

**Detecting stochastic gravitational waves with binary resonance**Diego Blas<sup>1,2,3</sup> and Alexander C. Jenkins<sup>3,\*</sup><sup>1</sup>*Grup de Física Teòrica, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain*<sup>2</sup>*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra, Spain*<sup>3</sup>*Theoretical Particle Physics and Cosmology Group, Physics Department, King's College London, University of London, Strand, London WC2R 2LS, United Kingdom*

(Received 7 September 2021; accepted 15 February 2022; published 11 March 2022)

LIGO and Virgo have initiated the era of gravitational-wave (GW) astronomy; but in order to fully explore GW frequency spectrum, we must turn our attention to innovative techniques for GW detection. One such approach is to use binary systems as dynamical GW detectors by studying the subtle perturbations to their orbits caused by impinging GWs. We present a powerful new formalism for calculating the orbital evolution of a generic binary coupled to a stochastic background of GWs, deriving from first principles a secularly-averaged Fokker-Planck equation which fully characterizes the statistical evolution of all six of the binary's orbital elements. We also develop practical tools for numerically integrating this equation, and derive the necessary statistical formalism to search for GWs in observational data from binary pulsars and laser-ranging experiments.

DOI: [10.1103/PhysRevD.105.064021](https://doi.org/10.1103/PhysRevD.105.064021)**I. INTRODUCTION**

We have entered the era of gravitational-wave (GW) astronomy. The advanced laser interferometers LIGO [1], Virgo [2], and KAGRA [3] now form a global GW observatory network [4], which has already proved its capabilities by detecting dozens of GW signals in the 10–1000 Hz band from coalescing compact binaries [5]. Concurrently, GW searches in the nHz band with pulsar timing arrays (PTAs) [6–9] are, after more than a decade of continuous operation, widely expected to make their first detections in the coming few years, with the NANOGrav Collaboration having possibly already seen the first hints of a GW detection in their 12.5-yr dataset [10]. By the mid-2030s, these experiments are expected to be superseded by third generation interferometers such as Einstein Telescope [11] and PTA campaigns by next-generation radio telescopes such as the Square Kilometre Array [12], as well as the space-based interferometer LISA [13], which will survey the as-yet-unexplored mHz frequency band. Further proposals to cover the intermediate frequency band between LISA and LIGO/Virgo/KAGRA include the atom interferometers AION [14] and MAGIS [15].

These experiments each hold enormous scientific potential, and together will probe a dizzying range of exotic astrophysical and cosmological phenomena throughout the

history of the Universe. However, practical and technical limitations restrict the sensitivity of each experiment to a narrow frequency band, leaving broad swathes of the GW frequency spectrum essentially unexplored. These gaps in the GW spectrum could contain signals which are inaccessible to any current or planned GW observatory; perhaps the clearest example is the GW signal expected from a cosmological first-order phase transition [16–18], which is sharply peaked and could, for a broad range of the physical parameter space, be missed by all of the GW experiments listed above. There is also the distinct possibility of such unexplored frequency bands containing unexpected signals, with the potential for exciting discoveries beyond just those models that have been proposed in the literature.

This problem motivates us to explore alternatives that can bridge these gaps in the GW spectrum. One possibility is to study the interaction between GWs and astronomical binary systems. Rather than searching for oscillations in the proper distance between the test masses in an interferometer (as in LIGO/Virgo/KAGRA, etc.), or between pulsars and the Earth (as in PTAs), we consider here the GW-induced oscillations between two freely-falling masses in a gravitationally-bound orbit. While these oscillations are extremely challenging to observe directly for any realistic binary, they can leave lasting imprints on the binary's orbit, particularly if they occur at an integer multiple of the binary's orbital frequency, as this causes the perturbations to be resonantly amplified. For long-duration GW signals, these imprints accumulate over time, eventually giving rise

\*alex.jenkins@ucl.ac.uk

Present address: Department of Physics and Astronomy, University College London, London WC1E 6BT, United Kingdom.

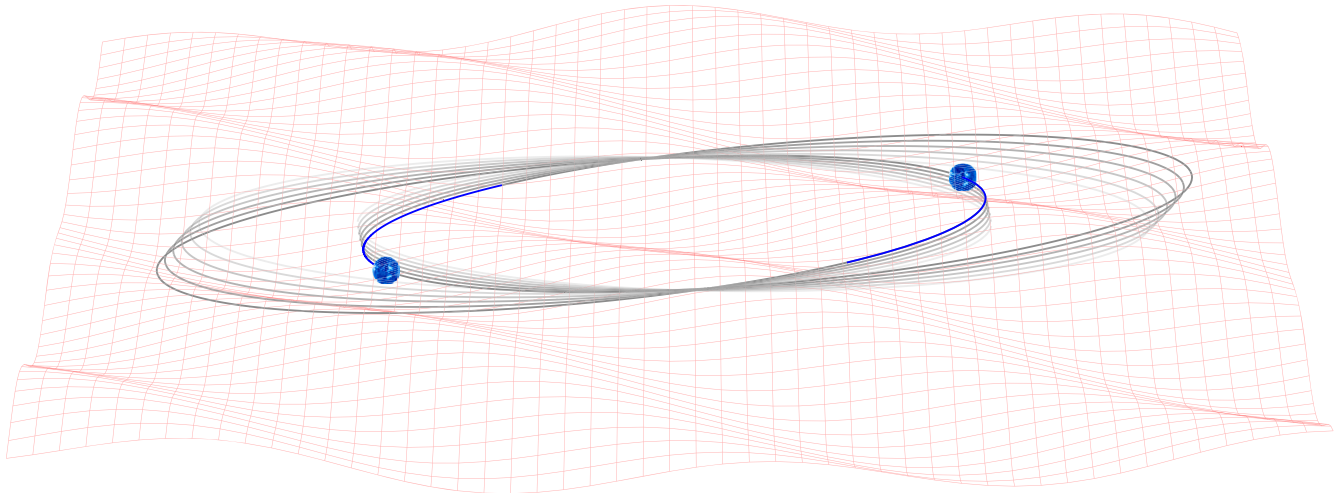


FIG. 1. Cartoon illustration of the binary resonance effect. Stochastic fluctuations in the background spacetime geometry due to incoming GWs perturb the trajectories of two orbiting masses, causing cumulative changes to their orbital elements.

to observable deviations which can be used to infer the GW amplitude, turning the binary into a dynamical GW detector (as illustrated in Fig. 1). This idea has a long history [19–33], and has been used to search for GWs with the orbit of the binary pulsar B1913 + 16 [34].<sup>1</sup> Nonetheless, this binary GW resonance effect has received relatively little attention in the GW community, and, we feel, has not yet been exploited to its full potential.

In this paper we develop, from first principles, a new formalism for calculating the GW-induced evolution of a generic binary system. Motivated by the requirements for the GW signal to be persistent (so that the orbital deviations accumulate over time) and broadband (so that the frequency content of the signal overlaps with the narrow resonant frequencies bands of the binary), we focus on the *stochastic GW background* (SGWB) [40–45]: a pseudo-random signal formed from the incoherent superposition of many independent GW sources throughout cosmic history. The stochastic nature of this signal means that we cannot deterministically calculate the evolution of a given binary; instead we treat each of the binary’s orbital elements as a time-dependent random variable, and study the statistics of the orbital perturbations over time. Our key result is a secularly-averaged *Fokker-Planck equation* (FPE) which fully specifies the evolution of the probability distribution for all six orbital elements. By comparing solutions of this FPE with high-precision orbital data from various binary systems, one can place stringent new constraints on the SGWB at frequencies that are inaccessible to all other current and future GW observatories, as we demonstrate in a companion paper [46].

The remainder of this paper is structured as follows. In Sec. II we give a brief, self-contained overview of

Keplerian orbits, orbital perturbations, and the formalism of osculating orbital elements. In Sec. III we specialize this formalism to perturbations from the SGWB, and develop a FPE for the orbital elements, expressing the coefficients of the equation in terms of GW transfer functions. In Sec. IV we derive these coefficients explicitly as functions of the orbital elements, and briefly discuss their properties. In Sec. V we obtain some exact late-time results for the case where the binary’s eccentricity is held fixed at zero, which greatly simplifies the FPE. In Sec. VI we tackle the more complicated general-eccentricity case, consider practical approaches for solving the FPE on observational timescales, and present some example results for the Hulse-Taylor binary pulsar B1913 + 16. In Sec. VII we develop the necessary statistical formalism to search for SGWB-induced orbital perturbations in observational data, discussing how to compute upper limits and sensitivity forecasts, and how to apply these tools to pulsar timing and laser-ranging experiments. We summarize our results in Sec. VIII. The Appendices give various technical details for the derivation of the FPE coefficients. We use units where  $c = 1$  throughout, but keep  $G \neq 1$ .

## II. BINARY DYNAMICS

In this section we introduce the machinery of osculating orbital elements. We start by recalling the basic properties of Keplerian orbits (i.e., orbits of two point masses interacting only through Newtonian gravity), before introducing the equations of motion (EOM) for the osculating elements, and discussing the most important contribution to these EOM from relativistic effects. We also introduce alternative sets of orbital elements that are useful in cases where the eccentricity or inclination of the orbit are small.

<sup>1</sup>Similar ideas have also been used to search for orbital changes induced by ultralight dark matter [35–39].

### A. Keplerian orbits

We start with a Keplerian binary, working in cylindrical coordinates  $(\hat{r}, \hat{\theta}, \hat{\ell})$  in a frame with the center of mass fixed at the origin. We also introduce a fixed Cartesian reference frame  $(\hat{x}, \hat{y}, \hat{z})$ , such the line of sight of the observer is in the positive  $\hat{z}$  direction. The unperturbed EOM for the separation vector  $\mathbf{r}$  is

$$\ddot{\mathbf{r}} + \frac{GM}{r^2} \hat{\mathbf{r}} = \mathbf{0}, \quad (2.1)$$

where  $r \equiv |\mathbf{r}|$  is the radial separation,  $\hat{\mathbf{r}} \equiv \mathbf{r}/r$  is the radial unit vector, and  $M \equiv m_1 + m_2$  is the total mass. We define the energy and angular momentum of the binary (in units of the reduced mass  $\mu \equiv m_1 m_2 / M$ ) by

$$\begin{aligned} \mathcal{E} &= \frac{1}{2} \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} - \frac{GM}{r}, \\ \boldsymbol{\ell} &= \mathbf{r} \times \dot{\mathbf{r}} = r^2 \dot{\hat{\theta}} \hat{\ell}. \end{aligned} \quad (2.2)$$

For the total angular momentum we write  $\ell \equiv |\boldsymbol{\ell}| = r^2 \dot{\theta}$ . These are all conserved, since

$$\begin{aligned} \dot{\mathcal{E}} &= \dot{\mathbf{r}} \cdot \left( \ddot{\mathbf{r}} + \frac{GM}{r^2} \hat{\mathbf{r}} \right) = 0, \\ \dot{\boldsymbol{\ell}} &= \mathbf{r} \times \ddot{\mathbf{r}} = \mathbf{r} \times \left( -\frac{GM}{r^2} \hat{\mathbf{r}} \right) = \mathbf{0}, \end{aligned} \quad (2.3)$$

where we used  $\dot{\mathbf{r}} = \dot{\mathbf{r}} \cdot \hat{\mathbf{r}}$ . The fact that  $\dot{\boldsymbol{\ell}} = \mathbf{0}$  means that the binary orbit is confined to a fixed 2D plane, which is specified with respect to the  $(\hat{x}, \hat{y}, \hat{z})$  reference frame by two angles: the inclination  $I$ , which is the angle between the binary's angular momentum vector  $\boldsymbol{\ell}$  and the observer's line of sight  $\hat{z}$ , and the longitude of ascending node  $\Omega$ , which is the angle between  $\hat{x}$  and the point where the orbit passes through the reference plane with positive velocity in the  $\hat{z}$  direction. (These angles are illustrated in Fig. 2.)

We can find the shape of the orbit in this plane by integrating Eq. (2.1), yielding a family of elliptical solutions

$$r = \frac{\ell^2 / (GM)}{1 + e \cos \psi}, \quad \mathcal{E} = -\frac{GM}{2a}, \quad \ell = \sqrt{GMa(1 - e^2)}, \quad (2.4)$$

with the shape of the ellipse described by its semimajor axis  $a$  and eccentricity  $e$ . Here we have introduced the true anomaly  $\psi$  as the angular position of the orbit within the orbital plane, defined such that the pericenter (minimum separation) occurs at  $\psi = 0$ . The point at which this occurs is defined by the argument of pericenter  $\omega \equiv \theta - \psi$ , which is measured relative to the ascending node. Since  $\omega$  is constant, the orbit is closed and the motion of the binary is

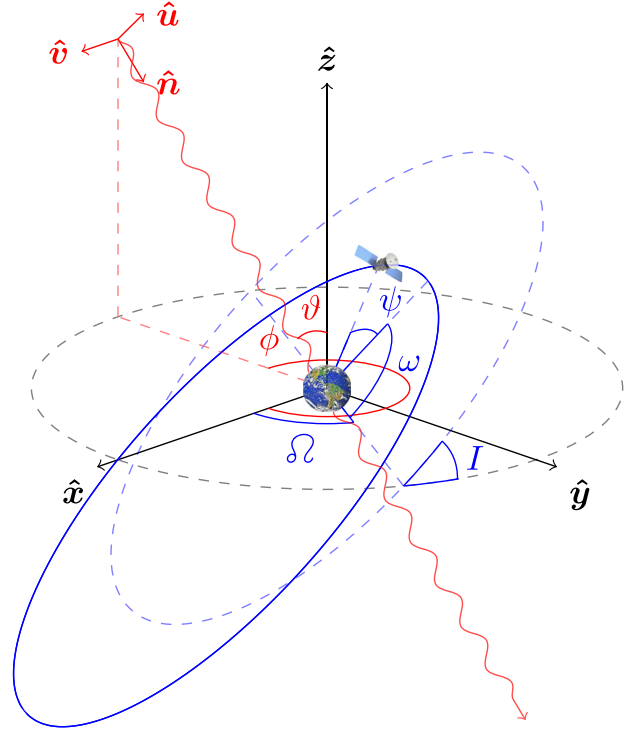


FIG. 2. Schematic diagram of a Keplerian orbit. The plane of the orbit (shown in blue) is defined relative to the fixed reference frame  $(\hat{x}, \hat{y}, \hat{z})$  (shown in black) by the inclination  $I$  and the longitude of ascending node  $\Omega$ , while the orientation of the orbit within this plane is specified by the argument of pericenter  $\omega$ . The true anomaly  $\psi$  acts as an angular coordinate for the position of the orbit within the orbital plane, measured relative to the pericenter. The polarization tensors describing the effects of an incoming plane GW are described in terms of the basis  $(\hat{n}, \hat{u}, \hat{v})$  (shown in red), which is related to the reference frame by the angles  $\vartheta$  and  $\phi$ . (Note that the wavelength of the GW is not shown to scale here—our analysis assumes a wavelength much larger than the size of the orbit.)

periodic in time, with period  $P$  related to the semimajor axis by Kepler's third law,

$$\frac{GM}{a^3} = \left( \frac{2\pi}{P} \right)^2. \quad (2.5)$$

(In what follows, we work entirely in terms of the period rather than the semimajor axis, as the former is more closely linked to the resonant frequencies of the orbit.)

The five constants  $(P, e, I, \Omega, \omega)$  are almost enough information to specify a particular Keplerian orbit; all that remains is to specify the time at which the binary is at pericenter,  $t_0$ . In practice, it is more convenient to replace  $t_0$  with the compensated mean anomaly,

$$\varepsilon \equiv \frac{2\pi}{P} (t - t_0) - \int_0^t dt' \frac{2\pi}{P(t')}, \quad (2.6)$$

as this is more well behaved when the orbit is perturbed [47]. Note that in the absence of perturbations this reduces to  $\varepsilon = -2\pi t_0/P$ . We call the set  $(P, e, I, \Omega, \omega, \varepsilon)$  the *orbital elements* of a binary.

### B. Perturbations and osculating orbits

Now, following Refs. [34,35,48,49], suppose the binary is acted upon by some small perturbing force

$$\delta\ddot{\mathbf{r}} = r(\mathcal{F}_r\hat{\mathbf{r}} + \mathcal{F}_\theta\hat{\boldsymbol{\theta}} + \mathcal{F}_\ell\hat{\boldsymbol{\ell}}), \quad (2.7)$$

so that the EOM is instead

$$\ddot{\mathbf{r}} + \frac{GM}{r^2}\hat{\mathbf{r}} = \delta\ddot{\mathbf{r}}. \quad (2.8)$$

[The factor of  $r$  in Eq. (2.7) is included for later convenience, as the GW perturbations we consider are proportional to the orbital separation.] Inserting this into Eq. (2.3) gives

$$\begin{aligned} \dot{\boldsymbol{\varepsilon}} &= \dot{\mathbf{r}} \cdot \delta\ddot{\mathbf{r}} = r\dot{r}\mathcal{F}_r + r^2\dot{\theta}\mathcal{F}_\theta, \\ \dot{\boldsymbol{\ell}} &= \mathbf{r} \times \delta\ddot{\mathbf{r}} = r^2\mathcal{F}_\theta\hat{\boldsymbol{\ell}} - r^2\mathcal{F}_\ell\hat{\boldsymbol{\theta}}, \end{aligned} \quad (2.9)$$

so the binary's energy and angular momentum are no longer constant. As a result, the binary is no longer described by a fixed set of orbital elements. However, the orbit is still tangent to some Keplerian ellipse at each moment in time. We therefore define  $(P, e, I, \Omega, \omega, \varepsilon)$  as functions of time which track the evolution of this tangent ellipse; these are called the *osculating* orbital elements.

A full derivation of the EOM for the osculating elements is given in, e.g., Refs. [48,49]; here we simply quote the resulting set of equations,

$$\begin{aligned} \dot{P} &= \frac{3P^2\gamma}{2\pi} \left[ \frac{e \sin \psi \mathcal{F}_r}{1 + e \cos \psi} + \mathcal{F}_\theta \right], \\ \dot{e} &= \frac{\dot{P}\gamma^2}{3Pe} - \frac{P\gamma^5 \mathcal{F}_\theta}{2\pi e(1 + e \cos \psi)^2}, \\ \dot{I} &= \frac{P\gamma^3 \cos \theta \mathcal{F}_\ell}{2\pi(1 + e \cos \psi)^2}, \\ \dot{\Omega} &= \frac{\tan \theta}{\sin I} \dot{I}, \\ \dot{\omega} &= \frac{P\gamma^3}{2\pi e} \left[ \frac{(2 + e \cos \psi) \sin \psi \mathcal{F}_\theta}{(1 + e \cos \psi)^2} - \frac{\cos \psi \mathcal{F}_r}{1 + e \cos \psi} \right] - \cos I \dot{\Omega}, \\ \dot{\varepsilon} &= -\frac{P\gamma^4 \mathcal{F}_r}{\pi(1 + e \cos \psi)^2} - \gamma(\cos I \dot{\Omega} + \dot{\omega}), \end{aligned} \quad (2.10)$$

where we have defined the dimensionless angular momentum,

$$\gamma \equiv \sqrt{1 - e^2} = \frac{\ell}{\sqrt{GMa}}. \quad (2.11)$$

Note that the size and shape of the orbit (determined by  $P$  and  $e$ ) is only affected by forces within the plane (i.e.,  $\mathcal{F}_r$  and  $\mathcal{F}_\theta$ ), while the plane of the orbit (determined by  $I$  and  $\Omega$ ) is only affected by forces normal to the plane (i.e.,  $\mathcal{F}_\ell$ ), and the radial and angular phases of the orbit (determined by  $\omega$  and  $\varepsilon$ ) are affected by both.

### C. Secular evolution due to relativistic effects

In Sec. III, we use Eq. (2.10) to calculate the perturbations to the osculating orbital elements caused by resonance with the SGWB. However, we can use the same set of equations to calculate the perturbations caused by relativistic corrections to the equations of motion, which are particularly important for binaries with short periods. Following Ref. [50], we write the relativistic force components to leading post-Newtonian (PN) order as

$$\begin{aligned} \mathcal{F}_r &= \left( \frac{2\pi}{P} \right)^2 \frac{v_P^2}{\gamma^8} (1 + e \cos \psi)^3 \left[ 3 - \eta - e^2(1 + 3\eta) \right. \\ &\quad \left. + e(2 - 4\eta) \cos \psi + e^2 \frac{8 - \eta}{2} \sin^2 \psi \right], \\ \mathcal{F}_\theta &= \left( \frac{2\pi}{P} \right)^2 \frac{2e v_P^2}{\gamma^8} \sin \psi (1 + e \cos \psi)^4 (2 - \eta), \\ \mathcal{F}_\ell &= 0, \end{aligned} \quad (2.12)$$

where

$$v_P \equiv \left( \frac{2\pi GM}{P} \right)^{1/3}, \quad (2.13)$$

is the binary's rms velocity, and

$$\eta \equiv \frac{\mu}{M} = \frac{m_1 m_2}{(m_1 + m_2)^2}, \quad (2.14)$$

is the dimensionless mass ratio. Inserting these expressions into Eq. (2.10), we average over the orbit to find the secular perturbations, using the fact that  $\dot{\psi} = \ell/r^2$  for Keplerian orbits to write

$$\dot{X}_{\text{sec}} \equiv \int_{t_0}^{t_0+P} \frac{dt}{P} \dot{X} = \int_0^{2\pi} \frac{d\psi}{2\pi} \frac{\gamma^3 \dot{X}}{(1 + e \cos \psi)^2}, \quad (2.15)$$

where  $X \in (P, e, I, \Omega, \omega, \varepsilon)$ . The only nonvanishing perturbations to the osculating elements in this case are then

$$\dot{\omega}_{\text{sec}} = \frac{6\pi v_P^2}{P\gamma^2}, \quad \dot{\varepsilon}_{\text{sec}} = \frac{2\pi v_P^2}{P} \left( 6 - 7\eta - \frac{15 - 9\eta}{\gamma} \right), \quad (2.16)$$

where the first equality is the famous perihelion precession. These are both 1PN corrections (i.e., order  $v_p^2$ ).

Note that the above implies that the binary's period and eccentricity are conserved at 1PN order. However, since these are typically the most precisely-measured orbital elements, it is worth including their leading-order relativistic evolution, even if this is at higher PN order and is thus much smaller than the terms in Eq. (2.16) for most binaries. This leading-order evolution of the period and eccentricity is due to GW radiation reaction, and can be calculated by applying simple energy-balance arguments to the unperturbed Keplerian orbit [51], giving

$$\begin{aligned}\dot{P}_{\text{sec}} &= -\frac{192\pi\eta v_p^5}{5\gamma^7} \left(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4\right), \\ \dot{e}_{\text{sec}} &= -\frac{608\pi\eta v_p^5}{15P\gamma^5} \left(e + \frac{121}{304}e^3\right).\end{aligned}\quad (2.17)$$

We see that these are both 2.5PN effects (i.e., order  $v_p^5$ ). The inclination and longitude of ascending node are both conserved at this PN order.

#### D. Small-eccentricity and small-inclination orbits

For binaries where the eccentricity is small ( $e \lesssim 10^{-3}$ ) the argument of pericenter  $\omega$  becomes ill defined, and this in turn means that the compensated mean anomaly  $\varepsilon$  becomes ill defined, as this is defined relative to the time at which the binary is at pericenter,  $t_0$ . These issues can be resolved by defining

$$\zeta \equiv e \sin \omega, \quad \kappa \equiv e \cos \omega, \quad \xi \equiv \omega + \varepsilon. \quad (2.18)$$

The first two quantities here are sometimes called the ‘‘Laplace-Lagrange eccentric parameters,’’ while the latter is the compensated mean argument. We can then describe the orbit of a near-circular binary in terms of the alternative set of osculating elements  $(P, \zeta, \kappa, I, \Omega, \xi)$  [49,52]. These evolve according to

$$\begin{aligned}\dot{\zeta} &= \dot{e} \sin \omega + \dot{\omega} e \cos \omega, \\ \dot{\kappa} &= \dot{e} \cos \omega - \dot{\omega} e \sin \omega, \\ \dot{\xi} &= \dot{\omega} + \dot{\varepsilon},\end{aligned}\quad (2.19)$$

with  $\dot{e}$ ,  $\dot{\Omega}$ ,  $\dot{\omega}$ , and  $\dot{\varepsilon}$  given by Eq. (2.10). For the relativistic perturbations (2.12), we thus have

$$\begin{aligned}\dot{\zeta}_{\text{sec}} &= \frac{6\pi v_p^2}{P} \left(\kappa - \frac{304}{45}\eta v_p^3 \zeta\right), \\ \dot{\kappa}_{\text{sec}} &= -\frac{6\pi v_p^2}{P} \left(\zeta + \frac{304}{45}\eta v_p^3 \kappa\right), \\ \dot{\xi}_{\text{sec}} &= -\frac{4\pi v_p^2}{P} (3 - \eta),\end{aligned}\quad (2.20)$$

where we have neglected  $\mathcal{O}(e^2)$  terms.

Similarly,  $\Omega$  is ill defined for orbits with very small inclination, so in this case we define

$$p \equiv I \sin \Omega, \quad q \equiv I \cos \Omega, \quad \lambda = \Omega + \xi, \quad (2.21)$$

and describe the orbit using  $(P, \zeta, \kappa, p, q, \lambda)$ .

### III. RESONANT GRAVITATIONAL-WAVE PERTURBATIONS

In this section we calculate the evolution of the osculating orbital elements of a binary system due to resonance with the SGWB. We start by specifying the perturbing force associated with an incoming plane GW in the limit where the wavelength is much larger than the size of the orbit. This allows us to write down a Langevin equation describing individual random realizations of the stochastic evolution of the osculating elements. Using the statistical properties of the SGWB, we then derive a secularly averaged FPE which describes the evolution of the full statistical distribution of the orbital elements over time-scales much longer than the binary period.

#### A. Coupling to the gravitational-wave polarization modes

The response of a binary to an impinging plane GW can be expressed in the proper detector frame (i.e., using Fermi normal coordinates to construct a freely falling frame with the binary's center of mass fixed at the origin, in which coordinate distances correspond to proper distances to a good approximation, and in which we can treat the GW as a perturbing Newtonian force) as [20,53]

$$\delta\ddot{r}^i = \frac{1}{2}\ddot{h}_{ij}r^j, \quad (3.1)$$

so that the resulting evolution of the binary is described in terms of the perturbing force terms

$$\mathcal{F}_r = \frac{1}{2}\ddot{h}_{ij}\hat{r}^i\hat{r}^j, \quad \mathcal{F}_\theta = \frac{1}{2}\ddot{h}_{ij}\hat{r}^i\hat{\theta}^j, \quad \mathcal{F}_\ell = \frac{1}{2}\ddot{h}_{ij}\hat{r}^i\hat{\ell}^j, \quad (3.2)$$

where  $h_{ij}(t)$  is the transverse-traceless part of the metric perturbation at the position of the binary's center of mass.

We decompose the GW strain in terms of plane waves of each polarization arriving from each direction on the sky,

$$h_{ij}(t) = \int_{S^2} d^2\hat{n} e_{ij}^A(\hat{n}) h_A(t, \hat{n}), \quad (3.3)$$

where  $A = +, \times$  are the two GW polarizations, and summation over the repeated polarization index is implied. We define the standard polarization tensors

$$e_{ij}^+ = \hat{u}^i\hat{u}^j - \hat{v}^i\hat{v}^j, \quad e_{ij}^\times = \hat{u}^i\hat{v}^j + \hat{v}^i\hat{u}^j, \quad (3.4)$$

with  $\hat{\mathbf{u}}, \hat{\mathbf{v}}$  unit vectors that are orthogonal to the GW propagation direction  $\hat{\mathbf{n}}$  and to each other, as illustrated in Fig. 2. We therefore write

$$\mathcal{F}_\alpha = \frac{1}{2} \int_{S^2} d^2\hat{\mathbf{n}} e^{A_{ij}} \hat{r}^i \hat{\alpha}^j \ddot{h}_A, \quad (3.5)$$

where  $\alpha$  runs over the cylindrical coordinates  $(r, \theta, \ell)$ .

Note that Eq. (3.1) is only correct in the limit where the GW wavelength  $\lambda$  is much larger than the size of the binary's orbit,  $\lambda \gg a$ . Using Kepler's third law (2.5), this condition can be rewritten as

$$fP \ll 1/v_p, \quad (3.6)$$

where  $f$  is the GW frequency. Since we are interested in GW frequencies which are harmonics of the binary period ( $f = n/P$  for some  $n \in \mathbb{Z}_+$ ), this tells us that the analysis below, based on Eq. (3.1), is only valid for harmonics that satisfy  $n \ll 1/v_p$ . This is not an impediment, since in the cases we are interested in the binary is subrelativistic,  $v_p \ll 1$ , and the strongest contribution typically comes from the lowest few harmonics anyway.

### B. Langevin formulation

The stochastic evolution of the binary is described by the evolution equations derived in Sec. II B, with the perturbing force terms given by Eq. (3.5). All this can be rewritten as a coupled set of nonlinear Langevin equations,

$$\dot{X}_i(\mathbf{X}, t) = V_i(\mathbf{X}) + \Gamma_i(\mathbf{X}, t), \quad (3.7)$$

where  $X_i$  runs over the set of orbital elements [either  $(P, e, I, \Omega, \omega, \varepsilon)$ , or the small-eccentricity and/or small-inclination alternatives described in Sec. II D],  $V$  is the deterministic drift term due to the relativistic effects described in Sec. II C, and  $\Gamma$  is the stochastic diffusion term due to resonance with the SGWB. The latter can be written as

$$\Gamma_i(\mathbf{X}, t) = \int_{S^2} d^2\hat{\mathbf{n}} T_i^A(\mathbf{X}, t, \hat{\mathbf{n}}) \ddot{h}_A(t, \hat{\mathbf{n}}), \quad (3.8)$$

where the  $T_i^A$  are transfer functions describing the coupling between the SGWB strain and the orbital elements. For example, using Eqs. (2.10) and (3.5), the transfer functions for the period  $P$  are

$$T_P^A = \frac{3P^2\gamma}{4\pi} \left( \frac{e \sin \psi}{1 + e \cos \psi} \hat{r}^i + \hat{\theta}^i \right) \hat{r}^j e_{ij}^A. \quad (3.9)$$

(The full set of transfer functions are written out in Appendix B in terms of their Fourier components.)

The explicit time-dependence in the transfer functions is via the true anomaly  $\psi(t)$ , due to the variation of the

binary's response over the course of each orbit. These variations occur only on timescales less than or equal to the orbital period  $P$ , which is much shorter than the timescales over which the orbital elements evolve,  $X/\dot{X} \gg P$ . The latter are *secular* effects, whereas the former are *resonant* ones, which contribute only at integer multiples of the fundamental frequency  $1/P$ . This property of the resonant spectrum, combined with the separation of scales between the resonant and secular evolution, allows us to approximate the transfer functions as Fourier series with period  $P$ ,

$$T_i^A(\mathbf{X}, t, \hat{\mathbf{n}}) = \sum_{n=-\infty}^{+\infty} e^{-2\pi i n t/P} T_{i,n}^A(\mathbf{X}, \hat{\mathbf{n}}). \quad (3.10)$$

In Sec. III D, we recast Eq. (3.7) as a FPE, allowing us to calculate the ensemble-averaged properties of the binary's orbital evolution. However, in order to do so, we must first describe the statistical properties of the SGWB strain, which we do in Sec. III C below.

### C. Statistical properties of the stochastic background

The stochastic GW strain components  $h_A$  are time-dependent random variables whose statistical properties are usually specified in terms of their Fourier transforms,

$$\tilde{h}_A(f, \hat{\mathbf{n}}) = \int_{-\infty}^{+\infty} dt e^{2\pi i f t} h_A(t, \hat{\mathbf{n}}). \quad (3.11)$$

Since the SGWB is generated by the superposition of a large number of statistically independent sources, it is usually assumed that the SGWB strain is Gaussian by the central limit theorem. This implies that the statistics of the strain Fourier components are fully specified by their first two moments. With the further standard assumptions that the SGWB is isotropic, stationary, unpolarized, and has zero phase correlation between different sky locations, these moments can be written as

$$\begin{aligned} \langle \tilde{h}_A(f, \hat{\mathbf{n}}) \rangle &= 0, \\ \langle \tilde{h}_A(f, \hat{\mathbf{n}}) \tilde{h}_{A'}^*(f', \hat{\mathbf{n}}') \rangle &= \frac{3H_0^2 \Omega(f)}{32\pi^3 |f|^3} \delta_{AA'} \delta(f - f') \delta(\hat{\mathbf{n}}, \hat{\mathbf{n}}'), \end{aligned} \quad (3.12)$$

with  $\Omega(f)$  the one-sided SGWB energy density spectrum,

$$\Omega(f) \equiv \frac{1}{\rho_{\text{crit}}} \frac{d\rho_{\text{gw}}}{d(\ln f)}, \quad (3.13)$$

which is normalized with respect to the cosmological critical density  $\rho_{\text{crit}} \equiv 3H_0^2/(8\pi G)$ , where  $H_0$  is the Hubble constant.

We are interested in the statistics of the second time derivative of the strain, which is related to the Fourier components by

$$\begin{aligned}\ddot{h}_A &= \frac{d^2}{dt^2} \int_{-\infty}^{+\infty} df e^{-2\pi ift} \tilde{h}_A \\ &= -4\pi^2 \int_{-\infty}^{+\infty} df e^{-2\pi ift} f^2 \tilde{h}_A.\end{aligned}\quad (3.14)$$

The corresponding moments are therefore

$$\begin{aligned}\langle \ddot{h}_A(t, \hat{\mathbf{n}}) \rangle &= 0, \\ \langle \ddot{h}_A(t, \hat{\mathbf{n}}) \ddot{h}_{A'}(t', \hat{\mathbf{n}}') \rangle &= 3\pi H_0^2 \delta_{AA'} \delta(\hat{\mathbf{n}}, \hat{\mathbf{n}}') \\ &\quad \times \int_0^\infty df \cos[2\pi f(t-t')] f \Omega(f).\end{aligned}\quad (3.15)$$

#### D. From the Langevin equation to the Fokker-Planck equation

Solving the Langevin equation (3.7) gives individual random trajectories of the binary through parameter space, corresponding to different random realizations of the SGWB. However, we are more interested in the ensemble of all possible random trajectories, which we describe in terms of the time-dependent distribution function (DF) for the orbital elements,  $W(\mathbf{X}, t)$ . This is defined such that the probability of the orbital elements  $\mathbf{X}$  belonging to any region  $\mathcal{X}$  of parameter space at time  $t$  is given by the corresponding integral over the DF,

$$\Pr(\mathbf{X} \in \mathcal{X} | t) = \int_{\mathcal{X}} d\mathbf{X} W(\mathbf{X}, t). \quad (3.16)$$

This DF can either be interpreted as the probability density function for the stochastic orbital elements of an individual binary, or as the cumulative distribution for the orbital elements of a population of multiple binaries.

Formally, we can write down the time evolution of the DF in terms of the Kramers-Moyal (KM) forward expansion [54,55],

$$\frac{\partial W}{\partial t} = \sum_{n=1}^{\infty} (-)^n \frac{\partial^n}{\partial X_{i_1} \partial X_{i_2} \cdots \partial X_{i_n}} (D_{i_1 i_2 \dots i_n}^{(n)} W), \quad (3.17)$$

where repeated indices are summed over. This is determined by the KM coefficients,

$$D_{i_1 i_2 \dots i_n}^{(n)}(\mathbf{X}, t) \equiv \lim_{\tau \rightarrow 0} \frac{1}{\tau n!} \left\langle \prod_{j=1}^n [X_{i_j}(t+\tau) - X_{i_j}(t)] \right\rangle, \quad (3.18)$$

with angle brackets indicating an ensemble average under the distribution  $W$  at time  $t$ . Since we take the SGWB as

Gaussian, the KM coefficients for orders  $n \geq 3$  all vanish,<sup>2</sup> leaving just the first two coefficients,

$$\begin{aligned}D_i^{(1)} &= \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle X_i(t+\tau) - X_i(t) \rangle, \\ D_{ij}^{(2)} &= \lim_{\tau \rightarrow 0} \frac{1}{2\tau} \langle [X_i(t+\tau) - X_i(t)][X_j(t+\tau) - X_j(t)] \rangle,\end{aligned}\quad (3.19)$$

which we call the *drift vector* and the *diffusion matrix*, respectively. The KM forward expansion (3.17) then becomes the *Fokker-Planck equation* (FPE),

$$\frac{\partial W}{\partial t} = -\partial_i (D_i^{(1)} W) + \partial_i \partial_j (D_{ij}^{(2)} W), \quad (3.20)$$

with  $\partial_i \equiv \partial / \partial X_i$ .

We can calculate the KM coefficients by directly integrating the Langevin equation (3.7), from some initial time  $t$  where the orbital elements are ‘‘sharp’’ (i.e., known exactly rather than randomly distributed),  $X_i(t) \equiv x_i$ , over some small time interval  $\tau$ ,

$$X_i(t+\tau) - x_i = \int_t^{t+\tau} dt' [V_i(\mathbf{X}(t')) + \Gamma_i(\mathbf{X}(t'), t')]. \quad (3.21)$$

(This derivation closely follows that in Sec. 3.3.2 of Ref. [54].) Both terms under the integral on the right-hand side (rhs) are random, due to the random spread in the orbital elements for all times  $t' > t$ . However, we can express these in terms of the sharp values  $x_i$  by Taylor expanding,

$$\begin{aligned}V_i(\mathbf{X}(t')) &= V_i(\mathbf{x}) + \partial_j V_i(\mathbf{x}) [X_j(t') - x_j] + \cdots, \\ \Gamma_i(\mathbf{X}(t'), t') &= \Gamma_i(\mathbf{x}, t') + \partial_j \Gamma_i(\mathbf{x}, t') [X_j(t') - x_j] + \cdots.\end{aligned}\quad (3.22)$$

Wherever  $X_i(t') - x_i$  appears on the rhs of Eq. (3.22) we can insert Eq. (3.21) and iterate, such that the only orbital elements that appear are the sharp initial values  $x_i$ . Then the only randomness is through the stochastic term  $\Gamma_i$ , whose statistics are completely specified in terms of the SGWB moments (3.15).

<sup>2</sup>One could derive this explicitly by using the approach described later in this section to show that  $D_{ijk}^{(3)} = 0$ , as the Pawula theorem [54,56] then implies that all higher-order coefficients  $n > 3$  must vanish in order to guarantee that the DF is normalized.

The first few terms in the iterative expansion are

$$X_i(t + \tau) - x_i = \int_t^{t+\tau} dt' V_i(\mathbf{x}) + \int_t^{t+\tau} dt' \partial_j V_i(\mathbf{x}) \int_t^{t'} dt'' [V_j(\mathbf{x}) + \Gamma_j(\mathbf{x}, t'')] + \dots \\ + \int_t^{t+\tau} dt' \Gamma_i(\mathbf{x}, t') + \int_t^{t+\tau} dt' \partial_j \Gamma_i(\mathbf{x}, t') \int_t^{t'} dt'' [V_j(\mathbf{x}) + \Gamma_j(\mathbf{x}, t'')] + \dots \quad (3.23)$$

By virtue of the time-independence of the sharp drift term  $V_i(\mathbf{x})$ , this immediately simplifies to

$$X_i(t + \tau) - x_i = \tau V_i(\mathbf{x}) + \frac{1}{2} \tau^2 V_j(\mathbf{x}) \partial_j V_i(\mathbf{x}) + \partial_j V_i(\mathbf{x}) \int_t^{t+\tau} dt' \int_t^{t'} dt'' \Gamma_j(\mathbf{x}, t'') + \dots \\ + \int_t^{t+\tau} dt' \Gamma_i(\mathbf{x}, t') + V_j(\mathbf{x}) \int_t^{t+\tau} dt' (t' - t) \partial_j \Gamma_i(\mathbf{x}, t') + \int_t^{t+\tau} dt' \partial_j \Gamma_i(\mathbf{x}, t') \int_t^{t'} dt'' \Gamma_j(\mathbf{x}, t'') + \dots \quad (3.24)$$

Taking the first moment of Eq. (3.24), all terms linear in  $\Gamma_i$  vanish, and we are left with

$$\langle X_i(t + \tau) - x_i \rangle \\ = \tau V_i(\mathbf{x}) + \frac{1}{2} \tau^2 V_j(\mathbf{x}) \partial_j V_i(\mathbf{x}) + \dots \\ + \int_t^{t+\tau} dt' \int_t^{t'} dt'' \langle \Gamma_j(\mathbf{x}, t'') \partial_j \Gamma_i(\mathbf{x}, t') \rangle + \dots, \quad (3.25)$$

so that the drift vector is given by

$$D_i^{(1)} = V_i + \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int_t^{t+\tau} dt' \int_t^{t'} dt'' \langle \Gamma_j(\mathbf{x}, t'') \partial_j \Gamma_i(\mathbf{x}, t') \rangle. \quad (3.26)$$

Similarly, when taking the second moment of Eq. (3.24) and taking the  $\tau \rightarrow 0$  limit, the only term that survives for the diffusion matrix is

$$D_{ij}^{(2)} = \lim_{\tau \rightarrow 0} \frac{1}{2\tau} \int_t^{t+\tau} dt' \int_t^{t'} dt'' \langle \Gamma_i(\mathbf{x}, t') \Gamma_j(\mathbf{x}, t'') \rangle. \quad (3.27)$$

By calculating the KM coefficients from Eqs. (3.26) and (3.27), we can then obtain the full DF for the orbital elements by integrating the FPE (3.20).

### E. Secular drift and diffusion

We rewrite the integrands of Eqs. (3.26) and (3.27) in terms of the GW strain and the binary transfer functions using Eq. (3.8). We can then use Eq. (3.15) to specify the SGWB strain covariance, and insert Eq. (3.10) to form a sum over harmonics of the binary's orbital frequency,  $f = n/P$ . The resulting expressions are

$$\langle \Gamma_i(\mathbf{x}, t') \Gamma_j(\mathbf{x}, t'') \rangle = 3\pi H_0^2 \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \int_{S^2} d^2 \hat{\mathbf{n}} T_{i,n}^{A*} T_{j,m}^A \int_0^\infty df e^{2\pi i(n t' - m t'')/P} \cos[2\pi f(t' - t'')] f \Omega(f), \\ \langle \Gamma_j(\mathbf{x}, t'') \partial_j \Gamma_i(\mathbf{x}, t') \rangle = 3\pi H_0^2 \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \int_{S^2} d^2 \hat{\mathbf{n}} T_{j,m}^A \partial_j T_{i,n}^{A*} \int_0^\infty df e^{2\pi i(n t' - m t'')/P} \cos[2\pi f(t' - t'')] f \Omega(f). \quad (3.28)$$

In order to derive the corresponding KM coefficients, we must evaluate two oscillatory time integrals,

$$\int_t^{t+\tau} dt' \int_t^{t'+\tau} dt'' e^{-2\pi i(f-n/P)t'} e^{2\pi i(f-m/P)t''}, \quad \int_t^{t+\tau} dt' \int_t^{t'} dt'' e^{-2\pi i(f-n/P)t'} e^{2\pi i(f-m/P)t''}. \quad (3.29)$$

To do so, we recall that the timescales  $\lesssim P$  associated with the binary's resonant frequencies are much shorter than the secular timescales  $X/\dot{X}$  we are interested in.<sup>3</sup> Even though we will later take the limit  $\tau \rightarrow 0$ , all we really require to derive a FPE valid on secular timescales  $\sim X/\dot{X}$  is that  $\tau \ll X/\dot{X}$ , and since we have  $P \ll X/\dot{X}$ , we can consistently also demand that  $\tau \gg P$ . In this limit, the first integral in Eq. (3.29) is approximated by

<sup>3</sup>It would also be interesting to study the dynamics on suborbital timescales, as was done in Refs. [38,57] for the case of ultralight dark matter. We leave this for future work.



$$\int_t^{t+\tau} dt' \int_t^{t'+\tau} dt'' e^{-2\pi i(f-n/P)t'} e^{2\pi i(f-m/P)t''} \simeq \tau \delta_{mn} \delta(f-n/P). \quad (3.30)$$

For the second integral, notice that we can write

$$\begin{aligned} & \tau \delta_{mn} \delta(f-n/P) \\ & \simeq \int_t^{t+\tau} dt' \int_t^{t'+\tau} dt'' e^{-2\pi i(f-n/P)t'} e^{2\pi i(f-m/P)t''} \\ & = \int_t^{t+\tau} dt' \int_t^{t'} dt'' e^{-2\pi i(f-n/P)t'} e^{2\pi i(f-m/P)t''} \\ & \quad + \int_t^{t+\tau} dt' \int_{t'}^{t'+\tau} dt'' e^{-2\pi i(f-n/P)t'} e^{2\pi i(f-m/P)t''}, \end{aligned} \quad (3.31)$$

and that in the limit  $\tau \gg P$  the latter two terms are approximately equal to each other, so that

$$\int_t^{t+\tau} dt' \int_t^{t'} dt'' e^{-2\pi i(f-n/P)t'} e^{2\pi i(f-m/P)t''} \simeq \frac{1}{2} \tau \delta_{mn} \delta(f-n/P). \quad (3.32)$$

Plugging this back in to Eqs. (3.26), (3.27), and (3.28), we are able to write the KM coefficients directly in terms of the GW transfer functions,

$$\begin{aligned} D_i^{(1)} &= V_i + \frac{3\pi}{2} H_0^2 \sum_{n=1}^{\infty} \int_{S^2} d^2 \hat{n} \frac{n \Omega_n}{P} \text{Re}(T_{j,n}^A \partial_j T_{i,n}^{A*}), \\ D_{ij}^{(2)} &= \frac{3\pi}{2} H_0^2 \sum_{n=1}^{\infty} \int_{S^2} d^2 \hat{n} \frac{n \Omega_n}{P} \text{Re}(T_{i,n}^{A*} T_{j,n}^A), \end{aligned} \quad (3.33)$$

where we have used  $T_{i,-n} = T_{i,n}^*$ , and where  $\Omega_n \equiv \Omega(n/P)$  is the SGWB energy density at the binary's  $n$ th harmonic frequency. (We always take only the real part when defining the KM coefficients, so for brevity we will leave this implicit from now on.) Equation (3.33) describes the secular drift and diffusion of the orbital elements over timescales much longer than the orbital period  $P$ .

Note that the drift vector includes not just the expected deterministic drift  $V_i$ , but also a stochastic term. As we show below, this “noise-induced drift” leads to a net evolution of the *mean values* of the orbital elements, and not just their variance. This is somewhat counterintuitive, and justifies the careful derivation presented in this section; otherwise, it would be tempting to assume that the SGWB affects only the variance of the orbital elements (as was assumed in Ref. [34], for example). One can understand this drift as being the result of a “diffusion gradient” due to the nonlinear coupling between the SGWB and the binary: the orbital elements change in response to the GW strain, thereby changing the values of the transfer functions and modifying the response to

the strain, with the resulting feedback loop creating a preferred direction in parameter space for the orbital elements to evolve toward. The partial derivative acting on the transfer function in Eq. (3.33) shows that this term would vanish in the linear case (i.e., where the transfer functions are constant), and therefore that this is a purely nonlinear effect.

The stochastic drift is also a result of our treatment of the SGWB strain, as we have started with a finite correlation time and taken the limit where this vanishes, rather than specifying zero correlation time from the start. (Physically, the SGWB strain at time  $t_1$  cannot be uncorrelated with that at  $t_2$  for arbitrarily small  $|t_1 - t_2|$ , as this would require  $\Omega(f)$  to remain finite as  $f \rightarrow \infty$ , which would make the total GW energy density diverge. However, the time over which the strain is correlated is still much shorter than the timescale over which the orbital elements evolve.) This is the Stratonovich prescription [58], which is appropriate for most physical applications, rather than the more formal Itô prescription [59]. For further discussion of the two prescriptions see, e.g., Ref. [54].

#### IV. CALCULATING THE KRAMERS-MOYAL COEFFICIENTS

In this section we use Eq. (3.33) to derive explicit expressions for the KM coefficients as functions of the binary orbital elements and the SGWB spectrum. We start with the most general expressions, before specializing to the cases of small eccentricity and small inclination.

##### A. General orbits

We present here the secular KM coefficients for general eccentricity  $e \in (0, 1)$ . The details of this calculation are lengthy and unimportant for the final results, so we quote only the final expressions here, with some further details given in the Appendices. In particular, Appendix A derives the necessary projections of the GW polarization tensors onto the binary's cylindrical coordinate basis; Appendix B uses these projections to write down the transfer functions for all six orbital elements ( $P, e, I, \Omega, \omega, \varepsilon$ ) in terms of functions of the eccentricity called *Hansen coefficients*, which we introduce below; Appendix C derives the equations describing how each of the KM coefficients transforms under a change of the reference frame, allowing us to select a particular frame which simplifies the calculations; Appendix D presents the values of the KM coefficients in this frame in terms of the Hansen coefficients; and Appendix E writes out all of the necessary Hansen coefficients explicitly as functions of eccentricity.

Putting all of these ingredients together, we find that the secular drift vector  $D_i^{(1)}$  is given in terms of the Hansen coefficients ( $C_n^{lm}, S_n^{lm}, E_n^{lm}$ ) by

$$\begin{aligned}
D_P^{(1)} &= V_P + \frac{9P^2\gamma^2}{4} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ |E_n^{02} + \frac{e}{2}(E_n^{11} - E_n^{13})|^2 - \frac{1+4e^2}{15}(S_n^{11})^2 - \frac{e\gamma^2}{15}S_n^{11'}S_n^{11} \right. \\
&\quad + \frac{\gamma^2}{10e} \left( E_n^{02} + \frac{e}{2}(E_n^{11} - E_n^{13}) \right) (3E_n^{11} + E_n^{13} + eE_n^{20} + 4E_n^{21} + 4eE_n^{22} - 4E_n^{23} - eE_n^{24} + 2E_n^{02'} + e(E_n^{11'} - E_n^{13'}))^* \\
&\quad \left. - \frac{\gamma^4}{10e} E_n^{22} (E_n^{11} - E_n^{13} + 2E_n^{02'} + e(E_n^{11'} - E_n^{13'}))^* \right], \\
D_e^{(1)} &= V_e - \frac{P\gamma^6}{20} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ S_n^{11'}S_n^{11} + \frac{3}{e^3} \left( E_n^{02} + \frac{e}{2}(E_n^{11} - E_n^{13}) - \gamma^2 E_n^{22} \right) \right. \\
&\quad \left. \times (E_n^{02} - E_n^{22} - e(E_n^{11} + E_n^{13} + 2E_n^{21} - 2E_n^{23} + E_n^{02'} - \gamma^2 E_n^{22'})) - \frac{e^2}{2} (E_n^{20} + 10E_n^{22} - E_n^{24} + E_n^{11'} - E_n^{13'})^* \right], \\
D_I^{(1)} &= \frac{\sin 2I}{2} D_{\Omega\Omega}^{(2)}, \quad D_{\Omega}^{(1)} = -2 \cot ID_{I\Omega}^{(2)}, \\
D_{\omega}^{(1)} &= V_{\omega} + \frac{3P\gamma^6}{40} \sin 2\omega \frac{2 - \sin^2 I}{\sin^2 I} \sum_{n=1}^{\infty} nH_0^2\Omega_n E_n^{20} (E_n^{22})^*, \quad D_{\varepsilon}^{(1)} = V_{\varepsilon}, \tag{4.1}
\end{aligned}$$

while the secular diffusion matrix  $D_{ij}^{(2)}$  is given by

$$\begin{aligned}
D_{PP}^{(2)} &= \frac{27P^3\gamma^2}{20} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ |E_n^{02} + \frac{e}{2}(E_n^{11} - E_n^{13})|^2 - \frac{(eS_n^{11})^2}{3} \right], \\
D_{Pe}^{(2)} &= \frac{\gamma^2 D_{PP}^{(2)}}{3Pe} - \frac{9P^2\gamma^6}{40} \sum_{n=1}^{\infty} nH_0^2\Omega_n E_n^{22} \left( \frac{2}{e} E_n^{02} + E_n^{11} - E_n^{13} \right)^*, \\
D_{ee}^{(2)} &= \frac{3P\gamma^6}{20e^2} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ |E_n^{02} + \frac{e}{2}(E_n^{11} - E_n^{13}) - \gamma^2 E_n^{22}|^2 - \frac{(eS_n^{11})^2}{3} \right], \\
D_{II}^{(2)} &= \frac{3P\gamma^6}{80} \sum_{n=1}^{\infty} nH_0^2\Omega_n [|E_n^{20}|^2 + |E_n^{22}|^2 + 2 \cos 2\omega E_n^{20} (E_n^{22})^*], \\
D_{I\Omega}^{(2)} &= \frac{3P\gamma^6 \sin 2\omega}{40 \sin I} \sum_{n=1}^{\infty} nH_0^2\Omega_n E_n^{20} (E_n^{22})^*, \\
D_{I\omega}^{(2)} &= -\frac{3P\gamma^6 \sin 2\omega}{40 \tan I} \sum_{n=1}^{\infty} nH_0^2\Omega_n E_n^{20} (E_n^{22})^*, \\
D_{\Omega\Omega}^{(2)} &= \frac{3P\gamma^6}{80\sin^2 I} \sum_{n=1}^{\infty} nH_0^2\Omega_n [|E_n^{20}|^2 + |E_n^{22}|^2 - 2 \cos 2\omega E_n^{20} (E_n^{22})^*], \\
D_{\Omega\omega}^{(2)} &= -\cos ID_{\Omega\Omega}^{(2)}, \\
D_{\omega\omega}^{(2)} &= \cos^2 ID_{\Omega\Omega}^{(2)} + \frac{3P\gamma^6}{80e^2} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ |E_n^{11} + E_n^{13} + 2E_n^{21} - 2E_n^{23} + \frac{e}{2}(E_n^{20} - E_n^{24})|^2 + \frac{4}{3}(C_n^{11})^2 \right], \\
D_{\omega\varepsilon}^{(2)} &= -\frac{3P\gamma^7}{80e^2} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ |E_n^{11} + E_n^{13} + 2(E_n^{21} - E_n^{23}) + \frac{e}{2}(E_n^{20} - 4E_n^{22} - E_n^{24})|^2 + \frac{4}{3}C_n^{11}(C_n^{11} - 2eC_n^{20}) - 4|eE_n^{22}|^2 \right], \\
D_{\varepsilon\varepsilon}^{(2)} &= \frac{3P\gamma^8}{80e^2} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ |E_n^{11} + E_n^{13} + 2(E_n^{21} - E_n^{23}) + \frac{e}{2}(E_n^{20} - 8E_n^{22} - E_n^{24})|^2 + \frac{4}{3}(C_n^{11} - 2eC_n^{20})^2 \right], \\
D_{PI}^{(2)} &= D_{P\omega}^{(2)} = D_{P\Omega}^{(2)} = D_{Pe}^{(2)} = D_{eI}^{(2)} = D_{e\omega}^{(2)} = D_{e\Omega}^{(2)} = D_{e\varepsilon}^{(2)} = D_{I\varepsilon}^{(2)} = D_{\Omega\varepsilon}^{(2)} = 0. \tag{4.2}
\end{aligned}$$

The Hansen coefficients are defined by

$$\begin{aligned} C_n^{lm}(e) &\equiv \int_{t_0}^{t_0+P} \frac{dt}{P} \exp\left(\frac{2\pi i n t}{P}\right) \frac{\cos m\psi}{(1+e\cos\psi)^l}, \\ S_n^{lm}(e) &\equiv \int_{t_0}^{t_0+P} \frac{dt}{P} \exp\left(\frac{2\pi i n t}{P}\right) \frac{\sin m\psi}{(1+e\cos\psi)^l}, \\ E_n^{lm}(e) &\equiv C_n^{lm}(e) + S_n^{lm}(e), \end{aligned} \quad (4.3)$$

with primes indicating eccentricity derivatives, and the deterministic drift terms are those given in Sec. II C,

$$\begin{aligned} V_P &= -\frac{192\pi\eta v_P^5}{5\gamma^7} \left(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4\right), \\ V_e &= -\frac{608\pi\eta v_P^5}{15P\gamma^5} \left(e + \frac{121}{304}e^3\right), \\ V_\omega &= \frac{6\pi v_P^2}{P\gamma^2}, \\ V_\varepsilon &= \frac{2\pi v_P^2}{P} \left(6 - 7\eta - \frac{15 - 9\eta}{\gamma}\right). \end{aligned} \quad (4.4)$$

The values of some of these drift and diffusion coefficients as functions of eccentricity and harmonic frequency are shown in Figs. 3–6; we focus on the coefficients pertaining to the period and eccentricity of the binary, as these are the most important for our envisaged observational applications. These coefficients, describing the evolution of the size and shape of the orbit, are independent of the remaining four orbital elements  $(I, \Omega, \omega, \varepsilon)$ , which all describe the orientation of the orbit in space. In Figs. 4 and 3 we plot the coefficients against eccentricity, assuming various power-law SGWB spectra  $\Omega(f) \sim f^\alpha$ ,  $\alpha \in [-2, +2]$ . There are two particularly striking features that are worth mentioning: the first is that  $D_{Pe}^{(2)}$  and  $D_{ee}^{(2)}$  both vanish as  $e \rightarrow 1$ , regardless of the SGWB spectrum; and the second is that  $D_e^{(1)}$  changes sign at an eccentricity  $e \approx 0.4$  that is approximately (though not exactly) independent of the SGWB spectrum; the significance of this near-universal crossover eccentricity is not immediately clear.

In Figs. 5 and 6 we plot the contributions from each of the binary's harmonic frequencies  $f = n/P$ . We see that the evolution of the period is driven by the  $n = 2$  harmonic for near-circular binaries, with all other harmonics having zero contribution to  $D_{PP}^{(2)}$  and  $D_P^{(1)}$  in the circular limit  $e \rightarrow 0$ . This mirrors the frequency content of GW emission from binaries, which is also dominated by the  $n = 2$  harmonic for small eccentricity. The significance of  $n = 2$  here can be understood by recognising that advancing a circular orbit in time by  $P/2$  (i.e. the inverse of the  $n = 2$  harmonic) is equivalent to exchanging the positions of the two bodies, resulting in a setup which has the same GW emission and absorption properties. The eccentricity evolution, on the

other hand, is dominated by the  $n = 1$  and  $n = 3$  harmonics when  $e$  is small. In all cases, the contribution from higher harmonics becomes stronger for larger eccentricities, as the Fourier spectrum of the binary's response to GWs becomes richer. (Again, this is the same qualitative pattern that one finds in the case of GW emission by the binary.) In the limit  $e \rightarrow 1$ , we see a very simple pattern emerge in which each coefficient approaches a power law in the harmonic number  $n$  (except for  $D_{ee}^{(2)}$ , which vanishes in this limit).

The qualitative link between the GW emission and absorption spectra of binaries noted above is intriguing, and raises the question of whether this relationship can be established more formally. This might lead to deeper insights into binary–GW interactions, for example by allowing us to prove something akin to a fluctuation-dissipation theorem for this system. It would also be interesting to make contact with existing results on the scattering of GWs by binaries [60]. However, on the face of it there are several important differences between the GW absorption and emission processes: for example, the masses of the orbiting bodies are crucial in determining the radiated GW flux, but have no influence at all on GW absorption, since the GW-induced oscillations in the orbital separation are independent of the masses. (This last statement is a manifestation of the equivalence principle.) We leave a more thorough exploration of these questions for future work.

Inserting the coefficients in Eqs. (4.1) and (4.2) into Eq. (3.20), we obtain a FPE which completely describes the secular evolution of a general binary system under SGWB resonance; this is the main result of our analysis. Sections V, VI, and VII explore various strategies for solving this equation, and for using these solutions to place constraints on the SGWB spectrum. Before moving on, we derive some simplified expressions for the KM coefficients in the cases where the eccentricity and/or the inclination are small.

### B. Small-eccentricity and small-inclination orbits

For binaries with very small eccentricity, we want to recast the results above in terms of the alternative orbital elements  $(P, \zeta, \kappa, I, \Omega, \xi)$ , as defined in Eq. (2.18). We do this using the coordinate transformation laws for the KM coefficients (see, e.g., Sec. 4.9 of Ref. [54] for a derivation),

$$\begin{aligned} D_i^{(1)} &= \frac{\partial X_i}{\partial X_{i'}} D_{i'}^{(1)} + \frac{\partial^2 X_i}{\partial X_{i'} \partial X_{j'}} D_{i'j'}^{(2)}, \\ D_{ij}^{(2)} &= \frac{\partial X_i}{\partial X_{i'}} \frac{\partial X_j}{\partial X_{j'}} D_{i'j'}^{(2)}, \end{aligned} \quad (4.5)$$

where summation over the primed indices is implied. Neglecting terms of order  $e^2 \sim \zeta^2 \sim \zeta\kappa \sim \kappa^2$ , we find the drift coefficients

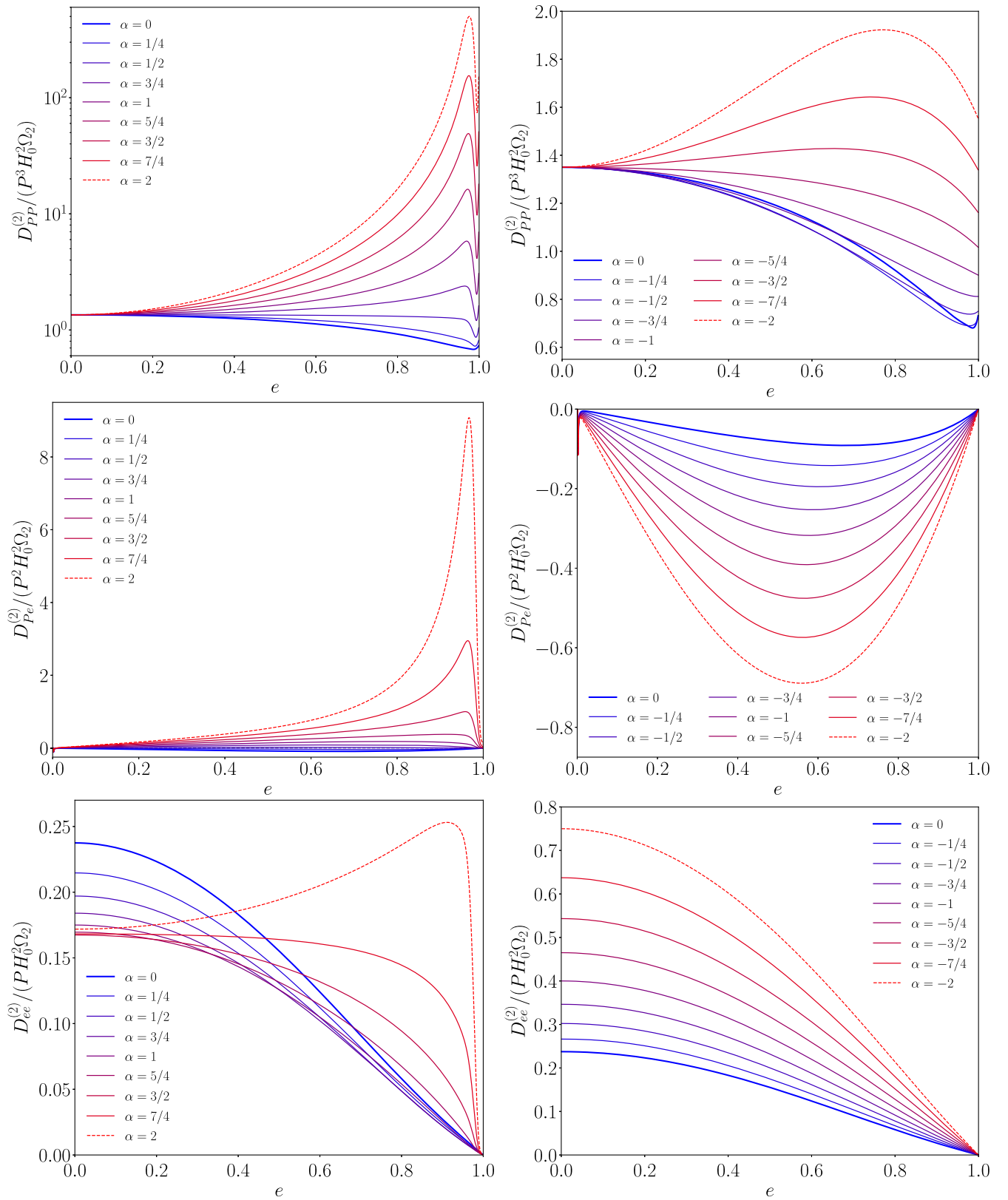


FIG. 3. The secular diffusion coefficients  $D_{PP}^{(2)}$ ,  $D_{Pe}^{(2)}$ , and  $D_{ee}^{(2)}$  (first, second, and third rows, respectively) as functions of eccentricity, for various power-law SGWB spectra,  $\Omega(f) \sim f^\alpha$ . The left column shows positive power-law indices,  $\alpha = 0, \dots, 2$ , while the right column shows negative indices,  $\alpha = 0, \dots, -2$ .

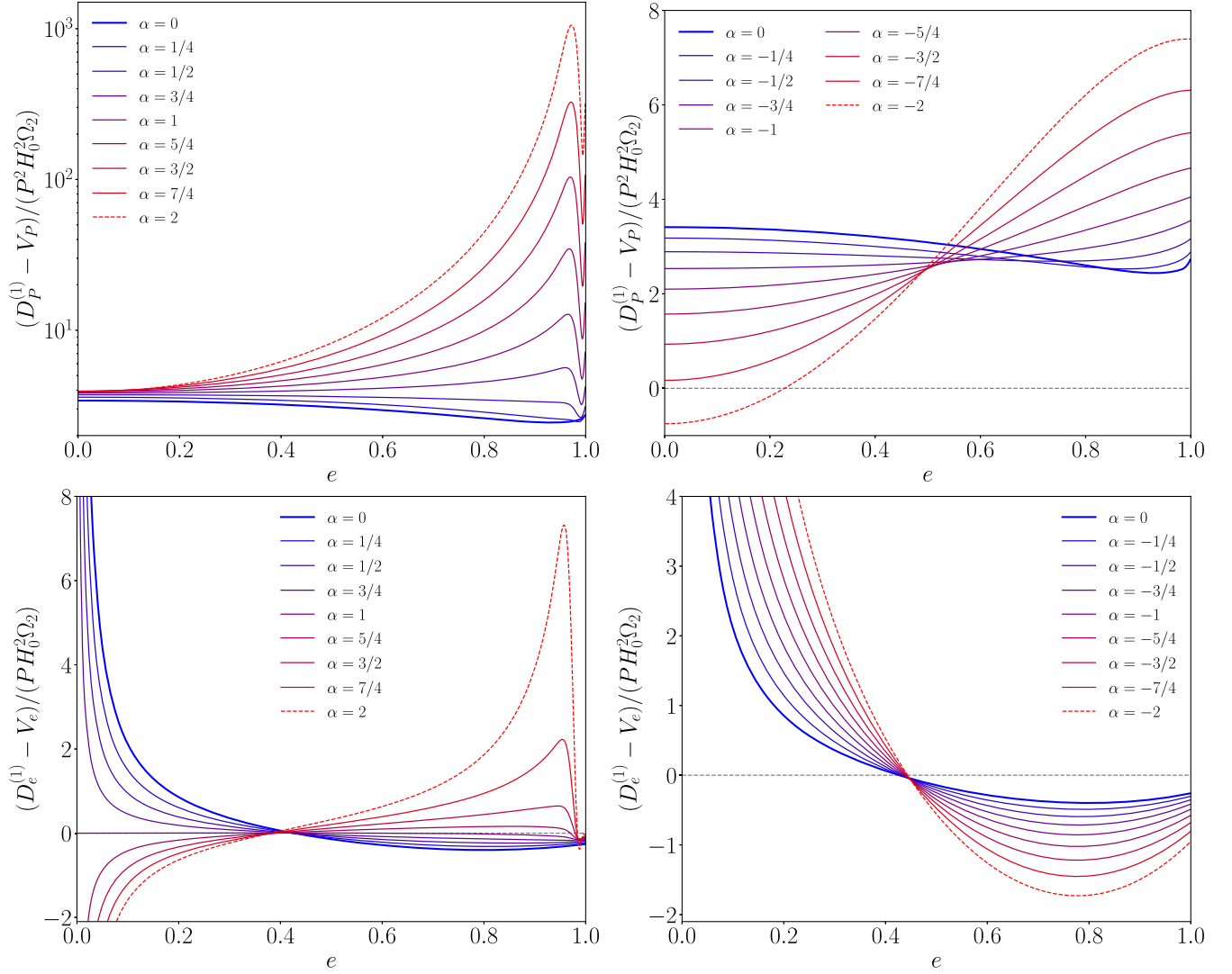


FIG. 4. Stochastic parts of the secular drift coefficients  $D_P^{(1)}$  (first row) and  $D_e^{(1)}$  (second row) as functions of eccentricity, for various power-law SGWB spectra,  $\Omega(f) \sim f^\alpha$ . The left column shows positive power-law indices,  $\alpha = 0, \dots, 2$ , while the right column shows negative indices,  $\alpha = 0, \dots, -2$ .

$$\begin{aligned}
 D_P^{(1)} &= V_P + \frac{3P^2}{160} H_0^2 (-79\Omega_1 + 288\Omega_2 - 27\Omega_3), \\
 D_\zeta^{(1)} &= V_\zeta + \frac{P}{160} \frac{\zeta}{\zeta^2 + \kappa^2} H_0^2 (25\Omega_1 - 27\Omega_3), \\
 D_\kappa^{(1)} &= V_\kappa + \frac{P}{160} \frac{\kappa}{\zeta^2 + \kappa^2} H_0^2 (25\Omega_1 - 27\Omega_3), \\
 D_I^{(1)} &= \frac{3P}{80} H_0^2 \Omega_2 \cot I, \\
 D_\Omega^{(1)} &= 0, \\
 D_\xi^{(1)} &= V_\xi,
 \end{aligned} \tag{4.6}$$

where the deterministic drift terms are given by

$$\begin{aligned}
 V_\zeta &= \frac{6\pi v_P^2}{P} \left( \kappa - \frac{304}{45} \eta v_P^3 \zeta \right), \\
 V_\kappa &= -\frac{6\pi v_P^2}{P} \left( \zeta + \frac{304}{45} \eta v_P^3 \kappa \right), \\
 V_\xi &= -\frac{4\pi v_P^2}{P} (3 - \eta),
 \end{aligned} \tag{4.7}$$

and the diffusion coefficients are

$$\begin{aligned}
 D_{PP}^{(2)} &= \frac{27P^3}{20} H_0^2 \Omega_2, \\
 D_{P\zeta}^{(2)} &= -\frac{3P\zeta}{160} H_0^2 (25\Omega_1 + 12\Omega_2 - 27\Omega_3), \\
 D_{P\kappa}^{(2)} &= -\frac{3P\kappa}{160} H_0^2 (25\Omega_1 + 12\Omega_2 - 27\Omega_3),
 \end{aligned}$$

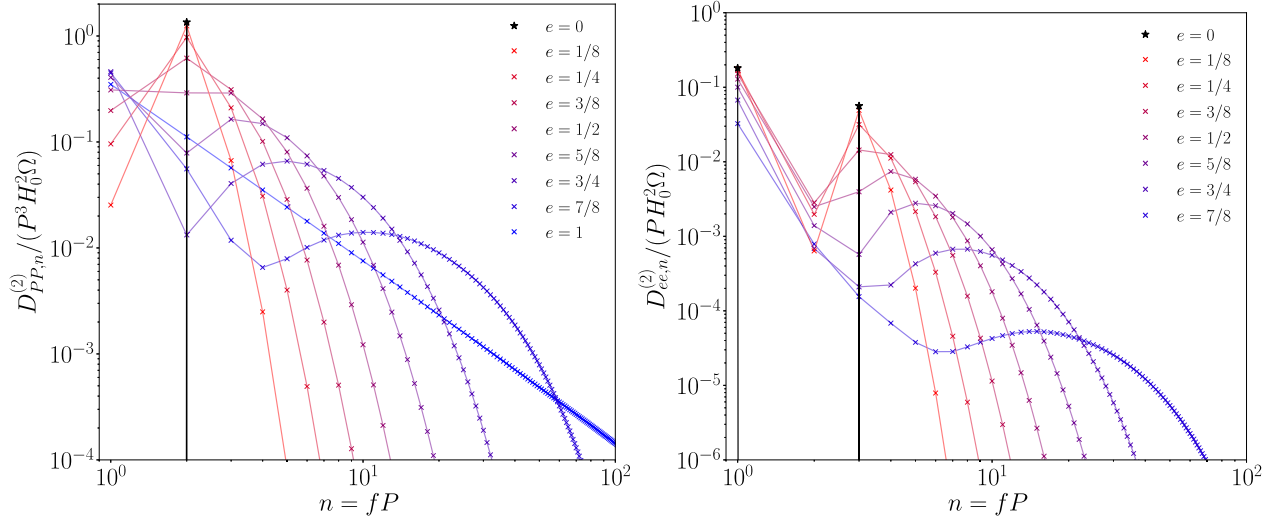


FIG. 5. Contributions to the secular diffusion coefficients  $D_{PP}^{(2)}$  (left panel) and  $D_{ee}^{(2)}$  (right panel) from different harmonic frequencies, for binaries with various eccentricities  $e = 0, \dots, 1$ . (The subscript “ $n$ ” here indicates that we have extracted the contribution from the  $n$ th harmonic.) We set  $\Omega(f) = \text{constant}$  here, so that each harmonic receives equal weighting; alternative SGWB spectra will give different weighting to each harmonic. Note that  $D_{ee}^{(2)} = 0$  when  $e \rightarrow 1$ .

$$\begin{aligned}
 D_{\zeta\zeta}^{(2)} &= D_{\kappa\kappa}^{(2)} = \frac{P}{160} H_0^2 (29\Omega_1 + 9\Omega_3), \\
 D_{\zeta\Omega}^{(2)} &= -\frac{3P\kappa}{80} H_0^2 \Omega_2 \frac{\cos I}{\sin^2 I}, \\
 D_{\zeta\xi}^{(2)} &= -\frac{P\kappa}{320} H_0^2 [203\Omega_1 - 12\Omega_2(20 + \cot^2 I) + 63\Omega_3], \\
 D_{\kappa\Omega}^{(2)} &= \frac{3P\zeta}{80} H_0^2 \Omega_2 \frac{\cos I}{\sin^2 I}, \\
 D_{\kappa\xi}^{(2)} &= \frac{P\zeta}{320} H_0^2 [203\Omega_1 - 12\Omega_2(20 + \cot^2 I) + 63\Omega_3], \\
 D_{II}^{(2)} &= \frac{3P}{80} H_0^2 \Omega_2, \\
 D_{\Omega\Omega}^{(2)} &= \frac{3P}{80} H_0^2 \Omega_2 \csc^2 I, \\
 D_{\Omega\xi}^{(2)} &= -\frac{3P}{80} H_0^2 \Omega_2 \frac{\cos I}{\sin^2 I}, \\
 D_{\xi\xi}^{(2)} &= \frac{3P}{80} H_0^2 \Omega_2 (16 + \cot^2 I), \\
 D_{PI}^{(2)} &= D_{P\Omega}^{(2)} = D_{P\xi}^{(2)} = D_{\zeta\kappa}^{(2)} \\
 &= D_{\zeta I}^{(2)} = D_{\kappa I}^{(2)} = D_{I\Omega}^{(2)} = D_{I\xi}^{(2)} = 0. \quad (4.8)
 \end{aligned}$$

(We have checked explicitly that the  $D_{PP}^{(2)}$  coefficient here matches the corresponding result in Ref. [34].)

We see from Eq. (4.6) that the stochastic drift term for the period  $P$  is usually positive, and can thus be interpreted physically as describing the softening of the binary due to the absorption of energy from the SGWB (the term can become negative if  $\Omega_2$  is significantly smaller than  $\Omega_1$  and/or  $\Omega_3$ , but SGWB spectra typically vary sufficiently slowly with frequency that this does not occur). Interestingly,

this implies that the net secular drift of the binary period (deterministic plus stochastic) generally changes sign at some critical value of  $P$ ; e.g., for a scale-invariant SGWB  $\Omega(f) = \text{constant}$ , this value is given by

$$\begin{aligned}
 P_{\text{crit}} &= \left( \frac{1024\pi\eta}{91H_0^2\Omega} \right)^{3/11} (2\pi GM)^{5/11} \\
 &\approx 95 \text{ yr} \times \left( \frac{\Omega}{10^{-6}} \frac{1/4}{\eta} \right)^{-3/11} \left( \frac{M}{M_\odot} \right)^{5/11}, \quad (4.9)
 \end{aligned}$$

which corresponds to a semimajor axis of

$$a_{\text{crit}} \approx 21 \text{ au} \times \left( \frac{\Omega}{10^{-6}} \frac{1/4}{\eta} \right)^{-2/11} \left( \frac{M}{M_\odot} \right)^{1/3}. \quad (4.10)$$

Binaries with  $P < P_{\text{crit}}$  will decay through GW emission, decreasing their period further, whereas binaries with  $P > P_{\text{crit}}$  undergo a net softening through SGWB absorption, leading to a further increase in their period. The point  $P = P_{\text{crit}}$  is thus an unstable fixed point of the Langevin equation for  $P$ .<sup>4</sup> Note however that random diffusion due to

<sup>4</sup>We can understand this instability through a thermodynamic analogy. In the absence of the SGWB, a binary system radiates GW energy with increasing intensity as it inspirals; its dynamical “temperature” grows as it loses energy, meaning that it has a *negative heat capacity*. Similarly, the SGWB acts as a heat reservoir and imparts energy to the binary (on average), therefore slowing the orbital motion and decreasing the system’s temperature. A conventional thermodynamic system with positive heat capacity would equilibrate at a point where the heat loss from GW radiation balanced the heat gain from the SGWB. Instead, the binary undergoes a runaway increase or decrease in its temperature due to its negative heat capacity.

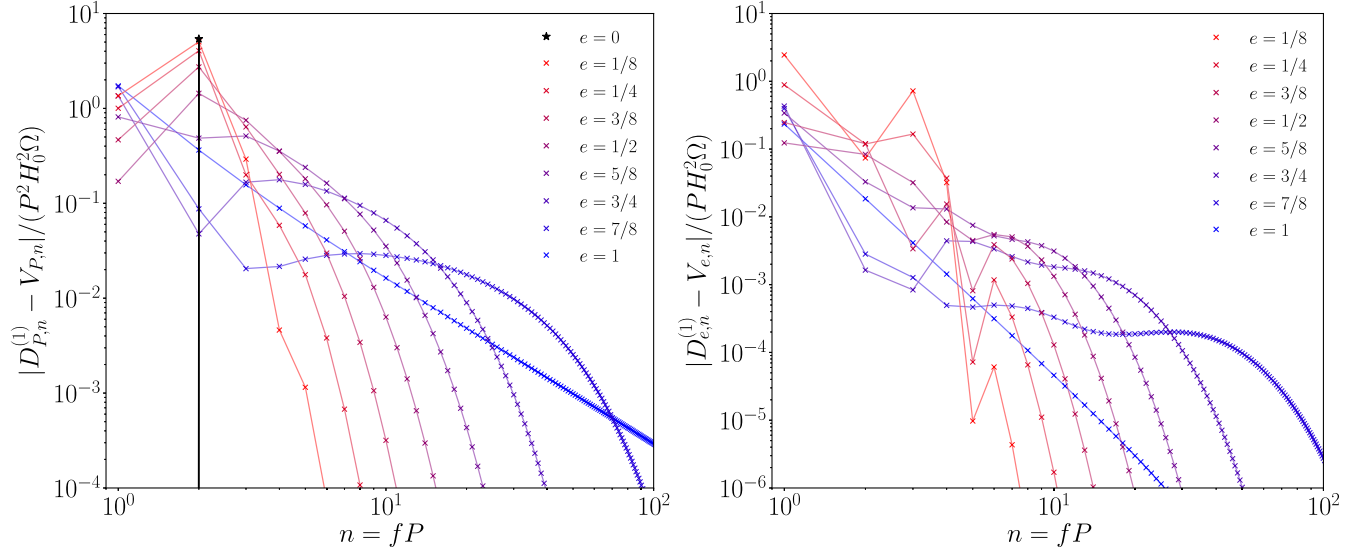


FIG. 6. Contributions to the stochastic parts of the secular drift coefficients  $D_p^{(1)}$  (left panel) and  $D_e^{(1)}$  (right panel) from different harmonic frequencies, for binaries with various eccentricities  $e = 0, \dots, 1$ . (The subscript “ $n$ ” here indicates that we have extracted the contribution from the  $n$ th harmonic.) We show the absolute values, as the drift coefficients have both positive and negative contributions. Note that  $D_e^{(1)} \rightarrow +\infty$  as  $e \rightarrow 0$ .

$D_{PP}^{(2)}$  acts on a similar timescale [see Fig. 7 and Eqs. (5.7) and (5.8)], and can easily push the system either side of this critical point.

We find a similar phenomenon for the eccentricity. Transforming back from the Laplace-Lagrange variables  $\zeta, \kappa$  for now, Eqs. (4.6) and (4.8) give

$$D_e^{(1)} = V_e + \frac{9P}{80e} H_0^2 (3\Omega_1 - \Omega_3), \quad (4.11)$$

so that the (usually positive) stochastic drift diverges as  $e \rightarrow 0$ , while the (always negative) deterministic part vanishes as  $e \rightarrow 0$ . The net eccentricity drift thus changes sign at a critical value, just as it does for the period. For example, assuming a scale-invariant SGWB spectrum, this critical value is

$$e_{\text{crit}} = \sqrt{\frac{27H_0^2\Omega}{4864\pi\eta v_p^5}} \approx 5.9 \times 10^{-5} \times \left(\frac{\Omega}{10^{-6}} \frac{1/4}{\eta}\right)^{1/2} \left(\frac{M}{M_\odot}\right)^{-5/6} \left(\frac{P}{\text{yr}}\right)^{11/6}. \quad (4.12)$$

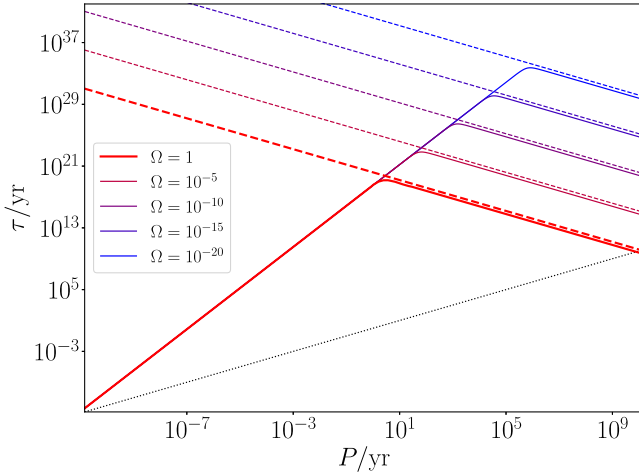


FIG. 7. Evolution timescales for the DF of a circular binary with total mass  $M = M_\odot$  immersed in a scale-invariant SGWB  $\Omega(f) = \text{constant}$ . Solid curves show the drift timescale (5.7), which is dominated by GW emission at “short” periods  $P \lesssim P_{\text{crit}}$  and by GW absorption from the SGWB at “long” periods  $P \gtrsim P_{\text{crit}}$ . Dashed curves show the diffusion timescale (5.8). The dotted black curve shows  $\tau = P$ ; the region below this curve corresponds to a timescale faster than the binary period, which breaks our secular averaging assumption and should thus be taken with a grain of salt.

Since  $\partial_e D_e^{(1)} < 0$  at this point,  $e_{\text{crit}}$  is a stable fixed point of the corresponding Langevin equation: binaries with larger eccentricity will tend to circularize through GW emission until they reach  $e_{\text{crit}}$ , while binaries with smaller eccentricity will on average have their eccentricity excited through SGWB resonance. This is particularly interesting from an observational point of view, as it suggests that eccentricities smaller than  $e_{\text{crit}}$  might be less frequently observed in sufficiently old systems, though random diffusion due to  $D_{ee}^{(2)}$  can still push systems below this point.

We also see from Eq. (4.6) that the stochastic drift term for the inclination changes sign at  $I = \pi/2$  [this is also true in the general-eccentricity case, see Eq. (4.1)]. Since  $\partial_I D_I^{(1)} < 0$  at  $I = \pi/2$ , this is a stable fixed point of the

corresponding Langevin equation; stochastic drift will, on average, drive binaries toward  $I = \pi/2$ . This effect is counteracted, however, by binaries diffusing away from  $I = \pi/2$ . As we show explicitly in Sec. V, on extremely long timescales these two effects balance each other, leaving an isotropic distribution for the inclination with mean  $\langle I \rangle = \pi/2$ .

In the limit where the binary's inclination is also small, we rewrite the KM coefficients in terms of the orbital elements  $(P, \zeta, \kappa, p, q, \lambda)$ . This gives the drift coefficients

$$\begin{aligned} D_p^{(1)} &= -\frac{Pp}{40} H_0^2 \Omega_2, \\ D_q^{(1)} &= -\frac{Pq}{40} H_0^2 \Omega_2, \end{aligned} \quad (4.13)$$

and the diffusion coefficients,

$$\begin{aligned} D_{\zeta\lambda}^{(2)} &= -\frac{P\kappa}{320} H_0^2 (203\Omega_1 - 240\Omega_2 + 63\Omega_3), \\ D_{\kappa\lambda}^{(2)} &= \frac{P\zeta}{320} H_0^2 (203\Omega_1 - 240\Omega_2 + 63\Omega_3), \\ D_{pp}^{(2)} &= D_{qq}^{(2)} = \frac{3P}{80} H_0^2 \Omega_2, \\ D_{p\lambda}^{(2)} &= \frac{3Pq}{160} H_0^2 \Omega_2 \\ D_{q\lambda}^{(2)} &= -\frac{3Pp}{160} H_0^2 \Omega_2 \\ D_{\lambda\lambda}^{(2)} &= \frac{3P}{5} H_0^2 \Omega_2, \\ D_{Pp}^{(2)} &= D_{Pq}^{(2)} = D_{P\lambda}^{(2)} = D_{\zeta\kappa}^{(2)} = D_{\zeta p}^{(2)} \\ &= D_{\zeta q}^{(2)} = D_{\kappa p}^{(2)} = D_{\kappa q}^{(2)} = D_{p q}^{(2)} = 0, \end{aligned} \quad (4.14)$$

where the coefficients not listed are identical to those in Eqs. (4.6) and (4.8). Note that this implies

$$D_I^{(1)} = \frac{3P}{80I} H_0^2 \Omega_2, \quad (4.15)$$

which diverges as  $I \rightarrow 0$ . This means that binaries are quickly excited away from zero inclination, similar to what happens for the eccentricity as  $e \rightarrow 0$ .

## V. SOME EXACT RESULTS FOR CIRCULAR BINARIES

We have shown that the osculating orbital elements of a binary coupled to the SGWB evolve according to a nonlinear six-dimensional FPE, for which no analytical solution is generally available. However, in the small-eccentricity limit  $e \rightarrow 0$ , we found in Eqs. (4.6) and (4.8) that the drift and diffusion of the binary period  $P$  are independent of all of the other orbital elements. This allows us to treat  $P$  separately by solving the one-dimensional FPE

$$\frac{\partial W}{\partial t} = -\frac{\partial J}{\partial P}, \quad (5.1)$$

where  $W(P, t)$  is now the single-variable DF for  $P$ , marginalized over the other orbital elements, and where

$$J(P, t) \equiv D^{(1)}W - \partial_P(D^{(2)}W) \quad (5.2)$$

is the *probability current* [54,55]. (We have suppressed the  $P$  subscripts on the drift and diffusion coefficients for this single-variable case.) In this section, we derive some exact results for this simplified equation. These results highlight the power of our Fokker-Planck formalism, which allows us to answer these questions about the full shape of the DF in a way that previous analyses are unable to.

### A. Quasistationary period distribution

The simplest kind of solution to look for is a stationary (i.e., time-independent) distribution, corresponding to constant probability current throughout the parameter space. Setting  $J = \text{constant}$  in Eq. (5.2), we can use the integrating factor (which is defined up to an arbitrary constant multiplicative factor)

$$I(P) \equiv \exp\left(\int dP \frac{D^{(1)}}{D^{(2)}}\right) \quad (5.3)$$

to obtain

$$W = \frac{I(P)}{D^{(2)}} \left( C - J \int \frac{dP}{I(P)} \right), \quad (5.4)$$

with  $C$  a constant which, for a given value of  $J$ , is fixed by the normalization of the DF. Clearly, the functional form of Eq. (5.4) depends on the SGWB energy spectrum  $\Omega(f)$ . As a simple example, consider a scale-invariant SGWB,  $\Omega(f) = \text{constant}$ , which has

$$\begin{aligned} D^{(1)} &= V_P + \frac{273}{80} P^2 H_0^2 \Omega, & D^{(2)} &= \frac{27}{20} P^3 H_0^2 \Omega, \\ I(P) &\propto P^{91/36} \exp\left[\frac{91}{132} \left(\frac{P_{\text{crit}}}{P}\right)^{11/3}\right]. \end{aligned} \quad (5.5)$$

A full exploration of spectra beyond this simple scale-invariant case is beyond our scope here, but we expect the qualitative results of this section to be reasonably robust to this choice.

We can fix  $J$  and  $C$  by imposing the appropriate boundary conditions. For some minimum value of  $P$  the binary merges or is tidally disrupted, whereas for some maximum value the binary becomes gravitationally unbound, so at both extremes we require absorbing boundary conditions—i.e., the DF must go to zero at both boundaries. Systems with absorbing boundary conditions do not admit nonzero stationary solutions [55]; formally, the conditional probability of the binary



having period  $P$  at time  $t$ , given that it initially had period  $P_0$  at time zero, obeys

$$\lim_{t \rightarrow \infty} W(P, t | P_0, 0) = 0 \quad (5.6)$$

across the entire parameter space. Intuitively, this is because all of the initial probability mass is eventually absorbed by one or other of the boundaries. We can also understand this in terms of the instability discussed in Sec. IV B; binaries either side of the critical period  $P_{\text{crit}}$  undergo a runaway evolution away from this point, reaching one of the two boundaries within finite time, thus leaving an empty distribution in the limit  $t \rightarrow \infty$ .

In practice, however, the timescale over which the binary evolves is set by its period, and binaries with short periods will approach stationarity much faster than binaries with long periods. More concretely, for a flat SGWB spectrum we have a drift timescale

$$\tau_{\text{drift}} \equiv \frac{P}{|D^{(1)}|} \simeq \begin{cases} \frac{80}{273PH_0^2\Omega}, & P \gg P_{\text{crit}}, \\ \frac{5P^{8/3}}{192\pi\eta(2\pi FM)^{5/3}}, & P \ll P_{\text{crit}}, \end{cases} \quad (5.7)$$

and a diffusion timescale

$$\tau_{\text{diff}} \equiv \frac{P^2}{|D^{(2)}|} = \frac{20}{27PH_0^2\Omega}. \quad (5.8)$$

As shown in Fig. 7, the fastest timescale is many orders of magnitude shorter near the lower boundary than it is near the upper boundary. It may be reasonable, therefore, to look for “quasistationary” solutions, where equilibrium is established near the lower boundary, but where the boundary condition for long periods is neglected. We achieve this by choosing  $J$  and  $C$  such that the DF goes to zero at the lower boundary, and such that the DF is normalized, but without enforcing any condition at the upper boundary.

As a simple example, consider a scale-invariant SGWB spectrum  $\Omega(f) = \text{constant}$ , and fix the DF to be zero at the period corresponding to the binary’s innermost stable circular orbit (ISCO),

$$P_{\text{ISCO}} \equiv 6^{3/2} \times 2\pi GM, \quad (5.9)$$

with the binary assumed to merge at periods shorter than this. The corresponding quasistationary distribution is then given in terms of the dimensionless variable  $q \equiv P/P_{\text{ISCO}}$  by

$$W_{\text{qs}} \propto \exp\left(\frac{\lambda}{q^{11/3}}\right) \left[ \frac{E_{77/132}(\lambda q^{-11/3})}{q^2} - \frac{E_{77/132}(\lambda)}{q^{17/36}} \right], \quad (5.10)$$

where  $E_n(z) \equiv \int_1^\infty dt e^{-zt} t^{-n}$  is the exponential integral function, and

$$\lambda \equiv \frac{91}{132} \left( \frac{P_{\text{crit}}}{P_{\text{ISCO}}} \right)^{11/3} = \frac{\sqrt{2}\eta(GMH_0)^{-2}}{8019\sqrt{3}\pi\Omega} \gg 1 \quad (5.11)$$

is a dimensionless constant which quantifies the strength of the deterministic drift  $V_P$  relative to the secular diffusion  $D^{(2)}$ . While the functional form of Eq. (5.10) is somewhat opaque, one can show that it approximately interpolates between  $W_{\text{qs}} \sim P^{5/3}$  for  $P_{\text{ISCO}} \ll P \ll P_{\text{crit}}$  and  $W_{\text{qs}} \sim P^{-17/36}$  for  $P \gg P_{\text{crit}}$  (this broken-power-law behavior is clearly seen in Fig. 8). We normalize Eq. (5.10) by integrating up to some maximum period, which we choose to be the age of the Universe. By substituting Eq. (5.10) into Eq. (5.2) and evaluating at  $q = 1$  where  $W = 0$  and  $\partial W/\partial P > 0$ , we see that the probability current is strictly negative,  $J < 0$ . Since we specified that  $J = \text{constant}$ , this means there is a uniform net flow toward shorter periods throughout the parameter space.

For hard binaries with  $P < P_{\text{crit}}$  it is obvious that we should have  $J < 0$ , as the negative deterministic drift is the most important effect and quickly drives the binary toward merger. For soft binaries with  $P > P_{\text{crit}}$  the negative probability current is less immediately obvious; it shows us that, while there is a net drift  $D^{(1)}$  pushing these binaries toward longer periods, on average they are nonetheless expected to flow toward shorter periods. We can understand this somewhat counterintuitive behavior by including

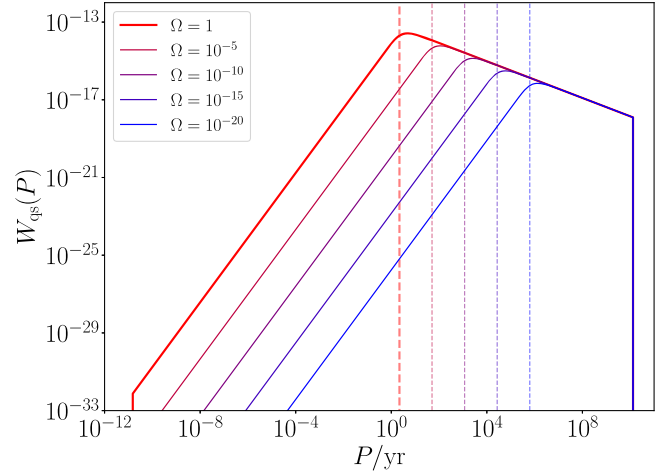


FIG. 8. The quasistationary distribution (5.10) for the period  $P$  of a binary with mass  $M = M_\odot$  coupled to a scale-invariant SGWB,  $\Omega(f) = \text{constant}$ . The upper and lower cutoffs are due to the age of the Universe and the binary’s ISCO period (5.9), respectively. The dashed vertical lines indicate the value of  $P_{\text{crit}}$  for each  $\Omega$ , as defined by Eq. (4.9)—this is roughly the peak of the distribution in each case. Binaries with periods  $P < P_{\text{crit}}$  decay deterministically through GW emission and are removed from the distribution, whereas binaries with  $P > P_{\text{crit}}$  are supported against decay by resonant absorption of the SGWB. As we lower the SGWB intensity  $\Omega$ , the resonance becomes weaker, and the binaries must have longer periods to avoid decay.

diffusion as well as drift effects. Indeed, note that by Eq. (5.2), the condition  $J < 0$  is equivalent to

$$\frac{\partial(\ln D^{(2)}W)}{\partial(\ln P)} > \frac{PD^{(1)}}{D^{(2)}} = \frac{91}{36}[1 - (P/P_{\text{crit}})^{-11/3}], \quad (5.12)$$

where the rhs tends to a positive constant value in the  $P \gg P_{\text{crit}}$  region we are interested in. What this is telling us is that, despite the positive drift coefficient, we can still have a negative probability current if and only if the diffusion coefficient  $D^{(2)}$  grows sufficiently quickly with  $P$ , as this makes it sufficiently likely for the binary's random walk to wander below  $P_{\text{crit}}$  and then rapidly approach short periods through GW emission. Interestingly, we find that for the quasistationary distribution we have  $\partial(\ln D^{(2)}W_{\text{qs}})/\partial(\ln P) \simeq 91/36$ , so this diffusive effect is *only just* strong enough to cause a net negative probability current.

We can repeat the calculations above for any given SGWB spectrum to write down a corresponding quasistationary solution for the period of a circular binary. However, the time taken to relax to this distribution is extremely long for typical SGWB spectra (see Fig. 7), so these solutions are physically uninteresting in most cases. Besides, the assumption of a perfectly circular binary is overly simplistic, as we have shown in Sec. IV B that the eccentricity distribution relaxes away from zero on shorter timescales. Nonetheless, the approach in this section is still useful for building intuitive understanding of the dynamics of the full DF, and may be useful for, e.g., studies of the orbital element distributions of old populations of binaries, where the full shape of the distribution is vitally important. Finding quasistationary solutions for the full multivariate FPE is much more challenging, and for eccentric binaries there is no guarantee that such a solution with the appropriate lower boundary condition even exists.

### B. Mean coalescence time

For any one-dimensional FPE, it is possible to write down an explicit formula for the *mean first passage time* at either of its boundaries—i.e., the average time taken for an individual random trajectory to reach that boundary, as a function of the initial position [55]. In our case, this is a useful tool for understanding how the presence of a particular SGWB spectrum impacts upon the eventual fate of a binary system. Applying this to our lower absorbing boundary at the ISCO, we thus have the mean coalescence time of a binary coupled to the SGWB, as a function of its initial period  $P_i$ ,

$$\langle \tau(P_i) \rangle = \int_{P_{\text{ISCO}}}^{P_i} dP \int_P^{P_{\text{max}}} dP' \frac{I(P')}{I(P)D^{(2)}(P')}, \quad (5.13)$$

where  $I(P)$  is the integrating factor defined in Eq. (5.3).

Returning to the example of a scale-invariant SGWB, this becomes

$$\langle \tau(P_i) \rangle = \int_1^{e_i} \frac{dq}{q^{91/36}} \int_q^{e_{\text{max}}} \frac{dq'}{q'^{17/36}} \frac{20e^{\lambda(q'^{-11/3} - q^{-11/3})}}{27P_{\text{ISCO}}H_0^2\Omega}. \quad (5.14)$$

This double integral is challenging to evaluate in general. However, we can easily verify Eq. (5.14) by showing that it reproduces the standard expression for the coalescence time due to deterministic GW emission in cases where the SGWB resonance is weak. Setting  $P_i \ll P_{\text{crit}}$ , we can safely take the limit  $\lambda \rightarrow \infty$  and extract the leading-order and next-to-leading-order terms,

$$\begin{aligned} \langle \tau(P_i) \rangle &\simeq \frac{405GM}{16\eta} \left( \frac{P_i}{P_{\text{ISCO}}} \right)^{8/3} \\ &\times \left[ 1 - \frac{96}{91} e^{-\frac{91}{132}(P_{\text{crit}}/P_i)^{11/3}} \frac{P_{\text{max}}^{151/36}}{P_i^{19/36} P_{\text{crit}}^{11/3}} \right]. \end{aligned} \quad (5.15)$$

The leading-order term agrees with the deterministic coalescence time one finds by integrating  $V_P$ , as expected. There is a small negative contribution from the next-to-leading-order term, which indicates that SGWB resonance slightly speeds up the coalescence in this regime. We can understand this by noticing that the term associated with  $D^{(2)}$  in Eq. (5.2) is always negative for the quasistationary distribution, so that diffusion always has a net negative contribution to the probability current and thus, on average, always help drive binaries toward merger.

### C. Including the remaining orbital elements

Having found a quasistationary period distribution, it is now relatively easy to obtain stationary distributions for the remaining orbital elements  $(I, \Omega, \xi)$ , so long as we hold  $\zeta$  and  $\kappa$  fixed at zero. To do so, we write the full FPE as

$$\frac{\partial W}{\partial t} = -\partial_i J_i, \quad J_i \equiv D_i^{(1)}W - \partial_j(D_{ij}^{(2)}W), \quad (5.16)$$

where  $W$  is now interpreted as the multivariate DF over the four orbital elements  $(P, I, \Omega, \xi)$ , and the four corresponding probability currents are

$$\begin{aligned} J_P &= D_P^{(1)}W - \partial_P(D_{PP}^{(2)}W), \\ J_I &= D_I^{(1)}W - D_{II}^{(2)}\partial_I W, \\ J_\Omega &= -D_{\Omega\Omega}^{(2)}\partial_\Omega W - D_{\Omega\xi}^{(2)}\partial_\xi W, \\ J_\xi &= V_\xi W - D_{\Omega\xi}^{(2)}\partial_\Omega W - D_{\xi\xi}^{(2)}\partial_\xi W. \end{aligned} \quad (5.17)$$

We have used Eqs. (4.6) and (4.8) to simplify these expressions, in particular using the independence of the diffusion terms from most of the orbital elements to take them outside the partial derivatives.

We first consider the inclination  $I$ . By definition, this is constrained to lie in the interval  $[0, \pi]$ . Unlike for the period  $P$ , a binary reaching one of the boundaries of this interval is not removed from the distribution; instead of absorbing boundary conditions, we have *reflecting* boundary conditions, i.e., the probability current  $J_I$  vanishes at both boundaries. However, if the distribution is stationary then  $J_I = \text{constant}$ , so the current must vanish everywhere on  $[0, \pi]$ . Setting  $J_I = 0$  in Eq. (5.17) gives

$$\partial_I \ln W = D_I^{(1)} / D_{II}^{(2)} = \cot I. \quad (5.18)$$

Integrating this, we find

$$W \propto \sin I, \quad (5.19)$$

which corresponds to an isotropic distribution (since  $\cos I$  is uniformly distributed). This makes intuitive sense: on long timescales, SGWB resonance causes the binary to “forget” its initial orbital plane, such that the resulting stationary distribution is spherically symmetric. This also agrees with our finding in Sec. IV B that  $D_I^{(1)} = 0$  at  $I = \pi/2$ , which is the mean inclination of an isotropic distribution.

For  $\Omega$  and  $\xi$  it is natural to impose periodic boundary conditions for the DF and for the probability currents,

$$\begin{aligned} W(P, I, \Omega, \xi, t) &= W(P, I, \Omega + 2\pi, \xi, t) \\ &= W(P, I, \Omega, \xi + 2\pi, t), \\ J_i(P, I, \Omega, \xi, t) &= J_i(P, I, \Omega + 2\pi, \xi, t) \\ &= J_i(P, I, \Omega, \xi + 2\pi, t). \end{aligned} \quad (5.20)$$

Stationarity requires  $\partial_\Omega J_\Omega = \partial_\xi J_\xi = 0$ . By inspection, we see that this is achieved if  $\partial_\Omega W = \partial_\xi W = 0$ , so that the DF depends only on  $P$  and  $I$ . This corresponds to  $\Omega$  and  $\xi$  being uniformly distributed, which also satisfies the periodic boundary conditions;  $\Omega$  then has zero probability current, while  $\xi$  has a uniform negative current due to the deterministic drift  $V_\xi$ , which depends only on the period.

## VI. SOLVING THE FULL FOKKER-PLANCK EQUATION

We now consider the full FPE for all six orbital elements  $(P, e, I, \Omega, \omega, \varepsilon)$ . Allowing nonzero eccentricity  $e > 0$  leads to much more complicated KM coefficients, and means that the results of Sec. V are no longer applicable. Nonetheless, we can use the fact that diffusion of the orbital elements due to the SGWB takes place on very long timescales  $\tau_{\text{diff}} \sim 1/(PH_0^2\Omega) \gg P$  (see Fig. 7). This allows us to develop some useful approximate solution schemes for much shorter observational timescales.

### A. Perturbative short-time solution

The FPE can be written as an operator equation

$$\frac{\partial W}{\partial t} = L_{\text{FP}} W, \quad (6.1)$$

defined by the linear differential operator,

$$L_{\text{FP}}(\mathbf{X}) \equiv -\partial_i D_i^{(1)} + \partial_i \partial_j D_{ij}^{(2)}. \quad (6.2)$$

This has the formal solution<sup>5</sup>

$$W(\mathbf{X}, t) = e^{tL_{\text{FP}}} W(\mathbf{X}, 0), \quad (6.3)$$

which, for short times  $t \ll \tau_{\text{diff}}$ , can be expanded as

$$W(\mathbf{X}, t) = [1 + tL_{\text{FP}}(\mathbf{X}, 0) + \mathcal{O}(t/\tau_{\text{diff}})^2] W(\mathbf{X}, 0), \quad (6.4)$$

where the rhs depends only on data at time zero.

Suppose that at time zero the binary’s orbital elements take on the “sharp” values  $x_i$ . The initial condition for the DF is then

$$W(\mathbf{X}, 0) = \delta^{(6)}(\mathbf{X} - \mathbf{x}). \quad (6.5)$$

By using a Fourier representation of the delta function, we can evaluate Eq. (6.4) to find [54]

$$\begin{aligned} W(\mathbf{X}, t) &= \frac{1}{\sqrt{\det 4\pi t D^{(2)}}} \\ &\times \exp \left\{ -\frac{[D^{(2)}]_{ij}^{-1}}{4t} (X_i - x_i - D_i^{(1)} t)(X_j - x_j - D_j^{(1)} t) \right\} \\ &+ \mathcal{O}(t/\tau_{\text{diff}})^2, \end{aligned} \quad (6.6)$$

i.e., on short timescales, the DF is a multivariate Gaussian with mean  $x_i + D_i^{(1)} t$  and covariance matrix  $2tD_{ij}^{(2)}$ . Here  $[D^{(2)}]_{ij}^{-1}$  represents the elements of the inverse of the diffusion matrix, and both the drift vector and diffusion matrix are evaluated at  $(\mathbf{X}, t) = (\mathbf{x}, 0)$ .

### B. Evolution of moments of the orbital elements

The short-time expansion shows that on observational timescales the DF of the orbital elements is approximately Gaussian, and is therefore completely characterized by its first two moments: the mean and the covariance matrix, which we write as

<sup>5</sup>Here we take advantage of the time-independence of the secular KM coefficients. For time-dependent coefficients, one would need to instead construct a time-ordered Dyson series [54], though this gives the same result for short times.

$$\begin{aligned}\bar{X}_i &\equiv \langle X_i \rangle, \\ C_{ij} &\equiv \text{Cov}[X_i, X_j] = \langle (X_i - \bar{X}_i)(X_j - \bar{X}_j) \rangle.\end{aligned}\quad (6.7)$$

It is therefore useful to take moments of the FPE to find the time evolution of these quantities, rather than attempting to calculate the full time-dependent DF. In doing so, we can calculate the backreaction of perturbations on the evolution of the binary, obtaining corrections to the linear growth found in Eq. (6.6).

The first moment of the FPE gives

$$\begin{aligned}\dot{\bar{X}}_i &= \frac{\partial}{\partial t} \int d\mathbf{X} X_i W = \int d\mathbf{X} X_i \frac{\partial W}{\partial t} \\ &= - \int d\mathbf{X} X_i \partial_j (D_j^{(1)} W) + \int d\mathbf{X} X_i \partial_j \partial_k (D_{jk}^{(2)} W).\end{aligned}\quad (6.8)$$

We integrate by parts, and assume that the DF falls off fast enough that all boundary terms vanish, leaving

$$\dot{\bar{X}}_i = \int d\mathbf{X} D_i^{(1)} W = \langle D_i^{(1)} \rangle. \quad (6.9)$$

This fall-off assumption is very reasonable here, as the diffusion rate is extremely small for realistic binaries, so the DF will only have support very near to the mean value. Doing the same for the second moment gives

$$\begin{aligned}\frac{d}{dt} \langle X_i X_j \rangle &= \int d\mathbf{X} X_i X_j \frac{\partial W}{\partial t} \\ &= - \int d\mathbf{X} X_i X_j \partial_k (D_k^{(1)} W) \\ &\quad + \int d\mathbf{X} X_i X_j \partial_k \partial_\ell (D_{k\ell}^{(2)} W) \\ &= \int d\mathbf{X} (X_i D_j^{(1)} + X_j D_i^{(1)}) W + 2 \int d\mathbf{X} D_{ij}^{(2)} W \\ &= \langle X_i D_j^{(1)} \rangle + \langle X_j D_i^{(1)} \rangle + 2 \langle D_{ij}^{(2)} \rangle.\end{aligned}\quad (6.10)$$

We can combine these to give the evolution equation for the covariance matrix,

$$\begin{aligned}\dot{C}_{ij} &= \frac{d}{dt} (\langle X_i X_j \rangle - \bar{X}_i \bar{X}_j) \\ &= \langle X_i D_j^{(1)} \rangle + \langle X_j D_i^{(1)} \rangle + 2 \langle D_{ij}^{(2)} \rangle \\ &\quad - \langle X_i \rangle \langle D_j^{(1)} \rangle - \langle X_j \rangle \langle D_i^{(1)} \rangle \\ &= \text{Cov}[X_i, D_j^{(1)}] + \text{Cov}[X_j, D_i^{(1)}] + 2 \langle D_{ij}^{(2)} \rangle.\end{aligned}\quad (6.11)$$

### C. The slow-diffusion approximation

Eqs. (6.9) and (6.11) fully describe the evolution of the mean and covariance of the orbital elements. However, they are given in terms of ensemble averages over nonlinear

functions of the orbital elements, which we cannot perform without knowing the full DF. Even if we were to evaluate them approximately by assuming a Gaussian distribution, the resulting expressions would be very cumbersome.

In the case where the variance is small and any given orbital element  $X_i$  is ‘‘close’’ to its mean value  $\langle X_i \rangle$  (in a probabilistic sense), one can instead Taylor expand an arbitrary function of the elements around the mean,

$$\begin{aligned}f(\mathbf{X}) &= f(\bar{\mathbf{X}}) + (X_i - \bar{X}_i) \partial_i f(\bar{\mathbf{X}}) \\ &\quad + \frac{1}{2} (X_i - \bar{X}_i)(X_j - \bar{X}_j) \partial_i \partial_j f(\bar{\mathbf{X}}) + \dots,\end{aligned}\quad (6.12)$$

so that the mean of the function is approximated by

$$\langle f(\mathbf{X}) \rangle \simeq f(\bar{\mathbf{X}}) + \frac{1}{2} C_{ij} \partial_i \partial_j f(\bar{\mathbf{X}}). \quad (6.13)$$

(Note that the first-order term in the expansion vanishes when taking the mean, as the first central moment is identically zero.)

We can justify using Eq. (6.13) by noting that the diffusion matrix calculated in Sec. IV is very small in most physical situations. To keep track of how this smallness propagates into the evolution equations, we introduce a formal small parameter  $\epsilon$  (which we will later set to unity), writing

$$D_{ij}^{(2)} \rightarrow \epsilon D_{ij}^{(2)}. \quad (6.14)$$

For sharp initial conditions  $\bar{X}_i = x_i$ ,  $C_{ij} = 0$ , we see from Eq. (6.11) that  $C_{ij} = \mathcal{O}(\epsilon)$ , as the  $\text{Cov}[X, D^{(1)}]$  terms are initially zero. We therefore also write

$$C_{ij} \rightarrow \epsilon C_{ij}. \quad (6.15)$$

We thus see that Eq. (6.13) is justified if we neglect terms of order  $\epsilon^2$ . We call this the *slow-diffusion* approximation, as it relies on the fact that the timescale  $\tau_{\text{diff}}$  over which the covariance grows is long compared to the observation time. Note that the stochastic contribution to the drift vector is generally of the same order as the diffusion matrix, so we also write

$$D_i^{(1)} = V_i + \epsilon \delta D_i^{(1)}, \quad (6.16)$$

where the stochastic term  $\delta D_i^{(1)}$  is suppressed by a factor of  $\epsilon$ .

Applying this approximation to Eqs. (6.9) and (6.11), we obtain the moment evolution equations to first order in  $\epsilon$ ,

$$\begin{aligned}\dot{\bar{X}}_i &= V_i + \epsilon \delta D_i^{(1)} + \frac{1}{2} \epsilon C_{jk} \partial_j \partial_k V_i + \mathcal{O}(\epsilon^2), \\ \epsilon \dot{C}_{ij} &= 2\epsilon D_{ij}^{(2)} + \epsilon C_{ik} \partial_k V_j + \epsilon C_{jk} \partial_k V_i + \mathcal{O}(\epsilon^2),\end{aligned}\quad (6.17)$$

where the drift vector and diffusion matrix are both evaluated at  $\bar{\mathbf{X}}$ .

We see that the evolution equation (6.17) for the mean orbital elements includes both  $\mathcal{O}(\epsilon^0)$  and  $\mathcal{O}(\epsilon^1)$  terms. For numerical reasons, it is convenient to separate these. We therefore write the mean orbital elements as

$$\bar{\mathbf{X}}(t) = \bar{\mathbf{X}}_0(t) + \epsilon \delta \bar{\mathbf{X}}(t), \quad (6.18)$$

where  $\bar{\mathbf{X}}_0$  represents the values the elements would take in the absence of the SGWB, which obey the deterministic evolution equation

$$\dot{\bar{X}}_{0,i} = V_i(\bar{\mathbf{X}}_0). \quad (6.19)$$

(Note that this separation of deterministic and stochastic parts of the drift is always possible, regardless of the size of the deterministic term.) Meanwhile  $\delta \bar{\mathbf{X}}$  represents the deviation in the mean due to SGWB resonance, and evolves according to

$$\begin{aligned}\epsilon \delta \dot{\bar{X}}_i &= V_i(\bar{\mathbf{X}}) - V_i(\bar{\mathbf{X}}_0) + \epsilon \delta D_i^{(1)}(\bar{\mathbf{X}}) \\ &+ \frac{1}{2} \epsilon C_{jk} \partial_j \partial_k V_i(\bar{\mathbf{X}}) + \mathcal{O}(\epsilon^2).\end{aligned}\quad (6.20)$$

We can Taylor-expand the terms that are evaluated at  $\bar{\mathbf{X}} = \bar{\mathbf{X}}_0 + \epsilon \delta \bar{\mathbf{X}}$  to give

$$\begin{aligned}\epsilon \delta \dot{\bar{X}}_i &= \epsilon \delta \bar{X}_j \partial_j V_i(\bar{\mathbf{X}}_0) + \epsilon \delta D_i^{(1)}(\bar{\mathbf{X}}_0) \\ &+ \frac{1}{2} \epsilon C_{jk} \partial_j \partial_k V_i(\bar{\mathbf{X}}_0) + \mathcal{O}(\epsilon^2).\end{aligned}\quad (6.21)$$

With the appropriate terms identified, we can send  $\epsilon \rightarrow 1$ . Our full set of evolution equations, to leading order in the slow-diffusion approximation, reads

$$\begin{aligned}\dot{\bar{X}}_{0,i} &= V_i, \\ \delta \dot{\bar{X}}_i &\simeq \delta D_i^{(1)} + \delta \bar{X}_j \partial_j V_i + \frac{1}{2} C_{jk} \partial_j \partial_k V_i, \\ \dot{C}_{ij} &\simeq 2D_{ij}^{(2)} + C_{ik} \partial_k V_j + C_{jk} \partial_k V_i,\end{aligned}\quad (6.22)$$

with all terms evaluated at the deterministic mean elements  $\bar{\mathbf{X}}_0$ .

It is interesting to note that, starting from sharp initial conditions, some elements of the covariance matrix remain fixed at zero under Eq. (6.22):

$$C_{Pl} = C_{P\Omega} = C_{eI} = C_{e\Omega} = C_{I\epsilon} = C_{\Omega\epsilon} = 0, \quad (6.23)$$

i.e., stochastic perturbations to these pairs of orbital elements remain statistically uncorrelated at first order in the slow-diffusion approximation. This is due to the vanishing of the corresponding elements of the diffusion matrix  $D_{ij}^{(2)}$  in Eq. (4.2), as well as the fact that  $V_I = V_\Omega = 0$ . However, all other elements of the covariance matrix generically grow over time.

In Fig. 9, we show an example of an integration of Eq. (6.22) for the Hulse-Taylor binary pulsar B1913 + 16 [61]. We see that, in this example, the period-eccentricity sector of the covariance matrix grows linearly with time,

$$C_{PP} \sim C_{Pe} \sim C_{ee} \sim t, \quad (6.24)$$

indicating that the right-hand side of Eq. (6.22) is dominated by the  $D^{(2)}$  term for these components. In contrast, many of the components involving the argument of pericenter and compensated mean anomaly (specifically  $C_{P\omega}$ ,  $C_{Pe}$ ,  $C_{e\omega}$ ,  $C_{ee}$ ,  $C_{\omega\omega}$ ,  $C_{\omega\epsilon}$ , and  $C_{e\epsilon}$ ) have relatively smaller values for the diffusion matrix, and are instead driven by the  $C\partial V$  terms, leading them to grow like  $\sim t^2$  (since the components of the covariance matrix sourcing them grow like  $C \sim t$ ). The components  $C_{II}$ ,  $C_{I\Omega}$ ,  $C_{I\omega}$ ,  $C_{\Omega\Omega}$ , and  $C_{\Omega\omega}$  evolve more erratically; this is due to the presence of  $\sin 2\omega$ ,  $\cos 2\omega$  terms in the corresponding components of the diffusion matrix [see Eq. (4.2)], which oscillate on a timescale  $\pi/\dot{\omega}_{\text{sec}} \approx 43$  yr due to the perihelion precession of the Hulse-Taylor system [62,63]. Finally, the remaining six components of the covariance matrix are zero throughout the integration time, as expected from Eq. (6.23).

In Fig. 10 we show the distribution of the orbital elements at the end of the integration in Fig. 9. We see that the orbital elements are, on the whole, weakly correlated with each other, with the notable exceptions of the pairs  $(P, \omega)$ ,  $(P, e)$ , and particularly  $(\omega, \epsilon)$ , which are highly covariant. Note that these distributions include the overall offset due to stochastic drift, but that this is less important than diffusion in this case, and is therefore harder to distinguish by eye.

#### D. Growth rate of non-Gaussianity

By using Eq. (6.22) we have neglected all higher-order moments of the distribution, and thus fail to capture any non-Gaussianities in the DF. We can measure the departure from Gaussianity by tracking the evolution of the third central moment,

$$S_{ijk} \equiv \langle (X_i - \bar{X}_i)(X_j - \bar{X}_j)(X_k - \bar{X}_k) \rangle, \quad (6.25)$$

as this vanishes identically for a Gaussian distribution. In particular, for each individual orbital element  $X_i$  we have, to leading order in the slow-diffusion approximation,

$$\dot{S}_{iii} \simeq -3\bar{X}_i C_{ij} \partial_j V_i, \quad (6.26)$$

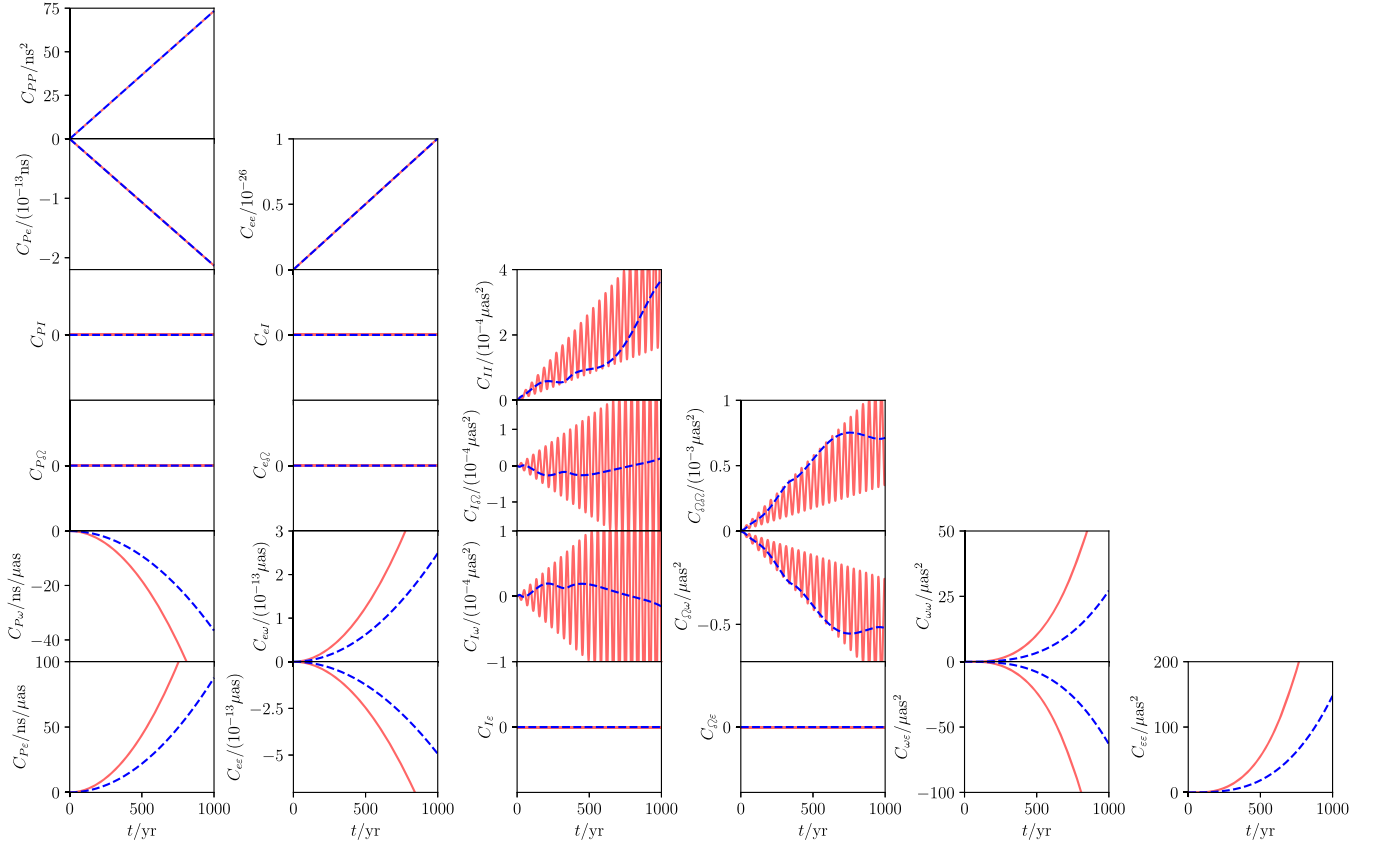


FIG. 9. The covariance matrix  $C_{ij}(t)$  of the orbital elements of the binary pulsar B1913 + 16 (the Hulse-Taylor system) over a 1 kyr interval, assuming a scale-invariant SGWB spectrum  $\Omega(f) = 10^{-5}$ , and including the first 400 harmonics. The blue dashed curves show numerical solutions of the evolution equations (6.22), while the pale red curves show the naive solution  $t \times \dot{C}_{ij}$ , which is exact only when  $\dot{C}_{ij} = \text{constant}$ .

(with no summation over  $i$ ). Unlike the equations for the first two moments, the rhs of Eq. (6.26) is initially zero—non-Gaussianity is only sourced once the covariance has had a chance to grow. We also see that non-Gaussianity can only be sourced at leading order for orbital elements which have a nonzero deterministic drift  $V_i$ ; the inclination  $I$  and longitude of ascending node  $\Omega$  remain Gaussian at leading order. This justifies our assumption that the distribution is Gaussian on observational timescales. However, studies of, e.g., the properties of old binary populations will require us to drop this assumption and solve for the full DF. The results of Sec. V are an important first step in studying cases like these.

## VII. OBSERVATIONAL APPLICATIONS

In this section we describe how to estimate the sensitivity of a given binary to a generic SGWB spectrum using two different high-precision probes of binary dynamics: timing of millisecond pulsars, and laser ranging experiments. This will allow us to compute forecasts for the upper limits that each observational probe will be able to place on various SGWB spectra.

### A. Distribution of the observed orbital elements

Consider a set of observations of a binary system, which are broken up into intervals much longer than the orbital period, but much shorter than the orbital diffusion timescale. These observations give us a discrete series of measurements of the orbital elements  $\mathbf{X}(t)$ , one for each interval, with the intervals being labeled by  $t$ . (It is necessary to split up the data into intervals, since each high-precision measurement of the orbital elements requires a large number of individual data points.) We assume these measurements are made with Gaussian noise and zero bias. The measured values  $\hat{\mathbf{X}}(t)$  are thus drawn from a multivariate Gaussian distribution centered on the true values  $\mathbf{X}(t)$ , with log-likelihood  $\mathcal{L} = \ln p$  given by

$$-2\mathcal{L}(\hat{\mathbf{X}}|\mathbf{X}) = \sum_t \ln \det 2\pi\mathbf{M} + (\hat{\mathbf{X}} - \mathbf{X})^\top \mathbf{M}^{-1} (\hat{\mathbf{X}} - \mathbf{X}), \quad (7.1)$$

where the measurement noise is described by the covariance matrix  $\mathbf{M}$ . The form of the covariance matrix depends

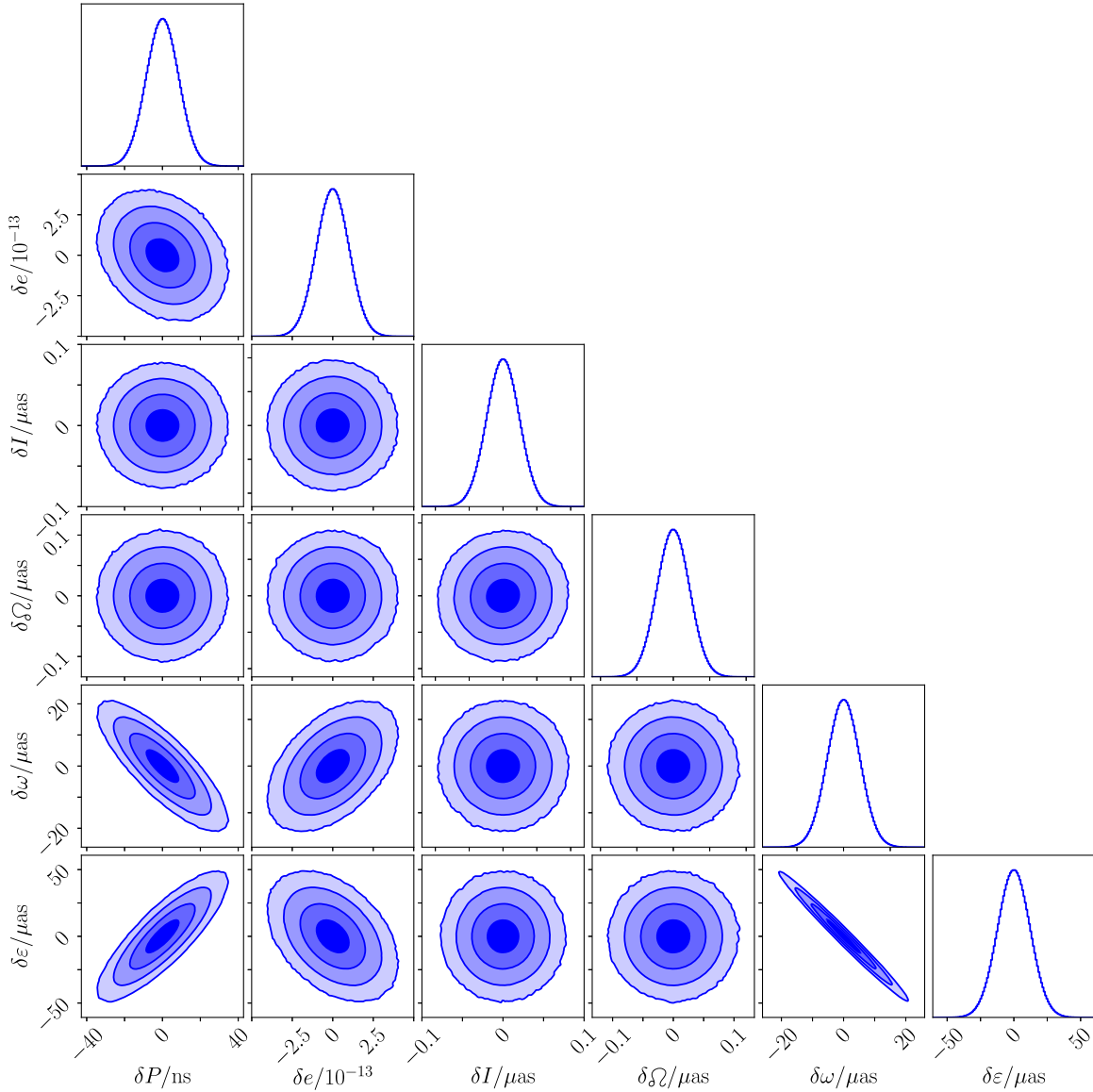


FIG. 10. Corner plot showing the distribution of the orbital elements of B1913 + 16 at the end of the 1 kyr numerical integration shown in Fig. 9.

on how sensitive the timing residuals are to each orbital element—we discuss this further below.

The true orbital elements  $\mathbf{X}(t)$  are themselves random due to the uncertainty caused by SGWB resonance. As discussed in Sec. VI, we can usually approximate the distribution of the elements as Gaussian (at least on observational timescales), so that the log-likelihood reads

$$-2\mathcal{L}(\mathbf{X}|\Omega) = \sum_t \ln \det 2\pi\mathbf{C} + (\mathbf{X} - \bar{\mathbf{X}})^\top \mathbf{C}^{-1} (\mathbf{X} - \bar{\mathbf{X}}), \quad (7.2)$$

where the mean values  $\bar{\mathbf{X}}(t)$  and covariance matrix  $\mathbf{C}(t)$  both depend on the SGWB spectrum  $\Omega(f)$ , and are computed as described in Sec. VI.

We can marginalize over the unknown “true” elements  $\mathbf{X}(t)$  to obtain a likelihood function for the measured elements  $\hat{\mathbf{X}}(t)$  for a given SGWB spectrum,

$$\begin{aligned} -2\mathcal{L}(\hat{\mathbf{X}}|\Omega) &\equiv -2 \ln \left[ \int d\mathbf{X} p(\hat{\mathbf{X}}|\mathbf{X}) p(\mathbf{X}|\Omega) \right] \\ &= \sum_t \ln \det 2\pi\mathbf{N} + (\hat{\mathbf{X}} - \bar{\mathbf{X}})^\top \mathbf{N}^{-1} (\hat{\mathbf{X}} - \bar{\mathbf{X}}), \end{aligned} \quad (7.3)$$

where

$$\mathbf{N} \equiv \mathbf{M} + \mathbf{C} \quad (7.4)$$

is the combined covariance matrix, incorporating the measurement uncertainty as well as the stochasticity of the orbital elements.

### B. Likelihood-ratio test

Given a set of measured orbital elements  $\hat{\mathbf{X}}$ , we can phrase the SGWB detection problem as a hypothesis test, where

- (i)  $\mathcal{H}_0$  is the *null hypothesis*, that there is no SGWB present,  $\Omega = 0$ .
- (ii)  $\mathcal{H}_\Omega$  is the *alternative hypothesis*, that there is a SGWB present,  $\Omega \neq 0$ .

The simplest version of this problem is when we are searching for a SGWB with a fixed spectral shape (e.g., a power law), in which case the only unknown is a single parameter setting the overall amplitude, which we denote  $\Omega$ . (Concretely, this parameter  $\Omega$  should then refer to the amplitude of the SGWB at some fixed reference frequency.)

A very natural way of carrying out such a hypothesis test is by using the log-likelihood-ratio statistic

$$\Lambda(\hat{\mathbf{X}}) \equiv 2 \max_{\Omega > 0} [\mathcal{L}(\hat{\mathbf{X}}|\Omega) - \mathcal{L}(\hat{\mathbf{X}}|0)], \quad (7.5)$$

i.e., we compare the likelihood (7.3) in the case of the null hypothesis  $\Omega = 0$  with the maximum value of the likelihood as a function of  $\Omega$  in the case of the alternative hypothesis. Since the maximum value of the likelihood over the entire range of values of  $\Omega$  is always greater than or equal to its value at  $\Omega = 0$ , we see that  $\Lambda \geq 0$ .

Using Eq. (7.3), along with the fact that  $\mathbf{N} = \mathbf{M}$  when  $\Omega = 0$ , we find

$$\begin{aligned} \Lambda(\hat{\mathbf{X}}) = \max_{\Omega > 0} \sum_t (\hat{\mathbf{X}} - \bar{\mathbf{X}}_0)^\top \mathbf{M}^{-1} (\hat{\mathbf{X}} - \bar{\mathbf{X}}_0) \\ - (\hat{\mathbf{X}} - \bar{\mathbf{X}})^\top \mathbf{N}^{-1} (\hat{\mathbf{X}} - \bar{\mathbf{X}}) - \ln \det(\mathbf{I} + \mathbf{M}^{-1} \mathbf{C}), \end{aligned} \quad (7.6)$$

where  $\bar{\mathbf{X}}_0$  is the value of  $\bar{\mathbf{X}}$  when  $\Omega = 0$ , and where we have used

$$\frac{\det \mathbf{N}}{\det \mathbf{M}} = \det \mathbf{M}^{-1} \mathbf{N} = \det(\mathbf{I} + \mathbf{M}^{-1} \mathbf{C}). \quad (7.7)$$

Given a set of measurements  $\hat{\mathbf{X}}(t)$ , we can therefore compute the likelihood ratio statistic using Eq. (7.6) by solving the FPE to find  $\bar{\mathbf{X}}(t)$  and  $\mathbf{C}(t)$  for a large number of possible values of  $\Omega$  in order to maximize the likelihood. If the resulting value of  $\Lambda$  is large enough, then we reject the null hypothesis and claim a detection of the SGWB.

An obvious question is: What value of  $\Lambda$  is “large enough?” To define the detection threshold, we need to know the distribution of  $\Lambda$  in the case where  $\mathcal{H}_0$  is true. Comparing the observed value of  $\Lambda$  with this null

distribution then allows us to directly infer the statistical significance of the results. Since  $\Lambda$  is a complicated function of the data  $\hat{\mathbf{X}}$ , it is difficult to find an exact distribution, even though we have fully specified how the data are distributed. However, in the limit where we have a large number of measurements (i.e., the data cover a large number of time intervals  $t$ ), Wilks’ theorem tells us that  $\Lambda$  follows a chi-squared distribution [64],

$$\lim_{n_t \rightarrow \infty} \Lambda \sim \chi_1^2, \quad (7.8)$$

where  $n_t$  is the number of time segments. (In this case the chi-squared distribution has one degree of freedom, as  $\mathcal{H}_\Omega$  has one free parameter compared to  $\mathcal{H}_0$ ; we could imagine using a more complicated model for the SGWB with  $k$  parameters, in which case the appropriate distribution would be  $\chi_k^2$ .) In this limit, we can therefore set the threshold for detecting the SGWB at a given confidence level according to the corresponding  $p$ -value of  $\chi_1^2$ ; e.g., a detection with 95% confidence would require  $\Lambda \geq 3.841$ .

### C. Sensitivity forecasts

We can also use the likelihood-ratio statistic discussed above to estimate the sensitivity of future observing campaigns to different SGWB spectra. To do so, we simply compute the expected value of  $\Lambda$  under the SGWB hypothesis  $\mathcal{H}_\Omega$ , and find the smallest value of  $\Omega$  for which this expectation surpasses the detection threshold—this tells us the weakest SGWB that we can expect to detect with a given set of observations.

Let us use  $\langle \cdots \rangle_\Omega$  to denote an expectation value under  $\mathcal{H}_\Omega$ . By definition, we have

$$\begin{aligned} \langle \hat{\mathbf{X}}_i \rangle_\Omega &= \bar{\mathbf{X}}_i, \\ \langle (\hat{\mathbf{X}}_i - \bar{\mathbf{X}}_i)(\hat{\mathbf{X}}_j - \bar{\mathbf{X}}_j) \rangle_\Omega &= \mathbf{N}_{ij}, \end{aligned} \quad (7.9)$$

where the mean vector  $\bar{\mathbf{X}}$  and covariance matrix  $\mathbf{C}$  here are computed using the true underlying value of  $\Omega$ . In principle, this true  $\Omega$  is different from the value  $\hat{\Omega}$  that maximizes the likelihood, and it is the latter which determines the values for  $\bar{\mathbf{X}}$  and  $\mathbf{C}$  that appear in the statistic  $\Lambda$ . However, in the  $n_t \rightarrow \infty$  limit discussed above we have  $\hat{\Omega} \rightarrow \Omega$  (in statistics parlance, the maximum-likelihood estimator is efficient [64]), so the two are interchangeable. It is thus straightforward to show that, in this limit, the expected value of  $\Lambda$  under  $\mathcal{H}_\Omega$  is

$$\langle \Lambda \rangle_\Omega = \sum_t \delta \bar{\mathbf{X}}^\top \mathbf{M}^{-1} \delta \bar{\mathbf{X}} + \text{tr} \mathbf{M}^{-1} \mathbf{C} - \ln \det(\mathbf{I} + \mathbf{M}^{-1} \mathbf{C}). \quad (7.10)$$

We see that the stochastic drift in the orbital elements appears quadratically here, which means that it typically



contributes less to the detectability of the SGWB than diffusion, which enters linearly through the covariance matrix.

Once we have specified the covariance matrix  $\mathbf{M}$  for our observations, we can compute Eq. (7.10) and find the smallest value of  $\Omega$  that we can expect to detect.

#### D. Application to pulsar timing

While the formalism we have developed above is applicable to a very broad class of astrophysical binary systems, one of the main applications is in the case where one member of the binary is a millisecond pulsar (MSP) [65]. Analysis of the timing data from this MSP then allows us to determine its orbit with incredible precision, with uncertainties as small as a few parts per billion in some cases. These precision measurements allow us to search for the very small stochastic perturbations to the orbit described above. (The same principle has been used to set novel constraints on ultralight dark matter, due to its resonant effects on binary pulsar orbits [35–39].) Here we follow the approach of Refs. [66–68] to estimate the covariance matrix  $\mathbf{M}$  for the orbital elements of a binary pulsar.

For each of the  $n_t$  observation intervals used to construct the likelihood ratio statistic above, one must observe some number  $n_{\text{obs}}$  of pulse arrival times, compare these arrival times with a timing formula for the binary, and thereby infer the binary's orbital elements at that time from the timing residuals. In the limit  $n_{\text{obs}} \rightarrow \infty$ , the resulting covariance matrix  $\mathbf{M}$  describing the uncertainty in the orbital elements is given by the inverse of the *Fisher matrix*,

$$\mathbf{M} = \mathbf{F}^{-1}, \quad \mathbf{F}_{ij} \equiv -\langle \partial_i \partial_j \mathcal{L}(t|\mathbf{X}) \rangle, \quad (7.11)$$

where  $\mathcal{L}$  is the log-likelihood describing the distribution of arrival times  $\mathbf{t} = (t_1, t_2, \dots, t_{n_{\text{obs}}})$  as a function of the orbital elements  $\mathbf{X}$ , and the angle brackets here denote an expectation value under that distribution. We make the standard assumptions that the pulse arrival times form a set of uncorrelated Gaussian random variables with constant variance  $\sigma^2$  (i.e., the timing noise is independent of the binary's orbit), and with mean values given by the timing formula,

$$\langle t_a \rangle \equiv \mathcal{T}_a(\mathbf{X}), \quad \text{Cov}[t_a, t_b] \equiv \delta_{ab} \sigma^2, \quad (7.12)$$

so that the log-likelihood is given by

$$-\mathcal{L}(\mathbf{t}|\mathbf{X}) = \sum_{a=1}^{n_{\text{obs}}} \ln 2\pi\sigma^2 + \frac{1}{\sigma^2} (t_a - \mathcal{T}_a)^2. \quad (7.13)$$

The Fisher matrix for a likelihood of this form (as derived in, e.g., Ref. [69]) is

$$\mathbf{F}_{ij} \equiv \frac{1}{\sigma^2} \sum_{a=1}^{n_{\text{obs}}} \frac{\partial \mathcal{T}_a}{\partial X_i} \frac{\partial \mathcal{T}_a}{\partial X_j}. \quad (7.14)$$

To evaluate the Fisher matrix, we therefore need the derivatives of the timing formula  $\mathcal{T}$  with respect to each of the orbital elements. Using the standard Blandford-Teukolsky timing formula [66], we write these as

$$\begin{aligned} \frac{\partial \mathcal{T}}{\partial P} &= \frac{v_P \sin I}{1 + m_1/m_2} \\ &\quad \times \frac{E - e \sin E}{2\pi(1 - e \cos E)} (\sin \omega \sin E - \gamma \cos \omega \cos E), \\ \frac{\partial \mathcal{T}}{\partial e} &= -\frac{P}{2\pi} \frac{v_P \sin I}{1 + m_1/m_2} \\ &\quad \times \left[ \sin \omega (1 + \sin^2 E) + \frac{\cos \omega \sin E}{\gamma} (e - \gamma^2 \cos E) \right], \\ \frac{\partial \mathcal{T}}{\partial I} &= \frac{P}{2\pi} \frac{v_P \cos I}{1 + m_1/m_2} [\sin \omega (\cos E - e) + \gamma \cos \omega \sin E], \\ \frac{\partial \mathcal{T}}{\partial \Omega} &= 0, \\ \frac{\partial \mathcal{T}}{\partial \omega} &= \frac{P}{2\pi} \frac{v_P \sin I}{1 + m_1/m_2} [\cos \omega (\cos E - e) - \gamma \sin \omega \sin E], \\ \frac{\partial \mathcal{T}}{\partial e} &= -\frac{P}{2\pi} \frac{v_P \sin I}{1 + m_1/m_2} \frac{\sin \omega \sin E - \gamma \cos \omega \cos E}{1 - e \cos E}, \end{aligned} \quad (7.15)$$

where the *eccentric anomaly*  $E(t)$  is defined by

$$r = a(1 - e \cos E), \quad \cos \psi = \frac{\cos E - e}{1 - e \cos E}, \quad (7.16)$$

and acts as an alternative to the true anomaly  $\psi$  as a way of parametrizing the orbital ellipse. We see that the timing formula does not depend on the longitude of ascending node,  $\partial \mathcal{T} / \partial \Omega = 0$ , which means that  $\Omega$  cannot be determined with pulsar timing (physically this is because  $\Omega$  corresponds to a rotation around the line-of-sight axis, and thus does not affect the observable motion parallel to the line of sight).

In the limit of many observed pulses,  $n_{\text{obs}} \rightarrow \infty$ , and assuming that these observations are distributed uniformly in time, we can replace the sum in Eq. (7.14) with an integral to obtain

$$\begin{aligned} \mathbf{F}_{ij} &\simeq \frac{n_{\text{obs}}}{T_{\text{obs}} \sigma^2} \int_0^{T_{\text{obs}}} dt \frac{\partial \mathcal{T}}{\partial X_i} \frac{\partial \mathcal{T}}{\partial X_j} \\ &\simeq \frac{n_{\text{obs}}}{T_{\text{obs}} \sigma^2} \frac{P}{2\pi} \int_0^{2\pi T_{\text{obs}}/P} dE (1 - e \cos E) \frac{\partial \mathcal{T}}{\partial X_i} \frac{\partial \mathcal{T}}{\partial X_j}, \end{aligned} \quad (7.17)$$

where  $T_{\text{obs}}$  is the time interval over which the  $n_{\text{obs}}$  pulse measurements are made, and where we have used Kepler's equation,

$$\frac{dE}{dt} = \frac{2\pi/P}{1 - e \cos E}. \quad (7.18)$$

For simplicity, we assume that  $T_{\text{obs}}$  is an integer multiple of the binary period  $P$ , and begins when the binary is at pericenter; if this is not the case, then the following formulas contain additional phase factors which do not affect the order of magnitude of the results.

We find that it is convenient to define

$$\mathbb{F}_{ij} = \frac{n_{\text{obs}} P^2}{16\pi^2 \sigma^2} \left( \frac{v_P \sin I}{1 + m_1/m_2} \right)^2 \tilde{\mathbb{F}}_{ij}, \quad (7.19)$$

pulling out some factors which appear in all of the Fisher matrix elements. Inserting the derivatives (7.15) into Eq. (7.17), we find that this factorized form of the Fisher matrix is given by

$$\begin{aligned} \tilde{\mathbb{F}}_{PP} &= \frac{8e}{P^2} \left( 1 + \frac{1}{4} e^2 \right) + \frac{\cos 2\omega}{P^2} f_1(e) \\ &\quad + \frac{2\pi T_{\text{obs}}}{P^3} \sin 2\omega f_2(e) \\ &\quad + \frac{8\pi^2 T_{\text{obs}}^2}{3P^4} \left[ 1 - \frac{1}{4} \cos 2\omega f_3(e) \right], \\ \tilde{\mathbb{F}}_{Pe} &= \frac{4}{P} \left( 1 + \frac{3}{8} e \right) \left( 1 - \frac{1}{6} e^2 \right) + \frac{\pi T_{\text{obs}} e}{P^2 \gamma} \sin 2\omega \\ &\quad - \frac{8 \cos 2\omega}{3P} \left( 1 + \frac{3}{4} e + \frac{1}{6} e^2 + \frac{1}{16} e^3 \right), \\ \tilde{\mathbb{F}}_{PI} &= \frac{\cot I}{P} \left[ 2e \left( 1 + \frac{1}{4} e \right) + \cos 2\omega \left( 1 - 2e - \frac{3}{2} e^2 \right) \right], \\ \tilde{\mathbb{F}}_{P\omega} &= -\frac{\sin 2\omega}{P} \left( 1 - 2e - \frac{3}{2} e^2 \right) - \frac{2\pi T_{\text{obs}} \gamma}{P^2}, \\ \tilde{\mathbb{F}}_{Pe} &= -\frac{\sin 2\omega}{P} f_2(e) - \frac{2\pi T_{\text{obs}}}{P^2} \left[ 1 - \frac{1}{4} \cos 2\omega f_3(e) \right], \\ \tilde{\mathbb{F}}_{ee} &= \frac{5}{\gamma^2} \left[ 1 - \frac{3}{4} e^2 - \frac{1}{20} e^4 \right. \\ &\quad \left. - \frac{9}{10} \cos 2\omega \left( 1 - \frac{23}{18} e^2 + \frac{1}{18} e^4 \right) \right], \\ \tilde{\mathbb{F}}_{eI} &= 3e \cot I \left[ 1 + \frac{1}{12} e^2 - \frac{11}{6} \cos 2\omega \left( 1 - \frac{1}{22} e^2 \right) \right], \\ \tilde{\mathbb{F}}_{e\omega} &= \frac{11}{2} e \sin 2\omega \left( 1 - \frac{1}{22} e^2 \right), \\ \tilde{\mathbb{F}}_{e\epsilon} &= \frac{e}{\gamma} \sin 2\omega, \end{aligned}$$

$$\tilde{\mathbb{F}}_{II} = 2 \cot^2 I \left( 1 + \frac{3}{2} e^2 - \frac{5}{2} e^2 \cos 2\omega \right),$$

$$\tilde{\mathbb{F}}_{I\omega} = 5e^2 \cot I \sin 2\omega,$$

$$\tilde{\mathbb{F}}_{I\epsilon} = 0,$$

$$\tilde{\mathbb{F}}_{\omega\omega} = 2 + 3e^2 + 5e^2 \cos 2\omega,$$

$$\tilde{\mathbb{F}}_{\omega\epsilon} = 2\gamma,$$

$$\tilde{\mathbb{F}}_{\epsilon\epsilon} = 4 \frac{\sin^2 \omega + \gamma \cos^2 \omega}{1 + \gamma}, \quad (7.20)$$

where the  $f_i(e)$  are functions of eccentricity, which are given to  $\mathcal{O}(e^{14})$  by

$$\begin{aligned} f_1(e) &\simeq 1 + \frac{8}{9} e - \frac{3}{16} e^2 - \frac{448}{225} e^3 - \frac{175}{288} e^4 - \frac{11584}{11025} e^5 \\ &\quad - \frac{2975}{9216} e^6 - \frac{67264}{99225} e^7 - \frac{96733}{460800} e^8 - \frac{5818432}{12006225} e^9 \\ &\quad - \frac{278579}{1843200} e^{10} - \frac{149726912}{405810405} e^{11} - \frac{20910823}{180633600} e^{12}, \\ f_2(e) &\simeq 1 + \frac{4}{3} e + \frac{5}{8} e^2 + \frac{2}{15} e^3 + \frac{1}{48} e^4 + \frac{1}{210} e^5 \\ &\quad - \frac{29}{768} e^6 - \frac{31}{1260} e^7 - \frac{359}{7680} e^8 - \frac{3559}{110880} e^9 \\ &\quad - \frac{469}{10240} e^{10} - \frac{19087}{576576} e^{11} - \frac{6099}{143360} e^{12}, \\ f_3(e) &\simeq e^2 + \frac{1}{2} e^4 + \frac{5}{16} e^6 + \frac{7}{32} e^8 + \frac{21}{128} e^{10} + \frac{33}{256} e^{12}. \end{aligned} \quad (7.21)$$

We have checked numerically that these expressions are accurate to within  $\approx 2\%$  even at large eccentricity,  $e = 0.95$ ; for smaller eccentricities, the accuracy is even better.

For binaries with small eccentricity, we can change to the alternative orbital elements  $(P, \zeta, \kappa, I, \Omega, \xi)$  to find

$$\begin{aligned} \tilde{\mathbb{F}}_{PP} &= \frac{1}{P^2} \left[ -\frac{\zeta^2 - \kappa^2}{\zeta^2 + \kappa^2} + \frac{4\pi T_{\text{obs}}}{P} \frac{\zeta \kappa}{\zeta^2 + \kappa^2} + \frac{8\pi^2 T_{\text{obs}}^2}{3P^2} \right], \\ \tilde{\mathbb{F}}_{P\zeta} &= \frac{4\zeta}{3P} \frac{5\zeta^2 + 6\kappa^2}{(\zeta^2 + \kappa^2)^{3/2}}, \\ \tilde{\mathbb{F}}_{P\kappa} &= \frac{4\kappa}{3P} \frac{\kappa^2}{(\zeta^2 + \kappa^2)^{3/2}}, \\ \tilde{\mathbb{F}}_{PI} &= -\frac{\cot I}{P} \frac{\zeta^2 - \kappa^2}{\zeta^2 + \kappa^2}, \\ \tilde{\mathbb{F}}_{P\xi} &= -\frac{2}{P} \left( \frac{\zeta \kappa}{\zeta^2 + \kappa^2} + \frac{\pi T_{\text{obs}}}{P} \right), \\ \tilde{\mathbb{F}}_{\zeta\zeta} &= \frac{19}{2}, \end{aligned}$$

$$\begin{aligned}
\tilde{F}_{\zeta\kappa} &= \frac{3}{4}\zeta\kappa, \\
\tilde{F}_{\zeta I} &= \frac{1}{2}\zeta \cot I \left( 15 + \frac{2\zeta^2}{\zeta^2 + \kappa^2} \right), \\
\tilde{F}_{\zeta\xi} &= 2\kappa \frac{\zeta^2 - \kappa^2}{\zeta^2 + \kappa^2}, \\
\tilde{F}_{\kappa\kappa} &= \frac{1}{2}, \\
\tilde{F}_{\kappa I} &= -\frac{1}{2}\kappa \cot I \left( 5 - \frac{2\zeta^2}{\zeta^2 + \kappa^2} \right), \\
\tilde{F}_{\kappa\xi} &= \zeta \left( \frac{5}{2} - \frac{\zeta^2}{\zeta^2 + \kappa^2} \right), \\
\tilde{F}_{II} &= 2\cot^2 I, \\
\tilde{F}_{I\xi} &= 0, \\
\tilde{F}_{\xi\xi} &= 2.
\end{aligned} \tag{7.22}$$

Assuming some values for the observation interval  $T_{\text{obs}}$ , number of observations  $n_{\text{obs}}$ , and rms timing noise  $\sigma$ , we can thus use the Fisher matrix elements given above to calculate the expected likelihood ratio (7.10) for a given SGWB spectrum, and therefore compute upper limit forecasts.

### E. Application to laser ranging

Another extremely precise observational probe of binary dynamics is *laser ranging* (LR), in which laser pulses are fired at retroreflectors on bodies orbiting the Earth; typically the Moon (lunar laser ranging, LLR) [70] or artificial satellites (satellite laser ranging, SLR) [71]. By measuring the round-trip times of these pulses, the size of the orbit can be tracked over time with millimeter precision.

We estimate the sensitivity of a generic LR experiment in a very similar way to our treatment of pulsar timing above. Individual range measurements are assumed to be unbiased, with uncorrelated Gaussian noise of variance  $\sigma^2$ . We write their mean value at time  $t$  in terms of the eccentric anomaly,

$$r(t) = a(1 - e \cos E(t)), \tag{7.23}$$

and define the Fisher matrix in the limit of many uniformly-spaced observations analogously to Eq. (7.17),

$$F_{ij} \simeq \frac{n_{\text{obs}}}{T_{\text{obs}}\sigma^2} \frac{P}{2\pi} \int_0^{2\pi T_{\text{obs}}/P} dE (1 - e \cos E) \frac{\partial r}{\partial X_i} \frac{\partial r}{\partial X_j}. \tag{7.24}$$

The necessary partial derivatives are given by

$$\begin{aligned}
\frac{\partial r}{\partial P} &= \frac{2a}{3P} (1 - e \cos E) - \frac{ae \sin E E - e \sin E}{P(1 - e \cos E)}, \\
\frac{\partial r}{\partial e} &= a \frac{e - \cos E}{1 - e \cos E}, \\
\frac{\partial r}{\partial \varepsilon} &= \frac{ae \sin E}{1 - e \cos E},
\end{aligned} \tag{7.25}$$

where we have used Kepler's equation,

$$E - e \sin E = \frac{2\pi t}{P} + \varepsilon. \tag{7.26}$$

Note that we have neglected the three angles describing the orientation of the orbital plane in space,  $(I, \Omega, \omega)$ , as these are not directly measured by the round-trip times of the laser pulses.

Inserting Eq. (7.25) into Eq. (7.24), we obtain the Fisher matrix,

$$\begin{aligned}
F_{PP} &\simeq \frac{4n_{\text{obs}}a^2}{9P^2\sigma^2} \left[ \frac{3\pi^2 T_{\text{obs}}^2}{2P^2} \left( e^2 + \frac{1}{4}e^4 \right) \right. \\
&\quad \left. + 1 + 3e + \frac{27}{16}e^2 + 4e^3 + \frac{315}{256}e^4 \right], \\
F_{Pe} &\simeq \frac{3n_{\text{obs}}a^2}{4P\sigma^2} \left( e + \frac{8}{9}e^2 + \frac{5}{8}e^3 + \frac{8}{45}e^4 \right), \\
F_{P\varepsilon} &\simeq -\frac{\pi n_{\text{obs}} T_{\text{obs}} a^2}{2P^2\sigma^2} \left( e^2 + \frac{1}{4}e^4 \right), \\
F_{ee} &= \frac{n_{\text{obs}}a^2}{\sigma^2} \left( 2 - \gamma - \frac{1 - \gamma}{e^2} \right), \\
F_{e\varepsilon} &= \frac{n_{\text{obs}}a^2}{\sigma^2} (1 - \gamma).
\end{aligned} \tag{7.27}$$

We have neglected  $\mathcal{O}(e^6)$  terms for the first three entries here, since the Moon and the artificial satellites that are typically tracked with SLR have small eccentricities (e.g., the lunar eccentricity is  $e_\zeta \approx 0.055$ ).

## VIII. SUMMARY AND OUTLOOK

In this paper we have developed a powerful new formalism for calculating the statistical evolution of binary systems coupled to the SGWB, deriving a secular FPE which captures the full probability distribution of all six orbital elements on timescales much longer than the binary period. The KM coefficients describing this FPE, given in Eqs. (4.1) and (4.2), and illustrated in Figs. 3–6, encapsulate the rich dynamical structure that arises from the interactions between tensorial GW perturbations and elliptical orbits.

The full FPE is a six-dimensional nonlinear PDE, and it is therefore challenging to find exact solutions. Nonetheless, we have extracted some qualitative features of the late-time behavior in Sec. V by fixing the eccentricity to zero. This analysis illustrates one of the key advantages of

our formalism over previous approaches: its ability to capture the full shape of the DF. We find that, while the stochastic drift effect due to the SGWB tends to increase the energy of the binary and counteract its orbital decay through GW emission, diffusion tends to have the opposite effect, so that the net influence of SGWB resonance is to drive the binary toward merger. At the same time, the SGWB perturbations act to erase any memory of the binary’s initial configuration in space, driving the orbit toward an isotropic distribution in the inclination  $I$ , and uniform distributions in the other angular variables. We also find, however, that there is a strong drive toward larger eccentricities in the  $e \rightarrow 0$  limit, so these results with  $e = 0$  must be taken with a pinch of salt.

We have also developed, in Sec. VI, a practical approach for numerically integrating the full FPE with general eccentricity  $e \in (0, 1)$  on observational timescales, taking advantage of the fact that these are typically much shorter than the diffusion timescale. The resulting set of equations (6.22) are illustrated in Figs. 9 and 10, using the binary pulsar B1913 + 16 as an example. Combined with the statistical tools developed in Sec. VII, this allows us to calculate SGWB sensitivity curves for probes of binary dynamics such as pulsar timing and lunar/satellite laser ranging, which we present in a companion paper [46], showing that these can improve upon existing bounds in the  $\mu\text{Hz}$  frequency band by several orders of magnitude.

Our results motivate further work to develop binary resonance into a precision tool for GW astronomy. The most important task will be to develop the necessary data analysis pipelines to conduct GW searches with pulsar-timing and laser-ranging data, in a way that fully utilizes the theoretical developments in this work. However, there also many possible avenues for developing our theoretical formalism further. One important problem is to develop practical numerical integration schemes that go beyond the approach in Sec. VI and capture non-Gaussian features in the DF, thereby taking full advantage of the Fokker-Planck approach (this will be particularly important for studies of populations of binaries [72,73]). It would also be interesting to relax some of our assumptions, for example abandoning the secular-averaging approach and attempting to capture the evolution of binaries within a single orbital period, or perhaps relaxing some of the usual assumptions about the GW strain statistics to develop searches for SGWBs that are non-Gaussian, anisotropic, or have non-standard polarization content. One could also explore the sensitivity of binaries to narrowband sources, using a similar approach to Ref. [39] to consider GW frequencies between the binary’s resonant frequencies. There is also no reason to restrict ourselves to just binaries; in future work, we plan to consider the SGWB-driven evolution of other gravitationally-bound systems such as hierarchical triples or many-body systems such as galaxies and globular clusters. Finally, our work could be extended

even further by considering other stochastic perturbing fields which may exist in the Universe, such as ultralight scalars [74–77].

## ACKNOWLEDGMENTS

A. C. J. thanks Andrew Pontzen for interesting discussions related to this work. We acknowledge the use of NUMPY [78] and SCIPY [79] in producing the numerical solutions in Figs. 9 and 10. Figures 4–9 were produced using MATPLOTLIB [80], while Fig. 10 was produced using CORNER.PY [81]. A. C. J. was supported by King’s College London through a Graduate Teaching Scholarship. D. B. is supported by a “Ayuda Beatriz Galindo Senior” from the Spanish “Ministerio de Universidades,” Grant No. BG20/00228. D. B. acknowledges support from the Fundación Jesus Serra and the Instituto de Astrofísica de Canarias under the Visiting Researcher Programme 2021 agreed between both institutions.

## APPENDIX A: POLARIZATION TENSORS IN THE BINARY’S COORDINATE FRAME

In order to describe the GW polarization tensors  $e_{ij}^A$ , we introduce the orthonormal frame  $(\hat{u}, \hat{v}, \hat{n})$ , where  $\hat{n}$  is the GW propagation direction (see Fig. 2). We want to find the components of these basis vectors in the frame of the binary,  $(\hat{r}, \hat{\theta}, \hat{\ell})$ , as this determines the binary’s response to the GW. First, we transform from the GW frame to the fixed reference frame  $(\hat{x}, \hat{y}, \hat{z})$  by applying the standard rotations with respect to the zenith  $\vartheta$  and azimuth  $\phi$ ,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{R}_\phi \mathbf{R}_\vartheta \begin{pmatrix} u \\ v \\ n \end{pmatrix}, \quad (\text{A1})$$

where

$$\mathbf{R}_\vartheta = \begin{pmatrix} \cos \vartheta & 0 & \sin \vartheta \\ 0 & 1 & 0 \\ -\sin \vartheta & 0 & \cos \vartheta \end{pmatrix},$$

$$\mathbf{R}_\phi = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (\text{A2})$$

This reference frame is transformed to the binary frame  $(\hat{r}, \hat{\theta}, \hat{\ell})$  with three further rotations, which specify the inclination  $I$ , the longitude of ascending node  $\Omega$ , and the argument of the binary in the orbital plane  $\theta = \psi + \omega$  [49],

$$\begin{pmatrix} r \\ \theta \\ \ell \end{pmatrix} = \mathbf{R}_\theta \mathbf{R}_I \mathbf{R}_\Omega \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad (\text{A3})$$

where

$$\begin{aligned} \mathbf{R}_\Omega &= \begin{pmatrix} \cos \Omega & \sin \Omega & 0 \\ -\sin \Omega & \cos \Omega & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ \mathbf{R}_I &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos I & \sin I \\ 0 & -\sin I & \cos I \end{pmatrix}, \\ \mathbf{R}_\theta &= \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \quad (\text{A4})$$

We thus obtain the desired relationship between the binary frame and the GW frame by applying all five rotations,

$$\begin{pmatrix} r \\ \theta \\ \ell \end{pmatrix} = \mathbf{R}_\theta \mathbf{R}_I \mathbf{R}_\Omega \mathbf{R}_\phi \mathbf{R}_\vartheta \begin{pmatrix} u \\ v \\ n \end{pmatrix}. \quad (\text{A5})$$

The various contractions with the polarization tensors are then given by

$$\begin{aligned} e_{ij}^+ \hat{r}^i \hat{r}^j &= -[\cos \varphi \cos I \sin \theta - \sin \varphi \cos \theta]^2 + [\sin \vartheta \sin I \sin \theta - \cos \vartheta (\cos \varphi \cos \theta + \sin \varphi \cos I \sin \theta)]^2, \\ e_{ij}^\times \hat{r}^i \hat{r}^j &= 2(\cos \varphi \cos I \sin \theta - \sin \varphi \cos \theta)[\cos \vartheta (\cos \varphi \cos \theta + \sin \varphi \cos I \sin \theta) - \sin \vartheta \sin I \sin \theta], \\ e_{ij}^+ \hat{r}^i \hat{\theta}^j &= (\sin \varphi \cos \theta - \cos \varphi \cos I \sin \theta)(\cos \varphi \cos I \cos \theta + \sin \varphi \sin \theta) \\ &\quad - [\cos \vartheta \cos \varphi \sin \theta + (\sin \vartheta \sin I - \cos \vartheta \sin \varphi \cos I) \cos \theta] \\ &\quad \times [\cos \vartheta (\cos \varphi \cos \theta + \sin \varphi \cos I \sin \theta) - \sin \vartheta \sin I \sin \theta], \\ e_{ij}^\times \hat{r}^i \hat{\theta}^j &= \cos \vartheta \cos 2\varphi \cos I \cos 2\theta + \sin \vartheta \sin I (\sin \varphi \cos 2\theta - \cos \varphi \cos I \sin 2\theta) \\ &\quad + \frac{1}{2} \cos \vartheta \sin 2\varphi (1 + \cos^2 I) \sin 2\theta, \\ e_{ij}^+ \hat{r}^i \hat{\ell}^j &= \cos \varphi \sin I (\cos \varphi \cos I \sin \theta - \sin \varphi \cos \theta) - (\sin \vartheta \cos I + \cos \vartheta \sin \varphi \sin I) \\ &\quad \times [\cos \vartheta (\cos \varphi \cos \theta + \sin \varphi \cos I \sin \theta) - \sin \vartheta \sin I \sin \theta], \\ e_{ij}^\times \hat{r}^i \hat{\ell}^j &= \sin \vartheta (\sin \varphi \cos I \cos \theta - \cos \varphi \cos 2I \sin \theta) - \cos \vartheta \sin I (\cos 2\varphi \cos \theta + \sin 2\varphi \cos I \sin \theta), \end{aligned} \quad (\text{A6})$$

where we define  $\varphi \equiv \phi - \Omega$ . We recall that  $\vartheta, \phi$  are the spherical coordinates of the incoming plane GW,  $\Omega$  is the longitude of ascending node, and  $\theta \equiv \psi + \omega$  is the orbital argument with respect to the ascending node, with  $\psi$  the true anomaly and  $\omega$  the argument of pericenter; these are all illustrated in Fig. 2.

## APPENDIX B: TRANSFER FUNCTIONS

The Fourier components of the transfer functions are defined in terms of the polarization tensor contractions discussed in Appendix A by

$$\begin{aligned} T_{P,n}^A &= \frac{3P^2\gamma}{4\pi} \left\langle \frac{e \sin \psi}{1 + e \cos \psi} e_{ij}^A \hat{r}_i \hat{r}_j + e_{ij}^A \hat{r}_i \hat{\theta}_j \right\rangle_n, \\ T_{e,n}^A &= \frac{\gamma^2 T_{P,n}^A}{3Pe} - \frac{P\gamma^5}{4\pi e} \left\langle \frac{e_{ij}^A \hat{r}_i \hat{\theta}_j}{(1 + e \cos \psi)^2} \right\rangle_n, \\ T_{I,n}^A &= \frac{P\gamma^3}{4\pi} \left\langle \frac{\cos \theta}{(1 + e \cos \psi)^2} e_{ij}^A \hat{r}_i \hat{\ell}_j \right\rangle_n, \end{aligned}$$

$$\begin{aligned} T_{\Omega,n}^A &= \frac{P\gamma^3}{4\pi \sin I} \left\langle \frac{\sin \theta}{(1 + e \cos \psi)^2} e_{ij}^A \hat{r}_i \hat{\ell}_j \right\rangle_n, \\ T_{\omega,n}^A &= \frac{P\gamma^3}{4\pi e} \left\langle \frac{\sin \psi (2 + e \cos \psi)}{(1 + e \cos \psi)^2} e_{ij}^A \hat{r}_i \hat{\theta}_j - \frac{\cos \psi e_{ij}^A \hat{r}_i \hat{r}_j}{1 + e \cos \psi} \right\rangle_n \\ &\quad - T_{\Omega,n}^A \cos I, \\ T_{\varepsilon,n}^A &= -\frac{P\gamma^4}{2\pi} \left\langle \frac{e_{ij}^A \hat{r}_i \hat{r}_j}{(1 + e \cos \psi)^2} \right\rangle_n - \gamma \cos I T_{\Omega,n}^A - \gamma T_{\omega,n}^A, \end{aligned} \quad (\text{B1})$$

where we have introduced the secular averaging operation

$$\langle \dots \rangle_n \equiv \int_0^P \frac{dt}{P} \exp\left(\frac{2\pi i n t}{P}\right) (\dots) \quad (\text{B2})$$

which extracts the  $n$ th-order Fourier coefficient of a given function of the true anomaly  $\psi(t)$ , holding the orbital elements fixed (as they vary over much longer timescales).

The subscript  $n$  distinguishes this from the ensemble average  $\langle \dots \rangle$ .

The Fourier components arising in the eccentric transfer functions can be expressed in terms of *Hansen coefficients*,  $C_n^{lm}$ ,  $S_n^{lm}$  [82]. These are functions of eccentricity that have been used for centuries in celestial mechanics to describe Keplerian motion. We define them here by

$$\begin{aligned} C_n^{lm}(e) &= \left\langle \frac{\cos m\psi}{(1 + e \cos \psi)^l} \right\rangle_n, \\ S_n^{lm}(e) &= \left\langle \frac{\sin m\psi}{(1 + e \cos \psi)^l} \right\rangle_n, \end{aligned} \quad (\text{B3})$$

with explicit expressions for particular sets of  $(l, m)$  given in Appendix E.

Inserting Eq. (A6) into Eq. (B1) and using Eq. (B3), we can thus write the transfer functions as linear combinations of the Hansen coefficients. These are then the input to computing the KM coefficients.

### APPENDIX C: TRANSFORMING THE REFERENCE FRAME

The polarization tensor contractions in Eq. (A6) are completely general and apply to any choice of reference frame. However, the corresponding expressions for the

$$\begin{aligned} \cos I &= s_{I'} \sin I' \sin \Omega' \hat{x}' \cdot \hat{z} - s_{I'} \sin I' \cos \Omega' \hat{y}' \cdot \hat{z} + \cos I' \hat{z}' \cdot \hat{z}, \\ \sin I \sin \Omega &= s_{I'} \sin I' \sin \Omega' \hat{x}' \cdot \hat{x} - s_{I'} \sin I' \cos \Omega' \hat{y}' \cdot \hat{x} + \cos I' \hat{z}' \cdot \hat{x}, \\ \sin I \sin \omega &= (\cos \Omega' \cos \omega' - \cos I' \sin \Omega' \sin \omega') \hat{x}' \cdot \hat{z} + (\sin \Omega' \cos \omega' + \cos I' \cos \Omega' \sin \omega') \hat{y}' \cdot \hat{z} + \sin I' \sin \omega \hat{z}' \cdot \hat{z}, \end{aligned} \quad (\text{C2})$$

where

$$s_{I'} \equiv \begin{cases} +1 & \text{if } \cos I' > 0, \\ -1 & \text{if } \cos I' < 0. \end{cases} \quad (\text{C3})$$

We are now free to specify the primed frame such that the KM coefficients are easier to compute. One particularly simple choice is to choose  $(\hat{x}', \hat{y}', \hat{z}')$  such that  $I' = \pi/4$  and  $\Omega' = \omega' = 0$ . (It may seem at first that  $I' = 0$  is a simpler choice, as the reference frame then coincides with the

transfer functions are very lengthy, making it prohibitively difficult to calculate the KM coefficients. We circumvent this difficulty by choosing a particular reference frame in which the transfer functions are much simpler, before transforming back to a general reference frame.

Given two reference frames,  $(\hat{x}, \hat{y}, \hat{z})$  and  $(\hat{x}', \hat{y}', \hat{z}')$ , we have two corresponding sets of orbital elements for the binary,  $X = (P, e, I, \Omega, \omega, \varepsilon)$  and  $X' = (P, e, I', \Omega', \omega', \varepsilon)$  (the period, eccentricity, and compensated mean anomaly are the same in both frames, as they do not depend on the orientation of the binary in space). The KM coefficients for the unprimed elements are given in terms of those for the primed ones by (see, e.g., Sec. 4.9 of Ref. [54] for a derivation)

$$\begin{aligned} D_i^{(1)} &= \frac{\partial X_i}{\partial X_{i'}} D_{i'}^{(1)} + \frac{\partial^2 X_i}{\partial X_{i'} \partial X_{j'}} D_{i'j'}^{(2)}, \\ D_{ij}^{(2)} &= \frac{\partial X_i}{\partial X_{i'}} \frac{\partial X_j}{\partial X_{j'}} D_{i'j'}^{(2)}, \end{aligned} \quad (\text{C1})$$

where primed indices run over the primed orbital elements. We therefore require the first and second partial derivatives of the unprimed elements with respect to the primed ones. These can be deduced from the following relations, which are derived from Sec. 2.8 of Ref. [49],

binary's frame. However, there is a coordinate singularity associated with  $I' = 0$  which makes some of the associated coefficients poorly behaved. Taking  $I' = \pi/4$  is much easier, particularly since we then have  $\sin I' = \cos I'$ , which simplifies many of the resulting expressions.)

Having specified the primed frame, we require the first and second derivatives to transform back to the unprimed frame, which is relevant for a general observer. Using Eq. (C2), we find that the nonzero derivatives, evaluated at  $(I', \Omega', \omega') = (\pi/4, 0, 0)$ , are given by

$$\begin{aligned} \frac{\partial I}{\partial I'} &= \cos \omega, & \frac{\partial I}{\partial \Omega'} &= -\frac{\sin \omega}{\sqrt{2}}, & \frac{\partial \Omega}{\partial I'} &= \frac{\sin \omega}{\sin I}, & \frac{\partial \Omega}{\partial \Omega'} &= \frac{\cos \omega}{\sqrt{2} \sin I}, \\ \frac{\partial \omega}{\partial I'} &= -\frac{\sin \omega}{\tan I}, & \frac{\partial \omega}{\partial \Omega'} &= \frac{1}{\sqrt{2}} \left( 1 - \frac{\cos \omega}{\tan I} \right), & \frac{\partial \omega}{\partial \omega'} &= 1, \\ \frac{\partial^2 I}{\partial I'^2} &= \frac{\sin^2 \omega}{\tan I}, & \frac{\partial^2 I}{\partial I' \partial \Omega'} &= \frac{\sin \omega}{\sqrt{2}} \left( \frac{\cos \omega}{\tan I} - 1 \right), & \frac{\partial^2 I}{\partial \Omega'^2} &= \frac{\cos \omega}{2} \left( \frac{\cos \omega}{\tan I} - 1 \right), \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \Omega}{\partial I^2} &= -\frac{\sin 2\omega}{\sin I \tan I}, & \frac{\partial^2 \Omega}{\partial I' \partial \Omega'} &= \frac{\cos \omega - \cos 2\omega \cot I}{\sqrt{2} \sin I}, & \frac{\partial^2 \Omega}{\partial \Omega'^2} &= \frac{\sin \omega}{\sin I} \left( \frac{\cos \omega}{\tan I} - \frac{1}{2} \right), \\
\frac{\partial^2 \omega}{\partial I^2} &= \sin 2\omega \frac{2 - \sin^2 I}{2 \sin^2 I}, & \frac{\partial^2 \omega}{\partial I' \partial \Omega'} &= \frac{1}{2\sqrt{2}} \left( \cos 2\omega \frac{2 - \sin^2 I}{\sin^2 I} - \frac{2 \cos \omega}{\tan I} - 1 \right), & \frac{\partial^2 \omega}{\partial \Omega'^2} &= \frac{\sin \omega}{\tan I} \left( \frac{1}{2} - \cos \omega \frac{2 - \sin^2 I}{\sin 2I} \right).
\end{aligned} \tag{C4}$$

It is straightforward to confirm that these are well-behaved and take on the appropriate values when  $(I, \Omega, \omega) \rightarrow (I', \Omega', \omega')$ . We thus find that the unprimed drift coefficients are given in terms of those in the primed frame by

$$\begin{aligned}
D_I^{(1)} &= \cos \omega \left[ D_{I'}^{(1)} + \left( \frac{\cos \omega}{\tan I} - 1 \right) \frac{D_{\Omega'\Omega'}^{(2)}}{2} \right] + \frac{\sin^2 \omega}{\tan I} D_{I'I'}^{(2)}, \\
D_\Omega^{(1)} &= \frac{1}{\sin I} \left[ \sin \omega D_{I'}^{(1)} - \frac{\sin 2\omega}{\tan I} D_{I'I'}^{(2)} + \sin \omega \left( \frac{\cos \omega}{\tan I} - \frac{1}{2} \right) D_{\Omega'\Omega'}^{(2)} \right], \\
D_\omega^{(1)} &= \frac{\sin \omega}{\sin I} \left[ -\cos I D_{I'}^{(1)} + \cos \omega \frac{2 - \sin^2 I}{\sin I} D_{I'I'}^{(2)} + \left( \cos I - \cos \omega \frac{2 - \sin^2 I}{\sin I} \right) \frac{D_{\Omega'\Omega'}^{(2)}}{2} \right],
\end{aligned} \tag{C5}$$

with the diffusion coefficients given by

$$\begin{aligned}
D_{XI}^{(2)} &= D_{X\Omega}^{(2)} = D_{X\omega}^{(2)} = D_{Xe}^{(2)} = D_{Ie}^{(2)} = D_{\Omega e}^{(2)} = 0, \\
D_{II}^{(2)} &= \cos^2 \omega D_{I'I'}^{(2)} + \frac{\sin^2 \omega}{2} D_{\Omega'\Omega'}^{(2)}, \\
D_{I\Omega}^{(2)} &= \frac{\sin 2\omega}{2 \sin I} \left( D_{I'I'}^{(2)} - \frac{1}{2} D_{\Omega'\Omega'}^{(2)} \right), \\
D_{I\omega}^{(2)} &= \sin \omega \left[ -\frac{\cos \omega}{\tan I} D_{I'I'}^{(2)} + \frac{1}{2} \left( \frac{\cos \omega}{\tan I} - 1 \right) D_{\Omega'\Omega'}^{(2)} - \frac{1}{\sqrt{2}} D_{\Omega'\omega'}^{(2)} \right], \\
D_{\Omega\Omega}^{(2)} &= \frac{1}{\sin^2 I} \left( \sin^2 \omega D_{I'I'}^{(2)} + \frac{\cos^2 \omega}{2} D_{\Omega'\Omega'}^{(2)} \right), \\
D_{\Omega\omega}^{(2)} &= \frac{1}{\sin I} \left[ -\frac{\sin^2 \omega}{\tan I} D_{I'I'}^{(2)} + \frac{\cos \omega}{2} \left( 1 - \frac{\cos \omega}{\tan I} \right) D_{\Omega'\Omega'}^{(2)} + \frac{\cos \omega}{\sqrt{2}} D_{\Omega'\omega'}^{(2)} \right], \\
D_{\omega\omega}^{(2)} &= \frac{\sin^2 \omega}{\tan^2 I} D_{I'I'}^{(2)} + \frac{1}{2} \left( 1 - \frac{\cos \omega}{\tan I} \right)^2 D_{\Omega'\Omega'}^{(2)} + \sqrt{2} \left( 1 - \frac{\cos \omega}{\tan I} \right) D_{\Omega'\omega'}^{(2)} + D_{\omega'\omega'}^{(2)}, \\
D_{\omega e}^{(2)} &= D_{\omega' e}^{(2)},
\end{aligned} \tag{C6}$$

where  $X$  here stands for  $P$  or  $e$ .

#### APPENDIX D: KRAMERS-MOYAL COEFFICIENTS IN THE PRIMED FRAME

Here we give the KM coefficients evaluated in the primed coordinate frame  $(I', \Omega', \omega') = (\frac{\pi}{4}, 0, 0)$ . To calculate these, we first evaluate the polarization tensor contractions from Appendix A in the primed frame, and insert these into Eq. (B1) to find the appropriate GW transfer functions, expressing the secular averages in terms of Hansen coefficients that are listed below in Appendix E. These secular transfer functions are then inserted into Eq. (3.33), integrating over the GW propagation direction  $\hat{n} = (\vartheta, \phi)$  to obtain the KM coefficients. The resulting expressions for the diffusion matrix are

$$\begin{aligned}
D_{PP}^{(2)} &= \frac{27P^3\gamma^2}{20} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ \left| E_n^{02} + \frac{e}{2}(E_n^{11} - E_n^{13}) \right|^2 - \frac{(eS_n^{11})^2}{3} \right], \\
D_{Pe}^{(2)} &= \frac{\gamma^2 D_{PP}^{(2)}}{3Pe} - \frac{9P^2\gamma^6}{40} \sum_{n=1}^{\infty} nH_0^2\Omega_n E_n^{22} \left( \frac{2}{e} E_n^{02} + E_n^{11} - E_n^{13} \right)^*, \\
D_{ee}^{(2)} &= \frac{3P\gamma^6}{20e^2} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ \left| E_n^{02} + \frac{e}{2}(E_n^{11} - E_n^{13}) - \gamma^2 E_n^{22} \right|^2 - \frac{(eS_n^{11})^2}{3} \right], \\
D_{I'I'}^{(2)} &= \frac{3P\gamma^6}{80} \sum_{n=1}^{\infty} nH_0^2\Omega_n |E_n^{20} + E_n^{22}|^2, \\
D_{\Omega'\Omega'}^{(2)} &= \frac{3P\gamma^6}{40} \sum_{n=1}^{\infty} nH_0^2\Omega_n |E_n^{20} - E_n^{22}|^2, \\
D_{\Omega'\omega'}^{(2)} &= -\frac{1}{\sqrt{2}} D_{\Omega'\Omega'}^{(2)}, \\
D_{\omega'\omega'}^{(2)} &= \frac{1}{2} D_{\Omega'\Omega'}^{(2)} + \frac{3P\gamma^6}{80e^2} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ \left| E_n^{11} + E_n^{13} + 2E_n^{21} - 2E_n^{23} + \frac{e}{2}(E_n^{20} - E_n^{24}) \right|^2 + \frac{4}{3}(C_n^{11})^2 \right], \\
D_{\omega'\epsilon}^{(2)} &= -\frac{3P\gamma^7}{80e^2} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ \left| E_n^{11} + E_n^{13} + 2(E_n^{21} - E_n^{23}) + \frac{e}{2}(E_n^{20} - 4E_n^{22} - E_n^{24}) \right|^2 + \frac{4}{3}C_n^{11}(C_n^{11} - 2eC_n^{20}) - 4e^2|E_n^{22}|^2 \right], \\
D_{\epsilon\epsilon}^{(2)} &= \frac{3P\gamma^8}{80e^2} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ \left| E_n^{11} + E_n^{13} + 2(E_n^{21} - E_n^{23}) + \frac{e}{2}(E_n^{20} - 8E_n^{22} - E_n^{24}) \right|^2 + \frac{4}{3}(C_n^{11} - 2eC_n^{20})^2 \right], \\
D_{P'I'}^{(2)} &= D_{P\Omega'}^{(2)} = D_{P\omega'}^{(2)} = D_{P\epsilon}^{(2)} = D_{eI'}^{(2)} = D_{e\Omega'}^{(2)} = D_{e\omega'}^{(2)} = D_{e\epsilon}^{(2)} = D_{I'\Omega'}^{(2)} = D_{I'\omega'}^{(2)} = D_{I'\epsilon}^{(2)} = D_{\Omega'\epsilon}^{(2)} = 0, \tag{D1}
\end{aligned}$$

while the drift vector is given by

$$\begin{aligned}
D_P^{(1)} &= V_P + \frac{9P^2\gamma^2}{80} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ 2E_n^{02}(10eE_n^{02} + (3 + 7e^2)E_n^{11} + (1 - 11e^2)E_n^{13} - 2e\gamma^2 E_n^{22})^* \right. \\
&\quad + E_n^{11}(e(3 + 2e^2)E_n^{11} - 2e(1 + 4e^2)E_n^{13} - 2\gamma^2 E_n^{22})^* - e(1 - 6e^2)|E_n^{13}|^2 + 2\gamma^2 E_n^{13}(E_n^{22})^* \\
&\quad + 8\gamma^2 \left( E_n^{02} + \frac{e}{2}(E_n^{11} - E_n^{13}) \right) \left( E_n^{21} - E_n^{23} + \frac{e}{4}(E_n^{20} + 6E_n^{22} - E_n^{24}) \right)^* \\
&\quad \left. + 4\gamma^2 \left( E_n^{02} + \frac{e}{2}(E_n^{11} - E_n^{13}) - \gamma^2 E_n^{22} \right) \left( E_n^{02'} + \frac{e}{2}(E_n^{11} - E_n^{13})' \right)^* - \frac{4e}{3} S_n^{11}((1 + 4e^2)S_n^{11} + e\gamma^2 S_n^{11'}) \right], \\
D_e^{(1)} &= V_e + \frac{3P\gamma^6}{20e^2} \sum_{n=1}^{\infty} nH_0^2\Omega_n \left[ \frac{1}{e} \left| E_n^{02} + \frac{e}{2}(E_n^{11} - E_n^{13}) \right|^2 + \frac{\gamma^4}{e} (S_n^{22})^2 - \frac{e^2}{3} S_n^{11'} S_n^{11} + \left( E_n^{02} + \frac{e}{2}(E_n^{11} - E_n^{13}) - \gamma^2 E_n^{22} \right) \right. \\
&\quad \left. \times \left( -\frac{2}{e} E_n^{02} + \frac{1}{2}(E_n^{11} + 3E_n^{13}) + 2(E_n^{21} - E_n^{23}) + E_n^{20'} - \gamma^2 E_n^{22'} + \frac{e}{2}(E_n^{20} + 12E_n^{22} - E_n^{24} + (E_n^{11} - E_n^{13})') \right)^* \right], \\
D_{I'}^{(1)} &= \frac{1}{2} D_{\Omega'\Omega'}^{(2)}, \\
D_{\Omega'}^{(1)} &= 0, \\
D_{\omega'}^{(1)} &= V_{\omega'}, \\
D_{\epsilon}^{(1)} &= V_{\epsilon}, \tag{D2}
\end{aligned}$$

where primes on the Hansen coefficients denote derivatives with respect to eccentricity. These KM coefficients can be transformed back to a general reference frame using Eqs. (C5) and (C6), resulting in the expressions given in Sec. IV.



## APPENDIX E: HANSEN COEFFICIENTS

Using various formulas given in Ref. [82], we write the Hansen coefficients as

$$\begin{aligned}
C_n^{lm} &= \frac{{}_2F_1(-l-m-1, -l+m-1; 1; \beta^2)}{2\gamma^{2l}(1+\beta^2)^{l+1}} [J_{n-m}(ne) + J_{n+m}(ne)] \\
&+ \sum_{k=1}^{\infty} \left\{ \frac{(-l+m-1)_k \beta^k}{2k! \gamma^{2l} (1+\beta^2)^{l+1}} {}_2F_1(-l-m-1, -l+m-1+k; 1+k; \beta^2) [J_{n-m-k}(ne) + J_{n+m+k}(ne)] \right. \\
&+ \left. \frac{(-l-m-1)_k \beta^k}{2k! \gamma^{2l} (1+\beta^2)^{l+1}} {}_2F_1(-l-m-1+k, -l+m-1; 1+k; \beta^2) [J_{n-m+k}(ne) + J_{n+m-k}(ne)] \right\}, \\
S_n^{lm} &= \frac{i {}_2F_1(-l-m-1, -l+m-1; 1; \beta^2)}{2\gamma^{2l}(1+\beta^2)^{l+1}} [J_{n-m}(ne) - J_{n+m}(ne)] \\
&+ \sum_{k=1}^{\infty} \left\{ \frac{i(-l+m-1)_k \beta^k}{2k! \gamma^{2l} (1+\beta^2)^{l+1}} {}_2F_1(-l-m-1, -l+m-1+k; 1+k; \beta^2) [J_{n-m-k}(ne) - J_{n+m+k}(ne)] \right. \\
&+ \left. \frac{i(-l-m-1)_k \beta^k}{2k! \gamma^{2l} (1+\beta^2)^{l+1}} {}_2F_1(-l-m-1+k, -l+m-1; 1+k; \beta^2) [J_{n-m+k}(ne) - J_{n+m-k}(ne)] \right\}, \quad (\text{E1})
\end{aligned}$$

where we define the expansion variable

$$\beta \equiv \frac{e}{1+\gamma}, \quad (\text{E2})$$

and where  $(\dots)_k$  is a rising Pochhammer symbol, defined by

$$(n)_k \equiv n(n+1)(n+2)\cdots(n+k-1), \quad (\text{E3})$$

while  ${}_2F_1(a, b; c; z)$  is a hypergeometric function, and  $J_n(z)$  is a Bessel function of the first kind. We see that in the circular case  $e = 0$  we have  $\beta = 0$ ,  ${}_2F_1(a, b; c; 0) = 1$ , and  $J_n(0) = \delta_{n,0}$ , and these expressions in Eq. (E1) simplify to

$$C_n^{lm} = \frac{1}{2}(\delta_{n,m} + \delta_{n,-m}), \quad S_n^{lm} = \frac{i}{2}(\delta_{n,m} - \delta_{n,-m}), \quad (\text{E4})$$

which can be confirmed by directly integrating Eq. (B3). Note also that for general eccentricity  $e \in (0, 1)$ , from the definition of the Pochhammer symbol Eq. (E3), the sums over  $k$  in Eq. (E1) terminate if and only if  $m \leq l + 1$ , otherwise the corresponding Hansen coefficients will have an infinite number of terms.

Using Eq. (E1), we can directly compute all the Hansen coefficients that appear in Eqs. (4.1) and (4.2), obtaining the cosine coefficients

$$\begin{aligned}
C_n^{02} &= -\left[ \frac{1+\gamma - \frac{2}{3}e^2(1+\frac{1}{4}\gamma)}{(1+\gamma)^4} \right] 12\gamma^2 J_n + \sum_{k=0}^{\infty} \frac{\beta^k(1-\beta^2)^3}{2(1+\beta^2)} (J_{n-k-2} + J_{n+k+2}), \\
C_n^{11} &= \frac{1}{2n\gamma^2} (J_{n-1} - J_{n+1}), \\
C_n^{13} &= -\left[ \frac{1+\gamma - \frac{9}{8}e^2(1+\frac{5}{9}\gamma) + \frac{1}{3}e^4}{\beta\gamma^2(1+\gamma)^6} \right] 20e^4 J_n - \left[ \frac{1-\gamma - \frac{9}{4}e^2(1-\frac{7}{9}\gamma) + \frac{3}{4}e^4(1-\frac{5}{3}\gamma)}{e^3\beta\gamma^2} \right] 2(J_{n-1} + J_{n+1}) \\
&\quad - \left[ \frac{1-\gamma - 3e^2(1-\frac{5}{6}\gamma) + 3e^4(1-\frac{5}{8}\gamma) - e^6(1-\frac{15}{32}\gamma)}{e^5\gamma^2} \right] 8(J_{n-2} + J_{n+2}) + \sum_{k=0}^{\infty} \frac{\beta^k(1-\beta^2)^5}{2(1+\beta^2)^2\gamma^2} (J_{n-k-3} + J_{n+k+3}), \\
C_n^{20} &= -\left[ \frac{n^2(1-\frac{1}{2}e^2) - 3n + 2}{n^4 e^2 \gamma^4} \right] 8J_{n-2} + \left( \frac{n-1}{n^3 e \gamma^4} \right) 4J_{n-3},
\end{aligned}$$

$$\begin{aligned}
C_n^{21} &= \left[ \frac{n^3(1 - \frac{7}{4}e^2 + \frac{3}{4}e^4) - n^2(1 - \frac{7}{2}e^2 + e^4) - 4n(1 + \frac{1}{2}e^2) + 4}{n^4 e^3 \gamma^4} \right] 4J_{n-2} - \left[ \frac{n^2(1 - \frac{3}{2}e^2 + \frac{1}{2}e^4) + n(1 + e^2) - 2}{n^3 e^2 \gamma^4} \right] 2J_{n-3}, \\
C_n^{22} &= - \left[ \frac{n^3(1 - \frac{7}{4}e^2 + \frac{3}{4}e^4) - 2n^2(1 - \frac{13}{8}e^2 + \frac{7}{16}e^4) - n(1 + \frac{1}{2}e^2) + 2 - e^2}{n^4 e^4 \gamma^4} \right] 16J_{n-2} \\
&\quad + \left[ \frac{n^2(1 - \frac{3}{2}e^2 + \frac{1}{2}e^4) + \frac{1}{2}ne^2 - 1 + \frac{1}{2}e^2}{n^3 e^3 \gamma^4} \right] 8J_{n-3}, \\
C_n^{23} &= \left[ \frac{n^2 \gamma^6 + 3n(1 - \frac{5}{4}e^2 + \frac{1}{4}e^4) + 2 - \frac{3}{2}e^2}{n^2 e^3 \gamma^4} \right] 4J_n - \left( \frac{1 - \frac{5}{4}e^2 + \frac{1}{4}e^4}{n^2 e^2 \gamma^4} \right) 12J_{n-1}, \\
C_n^{24} &= \frac{8\gamma^3}{(1+\gamma)^4} (J_{n-4} + J_{n+4}) \\
&\quad + \sum_{k=0}^{\infty} \frac{\beta^k (1+\gamma)^3}{16\gamma^4} \left[ (1-\beta^2)^7 (J_{n-k-4} + J_{n+k+4}) + \frac{(-7)_k}{k!} {}_2F_1(1, k-7; k+1; \beta^2) (J_{n+k-4} + J_{n-k+4}) \right], \tag{E5}
\end{aligned}$$

and the sine coefficients,

$$\begin{aligned}
S_n^{02} &= - \left[ \frac{1 - \gamma - 2e^2(1 - \frac{3}{4}\gamma) + e^4}{e^3} \right] 2i(J_{n-1} - J_{n+1}) + \sum_{k=0}^{\infty} \frac{i\beta^k (1-\beta^2)^3}{2(1+\beta^2)} (J_{n-k-2} - J_{n+k+2}), \\
S_n^{11} &= \frac{iJ_n}{ne\gamma}, \\
S_n^{13} &= - \left[ \frac{1 - 3e^2(1 - \frac{5}{6}\gamma) + 3e^4(1 - \frac{5}{8}\gamma) - e^6}{e^4 \gamma^2} \right] 4i(J_{n-1} - J_{n+1}) \\
&\quad - \left[ \frac{1 - \gamma - 3e^2(1 - \frac{5}{6}\gamma) + 3e^4(1 - \frac{5}{8}\gamma) - e^6(1 - \frac{15}{32}\gamma)}{e^5 \gamma^2} \right] 8i(J_{n-2} - J_{n+2}) + \sum_{k=0}^{\infty} \frac{i\beta^k (1-\beta^2)^5}{2(1+\beta^2)^2 \gamma^2} (J_{n-k-3} - J_{n+k+3}), \\
S_n^{21} &= \frac{-i}{n^2 e \gamma^3} [(n\gamma^2 + 2)J_n - 2eJ_{n-1}], \\
S_n^{22} &= \left[ \frac{n^3(1 - \frac{5}{4}e^2 + \frac{1}{4}e^4) - 2n^2(1 - \frac{9}{8}e^2) - n(1 + e^2) + 2}{n^4 e^4 \gamma^3} \right] 16iJ_{n-2} - \left[ \frac{n^3 \gamma^2 + n^2 e^2 - n}{n^4 e^3 \gamma^3} \right] 8iJ_{n-3}, \\
S_n^{23} &= \frac{15e^2}{8\gamma^3} i(J_{n-1} - J_{n+1}) - \frac{3e}{2\gamma^3} i(J_{n-2} - J_{n+2}) + \frac{1 - \frac{1}{4}e^2}{2\gamma^3} i(J_{n-3} - J_{n+3}), \\
S_n^{24} &= \frac{8i\gamma^3}{(1+\gamma)^4} (J_{n-4} - J_{n+4}) \\
&\quad + \sum_{k=0}^{\infty} \frac{\beta^k (1+\gamma)^3}{16\gamma^4} \left[ (1-\beta^2)^7 (J_{n-k-4} - J_{n+k+4}) + \frac{(-7)_k}{k!} {}_2F_1(1, k-7; k+1; \beta^2) (J_{n+k-4} - J_{n-k+4}) \right]. \tag{E6}
\end{aligned}$$

Note that all the Bessel functions appearing in Eqs. (E5) and (E6) have their argument equal to  $ne$ , even if they are not of order  $n$ ; we suppress this argument for brevity.

- 
- [1] J. Aasi *et al.* (LIGO Scientific Collaboration), Advanced LIGO, *Classical Quantum Gravity* **32**, 074001 (2015).  
[2] F. Acernese *et al.* (Virgo Collaboration), Advanced Virgo: A second-generation interferometric gravitational wave detector, *Classical Quantum Gravity* **32**, 024001 (2015).

- [3] T. Akutsu *et al.* (KAGRA Collaboration), KAGRA: 2.5 Generation interferometric gravitational wave detector, *Nat. Astron.* **3**, 35 (2019).  
[4] B.P. Abbott *et al.* (KAGRA, LIGO Scientific, Virgo Collaborations), Prospects for observing and localizing

- gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA, *Living Rev. Relativity* **21**, 3 (2018).
- [5] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run, *Phys. Rev. X* **11**, 021053 (2021).
- [6] M. Kramer and D. J. Champion, The European pulsar timing array and the large European array for pulsars, *Classical Quantum Gravity* **30**, 224009 (2013).
- [7] M. A. McLaughlin, The North American nanohertz observatory for gravitational waves, *Classical Quantum Gravity* **30**, 224008 (2013).
- [8] G. Hobbs, The parkes pulsar timing array, *Classical Quantum Gravity* **30**, 224007 (2013).
- [9] R. N. Manchester, The international pulsar timing array, *Classical Quantum Gravity* **30**, 224010 (2013).
- [10] Z. Arzoumanian *et al.* (NANOGrav Collaboration), The NANOGrav 12.5 yr Data Set: Search for an Isotropic Stochastic Gravitational-wave Background, *Astrophys. J. Lett.* **905**, L34 (2020).
- [11] M. Punturo *et al.*, The Einstein Telescope: A third-generation gravitational wave observatory, *Classical Quantum Gravity* **27**, 194002 (2010).
- [12] G. Janssen, G. Hobbs, M. McLaughlin, C. Bassa, A. T. Deller, M. Kramer, K. Lee, C. Mingarelli, P. Rosado, S. Sanidas, A. Sesana, L. Shao, I. Stairs, B. W. Stapper, and J. Verbiest, Gravitational wave astronomy with the SKA, *Proc. Sci., AASKA14 05* (2015) 037 [arXiv:1501.00127] 10.22323/1.215.0037.
- [13] P. Amaro-Seoane *et al.* (LISA Consortium Collaboration), Laser interferometer space antenna, arXiv:1702.00786.
- [14] L. Badurina *et al.*, AION: An Atom Interferometer Observatory and Network, *J. Cosmol. Astropart. Phys.* **05** (2020) 011.
- [15] M. Abe *et al.*, Matter-wave atomic gradiometer interferometric sensor (MAGIS-100), *Quantum Sci. Technol.* **6**, 044003 (2021).
- [16] M. Kamionkowski, A. Kosowsky, and M. S. Turner, Gravitational radiation from first order phase transitions, *Phys. Rev. D* **49**, 2837 (1994).
- [17] C. Caprini, M. Hindmarsh, S. Huber, T. Konstandin, J. Kozaczuk, G. Nardini, J. M. No, A. Petiteau, P. Schwaller, G. Servant, and D. J. Weir, Science with the space-based interferometer eLISA. II: Gravitational waves from cosmological phase transitions, *J. Cosmol. Astropart. Phys.* **04** (2016) 001.
- [18] C. Caprini, M. Chala, G. C. Dorsch, M. Hindmarsh, S. J. Huber, T. Konstandin, J. Kozaczuk, G. Nardini, J. M. No, K. Rummukainen, P. Schwaller, G. Servant, A. Tranberg, and D. J. Weir, Detecting gravitational waves from cosmological phase transitions with LISA: An update, *J. Cosmol. Astropart. Phys.* **03** (2020) 024.
- [19] B. Bertotti, Is the solar system gravitationally closed?, *Astrophys. J. Lett.* **14**, 51 (1973), <https://ui.adsabs.harvard.edu/abs/1973ApL....14...51B/abstract>.
- [20] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman, San Francisco, 1973).
- [21] V. N. Rudenko, Test bodies under the effect of gravitational radiation, *Sov. Astron.* **19**, 270 (1975), <https://ui.adsabs.harvard.edu/abs/1975AZh....52..444R/abstract>.
- [22] B. Mashhoon, On tidal resonance, *Astrophys. J.* **223**, 285 (1978).
- [23] M. S. Turner, Influence of a weak gravitational wave on a bound system of two point-masses., *Astrophys. J.* **233**, 685 (1979).
- [24] T. Futamase and T. Matsuda, Resonance between primordial gravitational waves and gravitationally bound systems, *Prog. Theor. Phys.* **61**, 86 (1979).
- [25] B. Mashhoon, B. J. Carr, and B. L. Hu, The influence of cosmological gravitational waves on a Newtonian binary system, *Astrophys. J.* **246**, 569 (1981).
- [26] B. Linet, Absorption of gravitational waves by nearly Newtonian systems, *Gen. Relativ. Gravit.* **14**, 479 (1982).
- [27] L. A. Nelson and W. Y. Chau, Orbital perturbations of a gravitationally bound two-body system with the passage of gravitational waves, *Astrophys. J.* **254**, 735 (1982), <https://ui.adsabs.harvard.edu/abs/1982ApJ...254..735N/abstract>.
- [28] C. Chicone, B. Mashhoon, and D. G. Retzliff, Gravitational ionization: Periodic orbits of binary systems perturbed by gravitational radiation, *Ann. Inst. H. Poincaré Phys. Theor.* **64**, 87 (1996), <https://ui.adsabs.harvard.edu/abs/1996AIHS...64..87C/abstract>.
- [29] C. Chicone, B. Mashhoon, and D. G. Retzliff, On the ionization of a Keplerian binary system by periodic gravitational radiation, *J. Math. Phys. (N.Y.)* **37**, 3997 (1996); **38**, 544(E) (1997).
- [30] C. Chicone, B. Mashhoon, and D. G. Retzliff, Gravitational ionization: A chaotic net in the Kepler system, *Classical Quantum Gravity* **14**, 699 (1997).
- [31] C. Chicone, B. Mashhoon, and D. G. Retzliff, Sustained resonance: A binary system perturbed by gravitational radiation, *J. Phys. A* **33**, 513 (2000).
- [32] L. Iorio, Orbital effects of a monochromatic plane gravitational wave with ultra-low frequency incident on a gravitationally bound two-body system, *ScienceOpen Res.* **0**, 13 (2014).
- [33] L. Iorio and M. L. Ruggiero, Perturbations of the orbital elements due to the magnetic-like part of the field of a plane gravitational wave, *Int. J. Mod. Phys. D* **30**, 2150088 (2021).
- [34] L. Hui, S. T. McWilliams, and I.-S. Yang, Binary systems as resonance detectors for gravitational waves, *Phys. Rev. D* **87**, 084009 (2013).
- [35] D. Blas, D. L. Nacir, and S. Sibiryakov, Ultralight Dark Matter Resonates with Binary Pulsars, *Phys. Rev. Lett.* **118**, 261102 (2017).
- [36] D. López Nacir and F. R. Urban, Vector fuzzy dark matter, fifth forces, and binary pulsars, *J. Cosmol. Astropart. Phys.* **10** (2018) 044.
- [37] J. M. Armaleo, D. López Nacir, and F. R. Urban, Binary pulsars as probes for spin-2 ultralight dark matter, *J. Cosmol. Astropart. Phys.* **01** (2020) 053.
- [38] V. Desjacques, E. Grishin, and Y. B. Ginat, Axion oscillations in binary systems: Angle-action surgery, *Astrophys. J.* **901**, 85 (2020).
- [39] D. Blas, D. López Nacir, and S. Sibiryakov, Secular effects of ultralight dark matter on binary pulsars, *Phys. Rev. D* **101**, 063016 (2020).
- [40] B. Allen, The stochastic gravity wave background: Sources and detection, in *Les Houches School of Physics:*

- Astrophysical Sources of Gravitational Radiation, [arXiv:gr-qc/9604033](https://arxiv.org/abs/1906.04033).
- [41] M. Maggiore, Gravitational wave experiments and early universe cosmology, *Phys. Rep.* **331**, 283 (2000).
- [42] T. Regimbau, The astrophysical gravitational wave stochastic background, *Res. Astron. Astrophys.* **11**, 369 (2011).
- [43] J. D. Romano and N. J. Cornish, Detection methods for stochastic gravitational-wave backgrounds: A unified treatment, *Living Rev. Relativity* **20**, 2 (2017).
- [44] N. Christensen, Stochastic gravitational wave backgrounds, *Rep. Prog. Phys.* **82**, 016903 (2019).
- [45] C. Caprini and D. G. Figueroa, Cosmological backgrounds of gravitational waves, *Classical Quantum Gravity* **35**, 163001 (2018).
- [46] D. Blas and A. C. Jenkins, companion Letter, Bridging the  $\mu\text{Hz}$  Gap in the Gravitational-Wave Landscape with Binary Resonance, *Phys. Rev. Lett.* **128**, 101103 (2022).
- [47] D. Brouwer and G. M. Clemence, *Methods of Celestial Mechanics* (Academic Press, New York, 1961).
- [48] J. A. Burns, Elementary derivation of the perturbation equations of celestial mechanics, *Am. J. Phys.* **44**, 944 (1976).
- [49] C. D. Murray and S. F. Dermott, *Solar System Dynamics* (Cambridge University Press, Cambridge, England, 2000).
- [50] E. Poisson and C. M. Will, *Gravity: Newtonian, Post-Newtonian, Relativistic* (Cambridge University Press, Cambridge, England, 2014).
- [51] P. C. Peters and J. Mathews, Gravitational radiation from point masses in a Keplerian orbit, *Phys. Rev.* **131**, 435 (1963).
- [52] C. Lange, F. Camilo, N. Wex, M. Kramer, D. C. Backer, A. G. Lyne, and O. Doroshenko, Precision timing measurements of PSR J1012+5307, *Mon. Not. R. Astron. Soc.* **326**, 274 (2001).
- [53] M. Maggiore, *Gravitational Waves. Vol. 1: Theory and Experiments*, Oxford Master Series in Physics (Oxford University Press, New York, 2007).
- [54] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd ed. (Springer-Verlag, Berlin, 1989).
- [55] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 3rd ed., Springer Series in Synergetics (Springer-Verlag, Berlin, 2004).
- [56] R. Pawula, Approximation of the linear Boltzmann equation by the Fokker-Planck equation, *Phys. Rev.* **162**, 186 (1967).
- [57] M. Rozner, E. Grishin, Y. B. Ginat, A. P. Igoshev, and V. Desjacques, Axion resonances in binary pulsar systems, *J. Cosmol. Astropart. Phys.* **03** (2020) 061.
- [58] R. L. Stratonovich, *Conditional Markov Processes and Their Application to the Theory of Optimal Control*, Modern analytic and computational methods in science and mathematics (Elsevier, New York, 1968).
- [59] K. Itô, Stochastic integral, *Proc. Imp. Acad.* **20**, 519 (1944).
- [60] L. Annulli, L. Bernard, D. Blas, and V. Cardoso, Scattering of scalar, electromagnetic and gravitational waves from binary systems, *Phys. Rev. D* **98**, 084001 (2018).
- [61] R. A. Hulse and J. H. Taylor, Discovery of a pulsar in a binary system, *Astrophys. J. Lett.* **195**, L51 (1975).
- [62] J. M. Weisberg, D. J. Nice, and J. H. Taylor, Timing measurements of the relativistic binary pulsar PSR B1913 + 16, *Astrophys. J.* **722**, 1030 (2010).
- [63] J. M. Weisberg and Y. Huang, Relativistic measurements from timing the binary pulsar PSR B1913 + 16, *Astrophys. J.* **829**, 55 (2016).
- [64] G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed. (Duxbury, Pacific Grove, California, 2002).
- [65] D. R. Lorimer, Binary and millisecond pulsars, *Living Rev. Relativity* **11**, 8 (2008).
- [66] R. Blandford and S. A. Teukolsky, Arrival-time analysis for a pulsar in a binary system, *Astrophys. J.* **205**, 580 (1976).
- [67] R. Epstein, The binary pulsar: Post-Newtonian timing effects, *Astrophys. J.* **216**, 92 (1977).
- [68] T. Damour and N. Deruelle, General relativistic celestial mechanics of binary systems II. The post-Newtonian timing formula, *Ann. Inst. Henri Poincaré Phys. Théor.* **44**, 263 (1986), <https://ui.adsabs.harvard.edu/abs/1986AIHS...44..263D/abstract>.
- [69] M. Tegmark, A. Taylor, and A. Heavens, Karhunen-Loeve eigenvalue problems in cosmology: How should we tackle large data sets?, *Astrophys. J.* **480**, 22 (1997).
- [70] T. W. Murphy, Lunar laser ranging: The millimeter challenge, *Rep. Prog. Phys.* **76**, 076901 (2013).
- [71] I. Ciufolini, A. Paolozzi, E. C. Pavlis, R. Koenig, J. Ries, V. Gurzadyan, R. Matzner, R. Penrose, G. Sindoni, C. Paris, H. Khachatryan, and S. Mirzoyan, A test of general relativity using the LARES and LAGEOS satellites and a GRACE Earth gravity model, *Eur. Phys. J. C* **76**, 120 (2016).
- [72] E. D. Barr, P. C. C. Freire, M. Kramer, D. J. Champion, M. Berezina, C. G. Bassa, A. G. Lyne, and B. W. Stappers, A massive millisecond pulsar in an eccentric binary, *Mon. Not. R. Astron. Soc.* **465**, 1711 (2017).
- [73] C. Y. Hui, K. Wu, Q. Han, A. K. H. Kong, and P. H. T. Tam, On the orbital properties of millisecond pulsar binaries, *Astrophys. J.* **864**, 30 (2018).
- [74] J. W. Foster, N. L. Rodd, and B. R. Safdi, Revealing the dark matter Halo with axion direct detection, *Phys. Rev. D* **97**, 123006 (2018).
- [75] T. Kalaydzhyan and N. Yu, Searching for stochastic background of ultra-light fields with atomic sensors, *Universe* **4**, 99 (2018).
- [76] G. P. Centers, J. W. Blanchard, J. Conrad, N. L. Figueroa, A. Garcon, A. V. Gramolin, D. F. Jackson Kimball, M. Lawson, B. Pelssers, J. A. Smiga, A. O. Sushkov, A. Wickenbrock, D. Budker, and A. Derevianko, Stochastic fluctuations of bosonic dark matter, *Nat. Commun.* **12**, 7321 (2021).
- [77] J. A. Dror, H. Murayama, and N. L. Rodd, Cosmic axion background, *Phys. Rev. D* **103**, 115004 (2021).
- [78] C. R. Harris *et al.*, Array programming with NumPy, *Nature (London)* **585**, 357 (2020).
- [79] P. Virtanen *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in Python, *Nat. Methods* **17**, 261 (2020).
- [80] J. D. Hunter, Matplotlib: A 2d graphics environment, *Comput. Sci. Eng.* **9**, 90 (2007).
- [81] D. Foreman-Mackey, CORNER.PY: Scatterplot matrices in PYTHON, *J. Open Source Software* **1**, 24 (2016).
- [82] V. A. Brumberg, *Analytical Techniques of Celestial Mechanics* (Springer-Verlag, Berlin, 1995).