

Strong Lensing of Gravitational Waves and Tests of General Relativity

A Thesis

Submitted to the

Tata Institute of Fundamental Research, Mumbai
for the degree of Doctor of Philosophy
in Physics

by

Srashti Goyal

International Centre for Theoretical Sciences,
Tata Institute of Fundamental Research, Bengaluru

October 11, 2023

Declaration

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions.

The work was done under the guidance of Professor Parameswaran Ajith, at the International Centre for Theoretical Sciences, Tata Institute of Fundamental Research, Bangalore.



Srashti Goyal

In my capacity as supervisor of the candidate's thesis, I certify that the above statements are true to the best of my knowledge.



Parameswaran Ajith

Date: 11 October, 2023

Acknowledgements

I am deeply grateful to my advisor, Ajith Parameswaran, for his immense support and guidance throughout the entire process of my thesis. His expertise and mentorship have been invaluable, and I am truly thankful for his contributions. I would also like to extend my heartfelt thanks to my collaborator, Shasvath Kapadia, who has relentlessly supported me with valuable advice and with whom I thoroughly enjoyed working.

Behind the scenes, the astrophysical relativity group at my institute, ICTS deserves special recognition. Their vibrant discussions and enthusiasm for GW research, even during the challenging times of the COVID-19 pandemic, played a crucial role in keeping my interest alive. I am particularly grateful for the support I received from (ex-) ICTS members such as Haris MK, Ajit Mehta, Pinak Mandal, and others, whose assistance has been immensely valuable throughout this journey. I am inspired by Prof. Bala Iyer, who has overseen all the group activities and helped in maintaining the research momentum. I also thank Vishal Vasani, Pravabati Chingambam and Rama Govindarajan for their valuable suggestions and support while monitoring the progress of this thesis. In addition to the beautiful campus, the excellent work culture at ICTS, with all professors, colleagues, and staff members, has made my PhD experience smooth and joyful. I extend my humble gratitude to the entire ICTS family for their contributions.

Furthermore, I would like to express my appreciation to direct collaborators, Harikrishnan D, Jose Maria Ezquiaga, Miguel Zumalacarregui, Jean-Rene Cudell, Aditya Vijaykumar, etc. Without their contributions, this thesis would not be complete. A significant portion of

this thesis was accomplished as a member of the LIGO Scientific Collaboration, closely connected to the lensing group. I am grateful to the collaboration for providing computational resources and the opportunity to work with researchers from around the world. I would like to give special mention to Otto Hannuksela, Tjonnie Li, and David Keitel for their valuable suggestions and for overseeing the administration of the group.

Finally, I would like to thank my family, friends and teachers for their unwavering love and support. I am especially thankful to all my yoga teachers whose motivation and guidance have helped me maintain a healthy work-life balance and navigate through the challenges that came along the way.

List of Figures

- 1.1 Top: GW strain time series of GW₁₅₀₉₁₄ waveform template projected onto H1 detector, showing the inspiral-merger-ringdown phases. Bottom: The BH separation in units of Schwarzschild radii (a/R_s), and the effective relative velocity v/c . Credits: LIGO scientific collaboration. 32
- 1.2 Measured noise PSDs in the three observing runs (O₁₋₃) of LIGO H1 detector, and design sensitivities for the future upgrades (Advanced LIGO, A+). Credits: Cahillane & Mansell [1] 34
- 1.3 Masses of CBCs from the LVK latest transient catalog, GWTC-3 [2]. Each circle represents a different compact object and the vertical scale indicates the mass as a multiple of the mass of our Sun. Blue circles represent BHs and orange circles represent NS. Half-blue / half-orange mixed circles are compact objects whose classification is uncertain. Each merger involves three compact objects: two merging objects and the final resulting object. Credits: LVK Collaboration/Frank Elavsky, Aaron Geller/Northwestern. 35
- 1.4 GW spectrum from its various kind of sources and frequency bands of the current and future GW detectors. Credits: P. Ajith. 36
- 1.5 Summary of EM counterparts of BNS and NSBH mergers and their dependence on the viewing angle with respect to the axis of the GRB jet. The kilonova, in contrast to the GRB and its afterglow, is relatively isotropic and thus represents the most promising counterpart for the majority of GW-detected mergers. Credits: Metzger and Berger (2012) [3], copyright by AAS 37

- 1.6 Simultaneous detection of GW signal by LIGO-Virgo detectors and a short-gamma ray burst signal by Fermi satellite from BNS merger GW170817. Credits: Astrobites magazine. 38
- 1.7 GW170817 measurement of H_0 . Marginalized posterior density for the Hubble constant H_0 (blue curve). Constraints at $1 - \sigma$ and $2 - \sigma$ from Planck (Planck Collaboration et al. 2016) and SHoES (Riess et al. 2016) are shown in green and orange. Credits: LVK Collaboration [4]. 39
- 1.8 Schematic diagram showing the gravitational potential and curvature scale probed by different tests of GR. The horizontal axis shows the curvature scale $R = \sqrt{M/L^3}$ due to the presence of a mass M of size L , while the vertical axis shows the compactness or the dimensionless gravitational potential, $\Phi = M/L$ probed by different experiments. For example, observation of the double binary pulsar (total mass of $2.6M_\odot$) and orbital separation of $\sim 10^6 km$ probes the curvature scale of $R \sim 2 \times 10^{-9} km^{-1}$ and compactness of $\Phi \sim 4 \times 10^{-6}$. In contrast, a black hole binary of total mass $10M_\odot$ at the time of merger probes a curvature scale of $R \sim 2.4 \times 10^{-2} km^{-1}$ and compactness of $\Phi \sim 0.5$. Credits: References [5, 6] 40
- 1.9 IMR consistency tests with GW signals. The final mass and spin posteriors from the inspiral and post-inspiral phases should match if the GR and its approximations capture the GW signal correctly. Credits: LVK collaboration [7]. 41
- 1.10 No-hair theorem of Black-holes. Credits: Astrobites magazine. 42
- 1.11 Types of EM lensing: strong, weak and micro lensing. (left) Observation of multiple images of a quasar are formed due to strong lensing by a galaxy. (middle) Distortions of the background galaxies caused due to weak lensing by a galaxy cluster, imaged by Hubble Space Telescope. (right) Schematic of the time-varying brightness of a star caused due to micro-lensing by a compact object. Credits: ESA, Hubble, NASA 44

- 1.12 Schematic representation of the GW lensing. On the top, a given signal is strongly lensed by a galaxy lens producing multiple images with different amplitudes, arrival times and phases. On the bottom, small compact lenses produce interference effects that distort the detected waveform. Credits: LVK collaboration 45
- 1.13 Angles involved in a thin gravitational lens system for the deflection of radiation in G.O limit. The distances from the observer to the lens and the source are D_L and D_S , respectively. The distance between the lens and source is D_{LS} . β is the angular position of the unlensed source, $\hat{\alpha}$ denotes the deflection angle and θ is the angular position of one of the images. The red and blue rectangles represent the parallel lens and source planes at D_L and D_S , respectively, which are perpendicular to the observer-lens axis (dotted) Credits: Reference [8]. 46
- 1.14 (blue) red region signifies the constraints on merger rate density of BBHs when (not) accounting for non-observation of lensing of the LVK O3a events. The high-redshift regions are more constrained due to the lensing magnification. Credits: LVK collaboration, O3a lensing [9] 50
- 1.15 An illustration of a sky localization of a quadruply lensed GW. The individual (colour) and the combined (black) sky localizations at 90 percent confidence are shown. Each lensed GW essentially gives us a new set of detectors with which to localize the event in the sky, allowing for improved sky localization. A dedicated follow-up of the narrowed sky region would then allow us to search for the lensed host galaxy from which the GW originates. Credits: Hannuksela et al. [10] 50
- 1.16 Estimation of lensing magnification for explaining the observed mass-gap nature of a compact nature, as a result of bias in redshift measurements. Credits: Pang et al [11]. 51

1.17 GW lensing beyond general relativity. A GW emitted by a binary black hole splits into its propagation eigenstates (waveforms in color) when it enters the region near the lens where modified gravity backgrounds are relevant. Depending on the time delays between the propagation eigenstates the signal detected could be scrambled or echoed. If the GW travels closer than the Einstein radius, multiple images could be formed as indicated by the grey solid trajectories. Credits: Eqzuiaga and Zummalacaragui [12]

52

2.1 *Top Panels:* A pair of lensed GW events detected by the H1 (Hanford) interferometer at design sensitivity. These events have time-frequency tracks with similar shapes. However, the signal energy in different time frequency bins along their tracks differ with respect to each other. *Bottom Panels:* A pair of unlensed GW events projected detected by the H1 interferometer at design sensitivity. These events have time-frequency tracks whose shapes are significantly different. 63

2.2 *Top Panels:* Bayestar skymaps of a pair of lensed events detected by the H1 (Hanford), L1 (Livingston), V1 (Virgo) network at design sensitivity. The skymaps of these events overlap. *Bottom Panels:* Bayestar skymaps of a pair of unlensed events detected by the H1, L1, V1 network at design sensitivity. The skymaps of these events do not overlap. 64

2.3 A visual representation of the overall flow of our ML classification scheme. Note that, in principle, one could have avoided the step that trains a second XGBOOST algorithm on features derived exclusively from the skymaps, and instead just used one XGBOOST that jointly trains on features from the skymaps and the outputs of the DENSENET algorithms. We found that both methods give similar results. We therefore choose to include the additional XGBOOST because it facilitates a stepwise analysis of the outputs of the individual components of the overall flow, trained separately on intrinsic and extrinsic parameters of the candidate pairs.

68

2.4 ROCs for DENSENET models trained on lensed and unlensed pairs of superimposed Q-transforms, for different cross-validation subsets of the DSTRV training set. ROCs for models trained on Q-transforms corresponding to individual detectors are evaluated, in addition to ROCs pertaining to the XGBOOST model trained on the outputs of the individual DENSENET models. For comparison, the ROC for the PO statistic that uses parameter estimation posteriors on the component masses, m_1, m_2 , is also plotted. At low false positive probabilities, the individual DensetNet models perform comparably to the posterior-overlap statistic. On the other hand, the XGBOOST model produces efficiencies that are 1.5 – 2 times better than the PO statistic at low FPPs, although there is some variation in the ROCs when the training set is changed, caused by small-number statistics. These improvements at low FPPs must therefore be interpreted with some caution.

75

2.5 ROCs for the XGBOOST model trained on metrics derived from pairs of Bayestar localisation skymaps, for different cross-validation subsets of the DSTRV training set. For comparison, the ROC for the PO statistic that uses parameter estimation posteriors on the skylocation coordinates, α, δ , is also plotted. The XGBOOST performs almost as well as the PO statistic, at low false positive probabilities.

76

- 2.6 ROCs for the overall classifier, for different cross-validation subsets of the DSTRV training set. Note that the output of the overall classifier is the output of the XGBOOST model trained on the outputs of the three DENSENET models pertaining to H1, L1 and V1, as well as the output of the first XGBOOST model trained on Bayestar skymaps. At low false positive probabilities, the classifier performs almost identically to the PO statistic, with mild variation in the ROCs when the training data set is varied. 77
- 3.1 The FPP of each lensed candidate pair constructed from the set of GW events that exceed an astrophysical probability [13, 14] threshold of 0.5, as evaluated using the PO and ML classification statistics. Orange dashed lines that correspond to an FPP threshold of 10^{-2} , are also placed. Pairs whose PO statistic-based or ML-based FPP fall below this threshold are selected for additional joint PE analyses. $< 10^{-6}$ has been mapped to an FPP of 1, which is reflected in the gap along the vertical axis between 0.4 and 1. 85
- 3.2 Posteriors obtained using the IMRPhenomXPHM waveform for GW191103 (blue) and GW191105 (orange). The overlap in the extrinsic parameters (e.g. sky location) is much larger than that for the intrinsic parameters (e.g. detector-frame chirp mass and spins). 87
- 3.3 The top five candidate strong lensing pairs from the PO analysis considering all the event pairs found based on the O3 data [2]. The dashed lines correspond to the 1σ and 2σ confidence levels for the combined PO statistic ($\mathcal{B}^{\text{overlap}} \times \mathcal{R}^{\text{gal}}$) with different lensing models computed from the background simulations. 88
- 3.4 Skymap of the event GW191105 from GraceDB created using LALInference (top) and from GWOSC created using BILBY after deglitching the Virgo data (bottom). The LALInference skymaps are narrower as compared to BILBY ones, likely due to the glitch.

- 3.5 Q-transform (or time-frequency maps) images that are input to ML QTs of the LENSID pipeline for the events GW₁₉₁₁₀₃ (top panel) and GW₁₉₁₁₀₅ (bottom panel). The chirping feature for GW₁₉₁₁₀₅ is broken in both the LIGO detectors, whereas for GW₁₉₁₁₀₃ the chirp signal is fairly visible in Hanford, and not so visible in Livingston. 90
- 3.6 Chirp mass estimates from PE and matched-filter-based searches for the GWTC-3 events. About 10 – 15% of events are significantly biased – the mean value of the detector-frame chirp mass lying outside the 90% confidence of the corresponding PE posteriors (vertical error bars) and matched-filter chirp mass (horizontal error bars) coming from Eq. 3.2. If these error bars cross the diagonal, then those events should not be considered as significantly biased. The lower chirp mass binaries are less biased than the high ones due to their longer inspiral in the frequency band of LIGO-Virgo detectors. 95
- 3.7 \mathcal{R}^{gal} statistic as a function of the time delay between the events, assuming the O₃ observation time. This is a model-dependent statistic – the galaxy lens is assumed to be SIE, and the lens parameters are fit to the SDSS catalog. Lens redshifts are also assumed to follow the SDSS catalog, while source masses assume a power-law + peak model. 96
- 3.8 Comparison of the rapid identification statistics (dashed) against the PE-based PO method (solid). There is only a partial loss in efficiency using these statistics relative to PO. 97
- 3.9 Inverse cumulative distribution function of the network SNRs. The sub-threshold triggers have lower SNRs as compared to super-threshold ones but all of them have $\text{SNR} > 7$. In fact, 40% of the subthreshold events have $\text{SNR} > 8$. 99
- 3.10 Mean chirp mass estimates for the O₃ super-sub event pairs and their $\mathcal{B}^{\text{masses}}$. As expected, the majority of events with large $\mathcal{B}^{\text{masses}}$ lie along the diagonal. The coefficient's value decreases for events situated away from the diagonal. 100

- 3.11 Inverse cumulative distribution function of the statistics for the O₃ super-sub pairs. The dashed lines correspond to the top 5 percentile of the estimated values, which is found to be at $\mathcal{R}^{\text{gal}} = 2.95$, $\mathcal{B}^{\text{sky}} = 0.47$, and $\mathcal{B}^{\text{masses}} = 0.87$ respectively. Only one event lies within this percentile, across all three statistics which is shown as a cross-mark (\times). 101
- 3.12 Statistics for the individual super-sub pairs which were found by TESLA during the LVK O₃ lensing searches [15]. Black dashed lines represent the statistic value corresponding to the top 5 percentile of all the pairs. An obvious, solitary, outlier is identified in the top-right corner of the plot. In addition to being in the top 5th percentile of the two statistics on the x and y axes, it is also in the top 5th percentile of the statistic on the color bar (see the dashed black line there). We also highlight the triggers corresponding to the target events GW191230 (green), GW191105 (blue) and GW191103 (black). 102
- 3.13 Skymaps of the individual events of the pairs (top) GW191230 - LGW200104, (bottom) GW191105-LGW191106. The dark (light) shaded region represents the 50 (90) % contour. There is a significant visual overlap between the two skymaps. This is further corroborated by the high skymap overlap statistic value. 103
- 3.14 Posteriors for GW191230_180458 (blue) and LGW200104_184028 (orange). The posteriors, though broad, have a significant overlap for both the intrinsic (left) and extrinsic (right) parameters. 105
- 4.1 The effect of various GW polarizations on a ring of test particles (tensor modes in the left, vector modes in the middle and scalar modes in the right). The wave is always traveling in the z direction. The dashed circles show the original configuration of the test particles before the arrival of the wave and the solid red/black circles and ellipses show the new position of the test particles during the two half cycles of the wave. 111

4.2 Correlation of optimal SNRs of the two images with the lensing time delay Δt between the images for each set of injections, created using the tensor (left), vector (middle) and scalar (right) polarization models. Highly magnified images correspond to shorter time delays, and will, typically, have larger SNRs. In contrast, images with much larger time delays will have low magnifications, and hence will be typically undetected. As a result, most of the detected events have time delays less than a day. 120

4.3 Marginalized posteriors of the parameters $m_1, m_2, \alpha, \sin \delta, \iota$ estimated from a lensed pair of tensor injections with tensor recovery (i.e., $\mathcal{H}_t^{[I]} - \mathcal{H}_t^{[R]}$ case). Gray lines show the injected values. Note that the posteriors estimated from the two images are overlapping and are consistent with the injected values. The images have comparable optimal SNR (8.2 for red and 9 for blue), but have a time delay of 3.31 between them. Since the sky location of the two images (with respect to the detectors) is different, the posteriors of the extrinsic parameters have some difference between them. 122

4.4 Same as Fig. 4.3, except that that the injection is performed using the tensor polarization model while parameter estimation is performed using the vector model (i.e., $\mathcal{H}_t^{[I]} - \mathcal{H}_v^{[R]}$ case). Note that the posteriors of the extrinsic parameters $\alpha, \sin \delta, \iota$, estimated from the two images, are not always overlapping and are not always consistent with the injected values. 123

4.5 Same as Fig. 4.4, except that the injection is performed using the tensor polarization model while parameter estimation is performed using the scalar model (i.e., $\mathcal{H}_t^{[I]} - \mathcal{H}_s^{[R]}$ case). Note that the posteriors of the extrinsic parameters $\alpha, \sin \delta, \iota$, estimated from the two images, are not always overlapping and are not always consistent with the injected values. 124

4.6 *Top panels:* Distribution of the Bayes factor between the “right” and “wrong” polarization hypotheses estimated from pairs of events (lensed or unlensed). Events are simulated assuming the tensor polarization hypothesis \mathcal{H}_t (left panel), vector polarization hypothesis \mathcal{H}_v (middle panel) as well as scalar polarization hypothesis \mathcal{H}_s (right panel). Each plot shows the distribution of the Bayes factors between the right and wrong polarization hypotheses (for e.g., T-V in the legends denote the Bayes factor \mathcal{B}_v^t). Note that the Bayes factors for the lensed pairs are almost always greater than the same computed from unlensed events with the same parameters. *Bottom panels:* Corresponding distribution of the ratios of the overlap factors \mathcal{B}_{U}^L assuming the “right” and “wrong” polarization hypotheses. Note that the overlap ratio is greater than 1 for 85 – 95% of events, suggesting that lensing improves the Bayes factors of the right hypothesis. 125

4.7 Correlation of the overlap ratios with the lensing time delay Δt between the images for each set of injections, created using the tensor (left), vector (middle) and scalar (right) polarization models. Different color markers show the overlap ratios between the “right” and “wrong” polarization models (for e.g., T-V denotes the overlap ratio between posteriors computed using the tensor and vector models). For the events below the black dashed lines, the PO ratio is less than one; hence lensing does not improve the polarization model selection. 126

5.1 GW polarisations (left) and detector strain (right) for a CBC ($30 + 30$) M_{\odot} with birefringent time delays $\Delta t_{12} = 5, 10, 100$ ms (top to bottom). The sky localization and detector orientation correspond to $F^+ = -0.38$, $F^{\times} = 0.71$ and LIB strain is given by Eq. 5.5

- 5.2 GR and LIB waveform amplitudes in frequency domain of GW₁₅₀₉₁₄-like CBC projected to the LIGO-Virgo detectors. The birefringence leads to additional frequency modulations and distorts the GR waveform. The magnitude of these distortions are however dependent on the two parameters: Δt_{12} and ϕ_{lens} . 138
- 5.3 Mismatch between GR and LIB waveforms for GW₁₅₀₉₁₄-like CBC (top) and GW₁₉₀₈₁₄-like CBC (bottom). left pannel: GR injection i.e. $\Delta t_{12}^{\text{inj}} = 0$ and $\phi_{\text{lens}}^{\text{inj}} = 0$. the mismatch is minimum for $\Delta t_{12}^{\text{rec}} \simeq 0$. right pannel: a LIB injection with $\Delta t_{12}^{\text{inj}} = 10$ ms and $\phi_{\text{lens}}^{\text{inj}} = \pi/5$. The mismatch is minimum at $\Delta t_{12}^{\text{rec}} \simeq \pm 10$ ms and $\phi_{\text{lens}}^{\text{rec}} \simeq \pi/5, \pi/4 + \pi/5$. 139
- 5.4 SNR dependence of Δt_{12} (ms) posteriors and the $\log \mathcal{B}_{\text{GR}}^{\text{LIB}}$ (upper-x axis) for the GW₁₅₀₉₁₄-like injections with different values of $\Delta t_{12}^{\text{inj}}$ (lower-x axis) and $\phi_{\text{lens}}^{\text{inj}} = \pi/5$. Time delays (Δt_{12}) as small as 1ms are recovered well with SNR 30 & 40 signals, and for SNR 10 signals time delays < 30 ms are not measurable. Both model selection and time delay measurements (without symmetry around $\Delta t_{12} = 0$) improve with the increase in SNR. 147
- 5.5 lens-induced birefringence (LIB) test of GWTC-3 events [16]. We show the posteriors on Δt_{12} (ms) and Bayes Factors $\log \mathcal{B}_{\text{GR}}^{\text{LIB}}$ (upper x -axis). Events with positive Bayes factors are highlighted in red. 149
- 5.6 Bayes factors distributions for GW₁₉₀₅₂₁-like CBC, calculated by doing PE with both the hypothesis, for ~ 100 GR injections from the GW₁₉₀₅₂₁ posteriors in different realisations of gaussian noise. The false alarm probability for the observed $\log \mathcal{B}_{\text{GR}}^{\text{LIB}} = 3.2$ is found to be 0.48. 149
- 5.7 GW₁₉₀₅₂₁($\log \mathcal{B}_{\text{GR}}^{\text{LIB}} = 3.21$) GR v/s LIB posteriors. *MaP* (maximum a posteriori) waveforms under GR and LIB ($\Delta t_{12} = 9.51$ ms, $\phi_{\text{lens}} = 0.06$ rad) hypothesis with the whitened strain as observed at the LIGO-Virgo detectors. 150

5.8 Birefringence optical depth for the phenomenological models considered here, using the parameters correspond to the 95% c.l. limit compatible with the non-observation of LIB. The dark (light) gray shaded histograms show the binned redshift distribution of analysed (all) GWTC-3 events. See Sec. 5.2.4 for details. 153

5.9 95% c.l. constraints on the parameters of a quartic Horndeski theory [12] using the lens-induced birefringence (LIB) test. Shaded regions are excluded according to GWTC-3 (this work, blue solid), GW170817 [17, 18, 19, 20] (green dashed) and GW190521 assuming an AGN binary [21] (red dotted, see Fig. 5.10). The GR limit corresponds to $p_{4\phi} \rightarrow 0, \Lambda_4 \rightarrow \infty$, when the scalar field is decoupled from gravity and its derivative interactions suppressed. See sections 5.4.2, 5.4.3 for details. If GW190521 is associated to an AGN, the upper shaded region improves the overall GWTC-3 constraints for $\Lambda_4 \gtrsim 3H_0$. If we further assume a detection of LIB, then the bottom red shaded region excludes GR. For reference, we also indicate Solar System constraints (gray horizontal) and the region there the GW frequencies at LIGO-Virgo detectors are larger than the (non-linear) energy scale of the effective field theory (magenta vertical). 155

5.10 Birefringent time delay for a source near a SMBH as a function of the angle of the observer, relative to the SMBH. Each line corresponds to a different source distance, for model parameters compatible with GW170817 (see. Eq. 5.20). The horizontal line corresponds to the lower bound on $\Delta t_{12} = 9.5$ ms from the analysis of GW190521. The region between the shaded areas encompasses 95% probability for a random observer. The lowest θ represents trajectories passing at 10 Schwarzschild radii of the SMBH.

- 6.1 Corner plot of posteriors from LIB recovery of GW150914-like BBH injection with $\text{SNR} = 30$, $\Delta t_{12}^{inj} = 10$ ms, $\phi_{\text{lens}}^{inj} = \pi/5$ rad. The orange lines correspond to the injection parameter values. The ϕ_{lens} posteriors are visibly correlated with ψ , ι and ϕ , leading to uninformative 1-D posteriors of ϕ_{lens} .

List of Tables

- 3.1 Rapid analysis of the subthreshold triggers found by TESLA for the target event GW191230_180458, whose median $\mathcal{M}_c^{\text{det}} [M_\odot] = 61.68$. LGW200104_184028 favours lensing through all three statistics. The rest of the triggers show no sufficient evidence of being lensed counterparts. 103
- 3.2 Rapid analysis of the subthreshold triggers found by TESLA for the target events: GW191103 (top) and GW191105(bottom). 104
- 5.1 Constraints on the phenomenological models Eq. (5.14) and Eq. (5.15), assuming no birefringence detected for analysed (all) GWTC-3 events. 152
- 6.1 GW150914-like (top) and GW190814-like (bottom) CBC parameters used during mismatch calculations in Sec. 5.2.2 and PE injection studies in Sec. 5.3. 165

Contents

1	Introduction and overview	31
1.1	GW physics and astronomy	31
1.1.1	Introduction	31
1.1.2	GW astronomy and science	33
1.2	Gravitational lensing	43
1.2.1	The geometrical optics limit	45
1.2.2	Strong lensing of GWs	48
1.2.3	Organisation of the thesis	54
2	Rapid identification of strongly-lensed gravitational wave signals using machine learning	57
2.1	Introduction	57
2.2	The posterior overlap statistic	60
2.3	Classification with Machine Learning	61
2.3.1	Data Representation	62
2.3.2	Data Preparation	64
2.3.3	Feature Construction	65
2.3.4	Overall Flow	67
2.3.5	Machine Learning Models	67
2.3.6	Training and optimisation	72
2.4	Results	73
2.4.1	Testing and Cross-validation	73
2.4.2	ROC Plots	74
2.5	Summary and Outlook	76
3	Search for strongly-lensed gravitational waves in the third ob-	

serving run of LIGO and Virgo	81
3.1 Introduction	81
3.2 Analysis of O ₃ catalog events	83
3.2.1 Lensing time delay statistic	85
3.2.2 PO analysis of GW ₁₉₁₁₀₃ -GW ₁₉₁₁₀₅	86
3.2.3 LENSID investigation of GW ₁₉₁₁₀₃ -GW ₁₉₁₁₀₅	88
3.3 Analysis of the targeted sub-threshold search triggers	89
3.3.1 A rapid method for preliminary identification of subthreshold strongly lensed counterparts to su- perthreshold gravitational-wave events	92
3.3.2 Assessment of the method with simulations	96
3.3.3 O ₃ targeted sub-threshold search triggers	98
3.3.4 Rapid Identification of the super-sub lensed can- didates in O ₃	99
3.3.5 PO analysis of GW ₁₉₁₂₃₀ -LGW ₂₀₀₁₀₄	103
3.4 Summary and Conclusions	106
4 Testing the nature of GW polarisations with strongly-lensed signals	109
4.1 Introduction	109
4.2 GW polarizations	112
4.3 Method	115
4.3.1 Model selection of polarizations	115
4.3.2 Model selection of polarizations using lensed GW events	116
4.3.3 Simulations	119
4.4 Results	122
4.5 Summary and Future Work	125
5 Lens-induced GW birefringence as a test of GR	131
5.1 Introduction	131
5.2 Method	134
5.2.1 Parameterized Lens-induced Birefringence Wave- forms	135
5.2.2 Template Mismatch Studies	137

5.2.3	Bayesian Inference	140
5.2.4	Lensing Probabilities	141
5.2.5	Beyond-Poisson statistics	145
5.3	Results	146
5.3.1	Injections	146
5.3.2	GWTC-3 Events	148
5.4	Implications	151
5.4.1	Constraints on generic LIB	152
5.4.2	GW birefringence in Horndeski theories	153
5.4.3	GW190521 as an AGN binary	156
5.5	Summary and Outlook	158
6	Conclusions and future work	161
	Annexures	164
	Appendix	164
	Bibliography	165
	Publication List	187

Abstract

Gravitational waves (GWs) offer an exciting new window to the Universe. Since the first detection of GWs from a binary black hole (BBH) merger in 2015, we have now observed about a hundred stellar mass compact binary coalescences (CBCs) with ground-based GW detectors. Similar to light, GWs could also be gravitationally lensed by intervening matter between the source and the detector. Depending on the mass of the lens and the GW wavelength, there could be either of the *strong*, *weak* and *microlensing* effects on the incoming GWs. Strong lensing would produce multiple copies of the GW signal with some time delay and relative magnification, weak lensing would (de)magnify the GW signal and induce a phase shift whereas microlensing would distort the GWs. About 0.1 – 1% of CBCs detectable by ground-based GW detectors are expected to be strongly lensed, however, till now there has been no confident detection of lensed GWs and is anticipated to be soon. This thesis will cover broadly two topics involving strongly lensed signals. a) Identifying the strongly lensed signals in the data of LIGO-Virgo GW detectors, b) Testing general relativity (GR) with lensed GWs.

In Chapter 1, I give an overview of GW astronomy. I will also describe gravitational lensing, with a special focus on GW strong lensing, where the multiple copies of GWs from a lensed CBC could appear as repeated events, separated in time. To search for these strongly lensed signals in the data, either Bayesian model selection techniques, as well as, data-driven approaches such as machine learning (ML) can be used. As the pairs of events to be analysed go as $\sim N^2$, for N observed events, the computational cost for the strong lensing identification in-

creases rapidly with more number of detected events. Hence rapid identification methods are needed for strong lensing identification, especially in the future. In Chapter 2, we develop an ML algorithm to classify a given pair of signals as lensed or unlensed and show that it performs comparably to a Bayesian model selection method called posterior overlap (PO) and is computationally much faster.

In the third observing run (O_3) of LIGO-Virgo detectors, dozens of CBCs have been detected. There is a non-negligible probability that some of them could be a strongly-lensed copies of each other, hence we search for the lensing signatures in the full O_3 data. In Chapter 3, we apply our ML algorithm to analyse all the O_3 event pairs as well as perform the PO analysis for identifying candidate strongly lensed GWs. Additionally, a lensed GW counterpart may appear below the signal-to-noise ratio (SNR) threshold used for GW detection due to demagnification (or some other reason) and can be missed by the regular CBC searches. Targeted sub-threshold searches have been performed to identify potential lensed counterparts of CBC signals detected by regular searches. However, posterior distributions of source parameters are not readily available for these sub-threshold signal candidates and therefore Bayesian methods like PO are computationally expensive for analysing large number of super-sub threshold event pairs. Hence, we develop a new rapid method to identify the most promising sub-threshold lensing counterparts of the events detected in O_3 . Overall, we find a couple of interesting candidates, but none of them are significant enough for the confident detection of strong lensing. As the detector sensitivities improve we may have a first confident detection in the coming years.

Strongly-lensed GW signals would have various astrophysical implications such as facilitating the tests of GR. In Chapter 4, we discuss our idea about how the multiple copies of GW signals due to strong lensing would help in better measurement of the GW polarisations. GR predicts only two polarisation modes, whereas alternative theories allow up to six polarisation modes of the GWs. The strain measured at each detector is a linear combination of the GW polarisations, and

therefore we are limited by the number of detectors for extracting all the six polarisations. Since strongly-lensed GW signals arrive at the detectors at different times (and hence will see different antenna responses), they provide additional linear combinations of the same GW polarisations and therefore improve our ability to distinguish between the polarisations.

The propagation of GWs near the lens is described in the weak-field limit. GR predicts that the GW polarisations travel with the speed of light and are disentangled from each other even near the lens. However, in scalar-tensor theories of gravity, additional scalar fields can interact with the gravitational potential of the lens and modify the GW propagation such that the GW polarisations mix and travel at different speeds, splitting the signal into different components. These components will then reach the detector at different times, can interfere and lead to the scrambling of GWs [12]. We call this effect lens-induced birefringence (LIB). In chapter 5, by modelling this effect phenomenologically we perform a test of GR in the Bayesian model selection framework using the latest catalog of GW events (GWTC-3). We find no significant deviation from GR and in turn, constrain the LIB probability and parameters of the alternative theories of gravity. Finally in chapter 6, we summarize the findings of this thesis along with the future research directions.

1 | Introduction and overview

1.1 GW physics and astronomy

1.1.1 Introduction

Gravitational waves (GWs) are the ripples in space-time that are caused by accelerating masses. The Newtonian notion of gravity as a force exerted by massive objects upon each other was redefined in Einstein's theory of general relativity (GR) in 1916 as a curvature of space-time induced by mass or radiation. Einstein's field equations when linearised, give rise to a wave equation predicting the existence of GWs, i.e. propagating perturbations of the space-time. In a vacuum, solutions to the resulting wave equation are simply plane waves travelling at the speed of light. In GR, in addition to being symmetric, the perturbation metric $h_{\mu\nu}$ satisfies traceless and transverse gauge conditions as a consequence of the conservation of energy and local Lorentz invariance [22]. Therefore, the GWs consist of only two degrees of freedom called the '+' and the '×' polarisations. At the leading order, the GWs strain amplitude h (norm of spatial components of $h_{\mu\nu}$) is proportional to the second time derivative of the mass quadrupole moment Q of the source, and for an equal mass binary system of total mass M (Schwarzschild radius $R_s = GM/c^2$) moving with velocity v goes roughly as [22],

$$h \sim G\ddot{Q}/(rc^4) \approx \left(\frac{GM/c^2}{r}\right) \cdot (v/c)^2 \approx \frac{R_s^2}{ra}, \quad (1.1)$$

where r is the distance to the source and a is the separation of the objects in the binary. Given a binary system with masses $30M_\odot$ BHs

each and located $400Mpc$ away, the typical GW strain amplitude is $h \sim 10^{-21}$ at the Earth. Even for the most promising astrophysical sources, GWs are incredibly weak and it took about a hundred years for their first direct detection. The two Laser Interferometer Gravitational-Wave Observatory (LIGO) [23] in USA Hanford (H1) and Livingston (L1) successfully detected GWs for the first time in 2015, from a binary black-hole (BBH) merger, called GW₁₅₀₉₁₄ [7]. These detectors measure the tiny changes in the length of two perpendicular arms of laser interferometer caused by the passing GW. To appreciate the sensitivity of the detectors, the induced change in length of the order $\Delta l = h.l \sim 10^{-18}$ m of the LIGO interferometer arms of length $l = 4$ km. When the two objects in the binary orbit around each other the energy is carried away by the GWs, which in turn reduces their angular momentum until they merge and become a heavier mass stable object. Fig. 1.1 shows the modelled GW strain of GW₁₅₀₉₁₄, along with the characteristic relative velocity and separation of the BHs.

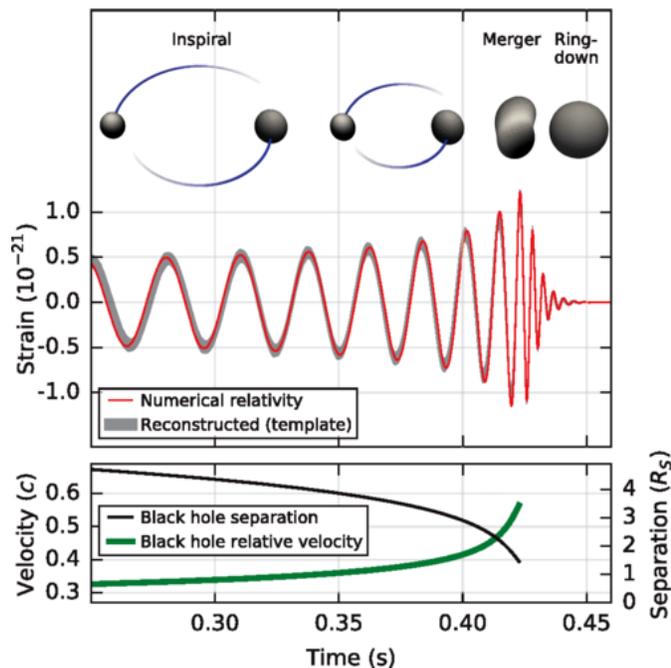


Figure 1.1: Top: GW strain time series of GW₁₅₀₉₁₄ waveform template projected onto H1 detector, showing the inspiral-merger-ringdown phases. Bottom: The BH separation in units of Schwarzschild radii (a/R_s), and the effective relative velocity v/c . Credits: LIGO scientific collaboration.

The GW waveform from a compact binary coalescence (CBC) has three phases, inspiral, merger and ringdown. For the inspiral, post-

Newtonian approximation is used to model the GW waveform and for the ringdown, perturbation theory around the final remnant object is used, whereas, for the merger phase these approximations are not valid and Einstein's field equations are solved numerically using supercomputers. Moreover, for almost all practical purposes like GW searches, semi-analytical or interpolated waveforms are constructed by combining approximation/perturbation methods and numerical relativity, so that they can be evaluated in seconds for a given set of CBC parameters.

1.1.2 *GW astronomy and science*

At present there are five ground-based GW detectors, Virgo [24] in Italy, KAGRA [25] in Japan and GEO 600 [26] in Germany in addition to the two LIGO detectors in the USA [23] mentioned before. Additionally, one more LIGO detector is being built in India [27] and is expected to join the GW detector network in a few years. A greater number of detectors in a network provide better localization of GW sources and increase the overall sensitivity for detecting GWs [28]. The ground-based detectors are sensitive in the $\sim 1 - 1000$ Hz frequency range of GWs, where they can detect various kinds of GWs signals, such as transients – short duration signals – from CBCs or supernovae explosions, continuous waves – long duration signals– from rotating neutron stars (NS) or axion clouds, and stochastic background of the population of GW sources. However, so far only GWs from CBCs have been detected.

GW signals are usually buried in the noise of detectors which makes the detection challenging. Fig. 1.2 shows the noise power spectral density (PSD, averaged over a long time) of the detectors. One of the most successful methods for GW detection is the matched-filtering technique, where the gravitational waveforms from the GW sources, like CBCs, are modelled to search for the GW signals in the data. The data is cross-matched with the waveform templates by convolution while calculating the matched-filter signal-to-noise ratio (SNR) as the noise-weighted inner product of the template and the data. Matched-

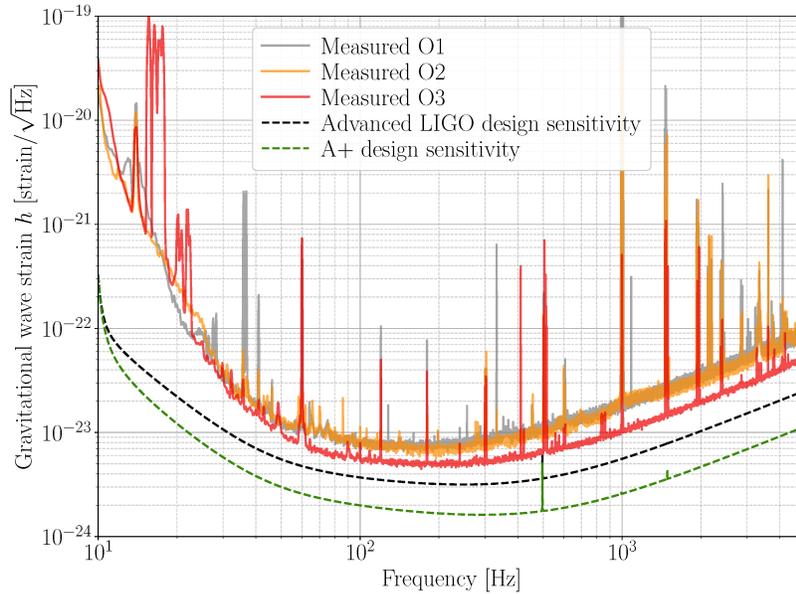
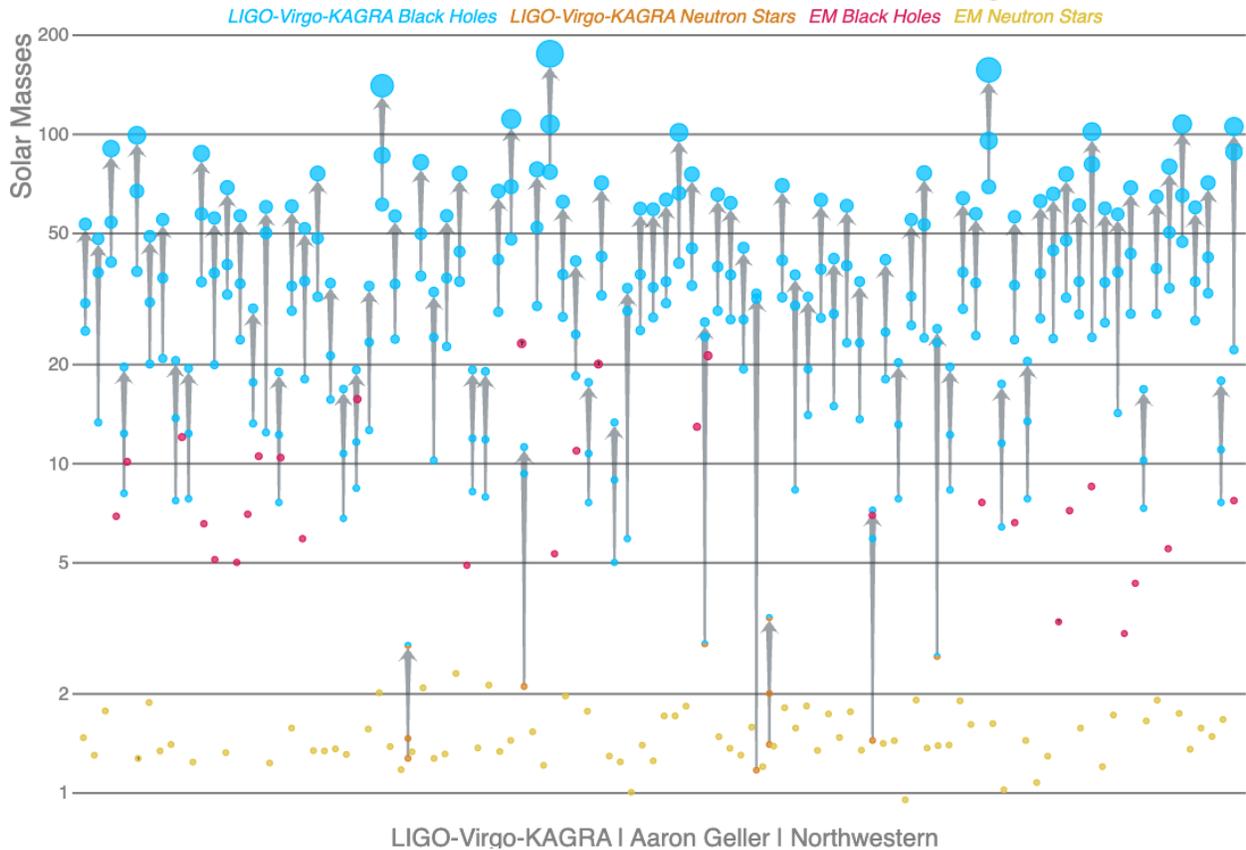


Figure 1.2: Measured noise PSDs in the three observing runs (O1-3) of LIGO H1 detector, and design sensitivities for the future upgrades (Advanced LIGO, A+). Credits: Cahillane & Mansell [1]

filter SNR is an optimal detection statistic for signals buried in Gaussian noise [29]. Additional signal consistency tests are used to distinguish true signals from non-Gaussian noise transients occasionally present in the data [30]. In practice, a finite number of templates are placed optimally in the parameter space, called as a template-bank, to speed up the matched-filtering process. Once the SNR is greater than a certain threshold for any template then a trigger is generated after which statistical tests are applied to estimate its false alarm rate, and post-processing and validation steps are performed to confirm the detections. With upgrades to the detectors, the noise floor is reduced and sensitivity is increased as shown in Fig. 1.2 [1].

Till now LIGO-Virgo-Kagra (LVK) detectors over the three observing runs have found about a hundred CBCs, most of which are BBHs, a couple of binary neutron stars (BNS) and a few mixed binaries consisting of BHs and NS or mass gap objects whose nature is unclear as they lie between the astrophysically expected mass ranges of BH and NS [2, 31, 32, 33, 34, 35]. Fig. 1.3 shows the CBCs from the latest catalog of LVK collaboration, GWTC-3. The first multimessenger event was detected in 2017 from a BNS merger called GW170817

where an electromagnetic (EM) counterpart, a short-gamma ray burst (GRB170817A) was detected by the Fermi satellite, just after the detection of GWs by the LIGO-Virgo detectors [36]. This led to many interesting science cases (mentioned later) and opened up a new field of GW-EM multi-messenger astronomy.



In future, Einstein Telescope [37] and Cosmic Explorer [38] are planned in Europe and USA with much greater sensitivities, such that they will detect $10^4 - 10^5$ CBCs per year up to the redshift of 10. There are also plans for space-based GW detectors which aim to observe GW from the various kinds of sources in different frequency bands, such as Laser Interferometer Space Antenna (LISA) [39] in the milli-Hz band to observe the massive BH binaries and galactic white dwarf binaries, Lunar GW Antenna (LGWA) [40] and DECIGO-BBO [41, 42] in the deci-Hz band to observe the long inspirals of the BBHs and BNSs etc., see Fig.

Figure 1.3: Masses of CBCs from the LVK latest transient catalog, GWTC-3 [2]. Each circle represents a different compact object and the vertical scale indicates the mass as a multiple of the mass of our Sun. Blue circles represent BHs and orange circles represent NS. Half-blue / half-orange mixed circles are compact objects whose classification is uncertain. Each merger involves three compact objects: two merging objects and the final resulting object. Credits: LVK Collaboration/Frank Elavsky, Aaron Geller/Northwestern.

1.4. Ref. [5] describes the future of GW astronomy and physics in the next decade. It is worth mentioning that at the time of writing this thesis, the first evidence of GW background from super-massive BHs occurred using the pulsar timing arrays, that are sensitive in the nano-Hz band [43].

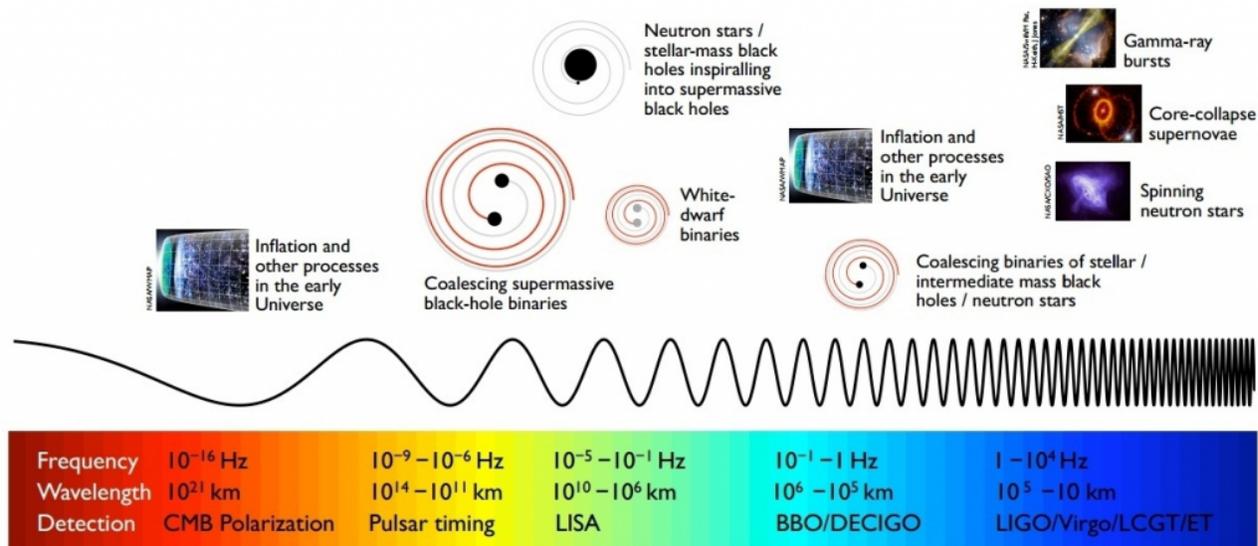


Figure 1.4: GW spectrum from its various kind of sources and frequency bands of the current and future GW detectors. Credits: P. Ajith.

The GW signals carry information about their source, the geometry of space-time and the nature of gravity itself. Therefore, they have a wide range of implications in various fields, such as:-

Compact objects and populations: The properties like mass, spins, rates etc. of compact objects like BHs and NSs can be estimated using the GW signals. These estimates help us to understand the binary object's formation channels, merger rates, populations and the underlying distribution of their masses, spins etc. [44, 45, 46, 2]. Observation of mass-gap objects, where one of the component masses of CBC is lying in the lower mass gap, $\sim 3 - 5M_{\odot}$ (see half blue half orange circles in Fig. 1.3) or in the upper mass gap, $\sim 50 - 120M_{\odot}$ (GW190521 event), in GWTC-3 has already challenged our theoretical understanding stellar evolution [47], and the physics of supernova explosion.

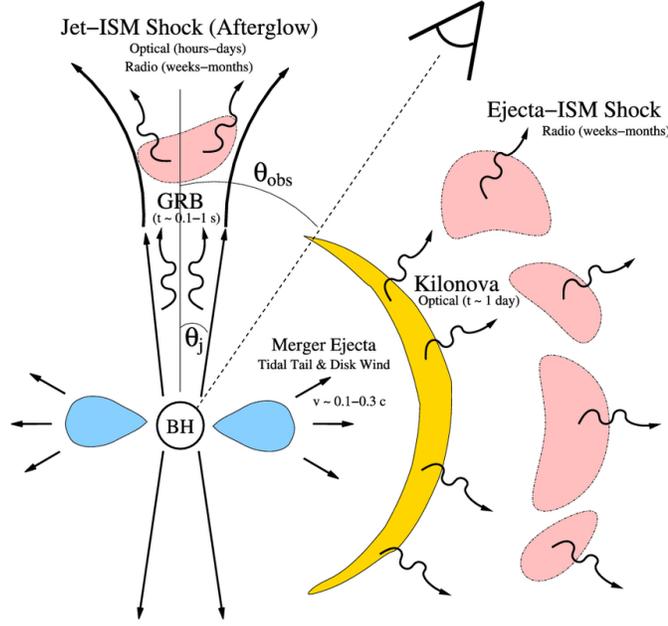


Figure 1.5: Summary of EM counterparts of BNS and NSBH mergers and their dependence on the viewing angle with respect to the axis of the GRB jet. The kilonova, in contrast to the GRB and its afterglow, is relatively isotropic and thus represents the most promising counterpart for the majority of GW-detected mergers. Credits: Metzger and Berger (2012) [3], copyright by AAS

Multi-Messenger Astronomy: BNS and NSBH mergers can have detectable EM counterparts like short-gamma ray bursts, kilonovae, optical afterglows etc., see Fig. 1.5. As soon as the GW signal is detected, and is likely to be BNS/NSBH an ‘alert’ is sent to the EM telescopes with the estimated source sky location from the GW signal for capturing the EM counterparts that may be emitted during the merger/post-merger phase. All this happens in a fraction of a second. If successful, the joint GW-EM observations can help us to study the violent merger, which has very rich physics and understand the properties of intermediate objects like hyper-massive NS that may be formed in the process. For example from GW170817 [48], we could conclude that a good fraction of heavy elements are manufactured in the BNS mergers. EM counterparts also allow us to localise the source to the host galaxy, which helps to measure the cosmological parameters and the speed of GWs precisely, and to constrain the alternative gravity theories invoked in the context of cosmology [4, 49, 17]. GW-EM observations may also allow us to probe the properties of the intergalactic medium in a unique way as the GWs do not interact with matter whereas EM radiation does [49].

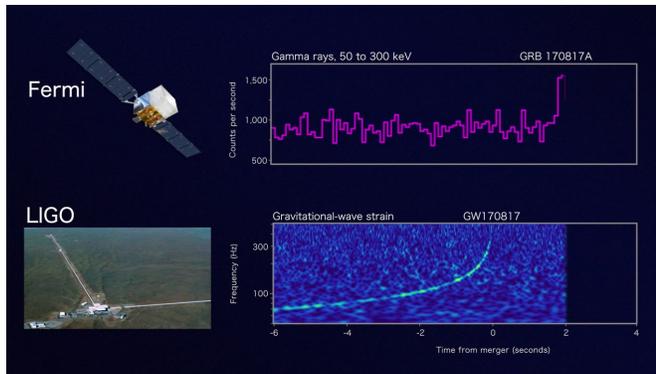


Figure 1.6: Simultaneous detection of GW signal by LIGO-Virgo detectors and a short-gamma ray burst signal by Fermi satellite from BNS merger GW170817. Credits: Astrobites magazine.

Cosmology: The CBCs are known as standard sirens for cosmology because they provide the distance to the source directly. The Hubble constant H_0 which describes the local rate of expansion of the universe was measured using multi-messenger event GW170817 [4], where the GW data provided the luminosity distance of the source while the EM counterpart provided the redshift. The current GW-based measurements are not precise enough to resolve the Hubble tension (see Fig. 1.7), in future with more (~ 50) such GW-EM observations it would be possible. It is also possible to use dark standard sirens i.e. CBCs without EM counterparts to estimate the cosmological parameters with the help of the galaxy catalogues [50].

The cosmological principle states that the universe is homogeneous and isotropic at large scales, however, we expect some anisotropies at small scales due to the Baryon acoustic oscillations, formation of large-scale structures by clustering of galaxies or dark matter filaments etc. The distribution of GW sources in the sky can provide an independent probe for studying the large-scale structure of the universe, especially with the future detectors when the sky-localisations will improve and the number of detections will be large enough [51, 52].

GWs can also help in probing the nature of dark matter and dark energy. GWs may be emitted and detectable from dark matter candidates like ultralight bosons [53], primordial BHs [54], etc. So far there has been no detection and hence constraints are put on their underlying models. In principle, we could also probe the early moments after

the Big Bang and cosmic inflation using the stochastic background of GWs [55].

Dense nuclear matter physics : NSs act as a natural lab for studying the properties of matter at ultra-high densities and pressure, which cannot be created artificially on the Earth [56]. Unlike BHs, NS can be tidally deformed and the GW signals from BNS mergers are sensitive to the tidal deformation parameters which can be then translated to the equation of state parameters and mass-radius relation of the NSs [57]. The current constraints are however poor due to a limitation that the BNS merging frequency ($\geq 1\text{kHz}$) is beyond the sensitivity frequency band of the ground-based GW detectors. An experiment called Neutron Star Extreme Matter Observatory (NEMO) is proposed with sensitivity in the frequency band 2–4 kHz to observe BNS merger and post-merger signals [58].

Gravitational lensing: Bending of light from the stars, galaxies and quasars due to the gravitational field of an intervening matter that acts as a lens, has been observed many times through EM telescopes and has allowed us to test GR, map dark matter distributions, etc. Likewise observing gravitational lensing of GWs is also a prediction of GR, which may distort or produce multiple copies of the GW signal [59, 60, 61, 62]. So far there has not been any detection of lensed signals but is expected soon. We will talk about the GW lensing phenomena in the upcoming sections in detail.

Testing GR: GWs act as a unique probe to the nature of gravity in the regions of high curvature and strong field, see Fig 1.8 [63, 5], which is inaccessible through other observations and laboratory tests. A direct comparison of GR versus any alternative theory of gravity with CBC signals is practically infeasible as the IMR waveforms for CBCs are unavailable for most of the alternative theories of gravity, owing to the complexities in their formulation and numerical relativity simulations. However, there are still a variety of tests of GR that are carried out

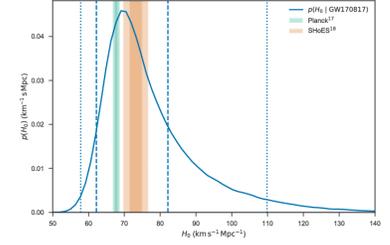


Figure 1.7: GW170817 measurement of H_0 . Marginalized posterior density for the Hubble constant H_0 (blue curve). Constraints at $1 - \sigma$ and $2 - \sigma$ from Planck (Planck Collaboration et al. 2016) and SHoES (Riess et al. 2016) are shown in green and orange. Credits: LVK Collaboration [4].

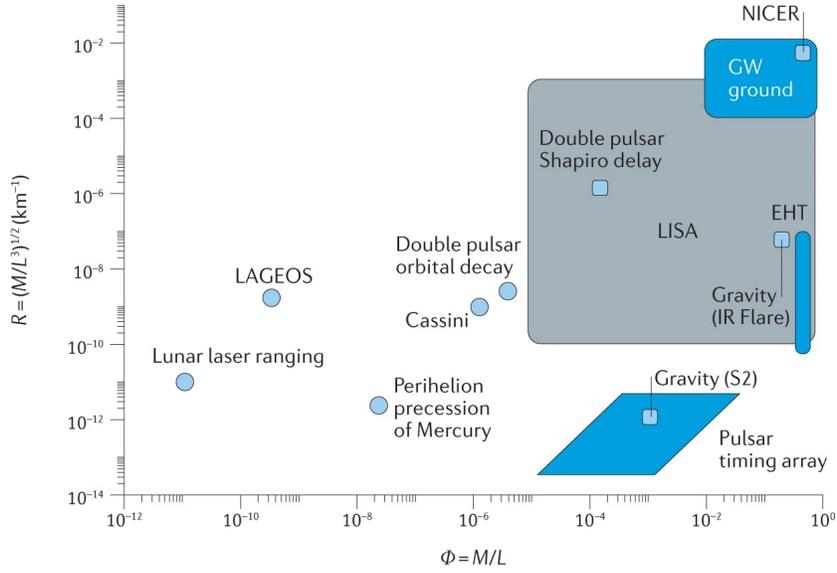


Figure 1.8: Schematic diagram showing the gravitational potential and curvature scale probed by different tests of GR. The horizontal axis shows the curvature scale $R = \sqrt{M/L^3}$ due to the presence of a mass M of size L , while the vertical axis shows the compactness or the dimensionless gravitational potential, $\Phi = M/L$ probed by different experiments. For example, observation of the double binary pulsar (total mass of $2.6M_{\odot}$) and orbital separation of $\sim 10^6 \text{ km}$ probes the curvature scale of $R \sim 2 \times 10^{-9} \text{ km}^{-1}$ and compactness of $\Phi \sim 4 \times 10^{-6}$. In contrast, a black hole binary of total mass $10M_{\odot}$ at the time of merger probes a curvature scale of $R \sim 2.4 \times 10^{-2} \text{ km}^{-1}$ and compactness of $\Phi \sim 0.5$. Credits: References [5, 6]

using GWs, such as:

- Waveform Consistency:* One of the primary tests is to compare the observed GW signals with the predicted waveforms based on GR. Inspiral-merger-ringdown (IMR) test compares the estimated final mass and spin of the remnant object using the data separately from the inspiral phase and post-inspiral i.e. merger-ringdown phase, see Fig. 1.9. As we mentioned above, different phases of GW waveform involve different theories or approximations of GR and different tools, IMR tests allow us to check the consistency of these using the GW data [64].
- Parameterized Waveforms:* The inspiral regime of GW waveform is modelled using the post-Newtonian theory, which adds correction terms as the lowest orders deviations to the Newtonian equation motion and quadrupole formula, expressed in orders of the ratio of the characteristic velocity of the compact objects to the speed of light v/c . Within GR, each of these terms is related to the properties of the binary system and their form is known. However, in alternative theories of gravity, there will be deviations appearing at the specific orders of v/c which could be directly measurable from the GW signals. Parameterized post-Newtonian tests measure these deviations to probe the weak field dynamics of the CBCs in a model-agnostic way [65, 66, 67, 68, 69, 70]. Similar to this a parameterised post-Einstein test is proposed to measure the deviations appearing at various orders of frequency [71] in the phase of the waveform predicted by GR. These parameterized deviations can also be translated to constraints on the specific alternative theories of gravity [6].
- Propagation:* GR predicts that GWs propagate at the speed of light and do not experience any dispersion. In alternative theories of gravity, for example, massive graviton theories, GWs can disperse and travel at a different speed. By measuring the phase of GW signals at different frequencies, one can test whether there is any

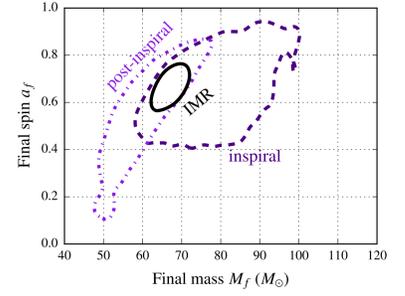


Figure 1.9: IMR consistency tests with GW signals. The final mass and spin posteriors from the inspiral and post-inspiral phases should match if the GR and its approximations capture the GW signal correctly. Credits: LVK collaboration [7].

dispersion or deviation from the expected speed and place an upper bound on the mass of the graviton [72, 68, 70]. According to GR and the standard cosmological model, the propagation of GWs away from the source is well described by the linear perturbations of the background Friedmann-Robertson-Walker (FRW) metric [73]. The measurements of GW speed (better constrained with GW signals with an EM counterpart), luminosity distance, phase etc. from the GW signals can also be used to constrain the modified gravity theories in which GW propagation departs from FRW propagation [17, 74, 75, 76, 77].

- *Ringdown/BH spectroscopy*: The properties of remnant BH from CBCs can be extracted from the ringdown signal. The ‘no-hair’ theorem suggests that a BH is fully defined by its mass, spin and charge. However, in certain scalar-tensor theories of gravity, BHs have additional scalar hair. According to GR, the GW ringdown signal is described by perturbations to the remnant object (Kerr BH), which can be modelled as a linear superposition of quasi-normal mode frequencies and their damping times. According to the no-hair theorem, each of these frequencies and damping times should depend only on the remnant BH’s mass and spin (assuming that the charge is zero, due to astrophysical reasons). Using the GW ringdown signals, any deviations in these frequencies or damping times will be a signature of a violation of GR or a non-BH nature for the remnant object. . [78, 79, 80].

Additionally, if the final remnant is an exotic compact object with a light-ring and reflective surface rather than an event horizon then, there may be signatures of GW echoes, i.e. occurrence of a repeated train of pulses separated in time after the ringdown. These GW echoes are modelled in a phenomenological way, and using the Bayesian model selection one compares the two hypotheses, GW and GW+echoes [46, 81, 82, 83].

- *Polarizations*: As we mentioned before, GR predicts only two (ten-

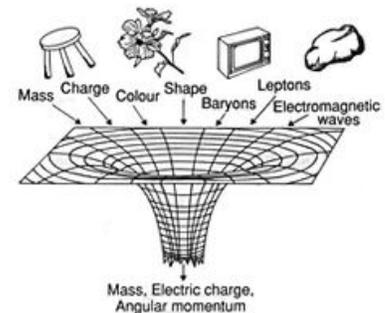


Figure 1.10: No-hair theorem of Black-holes. Credits: Astrobites magazine.

tor) polarization modes of the GWs. However, alternative theories allow up to six polarisation modes (additional scalar and vector modes). The GW signal at each detector is a unique linear combination of the GW polarisations. Therefore, using the data from multiple detectors we could extract the GW polarisations, and check if there are any additional modes than $+$ and \times . Currently, this test is limited by the number of detectors available in the network and the constraints exist for the relative strength of tensor v/s scalar and tensor v/s vector modes [84, 11, 49, 85, 2, 46]. We will see in Chapter 4, that strong lensing of GWs may help us in overcoming this limitation.

Using the GW observations till now, no compelling evidence of a violation of GR is found and in most cases, constraints are put on the extra parameters that describe any deviations from GR [86, 87, 68, 70]. In future as the detector sensitivities improve we expect to get better constraints on alternative theories, however, the systematic errors of GW waveform modelling may also become important while performing the tests of GR and especially while combining the measurements from multiple observations [88].

1.2 Gravitational lensing

Gravitational lensing is a phenomenon in which light, as well as GWs, are bent or distorted by the gravitational field of an intervening massive object which curves space-time [89]. In EM astronomy, based on the effect of lensing on the images, the following types of lensing phenomenon have been observed:

Strong Lensing: Strong gravitational lensing occurs when the gravitational field is strong enough to produce multiple images, arcs, or even complete rings around the lensing object. These images can be observed when the source, lens, and observer are suitably aligned.



Weak Lensing: Weak gravitational lensing refers to a more subtle effect where the gravitational field causes slight distortions in the shape and orientation of background objects. It is characterized by the statistical analysis of the shape and alignment of large numbers of galaxies or other distant objects. Weak lensing provides information about the distribution of matter in the universe and can be used to map the dark matter distribution.

Micro Lensing: Microlensing occurs when a massive object, such as a star, passes in front of a distant light source, such as a more distant star. The gravitational field of the foreground object acts as a lens and causes temporary brightening or magnification of the background source. Microlensing events are used to study the properties of individual stars, planets, and other compact objects.

Therefore in various astrophysical contexts, EM lensing not only helps us understand the properties of the lensing objects but also provides information about the large-scale structure of the universe, the nature of gravity, and the effects of space-time curvature. It serves as a powerful tool for astronomy and physics.

Similarly to light, GWs can also be lensed. However, the methods to detect and make use of GW lensing are entirely different. Instead of image shape distortions or transient brightening of stars, GW lensing

Figure 1.11: Types of EM lensing: strong, weak and micro lensing. (left) Observation of multiple images of a quasar are formed due to strong lensing by a galaxy. (middle) Distortions of the background galaxies caused due to weak lensing by a galaxy cluster, imaged by Hubble Space Telescope. (right) Schematic of the time-varying brightness of a star caused due to microlensing by a compact object. Credits: ESA, Hubble, NASA

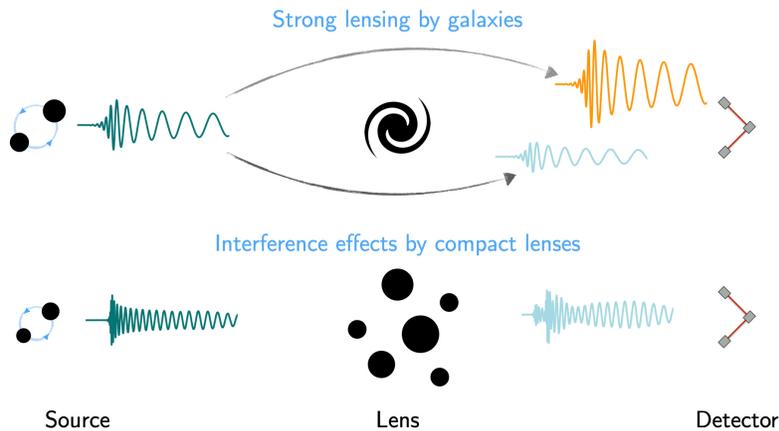


Figure 1.12: Schematic representation of the GW lensing. On the top, a given signal is strongly lensed by a galaxy lens producing multiple images with different amplitudes, arrival times and phases. On the bottom, small compact lenses produce interference effects that distort the detected waveform. Credits: LVK collaboration

results in temporally resolvable copies of GW events called *strong lensing* and/or frequency-dependent distortions in the GW waveforms, called *microlensing*. GWs can also get weakly lensed, where the GW signal from a binary would be magnified or de-magnified without resulting in multiple copies.¹

Till now there has been no confident detection of lensed GWs, however, in future, it is anticipated to be a very exciting discovery which will act as a new tool in astronomy and physics.

1.2.1 The geometrical optics limit

This section is mostly based on the references [90, 91, 92]. For any radiation of wavelength, λ the geometrical optics (G.O) limit applies when the length scale associated with the lens, $GM_{lens}/c^2 \gg \lambda$, where M_{lens} is the mass of the gravitational lens². For EM waves, this condition holds for any of the typical astrophysical lensing objects like planets, stars, galaxies, clusters etc. However, for the GWs the wavelengths can be larger and become comparable to the lens length scale, in which case wave optics effects become important. Here we are interested in the GO limit, where we can treat the radiation as rays travelling in straight lines and getting deflected due to the lens.

In the limit of a "thin lens", where the distances between the source, lens, and observer are much larger than the size of the lens (this is almost always true for astronomical objects), the source and lens po-

¹ Beware of the confusing terminology, the micro-lensing in GW case refers to the wave-optics effects whereas for EM it refers to time-varying magnification due to relative motion of source and lens w.r.t observer in the geometrical optics limit itself. Additionally, lensing in GWs may also lead to temporally unresolvable copies of GW signals called millilensing.

² Source size may also play a role, see [93]

sitions are defined by the vectors in their respective planes. Fig. 1.13 shows the geometry of the source-lens-observer system. The relation between the positions of the source (S), lens (L), and observer (O), is then given by the lens equation:

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \frac{D_{LS}}{D_S} \hat{\boldsymbol{\alpha}}(D_L \boldsymbol{\theta}) = \boldsymbol{\theta} - \nabla \psi(\boldsymbol{\theta}) \quad (1.2)$$

where $\boldsymbol{\beta}$ represents the true source position, $\boldsymbol{\theta}$ denotes the observed image position, $\boldsymbol{\alpha}(\boldsymbol{\theta})$ represents the deflection angle, and D_{LS} and D_S are the angular diameter distances between the lens and the source and between the observer and the source, respectively. In strong gravitational lensing, this equation can have multiple solutions, because a single source at $\boldsymbol{\beta}$ can be lensed into multiple images ($\boldsymbol{\theta}$).

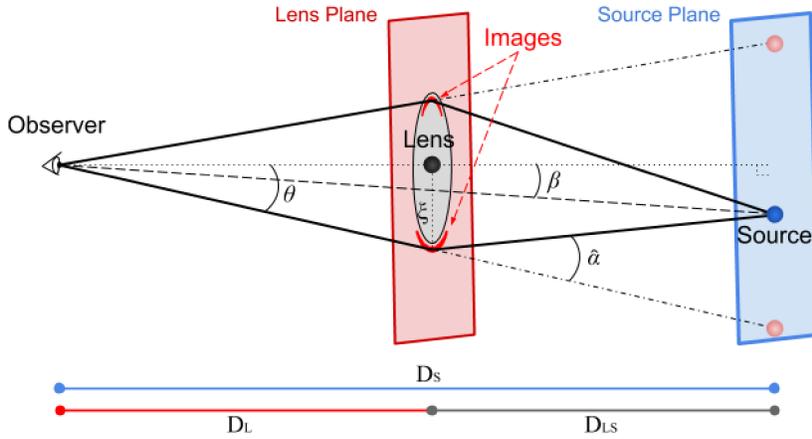


Figure 1.13: Angles involved in a thin gravitational lens system for the deflection of radiation in G.O limit. The distances from the observer to the lens and the source are D_L and D_S , respectively. The distance between the lens and source is D_{LS} . β is the angular position of the unlensed source, $\hat{\alpha}$ denotes the deflection angle and θ is the angular position of one of the images. The red and blue rectangles represent the parallel lens and source planes at D_L and D_S , respectively, which are perpendicular to the observer-lens axis (dotted). Credits: Reference [8].

The deflection potential ψ , is defined such that the scaled deflection angle is just the gradient of the potential ($\boldsymbol{\alpha} = \nabla \psi(\boldsymbol{\theta})$). It can be written as a scaled projection of the Newtonian gravitational potential Φ of the lens, and also relates to the lens surface mass density $\Sigma(\boldsymbol{\theta}')$ ³ as,

$$\psi(\boldsymbol{\theta}) = \frac{2D_{LS}}{D_L D_S c^2} \int dz \Phi(D_L \boldsymbol{\theta}, z) = \frac{1}{\pi \Sigma_{cr}} \int d\boldsymbol{\theta}' \Sigma(\boldsymbol{\theta}') \ln |\boldsymbol{\theta} - \boldsymbol{\theta}'| \quad (1.3)$$

where $\Sigma_{cr} = \frac{c^2 D_s}{4\pi G D_{ds} D_d}$ is called the critical surface density. Hence,

³ Using the Poisson equation and thin-lens approximation

the deflection angle and image positions vary with the mass profile of the lens. The most common lens profiles used are point mass, singular isothermal sphere/ellipsoid.

The magnification μ of each image is, $\mu = \frac{1}{\det(A_{ij})}$ where A_{ij} is the Jacobian matrix between the unlensed and lensed coordinate system,

$$A_{ij} = \frac{\partial \beta_i}{\partial \theta_j} = \delta_{ij} - \frac{\partial \alpha_i}{\partial \theta_j} = \delta_{ij} - \frac{\partial^2 \psi}{\partial \theta_i \partial \theta_j} \quad (1.4)$$

Because the matrix of second derivatives must be symmetric, the Jacobian can be decomposed into a diagonal term involving the convergence and a trace-free term involving the "shear". The term involving convergence magnifies the image by increasing its size while conserving surface brightness. The term involving the shear stretches the image tangentially around the lens.

There is an alternative way of deriving the lens equation, starting from the radiation arrival time t_d (Fermat surface), which is given by the sum of geometric and Shapiro time delays [92]

$$t_d = \frac{D_L D_S}{D_{LS} c} (1 + z_L) \left[\frac{(\boldsymbol{\theta} - \boldsymbol{\beta})^2}{2} - \psi \right] \quad (1.5)$$

Fermat's principle in geometric optics states that a light ray takes a path with a stationary path length. This immediately suggests that observed images should satisfy the following condition

$$0 = \nabla t_d = \boldsymbol{\theta} - \boldsymbol{\beta} - \nabla \psi(\boldsymbol{\theta}) \quad (1.6)$$

which gives back the lens equation 1.2. Likewise, the second derivative gives back the Jacobian matrix Eq. 1.4. This is equivalent to the stationary phase approximation to the diffraction integral formula for the amplification factor F of lensed GWs in the expanding universe as a function of frequency f and the source position $\boldsymbol{\beta}$ [92, 91], which is given by,

$$F(f, \boldsymbol{\beta}) = \frac{D_L D_S}{D_{LS}} \frac{1}{c} \frac{f}{i} (1 + z_L) \int d^2 \theta e^{i2\pi f t_d(\boldsymbol{\theta}, \boldsymbol{\beta})}. \quad (1.7)$$

In the stationary phase approximation, the integral over paths is

dominated by the stationary phase points (Eq. 1.6). The contribution to F from the j -th image is obtained by Taylor expanding t_d up to the quadratic level, which is valid at high enough frequencies such that $|f\Delta t_d| \gg 1$ where, Δt_d is the time delay difference between stationary points [91]. The final amplification factor in this approximation is then given by

$$F \approx \sum_j |\mu(\theta_j)|^{1/2} \exp\left(i2\pi f t_d(\theta_j) - i \text{sign}(f) \frac{n_j \pi}{2}\right), \quad (1.8)$$

where $n_j = 0$ for images formed at the minima of time delay surface called type I, $n_j = 1$ for images at the saddle called as type II, and $n_j = 2$ for images at the maxima called type III images. n_j , also known as the Morse index, accounts for the (Morse) phase shift that asymptotic waves far from the lens acquire due to crossing a given number of caustics between the source and the observer [90]. Note that this phase shift is an additional effect on top of magnification and time delay, and acts as a correction to the G.O. limit [94, 95].

1.2.2 Strong lensing of GWs

About 0.1 – 1% of events in the ground-based GW detectors are expected to be strongly lensed [96, 97]. For the transient GW signals such as that from CBCs, the multiple strongly lensed copies of GW signal from the same source appear as multiple temporally resolvable events [98, 94, 99, 100]. For ground-based GW detectors, this occurs when the lens is an intervening galaxy or a cluster of galaxies, and the GO limit applies. These events are (de)-magnified copies of the source separated by time delays that span minutes to weeks, if lensed by a galaxy and can go up to decades for lensing by galaxy clusters. Thus, they will have identical phase evolution, although their amplitudes will generally differ by a constant factor (see, e.g., [101]). Given a strongly lensed binary merger, each j^{th} copy of its GW signal ($\tilde{h}_j^L(f)$) with parameters Λ , has a different time delay (Δt_j), magnification (μ_j) and a constant phase difference ($\Delta\phi_j$), of either $0, \pi/2$ or π , depending on the image type (type I, II or III [94, 95]), as derived in the last

section.

$$\tilde{h}_j^L(f; \Lambda, \mu_j, \Delta t_j, \Delta \phi_j) = \sqrt{|\mu_j|} h(f; \Lambda, \Delta t_j) e^{i\Delta \phi_j \text{sign}(f)} \quad (1.9)$$

The sky localisation from the GWs signals is usually poor, $O(100 \text{ deg}^2)$ with current detectors. The change in apparent sky location due to lensing is expected to be of $O(\text{arcseconds})$, with the typical galaxy lenses. Unlike multiple images of lensed galaxies, the lensed GWs are therefore unresolvable in the sky. Hence, we expect the SL event pairs to have overlapping sky areas. Several techniques have been devised to search for such strongly lensed pairs of GW events, most of which rely on the identical phase evolution of the images, the overlapping GW localization sky areas of the images [102], and the Morse phase. These include two low-latency techniques. One is a machine learning (ML) based method that we develop in Ch. 2 which compares time-frequency maps and localization sky areas of individual events in a candidate lensed pair [103]. The other is a posterior-overlap (PO) method that compares the existing Bayesian posterior distribution of the (intrinsic and sky location) parameters of individual events acquired from Bayesian parameter estimation (PE) exercises [101], which is also described in 2. Other more comprehensive though computationally expensive methods involve sampling a joint-likelihood, constructed from the GW-likelihoods of the individual events in the candidate lensed pair while accounting for selection effects and population priors [104, 105, 106, 107]. The LVK collaboration has searched for strong lensing and microlensing signatures in the following LVK observing runs: O1–O2 [108], O3a [9], and the full O3 run [15], yielding no confident signatures. See Ch. 3 for the lensing searches in the full O3 run using low-latency methods like ML and PO. In parallel, other searches have been performed, confirming that no lensing features have been confidently detected so far [107, 109, 110, 111]. This is because most of the CBCs observed so far are from the low redshifts (till $z \sim 0.2$), hence have a low probability of lensing, $\lesssim 1/\text{yr}$. In future with ET and CE detectors, we can detect unlensed CBCs up to higher

redshifts ($z \sim 2 - 8$) and would be able to detect many lensed CBCs, $\sim 100/yr$. However, with the increasing number of events the probability of chance overlaps in the unrelated events also increase [112] therefore, it is a challenge for identifying just a handful of strongly-lensed GW signals among the thousands of unlensed ones.

Like EM lensing, GW lensing is also expected to be a powerful tool for studying astrophysics, cosmology, and fundamental physics. Some of the astrophysical implications of strong lensing of GWs that have been emerging with the arrival of GW astronomy are the following :

Astrophysics: The observation of multiple lensed signals can be used to infer properties of both source and lens. As the GWs sources are compact, the magnifications can be very high as compared to the extended sources like galaxies, quasars etc. Due to the magnification, lensing phenomena could allow us to probe the universe at higher redshifts and infer the population of sources and lenses. Even from the non-observation we could constrain the compact binary merger rates, especially at high redshifts as seen in Fig. 1.14 [113].

Hannuksela et al. [10] proposed that by combining the sky localisation posteriors of multiple (2 or more) lensed copies of GWs, one can reduce the uncertainty of the source sky location by a factor of 10 with quadruply lensed GWs [105, 106, 10, 114], see Fig. 1.15. This can also help in the identification of the host galaxy, by cross-matching with the EM galaxy lens catalogs[115].

The lensed signals could also be in principle used to explain the mass-gap events. The lensing magnification can make the source appear closer, and if not accounted for it can bias our estimation of the luminosity distance of the source, and eventually its redshift. Since the GW signals encode only the redshifted mass of the source, this magnification bias will propagate and lead to an overestimation of the intrinsic mass of the source. However, the mass-gap events observed till now, require lensing magnifications of the order 100 or even more to de-bias the estimation of intrinsic masses so that they become consistent with the astrophysically expected distribution, see Fig. 1.16.

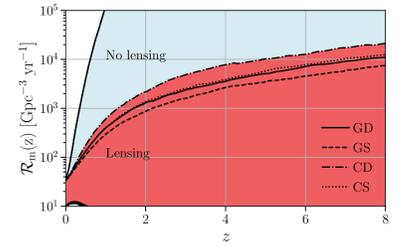


Figure 1.14: (blue) red region signifies the constraints on merger rate density of BBHs when (not) accounting for non-observation of lensing of the LVK O3a events. The high-redshift regions are more constrained due to the lensing magnification. Credits: LVK collaboration, O3a lensing [9]

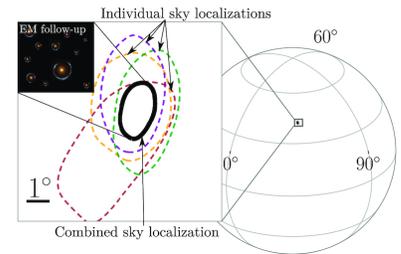


Figure 1.15: An illustration of a sky localization of a quadruply lensed GW. The individual (colour) and the combined (black) sky localizations at 90 percent confidence are shown. Each lensed GW essentially gives us a new set of detectors with which to localize the event in the sky, allowing for improved sky localization. A dedicated follow-up of the narrowed sky region would then allow us to search for the lensed host galaxy from which the GW originates. Credits: Hannuksela et al. [10]

Given the lensing rates and low probability of these high magnifications, it is unlikely to be the case [9, 11].

It is also a possibility to discover new galaxies and galaxy clusters from the observation of lensed copies of GWs. The merit of GW lensing, in addition to its sensitivity to high redshifts as compared to EM, is that the GWs detectors are sensitive to all-sky whereas EM telescopes have a limited field of view.

Cosmology: Strongly lensed transients offer a unique opportunity to do time delay cosmography and measure Hubble constant etc. very precisely. Lensed supernovae transients have turned out to be very useful for probing cosmology, as the time delays between the lensed counterparts can be estimated to O(day) precision. This is very challenging from the lensing of quasars, as the variation in their flux is at large timescales, typically tens of years and hence measuring them would require a longer period of observation. The GWs time of arrival can be measured to a $\sim ms$ precision which makes them an excellent candidate for time delay cosmography. Jana et al. [116] showed that with the future detectors ET and CE that cosmological parameters like H_0 and those related to the dark energy equation of state can be determined using the time delay distribution of the lensed BBHs.

The GW signals carry information about the luminosity distance directly and hence are called standard sirens. However, the GW strain contains only the multiplication of mass and redshift which make their separate measurements infeasible. Therefore, another means of getting the source redshift by identification of the host galaxy becomes important to do precision cosmology [117, 118, 119, 120, 121]. As discussed before, this can be achieved directly with BNS or NSBH mergers with EM counterparts. However, for the sources without an EM counterpart, precision cosmology is still achievable if they are lensed. As mentioned earlier, Hannuksela et. al [10] proposed using multiple copies of lensed GWs to narrow down the sky location and EM lensed catalogues to identify the host galaxy of the source.

Liu et al. [122] found that detecting 30 strongly lensed sources

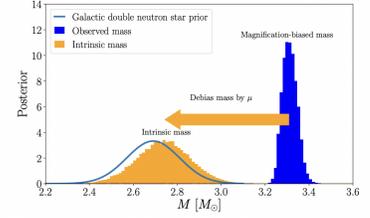


Figure 1.16: Estimation of lensing magnification for explaining the observed mass-gap nature of a compact nature, as a result of bias in redshift measurements. Credits: Pang et al [11].

jointly in GW and EM has the potential to enhance dark energy measurements by a factor of two compared to current measurements based on supernovae type Ia and cosmic microwave background observations.

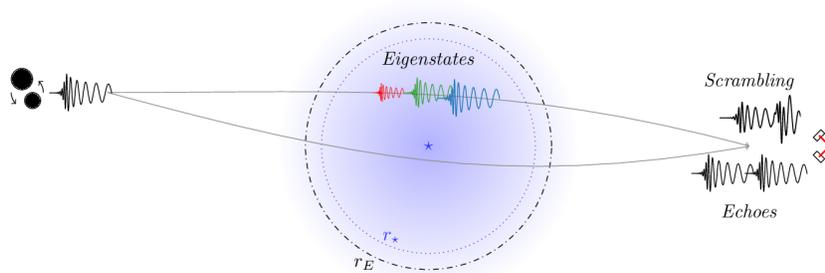


Figure 1.17: GW lensing beyond general relativity. A GW emitted by a binary black hole splits into its propagation eigenstates (waveforms in color) when it enters the region near the lens where modify gravity backgrounds are relevant. Depending on the time delays between the propagation eigenstates the signal detected could be scrambled or echoed. If the GW travels closer than the Einstein radius, multiple images could be formed as indicated by the grey solid trajectories. Credits: Ezquiaga and Zumalacaragui [12]

Testing GR: Observation of lensed GWs would itself be a validation of GR. Additionally, lensing probes the GW propagation near massive objects in the weak field limit. Some of the alternative theories of gravity like Brans-Dicke, Horndeski etc. predict the birefringence of GWs, where each GW component travels at a different speed and is deflected by lens differently [12]. Observation of birefringence would violate GR and on the other hand, non-observation will allow to constrain the beyond GR theories. Based on this phenomenon we propose a test of GR and apply it to the GWTC-3 events in chapter 5.

Strongly lensed copies of GWs can be used to improve the polarisation tests of GR. In GR there are only two polarisation modes $(+, \times)$, however, in alternative theories of gravity, there could be up to six polarisation modes. The GW signal at each detector is a linear combination of the GW polarisations, which depends on the sky location of the source relative to the detector. Currently, we are limited by the number of detectors in the network to extract each of the polarisations individually. With strong lensing, there would be multiple copies of GW signals, and each strongly lensed copy of GW signal contains a different linear combination of the polarisations, as they arrive at Earth at different times in which the relative location of the source changes.

We show in Ch. 4, how the strong lensing helps in distinguishing GW polarisations in a Bayesian model selection framework.

Strongly lensed GW events with detectable (also lensed) EM counterparts would also be a novel system to probe not only cosmology but also the nature of GWs. In massive graviton theories the speed of GWs is less than the speed of light, measuring the time delay between the lensed copies of EM and GW signals would allow us to accurately infer the difference in the speed of gravity and light while overcoming the uncertainty of intrinsic time delay between the emission of GW and EM radiation at the source. [123, 124]

Multi-messenger astronomy: The typical galaxy lens produces up to 4 copies of the radiation, and galaxy clusters can produce even more. To catch the EM radiation of a source, using its GW signals, early warning has to be sent to the EM observatories to point their telescopes in the direction of the source. This naturally causes some delays in the observations, lowering the chance to capture the EM radiation produced during the merger and post-merger phases. Magare et al. [125] proposed that using the information of detected lensed GWs of sources like BNS and NSBHs could help in predicting their upcoming lensed images (both GW and EM) and their sky locations by reconstructing the lens properties. This could give us the order days to months of early warning time and capture not only the merger/post-merger but also pre-merger radiation that may be produced during the multi-messenger events. This was the case for a lensed supernovae explosion observation, where the early warning due to predicted lensing delays helped in observing the shock-break phenomena which happen at the early stages of the supernovae explosion and last for only a few seconds[92].

Over the past few decades, an increasing number of fast radio bursts (FRBs) are being detected by radio telescopes. However, their origin and generation mechanism is still unknown. FRBs could also be lensed and appear in radio telescopes at different times. However, there has been no confident detection till now. Singh et al [126] proposed that

given the detection of a pair of FRB signals and a pair of GWs from a lensed NSBH source, just from the consistency of the time delays measured from the FRB and GWs alone, one could claim a 5σ association of the origin of FRB to the respective NSBH system.

Baker and Trodden [127] proposed that the time delay of lensed multimessenger signals (e.g., photons, GWs, massive neutrinos) from the same source can place bounds on the total neutrino mass and probe cosmological parameters. For sources at high redshift, the small relativistic corrections accumulated along the propagation may become measurable and carry information about the difference between null and non-null geodesics, giving insight into the expansion of the universe and properties of massive particles to trace geodesics.

Dark matter and sub-structures: GW signals, in addition to being strongly lensed by macro lenses (like galaxies, and clusters), can also get diffraction from the micro-lenses in their path. In such cases, each copy of a strongly lensed signal will also have frequency-dependent modulations, other than magnification. From the resultant signals, the masses of the micro-lenses could be inferred could help us in probing the substructures or the small-scale dark matter halos in the lensing galaxy [128, 129, 130, 131, 132, 133]. In principle, this could allow us to solve the ‘Missing-satellite-problem’, which originates from the low numbers of observed satellite dwarf galaxies in the Milky Way inconsistent with the cosmological simulation of dark and baryonic matter. Additionally, one could also search for Intermediate Mass BH that may be acting as a microlens [134] and also constrain the fraction of dark matter in the form of Massive Compact Halo Objects (MACHOs) [135].

1.2.3 Organisation of the thesis

The rest of the thesis is organised as follows. In Chapter 2, we develop an ML algorithm to classify a given pair of signals as lensed or unlensed and show that it performs comparably to a Bayesian model selection method and is computationally much faster. Chapter 3 deals with the search for strong lensing in the LVK O₃ data using both ML

and PO methods. We also develop and employ another rapid analysis method to identify weaker (sub-threshold) lensing counterparts in the O₃ data. In Chapter 4, we discuss our idea about how the multiple copies of GW signals due to strong lensing would help in better measurement of the GW polarisations. In chapter 5, by modelling this effect phenomenologically we perform a test of GR in the Bayesian model selection framework using the latest catalog of GW events (GWTC-3). Finally in chapter 6, we summarize the findings of this thesis along with the future research directions.

2 | Rapid identification of strongly-lensed gravitational wave signals using machine learning

Abstract

A small fraction of the GW signals that will be detected by second and third-generation detectors are expected to be strongly lensed by galaxies and clusters, producing multiple observable copies. While optimal Bayesian model selection methods are developed to identify lensed signals, processing tens of thousands (billions) of possible pairs of events detected with second (third) generation detectors is both computationally intensive and time-consuming. To mitigate this problem, we propose to use machine learning to rapidly rule out a vast majority of candidate lensed pairs. As a proof of principle, we simulate non-spinning BBH events added to Gaussian noise, and train the machine on their time-frequency maps (Q-transforms) and localisation skymaps (using Bayestar), both of which can be generated in seconds. We show that the trained machine is able to accurately identify lensed pairs with efficiencies comparable to existing Bayesian methods.

2.1 Introduction

With the arrival of GW astronomy over the past decade, observation of gravitational lensing of GWs is a highly anticipated in the near future. In this chapter, we concern ourselves with strong lensing of GWs assuming that the wavelength of the GWs is negligible in comparison to the Schwarzschild radius of the lenses, as is the case when GWs from coalescing stellar-mass BBHs are lensed by galaxies or galaxy clusters. As discussed in Ch. 1, in this so-called GO limit, strong lensing of GWs will result in the production of potentially resolvable

images. The resolvability of images in the sky is ultimately dependent on the resolution of the telescopes that observe these images. GW detectors typically have very poor angular resolution [136, 137] (at least in comparison to optical telescopes); the localisation skyarea for GW events detected by the LIGO-Virgo network in the second and third observing runs spanned tens of square degrees at best [46]. As a result, even strongly lensed GW events typically have images whose skyareas almost completely overlap each other. Indeed, one of the signatures that two GW events are lensed copies is that their skymaps overlap (see, e.g. [101, 138]). While strongly lensed GW events are completely unresolvable in the sky with current GW detectors, they are typically very well resolved in time. In the GO limit, these GW images would have different amplitudes, but their phase evolution would be identical [139, 140, 141, 131, 142, 95]. Thus, in principle, determining whether two non-overlapping GW events are lensed copies comes down to comparing the shapes of these signals with respect to each other.

In practice, however, such a comparison is non-trivial. Firstly, the observed GW signals are projections of the true GW signals onto the detectors; this projection depends on the location and orientation of the detector relative to the source, and would therefore be different for each of the temporally separated GW images. Furthermore, these images would be buried in detector noise. Even if the noise is assumed to be Gaussian and the corresponding PSD is assumed to be time invariant, each of the images would be buried in different realisations of this noise.

A robust alternative to such a direct comparison of the GW-signals is to work in the space of the inferred source parameters. Using optimal matched-filter based parameter inference techniques [143], Bayesian posterior distributions on the intrinsic parameters of the source (the masses and spins of the binary) and its extrinsic parameters (the sky-location of the binary) can be constructed. As mentioned earlier, the phase evolution of the GW images are expected to be identical, and therefore comparing the inferred posteriors on the intrinsic parameters

(which completely govern the phase evolution) of pairs of GW events should enable us to discriminate between lensed and unlensed pairs. This discriminability can be further enhanced by comparing the localisation skymaps which are expected to overlap almost entirely for lensed GW pairs [101].

Quantitatively, such a comparison can be achieved using Bayesian model selection [101, 10]. A Bayes factor derived from the overlap between the posteriors of pairs of events can be constructed and used to segregate these pairs as either lensed or unlensed. However, evaluating this discriminator is computationally expensive and time consuming. Bayesian parameter inference of BBH events can take hours to days. Additionally, constructing the Bayes factor can take up to a few minutes per event, and the number of such evaluations will grow as the square of the number GW events. This makes the estimation of the Bayes factor computationally challenging when large numbers of BBH events are expected to be detected in future observing runs.

Current estimates of the rate of stellar-mass binary black-hole (BBH) mergers [144] suggest that hundreds of BBH events are expected to be detected in LIGO-Virgo-Kagra’s next observing run (O4). Among these GW detections, up to a percent could be lensed copies of each other [145, 96], suggesting that there is a non-trivial chance that the first confirmed detection of a lensed GW pair could occur in O4. However, identifying such lensed pairs would require constructing $\mathcal{O}(10^2)$ posteriors on the GW events’ source-parameters and $\mathcal{O}(10^4)$ Bayes factors.

These numbers will get significantly larger with observing runs beyond O4, and astronomically large by the time the third generation (3G) network of ground-based detectors [38, 37, 146] completes its observations. The 3G network is expected to observe $\mathcal{O}(10^5 - 10^6)$ events, of which $\sim 0.3\%$ could be strongly lensed [96]. Therefore, $\mathcal{O}(10^5 - 10^6)$ event posteriors, and $\mathcal{O}(10^{10} - 10^{12})$ Bayes factors, would need to be evaluated.

This motivates the need to come up with a method to conduct a preliminary segregation of pairs of GW events to rapidly “weed out”

the vast majority of unlensed pairs. In this work, we propose to use machine learning algorithms, trained on time-frequency maps of the detector strain time series [147] and the (rapidly estimated) localisation skymaps [148], from both lensed and unlensed pairs of GW events, to construct a statistic to discriminate between lensed and unlensed pairs. Using synthetic, non-spinning BBH signals - both lensed and unlensed - injected in Gaussian noise, we show that our machine-learning-based statistic, performs almost as well as the optimal Bayes factor statistic described above, while reducing the computation time by orders of magnitude. The significant reduction in evaluation time is a direct consequence of the fact that time-frequency maps and localisation skymaps can be constructed in seconds, in contrast to GW inference posteriors which take hours to days to sample.

The rest of this chapter is organized as follows. Section 2.2 summarizes the evaluation of a Bayes factor statistic, section 2.3 introduces the machine learning algorithms we use, and delineates their training and validation. Section 2.4 describes our results in distinguishing between lensed and unlensed GW event pairs and compares them with the performance of a Bayesian method (posterior overlap statistic). Section 2.5 summarizes this work and discusses its potential benefits.

2.2 *The posterior overlap statistic*

Let $d(t)$ be the detector strain time series which is known to contain a gravitational wave signal $h(t, \vec{\theta})$ with shape (intrinsic and extrinsic) parameters $\vec{\theta}$, as well as one realisation of stochastic Gaussian noise as characterized by its power spectral density $S_n(f)$. A Bayesian inference of $\vec{\theta}$ from $d(t)$ can be achieved by sampling the posterior distribution on $\vec{\theta}$:

$$p(\vec{\theta} | d) = \frac{p(\vec{\theta})p(d | \vec{\theta})}{p(d)} \quad (2.1)$$

where [149]:

$$p(d | \vec{\theta}) \propto \exp [-(d - h | d - h)/2] \quad (2.2)$$

is the Gaussian likelihood, $p(\vec{\theta})$ is the prior distribution on the source parameters, $p(d)$ is the evidence, and $(\cdot | \cdot)$ symbolises the noise-weighted inner product:

$$(a | b) \equiv 2\Re \epsilon \int_{f_{\min}}^{f_{\max}} \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_n(f)} df \quad (2.3)$$

Here, \tilde{a}, \tilde{b} represent the Fourier transform of the time series $a(t), b(t)$; $[f_{\min}, f_{\max}]$ is the frequency range over which the inner product is evaluated; and $*$ represents complex conjugation.

Now consider two segments of data, $d_1(t)$ and $d_2(t)$, both of which are known to contain one GW signal each, $h_1(t)$ and $h_2(t)$, respectively. We now wish to determine which of the two hypotheses, \mathcal{H}_L and \mathcal{H}_U , is preferred by the data at hand. \mathcal{H}_L is the hypothesis that $h_1(t)$ and $h_2(t)$ are lensed copies of a GW signal originating from a single source. On the other hand, \mathcal{H}_U is the hypothesis that $h_1(t)$ and $h_2(t)$ are signals originating from two distinct, unrelated, sources.

As shown in [101], (in the absence of any prior knowledge of which of the hypotheses is preferred), the posterior overlap (PO) statistic to quantitatively determine the preferred hypothesis is the Bayes factor \mathcal{B}_U^L , defined as the ratio of the evidences of the joint data set $\{d_1, d_2\}$ given each of the hypotheses.

$$\mathcal{B}_U^L \equiv \frac{p(\{d_1, d_2\} | \mathcal{H}_L)}{p(\{d_1, d_2\} | \mathcal{H}_U)} = \int \frac{p(\vec{\theta} | d_1)p(\vec{\theta} | d_2)}{p(\vec{\theta})} d\vec{\theta} \quad (2.4)$$

This Bayes factor can be evaluated making use of the posteriors $p(\vec{\theta} | d_1)$ and $p(\vec{\theta} | d_2)$ estimated from the two data sets d_1 and d_2 , as well as the prior $p(\vec{\theta})$ employed in the parameter estimation.

2.3 Classification with Machine Learning

In the language of machine learning (ML), determining whether a pair of GW events are lensed copies of a single GW event, or unrelated (unlensed) to each other, is a binary classification problem. Using features derived from the data surrounding pairs of GW signals, we can in principle train an ML algorithm to classify them as either lensed

or unlensed. In this subsection we first describe the construction of the features we use, the ML algorithms we employ, along with their training, testing and optimisation.

2.3.1 Data Representation

The PO statistic crucially relies on a time-consuming way of representing the detector data, viz., the posterior distributions of source parameters inferred from the data surrounding the confirmed GW detections. To bypass this issue, we construct and train a machine learning model which takes as inputs time-frequency maps (*Q-transforms* of the GW event), as well as localisation skymaps (*Bayestar Skymaps*). Both of these can be produced within seconds, in contrast to sampling the full posterior on the source parameters which can take anywhere from several hours to several days.

Q-Transforms: *Q-transforms* [147] are a means by which time-frequency maps of generic transient signals can be produced. This is achieved by first representing the time-frequency plane as a collection of tiles (bins), and then reconstructing these generic signals as a combination of sine-Gaussians defined by their quality factor ‘*Q*’. The choice of ‘*Q*’ in each tile is determined from a matched-filter search across multiple ‘*Q*’ templates, and the template that produces the largest SNR is selected. Using the corresponding optimal sine-Gaussian, a spectrogram is generated. The time-frequency map is then plotted as colored tiles, where the color represents the so-called “normalized signal-energy”, which is proportional to the *Q*-transform magnitude (and related to the SNR).

As shown in Fig. 2.1, lensed events will have time-frequency maps whose shapes are similar, but whose signal energies across time-frequency tiles will differ in magnitude. This is a direct consequence of the fact that the phase evolution of strongly lensed pairs are expected to be identical, but the amplitudes will differ by a constant factor. On the other hand, unlensed signals will have distinct time-frequency maps with dissimilar shapes in general. ¹

Bayestar Skymaps: “*Bayestar*” [148] is the flagship low-latency skylo-

¹ A constant (additive) phase-factor called the Morse-phase, which is an integral multiple of $\pi/2$ depending on image type, will in general change the coalescence phase of the dominant GW mode [142, 150]. Note that *Q*-transforms are independent of coalescence phase, and are therefore unaffected by the Morse phase.

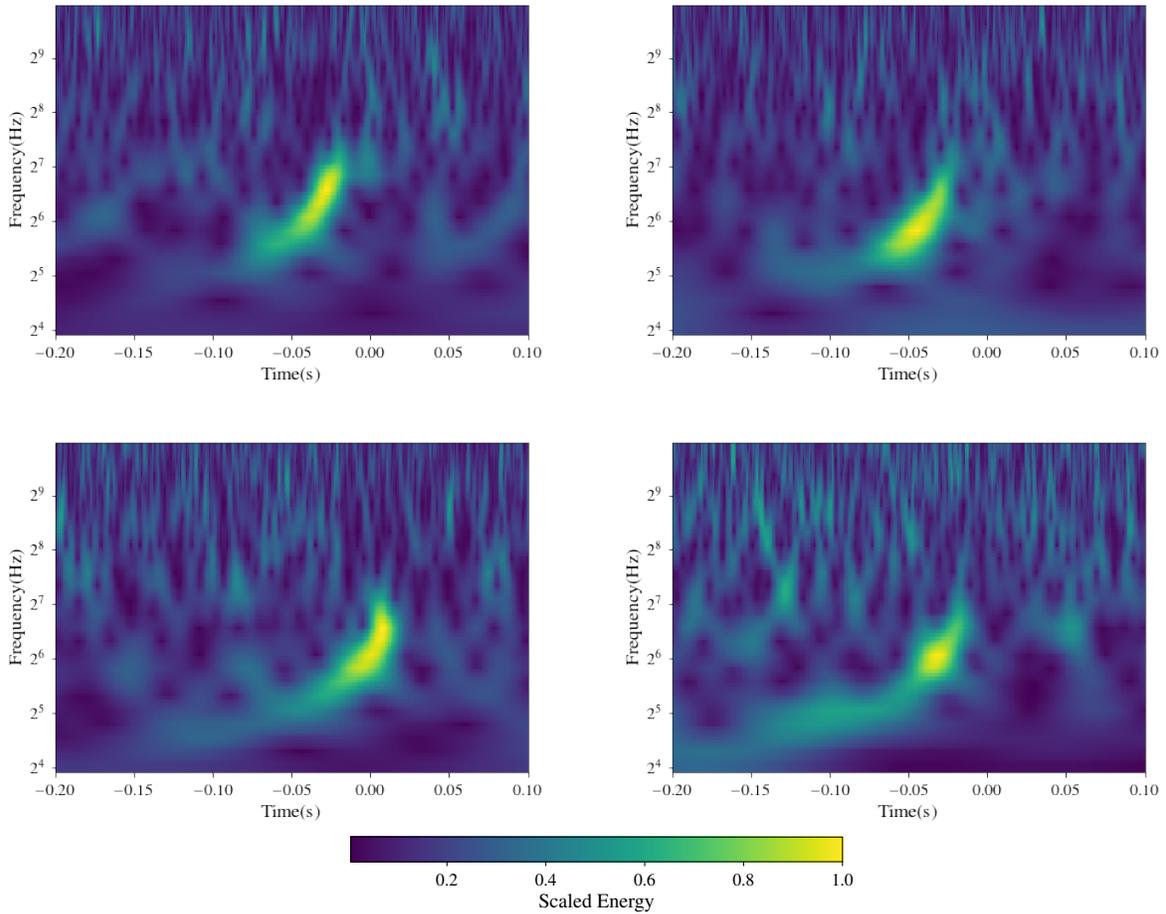


Figure 2.1: *Top Panels:* A pair of lensed GW events detected by the H1 (Hanford) interferometer at design sensitivity. These events have time-frequency tracks with similar shapes. However, the signal energy in different time frequency bins along their tracks differ with respect to each other. *Bottom Panels:* A pair of unlensed GW events projected detected by the H1 interferometer at design sensitivity. These events have time-frequency tracks whose shapes are significantly different.

calisation software of the LIGO-Virgo-Kagra (LVK) collaboration, used during the LVK's third observing run (O₃) to disseminate skymaps in real-time for electromagnetic follow-up of GW events [151]. These skymaps are produced in seconds, and are found to be comparable to those estimated from a full sampling of the joint posterior distribution of the source parameters. Bayestar exploits the fact that errors in sky localisation and the errors in the inference of the source masses, are semi-independent. Given that this software is exclusively focussed on providing localisation skyareas, it exploits this semi-independence to drastically reduce the dimensionality of the parameter estimation problem by fixing the intrinsic parameter values to those of the maximum likelihood template in the matched filter search that identified

the event. It is thus able to evaluate the (dimensionally-reduced) posterior on the extrinsic parameters rapidly, without significant loss in precision.

As shown in Fig. 2.2, lensed events are expected to have overlapping localisation skyareas, by virtue of the poor ($\mathcal{O}(10)$ sq. deg.) angular resolution of ground based GW detectors with respect to the typical angular separation of the images ($\mathcal{O}(1'')$). On the other hand, unlensed signals will generally have non-overlapping skymaps.

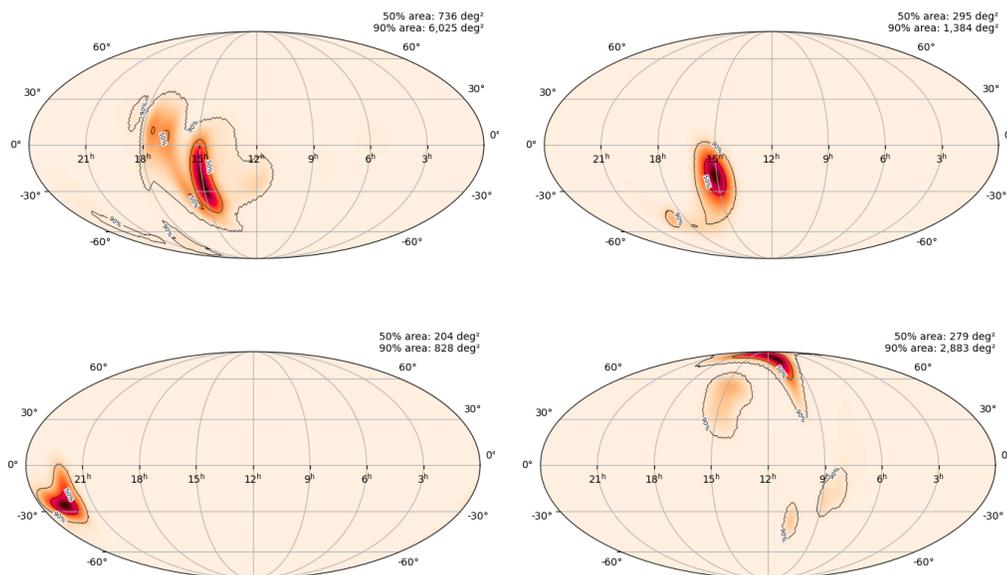


Figure 2.2: *Top Panels:* Bayestar skymaps of a pair of lensed events detected by the H1 (Hanford), L1 (Livingston), V1 (Virgo) network at design sensitivity. The skymaps of these events overlap. *Bottom Panels:* Bayestar skymaps of a pair of unlensed events detected by the H1, L1, V1 network at design sensitivity. The skymaps of these events do not overlap.

2.3.2 Data Preparation

In order to train, optimize, and test our machine learning models, we simulate the lensed and unlensed GW signals, and inject them in Gaussian noise. Our events consist of non-spinning binary black hole mergers detectable by the LIGO-Virgo network at design sensitivity, where detectability is defined by setting a threshold of 8 on the network SNR.

We follow [101] to generate a set of strongly lensed pairs of GW events, where the source BBH mergers follow a well-motivated distribution of masses and redshifts, and the lenses are assumed to be galaxies that can be modelled as singular isothermal ellipsoids whose

parameters are drawn from the SDSS galaxy population catalog [152]. We generated ≈ 2800 detectable lensed event pairs and ≈ 1000 unrelated events, which corresponds to half a million unlensed pairs. We subdivide this set into two sets; we use one for training, and the other for validation. For testing, we use another, distinct, set, although the general prescription still follows [101]². This set consists of ≈ 300 lensed pairs and ≈ 1000 unrelated events (half a million unlensed pairs). From here on out, we refer to the training and validation data set as “DSTrV”, and the testing set as “DSTe”.

² This data set is chosen for testing because the PO statistic was already evaluated for the candidate pairs in this set (and reported in [101]), which allows for a ready comparison with the ML statistic.

The waveforms are generated using the approximant IMRPHENOMPv2 [153, 154, 155], as implemented in the LALSIMULATION module of the LALSUITE software package [156]. The waveforms are then projected onto the LIGO and Virgo detectors using their antenna pattern functions, as implemented in the PYCBC [157] software package.

The detector noise is assumed to be Gaussian, and is generated using the zero-detuned high-power PSDs of Advanced LIGO and Advanced Virgo at their design sensitivities [158, 159], as implemented in PYCBC. The projected waveforms are then added to the detector noise strain to produce the total detector strain time series.

From the time series surrounding each GW event, we generate Q-transform images for each detector. For events whose primary mass $m_1 > 60M_\odot$, we set the range of quality factors to (3,7); otherwise, we set the range to (4,10). Further, using the same time series’, we use Bayestar to generate the localisation skymaps for all the events.

2.3.3 Feature Construction

Comparing the shapes of two time-frequency maps can be interpreted as a problem of image recognition, and therefore lends itself nicely to a machine-learning (ML) analysis designed for such problems. Motivated by the fact that the Q-transform based time frequency maps of lensed pairs will have similar shapes (though different signal energies across time-frequency tiles), while unlensed pairs while have dissimilar shapes in general, we superimpose the time-frequency maps of candidate pairs by aligning them along the time axis, which we pass

to our machine learning algorithm.

On the other hand, while lensed pairs will have overlapping skymaps and unlensed pairs will not, the *shapes* of these maps are not in general expected to be the same, since the relative position of the two images with respect to the detectors are, in general, different (due to the rotation of the earth). However, GW events' localisation skymaps are probability density functions in the space of right-ascension (α) and declination (δ). Thus, a skymap can be thought of as a two-dimensional matrix where each element gives the probability density evaluated at a given pixel in the skymap's image grid spanning the space of (α, δ) . The products of simple operations involving the matrices of candidate pairs can then be used as features that ML algorithms can employ to identify lensed events.

The Bayestar localisation skymaps are usually generated in .fits format, which contains the skylocalisation posterior information sampled over an adaptive HEALPIX grid [160]. We project them to cartesian coordinates using the HEALPY python library [161, 162], which gives us the localisation posterior evaluated over a 400×800 rectangular grid of pixels corresponding to (α, δ) pairs. Denoting the skylocalisation posteriors of each of the events pertaining to a candidate lensed pair as $P_{ij}^1 = P(\alpha_i, \delta_j | d_1)$ and $P_{ij}^2 = P(\alpha_i, \delta_j | d_2)$, we can construct the following metrics which can serve as features using which we can train an ML algorithm:

$$\begin{aligned}
 k_1 &= \sum_i \sum_j P_{ij}^1 P_{ij}^2, & k_2 &= \sum_i \sum_j |P_{ij}^1 - P_{ij}^2| \\
 k_3 &= \sqrt{\langle (P_{ij}^1 P_{ij}^2)^2 \rangle - \langle k_1 \rangle^2}
 \end{aligned}
 \tag{2.5}$$

k_1 is motivated by the PO statistic [101], k_2 is the absolute difference between the elements of the matrices, while k_3 is a standard deviation-like metric of the overlap between the skymaps. Note that angular brackets signify averaging over the total number of elements in each matrix.

2.3.4 Overall Flow

For simplicity, we build two sets of ML models - one that learns from Q-transforms and another that is fed with skymaps - to classify the event pairs as either lensed and unlensed. The models employ two different ML algorithms – DENSENET201 [163] and XGBOOST [164] (see Sec. 2.3).

The first set consists of three DENSENET201 ML models trained on superimposed QT (Q-Transform) images of the event pairs for each of the three detectors: H1 (Hanford), L1 (Livingston) and V1 (Virgo), operating at their design sensitivities. We further construct an XGBOOST model trained on the output of the DENSENET201 models. The output of this XGBOOST model gives us the probability of the lensing hypothesis, given the Q-transform images: $P(\mathcal{H}_L|QT_1, QT_2)$ ³.

We construct another XGBOOST model trained on the metrics derived from pairs of lensed and unlensed Bayestar skymaps. The output of this XGBOOST model gives us the probability of the lensing hypothesis, given the Bayestar skymaps: $P(\mathcal{H}_L|SM_1, SM_2)$.

The final output of our ML classifier is then given by:

$$P(\mathcal{H}_L|\{QT_1, QT_2\}; \{SM_1, SM_2\}) = P(\mathcal{H}_L|QT_1, QT_2) \cdot P(\mathcal{H}_L|SM_1, SM_2) \quad (2.6)$$

We summarize the overall flow of our classification scheme in Fig. 2.3.

2.3.5 Machine Learning Models

In this subsection, we briefly summarize the ML algorithms we use: DENSENET201 and XGBOOST.

DENSENET201: A number of supervised machine learning algorithms exist for binary classification problems. However, only a relatively small subset of these are particularly suited for image recognition. Among them is the DENSENET ML [163] algorithm, which is a kind of convolutional neural network (CNN) with important improvements to mitigate problems that typically plague CNNs. A CNN, in turn, is a category of artificial neural networks (see, e.g. [165]) often used

³A more complete notation for this probability would be as follows: $P(\mathcal{H}_L|\{QT_1^H, QT_2^H, QT_1^L, QT_2^L, QT_1^V, QT_2^V\})$. However, for notational simplicity, we omit the reference to the interferometers.

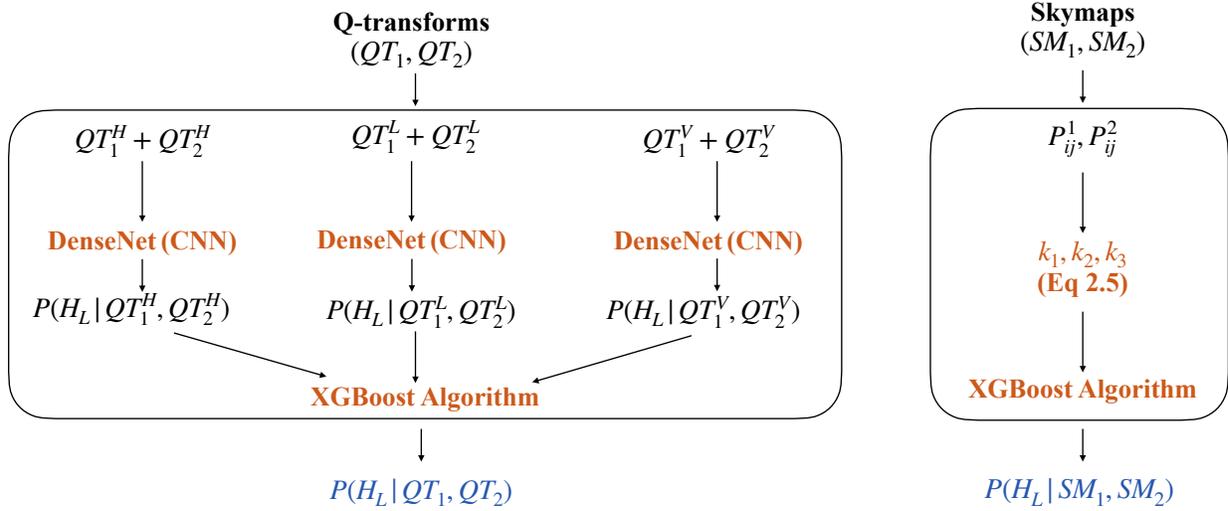


Figure 2.3: A visual representation of the overall flow of our ML classification scheme. Note that, in principle, one could have avoided the step that trains a second XGBOOST algorithm on features derived exclusively from the skymaps, and instead just used one XGBOOST that jointly trains on features from the skymaps and the outputs of the DENSENET algorithms. We found that both methods give similar results. We therefore choose to include the additional XGBOOST because it facilitates a stepwise analysis of the outputs of the individual components of the overall flow, trained separately on intrinsic and extrinsic parameters of the candidate pairs.

for classification problems that involve images, image recognition and computer vision (see, for example, [166]).

The basic architecture of a neural network consists of input/output layers of neurons, and a set of hidden layers in between [167]. Each neuron holds a number between in the range $[0, 1]$. An image passed to a neural network would fill the neurons of the input layer with values corresponding to the pixels of the image grid. The classification prediction of the neural network is recorded in the neurons of the output layer; specifically, in a binary classification problem such as ours, the output layer has one neuron representing the probability that the pair of superimposed Q-transforms corresponds to the “lensed” case.

The neurons in each hidden layer are derived using a two step process. The first step involves a linear operation between the vector of neurons \vec{a} in the previous layer, and a matrix of weights \overleftarrow{W} , and the second a non-linear operation that maps the output of the linear operation to numbers in the range $[0, 1]$:

$$\vec{a}_{n+1} = f(\overleftarrow{W}_n \cdot \vec{a}_n + \vec{b}_n) \quad (2.7)$$

Here, the non-linear function f is referred to as the “activation func-

tion”; common choices include the “sigmoid function” and “the rectified linear unit” (ReLU) function (see, e.g., [168]). Further, the vector \vec{b} is called the “bias”. This process is applied iteratively until the output layer is filled.

Training the neural network ultimately comes down to determining an optimal choice of weights matrices and bias vectors. This can be achieved by feeding the neural network with labelled data, and penalizing the network’s incorrect predictions using an appropriately defined cost function. The popular choice of cost function for binary classification is the binary cross entropy:

$$L_{\text{CE}} = -[y \log(p) + (1 - y) \log(1 - p)] \quad (2.8)$$

where y is the ground-truth (“lensed” = 1 or “unlensed” = 0) of the labelled data, and p is the neural-network’s predicted value for a given choice of weights and biases. Minimizing the loss function averaged over multiple training instances with distinct labelled data, using gradient descent, provides the required weights and biases.

In CNNs, some of the hidden layers perform convolution operations between the previous layer, and appropriately chosen filters, in place of the operation described in Eq. (2.7). The filter can be thought of as a matrix whose size is usually smaller than the matrix of pixels input to the CNN. The convolution operation then involves “sliding” the filter across the pixel grid matrix, which mathematically amounts to taking the product of the filter with each of the submatrices of the pixel grid matrix. The resulting output is sometimes referred to as a “feature map”.

A DENSENET is a type of deep CNN. In addition, its architecture has a few modifications to alleviate some of the problems commonly faced when using CNNs. DENSENET are based on the observation that CNNs can be substantially deeper, more accurate, and computationally efficient to train if there are shorter connections between the layers close to the input and those close to the output. Thus, in a typical DENSENET model, for each layer, the feature maps of all preceding

layers are used as inputs. Furthermore, the current layer’s own feature map is used as input to all the subsequent layers. Because of this type of architecture, DENSENET models have several advantages compared to other CNN models. They greatly reduce the number of parameters that define the architecture of the neural network, mitigate the vanishing-gradient problem, encourage feature reuse and strengthen the feature propagation through the network.

XGBOOST: eXtreme Gradient Boosting, (XGBOOST) [164] is a type of ensemble classifier that uses the combined output of a collection of trained decision trees to provide a probabilistic prediction of class-membership to data that needs to be segregated into discrete categories. A decision tree, in turn, learns from training data by iteratively placing linear cuts in feature-space which minimizes an appropriately chosen loss function. The repeated splits result in the segregated data being pushed down two separate branches at each leaf node in the tree, starting from the root-node where the first split in the training data takes place, and ending at leaf nodes where a terminating criterion (e.g: minimum number of samples in a leaf) has been satisfied.

“Bagging” (see, e.g., [169]) and “boosting” (see, e.g, [170]) are two ways in which the outputs of decision trees can be combined. In bagging, bootstrapped copies of the training data are passed to a collection of decision trees. The trees are then fitted, in parallel, to the training data they receive, and the final prediction of the classifier is an average over all the outputs across the ensemble of trees [171]. In contrast, boosting algorithms such as XGBOOST, fit decision trees to training data sequentially, where each subsequent tree improves on the errors in the predictions of class probability of the preceding tree.

In eXtreme Gradient Boosting, the iterative process of incrementally improving the prediction of the classifier with every fitted decision tree, reduces to minimizing the following objective function [164]:

$$\mathcal{L}_{t+1}^{\text{obj}} = \sum_i \mathcal{L}(y^i, p_{t+1}^i = p_t^i + O_t) + \gamma T + \frac{1}{2} \lambda O_t^2 \quad (2.9)$$

where, as before, y^i is the ground truth of training data point i , $p_t^i(p_{t+1}^i)$ is the classifier's predicted probability of class membership after the sequential fitting of t ($t + 1$) trees. For binary classification problems such as the one we are trying to tackle, the loss function \mathcal{L} is simply the binary cross-entropy defined in Eq. (2.8) (summed over the entire training set), and O_t is the output of the decision tree t with respect to which the objective function is to be minimized. The piece $\gamma T + \frac{1}{2}\lambda O_t^2$ in the objective function is a regularization term that controls the classifier's tendency towards overfitting by reducing its sensitivity to individual training data points. Here, T is the total number of leaves in a tree, and λ, γ are hyperparameters that can be appropriately set depending on the data at hand.

Minimizing \mathcal{L}^{obj} for each decision tree (which can have a vast variety of structures) is in general highly complicated. XGBOOST thus simplifies the minimization process in two ways. The first is that the loss function is approximated by a second-degree Taylor polynomial in O_t . The second is that within each tree, the objective function is repeatedly minimized at each leaf node. As a result, the process of fitting a decision tree reduces to maximizing the gain when splitting the training data at each leaf node. The gain is defined as the difference between the sum of the similarity scores of the two daughter nodes post the split, and the similarity score of the parent leaf. The similarity score at a leaf node l (containing N_l training samples) in tree $t + 1$ is defined as [164]:

$$S_{t+1}^l = \frac{(\sum_i^{N_l} R_t^i)^2}{\sum_i p_t^i(1 - p_t^i) + \lambda} \quad (2.10)$$

where the sum is taken over all the samples in the leaf node, and $R_t^i \equiv p_t^i - y^i$ is the residual of the i th training data point in the leaf node. The output of each tree, defined as $S_{t+1}^{\text{term}} / \sum_i R_t^i$ for the terminal leaf node, is then rescaled by a user defined learning rate η and then added to the log of the odds ratio corresponding to p_t^i , from which the probability estimate of tree $t + 1$ can be trivially computed.

As mentioned earlier, λ, γ are user defined regularization paramete-

ters that control overfitting. Specifically, γ sets a threshold on the gain; leaves along branches whose gains do not exceed γ are pruned. Thus, since positive values of λ tend to reduce the gain, λ effectively encourages pruning, which in turn reduces the sensitivity of the decision tree to individual training data points. ⁴

⁴ In ML literature, λ is often referred to as a “regularization parameter” and γ is referred to as a “tree complexity parameter”.

2.3.6 Training and optimisation

DENSENET201 : We use a DENSENET pretrained on the “Imagenet dataset” [172], which allows it to pick up features common to most images. We then add fully connected layers to it, along with the final layer of just one neuron, for our binary classification, and then retrain it with data specific to our problem (to wit, the superimposed Q-transforms). This method of pretraining with a generic data set and then retraining with a more specific one, is called “transfer learning”. The most significant benefit of this method is that it reduces the size of the dataset required for training and solving the problem at hand.

For each of the three detectors H1, L1 and V1, we train three individual DENSENET201 models using superimposed Q-transform pairs, where each image corresponds to a 3-dimensional array ($128 \times 128 \times 3$) of pixels ⁵. The DENSENET model is loaded with the imagenet weights using the neural network package [173]. To make it suitable for our binary classification task, its top layer is removed and a dense layer of 256 neurons with the ReLU activation function is added along with the final output layer of a single neuron with a sigmoid activation function. Each of the three models is trained on an equal number (1400) of lensed and unlensed Q -transform image pairs subselected from the DSTRV dataset using TPU (Tensor Processing Unit) hardware, which is available in a KAGGLE notebook [174]. In the top fully connected layer of the network, we use the sigmoid activation function (see. e.g., [175]) and we employ the Adam optimizer [176] for efficient gradient calculations. The model prediction is validated using a validation set sub-selected from the total training set.

⁵ Each pixel contains RGB values that correspond to the normalized signal energy at discrete time-frequency coordinates in the Q-transform image.

XGBOOST : As described in the previous section, XGBOOST has a number of tunable hyperparameters that need to be set based on the problem at hand.

The hyperparameter “n_estimators” sets the number of decision trees in the ensemble classifier that are to be fit to the training data sequentially. It can equivalently be thought of as the number of fitting iterations the model goes through as it sequentially improves the prediction of the ensemble classifier. We set n_estimators to 110. The learning rate, regularization parameter and tree complexity parameter are set to their default values of 0.3, 1, 0 respectively. The maximum depth of each decision tree is set using max_depth = 6.

In addition, we also set the “scale_pos_weight” parameter to 0.01. This hyperparameter serves as a weight to account for training data being biased towards one class – in our case, the unlensed class, for which we had about 100 times more data points than for the lensed class.

The first XGBOOST model is trained on the features derived from lensed and unlensed pairs of skymaps, described in Sec. 2.3.3, using the “DSTrV” dataset. Additionally, a second XGBOOST model is trained on the outputs of each of the three DENSENET models. The outputs of the two XGBOOST models are then combined (cf. Eq 2.6) to provide a ranking statistic for candidate lensed pairs.

2.4 Results

2.4.1 Testing and Cross-validation

We assess the performance of the trained ML models on the “DSTe” dataset. This allows us to compare their performance with the PO statistic, which is already computed for this dataset [101]. We summarize the performance of the ML models and the PO statistic with ROC⁶ plots of efficiency vs false positive probability (FPP), where efficiency is the ratio of accurately classified lensed events to the total number of lensed events, and FPP is the ratio of wrongly classified unlensed events to the total number of unlensed events.

⁶ Receiver Operating Characteristic

To check the robustness of the outputs of the machine learning models to changing training sets, we use stratified k -fold cross validation. We implement cross validation by doing a round-robin of dividing our dataset into $k = 3$ ($k = 10$) parts for the DENSENET (XGBOOST) models, using one part for validation and the rest for training. We test the k trained machines with the DSTe dataset.

2.4.2 ROC Plots

We evaluate the performance of the overall classifier and its different components using ROCs. For comparison, we also plot the ROCs for the PO statistic. We first test the performance of the individual DENSENET models trained on Q-transforms pertaining to each of the three detectors: H1, L1 and V1. We then test the XGBOOST model trained on the outputs of the DENSENET models. Since we used cross-validation to assess the robustness of the models, we trained and validated each of the models on the different cross-validation subsets of the DSTrV data set, and tested the differently trained models on the DSTe data set. This gives us an estimate of the variation of the ROCs due to differences in the training set.

Fig. 2.4 plots ROCs for the outputs of these models trained on Q-transforms. The ROC for the PO statistic constructed using parameter estimation posteriors on the component masses (m_1, m_2) , is also plotted for comparison. The ROCs pertaining to the individual DENSENET H1, L1, V1 models perform similarly to the ROC for the PO statistic, both at low and high false positive probabilities. The mean ROC corresponding to the XGBOOST model trained on the outputs of the individual DENSENET models performs comparably to the PO statistic. At very low FPPs, ML seems to perform about 1.5 – 2 times better than the PO statistic. However, there is some variation in the XGBOOST model's ROC due to the changing training set. These improvements must therefore be interpreted with some caution. As the variation in the ROCs at these FPPs suggests, low-number statistics are likely causing the ROC to be sensitive to changes in the training set.

Fig. 2.5 plots ROCs for the XGBOOST model trained on the features

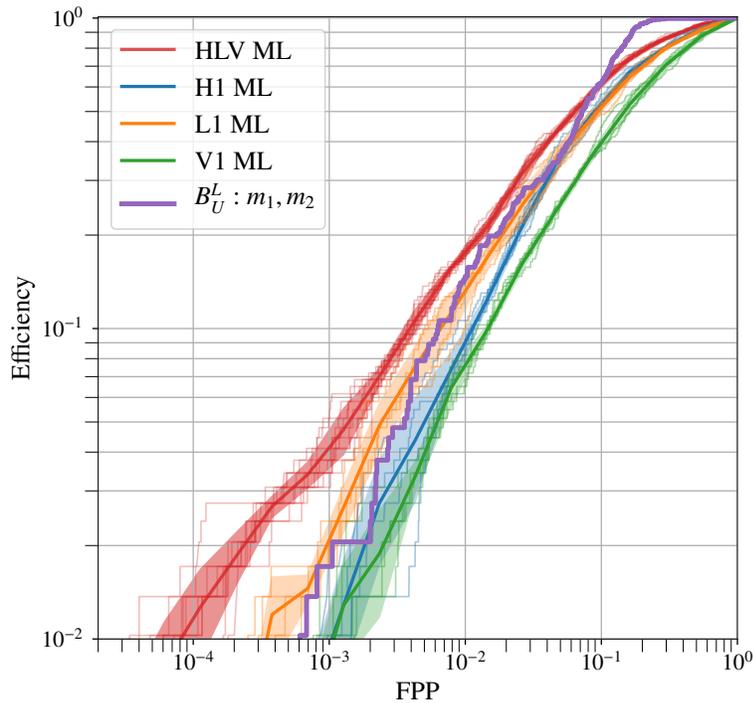


Figure 2.4: ROCs for DENSENET models trained on lensed and unlensed pairs of superimposed Q-transforms, for different cross-validation subsets of the DSTRV training set. ROCs for models trained on Q-transforms corresponding to individual detectors are evaluated, in addition to ROCs pertaining to the XGBOOST model trained on the outputs of the individual DENSENET models. For comparison, the ROC for the PO statistic that uses parameter estimation posteriors on the component masses, m_1, m_2 , is also plotted. At low false positive probabilities, the individual DensetNet models perform comparably to the posterior-overlap statistic. On the other hand, the XGBOOST model produces efficiencies that are 1.5 – 2 times better than the PO statistic at low FPPs, although there is some variation in the ROCs when the training set is changed, caused by small-number statistics. These improvements at low FPPs must therefore be interpreted with some caution.

(metrics) derived from pairs of Bayestar skymaps. Each ROC pertains to a different cross-validation subset of the DSTRV dataset. The ROC for the PO statistic evaluated using only the right-ascension (α) and declination (δ) is plotted for comparison. The XGBOOST performs as well as the PO statistic at low false positive probabilities, although at higher false positive probabilities the latter performs marginally better. As with the DENSENET models, there is some variation in the ROCs when the training set is varied.

Fig. 2.6 plots ROCs for the overall classifier, which is an XGBOOST model trained on the outputs of the DENSENET models and the first XGBOOST model. For comparison, the ROC for the PO statistic evaluated using the parameter estimation posterior on m_1, m_2, α, δ is also plotted. The mean ROC for the overall classifier performs almost identically to the PO statistic at low false positive probabilities, although at higher false positive probabilities the PO statistic performs marginally

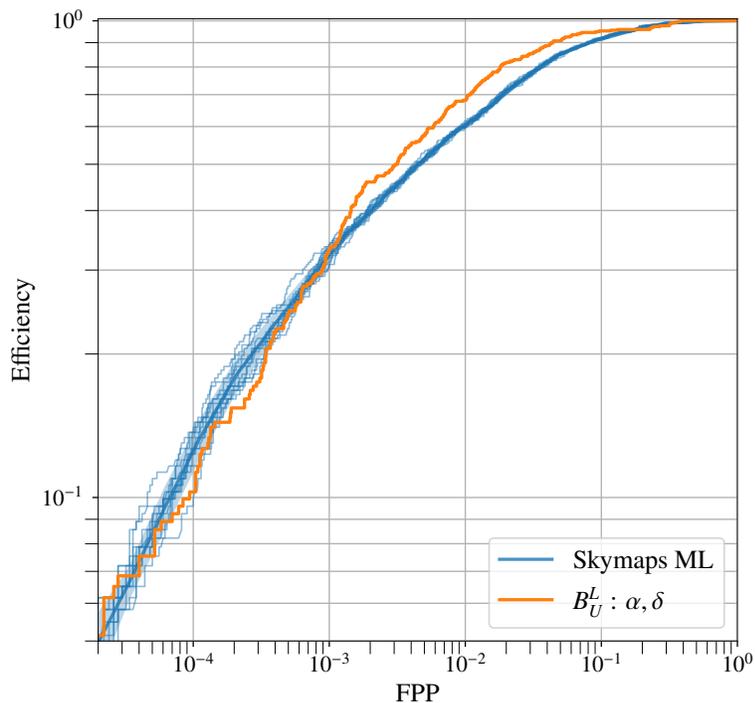


Figure 2.5: ROCs for the XGBoost model trained on metrics derived from pairs of Bayestar localisation skymaps, for different cross-validation subsets of the DStRv training set. For comparison, the ROC for the PO statistic that uses parameter estimation posteriors on the skylocation coordinates, α, δ , is also plotted. The XGBoost performs almost as well as the PO statistic, at low false positive probabilities.

better.

2.5 Summary and Outlook

GW observations of BBH events is expected to increase significantly in future observing runs, with $\mathcal{O}(10^2)$ events during O4 and $\mathcal{O}(10^5 - 10^6)$ during the 3G era. The number of candidate lensed pairs to classify could therefore be as high as $\mathcal{O}(10^4)$ and $\mathcal{O}(10^{10} - 10^{12})$, respectively. Current optimal Bayesian methods, such as the PO statistic, rely on the parameter estimation posterior on the source parameters, which could take anywhere from several hours to several days to sample.

This therefore motivates the need to come up with a preliminary classification scheme, that can rapidly rule out the vast majority of unlensed candidates. To that end, as a proof-of-principle, we construct a machine learning based classifier that can classify pairs of

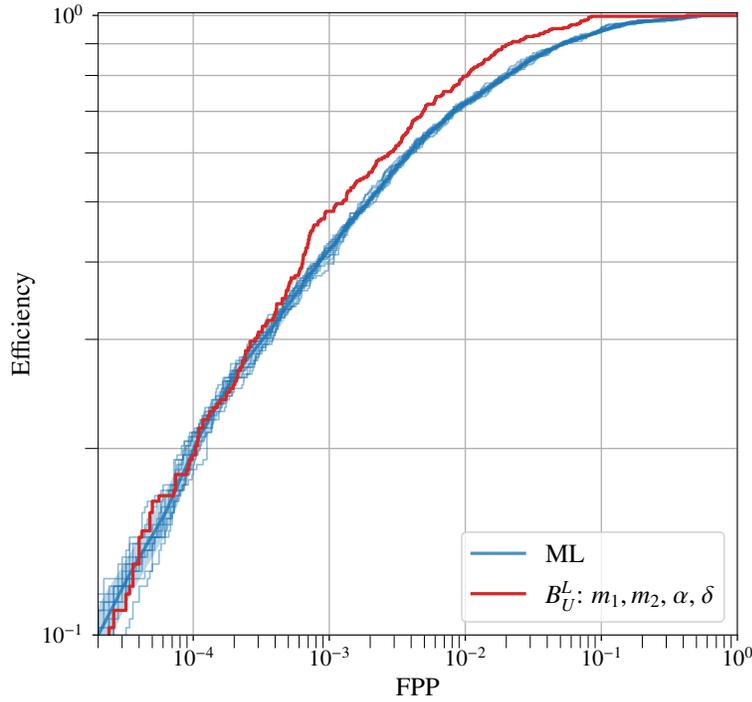


Figure 2.6: ROCs for the overall classifier, for different cross-validation subsets of the DStRV training set. Note that the output of the overall classifier is the output of the XGBOOST model trained on the outputs of the three DENSENET models pertaining to H1, L1 and V1, as well as the output of the first XGBOOST model trained on Bayestar skymaps. At low false positive probabilities, the classifier performs almost identically to the PO statistic, with mild variation in the ROCs when the training data set is varied.

non-spinning BBH events in seconds. We use two ML algorithms: DenseNet201 and XGBoost, to build models trained on time frequency maps and Bayestar skymaps of pairs of events. We construct 3 DenseNet models trained on GW events projected onto each of the three detectors in the LIGO-Virgo network at design sensitivity. The outputs of these models are fed to an XGBoost classifier to construct a corresponding model. The output of this model is then combined with the output of another XGBoost model trained on pairs of lensed and unlensed Bayestar skymaps, to produce the final ranking statistic of our overall ML classifier (cf. Fig. 2.3 and Eq. 2.6).

We train and validate the classifier on cross-validation subsets of the DStRV dataset, and test the performance of the ML classifier (including its different components) on the DStE dataset. We find that the overall ML classifier performs comparably to the PO statistic evaluated from the parameter estimation posterior on m_1, m_2, α, δ . More specifi-

cally, the performance of the ML classifier, as captured by ROC plots, shows that at low false positive probabilities, the classifier performs almost identically to the PO statistic, although at high false positive probabilities, the performance of the latter is marginally better.

Simple benchmarking tests suggest that our trained ML classifier is able to classify each event within 2 – 3 seconds ⁷. Including the time to produce the Q-transform images and Bayestar skymaps, the total classification time is still less than a minute. This is significantly faster than the PO statistic, which takes several minutes to classify once the parameter estimation posteriors are available. Since, in addition, these posteriors themselves can take hours to days to produce, per event, the benefit of using ML to perform a preliminary sweep of lensed candidate pairs to rule out the vast majority of them as unlensed, becomes manifestly evident.

Additionally, rapid ranking of candidate pairs makes estimating a background distribution computationally feasible. Such a distribution enables assigning statistics such as p-values/false positive probabilities, which are often the preferred statistics since they can be interpreted independently of the models used to analyze the pairs. Another potentially useful application of the rapid identification (and dissemination) of lensed GW events is in multi-messenger astronomy, since the joint GW-EM detection of lensed events could enable important tests of general relativity.

It might be worth mentioning that in addition to the PO statistic, there are more comprehensive Bayesian classification methods that take even longer to run. A fully Bayesian, joint parameter estimation scheme to identify lensed pairs by evaluating a coherence ratio that accounts for correlations between parameters of lensed events, and selection effects, currently takes of the order of weeks to complete, per candidate pair [104, 105, 106]. A more approximate joint parameter estimation method that neglects selection effects, is found to identify lensed pairs with similar efficiencies as the full joint parameter estimation method, but within hours instead of weeks [105]. Thus, identifying lensed pairs from the enormous number of candidate pairs in

⁷Note that this time is largely taken up in loading the necessary files for classification. The classification step itself takes less than a second.

future observing runs, can follow a step-wise procedure, where an ML classification method such as ours can rapidly rule out most of the candidate pairs as unlensed. The surviving pairs can then be followed up by the PO statistic, and then by joint parameter estimation methods.

Note that our work assumed stationary Gaussian noise, and that the candidate pairs consist of confirmed, high-significance non-spinning BBH events. We plan to systematically relax these assumptions in future work. Specifically, we are currently looking at the possibility of classifying confident GW events in real noise. We plan to train the machine on events injected in real noise, whiten the data so that the Q-transforms are less sensitive to varying PSDs, and investigate the possibility of using additional features. We are also working towards the classification of marginal BBH events, with an ML scheme similar to what was presented in this work. We hope to report the results of these investigations in the near future.

3 | Search for strongly-lensed gravitational waves in the third observing run of LIGO and Virgo

Abstract

Strong gravitational lensing by galaxy or galaxy clusters along the path of the GWs can give rise to multiple copies of the GW transient signals from CBCs. In the third observing run of LIGO and Virgo, dozens of GW signals have been detected [2]. There is a non-negligible astrophysical possibility that some of these events could be strongly lensed copies of each other. Based on our ML algorithm for lensing classification (see Ch. 2), we developed a pipeline called LENSID, as part of the LVK collaboration for lensing searches in the third observing run [15]. We use both PO and LENSID pipeline, to identify potential lensing candidate pairs from the detected events and then passed them to the more sophisticated pipelines for follow-up analysis. Overall we find no confident detection of strong lensing [15]. Additionally, we analyse the candidate sub-threshold lensing counterparts to the super-threshold events found by the Gst-LAL-based targeted search pipeline to rapidly identify the more interesting lensed super-sub candidate pairs using a new method that we developed. We also re-analyse the most significant candidate “super-sub” pair using PO and find it to be among the more significant candidate pairs, though not sufficiently significant to claim detection of lensing.

Some parts of this chapter (sec 3.2) are part of a publication [15] by the LVK collaboration. I was the primary contributor to this analysis.

3.1 Introduction

As discussed in Ch. 1, the strongly-lensed GWs would appear as repeated events in the LIGO and Virgo detectors, and therefore for

any given pair of detected GW events, we want to identify if they are strongly-lensed copies of each other or not. Given that the rate of strong lensing is non-negligible, we search for strong lensing signatures in the O₃ run data of LIGO and Virgo detectors ¹. Based on our ML model for identifying strongly lensed GW signals (see Ch. 2), we developed a pipeline called LENSID, as part of the LVK collaboration lensing searches in the O₃ data [15]. Here we report results from both PO analysis and ML pipeline for identifying the potential lensing candidate pairs from the detected events. The top $\sim 1\%$ of the candidates from either of the PO and ML pipeline—approximately a hundred pairs—were then passed on to the Joint-PE pipelines such as GOLUM [177] and HANABI [178]. We summarize the results in sec. 3.2. Though some interesting candidates have been found, eventually they are discarded by the follow-up analysis [15].

Additionally, a separate search is done to find the possible sub-threshold lensed counterparts for each of the targeted super-threshold events. This is done by constructing template banks using the targeted event’s posteriors of intrinsic parameters (masses, spins) which helps in reducing the false alarm rates of the event triggers as compared to the standard searches [109]. The Bayesian PE posteriors are not readily available for the subthreshold events, and given the computational costs involved it is practically infeasible to use the PE-based methods to rank the super-sub lensing candidates. Since our ML model was trained on super-threshold events ($\text{SNR} > 8$), we choose to instead do a preliminary analysis using the matched-filter mass estimates and Bayestar skymaps ² to rapidly find the most interesting lensing candidates. This new method and its results are described in Sec. 3.3. The most significant candidate “super-sub” pair deemed by this method is re-analyzed by PO in Sec. 3.3.5 and more sophisticated joint-PE methods [105, 106] and found to be among the more significant candidate pairs, though not sufficiently significant to claim observation of lensing. We point the reader to [179] for more detailed analyses of this event pair.

¹ Although KAGRA joined the third observing run for a short period, since its sensitivity was not comparable to LIGO and Virgo, KAGRA data was not used for lensing analysis.

² Bayestar analysis pipeline provides approximate, but fairly accurate posteriors of the sky location of GW signals making use of the arrival times, single-detector SNRs and the Fisher matrix of the template waveforms [148]

3.2 Analysis of O₃ catalog events

As mentioned in Ch. 1, there are several techniques to search for strongly lensed pairs of events. Among them, the low-latency ones are PO and ML, see Ch. 2. Both of these methods do not take into account the Morse phase information and the selection effects, whereas the more sophisticated joint-PE based pipelines, GOLUM [177] and HANABI [178] that samples over the joint GW-likelihood to compute the lensing Bayes factors do, but they are computationally more expensive. In the LVK O₃ lensing searches, a multi-stage follow-up strategy was adapted to reduce the computational costs for analysing all O₃ event pairs. Here we show the results from the ML and PO analysis that identified about a hundred pairs out of ~ 2200 GW event pairs to be worthy of follow-up analyses.

The PO statistic (given by Eq. 2.4) assesses the consistency between a lensed candidate pair’s posterior distributions of intrinsic parameters, sky location, and inclination angle (and thus acts as a discriminator between the lensed and unlensed hypotheses)[180]. To convert this statistic to an FPP,³ a background distribution of unlensed PO statistics needs to be estimated.

To that end, we conduct an injection campaign involving BBH only, in which we sample component masses $m_{1,2}$ from a power-law distribution [181] in the range $(10\text{--}50M_{\odot})$. We assume that the redshift distribution of BBH is similar to population synthesis simulations of isolated binary evolution [182]. All other parameters are sampled from uninformative prior distributions [180]. We inject the simulated signals into Gaussian noise with O₃a (first half of the O₃ run) representative PSD for a LIGO–Virgo detector network. We compute the PO statistic for all possible pairs in this injection set, and following [9], we assign an FPP to a candidate pair using its PO statistic. Candidate-lensed pairs involving BNS or NSBH events are not analyzed and ranked.

We additionally employ LENSID, an ML-based binary classification pipeline (see Ch. 2) to rapidly provide a probability of class membership (lensed or unlensed) for a given candidate BBH pair [183]. Such

³ FAR and FPP, while conceptually similar, pertain to different contexts in this work. In particular, we use FPP exclusively for significances associated with candidate lensed pairs to discriminate them from unlensed pairs. On the other hand, a FAR is associated with the significance assigned to individual candidate GW signal events.

an analysis not only serves as an independent method to rank candidate pairs but also provides a quantitative significance to pairs for which source-parameter inference samples are unavailable. To convert the LENSID output to an FPP, we construct a background distribution of ML outputs using a population of unlensed BBH events injected in Gaussian noise characterized by the O_{3a} representative PSD – the same as was used for the PO statistic. This PSD is found to be sufficiently similar to the averaged O₃ PSD for the estimation of the background distribution so as not to change the preliminary selection of candidate pairs. The BBH population is identical to the one used by the PO statistic analysis to construct its corresponding background distribution. Furthermore, the sky localisations used to rank candidate pairs come from the same PE analysis used to estimate the PO statistic ⁴.

A plot comparing the FPP assigned by the PO and ML analyses is shown in Fig. 3.1. Candidates that have either a PO-assigned FPP or ML-assigned-FPP, (or both), that are smaller than 1%, are selected for more comprehensive Bayesian analyses.

Out of all the ~ 2200 O₃ catalog event pairs, approximately ~ 100 pairs were passed for follow-up analysis whose PO statistic-based or ML-based FPP fall below the threshold of 10^{-2} , as seen in Fig. 3.1. However, these pairs were ultimately discarded after including the time delay prior information [179] and computing the Joint-PE Bayes factor [15], meaning that the observed overlap is unlikely to be coming from a lensed BBH and is more likely to be coincidental. One such pair, GW191103_012549–GW191105_143521 which we call simply GW191103–GW191105 from now on, was identified as one of the most likely candidates by the PO analysis (see Sec. 3.2.2) using the posteriors obtained with the *IMRPhenomXPHM* waveform [184] released publicly on the Gravitational Wave Open Science Centre (GWOSC) [185], whereas LENSID—which uses Q-transform images and BAYESTAR [148] skymaps—had not classified it as a candidate, see Sec. 3.2.3 for details. Ultimately after further investigations[179], the pair displayed no sufficient evidence to claim observation of strong lensing.

⁴Note that Bayestar, which is used to assign ML probabilities to real-event candidate pairs, is expected to provide sky localisations that are similar to those provided by this PE analysis.

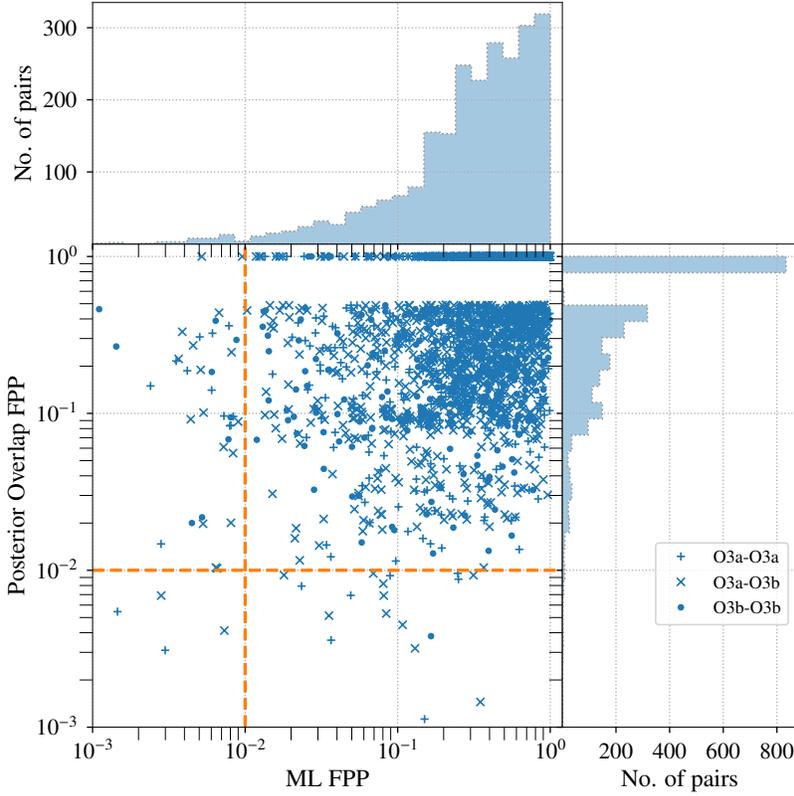


Figure 3.1: The FPP of each lensed candidate pair constructed from the set of GW events that exceed an astrophysical probability [13, 14] threshold of 0.5, as evaluated using the PO and ML classification statistics. Orange dashed lines that correspond to an FPP threshold of 10^{-2} , are also placed. Pairs whose PO statistic-based or ML-based FPP fall below this threshold are selected for additional joint PE analyses. $< 10^{-6}$ has been mapped to an FPP of 1, which is reflected in the gap along the vertical axis between 0.4 and 1.

3.2.1 *Lensing time delay statistic*

Complementary to the PO statistic an additional statistic \mathcal{R}^{gal} is constructed that uses arrival times as a means to discriminate between lensed and unlensed events. We know that for galaxy-scale lenses, the time delays could vary from several minutes to several weeks. For cluster-scale lenses, the time delays could even span months and years. Assuming an appropriate model for the distribution of galaxy-lens parameters, as well as the redshift distributions of lenses and sources, we can construct a distribution of time delays $P(\Delta t|H_L)$ [186] pertaining to detectable strongly lensed BBHs.

The distribution of time delays for unlensed pairs of GW signals $P(\Delta t|H_U)$ can be estimated by assuming the arrival times to follow a Poisson process (see Eq. 31 in [106]). Conversely, the distribution of time delays for detectable lensed pairs depends on the distribution of

lens parameters, source parameters, and the relative separations between the Earth, the lenses and the sources. The \mathcal{R}^{gal} statistic is constructed from time-delay distributions of lensed and unlensed events, assuming galaxy lenses, as [101]:

$$\mathcal{R}^{\text{gal}} = \frac{P(\Delta t_0 | H_L)}{P(\Delta t_0 | H_U)}, \quad (3.1)$$

where Δt_0 is the measured time delay between a given pair of signals. Though this is a model-dependent statistic and valid only for galaxy lenses, it improves our capability of identifying lensed events from unlensed ones. This and the PO statistic 2.4 can be used in combination as a product for efficient identification of galaxy-lensed GW events [101]. In the rest of this chapter we denote the PO statistic \mathcal{B}_U^L defined in Eq. 2.4 as the $\mathcal{B}^{\text{overlap}}$ to avoid confusion with other definitions. The overall PO statistic is $\mathcal{B}^{\text{overlap}} \times \mathcal{R}^{\text{gal}}$

3.2.2 PO analysis of GW191103-GW191105

GW191103 and GW191105 were BBHs detected during the third observing run [16]. In the main LVK analyses, the standard treatment of the signals revealed nothing out of the ordinary for these events. However, when treating the events as potential lensing candidates, the pair display some intriguing characteristics. There is a notable amount of overlap between some of the reported source parameters, such as the sky location and masses. Moreover, the two events have about two days delay between their merger times which is consistent with galaxy-scale lenses [187, 186].

For PO, the pair ended up having the highest overall PO statistic: $\log \mathcal{B}^{\text{overlap}} = 3.03$ and $\log \mathcal{R}^{\text{gal}} = 1.14$ for the SIS model giving a total of 4.17. Fig. 3.2 shows the posteriors of both events where one may see the varying degrees of overlap between the events.

Fig. 3.3 shows the candidate event pairs identified by PO analysis on the $\log \mathcal{B}^{\text{overlap}} - \log \mathcal{R}^{\text{gal}}$ plane considering both the SIS and SIE galaxy models. The choice of model affects only the $\log \mathcal{R}^{\text{gal}}$ value. The PO analysis is marginalised over phase and is, therefore, insen-

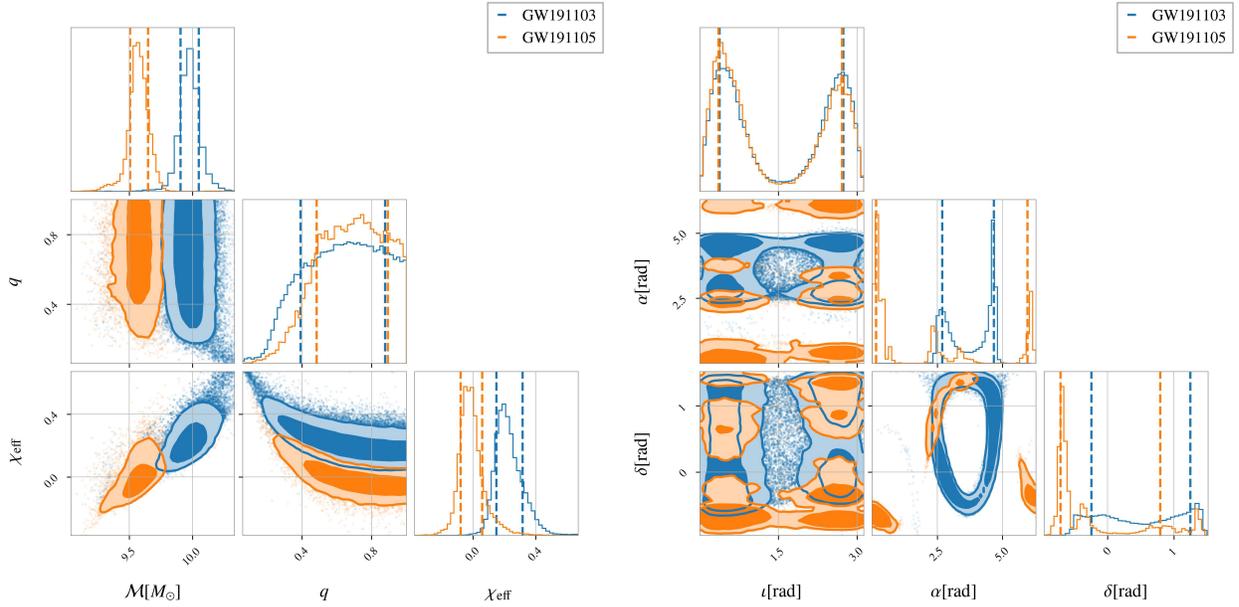


Figure 3.2: Posteriors obtained using the IMRPhenomXPHM waveform for GW191103 (blue) and GW191105 (orange). The overlap in the extrinsic parameters (e.g. sky location) is much larger than that for the intrinsic parameters (e.g. detector-frame chirp mass and spins).

sitive to the relative Morse phase ($\Delta\phi$) between the two events. As a result of this insensitivity, the SIE cases $\Delta\phi = 0$ and $\Delta\phi = \pi/2$ are considered separate models, hence we compute the \mathcal{R}^{gal} expected distributions for each case.

Posteriors of events detected by the LIGO and Virgo detectors can overlap by random coincidence meaning that unlensed pairs could also give high Bayes factors. For this reason, a background injection study with ~ 1000 unlensed events (the combinations of which yield about half a million pairs) is done to calculate the FAP [112] of the observed log Bayes factor for the candidate pair. The FAP per-pair (FAP_{pp}) for the candidate, hence the number of unlensed events with a Bayes factor higher than the one observed for the pair of interest, is found to be 1 in 10,000. Taking into consideration that a total of ~ 68 BBH events were detected in O3 the total FAP including the trials factor (given by $\text{FAP} = 1 - (1 - \text{FAP}_{\text{pp}})^{N_{\text{pairs}}}$) is found to be 0.3 i.e. a significance of slightly above 1σ . As seen in the figure the time delay for this event pair is more compatible with SIE $\Delta\phi = 0$ as compared to SIE $\Delta\phi = \pi/2$ and SIS. After this step, to extract more information about the event pair, it is passed to more extensive pipelines for fur-

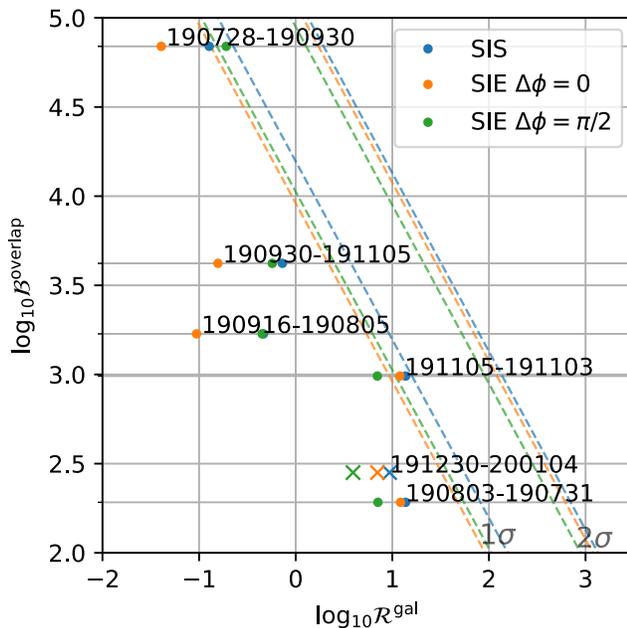


Figure 3.3: The top five candidate strong lensing pairs from the PO analysis considering all the event pairs found based on the O3 data [2]. The dashed lines correspond to the 1σ and 2σ confidence levels for the combined PO statistic ($\mathcal{B}^{\text{overlap}} \times \mathcal{R}^{\text{gal}}$) with different lensing models computed from the background simulations.

ther investigation, finding that the event pair is unlikely to be lensed [179]. Besides GW191103–GW191105, , GW190728–GW190930 is also close to 1σ (see Fig. 3.3). However, the pair has been discarded in previous searches with a lower overall significance than GW191103–GW191105 [9].

3.2.3 LENSID investigation of GW191103-GW191105

As mentioned in the last section, according to PO, the GW191103–GW191105 pair is found to be one of the most significant candidates however, with LENSID its FPP is found to be 0.16, finding it to be not significant for follow-up analysis. We don’t know the exact reason for this result, however, there are some clues as to why it would have happened. Before going into the clues, let us briefly revisit how LENSID works. LENSID is made up of two ML models, one which takes Q-transforms input — ML QTs, and another which takes skymaps as input — ML SMs. The ML QTs output the probability of each event being lensed individually for the three detectors. These outputs are then multiplied to get the overall ML QTs statistic. Likewise, the com-

binned ML statistic for each pair is calculated as the product of the statistics ML QTs and ML SMs. For more details kindly refer to Ch. 2.

The event GW191103 was observed only in two detectors, LIGO Hanford (H1) and LIGO Livingston (L1). Whereas, GW191105 was observed in all three detectors but there was contamination by a glitch in the Virgo detector. As seen in Fig. 3.4, the final PE skymap of the event, which is made after de-glitching the data, is different from the one that was uploaded on GraceDB. For ML SMs, we saw that the sky FPP reduced from 0.08 to 0.02 after using the PE skymap but still did not cross the threshold.

For ML QTs, only the H1 and L1 detector data are used as per the framework. We notice that the Qtransforms for the events, especially for GW191105 are visually poor. They seem to be broken in the middle, as shown in Fig. 3.5. Notice that the Qtransform of GW191105 in the L1 detector has a gap in the middle of the signal with peaks of power on both sides of the gap, which is not expected from a GW chirp signal. We checked that even though the SNRs are similar for both the events in H1 and L1 detectors, the ML QTs statistic varies a lot for the two detectors, 0.86 for the H1 and 0.12 for L1, where a statistic value of 1 means “lensed”, and 0 means “unlensed”. This indicates that the ML algorithm is not robust to the real noise fluctuations, which is expected as the machines were trained using the simulated Gaussian noise signals.

Additionally, from an injection study we found that compared to PO, LENSID is more prone to misclassifying lensed signals with low chirp masses ($< 20M_{\odot}$)⁵, which is the case here. In future, to mitigate this problem, the ML models shall be trained and tested in real detector noises and with signals of lower chirp masses or using the astrophysical population agnostic mass distributions.

3.3 Analysis of the targeted sub-threshold search triggers

It has been suggested in the literature that the rate of lensed events – where one event in a lensed pair is a superthreshold GW event, while

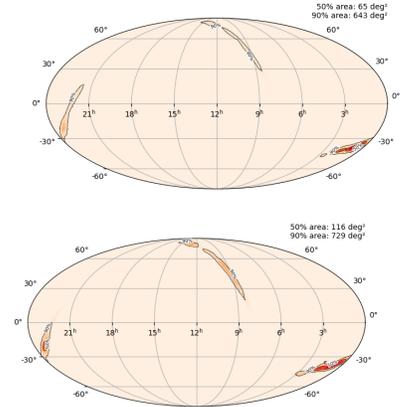


Figure 3.4: Skymap of the event GW191105 from GraceDB created using LALInference (top) and from GWOSC created using BILBY after de-glitching the Virgo data (bottom). The LALInference skymaps are narrower as compared to BILBY ones, likely due to the glitch.

⁵ It is reasonable to expect that the PO method, which relies on precisely estimated posteriors, works better at low masses. It is well known that template-based PE methods work better for low-mass binaries (long chirps) as compared to high-mass binaries at a given SNR. On the other hand, time-frequency detection methods, such as Q-transform, work better when the signal is localised into smaller time-frequency bins. Hence there are reasons to believe that the ML method based on Q-transform works better at high masses. Additionally, LENSID is trained on a realistic injection set of lensed events, which contain more high-mass binaries (as they dominate the lensed events). This may also be a reason for LENSID to work better at high masses.

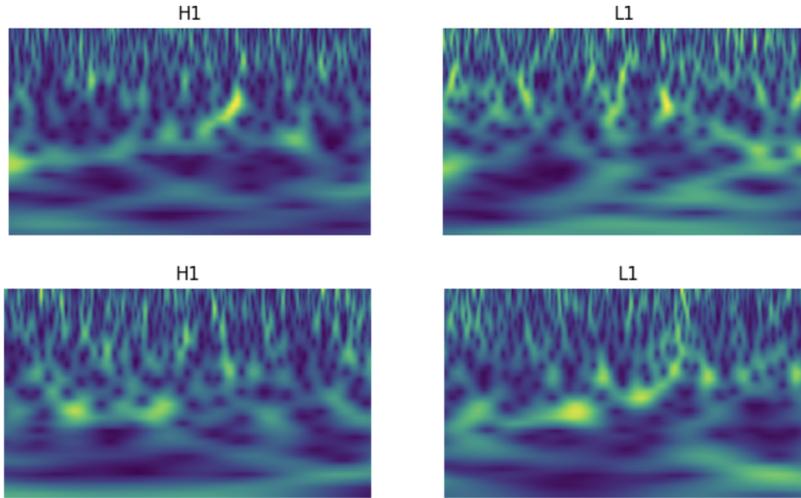


Figure 3.5: Q-transform (or time-frequency maps) images that are input to ML QTs of the LENSID pipeline for the events GW191103 (top panel) and GW191105 (bottom panel). The chirping feature for GW191105 is broken in both the LIGO detectors, whereas for GW191103 the chirp signal is fairly visible in Hanford, and not so visible in Livingston.

the other is a subthreshold event whose reduced significance could be due to one (or more) of several reasons, including de-magnification – is larger by a factor of few than the rate of lensed pairs where both GW events are superthreshold [188]. It is therefore worthwhile to search for such “super-sub” lensed pairs. A lensed GW counterpart can be missed by the usual matched-filter searches due to its low SNR as a result of lensing de-magnification or reduced sensitivity of the detector for certain sky location (relative to the detector). Hence, a separate matched-filter search is done to find the possible subthreshold lensed counterparts for each of the targeted super threshold events [189, 190]. This is done by utilizing the posteriors of intrinsic parameters of the superthreshold events, to construct a reduced template bank that enables a deeper search for subthreshold events by reducing the background noise. The candidate super-sub pairs have been found through these searches during O2 [190, 150] and O3 [15]. However, till now, no confident detection has been made.

While all the superthreshold candidates (events having search false alarm rate (FAR) $< 2/\text{day}$ and $p_{\text{astro}} > 0.5$ as per the GW transient catalogs, GWTC-2.1 [191] and GWTC-3 [2]) have PE posteriors readily available⁶, the subthreshold candidates generally do not. This is in part because there isn’t sufficient evidence to suggest that they are of astrophysical origin and therefore worthy of PE follow-up.

⁶ The posterior samples are found in the data releases [192] and [193].

Moreover, acquiring PE samples for all the subthreshold candidates is computationally taxing, which makes most of the existing lensing-identification methods mentioned above computationally expensive and time-consuming to be feasibly used. This is true even for the low-latency approaches. PO requires PE posteriors on the intrinsic parameters of both events in the lensed pair candidate to be analyzed [101]. On the other hand, the ML-based method needs extensive training and testing involving subthreshold events, which has yet to be completed [103]. To mitigate the increase in candidate pairs by including subthreshold events, we introduce another method that rapidly constructs interpretable, albeit approximate, statistics to rank the candidate lensing counterparts to the super-threshold events that are found by the targetted subthreshold searches. The method then provides a *preliminary* identification of super-sub lensed candidate pairs.

Our method is akin to PO, although the data products used can be generated rapidly without taxing computational resources. In particular, the GW-likelihood is approximated as a Gaussian using a Fisher information matrix [149]. This enables a rapid, though approximate, construction of posterior distributions on the chirp masses of the super/ subthreshold events. For each super-sub candidate pair, the chirp-mass posteriors of the events in the pair are compared using the Bhattacharyya Coefficient [194]. Another coefficient is produced by cross-correlating the Bayestar skymaps [148], generated in low latency, of each of the events in the pair. A third coefficient, exploiting the expected time delay distribution of detected strongly lensed events, is also constructed. The coefficient values are then sorted in descending order to identify any super-sub candidate pair with large values across all three coefficients. In sec. 3.3.1 we describe the method, and validate it with simulations in sec. 3.3.2 while comparing its performance with PO analysis. We finally apply the method to the O₃ super-sub candidate pairs in sec. 3.3.3 while summarizing its results thereafter.

All O₃ super-sub candidate pairs were found to be insignificant by one or more of these statistics, with the notable exception of one solitary outlier, GW191230_180458-LGW200104_180425, which we call

simply GW191230-LGW200104 from now on. This pair lies in the top 5th percentile of all super-sub pairs across all three statistics. We re-analyse the event pair using PO statistic in sec. 3.3.5 and it is also followed up by more sophisticated joint-PE methods [105, 106] in Ref. [179], finding to be among the more significant candidate pairs, though not sufficiently significant to claim observation of lensing.

3.3.1 *A rapid method for preliminary identification of subthreshold strongly lensed counterparts to superthreshold gravitational-wave events*

Given a pair of CBC events, we wish to determine whether they have a common provenance (lensed), or whether they are unrelated (unlensed). As discussed in Ch. 1, in the GO limit, each strongly lensed copy of the GW strain gets an overall magnification, time delay and a (Morse) phase shift, although the phase evolution remains unaffected. Hence, apart from luminosity distance, time of arrival, and coalescence phase, the Bayesian inference of intrinsic and extrinsic parameters should yield posterior distributions that overlap well. The PO statistic exploits this fact to distinguish between the lensed and unlensed candidate pairs, see Ch. 2.2. However, its reliance on the availability of PE posteriors makes it difficult to employ to identify sub-threshold image counterparts to super-threshold GW events. This is because PE posteriors are computationally expensive and time-consuming to produce and are thus usually unavailable for sub-threshold events.

On the other hand, the matched-filter-based search pipelines give point estimates of the intrinsic parameters and the GW’s network SNR. Furthermore, the chirp mass of a CBC is expected to be the best measured among all intrinsic parameters, certainly for those events whose in-band signal is dominated by the inspiral. We use this to construct a statistic, $\mathcal{B}^{\text{masses}}$, that estimates a “closeness” of the chirp-mass posteriors evaluated using a Fisher analysis.

Additionally, we use the Bayestar [148] sky localization software to rapidly re-construct the localization skymaps. We then cross-correlate them to evaluate another statistic, \mathcal{B}^{sky} , that measures the degree of

overlap between these skymaps.

And finally, we use the \mathcal{R}_{gal} [101, 186] statistic to assess if the time delay of the candidate is more consistent with the distribution of temporal separations of randomly distributed events within the observation time, or the distribution of simulated image time-delays constructed from an assumed distribution of galaxy lenses and sources.

The rest of this subsection describes these three statistics, $\mathcal{B}^{\text{masses}}$, \mathcal{B}^{sky} and \mathcal{R}_{gal} in detail.

Bayestar Skymaps: “Bayestar” is a sky localization software that can produce skymaps in seconds, by exploiting the fact that the intrinsic and extrinsic parameters of a CBC are semi-uncorrelated. It pins the intrinsic parameters to their matched-filter-search point estimates and rapidly marginalizes out the remaining nuisance parameters using Gaussian quadrature methods, to provide a posterior on the right ascension (α) and declination (δ) of the source. [148]. Following Sec. 2.3, we project the Bayestar skymaps to a Cartesian grid of the sky coordinates (α, δ). With this probability density $p(\alpha, \delta | d_i)$ for each of the two images (where d_i is the data pertaining to each of the images), we evaluate the overlap integral in Eq. 2.4, motivated by the PO statistic [101], with $\vec{\theta} = \{\alpha, \delta\}$, assuming isotropic sky priors: $p(\alpha, \delta) \propto \alpha \sin \delta$. This is our first statistic, which we call \mathcal{B}^{sky} . If the two skymaps don’t overlap at all $\mathcal{B}^{\text{sky}} = 0$.

Chirp Masses: The search pipelines report the detector frame matched-filtered chirp mass, $\mathcal{M}_c^{\text{det}}$, and SNR for the triggers based on the best match template at each detector. Following [112] we construct a posterior on the chirp mass for each image, $p(\mathcal{M}_c^{\text{det}} | d_i)$ as a Gaussian with mean as the average of matched-filtered chirp masses over detectors and standard deviation as,

$$\Delta(\log \mathcal{M}_c^{\text{det}}) = 0.08(\rho_{\text{thresh}})/\rho, \quad (3.2)$$

where ρ is the network SNR for the triggers and $\rho_{\text{thresh}} = 8$. Since chirp mass drives the frequency evolution in the inspiral phase, it is one of

the best measured parameters. Assuming it's errors as Gaussian that scale inversely as SNR is a good approximation. The Bhattacharyya coefficient in chirp mass ($\mathcal{B}^{\text{masses}}$) is defined as,

$$\mathcal{B}^{\text{masses}} = \int \sqrt{P(\mathcal{M}_c^{\text{det}}|d_1)P(\mathcal{M}_c^{\text{det}}|d_2)}d\mathcal{M}_c^{\text{det}}, \quad (3.3)$$

which is our second statistic for rapid lensing identification. It quantifies the degree of overlap between the two chirp mass posteriors. Note that unlike the PO statistic (Eq. 2.4), the Bhattacharyya coefficient is a prior-independent measure and is normalized ranging from 0 to 1. It measures the similarity of the two chirp mass posteriors without having to assume any priors. The integral in Eq. 3.3, is simplified for the Gaussian posteriors and can be written in terms of means and standard deviations of the two Gaussians.

In Gaussian noise, the matched-filter SNR is an optimal statistic. However, non-Gaussianities in real noise can produce spuriously large SNRs. Furthermore, even in Gaussian noise, noise fluctuations, and discreteness of the template bank, could result in signals being recovered by templates whose parameters are significantly biased with respect to the true source parameters.

We compare in Fig. 3.6 the estimates of chirp mass from PE with the matched filter estimates, for GWTC-3 events. The PE estimates are taken from GWOSC [195] whereas the matched filter estimates are taken from GraceDB, as reported by the search pipelines [157, 196, 197]. About 13 out of 81 events have $\mathcal{M}_c^{\text{det}}$ from the matched-filter values significantly different (mostly overestimated) with respect to the PE estimates. This is seen in the figure as points sufficiently deviated from the diagonal such that neither the vertical nor the horizontal error bar intersect the diagonal. Most of these events are in the chirp mass range $40 - 60M_\odot$. We do not find any correlation with the SNR for these biases. In addition to the chirp mass, the matched-filter searches also output the best-fit template's mass ratio and spins. This information can in principle be considered for lensing identification however, we leave it for future work.

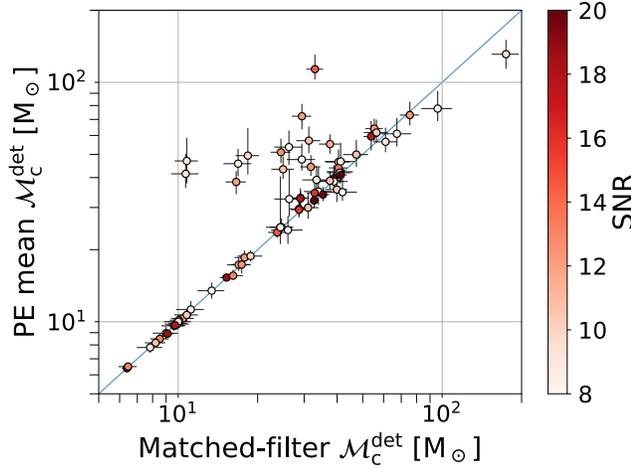


Figure 3.6: Chirp mass estimates from PE and matched-filter-based searches for the GWTC-3 events. About 10 – 15% of events are significantly biased – the mean value of the detector-frame chirp mass lying outside the 90% confidence of the corresponding PE posteriors (vertical error bars) and matched-filter chirp mass (horizontal error bars) coming from Eq. 3.2. If these error bars cross the diagonal, then those events should not be considered as significantly biased. The lower chirp mass binaries are less biased than the high ones due to their longer inspiral in the frequency band of LIGO-Virgo detectors.

The lower chirp mass binaries are less biased than the high ones due to their longer inspiral in the frequency band of LIGO-Virgo detectors. This is a caveat but is partially mitigated by the fact that the Bhattacharyya coefficient is not used in isolation but in conjunction with other statistics in this section which are not susceptible to template biases.

Time-Delay Distributions: The time of arrival of GW signals is measured at $\mathcal{O}(\text{ms})$ precision by matched-filter searches. Therefore we directly use the time delay statistic \mathcal{R}^{gal} to discriminate between lensed and unlensed events, as given by Eq. 3.1, considering only galaxy lenses. In particular, we assume to be the singular isothermal ellipsoid (SIE). The parameters, including redshifts, of the lens, are assumed to follow the SDSS catalog [198]. The BBH mergers are distributed over redshift following Oguri et. al. [93]. The mass spectrum of BBHs is assumed to be the power-law + peak model [44]. We set the detectability criteria to be network $\text{SNR} > 8$.⁷ Fig. 3.7 shows \mathcal{R}^{gal} statistic as a function of the time delay between the events, assuming the observation time of the O3 run. The \mathcal{R}^{gal} statistic favors small time delays and falls off rapidly with increasing time delays. Though this is a model-dependent statistic and valid only for galaxy lenses, it improves our capability of identifying lensed events from unlensed ones [101].

⁷ The sub-threshold events in O3 have $\text{SNR} > 7$ (see fig. 3.9) however, the time delay distributions do not change noticeably with this choice.

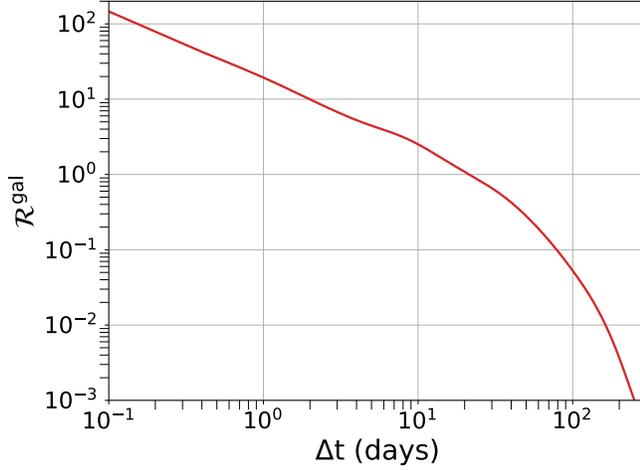


Figure 3.7: \mathcal{R}^{gal} statistic as a function of the time delay between the events, assuming the O3 observation time. This is a model-dependent statistic – the galaxy lens is assumed to be SIE, and the lens parameters are fit to the SDSS catalog. Lens redshifts are also assumed to follow the SDSS catalog, while source masses assume a power-law + peak model.

3.3.2 Assessment of the method with simulations

To assess the performance of our method, we first apply it to simulated lensed and unlensed events. These are injected in Gaussian noise, generated using the zero-detuned high-power PSDs of Advanced LIGO and Advanced Virgo at their design sensitivities [158, 159], as implemented in `pyCBC` [199, 157]. We compare the performance of our method with the PO statistic, using the injection set given in [101] which consists of roughly 300 lensed pairs and half a million unlensed pairs.

To calculate B^{masses} , we assume the posteriors in $\mathcal{M}_c^{\text{det}}$ to be Gaussian with mean as the maximum likelihood estimates from the PE runs (already performed over the mentioned injection set for the results of [101]) and use the SNR of the injection to calculate the standard deviation (see Eq. 3.2). We compare the performance of B^{masses} to the PO statistic $\mathcal{B}^{\text{overlap}}$ evaluated using the posteriors in component mass, setting $\vec{\theta} = \{m_1, m_2\}$ in Eq. 2.4. The maximum likelihood estimates of the masses are a proxy for the matched-filter search estimates – a proxy that is expected to be a good approximation for Gaussian noise. For a templated search involving real noise, apart from the non-Gaussian nature of the latter, we would have an additional source of error in chirp mass estimation incurred due to template-bank discreteness.

To construct B^{sky} , we generate the Bayestar skymaps using the same injection parameters and noise properties as used in the PE runs mentioned above, and fixing the intrinsic parameters to their true values. We compare B^{sky} to the PO statistic B^{overlap} evaluated using the PE posteriors in sky location, setting $\vec{\theta} = \{\alpha, \delta\}$ in Eq. 2.4.

We plot the ROCs graphs in Fig. 3.8 for each statistic. The ROCs display the efficiency (i.e. the fraction of lensed events truly identified as lensed) at a given false positive probability (FPP, the fraction of unlensed events falsely identified as lensed).

An ideal classifier would have an efficiency of 1 for all FPPs. However, during lensing identification, false positives can arise due to chance overlaps of the posteriors of unrelated events. Hence as we increase the threshold of a statistic to identify a pair as lensed, the number of false positives reduces but at the cost of reducing the efficiency. As seen in the figure, both the mass-based and skymap-based statistics produce slightly reduced efficiencies relative to the PO method.

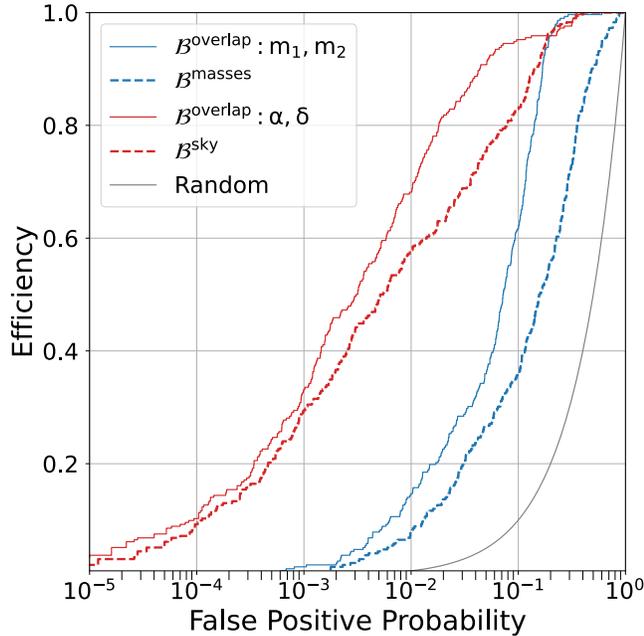


Figure 3.8: Comparison of the rapid identification statistics (dashed) against the PE-based PO method (solid). There is only a partial loss in efficiency using these statistics relative to PO.

3.3.3 *O3 targeted sub-threshold search triggers*

In the case of strongly-lensed gravitational waves, a sub-threshold event with the same intrinsic parameters as the super-threshold event is expected. We consider the GstLAL-based TargetEd Sub-threshold Lensing seArch (TESLA) method [109] to search for sub-threshold lensing counterparts. This is done using the posterior samples from targeted super-threshold events to reduce the background noise. Since the lensed counterpart can be de-magnified and obscured by noise, making it difficult to detect, it is necessary to minimize the impact of background noise. By reducing the amplitude and thus the optimal SNR, we generate these signals and inject them into actual data, which is analyzed using TESLA with a general template bank used in O3. We retain the templates that detect these injected signals and construct a reduced targeted bank. Finally, this targeted bank is used to identify potential lensed candidates for the targeted event from all possible data. For interested readers, please find the detailed description in [109].

As part of the LVK-collaboration-wide paper to search for lensing signatures in the full third observing [15], the TESLA method was applied to all super-threshold events with a probability of astrophysical origin $p_{\text{astro}} > 0.5$ documented in [2] to search for their possible sub-threshold lensed counterparts, should they exist. For each targeted search, we keep all candidates with a FAR < 1 in 30 days (i.e. $< 3.86 \times 10^{-7}$ Hz) and that pass a preliminary skymap overlap test [102] (See [15] for details).

For all the O3 events, altogether 472 possible sub-threshold lensed candidates were found as a deeper internal candidate list ⁸ of which only a small subset is reported in table I [15]. Fig. 3.9 shows the distribution of network SNRs of those superthreshold and subthreshold events. As expected the subthreshold signals have lower SNRs as compared to superthreshold ones. Nevertheless, they all have $\text{SNR} > 7$. Moreover, 40% of the subthreshold events have $\text{SNR} > 8$. This is not necessarily surprising, given that in real noise, SNR is known to

⁸ The candidates having 90% credible region skymap overlap > 0 , taken from table 1 of the data released in [200]

be suboptimal, relative to its performance in Gaussian noise. A more robust statistic, p_{astro} [13, 14], is therefore used to segregate signals of astrophysical and terrestrial origin.

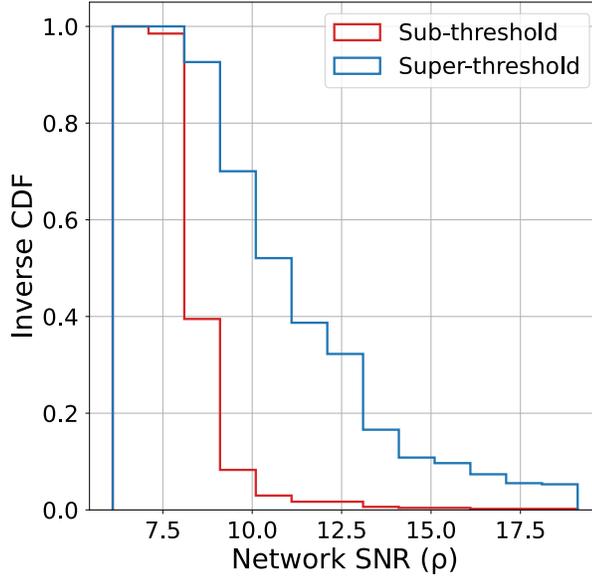


Figure 3.9: Inverse cumulative distribution function of the network SNRs. The sub-threshold triggers have lower SNRs as compared to super-threshold ones but all of them have $\text{SNR} > 7$. In fact, 40% of the subthreshold events have $\text{SNR} > 8$.

3.3.4 Rapid Identification of the super-sub lensed candidates in O_3

During the LVK O_3 lensing searches, [15], only 2 out of all the above-mentioned targeted super-sub lensed pair candidates, pertaining to subthreshold events with the lowest FAR (i.e rate at which noise can falsely trigger a GW-like event) were followed up by a joint-PE analysis as implemented in the GOLUM pipeline [105]. None of those candidates showed any signatures of lensing. Here we take all the 472 search triggers and perform a preliminary analysis to rapidly identify the most interesting super-sub lensed candidates using the matched-filter estimates and the Bayestar skymaps. The chirp mass and skymaps overlap for the GW event pairs is captured in the $\mathcal{B}^{\text{masses}}$ and \mathcal{B}^{sky} respectively, and the time delay prior information is captured in \mathcal{R}^{gal} (see sec. 3.3.1).

To calculate the chirp mass overlap statistic $\mathcal{B}^{\text{masses}}$, for the O_3

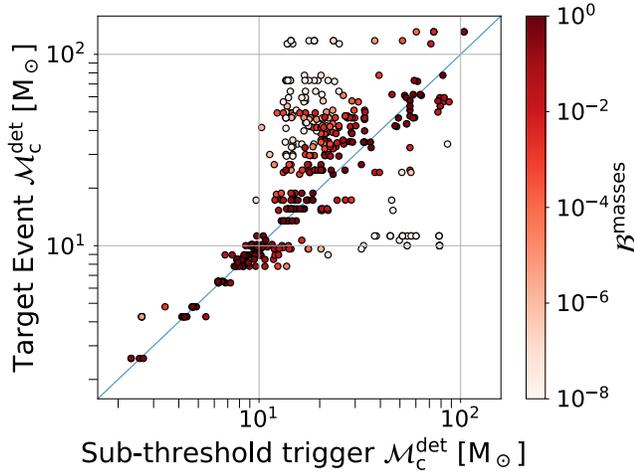


Figure 3.10: Mean chirp mass estimates for the O3 super-sub event pairs and their $\mathcal{B}^{\text{masses}}$. As expected, the majority of events with large $\mathcal{B}^{\text{masses}}$ lie along the diagonal. The coefficient's value decreases for events situated away from the diagonal.

super-sub pairs we use the matched-filter estimates of \mathcal{M}_c and ρ for the sub-threshold triggers, whereas for the super-threshold events we use the PE-based measurements from the open data available in GWOSC [185, 195]. Note that PE and matched-filter chirp mass estimates can deviate significantly for a small fraction of events, as shown in Fig. 3.6.

Assuming a Gaussian distribution of the chirp mass posteriors, using Eqs. 3.2-3.3, we calculate the Bhattacharyya coefficient $\mathcal{B}^{\text{masses}}$ for each of the super-sub pairs. Fig. 3.10 shows the mean estimates of chirp masses for the O3 super-sub pairs and their corresponding $\mathcal{B}^{\text{masses}}$. As expected, the $\mathcal{B}^{\text{masses}}$ is maximum along the diagonal i.e. when the masses of the super-sub pair are similar. Note that $\mathcal{B}^{\text{masses}}$ depends both on the mean and the standard deviation of the two Gaussians, therefore as we move away from the diagonal it falls off but not monotonically.

Next, we calculate the sky overlap statistic \mathcal{B}^{sky} (Eq. 2.4), using the Bayestar skymaps of the sub-threshold events and the more accurate PE skymaps of the super-threshold events. Finally, we estimate the \mathcal{R}^{gal} by using the trigger time information of the events. Fig. 3.12 shows the time delay statistic \mathcal{R}^{gal} , chirp-mass overlap $\mathcal{B}^{\text{masses}}$ and sky-overlap \mathcal{B}^{sky} statistics. for each of the super-sub pairs. The dashed lines correspond to the top 5 percentile statistic values of all

the O₃ pairs, which are found to be at $\mathcal{R}^{\text{gal}} = 2.95$, $\mathcal{B}^{\text{sky}} = 0.47$, and $\mathcal{B}^{\text{masses}} = 0.87$ respectively. These are estimated from the distribution of the statistics for O₃ super-sub pairs as shown in Fig. 3.11. Since lensing is a rare event (rates of strong lensing vary from 0.01% to 1% [96, 97]), we want to find the event pairs which are in the tails of the distribution of the three statistics as the lensing candidates. It should however be noted that the analysis could suffer from spurious biases (see Fig. 3.6) and hence should be followed up by a more comprehensive analysis.

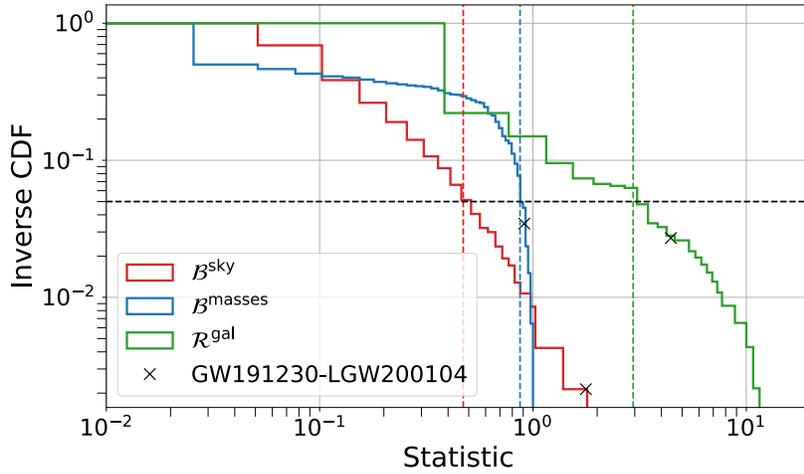
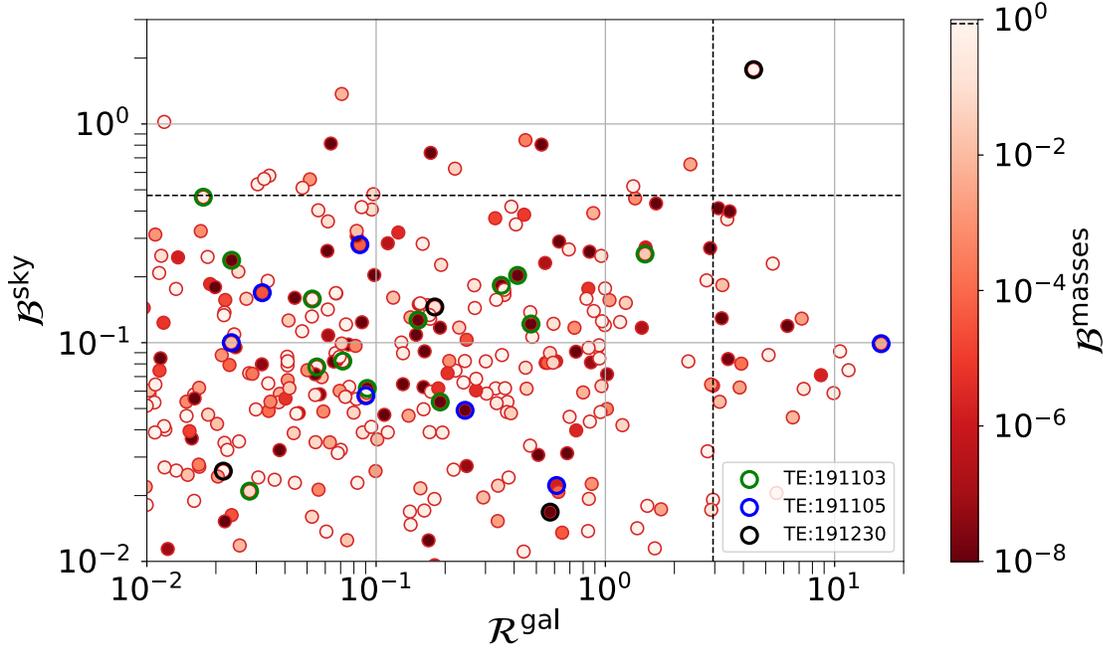


Figure 3.11: Inverse cumulative distribution function of the statistics for the O₃ super-sub pairs. The dashed lines correspond to the top 5 percentile of the estimated values, which is found to be at $\mathcal{R}^{\text{gal}} = 2.95$, $\mathcal{B}^{\text{sky}} = 0.47$, and $\mathcal{B}^{\text{masses}} = 0.87$ respectively. Only one event lies within this percentile, across all three statistics which is shown as a cross-mark (×).

The only pair which is in the top 5 percentile of all three statistics is GW191230-LGW200104, with LGW200104 as the sub-threshold counterpart to the super-threshold target event GW191230. Moreover, this pair has a time delay of only ~ 5 days, which is more consistent with a lensed time delay than the temporal separation of two unrelated events drawn randomly from a Poisson process. The statistics values for this pair is $\mathcal{R}^{\text{gal}} = 4.43$, $\mathcal{B}^{\text{sky}} = 1.77$, and $\mathcal{B}^{\text{masses}} = 0.9$. The good overlap of the skymaps for the pair is evident from Fig. 3.13. We conclude from our analysis that this pair is a (possibly lensed) outlier among the remaining population of (unlensed) O₃ super-sub threshold pairs.

This pair is followed up by PE-based lensing analysis such as PO in sec. 3.3.5 and joint-PE pipelines in the [179]. It turns out that after



incorporating the lensing models and selection effects [106], this pair has the highest significance (i.e. highest probability of being lensed), amongst all the O₃ pairs including the super-super threshold ones with a caveat that LGW200104 might be of terrestrial origin given that during the LVK online unlensed super-threshold searches, the event was found with the SPIIR [201, 202] and cWB [203] pipelines, which reported a p_{astro} of 1% and a FAR of 4824.6/yr. This is in contrast to the FAR of 6.59/yr estimated by TESLA. We report these numbers for the benefit of the reader and do not attempt to draw any conclusions about the nature (noise or signal) of LGW200104. Our method gives a *preliminary* assessment of whether LGW200104 is a lensed counterpart to GW191230, *assuming* the former is a signal, which it well may *not* be. We follow the LVK naming convention of the GW candidates i.e. GWYYMMDD_hhmmss, encoding the date and Coordinated Universal Time (UTC) of the signal [2]. Table 3.1 shows the rest of the triggers found in the searches with the target event GW191230. All of them have at least one of the statistics whose value is low enough to be relegated as unworthy of follow-up.

Figure 3.12: Statistics for the individual super-sub pairs which were found by TESLA during the LVK O₃ lensing searches [15]. Black dashed lines represent the statistic value corresponding to the top 5 percentile of all the pairs. An obvious, solitary, outlier is identified in the top-right corner of the plot. In addition to being in the top 5th percentile of the two statistics on the x and y axes, it is also in the top 5th percentile of the statistic on the color bar (see the dashed black line there). We also highlight the triggers corresponding to the target events GW191230 (green), GW191105 (blue) and GW191103 (black).

Sub-threshold Trigger	$\mathcal{M}_c^{\text{det}} [M_\odot]$	SNR	FAR [yr^{-1}]	Δt [days]	\mathcal{R}^{gal}	\mathcal{B}^{sky}	$\mathcal{B}^{\text{masses}}$
LGW200104_184028	58.20	8.48	6.59	5.02	4.43	1.77	0.91
LGW200301_075426	45.70	7.15	5.61	61.58	0.18	0.15	0.12
LGW200201_192756	20.50	7.80	5.11	33.06	0.57	0.02	0.00
LGW190818_232544	63.00	8.51	3.34	-133.78	0.02	0.03	0.99

Table 3.1: Rapid analysis of the sub-threshold triggers found by TESLA for the target event GW191230_180458,

Special target events: GW191103 and GW191105. As we mentioned previously in Sec. 3.2.2, during the LVK O3 strong lensing searches [15], events pair GW191103-GW191105 was found to be among the more significant ($\sim 1\sigma$) pairs, though unlikely to be lensed. The pair is also analysed in detail in [179] finding no conclusive evidence for lensing. Typically, galaxy lenses can produce more than two images of a GW source. Any additional sub-threshold lensed counterpart to the pair GW191103-GW191105 if found, can help us in increasing its significance of being lensed, in the reconstruction of lens configuration and in localising the source to a host galaxy through cross-matching of electromagnetically observed lensed galaxy catalogues [10, 115]. In the O3 data, 15 potential counterparts for GW191103 and 7 for GW191105 were found by TESLA but none of them is in common for both the targeted events. The rapid statistics for each of them are shown in tables 3.2. LGW191106_200820 was also considered to be a potentially lensed counterpart of GW191105, by virtue of having the highest $\mathcal{R}^{\text{gal}} = 15.95$. However, the poor skymap overlap ($\mathcal{B}^{\text{sky}} = 0.1$ which is worse than about half of all the super-sub pairs, see Fig. 3.11) and even poorer Bhattacharya coefficient value resulted in this event being relegated as unworthy of follow-up. The skymaps for this pair are shown in Fig. 3.13. In summary, we conclude that *no* obvious strongly lensed sub-threshold counterpart to the GW191103-GW191105 pair is found with our preliminary analysis.

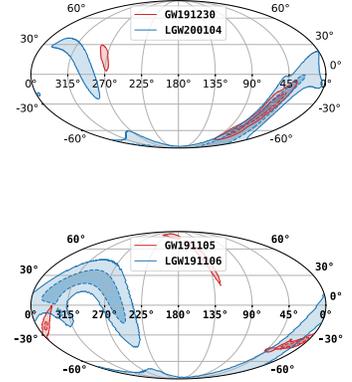


Figure 3.13: Skymaps of the individual events of the pairs (top) GW191230-LGW200104, (bottom) GW191105-LGW191106. The dark (light) shaded region represents the 50 (90) % contour. There is a significant visual overlap between the two skymaps. This is further corroborated by the high skymap overlap statistic value.

3.3.5 PO analysis of GW191230-LGW200104

After performing PE of the subthreshold trigger LGW200104, we find the PO statistic of the pair to be $\log_{10} \mathcal{B}^{\text{overlap}} = 2.45$. Since the com-

Triggers for the target event GW191105_143521 whose $\mathcal{M}_c^{det} [M_\odot] = 9.62$:-

Sub-threshold Trigger	$\mathcal{M}_c^{det} [M_\odot]$	SNR	FAR [yr^{-1}]	Δt [days]	\mathcal{R}^{gal}	\mathcal{B}^{sky}	$\mathcal{B}^{\text{masses}}$
LGW191106_200820	14.40	8.36	0.53	1.23	15.95	0.10	0.00
LGW200128_115458	17.60	7.80	7.57	83.89	0.08	0.28	0.00
LGW191207_050023	20.30	7.87	3.04	31.60	0.61	0.02	0.00
LGW191229_024823	32.20	9.08	12.05	53.51	0.24	0.05	0.00
LGW200303_074125	17.50	8.13	5.87	118.71	0.03	0.17	0.00
LGW200126_135203	11.90	7.72	6.75	81.97	0.09	0.06	0.17
LGW200315_070710	14.10	7.70	10.91	130.69	0.02	0.10	0.01

Triggers for the target event GW191103_012549 whose median $\mathcal{M}_c^{det} [M_\odot] = 10.01$:-

Sub-threshold Trigger	$\mathcal{M}_c^{det} [M_\odot]$	SNR	FAR [yr^{-1}]	Δt [days]	\mathcal{R}^{gal}	\mathcal{B}^{sky}	$\mathcal{B}^{\text{masses}}$
LGW191118_113217	14.00	8.08	6.02	15.42	1.49	0.25	0.02
LGW191213_164018	78.60	10.40	0.32	40.64	0.41	0.20	0.00
LGW190919_131654	33.00	9.15	2.89	-44.51	0.35	0.18	0.00
LGW190926_133040	78.60	7.10	7.57	-37.50	0.47	0.12	0.00
LGW190828_192315	54.30	7.37	6.24	-66.25	0.15	0.13	0.00
LGW200102_033257	78.60	7.18	0.57	60.09	0.19	0.05	0.00
LGW200211_024259	9.44	8.11	5.87	100.05	0.05	0.16	0.81
LGW200323_135352	8.59	8.25	9.02	141.52	0.02	0.46	0.33
LGW190805_134348	10.40	8.79	0.29	-89.49	0.07	0.08	0.87
LGW190813_125024	78.60	7.14	5.61	-81.52	0.09	0.06	0.00
LGW200312_144311	78.60	7.29	8.14	130.55	0.02	0.24	0.00
LGW190727_144658	8.97	8.39	2.76	-98.44	0.06	0.08	0.55
LGW190517_114359	78.60	7.40	0.85	-169.57	0.01	0.28	0.00
LGW190401_190150	54.30	8.31	5.68	-215.27	0.00	0.36	0.00
LGW200305_153119	13.00	8.61	4.92	123.59	0.03	0.02	0.06

Table 3.2: Rapid analysis of the sub-threshold triggers found by TESLA for the target events: GW191103 (top) and GW191105(bottom).

bined SNR of the sub-threshold trigger is 8, it is reasonable to treat the event pairs the same way we did for other candidates. Using the same time delay priors as for the super-threshold events we find $\log_{10} \mathcal{R}^{sal} = 0.97$ which makes the log of the overall PO statistic 3.43. Fig. 3.14 shows the posteriors for LGW200104 and GW191230. Visually, the degree of overlap in both extrinsic and intrinsic parameters is high. However, the intrinsic parameters posteriors are broader as compared to GW191103–GW191105. For events having high masses in the detector frame, such as these, the number of cycles in the waveform within the LIGO–Virgo frequency band is smaller. This leads to broader posteriors which in turn reduce the PO statistic while increasing the rate of coincidental overlaps [112]. In addition, lensed events are more likely to have higher detector frame masses than unlensed events due to their magnification. Hence, it is a challenge to identify high-mass lensed candidates. Including the population priors and selection effects might help [106], as accomplished by the follow-up analyses in [179].

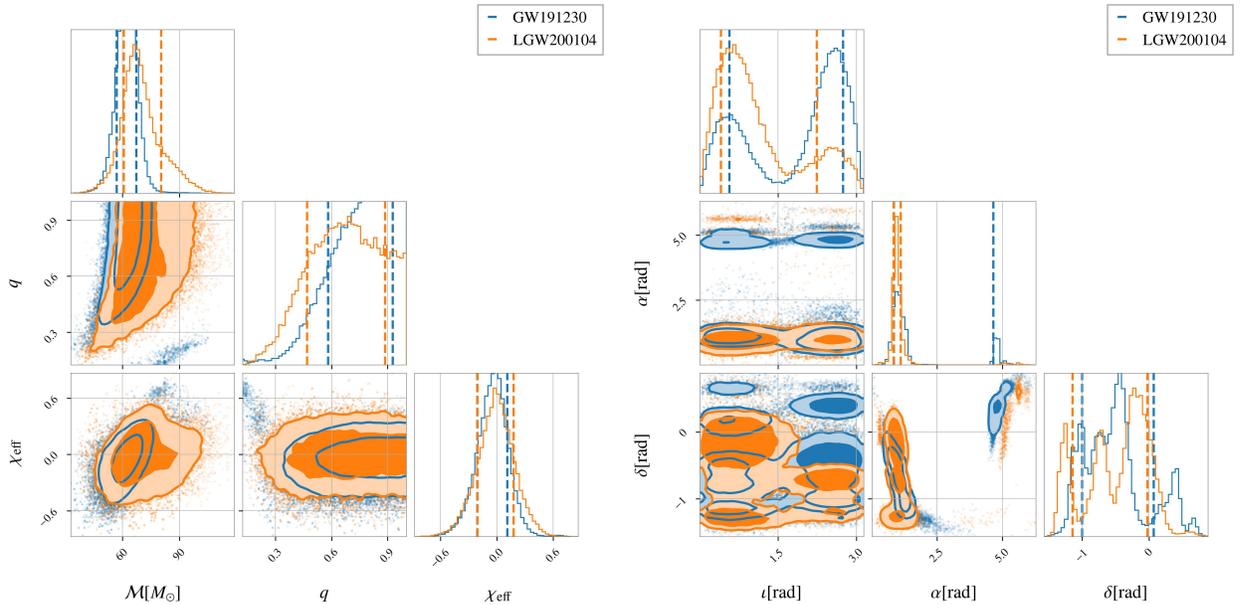


Figure 3.14: Posteriors for GW191230_180458 (blue) and LGW200104_184028 (orange). The posteriors, though broad, have a significant overlap for both the intrinsic (left) and extrinsic (right) parameters.

We also compute its significance using the super-threshold background and find it to be $\approx 1\sigma$, as shown in Fig. 3.3 as a cross symbol

(\times). In the follow-up analysis, described in [179], after including the morse phase information, selection effects and population priors the significance is slightly enhanced. This implies that this pair though not conclusively lensed, is one of the most significant candidates amongst all the O₃ event pairs.

3.4 Summary and Conclusions

We have extended the search for lensing signatures to all BBH candidates with a probability of astrophysical origin higher than 0.5 from O₃b [16], using the two approaches PO and ML, as a part of the LVK collaboration-wide searches for lensing signatures in the O₃ run [15]. For super-threshold/super-threshold pairs, we calculated the degree of overlap (PO statistic) between the posteriors of the intrinsic parameters and sky location, which were obtained from Bayesian inference. Moreover, we analyzed these pairs using a new analysis based on the comparison of spectrograms through ML.

Finally, pairs with FPP from either analysis smaller than 10^{-2} were further studied by conducting full joint Bayesian inference analyses that take population priors and selection effects into account. We found no pairs that show significant evidence for strong lensing [15, 179].

In the LVK O₃ lensing studies, the targetted searches of subthreshold lensed counterparts found a large no. of triggers out of which only the top 2 triggers that were having the lowest search FARs (high probability of being a GW signal) were followed up by the Bayesian model selection method for identifying them as lensed or unlensed. A large number of targeted subthreshold search triggers opens the possibility of one or more of them being lensed (de-magnified) counterparts of confidently detected superthreshold events. Identifying such subthreshold lensed counterparts using conventional methods involving large-scale PE exercises overburdens computational resources. Thus, a preliminary method that rapidly weeds out “obviously” unlensed super-sub candidate pairs is required.

In this Chapter, we have proposed one possible rapid and computationally inexpensive identification scheme. The method is conceptually similar to the PO statistic while relying on approximations to posteriors on chirp masses and sky location pertaining to the subthreshold counterpart. A Bhattacharyya coefficient, $\mathcal{B}^{\text{masses}}$, is constructed from the approximations of the chirp mass posteriors. A skymap overlap statistic, \mathcal{B}^{sky} , is constructed from Bayestar skymaps of subthreshold events and PE skymaps of superthreshold events. The method additionally uses prior information on expected lensed time delays, assuming a lens profile and a SDSS-catalog-fitted model for lens parameter distributions, to further enhance its discriminating abilities. Accordingly, an \mathcal{R}^{gal} statistic is constructed.

Applying the rapid identification method, to the O₃ super-sub pairs, we found that the vast majority of subthreshold events were unrelated to the superthreshold events that were targeted. However, one interesting super-sub candidate pair GW191230-LGW200104 was found to be an outlier – its $\mathcal{B}^{\text{masses, sky}}$ and \mathcal{R}^{gal} statistic values were *all* found to be within the highest 5th percentile of the super-sub candidate pair values analyzed. The follow-up analysis of the pair in [179] by more sophisticated joint-PE analyses, that also include selection effects [106], deemed this event to be among the more significant lensed candidates among all analyzed candidate pairs – super-super and super-sub, assuming a singular isothermal ellipsoid lens profile. While the significance of the event was still not sufficient to claim the detection of a lensed pair in part because LGW200104 had a very low p_{astro} , the identification of this event using our method motivates the need for rapid and computationally inexpensive analyses for future observing runs where the number of candidate pairs is expected to grow drastically.

In future work, we intend to further assess the performance of our method in more realistic simulated data sets. In particular, we plan to inject subthreshold lensed and unlensed events, in addition to superthreshold events, in real noise. We then plan to perform a matched filter search for these events to acquire matched-filter parameter values of chirp masses and arrival times. Using these, and the Bayestar

skymaps, we plan to reconstruct the ROCs to assess the performance in comparison to the PO statistic, to check if the loss in efficiency relative to PO is still acceptable.

We additionally plan to explore the possibility of using PE products from PE algorithms that can rapidly generate posterior samples, such as heterodyning methods [204, 205, 206] and ML based [207, 208, 103] methods, among others [209]. These products will likely be a non-trivial improvement over those that were used in the method presented in this work, although this conjecture needs to be tested, especially for subthreshold events.

Lastly, as we observe more and more events the chances of random coincidences of GW events having similar sky locations and intrinsic parameters will also increase [112]. Therefore, in addition to searching for additional lensed copies, the inclusion of prior information on expected magnifications, time delays, Morse phase differences [186] along with the selection effects and astrophysical population priors may be important for confident detection of strong lensing. To that end, proper modelling of lens and source populations is also required. We expect that as the detector sensitivities improve in upcoming observing runs of LVC, like O4, or when LIGO-India [28] joins the network to reduce the source localization uncertainties, we may soon have the first confident lensed GW detection.

4 | Testing the nature of GW polarisations with strongly-lensed signals

Abstract

GW observations by a network of ground-based laser interferometric detectors allow us to probe the nature of GW polarizations. This would be an interesting test of GR, since GR predicts only two polarization modes while there are theories of gravity that predict up to six polarization modes. The ability of GW observations to probe the nature of polarizations is limited by the available number of linearly independent detectors in the network. (To extract all polarization modes, there should be at least as many detectors as the polarization modes.) Strong gravitational lensing of GWs offers a possibility to significantly increase the effective number of detectors in the network. Due to strong lensing (e.g., by galaxies), multiple copies of the same signal can be observed with time delays of several minutes to weeks. Owing to the rotation of the earth, observation of the multiple copies of the same GW signal would allow the network to measure different combinations of the same polarizations. This effectively multiplies the number of detectors in the network. Focusing on strongly lensed signals from BBH mergers that produce two observable “images”, using Bayesian model selection and assuming simple polarization models, we show that our ability to distinguish between polarization models is significantly improved.

4.1 Introduction

As discussed in Ch. 1, recent observations of GWs [7, 210, 211, 212, 213] have offered new tests of GR in a regime inaccessible by other astronomical observations and laboratory tests [68, 70]. One set of interesting probes includes the nature of GWs themselves [63]. Amongst the various tests that are done using GWs, the measurement of GW polarisations provide yet another opportunity to test the predictions

of GR.

According to GR, GWs have only two independent polarization states — two transverse quadrupole (or, tensor) modes, while a general metric theory of gravity can admit up to six polarization modes (Fig. 4.1). For example, scalar-tensor theories admit two monopole (or, scalar) modes in addition to the tensor modes — massless scalar-tensor theories admit a transverse scalar (or, breathing) mode, while massive scalar-tensor theories admit both transverse and longitudinal scalar modes [63]. More general theories, such as bimetric theories [214], also admit two dipole (or, vector) modes.

GW polarizations can be constrained from the observation of long-lived signals from spinning NSs [215, 216] and stochastic sources [217, 218, 219] as well as from the observation of transient sources such as CBCs [84, 11, 49, 85]. While the detection probabilities of spinning NSs and stochastic background are uncertain, we are expecting the detection of hundreds to thousands of CBC signals in the next few years, using ground-based GW detectors such as LIGO [23], Virgo [24], KAGRA [25] and LIGO-India [27]. Note that each of these quadrupole detectors observes only *one* linear combination of these polarizations. The relative strength of each polarization mode in the observed signal in each detector depends on the response of the detector to the specific polarization. It turns out that the detector response to both the scalar modes (breathing and longitudinal modes) are identical, making them completely degenerate [220]. In summary, even in the ideal case, if we want to disentangle all the five non-degenerate polarization modes from the GW data, we need at least five detectors having different orientations. This would be challenging even when upcoming detectors such as KAGRA and LIGO-India join the international GW network, since the two LIGO detectors in Hanford and Livingston are nearly co-aligned with each other and hence measure largely the *same* linear combination of the polarizations.

Given the data from a network of GW detectors, we can compare the posterior probabilities of different hypotheses, for example, one hypothesis stating that the polarizations are exactly as predicted by

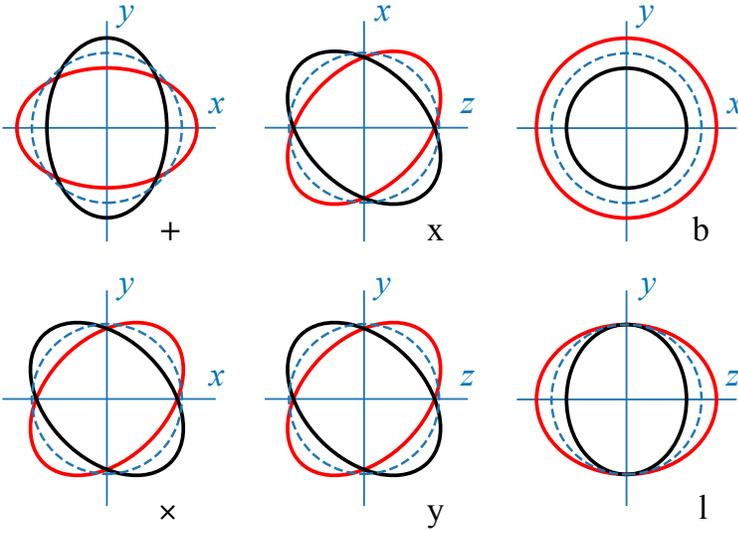


Figure 4.1: The effect of various GW polarizations on a ring of test particles (tensor modes in the left, vector modes in the middle and scalar modes in the right). The wave is always traveling in the z direction. The dashed circles show the original configuration of the test particles before the arrival of the wave and the solid red/black circles and ellipses show the new position of the test particles during the two half cycles of the wave.

GR, while the alternative hypothesis accommodating the presence of additional modes [84]. Motivated by the limited number of linearly independent detectors to observe the polarization modes, the current probes of the nature of GW polarizations have employed highly simplified hypotheses as alternatives to GR. That is, the alternative hypothesis assumes that the polarizations contain only scalar modes or only vector modes (no tensor modes). The analyses of some of the CBC observations by LVK detectors have concluded that the tensor-only hypothesis is preferred over scalar-only or vector-only hypotheses [85, 70, 49].

As discussed in Ch. 1, GWs can be gravitationally lensed by intervening matter distributions, such as galaxies and galaxy clusters. Recent estimates suggest that a small fraction ($\sim 0.1\% - 0.5\%$) of the BBH signals that we expect to detect using the LIGO-Virgo network will be strongly lensed by intervening galaxies [180, 98], producing multiple “images” of the signals ¹. These signals arrive at the detector with relative time delays ranging from several minutes to several weeks [180]. Lensing by galaxies or clusters is very well approximated by geometric optics since the mass scale of the lens is significantly larger than the wavelength of GWs ($GM_{\text{lens}}/c^2 \gg \lambda_{\text{GW}}$). Thus, multiple images would correspond to copies of the same signal with a

¹ Galaxy clusters will also cause strong lensing of the GW signals. However, the lensing probability due to clusters is likely to be significantly smaller than that by galaxies [221]. However, cluster lensing is expected to produce longer time delays between images, which will further improve our ability for polarization model selection from lensed GW signals.

relative magnification and time delay. Due to the rotation of the earth, observed signals from multiple images will involve different combinations of the same polarizations. As far as the polarization content is concerned, this is equivalent to observing the same signal with a multiplied number of detectors. For example, if two images of the merger are observed using a three-detector network, this is equivalent to observing the one merger signal with a six-detector network.

In this chapter, we explore the possibility of constraining the polarization content of GWs using BBH mergers that are strongly lensed by galaxies. We use the Bayesian model selection method proposed by [84] to identify the polarization content of simulated GW signals from BBHs. As In [84, 85, 70, 49], we make the simplified assumptions that polarizations contain only pure scalar, vector, or tensor modes. Under this assumption, we show that strongly lensed GW signals will enable us to constrain the polarization content significantly better than their unlensed counterparts.

Although alternative theories to GR also predict gravitational lensing, the resulting phenomenology is only beginning to be explored (see, e.g., [222, 12]). In this work, we assume that lensing in alternative theories does not significantly alter the polarization content of the GWs — to the very least, we require that in the geometric optics regime, multiple lensed images should contain the same combination of polarizations, except for an overall magnification; see, e.g., [12] for caveats.

The rest of the chapter is organized as follows: Section 4.3 summarizes our methodology, providing a brief introduction to the relevant theory, model selection formalism, as well as the details of the numerical simulations. Section 4.4 presents the results, while some concluding remarks and future work is discussed in Sec. 4.5.

4.2 GW polarizations

In the local Lorenz gauge, the spatial components of the metric perturbation h_{ij} at a given space-time point \vec{x} can be written in terms of six linearly independent polarization tensors, e^A ²

² In this Section, we denote four-vectors by the use of arrows (e.g., \vec{x}) and three-vectors by boldface (e.g., \mathbf{x}) and tensors by sans serif fonts (e.g., e). Repeated indices are assumed to be summed over.

$$h_{ij}(\vec{x}) = h_A(\vec{x}) \mathbf{e}_{ij}^A, \quad A \in \{+, \times, x, y, b, l\} \quad (4.1)$$

where, the index A stands for different polarizations: tensor “plus” (+) and “cross” (\times) modes, vector “x” and “y” modes and scalar “breathing” (b) and “longitudinal” (l) modes; and h_A is the amplitude for polarization A . The existence of six independent polarization modes (or, six linearly independent components of the metric perturbation) can be understood in the following way: the full metric perturbation $h_{\mu\nu}$ in four dimensions is symmetric and therefore has ten independent components. However, because of the Lorenz gauge condition, four degrees of freedom are taken away, leaving only six. (GR, in addition, satisfies the transverse-traceless gauge condition which takes away additional 4 degrees of freedom and hence allowing only two tensor polarization modes).

Further, the polarization tensors can be written in terms of the orthogonal basis vectors \mathbf{w}_x , \mathbf{w}_y , $\mathbf{w}_z \equiv \mathbf{w}_x \times \mathbf{w}_y$, where \mathbf{w}_z is the GW propagation direction [84, 219].

$$\begin{aligned} \mathbf{e}^+ &= \mathbf{w}_x \otimes \mathbf{w}_x - \mathbf{w}_y \otimes \mathbf{w}_y \\ \mathbf{e}^\times &= \mathbf{w}_x \otimes \mathbf{w}_y + \mathbf{w}_y \otimes \mathbf{w}_x \\ \mathbf{e}^x &= \mathbf{w}_x \otimes \mathbf{w}_z + \mathbf{w}_z \otimes \mathbf{w}_x \\ \mathbf{e}^y &= \mathbf{w}_y \otimes \mathbf{w}_z + \mathbf{w}_z \otimes \mathbf{w}_y \\ \mathbf{e}^b &= \mathbf{w}_x \otimes \mathbf{w}_x + \mathbf{w}_y \otimes \mathbf{w}_y \\ \mathbf{e}^l &= \mathbf{w}_z \otimes \mathbf{w}_z \end{aligned} \quad (4.2)$$

A ground-based laser interferometric detector measures a combination of these polarizations by the change in lengths of its perpendicular arms. This response is encoded in the detector tensor D , whose components are given by

$$D^{ij} = \frac{1}{2} \left(d_x^i d_x^j - d_y^i d_y^j \right) \quad (4.3)$$

where \mathbf{d}_x and \mathbf{d}_y are unit vectors along the detector arms, with a com-

mon origin. The strain, h_I measured by detector I , is then given as,

$$h_I(t) = h_{ij}(t, \mathbf{x}_I) D_I^{ij} = h_A(t, \mathbf{x}_I) e_{ij}^A D_I^{ij} = h_A(t, \mathbf{x}_I) F_I^A \quad (4.4)$$

where, $F_I^A \equiv D_I^{ij} e_{ij}^A$, are called the detector antenna pattern functions, which encode the response of the detector I to polarization A . Therefore, GW strain measured at the detector I can be written as the linear combination of polarization amplitudes multiplied with the corresponding antenna pattern functions. Expanding, F_A using Eqs. (4.2) and (4.3):

$$\begin{aligned} F_+ &= \frac{1}{2} \left[(\mathbf{w}_x \cdot \mathbf{d}_x)^2 - (\mathbf{w}_x \cdot \mathbf{d}_y)^2 - (\mathbf{w}_y \cdot \mathbf{d}_x)^2 + (\mathbf{w}_y \cdot \mathbf{d}_y)^2 \right] \\ F_\times &= (\mathbf{w}_x \cdot \mathbf{d}_x) (\mathbf{w}_y \cdot \mathbf{d}_x) - (\mathbf{w}_x \cdot \mathbf{d}_y) (\mathbf{w}_y \cdot \mathbf{d}_y) \\ F_x &= (\mathbf{w}_x \cdot \mathbf{d}_x) (\mathbf{w}_z \cdot \mathbf{d}_x) - (\mathbf{w}_x \cdot \mathbf{d}_y) (\mathbf{w}_z \cdot \mathbf{d}_y) \\ F_y &= (\mathbf{w}_y \cdot \mathbf{d}_x) (\mathbf{w}_z \cdot \mathbf{d}_x) - (\mathbf{w}_y \cdot \mathbf{d}_y) (\mathbf{w}_z \cdot \mathbf{d}_y) \\ F_b &= \frac{1}{2} \left[(\mathbf{w}_x \cdot \mathbf{d}_x)^2 - (\mathbf{w}_x \cdot \mathbf{d}_y)^2 + (\mathbf{w}_y \cdot \mathbf{d}_x)^2 - (\mathbf{w}_y \cdot \mathbf{d}_y)^2 \right] \\ F_l &= \frac{1}{2} \left[(\mathbf{w}_z \cdot \mathbf{d}_x)^2 - (\mathbf{w}_z \cdot \mathbf{d}_y)^2 \right] \end{aligned} \quad (4.5)$$

Evaluating these antenna pattern functions at a particular detector involves projecting the polarization tensors into the detector frame. This projection depends on the direction from which GW arrives for a particular detector and hence the sky-location of the GW source. We choose to describe the source using the equatorial coordinate system, in terms of right ascension α and declination δ . Additionally, antenna pattern functions also depend on polarization angle, ψ , which is due to the rotational freedom of the orthonormal vectors $(\mathbf{w}_x, \mathbf{w}_y)$ about the propagation direction \mathbf{w}_z . Further, the detector response and hence the antenna pattern functions also depend on time due to the rotation of the earth. Thus, the antenna pattern functions of the detector I for the polarization mode A can, in general, be written as $F_I^A(\alpha, \delta, \psi, t)$.

For three detector network, we have three strain measurements giving three different linear combinations of polarizations, as given by Eq.(4.4). However, due to strong lensing, multiple copies of the same GW signal would arrive at each detector with time delay Δt of minutes

to weeks. Due to the rotation of the earth, the antenna patterns during the arrival of, say two images, $F_I^A(\alpha, \delta, \psi, t)$ and $F_I^A(\alpha, \delta, \psi, t + \Delta t)$ can be considerably different from each other. This is equivalent to observing one signal with a six-detector network.

According to GR, in the geometrical optics limit, polarization tensors are parallelly propagated along the null geodesics, implying that lensing does not change the polarization content of a GW [223]. We assume this to be true in alternative theories also. As long as the metric perturbation follows a source-free wave equation, the polarization tensor should be conserved along the GW propagation.

4.3 Method

4.3.1 Model selection of polarizations

Bayesian model selection allows us to assign posterior probabilities for various hypotheses pertaining to the observed data. We formulate the polarization content of GWs as different Bayesian hypotheses. For e.g., GR is denoted as \mathcal{H}_t , as the theory only predicts tensor modes. The hypothesis that GWs contain only scalar (vector) modes is denoted as \mathcal{H}_s (\mathcal{H}_v). Following [215], we assume that the waveforms in \mathcal{H}_s and \mathcal{H}_v are the same as in \mathcal{H}_t ; the only change is in the antenna pattern functions. We do this as the BBH waveforms for alternative theories are presently not known. If available in the future they can be included in the same formalism.

Given the set of data $\{d\}$ from a network of detectors, the marginalized likelihood (or, Bayesian evidence) of the hypothesis \mathcal{H}_p can be computed by

$$P(\{d\}|\mathcal{H}_p) = \int d\theta P(\theta) P(\{d\}|\theta, \mathcal{H}_p), \quad (4.6)$$

where θ is a set of parameters that describe the signal under hypothesis \mathcal{H}_p (including the masses and spins of the compact objects in the binary, location and orientation of the binary and the arrival time and phase of the signal), $P(\theta)$ is the prior distribution of θ (which we take

to be independent of \mathcal{H}_p), and $P(\{d\}|\boldsymbol{\theta}, \mathcal{H}_p)$ is the likelihood of the data $\{d\}$, given the parameter vector $\boldsymbol{\theta}$ and hypothesis \mathcal{H}_p . Given the hypothesis \mathcal{H}_p and data $\{d\}$, we can sample and marginalize the likelihood over the parameter space using an appropriate stochastic sampling technique such as Nested Sampling [224].

Bayesian model selection allows us to compare multiple hypotheses. For e.g., the odds ratio \mathcal{O}_s^t is the ratio of the posterior probabilities of the two hypotheses \mathcal{H}_t and \mathcal{H}_s . When \mathcal{O}_s^t is greater than one then hypothesis \mathcal{H}_t is preferred over \mathcal{H}_s and vice versa. Using Bayes theorem, the odds ratio can also be written as the product of the ratio of the prior odds \mathcal{P}_s^t of the hypotheses and the likelihood ratio, or Bayes factor \mathcal{B}_s^t :

$$\mathcal{O}_s^t := \frac{P(\mathcal{H}_t|\{d\})}{P(\mathcal{H}_s|\{d\})} = \frac{P(\mathcal{H}_t)}{P(\mathcal{H}_s)} \times \frac{P(\{d\}|\mathcal{H}_t)}{P(\{d\}|\mathcal{H}_s)} = \mathcal{P}_s^t \times \mathcal{B}_s^t \quad (4.7)$$

Since GR has been tested well in a variety of settings, our prior odds are going to be highly biased towards tensor-only modes, i.e., $\mathcal{P}_s^t \gg 1$. Hence, in order to claim evidence of non-tensor modes the corresponding Bayes factor supporting the alternative hypothesis has to be very large. Since the Bayes factor is the only quantity that is derived from data, for the rest of the chapter, we focus on the Bayes factor. The Bayes factor from multiple, uncorrelated events $d^{(i)}$ can be combined as

$$\mathcal{B}_s^t = \prod_i \mathcal{B}_s^{t(i)}, \quad (4.8)$$

where

$$\mathcal{B}_s^{t(i)} = \frac{P(\{d^{(i)}\}|\mathcal{H}_t)}{P(\{d^{(i)}\}|\mathcal{H}_s)} \quad (4.9)$$

is the Bayes factor obtained from the i th event.

4.3.2 Model selection of polarizations using lensed GW events

When multiple GW events are produced by the strong lensing of a signal BBH merger, these events cannot be treated as uncorrelated events. Here we derive the Bayes factor between different polarization hypotheses \mathcal{H}_p using multiple lensed images of the same merger.

For simplicity, we will consider only two lensed images. However, the same formalism can be extended to more than two images also. Strong lensing of GWs from BBHs is expected to be dominated by galaxy lenses. Lensing by galaxies and galaxy clusters can be treated in geometric optics regime (wavelength of GWs significantly smaller than the mass scale of the lens). In this regime, lensing does not change the frequency profile of the waveform, and hence multiple images, arriving at the detector at different times, differ from each other only by a relative magnification and a constant phase shift³ [225, 226]. Hence the parameters describing the waveform, except for the luminosity distance (which is degenerate with the lensing magnification) and the time and phase at coalescence will be common between the two images.

Now consider the GW signals $d^{(1)}$ and $d^{(2)}$ produced by the strong lensing of a BBH merger. The Bayes factor between two polarization hypotheses \mathcal{H}_t and \mathcal{H}_s can be written as

$$\mathcal{B}_s^t := \frac{P(\{d^{(1)}, d^{(2)}\}|\mathcal{H}_t)}{P(\{d^{(1)}, d^{(2)}\}|\mathcal{H}_s)} \quad (4.10)$$

$$= \frac{\int d\theta_c P(\theta_c) P(d^{(1)}|\theta_c, \mathcal{H}_t) P(d^{(2)}|\theta_c, \mathcal{H}_t)}{\int d\theta_c P(\theta_c) P(d^{(1)}|\theta_c, \mathcal{H}_s) P(d^{(2)}|\theta_c, \mathcal{H}_s)}, \quad (4.11)$$

where θ_c is the vector of common parameters as the signals come from the same merger. Note that the probability distributions are marginalized over all the parameters except θ_c . Using the Bayes theorem, the likelihoods $P(\{d^{(i)}\}|\theta_c, \mathcal{H}_p)$ can be written in terms of the posteriors $P(\theta_c|\{d^{(i)}\}, \mathcal{H}_p)$ as

$$P(\{d^{(i)}\}|\theta_c, \mathcal{H}_p) = \frac{P(\theta_c|\{d^{(i)}\}, \mathcal{H}_p) P(\{d^{(i)}\}|\mathcal{H}_p)}{P(\theta_c)}, \quad (4.12)$$

where $P(\{d^{(i)}\}|\mathcal{H}_p)$ is the marginal likelihood of the hypothesis \mathcal{H}_p defined in Eq.(4.6). Using this, Eq.(4.11) can be rewritten as

$$\mathcal{B}_s^t = \mathcal{B}_s^{t(1)} \mathcal{B}_s^{t(2)} \frac{\mathcal{B}_U^L|\mathcal{H}_t}{\mathcal{B}_U^L|\mathcal{H}_s}, \quad (4.13)$$

where $\mathcal{B}_s^{t(1)}$ and $\mathcal{B}_s^{t(2)}$ are the Bayes factors of the polarization hypotheses obtained from event 1 and 2, respectively [see Eq.(4.9)], while \mathcal{B}_U^L

³Note that this is valid only for the quadrupole mode GW signals from non-precessing binaries, which is the case we are considering in this chapter.

is the lensing Bayes factor, defined as the PO statistic [180] (see Ch. 2.2). That is,

$$\mathcal{B}_{\mathcal{U}}^L|\mathcal{H}_p = \int \frac{d\boldsymbol{\theta}_c P(\boldsymbol{\theta}_c|d^{(1)}, \mathcal{H}_p) P(\boldsymbol{\theta}_c|d^{(2)}, \mathcal{H}_p)}{P(\boldsymbol{\theta}_c)}. \quad (4.14)$$

This is the Bayes factor between a different set of two hypotheses \mathcal{H}_L and \mathcal{H}_U , which are different from the hypotheses on polarization content. The lensing hypothesis \mathcal{H}_L states that this pair of events are lensed images of the same merger, while the unlensed hypothesis \mathcal{H}_U states that these are unrelated events. It can be seen from Eq.(4.14) that $\mathcal{B}_{\mathcal{U}}^L|\mathcal{H}_p$ is the prior-weighted inner product of the posteriors of the common parameters $\boldsymbol{\theta}_c$ obtained from the two images. These posteriors are computed assuming the polarization hypothesis \mathcal{H}_p . Note that these posteriors, and hence the lensing Bayes factor can be computed assuming different hypotheses for the polarization content $\mathcal{H}_p = \{\mathcal{H}_t, \mathcal{H}_s, \mathcal{H}_v\}$.

From Eq.(4.13) it is evident that, given a pair of lensed events, the combined Bayes factor between the two polarization hypotheses is the product of the Bayes factors computed from the individual events multiplied by an extra term $\frac{\mathcal{B}_{\mathcal{U}}^L|\mathcal{H}_t}{\mathcal{B}_{\mathcal{U}}^L|\mathcal{H}_s}$, which we call the *PO ratio*. We can do parameter estimation for the individual events assuming different polarization hypotheses \mathcal{H}_p to get the posteriors and marginal likelihoods (evidence) from each event. Later, from these posteriors, we compute the overlap factors $\mathcal{B}_{\mathcal{U}}^L$ of the two events. If the PO statistic using the correct polarization hypothesis, say \mathcal{H}_t , is larger than the same using the wrong hypothesis, say \mathcal{H}_s , (that is, if $\mathcal{B}_{\mathcal{U}}^L|\mathcal{H}_t > \mathcal{B}_{\mathcal{U}}^L|\mathcal{H}_s$), then the combined Bayes factor \mathcal{B}_S^t of the two lensed events will be larger than the same computed from two unlensed events with the same individual Bayes factors $\mathcal{B}_S^{t(1)}$ and $\mathcal{B}_S^{t(2)}$. This suggests that lensed events can improve our ability to identify the right polarization hypothesis, with an improvement factor given by the PO ratio $\frac{\mathcal{B}_{\mathcal{U}}^L|\mathcal{H}_t}{\mathcal{B}_{\mathcal{U}}^L|\mathcal{H}_s}$.

4.3.3 Simulations

Since there are no strong lensed events detected until now by the current detectors [108, 227]⁴, we use a simulated catalog of lensed BBH events presented in [180] to study the efficacy of polarization recovery. Simulations in [180] generated the observed parameters for different pairs of lensed BBH merger signals, assuming a well-motivated distribution of lens and source properties [180]. In the geometric optics regime, for a particular event, intrinsic parameters like the BHs' redshifted masses (m_1 and m_2) and spins as well as the extrinsic parameters like the sky-location (α and δ), orientation (ι and ψ) and the coalescence (orbital) phase ϕ_0 of the binary remain the same for multiple images⁵.

However, the strain amplitude will be magnified due to lensing, and hence the luminosity distance (d_L) estimated from the two images will be different. The observed time of coalescence (t_0) of the two signals will also be different due to the time delay between the arrival of the two signals (ranging from several minutes to several weeks). Figure 4.2 shows the optimal SNR of the two lensed images from the simulated GW events plotted against the time delay between them (left, middle and right plots correspond to tensor, vector and scalar polarizations, respectively). Highly magnified images correspond to shorter time delays, and will, typically, have larger SNRs. In contrast, images with much larger time delays will have lower magnifications, and hence will be typically undetected. When the source is close to but inside a caustic (infinite magnification curve) the images are highly magnified, but has small angular separation and time delay. On average, these magnified signals are detected with higher SNR. Therefore, the smaller time delay images correspond to high image magnifications and are more likely to be detectable⁶. As a result, most of the detected events have time delays less than a day.

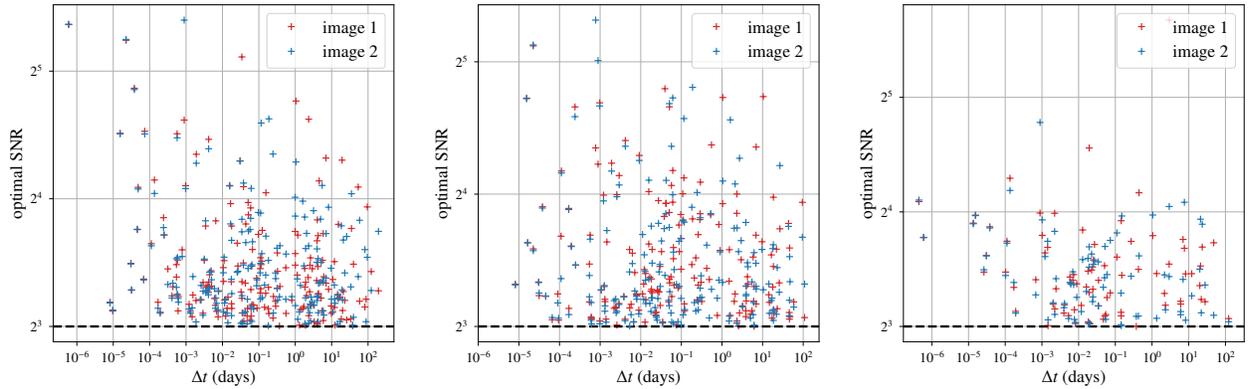
Except d_L and t_0 , all the other parameters can contribute to our common parameters vector θ_c . We limit ourselves to non-spinning binaries and perform Bayesian parameter estimation of the following parame-

⁴ We note that, during the late states of the preparation of this draft, [110] has identified an unusual lensing candidate in the data of the second observing run of LVK detectors. Additionally, [189] has identified possible sub-threshold candidates of strongly lensed images of GW150914 by a targeted search using LIGO's first observing run data.

⁵ Depending on the type of the image, the lensed waveform will contain an additional topological phase shift [225, 91]. However, it is straightforward to model this and to estimate the source parameters, which will be consistent between multiple images.

⁶ Kindly refer to Eqns. 4.47-4.48 of [228] and the discussion there for more details.

ters: $m_1, m_2, \alpha, \delta, d_L, t_0, \phi_0, \iota, \psi$. However, in order to compute the PO ratio, we consider only the following parameters: $m_1, m_2, \alpha, \delta, \iota$. That is, the posteriors are marginalized over all other parameters. This is based on our empirical observation that this choice of parameters typically provide the largest values of the PO ratios⁷. From the simulated



⁷Note that there might potentially be a more optimal choice for this set of parameters, depending on the model as-

parameters of the lensed events, we generate GW signals at different detectors for each polarization hypothesis – tensor \mathcal{H}_t , vector \mathcal{H}_v , and scalar \mathcal{H}_s , using the corresponding antenna pattern functions (see Sec. 4.2). The model signals are generated using the antenna patterns of the corresponding polarizations, but always assuming that the time evolution of the simulated waveform always follow the GR waveform. That is, $h_s(t) = h_\times(t) = h_+(t)$ and $h_1(t) = h_y(t) = h_\times(t)$.

For these simulations, we consider a three detector network consisting of two US-based Advanced LIGO detectors located in Hanford, WA and Livingston, LA and the Advanced Virgo detector located in Pisa, Italy. The LIGO detectors were assumed to have their design sensitivity with the power spectral density (PSD) given in [159] while the Virgo detector was assumed to have the PSD given in [229]. GW signals were simulated using the IMRPHENOMPv2 waveform approximant [230, 231, 232] coded in the LALSIMULATION module of the LALSUITE software package [156]. We select $\sim 100 - 200$ injections crossing a threshold of 8 for the network SNR. Once we have the injections, we use the Dynamic Nested Sampling [233] implementation

Figure 4.2: Correlation of optimal SNRs of the two images with the lensing time delay Δt between the images for each set of injections, created using the tensor (left), vector (middle) and scalar (right) polarization models. Highly magnified images correspond to shorter time delays, and will, typically, have larger SNRs. In contrast, images with much larger time delays will have low magnifications, and hence will be typically undetected. As a result, most of the detected events have time delays less than a day.

(DYNESTY) in BILBY package [234] to compute the posteriors of the binary parameters and the evidences of each polarization hypothesis \mathcal{H}_p . We have three simulated (injection) models $\mathcal{H}_t^{[I]}, \mathcal{H}_s^{[I]}, \mathcal{H}_v^{[I]}$ and three recovery models $\mathcal{H}_t^{[R]}, \mathcal{H}_s^{[R]}, \mathcal{H}_v^{[R]}$, allowing us to analyze the nine combinations of injection sets: $\mathcal{H}_t^{[I]} - \mathcal{H}_t^{[R]}, \mathcal{H}_t^{[I]} - \mathcal{H}_s^{[R]}, \mathcal{H}_t^{[I]} - \mathcal{H}_v^{[R]}, \mathcal{H}_s^{[I]} - \mathcal{H}_t^{[R]}, \mathcal{H}_s^{[I]} - \mathcal{H}_s^{[R]}, \mathcal{H}_s^{[I]} - \mathcal{H}_v^{[R]}, \mathcal{H}_v^{[I]} - \mathcal{H}_t^{[R]}, \mathcal{H}_v^{[I]} - \mathcal{H}_s^{[R]}$ and $\mathcal{H}_v^{[I]} - \mathcal{H}_v^{[R]}$.

We use the standard Gaussian likelihood model for estimating the posteriors of the parameters under different polarization hypotheses (see, e.g., [235]). For simplicity, no noise is added to the simulated signals. Further, we considered only non-spinning binaries. Thus, the likelihood is computed over the following parameters $m_1, m_2, \alpha, \delta, d_L, \iota, \psi, \phi_0, t_0$. We use uniform priors in redshifted component masses of the binary ($m_1, m_2 \in [3, 500]M_\odot$), isotropic sky location (uniform in $\alpha, \sin \delta$) and orientation (uniform in $\cos \iota, \phi_0$), uniform in polarization angle ψ , and a volumetric prior $\propto d_L^2$ on luminosity distance. Finally, the posteriors are marginalized over all the parameters except the ones that we consider for calculating the PO statistic, i.e., $m_1, m_2, \alpha, \delta, \iota$. Note that the parameter estimation is performed on each individual image, without imposing the constraint that multiple images share the same parameters.

As one would anticipate, the true (injected) parameters are recovered when the injection and recovery model are the same. As an example, Fig. 4.3 shows the estimated posterior distributions when the injections and recovery are performed using the same tensor hypothesis (i.e., the $\mathcal{H}_t^{[I]} - \mathcal{H}_t^{[R]}$ combination). In contrast, when parameter estimation on the tensor injection is performed using vector and scalar hypotheses, the intrinsic parameters (primary and secondary masses) are still recovered well, whereas extrinsic parameters such as the sky location and orientation are not recovered well (see Figs. 4.4 and 4.5). This is due to the fact that the recovery of the extrinsic parameters heavily depends on the antenna pattern functions, which are different for the injection and recovery models.

Further, note that sky location posteriors ($\alpha, \sin \delta$) of the lensed pairs

overlap well with the tensor model (Fig. 4.3) and not so well with vector and scalar models (Figs. 4.4 and 4.5 show the $\mathcal{H}_t^{[I]} - \mathcal{H}_v^{[R]}$ and $\mathcal{H}_t^{[I]} - \mathcal{H}_s^{[R]}$ combinations, respectively). As a result, in general, we would expect that the lensing Bayes factor \mathcal{B}_U^L will be larger for the tensor model. This is quantified in the next section.

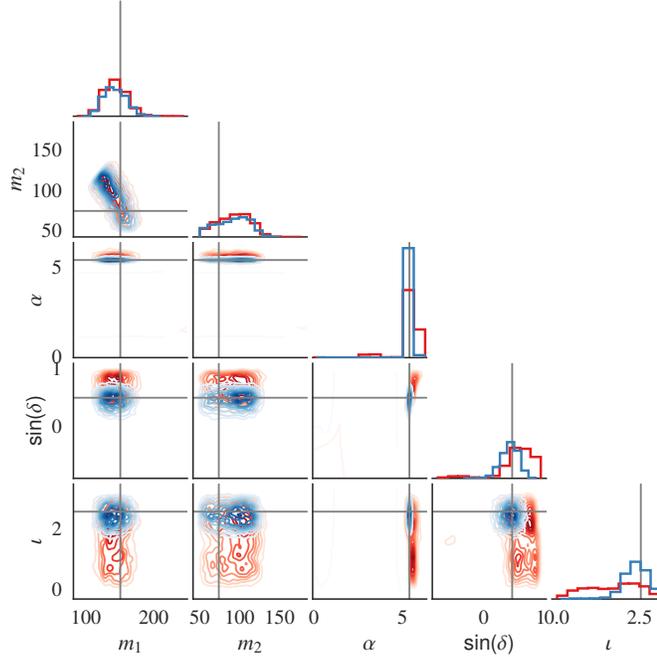


Figure 4.3: Marginalized posteriors of the parameters $m_1, m_2, \alpha, \sin \delta, l$ estimated from a lensed pair of tensor injections with tensor recovery (i.e., $\mathcal{H}_t^{[I]} - \mathcal{H}_t^{[R]}$ case). Gray lines show the injected values. Note that the posteriors estimated from the two images are overlapping and are consistent with the injected values. The images have comparable optimal SNR (8.2 for red and 9 for blue), but have a time delay of 3.31 between them. Since the sky location of the two images (with respect to the detectors) is different, the posteriors of the extrinsic parameters have some difference between them.

4.4 Results

Our aim is to quantify how well pairs of GW signals produced by the strong lensing of BBH mergers improve our ability to distinguish between polarization models, as compared to pairs of unrelated signals with similar strengths. If the two signals are unrelated (i.e., unlensed events) then the combined Bayes factor will just be the product of individual Bayes factors [Eq.(4.8)]. On the other hand, if the two events are lensed images of the same merger, then the combined Bayes factor is the product of the individual Bayes factors and additional factor, $\frac{\mathcal{B}_U^L | \mathcal{H}_{\text{right}}}{\mathcal{B}_U^L | \mathcal{H}_{\text{wrong}}}$, which we call the PO ratio [Eq.(4.13)]. If the PO ratio is greater than one then for the correct polarization hypothesis, this

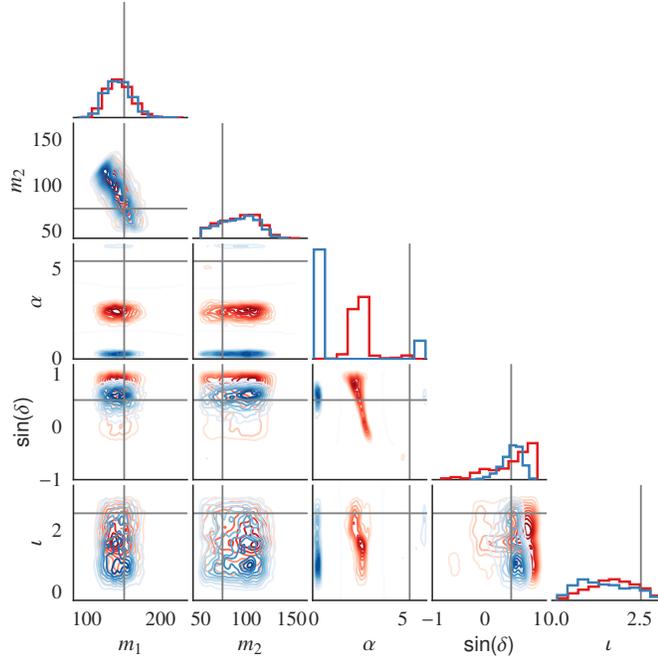


Figure 4.4: Same as Fig. 4.3, except that the injection is performed using the tensor polarization model while parameter estimation is performed using the vector model (i.e., $\mathcal{H}_t^{[I]} - \mathcal{H}_v^{[R]}$ case). Note that the posteriors of the extrinsic parameters $\alpha, \sin \delta, l$, estimated from the two images, are not always overlapping and are not always consistent with the injected values.

would show that lensing improves our ability to identify the right polarization hypothesis.

Figure 4.6 (top panels) shows the distribution of the polarization Bayes factors from the simulated GW signals. The fact that Bayes factors are almost always greater than 1 (log Bayes factors > 0) suggests that the right polarization hypothesis is almost always preferred. Note that, overall, the lensed Bayes factors are greater than unlensed ones, showing that the strong lensing improves the polarization models selection. This is also evident from the distribution of the PO ratios (bottom panel of Fig. 4.6). Note that the overlap ratios are greater than 1 (log overlap ratio > 0) for $\sim 85 - 95\%$ of the simulated events. The median value of the overlap ratio is $\sim 2 - 3$, which means that for more than 50% of the events lensing improves the polarization Bayes factor by a factor of e^2 or more.

The lower panel of Fig. 4.6 shows that, for a small fraction ($\sim 5 - 15\%$) of the simulated events, the PO ratio is less than (although very close to) one. That is, the lensing Bayes factor [Eq. (4.14)] assuming the right polarization hypothesis is slightly smaller than the same

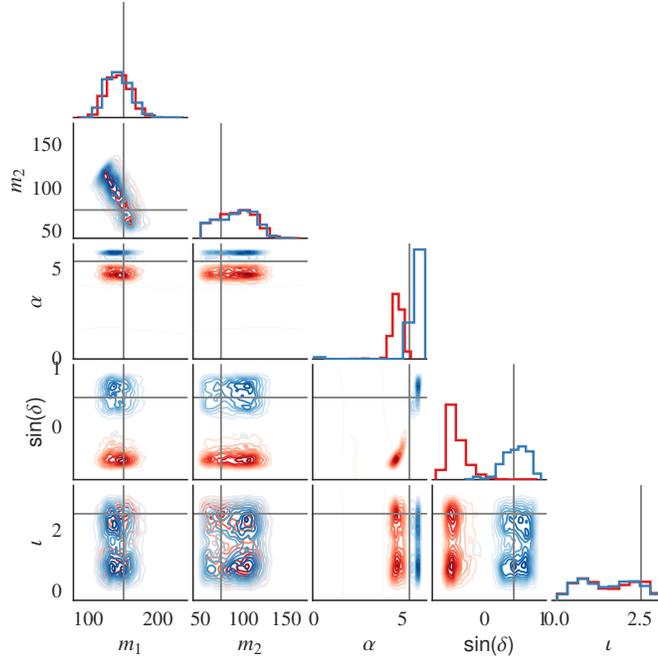


Figure 4.5: Same as Fig. 4.4, except that the injection is performed using the tensor polarization model while parameter estimation is performed using the scalar model (i.e., $\mathcal{H}_i^{[I]} - \mathcal{H}_s^{[R]}$ case). Note that the posteriors of the extrinsic parameters $\alpha, \sin \delta, \iota$, estimated from the two images, are not always overlapping and are not always consistent with the injected values.

assuming the wrong polarization hypothesis. These unusual event pairs do not show any significant correlations with the intrinsic or extrinsic parameters of the simulated BBHs. However, these event pairs have small lensing time delays (\sim less than half an hour). This is evident from Fig. 4.7, which plots the log overlap ratios for all the lensed event pairs against the lensing time delay between these images. This observation is broadly consistent with our expectation: during such short time delays (\sim less than half an hour) the change in the antenna patterns of the detectors due to the rotation of the earth is negligible. Hence the antenna patterns at the times of the two images will be very similar to each other. Thus, the linear combination of the polarizations measured from the two events will be practically the same. In other words, the posteriors estimated from the two events will have high overlaps, irrespective of the polarization model used. For such events, lensing is not expected to bring significant additional improvements. This is clear from Fig. 4.7, which shows that the overlap ratios from the small-time-delay events are modest (less than $\sim e^4$).

However, we would normally expect that the PO statistic of the right

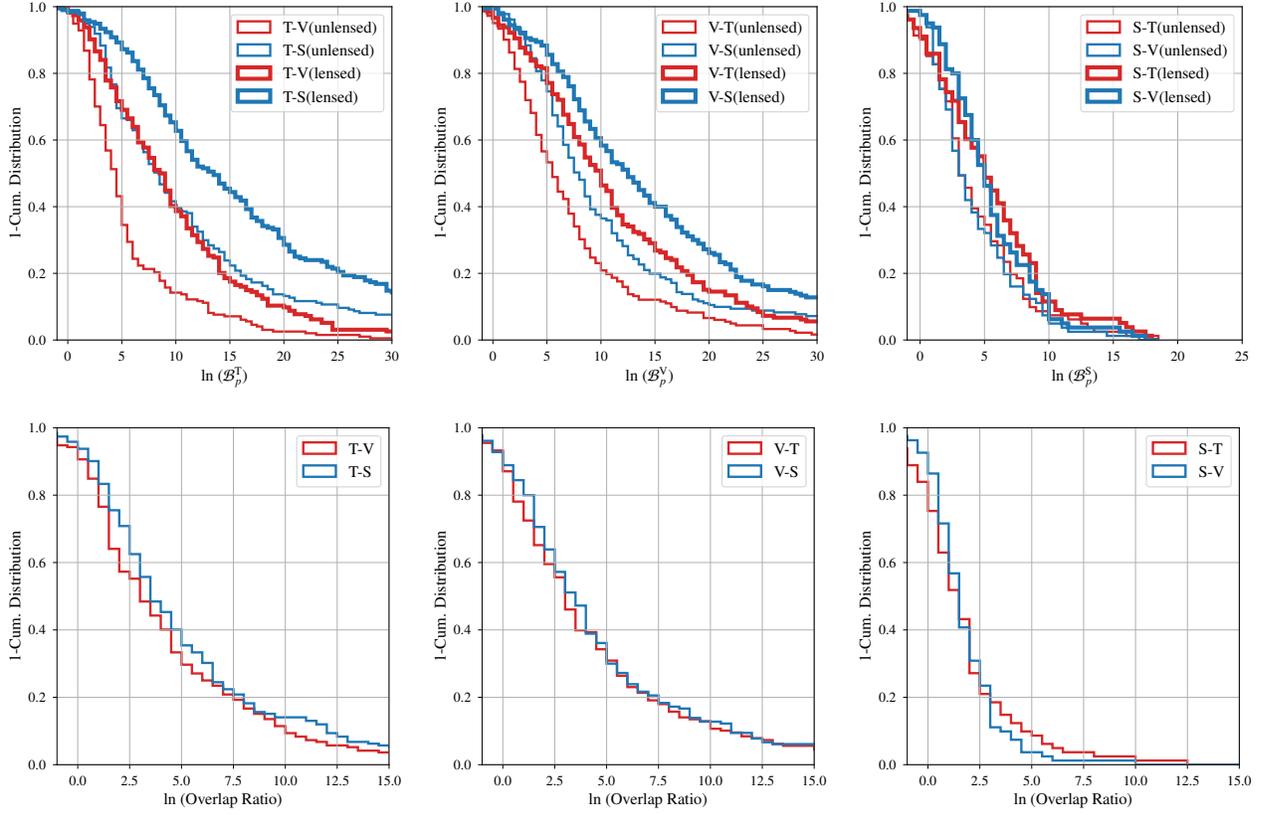


Figure 4.6: *Top panels:* Distribution of the Bayes factor between the “right” and “wrong” polarization hypotheses estimated from pairs of events (lensed or unlensed). Events are simulated assuming the tensor polarization hypothesis \mathcal{H}_t (left panel), vector polarization hypothesis \mathcal{H}_v (middle panel) as well as scalar polarization hypothesis \mathcal{H}_s (right panel). Each plot shows the distribution of the Bayes factors between the right and wrong polarization hypotheses (for e.g., T-V in the legends denote the Bayes factor \mathcal{B}_p^t). Note that the Bayes factors for the lensed pairs are almost always greater than the same computed from unlensed events with the same parameters. *Bottom panels:* Corresponding distribution of the ratios of the overlap factors \mathcal{B}_l^i assuming the “right” and “wrong” polarization hypotheses. Note that the overlap ratio is greater than 1 for 85 – 95% of events, suggesting that lensing improves the Bayes factors of the right hypothesis.

polarization model to be at least as large as the same using the wrong polarization models (in other words, the PO ratio should be greater than or equal to one). The reason for a small fraction of events to have overlap ratios slightly less than one is not well understood. It is likely that, when posteriors (using different polarizations) have very similar overlaps, the final results could be dominated by the numerical errors in our computations, such as the convergence of the parameter estimation and the inaccuracies in estimating the posterior distributions. We leave the detailed investigations on this as future work.

4.5 Summary and Future Work

Probing the polarization content of the GWs observed by a network of ground-based detectors offers an interesting probe on the nature of gravity. While GR predicts only two (tensor) polarization modes, there

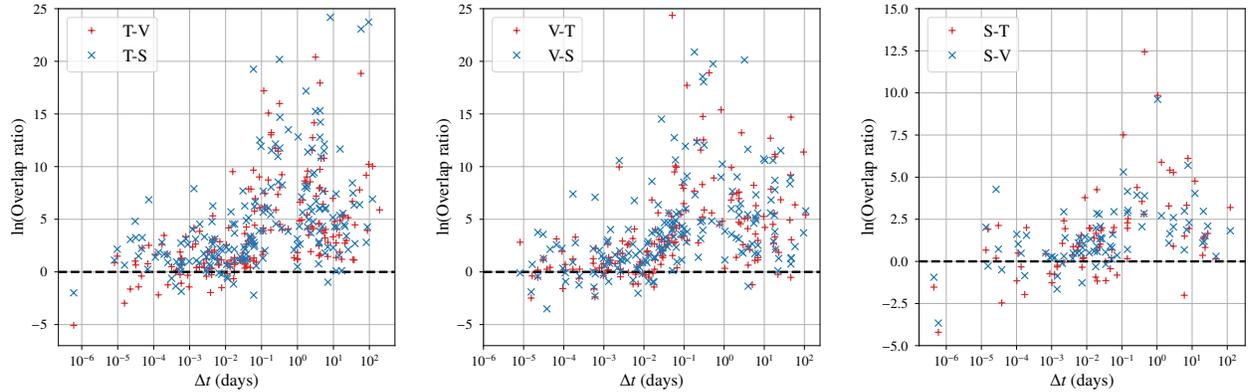


Figure 4.7: Correlation of the overlap ratios with the lensing time delay Δt between the images for each set of injections, created using the tensor (left), vector (middle) and scalar (right) polarization models. Different color markers show the overlap ratios between the “right” and “wrong” polarization models (for e.g., T-V denotes the overlap ratio between posteriors computed using the tensor and vector models). For the events below the black dashed lines, the PO ratio is less than one; hence lensing does not improve the polarization model selection.

are alternative theories that predict up to six polarizations (including scalar and vector modes, apart from the tensor modes predicted by GR). Each ground-based interferometer measures one particular linear combination of all these polarizations. Thus, if there are as many linearly independent detectors in the network as the number of independent polarizations, these polarizations modes can be extracted from the data, in principle. It turns out that the two scalar modes are degenerate as far as observations of ground-based detectors (which are quadrupolar antennas) are concerned. Thus five linearly independent polarization modes can be, in principle, extracted from the data of five linearly independent detectors. In practice, our ability to do this is limited by the presence of noise. In addition, the similar orientation of the two LIGO detectors in the USA makes this job difficult even with the upcoming network of five detectors including LIGO, Virgo, KAGRA and LIGO-India.

Strong lensing of GWs can significantly improve our ability to constrain GW polarizations. Recent estimates suggest that $\sim 0.1 - 0.5\%$ of the GW signals from BBHs that the Advanced GW detectors will observe in the next few years will be strongly lensed by intervening galaxies, producing multiple “images” (copies) of the same signal that arrive at the detector with relative time delays of several minutes to weeks. Since several hundred BBH detections are expected in the next few years, the first observation of lensed GW signals is likely to hap-

pen soon. Since the wavelength of the GWs is significantly smaller than the mass scale of these lenses, lensing effects can be calculated using geometric optics. In this limit, lensing does not affect the frequency profile of the GW signals. Thus, the multiple images of a single merger will be comprised of the same GW polarizations (albeit with a relative magnification). Due to the rotation of the earth, each detected image will allow the GW detector network to measure different linear combinations of the same polarizations. This is effectively equivalent to multiplying the number of detectors to observe a single GW signal.

We study the expected improvement, due to lensing, in our ability to probe the nature of GW polarizations making use of the Bayesian model selection formalism that was originally proposed by [84]. This uses a simplified model for the GW polarizations that are not present in GR: We assume that the time evolution of the additional polarizations (scalar and vector modes) follow that of the tensor modes. (Hence our ability to distinguish the polarization models depend greatly on the response of the GW detector network to different polarizations.) Additionally, we make a simplistic assumption that the GW polarizations consist of pure tensor, vector, or scalar modes. We show that strong lensing greatly improves our ability to distinguish the “right” and “wrong” polarization models for the GW signal. Note that, in our simulated GW observations, we do not take into account the effect of noise fluctuations in GW data (although we take into account the effect of the statistical properties of the noise by means of the power spectrum). The random noise fluctuations will cause additional scatter in the recovered Bayes factors and can further limit our ability to distinguish different polarization models, in real observations.

The joint Bayes factors (likelihood ratio between two polarization models) for multiple, unrelated events can be obtained as the product of Bayes factors computed for individual events as the noise and the signal in the individual data segments are unrelated. However, for a pair of strongly lensed events, though the noise is uncorrelated the GW signals present in the data segments are related. We show that the combined Bayes factor from such a lensed event is equal to the product

of the individual Bayes factors and an additional factor, namely, *PO ratio*, which is the ratio of the prior weighted overlaps of the posterior distributions of the GW parameters⁸ that are computed assuming the two polarization models under consideration. From simulated BBH events in the three detector network consisting of Advanced LIGO and Virgo detectors in design sensitivity, we show that the overlap ratio for the majority of lensed events ($> 50\%$) is greater than $e^2 - e^3$. This means that the Bayes factor supporting the right polarization hypothesis is improved by a factor of $\sim 7 - 20$ for most of the lensed events (as compared to pairs of unlensed signals with similar strengths). The improvement can be as large as several thousands for about 10% of the events. Note that, in this work, we only consider lensing by galaxies, under the assumption that lensing probability of galaxy clusters is negligible. Lensing by clusters will introduce much larger time delays between images, thus significantly improving our ability to distinguish between polarization models (see, e.g., Fig. 4.7). The simplistic polarization models that we use in this work can be extended to more realistic models, where the alternative model to GR would include scalar/vector modes *in addition to* the tensor modes. Even if we assume that the scalar/vector modes follow the same phase evolution as the tensor modes, this will require us to model the effect of the binary’s additional loss of energy and angular momentum (due to additional polarizations) on the orbital evolution itself. Additionally, the polarization model that is used in the model selection will require additional parameters that describe the relative strengths of the scalar, vector, and tensor modes (which will need to be marginalized away). Even then, the model selection described in Sec. 4.3.2 can be used to characterize the expected improvement due to lensing. Since the “right” polarization model is expected to produce larger overlaps between the posteriors estimated from multiple lensed images, we expect that strong lensing will provide similar improvements in our ability to do model selection. Note that, in this work we consider only double images produced by lensing, while $\sim 30\% - 40\%$ of the lensed events will also produce triple or quadruple images [180], potentially

⁸ To be precise, the posterior distributions of the parameters that are expected to be common between the images.

providing further improvements in the polarization model selection. Such improvements in the effective number of detectors in the network might also enable us to perform polarization reconstruction in a model agnostic way. We plan to explore these aspects as follow-up projects.

5 | Lens-induced GW birefringence as a test of GR

Abstract

Theories beyond GR modify the propagation of GWs. In some, inhomogeneities (aka. gravitational lenses) allow interactions between the metric and additional fields to cause lens-induced birefringence (LIB): a different speed of the two linear GW polarisations (+ and \times). Inhomogeneities then split the GW signal into two components whose relative time delay depends on the theory and lens parameters. Here we study the observational prospects for GW scrambling, i.e when the time delay between both GW polarisations is smaller than the signal's duration and the waveform recorded by a detector is distorted. We analyze the latest LVK catalog, GWTC-3, and find no conclusive evidence for LIB. The highest log Bayes factor that we find in favour of LIB is 3.21 for GW190521, a particularly loud but short event. However, when accounting for false alarms due to (Gaussian) noise fluctuations, this evidence is below $1-\sigma$. The tightest constraint on the time delay is < 0.51 ms (90% C.L.) from GW200311_115853. From the non-observation of GW scrambling, we constrain the optical depth for LIB, accounting for the chance of randomly distributed lenses (eg. galaxies) along the line of sight. Our LIB constraints on a (quartic) scalar-tensor Horndeski theory are more stringent than solar system tests for a wide parameter range and comparable to GW170817 in some limits. Interpreting GW190521 as an Active Galactic Nuclei (AGN) binary (i.e. taking an AGN flare as a counterpart to the merger) allows even more stringent constraints. Our results demonstrate the potential and high sensitivity achievable by tests of GR, based on GW lensing.

5.1 Introduction

As discussed in Chapter 1, the detection of GWs using the LVK detectors [23, 24, 25] from CBCs [7, 210, 211, 236, 16, 213, 237, 212] has enabled precision tests of GR in the strong field regime [68, 70, 238, 239].

Far away from the source, GR predicts that GWs are well described as linear perturbations of the background Friedmann-Robertson-Walker (FRW) metric [73]. Existing propagation tests hence typically consider modifications over the FRW background and its effect on the GW signals as measured at the detectors [75, 76, 77].

GR also dictates that GWs have only two tensor polarisations ($+$, \times) which propagate independently from each other at the speed of light. However, in alternative theories of gravity, extra degrees of freedom (tensor, vector, scalar, see Ch. 4.2) [63] can mix with GWs as they propagate, producing phenomena similar to neutrino oscillations (i.e. due to interactions between different neutrino flavours [240]). In Lorentz invariant theories, the symmetries of the FRW metric restrict mixing effects to tensor degrees of freedom, either fundamental (e.g. in bigravity) or composite (e.g. multiple vector fields) [241, 242, 243, 244, 245]. However, inhomogeneities spontaneously break Lorentz symmetry, allowing interaction between GWs and scalar or vector degrees of freedom [12, 246]. This leads to new, testable predictions, and opens new opportunities to probe the gravitational sector beyond the FRW limit.

The evolution of GWs on an inhomogeneous background is described via propagation eigenstates: linear combinations of the interaction eigenstates (h_+ , h_\times and perturbations of additional polarisations) with a well-defined dispersion relation (analogous to massive neutrinos). As the relation between interaction and propagation eigenstates and their speed depend on position and direction, an inhomogeneous region of space splits the original signal into several components, each arriving with a relative time delay [12]. Moreover, if deviations from GR are small, two eigenstates correspond to *mostly-tensorial polarizations* (linear combinations of h_+ , h_\times plus a negligible correction distinguishing both), with a very small speed difference.¹

We will refer to the difference in propagation speed between the $+$, \times polarizations as *lens-induced birefringence* (LIB). LIB is analogous to the way a non-isotropic crystal, such as calcite, splits light into two beams. This splitting is caused by a difference in the refractive index of the linear EM polarizations, which depends on the

¹We will ignore the remaining eigenstates (perturbations of beyond-GR fields plus negligible corrections) because 1) their emission needs to be suppressed to avoid dipolar radiation and 2) their speed can be substantially different, making an association with the mostly-tensorial part of the signal difficult [12].

alignment of the polarization vector with the crystal structure. In our case, birefringence is caused by the background configuration of additional, non-GR fields which spontaneously break Lorentz symmetry. Moreover, LIB splitting is independent of the frequency (in the high-frequency approximation assumed), which would correspond to a perfectly isochromatic birefringent crystal. Because GW detectors have excellent time resolution and bad sky localization, our main observable will be the time delay between split signals and not their angular separation.

If the arrival time difference between the mostly-tensorial polarisations is larger than the duration of the binary merger signal then we would see only one linear polarisation at a time [12], where one of them would be appearing as an edge-on binary. If the delay between the polarisations is larger than typical observing runs or the amplitude of one of the polarisations decays faster than the other, e.g. in Chern-Simons gravity [247], one would also expect an anisotropic inclination distribution. Current observations though show that the orientation distribution is consistent with being isotropic [248].

However, when the time delay is less than the duration of the signal, the GW waveform would be distorted or “scrambled” due to the interference of both polarisations. Note that this effect is frequency-independent, and hence distinguishable from a different dispersion relation for the $+$ and \times modes or the circularly polarized combinations (L-R), as predicted in GR [249, 250, 251] and alternative theories [252, 253, 254]. As it is not suppressed by the frequency, LIB is the dominant effect in the high-frequency limit for theories in which this effect is present.

Our study analyzes for the first time arrival time difference (Δt_{12}) between the two polarisation states due to different propagation speeds (frequency-independent dispersion relations) as a result of LIB. This is a new, model-independent test of a basic prediction of GR. We use these generic results to constrain GW lensing effects beyond GR, for example, in scalar-tensor theories with derivative interactions [12].

LIB signatures are not linked to a specific regime of gravitational

lensing in GR, such as strongly magnified or multiple images. The scale on which LIB can be observed is very sensitive to the theory parameters and independent of the Einstein radius R_E , which characterizes the regimes of gravitational lensing. Hence, for sufficiently strong deviations from GR, LIB can be detected for impact parameters much larger than R_E , typically associated with weak lensing. Therefore, LIB-tests can be applied to all the GW detections. In addition, LIB can be important for lenses very close to the source or the observer, for which R_E vanishes. This is particularly interesting for sources merging near massive objects (e.g. a supermassive BH) since the background configuration of the additional fields enhances LIB.

The rest of the chapter is organised as follows. In Sec. 5.2 we describe our LIB waveform model, and methods for data analysis and introduce parameterized LIB observation probabilities. In Sec. 5.3, we perform the birefringence test over a set of simulated GW events, and then to real events using the Bayesian model selection framework. In Sec. 5.4, we study the implications of the results in constraining LIB probabilities and beyond-GR theories. Finally, in Sec. 5.5 we summarize the main results and discuss future prospects.

5.2 Method

In GR, GWs have only two polarisations ($+$, \times) which propagate independently at the speed of light over the background FRW metric. A given ground-based detector I measures the GW signal, $h(t)$ as a linear combination of these polarisations [235],

$$h_I(t) = F_I^+ h_+(t) + F_I^\times h_\times(t) \quad (5.1)$$

where, F_I^+, F_I^\times are the detector antenna pattern functions. In the case of CBCs, the relative amplitude of the polarisation modes depends on the inclination and polarisation angles of the binary w.r.t to the line of sight, and also depends on the sensitivity of the detector for the source location at the time of arrival. The overall amplitude of the signal is inversely proportional to the luminosity distance of the source. Masses

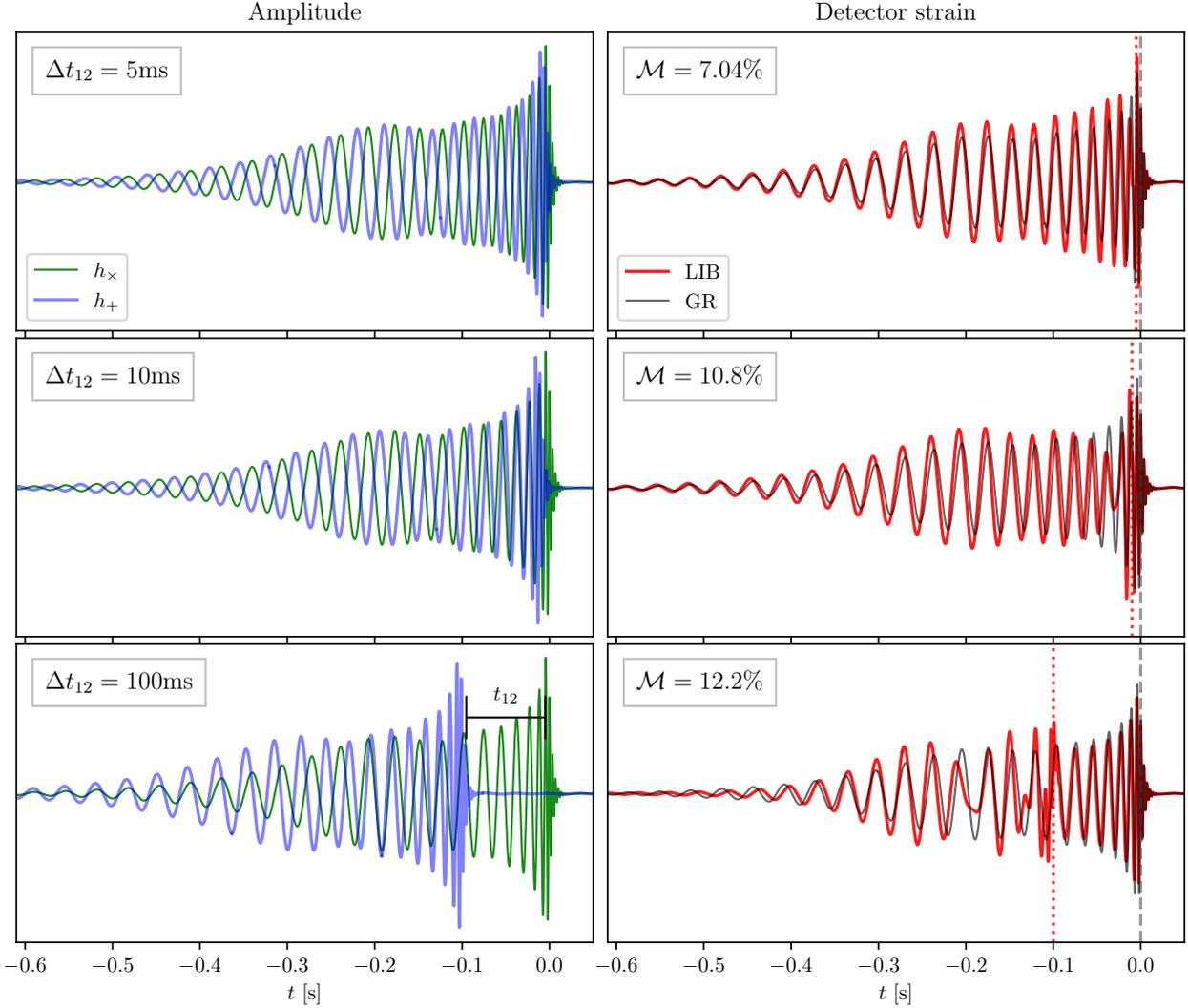


Figure 5.1: GW polarisations (left) and detector strain (right) for a CBC $(30 + 30)M_{\odot}$ with birefringent time delays $\Delta t_{12} = 5, 10, 100$ ms (top to bottom). The sky localization and detector orientation correspond to $F^+ = -0.38$, $F^{\times} = 0.71$ and LIB strain is given by Eq. 5.5

and spins of the source dictate the frequency evolution of the signal and its amplitude.

5.2.1 Parameterized Lens-induced Birefringence Waveforms

When there is any inhomogeneity along the travel path of a GW, e.g. an intervening galaxy, the GW can be gravitationally lensed. Gravitational lensing of a GW can produce multiple images of the original signal (strong lensing) or cause distortions (microlensing), but, in GR, both polarisations are affected in the same way, i.e. the polarisa-

tion rotation is negligible for any sensible astrophysical lens [91, 255]. However, in alternative theories of gravity the additional fields can couple with the tensor polarizations around the lens and modify the GW propagation eigenstates. These eigenstates are a linear combination of original GW polarisations that evolve independently, each with a different speed, thus reaching the detectors at different times. We will assume spherically symmetric lenses, focus on the limit of small deviations from GR, so the mostly-metric propagation eigenstates correspond to linear combinations of h_+, h_\times (depending on the projected angle between the lens and the source), and neglect the additional eigenstates (See Ref. [12] and footnote 1).

This class of LIB of GWs can be captured in a phenomenological manner as proposed in Ref. [12]. After diagonalizing the propagation equations, the propagation eigenstates can be computed and one gets the transformation matrix \mathcal{S} relating the polarisation amplitudes in GR and after the LIB:

$$[h_+, h_\times]_{\text{LIB}}^T = \mathcal{S}[h_+, h_\times]_{\text{GR}}^T \quad (5.2)$$

where

$$\mathcal{S} = \hat{\mathcal{M}} \text{diag}(1, \Delta) \hat{\mathcal{M}}^{-1}, \quad (5.3)$$

$$\hat{\mathcal{M}} = \begin{bmatrix} -\sin(2\phi_{\text{lens}}) & \cos(2\phi_{\text{lens}}) \\ \cos(2\phi_{\text{lens}}) & \sin(2\phi_{\text{lens}}) \end{bmatrix}, \quad (5.4)$$

and $\Delta = e^{-i\omega\Delta t_{12}}$ with Δt_{12} is the time delay between the polarisations and ϕ_{lens} is the angle that projects GW polarisations onto the lens plane, perpendicular to the direction of propagation.

It is easy to note that for $\phi_{\text{lens}} = \pi/2$, $\mathcal{S} = \text{diag}(1, \Delta)$, hence the signal observed by the detectors will just be superposition of $(+, \times)$ arriving at different times.

$$h_I^{\text{LIB}}(t) = F_I^+ h_+(t) + F_I^\times h_\times(t - \Delta t_{12}) \quad (5.5)$$

whereas, if $\Delta t_{12} = 0$ the LIB waveform morphology will be identical to the GR one, independent of ϕ_{lens} . Fig. 5.1 compares GR v/s

LIB waveform polarisations and the detector strains for various values Δt_{12} . Under LIB, the polarisations interfere leading to waveform distortions.

Since lensing is an environmental effect which can occur through any local inhomogeneity in the path of GWs, the parameters Δt_{12} and ϕ_{lens} are expected to vary between GW events. The time delay distribution depends on the theory and the (usually unknown) lens properties and the configuration relative to the source. In general, one can only predict the probability of the birefringence parameters given a gravitational theory and matter distribution (unless further information or assumptions are employed about the source's location or the signal's trajectory), see Sec. 5.2.4. This is in stark contrast to other tests of GW propagation (that are done with individual GW events) in which deviations represent a fundamental property of gravity (eg. massive graviton dispersion relations) and are thus the same across all events and only depend on their distance [63].

5.2.2 Template Mismatch Studies

In order to quantify distortions due to GW birefringence, we calculate the mismatch between the GR and LIB waveforms as seen the LIGO-Virgo detectors. At each detector (I), the mismatch between the injected waveform (h_I^{inj}) and the recovery waveform (h_I^{rec}) is given by:

$$M_I = 1 - \frac{(h_I^{\text{inj}} | h_I^{\text{rec}})}{\|h_I^{\text{inj}}\| \cdot \|h_I^{\text{rec}}\|} \quad (5.6)$$

where, $(\cdot | \cdot)$ symbolises the noise-weighted inner product:

$$(a | b) \equiv 2\Re \int_{f_{\min}}^{f_{\max}} \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_n(f)} df. \quad (5.7)$$

Here, \tilde{a}, \tilde{b} represent the Fourier transform of the time series $a(t), b(t)$; $[f_{\min}, f_{\max}]$ is the frequency range over which the inner product is evaluated; $*$ represents complex conjugation and $S_n(f)$ is the noise power spectral density (PSD) at the detector. The norm $\|h\| = \sqrt{(h|h)}$ is the optimal SNR of a waveform. We define the total mismatch \mathcal{M}

for network of detectors as the SNR (ρ_I) squared weighted average of individual detector matches,

$$\mathcal{M} = \frac{\sum_I \rho_I^2 M_I}{\sum_I \rho_I^2} \quad (5.8)$$

Note that the mismatch is a normalized quantity and is maximised over time and phase shifts. Thus, the mismatch quantifies differences in morphology between the signals. Whereas, during the parameter estimation (PE) from GW signals both the mismatch and SNR play a role as the log-likelihood $\propto \mathcal{M} \sum_I \rho_I^2$ around the maximum likelihood parameters. We first wish to quantify the overall detectability of the birefringence. Later, we will estimate parameters using Bayesian inference, accounting for correlations between all parameters.

Fig. 5.2 shows frequency domain LIB and GR waveforms for a GW₁₅₀₉₁₄-like CBC. The waveforms are generated using the approximant IMRPHENOMXPHM [256], as implemented in the LALSUITE software package [156]. The waveforms are then projected onto the LIGO and Virgo detectors using their antenna pattern functions, as implemented in the BILBY [234] software package. The LIB waveforms have additional frequency modulations which depend on the two parameters: Δt_{12} and ϕ_{lens} (see 5.2.1). Therefore we calculate the mismatch between the GR and LIB waveforms, keeping all the other parameters identical and fixed for the two waveforms. In practice, we calculate M_I using PYCBC.FILTER [199] module. The detector noise is generated using the zero-detuned high-power PSDs of Advanced LIGO and Advanced Virgo at their design sensitivities [158, 159].

We consider two systems of binaries, one whose parameters resemble to that of the first CBC detection GW₁₅₀₉₁₄ and other of a higher mass-ratio CBC GW₁₉₀₈₁₄ where the presence of higher-order modes (HoMs) of GWs are significant. For both the systems, we inject a GR and a LIB waveform and recover with LIB waveform to calculate the mismatch. The parameters for both the CBCs are mentioned in Appendix 5.5 Fig. 5.3 shows mismatches for a GW₁₅₀₉₁₄-like CBC (top) and GW₁₉₀₈₁₄-like CBC (bottom). As expected for a GR injection i.e.

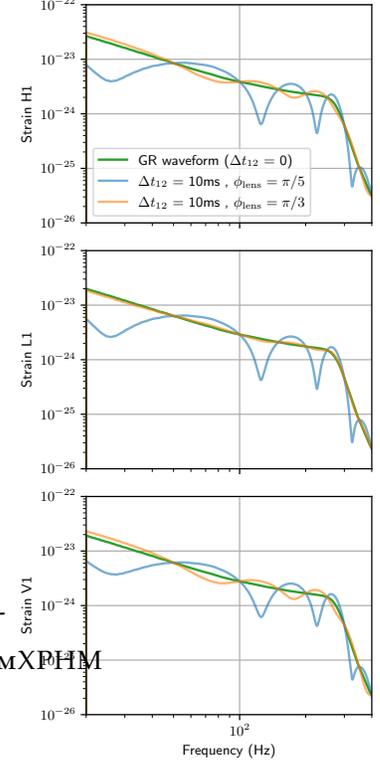
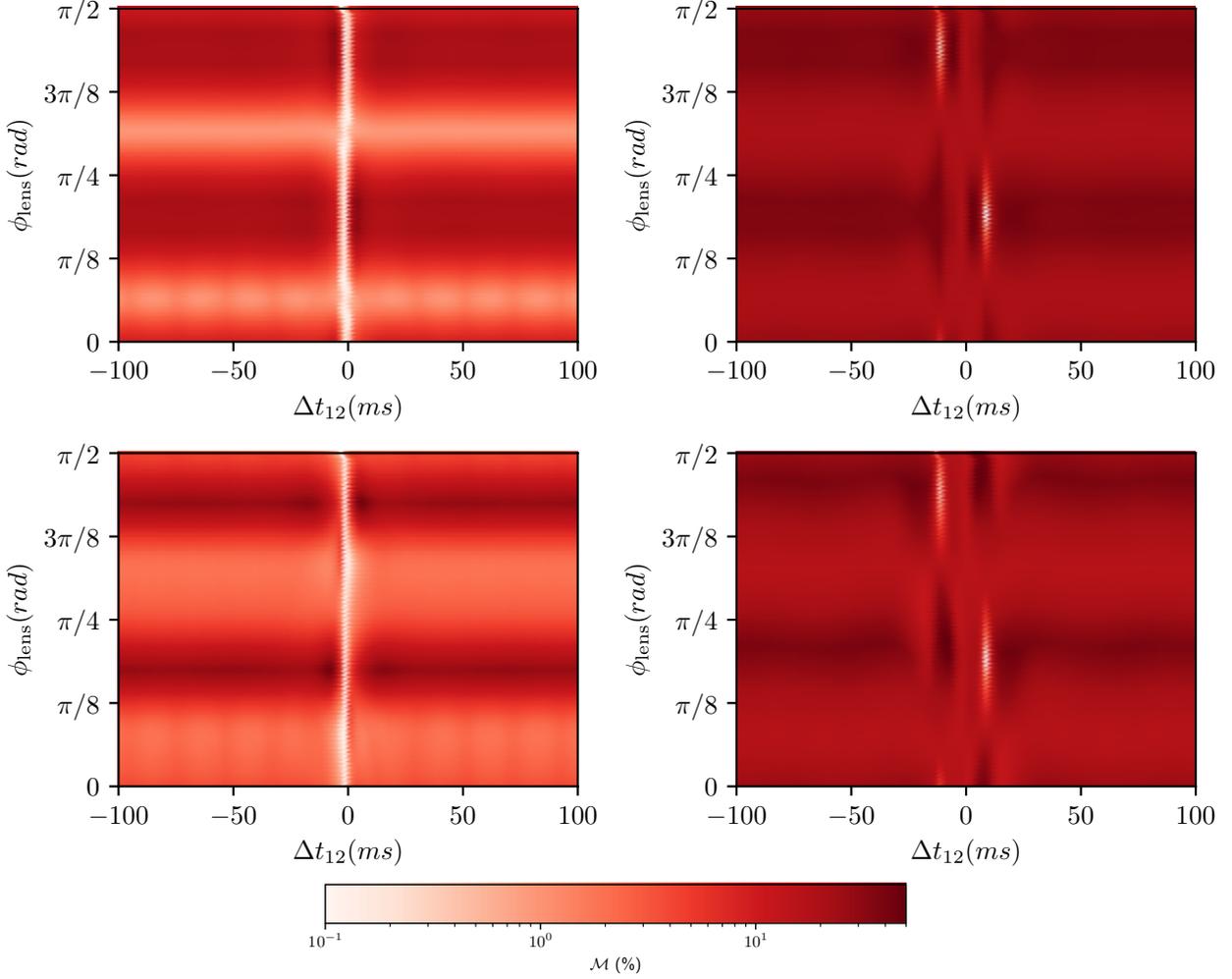


Figure 5.2: GR and LIB waveform amplitudes in frequency domain of GW₁₅₀₉₁₄-like CBC projected to the LIGO-Virgo detectors. The birefringence leads to additional frequency modulations and distorts the GR waveform. The magnitude of these distortions are however dependent on the two parameters: Δt_{12} and ϕ_{lens} .



$\Delta t_{12}^{\text{inj}} = 0$ and $\phi_{\text{lens}}^{\text{inj}} = 0$, the mismatch is minimum for $\Delta t_{12}^{\text{rec}} \simeq 0$, for all ϕ_{lens} as expected from Eq. 5.3. Additionally, the local minimum of mismatch is at $\phi_{\text{lens}}^{\text{rec}} \simeq \pi/3$, which could be because of vanishing polarisation (+ or \times) as seen at the detectors which further makes the mismatch independent of the time delay Δt_{12} . The waveform plots in Fig. 5.2 confirms this as the $\phi_{\text{lens}}^{\text{rec}} \simeq \pi/3$ waveform resembles the GR ones more as compared to the $\phi_{\text{lens}}^{\text{rec}} \simeq \pi/5$, especially in the Livingston (L1) detector for GW150914-like CBC.

We also checked the mismatch for the LIB injections (right panel Fig. 5.3) with $\Delta t_{12}^{\text{inj}} = 10$ ms and $\phi_{\text{lens}}^{\text{inj}} = \pi/5$ and the mismatch is minimum at, $\Delta t_{12}^{\text{rec}} \simeq \pm 10$ ms and $\phi_{\text{lens}}^{\text{rec}} \simeq \pi/5, \pi/4 + \pi/5$. We

Figure 5.3: Mismatch between GR and LIB waveforms for GW150914-like CBC (top) and GW190814-like CBC (bottom). left panel: GR injection i.e. $\Delta t_{12}^{\text{inj}} = 0$ and $\phi_{\text{lens}}^{\text{inj}} = 0$. the mismatch is minimum for $\Delta t_{12}^{\text{rec}} \simeq 0$. right panel: a LIB injection with $\Delta t_{12}^{\text{inj}} = 10$ ms and $\phi_{\text{lens}}^{\text{inj}} = \pi/5$. The mismatch is minimum at $\Delta t_{12}^{\text{rec}} \simeq \pm 10$ ms and $\phi_{\text{lens}}^{\text{rec}} \simeq \pi/5, \pi/4 + \pi/5$.

can infer the degeneracy between Δt_{12} and the coalescence time (t_c) as follows, from Eq. (5.1)-(5.4) if $\phi_{\text{lens}} \rightarrow \phi_{\text{lens}} + \pi/4$ and $\Delta \rightarrow 1/\Delta$ then, one finds $S \rightarrow S/\Delta$, which implies that the transformation matrix \mathcal{S} in Eq. (5.3) is invariant when $(\Delta t_{12}, \phi_{\text{lens}}, t_c)$ is transformed to $(-\Delta t_{12}, \phi_{\text{lens}} + \pi/4, t_c + \Delta t_{12})$. This degeneracy stems from the fact that we do not know the composition of $h_{+, \times}$ before it encounters the lens. Higher harmonics of the waveform can provide this additional information, as the amplitude of the $+, \times$ polarizations for each harmonic has a different dependence on the inclination angle (cf. Fig. 1 in Ref. [257]). Sources with a high mass ratio may thus distinguish the sign of Δt_{12} . We leave these investigations for the future.

5.2.3 Bayesian Inference

Bayesian model selection framework to compare the two hypotheses, the odds ratio $\mathcal{O}_{\text{GR}}^{\text{LIB}}$ is the ratio of the posterior probabilities of the two hypotheses LIB and GR can also be written as the product of the ratio of the prior odds $P_{\text{GR}}^{\text{LIB}}$ of the hypotheses and the likelihood ratio, or Bayes factor $\mathcal{B}_{\text{GR}}^{\text{LIB}}$:

$$\mathcal{O}_{\text{GR}}^{\text{LIB}} := \frac{P(\mathcal{H}^{\text{LIB}}|\{d\})}{P(\mathcal{H}^{\text{GR}}|\{d\})} = \frac{P(\mathcal{H}^{\text{LIB}})}{P(\mathcal{H}^{\text{GR}})} \times \frac{P(\{d\}|\mathcal{H}^{\text{LIB}})}{P(\{d\}|\mathcal{H}^{\text{GR}})} \quad (5.9)$$

$$= \mathcal{P}_{\text{GR}}^{\text{LIB}} \times \mathcal{B}_{\text{GR}}^{\text{LIB}} \quad (5.10)$$

Since GR has been tested well in a variety of settings, our prior odds are going to be highly biased towards it, i.e., $\mathcal{P}_{\text{GR}}^{\text{LIB}} \ll 1$. Hence, in order to claim evidence of birefringence the corresponding Bayes factor supporting the LIB hypothesis has to be very large. Since the Bayes factor is the only quantity that is derived from data, for the rest of the chapter, we focus on the Bayes factor, i.e. the ratio of evidence under the two hypotheses.

The waveforms under the GR and LIB hypotheses at each detector are defined in the same way as described in Sec. 5.2.2. We use the standard Gaussian likelihood model for estimating the posteriors of the parameters under different hypotheses (see, e.g., [235]). We use

uniform priors in redshifted component masses of the binary, isotropic sky location (uniform in $\alpha, \sin \delta$) and orientation (uniform in $\cos \iota, \phi_0$), uniform in polarisation angle ψ , and a prior $\propto d_L^2$ on luminosity distance. Additionally for the LIB hypothesis, we choose the priors on Δt_{12} as uniform $\in [-100, 100]$ ms and ϕ_{lens} as uniform $\in [0, \pi/2]$. To estimate the posterior distribution and evidences for the GR and LIB hypotheses, we use the open-source parameter estimation package `BILBY` package [234] coupled with the dynamical nested sampler `DYNESTY` [233].

5.2.4 *Lensing Probabilities*

The (non-)observation of birefringence can help us put constraints on theories beyond GR that predict LIB. According to GR, the strong lensing of GWs caused by galaxies occurs when sources lie inside the Einstein radius of the lens, which depends on lens mass and profile. This is the relevant scale determining the probability of lensing. However, birefringence beyond GR is in principle independent of the ratio between the impact parameter and the Einstein radius, changing the probability of observing LIB compared to strong lensing. It is thus possible to have LIB time delays without multiple images, but birefringence could also occur for the strongly lensed copies of GWs, in which case it applies to each image separately, as typical time delays between images will be larger than Δt_{12} [12].

Assuming that the lenses are randomly distributed, birefringence detection is described by Poisson statistics. A series of observations with L lensed and U unlensed GW events has an associated probability,

$$P = \exp\left(-\sum_i^U \lambda_i\right) \prod_j^L (1 - e^{-\lambda_j}). \quad (5.11)$$

The result depends on LIB rate for the i^{th} event:

$$\lambda_i = \int dz_s d\vec{p}_L d\vec{p}_S \tau(z_s, \vec{p}_L) \mathcal{P}_i(z_s, \vec{p}_S) P(\vec{p}_S, \vec{p}_L)$$

. (5.12)

Here S, L denote parameters corresponding to the source and lens/theory (i.e. beyond GR), \mathcal{P}_i is the posterior distribution of the source parameters and P is the prior, which includes relations between parameters (i.e. the measured Δt_{12} as a function of lens mass and beyond-GR parameters).² The *birefringence optical depth*, $\tau(z_s, \vec{p}_L)$ is the fraction of the sky for which LIB is detectable for sources at a redshift z_s . Hereafter we will assume the posterior to be sharply peaked at the mean source redshift z_s and include the integration on the lens model parameters (\vec{p}_L) in the definition of the optical depth, so $\lambda_i \approx \tau(z_{s,i})$. If birefringence is excluded in all events, the probability only depends on the total optical depth $\tau_{tot} \approx \sum_{i=1..N} \tau_i \approx \sum_{i=1..N} \lambda_i$.

The lensing optical depth $\tau(z_s)$ depends on the angular cross section $\hat{\sigma}(z_s, \vec{p}_L)$ and the density of lenses $\hat{n}(\vec{p}_L)$ [259]. In the following we will explicitly write the lens redshift z_L and let $\vec{p}_{L'}$ denote the remaining properties (i.e. lens mass & theory parameters). The total density of lenses at redshift z_L is then $\int \hat{n}(z_L, \vec{p}_{L'}) d\vec{p}_{L'}$. The optical depth is computed directly by adding-up the cross-sections weighted by the density at different redshifts, i.e.

$$\tau(z_s) = \int_0^{z_s} dz_L \int d\vec{p}_{L'} \frac{dV_c}{\delta\Omega dz_L} \hat{n}(z_L, \vec{p}_{L'}) \hat{\sigma}(z_s, z_L, \vec{p}_{L'}) \quad (5.13)$$

where $dV_c = \delta\Omega D_L^2 \frac{dz}{(1+z)H(z)}$ is the physical volume given the solid angle $\delta\Omega$, angular diameter distance to the lens D_L and the Hubble parameter $H(z)$. For simplicity, we will assume point mass lenses of mass M throughout. In GR, the lensing cross-section is $\sigma = \pi\theta_E^2$, where $\theta_E = R_E/D_L = (\frac{4GM D_{LS}}{c^2 D_L D_S})^{1/2}$ is the Einstein angle, D_S is the distance to the source from the earth and D_{LS} is the distance between lens and source.

The relation between the LIB-time delays, the theory parameters and the configuration of the lensed system is complex (See Sec. VI of Ref. [12] for a worked-out example in a viable Horndeski theory). For this reason we will first consider two phenomenological models with

² A more complete treatment should account for the selection function [258]. In terms of gravitational lensing, we expect that Δt_{12} correlates with magnification especially at sizeable impact parameters (e.g. single image regime, first magnified image if multiple images are formed), which dominate the lensing cross-section. Then events with larger Δt_{12} are more likely to be observed and neglecting this correlation is conservative.

generic dependences of lensing cross section. As a first example, we will assume that the relevant LIB scale is proportional to the Einstein angle, $\theta_X^E = \alpha_X \theta_E$, so that the cross section becomes

$$\sigma_X^E = \pi \alpha_X^2 \theta_E^2. \quad (5.14)$$

Then the optical depth is given by Eq. (74) in Ref. [12] where the lenses have been assumed point-like.

Under these assumptions, lensing probabilities are independent of the mass function.³

In our second example we assume that the relevant LIB scale is given by a constant physical scale associated to each halo. Moreover, we assume that this scale depends on the halo mass as a power law. Accordingly, the cross-section reads,

$$\sigma_{\text{ph}} = \pi \frac{R_{12}^2}{D_L^2} \left(\frac{M}{10^{12} M_\odot} \right)^{2n}. \quad (5.15)$$

The scale R_{12} fixes the probability of lensing for halos with $M = 10^{12} M_\odot$, while n allows us to extrapolate to different halo masses. Below we will discuss some cases of interest.

We now generalize the expression for the optical depth presented in Ref. [12] (Eq. 76) to include a realistic halo mass function. The optical depth from Eq. (5.15) is given by

$$\tau^{\text{ph}}(z_s, n) = \Omega_M h \left(\frac{R_{12}}{22 \text{kpc}} \right)^2 \hat{\tau}(z_s, n), \quad (5.16)$$

where

$$\begin{aligned} \hat{\tau}(z_s, n) = \int_0^{z_s} dz \frac{(1+z)^2}{H(z)/H_0} \int d \log(M) \\ \times \left(\frac{M}{10^{12} M_\odot} \right)^{2n-1} f(M, z). \end{aligned} \quad (5.17)$$

Here $f(M, z) = \frac{M^2}{\rho_0} \frac{d\hat{n}}{dM}$ is the scaled differential mass function (dimensionless) with ρ_0 as the matter density of universe at $z = 0$. We will use the Tinker *et al.* form [262] as implemented in the COLLOSSUS package [263]. As our approach is phenomenological, we assume

³ The mass independence also appears for the strong-lensing cross section for a distribution of point lenses [260, 261]. This differs from strong-lensing cross-section for extended lenses, where the lens mass affects the formation of multiple images [259].

a Planck Λ CDM cosmology [264]. The true optical depth of a consistent LIB model will typically depend more strongly on the theory parameters (e.g. entering Eq. (5.15) via R_{12}) than on the precise values of $H(z), f(M, z)$ of the underlying LIB cosmology, including the effects of deviations from GR in cosmological expansion and structure formation. This is the case for the example theory discussed in Sec. 5.4.2: GW lensing effects are orders of magnitude more sensitive than solar-system tests (cf. Fig. 5.9), in turn more stringent than current cosmological observations [265, 74] (for theories without a screening mechanism).

In addition to the Einstein radius scaling, Eq. (5.14), we will consider three cases of interest:

- $n = 1$: the physical scale is proportional to the total halo mass, much like the Schwarzschild radius. The rates are dominated by large masses and saturate at $z_s \gtrsim 1$, as the more massive halos are exponentially suppressed at early times. This case captures the dependence of the time delay in a Horndeski model (Sec. 5.4.2).
- $n = 1/2$: the scale has the same mass scaling as the Einstein radius and leads to rates independent of M . However, the overall redshift dependence is different, as R_E depends also on D_S, D_{LS} .
- $n = 1/3$: this mass scaling favors lighter halos and thus grows very rapidly with redshift. It is motivated by the mass-dependence of the Vainshtein radius R_V , i.e. the classical strong-coupling scale [266]. For $n = 1/3$ the contribution from lighter halos diverges and a low mass cutoff needs to be included (we will take $M > 10^7 M_\odot$). We will see this mass dependence when considering a binary merging near an active galactic nuclei in a Horndeski theory (Sec. 5.4.3).

The optical depths for each of the cases as a function of the source redshift are plotted in Fig. 5.8. Note that these phenomenological models assume that the cross section is independent of Δt_{12} , and thus common for all the analyzed events. Dependence in the time delay can be included, e.g. by multiplying Eqs. (5.14), (5.15) by a factor $(\Delta t_{12}/10\text{ms})^{-k}$. For the sake of simplicity, we will not include this

dependence and instead interpret the obtained values of $\alpha_X, R_{12}(n)$ at the median 95% c.l. from all analyzed events.

5.2.5 *Beyond-Poisson statistics*

The independent lens assumption fails to capture two circumstances that potentially enhance the detection of birefringence: the source environment and lensing by known objects. This situation is qualitatively different from strong lensing probabilities, which are weighted by the Einstein radius, which vanishes when $D_L \rightarrow D_S$ (near the source) or $D_L \rightarrow 0$ (near the detector). In contrast, birefringence probabilities do not suffer such suppression and can be sizeable for objects near the source or the observer. Our optical-depth framework (Sec. 5.2.4) does not consider this possibility.

Source environment may play a role for LIB, as GW sources will generally be located in regions denser than the cosmic average. In this case, the host galaxy (or objects within it) would have a much larger density compared to the cosmological average used in, e.g. Eq. (5.16). In addition, the projected cross-section $\propto 1/D^2$ will be larger for nearby objects, thus enhancing the probabilities. Given a distribution of GW sources near an object, the posterior on the theory parameters \vec{p} can be obtained as

$$P(\vec{p}) = \int dr d\theta P_s(r) \sin(\theta) P(\Delta t_{12}(r, \theta, \vec{p})) . \quad (5.18)$$

Here we have assumed a symmetric r -dependent distribution. The θ dependence corresponds to a uniform prior on the sphere. This simple dependence could be used to model the effect of the source's galaxy or nearby objects.

An extreme case of environmental enhancement is given by a binary merging in an AGN near a supermassive black hole (SMBH), will be discussed in Sec. 5.4.3, taking the multi-messenger scenario of GW190521 and its implications for the example Horndeski theory. Estimates for the rate of such events are uncertain. Nonetheless, in some cases it might be possible to associate an event with a SMBH

thanks to an EM counterpart [21], multiple images due to strong lensing [267, 268, 269] or strong-field propagation effects [251].

Another potential to improve the quoted result is by correlating GW arrival direction with known lenses. relevant in cases where the Milky Way (or perhaps even the Sun) may imprint an observable birefringence. Adding information on the GW direction, relative to known objects, will allow better constraints on those scenarios more effectively than assuming randomly located lenses. For instance, if stellar-scale lenses are relevant in a given theory and the cross-section scales as the physical radius (allowing nearby lenses to contribute), sources behind the milky way can probe a much larger effective cross-section than given by Eq. 5.17.

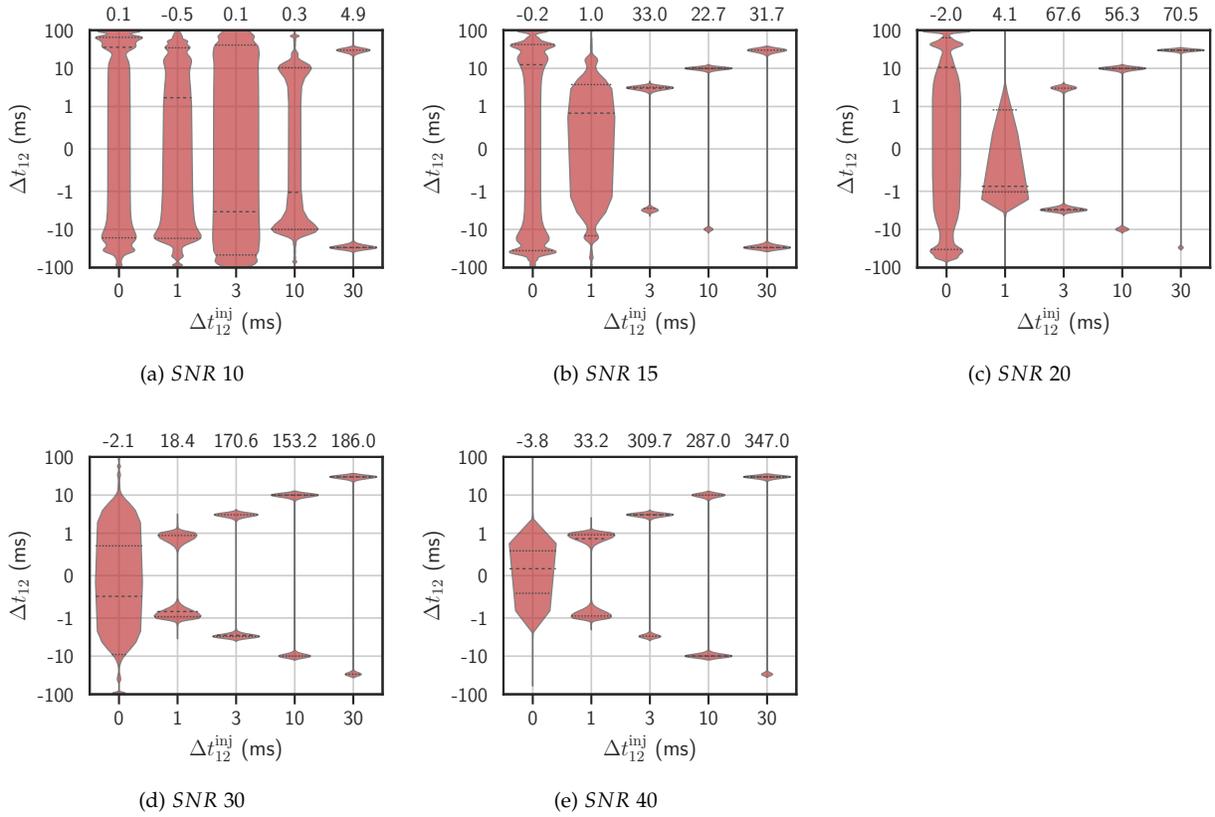
Finally, any confident detection of a lensed GW can be used to refine constraints within a given model. This would follow either through the identification of several GW detections as images of the same underlying source or through waveform distortions (millilensing). Both cases allow information about the lens mass and impact parameter to be recovered, at least when assuming a lens model [270, 271, 272]. That information can then place constraints within a specific theory of gravity.

5.3 Results

In order to test our method and understand the observing capabilities, we first apply our pipeline to injections. We then proceed to analyse the latest GW catalog (GWTC-3).

5.3.1 Injections

We inject GW150914-like signals in simulated Gaussian noise with $\Delta t_{12}^{\text{inj}} \in \{0, 1, 3, 10, 30\}$ ms and $\phi_{\text{lens}}^{\text{inj}} = \pi/5$ rad and recover them by running the parameter estimation routines under the GR and LIB hypothesis, as mentioned in Sec. 5.2.3. This allows us to compute the Bayes factors $\mathcal{B}_{\text{GR}}^{\text{LIB}}$ to compare the two hypothesis for each injection.



The injections are set to have $\text{SNR} \in \{10, 15, 20, 30, 40\}$. These SNRs are achieved by inversely scaling the luminosity distance (d_L) of the injections. Fig. 5.4 shows the violin plots and the log Bayes Factors for these injections. The posteriors on ϕ_{lens} are uninformative in all the cases and hence not shown in the figure. This is because ϕ_{lens} is highly correlated with other parameters like ψ, ϕ, ι , see Fig. 6.1 in the appendix. LIB time delays (Δt_{12}) as small as 1 ms are recovered well with SNR 30 and 40 signals, whereas for SNR 10 signals time delays < 30 ms are not measurable. As one would expect, only with $\Delta t_{12}^{\text{inj}} = 0$, i.e. GR injection the $\log \mathcal{B}_{\text{GR}}^{\text{LIB}} < 0$, i.e. consistent with the GR hypothesis, except for the SNR 10 case where $\log \mathcal{B}_{\text{GR}}^{\text{LIB}} = 0.1$ is within the intrinsic sampling error on the calculation of evidence. For $\Delta t_{12}^{\text{inj}} \in \{1, 3, 10, 30\}$ ms we find that $\log \mathcal{B}_{\text{GR}}^{\text{LIB}} > 0$, i.e. consistent with the LIB hypothesis for all the SNRs except 10. Hence, both model

Figure 5.4: SNR dependence of $\Delta t_{12}(\text{ms})$ posteriors and the $\log \mathcal{B}_{\text{GR}}^{\text{LIB}}$ (upper-x axis) for the GW150914-like injections with different values of $\Delta t_{12}^{\text{inj}}$ (lower-x axis) and $\phi_{\text{lens}}^{\text{inj}} = \pi/5$. Time delays (Δt_{12}) as small as 1ms are recovered well with SNR 30 & 40 signals, and for SNR 10 signals time delays < 30 ms are not measurable. Both model selection and time delay measurements (without symmetry around $\Delta t_{12} = 0$) improve with the increase in SNR.

selection and sensitivity to measure the time delays improve with an increase in SNR.

We also note that the time delays are measurable up to a symmetry around $\Delta t_{12} = 0$. This is because LIB waveform, Eq. (5.3), is identical at $(\Delta t_{12}, \phi_{\text{lens}}, t_c)$ and $(-\Delta t_{12}, \phi_{\text{lens}} + \pi/4, t_c + \Delta t_{12})$, which we also saw during mismatch studies with LIB injections, see right panel of Fig. 5.3. It is possible that for asymmetric and inclined binaries with significant HoMs a better measurement of ϕ_{lens} could break the Δt_{12} parity as well, however, this needs to be investigated further and is left for future studies.

Overall, as the sensitivity of the detectors improve we shall be able to measure the birefringence time delays as small as 1 ms. On the other hand, in the absence of birefringence we expect to see $\Delta t_{12}(\text{s})$ posteriors which are consistent with the GR value, i.e. zero and bayes factors that favour the GR hypothesis. Most events in GWTC-3 have $\text{SNR} < 30$. The time delay posteriors are hence expected to be broad, however the Bayes factors should already indicate whether LIB is present or not.

5.3.2 GWTC-3 Events

We now analyse 43 CBC events from the GWTC-3, that have low detection false alarm rate, $\text{FAR} \lesssim 10^{-3} \text{ yr}^{-1}$. These are also the events that are considered for other tests of GR performed previously [68, 70, 238, 239].

Fig. 5.5 shows the Δt_{12} posteriors and the log Bayes Factors for the real events. We find that for almost all the events the Δt_{12} posteriors are broad containing zero, i.e. consistent with GR. This is mostly due to the low SNRs of the events, as seen in our injection study. We also find the tightest 90% credible bounds on $|\Delta t_{12}| \lesssim 0.51 \text{ ms}$ coming from the event GW200311_115853 which has reasonably high SNR ($\simeq 17.8$) and moderate redshift ($z \sim 0.23$) as compared to other events. As expected ϕ_{lens} posteriors are uninformative for almost all the events.

38 out of 43 events resulted in $\log \mathcal{B}_{\text{GR}}^{\text{LIB}} < 0$, and hence consistent with the GR hypothesis. Only a few events showed preference to LIB hypothesis ($\log \mathcal{B}_{\text{GR}}^{\text{LIB}} > 0$), with highest one for GW190521 (3.21) and

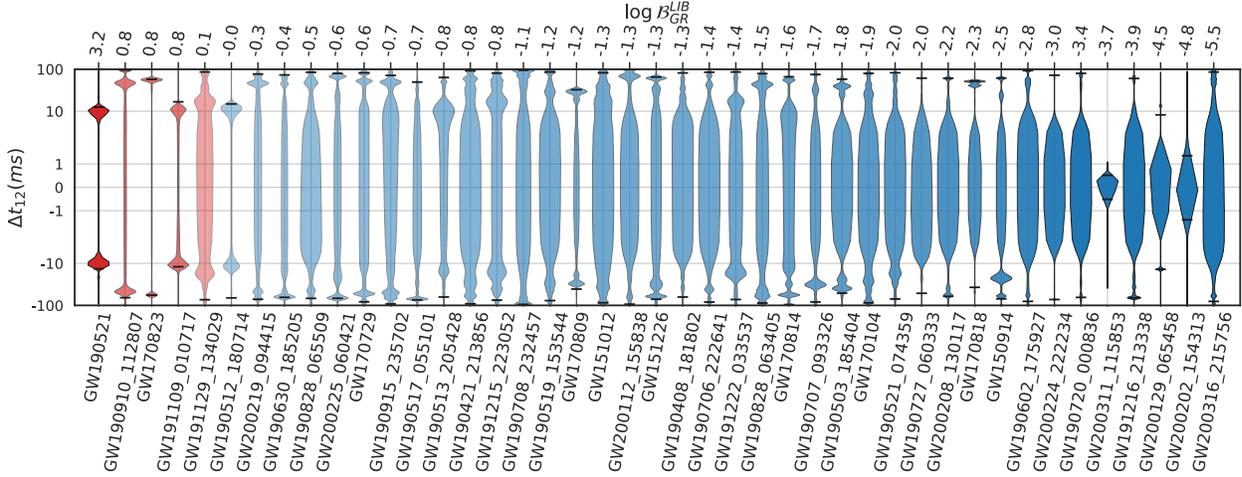


Figure 5.5: lens-induced birefringence (LIB) test of GWTC-3 events [16]. We show the posteriors on $\Delta t_{12}(\text{ms})$ and Bayes Factors $\log \mathcal{B}_{GR}^{LIB}$ (upper x -axis). Events with positive Bayes factors are highlighted in red.

then GW190910_112807 (0.8), GW170823 (0.8), GW191109_010717 (0.7) & GW191129_134029 (0.1).

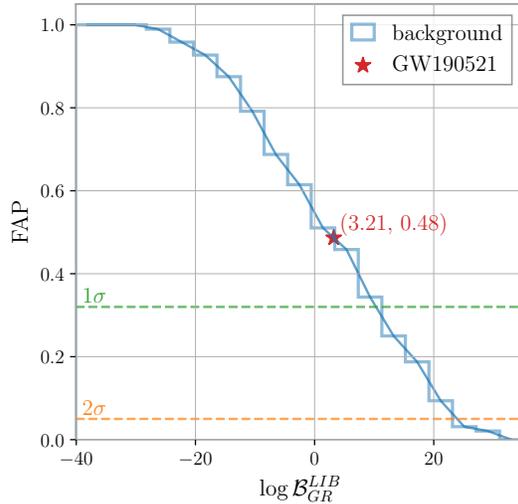
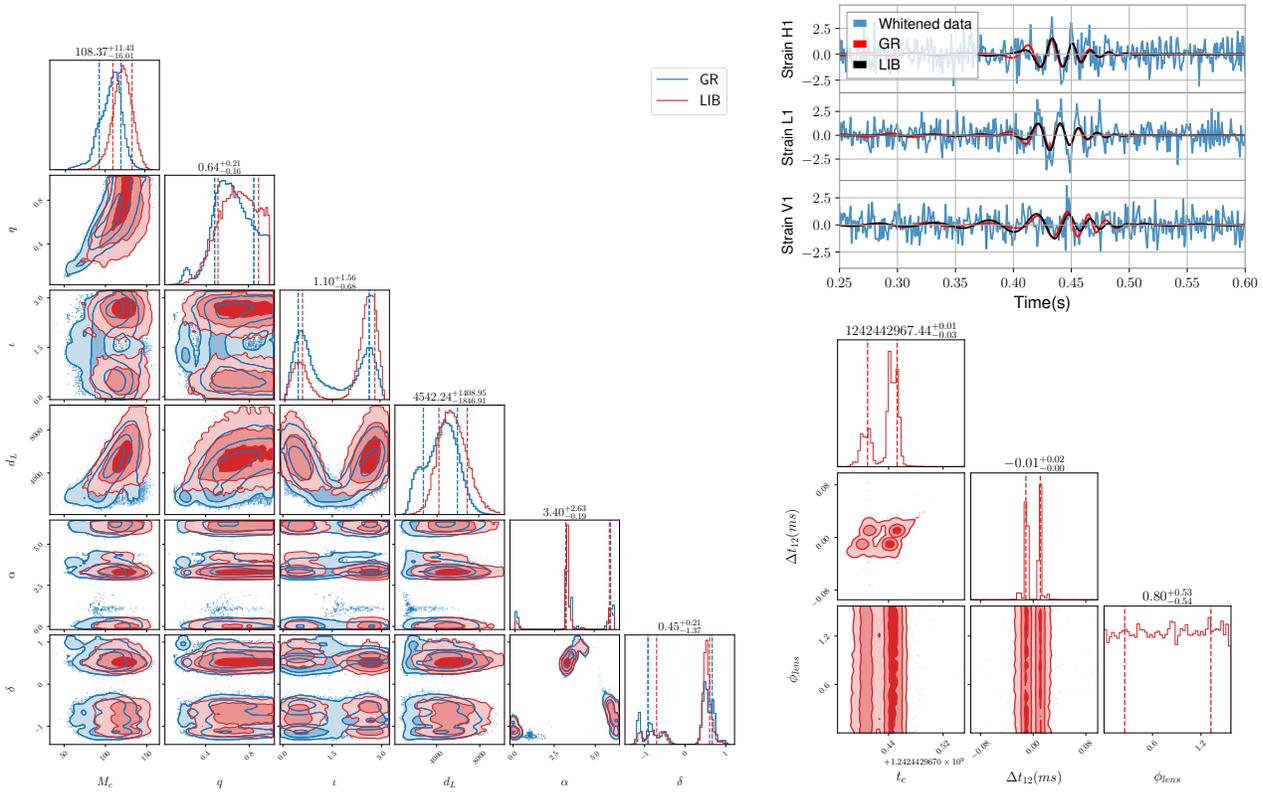


Figure 5.6: Bayes factors distributions for GW190521-like CBC, calculated by doing PE with both the hypothesis, for ~ 100 GR injections from the GW190521 posteriors in different realisations of gaussian noise. The false alarm probability for the observed $\log \mathcal{B}_{GR}^{LIB} = 3.2$ is found to be 0.48.

The Bayes factors are known to be prior dependent and its value does not signify the confidence in preferring one hypothesis over the other, but rather the preference of one hypothesis over the other given a set of prior assumptions. The model with extra parameters (LIB) could be either fitting the noise or the signal, therefore we take a fre-

quentist approach to determine the significance by considering different realisations of noise. We focus on the event with the highest Bayes factors (GW190521) and estimate its significance. We generate the background distribution of Bayes factors by injecting GR signals in Gaussian noise using the power spectral density around the trigger time. To calculate the false alarm probability corresponding to the observed Bayes Factor for the event GW190521, we simulate a hundred GR injections, whose parameters are taken from the posteriors of GW190521 event for the GR hypothesis. Fig. 5.6 shows the background distribution of the Bayes factors and the corresponding false alarm probability (FAP). The FAP corresponding to each $\mathcal{B}_{\text{GR}}^{\text{LIB}} = \kappa$ is calculated as the fraction of the background events having $\mathcal{B}_{\text{GR}}^{\text{LIB}} > \kappa$. We find that for the observed $\log \mathcal{B}_{\text{GR}}^{\text{LIB}} = 3.2$ for GW190521 is 0.48, i.e. its significance is less than 1σ .



It is to be noted that GW190521 is a remarkably loud but short (< 100 ms) signal, being easily fit by widely different hypotheses such

Figure 5.7: GW190521 ($\log \mathcal{B}_{\text{GR}}^{\text{LIB}} = 3.21$) GR v/s LIB posteriors. *MaP* (maximum a posteriori) waveforms under GR and LIB ($\Delta t_{12} = 9.51\text{ms}$, $\phi_{\text{lens}} = 0.06\text{rad}$) hypothesis with the whitened strain as observed at the LIGO-Virgo detectors.

as head on-collision of a boson star [273] or left-right (L-R), frequency-dependent birefringence [252]. In Fig. 5.7 we show the posteriors of the GW190521 event which has the highest log Bayes factor ($\ln \mathcal{B}_{\text{GR}}^{\text{LIB}} = 3.21$) from the PE runs of LIB and GR hypothesis. The two posteriors are consistent with each other with LIB favouring a slightly higher luminosity distance (d_L) and chirp mass (M_c). Additionally, posteriors under LIB are marginally narrower as compared to GR, which might be a reason for its $\ln \mathcal{B}_{\text{GR}}^{\text{LIB}} > 0$. It is worth noticing that Δt_{12} is degenerate with t_c , which is itself poorly measured due to low SNR in Virgo. We also plot the waveforms using maximum a posteriori (MaP) parameters along with the whitened time series data [274] as observed in the Hanford (H1), Livingston (L1) and Virgo (V1) detectors. It's easy to see that the signal duration is small and the two MaP waveforms are not very different from each other except for the tiny modulations in the LIB one. It can thus be concluded that the model selection favours the LIB hypothesis because it is fitting better the random noise at the detectors during the event GW190521. The other events with $\mathcal{B}_{\text{GR}}^{\text{LIB}} > 0$ are also show similar behavior and as their preference for LIB is marginal, we conclude that none of the events have any significant Bayes factor and find *no strong evidence* for birefringence.

5.4 Implications

In our analysis of the latest GW catalog, we have found that the majority of the events disfavor birefringence. For a subset of them (most notably GW190521) while the Bayesian inference prefers the LIB hypothesis, a follow-up background study indicates that most simulated GR signals give comparable Bayes factors. In the following, we present the implications of these results. First, we consider the implications for generic LIB. Then, we study the constraints on a specific scalar-tensor theory that predicts LIB. Finally, we entertain the possibility that GW190521 was emitted in an active galactic nucleus (AGN) and is displaying evidence of birefringence.

5.4.1 Constraints on generic LIB

From the non-observation of birefringence in the 43 events from GWTC-3 and using their median redshift values [16], we estimate the total optical depth for the LIB models discussed in Sec. 5.2.4. The non-observation of birefringence translates to constraints on the phenomenological model parameters, as summarized in Table 5.1. For reference, we also show the constraints obtained from the full GWTC-3 (90 events).

$\sigma_{\text{LIB}}^{1/2}$	95% c.l.	comment
$\propto M$	$R_{12} < 4.4 (2.9) \text{ kpc}$	Sec. 5.4.2
$\propto R_E$	$\alpha_E < 3.0 (1.6)$	
$\propto M^{1/2}$	$R_{12} < 20 (12) \text{ kpc}$	
$\propto M^{1/3}$	$R_{12} < 12 (6.9) \text{ kpc}$	

Table 5.1: Constraints on the phenomenological models Eq. (5.14) and Eq. (5.15), assuming no birefringence detected for analysed (all) GWTC-3 events.

The higher redshift events have higher optical depth. Non-observation of birefringence in distant sources leads to more stringent constraints, although the SNR scales with the inverse luminosity distance: hence some of the highest redshift events will not be considered because of our FAR threshold. The final results depend strongly on the model via source redshift and halo mass function. Figure 5.8 shows the redshift dependence of the optical depth for the parameterizations discussed, adopting the 95% c.l. values found by our analysis along with the observed GWTC-3 redshift distribution.

Future observations will increase in number of events and their SNRs, allowing better constrain the birefringence probabilities and ruling out more of the parameter space in the alternative theories of gravity. Higher-redshift observations above our FAR threshold will be especially valuable to constrain α_E and R_{12} for $n = 1/3, 1/2$ (see Fig. 5.8).

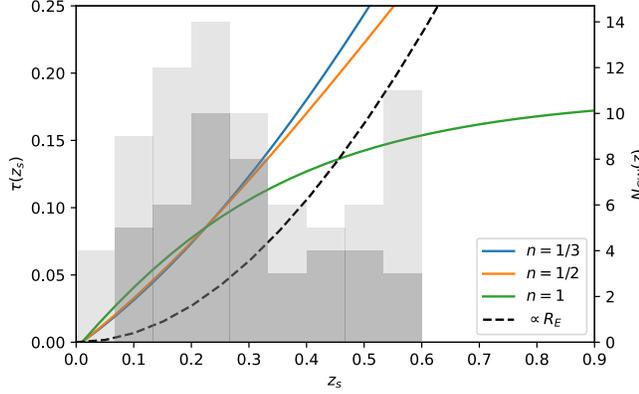


Figure 5.8: Birefringence optical depth for the phenomenological models considered here, using the parameters correspond to the 95% c.l. limit compatible with the non-observation of LIB. The dark (light) gray shaded histograms show the binned redshift distribution of analysed (all) GWTC-3 events. See Sec. 5.2.4 for details.

5.4.2 GW birefringence in Horndeski theories

Let us now use our results to a specific theory that predicts LIB. We will present the theory and translate the constraints of the phenomenological model (Table 5.1) into fundamental theory parameters. In the next subsection we will interpret a tentative detection of LIB in GW190521 as an AGN binary within the same theory. We will focus on a particular scalar-tensor theory within the Horndeski class [275], whose LIB predictions have been analyzed in detail, cf. Sec. 6 in Ref. [12]. The model is described by two parameters describing couplings between the Ricci scalar (R) and the new field ϕ : a linear coupling $p_{4\phi}$ and a derivative coupling suppressed by an energy scale Λ_4 . The Lagrangian of this theory can be written as [276, 277]

$$\mathcal{L} \sim -\frac{1}{2}(\nabla\phi)^2 + \frac{M_P^2}{2} \left(1 + \frac{p_{4\phi}\phi}{M_P}\right) R + \frac{\phi}{\Lambda_4^2} \nabla_\mu \nabla_\nu \phi G^{\mu\nu}, \quad (5.19)$$

where R is the Ricci scalar, $G_{\mu\nu}$ is the Einstein tensor, M_P is the Planck mass in units of $c = h = 1$, and ∇ the covariant derivative. The GR limit corresponds to $p_{4\phi} \rightarrow 0, \Lambda_4 \rightarrow \infty$. The parameters of this model are stringently constrained by the speed of GWs on the homogeneous FRW metric [17, 18, 19, 20] (see also [278, 279, 280]), as observed by the near-coincident arrival of GW170817 and its associated counterpart [281]: $|c_g/c - 1| \lesssim 10^{-15}$. Compliance with this limit requires

[12]

$$p_{4\phi} \lesssim 10^{-8} \Lambda_4 / H_0 \quad (\text{GW170817}). \quad (5.20)$$

While this constraint is extremely stringent, LIB allows comparable limits.

Specifying a model allows one to derive concrete predictions. The dependence of the time delay contributions (Shapiro, geometric) with the lens and theory parameters is complex. Nonetheless, we observed that the time delay decreases monotonically with the impact parameter. Moreover, its slope changes and becomes very sharp beyond the Vainshtein radius

$$r_V = 1.2 \text{Mpc} p_{4\phi}^{1/3} \left(\frac{M}{10^{12} M_\odot} \right)^{1/3} \left(\frac{H_0}{\Lambda_4} \right)^{2/3}. \quad (5.21)$$

r_V represents the scale at which the scalar field has a strong self-coupling near a massive object [266]⁴. In many scalar-tensor theories this leads to *screening*: a suppression of scalar field fluctuations for $r < r_V$, allowing the theory to approximately recover GR around massive bodies. However, screening is not necessary in this model given the stringent constraint from GW170817 (Eq. (5.20)). In this case, the strong-interaction within r_V represents a large coupling between the scalar field and the Riemann tensor, the kind of interaction producing LIB.

For simplicity, we will focus on the Shapiro time delay. The geometric time delay is usually dominant for massive halos at intermediate distances. (Fig 12 in Ref. [12]). It is proportional to the Einstein radius, and it could thus be captured generalizing Eq. 5.14 to extended lenses. Neglecting the geometric time delay is conservative but reasonable, since our constraints involve events at relatively low redshift ($z \lesssim 0.6$).

The LIB predictions have a simple dependence on the lens mass and theory parameters. We verified that $\Delta t_{12} \propto M \Lambda_4^{-4/3}$. The proportionality to the mass stems from the scaling with the Vainstein radius, as well as Δt_{12} , the impact parameter and the time spent by the GW on the region of sizeable birefringence are all $\propto r_V$. It allows us to directly connect the theory parameters to R_{12} with $n = 1$, as constrained in the

⁴For extended lenses one needs to consider the effective Vainshtein radius, such that $r_V(M(r_V^{\text{eff}})) = r_V^{\text{eff}}$, (see Eq. (186) and Fig. 14 in Ref. [12]).

phenomenological model (5.15). The scaling with Λ_4 allows us then to find R_{12} by equating $\Delta t_{12}(R_{12})$ to the constrained value for different $p_{4\phi}$, but keeping $M = 10^{12} M_\odot$, Λ_4 fixed. For simplicity, we will take a sensitivity of $\Delta t_{12} \sim 10$ ms to define R_{12} . Using the actual posteriors on Δt_{12} for each of the GWTC-3 events analyzed will not qualitatively affect these constraints in any significant manner.

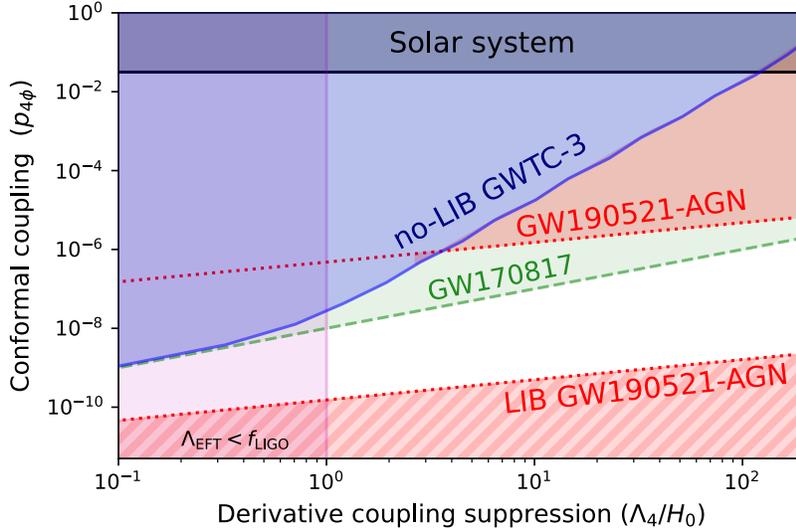


Figure 5.9: 95% c.i. constraints on the parameters of a quartic Horndeski theory [12] using the lens-induced birefringence (LIB) test. Shaded regions are excluded according to GWTC-3 (this work, blue solid), GW170817 [17, 18, 19, 20] (green dashed) and GW190521 assuming an AGN binary [21] (red dotted, see Fig. 5.10). The GR limit corresponds to $p_{4\phi} \rightarrow 0, \Lambda_4 \rightarrow \infty$, when the scalar field is decoupled from gravity and its derivative interactions suppressed. See sections 5.4.2, 5.4.3 for details. If GW190521 is associated to an AGN, the upper shaded region improves the overall GWTC-3 constraints for $\Lambda_4 \gtrsim 3H_0$. If we further assume a detection of LIB, then the bottom red shaded region excludes GR. For reference, we also indicate Solar System constraints (gray horizontal) and the region there the GW frequencies at LIGO-Virgo detectors are larger than the (non-linear) energy scale of the effective field theory (magenta vertical).

The excluded region is shown in Fig. 5.9, along with constraints from the GW speed on FRW and lunar laser ranging (no screening, $p_{4\phi} \ll 1$, see Sec. VBc in Ref. [265]). The change in slope at low Λ corresponds to a transition in which R_{12} surpasses the Vainsthein radius (5.21). For $\Lambda_4 \ll H_0$ the birefringence constraints approach those of the GW speed: this happens when r_V is so large that most GWs are effectively behind a lens. Then the constraints are satisfied in the limit $c_{\text{GW}} \rightarrow c$, equivalent to Eq. (5.20). For the sensitivity of GWTC-3, this happens for $\Lambda_4 \lesssim H_0$, where LVK frequencies lie beyond the validity of our framework as a classical effective field theory [282]. At increasing Λ_4 the constraints degrade, since probability becomes very suppressed, Eq. (5.21). For $\Lambda^4 \gtrsim 10^2 H_0$ solar system constraints become more efficient than birefringence.

5.4.3 *GW190521 as an AGN binary*

Let us now discuss the implications of a possible birefringence detection associated to GW190521. Given the constraints from the speed of GWs (5.20) on our example Horndeski theory, the chances of birefringence being caused by a lens in the line of sight are very small. We will instead interpret our result, $\Delta t_{12} \gtrsim 9.5$ ms as due to an environmental effect near the source. We will follow the scenario outlined in Ref. [21], where a candidate EM from an AGN J124942.3 + 344929, observed 34 days after the GW signal, suggests that the binary merged in the environment of a supermassive black hole (SMBH). Note that there are important uncertainties, both regarding the counterpart association (given large GW localization uncertainties [283]), and the significance of LIB detection (given our analysis of random noise realizations, Fig. 5.6). This discussion is therefore not a statement on the status of GR. Instead, it proves the potential of identifying environments of GW sources to test gravity theories.

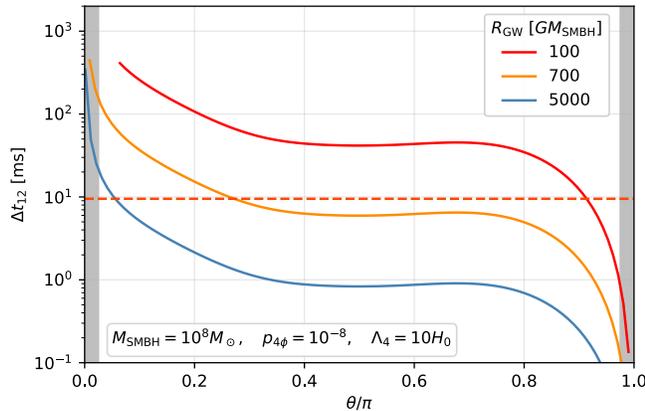


Figure 5.10: Birefringent time delay for a source near a SMBH as a function of the angle of the observer, relative to the SMBH. Each line corresponds to a different source distance, for model parameters compatible with GW170817 (see. Eq. 5.20). The horizontal line corresponds to the lower bound on $\Delta t_{12} = 9.5$ ms from the analysis of GW190521. The region between the shaded areas encompasses 95% probability for a random observer. The lowest θ represents trajectories passing at 10 Schwarzschild radii of the SMBH.

Following Ref. [21], we will assume an AGN binary scenario where the mass of the SMBH is $M_{\text{SMBH}} \sim 10^8 M_{\odot}$ and the source is located in a migration trap at $r \sim 700 GM_{\text{SMBH}}$. Then, using the framework of Ref. [12] allows us to compute the time delay as a function of the angle between the observer and the source, relative to the SMBH. The results are shown in Fig. 5.10 for $p_{4\phi} = 10^{-8}$, $\Lambda_4 = 10H_0$, compatible

GW170817 (5.20), and different distances to the SMBH (the dependence on M_{SMBH} is less pronounced, see below). The birefringent time delay becomes very large as $\theta \rightarrow 0^5$. Ultimately, the maximum time delay is limited by the existence of the horizon, $\theta_s \approx 2GM/r$. The birefringence also vanishes as $\theta \rightarrow \pi$ because of geometric cancellations in spherical symmetry.

We will translate these predictions into theory parameters and include the comparison to GW190521. We will take the values of M_{SMBH} and the source radius fixed, and consider the credible intervals as being determined by the angle θ , cf. Eq. (5.18) and Sec. 5.2.5. As we do not know the emission angle, we will assume a flat prior on the sphere $P(\theta) = \sin(\theta)$, and take the upper/lower 95% c.l. values based on $P(\theta)$ (excluding the shaded regions in Fig. 5.10). Limits on the theory parameters can be derived by noting that $\Delta t_{12}(\theta) \propto p_{4\phi}^{4/3} \Lambda_4^{-2/3} M_{\text{SMBH}}^{1/3}$ including different assumptions about the SMBH mass. Note that M_{SMBH} enters with a different scaling than the lens mass in Sec. 5.4.2, due to the source being at a fixed distance from the SMBH and within its Vainsthein radius, rather than randomly located.

The implications of GW190521 for the example theory (5.19) are shown in Fig. 5.9. The orange regions are excluded if we assume the AGN scenario as discussed above. The lower region excludes the GR limit $p_{4\phi} \rightarrow 0, \Lambda_4 \rightarrow \infty$ and relies on trusting the measured birefringence $\Delta t_{12} \gtrsim 9.5$ ms to be due to new gravitational physics. Even if the result is interpreted as noise (e.g. Fig. 5.6), assuming the AGN scenario leads to exclusion of the upper orange region (assuming sensitivity to $\Delta t_{12} \lesssim 9.5$ ms). Because of the different scaling with the theory parameters, the detection of an AGN binary becomes even more constraining than GW170817 for high Λ_4 . The beyond GR interpretation can be further probed not only by AGN events but by high-redshift multimessenger observations. In this case, the time delay between GWs and EM counterparts scales as $\approx 1\text{s} \left(10^8 p_{4\phi} \frac{H_0}{\Lambda_4}\right)^2 \frac{D}{40\text{Mpc}}$ and can be probed by distant BNS mergers.

⁵ Our calculation relies on small deviations from a straight trajectory. This assumption breaks down for small angles, where one needs to consider the geodesics of the SMBH space-time instead. However, our results are conservative since actual trajectories will bend toward the SMBH, thus increasing Δt_{12} relative to the straight propagation.

5.5 Summary and Outlook

In this chapter, we explored LIB as a test of GR using observations of GWs. LIB produces a difference in the arrival times of the GW polarisations in signals from the binary mergers, predicted by some alternatives to the GR. Using the Bayesian model selection framework, not only we can identify the signatures of birefringence, but also measure the time delay between the arrival of both polarisations (Δt_{12}). We show that this difference can be measured with high accuracy, of order few milliseconds with existing events and is likely to improve in the future following detector upgrades.

Using the latest GW catalog, GWTC-3, we find no strong evidence for the observation of the birefringence, with the highest $\log \mathcal{B}_{\text{GR}}^{\text{LIB}} = 3.21$ for the heaviest BBHs so far, GW190521. However, after simulating similar events under different noise realizations, we determine that there is a false alarm probability of 48%. This event has been associated with an AGN flare, possibly indicating that the merger occurred near an SMBH. This AGN scenario is especially favourable for the observation of LIB since the SMBH would act as a strong source of LIB. However, the AGN flare-GW association has been disputed, see e.g. [283]. Moreover, the loudness and shortness of this event makes it susceptible to different astrophysical and fundamental physics interpretations. It has also been found to be violating many tests of GR and mimicking many exotic scenarios of compact binary such as head-on-collision of a boson star [273] or left-right (L-R), frequency-dependent birefringence [252]. The latter effect is related to our flavour of LIB, with two important differences: first, L-R birefringence is defined in the basis of circularly polarized waves (left vs right, rather than + vs \times), and second, it depends on the GW frequency. Both features also appear in the Gravitational spin Hall effect in GR, although the L-R time-delay is very suppressed [250, 251].

Of the 43 analyzed events, we find that the tightest bounds on the time delay between the two polarisations is $\Delta t_{12} \sim 0.51$ ms at 90% credible intervals coming from the GW200311_115853 merger event,

while the median is $\Delta t_{12} \simeq 80$ ms. From the non-observation of LIB, we constrained the lensing optical depths in a phenomenological parameterization in which the lensing cross-section is proportional to the Einstein radius or a fixed physical radius with a power law scaling in the halo mass.

Our constraints can be translated to gravitational theories that predict LIB. As an example, we presented novel constraints on a Horndeski scalar-tensor theory featuring a new dynamical field and two free parameters. The theory is stringently constrained by the speed of GWs on the homogeneous FRW background following GW170817. Nevertheless, the lack of observed LIB places stringent bounds, which can be orders of magnitude better than Solar System tests and in some limits as tight as the GW speed bound. As a proof of principle of LIB due to a known inhomogeneity, we interpret GW190521 as an AGN binary (assuming that the signal originated in close proximity to a SMBH [21]) in terms of our example theory. Then, the large curvature is able to generate detectable LIB even when deviations from GR are minute. Our $|\Delta t_{12}| \gtrsim 9.5$ ms results would then exclude GR, placing a minimum value of the theory parameters. When interpreting this result as a fluctuation and GR to be correct, the AGN hypothesis is still able to produce very stringent bounds, that can even overcome those of the GW speed on FRW.

In future, the methods we developed here can be useful for studying new classes of events. Of particular interest will be signals where the merger is either near an SMBH or is known to have a lensed counterpart due to strong lensing. In such cases, the information about the lens may improve the constraints substantially, along the lines of the AGN-scenario we discussed. The increase in detection rate and a growing chance of strongly lensed identification makes LIB test also relevant for future runs of LVK detectors and upcoming GW detectors such as Einstein Telescope, Cosmic Explorer and LISA [284, 285, 286, 271]. Lastly, the addition of ground-based detectors such as LIGO-India and KAGRA can allow us to measure extra linear combinations of the GW polarisations and construct a null-stream [220]

to extract each of the polarisations individually. The extracted polarisations can then be used to test their consistency with GR or other theories of gravity directly.

Strongly lensed copies of GWs signals may allow us to measure additional linear combinations of the same GW polarisations and hence improve various tests of GR [287], including the one proposed here. Ultimately, developing LIB predictions for other alternative theories and generalizing the model-independent parameterizations presented here will allow our results to further test the landscape of theories beyond GR.

6 | Conclusions and future work

GW lensing is a new emerging field that has unique astrophysical implications and perhaps some are yet to be discovered. This thesis explores various themes related to the strong lensing of GWs and addresses some of the challenges in their detection while studying their implications in testing the theories of gravity.

We did a proof-of-principle study to demonstrate how ML can be used for rapid identification of strongly lensed pairs and can perform comparably to the PO method that is based on Bayesian model selection. Hence, Bayesian model selection and ML algorithms both can be used in synergy in a computationally efficient manner to classify a large number of events as lensed or unlensed. As a part of LVK collaboration, we perform both ML and PO analysis to search for strongly lensed signals in the full O₃ data. We find some interesting candidates and pass them on for further investigations like joint-PE methods that are computationally more expensive. Eventually, we find no confident detection of strong lensing in the O₃ data.

The lensed counterparts of detected GW events could also be sub-threshold, i.e. fall below the threshold used for GW detection. However, the Bayesian methods are computationally expensive and not an efficient way to analyse a large number of super-sub threshold pairs of events. ML also can be used for this purpose, which is our ongoing work. We instead developed a cheap and approximate way to analyse pairs using the match-filter-based mass estimates and reconstructed skymaps to narrow down the most interesting candidates. Using this method we find a super-sub threshold event pair, which turns out to be amongst one of the most significant candidates in whole O₃ data,

however not significant enough to claim a confident detection of strong lensing. The detection of additional sub-threshold lensed counterparts to a candidate event pair will also help in increasing the significance of a lensed event candidate to claim a confident detection. It may also allow one to localise the source to a host galaxy through cross-matching with EM catalogues and do precision cosmology [10]. In future, we would perform a study on lensing triples, their detection rates and efficiencies using the ML and PO methods.

Though challenging, we expect to have the first detection of GW strong lensing soon in the upcoming observing runs of LVK. Once detected, we showed that using the multiple copies of GWs one can probe the nature of GW polarisations much better. Though we restricted to only two lensed images, if we detect additional images they can be further used to improve the tests of GR polarisations. In future, we hope to make this test model-agnostic and extend it to more realistic polarisation models.

Strong lensing in alternative theories of gravity could also lead to birefringence, i.e. splitting the GWs into different components that travel at different speeds. To probe this phenomenon we proposed a new test of GR. Performing this test on GWTC-3 events we find no strong evidence for birefringence. While our test is theory-agnostic, we demonstrated its use to constrain the parameter space of beyond-GR theories. Moreover, for any strongly lensed events or AGN-associated events, this test will be directly relevant as we would have more information about the lens-source-observer system. Hence, in future, this test may provide more stringent or complimentary constraints as compared to other observations.

As the number of transient GW detections not only the computational costs for analysis will increase but the chances of random overlaps in measured parameters of binaries will also increase. Therefore, more robust identification methods are needed that are well-tested on mock datasets. PO and ML methods can be extended to incorporate selection effects, relative morse phase and lensing magnifications to improve their detection efficiencies.

In the case of type-2 images, which contain a (morse) phase shift of $\pi/2$ due to lensing, if there are higher order modes in the system then even from a single image we could detect lensing from the Bayesian model selection framework [288]. However, the higher modes are expected to make a significant contribution only for the inclined and asymmetric mass binary systems, and for current detectors, they are expected to be less probable.

In the 3G era, with Einstein Telescope [37] and Cosmic Explorer [38] detectors there is a good possibility of finding lensed BNS signals with EM counterparts. From such a multi-messenger lensing detection all the existing tests of GR could be improved [127, 123] and one could probe the source and lens environments more precisely. It will also help in precision cosmology.

As the GW sky localisation is poor for current GW detectors one cannot resolve the individual images in the sky and only in time. However, with multiband GW astronomy with future generation detectors like DECIGO, Einstein Telescope, and Cosmic Explorer detectors [289] we might be able to resolve the images. This may be particularly useful in identifying lensing by galaxy clusters, where, more images are formed with time delays of up to a few years. Identifying cluster-lensed GW signals is even harder as higher time delays lead to a larger number of potential lensing candidates and in turn, increase the false alarm probability [188]. Moreover, the galaxy cluster lens models are often much more complicated than galaxy lenses.

Another interesting possibility is micro-lens-induced wave-optics modulations in GW signals on top of strong lensing, for example, micro-lensing by the stars or compact objects of a strong lens galaxy [133]. The modelling of such signals is complicated and even for the simplest models the Bayesian analysis is very expensive. computationally. ML can help us in identifying such phenomena and we hope to work on this in future. Overall, the detection of lensed GWs requires development of both data analysis and waveform modelling techniques.

Appendix

GW Birefringence Injections

Here we list the injection parameters for the mismatch and the parameter estimation studies. Note that the luminosity distances are scaled as per the SNRs and hence are not mentioned in the table below. To demonstrate the correlations in the measurement of various

m_1	m_2	δ	α	ι	χ_1	χ_2	ψ	ϕ_c	t_c
38.3	33.19	-1.2	2.3	2.9	0.3	0.27	1.6	1.9	1126259462.414
24.4	2.7	-0.4	0.2	0.5	0.06	0.46	1.5	4.4	1249852256.99

parameters that may be leading to the uninformative ϕ_{lens} posteriors in our PE injection studies, we show a corner plot of the posteriors for a GW₁₅₀₉₁₄-like BBH injection with SNR = 30, $\Delta t_{12}^{\text{inj}} = 10$ ms, $\phi_{\text{lens}}^{\text{inj}} = \pi/5$ rad in Fig. 6.1.

Table 6.1: GW₁₅₀₉₁₄-like (top) and GW₁₉₀₈₁₄-like (bottom) CBC parameters used during mismatch calculations in Sec. 5.2.2 and PE injection studies in Sec. 5.3.

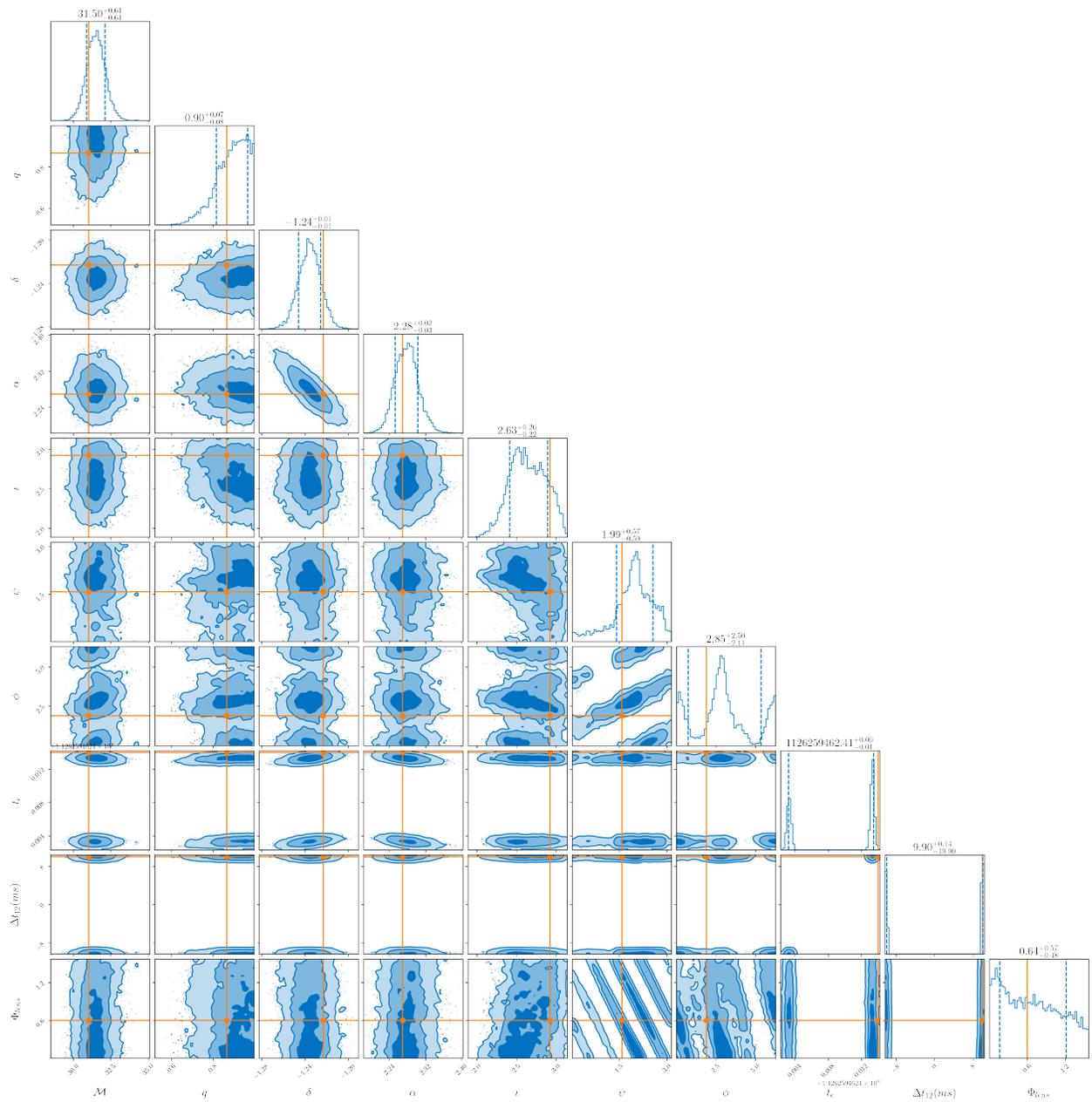


Figure 6.1: Corner plot of posteriors from LIB recovery of GW150914-like BBH injection with $\text{SNR} = 30$, $\Delta t_{12}^{\text{inj}} = 10$ ms, $\phi_{\text{lens}}^{\text{inj}} = \pi/5$ rad. The orange lines correspond to the injection parameter values. The ϕ_{lens} posteriors are visibly correlated with ψ , l and ϕ , leading to uninformative 1-D posteriors of ϕ_{lens} .

Bibliography

- [1] C. Cahillane and G. Mansell, *Review of the advanced ligo gravitational wave observatories leading to observing run four*, *Galaxies* **10** (2022) .
- [2] LIGO SCIENTIFIC, VIRGO, KAGRA collaboration, *GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run*, 2111.03606.
- [3] B. D. Metzger, *Kilonovae*, *Living Rev. Rel.* **20** (2017) 3 [1610.09381].
- [4] B. P. Abbott et al., *A gravitational-wave standard siren measurement of the Hubble constant*, *Nature* **551** (2017) 85 [1710.05835].
- [5] M. Bailes, B. K. Berger, P. R. Brady, M. Branchesi, K. Danzmann, M. Evans et al., *Gravitational-wave physics and astronomy in the 2020s and 2030s*, *Nature Reviews Physics* **3** (2021) 344.
- [6] N. Yunes, K. Yagi and F. Pretorius, *Theoretical Physics Implications of the Binary Black-Hole Mergers GW150914 and GW151226*, *Phys. Rev. D* **94** (2016) 084002 [1603.08955].
- [7] LIGO SCIENTIFIC COLLABORATION AND VIRGO COLLABORATION collaboration, *Observation of gravitational waves from a binary black hole merger*, *Phys. Rev. Lett.* **116** (2016) 061102.
- [8] Y. Tsapras, *Microlensing searches for exoplanets*, *Geosciences* **8** (2018) .
- [9] LIGO SCIENTIFIC, VIRGO collaboration, *Search for Lensing Signatures in the Gravitational-Wave Observations from the First Half of LIGO–Virgo’s Third Observing Run*, *Astrophys. J.* **923** (2021) 14 [2105.06384].
- [10] O. A. Hannuksela, K. Haris, K. K. Y. Ng, S. Kumar, A. K. Mehta, D. Keitel et al., *Search for Gravitational Lensing Signatures in LIGO–Virgo Binary Black Hole Events*, *apjl* **874** (2019) L2 [1901.02674].
- [11] P. T. Pang, R. K. Lo, I. C. Wong, T. G. Li and C. Van Den Broeck, *Generic searches for alternative gravitational wave polarizations with networks of interferometric detectors*, *Phys. Rev. D* **101** (2020) 104055 [2003.07375].

- [12] J. M. Ezquiaga and M. Zumalacárregui, *Gravitational wave lensing beyond general relativity: birefringence, echoes and shadows*, *Phys. Rev. D* **102** (2020) 124048 [2009.12187].
- [13] W. M. Farr, J. R. Gair, I. Mandel and C. Cutler, *Counting And Confusion: Bayesian Rate Estimation With Multiple Populations*, *Phys. Rev. D* **91** (2015) 023005 [1302.5341].
- [14] S. J. Kapadia et al., *A self-consistent method to estimate the rate of compact binary coalescences with a Poisson mixture model*, *Class. Quant. Grav.* **37** (2020) 045007 [1903.06881].
- [15] T. L. S. Collaboration, the Virgo Collaboration and the KAGRA Collaboration, *Search for gravitational-lensing signatures in the full third observing run of the ligo-virgo network*, Apr., 2023. 10.48550/arXiv.2304.08393.
- [16] LIGO SCIENTIFIC, VIRGO, KAGRA collaboration, *GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run*, 2111.03606.
- [17] J. M. Ezquiaga and M. Zumalacárregui, *Dark Energy After GW170817: Dead Ends and the Road Ahead*, *Phys. Rev. Lett.* **119** (2017) 251304 [1710.05901].
- [18] P. Creminelli and F. Vernizzi, *Dark Energy after GW170817 and GRB170817A*, *Phys. Rev. Lett.* **119** (2017) 251302 [1710.05877].
- [19] T. Baker, E. Bellini, P. G. Ferreira, M. Lagos, J. Noller and I. Sawicki, *Strong constraints on cosmological gravity from GW170817 and GRB 170817A*, *Phys. Rev. Lett.* **119** (2017) 251301 [1710.06394].
- [20] J. Sakstein and B. Jain, *Implications of the Neutron Star Merger GW170817 for Cosmological Scalar-Tensor Theories*, *Phys. Rev. Lett.* **119** (2017) 251303 [1710.05893].
- [21] M. J. Graham et al., *Candidate Electromagnetic Counterpart to the Binary Black Hole Merger Gravitational Wave Event S190521g*, *Phys. Rev. Lett.* **124** (2020) 251102 [2006.14122].
- [22] E. Barausse, *The physics of gravitational waves*, 3, 2023, 2303.11713.
- [23] LIGO SCIENTIFIC collaboration, *Advanced LIGO*, *Class. Quant. Grav.* **32** (2015) 074001 [1411.4547].
- [24] VIRGO collaboration, *Advanced Virgo: a second-generation interferometric gravitational wave detector*, *Class. Quant. Grav.* **32** (2015) 024001 [1408.3978].
- [25] KAGRA collaboration, *Interferometer design of the KAGRA gravitational wave detector*, *Phys. Rev.* **D88** (2013) 043007 [1306.6747].
- [26] K. L. D. and, *Status of GEO 600*, *Journal of Physics: Conference Series* **610** (2015) 012015.

- [27] B. Iyer et al., *LIGO-India Technical Report No. LIGOM1100296*, 2011.
- [28] M. Saleem, J. Rana, V. Gayathri, A. Vijaykumar, S. Goyal, S. Sachdev et al., *The science case for LIGO-india, Classical and Quantum Gravity* **39** (2021) 025004.
- [29] *Gravitational-wave data analysis*, in *Gravitational-Wave Physics and Astronomy*, pp. 269–347, John Wiley Sons, Ltd, (2011), <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9783527636037.ch7>, DOI.
- [30] B. Allen, *A χ^2 time-frequency discriminator for gravitational wave detection*, *Phys. Rev.* **D71** (2005) 062001 [gr-qc/0405045].
- [31] B. Zackay, L. Dai, T. Venumadhav, J. Roulet and M. Zaldarriaga, *Detecting Gravitational Waves With Disparate Detector Responses: Two New Binary Black Hole Mergers*, *arXiv e-prints* (2019) arXiv:1910.09528 [1910.09528].
- [32] T. Venumadhav, B. Zackay, J. Roulet, L. Dai and M. Zaldarriaga, *New binary black hole mergers in the second observing run of Advanced LIGO and Advanced Virgo*, *Phys. Rev. D* **101** (2020) 083030 [1904.07214].
- [33] B. Zackay, T. Venumadhav, L. Dai, J. Roulet and M. Zaldarriaga, *Highly spinning and aligned binary black hole merger in the Advanced LIGO first observing run*, *Phys. Rev. D* **100** (2019) 023007 [1902.10331].
- [34] S. Olsen, T. Venumadhav, J. Mushkin, J. Roulet, B. Zackay and M. Zaldarriaga, *New binary black hole mergers in the LIGO-Virgo O3a data*, *Phys. Rev. D* **106** (2022) 043009 [2201.02252].
- [35] A. H. Nitz, S. Kumar, Y.-F. Wang, S. Kastha, S. Wu, M. Schäfer et al., *4-OGC: Catalog of Gravitational Waves from Compact Binary Mergers*, *Astrophys. J.* **946** (2023) 59 [2112.06878].
- [36] VIRGO, LIGO SCIENTIFIC collaboration, *GW170817: Observation of gravitational waves from a binary neutron star inspiral*, *Phys. Rev. Lett.* **119** (2017) 161101.
- [37] S. Hild, S. Chelkowski and A. Freise, *Pushing towards the ET sensitivity using 'conventional' technology*, *arXiv e-prints* (2008) arXiv:0810.0604 [0810.0604].
- [38] D. Reitze, R. X. Adhikari, S. Ballmer, B. Barish, L. Barsotti, G. Billingsley et al., *Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO*, in *Bulletin of the American Astronomical Society*, vol. 51, p. 35, Sept., 2019, 1907.04833.
- [39] P. Amaro-Seoane, H. Audley, S. Babak, J. Baker, E. Barausse, P. Bender et al., *Laser interferometer space antenna*, 2017.
- [40] J. H. et al, *Lunar gravitational-wave antenna*, *The Astrophysical Journal* **910** (2021) 1.

- [41] S. Kawamura et al., *Current status of space gravitational wave antenna DECIGO and B-DECIGO*, *PTEP* **2021** (2021) 05A105 [2006.13545].
- [42] K. Yagi and N. Seto, *Detector configuration of decigo/bbo and identification of cosmological neutron-star binaries*, *Physical Review. D, Particles Fields* **83** (2011) .
- [43] EPTA collaboration, *The second data release from the European Pulsar Timing Array III. Search for gravitational wave signals*, **2306.16214**.
- [44] KAGRA, VIRGO, LIGO SCIENTIFIC collaboration, *Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3*, *Phys. Rev. X* **13** (2023) 011048 [2111.03634].
- [45] B. P. Abbott et al., *GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs*, *Physical Review X* **9** (2019) 031040 [1811.12907].
- [46] R. Abbott et al., *GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run*, *arXiv e-prints* (2020) arXiv:2010.14527 [2010.14527].
- [47] KAGRA, VIRGO, LIGO SCIENTIFIC collaboration, *Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3*, *Phys. Rev. X* **13** (2023) 011048 [2111.03634].
- [48] LIGO SCIENTIFIC, VIRGO, FERMI-GBM, INTEGRAL collaboration, *Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A*, *Astrophys. J. Lett.* **848** (2017) L13 [1710.05834].
- [49] B. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams et al., *Tests of general relativity with gw170817*, *Physical Review Letters* **123** (2019) .
- [50] M. Mancarella, N. Borghi, S. Foffa, E. Genoud-Prachex, F. Iacovelli, M. Maggiore et al., *Gravitational-wave cosmology with dark sirens: state of the art and perspectives for 3g detectors*, in *Proceedings of 41st International Conference on High Energy physics — PoS(ICHEP2022)*, Sissa Medialab, nov, 2022, DOI.
- [51] S. Kumar, A. Vijaykumar and A. H. Nitz, *Detecting baryon acoustic oscillations with third-generation gravitational wave observatories*, *The Astrophysical Journal* **930** (2022) 113.
- [52] G. Kashyap, N. K. Singh, K. S. Phukon, S. Caudill and P. Jain, *Dipole anisotropy in gravitational wave source distribution*, *Journal of Cosmology and Astroparticle Physics* **2023** (2023) 042.

- [53] R. Brito and P. Pani, *Black-hole superradiance: Searching for ultralight bosons with gravitational waves*, in *Handbook of Gravitational Wave Astronomy*, C. Bambi, S. Katsanevas and K. D. Kokkotas, eds., (Singapore), pp. 1377–1410, Springer Nature Singapore, (2022), DOI.
- [54] R. A. et al, *Search for subsolar-mass binaries in the first half of advanced LIGO's and advanced virgo's third observing run*, *Physical Review Letters* **129** (2022) .
- [55] *La Rivista del Nuovo Cimento* **39** (2016) 399–495.
- [56] R. Kumar, V. Dexheimer, J. Jahan, J. Noronha, J. Noronha-Hostler, C. Ratti et al., *Theoretical and experimental constraints for the equation of state of dense and hot matter*, .
- [57] LIGO SCIENTIFIC, VIRGO collaboration, *GW170817: Measurements of neutron star radii and equation of state*, *Phys. Rev. Lett.* **121** (2018) 161101 [1805.11581].
- [58] K. A. et al, *Neutron star extreme matter observatory: A kilohertz-band gravitational-wave detector in the global network*, *Publications of the Astronomical Society of Australia* **37** (2020) .
- [59] H. C. Ohanian, *On the Focusing of Gravitational Radiation*, *International Journal of Theoretical Physics* **9** (1974) 425.
- [60] S. Deguchi and W. D. Watson, *Diffraction in Gravitational Lensing for Compact Objects of Low Mass*, *ApJ* **307** (1986) 30.
- [61] Y. Wang, A. Stebbins and E. L. Turner, *Gravitational lensing of gravitational waves from merging neutron star binaries*, *Phys. Rev. Lett.* **77** (1996) 2875.
- [62] T. T. Nakamura, *Gravitational lensing of gravitational waves from inspiraling binaries by a point mass lens*, *Phys. Rev. Lett.* **80** (1998) 1138.
- [63] C. M. Will, *The Confrontation between General Relativity and Experiment*, *Living Rev. Rel.* **17** (2014) 4 [1403.7377].
- [64] A. Ghosh, N. K. Johnson-McDaniel, A. Ghosh, C. K. Mishra, P. Ajith, W. D. Pozzo et al., *Testing general relativity using gravitational wave signals from the inspiral, merger and ringdown of binary black holes*, *Classical and Quantum Gravity* **35** (2017) 014002.
- [65] K. G. Arun, B. R. Iyer, M. S. S. Qusailah and B. S. Sathyaprakash, *Probing the non-linear structure of general relativity with black hole binaries*, *Phys. Rev.* **D74** (2006) 024006 [gr-qc/0604067].
- [66] C. K. Mishra, K. G. Arun, B. R. Iyer and B. S. Sathyaprakash, *Parametrized tests of post-Newtonian theory using Advanced LIGO and Einstein Telescope*, *Phys. Rev.* **D82** (2010) 064010 [1005.0304].

- [67] T. Li, W. Del Pozzo, S. Vitale, C. Van Den Broeck, M. Agathos et al., *Towards a generic test of the strong field dynamics of general relativity using compact binary coalescence*, *Phys.Rev.* **D85** (2012) 082003 [1110.0530].
- [68] LIGO SCIENTIFIC AND VIRGO COLLABORATIONS collaboration, *Tests of general relativity with gw150914*, *Phys. Rev. Lett.* **116** (2016) 221101.
- [69] LIGO SCIENTIFIC COLLABORATION AND VIRGO COLLABORATION collaboration, *Gw151226: Observation of gravitational waves from a 22-solar-mass binary black hole coalescence*, *Phys. Rev. Lett.* **116** (2016) 241103.
- [70] LIGO SCIENTIFIC, VIRGO collaboration, *Tests of General Relativity with the Binary Black Hole Signals from the LIGO-Virgo Catalog GWTC-1*, *Phys. Rev. D* **100** (2019) 104036 [1903.04467].
- [71] N. Cornish, L. Sampson, N. Yunes and F. Pretorius, *Gravitational Wave Tests of General Relativity with the Parameterized Post-Einsteinian Framework*, *Phys. Rev. D* **84** (2011) 062003 [1105.2088].
- [72] D. Keppel and P. Ajith, *Constraining the mass of the graviton using coalescing black-hole binaries*, *Phys. Rev. D* **82** (2010) 122001 [1004.0284].
- [73] C. M. Will, *Bounding the mass of the graviton using gravitational wave observations of inspiralling compact binaries*, *Phys. Rev.* **D57** (1998) 2061 [gr-qc/9709011].
- [74] D. Alonso, E. Bellini, P. G. Ferreira and M. Zumalacárregui, *Observational future of cosmological scalar-tensor theories*, *Phys. Rev.* **D95** (2017) 063502 [1610.09290].
- [75] J. M. Ezquiaga and M. Zumalacárregui, *Dark Energy in light of Multi-Messenger Gravitational-Wave astronomy*, *Front. Astron. Space Sci.* **5** (2018) 44 [1807.09241].
- [76] S. Arai and A. Nishizawa, *Generalized framework for testing gravity with gravitational-wave propagation. ii. constraints on horndeski theory*, *Phys. Rev. D* **97** (2018) 104038.
- [77] A. Nishizawa, *Generalized framework for testing gravity with gravitational-wave propagation. i. formulation*, *Physical Review D* **97** (2018) .
- [78] M. Cabero, J. Westerweck, C. D. Capano, S. Kumar, A. B. Nielsen and B. Krishnan, *Black hole spectroscopy in the next decade*, *Phys. Rev. D* **101** (2020) 064044.
- [79] M. Isi, M. Giesler, W. M. Farr, M. A. Scheel and S. A. Teukolsky, *Testing the no-hair theorem with gw150914*, *Phys. Rev. Lett.* **123** (2019) 111102.
- [80] N. V. Krishnendu, K. G. Arun and C. K. Mishra, *Testing the binary black hole nature of a compact binary coalescence*, *Phys. Rev. Lett.* **119** (2017) 091101.

- [81] J. Abedi, H. Dykaar and N. Afshordi, *Echoes from the Abyss: Tentative evidence for Planck-scale structure at black hole horizons*, *Phys. Rev. D* **96** (2017) 082004 [1612.00266].
- [82] K. W. Tsang, M. Rollier, A. Ghosh, A. Samajdar, M. Agathos, K. Chatziioannou et al., *A morphology-independent data analysis method for detecting and characterizing gravitational wave echoes*, *Phys. Rev. D* **98** (2018) 024023 [1804.04877].
- [83] R. K. L. Lo, T. G. F. Li and A. J. Weinstein, *Template-based Gravitational-Wave Echoes Search Using Bayesian Model Selection*, *Phys. Rev. D* **99** (2019) 084052 [1811.07431].
- [84] M. Isi and A. J. Weinstein, *Probing gravitational wave polarizations with signals from compact binary coalescences*, 1710.03794.
- [85] LIGO SCIENTIFIC COLLABORATION AND VIRGO COLLABORATION collaboration, *Gw170814: A three-detector observation of gravitational waves from a binary black hole coalescence*, *Phys. Rev. Lett.* **119** (2017) 141101.
- [86] R. Abbott et al., *Tests of General Relativity with Binary Black Holes from the second LIGO-Virgo Gravitational-Wave Transient Catalog*, *arXiv e-prints* (2020) arXiv:2010.14529 [2010.14529].
- [87] LIGO SCIENTIFIC, VIRGO, KAGRA collaboration, *Tests of General Relativity with GWTC-3*, 2112.06861.
- [88] N. V. Krishnendu and F. Ohme, *Testing general relativity with gravitational waves: An overview*, *Universe* **7** (2021) 497.
- [89] S. Dodelson, *Gravitational Lensing*. Cambridge University Press, 2017.
- [90] P. Schneider, J. Ehlers and E. E. Falco, *Gravitational Lenses*. 1992, 10.1007/978-3-662-03758-4.
- [91] J. M. Ezquiaga, D. E. Holz, W. Hu, M. Lagos and R. M. Wald, *Phase effects from strong gravitational lensing of gravitational waves*, *Phys. Rev. D* **103** (2021) 064047.
- [92] M. Oguri, *Strong gravitational lensing of explosive transients*, *Reports on Progress in Physics* **82** (2019) 126901.
- [93] M. Oguri, *Effect of gravitational lensing on the distribution of gravitational waves from distant binary black hole mergers*, *Mon. Not. Roy. Astron. Soc.* **480** (2018) 3842 [1807.02584].
- [94] L. Dai, T. Venumadhav and K. Sigurdson, *Effect of lensing magnification on the apparent distribution of black hole mergers*, *Phys. Rev. D* **95** (2017) 044011 [1605.09398].
- [95] J. María Ezquiaga, D. E. Holz, W. Hu, M. Lagos and R. M. Wald, *Phase effects from strong gravitational lensing of gravitational waves*, *arXiv e-prints* (2020) arXiv:2008.12814 [2008.12814].

- [96] F. Xu, J. M. Ezquiaga and D. E. Holz, *Please repeat: Strong lensing of gravitational waves as a probe of compact binary and galaxy populations*, *arXiv e-prints* (2021) arXiv:2105.14390 [2105.14390].
- [97] A. R. A. C. Wierda, E. Wempe, O. A. Hannuksela, L. V. E. Koopmans and C. V. D. Broeck, *Beyond the detector horizon: Forecasting gravitational-wave strong lensing*, *The Astrophysical Journal* **921** (2021) 154.
- [98] K. K. Ng, K. W. Wong, T. Broadhurst and T. G. Li, *Precise LIGO Lensing Rate Predictions for Binary Black Holes*, *Phys. Rev. D* **97** (2018) 023012 [1703.06319].
- [99] G. P. Smith, M. Jauzac, J. Veitch, W. M. Farr, R. Massey and J. Richard, *What if ligo's gravitational wave detections are strongly lensed by massive galaxy clusters?*, *Monthly Notices of the Royal Astronomical Society* **475** (2018) 3823.
- [100] S.-S. Li, S. Mao, Y. Zhao and Y. Lu, *Gravitational lensing of gravitational waves: A statistical perspective*, *Mon. Not. Roy. Astron. Soc.* **476** (2018) 2220 [1802.05089].
- [101] K. Haris, A. K. Mehta, S. Kumar, T. Venumadhav and P. Ajith, *Identifying strongly lensed gravitational wave signals from binary black hole mergers*, 1807.07062.
- [102] H. W. Y. Wong, L. W. L. Chan, I. C. F. Wong, R. K. L. Lo and T. G. F. Li, *Using overlap of sky localization probability maps for filtering potentially lensed pairs of gravitational-wave signals*, 2112.05932.
- [103] S. Goyal, D. Harikrishnan, S. J. Kapadia and P. Ajith, *Rapid identification of strongly lensed gravitational-wave events with machine learning*, *Physical Review D* **104** (2021) 124057.
- [104] X. Liu, I. M. Hernandez and J. Creighton, *Identifying strong gravitational-wave lensing during the second observing run of advanced ligo and advanced virgo*, *The Astrophysical Journal* **908** (2021) 97.
- [105] J. Janquart, O. A. Hannuksela, K. Haris and C. V. D. Broeck, *Golum: A fast and precise methodology to search for, and analyze, strongly lensed gravitational-wave events*, *arXiv preprint arXiv:2203.06444* (2022).
- [106] R. K. L. Lo and I. M. Hernandez, *A bayesian statistical framework for identifying strongly-lensed gravitational-wave signals*, 2021.
- [107] C. McIsaac, D. Keitel, T. Collett, I. Harry, S. Mozzon, O. Edy et al., *Search for strongly lensed counterpart images of binary black hole mergers in the first two LIGO observing runs*, *Phys. Rev. D* **102** (2020) 084031 [1912.05389].
- [108] O. Hannuksela, K. Haris, K. Ng, S. Kumar, A. Mehta, D. Keitel et al., *Search for gravitational lensing signatures in LIGO-Virgo binary black hole events*, *Astrophys. J. Lett.* **874** (2019) L2 [1901.02674].

- [109] LIGO SCIENTIFIC, VIRGO collaboration, *Targeted subthreshold search for strongly lensed gravitational-wave events*, *Phys. Rev. D* **107** (2023) 123014 [1904.06020].
- [110] L. Dai, B. Zackay, T. Venumadhav, J. Roulet and M. Zaldarriaga, *Search for Lensed Gravitational Waves Including Morse Phase Information: An Intriguing Candidate in O2*, 2007.12709.
- [111] X. Liu, I. M. Hernandez and J. Creighton, *Identifying strong gravitational-wave lensing during the second observing run of Advanced LIGO and Advanced Virgo*, *Astrophys. J.* **908** (2021) 97 [2009.06539].
- [112] M. Caliskan, J. M. Ezquiaga, O. A. Hannuksela and D. E. Holz, *Lensing or luck? false alarm probabilities for gravitational lensing of gravitational waves*, *Phys. Rev. D* **107** (2023) 063023.
- [113] G. P. Smith, M. Bianconi, M. Jauzac, J. Richard, A. Robertson, C. P. L. Berry et al., *Deep and rapid observations of strong-lensing galaxy clusters within the sky localization of GW170814*, *MNRAS* **485** (2019) 5180 [1805.07370].
- [114] H. Yu, P. Zhang and F.-Y. Wang, *Strong lensing as a giant telescope to localize the host galaxy of gravitational wave event*, *Mon. Not. Roy. Astron. Soc.* **497** (2020) 204 [2007.00828].
- [115] E. Wempe, L. V. E. Koopmans, A. R. A. C. Wierda, O. A. Hannuksela and C. v. d. Broeck, *A lensing multi-messenger channel: Combining LIGO-Virgo-Kagra lensed gravitational-wave measurements with Euclid observations*, 2204.08732.
- [116] S. Jana, S. J. Kapadia, T. Venumadhav and P. Ajith, *Cosmography using strongly lensed gravitational waves from binary black holes*, *Phys. Rev. Lett.* **130** (2023) 261401.
- [117] M. Sereno, P. Jetzer, A. Sesana and M. Volonteri, *Cosmography with strong lensing of LISA gravitational wave sources*, *Mon. Not. Roy. Astron. Soc.* **415** (2011) 2773 [1104.1977].
- [118] K. Liao, X.-L. Fan, X.-H. Ding, M. Biesiada and Z.-H. Zhu, *Precision cosmology from future lensed gravitational wave and electromagnetic signals*, *Nature Commun.* **8** (2017) 1148 [1703.04151].
- [119] S. Cao, J. Qi, Z. Cao, M. Biesiada, J. Li, Y. Pan et al., *Direct test of the FLRW metric from strongly lensed gravitational wave observations*, *Sci. Rep.* **9** (2019) 11608 [1910.10365].
- [120] Y. Li, X. Fan and L. Gou, *Constraining Cosmological Parameters in the FLRW Metric with Lensed GW+EM Signals*, *Astrophys. J.* **873** (2019) 37 [1901.10638].
- [121] O. A. Hannuksela, T. E. Collett, M. Çalışkan and T. G. F. Li, *Localizing merging black holes with sub-arcsecond precision using gravitational-wave lensing*, *Mon. Not. Roy. Astron. Soc.* **498** (2020) 3395 [2004.13811].

- [122] B. Liu, Z. Li and Z.-H. Zhu, *Complementary constraints on dark energy equation of state from strongly lensed gravitational wave*, *Monthly Notices of the Royal Astronomical Society* **487** (2019) 1980.
- [123] X.-L. Fan, K. Liao, M. Biesiada, A. Piorkowska-Kurpas and Z.-H. Zhu, *Speed of Gravitational Waves from Strongly Lensed Gravitational Waves and Electromagnetic Signals*, *Phys. Rev. Lett.* **118** (2017) 091102 [1612.04095].
- [124] T. E. Collett and D. Bacon, *Testing the speed of gravitational waves over cosmological distances with strong gravitational lensing*, *Phys. Rev. Lett.* **118** (2017) 091101 [1602.05882].
- [125] S. Magare, S. J. Kapadia, A. More, M. K. Singh, P. Ajith and A. N. Ramprakash, *Gear-up for the Action Replay: Leveraging Lensing for Enhanced Gravitational-Wave Early-Warning*, 2302.02916.
- [126] M. K. Singh, S. J. Kapadia, S. Basak, P. Ajith and S. P. Tendulkar, *Déjà-vu et Déjà-entendu: Associating fast radio bursts with compact binary mergers via gravitational lensing*, 2304.02879.
- [127] T. Baker and M. Trodden, *Multimessenger time delays from lensed gravitational waves*, *Phys. Rev. D* **95** (2017) 063512 [1612.02004].
- [128] K.-H. Lai, O. A. Hannuksela, A. Herrera-Martín, J. M. Diego, T. Broadhurst and T. G. Li, *Discovering intermediate-mass black hole lenses through gravitational wave lensing*, *Phys. Rev. D* **98** (2018) 083005 [1801.07840].
- [129] M. Oguri and R. Takahashi, *Probing Dark Low-mass Halos and Primordial Black Holes with Frequency-dependent Gravitational Lensing Dispersions of Gravitational Waves*, *Astrophys. J.* **901** (2020) 58 [2007.01936].
- [130] S. Jung and C. S. Shin, *Gravitational-Wave Fringes at LIGO: Detecting Compact Dark Matter by Gravitational Lensing*, *Phys. Rev. Lett.* **122** (2019) 041103 [1712.01396].
- [131] R. Takahashi and T. Nakamura, *Wave Effects in the Gravitational Lensing of Gravitational Waves from Chirping Binaries*, *ApJ* **595** (2003) 1039 [astro-ph/0305055].
- [132] R. Takahashi, *Quasi-geometrical optics approximation in gravitational lensing*, *A&A* **423** (2004) 787 [astro-ph/0402165].
- [133] A. Mishra, A. K. Meena, A. More, S. Bose and J. S. Bagla, *Gravitational lensing of gravitational waves: effect of microlens population in lensing galaxies*, *mnras* **508** (2021) 4869 [2102.03946].
- [134] J. Diego, O. Hannuksela, P. Kelly, T. Broadhurst, K. Kim, T. Li et al., *Observational signatures of microlensing in gravitational waves at LIGO/Virgo frequencies*, *Astron. Astrophys.* **627** (2019) A130 [1903.04513].

- [135] S. Basak, A. Ganguly, K. Haris, S. Kapadia, A. K. Mehta and P. Ajith, *Constraints on Compact Dark Matter from Gravitational Wave Microlensing*, *Astrophys. J.* **926** (2022) L28 [2109.06456].
- [136] S. Fairhurst, *Triangulation of gravitational wave sources with a network of detectors*, *New Journal of Physics* **11** (2009) 123006 [0908.2356].
- [137] S. Fairhurst, *Localization of transient gravitational wave sources: beyond triangulation*, *Classical and Quantum Gravity* **35** (2018) 105002 [1712.04724].
- [138] L. P. Singer, D. A. Goldstein and J. S. Bloom, *The Two LIGO/Virgo Binary Black Hole Mergers on 2019 August 28 Were Not Strongly Lensed*, *arXiv e-prints* (2019) arXiv:1910.03601 [1910.03601].
- [139] S. Deguchi and W. D. Watson, *Diffraction in Gravitational Lensing for Compact Objects of Low Mass*, *The Astrophysical Journal* **307** (1986) 30.
- [140] Y. Wang, A. Stebbins and E. L. Turner, *Gravitational lensing of gravitational waves from merging neutron star binaries*, *Phys. Rev. Lett.* **77** (1996) 2875.
- [141] T. T. Nakamura, *Gravitational lensing of gravitational waves from inspiraling binaries by a point mass lens*, *Phys. Rev. Lett.* **80** (1998) 1138.
- [142] L. Dai and T. Venumadhav, *On the waveforms of gravitationally lensed gravitational waves*, *arXiv e-prints* (2017) arXiv:1702.04724 [1702.04724].
- [143] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale et al., *Parameter estimation for compact binaries with ground-based gravitational-wave observations using the lalinference software library*, *Phys. Rev. D* **91** (2015) 042003.
- [144] R. Abbott et al., *Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog*, *arXiv e-prints* (2020) arXiv:2010.14533 [2010.14533].
- [145] K. K. Y. Ng, K. W. K. Wong, T. Broadhurst and T. G. F. Li, *Precise LIGO lensing rate predictions for binary black holes*, *Phys. Rev. D* **97** (2018) 023012 [1703.06319].
- [146] S. Hild, M. Abernathy, F. Acernese, P. Amaro-Seoane, N. Andersson, K. Arun et al., *Sensitivity studies for third-generation gravitational wave observatories*, *Classical and Quantum Gravity* **28** (2011) 094013.
- [147] S. Chatterji, L. Blackburn, G. Martin and E. Katsavounidis, *Multiresolution techniques for the detection of gravitational-wave bursts*, *Classical and Quantum Gravity* **21** (2004) S1809.
- [148] L. P. Singer and L. R. Price, *Rapid bayesian position reconstruction for gravitational-wave transients*, *Phys. Rev. D* **93** (2016) 024013.

- [149] C. Cutler and É. E. Flanagan, *Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform?*, *Phys. Rev. D* **49** (1994) 2658 [gr-qc/9402014].
- [150] L. Dai, B. Zackay, T. Venumadhav, J. Roulet and M. Zaldarriaga, *Search for Lensed Gravitational Waves Including Morse Phase Information: An Intriguing Candidate in O2*, *arXiv e-prints* (2020) arXiv:2007.12709 [2007.12709].
- [151] R. Magee et al., *First Demonstration of Early Warning Gravitational-wave Alerts*, *ApJ* **910** (2021) L21 [2102.04555].
- [152] T. E. Collett, *THE POPULATION OF GALAXY-GALAXY STRONG LENSES IN FORTHCOMING OPTICAL IMAGING SURVEYS*, *The Astrophysical Journal* **811** (2015) 20.
- [153] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme et al., *Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms*, *Phys. Rev. Lett.* **113** (2014) 151101 [1308.3271].
- [154] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza et al., *Frequency-domain gravitational waves from nonprecessing black-hole binaries. i. new numerical waveforms and anatomy of the signal*, *Phys. Rev. D* **93** (2016) 044006.
- [155] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza et al., *Frequency-domain gravitational waves from nonprecessing black-hole binaries. ii. a phenomenological model for the advanced detector era*, *Phys. Rev. D* **93** (2016) 044007.
- [156] LIGO Scientific Collaboration, "LIGO Algorithm Library - LALSuite." free software (GPL), 2018. 10.7935/GT1W-FZ16.
- [157] A. Nitz, I. Harry, D. Brown, C. M. Biwer, J. Willis, T. D. Canton et al., *gwastro/pycbc: Pycbc release 1.16.4*, June, 2020. 10.5281/zenodo.3904502.
- [158] The Virgo Collaboration, *Advanced Virgo sensitivity curve study*, Tech. Rep. VIR-0073D-12, Virgo Collaboration, Nov., 2012.
- [159] *The updated advanced ligo design curve*, Tech. Rep. LIGO-T1800044-v5, LIGO Document Control Center, April, 2018.
- [160] M. R. Calabretta and B. F. Roukema, *Mapping on the HEALPix grid*, .
- [161] A. Zonca, L. Singer, D. Lenz, M. Reinecke, C. Rosset, E. Hivon et al., *healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in python*, *Journal of Open Source Software* **4** (2019) 1298.

- [162] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke et al., *HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere*, *ApJ* **622** (2005) 759 [arXiv:astro-ph/0409513].
- [163] G. Huang, Z. Liu, L. van der Maaten and K. Q. Weinberger, *Densely Connected Convolutional Networks*, *arXiv e-prints* (2016) arXiv:1608.06993 [1608.06993].
- [164] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, *arXiv e-prints* (2016) arXiv:1603.02754 [1603.02754].
- [165] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, A. M. Umar, O. U. Linus et al., *Comprehensive review of artificial neural network applications to pattern recognition*, *IEEE Access* **7** (2019) 158820.
- [166] J. Schmidhuber, *Deep Learning in Neural Networks: An Overview*, *arXiv e-prints* (2014) arXiv:1404.7828 [1404.7828].
- [167] T. Hastie, R. Tibshirani and J. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.
- [168] C. Nwankpa, W. Ijomah, A. Gachagan and S. Marshall, *Activation Functions: Comparison of trends in Practice and Research for Deep Learning*, *arXiv e-prints* (2018) arXiv:1811.03378 [1811.03378].
- [169] L. Breiman, *Bagging predictors*, *Machine learning* **24** (1996) 123.
- [170] Y. Freund and R. E. Schapire, *A short introduction to boosting*, in *In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, pp. 1401–1406, Morgan Kaufmann, 1999.
- [171] L. Breiman, *Random forests*, *Machine learning* **45** (2001) 5.
- [172] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, *Imagenet: A large-scale hierarchical image database*, in *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255, Ieee, 2009.
- [173] F. Chollet et al., “Keras.” <https://github.com/fchollet/keras>, 2015.
- [174] “Kaggle.” <https://www.kaggle.com>.
- [175] C. Nwankpa, W. Ijomah, A. Gachagan and S. Marshall, *Activation Functions: Comparison of trends in Practice and Research for Deep Learning*, *arXiv e-prints* (2018) arXiv:1811.03378 [1811.03378].
- [176] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, *arXiv e-prints* (2014) arXiv:1412.6980 [1412.6980].

- [177] J. Janquart, O. A. Hannuksela, H. K. and C. Van Den Broeck, *A fast and precise methodology to search for and analyse strongly lensed gravitational-wave events*, 2105.04536.
- [178] R. K. L. Lo and I. Magaña Hernandez, *A Bayesian statistical framework for identifying strongly-lensed gravitational-wave signals*, 2104.09339.
- [179] J. Janquart et al., *Follow-up analyses to the O3 LIGO-Virgo-KAGRA lensing searches*, 2306.03827.
- [180] K. Haris, A. K. Mehta, S. Kumar, T. Venumadhav and P. Ajith, *Identifying strongly lensed gravitational wave signals from binary black hole mergers*, 1807.07062.
- [181] VIRGO, LIGO SCIENTIFIC collaboration, *The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914*, *Astrophys. J.* **833** (2016) L1 [1602.03842].
- [182] M. Dominik, K. Belczynski, C. Fryer, D. E. Holz, E. Berti, T. Bulik et al., *Double Compact Objects II: Cosmological Merger Rates*, *Astrophys. J.* **779** (2013) 72 [1308.1546].
- [183] S. Goyal, H. D., S. J. Kapadia and P. Ajith, *Rapid Identification of Strongly Lensed Gravitational-Wave Events with Machine Learning*, 2106.12466.
- [184] G. Pratten et al., *Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes*, *Phys. Rev. D* **103** (2021) 104056 [2004.06503].
- [185] R. e. a. Abbott, *Open data from the first and second observing runs of advanced ligo and advanced virgo*, *SoftwareX* **13** (2021) 100658.
- [186] A. More and S. More, *Improved statistic to identify strongly lensed gravitational wave events*, *Mon. Not. Roy. Astron. Soc.* **515** (2022) 1044 [2111.03091].
- [187] A. R. A. C. Wierda, E. Wempe, O. A. Hannuksela, L. V. E. Koopmans and C. V. D. Broeck, *Beyond the detector horizon: Forecasting gravitational-wave strong lensing*, *The Astrophysical Journal* **921** (2021) 154.
- [188] A. R. A. C. Wierda, E. Wempe, O. A. Hannuksela, L. V. E. Koopmans and C. V. D. Broeck, *Beyond the detector horizon: Forecasting gravitational-wave strong lensing*, *The Astrophysical Journal* **921** (2021) 154.
- [189] LIGO SCIENTIFIC, VIRGO collaboration, *Targeted subthreshold search for strongly lensed gravitational-wave events*, *Phys. Rev. D* **107** (2023) 123014 [1904.06020].
- [190] C. McIsaac, D. Keitel, T. Collett, I. Harry, S. Mozzon, O. Edy et al., *Search for strongly lensed counterpart images of binary black hole mergers in the first two ligo observing runs*, *Phys. Rev. D* **102** (2020) 084031.
- [191] LIGO SCIENTIFIC, VIRGO collaboration, *GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run*, 2108.01045.

- [192] L. S. Collaboration and V. Collaboration, *GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run - Parameter Estimation Data Release*, May, 2022. 10.5281/zenodo.6513631.
- [193] L. S. Collaboration, V. Collaboration and K. Collaboration, *GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run — Parameter estimation data release*, Nov., 2021. 10.5281/zenodo.5546663.
- [194] A. Bhattacharyya, *On a measure of divergence between two multinomial populations*, *Sankhyā: The Indian Journal of Statistics (1933-1960)* **7** (1946) 401.
- [195] T. L. S. Collaboration, the Virgo Collaboration and the KAGRA Collaboration, *Open data from the third observing run of ligo, virgo, kagra and geo*, 2023.
- [196] K. Cannon, S. Caudill, C. Chan, B. Cousins, J. D. E. Creighton, B. Ewing et al., *Gstlal: A software framework for gravitational wave discovery*, 2020.
- [197] S. A. Usman et al., *The PyCBC search for gravitational waves from compact binary coalescence*, *Class. Quant. Grav.* **33** (2016) 215004 [1508.02357].
- [198] Y.-Y. Choi, C. Park and M. S. Vogeley, *Internal and Collective Properties of Galaxies in the Sloan Digital Sky Survey*, *Astrophys. J.* **658** (2007) 884 [astro-ph/0611607].
- [199] A. Nitz et al., *gwastro/pycbc: Pycbc release v1.16.8*, Aug., 2020. 10.5281/zenodo.3985815.
- [200] L. S. Collaboration, V. Collaboration and K. Collaboration, *Search for gravitational-lensing signatures in the full third observing run of the LIGO–Virgo network*, Apr., 2023. 10.5281/zenodo.7863629.
- [201] J. Luan, S. Hooper, L. Wen and Y. Chen, *Towards low-latency real-time detection of gravitational waves from compact binary coalescences in the era of advanced detectors*, *Phys. Rev. D* **85** (2012) 102002 [1108.3174].
- [202] Q. Chu et al., *SPIIR online coherent pipeline to search for gravitational waves from compact binary coalescences*, *Phys. Rev. D* **105** (2022) 024023 [2011.06787].
- [203] S. Klimenko et al., *Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors*, *Phys. Rev. D* **93** (2016) 042004 [1511.05999].
- [204] T. Islam, J. Roulet and T. Venumadhav, *Factorized Parameter Estimation for Real-Time Gravitational Wave Inference*, 2210.16278.
- [205] J. Roulet, S. Olsen, J. Mushkin, T. Islam, T. Venumadhav, B. Zackay et al., *Removing degeneracy and multimodality in gravitational wave source parameters*, *Physical Review D* **106** (2022) .

- [206] N. J. Cornish, *Rapid and robust parameter inference for binary mergers*, *Physical Review D* **103** (2021) .
- [207] M. Dax, S. R. Green, J. Gair, J. H. Macke, A. Buonanno and B. Schölkopf, *Real-time gravitational wave science with neural posterior estimation*, *Physical Review Letters* **127** (2021) .
- [208] H. Gabbard, C. Messenger, I. S. Heng, F. Tonolini and R. Murray-Smith, *Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy*, *Nature Physics* **18** (2021) 112.
- [209] L. Pathak, A. Reza and A. S. Sengupta, *Rapid reconstruction of compact binary sources using meshfree approximation*, 2210.02706.
- [210] VIRGO, LIGO SCIENTIFIC collaboration, *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral*, *Phys. Rev. Lett.* **119** (2017) 161101 [1710.05832].
- [211] LIGO SCIENTIFIC, VIRGO collaboration, *GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs*, *Phys. Rev.* **X9** (2019) 031040 [1811.12907].
- [212] T. Venumadhav, B. Zackay, J. Roulet, L. Dai and M. Zaldarriaga, *New binary black hole mergers in the second observing run of Advanced LIGO and Advanced Virgo*, *Phys. Rev. D* **101** (2020) 083030 [1904.07214].
- [213] A. H. Nitz, T. Dent, G. S. Davies, S. Kumar, C. D. Capano, I. Harry et al., *2-OGC: Open Gravitational-wave Catalog of binary mergers from analysis of public Advanced LIGO and Virgo data*, *Astrophys. J.* **891** (2019) 123 [1910.05331].
- [214] A. Schmidt-May and M. von Strauss, *Recent developments in bimetric theory*, *J. Phys. A* **49** (2016) 183001 [1512.00021].
- [215] M. Isi, M. Pitkin and A. J. Weinstein, *Probing Dynamical Gravity with the Polarization of Continuous Gravitational Waves*, *Phys. Rev. D* **96** (2017) 042001 [1703.07530].
- [216] LIGO SCIENTIFIC COLLABORATION AND VIRGO COLLABORATION collaboration, *First narrow-band search for continuous gravitational waves from known pulsars in advanced detector data*, *Phys. Rev. D* **96** (2017) 122006.
- [217] T. Callister, A. Biscoveanu, N. Christensen, M. Isi, A. Matas, O. Minazzoli et al., *Polarization-based Tests of Gravity with the Stochastic Gravitational-Wave Background*, *Phys. Rev. X* **7** (2017) 041058 [1704.08373].
- [218] LIGO SCIENTIFIC, VIRGO collaboration, *Search for Tensor, Vector, and Scalar Polarizations in the Stochastic Gravitational-Wave Background*, *Phys. Rev. Lett.* **120** (2018) 201102 [1802.10194].

- [219] A. Nishizawa, A. Taruya, K. Hayama, S. Kawamura and M.-a. Sakagami, *Probing non-tensorial polarizations of stochastic gravitational-wave backgrounds with ground-based laser interferometers*, *Phys. Rev. D* **79** (2009) 082002 [0903.0528].
- [220] K. Chatziioannou, N. Yunes and N. Cornish, *Model-Independent Test of General Relativity: An Extended post-Einsteinian Framework with Complete Polarization Content*, *Phys. Rev. D* **86** (2012) 022004 [1204.2585].
- [221] G. Smith et al., *Strong-lensing of Gravitational Waves by Galaxy Clusters*, *IAU Symp.* **338** (2017) 98 [1803.07851].
- [222] J. D. Bekenstein and R. H. Sanders, *Gravitational lenses and unconventional gravity theories*, *Astrophys. J.* **429** (1994) 480 [astro-ph/9311062].
- [223] J. M. Ezquiaga, D. E. Holz, W. Hu, M. Lagos and R. M. Wald, *Phase effects from strong gravitational lensing of gravitational waves*, 2020.
- [224] J. Skilling, *Nested sampling for general bayesian computation*, *Bayesian Anal.* **1** (2006) 833.
- [225] L. Dai and T. Venumadhav, *On the waveforms of gravitationally lensed gravitational waves*, 1702.04724.
- [226] R. Blandford and R. Narayan, *Fermat's principle, caustics, and the classification of gravitational lens images*, *Astrophys. J.* **310** (1986) 568.
- [227] L. P. Singer, D. A. Goldstein and J. S. Bloom, *The Two LIGO/Virgo Binary Black Hole Mergers on 2019 August 28 Were Not Strongly Lensed*, 1910.03601.
- [228] A. K. Mehta, *Exploring gravitational wave astrophysics Source modelling tests of general relativity and gravitational lensing*, Ph.D. thesis, Tata Institute of Fundamental Research, 2019.
- [229] The Virgo Collaboration, *Advanced Virgo Baseline Design*, Tech. Rep. VIR-0027A-09, Virgo Collaboration, Aug., 2009.
- [230] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme et al., *Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms*, *Phys. Rev. Lett.* **113** (2014) 151101 [1308.3271].
- [231] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza et al., *Frequency-domain gravitational waves from nonprecessing black-hole binaries. i. new numerical waveforms and anatomy of the signal*, *Phys. Rev. D* **93** (2016) 044006.
- [232] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza et al., *Frequency-domain gravitational waves from nonprecessing black-hole binaries. ii. a phenomenological model for the advanced detector era*, *Phys. Rev. D* **93** (2016) 044007.

- [233] J. S. Speagle, *DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences*, *MNRAS* **493** (2020) 3132 [1904.02180].
- [234] G. Ashton, M. Hübner, P. D. Lasky, C. Talbot, K. Ackley, S. Biscoveanu et al., *BILBY: A User-friendly Bayesian Inference Library for Gravitational-wave Astronomy*, *Apjs* **241** (2019) 27 [1811.02042].
- [235] J. Veitch et al., *Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library*, *Phys. Rev.* **D91** (2015) 042003 [1409.7215].
- [236] LIGO SCIENTIFIC, VIRGO collaboration, *GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run*, *Phys. Rev. X* **11** (2021) 021053 [2010.14527].
- [237] A. H. Nitz, S. Kumar, Y.-F. Wang, S. Kastha, S. Wu, M. Schäfer et al., *4-OGC: Catalog of gravitational waves from compact-binary mergers*, 2112.06878.
- [238] LIGO SCIENTIFIC, VIRGO collaboration, *Tests of general relativity with binary black holes from the second LIGO-Virgo gravitational-wave transient catalog*, *Phys. Rev. D* **103** (2021) 122002 [2010.14529].
- [239] LIGO SCIENTIFIC, VIRGO, KAGRA collaboration, *Tests of General Relativity with GWTC-3*, 2112.06861.
- [240] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.
- [241] K. Max, M. Platscher and J. Smirnov, *Gravitational Wave Oscillations in Bigravity*, *Phys. Rev. Lett.* **119** (2017) 111101 [1703.07785].
- [242] K. Max, M. Platscher and J. Smirnov, *Decoherence of Gravitational Wave Oscillations in Bigravity*, *Phys. Rev.* **D97** (2018) 064009 [1712.06601].
- [243] LISA COSMOLOGY WORKING GROUP collaboration, *Testing modified gravity at cosmological distances with LISA standard sirens*, *JCAP* **07** (2019) 024 [1906.01593].
- [244] J. B. Jiménez, J. M. Ezquiaga and L. Heisenberg, *Probing cosmological fields with gravitational wave oscillations*, *JCAP* **04** (2020) 027 [1912.06104].
- [245] J. M. Ezquiaga, W. Hu, M. Lagos and M.-X. Lin, *Gravitational wave propagation beyond general relativity: waveform distortions and echoes*, *JCAP* **11** (2021) 048 [2108.10872].
- [246] C. Dalang, P. Fleury and L. Lombriser, *Scalar and tensor gravitational waves*, *Phys. Rev. D* **103** (2021) 064075 [2009.11827].
- [247] M. Okounkova, W. M. Farr, M. Isi and L. C. Stein, *Constraining gravitational wave amplitude birefringence and Chern-Simons gravity with GWTC-2*, *Phys. Rev. D* **106** (2022) 044067 [2101.11153].

- [248] S. Vitale, S. Biscoveanu and C. Talbot, *The orientations of the binary black holes in GWTC-3*, 2204.00968.
- [249] C. de Rham and A. J. Tolley, *Speed of gravity*, *Phys. Rev. D* **101** (2020) 063518 [1909.00881].
- [250] L. Andersson, J. Joudioux, M. A. Oancea and A. Raj, *Propagation of polarized gravitational waves*, *Phys. Rev. D* **103** (2021) 044053 [2012.08363].
- [251] M. A. Oancea, R. Stiskalek and M. Zumalacárregui, *From the gates of the abyss: Frequency- and polarization-dependent lensing of gravitational waves in strong gravitational fields*, 2209.06459.
- [252] Y.-F. Wang, S. M. Brown, L. Shao and W. Zhao, *Tests of gravitational-wave birefringence with the open gravitational-wave catalog*, *Phys. Rev. D* **106** (2022) 084005 [2109.09718].
- [253] L. Haegel, K. O’Neal-Ault, Q. G. Bailey, J. D. Tasson, M. Bloom and L. Shao, *Search for anisotropic, birefringent spacetime-symmetry breaking in gravitational wave propagation from GWTC-3*, 2210.04481.
- [254] F. Bombacigno, F. Moretti, S. Boudet and G. J. Olmo, *Landau damping for gravitational waves in parity-violating theories*, *JCAP* **02** (2023) 009 [2210.07673].
- [255] G. Cusin and M. Lagos, *Gravitational wave propagation beyond geometric optics*, *Phys. Rev. D* **101** (2020) 044041 [1910.13326].
- [256] G. Pratten, C. García-Quirós, M. Colleoni, A. Ramos-Buades, H. Estellés, M. Mateu-Lucena et al., *Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes*, *Phys. Rev. D* **103** (2021) 104056 [2004.06503].
- [257] C. Mills and S. Fairhurst, *Measuring gravitational-wave higher-order multipoles*, *Phys. Rev. D* **103** (2021) 024042 [2007.04313].
- [258] S. Vitale, D. Gerosa, W. M. Farr and S. R. Taylor, *Inferring the properties of a population of compact binaries in presence of selection effects*, 2007.05579.
- [259] F. Xu, J. M. Ezquiaga and D. E. Holz, *Please repeat: Strong lensing of gravitational waves as a probe of compact binary and galaxy populations*, *The Astrophysical Journal* **929** (2022) 9.
- [260] Y. Pei, *Probability of lensing magnification by cosmologically-distributed point masses.*, *Frontiers of Astronomy in 1990s. Proc. Workshop held in Beijing* (1993) 88.
- [261] M. Zumalacárregui and U. Seljak, *Limits on stellar-mass compact objects as dark matter from gravitational lensing of type Ia supernovae*, *Phys. Rev. Lett.* **121** (2018) 141101 [1712.02240].
- [262] J. L. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. S. Warren, G. Yepes et al., *Toward a halo mass function for precision cosmology: The Limits of universality*, *Astrophys. J.* **688** (2008) 709 [0803.2706].

- [263] B. Diemer, *COLOSSUS: A python toolkit for cosmology, large-scale structure, and dark matter halos*, *Astrophys. J. Suppl.* **239** (2018) 35 [1712.04512].
- [264] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [1807.06209].
- [265] M. Zumalacárregui, *Gravity in the Era of Equality: Towards solutions to the Hubble problem without fine-tuned initial conditions*, *Phys. Rev. D* **102** (2020) 023523 [2003.06396].
- [266] A. I. Vainshtein, *To the problem of nonvanishing gravitation mass*, *Phys. Lett.* **B39** (1972) 393.
- [267] R. M. O’Leary, B. Kocsis and A. Loeb, *Gravitational waves from scattering of stellar-mass black holes in galactic nuclei*, *Mon. Not. Roy. Astron. Soc.* **395** (2009) 2127 [0807.2638].
- [268] B. Kocsis and J. Levin, *Repeated Bursts from Relativistic Scattering of Compact Objects in Galactic Nuclei*, *Phys. Rev. D* **85** (2012) 123005 [1109.4170].
- [269] D. J. D’Orazio and A. Loeb, *Repeated gravitational lensing of gravitational waves in hierarchical black hole triples*, *Phys. Rev. D* **101** (2020) 083031 [1910.02966].
- [270] R. Takahashi and T. Nakamura, *Wave effects in gravitational lensing of gravitational waves from chirping binaries*, *Astrophys. J.* **595** (2003) 1039 [astro-ph/0305055].
- [271] M. Çalışkan, L. Ji, R. Cotesta, E. Berti, M. Kamionkowski and S. Marsat, *Observability of lensing of gravitational waves from massive black hole binaries with lisa*, 2022. 10.48550/ARXIV.2206.02803.
- [272] G. Tambalo, M. Zumalacárregui, L. Dai and M. H.-Y. Cheung, *Gravitational wave lensing as a probe of halo properties and dark matter*, 2212.11960.
- [273] J. Calderón Bustillo, N. Sanchis-Gual, A. Torres-Forné, J. A. Font, A. Vajpeyi, R. Smith et al., *GW190521 as a Merger of Proca Stars: A Potential New Vector Boson of 8.7×10^{-13} eV*, *Phys. Rev. Lett.* **126** (2021) 081101 [2009.05376].
- [274] LIGO SCIENTIFIC, VIRGO collaboration, *GW190521: A Binary Black Hole Merger with a Total Mass of $150 M_{\odot}$* , *Phys. Rev. Lett.* **125** (2020) 101102 [2009.01075].
- [275] G. W. Horndeski, *Second-order scalar-tensor field equations in a four-dimensional space*, *Int. J. Theor. Phys.* **10** (1974) 363.
- [276] J. M. Ezquiaga, J. García-Bellido and M. Zumalacárregui, *Towards the most general scalar-tensor theories of gravity: a unified approach in the language of differential forms*, *Phys. Rev.* **D94** (2016) 024005 [1603.01269].

- [277] J. M. Ezquiaga, J. García-Bellido and M. Zumalacárregui, *Field redefinitions in theories beyond Einstein gravity using the language of differential forms*, *Phys. Rev.* **D95** (2017) 084039 [1701.05476].
- [278] P. Brax, C. Burrage and A.-C. Davis, *The Speed of Galileon Gravity*, *JCAP* **1603** (2016) 004 [1510.03701].
- [279] L. Lombriser and A. Taylor, *Breaking a Dark Degeneracy with Gravitational Waves*, *JCAP* **1603** (2016) 031 [1509.08458].
- [280] D. Bettoni, J. M. Ezquiaga, K. Hinterbichler and M. Zumalacárregui, *Speed of Gravitational Waves and the Fate of Scalar-Tensor Gravity*, *Phys. Rev.* **D95** (2017) 084029 [1608.01982].
- [281] VIRGO, FERMI-GBM, INTEGRAL, LIGO SCIENTIFIC collaboration, *Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A*, *Astrophys. J.* **848** (2017) L13 [1710.05834].
- [282] C. de Rham and S. Melville, *Gravitational Rainbows: LIGO and Dark Energy at its Cutoff*, *Phys. Rev. Lett.* **121** (2018) 221101 [1806.09417].
- [283] A. Palmese, M. Fishbach, C. J. Burke, J. T. Annis and X. Liu, *Do LIGO/Virgo Black Hole Mergers Produce AGN Flares? The Case of GW190521 and Prospects for Reaching a Confident Association*, *Astrophys. J. Lett.* **914** (2021) L34 [2103.16069].
- [284] V. Kalogera et al., *The Next Generation Global Gravitational Wave Observatory: The Science Book*, 2111.06990.
- [285] B. Sathyaprakash et al., *Scientific Objectives of Einstein Telescope*, *Class. Quant. Grav.* **29** (2012) 124013 [1206.0331].
- [286] X. Ding, M. Biesiada and Z.-H. Zhu, *Strongly lensed gravitational waves from intrinsically faint double compact binaries—prediction for the Einstein Telescope*, *JCAP* **12** (2015) 006 [1508.05000].
- [287] S. Goyal, K. Haris, A. K. Mehta and P. Ajith, *Testing the nature of gravitational-wave polarizations using strongly lensed signals*, *Physical Review D* **103** (2021) .
- [288] A. Vijaykumar, A. K. Mehta and A. Ganguly, *Detection and parameter estimation challenges of Type-II lensed binary black hole signals*, 2202.06334.
- [289] C. Liu, L. Shao, J. Zhao and Y. Gao, *Multiband observation of LIGO/Virgo binary black hole mergers in the gravitational-wave transient catalog GWTC-1*, *Monthly Notices of the Royal Astronomical Society* **496** (2020) 182 [<https://academic.oup.com/mnras/article-pdf/496/1/182/33385520/staa1512.pdf>].

Publications list

Publications relevant to thesis :-

1. **A rapid method for preliminary identification of subthreshold strongly lensed counterparts to superthreshold gravitational-wave events**, S Goyal, S J Kapadia, J R Cudell, A K Y Li, J C L Chan, 2023.
submitted to Physical Review D, arXiv Preprint: 2306.04397
2. **Follow-up Analyses to the O3 LIGO-Virgo-KAGRA Lensing Searches**, J Janquart, M Wright, S Goyal et al, 2023.
submitted to MNRAS, arXiv Preprint: 2306.03827
3. **Search for Gravitational Wave Lensing Signatures in LIGO/Virgo the Full O3 Data**, LVK Collaboration, 2023.
LIGO DCC : P2200031, arXiv Preprint: 2304.08393
4. **Probing Lens-induced Gravitational Wave Birefringence as a Test of General Relativity**, S Goyal, A Vijaykumar, J M Ezquiaga, M Zumalacarregi, A K Mehta, 2023.
accepted to Physical Review D, arXiv Preprint: 2301.04826
5. **Rapid Identification of Strongly Lensed Gravitational Wave Events with Machine Learning**, S Goyal, S J Kapadia, P Ajith, 2021.
Physical Review D: 104.124057 , arXiv Preprint: 2106.12466
6. **Testing the Nature of Gravitational-wave Polarizations using Strongly Lensed Sig-**

nals, S Goyal, K Haris, A K Mehta, P Ajith. 2020.

Physical Review D: 103.024038, arXiv Preprint: 2008.07060

Other publications :-

1. **The Science Case for LIGO-India**, M Saleem , J Rana, V Gayathri, A Vijaykumar, S Goyal, S Sachdev, J Suresh, S Sudhagar, A Mukherjee, G Gaur, B Sathyaprakash, A Pai, R Adhikari, P Ajith, S Bose. 2022.
Classical and Quantum Gravity: 39 025004, arXiv Preprint: 2105.01716
2. **Covid-19: Analysis of a Modified SEIR Model, a Somparison of Different Intervention Strategies and Projections for India**, A Das, A Dhar, S Goyal, A Kundu, S Pandey. 2021.
Chaos: 110595, medRxiv Preprint: 10.1101/2020.06.04.20122580
3. **Tests of General Relativity with GWTC-3** , LVK Collaboration, 2021.
arXiv Preprint: 2112.06861
4. **Search for Lensing Signatures in the Gravitational-Wave Observations from the First Half of LIGO–Virgo’s Third Observing Run**, LVK Collaboration, 2021. *Astrophys.J.* : 923 14, arXiv Preprint: 2105.06384