Gravitational waves from black hole-neutron star binaries: Effective Fisher matrices and parameter estimation using higher harmonics

Hee-Suk Cho, Evan Ochsner, Richard O'Shaughnessy, Chunglee Kim, and Chang-Hwan Lee

Gravitational waves from BH-NS binaries: Effective Fisher matrices and parameter estimation using higher harmonics

Hee-Suk Cho[1],[*] Evan Ochsner[2],[†] Richard O'Shaughnessy[2],[‡] Chunglee Kim[3], and Chang-Hwan Lee[1]

[1]*Department of Physics, Pusan National University, Busan 609-735, Korea*
[2]*Center for Gravitation and Cosmology, University of Wisconsin-Milwaukee, Milwaukee, WI 53211, USA and*
[3]*Department of Physics, West Virginia University, PO Box 6315, Morgantown, WV 26505, USA*

Inspiralling black hole-neutron star (BH-NS) binaries emit a complicated gravitational wave signature, produced by multiple harmonics sourced by their strong local gravitational field and further modulated by the orbital plane's precession. Some features of this complex signal are easily accessible to ground-based interferometers (e.g., the rate of change of frequency); others less so (e.g., the polarization content); and others unavailable (e.g., features of the signal out of band). For this reason, an ambiguity function (a diagnostic of dissimilarity) between two such signals varies on many parameter scales and ranges. In this paper, we present a method for computing an approximate, *effective* Fisher matrix from variations in the ambiguity function on physically pertinent scales which depend on the relevant signal to noise ratio. As a concrete example, we explore how higher harmonics improve parameter measurement accuracy. As previous studies suggest, for our fiducial BH-NS binaries and for plausible signal amplitudes, we see that higher harmonics at best marginally improve our ability to measure parameters. For non-precessing binaries, these Fisher matrices *separate* into intrinsic (mass, spin) and extrinsic (geometrical) parameters; higher harmonics principally improve our knowledge about the line of sight. For the precessing binaries, the extra information provided by higher harmonics is distributed across several parameters. We provide concrete estimates for measurement accuracy, using coordinates adapted to the precession cone in the detector's sensitive band.

I. INTRODUCTION

Ground based gravitational wave detector networks (notably LIGO [1] and Virgo [2]) are analyzing results of design-sensitivity searches for the signals expected from the inspiral and merger of double compact binaries. [3, 4]. For the lowest-mass compact binaries $M = m_1 + m_2 \leq 16 M_\odot$, the response of the detector to a binary merger with arbitrary masses, spins, and even eccentricity is well understood, particularly given the detectors' limited and low sensitive frequency band [5–16]. Though this complicated signal encodes all information about the binary's spacetime [17], the amount of accessible information depends on signal strength (or signal-to-noise ratio) [18]. Strong signals permit high-precision tests of general relativity; fainter signals allow high-precision constraints on some binary parameters; while very faint, short signals may only constrain the binary's mass. Qualitatively speaking, we can distinguish two configurations if they are separated by contours $1 - P \gtrsim 1/\rho^2$, where P (defined in Sec. III) is the (normalized) ambiguity function or "overlap" and ρ is the signal-to-noise ratio (SNR).

As higher-order corrections and new physics are added to our models for gravitational-wave signals, the functional dependence on various parameters (such as masses, spins and orientation angles) in the model grows in com-

plexity. On scales $\simeq 1/\rho^2$, for astrophysically plausible ρ, the ambiguity function is generally smooth. However, it may have more complicated fine-scale structure which may not be detectable for expected signal strengths. The Fisher matrix approach to estimating parameter errors is based on differentiating a waveform with respect to its parameters. These derivatives are defined in an infinitesimal patch of parameter space, and are thus measuring fine-scale structure, which could potentially be misleading about larger-scale, observable trends. This point is illustrated by Fig. 2, where we fit a quadratic through the ambiguity function. The ambiguity function changes shape, so fits to small ($P > 0.999$) and large ($P > 0.99$) regions of parameter space would give rather different estimates of posterior widths and thus parameter accuracy. Similarly, if the standard Fisher matrix were calculated via finite difference, step sizes on these scales would give different results, with the *smaller* step size giving a misleading estimate of parameter accuracy for a signal of expected strength.

In this paper we propose a simple effective procedure to identify relevant scales and parameter correlations, construct suitable "effective Fisher matrices", and estimate ambiguity functions at low but nontrivial signal to noise ratio. To demonstrate this technique, we examine the signal from selected black hole-neutron star (BH-NS) binaries, described in Section II. In this paper we use all available knowledge about the (post Newtonian) waveform, adopting a complete model for the adiabatic quasicircular inspiral of precessing BH-NS binaries. In particular, we employ all available harmonics and amplitude

---

[*]Electronic address: `chohs1439@pusan.ac.kr`
[†]Electronic address: `evano@gravity.phys.uwm.edu`
[‡]Electronic address: `oshaughn@gravity.phys.uwm.edu`

corrections, introducing small but non-negligible changes to the ambiguity function. In Section III we introduce our unconventional effective approach to the local ambiguity function. Using those tools, in Section IV we construct and approximate the overlap for signals similar to each reference binary. Motivated by parameter estimation, we provide explicit expressions for the Fisher matrix, correlations, and marginalized uncertainties for each configuration. For our fiducial configurations, higher harmonics principally allow us to improve our knowledge of the binary *orientation*, providing fairly little additional information about intrinsic parameters for the amplitude scales of immediate astrophysical interest.

Our results are complicated by coordinate-dependent effects, notably extreme sensitivity to the reference frequency at which parameters are specified. We show the choice of reference frequency can reduce (or introduce) fine-scale structure into the ambiguity function, similar to the effect of higher harmonics. Our effective approach can partially compensate for ill-chosen coordinates, such as the coalescence phase or initial spins. To reduce but not completely eliminate these systematic effects, we express our results using parameters specified near the peak sensitivity of the detector (here, 100 Hz).

A. Context and prior work

Several studies of gravitational wave detection from merging binaries have employed amplitude-corrected waveforms and higher harmonics. Investigations of space-based interferometers, such as the Laser Interferometer Space Antenna (LISA), have historically used complete signal models, accounting for both spin and precession [19]. As higher harmonics have a small effect, however, most previous studies of ground-based interferometers have omitted them, emphasizing spin. When included, higher harmonics were explored alone for non-precessing signals. Higher harmonics can allow detection of signals otherwise inaccessible due to the detector's limited bandwidth [20, 21]. The relative amplitudes of higher harmonics can probe astrophysical mechanisms for generating non-circularity [22]. Finally, higher harmonics (and precession) are well-known to break degeneracies and improve sky localization, particularly for LISA [23–25].

Several authors have explored the local ambiguity function "beyond the Fisher matrix", including higher-order correlation functions [26, 27] and projection effects due to the local shape of the signal manifold [28]. These methods still use explicit derivatives of the ambiguity function (via explicit derivatives of the signal) to construct their series approximations.

The Fisher matrix is often nearly or exactly singular, making inversion numerically challenging. Several authors have pointed out that a singular value implies an unconstrained parameter, limited only by the prior; see, e.g., Vallisneri [26]. In many cases, including those sin-

gular values addressed in the text, the singular value corresponds to a *bounded* parameter (e.g., an angle). The singular value simply indicates that parameter cannot be measured. In the phenomenological limit described in this paper, precisely zero eigenvalues never occur, unless a parameter is constrained by symmetry.

Our goal in this work is to understand the typical shape of the posterior $p(\lambda|n, \lambda_0)$ for n a noise realization and λ, λ_0 coordinates in the signal space. Using one notion of "typical" would produce an *average* posterior over all noise realizations. Such an average posterior, however, could be slightly wider than the posterior from any given noise realization. Instead, in this work we attempt to characterize the typical shape of any *one* noise realization. To do so, in effect we "transport" each posterior so their peak likelihoods lie at the same point in parameter space. In practice, our procedure amounts to ignoring noise-realization-dependent changes to the posterior.

II. SIMULATIONS

A. Amplitude-corrected precessing waveform

In this paper we construct the post Newtonian (pN) gravitational wave signal from a BH-NS binary using the `lalsimulation SpinTaylorT4` code [29], which is an implementation based on the waveforms described in [5, 6]. This time-domain code solves for the orbital dynamics of an adiabatic, quasicircular inspiralling binary by using the so-called TaylorT4 method (see [16] for an explanation of this and similar methods) of evolving the orbital phase and frequency supplemented with precession equations to track the motion of the spins and orbital plane [30]. The orbital phase and frequency evolution includes non-spinning corrections to 3.5pN order and spin corrections to 2pN order. The precession equations are given to 2pN order. This binary evolution is terminated prior to merger, either when it reaches the "minimum energy circular orbit", or when the orbital frequency ceases to increase monotonically.

At each time, the values of the gravitational wave polarizations measured by a distant observer can be constructed from the orbital phase, orbital frequency and the orientations of the spins and orbital plane. We can construct either the commonly used "restricted" (i.e. leading-order) polarizations which contain only the dominant second harmonic of the orbital phase, or we can construct amplitude-corrected polarizations which contain terms that oscillate at other harmonics of the orbital phase (and also higher-order corrections to the second harmonic). Expressions for the polarizations valid for quasi-circular, precessing binaries are currently known to 1.5pN order [15, 30, 31]. Throughout this work, when we refer to amplitude-corrected waveforms, we mean that we use the 1.5pN accurate polarizations.
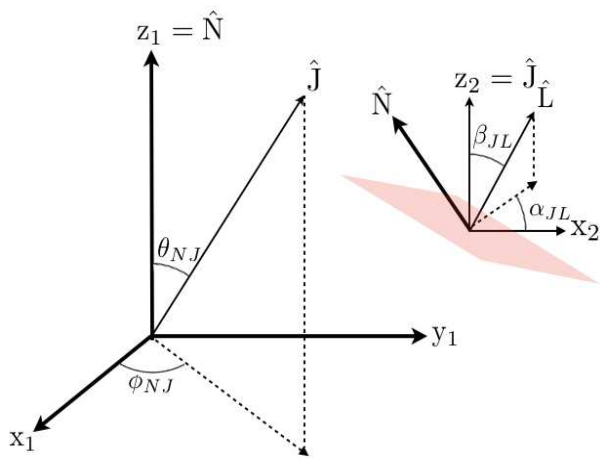
FIG. 1: **Coordinate system for the precessing binary**. The left coordinate corresponds to the conventional GW radiation frame. θ_{NJ} (ϕ_{NJ}) is a polar (azimuthal) angle of the total angular momentum (J) with respect to the radiation vector (N). In the left coordinate β_{JL} (α_{JL}) is a polar (azimuthal) angle of the orbital angular momentum (L) with respect to the total angular momentum (J). In the right coordinate, N, J, and x_2 are coplanar and the shaded region indicates the orbital plane.

B. Simulation coordinates

For the precessing binaries, LIGO-scale studies have been complicated by poor choice of coordinates, associated with the start of the waveform. The waveform generation code of the standard LIGO algorithm defines all the geometrical parameter values at the initial frequency (40 Hz for the initial LIGO and 10 Hz for the advanced LIGO), and evolve the binary system to get the full waveforms. Specifically, the orbit is described at some point, including the spin, orbital angular momentum vector, and the orbital phase. By contrast, the detector is more sensitive to higher frequencies. Allowing for the decreasing signal strength with frequency [32], the detector is most sensitive to the instantaneous binary configuration at 100 Hz (initial LIGO) and 40 Hz (advanced LIGO, e.g., see Fig. 2 in [32]). Motivated by this fact, we choose the reference frequency (f_{ref}), at which the instantaneous orientations of the spins and orbital plane are defined, to be 100 Hz.

One significant effect introduced by setting the reference frequency at 100 Hz is related to the orbital phase. The ambiguity function is dependent on how we choose the reference frequency. In appendix B, we describe an example and illustrate the significant effects in detail; see also Fig. 2.

Furthermore, following Brown et al.[32], we define the geometrical parameters to be angles between the radiation vector $\hat{N}$, total angular momentum axis $\hat{J}$ and orbital axis $\hat{L}$ as in Fig. 1. For comparison, other conventions specify some parameters using the line of sight, a vector $-\hat{N}$ pointing from our detector to the binary.

parameter	m_1	m_2	ι	ϕ	χ	M_c	η
non-spinning	10	1.4	$\pi/4$	0.0	0.0	2.994	0.1077
aligned-spin	10	1.4	$\pi/4$	0.0	1.0	2.994	0.1077

TABLE I: **Fiducial parameters for the non-spinning and aligned-spin binaries.** We adopt the chirp mass M_c and symmetric mass ratio η instead of individual masses. The orbital phase is defined at 100 Hz.

For non-spinning or aligned-spin binaries, the total angular momentum is parallel to the orbital angular momentum and the orbital axis is fixed. In effect, the conventional radiation frame is equivalent to the geometrical frame.

C. Fiducial simulations and local coordinates

In the case of a non-spinning binary, the binary is specified by 9 parameters. In this work, we choose masses (m_1, m_2) as intrinsic parameters, the polar (inclination) and azimuthal (polarization) angles (ι, ψ) of the orbital axis with respect to the radiation vector and an orbital phase ϕ as extrinsic parameters. Because we maximize the ambiguity function over the polarization, we need not take this parameter into explicit account henceforth. Remaining parameters are, the distance to the detector, sky position (two angles), and the coalescence time. The fiducial values of parameters are summarized in Table I. Mass components ($M = m_1 + m_2$) can be expressed by the symmetric mass ratio $\eta = m_1 m_2/M^2$ and chirp mass $M_c = M\eta^{3/5}$, we adopt these parameters in this work.

If the NS spin is assumed to be 0, the aligned-spin binary is specified by 10 parameters. 9 parameters are the same as the non-spinning case and the additional intrinsic parameter is dimensionless BH spin parameter χ. The fiducial values are also summarized in Table I.

The waveform of the precessing binary can be defined by 12 parameters if the NS spin is assumed to be 0. In this work, we consider η, M_c, BH spin χ, and the opening angle β_{JL} of the precessing cone as intrinsic parameters, α_{JL}, θ_{NJ}, ϕ_{NJ}, and the orbital phase ϕ as extrinsic parameters. Because we maximize over the polarization angle ψ, the parameter ϕ_{NJ} is eliminated from further consideration. Remaining parameters are the distance, sky position (two angles), the coalescence time. Throughout this paper the units are solar masses (for M_c); radians (for angles); or the natural dimensionless units (for η, χ).

Motivated by [32], we adopt a challenging reference configuration, where the polarization along the line of sight oscillates between circularly polarized (L along N) and linearly polarized (L perpendicular to N). Furthermore, to explore the extent to which higher-order harmonics allow measurement of parameters that only weakly impact the signal, we consider two possible sets of initial conditions for L along its precession cone. The fiducial values of the parameters are summarized in Table II. For case 1, the orbital axis is perpendicular to the

parameter	m_1	m_2	χ	β_{JL}	θ_{NJ}	α_{JL}	ϕ	configuration
case 1	10	1.4	1.0	$\pi/4$	$\pi/4$	0.0	0.0	$N \perp L$
case 2	10	1.4	1.0	$\pi/4$	$\pi/4$	π	0.0	$N \parallel L$

TABLE II: **Fiducial parameters for the precessing binary**: We adopt the chirp mass M_c and symmetric mass ratio η instead of individual masses. All the extrinsic parameters are defined at 100 Hz. For the case 1, L is perpendicular to N, and along for the case 2 at 100 Hz.

radiation vector at 100 Hz, for case 2 it is parallel to the radiation vector at 100 Hz. All the parameter values are the same between both cases except for α_{JL}.

D. Fiducial network

We assume two identical interferometers placed perpendicular to the incident signal, which is the optimal sky position of the source. We also assume the two interferometers are oriented by $\pi/4$ related to one another, giving comparable sensitivity to both polarizations. For the incident waveforms, we assume a zero noise limit to understand how similar the signals are. While not realistic, they avoid introducing complexity of the signal due to the source sky position.

III. DISTINGUISHING SIMULATIONS

A. Ambiguity function

In this work, we reorganize the two projections of the strain tensor $h_+ = e_+^{ab} h_{ab}/2$ and $h_\times = e_\times^{ab} h_{ab}/2$ into a complex function:

$$h(t) \equiv h_+(t) + i h_\times(t). \tag{1}$$

We coherently compare a fiducial signal $h_0(t, \lambda_0)$, where λ_0 indicates a fiducial source parameter set, to a nearby signal $h(t, \lambda)$, with parameters λ, by a complex overlap [33]

$$\langle h_0 | h \rangle \equiv 2 \int_{-\infty}^{\infty} \frac{df}{S_n(f)} [\tilde{h}_0(f) \tilde{h}(f)^*], \tag{2}$$

where $\tilde{h}(f)$ is the Fourier transform of $h(t)$ and $S_n(f)$ is a detector strain noise power spectrum. For simplicity, we adopt a semianalytic initial LIGO sensitivity [34, 35]. As pointed out by [33], this complex-valued expression characterizes the ability of a network to distinguish signals. The real part of the complex overlap corresponds to a linear sum of the conventional real overlaps of the two gravitational wave polarizations:

$$\mathrm{Re}\langle h_0 | h \rangle = (h_{0,+} | h_+) + (h_{0,\times} | h_\times), \tag{3}$$

where $(h_0 | h)$ indicates the conventional overlap of two real functions defined by

$$(h_0 | h) \equiv 4 \int_0^{\infty} \frac{df}{S_n(f)} \mathrm{Re}[\tilde{h}_0(f) \tilde{h}(f)^*]. \tag{4}$$

In appendix A, we summarize the differences between the real and complex overlaps.

We note that a change of the polarization angle, ψ, simple causes a rotation of the argument of the complex wave strain function, $h(\psi) = e^{-2i\psi} h(\psi = 0)$. Thus it is trivial to find the value of ψ which makes the complex overlap purely real, so that

$$\mathrm{Im}\langle h_0 | h' \rangle = 0, \tag{5}$$
$$\mathrm{Re}\langle h_0 | h' \rangle = (h_{0,+} | h'_+) + (h_{0,\times} | h'_\times),$$

and the value of the complex overlap maximized over polarization angle ψ is simply

$$\max_\psi \mathrm{Re}\langle h_0 | h \rangle = |\langle h_0 | h \rangle|. \tag{6}$$

The complex overlap (like the real-valued overlap) can also be maximized over the coalescence time t_c via an inverse Fourier transform as described in [36]. In particular, one uses the fact that

$$\tilde{h}(t_c = t) = \tilde{h}(t_c = 0) e^{-2\pi i f t} \tag{7}$$

and notes that the inverse Fourier transform of the complex overlap integrand in Eq. (2) will compute the complex overlap for all possible coalescence times of h at once

$$\langle h_0 | h(t_c = t) \rangle \equiv 2 \int_{-\infty}^{\infty} \frac{df}{S_n(f)} [\tilde{h}_0(f) \tilde{h}(f)^*] e^{2\pi i f t}. \tag{8}$$

The (normalized) ambiguity function $P(\lambda_0, \lambda)$ between two waveforms $h_0(t, \lambda_0)$ and $h(t, \lambda)$ is then defined as the complex overlap maximized over polarization angle and coalescence time,

$$P(\lambda_0, \lambda) \equiv \max_{t_c, \psi} \frac{|\langle h_0 | h \rangle|}{\sqrt{\langle h_0 | h_0 \rangle \langle h | h \rangle}}. \tag{9}$$

Unless otherwise noted, all overlaps are maximized in time and polarization. This is different from maximizing over orbital phase ϕ; see appendix A.

B. Likelihood

The detector noise $N(t)$ is assumed to be a stationary and Gaussian process. Given the detector output $S(t) = H(t, \lambda_0) + N(t)$ representing a real-valued signal in real-valued noise, the probability for the noise to have some realization N_0 is [37]

$$p(N = N_0) \propto e^{-(N_0 | N_0)/2}. \tag{10}$$

The posterior probability that the gravitational wave signal is characterized by the parameters λ, can be expressed by $p(\lambda | S) \propto p(\lambda) L(S | \lambda)$, where $p(\lambda)$ is the prior

probability that the signal is characterized by λ, $L(S|\lambda)$ is the likelihood, which can be written by [37]

$$L(S|\lambda) = C \times \exp\left[-\frac{1}{2}(S - H(\lambda)|S - H(\lambda)) \right], \quad (11)$$

where C is a proportional factor which, for simplicity, we assume to be 1 in this work.

Since we consider the complex strain, by choosing the appropriate polarization angle we shall write the detector output for the detector 1 and 2.

$$S_1 = H_+ + N_1, \quad S_2 = H_\times + N_2, \quad (12)$$

also

$$s = S_1 + iS_2, \quad h_0 = H_+ + iH_\times, \quad n_0 = N_1 + iN_2. (13)$$

The probability for the noise to have both realizations N_1 and N_2 is

$$\begin{aligned} p(N_1, N_2) &\propto e^{-\langle N_1|N_1\rangle/2} e^{-\langle N_2|N_2\rangle/2} \\ &= e^{-\mathrm{Re}\langle N_1 + iN_2|N_1 + iN_2\rangle/2}. \end{aligned} \quad (14)$$

Finally, using Eqs. (12 - 14), Eq. (11) can be expressed by the *complex* signals:

$$L(s|\lambda) = \exp\left[-\frac{1}{2}\mathrm{Re}\langle s - h(\lambda)|s - h(\lambda)\rangle \right]. \quad (15)$$

Substituting $s = h_0 + n_0$ into this equation [38], the likelihood is

$$\begin{aligned} L &= \exp\left[-\frac{1}{2}\mathrm{Re}\langle n_0 + h_0 - h|n_0 + h_0 - h\rangle \right] \quad (16) \\ &= \exp\left[-\frac{1}{2}\mathrm{Re}\{\langle h_0 - h|h_0 - h\rangle + 2\langle n_0|h_0 - h\rangle \right. \\ &\qquad \left. + \langle n_0|n_0\rangle\} \right], \end{aligned}$$

where the second and third terms in the a square bracket depend on the noise realization. They shift the position of the maximum likelihood but only weakly change the shape of the likelihood curve. In the limit of high SNR, these noise-dependent terms can be neglected, so,

$$\begin{aligned} L &= \exp\left[-\frac{1}{2}\mathrm{Re}\langle h_0 - h|h_0 - h\rangle \right] \quad (17) \\ &= \exp\left[-\frac{1}{2}\{\langle h_0|h_0\rangle + \langle h|h\rangle - 2\mathrm{Re}\langle h_0|h\rangle\} \right]. \end{aligned}$$

This equation corresponds to the case where two detectors are placed to have the maximum response to the incident two polarizations [for the detector placement, see Section II D]. While, Eq. (11) corresponds to one detector response to one polarization; see appendix A.

Using Eqs. (6) and (9), and the SNR defined by $\rho^2 = \langle h_0|h_0\rangle = \langle h|h\rangle$, the log likelihood can be expressed by our complex overlap convention:

$$\ln L = -\rho^2(1 - P) \quad (18)$$

where we assume the same strength for both signals[1].

For a given log likelihood, the scale of interest of the ambiguity function depends on the signal strength ρ^2:

$$1 - P \leq \frac{1}{\rho^2}. \quad (19)$$

By approximately identifying the $L \simeq 1/e$ surface of the likelihood, this condition allows us to estimate the set of parameters λ which cannot be distinguished from λ_0 with a signal amplitude of ρ using a signal model and noise curve that produces an overlap $P(\lambda_0, \lambda)$.[2]

C. Fisher matrix

If λ is close to λ_0, we can write $h_0 - h$ to the first order in the error $\Delta\lambda_i \equiv \lambda_0 - \lambda$

$$h_0 - h \sim \frac{\partial h}{\partial \lambda_i}\Delta\lambda_i. \quad (20)$$

So, in the limit of high SNR, the likelihood [Eq. (17)] is given as $L = \exp(-\Gamma_{ij}\Delta\lambda_i\Delta\lambda_j/2)$, where Γ_{ij} is

$$\Gamma_{ij} = \mathrm{Re}\left\langle \frac{\partial h}{\partial \lambda_i} \middle| \frac{\partial h}{\partial \lambda_j} \right\rangle. \quad (21)$$

This definition is analogous to the standard Fisher matrix, except that it is derived from the complex overlap and therefore contains information about both polarizations. If we assume that the prior $p(\lambda)$ is uniform, the parameter estimation errors $\Delta\lambda_i$ (i. e., the posterior probability density function) can be expressed by the Gaussian distribution

$$p(\Delta\lambda_i) = Ne^{-\Gamma_{ij}\Delta\lambda_i\Delta\lambda_j/2} \quad (22)$$

where $N = \sqrt{\det(\Gamma/2\pi)}$ is the corresponding normalization factor.

Using another expression relating the Fisher matrix to the log likelihood [26, 39] and Eq. (18), we can write

$$\begin{aligned} \Gamma_{ij} &= -\frac{\partial^2 \ln L(\lambda)}{\partial \lambda_i \partial \lambda_j}\bigg|_{\lambda_i = \lambda_j = \lambda_0} \quad (23) \\ &= \rho^2 \frac{\partial^2 (1 - P)}{\partial \lambda_i \partial \lambda_j}\bigg|_{\lambda_i = \lambda_j = \lambda_0} = \rho^2 \hat{\Gamma}_{ij}, \end{aligned}$$

---

[1] For different signal amplitudes, the log likelihood can be expressed in terms of ρ_0, ρ and P. For a suitable prior, we can analytically marginalize over the signal amplitude, producing a posterior of the form of Eq. (18). In this work, we focus attention on modeling P.

[2] More properly, the probability $F(L_0) = \int_{L > L_0} p(x)d\lambda$ of having likelihood $> L_o$ lets us create a confidence volume for any target confidence level. Because the probability $p \equiv F(L_o)$ depends sensitively on L, we anticipate the edge of this confidence interval will depend weakly (e.g., as $F^{-1}[p] \propto \sqrt{\ln p}$) on the precise probability used to define the threshold.

where λ_0 is the fiducial value of source parameter. Here, we define the normalized Fisher matrix $\hat{\Gamma}_{ij}$.

For a Gaussian noise and high SNR, the inverse of the Fisher matrix is the covariance matrix (Σ_{ij}) of parameter errors. The measurement error (σ_i) of each parameter and correlation coefficient (c_{ij}) between two parameters are defined as

$$\sigma_i = \sqrt{\Sigma_{ii}}, \quad c_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}, \tag{24}$$

The correlation coefficients c_{ij} are ρ-independent but often sensitive to small changes in $\hat{\Gamma}$. Conversely, the measurement error is inversely proportional to ρ. For the purposes of illustration, we adopt $\rho = 10$ whenever we calculate σ_i.

D. Relevant scales and effective approach

The Fisher matrix formally involves derivatives, i.e., infinitesimal variations of a parameter $d\lambda$. In this work, we compute an effective Fisher matrix by considering finite variations $\delta\lambda$ on scales which give physically observable changes to the ambiguity function, P.

To understand the variability on multiple scales we plot a one-dimensional ambiguity function of M_c for the leading-order amplitude, non-spinning binary in Fig. 2. In this figure, the ambiguity function P is calculated via Eq. (9), changing only M_c and fixing all other parameters to be the same for both signals. For comparison, we plot quadratic fits[3] to the ambiguity curve at different scales of $P > 0.99$ and $P > 0.999$. These are the scales of interest for signals with strength $\rho^2 \sim 10^2$ or 10^3 (see, Eq. (19)). [Note that except for the upper panel of Fig. 2, we choose the reference frequency at 100 Hz. The structure illustrated here is present but less significant at 100 Hz; see appendix B.]

The shape of the ambiguity function has structure on multiple scales. The neighborhood of $P > 0.999$ suggests a much sharper peak than what is seen at the $P > 0.99$ scale. A Fisher matrix computed from formal waveform parameter derivatives (defined in the limit $d\lambda \to 0$) or finite difference such that $P(\lambda_0, \lambda_0 + \delta\lambda) \gtrsim 0.999$ can be overly optimistic about how well λ can be measured for a signal with $\rho^2 \sim 10^2$.

Therefore, we wish to define an effective Fisher matrix from the curvature of the ambiguity function on the scales of interest. For example, in the case of two parameters, the fitting function P^* is

$$P^* = P_{\max} + p_1\delta\lambda_1^2 + p_2\delta\lambda_2^2 + p_{12}\delta\lambda_1\delta\lambda_2 , \tag{25}$$

---

[3] Since the posterior function is a normal distribution in the limit of high SNR (see Eq. (22)), a quadratic fitting function usually best fits the log likelihood function with a flat prior.
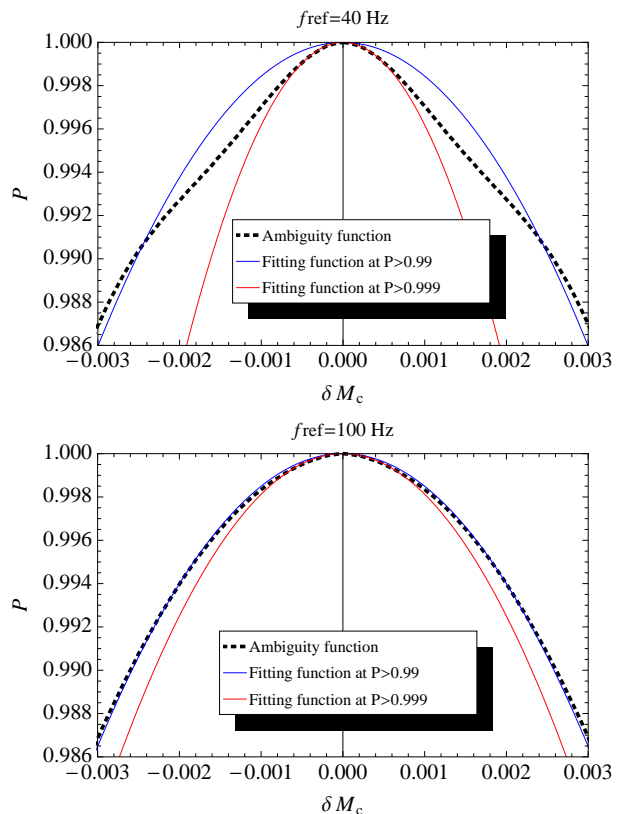

FIG. 2: **One dimensional ambiguity function for the non-spinning binary, showing variability on multiple scales**. Dotted line shows P calculated from Eq. (9) for the leading-order waveforms, where we compare our fiducial simulation with slightly offset analogs, changing just the chirp mass. For comparison, the two solid curves show quadratic fits to the ambiguity function, on scales of $P > 0.99$ (blue) and $P > 0.999$ (red). Note that the upper panel is computed with the reference frequency at 40 Hz, a default choice that accentuates this scale dependence. In the lower panel, we adopt a reference frequency at 100 Hz to mitigate the change in scale; see appendix B. In this paper we adopt a reference frequency at 100 Hz.

where p_i and p_{ij} are fitting coefficients and $P_{\max} = 1$. We calculate the effective Fisher matrix as

$$(\hat{\Gamma}_{ij})_{\text{eff}} = -\frac{\partial^2 P^*}{\partial\lambda_i\partial\lambda_j} . \tag{26}$$

In some cases, especially when a parameter is poorly determined, the variation of the ambiguity function with a parameter may not be well-described by a quadratic. See, for example, the top panel of Fig. 5. Therefore, we find it useful to employ an "iterative" procedure to find the parameters that are amenable to a quadratic fit. For each parameter x, we compute the one-dimensional curve $P(x, x + \delta x)$ and fit it against $1 - (\hat{\Gamma}_{xx})_{\text{eff}}\delta x^2/2$. This determines the diagonal elements of the effective Fisher matrix, $(\hat{\Gamma}_{xx})_{\text{eff}}$. We discard any parameters that are poorly fit by this quadratic (checked either "by eye"

or a with quantitative threshold on the goodness of fit). For the well-fit parameters, we determine the off-diagonal elements of the Fisher matrix by computing the two-dimensional surface $P(x, y; x+\delta x, y+\delta y)$ and fitting it to a function of the form like Eq. (25) while using the values from the one-dimensional fits, i.e. $p_x = -(\hat{\Gamma}_{xx})_{\rm eff}/2$. We note this method is used primarily as a sanity check to identify any parameters which induce obviously non-quadratic variations in P.

Once we have identified the space of all reasonably quadratic parameters, the ambiguity function on that space can be approximated as

$$P^*(\lambda_0, \lambda_0 + \delta\lambda) = 1 - (\hat{\Gamma}_{ij})_{\rm eff} \ \delta\lambda_i \ \delta\lambda_j/2 \ . \qquad (27)$$

Rather than using the iterative procedure described above to find the elements of $(\hat{\Gamma}_{ij})_{\rm eff}$ one at a time, it is straightforward to use a standard least-squares fitting technique to simultaneously solve for all of the $(\hat{\Gamma}_{ij})_{\rm eff}$. This will also give a better global approximation to the ambiguity function than the iterative approach. As an example of the small yet noticeable differences between these two procedures, Table VI compares the results for computing the effective Fisher matrix from both "iterative" and "simultaneous" fits to the ambiguity function. Everywhere else in this work (Tables III, IV, V, VI, VII, and IX) the effective Fisher matrix is computed by simultaneously fitting all coefficients.

In the cases where P is not well-described by a quadratic (e.g., see Fig. 5), we can adopt more complicated expressions to characterize the functional dependence of P when these parameters are varied, both in isolation and in correlation with well-constrained variables. As a concrete example, in the absence of higher harmonics the line of sight from the binary is both *weakly constrained* by observations and nearly *separable* in $P(\lambda_0, \lambda)$ from other degrees of freedom[4]. Specifically, ignoring maximization in time and phase, the overlap $P(\hat{N}, \hat{N}')$ between any two lines of sight can be well-approximated by Eq. (B1c) from [33]:

$$P_{\rm angles} \simeq \frac{|e^{2i\phi}Y_2(\theta)Y_2(\theta') + e^{-2i\phi}Y_{-2}(\theta)Y_{-2}(\theta')|}{\sqrt{(|Y_2(\theta)|^2 + |Y_{-2}(\theta)|^2)(|Y_2(\theta')|^2 + |Y_{-2}(\theta')|^2)}} \qquad (28)$$

where we use the shorthand $Y_m \equiv Y_{2m}^{(-2)}(\hat{n})$ and similarly for Y_m' to reduce superfluous subscripts and where we factor out the common $e^{im\phi}$ from $Y_{lm}^{(-2)}$. This function has wide, nearly flat extrema in $\hat{n}$ for each fixed $\hat{n}'$. On the other hand, in the absence of higher harmonics the line of sight has little impact on the waveform phase versus time away from the orbital plane. We can therefore approximate the ambiguity function for $P > 0.99$ in the top panel of Fig. 5 by

$$P \simeq P_{\rm angles}[1 - \frac{1}{2}(\hat{\Gamma}_{ab})_{\rm eff}\delta\lambda_a\delta\lambda_b] - \Gamma_{aN}\delta\lambda_a\delta\lambda_N \ , \quad (29)$$

where the N index varies over the line-of-sight parameters (θ, ϕ) and the a, b indices vary over the other parameters and $\Gamma_{aN} \simeq 0$ for $a \notin N$. With higher harmonics, the functional form above [Eq. (29)] is weakly perturbed by additional angular terms of the form

$$\begin{aligned} P(\lambda_0, \lambda) \simeq & P_{\rm angles}[1 - \frac{1}{2}(\hat{\Gamma}_{ab})_{\rm eff}\delta\lambda_a\delta\lambda_b] - \Gamma_{aN}\delta\lambda_a\delta\lambda_N \\ & - \frac{1}{2}(1 - \cos\iota)G_{\phi\phi}(\phi - \phi_0)^2 \qquad (30) \\ & - \frac{1}{2}G_{cc}(\cos\iota - \cos\iota_0)^2 \\ & - G_{\phi c}(\phi - \phi_0)(\cos\iota - \cos\iota_0) \end{aligned}$$

where G_{ab} is a matrix with $G_{\phi,c} \simeq 0$. This approximation both factors out the leading-order angular dependence and adds additional angular terms with parameter-dependent coefficients, designed to correctly reproduce a ϕ-independent result when $\cos\iota \simeq 0$. Although these terms allow us to correctly reproduce the non-ellipsoidal contours seen in the bottom panel of Fig. 5, Tables VI and VII show that this complicated structure only marginally improves the overall fit compared to a purely quadratic approximation $P \simeq 1 - (\hat{\Gamma}_{ab})_{\rm eff}\delta\lambda_a\delta\lambda_b/2$. Fit parameters for this more complicated functional dependence are not presented here.

We also considered a more general fit, treating $P_{\rm max}$ as a parameter. While this parameterization has a significant aesthetic advantage – its effective Fisher matrix is roughly scale-independent when $f_{\rm ref} \simeq 100$ Hz and agrees with analytic calculations – it systematically underestimates P in the neighborhood of the maximum. As both fits work well globally, we favor the simpler procedure and adopt $P_{\rm max} = 1$ except for Table V.

E. Comparing to standard Fisher matrix results

Despite subtle differences associated with time domain versus frequency domain waveforms, the complex overlap, higher harmonics, and the line of sight, our results for the effective Fisher matrix are directly comparable to earlier results calculated with the stationary phase approximation [40]. For example, for emission along the $\hat{z}$ axis, both the real and complex strain have the form[5] $\tilde{h} = Ae^{-2i\Phi}$ for $f > 0$. As a result, the Fisher matrix $\hat{\Gamma}_{ij}$ can be well-approximated by an *identical* average over

---

[4] Approximate separability of the line of sight from other degrees of freedom follows only in our well-chosen coordinates, where the binary configuration is specified at 100 Hz.

[5] The two differ for $f < 0$: the complex strain has $\tilde{h}(f < 0, \hat{z}) = 0$, while the real strain has $\tilde{h}(-f) = \tilde{h}(f)^*$.

frequency:

$$\hat{\Gamma}_{ij} = \frac{\int_{-\infty}^{\infty} df |\tilde{h}|^2 (\partial_a \Psi)(\partial_b \Psi)/S_h}{\int_{-\infty}^{\infty} df |\tilde{h}|^2 / S_h} \tag{31}$$

where we neglect derivatives $\partial_a A$ as small compared to the leading-order phase dependence. Thus, *each component* of our Fisher matrix must resemble previous results. In fact, as the general definition [Eq. (21)] suggests, each component of Γ_{ab} (unlike $\Sigma = \Gamma^{-1}$) depends only on the local response to changing *two* parameters λ_a, λ_b, no matter how many parameters exist in the model. Therefore, the Fisher matrix for an identical model with more parameters will have, as a submatrix, the Fisher matrix for the smaller model. By contrast, other methods for expressing uncertainty like the covariance Σ depend simultaneously on *all* terms in Γ. For low-mass binaries, the Fisher matrix Γ is well-known to be poorly conditioned, with eigenvalues spanning several orders of magnitude. We therefore preferentially compare the *component-by-component Fisher matrix*, rather than the covariances Σ, when comparing results. When presenting results, we provide several significant figures to insure all eigenvalues of Γ remain positive-definite.

Our complex overlap maximizes over time and polarization. The analytic Fisher matrix calculated from the stationary phase approximation [Eq. (31)] has time and phase as parameters. To account for maximizing over those parameters, we transform the full Fisher matrix Γ to a smaller-dimensional matrix which projects out those dimensions:

$$(\hat{\Gamma}_{ab})_{\max} = \hat{\Gamma}_{ab} - \hat{\Gamma}_{aC} Q_{CD} \hat{\Gamma}_{Cb} \tag{32}$$

$$Q_{CD} \equiv [\hat{\Gamma}_{CD}]^{-1} \tag{33}$$

where C, D run over the t, ϕ variables and a, b all other variables. In these expressions, the matrix Q is the inverse of the projection of $\hat{\Gamma}_{ab}$ into the t, ϕ subspace.

F. Comparing to posteriors

Standard parameter estimation techniques like Markov-Chain Monte Carlo produce samples of the full posterior probability distributions, including post-processed data products like one-dimensional standard deviations σ_a and two-dimensional covariances c_{ij}. For *strong signals* with well-isolated probability distributions, our one-dimensional standard deviations and covariances are directly comparable, for identical binaries. For fainter signals with broad probability distributions, our results will describe part of the posterior, in the neighborhood of one extremum.

For brevity, by maximizing over the event time and polarization, we have explicitly eliminated the two parameters and make no predictions about any correlation
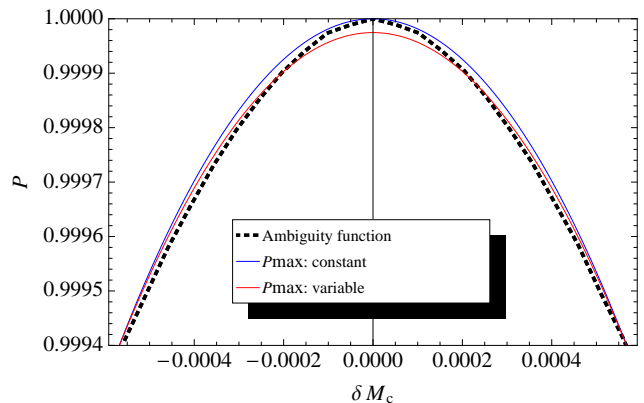

FIG. 3: **Two methods for fitting the ambiguity function**: Illustration of differences between our default fitting technique (blue curve), which fixes $P_{\max} = 1$, and an alternate fitting technique that lets $P_{\max} \neq 1$, shown here for the $P > 0.999$ scale (dotted curve). These two methods produce comparable global fit *functions*, but the fitting *parameters* for Γ differ by tens of percent; see, e.g., Table V. While numerical error can produce fluctuations in the overlap (e.g., due to insufficient sampling rate; we use 65536 Hz), the change in shape shown here is resolved.

including them[6]. We have also neglected other physical effects such as the fact that a real detector network will have an anisotropic response to the plus and cross polarizations as the sky position and orientation are varied. In reality, time of arrival, source orientation (polarization angle), sky position and distance are correlated and should also be addressed by the effective Fisher matrix, or direct ambiguity contour method. However we have entirely circumvented any discussion about these parameters by adopting an idealized network model and fixed SNRs. We will revisit these parameters and describe how to interpolate between our idealized model and a real anisotropic network in a subsequent publication, which compares MCMC results to the effective Fisher matrix predictions.

G. Numerical and systematic effects

At the very smallest scales, delicate implementation-dependent choices can also introduce artificial structure into the ambiguity function. We have already extensively described how the choice of reference frequency introduces (coordinate-dependent) structure. Less physically, the sampling rate for the waveforms can produce artifi-

---

[6] For a Gaussian posterior, the process of marginalizing a likelihood over a parameter is formally equivalent to maximizing the likelihood in that parameters, up to an irrelevant normalization constant. This approximation breaks down when the likelihood is not well approximated by a Gaussian.
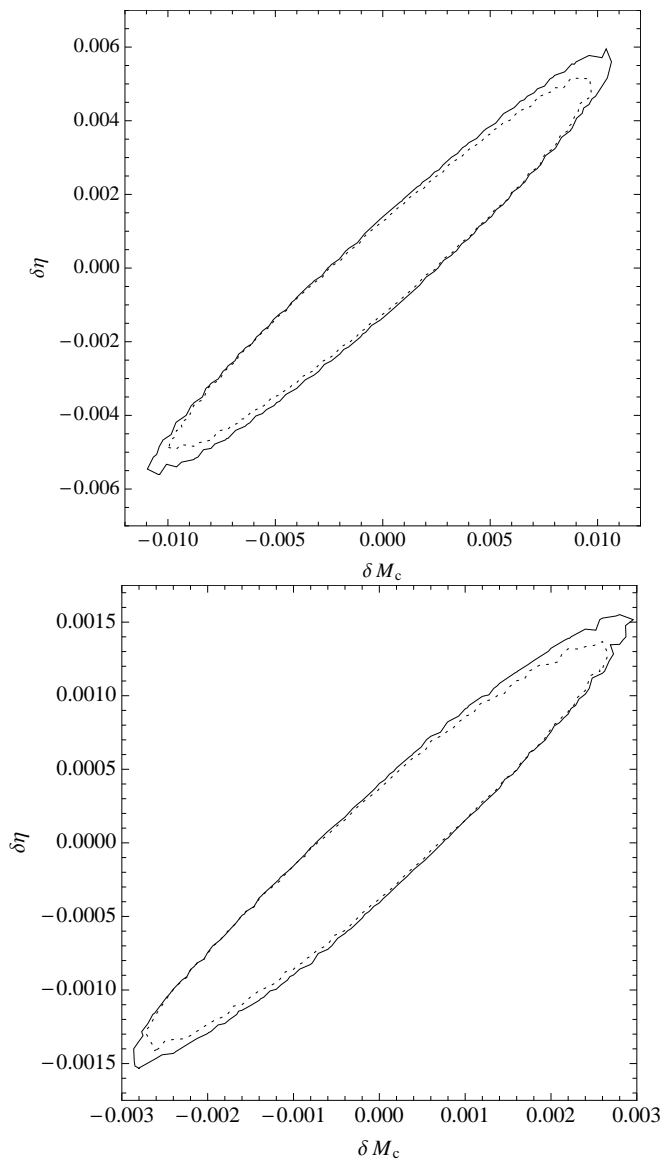
FIG. 4: **Comparison of the ambiguity contours between the leading-order (solid line) and higher-order (dotted line) waveforms for the non-spinning binaries.** $\delta\lambda$ is defined by a difference from the fiducial value in the Table I. The lines correspond to $P = 0.99$ (top panel) and $P = 0.999$ (bottom panel). Higher-order waveforms only marginally reduce the ambiguity contours.

cial small-scale structure; to avoid this effect we sample at a variety of data rates, typically either 8192 Hz (for $P > 0.99$) or 16384 Hz (for $P > 0.999$). Finally, the ambiguity function can also be impacted by our choice for the starting and ending frequency. For our calculations, we start integrating the waveform and integrate over all power above $f_{\mathrm{start}} = 40$ Hz. In our experience, this procedure best mimics the real data processing used in initial LIGO searches. However, a not-insignificant amount of power is present between 30 and 40 Hz; if included in the integral, the overlap would differ by $\Delta P \simeq 10^{-3}$,

comparable to some fine-scale structures of interest. At the other extreme, we terminate our evolution at the minimum-energy circular orbit (MECO), where the binary energy ceases to decrease monotonically.

One small, subtle, but important effect is the *nonzero overlap* of the waveforms along the $\pm\hat{z}$ axis[7]. For example, for our non-precessing binary, we find that for otherwise identical parameters

$$\max_{t_{\mathrm{c}},\psi}|\langle h(\hat{z})|h(-\hat{z}|t_{\mathrm{c}},\psi)\rangle| \simeq 1.7 \times 10^{-3} \qquad (34)$$

Equivalently, in the language of single-detector real overlaps, the sine and cosine chirps are *not precisely orthogonal*, for the same orbital phase[8]. To a first approximation the $h(\pm\hat{z})$ signals are *basis waveforms*; the waveform along any line of sight is a superposition of the two. Because these two signals are not orthogonal, the ambiguity function generally has fine-scale structure with $\Delta P \simeq 10^{-3}$, associated with the overlap of these two directions. For example, on this scale and below, the overlap between two non-precessing waveforms with just $l = |m| = 2$ emission is no longer well-described by Eq. (28). Instead, the ambiguity function gains additional fine-scale structure in angle. While extremely useful, for our purposes this result means that on sufficiently small scales $1 - P \lesssim 10^{-3}$, the complex overlap will have additional structure compared to "conventional" investigations of single-detector, optimally-oriented overlaps (i.e., overlaps of two real h_+ signals, extracted along $\hat{z}$). In particular, this nonzero overlap is partially responsible for the small-scale structure seen in Fig. 3.

Because of the many subtle interpretation and implementation issues associated with the smallest ambiguity scales, while we investigate the value of effective fitting to fine scales (e.g., $P > 0.999$), for simplicity we emphasize results for the scale relevant to most detection events ($P > 0.99$).

IV. RESULTS

Using a small set of fiducial simulations, we compare non-precessing and precessing signals against their immediate neighbors, mapping out an ambiguity function in each n-dimensional parameter space.

Higher harmonics perturb the ambiguity function by a quantifiable amount (i.e., $\delta\Gamma_{ab} \propto v^p$ for p depending on a, b and the harmonic). As has previously been shown elsewhere, we find that higher harmonics break degeneracies present in non-precessing, leading-order signals [19, 21, 23].

---

[7] Such waveforms would have a real overlap of unity, but the complex overlap is expected to be zero. See Appendix A.

[8] The real overlap is usually performed in conjunction with an explicit maximization over time and orbital phase, for a single harmonic. For that situation, this subtlety does not occur.
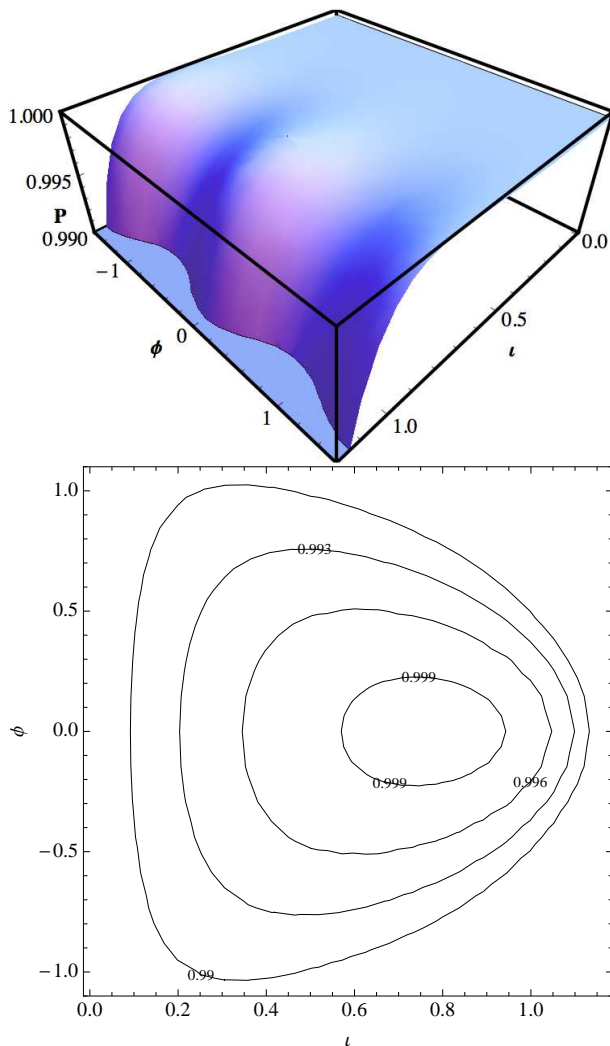
FIG. 5: **Comparison of the ambiguity surfaces between the leading-order (top panel) and higher-order (bottom panel) waveforms for the non-spinning binaries.** $\delta\lambda$ is defined by a difference from the fiducial value in the Table I. Higher-order waveforms change the ambiguity surface and break degeneracies related to the *inclination* and *orbital phase* so these parameters can be marginally observable at the scale of $P > 0.99$.

As described below, we generally find small but significant *scale-dependent* disagreement with the conventional stationary-phase Fisher matrix calculation, even in the absence of higher harmonics. Motivated by Fig. 2 and appendix B, we suspect that most scale dependence is introduced by suboptimal coordinates and can be minimized by a better choice of reference frequency. In Fig. 2, when we choose the reference frequency to be 40 Hz, we can directly see the fine structure depending on the scales of interest. In this case, the Fisher components $(\hat{\Gamma}_{M_c})_{\text{eff}}$ are 3634 ($P > 0.99$) and 7708 ($P > 0.999$). By using better coordinates ($f_{\text{ref}} \sim 100$ Hz), we can reduce the coordinate-dependent distortions from their originally considerable extent. Despite our best attempts

to find coordinates well-adapted to the problem, however, the change in Γ going from $P \simeq 0.99$ to $P \simeq 0.999$ for leading-order waveforms is comparable to the change in Γ going from leading-order waveforms to higher harmonics for the specific binary and line of sight considered here.

A. Zero spin

For a system without spin, higher harmonics principally provide information about the line of sight. For clarity, we will first discuss the most immediately relevant scale ($P \gtrsim 0.99$). Comparing the solid (without higher harmonics) and dotted (with higher harmonics) curves on the top panel of Fig. 4, we immediately see that higher harmonics provide little new information about intrinsic parameters, all other things being equal. Equivalently, looking at Table III, the effective Fisher matrix $\hat{\Gamma}$ on the two-dimensional parameters M_c, η without and with higher harmonics are similar to each other, as well as to a standard Fisher matrix computed using stationary phase approximation waveforms (labeled as $P \sim 1$). By contrast, as illustrated by the dramatic difference between the top and bottom panel in Fig. 5, higher harmonics produce a dramatic qualitative change in how well the line of sight $\hat{n}$ can be measured.

Both with and without higher harmonics [Fig. 5], the line of sight is very difficult to measure, particularly at the expected relevant scale $P \simeq 0.99$ (i.e., SNR of around 10). In both cases, the ambiguity function has a broad, extended, asymmetric extremum. In the absence of higher harmonics, the ambiguity function *cannot* be usefully described by a locally quadratic approximation, even a effective one. Nonetheless, by understanding the expected dependence on angle (and by adopting coordinates in band), we can propose a physically-well-motivated fitting function, both for the purely angular dependence and for the correlations between line of sight and other parameters [Eq. (30)]. This fitting function works extremely well when higher harmonics are neglected. When higher harmonics are included, a quadratic approximation suffices, as the local extremum is much narrower. In the latter case, Table IV provides the fitting parameters needed to reconstruct the full multidimensional fit. In fact, our well-chosen reference frequency produces a *nearly separable fit*, with zero off-diagonal terms [e.g., $\Gamma_{aN} \simeq 0, G_{\phi c} \simeq 0$ in Eq. (30)]. We will return to this simple structure frequently below.

Several effects besides higher harmonics also introduce fine-scale structure into the ambiguity function. As demonstrated in Fig. 2, the choice of reference frequency can introduce strong, scale-dependent features into the ambiguity function. We chose a reference frequency at 100 Hz to reduce its effect, but have not eliminated it completely. The nonzero overlap between the $(l, m) = (2, \pm 2)$ modes is another such effect. Hence, we are not surprised that our effective fitting parameters change as we reduce the range of P used in the fit, even

pN order		leading-order						higher-order			
fitting scale		P > 0.99		P > 0.999		P ∼ 1		P > 0.99		P > 0.999	
parameter		$M_{\rm c}$	η	$M_{\rm c}$	η	$M_{\rm c}$	η	$M_{\rm c}$	η	$M_{\rm c}$	η
$(\hat{\Gamma}_{ij})_{\rm eff}$	$M_{\rm c}$	3012±7	-5505±13	3621±69	-6474±125	2547	-5314	3243±8	-6070±13	3932±77	-7224±143
	η	-	10675±25	-	12478±228	-	11954	-	12151±27	-	14367±270
c_{ij}	$M_{\rm c}$	1.00	0.971	1.00	0.963	1.00	0.963	1.00	0.967	1.00	0.961
	η	-	1.00	-	1.00	-	1.00	-	1.00	-	1.00
σ_i		0.00760	0.00404	0.00618	0.00333	0.00735	0.00339	0.00689	0.00356	0.00578	0.00302

TABLE III: **Effective fitting parameters for a non-spinning binary, at different scales.** The fiducial values of parameters are $(M_{\rm c}, \eta)$=(2.994, 0.1077). Fitting parameters and uncertainties are calculated from a least-squares fit, treating the line of sight as fixed. For the one-dimensional errors σ_i, we adopt $\rho = 10$. Result for $P \sim 1$ is calculated by Eq. (21) using h_+ polarization of the SPA waveforms in $M_{\rm c}, \eta, t_{\rm c}$ and $\phi_{\rm ref}$, followed by analytic maximization over $t_{\rm c}$ and $\phi_{\rm ref}$; see the text for details. [The effective Fisher matrix corresponds to derivatives of an overlap maximized over $t_{\rm c}$ and ψ.] Systematic differences exist between the SPA waveforms and real overlap used in our numerical calculations and the analytic result shown for $P \sim 1$. Nonetheless, all methods largely agree: higher harmonics provide fairly little additional information about the chirp mass and mass ratio correlations, with all other parameters fixed; see Fig. 4. For comparison, this table provides fitting parameters for P for two volumes ($P > 0.99$ and $P > 0.999$). Because of the effects described in Fig. 2, Section III G, and Appendix B, the two fits disagree.

fitting scale			P > 0.99		
parameter		$M_{\rm c}$	η	ι	ϕ
$(\hat{\Gamma}_{ij})_{\rm eff}$	$M_{\rm c}$	3883±38	-7236±66	-0.08350±0.48	0.08169±0.23
	η	-	14209 ±138	0.1181±0.93	1.172±0.37
	ι	-	-	0.02723±0.0009	-0.001413 ±0.00052
	ϕ	-	-	-	0.03196 ±0.0003
c_{ij}	$M_{\rm c}$	1.00	0.976	-0.00313	-0.271
	η	-	1.00	-0.00494	-0.276
	ι	-	-	1.00	0.0474
	ϕ	-	-	-	1.00
σ_i		0.00740	0.00387	0.607	0.583

TABLE IV: **Effective fitting parameters for a non-spinning binary using higher-order waveforms.** Fitting parameters in this table are calculated by least-squares to a quadratic form $P \simeq 1 - (\hat{\Gamma}_{ij})_{\rm eff}\delta\lambda^i\delta\lambda^j/2$. Measurements of intrinsic $(M_{\rm c}, \eta)$ and extrinsic (line of sight) parameters *separate*: with only one exception $(\hat{\Gamma}_{\eta\phi})_{\rm eff}$, all off-diagonal terms coupling the line of sight and intrinsic parameters are consistent with 0. We anticipate a slightly different choice of reference frequency will eliminate the small residual correlation that remains. Fitting a general form described in Eq. (30) that accounts for the manifestly nonquadratic behavior shown in Fig. 5 leads to comparable results: a separable fit (i.e., $G_{c\phi} = 0$ and $\Gamma_{aN} \simeq 0$) that performs little better than the quadratic form used above.

pN order		leading-order						higher-order			
fitting scale		P > 0.99		P > 0.999		P ∼ 1		P > 0.99		P > 0.999	
$P_{\rm max}$		0.999767±1.09 × 10⁻⁵		0.999884±0.78 × 10⁻⁵		n/a		0.999775±1.01 × 10⁻⁵		0.999888±0.86 × 10⁻⁵	
parameter		$M_{\rm c}$	η	$M_{\rm c}$	η	$M_{\rm c}$	η	$M_{\rm c}$	η	$M_{\rm c}$	η
$(\hat{\Gamma}_{ij})_{\rm eff}$	$M_{\rm c}$	2899±8	-5295±14	2980±62	-5325±111	2546	-5313	3125±8	-5848±14	3274±72	-6014±133
	η	-	10261±28	-	10276±207	-	11954	-	11701±29	-	11974±257
c_{ij}	$M_{\rm c}$	1.00	0.971	1.00	0.962	1.00	0.963	1.00	0.967	1.00	0.960
	η	-	1.00	-	1.00	-	1.00	-	1.00	-	1.00
σ_i		0.00776	0.00413	0.00673	0.00363	0.00735	0.00339	0.00704	0.00364	0.00628	0.00328

TABLE V: **Non-spinning binary fit, using an alternate technique:** As Table III, except the effective fitting function allows $P_{\rm max}$ to be a parameter. In this example, this more generic fit produces effective Fisher matrices that more closely correspond to conventional stationary-phase results for leading-order emission.

11

pN order	leading-order									higher-order					
method	iterative			simultaneous			$P \sim 1$			iterative			simultaneous		
parameter	$M_{\rm c}$	η	χ	$M_{\rm c}$	η	χ	$M_{\rm c}$	η	χ	$M_{\rm c}$	η	χ	$M_{\rm c}$	η	χ
$(\hat{\Gamma}_{ij})_{\rm eff}$ $M_{\rm c}$	3686	-1652	-1007	3567±22	-1570±14	-975.6±5.9	2837	-2423	-669.6	4217	-2147	-1177	4129±34	-2083±25	-1157±9.3
η	-	1237	515.5	-	1170±9	492.4±3.5	-	2357	612.3	-	1806	685.6	-	1765±18	670.5±6.4
χ	-	-	283.8	-	-	275.5 ±1.6	-	-	163.8	-	-	340.0	-	-	335.5±2.6
c_{ij} $M_{\rm c}$	1.00	-0.947	0.996	1.00	-0.957	0.997	1.00	-0.981	0.994 c	1.00	-0.929	0.994	1.00	-0.935	0.995
η	-	1.00	-0.969	-	1.00	-0.974	-	1.00	-0.995	-	1.00	-0.958	-	1.00	-0.962
χ	-	-	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-	1.00
σ_i	0.0291	0.0180	0.135	0.0323	0.0201	0.149	0.0512	0.0621	0.440	0.0227	0.0131	0.104	0.0238	0.0137	0.108

TABLE VI: **Effective fitting parameters for an aligned-spin binary:** As Table III for a BH-NS binary with $\chi = 1$. This binary has one new parameter (χ). For all parameters shown here, a local quadratic form is a good approximation to the ambiguity function; this table provides the fitting parameters, treating the line of sight as fixed. For comparison purposes, this table provides both our standard simultaneous least-squares technique (applied to $M_{\rm c}, \eta, \chi$, with all other parameters fixed) and an iterative technique applied to successive low-order subspaces. Table VII provides the results for a more generic fit including line-of-sight dependence.

fitting scale	$P > 0.99$				
parameter	$M_{\rm c}$	η	χ	ι	ϕ
$(\hat{\Gamma}_{ij})_{\rm eff}$ $M_{\rm c}$	4388±54	-2256±40	-1230±15	-0.2527±0.47	-0.5585 ±0.38
η	-	1869±28	719.0±9.8	0.3000±0.31	1.122±0.23
χ	-	-	356.2±4.2	0.1098 ±0.14	0.3396±0.11
ι	-	-	-	0.03279±0.0009	-0.001972±0.0007
ϕ	-	-	-	-	0.02530±0.00025
c_{ij} $M_{\rm c}$	1.00	-0.960	0.997	-0.206	-0.622
η	-	1.00	-0.976	0.189	0.567
χ	-	-	1.00	-0.204	-0.616
ι	-	-	-	1.00	0.186
ϕ	-	-	-	-	1.00
σ_i	0.0324	0.0177	0.147	0.566	0.818

TABLE VII: **Effective fitting parameters for an aligned-spin binary using higher-order waveforms:** As Table IV, but for a BH-NS binary with $\chi = 1$. To a first approximation, the Fisher matrix roughly separates the intrinsic and extrinsic parameters. The weak covariances shown here are extremely susceptible to small changes in the Fisher matrix coefficients; only a single preferred value is shown.

in the absence of harmonics; see Table III.

The effect of scale dependence is fairly mild: the eigendirections for $P > 0.99$ and $P > 0.999$ agree, only the eigenvalue scale changes. For comparison, we also considered an alternate fitting technique that allowed the single best fit point to have $P_{\rm max} \neq 1$; see Fig. 4, and Table V. While this method leads to aesthetically pleasing results similar to analytic calculations, this fit systematically underestimates P near the maximum-likelihood point and does not completely eliminate the trend towards different fitting parameters on the smallest scales. We henceforth adopt $P_{\rm max} = 1$.

To facilitate approximate comparisons with prior work, Tables III and IV provide one-dimensional standard deviations σ and correlation coefficients c_{ij}. Unless otherwise stated, these quantities are derived *solely from the Fisher matrix fits from that same table*. For example, in Table III, the "measurement errors" σ_i follow from inverting the 2×2 matrix shown, while in Table IV they follow from inverting a full 4×4 matrix.

B. Aligned-spin

Repeating our effective Fisher matrix calculation for an aligned-spin BH-NS binary leads to results qualitatively similar to the zero-spin case. As previously, higher harmonics provide little additional information about intrinsic parameters, here $M_{\rm c}, \eta, \chi$; see Fig. 6. For example, looking at data along a fixed line of sight, Table VI shows that, on a component-by-component basis (Γ) and overall (σ_i), the two ambiguity functions with and without higher harmonics resemble one another. More directly, Fig. 6 shows three-dimensional contours of nearly constant P as a function of $M_{\rm c}, \eta, \chi$ for both leading-order emission (blue) and higher harmonics (red). While higher harmonics clearly do provide more information about intrinsic parameters – the red surface is nested inside the blue – the addition of higher harmonics only marginally improves our ability to measure the least-well-constrained combination of $M_{\rm c}, \eta, \chi$. Precisely as in the non-spinning case, however, higher harmonics provide more information about the line of sight. Despite our

pN order				leading-order			
parameter		M_c	η	χ	β	θ_{NJ}	α_{JL}	ϕ
	M_c	3234	-281.1	-612.8	868.8	-1.359	3.726	-10.18
	η	-	929.9	-3.089	-171.7	1.252	-6.253	11.33
	χ	-	-	132.0	-177.3	0.3616	-0.1832	2.147
$(\hat{\Gamma}_{ij})_{\rm eff}$	β	-	-	-	295.2	0.06336	1.674	-6.945
	θ_{NJ}	-	-	-	-	0.7774	-0.003074	0.006970
	α_{JL}	-	-	-	-	-	0.3349	-0.1021
	ϕ	-	-	-	-	-	-	0.3829
	M_c	1.00	0.407	0.730	0.144	-0.0729	-0.109	-
	η	-	1.00	0.742	0.729	-0.183	0.109	-
c_{ij}	χ	-	-	1.00	0.752	-0.180	-0.133	-
	β	-	-	-	1.00	-0.191	-0.106	-
	θ_{NJ}	-	-	-	-	1.00	0.0126	-
	α_{JL}	-	-	-	-	-	1.00	-
σ_i		0.00592	0.00573	0.0433	0.0212	0.116	0.187	-
pN order				higher-order			
parameter		M_c	η	χ	β	θ_{NJ}	α_{JL}	ϕ
	M_c	3401	-408.2	-644.3	911.9	-1.745	4.515	-11.89
	η	-	1055	7.249	-208.3	1.495	-7.229	12.58
	χ	-	-	139.3	-186.3	0.5298	-0.1706	2.471
$(\hat{\Gamma}_{ij})_{\rm eff}$	β	-	-	-	307.9	0.02350	1.998	-7.544
	θ_{NJ}	-	-	-	-	0.8305	0.003262	0.006150
	α_{JL}	-	-	-	-	-	0.3510	-0.1029
	ϕ	-	-	-	-	-	-	0.3951
	M_c	1.00	0.543	0.778	0.309	-0.154	-0.177	-
	η	-	1.00	0.815	0.786	-0.265	-0.00409	-
c_{ij}	χ	-	-	1.00	0.815	-0.270	-0.230	-
	β	-	-	-	1.00	-0.271	-0.204	-
	θ_{NJ}	-	-	-	-	1.00	0.0381	-
	α_{JL}	-	-	-	-	-	1.00	-
σ_i		0.00626	0.00626	0.0495	0.0237	0.115	0.188	-

TABLE VIII: **Effective fit for a precessing binary at the scale $P > 0.99$ using an iterative method:** The effective Fisher matrix needed to reproduce the calculated ambiguity function, derived from an iterative fit; see Table IX for a full 7-dimensional fit. Both of these matricies have comparable lists of eigenvalues $(3625, 908, 54, 4.9, 0.74, 0.29, \simeq 0)$ and $(3844, 999, 57, 4.1, 0.75, 0.28, \simeq 0)$ [bottom] and with the nearly zero eigenvalue roughly corresponding to the ϕ direction. Since the eigenvectors span several orders of magnitude, the correlation coefficients of this poorly-conditioned matrix are extraordinarily sensitive to small changes in the coefficients and have not been provided. Unlike the non-precessing case, the geometric and intrinsic parameters do not completely separate, even for these well-adapted coordinates. The one-parameter uncertainties σ_i shown are calculated by omitting the (nearly unmeasurable) ϕ direction from the fit.

line of sight providing some sensitivity to a symmetry-breaking $(2,1)$ harmonic, a waveform with harmonics encodes roughly similar information as a waveform without harmonics. Tables VI and VII provide the effective fitting parameters we used to reproduce their ambiguity function. As in the non-spinning case, we find an *approximately* separable fit, though less so than before [see c_{ij} in Table VII]. Even allowing for weak correlations between intrinsic parameters and the line of sight, the overall parameter covariances σ_i with harmonics [Table VII] are nearly unchanged from a model with only leading-order

emission [Table VI].[9]

The aligned-spin results at $\chi = 1$ cannot be easily compared with the corresponding zero-spin results $(\chi = 0)$. On the one hand, the component-by-component Fisher

---

[9] For leading-order emission, measurements of orientation and intrinsic parameters separate completely. The intrinsic parameter uncertainties computed in Table VI are therefore identical to the uncertainties for the corresponding parameters derived from a full 5×5 Fisher matrix.

pN order	leading-order						
parameter	$M_{\rm c}$	η	χ	β	$\theta_{\rm NJ}$	$\alpha_{\rm JL}$	ϕ
$(\hat{\Gamma}_{ij})_{\rm eff}$ $M_{\rm c}$	3279±12	-291.4±20	-618.8±2.7	878.7±3.4	-2.168±0.68	3.982±0.38	-11.58±0.32
η	-	922.8 ±5.3	-5.187±4.5	-175.5±4.7	1.648±0.34	-6.132 ±0.17	11.80±0.12
χ	-	-	132.3±0.52	-177.9±0.69	0.4429±0.13	-0.2340±0.079	2.321±0.063
β	-	-	-	295.1±1.5	0.1488±0.20	1.674 ±0.11	-7.263 ±0.058
$\theta_{\rm NJ}$	-	-	-	-	0.8038±0.0046	-0.004911 ±0.0063	0.005865±0.0069
$\alpha_{\rm JL}$	-	-	-	-	-	0.3299 ±0.0018	-0.1100±0.0035
ϕ	-	-	-	-	-	-	0.3850±0.0020
c_{ij} $M_{\rm c}$	1.00	0.477	0.747	0.238	-0.0961	-0.104	-
η	-	1.00	0.807	0.792	-0.265	0.0961	-
χ	-	-	1.00	0.801	-0.251	-0.108	-
β	-	-	-	1.00	-0.284	-0.0817	-
$\theta_{\rm NJ}$	-	-	-	-	1.00	0.0119	-
$\alpha_{\rm JL}$	-	-	-	-	-	1.00	-
σ_i	0.00622	0.00661	0.0495	0.0241	0.117	0.187	-

pN order	higher-order						
parameter	$M_{\rm c}$	η	χ	β	$\theta_{\rm NJ}$	$\alpha_{\rm JL}$	ϕ
$(\hat{\Gamma}_{ij})_{\rm eff}$ $M_{\rm c}$	3403±13	-420.2 ± 21	-640.4±2.4	916.4 ±3.6	-2.845±0.74	5.016 ±0.40	-12.35±0.31
η	-	1057±6.4	18.55 ±4.9	-212.3±5.3	1.533 ±0.38	-7.488 ±0.20	12.73±0.13
χ	-	-	137.2 ±0.55	-185.2 ±0.73	0.6655 ±0.14	-0.2104±0.090	2.472±0.061
β	-	-	-	308.5±1.2	-0.01406±0.22	2.022±0.12	-7.594±0.060
$\theta_{\rm NJ}$	-	-	-	-	0.8481 ±0.0048	0.003203 ±0.0067	0.01161±0.0068
$\alpha_{\rm JL}$	-	-	-	-	-	0.3456 ±0.0018	-0.1035±0.0038
ϕ	-	-	-	-	-	-	0.3911±0.0020
c_{ij} $M_{\rm c}$	1.0	0.391	0.726	0.127	-0.0705	-0.193	-
η	-	1.00	0.718	0.712	-0.226	0.0670	-
χ	-	-	1.00	0.744	-0.224	-0.234	-
β	-	-	-	1.00	-0.252	-0.170	-
$\theta_{\rm NJ}$	-	-	-	-	1.00	0.0224	-
$\alpha_{\rm JL}$	-	-	-	-	-	1.00	-
σ_i	0.00584	0.00529	0.0421	0.0210	0.113	0.191	-

TABLE IX: **Effective fit for a precessing binary at the scale $P > 0.99$ using a simultaneous method.** The effective Fisher matrix needed to reproduce the calculated ambiguity function, derived from a full 7-dimensional least squares; see Table VIII for an alternative iterative fit. Both methods produce comparable results, with comparable eigenvalue distributions. The difference between the leading-order result and a model including higher harmonics is small, though usually significantly in excess of our fitting parameter error (e.g., several standard deviations) and of the systematic differences between the two fitting methods. As expected, higher harmonics lead to *smaller* parameter correlations

matrix coefficients like $\Gamma_{\eta\eta}$ will differ significantly, as the waveform phasing changes as a function of χ and hence so does $\langle(\partial_\eta\Psi)^2\rangle$. On the other hand, the aligned-spin results allow a *new parameter* (spin) that was treated as fixed for the $\chi = 0$ case, with nontrivial coupling to the other intrinsic parameters. The one-parameter uncertainties σ_i are *dramatically* increased by including this previously-neglected systematic effect.

Both at leading and higher-order, our effective fit to the ambiguity function is complicated by the wide range of scales in Γ, even for fixed line of sight. As is well-known from previous Fisher matrix calculations with aligned-spins [40, 41], the ambiguity function in $M_{\rm c}, \eta, \chi$ has strong correlations, producing a narrow and extended

extremum. For the specific example described by our effective Fisher matrix, the $M_{\rm c}, \eta, \chi$ submatrix has eigenvalues $\simeq 4200, 400, 30$, describing a strong hierarchy of scales. For our purposes, Fig. 6 demonstrates our fiducial aligned-spin binary cannot be distinguished from binaries with spin $\chi \gtrsim 0.8$: for each χ in this range, suitable combinations of $M_{\rm c}, \eta$ exist with high overlap. For these extremely extended ambiguity ellipsoids, a fit that reproduces $P(\lambda_0, \lambda)$ over the full range in χ *might* require a more generic functional form than the one adopted so far: a quadratic with constant coefficients *in the neighborhood of the fiducial binary*. Effectively speaking, however, these additional degrees of freedom add little information with considerable expense. We will explore more
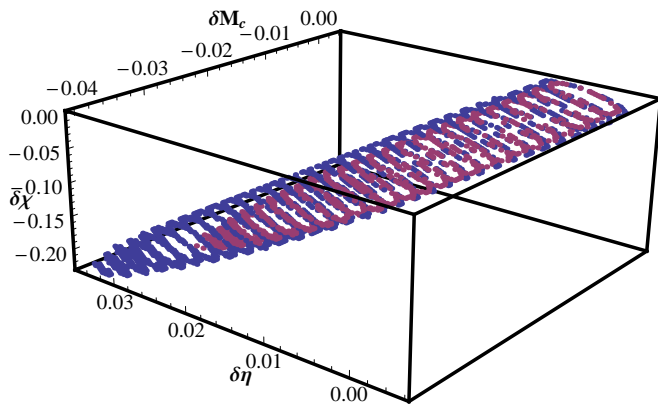
FIG. 6: **Ambiguity function for aligned-spin**: Points in M_c, η, χ with $P \in [0.99, 0.991]$, shown as blue (leading-order) and red (with higher harmonics). The two long, narrow surfaces (ambiguity ellipsoids) these points cover illustrates gravitational wave measurements can constrain one combination of M_c, η, χ tightly (e.g., M_c); one less so (e.g., η); and one almost not at all. The two surfaces nearly agree, with the most significant change being a slight reduction in the least-well-determined direction. The close agreement between these two surfaces shows higher harmonics provide fairly little additional information to break this degeneracy.

complicated effective dependence in a subsequent publication.

As in the zero-spin case, we find significant differences on the smallest scales in P, in a fashion that depends on the reference frequency. Given the number of dimensions, complex functional form, sensitivity to numerical implementation like the sampling rate, and less immediate observational relevance, we defer a detailed discussion of fine-scale effects to a subsequent paper.

C. Precessing spin: Case 1

For the first of our two fiducial precessing binaries, we find higher harmonics provide little added information beyond the constraints produced in the non-precessing case. This unfortunate but expected result can be seen, for example, from the one-dimensional covariances in Table IX; from the effective Fisher matrix coefficients $\hat{\Gamma}_{ij}$; or from their comparable sequences of eigenvalues. That said, even in the absence of higher harmonics, the ambiguity function for a *precessing* binary has simpler structure than the non-precessing result, with reduced correlations among the "intrinsic" parameters; a somewhat less extreme hierarchy of scales (i.e., eigenvalues)[10]; and roughly speaking a more quadratic ambiguity function.

In fact, for a precessing binary the previous clear separation between "intrinsic" and "geometric" parameters breaks down. As each instant the opening angle β of the precession cone of L around J depends on the relative magnitude of L and $S = J - L$, as well as on their (nearly conserved) misalignment angle $\hat{L} \cdot \hat{S}$. The magnitudes of L and S are essentially *intrinsic* parameters, characterizing the binary masses and BH spin; therefore, we expect the precession cone opening angle β to be intimately correlated with the intrinsic parameters. At the same time, the precession cone opening angle must be intimately connected to the "geometric" parameters that define the orientation of the binary at the reference frequency: $\theta_{NJ}, \alpha_{JL}, \phi$. Specifically, the orientation of the precession of L relative to the line of sight is characterized by the two angles θ_{NJ} (setting the orientation of J)[11] and α_{JL} (fixing the orientation of L along the precession cone at the reference frequency). The orbital phase ϕ at the reference frequency fixes the binary's geometry in band. This cross-coupling between intrinsic and extrinsic parameters has quantitative consequences for the Fisher matrix. Roughly speaking, the two new eigenvalues introduced into Γ by allowing precession (here with values $\simeq 5, 0.7$, associated with the β and α_{JL} parameters) can be expected to lie *between* the very large $(3000, 900, 100)$ and very small $(\simeq 0.3, \simeq 0)$ eigenvalues associated with the manifestly intrinsic (M_c, η, χ) and extrinsic (θ_{NJ}, ϕ) parameters.

For non-precessing binaries, the choice of a 100 Hz reference frequency nearly separated intrinsic and extrinsic parameters. A suitable choice of reference frequency may yet further reduce the off-diagonal terms in our effective Fisher matrix. For the present coordinates, however, we cannot cleanly decompose parameters into "intrinsic" and "geometric" parameters. Table IX shows correlation coefficients c_{ij} calculated by omitting the (nearly unmeasurable) ϕ coordinate in Γ; no obvious block-diagonal form occurs.

D. Precessing spin: Case 2

By contrast to the relatively simple ambiguity functions seen so far, our second set of binary parameters produces a significantly more complicated ambiguity function, particularly in geometric parameters. For example, Fig. 7 shows the ambiguity surfaces for $(\theta_{NJ}, \alpha_{JL})$ and $(\beta_{JL}, \alpha_{JL})$ fixing all other parameters, for "case 2". As in case 1, we have a highly symmetric binary starting with L, J, and the line of sight in the same plane at our ref-

---

[10] We can always rescale our eigenvalues by rescaling our coordinate units. However, in these units all parameters have a prior range of order unity, and physical meaning. As a result, our eigenvalues also have meaning: the coordinate combinations corresponding

to the smallest eigenvalues have minimal impact on the overlap and can be ignored.

[11] The other angle needed to specify the orientation of J is equivalent to a rotation around the line of sight, i.e. the polarization angle. As the complex overlap maximizes over this angle, it is explicitly removed as a parameter.
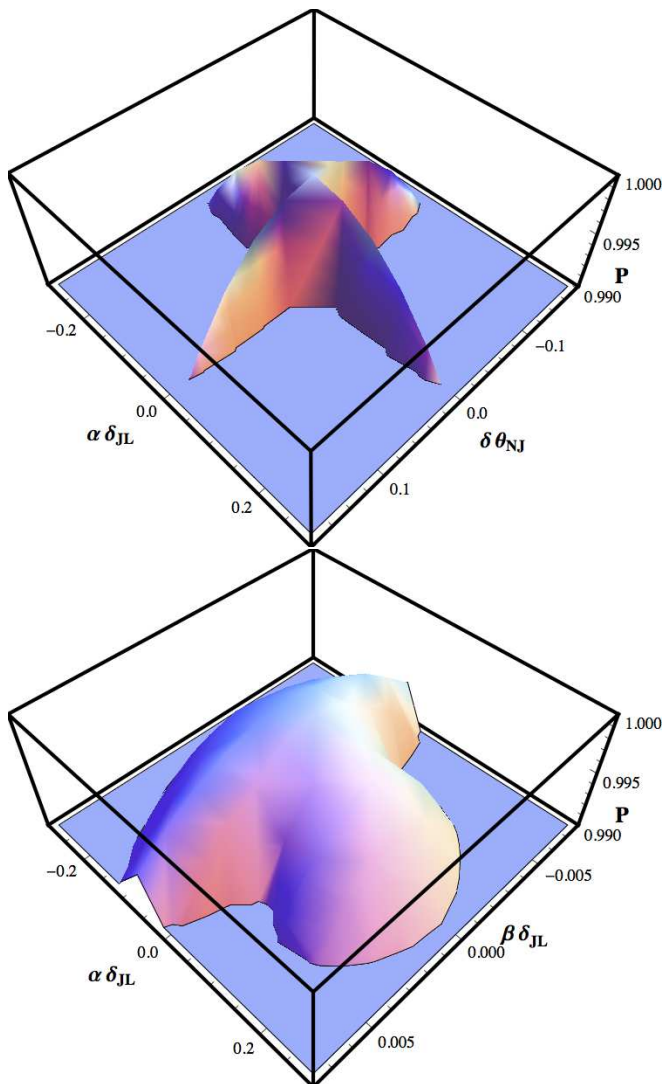
FIG. 7: **Symmetry can produce highly nonquadratic behavior**: Plot of the ambiguity function when changing only θ_{NJ} and α_{JL} (top panel) and β_{JL}, α_{JL} (bottom panel), for the high-symmetry binary parameter set "case 2".

erence frequency. However, in this case we start with L parallel to the line of sight, rather than perpendicular to it. The ambiguity function shows extreme sensitivity to the initial conditions and highly nonquadratic behavior. These differences occur despite the considerable similarity between case 1 and case 2: the two are, to an excellent approximation, the same configurations, just slightly offset in time. By contrast, the change in ambiguity versus $M_{\rm c}, \eta, \chi$ is well-described by a quadratic form.

This extreme scenario demonstrates that even an *effective* Fisher matrix has limits: sometimes, a more generic functional form including higher-order correlations must be used on relevant scales. That said, highly nonquadratic behavior only occurred for a high-symmetry binary. We expect typical binary initial conditions will produce nearly quadratic ambiguity functions.

V. CONCLUSIONS

In this paper we have used case studies of a coherent, two-detector ambiguity function $P(\lambda_0, \lambda)$ to estimate how much and what kind of information higher harmonics provide about BH-NS binaries. Given the high dimension of and severe degeneracies that plague the problem, we perform a tractable, idealized calculation instead of the straightforward but less easily understood explicit source in a real multi-detector network. Specifically, we place a single source directly overhead an idealized detector pair, equally sensitive to both polarizations. For all binaries, we find that higher harmonics provide little additional information about the binary's intrinsic parameters. Instead, at best they provide information about the orientation of the source relative to the line of sight (e.g., ι, ϕ for non-precessing binaries). Notably, higher harmonics make it easier to exclude a non-precessing binary with $\hat{L} \simeq \pm \hat{n}$.

When possible, we estimate the two-point function $P(\lambda_0, \lambda)$ on scales of interest $(1 - P < 1/\rho^2)$ by a locally quadratic function $1 - \Gamma_{\alpha\beta} \delta\lambda^\alpha \delta\lambda^\beta / 2$, whose coefficients we denote the "effective Fisher matrix." Using coordinates λ adapted to the most sensitive frequencies of our network, we show that to an excellent approximation the Fisher matrix for non-precessing binaries separates into intrinsic (masses and spins) and extrinsic (geometric) parameters. For precessing binaries, the angles describing the precession cone are intimately connected with the intrinsic masses and spins as well as the geometric parameters. Therefore, they correlate with both intrinsic and extrinsic parameters and entangle the two previously distinct sets of parameters. In future work we will compare our results to detailed posterior parameter distributions, as computed by Markov-chain Monte Carlo simulations. Our results suggest suitable coordinates will substantially simplify the interpretation of these posteriors.

In addition to our main results, this paper also provides technical suggestions of broad interest to the data analysis community. First, rather than performing the conventional (single-polarization, real) overlap maximized over time and orbital phase, we use a coherent two-polarization overlap maximized over time and polarization [42]. This overlap is far more discriminating than the single-polarization result and, lacking a maximum over the polar emission direction ϕ, is well adapted to asymmetric situations such as precession or the presence of higher harmonics. Second, we illustrate the importance of choosing coordinates adapted to the network's sensitive band and to the binary's in-band geometry. We have provided concrete examples of the consequences of poor choices for coordinate conventions, which can lead to pathological behavior in the complex overlap and ambiguity function. Finally, we recommend adopting an effective Fisher matrix, derived by fitting the ambiguity function on the scale at which variations in the parameters could be plausibly detected. We have shown that the conventional approach to the Fisher matrix can be

sensitive to unobservable fine-scale structure which can give misleading results. By contrast, simply by computing the ambiguity function over the area of interest, then fitting, one can explicitly verify whether a quadratic approximation even applies, as well as assess its error.

Appendix A: Comparison between the Real and Complex overlaps

One detector can not be sensitive to both polarizations, as it measures a single (real) strain variable h. The conventional definition of the overlap [Eq. (4)] therefore provides an inner product for a single real data sequence.

The complex overlap [Eq. (2)], however, uses information about both polarizations [33]. To illustrate the differences between these two diagnostics, Figs. 8 and 9 show the overlaps $\langle h_0|h \rangle$ between our fiducial non-spinning waveform and an identical binary, except for one parameter. In both cases, these overlaps are maximized over event time and polarization (for the complex overlap) or orbital phase (for the real overlap).

First and foremost, Fig. 8 shows that the *real* overlap does not change as inclination (ι) or the orbital phase (ϕ) are varied. In other words, with one detector, we cannot identify the inclination, which measures the relative amplitude of h_+ to $h_\times$. Also, by maximizing over ϕ, we lose information about it. The conventional single-detector overlap therefore provides no information about how well we can measure these parameters in a network sensitive to two polarizations.

By contrast, because the complex overlap explicitly uses two polarizations, it can identify the inclination; see the dotted line in Fig. 8. For example, because it can distinguish between left- and right-handed sources, the overlap between antipodal directions (i.e., $\iota = 0, \iota = \pi/2$) is zero. Moreover, if the source orientation is different from 0, then the waveform carries information about the orbital phase, which the complex overlap easily identifies. As a concrete example, if $\iota = \pi/2$ for the fiducial signal, the orbital phase ϕ is measurable at high SNR. When $\iota = \pi/4$, ϕ is unmeasurable for the leading-order wave-forms (for the analytic formula, see Eq. (B1) in [33]) but marginally measurable for the higher-order waveforms (see Fig. 5).

On the other hand, when *intrinsic* parameters are varied, the complex and real overlaps largely agree, even though the complex overlap has more information available; see Fig. 9.

For the figures and discussion in this paper, we compare the real and complex *normalized* overlaps, dividing each network response by the network signal to noise. For the two identical detectors used in the complex overlap, the network SNR is just a quadrature sum of two detectors; see Eq. (3),

$$\rho^2 = \rho_+^2 + \rho_\times^2. \tag{A1}$$

By contrast, for the real overlap, the relevant SNR is just ρ_+^2. For an identical source, a two-detector network has higher overall SNR.

One point of this paper is to compare the real and complex overlaps from a parameter estimation point of view. Starting from Eq. (11), the real overlap enters directly into the expression for the posterior. As a result, contours of the real ambiguity function $(h_0|h)$ should closely correspond to contours of the posterior parameter distribution, for measurements limited to a single polarization and known sky location. By contrast, for a source with known sky location seen by a network with comparable sensitivity to both polarizations, the *complex* overlap enters directly into the expression for the posterior. A real network will have unequal sensitivity to two polarizations. We therefore expect the real posterior will resemble some average between the posteriors estimated using the single-detector and network overlaps. In a subsequent publication we will compare our results with posteriors computed by MCMC, to quantify how well our simple estimates do at characterizing measurement accuracy.

Appendix B: Reference frequency and fine-scale structure in the ambiguity function and systematic errors

Even for a zero-spin binary, the choice of reference frequency at which the parameters are defined significantly influences the structure of the ambiguity function $P(\lambda_0, \lambda)$. As a concrete example, Fig. 10 shows that specifying the orbital phase at the start or end of the waveform introduces additional structure on physical scales into the ambiguity function for the most well-determined and physical parameters, the chirp mass M_c and mass ratio η. Even more troubling, Fig. 11 shows that an ill-chosen reference frequency can introduce *extremely* fine-scale structure into the ambiguity function, for orientations away from $\hat{z}$. By contrast, a reference frequency $f_{\rm ref}$ close to the half-power point of the detector
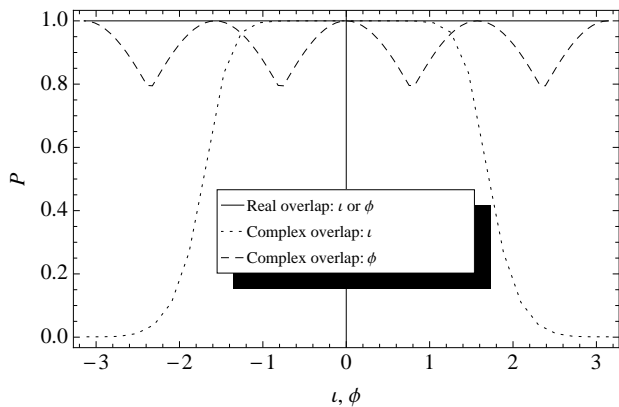
FIG. 8: **Comparison of the ambiguity functions between a real and complex overlaps for the extrinsic parameters.** Dotted line is calculated by changing ι where the fiducial value is 0 and other parameters are the same as in Table I. Dashed line is calculated by changing ϕ where $\iota = \pi/2$ and other parameters are the same as in Table I. The ambiguity surface for the real overlap is flat.
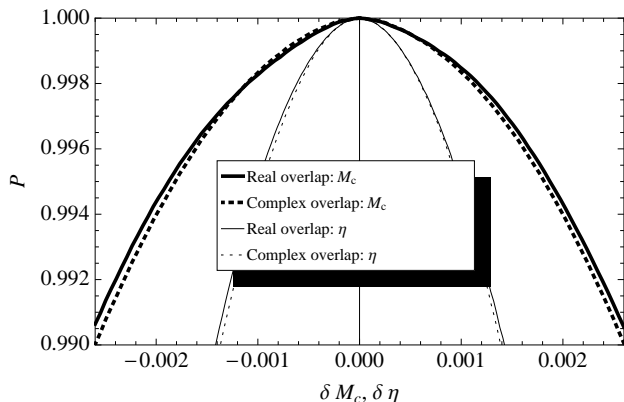

FIG. 9: **Comparison of the ambiguity functions between a real and complex overlaps for the intrinsic parameters.** Parameter values are summarized in Table I. Thick lines are calculated by changing M_c, others by changing η.

reduces these effects, where $f_{\rm ref}$ is estimated by

$$\frac{d\rho^2}{df} \equiv 4\frac{|h(f)|^2}{S_h} \qquad (B1)$$

$$\int_0^{f_{\rm ref}} \frac{d\rho^2}{df} = \int_{f_{\rm ref}}^\infty \frac{d\rho^2}{df} \qquad (B2)$$

for $|h(f)| \propto f^{-7/6}$ the standard restricted amplitude. For our binary model with $\iota = \pi/4$, when we choose the reference frequency lower than the optimal frequency, the change of the structure of the ambiguity function can be significant depending on the frequency, while these changes are relatively small when we choose the higher reference frequency; see Fig. 12.

What introduces this severe dependence for such a vanilla waveform? The significant accumulation of orbital and hence waveform phase between the coordinates' base point and the detector's sensitive band. Roughly speaking, the detector is sensitive to the configuration of the binary as it crosses through its sensitive band. This trajectory can be characterized by some instantaneous parameters $\bar{\lambda}$. By contrast, the waveform at a significantly earlier or later time has rotated roughly $\Delta\Phi/\pi$ times between that frequency and the observed one. As a result, derivatives of the waveform relative to λ differ from derivatives relative to $\bar{\lambda}$ by a term of order $h\partial\delta\Phi/\partial\bar{\lambda}^a$. As a result, the more sensitive the waveform is to a parameter, the more severe the *absolute* impact of adopting a poor reference frequency.

While we have chosen a single reference frequency, we have not optimized it, as we anticipate no single choice will work well for all parameters. The early part of the inspiral carries far more information about the chirp mass and mass ratio, as these both impact the total number of cycles. For example, to leading-order $\Gamma_{M_c M_c} \simeq \left\langle (\partial_{M_c}\Psi)^2 \right\rangle$ where the average $\langle\rangle$ corresponds to averaging in signal power (i.e., averaging over a distribution $\propto d\rho^2/df$). Evaluating the derivative and writing as an integral over frequency, we see the corresponding integrand to $d\rho^2/df$ in the definition of $\Gamma_{M_c M_c}$ is weighted by a polynomial in f that is *heavily* biased towards low frequencies:

$$\Gamma_{M_c M_c} \propto \int [f^{-10/3} + \text{lower powers}]d\rho^2/df \qquad (B3)$$

By contrast, spin effects enter at $v^3 \propto f$ past leading-order in phase, suggesting the spin-spin Fisher matrix components are less severely biased towards low frequency;

$$\Gamma_{SS} \propto \int f^{-10/3+2}d\rho^2/df \qquad (B4)$$

Other components of the Fisher matrix arise from other averages over frequency and have other break-even points. In other words, for each measurable quantity, a generally distinct epoch of the signal provides the most information. No one break-even point works for all.
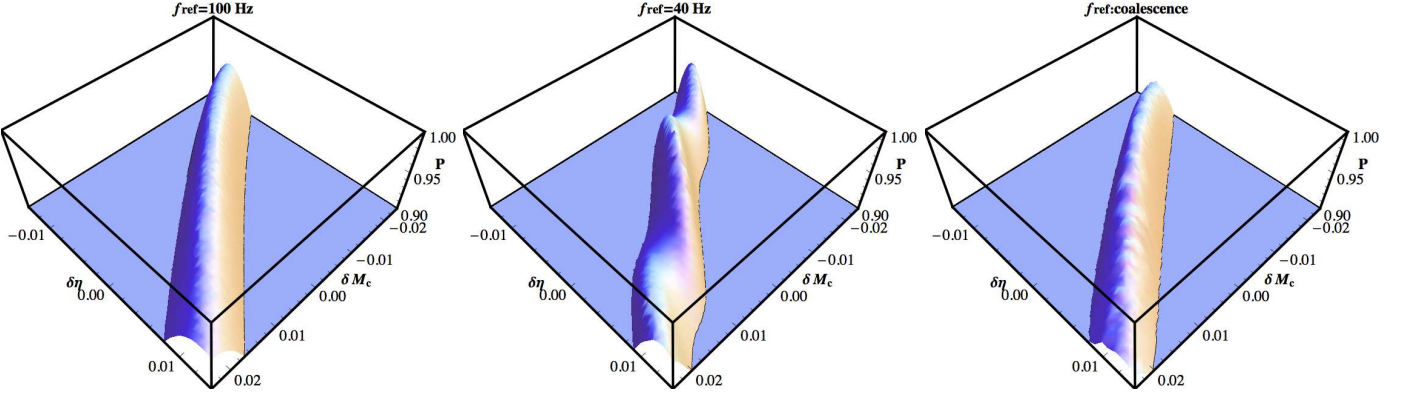
FIG. 10: **Examples for a dependence of the ambiguity function on the reference frequency using higher-order waveforms.**
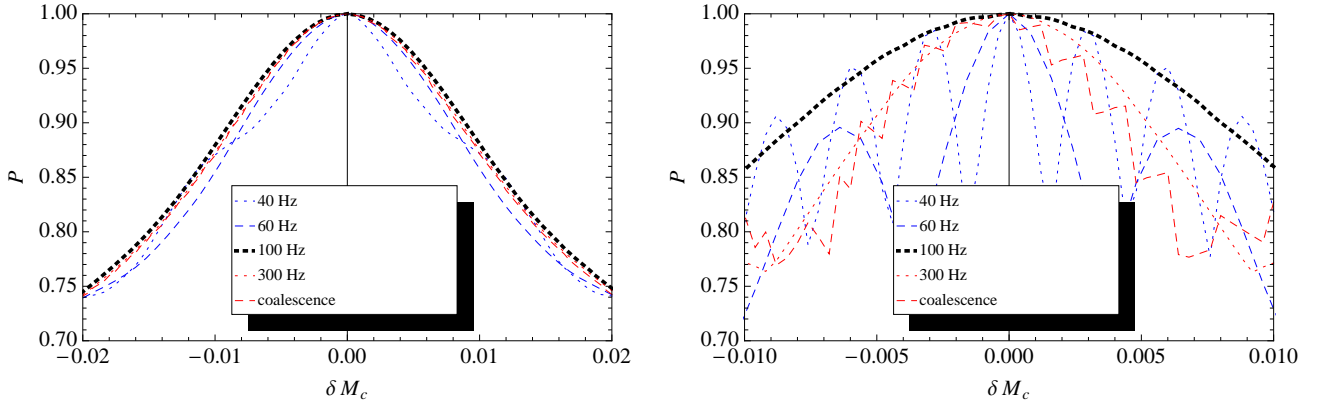

FIG. 11: **Examples for a dependence of the ambiguity function on the reference frequency.** Left: higher-order waveforms, $\iota = \pi/4$. Right: leading-order waveforms, $\iota = \pi/2$.
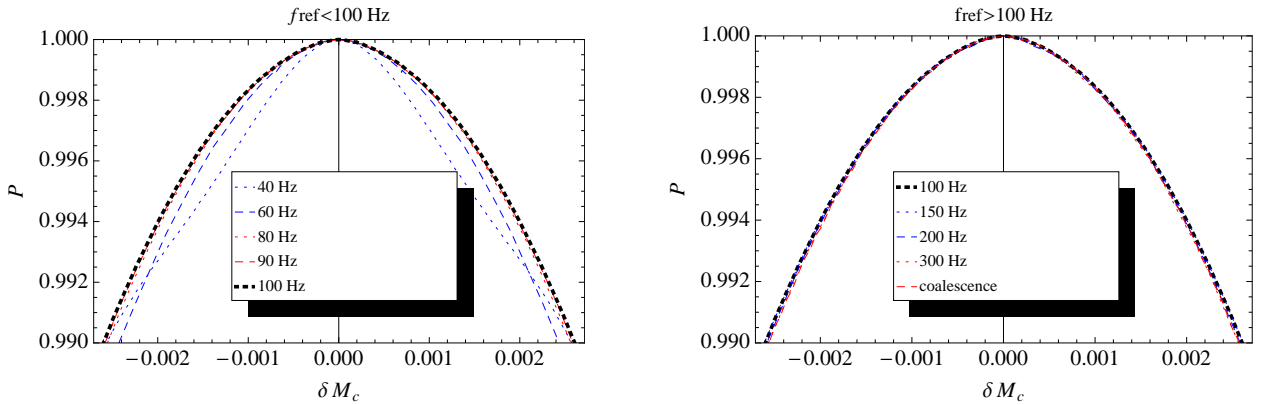

FIG. 12: **One dimensional ambiguity function with various reference frequencies for $\iota = \pi/4$, using leading-order waveforms.**

[1] Abbott et al. (The LIGO Scientific Collaboration). Nuclear Instruments and Methods in Physics Research A, **517**, 154 (2004).
[2] http://www.virgo.infn.it
[3] Abbott et al. (The LIGO Scientific Collaboration). Phys. Rev. D **77**, 062002 (2008).
[4] Abbott et al. (The LIGO Scientific Collaboration). Phys. Rev. D **79**, 122001 (2009).
[5] A. Buonanno, Y. Chen, and M. Vallisneri, Phys. Rev. D **67**, 104025 (2003).
[6] A. Buonanno, Y. Chen, Y. Pan, and M. Vallisneri, Phys. Rev. D **70**, 104003 (2004).
[7] T. Damour, A. Gopakumar, and B. R. Iyer, Phys. Rev. D **70**, 064028 (2004).
[8] Y. Pan, A. Buonanno, Y. Chen, and M. Vallisneri, Phys. Rev. D **69**, 104017 (2004).
[9] C. Königsdörffer and A. Gopakumar, Phys. Rev. D **71**, 024039 (2005).
[10] A. Buonanno, Y. Chen, Y. Pan, H. Tagoshi, and M. Vallisneri, Phys. Rev. D **72**, 084027 (2005).
[11] C. Königsdörffer and A. Gopakumar, Phys. Rev. D **73**, 124012 (2006).
[12] M. Tessmer and A. Gopakumar, MNRAS, **374**, 721 (2007).
[13] M. Hannam, S. Husa, B. Brügmann, and A. Gopakumar, Phys. Rev. D **78**, 104007 (2008).
[14] I. Hinder, F. Herrmann, P. Laguna, and D. Shoemaker, arXiv:0806.1037.
[15] K. G. Arun, A. Buonanno, G. Faye, and E. Ochsner, Phys. Rev. D **79**, 104023 (2009).
[16] A. Buonanno, B. R. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, Phys. Rev. D **80**, 084043 (2009).
[17] F. D. Ryan, Phys. Rev. D **52**, R3159 (1995).
[18] P. Jaranowski and A. Królak, Living Rev. Relativity **15**, 4 (2012).
[19] E. K. Porter and N. J. Cornish, Phys. Rev. D **78**, 064005 (2008).
[20] K. G. Arun, B. R. Iyer, B. S. Sathyaprakash, and S. Sinha, Phys. Rev. D **75**, 124002 (2007).
[21] C. Van den Broeck and A. S. Sengupta, Class. Quantum Grav. **24**, 1089 (2007).
[22] J. S. Key and N. J. Kornish, arXiv:1006.3759.
[23] R. N. Lang and S. A. Hughes, Phys. Rev. D **74**, 122001 (2006).
[24] A. Klein, P. Jetzer, and M. Sereno, Phys. Rev. D **80**, 064027 (2009).
[25] R. N. Lang, S. A. Hughes, and N. J. Cornish, Phys. Rev. D **84**, 022002 (2011).
[26] M. Vallisneri, Phys. Rev. D **77**, 042001 (2008).
[27] S. Vitale and M. Zanolin, Phys. Rev. D **82**, 124065 (2010).
[28] M. Vallisneri, Physical Review Letters **107**, 191104 (2011).
[29] https://www.lsc-group.phys.uwm.edu/daswg/projects/lal/nightly/docs/html/
[30] L. E. Kidder, Phys. Rev. D **52**, 821-847 (1995).
[31] C. M. Will and A. G. Wiseman, Phys. Rev. D **54**, 4813 (1996).
[32] D. A. Brown, A. Lundgren, and R. O'Shaughnessy, arXiv:1203.6060v2.
[33] R. O'Shaughnessy, J. Healy, L. London, Z. Meeks, and D. Shoemaker, Phys. Rev. D **85**, 084003 (2012).
[34] T. Damour, B. R. Iyer, and B. S. Sathyaprakash, Phys. Rev. D **63**, 044023 (2001).
[35] P. Ajith and S. Bose, Phys. Rev. D **79**, 084032 (2009).
[36] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown and J. D. E. Creighton, arXiv:gr-qc/0509116.
[37] L. S. Finn, Phys. Rev. D **46**, 5236 (1992).
[38] C. Cutler and M. Vallisneri, Phys. Rev. D **76**, 104018 (2007).
[39] P. Jaranowski and A. Królak, Phys. Rev. D **49**, 1723 (1994).
[40] E. Poisson and C. M. Will, Phys. Rev. D **52**, 848 (1995).
[41] C. Cutler and E. É. Flanagan, Phys. Rev. D **49**, 2658 (1994).
[42] R. O'Shaughnessy, B. Vaishnav, J. Healy, Z. Meeks, and D. Shoemaker, Phys. Rev. D **84**, 124002 (2011).