

The return of GOLUM: improving distributed joint parameter estimation for strongly lensed gravitational waves

Justin Janquart,^{1,2}★ K. Haris,^{1,2} Otto A. Hannuksela³ and Chris Van Den Broeck^{1,2}

¹*Nikhef – National Institute for Subatomic Physics, Science Park, NL-1098 XG Amsterdam, the Netherlands*

²*Institute for Gravitational and Subatomic Physics (GRASP), Department of Physics, Utrecht University, Princetonplein 1, NL-3584 CC Utrecht, the Netherlands*

³*Department of Physics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong*

Accepted 2023 September 7. Received 2023 September 7; in original form 2023 June 16

ABSTRACT

Owing to the forecasted improved sensitivity of ground-based gravitational-wave detectors, new research avenues will become accessible. This is the case for gravitational-wave strong lensing, predicted with a non-negligible observation rate in the coming years. However, because one needs to investigate all the event pairs in the data, searches for strongly lensed gravitational waves are often computationally heavy, and one faces high false-alarm rates. In this paper, we present upgrades made to the GOLUM software, making it more reliable while increasing its speed by re-casting the look-up table, imposing a sample control, and implementing symmetric runs on the two lensed images. We show how the recovered posteriors have improved coverage of the parameter space and how we increase the pipeline’s stability. Finally, we show the results obtained by performing a joint analysis of all the events reported until the GWTC-3 catalogue, finding similar conclusions to the ones presented in the literature.

Key words: gravitational lensing; strong – gravitational waves.

1 INTRODUCTION

Since the first detection of a binary black hole (BBH) merger through gravitational wave (GW) observation (Abbott et al. 2016), the detectors have been upgraded continually, leading to the detection of close to 100 compact binary coalescences (CBCs) reported after the third observing run (The LIGO Scientific Collaboration 2021b). These detections enabled us to study the population of compact binary mergers (The LIGO Scientific Collaboration 2023b), cosmology (The LIGO Scientific Collaboration 2023a), and test general relativity (The LIGO Scientific Collaboration 2021c). In the coming years, the LIGO (Aasi et al. 2015) and Virgo (Acernese et al. 2015) detectors should be upgraded further, improving their sensitivity even more. In addition, the KAGRA (Somiya 2012; Aso et al. 2013; Akutsu et al. 2019, 2021) detector should join the detector network soon, and LIGO-India is under construction (Iyer et al. 2011). The combination of increased detector number and improved sensitivity will open the way to new scientific avenues.

A new type of observation, forecast for the coming years (Ng et al. 2018; Oguri 2018; Wierda et al. 2021), is that of strongly lensed GWs. Strong lensing happens when the characteristic wavelength is larger than the typical size of the lens (often described by its Schwarzschild radius), which leads to multiple images with the same frequency evolution. The different images are magnified, phase shifted, and delayed in time (Ohanian 1974; Deguchi & Watson 1986; Wang, Stebbins & Turner 1996; Nakamura 1998; Takahashi & Nakamura

2003; Dai & Venumadhav 2017; Oguri 2018; Ezquiaga et al. 2021; Liu, Magaña Hernandez & Creighton 2021; Lo & Magana Hernandez 2023; Janquart et al. 2021a). If the lens is a galaxy, the images are typically separated from minutes to months (Dai, Venumadhav & Sigurdson 2017; Li et al. 2018; Ng et al. 2018; Oguri 2018; Wierda et al. 2021). If it is a galaxy cluster, the time delay can grow to years (Smith et al. 2018, 2017, 2019; Robertson et al. 2020; Ryczanowski et al. 2020). If not accounted for, lensing effects can lead to biased results. The magnification of the GW leads to a modified amplitude which can bias the luminosity distance (Dai et al. 2017; Ng et al. 2018; Pang et al. 2020). When higher order modes (HOMs) are present, the overall phase-shift can lead to biased posteriors (Dai & Venumadhav 2017; Ezquiaga et al. 2021; Janquart et al. 2021b; Vijaykumar, Mehta & Ganguly 2023). Consequently, if a lensed event has a significant HOM contribution, strong lensing can be identified more easily (Lo & Magana Hernandez 2023; Wang et al. 2021; Janquart et al. 2021b; Vijaykumar et al. 2023). Other types of lensing are also possible. If the typical size of the lens is comparable to the GW wavelength, one gets frequency-dependent beating patterns, often referred to as microlensing (Takahashi & Nakamura 2003; Cao, Li & Wang 2014; Christian, Vitale & Loeb 2018; Jit Singh et al. 2018; Lai et al. 2018; Hannuksela et al. 2019; Kim et al. 2021; Meena & Bagla 2020; Pagano, Hannuksela & Li 2020; Cheung et al. 2021; Wright & Hendry 2022).

Strongly lensed events would lead to new possibilities to probe fundamental physics. Because one gets multiple similar images, the effect is similar to an enhanced detector network made by the sum of the detectors observing each image. An extended network

* E-mail: j.janquart@uu.nl

greatly improves the sky localization for BBHs (Hannuksela et al. 2020; Janquart et al. 2021a; Lo & Magana Hernandez 2023), would allow for better probes of GW polarization (Goyal et al. 2021; Magaña Hernandez 2022), and to better probe HOMs should they be present (Janquart et al. 2021b). In addition, it would be possible to match the lensed GW data with electromagnetic (EM) data (Hannuksela et al. 2020; Wempe et al. 2022), leading to sub-arcsecond precision localization capabilities (Hannuksela et al. 2020; Wempe et al. 2022). For galaxy lenses, finding the corresponding lens–source system in the EM band generally requires the detection of four images. For galaxy cluster lenses, thanks to their scarcity, the same could be obtained using only two images (Serenio et al. 2011; Smith et al. 2017; Ryczanowski et al. 2020; Yu, Zhang & Wang 2020). Finding an EM counterpart to a GW-lensed event would also confirm the lensed nature of the event. Observing the two channels is also of interest for other tests of fundamental physics (Baker & Trodden 2017; Fan et al. 2017). The milli-second precision offered by strongly lensed events is complementary with the information given by the EM data, leading to the possibility of performing precision cosmography if a joint detection was made (Serenio et al. 2011; Liao et al. 2017; Cao et al. 2019; Li, Fan & Gou 2019). Even on its own, GW strong lensing offers interesting perspectives.

GW lensing searches are already ongoing (Hannuksela et al. 2019; Abbott et al. 2021; The LIGO Scientific Collaboration 2023c). So far, no confident detection has been reported, even if some intriguing candidates have been identified (Dai et al. 2020). The main philosophy behind strong lensing searches is identifying event pairs with the same intrinsic parameters, and originating from the same sky location. Several approaches have been developed to address various challenges. An intuitive approach, called posterior overlap, compares the posteriors obtained from unlensed searches (Haris et al. 2018). One takes the posteriors for (a subgroup of) the binaries’ parameters, makes a KDE reconstruction, and computes how (di)similar they are. For a lensed pair, one expects the two to overlap significantly. This method is relatively fast. However, overlaps can happen by chance, leading to risks of false-alarms (Wierda et al. 2021; Caliskan et al. 2022; Janquart, More & Van Den Broeck 2023). A significantly more accurate method is joint parameter estimation, where two data streams are analysed simultaneously, assuming they have the same intrinsic parameters and are linked through their relative magnification, time delay, and overall phase shift (Liu et al. 2021; Lo & Magana Hernandez 2023). A faster alternative was presented in Janquart et al. (2021a). The idea is to distribute the runs rather than doing them simultaneously, hence analysing the first image under the lensed hypothesis and using the results to analyse the second image. This leads to a faster parameter estimation scheme, equivalent under the lensed hypothesis. This approach is implemented in a framework called GOLUM (Janquart et al. 2021a; Janquart, Haris & Hannuksela 2022). In this work, we show developments made to this framework to enhance its speed and reliability.

One of the main issues faced when searching for strongly lensed GWs is the rapidly growing false alarm probability (FAP; Wierda et al. 2021; Caliskan et al. 2022; Janquart et al. 2023) related to the fast increase in the number of pairs to analyse (Ng et al. 2018). Indeed, when searching for strong lensing, one should consider all the pairs of events one can make from the single events, giving a quadratically growing number of pairs. In past analyses (Abbott et al. 2021; The LIGO Scientific Collaboration 2023c), the technique has been to analyse the data in multiple steps, going from the fastest and least accurate method (posterior overlap) to the most accurate

and computationally heavy one. However, in upcoming runs, the false alarms given by posterior overlap could lead to many events to follow up using joint parameter estimation as this analysis is quite heavy, taking up to weeks to analyse pairs of lensed BBH signals. It could become impossible to follow, requiring an enhanced intermediate step. Including a lens model in the detection statistic is one way to decrease the false-alarm risk (Haris et al. 2018; Wierda et al. 2021; Janquart et al. 2023). However, using the wrong model can mean one misses a genuinely lensed event (Janquart et al. 2023). The first way to decrease the FAP is to use more accurate methods (Janquart et al. 2023). Unfortunately, such methods are generally computationally heavier than less accurate ones, leading to issues when trying to follow up. It has been suggested in Janquart et al. (2023) that a method like GOLUM could bypass the most expensive step – the first image run – under the condition that HOMs are not too strong. Then, the analysis of the first image could be skipped, opening the door to the direct use of a more precise methodology.

In this work, we demonstrate some improvements added to the GOLUM software to address some of the challenges still present in the strong lensing data analysis. The work is structured as follows. In Section 2, we present the basics of strong lensing. In Sections 3 and 4, we present how one can perform Bayesian inference for strongly lensed GWs and how it can be tweaked to lead to a faster method. Sections 5–8 outline the different updates implemented in GOLUM. We then compare the updated pipeline with joint parameter estimation in Section 9, and show our results for the GW data released by the LVK in Section 10. Finally, we give our conclusions in Section 11.

2 STRONGLY LENSED GRAVITATIONAL WAVES

Strong lensing leads to several possible detectable images. These images are modified in three ways. First, they can be (de)magnified, translated by a magnification factor μ_j . Formally, the magnification is the inverse of the determinant of the lensing Jacobian matrix (Schneider, Ehlers & Falco 1992; Haris et al. 2018). Additionally, the image can undergo an overall phase shift (Dai & Venumadhav 2017; Ezquiaga et al. 2021), represented by a Morse factor n_j . It can only take three different values, making for three types of lensed images. One has type I images when $n_j = 0$. This happens when the image passes through the minimum of the Fermat potential and is equivalent to no phase shift. If the image passes through a saddle point, $n_j = 0.5$, and the image is of type II. Finally, if $n_j = 1$, one has a type III image, and the lensed GW passes through a maximum of the Fermat potential. When HOMs are present, the overall phase-shift induced by the Morse factor for a type II image cannot be mistaken for change in the binary’s phase value. Therefore, it can lead to the image identification and smoking-gun evidence for strong lensing (Ezquiaga et al. 2021; Janquart et al. 2021b; Vijaykumar et al. 2023). Since different images follow a different geometric path, they have an additional time delay t_j . It depends on the nature of the lens. For a galaxy lens, one expects minutes to month time delays (Takahashi & Nakamura 2003; Dai & Venumadhav 2017; Ng et al. 2018; Oguri 2018), while it can reach years for galaxy cluster lenses (Smith et al. 2018, 2017, 2019; Robertson et al. 2020; Ryczanowski et al. 2020).

Accounting for these different effects, the lensed h_L^j and unlensed h_U waveforms for an image j are linked as

$$\tilde{h}_L^j(f; \Theta, \Lambda_j) = \sqrt{\mu_j} \tilde{h}_U(f; \Theta) e^{(2i\pi f t_j - i\pi n_j \text{sign}(f))}, \quad (1)$$

where Θ represents the usual BBH parameters, $\Lambda_j = \{\mu_j, t_j, n_j\}$ represents the lensing parameters, and the tilde specifies we are working in the frequency domain.

Considering solely geometric optics,¹ the magnification and the time delay are generally not measurable on their own. For one image, the magnification can be absorbed in an observed luminosity distance

$$D_L^{\text{obs},j} = \frac{D_L}{\sqrt{\mu_j}}, \quad (2)$$

where D_L is the source luminosity distance. The time delay can also be absorbed in an observed time of coalescence

$$t_c^{\text{obs},j} = t_c + t_j, \quad (3)$$

where t_c is the unlensed time of coalescence. Formally, the Morse factor cannot be absorbed in an effective phase as the presence of HOMs would lift such a degeneracy (Ezquiaga et al. 2021; Janquart et al. 2021b; Vijaykumar et al. 2023). Due to these degeneracies, one studies the images by pairs and relates them via the relative lensing parameters $\Phi = \{\mu_{21}, t_{21}, n_{21}\}$, with

$$\begin{aligned} t_c^{\text{obs},2} &= t_c^{\text{obs},1} + t_{21}, \\ D_L^{\text{obs},2} &= \sqrt{\mu_{21}} D_L^{\text{obs},1}, \\ n_2 &= n_1 + n_{21}. \end{aligned} \quad (4)$$

With these parameters, it is easy to express the waveform for one image as a function of the waveform of another one.

An alternative possibility for the analysis is to sample directly the observed parameters (apparent luminosity distance, Morse factor, and coalescence time) for the two images. This does not account entirely for the lensing hypothesis since it does not link the observed parameters. However, this can be easier when including population models, and the correlation is added in the post-processing step when the models are added (Lo & Magana Hernandez 2023). This parametrization has been added to the GOLUM framework to incorporate selection effects in future studies.

3 BAYESIAN ANALYSIS FOR STRONGLY LENSED GRAVITATIONAL WAVES

The search for strong lensing is a Bayesian model selection problem. One has to decide whether it is more likely to observe the data under the lensed or the unlensed hypothesis. In GW lensing, one analyzes two data streams jointly to determine whether the GWs present are lensed images of each other. One of the data streams can be written

$$d_j = n_j(t) + h_H^j(\Xi_j), \quad (5)$$

where $n_j(t)$ is the noise realization for the stretch of data, and $h_H^j(\Xi_j)$ is the GW signal buried in the noise. It can either be a lensed image ($h_H^j = h_L^j$ and $\Xi_j = \{\Theta, \Lambda_j\}$) of a multiplet or an unlensed GW event ($h_H^j = h_U^j$ and $\Xi_j = \Theta_j$).

The goal is then to determine if it is more likely to be in the lensed hypothesis \mathcal{H}_L , hence the two observed GWs are images of each other described by $\{\Theta, \Lambda_1, \Lambda_2\}$, or to be in the unlensed hypothesis \mathcal{H}_U , thus the two GWs are independent and described by $\{\Theta_1, \Theta_2\}$. Typically, the comparison between the two hypotheses is done using the Odds ratio

$$\mathcal{O}_{\mathcal{H}_U}^{\mathcal{H}_L} = \frac{P(\mathcal{H}_L|d_1, d_2)}{P(\mathcal{H}_U|d_1, d_2)} = \frac{P(\mathcal{H}_L)}{P(\mathcal{H}_U)} \frac{P(d_1, d_2|\mathcal{H}_L)}{P(d_1, d_2|\mathcal{H}_U)}, \quad (6)$$

¹We do not consider inference from micro-lensing or effects such as the lifting of the mass-sheet degeneracy (Cremonese, Ezquiaga & Salzano 2021).

where we have just developed the Odds ratio's definition using Bayes theorem. $P(\mathcal{H}_L)/P(\mathcal{H}_U)$ is the prior odds and tells how likely it is to be in one of the two hypotheses before collecting any data. While it may not always be the correct thing to do (Hannuksela et al., in preparation), it often is disregarded, and only the second ratio, usually called the Bayes factor, is used to decide whether an event is lensed or not. In this work, we will use the conventions set by Lo & Magana Hernandez (2023) and Janquart et al. (2021a), and call the ratio of evidence

$$c_U^L = \frac{P(d_1, d_2|\mathcal{H}_L)}{P(d_1, d_2|\mathcal{H}_U)}, \quad (7)$$

the coherence ratio when it does not account for selection effects.² The term Bayes factor refers to the case where selection effects are included.

Under the lensed hypothesis and neglecting selection effects, the joint evidence for two data streams

$$\begin{aligned} p(d_1, d_2|\mathcal{H}_L) &= \int p(d_1|\Theta, \Lambda_1)p(d_2|\Theta, \Lambda_2) \\ &\quad \times p(\Theta, \Lambda_1, \Lambda_2)d\Theta d\Lambda_1 d\Lambda_2, \end{aligned} \quad (8)$$

where $p(\Theta, \Lambda_1, \Lambda_2)$ is the joint prior on the binary parameters and the lensing parameters, and $p(d_j|\Theta, \Lambda_j)$ with $j = \{1, 2\}$ are the individual likelihoods. When performing joint parameter estimation for lensed signals, equation (8) is the equation one seeks to solve (Dai et al. 2020; Liu et al. 2021; Lo & Magana Hernandez 2023).

Equation (8) can be recast by using the observed parameters (2), (3) and the relative lensing parameters (4)

$$\begin{aligned} p(d_1, d_2|\mathcal{H}_L) &= \int p(d_1|\vartheta)p(d_2|\vartheta, \Phi) \\ &\quad \times p(\vartheta, \Phi)d\vartheta d\Phi, \end{aligned} \quad (9)$$

where ϑ represents the binary parameters with the Morse factor and the observed luminosity distance and time delay for the first image. Φ is the set of relative lensing parameters linking the two images. Equations (8) and (9) are equivalent and represent simply a different way of expressing the waveforms and the link between them. This form has the main advantage of being easier to distribute in the GOLUM framework (see Section 4).

Under the unlensed hypothesis, the two data streams are simply uncorrelated. Therefore,

$$\begin{aligned} p(d_1, d_2|\mathcal{H}_U) &= \int p(d_1|\Theta_1)p(d_2|\Theta_2)p(\Theta_1, \Theta_2)d\Theta_1 d\Theta_2 \\ &= \int p(d_1|\Theta_1)p(\Theta_1)d\Theta_1 \int p(d_2|\Theta_2)p(\Theta_2)d\Theta_2 \\ &= p(d_1|\mathcal{H}_U)p(d_2|\mathcal{H}_U). \end{aligned} \quad (10)$$

Here, the priors are unrelated, and we directly get the product of the evidence under the unlensed hypothesis. This can, in principle, directly be obtained using the unlensed data analysis, generally performed by the LVK (The LIGO Scientific Collaboration 2021b).

Using equations (8) and (10) (or alternatively equations 9 and 10) one can express the coherence ratio (7). Evaluating the coherence ratio in this form is the task performed by joint parameter estimation tools (Dai et al. 2020; Liu et al. 2021; Lo & Magana Hernandez 2023), where one formally evaluates the joint likelihood. To make the comparison with our faster approach possible, we also added

²See Lo & Magana Hernandez (2023) for a detailed discussion on the subject. In this work, the different quantities neglect them because they can simply be added in a post-processing step.

a joint likelihood inference tool in the GOLUM package (Janquart et al. 2022).

4 GOLUM'S TRICK

Distributed joint parameter estimation makes use of the possibility to re-express the joint evidence (equation 9)³ as (Janquart et al. 2021a)

$$p(d_1, d_2 | \mathcal{H}_L) = p(d_2 | d_1, \mathcal{H}_L) p(d_1 | \mathcal{H}_L), \quad (11)$$

a product of the evidence for the first image only under the lensed hypothesis, hence accounting for the Morse factor, and the conditional evidence of the second data set given the observation of the first. Mathematically,

$$p(d_2 | d_1, \mathcal{H}_L) = \int \left[\int p(d_2 | \boldsymbol{\vartheta}, \Phi, \mathcal{H}_L) p(\boldsymbol{\vartheta} | d_1, \mathcal{H}_L) d\boldsymbol{\vartheta} \right] \times p(\Phi | \mathcal{H}_L) d\Phi. \quad (12)$$

Now, $p(\boldsymbol{\vartheta} | d_1, \mathcal{H}_L)$ is the probability of having a given observed parameter in the first image given its data. It is nothing less than the posteriors one would get by analyzing the first image under the lensed hypothesis. Equation (12) has the main advantage of leading to a faster evaluation than the full likelihood, expressing it in terms of a ‘marginalized’ likelihood

$$p(d_2 | d_1, \mathcal{H}_L) = \int L(\Phi | \mathcal{H}_L) p(\Phi | \mathcal{H}_L) d\Phi, \quad (13)$$

where

$$L(\Phi | \mathcal{H}_L) = \langle p(d_2 | \boldsymbol{\vartheta}, \Phi, \mathcal{H}_L) \rangle_{p(\boldsymbol{\vartheta} | d_1, \mathcal{H}_L)}. \quad (14)$$

While being fully equivalent to the joint likelihood, equation (13) has the advantage of having a faster evaluation. Indeed, since we use the posterior of the first image, the parameters are already sampled around the correct region of the parameter space, making for a better convergence. Moreover, these posteriors are known in advance, making it possible to generate a look-up table before the run, avoiding the computationally expensive step of waveform generation during the sampling process.

One issue at this stage is that we have not impacted the second image observation on the posteriors of the first image. This can be done by applying an additional reweighting step using (Janquart et al. 2021a)

$$p(\boldsymbol{\vartheta}, \Phi | d_1, d_2) \propto \frac{p(d_2 | \boldsymbol{\vartheta}, \Phi)}{L(\Phi)} p(\boldsymbol{\vartheta} | d_1) p(\Phi | d_1, d_2), \quad (15)$$

where only $p(d_2 | \boldsymbol{\vartheta}, \Phi)$ is unknown since the other terms are the posteriors for the first and second image runs, and the denominator is the lensed likelihood evaluated for the various samples during the second image run.

Using this method, Janquart et al. (2021a) showed the possibility of getting accurate posteriors with a run time of about 1 CPU hour for the second image. In the end, the time needed to analyze the first image – corresponding to the time needed to analyse a single unlensed BBH – is the dominating factor in the analysis. It is a major improvement compared to joint parameter estimation. However, using the GOLUM approach in various works has shown that it sometimes has shortcomings due to assumptions made in the derivation. In the coming sections, we show how to overcome some

of these. So we update GOLUM to get a pipeline with (at least) the same speed but an improved convergence and reliability.

5 MAKING GOLUM MORE EFFICIENT

The lookup table presented in Janquart et al. (2021a) gives a faster way to evaluate the likelihood for the second-image analysis. The main idea is to pre-compute the weighted inner products for each sample coming from the first image run for each detector. Then, during the nested sampling run, it suffices to correct these sums by the relative lensing parameters and put together the contribution of the different detectors. The evidence for the second event can be written as (Janquart et al. 2021a)

$$2 \ln(p(d_2 | \boldsymbol{\vartheta}, \Phi)) = \sum_{\text{ifo}} \left[\langle d_2^{\text{ifo}} | d_2^{\text{ifo}} \rangle + \frac{1}{\mu_{21}} \langle h_2^{\text{ifo}}(\boldsymbol{\vartheta}) | h_2^{\text{ifo}}(\boldsymbol{\vartheta}) \rangle - \frac{2e^{i\pi n_{21}}}{\sqrt{\mu_{21}}} \langle d_2^{\text{ifo}} | e^{2i\pi f t_{21}} h_2^{\text{ifo}}(\boldsymbol{\vartheta}) \rangle \right], \quad (16)$$

where ‘ifo’ runs over the detectors in the network, and the weighted inner product

$$\langle k | l \rangle = 4 \int_{-\infty}^{+\infty} \frac{\tilde{k}(f) \tilde{l}^*(f)}{S_n(f)}. \quad (17)$$

* means the complex conjugate and S_n is the power spectral density (PSD). This is an intuitive way of constructing the lookup table as it is a direct translation of the maths one would do. In the frequency domain, one needs to make one such table for each possible value of the Morse factor difference. So, we compute a value to store for each time sample in each detector for a given Morse factor difference and repeat for all the possible differences. Two issues were encountered with this version of the lookup table: it is memory extensive and not optimal. To reduce both of these issues, it suffices to realize the lensing parameters (except the time delay) are just prefactor of the various sums

$$2 \ln(p(d_2 | \boldsymbol{\vartheta}, \Phi)) = \sum_{\text{ifo}} \langle d_2^{\text{ifo}} | d_2^{\text{ifo}} \rangle + \frac{1}{\mu_{21}} \sum_{\text{ifo}} \langle h_2^{\text{ifo}}(\boldsymbol{\vartheta}) | h_2^{\text{ifo}}(\boldsymbol{\vartheta}) \rangle - \frac{2e^{i\pi n_{21}}}{\sqrt{\mu_{21}}} \sum_{\text{ifo}} \langle d_2^{\text{ifo}} | e^{2i\pi f t_{21}} h_2^{\text{ifo}}(\boldsymbol{\vartheta}) \rangle. \quad (18)$$

The first term in this equation is independent of the parameters as it is the inner product of the data with itself. This can be computed once and for all at the start and used as a correction factor later on. For the two other terms, instead of storing one value for each detector and each time delay, we can save the sum over the detectors directly, reducing the memory. In addition, since we do not need to perform the various sums for each sample in each iteration during the sampling process, we also reduce the number of operations and, thus, the computation time. Using this version of the lookup table and more efficient coding, we reduced the second image run by a factor of 5 on average.⁴

Finally, since t_{21} is convolved with the data, one has to compute each sum for each time-step in the data stretch. However, this makes for a large memory consumption, especially for longer duration, hence lower mass, events. Computing the sums for so many time samples is not necessary. Indeed, one typically has quite a good prior knowledge of the time delay, and the time-prior is often very restricted (typically a uniform distribution in $t_c^{\text{measured}} \pm 0.2s$ at most)

³The same developments work when parametrizing the evidence as in equation (8) but, for simplicity, we focus on one way to cast the problem. The results for the other parametrization are obtained similarly.

⁴The comparison in computational speed is done using the same *Intel(R) Core(TM) i7-9750H CPU @ 2.60 GHz* processor as in Janquart et al. (2021a).

and values outside of the priors are unexplored. Therefore, we updated the lookup table to take in the time-prior and pre-compute the different sums only for times included in the prior. The memory gain depends on the data duration, but it is a minimum factor of 10, making GOLUM more accessible for more extensive studies, lower mass systems and lower memory systems.

The combination of the two improvements also enables one to easily use a higher number of samples from the first image posterior without drastically increasing the computational time and memory requirements. In turn, it can be used to improve the analysis' accuracy.

6 SAMPLE CONTROL

When passing from the first to the second image run, it is common to sub-sample the posteriors to have a faster run. This approach can be applied without hurdle as long as the samples cover the parameter space. Afterward, each sample is used in equation (14), where it contributes to a weight in the sum. Therefore, each selected sample does not have the same final importance and contributes differently to the analysis. As for a reweighting process, we can compute the number of effective samples (Elvira, Martino & Robert 2018; Farr 2019; Payne, Talbot & Thrane 2019; Golomb & Talbot 2022) we have when drawing N samples from the first image run.

The generic expression for the number of effective samples N_{eff} for a process with weights w_j is (Elvira et al. 2018; Payne et al. 2019)

$$N_{\text{eff}} = \frac{\left(\sum_{j=1}^N w_j \right)^2}{\sum_{j=1}^N w_j^2}, \quad (19)$$

where N is the total number of samples selected.

For GOLUM,

$$w_j = p(d_2 | \boldsymbol{\vartheta}_j, \Phi). \quad (20)$$

In equation (20), the weights do not depend only on the parameters we are sub-sampling but also on the parameters we still need to infer. Therefore, we cannot directly evaluate the number of effective samples. Still, one can estimate the number of effective samples for a given number of total samples as

$$N_{\text{eff}} \simeq \frac{1}{N_\Phi} \sum_{i=1}^{N_\Phi} \left[\frac{\left(\sum_{j=1}^N p(d_2 | \boldsymbol{\vartheta}_j, \Phi_i) \right)^2}{\sum_{j=1}^N p(d_2 | \boldsymbol{\vartheta}_j, \Phi_i)^2} \right], \quad (21)$$

where we pre-select N_Φ lensing parameters from the prior and see the number of effective samples obtained on average for the different sets of prior values. Even if it provides an approximate number of effective samples, this relation is already useful to guide the number of samples one should select from the initial distributions to have a given number of effective samples in the analysis. So, by iteratively increasing N in equation (21), one can find the number of samples such that the estimated N_{eff} is bigger or equal to a target value. Typically, having about or more than 1000 effective samples is enough for a proper convergence of the likelihood. However, the corresponding number of samples to take from the posteriors of the first image is not necessarily 1000. For nicely lensed images (high SNR, well-behaved noise), taking around 1000–1500 samples is enough. However, if one wants stability when analysing unlensed events or fainter signals, the number of samples to take from the prior can grow to much more samples. Fig. 1 shows the stability as a function of the number of selected samples. Using 1000 samples leads to a likelihood close to the converged value. Using 2000 samples already leads to much better and more stable results. The first case typically has a corresponding

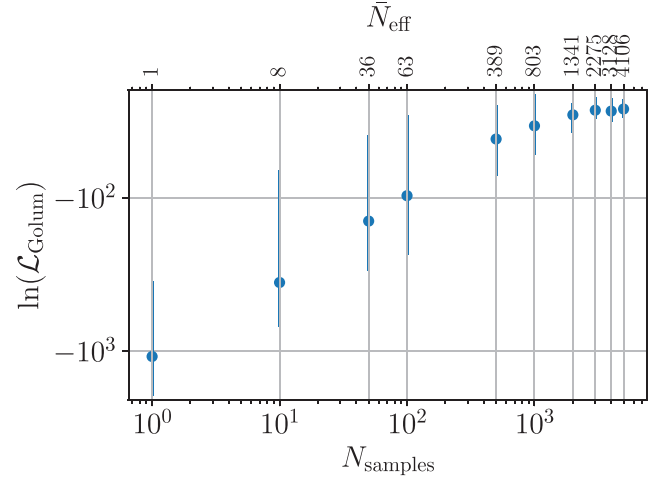


Figure 1. Evolution of the variability on the likelihood in GOLUM as a function of the number of samples and the corresponding mean number of effective samples. More samples lead to better stability from one iteration to the other. We reach stability once we are above 1000 effective samples.

effective number of samples between 700 and 900, while the second case always has more than 1000 effective samples. So, from these experiences, GOLUM gives stable evidence evaluation when using more than 1000 effective samples. For a typical lensed event pair, using 2000 samples ensures this condition.

Therefore, when analysing a large number of events, it is recommended to choose the number of effective samples one wants to use and adjust the number of sub-samples taken from the first image posterior for each pair accordingly, ensuring consistency from one analysis to the other. The option to compute the number of effective samples using equation (21) has been adapted to the software (Janquart et al. 2022).

7 ADAPTING GOLUM TO LOW-LATENCY

Using the lookup table presented in Section 5, the second image run takes $\mathcal{O}(5 \text{ min})$ for 1000 samples, and about $\mathcal{O}(20 \text{ min})$ for 5000 samples applying the same experimental conditions as in Janquart et al. (2021a). This is to compare with $\mathcal{O}(1 \text{ hr})$ for 1000 samples in GOLUM's previous implementation. Therefore, the second image run is much closer to the posterior overlap computation time, and the GOLUM run time is dominated by the first image analysis, taking the same computational time as traditional single BBH parameter estimation. However, using the distributed evidence (11), the coherence ratio (Janquart et al. 2023)

$$c_{\text{U}}^{\text{L}} = \frac{p(d_2 | d_1, \mathcal{H}_{\text{L}})}{p(d_2 | \mathcal{H}_{\text{U}})} \frac{p(d_1 | \mathcal{H}_{\text{L}})}{p(d_1 | \mathcal{H}_{\text{U}})} \approx \frac{p(d_2 | d_1, \mathcal{H}_{\text{L}})}{p(d_2 | \mathcal{H}_{\text{U}})}, \quad (22)$$

where the ratio of evidence under the lensed and the unlensed hypotheses for d_1 is approximated at one. This approximation is valid only when no strong HOM contributions are present in the data. This is generally the case for the current observation. Indeed, up to now, only two BBH events have been detected with a significant HOM contribution: GW190412 (Abbott et al. 2020a) and GW190814 (Abbott et al. 2020b), hence about one event out of 50. Moreover, it is not enough to have HOMs in the data. Their impact on the c_{U}^{L} would be significant only if one has a type II image, leading to a significant degeneracy lifting between the Morse phase and the signal's phase (Ezquiaga et al. 2021; Janquart et al. 2021b; Vijaykumar et al. 2023). This can potentially lead

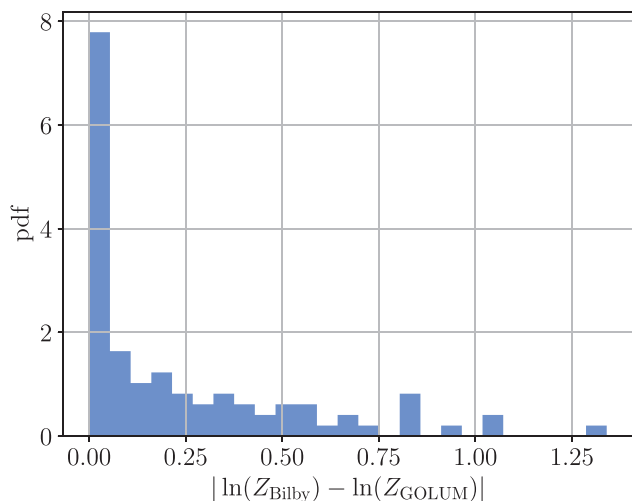


Figure 2. Difference between the log evidence recovered for type II images using BILBY and GOLUM, corresponding the unlensed and lensed hypothesis, respectively. One sees that most of the evidence is well below 1, meaning that the error made is not big enough to miss a lensed pair due to this approximation.

to the detection of the image type (Janquart et al. 2021b; Wang et al. 2021; Vijaykumar et al. 2023), and biases in the recovered posteriors (Vijaykumar et al. 2023). However, looking attentively into the details of Vijaykumar et al. (2022), the difference in evidence between lensed and unlensed events for second-generation detectors is large only for specific events, with a very low mass ratio and high total mass. Such systems are quite unlikely to be observed according to current population models (The LIGO Scientific Collaboration 2023b). As a consequence, the approximation above is generally valid, and one can neglect the evidence ratio for the first image.

We also show this in more detail ourselves. We injected 100 BBHs, selected from the observed mass and redshift distributions given in The LIGO Scientific Collaboration (2023b) and with a network signal-to-noise ratio (SNR) higher than 8 for a network of two LIGO detectors (Aasi et al. 2015) and the Virgo detector (Acernese et al. 2015) at design sensitivity. We then give an extra Morse phase corresponding to a type II image to all signals and inject them into noise. We then do the recovery under \mathcal{H}_L , hence accounting for a Morse factor, and once under \mathcal{H}_U , hence not accounting for the Morse factor. We account for HOMs and precession by injecting and recovering the signals using the *IMRPhenomXPHM* waveform (Pratten et al. 2021). Fig. 2 shows the difference in evidence under the lensed (noted Z_{GOLUM}) and unlensed (noted Z_{Bilby}) hypotheses. The unlensed run uses BILBY (Ashton et al. 2019), while the lensed runs use GOLUM. In the two cases, we use the DYNESTY sampler (Speagle 2020). All the values are distributed around zero, showing that the approximation is valid in most cases when following the expected BBH population. Out of all the events, only three have a deviation of more than one in $\ln(\mathcal{Z})$, which are events with more prominent HOMs. Looking at the coherence ratios found in other studies for lensed events (Janquart et al. 2023), one sees lensed events typically have $\ln(C_{\text{I}}^{\text{I}}) \geq 5$. Therefore, a lensed event should not be discarded with this approximation, even if it has some non-negligible HOM contribution. This was already hinted at by Janquart et al. (2023). We also note that outliers are possible and HOMs should formally be accounted for. Even if they would probably not modify the coherence ratio too much, one should treat them cautiously. Checking the results using the first image run would be helpful in

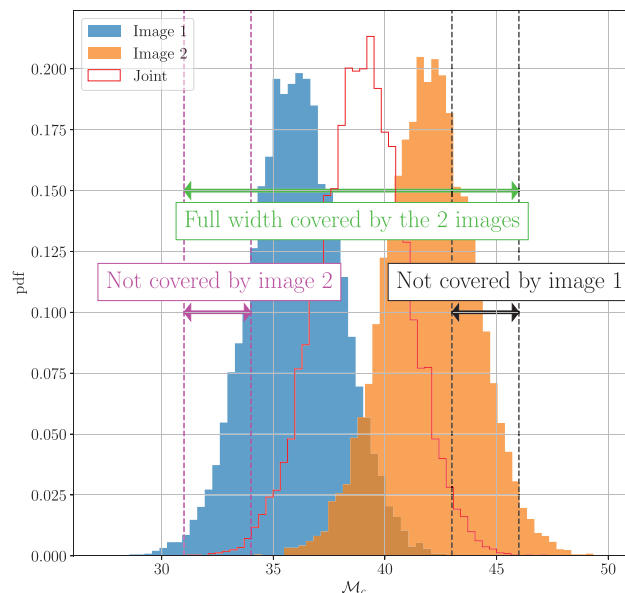


Figure 3. Illustration of the coverage of the joint parameter space for fiducial chirp mass recoveries. We straightforwardly use Gaussians to show the effect. Sampling only one of the two images is not enough to fully cover the joint results. However, using the two images works, motivating the use of a symmetric version of GOLUM.

such a case. Still, to follow the pace of the detections in low-latency, neglecting the first image can be a valuable trick.

Posterior overlap (Haris et al. 2018) also assumes that the posteriors obtained through the unlensed analyses are unbiased and representative of the ones obtained under the lensed hypothesis. An event breaking the approximation made in equation (22) would also break the approximation made in the posterior overlap analysis. Therefore, our suggested method is not necessarily more keen on issues when confronted with HOMs. On the other hand, our framework still accounts for more correlation between the events since we do not have to perform KDE reconstruction or take a sub-part of the parameter space. Therefore, using this fast GOLUM should reduce the number of false alarms and, thus, the number of events that need to be followed up by more complete joint parameter estimation methods.

We also note it should be possible to convert the unlensed results into lensed ones using hybrid sampling (Wolfe, Talbot & Golomb 2023), making it possible to change the dimensionality of the problem and rapidly correct for the strong-lensing effects. This could make it possible to use the exact GOLUM framework in low-latency. This is left for future work.

8 SYMMETRIC GOLUM

One remaining possible issue is the parameter space coverage. In principle, one expects the posteriors for the two events to match perfectly. However, in reality, noise effects, the number of online detectors, etc. can lead to variations in the posteriors with changes in the widths and shifts between the posteriors. Because of that, GOLUM can sometimes give slightly different posteriors than a joint parameter estimation tool. This does not mean we miss the correct value, but more that we can get overconfident in the final posteriors and miss a part of the lower significance region. An illustration of this phenomenon is given in Fig. 3, where we illustrate the different

posteriors as simple Gaussians. When considering only one of the two images, the final samples can miss a small part of the region obtained for the joint analysis. Once we have acquired the posteriors for the first image, GOLUM uses them to probe the lensing parameters, but they are not re-sampled. Therefore, we cannot create new samples, and regions not covered by the samples taken for the first image are not considered in the joint distribution. However, it can be the case that the joint parameter estimation extends to a region compatible with the second image but not covered by the first image.⁵

So, while in theory, the labels 1 and 2 in equation (11) should not matter, they do have a consequence on the final result due to the difference in observation condition. However, one can easily show that

$$\begin{aligned} p(d_1, d_2 | \mathcal{H}_L) &= p(d_2 | d_1, \mathcal{H}_L) p(d_1 | \mathcal{H}_L) \\ &= p(d_1 | d_2, \mathcal{H}_L) p(d_2 | \mathcal{H}_L) \\ &= \frac{1}{2} \left(p(d_2 | d_1, \mathcal{H}_L) p(d_1 | \mathcal{H}_L) \right. \\ &\quad \left. + p(d_1 | d_2, \mathcal{H}_L) p(d_2 | \mathcal{H}_L) \right), \end{aligned} \quad (23)$$

where the conditional likelihood is re-expressed simply as the conditional of d_1 given d_2 . Equation (23) is symmetric in the two events. Therefore, the parameter space is explored by the two events. For any of the two terms, the conditional evidence is computed using equation (13), once using the samples from image 1 to analyse image 2, and once using samples from image 2 to analyse image 1.⁶ So, this conditioned likelihood now requires two usual BBH runs (accounting for the Morse factor) and two GOLUM runs. Since the runs can be done in parallel, the total time of the run is the time needed to analyse a single BBH. In this case, the coherence ratio takes the more complex form

$$\begin{aligned} c_U^L &= \frac{1}{2} \frac{p(d_2 | d_1, \mathcal{H}_L) p(d_1 | \mathcal{H}_L) + p(d_1 | d_2, \mathcal{H}_L) p(d_2 | \mathcal{H}_L)}{p(d_1 | \mathcal{H}_U) p(d_2 | \mathcal{H}_U)} \\ &= \frac{1}{2} \left[\frac{p(d_2 | d_1, \mathcal{H}_L)}{p(d_2 | \mathcal{H}_U)} \frac{p(d_1 | \mathcal{H}_L)}{p(d_1 | \mathcal{H}_U)} + \frac{p(d_1 | d_2, \mathcal{H}_L)}{p(d_1 | \mathcal{H}_U)} \frac{p(d_2 | \mathcal{H}_L)}{p(d_2 | \mathcal{H}_U)} \right] \\ &\simeq \frac{1}{2} \left[\frac{p(d_2 | d_1, \mathcal{H}_L)}{p(d_2 | \mathcal{H}_U)} + \frac{p(d_1 | d_2, \mathcal{H}_L)}{p(d_1 | \mathcal{H}_U)} \right], \end{aligned} \quad (24)$$

where the last line assumes a low HOM hypothesis, as detailed in Section 7. In this case, only two GOLUM runs are needed, leading to something close to joint parameter estimation in less than an hour.

In addition to having evidence better accounting for the parameter space covered by the two images, one is generally also interested in the posteriors from the joint analysis. The posteriors can be expressed as

$$\begin{aligned} p(\boldsymbol{\vartheta}, \Phi | d_1, d_2) &= p(\boldsymbol{\vartheta} | d_1, d_2) p(\Phi | d_1, d_2) \\ &= \frac{1}{2} \left(p(\boldsymbol{\vartheta} | d_1, d_2) p(\Phi | d_1, d_2) \right. \\ &\quad \left. + p(\boldsymbol{\vartheta} | d_1, d_2) p(\Phi | d_1, d_2) \right). \end{aligned} \quad (25)$$

In this expression, we write $(\boldsymbol{\vartheta}, \Phi)$ without indices as $(\boldsymbol{\vartheta}_1, \Phi_{21})$ and $(\boldsymbol{\vartheta}_2, \Phi_{12})$ are related by a simple deterministic operation, meaning their probability distributions are trivially related. Each $p(\boldsymbol{\vartheta} | d_1, d_2)$ in equation (25) can be expressed using equation (15). In our context,

one of the terms is computed using $p(\boldsymbol{\vartheta}_1 | d_1)$ and $p(d_2 | \boldsymbol{\vartheta}_1, \Phi_{21})$ and the other is obtained by swapping images 1 and 2 in the computation. Hence, we account for $\boldsymbol{\vartheta}$ samples coming from the two posteriors, and we more largely cover the parameter space. This avoids us from missing a part of the parameter space because of narrower priors in the first image analysis.

This process pushes GOLUM closer to formal joint parameter estimation while not significantly increasing the computational time, especially when accounting for the improvements detailed in Section 5.

9 CONFRONTATION WITH JOINT PARAMETER ESTIMATION

In this section, we compare the results from the GOLUM analyses using the above-mentioned updates. We require 2000 effective samples. The 15 analysed BBHs are generated following the rates and population models given in The LIGO Scientific Collaboration (2023b). The lensing parameters linking the two events are sampled from the expected distribution for a singular isothermal ellipsoid lens model (More & More 2022). We use a uniform prior between 5 and 100 M_\odot for the chirp mass ($\mathcal{M}_c = (m_1 m_2)^{3/5} / (m_1 + m_2)^{1/5}$) and between 0.1 and 1 for the mass ratio ($q = m_2 / m_1$). The individual masses are constrained between 1 and 1000 M_\odot . For the other parameters, we use the usual priors. When considering individual Morse factors, we use a discrete prior on the three possible values, while for the Morse factor difference, we consider $n_{21} \in \{0, 0.5, 1, 1.5\}$. The prior on the relative magnification is uniform between 0.1 and 50, while the prior on the time delay is $\mathcal{U}(t_{21}^{\text{inj}} - 0.2, t_{21}^{\text{inj}} + 0.2)$. The events are injected in a network made of the two LIGO detectors and the Virgo detector at design sensitivity. We require the network SNR to be higher than 8 for the two images⁷ and perform the analysis. The unlensed runs are done using BILBY (Ashton et al. 2019) with the DYNesty sampler (Speagle 2020). The joint parameter inference is performed using the framework added into GOLUM (Janquart et al. 2022), and the distributed runs use the updated version of GOLUM explained above, using the same sampler. Again, the injections and analyses are performed using the *IMRPhenomXPHM* waveform (Pratten et al. 2021).

First, in Fig. 4, we compare the difference in likelihood for the symmetric GOLUM framework and the joint parameter framework and the old GOLUM implementation. With the new design, we obtain a difference in evidence between joint parameter estimation and GOLUM below the typical error on the evidence made by GOLUM.⁸ So, there is a real improvement in the stability of the framework obtained through symmetrization. Looking at the results for the old implementation, one sees it also performs relatively well, and can be an effective filter. We also note that, in the first approximation, one should consider as a lensed candidate all events with a coherence ratio superior to a threshold value [typically, 2 is a conservative value (Janquart et al. 2023)]. Even with the largest difference, the GOLUM approximate version using only one image is not discarding any lensed event. So, it can be used as a filtering algorithm and is indicative when studying a large-scale population.

A second problem tackled by the use of symmetrization is posterior coverage. The main idea is that it enables one to cover the parameter

⁵The effect can also be observed in fig. 9 of Lo & Magana Hernandez (2023).

⁶Images 1 and 2 here are just arbitrary labels to make the difference between them.

⁷Application of GOLUM to sub-threshold pairs was illustrated in Janquart et al. (2021a).

⁸Since we sum contributions from different runs, there is an error cumulation in GOLUM leading larger uncertainties on the final value.

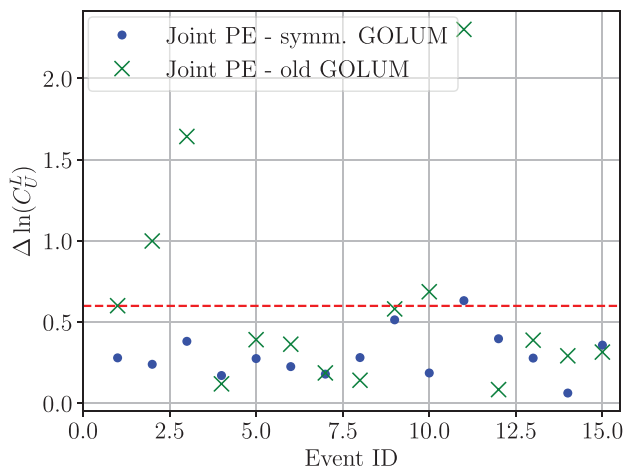


Figure 4. Comparison between the error on the evidence obtained through the symmetric GOLUM implementation and the old implementation. The 1-image implementation already performs sensibly well. However, there is an improvement when considering the symmetric version, where nearly all the differences are below the evidence error present in GOLUM.

space better than when using only one image. One of the issues found with the usual GOLUM is that it sometimes does not fully cover the parameter space. Jointly, the sampling process has inherent issues, such as vanishing samples. The symmetric implementation naturally addresses these issues. As represented in Fig. 3, it helps in covering the entire parameter space. Secondly, since it uses two GOLUM runs, more independent samples are directly available, and one reduces the issue of vanishing samples. Therefore, we get a better match between the symmetric GOLUM posteriors and those coming from joint parameter estimation. This is shown in Fig. 5, top panel, where we represent the sky location recovery. Another issue that can arise, and only partially solved with symmetrization, is the domination of a reduced number of samples in the reweighting process. This can happen when a couple of data points have weights $p(d_i|\vartheta, \Phi)/p(d_j|\Phi)$ significantly higher than the others. In such a case, the selection process latches on them, and one gets scattered posteriors with biased dominant regions, leading to high-density points. This is less likely to happen when considering more samples, which is why symmetrization is helpful. However, in some cases, scattering is still possible. The dominant value may be slightly away from the actual injected ones for some parameters. Therefore, the injected value is not in the highest density region. The posteriors are then more different between joint parameter estimation and GOLUM. Still, we note that, with symmetrization, we have not seen cases where the injected value had no samples when the joint parameter estimation had. Several possibilities to correct this exist. First, one could simply use more samples. However, covering the full space requires testing millions of combinations, which becomes computationally expensive. On the other hand, since there are samples in the correct region, one can think about using importance sampling using the formal joint parameter estimation likelihood. An example of a case where the biased higher density region leads to different posteriors between joint parameter estimation and GOLUM is given in Fig. 5, bottom panel. While the injected value is not in the 90 per cent confidence interval, it is in the 95 per cent one. The scattered posteriors problem was also present in the old GOLUM framework and was more pro-eminent. So, even if it does not completely solve the issue, symmetrization helps in significantly reducing it.

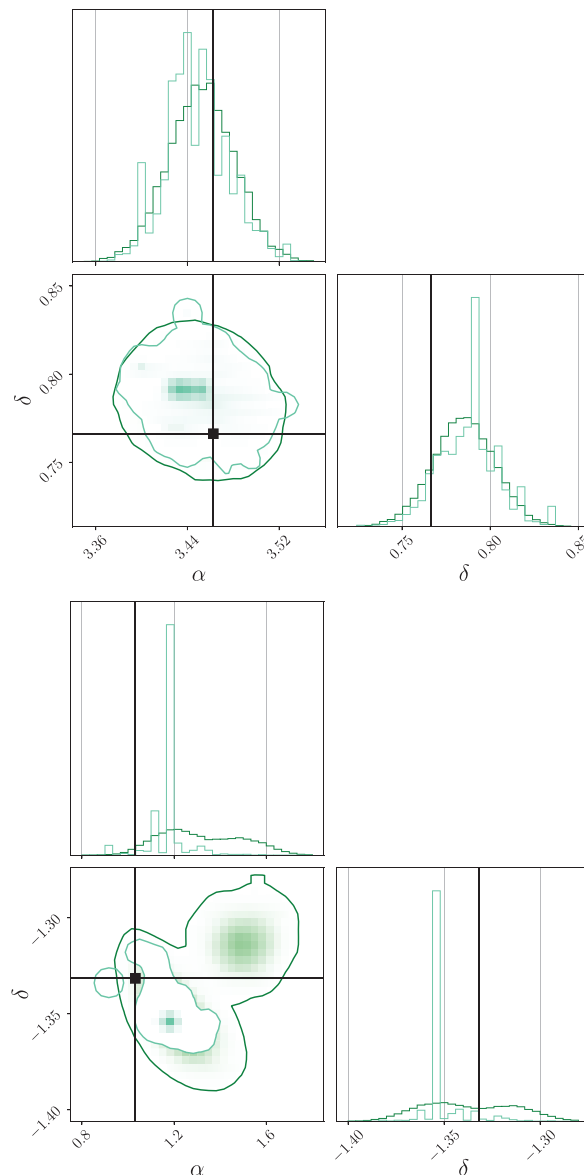


Figure 5. Sky recovery for the symmetric GOLUM framework (light blue) and complete joint inference (dark green). *Top:* GOLUM and joint parameter estimation have matching 90 per cent confidence interval recovery. *Bottom:* Scenario in which there are dominating samples in the reweighting process leading to a high-density region peaking away of the 90 per cent interval, artificially shrinking the region and excluding the injected value. One can also see the clear presence of dominant samples in the 1D posteriors, where there are a couple of bins with much more importance. This issue is reduced when using symmetric GOLUM.

10 CONFRONTATION WITH REAL DATA

The previous sections have shown symmetrization helps in getting GOLUM closer to the actual joint parameter estimation, and that one can generally skip the first image run to save time. In this section, we use the symmetric GOLUM and the publicly available data (The LIGO Scientific Collaboration, Virgo Collaboration & KAGRA Collaboration 2021a) from the GWTC-3 LVK catalogue (The LIGO Scientific Collaboration 2021b). We use the public samples and evidence obtained with BILBY (Ashton et al. 2019) and the IMRPHEMOPHM (Pratten et al. 2021) waveform as results for the

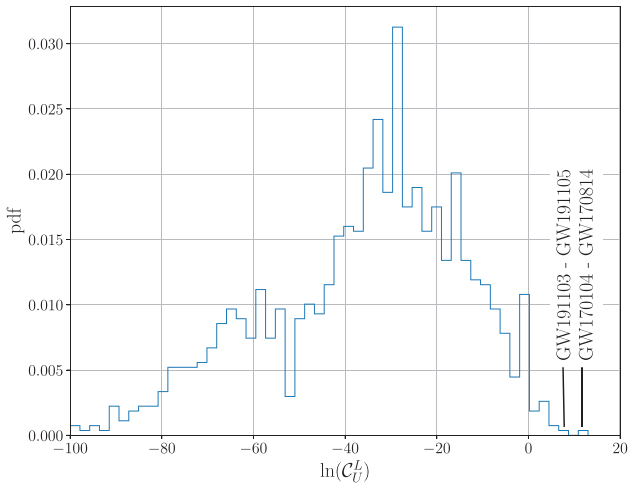


Figure 6. Histogram of the $\ln(C_U^L)$ values found for all the GW events using the symmetric GOLUM approach. Most of the events have a coherence ratio below zero. We also recover the different intriguing candidates reported in the literature: GW191103–GW191105 and GW170104–GW170814. While probably not lensed, recovering them shows that our method is trustworthy.

unlensed analysis. We also assume the HOMs to be negligible for all the events and do the approximate symmetric GOLUM analysis. To make a fair comparison, we also reweight all the posteriors and evidence from the initial data to correspond to values with the same priors. This avoids us from favoring some events because of the initial prior volume. In total, 87 BBHs are considered hence we analyse 3741 event pairs. The histogram of the $\ln(C_U^L)$ values, zoomed on the -100 to 14 value region, is presented in Fig. 6. There is a tail of events going down to a few hundred, but it is not so informative. We see that the vast majority of the events are well below zero. Only 43 event pairs have a positive coherence ratio. Out of those, most have $\ln(C_U^L) < 5$.

We note, however, the presence of two particular events in the higher tail, matching previously reported events. First, the GW191103–GW191105 event pair has a relatively high coherence ratio, $\ln(C_U^L) = 6.012$. This event is also reported in The LIGO Scientific Collaboration (2023c), where it is seen as an interesting candidate due to its characteristics. It also has a time delay of only about two days and a relative magnification close to 1. In addition, the two images are likely to be the same type. Having two type-I images close to each other is a probable scenario for galaxy lensing (More & More 2022). Therefore, the values are compatible with the lensing models. However, The LIGO Scientific Collaboration (2023c) show that it is not lensed when accounting for the population models. The posterior distribution we obtain for the relative magnification and the difference in Morse factor are represented in Fig. 7.

The other high value found is for GW170104–GW170814, where $\ln(C_U^L) = 11.153$, which is much higher than all the others. This event was already reported as significant in the literature (Hannuksela et al. 2019; Dai et al. 2020; Liu et al. 2021). However, it has a long time delay (we find 222.01 d) and a large difference in magnifications ($\mu_{\text{GW170814-GW170104}} = 0.37_{-0.07}^{+0.08}$, such that $D_L^{\text{GW170814}} = \sqrt{\mu_{\text{GW170814-GW170104}}} D_L^{\text{GW170104}}$). In addition, we find that the Morse factor difference between the two events is 1. Therefore, one of the two images is a type III image. This is not compatible with the expectations for galaxy lenses (see Collett & Bacon 2016; Collett et al. 2017; Wierda et al. 2021; More & More 2022, for example). This situation could be a bit more likely for a galaxy cluster lens,

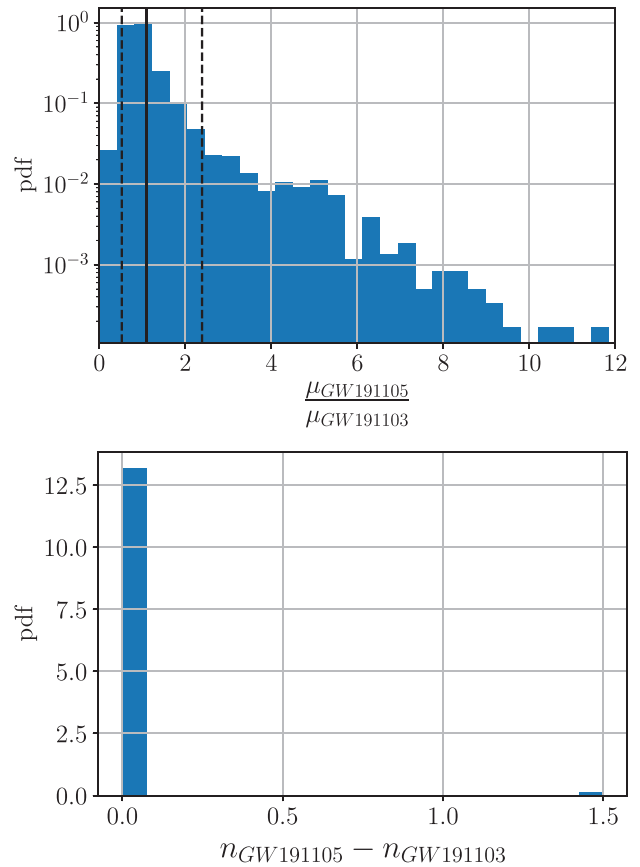


Figure 7. *Top:* Posterior recovery for the relative magnification for GW191103–GW191105. *Bottom:* Posterior recovery for the difference in the Morse factor. The two events also have a time delay of ~ 2.5 d. Therefore, the recovered parameters are compatible with the expectation for a galaxy lens (More & More 2022).

but this type of lensing has a low probability (see e.g. Smith et al. 2017, 2018, 2019; Robertson et al. 2020; Ryczanowski et al. 2020). Therefore, it seems unlikely for the event to be lensed and we do not push its study further in this work.

We do not report any new detections or claims. However, we have shown that our method can reproduce the results presented in literature over the years on a much shorter time-scale, owing to its enhanced speed.

11 CONCLUSIONS

In this work, we presented upgrades to the GOLUM framework, a tool to analyse strongly lensed GWs. We have shown how recasting the lookup table could reduce memory usage and computation time and how computing the number of effective samples leads to better stability. We then showed how the pipeline can be used in low latency by showing the generally small impact of the first image contribution to the coherence ratio. Finally, we have introduced a symmetrization of the pipeline, solving some issues found in the framework, such as difficulties in fully covering the parameter space when using only one image and vanishing samples in the reweighing process.

For each of the upgrades, we have shown how they work and what their effects are. In particular, we have compared the performances of the entire symmetric GOLUM framework with results coming from joint parameter estimation, also implemented in Janquart et al.

(2022). We started by showing that, generally, the first image run does not impact too much the evidence computation. Therefore, one is not forced to do it and can use the samples from the unlensed searches. Since the most computationally heavy part of GOLUM is the first image evaluation, dropping the first image analysis significantly reduces the computation time. In addition, the run time for the second image has been improved by a factor of 10 thanks to the different upgrades, making GOLUM quite suited for low-latency applications. This approach contains a certain degree of risk, as it can fail if important HOM contributions are present in the data. However, this would be flagged in the usual analyses, making the approximation evident. We also note that other low-latency methods, like posterior overlap, do not account for the Morse factor and are, consequently, sensitive to HOMs.

Using symmetrization leads to a clear improvement in the evidence compared to a single image, with values much closer to those given by joint analyses. In addition, we have shown that we get a better sample coverage with less scattered sky recovery. A remaining issue, only partially solved via the symmetric run, is the domination of a few samples in the weights. Indeed, a reduced number of samples can have much higher probability ratios in the reweighting process, leading to an artificial high-density region in the final posterior. This problem is already significantly reduced when using symmetrization, but there are still some cases where the final posterior is biased due to dominant weights. We suggest two possible avenues to solve this: use more samples to cover the parameters space better or perform importance sampling to correct for the bias using the formal joint likelihood. The main issue is that such methods are computationally heavy when applied to numerous samples. Still, it already gives a good idea about the sky localization of the event in case one would like to perform a follow-up analysis.

Finally, we have re-analysed all the GW BBH candidates released by the LVK since the first detection. Using the public samples and set-up, we performed (approximate) symmetric GOLUM runs for all the possible pairs. Our results match those of the LVK collaboration and other results given in the literature. Most events are certainly unlensed, and there are two notable event pairs: GW191103–GW191105 and GW170104–GW170814. The first has the second highest C_{L}^{L} but also has the particular property of having apparent lensing parameters compatible with the current galaxy lens models. The O2 pair is the one with the highest coherence ratio in all the pairs: $\ln(C_{\text{L}}^{\text{L}}) = 11.153$. While intriguingly high, other works have shown why this event pair is probably not lensed.

In the end, we have shown how improving our pipeline enables it to be faster and more reliable, making it possible to analyse large chunks of data in an approximate (but still accurate) form. This is a crucial step for future studies, where the number of events will grow steadily, and more and more pairs will have to be tracked. Therefore, using a more precise method with comparable speed is an important step forward to reduce the false-alarm probability. Let us also note that, with this approach, only a few events would have to be followed up using joint parameter estimation and low-latency results could be issued quickly in case an exceptional event is detected.

ACKNOWLEDGEMENTS

The authors thank Harsh Narola, Sylvia Biscoveanu, and John Vietch for useful discussions. We also thank Mick Wright for carefully re-reading the manuscript. JJ, KH, and CVDB are supported by the research program of the Netherlands Organisation for Scientific Research (NWO). The authors are grateful for computational resources provided by the LIGO Laboratory and supported by the

National Science Foundation Grant Nos. PHY-0757058 and PHY-0823459. This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation.

DATA AVAILABILITY

The data underlying this article will be shared in reasonable request to the corresponding authors.

REFERENCES

- Aasi J. et al., 2015, *Class. Quantum Gravity*, 32, 074001
 Abbott B. et al., 2016, *Phys. Rev. Lett.*, 116, 241102
 Abbott R. et al., 2020a, *Phys. Rev. D*, 102, 043015
 Abbott R. et al., 2020b, *ApJ*, 896, L44
 Abbott R. et al., 2021, *ApJ*, 923, 14
 Acernese F. et al., 2015, *Class. Quantum Gravity*, 32, 024001
 Akutsu T. et al., 2019, *Nat. Astron.*, 3, 35
 Akutsu T. et al., 2021, *Prog. Theor. Exp. Phys.*, 2021, 05A101
 Ashton G. et al., 2019, *ApJS*, 241, 27
 Aso Y., Michimura Y., Somiya K., Ando M., Miyakawa O., Sekiguchi T., Tatsumi D., Yamamoto H., 2013, *Phys. Rev. D*, 88, 043007
 Baker T., Trodden M., 2017, *Phys. Rev. D*, 95, 063512
 Caliskan M., Ezquiaga J. M., Hannuksela O., Holz D., 2022, APS April Meeting Abstracts, Y15.006
 Cao Z., Li L.-F., Wang Y., 2014, *Phys. Rev. D*, 90, 062003
 Cao S., Qi J., Cao Z., Biesiada M., Li J., Pan Y., Zhu Z.-H., 2019, *Sci. Rep.*, 9, 11608
 Cheung M. H. Y., Gais J., Hannuksela O. A., Li T. G. F., 2021, *MNRAS*, 503, 3326
 Christian P., Vitale S., Loeb A., 2018, *Phys. Rev. D*, 98, 103022
 Collett T. E., Bacon D. J., 2016, *MNRAS*, 456, 2210
 Collett T. E. et al., 2017, *ApJ*, 843, 148
 Cremonese P., Ezquiaga J. M., Salzano V., 2021, *Phys. Rev. D*, 104, 023503
 Dai L., Venumadhav T., 2017, preprint([arXiv:1702.04724](https://arxiv.org/abs/1702.04724))
 Dai L., Venumadhav T., Sigurdson K., 2017, *Phys. Rev. D*, 95, 044011
 Dai L., Zackay B., Venumadhav T., Roulet J., Zaldarriaga M., 2020, preprint([arXiv:2007.12709](https://arxiv.org/abs/2007.12709))
 Deguchi S., Watson W. D., 1986, *Phys. Rev. D*, 34, 1708
 Elvira V., Martino L., Robert C. P., 2018, *Int. Stat. Rev.*, 90, 525
 Ezquiaga J. M., Holz D. E., Hu W., Lagos M., Wald R. M., 2021, *Phys. Rev. D*, 103, 064047
 Fan X.-L., Liao K., Biesiada M., Piorowska-Kurpas A., Zhu Z.-H., 2017, *Phys. Rev. Lett.*, 118, 091102
 Farr W. M., 2019, *Res. Notes Am. Astron. Soc.*, 3, 66
 Golomb J., Talbot C., 2022, *ApJ*, 926, 79
 Goyal S., Haris K., Mehta A. K., Ajith P., 2021, *Phys. Rev. D*, 103, 024038
 Hannuksela O. A., Haris K., Ng K. K. Y., Kumar S., Mehta A. K., Keitel D., Li T. G. F., Ajith P., 2019, *ApJ*, 874, L2
 Hannuksela O. A., Collett T. E., Çalıřkan M., Li T. G., 2020, *MNRAS*, 498, 3395
 Haris K., Mehta A. K., Kumar S., Venumadhav T., Ajith P., 2018, preprint([arXiv:1807.07062](https://arxiv.org/abs/1807.07062))
 Iyer B., Souradeep T., Unnikrishnan C., Dhurandhar S., Raja S., Sengupta A., 2011, LIGO-India Tech. rep.. <https://dcc.ligo.org/LIGO-M1100296/public>
 Janquart J., Hannuksela O. A. K. H., Van Den Broeck C., 2021a, *MNRAS*, 506, 5430
 Janquart J., Seo E., Hannuksela O. A., Li T. G. F., Broeck C. V. D., 2021b, *ApJ*, 923, L1
 Janquart J., Haris K., Hannuksela O., 2022, GOLUM: A Software For Rapid Strongly-Lensed Gravitational Wave Parameter Estimation, Available at: <https://github.com/lemnis12/golum>
 Janquart J., More A., Van Den Broeck C., 2023, *MNRAS*, 519, 2046
 Jit Singh A., Li I. S. C., Hannuksela O. A., Li T. G. F., Kim K., 2018, preprint ([arXiv:1810.07888](https://arxiv.org/abs/1810.07888))

- Kim K., Lee J., Yuen R. S. H., Akseli Hannuksela O., Li T. G. F., 2021, *ApJ*, 915, 119
- Lai K.-H., Hannuksela O. A., Herrera-Martín A., Diego J. M., Broadhurst T., Li T. G. F., 2018, *Phys. Rev. D*, 98, 083005
- Li Y., Fan X., Gou L., 2019, *ApJ*, 873, 37
- Liao K., Fan X.-L., Ding X.-H., Biesiada M., Zhu Z.-H., 2017, *Nat. Commun.*, 8, 1148
- Li S.-S., Mao S., Zhao Y., Lu Y., 2018, *MNRAS*, 476, 2220
- Liu X., Magaña Hernandez I., Creighton J., 2021, *ApJ*, 908, 97
- Lo R. K. L., Magana Hernandez I., 2023, *Phys. Rev. D.*, 107, 123015
- Magaña Hernandez I., 2022, preprint (arXiv:2211.01272)
- Meena A. K., Bagla J. S., 2020, *MNRAS*, 492, 1127
- More A., More S., 2022, *MNRAS*, 515, 1044
- Nakamura T. T., 1998, *Phys. Rev. Lett.*, 80, 1138
- Ng K. K. Y., Wong K. W. K., Broadhurst T., Li T. G. F., 2018, *Phys. Rev. D*, 97, 023012
- Oguri M., 2018, *MNRAS*, 480, 3842
- Ohanian H. C., 1974, *Int. J. Theor. Phys.*, 9, 425
- Pagano G., Hannuksela O. A., Li T. G. F., 2020, *A&A*, 643, A167
- Pang P. T. H., Hannuksela O. A., Dietrich T., Pagano G., Harry I. W., 2020, *MNRAS*, 495, 3740
- Payne E., Talbot C., Thrane E., 2019, *Phys. Rev. D*, 100, 123017
- Pratten G. et al., 2021, *Phys. Rev. D*, 103, 104056
- Robertson A., Smith G. P., Massey R., Eke V., Jauzac M., Bianconi M., Rychanowski D., 2020, *MNRAS*, 495, 3727
- Rychanowski D., Smith G. P., Bianconi M., Massey R., Robertson A., Jauzac M., 2020, *MNRAS*, 495, 1666
- Schneider P., Ehlers J., Falco E., 1992, *Gravitational Lenses*. Springer-Verlag, Berlin
- Sereno M., Jetzer P., Sesana A., Volonteri M., 2011, *MNRAS*, 415, 2773
- Smith G. et al., 2017, Proc. IAU Symp. 338, *Gravitational Wave Astrophysics: Early Results from Gravitational Wave Searches and Electromagnetic Counterparts*. Kluwer, Dordrecht, p. 98
- Smith G. P., Jauzac M., Veitch J., Farr W. M., Massey R., Richard J., 2018, *MNRAS*, 475, 3823
- Smith G. P., Robertson A., Bianconi M., Jauzac M., 2019, preprint(arXiv:1902.05140)
- Somiya K., 2012, *Class. Quantum Gravity*, 29, 124007
- Speagle J. S., 2020, *MNRAS*, 493, 3132
- Takahashi R., Nakamura T., 2003, *ApJ*, 595, 1039
- The LIGO Scientific Collaboration, Virgo Collaboration, KAGRA Collaboration, 2021a, preprint(arXiv:2111.03606)
- The LIGO Scientific Collaboration, 2021b, preprint(arXiv:2111.03606)
- The LIGO Scientific Collaboration, 2021c, preprint(arXiv:2112.06861)
- The LIGO Scientific Collaboration, 2023a, *ApJ*, 949, 76
- The LIGO Scientific Collaboration, 2023b, *Phys. Rev. X*, 13, 011048
- The LIGO Scientific Collaboration, 2023c, preprint (arXiv:2304.08393)
- Vijaykumar A., Mehta A. K., Ganguly A., 2023, *Phys. Rev. D*, 108, 043036
- Wang Y., Stebbins A., Turner E. L., 1996, *Phys. Rev. Lett.*, 77, 2875
- Wang Y., Lo R. K. L., Li A. K. Y., Chen Y., 2021, *Phys. Rev. D*, 103, 104055
- Wempe E., Koopmans L. V. E., Wierda A. R. A. C., Akseli Hannuksela O., van den Broeck C., 2022, preprint (arXiv:2204.08732)
- Wierda A. R. A. C., Wempe E., Hannuksela O. A., Koopmans L. e. V. E., Van Den Broeck C., 2021, *ApJ*, 921, 154
- Wolfe N. E., Talbot C., Golomb J., 2023, *Phys. Rev. D*, 107, 104056
- Wright M., Hendry M., 2022, *ApJ*, 935, 68
- Yu H., Zhang P., Wang F.-Y., 2020, *MNRAS*, 497, 204

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.