


Effective two-body approach to the hierarchical three-body problem: Quadrupole to 1PN

Adrien Kuntz^{✉,*}, Francesco Serra,[†] and Enrico Trincherini^{✉‡}
Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
and INFN Sezione di Pisa, Largo Pontecorvo 3, 56127 Pisa, Italy

 (Received 28 October 2022; accepted 12 December 2022; published 6 February 2023)

Many binary systems of interest for gravitational-wave astronomy are orbited by a third distant body, which can considerably alter their relativistic dynamics. Precision computations are needed to understand the interplay between relativistic corrections and three-body interactions. We use an effective field theory approach to derive the effective action describing the long-timescale dynamics of hierarchical three-body systems up to 1PN quadrupole order. At this level of approximation, computations are complicated by the backreaction of small oscillations on orbital timescales as well as deviations from the adiabatic approximation. We address these difficulties by eliminating the fast modes through the method of near-identity transformations. This allows us to compute for the first time the complete expression of the 1PN quadrupole cross terms in generic configurations of three-body systems. We numerically integrate the resulting equations of motion and show that 1PN quadrupole terms can affect the long term dynamics of relativistic three-body systems.

DOI: [10.1103/PhysRevD.107.044011](https://doi.org/10.1103/PhysRevD.107.044011)

I. INTRODUCTION

The evolution of binary systems can be significantly affected by relativistic effects as well as by the presence of a third celestial body. The interplay between these effects is especially relevant in view of recent observations of triple systems with pulsars [1] as well as future detections of relativistic three-body systems expected with the gravitational-wave (GW) spatial interferometer LISA [2,3]. Three-body effects have been studied to understand the merger rate of compact objects [4–10], the transits of exoplanets [11–13], or the evolution of triple star systems [14–17]. Triple systems often come in a hierarchical setting where one of the constituents is far away from the other two. These configurations are conveniently described in terms of an “inner binary” elliptic motion composed of the two closest objects, and an “outer binary” elliptic motion made of the inner binary itself and the outer object, as depicted schematically in Fig 1. The long-timescale evolution of these orbits presents many interesting features already at the Newtonian level [18], the most well-known example being the Kozai-Lidov (KL) oscillations of the eccentricity of the inner binary [19,20]. These effects are usually studied through the method of double averaging [18], in which quick oscillations on timescales of order of the orbital periods are eliminated from the dynamics, leaving

only information about evolution on timescales larger than the orbital periods. This method is known to work whenever the system is away from resonances; the latter makes the whole analysis much less straightforward (see, e.g., [21]).

When relativistic corrections to the Newtonian dynamics are considered, in the so-called post-Newtonian (PN) expansion, the long-timescale dynamics undergoes radical changes, even when the system remains in a hierarchical configuration. For instance, the periastron precession can suppress high eccentricity oscillations when the precession timescale is much shorter than the timescale characterizing KL oscillations [4,22–25]. Even if the system can be described by the inner and the outer orbit, studying the effects of relativistic corrections on its evolution over long timescales remains in general quite difficult especially due to effects at higher orders in the PN expansion [9], which will usually leave an imprint on the dynamics after a parametrically large time. It is thus important to compute three-body relativistic interactions to a high accuracy in order to evolve these systems on long timescales. One option is to approach the problem with a numerical relativistic three-body solver [26–30]; however, this method is heavily time consuming. On the other hand, analytic methods are sparse at best [9,24,25,31–34].

Motivated by these developments and by the ensuing difficulties, we recently introduced a new effective field theory (EFT) approach to the relativistic, hierarchical three-body problem [35]. The central idea of this method is to take advantage of the two small parameters characterizing

*adrien.kuntz@sns.it

†francesco.serra@sns.it

‡enrico.trincherini@sns.it

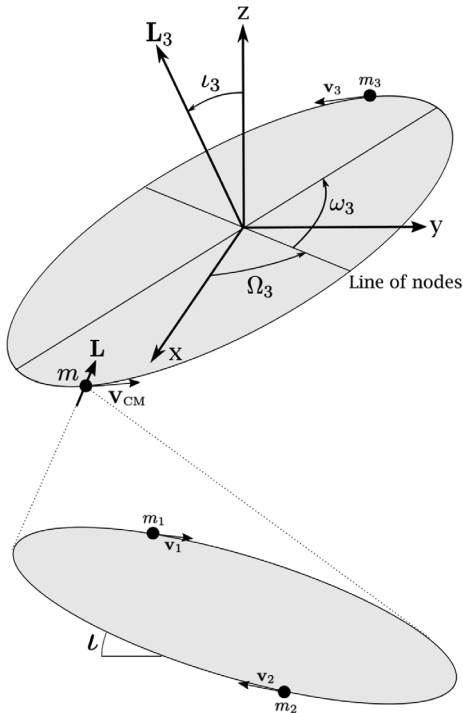


FIG. 1. The effective two-body description and the osculating elements. The inner binary is replaced with a point particle whose spin and multipole moments are related to the osculating elements of the inner orbit.

hierarchical three-body systems: the typical velocity of the inner bodies divided by light speed, v ($c = 1$), and the ratio $\varepsilon = a/a_3$ between the semimajor axis of the inner binary a and the one of the outer binary a_3 . Thanks to the double perturbative expansion with respect to these parameters, it is possible to match the dynamics of a hierarchical three-body system to a simpler two-body interaction, in which the inner binary is described as a single point particle endowed with multipole moments. This is achieved by employing the EFT techniques developed for the relativistic two-body problem [36,37] and by performing the 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 set up a systematic study of the three-body system.

In [35] we presented the EFT setup and derived the effective Lagrangian describing the system on long timescales up to 1PN dipole order, i.e., up to order $v^2\varepsilon^{3/2}$ beyond the leading Newtonian interaction. Instead, in the present work 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 our previous study [35]. 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, contrary to what happened at lower orders. We address these complications by following the method of near-identity transformations [38], which allows one to consistently implement the averaging procedure to any order of accuracy. While several authors already studied quadrupolar couplings at 1PN order [9,31,32,34,39–42], we are aware of only three which took into account these deviations from the adiabatic approximation [31–34]. 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 [31–33] neglecting some PN interactions that we describe in this work. On the other hand, the derivation in [34] reports a puzzling result that we will mention in Sec. IV.

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 parameters of the effective point particle is subtle in practice. For example, in our previous work [35] we showed how the choice of a spin supplementary condition (a gauge condition on the spin tensor of the effective point particle [43,44]) 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 [45]. Since PN corrections induce a nontrivial relation between momentum and velocity, these two sets of osculating elements will differ in general. We will give in Sec. II A the precise definition of contact elements, and we will elaborate more on their difference with respect to orbital elements in Appendix G. 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 [18]), their orbital counterpart features small variations over long timescales, which offers a new point of view on earlier findings of [31–34].

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 nontrivial 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 Sec. III B. 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 Sec. III D.

Let us now describe in more detail the organization of this article. We will begin by summarizing our “effective two-body” approach in Sec. II, which is defined by four main steps. Steps 1 and 2 will then be performed in Sec. III, where the core of our matching procedure is explained. To keep the discussion as simple as possible, we have deferred the computation of beyond-adiabatic corrections to Appendix C and use only the final result of this appendix in the main text. We then perform the two final steps of our approach in Sec. IV, where we integrate out the gravitational field due to the outer object. Finally, in Sec. V we provide a numerical solution implementing the new relativistic interaction derived in the present work, and we show how it influences the long-timescale dynamics in the case of a particular three-body system. The rest of the appendixes are devoted to the following: a presentation of the averaging procedure that we employ (Appendix A); a discussion of the conservation of the semimajor axis (Appendix B); a review of the Lagrange planetary equations (Appendix D); 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 E); an independent computation of the so-called quadrupole-squared terms of [46] (Appendix F); and the relation between contact and orbital elements (Appendix G).

A. Dictionary of symbols

As a preliminary, we define the quantities that characterize the three-body problem and list the symbols that we are going to use through the work. 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 Sec. II and examine their difference from orbital elements in Appendix G.

- (i) $\mathbf{y}_1, \mathbf{v}_1, \mathbf{y}_2, \mathbf{v}_2$: positions and velocities of the two constituents of the inner orbit, of masses m_1 and m_2 ;
- (ii) $\mathbf{y}_3, \mathbf{v}_3$: position and velocity of the external perturber, of mass m_3 ;

- (iii) $\mathbf{Y}_{\text{CM}}, \mathbf{V}_{\text{CM}}$: position and velocity of the center of mass of the inner binary, defined in Eq. (E1);
- (iv) \mathbf{r}, \mathbf{v} : relative variables in the inner binary instantaneous rest frame, defined in Sec. III B. Note that $\mathbf{r} = \mathbf{y}_1 - \mathbf{y}_2$ and $\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2$ only at lowest PN order, as explained in Sec. III B;
- (v) \mathbf{p}_A ($A = 1, 2, 3$), $\mathbf{P}_{\text{CM}}, \mathbf{p}$: conjugate momenta to the positions of the three bodies, position of the center of mass, and relative separation between two inner bodies;
- (vi) $r = |\mathbf{r}|$, $\mathbf{n} = \mathbf{r}/r$, $\mathbf{R} = \mathbf{Y}_{\text{CM}} - \mathbf{y}_3$, $R = |\mathbf{R}|$, $\mathbf{N} = \mathbf{R}/R$, $\mathbf{V} = \mathbf{V}_{\text{CM}} - \mathbf{v}_3$;
- (vii) $m = m_1 + m_2$ is the mass of the inner binary, $\mu = m_1 m_2 / m$ is the reduced mass of the inner binary, $X_1 = m_1 / m$, $X_2 = m_2 / m$ and $\nu = \mu / m$ are its characteristic mass ratios;
- (viii) a [a_3]: contact semimajor axis of the inner [outer] orbit;
- (ix) e [e_3]: contact eccentricity of the inner [outer] orbit;
- (x) $\mathcal{E} = m - G_N m \mu / (2a)$: total (relativistic) energy of the inner binary;
- (xi) $M = \mathcal{E} + m_3$ is the total mass of the effective two-body system. Similarly, $\mu_3 = m_3 \mathcal{E} / M$ is its reduced mass, $X_3 = m_3 / M$, $X_{\text{CM}} = \mathcal{E} / M$, and $\nu_3 = \mu_3 / M$ are the mass ratios characterizing the outer orbit;
- (xii) $n = \sqrt{G_N m / a^3}$ [$n_3 = \sqrt{G_N M / a_3^3}$]: contact frequency of the inner [outer] orbit;
- (xiii) $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ [$\boldsymbol{\alpha}_3, \boldsymbol{\beta}_3, \boldsymbol{\gamma}_3$]: contact orthonormal basis of vectors characterizing the inner [outer] orbit, aligned respectively along the contact semimajor axis (pointing toward the pericenter), the semiminor axis, and the angular momentum;
- (xiv) Ω, ω, ι [$\Omega_3, \omega_3, \iota_3$]: contact angles characterizing the orientation of the inner [outer] orbit, defined by $\boldsymbol{\alpha} = R_z(\Omega) R_x(\iota) R_z(\omega) \mathbf{u}_x$ where the R_{x_i} 's are rotation matrices along the given axis x_i , and similarly for the other vectors;
- (xv) u and η [u_3 and η_3]: contact mean and eccentric anomaly of the inner [outer] orbit. They are defined by $\eta - e \sin \eta = u$ and $u = nt + \sigma$ [$\eta_3 - e_3 \sin \eta_3 = u_3$ and $u_3 = n_3 t + \sigma_3$] where σ and σ_3 are the contact mean anomalies at initial time;
- (xvi) L, G, H [L_3, G_3, H_3]: contact conjugate momenta to u, ω , and Ω , respectively [u_3, ω_3 , and Ω_3 , respectively], defined in Eq. (C2);
- (xvii) $\mathbf{J} = \mu \sqrt{G_N m a (1 - e^2)} \boldsymbol{\gamma}$ [$\mathbf{J}_3 = \mu_3 \sqrt{G_N M a_3 (1 - e_3^2)} \boldsymbol{\gamma}$]: 1PN corrected (i.e., contact) Newtonian angular momentum vector of the inner [outer] orbit.

In addition to these symbols, we will use tilded contact elements to indicate orbital elements, which will only appear in Appendix G. We use the mostly positive metric signature and set $c = 1$. Greek letters refer to spacetime indices ranging from 0 to 3, while latin indices are spatial and range from 1 to 3. $\epsilon_{\alpha\beta\gamma\delta}$ denotes the totally

antisymmetric Levi-Civita tensor, not to be confused with our hierarchical expansion parameter $\varepsilon \equiv a/a_3$. The symmetric and antisymmetric parts of a tensor are defined with parentheses and brackets, respectively, $A_{(\mu\nu)} = (A_{\mu\nu} + A_{\nu\mu})/2$ and $A_{[\mu\nu]} = (A_{\mu\nu} - A_{\nu\mu})/2$.

II. THE EFT APPROACH TO THE RELATIVISTIC HIERARCHICAL THREE-BODY PROBLEM

We consider a hierarchical system in which the inner binary distance vector r is much smaller than its outer binary counterpart R . The idea at the core of [35] is that one can establish a correspondence between the Lagrangian of a hierarchical three-body system and an effective Lagrangian describing a two-body system, where one of the two bodies is a spinning point-particle endowed with multipole moments representing the inner binary. This identification was performed up to dipolar order in [35]. Here we summarize briefly the main steps of the computation:

- (1) We start with the GR action describing two bodies (the inner system) in the presence of a gravitational field:

$$S = \frac{M_{\text{P}}^2}{2} \int d^4x \sqrt{-g} R - \sum_{A=1,2} m_A \int dt \sqrt{-g_{\mu\nu} v_A^\mu v_A^\nu}, \quad (1)$$

where $v_A^\mu = (1, \mathbf{v}_A)$ is the coordinate velocity of the point particle. The line element of our spacetime is decomposed in the nonrelativistic limit following the space + time splitting presented in [47,48]

$$ds^2 = -e^{2\phi} (dt - A_i dx^i)^2 + e^{-2\phi} \gamma_{ij} dx^i dx^j, \quad (2)$$

where it is sufficient to consider the spatial part of the metric as diagonal for our 1PN computations: $\gamma_{ij} \simeq \delta_{ij}$. The two metric fields are decomposed using the background field method as $\phi = \bar{\phi} + \tilde{\phi}$, $A_i = \bar{A}_i + \tilde{A}_i$, where the tilde quantities correspond to an external arbitrary field (which will be related to the gravitational field of the third distant object in Step 3), while we integrate out the barred quantities corresponding to gravitons exchanges between the two bodies. At 1PN order, the resulting effective Lagrangian is

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

where \mathcal{L}_{EIH} is the Einstein-Infeld-Hoffmann Lagrangian given by [49]

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

- (2) The Lagrangian (3) is expanded in multipoles around the center of mass of the inner orbit, and then averaged over one period of the inner orbit to describe the dynamics on longer timescales. As discussed in Sec. III, in this way we obtain the classical effective Lagrangian for the long-timescale modes of the system. This Lagrangian describes a composite spinning point particle coupled to an arbitrary external gravitational field. The multipole moments of the composite object can be described in terms of either its osculating elements or its contact elements (see Sec. II A). For instance, the orbital elements would be defined through the position and velocity of the inner binary by

$$\begin{aligned} \tilde{a} &= -\frac{G_N m}{2} \left(\frac{v^2}{2} - \frac{G_N m}{r} \right)^{-1}, \\ \tilde{\boldsymbol{\alpha}} &= \frac{1}{G_N m} \mathbf{v} \times (\mathbf{r} \times \mathbf{v}) - \frac{\mathbf{r}}{r}, \\ \tilde{\boldsymbol{\gamma}} &= \frac{\mathbf{r} \times \mathbf{v}}{\sqrt{G_N m \tilde{a} (1 - \tilde{e}^2)}}. \end{aligned} \quad (5)$$

In this equation, \tilde{a} and $\tilde{\boldsymbol{\alpha}}$ are the semimajor axis and eccentricity of the orbit, while $\tilde{\boldsymbol{\gamma}}$ and $\tilde{\boldsymbol{\alpha}}$ are two unit vectors, respectively, directed along the Newtonian angular momentum and the Newtonian perihelion of the inner orbit. In this work, to handle higher-order corrections in a compact way, we find it convenient to use contact elements, which we indicate with untilded symbols and define below [see Eq. (11)].

- (3) We add the action of a third point particle and, as in the first step, integrate the external fields $\tilde{\phi}$, \tilde{A}_i mediating interactions between the third point particle and the effective spinning point particle representing the inner binary. The number of terms to take into account is dictated by a set of power-counting rules generalizing the non-relativistic general relativity (NRGR) approach [36,37] to the hierarchical three-body problem and is described in [35]. The two dimensionless parameters controlling the perturbative expansion are the usual PN parameter $v^2 = G_N m/a$ and the ratio of semimajor axes $\varepsilon = a/a_3$, where a_3 is the semimajor axis of the outer orbit.
- (4) We finally average the Lagrangian over the outer binary timescale. At this final step, the Lagrangian depends only on the osculating elements of both

inner and outer orbits. Note that the osculating elements of the outer orbit are defined similarly as in (5), replacing all inner quantities with outer ones. In particular, at first order in the PN expansion, the total mass of the outer binary will be $m_3 + \mathcal{E}$, where $\mathcal{E} = m - G_N m / (2a)$ is the total energy of the inner binary to Newtonian order. Thus, the osculating elements of the outer orbit implicitly contain

post-Newtonian corrections, which is a difference of our approach compared to previous ones (e.g., [9,31,34]) which use the total mass $m + m_3$ to define the outer orbit osculating elements.

After these four steps, we obtain a Lagrangian describing the interactions between the inner and outer orbits up to order $v^2 \epsilon^{5/2}$ on timescales much larger than the orbital periods:

$$\mathcal{L} = \mathbf{J} \cdot \boldsymbol{\Omega} + \mathbf{J}_3 \cdot \boldsymbol{\Omega}_3 + 3\mu \frac{G_N^2 m^2}{a^2 \sqrt{1-e^2}} + 3\mu_3 \frac{G_N^2 M^2}{a_3^2 \sqrt{1-e_3^2}} - \frac{4m + 3m_3}{2m} \Omega_{\text{prec}} \mathbf{J} \cdot \mathbf{J}_3 + \langle \mathcal{L}_{\text{quad}}^{\leq v^2 \epsilon^{5/2}} \rangle. \quad (6)$$

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

$$\mathbf{J} = \mu \sqrt{G_N m a (1-e^2)} \boldsymbol{\gamma}, \quad \boldsymbol{\Omega} = \boldsymbol{\alpha} \times \dot{\boldsymbol{\alpha}}, \quad (7)$$

with analogous formulas for the outer orbit. Let us comment on each component of Eq. (6). 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, named Lagrange planetary equations (LPE) (see Appendix D). 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 (6) 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 [50]. We refer to them as internal Lagrangians since they would be present even without any interaction between the two orbits. The third term is a coupling between the angular momenta of the two orbits at dipole order $v^2 \epsilon^{3/2}$. This term, whose effect on the orbits of hierarchical systems was already studied, e.g., in [39–41], was the highest-order coupling in the ϵ expansion that we described in [35]. It involves a precession frequency given by

$$\Omega_{\text{prec}} = \frac{G_N}{a_3^3 (1-e_3^2)^{3/2}}. \quad (8)$$

Finally, the term $\mathcal{L}_{\text{quad}}^{\leq v^2 \epsilon^{5/2}}$ encodes the contributions from quadrupole-suppressed 0PN and 1PN interactions to the long-timescale dynamics, which we will compute in the next sections and which is the main result of the present work.

A. Contact elements

The quantities above are given 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 in the present work. 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., [45]). This allows one 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

$$\mathcal{L} = \frac{1}{2} \mu v^2 + \frac{G_N m \mu}{r} + \mu \mathcal{R}, \quad (9)$$

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

$$\mathbf{p} = \mu \mathbf{v} + \mu \partial \mathcal{R} / \partial \mathbf{v}. \quad (10)$$

Then, rather than using Eq. (5), we will define the contact elements as follows:

$$\begin{aligned} a &= -\frac{G_N m}{2} \left(\frac{p^2}{2\mu^2} - \frac{G_N m}{r} \right)^{-1}, \\ e\boldsymbol{\alpha} &= \frac{1}{G_N m \mu^2} \mathbf{p} \times (\mathbf{r} \times \mathbf{p}) - \frac{\mathbf{r}}{r}, \\ \boldsymbol{\gamma} &= \frac{\mathbf{r} \times \mathbf{p}}{\sqrt{G_N m \mu^2 a (1-e^2)}}, \end{aligned} \quad (11)$$

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

$$\mathbf{r} = a((\cos\eta - e)\boldsymbol{\alpha} + \sqrt{1 - e^2} \sin\eta\boldsymbol{\beta}),$$

$$\mathbf{p} = \sqrt{\frac{G_N m}{a}} \frac{\mu}{1 - e \cos\eta} (-\sin\eta\boldsymbol{\alpha} + \sqrt{1 - e^2} \cos\eta\boldsymbol{\beta}), \quad (12)$$

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

$$u \equiv nt + \phi = \sqrt{\frac{G_N m}{a^3}} t + \phi = \eta - e \sin\eta, \quad (13)$$

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 G; see Eqs. (G13)–(G17). An important outcome of Appendix G, 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).

III. THE POINT-PARTICLE EFT TO QUADRUPOLAR ORDER

We now derive an effective Lagrangian describing the inner binary coupled to an external gravitational field on timescales 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 A. 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 C. We close the section by matching the result to a worldline action for the inner binary.

A. Integrating out fast modes

In field theoretic terms, we wish to derive the effective action for the slow modes (the long-timescale 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-timescale 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 timescales. Intuitively, a method to achieve this is considering the average and the average-free part (with respect to the orbital timescale) of the original equations of

motion. The splitting allows one 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, Refs. [38,51]).

Despite its intuitiveness, averaging presents a few subtleties that 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 setup 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 $T[\mathcal{L}]$ the period as a function of the slow variables \mathcal{L} , the average of a quantity $A(\mathcal{L}, u)$ can be defined in two different ways:

$$\langle A \rangle_u = \int_{u_0}^{u_0+2\pi} \frac{du}{2\pi} A(\mathcal{L}, u), \quad (14)$$

$$\begin{aligned} \langle A \rangle_t &= \frac{1}{T[\mathcal{L}(t)]} \int_t^{t+T[\mathcal{L}(t)]} dt' A(\mathcal{L}, u(t')) \\ &= \frac{1}{T[\mathcal{L}(t)]} \int_{u(t)}^{u(t)+2\pi} du \frac{dt}{du} A(\mathcal{L}, u). \end{aligned} \quad (15)$$

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., [51]), 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, closely following [38] 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 A.

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 it 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 A, 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.

B. Center of mass and relative coordinates in boosted frame

To carry out the multipole expansion, we express the two inner bodies' coordinates y_1, y_2 as functions of the position of the inner binary center of mass and of the relative distance between the two inner bodies, \mathbf{Y}_{CM} and \mathbf{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. 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. (12). These two frames are connected by a boost plus a translation, a transformation that is nontrivial 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 [35,44]. For this reason, we will express the terms that appear in the effective Lagrangian, a

functional evaluated in the global three-body rest frame \mathbf{R} , in terms of the coordinates of the inner binary rest frame \mathbf{R}' , so as to obtain a functional depending on the *intrinsic* contact elements of the inner binary. As studied in [35], 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 E, let us just show here the final relation between the absolute coordinates (y_1, y_2) and the relative ones, $\mathbf{Y}_{\text{CM}}, \mathbf{V}_{\text{CM}}$ (center-of-mass position and velocity of the inner binary to 1PN order) and \mathbf{r}', \mathbf{p}' (relative distance and momentum in the inner binary rest frame to 1PN order):

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{Y}_{\text{CM}} + (X_2 + \delta)\mathbf{r}' + X_2(\mathbf{V}_{\text{CM}} \cdot \mathbf{r}') \\ &\quad \times \left[(X_1 - X_2) \frac{\mathbf{p}'}{\mu} - \frac{\mathbf{V}_{\text{CM}}}{2} \right], \\ \mathbf{y}_2 &= \mathbf{Y}_{\text{CM}} + (-X_1 + \delta)\mathbf{r}' - X_1(\mathbf{V}_{\text{CM}} \cdot \mathbf{r}') \\ &\quad \times \left[(X_1 - X_2) \frac{\mathbf{p}'}{\mu} - \frac{\mathbf{V}_{\text{CM}}}{2} \right], \end{aligned} \quad (16)$$

where δ is a 1PN quantity defined by

$$\delta = -\frac{1}{m} \mathbf{V}_{\text{CM}} \cdot \mathbf{p}' + \nu(X_1 - X_2) \left(\frac{p'^2}{2\mu} - \frac{G_N m}{2r'} \right). \quad (17)$$

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

$$\mathbf{r} = \mathbf{r}' - (\mathbf{V}_{\text{CM}} \cdot \mathbf{r}') \left[\frac{\mathbf{V}_{\text{CM}}}{2} + (X_2 - X_1) \frac{\mathbf{p}'}{\mu} \right]. \quad (18)$$

The final step is just to express \mathbf{r}', \mathbf{p}' in terms of the contact elements as in Eq. (12). To avoid clutter, in the rest of the article we will suppress the primed label on \mathbf{r}', \mathbf{p}' .

C. Averaging the Lagrangian

In this section, we start the computation of the quadrupolar Lagrangian by carrying out Step 2 described in Sec. II. Specifically, we expand the Lagrangian (3) in multipoles around the center of mass of the inner binary, using the formulas given in Eq. (16). When carrying out computations up to order $\nu^2 \varepsilon^{5/2}$, it is crucial to include corrections to the leading order averaging procedure besides the quadrupolar order of the multipole expansion. These are due to the 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-timescale dynamics. To simplify the presentation, we leave the detailed analysis of these corrections to Appendix C. 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, $\mathcal{L}_{\text{quad},12}^{\leq v^2 \varepsilon^{5/2}}$, 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 Appendix C:

$$\begin{aligned} \mathcal{L}_{\text{quad}}^{(0)} = & -\frac{1}{2}\mu r^i r^j \partial_i \partial_j \tilde{\phi} \left[1 - \frac{G_{Nm}}{r} (1 - 2\nu) + \frac{3}{2}(2(X_2 - X_1)\mathbf{V}_{\text{CM}} \cdot \mathbf{v} + (1 - 3\nu)v^2) \right. \\ & \left. + (X_1 - X_2)\mathbf{V}_{\text{CM}} \cdot \mathbf{r} (r^i v^j + r^j v^i) \right] + \frac{1}{2}\mu r^i r^j \partial_i \partial_j \tilde{A}_k [V_{\text{CM}}^k + (X_2 - X_1)v^k], \end{aligned} \quad (20)$$

where all the relevant quantities used in this equation have been defined in Sec. III B and \mathbf{r} , \mathbf{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 to get to Eq. (20). First, we have dropped terms nonlinear in $\tilde{\phi}$ as well as V_{CM}^2 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 this paper. Second, although at this point the coupling $\partial_i \partial_j \tilde{A}_k V_{\text{CM}}^k$ is of $\varepsilon^{5/2} v$ order, it cannot contribute to the final $\varepsilon^{5/2} v^2$ Lagrangian when \tilde{A}_k has been integrated out. This is because the lowest-order coupling of \tilde{A}_k to m_3 involves also v_3 , so that

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 $\mathcal{L}_{\text{quad}}^{\leq v^2 \varepsilon^{5/2}}$ presented in Step 4 of Sec. II.

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

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 $\partial_i \partial_j \tilde{A}_k V_{\text{CM}}^k$ 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 Sec. III B, manifests itself in the last factor that multiplies $\partial_i \partial_j \tilde{\phi}$.

Following the method of near-identity transformations, presented in Appendix A, we can implement the averaging procedure and eliminate the short-timescale dynamics from the Lagrangian. We leave to Appendix C the corrections due to the breaking of adiabaticity and backreaction of short modes, $\langle \mathcal{L}_S \rangle$ in Eq. (19), while here we average Eq. (20) over one orbit of the inner binary, using the definition of contact elements given in Eq. (11). We obtain

$$\begin{aligned} \langle \mathcal{L}_{\text{quad}}^{(0)} \rangle = & -\frac{\mu a^2}{4} \left[(1 + 4e^2)\alpha^i \alpha^j + (1 - e^2)\beta^i \beta^j + \frac{G_{Nm}}{2a} ((1 - 5\nu + 4e^2(2\nu - 1))\alpha^i \alpha^j + (1 - 5\nu)(1 - e^2)\beta^i \beta^j) \right] \partial_i \partial_j \tilde{\phi} \\ & - \frac{1}{2}(X_2 - X_1) a e \varepsilon_{lik} J^l \alpha^j (\partial_i \partial_j \tilde{A}_k - 4V_{\text{CM}}^k \partial_i \partial_j \tilde{\phi}), \end{aligned} \quad (21)$$

where we recall that \mathbf{J} is the angular momentum of the binary defined in Eq. (7).

At this point we can compute the whole contribution $\mathcal{L}_{\text{quad},12}^{\leq v^2 \varepsilon^{5/2}}$ by adding the result from our analysis in Appendix C, i.e., Eq. (C16), to the contribution of Eq. (21):

$$\mathcal{L}_{\text{quad},12}^{\leq v^2 \varepsilon^{5/2}} = -\frac{1}{2} Q_E^{ij} \partial_i \partial_j \tilde{\phi} - \frac{1}{2} (X_2 - X_1) a e \varepsilon_{lik} J^l \alpha^j (\partial_i \partial_j \tilde{A}_k - 4V_{\text{CM}}^k \partial_i \partial_j \tilde{\phi}). \quad (22)$$

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

$$Q_E^{ij} = \frac{\mu a^2}{2} \left(f_\alpha(e) \alpha^i \alpha^j + f_\beta(e) \beta^i \beta^j - \frac{f_\alpha(e) + f_\beta(e)}{3} \delta^{ij} \right), \quad (23)$$

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

and the two functions of the eccentricity read

$$f_\alpha(e) = 1 + 4e^2 - \frac{G_{Nm}}{2a(1-e^2 + \sqrt{1-e^2})} [17 + 13\sqrt{1-e^2} + 5\nu(1 + \sqrt{1-e^2}) + e^2(56 + 15\sqrt{1-e^2} - \nu(13 + 8\sqrt{1-e^2})) + 4e^4(3 + 2\nu)], \quad (24)$$

$$f_\beta(e) = 1 - e^2 - \frac{G_{Nm}}{2a(1-e^2 + \sqrt{1-e^2})} [13 + 17\sqrt{1-e^2} + 5\nu(1 + \sqrt{1-e^2}) + e^2(31 + 18\sqrt{1-e^2} - 5\nu(2 + \sqrt{1-e^2})) + e^4(5\nu - 9)]. \quad (25)$$

The first term of the Lagrangian (22) contains both Newtonian e^2 and PN $v^2 e^2$ scalings, while its second term proportional to the difference of masses contains only the PN $v^2 e^{5/2}$ 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 G [see Eq. (G20)]. 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. (G20) implies that while the contact element a is constant by virtue of our averaging procedure (see Appendix B), the respective orbital element features post-Newtonian variations. This confirms the findings of the author of [31,33], who showed for the first time that cross terms induce variations in the (orbital) semimajor 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 [31,33].

D. Matching

We will now express the quadrupolar coupling (22) in a gauge-invariant way. As shown in, e.g., [36,53], the quadrupolar part of the effective action can be written in terms of two interactions, the electric-type and magnetic-type quadrupole terms,

$$\mathcal{S}_{\text{quad}} = -\frac{1}{2} \int d\tau E_{ij} Q_E^{ij} + \frac{2}{3} \int d\tau B_{ij} Q_B^{ij}, \quad (26)$$

²Even 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, similar to what is discussed in [52], since it amounts to neglecting a contact term.

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

$$E_{\mu\nu} = C_{\mu\alpha\beta} V_{\text{CM}}^\alpha V_{\text{CM}}^\beta, \\ B_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\alpha\beta\sigma} C^{\alpha\beta}{}_{\nu\rho} V_{\text{CM}}^\sigma V_{\text{CM}}^\rho. \quad (27)$$

Note that in Eq. (26) these tensors have been projected to a locally flat frame.

Being only interested in terms of order $v^2 e^{5/2}$, 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} 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 $\partial_i \tilde{A}^i = -4\dot{\tilde{\phi}}$ [47,48]. Thus, the tensors are given by

$$E_{ij} = \partial_i \partial_j \tilde{\phi}, \quad (28)$$

$$B_{ij} = \frac{1}{2} \epsilon_{mn(i} \partial_{j)} [\partial^m \tilde{A}^n + 4V_{\text{CM}}^m \partial^n \tilde{\phi}]. \quad (29)$$

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. (22)⁴:

³As 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.

⁴To get the coupling in Eq. (22) from Eqs. (29) and (30), one has to make use of the gauge condition on \tilde{A}_i and $\tilde{\phi}$ and the fact that

$$\frac{d\tilde{\phi}}{dt} = \dot{\tilde{\phi}} + V_{\text{CM}}^i \partial_i \tilde{\phi}$$

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

$$Q_B^{ij} = \frac{3}{2}(X_1 - X_2)aeJ^{(i}\alpha^{j)}. \quad (30)$$

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

$$Q_B^{ij} = \epsilon^{mn(i} \sum_A m_A x_A^j) x_A^m v_A^n, \quad (31)$$

(see, e.g., [36]), 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 Sec. III B, is crucial to obtain the correct magnetic-type interaction. Had we used relative variables naively defined in the three-body rest frame (as is the case in other studies [31–34]), 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 that 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:

$$S = \int d\tilde{\tau} \left[-\mathcal{E}_{1\text{PN}} + \frac{1}{2} J_{\mu\nu} \Omega^{\mu\nu} - \frac{1}{2} E_{ij} Q_E^{ij} + \frac{2}{3} B_{ij} Q_B^{ij} \right], \quad (32)$$

where Q_B^{ij} and Q_E^{ij} have been defined in Eqs. (30) and (23), $d\tilde{\tau} = dt \sqrt{-\tilde{g}_{\mu\nu} V_{\text{CM}}^\mu V_{\text{CM}}^\nu}$, and the first two terms represent the Lagrangian of the inner binary to dipole order, with $\mathcal{E}_{1\text{PN}}$ its energy to 1PN order, $J_{\mu\nu}$ its spin tensor, and $\Omega^{\mu\nu}$ its angular velocity rotation vector. These quantities have been computed in [35]; see Eq. (45) there. We can now move on to Step 3 of Sec. II and integrate out the external fields $\tilde{\phi}$, \tilde{A} in the presence of a third point-particle m_3 .

IV. DOUBLE-AVERAGED LAGRANGIAN UP TO ORDER $v^2\epsilon^{5/2}$

In the previous sections we have derived the Lagrangian $\mathcal{L}_{\text{quad},12}^{\leq v^2\epsilon^{5/2}}$ describing the dynamics of the inner binary coupled to an external gravitational field over timescales longer than the inner orbital period, up to order $v^2\epsilon^{5/2}$. 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 timescale to the problem: the orbital period of the outer orbit. Being interested in the evolution on much longer timescales, 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.

TABLE I. Power-counting rules for the quadrupolar vertices contained in the effective point-particle action (32) up to 1PN order, with $J_3 = (G_N M^3 a_3)^{1/2}$, $v^2 = G_N m/a$, and $\epsilon = a/a_3$. The rules are obtained using the scaling presented in Sec. IV A of [35]. 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 [37]).

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

The action of the third point particle is given by

$$\mathcal{L}_3 = -m_3 \sqrt{-\tilde{g}_{\mu\nu} v_3^\mu v_3^\nu}, \quad (33)$$

and we integrate out the gravitational field $\tilde{g}_{\mu\nu}$. Since the corresponding Lagrangian up to dipole order $v^2\epsilon^{3/2}$ has already been described in [35], we will concentrate on the quadrupolar contributions.⁵ From Eq. (32), we see that three new vertices appear at quadrupolar order, whose power-counting rules are summarized in Table I. Integrating out the external fields $\tilde{\phi}$, \tilde{A} as shown in the Feynman diagrams of Fig. 2, we find

$$\mathcal{L}_{\text{quad}}^{\leq v^2\epsilon^{5/2}} = \frac{G_N m_3}{2R^3} (3N^i N_j - \delta^{ij}) \times [Q_{E,ij} + 4(X_2 - X_1)ae\epsilon_{lki} J^l V^k \alpha_j], \quad (34)$$

where we have moved to the center-of-mass frame of the triple system by setting $\mathbf{V}_{\text{CM}} = X_3 \mathbf{V}$, $\mathbf{v}_3 = -X_{\text{CM}} \mathbf{V}$, and we recall that Q_E^{ij} has been defined in Eq. (23). This Lagrangian contains terms with three different scalings: Newtonian quadrupolar (ϵ^2) in the Newtonian part of Q_E^{ij} , 1PN quadrupolar ($v^2\epsilon^2$) in the 1PN part of Q_E^{ij} , and magnetic-type quadrupolar ($v^2\epsilon^{5/2}$) in the second term proportional to the difference of masses.

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

⁵Being 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.

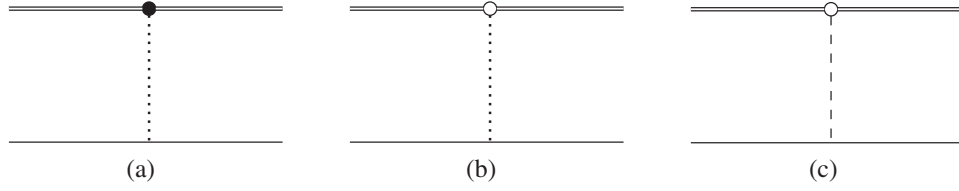


FIG. 2. Feynman diagrams obtained by integrating out the outer fields $\tilde{\phi}$, $\tilde{\mathbf{A}}$ and contributing to the quadrupolar Lagrangian (34) up to order $v^2\epsilon^{5/2}$. The double line represents the inner binary system, while the single line stands for the third particle m_3 . The black [white] dot represents the insertion of an “electric-type” [“magnetic-type”] quadrupolar coupling from Eq. (32). The dotted line represents the exchange of a scalar $\tilde{\phi}$, while the dashed line stands for the exchange of a vector $\tilde{\mathbf{A}}$.

$$\langle \mathcal{L}_{\text{quad}}^{\leq v^2\epsilon^{5/2}} \rangle = \frac{3\tilde{G}_N m_3}{4a_3^3(1-e_3^2)^{3/2}} \mathcal{Q}_E^{ij}(\alpha_3^i \alpha_3^j + \beta_3^i \beta_3^j) + \frac{3(X_2 - X_1)}{X_{\text{CM}}} \Omega_{\text{prec}} \frac{a}{a_3} \frac{ee_3}{1-e_3^2} ((\mathbf{J} \times \boldsymbol{\alpha}) \cdot (\mathbf{J}_3 \times \boldsymbol{\alpha}_3) - 2(\mathbf{J} \cdot \mathbf{J}_3)(\boldsymbol{\alpha} \cdot \boldsymbol{\alpha}_3)), \quad (35)$$

where Ω_{prec} is the precession frequency already defined in Eq. (8). This completes our derivation of the double-averaged Lagrangian up to order $v^2\epsilon^{5/2}$. 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, similar to what is discussed in Appendix C, would contribute only to quadrupole times v_3^2 order, which means starting from octupole order, since $v_3 \sim \epsilon^{1/2}v$.

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. (D7)–(D12). Note that the angular dependence of the quadrupolar $v^2\epsilon^2$ terms is quite similar to the Newtonian quadrupole at ϵ^2 , so that we expect that these $v^2\epsilon^2$ terms will not induce a qualitatively different behavior in the long-term evolution of the system. On the other hand, the angular structure of the $v^2\epsilon^{5/2}$ terms is more involved and somewhat similar to the Newtonian octupole, which means that similar to the octupole such terms can give rise to new behaviors at long times (see, e.g., [18,54]) 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-timescale dynamics obtained from the quadrupolar Lagrangian up to order $v^2\epsilon^{5/2}$.

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 that 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 [31–33]. As shown by Eq. (35), this causes one to neglect the presence of magnetic-type quadrupolar terms of order

$v^2\epsilon^{5/2}$, which can lead to new interesting behaviors in the long-term evolution of the system as we show in Sec. V. 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 G. On the other hand, the work in [34] 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 $v^2\epsilon^{1/2}$ within our power-counting rules. Such a term is absent from our derivation, and we believe that it should not be present.⁶

V. NUMERICAL SOLUTION TO THE LPE

In this section, we will numerically integrate the equations of motion stemming from the averaged Lagrangian (6) 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., Ref. [9] with lower-order perturbations in the Lagrangian, is out of the scope of this paper. Instead, we content ourselves with showing that the quadrupolar terms derived in this paper can have some nontrivial consequences on the long-term evolution of relativistic three-body hierarchical systems.

Varying the total averaged Lagrangian (6) over planetary elements as described in Appendix D, we obtain the so-called LPE which dictate the evolution of orbital

⁶Note that before averaging over the inner orbit, the Lagrangian indeed contains terms scaling as $v^2\epsilon^{1/2}$, 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 [34], the average of these $v^2\epsilon^{1/2}$ terms over the inner orbit still vanishes.

elements over long timescales. For the inner orbit, they are given by

$$\dot{a} = 0, \quad (36)$$

$$\dot{e} = -\sqrt{\frac{1-e^2}{G_N m a e^2}} \frac{\partial \mathcal{R}}{\partial \omega}, \quad (37)$$

$$\dot{i} = -\frac{1}{\sqrt{G_N m a (1-e^2)} \sin i} \frac{\partial \mathcal{R}}{\partial \Omega} + \frac{\cos i}{\sqrt{G_N m a (1-e^2)} \sin i} \frac{\partial \mathcal{R}}{\partial \omega}, \quad (38)$$

$$\dot{\omega} = \sqrt{\frac{1-e^2}{G_N m a e^2}} \frac{\partial \mathcal{R}}{\partial e} - \frac{\cos i}{\sqrt{G_N m a (1-e^2)} \sin i} \frac{\partial \mathcal{R}}{\partial i}, \quad (39)$$

$$\dot{\Omega} = \frac{1}{\sqrt{G_N m a (1-e^2)} \sin i} \frac{\partial \mathcal{R}}{\partial i}, \quad (40)$$

where $\mathcal{R} = \delta\mathcal{L}/\mu$ is the so-called *perturbing function*, with $\delta\mathcal{L}$ containing all terms beyond the first two kinetic parts in the total Lagrangian (6). 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. (6).

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 (6) gives that the total momentum, $\mathbf{J} + \mathbf{J}_3$ is conserved. We can thus follow the conventional elimination of the nodes procedure (see, e.g., Ref. [18]) to choose the orientation of the axis with the z axis parallel to $\mathbf{J} + \mathbf{J}_3$, in which the following relations between planetary elements hold true:

$$\Omega_3 = \Omega + \pi, \quad \|\mathbf{J}\| \sin i = \|\mathbf{J}_3\| \sin i_3. \quad (41)$$

This allows us to eliminate two variables from the eight dynamical variables that we are solving for by, e.g., expressing all quantities in terms of only Ω and $i_{\text{tot}} = i + i_3$. As a relevant remark, let us highlight two subtleties concerning the elimination of nodes. First, as discussed in [54], one cannot directly use the relations (41) 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 [19]) concluded that the z components of both angular momentum are conserved *independently*, which is incorrect as shown in [54]. 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 G. 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, $\|\mathbf{J}\| \cos i + \|\mathbf{J}_3\| \cos i_3$, 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 mass $m = 20M_\odot$ and mass ratio $\nu = 0.15$, while the outer perturber 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 [6,55–57]. The initial conditions of the system are described in Fig. 3, where we plot the total inclination $i_{\text{tot}} = i + i_3$ 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 $v^2 \epsilon^{3/2}$, i.e., dipole order) quenches the maximal eccentricity as well as the orbital flips. This behavior is well-known in the literature [9,25,58]. However, we find that adding the new 1PN quadrupolar terms that we computed in this article can retrigger the flips in inclination. This is shown in Fig. 3. Whether 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 quasiperiodic oscillations as a proxy for the absence of any resonant behavior (see discussion in Appendix B). 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 that is parametrically larger than the interval explored in the simulation.

⁷The fact that $\dot{\Omega} \neq \dot{\Omega}_3$ for osculating elements was already noticed in [34]. In comparison, our discussion adds that the elimination of nodes can be carried out consistently at the level of contact elements.

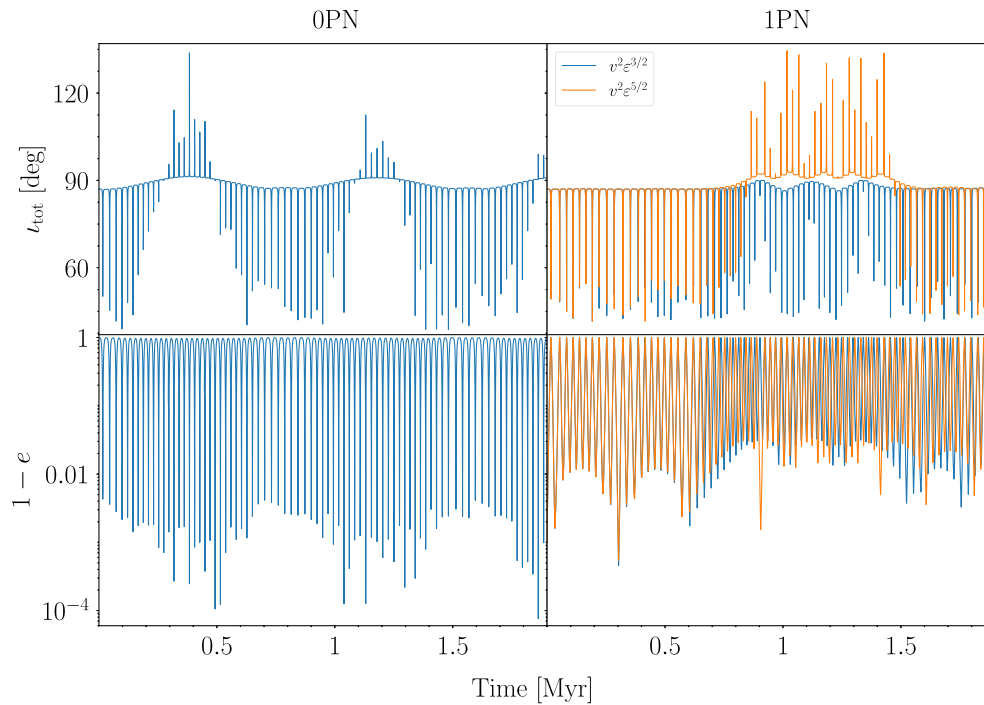


FIG. 3. *Impact of the quadrupole-1PN terms on the evolution of a three-body system.* We solve the LPE (36) with the following parameters for a three-body system: $m = 50M_{\odot}$, $\nu = 0.15$, $a = 5$ AU, $m_3 = 50M_{\odot}$, $a_3 = 350$ AU. The initial conditions are $e = 0.001$, $e_3 = 0.7$, $i_{\text{tot}} = 87^\circ$, $\omega = 240^\circ$, $\omega_3 = 0^\circ$, $\Omega = 0^\circ$. In the left plots we show the evolution of the total inclination $i_{\text{tot}} = i + i_3$ 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., [18]. The system features orbital flips and eccentricities as high as $1 - e \sim 10^{-4}$, which are typical of the octupolar Kozai-Lidov mechanism [5,6,18,54]. 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 $v^2\epsilon^{3/2}$ order (in blue) and up to $v^2\epsilon^{5/2}$ 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 [9,25,39–41,58]) generically quench the orbital flips and eccentricities. However, the new $v^2\epsilon^{5/2}$ terms computed in this article retrigger the orbital flips.

VI. CONCLUSIONS AND OUTLOOK

In this work 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 us 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 one to build

toward a systematic understanding of the long-timescale 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-timescale 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 one 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 timescales. 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.

We envision a few possible directions for future work. An obvious possibility is to derive the EFT for the long-timescale dynamics to higher order in either PN or multipole expansion. Several works have derived the three-body potential for gravitational systems in the PM expansion (see, e.g., [59]), computing also the first orders of the PN expansion. These results could be used as a starting point to derive the effective action for the long-timescale dynamics of the hierarchical three-body problem to higher PN orders. It would also be very interesting to use our effective two-body approach to study gravitational-wave generation in three-body systems; indeed, it is known that a blind application of the quadrupole formula in the center-of-mass frame of a hierarchical three-body system can lead to unphysical results [60]. In contrast, our approach could be used to consistently derive the radiation-reaction force in these systems.

Another possible direction of work would be to develop a systematic study of the presence or absence of resonances for given initial data. In this regard, our procedure of integrating out microscopic degrees of freedom in two different steps, despite its simplicity, has the drawback of forcing us to ignore possible resonant behaviors. In contrast, one might integrate out gravitons and perform the averaging in a single step, without inputting from the start the hierarchy between the semimajor axes. For instance, the method adopted in [61] of performing the hierarchical expansion by using the method of regions might be employed also in the averaging procedure. In any case, having a result that does not rely on the hierarchy between the two orbits might make it possible to detect resonant behaviors as well as to eventually derive our convenient description away from resonances. Further, it would be interesting to make more explicit the connection between the averaging procedure and the underlying field theory. One possibility could be describing the inner and outer orbits through coherent states of the Kepler problem (see, e.g., [62]). This might make it possible to derive our results diagrammatically, without writing the equations of motion for fast and slow modes.

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-timescale 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 [63,64]; taking into account the relativistic three-body effects that we computed in this work may be

crucial for the parameter estimation of these kind of systems.

ACKNOWLEDGMENTS

This research has been partly supported by the Italian MIUR under Contract No. 2017FMJFMW (PRIN2017).

APPENDIX A: AVERAGING THROUGH NEAR-IDENTITY TRANSFORMATIONS

We here present the averaging procedure that we adopt in our computations, the so-called averaging by near-identity transformations [38]. 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 :

$$\begin{aligned}\frac{d}{dt}x &= \epsilon g_1(x, u) + \epsilon^2 g_2(x, u) + \dots, \\ \frac{d}{dt}u &= H(x) + \epsilon h_1(x, u) + \epsilon^2 h_2(x, u) + \dots,\end{aligned}\quad (\text{A1})$$

where $\epsilon \ll 1$, the dots indicate terms with $i > 2$, and the functions g_i and h_i are periodic in u with period 2π . Physically, we can think of ϵ as a ratio of timescales, since the slow variable x has excursions of order 1 over times that are $1/\epsilon$ 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:

$$\begin{aligned}u(t) &= u_L(t) + \epsilon u_S(x_L, u_L), \\ x(t) &= x_L(t) + \epsilon x_S(x_L, u_L),\end{aligned}\quad (\text{A2})$$

where u_S, x_S encode the short-timescale dynamics and are chosen to be periodic in u_L with period 2π . Equation (A2) is called a near-identity transformation, since for $\epsilon \rightarrow 0$ it reduces to an identity. Physically, this ansatz encodes the fact that fast oscillations will have a 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 Eqs. (A1) order by order in ϵ , 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 timescale, up to corrections that are of arbitrarily high order (which are neglected after the truncation). Plugging the transformation in Eq. (A1) we have to second order

$$\begin{aligned}
\dot{x}_L + \epsilon(\dot{x}_L \partial_{x_L} + \dot{u}_L \partial_{u_L})x_S &= \epsilon g_1(x_L, u_L) + \epsilon^2(x_S \partial_{x_L} + u_S \partial_{u_L})g_1(x_L, u_L) + \epsilon^2 g_2(x_L, u_L) + \mathcal{O}(\epsilon^3), \\
\dot{u}_L + \epsilon(\dot{x}_L \partial_{x_L} + \dot{u}_L \partial_{u_L})u_S &= H(x_L) + \epsilon H'(x_L)x_S + \frac{\epsilon^2}{2}H''(x_L)x_S^2 + \epsilon h_1(x_L, u_L) \\
&\quad + \epsilon^2(x_S \partial_{x_L} + u_S \partial_{u_L})h_1(x_L, u_L) + \epsilon^2 h_2(x_L, u_L) + \mathcal{O}(\epsilon^3).
\end{aligned} \tag{A3}$$

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 ϵ^3 :

$$\begin{aligned}
\dot{x}_L &= \epsilon G_1(x_L) + \epsilon^2 G_2(x_L) + \mathcal{O}(\epsilon^3), \\
\dot{u}_L &= H(x_L) + \epsilon H_1(x_L) + \epsilon^2 H_2(x_L) + \mathcal{O}(\epsilon^3).
\end{aligned} \tag{A4}$$

To obtain this, we write $x_S = x_S^{(0)} + \epsilon x_S^{(1)} + \dots$ and $u_S = u_S^{(0)} + \epsilon u_S^{(1)} + \dots$. Then, comparing Eq. (A4) with Eq. (A3) we find that at first order it must be

$$\begin{aligned}
\partial_{u_L} x_S^{(0)} &= \frac{1}{H(x_L)}(g_1(x_L, u_L) - G_1(x_L)), \\
\partial_{u_L} u_S^{(0)} &= \frac{1}{H(x_L)}(h_1(x_L, u_L) + H'(x_L)x_S^{(0)}(x_L, u_L) \\
&\quad - H_1(x_L)).
\end{aligned} \tag{A5}$$

Integrating the first equation at fixed x_L , we find

$$x_S^{(0)}(x_L, u_L) = \int_0^{u_L} \frac{ds}{H(x_L)}(g_1(x_L, s) - G_1(x_L)) + C_0(x_L). \tag{A6}$$

From this, since by assumption $x_S^{(0)}$ is periodic in u_L with period 2π , we see that it must be

$$\int_0^{2\pi} \frac{ds}{H(x_L)}(g_1(x_L, s) - G_1(x_L)) = 0, \tag{A7}$$

which in turn implies that G_1 must be chosen to be the average of g_1 :

$$G_1 = \langle g_1 \rangle_{u_L} = \int_0^{2\pi} \frac{ds}{2\pi} g_1(x_L, s). \tag{A8}$$

We can further set $C_0(x_L)$ in such a way that $x_S^{(0)}$ has zero average. Schematically we will write

$$x_S^{(0)}(x_L, u_L) = \frac{1}{H(x_L)} \text{Af}_{u_L} \left(\int^{u_L} ds \text{Af}_s(g_1(x_L, s)) \right), \tag{A9}$$

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 timescale 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 ϵ^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

$$u_S^{(0)}(x_L, u_L) = \frac{1}{H(x_L)} \text{Af}_{u_L} \left(\int^{u_L} ds [\text{Af}_s(h_1(x_L, s)) + H'(x_L)x_S^{(0)}(x_L, s)] \right). \tag{A10}$$

This first order truncation of the approximation, when applied to the Newtonian hierarchical three-body problem, gives the Kozai-Lidov long-timescale dynamics. Turning to second order, we have

$$\begin{aligned}
\partial_{u_L} x_S^{(1)} &= \frac{1}{H(x_L)}(g_2 + (x_S^{(0)} \partial_{x_L} + u_S^{(0)} \partial_{u_L})g_1 - H_1 \partial_{u_L} x_S^{(0)} - G_2(x_L)), \\
\partial_{u_L} u_S^{(1)} &= \frac{1}{H(x_L)} \left(h_2 + H'x_S^{(1)} + \frac{H''}{2}(x_S^{(0)})^2 + (x_S^{(0)} \partial_{x_L} + u_S^{(0)} \partial_{u_L})h_1 - H_1 \partial_{u_L} u_S^{(0)} - H_2(x_L) \right),
\end{aligned} \tag{A11}$$

where $g_{1,2}$ and $h_{1,2}$ 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 (RHS). 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 $x_S^{(1)}$ and $u_S^{(1)}$ average-free as well. In particular, we find

$$\begin{aligned}
G_2 &= \int_0^{2\pi} \frac{ds}{2\pi} (g_2(x_L, s) + (x_S^{(0)} \partial_{x_L} + u_S^{(0)} \partial_{u_L}) g_1(x_L, s)), \\
H_2 &= \int_0^{2\pi} \frac{ds}{2\pi} (h_2(x_L, s) + (x_S^{(0)} \partial_{x_L} + u_S^{(0)} \partial_{u_L}) h_1(x_L, s) + \frac{H''}{2} (x_S^{(0)})^2),
\end{aligned} \tag{A12}$$

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 into the equations for x_L and u_L , and the slow dynamics can be determined up to terms of order ϵ^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 then be 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 it 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:

$$u = \eta - e \sin \eta, \tag{A13}$$

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 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

$$e = e_L(x_L) + \epsilon e_S(x_L, u_L); \tag{A14}$$

then we see that it must hold

$$u_L + \epsilon u_S(x_L, u_L) = \eta - (e_L(x_L) + \epsilon e_S(x_L, u_L)) \sin \eta. \tag{A15}$$

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:

$$\eta = \eta_L + \epsilon \eta_S(x_L, u_L). \tag{A16}$$

Then, it is possible to fix η_S so as to retain, to all orders, the relation

$$u_L = \eta_L - e_L \sin \eta_L. \tag{A17}$$

This choice makes it possible to express the integrands evaluated in u_L as simple functions of η_L . Moreover, it makes it clear that at all orders the change of variables will be given by

$$du_L = d\eta_L (1 - e_L \cos \eta_L). \tag{A18}$$

Note that crucially, as a result of integrating at fixed x_L , the Jacobian does not have the denominator factor $\dot{u} + \dot{\epsilon} \sin \eta$ 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 [65,66]. The idea behind this method is to introduce fictitious variables corresponding to long timescales and to determine the dependence on these long timescales by imposing the cancellation of terms that display a secular growth, i.e., terms growing linearly with time that 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 interested in estimating the range of validity in time of the approximate solution [38]. Besides the multiple scale analysis, we quote the method of Von Zeipel transformations, [9,67], which works at the level of the Hamiltonian implementing canonical transformations that eliminate the dependence on short-timescale modes, very similar to what we have done above. Finally, we quote the method of the dynamical renormalization group [68], which operates in a way similar to the multiple scale analysis, removing secularly growing terms by means of counterterms.

APPENDIX B: CONSERVATION OF CONTACT SEMIMAJOR AXIS

We now turn to the question of whether the semimajor axes of the two orbits remain constant over long timescales. Despite the simplicity of this question, to our understanding

the answer is quite involved. There is an intuitive reasoning to argue that the semimajor axes are constant over long timescales. That is, after the averaging procedure is carried out, the Lagrangian becomes independent of the mean anomaly of the corresponding orbit, therefore making the corresponding conjugate momentum constant. The latter, as shown in Eq. (C2), depends only on the semimajor axis. However, when considering PN corrections this argument can hold only for the contact element a , rather than for the orbital element \tilde{a} ; see discussion in Appendix G. Moreover, as we discuss in Appendix A, the dynamical variables left after the averaging procedure are the long-timescale 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 [9]. The result is that, indeed, as long as perturbation theory goes, the semimajor axes remain constant over long timescales. 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 (A1), 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. (A1) as partial derivatives of the Hamiltonian of the system $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1 + \dots$:

$$\begin{aligned} g_1 &= -\frac{\partial \mathcal{H}_1}{\partial u}, & g_2 &= -\frac{\partial \mathcal{H}_2}{\partial u}, \\ H &= \frac{\partial \mathcal{H}_0}{\partial x}, & h_1 &= \frac{\partial \mathcal{H}_1}{\partial x}, \end{aligned} \quad (\text{B1})$$

where we assume \mathcal{H} to be periodic in u , meaning that the functions g_i will be average-free. This allows 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-timescale source terms G_i , in Eq. (A4). For instance, it is evident that $G_1 = 0$, due to g_1 being average-free. For G_2 , given in Eq. (A12), determining the answer is less straightforward. Schematically, we have

$$\begin{aligned} G_2 &= \int_0^{2\pi} \frac{ds}{2\pi} \left\{ \frac{1}{H} \text{Af}(\mathcal{H}_1) \partial_x \partial_u \mathcal{H}_1 - \partial_x \left(\frac{1}{H} \text{Af} \left(\int^u \text{Af}(\mathcal{H}_1) \right) \right) \partial_u^2 \mathcal{H}_1 \right\} \\ &= \int_0^{2\pi} \frac{ds}{2\pi} \left\{ \partial_x \left(\frac{1}{H} \text{Af}(\mathcal{H}_1) \partial_u \mathcal{H}_1 \right) - \partial_u \left[\partial_x \left(\frac{1}{H} \text{Af} \left(\int^u \text{Af}(\mathcal{H}_1) \right) \right) \partial_u \mathcal{H}_1 \right] \right\}. \end{aligned} \quad (\text{B2})$$

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:

$$\frac{1}{H} \text{Af}(\mathcal{H}_1) \partial_u \mathcal{H}_1 = \frac{1}{2H} \partial_u (\mathcal{H}_1^2) - \frac{1}{H} \langle \mathcal{H}_1 \rangle \partial_u \mathcal{H}_1. \quad (\text{B3})$$

These terms give a vanishing contribution to the average; therefore, we find

$$G_2 = 0. \quad (\text{B4})$$

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-timescale part of x and u are still conjugate variables described by a

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 [38]. 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 nontrivial evolution of the semimajor axes on timescales 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 [21].

APPENDIX C: BACKREACTION AND DEVIATIONS FROM ADIABATICITY

As outlined in Appendix A, the averaging procedure can be conveniently defined as an average over the values of the slowly evolving part of the mean anomaly, and then expressed in terms of the eccentric anomaly as in

Eq. (A17). Here we apply this procedure to the three-body problem.

1. Long-timescale and short-timescale Lagrangians

To describe corrections to the adiabatic approximation, we should return to the Lagrangian (3). It will be useful to split it between a Newtonian term and a perturbation:

$$\mathcal{L} = \frac{1}{2}v^2 + \frac{G_N m}{r} + \mathcal{R} = \frac{(G_N m)^2}{2L^2} + L\dot{u} + G\dot{\omega} + H\dot{\Omega} + \mathcal{R}, \quad (\text{C1})$$

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 appendixes. 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 Sec. II A. It can be checked that this Lagrangian indeed gives the Lagrange planetary equations presented in Sec. D and that are usually presented in the Hamiltonian formalism (see also Refs. [35,45]).

Finally, the conjugate momenta are given by

$$L = \sqrt{G_N m a}, \quad G = L\sqrt{1-e^2}, \quad H = G \cos i. \quad (\text{C2})$$

Given Eq. (C1), we can perform near-identity transformations for all the contact elements as outlined in Appendix A. Once we determine the fast oscillating terms in each of the near-identity transformations to a desired order in both ϵ , the ratio of the semimajor axes, and v , the velocity of the bodies, we can simply plug back into the Lagrangian these values, obtaining a classical effective Lagrangian. Once we expand the resulting equations of motion in ϵ and v , we will recover the slow dynamics, as in Eq. (A4). 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-timescale Kozai-Lidov dynamics.

More explicitly, considering first the evolution of L and u alone, we employ the following two near-identity transformations:

$$\begin{aligned} u &= u_L + u_S(L_L, u_L), \\ L &= L_L + L_S(L_L, u_L), \end{aligned} \quad (\text{C3})$$

where the subscripts $_L$ and $_S$ stand for the long and short-timescale variables, respectively, and the $_S$ variables are suppressed by powers of both ϵ and v . Using the LPE for u and L (see Appendix D), we determine the equations for the short-timescale variables as discussed in Appendix A, finding at lowest order

$$\begin{aligned} \partial_{u_L} L_S &= \frac{L_L^3}{(G_N m)^2} \text{Af}_u \left(\frac{\partial \mathcal{R}}{\partial u} \right), \\ \partial_{u_L} u_S &= -\frac{L_L^3}{(G_N m)^2} \left[\text{Af}_u \left(\frac{\partial \mathcal{R}}{\partial L} \right) + \frac{3(G_N m)^2 L_S}{L_L^4} \right], \end{aligned} \quad (\text{C4})$$

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

$$\begin{aligned} L_S &= \frac{L_L^3}{(G_N m)^2} \text{Af}_{u_L} \left(\int_0^{u_L} du \text{Af}_u \left(\frac{\partial \mathcal{R}}{\partial u} \right) \right), \\ u_S &= -\text{Af}_{u_L} \left(\int_0^{u_L} du \left[\frac{L_L^3}{(G_N m)^2} \text{Af}_u \left(\frac{\partial \mathcal{R}}{\partial L} \right) + \frac{3L_S}{L_L} \right] \right). \end{aligned} \quad (\text{C5})$$

This is analogous to what was obtained in Eqs. (A9) and (A10), 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 A is the fact that now there are two small parameters, i.e., ratio of timescales, which we have not factorized explicitly. For convenience, in the following we will indicate $\mathcal{F} = \text{Af}_{u_L}(\mathcal{R})$ and, using that partial derivatives commute with the average over u_L , we will write $\text{Af}_{u_L} \left(\frac{\partial \mathcal{R}}{\partial X} \right) = \frac{\partial \mathcal{F}}{\partial X}$.

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

$$\begin{aligned} \partial_{u_L} G_S &= \frac{L_L^3}{(G_N m)^2} \frac{\partial \mathcal{F}}{\partial \omega}, & \partial_{u_L} \omega_S &= -\frac{L_L^3}{(G_N m)^2} \frac{\partial \mathcal{F}}{\partial G}, \\ \partial_{u_L} H_S &= \frac{L_L^3}{(G_N m)^2} \frac{\partial \mathcal{F}}{\partial \Omega}, & \partial_{u_L} \Omega_S &= -\frac{L_L^3}{(G_N m)^2} \frac{\partial \mathcal{F}}{\partial H}. \end{aligned} \quad (\text{C6})$$

To convert derivatives with respect to canonical momenta in derivatives with respect to osculating contact elements, we use the formulas given in Appendix D. As prescribed by the near-identity transformation approach, the quantities entering the right-hand side of Eqs. (C6) are expressed only in terms of the long-timescale variables e_L, a_L, \dots , as derived in Eq. (A5). 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

backreaction of the fast oscillations on the long-timescale dynamics, using the following procedure. At the level of the Lagrangian, we can substitute the near-identity transformations and write the expression in (C1) as

$$\mathcal{L} = \mathcal{L}_L + \mathcal{L}_S, \quad (\text{C7})$$

the first term corresponding to the Lagrangian evaluated on the long-timescale variables only, while the second, \mathcal{L}_S , corresponding to the remaining part, then to be averaged. For instance, we have

$$L\dot{u} = (L_L + L_S)(\dot{u}_L + \dot{u}_S). \quad (\text{C8})$$

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

$$\mathcal{L}_L = L_L\dot{u}_L + G_L\dot{\omega}_L + H_L\dot{\Omega}_L + \frac{(G_N m)^2}{2L_L^2} + \langle \mathcal{R} \rangle, \quad (\text{C9})$$

$$\begin{aligned} \mathcal{L}_S = & \frac{(G_N m)^2}{2} \left(\frac{1}{(L_L + L_S)^2} - \frac{1}{L_L^2} \right) + L_S\dot{u}_S \\ & + G_S\dot{\omega}_S + H_S\dot{\Omega}_S + \sum_X \frac{\partial \mathcal{F}}{\partial X} X_S, \end{aligned} \quad (\text{C10})$$

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 ϵ^2 or by v^2 .

Plugging back the equations of motion as well as the solutions obtained as in (C5), we find

$$\begin{aligned} \mathcal{L}_S = & -\frac{L_L^3}{(G_N m)^2} \left[\frac{\partial \mathcal{F}}{\partial u} \text{Af}_{u_L} \left(\int_0^{u_L} du \frac{\partial \mathcal{F}}{\partial L} \right) + \frac{\partial \mathcal{F}}{\partial \omega} \text{Af}_{u_L} \left(\int_0^{u_L} du \frac{\partial \mathcal{F}}{\partial G} \right) + \frac{\partial \mathcal{F}}{\partial \Omega} \text{Af}_{u_L} \left(\int_0^{u_L} du \frac{\partial \mathcal{F}}{\partial H} \right) \right] \\ & + \frac{3L_L^2}{(G_N m)^2} \left[\frac{\mathcal{F}^2}{2} - \mathcal{F}^2 - \frac{\partial \mathcal{F}}{\partial u} \text{Af}_{u_L} \left(\int_0^{u_L} du \mathcal{F} \right) \right], \end{aligned} \quad (\text{C11})$$

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. (C5). 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 end points $u_L = 0, 2\pi$. Moreover, we can simplify Eq. (C11) using that $\langle \text{Af}(M)\text{Af}(N) \rangle = \langle \text{Af}(M)N \rangle$ for any arbitrary functions M and N . Thus we find

$$\begin{aligned} \langle \mathcal{L}_S \rangle = & \frac{1}{2\pi} \int_0^{2\pi} du_L \left[\frac{L_L^3}{(G_N m)^2} \left(\frac{\partial \mathcal{F}}{\partial L} \mathcal{F} + \frac{\partial \mathcal{F}}{\partial G} \int_0^{u_L} du \frac{\partial \mathcal{F}}{\partial \omega} \right. \right. \\ & \left. \left. + \frac{\partial \mathcal{F}}{\partial H} \int_0^{u_L} du \frac{\partial \mathcal{F}}{\partial \Omega} \right) + \frac{3L_L^2}{2(G_N m)^2} \mathcal{F}^2 \right]. \end{aligned} \quad (\text{C12})$$

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 [46], which are a purely Newtonian contribution of postadiabatic corrections. As a proof of concept, we use our procedure to compute

such terms in Appendix F and show that they give back the exact same equations as the ones displayed in [46].

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 expanding the Lagrangian (3) in the center-of-mass frame. In doing so, one obtains a quadrupolar part $\mathcal{R}_{\text{quad}}$, of order ϵ^2 within our power-counting rules, and a post-Newtonian part \mathcal{R}_{1PN} . The latter is itself composed of two terms $\mathcal{R}_{\text{1PN}} = \mathcal{R}_{v^2} + \mathcal{R}_{v^2\epsilon^{1/2}}$ scaling differently: the first one is of order v^2 and corresponds to the usual Einstein-Infeld-Hoffmann (EIH) Lagrangian in the center-of-mass frame, and the second one is linear in V_{CM} and thus of order $v^2\epsilon^{1/2}$. 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 Sec. III B. Indeed, once we express the Lagrangian in terms of these coordinates, the contributions of order $v^2\epsilon^{1/2}$ 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. (E2), this term would have been nonzero and would have led to cumbersome formulas that would have prevented us from performing the matching of the inner binary to a point particle as we did in Sec. III D. We now give the expression of the different perturbing functions in terms of r and v as

$$\begin{aligned}\mathcal{R}_{v^2} &= \frac{1}{8}v^4(1-3\nu) \\ &\quad + \frac{G_N m}{2r} \left((3+\nu)v^2 + \nu(\mathbf{v} \cdot \mathbf{n})^2 - \frac{G_N m}{r} \right), \\ \mathcal{R}_{\text{quad}} &= -\frac{1}{2}r^i r^j \partial_i \partial_j \tilde{\phi},\end{aligned}\quad (\text{C13})$$

where we recall that $\tilde{\phi}$ is an arbitrary external field, $X_A = m_A/m$ and $\nu = X_1 X_2$. 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

$$\begin{aligned}\langle \mathcal{L}_S \rangle &= \frac{G_N m a}{8(1-e^2+\sqrt{1-e^2})} \left\{ [2(9+7\sqrt{1-e^2}) + e^2(51+11\sqrt{1-e^2}) + 16e^4] \alpha^i \alpha^j \right. \\ &\quad \left. + [2(7+9\sqrt{1-e^2}) + e^2(29+17\sqrt{1-e^2}) - 8e^4] \beta^i \beta^j \right\} \partial_i \partial_j \tilde{\phi}.\end{aligned}\quad (\text{C16})$$

This term will be part of the effective Lagrangian for the long-timescale dynamics:

$$\mathcal{L}_{\text{eff}} = \langle \mathcal{L}_L \rangle + \langle \mathcal{L}_S \rangle = L_L \dot{u}_L + G_L \dot{\omega}_L + H_L \dot{\Omega}_L + \langle \mathcal{R} \rangle + \langle \mathcal{L}_S \rangle.\quad (\text{C17})$$

This effective Lagrangian contains the term $\mathcal{L}_{\text{quad},12}^{\leq v^2 \varepsilon^{5/2}}$ computed in Sec. III, 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 A, was to differ with respect to the change of variables derived from the Kepler equation (A13). As discussed, we change the variable from the long-timescale part of the mean anomaly to the long-timescale 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 (C17) will not get contributions

$$\begin{aligned}R_{v^2} &= \frac{G_N^2 m^2}{2a^2(1-e\cos\eta)^2} \left[\frac{1-3\nu}{4}(1+e\cos\eta)^2 \right. \\ &\quad \left. + (3+\nu)(1+e\cos\eta) + \nu e^2 \frac{\sin^2\eta}{1-e\cos\eta} - 1 \right],\end{aligned}\quad (\text{C14})$$

$$\begin{aligned}\mathcal{R}_{\text{quad}} &= -\frac{a^2}{2} \left((\cos\eta - e)\alpha^i + \sqrt{1-e^2} \sin\eta \beta^i \right) \\ &\quad \times \left((\cos\eta - e)\alpha^j + \sqrt{1-e^2} \sin\eta \beta^j \right) \partial_i \partial_j \tilde{\phi}.\end{aligned}\quad (\text{C15})$$

We can now compute the cross terms following Eq. (C12). The Lagrangian displayed in Eq. (C12) 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 either of these two variables. Computations are performed using the *Mathematica* software, giving the full expression of cross terms in terms of osculating elements:

due to subleading oscillating parts of the near-identity transformations (e.g., short-timescale functions with amplitudes of order quadrupole-1PN, similar to $x_S^{(1)}$ in the notation of Appendix A), since these will be average-free and will have to multiply terms of \mathcal{L}_L , leading to average-free contributions, if any.

APPENDIX D: LAGRANGE PLANETARY EQUATIONS

The LPE express the time derivatives of all osculating (contact) elements in terms of derivatives of the Hamiltonian [45]. They can be obtained as the equations of motion stemming from the Lagrangian (C1). The derivatives of \mathcal{R} with respect to canonical momenta can be transformed in derivatives with respect to planetary elements by inverting the relations (C2):

$$a = \frac{L^2}{Gm}, \quad e = \sqrt{1 - \frac{G^2}{L^2}}, \quad \cos \iota = \frac{H}{G}.\quad (\text{D1})$$

Thus, one has

$$\frac{\partial \mathcal{R}}{\partial L} = 2\sqrt{\frac{a}{G_N m}} \frac{\partial \mathcal{R}}{\partial a} + \frac{1-e^2}{e\sqrt{G_N m a}} \frac{\partial \mathcal{R}}{\partial e}, \quad (\text{D2})$$

$$\frac{\partial \mathcal{R}}{\partial G} = -\frac{\sqrt{1-e^2}}{e\sqrt{G_N m a}} \frac{\partial \mathcal{R}}{\partial e} + \frac{\cos \iota}{\sqrt{G_N m a(1-e^2)} \sin \iota} \frac{\partial \mathcal{R}}{\partial \iota}, \quad (\text{D3})$$

$$\frac{\partial \mathcal{R}}{\partial H} = -\frac{1}{\sqrt{G_N m a(1-e^2)} \sin \iota} \frac{\partial \mathcal{R}}{\partial \iota}. \quad (\text{D4})$$

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,

$$\left. \frac{\partial \eta}{\partial e} \right|_{\sigma, a \text{ fixed}} = \frac{\sin \eta}{1 - e \cos \eta}, \quad \left. \frac{\partial \eta}{\partial a} \right|_{\sigma, e \text{ fixed}} = -\frac{3nt}{2a(1 - e \cos \eta)}. \quad (\text{D5})$$

Thus,

$$\begin{aligned} \frac{\partial \mathcal{R}}{\partial e} &= \left. \frac{\partial \mathcal{R}}{\partial e} \right|_{\eta \text{ fixed}} + \frac{\sin \eta}{1 - e \cos \eta} \frac{\partial \mathcal{R}}{\partial \eta}, \\ \frac{\partial \mathcal{R}}{\partial a} &= \left. \frac{\partial \mathcal{R}}{\partial a} \right|_{\eta \text{ fixed}} - \frac{3nt}{2a(1 - e \cos \eta)} \frac{\partial \mathcal{R}}{\partial \eta}. \end{aligned} \quad (\text{D6})$$

Apart from this subtlety, one can now easily derive the LPE from the Lagrangian (C1), and they read

$$\dot{a} = \sqrt{\frac{4a}{G_N m}} \frac{\partial \mathcal{R}}{\partial u}, \quad (\text{D7})$$

$$\dot{e} = -\sqrt{\frac{1-e^2}{G_N m a e^2}} \frac{\partial \mathcal{R}}{\partial \omega} + \frac{1-e^2}{\sqrt{G_N m a e}} \frac{\partial \mathcal{R}}{\partial u}, \quad (\text{D8})$$

$$\dot{\iota} = -\frac{1}{\sqrt{G_N m a(1-e^2)} \sin \iota} \frac{\partial \mathcal{R}}{\partial \Omega} + \frac{\cos \iota}{\sqrt{G_N m a(1-e^2)} \sin \iota} \frac{\partial \mathcal{R}}{\partial \omega}, \quad (\text{D9})$$

$$\dot{u} = \sqrt{\frac{G_N m}{a^3}} - \sqrt{\frac{4a}{G_N m}} \frac{\partial \mathcal{R}}{\partial a} - \frac{1-e^2}{\sqrt{G_N m a e}} \frac{\partial \mathcal{R}}{\partial e}, \quad (\text{D10})$$

$$\dot{\omega} = \sqrt{\frac{1-e^2}{G_N m a e^2}} \frac{\partial \mathcal{R}}{\partial e} - \frac{\cos \iota}{\sqrt{G_N m a(1-e^2)} \sin \iota} \frac{\partial \mathcal{R}}{\partial \iota}, \quad (\text{D11})$$

$$\dot{\Omega} = \frac{1}{\sqrt{G_N m a(1-e^2)} \sin \iota} \frac{\partial \mathcal{R}}{\partial \iota}. \quad (\text{D12})$$

The averaged equations can be obtained as explained in Appendix A, through a near-identity transformation. The LPE for the outer orbit are obtained in the very same way, 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. (6).

Finally, let us make two technical remarks. First, in order to express derivatives of the quadrupolar Lagrangian in Eq. (C15) with respect to angles, we use the following derivatives of the basis vectors:

$$\begin{aligned} \frac{\partial \alpha^i}{\partial \omega} &= \beta^i, & \frac{\partial \beta^i}{\partial \omega} &= -\alpha^i, \\ \frac{\partial \alpha^i}{\partial \Omega} &= -\cos \omega \sin \iota \gamma^i + \cos \iota \beta^i, \\ \frac{\partial \beta^i}{\partial \Omega} &= -\cos \iota \alpha^i + \sin \iota \sin \omega \gamma^i, \\ \frac{\partial \alpha^i}{\partial \iota} &= \sin \omega \gamma^i, & \frac{\partial \beta^i}{\partial \iota} &= \cos \omega \gamma^i. \end{aligned} \quad (\text{D13})$$

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 α , β , γ . In this gauge, the derivatives written above simplify greatly:

$$\frac{\partial \alpha^i}{\partial \omega} = \beta^i, \quad \frac{\partial \beta^i}{\partial \omega} = -\alpha^i, \quad \frac{\partial \alpha^i}{\partial \Omega} = -\gamma^i, \quad \frac{\partial \beta^i}{\partial \Omega} = 0, \quad \frac{\partial \alpha^i}{\partial \iota} = 0, \quad \frac{\partial \beta^i}{\partial \iota} = \gamma^i, \quad (\text{D14})$$

$$\frac{\partial \mathcal{R}}{\partial L} = 2\sqrt{\frac{a}{G m}} \frac{\partial \mathcal{R}}{\partial a} + \frac{1-e^2}{e\sqrt{G m a}} \frac{\partial \mathcal{R}}{\partial e}, \quad \frac{\partial \mathcal{R}}{\partial G} = -\frac{\sqrt{1-e^2}}{e\sqrt{G m a}} \frac{\partial \mathcal{R}}{\partial e}, \quad \frac{\partial \mathcal{R}}{\partial H} = -\frac{1}{\sqrt{G m a(1-e^2)}} \frac{\partial \mathcal{R}}{\partial \iota}. \quad (\text{D15})$$

To give further support to the validity of this simplification we refer the reader to Appendix F, where we explicitly derive that the spurious dependence on angles contained in the derivatives (D13) exactly cancel when averaging the high-energy modes in order to obtain the quadrupole-squared terms.

APPENDIX E: 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 (y_1, y_2) of the inner binary and the relative distance \mathbf{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 Sec. III B. 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

$$E\mathbf{Y}_{\text{CM}} = E_1\mathbf{y}_1 + E_2\mathbf{y}_2, \\ E = E_1 + E_2, \quad E_A = m_A + \frac{1}{2}m_A v_A^2 - \frac{G_N m_1 m_2}{2r}, \quad (\text{E1})$$

for $A = 1, 2$. In the three-body rest frame \mathbf{R} , the relation $\mathbf{r} = \mathbf{y}_1 - \mathbf{y}_2$ leads to

$$\mathbf{y}_1 = \mathbf{Y}_{\text{CM}} + (X_2 + \delta)\mathbf{r}, \quad \mathbf{y}_2 = \mathbf{Y}_{\text{CM}} + (-X_1 + \delta)\mathbf{r}, \quad (\text{E2})$$

where we have defined

$$X_A = \frac{m_A}{m}, \quad m = m_1 + m_2, \quad \mu = \frac{m_1 m_2}{m}, \quad \nu = \frac{\mu}{m}, \\ \delta = -\nu \mathbf{V}_{\text{CM}} \cdot \mathbf{v} + \nu (X_1 - X_2) \left(\frac{v^2}{2} - \frac{G_N m}{2r} \right). \quad (\text{E3})$$

However, the relative distance \mathbf{r} in the three-body rest frame \mathbf{R} cannot be expressed in terms of the *intrinsic* contact elements using Eq. (12). Rather, it will be the relative distance in the inner binary rest frame \mathbf{R}' , $\mathbf{r}' = \mathbf{y}'_1 - \mathbf{y}'_2$, and the respective momentum to be related to these convenient variables by Eq. (12). For this reason, we want to express the relative coordinates in the rest frame \mathbf{R} in terms of those defined in \mathbf{R}' . Concretely, around a time t_0 , the two coordinate systems are related by the following transformation illustrated in Fig. 4:

$$\begin{pmatrix} t' - t_0 \\ \mathbf{y}' \end{pmatrix} = \mathcal{B} \begin{pmatrix} t - t_0 \\ \mathbf{y} - \mathbf{Y}_{\text{CM}}(t_0) \end{pmatrix}, \quad (\text{E4})$$

where \mathcal{B} represents a Lorentz boost of velocity $V_{\text{CM}}(t_0)$. To 1PN order, the coordinates of the two inner bodies \mathbf{y}'_A (with $A = 1, 2$) in the frame \mathbf{R}' are related to those of the three-body rest frame \mathbf{R} by

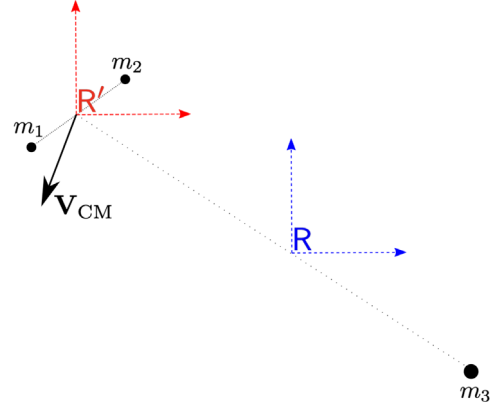


FIG. 4. The change of referential (E4): the rest frame of the inner binary \mathbf{R}' is obtained from the total center-of-mass frame \mathbf{R} by translating it by \mathbf{Y}_{CM} and boosting it by \mathbf{V}_{CM} .

$$\mathbf{y}'_A(t') = \mathbf{y}_A(t') - \mathbf{Y}_{\text{CM}} - \mathbf{V}_{\text{CM}}(t' - t_0) \\ + \mathbf{V}_{\text{CM}} \cdot (\mathbf{y}_A(t') - \mathbf{Y}_{\text{CM}}) \left(\mathbf{v}_A(t') - \frac{\mathbf{V}_{\text{CM}}}{2} \right), \quad (\text{E5})$$

where in the last equation it is understood that \mathbf{Y}_{CM} and \mathbf{V}_{CM} are evaluated at t_0 . Inserting this relation in the center-of-mass definition (E1) to 1PN order, we obtain

$$E_1 \mathbf{y}'_1 + E_2 \mathbf{y}'_2 = (\mathbf{V}_{\text{CM}} \cdot \mathbf{r}') \mathbf{p}', \quad (\text{E6})$$

where we have substituted $\mu \mathbf{v}' \simeq \mathbf{p}'$. From this, we find that the relation between absolute coordinates \mathbf{y}'_A and relative coordinates \mathbf{r}' , \mathbf{v}' in rest frame of the inner binary is

$$\mathbf{y}'_1 = (X_2 + \delta) \mathbf{r}' + (\mathbf{V}_{\text{CM}} \cdot \mathbf{r}') \mathbf{p}' / m, \\ \mathbf{y}'_2 = (-X_1 + \delta) \mathbf{r}' + (\mathbf{V}_{\text{CM}} \cdot \mathbf{r}') \mathbf{p}' / m, \quad (\text{E7})$$

where δ is the 1PN quantity already defined in Eq. (E3) (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 \mathbf{y}'_1 , \mathbf{y}'_2 , \mathbf{r}' , and \mathbf{v}' are evaluated at the same time t' .

This result allows one to express the three-body frame coordinates \mathbf{y}_A in terms of the relative coordinates in the rest frame of the inner binary, \mathbf{r}' , \mathbf{p}' ,⁸ which can in turn be expressed in terms of *intrinsic* contact elements using Eq. (12).⁹ We obtain

⁸Note that \mathbf{p}' is defined through the relative velocity \mathbf{v}' , the t' derivative of \mathbf{r}' .

⁹Although Eq. (12) 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 \mathbf{r}' and \mathbf{p}' .

$$\begin{aligned}
\mathbf{y}_1 &= \mathbf{Y}_{\text{CM}} + \mathbf{V}_{\text{CM}}(t - t_0) + (X_2 + \delta)\mathbf{r}' + X_2(\mathbf{V}_{\text{CM}} \cdot \mathbf{r}') \left[(X_1 - X_2) \frac{\mathbf{p}'}{\mu} - \frac{\mathbf{V}_{\text{CM}}}{2} \right], \\
\mathbf{y}_2 &= \mathbf{Y}_{\text{CM}} + \mathbf{V}_{\text{CM}}(t - t_0) + (-X_1 + \delta)\mathbf{r}' - X_1(\mathbf{V}_{\text{CM}} \cdot \mathbf{r}') \left[(X_1 - X_2) \frac{\mathbf{p}'}{\mu} - \frac{\mathbf{V}_{\text{CM}}}{2} \right],
\end{aligned} \tag{E8}$$

where \mathbf{r}' and \mathbf{p}' are evaluated at time t , while \mathbf{Y}_{CM} and \mathbf{V}_{CM} are evaluated at t_0 .

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

$$\mathbf{r} = \mathbf{r}' - (\mathbf{V}_{\text{CM}} \cdot \mathbf{r}') \left[\frac{\mathbf{V}_{\text{CM}}}{2} + (X_2 - X_1) \frac{\mathbf{p}'}{\mu} \right]. \tag{E9}$$

The final step is just to evaluate the above expressions (E8) and their time derivatives at $t = t_0$ and to express \mathbf{r}' , \mathbf{p}' in terms of the contact elements as in Eq. (12). To avoid clutter, in the main text and the rest of the appendixes we will suppress the primed label on \mathbf{r}' , \mathbf{p}' .

APPENDIX F: QUADRUPOLE-SQUARED TERMS

In this appendix we use the methodology developed in Appendix C to compute the so-called quadrupole-squared terms presented in [46]. These purely Newtonian contributions to the evolution of a hierarchical system come from deviations to the adiabatic approximation as well as back-reaction 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 C, 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 kinds 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; and (iv) fluctuations of the outer planetary elements on an outer binary timescale. In all these cases, a generalization of the methodology developed in Appendix C 1 shows that the cross terms in L_S (C12) scale as [see, in particular, Eq. (C10) and Eq. (C6)]

$$\text{(i)} \quad \frac{\mathcal{F}^2}{L} \frac{L^3}{G_N^2 m^2} \sim \mathcal{F}^2 \frac{a}{G_N m}, \tag{F1}$$

$$\text{(ii)} \quad \frac{\mathcal{F}^2}{L_3} \frac{L^3}{G_N^2 m^2} \sim \mathcal{F}^2 \frac{a}{G_N m} \varepsilon^{1/2}, \tag{F2}$$

$$\text{(iii)} \quad \frac{\mathcal{F}^2}{L} \frac{L_3^3}{G_N^2 m^2} \sim \mathcal{F}^2 \frac{a}{G_N m} \varepsilon^{-3/2}, \tag{F3}$$

$$\text{(iv)} \quad \frac{\mathcal{F}^2}{L_3} \frac{L_3^3}{G_N^2 m^2} \sim \mathcal{F}^2 \frac{a}{G_N m} \varepsilon^{-1}. \tag{F4}$$

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. [46]). This scaling is perfectly valid for the quadrupole-squared terms that 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 that explains that the dominant contribution to cross terms comes from case (i) as computed in Appendix C. 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 $\mathcal{R}_{\text{1PN}} = \mathcal{R}_{v^2} + \mathcal{R}_{v^2 \varepsilon^{1/2}} + \mathcal{R}_{v^2 \varepsilon} + \mathcal{R}_{v^2 \varepsilon^{3/2}} + \mathcal{R}_{v^2 \varepsilon^2}$. This splitting, as already emphasized in Appendix C 2, comes from introducing the center-of-mass coordinates (16) in the EIH Lagrangian (4). With our center-of-mass choice, $\mathcal{R}_{v^2 \varepsilon^{1/2}}$, $\mathcal{R}_{v^2 \varepsilon}$, and $\mathcal{R}_{v^2 \varepsilon^{3/2}}$ vanish after averaging on the inner binary timescale. Thus, the only term featuring a nontrivial 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 C.

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 timescales 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 timescales. Therefore we will generally write $X = X_{\bar{L}} + X_{\bar{S}}$, indicating a near-identity transformation having the period of the outer binary as the reference timescale. 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 $e^{1/2}$ factor.

At this point we can already check that the quadrupole-square cross terms Lagrangian $\mathcal{L}_{\bar{S}}$ scales as

$$\frac{\mathcal{L}_{\bar{S}}}{\mathcal{L}_{\bar{L}}} \sim \frac{n_3}{n} \sim \sqrt{\frac{Ma^3}{ma_3^3}}. \quad (\text{F5})$$

Thus, despite the quadrupole-squared terms are formally a $e^{3/2}$ 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. (34) to Newtonian order as

$$\begin{aligned} \mathcal{R}_{Q3} &= \frac{3G_N m_3}{2\mu} Q^{ij} \frac{N_i N_j}{R^3}, \\ Q^{ij} &= \frac{\mu a^2}{2} \left((1 + 4e^2) \alpha^i \alpha^j + (1 - e^2) \beta^i \beta^j - \frac{2 + 3e^2}{3} \delta^{ij} \right), \end{aligned} \quad (\text{F6})$$

and its average-free part, taken with respect to the long-timescale part of the outer mean anomaly $u_{3\bar{L}}$, as

$$\mathcal{F}_{Q3} = \mathcal{R}_{Q3} - \frac{1}{2\pi} \int_0^{2\pi} du_{3\bar{L}} \mathcal{R}_{Q3}, \quad (\text{F7})$$

where we recall that \mathbf{R} is the distance vector of the outer orbit, $R = |\mathbf{R}|$ and $\mathbf{N} = \mathbf{R}/R$. With this notation, using the equations of motion and expanding the perturbing function, we obtain

$$\begin{aligned} \mathcal{L}_{\bar{S}} &= L_{\bar{S}} \dot{u}_{\bar{S}} + G_{\bar{S}} \dot{\omega}_{\bar{S}} + H_{\bar{S}} \dot{\Omega}_{\bar{S}} + \sum_X \frac{\partial \mathcal{F}_{Q3}}{\partial X} X_{\bar{S}} \\ &= \text{Af} \left(\int^{u_{3\bar{L}}} ds \frac{\partial \mathcal{F}_{Q3}}{\partial G} \right) \frac{\partial \mathcal{F}_{Q3}}{\partial \omega} \\ &\quad + \text{Af} \left(\int^{u_{3\bar{L}}} ds \frac{\partial \mathcal{F}_{Q3}}{\partial H} \right) \frac{\partial \mathcal{F}_{Q3}}{\partial \Omega}. \end{aligned} \quad (\text{F8})$$

Then, performing the averages with respect to the long-timescale part of the outer mean anomaly, we find

$$\begin{aligned} \langle \mathcal{L}_{\bar{S}} \rangle &= -\frac{9G_N^2 m_3^2}{4\mu^2} \left(\frac{\partial Q^{ij}}{\partial \omega} \frac{\partial Q^{kl}}{\partial G} + \frac{\partial Q^{ij}}{\partial \Omega} \frac{\partial Q^{kl}}{\partial H} \right) \mathcal{A}_{ij;kl} \\ &\equiv \mathcal{B}^{ij;kl} \mathcal{A}_{ij;kl}, \end{aligned} \quad (\text{F9})$$

where the tensor $\mathcal{A}_{ij;kl}$ is a sort of variance given by

$$\mathcal{A}_{ij;kl} = \left\langle \left[\frac{N_i N_j}{R^3} - \left\langle \frac{N_i N_j}{R^3} \right\rangle \right] \times \int dt \left[\frac{N_k N_l}{R^3} - \left\langle \frac{N_k N_l}{R^3} \right\rangle \right] \right\rangle. \quad (\text{F10})$$

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. (F9). We will use the derivatives given in Appendix D, Eqs. (D2) and (D13). These derivatives generically depend on the angles Ω , ω , and ι . However, such angles cannot remain in the final result for $\langle \mathcal{L}_{\bar{S}} \rangle$, 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 nontrivial 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 D.

We now go on for the final computation. In the $\mathcal{B}^{ij;kl}$ tensor, the dependence on the angle ι nicely factors out:

$$\begin{aligned} \frac{1}{\sin \iota} \frac{\partial Q^{kl}}{\partial \iota} \left(\cos \iota \frac{\partial Q^{ij}}{\partial \omega} - \frac{\partial Q^{ij}}{\partial \Omega} \right) &= \frac{\mu^2 a^4}{4} [(1 + 4e^2) \cos \omega (\alpha^i \gamma^j + \text{sym}) - (1 - e^2) \sin \omega (\beta^i \gamma^j + \text{sym})] \\ &\quad \times [(1 + 4e^2) \sin \omega (\alpha^k \gamma^l + \text{sym}) + (1 - e^2) \cos \omega (\beta^k \gamma^l + \text{sym})]. \end{aligned} \quad (\text{F11})$$

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 antisymmetrization of the above expression, we are led to

$$\frac{\mu^2 a^4 (1 - e^2)(1 + 4e^2)}{8} ((\alpha^i \gamma^j + \text{sym})(\gamma^k \beta^l + \text{sym}) - (\gamma^i \beta^j + \text{sym})(\alpha^k \gamma^l + \text{sym})), \quad (\text{F12})$$

so that all angular dependence drops out. Putting all together, we find the final expression for $\mathcal{B}_{ij;kl}$:

$$\mathcal{B}_{ij;kl} = \frac{9G_N^2 m_3^2 a^4}{32\sqrt{G_N m a}(1 - e^2)} [-20e^2(1 - e^2)(\alpha^i \beta^j + \text{sym})(4\alpha^k \alpha^l - \beta^k \beta^l - \delta^{kl}) \\ + (1 - e^2)(1 + 4e^2)((\alpha^i \gamma^j + \text{sym})(\beta^k \gamma^l + \text{sym}) - (\beta^i \gamma^j + \text{sym})(\alpha^k \gamma^l + \text{sym}))], \quad (\text{F13})$$

while the other tensor $\mathcal{A}_{ij;kl}$ is given by

$$\mathcal{A}^{ij;kl} = \frac{g_1(e_3)}{\sqrt{G_N M a_3^9}} [\alpha_3^k \alpha_3^l (\alpha_3^i \beta_3^j + \text{sym}) - \alpha_3^i \alpha_3^j (\alpha_3^k \beta_3^l + \text{sym})] + \frac{g_2(e_3)}{\sqrt{G_N M a_3^9}} [\beta_3^i \beta_3^j (\alpha_3^k \beta_3^l + \text{sym}) - \beta_3^k \beta_3^l (\alpha_3^i \beta_3^j + \text{sym})], \quad (\text{F14})$$

where the two dimensionless functions of the outer eccentricity e_3 are given by

$$g_1(e_3) = \frac{1}{48e_3^2(1 - e_3^2)^{7/2}} [4(1 - \sqrt{1 - e_3^2}) + e_3^2(-8 + 9\sqrt{1 - e_3^2}) + e_3^4(4 + 5\sqrt{1 - e_3^2})], \quad (\text{F15})$$

$$g_2(e_3) = \frac{1}{48e_3^2(1 - e_3^2)^{5/2}} [4(-1 + \sqrt{1 - e_3^2}) + e_3^2(4 + \sqrt{1 - e_3^2})]. \quad (\text{F16})$$

Plugging the expressions of the basis vectors in terms of osculating angles, we have checked that our formula (F9) exactly recovers the LPE with quadrupole-squared terms derived in [46].

APPENDIX G: 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 one 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 [31,33] concerning the evolution of conserved quantities.

We will now write down explicitly the relation between these two sets of elements to 1PN order, following [69]. As discussed in Sec. II A, 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:

$$\mathbf{p} = \mathbf{v} + \frac{\partial \mathcal{R}_{1\text{PN}}}{\partial \mathbf{v}} \\ = \mathbf{v} + \left(\frac{1 - 3\nu}{2} v^2 + \frac{G_N m (3 + \nu)}{r} \right) \mathbf{v} + \frac{G_N m \nu}{r} (\mathbf{n} \cdot \mathbf{v}) \mathbf{n}, \quad (\text{G1})$$

where we recall that in all appendixes we use the convention of dividing the Lagrangian by the reduced mass μ , so that \mathbf{p} and \mathbf{v} have the 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{i} . Indeed, since the momentum \mathbf{p} is still in the plane of the motion, the definition of $\tilde{\Omega}$ and \tilde{i} do not get affected by the difference between momentum and velocity. Thus,

$$\tilde{\Omega} = \Omega, \quad (\text{G2})$$

$$\tilde{i} = i. \quad (\text{G3})$$

Let us start the nontrivial computations with the semimajor axis. We have the following definitions:

$$-\frac{G_N m}{2\tilde{a}} = \frac{v^2}{2} - \frac{G_N m}{r}, \quad (\text{G4})$$

$$-\frac{G_N m}{2a} = \frac{p^2}{2} - \frac{G_N m}{r}. \quad (\text{G5})$$

Combining Eqs. (G1), (G4), and (G5) we get to $\tilde{a} = a + \Delta a$ with

$$\Delta a = -\frac{2G_N m}{(1 - \tilde{e} \cos \tilde{\eta})^2} \left[\frac{1}{2} (1 - 3\nu)(1 + \tilde{e} \cos \tilde{\eta})^2 \\ + (3 + \nu)(1 + \tilde{e} \cos \tilde{\eta}) + \frac{\nu \tilde{e}^2 \sin^2 \tilde{\eta}}{1 - \tilde{e} \cos \tilde{\eta}} \right]. \quad (\text{G6})$$

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

$$|\mathbf{r} \times \mathbf{v}| = \sqrt{G_N m \tilde{a}(1 - \tilde{e}^2)}, \quad (\text{G7})$$

$$|\mathbf{r} \times \mathbf{p}| = \sqrt{G_N m a(1 - e^2)}. \quad (\text{G8})$$

Splitting $\tilde{e} = e + \Delta e$ we get to

$$\begin{aligned} \Delta e &= \frac{G_N m (1 - \tilde{e}^2)}{\tilde{a}(1 - \tilde{e} \cos \tilde{\eta})^3} [\cos \tilde{\eta}(1 - \tilde{e} \cos \tilde{\eta}) \\ &\quad \times (-7 + \nu - \tilde{e}(1 - 3\nu) \cos \tilde{\eta}) - \tilde{e} \nu \sin^2 \tilde{\eta}]. \end{aligned} \quad (\text{G9})$$

Finally, it 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:

$$r = \frac{a(1 - e^2)}{1 + e \cos f} = \frac{\tilde{a}(1 - \tilde{e}^2)}{1 + \tilde{e} \cos \tilde{f}}. \quad (\text{G10})$$

This gives $\tilde{f} = f + \Delta f$ with

$$\begin{aligned} \Delta f &= \frac{G_N m \sqrt{1 - \tilde{e}^2} \sin \tilde{\eta}}{2\tilde{a} \tilde{e}(1 - \tilde{e} \cos \tilde{\eta})^3} [14 - \tilde{e}^2 - 2\nu + 5\tilde{e}^2 \nu - 6\tilde{e}(2 + \nu) \\ &\quad \times \cos \tilde{\eta} - \tilde{e}^2(1 - 3\nu) \cos 2\tilde{\eta}]. \end{aligned} \quad (\text{G11})$$

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

$$\begin{aligned} \Delta u &= -\frac{G_N m}{8\tilde{a} \tilde{e}(1 - \tilde{e} \cos \tilde{\eta})^3} [2(-6\tilde{e}^4 + \tilde{e}^2(1 - 15\nu) - 4(7 - \nu)) \sin \tilde{\eta} + 2\tilde{e}(6(2 + \nu) + 2\tilde{e}^2(7 + 2\nu) - \tilde{e}^4(1 - 3\nu)) \sin 2\tilde{\eta} \\ &\quad + 2\tilde{e}^2(1 - 3\nu - \tilde{e}^2(6 + 4\nu)) \sin 3\tilde{\eta} - \tilde{e}^5(1 - 3\nu) \sin 4\tilde{\eta}]. \end{aligned} \quad (\text{G12})$$

To sum up our results, here are all the modifications to the osculating elements:

$$\Delta a = -\frac{2G_N m}{(1 - \tilde{e} \cos \tilde{\eta})^2} \left[\frac{1}{2} (1 - 3\nu)(1 + \tilde{e} \cos \tilde{\eta})^2 + (3 + \nu)(1 + \tilde{e} \cos \tilde{\eta}) + \frac{\nu \tilde{e}^2 \sin^2 \tilde{\eta}}{1 - \tilde{e} \cos \tilde{\eta}} \right], \quad (\text{G13})$$

$$\Delta e = \frac{G_N m (1 - \tilde{e}^2)}{\tilde{a}(1 - \tilde{e} \cos \tilde{\eta})^3} [\cos \tilde{\eta}(1 - \tilde{e} \cos \tilde{\eta})(-7 + \nu - \tilde{e}(1 - 3\nu) \cos \tilde{\eta}) - \tilde{e} \nu \sin^2 \tilde{\eta}], \quad (\text{G14})$$

$$\Delta \omega = -\frac{G_N m \sqrt{1 - \tilde{e}^2} \sin \tilde{\eta}}{2\tilde{a} \tilde{e}(1 - \tilde{e} \cos \tilde{\eta})^3} [14 - \tilde{e}^2 - 2\nu + 5\tilde{e}^2 \nu - 6\tilde{e}(2 + \nu) \cos \tilde{\eta} - \tilde{e}^2(1 - 3\nu) \cos 2\tilde{\eta}], \quad (\text{G15})$$

$$\begin{aligned} \Delta u &= -\frac{G_N m}{8\tilde{a} \tilde{e}(1 - \tilde{e} \cos \tilde{\eta})^3} [2(-6\tilde{e}^4 + \tilde{e}^2(1 - 15\nu) - 4(7 - \nu)) \sin \tilde{\eta} + 2\tilde{e}(6(2 + \nu) + 2\tilde{e}^2(7 + 2\nu) - \tilde{e}^4(1 - 3\nu)) \sin 2\tilde{\eta} \\ &\quad + 2\tilde{e}^2(1 - 3\nu - \tilde{e}^2(6 + 4\nu)) \sin 3\tilde{\eta} - \tilde{e}^5(1 - 3\nu) \sin 4\tilde{\eta}], \end{aligned} \quad (\text{G16})$$

$$\Delta \Omega = \Delta i = 0, \quad (\text{G17})$$

$$\Delta \boldsymbol{\alpha} = \boldsymbol{\beta} \Delta \boldsymbol{\omega}, \quad \Delta \boldsymbol{\beta} = -\boldsymbol{\alpha} \Delta \boldsymbol{\omega}. \quad (\text{G18})$$

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 C), 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

$$\tilde{e}_L = e_L + \langle \Delta e \rangle, \quad \tilde{e}_S = e_S + \Delta e - \langle \Delta e \rangle, \quad (\text{G19})$$

and similarly for the other osculating elements. This gives the final relation for the orbit-averaged contact and orbital elements:

$$\begin{aligned}
\tilde{a}_L &= a_L + G_N m \left[9 - \frac{16}{\sqrt{1-\tilde{e}_L^2}} - \nu \left(5 - \frac{6}{\sqrt{1-\tilde{e}_L^2}} \right) \right], \\
\tilde{e}_L &= e_L - \frac{G_N m \tilde{e}_L}{\tilde{a}_L} (8 - 3\nu) \frac{\sqrt{1-\tilde{e}_L^2}}{1+\sqrt{1-\tilde{e}_L^2}}, \\
\tilde{\omega}_L &= \omega_L, \quad \tilde{\Omega}_L = \Omega_L, \quad \tilde{l}_L = l_L, \quad \tilde{u}_L = u_L.
\end{aligned}
\tag{G20}$$

We can now describe the main physical effects stemming from this post-Newtonian shift of the osculating elements:

- (i) The first and most important point to notice is that the shifts written in Eq. (G20) 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,¹⁰ as it happens for quadrupolar and post-Newtonian perturbations in the LPE; rather, the two sets of osculating elements will differ by a quantity that 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.
- (ii) 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. (G20) in the kinetic term of the Lagrangian (C1) shifts the canonical momenta G and H as

$$\frac{\Delta G}{\tilde{G}} = \frac{\Delta H}{\tilde{H}} = \frac{e \langle \Delta e \rangle}{1-\tilde{e}^2} - \frac{\langle \Delta a \rangle}{2\tilde{a}} = \frac{G_N m}{2\tilde{a}} (7-\nu). \tag{G21}$$

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

$$\begin{aligned}
\mathbf{J}_{\text{1PN}} &= \mu \sqrt{G_N m \tilde{a} (1-\tilde{e}^2)} \tilde{\boldsymbol{\gamma}} \\
&= \mu \sqrt{G_N m \tilde{a} (1-\tilde{e}^2)} \boldsymbol{\gamma} \left(1 + \frac{G_N m}{2\tilde{a}} (7-\nu) \right).
\end{aligned}
\tag{G22}$$

On the other hand, the above expression corresponds to the 1PN expression of the conserved angular

¹⁰This is, as long as the elements are computed within the time interval in which our effective field theory is valid.

momentum that one can find in, e.g., [70], averaged over one orbit of the binary system. This is a nontrivial check of the validity of our computation.

- (iii) 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 that is conserved in time, but it is related to the osculating \tilde{a} with a formula involving the eccentricity, Eq. (G20). 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 [31,33] 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 [31,33] 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 [31,33]. Indeed, the following averages differ at 1PN quadrupolar order:

$$\frac{d\langle a \rangle}{dt} \neq \left\langle \frac{da}{dt} \right\rangle. \tag{G23}$$

The left-hand side (LHS) of this equation corresponds to the quantity that we compute in this article. On the other hand, the variation computed in [31,33] corresponds to the RHS of this equation, which by definition is equal to $(a(T) - a(0))/T$. Thus, the results of [31,33] 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 [31,33] proves exact conservation of *initial* PN energy and angular momentum, we are able to prove conservation of an *averaged* PN energy and angular momentum.

- [1] S. M. Ransom *et al.*, A millisecond pulsar in a stellar triple system, *Nature (London)* **505**, 520 (2014).
- [2] Travis Robson, Neil J. Cornish, Nicola Tamanini, and Silvia Toonen, Detecting hierarchical stellar systems with lisa, *Phys. Rev. D* **98**, 064012 (2018).
- [3] 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**, 063014 (2017).
- [4] 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, *Mon. Not. R. Astron. Soc.* **439**, 1079 (2014).
- [5] Naoki Seto, Highly Eccentric Kozai Mechanism and Gravitational-Wave Observation for Neutron Star Binaries, *Phys. Rev. Lett.* **111**, 061106 (2013).
- [6] Alexander P. Stephan, Smadar Naoz, Andrea M. Ghez, Gunther Witzel, Breann N. Sitarski, Tuan Do, and Bence Kocsis, Merging binaries in the galactic center: The eccentric kozai–lidov mechanism with stellar evolution, *Mon. Not. R. Astron. Soc.* **460**, 3494 (2016).
- [7] David Merritt, Tal Alexander, Seppo Mikkola, and Clifford M. Will, Stellar dynamics of extreme-mass-ratio inspirals, *Phys. Rev. D* **84**, 044024 (2011).
- [8] 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, *Astrophys. J.* **856**, 140 (2018).
- [9] Smadar Naoz, Bence Kocsis, Abraham Loeb, and Nicolas Yunes, Resonant post-Newtonian eccentricity excitation in hierarchical three-body systems, *Astrophys. J.* **773**, 187 (2013).
- [10] 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, *Astrophys. J.* **901**, 125 (2020).
- [11] David V. Martin and Amaury H. M. J. Triaud, Circumbinary planets—why they are so likely to transit, *Mon. Not. R. Astron. Soc.* **449**, 781 (2015).
- [12] David V. Martin and Amaury H. M. J. Triaud, Kozai–Lidov cycles towards the limit of circumbinary planets, *Mon. Not. R. Astron. Soc.* **455**, L46 (2015).
- [13] Tsevi Mazeh, Yuval Krymowolski, and Gady Rosenfeld, The high eccentricity of the planet orbiting 16 cygni b, *Astrophys. J.* **477**, L103 (1997).
- [14] T. Mazeh and J. Shaham, The orbital evolution of close triple systems: The binary eccentricity, *Astron. Astrophys.* **77**, 145 (1979), <https://ui.adsabs.harvard.edu/abs/1979A&A....77..145M>.
- [15] Bin Liu, Diego J. Muñoz, and Dong Lai, Suppression of extreme orbital evolution in triple systems with short-range forces, *Mon. Not. R. Astron. Soc.* **447**, 747 (2014).
- [16] Gongjie Li, Smadar Naoz, Bence Kocsis, and Abraham Loeb, Implications of the eccentric Kozai-Lidov mechanism for stars surrounding supermassive black hole binaries, *Mon. Not. R. Astron. Soc.* **451**, 1341 (2015).
- [17] Gongjie Li, Smadar Naoz, Bence Kocsis, and Abraham Loeb, Eccentricity growth and orbit flip in near-coplanar hierarchical three-body systems, *Astrophys. J.* **785**, 116 (2014).
- [18] Smadar Naoz, The eccentric kozai-lidov effect and its applications, *Annu. Rev. Astron. Astrophys.* **54**, 441 (2016).
- [19] Yoshihide Kozai, Secular perturbations of asteroids with high inclination and eccentricity, *Astron. J.* **67**, 591 (1962).
- [20] M. L. Lidov, The evolution of orbits of artificial satellites of planets under the action of gravitational perturbations of external bodies, *Planet. Space Sci.* **9**, 719 (1962).
- [21] Adrien Kuntz, Precession resonances in hierarchical triple systems, *Phys. Rev. D* **105**, 024017 (2022).
- [22] Omer Blaes, Man Hoi Lee, and Aristotle Socrates, The kozai mechanism and the evolution of binary supermassive black holes, *Astrophys. J.* **578**, 775 (2002).
- [23] Loren Hoffman and Abraham Loeb, Dynamics of triple black hole systems in hierarchically merging massive galaxies, *Mon. Not. R. Astron. Soc.* **377**, 957 (2007).
- [24] 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. R. Astron. Soc.* **428**, 2295 (2013).
- [25] Eric B. Ford, Boris Kozinsky, and Frederic A. Rasio, Secular evolution of hierarchical triple star systems, *Astrophys. J.* **535**, 385 (2000).
- [26] Pablo Galaviz and Bernd Brügmann, Characterization of the gravitational wave emission of three black holes, *Phys. Rev. D* **83**, 084013 (2011).
- [27] Pablo Galaviz, Stability and chaos of hierarchical three black hole configurations, *Phys. Rev. D* **84**, 104038 (2011).
- [28] Matteo Bonetti, Francesco Haardt, Alberto Sesana, and Enrico Barausse, Post-Newtonian evolution of massive black hole triplets in galactic nuclei—I. Numerical implementation and tests, *Mon. Not. R. Astron. Soc.* **461**, 4419 (2016).
- [29] Carlos O. Lousto and Hiroyuki Nakano, Three-body equations of motion in successive post-Newtonian approximations, *Classical Quantum Gravity* **25**, 195019 (2008).
- [30] Priti Gupta, Haruka Suzuki, Hirotada Okawa, and Kei-ichi Maeda, Gravitational waves from hierarchical triple systems with Kozai-Lidov oscillation, *Phys. Rev. D* **101**, 104053 (2020).
- [31] Clifford M. Will, Incorporating post-Newtonian effects in n -body dynamics, *Phys. Rev. D* **89**, 044043 (2014).
- [32] Clifford M. Will, New General Relativistic Contribution to Mercury’s Perihelion Advance, *Phys. Rev. Lett.* **120**, 191101 (2018).
- [33] Clifford M. Will, Post-Newtonian effects in N -body dynamics: Conserved quantities in hierarchical triple systems, *Classical Quantum Gravity* **31**, 244001 (2014).
- [34] Halston Lim and Carl L. Rodriguez, Relativistic three-body effects in hierarchical triples, *Phys. Rev. D* **102**, 064033 (2020).
- [35] Adrien Kuntz, Francesco Serra, and Enrico Trinchineri, Effective two-body approach to the hierarchical three-body problem, *Phys. Rev. D* **104**, 024016 (2021).
- [36] Rafael A. Porto, The effective field theorist’s approach to gravitational dynamics, *Phys. Rep.* **633**, 1 (2016).
- [37] Walter D. Goldberger and Ira Z. Rothstein, An effective field theory of gravity for extended objects, *Phys. Rev. D* **73**, 104029 (2006).

- [38] James A. Murdock, *Perturbations: Theory and Methods* (Wiley, New York, 1991).
- [39] 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, *Astrophys. J.* **887**, 210 (2019).
- [40] 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 (2020).
- [41] Bin Liu, Dong Lai, and Yi-Han Wang, Binary mergers near a supermassive black hole: Relativistic effects in triples, *Astrophys. J.* **883**, L7 (2019).
- [42] Cezary Migaszewski and Krzysztof Goździewski, Secular dynamics of coplanar, non-resonant planetary system under the general relativity and quadrupole moment perturbations, *Mon. Not. R. Astron. Soc.* **392**, 2 (2009).
- [43] A. J. Hanson and T. Regge, The relativistic spherical top, *Ann. Phys. (N.Y.)* **87**, 498 (1974).
- [44] Michele Levi and Jan Steinhoff, Spinning gravitating objects in the effective field theory in the post-Newtonian scheme, *J. High Energy Phys.* **09** (2015) 219.
- [45] V. Brumberg, *Essential Relativistic Celestial Mechanics* (Adam Hilger, Bristol, England and New York, 1991), p. 271.
- [46] Clifford M. Will, Higher-order effects in the dynamics of hierarchical triple systems: Quadrupole-squared terms, *Phys. Rev. D* **103**, 063003 (2021).
- [47] Barak Kol and Michael Smolkin, Non-relativistic gravitation: From Newton to Einstein and back, *Classical Quantum Gravity* **25**, 145011 (2008).
- [48] Barak Kol and Michael Smolkin, Einstein's action and the harmonic gauge in terms of Newtonian fields, *Phys. Rev. D* **85**, 044029 (2012).
- [49] A. Einstein, L. Infeld, and B. Hoffmann, The gravitational equations and the problem of motion, *Ann. Math.* **39**, 65 (1938).
- [50] Charles W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman, San Francisco, 1973).
- [51] Jan Sanders, Ferdinand Verhulst, and James Murdock, *Averaging Methods in Nonlinear Dynamical Systems* (Springer, New York, 2007).
- [52] Walter D. Goldberger, in *Les Houches Lectures on Effective Field Theories and Gravitational Radiation* (2007), <https://ui.adsabs.harvard.edu/abs/2007hep.ph....1129G>.
- [53] Walter D. Goldberger and Andreas Ross, Gravitational radiative corrections from effective field theory, *Phys. Rev. D* **81**, 124015 (2010).
- [54] Smadar Naoz, Will M. Farr, Yoram Lithwick, Frederic A. Rasio, and Jean Teyssandier, Secular dynamics in hierarchical three-body systems, *Mon. Not. R. Astron. Soc.* **431**, 2155 (2013).
- [55] Lisa Randall and Zhong-Zhi Xianyu, An analytical portrait of binary mergers in hierarchical triple systems, *Astrophys. J.* **864**, 134 (2018).
- [56] 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, *Mon. Not. R. Astron. Soc.* **474**, 5672 (2017).
- [57] Fabio Antonini and Frederic A. Rasio, Merging black hole binaries in galactic nuclei: Implications for Advanced-LIGO detections, *Astrophys. J.* **831**, 187 (2016).
- [58] Daniel Fabrycky and Scott Tremaine, Shrinking binary and planetary orbits by kozai cycles with tidal friction, *Astrophys. J.* **669**, 1298 (2007).
- [59] F. Loebbert, J. Plefka, C. Shi, and T. Wang, Three-body effective potential in general relativity at second post-Minkowskian order and resulting post-Newtonian contributions, *Phys. Rev. D* **103**, 064010 (2021).
- [60] Matteo Bonetti, Enrico Barausse, Guillaume Faye, Francesco Haardt, and Alberto Sesana, About gravitational-wave generation by a three-body system, *Classical Quantum Gravity* **34**, 215004 (2017).
- [61] C. R. T. Jones and M. Solon, *Scattering Amplitudes and N-Body Post-Minkowskian Hamiltonians in General Relativity and Beyond* (2022), <https://ui.adsabs.harvard.edu/abs/2022arXiv220802281J>.
- [62] C. C. Gerry, Coherent states and the Kepler-Coulomb problem, *Phys. Rev. A* **33**, 6 (1986).
- [63] Lisa Randall and Zhong-Zhi Xianyu, Observing eccentricity oscillations of binary black holes in LISA, *arXiv:1902.08604*.
- [64] Rohit S. Chandramouli and Nicolás Yunes, Ready-to-use analytic model for gravitational waves from a hierarchical triple with Kozai-Lidov oscillations, *Phys. Rev. D* **105**, 064009 (2022).
- [65] Carl Bender and Steven Orszag, *Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory* (Springer, New York, 1999), Vol. 1.
- [66] N. Van Kampen, Elimination of fast variables, *Phys. Rep.* **124**, 69 (1985).
- [67] Yusuke Hagihara, *Celestial Mechanics, Volume II: Perturbation Theory* (MIT Press, Cambridge, MA, 1972).
- [68] C. R. Galley and I. Z. Rothstein, Deriving analytic solutions for compact binary inspirals without recourse to adiabatic approximations, *Phys. Rev. D* **10**, 104054 (2017).
- [69] D. L. Richardson and T. J. Kelly, Two-body motion in the post-Newtonian approximation, *Celest. Mech.* **43**, 193 (1988).
- [70] Thibault Damour and Nathalie Deruelle, General relativistic celestial mechanics of binary systems. I. The post-Newtonian motion, *Ann. I. H. P.: Phys. Theor.* **43**, 107 (1985), <https://ui.adsabs.harvard.edu/abs/1985AIHS...43..107D>.