Systematic challenges for future gravitational wave measurements of precessing binary black holes

A. R. Williamson, J. Lange, R. O'Shaughnessy, J. A. Clark, Prayush Kumar, J. Calderón Bustillo, and J. Veitch

Phys. Rev. D **96**, 124041 — Published 29 December 2017

DOI: 10.1103/PhysRevD.96.124041

Systematic challenges for future gravitational wave measurements of precessing binary black holes

A. R. Williamson,[1,2] J. Lange,[1] R. O'Shaughnessy,[1] J. A. Clark,[3]
Prayush Kumar,[4,5] J. Calderón Bustillo,[3] and J. Veitch[6]

[1] *Center for Computational Relativity and Gravitation, Rochester Institute of Technology, Rochester, New York 14623, USA*
[2] *Department of Astrophysics/IMAPP, Radboud University, Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands*
[3] *Center for Relativistic Astrophysics and School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, USA*
[4] *Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario M5S 3H8, Canada*
[5] *Cornell Center for Astrophysics and Planetary Science, Cornell University, Ithaca, New York 14853, USA*
[6] *SUPA, University of Glasgow, Glasgow G12 8QQ, United Kingdom*

The properties of precessing, coalescing binary black holes are presently inferred through comparison with two approximate models of compact binary coalescence. In this work we show these two models often disagree substantially when binaries have modestly large spins ($a \gtrsim 0.4$) and modest mass ratios ($q \gtrsim 2$). We demonstrate these disagreements using standard figures of merit and the parameters inferred for recent detections of binary black holes. By comparing to numerical relativity, we confirm these disagreements reflect systematic errors. We provide concrete examples to demonstrate that these systematic errors can significantly impact inferences about astrophysically significant binary parameters. For the immediate future, parameter inference for binary black holes should be performed with multiple models (including numerical relativity), and carefully validated by performing inference under controlled circumstances with similar synthetic events.

I. INTRODUCTION

The Advanced Laser Interferometer Gravitational Wave Observatory (LIGO) detector [1] has reported the discovery of five binary black hole (BBH) mergers to date – GW150914 [2], GW151226 [3], GW170104 [4], GW170608 [5], and GW170814 [6] – the latter discovered jointly with the Advanced Virgo instrument [7]. Additionally, an astrophysically plausible candidate BBH signal, LVT151012, has been reported [8]. At this early stage, observations cannot firmly distinguish between a number of possible BBH formation mechanisms [9]. These include the evolution of isolated pairs of stars [10–15], dynamic binary formation in dense clusters [16], and pairs of primordial black holes (BHs) [17]; see, e.g., [9] and references therein.

One way to possibly distinguish between isolated and dynamic formation mechanisms is to measure the spin properties of the BHs [9, 18–23]. The presence of a component of the BH spins in the plane of the orbit leads to precession of that plane. If suitably massive and significantly spinning, such binaries will strongly precess within the LIGO sensitive band. If BBHs are the end points of isolated binary star systems, they would be expected to contain BHs with spins preferentially aligned with the orbital angular momentum [22, 24], and therefore rarely be strongly precessing. If, however, BBHs predominantly form as a result of gravitational interactions inside dense populations of stellar systems, the relative orientations of the BH spins with their orbits will be random, and some gravitational wave signals may be very strongly precessing. Precise measurements of their properties will provide unique clues into how BHs and massive stars evolve [19, 25–30].

The gravitational wave signals produced by strongly precessing systems are challenging to model. Direct numerical simulations of Einstein's equations are possible for these and other generic orbits, but are time consuming to produce. As a result, approximate models to numerical relativity have been developed, and recent models feature ways to mimic signals from precessing systems [31–37]. These approximate models have been used to infer the properties of observed systems [38, 39].

In this paper we demonstrate by example several systematic issues which can complicate the interpretation of rapidly-spinning and precessing binaries. First, we provide one of the first systematic head-to-head comparisons of these models for precessing, coalescing binaries, using physically equivalent parameters for both waveforms; see also [40, 41]. We show that the two models disagree frequently for precessing systems, including parameters within the posterior distributions of gravitational wave events like GW151226 and GW170104. We focus on these two events since their posteriors are considerably different from one another, but each has similarities with other detected BBH mergers; GW151226 is similar to GW170608, and GW170104 is broadly similar to both GW170814 and GW150914. Our study differs from several previous investigations of waveform fidelity [42–44] by focusing on precessing systems and observationally-motivated parameters. The two models principally disagree when the spins are both large and precessing. GW measurements like LIGO's have not strongly prescribed whether such strongly-precessing systems are consistent with any individual observation. Using concrete examples, we remind the reader that the posterior distributions for BH spins can depend significantly on the assumed prior distributions, particularly since these distributions are often broad and nongaussian [see, e.g. 45–47]. One astrophysically plausible prior distribution is significant BH natal spin (e.g., as motivated by some X-ray

ID	Model/Numerical Relativity	q	M ($M_\odot$)	χ_{1x}	χ_{1y}	χ_{1z}	χ_{2x}	χ_{2y}	χ_{2z}
A1	SEOBNRv3	1.91	60.0	-0.390	0.552	-0.346	0.174	-0.079	-0.052
A2	SEOBNRv3	3.01	26.5	0.951	-0.115	0.124	0.510	0.298	0.760
SXS:BBH:0165	Numerical Relativity	6.00	80.0	-0.058	0.776	-0.470	0.076	-0.172	-0.234
SXS:BBH:0112	Numerical Relativity	5.00	80.0	0	0	0	0	0	0

TABLE I: **Parameters of synthetic sources**: This table shows the parameters of all the synthetic sources (waveform approximant models and numerical relativity) used in this paper. q is the mass ratio defined with $q > 1$, M is the total mass, and χ_* are the components of the normalized spins.

observations) and random BH spin-orbit alignment (e.g., as implied by dynamical formation scenarios). We show that, if these prior assumptions are adopted, the posterior distribution is dominated by parameters for which the models disagree even more frequently.

We perform parameter estimation on synthetic signals to demonstrate quantitatively that these disagreements lead to biases in, and different conclusions about, astrophysically relevant quantities. These synthetic signals have parameters and detector configurations consistent with observed events. Extending the study of [48], which focused on weakly precessing systems, we show that inferences about GW sources derived using the conventional configuration can frequently be biased, particularly in certain regions of the parameter space and about observationally-relevant pairs of parameters. We show that the conclusions reached can be strongly dependent on the model used. We point out that extensive followup studies – using multiple models and numerical relativity – were performed on GW150914 [38, 48, 49] and GW170104 [4, 50], producing good agreement across multiple independent calculations.

This paper is organized as follows. In Section II we compare the predictions of two models for the radiation emitted by gravitational waves from precessing BBHs. To make our discussion extremely concrete and observationally relevant, we perform these comparisons on parameters drawn from LIGO's inferences about GW151226, and from our inferences about synthetic events designed to mimic GW170104 and GW151226. Under the conventional assumptions used in this analysis, we find the two models disagree, principally when the inferred binary parameters involve large precessing spins. Because the relative probability of large and precessing spins depends on our prior assumptions, we then repeat these comparison again, adopting prior assumptions that do not disfavor BBHs with two significant, precessing spins. To illustrate the implications of these disagreements, in Section III, we perform several proof-of-concept parameter inference calculations using synthetic gravitational wave signals. Again, using parameters consistent with real observations (i.e., drawn from observed posterior distributions of observed BBH mergers), we show that parameter inferences performed with the two models can disagree substantially about astrophysically relevant correlated parameters, like the mass and spin of the most massive BH. To highlight the fact that these disagreements occur frequently, not merely for systems viewed in rare edge-on lines of sight, we choose synthetic binaries which are inclined by $\pi/4$ to the line of sight. In Section IV, we discuss how our results extend the broadening appreciation of potential sources of systematic error in gravitational wave measurements.

II. MODELS FOR COMPACT BINARY COALESCENCE DISAGREE

A. Models for radiation from binary black holes

When inferring properties of coalescing BBHs [4, 8, 38, 39], LIGO has so far favored two approximate models for their gravitational radiation: an effective one body (EOB) model, denoted SEOBNRv3 [31, 32], and a phenomenological frequency-domain inspiral and merger model, denoted IMRPhenomPv2 [35].

The SEOBNRv3 model extends a long, incremental tradition to modeling the inspiral and spin dynamics of coalescing binaries via an ansatz for the two-body Hamiltonian [37]. In this approach, equations of motion for the BH locations and spins are evolved in the time domain. For nonprecessing binaries, outgoing gravitational radiation during the inspiral phase is generated using an ansatz for resumming the post-Newtonian expressions for outgoing radiation including non-quasicircular corrections, for the leading-order $\ell = 2$ subspace. For the merger phase of nonprecessing binaries, the gravitational radiation is generated via a resummation of many quasinormal modes, with coefficients chosen to ensure smoothness. The final BH's mass and spin, as well as some parameters in the nonprecessing inspiral model, are generated via calibration to numerical relativity simulations of BBH mergers. For precessing binaries, building off the post-Newtonian ansatz of separation of timescales and orbit averaging [51–54], gravitational radiation during the insprial is modeled as if from an instantaneously nonprecessing binary (with suitable nonprecessing spins), in a frame in which the binary is not precessing [55–57]. During the merger, the radiation is approximated using the same final BH state, with the same precession

frequency.[1] With well-specified initial data in the time domain, this method can be directly compared to the trajectories [59] and radiation [60] of numerical BBH spacetimes.

The IMRPhenomPv2 model is a part of an approach that attempts to approximate the leading-order gravitational wave radiation using phenomenological fits to the fourier transform of this radiation, computed from numerical relativity simulations and post-newtonian calculation [33–35]. Also using information about the final BH state, this phenomenological frequency-domain approach matches standard approximations for the post-Newtonian gravitational wave phase to an approximate, theoretically-motivated spectrum characterizing merger and ringdown. Precession is also incorporated by a "corotating frame" ansatz, here implemented via a stationary-phase approximation to the time-domain rotation operations performed for SEOBNRv3.

We make use of the `lalsimulation` implementations of these two approximations, provided and maintained by their authors in the same form as used in LIGO's O1 and O2 investigations.

The coalescence time and orientation (i.e., Euler angles) of a binary are irrelevant for the inference of intrinsic parameters from gravitational wave data. As a result, and following custom in stationary-phase calculations, the IMRPhenomPv2 model does not calibrate the reference phases and time. This makes easy head-to-head comparison with time-domain calculations somewhat more difficult. Specifically, two different sets of parameters are needed to generate the same gravitational radiation in SEOBNRv3 and IMRPhenomPv2, connected by (a) a change in the overall orbital phase $\phi_{\rm orb}$; (b) a change in the precession phase of the orbital angular momentum $\phi_{\mathbf{JL}}$; and (c) a change in the overall coalescence time t. In the approximations adopted by IMRPhenomPv2, these time and phase shifts do not qualitatively change the underlying binary or its overall orientation-dependent emission, just our perspective on it.

B. Binary black hole observations and model-based inference

BBH parameters are inferred through the use of Bayesian analysis with standard Monte Carlo techniques; see, e.g., [38, 61] and references therein.

For any BBH event, fully characterized by parameters x, we can compute the (Gaussian) likelihood function $p(d|x)$ for detector network data d containing a signal by using waveform models and an estimate of the (approximately Gaussian) detector noise on short timescales (see, e.g., [38, 49, 61] and references therein). In this expression x is shorthand for the set of 15 parameters needed to fully specify a quasicircular BBH. The posterior probability distribution is therefore $p(x|d) \propto p(d|x)p(x)$, where $p(x)$ is the prior probability of finding a BBH merger with different masses, spins, and orientations somewhere in the universe. LIGO–Virgo analyses have adopted a fiducial prior $p_{\rm ref}(x)$ that is uniform in orientation, in comoving volume, in mass, in spin direction (on the sphere), and, importantly for us, in spin magnitude [38, 61].

Using standard Bayesian tools [38, 61], one can produce a sequence of independent, identically distributed samples $x_{n,s}$ $(s = 1, 2, \ldots, S)$ from the posterior distribution $p(x|d)$ for each event n; that is, each $x_{n,s}$ is drawn from a distribution proportional to $p(d_n|x_n)p_{\rm ref}(x_n)$. Typical calculations of this type provide $\lesssim 10^4$ samples [38, 61] from which the posterior probability distribution is inferred.

Quite clearly the choice of prior $p_{\rm ref}$ directly influences the posterior, most significantly for parameters not well constrained by the data (e.g., due to weak dependence or strong degeneracies). As a concrete example, in the left panel of Figure 1 we show the cumulative distribution of $\chi_{1,z}$, the component of the primary BH's dimensionless spin in the z-direction, for a synthetic source similar to GW151226. The black curve corresponds to results evaluated using the fiducial prior, where χ_1, χ_2 are distributed independently and uniformly. The red curve is computed by drawing χ_1 from the cumulative distribution $P(<\chi_1) = \chi_1^3$, and similarly for χ_2. Henceforth we denote this the *volumetric (spin)* prior.

In the context of systematic errors and astrophysical measurements of BBHs, the choice of prior is important. Within the context of a specific astrophysical scenario or question of interest, a prior favoring large spins (or significant precession) can be appropriate. As we show later, these changes in prior can significantly increase the posterior probability of the region where model disagreement is substantial (e.g., large transverse spins, high mass ratio, and long signals).

When assessing the impact of modified priors, we exercise an abundance of caution and replicate the Bayesian inference calculations in full. In principle, with sufficiently many samples, we could estimate the posterior distribution for any prior $p(x)$ by using weighted samples. For example, we could estimate $P(<X)$ according to the modified prior $p(x)$ via the weighted empirical cumulative distribution $\hat{P}(<X) = \sum_k \Theta(X - X_k)p(x_k)/(N p_{\rm ref}(x_k))$. The approach of reweighted posterior samples is widely proposed in hierarchical model selection [27, 63]. In practice, however, this method is reliable if and only if x_k cover the parameter space completely and sufficiently densely. In our specific circumstances, the fiducial prior $p_{\rm ref}(x)$ associates substantial prior weight near $\chi_1, \chi_2 \simeq 0$ and little probability to configurations with two large spins. As a result, rescaling from the fiducial to the volumetric prior can introduce biases into astrophysical conclusions. As a concrete

[1] This choice of merger phase behavior is known to be inconsistent with precessional dynamics during merger [41, 58].

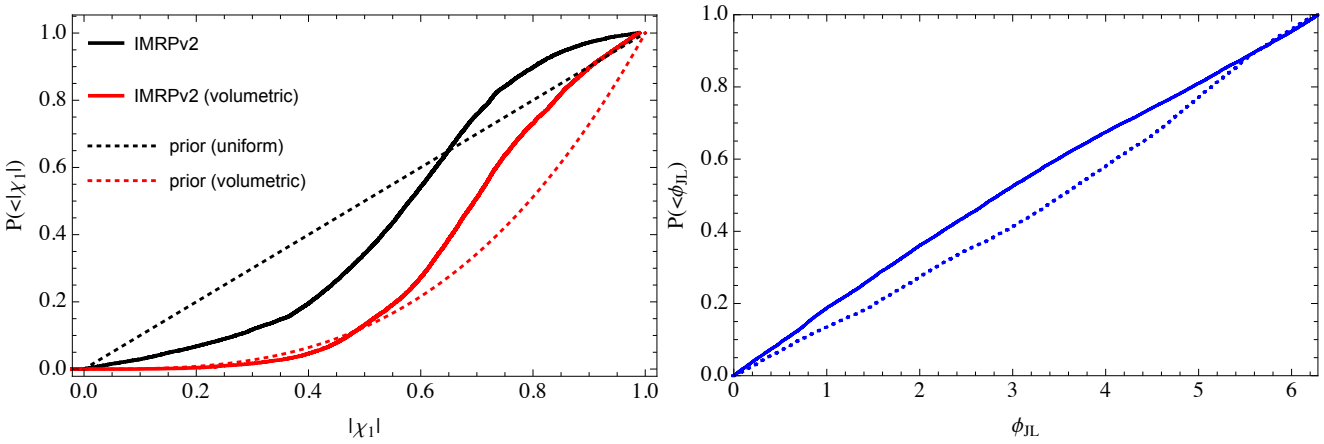

FIG. 1: **Priors and the relative significance of large spins** *Left panel*: For a synthetic GW151226-like event (A2), the inferred cumulative posterior distribution for $|\chi_1|$ using a prior $P(<|\chi_i|) = \chi_i$ (black) and $P(<|\chi_i|) = \chi_i^3$ (red), for $i = 1, 2$. For comparison, the two priors are indicated with dotted black and red lines. The posterior probability that this synthetic event has two significant, precessing spins depends on the prior. *Right panel*: Inferred cumulative posterior distribution for $\phi_{\mathbf{JL}}$, the polar angle of $\mathbf{L}$ relative to $\mathbf{J}$, for the volumetric prior $P(<|\chi_i|) = \chi_i^3$. The solid blue line shows the results of repeating a full PE calculation, including the modified prior. The dotted blue line shows the estimated distribution calculated by weighting the posterior samples. This synthetic event was generated with parameters similar to GW151226 and analyzed with a PSD appropriate to GW150914, generated in the manner of [62].

example, the right panel of Figure 1 shows the cumulative distribution of $\phi_{\mathbf{JL}}$, the polar angle of $\mathbf{L}$ relative to $\mathbf{J}$. The solid line shows the result of a full calculation with the volumetric prior. The dotted line shows the result derived using reweighted posterior samples, starting from the fiducial uniform-magnitude prior. While the two distributions are approximately consistent in extent, the two disagree in details. If used uncritically in (hierarchical) model selection, reweighted posterior samples could lead to biased conclusions about model inference, and (in the context of our study) to biased conclusions about the relative impact of model-model systematics. Of course, a careful treatment of reweighted posterior systematics would identify this potential problem, and the need for more samples to insure a reliable answer in any reweighted application (i.e., the expected variance of the Monte Carlo integral estimate for $\hat{P}$ is large, because $p/p_{\rm ref}$ is often large).

C. Model-model comparisons

To quantify the difference between two predicted gravitational waves from the same binary with the same spacetime coordinates and location, we use a standard data-analysis-motivated figure of merit: the mismatch. Like other figures of merit, the mismatch is calculated using an inner product between two (generally complex-valued) timeseries $a(t), b(t)$:

$$\langle a|b\rangle = 2\int_{|f|\geq f_{min}} \frac{\tilde{a}^*(f)\,\tilde{b}(f)}{S_n(|f|)} df\,, \qquad (1)$$

where $S_n(|f|)$ is the noise power spectral density of a fiducial detector, f_{min} is a chosen lower frequency cutoff (typically a few tens of Hz), and the integral includes both positive and negative frequencies. Usually these comparisons also involve parameterized signals $a(\lambda, \theta)$ and $b(\lambda', \theta')$, with maximization of the (normalized) inner product between a, b over some set of parameters Θ:

$$\langle a|b\rangle_\Theta = \max_\Theta \frac{\mathrm{Re}\,\langle a(\theta)|b(\theta')\rangle}{\sqrt{\langle a(\theta)|a(\theta)\rangle\,\langle b(\theta')|b(\theta')\rangle}} \qquad (2)$$

where Θ denotes the names of the parameters in θ over which we maximize. Maximization is asymmetric; we change the parameters of only one of the two signals, effectively considering the other as "the source". When the signals a and b are real-valued single-detector response timeseries and when Θ includes only t and $\phi_{\rm orb}$, this expression is known as the *match*.

In our comparisons, we fix one of the two timeseries a generated by model A, as if it was some known detector response (e.g., from another model's prediction). The other timeseries is a predicted single-detector response $b = \mathrm{Re}F^*h$ where F is a complex-valued antenna response function and h is the gravitational wave strain. Ideally, we should evaluate b using model B and precisely the same intrinsic and extrinsic parameters, calculating the *faithfulness* [64]. The precessing models considered in this work have different time and phase conventions. In order to specify the astrophysically equivalent binary to some configuration as evolved by SEOBNRv3, we need to adopt different t, $\phi_{\rm orb}$, and $\phi_{\mathbf{JL}}$. Reconciling the phase conventions adopted by these models is far beyond the scope of this work. However, we can find the most optimistic possible answer by maximizing the inner product

over t, $\phi_{\rm orb}$, and $\phi_{\mathbf{JL}}$, using differential evolution [65] to evolve towards the best-fitting signal. In other words, we use a figure of merit

$$\langle a|b\rangle_{t,\phi_{\rm orb},\phi_{\mathbf{JL}}} \ . \tag{3}$$

Note that, since our gravitational wave signals include higher modes, we do not simply maximize the match over $\phi_{\mathbf{JL}}$, as this would neglect the contibution to the inner product from these higher modes.

D. Comparison on posterior distributions

To investigate systematic errors in observationally relevant regions of parameter space,we perform model-model comparisons using samples drawn from the posterior parameter distributions for several of LIGO's detections to date, as well as for synthetic sources. Unless otherwise noted, these comparisons are performed on the expected Hanford detector response.

Figure 2 illustrates our comparisons for our synthetic GW170104-like event (A1 in Table I), using the fiducial spin prior. For the short waveforms needed to explain this signal, disagreement between the models is primarily associated with higher levels of precession, viewed in an orientation near the orbital plane where the effects of precession dominate.

Since lower mass systems take longer to evolve from some lower frequency to the merger, waveforms drawn from the posterior of GW151226 are significantly longer in duration. The two waveform models have significantly greater opportunity to dephase, leading to lower inner products. Figure 3 shows the distribution of these mismatches. The disagreement is significant over a larger portion of the parameter space than for our GW170104-like event, and is less strongly correlated with the orbital inclination $\theta_{\mathbf{JN}}$. It is clear that the spins play a leading role in producing these differences.

In Figure 4 we show two results for a synthetic GW151226-like event (A2 in Table I), one that adopts a uniform spin prior, and the other the volumetric spin prior. As in Figure 3 we see that more moderate-to-highly spinning systems in the posteriors show a greater degree of diagreement between the models. The volumetric spin prior increases the support for large spins, and so increases the proportion of the posterior where model disagreement is significant.

III. EXAMPLES OF BIASED INFERENCE OF BH PARAMETERS

To illustrate the discrepancies in inferred parameters which such disagreements can cause, we select points with significant differences and generate the associated waveforms with one model, before running the full parameter estimation analysis on these waveforms using the other model. We do not add any simulated instrumental noise to the model signal in this process.

In these demonstrations of the practical differences between models from each other and from numerical relativity, we use the same parameter estimation techniques and models applied by LIGO to infer the parameters of the first two observed BBHs [8, 66, 67]. Figures 5 and 6 show concrete examples of biased parameter inference. In Figure 5, the red dot shows the parameters of a synthetic signal, generated with SEOBNRv3 using intrinsic and extrinsic parameters corresponding to a sample point in the posterior distribution for GW151226 that showed significant mismatch between waveform models. The binary parameters chosen correspond to a configuration where the models disagree (i.e., low inner product); see Figure 3. Our synthetic data contains only the expected detector response (the "zero noise" realization), which we interpret in the context of synthetic off-source noise with the observed frequency-dependent form.[2] The black curves show the 90% posterior confidence intervals, derived using the LALINFERENCE parameter inference engine. In Figure 6, we generate a synthetic source signal from a numerical relativity simulation produced by the SXS collaboration [74], using an extension to LIGO's infrastructure to designed for this purpose [75, 76]. This figure demonstrates by concrete example that the two models' disagreement can propagate into biased inference about astrophysically important binary parameters, even now in a regime of low signal amplitude and large statistical error.

IV. DISCUSSION

A. Mismatch does not imply bias: Examples with high mass ratio and zero spin

Due to their neglect of higher-order modes, the two models disagree significantly with numerical relativity at high mass ratio, even in the absence of spin. Several previous studies have demonstrated these modes have a significant impact on the match [77–80]. For example, Figure 7 illustrates the mismatch introduced due to the neglect of higher order modes for non-spinning systems of varying mass ratios, with total masses $M = 80M_\odot$ and inclinations $\theta_{\mathbf{JN}} = \pi/4$.

A large mismatch, however, does not imply a large *bias*. As an example, Figure 8 shows two sets of inferences for a signal generated from a numerical relativity simulation. This had significant mismatch with both models

[2] A commonly-used technique to investigate the implications of parameter inference [68–73], the use of the "zero noise" realization allows us to compute trivially-reproducible posteriors which are centered on the true parameters yet also have the same structure (e.g., width and correlations) as would be expected from any realization of detector noise.

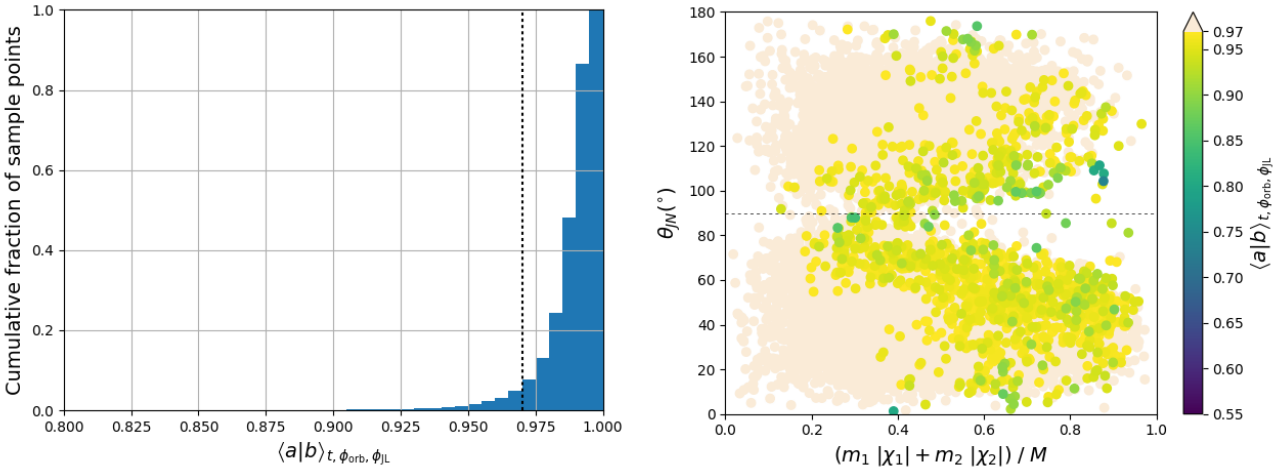

FIG. 2: **Model-model comparison on our synthetic GW170104-like event**: Using posterior samples from our synthetic GW170104-like event (A1), we calculate model-model inner products between IMRPhenomPv2 and SEOBNRv3 waveforms, maximized over t, $\phi_{\rm orb}$, and $\phi_{\mathbf{JL}}$. This analysis adopts the fiducial (uniform) prior on spin magnitude. In the left panel is a cumulative histogram of the maximized inner products. In the right panel the posterior samples are plotted in terms of $\theta_{\mathbf{JN}}$, the inclination of the observer relative to the total angular momentum, and a measure of the net binary BH spin. The color scale indicates the maximized inner product, with the lowest values occurring for large binary spins and preferentially near the orbital plane. The noise curve used for these calculations was the same as used in Figure 1.

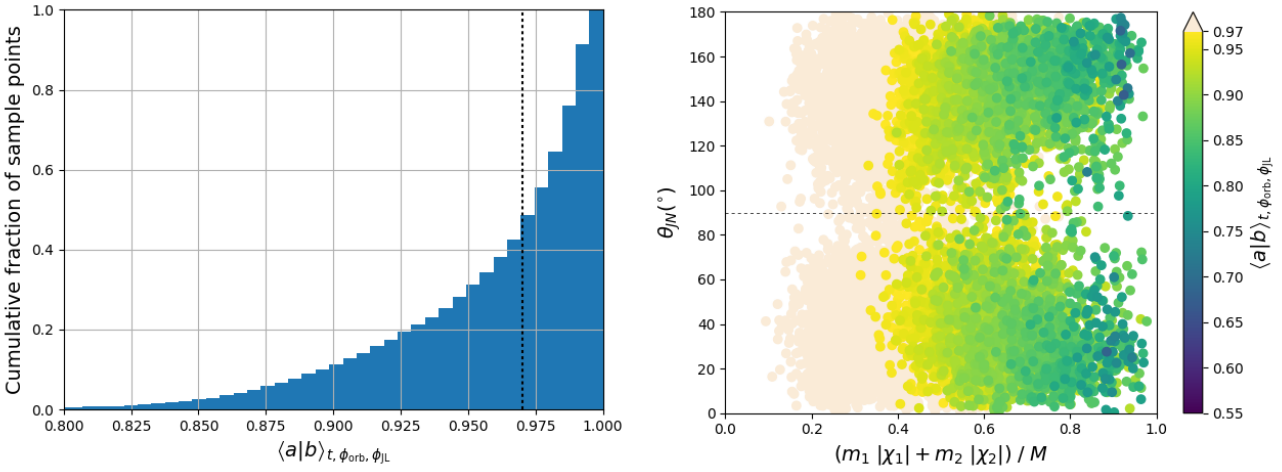

FIG. 3: **Model-model comparison on GW151226**: As Figure 2 but for GW151226. The intrinsic and extrinsic parameters used in this comparison are from LIGO's O1 posterior distribution. Frequent and significant disagreement is apparent. IMRPhenomPv2 produces waveforms that are somewhat longer than SEOBNRv3 for these modest masses, leading to dephasing due to a slight difference in the rate of frequency evolution integrating over such long waveforms. This effect correlates strongly with the binary spin. The noise curve used for these calculations was calculated from data near the time to GW151226.

used to perform these inferences when evaluated at the exact binary parameters of the simulation, except t, $\phi_{\rm orb}$, and $\phi_{\mathbf{JL}}$, which again are maximized over. Despite this discrepancy, the two models draw qualitatively similar conclusions.

Another reason for mismatch that may not result in large bias is the presence of higher modes. Models which omit or include higher harmonics can often disagree substantially when judged by mismatch. In the limit of a long signal, the different harmonics have distinct time-frequency trajectories and transfer information with minimal cross-contamination [83–85]. As a result, an analysis using only *one* mode will find similar best-fitting parameters to one using multiple higher modes, but with a wider posterior due to neglected information. At higher mass and near the end of the merger, however, multiple modes are both significant and, due to their brevity, harder to distinguish. Using a simple matched-based analysis applied to hybridized nonprecessing multimodal numerical relativity simula-

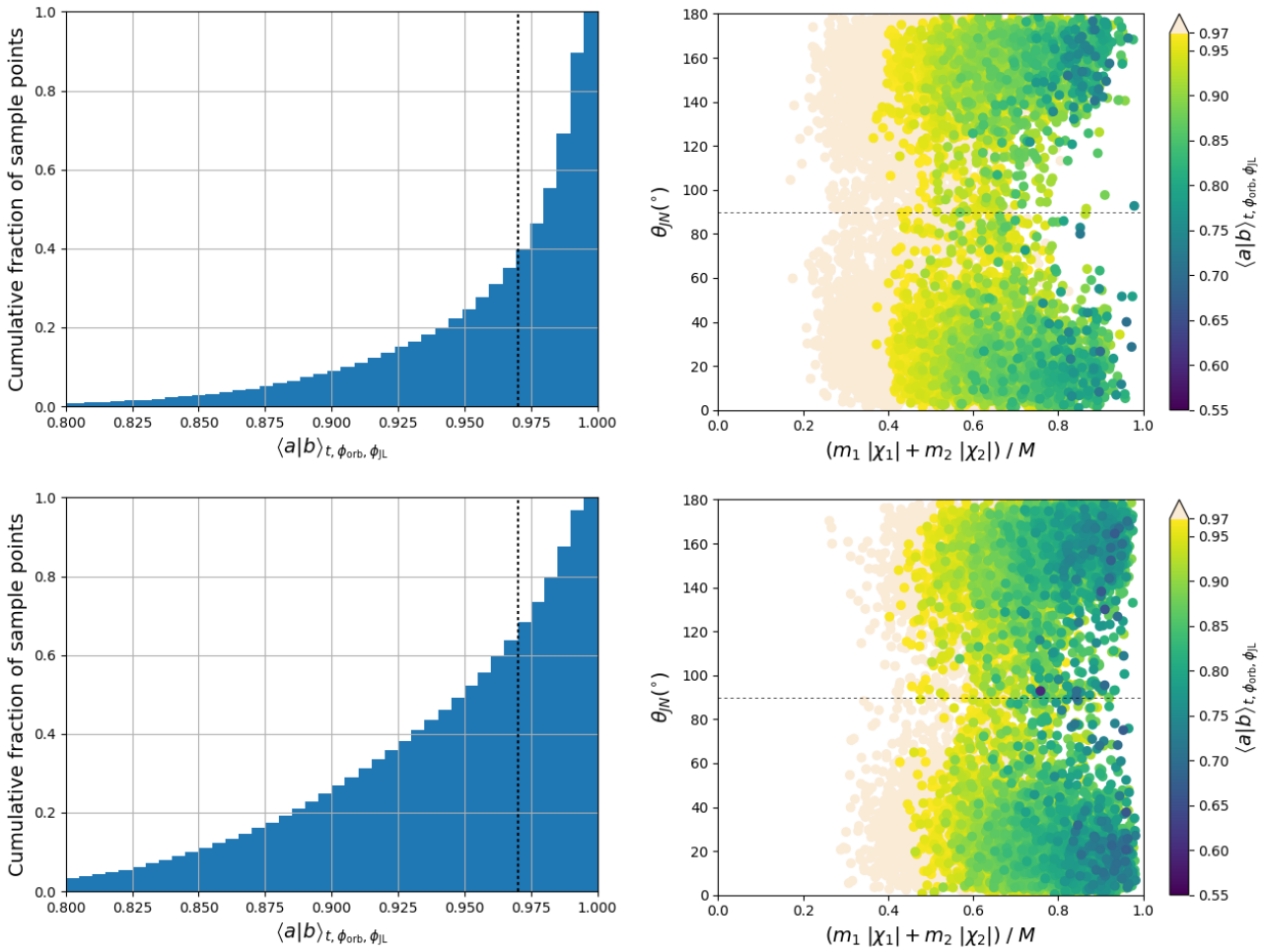

FIG. 4: **Model-model comparison on a synthetic GW151226-like event**: As Figures 2 and 3 but for a synthetic GW151226-like event (A2). As in Figure 2, the top and bottom panels show the results assuming a uniform and volumetric spin prior, respectively. Adopting a volumetric spin prior noticably increases the posterior support for large spins and hence the fraction of the posterior associated with parameters where the two models disagree significantly.

tions, [79] argued that for moderate-mass binaries, inferences based on the leading-order quadrupolar model alone would not be significantly biased, compared to the (large) statistical error expected at modest SNR; see the right panel of their Figure 1. For nonprecessing zero-spin binaries, we confirm by example that inferences about the binary are not biased. Figure 9 shows the posterior distributions inferred using two EOB models, one including higher-order modes (EOBNRv2HM), and the other omitting them (SEOBNRv4). The synthesised signal is a nonprecessing binary with $q = 5$ and $M = 80M_\odot$, generated via numerical relativity (i.e., a signal including higher order modes). Due to model limitations, these inferences are performed assuming both BHs have zero spin. This figure shows that both sets of parameter inferences are consistent with the true binary parameters used, and that inferences constructed with higher modes (via EOBNRv2HM) are both sharper and less biased than inferences that omit higher modes (via SEOBNRv4).

A large mismatch does imply, however, that the analysis is not using all available information. For example, searches for gravitational waves which neglect higher modes cannot fully capture all available signal power and a priori are somewhat less sensitive [78–80, 86]; but cf. [87]. Parameter inference calculations that use higher modes are well-known to be more discriminating about binary parameters [68, 69, 73, 88–92, 92]. Even for short signals associated with heavy BBHs, analyses with higher modes can draw tigher inferences about binary parameters [49, 69, 73], depending on the source; see, e.g., Figure 8.

B. Marginal distributions, degeneracy, and biases

Fortunately or not, nature and LIGO's instruments have conspired to produce short GW signals with modest

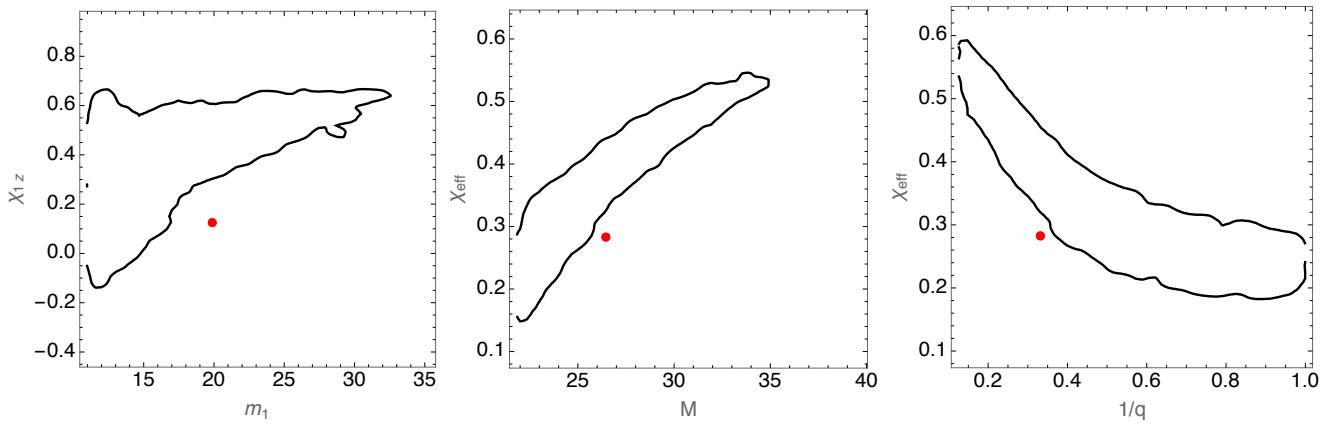

FIG. 5: **Biased parameter recovery with IMRPhenomPv2 I: SEOBNRv3 source (A2)**: Red dot shows the parameters of a synthetic coalescing binary, whose radiation is modeled with SEOBNRv3. Binary parameters are drawn from the posterior distribution of GW151226, and are summarized in Table I as source A2. The inclination of the orbital angular momentum relative to our line of sight is $\theta_{\mathbf{JN}} = 2.48$. No synthetic noise is added to the signal. For this source, the match between the detector response predicted using IMRPhenomPv2 and SEOBNRv3 is 0.817 in Hanford, after maximizing in $t, \phi_{orb}, \phi_{\mathbf{JL}}$. Black curve shows the 90% confidence interval derived from a detailed parameter inference calculation using the IMRPhenomPv2 approximation. Calculations are performed using a network of detectors whose noise power spectra are identical to the estimates derived for GW150914 [8, 62], using frequencies above 20Hz.

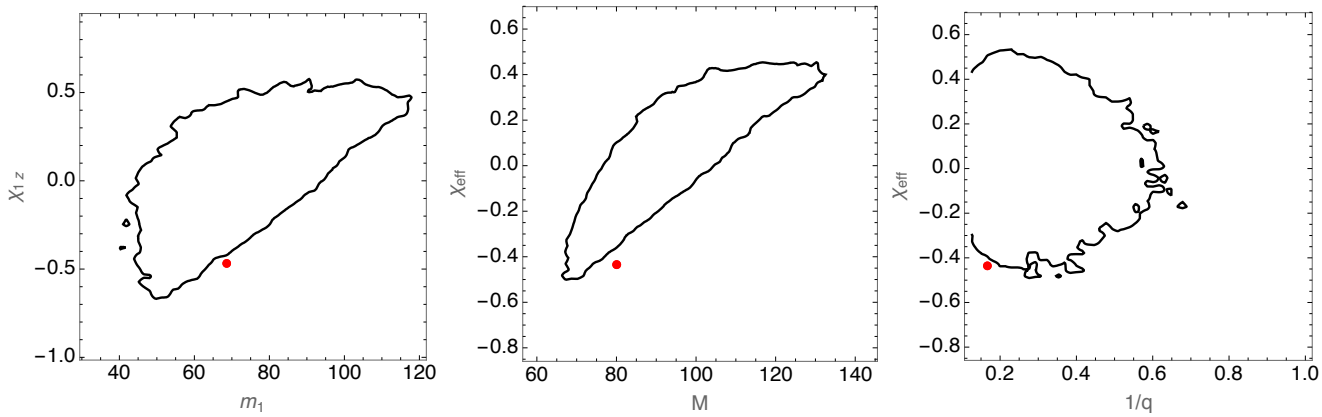

FIG. 6: **Biased parameter recovery with IMRPhenomPv2 II: NR source**: Red dot shows the parameters of a synthetic coalescing binary, whose radiation is modeled with a numerical relativity simulation SXS BBH:0165. All simulated modes $\ell \leq 8$ are included in our synthetic signal. The detector response is calculated assuming a signal at angle $\theta_{\mathbf{JN}} = \pi/4$, at a distance so the network SNR is $\sim$10. No synthetic noise is added. The black curves show the 90% confidence interval derived from a detailed parameter inference calculation using the IMRPhenomPv2 approximation. Calculations are performed using a network of detectors whose noise power spectra are identical to the estimates derived for GW150914 [62]. Because the $(2, 2)$ mode of this source starts at 27Hz, we only use frequencies greater than 30Hz in our analysis

amplitude to date. As illustrated by LIGO's results [4, 8] and our Figure 5, when using current methods (e.g., IMRPhenomPv2 and SEOBNRv3), the inferred posterior distributions for most parameters are quite broad, dominated by substantial statistical error. Inferences about individual parameters are also protected by strong degeneracies in these approximate models (e.g., in the neglect of higher-order modes) and in the physics of binary mergers (e.g., in the dependence of merger trajectories on net aligned spin). For example, the rightmost panel of Figure 5 shows the posterior distribution in mass ratio, q, and effective spin, $\chi_{\rm eff} = \frac{c}{G}\left(\frac{\chi_1}{m_1} + \frac{\chi_2}{m_2}\right) \cdot \frac{\hat{\mathbf{L}}}{M}$ [93, 94]; the joint posterior is tightly correlated (and strongly biased), but the individual marginal distributions for q and $\chi_{\rm eff}$ are broad and contain the true parameters.

In principle, inference with higher modes and precession can more efficiently extract information from and produce significantly narrower posteriors for BBH mergers; see, e.g., [69, 73] as well as our Figure 9. Proof-of-concept new models containing these modes for precessing binaries have only recently been introduced [40, 41], and have not yet been extensively applied to parameter inference.

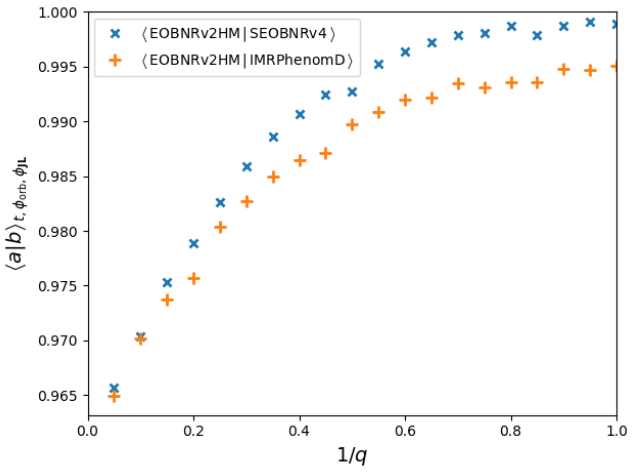

FIG. 7: **The effect on inner product due to neglecting higher modes**: Here we generate a series of non-spinning waveforms with $M = 80M_\odot$ and $\theta_{\mathbf{JN}} = \pi/4$ using an EOB model that includes higher modes, EOBNRv2HM, then use the same parameters to generate waveforms with two models that do not include these higher modes, one EOB and one phenomenological – SEOBNRv4 [81] and IMRPhenomD [82]. Again we calculate the inner product maximized over $\phi_{\rm orb}$, $\phi_{\mathbf{JL}}$, and t. As higher modes are most important for heavier and unequal mass binaries, these large mismatches may be responsible for disagreements seen in Figure 6. Conversely, higher modes are not significant for and not included in models compared Figures 3 and 5, so are unlikely to be responsible for the large discrepancies seen there.

V. CONCLUSIONS

Using concrete BBH parameters consistent with LIGO's observations to date, we have demonstrated that the two waveform models used to infer BBH parameters can often be significantly inconsistent with one another, as measured by the inner product of waveforms generated with identical parameters except coalescence time, orbital phase, and precession phase, which we maximize over. Differences are most significant for parameters corresponding to strongly precessing BBHs, viewed from directions where modulations from precession are strongly imprinted on the outgoing radiation. While significant model-model mismatch is a necessary but not sufficient condition for parameter biases, we demonstrate that for unexceptional binary parameters – at the signal strengths, masses, and spins corresponding to current observations – these model differences are more than sufficient to significantly bias parameter inference for astrophysically interesting quantities, like the joint distribution of the most massive BH's mass and spin.

At present, the posterior distribution for any individual astrophysically interesting parameter is often large, due to a combination of modest signal strength, brevity, and some degree of model incompleteness. In particular, even for the most extreme examples of synthetic inference studied here, where model disagreements were most substantial, the one-dimensional posterior probability distributions still contained the known value. While these biases described in this work are not always large compared to the posterior's extent, these biases could complicate attempts to use multiple events to draw astrophysical inferences about compact binary populations.

In principle, systematic differences between models could be identified by performing parameter inference with both models and identifying regions of disagreement. In practice, however, for long BBH merger signals like GW151226, the current computational cost of large-scale parameter estimation with SEOBNRv3 and conventional parameter inference tools remains cost-prohibitive. Work on optimizing the generation of waveforms is ongoing (see e.g. [95]) and will hopefully enable this approach in the future.

In the immediate future, however, parameter inference for BBH mergers should be performed with multiple models (including numerical relativity), and carefully validated by performing inference under controlled circumstances with similar synthetic events. After the discovery of GW150914, Abbott et al. [48] performed a systematic parameter investigation study, assessing how reliably the (known) parameters of synthetic signals were recovered. That investigation used IMRPhenomPv2 for parameter inference, full numerical relativity simulations as sources, and emphasized source parameters similar to GW150914. Our study complements this initial investigation by directly comparing the two models used for inference, by using both model- and numerical relativity-based synthetic sources, and by using source parameters consistent with subsequent LIGO observations. Extensive followup studies of this kind were also performed for GW170104 [4, 50], and the general parameters of these events are not in doubt.

Acknowledgments

ARW acknowledges support from the RIT's Office of the Vice President for Research through the FGWA SIRA initiative. ROS is supported by NSF AST-1412449, PHY-1505629, and PHY-1607520. JAC and JCB are supported by NSF PHY-1505824, PHY-1505524, and PHY-1333360. PK gratefully acknowledges support for this research at CITA from NSERC of Canada, the Ontario Early Researcher Awards Program, the Canada Research Chairs Program, and the Canadian Institute for Advanced Research. JV was supported by the UK STFC grant ST/K005014/1. ROS acknowledges the hospitality of the Aspen Center for Physics, supported by NSF PHY-1607611, where this work was completed. The authors thank Katerina Chatziioannou for comments on the manuscript. We note that after our limited investigations on the effect of priors on GW170104 was performed, a thorough study on the effect of priors was performed by Vitale et al [47]. The authors thank the

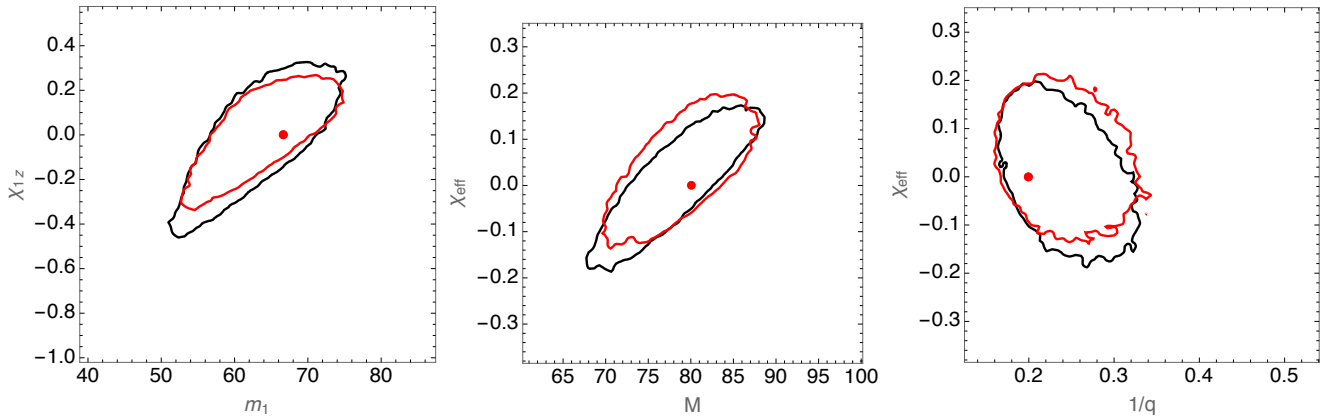

FIG. 8: **Omitting higher modes: Unbiased parameter inference, despite a high mismatch**: Red dot shows the parameters of a synthetic nonprecessing binary, whose radiation is modeled with a numerical relativity simulation SXS BBH:0112. All simulated modes $\ell \leq 8$ are included in our synthetic signal. The detector response is calculated assuming a source with total mass $80M_\odot$ oriented at angle $\theta_{\mathbf{JN}} = \pi/4$, at a distance so the network SNR is 20. No synthetic noise is added. For this source, the best match with the IMRPhenomPv2 and SEOBNRv3 approximations is $\simeq 0.96$. The black and red curves shows the 90% confidence interval derived from a detailed parameter inference calculation using the IMRPhenomD and SEOBNRv4 approximations, respectively. Calculations are performed using a network of detectors whose noise power spectra are identical to the estimates derived for GW150914 [62].

LIGO Scientific Collaboration for access to the data and gratefully acknowledge the support of the United States National Science Foundation (NSF) for the construction and operation of the LIGO Laboratory and Advanced LIGO as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, and the Max-Planck-Society (MPS) for support of the construction of Advanced LIGO. Additional support for Advanced LIGO was provided by the Australian Research Council. Extensive use was made of the `Matplotlib` [96], and `NumPy/SciPy` [97] software packages throughout the course of this work.

We are grateful for computational resources used for the parameter estimation and match calculations provided by Cardiff University, funded by an STFC grant supporting UK Involvement in the Operation of Advanced LIGO, and by California Institute of Technology at Pasadena, California.

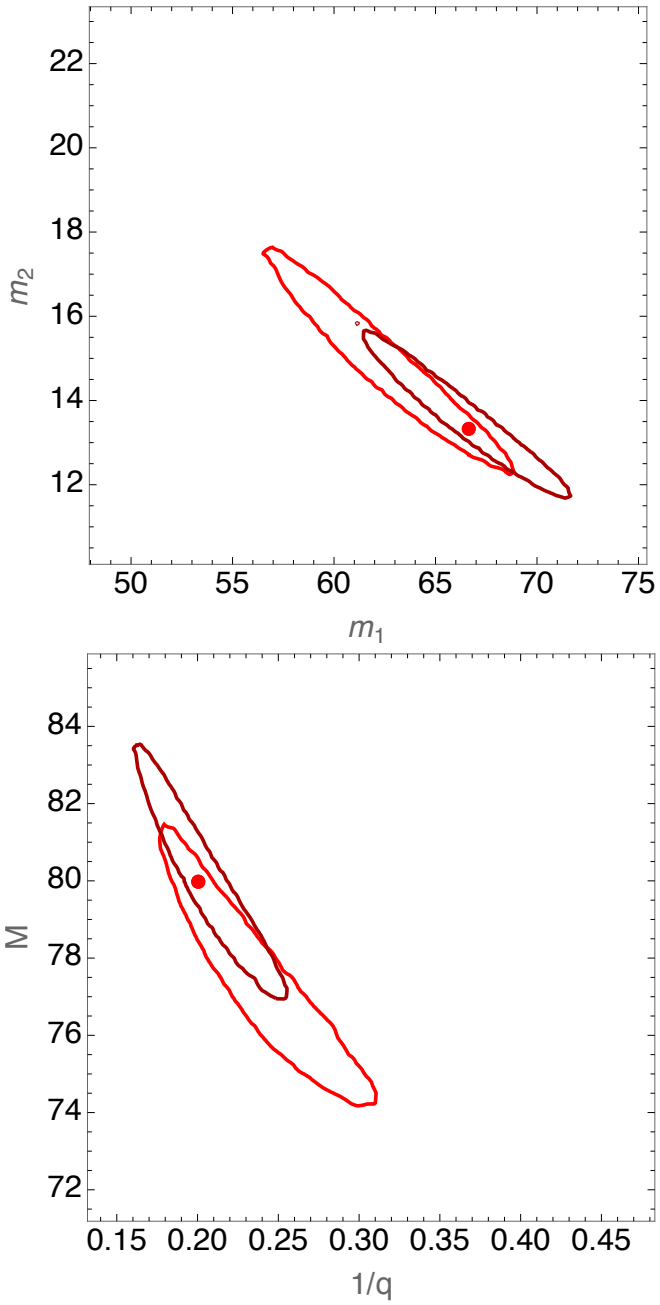

FIG. 9: **Parameter recovery with and without higher modes (assuming zero spin)**: Red dot shows the parameters of a synthetic nonprecessing binary, generated as in Figure 8. The dark red contour shows inference using EOBNRv2HM (a nonspinning model including higher modes); the light red contour shows parameter inferences drawn using SEOBNRv4, assuming both BHs have zero spin. The former region is smaller than the latter, and more closely centered on the true parameters. This figure illustrates the previously-appreciated fact that inference including higher modes draws sharper conclusions with smaller biases, using the examples previously used in this work.

[1] J. Aasi et al. (LIGO Scientific), Class. Quant. Grav. **32**, 074001 (2015), 1411.4547.

[2] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, et al., Physical Review Letters **116**, 061102 (2016), 1602.03837.

[3] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, et al., Physical Review Letters **116**, 241103 (2016), 1606.04855.

[4] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, et al., Physical Review Letters **118**, 221101 (2017), 1706.01812.

[5] B. P. Abbott et al. (Virgo, LIGO Scientific) (2017), 1711.05578.

[6] B. P. Abbott et al. (Virgo, LIGO Scientific), Phys. Rev. Lett. **119**, 141101 (2017), 1709.09660.

[7] F. Acernese et al. (VIRGO), Class. Quant. Grav. **32**, 024001 (2015), 1408.3978.

[8] B. Abbott et al. (The LIGO Scientific Collaboration and the Virgo Collaboration), Phys. Rev. X **6**, 041015 (2016).

[9] B. Abbott et al. (The LIGO Scientific Collaboration and the Virgo Collaboration), Astrophysical Journal **818**, L22 (2016), 1602.03846, URL `https://dcc.ligo.org/LIGO-P1500262/public/main`.

[10] K. Belczynski, D. Holz, T. Bulik, and R. O'Shaughnessy, Nature **534**, 512 (2016), URL `http://www.nature.com/nature/journal/v534/n7608/full/nature18322.html`.

[11] K. Belczynski, M. Dominik, T. Bulik, R. O'Shaughnessy, C. L. Fryer, and D. E. Holz, Astrophysical Journal **715**, L138 (2010), URL `http://xxx.lanl.gov/abs/arXiv:1004.0386`.

[12] R. O'Shaughnessy, R. Kopparapu, and K. Belczynski, Class. Quant. Grav. **29**, 145011 (2012), URL `http://xxx.lanl.gov/abs/0812.0591`.

[13] M. Dominik, K. Belczynski, C. Fryer, D. E. Holz, E. Berti, T. Bulik, I. Mandel, and R. O'Shaughnessy, Astrophysical Journal **759**, 52 (2012).

[14] I. Mandel and S. E. de Mink, MNRAS **458**, 2634 (2016).

[15] P. Marchant, N. Langer, P. Podsiadlowski, T. Tauris, and T. Moriya, A&A **588**, A50 (2016), URL `http://xxx.lanl.gov/abs/arXiv:1601.03718`.

[16] C. L. Rodriguez, C.-J. Haster, S. Chatterjee, V. Kalogera, and F. A. Rasio, Astrophysical Journal **824**, L8 (2016).

[17] S. Bird, I. Cholis, J. B. Muñoz, Y. Ali-Haïmoud, M. Kamionkowski, E. D. Kovetz, A. Raccanelli, and A. G. Riess, Phys. Rev. Lett **116**, 201301 (2016).

[18] R. O'Shaughnessy, presentation at Syracuse University (2009), URL `http://ccrg.rit.edu/~oshaughn/Talks/2009-05-29-Syracuse.pdf`.

[19] C. L. Rodriguez, M. Zevin, C. Pankow, V. Kalogera, and F. A. Rasio, Astrophysical Journal **832**, L2 (2016), 1609.05916, URL `http://xxx.lanl.gov/abs/arXiv:1609.05916`.

[20] S. Vitale, R. Lynch, R. Sturani, and P. Graff, Classical and Quantum Gravity **34**, 03LT01 (2017), 1503.04307, URL `http://xxx.lanl.gov/abs/arXiv:1503.04307`.

[21] S. Stevenson, C. P. L. Berry, and I. Mandel, ArXiv e-prints (2017), 1703.06873.

[22] R. O'Shaughnessy, D. Gerosa, and D. Wysocki, Phys. Rev. Lett. **119**, 011101 (2017), 1704.03879.

[23] C. Talbot and E. Thrane, ArXiv e-prints (2017), 1704.08370.

[24] V. Kalogera, Astrophysical Journal **541**, 319 (2000), astro-ph/9911417.

[25] S. Vitale, R. Lynch, R. Sturani, and P. Graff, Classical and Quantum Gravity **34**, 03LT01 (2017), 1503.04307.

[26] S. Stevenson, C. P. L. Berry, and I. Mandel, ArXiv e-prints (2017), 1703.06873.

[27] W. M. Farr, S. Stevenson, M. C. Miller, I. Mandel, B. Farr, and A. Vecchio, ArXiv e-prints (2017), 1706.01385.

[28] D. Wysocki, D. Gerosa, R. O'Shaughnessy, K. Belczynski, and et al, In preparation (2017).

[29] K. Breivik, C. L. Rodriguez, S. L. Larson, V. Kalogera, and F. A. Rasio, Astrophysical Journal **830**, L18 (2016), 1606.09558.

[30] A. Nishizawa, E. Berti, A. Klein, and A. Sesana, Phys. Rev. D **94**, 064020 (2016), 1605.01341.

[31] A. Taracchini, A. Buonanno, Y. Pan, T. Hinderer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, A. H. Mroué, H. P. Pfeiffer, et al., Phys. Rev. D **89**, 061502 (2014), 1311.2544.

[32] Y. Pan, A. Buonanno, A. Taracchini, L. E. Kidder, A. H. Mroué, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi, Phys. Rev. D **89**, 084006 (2014), 1307.6232.

[33] P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, J. T. Whelan, B. Brügmann, P. Diener, J. Gonzalez, M. Hannam, et al., Classical and Quantum Gravity **24**, 689 (2007), 0704.3764, URL `http://xxx.lanl.gov/abs/arxiv:0704.3764`.

[34] L. Santamaría, F. Ohme, P. Ajith, B. Brügmann, N. Dorband, M. Hannam, S. Husa, P. Mösta, D. Pollney, C. Reisswig, et al., Phys. Rev. D **82**, 064016 (2010), 1005.3306, URL `http://xxx.lanl.gov/abs/arXiv:1005.3306`.

[35] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, Phys. Rev. Lett **113**, 151101 (2014), 1308.3271.

[36] J. Blackman, S. E. Field, C. R. Galley, B. Szilágyi, M. A. Scheel, M. Tiglio, and D. A. Hemberger, Physical Review Letters **115**, 121102 (2015), 1502.07758.

[37] A. Taracchini, Y. Pan, A. Buonanno, E. Barausse, M. Boyle, T. Chu, G. Lovelace, H. P. Pfeiffer, and M. A. Scheel, Phys. Rev. D **86**, 024011 (2012), 1202.0790.

[38] B. Abbott et al. (The LIGO Scientific Collaboration and the Virgo Collaboration), Phys. Rev. Lett **116**, 241102 (2016), URL `http://link.aps.org/doi/10.1103/PhysRevLett.116.241102`.

[39] B. Abbott et al. (The LIGO Scientific Collaboration and the Virgo Collaboration), Phys. Rev. X **6**, 041014 (2016), URL `http://adsabs.harvard.edu/abs/2016arXiv160601210T`.

[40] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, D. A. Hemberger, P. Schmidt, and R. Smith, Phys. Rev. D **95**, 104023 (2017), 1701.00550.

[41] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, C. D. Ott, M. Boyle, L. E. Kidder, H. P. Pfeiffer, and B. Szilágyi, ArXiv e-prints (2017), 1705.07089.

[42] I. Hinder, A. Buonanno, M. Boyle, Z. B. Etienne,

J. Healy, N. K. Johnson-McDaniel, A. Nagar, H. Nakano, Y. Pan, H. P. Pfeiffer, et al., Classical and Quantum Gravity **31**, 025012 (2013), 1307.5307.

[43] P. Kumar, K. Barkett, S. Bhagwat, N. Afshari, D. A. Brown, G. Lovelace, M. A. Scheel, and B. Szilágyi, Phys. Rev. D **92**, 102001 (2015), 1507.00103.

[44] P. Kumar, T. Chu, H. Fong, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, M. A. Scheel, and B. Szilagyi, Phys. Rev. D **93**, 104050 (2016), 1601.05396.

[45] K. Chatziioannou, N. Cornish, A. Klein, and N. Yunes, Astrophysical Journal **798**, L17 (2015), 1402.3581.

[46] B. Farr, C. P. L. Berry, W. M. Farr, C.-J. Haster, H. Middleton, K. Cannon, P. B. Graff, C. Hanna, I. Mandel, C. Pankow, et al., Astrophysical Journal **825**, 116 (2016), 1508.05336.

[47] S. Vitale, D. Gerosa, C.-J. Haster, K. Chatziioannou, and A. Zimmerman, ArXiv e-prints (2017), 1707.04637.

[48] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, et al., Classical and Quantum Gravity **34**, 104002 (2017), 1611.07531, URL `https://arxiv.org/abs/1611.07531`.

[49] B. Abbott et al. (The LIGO Scientific Collaboration and the Virgo Collaboration), Phys. Rev. D **94**, 064035 (2016), URL `http://link.aps.org/doi/10.1103/PhysRevD.94.064035`.

[50] J. Healy, J. Lange, A. R. Williamson, M. Campanelli, C. Lousto, R. O'Shaughnessy, Y. Zlochower, J. Clark, J. Calderón Bustillo, B. Khamesra, et al., in preparation (2017).

[51] C. M. Will and A. G. Wiseman, Phys. Rev. D **54**, 4813 (1996), gr-qc/9608012.

[52] L. E. Kidder, Phys. Rev. D **52**, 821 (1995), arXiv:gr-qc/9506022.

[53] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, Phys. Rev. D **49**, 6274 (1994).

[54] Y. Pan, A. Buonanno, Y. Chen, and M. Vallisneri, Phys. Rev. D **69**, 104017 (2004), URL `http://xxx.lanl.gov/abs/gr-qc/0310034`.

[55] P. Schmidt, M. Hannam, S. Husa, and P. Ajith, Phys. Rev. D **84**, 024046 (2011), 1012.2879.

[56] L. Pekowsky, R. O'Shaughnessy, J. Healy, and D. Shoemaker, Phys. Rev. D **88**, 024040 (2013), URL `http://link.aps.org/doi/10.1103/PhysRevD.88.024040`.

[57] M. Boyle, Phys. Rev. D **87**, 104006 (2013), 1302.2919.

[58] R. O'Shaughnessy, L. London, J. Healy, and D. Shoemaker, Phys. Rev. D **87**, 044038 (2013), 1209.3712.

[59] S. Ossokine, M. Boyle, L. E. Kidder, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi, Phys. Rev. D **92**, 104028 (2015), 1502.01747.

[60] S. Babak, A. Taracchini, and A. Buonanno, Phys. Rev. D **95**, 024010 (2017), 1607.05661.

[61] J. Veitch, V. Raymond, B. Farr, W. M. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, et al., Phys. Rev. D **91**, 042003 (2015), URL `http://link.aps.org/doi/10.1103/PhysRevD.91.042003`.

[62] B. P. Abbott et al. (Virgo, LIGO Scientific), Phys. Rev. **D94**, 064035 (2016), 1606.01262.

[63] D. W. Hogg, A. D. Myers, and J. Bovy, Astrophysical Journal **725**, 2166 (2010), 1008.4146.

[64] T. Damour, B. R. Iyer, and B. S. Sathyaprakash, Phys. Rev. D **57**, 885 (1998), gr-qc/9708034.

[65] R. Storn and K. Price, Journal of Global Optimization **11**, 341 (1997), ISSN 1573-2916, URL `https://doi.org/10.1023/A:1008202821328`.

[66] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, et al., Phys. Rev. D **91**, 042003 (2015), 1409.7215.

[67] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, et al., Physical Review Letters **116**, 241102 (2016), 1602.03840.

[68] R. O'Shaughnessy, B. Farr, E. Ochsner, H.-S. Cho, C. Kim, and C.-H. Lee, Phys. Rev. D **89**, 064048 (2014), 1308.4704.

[69] P. B. Graff, A. Buonanno, and B. S. Sathyaprakash, Phys. Rev. D **92**, 022002 (2015), 1504.04766.

[70] B. Miller, R. O'Shaughnessy, T. B. Littenberg, and B. Farr, Phys. Rev. D **92**, 044056 (2015), 1506.06032.

[71] M. Agathos, J. Meidam, W. Del Pozzo, T. G. F. Li, M. Tompitak, J. Veitch, S. Vitale, and C. Van Den Broeck, Phys. Rev. **D92**, 023012 (2015), 1503.05405.

[72] D. Trifirò, R. O'Shaughnessy, D. Gerosa, E. Berti, M. Kesden, T. Littenberg, and U. Sperhake, Phys. Rev. D **93**, 044071 (2016), 1507.05587.

[73] R. O'Shaughnessy, J. Blackman, and S. E. Field, Classical and Quantum Gravity **34**, 144002 (2017), 1701.01137.

[74] A. H. Mroué, M. A. Scheel, B. Szilágyi, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, S. Ossokine, N. W. Taylor, et al., Physical Review Letters **111**, 241104 (2013), 1304.6077.

[75] P. Schmidt, I. W. Harry, and H. P. Pfeiffer (2017), 1703.01076.

[76] C. R. Galley and P. Schmidt, ArXiv e-prints (2016), 1611.07529.

[77] V. Varma, P. Ajith, S. Husa, J. C. Bustillo, M. Hannam, and M. Pürrer, Phys. Rev. D **90**, 124004 (2014), 1409.2349.

[78] J. Calderón Bustillo, S. Husa, A. M. Sintes, and M. Pürrer, Phys. Rev. D **93**, 084019 (2016), 1511.02060.

[79] V. Varma and P. Ajith, ArXiv e-prints (2016), 1612.05608.

[80] J. Calderón Bustillo, P. Laguna, and D. Shoemaker, Phys. Rev. D **95**, 104038 (2017), 1612.02340.

[81] A. Bohé et al., Phys. Rev. **D95**, 044028 (2017), 1611.03703.

[82] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, Phys. Rev. **D93**, 044007 (2016), 1508.07253.

[83] D. A. Brown, A. Lundgren, and R. O'Shaughnessy, Phys. Rev. D **86**, 064020 (2012), 1203.6060, URL `http://arxiv.org/abs/1203.6060`.

[84] A. Lundgren and R. O'Shaughnessy, Phys. Rev. D **89**, 044021 (2014), URL `http://link.aps.org/doi/10.1103/PhysRevD.89.044021`.

[85] R. O'Shaughnessy, P. Nepal, and A. Lundgren, (arXiv:1509.06581) (2015), URL `http://xxx.lanl.gov/abs/arXiv:1509.06581`.

[86] R. O'Shaughnessy, B. Vaishnav, J. Healy, and D. Shoemaker, Phys. Rev. D **82**, 104006 (2010), 1007.4213.

[87] C. Capano, Y. Pan, and A. Buonanno, Phys. Rev. D **89**, 102003 (2014), 1311.1286.

[88] C. Van Den Broeck and A. S. Sengupta, Classical and Quantum Gravity **24**, 155 (2007), gr-qc/0607092.

[89] R. N. Lang and S. A. Hughes, Phys. Rev. D **74**, 122001 (2006), gr-qc/0608062.

[90] A. Klein, P. Jetzer, and M. Sereno, Phys. Rev. D **80**, 064027 (2009), 0907.3318.
[91] E. K. Porter and N. J. Cornish, Phys. Rev. D **78**, 064005 (2008), 0804.0332.
[92] R. N. Lang, S. A. Hughes, and N. J. Cornish, Phys. Rev. D **84**, 022002 (2011), 1101.3591.
[93] E. Racine, Phys. Rev. **D78**, 044021 (2008), 0803.1820.
[94] P. Ajith et al., Phys. Rev. Lett. **106**, 241101 (2011), 0909.2867.
[95] C. Devine, Z. B. Etienne, and S. T. McWilliams, Class. Quant. Grav. **33**, 125025 (2016), 1601.03393.
[96] J. D. Hunter, Computing In Science & Engineering **9**, 90 (2007).
[97] S. van der Walt, S. C. Colbert, and G. Varoquaux, Computing in Science & Engineering **13**, 22 (2011).