Parameter estimation for compact binary coalescence signals with the first generation gravitational-wave detector network
PARAMETER ESTIMATION FOR COMPACT BINARY ...

J. Worden,\textsuperscript{17} J. Yablon,\textsuperscript{66} I. Yakushin,\textsuperscript{6} H. Yamamoto,\textsuperscript{1} K. Yamamoto,\textsuperscript{64b,64d} C. C. Yancey,\textsuperscript{42} H. Yang,\textsuperscript{51} D. Yeaton-Massey,\textsuperscript{1} S. Yoshida,\textsuperscript{94} M. Yvert,\textsuperscript{4} A. Zadrożny,\textsuperscript{27e} M. Zanolin,\textsuperscript{72} J.-P. Zendri,\textsuperscript{64c} F. Zhang,\textsuperscript{46} L. Zhang,\textsuperscript{1} C. Zhao,\textsuperscript{33} N. Zotov,\textsuperscript{89} M. E. Zucker,\textsuperscript{23} and J. Zweizig\textsuperscript{1}

(LIGO-Virgo Scientific Collaboration)

\textsuperscript{1}LIGO-California Institute of Technology, Pasadena, California 91125, USA
\textsuperscript{2}California State University Fullerton, Fullerton, California 92831, USA
\textsuperscript{3}SUPA, University of Glasgow, Glasgow, G12 8QQ, United Kingdom
\textsuperscript{4}Laboratoire d’Annecy-le-Vieux de Physique des Particules (LAPP), Université de Savoie, CNRS/IN2P3, F-74941 Annecy-Le-Vieux, France
\textsuperscript{5}INFN, Sezione di Napoli, I-80126 Napoli, Italy
\textsuperscript{5}Univesità di Napoli ‘Federico II’, Compresso Universitario di Monte S.Angelo, I-80126 Napoli, Italy
\textsuperscript{5a}INFN, Sezione di Napoli, I-80126 Napoli, Italy
\textsuperscript{5b}Universita` di Napoli ‘Federico II’, Complesso Universitario di Monte S.Angelo, I-80126 Napoli, Italy
\textsuperscript{5c}Universita` di Salerno, Fisciano, I-84084 Salerno, Italy
\textsuperscript{6}LIGO-Livingston Observatory, Livingston, Louisiana 70754, USA
\textsuperscript{7}Cardiff University, Cardiff, CF24 3AA, United Kingdom
\textsuperscript{8}University of Sannio at Benevento, I-82100 Benevento, Italy and INFN (Sezione di Napoli), Italy
\textsuperscript{9}Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany
\textsuperscript{10}Leibniz Universität Hannover, D-30167 Hannover, Germany
\textsuperscript{11}Nikhef, Science Park, 105 1098 XG Amsterdam, The Netherlands
\textsuperscript{12}University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53201, USA
\textsuperscript{13}University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53201, USA
\textsuperscript{13a}INFN, Sezione di Pisa, I-56127 Pisa, Italy
\textsuperscript{13b}Università di Pisa, I-56127 Pisa, Italy
\textsuperscript{13c}Universita` di Siena, I-53100 Siena, Italy
\textsuperscript{14}University of Florida, Gainesville, FL 32611, USA
\textsuperscript{15}National Astronomical Observatory of Japan, Tokyo 181-8588, Japan
\textsuperscript{15a}University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53201, USA
\textsuperscript{15b}University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53201, USA
\textsuperscript{16}University of Western Australia, Crawley, Washington 6009, Australia
\textsuperscript{17}LIGO-Hanford Observatory, Richland, Washington 99352, USA
\textsuperscript{18}University of Birmingham, Birmingham, B15 2TT, United Kingdom
\textsuperscript{19}Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-14476 Golm, Germany
\textsuperscript{20}Montana State University, Bozeman, Montana 59717, USA
\textsuperscript{21}European Gravitational Observatory (EGO), I-56021 Cascina (PI), Italy
\textsuperscript{22}Syracuse University, Syracuse, New York 13244, USA
\textsuperscript{23}LIGO-Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
\textsuperscript{24}APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/Irfu, Observatoire de Paris, Sorbonne Paris Cité, 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France
\textsuperscript{25}Columbia University, New York, New York 10027, USA
\textsuperscript{26}Stanford University, Stanford, California 94305, USA
\textsuperscript{27}Astronomical Observatory Warsaw University, 00-478 Warsaw, Poland
\textsuperscript{27a}Astronomical Observatory Warsaw University, 00-478 Warsaw, Poland
\textsuperscript{27b}CAMK-PAN, 00-716 Warsaw, Poland
\textsuperscript{27c}Bialystok University, 15-424 Bialystok, Poland
\textsuperscript{27d}Bialystok University, 15-424 Bialystok, Poland
\textsuperscript{27e}NCBJ, 05-400 Świdnik-Otwock, Poland
\textsuperscript{27f}Institute of Astronomy, 65-265 Zielona Góra, Poland
\textsuperscript{29}The University of Texas at Brownsville, Brownsville, Texas 78520, USA
\textsuperscript{29}San Jose State University, San Jose, California 95192, USA
\textsuperscript{30}Moscow State University, Moscow, 119992, Russia
\textsuperscript{31}LAL, Université Paris-Sud, IN2P3, F-91898 Orsay, France
\textsuperscript{31b}ESPCI, CNRS, F-75005 Paris, France
\textsuperscript{32}NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA
\textsuperscript{33}University of Western Australia, Crawley, Washington 6009, Australia
\textsuperscript{34}The Pennsylvania State University, University Park, Pennsylvania 16802, USA
\textsuperscript{35}Université Nice-Sophia-Antipolis, CNRS, Observatoire de la Côte d’Azur, F-06304 Nice, France
\textsuperscript{35b}Institut de Physique de Rennes, CNRS, Université de Rennes 1, 35042 Rennes, France
\textsuperscript{35a}INFN, Sezione di Perugia, 06123 Perugia, Italy
\textsuperscript{36}Laboratoire des Matériaux Avancés (LMA), IN2P3/CNRS, Université de Lyon, F-69622 Villeurbanne, Lyon, France
\textsuperscript{37}Washington State University, Pullman, Washington 99164, USA
\textsuperscript{38}INFN, Section di Perugia, 06123 Perugia, Italy

062001-3
Compact binary systems with neutron stars or black holes are one of the most promising sources for ground-based gravitational-wave detectors. Gravitational radiation encodes rich information about source physics; thus parameter estimation and model selection are crucial analysis steps for any detection candidate events. Detailed models of the anticipated waveforms enable inference on several parameters, such as component masses, spins, sky location and distance, that are essential for new astrophysical studies of these sources. However, accurate measurements of these parameters and discrimination of models describing the underlying physics are complicated by artifacts in the data, uncertainties in the waveform models and in the calibration of the detectors. Here we report such measurements on a selection of simulated signals added either in hardware or software to the data collected by the two LIGO instruments and the Virgo detector during their most recent joint science run, including a "blind injection" where the signal was not initially revealed to the collaboration. We exemplify the ability to extract information about the source physics on signals that cover the neutron-star and black-hole binary parameter space over the component mass range \(1M_\odot\rightarrow25M_\odot\) and the full range of spin parameters. The cases reported in this study provide a snapshot of the status of parameter estimation in preparation for the operation of advanced detectors.

I. INTRODUCTION

General relativity predicts that binary systems of compact objects lose energy through the emission of gravitational radiation, a prediction confirmed through binary-pulsar observations [1–3]. During this process they emit a characteristic "chirping" gravitational-wave (GW) signal of predominantly increasing amplitude and frequency. For neutron stars and black holes the signal enters the observational band of the initial (at 40 Hz) and advanced (at 10–20 Hz) ground-based laser-interferometric detectors (e.g., Refs. [4–6]), sweeping through the detection band for a few seconds (depending on the masses of the objects) until coalescence.

The search for GW signatures of compact binary coalescence in LIGO’s most recent sixth science run and Virgo’s science runs 2 and 3 is described in Refs. [7,8]. No detection was reported. However, the operation of advanced instruments—Advanced LIGO and Advanced Virgo—from 2015+ [6,9,10] suggests that the coalescence of compact binaries could be observed in the not too distant future [11]. Once a detection candidate has been identified, the next step in the analysis is to measure the source parameters and test models that describe the underlying physics. This step is critical for enabling studies in astrophysics and fundamental physics in which GW observations are expected to provide a new view of relativistic phenomena. One of the signals studied in this work represents an end-to-end test of this process: a detection candidate was identified with a false alarm rate of 1 in 7000 years \((1.4 \times 10^{-4} \text{ yr}^{-1})\) [7] and was fully characterized in terms of its physical properties before it was revealed to be a "blind" hardware injection, i.e., a simulated signal added by coherently actuating the mirrors of the LIGO and Virgo detectors, without the knowledge of the data analysts [12].

The ability to accurately estimate the parameters of coalescing binaries—including the masses of the components, their spins, the location of a binary on the sky and its distance, and the implications for new insights into the underlying physical processes—have been at the center of several studies. Accurate measurements of the masses provide information about the mass distribution of (binary) black holes and neutron stars, clues to determining the maximum mass of neutron stars (e.g., Ref. [13]), the underlying neutron star equation of state, the minimum mass of stellar-mass black holes and the presence or absence of the so-called "mass gap" [14–16]. Spin measurements are a direct window onto the critical stage of common envelope evolution, and the details of the supernova processes and kicks. Localizing the merger of compact binary systems on the sky may permit the identification of the host, and therefore an opportunity for in-depth studies of the environments in which compact binaries form (e.g., Refs. [17–22]). If the localization region in the sky is sufficiently small, a possible electromagnetic counterpart could be found and multimessenger studies of coalescing compact objects would become possible (e.g., Refs. [23–25]). For instance, the question of whether compact binaries are the progenitors of short gamma-ray bursts could be definitively answered. The direct determination of the luminosity distance to the source opens new possibilities for low-redshift cosmography. In general, coalescing binaries are a laboratory for tests of the behavior of gravity in the strong-field regime. Eventually, when multiple detections are available, studies of source populations will be possible (e.g., Refs. [26,27]).
as well as more stringent tests of the underlying source dynamics [28,29].

Most of these studies use theoretical estimates of parameter uncertainty based on the Cramer-Rao bound [30], which should be valid in the limit of high signal-to-noise ratio (SNR). Initial detections may be too weak for this bound to provide useful guidance. Therefore, a complete Bayesian analysis like the one described below must be used to quantify parameter uncertainties. Other studies have relied on injections into synthetic data. In this paper, we will use injections into real data, which introduces a new set of challenges, including non-Gaussianity and nonstationarity.

The fact that gravitational waveforms used in the analysis are an approximation to the actual radiation produced by astrophysical sources and that the measured strain is affected by the uncertainties in the instrument calibration [31–33] represent additional challenges for making robust inferences on the underlying physics. To study parameter estimation in this regime, we have analyzed several artificial compact binary coalescence (CBC) signals added to real detector data, including the “blind” injection described above, added both in hardware and software to the data collected by the two LIGO instruments (Hanford and Livingston) and the Virgo detector during the most recent joint science run, S6/VSR2-3. The use of injections has been, and continues to be, an essential means to validate the detection process, and as we report here, has been naturally extended to the source-characterization stage of the analysis. Here we exemplify the ability to extract information about the source physics on a selected number of injections that cover the neutron-star and black-hole parameter space over the component mass range 1 M⊙–25 M⊙ and the full range of spin parameters. We consider a spectrum of realistic signal strengths, from candidates observed close to the detection threshold to high-SNR events, and various relative strengths across the instruments of the network. We analyze the signals using a range of waveform models that demonstrate the interplay between (some) systematic bias and statistical uncertainty. To help validate our results, we carry out the analysis with several independent techniques; these are implemented within a specially developed software package part of the LSC Algorithm Library, LALINFERNE [34].

The paper is organized as follows. In Sec. II we give a brief overview of the analysis method. While no detections were claimed in Ref. [7], simulated signals (“injections”) were added to the data, both at a hardware level as the data was being taken and in software afterwards. The hardware injections were performed to validate the end-to-end analysis, including parameter estimation on detection candidates, whereas the injections in software serve as a useful comparison, free of any calibration error in the detectors. Here we report on the analysis using six hardware and software injections, including waveform models for binary neutron star (BNS), neutron star–black hole binary (NSBH) and binary black-hole (BBH) simulations, described in Sec. III. One of these hardware injections was performed without the knowledge of the data analysis teams as part of the “blind injection challenge”; it was successfully detected, as reported in Ref. [7]. We use these injections to illustrate the possible implications for GW astronomy in Sec. IV, and we conclude in Sec. V.

II. ANALYSIS

A. Bayesian inference

Each data segment containing an injected signal was analyzed using a Bayesian parameter estimation pipeline to calculate the probability density function (PDF) of the unknown parameters of the waveform model. We will call θ the vector containing these parameters. The actual content and dimension of θ, i.e. the dimensionality of the parameter space, depend on the waveform model used for the analysis (see Sec. II B).

The posterior distribution of θ given a model H is given by Bayes’ theorem,

\[ p(\theta | \{d\}, H) = \frac{p(\theta | H) p(\{d\} | \theta, H)}{p(\{d\} | H)} , \]

(1)

where \( p(\theta | H) \) is the prior distribution of \( \theta \), describing knowledge about the parameters within a model \( H \) before the data is analyzed, and \( p(\{d\} | \theta, H) \) is the likelihood function, denoting the probability under model \( H \) of obtaining the data set \( \{d\} \) for a given parameter set \( \theta \). The likelihood is a function of the noise-weighted residuals after subtracting the model from the data, and is thus a direct measure of the goodness of fit of the model to the data.

The optimal network SNR is defined as

\[ \text{SNR} = \sqrt{\sum_{\text{det}} \int_{f_{\text{low}}}^{f_{\text{high}}} \frac{|s_{\text{det}}(f, \theta)|^2}{S_{\text{det}}(f)} df} , \]

(2)

where the sum is taken over each detector \( \text{det} \) with \( s_{\text{det}} \) the signal in that detector and \( S_{\text{det}}(f) \) its noise power spectral density (PSD).

Our model for the likelihood function is based on the assumption that the noise is stationary and Gaussian, and uncorrelated in different frequency bins. Although we do not expect this assumption to be precisely true for real detector noise, limited investigations suggest that this is an acceptable approximation when the data is of good quality [35].

The denominator of Eq. (1), \( P(\{d\} | H) = Z_H \), is the evidence for the model \( H \). As it is a normalization constant, the evidence does not affect the estimation of the parameters for a particular model \( H \), but it does allow us to compare the ability of different models to describe the data. The Bayes factor between two models,
which quantifies the relative support given to each model by the data, is given by

\[
Z_1 = \frac{P(d|H_1)}{P(d|H_2)} = \frac{\int_\theta p(\theta | H_1) P(d|\theta, H_1) d\theta}{\int_\lambda p(\lambda | H_2) P(d|\lambda, H_2) d\lambda},
\]

(3)

where \(H_1\) and \(H_2\) represent the two different models and \(\theta, \lambda\) their respective model-parameter vectors. The parameter spaces can be completely unrelated, even of different dimensionality, which allows us to compare models which range from nine parameters in the case of nonspinning systems, up to 15 in the case of a system with two spinning components. Throughout we will quote evidence relative to the Gaussian noise model, \(\ln(Z) = \ln(Z_H) - \ln(Z_{\text{Gaussian}})\).

The high dimensionality of the parameter space and the complicated structure of the likelihood function make it impractical to exhaustively calculate posterior quantities. Instead we rely on stochastic sampling of the posterior PDF, which provides us with an approximation of the underlying true distributions, e.g. by binning the samples to produce histograms. The analysis was performed using LALINFERENCE [34], which allows calculation of the prior, likelihood and templates using standardized functions. Because accurate sampling of the posterior PDF is a difficult task, we used two independent sampling algorithms on all but a few cases, cross-comparing results to confirm convergence. These were based on Markov Chain Monte Carlo (MCMC) [36,37] and nested sampling [38] techniques, and were included in the LALINFERENCE package. We found good agreement between the two techniques, giving us confidence in the results. For a sample of the cases we used a third algorithm, MultiNest [39], as an additional check. The nested sampling and MultiNest algorithms directly produce an estimate of the evidence [40]; we also computed the evidence from MCMC results using direct integration [41].

B. Waveforms and source parameters

Unless otherwise specified we shall use a system of units in which \(c = G = 1\). The waveform as measured at a generic detector can be written in the frequency domain as [42]

\[
s(f, \tilde{\theta}) = [F_+(\alpha, \delta, \psi) h_+(f, \tilde{\theta}) + F_\times(\alpha, \delta, \psi) h_\times(f, \tilde{\theta})]e^{2\pi if\Delta t},
\]

(4)

where

(i) \(F_+(\alpha, \delta, \psi)\) and \(F_\times(\alpha, \delta, \psi)\) are the known antenna beam pattern of the detector, which gives the amplitude response of the antenna to the + and × polarizations;

(ii) \(\alpha\) and \(\delta\) are the right ascension and declination of the source;

(iii) \(\psi\) is the polarization angle (see, e.g., Ref. [43]);

(iv) \(\Delta t\) is the time delay between the arrival of the signal at the detector and at a common reference frame (e.g., the center of the Earth), a known function of \(\alpha\) and \(\delta\);

(v) \(h_+(f, \tilde{\theta})\) and \(h_\times(f, \tilde{\theta})\) are the two independent polarizations of the signal, with \(\tilde{\theta} = \tilde{\theta} \setminus \{\alpha, \delta, \psi\}\) (i.e. \(\tilde{\theta}\) includes all the waveform parameters except for right ascension, declination and the polarization angle).

The actual form of \(h_+(f, \tilde{\theta})\) and \(h_\times(f, \tilde{\theta})\) depends on which model one is considering.

The GW signal can be written using the post-Newtonian (pN) expansion [44]. We used a total of nine different waveform models (or approximants) in our analysis, including the TaylorF2 (which gives a pN-based, analytic expression for the frequency-domain waveform) [45,46], SpinTaylorT4 (which is a time-domain pN model supplemented with spin-precession equations) [47] and IMRPhenomB (which is a TaylorF2-like waveform model supplemented with a merger-ringdown and

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
Model & Name & Spin effects & Merger and ring-down & Reference \\
\hline
TF2 & TaylorF2 & no & no & [45] \\
TF2_2 & TaylorF2 @ 2 pN & no & no & [45] \\
TF2_RS & TaylorF2 RedSpin & aligned & no & [46] \\
ST_NS & Nonspinning STPN & no & no & [47] \\
ST_SA & Aligned spin STPN & aligned & no & [47] \\
ST & Full STPN & yes & no & [47] \\
IMRPB & IMRPhenomB & yes & yes & [48] \\
TF2_25\(a\) & TaylorF2 @ 2.5 pN & no & no & [45] \\
ST_25\(a\) & Full STPN @ 2.5 pN & yes & no & [47] \\
\hline
\end{tabular}
\end{table}

\(a\)Used only for the blind hardware injection.
calibrated to agree with numerical relativity simulations) [48] approximants with binary spin parameters either disabled, set aligned to the orbital angular momentum, or enabled and allowed full freedom of orientation. Table I lists the approximants used. For example, the TaylorF2 waveform at lowest amplitude order can be written in the frequency domain combined with the stationary-phase approximation [42],

$$h_\times(f) = \frac{1 + \cos^2 \iota}{2 d_L} h(f), \quad \text{(5)}$$

$$h_\times(f) = i \cos \iota \frac{d_L}{c} h(f), \quad \text{(6)}$$

where

$$h(f) = \frac{M^5}{\pi^3} \sqrt{\frac{5}{24}} e^{-7/6} e^{-i \phi(f)}, \quad \text{(7)}$$

$$\psi(f) = 2 \pi f t_c - \phi_c - \frac{\pi}{4} + \frac{3}{128 \eta v^3} \sum_{k=0}^N \alpha_k v^k, \quad \text{(8)}$$

with

$$v = (\pi (m_1 + m_2) f)^{1/3}.$$

We see that the two polarizations depend on six parameters (the $\alpha_k$ are known functions of $M$ and $\eta$ [45], up to $N = 7$).

(i) $M$. The chirp mass of the system: $M = (m_1 m_2)^{3/5} (m_1 + m_2)^{-1/5}$, where $m_1$ and $m_2$ are the component masses.

(ii) $\eta$. The symmetric mass ratio, defined as $\eta = (m_1 m_2)/(m_1 + m_2)^2$.

(iii) $\iota$. The inclination angle, i.e. the angle between the orbital angular momentum vector and the line of sight.

(iv) $d_L$. The luminosity distance to the source.

(v) $t_c$. An arbitrary reference time, usually chosen to be the time of coalescence of the binary.

(vi) $\phi_c$. The phase of the waveform at the reference time $t_c$.

Together with $\alpha$, $\delta$ and $\psi$, these six parameters form the nine-dimensional vector $\hat{\theta}$ for the TaylorF2 model. The upper value of the index $k$ in Eq. (8) determines the phase pN order, with $N = 0$ being the lowest order (0 pN) and $N = 7$ the 3.5 pN order. We assume that the eccentricity of the binary system is negligible, as radiation reaction circularizes orbits efficiently before the signal frequency enters the detector’s observational band [49]. We can also see from Eqs. (5) and (6) that the inclination $\iota$ and distance $d_L$ are strongly correlated. This correlation is illustrated in Sec. III.

When spins are present and we constrain the spin vectors to be aligned with the orbital angular momentum of the binary, we need two additional parameters, bringing $\hat{\theta}$ to 11 parameters. We use the spin magnitudes $a_1$ and $a_2$, defined as $a_i = (\hat{\mathbf{s}}_i \cdot \hat{L})/m_i^2$, where $\hat{\mathbf{s}}_i$ and $m_i$ are the spin and mass of the object $i$, and $\hat{L}$ is the unit vector along the orbital angular momentum. To account for both aligned and antialigned cases, we allow $a_i$ to be in the range $[-1, 1]$. When generic spins are considered, we need six additional parameters, and $\hat{\theta}$ becomes a 15-dimensional vector. In this case the spin magnitudes $a_i$, defined as $a_i = |\hat{\mathbf{s}}_i|/m_i^2$ and in the range $[0, 1]$, and four angles specify the orientations.

The presence of arbitrary spins will alter both the amplitude and phase of the signal. The pN expansion of the phase evolution is changed to include additional terms dependent on the components of the spin vectors. If the spins are not aligned with the orbital angular momentum, then the spin-orbit and spin-spin coupling will cause the orbital angular momentum vector to precess around the total angular momentum vector. As this implies that $\iota$ (and $\psi$) become time dependent, the values of these quantities at the GW frequency 40 Hz are used for parameter estimation. We refer the reader to Ref. [47] for more details about SpinTaylorT4.

The pN approximation is valid for the inspiral phase of the binary, when the two objects are still far apart [50,51]. For this reason, the TaylorF2 and SpinTaylorT4 waveforms must terminate at an approximate point where they begin to break down, with the TaylorF2 ending at the innermost stable circular orbit [45,46] frequency of a test particle orbiting a Schwarzschild black hole, and the SpinTaylorT4 terminating at the minimum energy circular orbit [47]. IMRPhenomB waveforms extend the signal into the merger and ring-down stages of the binary coalescence, making use of a phenomenological model tuned by comparison with numerical relativity simulations. These later stages take place at higher frequencies than the inspiral, and hence are more important for higher-mass systems where the frequency scale is lower, and where the merger and ring-down contribute a larger proportion of SNR by being in a more sensitive band of the detectors.

For the analysis of nonblind hardware and software injections (the first seven models of Table I), we considered the pN expansion either up to the 2 pN order in phase or up to the 3.5 pN order in phase (2.5 pN for spin effects when included), and 0 pN in amplitude. For the blind hardware injection we used approximants at 2.5 pN order in phase, as the blind injection only included terms up to this order [7,12] (see Sec. III C below). The different waveforms use different post-Newtonian expansions of the binary phase, and do not give identical results in general. When analyzing a real GW signal, the ignorance of the higher-order pN terms will cause a systematic bias in the inference of the binary parameters, and using a range of different approximants here will give us an indication of the size of this error compared to the statistical uncertainty in the results.
C. Priors

As shown in Eq. (1), the posterior distribution of $\hat{\theta}$ depends on both the likelihood and prior distributions of $\hat{\theta}$. We used the same prior density functions (and range) for all the analyses, uniform in the component masses with the range $1 M_\odot \leq m_1, m_2 \leq 30 M_\odot$, and with the total mass constrained by $m_1 + m_2 \leq 35 M_\odot$. This range encompasses the low-mass search range used in Ref. [7], where $1 M_\odot \leq m_1 \leq m_2 \leq 24 M_\odot$ and $m_1 + m_2 \leq 25 M_\odot$. The prior density function on the location of the source was taken to be uniform in volume, constrained between luminosity distances $d_L \in [1, 100] \text{ Mpc}$. We used an isotropic prior on the orientation of the orbital plane of the binary. For analyses using waveform models that account for possible spins, the prior on the spin magnitudes, $a_1, a_2$, was taken to be uniform in the range $[0,1]$ (range $[-1, 1]$ in the spin-aligned cases), and the spin angular momentum vectors were taken to be isotropic.

The computational cost of the parameter estimation pipeline precludes us from running it on all times; therefore, the parameter estimation analysis relies on an estimate of the coalescence-time as provided by the detection pipeline [7]. In practice, a 200 ms window centered on this trigger time is sufficient to guard against the uncertainty in the coalescence-time estimates from the detection pipeline [7]. The PSD of the instrumental noise was estimated using 1024 s of data after the end of the analyzed segment. We verified that varying the methods of PSD estimation (using $\pm 512$ seconds spanning the signal trigger time, and using median and mean estimation methods) had a negligible effect on the parameter-estimation results—smaller than the systematic uncertainties of parameter estimation—although varying the time during which the PSD is estimated can have a significant effect (see Sec. III.D).

Calibration errors, which can influence the reconstructed amplitudes, phase, and timing of the data [31], have the potential to affect parameter-estimation results. An analysis of the effect of calibration errors, which considered mock errors similar to those expected during the S6 and VSR2/3 runs, concluded that such errors are unlikely to cause a significant deterioration in parameter-estimation accuracy [33] at the moderate SNRs considered in this paper.

III. SIMULATIONS

Over the course of the LIGO S6 and Virgo VSR2/3 science runs a series of hardware injections have been carried out where the arm lengths of the three detectors were physically changed to simulate the passing of a GW. As an end-to-end test of the search pipeline during the science runs, a signal was added to the data via a hardware injection, without the data analysts’ knowledge (“blind”). The parameters and template family were revealed only after the search was complete, and we include a retrospective analysis of this blind injection in Sec. III.C (see Ref. [12] for the parameter estimation carried out before the injection was revealed).

We have also added several simulated software injections into real detector noise from those runs. Below, we describe the results of parameter-estimation analysis on several of these injections, whose parameters are listed in Table II. We present the posterior probability density functions on the source parameters using a coherent, multi-detector data model with the seven waveform families except IMRPhenomB, which has a negligible contribution to the signal-to-noise ratio at higher frequencies. The 32-second segments of time-domain data were Tukey windowed, with a 0.4 s roll-off on either side of the segment. The PSD of the instrumental noise was estimated using 1024 s of data after the end of the analyzed segment. We verified that varying the methods of PSD estimation (using $\pm 512$ seconds spanning the signal trigger time, and using median and mean estimation methods) had a negligible effect on the parameter-estimation results—smaller than the systematic uncertainties of parameter estimation—although varying the time during which the PSD is estimated can have a significant effect (see Sec. III.D).

Calibration errors, which can influence the reconstructed amplitudes, phase, and timing of the data [31], have the potential to affect parameter-estimation results. An analysis of the effect of calibration errors, which considered mock errors similar to those expected during the S6 and VSR2/3 runs, concluded that such errors are unlikely to cause a significant deterioration in parameter-estimation accuracy [33] at the moderate SNRs considered in this paper.

D. Data description

Data from the multidetector network consisting of two LIGO instruments (H1 and L1) and Virgo (V1) were used coherently in the analysis. From each detector a total of 32 seconds of data $[-30, +2] \text{s}$ around the global positioning system time of the injection were analyzed. The initial frequency of the analysis was 40 Hz, which is low enough so that the detectors are not significantly sensitive to the start of the signals and template. The sampling rate was 2048 Hz (4096 Hz for the BNS, Sec. III.A.2), corresponding to a 1024 Hz (2048 Hz) Nyquist frequency, high enough to include the entire waveform for all models.
described above. In Appendix A we present 90% credible intervals obtained using the waveform models in Table I for each injection in Table II.

A. Nonspinning hardware injections

1. Binary black hole

We first describe the analysis of a hardware injection corresponding to a binary black hole with nonspinning $4.91M_\odot$ and $4.02M_\odot$ components, injected at a network SNR of 13 (SNRs of 8.7 in the Hanford and Livingston detectors and 4.4 in the Virgo detector). This injection was made with effective one-body numerical relativity (EOBNR) waveforms, using the EOBNRv1 version as described in Ref. [54].

Figure 1 shows a comparison of the posterior PDF of the mass parameters for different models. The chirp mass has a low statistical uncertainty of $\sim 1\%$, with the greatest statistical uncertainty claimed by models that allow for non-zero spin magnitudes due to interparameter degeneracies. The 90% credible intervals on the chirp mass obtained with TaylorF2 templates just exclude the true value, an indication of the systematic bias due to the waveform differences between the injected EOBNR waveform and these

![FIG. 1](image-url)

FIG. 1 (color online). (Left) Posterior probability distributions for the chirp mass $\mathcal{M}$ of the nonspinning BBH hardware injection (Sec. III A 1) for the seven signal models considered. The injected value is marked with a vertical red line. (Right) Overlay of 90% probability regions for the joint posterior distribution on the component masses $m_1$, $m_2$ of the binary. The true value is marked by the blue star. Models which allow for nonzero spins find wider PDFs for the coupled mass parameters.

![FIG. 2](image-url)

FIG. 2 (color online). Joint posterior probability regions for the location and inclination angle of the nonspinning BBH hardware injection (Sec. III A 1) for the seven signal models considered. (Left) The binary is constrained to two neighboring regions of the sky. (Right) The distance and inclination, like the sky location, are estimated with a similar accuracy in models that include or exclude spins.
templates. These differences are also responsible for the particularly strong bias in the mass ratio for 2.0 pN TaylorF2 templates, exceeding the statistical measurement uncertainty in this parameter. In contrast, we find that by using the 3.5 pN TaylorF2 templates the systematic bias is reduced to less than the statistical uncertainty at the 90% credible interval. The uncertainty in the mass ratio is typically larger than that in the chirp mass, leading to the characteristic thin, correlated joint distribution for $m_1$ and $m_2$ evident in Fig. 1 (right).

Figure 2 shows a selection of PDFs for the extrinsic parameters of the binary black hole (BBH) source: the recovered sky position, distance and inclination angle. The binary can be constrained to two neighboring regions of the sky representing the reflection of the source location through the plane of the detectors, which cuts this contiguous region in two. The correlations between intrinsic parameters (masses and spins, see Refs. [55–57]) and extrinsic parameters are generally weaker than correlations between different parameters within the same group. Thus both models that include and models that exclude nonzero spins have similar statistical measurement uncertainties for extrinsic parameters.

Finally, Fig. 3 shows individual posteriors for the two dimensionless spin magnitudes, obtained using the model ST (Table I), full-spin STPN waveforms. Although spin is not very strongly constrained, particularly for the lower-mass secondary, both spin measurements are consistent with the true value of 0 spin. The absence of a strong constraint is due to both the small difference in masses and nearly face-on inclination. This inability to constrain the spin is reflected in the evidence for each template family, shown in Table III, where all models have the same evidence within the error bars.

2. Binary neutron star

We also analyzed a hardware injection simulating a nonspinning neutron-star binary system with $1.81M_\odot$ and $1.65M_\odot$ components, injected at a network SNR of 36 (SNR of 26 in the Hanford detector, 17 in the Livingston detector and 20 in the Virgo detector). This injection was

<table>
<thead>
<tr>
<th>Sec.</th>
<th>TF2</th>
<th>TF2_2</th>
<th>TF2_25</th>
<th>TF2_RS</th>
<th>ST_NS</th>
<th>ST_SA</th>
<th>ST</th>
<th>ST_25</th>
<th>IMRPB</th>
</tr>
</thead>
<tbody>
<tr>
<td>III A 1</td>
<td>69</td>
<td>70</td>
<td>-</td>
<td>70</td>
<td>69</td>
<td>71</td>
<td>71</td>
<td>-</td>
<td>73</td>
</tr>
<tr>
<td>III A 2</td>
<td>686</td>
<td>699</td>
<td>-</td>
<td>694</td>
<td>685</td>
<td>697</td>
<td>694</td>
<td>-</td>
<td>668</td>
</tr>
<tr>
<td>III B 1</td>
<td>154</td>
<td>146</td>
<td>-</td>
<td>153</td>
<td>158</td>
<td>158</td>
<td>157</td>
<td>-</td>
<td>155</td>
</tr>
<tr>
<td>III B 2</td>
<td>48</td>
<td>48</td>
<td>-</td>
<td>50</td>
<td>50</td>
<td>52</td>
<td>64</td>
<td>-</td>
<td>52</td>
</tr>
<tr>
<td>III C</td>
<td>-</td>
<td>-</td>
<td>136</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>213</td>
<td>-</td>
</tr>
</tbody>
</table>
made with the same EOBNR waveform family as in Sec. III A 1 above.

Figure 4 shows a comparison of the posterior PDF of the mass parameters for the different models. The chirp mass has a low statistical uncertainty of \( \lesssim 0.1\% \) thanks to the high network SNR, with the greatest statistical uncertainty claimed by models that allow for nonzero spin magnitudes due to interparameter degeneracies, similar to Sec. III A 1. The greater statistical uncertainty of the IMRPhenomB waveform model stems from the fact that this model does not describe such a low-mass system \((m_1 + m_2 < 10 M_\odot)\) very well [48]. Correspondingly, it has the lowest Bayes factor in Table III. The 90% credible intervals on the chirp mass obtained with TaylorF2 templates include the true value. Unlike in Sec. III A 1, the merger present in the injected EOBNR waveform happens at a frequency (1.2 kHz) where the detectors are not very sensitive. Figure 5 shows the extrinsic parameters of the BNS source: sky position, distance and inclination angle. In this case, the data were such that the injected distance and inclination angle lie far outside the 90% credible intervals. This

![FIG. 4 (color online).](image1)

Left: Posterior probability distributions for the chirp mass \( M \) of the nonspinning BNS hardware injection (Sec. III A 2) for the seven signal models considered. The injected value is marked with a vertical red line. Right: Overlay of 90% probability regions for the joint posterior distribution on the component masses \( m_1, m_2 \) of the binary. The true value is marked by the blue star.

![FIG. 5 (color online).](image2)

Left: Joint posterior probability regions for the location and inclination angle of the nonspinning BNS hardware injection (Sec. III A 2) for the seven signal models considered. The binary can be constrained to two regions of the sky representing the reflection of the source location through the plane of the detectors. Right: The distance and inclination, like the sky location, are estimated with a similar accuracy in models that include or exclude spins. The data were such that the injected distance and inclination angle are far outside the 90% credible intervals.
could be due to a particularly unlikely noise realization, some nonstationarity or non-Gaussianity in the noise (i.e., a “glitch”), or some issue with the way the hardware injection was carried out.

**B. Spinning software injections**

1. **Binary black hole**

We simulated the signal from a binary black hole with misaligned spinning components with a SpinTaylor software injection. This binary consists of 6M$_\odot$ and 5M$_\odot$ black holes, with dimensionless spin magnitudes of 0.6 and 0.8, misaligned with the orbital angular momentum by angles of 40° and 150°, respectively, and an optimal network SNR of 19 (SNR of 17 in the Hanford detector, 6.5 in the Livingston detector and 4.9 in the Virgo detector). The misalignment between the spins and the orbital angular momentum causes the plane of the binary to precess, producing both amplitude and phase modulations in the received GW signal (see Sec. II B).

Figure 6 shows a comparison of the posterior PDFs of the mass parameters inferred with different template models. The mass ratio is again severely biased for the 2 pN TaylorF2 model, which is not surprising for a 3.5 pN injection with spinning components as both the pN order and the spin alter the phase evolution of the signal.

FIG. 6 (color online). (Left) Posterior probability distributions for the chirp mass $\mathcal{M}$ of the spinning BBH software injection (Sec. III B 1) for the seven signal models considered. The injected value is marked with a vertical red line. (Right) Overlay of 90% probability regions for the joint posterior distribution on the component masses $m_1$, $m_2$ of the binary.

FIG. 7 (color online). Joint posterior probability regions for the location and inclination angle of the spinning BBH software injection (Sec. III B 1). (Left) The binary’s true location lies just outside of the 90% credible interval. (Right) The degeneracy in distance and inclination prevents either parameter from being accurately constrained individually.
Figure 7 shows the recovered sky location, distance and inclination angle of the source. For all these parameters the injection value lies just outside of the 90% credible regions.

The spin magnitudes are again poorly constrained, with nonzero support across the entire allowed range, as shown in Fig. 8. Even though the injected signal was simulated from a system with high spin magnitudes (0.6 and 0.8), the low difference in masses, the near antialignment of the spins with the orbital angular momentum and especially the face-on inclination conspire to give the poor spin estimates. In addition, with this weak precession effect, the spin tilts (angles between the spin vector and the orbital angular momentum) are also poorly constrained. A comparison of the evidences in Table III indicates that all models have the same evidence within the error bars.

2. Neutron star–black hole

We simulated the signal from a neutron star–black hole system with misaligned black-hole spin components with a SpinTaylor software injection. This binary consists of a 10 M$_\odot$ black hole with dimensionless spin magnitude of 0.7, misaligned with the orbital angular momentum by an angle of 130°, and a 1.4 M$_\odot$ nonspinning neutron star, at an optimal network SNR of 13 (SNR of 7.9 in the Hanford detector, 9.2 in the Livingston detector and 3.6 in the Virgo detector).

Figure 9 shows a comparison of the posterior PDFs of the mass parameters inferred with different template models. Figure 10 shows the recovered sky location, distance and inclination angle of the source. The joint distance-inclination posterior degeneracy prevents either parameter from being measured precisely. The model describing the data most accurately—with the highest Bayes factor in Table III—is the Full STPN, which delivers the most accurate parameter estimates, as is expected since this model was used for the injection.

Figure 11 shows the posterior PDFs for the spin parameters. The spin magnitude of the neutron star is unconstrained due to the large difference in masses. The spin of the more massive black hole encodes more information in the waveform, but is also poorly constrained due to the almost face-on inclination.

C. Blind hardware injection

The search pipeline described in Ref. [7] identified a GW candidate occurring on 16 September 2010 at 06:42:23 UTC. A Bayesian analysis was performed using the algorithms and implementations described above, where parameter estimates varied significantly depending on the exact model used for the gravitational waveform.

Following the completion of the analysis, the event was revealed to be a blind injection. Further investigation revealed several problems with the pre-un-blinding parameter estimation:

(i) The template signal included phase corrections only up to 2.5 pN order, which is an outlier in the post-Newtonian expansion; see Ref. [58]. At that time (before the blind injection was revealed as such) parameter estimation was not carried out with templates at this order, leading to a significant bias in the mass ratio, and hence the component masses.

(ii) The signal to be injected in the Hanford and Livingston sites had the wrong sign, making the signal incoherent between the LIGO and Virgo detectors. This caused strong biases in the estimated values of the extrinsic parameters.
(iii) Lastly, an injection software bug artificially set to zero one of the phase terms in the injected waveform.

We present here an analysis using templates at 2.5 pN order in phase, performed after the un-blinding. We artificially introduce a sign flip in the templates for the Hanford and Livingston detectors in order to match the injected waveform. Figure 12 shows a comparison of the posterior PDFs of the mass parameters for TaylorF2 and SpinTaylor waveform models. The fully spinning, pre-processing analysis with SpinTaylor templates at 2.5 pN order in phase (i.e., the same template as used in the injection, up to the injection software bug) centers on the correct masses for this neutron star–black hole system, $m_1 = 24.81 \text{ M}_\odot$ and $m_2 = 1.74 \text{ M}_\odot$. However, the systematic bias due to fixing the spin to zero in the analysis of this spinning injection is very significant, and leads to the wrong conclusion that a BBH system is observed if the TaylorF2 model is used.

Figure 13 shows the recovered sky position, distance and inclination angle of the source. The sky location recovered is contained in several distinct regions, spread in an arc on the sky. This behavior is consistent with two detectors contributing the majority of the SNR 15 for this source (an SNR of 11 in the Hanford detector, 9.8 in the Livingston detector and 4.1 in the Virgo detector). On top of this first-order constraint, spin projection effects can break the degeneracy, as discussed in Ref. [59].

The large difference in masses, $m_1 = 24.81 \text{ M}_\odot$, $m_2 = 1.74 \text{ M}_\odot$, in combination with the high inclination allows for the magnitude of the spin of the massive component to...
be measured with an accuracy \( \sim 10\% \), as shown in Fig. 14. Meanwhile, the spin of the light component is unconstrained, and the posterior on its magnitude is consistent with the prior. Correspondingly, the evidences in Table III favor the spinning model. Signals from face-on binaries are louder than signals from high-inclination sources, making a face-on binary more likely to be detected. As such the poor spin constraints in the previous sections are more representative of the expected observations. However, the study presented here does not include a statistically sufficient number of signals to cover the range of possible measurements. A wider study is ongoing to predict more accurately the constraints the LIGO/Virgo network will be able to put on spins.

A deeper analysis of the data around the blind hardware injection revealed some non-Gaussian behavior of the Livingston detector before and during the injected signal. From Fig. 15 it can be seen that the analysis of the Hanford detector data or Livingston detector data alone produces mass estimates consistent with the full network analysis using data from all three detectors, but the Livingston detector data cannot constrain the parameters as precisely as the Hanford detector data, showing multiple modes. This leads us to conclude that detector glitches have the potential to reduce parameter-estimation accuracy. However, the coherent analysis of the data from the multi-detector network increases robustness against the effect of a glitch in a single instrument by requiring that the

![Figure 11](image1.png)

**FIG. 11** (color online). Posterior probability distributions for the dimensionless spin magnitude of the heavier (left) and lighter (right) components of the binary from the spinning NSBH software injection (Sec. III B 2), as inferred in the model ST (Table I), full-spin STPN; the true values are shown with vertical red lines.

![Figure 12](image2.png)

**FIG. 12** (color online). (Left) Posterior probability distributions for the chirp mass \( M \) of the blind injection (Sec. III C) for signal models at 2.5 pN. The injected value is marked with a vertical line. (Right) Overlay of 90% probability regions for the joint posterior distribution on the component masses \( m_1, m_2 \) of the binary. The bias introduced by an analysis with a model which disallows spin is clear.
recovered parameters are consistent with all data sets. This effect could be mitigated by including such glitches in the model of the data, as suggested in Ref. [60].

**D. The effects of different noise realizations**

To exemplify the effect of noise on the recovered PDFs, ten software injections (see Fig. 16) with parameters identical to the nonspinning BBH hardware injection (Sec. III A 1) were created with injection times in succession at 10-second intervals. With each injection at a different time, the estimated PSDs were also (slightly) different (see Fig. 17), reflecting the slow change of the PSD as a function of time. All the signals were both injected and recovered using the TaylorF2 signal model to eliminate systematic effects from using the wrong waveform family. Figure 16 shows the joint distributions of the recovered component masses as well as the distance and inclination. The spread of the 90% credible intervals for different injection times is comparable to the spread observed using different waveform models, as shown in Figs. 1 and 2, illustrating the dependence on the specific model and realization of the noise used. Fixing the injection time for a series of different PSDs gives narrower PDFs than in Fig. 16. For the particular data segment considered here, the same is also true for a fixed PSD and a series of injection times but to a lesser extent. We plan to address

![Figure 13](image13.png)

**FIG. 13 (color online).** Joint posterior probability regions for the location and inclination angle of the blind injection (Sec. III C). (Left) The sky location is constrained to several distinct regions lying along a half circle on the sky. (Right) Again, the characteristic V-shape degeneracy in distance and inclination is evident.

![Figure 14](image14.png)

**FIG. 14 (color online).** Posterior probability distributions for the dimensionless spin magnitude (left) and tilt angle (angle between the spin vector and the orbital angular momentum, measured at 40 Hz, right) of the heavier component of the binary from the blind injection (Sec. III C), as inferred with model full-spin STPN at 2.5 pN order in phase (Table I); the true values are shown with vertical red lines. The large difference in masses allows for the spin of the massive component to be measured. On the other hand, the spin of the light component is unconstrained.
this in the future by performing the analysis with a model in which the PSD is not fixed, but parametrized as a function of a suitable set of unknown parameters to be marginalized over. For the ten software injections the true injected parameter is found within the given 90% credible interval in 90% of the injections, as expected. Figure 18 shows a direct comparison between a nonspinning BBH hardware injection and three software injections replicated with the same parameters. All are evaluated using the PSD of the hardware injection with the software injections at 100, 500 and 1000 seconds before the hardware injection.

Again, there are variations in the recovered PDFs but the hardware injection is not an outlier. The variations are caused by the inherent dependency of the analysis on the particular noise realization used. Statistical fluctuations appear to be able to dominate any systematic bias introduced by the method of performing the hardware injection, such as the actuation function used in modeling the impulse applied to the mirrors to produce the desired signal.

E. Computational cost

The choice of approximant has strong implications for the computational cost of the analysis, which varied from a few hours to three weeks of real time, using up to three 8-core CPUs. Frequency-domain templates, such as TaylorF2 and IMRPhenomB, generally require much less computational time than the time domain SpinTaylorT4, as they do not require a numerical solution of the orbital frequency and spin evolution equations or a Fourier transform before computing the likelihood. The time domain SpinTaylorT4 can include the full six spin parameters, allowing for precession of the orbital plane, whereas the frequency-domain templates can only include spin aligned with the orbit. The computational cost is also strongly influenced by the length of the templates, which ranged from 2.16 s for the blind injection (Sec. III C) to 18 s for the BNS hardware injection (Sec. III A 2), using a 40 Hz starting frequency. An additional effect is the higher frequency at which a BNS will merge, forcing us to analyze the data at a higher sampling rate. A higher signal-to-noise ratio also increases the run time with more iterations required to reach the more contrasted posterior distribution. This effect can be somewhat managed by increasing the number of parallel tempering chains in the MCMC algorithm [36,37], or by using parallel nested sampling runs, reducing the overall run time at the expense of using more computing.

FIG. 15 (color online). Overlay of 90% probability regions for the joint posterior distribution on the component masses $m_1, m_2$ of the binary for the blind injection (Sec. III C), as inferred with model ST_25 (Table I), full-spin STPN (at 2.5 pN order in phase), using data from the Hanford detector only, the Livingston detector only or the whole Hanford-Livingston-Virgo network.

FIG. 16 (color online). Overlay of 90% probability regions for the joint posterior probability distribution of the component masses $m_1, m_2$ of the binary (left) and the distance and inclination (right) for a series of signal injection times. Both sets of distributions show a spread comparable to the spread observed between different waveform models, as shown in Figs. 1 and 2.
FIG. 17 (color online). The plots show in red ten overlaid noise power spectral densities (top left: LIGO-Hanford; top right: LIGO-Livingston; bottom: Virgo), each computed over 1024 s of data, and separated by 10 s. In black is the PSD used in Figs. 1 and 2. The PSDs were used in the analyses whose results are shown in Figs. 16 and 18 and provide an indication of the fluctuation of the noise in the instruments over a time span of minutes.

FIG. 18 (color online). (Left) Posterior probability distributions for the chirp mass $\mathcal{M}$ of the nonspinning BBH injection for signal model TF2 (Table I), TaylorF2 at 3.5 pN order, for the hardware injection (Sec. III A 1) and three software injections of the same parameters injected 100, 500 and 1000 seconds before the hardware injection (see Sec. III D). The injected value is marked with a vertical red line. (Right) Overlay of 90% probability regions for the joint posterior distribution on the inclination and distance of the binary. Different realizations of the noise, rather than differences between hardware and software injections, are the likely reason for the variations in recovered PDFs.
units in parallel. Other promising cost-saving techniques include waveform interpolation to ease the burden of computing millions of templates (e.g., Ref. [61]).

In a future triggered analysis, where the parameter estimation is run as a follow-up to a search algorithm, some guidance on the chirp mass from the search pipeline could be used to choose the length of the data segment to be analyzed so as to minimize the cost while allowing enough time for the full waveform to be generated. In the analysis presented here, the run time varied from hours for the low-SNR, short segment (high mass) injections using frequency-domain templates, up to several weeks for the high-SNR, long segments (low mass) using time-domain fully precessing templates.

IV. IMPLICATIONS FOR GRAVITATIONAL-WAVE ASTRONOMY

Although the five injections in Table II are not statistically sufficient to cover the range of source possibilities, we do have samples representing the three expected classes—BNS, NSBH and BBH—at a range of SNR and some including spin. This allows us to highlight some of the general features of parameter estimates that are likely to be encountered with future analyses of signal candidates in advanced gravitational-wave detectors.

We have found, in accordance with expectations, that the chirp mass parameter can be resolved with a very small statistical uncertainty of only a few percent or even lower, even when the SNR is relatively low, as in the case of the BBH hardware injection (Sec. IIIA.1) and the NSBH software injection (Sec. IIIB.2), which had network SNRs of 13 (see Figs. 1 and 9). With the lower-mass and higher-SNR BNS hardware injection, the systematic differences between waveform models dominate the statistical uncertainty of the chirp mass, as the signal accumulates more cycles in the sensitive band of the detectors.

Previous parameter-estimation studies have mainly focused on estimating the typical size of the errors, primarily using the Fisher information matrix [62]. The inverse of the Fisher matrix provides lower bounds for the covariance matrix of maximum likelihood estimators (Cramer-Rao lower bound) [30]. In Ref. [63] a single detector analysis of nonspinning signals suggested that the chirp mass $M$ could be estimated by LIGO with a fraction of a percent accuracy and the mass ratio $\eta$ with an accuracy of a few percent. We find in our analysis (Sec. III) that for the five injections considered this accuracy is reached only for the high-SNR example in Sec. IIIA.2. In general, the full analysis recovers the chirp mass $M$ and the mass ratio $\eta$ at the percent and tens of percent levels, respectively. In addition, Ref. [63] emphasized the dependence of the accuracy on the phase order of the post-Newtonian expansion, which we illustrate in the figures of Sec. III.

Multiple detector studies, based on both Fisher matrix estimates and numerical simulations [64,65] showed the different accuracies in parameter estimation between intrinsic parameters (e.g., masses) and extrinsic parameters (e.g., sky position, the strongly correlated inclination and distance). The estimation of the extrinsic parameters mainly depends on the number and positions of the GW detectors, while the accuracy for intrinsic parameters depends mainly on the SNR, and they are nearly equally well estimated with a single interferometer, for the same SNR. The Fisher matrix calculation has known limitations [66], and tends to underestimate the errors at low SNR [67], which our results confirm.

In all examples, the mass ratio parameter is strongly affected by systematic differences between the waveform models. For those models which included spin, the mass ratio parameter had a larger uncertainty due to correlations between the spins and the mass ratio (see Refs. [55–57]), but these broader posteriors did encompass the true values. This indicates that even for sources which are expected to have an insignificant spin, i.e. BNS systems, a conservative analysis with fully spinning templates is desirable. For systems involving black holes, performing the full spinning parameter estimation is essential, and further improvements in waveform modeling, informed by numerical relativity simulations, are crucial for reducing systematic biases and therefore providing more precise mass and spin estimates. Particularly necessary are quickly computed and accurate waveform models which include arbitrary spins, and inspiral, merger and ring-down phases.

For the extrinsic parameters, of which the most interesting are the source location and distance, the accuracy of inference is limited by the statistical uncertainty coming from the finite SNR of the source. The overall signal amplitude in each detector may be measured with a fractional accuracy $\sim 1/\text{SNR}$, but the amplitude in each detector is a complicated function of the location, distance and polarization angle [see Eq. (4)], which produces the complicated credible interval shapes seen above. Most notably, the precision to which the distance can be measured is limited by the presence of correlation with the inclination angle, which is true for any detector network. On the other hand, the degeneracy in sky position can be greatly reduced by observing the a strong signal in multiple detectors; in the case of the three-detector network used here the sky position is restricted to two locations, symmetric when reflected across the plane containing the detectors, and with opposite inclinations (see Fig. 5). This degeneracy is broken in the case of the spinning BBH signal (Sec. IIIB.1) where the inclination angle can be determined from the precession of the orbital plane, eliminating one of the two sky locations (see Fig. 7). As the size of the two sky locations is primarily determined by the location of the detectors and the SNR, these results are expected to be qualitatively similar for the advanced detector network. When the signal amplitude is relatively low in one detector...
(see Fig. 10), more of the ring-like structure is observable, and the location cannot be well determined.

The evidence values listed in Table III can in principle allow us to discriminate between different waveform models, e.g. whether the data support the presence of spins in the compact objects that generate the detected signal. However, in the analysis reported here the noise model that was used (Gaussian and stationary with a PSD estimated using nearby data) is likely not realistic enough to draw sufficiently robust conclusions, with the exception of the most clear cases (for example the dominance of the spinning model for the injection of Sec. III C). Work is in progress to include in the Bayesian analysis our uncertainty about the PSD and a more flexible noise model [60,68,69].

V. CONCLUSIONS

In this paper, we have applied a suite of Bayesian parameter-estimation tools to hardware and software injections from the last initial LIGO and Virgo science runs. The primary challenge of parameter estimation for CBC signals lies in efficiently locating and sampling the modes of the posterior PDF in a multidimensional parameter space. We have shown that we are able to explore and compare these complicated, correlated and degenerate PDFs for a variety of waveform models which, taken together, include the full range of physical effects expected to be of importance in a real analysis: inclusion of spins, precession of the orbital plane, and the merger–ring down signal. We have verified the consistency of our results by cross comparison among different sampling algorithms.

Although there has been rapid progress in the field of GW parameter estimation, a number of key questions remain. Some of these are already being addressed, or will be addressed over the next few years before advanced detectors come online. Some of the most pressing questions are:

(i) How precisely will we be able to measure spin magnitudes and spin tilt angles in precessing NSBH and BBH systems?

(ii) How important are systematic waveform biases due to imperfect waveform knowledge for various system classes? What are the accuracy requirements on waveform families for parameter estimation?

(iii) What is the best way to handle the long-duration signals in advanced detectors? Their extended bandwidth can cause current implementations to increase in runtime by factors of 25 or more.

(iv) How accurately can we measure finite-size and tidal-dissipation effects for systems involving neutron stars?

(v) How is parameter estimation affected when the background noise is not stationary and Gaussian?

We anticipate that the analysis methods used in this paper will be directly applicable to the first detections from the advanced detector era, allowing us to carry out astrophysical inference and to measure masses, spins, and sky locations of coalescing compact binaries. This, in turn, will improve searches for electromagnetic counterparts, and ultimately allow us to solve a key inverse problem of GW astrophysics: to reconstruct binary evolution and dynamical binary formation from the observed distributions of masses and spins of merging compact binaries.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the United States National Science Foundation for the construction and operation of the LIGO Laboratory, the Science and Technology Facilities Council of the United Kingdom, the Max-Planck-Society, and the State of Niedersachsen/Germany for support of the construction and operation of the GEO600 detector, and the Italian Istituto Nazionale di Fisica Nucleare and the French Centre National de la Recherche Scientifique for the construction and operation of the Virgo detector. The authors also gratefully acknowledge the support of the research by these agencies and by the Australian Research Council, the International Science Linkages program of the Commonwealth of Australia, the Council of Scientific and Industrial Research of India, the Istituto Nazionale di Fisica Nucleare of Italy, the Spanish Ministerio de Economía y Competitividad, the Conselleria d’Economia Hisenda i Innovació of the Govern de les Illes Balears, the Foundation for Fundamental Research on Matter supported by the Netherlands Organisation for Scientific Research, the Polish Ministry of Science and Higher Education, the FOCUS Programme of Foundation for Polish Science, the Royal Society, the Scottish Funding Council, the Scottish Universities Physics Alliance, the National Aeronautics and Space Administration, 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 Carnegie Trust, the Leverhulme Trust, the David and Lucile Packard Foundation, the Research Corporation, and the Alfred P. Sloan Foundation.

APPENDIX: CREDIBLE INTERVALS

In Tables IV, V, VI, VII, and VIII we report the median value and 90% credible intervals for the one-dimensional posteriors in $M$, $\eta$, $m_1$, $m_2$, and $d_f$ obtained in the context of the different waveform models for the injections listed in Table II. Our reported credible intervals are symmetric, excluding an equal posterior mass (in this case 0.05) below and above the interval: a symmetric $(100 - \alpha)$% credible interval $[\theta_{\min}, \theta_{\max}]$ satisfies

$$ \int_{\theta_{\min}}^{\theta_{\max}} d\theta \, p(\theta | d, H) = \int_{\theta_{\min}}^{\theta_{\max}} d\theta \, p(\theta | d, H) = \frac{1}{2} \alpha. \quad (A1) $$

062001-21
### TABLE IV. Median value and 90% credible intervals on chirp mass ($\mathcal{M}$), symmetric mass ratio ($\eta$), masses ($m_1$ and $m_2$) and distance ($d_L$) obtained in the context of various waveform models (see Table I) for the injection discussed in Sec. III A 1.

<table>
<thead>
<tr>
<th>$\mathcal{M}$</th>
<th>$\eta$</th>
<th>$m_1$</th>
<th>$m_2$</th>
<th>$d_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST</td>
<td>3.85^{+0.89}_{-0.77}</td>
<td>0.236^{+0.025}_{-0.019}</td>
<td>5.67^{+0.56}_{-0.61}</td>
<td>3.49^{+0.27}_{-0.21}</td>
</tr>
<tr>
<td>ST_SA</td>
<td>3.88^{+0.98}_{-0.82}</td>
<td>0.224^{+0.025}_{-0.019}</td>
<td>6.27^{+0.21}_{-0.18}</td>
<td>3.23^{+0.31}_{-0.28}</td>
</tr>
<tr>
<td>ST_NS</td>
<td>3.85^{+0.86}_{-0.83}</td>
<td>0.244^{+0.025}_{-0.018}</td>
<td>5.13^{+0.34}_{-0.29}</td>
<td>3.83^{+0.37}_{-0.33}</td>
</tr>
<tr>
<td>TF2</td>
<td>3.85^{+0.86}_{-0.83}</td>
<td>0.244^{+0.025}_{-0.018}</td>
<td>5.13^{+0.34}_{-0.29}</td>
<td>3.83^{+0.37}_{-0.33}</td>
</tr>
<tr>
<td>TF2_2</td>
<td>3.84^{+0.86}_{-0.83}</td>
<td>0.161^{+0.018}_{-0.015}</td>
<td>9.15^{+0.10}_{-0.08}</td>
<td>2.32^{+0.14}_{-0.12}</td>
</tr>
<tr>
<td>IMRBP</td>
<td>3.88^{+0.93}_{-0.86}</td>
<td>0.190^{+0.024}_{-0.001}</td>
<td>7.47^{+1.4}_{-1.3}</td>
<td>2.78^{+0.19}_{-0.18}</td>
</tr>
</tbody>
</table>

### TABLE V. Median value and 90% credible intervals on chirp mass ($\mathcal{M}$), symmetric mass ratio ($\eta$), masses ($m_1$ and $m_2$) and distance ($d_L$) obtained in the context of various waveform models (see Table I) for the injection discussed in Sec. III A 2.

<table>
<thead>
<tr>
<th>$\mathcal{M}$</th>
<th>$\eta$</th>
<th>$m_1$</th>
<th>$m_2$</th>
<th>$d_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST</td>
<td>1.503^{+0.503}_{-0.504}</td>
<td>0.248^{+0.025}_{-0.024}</td>
<td>1.88^{+0.17}_{-0.16}</td>
<td>1.59^{+0.69}_{-0.65}</td>
</tr>
<tr>
<td>ST_SA</td>
<td>1.503^{+0.504}_{-0.503}</td>
<td>0.248^{+0.025}_{-0.024}</td>
<td>1.88^{+0.17}_{-0.16}</td>
<td>1.59^{+0.69}_{-0.65}</td>
</tr>
<tr>
<td>ST_NS</td>
<td>1.501^{+0.502}_{-0.501}</td>
<td>0.248^{+0.025}_{-0.024}</td>
<td>1.88^{+0.17}_{-0.16}</td>
<td>1.59^{+0.69}_{-0.65}</td>
</tr>
<tr>
<td>TF2</td>
<td>1.501^{+0.502}_{-0.501}</td>
<td>0.248^{+0.025}_{-0.024}</td>
<td>1.88^{+0.17}_{-0.16}</td>
<td>1.59^{+0.69}_{-0.65}</td>
</tr>
<tr>
<td>TF2_2</td>
<td>1.500^{+0.500}_{-0.498}</td>
<td>0.185^{+0.018}_{-0.017}</td>
<td>3.12^{+0.06}_{-0.06}</td>
<td>1.01^{+0.03}_{-0.03}</td>
</tr>
<tr>
<td>IMRBP</td>
<td>1.509^{+0.512}_{-0.506}</td>
<td>0.218^{+0.022}_{-0.021}</td>
<td>2.56^{+0.11}_{-0.11}</td>
<td>1.21^{+0.93}_{-0.93}</td>
</tr>
</tbody>
</table>

### TABLE VI. Median value and 90% credible intervals on chirp mass ($\mathcal{M}$), symmetric mass ratio ($\eta$), masses ($m_1$ and $m_2$) and distance ($d_L$) obtained in the context of various waveform models (see Table I) for the injection discussed in Sec. III B 1.

<table>
<thead>
<tr>
<th>$\mathcal{M}$</th>
<th>$\eta$</th>
<th>$m_1$</th>
<th>$m_2$</th>
<th>$d_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST</td>
<td>4.76^{+0.7}_{-0.7}</td>
<td>0.244^{+0.025}_{-0.024}</td>
<td>6.41^{+0.53}_{-0.53}</td>
<td>4.68^{+0.38}_{-0.38}</td>
</tr>
<tr>
<td>ST_SA</td>
<td>4.76^{+0.7}_{-0.7}</td>
<td>0.244^{+0.025}_{-0.024}</td>
<td>6.41^{+0.53}_{-0.53}</td>
<td>4.68^{+0.38}_{-0.38}</td>
</tr>
<tr>
<td>ST_NS</td>
<td>4.76^{+0.7}_{-0.7}</td>
<td>0.244^{+0.025}_{-0.024}</td>
<td>6.41^{+0.53}_{-0.53}</td>
<td>4.68^{+0.38}_{-0.38}</td>
</tr>
<tr>
<td>TF2</td>
<td>4.75^{+0.7}_{-0.7}</td>
<td>0.246^{+0.025}_{-0.024}</td>
<td>6.27^{+0.31}_{-0.31}</td>
<td>4.78^{+0.41}_{-0.41}</td>
</tr>
<tr>
<td>TF2_2</td>
<td>4.76^{+0.7}_{-0.7}</td>
<td>0.172^{+0.015}_{-0.015}</td>
<td>10.7^{+0.13}_{-0.13}</td>
<td>3.02^{+0.59}_{-0.59}</td>
</tr>
<tr>
<td>IMRBP</td>
<td>4.74^{+0.7}_{-0.7}</td>
<td>0.241^{+0.015}_{-0.016}</td>
<td>6.62^{+0.54}_{-0.54}</td>
<td>4.52^{+0.37}_{-0.37}</td>
</tr>
</tbody>
</table>

### TABLE VII. Median value and 90% credible intervals on chirp mass ($\mathcal{M}$), symmetric mass ratio ($\eta$), masses ($m_1$ and $m_2$) and distance ($d_L$) obtained in the context of various waveform models (see Table I) for the injection discussed in Sec. III B 2.

<table>
<thead>
<tr>
<th>$\mathcal{M}$</th>
<th>$\eta$</th>
<th>$m_1$</th>
<th>$m_2$</th>
<th>$d_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST</td>
<td>3.16^{+0.31}_{-0.29}</td>
<td>0.10^{+0.017}_{-0.0092}</td>
<td>11.7^{+0.5}_{-0.13}</td>
<td>1.44^{+0.5}_{-0.66}</td>
</tr>
<tr>
<td>ST_SA</td>
<td>3.24^{+0.35}_{-0.34}</td>
<td>0.18^{+0.027}_{-0.011}</td>
<td>6.61^{+0.24}_{-0.11}</td>
<td>2.21^{+0.5}_{-0.66}</td>
</tr>
<tr>
<td>ST_NS</td>
<td>3.22^{+0.35}_{-0.34}</td>
<td>0.13^{+0.019}_{-0.013}</td>
<td>9.1^{+0.4}_{-0.17}</td>
<td>1.71^{+0.5}_{-0.66}</td>
</tr>
<tr>
<td>TF2</td>
<td>3.25^{+0.34}_{-0.33}</td>
<td>0.15^{+0.023}_{-0.011}</td>
<td>8.02^{+0.9}_{-0.35}</td>
<td>1.9^{+0.5}_{-0.66}</td>
</tr>
<tr>
<td>TF2_2</td>
<td>3.25^{+0.34}_{-0.33}</td>
<td>0.07^{+0.0163}_{-0.0054}</td>
<td>13.5^{+0.8}_{-0.9}</td>
<td>1.35^{+0.5}_{-0.66}</td>
</tr>
<tr>
<td>IMRBP</td>
<td>3.25^{+0.34}_{-0.33}</td>
<td>0.18^{+0.02}_{-0.012}</td>
<td>6.88^{+0.08}_{-0.08}</td>
<td>2.16^{+0.5}_{-0.66}</td>
</tr>
</tbody>
</table>
TABLE VIII. Median value and 90% credible intervals on chirp mass ($\mathcal{M}$), symmetric mass ratio ($\eta$), masses ($m_1$ and $m_2$) and distance ($d_\star$) obtained in the context of various waveform models (see Table I) for the injection discussed in Sec. III C.

<table>
<thead>
<tr>
<th>$\mathcal{M}$</th>
<th>$\eta$</th>
<th>$m_1$</th>
<th>$m_2$</th>
<th>$d_\star$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF2_25</td>
<td>4.63$^{+6.6}_{-4.6}$</td>
<td>0.249$^{+0.25}_{-0.32}$</td>
<td>5.67$^{+35}_{-35}$</td>
<td>58.3$^{+3}_{-4.45}$</td>
</tr>
<tr>
<td>ST_25</td>
<td>4.94$^{+9.7}_{-4.9}$</td>
<td>0.060$^{+0.070}_{-0.0534}$</td>
<td>25$^{+23}_{-22}$</td>
<td>1.71$^{+84}_{-1.62}$</td>
</tr>
</tbody>
</table>