Directly comparing GW150914 with numerical solutions of Einstein’s equations for binary black hole coalescence

B. P. Abbott et al.

We compare GW150914 directly to simulations of coalescing binary black holes in full general relativity, including several performed specifically to reproduce this event. Our calculations go beyond existing semianalytic models, because for all simulations – including sources with two independent, precessing spins – we perform comparisons which account for all the spin-weighted quadrupolar modes, and separately which account for all the quadrupolar and octopolar modes. Consistent with the posterior distributions reported in LVC-PE[1] (at the 90% credible level), we find the data are compatible with a wide range of nonprecessing and precessing simulations. Followup simulations performed using previously-estimated binary parameters most resemble the data, even when all quadrupolar and octopolar modes are included. Comparisons including only the quadrupolar modes constrain the total redshifted mass \( M_\text{r} \in [64M_\odot - 82M_\odot] \), mass ratio \( 1/q = m_2/m_1 \in [0.6, 1] \), and effective aligned spin \( \chi_{\text{eff}} \in [-0.3, 0.2] \), where \( \chi_{\text{eff}} = (S_1/m_1 + S_2/m_2) \cdot \hat{\mathbf{L}}/M \). Including both quadrupolar and octopolar modes, we find the mass ratio is even more tightly constrained. Even accounting for precession, simulations with extreme mass ratios and effective spins are highly inconsistent with the data, at any mass. Several nonprecessing and precessing simulations with similar mass ratio and \( \chi_{\text{eff}} \) are consistent with the data. Though correlated, the components’ spins (both in magnitude and directions) are not significantly constrained by the data: the data is consistent with simulations with component spin magnitudes \( a_{1,2} \) up to at least 0.8, with random orientations. Further detailed followup calculations are needed to determine if the data contain a weak imprint from transverse (precessing) spins. For nonprecessing binaries, interpolating between simulations, we reconstruct a posterior distribution consistent with previous results. The final black hole’s redshifted mass is consistent with \( M_\text{r} = 64.0M_\odot \sim 73.5M_\odot \) and the final black hole’s dimensionless spin parameter is consistent with \( a_f = 0.62 - 0.73 \). As our approach invokes no intermediate approximations to general relativity and can strongly reject binaries whose radiation is inconsistent with the data, our analysis provides a valuable complement to LVC-PE[1].

I. INTRODUCTION

On September 14, 2015 09:50:45 UTC, gravitational waves were observed in coincidence by the twin instruments of the Laser Interferometer Gravitational-wave Observatory (LIGO) located at Hanford, Washington, and Livingston, Louisiana, in the USA, an event known as GW150914[2]. LVC-detect[2], LVC-PE[1], LVC-TestGR[3], and LVC-Burst[4] demonstrate consistency between GW150914 and selected individual predictions for a binary black hole coalescence, derived using numerical solutions of Einstein’s equations for general relativity. LVC-PE[1] described a systematic, Bayesian method to reconstruct the properties of the coalescing binary, by comparing the data with the expected gravitational wave signature from binary black hole coalescence [5], evaluated using state-of-the-art semianalytic approximations to its dynamics and radiation [6,8].

In this paper, we present an alternative method of reconstructing the binary parameters of GW150914, without using the semianalytic waveform models employed in LVC-PE[1]. Instead, we compare the data directly with the most physically complete and generic predictions of general relativity: computer simulations of binary black hole coalescence in full nonlinear general relativity (henceforth referred to as numerical relativity, or NR). Although the semianalytic models are calibrated to NR simulations, even the best available models only imperfectly reproduce the predictions of numerical relativity, on a mode-by-mode basis [9]. Furthermore, typical implementations of these models, such as those used in LVC-PE[1], consider only the dominant spherical-harmonic mode of the waveform (in a corotating frame). For all NR simulations considered here—including sources with two independent, precessing spins—we perform comparisons that account for all the quadrupolar spherical-harmonic waveform modes, and separately comparisons that account for all the quadrupolar and octopolar spherical harmonic modes.

The principal approach introduced in this paper is different from LVC-PE[1], which inferred the properties of GW150914 by adopting analytic waveform models. Qualitatively speaking, these models interpolate the outgoing gravitational wave strain (waveforms) between the well-characterized results of numerical relativity, as provided by a sparse grid of simulations. These interpolated or analytic waveforms are used to generate a continuous posterior distribution over the binary’s parameters. By contrast, in this study, we compare numerical relativity to the data first, evaluating a single scalar quantity (the marginalized likelihood) on the grid of binary parameters prescribed and provided by all available NR simulations. We then construct an approximation to the marginalized likelihood that interpolates between NR simulations with different parameters. To the extent that the likelihood is a simpler function of parameters than the waveforms, this method may require fewer NR simulations and fewer modeling assumptions. Moreover, the interpolant for the likelihood needs to be accurate only near its peak value, and not everywhere in parameter space. A similar study was conducted on GW150914 using a
II. MOTIVATION FOR THIS STUDY

This paper presents an alternative analysis of GW150914 and an alternative determination of its intrinsic parameters. The methods used here differ from those in LVC-PE[1] in two important ways. First, the statistical analysis here is performed in a manner different than and independent of the one in LVC-PE[1]. Second, the gravitational waveform models used in LVC-PE[1] are analytic approximations of particular functional forms, with coefficients calibrated to match selected NR simulations; in contrast, here we directly use waveforms from NR simulations. Despite these differences, our conclusions largely corroborate the quantitative results found in LVC-PE[1].

Our study also addresses key challenges associated with gravitational wave parameter estimation for black hole binaries with total mass $M > 50M_\odot$. In this mass regime, LIGO is sensitive to the last few dozens of cycles of coalescence, a strongly nonlinear epoch that is the most difficult to approximate with analytic (or semi-analytic) waveform models [6, 11, 13, 14]. Figure 1 illustrates the dynamics and expected detector response in this regime, for a source like GW150914. For these last dozen cycles, existing analytic waveform models have only incomplete descriptions of precession, lack higher-order spherical-harmonic modes, and do not fully account for strong nonlinearities. Preliminary investigations have shown that inferences about the source drawn using these existing analytic approximations can be slightly or significantly biased [9, 13, 15, 16]. Systematic studies are underway to assess how these approximations influence our best estimates of a candidate binary’s parameters. At present, we can only summarize the rationale for these investigations, not their results. To provide three concrete examples of omitted physics, first and foremost, even the most sophisticated models for binary black hole coalescence [6] do not yet account for the asymmetries [15] responsible for the largest gravitational-wave recoil kicks [16, 17]. Second, the analytic waveform models adopted in LVC-PE[1] adopted simple spin treatments (e.g., a binary with aligned spins, or a binary with single effective precessing spin) that cannot capture the full spin dynamics [18–20]. A single precessing spin is often a good approximation, particularly for unequal masses where one spin dominates the angular momentum budget [6, 18, 21, 22]. However, for appropriate comparable-mass sources, two-spin effects are known to be observationally accessible [15, 24] and cannot be fully captured by a single spin. Finally, LVC-PE[1] and LVC-TestGR [3] made an additional approximation to connect the inferred properties of the binary black hole with the final black hole state [25, 26]. The analysis presented in this work does not rely on these approximations: observational data are directly compared against a wide range of NR simulations and the final black hole properties are extracted directly from these simulations, without recourse to estimated relationships between the initial and final state. By circumventing these approximations, our analysis can corroborate conclusions about selected physical properties of GW150914 presented in those papers.

Despite the apparent simplicity of GW150914, we find that a range of binary black hole masses and spins, including strongly precessing systems with significant misaligned black hole spins [27], are reasonably consistent with the data. The reason the data cannot distinguish between sources with qualitatively different dynamics is a consequence of both the orientation of the source relative to the line of sight and the timescale of GW150914. First, if the line of sight is along or opposite the total angular momentum vector of the source, even the most strongly precessing black hole binary emits a weakly-modulated inspiral signal, lacking unambiguous signatures of precession and easily mistaken for a nonprecessing binary [24, 28]. Second, because GW150914 has a large total mass, very little of the inspiral lies in LIGO’s frequency band, so the signal is short, with few orbital cycles and even fewer precession cycles prior to or during coalescence. The short duration of GW150914 provides few opportunities for the dynamics of two precessing spins to introduce distinctive amplitude and phase modulations into its gravitational wave inspiral signal [13, 14].

Although the orientation of the binary and the short du-
ration of the signal make it difficult to extract spin information from the \textit{inspiral}, comparable-mass binaries with large spins can have exceptionally rich dynamics with waveform signatures that extend into the late inspiral and the strong-field merger phase [8, 29]. By utilizing full NR waveforms instead of the single-spin (precessing) and double-spin (nonprecessing) models applied in LVC-PE[1], the approach described here provides an independent opportunity to extract additional insight from the data, or to independently corroborate the results of LVC-PE[1].

Our study employs a simple method to carry out our Bayesian calculations: for each NR simulation, we evaluate the marginalized likelihood of the simulation parameters given the data. The likelihood is evaluated via an adaptive Monte Carlo integrator. This method provides a quantitative ranking of simulations; with judicious interpolation in parameter space, the method also allows us to identify candidate parameters for followup numerical relativity simulations. To estimate parameters of GW150914, we can simply select the subset of simulations and masses that have a marginalized likelihood greater than an observationally-motivated threshold (i.e., large enough to contribute significantly to the posterior). Even better, with a modest approximation to fill the gaps between NR simulations, we can reproduce and corroborate the results in LVC-PE[1] with a completely independent method. We explicitly construct an approximation to the likelihood that interpolates between simulations of precessing binaries, and demonstrate its validity and utility.

It is well-known that the choice of prior may influence conclusions of Bayesian studies when the data do not strongly constrain the relevant parameters. For example, the results of LVC-PE[1] suggest that GW150914 had low to moderate spins, but this is due partly to the conventional prior used in LVC-PE[1] and earlier studies [5]. This prior is uniform in spin magnitude, and therefore unfavorable to the most dynamically interesting possibilities: comparable-mass binaries with two large, dynamically-significant precessing spins [24]. In contrast, by directly considering the (marginalized) likelihood, the results of our study are independent of the choice of prior. For example, we find here that GW150914 is consistent with two large, dynamically significant spins.

Finally, our efforts to identify even subtle hints of spin precession are motivated by the astrophysical opportunities afforded by spin measurements with GW150914; see LVC-Astro[30]. Using the geometric spin prior adopted in LVC-PE[1], the data from GW150914 are just as consistent with an origin from a nonprecessing or precessing binary, as long as as the sum of the components of the spins parallel to the orbital angular momentum \( \mathbf{L} \) is nearly zero. If the binary black hole formed from isolated stellar evolution, one could reasonably expect all angular momenta to be strictly and positively aligned at coalescence; see LVC-Astro[30] and [31]. Hence, if we believe GW150914 formed from an isolated binary, our data would suggest black holes are born with small spins: \( a_1 = |S_1|/m_1^2 \leq 0.22 \) and \( a_2 \leq 0.25 \), where \( S_1 \) and \( m_1 \) are the black hole spins and masses [LVC-PE[1]]. If these strictly-aligned isolated evolution formation scenarios are true, then a low black hole spin constrains the relevant accretion, angular momentum transport, and tidal interaction processes in the progenitor binary; cf. [31-33]. On the other hand, the data are equally consistent with a strongly-precessing black hole binary with large component spins, formed in a densely interacting stellar cluster (LVC-Astro[30]). Measurements of the binary black holes’ transverse spins will therefore provide vital clues as to the processes that formed GW150914. In this work we use numerical relativity to check for any evidence for or against spin precession that might otherwise be obscured by model systematics. Like LVC-PE[1], we find results consistent with but with no strong support for precessing spins.

III. METHODS

A. Numerical relativity simulations of binary black hole coalescence

The first attempts to solve the field equations of general relativity numerically began in the 1960s, by Hahn and Lindquist [33], followed with some success by Smarr [35, 36]. In the 1990s, a large collaboration of US universities worked together to solve the “Grand Challenge” of evolving binary black holes [37, 39]. In 2005, three groups [40-42] developed two completely independent techniques that produced the first collisions of orbiting black holes. The first technique [40] involved the use of generalized harmonic coordinates and excision of the black hole horizons, while the second technique [41, 42], dubbed “moving punctures approach”, used singularity avoiding slices of the black hole spacetimes.

Since then, the field has seen an explosion of activity and improvements in methods and capabilities; see, e.g., [43-46]. Multiple approaches have been pursued and validated against one another [47, 48]. Binaries can now be evolved in wide orbits [49]; at high mass ratios up to 100:1 [50, 51]; with near-maximal black hole spin [52, 54]; and for many orbits before coalescence [27, 55]. At sufficiently large separations, despite small gauge and frame ambiguities, the orbital and spin dynamics evaluated using numerical relativity agrees with post-Newtonian calculations [6, 56-59]. The stringent phase and amplitude needs of gravitational wave detection and parameter estimation prompted the development of revised standards for waveform accuracy [60, 61]. Several projects have employed numerical relativity-generated waveforms to assess gravitational-wave detection and parameter estimation strategies [11, 62-66]. These results have been used to calibrate models for the leading-order radiation emitted from binary black hole coalescence [6, 8, 9, 12, 57, 67, 68]. Our study builds on this past decade’s experience with modeling the observationally-relevant dynamics and radiation from comparable-mass coalescing black holes.

In this and most NR work, the initial data for a simulation of coalescing binaries are characterized by the properties and initial orbit of its two component black holes: by initial black holes masses \( m_1, m_2 \) and spins \( S_1, S_2 \), specified in a quasicircular orbit such that the (coordinate) orbital angular momentum is aligned with the \( \hat{z} \) axis and the initial separation is along the \( \hat{x} \) axis. In this work, we characterize these simulations by
the dimensionless mass ratio \( q = m_1/m_2 \) (adopting the convention \( m_1 \geq m_2 \)); the dimensionless spin parameters \( \chi_i = S_i/m_i^2 \); and an initial dimensionless orbital frequency \( M_0 \omega_0 \). For each simulation, the orientation-dependent gravitational wave strain \( h(t, r, \hat{n}) \) at large distances can be efficiently characterized by a (spin-weighted) spherical harmonic decomposition of \( h(t, r, \hat{n}) \) as \( h(t, r, \hat{n}) = \sum_{l \geq 2} \sum_{m=-l}^{l} h_{lm}(t, r) Y_{lm}(\hat{n}) \). To a good first approximation, only a few terms in this series are necessary to characterize the observationally-accessible radiation in any direction [14, 69, 72]. For example, when a binary is widely separated, only two terms dominate this sum: \( (l, |m|) = (2, 2) \). Conversely, however, several terms (modes) are required for even nonprecessing binaries, viewed along a generic line of sight; more are needed to capture the radiation from precessing binaries.

For nonprecessing sources with a well-defined axis of symmetry, individual modes \( (l, m) \) have distinctive characters, and can be easily isolated numerically and compared with analytic predictions. For precessing sources, however, rotation mixes modes with the same \( l \). To apply our procedure self-consistently to both nonprecessing and precessing sources, we include all modes \( (l, m) \) with the same \( l \). However, at the start of each simulation, the \( (l, m) \) mode oscillates at \( m \) times the orbital frequency. For \( m > 3 \), scaling our simulations to the inferred mass of the source, this initial mode frequency is often well above 30 Hz, the minimum frequency we adopt in this work for parameter estimation. We therefore cannot safely and self-consistently compare all modes with \( l > 3 \) to the data using numerical relativity alone: an approximation would be required to go to higher order (i.e., hybridizing each NR simulation with an analytic approximation at early times).

Therefore, in this paper, we use all five of the \( l = 2 \) modes to draw conclusions about GW150914, necessary and sufficient to capture the leading-order influence of any orbital precession. To incorporate the effect of higher harmonics, we repeat our calculations, using all of the \( l \leq 3 \) modes. We defer a careful treatment of higher-order modes and the \( m = 0 \) modes to PE+NR-Methods [10] and subsequent work.

B. Numerical relativity simulations used

Our study makes use of 1139 distinct simulations of binary black hole quasicircular inspiral and coalescence. Table I summarizes the salient features of this set: mass ratio and initial spins for the simulations used here, all initially in a quasicircular orbit with orbital separation along the \( \hat{x} \) axis and velocities along \( \pm \hat{y} \).

The RIT group provided 394 simulations [17, 27, 73]. The simulations include binaries with a wide range of mass ratios, as well as a wide range of black hole angular momentum (spin) magnitudes and directions [17, 27, 73, 74, 74], including a simulation with large transverse spins and several spin precession cycles which fits GW150914 well [27], as described below. The SXS group has provided both a publicly-available catalog of coalescing black hole binary mergers [75], a new catalog of nonprecessing simulations [76], and selected supplementary simulations described below. Currently extended to 310 members in the form used here, this catalog includes many high-precision zero- and aligned-spin sources; selected precessing systems; and simulations including extremely high black hole spin. The Georgia Tech group (GT) provided 406 simulations; see [15] and [77] for further details. This extensive archive covers a wide range of spin magnitudes and orientations, including several systematic one- and two-parameter families. The Cardiff-UIB group provided 29 simulations, all specifically produced to follow up GW150914 via a high-dimensional grid stencil, performed via the BAM code [78, 79]. These four sets of simulations explore the model space near the event in a well-controlled fashion. In addition to previously-reported simulations, several groups performed new simulations (108 in total) designed to reproduce the parameters of the event, some of which were applied to our analysis. These simulations are denoted in Table I and our other reports by an asterisk (*). These followup simulations include three independent simulations of the same parameters drawn from the distributions in LVC-PE [1], from RIT, SXS, and GT, allowing us to assess our systematic error. These simulations were reported in LVC-detect [2] and LVC-Burst [4], and are indicated by (+) in our tables.

The simulations used here have either been published previously, or were produced using one of three well-tested procedures operating in familiar circumstances. For reference, in Appendix A we outline the three groups’ previously-established methods and results. For this application, we trust these simulations’ accuracy, based on their past track record of good performance. By incorporating simulations of identical physics provided by different groups, our methods provide limited direct corroboration: simulations with similar physics produce similar results.

C. Directly comparing NR with data

For each simulation, each choice of seven extrinsic parameters \( \theta \) (4 spacetime coordinates for the coalescence event; three Euler angles for the binary’s orientation relative to the Earth), and each choice for the redshifted total binary mass \( M_* = (1 + z)M \), we can predict the response \( h_{\ell k} \) of both of the \( k = 1, 2 \) LIGO instruments to the implied gravitational wave signal. Using \( \lambda \) to denote the combination of redshifted mass \( M_* \) and the numerical relativity simulation parameters needed to uniquely specify the binary’s dynamics, we can therefore evaluate the likelihood of the data given the noise:

\[
\ln \mathcal{L}(\lambda; \theta) = -\frac{1}{2} \sum_k \langle h_{\ell k}(\lambda, \theta) - d_{\ell k} h_{\ell k}(\lambda, \theta) - d_{\ell k} \rangle_k - \langle d_{\ell k} | d_{\ell k} \rangle_k \tag{1}
\]

where \( h_{\ell k} \) are the predicted response of the \( k \)th detector due to a source with parameters \( \lambda, \theta; d_{\ell k} \) are the detector data in instrument \( k \); and \( \langle a|b\rangle_k \equiv \int_0^\infty 2d/f(\omega)\tilde{f}(\omega)/S_{\omega k}(\omega) \) is an inner product implied by the \( k \)th’s detector’s noise power spectrum \( S_{\omega k}(\omega) \); see, e.g., [80] for more details. In practice, as discussed in the next section, we adopt a low-frequency cutoff
The gray line shows the idealized strain response $h(t) = F_s h_s(t) + F_h h_h(t)$, while the solid black line shows the whitened strain response, using the same noise power spectrum as LVC-detect[2].

Except for an overall normalization constant and a different choice for low-frequency cutoff, our expression agrees with Eq. (1) in LVC-PE[1].

The joint posterior probability of $\lambda, \theta$ follows from Bayes theorem:

$$p_{\text{post}}(\lambda, \theta) = \frac{\mathcal{L}(\lambda, \theta) p(\theta) p(\lambda)}{\int d\lambda d\theta \mathcal{L}(\lambda, \theta) p(\theta) p(\lambda)},$$

where $p(\theta)$ and $p(\lambda)$ are priors on the (independent) variables $\theta, \lambda$[1]. For each $\lambda$ — that is, for each simulation and each redshifted mass $M_z$ — we evaluate the marginalized likelihood

$$\mathcal{L}_{\text{marg}}(\lambda) \equiv \int \mathcal{L}(\lambda, \theta) p(\theta) d\theta$$

via direct Monte Carlo integration, where $p(\theta)$ is uniform in 4-volume and source orientation [80][2]. The marginalized likelihood measures the similarity between the data and a source with parameters $\lambda$ and enters naturally into full Bayesian posterior calculations. In terms of the marginalized likelihood and some assumed prior $p(\lambda)$ on intrinsic parameters like masses and spins, the posterior distribution for intrinsic parameters is

$$p_{\text{post}}(\lambda) = \frac{\mathcal{L}_{\text{marg}}(\lambda) p(\lambda)}{\int d\lambda \mathcal{L}_{\text{marg}}(\lambda) p(\lambda)}.$$

If we can evaluate $\mathcal{L}_{\text{marg}}$ on a sufficiently dense grid of intrinsic parameters, Eq. (5) implies that we can reconstruct the full posterior parameter distribution via interpolation or other local approximations. This reconstruction needs to accurately reproduce $\mathcal{L}_{\text{marg}}$ only near its peak value; for example, if $\mathcal{L}_{\text{marg}}(\lambda)$ can be approximated by a $d$-dimensional Gaussian, then we anticipate only configurations $\lambda$ with

$$\ln \mathcal{L}_{\text{max}} / \mathcal{L}_{\text{marg}}(\lambda) > \chi^2_{\text{d,s}} / 2$$

contribute to the posterior distribution at the $1-\epsilon$ credible interval, where $\chi^2_{\text{d,s}}$ is the inverse-$\chi^2$ distribution.

Based on similarity of our distribution to a suitably-parameterized multidimensional Gaussian, we anticipate that only the region of parameter space with $\ln \mathcal{L}_{\text{max}} - \ln \mathcal{L}_{\text{marg}}(\lambda) \lesssim 6.7$ can potentially impact our conclusions regarding the 90% credible level for $d = 8$ (i.e., two masses and two precessing spins); for $d = 4$, more relevant to the most strongly accessible parameters (i.e., two masses and two aligned spins), the corresponding interval is $\ln \mathcal{L}_{\text{max}} - \ln \mathcal{L}_{\text{marg}}(\lambda) \lesssim 4$.

Each NR simulation corresponds to a particular value of $d$ or $s$, of the intrinsic parameters (mass ratio and the three components of each spin vector) but can be scaled to an arbitrary value of the total redshifted mass $M$. Therefore each NR simulation represents a one-parameter family of points in the $d$-dimensional parameter space of all possible values of $\lambda$. For each simulation, we evaluate the marginalized log likelihood versus redshifted mass $\mathcal{L}_{\text{marg}}(M_z)$ on an array of masses, adaptively exploring each one-parameter family to cover the interval $\ln \mathcal{L}_{\text{max}} - \ln \mathcal{L}_{\text{marg}}(\lambda) < 10$. To avoid systematic bias introduced by interpolation or fitting, our principal results are simply these tabulated function values, explored almost continuously in mass $M_z$ and discretely, as our fixed simulation archive permits, in other parameters. The set of intrinsic parameters $V_C \equiv \{\lambda : \ln \mathcal{L}_{\text{marg}} > C\}$ above a cutoff $C$ identifies a subset of binary configurations whose gravitational wave emission is consistent with the data[2]. Though this approach provides a powerful model-independent approach to gravitational-wave parameter estimation, as described above it is restricted to the discrete grid of NR simulation values. Fortunately, the brevity and simplicity of the signal — only a few chirping and little-modulated cycles — requires the posterior

\[\text{In Figure 1. Simulated waveform: Predicted strain in H1 for a source with parameters } q = 1.22, y_1z = 0.33, y_2z = 0.44, \text{ simulated in full general relativity; compare to Figure 2 in LVC-detect[2]. The gray line shows the idealized strain response } h(t) = F_s h_s(t) + F_h h_h(t), \text{ while the solid black line shows the whitened strain response, using the same noise power spectrum as LVC-detect[2].}\\]
distribution to be broad and smooth, extending over many numerical relativity simulations’ parameters. This allows us to go beyond comparisons on a discrete grid of NR simulations, and instead interpolate between simulations to reconstruct the entire distribution.

To establish a sense of scale, we can use a simple order-of-magnitude calculation for \( \ln L_{\text{marg}} \). The signal to noise ratio \( \rho \) and peak likelihood of any assumed signal are related: \( \rho = \sqrt{2 \max \ln L} \). Even at the best intrinsic parameters \( \lambda \), the marginalized log-likelihood \( \ln L_{\text{marg}} \) will be well below the peak value \( \max \ln L \), because only a small fraction of extrinsic parameters \( \theta \) have support from the data \[82\]. Using GW150914’s previously-reported signal amplitude [\( \rho = 23.5\text{–}26.8 \)], its extrinsic parameters and their uncertainty \([\text{LVC-PE}[1]]\), and our prior \( p(\theta) \), we expect the peak value of \( \ln L_{\text{marg}} \) to be of order 240–330. The interval of \( \ln L_{\text{marg}} \) selected by Eq. (6) is a small fraction of the full range of \( \ln L_{\text{marg}} \), identifying a narrow range of parameters \( \lambda \) which are consistent with the data.

Our analysis of this event, as well as synthetic data, suggests that \( \ln L_{\text{marg}} \) is often well-approximated by simple low-order series in intrinsic parameters \( \lambda \). This simple behavior is most apparent versus total mass \( M_z \). Figure 2 shows examples of the marginalized log likelihood evaluated using two of our most promising simulation candidates; they are well-approximated by a quadratic over the entire observationally-interesting range. We approximate \( \ln L_{\text{marg}}(M_z) \) as a second-order Taylor series,

\[
\ln L_{\text{marg}}(M_z) \approx \ln L - \frac{1}{2} \Gamma_{\text{MM}}(M_z - M_{z,*})^2, \tag{7}
\]

where the constants \( \ln L, M_{z,*}, \) and \( \Gamma_{\text{MM}} \) represent the largest value of \( \ln L_{\text{marg}} \), the redshifted mass at which this maximum occurs, and the second derivative at the peak value. Even in (rare) cases when a locally quadratic approximation slightly breaks down, we still use \( \ln L \) to denote our estimate of the peak of \( \ln L_{\text{marg}}(M_z) \). As a means of efficiently communicating trends in the quality of fit versus intrinsic parameters, the two quantities \( \ln L \) and \( M_{z,*} \) are reported in Table III.

Motivated by the success of this approximation, in Section IV B we also supply a quadratic approximation to \( \ln L_{\text{marg}} \) near its peak, under the restrictive approximation that all angular momenta are parallel, using information from only non-precessing simulations. Using this quadratic approximation, we can numerically estimate \( \ln L_{\text{marg}} \) and hence the posterior [Eq. (5)] for arbitrary aligned-spin binaries. For any coordinate transformation \( z = Z(\lambda) \), we can use suitable supplementary coordinates and direct numerical quadrature to determine the marginal posterior density \( p_{\text{post}}(z) = \int p_{\text{post}}(\lambda) \delta(z - Z(\lambda)). \) As shown below, this procedure yields results comparable to LVC-PE[1] for nonprecessing binaries.

Because of finite computational resources, NR simulations of binary black holes cannot include an arbitrary number of orbits before merger. Instead, they start at some finite initial orbital frequency. While many NR simulations follow enough binary orbits to be compared with GW150914 over the entire LIGO frequency band, some NR simulations miss some early-time information. Therefore, in this section we describe a simple approximation (a low frequency cutoff) we apply to ensure that simulations with similar physics (but different initial orbital frequencies) lead to similar results.

At the time of GW150914, the instruments had relatively poor sensitivity to frequencies below 30 Hz and almost no sensitivity below 20 Hz. For this reason, the interpretations adopted in LVC-PE[1] adopted a low-frequency cutoff of 20 Hz. Because of the large number of cycles accumulated at low frequencies, a straightforward Fisher matrix estimate \[83, 84\] suggests these low frequencies (20–30 Hz) provide a nontrivial amount of information, particularly about the binary’s total mass. Equivalently, using the techniques described in this paper, the function \( \ln L_{\text{marg}}(M_z) \) will have a slightly higher and narrower peak when including all frequencies than when truncating the signal to only include frequencies above 30 Hz; see PE+NR-Methods[10].

Because of limited computational resources, relatively few simulations start in a sufficiently wide orbit such that, for \( M_z = 70M_\odot \), their radiation in the most significant harmonics of the orbital frequency will be at or below the lowest frequency (20 Hz) adopted in

---

4 We find similar results using more sophisticated nonparametric interpolation schemes. The results reported in Table III use one-dimensional Gaussian process interpolation to determine the peak value.
If $f_{\text{min}}$ is the low-frequency cutoff, $M_0 \omega_0/m \leq 0.02 (M_*/70 M_\odot)(f_{\text{min}}/20 \text{Hz}) (2/m)$, where $M_0 \omega_0$ is the initial orbital frequency of the simulation reported in Table III can be safely used to analyze a signal containing a significant contribution from the $m$th harmonic of the orbital frequency. Figure 3 shows examples of the strain in LIGO-Hanford, predicted using simulations of different intrinsic duration, superimposed with lines approximately corresponding to different gravitational wave frequencies. To facilitate an apples-to-apples comparison incorporating the widest range of available simulations, in this work we principally report on comparisons calculated by adopting a low-frequency cutoff of 30 Hz; see, e.g., Table III. We also briefly report on comparisons performed using a low-frequency cutoff of 10 Hz.) As we describe in subsequent sections, while this choice of 30 Hz slightly degrades our ability to make subtle distinctions between different precessing configurations, it does not dramatically impair our ability to reconstruct parameters of the event, given other significant degeneracies.

Even this generous low-frequency cutoff is not perfectly safe: for each simulation, a minimum mass exists at which the starting gravitational wave frequency is 30 Hz or larger. In the plots and numerical results reported here, we have eliminated simulation and mass choices that correspond to scaling an NR simulation to a starting frequency above 30 Hz. The inclusion or suppression of these configurations does not significantly change our principal results.

This paper uses enough NR simulations to adequately sample the four-dimensional space of nonprecessing spins, particularly for comparable masses. As described below, this high simulation density insures we can reliably approximate the marginalized likelihood $\log L_{\text{marg}}$ for nonprecessing systems. On the other hand, the eight-dimensional parameter space of precessing binaries is much more sparsely explored by the simulations available to us. But because the reconstructed gravitational wave signals in LVC-detect [2] and LVC-PE [1] exhibit little to no modulation, we expect that the remaining four parameters must have at best a subtle effect on the signal: the likelihood and posterior distribution should depend only weakly on any additional subdominant parameters. Having identified dominant trends using nonprecessing simulations, we can use controlled sequences of simulations with similar parameters to determine the residual impact of transverse spins. Even if the marginalized likelihood cannot be safely approximated in general, a simulation’s value of $\log L$ provides insight into the parameters of the event.

Motivated by the parameters reported in LVC-PE [1] and our results in Table III, several followup simulations were performed to reproduce GW150914. These simulations are responsible for most of the best-fitting aligned-spin results reported in Table III.

E. Impact of instrumental uncertainty

For simplicity, our analysis does not automatically account for instrumental uncertainty (i.e., in the detector noise power spectrum or instrument calibration), as do the methods in LVC-PE [1]. LVC-PE [1] suggests that, for the intrinsic parameters $\lambda$ of interest here, the impact of these systematic instrumental uncertainties effects are relatively small. We have repeated our analysis using two versions of the instrumental calibration; we find no significant change in our results.

F. Comparison with other methods

LVC-Burst [4] reported on direct comparisons between radiation extracted from NR simulations and nonparametrically reconstructed estimates of the gravitational wave signal; see, e.g., their Fig. 12. Their comparisons quickly identified masses, mass ratios, and spins that were consistent with the data. Our study, which attempts a fully Bayesian direct comparison between the data and the multimodal predictions of NR, produces results consistent with those of LVC-Burst [4]; see, e.g., Figure 4 described below.

LVC-PE [1] performed Bayesian inference on the data using semianalytic models for the gravitational waves from a coalescing compact binary. We directly compare our posterior distribution with that of LVC-PE [1] for the special case of aligned spins. Differences between these posterior distributions can be due to many factors: our choice of starting frequency is slightly higher (30 Hz versus 20 Hz); our approach does not account for calibration uncertainty; and of course we employ NR instead of a semianalytic waveform.
model. To isolate the effects of NR, we have repeated our analysis but with the same nonprecessing waveform model used in LVC-PE\cite{10} rather than with NR waveforms. Using the same input waveforms, our method and that of LVC-PE\cite{10} produce very similar results; see PE+NR-Methods\cite{10} for details. To isolate the effects of the low-frequency cutoff, we performed the nonprecessing analysis reported in LVC-PE\cite{10} with a low frequency cutoff of 30 Hz; we found results similar to LVC-PE\cite{10}.

IV. RESULTS I: PRE-COALESCENCE PARAMETERS

We present two types of results. For generic, precessing NR simulations, we evaluate the marginalized likelihood of source parameters given the data, but because the parameter-space coverage of NR simulations is so sparse, we do not attempt to construct an interpolant for the likelihood as a function of source parameters. For nonprecessing sources, we construct such an interpolant, and we compare with the results of LVC-PE\cite{10}. Using the computed likelihoods, we quantify whether the data are consistent with or favor a precessing source.

A. Results for generic sources, without interpolation

Because the available generic NR simulations represent only a sparse sampling of the parameter space, for generic sources we adopt a conservative approach: we rely only on our estimates of the marginalized likelihood \( \ln L_{\text{marg}} \), and we do not interpolate the likelihood between intrinsic parameters. For nonprecessing sources, we do not interpolate the likelihood between intrinsic parameters, and we compare with the results of LVC-PE\cite{10}. Using the computed likelihoods, we quantify whether the data are consistent with or favor a precessing source.

\[ \chi_{\text{eff}} = \frac{\left| S_1/m_1 + S_2/m_2 \right| \cdot \hat{L}/M}{\text{data}} \]

\[ \chi_{\text{eff}} \geq 0 \]

For the progenitor black hole parameters, our results using \( l = 2 \) modes are summarized in Figures\cite{4, 5, 6, 7} (for nonprecessing sources), as well as by Figures\cite{6, 7} (for nonprecessing sources). For comparison, these figures also include the results obtained in LVC-PE\cite{10}, using approximations appropriate for nonprecessing (black curves) and simply the precessing (blue curves) binaries. The first column of Table\cite{10} shows the one-dimensional range inferred for each parameter by our threshold-based method, using \( l = 2 \) modes only.

Before describing our results, we first demonstrate why our strategy is effective: Figures\cite{4, 6, 7, 8} show that the likelihood is smooth and slowly varying, dominated by a few key parameters. As seen in the right panel of Figure\cite{4}, even our large NR array is relatively sparse. However, as the color scale on this and other figures indicate, the marginalized likelihood varies smoothly with parameters, over a range of more than \( e^{100} \). The simplicity of \( \ln L_{\text{marg}} \) is most apparent using controlled one- and two-parameter subspaces; for example, Figure\cite{3} shows that \( \ln L \) (i.e., the peak of the \( \ln L_{\text{marg}}(M) \)) varies smoothly as a function of \( \chi_{1z} \), \( \chi_{2z} \), for nonprecessing binaries of different mass ratios \( q = m_1/m_2 \). Targeted NR simulations have corroborated the simple dependence of the likelihood seen here. Despite employing simulations with two strongly precessing spins and including higher harmonics, two factors which have been previously shown to be able to break degeneracies\cite{24, 81, 83, 89}. Table\cite{10} reveals that simulations with the same values of \( q \) and \( \chi_{\text{eff}} \) almost always have similar values of \( \ln L \). In other words, these two simple parameters explain most of the variation in \( \ln L \), even when \( L \) changes by up to a factor of \( e^{100} \). Finally and critically, simulations with similar physics produce very similar results. By adopting \( f_{\text{low}} = 30 \) Hz and thereby largely standardizing simulation duration, we find similar values of \( \ln L \) when comparing the data to simulations performed by different groups with similar (or even identical) parameters.

Our results and that of LVC-PE\cite{10} constrain the progenitor binary’s redshifted mass, mass ratio, and aligned effective spin \( \chi_{\text{eff}} \); see Table\cite{11}. The effective spin is defined as\cite{90, 91}

For example, the color scale in Figure\cite{4} provides a graphical representation of \( \ln L \) versus \( \chi_{\text{eff}} \); large values of \( \chi_{\text{eff}} \) (only possible for spin-aligned systems) are inconsistent with the data. The agreement between our results and LVC-PE\cite{10} persists despite using a much larger simulation set than those used to calibrate the models used in LVC-PE\cite{10}; despite employing simulations with black hole spins that are both precessing and with magnitude significantly outside the range \( \chi < 0.5 \) – 0.8 for which these models were calibrated\cite{11, 92, 93}.

The three parameters \( M_z, q, \) and \( \chi_{\text{eff}} \) are well-known to have a strong and tightly-correlated impact on the gravitational wave signal and hence on implied posterior distributions\cite{82, 84, 94, 95}. Since general relativity is scale-free, the total redshifted binary mass \( M \), sets the characteristic physical timescale of the coalescence. Due to strong spin-orbit coupling, aligned spins \( (\chi_{\text{eff}} > 0) \) extend the temporal duration of the inspiral\cite{13} and coalescence of the two black holes (e.g., the hangup effect\cite{96}); aligned spins also increase the final black hole spin and hence extend the duration of the post-
TABLE I. Constraints on $M_\text{t}$, $q$, $\chi_{\text{eff}}$: Constraints on selected parameters of GW150914 derived by directly comparing the data to numerical relativity simulations. The first column reports the extreme values of each parameter consistent with $\ln L_{\text{marg}} > 268.6$ [Eq. (6), with $d = 4$], corresponding to the black points shown in Figures 4, 7, and 10. These are computed using all the $l = 2$ modes of the NR waveforms. Because these extreme values are evaluated only on a sparse discrete grid of NR simulations, this procedure can underestimate the extent of the allowed range of each parameter. The second column reports the 90% credible interval derived by fitting $\ln L_{\text{marg}}$ versus these parameters for nonprecessing binaries, to enable interpolation between points on the discrete grid in $\lambda$; see Section IV B for details. The third column is the union of the two intervals. For comparison, the fourth column provides the interval reported in LVC-PE\cite{1}, including precession and systematics. The remaining three columns show our results derived using all $l \leq 3$ modes.

![Fig. 4. Mass, mass ratio, and effective spin are constrained and correlated.](image)

Merger quasinormal ringdown\cite{27}. More extreme mass ratio extends the duration of the pre-merger phase while dramatically diminishing the amplitude and frequency of post-merger oscillations\cite{67,68,98,99}. As noted above, the data tightly constrain one of these combinations (e.g., the total redshifted mass at fixed simulation parameters). Hence, our ability to constrain any individual parameter $M_\text{t}$, $q$, or $\chi_{\text{eff}}$ is limited not by the accuracy to which $M_\text{t}$ is determined for each simulation (i.e., the width $1/\sqrt{\chi_{\text{eff}}}$), but rather by differences between simulations (i.e., trends in $\ln L$ versus $\chi_{\text{eff}}, q$) which break the
degeneracy between these tightly correlated parameters.

Simulations with a variety of physics fit the data, including strongly precessing systems. In Table III, several simulations with large transverse spins but nearly zero net aligned spin fit the data almost as well as the best-fitting nonprecessing simulations (e.g., SXS:BBH:3; RIT simulation D10_\(q\mathbf{0.75} - \mathbf{0.8}\_x\mathbf{10}^\mathbf{n100}\)). As described below, Table V shows that these and other long precessing simulations fit even better when more low-frequency content is included.

The correspondence between our results and those presented in LVC-PE merits further reflection: by construction our fiducial analysis (Table III) omitted nontrivial early-time information (i.e., \(f < 30\) Hz) which, for each simulation, more tightly constrains the range of masses that could be consistent with the data. In fact, as we show below, strong degeneracies in the gravitational wave signal between mass, mass ratio, and spin imply that our ability to break these degeneracies dominates our reconstruction of source parameters. Omitting information from low frequencies marginally reduces our ability to identify the range of masses that are consistent with the data for one simulation; however, this omission does not impair our ability to draw conclusions overall, after accounting for uncertainties in mass and mass ratio.

Directly comparable to Fig. 12 in LVC-Burst[4], the right panel of Figure 4 provides a visual representation of one key correlation between \(q\) and \(\chi_{\text{eff}}\): only a narrow range of mass ratios and aligned effective spin \(\chi_{\text{eff}}\) are consistent with the data. This range includes both nonprecessing and precessing simulations. Most other parameters have a subdominant effect. For example, restricting attention to nonprecessing binaries for clarity, the data do not strongly discriminate between systems with similar \(\chi_{\text{eff}}\) but different \(\chi_{1, z}, \chi_{2, z}\); see, e.g., Figure 6.

B. Results for nonprecessing sources, including interpolation

Both LVC-PE[1] and our highly-ranked simulations in Table III demonstrate that binary black holes with nonprecessing spins can reproduce GW150914. Only four parameters characterize a nonprecessing binary: the two component masses \(m_1, m_2\) and the components of each BH’s dimensionless spin \(\chi_i\), projected perpendicular to the orbital plane (\(\chi_{1, z}, \chi_{2, z}\)). Nonprecessing binary black hole coalescences have been extensively simulated [45]; see, e.g., Table I. Several models have been developed to reproduce the leading-order gravitational wave emission (the \((l, m) = (2, 2)\) mode) [6, 8, 12, 57, 67, 68]; one, the SEOBNRv2 model [22], is adopted as a fiducial reference by LVC-PE. While this model has not been calibrated to NR for large values of \(\chi_{\text{eff}}\) and \(q\) [11], it has been shown to accurately reproduce the \((2, 2)\) mode from binaries with comparable mass and low spins [9, 11, 61]. Because of degeneracies, data from GW150914 do not easily distinguish between different points in parameter space that have the same values of \(M_2, q, \chi_{\text{eff}}\); in particular, it is difficult to individually measure \(\chi_{1, z}\) and \(\chi_{2, z}\) when \(q \approx 1, \chi_{\text{eff}} \approx 0\) and \(\chi_{1, z} = -\chi_{2, z}\); see, e.g., [99]. Because GW150914 has comparable masses and is oriented face-off with respect to the line of sight, even including higher-order modes in the gravitational waveform (which we do in our approach here but is not done for the analytic waveform models) does not strongly break these degeneracies and allow us to distinguish individual spins.

By stitching together our fits for \(\ln L_{\text{marg}}(M_2)\) and reconstructing the relevant parts of the likelihood for all masses and aligned spins, we can estimate the full posterior distribution for \(M_2, q, \chi_{1, z}, \chi_{2, z}\) using Eq. (5). Due to inevitable systematic modeling errors in the fit, as described below, this approximation may not have the statistical purity of the method presented in LVC-PE: any credible intervals or deductions drawn from it should be interpreted with judicious skepticism.

On the other hand, this method enables the reader to recalculate the posterior distribution using any prior \(p(\lambda)\), including astrophysically-motivated choices. Fitting to nonprecessing simulations, we find \(\ln L_{\text{marg}}\) for \(\ln L_{\text{marg}} > 262\) is reasonably well-approximated by a quadratic function of the intrinsic parameters \(M_\ast = (m_1 m_2)^{3/5}/(m_1 + m_2)^{3/5}\), \(\eta = (m_1 m_2)/(m_1 z_1^2 + m_2 z_2^2)\), \(\delta = (m_1 - m_2)/M_\ast\), \(\chi_{\text{eff}}\), and \(\chi_\ast \equiv (m_1 \chi_1 - m_2 \chi_2)/M_\ast\):

\[
\ln L_{\text{marg}} \approx 268.4 - \frac{1}{2} (\chi - \chi_\ast) \Gamma \chi
\]

where the indices \(a, b\) run over the variables \(M_\ast, \eta, \chi_\ast, \chi\). In this expression, \(\chi_\ast\) represents the vector \((M_\ast, \eta, \chi_\ast, \chi)\). \(\chi_\ast\) corresponds to the vector \((M_\ast = 31.76 M_\odot, \eta = 0.255, \chi_{\text{eff}} = -0.037, \chi_\ast = 0)\) of parameters which maximize \(\ln L_{\text{marg}}\) and \(\Gamma\) is a matrix (indexed by \(M_\ast, \eta, \chi_{\text{eff}}, \chi\)), \(\chi\) with numerical values

\[
\Gamma = \begin{bmatrix}
3.75 & -224.2 & -52.0 & 0 & 0 \\
-224.2 & 2269.2 & 2692 & 0 & 0 \\
-52.0 & 2692 & 846.9 & 0 & 0 \\
0 & 0 & 0 & 2.57 & -16.3 \\
0 & 0 & 0 & -16.3 & 0
\end{bmatrix}
\]

Here we retain many significant digits to account for structure in \(\Gamma\), which is nearly singular. Equation (9) respects exchange symmetry \(m_1, \chi_1 \leftrightarrow m_2, \chi_2\). Our results do not sensitively depend on the value of \(\Gamma_{\chi, \chi}\), indicating that this quantity is not strongly constrained by the data. Conversely, the posteriors do depend on \(\Gamma_{\chi, \chi}\). As the contrast between the first term in Eq. (9) and the data Table III makes immediately apparent, this coarse approximation can differ from the simulated results by of order 1.7 in the log (rms residual). This reflects the combined impact of Monte Carlo error, systematic error caused by too few orbits in some simulations, and systematic errors caused by sparse placement of NR simulations and non-quadaric behavior of \(\ln L_{\text{marg}}\) with respect to parameters. Repeating our calculation while including all the \(l \leq 3\) modes, we find the same functional form as Eq. (9), but with a different vector \(\chi_{\ast(3)}^\text{NR} = (M_\ast = 38.1 M_\odot, \eta = 0.32, \chi_{\text{eff}} = 0.11, \chi_\ast = 0)\), and a different matrix

\[
\Gamma_{\ast(3)} = \begin{bmatrix}
3.746 & -235.5 & -51.5 & 0 & 0 \\
-235.5 & 17970 & 2941 & 0 & 0 \\
-51.5 & 2941 & 833.2 & 0 & 0 \\
0 & 0 & 0 & 0.57 & -12.57 \\
0 & 0 & 0 & -12.57 & 0
\end{bmatrix}
\]
FIG. 5. **Distributions agree [nonprecessing case]:** Comparison between the posterior distributions reported in LVC-PE[1] for nonprecessing binaries (solid) and the posterior distributions implied by a leading-order approximation to $\ln L_{\text{prior}}$ [Eq. (9)] derived using $l \leq 2$ (dotted) and $l \leq 3$ (dashed). **Left panel:** $m_{1,2}$ (black) and $m_{2,3}$ (red). **Center panel:** Mass ratio $1/q = m_{2,3}/m_{1,2}$. The data increasingly favor comparable-mass binaries as higher-order harmonics are included in the analysis. **Right panel:** Aligned effective spin $\chi_{\text{eff}}$. The noticeable differences between our $\chi_{\text{eff}}$ distributions and the solid curve are also apparent in Figures 7 and 10; our analysis favors a slightly higher effective spin.

We label $\lambda^{(3)}$ and $\Gamma^{(3)}$ with a superscript “3” to distinguish this result from the corresponding result using only $l = 2$ modes shown in Eq. (9).

For nonprecessing sources, using Eq. (5) and a uniform prior in $\chi_{1,2}, \chi_{2,3}$ and the two component masses, we can evaluate the marginal posterior probability $p(z)$ for any intrinsic parameter(s) $z$. The two-dimensional marginal posterior probability is shown as a green solid ($l = 2$) and dashed ($l \leq 3$) line in Figures 4 and 10. Both the $l = 2$ and $l \leq 3$ two-dimensional distributions are in reasonable agreement with the posterior distributions reported in LVC-PE[1] for nonprecessing binaries, shown as a black curve in these figures. These two-dimensional distributions are also consistent with the distribution of simulations with $\ln L_{\text{marg}} > 268.6$ (i.e., black points). Additionally, Figure 5 shows several one-dimensional marginal probability distributions ($m_{1,2}, m_{2,3}, q, \chi_{\text{eff}}$), shown as dotted ($l = 2$) or dashed lines ($l \leq 3$); for comparison, the solid line shows the corresponding distribution from LVC-PE[1] for nonprecessing binaries.

Despite broad qualitative agreement, these comparisons highlight several differences between our results and LVC-PE[1], and between results including $l = 2$ modes and those including all $l \leq 3$ modes. For example, Figure 4 shows that the distribution in $M_e$, $q$, $\chi_{\text{eff}}$, computed using our method (solid green lines and black points) is slightly different than the corresponding distributions in LVC-PE[1]. As seen in this figure and in Figure 5, the posterior distribution in LVC-PE[1] includes binaries with low effective spin, outside the support of the distributions reported here. These differences are directly reflected in the marginal posterior $p(\chi_{\text{eff}})$ (right panel of Figure 5) and in Table 1. Our results for the component spins $\chi_{1,2}, \chi_{2,3}$, the effective spin $\chi_{\text{eff}}$, the total mass $M_z$, and the mass of the more massive object $m_{1,2}$ do not change significantly when $l = 3$ modes are included. The mass ratio distribution $p(q)$ is also slightly different from LVC-PE[1] when $l = 3$ modes are included; see Figure 5. Compared to prior work, this analysis favors comparable-mass binaries when higher modes are included; see, e.g., the center panel of
FIG. 7. Aligned spin components not constrained [aligned only shown]: Colors represent the marginalized log likelihood as a function of the aligned spin components $\chi_{1z}$ and $\chi_{2z}$. Each point represents an NR simulation; only nonprecessing simulations are included. Points with $265.8 < \ln L < 268.6$ are shown in light gray, with $\ln L > 268.6$ are shown in black, and with $\ln L < 265.8$ are shown according to the color scale on the right (points with $\ln L_{\text{marg}} < 172$ have been suppressed to increase contrast). The quantity $\ln L$ is the maximum value of $\ln L_{\text{marg}}$ with respect to mass; see Eq. (7). Consistent with our other results, $f_{\text{low}} = 30\,\text{Hz}$. For comparison, the solid black contours show the 90% credible intervals derived in LVC-PE[1], assuming spin-orbit alignment and omitting corrections for waveform systematics. The solid and dashed green contours are the nominal 90% credible interval derived using an approximation to our data for $\ln L_{\text{marg}}$, assuming both spins are exactly parallel to the orbital angular momentum, for $l = 2$ (solid) and $l = 3$ (dashed), respectively; see Section IV B for more details.

Figure 5

The differences between the results reported here and LVC-PE[1] should be considered in context: not only does our study employ numerical relativity without analytic waveform models, but it also adopts a slightly different starting frequency, omits any direct treatment of calibration uncertainty, and employs a quadratic approximation to the likelihood. That said, comparisons conducted under similar limitations and using real data, differing only in the underlying waveform model, reproduce results from LALINFERENCE; see PE+NR-Methods[10] for details.

By assuming the binaries are strictly aligned but permitting generic spin magnitudes, our analysis (and that in LVC-PE[1]) neglects prior information that could be used to significantly influence the posterior spin distributions. For example, the part of the posterior in the bottom right quadrant of Figure 7 is unstable to large angle precession[100]: if a comparable-mass binary formed at large separation with $\chi_{1z} > 0$ and $\chi_{2z} < 0$, it could not remain aligned during the last few orbits. Likewise, the astrophysical scenarios most likely to produce strictly aligned binaries — isolated binary evolution — are most likely to result in both $\chi_{1z}, \chi_{2z} > 0$: both spins would be strictly and positively aligned (see, e.g., [101]). In that case, only the top right quadrant of Figure 2 would be relevant. Using the analytic tools provided here, the reader can regenerate approximate posterior distributions employing any prior assumptions, including these two considerations.

C. Transverse and precessing spins

Figure 8 shows the maximum likelihood for the available NR simulations, plotted as a function of the magnitude of the aligned and transverse spin components. The figure shows that there are both precessing and nonprecessing simulations that have large likelihoods (black points), indicating that many precessing simulations are as consistent with the data as nonprecessing simulations. Moreover, simulations with large precessing spins are consistent with the GW150914; many configurations have $\chi_{\text{eff}} \approx 0$ but large spins on one or both BHs in the binary. Keeping in mind the limited range of simulations available, the magnitude and direction of either BHs spin cannot be significantly constrained by our method.

Not all precessing simulations with suitable $q, \chi_{\text{eff}}$ are consistent with GW150914; some have values of $\ln L$ that are not within 10 of the peak; see the right panel of Fig. 8. The marginal log-likelihood $\ln L$ depends on the transverse spins, not just the dominant parameters ($q, \chi_{\text{eff}}, M_c$). As a concrete illustration, Figure 9 shows that the marginalized log likelihood depends on the specific direction of the transverse spin, in the plane perpendicular to the angular momentum axis. Specifically, this figure compares the peak marginalized log likelihood ($\ln L$) calculated for each simulation with the value of $\ln L$ predicted from our fit to nonprecessing binaries. For precessing binaries, $\ln L$ is neither in perfect agreement with the nonprecessing prediction, nor independent of rotations of the initial spins about the initial orbital angular momentum by an angle $\phi$.

While the transverse spins do influence the likelihood, slightly, the data do not favor any particular precessing configurations. No precessing simulations had marginalized likelihoods that were both significant overall and significantly above the value we predicted assuming aligned spins. In other words, the data do not seem to favor precessing systems, when analyzed using only information above 30 Hz.

Our inability to determine the most likely transverse spin components is expected, given both our self-imposed restrictions ($f_{\text{low}} = 30\,\text{Hz}$) and the a priori effects of geometry. For example, the lack of apparent modulation in the signal reported in LVC-detect[2] and LVC-Burst[4] points to an orientation with $J$ parallel to the line of sight, along which precession-induced modulations are highly suppressed. In addition, the high mass and hence extremely short observation timescales involved are particularly unfavorable to attempts to extract precession-induced modulation from the pre-merger signal: the pre-coalescence precession rate for these sources is low ($\Omega_p \approx (2 + 3m_2/m_1)J/2r^3 \approx 2\pi \times 1\,\text{Hz}/40\,\text{Hz}\times\text{yr}^{3/2}$ for this system, where $J$ is the magnitude of the total orbital angu-
lar momentum and we assume $J \approx L$; see [13]), implying at best two pre-merger precession cycles could be accessible from the early signal; see [LVC-PE1]. As with the total binary mass, spin information will be accessible at lower frequencies (i.e., between $10-30$ Hz); however, our fiducial analysis using $f_{\text{low}} = 30$ Hz is not well-suited to extract it.

For a suitably-oriented source, the strongly nonlinear merger phase can in principle encode significant information about the coalescing binary’s precessing spins. Qualitatively speaking, this information is encoded in the relative amplitude and phase of subdominant quasilinear perturbations, causing the radiation from the final black hole to appear precess [15, 29]. This information also influences the final black hole mass and spin. The model used in LVC-PE [1] adopted a geometric ansatz to incorporate these effects at leading order, using a lower-dimensional effective model for a single precessing spin. However, in this work, despite including higher modes and having direct access to as-yet unmodeled effects, our analysis shows no significant difference from the previously reported conclusions regarding the transverse spin distribution.

The low frequency content of GW150914 may contain some further signature consistent with two precessing spins. Simultaneously with this work, an analysis has been performed using semianalytic models that can fully capture both spins’ dynamics LVC-SEOBNRv3 [102]. Within the context of this study, Table V shows an analysis without an artificially imposed low frequency cut-
off. As expected, the best-fitting long simulations seen in our previous report fit equally well and agree. Notably, however, we find an increase in the marginalized likelihood for precessing simulations like SXS:BBH:308 and D21.5_q1_a0.2_0.8_th104.4775_n100. More broadly, when we include low-frequency content, many precessing simulations that previously had not fit the high-frequency content as well become more significant. However, to extract low-frequency content reliably, we will need to both hybridize these precessing simulations and interpolate the likelihood as a function of both precessing spins. These further investigations are beyond the scope of the present study.

V. RESULTS II: STRONG-FIELD PROPERTIES AND POST-COALESCENCE PARAMETERS

The numerical relativity simulations listed in Table [1] have been previously used to develop accurate models for the final black hole mass and spin [25, 26, 105, 104]. The relationships developed in [25] for nonprecessing binaries were used in LVC-PE[1] and LVC-TestGR[3] to infer the final black hole mass and spin, based on the pre-coalescence spins. By construction, this approximation neglects the impact of transverse spins. Both this work and in LVC-PE[1] have shown that GW150914 is consistent both with nonprecessing and precessing pre-coalescence spins. When large, these spins are well-known to significantly impact the final black hole mass and spin [17, 105, 109].

With direct access to both an accurate multimodal waveform for generic precessing systems and the final black hole state, the method applied in this work is uniquely well equipped to identify the final black hole mass and spin. Figure 10 shows our results. Rather than approximate a posterior distribution — a significant challenge in 8 dimensions — we simply report sets of points $M_{f}, a_{f}$ corresponding to simulations and initial redshifted masses $M_{i}$ so $\ln L_{\text{marg}}(M_{i})$ is greater than some cutoff. When we include only nonprecessing simulations, we find results consistent with the reported values in LVC-PE[1] and LVC-TestGR[3]. While many simulations listed in Table [1] have some transverse spin, many also have zero transverse spin, so overall the transverse spin distribution of our simulations is more concentrated towards zero than the prior adopted in LVC-PE[1]. Given the excellent agreement between our results and LVC-PE[1] for pre-coalescence parameters, particularly in the subset of spin-aligned binaries, we cannot identify any nonzero difference for final parameters that is introduced by our methodology (e.g. our restriction to $f_{\text{low}} = 30 \text{ Hz}$).

VI. CONCLUSIONS

Using a full Bayesian parameter estimation technique, we directly compare GW150914 with a large set of binary black hole simulations produced using full numerical relativity. Our comparisons employ physics and radiation content ($l \leq 3$ modes) not available or only partially captured by the two semianalytic models used in LVC-PE[1]. Using our completely independent approach, we nonetheless arrive at results similar to those of LVC-PE[1]. Comparisons including only the dominant modes (all $l \leq 2$) constrain the total redshifted mass $M_{i}$ [64 − 82 $M_{\odot}$], mass ratio $1/q = m_{1}/m_{2} \in [0.6, 1]$, and effective aligned spin $\chi_{\text{eff}} \in [-0.3, 0.2]$. Including $l = 3$ modes, we find the mass ratio is even more tightly constrained. Both nonprecessing and precessing simulations fit the data; no compelling evidence exists for or against a precessing origin. Even accounting for precession, simulations with extreme mass ratios and effective spins are highly inconsistent with the data, at any mass. Several nonprecessing and precessing simulations with similar mass ratio and $\chi_{\text{eff}}$ are consistent with the data. Though correlated, the component spins (both in magnitude and direction) are not significantly constrained by the data: the data are consistent with simulations with component spin magnitudes $a_{1,2}$ up to at least 0.8, with random orientations.

This paper also provides the first concrete illustration, using real gravitational wave data, of several methods to aid the interpretation of gravitational wave observations using numerical relativity. First and foremost, this method demonstrates that the marginalized likelihood can be efficiently evaluated on a grid [80, 81]. Straightforward reconstructions (e.g., fits, interpolation) allow us to reconstruct the posterior at low cost. Further, NR simulations are sufficiently dense, and the marginal log-likelihood $\ln L_{\text{marg}}$ sufficiently simple, that $\ln L_{\text{marg}}$ can be effectively approximated using available catalogs of NR simulations. Second, we provide and employ a simple but effective approximation to the marginalized likelihood. A particularly efficient way to communicate results, this data product enables further investigations, including the impact of the prior on our conclusions; the ability to
incorporate the spin-precession instability into our posterior [100]; and anything involving conditional distributions, which are trivially produced using the fit. Third, this investigation has demonstrated the critical role that numerical relativity can play in data analysis while simultaneously illuminating a path forward in the era of frequent detections. We demonstrate that NR results can be directly applied to data analysis, without intervening approximations. In the future, while low-frequency sensitivity will improve, so will our ability to effectively hybride these simulations, so this approach will remain valuable even when very long signal models are required to reproduce the data. Targeted followup can be performed guided by $\ln L$, our measure of overall fit (maximizing $\ln L_{\text{mag}}$ over mass). Fourth, as described in PE+NR-Methods[10], this method provides a direct and unambiguous method to assess the relative impact of higher harmonics, waveform extraction, and modeling uncertainty on a point-by-point basis. Investigations using this technique will provide a valuable complement to parallel studies with LALInference [13].

As noted in LVC-Astro[30], the inferred spin magnitudes and misalignments provide unique and distinctive clues to the astrophysical origin of GW150914. Notably, strongly misaligned spins require a violent origin, either through exceptionally dynamic stellar processes or a cluster origin. Our analysis cannot definitively support or rule out such an origin. We recommend further analysis of GW150914 with improved models for binary inspiral and coalescence, whether derived semianalytically or via hybridization and/or interpolation of pure numerical relativity. For example, LVC-SEOBNRv3[102] reports marginally tighter constraints on (two) precessing spins, by comparing GW150914 against a model for the emitted radiation including the very early inspiral, which by necessity NR simulations must omit. Combined with this method, we further anticipate a large-scale simulation campaign in full numerical relativity to explore simulations comparable to GW150914 could allow us to extract more insight into its nature. PE+NR-Methods[10] will provide further details on and examples with the method employed in this work.

ACKNOWLEDGMENTS

The authors 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, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. The authors gratefully acknowledge the Italian Istituto Nazionale di Fisica Nucleare (INFN), the French Centre National de la Recherche Scientifique (CNRS) and the Foundation for Fundamental Research on Matter supported by the Netherlands Organisation for Scientific Research, for the construction and operation of the Virgo detector and the creation and support of the EGO consortium. The authors also gratefully acknowledge research support from these agencies as well as by the Council of Scientific and Industrial Research of India, Department of Science and Technology, India, Science & Engineering Research Board (SERB), India, Ministry of Human Resource Development, India, the Spanish Ministerio de Economía y Competitividad, the Conselleria d’Economia i Competitivitat and Conselleria d’Educatió, Cultura i Universitats of the Govern de les Illes Balears, the National Science Centre of Poland, the European Commission, the Royal Society, the Scottish Funding Council, the Scottish Universities Physics Alliance, the Hungarian Scientific Research Fund (OTKA), the Lyon Institute of Origins (LIO), the National Research Foundation of Korea, Industry Canada and the Province of Ontario through the Ministry of Economic Development and Innovation, the National Science and Engineering Research Council Canada, the Brazilian Ministry of Science, Technology, and Innovation, the Leverhulme Trust, the Research Corporation, Ministry of Science and Technology (MOST), Taiwan and the Kavli Foundation. The authors gratefully acknowledge the support of the NSF, STFC, MPS, INFN, CNRS, and the State of Niedersachsen/Germany for provision of computational resources.

The SXS collaboration also gratefully acknowledges Compute Canada, the Research Corporation, and California State University Fullerton for computational resources, as well as the the support of the National Science Foundation, the Research Corporation for Science Advancement,and the Sherman Fairchild Foundation.

The RIT team gratefully acknowledges the NSF for financial support, as well as Blue Waters and XSEDE for computational resources.

This paper has been assigned the document number LIGO-P1500263.


Appendix A: Simulation list

In this section, we enumerate the simulations used in this work [Table II], providing a more detailed description of the simulations performed and their relationship to the literature. Unless otherwise noted, we extract $\hat{m}_{\text{irr}}(f)$ [and therefore $h_{\text{irr}}(f)$] at infinity using a perturbative extrapolation [110] re-expressed in the Fourier domain; see PE+NR-Methods [10] for further details.

**RIT simulations**: BBH data were evolved using the LazEv [111] implementation of the moving puncture approach [41,42] with the conformal function $W = \sqrt{\chi} = \exp(-2\phi)$ suggested by Ref. [112]. For the run presented here, we use centered, sixth-order finite differencing in space [113] and a fourth-order Runge Kutta time integrator. (Note that we do not upwind the advection terms.) This code uses the EinsteinToolKit [114] / Cactus [115] / Carpet [116] infrastructure. The Carpet mesh refinement driver provides a “moving boxes” style of mesh refinement. In this approach, refined grids of fixed size are arranged about the coordinate centers of both holes. The Carpet code then moves these fine grids about the computational domain by following the trajectories of the two BHs. The RIT group used AHFinderDirect [117] to locate apparent horizons. The magnitude of the horizon spin is computed using the isolated horizon (IH) algorithm detailed in Ref. [118] and as implemented in Ref. [119]. Note that once we have the horizon spin, we can calculate the horizon mass via the Christodoulou formula

$$m_H = \sqrt{m_{\text{irr}}^2 + S_H^2/(4m_{\text{irr}}^2)},$$  \hspace{1cm} (A1)

where $m_{\text{irr}} = \sqrt{\lambda/(16\pi)}$, $\lambda$ is the surface area of the horizon, and $S_H$ is the spin angular momentum of the BH (in units of $M^2$).

The 128 simulations reported in Zlochower and Lousto [17], denoted in Table II by RIT-Kicks, have only one black hole spinning with $|\chi| = 0.8$. For a handful of these simulations, the estimate of the final black hole mass and spin has been updated since the original publication.

These simulations include (a) a simulation with large transverse spins and several spin precession cycles which fits the data well [24]; (b) a wide range of simulations with large aligned and antialigned spins for mass ratios near and far from unity [73]; (c) a set of simulations with targeted mass ratios and spins, designed to systematically explore the parameter space and reconstruct generic recoil kicks when $q > 1$ [17]; (d) and a set of equal mass simulations with large spins (0.8) and generic orientations, designed to systematically explore the parameter space and reconstruct recoil when $q = 1$ [74].

**SXS simulations**: SXS provided simulations from their public catalog – initially reported in [58] – as well as several selected followup simulations. The SXS collaboration uses the Spectral Einstein Code (SpEC) [120] for evolution. Quasiequilibrium initial data are constructed in the extended conformal thin-sandwich formalism using a pseudo-spectral elliptic solver [121,122]. The evolution occurs on a grid extending from inner excision boundaries, slightly inside the apparent horizons, to an outer boundary on which constraint-preserving boundary conditions are imposed [123]. The code uses a first-order generalized harmonic representation of Einstein’s equations with damped harmonic gauge [124,128]. After merger, the grid is updated to include only one excision boundary [129,130]. The excision boundaries are dynamically adjusted to conform to the shapes of the apparent horizons [129,130]. The initial orbital eccentricity is reduced with an iterative procedure [131,132]. Other improvements have been applied to enable long simulations [133] and simulations of highly-spinning black holes [134].

**GT simulations**: Initial data was evolved with F{\textsc{t}}aya, which was used in previous black hole-black hole (BH-BH) studies [135,142]. The grid structure for each run consisted of 10 levels of refinement provided by Carpet [116], a mesh refinement package for Cactus [115]. Each successive level’s resolution decreased by a factor of 2. Sixth-order spatial finite differencing was used with the BSSN equations implemented with Kranc [143].

Simulations denoted by GT-Aligned refer to the $z$, $zq$, and $zU$ series in [135,139]; the GT-Misaligned case refers to the S and Sq series; and GT-Tilting refers to the T and Tq series. Where available, we adopt the naming convention used in [77]. In particular, the 452 simulations in [77] survey the most extensive parameter space of binary black hole (BBH) systems with 49 non-spinning, 81 aligned-spinning and 324 generic precessing spins BBH simulations. They cover mass-ratios ranging from $q \leq 15$ for non-spinning and $q \leq 8$ for precessing spinning BBH systems, and include generic spin orientations and spin magnitudes, $|\alpha| < 0.8$.

**BAM simulations**: The Cardi{\textsc{f}}-UIB group produced 29 simulations using parameters similar to the event, with approximately random initial configurations within the 99% credible region inferred for GW150914 in LVC-PE[1]. These BAM simulations were produced by the bifunctional adaptive mesh (BAM) NR code [78,79]. The BAM code solves the Einstein evolutions equations using the BSSN [143,145] formulation of the $3+1$ decomposed Einstein field equations. The BSSN equations are integrated with a fourth order finite-difference Runge-Kutta time integrator, with a fixed time step along with a sixth order accurate finite difference algorithm based on the method-of-lines for spatial derivatives. The $\chi$ variation of the moving-puncture method is used where a new conformal factor defined as $\chi = \psi^{-4}$ which is finite at the puncture. The lapse and shift gauge functions are evolved using the $1 + \log$ slicing condition and the Gamma driver shift condition respectively. Conformally flat puncture initial data [146,148] are calculated using the pseudospectral elliptic solver described in [149].

For a handful of these simulations, the estimate of the final black hole mass and spin has been updated since the original publication.
1. Followup simulations

Several groups performed new simulations in response to GW150914, indicated in Table II by an asterisk (*). Some of these simulations were made available for this analysis. The SXS group performed 8 targeted simulations near the maximum mass nonprecessing binaries, targeting the mass ratio and spin range favored by LVC-PE [1]. The RIT group performed a systematic followup campaign on nonprecessing binaries, targeting the mass ratio and spin range favored by LVC-PE [1]. This campaign included 52 new simulations of nonprecessing binaries in the range of mass ratio 1/2 ≤ q ≤ 1 for spinning binaries and up to q = 1/6 without spin. So far the sequence also includes 11 new precessing simulations in the observationally-relevant mass range of 1/3 ≤ q ≤ 3/4 to further calibrate results.

<table>
<thead>
<tr>
<th>Name</th>
<th>Key</th>
<th>q</th>
<th>$\chi_1.s$</th>
<th>$\chi_1.y$</th>
<th>$\chi_1.z$</th>
<th>$\chi_2.s$</th>
<th>$\chi_2.y$</th>
<th>$\chi_2.z$</th>
<th>$\chi_{a\theta}$</th>
<th>$M_{\text{chirp}}$</th>
<th>$M_1/M_a$</th>
<th>$a_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIT-Generic</td>
<td>D18.08_1107.a.e.b.s.a107(*)</td>
<td>6.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.025</td>
<td>0.986</td>
<td>0.372</td>
</tr>
<tr>
<td>RIT-Generic</td>
<td>D18.08_1108.a.e.b.s.a107(*)</td>
<td>2.990</td>
<td>0.757</td>
<td>0.030</td>
<td>0.259</td>
<td>-</td>
<td>-</td>
<td>-0.800</td>
<td>-0.006</td>
<td>0.029</td>
<td>0.965</td>
<td>0.756</td>
</tr>
<tr>
<td>RIT-Generic</td>
<td>D18.08_1109.a.e.b.s.a107(*)</td>
<td>2.990</td>
<td>0.754</td>
<td>0.031</td>
<td>-0.268</td>
<td>-</td>
<td>0.800</td>
<td>-0.001</td>
<td>0.029</td>
<td>0.970</td>
<td>0.607</td>
<td></td>
</tr>
<tr>
<td>RIT-Generic</td>
<td>D18.08_1110.a.e.b.s.a107(*)</td>
<td>2.000</td>
<td>-</td>
<td>-</td>
<td>0.500</td>
<td>-</td>
<td>-</td>
<td>-0.500</td>
<td>0.167</td>
<td>0.288</td>
<td>0.953</td>
<td>0.751</td>
</tr>
<tr>
<td>RIT-Generic</td>
<td>D18.08_1111.a.e.b.s.a107(*)</td>
<td>2.000</td>
<td>0.686</td>
<td>0.059</td>
<td>0.392</td>
<td>-</td>
<td>-0.006</td>
<td>-0.801</td>
<td>-0.005</td>
<td>0.030</td>
<td>0.956</td>
<td>0.768</td>
</tr>
</tbody>
</table>

**TABLE II: List of simulations**: Table of simulations used in this work. Columns indicate the group; an (internal) shorthand for the simulation; the mass ratio; and the components of the dimensionless spin $\chi_1 = S_i/m_i^2$; the effective aligned spin $\chi_z$; the estimated initial starting orbital frequency $\omega_{\text{chirp}}$; and (where available) the final black mass and spin. [We indicate where the black hole mass and spin was unavailable by using X for the corresponding entry.] (The printed table only shows a few entries from each group; the full table is available as online supplementary material.)

Continued on next page
In this section, we enumerate the simulations used in this work, ordered by one measure of their similarity with the data [\( \ln L \), in Table III]. For nonprecessing binaries, Figure 6 provides a visual illustration of some trends in \( \ln L \) versus mass ratio and the two component spins.

**TABLE III: Peak Marginalized \( \ln L \): Consistency between simulations**

Peak value of the marginalized log likelihood \( \ln L \) [Eq. (7)] evaluated using a lower frequency \( f_{\text{low}} = 30 \) Hz and all modes with \( l \leq 2 \); the simulation key, described in Table II [an asterisk (*) denotes a new simulation motivated by GW150914, and a (+) denotes one of the simulations reported in LVC-detect][4]; the initial spins of the simulation (using \( - \) to denote zero, to enhance readability); the initial \( \chi_0 \); the total (redshifted) mass of the best fit; and the starting frequency (in Hz) of the best fit. Though omitting information accessible to the longest simulations, this choice of low-frequency cutoff eliminates systematic biases associated with simulation duration, which differs across our archive, as seen by the last column.

<table>
<thead>
<tr>
<th>( \ln L )</th>
<th>Key</th>
<th>( q )</th>
<th>( \chi_{1,x} )</th>
<th>( \chi_{1,y} )</th>
<th>( \chi_{2,x} )</th>
<th>( \chi_{2,y} )</th>
<th>( \chi_{2,z} )</th>
<th>( M_z / M_\odot )</th>
<th>( f_{\text{start}} ) (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>272.2</td>
<td>SXS:BBH:0316(*)</td>
<td>1.221</td>
<td>-</td>
<td>-</td>
<td>0.250</td>
<td>-</td>
<td>-0.250</td>
<td>0.00</td>
<td>73.0</td>
</tr>
<tr>
<td>272.1</td>
<td>D12_q.1.00_a.0.25_d.0.25_n100(*)</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>0.250</td>
<td>-</td>
<td>-0.250</td>
<td>0.00</td>
<td>73.2</td>
</tr>
<tr>
<td>272.1</td>
<td>SXS:BBH:0002</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>73.2</td>
</tr>
<tr>
<td>271.8</td>
<td>D11_q.8.75_a.0.0_d.0.0_n100(*)</td>
<td>1.333</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>72.1</td>
</tr>
<tr>
<td>271.8</td>
<td>SXS:BBH:0305(+*)</td>
<td>1.221</td>
<td>-</td>
<td>-</td>
<td>0.330</td>
<td>-</td>
<td>-0.440</td>
<td>-0.02</td>
<td>74.2</td>
</tr>
<tr>
<td>271.6</td>
<td>SXS:BBH:0218</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>-0.500</td>
<td>-</td>
<td>0.500</td>
<td>0.00</td>
<td>73.3</td>
</tr>
<tr>
<td>271.6</td>
<td>SXS:BBH:0198</td>
<td>1.202</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>73.4</td>
</tr>
<tr>
<td>271.6</td>
<td>SXS:BBH:0307(*)</td>
<td>1.228</td>
<td>-</td>
<td>-</td>
<td>0.320</td>
<td>-</td>
<td>-0.580</td>
<td>-0.08</td>
<td>70.0</td>
</tr>
<tr>
<td>271.6</td>
<td>GT:BBH:476</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>67.9</td>
</tr>
<tr>
<td>271.6</td>
<td>S0_D10.04_q.1.3333_a.0.45_d.0.80_n100</td>
<td>1.334</td>
<td>-</td>
<td>-</td>
<td>0.450</td>
<td>-</td>
<td>-0.801</td>
<td>-0.09</td>
<td>71.9</td>
</tr>
<tr>
<td>271.5</td>
<td>D12_q.8.05_d.0.0_d.0.0_n100(*)</td>
<td>1.176</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>73.0</td>
</tr>
<tr>
<td>271.5</td>
<td>D12_q.25_d.0.08_a.0.04_d.0.33_n100(+*)</td>
<td>1.22</td>
<td>-</td>
<td>-</td>
<td>0.330</td>
<td>-</td>
<td>-0.440</td>
<td>-0.02</td>
<td>72.9</td>
</tr>
<tr>
<td>271.5</td>
<td>SXS:BBH:0312(*)</td>
<td>1.203</td>
<td>-</td>
<td>-</td>
<td>0.390</td>
<td>-</td>
<td>-0.480</td>
<td>-0.00</td>
<td>73.9</td>
</tr>
<tr>
<td>271.4</td>
<td>SXS:BBH:0127</td>
<td>1.34</td>
<td>0.010</td>
<td>-0.077</td>
<td>-0.017</td>
<td>-0.061</td>
<td>-0.056</td>
<td>-0.179</td>
<td>-0.09</td>
</tr>
<tr>
<td>271.4</td>
<td>SXS:BBH:0115</td>
<td>1.07</td>
<td>0.019</td>
<td>0.013</td>
<td>-0.204</td>
<td>0.243</td>
<td>-0.067</td>
<td>0.291</td>
<td>0.04</td>
</tr>
<tr>
<td>271.3</td>
<td>SXS:BBH:0213</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>73.2</td>
</tr>
<tr>
<td>271.3</td>
<td>UD_D10.01_q.1.00_a.0.4_d.0.100</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>0.400</td>
<td>-</td>
<td>-0.400</td>
<td>-0.00</td>
<td>73.4</td>
</tr>
<tr>
<td>271.2</td>
<td>D12_q.1.00_a.0.25_d.0.00_n100(*)</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>73.1</td>
</tr>
<tr>
<td>271.2</td>
<td>SXS:BBH:0222</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>-0.300</td>
<td>-</td>
<td>-</td>
<td>-0.15</td>
<td>69.1</td>
</tr>
<tr>
<td>271.2</td>
<td>SXS:BBH:0217</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>-0.600</td>
<td>-</td>
<td>0.600</td>
<td>0.00</td>
<td>73.2</td>
</tr>
<tr>
<td>271.1</td>
<td>D10_q.8.75_a.0.5_d.0.25_n100(*)</td>
<td>1.333</td>
<td>-</td>
<td>-</td>
<td>0.250</td>
<td>-</td>
<td>-0.500</td>
<td>-0.07</td>
<td>71.9</td>
</tr>
<tr>
<td>271.0</td>
<td>BAM150914:24(*)</td>
<td>1.2</td>
<td>0.151</td>
<td>0.396</td>
<td>0.017</td>
<td>-0.278</td>
<td>-0.605</td>
<td>-0.085</td>
<td>-0.03</td>
</tr>
<tr>
<td>270.9</td>
<td>GW15_D12_q.22_d.0.33_d.0.44_m140(*)</td>
<td>1.228</td>
<td>0.072</td>
<td>0.027</td>
<td>0.325</td>
<td>0.201</td>
<td>0.285</td>
<td>-0.571</td>
<td>-0.08</td>
</tr>
<tr>
<td>270.9</td>
<td>SXS:BBH:0308<a href="*">Lev3</a></td>
<td>1.12</td>
<td>0.138</td>
<td>-0.200</td>
<td>-0.008</td>
<td>-0.065</td>
<td>-0.241</td>
<td>-0.099</td>
<td>-0.05</td>
</tr>
</tbody>
</table>

Continued on next page
TABLE IV: Peak Marginalized ln L I: Consistency between simulations: Peak value of the marginalized log likelihood ln L evaluated using a lower frequency $f_{\text{low}} = 30$ Hz [Eq. (7)] and all modes with $l \leq 3$; the simulation key, described in Table II [an asterisk (*) denotes a new simulation motivated by GW150914]; the initial spins of the simulation (using $-\chi$ to denote zero, to enhance readability); the initial $\chi_{\text{eff}}$; the total (redshifted) mass of the best fit; and the starting frequency (in Hz) of the best fit. Though omitting information accessible to the longest simulations, this choice of low-frequency cutoff eliminates systematic biases associated with simulation duration, which differs across our archive.

<table>
<thead>
<tr>
<th>ln L</th>
<th>Key</th>
<th>$q$</th>
<th>$\chi_{1,x}$</th>
<th>$\chi_{1,y}$</th>
<th>$\chi_{1,z}$</th>
<th>$\chi_{2,x}$</th>
<th>$\chi_{2,y}$</th>
<th>$\chi_{2,z}$</th>
<th>$\chi_{\text{eff}}$</th>
<th>$M_1/M_\odot$</th>
<th>$f_{\text{low}}$(Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>272.0</td>
<td>SXS:BBH:0307(*)</td>
<td>1.228</td>
<td>-</td>
<td>-0.320</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.580</td>
<td>70.9</td>
<td>17.3</td>
</tr>
<tr>
<td>271.8</td>
<td>SXS:BBH:0115</td>
<td>1.07</td>
<td>0.019</td>
<td>0.013</td>
<td>-0.204</td>
<td>0.243</td>
<td>0.067</td>
<td>0.291</td>
<td>0.04</td>
<td>74.0</td>
<td>14.6</td>
</tr>
<tr>
<td>273.8</td>
<td>SXS:BBH:0305(*)</td>
<td>1.221</td>
<td>-</td>
<td>-0.330</td>
<td>-</td>
<td>-0.440</td>
<td>-</td>
<td>-</td>
<td>-0.02</td>
<td>73.5</td>
<td>15.4</td>
</tr>
<tr>
<td>271.8</td>
<td>D12_q1.00_a-0.25-0.25_n100(*)</td>
<td>1.0</td>
<td>-</td>
<td>-0.250</td>
<td>-</td>
<td>-0.250</td>
<td>-</td>
<td>-</td>
<td>-0.00</td>
<td>73.6</td>
<td>20.4</td>
</tr>
<tr>
<td>271.5</td>
<td>SXS:BBH:0310(*)</td>
<td>1.221</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>72.4</td>
<td>15.4</td>
</tr>
<tr>
<td>271.0</td>
<td>GW15_D12_q1.22_a0.33-0.44_n140(*)</td>
<td>1.22</td>
<td>-</td>
<td>-0.330</td>
<td>-</td>
<td>-0.440</td>
<td>-</td>
<td>-</td>
<td>-0.02</td>
<td>72.8</td>
<td>19.4</td>
</tr>
<tr>
<td>270.8</td>
<td>SXS:BBH:0116</td>
<td>1.08</td>
<td>-0.078</td>
<td>0.065</td>
<td>0.033</td>
<td>0.185</td>
<td>0.007</td>
<td>0.103</td>
<td>0.07</td>
<td>75.1</td>
<td>13.9</td>
</tr>
<tr>
<td>270.7</td>
<td>SXS:BBH:0902</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>75.1</td>
<td>9.7</td>
</tr>
<tr>
<td>270.6</td>
<td>SXS:BBH:0316(*)</td>
<td>1.186</td>
<td>0.241</td>
<td>0.170</td>
<td>0.299</td>
<td>-0.203</td>
<td>-0.172</td>
<td>-0.271</td>
<td>0.04</td>
<td>74.1</td>
<td>15.2</td>
</tr>
<tr>
<td>270.4</td>
<td>D12_q0.85_a0.00.0_n100(*)</td>
<td>1.176</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.00</td>
<td>74.5</td>
<td>20.2</td>
</tr>
<tr>
<td>270.0</td>
<td>SXS:BBH:0120</td>
<td>1.12</td>
<td>0.138</td>
<td>-0.200</td>
<td>-0.008</td>
<td>-0.065</td>
<td>-0.241</td>
<td>-0.099</td>
<td>-0.05</td>
<td>70.0</td>
<td>14.9</td>
</tr>
<tr>
<td>270.4</td>
<td>SXS:BBH:0198</td>
<td>1.202</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>74.7</td>
<td>12.7</td>
</tr>
<tr>
<td>270.3</td>
<td>SXS:BBH:0125</td>
<td>1.27</td>
<td>0.012</td>
<td>0.045</td>
<td>-0.058</td>
<td>0.389</td>
<td>0.241</td>
<td>0.070</td>
<td>-0.00</td>
<td>73.5</td>
<td>13.6</td>
</tr>
<tr>
<td>270.3</td>
<td>GT:BBH:900(*)</td>
<td>1.2</td>
<td>0.400</td>
<td>-</td>
<td>0.400</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>72.7</td>
<td>22.3</td>
</tr>
<tr>
<td>270.2</td>
<td>D10_q0.75_a0.0.0_0.25_n100(*)</td>
<td>1.333</td>
<td>-</td>
<td>-0.250</td>
<td>-</td>
<td>-0.500</td>
<td>-</td>
<td>-</td>
<td>-0.07</td>
<td>73.7</td>
<td>26.7</td>
</tr>
<tr>
<td>270.1</td>
<td>GT:BBH:898(*)</td>
<td>1.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>74.8</td>
<td>17.9</td>
</tr>
</tbody>
</table>

TABLE V: Peak Marginalized ln L: Low frequency included: Peak value of the marginalized log likelihood ln L evaluated using a lower frequency $f_{\text{low}} = 10$ Hz and all modes with $l \leq 2$; the simulation key, described in Table II [an asterisk (*) denotes a new simulation motivated by GW150914, and a (+) denotes one of the simulations reported in LVC-detect]; the initial spins of the simulation (using $-\chi$ to denote zero, to enhance readability); the initial $\chi_{\text{eff}}$; the total (redshifted) mass of the best fit. This choice of low-frequency cutoff ensures that long simulations can make the best use of low-frequency information in the data, significantly improving our constraints on $M$ and spin precession.

<table>
<thead>
<tr>
<th>ln L</th>
<th>Key</th>
<th>$q$</th>
<th>$\chi_{1,x}$</th>
<th>$\chi_{1,y}$</th>
<th>$\chi_{1,z}$</th>
<th>$\chi_{2,x}$</th>
<th>$\chi_{2,y}$</th>
<th>$\chi_{2,z}$</th>
<th>$\chi_{\text{eff}}$</th>
<th>$M_1/M_\odot$</th>
<th>$f_{\text{low}}$(Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>277.3</td>
<td>SXS:BBH:0002</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>72.7</td>
<td>10.0</td>
</tr>
<tr>
<td>277.2</td>
<td>SXS:BBH:0931(*)</td>
<td>1.217</td>
<td>-</td>
<td>-0.380</td>
<td>-</td>
<td>-0.520</td>
<td>-</td>
<td>-</td>
<td>-0.03</td>
<td>72.9</td>
<td>15.5</td>
</tr>
<tr>
<td>276.9</td>
<td>SXS:BBH:0305(*)</td>
<td>1.221</td>
<td>-</td>
<td>-0.330</td>
<td>-</td>
<td>-0.440</td>
<td>-</td>
<td>-</td>
<td>-0.02</td>
<td>73.2</td>
<td>15.5</td>
</tr>
<tr>
<td>276.6</td>
<td>D11_q0.75_a0.0_0.0_n100(*)</td>
<td>1.333</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.00</td>
<td>72.6</td>
<td>22.9</td>
</tr>
<tr>
<td>276.1</td>
<td>SXS:BBH:0006</td>
<td>1.345</td>
<td>0.234</td>
<td>0.148</td>
<td>-0.161</td>
<td>0.091</td>
<td>0.064</td>
<td>-0.101</td>
<td>-0.14</td>
<td>69.0</td>
<td>13.7</td>
</tr>
</tbody>
</table>

Continued on next page
J. Zweizig¹ (LIGO Scientific Collaboration and Virgo Collaboration)


¹LIGO, California Institute of Technology, Pasadena, CA 91125, USA
²Louisiana State University, Baton Rouge, LA 70803, USA
³American University, Washington, D.C. 20016, USA
⁴Università di Salerno, Fisciano, I-84084 Salerno, Italy
⁵INFN, Sezione di Napoli, Complesso Universitario di Monte S.Angelo, I-80126 Napoli, Italy
⁶University of Florida, Gainesville, FL 32611, USA
⁷LIGO Livingston Observatory, Livingston, LA 70754, USA
⁸Laboratoire d’Annecy-le-Vieux de Physique des Particules (LAPP), Université Savoie Mont Blanc, CNRS/IN2P3, F-74941 Annecy-le-Vieux, France
⁹University of Sannio at Benevento, I-82100 Benevento, Italy and INFN, Sezione di Napoli, I-80100 Napoli, Italy
¹⁰Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany
¹¹Nikhef, Science Park, 1098 XG Amsterdam, The Netherlands
¹²LIGO, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
¹³Instituto Nacional de Pesquisas Espaciais, 12227-010 São José dos Campos, São Paulo, Brazil
¹⁴INFN, Gran Sasso Science Institute, I-67100 L’Aquila, Italy
¹⁵INFN, Sezione di Roma Tor Vergata, I-00133 Roma, Italy
¹⁶Inter-University Centre for Astronomy and Astrophysics, Pune 411007, India
¹⁷International Centre for Theoretical Sciences, Tata Institute of Fundamental Research, Bangalore 560012, India
¹⁸University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA
¹⁹Leibniz Universität Hannover, D-30167 Hannover, Germany
²⁰Università di Pisa, I-56127 Pisa, Italy
²¹INFN, Sezione di Pisa, I-56127 Pisa, Italy
²²Australian National University, Canberra, Australian Capital Territory 0200, Australia
²³The University of Mississippi, University, MS 38677, USA
²⁴California State University Fullerton, Fullerton, CA 92831, USA
²⁵LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France
²⁶Chennai Mathematical Institute, Chennai 603103, India
²⁷Università di Roma Tor Vergata, I-00133 Roma, Italy
²⁸University of Southampton, Southampton SO17 1BJ, United Kingdom
²⁹Universität Hamburg, D-22761 Hamburg, Germany
³⁰INFN, Sezione di Roma, I-00185 Roma, Italy
³¹Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-14476 Potsdam-Golm, Germany
³²APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/Irfu, Observatoire de Paris, Sorbonne Paris Cité, F-75205 Paris Cedex 13, France
³³Montana State University, Bozeman, MT 59717, USA
³⁴Università di Perugia, I-06123 Perugia, Italy
³⁵INFN, Sezione di Perugia, I-06123 Perugia, Italy
³⁶European Gravitational Observatory (EGO), I-56021 Cascina, Pisa, Italy
³⁷Syracuse University, Syracuse, NY 13244, USA
³⁸SUPA, University of Glasgow, Glasgow G12 8QQ, United Kingdom
³⁹LIGO Hanford Observatory, Richland, WA 99352, USA
⁴⁰Wigner RCP, RMKI, H-1121 Budapest, Konkoly Thege Miklós út 29-33, Hungary
⁴¹Columbia University, New York, NY 10027, USA
⁴²Stanford University, Stanford, CA 94305, USA
⁴³Università di Padova, Dipartimento di Fisica e Astronomia, I-35131 Padova, Italy
⁴⁴INFN, Sezione di Padova, I-35131 Padova, Italy
⁴⁵CAMK-PAN, 00-716 Warsaw, Poland
⁴⁶University of Birmingham, Birmingham B15 2TT, United Kingdom
⁴⁷Università degli Studi di Genova, I-16146 Genova, Italy
⁴⁸INFN, Sezione di Genova, I-16146 Genova, Italy