Supplement to "GW250114: testing Hawking's area law and the Kerr nature of black holes"

(The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration)*

I. OBSERVATION AND INFERENCE DETAILS

Here we describe additional details relevant to the detection and parameter inference of GW250114. See also Ref. [1] for more details on the methods for identifying and characterizing gravitational-wave transients.

The GstLAL [2–13] search pipeline was the first to report that it had identified GW250114 with high confidence 15 seconds after the signal arrived. Within 75 s of the signal crossing the Earth, all operating low-latency search pipelines reported detections, including those targeting unmodelled transient events: MLy [14], SPIIR [15], MBTA [16, 17], PyCBC [18], and cWB [19, 20]. All pipelines assigned a false alarm rate of 1 per 100 years or lower. The modeled pipelines reported network matched-filter SNRs [21] ranging from 77 to 80. The variation is due to the discreteness in the sets of filter waveforms used [1] as well as differences in the estimation of the detector noise power spectral density at that time. This signal has by far the highest SNR yet recorded, with the previous highest being the single-detector signal GW230814_230901 with a matched-filter SNR of 42 [22, 23]. The two-detector sky localization was released publicly at 08:22:35 UTC, 32 s after the signal reached the detectors. Based on the classification reported by the search pipelines [24], GW250114 was determined to have a greater than 99% probability of having a binary black hole origin. The data-quality report noted no issues in LIGO Hanford. Some excess power was reported in LIGO Livingston, although glitch subtraction [25–27] was not needed due to the power being confined to low frequencies.

We infer the source properties following standard procedures [1] using the Asimov workflow package [28]. We analyze 8 s of data in the 20−896 Hz frequency range. The detector noise power spectral density is obtained with BayesWave [25, 29]. We use the inference library bilby [30, 31] and the dynesty nested sampler [32] to sample from the posterior distribution of the source parameters for the main analyses that assume quasicircular orbits. We further use bilby and RIFT [33–35] for analyses that allow for eccentric orbits, but are restricted to aligned spins. We marginalize over the luminosity distance with a prior that corresponds to a uniform merger rate in co-moving volume using the Planck 2015 ΛCDM cosmology model [36]. We further marginalize over uncertainties in the detector calibration; the 1σ bounds on the calibration uncertainty remain within 4% in the amplitude and 2 degrees in phase for both detectors within the analysis band.

We model the signal with four waveform models that include the effects of spin-precession and higher-order radiation modes but are restricted to quasi-circular orbits: NRSur7dq4 [37], PhenomXPHM [38, 39], SEOBNRv5PHM [40–43], and PhenomXO4a [44, 45]. Analysis settings for PhenomXPHM, SEOBNRv5PHM, and PhenomXO4a ensure that modes up to $\ell = 4$ are fully accounted for. Due to the finite length of NRSur7dq4, the $\ell = 3$ and $\ell = 4$ modes enter above 22.5 Hz and 30 Hz respectively. Priors are uniform in detector-frame component masses and spin magnitudes, isotropic in spin orientations, isotropic in the binary's orientation, uniform in merger time and coalescence phase, uniform in comoving volume, and isotropic in sky location. Figure 1 shows the two-dimensional posterior for the component masses and effective inspiral and precessing spin parameters obtained from each model. Even though the posteriors are not identical, they all lead to a similar interpretation of GW250114 as an equal-mass system with small spins. We list the 90% credible intervals for further parameters in Table I.

We additionally consider two models that allow for eccentric orbits but are restricted to aligned spins: SEOBNRv5EHM [46] and TEOBResumS-DALI [47]. We use identical settings to the main runs, and adopt uniform priors for the orbital eccentricity and relativistic (mean) anomaly for SEOBNRv5EHM (TEOBResumS-DALI). Both analyses yield no evidence for eccentricity, with 90% upper limits of $e \leq 0.03$ at an orbit-averaged gravitational-wave frequency [48] of 13.33 Hz.

II. TECHNICAL DETAILS AND FURTHER RESULTS FOR THE POST-MERGER ANALYSIS

IIA. Reference parameters

The ringdown analyses in the main text are set up based on a number of reference parameters derived from a preliminary full-signal NRSur7dq4 analysis, which are also consistent with the final production analysis. The first input needed is some guidance on the inferred merger time. As a proxy for this, we adopt the inferred peak of the gravitational-wave strain, integrated over the celestial sphere around the source; this quantity is independent of the observer's orientation and is commonly adopted as a reference in numerical-relativity studies [e.g., 49, 50]. This is also the quantity adopted by the NRSur7dq4 approximant as the definition of the coalescence time [37]. For a waveform decomposed in terms of spin-weighted spherical harmonics ${}_{-2}Y_{\ell m}(\theta, \phi)$ such that

$$h_+ - i h_\times = \sum_{\ell m} h_{\ell m}(t)\, {}_{-2}Y_{\ell m}(\theta, \phi)\,, \tag{1}$$

TABLE I. Source properties of GW250114 for various parameters and four waveform models for spin-precessing, quasicircular systems. We report the median values together with the 90% symmetric credible intervals at a reference frequency of 20 Hz. For parameters that rail against the minimum (maximum) possible values we display upper (lower) limits at the 90% credible level.

Parameter	NRSur7dq4	SEOBNRv5PHM	PhenomXPHM	PhenomXO4a
Primary mass $m_1/M_\odot$	$33.6^{+1.2}_{-0.8}$	$33.5^{+1.2}_{-0.8}$	$33.7^{+1.2}_{-0.9}$	$33.5^{+1.3}_{-0.9}$
Secondary mass $m_2/M_\odot$	$32.2^{+0.8}_{-1.3}$	$32.2^{+0.9}_{-1.3}$	$32.3^{+0.9}_{-1.4}$	$32.1^{+0.9}_{-1.5}$
Mass ratio $q = m_2/m_1$	≥ 0.91	≥ 0.91	≥ 0.91	≥ 0.91
Total mass $M/M_\odot$	$65.8^{+1.1}_{-1.2}$	$65.7^{+1.1}_{-1.1}$	$66.0^{+1.2}_{-1.1}$	$65.5^{+1.3}_{-1.3}$
Detector-frame total mass $(1+z)\,M/M_\odot$	$71.5^{+0.9}_{-1.0}$	$71.1^{+1.1}_{-1.1}$	$71.6^{+1.0}_{-1.1}$	$70.9^{+1.3}_{-1.2}$
Chirp mass $\mathcal{M}/M_\odot$	$28.6^{+0.5}_{-0.5}$	$28.6^{+0.5}_{-0.5}$	$28.7^{+0.5}_{-0.5}$	$28.5^{+0.6}_{-0.6}$
Detector-frame chirp mass $(1+z)\,\mathcal{M}/M_\odot$	$31.1^{+0.4}_{-0.4}$	$30.9^{+0.5}_{-0.5}$	$31.2^{+0.4}_{-0.5}$	$30.8^{+0.6}_{-0.5}$
Final mass $M_\mathrm{f}/M_\odot$	$62.7^{+1.0}_{-1.1}$	$62.6^{+1.0}_{-1.0}$	$62.9^{+1.1}_{-1.0}$	$62.5^{+1.2}_{-1.1}$
Detector-frame final mass $(1+z)\,M_f/M_\odot$	$68.1^{+0.8}_{-0.9}$	$67.8^{+1.0}_{-1.0}$	$68.2^{+0.9}_{-1.0}$	$67.6^{+1.1}_{-1.1}$
Primary spin magnitude χ_1	≤ 0.24	≤ 0.32	≤ 0.32	≤ 0.30
Secondary spin magnitude χ_2	≤ 0.26	≤ 0.34	≤ 0.36	≤ 0.35
Effective inspiral-spin χ_eff	$-0.03^{+0.03}_{-0.04}$	$-0.05^{+0.04}_{-0.05}$	$-0.03^{+0.04}_{-0.05}$	$-0.07^{+0.06}_{-0.05}$
Effective precessing-spin χ_p	$0.11^{+0.15}_{-0.11}$	$0.15^{+0.19}_{-0.15}$	$0.15^{+0.21}_{-0.15}$	$0.16^{+0.16}_{-0.13}$
Final spin χ_f	$0.68^{+0.01}_{-0.01}$	$0.67^{+0.01}_{-0.01}$	$0.68^{+0.01}_{-0.01}$	$0.67^{+0.02}_{-0.01}$
Luminosity distance $D_\mathrm{L}/\mathrm{Mpc}$	403^{+74}_{-70}	385^{+75}_{-69}	399^{+79}_{-72}	381^{+77}_{-72}
Viewing angle Θ/rad	$0.78^{+0.19}_{-0.23}$	$0.82^{+0.19}_{-0.22}$	$0.78^{+0.20}_{-0.23}$	$0.82^{+0.21}_{-0.23}$
Source redshift z	$0.09^{+0.01}_{-0.01}$	$0.08^{+0.01}_{-0.01}$	$0.09^{+0.02}_{-0.01}$	$0.08^{+0.02}_{-0.01}$
Network ($\ell = 4, \lvert m \rvert = 4$) mode SNR ρ_{44}	$3.6^{+1.4}_{-1.5}$	$3.9^{+1.4}_{-1.5}$	$3.6^{+1.5}_{-1.6}$	$4.0^{+1.5}_{-1.6}$

the peak of the strain over the celestial sphere is given by

$$t_\mathrm{peak} = \mathrm{argmax}_t \left[\sum_{\ell m} |h_{\ell m}(t)|^2 \right], \tag{2}$$

which is manifestly invariant under rotations. This is equivalent to the peak of the strain norm integrated over the celestial sphere because the angular harmonics are orthonormal.

For each sample in the NRSur7dq4 reference posterior, we use Eq. (2) to infer the arrival time of the peak strain at geocenter, or any given detector. This results in a posterior on the merger time at each detector. For the NRSur7dq4 run in the main text, the measured GPS peak times are $t^\mathrm{LHO}_\mathrm{peak} = 1420878141.2190^{+0.0001}_{-0.0001}$ s at LIGO Hanford and $t^\mathrm{LLO}_\mathrm{peak} = 1420878141.2165^{+0.0001}_{-0.0001}$ s at LIGO Livingston. We select the maximum-likelihood sample from the preliminary reference posterior as a representative value for the ringdown analysis. This is the reference time $t_\mathrm{peak} = 1420878141.235932$ s at geocenter, with the corresponding sky location ($\alpha = 2.333$, $\delta = 0.190$) implicitly encoding the individual detector times. We list all reference values in Table II.

To guide the analysis and report results, it is useful to define a reference timescale in units of the final black hole mass. We derive a posterior on the redshifted final black hole mass $(1+z)M_\mathrm{f}$ and spin χ_f by applying the NRSur7dq4Remnant model [37] to NRSur7dq4 posterior samples. As reference to quote timescales in the ringdown analysis, we adopt the same maximum-likelihood sample mentioned above, which is quoted in Table II together with the corresponding timescale t_{M_f}. Relative to the chosen reference times, the peak distributions inferred by the NRSur7dq4 analysis in the main text are $t^\mathrm{LHO}_\mathrm{peak} - t^\mathrm{LHO}_\mathrm{ref} = -0.07^{+0.37}_{-0.37}\, t_{M_\mathrm{f}}$ at LIGO Hanford and $t^\mathrm{LLO}_\mathrm{peak} - t^\mathrm{LLO}_\mathrm{ref} = 0.07^{+0.38}_{-0.38}\, t_{M_\mathrm{f}}$ at LIGO Livingston. The standard deviations of those distributions are $0.22\, t_{M_\mathrm{f}}$ and $0.23\, t_{M_\mathrm{f}}$ respectively.

IIB. Ringdown analyses

IIB1. Parameterization and priors

Each quasinormal mode is described by four parameters besides its frequency and damping rate: an overall amplitude $A_{\ell mn}$, a polarization ellipticity $\epsilon_{\ell mn}$, a polarization angle $\theta_{\ell mn}$, and a fiducial phase $\phi_{\ell mn}$ [51, 52]. These parameters control the amplitude

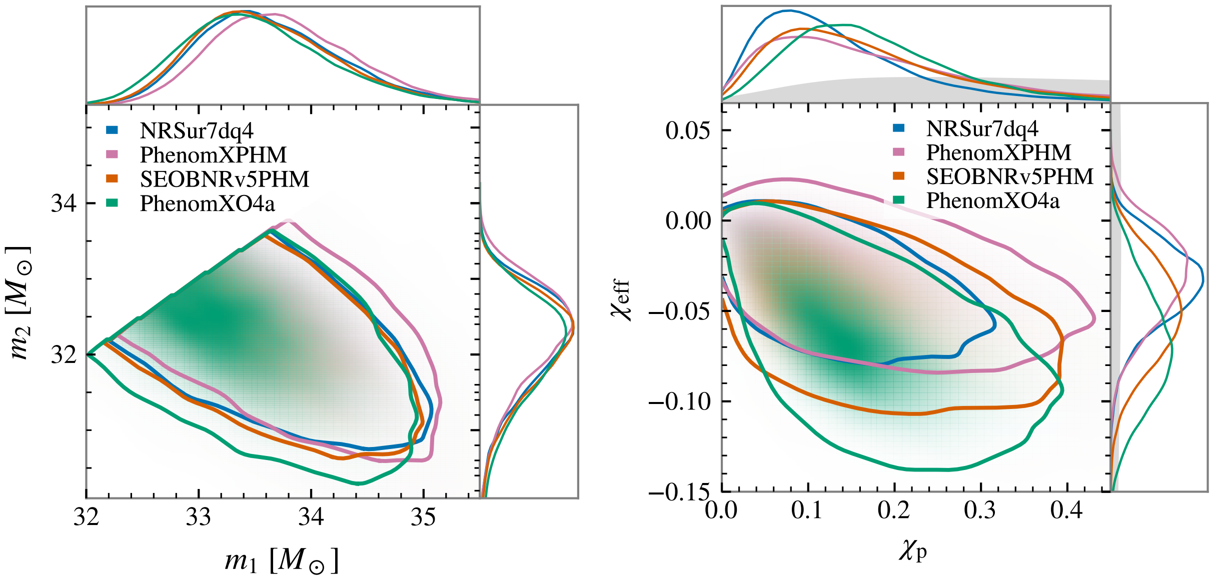

FIG. 1. Marginal posterior distributions for the source-frame component masses (left) and effective inspiral and precessing spin parameters (right) of GW250114 using four waveform models for spin-precessing, quasicircular systems. All models favor an equal-mass system with small spins.

TABLE II. Reference parameters for the ringdown analysis.

Parameter	Value	Description
$t_{\rm peak}$ (geocenter)	1420878141.235932 s	Reference time at geocenter
α	2.333 rad	Right ascension of source
δ	0.190 rad	Declination of source
ψ	1.329 rad	Polarization angle
$t_{\rm peak}^{\rm LHO}$	1420878141.2190118 s	Reference time at LIGO Hanford
$t_{\rm peak}^{\rm LLO}$	1420878141.2165165 s	Reference time at LIGO Livingston
$(1+z)\,M_{\rm f}$	68.409 $M_\odot$	Reference final black hole mass
$t_{M_{\rm f}}$	0.337 ms	Reference $(1+z)M_{\rm f}$ in units of time
$(1+z)\,M$	71.849 $M_\odot$	Reference total mass
t_M	0.354 ms	Reference $(1+z)M$ in units of time

and phase of the two gravitational-wave polarizations (+ and ×) for each mode:

$$h_+ = A\left[\cos\theta\cos(2\pi f t-\phi) - \epsilon\sin\theta\sin(2\pi f t-\phi)\right]\exp(-\gamma t)\,, \tag{3a}$$

$$h_\times = A\left[\sin\theta\cos(2\pi f t-\phi) + \epsilon\cos\theta\sin(2\pi f t-\phi)\right]\exp(-\gamma t)\,, \tag{3b}$$

suppressing mode indices $(\ell, |m|, n)$ for brevity. This is the most generic expression for a quasinormal-mode signal and subsumes both positive and negative frequency contributions, which jointly encode the polarization content of each mode [51, 52]. Assuming prograde modes with nonvanishing m, the sign of the frequency is related to the sign of the azimuthal number by $\mathrm{sgn}(m) = \mathrm{sgn}(f)$, so that the positive and negative frequencies encode the right- and left-handed polarized components of the mode. The same would be true for retrograde modes, except that $\mathrm{sgn}(m) = -\mathrm{sgn}(f)$. This paper only considers prograde modes.

Without information about the expected intrinsic amplitudes of the quasinormal modes, ringdown analyses cannot infer a luminosity distance and the only mass scale to which they are sensitive is the product $(1+z)M_{\rm f}$. In the case of a Kerr fit, all frequencies and damping rates are derived from a given $(1+z)M_{\rm f}$ and $\chi_{\rm f}$, such that $f_{\ell|m|n} = f_{\ell|m|n}[(1+z)M_{\rm f}, \chi_{\rm f}]$ and $\gamma_{\ell|m|n} = \gamma_{\ell|m|n}[(1+z)M_{\rm f}, \chi_{\rm f}]$. For beyond-Kerr fits, we introduce additional parameters $\delta f_{\ell|m|n}$ and $\delta\gamma_{\ell|m|n}$ such that $f_{\ell|m|n} = f_{\ell|m|n}[(1+z)M_{\rm f}, \chi_{\rm f}]\exp(\delta f_{\ell|m|n})$ and $\gamma_{\ell|m|n} = \gamma_{\ell|m|n}[(1+z)M_{\rm f}, \chi_{\rm f}]\exp(\delta\gamma_{\ell|m|n})$. The exponential parameterization avoids a singularity as the deviation parameter approaches −1 [51].

The RINGDOWN code places priors that are flat in $(1+z)M_{\rm f}$, $\chi_{\rm f}$, A, θ, and ϕ as well as, when applicable, δf_{221} and $\delta\gamma_{221}$; the ellipticity prior peaks at $\epsilon = 0$ but has broad support over the entire domain (Fig. 16 in [52]). The amplitude prior is flat in

A over the interval $[0, 5 \times 10^{-20}]$ and is broad enough to always offer full support to the posterior without truncating it. The priors in δf_{221} and $\delta\gamma_{221}$ are flat over the intervals $[-0.8, 0.8]$ and $[-0.5, 0.5]$ respectively. The quasinormal-mode polarization angle θ is fully degenerate with the source polarization angle ψ [51, 52], which defines the orientation of the source's angular momentum relative to the celestial North pole and is used to compute the antenna pattern response functions for each detector [53, 54]. Given this exact degeneracy, the ringdown analysis chooses a fiducial angle $\psi = 1.329$ to compute antenna patterns, based on the same maximum-likelihood reference sample used to derive $t_{\rm peak}$ as explained above, although any arbitrary choice of ψ would be valid. The prior in the ringdown code is independent for each mode in all parameters.

The pyRing code natively parameterizes the modes slightly differently. For a given $(\ell, |m|, n)$ mode, it uses

$$\begin{aligned} h_+ - i h_\times = \; & C_{\ell,+m,n} \, {}_{-2}Y_{\ell+m}(\iota, \varphi = 0) \exp\left[i\left(2\pi f_{\ell|m|n} t + \phi_{\ell,+m,n}\right)\right] \exp\left[-\gamma_{\ell|m|n} t\right] + \\ & C_{\ell,-m,n} \, {}_{-2}Y_{\ell-m}(\iota, \varphi = 0) \exp\left[i\left(-2\pi f_{\ell|m|n} t + \phi_{\ell,-m,n}\right)\right] \exp\left[-\gamma_{\ell|m|n} t\right] \,, \end{aligned} \tag{4}$$

where ι is an inclination parameter that is sampled from a prior uniform in $\cos\iota$, at the same time as the free amplitudes $C_{\ell,+m,n}$ and $C_{\ell,-m,n}$, which can be interpreted as the amplitude of the right and left-handed polarized contributions to the mode respectively; the corresponding phases, $\phi_{\ell,+m,n}$ and $\phi_{\ell,-m,n}$, combine in difference and sum to produce θ and ϕ in Eq. (3) [52], and are fully degenerate with the polarization angle ψ, which pyRing also samples over a flat prior. The pyRing prior is also flat on $(1+z)M_{\rm f}$ and $\chi_{\rm f}$.

The pyRing prior on the mode amplitude A is implicitly defined by the priors on $C_{\ell,\pm m,n}$, which are uniform over the interval $[0, 5 \times 10^{-20}]$, and the prior on ι; because of prior volume effects, it amounts to a density that has no support at the origin and disfavors $A \to 0$ (cf. Fig. 16 in [52]). Additionally, since all modes share $\cos\iota$, the prior correlates the different modes. For a two-mode model with $\ell = |m| = 2$ and $n = 0, 1$, the Jacobian to a flat prior in $\{A, \epsilon\}$ can be computed analytically and is given by

$$J = \frac{(1-\cos^2\iota)^4}{4\left[C_{2,-2,0}\,(1-\cos\iota)^2 + C_{2,+2,0}\,(1+\cos\iota)^2\right]\left[C_{2,-2,1}\,(1-\cos\iota)^2 + C_{2,+2,1}\,(1+\cos\iota)^2\right]} \,. \tag{5}$$

In the main text, we reweight the pyRing posterior to a flat prior in $\{A, \epsilon\}$ by applying this Jacobian and truncating to the appropriate bounds.

The pyRing code implements deviations from the Kerr spectrum by writing

$$\omega_{221} = \omega_{221}[(1+z)M_{\rm f}, \chi_{\rm f}]\,(1 + \delta\omega_{221}) \,, \tag{6a}$$

$$\tau_{221} = \tau_{221}[(1+z)M_{\rm f}, \chi_{\rm f}]\,(1 + \delta\tau_{221}) \,, \tag{6b}$$

where $\omega_{221} \equiv 2\pi f_{221}$ and $\tau_{221} \equiv 1/\gamma_{221}$. For small deviation parameters, this parameterization is equivalent to the ringdown one, except with a singularity at $\delta\omega = -1$ and $\delta\tau = -1$. The above implies the following relationship between the ringdown parameters $(\delta f_{221}, \delta\gamma_{221})$ and the pyRing parameters $(\delta\omega_{221}, \delta\tau_{221})$:

$$\delta f_{221} = \log(1 + \delta\omega_{221}) \;\;, \;\; \delta\gamma_{221} = -\log(1 + \delta\tau_{221}) \,. \tag{7}$$

To go from a uniform prior in $(\delta\omega_{221}, \delta\tau_{221})$ to a uniform prior in $(\delta f_{221}, \delta\gamma_{221})$ we must apply a Jacobian given by

$$J = |1 + \delta\omega_{221}|^{-1}\, |1 + \delta\tau_{221}|^{-1} \,. \tag{8}$$

In the main text, we reweight the pyRing posterior to a flat prior in δf_{221} and $\delta\gamma_{221}$ by applying this Jacobian and truncating to the appropriate bounds.

IIB2. Data conditioning

Both the pyRing and ringdown analyses are based on data sampled at 4096 Hz with a covariance matrix derived from the same estimate of the power spectral density used in the main analysis described in Sec. I of this Supplement. Before obtaining the covariance matrix, the power spectral density is treated to censor frequencies below 20 Hz and above 1830 Hz to match the integration band for the likelihood in the full-signal analysis [55].

Additionally, pyRing applies a Butterworth bandpass filter to the data, suppressing frequencies below 20 Hz and above 2043 Hz; the filtering is applied to 64 s of data around the event. ringdown only applies a high-pass Butterworth filter at 10 Hz to remove zero-frequency offsets; there is no low-pass filtering other than truncation of the frequency series at Nyquist (the digital filter described in Ref. [55]); the conditioning is applied to 634 s of data around the event time. Neither code applies any filtering to the signal templates in the likelihood calculation [55]. The difference in conditioning at high frequencies between ringdown and pyRing is understood to cause a subdominant (but measurable) systematic difference in the two posterior distributions.

Estimates of the SNR accumulated after the signal peak indicate that an integration time of $T = 0.6\,\mathrm{s}$ is sufficient to capture the entirety of the post-merger signal; therefore, this is the analysis duration used in all runs by the ringdown code. For reasons of computational efficiency, the pyRing analysis is run with a shorter integration time of $T = 0.2\,\mathrm{s}$, leading to slightly broader posterior distributions. Also to reduce computational cost, the pyRing analysis was run on a sparser grid of start times.

The main root of the systematic differences between the two codes is understood to be in the selection of the analysis data. For each choice of analysis start time $t_>$, the ringdown code selects the first sample of the data to be analyzed at each detector based on the native sampling rate of 16384 Hz provided by the LIGO detectors; once the sample closest to the requested start time is identified in the 16384 Hz data, the ringdown code downsamples to 4096 Hz while preserving the selected sample. The effective timing precision of ringdown is thus $\delta t \approx 1/(16384\,\mathrm{Hz}) = 0.06\,\mathrm{ms} \approx 0.18\,t_{M_\mathrm{f}}$. The pyRing code, on the other hand, first downsamples the data to 4096 Hz and then selects the first sample of the data to be analyzed at each detector, meaning that the effective timing precision is $\delta t \approx 1/(4096\,\mathrm{Hz}) = 0.24\,\mathrm{ms} \approx 0.72\,t_{M_\mathrm{f}}$. The effect of this coarse graining varies for each requested start time and for each detector: if the start time happens to fall on a sample at a given detector, it is unlikely to also fall on a sample at the other.

The above means that the ringdown and pyRing results cannot be made to match by a uniform relabeling of the start times. For example, for runs requesting $t_> = 10.5\,t_{M_\mathrm{f}}$ in the main text, the ringdown code starts the analysis at $t_\mathrm{LHO} = 1420878141.222534\,\mathrm{s}$ (GPS) and $t_\mathrm{LLO} = 1420878141.220032\,\mathrm{s}$ (GPS) for Hanford and Livingston respectively, which is $-0.046\,t_{M_\mathrm{f}}$ and $-0.067\,t_{M_\mathrm{f}}$ relative to the requested $t_>$ in each detector respectively. The closest available pyRing run is the one that requested $t_> = 10\,t_{M_\mathrm{f}}$; this starts the analysis at $t_\mathrm{LHO} = 1420878141.222412\,\mathrm{s}$ (GPS) and $t_\mathrm{LLO} = 1420878141.219971\,\mathrm{s}$ (GPS) for Hanford and Livingston respectively, which is $-0.41\,t_{M_\mathrm{f}}$ and $-0.25\,t_{M_\mathrm{f}}$ relative to the target time of $t_> = 10.5\,t_{M_\mathrm{f}}$ in each detector respectively. In other words, the start time of the pyRing $10\,t_{M_\mathrm{f}}$ run differs from that in the ringdown $10.5\,t_{M_\mathrm{f}}$ run by $-0.36\,t_{M_\mathrm{f}}$ for Hanford and $-0.18\,t_{M_\mathrm{f}}$ in Livingston. A similar calculation shows that the pyRing $6\,t_{M_\mathrm{f}}$ start time is $0.54\,t_{M_\mathrm{f}}$ after the ringdown $6\,t_{M_\mathrm{f}}$ run in both detectors. The fact that pyRing uses less data than ringdown explains the systematic differences in the posteriors presented in the main text, as was verified by running ringdown with the exact same data as pyRing and reproducing that code's results.

Just as the inspiral time-domain analysis, the two sets of post-merger analyses also ignore the uncertainty over the detector calibration, which is expected to have a negligible impact on ringdown analyses at this SNR [56].

IIB3. Computation of amplitude significance

In the main text, we provide estimates for the significance with which we can establish that the amplitude A of a given quasinormal mode is greater than zero. This entails estimating the posterior probability density at $A = 0$, which represents the boundary of the amplitude parameter space ($A \geq 0$) and therefore will never be directly represented in the set of posterior samples. Our significance estimates are based on the smallest probability p such that the highest posterior density (HPD) interval enclosing probability mass p includes the origin. HPD intervals can produce counter-intuitive significance results when the posterior has a sharp truncation at large amplitude, either because the prior cuts off the posterior or due to the structure of the likelihood function, but our amplitude posteriors are not of this shape.

We choose to estimate the significance of the amplitude $A > 0$ by direct integration over an HPD interval of a KDE-based representation of the posterior density, $p(A)$. This method has an advantage over sample-based methods in that it can estimate arbitrarily small values of $p(0)$ (i.e. arbitrarily high significance), while direct sample-based methods bottom out at $p(0) \sim 1/N_\mathrm{samples}$. Compared to other, simpler estimates of the significance such as calculating the z-score $z = \mu_A/\sigma_A$ where μ_A and σ_A are the mean and standard deviation of the posterior samples, this method has the advantage that it can account for non-Gaussian shapes of the posterior density. We first form a standard KDE estimate of the posterior density from the samples of the amplitude A, using automatic bandwidth estimation from the `scipy.stats.gaussian_kde` function. This density estimate, $k_\mathrm{raw}(A)$ is normalized such that it integrates to unity over $-\infty < A < \infty$. To account for the boundary at $A = 0$, we reflect the density estimate about the origin, defining

$$k(A) = k_\mathrm{raw}(A) + k_\mathrm{raw}(-A); \tag{9}$$

this ensures that $k(A)$ integrates to unity over $0 \leq A < \infty$, and so is a suitable density estimate over our domain.

We then evaluate $k(0)$, the (estimated) posterior density at $A = 0$. We evaluate the significance of the amplitude $A > 0$ by integrating the posterior density estimate, k, over A values such that $k(A) > k(0)$, i.e., computing the smallest p such that the highest-posterior-density interval containing probability mass p includes $A = 0$. Thus,

$$p = \int_{\{A|k(A)>k(0)\}} \mathrm{d}A\, k(A). \tag{10}$$

By construction, $0 \leq p \leq 1$.

Once we have an estimate for p, we communicate this as a number of σ by expressing this value in terms of tail probabilities for the Gaussian distribution, i.e., we establish that $A > 0$ at $x\,\sigma$ significance, where x is defined by

$$\int_{-x}^{x} \mathrm{d}x' \, \phi\,(x') = p\,, \tag{11}$$

where ϕ is the probability density function for the standard normal distribution. Equivalently (but more stably numerically), we can define x by

$$\int_{-\infty}^{-x} \mathrm{d}x' \, \phi\,(x') = \frac{1-p}{2}\,. \tag{12}$$

This is the quantity that we estimate and report in the main text. The quantity $1 - p$ can be computed by

$$1 - p = \int_{\{A|k(A)\leq k(0)\}} \mathrm{d}A\, k(A), \tag{13}$$

which is more numerically stable to evaluate when $p \simeq 1$.

IIC. Further parameter posteriors

Figure 5 in the main text shows posteriors for the amplitude of the fundamental and overtone modes as a function of the analysis start time. Amplitudes are referenced to the start time of each analysis and are presented for both the single-mode and two-mode models. For completeness, Fig. 2 shows the posteriors for the frequency and damping rate of each mode in similar style. In the bottom panel, we further show the network matched-filtered SNR as a function of time for both models. For the two-mode analysis, we present the SNR of the full model, as the mode nonorthogonality makes defining the SNR of each mode ambiguous. The inferred frequencies and damping rates are consistent across all times of applicability of their respective model, i.e., after $6\,t_{M_f}$ for the two-mode model. Uncertainties increase and the SNR decreases as the analysis start time is moved later, as expected. Moreover, the SNR recovered by the two-mode and the one-mode models in their overlapping time region are highly consistent.

III. TECHNICAL DETAILS AND FURTHER TESTS FOR THE PRE-MERGER ANALYSIS

To enable the sharp truncation of the data and gravitational-wave model at a time before the binary merger [57–59], we carry out the analysis in the time-domain using TDinf [60, 61], guided by the reference parameters from Table II. The TDinf inference package samples the 15-dimensional parameter space of a quasicircular black-hole binary using the emcee sampler [62]. We adopt the same settings as the full-signal frequency-domain analyses with bilby (in terms of the data, trigger time, power spectral density, and bandwidth), other than the amount of data considered, certain priors, and the treatment of calibration. Time-domain analysis of the full signal is based on 1.4 s of data, which are appropriately truncated. The priors are the same as the bilby analysis other than the masses (uniform in total mass and mass ratio), distance (uniform in luminosity distance), and time (Gaussian centered around the geocenter trigger time with a standard deviation of 0.01 s). The distance prior has a minimal effect on the inferred area, which is based on redshifted masses that are minimally correlated with the distance. The composite area law prior is broad and relatively flat, as seen in Fig. 5 in the main text. We use the autocorrelation length (ACL) from the full ensemble of walkers to determine the burn-in period and thinning of the chains. Depending on the truncation time, the lengths of these chains range between 100,000–500,000 steps, each with 512 walkers. We use a burn-in of at least five times the maximum ACL across sampled parameters, corresponding to > 40% the chain length, and thin by half of the minimum ACL. For some of the truncation times, a handful of chains are not converged to the bulk of the posterior at the end of the burn-in, instead representing a secondary mode in likelihood and sky location. For these chains, we manually extend the burn-in period until they converge to the bulk of the posterior. We have verified that the posterior is not restricted by any of the prior edges. We have also verified that the time-domain and bilby analyses of the full signal yield statistically identical posteriors when reweighted to the same prior for both NRSur7dq4 and PhenomXPHM.

An additional difference is that the time-domain analysis neglects the uncertainty over the detector calibration, which bilby marginalizes over. We have verified that this has a minimal impact by repeating the bilby analysis while neglecting calibration uncertainty. The posteriors for the detector-frame total mass with and without marginalizing over the calibration uncertainty differ by a Jensen–Shannon divergence [63] of 0.002 nat. For reference, changing the waveform model from NRSur7dq4 to PhenomXPHM yields a Jensen–Shannon divergence of 0.018 nat, to SEOBNRv5PHM of 0.041 nat, and to PhenomXO4a of

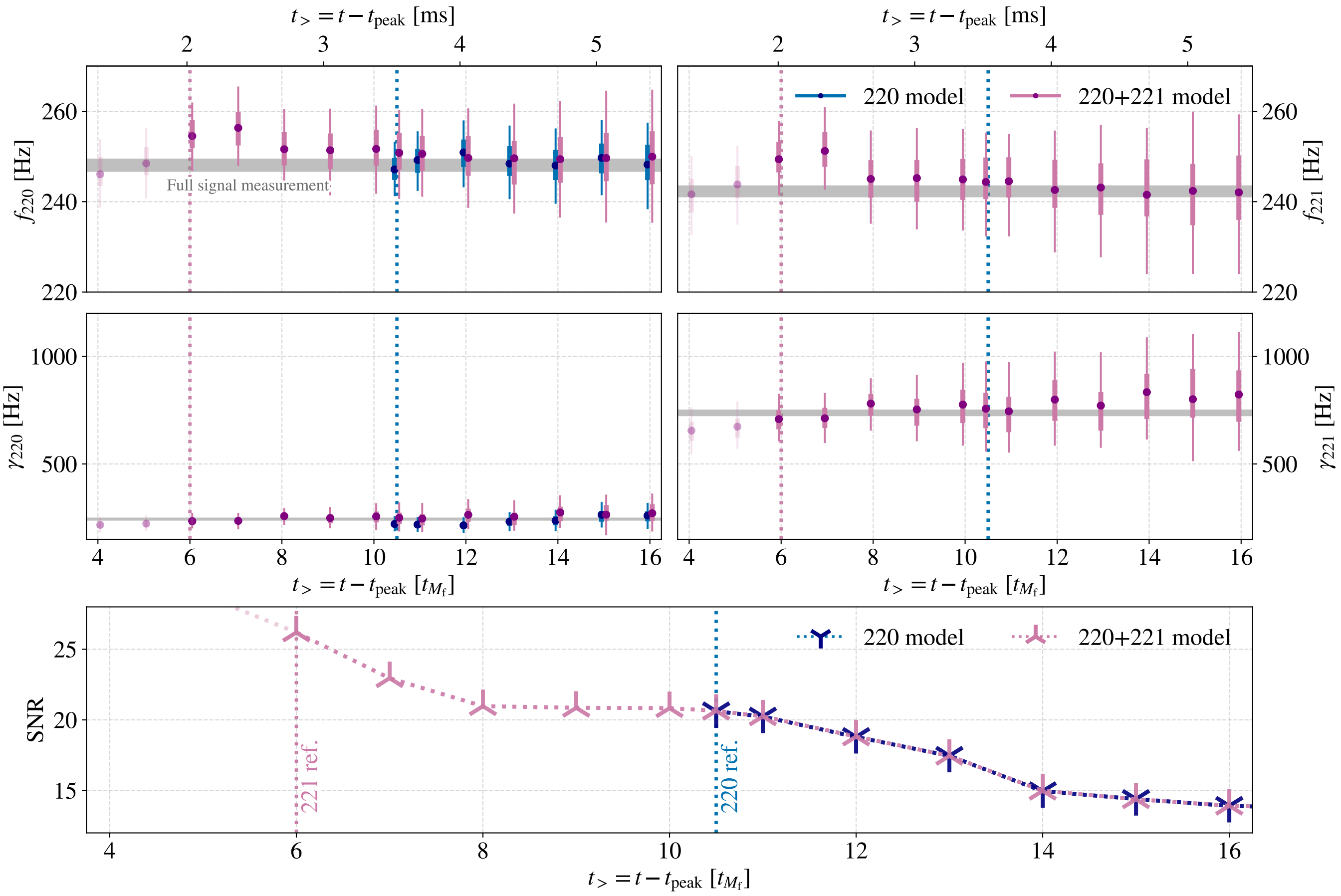

FIG. 2. Similar to Fig. 3 in the main text, but showing the frequency (left) and damping rate (right) of the fundamental (top) and overtone (middle) modes as a function of the analysis start time. The bottom panel shows the network matched-filtered SNR as a function of time for both the single-mode and two-mode models, which are highly consistent. In the two-mode model, both frequencies and damping rates are inferred jointly from a common mass and spin assuming a Kerr spectrum. We show the 90% credible intervals for the frequencies and damping times inferred from the full signal analysis as grey horizontal shaded areas. As in Fig. 3 in the main text, we show results with RINGDOWN as there is not exact timing correspondence between the RINGDOWN and PYRING posteriors; we nonetheless obtain similar results with PYRING.

0.097 nat. The impact of calibration uncertainty is therefore subdominant compared to other potential sources of systematic errors.

The main text presents results for a number of truncation times in Fig. 5. We use the gravitational-wave luminosity inferred from the full-signal analysis (a measure of how dynamical and relativistic the system is [64]) to determine that most of the emission occurs during the late coalescence stages. While the averaged gravitational-wave strain, defined as $\left(\sum_{\ell,m} |h_{\ell m}(t)|^2\right)^{1/2}$, peaks at $t_{\rm peak}$ by definition, the flux, proportional to $\left(\sum_{\ell,m} |\dot{h}_{\ell m}(t)|^2\right)^{1/2}$ (where a dot denotes differentiation with respect to time), peaks at $6\, t_M$ after $t_{\rm peak}$. Specifically, the flux is 10% of its maximum $-36\, t_M$ before $t_{\rm peak}$ and 1% of its maximum $-232\, t_M$ before $t_{\rm peak}$. The former is also comparable to common estimates for the transition to merger [65–67]. Based on these time estimates, Fig. 5 in the main text presents results for an array of truncation times after $t_< = -250\, t_M$ and more detailed results for $t_< = -40\, t_M$.

IV. TECHNICAL DETAILS AND FURTHER RESULTS FOR THE AREA LAW ANALYSIS

The area law test relies on the independent analyses of the pre-merger data described in Sec. III and the post-merger data described in Sec. II of this Supplement. Both tests are carried out in the time-domain and avoid quantitative reference to the full-signal results, which are based on waveform models within general relativity that obey the area law by construction. Since the pre- and post-merger analyses are also independent of each other, there is no information shared between the two sets of results and no assumption about the initial and final black holes sharing the same location in the sky. The only slight exception to this is that, by comparing initial and final quantities in the detector frame, we implicitly assume that the pre- and post-merger signals were redshifted by the same factor. The recoil kick of the remnant can break this assumption by inducing a redshift; however, this is at most 10% and only in the most extreme cases with large spins inconsistent with GW250114 [68]. An analysis in which the extrinsic parameters are sampled jointly for both the pre- and post-merger data, such as [59], could gain by imposing the same location in the sky for the initial and final black holes.

Our test yields independent measurements of the initial and final areas. If the posterior distributions on the final and initial area are denoted as $p_f(\mathcal{A}_\mathrm{f})$ and $p_i(\mathcal{A}_\mathrm{i})$ respectively, the significance σ of a non-detection of a violation of Hawking's area law is given by the separation between $p_f(\mathcal{A}_\mathrm{f})$ and $p_i(\mathcal{A}_\mathrm{i})$. Following the cosmology literature estimating the significance of the Hubble tension [69, 70], we estimate the significance as $X\sigma$ with

$$X = \frac{\mu_\mathrm{f} - \mu_\mathrm{i}}{\sqrt{\sigma_\mathrm{f}^2 + \sigma_\mathrm{i}^2}}\,, \tag{14}$$

where $\mu_\mathrm{i}/\mu_\mathrm{f}$ and $\sigma_\mathrm{i}/\sigma_\mathrm{f}$ are the means and standard deviations respectively of the initial and final area distributions. Since this estimate relies only on the first two cumulants, it is less sensitive to sampling error at the tails of the distribution. Empirically, sampling the posterior tails beyond the $\sim 4\sigma$ level is unreliable and highly sensitive to sampler settings and minor analysis choices, so we avoid using tail samples.

In the main text, Fig. 5 shows results where the final area has been measured via a single $(\ell = 2, |m| = 2, n = 0)$ mode starting at $10.5\, t_{M_\mathrm{f}}$, which is the earliest time that the overtone significance falls below 1σ. Identification of this time is based on the post-merger data alone, without requesting consistency with the full-signal analysis. Here we show similar results obtained for different pre- and post-merger start times, for both the single-mode and two-mode ringdown models. Figure 3 displays the inferred initial and final areas as a function of inspiral end time $t_<$ and ringdown start time $t_>$, respectively. Similarly to Fig. 5, we show the 90% credible intervals for the initial and final areas as inferred from the full-signal analysis for reference. These are consistent with both the initial and final areas within statistical uncertainties, although this comparison does not guide any of the analysis choices in testing the area law. The significance, as defined in Eq. (14), can be assessed via the separation of the distributions relative to their widths.

As expected, the uncertainty in the remnant area grows as the start time is pushed to later parts of the data where the signal is weaker; for the initial area, the opposite is true and the uncertainty grows as the analysis end time is pushed earlier. The single-mode ringdown result used in the main text corresponds to a start time of $t_> = 10.5\, t_{M_\mathrm{f}}$ and is highlighted by a vertical blue band. Including the overtone in the model at this and later times necessarily broadens the uncertainty in the final area. This is both because the overtone is not required to explain the data and because the overtone and the fundamental are not orthogonal: introducing additional, unconstrained degrees of freedom broadens the posterior distribution.

The reference start time showcased in the main text was chosen independently of the full-signal analysis purely as the earliest time at which the data are consistent with a single mode, according to our 1σ criterion. Had the SNR of this event been higher, we would likely not have found the data to be sufficiently well-described by a single mode at $t_> = 10.5\, t_{M_\mathrm{f}}$. This would have led us to use a later start time for the reference ringdown analysis with a single mode. Carrying out the area law test with the two-mode analysis instead selects $6\, t_{M_\mathrm{f}}$ as the earliest time at which this model explains the observed data. The area law is again confidently satisfied at 3.6σ significance.

---

[*] lvc.publications@ligo.org

[1] A. G. Abac *et al.* (LIGO Scientific, Virgo, KAGRA), GWTC-4.0: Methods for Identifying and Characterizing Gravitational-wave Transients (2025), arXiv:2508.18081 [gr-qc].

[2] K. Cannon, C. Hanna, and D. Keppel, Interpolating compact binary waveforms using the singular value decomposition, Phys. Rev. D **85**, 081504 (2012), arXiv:1108.5618 [gr-qc].

[3] C. Messick *et al.*, Analysis Framework for the Prompt Discovery of Compact Binary Mergers in Gravitational-wave Data, Phys. Rev. D **95**, 042001 (2017), arXiv:1604.04324 [astro-ph.IM].

[4] S. Sachdev *et al.*, The GstLAL Search Analysis Methods for Compact Binary Mergers in Advanced LIGO's Second and Advanced Virgo's First Observing Runs (2019), arXiv:1901.08580 [gr-qc].

[5] C. Hanna *et al.*, Fast evaluation of multidetector consistency for real-time gravitational wave searches, Phys. Rev. D **101**, 022003 (2020), arXiv:1901.02227 [gr-qc].

[6] K. Cannon *et al.*, GstLAL: A software framework for gravitational wave discovery, SoftwareX **14**, 100680 (2021), arXiv:2010.05082 [astro-ph.IM].

[7] L. Tsukada *et al.*, Improved ranking statistics of the GstLAL inspiral search for compact binary coalescences, Phys. Rev. D **108**, 043004 (2023), arXiv:2305.06286 [astro-ph.IM].

[8] B. Ewing *et al.*, Performance of the low-latency GstLAL inspiral search towards LIGO, Virgo, and KAGRA's fourth observing run, Phys. Rev. D **109**, 042008 (2024), arXiv:2305.05625 [gr-qc].

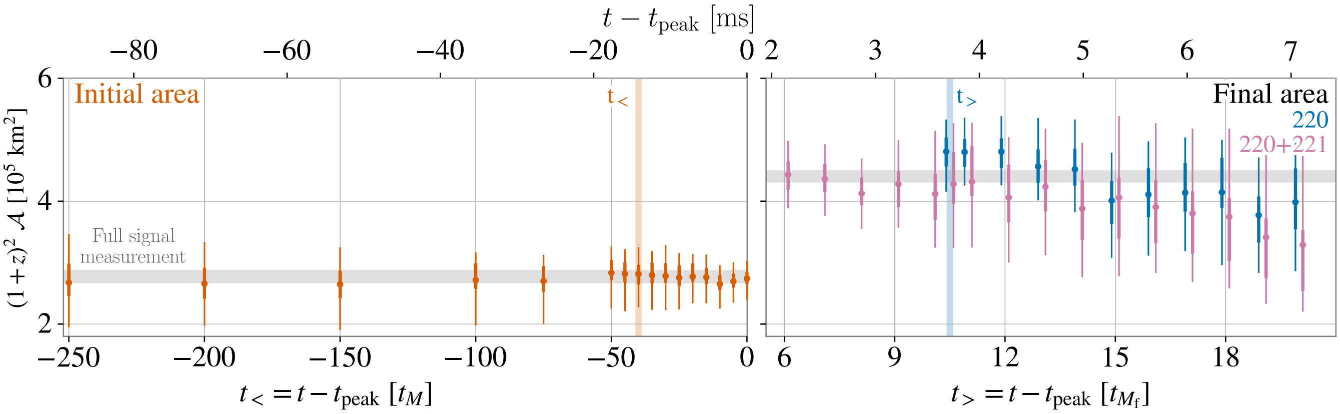

FIG. 3. Measurements for the initial and final black hole areas as a function of fit time. (Left panel) Initial black hole area inferred by analyzing the pre-merger signal with TDinf, truncated at the time indicated on the x-axis. (Right panel) Final black hole areas inferred by analyzing the post-merger signal with ringdown, with the analysis start time indicated on the x-axis. We show results for the two-mode analysis from $6\,t_{M_f}$ after the peak strain and the single-mode only from $10.5\,t_{M_f}$ after the peak strain. We highlight with vertical bands the reference areas quoted in this paper, at $-40\,t_M$ before the peak strain for the inspiral analysis and $10.5\,t_{M_f}$ after the peak strain for the single-mode ringdown analysis. Horizontal grey bands indicate the 90% credible intervals for the initial and final areas as inferred from the full inspiral-merger-ringdown analysis.

[9] S. Sakon *et al.*, Template bank for compact binary mergers in the fourth observing run of Advanced LIGO, Advanced Virgo, and KAGRA, Phys. Rev. D **109**, 044066 (2024), arXiv:2211.16674 [gr-qc].

[10] P. Joshi *et al.*, New Methods for Offline GstLAL Analyses (2025), arXiv:2506.06497 [gr-qc].

[11] P. Joshi *et al.*, How Many Times Should We Matched Filter Gravitational Wave Data? A Comparison of GstLAL's Online and Offline Performance (2025), arXiv:2505.23959 [gr-qc].

[12] P. Joshi, L. Tsukada, and C. Hanna, Method for removing signal contamination during significance estimation of a GstLAL analysis, Phys. Rev. D **108**, 084032 (2023), arXiv:2305.18233 [gr-qc].

[13] A. Ray *et al.*, When to Point Your Telescopes: Gravitational Wave Trigger Classification for Real-Time Multi-Messenger Followup Observations (2023), arXiv:2306.07190 [gr-qc].

[14] V. Skliris, M. R. K. Norman, and P. J. Sutton, Toward real-time detection of unmodeled gravitational wave transients using convolutional neural networks, Phys. Rev. D **110**, 104034 (2024), arXiv:2009.14611 [astro-ph.IM].

[15] Q. Chu *et al.*, SPIIR online coherent pipeline to search for gravitational waves from compact binary coalescences, Phys. Rev. D **105**, 024023 (2022), arXiv:2011.06787 [gr-qc].

[16] F. Aubin *et al.*, The MBTA pipeline for detecting compact binary coalescences in the third LIGO–Virgo observing run, Class. Quant. Grav. **38**, 095004 (2021), arXiv:2012.11512 [gr-qc].

[17] C. Alléné *et al.*, The MBTA pipeline for detecting compact binary coalescences in the fourth LIGO-Virgo-KAGRA observing run, Class. Quant. Grav. **42**, 105009 (2025), arXiv:2501.04598 [gr-qc].

[18] T. Dal Canton, A. H. Nitz, B. Gadre, G. S. Cabourn Davies, V. Villa-Ortega, T. Dent, I. Harry, and L. Xiao, Real-time Search for Compact Binary Mergers in Advanced LIGO and Virgo's Third Observing Run Using PyCBC Live, Astrophys. J. **923**, 254 (2021), arXiv:2008.07494 [astro-ph.HE].

[19] S. Klimenko *et al.*, Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors, Phys. Rev. D **93**, 042004 (2016), arXiv:1511.05999 [gr-qc].

[20] T. Mishra, S. Bhaumik, V. Gayathri, M. J. Szczepańczyk, I. Bartos, and S. Klimenko, Gravitational waves detected by a burst search in LIGO/Virgo's third observing run, Phys. Rev. D **111**, 023054 (2025), arXiv:2410.15191 [astro-ph.HE].

[21] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton, FINDCHIRP: An Algorithm for detection of gravitational waves from inspiraling compact binaries, Phys. Rev. D **85**, 122006 (2012), arXiv:gr-qc/0509116.

[22] A. G. Abac *et al.* (LIGO Scientific, Virgo, KAGRA), GW230814: investigation of a loud gravitational-wave signal observed with a single detector, LIGO-P230814 (2025).

[23] A. G. Abac *et al.* (LIGO Scientific, VIRGO, KAGRA), GWTC-4.0: Updating the Gravitational-Wave Transient Catalog with Observations from the First Part of the Fourth LIGO-Virgo-KAGRA Observing Run (2025), arXiv:2508.18082 [gr-qc].

[24] LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration, LIGO/Virgo/KAGRA S250114ax: Identification of a GW compact binary merger candidate, GCN Circular 38932 (2025).

[25] N. J. Cornish, T. B. Littenberg, B. Bécsy, K. Chatziioannou, J. A. Clark, S. Ghonge, and M. Millhouse, BayesWave analysis pipeline in the era of gravitational wave observations, Phys. Rev. D **103**, 044006 (2021), arXiv:2011.09494 [gr-qc].

[26] S. Hourihane, K. Chatziioannou, M. Wijngaarden, D. Davis, T. Littenberg, and N. Cornish, Accurate modeling and mitigation of overlapping signals and glitches in gravitational-wave data, Phys. Rev. D **106**, 042006 (2022), arXiv:2205.13580 [gr-qc].

[27] S. Hourihane and K. Chatziioannou, Glitches far from transient gravitational-wave events do not bias inference (2025), arXiv:2506.21869 [gr-qc].

[28] D. Williams, J. Veitch, M. L. Chiofalo, P. Schmidt, R. P. Udall, A. Vajpeji, and C. Hoy, Asimov: A framework for coordinating parameter estimation workflows, J. Open Source Softw. **8**, 4170 (2023), arXiv:2207.01468 [gr-qc].

[29] T. B. Littenberg and N. J. Cornish, Bayesian inference for spectral estimation of gravitational wave detector noise, Phys. Rev. D **91**, 084034 (2015), arXiv:1410.3852 [gr-qc].

[30] G. Ashton *et al.*, BILBY: A user-friendly Bayesian inference library for gravitational-wave astronomy, Astrophys. J. Suppl. **241**, 27 (2019), arXiv:1811.02042 [astro-ph.IM].

[31] I. M. Romero-Shaw *et al.*, Bayesian inference for compact binary coalescences with bilby: validation and application to the first LIGO–Virgo gravitational-wave transient catalogue, Mon. Not. Roy. Astron. Soc. **499**, 3295 (2020), arXiv:2006.00714 [astro-ph.IM].

[32] J. S. Speagle, dynesty: a dynamic nested sampling package for estimating Bayesian posteriors and evidences, Mon. Not. Roy. Astron. Soc. **493**, 3132 (2020), arXiv:1904.02180 [astro-ph.IM].

[33] C. Pankow, P. Brady, E. Ochsner, and R. O'Shaughnessy, Novel scheme for rapid parallel parameter estimation of gravitational waves from compact binary coalescences, Phys. Rev. D **92**, 023002 (2015), arXiv:1502.04370 [gr-qc].

[34] J. Lange *et al.*, Parameter estimation method that directly compares gravitational wave observations to numerical relativity, Phys. Rev. D **96**, 104041 (2017), arXiv:1705.09833 [gr-qc].

[35] D. Wysocki, R. O'Shaughnessy, J. Lange, and Y.-L. L. Fang, Accelerating parameter inference with graphics processing units, Phys. Rev. D **99**, 084026 (2019), arXiv:1902.04934 [astro-ph.IM].

[36] P. A. R. Ade *et al.* (Planck), Planck 2015 results. XIII. Cosmological parameters, Astron. Astrophys. **594**, A13 (2016), arXiv:1502.01589 [astro-ph.CO].

[37] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, Surrogate models for precessing binary black hole simulations with unequal masses, Phys. Rev. Research. **1**, 033015 (2019), arXiv:1905.09300 [gr-qc].

[38] G. Pratten *et al.*, Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes, Phys. Rev. D **103**, 104056 (2021), arXiv:2004.06503 [gr-qc].

[39] M. Colleoni, F. A. R. Vidal, C. García-Quirós, S. Akçay, and S. Bera, Fast frequency-domain gravitational waveforms for precessing binaries with a new twist, Phys. Rev. D **111**, 104019 (2025), arXiv:2412.16721 [gr-qc].

[40] L. Pompili *et al.*, Laying the foundation of the effective-one-body waveform models SEOBNRv5: Improved accuracy and efficiency for spinning nonprecessing binary black holes, Phys. Rev. D **108**, 124035 (2023), arXiv:2303.18039 [gr-qc].

[41] M. Khalil, A. Buonanno, H. Estelles, D. P. Mihaylov, S. Ossokine, L. Pompili, and A. Ramos-Buades, Theoretical groundwork supporting the precessing-spin two-body dynamics of the effective-one-body waveform models SEOBNRv5, Phys. Rev. D **108**, 124036 (2023), arXiv:2303.18143 [gr-qc].

[42] M. van de Meent, A. Buonanno, D. P. Mihaylov, S. Ossokine, L. Pompili, N. Warburton, A. Pound, B. Wardell, L. Durkan, and J. Miller, Enhancing the SEOBNRv5 effective-one-body waveform model with second-order gravitational self-force fluxes, Phys. Rev. D **108**, 124038 (2023), arXiv:2303.18026 [gr-qc].

[43] A. Ramos-Buades, A. Buonanno, H. Estellés, M. Khalil, D. P. Mihaylov, S. Ossokine, L. Pompili, and M. Shiferaw, Next generation of accurate and efficient multipolar precessing-spin effective-one-body waveforms for binary black holes, Phys. Rev. D **108**, 124037 (2023), arXiv:2303.18046 [gr-qc].

[44] E. Hamilton, L. London, J. E. Thompson, E. Fauchon-Jones, M. Hannam, C. Kalaghatgi, S. Khan, F. Pannarale, and A. Vano-Vinuales, Model of gravitational waves from precessing black-hole binaries through merger and ringdown, Phys. Rev. D **104**, 124027 (2021), arXiv:2107.08876 [gr-qc].

[45] J. E. Thompson, E. Hamilton, L. London, S. Ghosh, P. Kolitsidou, C. Hoy, and M. Hannam, PhenomXO4a: a phenomenological gravitational-wave model for precessing black-hole binaries with higher multipoles and asymmetries, Phys. Rev. D **109**, 063012 (2024), arXiv:2312.10025 [gr-qc].

[46] A. Gamboa *et al.*, Accurate waveforms for eccentric, aligned-spin binary black holes: The multipolar effective-one-body model SEOBNRv5EHM, Phys. Rev. D , (2025), arXiv:2412.12823 [gr-qc].

[47] A. Nagar, R. Gamba, P. Rettegno, V. Fantini, and S. Bernuzzi, Effective-one-body waveform model for noncircularized, planar, coalescing black hole binaries: The importance of radiation reaction, Phys. Rev. D **110**, 084001 (2024), arXiv:2404.05288 [gr-qc].

[48] A. Ramos-Buades, A. Buonanno, and J. Gair, Bayesian inference of binary black holes with inspiral-merger-ringdown waveforms using two eccentric parameters, Phys. Rev. D **108**, 124063 (2023), arXiv:2309.15528 [gr-qc].

[49] K. Mitman *et al.*, Probing the ringdown perturbation in binary black hole coalescences with an improved quasi-normal mode extraction algorithm (2025), arXiv:2503.09678 [gr-qc].

[50] L. Gao *et al.*, Robustness of extracting quasinormal mode information from black hole merger simulations, Phys. Rev. D **112**, 024025 (2025), arXiv:2502.15921 [gr-qc].

[51] M. Isi and W. M. Farr, Analyzing black-hole ringdowns (2021), lIGO-P2100227, arXiv:2107.05609 [gr-qc].

[52] M. Isi, Parametrizing gravitational-wave polarizations, Class. Quant. Grav. **40**, 203001 (2023), arXiv:2208.03372 [gr-qc].

[53] W. G. Anderson, P. R. Brady, D. Chin, J. D. E. Creighton, K. Riles, and J. T. Whelan, *Beam pattern response functions and times of arrival for earthbound interferometer*, Tech. Rep. LIGO-T010110 (LIGO Scientific Collaboration, 2002).

[54] L. S. Finn, The Response of interferometric gravitational wave detectors, Phys. Rev. D **79**, 022002 (2009), arXiv:0810.4529 [gr-qc].

[55] H. Siegel, M. Isi, and W. M. Farr, Ringdown of GW190521: Hints of multiple quasinormal modes with a precessional interpretation, Phys. Rev. D **108**, 064008 (2023), arXiv:2307.11975 [gr-qc].

[56] M. R. Sinha, L. Sun, and S. Ma, Impact of Detector Calibration Accuracy on Black Hole Spectroscopy (2025), arXiv:2506.15979 [gr-qc].

[57] M. Isi, W. M. Farr, M. Giesler, M. A. Scheel, and S. A. Teukolsky, Testing the Black-Hole Area Law with GW150914, Phys. Rev. Lett. **127**, 011103 (2021), arXiv:2012.04486 [gr-qc].

[58] M. Cabero, C. D. Capano, O. Fischer-Birnholtz, B. Krishnan, A. B. Nielsen, A. H. Nitz, and C. M. Biwer, Observational tests of the black hole area increase law, Phys. Rev. D **97**, 124069 (2018), arXiv:1711.09073 [gr-qc].

[59] A. Correia and C. D. Capano, Sky marginalization in black hole spectroscopy and tests of the area theorem, Phys. Rev. D **110**, 044018

(2024), arXiv:2312.15146 [gr-qc].
[60] S. J. Miller, M. Isi, K. Chatziioannou, V. Varma, and I. Mandel, GW190521: Tracing imprints of spin-precession on the most massive black hole binary, Phys. Rev. D **109**, 024024 (2024), arXiv:2310.01544 [astro-ph.HE].
[61] S. J. Miller, M. Isi, K. Chatziioannou, V. Varma, and S. Hourihane, Measuring spin precession from massive black hole binaries with gravitational waves: insights from time-domain signal morphology (2025), arXiv:2505.14573 [gr-qc].
[62] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, emcee: The MCMC Hammer, PASP **125**, 306 (2013), arXiv:1202.3665 [astro-ph.IM].
[63] J. Lin, Divergence measures based on the shannon entropy, IEEE Transactions on Information Theory **37**, 145 (1991).
[64] L. Blanchet, Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries, Living Rev. Rel. **17**, 2 (2014), arXiv:1310.1528 [gr-qc].
[65] G. B. Cook, Three-dimensional initial data for the collision of two black holes. ii. quasicircular orbits for equal-mass black holes, Phys. Rev. D **50**, 5025 (1994).
[66] T. W. Baumgarte, The Innermost stable circular orbit of binary black holes, Phys. Rev. D **62**, 024018 (2000), arXiv:gr-qc/0004050.
[67] L. Blanchet, D. Langlois, and E. Ligout, Innermost stable circular orbit of arbitrary-mass compact binaries at fourth post-Newtonian order (2025), arXiv:2505.01278 [gr-qc].
[68] J. Healy and C. O. Lousto, Ultimate Black Hole Recoil: What is the Maximum High-Energy Collision Kick?, Phys. Rev. Lett. **131**, 071401 (2023), arXiv:2301.00018 [gr-qc].
[69] M. Raveri and W. Hu, Concordance and Discordance in Cosmology, Phys. Rev. D **99**, 043506 (2019), arXiv:1806.04649 [astro-ph.CO].
[70] L. Verde, T. Treu, and A. G. Riess, Tensions between the Early and the Late Universe, Nature Astron. **3**, 891 (2019), arXiv:1907.10625 [astro-ph.CO].