

Efficient waveforms for asymmetric-mass eccentric equatorial inspirals into rapidly spinning black holes

Christian E. A. Chapman-Bird^{1,2,*}, Lorenzo Speri³, Zachary Nasipak⁴, Ollie Burke^{2,5},
Michael L. Katz⁶, Alessandro Santini⁷, Shubham Kejriwal⁸, Philip Lynch⁷, Josh Mathews⁸,
Hassan Khalvati^{9,10}, Jonathan E. Thompson^{4,11}, Soichiro Isoyama⁸, Scott A. Hughes¹²,
Niels Warburton¹³, Alvin J. K. Chua^{8,14} and Maxime Pigou⁵

¹*Institute for Gravitational Wave Astronomy and School of Physics and Astronomy,
University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom*

²*School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom*

³*European Space Agency (ESA),
European Space Research and Technology Centre (ESTEC),
Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands*

⁴*School of Mathematical Sciences and STAG Research Centre, University of Southampton,
Southampton, United Kingdom, SO17 1BJ*

⁵*Laboratoire des 2 Infinis - Toulouse (L2IT-IN2P3),
Université de Toulouse, CNRS, F-31062 Toulouse Cedex 9, France*

⁶*NASA Marshall Space Flight Center, Huntsville, Alabama 35811, USA*

⁷*Max Planck Institute for Gravitational Physics (Albert Einstein Institute),
Am Mühlenberg 1, 14476 Potsdam, Germany*

⁸*Department of Physics, National University of Singapore, 21 Lower Kent Ridge Rd, Singapore 119077*

⁹*Perimeter Institute for Theoretical Physics, Ontario, N2L 2Y5, Canada*

¹⁰*University of Guelph, Guelph, Ontario N1G 2W1, Canada*

¹¹*Theoretical Astrophysics Group, California Institute of Technology, Pasadena, California 91125, USA*

¹²*Department of Physics and MIT Kavli Institute, MIT, Cambridge, Massachusetts 02139 USA*

¹³*School of Mathematics & Statistics, University College Dublin, Belfield, Dublin 4, Ireland*

¹⁴*Department of Mathematics, National University of Singapore, Singapore 119076*



(Received 12 June 2025; accepted 26 September 2025; published 12 November 2025)

Observations of gravitational-wave signals emitted by compact binary inspirals provide unique insights into their properties, but their analysis requires accurate and efficient waveform models. Intermediate- and extreme-mass-ratio inspirals (I/EMRIs), with mass ratios $q \gtrsim 10^2$, are promising sources for future detectors such as the Laser Interferometer Space Antenna (LISA). Modeling waveforms for these asymmetric-mass binaries is challenging, entailing the tracking of many harmonic modes over thousands to millions of cycles. The FastEMRIWaveforms (FEW) modeling framework addresses this need, leveraging precomputation of mode data and interpolation to rapidly compute adiabatic waveforms for eccentric inspirals into zero-spin black holes. In this work, we extend FEW to model eccentric equatorial inspirals into black holes with spin magnitudes $|a| \leq 0.999$. Our model supports eccentricities $e \leq 0.9$ and semilatus recta $p \leq 200$, enabling the generation of long-duration IMRI waveforms, and produces waveforms in ~ 100 ms with hardware acceleration. Characterizing systematic errors, we estimate that our model attains mismatches of $\sim 10^{-5}$ (for LISA sensitivity) with respect to error-free adiabatic waveforms over the majority of the parameter space. We find that kludge models can introduce errors in signal-to-noise ratios (SNRs) as great as $^{+60\%}_{-40\%}$ and induce marginal biases of up to $\sim 1\sigma$ in parameter estimation. We show that LISA's horizon redshift for I/EMRI signals varies significantly with a , reaching a redshift of 3 (15) for EMRIs (IMRIs) with only minor ($\sim 10\%$) dependence on e for an SNR threshold of 20. For signals with SNR ~ 50 , spin and eccentricity at plunge are measured with uncertainties of $\delta a \sim 10^{-7}$ and $\delta e_f \sim 10^{-5}$. This work advances the state of the art in waveform generation for asymmetric-mass binaries, providing open-source tools for the investigation of I/EMRI astrophysics and data analysis.

DOI: [10.1103/scbp-75pf](https://doi.org/10.1103/scbp-75pf)

*Contact author: c.chapman-bird@bham.ac.uk

I. INTRODUCTION

The advancement of gravitational wave (GW) astronomy has significantly enhanced our understanding of astrophysical phenomena, particularly through the analysis of compact binary mergers identified by the ground-based LIGO-Virgo-KAGRA GW detector network [1–5]. To facilitate these efforts, models for the gravitational waveforms emitted by these coalescences have been developed and refined over time [6–8]. However, there exists a region in the parameter space that has not yet been explored by GW detectors: systems where the masses $m_{1,2}$ of the binary components differ substantially (i.e., $m_1 \gg m_2$) [9]. These asymmetric-mass binaries are not the primary focus of current ground-based detectors, which have identified sources with mass ratios of up to $q = m_1/m_2 \sim 10$ (in the case of GW190814 [10–12]), and even up to $q \sim 26$ in the case of the event GW191219_163120 [1] (although this mass ratio estimate may be unreliable due to the inaccuracy of the waveform models used in this region of parameter space).

Asymmetric-mass binaries are anticipated to be key sources for future observatories. When the smaller body is a intermediate-mass black hole (IMBH) with $m_1 \sim 10^3 M_\odot$, their GWs are in the band of next-generation ground-based detectors such as Cosmic Explorer [13] or the Einstein Telescope [14] and deci-Hertz detectors [15] such as (B-)DECIGO [16–18] and lunar GW detectors [19]. Binaries in which the larger body is a massive black hole, $m_1 \in [10^4, 10^8] M_\odot$, radiate in the milli-Hertz band targeted by detectors such as Laser Interferometer Space Antenna (LISA) [20,21], TianQin and Taiji [22–24].

One of LISA’s primary scientific objectives is to study the properties and environments of black holes in the local Universe with observations of intermediate-mass-ratio inspirals (IMRIs) and extreme-mass-ratio inspirals (EMRIs), characterized by mass ratios of $q \in [10^2, 10^4]$ and $q \in [10^4, 10^6]$, respectively [25]. Binaries with even larger mass ratio are also of interest [26,27]. While measurement prospects for IMRIs are less well explored due to a lack of accurate and efficient waveform models in this regime, the analysis of LISA observations of EMRI signals will enable measurements of binary parameters with subpercent precision [28,29]. Such precision facilitates rigorous tests of general relativity [30–35] and offers insights into the environments surrounding massive black holes [36–40]. This, in turn, enhances our understanding of the mass function of massive black holes (MBHs) [41,42], the dense stellar environments in galactic cores [20] and the gas disks surrounding these black holes [40,43–50], providing insights into the many proposed formation channels of these systems [51–54]. Moreover, EMRI signals are effective probes of cosmic expansion [55–59] and may even enable LISA calibration errors to be constrained [60].

Achieving these scientific goals entails the unbiased analysis of these signals, a procedure reliant upon accurate waveform models. Waveforms from IMRIs and EMRIs are among the most challenging to construct accurately of any compact binary system [61–63]. As the trajectory of the secondary object is (in general) expected to be eccentric and inclined with respect to the primary object’s spin-momentum vector [64] (leading to Lense–Thirring precession of the orbital plane), the inspiral is triperiodic [65], which manifests as a rich structure of tens of thousands of strong harmonic (sideband) modes in the gravitational waveform [66,67] (as shown in Fig. 1). Models must accurately track the GW phase evolution of these modes to subradian precision over the thousands to millions of orbital cycles that the secondary object typically completes in band, which is an infeasible task for the numerical relativity (NR) techniques that underpin existing models of comparable-mass binaries (despite ongoing efforts to address scale disparity in NR simulations [68–72]). Moreover, waveform models must also be highly efficient; computational wall times of less than one second are vital for the analysis of IMRIs and EMRIs to be a feasible prospect, especially in the context of a global LISA analysis framework (e.g., Refs. [73–75]).

The most promising approach that satisfies these stringent modeling constraints is the gravitational self-force (GSF) paradigm [76–82], where the metric of the binary is found by perturbatively solving the Einstein field equations in powers of the small mass ratio around the metric of the primary black hole; we refer readers to Refs. [61,62,83] for reviews. In particular, the modern waveform generation scheme in the GSF paradigm is built upon the multiscale framework [61,84–86], leveraging the quasiperiodic orbital dynamics of asymmetric-mass binaries (see, e.g., Refs. [87–105] for other efforts toward IMRIs and EMRIs waveform generation). An important feature of this framework is that it enables waveform generation to be divided into an expensive “offline” step (in which many numerical data products are computed once and stored) and a fast “online” step in which these data products are interpolated and integrated to build waveforms quickly and accurately. In the case of the inspiral dynamics,¹ at leading order, one can approximate the change of energy, angular momentum, and Carter constant [111] of the binary due to the

¹In this work, we exclusively focus on the inspiral stage of waveform modeling, as the merger-ringdown stage does not significantly contribute to the measurement precision of EMRI parameters. While our waveform model does not incorporate these stages of coalescence, they may become more impactful for IMRIs and their inclusion is worth investigating in this regard. Recent advancements have been made in terms of merger-ringdown modeling in the GSF framework (e.g., Refs. [106–110]), highlighting the potential for the future extension of our model to include this stage; see also our concluding discussion in Sec. VI.

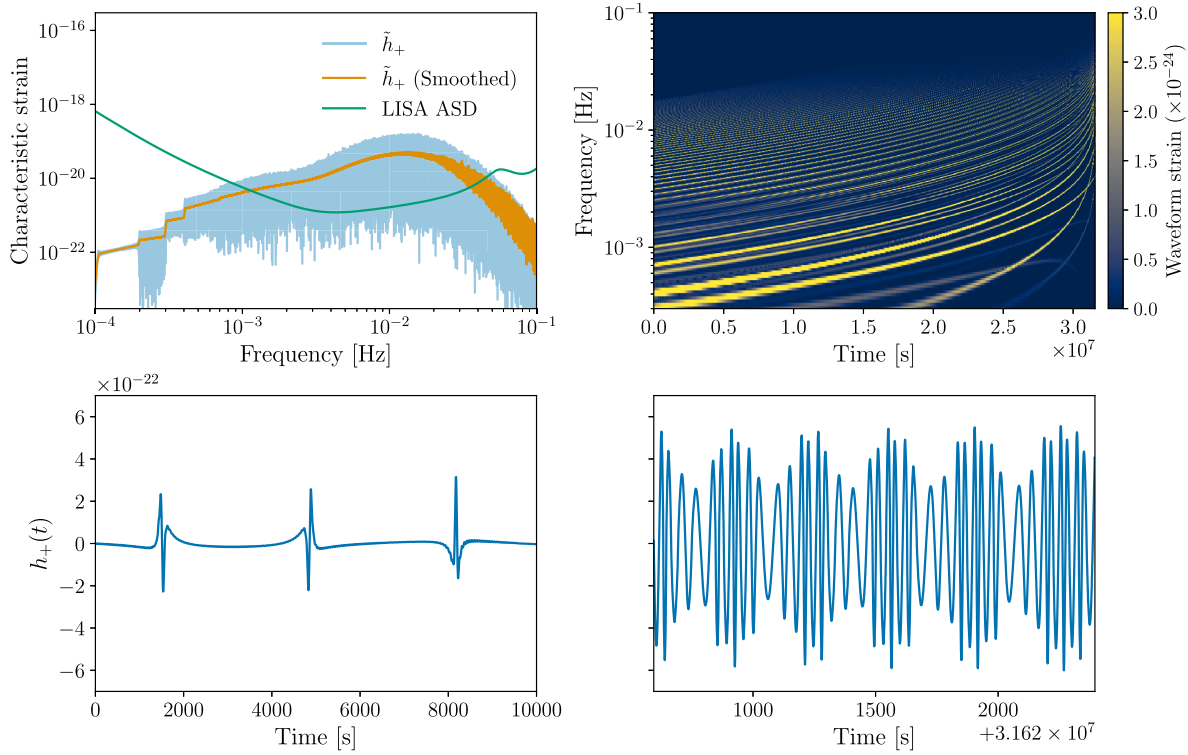


FIG. 1. Waveform for a retrograde eccentric EMRI into a spinning MBH in the frequency (upper left), time-frequency (upper right) and time domain (early and late times in lower left and lower right, respectively). This system has parameters $\{m_1, m_2, a, p_0, e_0, d_L\} = \{10^5 M_\odot, 30 M_\odot, -0.998, 28.3, 0.85, 1 \text{ Gpc}\}$, and plunges after one year. When observed with LISA over this duration, this signal has an SNR of 41. The rich harmonic mode structure of the waveform evolves as the binary circularizes and the trajectory enters the strong-field region of the MBH (see also Sec. III B). Despite the waveform’s complexity and size, it is generated by FEW in less than 100 ms of wall time for an A100 GPU, sufficiently fast enough for full parameter estimation studies to be performed on a timescale of hours.

emission of GWs using so-called “flux-balance formulas” [112–117] in what is known as the adiabatic approximation [66,118–124]. These fluxes have been calculated using both numerical codes [67,125–137], and in analytic post-Newtonian (PN) expansions [138–148]. It has been determined through a multiscale analysis (and subsequently verified with data analysis simulations in Ref. [149]) that one must go beyond leading order adiabatic waveforms and include the first postadiabatic corrections to the dynamics of the binary to achieve the subradian accuracy required for LISA’s scientific goals. This requires knowledge of not only the first order in $\epsilon = 1/q$ GSF [63,150–157] but also the second order GSF [158–160].² When included, GSF waveforms show remarkable agreement with NR even at $q \sim 10$ [164]. These corrections can be included in an efficient way by leveraging averaging techniques [115,120,133], the multi-scale expansion [61,84–86], or near-identity averaging transformations [161,165–169]. This leads to the

²The effects of the secondary object’s spin also enter at this first postadiabatic order [86,161–163].

expensive offline step mentioned above, where one densely tiles the intrinsic EMRI parameter space such that it may be rapidly interpolated during waveform generation. As the GSF data varies smoothly over parameter space, techniques such as cubic spline interpolation can be applied at low computational cost, facilitating rapid modeling of the inspiral dynamics and trajectory evolution. Unlike the dynamics, the amplitude evolution of each waveform mode needs only to be known to leading order for EMRIs (see, e.g., Appendix C of Ref. [149]). These amplitudes also vary smoothly over parameter space and are readily interpolated.

Even once the vast majority of computational effort is shifted into these up-front data-generation procedures, interpolating many thousands of waveform modes at the density of the LISA data stream and summing over them to construct the full waveform is still a computationally expensive and memory-intensive procedure. The FEW project and modeling framework is a pioneering effort in addressing these computational costs. It is a software package capable of rapidly computing LISA-scale, fully relativistic adiabatic waveforms in both the time and frequency domains in subsecond speeds [170–173].

It integrates stand-alone modules to generate EMRI waveforms on both GPUs and central processing units (CPUs), with the GPU version demonstrating a speedup of more than three orders of magnitude compared to its CPU counterpart for eccentric inspirals into nonspinning black holes. This significant acceleration rendered the Bayesian parameter estimation of high-fidelity EMRI signals feasible on timescales of hours/days for the first time; FEW was applied in Ref. [171] to perform ~ 100 such analyses, and has since seen extensive application in the field of EMRI data analysis (e.g., Refs. [35,40,48–50,59,149]).

A significant limitation of the original FEW implementation is the restriction of its fully relativistic waveform generation to zero-spin systems. It is anticipated that IMBHs and MBHs in nature will exhibit a wide variety of spin magnitudes that are closely linked to their formation and evolutionary history (see, e.g., the review in Ref. [174]). A rapidly spinning primary object significantly alters waveform morphology; in the case of prograde systems (where the primary’s spin vector is aligned with the orbital angular momentum vector), the inspiral extends deep into the strong-field region of the central object, enhancing signal detectability, harmonic mode content, and the precision of parameter estimates [28,175]. The development of accurate and rapid waveform models that incorporate such features is therefore vital for both the development of analysis techniques and studies of scientific prospects for realistic EMRI signals.

In this work, we address this need with the release of FEWv2, extending its domain of validity to spinning systems and covering a large portion of the parameter space of Kerr eccentric equatorial inspirals at adiabatic order. Specifically, we support dimensionless primary spin magnitudes of up to 0.999 that exceed the Thorne limit (~ 0.998 [176]) which represents the astrophysical “limit” of Kerr-spin parameters accounting for spin-up (and spin-down) mechanisms arising from accretion.³ Similar to expanding coverage of parameter space in primary spin, we also extend our coverage of semilatus rectum and eccentricity to ~ 200 and 0.9, respectively (cf., ~ 16 and 0.7 in FEW v1). This provides the necessary support for the generation of long-duration IMRI waveforms. We also take this opportunity to thoroughly assess the systematics of waveform generation with FEW (with specific focus on the impact of interpolation errors), ensuring that our model produces waveforms that meet the accuracy requirements imposed by LISA (when neglecting post-adiabatic effects). From the results of this assessment, we conclude that our model is most robust for spins of up to ~ 0.998 and initial eccentricities of up to ~ 0.85 . While our

model does not include the postadiabatic corrections necessary for the unbiased analysis of real signals (see Ref. [149] for a discussion on this point), these additional contributions may be readily folded in as they become available, and it reproduces many features of these signals (such as detectability and parameter estimation precision) sufficiently well for use in studies of IMRI and EMRI science. As our model also accurately produces the rich harmonic mode spectrum of these sources (Fig. 1), it will also serve as an effective test bed for the development of search and identification algorithms for these signals (which is an area of active research [183–189]) and will facilitate the construction of realistic test datasets as part of upcoming LISA Data Challenges [190].

This paper is organized as follows. In Sec. II, we provide an overview of the FEW framework for fast waveform generation. We then describe the extensions to this framework (both in the incorporation of primary spin and in the general improvement of the software) that define our waveform model in Sec. III. To verify that our model is accurate and robust, we perform an extensive validation study of our systematic assumptions in Sec. IV. After assessing the computational cost of our model, we apply it in the examination of multiple facets of EMRI science in Sec. V. Our conclusions and outlook of future work are then presented in Sec. VI.

Some incidental matters are relegated to Appendixes. We explain the new mass convention that FEWv2 adopts in Appendix A. The data grid and coordinates that we use for the model implementation are explained in Appendix B. The source parameter prior probability distribution for the Monte Carlo studies in this work is summarized in Appendix C, and we provide a concise description of our GW data analysis methods (which we apply throughout this work) in Appendix D. In Appendix E, we provide additional validation of our waveform model by comparing it to existing models and data from a specific case. Finally, the marginal posterior distributions corresponding to the analyses performed in Sec. V F are provided in Appendix F.

Access information for repositories containing both the FEW package and the resources required to reproduce the results presented in this work can be found in the Data Availability statement. Throughout this manuscript, we employ geometrized units unless otherwise stated, setting $G = 1 = c$.

II. WAVEFORMS FOR ECCENTRIC EQUATORIAL INSPIRALS

We consider a binary with masses $m_1 \gg m_2$ and spins $\vec{\chi}_1$ and $\vec{\chi}_2$. In our current treatment, we set $\vec{\chi}_2 = \vec{0}$: in Refs. [86,161,191–194] it is demonstrated that effects due to the higher multipole moments of the compact object (CO), including its spin, enter the inspiral dynamics at

³Methods for spinning up MBHs beyond the Thorne limit are discussed in [177–179]; further investigations regarding EMRIs into these near-extremal black holes can be found in [175,180–182].

postadiabatic orders and can therefore be ignored in our adiabatic model. We choose our source frame so that m_1 sits at the origin with its spin $\vec{\chi}_1$ aligned with the z axis. The magnitude of $\vec{\chi}_1$ is then parametrized in terms of the (dimensionless) spin parameter $a = |\vec{\chi}_1|/m_1^2$. Following [66,122], we construct the waveform strain at null infinity $h = h_+ - ih_\times$ by summing over harmonic modes. Each mode is defined by its phase Φ_{mkn} and amplitude $H_{\ell mkn}$, which evolve over the course of an inspiral. The mode indices (ℓ, m, k, n) refer to the multipole of the gravitational radiation and the harmonic of the fundamental frequencies of the orbit (with m, k and n corresponding to azimuthal, polar, and radial motion, respectively). For inspirals restricted to the equatorial plane, which have no spin-induced orbital precession, the modes with $k \neq 0$ vanish due to symmetry such that

$$h(t) = \frac{\mu}{d_L} \sum_{\ell mn} H_{\ell mn}(t, \theta, \phi) e^{-i\Phi_{mn}(t)}, \quad (1)$$

where $\mu = m_1 m_2 / M$ is the reduced mass of the binary, $M = m_1 + m_2$ is its total mass, d_L is the luminosity distance to the source, and (θ, ϕ) are, respectively, the source-frame polar and azimuthal viewing angles of the observer. Our notation for the masses in our waveform model deviates from the majority of previous adiabatic waveform models in the GSF literature, e.g., [48,66,122,124,136,171], which define $\mu = m_2$ and $M = m_1$. Instead, our mass convention is chosen to align with what is commonly used when modeling comparable-mass binaries, e.g., [195–198]. Further discussion of this change in convention is provided in Appendix A.

In the following Section, we review how Eq. (1) is constructed in the FEW framework, highlighting where assumptions made for nonspinning systems must be modified in the extension of this framework to spinning systems.

A. Inspiral trajectories into spinning black holes

In order to compute the amplitude and phase of each mode in Eq. (1), we must first obtain the orbital trajectory of the CO by integrating its equations of motion. Operating in the multiscale framework at adiabatic order [61,84,122,199], we parametrize the inspiral trajectory at each moment in time by three quasi-Keplerian orbital elements— $p(t)$ (semilatus rectum), $e(t)$ (eccentricity), and $x_I(t)$ (cosine of the orbital inclination)—along with the fundamental orbital phases $\Phi_A(t)$ with $A = \{r, \theta, \phi\}$ (up to initial conditions). In this work, we specifically choose these elements such that they satisfy the geodesic relations among semilatus rectum, eccentricity, inclination and the (dimensionless) orbital frequencies $\hat{\Omega}_A(a, p, e, x_I)$, as described in [200] (built on works [65,131,201,202]), so that it adheres to the convention set out in the LISA

convention document [203].⁴ The equations of motion then describe how these quantities evolve with respect to time. At adiabatic order, the evolution of the orbital elements is driven solely by the orbit-averaged backreaction due to GW emission, and the evolution of the fundamental phases is equivalent to that of the tangent geodesic (which changes with respect to time due to the slowly evolving orbital elements). The adiabatic equations of motion are therefore a system of six ordinary differential equations (ODEs) that must be solved numerically. Three of these ODEs are associated with the evolution of the (dimensionless) orbital elements $\alpha \in \{p, e, x_I\}$,

$$\frac{d\alpha}{dt} = \frac{\nu}{M} [\hat{f}_\alpha^{(0)}(a, p, e, x_I) + \mathcal{O}(\nu)], \quad (2)$$

where $\nu = \mu/M$ is the symmetric mass ratio, and we define our time parameter t to scale with M . Note that in most prior EMRI literature, including the first version of FEW [171], the small parameter is taken to be the (small) mass ratio $\epsilon = m_2/m_1 \leq 1$, and t is defined in terms of m_1 , not M . Here we opt to switch to the symmetric-mass convention, as doing so is known to improve agreement with intermediate and comparable mass binaries [164]. As $\epsilon = \nu + \mathcal{O}(\nu^2)$, changing from the small to the symmetric mass ratio (and likewise scaling time from m_1 to M) only affects the waveform phase at postadiabatic order, which we do not consider in this work. A detailed discussion of this change in mass convention can be found in Appendix A.

At adiabatic order, the forcing functions $\hat{f}_\alpha^{(0)}$ can be related to the asymptotic fluxes of energy and angular momentum (as well as a rate of change of the Carter constant, which can be deduced from asymptotic radiation fields and locally defined quantities along the orbit) via flux balance laws [113–117]. Due to axial symmetry, $\hat{f}_{x_I}$ is always equal to zero for equatorial inspirals. The remaining three ODEs describe the evolution of the fundamental phases $(\Phi_\phi, \Phi_\theta, \Phi_r)$, which are obtained by integrating their corresponding fundamental frequencies

$$\frac{d\Phi_A}{dt} = \frac{1}{M} \hat{\Omega}_A(a, p, e, x_I), \quad (3)$$

with $A = \{r, \theta, \phi\}$ and the dimensionless frequencies $\hat{\Omega}_A(a, p, e, x_I)$ found in, e.g., Refs. [65,131,201,202].

These phases can then be related to the $\{r, \theta, \phi\}$ Boyer-Lindquist coordinates of the secondary using the relations

⁴Some works treat p as dimensionless (e.g., [201]), while others take p to scale with m_1 (e.g., [200]). In this work, we make all quantities, including p , dimensionless and explicitly introduce the dependence on mass through the equations of motion. We discuss our choice of mass convention in Appendix A.

found in Ref. [200]. The phase evolution of each mode in Eq. (1) is then

$$\Phi_{mkn}(t) = m\Phi_\phi(t) + k\Phi_\theta(t) + n\Phi_r(t). \quad (4)$$

However, as only modes with $k=0$ contribute to the waveform for equatorial inspirals, we may write $\Phi_{mn} = m\Phi_\phi + n\Phi_r$ and ignore Φ_θ . By combining Eqs. (2) and (3), and integrating over a radiation reaction timescale ($\sim M/\nu$), we can see that the truncation in the evolution of the orbital elements leads to an error in the orbital phase of $O(\nu^0)$ [84].⁵ This is sufficient to demonstrate the FEW framework for IMRIs and EMRIs, and to obtain the quantitative results of the studies performed in this work. Once postadiabatic corrections to the evolution of the orbital elements are known, they can easily be incorporated as additional forcing terms into Eqs. (2) and (3) and quickly evaluated via multiscale expansion [61,86] or near-identity transform algorithms [166,168].

We generate inspiral trajectories by numerically integrating Eqs. (2) and (3) from a set of initial conditions $\{p_0, e_0, \Phi_{\phi 0}, \Phi_{r 0}\}$ with an adaptive ODE solver, returning a sparse trajectory consisting of ~ 100 points. Formally, the adiabatic approximation breaks down as the trajectory approaches the separatrix $p_{\text{sep}}(a, e, x_I)$, which represents the innermost stable orbit in Kerr spacetime [208]. As $p \rightarrow p_{\text{sep}}(a, e, x_I)$, the CO “transitions” across the boundary of the separatrix and plunges into the MBH on timescales where the adiabatic approximation of the inspiral trajectory is no longer valid [209]. In practice, it is sensible to terminate the trajectory at an intermediate point $p_{\text{sep}} + \Delta p_{\text{buf}}$ prior to crossing the separatrix. Not only is this regime better modeled by a multiscale transition-to-plunge expansion (e.g., Refs. [107–110,210,211]), but in our parametrization the forcing functions $\hat{f}_a^{(0)}$ diverge as $p \rightarrow p_{\text{sep}}$ [129,212]⁶ which leads to numerical instabilities in ODE integration techniques such as Runge-Kutta methods. The point at which one should switch from an inspiral to a transitional model (in terms of matching at first postadiabatic order) is not yet fully understood for generic trajectories. We can get some guidance by considering quasicircular orbits, for which it is well understood that the transition away from adiabatic inspiral happens when

⁵In this work, we are ignoring the GSF effects of transient resonances [204], which first enter the inspiral evolution at $O(\nu^{1/2})$ past the adiabatic order [169,205,206]. Transient resonances for equatorial eccentric inspirals are even weaker, entering at $O(\nu^1)$ (i.e., first postadiabatic order) [207] and can therefore be neglected in our adiabatic model.

⁶This divergence is due to the Jacobian between the parameters (p, e, x_I) and the integrals of motion (E, L_z, Q) ; see Eq. (7). As discussed in Ref. [212], one can avoid this issue by changing the parametrization, though at the cost introducing other numerical complications.

$p - p_{\text{sep}} \approx K(\nu/10^{-5})^{2/5}$ [106,107,209]; the numerical coefficient K is roughly 0.05 for zero-spin systems, and varies with primary spin and orbital inclination by a factor of order several tens of percent. As our inspiral model extends to $p - p_{\text{sep}} = 2 \times 10^{-3}$ (see Sec. III A), we expect it will be capable of incorporating such a transitional component in the future, though extending these transition models to eccentric orbits is still an area of active research [108,109].

In principle, this mathematical framework for obtaining inspiral trajectories is the same as that described in [171], which focused on inspirals into Schwarzschild black holes. However, extending this framework to incorporate spin on the primary component introduces additional complexity that raises computational costs and accentuates any systematic errors present (such as interpolation error) for the following reasons:

- (i) Introducing a increases the dimensionality of the $\hat{f}_{p,e}^{(0)}$ data grids from two to three dimensions. This increases both the computational cost of interpolating these grids and the error incurred in this interpolation, the latter of which must be carefully controlled to ensure waveforms are accurate.
- (ii) For $a \neq 0$, the separatrix no longer takes the simple form $p_{\text{sep}} = 6 + 2e$ and a more costly root-finding operation is often required [208,213].⁷
- (iii) The separatrix moves to smaller values of p as $a \rightarrow 1$. As $\hat{f}_{p,e}^{(0)}$ and $\hat{\Omega}_{\phi,r}$ vary increasingly rapidly as p decreases, the ODE integrator must take smaller steps to maintain a given error tolerance, further increasing computational costs.

A description of how the FEW inspiral generation framework must be modified to address these issues is presented in Sec. III A.

B. Mode amplitudes for Kerr inspirals

The amplitude of each mode may be written as

$$H_{\ell mn}(t) = A_{\ell mn}(t) {}_{-2}S_{\ell mn}(\theta; \hat{\omega}_{mn}) e^{im\phi}, \quad (5)$$

where ${}_{-2}S_{\ell mn}$ are spheroidal harmonics with spin-weight -2 , and

$$A_{\ell mn}(t) = -2 \frac{\hat{Z}_{\ell mn}^\infty(a, p, e, x_I)}{\hat{\omega}_{mn}^2(a, p, e, x_I)}. \quad (6)$$

⁷In principle, it is indeed possible to express the Kerr equatorial separatrix in a closed-form analytical expression $p_{\text{sep}}(a, e)$ through the solution of degree 4 polynomials for p (cf. Ref. [214]). However, the expression is relatively lengthy (involving nested radicals) and it is found that the computational efficiency is similar to that of the preexisting numerical root finding method. For that reason, we shall not employ this analytical expression in this work.

In analogy to Eq. (4), $\hat{\omega}_{mn} = m\hat{\Omega}_\phi + n\hat{\Omega}_r$ is the frequency of the mode, and the amplitude $\hat{Z}_{\ell mn}^\infty$ is obtained by solving the Teukolsky equation [66]. As the orbital elements (p, e, x_I) evolve during an inspiral, the dependent quantities $\hat{\omega}_{mn}$, $\hat{Z}_{\ell mn}^\infty$ and $S_{\ell mn}$ evolve as well; we indicate this by expressing $A_{\ell mn}$ and $H_{\ell mn}$ as functions of time.

The evolution of the spheroidal harmonics has the potential to greatly increase the cost of computing each mode, since it appears that one must continually recompute these functions as $\hat{\omega}_{mn}$ evolves along the inspiral. To avoid this cost, we use the fact that the spin-weighted spheroidal harmonics $_{-2}S_{\ell mn}(\theta; \hat{\omega}_{mn})e^{im\phi}$ can be expanded into spin-weighted *spherical* harmonics $_{-2}Y_{\ell m}(\theta, \phi)$ very efficiently (indeed, the spheroidal harmonics exactly reduce to spherical harmonics when $a = 0$ [215]). These functions do not vary during an inspiral and are inexpensive to compute. We take advantage of this behavior by projecting the amplitudes $A_{\ell mn}$ onto the spherical harmonic basis, allowing us to build each wave mode using angular functions that do not vary along the trajectory. The mode spectrum which results is generally broader in the deep strong field accessible to prograde inspirals (i.e., higher multipoles contribute more significantly to the overall waveform), which places more stringent requirements on the framework used to generate these amplitudes. The remapping procedure and subsequent modifications to the FEW amplitude module that are required in order to accommodate this increased complexity are described in Sec. III B.

III. MODEL IMPLEMENTATION

A. Inspiral trajectory

The accurate and efficient integration of Eqs. (2) and (3) is an essential component of the FEW framework. However, as pointed out in Sec. II A, incorporating the effects of spin presents additional challenges that must be addressed if accuracy and efficiency are to be retained. In this Section, we will demonstrate how the existing FEW framework for trajectory generation (which previously modeled zero-spin systems) can be modified for use as part of a Kerr equatorial eccentric waveform model. We will also discuss general modifications to the trajectory module of FEW we have made that improve the accuracy and efficiency of trajectories and waveforms, but are not limited in scope to the introduction of MBH spin into the model.

1. Computation of trajectory fluxes

To evolve the trajectories, we must compute the forcing functions $\hat{f}_{p,e}^{(0)}$ on the right-hand side of Eq. (2). At adiabatic order, $\hat{f}_{p,e}^{(0)}$ are related to the flux of energy and angular momentum lost by the system (both to null infinity and through the event horizon of the primary) due to GW emission,

$$\hat{f}_\alpha^{(0)} = -\frac{m_1^2}{m_2} \left[\frac{\partial \alpha}{\partial E} \left\langle \frac{dE}{dt} \right\rangle_{\text{GW}} + \frac{\partial \alpha}{\partial L} \left\langle \frac{dL}{dt} \right\rangle_{\text{GW}} \right], \quad (7)$$

where $\alpha = (p, e)$ and the Jacobian elements $\partial \alpha / \partial E$ and $\partial \alpha / \partial L$ are analytic functions of (a, p, e, x_I) as given in Appendix B of Ref. [66]. The time-averaged GW fluxes $\langle dE/dt \rangle_{\text{GW}}$ and $\langle dL/dt \rangle_{\text{GW}}$ are calculated from the Teukolsky equation using techniques in black hole perturbation theory, as described in, e.g., Refs. [66, 136, 216]. We omitted the Carter flux $\langle dQ/dt \rangle_{\text{GW}}$ in Eq. (7) as it is zero for equatorial orbits. The fluxes are also functions of (a, p, e, x_I) , but must be determined numerically.

In this work, we use the PYTHON library PYBHPT [124, 217] to compute $\langle dE/dt \rangle_{\text{GW}}$ and $\langle dL/dt \rangle_{\text{GW}}$. Fluxes are computed to a requested precision of 10^{-8} . We chose this value as a balance between computational cost and model accuracy, since multiscale analyses suggest that errors in the fluxes induce $O(\nu^{-1})$ errors in the GW phasing [84]. The computational cost of each flux calculation ranges from ~ 1 to ~ 1000 seconds and is highly dependent on the values of (a, p, e, x_I) . As PYBHPT performs computations in the frequency domain, we find that it is most efficient and accurate for orbits where $p > p_{\text{sep}} + 1$ and $e < 0.1$, since these orbits have a very narrow frequency spectrum. The computational cost of the flux calculations then grows as $p \rightarrow p_{\text{sep}}$ and $e \rightarrow 1$. (See Appendix B 4 for further details on the timing of flux calculations.) Compared to the rest of the waveform generation, the fluxes are incredibly expensive to compute. In the following Section, we describe how this computational cost is circumvented in the FEW framework.

2. Rapid computation of ODE derivatives

In order to integrate Eqs. (2) and (3), we must first specify the form of $\hat{f}_{p,e}^{(0)}$ and $\hat{\Omega}_{\phi,r}$. The latter terms are simple to incorporate; analytic expressions for the dimensionless orbital frequencies of geodesics in generic Kerr, which are equal to $\hat{\Omega}_{\phi,r}$ at adiabatic order, have existed in the literature for some time (see, e.g., [201, 218] and references therein) and are inexpensive to evaluate. We employ these expressions in our framework, simplifying terms involving x_I (given $x_I = 1$) where possible for efficiency and numerical stability.

Obtaining the forcing functions is less straightforward, as the data generation procedure in Sec. III A 1 is orders of magnitude too expensive to be applied on the fly during trajectory integration. However, these fluxes vary smoothly over the parameter space and are ideal targets for interpolation schemes. We therefore opt to generate forcing function values *en masse*, producing data grids in parameter space over which to perform this interpolation. Producing these data grids is computationally expensive, but must only be performed once; as the methods described in Sec. III A 1 are straightforward to parallelize, this can be

achieved relatively quickly on distributed computing resources. Of particular note is the accuracy to which the fluxes must be estimated, which is $\mathcal{O}(\nu)$ for the subradian accumulated phase error over the course of a typical trajectory evaluation.

Setting up the forcing function interpolants entails three main decisions that affect the interpolation accuracy and computational cost. The first is to choose an interpolation scheme. To avoid bottlenecking trajectory generation, the cost of evaluating the interpolants should be as low as possible while retaining accuracy (provided that the number of data grid points required to attain this performance is not prohibitively large). We therefore choose to employ a tricubic spline interpolant over a uniform grid, as implemented in the MULTISPLINE package [219]. While uniform spline methods enforce more stringent requirements on the grid spacing parameters (requiring a uniform step-size in all three dimensions), they are also far more efficient as they eliminate costly index-solving operations that are required when spacing is nonuniform. We select $E(3)$ boundary conditions for all parameters [220], as this has been shown to improve forcing function interpolation accuracy for quasicircular inspirals in previous work [124]. With this choice of interpolating function, each set of two flux values is computed in $< 1 \mu\text{s}$, which is sufficient to obtain year-duration trajectories in ~ 10 ms.

Next, we must specify our interpolation variables. This is essential for our framework because a regular grid cannot be constructed in (a, p, e) space (recalling that $x_I = 1$ for equatorial inspirals) due to the boundary imposed by p_{sep} that varies across the parameter space. Choosing sensible interpolation variables can also spread out rapid variation in the fluxes across the parameter space (especially near the separatrix), improving the accuracy of interpolation in these regions. To this end, we format our data grids in terms of three interpolation variables (u, w, z) that are related to (a, p, e) by bijective and analytically invertible transformations. For brevity, we will not state these transformations in the main text here (they are given in full in Appendix B) but instead summarize the important features of the resulting grid structure below:

- (i) Our grids span $a \in [-0.999, 0.999]$ and $p \in [p_{\text{sep}} + 10^{-3}, 200]$. For $p - p_{\text{sep}} > 9.001$, $e \in [0, 0.9]$.
- (ii) For smaller values of p , we employ an eccentricity taper such that the grids span a smaller range of eccentricities as p decreases. Data with high e and low p are extremely computationally expensive to compute accurately, but sources in this region of the parameter space are short lived and therefore are less likely to be detectable by LISA. In order to reduce computational costs, we exclude this region of parameter space from our model.
- (iii) We separate the parameter space into “inner” and “outer” regions (with some smaller overlap between them) such that this taper finishes at

$p - p_{\text{sep}} = 9.001$ without introducing a discontinuity into our grid coordinate transformations (which can lead to interpolation artifacts).

The shape and grid-point distribution of these grids is also shown in Fig. 20 and Fig. 21 in Appendix B, which clearly displays the structure of the eccentricity taper. Functions for performing these mappings are implemented in FEW and are readily applicable in the construction of user-specified grids.

Finally, we must choose the dimensions of the grid. With all other aspects of the problem fixed, this is essentially a trade-off between the accuracy of the interpolant and the computational cost of producing the grids. However, assessing how interpolation error in trajectory modeling propagates to the accuracy of the resulting waveform is not straightforward. We therefore opt to identify empirically what flux grid dimensions are required to satisfy waveform accuracy requirements. Based on this process, we settle on the dimensions $(N_u, N_w, N_z) = (129, 65, 65)$, $(65, 33, 33)$ for the inner and outer grid regions, respectively. As we will later demonstrate in Sec. IV, this ensures that grid point density (and therefore, the accuracy of the interpolated forcing functions) is sufficient for our model to be highly robust over the majority of the parameter space.

3. ODE integration with continuous solution

In the previous iteration of FEW, trajectories were integrated numerically with routines from the ODE module of the GSL software package [221]. This provided the necessary tools to obtain trajectories in milliseconds without the need to develop and test ODE integrator codes. However, this choice came with the significant drawback of inflexibility; it required that trajectory models be implemented in C (such that end users were forced to rebuild the FEW package for any nontrivial modifications to the inspiral model), and prohibits extension of the integration scheme to the GPU.

In FEWv2, we address these limitations with a bespoke implementation of the explicit embedded Dormand-Prince 8 (5,3) Runge-Kutta method [222]. The integrator logic is implemented entirely in PYTHON, consisting of vectorized array operations in order to remain performant. It accesses the ODE derivatives via a new class interface that the user can readily adapt in order to substitute these derivatives for any PYTHON function they choose. We anticipate this flexibility will greatly streamline the process of extending the FEW framework to investigate fundamental physics or environmental effects, which typically entails the modification of the trajectory fluxes (see e.g., Refs. [35,40,46,49,223] for examples of such modifications).

One key component of this new integrator framework is access to the *continuous solution* of the ODE solver. This allows for the construction of a C^1 -continuous interpolation of the evolution of the ODE variables as a seventh order piecewise polynomial [222]. As an example, the

azimuthal phase between time t_i and t_{i+1} , where t_i is the time of the i th node of the numerical ODE solution, can be interpolated via

$${}^i\Phi_\alpha(t) = {}^i c_0 + s({}^i c_1 + \bar{s}({}^i c_2 + s({}^i c_3 + \bar{s}({}^i c_4 + s({}^i c_5 + \bar{s}({}^i c_6 + {}^i c_7 s)))))), \quad (8)$$

where $\bar{s} = 1 - s$ and $s = (t - t_i)/(t_{i+1} - t_i) \in [0, 1]$. The coefficients ${}^i c_j$ for $j \in [0, 7]$ are functions of 12 intermediate evaluations of the ODE derivatives performed during an integration step (along with three further evaluations) and are therefore relatively inexpensive to compute. The explicit relation between the intermediate steps and the ${}^i c_j$'s can be found in Ref. [222] in terms of a constrained system of linear equations in terms of the Runge-Kutta coefficients.

The benefits greatly outweigh the additional computational cost of three additional derivative evaluations. A continuous approximation of $p(t)$ and $e(t)$ greatly simplifies the final root-finding step of trajectory evaluation (which was previously performed with an iterative Euler-step method) and retains the desired error tolerance of the integrator, improving the stability of the trajectory with respect to small perturbations in initial conditions. In FEWv1, cubic splines of the orbital phases are constructed as this continuous solution is not available, resulting in unstable waveform derivatives (and therefore, information matrices [224]) and (somewhat unexpectedly) led to relatively poor reconstruction of the more slowly varying waveform phasing in the early inspiral where trajectory points are typically very sparsely distributed. Access to an accurate interpolation of the orbital phases remedies this behavior, ensuring that waveform numerical derivatives remain smooth and phasing is consistent for the entire waveform.

A continuous approximation for the phase also yields similar piecewise polynomials for the orbital frequencies and their derivatives with respect to time; as shown in Fig. 2, this approximation is very accurate (exceeding nine decimal places in the frequencies for the majority of the inspiral). This is particularly useful in the construction of frequency-domain waveforms (which require this information as an input) and will enable resonance effects (which require careful monitoring of these frequencies, as well as accurate frequency derivatives) to be seamlessly folded into the FEW framework in the future.

4. Numerical integration of ODE system

With our framework for obtaining the derivatives in Eqs. (2) and (3) and integrating the resulting ODE system in place, we are ready to evolve a trajectory from a set of initial conditions $\{p_0, e_0, \Phi_{\phi 0}, \Phi_{\theta 0}\}$ given specific values for M , ν , and a . To improve the numerical stability of the

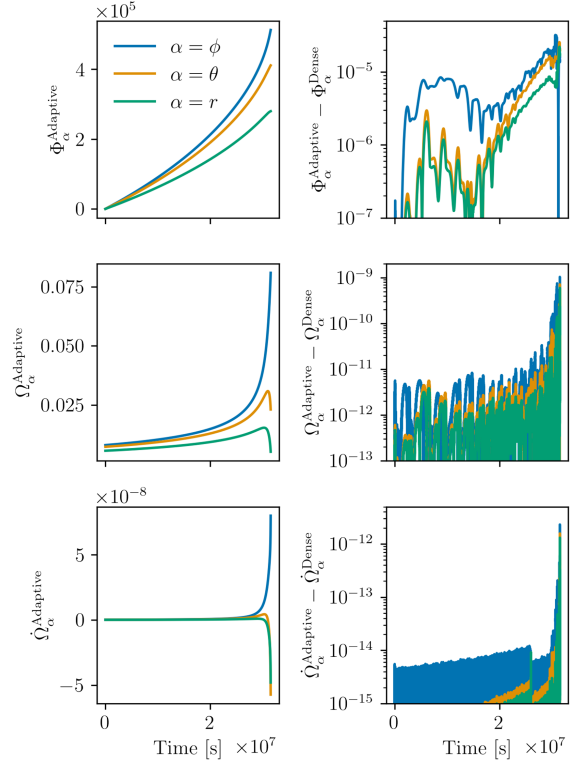


FIG. 2. Continuous solutions for the orbital phase, frequency, and frequency derivative (left) of the trajectory obtained with an adaptive stepping integrator, and their absolute difference (right) with respect to a densely stepped (fixed step size) integration. The considered binary system has component masses $(10^6, 10M_\odot)$ and $(a, e_0) = (0.998, 0.5)$. The polar phase and its derivatives (denoted here by θ) do not enter waveform generation for equatorial inspirals, but are included here for completeness.

integrator, we first rescale the problem such that the integration variables vary by $\mathcal{O}(1)$ across an inspiral. We achieve this by integrating the system with respect to the radiation-reaction timescale $t_{\text{rr}} = \nu t$ and eliminating the prefactor ν from Eq. (2) accordingly. Rather than scaling Eq. (3) similarly, we leave it unchanged [in order to achieve the desired property of $\mathcal{O}(1)$ scaling for all parameters] and instead rescale the orbital phase evolution postintegration by ν^{-1} . This approach is valid in the adiabatic approximation because (by definition) the $\hat{f}_{p,e}^{(0)}$ are independent of the orbital phases. A consequence of this choice is that, for an integrator absolute error tolerance σ_{tol} , the effective error tolerance on the orbital phases is σ_{tol}/ν . This is in fact a desirable outcome: we are mainly interested in tuning the accuracy with which $p(t)$ and $e(t)$ are evolved (as small errors in these parameters will gradually accumulate over an inspiral, coupling in turn to the orbital phase evolution), whereas the error in the reconstructed phase will remain of order σ_{tol}/ν , which is small and has negligible impact on waveform phase accuracy. This scaling is demonstrated in Fig. 3, where

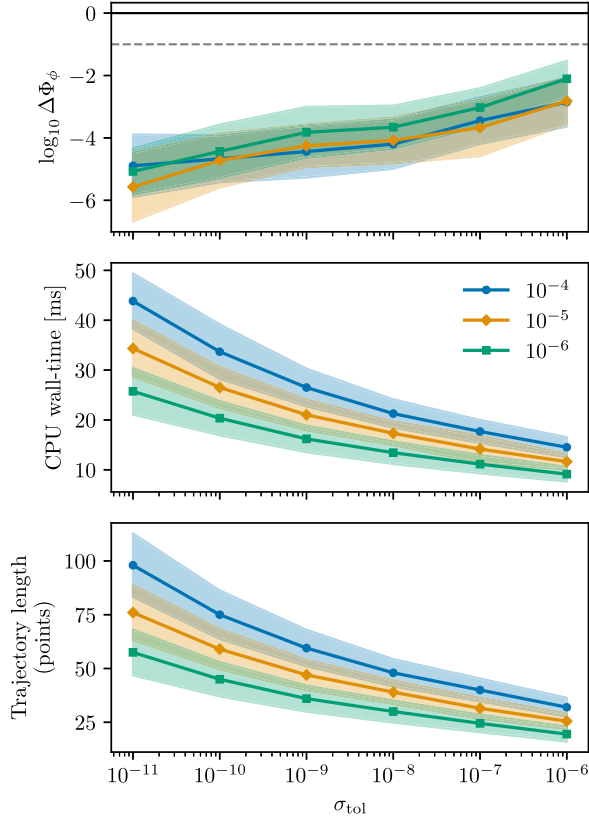


FIG. 3. Trajectory characteristics as a function of ODE integrator absolute tolerance σ_{tol} for 100 draws of four-year duration inspirals with ϵ of 10^{-4} (blue circles), 10^{-5} (orange diamonds) and 10^{-6} (green squares) respectively. *Top panel:* Deviation of the phase Φ_ϕ at end of inspiral with respect to a trajectory with $\sigma_{\text{tol}} = 10^{-13}$. *Middle panel:* Wall time per trajectory evaluation. *Bottom panel:* Number of adaptive points in the trajectory solution.

phase error grows in proportion to (but is several orders of magnitude larger than) the error tolerance of the ODE integrator. Trajectory wall time varies significantly with respect to ODE error; to ensure trajectories are accurate to within 10^{-3} radians, we set $\sigma_{\text{tol}} = 10^{-11}$ by default, but this can be relaxed by the end user according to their accuracy requirements. We also note that trajectory wall time will be hardware dependent; the timings presented in Fig. 3 were obtained with an Apple M3 processor.

Trajectory integration can be performed either forward or backward with respect to time [the latter achieved by negating Eqs. (2) and (3)]. The size of the first integration step from the starting point is chosen automatically with standard techniques [225]. The integrator then steps adaptively, producing sparse trajectories of ~ 100 points in length. During forward integration, we terminate adaptive time stepping once the integrator either attempts to step to a point with $p < p_{\text{stop}}$, where $p_{\text{stop}} = p_{\text{sep}} + \Delta p_{\text{buf}}$ with $\Delta p_{\text{buf}} = 2 \times 10^{-3}$, or exceeds the maximum duration T

set by the user.⁸ In the former case, we must then place a point at $p = p_{\text{stop}}$ to avoid truncation errors in the results of data analysis. As the adaptive stepping procedure first attempted to step past p_{stop} , we have access to a continuous solution in this interval; we therefore perform a numerical root-finding operation via Brent’s method to find the value of t where

$$\delta p_{\text{stop}}(t) = p(t) - p_{\text{stop}}(p(t), e(t)) \quad (9)$$

is equal to zero. This typically converges to within σ_{tol} of the root in ~ 5 iterations and contributes little to the overall cost of a trajectory evaluation. Once the root t_{root} has been found, the final trajectory point is obtained by evaluating the continuous solution at t_{root} . In the latter case of a trajectory duration exceeding T , this root-finding operation is not required and the continuous solution is simply evaluated at $t = T$ in order to obtain the last trajectory point. Backward integration proceeds similarly, except that p_{stop} is now determined by the boundaries of the forcing function data grids (see Appendix B for these definitions). If backward integration reaches this grid boundary before the duration exceeds T , a final point is placed at the boundary and integration is terminated.

The outputs of the trajectory integration are then fed along to later stages of waveform generation. The intermediate evaluation coefficients for the orbital phases are passed directly to the waveform summation stage (in a similar manner to which cubic spline coefficients for these quantities were inserted in the original FEW implementation). As the mode amplitudes are functions of $p(t)$ and $e(t)$, they cannot be interpolated with the continuous ODE solution; we instead retain the existing behavior, passing the sparsely distributed $p(t)$ and $e(t)$ points to the amplitude module to be evaluated (and subsequently interpolated with cubic splines). Improvements made to this amplitude module are discussed in the next Subsection.

B. Waveform amplitudes

1. Teukolsky amplitude data grids

As with the orbital-element forcing functions, the mode amplitudes $A_{\ell mn}$ are too expensive to compute as part of a rapid waveform generation framework. Instead, we may obtain $A_{\ell mn}$ via interpolation over precomputed data grids. Extending the amplitude data from Schwarzschild spacetimes to Kerr spacetimes presents two key complications:

⁸Recall that we terminate our grid at $p_{\text{sep}} + 10^{-3}$, both to avoid numerical instabilities in the adiabatic equations of motion near the separatrix and because the dynamics in this regime are better modeled in a transition-to-plunge (rather than inspiral) framework. We then choose $\Delta p_{\text{buf}} = 2 \times 10^{-3}$ to prevent trajectories from reaching this grid boundary.

- (1) Kerr trajectories evolve more “deeply” in the strong field of the MBH (i.e., they can have smaller values of p), leading to more (ℓ, m, n) harmonics contributing to the GW strain.
- (2) The GW strain naturally decomposes onto a basis of *spheroidal* harmonics that, like the amplitudes, also evolve in time—note the dependence on $\hat{\omega}_{mn}$ in the spheroidal harmonic in Eq. (5).

To address the former issue, we simply precompute more waveform harmonics for the Kerr model, specifically all (ℓ, m, n) modes within the limits $2 \leq \ell \leq \ell_{\max} = 10$, $|m| \leq \ell$, and $|n| \leq n_{\max} = 55$. For computational efficiency, we only compute mode amplitudes with $m \geq 0$ and infer amplitudes for $m < 0$ with a conjugate mode symmetry [130,136], as was performed in FEWv1 [171].

To circumvent the second complication, we expand the spheroidal harmonics in terms of the spherical harmonics, as proposed in [171] and implemented in other Kerr models [48,124]. This is expressed by the relation

$$S_{\ell mn}(\theta, \hat{\omega}_{mn})e^{im\phi} = \sum_{j=\ell_{\min}}^{\infty} b_{\ell mn}^j(t) {}_{-2}Y_{jm}(\theta, \phi), \quad (10)$$

where $\ell_{\min} = \max(2, |m|)$ and $b_{\ell mn}^j(t)$ are spherical-spheroidal mixing coefficients. Further discussion regarding this expansion [and the computation of $b_{\ell mn}^j(t)$] can be found in, e.g., Appendix A of [130]. The time dependence of the spheroidal harmonics is then absorbed by $b_{\ell mn}^j(t)$. Inserting Eq. (10) into Eq. (1), we then sum over the spheroidal harmonic mode amplitudes $A_{\ell mn}$ to obtain their corresponding *spherical* harmonic amplitudes,⁹

$$\mathcal{A}_{\ell mn} = \sum_{j=\ell_{\min}}^{\infty} b_{jmn}^{\ell} A_{jmn}. \quad (11)$$

After this remapping, the waveform strain is then written as

$$h(t) = \frac{\mu}{d_L} \sum_{\ell mn} \mathcal{A}_{\ell mn}(t) {}_{-2}Y_{\ell m}(\theta, \phi) e^{-i\Phi_{mn}(t)}. \quad (12)$$

As Eq. (12) expresses the waveform strain for eccentric equatorial Kerr inspirals in the same form as that of Schwarzschild inspirals, we are free to apply FEW’s rapid mode-summation framework without modification. This approach is readily extendable to eccentric and inclined Kerr inspirals once flux and mode amplitude data for these systems are available [66,123,134,137,171,226].

⁹See also, e.g., Appendix A of [87], Sec. III B of [155], and Sec. II B of [124] for explicit derivations of this procedure.

2. Interpolation of mode amplitudes

Due to the large number of harmonic modes that contribute significantly to a typical EMRI waveform, obtaining their amplitudes efficiently via interpolation is a vital component of the FEW waveform generation framework. As was performed in FEWv1, we opt to interpolate the complex amplitudes in terms of their real and imaginary components. These quantities vary smoothly over the parameter space, and this allows for both the leading-order term $|\mathcal{A}_{\ell mn}|$ and the postadiabatic contribution $\psi_{\ell mn} = \arg(\mathcal{A}_{\ell mn})$ to be obtained directly, at the cost of the additional memory needed to store data grids for both components at full resolution. While the latter term is not strictly speaking necessary for an adiabatic model, we seek to demonstrate that all effects at first postadiabatic order can readily be included in the FEW framework, and so include this for completeness. While it is tempting to interpolate $|\mathcal{A}_{\ell mn}|$ and $\psi_{\ell mn}$ directly (with the accuracy requirements on the latter being relaxed due to its post-adiabatic nature), the former is nonsmooth (due to zero crossings in the amplitude surfaces) and the latter must be unwrapped to avoid discontinuities (which is challenging in multiple dimensions). We plan to investigate whether a form of amplitudes exists that retains the benefits of both conventions in more detail in future work.

An important feature of this interpolation problem is that it is highly parallelizable with respect to both the mode indices (ℓ, m, n) and the parameters (a, p, e) . It is therefore sensible to design an interpolation framework that exploits this by computing mode amplitudes over many mode index combinations and parameter values simultaneously. For FEWv1, this was achieved by fitting a reduced-order model (in the form of a neural network) that rapidly evaluated a set of basis coefficients for multiple sets of input parameters. These coefficients were then converted into a complete set of mode amplitudes (i.e., over all sets of mode indices) via a matrix dot product; see Section 4 of [171] for further discussion. This approach is computationally efficient and well-suited to acceleration via GPU vectorization techniques. However, it suffers two significant drawbacks. First, the model struggles to accurately estimate the amplitudes of weaker (but still significant) modes. This was identified in previous work and was observed to limit the accuracy of the waveforms obtained to mismatches $\sim 10^{-5}$ [171], which is on the order of the accuracy required for unbiased analysis of the loudest expected astrophysical EMRI signals (which may have SNRs in the high hundreds [28]; see Sec. IV C for further discussion). Note that for quieter sources, larger mismatches than this will not necessarily lead to parameter biases [171]; such a requirement is only enforced by the (potential) existence of particularly loud (and therefore informative) signals.

Even if this limitation can be overcome (e.g., by tuning the loss function used to train the neural network), a second drawback to this approach is that it is completely inflexible

with respect to the requested mode content of the waveform. For instance, if the user wishes to generate a waveform containing only a handful of modes (e.g., as part of the early stage of a signal extraction algorithm), the model must first generate every mode amplitude before discarding the majority of them. Potential improvements to the mode selection process that reduce the number of mode amplitudes to be generated, such as a parametrized “mask” in mode index space, will similarly be affected (albeit to a lesser extent). We therefore seek an alternative approach that addresses these limitations, accurately and adaptively constructing the mode amplitude spectrum without compromising efficiency.

As each $\mathcal{A}_{\ell mn}$ varies smoothly over the parameter space, in a similar manner to the fluxes, they can be accurately modeled with interpolation methods. However, the accuracy with which each mode amplitude must be recovered is far less stringent than that of the forcing functions, as systematic errors in these amplitudes do not accumulate over the course of the inspiral, whereas inaccurate ODE derivatives lead to secularly growing errors in the inspiral trajectory (and hence phasing). Estimating each mode amplitude with a relative error of $\sim 10^{-2}$ is sufficient for systematic biases to remain small with respect to measurement precision for astrophysically relevant systems with SNRs of ~ 100 [227], with some dependence on how SNR is distributed between modes (which varies as a function of spin, eccentricity, and orbital separation). What is important is to attain this relative error across the spectrum of mode amplitudes (which spans several orders of magnitude). Basis-compression approaches like reduced-order modeling can struggle to achieve this global relative error.

We instead opt to perform spline interpolation, which is particularly well suited to this task and addresses the limitations outlined in the previous paragraph. Interpolating each mode amplitude separately ensures that the weaker modes in the spectrum are still estimated accurately. Mode amplitudes are readily interpolated in parallel, but as each interpolation operation is easily separable from the others, it is straightforward to evaluate only a requested sub-set of the full mode spectrum. The relationship between $\mathcal{A}_{\ell mn}$ and a is sufficiently weak (at least for nonextremal systems [124,175,180]) that linear interpolation with respect to this parameter performs adequately over the majority of the parameter space, further reducing computational costs. We therefore interpolate mode amplitude data with a “bicubic+linear” spline interpolation framework, striking a balance between accuracy and efficiency. Similarly to the case of the fluxes, we define data grid coordinates (u, w, z) connected to (p, e, a) by invertible transformations that are described in Appendix B. These transformations are of similar form to those used for the forcing function grid setup, but with adjusted coefficients to ensure the strong-field region is sufficiently densely sampled to capture the rapid amplitude

variations that are characteristic of this region of the parameter space. The form of the eccentricity tapering is correspondingly altered to ensure it retains the same shape as the flux grids, such that the forcing function and amplitude interpolation domains (and therefore, the domain over which the waveform is defined) are consistent. The value of the mode amplitude $\mathcal{A}_{\ell mn}$ given grid coordinates (u, w, z) is then written as

$$\mathcal{A}_{\ell mn} = (\mathcal{A}_{\ell mn}^+ - \mathcal{A}_{\ell mn}^-) \frac{z - \hat{z}_-}{\hat{z}_+ - \hat{z}_-} + \mathcal{A}_{\ell mn}^-, \quad (13)$$

where $\mathcal{A}_{\ell mn}^{\pm} \equiv \mathcal{A}_{\ell mn}(\hat{z}_{\pm})$ are obtained via a bicubic spline interpolant over (u, w) , and $(\hat{z}_-, \hat{z}_+)$ are the values of the grid points immediately (below, above) z .

To justify the wide range of modes we include in our waveform model, in Fig. 4 we apply our amplitude interpolant to demonstrate how the morphology of the waveform mode spectrum changes over an eccentric ($e_0 = 0.8$) inspiral into a rapidly spinning ($a = 0.998$) MBH, plotting the mode power $P_{\ell mn} = |\mathcal{A}_{\ell mn}|^2$ as a function of (ℓ, m, n) . At early times, the waveform contains rich n -mode content due to the high initial eccentricity, but the mode spectrum rapidly tails off with respect to ℓ as the CO is orbiting in the relatively weak field. As the inspiral circularizes due to GW emission, the higher-order n -mode contribution gradually weakens; meanwhile, the high spin of the MBH allows the CO trajectory to extend deep into the strong field of the MBH, significantly accentuating the higher ℓ -mode content of the waveform. Typically, for systems with higher eccentricities at lower semilatus rectum values than what is shown in Fig. 4, more power is contained in the higher n -mode content of the spectrum (particularly for higher values of ℓ) [67,134].

One significant drawback of interpolating each mode amplitude with cubic splines (as opposed to a reduced-basis representation) is that the size of the spline coefficient arrays that must be stored in memory grows linearly with the number of modes supported in the model. For the 6993 modes that we interpolate in our waveform model, these coefficients arrays total ~ 5 GB in size, which is sufficiently compact to fit in the memory of most GPUs. However, some care will be required once FEW is extended to inclined and eccentric inspirals in future work, as the number of waveform modes (and therefore the memory requirements) will increase significantly (with typically as many as $\sim 10^5$ – 10^6 waveform modes [66,122,171]).

C. Mode selection and waveform summation

While the inclusion of spin on the primary introduces additional complexity to the generation of inspiral trajectories and mode amplitudes, once these have been addressed, the waveform may be constructed in much the same way as for spinless systems. We review this

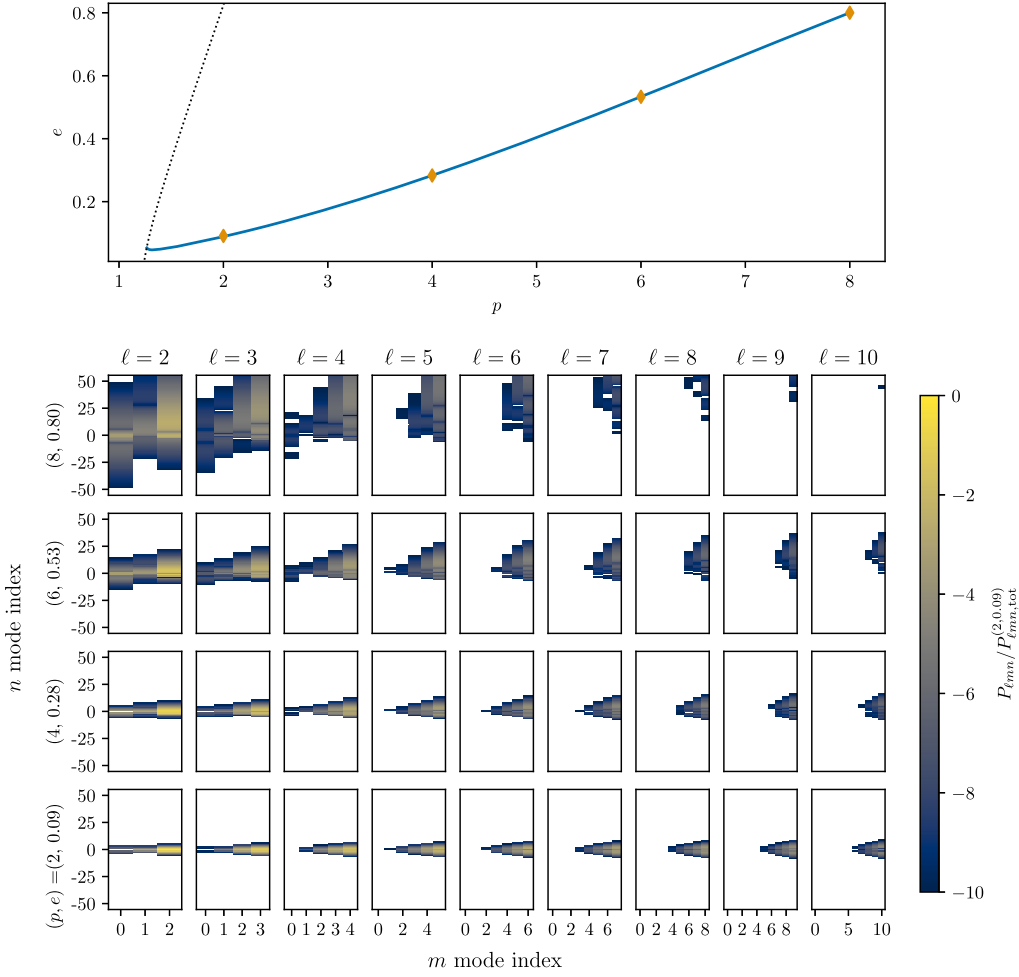


FIG. 4. Top panel: The blue curve represents a highly eccentric and equatorial inspiral trajectory for $\epsilon = 10^{-5}$ and $a = 0.998$. The black dotted line describes separatrix p_{sep} , which evolves as a function of eccentricity. The orange diamonds represent individual snapshots of the trajectory. Bottom panel: A plot of the mode spectrum normalized by the total mode power $P_{\ell mn, \text{tot}}^{(p=2, e=0.09)}$ at the points in the trajectory indicated by the orange diamonds. Modes with a normalized power below 10^{-10} are not shown. As the orbital parameters evolve, the distribution of the mode power with respect to mode index changes significantly, presenting the need for wide mode index coverage in a typical inspiral. The number of modes per row that account for 99% of the total power in that row are (129, 40, 24, 26); the union of these mode sets consists of 160 elements (due to the evolving shape of the mode spectrum). For 99.999% of the total power (our default for waveform generation) these counts become (718, 345, 199, 174) and 988, respectively.

procedure below, which is largely unchanged except for the incorporation of the continuous ODE solution described in Sec. III A 3; see [171] and [173] for more information regarding these stages of waveform generation in the time domain and frequency domain, respectively.

Before performing the summation in Eq. (1), we apply a mode selection process to reduce the number of modes that must be summed without sacrificing accuracy, which proceeds as follows. First, at each sparse trajectory point, we compute the mode power $|A_{\ell mn}|^2$ for each waveform mode in our amplitude data grids. We then sort the list of mode power values at each individual point, perform a cumulative summation over this list and truncate the sum once it exceeds a fraction $(1 - \kappa)$ of the total power. We assume that modes that do not pass this cutoff do not

strongly contribute to the overall waveform. However, to ensure that the waveform is smooth with respect to time, we take the union of the sets of modes that pass the individual cutoffs and retain only these modes for waveform summation. The tuning parameter κ controls the accuracy of the waveform with respect to this procedure; in line with previous work, we set $\kappa = 10^{-5}$ by default. We have verified that this value is a sensible choice for a variety of EMRIs and IMRIs, as discussed in Sec. V F.

After mode selection, we perform the summation in Eq. (1) over the remaining sets of mode indices. For time-domain waveforms, we interpolate the sparse quantities $\mathcal{A}_{\ell mn}$ and Φ_{mn} (with cubic splines and the continuous ODE solution, respectively) on a uniform grid of time values of spacing dt , summing the interpolated waveform modes

together according to Eq. (1). Frequency-domain waveforms are obtained via the approach described in [173], making use of the stationary phase approximation. As this approach assumes that the frequency evolution of each mode is described by cubic splines (such that the stationary phase relation has analytic roots), we retain this structure rather than incorporating the continuous solution depicted in Fig. 2. The frequency-domain waveform also requires higher-order frequency derivatives to approximate the Fourier phase in each bin; as the derivatives of this cubic spline are less accurate than those obtained with the continuous solution, we rewrite this part of the frequency-domain summation to use the coefficients output by the ODE solver. We found that this significantly improved the consistency of the time-domain and frequency-domain waveforms. Modifying the root-finding procedure of the stationary phase approximation to incorporate our more accurate solution for the frequency evolution (e.g., by taking the analytic root as an input to Newton’s method, which will converge in 1–2 iterations) will further improve this consistency at minimal computational cost, and is left to future work.

Once either time- or frequency-domain waveforms have been obtained in the source frame, they are transformed to the detector frame (as described in [228]) for use in data analysis.

IV. VALIDATION OF MODEL ACCURACY

With our waveform generation framework fully specified, we will now explore the accuracy of our waveform model at adiabatic order. This is made difficult by lack of accurate waveform templates to compare against: numerical relativity simulations have not been computed (however, see recent efforts in Refs. [70–72]) in this region of the parameter space (especially for the number of orbital cycles [$\sim \mathcal{O}(\nu^{-1})$] required for useful comparisons to be made), and PN models are not trustworthy in the strong-field region of our domain where confirming the accuracy of our model is most pertinent. We therefore opt for a different approach in which we quantify significant sources of systematic error in our model (which in combination are responsible for any differences between our model and “exact” adiabatic EMRI waveforms), with the aim of understanding how these errors propagate to the accuracy of our constructed waveforms. In this way, we can roughly quantify the faithfulness of our model with respect to an “error-free” adiabatic model without the need for independent waveforms to compare against. Throughout this section, we will examine each component of our waveform generation in turn: we first consider the trajectory module (Sec. IV A), followed by the amplitude module (Sec. IV B) and conclude with an examination of how these systematics impact accuracy at the waveform template level (Sec. IV C).

In addition to this self-contained examination of our model, we also verify that it correctly matches other models in limiting cases in Appendix E. In combination with the systematic validation that we perform throughout this section, the results of these comparisons confirm that our model produces accurate waveforms for spinning and/or equatorial IMRIs and EMRIs across its domain of validity of $|a| \leq 0.999$ and $e_0 \leq 0.9$. In this process, we identify that our model is least accurate for: $a > 0.99$ and $p - p_{\text{sep}} \lesssim 1$; and $e \gtrsim 0.85$. These are regions where self-force data vary rapidly and are most challenging to interpolate.

A. Trajectory validation

In this Subsection, we verify that our model for the inspiral trajectory of the CO accurately traces the orbital phasing of the system to adiabatic order over the long timescales typical of the asymmetric-mass binary systems observable by LISA. As our model only incorporates adiabatic contributions to the inspiral evolution, one could argue that we need only construct the phase evolution to an accuracy of $\mathcal{O}(1)$ radian, as we cannot hope to match trajectories that include postadiabatic contributions to better than this scale (and typically, ignoring postadiabatic terms regularly leads to 10–100 radians of phase error [122,149,164,169], at least for the quasicircular EMRIs). However, we can assume that postadiabatic effects will be implemented in our framework once they are readily obtainable (as they are necessary to achieve sufficient accuracy for data analysis purposes), so this is not a useful requirement for us to satisfy. In this work, our focus remains on the accuracy of our model for the adiabatic contribution to the inspiral. Quantitatively assessing how phase errors translate into systematic biases in EMRI parameter estimation is not straightforward. However, in Ref. [149], which examined quasicircular inspirals into nonspinning MBHs, a phase error of $\sim 10^{-1}$ radians was sufficiently small for any resulting biases to remain subdominant to statistical uncertainties. Our aim therefore is to ensure that our trajectory model reproduces inspiral phasing with at least this level of accuracy.

The faithfulness of the phase evolution of our trajectory model with respect to error-free adiabatic trajectories is wholly dependent on the accuracy to which the forcing functions $\hat{f}_{p,e}^{(0)}$ are estimated. In our framework, there are two potential sources of systematic error at this stage:

- (1) The accuracy with which the $\hat{f}_{p,e}^{(0)}$ data grids are computed; as $\hat{f}_{p,e}^{(0)}$ are obtained from infinite mode-sums of oscillatory integrals, they are expected to contain errors due to limitations in numerically resolving the integrands and from the finite truncation of the mode-sum.

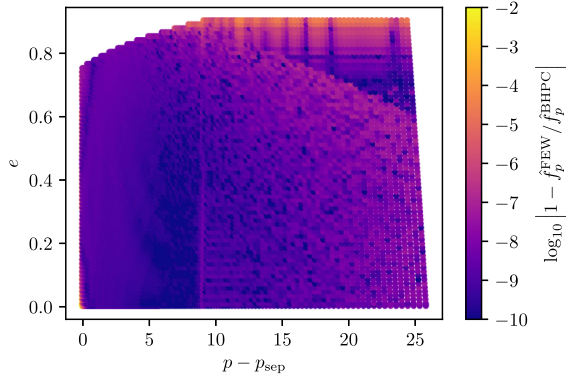


FIG. 5. Relative error in $\hat{f}_p^{(0)}$ between an interpolation of our data grid and the corresponding data of Ref. [136], available from the BHPC data lake [231]. The two datasets agree to better than eight digits over the majority of the parameter space. Here we examine prograde orbits with $a = 0.5$; we observe similar agreement for other spin values $a \in [-0.9, 0.9]$.

- (2) The spline interpolation of these data grids, which is closely tied to the grid-point density and choice of parameter conventions.

We begin with the first of these error sources by comparing our computed data grids with two independent data sets. In order for this comparison to be made, it is still necessary for us to interpolate our data—the results we present here are therefore a combination of interpolation error and data computation errors. The purpose of these comparisons is to ascertain whether any regions have been computed with significant errors compared to the errors incurred in this interpolation (we will examine interpolation error separately later in this Subsection).

The first of these datasets was computed by the authors of Ref. [136] (built on the black hole perturbation club (BHPC) code developed in Refs. [134,229,230]) and is publicly available on the BHPC website [231]. It broadly and densely covers a large fraction of our parameter space. The authors accompany their data with an error estimate informed by their truncation of mode summations. For $p \lesssim p_{\text{sep}} + 1$, the estimated error becomes prohibitively large for validation purposes—in this region, the discrepancy between our data and theirs grows proportionally to the predicted error. We therefore excise this region of the parameter space from this comparison (and will examine it separately below with a separate dataset). The region of the BHPC data we retain extends to $p \sim 25$ and $e = 0.925$, providing good coverage of a large fraction of our grid. Beyond $p \sim 25$, we validate our data separately with BHPC’s PN results (see Appendix E 2). The relative error between an interpolation of our data and the BHPC data for $\hat{f}_p^{(0)}$ is shown in Fig. 5—for brevity, we do not include the comparison with $\hat{f}_e^{(0)}$, which follows similar behavior. We attain excellent agreement with this dataset across the

parameter space, with relative error remaining below $\sim 10^{-8}$ over the vast majority of the grid. Above $e \sim 0.8$, errors steadily rise to $\sim 10^{-4}$. This is expected: mode-sums with respect to n converge slowly in this region, requiring one to compute and sum over hundreds of n modes. The large n harmonics become increasingly more difficult to resolve at floating point precision and introduce their own source of error if they are under-resolved (similar to how it is incredibly difficult to calculate the 100th harmonic in a Fourier series expansion). For a single n mode, this error is typically subdominant to the total error, but if the mode-sum is not truncated early enough, the accumulation of these small errors can saturate the overall relative accuracy of the results. As both our data and that of the BHPC are impacted by this limitation, it is not clear which of the two has been computed more accurately in this high eccentricity region, only that the two datasets are consistent to ~ 4 digits of precision. This comparison is performed for prograde orbits with $a = 0.5$ —we find similarly good agreement across all spin values ($a = 0.1, 0.3, 0.5, 0.7, 0.9$) for both prograde and retrograde configurations, confirming that we have accurately computed and interpolated our $\hat{f}_{p,e}^{(0)}$ data grids over this part of the parameter space.

To inspect the accuracy of our data grids in the strong-field region $p \rightarrow p_{\text{sep}}$, we also compare against data computed with the GREMLIN code, independently developed by author Hughes and previous collaborators. (See Ref. [67] for an overview of this code’s foundations, and Refs. [232,233] for discussion of updates to notation and methods that this code uses. Additional recent optimizations and algorithmic improvements, including a port to the Open Science Grid [234], will be described in a future publication, accompanied by a public release after removing proprietary code currently used by this package.) This data is more densely sampled in this region than our grids in a and p ; comparisons are done on four slices in eccentricity that follow a taper similar to the one we include in our grid reparametrization. We distinguish forcing functions computed with FEW and GREMLIN with superscript “FEW” and “SAH” respectively. The relative error between an interpolation of our data and each of these slices is shown in Fig. 6 for the case of $\hat{f}_p^{(0)}$. We observe excellent agreement (with relative errors less than one part in a million) for $p - p_{\text{sep}} \gtrsim 10^{-2}$. Closer to the separatrix, we observe a larger discrepancy between the datasets, with relative errors reaching $\sim 10^{-3}$; this is expected as (especially at larger eccentricities) it is computationally challenging to compute $\hat{f}_{p,e}^{(0)}$ in this region, and our interpolation error begins to dominate at high values of a (as indicated by the alternating bands of larger error with respect to a), as we will examine further below. However, EMRIs typically complete relatively few orbital cycles in this region of the parameter space (compared to their total number of in-band cycles) before

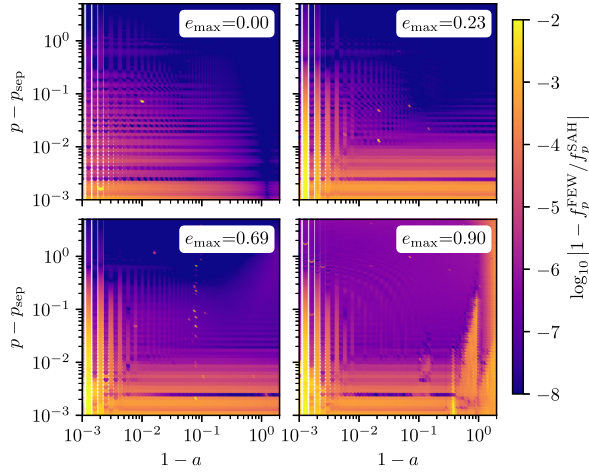


FIG. 6. Relative error in $\hat{f}_p^{(0)}$ between an interpolation of our data grid and a dataset computed with the GREMLIN code. The four panels shown correspond to four slices in eccentricity that follow the tapering scheme described in Appendix B. The two datasets generally agree to ~ 6 digits for $p - p_{\text{sep}} \gtrsim 10^{-2}$, increasing to ~ 4 digits below this value. Some data quality issues are also evident in the lower right panel; see main text for discussion.

crossing the separatrix, so even relatively large errors will not significantly impact the phase accuracy of the trajectory.

While the deviations observed are due in part to errors in the interpolation of our data grids, they are also impacted by differences in the convergence criteria associated with each software. The GREMLIN computations truncate at either a relative precision of 10^{-7} or up to $n = 250$ closest to the separatrix, whereas PYBHPT cuts off at either a relative precision of 10^{-8} or $n = 300$; as these higher n modes can contain significant errors due to limited numerical precision, truncating the mode-sum at different values of n may yield results that differ substantially beyond the first few digits. Moreover, relative precision cutoffs may not be directly comparable between datasets, as this quantity may be estimated in different numerical frameworks. Likewise, the BHPC data were computed with mode-sums truncated at up to $n = 1000$ or $\ell = 25$ for the (highly eccentric) strongest-field orbits to ensure a relative precision at a level of 10^{-6} [136], which may again yield significant differences with our data (such as those observed for the highest eccentricities in Fig. 5). Errors due to these differences in methodology as a function of parameter space are not well understood at present; further investigation of these systematics is beyond the scope of this work and reported in a separate work [235] (but is worthy of future study, particularly for eccentric and inclined orbits).

Two other significant features are present in the heat maps shown in Fig. 6. The first is that for $a \rightarrow 0.999$, we observe a significant increase in relative error in alternating

bands of constant a . These bands correspond to a values lying on grid nodes (low interpolation error) and exactly between grid nodes (high interpolation error) respectively, and their existence suggests that our grid density with respect to a is insufficient near p_{sep} (even despite our reparametrization, which concentrates grid points in this region). As we will see later in this Subsection, this has little impact on the accuracy of our trajectories. The second feature is observed for the largest eccentricity slice (bottom-right panel of Fig. 6), $a \in [0, 0.5]$ and $p - p_{\text{sep}} \lesssim 1$, where a patch of larger relative error is observed. These points correspond to $e \sim 0.5$ near the separatrix (due to the eccentricity taper with respect to $e_{\text{max}} = 0.9$, see Appendix B) and the larger errors are a consequence of ill-converged mode-sums typical of more eccentric orbits near $p = p_{\text{sep}}$. As this eccentricity slice is at the edge of our grid, trajectories typically do not perform many orbital cycles in this region before their eccentricities decay to lower values where this numerical precision issue is absent; we therefore do not anticipate this will significantly impact the accuracy of our model. In this eccentricity slice, a steadily-growing relative error is also observed for $a \rightarrow -0.999$, which corresponds to $e \rightarrow 0.9$ and is explained by the differences in convergence criteria between implementations (as described above).

Taking these comparisons with independent datasets as a reference point, we assess the impact of interpolation error on our estimation of $f_{p,e}^{(0)}$ as a function of the inspiral parameters. As our data grids have dimensions of the form $2^n + 1$ for some positive integer n , we can construct a grid of dimension $2^{n-1} + 1$ (halving the point density) by omitting every second point in the grid, which we will refer to as “down-sampling” the grid by a factor of 2. We can then evaluate this down-sampled grid at the omitted points, placing the upper bound on the interpolation error for the full grid across the parameter space. The comparison is performed in Fig. 7 for $a \sim 0.998$, the spin value at which we find the interpolation errors to be most significant (in line with what is observed in Fig. 6); again, we omit our comparison with $\hat{f}_e^{(0)}$ for brevity, which yields similar error estimates. Over the majority of the parameter space, this interpolation error estimate exceeds those shown in Figs. 5 and 6 where we compare to independent datasets. This means that for a trajectory model built with data grids down-sampled by a factor of 2 (or higher), interpolation error due to insufficient point density will be the dominant contribution to orbital dephasing.

We will now use this result to roughly quantify the accuracy of our model. We consider models built from data grids down-sampled by factors of 2, 4 and 8, and compare phase trajectories constructed with these down-sampled models to those produced by our full-scale model (which has the full grid density). In the top panel of Fig. 8, we plot histograms for dephasings between each down-sampling

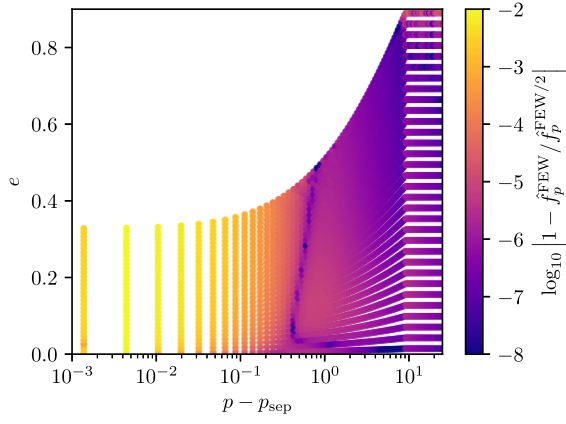


FIG. 7. Relative error in $\hat{f}_p^{(0)}$ between data grid values and those obtained via interpolation over this grid down-sampled by a factor of 2 (denoted “FEW” and “FEW/2”). The relative errors shown provide an upper bound on the interpolation error incurred by the full-resolution interpolant. Shown is a slice in (p, e) with $a = 0.9985$, which is the spin value for which our interpolant is the most inaccurate.

scenario and our full-resolution model for a random sample of 10^3 inspirals with p_0 tuned such that they inspiral after 4 years (see Appendix C for details of the random-draw procedure). We observe that these dephasings display power-law convergence (i.e., in line with the power of 2 of the down-sampling factor) as a function of grid density. By fitting this convergence trend between grid density and dephasing for each data point, we then extrapolate to obtain an estimate of the phase error at the full grid density, resulting in the dashed histogram in the top panel of Fig. 8. Taking this extrapolation as a rough estimate of our model’s dephasing with respect to a “perfect” adiabatic inspiral model, we predict that our model attains a phase accuracy of $\sim 10^{-4}$ rad over the majority of the parameter space, with a higher tail extending to $\sim 10^{-2}$ rad. This is significantly smaller than our original target of 10^{-1} rad, and should be sufficiently accurate for the analysis of EMRI signals that LISA is expected to observe. We emphasize that this extrapolation assumes that our interpolation errors are significantly larger than the numerical precision errors in our computed data. In general, the worst-behaved systems have very high spins ($a > 0.99$) and/or very high initial eccentricities ($e_0 > 0.8$). This highlights that these systems may require a more careful treatment (both in the generation and interpolation of forcing function data), an exploration of which we leave for future work.

It is also useful to understand whether insufficient sampling density in one of u , w or z (corresponding to p , e and a) contributes significantly to interpolation error in $\hat{f}_{p,e}^{(0)}$. To examine this more closely, we also construct trajectory models from grids down-sampled by a factor of 2 in a single dimension at a time. Histograms of dephasings between trajectories from these models and

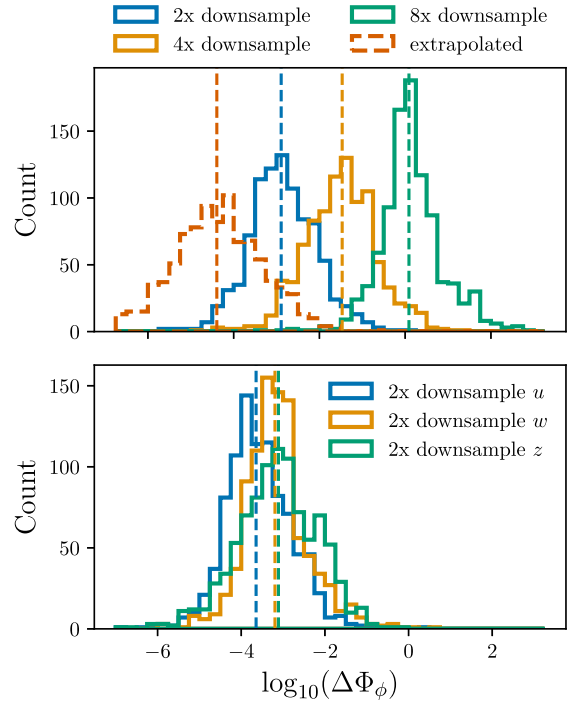


FIG. 8. Dephasing due to down-sampling of $\hat{f}_{p,e}^{(0)}$ interpolant data grids for a sample of 10^3 four-year inspirals. In all cases, dashed vertical lines indicate median values. Top panel: Convergence of dephasing when one down-samples by a factor of 2^1 (blue), 2^2 (yellow), or 2^3 (green) in all grid dimensions; the red dashed histogram is obtained by extrapolating down-sampled phase errors to the full grid density (2^0). Bottom panel: Dephasing when halving the sampling density in the u (blue), w (yellow), or z (green) grid dimension only (with $\{u, w, z\}$ corresponding to $\{p, e, a\}$, respectively). The histograms overlap significantly, indicating appropriately proportioned grid dimensions, and lie between the extrapolated and lowest-order down-sampled histograms in the top panel.

our full-resolution model are shown in the bottom panel of Fig. 8. As these histograms overlap significantly, we conclude that the point density of our grids is well proportioned with respect to grid dimension. As expected, these histograms also lie between those corresponding to the extrapolated dephasings and the model constructed with grids down-sampled by 2 in all dimensions.

B. Mode amplitude validation

We now turn to examine the faithfulness of the amplitude module of our waveform generation framework. In general, any systematic errors in the mode amplitudes $A_{\ell mn}$ should be significantly less impactful than those present in the forcing functions. Contributions from the latter grow secularly over inspiral, whereas amplitude errors remain of fixed order. For systematic biases in EMRI parameter estimation to be sufficiently small (assuming no orbital phase errors), the relative error in each mode amplitude must be smaller (in an order-of-magnitude sense) than the

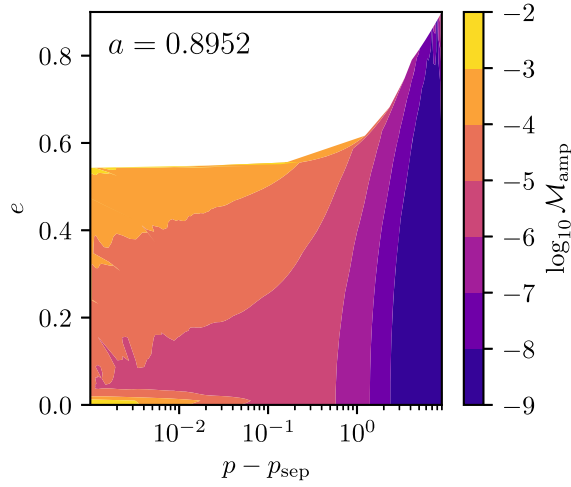


FIG. 9. Mode amplitude (approximate) mismatch $\mathcal{M}_{\text{amp}}$ [Eq. (14)] between interpolated amplitudes from our model ($\mathcal{A}_{\ell mn}^{\text{FEW}}$) and those produced with the GREMLIN code ($\mathcal{A}_{\ell mn}^{\text{SAH}}$). Mismatch due to amplitude interpolation error increases significantly as p decreases below $\sim p_{\text{sep}} + 1$, and is slightly worse for higher eccentricities. This behavior is due to our linear interpolation of amplitudes with respect to a (explored in Fig. 10).

inverse of the SNR of the mode [227]. It is therefore important to verify that the scale of these errors is small. In this Subsection, we will compare the outputs of our amplitude module with an independent dataset to quantify the impact of errors in our interpolated amplitudes. We will also explore where our choice to perform linear interpolation with respect to a (as described in Sec. III B 2) limits the accuracy of our mode amplitude values when compared to spline interpolation along this dimension.

As in Sec. IV A, we will validate our model by comparing interpolated mode amplitudes ($\mathcal{A}_{\ell mn}^{\text{FEW}}$) against those produced with the GREMLIN code ($\mathcal{A}_{\ell mn}^{\text{SAH}}$). As our amplitude module produces 6993 modes at each point in parameter space, examining any particular mode in isolation will not be particularly informative with regard to the accuracy of our amplitude model as a whole (especially as the relative importance of each mode is also a function of the parameter space). Instead, in a similar vein to Ref. [122], we will define a suitable figure of merit that encapsulates all mode amplitudes. It is similar in form to that of the waveform mismatch (which we will apply in Sec. IV C; see Appendix D for details), and is written as

$$\mathcal{M}_{\text{amp}} = \frac{\sum_{\ell mn} (\mathcal{A}_{\ell mn}^{\text{FEW}})^* \mathcal{A}_{\ell mn}^{\text{SAH}}}{|\mathcal{A}_{\ell mn}^{\text{FEW}}| |\mathcal{A}_{\ell mn}^{\text{SAH}}|}, \quad (14)$$

where $\star$ denotes complex conjugation and $|\cdot|$ is the amplitude vector norm (over the indices ℓ , m and n). This quantity corresponds to a mismatch between two snapshot waveforms (i.e., constant in amplitude and frequency) with amplitudes $\mathcal{A}_{\ell mn}^{\text{FEW}}$ and $\mathcal{A}_{\ell mn}^{\text{SAH}}$, assuming that

overlap terms between modes are small. While this is not generally true, the inclusion of these terms is not required for us to explore how the accuracy of our modes varies over parameter space, and their exclusion enables us to avoid performing waveform computations at this stage. This also ignores the spin-weighted spherical harmonics, as we wish to investigate differences independent of the viewing angle. The $\mathcal{A}_{\ell mn}^{\text{SAH}}$ data we compare against only extends to $p \sim 10$; see Appendix E 2 for a comparison between our amplitude interpolation and PN results in the weak-field limit. In Fig. 9, we show how $\mathcal{M}_{\text{amp}}$ varies over (p, e) for $a = 0.8952$ (which lies between two nodes in our amplitude data grids, maximizing interpolation errors). For $p \gtrsim p_{\text{sep}} + 1$, our interpolated amplitudes agree well with the GREMLIN data for all eccentricities, with estimated mismatches $\leq 10^{-5}$. Below this value, $\mathcal{M}_{\text{amp}}$ grows rapidly with respect to p , reaching $\sim 10^{-3}$ at higher eccentricities for $p \gtrsim p_{\text{sep}} + 10^{-2}$. This suggests that our amplitude interpolation errors are significant near the separatrix. In Sec. IV C, we investigate the impact of amplitude interpolation on the waveform level and obtain results that agree quantitatively with those presented in Fig. 9.

Recall that, in our amplitude interpolation framework (Sec. III B), we opt to perform linear interpolation with respect to a for the purposes of efficiency and to reduce memory usage. This is motivated by the weak relationship between mode amplitude and spin (given constant p and e). However, our grid coordinate parametrization (which maps the curved inner surface of both the separatrix and the eccentricity taper to rectilinear coordinates) is highly non-linear near the separatrix, especially as $a \rightarrow 1$. This effect is the main contributor to the larger amplitude mismatches we observe in Fig. 9. To demonstrate this point and explore the impact of such behavior on linear interpolation accuracy near the separatrix, we construct tricubic spline interpolants (identical in form to those applied to the forcing functions in Sec. III A 1) and compare their outputs to those of the bicubic + linear scheme. For regions where the linear interpolation is a good approximation, we expect the outputs of these two interpolation methods to be in strong agreement. We consider the $\mathcal{A}_{220}$ mode, which is typically the most significant mode in terms of SNR contribution for most of the inspiral: we found that comparisons performed for other strongly contributing modes yielded similar results to this case. Relative differences between the (complex) outputs of the bicubic + linear and tricubic interpolation configurations for this mode are shown in Fig. 10. We find that (except for regions of the parameter space where $|\mathcal{A}_{220}| \rightarrow 0$) a linear interpolation in a is sufficient for mode amplitudes to be computed accurately (better than three digits of precision) for $p - p_{\text{sep}} \gtrsim 1$. Larger relative differences occur where $|\mathcal{A}_{220}| \sim 0$, but as the amplitude is small here regardless of the interpolant used we do not expect this to significantly impact waveform accuracy. For lower values of p and $a \gtrsim 0.9$, the

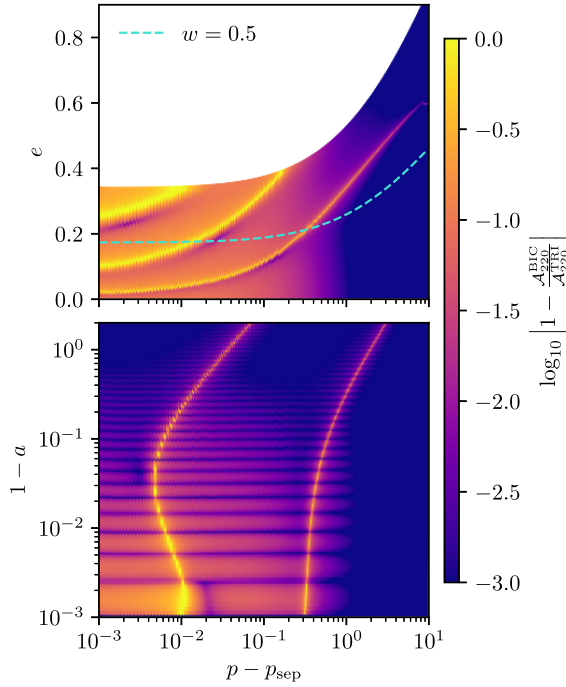


FIG. 10. Relative difference between bicubic + linear and tricubic interpolation of the A_{220} mode amplitude with respect to: (p, e) with $a = 0.998$ (top panel); (p, a) with grid coordinate $w = 0.5$ (bottom panel). The eccentricity corresponding to this value of w for $a = 0.998$ is shown in the top panel as a cyan dashed line; for lower spins, e at $p = p_{\text{sep}} + 10^{-3}$ will be higher than this line (see Appendix B). The outputs of the two interpolants are in close agreement (better than *three* decimal places), with the exception of the strongest-field region and spins greater than 0.9 (regardless of eccentricity). Note that the real and imaginary components of A_{220} are interpolated separately; here we have plotted the relative difference between the complex amplitudes for visual clarity, as the components pass through zero more frequently.

relative difference grows, reaching a few tens of percent at the highest spins and lowest orbital separations. However, it is reasonable to expect that the impact of these errors on waveform accuracy will be subdominant with respect to errors of similar scale at larger orbital separations, as EMRIs typically spend only a small fraction of their number of orbital cycles in this region.

For systems with particularly small mass ratios $\epsilon \lesssim \sim 10^{-6}$, which inspiral sufficiently slowly to accumulate hundreds of orbital cycles in the region $p - p_{\text{sep}} < 1$, the impact of these amplitude interpolation errors is likely to be more significant. As the primary focus of our waveform model is mass ratios $\epsilon \gtrsim 10^{-6}$, this is an acceptable limitation of our model (but we will examine a candidate source with $\epsilon = 10^{-6}$ to understand this in more detail). In Sec. IV C, we will demonstrate that this reasoning holds for such systems by performing comparisons between amplitude models on the waveform level.

C. Waveform validation

Now that we have characterized the dominant sources of error in our trajectory and mode amplitude generation frameworks, we will explore how these errors manifest in waveform templates produced with our model. The gold standard for understanding the impact of systematic errors is to examine them through the lens of Bayesian parameter estimation, identifying how these errors propagate to the structure of the posterior distribution [Eq. (D9)]. However, this procedure requires millions of likelihood evaluations (and therefore hours of GPU wall time) per comparison, and is therefore prohibitively computationally expensive for a global examination of the parameter space. Techniques for approximating linear biases in parameter estimates [236–238] (such as the approach we apply in Sec. V C for this purpose) reduce this to $\sim 10^2$ waveform evaluations per point in parameter space (which are required for computing waveform derivatives). However, the resulting bias vectors in 13-dimensional parameter space are challenging to interpret in the context of a random draw of sample points over the EMRI parameter space.

We instead opt to quantify the fidelity of our model in terms of the waveform mismatch $\mathcal{M}$, which is defined in terms of the noise-weighted inner product and therefore incorporates the sensitivity curve of the detector (in this case, LISA). Definitions for the inner product and mismatch are given in Appendix D. A mismatch of zero implies that two templates are identical, whereas a mismatch of one indicates that the two waveforms are orthogonal. The accuracy requirements for a waveform model to be applied in data analysis without introducing bias must necessarily be expressed relative to the SNR ρ of the signal being analyzed (with louder signals requiring more faithful models—see [238] for more discussion). A useful (albeit typically strongly conservative) criterion derived in Refs. [227, 239–241] relates the SNR of the template signal s (ρ_s) defined in Eq. (D8) to the minimum mismatch between this signal and a waveform template h such that the two are indistinguishable. It states that if

$$\mathcal{M}_{\min} := \mathcal{M}(s, h) \leq \frac{D}{2\rho_s^2}, \quad (15)$$

with $D = 12$ being the number of parameters in the waveform model that are estimated during inference (cf. Table I), then s and h are identical for the purposes of data analysis (at a confidence of 1σ). Choosing what constitutes an acceptable value of $\mathcal{M}_{\min}$ for our model therefore requires us to specify ρ_s . For EMRIs with $\epsilon \lesssim 10^{-4}$ situated at astrophysical distances ($d_L \gtrsim 1$ Gpc), LISA accumulates an SNR of up to $\sim 10^3$ over a few years of observation, given source parameters that maximize detectability (such as high MBH spin). See Table I for some example systems, which we will study in more detail in

TABLE I. Parameters for investigated EMRI and IMRI sources. Waveforms for these systems are shown in Fig. 19. (From left to right columns): Detector-frame masses m_1 and m_2 , dimensionless MBH spin a , initial semilatus rectum p_0 , initial (final) eccentricity e_0 (e_f), luminosity distance d_L , redshift z_r , and the optimal SNR ρ . Each source is observed for two years and plunges just after the end of the observation window. For sources with $m_1 = 10^5 M_\odot$, we use a finer sample cadence $dt = 2$ s to resolve high-frequency GWs near plunge; all other sources use our fiducial sampling interval of $dt = 5$ s. In all cases, we fix the angular parameters $(\theta_S, \phi_S, \theta_K, \phi_K) = (0.8, 2.2, 1.6, 1.2)$ and initial orbital phases $(\Phi_{\phi_0}, \Phi_{r_0}) = (2.0, 3.0)$.

	Summary	$m_1(M_\odot)$	$m_2(M_\odot)$	a	p_0	$e_0(e_f)$	d_L (Gpc)	z_r	ρ
1	EMRI (Prograde)	10^6	10^1	0.998	7.728	0.730(0.045)	2.204	0.394	50
2	Strong-field EMRI	10^7	10^1	0.998	2.120	0.425(0.261)	3.590	0.593	30
3	Heavy IMRI	10^7	10^5	0.950	23.425	0.850(0.023)	7.250	1.058	500
4	Light IMRI	10^5	10^3	0.950	74.383	0.850(0.004)	3.500	0.581	200
5	EMRI (Retrograde)	10^5	10^1	-0.500	26.190	0.800(0.195)	1.081	0.212	30

Sec. V F. Informed by this order-of-magnitude estimate, we choose the value $\mathcal{M}_{\min} = 10^{-5}$; this is compatible with signal SNRs of several hundreds and [considering the conservative nature of Eq. (15)] should represent sufficient accuracy for the purposes of EMRI data analysis with LISA. In what follows, we stress that we do not maximize our mismatch calculations over phase, coalescence time, or other extrinsic parameters (such as source orientation); as this would always improve the mismatches we would observe, our results are therefore slightly conservative. However, the impact of this choice is small for EMRIs, as these mismatches are the result of small differences between many sets of harmonic modes that (in most cases) cannot all be significantly reduced by modifying such parameters.

1. Trajectory dephasing

Based on the results presented in Secs. IV A and IV B, the most significant source of systematic error in the constituent components of our waveform model (i.e., the trajectory and mode amplitude modules) is orbital dephasing due to interpolation errors in the forcing functions. We will now quantify how this dephasing propagates to the accuracy of the computed waveforms, following a similar approach as was applied in Sec. IV A but with a focus on waveform mismatch (rather than orbital dephasing). Randomly drawing 10^4 points in parameter space (obtained according to the procedure detailed in Appendix C), we compute waveforms with both our (full-resolution) model and lower-resolution models with down-sampled forcing function data grids. In Fig. 11, we summarize the mismatches between the full-resolution and lower-resolution models at these points in parameter space. As in Fig. 8, a clear trend is observed, with mismatches decreasing in proportion to grid resolution. Fitting this exponential trend between grid resolution and mismatch for each data point (similarly to in Sec. IV A) and extrapolating this fit to our full grid density yields the red dashed histogram in Fig. 11. Due to the poor phase accuracy of the lowest-resolution model (which regularly exhibits in excess of one radian of orbital dephasing) we do not include

it when constructing this extrapolation. This should be considered a rough, order-of-magnitude estimate of the accuracy of our full-resolution model; it is not completely representative of the mismatch between our model and an error-free adiabatic model (especially for the smaller mismatches $\lesssim 10^{-8}$, as we expect other minor sources of systematic error to become important at this level). In general, we observe mismatches between waveforms from the full- and half-resolution models (h^{1DS} and h^{2DS} respectively) of less than 10^{-5} over the majority of the parameter space, with extrapolated mismatches improving upon this by roughly two orders of magnitude. When compared to our mismatch requirement of 10^{-5} , these results suggest that our waveform model is highly robust over a large fraction of the EMRI parameter space.

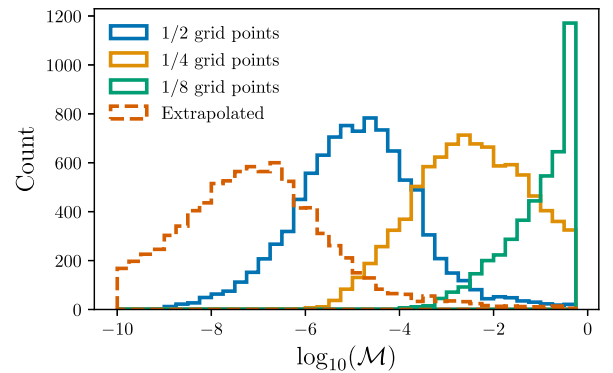


FIG. 11. Similar to the top panel of Fig. 8, but in terms of waveform mismatches instead of orbital dephasings. We consider four waveform models built with forcing function data grids of successively coarser resolutions. All mismatches are computed with respect to waveforms generated with the finest-resolution model. The majority of mismatches decrease exponentially (in proportion to the data grid density); the red dashed histogram is obtained by extrapolating this exponential scaling to the actual resolution of our data grid. A long upper tail of mismatches is evident, which corresponds to systems with high eccentricities (see Fig. 12 and main text for discussion).

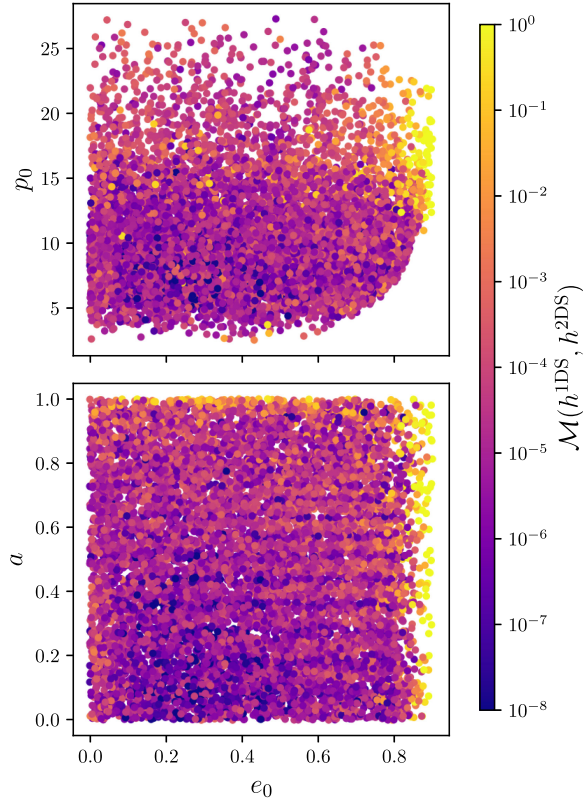


FIG. 12. Mismatches between waveforms computed with models using full-resolution ($h^{1\text{DS}}$) and half-resolution ($h^{2\text{DS}}$) forcing function data grids (blue histogram in Fig. 11), with respect to (e_0, p_0) (top panel) and (e_0, a) (bottom panel). The mismatches shown can be interpreted as upper bounds on the accuracy of our full-resolution model at each point in parameter space. The accuracy of our model decreases significantly for $e_0 \gtrsim 0.85$.

We also find a small tail of higher mismatches $\mathcal{M} \rightarrow 1$ that are easily identified when $\mathcal{M}$ is examined as a function of (a, p, e) , as shown in Fig. 12. Mismatch increases rapidly with respect to eccentricity beyond $e \sim 0.85$, which is not unexpected given the larger errors in the forcing functions in this region identified in Fig. 5. It is not clear whether the relatively poor performance of our model in this region is the result of data quality issues or insufficient interpolation density (or more likely, some combination of the two); regardless of the cause, these results suggest that users of our waveform model should exercise caution when choosing $e_0 \gtrsim 0.85$. Some poorer mismatches are also observed for very high spins ($a \gtrsim 0.998$), which we attribute to the impact of insufficient grid density with respect to a nearer to the separatrix (as was identified in Fig. 6). Both of these problematic regions are near the edges of our domain of validity, where the forcing functions vary rapidly, are computationally expensive, and are challenging to obtain accurately. Given the resulting challenge in modeling systems with such high eccentricities and/or spins, treating these areas of the parameter space separately (e.g., by

constructing separate data grids bespoke to these regions) is an idea worthy of investigation in future work.

2. Amplitude interpolation errors

As discussed in Sec. IV B, we do not expect that systematic errors in our mode amplitude interpolation will significantly impact the accuracy of our waveform model. We will now test the validity of this assertion by comparing waveforms generated with two models that differ only in the generation of their amplitudes at each point output by the ODE solver: one uses amplitudes computed with our bicubic + linear interpolant (Sec. III B 2), whereas the other directly solves the Teukolsky equation to obtain the mode amplitudes (Sec. III B 1) using the GREMLIN code. In what follows, we will refer to waveforms from these models as “FEW” and “SAH” respectively. As the latter model is orders of magnitude more computationally expensive (despite amplitudes only being computed at ~ 100 points along the trajectory), we will consider three EMRI systems in lieu of a more comprehensive study over the EMRI parameter space.

The first system has parameters $\{m_1, \epsilon, a, p_0, e_0\} = \{10^6 M_\odot, 10^{-5}, 0.998, 7.81, 0.7\}$ (with p_0 chosen such that the CO plunges after two years), and is representative of the region of our parameter space for which our mode amplitude interpolant is least accurate (i.e., high spin and eccentricity) for inspirals of this mass ratio and duration. From Fig. 10, we do not expect weak-field amplitude interpolation errors to contribute significantly to waveform mismatch, so our choice of a two-year inspiral (over e.g., a four-year inspiral) will have little bearing on the results of this comparison. In Fig. 13, we show the cumulative mismatch (i.e., the mismatch computed over increasingly large intervals of time) between the corresponding waveforms from these two models as a function of both t and p . Also shown in Fig. 13 is the time-domain waveform strain near the beginning and end of inspiral, with differences in mode amplitudes visible by eye in the latter case. At earlier times (in the weaker field), $\mathcal{M} \sim 10^{-8}$; this rises to $\mathcal{M} \sim 10^{-5}$ as $p \rightarrow p_{\text{sep}}$. This behavior is expected based on the poorer performance of our interpolant close to the separatrix at high spins (Fig. 10). The majority of the mismatch accumulates in the interval $p - p_{\text{sep}} \in [1, 0.1]$, after which the inspiral completes the few remaining orbital cycles and the cumulative mismatch levels off. To confirm that the primary cause of waveform inaccuracy near the separatrix is a result of our interpolation scheme (rather than issues of data quality), we also construct the mode amplitudes for this system with a tricubic interpolant and compare the resulting waveform with that of the “SAH” model. The cumulative mismatch we obtain (shown in the bottom panel of Fig. 13 in red) is lower than that obtained with our fiducial interpolation scheme by more than an order of magnitude, confirming

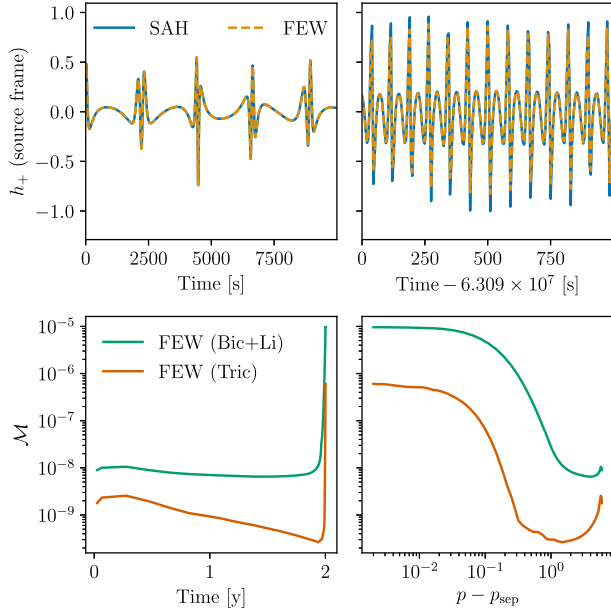


FIG. 13. Comparisons between waveforms constructed with either interpolated (“FEW”) or directly computed amplitudes (“SAH”) at each sparse trajectory point. This EMRI has parameters $(a, p_0, e_0) = (0.998, 7.81, 0.7)$. Top panels: Time-domain strains at early and late times for waveforms built with exact Teukolsky mode amplitudes (blue line) and our bicubic + linear interpolation (yellow dashed line). Amplitude differences are evident near the end of inspiral. Bottom panels: Cumulative mismatches with respect to either time or semilatus rectum between “FEW” and “SAH” waveforms. Here we examine both our fiducial bicubic + linear interpolation scheme (green line) and a tricubic interpolation scheme (red line). Mismatch increases as the inspiral approaches the separatrix; the larger mismatch in the green line suggests that this is a result of the linear interpolation in a becoming increasingly inaccurate near the separatrix.

that linear interpolation with respect to a is indeed the limiting factor for the accuracy of our amplitude module near the separatrix.

While this sharp increase in mismatch may be alarming, it is important to note that the parameters of this EMRI were chosen to probe the impact of the largest amplitude interpolation errors our model exhibits (given this mass ratio and duration). Additionally, as previously highlighted in Sec. II A, the multiscale expansion underpinning our model begins to break down in these final stages of inspiral; this region may therefore be more accurately captured by a transition-to-plunge framework [107]. As mismatch only accumulates significantly near the separatrix, the deviation should be larger for more extreme mass ratios. To confirm this, we perform a similar analysis for an EMRI of mass ratio 10^{-6} , with parameters given by the second row of Table I. Note that $p \lesssim p_{\text{sep}} + 0.5$ for this entire inspiral, so the two-year duration of this observation uses wave amplitudes from the region of parameter space where our

amplitude interpolation is least accurate. We do not show a visualization of the results of this comparison for brevity, but observe a mismatch of 10^{-3} that steadily grows by a factor of ~ 2 over the course of the inspiral. As expected, this is significantly larger than what was observed for the $\epsilon = 10^{-5}$ system. This highlights the inherent difficulty in modeling these EMRIs that spend many years in the neighborhood of the separatrix; as self-force data varies rapidly in this region, it presents a significant challenge for interpolation schemes. However, for LISA, these sources will not have particularly high SNRs at typical astrophysical distances—the SNR of this source is 30—meaning that [according to Eq. (15)] even given the poorer performance of our model in this region of the parameter space, we still do not expect to observe significant systematic biases during parameter estimation as a result. While good performance over the entire parameter space is obviously desirable, there is also scope in future work to optimize waveform models based on the expected SNRs of astrophysical sources, which may improve the efficiency of these models without introducing significant biases into the results of data analysis.

To demonstrate the accuracy of our amplitude model away from the edges of our domain validity, we also consider a third EMRI with arbitrarily chosen parameters $\{m_1, \epsilon, a, p_0, e_0\} = \{10^5 M_\odot, 10^{-4}, 0.5, 27.86, 0.4\}$. We find that the waveforms from each model are in extremely close agreement: the mismatch between them remains at $\sim 10^{-10}$ throughout the entire inspiral. This confirms that for inspirals with more typical spins and/or eccentricities, the amplitudes produced by our model are sufficiently accurate so as not to limit waveform accuracy (compared to other sources of error, such as forcing function interpolation).

Based on the results of these comparisons, we are confident that our waveform mode amplitude framework is highly robust over the majority of the parameter space. Future development efforts in this area should be focused on the accurate interpolation of mode amplitudes near the separatrix, as well as the investigation of how close to the separatrix our multiscale framework must extend to before the switch to a transitional framework is necessary. These efforts will improve the accuracy of our framework for systems with more extreme mass ratios ($\epsilon \lesssim 10^{-6}$).

V. RESULTS AND DISCUSSION

In previous Sections, we have described the implementation of our waveform model and demonstrated that it is robust over its domain of validity. Here, we apply our model to explore scientific and data analysis prospects for asymmetric-mass binary observations with LISA. We first characterize the computational cost of our waveform model (Sec. V A), which is closely tied to the feasibility of data analysis and parameter estimation. We then examine the

impact of semirelativistic models for waveform mode amplitudes (which have been applied extensively in the literature) on the SNR of EMRI sources (Section V B), and explore how these approximate models can induce biases during parameter recovery in Sec. V C. Finding that SNRs computed with approximate models are subject to significant errors, in Section V D we apply our model to the exploration of the sky-averaged detection horizon for eccentric and rapidly spinning EMRI and IMRI systems, examining its dependence on the component masses of the system. We then examine the detectability of small eccentricities for a representative, rapidly spinning EMRI system in Sec. V E, which is an area of great astrophysical interest with respect to the formation of EMRIs in gas-dominated environments such as accretion disks. We conclude this Section with an exploration of parameter recovery for EMRI and IMRI sources (with parameters listed in Table I) in the Bayesian inference context in Sec. V F, highlighting similarities and differences in the results obtained between these two cases. Unless otherwise stated, our analysis procedures throughout this Section follow the conventions outlined in Appendix D.

A. Waveform computational cost analysis

The principal objective of the FEW package is to provide tools for the rapid generation of adiabatic waveforms by leveraging vectorized operations on GPU hardware. In this Subsection, we will examine the computational cost of waveform generation in both the time domain (TD) and the frequency domain (FD). The procedures for constructing waveforms in these two domains differ only at the stage of the waveform summation modules, which have not changed significantly since the TD (Refs. [171,242]) and FD (Ref. [173]) releases of the package. The differences in timing with respect to these previous works can therefore be attributed to: the expansion of the parameter-space coverage of the model, particularly the inclusion of spin and the evolution of the inspiral closer to the separatrix; and changes made to the trajectory and amplitude generation modules. We warn the reader that the computational cost of the waveform can vary depending on the computing resources used; all computational wall times reported in this Section were obtained using an NVIDIA A100 GPU and a 2 GHz AMD EPYC-7713 CPU. Furthermore, the timing distribution will depend on our procedure for sampling source parameters—this is described in Appendix C.

We will examine the cost of our model for four-year inspirals with sampling cadence $dt = 5$ s. This sampling cadence is based roughly on what is required for systems with $M \sim 10^6 M_\odot$; the computational cost for other choices of cadence can be roughly estimated by scaling the wall time with respect to this value of dt . We consider two mode selection thresholds, $\kappa = (10^{-2}, 10^{-5})$, to probe how this cost depends on the mode content requested in waveform generation. Drawing 10^4 sets of parameters, we show the

corresponding distribution of computational wall times per waveform evaluation in Fig. 14. We obtain median wall times of 0.11 s (0.13 s) per TD evaluation and 0.13 s (0.15 s) per FD evaluation, for $\kappa = 10^{-2}$ (10^{-5}). The FD wall time distribution presents a long upper tail that extends to ~ 1 s; this feature corresponds to primary masses $m_1 \lesssim 10^6 M_\odot$ with $e_0 \gtrsim 0.7$. The reason for this discrepancy in timings between output domains is as follows (based on arguments presented in Ref. [173]). The computational cost of the waveform scales with the number of data points at which each waveform mode must be evaluated. For TD generation, each array element receives contributions from all the selected harmonics (such that the overall summation cost scales linearly with the number of waveform modes). In contrast, the cost per mode of the FD generation depends on the frequency evolution of that mode, as each mode is only evaluated in frequency bins it evolves through. Therefore, the FD waveform is most efficient when the modes of the signal span a compact range of frequencies, which corresponds to higher primary mass values; conversely, at lower primary masses and higher eccentricities many modes span a wide range of frequencies, leading to the tail observed in Fig. 14. In terms of performance with respect to eccentricity at higher primary masses, it is argued in [173] that FD waveform generation is more efficient than its TD counterpart for $e_0 \sim 0.7$ when $m_1 \sim 10^7$, and this remains true in our results. It is worth noting that frequency-domain waveforms generated via the stationary phase approximation can be down-sampled to reduce computational cost, potentially improving efficiency for certain analyses [173]. However, such down-sampling reduces the effective signal-to-noise ratio, leading to biased parameter posteriors in noisy data, limiting its applicability for realistic EMRI inference.

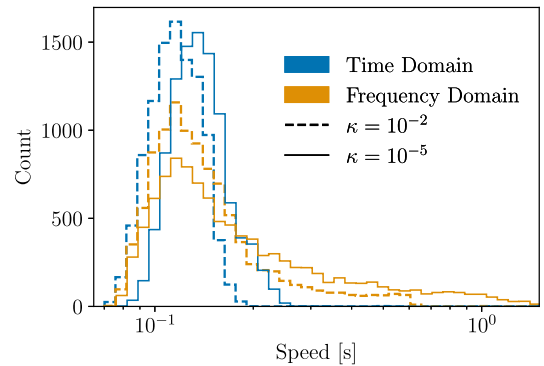


FIG. 14. Wall time of waveform generation with FEW in either the TD (blue) or FD (orange) output domains, for 10^4 randomly sampled sets of source parameters. We show wall times for $\kappa = 10^{-5}$ (dashed line) and $\kappa = 10^{-2}$ (solid line) to explore the impact of waveform mode content on computational costs. The sampling interval is chosen to be $dt = 5$ s, with all inspirals plunging after four years.

The inclusion of primary spin increases the azimuthal frequency $\hat{\Omega}_\phi$ close to the separatrix by up to a factor of six (for $a = 0.999$, $e = 0$), which in turn leads to many modes evolving over a much larger range of frequencies in a typical waveform than for the zero-spin case. This causes the FD waveform generation to exceed that of TD generation for the majority of spins $a > 0$ (which is why the lower ends of the FD timing distributions do not extend below their TD counterparts). We remind the reader that Fig. 14 does not include retrograde orbits ($a < 0$), where we expect the opposite behavior to occur and the relative cost of FD waveform generation to decrease.

We also examine the breakdown in wall time between different waveform generation modules for the two output domains. For a four-year inspiral with source parameters $\{m_1, e, a, e_0\} = \{10^6 M_\odot, 10^{-5}, 0.9, 0.1\}$, we measure wall times of 0.10 s for the trajectory module, 0.02 s for the amplitude module, and 0.15 s (0.60 s) for the TD (FD) summation module (for $\kappa = 10^{-5}$). The trajectory and summation modules therefore constitute the bulk of waveform generation computational cost.

To verify that the TD and FD waveforms are consistent with each other, we conduct a mismatch analysis over the 10^4 sets of drawn parameters. We find the median mismatch between the two to be $\sim 10^{-3}$, suggesting that the waveforms are in good agreement. We note that computing mismatches across output domains can be misleading due to the windowing/leakage effects, and that different window choices can significantly impact the mismatches obtained. This was explored in more detail in Ref. [173]; note that in that work, a Hann window was used for these comparisons, whereas here we instead adopt a Tukey window with shape parameter 0.005. Our choice of window does not affect any of the conclusions drawn in this Subsection.

B. Impact of relativistic amplitudes on waveform SNR

Semirelativistic waveform models known as “kludges” [243–245] have been commonly used tools to make qualitative statements describing detection rates, to perform large-scale parameter measurement studies, and to examine the impact of stochastic backgrounds composed of unresolved EMRI signals on LISA data analysis [28,42,246–248]. These studies perform SNR calculations with kludge models that only incorporate a quadrupolar approximation (i.e., $\ell = 2 = m$) to the mode structure of the waveform. As was identified in Refs. [48,136], these approximate quadrupolar amplitudes may be insufficiently accurate to adequately describe the EMRI mode spectrum, particularly in the strong-field regime of gravity accessible to eccentric inspirals into rapidly spinning black holes. As we have already shown in Fig. 4, a large number of (ℓ, m, n) modes are required to faithfully reconstruct the true EMRI signal. In this Subsection, we will investigate through SNR comparisons (which are critical for

determining detection rates) the impact of approximating the mode amplitudes of adiabatic EMRI waveforms with those of the augmented analytic kludge (AAK), a kludge model that has seen extensive application in the literature. Understanding the regions of parameter space where the SNR is inaccurately estimated, and to what extent, is crucial for avoiding misrepresentations of the detection rate of EMRIs.

Exploiting the modular nature of FEW, we combine our relativistic trajectory model (Sec. III A) with two different amplitude models. This results in two waveform models: the first uses our adiabatic mode amplitudes (Sec. III B) and is the model we introduce in this work; the second uses the AAK mode amplitudes. We will refer to these models as KERR and AAK, respectively, throughout this subsection. The KERR amplitudes, described in Sec. III B, consider the harmonics (ℓ, m, n) for $2 \leq \ell \leq 10$, $|m| \leq \ell$, and $|n| \leq 55$. For the AAK model, which is quadrupolar [195] (i.e., $\ell = 2 = m$), we set the number of Fourier modes to $n_{\max} = 50$; note that these modes do not have a one-to-one correspondence with those of our model (see Ref. [245] for definitions). We confirmed that this value of $n_{\max}$ is sufficiently high to represent the AAK for the purposes of this analysis (i.e., our results are convergent with respect to $n_{\max}$). For the KERR model, we use the default mode-selection threshold $\kappa = 10^{-5}$; by construction, the SNRs computed with the KERR model for $\kappa = 0$ and $\kappa = 10^{-5}$ will differ by $\sim 10^{-5}$, which is negligible on the scale of the SNR variations we identify in our results.

For each set of source parameters, we choose p_0 such that inspirals plunge after four years and fix extrinsic parameters (such as orientation angles and initial phases) to those given in the caption of Table I. Our results are largely independent of this choice of extrinsic parameters.

In Fig. 15, we plot the SNR ρ as a function of e_0 and a for EMRIs with $(m_1, m_2) = (10^6, 10)M_\odot$. The solid white line indicates the boundary where the two waveform models provide a consistent prediction of the signal-to-noise ratio $\rho_{\text{AAK}}/\rho_{\text{KERR}} = 1$. Above this line (at higher spins), the AAK overestimates the SNR by up to $\sim 30\%$, whereas below it underestimates it by up to $\sim 20\%$. This discrepancy between the models increases with eccentricity, particularly at lower values of a . Two main factors contribute to the differences observed. First, as the AAK is quadrupolar, it will become increasingly inaccurate as the orbital separation decreases and missing modes with $\ell > 2$ become more significant. Second, as the mode amplitudes of the AAK are based upon a low-velocity approximation [249], they will also become increasingly inaccurate as the orbital separation decreases. What we observe in Fig. 15 is a combination of these two effects.

The results of such a comparison will change depending on the masses of the system, due to the shape of the LISA power spectral density (PSD) and the change in p_0 required to hold the inspiral duration fixed. To investigate this

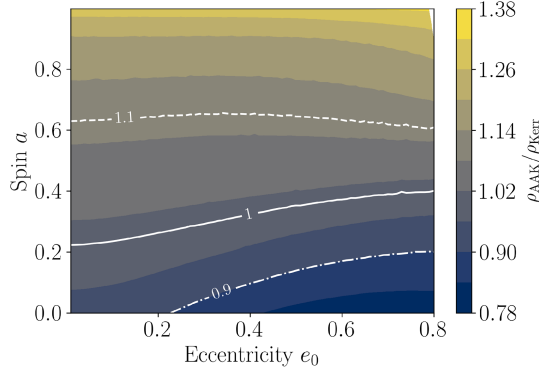


FIG. 15. Ratio of the optimal SNRs computed with the semi-relativistic AAK mode amplitude model (ρ_{AAK}) and the adiabatic model presented in this work (ρ_{Kerr}), as a function of initial eccentricity $e_0 \in [0, 0.7]$ and primary spin $a \in [0, 0.998]$. For all systems, we consider four-year inspirals with masses $(m_1, m_2) = (10^6, 10)M_\odot$. Waveforms for both models are built with the same adiabatic trajectory model (Sec. III A). The (dashed, solid, dot-dashed) white lines indicate the ratio of the SNRs for three reference values $\rho_{AAK}/\rho_{Kerr} \in \{1.1, 1, 0.9\}$.

behavior, we also perform this analysis for other values of m_1 and m_2 (keeping the mass ratio $\epsilon = 10^{-5}$ fixed), but do not show the corresponding figures for brevity. For $m_1 = 10^5$, the faster rate of inspiral requires us to set p_0 higher (for a two-year inspiral) than if we were to have larger primary masses. As most of the inspiral takes place in this weaker-field region (where the assumptions of the AAK fare better), and the strong-field emission is shifted to higher frequencies due to the lower total mass (where LISA is less sensitive), we find only moderate deviations in the SNR of $\sim 9\%$ at most. Conversely, discrepancies between the models become more pronounced for $m_1 > 10^6$, with a lower p_0 leading to inspirals in the stronger-field regime where the amplitudes of the AAK fare worse. Indeed, for $m_1 = 10^7 M_\odot$, we see an overestimation of the SNR of $\sim 60\%$ for higher a and lower e_0 , and an underestimation of $\sim 40\%$ for lower a and higher e_0 , deviations roughly twice as large as in the $m_1 = 10^6 M_\odot$ case. For all three cases, the differences between the SNRs computed with each model increased for larger eccentricities, with a larger overall difference for larger values of m_1 .

From these results, we can conclude that one should be careful when employing the AAK (or other similar kludge waveform models) in prospective studies of LISA’s scientific capabilities with EMRIs. This conclusion is qualitatively similar to that of Ref. [48], which performed similar investigations in the quasicircular case. Our work reinforces this conclusion in the presence of eccentricity, and identifies that the AAK typically fares worse for eccentric systems than for quasicircular ones. As mentioned earlier in this Subsection, the systematic errors in the SNRs computed with such models will impact the quantitative conclusions of

EMRI rate estimation studies. In Ref. [28], detection rates and catalogs (that have since been applied extensively in the literature) were obtained using a kludge-based waveform model [30]. A similar analysis conducted using relativistic waveform models, such as ours, would likely yield lower detection rates, since many of these catalogs consisted solely of systems containing rapidly rotating MBHs with $a \in (0.9, 0.998]$, and our results indicates that kludge models significantly overestimate the SNRs of these systems. A more detailed study characterizing these systematics will be explored in future work.

C. Parameter recovery with approximate mode amplitude models

In addition to inaccurate SNR calculations, approximate waveform models can also induce systematic biases in the inference of EMRI parameters, e.g., due to missing physics [35,250] or insufficient model accuracy [171,238,245]. In this Subsection, we explore the latter by studying the impact of incomplete or approximate waveform mode amplitude models on biases in parameter estimation. We consider our fully relativistic waveform model constructed in Sec. III (and with the default mode-selection criterion as described in Sec. III C) as the representative “true” model, and denote it as KERR. Similarly to Section V B, we consider two approximate models with different mode amplitudes but the same relativistic trajectory as this model (cf. Sec. III A). Our first approximate model, KERR $\ell 2$, is composed of only the dominant $\ell = 2$ modes (but accounts for all corresponding m, n modes, where $|m| \leq \ell$ and $|n| \leq n_{\max} = 55$; cf. Sec. III B) while the second model, the AAK, uses semirelativistic (“kludge”) amplitudes formulated in Ref. [245], which are based on the quadrupole moment formalism (cf. [195]). Unlike Section V B, we do not control the mode content of the AAK waveform model. The modes are chosen instead according to the default mode-selection criteria in FEW, i.e., $n_{\max} = \max(4, \lfloor 30e_0 \rfloor)$. We then infer the injected KERR signal with these approximate models and assess their impact on parameter recovery.

Our analysis is set up as follows. We consider signals in a 2D grid of parameter points on the surface defined by (a, e_0) . We select $N = 10$ points each along prograde inspirals $a \in [0.1, 0.9]$ and $e_0 \in [0.1, 0.5]$ for a total of 100 gridpoints. We choose only prograde inspirals ($a > 0.0$) because the effect of higher multipole modes of $A_{\ell mn}$ is less pronounced in retrograde inspirals; The mode amplitude $A_{\ell mn}$ of a given ℓ mode is $O(p^{-\ell/2})$ (or higher) for equatorial inspirals [122,251,252], and the retrograde trajectory is completed within the weak field $p \gg 1$ in the majority of cases. We also restrict the analysis to moderate eccentricity values for computational feasibility, as waveforms with $e_0 > 0.5$ are significantly more expensive due to their larger mode content.

We choose three detector-frame mass combinations, with $m_1 = \{0.5, 1.0, 1.5\} \times 10^6 M_\odot$ paired with $m_2 = \{5.0, 10.0, 15.0\} M_\odot$, respectively, to gauge the impact of approximate models across different detector frequencies. All other parameters are fixed to the following values: the sky location parameters $(\theta_S, \phi_S) = (\pi/5, \pi/6)$, the spin orientation parameters $(\theta_K, \phi_K) = (\pi/3, \pi/4)$, and the initial azimuthal and radial phases $\Phi_{\phi_0} = 0.0 = \Phi_{r_0}$. The initial semilatus rectum p_0 is fixed to $p_{\text{plunge}} + 0.5$ where p_{plunge} is the initial semilatus rectum value that leads to plunge after one year of inspiral. Finally, the luminosity distance d_L is fixed such that each source has an SNR of exactly 100 when measured using the true waveform. Note that this distance rescaling introduces a gradient in the source-frame masses $(m_1^{(s)}, m_2^{(s)})$ across the grid, which are related to the redshifted (detector-frame) masses as $(m_1^{(s)}, m_2^{(s)}) = (m_1, m_2)/(1 + z_r)$ where z_r is the redshift which varies with d_L . However, this does not influence our analysis, since we exclusively work in terms of detector-frame masses. We perform the analysis in the log-mass parametrization and vary all model parameters such that the parameter vector is $\theta := (\ln(m_1), \ln(m_2), a, p_0, e_0, d_L, \theta_S, \phi_S, \theta_K, \phi_K, \Phi_{\phi_0}, \Phi_{r_0})$ with dimension $D = 12$.

At each gridpoint, inference of the true signal with the approximate model generically introduces systematic biases in the best-fit parameter estimate that can be approximated in the high-SNR limit by the linear signal approximation given by Cutler and Vallisneri [238] [Eq. (D10)]. To quantify these biases, we calculate the sigma-contour level between the injected and the best-fit points, given by the Mahalanobis distance [253], notated as D_{Maha} and defined explicitly in Eq. (D13). D_{Maha} quantifies how many sigma levels away the biased parameter recovery point is from the injected parameters in the full D -dimensional space. It scales proportionately with the induced biases on the recovered parameters, such that $D_{\text{Maha}} \in [0.0, \infty)$ and the null value is recovered at the injected parameter point. See discussion below Eq. (D13) for more details. In the rest of this Section, we use “Mahalanobis distance” and “sigma contours” interchangeably. We present our main results in the following two Subsections. The first Subsection (results I) quantifies the sigma contour levels in the full-dimensional and the marginalized one-dimensional spaces across the (a, e_0) grid for a fixed total mass binary, while the second Subsection (results II) assesses the impact that different total EMRI masses $(m_1 + m_2)$ have on the biases across the grid. We provide qualitative reasoning for the observed trends in both subsections and draw our conclusions in the final paragraph.

1. Results I: Sigma contour levels

Our results are presented in Fig. 16 for the $(m_1, m_2) = (10^6, 10) M_\odot$ case. For the KERR ℓ 2 case (top panel), we find that the sigma contours categorically scale with both a

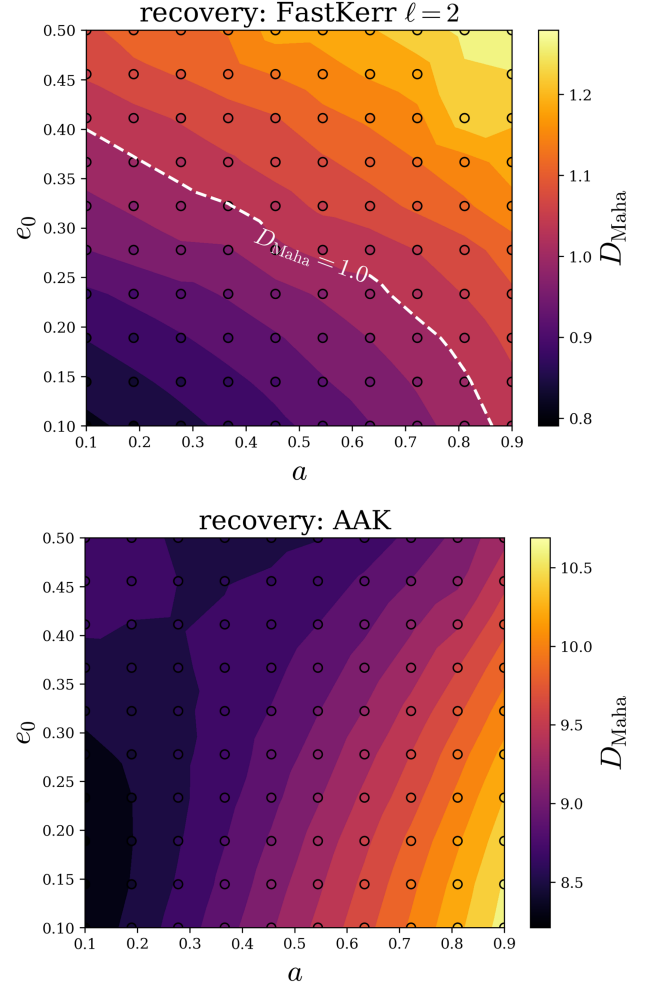


FIG. 16. Distribution of sigma contour levels D_{Maha} of the best-fit θ_{bf} parameters with respect to the truths θ_{tr} visualized on a grid of a and e_0 values. At each gridpoint (black circles), the injected signal is generated using the default model and inferred with the approximate KERR ℓ 2 (top panel) and AAK (bottom panel) waveforms as described in the text, showing the impact of higher-multipole ($\ell > 2$) and relativistic mode amplitudes on parameter recovery, respectively. The plot corresponds to injected MBH and CO (redshifted) masses $(m_1, m_2) = (10^6, 10) M_\odot$. The white line in the top panel represents $D_{\text{Maha}} = 1.0$. Also note the difference in colour scale between the two panels.

and e_0 . On the contrary, in the case of AAK (bottom panel), we observe an inverse scaling with e_0 , with the typical sigma-contour levels ~ 8 – 10 times higher than KERR ℓ 2. We now qualitatively explain these scalings.

Scaling with a . At larger a ’s, the separatrix of the orbit is closer to the horizon of the MBH, which will increase the number of highly relativistic orbits. In this region, the importance of higher-order and fully relativistic mode amplitudes increases, and consequently, parameter recovery with the KERR ℓ 2 and AAK models is worsened.

Scaling with e_0 . For the KERR ℓ 2 model, as the eccentricity increases, higher ℓ -mode amplitudes are no longer

suppressed in general and are more similar in magnitude to the $\ell = 2$ mode (cf. Refs. [66,67,134].) This leads to worsened parameter recovery with the KERR ℓ 2 model. On the other hand, the AAK model does not scale as strongly with eccentricities, neither is the scaling monotonic across the grid. While this trend is qualitatively consistent with the findings in Fig. 15 (see also Section VB), our results warrant a detailed analysis in the future. We argue that at smaller initial eccentricities where the separatrix is closer to the MBH, the CO completes more strong-field orbits, and thus the AAK model incurs larger biases.

We also calculated the one-dimensional marginalized sigma biases (z scores) on each parameter explicitly and found that the luminosity distance d_L incurs the largest biases among all parameters with a median value $\sim 0.3\sigma$ for KERR ℓ 2 and $\sim 1.5\sigma$ for AAK when $(m_1, m_2) = (10^6, 10)M_\odot$. In both cases, the shift in d_L can be understood to make up for the loss of SNR compared to the true signal when it is recovered using the approximate waveforms (also see Section VB). In the KERR ℓ 2 model, all other parameters are typically biased to $\lesssim 0.1\sigma$, showing robust recovery. However, the biases in the AAK model are a factor ~ 3 – 5 times higher (but still within 1σ), consistent with the trend in Fig. 16. Two notable exceptions are the recovered phases $(\Phi_{\phi_0}, \Phi_{r_0})$, with biases a factor ~ 10 – 12 higher compared to the KERR ℓ 2 model, hinting at severe dephasing of the AAK waveform compared to the injected signal.

2. Results II: Other mass pairs

Finally, we analyzed two additional EMRI sources with (i) $(m_1, m_2) = (0.5 \times 10^6, 5)M_\odot$ (lower mass), and (ii) $(m_1, m_2) = (1.5 \times 10^6, 15)M_\odot$ (higher mass). In the KERR ℓ 2 case, we found the qualitative scaling of sigma contours with a and e_0 to be the same. However, for AAK, the scaling with e_0 significantly depended on the total mass of the system: biases incurred by the lower-mass EMRI were largely insensitive to e_0 and scaled with a , while the higher-mass EMRI showed stronger dependence on both a and e_0 . While a detailed analysis is beyond the scope of this work, these scalings can be qualitatively attributed to the relation between the number of relativistic orbits completed, the total mass of the system, and the LISA sensitivity curve, similarly to the discussion in Section VB.

Assessing the scaling of the average incurred biases with the total EMRI masses (and consequently the detector's frequency band), we found opposing trends for the two approximate models, as described below.

KERR ℓ 2 case. On average, the lower-mass EMRI incurred smaller biases while the higher-mass system was more biased. This behavior is consistent with the results of Sec. VD of Ref. [171]. They argue that as the total mass of the binary pair increases, the GW signal slides to lower detector frequencies where the instrument's noise suppresses contributions from lower multipole modes.

Consequently, higher modes become important in the signal and inference biases are introduced in parameter recovery with the KERR ℓ 2 model.

AAK case. The typical biases incurred by the lower-mass EMRI were the largest in this case, opposite to the trend observed in KERR ℓ 2. This may be attributed to the larger number of strong-field orbits completed by the lower-mass EMRI, leading to a larger accumulation of phase errors. The marginalized one-dimensional sigma biases showed that, while other parameters typically incurred similar biases across all three systems, the lower-mass binary incurred $\approx 40\%$ larger biases in the initial phases $(\Phi_{\phi_0}, \Phi_{r_0})$ than the higher-mass system. This is consistent with our interpretation.

Overall, we find that the biases incurred using approximate models are small across a broad grid of a and e_0 values. Even though the AAK waveforms led to sigma contour level biases of factor 8–10 larger than KERR ℓ 2, both models were able to recover all inferred parameters with one-dimensional biases at $\lesssim 1\sigma$. This is useful, e.g., for coarse-grain recovery of the parameters during the search stage of EMRI inference, for which faster and approximate waveforms may be necessary and sufficient [254]. We caution, however, that inference with approximate waveforms may lead to multimodalities in the likelihood surface [228,255], potentially compromising robust recovery. The linear signal approximation framework adopted in our study does not capture such features, and we leave its detailed analysis to future work. Additionally, the absence of accurate models may significantly bias the inference of fully relativistic signals emitted from systems with high primary spins. The accuracy of parameter recovery also strongly depends on the frequency band of the signal, the orbital eccentricity, and the choice of the approximate model.

D. Horizon redshift of asymmetric-mass binaries with LISA

LISA is expected to be sensitive to a wide range of GW signals emitted by the inspiral of both stellar-mass and intermediate-mass COs into MBHs. Depending on the mass ratios of these sources, observations of IMRIs and EMRIs will be effective probes of the origin and evolution of MBHs and their environments in mass ranges currently inaccessible via electromagnetic means [21,256]. In this Subsection, we will investigate the range of distances (and therefore how far back in cosmic history) over which LISA can detect these sources. Figure 17 shows sky-averaged horizon redshifts $\bar{z}$ for asymmetric-mass binaries of various component masses, MBH spins and initial eccentricities. Each line consists of 20 points, with detector-frame primary masses spaced uniformly in their logarithm between 5×10^4 and $5 \times 10^8 M_\odot$. We restrict the two panels in Fig. 17 to the region of primary source mass $m_1^{(s)} \in [10^5, 10^7]M_\odot$. For each combination of intrinsic

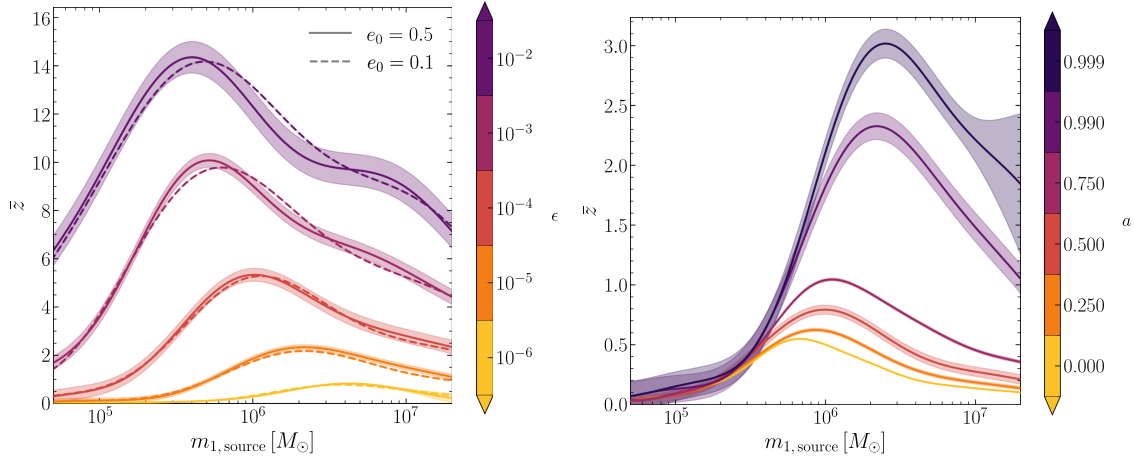


FIG. 17. Here we show the evolution of the horizon redshift $\bar{z}$ at which the sky-averaged SNR of an eccentric, equatorial I/EMRI reaches a threshold SNR $\bar{\rho} = 20$, for various mass ratios (*left panel*) and spins (*right panel*), as a function of the source-frame primary mass. All systems shown are two-year inspirals, and we set an initial eccentricity $e_0 = 0.5$ for all systems represented by solid lines. Shaded regions represent the 1σ uncertainty region predicted by a GPR trained on the produced data. *Left panel*: Systems with $a = 0.99$ and $\epsilon \in [10^{-6}, 10^{-2}]$. Dashed lines represent systems with the same properties, but initial eccentricity $e_0 = 0.1$. We do not plot the associated uncertainty intervals for visual clarity, but they are similar in scale to those of the $e_0 = 0.5$ case. *Right panel*: Systems with $\epsilon = 10^{-5}$ and $a \in [0.0, 0.999]$.

parameters, we draw 100 sets of extrinsic parameters (initial phases and orientation/sky-position angles) according to Appendix C, and average the SNRs of the corresponding systems.

We then compute the luminosity distance d_L at which each system should be placed to achieve an SNR detection threshold of 20 (a value commonly assumed in the literature [245,248,257]) and convert it to $\bar{z}$ assuming PLANCK18 cosmology [258,259]. Finally, to take into account the dependence on the number of generated sources during the averaging procedure and the discreteness of the mass grid used for the data generation, we fit a Gaussian process regressor [260,261] on the computed $(m_1^{(s)}, \bar{z})$ points. This allows us to provide an estimate of the uncertainty associated with each horizon redshift curve, and to extrapolate to regions beyond the parameter space coverage of our model for two-year inspirals. This extrapolation applies for large masses, large spins, and high eccentricities (which correspond to highly eccentric inspirals at low orbital separations, which are outside of the domain of validity of our model), and leads to the large uncertainty associated with the high-mass tail of the $a = 0.999$ curve in the right-hand panel of Fig. 17.

We find that the horizon redshift only mildly depends on the initial eccentricity for all values of a explorably with our waveform model. The left panel of Fig. 17 shows the mass-ratio dependence of the horizon redshift for two different values of e_0 , namely 0.5 (higher eccentricity; solid lines) and 0.1 (lower eccentricity; dashed lines). We find that the maximum horizon redshift of the lower-eccentricity curve is only slightly smaller than that of the higher-eccentricity curve, with the relative difference between the maxima of

the two eccentricity curves being less than 6% across all mass ratios considered. More interestingly, the two show different correlations with the primary mass; in particular, the lower-eccentricity curves do not seem to show the SNR “hump” we see in the higher-eccentricity curves. This is especially evident in the $\epsilon \gtrsim 10^{-3}$ cases, and can be interpreted as the contribution of higher-frequency waveform modes with large n ; at larger total masses, these modes lie in the most sensitive region of the LISA sensitivity curve (shown in Fig. 1) and contribute significantly to the total SNR of the system. As the amplitudes of these modes scale strongly with eccentricity, this feature is more readily apparent for particularly eccentric systems. Additionally, the value of $m_1^{(s)}$ at which $\bar{z}$ is maximized varies with respect to e_0 . The deviation is reasonably small, with a relative difference between the two eccentricity cases (over all primary mass values) of $\sim 18\%$ at its largest. This effect depends on the mass ratio: the low-eccentricity peak is shifted toward smaller masses for the $\epsilon = 10^{-6}$ systems, while it is moved toward larger values for the remaining systems considered, with larger shifts for larger mass ratios. The most significant contribution to the SNR occurs as the CO enters the strong field (toward the end of inspiral); due to our initial conditions on p_0 in this analysis being fixed by the time to plunge, the eccentricity at this stage of the inspiral decreases as a function of mass ratio (as the inspiral will circularize more when starting from larger orbital separations). This accentuates the impact of eccentricity on detectability as mass ratio increases, as depicted in the left panel of Fig. 17.

We also find that the mass location of the detectability peak shows different trends in the two panels, moving

toward lower (higher) source frame primary masses with larger mass ratios (MBH spins). The relationship between mass ratio and SNR (left panel), for a given value of m_1 , is largely driven by the proportionality of the waveform amplitudes to $\mu \sim m_2$ [Eq. (1)]. As the rate of inspiral also increases with mass ratio, our fixed time to plunge leads to higher values of p_0 ; as this decreases the frequencies of strong harmonics, this in turn slightly accentuates the detectability of lower-mass systems, which manifests as a gradual shift in the location of the maximum horizon redshift in the left panel of Fig. 17. The impact of MBH spin (right panel) is also highly pronounced, particularly at larger values of m_1 . Raising MBH spin increases the number of strong-field orbital cycles completed by the CO before it plunges, where both the amplitude and frequency of the GWs emitted are largest. For lower- m_1 systems, this high-frequency radiation sits toward the upper end of LISA's sensitivity and therefore does not contribute strongly to the SNR (and therefore the detectability) of these systems. As m_1 increases, this strong-field emission decreases toward the frequencies where LISA is most sensitive, significantly enhancing the overall SNR. This trend continues until the strong-field emission shifts below the minimum in the LISA sensitivity curve (~ 3 mHz; see Fig. 1) at which point the SNR begins to decrease.

To summarize quantitatively the prospects for probing I/EMRI sources over cosmological distances with LISA, we find that for an SNR detection threshold of 20, we can detect EMRIs with mass ratios of 10^{-5} at redshifts $\lesssim 3$, with the maximum horizon redshift at $m_1^{(s)} \sim 2 \times 10^6$ and $a = 0.999$. For systems with $a = 0.99$, the horizon redshift increases to ~ 5 for EMRIs with mass ratios of 10^{-4} , and to ~ 14 for IMRIs with mass ratios of 10^{-2} .

E. Distinguishing quasicircular and mildly eccentric systems

The astrophysical formation scenarios of EMRI systems are one of the key areas that GW observations of these systems can address. One crucial parameter that can help distinguish between different formation models is eccentricity. While event rates are largely uncertain [28,262,263], a significant fraction of EMRI systems may form in gas-dominated environments such as accretion disks [51]. The astrophysics underpinning the formation and evolution of these “wet” EMRIs is poorly constrained at present [52]. In general, interactions between the disk and the inspiralling object are expected to significantly dampen orbital eccentricity such that the system is quasicircular and nearly equatorial once its GW emission enters the LISA band. However, more complicated models of these disk-inspiral interactions or other environmental effects (such as interactions with other bodies in the disk) may yield LISA-band EMRIs in disks with small (and potentially measurable) eccentricities

[264,265]. Identifying these small eccentricities in EMRI signals may therefore provide significant insights into the astrophysical processes that drive the formation of these systems. To this end, in this short Subsection we will explore and identify a lower bound for measurable eccentricities in eccentric equatorial EMRI systems with a rapidly spinning MBH.

Intuition on measurable eccentricities can be attained through a simple mismatch case study. Consider an EMRI system $h^{\text{circ}} \equiv h(t; e = 0)$, with masses $(m_1, m_2) = (10^6, 10)M_\odot$, dimensionless primary spin $a = 0.998$ and initial semilatus rectum $p_0 = 10.628$ that is observed by LISA for two years prior to the plunge of the secondary object. When attempting to measure the system's eccentricity during data analysis, we will propose a template $h^{\text{ecc}} \equiv h(t; e)$ with nonzero eccentricity and compare it with the observed signal h^{circ} . To study the measurability of small eccentricities, we will examine how the mismatch $\mathcal{M}(h^{\text{circ}}, h^{\text{ecc}})$ behaves in the limit as $e \rightarrow 0^+$. A more complete study of the detectability of a small nonzero eccentricity would include comparison of Bayes' factors for the quasicircular and eccentric hypotheses, which is beyond the scope of this paper; we simply seek to establish an order-of-magnitude estimate of what eccentricities are measurable for a typical EMRI system.

We compute mismatches $\mathcal{M}(h^{\text{circ}}, h^{\text{ecc}})$ for $e_0 \in [10^{-8}, 10^{-3}]$. Our results are given in Fig. 18. For very low eccentricities $e_0 \in [10^{-8}, 10^{-5})$, the mismatch between circular and eccentric waveforms is dominated by ODE

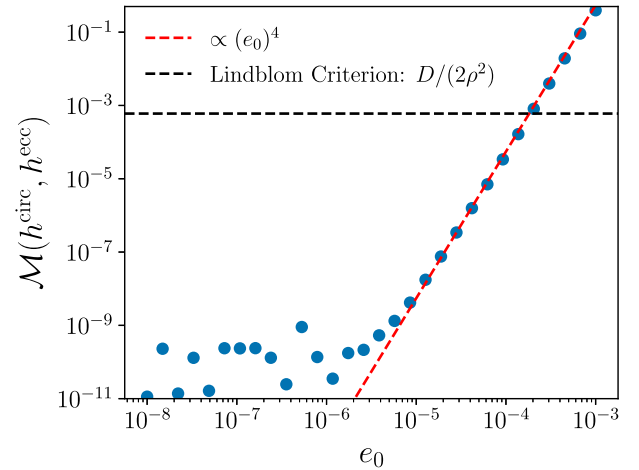


FIG. 18. Mismatches (blue dots) between quasicircular (h^{circ}) and eccentric h^{ecc} waveforms as $e_0 \rightarrow 0$. Other source parameters $\{m_1, m_2, a\} = \{10^6 M_\odot, 25 M_\odot, 0.998\}$, with SNR $\rho = 100$. The black dashed line indicates the Lindblom distinguishability criterion [Eq. (15)] for this SNR, with $D = 12$ (12 sampled parameters). For $e_0 \gtrsim 10^{-5}$, mismatches grow as $(e_0)^4$ (red dashed line) in line with PN scaling arguments. At lower initial eccentricities, other systematics in waveform generation obfuscate any physical relationship between mismatch and initial eccentricity.

numerical errors. However, for $e_0 \gtrsim 10^{-5}$, the mismatch grows by an amount proportional to $(e_0^2)^2$ as given by the red dashed line. The black dashed line indicates the Lindblom distinguishability criterion [Eq. (15)], where mismatch values $(<) > D/2\rho^2$ indicate that the true and approximate waveforms are (in)distinguishable. As in Sec. IV C, we set the number of sampled model parameters $D = 12$. From Fig. 18, we see that the two models are distinguishable for $e_0 \gtrsim 2 \times 10^{-4}$, highlighting that for $\rho < 100$ we cannot resolve eccentricities smaller than $e_0 < 2 \times 10^{-4}$ for this set of source parameters. We have checked that our conclusions remain unchanged when decreasing the ODE integrator error σ_{tol} .

To understand why the mismatch $\mathcal{M}(h^{\text{ecc}}, h^{\text{circ}}) \propto (e_0^2)^2$, it helps to understand what drives the dephasing in the first place. If $h^{\text{ecc}} \sim h^{\text{circ}} e^{i\Delta\Phi}$ for $\Delta\Phi$ the dephasing between h^{circ} and h^{ecc} , then $\mathcal{M} \approx 1 - \cos(\Delta\Phi) \approx \Delta\Phi^2/2$. We understand from the stability of quasicircular inspirals [129,138,266–268] that, in the low eccentricity limit, the fluxes scale as $\sim \dot{E} \propto \dot{E}_{\text{circ}} + \mathcal{O}(e^2)$. The dephasings between h^{circ} and h^{ecc} are therefore driven by differences in the fluxes that are proportional to e^2 , implying that $\Delta\Phi \sim e^2$. From these simple scaling arguments, it follows that $\mathcal{M}(h^{\text{circ}}, h^{\text{ecc}})$ should grow proportionally to $\Delta\Phi^2 \sim (e_0^2)^2 = (e_0)^4$, in line with our observations.

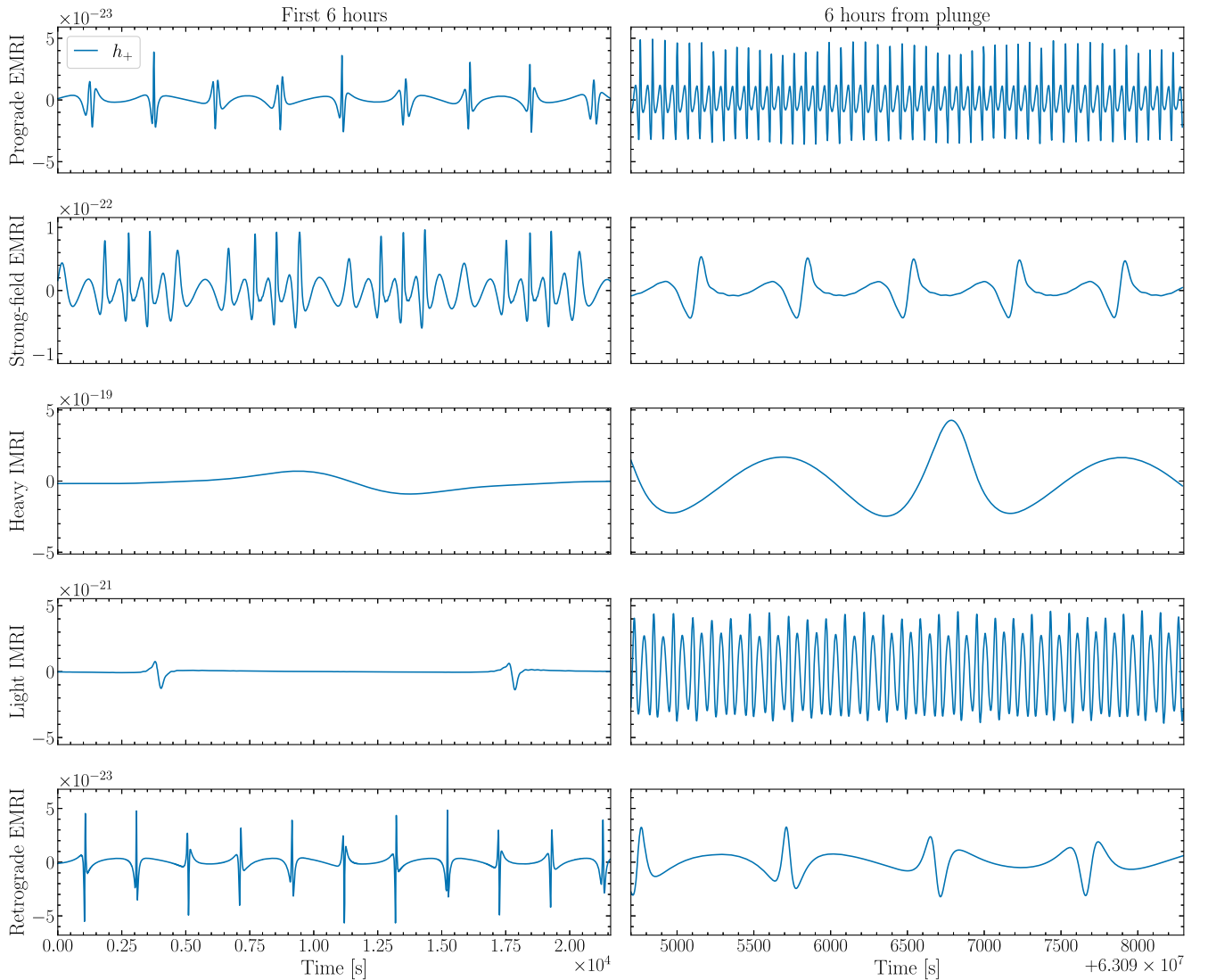


FIG. 19. Time-domain snapshots of the science-case waveforms described in Table I during the early (*left panel*) and late (*right panel*) periods of a two-year inspiral. This set of representative sources demonstrates the wide range of waveform morphologies captured by our model. In all the panels, only the plus polarization h_+ is shown; for the two IMRI systems, initial orbital phases have also been adjusted for visual clarity.

F. Inference prospects for asymmetric-mass binaries with LISA

In this Subsection, we investigate parameter estimation prospects for a range of IMRI and EMRI sources. As FEW is capable of generating waveforms in ~ 100 ms (Sec. VA), we are able to perform Bayesian inference studies on a timescale of hours, directly exploring the structure and scale of the posterior distributions of these sources. The motivation for performing these analyses is twofold. First, we seek to verify that our model reproduces features expected of posterior distributions for these systems: in particular, that they are approximately Gaussian, and that parameters intrinsic to the inspiral dynamics (such as spin and eccentricity) are recovered with high precision. Second, it was identified in Ref. [171] that the mode content of EMRI waveform models can be reduced significantly without significantly biasing inference results; our analyses will reexamine this conclusion with the inclusion of MBH spin and for IMRI sources. To our knowledge, this is the first full Bayesian investigation of rapidly spinning and highly eccentric IMRI and EMRI systems to appear in the literature to date (however, see Ref. [269] for a parameter estimation study of mildly eccentric binaries with $q \leq 10$).

We perform parameter estimation on simulations of each of the sources listed in Table I. The parameters of these sources were chosen to represent the wide range of waveform morphologies accessible with our waveform model Fig. 19, with luminosity distances set to astrophysically motivated values. A notable exception to our analysis conventions is that we set a finer sampling cadence of $dt = 2$ s for Sources 4 and 5 in order to better resolve their high-frequency mode content; for other sources, we use our fiducial sampling cadence of $dt = 5$ s. While these sampling cadences are not quite high enough to resolve the highest-frequency harmonics near plunge for some of these systems (particularly Sources 1 and 4), we found that these cadences were sufficient to obtain accurate parameter constraints. For each simulation, we generate a waveform with mode-selection parameter $\kappa = 0$ (i.e., all modes in our amplitude model are included). During inference, we consider $\kappa = 10^{-5}$ and $\kappa = \{10^{-3}, 10^{-2}\}$ (the latter value depending on the SNR of the injection), demonstrating how parameter inference results are impacted when weaker harmonics (which are more important for high eccentricities and spins; see Fig. 4) are neglected. We do not show the (marginal) posteriors obtained from our inference runs in the main text for brevity; they can be found in Appendix F, and are referenced in the rightmost column of Table II, where we summarize the relative measurement precisions for parameters of astrophysical interest. These include the source frame masses $m_i^{(s)}$, spin a , final eccentricity e_f , and luminosity distance d_L . We also report the sky-localization

area of each source: given the covariance matrix of the sky-angle posterior samples $\Sigma = \text{cov}(\theta_s, \phi_s)$, and $\bar{\theta}_s = \text{median}(\theta_s)$, we compute the 99% quantile of the sky-localization area in square degrees, given by

$$\Delta\Omega^{\text{sky}} = 9 \cdot 2\pi \sin \bar{\theta}_s \sqrt{\det \Sigma} \left(\frac{180}{\pi} \right)^2. \quad (16)$$

We remind the reader that all parameter inference was conducted using second-generation time-delay interferometry (TDI) variables as implemented by the FASTLISARESPONSE software [270], assuming constant and equal-armlength orbits; we sampled all posterior distributions with Markov chain Monte-Carlo (MCMC) using the ERYN package [271–273] and default settings.

1. EMRIs

We investigate three EMRIs with mass ratios $\epsilon \in \{10^{-4}, 10^{-5}, 10^{-6}\}$ and parameters given by the fifth, first, and second rows of Table I, respectively. As confirmed by our horizon redshift study in Section VD, these sources are indeed observable and each is unique in its own way. Source 1 has high spin and initial eccentricity $(a, e_0) = (0.998, 0.730)$, and represents a fairly typical EMRI, assuming that many MBHs in nature are spinning rapidly [28, 274]. The more extreme mass-ratio of Source 2 causes its inspiral trajectory to evolve on a timescale much slower than that of Source 1. This is reflected by its two-year inspiral starting deep in the strong field of the primary ($p_0 \sim 2.12$) for $(a, e_0) = (0.998, 0.425)$, terminating ~ 0.15 away from the separatrix. Notice that the final eccentricity $e_f \sim 0.261$; the orbit is still quite eccentric close to plunge, and therefore lies in the upper tail of the eccentricity-at-plunge population for EMRIs formed via capture in Ref. [28]. The final EMRI we consider, (Source 5) is a retrograde inspiral that (due to its mass ratio $\epsilon = 10^{-4}$) begins in the weak-field regime ($p_0 = 26.19$) with high $e_0 = 0.8$ and moderate spin $a = -0.5$, terminating near the separatrix with nontrivial eccentricity $e_f = 0.195$. Based on a prior examination of EMRI astrophysics in Ref. [28], we believe these are characteristic asymmetric-mass binaries that are readily observable by LISA to plausible luminosity distances (and therefore redshifts).

Measurement precisions on the parameters of these sources are shown in the corresponding rows of Table II. For Source 1, we see that we can constrain the intrinsic parameters to usual levels ($\Delta\theta/\theta \sim 10^{-6}$) for a source with an SNR of 30. From the marginal posteriors (shown in Fig. 28), we find that setting $\kappa = 10^{-2}$ (retaining 235 modes) yields identical inference results to $\kappa = 10^{-5}$ (1228 modes), with any systematic biases present being too small to identify. Similar results are obtained for Source 2 in terms of measurement precision (accounting for the lower

TABLE II. Relative precisions obtained in the parameter estimation of the sources listed in Table I (where δx denotes the relative precision in parameter x , defined as the 1σ width of the marginal posterior on x normalized by the median value). The ^(s) superscripts indicates that masses have been converted in the source frame. For each sample, we recover the redshift from the luminosity distance d_L assuming the PLANCK18 cosmology. We also report $\Delta\Omega^{\text{sky}}$ [Eq. (16)], which is 99% of the sky-localization area of each source in square degrees. In the last column, we provide the reference for the corresponding posterior distributions in Appendix F. Notably, $\delta m_1^{(s)}$ and $\delta m_2^{(s)}$ are a factor of ~ 3 orders of magnitude larger than the equivalent quantities computed in the detector frame, $\delta m_i^{(d)}$. This is a consequence of the mass relation $m_i^{(d)} = (1 + z_r)m_i^{(s)}$, which introduces the (much larger) uncertainty in d_L due to its relationship with z_r via our assumed cosmology.

	$\delta m_1^{(s)} (10^{-3})$	$\delta m_2^{(s)} (10^{-3})$	$\delta a (10^{-5})$	$\delta e_f (10^{-4})$	$\delta d_L (10^{-2})$	$\Delta\Omega^{\text{sky}} (\text{deg}^2)$	Figure
1	9.181	9.181	0.023	0.785	3.802	1.049	(28)
2	16.506	16.505	0.040	0.295	5.326	3.893	(29)
3	2.459	2.455	2.600	2.545	0.776	17.006	(32)
4	2.650	2.650	5.234	69.845	0.870	1.079	(30)
5	9.200	9.200	11.830	0.154	5.847	1.540	(31)

SNR of this source), except for the improved determination of e_f , which is expected due to the significant final eccentricity of this source. Correlations between parameters are more pronounced for this source than the other two, which is likely a feature of the mass ratio of this source leading to slower orbital parameter evolution. Similarly to Source 1, there is an absence of any systematic biases in Fig. 29 between the posteriors for $\kappa = 10^{-5}$ (809 modes) and $\kappa = 10^{-2}$ (186 modes). For Source 3–5, the spin constraint are wider than for the other two sources, which is expected as the effects of spin are less pronounced for this source (as the inspiral spends a large fraction of the two-year duration in the weaker field). The same can not be said for eccentricity effects (which are still significant in the weak field), which is reflected in the recovery of e_f . Once again, no systematic biases in Fig. 31 are observable between posteriors for $\kappa = 10^{-5}$ (928 modes) and $\kappa = 10^{-2}$ (208 modes).

Our results confirm that EMRI posterior distributions continue to be extremely narrow with the inclusion of MBH spin. In all cases, we recover Gaussian posteriors, which in turn implies that information matrices are an effective tool for probing the measurement precision for these sources. These conclusions do not change when waveform mode content is reduced by 70%–75% in all three cases, which in turn reduced inference times by more than a factor of two. Despite the poorer measurement precisions obtained when analyzing the retrograde source, luminosity distance and sky localization area were constrained to roughly equal precision for all three analyses, which is expected due to the similar SNRs of these sources. Notably, this leads to source-frame mass measurements of similar precision in all cases, despite the wider uncertainties in detector-frame masses for the retrograde source. This highlights that while retrograde systems may not provide as strong a constraint on the MBH spin, they will still be as effective as prograde EMRIs for probing the MBH mass population [41,42]. With sky localization areas of $\sim 1 \text{ deg}^2$, both prograde and retrograde

EMRIs will also be valuable signals for dark-siren cosmological analyses [56,275,276].

2. IMRIs

As demonstrated in our horizon redshift study (Fig. 17), IMRIs will be observable by LISA with high SNRs to high redshifts. To our knowledge, we have developed the first waveform model capable of modeling highly eccentric IMRIs into rapidly spinning MBHs (at the adiabatic order) that is sufficiently rapid for use in Bayesian inference studies. To demonstrate these capabilities and perform an initial investigation of IMRI posterior structure, we examine two IMRIs with mass ratios $\epsilon = 10^{-2}$ and parameters given by rows 3 and 4 in Table I. Due to the large mass ratio, we see that the orbital parameters evolve far more rapidly than for the EMRI sources we examined previously. Indeed, despite their high initial eccentricities $e_0 = 0.85$, the binaries have circularized significantly prior to plunge. Notably, the SNRs $\rho = \{500, 200\}$ of these sources are also significantly larger, mainly due to the waveform amplitudes being proportional to $\mu \sim m_2$ [Eq. (1)]. However, having such high SNRs may lead to more significant systematic biases when waveform models with incomplete mode content are used in parameter inference, as we explore below.

For both Source 3 and Source 4, waveforms with $\kappa = 10^{-5}$ (949 and 626 modes, respectively) are sufficiently complete for performing parameter estimation with no observable biases, as shown in Figs. 30 and 32. However, when attempting to infer the parameters with $\kappa = 10^{-2}$ (not shown in Figures; 223 and 156 modes, respectively) we observed statistically significant biases across all the IMRI parameters. Nearly all parameters were recovered with biases of $\gtrsim 3\sigma$ away from the truth (with σ computed as the standard deviation of the posterior samples) that were readily identifiable compared to the scale of the posterior. Even for $\kappa = 10^{-3}$ (412 and 291 modes, respectively), we observe unacceptable biases in

the luminosity distance and sky location on the order of $\sim 4\sigma$ and $\sim 7\sigma$ biases away from the true parameters, respectively. Biases of this scale would propagate to population/cosmological analyses and may contaminate the results of these studies, especially when considering that luminosity distance biases will affect source-frame mass estimates as well. This behavior is to be expected considering the higher SNRs of these sources—from Eq. (15), we can expect mismatches of 10^{-3} (which $\kappa = 10^{-3}$ roughly leads to) to be insufficient for the analysis of signals with SNRs of hundreds. The fact that extrinsic parameters are most affected by the omission of impactful waveform modes is also expected, as this has been previously identified in the analysis of comparable-mass case binaries observed with LISA [277].

We report measurement precisions of parameters of astrophysical interest for these sources in Table II. Despite the higher SNRs of these signals, their spins are recovered with larger uncertainties than for the prograde EMRI sources we examined. This can be understood in terms of the rapid evolution of the IMRI trajectories, which complete relatively few cycles in the strong-field regime where the effects of spin are most important. Eccentricity at plunge is also recovered more poorly; this has similar justification, in addition to the IMRI trajectories beginning at larger orbital separations and therefore entering the strong-field regime with lower eccentricities than the EMRI sources. Luminosity distances are recovered precisely (to better than 1% precision) in both cases, which is to be expected as amplitude measurements depend strongly on the SNR. This has the interesting consequence that despite the recovery of detector-frame masses for the IMRI sources being less precise than for the EMRI sources, they still provide more precise recovery of source-frame masses due to the improvement in the luminosity distance measurement. Last, we note that the 99% sky-localization area for Source 3 ($\sim 17 \text{ deg}^2$) is more than an order of magnitude larger than that of Source 4. This occurs because during early inspiral, the bulk of the mode spectrum for this source sits mainly at lower frequencies ($\lesssim 10^{-4}$ Hz) that are poorly measured by LISA. As the inspiral evolves during the observation, the mode spectrum shifts to higher frequencies (“entering” the lower end of the LISA band) and is measured well. While this still enables many parameters to be recovered precisely, the sky position of the source is determined by long-timescale variations in the LISA response function; as the waveform of this source only has significant SNR for part of the observation (such that fewer of these variations are measurable), this in turn leads to a poorer localization of the source on the sky.

These analyses highlight that LISA will be capable of measuring the parameters of IMRI systems very precisely, due in part to the high SNRs of these systems. Comparing between the extreme-mass-ratio limit (where many harmonic modes are necessary for unbiased inferences) and the

comparable-mass limit (where current waveform models include only a handful of higher-order modes [278,279]), it has not been explored in detail in the literature where IMRI modeling lies on this scale (particularly with the inclusion of significant eccentricity). Our results indicate definitively that IMRIs (at least, of mass ratios $\sim 10^{-2}$) lie toward the EMRI end of this scale—models will need to accurately produce the harmonic mode structure of waveforms for these sources if their parameters are to be faithfully recovered. We note that (similarly to EMRIs) it is likely that models with limited mode coverage will be sufficient for detecting these sources in LISA data, especially given the larger SNRs of these sources. The computational benefits of doing so will be significant: for our analyses, the $\sim 55\%$ reduction in mode content achieved by setting $\kappa = 10^{-3}$ halved the parameter estimation runtime.

VI. CONCLUSIONS AND FUTURE WORK

A. Summary of this work

In this work, we have presented the first GSF-based waveform model for asymmetric-mass binaries capable of efficiently generating analysis-length eccentric equatorial inspirals into rapidly rotating black holes (at the adiabatic order). This model is housed in the new v2 release of the FEW package, which provides the constituent components of this waveform model—new inspiral trajectory and mode amplitude models—as part of a modular framework that is readily adaptable to meet the requirements of the end user. In addition to the introduction of primary spins $|a| \leq 0.999$, our model also extends semilatus rectum and eccentricity support from $(p_0, e_0) = (\sim 16, 0.7)$ to $(p_0, e_0) = (200, 0.9)$ with respect to the previous adiabatic FEW model. This expansion of parameter-space coverage extends FEW to the modeling of long-duration IMRIs, providing the first efficient adiabatic-order waveforms for these systems that incorporate large spins and/or eccentricities to appear in the literature. In recent work, PN- and Effective-One-Body-based waveform models have been developed that may be extended to the IMRI regime (albeit for smaller or moderate eccentricities), e.g., Refs. [88,90,95,98,100,103,104]. Similarly to our case, these models do not have long-duration IMRI simulations against which to investigate their accuracy; comparison between these models and the one we have presented here (particularly with the inclusion of post-adiabatic effects) is therefore worthy of future investigation.

We examined significant sources of systematic error in our model—the precise computation and interpolation of GSF data products—with comparisons against independent datasets and analyses of interpolation error convergence, and estimated that our model attains mismatches of $\sim 10^{-5}$ with respect to error-free adiabatic waveforms over the majority of the parameter space. The most significant source of systematic error we identified was the interpolation of forcing functions. Errors due to the linear

interpolation of mode amplitudes with respect to spin were also found to increase significantly near the separatrix. During this process, we identified that our model is least robust for eccentricities $e_0 \gtrsim 0.85$, and for spins $a \sim 0.998$ near the separatrix ($p \sim p_{\text{sep}} + 1$). Waveform generation wall-times were found to be ~ 100 ms, and were slightly larger than those obtained in Ref. [171] for zero-spin systems. This increase in cost is due primarily to the inclusion of larger n modes in the waveform (especially for prograde inspirals with larger spins), but it also due to the implementation of the FEW trajectory module entirely in PYTHON (which is less efficient than the previous C implementation). The latter choice is worth the performance penalty, as it significantly improves the customizability of the FEW framework via modifications to the inspiral model, and access to the continuous ODE solution is an asset to waveform generation and data analysis schemes.

By comparing with both the AAK and a quadrupolar variant of our waveform model, we demonstrated the importance of relativistic amplitudes and higher-order modes when performing SNR computations and parameter estimation. We found that SNRs computed with the AAK are more inaccurate for stronger-field high-mass systems, where the missing higher-order modes are most important and the semirelativistic amplitudes of the AAK are least accurate. For all component masses considered, the AAK overestimated the SNR for large spins (by up to 60% in the high-mass case), and exhibited larger SNR errors for larger eccentricities. This has ramifications for quantitative studies of the EMRI detection rate in the literature (such as Ref. [28]), which should be revisited with relativistic waveform models in order to obtain better rate estimates. Exploring inference biases under the linear signal approximation, we identify biases of $\sim 1\sigma$ with the AAK and $\sim 0.1\sigma$ with the quadrupolar relativistic model. While these biases may be identifiable in the parameter estimation of EMRIs, they are sufficiently small that these approximate models may be readily applicable in the search for EMRI signals in detector data (where any corresponding reductions in computational cost would be highly beneficial). An exploration of this possibility is left to future work.

To quantify and explore LISA's sensitivity to GWs from asymmetric-mass binaries, we computed sky-averaged horizon (maximum) redshifts at which LISA can detect sources with varying component masses, spins and eccentricities. We found that LISA can detect EMRIs of mass ratio 10^{-6} (10^{-4}) at redshifts of up to 1 (5), and IMRIs of mass ratio 10^{-3} (10^{-2}) at redshifts of up to 10 (15), with only $\sim 10\%$ variations with respect to eccentricity. In the case of EMRIs with mass ratio 10^{-5} , horizon redshifts varied from 0.5 to 3 as spin increased from 0 to 0.999. For all configurations, we found that LISA was most sensitive to binaries with source-frame primary masses $m_1^{(s)} \sim 10^6 M_\odot$. Our horizon redshift results highlight that the detection and

characterization of asymmetric-mass binaries will enable LISA to effectively probe the evolution of MBH, intermediate-mass black hole and stellar-mass black hole populations over cosmic time, in line with LISA's science objectives. The waveform model we have developed (as a component of the FEW framework) is unique in the respect that it is able to accurately reproduce the physical characteristics of this wide range of GW sources, providing the tools necessary for the exploration of LISA science with IMRI and EMRI observations.

In a similar vein, we also examined the minimum eccentricity measurable by LISA for a representative (and rapidly spinning) EMRI system. Understanding this limit is highly relevant to models of EMRI formation in gas-dominated environments (and therefore the population of these systems as a whole). We find that the relationship between eccentricity and mismatch (with respect to a quasicircular waveform) scales as e_0^4 (in line with low-eccentricity expansions around quasicircular inspirals), and that eccentricities less than $\sim 10^{-4}$ are unlikely to be measurable by LISA for these systems. This is significantly larger than the typical measurement precision achievable for EMRIs with larger eccentricities, serving as a reminder that measurement precisions on inspiral parameters are not synonymous with the minimum values of these parameters that are measurable. We only examined a single system in our analysis, as our goal was simply to obtain a rough intuition for the measurability of small eccentricities; characterizing how the detectability of these small eccentricities varies across the EMRI parameter space is an area worthy of future investigation.

Last, we performed Bayesian parameter estimation of several EMRI and IMRI sources with MCMC methods, attaining relative precisions of $10^{-2} - 10^{-3}$ for source-frame mass measurements. The spins and final eccentricities of these systems were measured with precisions of $10^{-7} - 10^{-4}$ and $10^{-5} - 10^{-3}$, respectively. In most cases, luminosity distances and 99% sky-localization areas were measured with $\sim 1\%$ precision and errors of $\sim 1 \text{ deg}^2$, respectively, for two-year inspirals. We identified significant biases when performing inference on IMRI observations using waveforms with reduced mode content. This is the first demonstration of importance of higher modes in the context of IMRIs for LISA, highlighting that the analysis of these signals is more similar in principle to that of EMRIs than comparable-mass binaries. While the systematic biases we observed when using fewer harmonic modes were significant in the context of inference, they were sufficiently small that such approximate models may be readily applicable in the detection and identification of these signals (especially given the factor of ~ 2 speedup we observed in our analyses when using these models). From our inference results, it is clear that significant effort is required to properly explore the scientific potential of IMRI observations with LISA. As we have demonstrated, these

sources are readily modeled and analyzed with the FEW framework, which we expect will be foundational to this future work.

B. Prospects for future work

The implementation of this new waveform model and the general improvements to the FEW framework lays the foundations for many extensions to incorporate other physical features necessary for a complete description of generic I/EMRI systems. In terms of parameter space coverage this includes extending to precessing and then generic (eccentric and precessing) orbits. Further extensions to the model include adding the merger and ringdown, resonance effects, and GW memory. In terms of improving phase accuracy, postadiabatic corrections from second-order self-force calculations and due to the spin on the secondary need to be incorporated. The FEW framework is designed to incorporate all of these effects whilst maintaining rapid waveform generation. We now briefly discuss how each of these extensions can be made.

1. Extension to precessing orbits

Precessing orbits, sometimes known as spherical orbits in the EMRI literature (e.g., Refs. [168,280–282]), are orbits with $e = 0$ but $|x_I| \neq 1$. Just as with eccentric, equatorial orbits the parameter space is three-dimensional consisting of (a, p, x_I) . Thus all the framework in this paper for interpolation the fluxes and waveform amplitudes will carry over to the precessing case. The main challenge is to compute the fluxes and amplitudes across the parameter space. Codes exist to make these calculations, e.g., Refs. [130,231,283] and the overall computational cost will likely be less than was required for the present work as the number of polar harmonics needed to accurately compute the flux is typically less than the number of radial harmonics needed for eccentric orbits.

Extension to generic orbits. The FEW framework is designed with generic, eccentric and precessing orbits in mind but some technical limitations must be overcome before such a model can be built. Of particular note is the increased difficulty in interpolating data for these inspirals; the parameter-space dimensionality increases to four which greatly expands the (already large) memory requirements for mode amplitude interpolation. Given that polar ($k \neq 0$) mode amplitudes must also be computed (increasing the total number of modes by more than an order of magnitude [226]), extrapolating from the domain of our model leads to mode amplitude grids of ~ 500 GB in size or more, which is several times too large to fit in the memory of existing GPU hardware. The overall cost of computing the underlying data for flux and amplitude interpolation also grows significantly as the higher dimensional parameter space necessitates sampling the space at more points. Furthermore, at each point in the parameter space the perturbation theory computation is more expensive as one must integrate over

both an orbit’s radial and polar motions, and one must sum over both radial and polar frequency harmonics. This increases the computational cost of each mode which contributes to the flux by a factor of about 20–30, and increases the number of modes that contribute to the flux by a similar factor. The problem remains highly parallelizable, so it will not be difficult to exploit efficient codes which can compute these data [67,123,134,137,156]. However, the sheer quantity of data needed means that care is needed in designing the grid spacing across the parameter space before committing significant computational resources.

2. Inspiral-merger-ringdown models

For typical EMRIs, very little SNR is accumulated during the merger and ringdown portions of the waveform. With recent work showing that GSF results can also be used to model IMRIs [159,164,284,285], it becomes more important to include the merger and ringdown [106,108,109,286]. Previous approaches to computing these in perturbation theory involved computing the waveform using time-domain codes [105,287,288] and thus the results were not suitable for the FEW framework. To overcome this, recent work has cast the transition from inspiral to plunge and the plunging portions of the waveform within the multiscale framework [107,110]. Within this approach the online computation of the transition and the plunge involves solving new trajectory ODEs and the waveform is computed from a new set of amplitudes. Structurally, this is exactly the same as the present online inspiral computation and, as such, the results of these calculations can be readily incorporated into FEW.

3. Resonances

Generic inspirals can experience resonant phenomena when any two of three orbital frequencies ($\hat{\Omega}_\phi, \hat{\Omega}_\theta, \hat{\Omega}_r$) are a low-integer multiple of each other [115,133,204,207,289–292].¹⁰ In the context of the resonance caused by GSF effects, resonances involving both the radial and polar frequencies are particularly significant. When resonances occur, they give a small “kick” to the orbital elements. After this kick, a phase difference with respect to an adiabatic model (that does not include resonances) accumulates over a fixed frequency window as $\nu^{-1/2}$ (e.g., Refs. [122,169,204–206,297]); resonances therefore enter between adiabatic and postadiabatic order, and constitute the most significant correction to adiabatic inspirals in the multiscale framework. In order to incorporate these effects, FEW will need to detect the resonant surfaces and apply the appropriate precomputed jumps to

¹⁰Similar resonances in I/EMRI systems can also occur as a consequence of the existence of a third body [36,37,293,294] or the presence of gravitational perturbations in the environment more generally [295,296]. In such cases, the azimuthal frequency is implicated in the resonance.

the orbital elements. The new version of trajectory module presented in this work provides access to a high-order interpolation of the phase space trajectory and its derivatives, which will make root finding for these resonance surfaces inexpensive and straightforward to implement.

4. GW memory

In the comparable-mass regime, it was recently shown that including GW memory effects can help to break degeneracies between parameters [298], and that not including memory effects in waveform models can lead to biases in parameter estimation [299]. Although unlikely to be important in the EMRI regime, for IMRIs it may be important to model GW memory effects [300]. These effects have recently been computed in the small-mass-ratio limit for quasicircular inspirals [301,302]. In general, including these effects in FEW is a matter of interpolating over an additional set of amplitudes.

5. Postadiabatic accuracy: Including the second-order GSF

Current second-order GSF calculations are carried out via a multiscale framework [61,84–86]. Within this approach, the trajectory through the parameter space is computed using an extended version of Eq. (2) that takes the form

$$\frac{d\alpha}{dt} = \frac{\nu}{M} [\hat{f}_\alpha^{(0)}(a, p, e, x_I) + \nu \hat{f}_\alpha^{(1)}(a, p, e, x_I) + \mathcal{O}(\nu^2)]. \quad (17)$$

The postadiabatic forcing functions $\hat{f}_\alpha^{(1)}$ are very computationally expensive to compute, but they can be precomputed and interpolated in the same manner as the adiabatic forcing functions. Second-order GSF calculations also compute the second-order waveform amplitudes. It is expected that these are not necessary to include in EMRI waveforms [149], but they may be required when modeling low- q IMRIs to avoid biases in parameter estimation. In that case, the second-order amplitudes will be interpolated along the sparse trajectory during online waveform generation alongside the adiabatic amplitudes. This means that online waveform generation of postadiabatic waveforms will be almost as fast as adiabatic models. At present, the $\hat{f}_\alpha^{(1)}$ are only known for quasicircular inspirals into a nonrotating black hole [164], and these results are ready to be incorporated into FEW (as was initially investigated in Ref. [149]). It is an active area of research to compute $\hat{f}_\alpha^{(1)}$ for more complex orbital configurations [303–306]. In addition to post adiabatic contributions to the forcing functions, second-order corrections to the mode amplitudes may also be important for loud IMRIs (but not for EMRIs, see Ref. [149] for discussion). Provided that GW fluxes are known to second order, the extension to second-order amplitudes is straightforward, being of the form

$$\mathcal{A}_{\ell m k n} = \mathcal{A}_{\ell m k n}^{(0)} + \nu \mathcal{A}_{\ell m k n}^{(1)}, \quad (18)$$

where $\mathcal{A}_{\ell m k n}^{(0)}$ are the adiabatic amplitudes described in Sec. III B. Postadiabatic amplitude corrections may also include contributions from the spin of the secondary object.

6. Post-adiabatic accuracy: Including the spin on the secondary

The effect of the spin on the secondary enters the waveform phase at postadiabatic order. In terms of waveform generation, a spinning secondary introduces another forcing term to Eq. (17) and the orbital frequencies are also modified. Both of these changes can readily be incorporated into FEW. At present, codes exist that can compute the necessary forcing terms for circular [307–311], eccentric [216], and generic orbits [161–163]. Last, while spin precession of the secondary object enters at second post-adiabatic order, it also contributes to mode phasing and amplitude evolution differently from other terms of this order [86]; the FEW framework provides the necessary tools for the importance of secondary spin precession to be examined in future work. With the exception of GW memory and generic orbits, work is already underway on all of these extensions (and we note that FEW currently can generate generic inspirals using a 5PN model [122]). In addition to the above, FEW can also be extended to include environmental and beyond-GR physics. So long as any additional physics can be modeled as a forcing function added to Eq. (2) or Eq. (17) that do not depend on the orbital phase (or if such dependence can be transformed away [165,166,169,200]) it can readily be incorporated into our framework. As the FEW trajectory module now exists entirely in PYTHON, if software for computing these modifications has already been developed it can be used directly to rapidly produce a new inspiral (and hence, waveform) model suitable for subsequent data analysis investigations. See, e.g., Refs. [50] and [35] for first efforts in this direction.

Last, in addition to these avenues for the development of more accurate and generic models, there is also significant scope to further accelerate waveform generation with the FEW framework. As we identified in Sec. VA, $\sim 90\%$ of the computational cost of the waveform derives from trajectory integration and waveform summation; there is scope to accelerate these two components in future work. The FEW integrator has been constructed entirely of vectorized array operations in preparation for batched trajectory integration on either CPU or GPU (depending on the batch size). While ODE integration is inherently serial, by computing derivatives for many trajectories in parallel the time per trajectory can be reduced significantly (potentially by orders of magnitude with GPUs and batch sizes $\gtrsim 10^2$). There are also multiple potential avenues toward improving the efficiency of the mode summation. While FEW does perform online mode selection, this operation is both expensive

(requiring a full set of amplitudes to be computed) and suboptimal (as can be verified with a slow, greedy-type minimization of waveform mismatch). Improving this procedure may yield reduced computational cost without compromising the results of inference. Additionally, as is highlighted by the compact nature of each waveform mode in Fig. 1, time-frequency methods may significantly improve the mode summation for I/EMRI waveforms by more than an order of magnitude. This also presents an opportunity to integrate the application of the LISA response—itself an expensive operation—with this summation operation, directly computing TDI variables and further reducing computational costs. Development of these techniques for implementation in FEW is currently ongoing [312]. Bringing together the myriad potential improvements to be made to the efficiency of FEW, per-waveform wall times of ~ 1 ms are a realistic possibility, the achievement of which would greatly impact the identification and analysis of IMRI and EMRI signals.

We thank Ryuichi Fujita and Adam Pound for interesting discussions. We also thank Barry Wardell for support in setting up the FEW data download server, and Patrick Bourq and Christopher Whittall for improving the FEW tutorials. We also thank the participants of the 2025 FEW Hackathon, hosted by the University of Southampton and funded by A. Pound’s Royal Society University Research Fellowship, for supporting key code developments in FEW and the production of this paper. O. B. acknowledges computational resources from the French space agency CNES in the framework of LISA. O. S. K. acknowledges the computing resources accessed from NUS IT Research Computing group. Part of the analysis was performed on the “Symmetry” HPC at Perimeter Institute, using NVIDIA H200 GPU devices. S. A. H. acknowledges the use of MIT Kavli Institute resources at MIT’s ENGAGING cluster, SUBMIT resources of the MIT Department of Physics, and the Open Science Grid.

ACKNOWLEDGMENTS

We are also indebted to BHPC (in particular Ryuichi Fujita, Hiroyuki Nakano, Norichika Sago, and Masaru Shibata) for providing us with independent analytical and numerical data for the forcing functions and mode amplitudes used in Sec. IV and Appendix E 2 for the validation, upon request. C. E. A. C.-B. acknowledges support from UKSA Grant No. UKRI971. L. S. would like to acknowledge the support of the European Space Agency through ESA’s postdoctoral Research Fellowship programme. Z. N. acknowledges support from the ERC Consolidator/UKRI Frontier Research Grant GWModels (selected by the ERC and funded by UKRI [Grant No. EP/Y008251/1]). O. B. acknowledges financial support from the Grant No. UKRI972 awarded via the UK Space Agency. O. S. K. acknowledges the support of the NUS Research Scholarship. J. M. and A. J. K. C. acknowledge support

from the NUS Faculty of Science, under the research Grant No. 22-5478-A0001. H. K. acknowledges the support from the Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Colleges and Universities. J. E. T. and A. J. K. C. acknowledge support from the NASA LISA Preparatory Science Grant No. 20-LPS20-0005. S. I. and A. J. K. C. acknowledge support from the Ministry of Education, Singapore, under the Academic Research Fund Tier 1 A-800149200-00 (FY2023). S. A. H. acknowledges support from NSF Grants No. PHY-2110384 and No. PHY-2409644. N. W. acknowledges support from a Royal Society—Research Ireland University Research Fellowship. This publication has emanated from research conducted with the financial support of Research Ireland under Grants No. 16/RS-URF/3428, No. 17/RS-URF-RG/3490 and No. 22/RS-URF-R/3825. This research was also done using services provided by the OSG Consortium [234,313–315], which is supported by the National Science Foundation Awards No. 2030508 and No. 2323298.

C. E. A. C.-B.: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing—original draft, writing—review and editing. L. S.: Conceptualization, formal analysis, investigation, methodology, supervision, validation, visualization, writing—original draft, writing—review and editing. Z. N.: Conceptualization, data curation, formal analysis, investigation, methodology, resources, software, validation, visualization, writing—original draft, writing—review and editing. O. B.: Formal analysis, investigation, methodology, validation, visualization, writing—original draft, writing—review and editing. M. L. K.: Conceptualization, investigation, methodology, resources, software, supervision, writing—review and editing. A. S.: Formal analysis, investigation, methodology, visualization, writing—original draft. S. K.: Formal analysis, investigation, validation, visualization, software, writing—original draft, writing—review and editing. P. L.: Investigation, validation, visualization, software, writing—original draft, writing—review and editing. J. M.: Formal analysis, investigation, validation, visualization, writing—original draft, writing—review and editing. H. K.: Investigation, validation, visualization, writing—original draft, writing—review and editing. J. E. T.: Validation, visualization, writing—reviewing and editing. S. I.: Data curation, formal analysis, investigation, supervision, validation, writing—original draft, writing—review and editing. S. A. H.: Conceptualization, data curation, formal analysis, investigation, methodology, validation, writing—review and editing. N. W.: Conceptualization, writing—original draft, writing—review and editing. A. J. K. C.:

Conceptualization, supervision, writing—review and editing. M. P.: Software.

DATA AVAILABILITY

The data that support the findings of this article are openly available [316,317], embargo periods may apply.

APPENDIX A: MASS CONVENTIONS

The previous implementation of FEW scaled quantities with respect to the primary mass m_1 and the small mass ratio ϵ . As aforementioned, in this work we choose a symmetric mass convention, which instead scales our model with respect to the total mass M and the symmetric mass ratio ν . To relate these different conventions, we first reexpress our equations of motion so that they are independent of mass

$$\frac{d\alpha}{d\hat{t}} = \hat{f}_\alpha(a, p, e, x_I) + O(\epsilon), \quad (\text{A1a})$$

$$\frac{d\hat{\Phi}_A}{d\hat{t}} = \hat{\Omega}_A(a, p, e, x_I) + O(\epsilon), \quad (\text{A1b})$$

where

$$\hat{t} = \epsilon t / m_1, \quad \hat{\Phi} = \epsilon \Phi, \quad (\text{A2a})$$

$$\hat{\Omega} = m_1 \Omega, \quad \hat{f}_\alpha = m_1 f_\alpha. \quad (\text{A2b})$$

Additionally, we take (p, e, x_I) to be dimensionless and define them by their implicit relationships to the orbital frequencies $\Omega(a, p, e, x_I)$. More concretely, (p, e, x_I) is related to the orbital parameters in Ref. [201], which we denote as $(p^F, e^F, \theta_{\text{inc}}^F)$, by

$$p = p^F, \quad e = e^F, \quad x_I = \cos \theta_{\text{inc}}^F, \quad (\text{A3})$$

while it is related to those in [218], which we denote as (r_1^V, r_2^V, z_1^V) , by

$$p = \frac{2r_1^V r_2^V}{m_1(r_1^V + r_2^V)}, \quad e = \frac{r_1^V - r_2^V}{r_1^V + r_2^V}, \quad x_I = \sqrt{1 - (z_1^V)^2}, \quad (\text{A4})$$

and, to those in [200], which we refer to as $(p^{\text{LB}}, e^{\text{LB}}, x_I^{\text{LB}})$, by

$$p = \frac{p^{\text{LB}}}{m_1}, \quad e = e^{\text{LB}}, \quad x_I = x_I^{\text{LB}}. \quad (\text{A5})$$

If we now consider the limit $m_1 \gg m_2$, then $\epsilon = \nu + O(\nu)$ and $m_1 = M[1 + O(\nu)]$, giving us

$$\hat{t} = \nu t / M + O(\nu^2), \quad \hat{\Phi} = \nu \Phi + O(\nu), \quad (\text{A6a})$$

$$\hat{\Omega} = M\Omega + O(\nu), \quad \hat{f}_\alpha = Mf_\alpha + O(\nu). \quad (\text{A6b})$$

Inserting Eq. (A6) into Eq. (A1) and dropping all higher-order terms, we arrive at a new set of equations of motion, given by Eqs. (2) and (3). These equations are consistent with Eq. (A1) at leading adiabatic order, but introduce new higher-order postadiabatic corrections. Note that, based on Eq. (A1), if we compute an adiabatic trajectory using our symmetric mass ratio convention, we can then relate it to an adiabatic trajectory produced via Eq. (A1) by rescaling time by $m_1/M \times \nu/\epsilon$ and the phase by ν/ϵ . Alternatively, one can simply replace m_1 and m_2 by M and μ in the small mass ratio model to produce the same trajectory as FEW.

Our mass convention also affects the amplitudes of the waveform, which we scale in terms of the reduced mass $\mu = m_2[1 + O(\nu)]$. In practice, this rescaling of the strain induces a change that is on the same order as the interpolation error in our amplitudes. Therefore, it has a negligible impact when compared to the amplitudes of other adiabatic models. Ultimately, if one wants to compare other models to FEW, but these models use the small mass ratio convention, then one needs to use the masses M and μ as input for the small mass ratio model and m_1 and m_2 as input in FEW to produce comparable waveforms.

APPENDIX B: DATA GRIDS AND COORDINATE PARAMETRIZATIONS

In our model, the simulation domain is divided into two regions: an inner region (*Region A*) and an outer region (*Region B*), following the approach in [318]. Region A is compact in the p dimension and designed to densely sample amplitude and flux data near the separatrix, where these quantities exhibit rapid variation. To manage computational cost, we taper the maximum eccentricity as we approach the separatrix, thereby avoiding highly eccentric orbits near plunge, which are exponentially more expensive to simulate (see Appendix B 4 and Fig. 21). This tapering is also astrophysically motivated: gravitational radiation tends to circularize orbits, so we expect eccentricities to decrease as the separatrix is approached. Region B spans a much broader range in p and e , and is sampled according to the PN scaling behavior of fluxes and amplitudes. In the following Subsections, we detail the coordinate systems and sampling strategies used to construct our flux and amplitude datasets.

1. Fluxes

For the fluxes, the inner grid in Region A has bounds

$$\delta p_{\text{min}}^A \leq p - p_{\text{sep}}(a, e) \leq \delta p_{\text{max}}^A, \quad (\text{B1})$$

$$0 \leq e \leq S_{\text{ecc}}(a, p, e; e_{\text{sep}}, e_{\text{max}}), \quad (\text{B2})$$

$$a_{\text{min}} \leq a \leq a_{\text{max}}, \quad (\text{B3})$$

with $x_I = 1$ fixed and

$$\delta p_{\min}^A = 0.001, \quad \delta p_{\max}^A = 9 + \delta p_{\min}^A, \quad (\text{B4})$$

$$e_{\text{sep}} = 0.25, \quad e_{\max} = 0.9, \quad (\text{B5})$$

$$a_{\min} = -0.999, \quad a_{\max} = 0.999. \quad (\text{B6})$$

As aforementioned, the curve $S_{\text{ecc}}(a, p, e)$ tapers the maximum value of the eccentricity from $e = 0.9$ at larger p values to $e = 0.25$ as we approach the separatrix. To define S_{ecc} , we first transform to a new coordinate system (u, w, y, z) , which—for generic orbits—is related to (a, p, e, x_I) via

$$u = \left[\frac{\ln(p - p_{\text{sep}}(a, e, x_I) + C_p) - C_\Delta}{\ln 2} \right]^\alpha, \quad (\text{B7a})$$

$$w = \frac{e}{S_{\text{gen}}(a, p, e, x_I)}, \quad (\text{B7b})$$

$$y = \frac{x_I - x_{\min}}{1 - x_{\min}}, \quad (\text{B7c})$$

$$z = \frac{\hat{\chi}(a) - \hat{\chi}_{\min}}{\hat{\chi}_{\max} - \hat{\chi}_{\min}}, \quad (\text{B7d})$$

where

$$\hat{\chi}(a) = (1 - a)^{1/3}, \quad \hat{\chi}_{\min} = \hat{\chi}(a_{\max}), \quad \hat{\chi}_{\max} = \hat{\chi}(a_{\min}),$$

and $S_{\text{gen}}(a, p, e, x_I) = \tilde{S}[u(a, p, e, x_I), z(a)]$ with

$$\tilde{S}(u, z) = e_{\text{sep}} + (e_{\max} - e_{\text{sep}}) \sqrt{z + u^\beta(1 - z)}, \quad (\text{B8})$$

providing the eccentricity tapering mentioned at the beginning of Appendix B. Furthermore, we have introduced the constants

$$C_p = \delta p_{\max}^A - 2\delta p_{\min}^A, \quad C_\Delta = \ln(\delta p_{\max}^A - \delta p_{\min}^A),$$

$$\alpha = 1/2, \quad \beta = 1/2,$$

and $x_{\min}$ is a free parameter that only has to be defined for inclined orbits. Note that the coordinates are defined so that (u, w, y, z) all lie in the domain $[0, 1]$.

For this work, we set $x_I = 1 = y$, and to handle retrograde orbits, we set $a < 0$ rather than switching the sign of x_I . Then we have $S_{\text{ecc}}(a, p, e) = \tilde{S}_{\text{gen}}(a, p, e, x_I = 1)$. We compute the fluxes on a three-dimensional rectilinear grid by uniformly sampling the coordinates (u, w, z) over the unit cube (i.e., $[0, 1]^3$). The grid consists of $N_u \times N_w \times N_z$ points, with $(N_u, N_w, N_z) = (129, 65, 65)$.

The outer grid in Region B is defined by the bounds

$$\delta p_{\min}^B \leq p - p_{\text{sep}}(a, e, x_I) \leq \delta p_{\max}^B, \quad (\text{B9})$$

$$0 \leq e \leq e_{\max}, \quad (\text{B10})$$

$$-a_{\max} \leq a \leq a_{\max}, \quad (\text{B11})$$

with $e_{\max}, a_{\max}$ being the same values used in Region A, and

$$\delta p_{\min}^B = \delta p_{\max}^A - \delta p_{\min}^A, \quad (\text{B12})$$

$$\delta p_{\max}^B = 200. \quad (\text{B13})$$

Like Region A, we construct a new coordinate system (U, W, Y, Z) , which is related to (a, p, e, x_I) via

$$U = \frac{(\delta p_{\min}^B)^{-1/2} - [p - p_{\text{LSO}}(a, e, x_I)]^{-1/2}}{(\delta p_{\min}^B)^{-1/2} - [p_{\max}^B - p_{\text{LSO}}(a, e, x_I)]^{-1/2}}, \quad (\text{B14})$$

$$W = \frac{e}{e_{\max}}, \quad (\text{B15})$$

$$Y = \frac{x_I - x_{\min}}{1 - x_{\min}}, \quad (\text{B16})$$

$$Z = \frac{\chi(a) - \chi_{\min}}{\chi_{\max} - \chi_{\min}}. \quad (\text{B17})$$

Once again, with $Y = 1$ fixed, we produce a uniformly spaced rectilinear grid with $(N_U, N_W, N_Z) = (65, 33, 33)$ points in (u, w, z) .

2. Amplitude grid

For the amplitudes, we use the same bounds for Region A as we did for the fluxes [which are given in Eq. (B4)]. To construct the inner amplitude grid, we also use the same coordinates as the fluxes, (u, w, y, z) in Eq. B 1, but with the modified constants

$$\alpha = 1/3, \quad \beta = 3. \quad (\text{B18})$$

As before, we sample amplitudes on a uniformly spaced rectilinear grid with $(N_u, N_w, N_z) = (33, 33, 33)$ points in (u, w, z) .

When constructing amplitudes in the outer domain, we use slightly different bounds for Region B as compared to the fluxes,

$$\delta p_{\min}^B = \delta p_{\max}^A - 0.001,$$

$$\delta p_{\max}^B = 200 + \delta p_{\min}^B.$$

We also use modified grid coordinates

$$\hat{U} = \frac{\hat{U}_{p,\max}(a, e, x_I) - \hat{U}_p(p)}{\hat{U}_{p,\max}(a, e, x_I) - \hat{U}_{p,\min}(a, e, x_I)}, \quad (\text{B19})$$

$$\hat{W} = \frac{e}{e_{\max}}, \quad (\text{B20})$$

$$\hat{Y} = y(x_I), \quad (\text{B21})$$

$$\hat{Z} = z(a), \quad (\text{B22})$$

with $\hat{U}_p(p) = p^{-1/2}$, $\hat{U}_{p,\max}(a, e, x_I) = \hat{U}_p[\delta p_{\min}^B + p_{\text{sep}}(a, e, x_I)]$, $\hat{U}_{p,\min}(a, e, x_I) = \hat{U}_p[\delta p_{\max}^B + p_{\text{sep}}(a, e, x_I)]$. Note that $(\hat{W}, \hat{Y}, \hat{Z})$ are identical to the outer grid coordinates (W, Y, Z) used for the fluxes. To reduce the memory footprint of our interpolant by 37.5%, we reduce the sampling density with respect to Region A, such that $(N_{\hat{U}}, N_{\hat{W}}, N_{\hat{Z}}) = (17, 17, 33)$ points in $(\hat{U}, \hat{W}, \hat{Z})$ —the density with respect to Z remains fixed for the purposes of vectorized linear interpolation.

There are minor differences between the flux and amplitude grids, stemming from the fact that the data were generated in two stages: amplitudes first, followed by fluxes. In the initial amplitude grid, we used $\alpha = 1/3$, which concentrated sampling density in Region A near the separatrix. This was necessary due to the sparse nature of the amplitude data in that region. However, when applying this same sampling strategy to the fluxes, which were sampled more densely overall, it led to oversampling near the separatrix with $\Delta p \lesssim 10^{-5}$ for the first several points in u . We found that increasing α to 1, while keeping the product $\alpha\beta = 1$ fixed, provided a more balanced sampling of the fluxes across all of Region A. This adjustment preserved the same eccentricity tapering used for the amplitude grid, ensuring consistency between the two datasets.

Furthermore, we set $\delta p_{\min}^B = \delta p_{\max}^A$ for the fluxes to create a small region of overlap between A and B, which we used for preliminary cross-validation of interpolated flux values in the shared domain. We also found that using U for the fluxes, instead of $\hat{U}$, and truncating Region B at $\delta p_{\min}^B$ and $\delta p_{\max}^B$ in (B12) provided coordinates that much more closely align with the PN behavior of the fluxes at large p . This change resulted in a modest but noticeable improvement in flux interpolation in the outer region.

3. Solving for edges of domain

When evaluating interpolated functions of our data grids, it is useful to identify the edges of our domain. If we truncate the evaluation of our amplitude grid at the same maximum value of p as the fluxes, then we have

$$p_{\text{lower}}(a, e) = p_{\text{sep}}(a, e, x_I = 1) + \delta p_{\min}^A, \quad (\text{B23})$$

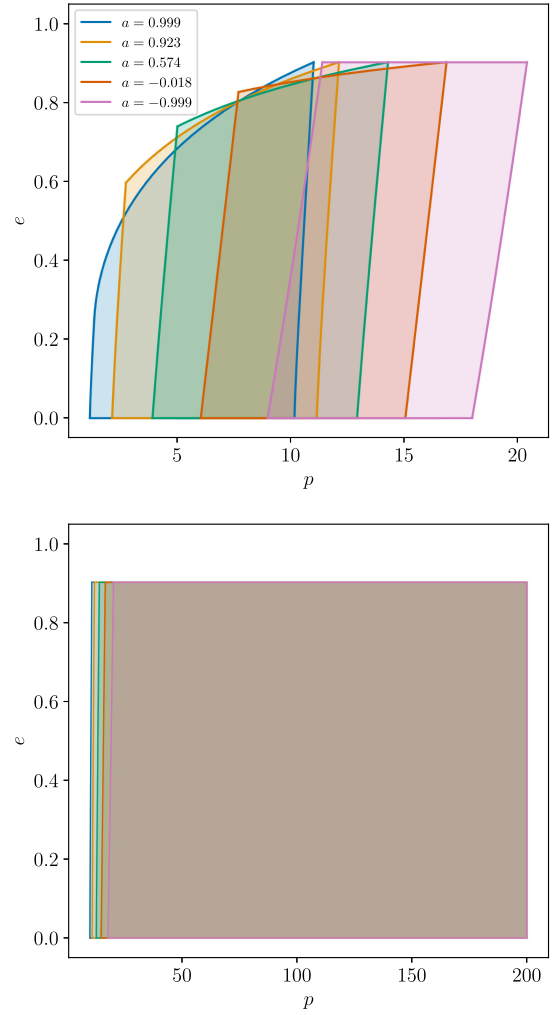


FIG. 20. Visualization of the flux grid domains, plotted in the (p, e) plane for various fixed spin parameters $a = \{0.999, 0.923, 0.574, -0.018, -0.999\}$. The top panel shows the inner grid (Region A), and the bottom panel shows the outer grid (Region B).

$$p_{\text{upper}}(a, e) = 200, \quad (\text{B24})$$

$$a_{\text{lower}}(p, e) = a_{\min}, \quad (\text{B25})$$

$$a_{\text{upper}}(p, e) = a_{\max}, \quad (\text{B26})$$

$$e_{\text{lower}}(a, p) = 0, \quad (\text{B27})$$

where the “lower” and “upper” subscripts refer to the minimum and maximum values of a , p , or e , respectively, on our grid when the two other orbital parameters are fixed.

The process for finding maximum or “upper” e given a constant (a, p) is slightly more involved due to the tapered eccentricity bound used in Region A. To identify $e_{\max}(a, p)$, we use the following method:

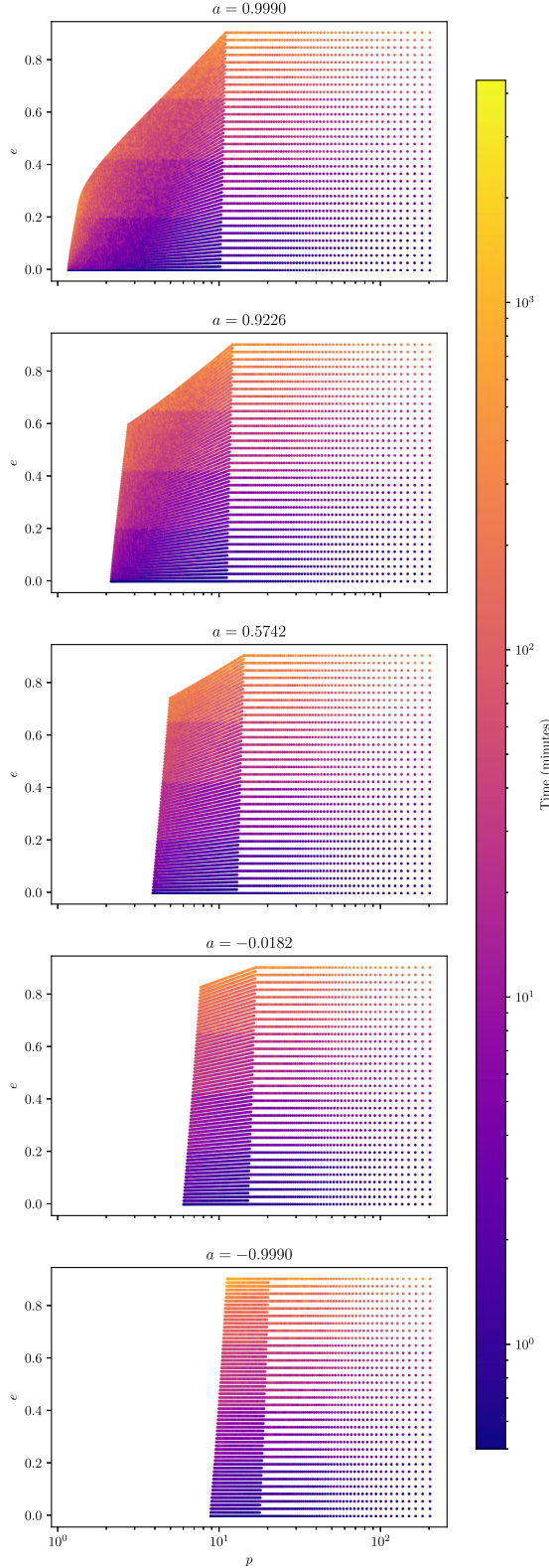


FIG. 21. Number of CPU minutes required to calculate the fluxes at each grid point for the same slices of a visualized in Fig. 20.

- (1) if $p \geq p_{\text{sep}} + \delta p_{\text{min}}^B$, then $e_{\text{upper}} = 0.9$, otherwise.
- (2) we solve for the root e_{lower}^0 that satisfies

$$w(a, p, e_{\text{lower}}^0) - 1 = 0$$

$$= e_{\text{lower}}^0 / \tilde{S}[u(a, p, e_{\text{lower}}^0, 1), z(a)],$$

in the interval $e_{\text{lower}}^0 \in (0, e_{\text{max}})$;

- (3) if $p_{\text{sep}}(a, e_{\text{lower}}^0, 1) < p$, then $e_{\text{lower}} = e_{\text{lower}}^0$;
- (4) else, this is a false root, and e_{lower} is instead the root of the equation $p - p_{\text{sep}}(a, e_{\text{lower}}, 1) = 0$ in the interval $e_{\text{lower}} \in (0, e_{\text{max}}^0)$.

4. Timing of flux calculations

Based on the grids and sampling rates described above, we compute the fluxes at 615,810 points in the (a, p, e) parameter space. We perform these calculations on the Open Science Grid, a free open-access high throughput computing infrastructure [234,313–315]. In Fig. 21 we plot the computational cost of each flux calculation for constant values of a . The computational cost is most significantly impacted by the value of the eccentricity and grows exponentially with e . However, this growth is not smooth due to the nature of the flux calculations performed by PYBHPT. As the eccentricity grows, PYBHPT needs to increase the numerical resolution of various computations, which it does by increasing the sampling rate by factors of two. Thus, at certain regions in parameter space, the calculation effectively doubles in computational cost because the calculation is hard coded to run at double the numerical resolution to meet precision requirements. This leads to the sharp horizontal bands in Fig. 21. Furthermore, particularly for large prograde spins, we observe a rapid growth in computational cost as we approach the separatrix. All together, the flux grid took just under 605,000 CPU hours to calculate.

APPENDIX C: SETUP FOR MONTE CARLO STUDIES OF WAVEFORM CHARACTERISTICS

At multiple points throughout this work, we examine the accuracy and performance of our waveform model by performing Monte Carlo draws of source parameters. For consistency, we randomly sample these parameters from the same distributions in each case; these are summarized in Table III. Given a set of these drawn parameters, p_0 is chosen according to the desired time to plunge of the inspiral in question (typically either two or four years). For parameter sets where no such point can be found in our domain of validity (e.g., highly eccentric systems with $e \lesssim 10^{-6}$), a new point is drawn to replace it. This rejection procedure yields the more complicated distribution that appears in the top panel of Fig. 12.

TABLE III. Distributions used for the Monte Carlo studies of waveform characteristics performed in this work. Given a set of these drawn parameters, p_0 is chosen according to the desired time to plunge of the inspiral in question; see text of Appendix C for discussion.

Parameter	Unit	Min	Max	Distribution
m_1	$M_\odot$	10^5	10^7	Uniform in $\log m_1$
m_2	$M_\odot$	10^5	10^7	Uniform in $\log m_2$
a	$\dots$	0	0.999	Uniform
e_0	$\dots$	0	0.9	Uniform

For waveform-level comparisons, angular parameters are distributed in the typical manner: azimuthal and polar angles are distributed uniformly on the surface of the unit sphere, and phase parameters are uniformly distributed in the range $[0, 2\pi)$. As none of our Monte Carlo analyses entail random sampling of d_L , we set this parameter to a fiducial value of 1 Gpc in all cases.

Notably, Table III only covers prograde inspirals. The retrograde half of the parameter space is significantly less challenging to model accurately at all stages of waveform generation. We therefore exclude this region from many of our Monte Carlo studies and focus primarily on prograde inspirals. None of the conclusions presented in this work change significantly when considering the retrograde case (to the contrary, the median performance of our model over the parameter space improves with such an extension).

APPENDIX D: GW DATA ANALYSIS FRAMEWORK

The strain data stream observed by LISA will be a superposition of multiple GW signals of various source types that overlap in both time and frequency [319]. These signals will be buried in a mixture of both instrumental noise and astrophysical confusion noise (particularly the Galactic foreground [53]). For simplicity, we assume that the data stream $d(t)$ contains a single EMRI signal in additive noise, such that

$$d(t) = h(t; \theta_{\text{tr}}) + n(t), \quad (\text{D1})$$

where $h(t; \theta_{\text{tr}})$ is the EMRI signal described by the parameters $\theta = \theta_{\text{tr}}$ (which we wish to infer in data analysis). This assumption is well motivated by the negligible overlap between EMRIs and other sources in the data stream [255].

The probabilistic noise process $n(t)$ determines the form of the likelihood function used in the inference of θ . In our analysis, we assume that $n(t)$ is Gaussian, (weakly) stationary, and circulant. While the cyclostationarity of the Galactic foreground signal invalidates this assumption [320,321], including this feature in our noise model will significantly increase the computational cost of our analyses

without greatly affecting any of our conclusions, so we do not consider this noise feature in this work. Our assumptions regarding the noise process leads to a diagonal noise-covariance matrix in the Fourier domain [322],

$$\Sigma(f, f') = \langle \tilde{n}(f) \tilde{n}^*(f') \rangle = \frac{1}{2} \delta(f - f') S_n(f') df', \quad (\text{D2})$$

with Fourier-domain convention

$$\tilde{a}(f) = \int_{-\infty}^{\infty} a(t) \exp(-2\pi i f t) dt. \quad (\text{D3})$$

The PSD $S_n(f)$ describes the average mean-square fluctuations of $\tilde{n}(f)$ as a function of f . The diagonal noise covariance matrix defined in Eq. (D2) gives rise to the Whittle likelihood function that is ubiquitous in the field of GW data analysis. The log of this likelihood function is

$$\log \mathcal{L}(d|\theta) \propto -\frac{1}{2} (d - h_m | d - h_m), \quad (\text{D4})$$

with the noise-weighted inner product

$$(a|b) = 4\text{Re} \int_0^\infty \frac{\tilde{a}^*(f) \tilde{b}(f)}{S_n(f)} df, \quad (\text{D5})$$

where $\star$ denotes complex conjugation. The quantity $h_m(t; \theta) \equiv h_m$ represents a template from the waveform model used to perform inference on the signal h in the data stream. Deviations between models for the signal and template waveforms yield systematic biases: for $h_m(t; \theta) \neq h(t; \theta)$, a parameter estimation analysis would recover biased parameters θ_{bf} . This will be discussed later. The difference between two waveforms h_1 and h_2 can be expressed in terms of the normalized overlap and mismatch, respectively,

$$O(h_1, h_2) = \frac{(h_1|h_2)}{\sqrt{(h_1|h_1)(h_2|h_2)}}, \quad (\text{D6})$$

$$\mathcal{M}(h_1, h_2) = 1 - O(h_1, h_2). \quad (\text{D7})$$

The mismatch $\mathcal{M}$ is equal to zero for $h_1 \propto h_2$ and equal to one if the two waveforms h_1 and h_2 are orthogonal. We will use Eq. (D7) extensively in Sec. IV C as a figure of merit for the faithfulness of our waveform model.

A useful quantity that determines the strength of an EMRI signal h with respect to the noise is the optimal matched-filter SNR

$$\rho = \sqrt{(h|h)}. \quad (\text{D8})$$

Equation (D8) is optimal in the sense that it is the mean SNR (averaging over all noise realizations) under the assumption that the template waveform model is correct.

Statistical biases in measured parameters due to noise fluctuations scale inversely with ρ , whereas systematic biases due to modeling errors do not; considering the range of expected SNRs for typical EMRIs is therefore an important facet of our analysis of waveform systematics in Sec. IV C.

Data analysis in GW astronomy is typically performed with Bayesian inference techniques, at the heart of which lies Bayes's theorem,

$$p(\theta|d) = \frac{\pi(\theta)\mathcal{L}(\theta|d)}{\mathcal{Z}(d)}, \quad (\text{D9})$$

where $\pi(\theta)$ is the prior probability, and the normalizing constant $\mathcal{Z}$ is usually denoted as the evidence (and is independent of θ). The posterior distribution $p(\theta|d)$ encodes all information regarding the values of θ that are best supported by the observational data, given our prior expectations; insight regarding the form of $p(\theta|d)$ is attained via parameter estimation. In Sec. V F 1, we perform MCMC analyses (as implemented by the `ERYN` package [271–273], with default settings) to sample EMRI posterior distributions, from the results of which we compute parameter variances as summary statistics. We assume wide uniform priors on all sampled parameters (such that they do not truncate the posterior bulk). Initial MCMC walker positions are chosen to be near the injected EMRI parameters, as our goal is not to perform a search, only to probe the form of the posterior bulk. To incorporate the response function of LISA in our analyses, we use the `FASTLISARESPONSE` package [270] to generate second-generation $\{A, E, T\}$ TDI variables [323] from our EMRI waveforms, assuming spacecraft orbits with constant and equal arm lengths. Our model for $S_n(f)$ is the `SCIRDV1` sensitivity curve [324] for this TDI variable convention, including the effects of the Galactic confusion noise [325] as a stationary noise process [326].

Despite the computational efficiency of our model on GPUs, sampling Eq. (D9) still typically takes ~ 12 hr per EMRI source. This is prohibitively expensive for more exhaustive studies of parametric biases due to waveform model differences, such as those performed in Sec. V C, in which a signal with parameters θ_{tr} is injected with one model and analyzed with another, returning (potentially biased) best-fit parameters θ_{bf} . To facilitate such an analysis, we estimate these biases via the linear signal approximation given by Cutler and Vallisneri [238],

$$\Delta\theta \approx (\Gamma^{-1}(\theta_{\text{tr}}))^{ij}(\partial_j h_{\text{ap}}(\theta_{\text{tr}})|h_{\text{tr}}(\theta_{\text{tr}}) - h_{\text{ap}}(\theta_{\text{tr}})), \quad (\text{D10})$$

where $\Delta\theta := \theta_{\text{bf}} - \theta_{\text{tr}}$ and $\partial_j h := \partial h / \partial \theta_j$, and the $\approx$ indicates we have ignored higher-order terms. The information matrix Γ has the elements

$$\Gamma_{ij} := (\partial_i h_{\text{ap}} | \partial_j h_{\text{ap}}) \quad (\text{D11})$$

evaluated in Eq. (D10) at θ_{tr} .

Parameters estimated from noisy observations will be modified by a probabilistic shift away from the true parameter values. In a similar vein to Eq. (D10), at leading order this shift is approximately

$$\Delta\theta_{\text{noise}} \approx (\Gamma^{-1}(\theta_{\text{tr}}))^{ij}(\partial_j h_{\text{ap}}(\theta_{\text{tr}})|n_{\text{nr}}), \quad (\text{D12})$$

where n_{nr} is a specific noise realization. Since n_{nr} has zero mean, $\Delta\theta_{\text{noise}}$ is unbiased (i.e., $\mathbb{E}[\Delta\theta_{\text{noise}}] = 0$). For all of the parameter estimation analyses we perform in this work—both in the form of Cutler-Vallisneri bias estimates (Sec. V C) and MCMC explorations of posterior distributions (Sec. V F)—we assume a “zero-noise” realization, such that n_{nr} is the zero-vector. We make this choice to eliminate the statistical parameter shifts due to noise, leaving only the systematic shifts due to modeling errors (the characterization of which is the goal of these analyses). Our results should therefore be interpreted as the expectation over all noise realisations. A pertinent question is why the posteriors recovered in Sec. V F have nonzero widths, despite the apparent absence of noise. It is important to emphasize that the “zero-noise” realization is a perfectly valid draw from the distribution of all possible noise realizations, and that we should therefore expect our parameter estimates to be uncertain for zero noise according to our noise model (given by the PSD), just as we would expect to occur for any other realization.

The linear signal approximation is typically valid for EMRIs with detectable SNRs $\rho \gtrsim 20$ (which gives rise to the Gaussian posterior distributions obtained in Sec. V F). We use the `STABLEEMRIFISHER` package [327] to obtain robust information matrices given a set of EMRI parameters. Once the best-fit parameter point is obtained, we quantify its bias with respect to θ_{tr} by its corresponding sigma contour level. It is given by the Mahalanobis distance D_{Maha} [253], i.e. the multivariate generalization of the z score, and is defined as

$$D_{\text{Maha}}^2 := \frac{\Delta\theta^T \cdot \Gamma \cdot \Delta\theta}{D}. \quad (\text{D13})$$

We scale by $1/D$ here compared to the original definition since the unscaled Mahalanobis distance increases with dimensionality, which can lead to deceptively large sigma-contour levels. Consider, e.g., Γ to be a $D \times D$ identity matrix and $\Delta\theta$ a length D vector of 1's, representing a 1σ -biased point from the peak of a standard normal in D dimensions. The numerator of Eq. (D13) is D in this case, justifying the scaling by $1/D$ to recover $D_{\text{Maha}} = 1$. Finally, note that parameter space correlations also impact the Mahalanobis distance through Γ .

APPENDIX E: COMPARISON AGAINST OTHER IMPLEMENTATIONS IN LIMITING CASES

In addition to the self-contained examination of our FEWv2 carried out in Sec. IV, we can also verify that it correctly matches other models in limiting cases. Specifically, we compare against accurate models for quasicircular inspirals into spinning MBHs [48,124] (Appendix E 1), and verify that elements of our model approach PN-GSF based results in the weak-field limit ($p \gg p_{\text{sep}}$) with moderate eccentricities (Appendix E 2).

The results of the comparisons conducted in this Appendix, when combined with the systematic validation performed throughout the main body of the text, serve to confirm that our model produces accurate waveforms for spinning and/or equatorial IMRI and EMRI across its domain of validity.

1. Asymmetric-mass, quasicircular inspirals into rapidly spinning black holes

In order to compare the model developed in this work (FEWv2) against other implementations that are expected to be faithful in the strong-field regime, in this Appendix we benchmark its performance in the noneccentric limit against two existing Kerr quasicircular equatorial waveform models that are similarly accurate to adiabatic order. The first reference model KERRCIRC, detailed in [48], is an independent extension to earlier versions of FEW but using flux and amplitude data generated by a distinct Teukolsky solver. The second reference model is BHPWAVE [124], which provides waveforms from a completely independent generative framework. Our comparison focuses on evaluating the orbital phase evolution by comparing trajectories, examining interpolated waveform amplitudes, and analyzing the mismatch between the full waveforms as a function of the primary black hole's spin parameter.

Before computing the phase shifts, we rescaled the masses as discussed in Appendix A to have a similar mass convention across all three models. Moreover, since the overall scaling of the phases and phase errors are known with the mass ratio and with the primary mass, we only report the comparison as a function of the primary spin. We chose masses of $(m_1, m_2) = (10^6, 10)M_\odot$. Throughout this section, for each value of the primary spin parameter a , the inspiral is started at a corresponding p_0 such that the evolution lasts four years before reaching the plunge.

The absolute value of the final phase shift as a function of primary spin compared to both models is shown in Fig. 22. The phase shift $\Delta\Phi_\phi \lesssim 1$ for all spin values; we observe some oscillations alongside the overall increase for higher spins in prograde orbits. The primary source of these phase shifts lies in differences in flux and interpolation accuracy. These discrepancies arise from several factors, such as the accuracy of the flux data, choice of parametrization, the scaling applied to the flux data prior to interpolation, the

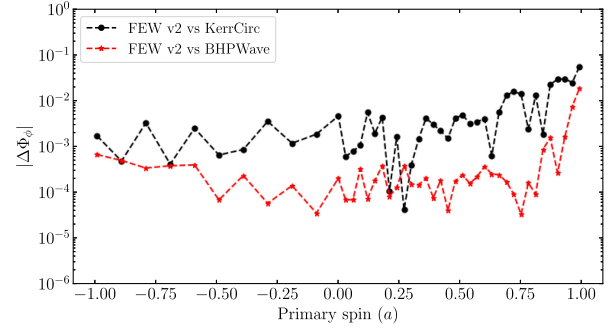


FIG. 22. Absolute value of the final orbital phase shift between our model and two other models, KERRCIRC (black dashed line) and BHPWAVE (red dashed line), in the circular orbit limit for different values of primary spin, $a \in [-0.99, +0.99]$. The phase shifts are for four-year inspirals and the masses are fixed to $(m_1, m_2) = (10^6, 10)M_\odot$.

type of interpolant used, and the interpolation grid itself, specifically, the density of said grid.

The KERRCIRC model employs a uniform grid in spin with $\Delta a = 0.01$, whereas BHPWAVE uses a nonuniform grid with a higher density of points concentrated at larger spin values. The latter setup is more similar to ours, in which the grid resolution is denser at higher spin and sparser at lower spin. This grid structure likely explains why BHPWAVE shows better agreement with our model. A detailed study of grid-induced and interpolation-induced errors as a function of spin is reported in [235].

Furthermore, we compared the interpolated amplitudes in our model for four different primary spin values, $a \in \{0.99, 0.5, -0.5, -0.99\}$ and three different modes, $(\ell, m) \in \{(2, 2), (5, 5), (10, 10)\}$ as a function of orbital separation, p , with both KERRCIRC, and BHPWAVE in Fig. 23. We dropped the index n here because in the circular orbit limit, there are no radial harmonics and $n = 0$. We plot $|\Delta\mathcal{A}_{\ell m 0}|$ instead of the fractional error, as the latter can appear misleadingly large for higher modes due to their smaller amplitudes. We observe that the mode amplitudes agree more closely between KERRCIRC and BHPWAVE than with our model. This is primarily because we do not employ a very dense grid in p for computing the amplitudes (Appendix B). We remind the reader that amplitude errors do not accumulate over the inspiral and therefore do not lead to significant deviations in the total mismatch between waveforms (this is observed during amplitude validation in Sec. IV B).

Finally, to provide an overall comparison of the full waveforms, we computed the mismatch (see Appendix D for definitions) between our waveform and both KERRCIRC and BHPWAVE as a function of the primary spin for $a \in [-0.99, 0.99]$ in the circular orbit limit. The mismatches were calculated using a flat PSD, such that differences between models are captured uniformly across all frequencies. As for the phase shifts, we fixed the source masses to

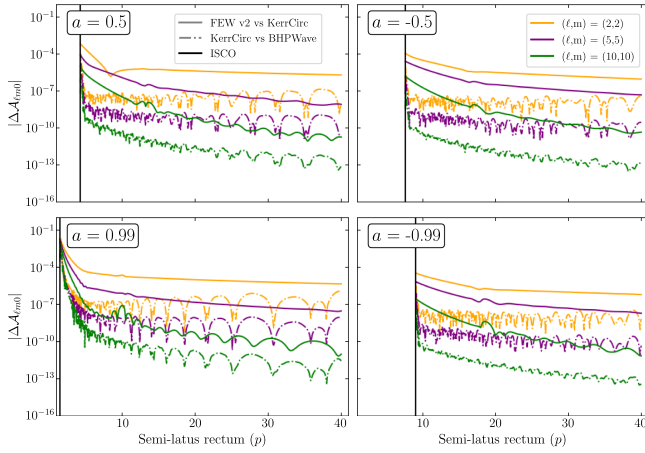


FIG. 23. Comparison of the real part of the mode amplitude differences, $|\Delta A_{\ell m 0}|$, as a function of orbital separation p . Each panel corresponds to a fixed primary spin value, $a \in 0.99, 0.5, -0.5, -0.99$, and shows three different modes, $(\ell, m) \in (2, 2), (5, 5), (10, 10)$. The solid lines compare our model to KERRCIRC, while the dashed lines compare KERRCIRC to BHPWAVE. The innermost stable circular orbit is indicated by a vertical marker.

$(m_1, m_2) = (10^6, 10)M_\odot$ and viewing angles are fixed to typical value for the viewing angles $(\theta, \phi) = (\pi/5, \pi/3)$. The results are shown in Fig. 24. The mismatch between our waveform and BHPWAVE remains below 10^{-5} for almost all spin values, in both prograde and retrograde orbits, except for a sharp increase as $a \rightarrow 0.99$. This increase coincides with the region where the phase shifts and amplitude differences are largest, as previously discussed. Typically, for larger spin values, as the system enters deeper into the strong-field regime, any discrepancies sourced from differences in amplitudes, fluxes, grid structures, or interpolation schemes become more pronounced.

In comparison, the mismatch with KERRCIRC is systematically larger, which is consistent with the phase shift trends observed in Fig. 22. It is worth mentioning that these mismatch values do not necessarily reflect the absolute accuracy of any of the models considered here, but rather indicate the level of agreement between them. In most cases, the discrepancies are small enough to be below the resolvable threshold of the detector and thus unlikely to impact observational outcomes. A detailed investigation into the sources of these differences is beyond the scope of this article and is presented separately in a study focusing on the systematics errors in circular equatorial waveforms [235].

2. Weak-field behavior against PN-GSF approaches

The combined PN-GSF approach of black hole perturbation theory yields closed-form analytic solutions to the (frequency-domain) Teukolsky equation and thus the inspiral forcing functions in Eq. (2) and the GW strain

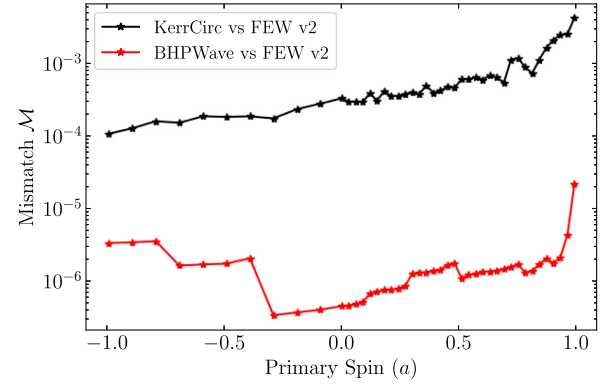


FIG. 24. Mismatch between our waveform and KERRCIRC (solid line) or BHPWAVE (dashed line) as a function of the primary spin a , in the circular orbit limit. Mismatches were computed using a flat PSD, with a four-year observation time before plunge (i.e. reaching the innermost stable circular orbit). The source parameters were fixed at $(M, \mu) = (10^6, 10)M_\odot$, and an arbitrary viewing configuration with $(\theta, \phi) = (\pi/5, \pi/3)$.

amplitudes in Eq. (5) in turn; we refer a reader to Refs. [138,139] for reviews. In the context of the EMRI and IMRI modeling, this approach is dependent upon order-by-order expansions in powers of the (dimensionless) velocity parameter: $v = \sqrt{1/p}$ (or similar parameters) and often eccentricity: e (while keeping the inclination and primary spin arbitrary), limiting the accuracy of the analytic expressions in the stronger field ($v \simeq 1$) and at larger eccentricities ($e \simeq 1$). At the same time, this method is particularly advantageous to account for a vast region of the weak-field parameter space ($v \ll 1$) that would be expensive to cover with numerical GSF computations.

Building upon the results reported in [114,140,328], Ref. [136] derived the analytic expressions for the inspiral forcing functions equivalent to Eq. (2) for generic (eccentric and precessing) adiabatic inspirals, valid to $\mathcal{O}(v^{11}, e^{11})$ relative to the leading “Newtonian-circular” terms. Shortly afterward, Ref. [122] combined that result with the analytic GW strain amplitudes (derived to the same PN accuracy) to produce a generic adiabatic waveform model. In this work, we label these analytical results simply “PN5” for brevity but we highlight the additional eccentricity expansion employed in their derivation. The PN5 expressions for both the spheroidal harmonic modes of the Teukolsky amplitudes and the inspiral forcing functions are publicly available from the Black Hole Perturbation Club (BHPC) website [231].¹¹

¹¹There are a number of other PN-GSF calculations that explicitly consider the eccentric equatorial inspirals in Kerr spacetimes [251,329–331]. However, the results obtained are lower orders in p and e expansion (sometimes with slow primary spin $a \ll 1$ assumed), or have yet to be implemented in FEW. For this reason, we have not used them as a weak-field benchmark for the results in this work.

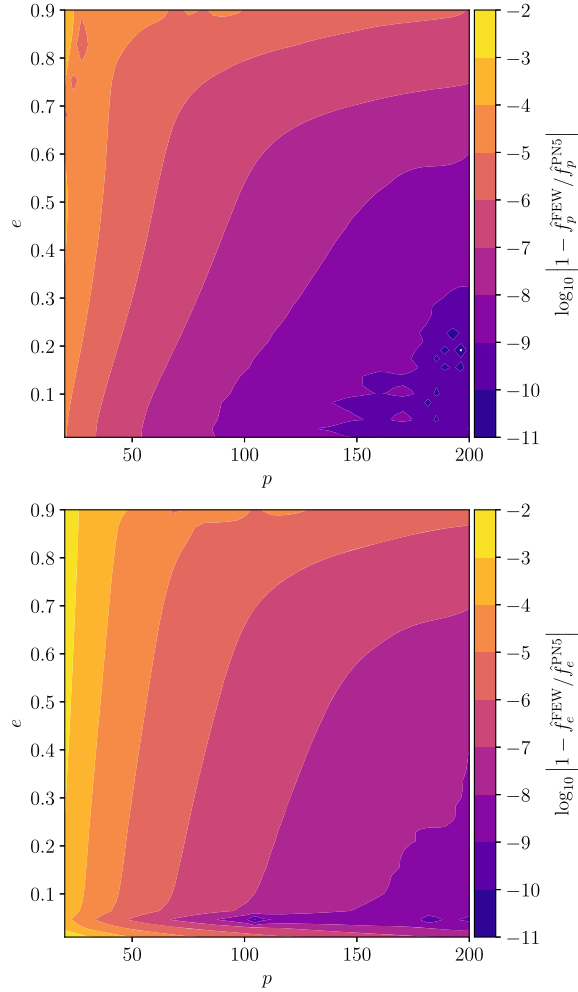


FIG. 25. Relative error in $\hat{f}_{p,e}^{(0)}$ between the interpolated forcing function in this work (FEW, $\hat{f}_{p,e}^{\text{FEW}}$) and the PN forcing function in Appendix E 2 (PN5, $\hat{f}_{p,e}^{\text{PN5}}$) in the weak field for $a = 0.998$ (we find similar behavior for other values of a).

Some of these PN5 results have already been integrated into FEW. For instance, the PN5-AAK waveform model combines the PN5 inspiral trajectory with the semirelativistic (“kludge”) amplitudes [245] as described in Ref. [171], providing a semirelativistic waveform model defined over the generic inspiral parameter space: the FEW implementation of the PN5 mode amplitudes is also under development (cf. Sec. VIB).

We have used the equatorial limit of PN5 results to validate our model (FEWv2) in the weak field by:

- (1) Validating the interpolations of the forcing functions $\hat{f}_{p,e}^{(0)}$ in the trajectory module against their PN5 equivalents (Fig. 25).
- (2) Validating the interpolations of the mode amplitudes $\mathcal{A}_{\ell mn}$ against their PN5 equivalents for sample harmonic modes; the PN5 spherical harmonic waveform mode amplitudes were obtained from the PN5 Teukolsky spheroidal harmonic mode amplitudes

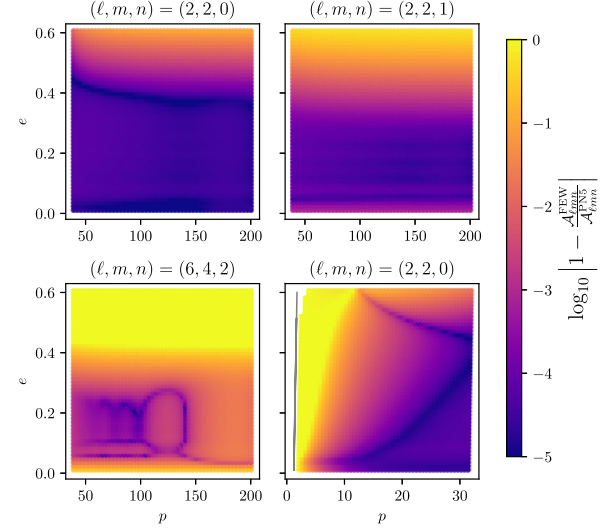


FIG. 26. Absolute relative error for three selected modes of the interpolated (complex) mode amplitude in this work (FEW, $\mathcal{A}_{\ell mn}^{\text{FEW}}$) and the PN5 waveform mode amplitude in Appendix E 2 (PN5, $\mathcal{A}_{\ell mn}^{\text{PN5}}$) with $a = 0.998$ (prograde); note the difference in the range of p in the right bottom panel, which depicts the error in the strong field region. We find similar levels of agreement across other values of a .

via Eq. (6) and projected onto spherical harmonics as described in Eq. (11) (Fig. 26).

- (3) Assessing the orbital dephasing (over four years) between the (fully relativistic) trajectory model in FEW v2 and the PN5 inspiral model scales (Fig. 27).

To do this, we have since implemented two minor improvements to the PN5 trajectory module. Firstly, we have corrected minor typos in the (adiabatic) PN5 fluxes of the angular momentum and Carter constant and the associated PN5 forcing functions $\hat{f}_{p,e}^{\text{PN5}}$. Secondly, unlike the public database at BHPC [231], $\hat{f}_{p,e}^{\text{PN5}}$ here are computed from the PN5 fluxes of energy and angular momentum (in the equatorial limit) *without* expanding the Jacobian elements in Eq. (7) in powers of v or e , as doing so enhances the accuracy of $\hat{f}_{p,e}^{\text{PN5}}$ in comparison to their numerical counterparts (cf. Ref. [122]).

We find in Fig. 25 that the agreement between the interpolations of the forcing functions $\hat{f}_{p,e}^{\text{FEW}}$ and the PN5 forcing functions $\hat{f}_{p,e}^{\text{PN5}}$ improves as p increases and e decreases. This agreement in $\hat{f}_e^{(0)}$ breaks down for small e , since both magnitudes of $\hat{f}_e^{\text{PN5,FEW}}$ become very small. A similar trend can be observed in the amplitude comparisons displayed in the top two panels of Fig. 26, aside from the bottom left panel where the magnitude of the higher multipole amplitude varies rapidly and become small, and the bottom right panel where the accuracy of the PN5 amplitudes decrease closer to the separatrix. However, we find the interpolation error is predominant at smaller

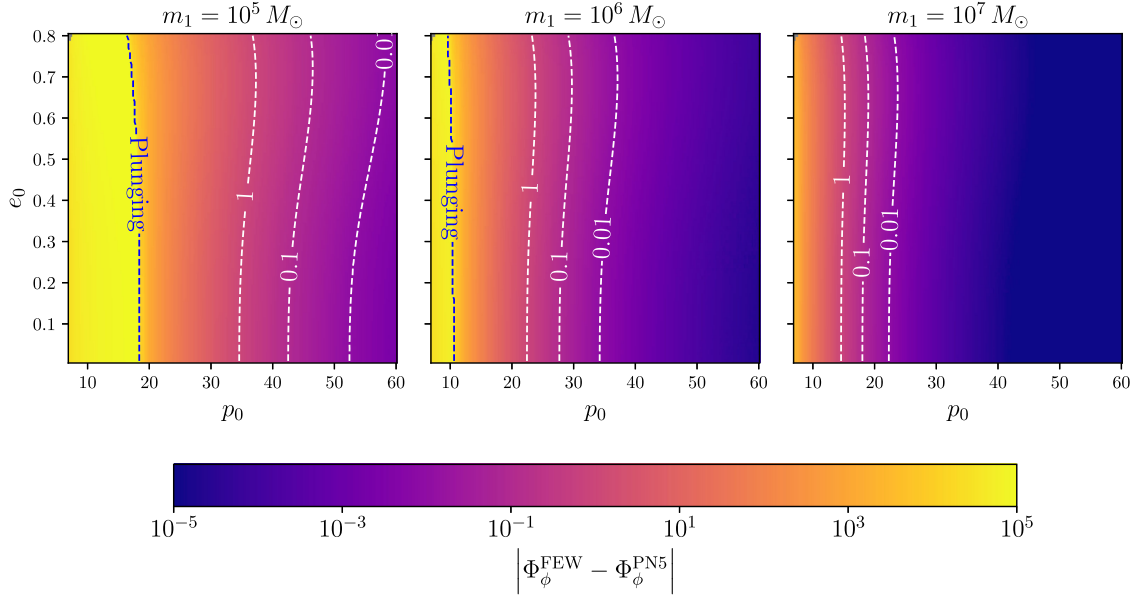


FIG. 27. (From left to right): Orbital dephasings between the trajectory model presented in this work (Kerr) and the PN trajectory discussed in Appendix E 2 (PN5) for $m_1 \in \{10^5, 10^6, 10^7\} M_\odot$. All inspirals are evolved for four years, with $a = 0.998$ and $\epsilon = 10^{-5}$. The white dashed contours indicate regions where the dephasing crosses $\{0.01, 0.1, 1\}$ rad. The blue dashed line indicates the outer boundary in (p_0, e_0) where the body plunges within four years in the PN5 model.

values of e in this case. The deterioration of the agreement in $\mathcal{A}_{\ell mn}$ as the eccentricity becomes larger (typically $e \gtrsim 0.4$) is a result of the eccentricity expansion in the PN5 amplitudes, which become significantly less accurate in that region. This observation has been corroborated through a comparison of the PN5 amplitudes with numerical amplitude data from both the black hole perturbation toolkit (BHPT) [283] and the BHPC [122,231]. For weak-field trajectories with initial separation $p_0 \gg p_{\text{sep}}$ in Fig. 27, we observe that the dephasing between both models tends to zero (independent of the initial eccentricity e_0), indicating that the two models converge in the weak field limit, as expected.

Moving forward, in addition to validating the accuracy of the interpolated amplitudes and forcing functions, these comparisons underpin the integration of PN-GSF results into future FEW frameworks. The current model could be extended much deeper into the weak field and the accuracy of the amplitude and forcing functions may be improved by switching from interpolated data to their PN-GSF counterparts at sufficiently large p (and lower values of e). This motivates further developments in the PN-GSF framework such as extending the analytic expressions to higher PN orders and in turn improving their accuracy at smaller separations. See Refs. [147,148,331,332] for recent efforts in this direction. Furthermore, deriving the PN expressions

without expanding in powers of the eccentricity might be a step-function improvement for the use of PN-GSF information when $e \gtrsim 0.4$ [122,226,333]. We leave these improvements to future work.

APPENDIX F: MARGINAL POSTERIOR DISTRIBUTIONS FOR SCIENCE-CASE SOURCES

In this Appendix, we show corner plots of the marginal posterior distributions obtained from the parameter estimation analyses performed in Sec. V F. Figures 28, 29 and 31 correspond to the analysis of EMRIs with parameters given by rows 1, 2 and 5 of Table I, respectively. We do not observe any significant biases in the intrinsic parameters of these sources for $\kappa = 10^{-2}$, which is expected [based on Eq. (15)] given that these sources have SNRs ≤ 50 .

Similarly, Figs. 30 and 32 correspond to the analysis of IMRIs with parameters given by rows 3 and 4 of Table I, respectively. As the SNRs of these sources is higher (500 and 300, respectively), analyses with $\kappa = 10^{-2}$ failed to converge due to large mismodeling errors shifting the posterior bulk far from the parameters of the injections. For $\kappa = 10^{-3}$, biases in intrinsic parameters are generally below 1σ , but more significant biases are observed in the recovery of extrinsic parameters.

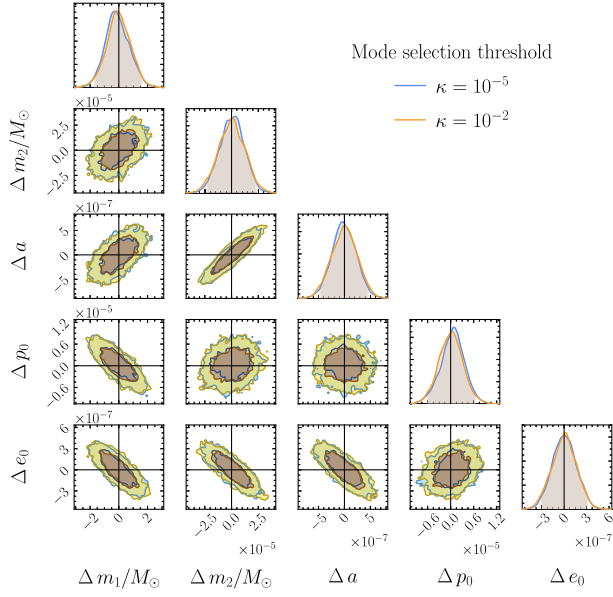


FIG. 28. Posterior distributions of the intrinsic parameters for an EMRI with parameters given in the first row of Table I for two values of the mode selection threshold. The blue (orange) posteriors refer to a threshold set to $\kappa = 10^{-5}$ (10^{-2}). 2D (1D) shaded areas refer to 1- and 2-sigma (95%) credible regions. For the intrinsic parameters, we plot the quantities $\Delta X = X - X_{\text{true}}$ for visual clarity.

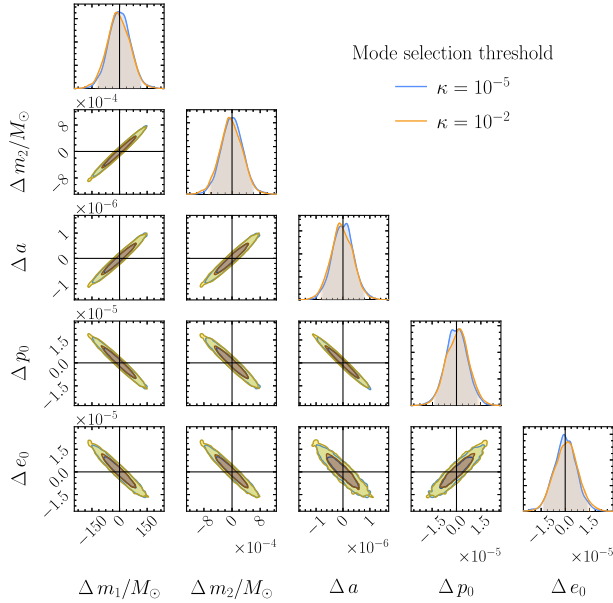


FIG. 29. Posterior distributions of the intrinsic parameters for an EMRI with parameters given in the second row of Table I for two values of the mode selection threshold. The blue (orange) posteriors refer to a threshold set to $\kappa = 10^{-5}$ (10^{-2}). 2D (1D) shaded areas refer to 1- and 2-sigma (95%) credible regions. For the intrinsic parameters, we plot the quantities $\Delta X = X - X_{\text{true}}$ for visual clarity.

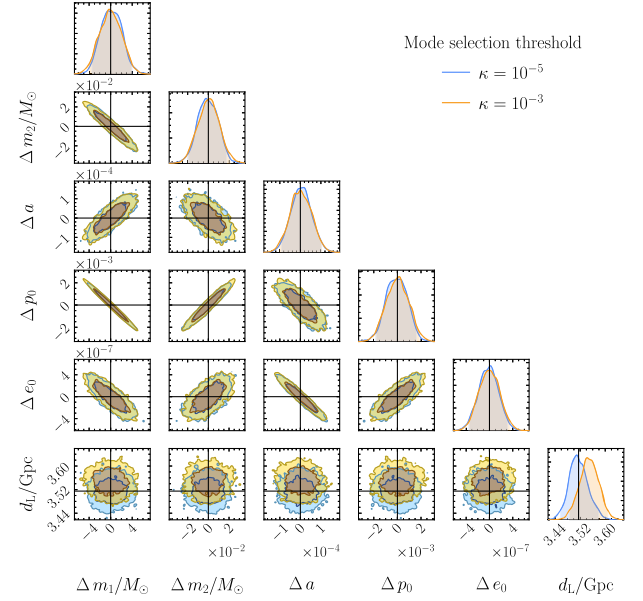


FIG. 30. Posterior distributions of the intrinsic parameters for an IMRI with parameters given in the fourth row of Table I for two values of the mode selection threshold. The blue (orange) posteriors refer to a threshold set to $\kappa = 10^{-5}$ (10^{-3}). 2D (1D) shaded areas refer to 1- and 2-sigma (95%) credible regions. For the intrinsic parameters, we plot the quantities $\Delta X = X - X_{\text{true}}$ for visual clarity. We also include the luminosity distance d_L to show the bias induced by the mode selection threshold.

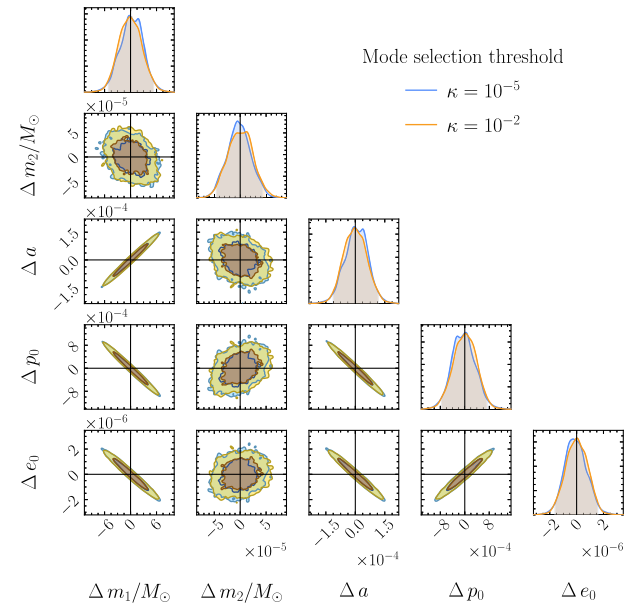


FIG. 31. Posterior distributions of the intrinsic parameters for an EMRI with parameters given in the fifth row of Table I for two values of the mode selection threshold. The blue (orange) posteriors refer to a threshold set to $\kappa = 10^{-5}$ (10^{-2}). 2D (1D) shaded areas refer to 1- and 2-sigma (95%) credible regions. For the intrinsic parameters, we plot the quantities $\Delta X = X - X_{\text{true}}$ for visual clarity.

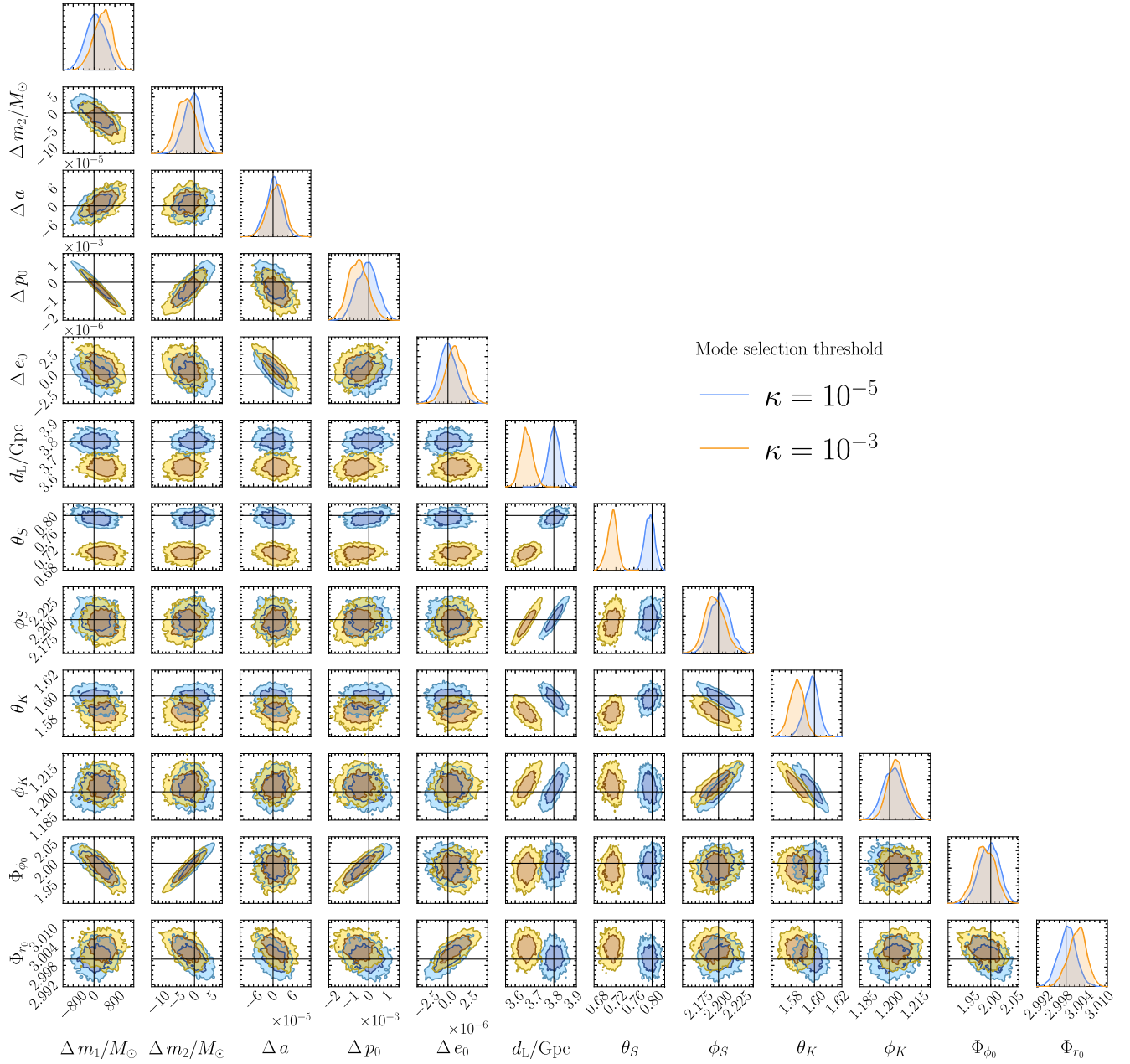


FIG. 32. Posterior distributions for an IMRI with parameters given in the third row of Table I for two values of the mode selection threshold. The blue (orange) posteriors refer to a threshold set to $\kappa = 10^{-5}$ (10^{-3}). 2D (1D) shaded areas refer to 1- and 2-sigma (95%) credible regions. For the intrinsic parameters, we plot the quantities $\Delta X = X - X_{\text{true}}$ for visual clarity. The $\kappa = 10^{-3}$ posteriors yield significant biases in the extrinsic parameters due to the high SNR of this source (500).

- [1] R. Abbott *et al.*, GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run, *Phys. Rev. X* **13**, 041039 (2023).
- [2] J. Aasi *et al.*, Advanced LIGO, *Classical Quantum Gravity* **32**, 074001 (2015).
- [3] F. Acernese *et al.*, Advanced Virgo: A second-generation interferometric gravitational wave detector, *Classical Quantum Gravity* **32**, 024001 (2015).
- [4] T. Akutsu *et al.*, KAGRA: 2.5 generation interferometric gravitational wave detector, *Nat. Astron.* **3**, 35 (2019).
- [5] R. Abbott *et al.*, Population of merging compact binaries inferred using gravitational waves through GWTC-3, *Phys. Rev. X* **13**, 011048 (2023).
- [6] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, Surrogate models for precessing binary black hole simulations with unequal masses, *Phys. Rev. Res.* **1**, 033015 (2019).
- [7] A. Ramos-Buades, A. Buonanno, H. Estellés, M. Khalil, D. P. Mihaylov, S. Ossokine, L. Pompili, and M. Shiferaw, Next generation of accurate and efficient multipolar precessing-spin effective-one-body waveforms for binary black holes, *Phys. Rev. D* **108**, 124037 (2023).
- [8] G. Pratten *et al.*, Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes, *Phys. Rev. D* **103**, 104056 (2021).
- [9] N. Afshordi *et al.*, Waveform modelling for the laser interferometer space antenna, *arXiv:2311.01300*.
- [10] R. Abbott *et al.*, GW190814: Gravitational waves from the coalescence of a 23 solar mass black hole with a 2.6 solar mass compact object, *Astrophys. J. Lett.* **896**, L44 (2020).
- [11] A. G. Abac *et al.*, Observation of gravitational waves from the coalescence of a $2.5\text{--}4.5M_{\odot}$ compact object and a neutron star, *Astrophys. J. Lett.* **970**, L34 (2024).
- [12] R. Abbott *et al.*, GW190412: Observation of a binary-black-hole coalescence with asymmetric masses, *Phys. Rev. D* **102**, 043015 (2020).
- [13] M. Punturo *et al.*, The Einstein telescope: A third-generation gravitational wave observatory, *Classical Quantum Gravity* **27**, 194002 (2010).
- [14] D. Reitze *et al.*, Cosmic explorer: The U.S. contribution to gravitational-wave astronomy beyond LIGO, *Bull. Am. Astron. Soc.* **51**, 035 (2019).
- [15] M. A. Sedda *et al.*, The missing link in gravitational-wave astronomy: Discoveries waiting in the decihertz range, *Classical Quantum Gravity* **37**, 215011 (2020).
- [16] N. Seto, S. Kawamura, and T. Nakamura, Possibility of direct measurement of the acceleration of the universe using 0.1-Hz band laser interferometer gravitational wave antenna in space, *Phys. Rev. Lett.* **87**, 221103 (2001).
- [17] T. Nakamura *et al.*, Pre-DECIGO can get the smoking gun to decide the astrophysical or cosmological origin of GW150914-like binary black holes, *Prog. Theor. Exp. Phys.* **2016**, 093E01 (2016).
- [18] S. Kawamura *et al.*, Current status of space gravitational wave antenna DECIGO and B-DECIGO, *Prog. Theor. Exp. Phys.* **2021**, 05A105 (2021).
- [19] M. Branchesi *et al.*, Lunar gravitational-wave detection, *Space Sci. Rev.* **219**, 67 (2023).
- [20] P. Amaro-Seoane, J. R. Gair, M. Freitag, M. Coleman Miller, I. Mandel, C. J. Cutler, and S. Babak, Astrophysics, detection and science applications of intermediate- and extreme mass-ratio inspirals, *Classical Quantum Gravity* **24**, R113 (2007).
- [21] M. Colpi *et al.*, LISA definition study report, *arXiv:2402.07571*.
- [22] J. Mei *et al.*, The TianQin project: Current progress on science and technology, *Prog. Theor. Exp. Phys.* **2021**, 05A107 (2021).
- [23] Z. Luo, Y. Wang, Y. Wu, W. Hu, and G. Jin, The Taiji program: A concise overview, *Prog. Theor. Exp. Phys.* **2021**, 05A108 (2021).
- [24] Y. Gong, J. Luo, and B. Wang, Concepts and status of Chinese space gravitational wave detection projects, *Nat. Astron.* **5**, 881 (2021).
- [25] P. Amaro-Seoane, Relativistic dynamics and extreme mass ratio inspirals, *Living Rev. Relativity* **21**, 4 (2018).
- [26] Pau Amaro-Seoane, Extremely large mass-ratio inspirals, *Phys. Rev. D* **99**, 123025 (2019).
- [27] E. Gourgoulhon, A. Le Tiec, F. H. Vincent, and N. Warburton, Gravitational waves from bodies orbiting the Galactic Center black hole and their detectability by LISA, *Astron. Astrophys.* **627**, A92 (2019).
- [28] S. Babak, J. Gair, A. Sesana, E. Barausse, C. F. Sopuerta, C. P. L. Berry, E. Berti, P. Amaro-Seoane, A. Petiteau, and A. Klein, Science with the space-based interferometer LISA. V: Extreme mass-ratio inspirals, *Phys. Rev. D* **95**, 103012 (2017).
- [29] C. P. L. Berry, S. A. Hughes, C. F. Sopuerta, A. J. K. Chua, A. Heffernan, K. Holley-Bockelmann, D. P. Mihaylov, M. C. Miller, and A. Sesana, The unique potential of extreme mass-ratio inspirals for gravitational-wave astronomy, *Bull. Am. Astron. Soc.* **51**, 42 (2019).
- [30] L. Barack and C. Cutler, LISA capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy, *Phys. Rev. D* **69**, 082005 (2004).
- [31] J. R. Gair, M. Vallisneri, S. L. Larson, and J. G. Baker, Testing general relativity with low-frequency, space-based gravitational-wave detectors, *Living Rev. Relativity* **16**, 7 (2013).
- [32] A. Maselli, N. Franchini, L. Gualtieri, and T. P. Sotiriou, Detecting scalar fields with extreme mass ratio inspirals, *Phys. Rev. Lett.* **125**, 141101 (2020).
- [33] A. Maselli, N. Franchini, L. Gualtieri, T. P. Sotiriou, S. Barsanti, and P. Pani, Detecting fundamental fields with LISA observations of gravitational waves from extreme mass-ratio inspirals, *Nat. Astron.* **6**, 464 (2022).
- [34] S. Barsanti, N. Franchini, L. Gualtieri, A. Maselli, and T. P. Sotiriou, Extreme mass-ratio inspirals as probes of scalar fields: Eccentric equatorial orbits around Kerr black holes, *Phys. Rev. D* **106**, 044029 (2022).
- [35] L. Speri, S. Barsanti, A. Maselli, T. P. Sotiriou, N. Warburton, M. van de Meent, A. J. K. Chua, O. Burke, and J. Gair, Probing fundamental physics with extreme mass ratio inspirals: A full Bayesian inference for scalar charge, *arXiv:2406.07607*.
- [36] B. Bonga, H. Yang, and S. A. Hughes, Tidal resonance in extreme mass-ratio inspirals, *Phys. Rev. Lett.* **123**, 101103 (2019).

- [37] H. Yang, B. Bonga, Z. Peng, and G. Li, Relativistic mean motion resonance, *Phys. Rev. D* **100**, 124056 (2019).
- [38] Y. Yin, J. Mathews, A. J. K. Chua, and X. Chen, Relativistic model of binary extreme-mass-ratio inspiral systems and their gravitational radiation, *Phys. Rev. D* **111**, 103007 (2025).
- [39] C. Dyson, T. F. M. Spieksma, R. Brito, M. van de Meent, and S. Dolan, Environmental effects in extreme-mass-ratio inspirals: Perturbations to the environment in Kerr spacetimes, *Phys. Rev. Lett.* **134**, 211403 (2025).
- [40] L. Speri, A. Antonelli, L. Sberna, S. Babak, E. Barausse, J. R. Gair, and M. L. Katz, Probing accretion physics with gravitational waves, *Phys. Rev. X* **13**, 021035 (2023).
- [41] J. R. Gair, C. Tang, and M. Volonteri, LISA extreme-mass-ratio inspiral events as probes of the black hole mass function, *Phys. Rev. D* **81**, 104014 (2010).
- [42] C. E. A. Chapman-Bird, C. P. L. Berry, and G. Woan, Rapid determination of LISA sensitivity to extreme mass ratio inspirals with machine learning, *Mon. Not. R. Astron. Soc.* **522**, 6043 (2023).
- [43] E. Barausse, L. Rezzolla, D. Petroff, and M. Ansorg, Gravitational waves from extreme mass ratio inspirals in non-pure Kerr spacetimes, *Phys. Rev. D* **75**, 064026 (2007).
- [44] E. Barausse and L. Rezzolla, The influence of the hydrodynamic drag from an accretion torus on extreme mass-ratio inspirals, *Phys. Rev. D* **77**, 104027 (2008).
- [45] N. Yunes, B. Kocsis, A. Loeb, and Z. Haiman, Imprint of accretion disk-induced migration on gravitational waves from extreme mass ratio inspirals, *Phys. Rev. Lett.* **107**, 171103 (2011).
- [46] E. Barausse, V. Cardoso, and P. Pani, Can environmental effects spoil precision gravitational-wave astrophysics?, *Phys. Rev. D* **89**, 104059 (2014).
- [47] E. Barausse, V. Cardoso, and P. Pani, Environmental effects for gravitational-wave astrophysics, *J. Phys. Conf. Ser.* **610**, 012044 (2015).
- [48] H. Khalvati, A. Santini, F. Duque, L. Speri, J. Gair, H. Yang, and R. Brito, Impact of relativistic waveforms in LISA's science objectives with extreme-mass-ratio inspirals, *Phys. Rev. D* **111**, 082010 (2025).
- [49] F. Duque, S. Kejriwal, L. Sberna, L. Speri, and J. Gair, Constraining accretion physics with gravitational waves from eccentric extreme-mass-ratio inspirals, *Phys. Rev. D* **111**, 084006 (2025).
- [50] L. Copparoni, E. Barausse, L. Speri, L. Sberna, and A. Derdzinski, Implications of stochastic gas torques for asymmetric binaries in the LISA band, *Phys. Rev. D* **111**, 104079 (2025).
- [51] Z. Pan, Z. Lyu, and H. Yang, Wet extreme mass ratio inspirals may be more common for spaceborne gravitational wave detection, *Phys. Rev. D* **104**, 063007 (2021).
- [52] Z. Pan and H. Yang, Formation rate of extreme mass ratio inspirals in active galactic nuclei, *Phys. Rev. D* **103**, 103018 (2021).
- [53] P. A. Seoane *et al.*, Astrophysics with the laser interferometer space antenna, *Living Rev. Relativity* **26**, 2 (2023).
- [54] S. Naoz, S. C. Rose, E. Michaely, D. Melchor, E. Ramirez-Ruiz, B. Mockler, and J. D. Schnittman, The combined effects of two-body relaxation processes and the eccentric Kozai–Lidov mechanism on the extreme-mass-ratio inspirals rate, *Astrophys. J. Lett.* **927**, L18 (2022).
- [55] C. L. MacLeod and C. J. Hogan, Precision of Hubble constant derived using black hole binary absolute distances and statistical redshift information, *Phys. Rev. D* **77**, 043512 (2008).
- [56] D. Laghi, N. Tamanini, W. Del Pozzo, A. Sesana, J. Gair, S. Babak, and D. Izquierdo-Villalba, Gravitational-wave cosmology with extreme mass-ratio inspirals, *Mon. Not. R. Astron. Soc.* **508**, 4512 (2021).
- [57] C. Liu, D. Laghi, and N. Tamanini, Probing modified gravitational-wave propagation with extreme mass-ratio inspirals, *Phys. Rev. D* **109**, 063521 (2024).
- [58] P. Auclair *et al.*, Cosmology with the laser interferometer space antenna, *Living Rev. Relativity* **26**, 5 (2023).
- [59] M. Toscani, O. Burke, C. Liu, N. B. Zamel, N. Tamanini, and F. Pozzoli, Strongly lensed extreme mass-ratio inspirals, *Phys. Rev. D* **109**, 063505 (2024).
- [60] E. Savalle, J. Gair, L. Speri, and S. Babak, Assessing the impact of instrumental calibration uncertainty on LISA science, *Phys. Rev. D* **106**, 022003 (2022).
- [61] A. Pound and B. Wardell, Black hole perturbation theory and gravitational self-force, in *Handbook of Gravitational Wave Astronomy* (Springer, Singapore, 2022), pp. 1411–1530.
- [62] L. Barack and A. Pound, Self-force and radiation reaction in general relativity, *Rep. Prog. Phys.* **82**, 016904 (2019).
- [63] L. Barack, Gravitational self force in extreme mass-ratio inspirals, *Classical Quantum Gravity* **26**, 213001 (2009).
- [64] T. Alexander, EMRIs and the relativistic loss-cone: The curious case of the fortunate coincidence, *J. Phys. Conf. Ser.* **840**, 012019 (2017).
- [65] W. Schmidt, Celestial mechanics in Kerr space-time, *Classical Quantum Gravity* **19**, 2743 (2002).
- [66] S. A. Hughes, N. Warburton, G. Khanna, A. J. K. Chua, and M. L. Katz, Adiabatic waveforms for extreme mass-ratio inspirals via multivoice decomposition in time and frequency, *Phys. Rev. D* **103**, 104014 (2021); *Phys. Rev. D* **107**, 089901(E) (2023).
- [67] S. Drasco and S. A. Hughes, Gravitational wave snapshots of generic extreme mass ratio inspirals, *Phys. Rev. D* **73**, 024027 (2006); *Phys. Rev. D* **88**, 109905(E) (2013); **90**, 109905(E) (2014).
- [68] C. O. Lousto and J. Healy, Exploring the small mass ratio binary black hole merger via Zeno's dichotomy approach, *Phys. Rev. Lett.* **125**, 191102 (2020).
- [69] N. Rosato, J. Healy, and C. O. Lousto, Adapted gauge to small mass ratio binary black hole evolutions, *Phys. Rev. D* **103**, 104068 (2021).
- [70] C. O. Lousto and J. Healy, Study of the intermediate mass ratio black hole binary merger up to 1000:1 with numerical relativity, *Classical Quantum Gravity* **40**, 09LT01 (2023).
- [71] N. A. Wittek, A. Pound, H. P. Pfeiffer, and L. Barack, Worldtube excision method for intermediate-mass-ratio inspirals: Self-consistent evolution in a scalar-charge model, *Phys. Rev. D* **110**, 084023 (2024).
- [72] N. A. Wittek, L. Barack, H. P. Pfeiffer, A. Pound, N. Deppe, L. E. Kidder, A. Macedo, K. C. Nelli, W. Throwe, and N. L. Vu, Relieving scale disparity in binary black hole simulations, [arXiv:2410.22290](https://arxiv.org/abs/2410.22290).

- [73] M. L. Katz, N. Karnesis, N. Korsakova, J. R. Gair, and N. Stergioulas, Efficient GPU-accelerated multisource global fit pipeline for LISA data analysis, *Phys. Rev. D* **111**, 024060 (2025).
- [74] S. Deng, S. Babak, M. Le Jeune, S. Marsat, E. Plagnol, and A. Sartirana, Modular global-fit pipeline for LISA data analysis, *Phys. Rev. D* **111**, 103014 (2025).
- [75] T. B. Littenberg and N. J. Cornish, Prototype global analysis of LISA data with multiple source types, *Phys. Rev. D* **107**, 063004 (2023).
- [76] Y. Mino, M. Sasaki, and T. Tanaka, Gravitational radiation reaction to a particle motion, *Phys. Rev. D* **55**, 3457 (1997).
- [77] T. C. Quinn and R. M. Wald, An axiomatic approach to electromagnetic and gravitational radiation reaction of particles in curved space-time, *Phys. Rev. D* **56**, 3381 (1997).
- [78] S. E. Gralla and R. M. Wald, A rigorous derivation of gravitational self-force, *Classical Quantum Gravity* **25**, 205009 (2008).
- [79] S. Detweiler, Gravitational radiation reaction and second order perturbation theory, *Phys. Rev. D* **85**, 044048 (2012).
- [80] S. E. Gralla, Second order gravitational self force, *Phys. Rev. D* **85**, 124011 (2012).
- [81] A. Pound, Second-order gravitational self-force, *Phys. Rev. Lett.* **109**, 051101 (2012).
- [82] A. I. Harte, Motion in classical field theories and the foundations of the self-force problem, *Fund. Theor. Phys.* **179**, 327 (2015).
- [83] E. Poisson, A. Pound, and I. Vega, The motion of point particles in curved spacetime, *Living Rev. Relativity* **14**, 7 (2011).
- [84] T. Hinderer and E. E. Flanagan, Two timescale analysis of extreme mass ratio inspirals in Kerr. I. Orbital motion, *Phys. Rev. D* **78**, 064028 (2008).
- [85] J. Miller and A. Pound, Two-timescale evolution of extreme-mass-ratio inspirals: Waveform generation scheme for quasicircular orbits in Schwarzschild space-time, *Phys. Rev. D* **103**, 064048 (2021).
- [86] J. Mathews and A. Pound, Post-adiabatic waveform-generation framework for asymmetric precessing binaries, *arXiv:2501.01413*.
- [87] N. Yunes, A. Buonanno, S. A. Hughes, Y. Pan, E. Barausse, M. C. Miller, and W. Throwe, Extreme mass-ratio inspirals in the effective-one-body approach: Quasicircular, equatorial orbits around a spinning black hole, *Phys. Rev. D* **83**, 044044 (2011); *Phys. Rev. D* **88**, 109904 (E) (2013).
- [88] D. Chiamerello and A. Nagar, Faithful analytical effective-one-body waveform model for spin-aligned, moderately eccentric, coalescing black hole binaries, *Phys. Rev. D* **101**, 101501 (2020).
- [89] C. Zhang, W.-B. Han, and S.-C. Yang, Analytical effective one-body formalism for extreme-mass-ratio inspirals with eccentric orbits, *Commun. Theor. Phys.* **73**, 085401 (2021).
- [90] A. Ramos-Buades, A. Buonanno, M. Khalil, and S. Ossokine, Effective-one-body multipolar waveforms for eccentric binary black holes with nonprecessing spins, *Phys. Rev. D* **105**, 044035 (2022).
- [91] C. Zhang, W.-B. Han, X.-Y. Zhong, and G. Wang, Geometrized effective-one-body formalism for extreme-mass-ratio limits: Generic orbits, *Phys. Rev. D* **104**, 024050 (2021).
- [92] S. Albanesi, A. Nagar, and S. Bernuzzi, Effective one-body model for extreme-mass-ratio spinning binaries on eccentric equatorial orbits: Testing radiation reaction and waveform, *Phys. Rev. D* **104**, 024067 (2021).
- [93] A. Nagar and S. Albanesi, Toward a gravitational self-force-informed effective-one-body waveform model for nonprecessing, eccentric, large-mass-ratio inspirals, *Phys. Rev. D* **106**, 064049 (2022).
- [94] A. Albertini, A. Nagar, A. Pound, N. Warburton, B. Wardell, L. Durkan, and J. Miller, Comparing second-order gravitational self-force and effective one body waveforms from inspiralling, quasicircular and nonspinning black hole binaries. II. The large-mass-ratio case, *Phys. Rev. D* **106**, 084062 (2022).
- [95] A. Albertini, R. Gamba, A. Nagar, and S. Bernuzzi, Effective-one-body waveforms for extreme-mass-ratio binaries: Consistency with second-order gravitational self-force quasicircular results and extension to nonprecessing spins and eccentricity, *Phys. Rev. D* **109**, 044022 (2024).
- [96] S. Albanesi, S. Bernuzzi, T. Damour, A. Nagar, and A. Placidi, Faithful effective-one-body waveform of small-mass-ratio coalescing black hole binaries: The eccentric, nonspinning case, *Phys. Rev. D* **108**, 084037 (2023).
- [97] C. Cheung, J. Parra-Martinez, I. Z. Rothstein, N. Shah, and J. Wilson-Gerow, Effective field theory for extreme mass ratio binaries, *Phys. Rev. Lett.* **132**, 091402 (2024).
- [98] A. Gamboa *et al.*, Accurate waveforms for eccentric, aligned-spin binary black holes: The multipolar effective-one-body model SEOBNRv5EHM, *arXiv:2412.12823*.
- [99] N. Loutrel, S. Mukherjee, A. Maselli, and P. Pani, Analytical model of precessing binaries using post-Newtonian theory in the extreme mass-ratio limit: General formalism, *Phys. Rev. D* **110**, 024006 (2024).
- [100] K. Paul, A. Maurya, Q. Henry, K. Sharma, P. Satheesh, Divyajyoti, P. Kumar, and C. K. Mishra, Eccentric, spinning, inspiral-merger-ringdown waveform model with higher modes for the detection and characterization of binary black holes, *Phys. Rev. D* **111**, 084074 (2025).
- [101] C. Cheung, J. Parra-Martinez, I. Z. Rothstein, N. Shah, and J. Wilson-Gerow, Gravitational scattering and beyond from extreme mass ratio effective field theory, *J. High Energy Phys.* **10** (2024) 005.
- [102] D. Akpinar, V. del Duca, and R. Gonzo, The spinning self-force EFT: 1SF waveform recursion relation and Compton scattering, *arXiv:2504.02025*.
- [103] G. Morras, G. Pratten, and P. Schmidt, Improved post-Newtonian waveform model for inspiralling precessing-eccentric compact binaries, *Phys. Rev. D* **111**, 084052 (2025).
- [104] M. d. L. Planas, A. Ramos-Buades, C. García-Quirós, H. Estellés, S. Husa, and M. Haney, Time-domain phenomenological multipolar waveforms for aligned-spin binary black holes in elliptical orbits, *arXiv:2503.13062*.
- [105] I. Strusberg, B. Rom, and R. Sari, Universal Waveforms for Extreme Mass-Ratio Inspiral, *arXiv:2505.07941*.

- [106] A. Apte and S. A. Hughes, Exciting black hole modes via misaligned coalescences: I. Inspiral, transition, and plunge trajectories using a generalized Ori-Thorne procedure, *Phys. Rev. D* **100**, 084031 (2019).
- [107] L. K  chler, G. Comp  re, L. Durkan, and A. Pound, Self-force framework for transition-to-plunge waveforms, *SciPost Phys.* **17**, 056 (2024).
- [108] D. R. Becker and S. A. Hughes, Transition from adiabatic inspiral to plunge for eccentric binaries, *Phys. Rev. D* **111**, 064003 (2025).
- [109] G. Lhost and G. Comp  re, Approach to the separatrix with eccentric orbits, *arXiv:2412.04249*.
- [110] L. K  chler, G. Comp  re, and A. Pound, Self-force framework for merger-ringdown waveforms, *arXiv:2506.02189*.
- [111] B. Carter, Global structure of the Kerr family of gravitational fields, *Phys. Rev.* **174**, 1559 (1968).
- [112] Y. Mino, Perturbative approach to an orbital evolution around a supermassive black hole, *Phys. Rev. D* **67**, 084027 (2003).
- [113] S. Drasco, E. E. Flanagan, and S. A. Hughes, Computing inspirals in Kerr in the adiabatic regime. I. The scalar case, *Classical Quantum Gravity* **22**, S801 (2005).
- [114] N. Sago, T. Tanaka, W. Hikida, K. Ganz, and H. Nakano, The adiabatic evolution of orbital parameters in the Kerr spacetime, *Prog. Theor. Phys.* **115**, 873 (2006).
- [115] T. Tanaka, Gravitational radiation reaction, *Prog. Theor. Phys. Suppl.* **163**, 120 (2006).
- [116] S. Isoyama, R. Fujita, H. Nakano, N. Sago, and T. Tanaka, ‘‘Flux-balance formulae’’ for extreme mass-ratio inspirals, *Prog. Theor. Exp. Phys.* **2019**, 013E01 (2019).
- [117] A. M. Grant, Flux-balance laws for spinning bodies under the gravitational self-force, *Phys. Rev. D* **111**, 084015 (2025).
- [118] S. A. Hughes, Evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational wave emission. II. Inspiral trajectories and gravitational wave forms, *Phys. Rev. D* **64**, 064004 (2001); *Phys. Rev. D* **88**, 109902(E) (2013).
- [119] L. S. Finn and K. S. Thorne, Gravitational waves from a compact star in a circular, inspiral orbit, in the equatorial plane of a massive, spinning black hole, as observed by LISA, *Phys. Rev. D* **62**, 124021 (2000).
- [120] Y. Mino, Adiabatic expansion for metric perturbation and the condition to solve the gauge problem for gravitational radiation reaction problem, *Prog. Theor. Phys.* **115**, 43 (2006).
- [121] P. A. Sundararajan, G. Khanna, S. A. Hughes, and S. Drasco, Towards adiabatic waveforms for inspiral into Kerr black holes: II. Dynamical sources and generic orbits, *Phys. Rev. D* **78**, 024022 (2008).
- [122] S. Isoyama, R. Fujita, A. J. K. Chua, H. Nakano, A. Pound, and N. Sago, Adiabatic waveforms from extreme-mass-ratio inspirals: An analytical approach, *Phys. Rev. Lett.* **128**, 231101 (2022).
- [123] M. Kerachian, L. Polcar, V. Skoup  y, C. Efthymiopoulos, and G. Lukes-Gerakopoulos, Action-angle formalism for extreme mass ratio inspirals in Kerr spacetime, *Phys. Rev. D* **108**, 044004 (2023).
- [124] Z. Nasipak, Adiabatic gravitational waveform model for compact objects undergoing quasicircular inspirals into rotating massive black holes, *Phys. Rev. D* **109**, 044020 (2024).
- [125] T. Nakamura, K. Oohara, and Y. Kojima, General relativistic collapse to black holes and gravitational waves from black holes, *Prog. Theor. Phys. Suppl.* **90**, 1 (1987).
- [126] M. Shibata, Gravitational waves induced by a particle orbiting around a rotating black hole: Spin orbit interaction effect, *Phys. Rev. D* **48**, 663 (1993).
- [127] M. Shibata, Gravitational waves induced by a particle orbiting around a rotating black hole: Effect of orbital precession, *Prog. Theor. Phys.* **90**, 595 (1993).
- [128] M. Shibata, Gravitational waves by compact stars orbiting around rotating supermassive black holes, *Phys. Rev. D* **50**, 6297 (1994).
- [129] D. Kennefick, Stability under radiation reaction of circular equatorial orbits around Kerr black holes, *Phys. Rev. D* **58**, 064012 (1998).
- [130] S. A. Hughes, The evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational wave emission, *Phys. Rev. D* **61**, 084004 (2000); *Phys. Rev. D* **63**, 049902(E) (2001); **65**, 069902(E) (2002); **67**, 089901(E) (2003); **78**, 109902(E) (2008); **90**, 109904(E) (2014).
- [131] K. Glampedakis and D. Kennefick, Zoom and whirl: Eccentric equatorial orbits around spinning black holes and their evolution under gravitational radiation reaction, *Phys. Rev. D* **66**, 044002 (2002).
- [132] S. A. Hughes, S. Drasco, E. E. Flanagan, and J. Franklin, Gravitational radiation reaction and inspiral waveforms in the adiabatic limit, *Phys. Rev. Lett.* **94**, 221101 (2005).
- [133] Y. Mino, Self-force in the radiation reaction formula, *Prog. Theor. Phys.* **113**, 733 (2005).
- [134] R. Fujita, W. Hikida, and H. Tagoshi, An efficient numerical method for computing gravitational waves induced by a particle moving on eccentric inclined orbits around a Kerr black hole, *Prog. Theor. Phys.* **121**, 843 (2009).
- [135] E. Harms, S. Bernuzzi, A. Nagar, and A. Zenginoglu, A new gravitational wave generation algorithm for particle perturbations of the Kerr spacetime, *Classical Quantum Gravity* **31**, 245004 (2014).
- [136] R. Fujita and M. Shibata, Extreme mass ratio inspirals on the equatorial plane in the adiabatic order, *Phys. Rev. D* **102**, 064005 (2020).
- [137] C. Chen and J. Jing, Gravitational wave fluxes on generic orbits in near-extreme Kerr spacetime: Higher spin and large eccentricity, *Sci. China Phys. Mech. Astron.* **67**, 110411 (2024).
- [138] Y. Mino, M. Sasaki, M. Shibata, H. Tagoshi, and T. Tanaka, Black hole perturbation: Chapter 1, *Prog. Theor. Phys. Suppl.* **128**, 1 (1997).
- [139] M. Sasaki and H. Tagoshi, Analytic black hole perturbation approach to gravitational radiation, *Living Rev. Relativity* **6**, 6 (2003).
- [140] K. Ganz, W. Hikida, H. Nakano, N. Sago, and T. Tanaka, Adiabatic evolution of three ‘constants’ of motion for greatly inclined orbits in Kerr spacetime, *Prog. Theor. Phys.* **117**, 1041 (2007).

- [141] R. Fujita, Gravitational waves from a particle in circular orbits around a Schwarzschild black hole to the 22nd post-Newtonian order, *Prog. Theor. Phys.* **128**, 971 (2012).
- [142] A. G. Shah, Gravitational-wave flux for a particle orbiting a Kerr black hole to 20th post-Newtonian order: A numerical approach, *Phys. Rev. D* **90**, 044025 (2014).
- [143] R. Fujita, Gravitational waves from a particle in circular orbits around a rotating black hole to the 11th post-Newtonian order, *Prog. Theor. Exp. Phys.* **2015**, 33E01 (2015).
- [144] C. Munna, Analytic post-Newtonian expansion of the energy and angular momentum radiated to infinity by eccentric-orbit nonspinning extreme-mass-ratio inspirals to the 19th order, *Phys. Rev. D* **102**, 124001 (2020).
- [145] C. Munna and C. R. Evans, Eccentric-orbit extreme-mass-ratio-inspiral radiation II: 1PN correction to leading-logarithm and subleading-logarithm flux sequences and the entire perturbative 4PN flux, *Phys. Rev. D* **102**, 104006 (2020).
- [146] C. Munna, C. R. Evans, and E. Forseth, Tidal heating and torquing of the primary black hole in eccentric-orbit, nonspinning, extreme-mass-ratio inspirals to 22PN order, *Phys. Rev. D* **108**, 044039 (2023).
- [147] N. Sago, R. Fujita, and H. Nakano, Post-Newtonian templates for phase evolution of spherical extreme mass ratio inspirals, *Phys. Rev. D* **111**, 064043 (2025).
- [148] J. C. Castillo, C. R. Evans, C. Kavanagh, J. Neef, A. Ottewill, and B. Wardell, Post-Newtonian expansion of gravitational energy and angular momentum fluxes: Inclined spherical orbits about a Kerr black hole, *Phys. Rev. D* **111**, 084004 (2025).
- [149] O. Burke, G. A. Piovano, N. Warburton, P. Lynch, L. Speri, C. Kavanagh, B. Wardell, A. Pound, L. Durkan, and J. Miller, Assessing the importance of first postadiabatic terms for small-mass-ratio binaries, *Phys. Rev. D* **109**, 124048 (2024).
- [150] L. Barack and N. Sago, Gravitational self force on a particle in circular orbit around a Schwarzschild black hole, *Phys. Rev. D* **75**, 064021 (2007).
- [151] L. Barack and N. Sago, Gravitational self-force on a particle in eccentric orbit around a Schwarzschild black hole, *Phys. Rev. D* **81**, 084021 (2010).
- [152] T. Osburn, E. Forseth, C. R. Evans, and S. Hopper, Lorenz gauge gravitational self-force calculations of eccentric binaries using a frequency domain procedure, *Phys. Rev. D* **90**, 104031 (2014).
- [153] S. Isoyama, L. Barack, S. R. Dolan, A. Le Tiec, H. Nakano, A. G. Shah, T. Tanaka, and N. Warburton, Gravitational self-force correction to the innermost stable circular equatorial orbit of a Kerr black hole, *Phys. Rev. Lett.* **113**, 161101 (2014).
- [154] B. Wardell, Self-force: Computational strategies, *Fund. Theor. Phys.* **179**, 487 (2015).
- [155] M. van de Meent, Gravitational self-force on eccentric equatorial orbits around a Kerr black hole, *Phys. Rev. D* **94**, 044034 (2016).
- [156] M. van de Meent, Gravitational self-force on generic bound geodesics in Kerr spacetime, *Phys. Rev. D* **97**, 104033 (2018).
- [157] S. R. Dolan, L. Durkan, C. Kavanagh, and B. Wardell, Metric perturbations of Kerr spacetime in Lorenz gauge: Circular equatorial orbits, *Classical Quantum Gravity* **41**, 155011 (2024).
- [158] A. Pound, B. Wardell, N. Warburton, and J. Miller, Second-order self-force calculation of gravitational binding energy in compact binaries, *Phys. Rev. Lett.* **124**, 021101 (2020).
- [159] N. Warburton, A. Pound, B. Wardell, J. Miller, and L. Durkan, Gravitational-wave energy flux for compact binaries through second order in the mass ratio, *Phys. Rev. Lett.* **127**, 151102 (2021).
- [160] L. Durkan and N. Warburton, Slow evolution of the metric perturbation due to a quasicircular inspiral into a Schwarzschild black hole, *Phys. Rev. D* **106**, 084023 (2022).
- [161] L. V. Drummond, P. Lynch, A. G. Hanselman, D. R. Becker, and S. A. Hughes, Extreme mass-ratio inspiral and waveforms for a spinning body into a Kerr black hole via osculating geodesics and near-identity transformations, *Phys. Rev. D* **109**, 064030 (2024).
- [162] V. Skoupy, G. Lukes-Gerakopoulos, L. V. Drummond, and S. A. Hughes, Asymptotic gravitational-wave fluxes from a spinning test body on generic orbits around a Kerr black hole, *Phys. Rev. D* **108**, 044041 (2023).
- [163] G. A. Piovano, C. Pantelidou, J. Mac Uilliam, and V. Witzany, Spinning particles near Kerr black holes: Orbits and gravitational-wave fluxes through the Hamilton-Jacobi formalism, *Phys. Rev. D* **111**, 044009 (2025).
- [164] B. Wardell, A. Pound, N. Warburton, J. Miller, L. Durkan, and A. Le Tiec, Gravitational waveforms for compact binaries from second-order self-force theory, *Phys. Rev. Lett.* **130**, 241402 (2023).
- [165] M. Van De Meent and N. Warburton, Fast self-forced inspirals, *Classical Quantum Gravity* **35**, 144003 (2018).
- [166] P. Lynch, M. van de Meent, and N. Warburton, Eccentric self-forced inspirals into a rotating black hole, *Classical Quantum Gravity* **39**, 145004 (2022).
- [167] P. Lynch, Efficient trajectory calculations for extreme mass-ratio inspirals using near-identity (averaging) transformations, Ph.D. dissertation, University College Dublin, 2022, available electronically at <http://hdl.handle.net/10197/13347>.
- [168] P. Lynch, M. van de Meent, and N. Warburton, Self-forced inspirals with spin-orbit precession, *Phys. Rev. D* **109**, 084072 (2024).
- [169] P. Lynch, V. Witzany, M. van de Meent, and N. Warburton, Fast inspirals and the treatment of orbital resonances, *Classical Quantum Gravity* **41**, 225002 (2024).
- [170] A. J. K. Chua, M. L. Katz, N. Warburton, and S. A. Hughes, Rapid generation of fully relativistic extreme-mass-ratio-inspiral waveform templates for LISA data analysis, *Phys. Rev. Lett.* **126**, 051102 (2021).
- [171] M. L. Katz, A. J. K. Chua, L. Speri, N. Warburton, and S. A. Hughes, Fast extreme-mass-ratio-inspiral waveforms: New tools for millihertz gravitational-wave data analysis, *Phys. Rev. D* **104**, 064047 (2021).
- [172] C. Chapman-Bird, A. J. K. Chua, S. Hughes, M. Katz, Z. Nasipak, M. Pigou, L. Speri, and N. Warburton, Fastemriwaveforms, [10.5281/zenodo.15630565](https://doi.org/10.5281/zenodo.15630565) (2025).

- [173] L. Speri, M. L. Katz, A. J. K. Chua, S. A. Hughes, N. Warburton, J. E. Thompson, C. E. A. Chapman-Bird, and J. R. Gair, Fast and Fourier: Extreme mass ratio inspiral waveforms in the frequency domain, *Front. Appl. Math. Stat.* **9** (2024).
- [174] C. S. Reynolds, Observational constraints on black hole spin, *Annu. Rev. Astron. Astrophys.* **59**, 117 (2021).
- [175] O. Burke, J. R. Gair, J. Simón, and M. C. Edwards, Constraining the spin parameter of near-extremal black holes using LISA, *Phys. Rev. D* **102**, 124054 (2020).
- [176] K. S. Thorne, Disk accretion onto a black hole. 2. Evolution of the hole, *Astrophys. J.* **191**, 507 (1974).
- [177] J. M. Bardeen, Kerr metric black holes, *Nature (London)* **226**, 64 (1970).
- [178] A. Sadowski, M. Bursa, M. Abramowicz, W. Kluzniak, J.-P. Lasota, R. Moderski, and M. Safarzadeh, Spinning up black holes with super-critical accretion flows, *Astron. Astrophys.* **532**, A41 (2011).
- [179] A. Arbey, J. Auffinger, and J. Silk, Evolution of primordial black hole spin due to Hawking radiation, *Mon. Not. R. Astron. Soc.* **494**, 1257 (2020).
- [180] S. E. Gralla, S. A. Hughes, and N. Warburton, Inspiral into Gargantua, *Classical Quantum Gravity* **33**, 155002 (2016); *Classical Quantum Gravity* **37**, 109501(E) (2020).
- [181] G. Compère, K. Fransen, T. Hertog, and J. Long, Gravitational waves from plunges into Gargantua, *Classical Quantum Gravity* **35**, 104002 (2018).
- [182] O. Burke, J. R. Gair, and J. Simón, Transition from inspiral to plunge: A complete near-extremal trajectory and associated waveform, *Phys. Rev. D* **101**, 064026 (2020).
- [183] P. S. Cole, J. Alvey, L. Speri, C. Weniger, U. Bhardwaj, D. Gerosa, and G. Bertone, Sequential simulation-based inference for extreme mass ratio inspirals, [arXiv:2505.16795](#).
- [184] S. H. Strub, L. Speri, and D. Giardini, Searching for extreme mass ratio inspirals in LISA: From identification to parameter estimation, [arXiv:2505.17814](#).
- [185] C.-Q. Ye, H.-M. Fan, A. Torres-Orjuela, J.-d. Zhang, and Y.-M. Hu, Identification of gravitational waves from extreme-mass-ratio inspirals, *Phys. Rev. D* **109**, 124034 (2024).
- [186] A. J. K. Chua, One-stop function for gravitational-wave detection, identification, and inference, *Phys. Rev. D* **106**, 104051 (2022).
- [187] C. Badger, J. A. Font, M. Sakellariadou, and A. Torres-Forné, High-speed reconstruction of long-duration gravitational waves from extreme-mass-ratio inspirals using sparse dictionary learning, *Phys. Rev. D* **110**, 064074 (2024).
- [188] X.-T. Zhang, C. Messenger, N. Korsakova, M. L. Chan, Y.-M. Hu, and J.-d. Zhang, Detecting gravitational waves from extreme mass ratio inspirals using convolutional neural networks, *Phys. Rev. D* **105**, 123027 (2022).
- [189] Q. Yun, W.-B. Han, Y.-Y. Guo, H. Wang, and M. Du, The detection, extraction and parameter estimation of extreme-mass-ratio inspirals with deep learning, *Sci. China Phys. Mech. Astron.* **68**, 210413 (2025).
- [190] Q. Baghi, The LISA data challenges, in *56th Rencontres de Moriond on Gravitation* (2022).
- [191] L. V. Drummond, A. G. Hanselman, D. R. Becker, and S. A. Hughes, Extreme mass-ratio inspiral of a spinning body into a Kerr black hole I: Evolution along generic trajectories, [arXiv:2305.08919](#).
- [192] X. Feng, Z. Lyu, and H. Yang, Black-hole perturbation theory with post-Newtonian theory: Towards hybrid waveforms for neutron-star binaries, *Phys. Rev. D* **105**, 104043 (2022).
- [193] M. Rahman and A. Bhattacharyya, Prospects for determining the nature of the secondaries of extreme mass-ratio inspirals using the spin-induced quadrupole deformation, *Phys. Rev. D* **107**, 024006 (2023).
- [194] L. Xu, S. Yang, W. Han, X. Zhong, R. Tang, and Y. Zhang, Distinguishing compact objects in extreme-mass-ratio inspirals by gravitational waves, *Universe* **11**, 18 (2025).
- [195] L. Blanchet, Post-Newtonian theory for gravitational waves, *Living Rev. Relativity* **17**, 2 (2014).
- [196] L. Blanchet, G. Faye, Q. Henry, F. Larrouturou, and D. Trestini, Gravitational-wave phasing of quasicircular compact binary systems to the fourth-and-a-half post-Newtonian order, *Phys. Rev. Lett.* **131**, 121402 (2023).
- [197] L. Pompili *et al.*, Laying the foundation of the effective-one-body waveform models SEOBNRv5: Improved accuracy and efficiency for spinning nonprecessing binary black holes, *Phys. Rev. D* **108**, 124035 (2023).
- [198] S. Albanesi, R. Gamba, S. Bernuzzi, J. Fontbuté, A. Gonzalez, and A. Nagar, Effective-one-body modeling for generic compact binaries with arbitrary orbits, [arXiv:2503.14580](#).
- [199] S. A. Hughes, Adiabatic and post-adiabatic approaches to extreme mass ratio inspiral, in *14th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories* (2017), Vol. 2, pp. 1953–1959.
- [200] P. Lynch and O. Burke, A note on the conversion of orbital angles for extreme mass ratio inspirals, [arXiv:2411.04955](#).
- [201] R. Fujita and W. Hikida, Analytical solutions of bound timelike geodesic orbits in Kerr spacetime, *Classical Quantum Gravity* **26**, 135002 (2009).
- [202] N. Warburton, B. Wardell, O. Long, S. Upton, P. Lynch, Z. Nasipak, and L. C. Stein, Kerrgeodesics, [10.5281/zenodo.8108265](#) (2023).
- [203] S. Babak *et al.*, LISA Rosetta Stone, LISA DDPC Convention Guide, 2025, internal DDPC document (to be published).
- [204] E. E. Flanagan and T. Hinderer, Transient resonances in the inspirals of point particles into black holes, *Phys. Rev. Lett.* **109**, 071102 (2012).
- [205] M. van de Meent, Conditions for sustained orbital resonances in extreme mass ratio inspirals, *Phys. Rev. D* **89**, 084033 (2014).
- [206] C. P. L. Berry, R. H. Cole, P. Cañizares, and J. R. Gair, Importance of transient resonances in extreme-mass-ratio inspirals, *Phys. Rev. D* **94**, 124042 (2016).
- [207] M. van de Meent, Resonantly enhanced kicks from equatorial small mass-ratio inspirals, *Phys. Rev. D* **90**, 044027 (2014).
- [208] L. C. Stein and N. Warburton, Location of the last stable orbit in Kerr spacetime, *Phys. Rev. D* **101**, 064007 (2020).

- [209] A. Ori and K. S. Thorne, The transition from inspiral to plunge for a compact body in a circular equatorial orbit around a massive, spinning black hole, *Phys. Rev. D* **62**, 124022 (2000).
- [210] G. Compère and L. Küchler, Self-consistent adiabatic inspiral and transition motion, *Phys. Rev. Lett.* **126**, 241106 (2021).
- [211] G. Compère and Lorenzo Küchler, Asymptotically matched quasi-circular inspiral and transition-to-plunge in the small mass ratio expansion, *SciPost Phys.* **13**, 043 (2022).
- [212] S. A. Hughes, Parameterizing black hole orbits for adiabatic inspiral, *Phys. Rev. D* **109**, 064077 (2024).
- [213] T. Ng and E. Teo, Parametric solutions to the Kerr separatrix, *Classical Quantum Gravity* **42**, 075003 (2025).
- [214] G. Compère, Y. Liu, and J. Long, Classification of radial Kerr geodesic motion, *Phys. Rev. D* **105**, 024075 (2022).
- [215] S. A. Teukolsky, Perturbations of a rotating black hole. I. Fundamental equations for gravitational, electromagnetic, and neutrino-field perturbations, *Astrophys. J.* **185**, 635 (1973).
- [216] V. Skoupý and G. Lukes-Gerakopoulos, Spinning test body orbiting around a Kerr black hole: Eccentric equatorial orbits and their asymptotic gravitational-wave fluxes, *Phys. Rev. D* **103**, 104045 (2021).
- [217] Z. Nasipak, znasipak/pybhpt: v0.9.0, [10.5281/zenodo.15627818](https://doi.org/10.5281/zenodo.15627818) (2025).
- [218] M. van de Meent, Analytic solutions for parallel transport along generic bound geodesics in Kerr spacetime, *Classical Quantum Gravity* **37**, 145007 (2020).
- [219] C. Chapman-Bird, Z. Nasipak, and M. PIGOU, znasipak/multispline: v0.8.3, [10.5281/zenodo.15188413](https://doi.org/10.5281/zenodo.15188413) (2025).
- [220] G. Behforooz, End conditions for cubic spline interpolation derived from integration, *Appl. Math. Comput.* **29**, 231 (1989).
- [221] GSL—GNU Scientific Library, (<https://www.gnu.org/software/gsl/>) (1996).
- [222] E. Hairer, S. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff Problems*, ser. Springer Series in Computational Mathematics (Springer, Berlin, Heidelberg, 1993), pp. 192–193.
- [223] B. Kocsis, N. Yunes, and A. Loeb, Observable signatures of EMRI black hole binaries embedded in thin accretion disks, *Phys. Rev. D* **84**, 024032 (2011).
- [224] M. Vallisneri, Use and abuse of the Fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects, *Phys. Rev. D* **77**, 042001 (2008).
- [225] I. Gladwell, L. Shampine, and R. Brankin, Automatic selection of the initial step size for an ode solver, *J. Comput. Appl. Math.* **18**, 175 (1987).
- [226] R. Fujita (private communication).
- [227] L. Lindblom, B. J. Owen, and D. A. Brown, Model waveform accuracy standards for gravitational wave data analysis, *Phys. Rev. D* **78**, 124020 (2008).
- [228] M. L. Katz, A. J. K. Chua, L. Speri, N. Warburton, and S. A. Hughes, Fast extreme-mass-ratio-inspiral waveforms: New tools for millihertz gravitational-wave data analysis, *Phys. Rev. D* **104**, 064047 (2021).
- [229] R. Fujita and H. Tagoshi, New numerical methods to evaluate homogeneous solutions of the Teukolsky equation, *Prog. Theor. Phys.* **112**, 415 (2004).
- [230] R. Fujita and H. Tagoshi, New numerical methods to evaluate homogeneous solutions of the Teukolsky equation II. Solutions of the continued fraction equation, *Prog. Theor. Phys.* **113**, 1165 (2005).
- [231] Black Hole Perturbation Club (BHPC), (<https://sites.google.com/view/bhpc1996/home>) (2003).
- [232] W. T. Thrope, High precision calculation of generic extreme mass ratio inspirals, Cambridge, MA, May 2010. [Online]. Available: <http://hdl.handle.net/1721.1/61270>.
- [233] S. O’Sullivan and S. A. Hughes, Strong-field tidal distortions of rotating black holes: Formalism and results for circular, equatorial orbits, *Phys. Rev. D* **90**, 124039 (2014); *Phys. Rev. D* **91**, 109901(E) (2015).
- [234] R. Pordes, D. Petravick, B. Kramer, D. Olson, M. Livny, A. Roy, P. Avery, K. Blackburn, T. Wenaus, F. Würthwein, I. Foster, R. Gardner, M. Wilde, A. Blatecky, J. McGee, and R. Quick, The open science grid, *J. Phys. Conf. Ser.* **78**, 012057 (2007).
- [235] H. Khalvati, P. Lynch, O. Burke, L. Speri, M. van de Meent, and Z. Nasipak, Systematic errors in fast relativistic waveforms for Extreme Mass Ratio Inspirals, [arXiv:2509.08875](https://arxiv.org/abs/2509.08875).
- [236] M. A. Miller, Accuracy requirements for the calculation of gravitational waveforms from coalescing compact binaries in numerical relativity, *Phys. Rev. D* **71**, 104016 (2005).
- [237] E. E. Flanagan and S. A. Hughes, Measuring gravitational waves from binary black hole coalescences: 2. The waves’ information and its extraction, with and without templates, *Phys. Rev. D* **57**, 4566 (1998).
- [238] C. Cutler and M. Vallisneri, LISA detections of massive black hole inspirals: Parameter extraction errors due to inaccurate template waveforms, *Phys. Rev. D* **76**, 104018 (2007).
- [239] K. Chatziioannou, J. A. Clark, A. Bauswein, M. Millhouse, T. B. Littenberg, and N. Cornish, Inferring the post-merger gravitational wave emission from binary neutron star coalescences, *Phys. Rev. D* **96**, 124035 (2017).
- [240] M. Pürrer and C.-J. Haster, Gravitational waveform accuracy requirements for future ground-based detectors, *Phys. Rev. Res.* **2**, 023151 (2020).
- [241] A. Toubiana and J. R. Gair, Indistinguishability criterion and estimating the presence of biases, [arXiv:2401.06845](https://arxiv.org/abs/2401.06845).
- [242] A. J. K. Chua, M. L. Katz, N. Warburton, and S. A. Hughes, Rapid generation of fully relativistic extreme-mass-ratio-inspiral waveform templates for LISA data analysis, *Phys. Rev. Lett.* **126**, 051102 (2021).
- [243] L. Barack and C. Cutler, LISA capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy, *Phys. Rev. D* **69**, 082005 (2004).
- [244] S. Babak, H. Fang, J. R. Gair, K. Glampedakis, and S. A. Hughes, ‘Kludge’ gravitational waveforms for a test-body orbiting a Kerr black hole, *Phys. Rev. D* **75**, 024005 (2007); *Phys. Rev. D* **77**, 04990(E) (2008).

- [245] A. J. K. Chua, C. J. Moore, and J. R. Gair, Augmented kludge waveforms for detecting extreme-mass-ratio inspirals, *Phys. Rev. D* **96**, 044005 (2017).
- [246] M. Bonetti and A. Sesana, Gravitational wave background from extreme mass ratio inspirals, *Phys. Rev. D* **102**, 103023 (2020).
- [247] M. Piarulli, R. Busicchio, F. Pozzoli, O. Burke, M. Bonetti, and A. Sesana, Test for LISA foreground Gaussianity and stationarity: Extreme mass-ratio inspirals, *Phys. Rev. D* **111**, 103047 (2025).
- [248] F. Pozzoli, S. Babak, A. Sesana, M. Bonetti, and N. Karnesis, Computation of stochastic background from extreme-mass-ratio inspiral populations for LISA, *Phys. Rev. D* **108**, 103039 (2023); *Phys. Rev. D* **110**, 049903(E) (2024).
- [249] P. C. Peters and J. Mathews, Gravitational radiation from point masses in a Keplerian orbit, *Phys. Rev.* **131**, 435 (1963).
- [250] S. Kejriwal, L. Speri, and A. J. K. Chua, Impact of correlations on the modeling and inference of beyond vacuum-general relativistic effects in extreme-mass-ratio inspirals, *Phys. Rev. D* **110**, 084060 (2024).
- [251] H. Tagoshi, PostNewtonian expansion of gravitational waves from a particle in slightly eccentric orbit around a rotating black hole, *Prog. Theor. Phys.* **93**, 307 (1995); *Prog. Theor. Phys.* **118**, 577(E) (2007).
- [252] R. Fujita and B. R. Iyer, Spherical harmonic modes of 5.5 post-Newtonian gravitational wave polarisations and associated factorised resummed waveforms for a particle in circular orbit around a Schwarzschild black hole, *Phys. Rev. D* **82**, 044051 (2010).
- [253] P. C. Mahalanobis, Reprint of: Mahalanobis, p.c. (1936) “on the generalised distance in statistics”, *Sankhya A* **80**, 1 (2018).
- [254] J. R. Gair, L. Barack, T. Creighton, C. Cutler, S. L. Larson, E. S. Phinney, and M. Vallisneri, Event rate estimates for LISA extreme mass ratio capture sources, *Classical Quantum Gravity* **21**, S1595 (2004).
- [255] A. J. K. Chua and C. J. Cutler, Nonlocal parameter degeneracy in the intrinsic space of gravitational-wave signals from extreme-mass-ratio inspirals, *Phys. Rev. D* **106**, 124046 (2022).
- [256] J. E. Greene, J. Strader, and L. C. Ho, Intermediate-mass black holes, *Annu. Rev. Astron. Astrophys.* **58**, 257 (2020).
- [257] M. Bonetti and A. Sesana, Gravitational wave background from extreme mass ratio inspirals, *Phys. Rev. D* **102**, 103023 (2020).
- [258] N. Aghanim *et al.*, Planck 2018 results. VI. Cosmological parameters, *Astron. Astrophys.* **641**, A6 (2020); *Astron. Astrophys.* **652**, C4(E) (2021).
- [259] O. Lahav and A. R. Liddle, The cosmological parameters, [arXiv:2403.15526](https://arxiv.org/abs/2403.15526).
- [260] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (The MIT Press, Cambridge, MA, 2005), [10.7551/mitpress/3206.001.0001](https://doi.org/10.7551/mitpress/3206.001.0001).
- [261] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, Scikit-learn: Machine learning in Python, *J. Mach. Learn. Res.* **12**, 2825 (2011).
- [262] B. Rom, I. Linial, K. Kaur, and R. Sari, Dynamics around supermassive black holes: Extreme-mass-ratio Inspirals as gravitational-wave sources, *Astrophys. J.* **977**, 7 (2024).
- [263] L. Broggi, E. Bortolas, M. Bonetti, A. Sesana, and M. Dotti, Extreme mass ratio inspirals and tidal disruption events in nuclear clusters—I. Time-dependent rates, *Mon. Not. R. Astron. Soc.* **514**, 3270 (2022).
- [264] Y.-P. Li, H. Yang, and Z. Pan, Extreme mass-ratio inspirals in active galactic nucleus disks: The role of circumsingle disks, *Phys. Rev. D* **111**, 063074 (2025).
- [265] T. F. M. Spieksma and E. Cannizzaro, In the grip of the disk: Dragging the companion through an AGN, [arXiv:2504.08033](https://arxiv.org/abs/2504.08033).
- [266] F. D. Ryan, Effect of gravitational radiation reaction on circular orbits around a spinning black hole, *Phys. Rev. D* **52**, R3159 (1995).
- [267] D. Kennefick and A. Ori, Radiation reaction induced evolution of circular orbits of particles around Kerr black holes, *Phys. Rev. D* **53**, 4319 (1996).
- [268] R. Fujita, S. Isoyama, A. Le Tiec, H. Nakano, N. Sago, and T. Tanaka, Hamiltonian formulation of the conservative self-force dynamics in the Kerr geometry, *Classical Quantum Gravity* **34**, 134001 (2017).
- [269] G. Huez, S. Bernuzzi, M. Breschi, and R. Gamba, Gravitational waves from eccentric binary neutron star mergers: Systematic biases induced by quasi-circular templates, [arXiv:2504.18622](https://arxiv.org/abs/2504.18622).
- [270] M. L. Katz, J.-B. Bayle, A. J. K. Chua, and M. Vallisneri, Assessing the data-analysis impact of LISA orbit approximations using a GPU-accelerated response model, *Phys. Rev. D* **106**, 103001 (2022).
- [271] N. Karnesis, M. L. Katz, N. Korsakova, J. R. Gair, and N. Stergioulas, Eryn: A multipurpose sampler for Bayesian inference, *Mon. Not. R. Astron. Soc.* **526**, 4814 (2023).
- [272] M. Katz, N. Karnesis, and N. Korsakova, [mikekatz04/eryn: First full release, 10.5281/zenodo.7705496](https://arxiv.org/abs/2305.10528) (2023).
- [273] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, emcee: The MCMC Hammer, *Astron Soc Pac.* **125**, 306 (2013).
- [274] E. Barausse, The evolution of massive black holes and their spins in their galactic hosts, *Mon. Not. R. Astron. Soc.* **423**, 2533 (2012).
- [275] R. Gray *et al.*, Joint cosmological and gravitational-wave population inference using dark sirens and galaxy catalogues, *J. Cosmol. Astropart. Phys.* **12** (2023) 023.
- [276] B. F. Schutz, Determining the Hubble constant from gravitational wave observations, *Nature (London)* **323**, 310 (1986).
- [277] S. Yi, F. Iacovelli, S. Marsat, D. Wadekar, and E. Berti, Systematic biases from the exclusion of higher harmonics in parameter estimation on LISA binaries, [arXiv:2502.12237](https://arxiv.org/abs/2502.12237).
- [278] S. Marsat, J. G. Baker, and T. Dal Canton, Exploring the Bayesian parameter estimation of binary black holes with LISA, *Phys. Rev. D* **103**, 083011 (2021).
- [279] A. Toubiana, L. Pompili, A. Buonanno, J. R. Gair, and M. L. Katz, Measuring source properties and quasinormal

- mode frequencies of heavy massive black-hole binaries with LISA, *Phys. Rev. D* **109**, 104019 (2024).
- [280] D. C. Wilkins, Bound geodesics in the Kerr metric, *Phys. Rev. D* **5**, 814 (1972).
- [281] E. Teo, Spherical photon orbits around a Kerr black hole, *Gen. Relativ. Gravit.* **35**, 1909 (2003).
- [282] E. Teo, Spherical orbits around a Kerr black hole, *Gen. Relativ. Gravit.* **53**, 10 (2021).
- [283] Black Hole Perturbation Toolkit, <https://bhptoolkit.org/> (2018).
- [284] A. Le Tiec, A. H. Mroue, L. Barack, A. Buonanno, H. P. Pfeiffer, N. Sago, and A. Taracchini, Periastron advance in black hole binaries, *Phys. Rev. Lett.* **107**, 141101 (2011).
- [285] M. van de Meent and H. P. Pfeiffer, Intermediate mass-ratio black hole binaries: Applicability of small mass-ratio perturbation theory, *Phys. Rev. Lett.* **125**, 181101 (2020).
- [286] B. Leather, A. Buonanno, and M. van de Meent, Inspiral-merger-ringdown waveforms with gravitational self-force results within the effective-one-body formalism, *arXiv:2505.11242*.
- [287] A. Taracchini, A. Buonanno, G. Khanna, and S. A. Hughes, Small mass plunging into a Kerr black hole: Anatomy of the inspiral-merger-ringdown waveforms, *Phys. Rev. D* **90**, 084025 (2014).
- [288] H. Lim, G. Khanna, A. Apte, and S. A. Hughes, Exciting black hole modes via misaligned coalescences: II. The mode content of late-time coalescence waveforms, *Phys. Rev. D* **100**, 084032 (2019).
- [289] U. Ruangsri and S. A. Hughes, Census of transient orbital resonances encountered during binary inspiral, *Phys. Rev. D* **89**, 084036 (2014).
- [290] J. Brink, M. Geyer, and T. Hinderer, Astrophysics of resonant orbits in the Kerr metric, *Phys. Rev. D* **91**, 083001 (2015).
- [291] E. Levati, A. Cárdenas-Avendaño, K. Destounis, and P. Pani, Cumulative effect of orbital resonances in extreme-mass-ratio inspirals, *Phys. Rev. D* **111**, 104006 (2025).
- [292] C. M. Hirata, Resonant recoil in extreme mass ratio binary black hole mergers, *Phys. Rev. D* **83**, 104024 (2011).
- [293] P. Gupta, B. Bonga, A. J. K. Chua, and T. Tanaka, Importance of tidal resonances in extreme-mass-ratio inspirals, *Phys. Rev. D* **104**, 044056 (2021).
- [294] P. Gupta, T. Kakehi, and T. Tanaka, Resonant jumps induced by stationary tidal perturbation: A two-for-one deal, *Classical Quantum Gravity* **39**, 245005 (2022).
- [295] T. A. Apostolatos, G. Lukes-Gerakopoulos, and G. Contopoulos, How to observe a non-Kerr spacetime using gravitational waves, *Phys. Rev. Lett.* **103**, 111101 (2009).
- [296] Z. Pan, H. Yang, L. Bernard, and B. Bonga, Resonant dynamics of extreme mass-ratio inspirals in a perturbed Kerr spacetime, *Phys. Rev. D* **108**, 104026 (2023).
- [297] L. Speri and J. R. Gair, Assessing the impact of transient orbital resonances, *Phys. Rev. D* **103**, 124032 (2021).
- [298] Y. Xu, M. Rosselló-Sastre, S. Tiwari, M. Ebersold, E. Z. Hamilton, C. García-Quirós, H. Estellés, and S. Husa, Enhancing gravitational wave parameter estimation with nonlinear memory: Breaking the distance-inclination degeneracy, *Phys. Rev. D* **109**, 123034 (2024).
- [299] M. Rosselló-Sastre, S. Husa, S. Bera, and Y. Xu, Impact of the ($\ell = 2, m = 0$) spherical harmonic mode with memory on parameter estimation for ground-based detectors, *arXiv:2506.05859*.
- [300] T. Islam, S. E. Field, G. Khanna, and N. Warburton, Survey of gravitational wave memory in intermediate mass ratio binaries, *Phys. Rev. D* **108**, 024046 (2023).
- [301] K. Cunningham, C. Kavanagh, A. Pound, D. Trestini, N. Warburton, and J. Neef, Gravitational memory: New results from post-Newtonian and self-force theory, *arXiv:2410.23950*.
- [302] A. Elhashash and D. A. Nichols, Waveform models for the gravitational-wave memory effect: Extreme mass-ratio limit and final memory offset, *Phys. Rev. D* **111**, 044052 (2025).
- [303] A. Spiers, A. Pound, and J. Moxon, Second-order Teukolsky formalism in Kerr spacetime: Formulation and nonlinear source, *Phys. Rev. D* **108**, 064002 (2023).
- [304] B. Leather and N. Warburton, Applying the effective-source approach to frequency-domain self-force calculations for eccentric orbits, *Phys. Rev. D* **108**, 084045 (2023).
- [305] T. Osburn and N. Nishimura, New self-force method via elliptic partial differential equations for Kerr inspiral models, *Phys. Rev. D* **106**, 044056 (2022).
- [306] Y.-X. Wei, J.-d. Zhang, and J. Mei, Second-order self-force for eccentric extreme-mass ratio inspirals on schwarzschild spacetime, *arXiv:2504.09640*.
- [307] E. Harms, G. Lukes-Gerakopoulos, S. Bernuzzi, and A. Nagar, Asymptotic gravitational wave fluxes from a spinning particle in circular equatorial orbits around a rotating black hole, *Phys. Rev. D* **93**, 044015 (2016); *Phys. Rev. D* **100**, 129901(A) (2019).
- [308] S. Akcay, S. R. Dolan, C. Kavanagh, J. Moxon, N. Warburton, and B. Wardell, Dissipation in extreme-mass ratio binaries with a spinning secondary, *Phys. Rev. D* **102**, 064013 (2020).
- [309] G. A. Piovano, A. Maselli, and P. Pani, Extreme mass ratio inspirals with spinning secondary: A detailed study of equatorial circular motion, *Phys. Rev. D* **102**, 024041 (2020).
- [310] J. Mathews, A. Pound, and B. Wardell, Self-force calculations with a spinning secondary, *Phys. Rev. D* **105**, 084031 (2022).
- [311] A. Albertini, A. Nagar, J. Mathews, and G. Lukes-Gerakopoulos, Comparing second-order gravitational self-force and effective-one-body waveforms from inspiralling, quasicircular black hole binaries with a nonspinning primary and a spinning secondary, *Phys. Rev. D* **110**, 044034 (2024).
- [312] C. Chapman-Bird (to be published).
- [313] I. Sfiligoi, D. C. Bradley, B. Holzman, P. Mhashilkar, S. Padhi, and F. Wurthwein, The pilot way to grid resources using glideinwms, in *2009 WRI World Congress on Computer Science and Information Engineering* (2009), Vol. 2, pp. 428–432.
- [314] OSG, Ospoool, 2006. [Online]. Available: https://osg-htc.org/services/open_science_pool.html.
- [315] Open science data federation, 2015. [Online]. Available: <https://osdf.osg-htc.org/>.
- [316] C. Chapman-Bird, L. Speri, Z. Nasipak, O. Burke, M. L. Katz, A. Santini, S. Kejriwal, P. Lynch, J. Mathews, H.

- Khalvati, J. Thompson, S. Isoyama, S. A. Hughes, N. Warburton, A. J. K. Chua, and M. Pigou, Data release for “efficient waveforms for asymmetric-mass eccentric equatorial inspirals into rapidly-spinning black holes”, [10.5281/zenodo.15631641](https://doi.org/10.5281/zenodo.15631641) (2025).
- [317] C. Chapman-Bird, A. J. Chua, S. A. Hughes, M. L. Katz, Z. Nasipak, and N. Warburton, Data for fast emri waveforms, [10.5281/zenodo.15624459](https://doi.org/10.5281/zenodo.15624459) (2025).
- [318] J. C. Bardwell, Improving adiabatic trajectories of eccentric-equatorial extreme mass ratio inspirals with spinning massive black hole, Ph.D. dissertation, San Diego State University, 2024, copyright—Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated—2025-03-14. [Online]. Available: <https://www.proquest.com/dissertations-theses/improving-adiabatic-trajectories-eccentric/docview/3152477607/se-2>.
- [319] L. Speri, N. Karnesis, A. I. Renzini, and J. R. Gair, A roadmap of gravitational wave data analysis, *Nat. Astron.* **6**, 1356 (2022).
- [320] F. Pozzoli, R. Buscicchio, A. Klein, V. Korol, A. Sesana, and F. Haardt, Cyclostationary signals in LISA: A practical application to Milky Way satellites, *Phys. Rev. D* **111**, 063005 (2025).
- [321] M. C. Digman and N. J. Cornish, LISA gravitational wave sources in a time-varying galactic stochastic background, *Astrophys. J.* **940**, 10 (2022).
- [322] P. Whittle, Curve and periodogram smoothing, *J. R. Stat. Soc. Ser. B (Statistical Methodology)* **19**, 38 (1957).
- [323] T. A. Prince, M. Tinto, S. L. Larson, and J. W. Armstrong, The LISA optimal sensitivity, *Phys. Rev. D* **66**, 122002 (2002).
- [324] S. Babak, M. Hewitson, and A. Petiteau, LISA sensitivity and SNR calculations, [arXiv:2108.01167](https://arxiv.org/abs/2108.01167).
- [325] N. Karnesis, S. Babak, M. Pieroni, N. Cornish, and T. Littenberg, Characterization of the stochastic signal originating from compact binary populations as measured by LISA, *Phys. Rev. D* **104**, 043019 (2021).
- [326] T. Robson, N. J. Cornish, and C. Liu, The construction and use of LISA sensitivity curves, *Classical Quantum Gravity* **36**, 105011 (2019).
- [327] S. Kejriwal, O. Burke, and C. Chapman-Bird, Stableemri-fisher (sef), manuscript in preparation. [Online]. Available: <https://github.com/perturber/StableEMRIFisher>.
- [328] N. Sago and R. Fujita, Calculation of radiation reaction effect on orbital parameters in Kerr spacetime, *Prog. Theor. Exp. Phys.* **2015**, 073E03 (2015).
- [329] D. Bini, T. Damour, and A. Geralico, High post-Newtonian order gravitational self-force analytical results for eccentric equatorial orbits around a Kerr black hole, *Phys. Rev. D* **93**, 124058 (2016).
- [330] D. Bini and A. Geralico, New gravitational self-force analytical results for eccentric equatorial orbits around a Kerr black hole: Redshift invariant, *Phys. Rev. D* **100**, 104002 (2019).
- [331] C. Munna, High-order post-Newtonian expansion of the generalized redshift invariant for eccentric-orbit, equatorial extreme-mass-ratio inspirals with a spinning primary, *Phys. Rev. D* **108**, 084012 (2023).
- [332] V. Skoupý and V. Witzany, Post-Newtonian expansions of extreme mass ratio inspirals of spinning bodies into Schwarzschild black holes, *Phys. Rev. D* **110**, 084061 (2024).
- [333] D. Trestini, Z. Nasipak, and A. Pound (private communication).