

# Testing the validity of the single-spin approximation in inspiral-merger-ringdown waveforms

Michael Pürrer,<sup>1</sup> Mark Hannam,<sup>1</sup> P. Ajith,<sup>2,3,4</sup> and Sascha Husa<sup>5</sup>

<sup>1</sup>*School of Physics and Astronomy, Cardiff University, Queens Building, CF24 3AA, Cardiff, United Kingdom*

<sup>2</sup>*International Centre for Theoretical Sciences, Tata Institute of Fundamental Research, IISc Campus, Bangalore 560012, India*

<sup>3</sup>*LIGO Laboratory, California Institute of Technology, Pasadena, CA 91125, USA*

<sup>4</sup>*Theoretical Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA*

<sup>5</sup>*Departament de Física, Universitat de les Illes Balears, Crta. Valldemossa km 7.5, E-07122 Palma, Spain*

Gravitational-wave signals from black-hole binaries with non-precessing spins are described by four parameters – each black hole’s mass and spin. It has been shown that the dominant spin effects can be modeled by a *single* spin parameter, leading to the development of several *three-parameter* waveform models. Previous studies indicate that these models should be adequate for gravitational-wave detection. In this paper we focus on the systematic biases that would result from using them to estimate binary parameters, and consider a one-parameter family of configurations at one choice of mass ratio and effective single spin. We find that for low-mass binaries within that family of configurations, where the observable waveform is dominated by the inspiral, the systematic bias in all physical parameters is smaller than the parameter uncertainty due to degeneracies between the mass ratio and the spins, at least up to signal-to-noise ratios (SNRs) of 50. For higher-mass binaries, where the merger and ringdown make a greater contribution to the observed signal, the bias in the mass ratio is comparable to its uncertainty at SNRs of only  $\sim 30$ , and the bias in the measurement of the total spin is *larger* than the uncertainty defined by the 90% confidence region even at an SNR of only 10. Although this bias may be mitigated in future models by a better choice of single-effective-spin parameter, these results suggest that it may be possible to accurately measure *both* black-hole spins in intermediate-mass binaries.

PACS numbers: 04.25.Dg, 04.25.Nx, 04.30.Db, 04.30.Tv

## I. INTRODUCTION

The inspiral and merger of black-hole and neutron-star binaries are the most promising sources for the first direct detection of gravitational waves (GWs) with the Advanced LIGO (aLIGO) and Virgo (AdV) detectors [1–3], and are expected to provide a wealth of astrophysical information (see, e.g., [4]). The optimal technique to locate their signals in the detector data is to cross-correlate the data against a large bank of theoretical signal templates. A search across the full parameter space of component masses, spins, sky locations, orientations and the distance is computationally extremely challenging. Partly for this reason searches in data from the initial LIGO and Virgo detectors [5–8] focussed on binaries with non-spinning components, for which a two-dimensional template bank suffices, greatly reducing the computational cost. The two dimensions are defined by combinations of the component masses, and the effects of the sky-location, orientation and distance on the signal observed by a single detector can be absorbed into an overall amplitude scale factor.

Such simplifications are not possible for generic spinning binaries, where the components’ spins cause precession of the orbital plane and of the spins themselves, leading to far more complex GW signals. However, if we consider only spins aligned/anti-aligned to the binary’s orbital angular momentum, then the only spin effects are on the inspiral rate and the signal amplitude – the basic waveform structure is unchanged from the non-spinning case. Including the aligned/anti-aligned spin effects in the waveform templates makes it possible to detect a much larger volume of the binary parameter space, including in some cases a significant fraction of precessing binaries [9–11]. It is also possible that non-precessing-binary models can be used as the basis for con-

structing generic waveform models [12, 13]. Note that this study analyzes only the  $l = 2, m = \pm 2$  modes of the gravitational wave signal.

The inclusion of the (non-precessing) black-hole spins doubles the dimensionality of the search parameter space over non-spinning searches. However, studies of inspiral dynamics using post-Newtonian expansions, and of merger and ringdown with numerical solutions of Einstein’s equations, show that the dominant spin effects can be modeled with a single parameter [9, 10, 14]. This has motivated the development of waveform models parametrized by only the binary’s mass ratio and effective total spin (the binary’s total mass appears as a simple overall scale factor) [9, 10, 14–16]. Also, recent work on implementing non-precessing-spin template banks has exploited the partial degeneracy between the two spins [11, 17].

The use of a single effective spin parameter is also motivated by the high computational cost of fully general relativistic numerical simulations. Work to date on phenomenological waveform models suggests that we require at least four simulations in each direction of parameter space that we wish to model. A model of the full seven-dimensional parameter space of generic binary waveforms would require  $4^7 \approx 16,000$  simulations, which are not feasible before the commissioning of aLIGO and AdV [9, 14, 18–21]. The most ambitious study to date includes “only”  $\sim 200$  waveforms, at moderate mass ratios and black-hole spins [22]. It is therefore important that we exploit any degeneracies that reduce the dimensionality of the parameter space that we must model.

While single-effective-spin models are believed to capture the phenomenology of non-precessing-binary signals with sufficient fidelity for GW detection, little is known about how well they would perform if used to estimate the source parameters following a detection. The single-spin approximation is

only valid in the leading-order post-Newtonian spin terms (although it holds to higher order when both masses are equal, and for extreme mass ratios, where the influence of the smaller black hole’s spin is negligible), and does not hold through merger, where the appropriate single spin becomes the total spin angular momentum of the two black holes [23–26]. If we use a single-effective-spin waveform model for parameter estimation, what will be the bias in the measurement of the black-hole masses, and of the spin parameter itself? Obviously, if we approximate the two black-hole spins with a single spin, then we cannot use this model to measure the individual black-hole spins; on the other hand, if a single spin parameter models the dominant spin effects, then both spins will be difficult to measure even if we *did* use a double-spin model. We will return to this point later.

In this paper we explore the parameter biases due to the use of a single-effective-spin model. Since we expect the single-effective-spin approximation to become less valid for higher mass ratios, we consider a set of configurations at the highest mass ratio of the numerical simulations that were used to calibrate current phenomenological models, 1:4 [9, 14]. The computational cost of numerical simulations precludes an exhaustive study, so we focus on one value of the effective spin parameter,  $\chi_{\text{IMR}} := (m_1\chi_1 + m_2\chi_2)/(m_1 + m_2) = 0.45$ . (Here  $m_1$  and  $m_2$  are the masses, and  $\chi_1$  and  $\chi_2$  are the Kerr parameters of the black holes). We produce a set of five simulations with differing values of the individual black-hole spins ( $\chi_1$  and  $\chi_2$ ), but with the same value of  $\chi_{\text{IMR}}$ . From the numerical-relativity waveforms we construct hybrid PN-NR waveforms, which are in turn compared against one of the phenomenological models, “IMRPhenomC” (see Sec. II B for a more detailed description of the waveform model). By identifying the IMRPhenomC waveform that agrees best with each hybrid, we estimate the parameter biases due to the use of a single-effective-spin model.

There are a number of issues that make it difficult to draw conclusions from this procedure. The results will be skewed by artifacts in the construction of the particular waveform model that we use (the details of the phenomenological ansatz, the coverage of the parameter space by numerical waveforms, and the accuracy of the waveforms), which may swamp the errors due to the single-effective-spin approximation. Previous studies have shown that the main source of uncertainty in hybrid PN-NR waveforms is in the PN regime [27–30], and as such our results will depend on the PN approximant we use in our hybrids, and on the hybridization frequency. We discuss these issues further, and the steps we have taken to mitigate them, in Sec. V.

The layout of the paper is as follows. We summarize the single-spin approximation, waveform models and our numerical waveforms in Secs. II and III. In Sec. IV we make a preliminary study of biases in the inspiral regime, where we can compare single- and double-spin PN models using the same PN approximant, and do not have to concern ourselves with issues of hybridization or phenomenological modeling. In addition to quantifying the parameter biases due to the single-effective-spin approximation for low-mass binaries (for mass-ratio 1:4 and moderate spins), this section also provides con-

text and contrast to the full inspiral-merger-ringdown results, which are in Sec. V.

## II. PRELIMINARIES

### A. The single-spin approximation

We consider black-hole binaries where the spins are aligned or anti-aligned with the orbital angular momentum. Then the spins and the angular momentum do not precess, which leads to a considerable simplification of the GW signal over generic configurations. These aligned-spin waveforms are parametrized by the black-hole masses and spins.

A single effective spin parameter  $\chi_{\text{IMR}} := (m_1\chi_1 + m_2\chi_2)/M$  has been used in the construction of the non-precessing phenomenological inspiral-merger-ringdown (IMR) models presented in Refs. [9, 14]. These models parametrize the waveforms by their mass  $M$ , symmetric mass ratio  $\eta = m_1m_2/M^2$ , and the effective spin parameter  $\chi_{\text{IMR}}$ . They incorporate a PN description of the inspiral, while the merger and ringdown regimes are tuned using the results of numerical simulations. A recent study [10] has addressed how well a related “reduced spin” parameter motivated by PN theory works for inspiral searches. This PN model has been shown to be sufficiently accurate for GW searches (“effectual”), and to agree well with the full two-spin waveforms (“faithful”) when either the spins or the masses are equal.

In constructing the PN reduced-spin parameter, we note that all spin effects can be described by two parameters  $\chi_s \equiv \chi_s \cdot \hat{\mathbf{L}}_N \equiv (\chi_1 + \chi_2)/2$  and  $\chi_a \equiv \chi_a \cdot \hat{\mathbf{L}}_N \equiv (\chi_1 - \chi_2)/2$ , which remain constant throughout the evolution. The dimensionless spin parameters  $\chi_i$  are defined as  $\chi_i = S_i/m_i^2$ , where  $S_i$  is the spin of black hole  $i$ . The leading order spin term due to spin-orbit coupling appearing at 1.5PN order in the amplitude and the phase can be represented by a single “reduced spin” parameter (see e.g. [10, 31])

$$\chi_{\text{PN}} \equiv \chi_s + \delta\chi_a - \frac{76\eta}{113}\chi_s, \quad (2.1)$$

where  $\eta = m_1m_2/M^2$  and  $M = m_1 + m_2$ .

In contrast, the “effective spin” parameter used in the phenomenological models for black-hole binaries with non-precessing spins [9, 14] is defined as a simple mass-weighted linear combination of the spins

$$\chi_{\text{IMR}} \equiv (m_1\chi_1 + m_2\chi_2)/M = \chi_s + \delta\chi_a. \quad (2.2)$$

For equal masses both spin parameters are a function of the symmetric combination of the spins  $\chi_s$  only. The only difference is an overall factor. For unequal masses both spin parameters depend on the symmetric and anti-symmetric spin combinations. The difference between the spin parameters depends linearly on  $\eta$  and therefore goes to zero for infinite mass-ratio.

Historically, the non-precessing phenomenological IMR models to date have used the effective spin parameter  $\chi_{\text{IMR}}$  defined in Eq. (2.2) due to its simple form. While this choice was

sufficient to build effectual non-precessing waveform models, we will present evidence that suggests that  $\chi_{\text{PN}}$  is a better choice and should be used for future models.

### B. Phenomenological single-spin models

We can quantify the agreement between families of waveforms with the same value of  $\chi_{\text{PN}}$  or  $\chi_{\text{IMR}}$ , but in order to estimate the parameter bias that would result from the single-spin approximation, individual waveforms are not sufficient; we require a waveform family. In this study we compare our PN/NR hybrids with the phenomenological model for black-hole binaries with non-precessing spins presented in Ref. [14]. For consistency with the labeling used within the LIGO-Virgo Collaboration [32] we refer to this model as “IMRPhenomC”. (“IMRPhenomA” refers to a model of non-spinning binaries [18–20], and “IMRPhenomB” to an earlier model of non-precessing binaries [9]; we choose to use “IMRPhenomC” because it incorporates higher-order PN information in the inspiral phasing, but also make cross-checks against the IMRPhenomB model.)

The model waveforms are parametrized by their total mass  $M = m_1 + m_2$ , symmetric mass ratio  $\eta = m_1 m_2 / M^2$ , and the effective total spin parameter  $\chi_{\text{IMR}}$  defined in Eq. (2.2). The waveform is represented in the Fourier domain as  $h(f) = A(f)e^{i\Psi(f)}$ . The amplitude  $A(f)$  and phase  $\Psi(f)$  are modeled separately. The IMRPhenomC amplitude is constructed from two parts: a PN inspiral amplitude with the addition of a higher order frequency term, and a ringdown portion, both of which are fit to the model hybrids. For the inspiral portion of the phase IMRPhenomC uses the complete TaylorF2 [33–36] PN inspiral phasing (up to 3.5PN order, although the spin terms are complete only up to 2.5PN). Only the late inspiral/merger phase is fitted in a narrow frequency range  $[0.1f_{\text{RD}}, f_{\text{RD}}]$  to numerical simulations, while the ringdown waveform is obtained from analytically derived quasinormal mode expressions for the frequency and attached continuously to the merger phase. For both the amplitude and the phase smooth tanh-window functions are used to connect the individual parts.

The model is a power series in the frequency  $f$ , and the coefficients in the model are written as polynomials in the two physical parameters  $\eta$  and  $\chi_{\text{IMR}}$  (the total mass is an overall scale factor), and it is the coefficients of these polynomials that are then calibrated to hybrids of PN and NR waveforms. There are 45 free parameters in IMRPhenomC, although the final model is a function of only  $\{M, \eta, \chi_{\text{IMR}}\}$ . The hybrids used to construct IMRPhenomC were produced in the frequency domain, using TaylorF2 for the PN part and a rather broad fitting window  $Mf \in [0.01, 0.02]$ . The construction of frequency-domain PN/NR hybrids is discussed in more detail in Sec. III A.

### C. Matches, fitting factors and confidence regions

We quantify the agreement between two waveforms,  $h_1(f)$  and  $h_2(f)$ , with the standard inner product weighted by the power spectral density  $S_n(f)$  of a detector [37], called the *overlap*

$$\langle h_1 | h_2 \rangle = 4 \text{Re} \int_{f_{\min}}^{f_{\max}} \frac{\tilde{h}_1(f) \tilde{h}_2^*(f)}{S_n(f)} df. \quad (2.3)$$

The inner product is calculated in terms of the frequency-domain waveforms  $\tilde{h}(f)$ . The frequency range in which the detector is deemed sensitive is  $[f_{\min}, f_{\max}]$ . Let  $\hat{h}(f) \equiv \tilde{h}(f) / \sqrt{\langle \tilde{h} | \tilde{h} \rangle}$  be the normalized frequency-domain waveform. The match between two normalized waveforms is then defined as their inner product, maximized over time and phase shifts of the waveform,

$$M(h_1, h_2) = \max_{\Delta t, \Delta \phi} \langle \hat{h}_1 | \hat{h}_2 \rangle. \quad (2.4)$$

Given a signal waveform  $h(\Lambda)$  with physical parameters  $\Lambda$  and a template  $x(\Lambda)$  with physical parameters  $\Lambda$  we define the *fitting factor*

$$\text{FF} = \max_{\Delta t, \Delta \phi, \Lambda} \langle \hat{x}(\Lambda) | \hat{h}(\Lambda) \rangle. \quad (2.5)$$

Instead of the fitting factor we will often quote the fully optimized mismatch

$$\mathcal{M} = 1 - \text{FF}. \quad (2.6)$$

The match quantifies the physical agreement between two waveforms (since the time of arrival and overall phase of the waveform do not change the underlying physics of the binary). The fitting factor is a measure of how well a matched-filter search with a given waveform family can perform in detecting a particular signal; a fitting factor greater than 0.965 indicates that no more than 10% of signals will be lost in a search. It does not tell us, however, how well the parameters of the best-match template will agree with the true source parameters of the signal.

A PN-NR hybrid binary waveform that has been produced for a given total mass  $M$  can be trivially rescaled to a different mass. Therefore, the match between two such waveforms can be optimized over the total mass. The phenomenological model “IMRPhenomC” used in this study depends in addition on the symmetric mass-ratio  $\eta$  and the effective spin  $\chi_{\text{IMR}}$ , which allows us to compute fitting factors by optimizing matches over  $\Lambda = \{M, \eta, \chi\}$ .

In this paper we compare PN-NR hybrids signals with “IMRPhenomC” with reference to the expected sensitivity of the Advanced LIGO detector [38–40]. Early science runs are expected around 2015 [41]. At its optimum sensitivity several years later, the anticipated sensitivity is given by the “zero-detuned high-power” noise curve [42]. We use a linear interpolation of this expected PSD and choose  $f_{\min} = 15$  Hz, and  $f_{\max} = 8$  kHz.

For PN matches we choose the upper frequency of the overlap integral as the frequency of the innermost stable circular

orbit (ISCO) of a test particle around a Schwarzschild black hole  $f_{\text{ISCO}} = v_{\text{ISCO}}^3/(\pi M)$ , where  $v_{\text{ISCO}} = 1/\sqrt{6}$  just as in [10]. The Schwarzschild ISCO is an arbitrary point at which to terminate the PN waveform, but it corresponds to the choice commonly made in detector searches [43].

The model parameters for the waveform that best matches the signal correspond to the parameters that are most likely to be recovered in a GW measurement. We are also interested in the range of parameters that would be recovered in 90% of observations at a given SNR, i.e., the 90% confidence region for that SNR, which illustrates the statistical uncertainty in the measurement.

At high SNRs the confidence region can be estimated by Fisher-matrix methods [31, 34, 44, 45], while in general one should construct the full posterior probability distribution function [46–50]. The latter is computationally very expensive, but Ref [51] shows that it is possible to produce a good approximation to the correct confidence region by computing matches between the model waveform with the physical parameters of the signal, and model waveforms with a range of neighboring parameters. All neighbouring waveforms that have a match greater than some threshold are within the 90% confidence region. The threshold for a given SNR  $\rho$  assuming a 3-dimensional parameter space is [51]

$$M(h_m(\theta), h_m(\theta_0)) \geq 1 - 3.12/\rho^2, \quad (2.7)$$

where  $\theta$  are arbitrary waveform parameters (in this case  $M$ ,  $\eta$ , and  $\chi_{\text{PN}}$  or  $\chi_{\text{IMR}}$ ), and  $\theta_0$  are the correct parameters, and  $h_m(\theta)$  are the model waveforms.

We have computed fitting factors and the associated best parameters with two different methods. The Nelder-Mead Amoeba [52] simplex method has been used to compute fitting factors for a range of masses (see Sec. V B). For selected masses we have computed 90% confidence regions by sampling the matches on a suitably fine grid in  $(M_c, \eta, \chi)$  space. The latter computation is a lot more expensive, but more reliable — in some cases the amoeba calculation can be trapped in a local minimum, especially when the confidence region in question is not simply connected. At low masses and high SNRs confidence regions can be very elongated filaments and a transformation to rotate and squash the region into a more compact form is then helpful to keep the computation within a reasonable cost. This is related to the alternative parameter-space coordinates that are being used in placing waveforms in search template banks [11, 53–55].

### III. NUMERICAL WAVEFORMS

To fully test the single-spin approximation across the binary parameter space, we would need to perform, for each of a wide range of choices of  $\eta$  and  $\chi_{\text{PN}}$  (or  $\chi_{\text{IMR}}$ ), a series of simulations for choices of different black-hole spins that correspond to the same value of  $\chi_{\text{PN}}$  ( $\chi_{\text{IMR}}$ ). In doing so, we would have produced enough simulations to construct a complete two-spin waveform model — but the high computational cost of doing so is one of the motivations for producing a single-effective-spin model in the first place!

We expect the single-effective-spin approximation to become less accurate as the binary mass ratio increases, and so in this study we focus on the largest mass ratio that was considered in the numerical simulations used to calibrate current phenomenological models [9, 14],  $q = m_2/m_1 = 4$ . We choose an effective total spin of  $\chi_{\text{IMR}} = 0.45$ ; this is a relatively large total effective spin for which we can also choose a wide range of individual black-hole spins.

In our simulations we set the total mass  $M = 1$  and using the convention  $m_1 < m_2$  have  $m_1 = 0.2$  and  $m_2 = 0.8$ . We then have  $\chi_{\text{IMR}} = 0.2\chi_1 + 0.8\chi_2$ . With our choice of  $\chi_{\text{IMR}} = 0.45$  we let  $\chi_1$  (the spin of the smaller BH), vary between  $-0.75$  and  $+0.75$ . (Due to the large junk-radiation content in Bowen-York initial data for highly spinning binaries, we do not consider spins higher than  $0.75$ .) Along this line of  $\chi_{\text{IMR}} = 0.45$  in the  $(\chi_1, \chi_2)$  plane we pick configurations for  $\chi_1 = -0.75, -0.25, 0.25, 0.75$ . In addition we also add the configuration with equal spins  $\chi_1 = \chi_2 = 0.45$ . The latter configuration is important since IMRPhenomC [14] assumes equal spins for its PN part. Our chosen configurations are summarized in table I (also see Fig. 1).

Run	$\chi_1$	$\chi_2$	$m_1/h_{\text{min}}$	Cycles	$D/M$	$e_{\phi, \text{GW}}$	$M_f$	$a_f/M$
1	-0.75	0.75	44.4	29	10.739	0.0003	0.966	0.84
2	-0.25	0.625	38.5	28	10.782	0.0006	0.969	0.79
3	0.25	0.5	38.5	28	10.831	0.0007	0.971	0.74
4	0.45	0.45	38.5	28	10.853	0.0027	0.973	0.72
5	0.75	0.375	44.4	28	10.889	0.0014	0.972	0.68

TABLE I: The series of  $q = 4$ ,  $\chi_{\text{IMR}} = 0.45$  configurations used in this study (also see Fig. 1). We show the spin parameters of the individual black holes, the resolution on the finest grid level with respect to the smallest black hole,  $m_1/h_{\text{min}}$ , the number of GW cycles before the amplitude maximum at merger, the initial separation  $D/M$ , the eccentricity  $e_{\phi, \text{GW}}$  measured from the GW phase [56] and the final mass and spin. For the equal-spin configuration only one step of eccentricity reduction was performed.

The numerical simulations were performed with the BAM code [57, 58], which evolves black-hole-binary puncture initial data [59, 60] (generated using a pseudo-spectral elliptic solver [61]), and evolves them with the  $\chi$ -variant of the moving-puncture [62–64] version of the BSSN [65, 66] formulation of the 3+1 Einstein evolution equations. Spatial finite-difference derivatives are sixth-order accurate in the bulk [58], Kreiss-Oliger dissipation terms converge at fifth order, and a fourth-order Runge-Kutta algorithm is used for the time evolution. The gravitational waves emitted by the binary are calculated from the Newman-Penrose scalar  $\Psi_4$ , and the details of our implementation of this procedure are given in [57].

The basic grid setup for this study is based on a setup used for a convergence series of  $q = 4$ ,  $\chi_1 = \chi_2 = 0.75$  simulations. In general, the mass-ratio has the most significant impact on choosing a grid setup, since the apparent horizons (AH) of the black holes must be resolved by the finest grid level. We have found that choosing the size of the innermost box about a factor  $1.5\times$  the size of the AH of the smaller BH leads to good accuracy. Since the spin of the smaller BH varies in

the five configurations considered in this study, the individual grid setup is tuned differently for each simulation. We found that sufficient accuracy could be achieved by using a minimum of 16 buffer zones (rather than the formal requirement of 32) between mesh-refinement boxes. The Courant factor has been reduced to  $C = 0.4$  (from  $C = 0.5$  in our previous work) to curb the contribution of time-integration error to the overall NR error. A detailed study of the accuracy of these simulations, and the effects of the errors due to both the spatial finite-differencing and the timestepping, will be presented in a forthcoming paper [67].

We have performed two iterations of the eccentricity reduction method detailed in Ref. [56] for each of the configurations. The final eccentricities are below 0.0015. The numerical waveforms are obtained by extracting the  $l = m = 2$  mode of  $\psi_4$  at  $r = 180M$ .

While we have not performed convergence tests for the NR waveforms, we expect that the results are robust because we have used a “safe” grid setup that has led to accurate results for related aligned-spin configurations with  $q=4$  and  $\chi_1 = \chi_2 = 0.75$ . In Sec. VB we show how the numerical resolution affects fitting factors and biases of the (0.45, 0.45) waveform with IMRPhenomC, and find that it does not affect our overall results.

#### A. Construction and accuracy of PN-NR hybrid waveforms

The NR waveforms we have computed for this study do not cover the full aLIGO sensitivity band for lower-mass binaries. Therefore we need to hybridize the NR waveform with a PN approximant. Different choices are possible for constructing such PN-NR hybrids (see Ref. [36] for a summary). We wish to compare our IMR hybrids with the IMRPhenomC model, in which the inspiral is modeled by the TaylorF2 frequency-domain PN approximant. In order to minimize effects arising from differing PN approximants, we choose to create hybrids with TaylorF2.

We use a frequency domain hybridization method as described in [14]. We include the NR  $\psi_4$  waveform data from the time immediately after the passage of the burst of initial junk radiation, up to the point where the ringdown is dominated by numerical noise, and apply a Planck tapering window [68] of width 300 M at the start of the dataset. The waveform is further padded with zeroes before computing the FFT to increase the frequency resolution. The NR strain is calculated in the Fourier domain as  $\tilde{h}_{\text{NR}}(f) = \tilde{\psi}_4^{\text{NR}}(f)/(2\pi f)^2$ . The matching procedure aligns  $\tilde{h}_{\text{PN}}(f; t_0, \phi_0) = \tilde{h}_{\text{F2}}(f) + 2\pi f t_0 + \phi_0$  and  $\tilde{h}_{\text{NR}}(f)$  by a least squares fit over a fitting interval  $[f_1, f_2]$ , which we discuss in more detail below. We then determine the matching frequency  $f_m \in [f_1, f_2]$  at which the NR and PN phases coincide by a root-finding algorithm. The PN and NR amplitudes are aligned separately without any freedom to adjust parameters.

After having settled on an approximant and choice of hybridization method we can still choose the frequency region over which the hybridization is performed. Intuitively it makes sense to hybridize at as low a frequency as possible so

as to extract as much useful information from the NR data as we can, and to minimize errors in the PN approximant, which increase with GW frequency. It also stands to reason that the matching interval  $[f_1, f_2]$  should not be too narrow as the fit would then be prone to pick up spurious oscillations in the NR data. For the hybrids considered in this study we have chosen an interval length of  $M\Delta\omega \sim 0.01$  and a matching frequency of about  $M\omega_m \sim 0.07$ . This results in a relative matching width of  $\Delta\omega/\omega_m \sim 0.14$  and is consistent with the choice advocated in Ref. [28].

One accessible measure to gauge the quality of a hybrid are the parameter errors given by the least squares fit. At first glance a high quality fit appears to be very desirable. However, the PN model used in the fit gives a worse approximation of the NR phase as we go to higher frequencies and thus small standard errors in the fitting parameters do not necessarily imply that the hybrid will be very faithful.

In a practical sense it is useful to think of the difference between hybrids with varying hybridization regions (assuming reasonable comparison interval, i.e. not unreasonably high frequencies) as a way of quantifying the error in the hybrid caused by the PN and NR data. We would like to know how such variations in the hybrid construction manifest themselves in biases and uncertainties. In Sec. VB we verify that our results are robust with respect to hybridization artifacts.

## IV. RESULTS FOR PN SINGLE-EFFECTIVE-SPIN MODELS

A single-effective-spin PN model has recently [10] been shown to be an effective search template (i.e., fitting factors  $> 0.97$ ) as well as “faithful” to two-spin signals (i.e., non-optimized matches are also  $> 0.97$ ), when either the spins or the masses are equal. Here we address the question of biases and uncertainties incurred by the single-effective-spin approximation. While there is a preferred “reduced-spin” parameter  $\chi_{\text{PN}}$  in the PN regime, we also generalize the model used in Ref. [10] to arbitrary definitions of an effective spin parameter and compare with a model built from  $\chi_{\text{IMR}}$ , the effective spin parameter used by current phenomenological waveform models.

The construction of a single-effective-spin model is straightforward. We choose one based on the TaylorF2 approximant. The model is based on the mapping  $(\chi_1, \chi_2, \eta; f) \mapsto (\chi_{\text{eff}}, \eta; f)$ , where  $\chi_{\text{eff}} = \chi_{\text{eff}}(\chi_1, \chi_2, \eta)$  is an arbitrary effective spin parameter. To build the model an inverse of this mapping is needed which requires a relation between  $\chi_1$  and  $\chi_2$ . We choose to use only the symmetric part of the input spins (i.e., setting  $\chi_a = 0$ ) and define the frequency domain single spin model strain as

$$\tilde{h}_{\text{M}}(\chi_{\text{eff}}(\chi_1, \chi_2, \eta), \eta; f) := \tilde{h}_{\text{F2}}(\chi_s, \chi_s, \eta; f). \quad (4.1)$$

With this definition the model represents equal spin configurations exactly. For the choice  $\chi_{\text{eff}} = \chi_{\text{PN}}$  this model is identical to the one defined in Ref. [10].

We choose the following mass-ratio  $q = 4$  configurations for the comparison of single spin PN models (see Fig. 1). For

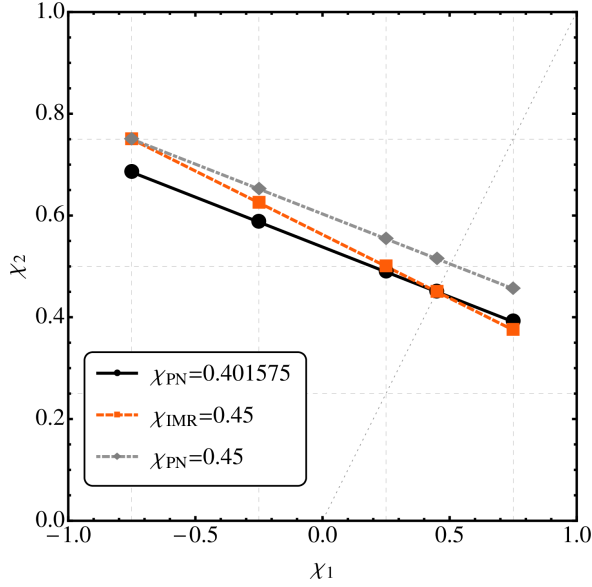


FIG. 1: Configurations on lines of constant “reduced” spin parameter  $\chi_{\text{PN}}$  and “effective” spin parameter  $\chi_{\text{IMR}}$  chosen in this study for mass-ratio  $q = 4$ . Two of these configurations lie on the line  $\chi_1 = \chi_2$  (thin dotted line).

the  $\chi_{\text{IMR}}$  model we select the same cases along  $\chi_{\text{IMR}} = 0.45$  (orange, dashed) as used for NR simulations. The configurations for the  $\chi_{\text{PN}}$  model are chosen along a line  $\chi_{\text{PN}} = 0.401575$  (black solid), rather than  $\chi_{\text{PN}} = 0.45$  (gray, dot-dashed). The reason is that  $\chi_{\text{PN}} = 0.401575$  intersects  $\chi_{\text{IMR}} = 0.45$  at the equal spin  $(\chi_1, \chi_2) = (0.45, 0.45)$  configuration, for which both models are exact. For  $\chi = 0.45$  and  $\chi_1 = -0.75$  the symmetric combination of the spins  $\chi_s$  vanishes and thus  $\chi_{\text{IMR}} = \chi_{\text{PN}}$  at this point. The largest deviation between the spin parameters happens at the largest positive spin of the smaller BH  $\chi_1 = 0.75$  and is about 20%.

It is well known that  $\chi_{\text{PN}}$  is the (almost) optimal single-effective-spin parameter in the PN regime (this combination appears explicitly in the leading-order spin-orbit coupling) [10, 31]. The superiority of  $\chi_{\text{PN}}$  over  $\chi_{\text{IMR}}$  is illustrated in Fig. 2 by how quickly matches between a single spin model (based on either  $\chi_{\text{PN}}$  or  $\chi_{\text{IMR}}$ ) and TaylorF2 signals (again at constant  $\chi_{\text{PN}}$  or  $\chi_{\text{IMR}}$ , respectively) degrade when one moves away from the point  $\chi_1 = \chi_2$ , where the models are exact, along a line of the respective  $\chi_{\text{eff}} = \text{const}$ .

As an example we compute fully optimized matches (fitting factors) and parameter biases for TaylorF2 signal waveforms chosen as above with each of the single spin models for a system mass of  $M = 7M_\odot$ . To quantify the statistical uncertainty we also calculate 90% confidence regions in the 3-dimensional space of model parameters  $(\mathcal{M}_c, \eta, \chi)$ .

From an astrophysical point of view, a compact binary with  $M = 7M_\odot$  could correspond to an actual NS-BH binary with component masses (1.4, 5.6). The NS would be expected to have very small spin which is contrary to some of the configurations chosen here. However, our goal is to compare how well the single spin approximation works in the PN and IMR regimes (see Sec. V B) and therefore we choose the same

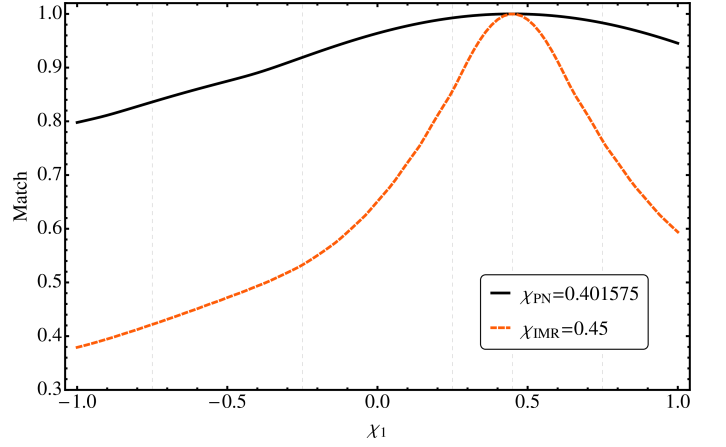


FIG. 2: Matches of TaylorF2 signal waveforms along lines of  $\chi_{\text{eff}} = \text{const}$  with single spin PN models with parameters  $\chi_{\text{IMR}} = 0.45$  or  $\chi_{\text{PN}} = 0.401575$  at  $7M_\odot$ . The signal waveforms lie on lines  $\chi_{\text{IMR}} = 0.45$  or  $\chi_{\text{PN}} = 0.401575$ , respectively.

range of spin values for both PN and IMR models.

The results are summarized in Tabs. II and III. The fully optimized mismatches (see Eq. (2.6)) have been computed from 90% confidence regions at SNR 50 which will be discussed later. We have performed a least-squares fit to the elongated direction of the filament-like confidence regions and subsequently carried out a local minimization starting from the best match found along the curve fit. Although the knowledge of the confidence regions is not needed to compute fitting factors this method leads to more reliable results than simpler optimization methods.

We define biases as  $\Delta\Lambda := \Lambda_{\text{recovered}} - \Lambda_{\text{true}}$ , where  $\Lambda$  is one of the model parameters  $(\mathcal{M}_c, \eta, \chi)$ . Note that in tabs. II and III we have dropped the subscript PN/IMR in the relative bias of the spin parameters. For ease of comparison we also give the absolute bias in the spin.

Case $(\chi_1, \chi_2)$	Mismatch $\mathcal{M}$	$\Delta\mathcal{M}_c/\mathcal{M}_c[\%]$	$\Delta\eta/\eta[\%]$	$\Delta\chi/\chi[\%]$	$\Delta\chi_{\text{PN}}$
(-0.75, 0.685104)	$6 \times 10^{-4}$	0.07	-13.7	13.4	0.054
(-0.25, 0.587144)	$2 \times 10^{-4}$	0.04	-8.8	8.5	0.034
(0.25, 0.489184)	$2 \times 10^{-5}$	0.01	-2.8	2.7	0.011
(0.45, 0.45)	0	0	0	0	0
(0.75, 0.391224)	$4 \times 10^{-5}$	-0.01	1.3	-2.0	-0.008

TABLE II: Fully optimized mismatches and biases between single spin PN model using  $\chi_{\text{PN}}$  and TaylorF2 signals along  $\chi_{\text{PN}} = 0.401575$  for  $7M_\odot$ .

Case $(\chi_1, \chi_2)$	Mismatch $\mathcal{M}$	$\Delta\mathcal{M}_c/\mathcal{M}_c[\%]$	$\Delta\eta/\eta[\%]$	$\Delta\chi/\chi[\%]$	$\Delta\chi_{\text{IMR}}$
(-0.75, 0.75)	$5 \times 10^{-4}$	0.09	-20.5	25.2	0.113
(-0.25, 0.625)	$2 \times 10^{-4}$	0.06	-14.1	16.4	0.074
(0.25, 0.5)	$3 \times 10^{-5}$	0	1.1	2.2	0.01
(0.45, 0.45)	0	0	0	0	0
(0.75, 0.375)	$5 \times 10^{-5}$	0	-0.5	-3.9	-0.018

TABLE III: Fully optimized mismatches and biases between single spin PN model using  $\chi_{\text{IMR}}$  and TaylorF2 signals along  $\chi_{\text{IMR}} = 0.45$  for  $7M_\odot$ .

The fitting factors for both single-effective-spin models are very high, with the fully optimized mismatch below 0.1%. The models are exact at the equal-spin configuration and therefore the true parameters are recovered. As we move away from the equal-spin configuration the mismatch becomes larger and the parameter biases increase. The bias in the chirp mass  $\mathcal{M}_c = M\eta^{3/5}$  is overall very small, below 0.1%, consistent with standard results. This is expected since the leading factor in the PN phase evolution for non-precessing binaries is proportional to  $1/(\mathcal{M}_c\pi f)^{5/3}$ , which is dominated by the chirp mass [44, 45, 51]. In contrast, there is considerable bias in the spin parameter and symmetric mass-ratio  $\eta$  for the very unequal spin configurations. For the  $\chi_{\text{PN}}$ -model the modulus of the biases in  $\eta$  and  $\chi$  increases to about 15% for the configuration with  $\chi_1 = -0.75$ , which is the farthest from the equal-spin case. The biases are worse for the  $\chi_{\text{IMR}}$ -model and reach about 23% for the  $\chi_1 = -0.75$  configuration. The absolute spin bias is at most  $\Delta\chi \sim 0.05$  for  $\chi_{\text{PN}}$ , while it rises to twice that value for  $\chi_{\text{IMR}}$ .

We know from PN theory that  $\chi_{\text{PN}}$  provides a better single-effective-spin approximation at low masses, but these results quantify the difference in parameter biases in using either  $\chi_{\text{PN}}$  or  $\chi_{\text{IMR}}$ . The bias resulting from the use of  $\chi_{\text{IMR}}$  is up to twice as large as that due to using a model parametrized by  $\chi_{\text{PN}}$ .

At first glance, the results in Tabs. II and III suggest that the single-effective-spin approximation is entirely inappropriate for parameter estimation: the uncertainty in the mass ratio and the spin parameter can be as high as  $\sim 15\%$ , even if we use the  $\chi_{\text{PN}}$  approximation. This in turn suggests that, if a single-effective-spin approximation behaves poorly, then we may be able to accurately resolve the *individual* black-hole spins with a complete two-spin model. Before making this conclusion, we should consider the statistical uncertainty in the parameter measurement for likely aLIGO and AdV SNRs.

The non-detection of signals in first-generation detectors suggests that events with SNRs higher than 30 will be rare in second-generation detectors [69]. This should be borne in mind when we consider Fig. 3, which shows the 90% confidence regions (see Eq. (2.7)) for a much higher SNR of 50. The confidence regions correspond to a  $7M_\odot$  binary, and show results for both single-effective-spin PN models. We project the 3-dimensional confidence regions in  $(\mathcal{M}_c, \eta, \chi)$  onto the symmetric mass-ratio  $\eta$  and effective spin  $\chi$ . The optimal parameters are denoted by colored symbols, while the true parameters are shown by red stars. At this mass and high SNR the confidence regions are very elongated filaments and we choose to depict them by a curve fit through center of the regions. The regions shown for the equal-spin configuration  $(\chi_1, \chi_2) = (0.45, 0.45)$  (solid dark green) are the confidence regions in the proper sense as the signal is exactly represented by the models at this point. For the waveforms that do not lie in the model subspaces the confidence regions are computed with the model waveform that has the best match with the given signal. The very large uncertainty given by the extent of the confidence regions in Fig. 3 is due to the (approximate) degeneracy between mass-ratio and spin [31, 34, 51, 70]. One can see from the leading order spin-orbit term that it is possible to mimic the effect of spin by modifying the mass-ratio at

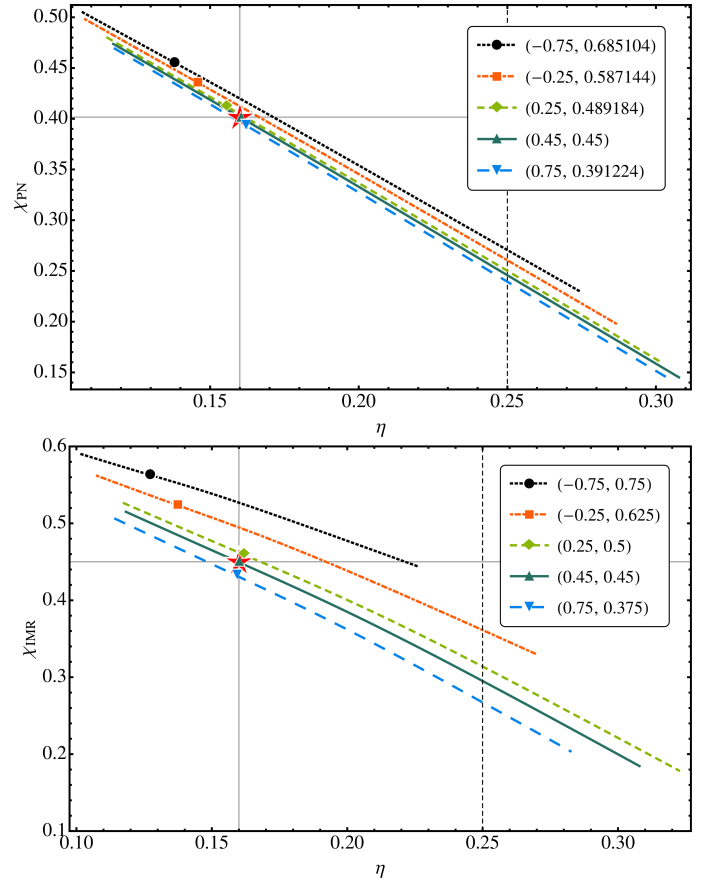


FIG. 3: 90% confidence regions represented as lines at SNR = 50 and  $7M_\odot$  for a single spin PN models using  $\chi_{\text{PN}}$  (top panel) and  $\chi_{\text{IMR}}$  (bottom panel) and TaylorF2 signals along  $\chi_{\text{PN}} = 0.401575$ . The recovered optimal parameters are denoted by colored symbols, while the true parameters are shown by a red star.

constant chirp mass.

The uncertainties for the  $\chi_{\text{PN}}$  model are about  $\Delta\chi \sim 0.35$  and  $\Delta\eta \sim 0.2$ . For the  $\chi_{\text{IMR}}$  model the uncertainties are comparable, but a bit larger, roughly  $\Delta\chi \sim 0.45$  and  $\Delta\eta \sim 0.25$ . For both models the confidence regions extend into the region of unphysical  $\eta > 0.25$ . If we project the confidence regions onto the plane of component masses  $(m_1, m_2)$  all configurations lie on top of a line of constant chirp mass  $\mathcal{M}_c \sim 2.33$ . For both models the configurations range from an equal-mass binary with total mass  $5.4M_\odot$  up to a mass-ratio  $q = 7$  binary with total mass  $9M_\odot$ , as opposed to the true parameters  $(m_1, m_2) = (1.4, 5.6)$ . These results are consistent with those shown in Ref. [70], and illustrate the point made in that work, that we would not be able to determine if such a source was a binary containing two black holes, or a black hole and a neutron star.

We can quantify the additional uncertainty introduced by the single spin approximation with a given parameter  $\chi_{\text{PN}}$  or  $\chi_{\text{IMR}}$  by comparing the “spread” in  $\chi$  between the recovered parameters with the extent of the model confidence region in the  $\chi$  direction. For the  $\chi_{\text{PN}}$ -model we find a spread  $\Delta\chi_{\text{PN}} \sim 0.06$  vs a spread of  $\Delta\chi_{\text{IMR}} \sim 0.13$  for the  $\chi_{\text{IMR}}$  model. The

extent of the equal-spin confidence regions in  $\chi$  is roughly 0.32 for  $\chi_{\text{PN}}$  and 0.34 for  $\chi_{\text{IMR}}$ . For both models at the chosen mass of  $7M_{\odot}$  the statistical uncertainties dwarf the spread in the biases, even at this high SNR of 50. In addition, note that all of the recovered parameters for the  $\chi_{\text{PN}}$  model are within the statistical error bars of the “true” parameters.

These results demonstrate that, while the systematic parameter biases from the single-effective-spin models may appear large, they are in fact much smaller than the statistical errors, even at high SNR. We conclude, then, that the reduced-spin model presented in Ref. [10] is likely to be sufficient for parameter estimation of low-mass signals from aLIGO and AdV.

## V. RESULTS FOR IMR WAVEFORMS

### A. Matches between IMR waveforms

We now consider the family of  $\chi_{\text{IMR}} = 0.45$  PN-NR hybrid waveforms summarized in Tab. I.

In Fig. 4 we show matches between the TaylorF2 frequency-domain hybrid (see Sec. III A) of the reference  $(\chi_1, \chi_2) = (0.45, 0.45)$  case, with each of the other configurations listed in Tab. I. As we would expect, the further away the individual spins are from the fiducial  $(0.45, 0.45)$  waveform the worse the matches become.

The degradation of the matches for low masses is expected due to the use of the  $\chi_{\text{IMR}}$  parameter; the configurations used in this study lie along a line of  $\chi_{\text{IMR}} = 0.45$ . The  $(0.45, 0.45)$  configuration corresponds to  $\chi_{\text{PN}} \approx 0.402$ . Fig. 1 shows how these lines diverge. Those configurations for which the spread between these lines is the largest (i.e., those that are the farthest away from the fiducial  $(0.45, 0.45)$  configuration) are therefore expected to have the worst match. In fact, for low masses, the matches between these IMR waveforms are very close to the PN matches computed at  $7M_{\odot}$  in Fig. 2, as expected.

Around  $50M_{\odot}$  hybridization artifacts lead to a visible kink in the matches. Note that these are not artifacts of the hybridization procedure itself, but rather a result of the disagreement between the TaylorF2 and fully general relativistic NR waveforms at the matching frequency; these artifacts could be made arbitrarily small if we produced NR waveforms of sufficient length to match to PN at arbitrarily low frequencies. Previous studies of NR waveform length requirement (in particular Ref. [30]), and the estimates of statistical uncertainties in this study, suggest that the measurement errors due to these effects do not have a significant impact on the scientific information that can be extracted from aLIGO and AdV GW observations.

As we go to higher masses the matches improve considerably, which indicates that in the merger regime  $\chi_{\text{IMR}}$  performs well. Beyond  $200M_{\odot}$  the matches drop off as we move away from the fiducial  $(0.45, 0.45)$  configuration due to the different final spin and thus different ringdown frequency of the remnant BHs.

It is very likely that the matches between waveforms along a line of constant  $\chi_{\text{PN}}$  would be far higher even through the

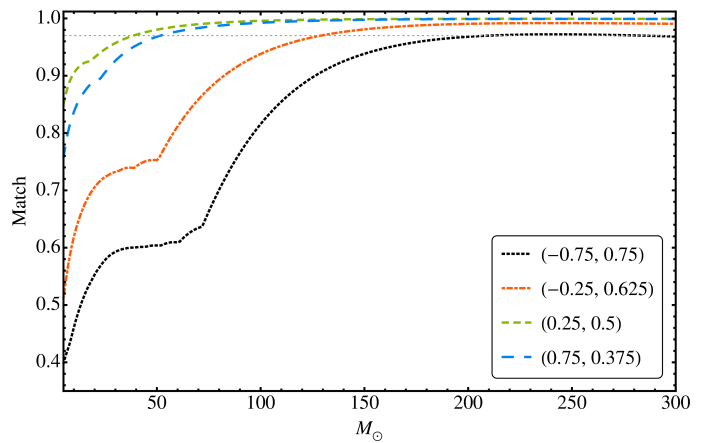


FIG. 4: Matches between unequal spin TaylorF2 frequency-domain PN-NR hybrids with the equal spin  $(0.45, 0.45)$  hybrid.

merger and ringdown. Again, our results are consistent with our expectation from PN theory that future phenomenological models should be parametrized by  $\chi_{\text{PN}}$ .

### B. Biases and uncertainties against IMRPhenomC

As with the TaylorF2 reduced-spin PN model in Sec. IV, we now wish to study the parameter biases due to the use of the single-effective-spin  $\chi_{\text{IMR}}$  in full inspiral-merger-ringdown waveform models. For this purpose we will compare our family of constant  $\chi_{\text{IMR}} = 0.45$  PN-NR hybrids against one of the current phenomenological IMR models, IMRPhenomC. Our goal is complicated by a number of factors. One class of error sources in our analysis is the accuracy of the hybrid waveforms: this will depend on the accuracy of the NR waveforms, on the frequency at which they are hybridized to PN waveforms, and the accuracy of the PN approximant that was used. These errors are mitigated by choosing the same PN approximant, TaylorF2, as was used for the inspiral part of IMRPhenomC. The two other sources of uncertainty (hybridization frequency and NR-waveform accuracy), can be quantified, and we will show (in Figs. 12 and 13) that they do not affect our conclusions.

A more serious source of error is in the IMRPhenomC model itself. This model was calibrated to NR waveforms up to mass ratio 1:4, but *not* spinning-binary waveforms at that mass ratio. The hybrids that we compare with the model are therefore at the very edge of the region of parameter space over which the model was calibrated. In addition, IMRPhenomC (and all current phenomenological models in general) were designed with detection in mind, and not as a tool for parameter estimation. There are certainly errors in how well IMRPhenomC represents the hybrid waveforms from which it was built, i.e. *modeling errors*, as well as hybridization artifacts due to the waveform length, and errors in the NR simulations that were used. From these combined error sources we expect a bias in the parameters that we estimate using this model, which will be nontrivially combined with

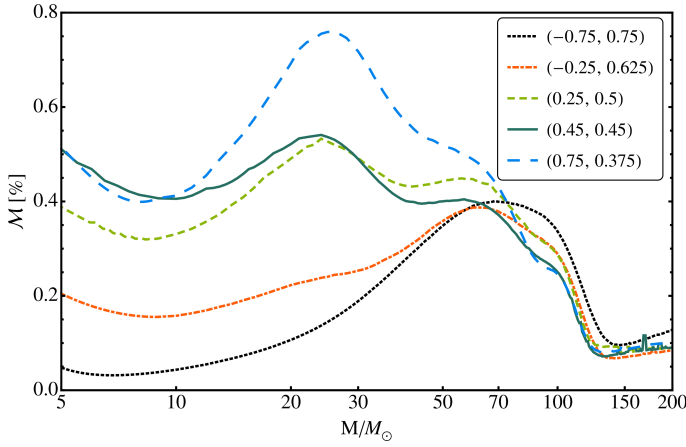


FIG. 5: Fully optimized mismatches between IMRPhenomC and TaylorF2 PN-NR hybrids.

the errors due to the use of the single-effective-spin approximation, which are the biases we wish to measure in this study. The other error sources can be reduced in future phenomenological models, while the bias due to the single-effective-spin approximation will be inherent in all such models.

Despite these complications, we are able to draw a number of important conclusions from our results, which we discuss in this section.

The first point to emphasize is that, despite all of the shortcomings outlined above, the IMRPhenomC model achieves its main purpose as a search template family. We show fully optimized mismatches with IMRPhenomC in Fig. 5. All of them are below 1%. This confirms the effectualness of IMRPhenomC in this region of the parameter space, and the suitability of the model for aligned-spin GW searches. Let us contrast these results of IMRPhenomC with its predecessor IMRPhenomB: Due to its simpler PN part IMRPhenomB is only effectual for masses  $> 20 M_\odot$  for the configurations considered here.

Biases of the total binary mass, as a function of the total mass of the signal, are shown in Fig. 6. This figure illustrates well the complications that were discussed above in our comparisons with IMRPhenomC. The reference waveform has spins (0.45, 0.45), and should be identical to the  $\chi_{\text{IMR}} = 0.45$  waveform in the model. The bias in the total mass for this case is therefore most likely due to modeling artifacts in IMRPhenomC. If we are to assess the bias due only to the use of the single-effective-spin approximation, then we must look at the *spread* of the parameters away from the (solid) (0.45, 0.45) line in the figure. We then see that the spread in the total mass can be as high as  $\sim 5\%$  at low masses (due in most part, once again, to the parametrization by  $\chi_{\text{IMR}}$  instead of  $\chi_{\text{PN}}$ ) and no more than  $\sim 2\%$  at intermediate masses. At high masses the spread in the total mass is around 1%.

The bias in the chirp mass is shown in Fig. 7. For inspiral signals, we expect to be able to measure the chirp mass extremely accurately, because it is the leading-order PN contribution; that is the motivation for the definition of the chirp mass. And indeed the bias in the chirp mass is below 0.1%

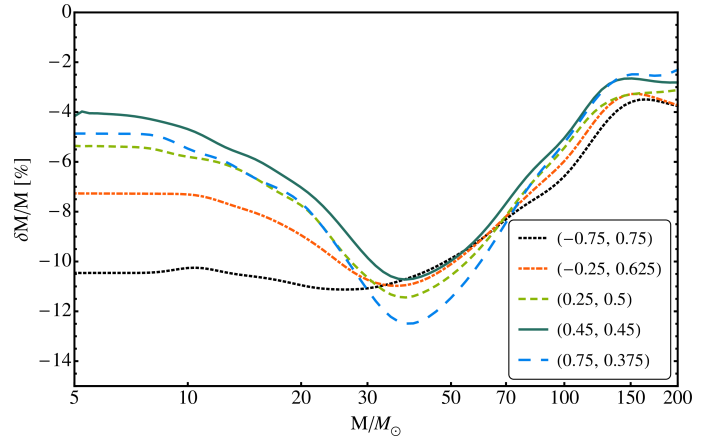


FIG. 6: Biases in the binary mass for IMRPhenomC and TaylorF2 PN-NR hybrid waveforms.

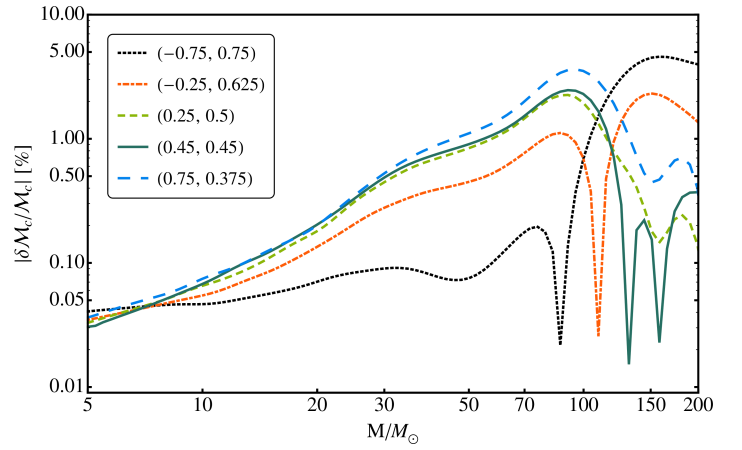


FIG. 7: Biases in the chirp mass  $M_c$  for IMRPhenomC and TaylorF2 PN-NR hybrid waveforms.

for masses below  $10 M_\odot$ . However, the chirp mass has little significance during merger and ringdown, and the chirp-mass bias increases for higher-mass binaries; the fractional error in the chirp mass is comparable to (and in some cases larger than) that in the total mass above  $50 M_\odot$ . Note that the dips in the curves in the figure are due to the use of a logarithmic scale; these are points where  $\delta M/M$  changes sign.

The biases in mass ratio and the symmetric mass-ratio are shown in Fig. 8. Once again, to interpret these figures in terms of the bias due to the single-effective-spin approximation, we should consider the spread of values around the (0.45, 0.45) lines. We see for the mass ratio, the spread in values is around 20% for high-mass systems. (At low masses, the results are again exaggerated by the use of the  $\chi_{\text{IMR}}$  parameter.)

The biases in the spin parameter and the recovered spin values themselves are shown in Fig. 9. At low masses it makes sense to compare the results with the findings of Sec. IV where we studied single spin PN models. The PN model using  $\chi_{\text{IMR}}$  is the relevant one to compare with. The biases are in general consistent between our PN and IMR studies: very low in  $M_c$ , of similar magnitude in  $\eta$  and  $\chi$ . The spread in the  $\chi$  biases

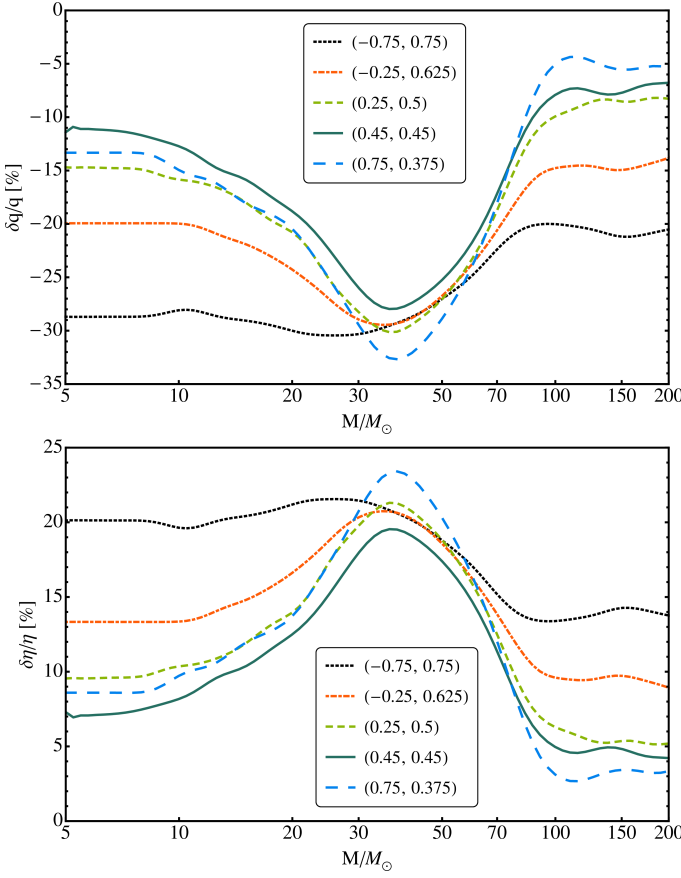


FIG. 8: Biases in the mass-ratio and the symmetric mass-ratio for IMRPhenomC and TaylorF2 PN-NR hybrid waveforms.

Case $(\chi_1, \chi_2)$	Mismatch $\mathcal{M}$	$\Delta M_c/M_c$ [%]	$\Delta \eta/\eta$ [%]	$\Delta \chi/\chi$ [%]	recovered $\chi_{\text{IMR}}$
$(-0.75, 0.75)$	$4 \times 10^{-3}$	-0.1	17.65	5.15	0.473
$(-0.25, 0.625)$	$4 \times 10^{-3}$	-0.34	15.35	-6.67	0.42
$(0.25, 0.5)$	$4 \times 10^{-3}$	-0.88	17.51	-28.93	0.32
$(0.45, 0.45)$	$4 \times 10^{-3}$	-0.89	15.56	-32.07	0.306
$(0.75, 0.375)$	$5 \times 10^{-3}$	-1.06	17.83	-44.06	0.252

TABLE IV: Fully optimized mismatches and biases between IMRPhenomC and IMR signals (frequency domain TaylorF2 hybrids) along  $\chi_{\text{IMR}} = 0.45$  computed using local minimization in the center of the confidence region for  $50 M_\odot$ .

is a little smaller, about 15% for the IMR results, as opposed to 25% for PN. The difference in the PN and IMR confidence regions, even at low masses, was also noted in Ref. [70], and will be studied further in future work.

As in the PN study in Sec. IV, the spread in biases needs to be put into context with the statistical uncertainty in the parameter measurements. For comparison, Fig. 10 shows the 90% confidence regions (see Eq. (2.7)) for a  $50 M_\odot$  binary and table IV summarizes the biases at this mass. While the extent of the confidence regions is much more confined than was the case for the PN models, the regions are still elongated in a diagonal direction in  $\eta$  and  $\chi$  and illustrate the degeneracy between mass-ratio and spin [51]. We see that the uncertainty

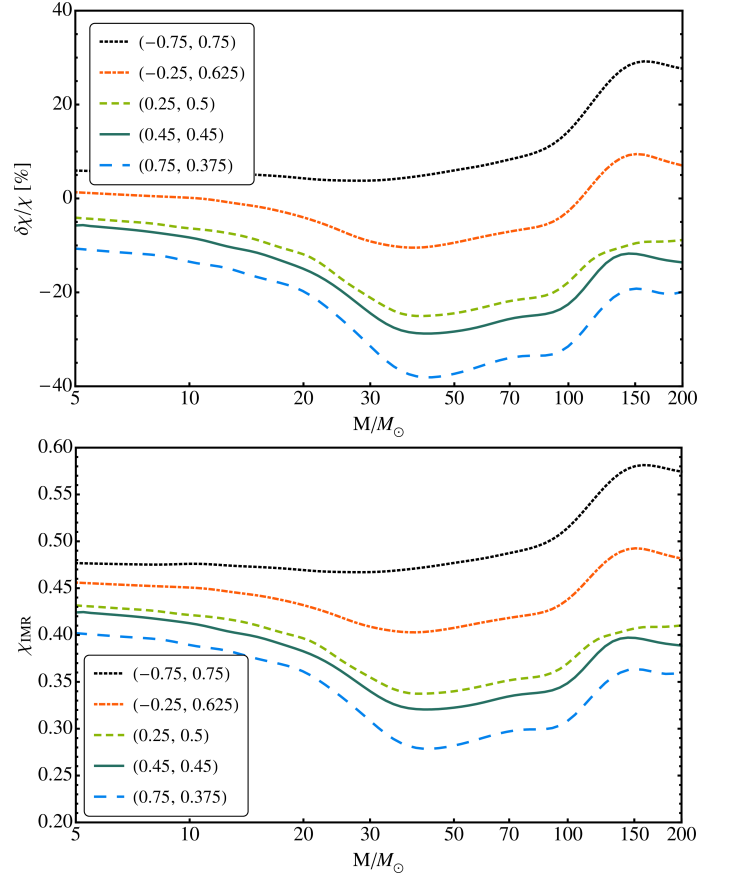


FIG. 9: Biases in the effective spin  $\chi_{\text{IMR}}$  for IMRPhenomC and TaylorF2 PN-NR hybrid waveforms (top) and recovered effective spin  $\chi_{\text{IMR}}$  (bottom). At SNR 10 the statistical uncertainty from the 90% confidence region for masses 20, 50,  $100 M_\odot$  is roughly 0.15, 0.2 and 0.4, respectively.

Mass	$20 M_\odot$	$50 M_\odot$	$100 M_\odot$
Spread in $\chi$ biases	0.11	0.22	0.2
Extent of CR in $\chi$ SNR 10	0.16	0.22	0.41
Extent of CR in $\chi$ SNR 20	0.07	0.1	0.2
Comparable SNR	13	10	20

TABLE V: We compare the spread in the  $\chi$  biases for the 5 hybrids with IMRPhenomC against the extent of the IMRPhenomC confidence region in the  $\chi$  direction at SNR 10 and 20 for  $M = 20, 50, 100 M_\odot$ . We also show at which SNR these numbers become comparable.

in the masses is in general larger than the spread in the parameter biases. For example, the spread in  $\eta$  values at  $50 M_\odot$  was  $\sim 5\%$ , while the statistical uncertainty in the mass ratio at SNR 30 is  $\sim 10\%$  (see Fig. 10). We can conclude, then, that the single-effective-spin approximation does not adversely affect estimation of the black-hole masses at likely advanced-detector SNRs.

The situation is quite different for the spin parameter. At  $50 M_\odot$ , the recovered spin parameter has a spread of  $\Delta \chi_{\text{IMR}} \approx 0.2$ , while the statistical uncertainty in  $\chi_{\text{IMR}}$  becomes comparable at SNR 10. For higher SNRs the spread in the spin bias

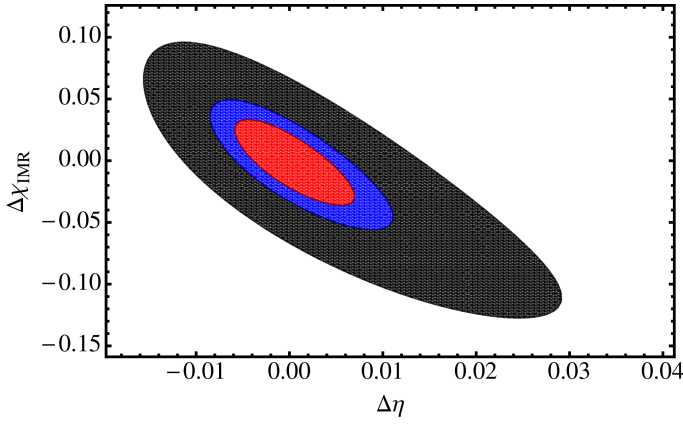


FIG. 10: 90% confidence regions at SNR = 10, 20, 30 (black, blue, red) and  $50M_\odot$  for IMRPhenomC with IMRPhenomC ( $\eta = 0.16$ ,  $\chi = 0.45$ ).

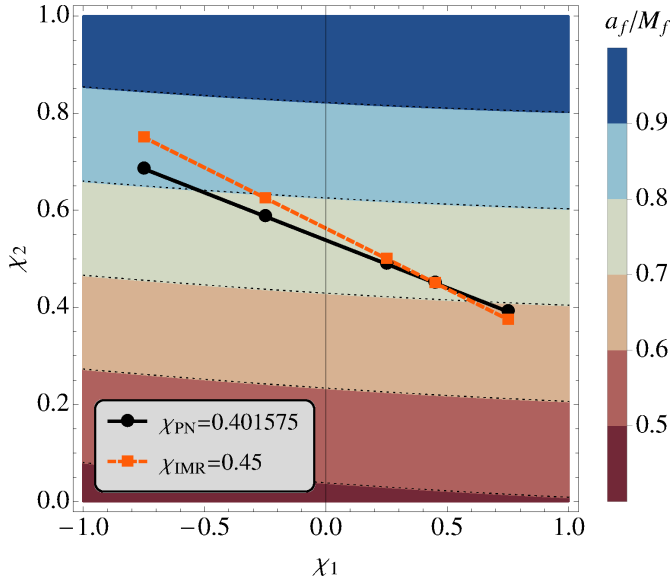


FIG. 11: Contours of constant final spin  $a_f/M_f$  compared to lines of  $\chi = \text{const}$  as shown in Fig. 1.

dominates.

We focus on the bias in  $\chi_{\text{IMR}}$  in Tab. V. We concentrate on the spread of the biases and therefore ignore the offset in the average bias of the IMR configurations from the true value, due to the model not being faithful in this region of the parameter space. The spread of recovered parameter values gives us an indication of the additional uncertainties introduced by the single-effective-spin approximation. We find that the single spin approximation is valid for  $\chi_{\text{IMR}}$  to SNR 10 for masses 7, 20,  $50M_\odot$  and up to SNR 20 for  $M > 100M_\odot$ . Results for IMRPhenomB lead to comparable conclusions.

We have noted in Sec. V A that the waveform from the ring-down of the final black hole will be characterized by the final spin, and not by either  $\chi_{\text{PN}}$  or  $\chi_{\text{IMR}}$ . We illustrate this point in Fig. 11, where we overlay the curves of constant  $\chi_{\text{PN}} = 0.401575$  and  $\chi_{\text{IMR}} = 0.45$  on the contours of constant

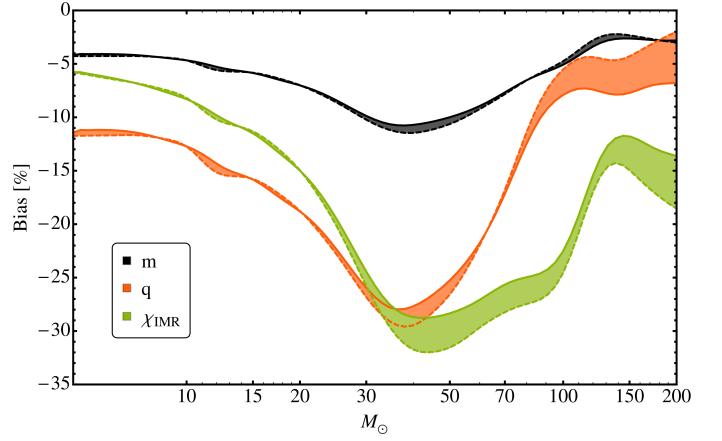


FIG. 12: Biases for the (0.45, 0.45) configuration using a numerical resolution of  $N = 80$  (solid) vs  $N = 64$  (dashed) gridpoints.

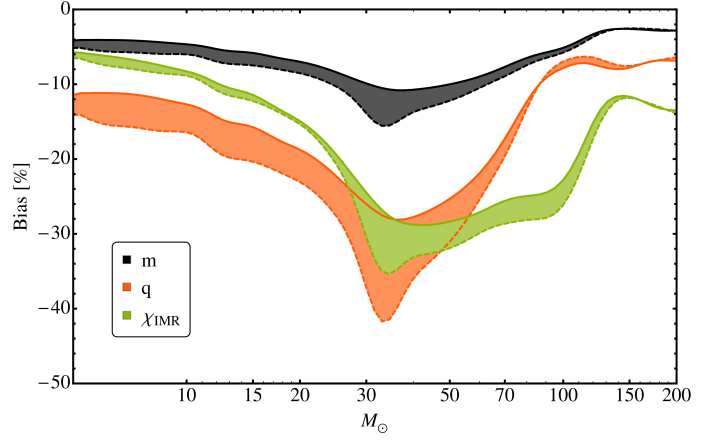


FIG. 13: Biases for the (0.45, 0.45) configuration using TaylorF2-hybrids with matching frequencies of  $M\omega_m = 0.07$  (solid) vs  $M\omega_m = 0.08$  (dashed).

*final spin*, which can be predicted by a number of formulas in the literature [23, 25, 26, 71, 72]. The results of the various final-spin formulas agree to within a few percent with our numerical results for the final spins (see Tab. I) with the largest disagreement at the  $(-0.75, 0.75)$  configuration. We see that configurations with the same value of either spin parameter during inspiral can lead to a black hole with a wide range of final spins, depending on the individual spins of the progenitor black holes. (In Tab. I we see that final spins for our family of  $\chi_{\text{IMR}} = 0.45$  numerical simulations ranges from 0.68 up to 0.84.)

Finally, we verify that our results are not qualitatively changed by errors in the numerical-relativity waveforms or artifacts due to the choice of hybridization frequency. Fig. 12 compares the variation in biases caused by changing the resolution of the numerical waveform used to construct the hybrid, while Fig. 13 shows the effect of changing the hybridization frequency. Changing the NR resolution only manifests itself at higher masses (both hybrids have been constructed at comparable matching frequencies  $M\omega \sim 0.07$ ), whereas changing

the hybrid parameter affects predominantly low to medium masses. While this is not an in-depth error analysis, it gives an indication of how sensitive our results are to these two sources of errors that we can control. It is clear from the figures that neither of these error sources appears to be a serious issue in most cases, but does warrant further study in the future.

## VI. CONCLUSIONS

Several non-precessing-binary waveform models make use of the observation that the effects of the black-hole spins on the inspiral rate can be approximated by a single effective spin parameter, either  $\chi_{\text{PN}}$  in the case of a reduced-spin inspiral model [10], or  $\chi_{\text{IMR}}$  for phenomenological inspiral-merger-ringdown models [9, 14]. (The two parameters are defined in Eqns. (2.1) and (2.2).) These models were developed primarily for use in template banks for gravitational-wave searches, but they have also been used for parameter estimation [50]. We have investigated the systematic bias in parameter measurement due to the use of these single-effective-spin approximations.

Our primary goal has been to explore the parameter bias from phenomenological waveform models, and for this purpose we focussed on a family of numerical-relativity simulations of binaries with mass ratio 1:4, and all with  $\chi_{\text{IMR}} = 0.45$  (but with different values of individual spins). Fig. 4 shows that the noise-weighted inner product (match) between these waveforms is in general larger than 0.97 for masses greater than  $200 M_\odot$ , but the match degrades significantly at lower masses. We note that if we had parameterized these waveforms instead by a constant value of  $\chi_{\text{PN}}$ , then their matches would be much better at low masses, since we know that the waveforms are partially degenerate in  $\chi_{\text{PN}}$  in the low-frequency (post-Newtonian) regime.

In assessing systematic parameter biases, we first consider the inspiral regime, and study two related families of post-Newtonian waveforms, one with  $\chi_{\text{PN}} = 0.45$  and the other  $\chi_{\text{IMR}} = 0.45$  (again, with varying individual spins). Here we find that, for signal-to-noise ratios below 50 (and we expect most observations in aLIGO and AdV to be below 30), the statistical uncertainty in the measurement of both the masses and the spin ( $\chi_{\text{PN}}$ ) is significantly larger than the parameter bias incurred by using the single-effective-spin-approximation model to estimate the parameters. This result is discussed in detail in Sec. IV, and summarized here in the upper panel of Fig. 14. Although limited to only one point in the  $(\eta, \chi_{\text{PN}})$  parameter space, this result suggests that the reduced-spin inspiral model is sufficiently accurate for parameter estimation from GW observations in aLIGO and AdV.

We then consider the hybrid PN-NR inspiral-merger-ringdown waveforms (with constant  $\chi_{\text{IMR}} = 0.45$ ), and compare them against the IMRPhenomC model. Our first observation is that, while the IMRPhenomC model performs well for detection purposes, with all fitting factors  $>0.99$  (see Fig. 5), artifacts in the construction of the waveform model cause a significant systematic bias even for the equal-spin waveform that should be reproduced by the model with  $\chi_{\text{IMR}} = 0.45$ .

This problem can be removed in future by producing a model calibrated against NR waveforms across a larger volume of parameter space, and such work is already underway. The bias we observe at low masses (see Figs. 6, 8, and 9) is consistent with our choice of  $\chi_{\text{IMR}}$  as our single-effective-spin parameter, and suggests that future phenomenological models should be parameterized instead with  $\chi_{\text{PN}}$ .

Our main observation from these results is that at intermediate masses (around  $50 M_\odot$ ) the *spread* in recovered spin values is far larger than the statistical uncertainty in  $\chi_{\text{IMR}}$ , even at an SNR of 10, which is close to the detection threshold. This can be seen in Figs. 9 and 10, and is summarized here in the lower panel of Fig. 14, which shows the deviation in the parameter measurement from the value obtained for the reference  $\chi_1 = \chi_2 = 0.45$  waveform.

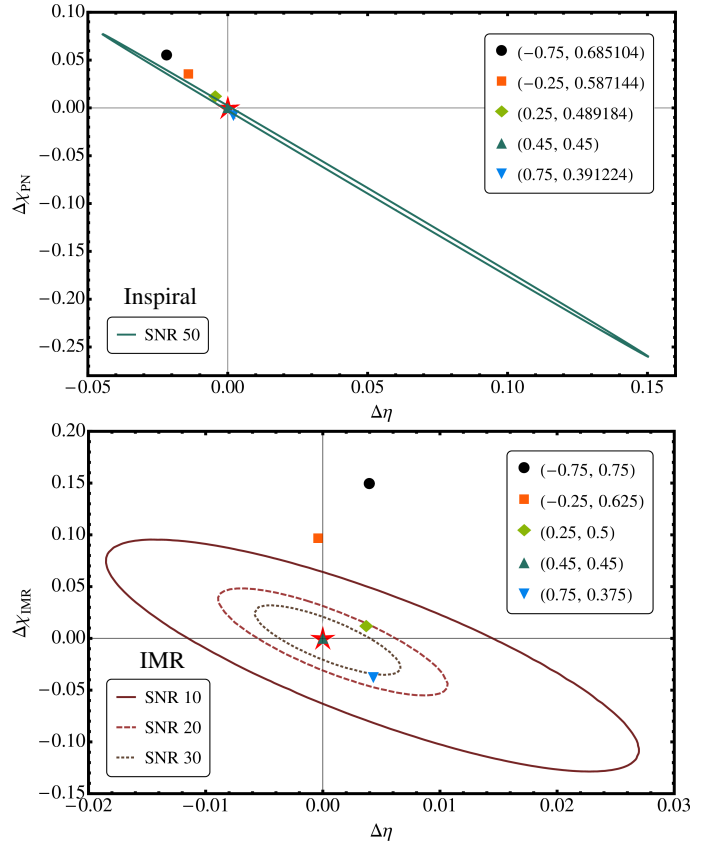


FIG. 14: These plots compare the systematic biases in measuring the effective spin parameter  $\chi$  and the symmetric mass ratio  $\eta$  with the corresponding statistical uncertainties in the case of the *reduced-spin* PN model (top) and the phenomenological IMR model IMR-PhenomC (bottom). In each panel, the star corresponds to the actual value of parameters ( $\eta = 0.16, \chi_{\text{PN}} = 0.401575$  in the top panel and  $\eta = 0.16, \chi_{\text{IMR}} = 0.45$  in the bottom panel) while the other markers correspond to the parameters of the best-fit single-spin model. The ellipses correspond to the statistical uncertainty in measuring the parameters (90% confidence region) at different SNRs. In the bottom panel, the recovered parameter values have been shifted so that they coincide with the true parameters for the equal-spin configuration.

The small parameter bias due to the single-effective-spin approximation at low masses implies that the approximation

holds well at these masses, and therefore that *it will be difficult to measure the component spins* in low-mass binaries. This is the unfortunate corollary of the validity of the single-effective-spin approximation: GW searches can be more efficient, but parameter estimation is less accurate.

Since we observe large parameter biases at intermediate masses, this implies that in these cases we *may be* able to measure the individual spins. It is quite likely that there is a strong degeneracy between the two spins and the mass ratio at all stages in the binary's evolution, but if we observe a waveform in which the early part (the inspiral) is characterized by  $\chi_{\text{PN}}$ , and the late part (the ringdown) is characterized by the final spin, then it is likely that to describe the full waveform we require knowledge of both black-hole spins. If this is the case, then it follows that accurate measurements of both spins may be possible. This is an interesting topic for further work.

This study analyzes only the  $l = 2, m = \pm 2$  modes of the gravitational wave signal. The amplitude of higher order modes is sensitive to the mass-ratio, and thus could serve to break degeneracies between  $\chi$  and  $\eta$  in parameter estimation. However, it should be borne in mind that at the mass-ratios that this model is calibrated at the higher modes are all much weaker than the 22-mode and at the low SNRs considered here we do not expect our results to change appreciably.

Do these results imply that we should construct a two-spin non-precessing model? Not necessarily. A single-effective-spin model is sufficient for detection, and, following detection, if the source is a black-hole-binary with total mass  $\sim 50 M_{\odot}$ , then we could perform addition simulations with varying individual black-hole spins, and produce a localized model for parameter estimation purposes. We note that although it is possible to produce sufficient waveforms to cover the non-

precessing-binary parameter space in time for the commissioning of aLIGO and AdV, a higher priority may be to produce models that approximately cover the *full* precessing-binary parameter space, for which it would be more efficient to model the spins parallel to the orbital angular momentum by only a single parameter.

Further work is needed to verify the spread of the parameter biases in the spin in the IMR waveforms, using updated phenomenological waveform models, and studies across a larger volume of the parameter space. In assessing our ability to measure either an effective single spin parameter, or both black-hole spins, we also need to quantify the influence of harmonics beyond the dominant ( $\ell = 2, |m| = 2$ ) modes.

## VII. ACKNOWLEDGEMENTS

We thank S. Fairhurst, F. Ohme, H. Pfeiffer, S. Khan, B. Sathyaprakash, and P. Schmidt for useful discussions and comments. M. Hannam was supported Science and Technology Facilities Council grants ST/H008438/1 and ST/I001085/1. S. Husa was supported by grant FPA-2007-60220 from the Spanish Ministry of Science and the Spanish MICINN's Consolider-Ingenio 2010 Programme under grant MultiDark CSD2009-00064. M. Pürrer thanks Caltech and the Universitat de les Illes Balears (UIB) for hospitality. BAM simulations were carried out at Advanced Research Computing (ARCCA) at Cardiff, and as part of the European PRACE petascale computing initiative on the clusters Hermit, Curie and SuperMUC.

- 
- [1] G. M. Harry et al. (LIGO Scientific Collaboration), *Class.Quant.Grav.* **27**, 084006 (2010).
  - [2] F. Acernese et al. (The Virgo Collaboration), *Advanced Virgo Baseline Design* (2009), [Virgo Technical Document VIR-0027A-09].
  - [3] J. Abadie et al. (LIGO Scientific), *Class. Quant. Grav.* **27**, 173001 (2010), 1003.2480.
  - [4] B. Sathyaprakash and B. Schutz, *Living Rev.Rel.* **12**, 2 (2009), <http://www.livingreviews.org/lrr-2009-2>.
  - [5] J. Aasi et al. (LIGO Scientific Collaboration, Virgo Collaboration), *Phys. Rev. D* **87**, 022002, 022002 (2013), 1209.6533.
  - [6] J. Abadie et al. (LIGO Collaboration, Virgo Collaboration), *Phys.Rev.* **D85**, 082002 (2012), 1111.7314.
  - [7] J. Abadie et al. (LIGO Scientific Collaboration, Virgo Collaboration), *Phys.Rev.* **D83**, 122005 (2011), 1102.3781.
  - [8] J. Abadie et al. (LIGO Scientific Collaboration, Virgo Collaboration), *Phys.Rev.* **D82**, 102001 (2010), 1005.4655.
  - [9] P. Ajith, M. Hannam, S. Husa, Y. Chen, B. Bruegmann, et al., *Phys.Rev.Lett.* **106**, 241101 (2011), 0909.2867.
  - [10] P. Ajith, *Phys.Rev.* **D84**, 084037 (2011), 1107.1267.
  - [11] D. A. Brown, I. Harry, A. Lundgren, and A. H. Nitz, *Phys.Rev.* **D86**, 084017 (2012), 1207.6406.
  - [12] P. Schmidt, M. Hannam, and S. Husa, *Phys.Rev.* **D86**, 104063 (2012), 1207.3088.
  - [13] L. Pekowsky, R. O'Shaughnessy, J. Healy, and D. Shoemaker (2013), 1304.3176.
  - [14] L. Santamaria et al., *Phys. Rev.* **D82**, 064016 (2010), 1005.3306.
  - [15] B. Vaishnav, I. Hinder, F. Herrmann, and D. Shoemaker, *Phys. Rev.* **D76**, 084020 (2007), 0705.3829.
  - [16] C. Reisswig et al., *Phys. Rev.* **D80**, 124026 (2009), 0907.0462.
  - [17] P. Ajith, N. Fotopoulos, S. Privitera, A. Neunzert, and A. Weinstein (2012), 1210.6666.
  - [18] P. Ajith et al., *Class. Quant. Grav.* **24**, S689 (2007), 0704.3764.
  - [19] P. Ajith et al., *Phys. Rev.* **D77**, 104017 (2008), 0710.2335.
  - [20] P. Ajith, *Class. Quant. Grav.* **25**, 114033 (2008), 0712.0343.
  - [21] M. Hannam and I. Hawke, *Gen.Rel.Grav.* **43**, 465 (2011), 0908.3139.
  - [22] A. H. Mroue, M. A. Scheel, B. Szilagyi, H. P. Pfeiffer, M. Boyle, et al. (2013), 1304.6077.
  - [23] A. Buonanno, L. E. Kidder, and L. Lehner, *Phys. Rev.* **D77**, 026004 (2008).
  - [24] L. Rezzolla et al., *Astrophys. J.* **674**, L29 (2008), 0710.3345.
  - [25] W. Tichy and P. Marronetti, *Phys.Rev.* **D78**, 081501 (2008), 0807.2985.
  - [26] C. O. Lousto, M. Campanelli, and Y. Zlochower, *Class. Quant. Grav.* **27**, 114006 (2010), 0904.3541.
  - [27] M. Hannam, S. Husa, F. Ohme, and P. Ajith, *Phys.Rev.* **D82**, 124052 (2010), 1008.2961.
  - [28] I. MacDonald, S. Nissanke, H. P. Pfeiffer, and H. P. Pfeiffer,

- Class.Quant.Grav. **28**, 134002 (2011), 1102.5128.
- [29] M. Boyle, Phys.Rev. **D84**, 064013 (2011), 1103.5088.
- [30] F. Ohme, M. Hannam, and S. Husa, Phys.Rev. **D84**, 064029 (2011), 1107.0996.
- [31] E. Poisson and C. M. Will, Phys.Rev. **D52**, 848 (1995), gr-qc/9502040.
- [32] *LSC algorithm library (LAL)*, URL <http://www.lsc-group.phys.uwm.edu/lal>.
- [33] B. Sathyaprakash and S. Dhurandhar, Phys.Rev. **D44**, 3819 (1991).
- [34] C. Cutler and E. E. Flanagan, Phys.Rev. **D49**, 2658 (1994).
- [35] S. Droz, D. J. Knapp, E. Poisson, and B. J. Owen, Phys. Rev. D **59**, 124016 (1999), gr-qc/9901076.
- [36] P. Ajith, M. Boyle, D. A. Brown, B. Brügmann, L. T. Buchman, et al., Class.Quant.Grav. **29**, 124001 (2012), 1201.5319.
- [37] C. Cutler and E. E. Flanagan, Phys. Rev. D **49**, 2658 (1994).
- [38] B. Abbott et al. (LIGO Scientific), Rept. Prog. Phys. **72**, 076901 (2009), 0711.3041.
- [39] D. Shoemaker (the Advanced LIGO Team), *Advanced LIGO Reference Design* (2009), [LIGO-M060056].
- [40] G. M. Harry and the LIGO Scientific Collaboration, Class. Quant. Grav. **27**, 084006 (2010).
- [41] J. Aasi et al. (LIGO Scientific Collaboration, Virgo Collaboration) (2013), 1304.0670.
- [42] The LIGO Scientific Collaboration, Tech. Rep. LIGO-T0900288-v3, LIGO Project (2009), URL <https://dcc.ligo.org/DocDB/0002/T0900288/003/AdvLIGO%20noise%20curves.pdf>.
- [43] S. Babak, R. Biswas, P. Brady, D. Brown, K. Cannon, et al., Phys.Rev. **D87**, 024033 (2013), 1208.3491.
- [44] L. S. Finn and D. F. Chernoff, Phys.Rev. **D47**, 2198 (1993), gr-qc/9301003.
- [45] K. Arun, B. R. Iyer, B. Sathyaprakash, and P. A. Sundararajan, Phys.Rev. **D71**, 084008 (2005).
- [46] M. van der Sluys et al., Classical and Quantum Gravity **25**, 184011 (2008), 0805.1689.
- [47] M. V. van der Sluys et al., The Astrophysical Journal Letters **688**, L61 (2008), 0710.1897.
- [48] J. Veitch and A. Vecchio, Phys. Rev. D **81**, 062003 (2010), 0911.3820.
- [49] F. Feroz, M. P. Hobson, and M. Bridges, Monthly Notices of the Royal Astronomical Society **398**, 1601 (2009), 0809.3437.
- [50] J. Aasi et al. (LIGO Collaboration, Virgo Collaboration) (2013), 1304.1775.
- [51] E. Baird, S. Fairhurst, M. Hannam, and P. Murphy, Phys.Rev. **D87**, 024035 (2013), 1211.0546.
- [52] J. Nelder and R. Mead, Comput.J. **7**, 308 (1965).
- [53] F. Ohme, A. B. Nielsen, D. Keppel, and A. Lundgren (2013), 1304.7017.
- [54] T. Tanaka and H. Tagoshi, Phys.Rev. **D62**, 082001 (2000), gr-qc/0001090.
- [55] A. Pai and K. Arun, Class.Quant.Grav. **30**, 025011 (2013), 1207.1943.
- [56] M. Pürrer, S. Husa, and M. Hannam, Phys.Rev. **D85**, 124051 (2012), 1203.4258.
- [57] B. Brügmann et al., Phys. Rev. **D77**, 024027 (2008), gr-qc/0610128.
- [58] S. Husa, J. A. González, M. Hannam, B. Brügmann, and U. Sperhake, Class. Quant. Grav. **25**, 105006 (2008), 0706.0740.
- [59] S. Brandt and B. Brügmann, Phys. Rev. Lett. **78**, 3606 (1997), gr-qc/9703066.
- [60] J. M. Bowen and J. W. York, Jr., Phys. Rev. **D21**, 2047 (1980).
- [61] M. Ansorg, B. Brügmann, and W. Tichy, Phys. Rev. **D70**, 064011 (2004), gr-qc/0404056.
- [62] M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower, Phys. Rev. Lett. **96**, 111101 (2006), gr-qc/0511048.
- [63] J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. van Meter, Phys. Rev. Lett. **96**, 111102 (2006), gr-qc/0511103.
- [64] M. Hannam, S. Husa, D. Pollney, B. Brügmann, and N. Ó Murchadha, Phys. Rev. Lett. **99**, 241102 (2007), gr-qc/0606099.
- [65] M. Shibata and T. Nakamura, Phys. Rev. **D52**, 5428 (1995).
- [66] T. W. Baumgarte and S. L. Shapiro, Phys. Rev. **D59**, 024007 (1999), gr-qc/9810065.
- [67] S. Husa et al. (2013), in preparation.
- [68] D. McKechnan, C. Robinson, and B. Sathyaprakash, Class.Quant.Grav. **27**, 084020 (2010), 1003.2939.
- [69] L. Lindblom, B. J. Owen, and D. A. Brown, Phys. Rev. **D78**, 124020 (2008), 0809.3844.
- [70] M. Hannam, D. A. Brown, S. Fairhurst, C. L. Fryer, and I. W. Harry, Astrophys.J. **766**, L14 (2013), 1301.5616.
- [71] E. Barausse and L. Rezzolla, Astrophys.J. **704**, L40 (2009), 0904.2577.
- [72] L. Rezzolla, E. Barausse, E. N. Dorband, D. Pollney, C. Reisswig, J. Seiler, and S. Husa, Phys. Rev. **D78**, 044002 (2008), 0712.3541.