scale modes, very similarly to what we have done above. Finally, we quote the method of dynamical renormalization group [320], which operates in a way similar to the multiple scale analysis, removing secularly growing terms by means of counterterms.

B.4 CONSERVATION OF CONTACT SEMI-MAJOR AXIS

We now turn to the question of whether the semi-major axes of the two orbits remain constant over long time-scales. Despite the simplicity of this question, to our understanding the answer is quite involved. There is an intuitive reasoning to argue that the semi-major axes are constant over long time-scales. That is, after the averaging procedure is carried out, the Lagrangian becomes independent on the mean anomaly of the corresponding orbit, therefore making the corresponding conjugate momentum constant. The latter, as shown in Eq. (B.53), depends only on the semimajor axis. However, when considering PN corrections this argument can only hold for the contact element a, rather than for the orbital element $\tilde{a}$, see discussion in Appendix B.9. Moreover, as we discuss in Appendix B.3, the dynamical variables left after the averaging procedure are the long time-scale modes of the original, full contact elements. Therefore the statement will not hold for the full contact element, but only for its slowly evolving part.

The way to make this intuition rigorous is to implement a canonical transformation of the Hamiltonian, eliminating order by order the dependence on the mean anomalies. This is achieved through the Von Zeipel transformations, as discussed in [249]. The result is that indeed, as long as perturbation theory goes, the semimajor axes remain constant over long time-scales.

The systematic control over each order in perturbation theory given by the near identity transformations allows for an independent check of this statement. Explicitly, using the formulas derived for the toy model (B.30), we can check that, for instance, to second order we expect the semimajor axis to be constant on long timescales. To argue this, we can think of

the functions appearing as sources on the right hand side of Eq. (B.30) as partial derivatives of the Hamiltonian of the system $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1 + \dots$:

$$g_1 = -\frac{\partial \mathcal{H}_1}{\partial u} \, , \; g_2 = -\frac{\partial \mathcal{H}_2}{\partial u} \, , \qquad H = \frac{\partial \mathcal{H}_0}{\partial x} \, , \; h_1 = \frac{\partial \mathcal{H}_1}{\partial x} \, , \tag{B.48}$$

where we assume $\mathcal{H}$ to be periodic in u, meaning that the functions g_i will be average free. This allow us to inspect order by order whether the time derivative of the conjugate momentum to the mean anomaly vanishes or not, by computing the long time-scale source terms G_i, in Eq. (B.33). For instance, it is evident that $G_1 = 0$, due to g_1 being average free. For G_2, given in Eq.(B.41), determining the answer is less straightforward. Schematically, we have:

$$\begin{aligned} G_2 &= \int_0^{2\pi} \frac{ds}{2\pi} \left\{ \frac{1}{H} \mathrm{Af}(\mathcal{H}_1) \partial_x \partial_u \mathcal{H}_1 - \partial_x \left(\frac{1}{H} \mathrm{Af}\left(\int^u \mathrm{Af}(\mathcal{H}_1) \right) \right) \partial_u^2 \mathcal{H}_1 \right\} \\ &= \int_0^{2\pi} \frac{ds}{2\pi} \left\{ \partial_x \left(\frac{1}{H} \mathrm{Af}(\mathcal{H}_1) \partial_u \mathcal{H}_1 \right) - \partial_u \left[\partial_x \left(\frac{1}{H} \mathrm{Af}\left(\int^u \mathrm{Af}(\mathcal{H}_1) \right) \right) \partial_u \mathcal{H}_1 \right] \right\} \, . \end{aligned} \tag{B.49}$$

Here we have used that $H = H(x)$ and that the integrals over u are performed at fixed x, as dictated by the method of near identity transformations. In this expression, the total u derivative gives a vanishing contribution thanks to periodicity of the functions, while in the first term we can recognize a total derivative plus an average free term:

$$\frac{1}{H} \mathrm{Af}(\mathcal{H}_1) \partial_u \mathcal{H}_1 = \frac{1}{2H} \partial_u (\mathcal{H}_1^2) - \frac{1}{H} \langle \mathcal{H}_1 \rangle \partial_u \mathcal{H}_1 \, . \tag{B.50}$$

These terms give a vanishing contribution to the average, therefore we find

$$G_2 = 0 \, . \tag{B.51}$$

Moving to higher orders, we will have to handle increasingly complex expressions. In the end, also from the point of view of near identity transformations, the simplest route might be showing that the long time-scale part of x and u are still conjugate variables described by an Hamiltonian.

Despite these results, the perturbative expansion can fail due to resonances between modes of the two orbits. In practice, if the orbits have commensurable periods, then some terms of the expansion can be enhanced by inverse powers of the expansion parameters, usually called small divisors [72]. As already remarked, in the analysis of this work we have discarded such cases by performing the averages over the two orbits independently. Generally however, small divisors will appear at high enough orders in perturbation theory. Their presence will determine a loss of validity of the predictions that we have obtained and a corresponding non trivial evolution of the semimajor axes on time-scales that are parametrically larger than those characterizing the effects described by lower orders in perturbation theory. Instead if the system is studied close to a resonance, then the standard perturbation theory will stop working already from low orders and resonant behavior will appear early on in the evolution of the system. As an example, the effects of a resonance on the evolution of a hierarchical triple were studied in [321].

B.5 BACKREACTION AND DEVIATIONS FORM ADIABATICITY

As outlined in Appendix B.3, the averaging procedure can be conveniently defined as an average over the values of the slowly evolving part of the mean anomaly, then expressed in terms of the eccentric anomaly as in Eq. (B.46). Here we apply this procedure to the three body problem.

B.5.1 *Long-timescale and short-timescale Lagrangians*

In order to describe corrections to the adiabatic approximation, we should return to the Lagrangian (4.12). It will be useful to split it between a Newtonian term and a perturbation:

$$\mathcal{L} = \frac{1}{2}v^2 + \frac{G_N m}{r} + \mathcal{R} = \frac{(G_N m)^2}{2L^2} + L\dot{u} + G\dot{\omega} + H\dot{\Omega} + \mathcal{R} \,, \tag{B.52}$$

where the perturbing function $\mathcal{R}$ contains the kinetic term of the center of mass as well as any other term beyond the Newtonian two-body interaction of the inner binary. For simplicity we have divided the Lagrangian by μ, a notation that we will use throughout the appendices. The second equality expresses the Newtonian part as a first-order Lagrangian depending on the (osculating) contact elements of the orbit, which are defined in Section 5.1. It can be checked that this Lagrangian indeed gives the Lagrange Planetary Equations presented in Section B.1 and which are usually presented in the Hamiltonian formalism, see also [304, 322].

As mentioned in App. B.1, the conjugate momenta are given by

$$L = \sqrt{G_N m a} \,, \; G = L\sqrt{1-e^2} \,, \; H = G\cos\iota \,. \tag{B.53}$$

Given Eq. (B.52), we can perform near-identity transformations for all the contact elements as outlined in Appendix B.3. Once we determine the fast oscillating terms in each of the near-identity transformations to a desired order in both e, the ratio of the semimajor axes, and v, the velocity of the bodies, we can simply plug back in the Lagrangian these values, obtaining a classical effective Lagrangian. Once we expand the resulting equations of motion in e and v, we will recover the slow dynamics, as in Eq. (B.33). To simplify even more the Lagrangian, we can take its average over the mean anomaly (keeping fixed slow variables, as prescribed by the near-identity transformation). Doing so will not alter the equations of motion for the slowly evolving variables, since the average commutes with variations of the Lagrangian with respect to these variables. It will simply remove the average-free part of the effective Lagrangian, which does not carry dynamical information. As remarked in the previous Appendix, following this procedure at leading order will lead to the long time-scale Kozai-Lidov dynamics.

More explicitly, considering first the evolution of L and u alone, we employ the following two near-identity transformations:

$$\begin{aligned} u &= u_L + u_S(L_L, u_L) \,, \\ L &= L_L + L_S(L_L, u_L) \,, \end{aligned} \tag{B.54}$$

where the subscripts $_L$, $_S$ stand for the long and short time-scale variable respectively, and the $_S$ variables are suppressed by powers of both ε and v. Using the LPE for u and L, see Appendix B.1, we determine the equations

for the short time-scale variables as discussed in Appendix B.3, finding at lowest order:

$$\begin{aligned}\partial_{u_L} L_S =& \frac{L_L^3}{(G_N m)^2}\mathrm{Af}_u\Big(\frac{\partial \mathcal{R}}{\partial u}\Big)\,,\\ \partial_{u_L} u_S =& -\frac{L_L^3}{(G_N m)^2}\Big[\mathrm{Af}_u\Big(\frac{\partial \mathcal{R}}{\partial L}\Big)+\frac{3(G_N m)^2 L_S}{L_L^4}\Big]\,,\end{aligned} \tag{B.55}$$

where, following the near-identity transformation procedure, at leading order we use $\dot{X} \simeq \frac{(G_N m)^2}{L_L^3}\partial_u X$, for a generic quantity X. These equations can be solved to give:

$$\begin{aligned}L_S =& \frac{L_L^3}{(G_N m)^2}\mathrm{Af}_{u_L}\Big(\int_0^{u_L} du\,\mathrm{Af}_u\Big(\frac{\partial \mathcal{R}}{\partial u}\Big)\Big)\,,\\ u_S =& -\mathrm{Af}_{u_L}\Big(\int_0^{u_L} du\,\Big[\frac{L_L^3}{(G_N m)^2}\mathrm{Af}_u\Big(\frac{\partial \mathcal{R}}{\partial L}\Big)+\frac{3L_S}{L_L}\Big]\Big)\,.\end{aligned} \tag{B.56}$$

This is analogous to what obtained in Eq. (B.38), (B.39), identifying $h_1 \mapsto -\partial\mathcal{R}/\partial L$, $g_1 \mapsto \partial\mathcal{R}/\partial u$ and $\chi_0 H'/H \mapsto -3L_S/L_L$. A difference with respect to the toy model presented in Appendix B.3 is the fact that now there are two small parameters, i.e. ratio of time-scales, which we have not factorized explicitly. For convenience, in the following we will indicate $\mathcal{F} = \mathrm{Af}_{u_L}(\mathcal{R})$ and, using that partial derivatives commute with the average over u_L, we will write $\mathrm{Af}_{u_L}(\frac{\partial \mathcal{R}}{\partial X}) = \frac{\partial \mathcal{F}}{\partial X}$.

Similar results will follow for the other contact elements and conjugate momenta. At leading order, the equations of motion for the short-timescale variables, obtained through the near-identity transformation, will read

$$\begin{aligned}\partial_{u_L} G_S &= \frac{L_L^3}{(G_N m)^2}\frac{\partial \mathcal{F}}{\partial \omega}\,, \quad \partial_{u_L}\omega_S = -\frac{L_L^3}{(G_N m)^2}\frac{\partial \mathcal{F}}{\partial G}\,,\\ \partial_{u_L} H_S &= \frac{L_L^3}{(G_N m)^2}\frac{\partial \mathcal{F}}{\partial \Omega}\,, \quad \partial_{u_L}\Omega_S = -\frac{L_L^3}{(G_N m)^2}\frac{\partial \mathcal{F}}{\partial H}\,.\end{aligned} \tag{B.57}$$

In order to convert derivatives with respect to canonical momenta in derivatives with respect to osculating contact elements, we use the formulas given in Appendix B.1. As prescribed by the near-identity transformation approach, the quantities entering the right-hand side of Eqs. (B.57) are expressed only in terms of the long timescale variables e_L, a_L..., as derived in Eq. (B.34). For simplicity, we will drop the $_L$ subscript from now on, so that it will be understood that all osculating contact elements appearing in the perturbing function do not include quickly oscillating parts. From the knowledge of these leading order oscillating parts of our variables, we can now compute the leading effect of back-reaction of the fast oscillations on the long time-scale dynamics, using the following procedure. At the level of the Lagrangian, we can substitute the near-identity transformations and write the expression in (B.52) as

$$\mathcal{L} = \mathcal{L}_L + \mathcal{L}_S\,, \tag{B.58}$$

the first term corresponding to the Lagrangian evaluated on the long time-scale variables only, while the second, $\mathcal{L}_S$, corresponding to the remaining part, then to be averaged. For instance we have:

$$L\dot{u} = (L_L + L_S)(\dot{u}_L + \dot{u}_S)\,. \tag{B.59}$$

When we take the average of this quantity, the mixed terms $L_L \dot{u}_S$ and $L_S \dot{u}_L$, being exactly average free, will not contribute. Therefore, we only need to keep track of the two contributions $L_L \dot{u}_L$ in $\mathcal{L}_L$ and $L_S \dot{u}_S$ in $\mathcal{L}_S$. This means that, up to unimportant average-free terms, the long-timescale and short-timescale parts of the Lagrangian read

$$\mathcal{L}_L = L_L \dot{u}_L + G_L \dot{\omega}_L + H_L \dot{\Omega}_L + \frac{(G_N m)^2}{2L_L^2} + \langle \mathcal{R} \rangle \,, \tag{B.60}$$

$$\mathcal{L}_S = \frac{(G_N m)^2}{2}\Big(\frac{1}{(L_L + L_S)^2} - \frac{1}{L_L^2}\Big) + L_S \dot{u}_S + G_S \dot{\omega}_S + H_S \dot{\Omega}_S + \sum_X \frac{\partial \mathcal{F}}{\partial X} X_S \,, \tag{B.61}$$

where X represent contact elements and conjugate momenta, and we have used the splitting $\mathcal{R} = \mathcal{F} + \langle \mathcal{R} \rangle$ together with the fact that $\partial \langle \mathcal{R} \rangle / \partial X \; X_S$ is an average-free quantity. In this equation, it is understood as before that $\mathcal{R}$ and $\mathcal{F}$ are evaluated on the long-timescale variables only. We will approximate the first term as $3/2((G_N m)^2/L_L^4) L_S^2$, since L_S is average-free and all the short timescale variables are suppressed either by e^2 or by v^2.

Plugging back the equations of motion as well as the solutions obtained as in (B.56), we find

$$\begin{aligned} \mathcal{L}_S = &- \frac{L_L^3}{(G_N m)^2}\Big[\frac{\partial \mathcal{F}}{\partial u}\mathrm{Af}_{u_L}\Big(\int^{u_L} du \frac{\partial \mathcal{F}}{\partial L}\Big) + \frac{\partial \mathcal{F}}{\partial \omega}\mathrm{Af}_{u_L}\Big(\int_0^{u_L} du \frac{\partial \mathcal{F}}{\partial G}\Big) \\ &+ \frac{\partial \mathcal{F}}{\partial \Omega}\mathrm{Af}_{u_L}\Big(\int_0^{u_L} du \frac{\partial \mathcal{F}}{\partial H}\Big)\Big] + \frac{3L_L^2}{(G_N m)^2}\Big[\frac{\mathcal{F}^2}{2} - \mathcal{F}^2 - \frac{\partial \mathcal{F}}{\partial u}\mathrm{Af}_{u_L}\Big(\int^{u_L} du \mathcal{F}\Big)\Big] \,, \end{aligned} \tag{B.62}$$

where in the second line, the first term comes from the expansion of $1/(L_L + L_S)^2$, the second from substituting $L_S \dot{u}_S$ and the third from substituting the solution found for u_S in Eq. (B.56). In order to simplify further this expression, it is useful to consider its average over u_L. Considering the average allows us to perform integration by parts without having to deal with boundary terms, thanks to the fact that these are obtained subtracting the values of periodic functions at the endpoints $u_L = 0\,, 2\pi$. Moreover, we can simplify Eq. (B.62) using that $\langle \mathrm{Af}(M)\mathrm{Af}(N) \rangle = \langle \mathrm{Af}(M) N \rangle$ for any arbitrary functions M, N. Thus we find:

$$\begin{aligned} \langle \mathcal{L}_S \rangle = \frac{1}{2\pi}\int_0^{2\pi} du_L \Big[\frac{L_L^3}{(G_N m)^2}\Big(\frac{\partial \mathcal{F}}{\partial L}\mathcal{F} + \frac{\partial \mathcal{F}}{\partial G}\int_0^{u_L} du \frac{\partial \mathcal{F}}{\partial \omega} + \frac{\partial \mathcal{F}}{\partial H}\int_0^{u_L} du \frac{\partial \mathcal{F}}{\partial \Omega}\Big) \\ + \frac{3L_L^2}{2(G_N m)^2}\mathcal{F}^2\Big] \,. \end{aligned} \tag{B.63}$$

In the next Subsection, we will compute the quadrupolar post-Newtonian cross-terms given by this procedure. Using the same procedure, we can also obtain the so-called "quadrupole-squared" terms studied in [298], which are a purely Newtonian contribution of post-adiabatic corrections. As a proof-of-concept, we use our procedure to compute such terms in Appendix B.8 and show that they give back the exact same equations as the ones displayed in [298].

B.5.2 *1PN quadrupolar cross-terms*

We now need the expression of the perturbing function $\mathcal{R}$, which contains both post-Newtonian and quadrupolar terms. They are obtained by expand-

ing the Lagrangian (4.12) in the center-of-mass frame. In doing so, one obtains a quadrupolar part $\mathcal{R}_{\rm quad}$, of order ε^2 within our power-counting rules, and a post-Newtonian part $\mathcal{R}_{\rm 1PN}$. The latter is itself composed of two terms $\mathcal{R}_{\rm 1PN} = \mathcal{R}_{v^2} + \mathcal{R}_{v^2\varepsilon^{1/2}}$ scaling differently: the first one is of order v^2 and corresponds to the usual EIH Lagrangian in the center-of-mass frame, and the second one is linear in $V_{\rm CM}$ and thus of order $v^2\varepsilon^{1/2}$. Higher-order terms in the ε expansion can be safely neglected for the precision we aim to.

Now, a helpful simplification comes directly from using the relative coordinates of the inner binary center of mass as explained in Section 5.2.2. Indeed, once we express the Lagrangian in terms of these coordinates, the contributions of order $v^2\varepsilon^{1/2}$ precisely cancel each other, so that they do not contribute to cross-terms. This cancellation is made possible by the fact that the change of reference frame mixes different orders of the expansion. Had we defined the contact elements through the three-body center-of-mass relative coordinates in Eq. (B.72), this term would have been non-zero and would have lead to cumbersome formulas which would have prevented us from performing the matching of the inner binary to a point-particle as we did in Section 5.2.4. We now give the expression of the different perturbing functions in terms of r and v as

$$\begin{aligned}\mathcal{R}_{v^2} &= \frac{1}{8}v^4(1-3\nu) + \frac{G_N m}{2r}\left((3+\nu)v^2 + \nu(\mathbf{v}\cdot\mathbf{n})^2 - \frac{G_N m}{r}\right) ,\\ \mathcal{R}_{\rm quad} &= -\frac{1}{2}r^i r^j \partial_i\partial_j\tilde{\phi} ,\end{aligned} \tag{B.64}$$

where we recall that $\tilde{\phi}$ is an arbitrary external field, $X_A = m_A/m$ and $\nu = X_1X_2$. Substitution of near identity transformation in these expressions will produce 1PN quadrupolar terms, i.e of order $v^2\varepsilon^2$. Using the expression of the osculating elements, we find

$$\begin{aligned}\mathcal{R}_{v^2} = \frac{G_N^2 m^2}{2a^2(1-e\cos\eta)^2}\Bigg[&\frac{1-3\nu}{4}(1+e\cos\eta)^2 + (3+\nu)(1+e\cos\eta)\\ &+\nu e^2\frac{\sin^2\eta}{1-e\cos\eta} - 1\Bigg] ,\end{aligned} \tag{B.65}$$

$$\begin{aligned}\mathcal{R}_{\rm quad} = -\frac{a^2}{2}&\left((\cos\eta - e)\,\alpha^i + \sqrt{1-e^2}\sin\eta\,\beta^i\right)\times\\ &\left((\cos\eta - e)\,\alpha^j + \sqrt{1-e^2}\sin\eta\,\beta^j\right)\partial_i\partial_j\tilde{\phi} .\end{aligned} \tag{B.66}$$

We can now compute the cross-terms following Eq. (B.63). The Lagrangian displayed in Eq. (B.63) contains cross-terms of order $v^2\varepsilon^2$. It also contains 2PN (v^4) and quadrupole-squared (ε^4) terms, which we will ignore since we limit ourselves to 1PN quadrupolar order. Cross-terms of order $v^2\varepsilon^2$ will not receive contributions from the term containing both H and Ω derivatives, since $\mathcal{R}_{v^2}$ does not depend on neither of these two variables. Computations are performed using the *Mathematica* software, giving the full expression of cross-terms in terms of osculating elements:

$$\begin{aligned}&\langle\mathcal{L}_S\rangle =\\ &\frac{G_N m a}{8(1-e^2+\sqrt{1-e^2})}\Big\{[2(9+7\sqrt{1-e^2}) + e^2(51+11\sqrt{1-e^2}) + 16e^4]\alpha^i\alpha^j\\ &+[2(7+9\sqrt{1-e^2}) + e^2(29+17\sqrt{1-e^2}) - 8e^4]\beta^i\beta^j\Big\}\partial_i\partial_j\tilde{\phi} .\end{aligned} \tag{B.67}$$

This term will be part of the effective Lagrangian for the long time-scale dynamics:

$$\mathcal{L}_{eff} = \langle \mathcal{L}_L \rangle + \langle \mathcal{L}_S \rangle = L_L \dot{u}_L + G_L \dot{\omega}_L + H_L \dot{\Omega}_L + \langle \mathcal{R} \rangle + \langle \mathcal{L}_S \rangle \,. \tag{B.68}$$

This effective Lagrangian contains the term $\mathcal{L}^{\leq v^2\varepsilon^{5/2}}_{\text{quad},12}$ computed in Section 5.2, but it also includes the terms at lower order in the multipole expansion.

Note that other cross-terms, besides those contained in $\langle \mathcal{L}_S \rangle$ and those coming from terms of order quadrupole-1PN in $\mathcal{R}$, might come in principle if the change of variables from the mean anomaly to the eccentric anomaly, discussed in Appendix B.3, was to differ with respect to the change of variables derived from the Kepler equation (B.42). As discussed, we change variable from the long time-scale part of the mean anomaly, to the long time-scale part of the eccentric anomaly, defined so as to be related to the former by the Kepler equation. This removes any potential cross term contribution due to the change of variables in the integration.

Finally, note also that to this order cross terms in the effective Lagrangian (B.68) will not get contributions due to subleading oscillating parts of the near-identity transformations (e.g. short time-scale functions with amplitude of order quadrupole-1PN, similar to $x^{(1)}_S$ in the notation of Appendix B.3), since these will be average free and will have to multiply terms of $\mathcal{L}_L$, leading to average-free contributions, if any.

B.6 LAGRANGE PLANETARY EQUATIONS BEYOND LEADING AVERAGING

As seen in the previous section, in order to derive the averaged Lagrangian to order $v^2\varepsilon^{5/2}$ we need to compute the averages following the method presented in Appendix B.3, the near-identity transformations. The LPE for the outer orbit are obtained in the very same way as described in App. B.1, replacing all inner quantities with outer ones with the only caveat of substituting the mass m with the sum $m_3 + \mathcal{E}$, with $\mathcal{E} = m - G_N m/2a$.

To follow through the computations outlined in App. B.5.1, two technical remarks are in order. First, in order to express derivatives of the quadrupolar Lagrangian in Eq. (B.66) with respect to angles, we use the following derivatives of the basis vectors:

$$\begin{aligned}
&\frac{\partial \alpha^i}{\partial \omega} = \beta^i \,, \quad \frac{\partial \beta^i}{\partial \omega} = -\alpha^i \,, \\
&\frac{\partial \alpha^i}{\partial \Omega} = -\cos\omega \sin\iota\, \gamma^i + \cos\iota\, \beta^i \,, \quad \frac{\partial \beta^i}{\partial \Omega} = -\cos\iota\, \alpha^i + \sin\iota \sin\omega\, \gamma^i \,, \\
&\frac{\partial \alpha^i}{\partial \iota} = \sin\omega\, \gamma^i \,, \quad \frac{\partial \beta^i}{\partial \iota} = \cos\omega\, \gamma^i \,.
\end{aligned} \tag{B.69}$$

Second, we will use (spatial) gauge-invariance in order to simplify the computations as much as possible. Indeed, we know that the Lagrangian is a rotation-invariant quantity. Thus, when carrying out the average of the first-order Lagrangian obtained by integrating out the high-energy modes, we use a particular gauge choice for the angles: $\iota = \pi/2$, $\omega = 0$. The resulting averaged Lagrangian will be gauge-invariant provided we express it only

in terms of the basis vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$. In this gauge, the derivatives written above simplify greatly:

$$\begin{aligned}
&\frac{\partial \alpha^i}{\partial \omega} = \beta^i \,, \quad \frac{\partial \beta^i}{\partial \omega} = -\alpha^i \,, \quad \frac{\partial \alpha^i}{\partial \Omega} = -\gamma^i \,, \quad \frac{\partial \beta^i}{\partial \Omega} = 0 \,, \\
&\frac{\partial \alpha^i}{\partial \iota} = 0 \,, \quad \frac{\partial \beta^i}{\partial \iota} = \gamma^i \,, \\
&\frac{\partial \mathcal{R}}{\partial L} = 2\sqrt{\frac{a}{Gm}}\frac{\partial \mathcal{R}}{\partial a} + \frac{1-e^2}{e\sqrt{Gma}}\frac{\partial \mathcal{R}}{\partial e} \,, \quad \frac{\partial \mathcal{R}}{\partial G} = -\frac{\sqrt{1-e^2}}{e\sqrt{Gma}}\frac{\partial \mathcal{R}}{\partial e} \,, \\
&\frac{\partial \mathcal{R}}{\partial H} = -\frac{1}{\sqrt{Gma(1-e^2)}}\frac{\partial \mathcal{R}}{\partial \iota} \,.
\end{aligned} \tag{B.70}$$

To give further support to the validity of this simplification we refer the reader to Appendix B.8, where we explicitly derive that the spurious dependence on angles contained in the derivatives (B.69) exactly cancel when averaging the high-energy modes in order to obtain the quadrupole-squared terms.

B.7 FROM THE THREE-BODY CENTER-OF-MASS FRAME TO THE INNER BINARY REST FRAME

In this Appendix we derive the explicit relation between the absolute coordinates $(\boldsymbol{y}_1, \boldsymbol{y}_2)$ of the inner binary and the relative distance $\boldsymbol{r}'$ defined in the inner binary rest frame, which is the natural quantity in terms of which one can express osculating elements, as discussed in Section 5.2.2. Let us begin by recalling the post-Newtonian definition of the center-of-mass of the relativistic two-body system composed by the inner binary, given to 1PN order by:

$$\begin{aligned}
E\mathbf{Y}_{\rm CM} &= E_1\mathbf{y}_1 + E_2\mathbf{y}_2 \,, \\
E = E_1 + E_2 \quad , \quad E_A &= m_A + \frac{1}{2}m_A v_A^2 - \frac{G_N m_1 m_2}{2r} \,,
\end{aligned} \tag{B.71}$$

for $A = 1, 2$. In the three-body rest frame R, the relation $\boldsymbol{r} = \boldsymbol{y}_1 - \boldsymbol{y}_2$ leads to

$$\mathbf{y}_1 = \mathbf{Y}_{\rm CM} + (X_2 + \delta)\mathbf{r} \,, \quad \mathbf{y}_2 = \mathbf{Y}_{\rm CM} + (-X_1 + \delta)\mathbf{r} \,, \tag{B.72}$$

where we have defined

$$\begin{aligned}
X_A = \frac{m_A}{m} \,, \quad m = m_1 + m_2 \,, \quad \mu = \frac{m_1 m_2}{m} \,, \quad \nu = \frac{\mu}{m} \,, \\
\delta = -\nu \mathbf{V}_{\rm CM}\cdot\mathbf{v} + \nu(X_1 - X_2)\left(\frac{v^2}{2} - \frac{G_N m}{2r}\right) .
\end{aligned} \tag{B.73}$$

However, the relative distance $\boldsymbol{r}$ in the three-body rest frame R cannot be expressed in terms of the *intrinsic* contact elements using Eq. (5.6). Rather, it will be the relative distance in the inner binary rest frame R′, $\boldsymbol{r}' = \boldsymbol{y}_1' - \boldsymbol{y}_2'$, and the respective momentum to be related to these convenient variables by Eq. (5.6). For this reason, we want to express the relative coordinates in the rest frame R in terms of those defined in R′. Concretely, around a time t_0, the two coordinate systems are related by the following transformation:

$$\begin{pmatrix} t' - t_0 \\ \boldsymbol{y}' \end{pmatrix} = \mathcal{B}\begin{pmatrix} t - t_0 \\ \boldsymbol{y} - \mathbf{Y}_{\rm CM}(t_0) \end{pmatrix} \,, \tag{B.74}$$

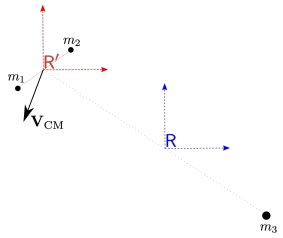

Figure B.1: Illustration of the change of referential (B.74): the rest frame of the inner binary R′ is obtained from the total center-of-mass frame R by translating it by $\boldsymbol{X}_{\rm CM}$ and boosting it by $\boldsymbol{V}_{\rm CM}$.

where $\mathcal{B}$ represents a Lorentz boost of velocity $V_{\rm CM}(t_0)$. To 1PN order, the coordinates of the two inner bodies $\boldsymbol{y}'_A$ (with $A = 1, 2$) in the frame R′ are related to those of the three-body rest frame R by:

$$\begin{aligned}\boldsymbol{y}'_A(t') =& \boldsymbol{y}_A(t') - \boldsymbol{Y}_{\rm CM} - \boldsymbol{V}_{\rm CM}(t' - t_0) \\ &+ \boldsymbol{V}_{\rm CM}\cdot\left(\boldsymbol{y}_A(t') - \boldsymbol{Y}_{\rm CM}\right)\left(\boldsymbol{v}_A(t') - \frac{\boldsymbol{V}_{\rm CM}}{2}\right),\end{aligned} \tag{B.75}$$

where in the last equation it is understood that $\boldsymbol{Y}_{\rm CM}$ and $\boldsymbol{V}_{\rm CM}$ are evaluated at t_0. Inserting this relation in the center-of-mass definition (B.71) to 1PN order, we obtain

$$E_1\boldsymbol{y}'_1 + E_2\boldsymbol{y}'_2 = \left(\boldsymbol{V}_{\rm CM}\cdot\boldsymbol{r}'\right)\boldsymbol{p}', \tag{B.76}$$

where we have substituted $\mu\boldsymbol{v}' \simeq \boldsymbol{p}'$. From this, we find that the relation between absolute coordinates $\boldsymbol{y}'_A$ and relative coordinates $\boldsymbol{r}'$, $\boldsymbol{v}'$ in rest frame of the inner binary is:

$$\begin{aligned}\boldsymbol{y}'_1 &= \left(X_2 + \delta\right)\boldsymbol{r}' + \left(\boldsymbol{V}_{\rm CM}\cdot\boldsymbol{r}'\right)\boldsymbol{p}'/m, \\ \boldsymbol{y}'_2 &= \left(-X_1 + \delta\right)\boldsymbol{r}' + \left(\boldsymbol{V}_{\rm CM}\cdot\boldsymbol{r}'\right)\boldsymbol{p}'/m,\end{aligned} \tag{B.77}$$

where δ is the 1PN quantity already defined in Eq. (B.73) (the use of primed or unprimed quantities in δ does not matter since the difference would be of 2PN order), and it is understood that $\boldsymbol{y}'_1$, $\boldsymbol{y}'_2$, $\boldsymbol{r}'$ and $\boldsymbol{v}'$ are evaluated at the same time t'.

This result allows to express the three-body frame coordinates $\boldsymbol{y}_A$ in terms of the relative coordinates in the rest frame of the inner binary, $\boldsymbol{r}'$, $\boldsymbol{p}'$ [2], which can be in turn expressed in terms of *intrinsic* contact elements using Eq. (5.6)[3]. We obtain:

$$\begin{aligned}\boldsymbol{y}_1 =& \mathbf{Y}_{\rm CM} + \mathbf{V}_{\rm CM}(t - t_0) + (X_2 + \delta)\mathbf{r}' \\ &+ X_2\left(\boldsymbol{V}_{\rm CM}\cdot\boldsymbol{r}'\right)\left[\left(X_1 - X_2\right)\frac{\boldsymbol{p}'}{\mu} - \frac{\boldsymbol{V}_{\rm CM}}{2}\right], \\ \boldsymbol{y}_2 =& \mathbf{Y}_{\rm CM} + \mathbf{V}_{\rm CM}(t - t_0) + (-X_1 + \delta)\mathbf{r}' \\ &- X_1\left(\boldsymbol{V}_{\rm CM}\cdot\boldsymbol{r}'\right)\left[\left(X_1 - X_2\right)\frac{\boldsymbol{p}'}{\mu} - \frac{\boldsymbol{V}_{\rm CM}}{2}\right],\end{aligned} \tag{B.78}$$

where $\boldsymbol{r}'$ and $\boldsymbol{p}'$ are evaluated at the time t, while $\mathbf{Y}_{\rm CM}$ and $\mathbf{V}_{\rm CM}$ are evaluated at t_0.

2 Note that $\boldsymbol{p}'$ is defined through the relative velocity $\boldsymbol{v}'$, the t' derivative of $\boldsymbol{r}'$.

3 Although Eq. (5.6) could be used to define osculating elements whatever the frame, now we will apply it to the relative distance and momentum in the binary rest frame, which we are calling $\boldsymbol{r}'$ and $\boldsymbol{p}'$.

In particular, this result implies the following relation between $\boldsymbol{r} = \boldsymbol{y}_1 - \boldsymbol{y}_2$ and $\boldsymbol{r}'$:

$$\boldsymbol{r} = \boldsymbol{r}' - (\boldsymbol{V}_{\rm CM}\cdot\boldsymbol{r}')\left[\frac{\boldsymbol{V}_{\rm CM}}{2} + (X_2 - X_1)\frac{\boldsymbol{p}'}{\mu}\right] . \tag{B.79}$$

The final step is just to evaluate the above expressions (B.78) and their time-derivatives at $t = t_0$ and to express $\boldsymbol{r}'$, $\boldsymbol{p}'$ in terms of the contact elements as in Eq. (5.6). To avoid clutter, in the main text and the rest of the Appendices we will suppress the primed label on $\boldsymbol{r}'$, $\boldsymbol{p}'$.

B.8 QUADRUPOLE-SQUARED TERMS

In this Appendix we use the methodology developed in Appendix B.5 to compute the so-called quadrupole-squared terms presented in [298]. These purely Newtonian contributions to the evolution of a hierarchical system come from deviations to the adiabatic approximation as well as backreaction of quickly oscillating modes when averaging quadrupolar terms. The magnitude of these contributions can be greater than octupole order terms so that they could induce interesting deviations to the Kozai-Lidov mechanism.

As stated in Appendix B.5, corrections to the leading order averaging could come at the level of either the inner binary or the outer binary orbital motion, because of short-timescale fluctuations of the form $X = X_L + X_S$ where X is any osculating element or canonical momentum of the inner or outer binary. More precisely, there are four kind of short-timescale fluctuations of osculating elements to consider: (*i*) fluctuations of the inner planetary elements on an inner binary timescale; (*ii*) fluctuations of the outer planetary elements on an inner binary timescale; (*iii*) fluctuations of the inner planetary elements on an outer binary timescale; (*iv*) fluctuations of the outer planetary elements on an outer binary timescale. In all these cases, a generalization of the methodology developed in Appendix B.5.1 shows that the cross-terms in L_S (B.63) scale as (see in particular Eq. (B.61) and Eq. (B.57)):

$$(i)\quad \frac{\mathcal{F}^2}{L}\frac{L^3}{G_N^2 m^2} \sim \mathcal{F}^2\frac{a}{G_N m} , \tag{B.80}$$

$$(ii)\quad \frac{\mathcal{F}^2}{L_3}\frac{L^3}{G_N^2 m^2} \sim \mathcal{F}^2\frac{a}{G_N m}\varepsilon^{1/2} , \tag{B.81}$$

$$(iii)\quad \frac{\mathcal{F}^2}{L}\frac{L_3^3}{G_N^2 m^2} \sim \mathcal{F}^2\frac{a}{G_N m}\varepsilon^{-3/2} , \tag{B.82}$$

$$(iv)\quad \frac{\mathcal{F}^2}{L_3}\frac{L_3^3}{G_N^2 m^2} \sim \mathcal{F}^2\frac{a}{G_N m}\varepsilon^{-1} . \tag{B.83}$$

This makes it clear that, in general, case *(iii)* gives the largest contribution of the cross-terms (see also the interesting discussion in Appendix B of Ref. [298]). This scaling is perfectly valid for the quadrupole-squared terms which we want to compute in this Appendix; instead, in the previous case of quadrupole-1PN cross-terms, there is an additional suppression on the outer binary timescale which explains that the dominant contribution to cross-terms comes from case *(i)* as computed in Appendix B.5. This additional suppression comes from the fact that the PN terms in the function $\mathcal{F}$ come with an additional ε^2 multiplicative factor for cases (*iii*) and (*iv*) — i.e. when averaging over the outer binary timescale. To see this,

remark that the 1PN perturbing function can be schematically written as: $\mathcal{R}_{1\text{PN}} = \mathcal{R}_{v^2} + \mathcal{R}_{v^2\varepsilon^{1/2}} + \mathcal{R}_{v^2\varepsilon} + \mathcal{R}_{v^2\varepsilon^{3/2}} + \mathcal{R}_{v^2\varepsilon^2}$. This splitting, as already emphasized in Appendix B.5.2, comes from introducing the center-of-mass coordinates (5.10) in the EIH Lagrangian (4.11). With our center-of-mass choice, both $\mathcal{R}_{v^2\varepsilon^{1/2}}$, $\mathcal{R}_{v^2\varepsilon}$ and $\mathcal{R}_{v^2\varepsilon^{3/2}}$ vanish after averaging on the inner binary timescale. Thus, the only term featuring a non-trivial dependence on the outer binary timescale is $\mathcal{R}_{v^2\varepsilon^2}$, since $\mathcal{R}_{v^2}$ is just the standard EIH Lagrangian for the inner binary and does not depend on the outer binary period. This proves the additional ε^2 suppression of outer binary averages, justifying the use of case *(i)* in Appendix B.5.

Let us now derive the quadrupole-squared terms using the scaling *(iii)*. Since we are only interested in deviations from adiabaticity in the *outer* average, we only perform leading order averaging for the inner binary timescale. Having removed the dependence of the Lagrangian on the mean anomaly of the inner orbit, the semimajor axis of the inner orbit will remain constant. This means that this variable will not have short oscillations on time-scales of the order of the period of the outer binary. The other variables instead will have modes that evolve during the period of the outer orbit and modes that evolve on much longer time-scales. Therefore we will generally write $X = X_{\bar{L}} + X_{\tilde{S}}$, indicating a near identity transformation having the period of the outer binary as reference time-scale. As indicated by the scaling *(iii)*, we only consider such a decomposition for X being an inner osculating element, because fluctuations of outer osculating elements are suppressed by an additional $\varepsilon^{1/2}$ factor.

At this point we can already check that the quadrupole-square cross terms Lagrangian $\mathcal{L}_{\tilde{S}}$ scales as

$$\frac{\mathcal{L}_{\tilde{S}}}{\mathcal{L}_{\bar{L}}} \sim \frac{n_3}{n} \sim \sqrt{\frac{Ma^3}{ma_3^3}} \,. \tag{B.84}$$

Thus, despite the quadrupole-squared terms are formally a $\varepsilon^{3/2}$ perturbation to the quadrupole (i.e between octupole and hexadecapole), they could be enhanced to a greater magnitude than the octupole if the ratio M/m is large.

To proceed with the computation, we will indicate the quadrupolar part of the perturbing function in Eq. (5.27) to Newtonian order as:

$$\begin{aligned} \mathcal{R}_{Q3} &= \frac{3G_N m_3}{2\mu} Q^{ij} \frac{N_i N_j}{R^3} \,, \\ Q^{ij} &= \frac{\mu a^2}{2}\left((1+4e^2)\alpha^i\alpha^j + (1-e^2)\beta^i\beta^j - \frac{2+3e^2}{3}\delta^{ij}\right) , \end{aligned} \tag{B.85}$$

and its average free part, taken with respect to the long time-scale part of the outer mean anomaly $u_{3\,\tilde{L}}$, as:

$$\mathcal{F}_{Q3} = \mathcal{R}_{Q3} - \frac{1}{2\pi}\int_0^{2\pi} du_{3\,\tilde{L}}\mathcal{R}_{Q3} \,, \tag{B.86}$$

where we recall that $\mathbf{R}$ is the distance vector of the outer orbit, $R = |\boldsymbol{R}|$ and $\boldsymbol{N} = \boldsymbol{R}/R$. With this notation, using the equations of motion and expanding the perturbing function, we obtain:

$$\begin{aligned} \mathcal{L}_{\tilde{S}} &= L_{\tilde{S}}\dot{u}_{\tilde{S}} + G_{\tilde{S}}\dot{\omega}_{\tilde{S}} + H_{\tilde{S}}\dot{\Omega}_{\tilde{S}} + \sum_X \frac{\partial \mathcal{F}_{Q3}}{\partial X} X_{\tilde{S}} \\ &= \mathrm{Af}\Big(\int^{u_{3\bar{L}}} ds \frac{\partial \mathcal{F}_{Q3}}{\partial G}\Big)\frac{\partial \mathcal{F}_{Q3}}{\partial \omega} + \mathrm{Af}\Big(\int^{u_{3\bar{L}}} ds \frac{\partial \mathcal{F}_{Q3}}{\partial H}\Big)\frac{\partial \mathcal{F}_{Q3}}{\partial \Omega} \,. \end{aligned} \tag{B.87}$$

Then, performing the averages with respect to the long time-scale part of the outer mean anomaly, we find:

$$\langle \mathcal{L}_S \rangle = -\frac{9G_N^2 m_3^2}{4\mu^2}\left(\frac{\partial Q^{ij}}{\partial \omega}\frac{\partial Q^{kl}}{\partial G} + \frac{\partial Q^{ij}}{\partial \Omega}\frac{\partial Q^{kl}}{\partial H}\right)\mathcal{A}_{ij;kl} \equiv \mathcal{B}^{ij;kl}\mathcal{A}_{ij;kl}\,, \tag{B.88}$$

where the tensor $\mathcal{A}_{ij;kl}$ is a sort of variance given by

$$\mathcal{A}_{ij;kl} = \left\langle \left[\frac{N_i N_j}{R^3} - \left\langle \frac{N_i N_j}{R^3}\right\rangle\right] \times \int \mathrm{d}t \left[\frac{N_k N_l}{R^3} - \left\langle \frac{N_k N_l}{R^3}\right\rangle\right]\right\rangle\,. \tag{B.89}$$

Note that the choice of a constant in the integration does not matter since it multiplies a zero-mean term. $\mathcal{A}_{ij;kl}$ is symmetric in the (i,j) as well as in the (k,l) indices, and it is antisymmetric by exchange of the pairs (i,j) and (k,l) (by an integration by parts). There now just remains to compute the derivatives in Eq. (B.88). We will use the derivatives given in Appendix B.1, Eqs. (B.15) and (B.69). These derivatives generically depend on the angles Ω, ω and ι. However, such angles cannot remain in the final result for $\langle \mathcal{L}_S \rangle$, since the contraction in the spatial indices needs to transform correctly under a rotation. This is analogous to the 'background field method' in EFTs: if we integrate out some fluctuating field in some given gauge, then the computational steps can be gauge-dependent but the final result should be gauge-invariant since it is expressed as a (gauge-invariant) long-wavelength Lagrangian. The cancellation of the angular dependence will be a very non-trivial check of our computation. In the more complicated computation of the PN quadrupolar cross-terms, instead, we have chosen to reverse the argument and choose a particular gauge (i.e, a particular value for Ω, ω and ι) to simplify the calculations, as explained at the end of Appendix B.1.

We now go on for the final computation. In the $\mathcal{B}^{ij;kl}$ tensor, the dependence on the angle ι nicely factors out:

$$\begin{aligned}
&\frac{1}{\sin\iota}\frac{\partial Q^{kl}}{\partial \iota}\left(\cos\iota\frac{\partial Q^{ij}}{\partial \omega} - \frac{\partial Q^{ij}}{\partial \Omega}\right) = \\
&\frac{\mu^2 a^4}{4}\Big[(1+4e^2)\cos\omega\left(\alpha^i\gamma^j + \mathrm{sym}\right) - (1-e^2)\sin\omega\left(\beta^i\gamma^j + \mathrm{sym}\right)\Big] \\
&\times\Big[(1+4e^2)\sin\omega\left(\alpha^k\gamma^l + \mathrm{sym}\right) + (1-e^2)\cos\omega\left(\beta^k\gamma^l + \mathrm{sym}\right)\Big]\,.
\end{aligned} \tag{B.90}$$

However, there seems to remain an additional dependence on ω which we do not expect. This spurious dependence can be removed completely by using the fact that $\mathcal{A}_{ij;kl}$ is antisymmetric under the exchange of the pairs (i,j) and (k,l). Taking the anti-symmetrization of the above expression, we are led to

$$\begin{aligned}
&\frac{\mu^2 a^4(1-e^2)(1+4e^2)}{8}\times \\
&\Big((\alpha^i\gamma^j + \mathrm{sym})(\gamma^k\beta^l + \mathrm{sym}) - (\gamma^i\beta^j + \mathrm{sym})(\alpha^k\gamma^l + \mathrm{sym})\Big)\,,
\end{aligned} \tag{B.91}$$

so that all angular dependence drops out. Putting all together, we find the final expression for $\mathcal{B}_{ij;kl}$:

$$\begin{aligned}\mathcal{B}_{ij;kl} =& \frac{9G_N^2 m_3^2 a^4}{32\sqrt{G_N m a(1-e^2)}} \times \\ & \Big[-20e^2(1-e^2)(\alpha^i\beta^j+\text{sym})(4\alpha^k\alpha^l-\beta^k\beta^l-\delta^{kl})+(1-e^2)(1+4e^2)\times \\ & \Big((\alpha^i\gamma^j+\text{sym})(\beta^k\gamma^l+\text{sym})-(\beta^i\gamma^j+\text{sym})(\alpha^k\gamma^l+\text{sym})\Big)\Big]\,,\end{aligned} \tag{B.92}$$

while the other tensor $\mathcal{A}_{ij;kl}$ is given by

$$\begin{aligned}\mathcal{A}^{ij;kl} =& \frac{g_1(e_3)}{\sqrt{G_N M a_3^9}}\Big[\alpha_3^k\alpha_3^l(\alpha_3^i\beta_3^j+\text{sym})-\alpha_3^i\alpha_3^j(\alpha_3^k\beta_3^l+\text{sym})\Big] \\ & + \frac{g_2(e_3)}{\sqrt{G_N M a_3^9}}\Big[\beta_3^i\beta_3^j(\alpha_3^k\beta_3^l+\text{sym})-\beta_3^k\beta_3^l(\alpha_3^i\beta_3^j+\text{sym})\Big]\,,\end{aligned} \tag{B.93}$$

where the two dimensionless functions of the outer eccentricity e_3 are given by

$$\begin{aligned}g_1(e_3) =& \frac{1}{48e_3^2(1-e_3^2)^{7/2}}\Big[4(1-\sqrt{1-e_3^2})+e_3^2(-8+9\sqrt{1-e_3^2}) \\ & +e_3^4(4+5\sqrt{1-e_3^2})\Big]\,,\end{aligned} \tag{B.94}$$

$$g_2(e_3) = \frac{1}{48e_3^2(1-e_3^2)^{5/2}}\Big[4(-1+\sqrt{1-e_3^2})+e_3^2(4+\sqrt{1-e_3^2})\Big]\,. \tag{B.95}$$

Plugging the expressions of the basis vectors in terms of osculating angles, we have checked that our formula (B.88) exactly recovers the LPE with quadrupole-squared terms derived in [298].

B.9 FROM CONTACT ELEMENTS TO ORBITAL ELEMENTS

In this Appendix we explore the difference between contact elements and orbital elements, the two kinds of osculating elements that allow to describe the three-body system efficiently. As already remarked, the difference between these two sets of elements is of 1PN order, with the contact terms being particularly useful to repackage various PN corrections in the effective action. This analysis allows us to compare our results with some of the results of [268, 281] concerning the evolution of conserved quantities.

We will now write down explicitly the relation between these two sets of elements to 1PN order, following [323]. As discussed in Section 5.1, the key difference between the two sets of variables is the fact that the momentum is not simply proportional to the velocity in the PN expansion. Therefore it is useful to inspect the relation between $\mathbf{p}$ and $\mathbf{v}$ at 1PN:

$$\mathbf{p} = \mathbf{v} + \frac{\partial \mathcal{R}_{1\text{PN}}}{\partial \mathbf{v}} = \mathbf{v} + \left(\frac{1-3\nu}{2}v^2 + \frac{G_N m(3+\nu)}{r}\right)\mathbf{v} + \frac{G_N m\nu}{r}(\mathbf{n}\cdot\mathbf{v})\mathbf{n}\,, \tag{B.96}$$

where we recall that in all Appendices we use the convention of dividing the Lagrangian by the reduced mass μ, so that $\mathbf{p}$ and $\mathbf{v}$ have same dimension. In the 1PN term we could use indifferently $\mathbf{v}$ or $\mathbf{p}$ since the difference would

be of 2PN order. The easiest orbital elements to relate to contact elements are the angles $\tilde{\Omega}$ and $\tilde{\imath}$. Indeed, since the momentum $\mathbf{p}$ is still in the plane of the motion, the definition of $\tilde{\Omega}$ and $\tilde{\imath}$ do not get affected by the difference between momentum and velocity. Thus,

$$\tilde{\Omega} = \Omega \,, \tag{B.97}$$

$$\tilde{\imath} = \imath \,. \tag{B.98}$$

Let us start the non-trivial computations with the semimajor axis. We have the following definitions:

$$-\frac{G_N m}{2\tilde{a}} = \frac{v^2}{2} - \frac{G_N m}{r} \,, \tag{B.99}$$

$$-\frac{G_N m}{2a} = \frac{p^2}{2} - \frac{G_N m}{r} \,. \tag{B.100}$$

Combining Eqs.(B.96), (B.99) and (B.100) we get to $\tilde{a} = a + \Delta a$ with

$$\begin{aligned}\Delta a = -\frac{2G_N m}{(1-\tilde{e}\cos\tilde{\eta})^2}\Bigg[&\frac{1}{2}(1-3\nu)(1+\tilde{e}\cos\tilde{\eta})^2 \\ &+ (3+\nu)(1+\tilde{e}\cos\tilde{\eta}) + \frac{\nu\tilde{e}^2\sin^2\tilde{\eta}}{1-\tilde{e}\cos\tilde{\eta}}\Bigg] \,.\end{aligned} \tag{B.101}$$

In the above expression, using the contact or the osculating elements in the RHS makes no difference since we neglect terms of order 2PN (Δa is a 1PN quantity). Let us now focus on the eccentricity. We have:

$$|\mathbf{r}\times\mathbf{v}| = \sqrt{G_N m\tilde{a}(1-\tilde{e}^2)} \,, \tag{B.102}$$

$$|\mathbf{r}\times\mathbf{p}| = \sqrt{G_N m a(1-e^2)} \,. \tag{B.103}$$

$$\tag{B.104}$$

Splitting $\tilde{e} = e + \Delta e$ we get to

$$\begin{aligned}\Delta e =& \frac{G_N m(1-\tilde{e}^2)}{\tilde{a}(1-\tilde{e}\cos\tilde{\eta})^3}\times \\ &\Big[\cos\tilde{\eta}(1-\tilde{e}\cos\tilde{\eta})(-7+\nu-\tilde{e}(1-3\nu)\cos\tilde{\eta}) - \tilde{e}\nu\sin^2\tilde{\eta}\Big] \,.\end{aligned} \tag{B.105}$$

There finally remains to find the argument of perihelion ω and mean anomaly u. Let us denote by f the true anomaly, representing the angle of the object along its trajectory on the ellipse, measured from perihelion. Then one has $f + \omega = \tilde{f} + \tilde{\omega}$ since this corresponds to the true physical angle of the object, and should not depend on whether we use tilde quantities or not. Thus, $\Delta\omega = -\Delta f$. To find Δf, one can use the definition of the radius vector:

$$r = \frac{a(1-e^2)}{1+e\cos f} = \frac{\tilde{a}(1-\tilde{e}^2)}{1+\tilde{e}\cos\tilde{f}} \,. \tag{B.106}$$

This gives $\tilde{f} = f + \Delta f$ with

$$\begin{aligned}\Delta f =& \frac{G_N m\sqrt{1-\tilde{e}^2}\sin\tilde{\eta}}{2\tilde{a}\tilde{e}(1-\tilde{e}\cos\tilde{\eta})^3}\times \\ &\Big[14-\tilde{e}^2-2\nu+5\tilde{e}^2\nu-6\tilde{e}(2+\nu)\cos\tilde{\eta}-\tilde{e}^2(1-3\nu)\cos 2\tilde{\eta}\Big] \,.\end{aligned} \tag{B.107}$$

Finally, to find Δu (defined by $\tilde{u} = u + \Delta u$) one can use again the definition of $r = a(1 - e\cos\eta)$ together with $\eta - e\sin\eta = u$ to find

$$\begin{aligned}\Delta u = &-\frac{G_N m}{8\tilde{a}\tilde{e}(1-\tilde{e}\cos\tilde{\eta})^3}\Big[2\big(-6\tilde{e}^4+\tilde{e}^2(1-15\nu)-4(7-\nu)\big)\sin\tilde{\eta}\\ &+2\tilde{e}\big(6(2+\nu)+2\tilde{e}^2(7+2\nu)-\tilde{e}^4(1-3\nu)\big)\sin 2\tilde{\eta}\\ &+2\tilde{e}^2\big(1-3\nu-\tilde{e}^2(6+4\nu)\big)\sin 3\tilde{\eta}-\tilde{e}^5(1-3\nu)\sin 4\tilde{\eta}\Big]\,. \end{aligned} \tag{B.108}$$

To sum up our results, here are all the modifications to the osculating elements:

$$\begin{aligned}\Delta a = &-\frac{2G_N m}{(1-\tilde{e}\cos\tilde{\eta})^2}\times\\ &\left[\frac{1}{2}(1-3\nu)(1+\tilde{e}\cos\tilde{\eta})^2+(3+\nu)(1+\tilde{e}\cos\tilde{\eta})+\frac{\nu\tilde{e}^2\sin^2\tilde{\eta}}{1-\tilde{e}\cos\tilde{\eta}}\right], \end{aligned} \tag{B.109}$$

$$\begin{aligned}\Delta e = &\frac{G_N m(1-\tilde{e}^2)}{\tilde{a}(1-\tilde{e}\cos\tilde{\eta})^3}\times\\ &\Big[\cos\tilde{\eta}(1-\tilde{e}\cos\tilde{\eta})(-7+\nu-\tilde{e}(1-3\nu)\cos\tilde{\eta})-\tilde{e}\nu\sin^2\tilde{\eta}\Big], \end{aligned} \tag{B.110}$$

$$\begin{aligned}\Delta\omega = &-\frac{G_N m\sqrt{1-\tilde{e}^2}\sin\tilde{\eta}}{2\tilde{a}\tilde{e}(1-\tilde{e}\cos\tilde{\eta})^3}\times\\ &\Big[14-\tilde{e}^2-2\nu+5\tilde{e}^2\nu-6\tilde{e}(2+\nu)\cos\tilde{\eta}-\tilde{e}^2(1-3\nu)\cos 2\tilde{\eta}\Big], \end{aligned} \tag{B.111}$$

$$\begin{aligned}\Delta u = &-\frac{G_N m}{8\tilde{a}\tilde{e}(1-\tilde{e}\cos\tilde{\eta})^3}\Big[2\big(-6\tilde{e}^4+\tilde{e}^2(1-15\nu)-4(7-\nu)\big)\sin\tilde{\eta}\\ &+2\tilde{e}\big(6(2+\nu)+2\tilde{e}^2(7+2\nu)-\tilde{e}^4(1-3\nu)\big)\sin 2\tilde{\eta}\\ &+2\tilde{e}^2\big(1-3\nu-\tilde{e}^2(6+4\nu)\big)\sin 3\tilde{\eta}-\tilde{e}^5(1-3\nu)\sin 4\tilde{\eta}\Big], \end{aligned} \tag{B.112}$$

$$\Delta\Omega = \Delta\iota = 0\,, \tag{B.113}$$

$$\Delta\boldsymbol{\alpha} = \boldsymbol{\beta}\Delta\omega\,, \quad \Delta\boldsymbol{\beta} = -\boldsymbol{\alpha}\Delta\omega\,. \tag{B.114}$$

These formulas relate the instantaneous values of these two sets of elements. However, it is more interesting to analyze the difference in the orbit-averaged elements, denoted with an $_L$ subscript. Using our previous splitting between long-timescale and short-timescale variables (see Appendix B.5), one can write e.g for the eccentricity $\tilde{e} = e + \Delta e = e_L + e_S + \Delta e$ with $\langle e_S\rangle = 0$. Thus, we can identify:

$$\tilde{e}_L = e_L + \langle\Delta e\rangle\,, \quad \tilde{e}_S = e_S + \Delta e - \langle\Delta e\rangle\,, \tag{B.115}$$

and similarly for the other osculating elements. This gives the final relation for the orbit-averaged contact and orbital elements:

$$\begin{aligned}\tilde{a}_L &= a_L + G_N m\left[9-\frac{16}{\sqrt{1-\tilde{e}_L^2}}-\nu\left(5-\frac{6}{\sqrt{1-\tilde{e}_L^2}}\right)\right],\\ \tilde{e}_L &= e_L - \frac{G_N m\tilde{e}_L}{\tilde{a}_L}(8-3\nu)\frac{\sqrt{1-\tilde{e}_L^2}}{1+\sqrt{1-\tilde{e}_L^2}}\,,\\ \tilde{\omega}_L &= \omega_L\,, \quad \tilde{\Omega}_L = \Omega_L\,, \quad \tilde{\iota}_L = \iota_L\,, \quad \tilde{u}_L = u_L\,. \end{aligned} \tag{B.116}$$

We can now describe the main physical effects stemming from this post-Newtonian shift of the osculating elements:

- The first and most important point to notice is that the shifts written in Eq. (B.116) will stay small at *any* moment in time. In other words, they cannot accumulate a small effect over long timescales to get an important effect[4], like it happens for quadrupolar and post-Newtonian perturbations in the LPE; rather, the two sets of osculating elements will differ by a quantity which is of post-Newtonian order at any time. Thus, replacing contact elements by orbital ones will not make a qualitative difference concerning the long-timescale evolution of the binary system.

- Plugging the shifts $\tilde{e}_L = e_L - \langle \Delta e \rangle$ and $\tilde{a}_L = a_L - \langle \Delta a \rangle$ implied by Eq. (B.116) in the kinetic term of the Lagrangian (B.52) shifts the canonical momenta G and H as

$$\frac{\Delta G}{\tilde{G}} = \frac{\Delta H}{\tilde{H}} = \frac{e\langle \Delta e \rangle}{1-\tilde{e}^2} - \frac{\langle \Delta a \rangle}{2\tilde{a}} = \frac{G_N m}{2\tilde{a}}(7-\nu) \,. \tag{B.117}$$

  In other words, this replaces the effective Newtonian spin of the binary system $\mathbf{J} = \mu\sqrt{G_N m \tilde{a}(1-\tilde{e}^2)}\boldsymbol{\gamma}$ with its 1PN counterpart given by

$$\mathbf{J}_{1\mathrm{PN}} = \mu\sqrt{G_N m a(1-e^2)}\tilde{\boldsymbol{\gamma}} = \mu\sqrt{G_N m \tilde{a}(1-\tilde{e}^2)}\boldsymbol{\gamma}\left(1 + \frac{G_N m}{2\tilde{a}}(7-\nu)\right) . \tag{B.118}$$

  On the other hand, the above expression corresponds to the 1PN expression of the conserved angular momentum that one can find in e.g. [300], averaged over one orbit of the binary system. This is a non-trivial check of the validity of our computation.

- The most interesting effect coming from these shifts of a and e is that the semimajor axis $\tilde{a}$ is not conserved in time. Indeed, it is the contact element a which is conserved in time, but it is related to the osculating $\tilde{a}$ with a formula involving the eccentricity, eq. (B.116). Since the eccentricity is itself allowed to vary, one should have a variation of $\tilde{a}$ over time at 1PN order. This effect has already been discussed in [268, 281] where it was derived using the 1PN equations of motion. However, our treatment makes clear the fact that such an effect cannot accumulate over a long timescale and give appreciable variation in the semimajor axis a as discussed before.

  On the other hand, a detailed comparison shows that our formula for the variation of $\tilde{a}$ and the one given in [268, 281] seem to be in disagreement. The source of this apparent incompatibility can be traced back to the fact that our averaging procedure is somewhat different than the one discussed in [268, 281]. Indeed, the following averages differ at 1PN quadrupolar order:

$$\frac{\mathrm{d}\langle a \rangle}{\mathrm{d}t} \neq \left\langle \frac{\mathrm{d}a}{\mathrm{d}t} \right\rangle \,. \tag{B.119}$$

  The l.h.s. of this equation corresponds to the quantity that we compute in this article. On the other hand, the variation computed in [268, 281] corresponds to the RHS of this equation, which by definition is equal

4 This is, as long as the elements are computed within the time interval in which our effective field theory is valid.

to $(a(T) - a(0))/T$. Thus, the results of [268, 281] concern the evolution of the *initial value* of osculating elements after one inner period, while we are interested in the evolution of the *mean value* of osculating elements. This difference also shows itself in the conservation of energy and angular momentum: while the work in [268, 281] proves exact conservation of *initial* PN energy and angular momentum, we are able to prove conservation of an *averaged* PN energy and angular momentum.

BIBLIOGRAPHY

[1] B. P. Abbott and others. GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. *Phys. Rev. X*, 9(3):031040, 2019. _eprint: 1811.12907.

[2] R. Abbott and others. GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. *Phys. Rev. X*, 11:021053, 2021. _eprint: 2010.14527.

[3] R. Abbott and others. GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run. November 2021. _eprint: 2111.03606.

[4] Edwin Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proc. Nat. Acad. Sci.*, 15:168–173, 1929.

[5] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020. [Erratum: Astron.Astrophys. 652, C4 (2021)].

[6] Irwin I. Shapiro. Fourth Test of General Relativity. *Phys. Rev. Lett.*, 13:789–791, 1964.

[7] I. I. Shapiro, M. E. Ash, R. P. Ingalls, W. B. Smith, D. B. Campbell, R. B. Dyce, R. F. Jurgens, and G. H. Pettengill. Fourth test of general relativity - new radar result. *Phys. Rev. Lett.*, 26:1132–1135, 1971.

[8] Matthias Bartelmann and Peter Schneider. Weak gravitational lensing. *Phys. Rept.*, 340:291–472, 2001. _eprint: astro-ph/9912508.

[9] J. O. Dickey and others. Lunar Laser Ranging: A Continuing Legacy of the Apollo Program. *Science*, 265:482–490, 1994.

[10] James G. Williams, Slava G. Turyshev, and Dale H. Boggs. Progress in lunar laser ranging tests of relativistic gravity. *Phys. Rev. Lett.*, 93:261101, 2004. _eprint: gr-qc/0411113.

[11] Ingrid H. Stairs. Testing General Relativity with Pulsar Timing. *Living Reviews in Relativity*, 6(1):5, December 2003. arXiv:astro-ph/0307536.

[12] M. Kramer, I. H. Stairs, R. N. Manchester, M. A. McLaughlin, A. G. Lyne, R. D. Ferdman, M. Burgay, D. R. Lorimer, A. Possenti, N. D'Amico, J. M. Sarkissian, G. B. Hobbs, J. E. Reynolds, P. C. C. Freire, and F. Camilo. Tests of general relativity from timing the double pulsar. *Science*, 314(5796):97–102, October 2006. arXiv:astro-ph/0609417.

[13] B. P. Abbott and others. Multi-messenger Observations of a Binary Neutron Star Merger. *Astrophys. J. Lett.*, 848(2):L12, 2017. _eprint: 1710.05833.

[14] B. P. Abbott et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Phys. Rev. Lett.*, 119(16):161101, 2017.

[15] Luc Blanchet. Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries. *Living Rev. Rel.*, 17:2, 2014.

[16] Emanuele Berti, Vitor Cardoso, and Andrei O. Starinets. Quasinormal modes of black holes and black branes. *Class. Quant. Grav.*, 26:163001, 2009. _eprint: 0905.2975.

[17] Takashi Nakamura, Kenichi Oohara, and Yasufumi Kojima. General Relativistic Collapse to Black Holes and Gravitational Waves from Black Holes. *Progress of Theoretical Physics Supplement*, 90:1–218, January 1987.

[18] Thomas W. Baumgarte and Stuart L. Shapiro. On the Numerical Integration of Einstein's Field Equations. *Physical Review D*, 59(2):024007, December 1998. arXiv:gr-qc/9810065.

[19] Philippe Grandclément and Jérôme Novak. Spectral Methods for Numerical Relativity. *Living Reviews in Relativity*, 12(1):1, Jan 2009.

[20] Robert A. Eisenstein. Numerical Relativity and the Discovery of Gravitational Waves. *Annalen der Physik*, 531(8):1800348, 2019. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.201800348.

[21] K. Schwarzschild. On the gravitational field of a mass point according to Einstein's theory, May 1999. arXiv:physics/9905030.

[22] Werner Israel. Event horizons in static vacuum space-times. *Phys. Rev.*, 164:1776–1779, 1967.

[23] Jacob D. Bekenstein. Nonexistence of Baryon Number for Static Black Holes. *Physical Review D*, 5(6):1239–1246, March 1972. Publisher: American Physical Society.

[24] Jacob D. Bekenstein. Nonexistence of Baryon Number for Black Holes. II. *Phys. Rev. D*, 5:2403–2412, May 1972.

[25] D. C. Robinson. Uniqueness of the Kerr black hole. *Phys. Rev. Lett.*, 34:905–906, 1975.

[26] B. Carter. Axisymmetric Black Hole Has Only Two Degrees of Freedom. *Phys. Rev. Lett.*, 26:331–333, 1971.

[27] Jacob D. Bekenstein. Novel "no-scalar-hair" theorem for black holes. *Phys. Rev. D*, 51:R6608–R6611, Jun 1995.

[28] Markus Heusler. No hair theorems and black holes with hair. *Helv. Phys. Acta*, 69(4):501–528, 1996.

[29] Roger Penrose. Gravitational Collapse and Space-Time Singularities. *Physical Review Letters*, 14(3):57–59, January 1965. Publisher: American Physical Society.

[30] R. Penrose. "Golden Oldie": Gravitational Collapse: The Role of General Relativity. *General Relativity and Gravitation*, 34(7):1141–1165, July 2002.

[31] Kostas D. Kokkotas and Bernd G. Schmidt. Quasinormal modes of stars and black holes. *Living Rev. Rel.*, 2:2, 1999. _eprint: gr-qc/9909058.

[32] Emanuele Berti, Vitor Cardoso, and Clifford M. Will. On gravitational-wave spectroscopy of massive black holes with the space interferometer LISA. *Phys. Rev.*, D73:064030, 2006.

[33] R. Abbott et al. Tests of General Relativity with GWTC-3. 12 2021.

[34] Leor Barack and Curt Cutler. LISA capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy. *Physical Review D*, 69(8):082005, April 2004.

[35] Pau Amaro-Seoane et al. Low-frequency gravitational-wave science with eLISA/NGO. *Class. Quant. Grav.*, 29:124016, 2012.

[36] Pau Amaro-Seoane et al. eLISA/NGO: Astrophysics and cosmology in the gravitational-wave millihertz regime. *GW Notes*, 6:4–110, 2013.

[37] B Sathyaprakash, M Abernathy, F Acernese, P Ajith, B Allen, P Amaro-Seoane, N Andersson, S Aoudia, K Arun, P Astone, and et al. Scientific objectives of einstein telescope. *Classical and Quantum Gravity*, 29(12):124013, Jun 2012.

[38] M. Punturo and others. The Einstein Telescope: A third-generation gravitational wave observatory. *Class. Quant. Grav.*, 27:194002, 2010.

[39] David Reitze et al. Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO, July 2019. arXiv:1907.04833 [astro-ph, physics:gr-qc].

[40] Travis Robson, Neil Cornish, and Chang Liu. The construction and use of LISA sensitivity curves. *Classical and Quantum Gravity*, 36(10):105011, May 2019. arXiv:1803.01944 [astro-ph, physics:gr-qc].

[41] Steven L. Detweiler. Pulsar timing measurements and the search for gravitational waves. *Astrophys.J.*, 234:1100–1104, 1979.

[42] Gabriella Agazie et al. The NANOGrav 15-year Data Set: Search for Anisotropy in the Gravitational-Wave Background, June 2023. arXiv:2306.16221 [astro-ph, physics:gr-qc].

[43] Daniel J. Reardon et al. Search for an Isotropic Gravitational-wave Background with the Parkes Pulsar Timing Array. *The Astrophysical Journal*, 951:L6, July 2023. ADS Bibcode: 2023ApJ...951L...6R.

[44] Gabriella Agazie et al. The NANOGrav 15-year Data Set: Constraints on Supermassive Black Hole Binaries from the Gravitational Wave Background, June 2023. arXiv:2306.16220 [astro-ph, physics:gr-qc].

[45] Gabriella Agazie et al. The NANOGrav 15-year Data Set: Evidence for a Gravitational-Wave Background, June 2023. arXiv:2306.16213 [astro-ph, physics:gr-qc].

[46] Asher Berlin, Diego Blas, Raffaele Tito D'Agnolo, Sebastian A. R. Ellis, Roni Harnik, Yonatan Kahn, and Jan Schütte-Engel. Detecting High-Frequency Gravitational Waves with Microwave Cavities. *Physical Review D*, 105(11):116011, June 2022. arXiv:2112.11465 [astro-ph, physics:hep-ex, physics:hep-ph].

[47] Asher Berlin, Diego Blas, Raffaele Tito D'Agnolo, Sebastian A. R. Ellis, Roni Harnik, Yonatan Kahn, Jan Schütte-Engel, and Michael Wentzel. MAGO\,2.0: Electromagnetic Cavities as Mechanical Bars for Gravitational Waves. March 2023. _eprint: 2303.01518.

[48] Michael A. Fedderke, Peter W. Graham, Bruce Macintosh, and Surjeet Rajendran. Astrometric gravitational-wave detection via stellar interferometry. *Phys. Rev. D*, 106(2):023002, 2022. _eprint: 2204.07677.

[49] Kazunori Akiyama and others. First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole. *Astrophys. J. Lett.*, 875:L1, 2019. _eprint: 1906.11238.

[50] Collaboration Event Horizon Telescope. First Sagittarius A* Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole in the Center of the Milky Way. *The Astrophysical Journal Letters*, 930(2):L12, May 2022. Publisher: The American Astronomical Society.

[51] Michele Maggiore. Gravitational wave experiments and early universe cosmology. *Physics Reports*, 331(6):283–367, July 2000.

[52] Clifford M. Will. The Confrontation between General Relativity and Experiment. *Living Rev. Rel.*, 17:4, 2014.

[53] Emanuele et al. Berti. Testing General Relativity with Present and Future Astrophysical Observations. *Classical and Quantum Gravity*, 32(24):243001, December 2015. arXiv:1501.07274 [astro-ph, physics:gr-qc, physics:hep-ph, physics:hep-th].

[54] Leor Barack and others. Black holes, gravitational waves and fundamental physics: a roadmap. *Class. Quant. Grav.*, 36(14):143001, 2019. _eprint: 1806.05195.

[55] Pierre Auclair et al. Cosmology with the Laser Interferometer Space Antenna. 4 2022.

[56] Alberto Nicolis, Riccardo Penco, Federico Piazza, and Rachel A. Rosen. More on gapped Goldstones at finite density: More gapped Goldstones. *Journal of High Energy Physics*, 2013(11):55, November 2013. arXiv:1306.1240 [cond-mat, physics:hep-th, physics:quant-ph].

[57] Alberto Nicolis, Riccardo Penco, and Rachel A. Rosen. Relativistic Fluids, Superfluids, Solids and Supersolids from a Coset Construction. *Physical Review D*, 89(4):045002, February 2014. arXiv:1307.0517 [hep-th].

[58] Alberto Nicolis, Riccardo Penco, Federico Piazza, and Riccardo Rattazzi. Zoology of condensed matter: Framids, ordinary stuff, extraordinary stuff, January 2015. arXiv:1501.03845 [cond-mat, physics:gr-qc, physics:hep-th].

[59] Tomas Brauner et al. Snowmass White Paper: Effective Field Theories for Condensed Matter Systems, March 2022. Publication Title: arXiv e-prints ADS Bibcode: 2022arXiv220310110B.

[60] H. Georgi, H. R. Quinn, and S. Weinberg. Hierarchy of Interactions in Unified Gauge Theories. *Physical Review Letters*, 33(7):451–454, August 1974.

[61] Steven Weinberg. Phenomenological Lagrangians. *Physica A*, 96(1-2):327–340, 1979.

[62] J. Gasser and H. Leutwyler. Chiral Perturbation Theory to One Loop. *Annals Phys.*, 158:142, 1984.

[63] G. D'Ambrosio, G. F. Giudice, G. Isidori, and A. Strumia. Minimal Flavour Violation: an effective field theory approach. *Nuclear Physics B*, 645(1-2):155–187, November 2002. arXiv:hep-ph/0207036.

[64] Kaustubh Agashe, Roberto Contino, and Alex Pomarol. The Minimal Composite Higgs Model. *Nuclear Physics B*, 719(1-2):165–187, July 2005. arXiv:hep-ph/0412089.

[65] G. F. Giudice, C. Grojean, A. Pomarol, and R. Rattazzi. The Strongly-Interacting Light Higgs. *Journal of High Energy Physics*, 2007(06):045–045, June 2007. arXiv:hep-ph/0703164.

[66] Ben Gripaios, Alex Pomarol, Francesco Riva, and Javi Serra. Beyond the Minimal Composite Higgs Model. *Journal of High Energy Physics*, 2009(04):070–070, April 2009. arXiv:0902.1483 [hep-ph].

[67] Roberto Contino, Margherita Ghezzi, Christophe Grojean, Margarete Muhlleitner, and Michael Spira. Effective Lagrangian for a light Higgs-like scalar. *Journal of High Energy Physics*, 2013(7):35, July 2013. arXiv:1303.3876 [hep-ph].

[68] Matteo Cataneo, Simon Foreman, and Leonardo Senatore. Efficient exploration of cosmology dependence in the EFT of LSS. *Journal of Cosmology and Astroparticle Physics*, 2017(04):026–026, April 2017. arXiv:1606.03633 [astro-ph].

[69] R. P. Feynman and M. Gell-Mann. Theory of the Fermi Interaction. *Physical Review*, 109(1):193–198, January 1958.

[70] Walter D. Goldberger and Ira Z. Rothstein. An Effective Field Theory of Gravity for Extended Objects. *Physical Review D*, 73(10), May 2006. arXiv: hep-th/0409156.

[71] Rafael A. Porto. Post-newtonian corrections to the motion of spinning bodies in nonrelativistic general relativity. *Physical Review D*, 73(10), May 2006.

[72] James A. Murdock. *Perturbations: Theory and Methods*. Wiley, 1991.

[73] Marc H. Goroff and Augusto Sagnotti. The ultraviolet behavior of Einstein gravity. *Nuclear Physics B*, 266(3-4):709–736, March 1986.

[74] Steven Weinberg. Photons and gravitons in perturbation theory: Derivation of Maxwell's and Einstein's equations. *Phys. Rev.*, 138:B988–B1002, 1965.

[75] Robert M. Wald. Spin-2 Fields and General Covariance. *Phys. Rev. D*, 33:3613, 1986.

[76] R. P. Feynman. *Feynman lectures on gravitation*. 1996.

[77] Robert M. Wald. *General Relativity*. Chicago Univ. Pr., Chicago, USA, 1984.

[78] John F. Donoghue. General relativity as an effective field theory: The leading quantum corrections. *Phys. Rev. D*, 50:3874–3888, 1994.

[79] John F. Donoghue. The effective field theory treatment of quantum gravity. *AIP Conf. Proc.*, 1483(1):73–94, 2012.

[80] M. Ostrogradsky. Mémoires sur les équations différentielles, relatives au probléme des isopérimétres. *Mem. Acad. St. Petersbourg*, 6(4):385–517, 1850.

[81] Richard P. Woodard. Ostrogradsky's theorem on Hamiltonian instability. *Scholarpedia*, 10(8):32243, 2015.

[82] F. Rohrlich. The self-force and radiation reaction. *American Journal of Physics*, 68(12):1109–1112, December 2000.

[83] Samuel E. Gralla, Abraham I. Harte, and Robert M. Wald. A Rigorous Derivation of Electromagnetic Self-force. *Physical Review D*, 80(2):024031, July 2009. arXiv:0905.2391 [gr-qc, physics:hep-th, physics:physics].

[84] Austin Joyce, Bhuvnesh Jain, Justin Khoury, and Mark Trodden. Beyond the cosmological standard model. *Physics Reports*, 568:1–98, March 2015.

[85] Johannes Noller, Luca Santoni, Enrico Trincherini, and Leonardo G. Trombetta. Black Hole Ringdown as a Probe for Dark Energy. *Phys. Rev. D*, 101:084049, 2019.

[86] Lam Hui and Alberto Nicolis. No-Hair Theorem for the Galileon. *Phys. Rev. Lett.*, 110:241104, 2013.

[87] Eugeny Babichev, Christos Charmousis, and Antoine Lehébel. Black holes and stars in Horndeski theory. *Class. Quant. Grav.*, 33(15):154002, 2016.

[88] Eugeny Babichev, Christos Charmousis, and Antoine Lehébel. Asymptotically flat black holes in Horndeski theory and beyond. *JCAP*, 1704(04):027, 2017.

[89] Carlos A. R. Herdeiro and Eugen Radu. Asymptotically flat black holes with scalar hair: a review. *Int. J. Mod. Phys. D*, 24(09):1542014, 2015.

[90] Alexandre Yale and T. Padmanabhan. Structure of Lanczos-Lovelock Lagrangians in Critical Dimensions. *Gen. Rel. Grav.*, 43:1549–1570, 2011.

[91] Tsutomu Kobayashi, Masahide Yamaguchi, and Jun'ichi Yokoyama. Generalized G-inflation: Inflation with the most general second-order field equations. *Prog. Theor. Phys.*, 126:511–529, 2011.

[92] Thomas P. Sotiriou and Shuang-Yong Zhou. Black hole hair in generalized scalar-tensor gravity: An explicit example. *Phys.Rev.D*, 90:124063, 2014.

[93] Kartik Prabhu and Leo C. Stein. Black hole scalar charge from a topological horizon integral in Einstein-dilaton-Gauss-Bonnet gravity. *Phys. Rev.*, D98(2):021503, 2018.

[94] Justin L. Ripley and Frans Pretorius. Scalarized Black Hole dynamics in Einstein dilaton Gauss-Bonnet Gravity. *Phys. Rev.*, D101(4):044015, 2020.

[95] Justin L. Ripley and Frans Pretorius. Gravitational Collapse in Einstein dilaton Gauss-Bonnet Gravity. *Classical and Quantum Gravity*, 36(13):134001, July 2019. arXiv:1903.07543 [astro-ph, physics:gr-qc, physics:hep-th].

[96] Justin L. Ripley and Frans Pretorius. Hyperbolicity in Spherical Gravitational Collapse in a Horndeski Theory. *Phys. Rev.*, D99(8):084014, 2019.

[97] Helvi Witek, Leonardo Gualtieri, Paolo Pani, and Thomas P. Sotiriou. Black holes and binary mergers in scalar Gauss-Bonnet gravity: scalar field dynamics. *Phys. Rev. D*, 99(6):064035, 2019.

[98] Zhenwei Lyu, Nan Jiang, and Kent Yagi. Constraints on Einstein-dilation-Gauss-Bonnet gravity from black hole-neutron star gravitational wave events. *Phys. Rev. D*, 105(6):064001, 2022.

[99] Xian O. Camanho, Jose D. Edelstein, Juan Maldacena, and Alexander Zhiboedov. Causality Constraints on Corrections to the Graviton Three-Point Coupling. *JHEP*, 02:020, 2016.

[100] Stephon Alexander and Nicolas Yunes. Chern-Simons Modified General Relativity. *Phys. Rept.*, 480:1–55, 2009.

[101] Hector O. Silva, Jeremy Sakstein, Leonardo Gualtieri, Thomas P. Sotiriou, and Emanuele Berti. Spontaneous scalarization of black holes and compact stars from a Gauss-Bonnet coupling. *Phys. Rev. Lett.*, 120(13):131104, 2018.

[102] Caio F. B. Macedo, Jeremy Sakstein, Emanuele Berti, Leonardo Gualtieri, Hector O. Silva, and Thomas P. Sotiriou. Self-interactions and Spontaneous Black Hole Scalarization. *Phys. Rev. D*, 99(10):104041, 2019.

[103] Alexandru Dima, Enrico Barausse, Nicola Franchini, and Thomas P. Sotiriou. Spin-induced black hole spontaneous scalarization. *Phys. Rev. Lett.*, 125(23):231101, 2020.

[104] Richard John Eden, Peter V. Landshoff, David I. Olive, and John Charlton Polkinghorne. The analytic S-matrix. March 2017.

[105] Marcel Froissart. Asymptotic behavior and subtractions in the Mandelstam representation. *Phys. Rev.*, 123:1053–1057, 1961.

[106] A. Martin. Unitarity and high-energy behavior of scattering amplitudes. *Phys. Rev.*, 129:1432–1436, 1963.

[107] Alexander Zhiboedov. Notes on the analytic S-matrix (under construction).

[108] Allan Adams, Nima Arkani-Hamed, Sergei Dubovsky, Alberto Nicolis, and Riccardo Rattazzi. Causality, analyticity and an IR obstruction to UV completion. *JHEP*, 10:014, 2006.

[109] Alberto Nicolis, Riccardo Rattazzi, and Enrico Trincherini. Energy's and amplitudes' positivity. *JHEP*, 05:095, 2010. [Erratum: JHEP 11, 128 (2011)].

[110] Luca Vecchi. Causal versus analytic constraints on anomalous quartic gauge couplings. *JHEP*, 11:054, 2007.

[111] Brando Bellazzini, Luca Martucci, and Riccardo Torre. Symmetries, Sum Rules and Constraints on Effective Field Theories. *JHEP*, 09:100, 2014.

[112] Brando Bellazzini. Softness and amplitudes' positivity for spinning particles. *JHEP*, 02:034, 2017.

[113] Claudia de Rham, Scott Melville, Andrew J. Tolley, and Shuang-Yong Zhou. Positivity bounds for scalar field theories. *Phys. Rev. D*, 96(8):081702, 2017.

[114] Claudia de Rham, Scott Melville, Andrew J. Tolley, and Shuang-Yong Zhou. UV complete me: Positivity Bounds for Particles with Spin. *JHEP*, 03:011, 2018.

[115] Andrew J. Tolley, Zi-Yue Wang, and Shuang-Yong Zhou. New positivity bounds from full crossing symmetry. *JHEP*, 05:255, 2021.

[116] Nima Arkani-Hamed, Tzu-Chen Huang, and Yu-Tin Huang. The EFT-Hedron. *JHEP*, 05:259, 2021.

[117] Simon Caron-Huot and Vincent Van Duong. Extremal Effective Field Theories. *JHEP*, 05:280, 2021.

[118] Simon Caron-Huot, Dalimil Mazac, Leonardo Rastelli, and David Simmons-Duffin. Sharp Boundaries for the Swampland. *JHEP*, 07:110, 2021.

[119] Brando Bellazzini, Joan Elias Miró, Riccardo Rattazzi, Marc Riembau, and Francesco Riva. Positive moments for scattering amplitudes. *Phys. Rev. D*, 104(3):036006, 2021.

[120] Brando Bellazzini, Marc Riembau, and Francesco Riva. IR side of positivity bounds. *Phys. Rev. D*, 106(10):105008, 2022.

[121] Zvi Bern, Dimitrios Kosmopoulos, and Alexander Zhiboedov. Gravitational effective field theory islands, low-spin dominance, and the four-graviton amplitude. *J. Phys. A*, 54(34):344002, 2021.

[122] Zvi Bern, Enrico Herrmann, Dimitrios Kosmopoulos, and Radu Roiban. Effective Field Theory islands from perturbative and non-perturbative four-graviton amplitudes. *JHEP*, 01:113, 2023. _eprint: 2205.01655.

[123] Li-Yuan Chiang, Yu-tin Huang, Laurentiu Rodina, and He-Chen Weng. De-projecting the EFThedron. 4 2022.

[124] Tejas Dethe, Harmeet Gill, Dylan Green, Andrew Greensweight, Luis Gutierrez, Muyuan He, Toshiki Tajima, and Kevin Yang. Causality and dispersion relations. *American Journal of Physics*, 87(4):279–290, April 2019.

[125] Sebastian Mizera. Physics of the Analytic S-Matrix. June 2023. _eprint: 2306.05395.

[126] Mariana Carrillo Gonzalez, Claudia de Rham, Victor Pozsgay, and Andrew J. Tolley. Causal effective field theories. *Phys. Rev. D*, 106(10):105018, 2022. _eprint: 2207.03491.

[127] Simon Caron-Huot, Yue-Zhou Li, Julio Parra-Martinez, and David Simmons-Duffin. Causality constraints on corrections to Einstein gravity. 1 2022.

[128] Dong-Yu Hong, Zhuo-Hui Wang, and Shuang-Yong Zhou. Causality bounds on scalar-tensor EFTs. 4 2023.

[129] Brando Bellazzini, Clifford Cheung, and Grant N. Remmen. Quantum Gravity Constraints from Unitarity and Analyticity. *Phys. Rev. D*, 93(6):064076, 2016.

[130] C. Beadle, G. Isabella, D. Perrone, S. Ricossa, F. Riva, and F. Serra. Dispersion relations without forward limits – in preparation.

[131] Thomas P. Sotiriou and Shuang-Yong Zhou. Black hole hair in generalized scalar-tensor gravity. *Phys. Rev. Lett.*, 112:251102, 2014.

[132] Gregory Walter Horndeski. Second-order scalar-tensor field equations in a four-dimensional space. *Int. J. Theor. Phys.*, 10:363–384, 1974.

[133] C. Deffayet, Xian Gao, D. A. Steer, and G. Zahariade. From k-essence to generalised Galileons. *Phys. Rev.*, D84:064039, 2011.

[134] Mehdi Saravani and Thomas P. Sotiriou. Classification of shift-symmetric Horndeski theories and hairy black holes. *Phys. Rev.*, D99(12):124004, 2019.

[135] Kent Yagi, Leo C. Stein, and Nicolas Yunes. Challenging the Presence of Scalar Charge and Dipolar Radiation in Binary Pulsars. *Phys. Rev.*, D93(2):024010, 2016.

[136] Alberto Nicolis, Riccardo Rattazzi, and Enrico Trincherini. The Galileon as a local modification of gravity. *Phys. Rev.*, D79:064036, 2009.

[137] Jérôme Gleyzes, David Langlois, Federico Piazza, and Filippo Vernizzi. Healthy theories beyond Horndeski. *Phys. Rev. Lett.*, 114(21):211101, 2015.

[138] Jérôme Gleyzes, David Langlois, Federico Piazza, and Filippo Vernizzi. Exploring gravitational theories beyond Horndeski. *JCAP*, 1502:018, 2015.

[139] David Langlois and Karim Noui. Degenerate higher derivative theories beyond Horndeski: evading the Ostrogradski instability. *JCAP*, 1602(02):034, 2016.

[140] Marco Crisostomi, Matthew Hull, Kazuya Koyama, and Gianmassimo Tasinato. Horndeski: beyond, or not beyond? *JCAP*, 1603(03):038, 2016.

[141] Jibril Ben Achour, David Langlois, and Karim Noui. Degenerate higher order scalar-tensor theories beyond Horndeski and disformal transformations. *Phys. Rev.*, D93(12):124005, 2016.

[142] Jibril Ben Achour, Marco Crisostomi, Kazuya Koyama, David Langlois, Karim Noui, and Gianmassimo Tasinato. Degenerate higher order scalar-tensor theories beyond Horndeski up to cubic order. *JHEP*, 12:100, 2016.

[143] Jibril Ben Achour, Hongguang Liu, and Shinji Mukohyama. Hairy black holes in DHOST theories: Exploring disformal transformation as a solution-generating method. *JCAP*, 2002(02):023, 2020.

[144] Maria Okounkova, Leo C. Stein, Mark A. Scheel, and Daniel A. Hemberger. Numerical binary black hole mergers in dynamical Chern-Simons gravity: Scalar field. *Phys. Rev.*, D96(4):044020, 2017.

[145] Lodovico Capuano, Luca Santoni, and Enrico Barausse. Black hole hairs in scalar-tensor gravity (and lack thereof). 4 2023.

[146] Nicolas Yunes and Frans Pretorius. Dynamical Chern-Simons Modified Gravity. I. Spinning Black Holes in the Slow-Rotation Approximation. *Phys. Rev. D*, 79:084043, 2009.

[147] J. G. Lee, E. G. Adelberger, T. S. Cook, S. M. Fleischer, and B. R. Heckel. New Test of the Gravitational $1/r^2$ Law at Separations down to 52 μm. *Phys. Rev. Lett.*, 124(10):101101, 2020.

[148] I. T. Drummond and S. J. Hathrell. QED Vacuum Polarization in a Background Gravitational Field and Its Effect on the Velocity of Photons. *Phys. Rev. D*, 22:343, 1980.

[149] Manuel Accettulli Huber, Andreas Brandhuber, Stefano De Angelis, and Gabriele Travaglini. Eikonal phase matrix, deflection angle and time delay in effective field theories of gravity. *Phys. Rev. D*, 102(4):046014, 2020.

[150] Sijie Gao and Robert M. Wald. Theorems on gravitational time delay and related issues. *Class. Quant. Grav.*, 17:4999–5008, 2000.

[151] Jose D. Edelstein, Rajes Ghosh, Alok Laddha, and Sudipta Sarkar. Causality constraints in Quadratic Gravity. *JHEP*, 09:150, 2021.

[152] Garrett Goon and Kurt Hinterbichler. Superluminality, black holes and EFT. *JHEP*, 02:134, 2017.

[153] Claudia de Rham and Andrew J. Tolley. Speed of gravity. *Phys. Rev. D*, 101(6):063518, 2020.

[154] Claudia de Rham and Andrew J. Tolley. Causality in curved spacetimes: The speed of light and gravity. *Phys. Rev. D*, 102(8):084048, 2020.

[155] Brando Bellazzini, Giulia Isabella, Matthew Lewandowski, and Francesco Sgarlata. Gravitational Causality and the Self-Stress of Photons. 8 2021.

[156] Calvin Y. R. Chen, Claudia de Rham, Aoibheann Margalit, and Andrew J. Tolley. A cautionary case of casual causality. *JHEP*, 03:025, 2022.

[157] Claudia de Rham, Andrew J. Tolley, and Jun Zhang. Causality Constraints on Gravitational Effective Field Theories. 12 2021.

[158] D. Amati, M. Ciafaloni, and G. Veneziano. Planckian scattering beyond the semiclassical approximation. *Phys. Lett. B*, 289:87–91, 1992.

[159] Daniel N. Kabat and Miguel Ortiz. Eikonal quantum gravity and Planckian scattering. *Nucl. Phys. B*, 388:570–592, 1992.

[160] Ratindranath Akhoury, Ryo Saotome, and George Sterman. High Energy Scattering in Perturbative Quantum Gravity at Next to Leading Power. *Phys. Rev. D*, 103(6):064036, 2021.

[161] Brando Bellazzini, Matthew Lewandowski, and Javi Serra. Positivity of Amplitudes, Weak Gravity Conjecture, and Modified Gravity. *Phys. Rev. Lett.*, 123(25):251103, 2019.

[162] Zvi Bern, Andres Luna, Radu Roiban, Chia-Hsien Shen, and Mao Zeng. Spinning black hole binary dynamics, scattering amplitudes, and effective field theory. *Phys. Rev. D*, 104(6):065014, 2021. _eprint: 2005.03071.

[163] Murat Kologlu, Petr Kravchuk, David Simmons-Duffin, and Alexander Zhiboedov. Shocks, Superconvergence, and a Stringy Equivalence Principle. *JHEP*, 11:096, 2020.

[164] G. Veneziano. Large N bounds on, and compositeness limit of, gauge and gravitational interactions. *JHEP*, 06:051, 2002.

[165] Gia Dvali. Black Holes and Large N Species Solution to the Hierarchy Problem. *Fortsch. Phys.*, 58:528–536, 2010.

[166] Nima Arkani-Hamed, Lubos Motl, Alberto Nicolis, and Cumrun Vafa. The String landscape, black holes and gravity as the weakest force. *JHEP*, 06:060, 2007.

[167] Da Liu, Alex Pomarol, Riccardo Rattazzi, and Francesco Riva. Patterns of Strong Coupling for LHC Searches. *JHEP*, 11:141, 2016.

[168] Yuta Hamada, Toshifumi Noumi, and Gary Shiu. Weak Gravity Conjecture from Unitarity and Causality. *Phys. Rev. Lett.*, 123(5):051601, 2019.

[169] Subham Dutta Chowdhury, Abhijit Gadde, Tushar Gopalka, Indranil Halder, Lavneet Janagal, and Shiraz Minwalla. Classifying and constraining local four photon and four graviton S-matrices. *JHEP*, 02:114, 2020.

[170] Junsei Tokuda, Katsuki Aoki, and Shin'ichi Hirano. Gravitational positivity bounds. *JHEP*, 11:054, 2020.

[171] Lasma Alberte, Claudia de Rham, Sumer Jaitly, and Andrew J. Tolley. QED positivity bounds. *Phys. Rev. D*, 103(12):125020, 2021.

[172] Nima Arkani-Hamed, Yu-tin Huang, Jin-Yu Liu, and Grant N. Remmen. Causality, Unitarity, and the Weak Gravity Conjecture. 9 2021.

[173] Subham Dutta Chowdhury, Kausik Ghosh, Parthiv Haldar, Prashanth Raman, and Aninda Sinha. Crossing Symmetric Spinning S-matrix Bootstrap: EFT bounds. 12 2021.

[174] Li-Yuan Chiang, Yu-tin Huang, Wei Li, Laurentiu Rodina, and He-Chen Weng. (Non)-projective bounds on gravitational EFT. 1 2022.

[175] Murray Gell-Mann, M. L. Goldberger, and Walter E. Thirring. Use of causality conditions in quantum theory. *Phys. Rev.*, 95:1612–1627, 1954.

[176] Venkatesa Chandrasekaran, Grant N. Remmen, and Arvin Shahbazi-Moghaddam. Higher-Point Positivity. *JHEP*, 11:015, 2018.

[177] Solomon Endlich, Victor Gorbenko, Junwu Huang, and Leonardo Senatore. An effective formalism for testing extensions to General Relativity with gravitational waves. *JHEP*, 09:122, 2017.

[178] Vitor Cardoso, Masashi Kimura, Andrea Maselli, and Leonardo Senatore. Black Holes in an Effective Field Theory Extension of General Relativity. *Phys. Rev. Lett.*, 121(25):251105, 2018.

[179] Noah Sennett, Richard Brito, Alessandra Buonanno, Victor Gorbenko, and Leonardo Senatore. Gravitational-Wave Constraints on an Effective Field-Theory Extension of General Relativity. *Phys. Rev. D*, 102(4):044056, 2020.

[180] Maximilian Ruhdorfer, Javi Serra, and Andreas Weiler. Effective Field Theory of Gravity to All Orders. *JHEP*, 05:083, 2020.

[181] Johan Henriksson, Brian McPeak, Francesco Russo, and Alessandro Vichi. Bounding Violations of the Weak Gravity Conjecture. 3 2022.

[182] Zong-Zhe Du, Cen Zhang, and Shuang-Yong Zhou. Triple crossing positivity bounds for multi-field theories. *JHEP*, 12:115, 2021.

[183] Deeksha Chandorkar, Subham Dutta Chowdhury, Suman Kundu, and Shiraz Minwalla. Bounds on Regge growth of flat space scattering from bounds on chaos. *JHEP*, 05:143, 2021.

[184] Kelian Häring and Alexander Zhiboedov. Gravitational Regge bounds. 2 2022.

[185] Claudia de Rham, Scott Melville, Andrew J. Tolley, and Shuang-Yong Zhou. Massive Galileon Positivity Bounds. *JHEP*, 09:072, 2017.

[186] Brando Bellazzini, Francesco Riva, Javi Serra, and Francesco Sgarlata. Beyond Positivity Bounds and the Fate of Massive Gravity. *Phys. Rev. Lett.*, 120(16):161101, 2018.

[187] Claudia de Rham, Scott Melville, and Andrew J. Tolley. Improved Positivity Bounds and Massive Gravity. *JHEP*, 04:083, 2018.

[188] Pietro Baratella, Dominik Haslehner, Maximilian Ruhdorfer, Javi Serra, and Andreas Weiler. RG of GR from On-shell Amplitudes. 9 2021.

[189] S. Mignemi and N. R. Stewart. Charged black holes in effective string theory. *Phys. Rev. D*, 47:5259–5269, 1993.

[190] Claudia de Rham, Jérémie Francfort, and Jun Zhang. Black Hole Gravitational Waves in the Effective Field Theory of Gravity. *Phys. Rev. D*, 102(2):024079, 2020.

[191] Manuel Accettulli Huber, Andreas Brandhuber, Stefano De Angelis, and Gabriele Travaglini. From amplitudes to gravitational radiation with cubic interactions and tidal effects. *Phys. Rev. D*, 103(4):045015, 2021.

[192] Pablo A. Cano and Alejandro Ruipérez. Leading higher-derivative corrections to Kerr geometry. *JHEP*, 05:189, 2019. [Erratum: JHEP 03, 187 (2020)].

[193] Pablo A. Cano, Kwinten Fransen, Thomas Hertog, and Simon Maenaut. Gravitational ringing of rotating black holes in higher-derivative gravity. *Phys. Rev. D*, 105(2):024064, 2022.

[194] Hector O. Silva, Abhirup Ghosh, and Alessandra Buonanno. Black-hole ringdown as a probe of higher-curvature gravity theories. 5 2022.

[195] Remya Nair, Scott Perkins, Hector O. Silva, and Nicolás Yunes. Fundamental Physics Implications for Higher-Curvature Theories from Binary Black Hole Signals in the LIGO-Virgo Catalog GWTC-1. *Phys. Rev. Lett.*, 123(19):191101, 2019.

[196] Hector O. Silva, A. Miguel Holgado, Alejandro Cárdenas-Avendaño, and Nicolás Yunes. Astrophysical and theoretical physics implications from multimessenger neutron star observations. *Phys. Rev. Lett.*, 126(18):181101, 2021.

[197] Pablo A. Cano, Kwinten Fransen, and Thomas Hertog. Ringing of rotating black holes in higher-derivative gravity. *Phys. Rev. D*, 102(4):044047, 2020.

[198] Pratik Wagle, Nicolas Yunes, and Hector O. Silva. Quasinormal modes of slowly-rotating black holes in dynamical Chern-Simons gravity. 3 2021.

[199] Manu Srivastava, Yanbei Chen, and S. Shankaranarayanan. Analytical computation of quasinormal modes of slowly rotating black holes in dynamical Chern-Simons gravity. *Phys. Rev. D*, 104(6):064034, 2021.

[200] Andrea Antonelli, Alessandra Buonanno, Jan Steinhoff, Maarten van de Meent, and Justin Vines. Energetics of two-body Hamiltonians in post-Minkowskian gravity. *Phys. Rev.*, D99(10):104004, 2019.

[201] Clifford Cheung, Ira Z. Rothstein, and Mikhail P. Solon. From Scattering Amplitudes to Classical Potentials in the Post-Minkowskian Expansion. *Phys. Rev. Lett.*, 121(25):251101, 2018.

[202] Zvi Bern, Julio Parra-Martinez, Radu Roiban, Michael S. Ruf, Chia-Hsien Shen, Mikhail P. Solon, and Mao Zeng. Scattering amplitudes and conservative dynamics at the fourth post-Minkowskian order. *PoS*, LL2022:051, 2022.

[203] Zvi Bern, Julio Parra-Martinez, Radu Roiban, Michael S. Ruf, Chia-Hsien Shen, Mikhail P. Solon, and Mao Zeng. Scattering Amplitudes, the Tail Effect, and Conservative Binary Dynamics at O(G4). *Phys. Rev. Lett.*, 128(16):161103, 2022. _eprint: 2112.10750.

[204] Zvi Bern, Clifford Cheung, Radu Roiban, Chia-Hsien Shen, Mikhail P. Solon, and Mao Zeng. Black Hole Binary Dynamics from the Double Copy and Effective Theory. *JHEP*, 10:206, 2019.

[205] Paolo Di Vecchia, Carlo Heissenberg, Rodolfo Russo, and Gabriele Veneziano. The eikonal approach to gravitational scattering and radiation at \mathcalO (G^3). *JHEP*, 07:169, 2021. _eprint: 2104.03256.

[206] Paolo Di Vecchia, Carlo Heissenberg, Rodolfo Russo, and Gabriele Veneziano. Classical Gravitational Observables from the Eikonal Operator. October 2022. _eprint: 2210.12118.

[207] Paolo Di Vecchia, Carlo Heissenberg, Rodolfo Russo, and Gabriele Veneziano. The gravitational eikonal: from particle, string and brane collisions to black-hole encounters, June 2023. arXiv:2306.16488 [gr-qc, physics:hep-ph, physics:hep-th].

[208] Gustav Uhre Jakobsen, Gustav Mogull, Jan Plefka, Benjamin Sauer, and Yingxuan Xu. Conservative scattering of spinning black holes at fourth post-Minkowskian order. June 2023. _eprint: 2306.01714.

[209] David A. Kosower, Ben Maybee, and Donal O'Connell. Amplitudes, observables, and classical scattering. *Journal of High Energy Physics*, 2019(2):137, February 2019.

[210] Gregor Kälin and Rafael A. Porto. From Boundary Data to Bound States. *JHEP*, 01:072, 2020.

[211] Gregor Kälin and Rafael A. Porto. From Boundary Data to Bound States II: Scattering Angle to Dynamical Invariants (with Twist). *Journal of High Energy Physics*, 2020(2):120, February 2020. arXiv:1911.09130 [gr-qc, physics:hep-th].

[212] Tim Adamo and Riccardo Gonzo. Bethe-Salpeter equation for classical gravitational bound states. *Journal of High Energy Physics*, 2023(5):88, May 2023. arXiv:2212.13269 [gr-qc, physics:hep-ph, physics:hep-th].

[213] Rafael A. Porto. The Effective Field Theorist's Approach to Gravitational Dynamics. *Physics Reports*, 633:1–104, May 2016. arXiv:1601.04914.

[214] Jonathan Thornburg. The Capra Research Program for Modelling Extreme Mass Ratio Inspirals, February 2011. arXiv:1102.2857 [astro-ph, physics:gr-qc].

[215] Eric Poisson, Adam Pound, and Ian Vega. The motion of point particles in curved spacetime. *Living Reviews in Relativity*, 14(1):7, December 2011. arXiv:1102.0529 [gr-qc].

[216] Chris Kavanagh, Adrian C. Ottewill, and Barry Wardell. Analytical high-order post-Newtonian expansions for spinning extreme mass ratio binaries. *Phys. Rev. D*, 93(12):124038, 2016.

[217] A. Buonanno and T. Damour. Effective one-body approach to general relativistic two-body dynamics. *Phys.Rev.D*, 59:084006, 1999.

[218] Thibault Damour. Gravitational self-force in a Schwarzschild background and the effective one-body formalism. *Physical Review D*, 81(2):024017, January 2010.

[219] Thibault Damour, Piotr Jaranowski, and Gerhard Schäfer. Fourth post-newtonian effective one-body dynamics. *Physical Review D*, 91(8), Apr 2015.

[220] Thibault Damour. Gravitational scattering, post-Minkowskian approximation, and effective-one-body theory. *Physical Review D*, 94(10), Nov 2016.

[221] Walter D. Goldberger and Ira Z. Rothstein. Dissipative Effects in the Worldline Approach to Black Hole Dynamics. *Physical Review D*, 73(10):104030, May 2006. arXiv:hep-th/0511133.

[222] Rafael A. Porto, Andreas Ross, and Ira Z. Rothstein. Spin induced multipole moments for the gravitational wave amplitude from binary inspirals to 2.5 Post-Newtonian order. *Journal of Cosmology and Astroparticle Physics*, 2012(09):028–028, September 2012. arXiv:1203.2962 [astro-ph, physics:gr-qc, physics:hep-ph, physics:hep-th].

[223] Michele Levi and Jan Steinhoff. Spinning gravitating objects in the effective field theory in the post-Newtonian scheme. *JHEP*, 09:219, 2015.

[224] Zvi Bern, Julio Parra-Martinez, Radu Roiban, Eric Sawyer, and Chia-Hsien Shen. Leading Nonlinear Tidal Effects and Scattering Amplitudes. *JHEP*, 05:188, 2021. _eprint: 2010.08559.

[225] Rafael Aoude, Kays Haddad, and Andreas Helset. Tidal effects for spinning particles. *Journal of High Energy Physics*, 2021(3):97, March 2021. arXiv:2012.05256 [gr-qc, physics:hep-ph, physics:hep-th].

[226] Alessandra Buonanno, Mohammed Khalil, Donal O'Connell, Radu Roiban, Mikhail P. Solon, and Mao Zeng. Snowmass White Paper: Gravitational Waves and Scattering Amplitudes. In *Snowmass 2021*, April 2022. _eprint: 2204.05194.

[227] Smadar Naoz. The eccentric kozai-lidov effect and its applications. *Annual Review of Astronomy and Astrophysics*, 54(1):441–489, Sep 2016.

[228] J. L. Margot, P. Pravec, P. Taylor, B. Carry, and S. Jacobson. *Asteroid Systems: Binaries, Triples, and Pairs*, pages 355–374. 2015.

[229] David Nesvorný, David Vokrouhlický, William F. Bottke, Keith Noll, and Harold F. Levison. Observed binary fraction sets limits on the extent of collisional grinding in the kuiper belt. *The Astronomical Journal*, 141(5):159, Apr 2011.

[230] Andrei Tokovinin. From Binaries to Multiples. II. Hierarchical Multiplicity of F and G Dwarfs. *aj*, 147(4):87, April 2014.

[231] Alexander P. Stephan et al. Merging binaries in the galactic center: the eccentric kozai–lidov mechanism with stellar evolution. *Monthly Notices of the Royal Astronomical Society*, 460(4):3494–3504, May 2016.

[232] Smadar Naoz and Daniel C. Fabrycky. Mergers and obliquities in stellar triples. *The Astrophysical Journal*, 793(2):137, Sep 2014.

[233] A. A. Tokovinin. On the multiplicity of spectroscopic binary stars. *Astronomy Letters*, 23(6):727–730, November 1997.

[234] Heather A. Knutson et al. Friends of Hot Jupiters. I. A Radial Velocity Search for Massive, Long-period Companions to Close-in Gas Giant Planets. *The Astrophysical Journal*, 785(2):126, April 2014.

[235] Dimitri Veras and Eric B. Ford. Secular orbital dynamics of hierarchical two-planet systems. *The Astrophysical Journal*, 715(2):803–822, May 2010.

[236] Yanqin Wu, Norman W. Murray, and J. Michael Ramsahai. Hot jupiters in binary star systems. *The Astrophysical Journal*, 670(1):820–825, Nov 2007.

[237] Henry Ngo et al. Friends of hot jupiters. ii. no correspondence between hot-jupiter spin-orbit misalignment and the incidence of directly imaged stellar companions. *The Astrophysical Journal*, 800(2):138, Feb 2015.

[238] S. M. Ransom et al. A millisecond pulsar in a stellar triple system. *Nature*, 505(7484):520–524, jan 2014.

[239] Lisa Randall and Zhong-Zhi Xianyu. An Analytical Portrait of Binary Mergers in Hierarchical Triple Systems. *Astrophys. J.*, 864(2):134, 2018.

[240] Naoki Seto. Highly Eccentric Kozai Mechanism and Gravitational-Wave Observation for Neutron Star Binaries. *Phys. Rev. Lett.*, 111:061106, 2013.

[241] Fabio Antonini, Sourav Chatterjee, Carl L. Rodriguez, Meagan Morscher, Bharath Pattabiraman, Vicky Kalogera, and Frederic A. Rasio. BLACK HOLE MERGERS AND BLUE STRAGGLERS FROM HIERARCHICAL TRIPLES FORMED IN GLOBULAR CLUSTERS. *The Astrophysical Journal*, 816(2):65, Jan 2016.

[242] Nathan W. C. Leigh and Aaron M. Geller. The dynamical significance of triple star systems in star clusters. *Monthly Notices of the Royal Astronomical Society*, 432(3):2474–2479, 05 2013.

[243] R. P. Deane et al. A close-pair binary in a distant triple supermassive black-hole system. *Nature*, 511:57, 2014.

[244] Travis Robson, Neil J. Cornish, Nicola Tamanini, and Silvia Toonen. Detecting hierarchical stellar systems with lisa. *Physical Review D*, 98(6), Sep 2018.

[245] Kohei Inayoshi, Nicola Tamanini, Chiara Caprini, and Zoltán Haiman. Probing stellar binary black hole formation in galactic nuclei via the imprint of their center of mass acceleration on their gravitational wave signal. *Phys. Rev. D*, 96(6):063014, 2017.

[246] Joe M. Antognini, Benjamin J. Shappee, Todd A. Thompson, and Pau Amaro-Seoane. Rapid eccentricity oscillations and the mergers of compact objects in hierarchical triples. *Monthly Notices of the Royal Astronomical Society*, 439(1):1079–1091, 02 2014.

[247] David Merritt, Tal Alexander, Seppo Mikkola, and Clifford M. Will. Stellar dynamics of extreme-mass-ratio inspirals. *Phys. Rev. D*, 84:044024, Aug 2011.

[248] Bao-Minh Hoang, Smadar Naoz, Bence Kocsis, Frederic A. Rasio, and Fani Dosopoulou. Black hole mergers in galactic nuclei induced by the eccentric kozai–lidov effect. *The Astrophysical Journal*, 856(2):140, apr 2018.

[249] Smadar Naoz, Bence Kocsis, Abraham Loeb, and Nicolas Yunes. Resonant Post-Newtonian Eccentricity Excitation in Hierarchical Three-body Systems. *Astrophys. J.*, 773:187, 2013.

[250] Barnabás Deme, Bao-Minh Hoang, Smadar Naoz, and Bence Kocsis. Detecting Kozai-Lidov Imprints on the Gravitational Waves of Intermediate-mass Black Holes in Galactic Nuclei. *Astrophysical Journal*, 901(2):125, October 2020.

[251] David V. Martin and Amaury H. M. J. Triaud. Circumbinary planets – why they are so likely to transit. *Monthly Notices of the Royal Astronomical Society*, 449(1):781–793, mar 2015.

[252] David V. Martin and Amaury H. M. J. Triaud. Kozai–Lidov cycles towards the limit of circumbinary planets. *Monthly Notices of the Royal Astronomical Society: Letters*, 455(1):L46–L50, 10 2015.

[253] Tsevi Mazeh, Yuval Krymolowski, and Gady Rosenfeld. The high eccentricity of the planet orbiting 16 cygni b. *The Astrophysical Journal*, 477(2):L103–L106, mar 1997.

[254] T. Mazeh and J. Shaham. The orbital evolution of close triple systems: the binary eccentricity. *Astronomy and Astrophysics*, 77:145, August 1979.

[255] Bin Liu, Diego J. Muñoz, and Dong Lai. Suppression of extreme orbital evolution in triple systems with short-range forces. *Monthly Notices of the Royal Astronomical Society*, 447(1):747–764, Dec 2014.

[256] Gongjie Li, Smadar Naoz, Bence Kocsis, and Abraham Loeb. Implications of the eccentric Kozai-Lidov mechanism for stars surrounding supermassive black hole binaries. *Monthly Notices of the Royal Astronomical Society*, 451(2):1341–1349, August 2015.

[257] Gongjie Li, Smadar Naoz, Bence Kocsis, and Abraham Loeb. Eccentricity Growth and Orbit Flip in Near-coplanar Hierarchical Three-body Systems. *Astrophysical Journal*, 785(2):116, April 2014.

[258] M.L. Lidov. The evolution of orbits of artificial satellites of planets under the action of gravitational perturbations of external bodies. *Planetary and Space Science*, 9(10):719–759, 1962.

[259] Ryan M. O'Leary, Yohai Meiron, and Bence Kocsis. Dynamical Formation Signatures of Black Hole Binaries in the First Detected Mergers by LIGO. *The Astrophysical Journal*, 824(1):L12, June 2016.

[260] Miguel A. S. Martinez et al. Black Hole Mergers from Hierarchical Triples in Dense Star Clusters. *The Astrophysical Journal*, 903(1):67, November 2020.

[261] Smadar Naoz, Will M. Farr, Yoram Lithwick, Frederic A. Rasio, and Jean Teyssandier. Hot Jupiters from secular planet-planet interactions. *Nature*, 473(7346):187–189, May 2011.

[262] Smadar Naoz, Will M. Farr, Yoram Lithwick, Frederic A. Rasio, and Jean Teyssandier. Secular dynamics in hierarchical three-body systems. *Monthly Notices of the Royal Astronomical Society*, 431(3):2155–2171, 03 2013.

[263] Eric B. Ford, Boris Kozinsky, and Frederic A. Rasio. Secular evolution of hierarchical triple star systems. *The Astrophysical Journal*, 535(1):385–401, may 2000.

[264] Omer Blaes, Man Hoi Lee, and Aristotle Socrates. The Kozai Mechanism and the Evolution of Binary Supermassive Black Holes. *The Astrophysical Journal*, 578(2):775–786, oct 2002.

[265] Loren Hoffman and Abraham Loeb. Dynamics of triple black hole systems in hierarchically merging massive galaxies. *Monthly Notices of the Royal Astronomical Society*, 377(3):957–976, 05 2007.

[266] Francesco Biscani and Sante Carloni. A first-order secular theory for the post-Newtonian two-body problem with spin – I: The restricted case. *Mon. Not. Roy. Astron. Soc.*, 428:2295–2310, 2013.

[267] Bin Liu, Dong Lai, and Yi-Han Wang. Binary Mergers near a Supermassive Black Hole: Relativistic Effects in Triples. *Astrophysical Journal*, 883(1):L7, September 2019.

[268] Clifford M. Will. Post-Newtonian effects in N-body dynamics: conserved quantities in hierarchical triple systems. *Class. Quant. Grav.*, 31(24):244001, 2014.

[269] Yusuke Hagihara. *Celestial Mechanics, Volume II: Perturbation Theory*. MIT Press, 1972.

[270] Kaze W. K. Wong, Vishal Baibhav, and Emanuele Berti. Binary radial velocity measurements with space-based gravitational-wave detectors. *Monthly Notices of the Royal Astronomical Society*, 488(4):5665–5670, October 2019.

[271] Alejandro Torres-Orjuela, Xian Chen, Zhoujian Cao, Pau Amaro-Seoane, and Peng Peng. Detecting the beaming effect of gravitational waves. *Physical Review D*, 100(6), Sep 2019.

[272] Yohai Meiron, Bence Kocsis, and Abraham Loeb. Detecting triple systems with gravitational wave observations. *The Astrophysical Journal*, 834(2):200, Jan 2017.

[273] Hang Yu and Yanbei Chen. Direct determination of supermassive black hole properties with gravitational-wave radiation from surrounding stellar-mass black hole binaries. *Physical Review Letters*, 126(2), Jan 2021.

[274] Vitor Cardoso, Francisco Duque, and Gaurav Khanna. Gravitational tuning forks and hierarchical triple systems, 2021.

[275] Alejandro Torres-Orjuela, Xian Chen, and Pau Amaro-Seoane. Phase shift of gravitational waves induced by aberration. *Physical Review D*, 101(8), Apr 2020.

[276] Huan Yang, Béatrice Bonga, Zhipeng Peng, and Gongjie Li. Relativistic mean motion resonance. *Physical Review D*, 100(12), Dec 2019.

[277] Wen-Biao Han and Xian Chen. Testing general relativity using binary extreme-mass-ratio inspirals. *Monthly Notices of the Royal Astronomical Society: Letters*, 485(1):L29–L33, Mar 2019.

[278] Béatrice Bonga, Huan Yang, and Scott A. Hughes. Tidal Resonance in Extreme Mass-Ratio Inspirals. *Physical Review Letters*, 123(10), Sep 2019.

[279] Yun Fang and Qing-Guo Huang. Three body first post-newtonian effects on the secular dynamics of a compact binary near a spinning supermassive black hole. *Phys. Rev. D*, 102:104002, Nov 2020.

[280] Yun Fang, Xian Chen, and Qing-Guo Huang. Impact of a Spinning Supermassive Black Hole on the Orbit and Gravitational Waves of a Nearby Compact Binary. *Astrophysical Journal*, 887(2):210, December 2019.

[281] Clifford M. Will. Incorporating post-newtonian effects in n-body dynamics. *Phys. Rev. D*, 89:044043, Feb 2014.

[282] Clifford M. Will. New general relativistic contribution to mercury's perihelion advance. *Phys. Rev. Lett.*, 120:191101, May 2018.

[283] Halston Lim and Carl L. Rodriguez. Relativistic three-body effects in hierarchical triples. *Phys. Rev. D*, 102(6):064033, 2020.

[284] Cezary Migaszewski and Krzysztof Gozdziewski. Secular dynamics of coplanar, non-resonant planetary system under the general relativity and quadrupole moment perturbations. *Mon. Not. Roy. Astron. Soc.*, 392:2, 2009.

[285] Michele Levi. Effective Field Theories of Post-Newtonian Gravity: A comprehensive review. 2018.

[286] Michele Levi and Jan Steinhoff. Leading order finite size effects with spins for inspiralling compact binaries. *JHEP*, 06:059, 2015.

[287] Michele Levi. Next to Leading Order gravitational Spin-Orbit coupling in an Effective Field Theory approach. *Phys. Rev. D*, 82:104004, 2010.

[288] Michele Levi and Jan Steinhoff. Next-to-next-to-leading order gravitational spin-squared potential via the effective field theory for spinning objects in the post-Newtonian scheme. *JCAP*, 1601:008, 2016.

[289] Michèle Levi, Andrew J. McLeod, and Matthew von Hippel. NNNLO gravitational spin-orbit coupling at the quartic order in G. 2020.

[290] Andreas Ross. Multipole expansion at the level of the action. *Physical Review D*, 85(12), June 2012. arXiv: 1202.4750.

[291] Walter D. Goldberger and Andreas Ross. Gravitational radiative corrections from effective field theory. *Physical Review D*, 81(12), June 2010. arXiv: 0912.4254.

[292] Barak Kol and Michael Smolkin. Non-Relativistic Gravitation: From Newton to Einstein and Back. *Class. Quant. Grav.*, 25:145011, 2008.

[293] Barak Kol and Michael Smolkin. Einstein's action and the harmonic gauge in terms of Newtonian fields. *Phys. Rev. D*, 85:044029, 2012.

[294] A. Einstein, L. Infeld, and B. Hoffmann. The Gravitational Equations and the Problem of Motion. *The Annals of Mathematics*, 39(1):65, January 1938.

[295] M. H. L. Pryce. The Mass-Centre in the Restricted Theory of Relativity and Its Connexion with the Quantum Theory of Elementary Particles. *Proceedings of the Royal Society of London Series A*, 195(1040):62–81, November 1948.

[296] A.J Hanson and T Regge. The relativistic spherical top. *Annals of Physics*, 87(2):498 – 566, 1974.

[297] E. Corinaldesi and A. Papapetrou. Spinning Test-Particles in General Relativity. II. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 209(1097):259–268, 1951.

[298] Clifford M. Will. Higher-order effects in the dynamics of hierarchical triple systems: Quadrupole-squared terms. *Physical Review D*, 103(6):063003, March 2021.

[299] Liantong Luo, Boaz Katz, and Subo Dong. Double-averaging can fail to characterize the long-term evolution of lidov–kozai cycles and derivation of an analytical correction. *Monthly Notices of the Royal Astronomical Society*, 458(3):3060–3074, Mar 2016.

[300] Thibault Damour and Nathalie Deruelle. General relativistic celestial mechanics of binary systems. i. the post-newtonian motion. *Annales de l'I.H.P. Physique théorique*, 43(1):107–132, 1985.

[301] Adrien Kuntz, Federico Piazza, and Filippo Vernizzi. Effective field theory for gravitational radiation in scalar-tensor gravity. *JCAP*, 1905(05):052, 2019.

[302] G. Schäfer. Acceleration-dependent lagrangians in general relativity. *Physics Letters A*, 100(3):128 – 129, 1984.

[303] Thibault Damour and Gerhard Schäfer. Redefinition of position variables and the reduction of higher-order lagrangians. *Journal of Mathematical Physics*, 32(1):127–134, 1991.

[304] V. Brumberg. *Essential Relativistic Celestial Mechanics*. 01 1991.

[305] Charles W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W. H. Freeman, San Francisco, 1973.

[306] Jan Sanders, Ferdinand Verhulst, and James Murdock. *Averaging Methods in Nonlinear Dynamical Systems*. Springer, 01 2007.

[307] Walter D. Goldberger. Les Houches lectures on effective field theories and gravitational radiation. In *Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime Les Houches, France, July 31-August 25, 2006*, 2007.

[308] Yoshihide Kozai. Secular perturbations of asteroids with high inclination and eccentricity. *The Astronomical Journal*, 67:591–598, November 1962.

[309] Daniel Fabrycky and Scott Tremaine. Shrinking binary and planetary orbits by kozai cycles with tidal friction. *The Astrophysical Journal*, 669(2):1298–1315, nov 2007.

[310] N W C Leigh, A M Geller, B McKernan, K E S Ford, M-M Mac Low, J Bellovary, Z Haiman, W Lyra, J Samsing, M O'Dowd, B Kocsis, and S Endlich. On the rate of black hole binary mergers in galactic nuclei due to dynamical hardening. *Monthly Notices of the Royal Astronomical Society*, 474(4):5672–5683, dec 2017.