Search for gravitational-lensing signatures in the full third observing run of the LIGO–Virgo network

The LIGO Scientific Collaboration, the Virgo Collaboration and the KAGRA Collaboration

ABSTRACT

Gravitational lensing by massive objects along the line of sight to the source causes distortions of gravitational wave-signals; such distortions may reveal information about fundamental physics, cosmology and astrophysics. In this work, we have extended the search for lensing signatures to all binary black hole events from the third observing run of the LIGO–Virgo network. We search for repeated signals from strong lensing by 1) performing targeted searches for subthreshold signals, 2) calculating the degree of overlap amongst the intrinsic parameters and sky location of pairs of signals, 3) comparing the similarities of the spectrograms amongst pairs of signals, and 4) performing dual-signal Bayesian analysis that takes into account selection effects and astrophysical knowledge. We also search for distortions to the gravitational waveform caused by 1) frequency-independent phase shifts in strongly lensed images, and 2) frequency-dependent modulation of the amplitude and phase due to point masses. None of these searches yields significant evidence for lensing. Finally, we use the non-detection of gravitational-wave lensing to constrain the lensing rate based on the latest merger-rate estimates and the fraction of dark matter composed of compact objects.

1. INTRODUCTION

Gravitational lensing occurs when a massive object bends spacetime in a way that alters the path or properties of a propagating wave. Gravitational lensing is expected to affect gravitational waves (GWs), resulting, for example, in repeated signals, (de-)magnification of the amplitude, phase shifts, and beating patterns (Ohanian 1974; Thorne 1982; Deguchi & Watson 1986; Wang et al. 1996; Nakamura 1998; Takahashi & Nakamura 2003). The exact alteration of the gravitational waveform depends on the nature of the lens system.

For massive lenses, gravitational lensing changes the GW amplitude without affecting the frequency evolution (Wang et al. 1996; Dai & Vennumadhav 2017; Ezquiaga et al. 2021). Moreover, such systems may also produce multiple signals observed as repeated events separated by a time delay of minutes to months for galaxies (Ng et al. 2018; Li et al. 2018; Oguri 2018), and up to years for galaxy clusters (Smith et al. 2018, 2017, 2019; Robertson et al. 2020; Ryczanowski et al. 2020). The current-generation GW detector network has a realistic chance of detecting the first lensed signal within its operation period (Ng et al. 2018; Li et al. 2018; Oguri 2018).

For low-mass lenses, such as stars or compact objects, microlensing introduces beating patterns in the waveform (Deguchi & Watson 1986; Nakamura 1998; Takahashi & Nakamura 2003; Cao et al. 2014; Jung & Shin 2019; Lai et al. 2018; Christian et al. 2018; Dai et al. 2018; Diego 2020). More generally, a field of light lenses may produce even more complex patterns on the gravitational waveform (Diego et al. 2019; Pagano et al. 2020; Cheung et al. 2021). Under the right conditions and with sufficient knowledge about the lens, these beating patterns may be observable with current-generation GW detectors.

The detection of lensed GWs paves the way for numerous scientific pursuits, including source localization (Hannuksela et al. 2020) and characterization (Lai et al. 2018; Diego 2020; Oguri & Takahashi 2020), precision cosmology (Sereno et al. 2011; Liao et al. 2017; Cao et al. 2019; Li et al. 2019b; Hannuksela et al. 2020), and tests of general relativity (Baker & Trodden 2017; Collett & Bacon 2017; Fan et al. 2017; Goyal et al. 2021; Ezquiaga & Zumalacárrregui 2020). Indeed, the prospects for fundamental physics and astrophysics have sparked a wide interest in searching for lensed GWs. Previous work from the LIGO–Virgo Collaboration has considered a range of strong and microlensing signatures for events in the first half of the third observing run (O3a) (Abbott et al. 2021a). Nevertheless, these studies have yielded no confident evidence for GW lensing.

In this work, we search for a variety of lensing signatures in the third LIGO Scientific, Virgo, and KAGRA (LVK) Collaboration Gravitational-Wave Transient Catalog (GWTC-3) (Abbott et al. 2021b) and study its implications for GW lensing. In particular, we expand on the lensing results presented for the first half of the third observing run of the LIGO–Virgo network (O3a) (Abbott et al. 2021a) by including the signals found in the second half of the third observing run (O3b) and by including additional analyses to further test the lensing hypothesis and interpret their outcomes. First, we search for the effects of strong lensing by studying the similarity and lensing evidence for pairs of binary black hole (BBH) mergers. We consider both pairs of detected mergers (super-threshold) and pairs formed by detected mergers and candidates that nominally fall below the detection threshold (sub-threshold).
with consistent waveform morphologies. Second, we search for evidence of microlensing induced by point-mass lenses. Finally, we constrain the expected rate of lensed signals, black hole (BH) merger-rate density, and the fraction of dark matter composed of compact objects.

It is important to note that GWTC-3 is a cumulative catalog describing all the GW transients found in all observing runs to date: O1, O2, O3a, and O3b. O1 made observations between 2015 September 12 00:00 UTC to 2016 January 19 16:00 UTC, O2 between 2016 November 30 16:00 UTC to 2017 August 25 22:00 UTC, O3a between 2019 April 1 15:00 UTC to 2019 October 1 15:00 UTC, and O3b between 2019 November 1 15:00 UTC to 2020 March 27 17:00 UTC.

Results of all analyses in this paper and associated data products can be found in Abbott et al. (2021c). GW strain data (GWOSC 2021) and posterior samples (Abbott et al. 2021d) for all events from GWTC-3 are available from the Zenodo platform or the Gravitational Wave Open Science Center (Abbott et al. 2021e).

2. DATA AND EVENTS

The analyses presented here expand on the lensing results from the first half of O3 (also referred to as O3a) by documenting new results from the second half of O3 (also referred to as O3b) using GWTC-3 (Abbott et al. 2021f). The O3a lensing results paper (Abbott et al. 2021a) used the GWTC-2 catalog (Abbott et al. 2021g). Since then, GWTC-2.1 (Abbott et al. 2021h) has reclassified 2 of the candidates used in the O3a lensing paper as having a probability of astrophysical origin of less than 0.5 and are not included in the results described here, specifically GW190424_180648 and GW190909_114149. GWTC-3 also includes 5 events that were identified by the O3a lensing sub-threshold counterpart image search, namely GW190925_233845, GW190426_190642, GW190725_184728, GW190805_211137, and GW190916_200658.

Various instrumental upgrades have led to more sensitive data in O3b, with a median binary neutron star (BNS) inspiral ranges (Finn & Chernoff 1993; Allen et al. 2012a) of 115 Mpc in O3b compared to 108 Mpc in O3a for LIGO Hanford, 133 Mpc in O3b compared to 135 Mpc O3a for LIGO Livingston, and 51 Mpc in O3b compared to 45 Mpc in O3a for Virgo (Abbott et al. 2021f). The duty factor for at least one detector being online was 96.6%; for any two detectors being online at the same time was 85.3%; and for all three detectors together was 51%. Further details regarding instrument performance and data quality for O3b are available in Abbott et al. (2021f); Davis et al. (2021a); Acernese et al. (2022).

The LIGO and Virgo detectors used a photon recoil-based calibration (Karki et al. 2016; Cahillane et al. 2017; Viets et al. 2018) resulting in a complex-valued, frequency-dependent detector response. Previous studies have documented the systematic error and uncertainty bounds for O3b strain calibration in LIGO (Sun et al. 2020, 2021) and Virgo (Acernese et al. 2021).

Transient noise sources, referred to as glitches, contaminate the data and can affect the confidence of candidate detections. Times affected by glitches and other data quality issues are identified so that searches for GW events can exclude (veto) these periods of poor data quality (Abbott et al. 2016a, 2020a; Davis et al. 2021b; Nguyen et al. 2021; Fiori et al. 2020). In addition, several known persistent noise sources are subtracted from the data using information from witness auxiliary sensors (Driggers et al. 2019; Davis et al. 2019).

Candidate events, including those reported in Abbott et al. (2021f) and the new candidates found by the search for sub-threshold counterpart images in Sec. 3.1 of this paper, have undergone a validation process to evaluate if instrumental artifacts could affect the analysis; this process is described in detail in Sec. 5.5 of Davis et al. (2021b). This process can also identify data quality issues that need further mitigation for individual events, such as the subtraction of glitches (Cornish et al. 2021; Davis et al. 2022) and non-stationary noise couplings (Vajente et al. 2020), before executing parameter estimation (PE) algorithms. See Table XIV of Abbott et al. (2021f) for the list of events requiring such mitigation.

The GWTC-3 catalog (Abbott et al. 2021f) contains 35 events from O3b in addition to the 55 previous events from previous observing runs (Abbott et al. 2021h) with a false-alarm rate (FAR) below two per year, and an expected rate of contamination from detector noise less than 10–15% (Abbott et al. 2021f). We neglect the potential contamination in this analysis. These events were identified by four search pipelines: one minimally modeled transient search cWB (Klimenko et al. 2004, 2005, 2006, 2011, 2016) and three matched-filter searches GsrtLAL (Sachdev et al. 2019; Hanna et al. 2020; Messick et al. 2017), Multi-Band Template Analysis (MBTA) (Adams et al. 2016; Aubin et al. 2021), and PyCBC (Allen et al. 2012b; Allen 2005; Dal Canton et al. 2014; Usman et al. 2016; Nitz et al. 2017). Their parameters were estimated through Bayesian inference using the bilby (Ashton et al. 2019; Smith et al. 2020; Romero-Shaw et al. 2020) and RIFT (Pankow et al. 2015; Lange et al. 2017; Wysocki et al. 2019) packages. Both the matched-filter searches and PE use a variety of BBH waveform models which generally combine knowledge from post-Newtonian theory, the effective-one-body formalism, and numerical relativity (for general introductions to these approaches, see Blanchet 2014; Damour & Nagar 2016; Palenzuela 2020; Schmidt 2020 and references therein). The analyses in this paper rely on the same methods, and the specific waveform models and analysis packages used are described in each section.

Of the 35 events from O3b, 31 are likely BBHs, while four have component masses consistent with being below...
3M⊙ (Abbott et al. 2021i,f), thus potentially containing a neutron star. We consider these 35 events in the analyses documented in this paper. Specifically, we use the following input data sets for each analysis. The searches for sub-threshold counterpart images in Sec. 3.1 cover the whole O3 strain data set, using the same data quality veto choices as in Abbott et al. (2021f) but a strain data set consistent with the PE analyses: the final calibration version of LIGO data (Sun et al. 2021) with additional noise subtraction (Vajente et al. 2020). The posterior-overlap analysis in Sec. 3.2 starts from the posterior samples released with GWTC-3 (GWOSC 2021). The joint-PE analyses in Sec. 3.3 and microlensing analysis in Sec. 4 reanalyze the strain data in short segments around the event times, available from the same data release, with data selection and noise mitigation choices matching those of the PE analyses in Abbott et al. (2021f).

3. STRONG LENSING

If a GW travels close enough to a massive lens, it will produce multiple images, with the number of images depending on the lens profile and source lens geometry. This regime is known as the strong-lensing limit. Each of these lensed images $h_j^L$ will have a change in its amplitude, arrival time and phase compared to the emitted signal $h$ (Schneider et al. 1992):

$$h_j^L(f) = \sqrt{\mu_j} \exp \left[ i 2 \pi f \Delta t_j - i \text{sign}(f)n_i \pi \right] h(f),$$

for $n_i = 0, 1/2, 1$ for type I, II and III images, which correspond to different minima of the lensing potential. While the magnification $\mu_j$ and time delay $\Delta t_j$ do not affect the waveform morphology (they are completely degenerate with the luminosity distance and coalescence time) the frequency-independent lensing phase shift $n_i \pi$ could induce distortions when the signal has multiple frequency components (Dai & Venumadhav 2017; Ezquiaga et al. 2021). In particular, this occurs for type II images since type I does not have a phase shift, and type III only flips the overall sign, which is degenerate with shifting the polarization angle by $\pi/2$. The $\text{sign}(f)$ term is only there to ensure that the time domain waveform is real.

Making a distinction between effects that do and do not change the waveform morphology, we divide our search into two parts. First, we search for pairs of events consistent with the strong-lensing hypothesis. Some of these pairs will have sufficiently strong amplitudes that can be identified as confident detections (super-threshold) by the search pipelines used in Abbott et al. (2021g,h,f), while others may have not been identified as signals (sub-threshold) because of the relative de-magnification. Our searches will include both sub- and super-threshold pairs. A pair is the minimum association, but higher multiplicities are also possible. Then, we search for strong lensing focusing on the distortion of type II images.

3.1. Sub-threshold Search

In this section, we describe the search for possible sub-threshold counterparts of super-threshold detections from O3. We perform searches over all O3 strain data following the rules for data selection described in Abbott et al. (2021h) and Abbott et al. (2021f). A general search for GWs uses a large template bank covering a broad parameter space as we have no prior information about the signal subspace, resulting in a high trials factor and hence incurring a high noise background. Sub-threshold (lensed) GWs with smaller amplitudes will therefore be easily buried in the noise without being identified as detections as they cannot pass the usual detection threshold.

To uncover these sub-threshold (lensed) signals, we have to effectively reduce the noise background while keeping the targeted foreground (i.e. the signals) constant (Li et al. 2019a; McIsaac et al. 2020; Dai et al. 2020). The strong lensing hypothesis asserts that lensed GWs, super-threshold or sub-threshold, coming from the same origin have identical waveforms apart from an overall scaling factor and a Morse phase factor as described in Eq. 1, and hence should have consistent inferred intrinsic masses and spins. Therefore, we can construct a reduced template bank with only templates that have masses and spins similar to those of a target super-threshold detection. Using the reduced bank lowers the trials factors and noise background and effectively searches for previously unidentified possible sub-threshold lensed counterparts to the target detection. For each known candidate from O3 with a probability of astrophysical origin $p_{astro} > 0.5$, we create a reduced template bank using their respective public posterior mass and spin samples released with GWTC-3 (GWOSC 2021), ensuring that the templates will match well with their respective target events while improving the ranking statistics of the search for similar events, and hence potentially returning new candidates that previously did not reach the threshold $p_{astro} > 0.5$ in GWTC-3. Details of how the reduced banks are constructed can be found in Li et al. (2019a).

Given these template banks, we proceed with configurations and procedures as outlined in Abbott et al. (2021a) to produce a priority list of potential lensed candidates matching each target event, using GstLAL (Messick et al. 2016; Sachdev et al. 2019) as the search pipeline. The list of candidates obtained is again further vetted using a sky location consistency check detailed in Wong et al. (2021); Abbott et al. (2021a) to ensure the candidates have consistent sky location with the target event. To avoid false dismissal at this step, we only veto candidates with an overlap in 90% credible region of the sky location $O_{90\%\text{CR}} = 0$. All candidates with non-vanishing

\footnote{The Morse phase factor for different image types has not been considered in the search described here. Should a GW include detectable higher-order multipole moments, then the Morse phase factor will cause complicated changes to the waveforms, inducing a loss in the search sensitivity.}
localization overlap are kept for further follow-up with data quality checks as discussed in Sec. 2.

In Table 1, we list the top five candidates from the individual targeted searches for counterparts of the detections reported in O3. As in the O3a lensing paper (Abbott et al. 2021a), we do not assess in detail the probability of astrophysical origin for each of these. It is also important to note that the reported false-alarm rates (FARs) do not indicate how likely each trigger is a lensed counterpart of the target event, but only how likely noise produces a trigger with a ranking statistic higher or equal to that of the candidate under consideration using these reduced template banks. Similar to Abbott et al. (2021a), we account for the fact that we have analyzed ~ 332 days of data multiple times for a total of 76 events, and set the FAR threshold to be 1 in 69 years (i.e. 4.59 × 10⁻¹⁰ Hz). We followed up on the top two candidates listed that passed the FAR threshold through golium’s joint PE analysis (Janquart et al. 2021, discussed in 3.3). The results are included in Table 1. Since both pairs of candidates have mildly negative log₁₀ coherence ratios, showing that there is no evidence supporting the lensing hypothesis for either of these pairs, we did not further follow them up with the more computationally intensive hanabi analysis (Lo & Magaña Hernandez 2021, discussed in 3.3).

3.2. Preliminary Identification of Lensed Pair Candidates

Multiple, non-overlapping images produced by strongly lensed GW signals have identical phase evolution, and therefore their intrinsic parameters (as well as their orbit’s inclination with respect to the line of sight) are expected to have overlapping posteriors. In addition, the angular separations of images (produced by galaxies or galaxy clusters) are several orders of magnitude smaller than the uncertainties associated with their GW sky location. As a result, their sky localisations will also overlap. (As in the previous section, the Morse phase for different image types is not considered here.)

Under these assumptions, a Bayes-factor statistic (B_L) that assesses the consistency between a lensed candidate pair’s posterior distributions of intrinsic parameters, sky location, and inclination angle (and thus acts as a discriminator between the lensed and unlensed hypotheses) can be constructed (Haris et al. 2018). To convert this statistic to a false-positive probability (FPP), a background distribution of unlensed B_L needs to be estimated.

To that end, we conduct an injection campaign involving BBHs only, in which we sample component masses m_{1,2} from a power-law distribution (Abbott et al. 2016b) in the range (10–50M_☉). We assume that the redshift distribution of BBHs is similar to population synthesis simulations of isolated binary evolution (Belczynski et al. 2008, 2010; Dominik et al. 2013; Eldridge et al. 2019; Bouffanais et al. 2021; Zevin et al. 2021). All other parameters are sampled from uninformative prior distributions (Haris et al. 2018). We inject the simulated signals into Gaussian noise with O3a representative power spectral density (PSD) for a LIGO–Virgo detector network. We compute B_L for all possible pairs in this injection set, and following Abbott et al. (2021a), we assign an FPP to a candidate pair using its B_L. Candidate lensed pairs involving BNS or neutron star-black hole (NSBH) events are not analyzed and ranked.

We additionally employ a machine learning (ML)-based binary classification scheme to rapidly provide a probability of class membership (lensed or unlensed) for a given candidate BBH pair (Goyal et al. 2021a). Such an analysis not only serves as an independent method to rank candidate pairs but also provides a quantitative significance to pairs for which source-parameter inference samples are unavailable.

Q-transform-based (Chatterji et al. 2004) time–frequency maps of strongly lensed BBHs are expected to have similar shapes, although the signal energy in each time-frequency tile will differ between images. Furthermore, as mentioned earlier, their sky localisations will overlap. Exploiting these facts, ML models that take Q-transforms and Bayestar (Singer & Price 2016) sky localisations as inputs, are built. These models use a DenseNet (Huang et al. 2016) architecture (with several layers pre-trained on the ImageNet dataset; Deng et al. 2009), and XGBoost (Chen & Guestrin 2016) algorithms, trained on lensed and unlensed BBH signals injected in Gaussian noise (for details on the ML training set, see Goyal et al. (2021a)). The outputs of the individual models are then combined to provide a probability that a candidate pair is lensed or unlensed.

To convert this probability to an FPP, we construct a background distribution of ML probabilities using a population of unlensed BBH events injected in Gaussian noise characterized by the O3a representative PSD – the same as was used for the posterior overlap statistic. This PSD is found to be sufficiently similar to the averaged O3 PSD for the estimation of the background distribution so as not to change the preliminary selection of candidate pairs. The BBH population is identical to the one used by the posterior overlap statistic analysis to construct its corresponding background distribution. Furthermore, the sky localisations used to rank candidate pairs come from the same PE analysis used to estimate the posterior overlap statistic.

Note that Bayestar, which is used to assign ML probabilities to real-event candidate pairs, is expected to provide sky localisations that are similar to those provided by this PE analysis.
## Table 1. Top 5 candidates from individual sub-threshold searches for strongly-lensed counterpart images of O3 events from GWTC-3.

<table>
<thead>
<tr>
<th>Target event</th>
<th>Lensed candidate (UTC)</th>
<th>$\Delta t$ [days]</th>
<th>($1 + z$)$M_\odot$ ($M_{\odot}$</th>
<th>($1 + z$)$M_{\mathrm{GW}}^\text{min}$($M_{\odot}$)</th>
<th>FAR yr$^{-1}$</th>
<th>$O_{\text{GW/CR}}$ [%]</th>
<th>$\log_{10}(C^1_{\text{LL}})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GW190930,133541</td>
<td>19-08-05 13:43:48</td>
<td>−56.0</td>
<td>10.2</td>
<td>9.86</td>
<td>0.002</td>
<td>61.000%</td>
<td>−6.4</td>
</tr>
<tr>
<td>GW191204,171526</td>
<td>19-08-05 13:43:48</td>
<td>−121.1</td>
<td>10.2</td>
<td>9.66</td>
<td>0.006</td>
<td>25.40%</td>
<td>−12.2</td>
</tr>
<tr>
<td>GW190828,065509</td>
<td>19-11-12 12:13:18</td>
<td>76.2</td>
<td>45.5</td>
<td>17.3</td>
<td>0.023</td>
<td>18.00%</td>
<td>−</td>
</tr>
<tr>
<td>GW190725,174728</td>
<td>19-08-05 13:43:48</td>
<td>10.8</td>
<td>10.2</td>
<td>8.88</td>
<td>0.038</td>
<td>39.50%</td>
<td>−</td>
</tr>
<tr>
<td>GW190828,065509</td>
<td>20-02-06 07:24:59</td>
<td>162.0</td>
<td>15.1</td>
<td>17.3</td>
<td>0.154</td>
<td>36.00%</td>
<td>−</td>
</tr>
</tbody>
</table>

Note—The first column lists the target event from O3. The second column shows the time (YY-MM-DD HH-MM-SS) in UTC of the found sub-threshold candidate. The third column shows the time difference (in days) between the candidate and the target event. The fourth column shows the redshifted chirp mass of the template that found the trigger. The fifth column shows the redshifted chirp mass of the target event. The sixth column shows the FARs from the individual search for the new candidate from the second column. The seventh column shows the percentage overlap of the 90% sky localization regions between the candidate and the target event. The eighth column shows the log$_{10}$ coherence ratio obtained from golum’s joint PE analysis.

A plot comparing the FPPs assigned by the posterior overlap and ML analyses is shown in Fig. 1. Candidates that have either a posterior-overlap-assigned FPP or ML-assigned-FPP, (or both), that are smaller than 1%, are selected for more comprehensive Bayesian analyses.

**Figure 1.** The FPPs of each lensed candidate pair constructed from the set of GW events that exceed an astrophysical probability (Farr et al. 2015; Kapadia et al. 2020) threshold of 0.5, as evaluated using the $B^C$ and ML classification statistics. Orange dashed lines that correspond to an FPP threshold of $10^{-4}$, are also placed. Pairs whose $B^C_{\text{L}}$-based or ML-based FPPs fall below this threshold are selected for additional joint parameter estimation analyses. $B^C_{\text{L}} < 10^{-4}$ has been mapped to an FPP of 1, which is reflected in the gap along the vertical axis between 0.4 and 1.

### 3.3. Joint Parameter Estimation

Similar to the analysis of O3a data (Abbott et al. 2021a), we perform a joint PE analysis for the most relevant candidate lensing pairs. We follow up on the pairs that display low FPP in their posterior overlap or ML classification scheme. These are pairs within the whole of O3, but we only consider here those with at least one event in O3b since pairs in O3a were studied in Abbott et al. (2021a). We use two complementary pipelines: *golum* (Janquart et al. 2021) and *hanabi* (Lo & Magaña Hernández 2021). Both pipelines use the nested sampling algorithm *dynesty* (Speagle 2020), and implement the joint PE with the help of *bilby* (Ashton et al. 2019; Romero-Shaw et al. 2020).

*golum* (Janquart et al. 2021) is a joint PE tool where the workload is reduced by analyzing the two images, under the lensed hypothesis, in two successive stages. The first image is characterized by the same parameters of the unlensed case (where the time of coalescence and the luminosity distance are the observed ones) with an additional Morse factor. The second image is then analyzed using (samples of) the posterior from the first image as the prior and linking the parameters modified by lensing through three lensing parameters: a time difference, a relative magnification, and a Morse factor difference. The final coherence ratio $C^1_{\text{LL}}$ is the ratio of the product of the evidences for the two runs under the lensed hypothesis and the product of evidences for the two images analyzed under the unlensed hypothesis.

*hanabi* (Lo & Magaña Hernández 2021) first performs a joint inference on a signal pair by constructing a joint likelihood function that is a product of the likelihood function for each individual event, with a joint prior distribution. The latter is defined for a set of joint parameters that can simultaneously describe both signals if they are truly lensed, for example, the masses and the spins, as well as a set of parameters that are different for each of the signals such as the time of arrival, the apparent luminosity distance, and the Morse phase factor associated to each of the lensed signals. The joint parameter space is explored with the package *hanabi.inference* (Lo & Magaña Hernández 2021). The inference result is then reweighted with an astrophysically motivated prior distribution; for example, the astrophysical prior distribution for the
redshifted component masses would be dependent on both the population model for the intrinsic BBH masses and the redshift distribution of the sources. However, the true source redshift cannot be determined from GW observations alone since the true source redshift is degenerate with the magnification from strong lensing. To compute the Bayes factor $B_U^L$, the source redshift, which serves as a hyper-parameter for the signal pair, must be marginalized over. Selection effects enter as a normalization constant to the marginal data likelihood. This procedure is implemented in hanabi.hierarchical with the help of gwpopulation (Talbot et al. 2019). The ratio of unnormalized evidences calculated under the lensed hypothesis and the unlensed hypothesis using this astrophysical prior is referred to as the population-weighted coherence ratio $C_U^{\text{pop}}$, while the ratio of normalized evidences that accounts for both population prior and selection effects is referred to as the Bayes factor $B_U^L$ in this analysis. We follow our fiducial singular isothermal sphere (SIS) lensing model when computing the magnification prior (Abbott et al. 2021a). This analysis however does not impose any informative prior on the time delay or the image types from the lensing model.

Both pipelines use IMPPhenomXPHM (Pratten et al. 2021) as the waveform model, with an additional Morse phase applied to each of the waveform polarizations in the frequency domain. Other inputs, such as the power spectral density estimates and the calibration envelopes, are chosen to match the analyses done in the GWTC-3 catalog paper (Abbott et al. 2021b). Following the same prescriptions of the other analyses, we fix the BBH population model to the Power-Law + Peak model for the primary masses and the merger rate history to Madau–Dickinson star-formation rate (Madau & Dickinson 2014) normalized by the median GWTC-3 rate (Abbott et al. 2021j).

Taking advantage of golum’s rapid joint PE, we analyze the 75 pairs of candidates highlighted by posterior overlap and ML. For each of them, we compute the coherence ratio, which accounts for the probability ratio of the lensed and unlensed hypotheses without including selection effects and population priors. We find that there is a wide range of $\log_{10}(C_U^L)$ values, with a peak slightly above zero. This comes from the fact that this analysis considers only triggers already flagged by the posterior overlap and ML analyses. As a consequence, the analysis is biased towards the higher values. Nevertheless, a significant proportion of events flagged with the ML pipeline and the posterior overlap pipeline are disfavored, having $\log_{10}(C_U^L) < 0$. When comparing the highest coherence ratio found in the data, $\log_{10}(C_U^L) = 2.5$, with a background of unlensed events, we find that it is well within the expected values, with $1\%$ of the background events having larger $C_U^L$. This background is computed for a population of compact binaries that follows the mass, spin and redshift distribution of GWTC-3 (Abbott et al. 2021j). This large number of possible $\log_{10}(C_U^L)$ is consistent with the high number of expected false alarms (Wierda et al. 2021; Çalışkan et al. 2022a). For those pairs with the highest coherence ratio, we follow up with the hanabi pipeline for a total of 17 pairs. Our main results are presented in Fig. 2, where the left column indicates the event pairs and the horizontal axis their $B_U^L$. There we can observe that none of the event pairs shows support for the lensing hypothesis, i.e. all $B_U^L < 1$. The pair with highest $B_U^L$ is GW190620_030421 – GW200216_220804, for an evidence against lensing of $\sim 1/100$ with the fiducial merger rate density model following the Madau-Dickinson star-formation rate. As a robustness check of how using different merger rate density models would change the results, we repeat the calculations using two more models, namely $R_{\min}(z)$ and $R_{\max}(z)$ from our previous O3a analysis (Abbott et al. 2021a) that minimally and maximally bracket many existing population-synthesis results (Belczynski et al. 2008, 2010; Dominik et al. 2013; Eldridge et al. 2019). We see that while the exact values for the Bayes factor change with the use of different merger rate density models, the conclusion remains that there is no support for the lensing hypothesis in any of the event pairs analyzed. To further assess the significance of these pairs we also include a color code to indicate the probability of having an astrophysical origin $p_{\text{astro}}^\text{pair}$, defined as the product of the highest $p_{\text{astro}}$ of each event reported in the GWTC-3 catalog paper (Abbott et al. 2021b) by different pipelines. In conclusion, we find no evidence of multiply imaged events.

### 3.4. Type II image search

In addition to the search for strong-lensing identifying multiple images, we also look for the distortions that lensing introduces in type II images (Ezquiaga et al. 2021). This is because the frequency-independent phase shift that each image acquires becomes a frequency-dependent time delay for different frequency components. Therefore, for signals containing different measurable spherical harmonic modes, as recently detected in GW190412 (Abbott et al. 2020c), GW190814 (Abbott et al. 2020d), and other events (Abbott et al. 2021b), the overall lensed waveform can be distorted. The extent of the distortion is subject to the power in modes beyond the quadrupole radiation. As a consequence, we do not expect to see these distortions in the majority of the lensed events with current sensitivities. However, if not searched for, they might be mistaken with deviations from general relativity (Ezquiaga et al. 2022).

To look for these distortions, we use golum (Januart et al. 2021). Within GWTC-3 we identify 10 events whose posterior has some information about the Morse phase, either by favoring or disfavoring the distortions of the type II image by more than $4\%$ with respect to normality, i.e. the probability of each image type $p(n_j)$ is $p(n_j) > 0.37$ or $p(n_j) < 0.29$. We summarize the evidence of one image type versus another in
Figure 2. Bayes factors $B^L_U$ from hanabi for the highest-ranked multiple-image candidate pairs. As a check on the robustness of our results, we show the Bayes factors calculated using three different merger rate density models, namely the fiducial model tracking the Madau–Dickinson star-formation rate (Madau & Dickinson 2014), and also the $R_{\min}(z)$ and $R_{\max}(z)$ model introduced in Abbott et al. (2021a). The color for each marker represents the value of $p_{\text{pair}}$ for each pair, which is the probability that both of the signals from a pair are of astrophysical origins and not from terrestrial sources.

Fig. 3. Since only type II images display waveform distortions, we only compute the Bayes factors of the type-II-vs-I and the type-II-vs-III hypotheses. As can be seen in Fig. 3, only a few events display a preference for one image type versus the other one. This is expected given the signal-to-noise ratio (SNR) of these events and their power in higher multipole moments. However, GW190412 and GW200129_065458 present higher evidence for type II images. For GW190412 we find a log10 Bayes factor for type II vs. I of 0.60 ± 0.16 and for type II vs. III of 0.22 ± 0.16. For GW200129_065458 we find 0.38 ± 0.14 and 0.24 ± 0.14 for type II vs. I and type II vs. III respectively. These events have possible super-threshold counterparts but those were discarded by the golum analysis. In addition, we have also searched for sub-threshold triggers associated with these events, but found none.

To assess the significance of the type II images, we follow up on GW190412 and GW200129_065458 performing a simulation campaign of type I and type II images. GW190412 simulations show that indeed this event has enough power in higher multipole moments to favor the type II hypothesis so that it could meaningfully test that hypothesis and would favor it if it were true. For GW200129_065458, however, that is not the case. Moreover, GW200129_065458 might have a significant glitch under subtraction (Payne et al. 2022). The preference of GW190412 for a type II image could be just a systematic effect due to the waveform modeling, especially since this event falls in challenging parts of the parameter space (Abbott et al. 2020c; Colleoni et al. 2021; Hannam et al. 2021). For this reason, we repeat the analysis with different waveform families from our fiducial IMRPhenomXPHM model (Pratten et al. 2021). We find that the preference for a type II image remains when using SEOBNRv4PHM (Ossokine et al. 2020) or IMRPhenomPv3HM (Khan et al. 2020). The same conclusion holds when using different noise realizations for the simulations. Details on these simulation campaigns can be found in Appendix A.
Although we find a mild preference for the type II image hypothesis in GW190412, we find that this analysis cannot provide conclusive evidence of strong lensing. However, our techniques and pipeline will be relevant for future observing runs when high-SNR events display stronger evidence of higher-order modes.

4. MICROLENSING EFFECTS

When the characteristic wavelengths of GWs are comparable to the Schwarzschild radius of a lens ($\lambda_{GW} \sim R_{Sch}$), we may observe frequency-dependent magnification of the waveform that can inform us about the lens model (Takahashi & Nakamura 2003; Cao et al. 2014; Jung & Shin 2019; Lai et al. 2018; Christian et al. 2018; Dai et al. 2018; Diego et al. 2019; Diego 2020; Pagano et al. 2020; Cheung et al. 2021; Cremonese et al. 2021; Çalışkan et al. 2022b). Since the GWs of sources such as BBHs sweeps through a wide range of frequencies, these beating patterns can reveal the presence of intervening microlenses. In the sensitive range of ground-based detectors, these effects are expected for objects up to $\sim 10^5 M_\odot$, which includes stellar-mass objects and intermediate-mass BHs.

Objects that can cause these microlensing effects are predominantly found in larger structures. Therefore we expect that realistic microlensing due to a field of microlenses embedded in an external macromodel potential such as galaxies and galaxy clusters causes complex effects on the unlensed waveforms (Diego et al. 2019). While the effects of these systems on GW signals have been studied (Diego 2020; Cheung et al. 2021; Mishra et al. 2021; Yeung et al. 2021), the resulting waveforms are computationally costly to evaluate. Nevertheless, in the absence of specific knowledge of the matter distribution along the travel path and to keep the problem computationally tractable, we assume that the beating patterns are caused by isolated point masses as a first approximation. In this case, the microlensed waveform $h^{ Micro}$ can be related to the unlensed waveform $h^U$ according to

$$h^{ Micro}(f; \theta, M^L_\ell, y) = h^U(f; \theta) F(f; M^L_\ell, y),$$

(2)

where $\theta$ represents the set of parameters defining an unlensed GW signal, $M^L_\ell = M_\ell(1 + z_l)$ is the redshifted lens mass, $y$ is the dimensionless impact parameter, and $F(f; M^L_\ell, y)$ is the frequency-dependent lensing magnification factor (e.g., Takahashi & Nakamura 2003).

Similar to Abbott et al. (2021a), we perform Bayesian inference on all events from O3b using the unlensed signal model $h^U$ and the microlensing signal model $h^{ Micro}$. In particular, we use bilby (Ashton et al. 2019; Romero-Shaw et al. 2020) and the nested sampling algorithm dynesty (Speagle 2020). Data products such as strain data and PSDs are the same as for GWTC-3 and between the two signal models (Abbott et al. 2021b) For the GW parameters, we use the same priors as GWTC-3, while the prior on the lens mass $M^L_\ell$ is log uniform in the range $[1-10^5 M_\odot]$ and the prior on the impact parameter is $p(y) \propto y$ between $[0, 1]$. All events were analyzed using IMRPhenomXPHM (Pratten et al. 2021).

The process yields posterior probability distributions of $\theta$ or $[\theta, M^L_\ell, y]$ for the unlensed and lensed signal models, respectively. Moreover, we compute the evidence ratio between the microlensed and unlensed signal models, better known as the Bayes factor $B^ Micro_U$. Fig. 4 shows the distribution of $\log_{10} B^ Micro_U$ for all the events in O3 and simulated unlensed signals (orange, dashed line) from Abbott et al. (2021a).

Figure 4. Distribution of microlensing $\log_{10}$ Bayes factors $B^ Micro_U$ for all events in O3 (blue, solid line) and simulated unlensed signals (orange, dashed line) from Abbott et al. (2021a).

5. IMPLICATIONS

In this section, we consider some of the implications that derive from the search for lensing signatures. We first forecast the number of detectable strongly lensed events based on the latest knowledge on the merger-rate density (Sec. 5.1). Next, we infer upper limits on the strong lensing rate using the non-detection of resolvable strongly lensed BBH events (Sec. 5.2). Finally, we use the non-detection of microlensing to infer the compact dark matter fraction in the Universe (Sec. 5.3).

5.1. Strong lensing rate

We predict the rate of lensing using the standard methods outlined in the literature (Ng et al. 2018; Li et al. 2018; Oguri 2018; Xu et al. 2021; Mukherjee et al. 2021a; Wierda et al. 2021), at galaxy and galaxy-cluster lens mass scales. To model the lens population, we need to choose a density profile
and a mass function. We adopt the SIS density profile for both galaxies and galaxy clusters. Moreover, we use the velocity dispersion function from the Sloan Digital Sky Survey (Choi et al. 2007) for galaxies and the halo mass function from Tinker et al. (2008) for clusters which have also been used in other lensing studies (e.g., Oguri & Marshall 2010; Robertson et al. 2020). The SIS profile can accurately describe lensing by galaxies, but the mass distribution of clusters tends to be more complicated. Nevertheless, Robertson et al. (2020) have demonstrated that the SIS model can reproduce the lensing rate predictions from a study of numerically simulated cluster lenses. Thus, we adopt the same model for both galaxies and galaxy clusters.

Under the SIS model, we obtain two images with different magnifications and arrival times. The rate of strong lensing is given by

$$\mathcal{R}_{\text{lens}} = \int \frac{dN(M_h, z_1)}{dM_h} \frac{dD_c}{dz_1} \mathcal{R}_m(z_m) \frac{dV_c}{dz_m} \sigma(M_h, z_1, z_m, \rho, \rho_c) \times p(\rho | z_m) \, d\rho \, dz_m \, dz_1 \, dM_h,$$

where $dN(M_h, z_1)/dM_h$ is the differential comoving number density of lensing halos in a halo mass shell at lens redshift $z_1$, $D_c$ and $V_c$ are the comoving distance and volume, respectively, at a given redshift, $\mathcal{R}_m(z_m)$ is the total comoving merger rate density at redshift $z_m$, $(1+z_m)$ accounts for the cosmological time dilation, $p(\rho | z_m)$ is the distribution of SNR at a given redshift, $\rho_c$ is the network SNR threshold, and $\sigma$ is the lensing cross-section which indicates, as a function of its various arguments, how efficiently strong lensing will occur. We model the mass distribution of BBHs following the results for the Power Law + Peak model of Abbott et al. (2021)). We consider a merger rate density model that assumes the Madau–Dickinson ansatz (Madau & Dickinson 2014) that is consistent with recent results from GWTC-3. Moreover, we make use of the absence of a detected stochastic gravitational-wave background (SGWB) to further constrain the merger rate density (Abbott et al. 2021)). For consistency with previous analyses (e.g., Abbott et al. 2021k), we take the Hubble constant from Planck 2015 observations to be $H_0 = 67.9 \text{ km s}^{-1} \text{ Mpc}^{-1} \text{ (Ade et al. 2016)}$. Furthermore, we choose $\rho_c = 8$ as a point estimator of the detectability of GW signals. We find this choice to be consistent with the search results in Abbott et al. (2021g) and Sec. 3.1, and we estimate its impact to be subdominant compared to other sources of uncertainty.

In Table 2, we show our estimates for the relative rate of lensing expected to be observed by the LIGO–Virgo network of detectors. The results are shown separately for galaxy-scale and cluster-scale lenses. Furthermore, these rates are calculated for events that are doubly lensed and for two cases: when only a single event (i.e., the brighter one) is detected (S), and when both of the doubly lensed events are detected (D). The expected fractional rate of lensing (i.e. the lensed to unlensed rate) spans the range $O(10^{-4}–10^{-3})$, depending on the merger rate density assumed. We estimate the fractional rate of observed double (single) events for galaxy-scale lenses to lie in the range $1.9–11.0 \times 10^{-4}$ ($5.0–19.5 \times 10^{-4}$). Similarly, for cluster-scale lenses, the fractional rate is estimated to be in the range of $0.8–4.4 \times 10^{-4}$ ($2.0–7.6 \times 10^{-4}$), typically lower than the rates on galaxy scales. These estimates suggest that observing a lensed double image is unlikely at the current sensitivity of the LIGO–Virgo network of detectors. Nevertheless, at design sensitivity and with future upgrades, standard forecasts suggest that the possibility of observing such events might become significant (Ng et al. 2018; Li et al. 2018; Oguri 2018; Xu et al. 2021; Mukherjee et al. 2021a; Wierda et al. 2021). Compared with other lens models, our lensing rates are consistent with those predicted for singular isothermal ellipsoid (SIE) models (e.g., Oguri 2018; Xu et al. 2021; Wierda et al. 2021).

### 5.2. Implications from the non-observation of strongly lensed events

The absence of any detections of strongly lensed GW events before and during O3 provides a complementary way to constrain the merger rates of compact objects at high redshift. The detection of individual GW events has enabled measurement of the low redshift ($z < 1$) merger rate (Abbott et al. 2021j). However, the high redshift merger rate of GW sources is not yet measured directly, and we have only been able to place an upper limit on it from the absence of a detection of the SGWB (Abbott et al. 2021)). The absence of such a detection naturally leads to a bound on the lensing rate expected from GWTC-3 (Mukherjee et al. 2021b; Buscicchio et al. 2020).

By using the same power-law form for the merger rate as that used in Sec. 5.1, but extended up to $z = 2$, we obtain limits on the merger rate at redshift $z > 1$ from the absence of detections of strongly lensed events. The corresponding constraints (90% credible intervals) are shown in Fig. 5 as the cross-hatched region bounded by the dash-dotted curves. The changes in the upper bound of the merger rates are driven by the absence of detected lensing events, whereas the lower bound is driven by the low-redshift constraints on the merger rate. For comparison, the current limits on the merger rate from GWTC-3 up to redshift $z = 1$ (Abbott et al. 2021j), with the bounding curves extrapolated to higher redshifts $z > 1$, are shown as the grey shaded region bounded by the dotted curves. For further comparison, we have also plotted the solid black curves which show the current constraints from the absence of detection of the SGWB (Abbott et al. 2021)). The upper bounds on the merger rate from lensing are more stringent than the bounds from GWTC-3 at high redshift (Abbott et al. 2021j), and are also comparable with the bounds from the SGWB for redshifts $z < 1.2$. The slight difference between
Table 2. Expected fractional rates of observable lensed double events at current LIGO–Virgo sensitivity.

<table>
<thead>
<tr>
<th>Merger Rate Density</th>
<th>Galaxies</th>
<th>Galaxy Clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>$R_D$</td>
<td>$R_S$</td>
</tr>
<tr>
<td>GWTC-3+Stochastic</td>
<td>$1.9 \times 10^{-4}$</td>
<td>$5.0 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

Note—This table lists the relative rates of lensed double events expected to be observed by LIGO–Virgo at the current sensitivity where both of the lensed events are detected ($R_D$) and only one of the lensed events is detected ($R_S$) above the SNR threshold. The rates encompass a 90 percent credible interval. We show the rate of lensing by galaxies ($\sigma_{vd} = 10–300 \text{ km s}^{-1}$) and galaxy clusters ($\log_{10}(M_{halo}/M_\odot) \sim 14–16$) separately.

The constraints on the merger rates at low redshift derived from the SGWB (Abbott et al. 2021) and from GWTC-3 (Abbott et al. 2021j) arise because the bounds from the SGWB are obtained here using the previous constraints on the merger rate at low redshift derived using GWTC-2 (Abbott et al. 2021m).

5.3. Constraints on compact dark matter from gravitational-wave microlensing

Objects whose size is comparable to their gravitational radius, and that cause microlensing effects on GW signals, could be candidates for dark matter. Although their abundance is heavily constrained by several astronomical observations (Carr & Kuhnel 2020; Carr et al. 2020), the possibility of their contributing to dark matter cannot be ruled out in several mass windows.

Here we use the non-observation of microlensing effects on the GW signals detected by LIGO and Virgo to constrain the fraction of dark matter contributed by compact objects in the mass range $\sim 10^2–10^3 \, M_\odot$ (Jung & Shin 2019; Urrutia & Vaskonen 2021; Basak et al. 2021). The essential idea is that if a significant fraction of dark matter is in the form of compact objects, they would introduce detectable microlensing signatures on the GW signals that we observe.

Assuming that lensed and unlensed events occur as Poisson processes, we compute the posterior distribution on the lensing fraction ($u \equiv \Lambda_L/\Lambda$), defined as the ratio of Poisson means of lensed events to the total number of detected events. This is then used to compute the posterior of the fraction of compact dark matter ($f_{DM} = \Omega_{CO}/\Omega_{DM}$) (Basak et al. 2021). We take that a total of $N = 67$ BBH mergers are detected during the O3 run $^4$, and none of them is lensed (i.e., $N_f = 0$). We then estimate the posterior distribution of the lensing fraction $u$. Finally, the posterior of $f_{DM}$ can be computed as

$$p(f_{DM} \mid \{N_f = 0, N\}) = p(u \mid \{N_f = 0, N\}) \left| \frac{du}{df_{DM}} \right|,$$  

where $du/df_{DM}$ is the Jacobian that relates the observed fraction $u$ of lensed events to the compact dark matter fraction $f_{DM}$ in the Universe.

We determine this Jacobian by simulating astrophysical populations of BBH mergers lensed by point mass lenses (Basak et al. 2021). The constraints we obtain depend upon the assumed distributions of the component masses, spins and the redshifts of the mergers, which have considerable uncertainties. We assume that the masses are distributed according to the Power-law + Peak model of Abbott et al. (2021j) while spins are assumed to be aligned/antialigned with the orbital angular momentum with magnitudes distributed uniformly in $(0, 0.99)$. We use the approximant IMRPhenomD (Khan et al. 2016) to produce the waveforms. We consider different redshift distributions of the mergers: uniform distribution in comoving volume, the power-law model of Abbott et al. (2021j), the Madau-Dickinson model (Madau & Dickinson 2014), as well as some representative population-synthesis models given by Dominik et al. (2013) and Belczynski et al. (2016). In our simulations, compact objects are approximated by point mass lenses and distributed uniformly in comoving volume.

$^4$ These are the events cataloged in GWTC-3 that do not contain a neutron star component.

$^5$ The simulations are done assuming the O3b representative PSD and Gaussian noise. The Jacobian is not expected to change significantly if real noise is used instead.
volume. Binaries producing a network SNR of 8 or above in the LIGO–Virgo detectors are deemed detectable. In order to reduce the computational cost of performing the simulations, we estimate $2U_{\text{Micro}}$ using an approximation to the Bayes factor that is expected to be accurate in the high-SNR regime (Cornish et al. 2011; Vallisneri 2012). We then compute the fraction of detected events that produce a $2U_{\text{Micro}}$ larger than the highest $2U_{\text{Micro}}$ obtained from real LIGO–Virgo events. This lensing fraction is computed as a function of the $f_{DM}$, which is used to compute the Jacobian $du/df_{DM}$.

The largest value of the microlensing likelihood ratio obtained from GWTC-3 events is $\log_{10}2U_{\text{Micro}} = 0.799$. We compute the fraction of simulated events with $\log_{10}2U_{\text{Micro}} \geq 0.799$, for different lens masses. This allows us to compute the Jacobian $du/df_{DM}$ and thus the posterior on $f_{DM}$. The 90% upper limits are shown as a function of the lens mass (assuming a monochromatic spectrum) in Fig. 6. The bounds we obtain are weaker than some of the existing constraints (Carr & Kuhnel 2020; Carr et al. 2020). The GW lensing bounds will improve significantly in the next few years as the sensitivity of GW detectors improve (Abbott et al. 2018). Assuming ~300 BBH detections in O4 and $O(1000)$ detections in O5, the constraints on $f_{DM}$ will improve to $\sim 10^{-1}$ and $\sim 10^{-2}$, respectively.

6. CONCLUDING REMARKS

We have extended the search for lensing signatures to all BBH candidates with a probability of astrophysical origin higher than 0.5 from O3b (Abbott et al. 2021b). While we have not observed any significant candidates for strongly lensed events, we updated the constraints on the rate of such events from several different analyses. First, we searched for sub-threshold repeated signals associated with super-threshold events using reduced template banks produced from the posterior probability distributions of the super-threshold events. Interesting sub-threshold/super-threshold pairs and pairs formed from two super-threshold events were further analyzed for their probability of being from a single, strongly lensed source. For super-threshold/super-threshold pairs, we calculated the degree of overlap between the posteriors of the intrinsic parameters and sky location, which were obtained from Bayesian inference. Moreover, we analyzed these pairs using a new analysis based on the comparison of spectrograms through machine learning. Finally, pairs with false-positive probability from either analysis smaller than $10^{-2}$ were further studied by conducting full joint Bayesian inference analyses that take population priors and selection effects into account. We found no pairs that show significant evidence for strong lensing.

The events from O3b were also analyzed for distortions caused by the lens on the gravitational waveform. First, we searched for the distortions that lensing introduces on type II signals, which are in the form of a frequency-independent phase shift (Morse phase). The Bayes factors for all events show no evidence for type II signal distortions. Similarly, we searched for the frequency-dependent distortions caused by point masses. None of the computed Bayes factors show any significant signs of microlensing. For both analyses, some events show interesting features in the posteriors for the Morse phase or lens mass. However, follow-up analyses using simulated signals show no further signs of the lensing nature of these features. Altogether, we found no significant evidence for distortions of the gravitational waveforms that can be attributed to lensing.

The lack of evidence for lensing is then used to infer properties of the lensing rates and to set constraints on the dark matter fraction of (dark) compact objects.

Finally, we note that our conclusions are based on estimates and assumptions that are in line with other analyses from the LIGO–Virgo–KAGRA Collaboration (Abbott et al. 2021f,j). It is possible to arrive at different conclusions and interpretations if assumptions are chosen differently. Examples include claims that almost all detections are strongly lensed if one assumes that heavy BHs do not exist (Broadhurst et al. 2018, 2020a,b). Data from the upcoming observing runs are ex-
pected to further expand the catalog of GW detections that can further shed light on the lensing of GWs (Abbott et al. 2020b). Moreover, multi-messenger astronomy may provide significant input in confirming and interpreting possible lensed GW signals (Wempe et al. 2022).

This material is based upon work supported by NSF’s LIGO Laboratory which is a major facility fully funded by the National Science Foundation. The authors also gratefully acknowledge the support of the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO 600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. The authors gratefully acknowledge the Italian Istituto Nazionale di Fisica Nucleare (INFN), the French Centre National de la Recherche Scientifique (CNRS) and the Netherlands Organization for Scientific Research (NWO), for the construction and operation of the Virgo detector and the creation and support of the EGO consortium. The authors also gratefully acknowledge research support from these agencies as well as by the Council of Scientific and Industrial Research of India, the Department of Science and Technology, India, the Science & Engineering Research Board (SERB), India, the Ministry of Human Resource Development, India, the Spanish Agencia Estatal de Investigación (AEI), the Spanish Ministry of Ciencia e Innovación and Ministerio de Universidades, the Conselleria de Fons Europeus, Universitat i Cultura and the Direcció General de Política Universitaria i Recerca del Govern de les Illes Balears, the Conselleria d’Innovació, Universitats, Ciència i Societat Digital de la Generalitat Valenciana and the CERCA Programme Generalitat de Catalunya, Spain, the National Science Centre of Poland and the European Union – European Regional Development Fund; Foundation for Polish Science (FNP), the Swiss National Science Foundation (SNSF), the Russian Foundation for Basic Research, the Russian Science Foundation, the European Commission, the European Social Funds (ESF), the European Regional Development Funds (ERDF), the Royal Society, the Scottish Funding Council, the Scottish Universities Physics Alliance, the Hungarian Scientific Research Fund (OTKA), the French Lyon Institute of Origins (LIO), the Belgian Fonds de la Recherche Scientifique (FRS-FNRS), Actions de Recherche Concertées (ARC) and Fonds Wetenschappelijk Onderzoek – Vlaanderen (FWO), Belgium, the Paris Île-de-France Region, the National Research, Development and Innovation Office Hungary (NKFIH), the National Research Foundation of Korea, the Natural Science and Engineering Research Council Canada, Canadian Foundation for Innovation (CFI), the Brazilian Ministry of Science, Technology, and Innovations, the International Center for Theoretical Physics South American Institute for Fundamental Research (ICTP-SAIFR), the Research Grants Council of Hong Kong, the National Natural Science Foundation of China (NSFC), the Leverhulme Trust, the Research Corporation, the National Science and Technology Council (NSTC), Taiwan, the United States Department of Energy, and the Kavli Foundation. The authors gratefully acknowledge the support of the NSF, STFC, INFN and CNRS for provision of computational resources.

This work was supported by MEXT, JSPS Leading-edge Research Infrastructure Program, JSPS Grant-in-Aid for specially Promoted Research 26000005, JSPS Grant-in-Aid for Scientific Research on Innovative Areas 2905: JP17H06358, JP17H06361 and JP17H06364, JSPS Core-to-Core Program A. Advanced Research Networks, JSPS Grant-in-Aid for Scientific Research (S) 17H06133 and 20H05639, JSPS Grant-in-Aid for Transformative Research Areas (A) 20A203; JP20H05854, the joint research program of the Institute for Cosmic Ray Research, University of Tokyo, National Research Foundation (NRF), Computing Infrastructure Project of Global Science experimental Data hub Center (GSDC) at KISTI, Korea Astronomy and Space Science Institute (KASI), and Ministry of Science and ICT (MSIT) in Korea, Academia Sinica (AS), AS Grid Center (ASGC) and the National Science and Technology Council (NSTC) in Taiwan under grants including the Rising Star Program and Science Vanguard Research Program, Advanced Technology Center (ATC) of NAOJ, and Mechanical Engineering Center of KEK.

**Software:** Analyses in this paper made use of LALSuite (LIGO Scientific Collaboration and Virgo Collaboration 2018), the GstLAL (Cannon et al. 2012; Messick et al. 2017; Hanna et al. 2020; Sachdev et al. 2019) pipeline; Bayesian inference with bilby (Ashton et al. 2019; Smith et al. 2020; Romero-Shaw et al. 2020); as well as the packages NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), Astropy (Robitaille et al. 2013; Price-Whelan et al. 2018), IPython (Perez & Granger 2007), and ligo.skymap (Singer 2019). Plots were produced with Matplotlib (Hunter 2007), and Seaborn (Waskom et al. 2020).

**APPENDIX**

**A. TYPE II SIMULATION CAMPAIGNS**

Given the mild evidence of GW190412 and GW200129_065458 towards being a type II strongly lensed image presented in Sec. 3.4, we follow up on these events by doing an injection campaign where we simulate type I and type II images similar to the events,
and verify whether the posteriors recovered are compatible with the distribution observed for the real events. These injections are performed in different noise realizations and with different waveform models. The observed feature could be caused by two main other effects than a type II image: noise artifacts or systematic effects in the waveform modeling. In the former, non-Gaussianities in the noise could be such that they lead to the observation of spurious features, while in the latter case, the specific combination of observed parameters could lead to some systematic issues in fitting with the waveform model. Waveform systematics might be especially important for these events since they lie in challenging parts of the parameter space (Abbott et al. 2020c; Colleoni et al. 2021; Hannam et al. 2021). Moreover, for GW200129_065458 Payne et al. (2022) reports that there could be a significant glitch under subtraction.

To test for the noise-related features, we generate colored Gaussian noise from the PSD around the time of the candidate and then inject the maximum likelihood parameters coming from the parameter estimation and take the Morse factor to be either the value for a type I or a type II image. In the first step, we do this for only one noise realization for each event to see whether we can reproduce similar features or not. For GW200129_065458, the injection shows that the effect is too weak to be distinguishable from one image type to the other, as can be seen in the uninformative posteriors of Fig. 7. As a consequence, no further investigation is done into this event. On the other hand, for GW190412, the feature seen in the real data is compatible with the one seen in the injection (see Fig. 8).

Given that the real-data results are compatible with the type II injection for GW190412, we investigate further the noise hypothesis. For this purpose, we take the maximum likelihood parameters and a Morse factor of 0 or 1/2 and inject the signal generated with the IMRPhenomXPHM (Pratten et al. 2021) model in ten different noise realizations. We then repeat the analysis in the same way as for the real signal and verify if we retrieve the same preference for a type II image. For all the noise realizations used here, we see the same behavior as in Fig. 8.

We perform an extra test by injecting the maximum likelihood parameters with a given image type in the generated noise for different waveform models. We use the IMRPhenomPv3HM (Khan et al. 2020) and the SEOBNRv4PHM (Ossokine et al. 2020) model to generate the signal and use the IMRPhenomXPHM (Pratten et al. 2021) model to recover it. This enables us to combine the two possible sources of systematics. This way, we can verify whether a different noise combined with a different model also leads to a preference for type II images. For all the noise realizations and the two models used for the injections, we find that the injections always recover the correct hypothesis, and the fact that the real event supports type II is unlikely to be a result of noise or waveform artifacts, as shown in Fig. 9.

Although these tests do not discard the type II image hypothesis, they cannot conclusively confirm it. To confirm the presence of lensing for this event with a mild preference for a type II image, we would need additional evidence. Therefore, we search for possible sub-threshold counterparts with the methodology explained in Sec. 3.1. However, we find only marginal triggers.

![Figure 7. Posterior distribution of the Morse phase for GW200129_065458. We compare the real event posterior (solid-blue) with an injection campaign of type I (dashed-orange) and type II (dotted-green) images. Type II images correspond to \( n_1 = 1/2 \). For this event, the differences between the distribution are small and make it difficult to learn anything additional about the event. The Kolmogorov–Smirnov statistic is 0.07 for type I vs real, and 0.08 for type II vs real.](image)
Figure 8. Posterior distribution of the Morse phase for GW190412. We compare the real event posterior (solid-blue) with an injection campaign of type I (dashed-orange) and type II (dotted-green) images. Type II images correspond to $n_1 = 1/2$. For this event, the peak seen in the real data and the one seen for the type II image are compatible, hinting at a possible type II image. In this case, the Kolmogorov–Smirnov statistic is 0.20 for type I vs real, and 0.13 for type II vs real.

Figure 9. Comparison of the Morse factor distribution for the real event (solid-purple) with the recovered posterior distribution for an injection made with IMRPhenomPv3HM for a type I image (dashed-blue) and a type II image (dotted-orange), and with an injection made with SEOBNRv4PHM for a type I image (dashed-green) and a type II image (dotted-red). In all the cases, the posterior distributions agree with the injected data, with the real event resembling a type II image.

In the end, these additional searches did not enable us to find any extra evidence for lensing, while still not ruling out the possibility for GW190412 to be a type II image.

B. MARGINALIZED POSTERIORS OF MICROLENSING PARAMETERS

As a supplement to the distribution of $\log_{10}$ Bayes factors $B_{\text{Micro}}^U$ shown in Fig. 4, we show the individual marginalized posterior distributions of redshifted lens mass $M_L^*$ and $\log_{10} B_{\text{Micro}}^U$ (right vertical axis) in Fig. 10. The Bayes factors individually do not show clear evidence for microlensing by point-mass lenses. However, several events show a narrow posterior distribution of the redshifted lens mass. An example is GW200208.130117 (with $\log_{10} B_{\text{Micro}}^U = 0.8$), for which the waveform corresponding to the maximum posterior for this event, with and without lensing, is shown in Fig. 11. The beating pattern introduced by the point-mass lens is most visible as a reduction of the amplitude for two cycles in the middle of the signal and an increase in the amplitude.
before and after this reduction. We hypothesize that short-duration noise fluctuations may have caused an apparent dip in the signal, which in turn may have led to a distortion similar to a point-mass lens beating pattern. This is corroborated by a low Bayes factor $B_M^{\text{Micro}}$, which concludes the data is inconclusive about the microlensing hypothesis.

REFERENCES

—. 2021c, Data release for "Search for gravitational-lensing signatures in the full third observing run of the LIGO-Virgo network". https://dcc.ligo.org/XXXXXX/public
Figure 11. The time-domain waveform corresponding to the maximum posterior of GW200208.130117, with and without the microlensing hypothesis.
Department of Information and Management Systems Engineering, Nagaoka University of Technology, Nagaoka City, Niigata 940-2188, Japan

Tata Institute of Fundamental Research, Mumbai 400005, India

Eindhoven University of Technology, 5600 MB Eindhoven, Netherlands

Department of Physics and Astronomy, Sejong University, Gwangjin-gu, Seoul 143-747, Republic of Korea

Concordia University Wisconsin, Mequon, WI 53097, USA

Department of Electrophysics, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

Department of Physics, Rikkyo University, Toshima-ku, Tokyo 171-8501, Japan

* Deceased, December 2021.