Probing dark-matter effects with gravitational waves using the parametrized post-Einsteinian framework

Eileen Wilcox, David A. Nichols,[*] and Kent Yagi[†]

Department of Physics, University of Virginia,
P.O. Box 400714, Charlottesville, Virginia 22904-4714, USA

A massive black hole can develop a dark-matter overdensity, and the dark matter changes the evolution of a stellar-mass compact object inspiraling around the massive black hole through the dense dark-matter environment. Specifically, dynamical friction speeds up the inspiral of the compact object and causes feedback on the dark-matter distribution. These intermediate mass-ratio inspirals with dark matter are a source of gravitational waves (GWs), and the waves can dephase significantly from an equivalent system in vacuum. Prior work has shown that this dephasing needs to be modeled to detect the GWs from these systems with LISA (the Laser Interferometer Space Antenna); it also showed that the density and distribution of dark matter can be inferred from a GW measurement. In this paper, we study whether the parametrized post-Einsteinian (ppE) framework can be used to infer the presence of dark matter in these systems. We confirm that if vacuum waveform templates are used to model the GWs from an inspiral in a dark-matter halo, then the resulting parameter estimation is biased. We then apply the ppE framework to determine whether it can reduce the parameter-estimation biases, and we find that adding one ppE phase term to a waveform template eliminates the parameter-estimation biases (statistical errors become larger than the systematic ones), but the effective post-Newtonian order in the ppE framework must be specified without uncertainties. When the post-Newtonian order has uncertainty, we find that the systematic errors on the ppE and the binary's parameters exceed the statistical errors. Thus, the simplest ppE framework would not give unbiased results for these systems, and a further extension of it, or dedicated parameter estimation with gravitational waveforms that include dark-matter effects would be needed.

I. INTRODUCTION

There is compelling evidence for the presence of dark matter (DM) over a wide range of length scales, and observations indicate that it is abundant, cold, and weakly interacting [1]. The Planck satellite, for example, determined that dark matter accounts for 26.8% of the mass-energy content in the Universe [2]. There have been extensive experimental and observational programs to detect the particles that compose dark matter, but given the feeble interactions between it and baryonic matter or its small self-interactions, these efforts have not yet determined the fundamental constituents of dark matter (see, e.g., the reviews [3–6]). This has motivated not only improving the existing searches for dark matter, but also exploring new methods to determine its nature [7]. One recent effort is using gravitational waves (GWs) to search for the presence of high densities of dark matter that can form around black holes [8].

LIGO's successful detection of GWs in 2015 [9] not only confirmed a long-standing prediction of Einstein's theory of relativity, but also created a new opportunity to study fundamental gravitational physics [10,11]. Nearly 100 confident detections of mergers of compact objects have now been made [12–14], which have tested the predictions of general relativity precisely using an large battery of tests [15–19]. The spectrum of GW observations has opened further with the likely detection of a stochastic GW background by pulsar timing arrays [20–23]. It will continue to expand in the next decade with next-generation GW detectors, such as the planned LISA [24] (in space) and Einstein Telescope [25,26] or Cosmic Explorer [27]. These new detectors will not only expand the GW frequency spectrum, but also provide new opportunities to study fundamental physics.

Our focus in this paper will be on the LISA detector and the possibility of using the GWs from the inspiral of a neutron star (NS) into an intermediate-mass black hole (IMBH) with a surrounding dense distribution of dark matter, as a means to learn about the surrounding dark-matter environment. The inspiral of a compact object into an IMBH—known as an intermediate mass-ratio inspiral

---

[*]Contact author: david.nichols@virginia.edu
[†]Contact author: kyagi@virginia.edu

(IMRI)—or a similar inspiral with a supermassive BH primary (an extreme mass-ratio inspiral) are a key GW source class for the LISA detector. These systems have been well studied in the absence of a surrounding DM distribution, and have been established to provide precise information about the spacetime geometry of the primary massive BH (see, e.g., [28]). A nonvacuum environment could alter the IMRI's rate of inspiral and affect the GWs emitted from these systems. In fact, prior work has confirmed this: if dark matter is present around an IMBH in sufficiently large amounts, the density of a spherically symmetric DM overdensity (referred to as a DM "spike" [29]) could be inferred from a GW measurement with LISA [30,31].

Dark-matter particles scatter gravitationally with the NS as it inspirals, and there is a net transfer of energy from the NS to the DM spike; this effect, known as "dynamical friction" [32–34], causes the NS to inspiral into the IMBH more rapidly than an equivalent IMRI in vacuum would. Previous work by Eda *et al.* [31] showed that dynamical friction led to a large change in the rate of inspiral that allowed the distribution of dark matter to be inferred precisely. For much of the parameter space of IMRIs with dark matter in [31], it was subsequently shown in [35] that neglecting feedback from dynamical friction onto the DM distribution significantly overestimated the increase in the rate of inspiral. Reference [35] introduced a formalism to evolve the DM spike that takes into account the energy imparted to the DM distribution through dynamical friction. The IMRI's orbit and the dark matter distribution were evolved simultaneously as a coupled set of ordinary and partial differential equations over the longer timescales associated with gravitational radiation reaction and dynamical friction. The DM density in the vicinity of the NS secondary in this dynamic case was transiently depleted during the inspiral, which decreased the effect of dynamical friction and the amount of dephasing from a vacuum IMRI with the same masses. The GW dephasing was also a more nontrivial function of frequency in the dynamic case than in the static (nonevolving) case considered in [31].

To determine how well the IMRIs with a dynamic DM distribution could be studied by LISA, Ref. [36] performed a simulation study of the prospects for measurement, detection, and distinction from vacuum IMRIs. It was shown in [36] that the detection horizon for IMRIs with DM had very little difference from vacuum binaries; however, vacuum IMRI waveforms were also not capable of fitting an IMRI with a dynamic DM distribution when optimized over the vacuum IMRI parameters. Thus, Ref. [36] created a phenomenological analytical model of the GW phase in the frequency domain to perform Bayesian parameter-estimation studies. In the two test cases considered in [36], the power law that determines the radial falloff of the initial DM density could be inferred to a few and a few tens of a percent, respectively. This established that properties of the DM could still be inferred in the dynamic case via a GW measurement with LISA, and that the previous studies that assumed the DM was static during the inspiral significantly overestimated the precision to which the DM distribution could be measured.

Since presence of dark matter in an IMRI can produce a measurable GW dephasing from analogous vacuum systems, it is useful to know how precisely this dephasing needs to be modeled to be detectable and have minimal bias in the inferred GW parameters. While using the true signal is the optimal choice for matched-filtering-based data analysis, there are more model-agnostic frameworks, such as the parametrized post-Einsteinian (ppE) approach [37–39], which aim to capture deviations from vacuum waveforms in general relativity (GR). The ppE framework has been shown to detect deviations from GR in simulation studies [40–42] and also has been applied to the existing GW events to constrain such deviations [10,11,17–19]. Parametrized waveforms similar to those used in the ppE framework have also been applied to study effects of the astrophysical environment in IMRIs, in particular those induced by dark matter (see in [43–45]). However, neither of the signal models used in [43,45] were as complex as the phase in the dynamic case in [36], so it has not been established how well the ppE framework can capture effects induced by frequency-domain signals with a more complicated morphology.

In this paper, we address this question by investigating how well the ppE framework can account for these dynamic DM-induced GW dephasing from vacuum binaries. As a first calculation, we study the amount of systematic bias in parameter estimation if the GW signal is an IMRI with a surrounding distribution of dark matter, but a GR template waveform is used to analyze the GW signals. Next, we use a simple ppE waveform to analyze the GWs from an IMRI with dark matter. Our ppE model modifies the vacuum gravitational waveform phase in the frequency domain by the addition of a single term involving a power of GW frequency. This term has a constant overall scaling β and the exponent is denoted by b; the b coefficient determines the post-Newtonian (PN) order of the ppE term. We use a Fisher-matrix formalism to compute the statistical errors on waveform parameters and address whether the ppE template is able to capture the DM effects on the gravitational waveform.

Throughout this paper we focus on IMRIs in which the secondary is a neutron star, and the dark matter has negligible interactions with the nuclear matter in the neutron star. This assumption is necessary, because recent work [46] showed that a black-hole secondary can accrete enough mass during the inspiral to induce a dephasing of hundreds of GW cycles with respect to systems in which the effects of only dynamical friction were modeled. It also leads to nontrivial changes in the distribution of dark matter after the merger. Other work [47] also generalized the prescription to IMRIs on eccentric orbits.

We now summarize our main findings. We first determined that if we use the vacuum GR template waveform to analyze the GW signal from an IMRI in a DM spike, then vacuum parameters, such as the chirp mass, are significantly biased. This is consistent with the results in [36]. We next studied these IMRIs using ppE templates. In this context, we first found the value of ppE exponent b by minimizing the systematic error on the IMRI's "vacuum" (those unrelated to the DM) parameters. We assumed perfect knowledge of b (i.e., b was fixed) and computed the uncertainties on the ppE parameter β. This procedure decreased the systematic bias on vacuum parameters below their systematic errors, and the systematic error on β was much smaller than the statistical error, too. The systematic error on the vacuum parameters, however, is very sensitive to the choice of b, which is not expected to be known *a priori* and requires fine tuning to minimize the systematic errors. Consequently, we also performed another analysis in which we compute uncertainties on both β and b; in this context, we found that systematic errors are now larger than statistical errors. This suggests that the ppE template has the potential to remove the bias and measure the GW effects induced by the DM spike, but future work is necessary to determine if more general ppE waveforms with the larger parameter sets could improve the level of systematic and statistical errors.

The rest of the paper is organized as follows. Section II explains the waveform models that we use for IMRIs in a dark-matter environment and for the ppE framework. In Sec. III, we review the Fisher framework that we use to compute statistical and systematic errors. Section IV presents the main questions that we address in this paper and the methods we use to answer them. Our main results are given in Sec. V, and they are followed by our conclusions in Sec. VI. We give a few supplemental results in two Appendixes. We use the geometric units $c = G_N = 1$ throughout this paper.

II. GRAVITATIONAL WAVES FROM A COMPACT-BINARY INSPIRAL IN A DARK-MATTER HALO

A schematic of the IMRIs that we will study in this paper is given in Fig. 1. We will assume that the IMBH has a mass m_1 between 10^3 and $10^5 M_\odot$, and will be the primary (more massive object) in the binary. The secondary (less massive object) will be a neutron star with mass $m_2 = 1.4 M_\odot$. Following Refs. [35,36], we assume that the initial DM density around the IMBH is a spherically symmetric and isotropic spike given by

$$\rho_{DM}(r) = \begin{cases} \rho_{sp}(r_{sp}/r)^{\gamma_{sp}} & r_{in} \leq r \leq r_{sp} \\ 0 & r < r_{in} \end{cases}. \qquad (2.1)$$

In Eq. (2.1), r denotes the distance from the IMBH, γ_{sp} is the radial power law, ρ_{sp} is the normalization of the density,
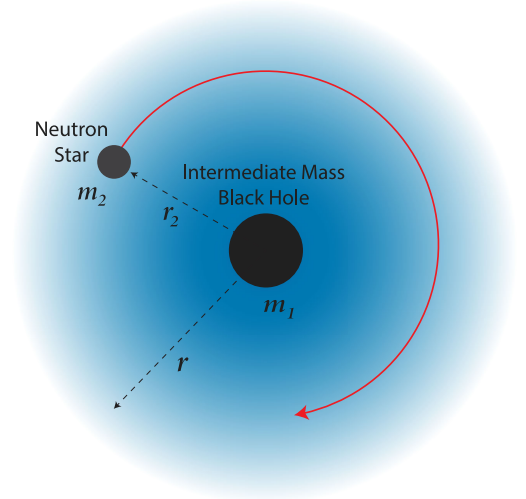

FIG. 1. Schematic of an IMRI with a DM spike. The IMBH has mass m_1 between 10^3 and $10^5 M_\odot$ and is orbited by a neutron star of mass $m_2 = 1.4 M_\odot$. The orbital radius is denoted by r_2, and r is the distance from the center of the IMBH.

$r_{in} = 4m_1$ is the inner radius of the spike, and r_{sp} is the outer radius of the spike.[1] Following [31,35], we do not treat r_{sp} as an independent parameter, but as being determined by $r_{sp} \approx 0.2r_h$, where r_h is defined by

$$\int_{r_{in}}^{r_h} \rho_{DM}(r) d^3 r = 2m_1. \qquad (2.2)$$

Throughout this paper, we will restrict to $\gamma_{sp} = 7/3$ and $\rho_{sp} = 226 M_\odot/\text{pc}^3$, which were the values used in the model of the spike in [30,31].

As discussed in Sec. I, the evolution of this IMRI must be performed simultaneously with the evolution of the dark matter. The HALOFEEDBACK code [49] produces the orbital or GW phase from such systems, but a single numerical evolution takes sufficiently long that it is inefficient to use it in parameter-estimation studies. Thus, Ref. [36] constructed a phenomenological waveform model for the time-domain phase (as a function of GW frequency), which was calibrated over different mass-ratio IMRIs for a range of different γ_{sp} and ρ_{sp} values for the initial density in Eq. (2.1). The phase is valid for IMRIs on circular orbits at Newtonian order; thus, it was used for just the dominant $l = 2$, $m = 2$ spherical-harmonic mode of the gravitational waveform.

We next review this phenomenological waveform, which we will use throughout the paper as the "true" GW signal.

---

[1]At radii $r > r_{sp}$, the spike will match onto the larger-scale DM halo, such as a Navarro-Frenk-White [48] profile; however, we will not treat the DM distribution at these larger radii in this paper.

A. Gravitational-wave phase in the time domain

The ansatz for the time-domain phenomenological phase in [36], when written as a function of the GW frequency, f, is given by

$$
\begin{aligned}
\Phi(f) \equiv \Phi^V(f)\{1 - \eta y^{-\lambda} \\
\times [1 - {}_2F_1(1, \vartheta, 1 + \vartheta, -y^{-5/(3\vartheta)})]\}.
\end{aligned}
\tag{2.3}
$$

Here, ${}_2F_1(a, b, c, z)$ is our notation for a Gaussian hypergeometric function, and the multiplicative factor $\Phi^V(f)$ is the vacuum phase:

$$
\Phi^V(f) = \frac{1}{16}(\pi \mathcal{M} f)^{-5/3}.
\tag{2.4}
$$

We used $\mathcal{M} = (m_1 + m_2)\nu^{3/5}$ to denote the chirp mass, and $\nu = m_1 m_2/(m_1 + m_2)^2$ for the symmetric mass ratio. The variable y is a dimensionless frequency, $y = f/f_t$, where f_t is a constant frequency scale.

To describe the parameter choices that reproduce the GW phase from IMRIs in static and dynamic DM distributions, it is useful to define the following quantities. We first introduce a frequency scale f_{eq} defined by

$$
f_{eq} = c_f^{3/(11-2\gamma_{sp})}.
\tag{2.5}
$$

The coefficient c_f is given by

$$
c_f = \frac{5}{8m_1^2} \pi^{2(\gamma_{sp}-4)/3} (m_1 + m_2)^{(1-\gamma_{sp})/3} r_{sp}^{\gamma_{sp}} \xi \rho_{sp} \log \Lambda.
\tag{2.6}
$$

The new parameters introduced in this expression are $\log \Lambda$, the Coulomb logarithm, where the value of Λ is given by $\Lambda = \sqrt{m_1/m_2}$ as in [35]. The factor ξ is the fraction of DM particles moving slower than the orbital speed of the secondary. It was shown in [46] to be given by

$$
\xi = 1 - I_{1/2}\left(\gamma_{sp} - \frac{1}{2}, \frac{3}{2}\right),
\tag{2.7}
$$

where $I_{\frac{1}{2}}$ is the regularized incomplete beta function. For $\gamma_{sp} = 7/3$, this has the numerical value $\xi \approx 0.58$ used in [35].

The ansatz for the phase in Eq. (2.3) is sufficiently general that it can be used to describe the GW phase for both static DM distributions in [31] and the dynamic ones in [35,36], for different choices of the parameters η, λ, ϑ, and f_t. By choosing

$$
\eta = 1, \quad \lambda = 0, \quad \vartheta = \frac{5}{11 - 2\gamma_{sp}}, \quad f_t = f_{eq},
\tag{2.8}
$$

one obtains the result for the GW phase $\Phi(f)$ for the IMRI in a static DM spike [Eq. (2.1)] at Newtonian order. The

Newtonian-order GW phase for the IMRI in a dynamic DM spike was obtained by choosing the parameters to be

$$
\eta = \frac{5}{8 - \gamma_{sp}} \left(\frac{f_{eq}}{f_b}\right)^{(11-2\gamma_{sp})/3},
\tag{2.9a}
$$

$$
\lambda = \frac{6 - 2\gamma_{sp}}{3},
\tag{2.9b}
$$

$$
\vartheta = 1,
\tag{2.9c}
$$

$$
f_t = f_b.
\tag{2.9d}
$$

An ansatz for the break frequency f_b was proposed in [36] to be

$$
f_b = \sigma \left(\frac{m_1}{1000 M_\odot}\right)^{-\alpha_1} \left(\frac{m_2}{M_\odot}\right)^{\alpha_2} \left(1 + \zeta \log \frac{\gamma_{sp}}{\gamma_r}\right),
\tag{2.10}
$$

and the five free parameters in this expression were fixed by fitting the GW phase to 80 simulations run by the HALOFEEDBACK code. The values of these best-fit parameters are $\alpha_1 = 1.4412$, $\alpha_2 = 0.4511$, $\sigma = 0.8163$ Hz, $\zeta = -0.4971$, and $\gamma_r = 1.4396$. For further discussion of the form of this ansatz for the fitting function, see [36].

To illustrate how the dark matter changes the GW phase, we show in Fig. 2 the dephasing ($\Phi^V - \Phi$) from a vacuum IMRI with the same parameters of the binary, though without the DM. The left panel reproduces a similar result in [36], and serves as a consistency check. The right panel shows how the phase differs as a function of different primary masses for a fixed secondary mass. The dephasing goes to zero at the inner-most stable circular orbit (ISCO) GW frequency, because the vacuum and DM GW phases were chosen to be equal at that value. Frequencies above that of the ISCOs are not shown, as the phase is valid just for the inspiral (and the plunge and merger would need to be modeled to compute the phase at these higher frequencies).

B. Gravitational-wave phase in the frequency domain

We next discuss the form of the GW strain in frequency domain. Because we will be working with the LISA noise curves that are averaged over the sky-position and polarization angles, we will also average the amplitude of the GW strain over these angles as well as the inclination angle and reference phase angle. This angle averaging removes the dependence of the strain $\tilde{h}$ on the antenna-response functions F_+ and $F_\times$. When we compute the noise-weighted inner products of two signals, the factors arising from angle averaging the signals will also cancel with equivalent factors in angle averaging the noise curve $S_n(f)$, as described in more detail in [50]. Thus, we will drop these factors from the sky-position and polarization average from
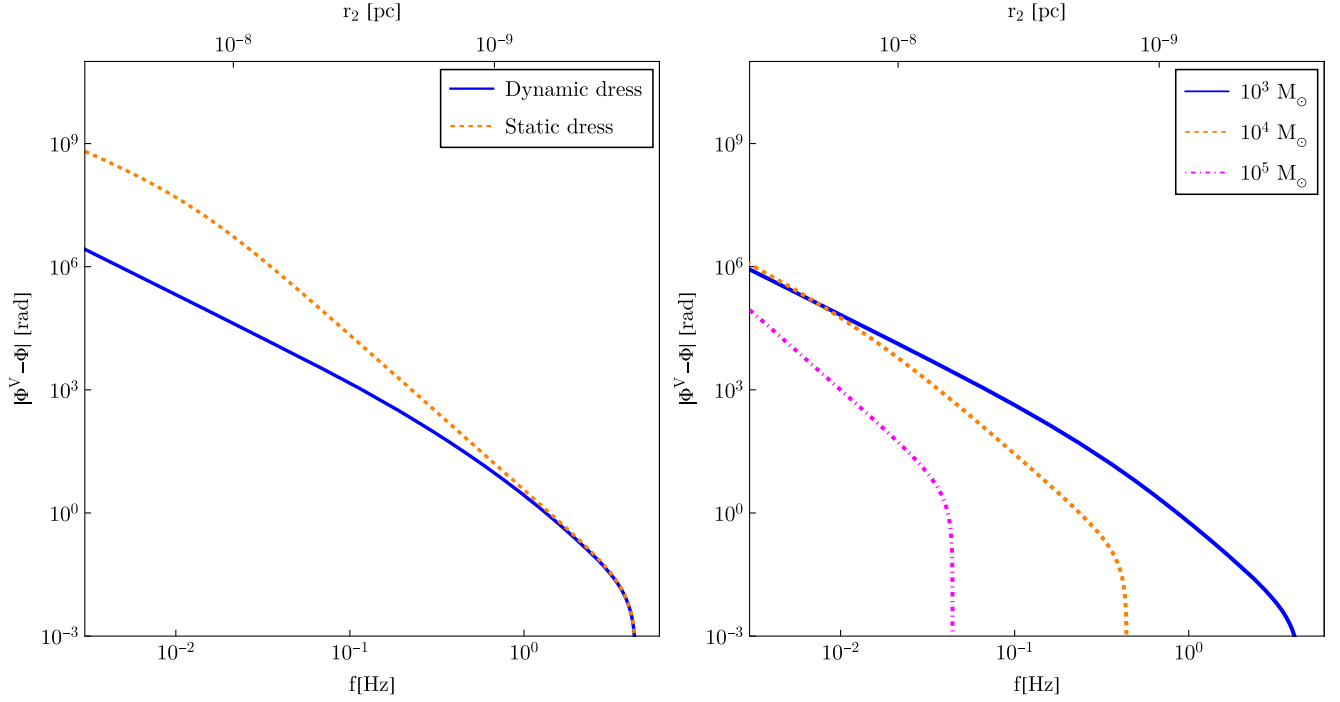
FIG. 2. Gravitational-wave dephasing for IMRIs with DM from those without. Left: a comparison of the GW dephasing for static and dynamic DM spikes versus frequency. The binary masses are given by $(m_1, m_2) = (10^3, 1.4)M_\odot$, and the DM parameters are those described in the introduction to Sec. II. The top horizontal axis shows the binary separation r_2 which is related to GW frequency f through the Kepler's law $\pi f = \sqrt{(m_1 + m_2)/r_2^3}$, for a circular orbit. In both static and dynamic cases, the dephasing resembles a broken power law, with breaks at different frequencies in each case. The inspiral proceeds from left to right and the dephasing goes to zero at the ISCO frequency, because the phases in the vacuum and nonvacuum cases are equated at this value. Right: similar to the left panel but only the dephasing for the dynamic DM case is shown for the three different values of m_1 given in the legend.

our waveform and denote this normalized, angle-averaged strain by just $\tilde{h}(f)$. It is common then to evaluate the Fourier transform in the stationary-phase approximation (SPA) and to write $\tilde{h}$ as

$$\tilde{h}(f) = A f^{-7/6} e^{i \Psi(f)}. \qquad (2.11)$$

Because the phase $\Phi(f)$ was computed using Newtonian physics, we use just the leading, quadrupole PN vacuum term for the amplitude. The expression for this normalized, angle-averaged amplitude is

$$A = \frac{\mathcal{M}^{5/6}}{\sqrt{6}\pi^{2/3} D_L}, \qquad (2.12)$$

where D_L is the luminosity distance to the source.[2]

---

[2]The numerical factor of $\sqrt{1/6}$ in A comes from the product of two square roots: $\sqrt{5/24}$ and $\sqrt{4/5}$. The factor of $\sqrt{5/24}$ is the square root that appears in the amplitude of the plus and cross polarizations of the strain [see, e.g., Eq. (20) of [50]]. The factor of $\sqrt{4/5}$ comes from the average over the inclination angle and reference phase [see, e.g., Eq. (16) of [50]]. We also do not include the DM-induced corrections to the amplitude (so that it has the same form as that in GR), because the amplitude corrections are not as significant as those in the phase (see, e.g., [39]).

Next, we turn to the frequency-domain SPA phase. The phase $\Phi(f)$ in Eq. (2.3) is the time-domain phase written in terms of frequency—i.e., $\Phi[t(f)]$. This differs from the SPA phase $\Psi(f)$ in frequency domain [51], but the two are related by

$$\Psi(f) = 2\pi f t(f) - \phi(f) - \frac{\pi}{4}. \qquad (2.13)$$

Here we have defined

$$\phi(f) = 2\pi \int^f f \frac{dt}{df} df = -\Phi[t(f)], \qquad (2.14)$$

which has a relative minus sign from the phase $\Phi(f)$ in Sec. II A.[3] To obtain Ψ, we need to compute $t(f)$. It can be obtained from differentiating and then integrating Eq. (2.14):

---

[3]The phases ϕ and Φ differ by a minus sign because the latter is defined by [35,36]

$$\Phi(f) = 2\pi \int_f f \frac{dt}{df} df. \qquad (2.15)$$

$$t(f) = \frac{1}{2\pi} \int^f \frac{1}{f} \frac{d\phi}{df} df. \tag{2.16}$$

From the time-domain phase in Eq. (2.3), we find the frequency-domain phase for the dynamic DM case is given by

$$\Psi = \Psi^V(f) - \frac{5\eta}{16(3\bar{\lambda}^2 + \bar{\lambda} - 2)}(\pi\mathcal{M}f_t)^{-5/3}y^{-\bar{\lambda}}$$
$$\times \left[{}_2F_1\left(1, \frac{3(\bar{\lambda}+1)}{5}; \frac{3\bar{\lambda}+8}{5}; -y^{-5/3}\right) \right.$$
$$\left. - \frac{3}{5}(\bar{\lambda}+1)y^{5/3}\log\left(y^{-5/3}+1\right) + \frac{3}{5}\bar{\lambda} - \frac{2}{5}\right]. \tag{2.17}$$

The vacuum phase, to leading PN order, is given by

$$\Psi^V(f) = 2\pi f t_c - \phi_c - \frac{\pi}{4} + \frac{3}{128}(\pi\mathcal{M}f)^{-5/3}, \tag{2.18}$$

and we have defined

$$\bar{\lambda} = \lambda + \frac{5}{3} = \frac{11 - 2\gamma_{\rm sp}}{3}. \tag{2.19}$$

The parameters t_c and ϕ_c are the time and phase of coalescence, respectively.

C. Parametrized post-Einsteinian framework

We next review the parametrized post-Einsteinian framework, which was originally developed for testing GR with GW observations with a theory-agnostic approach [37,52] (in the sense that the ppE framework was designed to make nonrestrictive assumptions about the gravitational-waveform model used to analyze a given GW event, and to permit the GW observation to determine if the ppE hypothesis or the GR hypothesis is favored). We will apply this framework in this paper to determine how well the ppE framework can reduce potential biases in parameter estimation and detect (or constrain) the DM effects in GWs from an IMRI immersed in an evolving DM spike.

The ppE template waveform for the $\ell = 2$, $m = 2$ harmonic that we will use is given by

$$\tilde{h}(f) = Af^{-7/6}e^{i(\Psi^V+\beta u^b)}, \tag{2.20}$$

where $u = (\pi\mathcal{M}f)^{1/3}$, β is the ppE parameter that controls the overall magnitude of the DM effects (in the context of this paper), and b is the exponent that is related to the "effective" PN order N of the correction term relative to Ψ^V by $b = 2N - 5$. As we will discuss in more detail below, the phase in Eq. (2.17) is not the vacuum GR phase plus a single power law in frequency as in Eq. (2.20), but there is a choice of the parameter b which will minimize errors, which will determine the effective PN order.
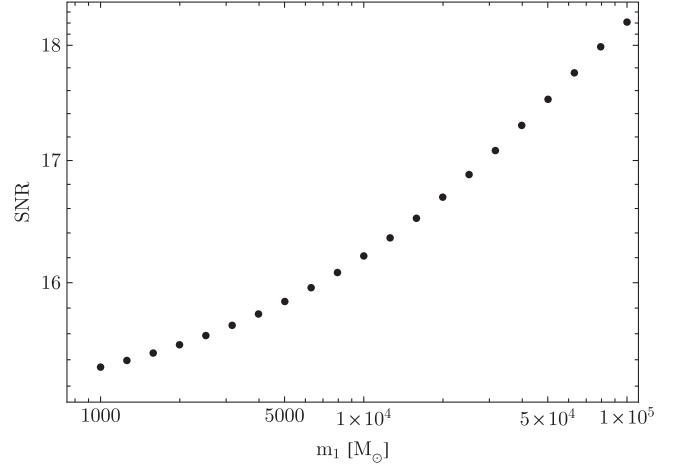

FIG. 3. SNR as a function of the primary mass. The mass of the secondary has been fixed to $m_2 = 1.4M_\odot$. Our choices of the luminosity distances that produce these SNRs are described in more detail in the text of Sec. III A.

III. DATA ANALYSIS METHODS AND MEASUREMENT ERRORS

In this section, we review the data-analysis methods that we use to compute the signal-to-noise ratio (SNR), the systematic errors and the statistical errors. We use matched filtering and Fisher-matrix methods [51,53] to compute the SNR and statistical errors, respectively. In the context of GW data analysis, the inverse Fisher matrix is most naturally interpreted as a Bayesian uncertainty of the posterior probability distribution for a single GW event under several assumptions [54]: for a high SNR signal with Gaussian noise, when the prior probabilities are constant over the relevant parameter space, the inverse Fisher matrix is the covariance matrix for the true source parameters.

As we show in more detail in Fig. 3, and describe in Sec. III A, the SNR of our signals will approximately 17, which is the SNR threshold for semicoherent searches derived in [55] (see also [56]). This reasonably high threshold, in addition to the well-constrained binary parameters in IMRIs and the reduced parameter set in our angle-averaged waveforms allows the assumptions underlying the Fisher-matrix formalism to be satisfied to a reasonable approximation. Because we are using a ppE waveform model rather than the exact dynamic DM phase, there will be biases (systematic errors) that arise from using this simpler waveform family. We estimate these systematic errors using the method described in [57], which we describe in more detail in Sec. III B.

A. Signal-to-noise ratio

To compute the SNR, we first define the noise-weighted inner product

$$(a|b) = 2\int_{f_{\rm low}}^{f_{\rm high}} \frac{\tilde{a}^*\tilde{b} + \tilde{a}\tilde{b}^*}{S_n(f)} df. \tag{3.1}$$

Here a and b represent time series, and $\tilde{a}$ and $\tilde{b}$ are their Fourier transforms. The noise spectral density, $S_n(f)$ will be taken to be the sky and polarization angle-averaged one, as given in [50].

We next discuss how we determine the lower and upper frequency limits of integration for the noise-weighted inner product in Eq. (3.1). To do so, we begin by noting that the binaries are at sufficiently large distances that we need to take into account the gravitational redshift z of the frequency between the source frame and the LISA detector frame; similarly, the source-frame chirp mass $\mathcal{M}$ must be scaled by $1 + z$ to convert to the detector-frame chirp mass $\mathcal{M}_z$ [i.e., $\mathcal{M}_z = (1 + z)\mathcal{M}$]. Because we will specify a luminosity distance D_L to the source, we need to choose a cosmology to obtain the redshift z (specifically, we will use the flat Planck 2015 cosmological parameters [58]).

We choose the lower frequency to be the frequency from which the IMRI will merge in a four-year period, assuming it inspirals adiabatically on a sequence of quasicircular orbits; We use the leading Newtonian-order expression for df/dt in vacuum to compute this value, and we specifically use the convenient expression given in [59], which uses an infinite frequency to define the merger:

$$f_{\text{low}} \approx 0.012 \text{ Hz} \left(\frac{\mathcal{M}_z}{50 M_\odot}\right)^{-\frac{5}{8}} \left(\frac{T_{\text{obs}}}{4 \text{ yr}}\right)^{-\frac{3}{8}}. \qquad (3.2)$$

We choose the upper frequency limit to be

$$f_{\text{high}} = \min(f_{\text{ISCO}}, 1 \text{ Hz}), \qquad (3.3)$$

where f_{ISCO} is the frequency at innermost stable circular orbit, which is also given in [59]:

$$f_{\text{ISCO}} = \frac{1}{6^{3/2}\pi M_z}. \qquad (3.4)$$

Here $M_z \approx m_{1,z}$ is the redshifted total mass of the binary (which is also approximately the shifted primary mass m_1). The factor of 1 Hz is our choice for the upper frequency cutoff for LISA.

The SNR is given in terms of the noise-weighted inner product by

$$\text{SNR} = \sqrt{(h|h)}. \qquad (3.5)$$

The values of the SNR for the IMRIs considered in this paper are shown in Fig. 3. In that figure, the primary mass m_1 is varied while the secondary mass is fixed to $m_2 = 1.4 M_\odot$. To choose the distance, we first computed each SNR with the corresponding source-frame masses and adjusted the luminosity distance to make the SNR equal to 15. With this luminosity distance, D_L for each binary, we then determined the corresponding redshift via the relationship

$$D_L = \frac{1 + z}{H_0} \int_0^z \frac{dz'}{\sqrt{\Omega_M(1 + z')^3 + \Omega_\Lambda}}. \qquad (3.6)$$

Here $H_0 = 67.8$ km/s/Mpc is the local Hubble constant, $\Omega_\Lambda = 0.7$ is the energy density of the cosmological constant (rounded to one significant figure) [58], and $\Omega_M = 1 - \Omega_\Lambda$ is the matter density. We then computed the SNR in Fig. 3 using the redshifted masses. The distances required to have these SNRs increase monotonically as a function of mass and are approximately 60 Mpc for the $10^3 M_\odot$ case and 820 Mpc for the $10^5 M_\odot$ case (rounded to the nearest tens digit).

B. Systematic and statistical errors

The systematic error on a given parameter θ^j is given by [57]

$$\Delta_{\text{sys}}\theta^j = (\Gamma^{-1})^{jk}(ih\Delta\Psi|\partial_k h), \qquad (3.7)$$

where $\Delta\Psi$ is the difference between the phase of the injected (true) and template waveforms.

Computing statistical errors makes use of the Fisher information matrix. It is computed from the noise-weighted inner product [Eq. (3.1)] of the derivatives of the waveform h with respect to each parameter [51,53]:

$$\Gamma_{jk} = \left(\frac{\partial h}{\partial \theta^j}\bigg| \frac{\partial h}{\partial \theta^k}\right). \qquad (3.8)$$

The inverse of the Fisher matrix is the covariance matrix (under the assumption noted above that the priors are constant over relevant parameter range). However, Gaussian priors centered around the maximum *a posteriori* values can be incorporated into the formalism. The standard deviations of these Gaussian priors will be denoted by $\sigma_{\theta_i}^{(0)}$, as in [59–61]. Assuming that there are no correlations among the prior parameters, one can then define an effective Fisher (inverse covariance matrix $\tilde{\Gamma}_{ij}$) by

$$\tilde{\Gamma}_{ij} = \Gamma_{ij} + \frac{1}{\left(\sigma_{\theta_i}^{(0)}\right)^2}\delta_{ij}, \qquad (3.9)$$

where δ_{ij} is the Kronecker delta. Thus $(\tilde{\Gamma}^{-1})_{ij}$ is a covariance matrix that takes into account the Gaussian priors. Our measure of the statistical error for a given parameter θ^i will be the square root of the diagonal elements of the effective Fisher matrix,

$$\Delta_{\text{stat}}\theta^i = \sqrt{(\tilde{\Gamma}^{-1})_{ii}}. \qquad (3.10)$$

The repeated index i is not summed over on the right-hand side. We only impose a prior on the phase of coalescence ϕ_c, which is periodic with period π. We approximate this in

this context by setting $\sigma_{\phi_c}^{(0)} = \pi$. The remaining parameters will have no prior imposed, which one can formally describe as setting $\sigma_{\theta_i}^{(0)} = \infty$ for these remaining parameters.

IV. OVERVIEW OF THE PPE TESTS

We give a brief overview of the different studies we perform to assess the performance of the ppE framework for capturing the dephasing effects induced by dark matter and the accuracy with which the vacuum and ppE parameters can be measured.

The first test that we perform examines whether a vacuum merger template is capable of accurately estimating the vacuum parameters for a GW waveform from an IMRI with a dynamic DM spike. The work in [36] examined this question, and determined that the inferred value of the chirp mass could be significantly biased, and the bias increased as the characteristic density of the spike became larger. The method used in [36] involved solving an optimization problem that became difficult to solve for large biases. Thus, it was unable to compute an estimate of the bias for the initial conditions of the DM spike that we treat in this paper. The method of computing the bias in [57], however, uses the true parameters of the IMRI waveform (which we assume are known), which allows us to obtain an estimate of the systematic error. For simplicity, we refer to this test as Question 1 (Q1).

The second test uses the ppE waveform templates under the assumption that the parameter β has zero as its true value, but can be nonzero to accommodate waveforms that deviate from vacuum GR ones. We will compute the statistical and systematic errors for the vacuum waveform parameters and β. A nonzero systematic error on β that is larger than the statistical error would be an indication that a nonzero ppE phase term would be needed to perform unbiased data analysis. In this test, we compute the systematic errors for a large number of fixed ppE exponent parameters b, and we choose the best one by minimizing the systematic error on the vacuum parameters (particularly, the chirp mass $\mathcal{M}_z$). An illustration of this procedure and further details about it are provided in Appendixes A and B; specifically, we also demonstrate the validity of this procedure by recovering the value of b in the case of a static spike, which is known analytically. Because we assume b is fixed in each individual computation of the systematic error on the vacuum parameters, we do not include b as a waveform-model parameter which has systematic or statistical uncertainties. This test will be called Q2.

Our third test involves computing the systematic and statistical errors after shifting the parameters by amounts that are equal to the systematic errors computed in Q2. We then evaluate the systematic and statistical errors with these updated "best" parameter values. We check to see if the statistical errors are larger or smaller than the systematic errors. If statistical errors dominate, then the ppE waveform

would not have significant bias; if the systematic ones do, then the ppE waveform would require further modification to robustly capture the GW effects of a dynamic DM spike. As in Q2, we determine b through the procedure described in Appendix B, and we assume that the optimized b is a fixed parameter with no uncertainties. This test is Q3.

Our fourth and final test is similar to Q3, but it allows the ppE exponent b to have uncertainty once it is fixed to the value used in Q3. In this case, we compute the systematic and statistical errors on all parameters in Q3 and also for b. The interpretation of this test, Q4, is the same as in the prior one (Q3).

For convenience, we summarize the main questions that the four tests are aimed to assess in the list below:
(Q1) How large is the bias on the vacuum parameters when using a vacuum template to analyze a GW signal with dynamic DM effects?
(Q2) Does adding a ppE phase term to the vacuum template reduce bias when the "fiducial" value of the ppE phase parameter β is assumed to vanish— and if so, by how much?
(Q3) How accurate is the ppE template when the fiducial parameter values are adjusted to account for the parameter biases, assuming the ppE exponent b has no uncertainty?
(Q4) How accurate is the ppE template at capturing DM effects in the gravitational waveform when the parameter b has uncertainty?

V. RESULTS OF THE PPE TESTS

We now present the results of the four tests that address the questions listed in Sec. IV. We assume a four-year observation period with LISA for all the results below.

A. Q1. Parameter bias from using vacuum waveforms

We begin by addressing Q1, which pertains to the amount of bias that arises from using vacuum GW templates to model IMRIs in a dynamical DM spike. In this case, the difference in the frequency-domain phase between the simulated signal and template is

$$\Delta\Psi = \Psi - \Psi^V, \tag{5.1}$$

where Ψ and Ψ^V are given in Eqs. (2.17) and (2.18), respectively. We choose the vacuum parameters for our study to be

$$\theta^i = (\ln \mathcal{M}_z, t_c, \phi_c, \ln A), \tag{5.2}$$

with A defined in Eq. (2.12). The true (injected) signal has the time and phase of coalescence set to be $t_c^{(\text{fid})} = \phi_c^{(\text{fid})} = 0$. For the chirp mass, we use $m_2^{(\text{fid})} = 1.4M_\odot$ in the source frame, and we vary the value of m_1 for the
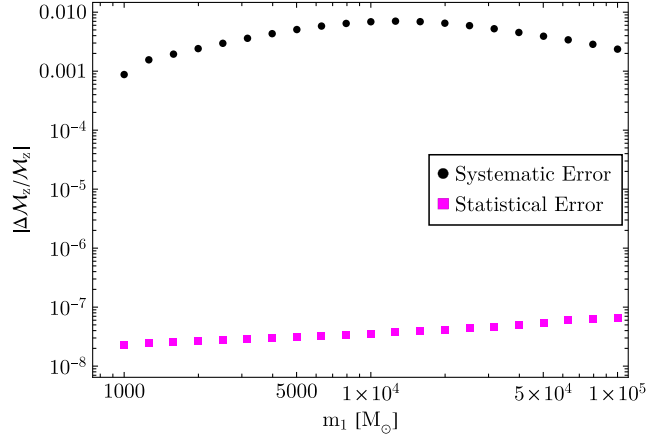
FIG. 4. (Q1) The relative systematic and statistical errors on the chirp mass against the primary BH mass. The errors were computed using a vacuum IMRI template assuming the true signal was an IMRI with a dynamic DM spike. The systematic error is significantly greater than the statistical error, which demonstrates the large bias that would arise if one were to attempt to use a vacuum template to analyze GWs from IMRIs in a DM spike.

primary black hole. The luminosity distances (and thus redshifts) are the same as those used to produce Fig. 3.

Figure 4 presents the fractional systematic and statistical error on the detector-frame chirp mass as a function of m_1. Note that the systematic error is consistently higher than the statistical error by at least four orders of magnitude, which is a significant parameter-estimation bias. The bias on the detector-frame chirp mass is the most pronounced bias, but the other parameters also have nontrivial biases, which we do not show.

The results above show that if a GW signal from an IMRI within a DM spike is detected through a matched-filtering search against vacuum templates, there will be large biases on the chirp mass. Given the results in Fig. 6 of [36], however, it will likely be challenging to detect such systems with vacuum templates. Thus, if a vacuum template is not capable of detecting the signal, then we would assume that the signal was detected by some other means (such as a coherent excess-power search) and was analyzed with vacuum templates.

Also note that we do not attempt to determine the source-frame chirp mass, as this would require measuring the luminosity distance (and thus the redshift). Because the statistical error of such measurements is not nearly as high as that for the chirp mass, propagation of errors would lead to a much larger statistical error on the source-frame chirp mass than that shown in Fig. 4.

B. Q2. Using ppE templates to reduce bias

Next, we address Q2, which relates to whether analyzing the event with a ppE template leads to less biased results than using vacuum templates (as in Q1). The phase difference $\Delta\Psi$ will now include a ppE term,

$$\Delta\Psi = \Psi - \Psi_V - \beta u^b. \tag{5.3}$$

When we compute the Fisher matrix and parameter bias, we now include one additional parameter, the ppE parameter β:

$$\theta^i = (\ln\mathcal{M}_z, t_c, \phi_c, \ln A, \beta). \tag{5.4}$$

The fiducial values for the vacuum parameters are the same as those in Q1 (Sec. VA) and the value of β is $\beta^{(\mathrm{fid})} = 0$. For the exponent b [or the effective PN order $N = (b+5)/2$], we determine the best fit value by minimizing the systematic errors on some of the vacuum parameters, as described in Appendix B. These minimum-error values of b depend upon the value of the mass m_1. For example, for $m_1 = 10^3 M_\odot$, we find that $N = -0.7121$ gives the minimum systematic error on the chirp mass.

The top panel of Fig. 5 shows the systematic and statistical error on the chirp mass with the ppE phase included in the templates. For all tested primary BH masses, the systematic errors are reduced by 8–9 orders of magnitude from the errors for the corresponding cases for Q1 (shown in Fig. 4). The statistical errors, however, have increased only by roughly an order of magnitude, which occurs because of the correlation between the chirp mass and ppE parameter β. Overall, the systematic error is consistently smaller than the statistical one by many orders of magnitude, which is an indication that the ppE template could, in principle, remove the bias in the vacuum parameter estimation for GWs from binary black holes in a DM halo. We only show the result for the chirp mass here, but we have checked that there are also similar decreases in the errors for the other vacuum parameters in the phase.

The bottom panel of Fig. 5 shows the systematic and statistical error for β. In this case, the systematic error is much larger than the statistical one; this occurs because the effects of DM on the GW phase are captured mainly by the ppE term in the ppE template, thereby producing a large deviation in the estimate for β from its fiducial value of zero (see Appendix A for further details). We also note that the statistical error on β for $m_1 = 10^3$–$10^4 M_\odot$ is roughly consistent with $\Delta\beta \sim 10^{-12}$ for intermediate-mass-ratio inspirals found in [62].

C. Q3. Updating the fiducial values in the ppE templates

The results in Q2 showed a large systematic error in β, because we used a fiducial value of $\beta = 0$. We can rerun our calculations by updating the fiducial values to

$$\theta^{i(\mathrm{fid,new})} = \theta^{i(\mathrm{fid,old})} + \Delta_{\mathrm{sys}}\theta^i, \tag{5.5}$$

where $\theta^{i(\mathrm{fid,old})}$ and $\Delta_{\mathrm{sys}}\theta^i$ are the fiducial values assumed in Q2 and the systematic errors are those determined in Q2. Gravitational waveforms with the above updated fiducial
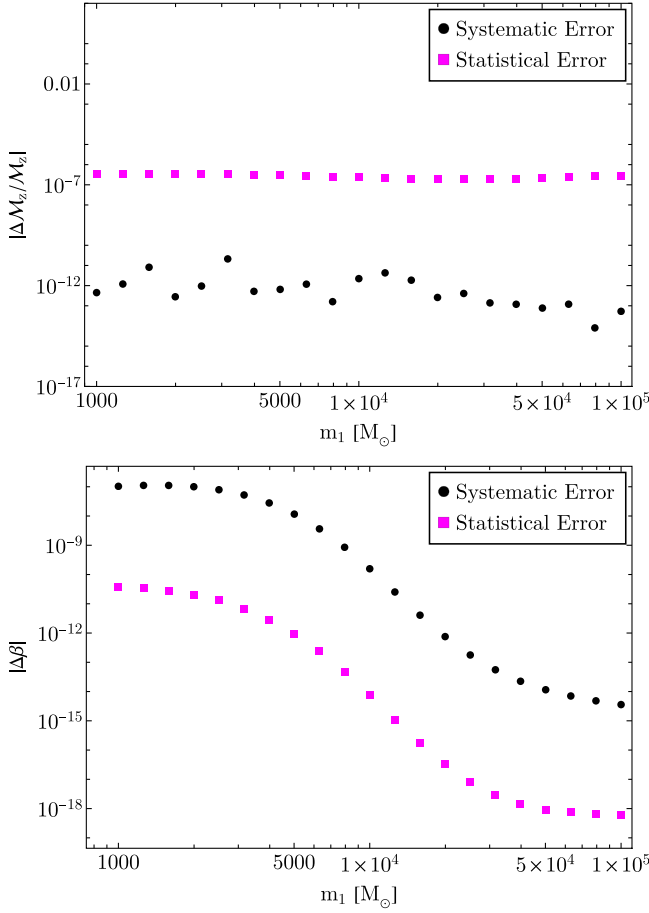
FIG. 5. (Q2) Errors with a ppE template. Top: errors on the detector-frame chirp mass, similar to those shown in Fig. 4, but now a ppE template is used and β is included in the parameter set. The systematic error has significantly decreased, which indicates that the ppE template can reduce the bias in the parameter estimation from an IMRI with a DM distribution. Bottom: the absolute systematic and statistical errors for the parameter β. The systematic error is several orders of magnitude larger; thus, the parameter estimation for β is significantly biased.
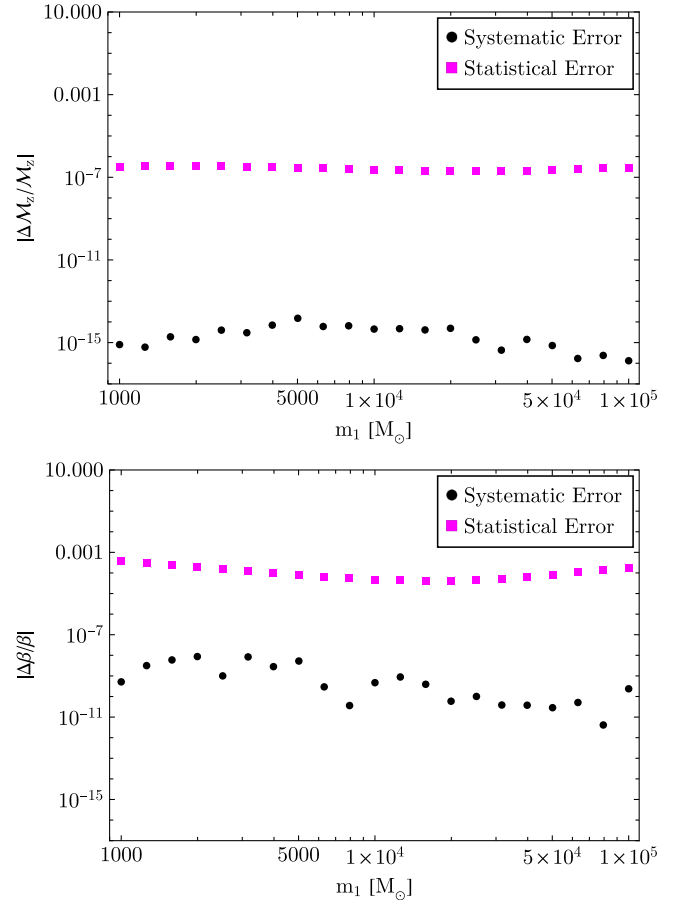

FIG. 6. (Q3) Errors with updated fiducial parameter values. Top: similar to the top panel in Fig. 5, but the fiducial values were updated as in Eq. (5.5). Bottom: similar to the top panel, but for the ppE parameter β. Note that unlike the bottom panel of Fig. 5, we plot the *fractional* error on β instead of the absolute error. In both the top and bottom panels, the systematic errors for both the chirp mass and ppE β are much smaller than their respective statistical errors.

values mimic more closely the "true" gravitational waveform from an IMRI with a DM halo. We keep the value of b to be the same as that used in Q2.

Figure 6 shows the errors on the chirp mass and ppE β with the updated fiducial values. For the chirp mass, the statistical error remains almost unaffected from the results in Q2, while the systematic error has been reduced by a few orders of magnitude. For ppE parameter, the fractional systematic error is $\mathcal{O}(10^{-10})$ in most cases; however, the fractional statistical error is $\mathcal{O}(10^{-4})$. The fractional statistical error is roughly consistent with the ratio of the absolute statistical and systematic errors in Q2. It is not precisely the same, however, because the partial derivative $\partial \tilde{h}/\partial \mathcal{M}_z$ in the inverse Fisher matrix now contains contributions from terms proportional to β, which is assumed to be nonzero. This will introduce additional correlations

between $\mathcal{M}_z$ and β, which would vanish when the fiducial value of β is set to be zero, as in Q2. Nevertheless, the systematic errors are now much smaller than the statistical errors for both parameters. This shows that, in principle, the ppE template is able to reduce the bias in vacuum parameters and to capture the effects of a dynamical DM distribution with a nonzero fiducial value of β.

D. Q4. Uncertainties in the ppE exponent b

In the previous two tests, we determined the optimal value of ppE exponent b that minimized the systematic error on the chirp mass, and we assumed that it was known without any uncertainty. This followed one viewpoint on the ppE framework, in which b is not considered a free-parameter of the framework; rather the ppE tests are run with fixed b for the values appropriate for a few common, discrete post-Newtonian orders at which different

deviations from GR could conceivably occur (based on specific examples of deviations that arise in theories of gravity that contain additional interactions not present in GR). However, the dark-matter effects can produce a dephasing that (at least locally in the frequency domain) would be best fit by a b parameter that is not constrained to be any particular PN order. The power law γ_{sp} in the initial DM density in Eq. (2.1) can take on a continuum of values in an interval on the real axis, and the best-fit b parameter in these cases would also be a continuum on another real interval, rather than at a few discrete values at a few distinct PN orders.

Thus, in the context of capturing dynamical DM effects, it is more natural to consider b to be a free parameter in the ppE framework, which would be determined simultaneously with the other parameters and would have systematic

and statistical uncertainties. This then leads to the final test, Q4, which addresses how the results in Q3 change if b is included in the parameter set:

$$\theta^i = (\ln \mathcal{M}_z, t_c, \phi_c, \ln A, \beta, b). \qquad (5.6)$$

We thus will fix the fiducial values for the vacuum parameters and the ppE parameter β to be the same as those in Q3. The fiducial values for b for different m_1 are the same as those used in Q2 and Q3.

Figure 7 summarizes the fractional statistical and systematic errors on the chirp mass, β, and b in Q4, in the right column. The results of Q1 in Fig. 4 are included in the left column and those of Q3 in Fig. 6 are shown in the middle column (so as to make a comparison of the three tests easier). By considering b as a waveform-model parameter,
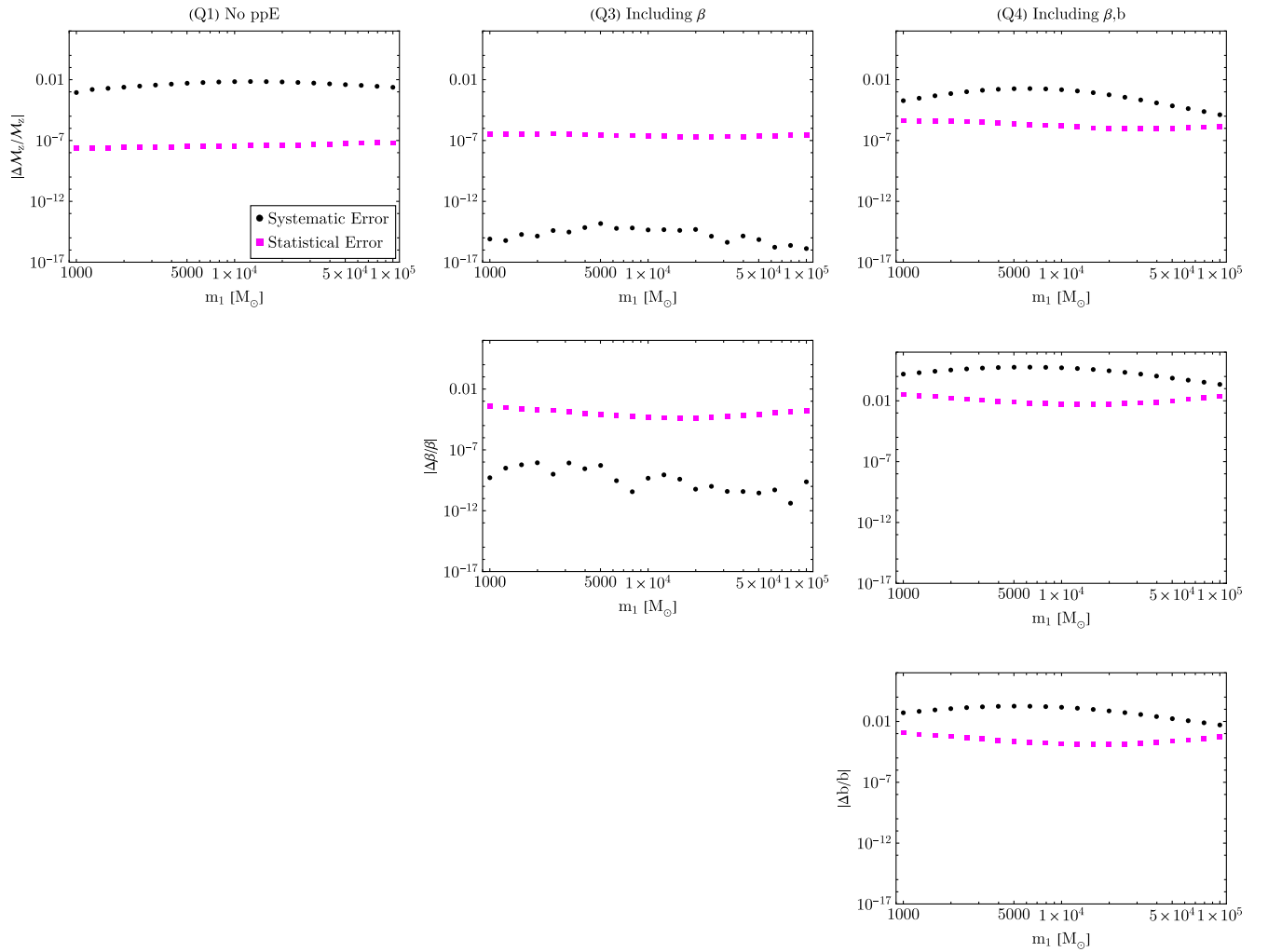

FIG. 7.   Comparison of the statistical and systematic errors in the different tests. The left column (Q1) and central column (Q3) reproduce the results in Figs. 4 and 6, respectively. They are given to make the comparison with the new results for Q4 in the right column easier. The first row allows for a comparison of the statistical and systematic errors on the chirp mass in tests Q1, Q3, and Q4. The second row gives a similar comparison for the ppE parameter β between Q3 and Q4. The third row shows the errors on the parameter b, which is computed only in Q4. Adding b as a free parameter with errors significantly increases the systematic error compared with other tests, and increases the statistical effor by around an order of magnitude.

the statistical errors on the chirp mass and β are about an order of magnitude higher than those in Q3, presumably because of increased correlations with the chirp mass and ppE parameter β. The systematic errors are significantly larger than the equivalent ones in Q3; thus $\Delta_{\mathrm{sys}}\mathcal{M}_z/\mathcal{M}_z$ and $\Delta_{\mathrm{sys}}\beta/\beta$ now exceed the statistical errors in Q4.

The reason for the increased systematic error can be understood from considering the linear equations in Eq. (3.7). For the five-parameter set θ^i in Q3, the value of b was optimized to make the systematic error (the product of the inverse Fisher matrix with the five-component vector of noise-weighted inner products) small. In Q4, however, the parameter set contains a sixth component, so the inverse Fisher matrix becomes a six-by-six matrix. It multiplies a vector with the same first five components as in Q3, but with a sixth component $(ih\Delta\Psi|\partial h/\partial b)$ added. There is not a simple relationship between the five-by-five block of the inverse Fisher matrix from Q4 (with the same parameters as those from Q3) and the inverse Fisher matrix in Q3. Thus, using the optimal values from Q3 in Q4 will not necessarily preserve the small systematic errors (and the results in Fig. 7 demonstrate that it does not). In principle, one should perform a new optimization (potentially over multiple parameters) to see if there exist other combinations of parameters that can reduce the systematic errors in Q4. However, we were unable to determine such parameters. Thus, we were not able to use this simplest ppE model to capture DM effects and obtain an unbiased estimate on the chirp mass and β parameter.

VI. CONCLUSIONS

In this paper, we investigated IMRI systems within an evolving DM spike, and we examined how well the ppE framework could model the GW effects of the DM in these systems. We first showed that using a vacuum template to model a merger within a dynamic DM spike causes the parameter estimation to be biased. Using vacuum templates to analyze systems with DM is most reasonable if the system is detectable with vacuum templates; however, Ref. [36] showed that the Bayes factor for the DM versus vacuum hypothesis strongly favored DM systems, thus indicating that vacuum templates would likely not be able to detect these systems for the systems that we considered. This study then supposed that the GW signal with DM could be detected by some other means.

We next performed additional tests within the context of the ppE framework. The ppE approach often makes the assumption that the true signal is that of vacuum GR, and it places constraints on a parameter β in the GW phase that permits a deviation from GR by computing its statistical error. Dark-matter effects in the GW phase, however, do not have a zero coefficient, so this approach would only be unbiased if the statistical error when using the ppE framework exceeded the systematic error from assuming vacuum GR is the true signal. After optimizing the exponent in the ppE phase, the statistical errors were several orders of magnitude smaller than the systematic ones; thus, a nonzero value of β is necessary to have unbiased results.

We then performed a similar study in which the fiducial value of the β ppE parameter was assumed to be nonzero. With an optimized ppE exponent b, the statistical errors were no larger than the systematic ones. This suggested that if the exponent was considered to be fixed (and not a parameter with its own errors) then the ppE framework could be used for parameter estimation without introducing bias.

The systematic errors on the vacuum parameters depend sensitively on the choice of the effective PN order; specifically, if the value is slightly off from the optimized one, the error goes up by several orders of magnitude. To run the ppE framework with a fixed exponent b, it would be necessary to perform the analysis for a large number of finely sampled b values. This may not be feasible, so it is more natural to treat the ppE parameter b as an undetermined model parameter that has its own systematic and statistical errors. Using the same optimized value from the previous test caused the systematic errors to exceed the statistical errors, and attempts at further optimization of the fiducial parameters did not significantly improve the results. We conclude that although the ppE template can, in principle, remove the bias in the parameter estimation for GWs from binary black holes with a DM halo, it may be difficult in practice to do so using the simplest ppE framework.

There are several different avenues for future work. One reason that the single-term ppE framework may not work well is that the GWs from IMRIs with evolving DM have deviations from vacuum waveforms that are not a single power law. Adding multiple ppE terms could produce a better fit to the phase, if they could be optimized to best match the phase from the IMRIs with DM. The case that we studied in this paper used just the leading-order Newtonain, quadurpolar waveform for an IMRI on a circular orbit, with a nonspinning primary BH surrounded by a DM distribution that evolves from the injection of energy from dynamical friction. Many of assumptions could be generalized. For example, if the secondary is a BH, then accretion onto the secondary introduces additional dephasing from systems that treat only dynamical friction. The GW effects enter at a different effective PN order from the dynamical-friction ones [46], and it would be interesting to investigate how well the ppE framework could account for both accretion and dynamical-friction feedback. IMRIs are strongly relativistic systems, so it would be useful to include higher PN terms and additional spherical-harmonic modes of the waveforms. Understanding how well the ppE framework would perform in this context would also be useful to investigate.

Finally, the intent of the original ppE framework is to test for theory-agnostic deviations from GR that arise at fixed PN orders, most frequently at half-integer values. Our results have shown that a nonzero ppE parameter β significantly reduces the systematic error in our Fisher-based parameter-estimation studies, which suggests that waveforms with nonzero β should be used to analyze the GW signals from such systems. In the context of the ppE framework, a nonzero β is associated with a deviation from GR, which might suggest that an environmental effect from a dynamic DM spike instead could be interpreted as a deviation from GR. However, the systematic error depended sensitively on the ppE exponent parameter b, which took on a continuum of values rather than those associated with a particular half-integer PN order. The fact the best value of b is not one of these discrete values is a likely indication that the origin of the nonzero β is an environmental effect rather than a beyond-GR effect. However, it may be the case that in practical ppE studies, the tests will be performed only a handful of half-integer PN orders. As a future study, it would also be interesting to investigate how well parameter estimation with ppE waveforms with these discrete ppE exponents b perform when the waveform of an IMRI with a dynamic DM spike is used as the injected signal. Such studies would determine whether or not the environmental effects arising from a dynamic DM spike could mimic those effects more commonly associated with gravitational theories beyond GR.

ACKNOWLEDGMENTS

DATA AVAILABILITY

The simulated data used to generate the figures in this paper could be made available upon request.

APPENDIX A: ANALYTIC CALCULATION OF THE SYSTEMATIC ERROR

When the phase difference from vacuum, $\Delta\Psi$, can be written in a ppE form, $\Delta\Psi = \alpha u^a$ (for some constants α and a), then the systematic error computed using the ppE templates has a simple form. Specifically, if we use the ppE template h in Eq. (2.20) to recover the true waveform, then the systematic error on each parameter θ^i becomes

$$\Delta_{\text{sys}}\theta^j = \alpha(\Gamma^{-1})^{jk}(iu^a h|\partial_k h). \qquad (\text{A1})$$

Next, we will choose the ppE exponent in the template to be $b = a$. This implies that $\partial_\beta \tilde{h} = iu^a h$, so the above systematic error becomes

$$\Delta\theta^i = \alpha(\Gamma^{-1})^{ij}(\partial_\beta h|\partial_j h). \qquad (\text{A2})$$

Using the fact that $(\partial_\beta h|\partial_j h) = \Gamma_{\beta j}$ are the β components of the Fisher matrix, we then have

$$\Delta\theta^i = \alpha(\Gamma^{-1})^{ij}\Gamma_{\beta j} = \alpha \mathbf{I}^{i\beta} = \begin{cases} \alpha & (i = \beta) \\ 0 & (i \neq \beta) \end{cases}, \qquad (\text{A3})$$

where $\mathbf{I}$ is the identity matrix.

The above analytic estimate shows that if the PN order (or the exponent) in the ppE term is chosen appropriately, then under the assumption made on $\Delta\Psi$, the systematic error appears in that of the ppE parameter β without affecting estimation of other parameters. This, in turn, can be used to determine the effective PN order of the DM effect. Namely, one can look at the systematic errors in the other parameters (e.g., $\Delta_{\text{sys}} \ln \mathcal{M}$) and identify the PN order that minimizes this systematic error. We illustrate this approach in the next appendix.

APPENDIX B: OPTIMAL PN ORDERS FOR DARK-MATTER-INDUCED DEPHASING

Here we illustrate how we use the results of Appendix A to identify the optimal PN-order N for the GW dephasing induced by DM, which minimizes the systematic error on the chirp mass. The order N is related to the ppE exponent by $b = 2N - 5$. We illustrate the procedure for both static and dynamic DM spikes.

1. Static DM case

We first illustrate the method described in Appendix A for a static DM distribution. For $f \gg f_{\text{eq}}$, with f_{eq} given in Eq. (2.5), the hypergeometric function in Eq. (2.3) can be expanded for small $1/y$, and the phase difference $\Delta\Psi$ has the following scaling with frequency:

$$\Delta\Psi \propto f^{-2(8-\gamma_{\text{sp}})/3}. \qquad (\text{B1})$$

For example, when $\gamma_{\text{sp}} = 7/3$, the above dephasing scales with f as $\Delta\Psi \propto f^{-34/9}$. Comparing this with the ppE term in the phase, we find that the ppE exponent b corresponds to $b = -34/3$, which means that the effective PN order should be $N = -19/6$.

The left panel of Fig. 8 shows the fractional error on the chirp mass as a function of the effective PN order N in the ppE term. The template used is the ppE waveform with the fiducial parameter of $\beta = 0$, and the masses are chosen to be $(m_1, m_2) = (10^3, 1.4)M_\odot$. Notice that the systematic error drops significantly around the PN order of $N = -19/6$ as estimated analytically above, which agrees with our analytic calculation in Appendix A. Notice also that the error is very sensitive to the choice of the PN order, in that it rises quickly if the PN order is slightly off from the optimal value. We found that we
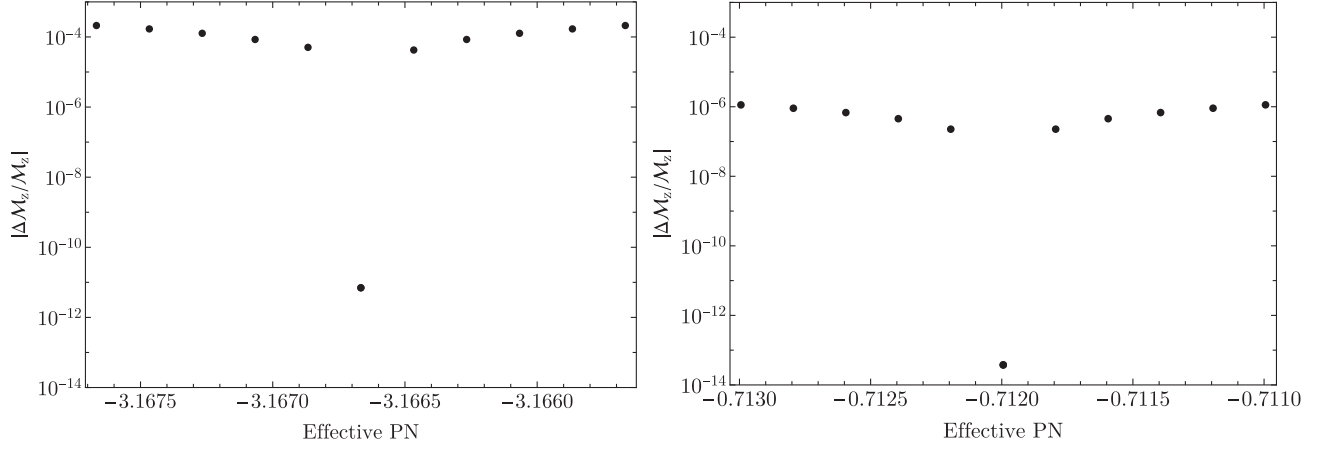
FIG. 8.   Optimizing over different PN orders. Left: the fractional systematic error on the chirp mass against the "effective" PN order $N = (b + 5)/2$ in the ppE term for a static DM distribution. Note the very sharp drop near the PN order of $N = -19/6 = -3.16667$ that can be computed analytically in this case when the DM effects are treated as a perturbation to those of gravitational radiation reaction. Right: a similar plot as that shown in the left panel, but the waveform phase now comes from the waveform for an IMRI in an evolving DM spike.

need a precision of at least five decimal places in N to find an accurate optimal value.

2. Dynamic DM case

We next examine at the dynamic DM case which has a more complicated frequency dependence in the phase than the static DM case. The right panel of Fig. 8 shows the fractional systematic error on the chirp mass for the dynamic DM case. Notice that the effective PN order has changed significantly from the static DM case and is now around $N = -0.7121$. Once again, the systematic error depends very sharply on the choice of the optimal PN order.

[1] G. Bertone and D. Hooper, History of dark matter, Rev. Mod. Phys. **90,** 045002 (2018).

[2] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. VI. Cosmological parameters, Astron. Astrophys. **641,** A6 (2020); **652,** C4(E) (2021).

[3] G. Bertone, D. Hooper, and J. Silk, Particle dark matter: Evidence, candidates and constraints, Phys. Rep. **405,** 279 (2005).

[4] J. L. Feng, Dark matter candidates from particle physics and methods of detection, Annu. Rev. Astron. Astrophys. **48,** 495 (2010).

[5] J. M. Gaskins, A review of indirect searches for particle dark matter, Contemp. Phys. **57,** 496 (2016).

[6] M. Schumann, Direct detection of WIMP dark matter: Concepts and status, J. Phys. G **46,** 103003 (2019).

[7] G. Bertone and T. M. P. Tait, A new era in the search for dark matter, Nature (London) **562,** 51 (2018).

[8] G. Bertone *et al.*, Gravitational wave probes of dark matter: Challenges and opportunities, SciPost Phys. Core **3,** 007 (2020).

[9] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Observation of gravitational waves from a binary black hole merger, Phys. Rev. Lett. **116,** 061102 (2016).

[10] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Tests of general relativity with GW150914, Phys. Rev. Lett. **116,** 221101 (2016); **121,** 129902(E) (2018).

[11] N. Yunes, K. Yagi, and F. Pretorius, Theoretical physics implications of the binary black-hole mergers GW150914 and GW151226, Phys. Rev. D **94,** 084002 (2016).

[12] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GWTC-1: A gravitational-wave transient catalog of compact binary mergers observed by LIGO and Virgo during the first and second observing runs, Phys. Rev. X **9,** 031040 (2019).

[13] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GWTC-2: Compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run, Phys. Rev. X **11,** 021053 (2021).

[14] R. Abbott *et al.* (KAGRA, Virgo, and LIGO Scientific Collaborations), 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).

[15] B. P. Abbott *et al.* (LIGO Scientific, Virgo, 1M2H, Dark Energy Camera GW-E, DES, DLT40, Las Cumbres Observatory, VINROUGE, and MASTER Collaborations),

A gravitational-wave standard siren measurement of the Hubble constant, Nature (London) **551,** 85 (2017).

[16] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), Constraints on the cosmic expansion history from GWTC–3, Astrophys. J. **949,** 76 (2023).

[17] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Tests of general relativity with GW170817, Phys. Rev. Lett. **123,** 011102 (2019).

[18] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Tests of general relativity with binary black holes from the second LIGO-Virgo gravitational-wave transient catalog, Phys. Rev. D **103,** 122002 (2021).

[19] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), Tests of general relativity with GWTC-3, arXiv:2112.06861.

[20] G. Agazie *et al.* (NANOGrav Collaboration), The NANO-Grav 15 yr data set: Evidence for a gravitational-wave background, Astrophys. J. Lett. **951,** L8 (2023).

[21] D. J. Reardon *et al.*, Search for an isotropic gravitational-wave background with the parkes pulsar timing array, Astrophys. J. Lett. **951,** L6 (2023).

[22] J. Antoniadis *et al.* (EPTA Collaboration and InPTA Collaboration), The second data release from the European pulsar timing array—III. Search for gravitational wave signals, Astron. Astrophys. **678,** A50 (2023).

[23] H. Xu *et al.*, Searching for the nano-hertz stochastic gravitational wave background with the Chinese pulsar timing array data release I, Res. Astron. Astrophys. **23,** 075024 (2023).

[24] P. Amaro-Seoane *et al.* (LISA Collaboration), Laser interferometer space antenna, arXiv:1702.00786.

[25] B. Sathyaprakash *et al.*, Scientific objectives of Einstein Telescope, Classical Quantum Gravity **29,** 124013 (2012); **30,** 079501(E) (2013).

[26] M. Maggiore *et al.*, Science case for the Einstein Telescope, J. Cosmol. Astropart. Phys. 03 (2020) 050.

[27] M. Evans *et al.*, A horizon study for Cosmic Explorer: Science, observatories, and community, arXiv:2109.09882.

[28] P. Amaro-Seoane, Relativistic dynamics and extreme mass ratio inspirals, Living Rev. Relativity **21,** 4 (2018).

[29] P. Gondolo and J. Silk, Dark matter annihilation at the Galactic Center, Phys. Rev. Lett. **83,** 1719 (1999).

[30] K. Eda, Y. Itoh, S. Kuroyanagi, and J. Silk, New probe of dark-matter properties: Gravitational waves from an intermediate-mass black hole embedded in a dark-matter minispike, Phys. Rev. Lett. **110,** 221101 (2013).

[31] K. Eda, Y. Itoh, S. Kuroyanagi, and J. Silk, Gravitational waves as a probe of dark matter minispikes, Phys. Rev. D **91,** 044045 (2015).

[32] S. Chandrasekhar, Dynamical friction. I. General considerations: The coefficient of dynamical friction, Astrophys. J. **97,** 255 (1943).

[33] S. Chandrasekhar, Dynamical friction. II. The rate of escape of stars from clusters and the evidence for the operation of dynamical friction, Astrophys. J. **97,** 263 (1943).

[34] S. Chandrasekhar, Dynamical friction. III. A more exact theory of the rate of escape of stars from clusters, Astrophys. J. **98,** 54 (1943).

[35] B. J. Kavanagh, D. A. Nichols, G. Bertone, and D. Gaggero, Detecting dark matter around black holes with gravitational

waves: Effects of dark-matter dynamics on the gravitational waveform, Phys. Rev. D **102,** 083006 (2020).

[36] A. Coogan, G. Bertone, D. Gaggero, B. J. Kavanagh, and D. A. Nichols, Measuring the dark matter environments of black hole binaries with gravitational waves, Phys. Rev. D **105,** 043009 (2022).

[37] N. Yunes and F. Pretorius, Fundamental theoretical bias in gravitational wave astrophysics and the parametrized post-Einsteinian framework, Phys. Rev. D **80,** 122003 (2009).

[38] K. Chatziioannou, N. Yunes, and N. Cornish, Model-independent test of general relativity: An extended post-Einsteinian framework with complete polarization content, Phys. Rev. D **86,** 022004 (2012); **95,** 129901(E) (2017).

[39] S. Tahura and K. Yagi, Parameterized post-Einsteinian gravitational waveforms in various modified theories of gravity, Phys. Rev. D **98,** 084042 (2018); **101,** 109902(E) (2020).

[40] N. Cornish, L. Sampson, N. Yunes, and F. Pretorius, Gravitational wave tests of general relativity with the parameterized post-Einsteinian framework, Phys. Rev. D **84,** 062003 (2011).

[41] L. Sampson, N. Cornish, and N. Yunes, Gravitational wave tests of strong field general relativity with binary inspirals: Realistic injections and optimal model selection, Phys. Rev. D **87,** 102001 (2013).

[42] L. Sampson, N. Cornish, and N. Yunes, Mismodeling in gravitational-wave astronomy: The trouble with templates, Phys. Rev. D **89,** 064037 (2014).

[43] V. Cardoso and A. Maselli, Constraints on the astrophysical environment of binaries with gravitational-wave observations, Astron. Astrophys. **644,** A147 (2020).

[44] G. Caneva Santoro, S. Roy, R. Vicente, M. Haney, O. J. Piccinni, W. Del Pozzo, and M. Martinez, First constraints on compact binary environments from LIGO-Virgo data, Phys. Rev. Lett. **132,** 251401 (2024).

[45] M. I. B. Rivera and R. C. Reyes, Measurable parameter combinations of environmentally-dephased EMRI gravitational-wave signals, New Astron. **112,** 102263 (2024).

[46] D. A. Nichols, B. A. Wade, and A. M. Grant, Secondary accretion of dark matter in intermediate mass-ratio inspirals: Dark-matter dynamics and gravitational-wave phase, Phys. Rev. D **108,** 124062 (2023).

[47] T. K. Karydas, B. J. Kavanagh, and G. Bertone, Sharpening the dark matter signature in gravitational waveforms I: Accretion and eccentricity evolution, arXiv:2402.13053.

[48] J. F. Navarro, C. S. Frenk, and S. D. M. White, The structure of cold dark matter halos, Astrophys. J. **462,** 563 (1996).

[49] Bradley J. Kavanagh, HALOFEEDBACK code, https://github .com/bradkav/HaloFeedback (2022).

[50] T. Robson, N. J. Cornish, and C. Liu, The construction and use of LISA sensitivity curves, Classical Quantum Gravity **36,** 105011 (2019).

[51] C. Cutler and E. E. Flanagan, Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral wave form?, Phys. Rev. D **49,** 2658 (1994).

[52] N. Yunes and X. Siemens, Gravitational-wave tests of general relativity with ground-based detectors and pulsar timing-arrays, Living Rev. Relativity **16,** 9 (2013).

[53] L. S. Finn, Detection, measurement and gravitational radiation, Phys. Rev. D **46**, 5236 (1992).

[54] 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).

[55] 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).

[56] 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).

[57] 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).

[58] P. A. R. Ade *et al.* (Planck Collaboration), Planck 2015 results. XIII. Cosmological parameters, Astron. Astrophys. **594**, A13 (2016).

[59] E. Berti, A. Buonanno, and C. M. Will, Estimating spinning binary parameters and testing alternative theories of gravity with LISA, Phys. Rev. D **71**, 084025 (2005).

[60] E. Poisson and C. M. Will, Gravitational waves from inspiraling compact binaries: Parameter estimation using second post-Newtonian wave forms, Phys. Rev. D **52**, 848 (1995).

[61] Z. Carson, Probing fundamental physics with gravitational waves, Ph.D. thesis, Virginia University, 2020, arXiv:2010.04745.

[62] K. Chamberlain and N. Yunes, Theoretical physics implications of gravitational wave observation with future detectors, Phys. Rev. D **96**, 084039 (2017).