GWTC-4.0: Searches for Gravitational-Wave Lensing Signatures

THE LIGO SCIENTIFIC COLLABORATION, THE VIRGO COLLABORATION, AND THE KAGRA COLLABORATION

(Compiled: 17 December 2025)

ABSTRACT

Gravitational waves can be gravitationally lensed by massive objects along their path. Depending on the lens mass and the lens–source geometry, this can lead to the observation of a single distorted signal or multiple repeated events with the same frequency evolution. We present the results for gravitational-wave lensing searches on the data from the first part of the fourth LIGO–Virgo–KAGRA observing run (O4a). We search for strongly lensed events in the newly acquired data by (1) searching for an overall phase shift present in an image formed at a saddle point of the lens potential, (2) looking for pairs of detected candidates with consistent frequency evolution, and (3) identifying sub-threshold counterpart candidates to the detected signals. Beyond strong lensing, we also look for lensing-induced distortions in all detected signals using an isolated point-mass model. We do not find evidence for strongly lensed gravitational-wave signals and use this result to constrain the rate of detectable strongly lensed events and the merger rate density of binary black holes at high redshift. In the search for single distorted lensed signals, we find one outlier: GW231123_135430, for which we report more detailed investigations. While this event is interesting, the associated waveform uncertainties make its interpretation complicated, and future observations of the populations of binary black holes and of gravitational lenses will help determine the probability that this event could be lensed.

1. INTRODUCTION

Massive astrophysical objects such as galaxies and galaxy clusters curve spacetime around them. Propagating waves are deflected by such massive objects as they pass in their vicinity, a phenomenon known as gravitational lensing. According to the equivalence principle, gravitational lensing does not only affect light but also gravitational-wave (GW) signals (Schneider et al. 1992). Strong gravitational lensing can produce repeated signals from the same source. These repeated signals (referred to as images) can vary in amplitude, arrival time, and phase. Due to the long wavelengths of GW signals detected by ground-based interferometers ($\sim 10^2$–10^3 km), wave-optics effects can also be apparent when the characteristic size of the lens ($GM_{\rm L}/c^2$, with G the gravitational constant and c the speed of light) is comparable to the wavelength. They may distort the waveform in characteristic ways (Takahashi & Nakamura 2003). The specific effects and intensity of gravitational lensing depend on the alignment of the source, lens, and observer as well as the mass of the lens and the GW wavelength (Schneider et al. 1992; Takahashi & Nakamura 2003). An introduction to GW lensing can also be found in Section 5.1.2. of Abac et al. (2025a).

Identifying lensed GW signals presents numerous opportunities for probing new scientific avenues. In cos-

mology, lensed GWs provide an independent method for measuring cosmological parameters (Liao et al. 2017; Hannuksela et al. 2020; Finke et al. 2021; Jana et al. 2023; Balaudo et al. 2023; Narola et al. 2024; Wempe et al. 2024; Jana et al. 2024a), can break the mass-sheet degeneracy (Cremonese et al. 2021; Chen et al. 2024), and help probing large-scale structures (Mukherjee et al. 2020; Savastano et al. 2023; Vujeva et al. 2025) and the nature of dark matter (Tambalo et al. 2023; Jana et al. 2024b; Jung & Shin 2019; Basak et al. 2022; Urrutia & Vaskonen 2021; Barsode et al. 2024). Additionally, repeated lensed GW signals allow for stringent tests of general relativity (GR; Goyal et al. 2021; Ezquiaga & Zumalacárregui 2020; Goyal et al. 2023; Chan et al. 2025a). The improved sky localization of lensed GWs, especially when paired with electromagnetic counterparts, enhances our ability to study their origin and environment (Hannuksela et al. 2020; Uronen et al. 2024), and to provide early warning of mergers (Magare et al. 2023), to give a few examples of the opportunities offered by multi-messenger lensing (Smith et al. 2025).

Continuous effort has been made to identify lensing signatures in GW data (e.g., Hannuksela et al. 2019; Dai et al. 2020), particularly during the third observing run (O3) of the LIGO–Virgo–KAGRA (LVK) net-

work (Abbott et al. 2021, 2023a). Dedicated studies, including the follow-up lensing analyses of O3 events (Janquart et al. 2023b), have been conducted to identify potential lensing candidates. However, no conclusive evidence for lensed GWs has yet been found. With the improving sensitivity of the detectors (Abbott et al. 2020; Capote et al. 2025; Soni et al. 2025; Abac et al. 2025a), the probability of detecting lensed GW signals is also increasing as a typical relative strong lensing rate is $\mathcal{O}(1/1000)$ (Ng et al. 2018; Li et al. 2018; Mukherjee et al. 2021a; Wierda et al. 2021; Xu et al. 2022). In this work, we present searches for GW lensing in the data collected during the first part of the fourth observing run (O4a; Abac et al. 2025b) and corresponding to new detections reported in GWTC-4.0 (Abac et al. 2025c).

This paper is structured as follows. In Section 2, we summarize the GW data collected during O4a, and explain the selection of compact binary coalescence (CBC) events analyzed in this work. Section 3 describes the analysis framework used to search for gravitational-lensing signatures, including pairwise analyses for strong-lensing and single-event analyses for strongly-lensed type II images and wave-optics effects. Section 4 presents the results from pair-wise strong lensing searches, while Section 5 details the search for lensing in single signals. Section 6 reports on detailed investigations for GW231123_135430 (hereafter referred to as GW231123) identified as an outlier for our analyses in Section 5 and Abac et al. (2025d). In Section 7, we interpret the implications of our results in the context of astrophysical strong lensing rates and merger rate density. Finally, Section 8 provides a summary of our findings and outlines prospects for future observing runs.

2. DATA AND EVENTS

GWTC-4.0 (Abac et al. 2025c) is a cumulative catalog of GW detections to date, covering the first three observing runs and O4a. It contains 128 new O4a candidates with a probability of astrophysical origin $p_{astro} > 0.5$ and satisfactory event-validation checks (Abac et al. 2025e). O4a ran from 2023 May 24 until 2024 January 16 with the two LIGO detectors (Capote et al. 2024; Soni et al. 2025) taking data. Of the O4a candidates, 86 pass a false alarm rate (FAR) $< 1\,\mathrm{yr}^{-1}$ threshold. The candidates were identified in offline searches by different search pipelines detailed in Abac et al. (2025e). The majority of these candidates are binary black holes (BBHs). In that period, 2 candidates, GW230529_181500 (Abac et al. 2024) and GW230518_125908, are consistent with sources containing a secondary mass $m_2 < 3M_\odot$, making them likely neutron star–black hole binary (NSBH) candidates. Candidates passing these BBH and FAR

criteria have their source properties characterized using parameter estimation (Abac et al. 2025e). Here, we mostly rely on parameter-estimation results obtained using IMRPhenomXPHM-SpinTaylor (Pratten et al. 2021; Colleoni et al. 2025) as base reference unlensed waveform, unless otherwise specified.

The analyses presented in this paper consider only candidates from O4a with FAR $< 1\,\mathrm{yr}^{-1}$ which are identified as BBHs, corresponding to a total of 84 new candidates. We do not include candidates from previous observing runs (Hannuksela et al. 2019; Abbott et al. 2021, 2023a) in our searches due to the computational cost, and because the time gap between the end of third observing run and the start of the fourth is too long to expect identifying strongly-lensed images spread between the two (Wierda et al. 2021; Çalışkan et al. 2023), even if strong lensing by galaxy clusters could produce such time delays in some scenarios (Smith et al. 2018; Robertson et al. 2020; Ryczanowski et al. 2020). We do not include non-BBH candidates as their lensing probability is low and GW230529_181500 (Abac et al. 2024) has been previously investigated (Janquart et al. 2025).

3. DATA-ANALYSIS FRAMEWORK

We use multiple analysis methods, each looking for specific lensing signatures. The communication between each part of the analyses is automated in the LensingFlow framework (Wright et al. 2025b), which relies on Asimov (Williams et al. 2023) and CBCFlow (Udall et al. 2024). Parameter-estimation-based frameworks (Lo & Magana Hernandez 2023; Janquart et al. 2021b, 2023a; Wright & Hendry 2021) rely on Bilby (Ashton et al. 2019) with the Dynesty sampler (Speagle 2020). For such analyses, both the power spectral density (PSD) and the priors on the source parameters were chosen to be consistent with the unlensed investigations carried out in Abac et al. (2025c).

As explained in Section 5.10 of Abac et al. (2025e), in the late preparation of this work, a normalization error in the likelihood used for inference was found. Depending on the analysis requirements, different approaches (reweighting older posteriors or rerunning the corrected code) are used to mitigate this issue. This is detailed in the relevant sections. In the future, we will update the results affected by this error for all analyzed candidates affected by this issue.

3.1. *Searches for Multiple Images*

If the lens is massive enough, lensing effects can be described in the geometric-optics approximation. If the source is well-aligned with the lens, the GW is split into multiple distinct images with the same frequency evolu-

tion. Due to the limited angular resolution of GW detector networks, they will appear to be originating from the same region in the sky. These images are delayed in time, (de)magnified, and can undergo an overall phase shift. So, the lensed waveform for image j (h_L^j) is linked to the unlensed waveform (h_U) as (Schneider et al. 1992)

$$\tilde{h}_L^j(f; \boldsymbol{\theta}, \boldsymbol{\phi}_j) = \sqrt{\mu_j}\, \tilde{h}_U(f; \boldsymbol{\theta})\, e^{-2i\pi f t_j + i\pi n_j \mathrm{sign}(f)}, \quad (1)$$

where $\boldsymbol{\theta}$ represents the usual BBH parameters, and $\boldsymbol{\phi}_j = \{\mu_j, t_j, n_j\}$ the lensing parameters, with μ_j the magnification, t_j the time delay, and n_j the Morse factor (Schneider et al. 1992). To get this relation, we follow the Fourier transform convention presented in Appendix B of Abac et al. (2025a). The Morse factor can only take one of the three following discrete values $\{0, 0.5, 1\}$, corresponding to an image forming at a minimum, a saddle point, or a maximum of the lensing time delay, referred to as type I, II, and III images, respectively. This additional phase shift can lead to detectable features in GW signals, provided the image is of type II and the signal has signatures of precession, higher-order multipole moments, and/or eccentricity (Dai & Venumadhav 2017; Ezquiaga et al. 2021; Wang et al. 2021; Janquart et al. 2021b; Vijaykumar et al. 2023).

We look for pairs of signals with matching characteristics. We do this for pairs of super-threshold candidates, i.e., new GWTC-4.0 BBH candidates with FAR $< 1\,\mathrm{yr}^{-1}$ (Section 3.1.1). Since lower magnification and changes in observing conditions can lead to images being below the detection threshold, we also look for counterpart sub-threshold images to the super-threshold candidates in the data (Section 3.1.2).

3.1.1. *Super-Threshold Image Pairs*

When doing pair-wise image searches for strong lensing, one looks for pairs of events with matching time–frequency evolution. This means that parameters unaffected by lensing, such as the detector-frame masses or the spins, should be similar.

Our workflow is split in multiple stages, referred to as tiers, with increasing computational cost and accuracy at each step. It starts with two tier-1 analyses, probing all the pairs made from the O4a BBH candidates, corresponding to a total of 3486 unique pairs. These are:

- POSTERIOR OVERLAP (PO; Haris et al. 2018), which looks at the consistency between the inferred properties of the two events for a subset of parameters (detector-frame component masses (Abac et al. 2025e), spin amplitudes and tilts, and the sky localization), as encoded in a

ranking statistic B_U^L, combined with another ranking statistic $\mathcal{R}_U^L$ based on the expected time-delay distribution for galaxy lensing (Haris et al. 2018; More & More 2022);

- PHAZAP (Ezquiaga et al. 2023), which looks at the consistency of the reconstructed phases of the $\ell = m = 2$ mode of the GW signal at a specific frequency of 40 Hz, the arrival time between detectors, and the frequency evolution of the phases.

Each of these methods produces a specific ranking statistic that can be used to assess whether the event pair under consideration is sufficiently significant to be further analyzed; see Appendix A for more detail. To make the comparison between different frameworks easier, we define a common false-positive probability (FPP)[1] which can be computed based on an unlensed background distribution by taking the fraction of unlensed pairs with a ranking statistic higher than for the real pair under consideration. Information about the background used in this work can be found below, and more details are given in Appendix B.

Pairs with a FPP below 1% are deemed interesting and passed to the next step of the workflow, the tier-2 analysis code FAST-GOLUM (Janquart et al. 2021a, 2023a). It evaluates the joint likelihood using samples obtained when analyzing the first image and linking them to the second image by sampling the relative magnification, time delay between the two images and the relative Morse factor difference. This recasting of joint parameter estimation (JPE) is then faster than full JPE in the next step since it is already focused on the region of interest, and is further sped up by using a look-up table (Janquart et al. 2021a, 2023a). The FAST-GOLUM analysis is done using the IMRPHENOMXPHM-SPINTAYLOR waveform (Pratten et al. 2021; Colleoni et al. 2025) to which we add the lensing effects as described in Eq. (1), and a uniform prior between 0.1 and 50 for the relative magnification, a uniform prior between $[t_{ij} - 0.2, t_{ij} + 0.2]$ s for the time delay (where t_{ij} is the difference in trigger times for the two events) and a discrete prior evenly distributed between the values $\{0, 0.5, 1, 1.5\}$ for the Morse factor difference. Priors on these parameters are chosen to cover a broad range of astrophysical lensing scenarios. Pairs are ranked using their coherence ratio

---

[1] Similarly to Abbott et al. (2023a), we use FPP for significance associated with the lensing hypothesis (as opposed to the unlensed hypothesis), while FAR is associated with the significance assigned to individual candidate GW signal events (as opposed to them being noise events).

$\mathcal{C}_{\mathrm{U}}^{\mathrm{L}} = p(d_1, d_2 | \mathrm{L}) / \big(p(d_1 | \mathrm{U}) p(d_2 | \mathrm{U}) \big)$, where d_i correspond to the data for the ith image and L (U) refers to the lensed (unlensed) hypothesis. So, the coherence ratio corresponds to the ratio between the lensed and unlensed evidence for each GW signal (Lo & Magana Hernandez 2023; Janquart et al. 2021a). If $\log_{10} \mathcal{C}_{\mathrm{U}}^{\mathrm{L}} > 0$, the pair is passed to the next step.

For complete JPE (Liu et al. 2021; Lo & Magana Hernandez 2023; Janquart et al. 2021a, 2023a), corresponding to tier 3, we compute the Bayes factor using HANABI (Lo & Magana Hernandez 2023) and the IMRPHENOMXPHM-SPINTAYLOR waveform (Pratten et al. 2021; Colleoni et al. 2025). It calculates the lensed evidence for the event pair by doing a joint analysis of the two data segments under the lensed hypothesis. The two signals are analyzed together with the same BBH parameters and assuming they differ only through the lensing effect, with a change in amplitude, a time delay and an overall phase shift. HANABI also reweighs the evidence to account for an astrophysical distribution of BBH and lens parameters using the maximum-likelihood estimation of the GWCT-4.0 population (Abac et al. 2025f) and the same singular isothermal sphere (SIS) model as in Abbott et al. (2023a), respectively. It also includes selection effects (Lo & Magana Hernandez 2023). We refer to the final quantity including population models and selections effects as the strong lensing Bayes factor, following Lo & Magana Hernandez (2023) and Abbott et al. (2023a).

3.1.2. *Sub-Threshold Image Search*

For a given super-threshold image, we can look for possible sub-threshold counterparts. To dig out such faint signals, one decreases the trials factor of the search by constructing a template bank containing a reduced number of waveforms. This is done by generating waveforms resembling the main signal (McIsaac et al. 2020; Li et al. 2023; Dai et al. 2020; Li et al. 2025a).

Sub-threshold searches are done with two approaches. One relies on PYCBC (Dal Canton et al. 2014; Usman et al. 2016; Nitz et al. 2017) to search for counterparts of each super-threshold event with a single template (McIsaac et al. 2020), which is chosen here as the aligned-spin projection of a maximum-posterior estimate from the IMRPHENOMXPHM-SPINTAYLOR results. Another method, TESLA-X, relies on GST-LAL (Messick et al. 2017; Cannon et al. 2021) to do the search and construct a targeted template bank based on a dedicated injection campaign (Li et al. 2025a, 2023). TESLA-X results are only presented for GW231123, as detailed in Section 6. Both methods matched-filter the HL-coincident times in the full O4a data set using the

SEOBNRv4 waveform model (Bohé et al. 2017). Further, to reduce the risk of spurious triggers, they veto triggers that do not overlap in 90% credible sky localization (i.e., we require $\mathcal{O}_{90\% \text{ CI}} > 0$) with their targeted super-threshold candidate (Wong et al. 2021). Pairs made of these triggers and their corresponding super-threshold event need to be followed up with the same methods as super-threshold pairs to assess the lensing hypothesis (Section 3.1.1).

3.2. *Searches for Single Distorted Lensed Signals*

We also report the results of analyses searching for single distorted signals among the O4a BBH candidates. One is a search for type II images produced by strong lensing in the geometric-optics limit. The second is an analysis in the wave-optics regime using an isolated point mass lens model. The third is a targeted phenomenological search for lensing near a fold caustic, which is applied only for specific candidates. In some cases, lensing-induced distortions can lead to the non-detection of signals with unlensed searches, leading to a selection effect (Chan et al. 2025b). This is not considered here.

3.2.1. *Type II Image Searches*

Type II image searches are done with the GOLUM framework (Janquart et al. 2021a, 2023a) as in Abbott et al. (2023a). Even though such signals are expected to be rare in current detections (Abac et al. 2025c), we still search for such signatures as they could be unambiguous evidence for lensing (Ezquiaga et al. 2021; Wong et al. 2021; Janquart et al. 2021b; Vijaykumar et al. 2023) or mimic beyond-GR effects (Mishra et al. 2023; Gupta et al. 2024; Wright et al. 2024; Narayan et al. 2024). Finding such a signal would imply the presence of other images. However, those may arrive at periods when detectors are offline or could be too faint to be detectable even by our sub-threshold searches.

3.2.2. *Searches for Wave-Optics Effects by Point Mass Lenses*

When the gravitational radius of the lens is comparable to the wavelength of the GW (wave-optics regime), lensing leads to frequency-dependent beating patterns in the signal. This can be modeled by an amplification factor $F(f; \boldsymbol{\phi})$ modifying the unlensed waveform (Takahashi & Nakamura 2003)

$$\tilde{h}_{\mathrm{L}}(f; \boldsymbol{\theta}, \boldsymbol{\phi}) = F(f; \boldsymbol{\phi}) \, \tilde{h}_{\mathrm{U}}(f; \boldsymbol{\theta}), \qquad (2)$$

where $\boldsymbol{\theta}$ represents the usual binary parameters, and $\boldsymbol{\phi}$ the lensing parameters. The latter can be phenomenological or depend on the lensing model which is as-

sumed. For the simplest mass profiles, like the isolated point-mass model used in this work (Schneider et al. 1992), it corresponds to the redshifted lens mass $M_{\rm L}^z = (1 + z_{\rm L})M_{\rm L}$, where $z_{\rm L}$ is the redshift of the lens and $M_{\rm L}$ its mass, and the impact parameter y. This model is chosen as it is expected to appropriately represent isolated compact lenses, ranging from isolated stars to black holes. We compare the Bayes factor $\mathcal{B}_{\rm U}^{\rm Mod}$ with those obtained for an unlensed background to assess the related FPP since imperfections in the data and models could lead to spurious favoring of the lensing hypothesis.

This analysis is done using the Gravelamps framework (Wright & Hendry 2021), which incorporates lens models into Bilby (Ashton et al. 2019). We use an updated framework compared to previous work (Abbott et al. 2021, 2023a; Janquart et al. 2023b); an inconsistency in the Fourier transform convention was identified in the results obtained from past observing runs (Wright et al. 2025a). Updated results for past observing runs will be reported in separate work.

3.2.3. *Phenomenological Searches for Single Distorted Signals*

The transition from wave- to geometric-optics effects can be seen as a continuous change from one regime to another. Therefore, for isolated lenses with larger masses, there is a regime where the lensed GW signal can be modeled using a superposition of multiple images, each described by Eq. (1) but where the time delay is at most a few tens of milliseconds. This leads to the observation of a single distorted signal, which can be modeled as a sum of all images, leading to a modification of the amplification factor in Eq. (2) as (Liu et al. 2023):

$$F(f; \boldsymbol{\phi}) = \sum_{j=1}^{K} \sqrt{\mu_j} \; e^{-2i\pi f t_j + i\pi n_j \,{\rm sign}(f)} , \qquad (3)$$

where $\boldsymbol{\theta}$ represents the usual binary parameters, and $\boldsymbol{\phi} = \{\boldsymbol{\phi}_j\}_{j=1,\dots,K}$ are the phenomenological lens parameters (magnification, time delay and Morse factor for each image), with K the total number of images formed.

A specific case of this kind of analysis, which is the one considered in this work, is that of lensing near a fold caustic, where one expects two lensed images ($K = 2$) with the same (large) magnification, opposite parities, and a millisecond time delay (Schneider et al. 1992; Lo et al. 2025; Ezquiaga et al. 2025). We do not perform this analysis on all events, but consider it in more detail when investigating GW231123 (see Section 6.2).

3.3. *Unlensed Background*

We often compare the statistics obtained for the lensed versus unlensed hypothesis for each real event with that obtained for an unlensed background. This is done to compute FPP values while accounting for fluctuations due to the noise present. Most of the lensing models have extra degrees of freedom compared to the unlensed one, potentially leading to a larger evidence for the lensed hypothesis even when the signal is not lensed but when additional features, like non-Gaussian noise fluctuations, are present.

Our background is constructed using real stretches of O4a data devoid of detected GW signals. We add an unlensed signal to the data and follow the usual analysis approaches (Abac et al. 2025e). We use 254 unlensed signals with masses sampled from the Power-Law + Peak mass distribution, their spin magnitude from a Beta distribution, and their spin tilts from an isotropic plus truncated half-Gaussian mixture (Abbott et al. 2023b). All other parameters are drawn from their usual priors (Abac et al. 2025e). Injection and parameter estimation are done using the IMRPhenomXPHM-SpinTaylor waveform (Pratten et al. 2021; Colleoni et al. 2025), and parameter estimation follows the usual approaches (Abac et al. 2025e). Details about the background construction and analysis can be found in Appendix B.

4. SEARCHES FOR MULTIPLE IMAGES

In this section, we report the results from searches for multiple images produced by strong lensing (Section 3.1).

4.1. *Super-Threshold Pairs*

The strong-lensing analyses are summarized in Figure 1, where the flowchart also reports the number of pairs analyzed and seen as interesting at the various steps in our workflow. These steps are further detailed below. Of the total 3486 initial pairs from all O4a events, only 50 were passed to our tier-3 analyses.

4.1.1. *Identification of Pairs with Tier-1 Analyses*

Here, we report the results found with PO (Haris et al. 2018) and Phazap (Ezquiaga et al. 2023), described in Section 3.1.1. To account for the Bilby likelihood issue reported in Sec. 5.10 of Abac et al. (2025e), we run these analyses on reweighted posteriors since changes are not expected to be large.

Figure 2 shows the FPP found for the two analyses run on all O4a events. Using a FPP threshold of 1% for the two tier-1 approaches, we find a total of 105 pairs that need to be followed up by the tier-2 analysis, representing a 97% reduction in the number of candidates. From all the pairs seen as interesting, there are 5 that are common to both analyses, 33 flagged only by posterior
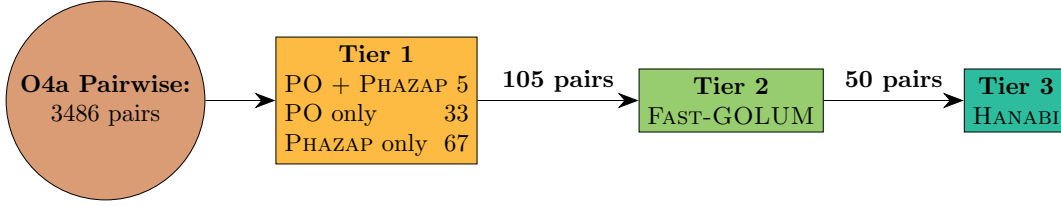
Figure 1. Flowchart showing the progression of the pairwise analyses and the numbers of pairs that passed each stage of the analysis process. With this approach, only a reduced number of pairs needs to be analyzed with JPE, significantly reducing the computational cost.
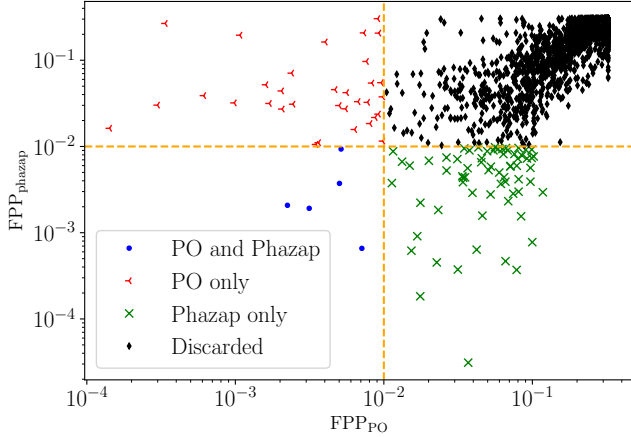
Figure 2. False-positive probability found for all O4a event pairs for POSTERIOR OVERLAP (PO) and PHAZAP. The dashed orange line shows the 1% threshold used to select pairs for follow-up. We pass to the second tier all pairs with an FPP below this threshold for at least one of the two approaches.

overlap (PO) and 67 found only by PHAZAP. Details on the statistics used to select the pairs and compute their FPPs can be found in Appendix A.1 and A.2.

4.1.2. Tier-2 Analysis

For this analysis, the runs are done using a likelihood fixed for the issue reported in Section 5.10 of Abac et al. (2025e). Out of the 105 pairs passed to FAST-GOLUM (Janquart et al. 2021a, 2023a), 55 pairs ($\sim$ 52%) were discarded, leading to 50 event pairs being passed to the tier-3 analyses. Details on the analysis and the results are given in Appendix A.3.

4.1.3. Tier-3 Analysis Results

For HANABI, analyses are done using a likelihood fixed for the issue reported in Sec. 5.10 of Abac et al. (2025e). Out of the 3486 initial O4a pairs, 50 were followed up by HANABI (Lo & Magana Hernandez 2023). HANABI finds that none of the event pairs shows a preference for the strong-lensing hypothesis over the null hypothesis that the events in a pair are unrelated.

Since the strong-lensing Bayes factors are sensitive to assumptions about the population of the sources, we repeat the same Bayes factor calculation using three different models for the merger rate density, namely one that tracks the star formation rate from Madau & Dickinson (2014), as well as $R_{\min}(z)$ and $R_{\max}(z)$ that are described in Abbott et al. (2023a), respectively. The corresponding markers in Figure 3 show the Bayes factors $\mathcal{B}_{\mathrm{U}}^{\mathrm{L}}$ computed using these merger rate density models. We see that our conclusion that there is no evidence for strongly lensed GW signals within the 50 event pairs passed to the analysis is robust against differences in the assumed merger rate density.

The analysis assumes all of the GW events in those 50 pairs have astrophysical origins. This might not be the case, and our assessment that an event is astrophysical in origin is encoded in p_{astro}, reported in Abac et al. (2025c). To take this into account, we color each marker in Figure 3 with $p_{\mathrm{astro}}^{\mathrm{pair}}$, which is the product of p_{astro} for each event in a pair. All of the pairs we considered in the analysis have $p_{\mathrm{astro}}^{\mathrm{pair}} > 0.9$.

4.2. Searches for Sub-Threshold Counterparts

Searches for sub-threshold candidates potentially matching known GWTC-4.0 events were performed over all O4a strain data following the rules for data selection described in Abac et al. (2025e) and using the PYCBC-based targeted search algorithm (McIsaac et al. 2020). For all sub-threshold searches, the templates are generated based on the posterior samples obtained prior to reweighing for the likelihood issue. This is not expected to change the results substantially; updated results are left as future work.

In Table 1, we list all triggers from the PYCBC targeted searches for counterparts of the 84 super-threshold targets. These have a single-template FAR below $1\,\mathrm{yr}^{-1}$ and a non-zero sky overlap. We have also removed those corresponding to GWTC-4.0 events with FAR $< 1\,\mathrm{yr}^{-1}$ and $p_{\mathrm{astro}} > 0.5$. Due to the high density of GW events in some regions of parameter space, many triggers are recovered multiple times by searches for different super-threshold target events. The table includes only the

Sub-Threshold Trigger (UTC) (yy:mm:dd hh:mm:ss)	Target Super-Threshold Event GWYYMMDD_HHMMSS	Time delay	Network SNR	FAR (yr^{-1})	$\mathcal{O}_{90\% \text{ CI}}$ (%)	$\mathcal{R}_{\mathrm{U}}^{\mathrm{L}}$
2023-09-03 07:43:35	GW231104_133418	$62^{\mathrm{d}}05^{\mathrm{h}}50^{\mathrm{m}}44^{\mathrm{s}}$	7.88	5.6×10^{-6}	0.11	0.31
2023-10-30 12:34:06	GW231223_075055	$53^{\mathrm{d}}19^{\mathrm{h}}16^{\mathrm{m}}49^{\mathrm{s}}$	7.99	9.6×10^{-6}	52	0.41
2023-11-20 02:21:03*	GW230904_051013	$76^{\mathrm{d}}21^{\mathrm{h}}10^{\mathrm{m}}50^{\mathrm{s}}$	9.62	2.5×10^{-3}	7.1	0.21
2023-10-02 14:39:16*	GW230824_033047	$39^{\mathrm{d}}11^{\mathrm{h}}08^{\mathrm{m}}29^{\mathrm{s}}$	9.34	0.018	14	0.68
2023-08-13 14:49:14*	GW231226_101520	$134^{\mathrm{d}}19^{\mathrm{h}}26^{\mathrm{m}}06^{\mathrm{s}}$	8.02	0.17	2.6	0.08
2023-12-31 12:01:47*	GW230831_015414	$122^{\mathrm{d}}10^{\mathrm{h}}07^{\mathrm{m}}34^{\mathrm{s}}$	9.05	0.21	6	0.09
2023-11-10 17:17:31*	GW231001_140220	$40^{\mathrm{d}}03^{\mathrm{h}}15^{\mathrm{m}}11^{\mathrm{s}}$	8.39	0.25	14	0.66
2023-12-20 17:34:06*	GW230803_033412	$139^{\mathrm{d}}13^{\mathrm{h}}59^{\mathrm{m}}54^{\mathrm{s}}$	7.60	0.33	22	0.07
2023-10-07 18:27:39	GW231118_071402	$41^{\mathrm{d}}12^{\mathrm{h}}46^{\mathrm{m}}22^{\mathrm{s}}$	12.76	0.41	7	0.62
2023-12-31 10:55:39*	GW230928_215827	$93^{\mathrm{d}}12^{\mathrm{h}}57^{\mathrm{m}}13^{\mathrm{s}}$	7.92	0.42	16	0.15
2023-10-26 13:07:04*	GW230927_043729	$29^{\mathrm{d}}08^{\mathrm{h}}29^{\mathrm{m}}35^{\mathrm{s}}$	7.97	0.62	4.9	1.02
2023-08-22 23:03:37*	GW230630_125806	$53^{\mathrm{d}}10^{\mathrm{h}}05^{\mathrm{m}}31^{\mathrm{s}}$	8.17	0.51	6	0.41
2023-08-17 21:23:49*	GW230726_002940	$22^{\mathrm{d}}20^{\mathrm{h}}54^{\mathrm{m}}09^{\mathrm{s}}$	7.75	0.54	71	1.37
2023-09-26 08:45:30	GW231001_140220	$5^{\mathrm{d}}05^{\mathrm{h}}16^{\mathrm{m}}50^{\mathrm{s}}$	7.91	0.63	45	5.04
2023-08-19 01:22:51*	GW231104_133418	$77^{\mathrm{d}}12^{\mathrm{h}}11^{\mathrm{m}}27^{\mathrm{s}}$	8.06	0.64	9.5	0.21
2023-07-03 11:41:11*	GW230814_061920	$41^{\mathrm{d}}18^{\mathrm{h}}38^{\mathrm{m}}09^{\mathrm{s}}$	8.27	0.67	20	0.62
2023-11-26 01:09:28*	GW231213_111417	$17^{\mathrm{d}}10^{\mathrm{h}}04^{\mathrm{m}}50^{\mathrm{s}}$	8.75	0.68	4.9	1.85
2023-08-30 04:32:38*	GW231129_081745	$91^{\mathrm{d}}03^{\mathrm{h}}45^{\mathrm{m}}07^{\mathrm{s}}$	8.01	0.81	11	0.15
2023-06-25 21:16:55*	GW231008_142521	$104^{\mathrm{d}}17^{\mathrm{h}}08^{\mathrm{m}}27^{\mathrm{s}}$	7.93	0.82	64	0.12
2023-09-04 04:12:45*	GW240107_013215	$124^{\mathrm{d}}21^{\mathrm{h}}19^{\mathrm{m}}31^{\mathrm{s}}$	7.67	0.93	62	0.09
2023-06-09 01:08:24*	GW230608_205047	$0^{\mathrm{d}}04^{\mathrm{h}}17^{\mathrm{m}}37^{\mathrm{s}}$	7.61	0.96	1.6	31.42
2023-09-02 22:45:55*	GW231119_075248	$77^{\mathrm{d}}09^{\mathrm{h}}06^{\mathrm{m}}52^{\mathrm{s}}$	7.69	0.99	15	0.21

Table 1. The top 22 triggers found by the PyCBC sub-threshold searches for the 84 target super-threshold events, in terms of lowest single-template FAR below $1\,\mathrm{yr}^{-1}$ and requiring sky overlap with the target event of $\mathcal{O}_{90\% \text{ CI}} > 0$. The columns are: the date and time of the sub-threshold trigger, the target event name from GWTC-4.0, the time delay between the two, the SNR and FAR of the new trigger in the single-template targeted search, the sky localization overlap statistic $\mathcal{O}_{90\% \text{ CI}}$, and the time-delay ranking statistic $\mathcal{R}_{\mathrm{U}}^{\mathrm{L}}$. Many of these triggers are found in multiple targeted searches. For these, we have only listed the lowest-FAR result with non-zero sky overlap, with the full set of triggers for all target events included in the data release for this paper. Triggers marked by an asterisk already appear in the full GWTC-4.0 candidate list, though they do not pass the FAR $< 1\,\mathrm{yr}^{-1}$ and $p_{\mathrm{astro}} > 0.5$ criteria.

lowest-FAR pairing for such cases. Results of searches focusing on GW231123 are reported separately in Section 6, and none of the PyCBC triggers for that event made this overall top list. An extended list of all triggers from all searches below a FAR of $12\,\mathrm{yr}^{-1}$ is included in the data release for this paper.

Most of the triggers in Table 1 and 69 out of the 177 unique triggers in the data release already appear in the full GWTC-4.0 candidate list, though they correspond to triggers not satisfying the FAR $< 1\,\mathrm{yr}^{-1}$ and $p_{\mathrm{astro}} > 0.5$ criteria, and generally do not have full data-quality and parameter-estimation information available. On the other hand, the two lowest-FAR triggers, found on 2023-09-03 07:43:35 (possible counterpart to GW231104_133418) and 2023-10-30 12:34:06 (possible counterpart to GW231223_075055), are not included in GWTC-4.0. Both appear with low FAR in only this targeted search.

As in previous studies (e.g., Abbott et al. 2023a), we do not perform in-depth data-quality checks for possible instrumental origins of any of these triggers. The reported FARs also do not indicate how likely each trigger is an actual lensed counterpart of the target signal, but only how likely noise produces a trigger with a ranking statistic higher or equal to the one under consideration using these single-template searches. For our two top triggers, we ran parameter estimation with a setup similar to the one used for online candidates and assessed the probability of them being lensed counterparts to their corresponding target super-threshold event, using PO and PHAZAP. We found that neither of these two pairs match each other well enough to pass the threshold from Section 4.1.1, and hence they are not interesting enough to be passed to further analyses.

More in-depth analyses could be carried out to select candidates of interest from this data release and
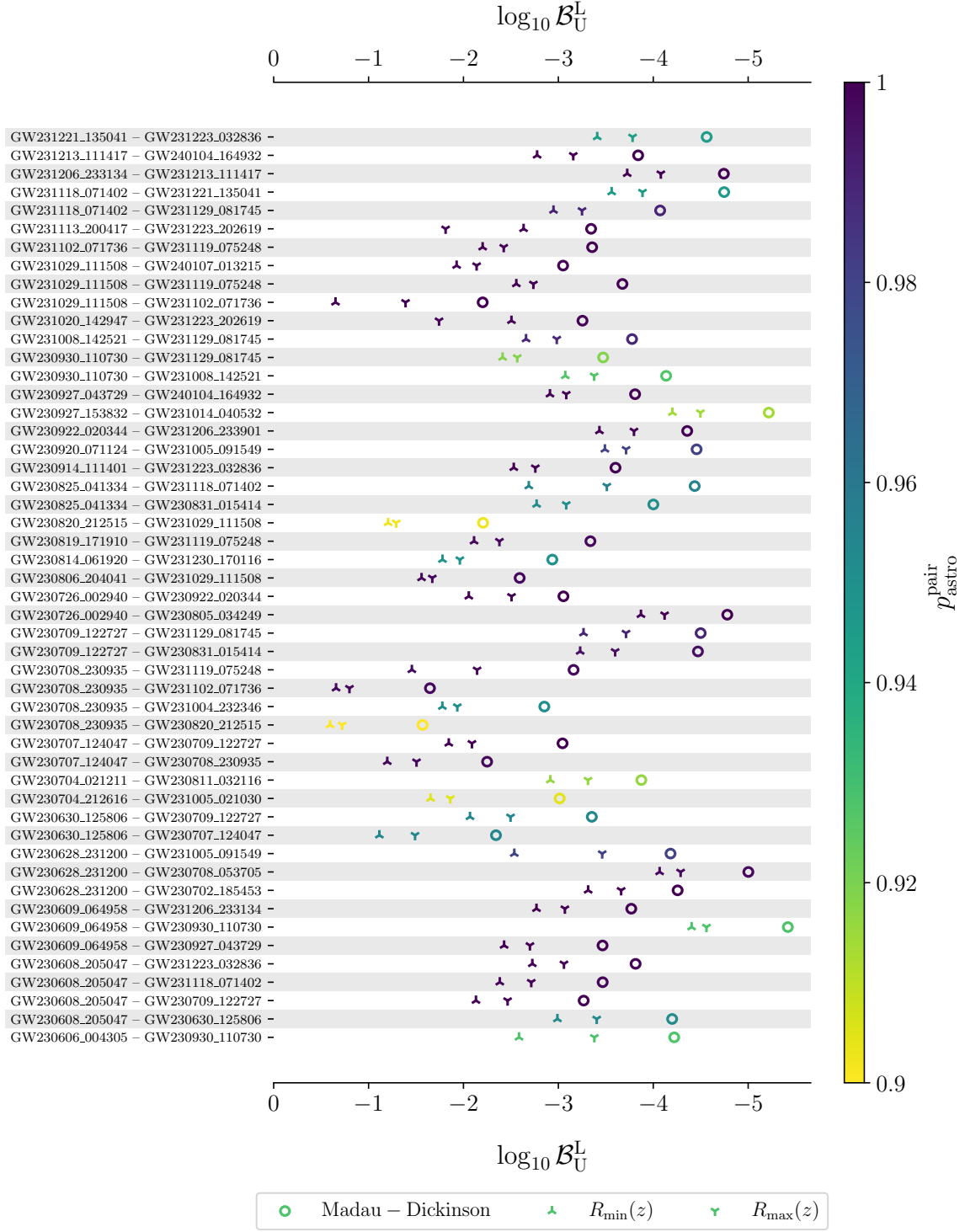
Figure 3. Bayes factors $\mathcal{B}_U^L$ comparing the lensed and unlensed hypotheses for the 50 event pairs passed to the HANABI analysis, computed using three different merger rate density models, namely one that tracks the star formation rate from Madau & Dickinson (2014), $R_{\min}(z)$ and $R_{\max}(z)$ that are described in Abbott et al. (2023a), respectively. The color of each marker indicates the value of $p_{\mathrm{astro}}^{\mathrm{pair}}$, which is the product of the probabilities that each event in a pair has an astrophysical origin as reported in Abac et al. (2025c). None of the event pairs shows a preference for the strong-lensing hypothesis over the null hypothesis that the events in a pair are unrelated.
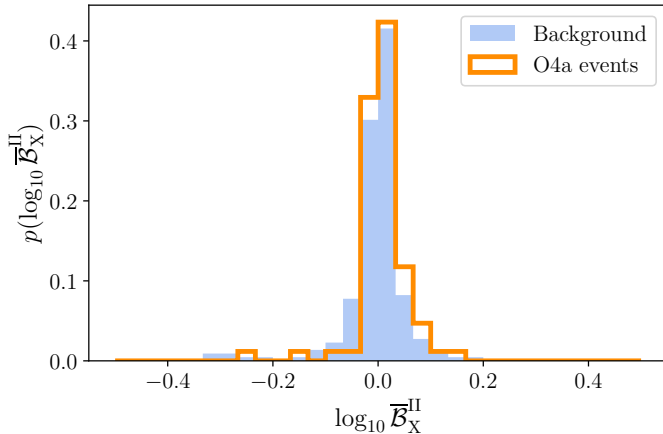
Figure 4. Distribution of averaged Bayes factors for the type II versus type I or III hypotheses obtained for real data and the unlensed background. No significant outlier candidates are identified.

study how much of the observed excess is due to noise artifacts, unlensed weak GW signals, or actual lensed counterparts. To facilitate such studies (Ng et al. 2024), we provide two additional statistics of interest for each super-threshold–sub-threshold pair: the sky localization overlap statistic $\mathcal{O}_{90\% \text{ CI}}$ (Wong et al. 2021) and the time-delay ranking statistic $\mathcal{R}_{\text{U}}^{\text{L}}$ based on galaxy-lensing priors (Haris et al. 2018; More & More 2022).

5. SEARCHES FOR SINGLE DISTORTED SIGNALS

5.1. *Searches for Type II Images*

Here, we report on the analyses looking for type II images detailed in Section 3.2.1. Reported results are reweighted to account for the issue with the likelihood previously explained. Figure 4 shows the distribution of averaged Bayes factors comparing the probability of a type II image as opposed to another image type, i.e., we take the mean value of the Bayes factor comparing type II vs type I images and that comparing type II vs type III images. This statistic is used because one cannot distinguish the difference between type I and III images, and any difference in Bayes factors is only due to sampling effects. Our results are overlaid with the distribution of Bayes factors from the unlensed background described in Section 3.3. We find no outliers, meaning that no candidate shows support for the type II image hypothesis beyond what can be expected from noise fluctuations.

5.2. *Point-Mass Lens Model Analyses*

All O4a events were analyzed using the isolated point-mass model for the amplification (Wright & Hendry 2021), and the IMRPHENOMXPHM-SPINTAYLOR waveform model (Pratten et al. 2021; Colleoni et al.

2025). We use a uniform prior between $1\,M_\odot$ and $10^4\,M_\odot$ for the redshifted lens mass (M_{L}^z), while the prior for the dimensionless source position (y) is $\propto y$ with $y \in [0.1, 3]$. The mass range corresponds to lens masses from stars $(\sim 1 M_\odot)$ to intermediate-mass black holes $(\sim 10^4 M_\odot)$, beyond which the time delay between images becomes large enough to resolve multiple signals.

To handle the likelihood issue mentioned in Sec. 3, for all GRAVELAMPS analyses and complementary unlensed analyses, results were obtained by reweighing the nested sampling chains rather than the posterior distributions. This leads to more robust calculation of the evidence.

For each event, we calculate a Bayes factor $\mathcal{B}_{\text{U}}^{\text{Mod}}$, corresponding to the evidence ratio between the two hypotheses. The resulting distribution is shown in Figure 5. To assess statistical fluctuations in the Bayes factor for unlensed signals, we also show the values of $\mathcal{B}_{\text{U}}^{\text{Mod}}$ for the astrophysical background described in Section 3.3 and Appendix B. The distribution of Bayes factors for O4a events is comparable to that of the background except for GW231123, which has $\log_{10}\mathcal{B}_{\text{U}}^{\text{Mod}} = 3.8$. This is the highest value measured so far (Abac et al. 2025d) and will be analyzed in more detail in Section 6. Moreover, since this value is outside of our astrophysical background, we can place an upper bound on its FPP of 0.39%, where the statistic is defined as the ratio between unlensed signals with a higher Bayes factor to all unlensed events in the background and does not include the trials factor. The p-value including trials factor is given by

$$p = 1 - (1 - \text{FPP})^N, \tag{4}$$

where N is the number of events. For our case, we find $p < 28\%$, showing the limitations related to the size of our astrophysical background. Another potential limitation of our background is the low abundance of GW231123-like events, for which there is a higher risk of false alarm due to degeneracies between spin and lensing effects (Liu & Kim 2024). Additionally, short signals are known to be more affected by non-Gaussian noise fluctuations (e.g., Abac et al. 2025g; Ray et al. 2025), which could then be captured by the lensing model. While our background uses real noise to account for such effects, the noise was cleaned before injections were done (see Section B), which could lead to over-cleaning and reduce the effect of non-Gaussian features. So, further studies are needed to quantify their impact.

Because of its outstanding properties, we investigate GW231123 in more detail in the next section.

6. GW231123_135430

This section presents more thorough analyses of GW231123, flagged in Abac et al. (2025d) as a potential
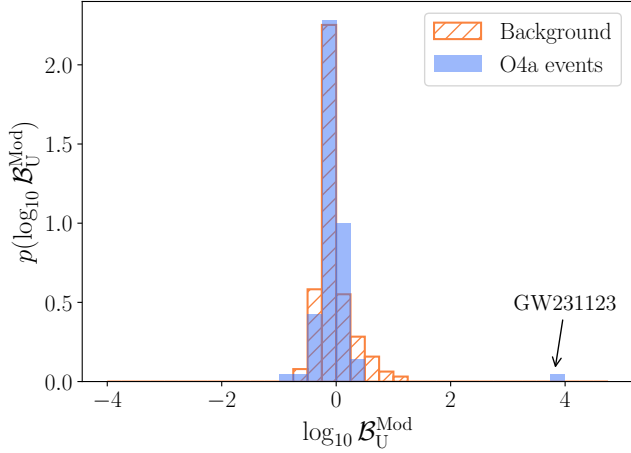
Figure 5. Histogram of $\log_{10}\mathcal{B}_{\mathrm{U}}^{\mathrm{Mod}}$ for O4a BBH candidates. Except for GW231123, all Bayes factors found with IMRPhenomXPHM-SpinTaylor fall within the range expected for unlensed events based on the background.

candidate for lensing distortions, which is also further shown in Section 5. We focus on comparing the lensed and unlensed hypotheses for a quasi-circular BBH coalescence.

Under the unlensed hypothesis, GW231123 is exceptional as its source corresponds to one of the most massive BBHs detected so far, with a total mass of 190–265 $M_\odot$ and high spins (Abac et al. 2025d). Analysis results also show variation with the waveform model. By standard accuracy measures, NRSur7dq4 is, on average, the most accurate available model in this region of parameter space, although Abac et al. (2025d) finds that in some cases all models can lead to biased parameter estimates.

Here, we present the properties of the source and point mass lens under the lensed hypothesis, evaluate the evidence of lensing by fold caustic, search for various strong lensing effects, discuss the astrophysical priors of lensing, investigate in more detail waveform consistency in the lensing context, and look at possible noise artifacts that could increase the support for the lensing hypothesis.

6.1. *Properties of GW231123 and the Corresponding Point Mass Lens*

For the single-distorted signal analyses done, GW231123 has the highest Bayes factor seen so far, above what is expected from our astrophysical background. Noting the significant disagreements between the source parameters estimated using different (unlensed) waveform approximants (Abac et al. 2025d), we similarly perform our analyses using multiple waveforms. This initial set of

Waveform Model	XPHM-ST	NRSur	XO4a
$\log_{10}(B_{\mathrm{U}}^{\mathrm{Mod}})$	3.8	1.5	0.1

Table 2. Bayes factors for GW231123 when comparing the point-mass lensed and unlensed models with three different waveform approximants. XPHM-ST, NRSur, and XO4a denote IMRPhenomXPHM-SpinTaylor, NRSur7dq4, and IMRPhenomXO4a, respectively.

tests is performed using the IMRPhenomXPHM-SpinTaylor (Pratten et al. 2021; Colleoni et al. 2025), IMRPhenomXO4a (Thompson et al. 2024), and NR-Sur7dq4 (Varma et al. 2019) waveform models. The corresponding recovered Bayes factors are shown in Table 2. Note that we consider only an isolated point mass lens and do not consider the case of e.g., a microlens embedded in the potential of a bigger lens, such as a galaxy, which could change the inferred quantities (Diego et al. 2019; Cheung et al. 2021; Yeung et al. 2023; Mishra et al. 2021, 2024; Shan et al. 2023, 2024, 2025a,b; Seo et al. 2022, 2025). While these Bayes factors indicate some favoring of the lensing hypothesis under the assumption of Gaussian and stationary noise for the IMRPhenomXPHM-SpinTaylor and NRSur7dq4 cases, support is more marginal for IMRPhenomXO4a. Spurious favoring of lensing can be caused by features unaccounted in the analyses, such as non-Gaussian noise effects or inaccuracy in waveforms used. The favoring of lensing decreases for waveform models more favored by the data (Abac et al. 2025d). This shows a possible support for lensing coming from compensating for differences between the waveform models and the data. We follow up on some of these aspects in this and the next subsections.

For the point-mass analysis, we infer the lens properties (the redshifted lens mass and the dimensionless impact parameter) using the three waveform models. Using IMRPhenomXPHM-SpinTaylor (NRSur7dq4), the redshifted lens mass and dimensionless impact parameter are $911^{+504}_{-292}M_\odot$ ($803^{+489}_{-299}M_\odot$) and $0.61^{+0.27}_{-0.21}$ ($0.69^{+0.37}_{-0.26}$), respectively. The IMRPhenomXO4a waveform yields broader posterior ranges for both lens parameters, with $687^{+13208}_{-356}M_\odot$ and $2.13^{+0.79}_{-1.54}$. Details of the lens properties are provided in Appendix C.1.

Turning to the inferred BBH properties, under the lensed hypothesis, we observe a broadening of the posteriors and a reduction in the differences between the waveform models considered. In Figure 6, we show that apparent differences in the recovered posterior in the unlensed analyses are reduced in the lensed ones for the detector-frame chirp mass, with similar trends found for all parameters. Broader posteriors are expected under the lensing hypothesis, because of the greater freedom

Parameter	$\mathcal{M}$		χ_1		χ_2	
	KS	JSD	KS	JSD	KS	JSD
Unlensed	0.647	0.294	0.150	0.044	0.051	0.002
Lensed	0.342	0.089	0.097	0.009	0.026	4×10^{-4}

Table 3. Kolmogorov–Smirnov statistics (KS) and Jensen–Shannon divergence (JSD, in nats) values between the chirp mass and spin magnitude posteriors obtained for individual detectors under the unlensed and lensed hypotheses using the IMRPHENOMXPHM-SPINTAYLOR waveform.

in fitting the lens. The broader posteriors extend to regions of parameter space where waveform models are more reliable.

Analyses were conducted examining only the data obtained by each of the LIGO Hanford (LHO) and LIGO Livingston (LLO) detectors. Results for each detector are then compared to ascertain whether any transient noise artifacts in a single detector could result in the apparent support for lensing, as was noted for GW200208_130117 in Janquart et al. (2023b). Results from both detectors indicate support for the lensing hypothesis, with stronger evidence in the LLO analysis than in the LHO analysis. Specifically, the $\log_{10}(B_{\mathrm{U}}^{\mathrm{Mod}})$ values are 0.62 for LHO and 2.64 for LLO analyses with the IMRPHENOMXPHM-SPINTAYLOR model. Such a difference could be driven by the higher SNR present in LLO compared to LHO, or to specific features present in one detector (see Section 6.6 for a discussion on this topic). Furthermore, we calculate Kolmogorov–Smirnov (KS) statistics and Jensen–Shannon (JS) divergence between the chirp mass and spin amplitude posteriors obtained from the unlensed LHO and LLO analyses and those from the corresponding lensed analyses. As shown in Table 3, both measures take lower values in the lensing case, showing the analyses under the lensed hypothesis show greater consistency for information obtained from each detector individually.

In the unlensed analysis, there is a distinct preference for extremal spins for this event. Lensing distortions and the effects of spin precession are highly correlated (e.g., Liu & Kim 2024). The comparison between the spins for the primary and secondary are shown in Figure 7 for the NRSUR7DQ4 waveform model. The observed behavior is qualitatively similar for all waveform models considered here. Under the lensing hypothesis, the higher values for the primary spin remain preferred albeit with a broader peak. The posterior on the secondary becomes much broader, granting more support for moderate spin values at the expense of support for extremal values. This greater support for moderate spin amplitudes may also lead to the greater agreement between the wave-
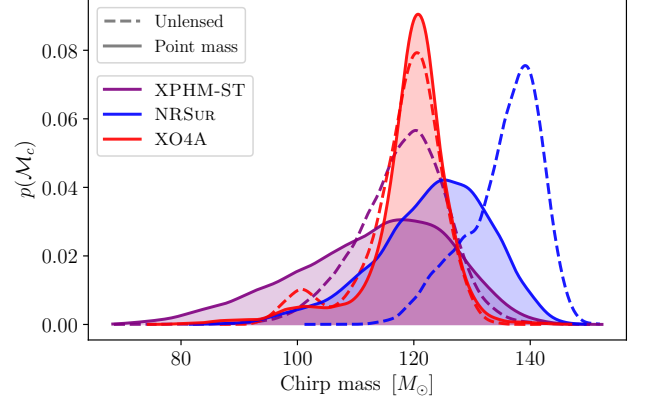
Figure 6. Comparison of the detector-frame chirp mass posteriors under the point mass lensing (solid line) and unlensing (dashed line) hypotheses using the three considered waveform models: IMRPHENOMXPHM (XPHM-ST), NR-SUR7DQ4 (NRSUR), and IMRPHENOMXO4A (XO4A). The difference between the waveforms is reduced when the analysis is done under the lensing hypothesis.
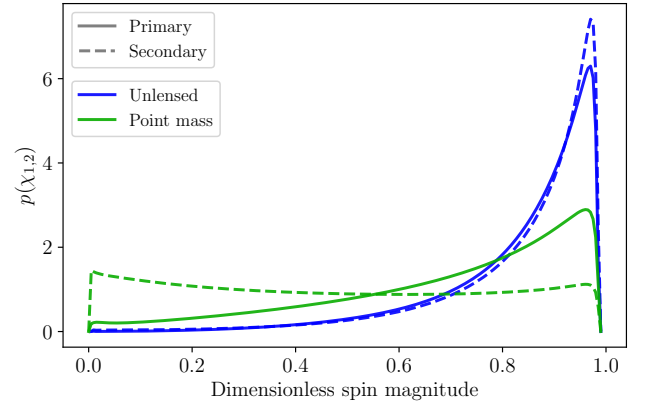
Figure 7. Posteriors of the dimensionless spin magnitudes recovered for GW231123 in the lensed (green) and unlensed (blue) analyses with the NRSUR7DQ4 waveform model. The solid(dashed) curves correspond to the primary (secondary) component. Under the lensing hypothesis, there is greater support for more moderate spins, particularly in the secondary component.

forms as it would bring the waveform models toward regions of better calibration (Varma et al. 2019; Pratten et al. 2021; Colleoni et al. 2025; Thompson et al. 2024).

Detector-frame masses remain high under the lensed hypotheses. When examining the source-frame masses under the isolated point mass model, these would be $109^{+15}_{-13}M_\odot$ for the primary and $59^{+13}_{-11}M_\odot$ for the secondary in the NRSUR7DQ4 analysis. This represents a modest drop in the primary mass and a significant decrease for the secondary compared with the $129^{+15}_{-14}M_\odot$ and $114^{+14}_{-17}M_\odot$ for the primary and secondary respec-

Waveform Model	$\log_{10} \mathcal{B}_{\mathrm{U}}^{\mathrm{Fold}}$
IMRPHENOMXPHM-SPINTAYLOR	2.05
NRSUR7DQ4	−0.05

Table 4. Log Bayes factor for the lensing near a fold caustic scenario compared to the unlensed BBH merger for GW231123, using two different models for the underlying BBH waveform.

Waveform	Detectors	n_j prior	$\log_{10} \mathcal{B}_{\mathrm{U}}^{\mathrm{L}}$	$\log_{10} \overline{\mathcal{B}}_X^{II}$
NRSUR	LHO–LLO	discrete	0.04	−0.80
NRSUR	LHO–LLO	continuous	−0.45	
XPHM-ST	LHO–LLO	discrete	−0.17	−0.35
XPHM-ST	LHO-only	continuous	0.27	
XPHM-ST	LHO–LLO	continuous	−0.14	
XPHM-ST	LHO-only	discrete	0.22	−0.01
XPHM-ST	LLO-only	continuous	0.29	
XPHM-ST	LLO-only	discrete	−0.52	0.51

Table 5. Summary table of the GOLUM type II analysis results for the IMRPHENOMXPHM-SPINTAYLOR (XPHM-ST) and NRSUR7DQ4 (NRSUR) waveforms. The third row corresponds to the result for GW231123 plotted in Fig. 4. There are discrepancies between results depending on the waveform and detector configuration chosen. However, overall, no consistent evidence for a type II image is seen.

tively reported from the unlensed analysis (Abac et al. 2025d).

6.2. Searching for Lensing Near a Fold Caustic

Here, we analyze GW231123 in the context of lensing near a fold caustic of a foreground lens, producing highly magnified images. This analysis has only one extra free parameter compared to the unlensed BBH case, which is the time delay t_{ij} between the two lensed signals. A uniform distribution from 0.002 s to 0.6 s is chosen to be the agnostic prior for t_{ij}. The lower bound is limited by the sampling rate of the data that we used, while the upper bound is chosen such that the entire lensed waveform will still fit inside the standard 8 s analysis window.

Table 4 shows the Bayes factor comparing the lensing near a fold caustic with the unlensed BBH merger using the same two underlying waveform models, IMRPHENOMXPHM-SPINTAYLOR (Pratten et al. 2021; Colleoni et al. 2025) and NRSUR7DQ4 (Varma et al. 2019). We find that the unlensed hypothesis is preferred over the fold-caustic lensing when using NRSUR7DQ4, while the opposite is true when using IMRPHENOMXPHM-SPINTAYLOR. This is in line with other unlensed analyses showing discrepancies between the two models (Abac et al. 2025d).

When comparing the recovered posterior distribution for the BBH parameters, we also see that the two waveform models have a better agreement under the fold-caustic lensing hypothesis than under the unlensed hypothesis, possibly partly fitting information missing under the unlensed hypothesis. However, the obtained Bayes factors are not significant enough to claim strong evidence supporting fold-caustic lensing.

6.3. Searching for Strong Lensing Effects

Here we search for signatures of strong lensing in GW231123. We evaluate the evidence of it being a Type II lensed image, an outlier in the population of BBHs, and search for any sub-threshold counterparts produced by strong lensing.

6.3.1. Type II Lensed Image

While no evidence for type II images were found in Section 5.1, here, we re-do the GOLUM type II analyses for GW231123 with adapted waveform models and using two different priors on the Morse factor: a discrete one ($n_j \in \{0, 0.5, 1\}$) and a continuous one (between 0 and 1). The latter is done because if a signal is of type II, the posterior should peak at 0.5 while other effects, like noise, waveform systematics and deviation from GR, can lead to a Morse factor value peaking away, which cannot be seen in the discrete case (Abbott et al. 2023a; Janquart et al. 2023b; Wright et al. 2024). We also do analyses on individual detectors to gauge the consistency in results between them. In addition to the Bayes factor comparing explicitly the image types ($\overline{\mathcal{B}}_X^{II}$), we also report the Bayes factor comparing the evidence obtained in the lensed and unlensed analyses ($\mathcal{B}_{\mathrm{U}}^{\mathrm{L}}$). The results are summarized in Table 5.

Generally, while the NRSUR7DQ4 waveform model disfavors type II images, the IMRPHENOMXPHM-SPINTAYLOR results show greater ambiguity with support for type II images found in LLO-only runs, as well as the combined-detector continuous prior run. However, both the LHO–LLO and LLO-only continuous prior runs do not peak distinctly around $n_j = 0.5$, with some skew towards a higher or lower phase shift, thus indicating that this support is likely to come from some other, unidentified effect rather than true type II effects (Janquart et al. 2023b; Wright et al. 2024; Narayan et al. 2024). Type II analyses with different waveforms recover different posterior distributions for intrinsic parameters, supporting the systematics identified in Abac et al. (2025d). In summary, we do not find consistent evidence for a type II image.

6.3.2. *Searches for Population Outlier*

If a strongly lensed image is not identified as lensed, its magnification will bias the inferred source properties (Oguri 2018). In particular, a magnified event ($\mu > 1$) will appear as closer and more massive. Therefore, it is possible to search for strongly lensed signals as potential outliers of the inferred astrophysical population (Hannuksela et al. 2019; Farah et al. 2025). To avoid false positives, it is necessary to account for the uncertainty in the inferred astrophysical model as well as event parameter uncertainty induced by noise in the detector (Fishbach et al. 2020). If inferred source properties are identified as outliers, the magnification needed to make it consistent with the population can be determined. Alternatively, one can assume the lensing hypothesis and then check if the inferred magnification is consistent with the strong-lensing likelihoods (Hannuksela et al. 2019; Abbott et al. 2021). However, large magnifications are needed to lead to observable outliers, and such occurrences are rare (Farah et al. 2025).

While GW231123 has exceptionally high mass, it is still consistent with the high-mass tail for the population distribution and is not an outlier (Abac et al. 2025d,f). This fact alone is not sufficient to gauge the possibility that the event is lensed (Farah et al. 2025).

6.3.3. *Sub-Threshold Counterparts*

If GW231123 is part of a multiplet of images, it is of interest to see if we can find sufficiently similar counterparts in the data. No super-threshold candidate pairs pass the selection criteria from Section 4.1, but it is possible that some sub-threshold counterparts exist (see Section 4.2). Here we report also lower-significance triggers for completeness. We report the results for these searches with the TESLA-X (Li et al. 2023, 2025a) and PyCBC-based (McIsaac et al. 2020) searches. Including triggers with FAR below 1 per 30 days, the first approach leads to two triggers being found, while the other leads to eight triggers that correspond to known super-threshold events and four additional sub-threshold triggers, only two of which have sky localization overlap with GW231123. None of these triggers are common between the two searches. The two TESLA-X triggers and one from PyCBC correspond to candidates included in the full GWTC-4.0 list but not as sufficiently significant for full data-quality and parameter-estimation studies. The triggers with non-zero sky overlap are reported in Table 6 with the same information as in Table 1. None of them appear to be of sufficient interest based on their FARs and we, therefore, do not perform further follow-up studies on them.

6.4. *Possible Astrophysical Systems*

To assess the significance using Bayesian approaches, one should ideally evaluate the Bayes factor against the plausibility of lensing. Here, we consider a few selected example systems where lensing may be expected, based on existing literature.

From an astrophysical perspective, for a black hole lens, one would expect such systems to be formed in dense environments, such as star clusters for a central massive lens, von Zeipel–Lidov–Kozai induced mergers (Ubach et al. 2025), or mergers happening in a migration trap around an active galactic nucleus (Leong et al. 2025). Such models require a particular formation channel for the BBHs. However, under those scenarios, Ubach et al. (2025) and Leong et al. (2025) find that a small fraction of events may be lensed. Alternatively, one can take a more empirical approach and consider lensing candidates with millisecond to second time delays reportedly observed via gam ma ray bursts (Paynter et al. 2021; Levan et al. 2025). While these candidates are being debated (Mukherjee & Nemiroff 2023), they can offer an avenue to constraining the plausibility of the lensing hypothesis empirically.

A comprehensive assessment of the astrophysical priors for lensing, including selection effects and considering specific lens models outlined above, should be carried out in the future to better assess the plausibility of the lensing hypothesis.

6.5. *Waveform Reconstruction*

One investigation to understand where the apparent support for the lensing hypothesis (under the point mass lens model presented in Section 6.1) could come from is to compare the inferred unlensed and lensed waveforms with the one reconstructed by an unmodelled search pipeline for GW signals (Shan et al. 2025a). Here, we use cWB (Drago et al. 2020), which uses a wavelet basis to reconstruct coherent power across detectors. A comparison of the waveforms from the cWB analysis, the unlensed parameter estimation, and the lensed parameter estimation (using the point mass lens model) are shown in Figure 8. Here and in the next Section, we focus on NRSur7dq4 since it was shown in Abac et al. (2025d) to be the most accurate waveform model on average for events in this region of the parameter space.

To quantify the difference between the two models, we use the mismatch, which represents the dissimilarity between two signals and is defined as (Owen 1996)

$$\bar{\mathcal{M}} = 1 - \max_{t_0, \phi_0} \frac{\langle h_1 | h_2 \rangle}{\sqrt{\langle h_1 | h_1 \rangle \langle h_2 | h_2 \rangle}} \,, \qquad (5)$$

Analysis	Sub-Threshold Trigger (UTC)	Absolute Time Delay	Network SNR	FAR	$\mathcal{O}_{90\% \text{ CI}}$	$\mathcal{R}_{\text{U}}^{\text{L}}$
	(yy:mm:dd hh:mm:ss)	(days)		(yr^{-1})	(%)	
PYCBC	2023-10-29 05:30:40*	25.35	8.15	4.13	16	1.22
PYCBC	2023-07-30 17:22:24	115.86	8.61	4.87	11	0.10
TESLA-X	2023-11-13 06:20:50*	10.32	8.64	6.52	23	3.10
TESLA-X	2023-10-02 14:39:16*	51.97	9.41	9.97	19	0.43

Table 6. Possible counterpart triggers found for GW231123 with our sub-threshold searches and with non-zero sky overlap, excluding those already considered as significant GWTC-4.0 events, down to FAR of 1 per 30 days. We find two such triggers with the TESLA-X search and four with the PYCBC-based search, though two of those do not have overlapping sky localizations with GW231123 and are not reported in the table. Triggers marked with an asterisk correspond to candidates included in the full GWTC-4.0 list, but due to FAR $> 1 \, \text{yr}^{-1}$ in the global search no data-quality or parameter-estimation studies were performed on those. PYCBC also recovered eight of the significant GWTC-4.0 events as possible counterparts for GW231123, but those were already excluded by the standard super-threshold lensing analyses discussed in this paper.
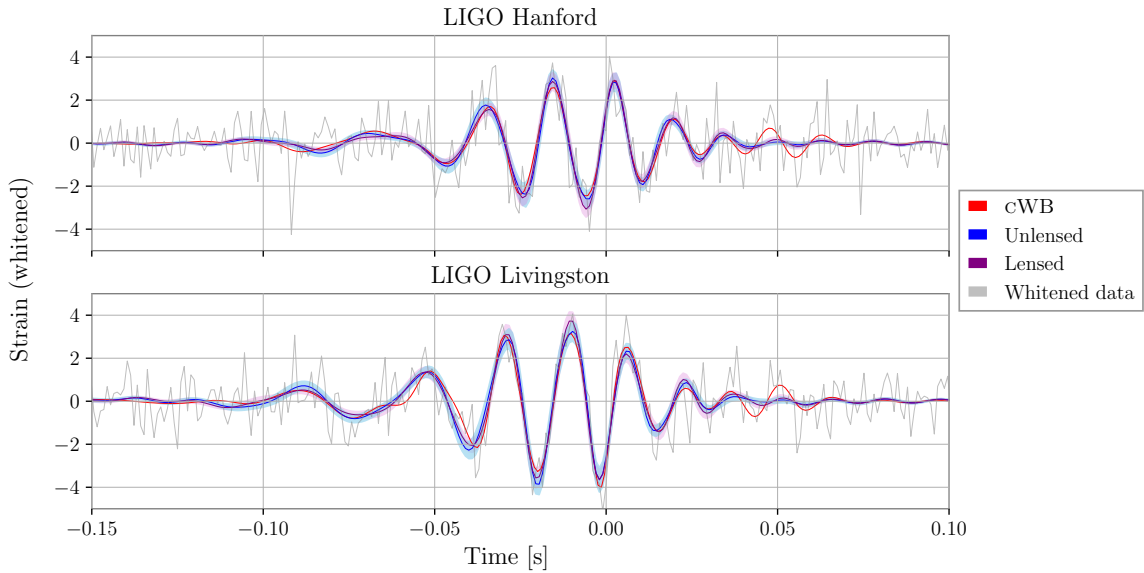
Figure 8. Comparison between the CWB waveform reconstruction and the 90% credible interval waveforms from the parameter-estimation analyses under the lensed and unlensed hypotheses with the NRSUR7DQ4 waveform model. The time represents the time relative to 2023 November 23 at 13:54:30.619 UTC. The lensed signal does not show an overall visibly closer match to the CWB waveform.

where $\langle h_1 | h_2 \rangle = \int df \, \tilde{h}_1^*(f) \tilde{h}_2(f) / S_n(f)$ is the noise-weighted inner product between two waveforms h_1 and h_2, with $\tilde{h}(f)$ being the Fourier transform of $h(t)$, $S_n(f)$ the one-sided power spectral density of the detector noise, and the maximization is done over time and phase of coalescence. Looking at the mismatch between the CWB reconstruction and the recovered lensed and unlensed waveforms when analyzing the signal with NRSUR7DQ4, we find the lensed waveform to generally better match the model-agnostic reconstructed signal. A distribution of mismatch values is given in Appendix C.2, and we find mismatches of $0.045^{+0.010}_{-0.007}$ and $0.051^{+0.010}_{-0.009}$ for LHO and LLO for the unlensed model, while they are of $0.035^{+0.007}_{-0.005}$ and $0.039^{+0.010}_{-0.007}$ for the lensed case. The difference between the two being small, they are likely to be caused by the extra degrees of free-

dom in the lensed model and are probably not significant.

6.6. Noise Considerations

A possible cause for the favoring of lensing would be the presence of non-Gaussian noise features. There are known glitches, corresponding to noise over-densities, around the time of the signal. Therefore, the LHO analyses are done using deglitched data (Abac et al. 2025d) but non-Gaussian noise features could still be present in the data biasing results, especially if considering population of microglitches (Ray et al. 2025). In this case, one would need the noise fluctuations to be similar in the two detectors to give the observed consistency with lensing characteristics.
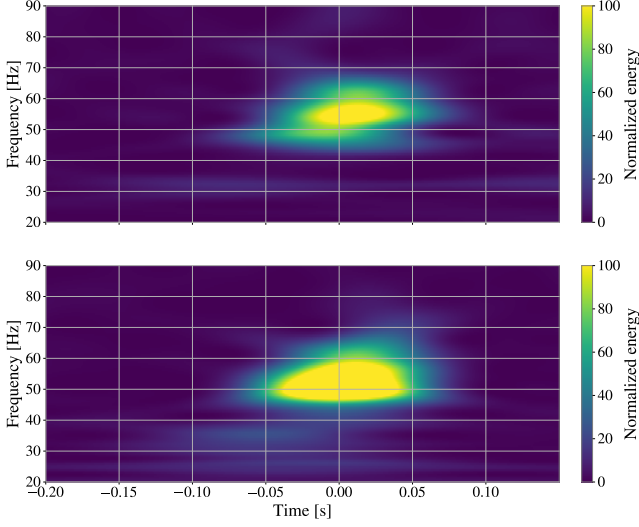
Figure 9. Normalized squared amplitude in time–frequency space for the lensed (top) and unlensed (bottom) results obtained with the NRSur7dq4 waveform model and the maximum-likelihood points. The time represents the offset relative to 2023 November 23 at 13:54:30.619 UTC. A comparison between the two waveforms reveals differences spread across multiple frequency bins rather than confined to a single bin, and these do not coincide with a particular spectral line.

While this is unlikely for random noise fluctuations, a possible candidate would be the 60 Hz line related to the power grid. In that case, one would expect the lensing support to be significantly reduced if the line is removed and the main excess power from lensing to be concentrated around that frequency. Doing such an analysis, we find that notching does not change the lensing support, which is also in line with findings from Abac et al. (2025d), where results were unchanged after notching the 60Hz line. Additionally, we verify whether the power is spread and not concentrated in a narrow band around 60 Hz by examining the power difference between lensed and unlensed results. Figure 9 shows the normalized squared amplitude in time-frequency space for the two cases at the maximum-likelihood points of the NRSur7dq4 analyses. As can be seen, the lensed case exhibits a broader power distribution compared to the unlensed case, indicating that the extra fitting introduced by lensing affects multiple frequency regions simultaneously rather than being confined to a single narrow band. Therefore, it is unlikely the support is due to a spectral line.

7. IMPLICATIONS

In this section, we present implications of the non-detection results of our strong-lensing search analyses, from which we can derive constraints on the detection rate of strong lensing and on the BBH merger rate density at high redshift.

7.1. *Strong-Lensing Event Rates*

We constrain the relative occurrence of lensed events due to galaxies and galaxy clusters compared to unlensed ones by following standard techniques (Ng et al. 2018; Li et al. 2018; Mukherjee et al. 2021a; Xu et al. 2022; Wierda et al. 2021; Phurailatpam et al. 2024). We simulate a large population of lensed and unlensed mergers based on the latest population constraints and determine the relative observation of such signals. The lens-mass model is chosen to be SIS for all types of lenses. For the galaxies, the velocity dispersion function follows from the Sloan Digital Sky Survey (Choi et al. 2007) with velocity dispersion $\sigma_{\mathrm{vd}} = 10$–$350\,\mathrm{km\,s^{-1}}$. For galaxy clusters, we use the halo-mass function from Tinker et al. (2008) with halo masses $\log_{10}(M_{\mathrm{halo}}/M_{\odot}) \sim$ 14–16. Although it could impact rates, we do not consider the effect of sub-substructures in clusters (Vujeva et al. 2025). The BBH population is taken from Abac et al. (2025f). We choose a detection criterion on the two-detector SNR of the lensed events to be above 8. Furthermore, this SNR cut is applied to (a) the weaker of the two lensed signals implying both are detectable (denoted D), or (b) the stronger of the two lensed signals implying at least one lensed signal is detectable (denoted S).

Table 7 gives the latest median and 90% strong-lensing event rates relative to the unlensed event detection rate for two choices of merger rate density models: (a) Power-Law model where the local rate R_0 and the power-law index κ are set to the median of the posteriors that are constrained by GWTC-4.0 (Abac et al. 2025f), and (b) Power-Law + Peak model where the posteriors of the parameters are jointly constrained from the GWTC-4.0 results as well as the non-detection of the stochastic GW background (SGWB; Abac et al. 2025h). In this case, the lensing rates are shown for the 90% credible interval. Our results are consistent with those based upon GWTC-3.0 (Abbott et al. 2023a) and have decreased uncertainties due to improved constraints on the merger rate density.

7.2. *Constraints on the Binary Black Hole Merger Rate Density at High Redshift*

The lensing optical depth is larger for binaries at high redshifts; thus the non-detection of strongly lensed GW signals provides a unique way of constraining the BBH merger rate at redshifts larger than the redshift horizon of the current detectors under the assumption that rate density models can be extrapolated to higher redshifts.

Merger Rate Density Model	Galaxy lenses		Galaxy Cluster lenses	
	R_{D}	R_{S}	R_{D}	R_{S}
Power-Law	18.1×10^{-4}	24.5×10^{-4}	5.4×10^{-4}	7.6×10^{-4}
Power-Law + Peak	3.2–9.9×10^{-4}	5.0–13.5×10^{-4}	0.9–3.8×10^{-4}	2.1–5.1×10^{-4}

Table 7. Expected fractional rates of observable lensed single or double events at current LVK sensitivity. Relative rates for lensed double events where R_{D} implies both of the lensed events are detected and R_{S} implies only the louder of the two events is detected above the SNR threshold and computed for galaxy- and cluster-scale lenses separately. In the Power-Law model, the model parameters are set to the median of the posteriors constrained by GWTC-4.0 (Abac et al. 2025f). In the Power-Law + Peak model, the parameters are obtained from the combined analysis of CBC detections and limits from SGWB (Abac et al. 2025h). Here, the ranges in the lensing rates encompass a 90% credible interval given the posteriors.

This is complementary to the constraints obtained from the non-detection of SGWB, since the relative contribution of binaries at different redshifts to the SGWB and lensing is, in general, different (Buscicchio et al. 2020a,b; Mukherjee et al. 2021b).

We start with the latest posteriors for the merger rate density model, assumed to have power-law redshift dependence and constrained by the GWTC-4.0 (Abac et al. 2025f), to retain those models that are consistent with non-detection of strongly lensed BBH signals (see Abbott et al. (2021, 2023a) for details). Figure 10 shows the original GWTC-4.0 only constraints and the revised upper bounds from absence of strongly lensed BBH events. Independent of the strong lensing analysis, we show constraints on the merger rate density obtained from the non detection of a SGWB for reference. The non-detection of lensing signatures provides upper limits comparable to those obtained from the SGWB. The SGWB results assume a Power-Law + Peak model for the redshift-dependence of the merger rate densities (as described in Section 7.1) and hence, a direct comparison between strong lensing and SGWB is not trivial. More stringent constraints from strong lensing are expected in the future as detectors become more sensitive and more signals are detected (Buscicchio et al. 2020a,b; Mukherjee et al. 2021b).

8. CONCLUSION

We have performed searches for GW lensing signatures in O4a data. Our analysis combined three complementary strands:

Pair-wise strong-lensing searches:—All 3486 possible pairs of O4a astrophysical BBH detections with FAR $<$ $1\,\mathrm{yr}^{-1}$ were processed with two independent tier-1 analyses (Posterior Overlap and Phazap) and an intermediate fast-joint-PE filtering process (tier 2). After filtering, 50 pairs remained and were followed up with a full joint parameter estimation (tier 3). All pairs considered have Bayes factors consistent with unlensed expectations, and none were further followed up.
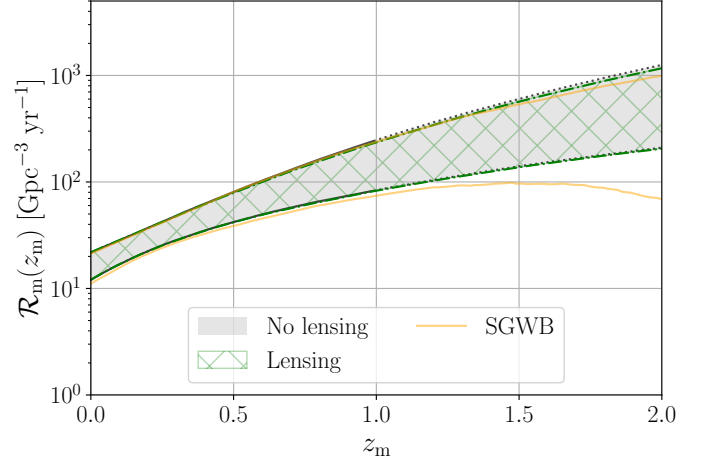
Figure 10. Constraints on the BBH merger rate density as a function of the merger redshift (z_{m}), obtained from the non-detection of strongly lensed GW signals during O4a (green cross-hatched region). For reference, we show the currently allowed merger rates constrained by the GWTC-4.0 results (grey shaded region, no lensing; Abac et al. 2025f). The solid orange curves enclose the constraints from the absence of detection of a SGWB in the O4a data (Abac et al. 2025h).

Sub-threshold counterpart searches:—Targeted searches for possible sub-threshold lensed counterparts were conducted for all BBH candidates found in O4a with FAR $< 1\,\mathrm{yr}^{-1}$. The general search was done with a PyCBC-based approach. We have presented the candidates recovered with single-template FAR $< 1\mathrm{yr}^{-1}$. We followed up on the new interesting candidates, not finding significant support for lensing. We also reported candidate counterparts for GW231123 found with two different approaches (TESLA-X and PyCBC).

Single-event analyses:—All BBH candidates with FAR $< 1\,\mathrm{yr}^{-1}$ were re-analysed with an isolated point-mass lens model, searching for frequency-dependent lensing distortions. With the exception of GW231123, no signal showed a significant lensing signature, and Bayes factors were undistinguishable from those of unlensed sim-

ulations. No signal displayed the unambiguous type-II phase structure predicted for strongly lensed images with large higher-mode content.

GW231123:—This event was further investigated in order to understand the significance of the lensing hypothesis and to understand what could be driving the large Bayes factors. The probability of an astrophysical background of unlensed signals producing this level of consistency with a lensed waveform is $< 0.39\%$ ($< 28\%$) without (with) the inclusion of the trials factor, where the significance is limited by the size of our background simulations. However, missing physics in the unlensed waveform models compensated by lensing effects could drive the large Bayes factors. Further observations of BBHs and of gravitational lenses will help determine the probability for GW231123 to be lensed.

Implication of the non-detection of strong lensing:—Assuming GW231123 is not affected by strong lensing, the non-detection of this lensing effect with BBH-population and lens-population models yields an observable lensed doubles per unlensed BBH detection rate of 3.2–9.9×10^{-4} (0.9–3.8×10^{-4}) for galaxy (galaxy-cluster) lenses at O4a sensitivity. Thus, the non-detection of strong lensing in the current data is not surprising. The absence of detected strongly lensed signals also enables constraints on the BBH merger rate density at a redshift larger than one. The upper limit found is comparable to that found via the non-detection of the SGWB.

Future runs with higher sensitivity and more detectors (Abbott et al. 2020) will also help in finding more confident detections of GW lensing signatures by reducing the risk of false matches with better sky localization and better measurement accuracy.

DATA AVAILABILITY

All strain data analysed as part of GWTC-4.0 are publicly available through GWOSC (Abac et al. 2025i). The details of this data release and information about the digital version of the GWTC are described in detail in (Abac et al. 2025b).

ACKNOWLEDGMENTS

Analyses in this catalog relied upon the LALSuite software library (LIGO Scientific Collaboration et al. 2018). Estimates of the noise spectra and glitch models were obtained using BayesWave (Cornish & Littenberg 2015; Littenberg et al. 2016; Cornish et al. 2021). The detection of sub-threshold candidates and their statistical significance were obtained using Gst-LAL (Messick et al. 2017; Sachdev et al. 2019; Hanna et al. 2020; Cannon et al. 2020) and PyCBC (Allen

et al. 2012; Dal Canton et al. 2014; Usman et al. 2016; Nitz et al. 2017; Davies et al. 2020) based software. Source-parameter estimation was performed with the Bilby library (Ashton et al. 2019; Smith et al. 2020; Romero-Shaw et al. 2020) using the Dynesty nested sampling package (Speagle 2020). PESummary was used to postprocess and collate parameter-estimation results (Hoy & Raymond 2021). The various stages of the parameter-estimation analysis were managed with the Asimov library (Williams et al. 2023). For lensing analyses, we used TESLA-X (Li et al. 2025b), Posterior Overlap (Barsode et al. 2025), Phazap (Ezquiaga & Lo 2023), GOLUM (Janquart et al. 2022b), and Gravelamps (Wright et al. 2021). Some of the software above rely on standard libraries such as Astropy (Price-Whelan et al. 2022), Cython (Behnel et al. 2011), and Pandas (Wes McKinney 2010). Plots were prepared with Matplotlib (Hunter 2007) and Seaborn (Waskom 2021). NumPy (Harris et al. 2020) and SciPy (Virtanen et al. 2020) were used in the preparation of the manuscript.

APPENDIX

A. TIER-1 AND TIER-2 SEARCHES FOR MULTIPLE IMAGES

In this Appendix, we present more details on the tier-1 methods used to analyze the whole set of GW pairs considered in this work.

A.1. *Posterior Overlap*

The PO approach (Haris et al. 2018), already used in past lensing searches (Hannuksela et al. 2019; Abbott et al. 2021, 2023a; Janquart et al. 2023b), directly looks at the consistency between the posterior probability distributions obtained by traditional parameter-estimation analyses on the two events. It then calculates the overlap between them, leading to a ranking statistic, denoted $B_{\rm U}^{\rm L}$. In parallel, PO also looks at the compatibility between the measured time delay and the one expected from galaxy lenses. Requiring such compatibility has been shown to efficiently reduce the FPP (Haris et al. 2018; Wierda et al. 2021; More & More 2022; Janquart et al. 2022a). This second ranking statistic is noted $\mathcal{R}_{\rm U}^{\rm L}$. The final statistic for the search is then obtained by multiplying the two. This product is then compared with the same statistic obtained from unlensed background events (see Appendix B), which is used to compute the FPP.

A.2. *Phazap*

PHAZAP (Ezquiaga et al. 2023) gauges the lensed nature of a system by looking for matching phases upon arrival at the detectors. The problem can be reduced to a six-dimensional parameter space, consisting of $\{\phi_{\rm H}, \phi_{\rm L}, \phi_{\rm V}, \tau_{\rm HL}, \tau_{\rm HV}, \Delta\phi_f\}$ representing the phases at each detector, the arrival-time difference and the frequency evolution of the phase. The phases $\phi_{\rm H,L,V}$ of the signals under consideration are reconstructed at a reference frequency (40 Hz) for each of the three detectors, while $\Delta\phi_f$, the difference in orbital phase for a given detector at two different frequency points, spans from 20 Hz to 100 Hz. If a pair of signals is indeed strongly lensed, these six phasing parameters should coincide for the two signals, modulo a possible overall phase offset at the detector phases of $\pi/2$ or π due to the Morse phase. In the presence of detector noise and measurement uncertainties, we have a distribution for each parameter instead of a point estimate which is derived directly from the posterior samples of each event. We test their consistency with the strong-lensing hypothesis by measuring their distance in the six-dimensional phase space. The compatibility between phases is then quantified by a p-value. A similar statistic is computed for the unlensed background events
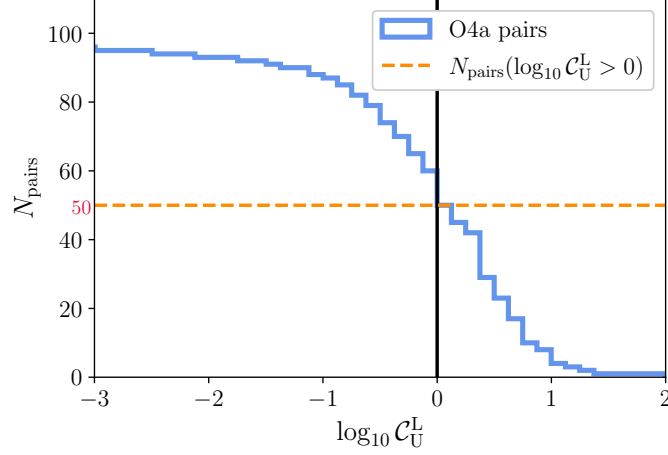
Figure 11. Cumulative histogram showing the distribution of the coherence ratio $\log_{10} \mathcal{C}_U^L$ of the lensed evidence against the unlensed evidence for the pairs passed to FAST-GOLUM. We only show pairs with $\log_{10} \mathcal{C}_U^L > -3$. This analysis keeps 50 pairs to be followed up by tier-3 analyses.

(see Appendix B), and the FPP is then computed as the fraction of unlensed pairs with a p-value smaller than the one obtained for the lensed pair.

A.3. *Fast-GOLUM*

FAST-GOLUM (Janquart et al. 2021a, 2023a) uses the posterior of the first image, obtained by analyzing it under the strong-lensing hypothesis, to compute the conditioned evidence. This is done by performing nested sampling on the relative lensing parameters. By multiplying the conditioned evidence with the evidence of the first image, one obtains the joint evidence for strong lensing. Since this first image posterior is already concentrated on the region of interest in parameter space and one can further speed up calculation by using a look-up table, computing the joint evidence is accelerated. Dividing this joint evidence by the product of the unlensed evidences of the two images gives the coherence ratio

$$\mathcal{C}_U^L = \frac{p(d_2|d_1, L)p(d_1|L)}{p(d_1|U)p(d_2|U)} , \tag{A1}$$

where d_i is the data for the ith signal, and L and U denote the lensed and unlensed hypotheses, respectively.

This approach is used for all event pairs flagged as interesting by tier-1 analyses. Figure 11 shows the distribution of the coherence ratio $\log_{10} \mathcal{C}_U^L$ of the 105 pairs analyzed by FAST-GOLUM.

B. ASTROPHYSICAL BACKGROUND

To construct our astrophysical background, we simulated around 40,000 BBH signals in the LIGO-only detector network, assuming the two detectors are operational for all detections and requiring network optimal SNR > 8. From these, a random subset of 350 signals was selected for further analysis. The mass and spin priors were derived from the inferred population model based on GWTC-3.0 data (Abbott et al. 2023b): the source-frame component masses were drawn from the POWERLAW+PEAK distribution with a higher-mass cut extended at 200 $M_\odot$, and the spins were assumed to be independent and identically distributed, with spin magnitudes following a Beta distribution and spin orientations (tilts) modeled as an isotropic plus truncated half-Gaussian mixture (Talbot et al. 2019). The merger rate density was assumed to follow the Madau–Dickinson profile (Madau & Dickinson 2014; Fishbach et al. 2018), with cosmological parameters consistent with Ade et al. (2016).

These BBH signals were then injected into detector data from the O4a observing run. To avoid loud glitches or the presence of other GW signals, we scanned the detector data using the OMICRON pipeline (Robinet et al. 2020) and excluded stretches with SNR > 6. The remaining data was divided into 32 s segments and further cleaned for residual loud non-Gaussian features using BAYESWAVE (Pankow et al. 2018; Cornish et al. 2021; Chatziioannou et al. 2021; Hourihane et al. 2022). This cleaning was required since glitches can persist in the data even where OMICRON trigger has SNR lower than 6, as can be seen in Figure 12. Here, cleaning is done before injections, which is done for computational efficiency, but could lead to an over-correction of non-Gaussian features. The waveforms, generated
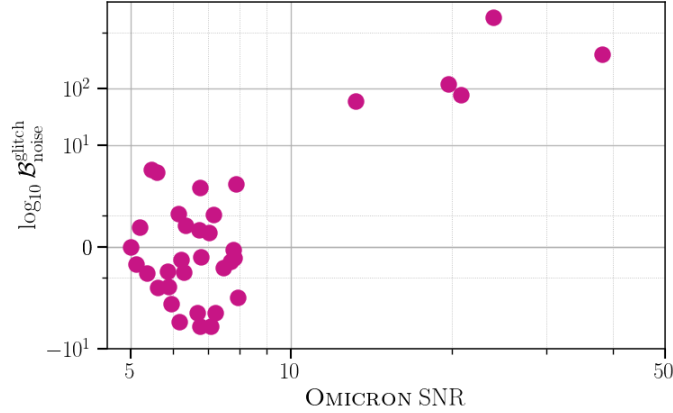
Figure 12. Comparison between the Bayes factor for the presence of a glitch versus purely Gaussian noise ($\log_{10} \mathcal{B}_{\mathrm{noise}}^{\mathrm{glitch}}$), computed using BAYESWAVE, and the Omicron trigger SNR across various data segments. For Omicron SNRs greater than 10, the Bayes factor is consistently high (> 100), indicating strong agreement between the two measures. However, for SNRs below 8, there are instances where the Omicron SNR is low, yet the Bayes factor remains significant (greater than 1), and in some cases even higher (exceeding 3), which suggests strong disagreement between the two measures according to the Kass–Raftery scale (Kass & Raftery 1995).
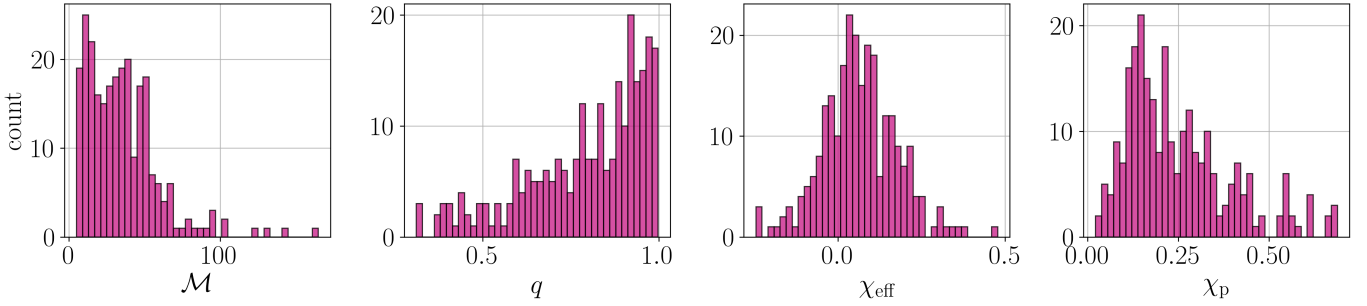
Figure 13. Distributions of the intrinsic parameters characterizing the background population: chirp mass ($\mathcal{M}$), mass ratio (q), effective inspiral spin (χ_{eff}), and effective precession spin (χ_{p}).

using IMRPHENOMXPHM-SPINTAYLOR waveform model (Pratten et al. 2021; Colleoni et al. 2025), were then added to these cleaned segments. The PSD estimated by BAYESWAVE was used for computing likelihoods during parameter estimation.

Out of the 350 injected and analyzed signals, 254 had a recovered network matched-filter SNR greater than 8. Those were considered detections, or simulated observations, for our astrophysical background. Figure 13 shows the distribution of intrinsic parameters characterizing this background population.

C. ADDITIONAL RESULTS FOR GW231123

C.1. *Posteriors Inferred From Point-Mass Analyses with Various Waveform Models*

The inferred lens properties from the point-mass analyses vary depending on the waveform model used. Figure 14 shows the posteriors obtained from the point-mass analyses using three waveforms.

C.2. *Mismatch Between* CWB *and Waveform Models*

In addition to the values reported for the mean waveforms in Section 6.5, here we also show the distribution of mismatch values between the CWB reconstruction and the recovered lensed and unlensed distributions for all waveforms in the 90% credible interval. Figure 15 shows the distribution of mismatch values for each LIGO detectors. The mismatch values between the CWB reconstruction and the mean reconstructed unlensed and lensed waveforms are 0.034 (0.039) and 0.028 (0.031) for LHO (LLO), respectively, indicating that the representative lensed reconstruction provides a closer match to the model-agnostic case. This is further shown by the median and 90% credible intervals
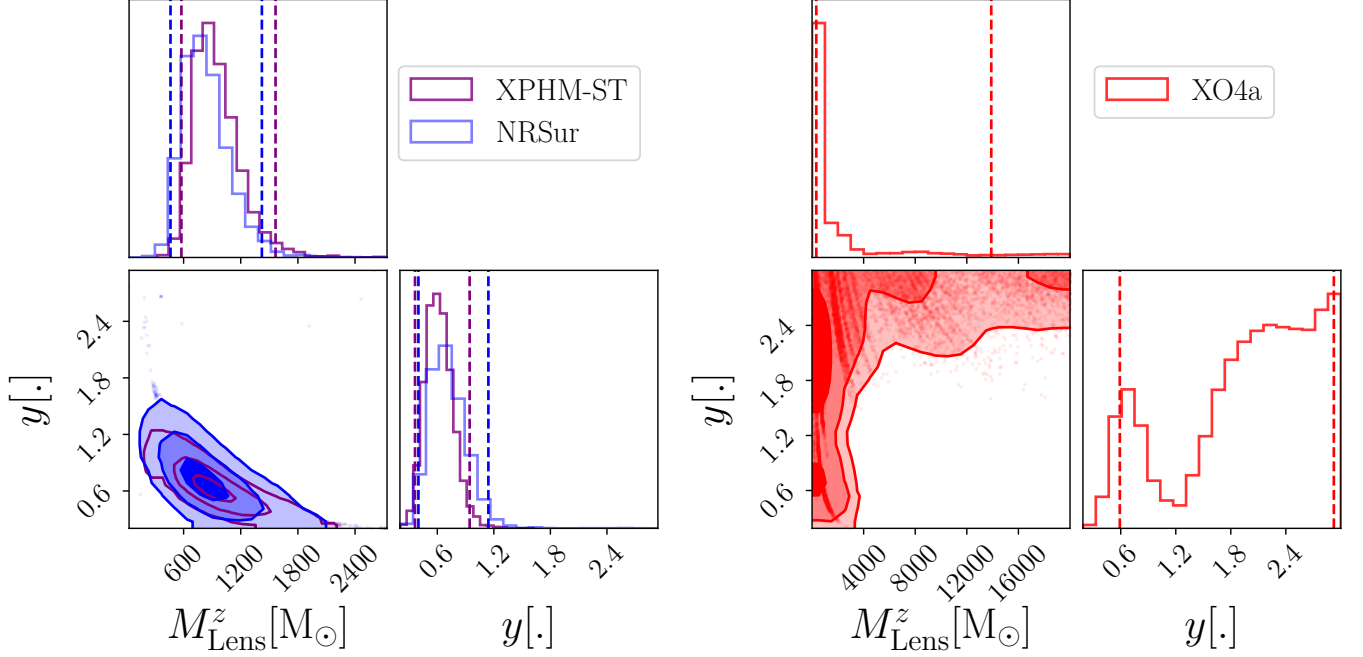
Figure 14. Posteriors of redshifted lens mass and dimensionless impact parameter inferred from the point-mass analyses with IMRPHENOMXPHM-SPINTAYLOR (XPHM-ST; purple), NRSUR7DQ4 (NRSUR; blue), and IMRPHENOMXO4A (XO4A; red) waveforms. The vertical dashed lines indicate the edges of the 90% credible intervals. Consistent with the lensing supports shown in Table 2, the XPHM-ST and NRSUR analyses show well-converged posteriors with a clear peak, whereas the analysis with the XO4A yields broader posteriors, reflecting weaker lensing support.

found for the mismatch distribution: $0.045^{+0.010}_{-0.007}$ and $0.051^{+0.010}_{-0.009}$ for LHO and LLO for the unlensed model, and $0.035^{+0.007}_{-0.005}$ and $0.039^{+0.010}_{-0.007}$ for the lensed case. The difference between the distributions in the lensed and unlensed case is further quantified using the KS statistic, yielding values of 0.83 for LHO and 0.76 for LLO, indicating that, although the distributions appear relatively close, they are statistically different. While such improvement shows that the lensed model is more consistent with the CWB reconstruction, it can be attributed to the additional degrees of freedom in the lensed model and is not a definitive proof of the lensed nature of the event.
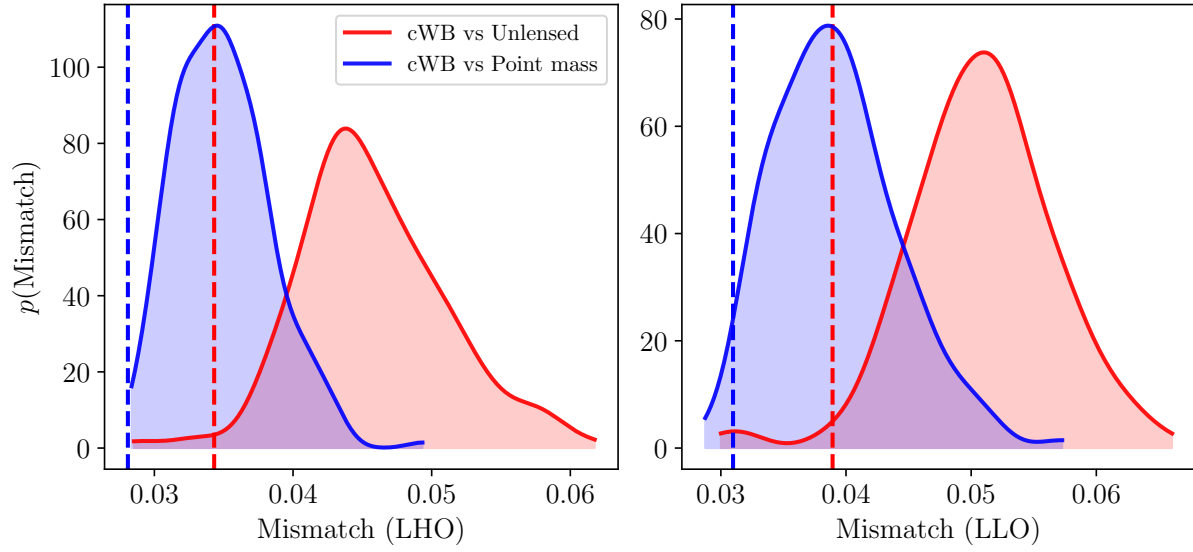
Figure 15. Distribution of the mismatch between the CWB reconstruction and the recovered lensed and unlensed distributions for all waveforms in the 90% credible interval. The vertical dashed lines indicate the mismatch values between the CWB reconstruction and mean unlensed (red) and mean lensed (blue) waveforms, respectively. The lensed reconstructed signal generally better matches the model-agnostic reconstructed signal than the unlensed one.

REFERENCES

Abac, A. G., et al. 2024, Astrophys. J. Lett., 970, L34, doi: 10.3847/2041-8213/ad5beb

—. 2025a. https://arxiv.org/abs/2508.18080

—. 2025b. https://arxiv.org/abs/2508.18079

—. 2025c. https://arxiv.org/abs/2508.18082

—. 2025d. https://arxiv.org/abs/2507.08219

—. 2025e. https://arxiv.org/abs/2508.18081

—. 2025f. https://arxiv.org/abs/2508.18083

—. 2025g. https://arxiv.org/abs/2509.07348

—. 2025h. https://arxiv.org/abs/2508.20721

—. 2025i, To be published in this issue. https://arxiv.org/abs/2508.18079

Abbott, B. P., et al. 2020, Living Rev. Rel., 23, 3, doi: https://doi.org/10.1007/lrr-2016-1

Abbott, R., et al. 2021, Astrophys. J., 923, 14, doi: 10.3847/1538-4357/ac23db

—. 2023a, arXiv:2304.08393. https://arxiv.org/abs/2304.08393

—. 2023b, Phys. Rev. X, 13, 011048, doi: 10.1103/PhysRevX.13.011048

Ade, P. A. R., et al. 2016, Astron. Astrophys., 594, A13, doi: 10.1051/0004-6361/201525830

Allen, B., Anderson, W. G., Brady, P. R., Brown, D. A., & Creighton, J. D. E. 2012, Phys. Rev. D, 85, 122006, doi: 10.1103/PhysRevD.85.122006

Ashton, G., et al. 2019, Astrophys. J. Suppl., 241, 27, doi: 10.3847/1538-4365/ab06fc

Balaudo, A., Garoffolo, A., Martinelli, M., Mukherjee, S., & Silvestri, A. 2023, JCAP, 06, 050, doi: 10.1088/1475-7516/2023/06/050

Barsode, A., Kapadia, S., & Ajith, P. 2024, The Astrophysical Journal, 975, 48

Barsode, A., Kumar, S., & Haris, K. 2025, PO: Posterior Overlap method for identifying strongly lensed gravitational wave signals from binary black hole mergers, https://git.ligo.org/ankur.barsode/posterior-overlap-o4a

Basak, S., Ganguly, A., Haris, K., et al. 2022, Astrophys. J., 926, L28, doi: 10.3847/2041-8213/ac4dfa

Behnel, S., Bradshaw, R., Citro, C., et al. 2011, Computing in Science Engineering, 13, 31, doi: 10.1109/MCSE.2010.118

Bohé, A., et al. 2017, Phys. Rev. D, 95, 044028, doi: 10.1103/PhysRevD.95.044028

Buscicchio, R., Moore, C. J., Pratten, G., et al. 2020a, Phys. Rev. Lett., 125, 141102, doi: 10.1103/PhysRevLett.125.141102

Buscicchio, R., Moore, C. J., Pratten, G., Schmidt, P., & Vecchio, A. 2020b, Phys. Rev. D, 102, 081501, doi: 10.1103/PhysRevD.102.081501

Çalışkan, M., Ezquiaga, J. M., Hannuksela, O. A., & Holz, D. E. 2023, Phys. Rev. D, 107, 063023, doi: 10.1103/PhysRevD.107.063023

Cannon, K., et al. 2020. https://arxiv.org/abs/2010.05082

Cannon, K., Caudill, S., Chan, C., et al. 2021, SoftwareX, 14, 100680, doi: 10.1016/j.softx.2021.100680

Capote, E., Dartez, L., & Davis, D. 2024, Class. Quant. Grav., 41, 185001, doi: 10.1088/1361-6382/ad694d

Capote, E., Jia, W., Aritomi, N., et al. 2025, Phys. Rev. D, 111, 062002, doi: 10.1103/PhysRevD.111.062002

Chan, J. C. L., Dyson, C., Garcia, M., Redondo-Yuste, J., & Vujeva, L. 2025a, Phys. Rev. D, 112, 064009, doi: 10.1103/6h6r-46cd

Chan, J. C. L., Seo, E., Li, A. K. Y., Fong, H., & Ezquiaga, J. M. 2025b, Phys. Rev. D, 111, 084019, doi: 10.1103/PhysRevD.111.084019

Chatziioannou, K., Cornish, N., Wijngaarden, M., & Littenberg, T. B. 2021, Phys. Rev. D, 103, 044013, doi: 10.1103/PhysRevD.103.044013

Chen, A., Cremonese, P., Ezquiaga, J. M., & Keitel, D. 2024, Phys. Rev. D, 110, 123015, doi: 10.1103/PhysRevD.110.123015

Cheung, M. H. Y., Gais, J., Hannuksela, O. A., & Li, T. G. F. 2021, Mon. Not. Roy. Astron. Soc., 503, 3326, doi: 10.1093/mnras/stab579

Choi, Y.-Y., Park, C., & Vogeley, M. S. 2007, Astrophys. J., 658, 884, doi: 10.1086/511060

Colleoni, M., Vidal, F. A. R., García-Quirós, C., Akçay, S., & Bera, S. 2025, Phys. Rev. D, 111, 104019, doi: 10.1103/PhysRevD.111.104019

Cornish, N. J., & Littenberg, T. B. 2015, Class. Quant. Grav., 32, 135012, doi: 10.1088/0264-9381/32/13/135012

Cornish, N. J., Littenberg, T. B., Bécsy, B., et al. 2021, Phys. Rev. D, 103, 044006, doi: 10.1103/PhysRevD.103.044006

Cremonese, P., Ezquiaga, J. M., & Salzano, V. 2021, Phys. Rev. D, 104, 023503, doi: 10.1103/PhysRevD.104.023503

Dai, L., & Venumadhav, T. 2017, arXiv:1702.04724. https://arxiv.org/abs/1702.04724

Dai, L., Zackay, B., Venumadhav, T., Roulet, J., & Zaldarriaga, M. 2020, arXiv:2007.12709. https://arxiv.org/abs/2007.12709

Dal Canton, T., et al. 2014, Phys. Rev. D, 90, 082004, doi: 10.1103/PhysRevD.90.082004

Davies, G. S., Dent, T., Tápai, M., et al. 2020, Phys. Rev. D, 102, 022004, doi: 10.1103/PhysRevD.102.022004

24

Diego, J. M., Hannuksela, O. A., Kelly, P. L., et al. 2019, Astron. Astrophys., 627, A130, doi: 10.1051/0004-6361/201935490

Drago, M., et al. 2020, doi: 10.1016/j.softx.2021.100678

Ezquiaga, J. M., Holz, D. E., Hu, W., Lagos, M., & Wald, R. M. 2021, Phys. Rev. D, 103, 064047, doi: 10.1103/PhysRevD.103.064047

Ezquiaga, J. M., Hu, W., & Lo, R. K. L. 2023, Phys. Rev. D, 108, 103520, doi: 10.1103/PhysRevD.108.103520

Ezquiaga, J. M., & Lo, R. K. L. 2023, phazap: gravitational wave phase reconstruction for low-latency identification of strongly lensed signals, https://github.com/ezquiaga/phazap

Ezquiaga, J. M., Lo, R. K. L., & Vujeva, L. 2025. https://arxiv.org/abs/2503.22648

Ezquiaga, J. M., & Zumalacárregui, M. 2020, Phys. Rev. D, 102, 124048, doi: 10.1103/PhysRevD.102.124048

Farah, A., Ezquiaga, J. M., Fishbach, M., & Holz, D. 2025. https://arxiv.org/abs/2507.07964

Finke, A., Foffa, S., Iacovelli, F., Maggiore, M., & Mancarella, M. 2021, Phys. Rev. D, 104, 084057, doi: 10.1103/PhysRevD.104.084057

Fishbach, M., Farr, W. M., & Holz, D. E. 2020, Astrophys. J. Lett., 891, L31, doi: 10.3847/2041-8213/ab77c9

Fishbach, M., Holz, D. E., & Farr, W. M. 2018, Astrophys. J. Lett., 863, L41, doi: 10.3847/2041-8213/aad800

Goyal, S., Haris, K., Mehta, A. K., & Ajith, P. 2021, Phys. Rev. D, 103, 024038, doi: 10.1103/PhysRevD.103.024038

Goyal, S., Vijaykumar, A., Ezquiaga, J. M., & Zumalacarregui, M. 2023, Phys. Rev. D, 108, 024052, doi: 10.1103/PhysRevD.108.024052

Gupta, A., et al. 2024, arXiv:2405.02197. https://arxiv.org/abs/2405.02197

Hanna, C., et al. 2020, Phys. Rev. D, 101, 022003, doi: 10.1103/PhysRevD.101.022003

Hannuksela, O. A., Collett, T. E., Çalışkan, M., & Li, T. G. 2020, Monthly Notices of the Royal Astronomical Society, 498, 3395, doi: 10.1093/mnras/staa2565

Hannuksela, O. A., Haris, K., Ng, K. K. Y., et al. 2019, Astrophys. J. Lett., 874, L2, doi: 10.3847/2041-8213/ab0c0f

Haris, K., Mehta, A. K., Kumar, S., Venumadhav, T., & Ajith, P. 2018, arXiv:1807.07062. https://arxiv.org/abs/1807.07062

Harris, C. R., et al. 2020, Nature, 585, 357, doi: 10.1038/s41586-020-2649-2

Hourihane, S., Chatziioannou, K., Wijngaarden, M., et al. 2022, Phys. Rev. D, 106, 042006, doi: 10.1103/PhysRevD.106.042006

Hoy, C., & Raymond, V. 2021, SoftwareX, 15, 100765, doi: 10.1016/j.softx.2021.100765

Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90, doi: 10.1109/MCSE.2007.55

Jana, S., Kapadia, S. J., Venumadhav, T., & Ajith, P. 2023, Phys. Rev. Lett., 130, 261401, doi: 10.1103/PhysRevLett.130.261401

Jana, S., Kapadia, S. J., Venumadhav, T., More, S., & Ajith, P. 2024a, arxiv:2405.17805. https://arxiv.org/abs/2405.17805

—. 2024b, arxiv:2408.05290. https://arxiv.org/abs/2408.05290

Janquart, J., Hannuksela, O. A., K., H., & Van Den Broeck, C. 2021a, Mon. Not. Roy. Astron. Soc., 506, 5430, doi: 10.1093/mnras/stab1991

Janquart, J., Haris, K., Hannuksela, O. A., & Van Den Broeck, C. 2023a, Mon. Not. Roy. Astron. Soc., 526, 3088, doi: 10.1093/mnras/stad2838

Janquart, J., More, A., & Van Den Broeck, C. 2022a, Mon. Not. Roy. Astron. Soc., 519, 2046, doi: 10.1093/mnras/stac3660

Janquart, J., Seo, E., Hannuksela, O. A., Li, T. G. F., & Broeck, C. V. D. 2021b, Astrophys. J. Lett., 923, L1, doi: 10.3847/2041-8213/ac3bcf

Janquart, J., Uronen, L. E., Heynen, J., Haris, K., & Hannuksela, O. 2022b, GOLUM: a software for rapid strongly-lensed gravitational wave parameter estimation, https://github.com/lemnis12/golum

Janquart, J., et al. 2023b, Mon. Not. Roy. Astron. Soc., 526, 3832, doi: 10.1093/mnras/stad2909

Janquart, J., Keitel, D., Lo, R. K. L., et al. 2025, Monthly Notices of the Royal Astronomical Society, 537, 1001, doi: 10.1093/mnras/staf049

Jung, S., & Shin, C. S. 2019, Phys. Rev. Lett., 122, 041103, doi: 10.1103/PhysRevLett.122.041103

Kass, R. E., & Raftery, A. E. 1995, J. Am. Statist. Assoc., 90, 773, doi: 10.1080/01621459.1995.10476572

Leong, S. H. W., Janquart, J., Sharma, A. K., et al. 2025, Astrophys. J. Lett., 979, L27, doi: 10.3847/2041-8213/ad9ead

Levan, A. J., Gompertz, B. P., Smith, G. P., et al. 2025, Phil. Trans. Roy. Soc. Lond. A, 383, 20240122, doi: 10.1098/rsta.2024.0122

Li, A. K. Y., Chan, J. C. L., Fong, H., et al. 2025a, Mon. Not. Roy. Astron. Soc., 542, 9981010, doi: 10.1093/mnras/staf1259

—. 2025b, TESLA-X: gstLAL based TargetEd Subthreshold Lensing seArch, https://git.ligo.org/alvinli/tesla/

Li, A. K. Y., Lo, R. K. L., Sachdev, S., et al. 2023, Phys. Rev. D, 107, 123014, doi: 10.1103/PhysRevD.107.123014

Li, S.-S., Mao, S., Zhao, Y., & Lu, Y. 2018, Mon. Not. Roy. Astron. Soc., 476, 2220, doi: 10.1093/mnras/sty411

Liao, K., Fan, X.-L., Ding, X.-H., Biesiada, M., & Zhu, Z.-H. 2017, Nature Commun., 8, 1148, doi: 10.1038/s41467-017-01152-9

LIGO Scientific Collaboration, Virgo Collaboration, & KAGRA Collaboration. 2018, LVK Algorithm Library - LALSuite, Free software (GPL), doi: 10.7935/GT1W-FZ16

Littenberg, T. B., Kanner, J. B., Cornish, N. J., & Millhouse, M. 2016, Phys. Rev. D, 94, 044050, doi: 10.1103/PhysRevD.94.044050

Liu, A., & Kim, K. 2024, Physical Review D, 110, 123008

Liu, A., Wong, I. C. F., Leong, S. H. W., et al. 2023, Mon. Not. Roy. Astron. Soc., 525, 4149, doi: 10.1093/mnras/stad1302

Liu, X., Magana Hernandez, I., & Creighton, J. 2021, Astrophys. J., 908, 97, doi: 10.3847/1538-4357/abd7eb

Lo, R. K. L., & Magana Hernandez, I. 2023, Phys. Rev. D, 107, 123015, doi: 10.1103/PhysRevD.107.123015

Lo, R. K. L., Vujeva, L., Ezquiaga, J. M., & Chan, J. C. L. 2025, Phys. Rev. Lett., 134, 151401, doi: 10.1103/PhysRevLett.134.151401

Madau, P., & Dickinson, M. 2014, Ann. Rev. Astron. Astrophys., 52, 415, doi: 10.1146/annurev-astro-081811-125615

Magare, S., Kapadia, S. J., More, A., et al. 2023, Astrophys. J. Lett., 955, L31, doi: 10.3847/2041-8213/acf668

McIsaac, C., Keitel, D., Collett, T., et al. 2020, Phys. Rev. D, 102, 084031, doi: 10.1103/PhysRevD.102.084031

Messick, C., Blackburn, K., Brady, P., et al. 2017, Phys. Rev. D, 95, 042001, doi: 10.1103/PhysRevD.95.042001

Messick, C., et al. 2017, Phys. Rev. D, 95, 042001, doi: 10.1103/PhysRevD.95.042001

Mishra, A., Krishnendu, N. V., & Ganguly, A. 2023, arXiv:2311.08446. https://arxiv.org/abs/2311.08446

Mishra, A., Meena, A. K., More, A., & Bose, S. 2024, Mon. Not. Roy. Astron. Soc., 531, 764, doi: 10.1093/mnras/stae836

Mishra, A., Meena, A. K., More, A., Bose, S., & Bagla, J. S. 2021, Mon. Not. Roy. Astron. Soc., 508, 4869, doi: 10.1093/mnras/stab2875

More, A., & More, S. 2022, Mon. Not. Roy. Astron. Soc., 515, 1044, doi: 10.1093/mnras/stac1704

Mukherjee, O., & Nemiroff, R. J. 2023, Mon. Not. Roy. Astron. Soc., 527, L132, doi: 10.1093/mnrasl/slad158

Mukherjee, S., Broadhurst, T., Diego, J. M., Silk, J., & Smoot, G. F. 2021a, Mon. Not. Roy. Astron. Soc., 506, 3751, doi: 10.1093/mnras/stab1980

—. 2021b, Mon. Not. Roy. Astron. Soc., 501, 2451, doi: 10.1093/mnras/staa3813

Mukherjee, S., Wandelt, B. D., & Silk, J. 2020, Mon. Not. Roy. Astron. Soc., 494, 1956, doi: 10.1093/mnras/staa827

Narayan, P., Johnson-McDaniel, N. K., & Gupta, A. 2024. https://arxiv.org/abs/2412.13132

Narola, H., Janquart, J., Haegel, L., et al. 2024, Phys. Rev. D, 109, 084064, doi: 10.1103/PhysRevD.109.084064

Ng, K. K. Y., Wong, K. W. K., Broadhurst, T., & Li, T. G. F. 2018, Phys. Rev. D, 97, 023012, doi: 10.1103/PhysRevD.97.023012

Ng, L. C. Y., Janquart, J., Phurailatpam, H., et al. 2024, Mon. Not. R. Astron. Soc., doi: 10.1093/mnras/staf874

Nitz, A. H., Dent, T., Dal Canton, T., Fairhurst, S., & Brown, D. A. 2017, Astrophys. J., 849, 118, doi: 10.3847/1538-4357/aa8f50

Oguri, M. 2018, Mon. Not. Roy. Astron. Soc., 480, 3842, doi: 10.1093/mnras/sty2145

Owen, B. J. 1996, Phys. Rev. D, 53, 6749, doi: 10.1103/PhysRevD.53.6749

Pankow, C., et al. 2018, Phys. Rev. D, 98, 084016, doi: 10.1103/PhysRevD.98.084016

Paynter, J., Webster, R., & Thrane, E. 2021, Nature Astron., 5, 560, doi: 10.1038/s41550-021-01307-1

Phurailatpam, H., More, A., Narola, H., et al. 2024, arxiv:2407.07526. https://arxiv.org/abs/2407.07526

Pratten, G., et al. 2021, Phys. Rev. D, 103, 104056, doi: 10.1103/PhysRevD.103.104056

Price-Whelan, A. M., et al. 2022, Astrophys. J., 935, 167, doi: 10.3847/1538-4357/ac7c74

Ray, A., Banagiri, S., Thrane, E., & Lasky, P. D. 2025. https://arxiv.org/abs/2510.07228

Robertson, A., Smith, G. P., Massey, R., et al. 2020, Mon. Not. Roy. Astron. Soc., 495, 3727, doi: 10.1093/mnras/staa1429

Robinet, F., Arnaud, N., Leroy, N., et al. 2020, SoftwareX, 12, 100620, doi: 10.1016/j.softx.2020.100620

Romero-Shaw, I. M., et al. 2020, Mon. Not. Roy. Astron. Soc., 499, 3295, doi: 10.1093/mnras/staa2850

Ryczanowski, D., Smith, G. P., Bianconi, M., et al. 2020, Mon. Not. Roy. Astron. Soc., 495, 1666, doi: 10.1093/mnras/staa1274

Sachdev, S., et al. 2019. https://arxiv.org/abs/1901.08580

Savastano, S., Tambalo, G., Villarrubia-Rojo, H., & Zumalacarregui, M. 2023, Phys. Rev. D, 108, 103532, doi: 10.1103/PhysRevD.108.103532

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses, Astronomy and Astrophysics Library (Springer), doi: 10.1007/978-3-662-03758-4

Seo, E., Hannuksela, O. A., & Li, T. G. 2022, The Astrophysical Journal, 932, 50

Seo, E., Shan, X., Janquart, J., et al. 2025, The Astrophysical Journal, 988, 159

Shan, X., Chen, X., Hu, B., & Li, G. 2024, Sci. China Phys. Mech. Astron., 67, 269511, doi: 10.1007/s11433-023-2334-9

Shan, X., Hu, B., Chen, X., & Cai, R.-G. 2025a, Nature Astron., 9, 916, doi: 10.1038/s41550-025-02519-5

Shan, X., Li, G., Chen, X., et al. 2025b, Sci. China Phys. Mech. Astron., 68, 219512, doi: 10.1007/s11433-024-2502-1

Shan, X., Li, G., Chen, X., Zheng, W., & Zhao, W. 2023, Sci. China Phys. Mech. Astron., 66, 239511, doi: 10.1007/s11433-022-1985-3

Smith, G. P., Jauzac, M., Veitch, J., et al. 2018, Mon. Not. Roy. Astron. Soc., 475, 3823, doi: 10.1093/mnras/sty031

Smith, G. P., et al. 2025, Phil. Trans. Roy. Soc. Lond. A, 383, 20240134, doi: 10.1098/rsta.2024.0134

Smith, R. J. E., Ashton, G., Vajpeyi, A., & Talbot, C. 2020, Mon. Not. Roy. Astron. Soc., 498, 4492, doi: 10.1093/mnras/staa2483

Soni, S., et al. 2025, Class. Quant. Grav., 42, 085016, doi: 10.1088/1361-6382/adc4b6

Speagle, J. S. 2020, MNRAS, 493, 3132, doi: 10.1093/mnras/staa278

Takahashi, R., & Nakamura, T. 2003, Astrophys. J., 595, 1039, doi: 10.1086/377430

Talbot, C., Smith, R., Thrane, E., & Poole, G. B. 2019, Phys. Rev. D, 100, 043030, doi: 10.1103/PhysRevD.100.043030

Tambalo, G., Zumalacárregui, M., Dai, L., & Cheung, M. H.-Y. 2023, Phys. Rev. D, 108, 103529, doi: 10.1103/PhysRevD.108.103529

Thompson, J. E., Hamilton, E., London, L., et al. 2024, Phys. Rev. D, 109, 063012, doi: 10.1103/PhysRevD.109.063012

Tinker, J. L., Kravtsov, A. V., Klypin, A., et al. 2008, Astrophys. J., 688, 709, doi: 10.1086/591439

Ubach, H., Gieles, M., & Miralda-Escudé, J. 2025. https://arxiv.org/abs/2505.04794

Udall, R., Ashton, G., & Yarbrough, Z. 2024, CBC Workflow, https://git.ligo.org/cbc/projects/cbcflow

Uronen, L. E., Li, T., Janquart, J., et al. 2024. https://arxiv.org/abs/2406.14257

Urrutia, J., & Vaskonen, V. 2021, Mon. Not. Roy. Astron. Soc., 509, 1358, doi: 10.1093/mnras/stab3118

Usman, S. A., et al. 2016, Class. Quant. Grav., 33, 215004, doi: 10.1088/0264-9381/33/21/215004

Varma, V., Field, S. E., Scheel, M. A., et al. 2019, Phys. Rev. Research., 1, 033015, doi: 10.1103/PhysRevResearch.1.033015

Vijaykumar, A., Mehta, A. K., & Ganguly, A. 2023, Phys. Rev. D, 108, 043036, doi: 10.1103/PhysRevD.108.043036

Virtanen, P., et al. 2020, Nature Meth., 17, 261, doi: 10.1038/s41592-019-0686-2

Vujeva, L., Ezquiaga, J. M., Lo, R. K. L., & Chan, J. C. L. 2025. https://arxiv.org/abs/2501.02096

Wang, Y., Lo, R. K. L., Li, A. K. Y., & Chen, Y. 2021, Phys. Rev. D, 103, 104055, doi: 10.1103/PhysRevD.103.104055

Waskom, M. 2021, J. Open Source Softw., 6, doi: 10.21105/joss.03021

Wempe, E., Koopmans, L. V., Wierda, A. R. A., Hannuksela, O. A., & Van Den Broeck, C. 2024, Monthly Notices of the Royal Astronomical Society, 530, 3368, doi: 10.1093/mnras/stae1023

Wes McKinney. 2010, in Proceedings of the 9th Python in Science Conference, ed. Stéfan van der Walt & Jarrod Millman, 56 – 61, doi: 10.25080/Majora-92bf1922-00a

Wierda, A. R. A. C., Wempe, E., Hannuksela, O. A., Koopmans, L. e. V. E., & Van Den Broeck, C. 2021, Astrophys. J., 921, 154, doi: 10.3847/1538-4357/ac1bb4

Williams, D., Veitch, J., Chiofalo, M. L., et al. 2023, J. Open Source Softw., 8, 4170, doi: 10.21105/joss.04170

Wong, H. W. Y., Chan, L. W. L., Wong, I. C. F., Lo, R. K. L., & Li, T. G. F. 2021, arXiv:2112.05932. https://arxiv.org/abs/2112.05932

Wright, M., & Hendry, M. 2021, Astrophys. J., 935, 68, doi: 10.3847/1538-4357/ac7ec2

Wright, M., Janquart, J., & Johnson-McDaniel, N. 2024, arXiv:2403.08957. https://arxiv.org/abs/2403.08957

Wright, M., Janquart, J., et al. 2025a, Microlensing Fourier Convention Inconsistency

Wright, M., Liu, A., Wong, I. C. F., et al. 2021, Gravelamps: Gravitiational Wave Lensing Mass Profile Model Selection, https://git.ligo.org/mick.wright/gravelamps

Wright, M., et al. 2025b. https://arxiv.org/abs/2507.20256

Xu, F., Ezquiaga, J. M., & Holz, D. E. 2022, Astrophys. J., 929, 9, doi: 10.3847/1538-4357/ac58f8

Yeung, S. M. C., Cheung, M. H. Y., Seo, E., et al. 2023, Mon. Not. Roy. Astron. Soc., 526, 2230, doi: 10.1093/mnras/stad2772