

## 8 Examples of Bayesian statistics in gravitational wave astronomy

In this section we will provide some examples of the application of Bayesian statistics in gravitational wave astronomy. In most cases we will briefly outline what is done, and provide references where further information can be obtained.

### 8.1 LIGO Parameter Estimation

Parameter estimation results for sources detected by the LIGO interferometers are obtained and summarised as posterior distributions using the Bayesian techniques described earlier in this course. Typically, LIGO parameter estimation results are quoted as posterior medians and symmetric credible intervals. Figure 31 gives an example of this, showing the summary of parameter estimation results for all of the events observed by LIGO and Virgo during the O1 and O2 observing runs (Abbott et al. (2019), *Phys. Rev. X* **9** 031040).

LIGO/Virgo parameter estimation results in O1 and O2 were computed using the *LALInference* software suite, which includes two separate parameter estimation codes. *LALInferenceMCMC* is a Markov Chain Monte Carlo code, which generates posterior distributions using the Metropolis-Hastings algorithm and proposal distributions that are tuned to features expected in the likelihood for gravitational wave observations of compact binary inspirals. Further details can be found in

- Röver, C., Meyer, R., and Christensen, N., *Bayesian Inference on Compact Binary Inspiral Gravitational Radiation Signals in Interferometric Data*, *Class. Quantum Grav.* **23**, 4895 (2006).
- van der Sluys, M., Raymond, V., Mandel, I., Röver, C., Christensen, N., Kalogera, V., Meyer, R., and Vecchio, A., *Parameter Estimation of Spinning Binary Inspirals Using Markov-Chain Monte Carlo*, *Class. Quantum Grav.* **25**, 184011 (2008).

*LALInferenceNest* is a nested sampling algorithm, which obtains candidate values for updates to the live point set by carrying out short MCMC chains originating at the current lowest likelihood point in the live point set. Further details can be found in

- Veitch, J., and Vecchio, A., *Phys. Rev. D* **81**, 062003 (2010).

A summary of the *LALInference* package can be found in

- Veitch, J., et al., *Parameter Estimation for Compact Binaries with Ground-Based Gravitational-Wave Observations Using the LALInference Software Library*, *Phys. Rev. D* **91**, 042003 (2015).

and the version used in the analysis of the O2 events can be downloaded from

- [https://git.ligo.org/lscsoft/lalsuite/tree/lalinference\\_o2](https://git.ligo.org/lscsoft/lalsuite/tree/lalinference_o2) .

From O3 onwards, an additional parameter estimation code, *Bilby*, has been developed and used to obtain posterior distributions for LIGO/Virgo detections. This code uses generic freely available Bayesian sampling codes to draw samples from the posterior distribution, such as *DYNesty* and *PTMCMC*. The rest of the code consists of wrappers and functions to compute the correct likelihood to feed to the sampling codes. The description of the software can be found in

Event	$m_1/M_\odot$	$m_2/M_\odot$	$\mathcal{M}/M_\odot$	$\chi_{\text{eff}}$	$M_f/M_\odot$	$a_f$	$E_{\text{rad}}/(M_\odot c^2)$	$\ell_{\text{peak}}/(\text{erg s}^{-1})$	$d_L/\text{Mpc}$	$z$	$\Delta\Omega/\text{deg}^2$
GW150914	$35.6^{+4.7}_{-3.1}$	$30.6^{+3.0}_{-4.4}$	$28.6^{+1.7}_{-1.5}$	$-0.01^{+0.12}_{-0.13}$	$63.1^{+3.4}_{-3.0}$	$0.66^{+0.05}_{-0.04}$	$3.1^{+0.4}_{-0.4}$	$3.6^{+0.4}_{-0.4} \times 10^{56}$	$440^{+150}_{-170}$	$0.09^{+0.03}_{-0.03}$	182
GW151012	$23.2^{+14.9}_{-5.5}$	$13.6^{+4.1}_{-4.8}$	$15.2^{+2.1}_{-1.2}$	$0.05^{+0.31}_{-0.20}$	$35.6^{+10.8}_{-3.8}$	$0.67^{+0.13}_{-0.11}$	$1.6^{+0.6}_{-0.5}$	$3.2^{+0.8}_{-1.7} \times 10^{56}$	$1080^{+550}_{-490}$	$0.21^{+0.09}_{-0.09}$	1523
GW151226	$13.7^{+8.8}_{-3.2}$	$7.7^{+2.2}_{-2.5}$	$8.9^{+0.3}_{-0.3}$	$0.18^{+0.20}_{-0.12}$	$20.5^{+6.4}_{-1.5}$	$0.74^{+0.07}_{-0.05}$	$1.0^{+0.1}_{-0.2}$	$3.4^{+0.7}_{-1.7} \times 10^{56}$	$450^{+180}_{-190}$	$0.09^{+0.04}_{-0.04}$	1033
GW170104	$30.8^{+7.3}_{-5.6}$	$20.0^{+4.9}_{-4.6}$	$21.4^{+2.2}_{-1.8}$	$-0.04^{+0.17}_{-0.21}$	$48.9^{+5.1}_{-4.0}$	$0.66^{+0.08}_{-0.11}$	$2.2^{+0.5}_{-0.5}$	$3.3^{+0.6}_{-1.0} \times 10^{56}$	$990^{+440}_{-430}$	$0.20^{+0.08}_{-0.08}$	921
GW170608	$11.0^{+5.5}_{-1.7}$	$7.6^{+1.4}_{-2.2}$	$7.9^{+0.2}_{-0.2}$	$0.03^{+0.19}_{-0.07}$	$17.8^{+3.4}_{-0.7}$	$0.69^{+0.04}_{-0.04}$	$0.9^{+0.0}_{-0.1}$	$3.5^{+0.4}_{-1.3} \times 10^{56}$	$320^{+120}_{-110}$	$0.07^{+0.02}_{-0.02}$	392
GW170729	$50.2^{+16.2}_{-10.2}$	$34.0^{+9.1}_{-10.1}$	$35.4^{+6.5}_{-4.8}$	$0.37^{+0.21}_{-0.25}$	$79.5^{+14.7}_{-10.2}$	$0.81^{+0.07}_{-0.13}$	$4.8^{+1.7}_{-1.7}$	$4.2^{+0.9}_{-1.5} \times 10^{56}$	$2840^{+1400}_{-1360}$	$0.49^{+0.19}_{-0.21}$	1041
GW170809	$35.0^{+8.3}_{-5.9}$	$23.8^{+5.1}_{-5.2}$	$24.9^{+2.1}_{-1.7}$	$0.08^{+0.17}_{-0.17}$	$56.3^{+5.2}_{-3.8}$	$0.70^{+0.08}_{-0.09}$	$2.7^{+0.6}_{-0.6}$	$3.5^{+0.6}_{-0.9} \times 10^{56}$	$1030^{+320}_{-390}$	$0.20^{+0.05}_{-0.07}$	308
GW170814	$30.6^{+5.6}_{-3.0}$	$25.2^{+2.8}_{-4.0}$	$24.1^{+1.4}_{-1.1}$	$0.07^{+0.12}_{-0.12}$	$53.2^{+3.2}_{-2.4}$	$0.72^{+0.07}_{-0.05}$	$2.7^{+0.4}_{-0.3}$	$3.7^{+0.4}_{-0.5} \times 10^{56}$	$600^{+150}_{-220}$	$0.12^{+0.03}_{-0.04}$	87
GW170817	$1.46^{+0.12}_{-0.10}$	$1.27^{+0.09}_{-0.09}$	$1.186^{+0.001}_{-0.001}$	$0.00^{+0.02}_{-0.01}$	$\leq 2.8$	$\leq 0.89$	$\geq 0.04$	$\geq 0.1 \times 10^{56}$	$40^{+7}_{-15}$	$0.01^{+0.00}_{-0.00}$	16
GW170818	$35.4^{+7.5}_{-4.7}$	$26.7^{+4.3}_{-5.2}$	$26.5^{+2.1}_{-1.7}$	$-0.09^{+0.18}_{-0.21}$	$59.4^{+4.9}_{-3.8}$	$0.67^{+0.07}_{-0.08}$	$2.7^{+0.5}_{-0.5}$	$3.4^{+0.5}_{-0.7} \times 10^{56}$	$1060^{+420}_{-380}$	$0.21^{+0.07}_{-0.07}$	39
GW170823	$39.5^{+11.2}_{-6.7}$	$29.0^{+6.7}_{-7.8}$	$29.2^{+4.6}_{-3.6}$	$0.09^{+0.22}_{-0.26}$	$65.4^{+10.1}_{-7.4}$	$0.72^{+0.09}_{-0.12}$	$3.3^{+1.0}_{-0.9}$	$3.6^{+0.7}_{-1.1} \times 10^{56}$	$1940^{+970}_{-900}$	$0.35^{+0.15}_{-0.15}$	1666

Figure 31: Parameter estimation results summary from the first Gravitational Wave Transient Catalogue published by the LIGO/Virgo collaboration (*Phys. Rev. X* **9** 031040 (2019)). Results are presented as the median and 90% symmetric credible interval of the Bayesian posterior distribution.

- Ashton, G., et al. (2019), *Astrophys. J. Supp.* **241**, 27

and the software can be downloaded from

- <https://git.ligo.org/lscsoft/bilby>

As well as providing tables summarising the median and symmetric credible intervals for the observed sources, LIGO papers typically include plots of the full Bayesian posterior distributions. These take various forms. Two-dimensional joint posterior distributions are often given for pairs of parameters that are correlated, such as the chirp mass and mass ratio or the final mass and spin of the remnant black hole produced by the merger or the sky location of the merger event. Examples of two-dimensional posterior distributions are shown in Figure 32 and Figure 33. One dimensional posteriors are often plotted as “violin plots” to allow comparison between the results for multiple events. The violin plot plots the parameter value on the  $y$ -axis and the posterior density on the  $x$ -axis, which is opposite to the usual convention. Additionally, the posterior is reflected in the  $y$ -axis so that it is symmetric about that axis for each event. The width of the resulting violin plot is proportional to the posterior probability for the corresponding value of the parameter. An example is shown in Figure 34. Posteriors in the spins of the black holes, which is fundamentally a three-dimensional quantity, are typically represented by semi-circular density plots such as those shown in Figure 35. The full 3D posterior is marginalised over the (poorly constrained) azimuthal direction of the spin, and the resulting 2D posterior is represented on a semi-circle with the spin-magnitude as the radial direction and the angle between the spin vector and the orbital angular momentum as the angular direction. The density of the colour in these plots is proportional to the posterior density for the corresponding spin vector.

*LALInference* is also used to obtain posterior deviations on parameters characterising deviations from general relativity, to facilitate tests of GR. More details can be found, along with results from analysis of the O1 and O2 events, in Abbott, B.P., et al., *Phys. Rev. D* **100**, 104036.

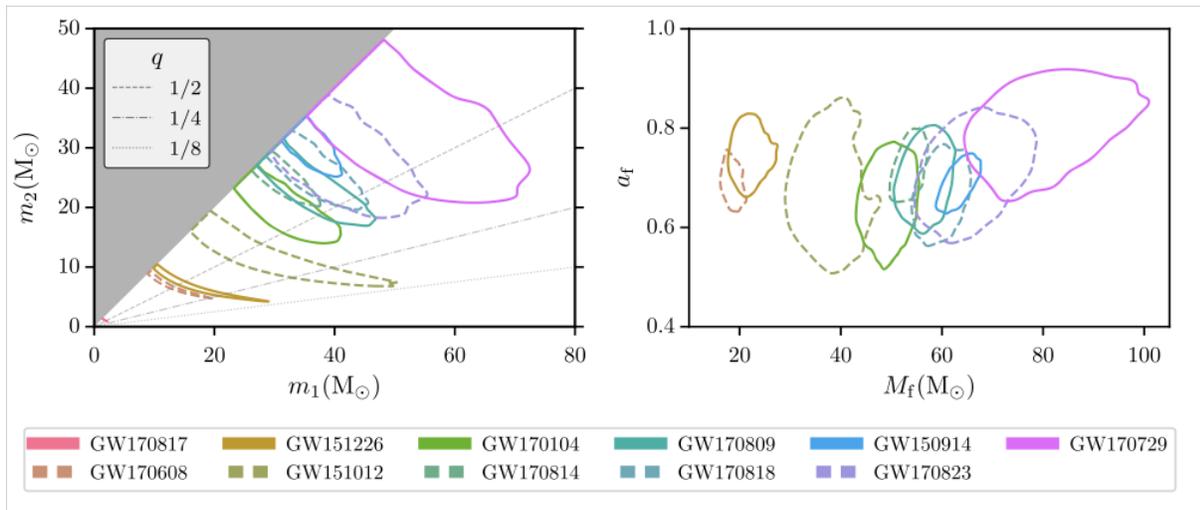


Figure 32: Joint two dimensional posterior on mass and mass ratio (left) and on final mass and spin (right) for all of the events observed by LIGO/Virgo during the O1 and O2 observing runs. Reproduced from Abbott et al. (2019), *Phys. Rev. X* **9** 031040.

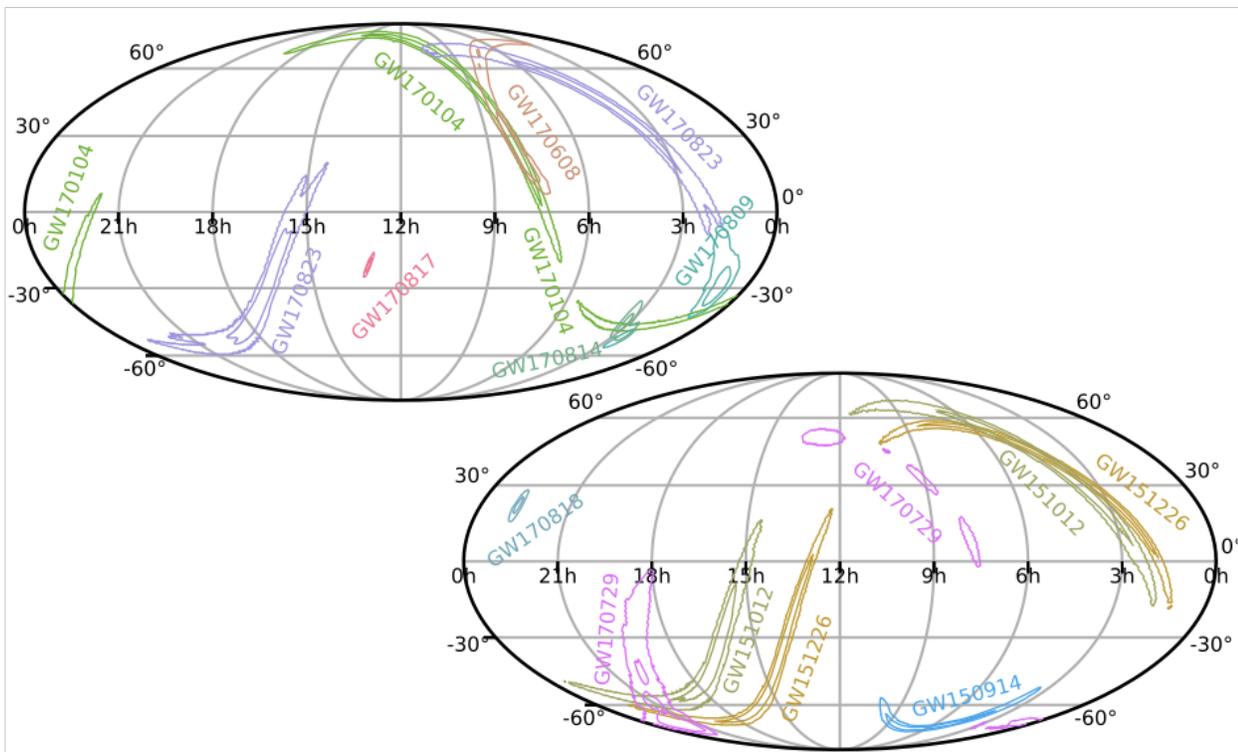


Figure 33: Sky location posterior distribution for all events observed by LIGO/Virgo during the O1 and O2 observing runs. Reproduced from Abbott et al. (2019), *Phys. Rev. X* **9** 031040.

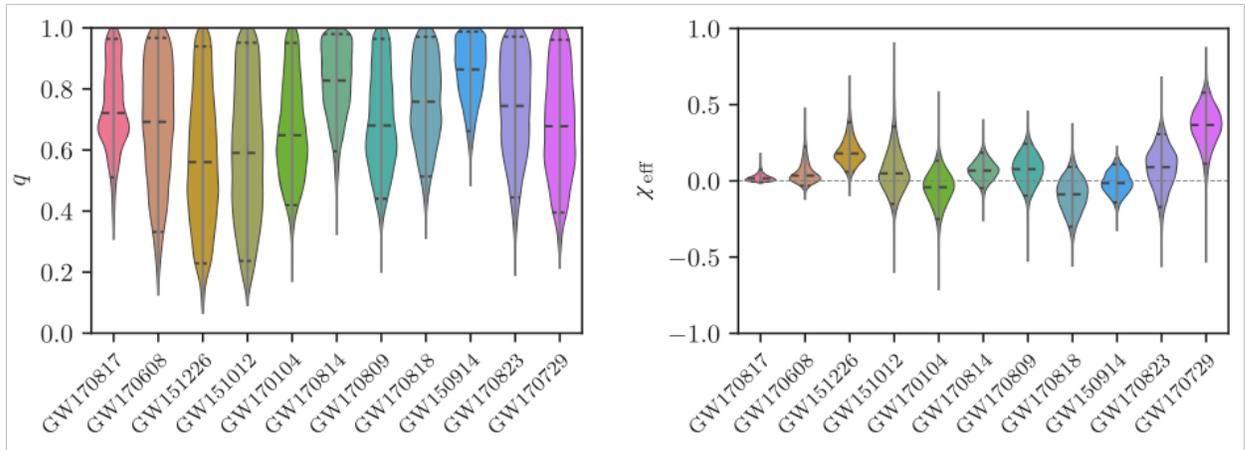


Figure 34: One-dimensional marginalised posteriors on the mass ratio (left) and effective spin (right) for all the events observed by LIGO/Virgo during the O1 and O2 observing runs. The one-dimensional posteriors are represented as “violin plots” as described in the text. Reproduced from Abbott et al. (2019), *Phys. Rev. X* **9** 031040.

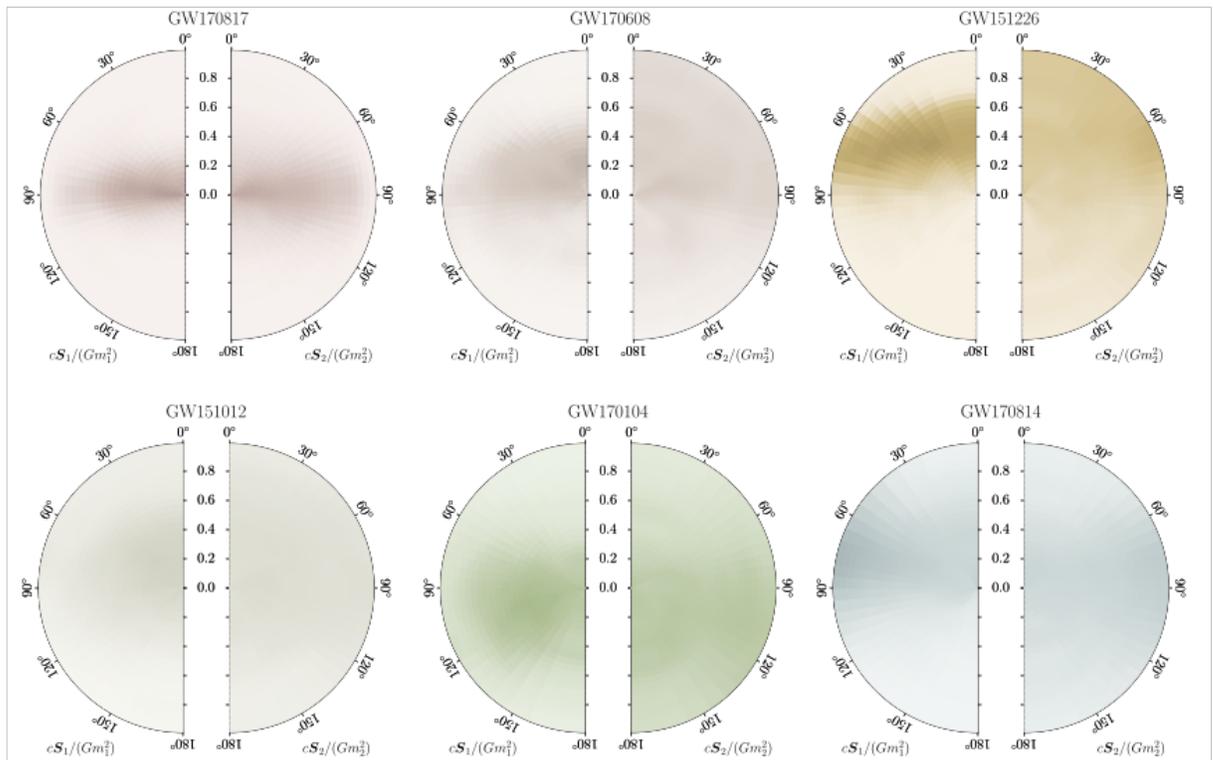


Figure 35: Posteriors on the spins of the two components in the binary for all of the events observed by LIGO/Virgo during the O1 and O2 observing runs. The distance from the origin represents the magnitude of the spin, and the angle represents the direction of the spin. The two halves of the plot are for the primary (left) and secondary (right) object in the binary. The density of colour is proportional to the posterior density for that spin value. Reproduced from Abbott et al. (2019), *Phys. Rev. X* **9** 031040.

## 8.2 Reduced order modelling

LIGO parameter estimation codes are computationally expensive, primarily due to the cost of evaluating models of the gravitational waveforms to compute likelihoods. To make inference more efficient, it is advantageous to have models of the signals that are quicker to evaluate. This has been achieved by building **reduced order models** and **surrogate models**. The principle of both approaches is quite similar. First, a basis for the space of waveforms is found that has lower dimensionality than the number of samples in the original waveforms. Then either a fast interpolant is constructed to map physical parameters to the weights of the basis functions (in the case of some surrogate models, the interpolant is built directly for the waveform itself) or a **reduced order quadrature** representation of the likelihood is constructed. In the latter approach, a projection of the target waveform onto the reduced basis is obtained not by using overlaps to find the best projection, but instead by requiring the target waveform to exactly match a linear combination of basis waveforms at a number of points, called **quadrature interpolation points**, equal to the number of functions in the basis. This allows the likelihood quadrature to be reduced to a sum over the target waveform evaluated at the quadrature points weighted by data-dependent constants that can be computed prior to running inference from overlaps of the basis functions with the data

$$\begin{aligned}
 \left( h(\vec{\lambda}) | d \right) &= 4\Re \int_0^\infty \frac{\tilde{h}(\vec{\lambda}) \tilde{d}^*(f)}{S_h(f)} df \\
 &\approx 4\Re \left[ \sum_{k=0}^{N/2} d^*(f_k) \vec{e}^T(f_k) \Delta f \mathbf{A}^{-1} \right] \vec{h}(\vec{\lambda}) \\
 &= 4\Re \sum_{k=1}^m \omega_k h(F_k; \vec{\lambda}).
 \end{aligned} \tag{109}$$

Reduced order quadrature approximations to likelihoods are the state of the art in LIGO parameter estimation, but they require being able to evaluate the target waveform at certain frequencies quickly and so can only really be used with frequency-domain waveform approximations. Surrogate models can be used to accelerate inference with time-domain waveform models.

## 8.3 Population inference

Inference on the properties of the population of sources from which the observed LIGO events are drawn also uses Bayesian methods, specifically Bayesian hierarchical modelling. We encountered one example of this in Section 4.9, which is the inference of cosmological parameters using gravitational wave observations of binary neutron star mergers with counterparts. Other examples include inference on the rate of mergers of different types of source in the Universe, and on the distributions of masses and spins of black holes and neutron stars. Full details on the range of population analyses carried out for the O1 and O2 events can be found in Abbott, B.P., et al., *Astrophys. J. Lett.* **882**, L24 (2019) and references therein, but we summarise some of the key analyses here.

### 8.3.1 Rate estimation

Accurate estimation of the rate of events in the Universe is complicated by confusion with detector noise, i.e., identifying which events are real gravitational wave events and which are instrumental artefacts, and by the need to make assumptions about the distribution of parameters of sources in the population. The first problem was tackled in Farr, W., Gair, J.R., Mandel, I., and Cutler, C., *Phys. Rev. D* **91**, 023005 (2014). If the output of the detector is represented by a sequence of values of a detection statistic,  $x$ , and any statistic value that exceeds some threshold,  $x_{\min}$ , is regarded as a detection, then the observed data is a set of detection statistic values above threshold,  $\{x_i\}$ . Some of these events correspond to real foreground events, while others arise due to noise fluctuations in the detector and are background. We introduce an (unobserved) parameter  $f_i$  for each event such that  $f_i = 1$  if it is a foreground event and  $f_i=0$  if it is background. The foreground and background events are assumed to be generated by independent Poisson processes with rates

$$\frac{dN_f}{dx} = R_f \hat{f}(x, \theta_f), \quad \frac{dN_b}{dx} = R_b \hat{b}(x, \theta_b)$$

and corresponding cumulative distributions  $\hat{F}(x, \theta_f)$ ,  $\hat{B}(x, \theta_b)$ . Here  $R_f$  and  $R_b$  are the foreground and background rates respectively and  $\theta_f$  and  $\theta_b$  represent any unknown parameters that characterise the foreground and background distributions. The combined posterior for the rates, event flags and distribution parameters is

$$p(f_i, R_f, R_b, \theta | d_{\text{to}}, N) = \frac{\alpha}{p(d_{\text{to}}, N) N!} \left[ \prod_{i|f_i=1} R_f \hat{f}(x_i, \theta) \right] \left[ \prod_{i|f_i=0} R_b \hat{b}(x_i, \theta) \right] \exp[-(R_f + R_b)] \frac{p(\theta)}{\sqrt{R_f R_b}}$$

where  $p(\theta)$  is the prior on the posterior parameters and we are using a Jeffreys' prior  $p(R) \propto 1/\sqrt{R}$  on the rates. The subscript on  $d_{\text{to}}$  indicates that we are using time-ordered data. The data could also be analysed ordered by ranking statistic. This posterior can be marginalised over the unknown flags to give posteriors on the rates, or over the rates to give posterior probabilities for  $f_i = 1$  for each event.

One complication with this approach is that it relies on a model for the foreground and background distributions. These can be estimated by injections and time-slides, but, since LIGO is not equally sensitive to all types of CBC event, the former requires imposing some model of the astrophysical population from which the events are drawn. One approach to this is to assume that all events in the Universe are the same as the one that has been observed. This approach was used in Kim, Kalogera and Lorimer (*Astrophys. J.* **584**, 985 (2003)) to estimate the rate of double neutron star mergers and so is often referred to as the ‘‘KKL method’’. In the first LIGO detection paper, for GW150914, the combination of the rate estimation accounting for confusion (FGMC) and the KKL method was used to infer the rate of binary black hole mergers. The application of this ‘‘alphabet soup’’ method was complicated by the fact that the data being analysed to infer the background for GW150914 contained a second CBC trigger, LVT151012. The parameters of this event were completely different to GW150914, so the KKL method could still be applied, but generalising to the case where all events in the Universe were either like GW150914 or LVT151012. Further details can be found in Abbott, B.P., et al. *Astrophys. J. Lett.* **833**, 1 (2016) and Abbott, B.P., et al. *Astrophys. J. Supp.* **227**, 14 (2016).

One additional trigger, GW151226, was present in the LIGO O1 data, and that again had sufficiently distinct parameters that the KKL approach could be used. In O2, the events

began to have much more posterior overlap and so this method could no longer be used. Now, a model of the population is assumed in event rate estimation. Recent analyses have used both a power-law mass distribution or a flat in log-mass distribution in an attempt to bound the range of possible rate, but future results are likely to shift towards a single combined analysis of the population parameters and rate.

### 8.3.2 Black hole mass distribution

The mass distribution of stellar-origin black holes in binaries can be inferred from LIGO/Virgo observations in a hierarchical analysis by placing a prior on the mass of individual events that depends on some unknown parameters that can be constrained from analysing the full set of events. In Abbott, B.P., et al., *Astrophys. J. Lett.* **882**, L24 (2019) three different models of the mass function were used. Models A and B assumed a power law distribution on mass and mass ratio

$$p(m_1, m_2 | m_{\min}, m_{\max}, \alpha, \beta_q) \propto \begin{cases} C(m_1) m_1^{-\alpha} q^{\beta_q} & \text{if } m_{\min} \leq m_2 \leq m_1 \leq m_{\max} \\ 0 & \text{otherwise} \end{cases} .$$

In model A,  $m_{\min} = 5M_{\odot}$ ,  $\beta_q = 0$  and the only free parameters are  $m_{\max}$  and  $\alpha$ . In model B, all four parameters are allowed to vary. The third model mixes a power-law component of the above form, with a Gaussian component, designed to fit any excess of events near the lower mass limit of the pair-instability supernova mass gap. The model is

$$p(m_1 | \theta) = \left[ (1 - \lambda_m) A(\theta) m_1^{-\alpha} \Theta(m_{\max} - m_1) + \lambda_m B(\theta) \exp\left(-\frac{(m_1 - \mu_m)^2}{2\sigma_m^2}\right) \right] S(m_1, m_{\min}, \delta m)$$

$$p(q = m_2/m_1 | m_1, \theta) = C(m_1, \theta) q^{\beta_q} S(m_2, m_{\min}, \delta m). \quad (110)$$

The mass distribution obtained by fitting these models to the O1 and O2 data is shown in Figure 36.

### 8.3.3 Black hole spin distribution

A hierarchical analysis of LIGO/Virgo events can also provide insight into the spin distribution. This can be done either parametrically or non-parametrically and both analyses were done for the O1 and O2 events in Abbott, B.P., et al., *Astrophys. J. Lett.* **882**, L24 (2019). The parametric approach models the spin magnitude using a Beta distribution

$$p(a_i | \alpha_a, \beta_a) = \frac{a_i^{\alpha_a - 1} (1 - a_i)^{\beta_a - 1}}{B(\alpha_a, \beta_a)}$$

while the non-parametric analysis models the spin-magnitude distribution as a set of heights of a binned distribution, with the bin heights free parameters to be determined by the observations. For example, a three-bin distribution (Farr, B., Holz, D., and Farr, W., *Astrophys. J.* **854**, L9 (2018))

$$p(a) = \begin{cases} A_1/3 & 0 \leq a \leq 1/3 \\ A_2/3 & 1/3 \leq a \leq 2/3 \\ 1 - (A_1 + A_2)/3 & 2/3 \leq a \leq 1 \end{cases} .$$

The posteriors obtained from applying these models to the O1 and O2 events are shown in Figure 37.

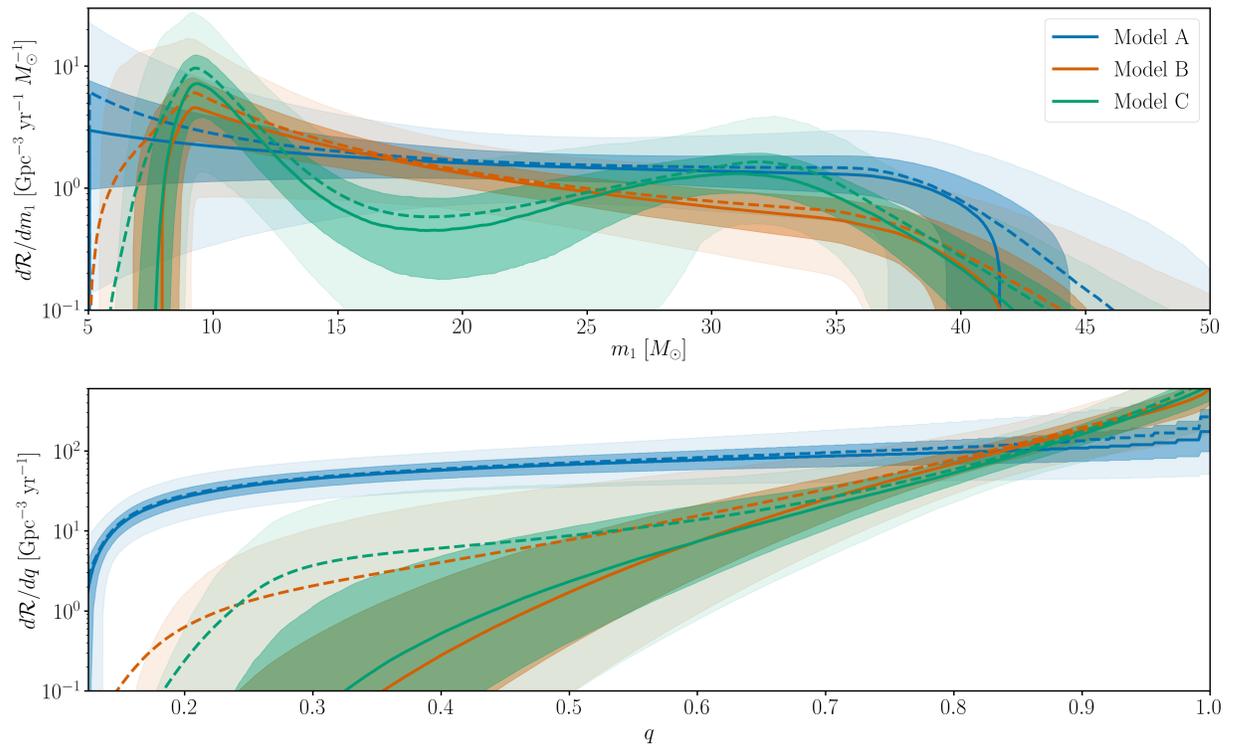


Figure 36: Black hole mass function inferred from LIGO/Virgo events observed in the O1 and O2 observing runs. Figure reproduced from Abbott, B.P., et al., *Astrophys. J. Lett.* **882**, L24 (2019).

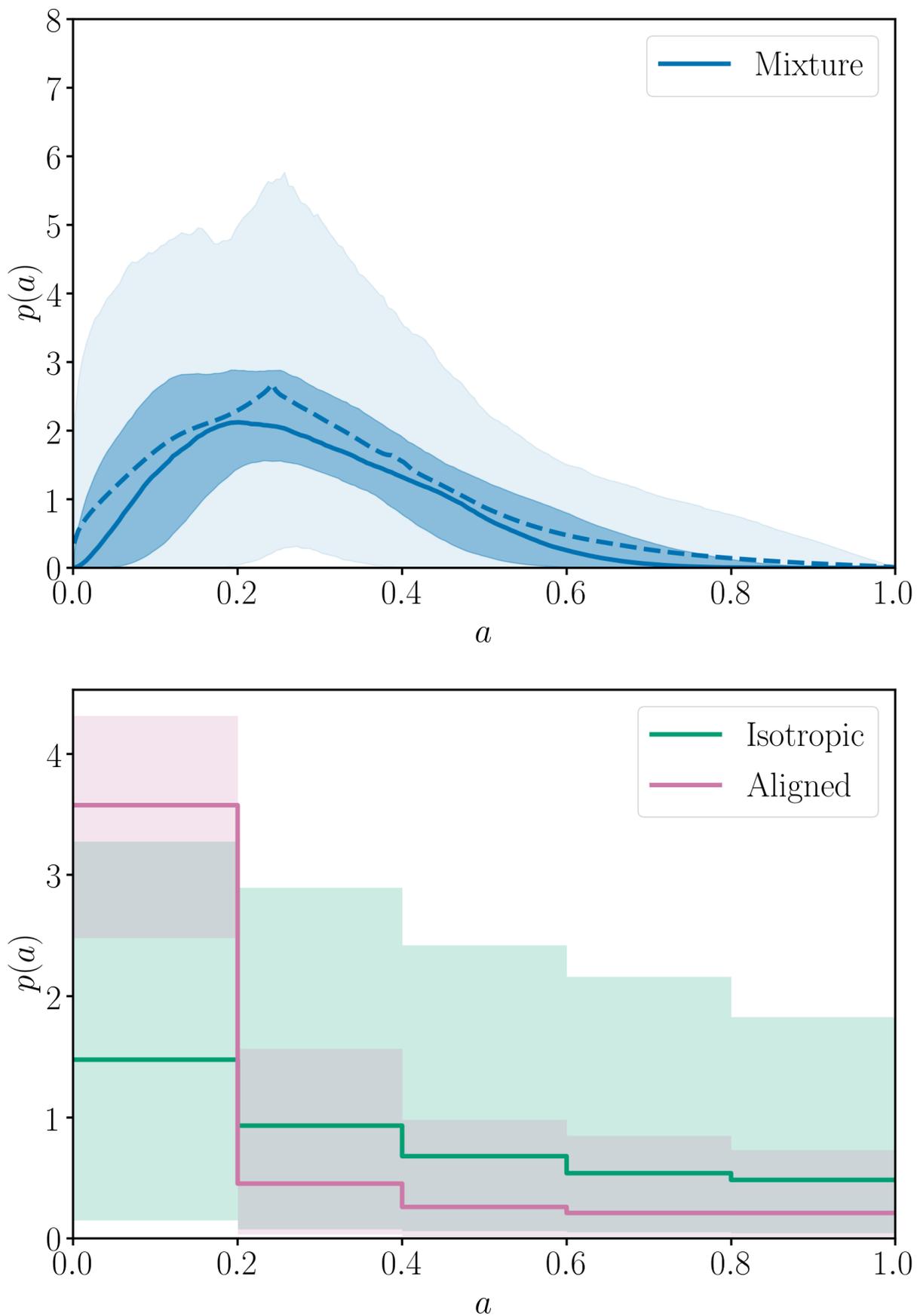


Figure 37: Black hole spin distribution inferred from LIGO/Virgo events observed in the O1 and O2 observing runs, using a parametric (top panel) or non parametric (bottom panel)

The spin direction is also a parameter of interest astrophysically, as different formation scenarios predict either isotropically distributed spin directions, or a preference for spins to be aligned with the angular momentum of the binary. To capture this, the analysis of the O1 and O2 data used a mixture model

$$p(\cos t_1, \cos t_2 | \sigma_1, \sigma_2, \zeta) = \frac{(1 - \zeta)}{4} + \frac{2\zeta}{\pi} \prod_{i \in \{1,2\}} \frac{\exp(-(1 - \cos t_i)^2 / 2\sigma_i^2)}{\sigma_i \operatorname{erf}(\sqrt{2}/\sigma_i)}.$$

At present, LIGO measurements are not sufficiently informative about spins to strongly constrain the parameters of the model.

### 8.3.4 Rate evolution

The FGMC+KKL method described earlier assumes that the rate of mergers is constant, but in principle this could evolve over cosmic history (the FGMC framework can handle this, but the interpretation of  $R_f$  is different, as the average rate over the sensitive volume of the detector). An evolution of the rate can be explicitly included and constrained by introducing an extra parameter into the rate density

$$\frac{dR}{d\xi}(z|\theta) = R_0 p(\xi|\theta) (1+z)^\lambda.$$

The analysis of the O1 and O2 events provided weak evidence for an evolution in rate with redshift, but this was mostly due to the event GW170729, which was the most marginal detection. The rate evolution will be better constrained by the order of magnitude increase in events expected in O3 and future observing runs.

## 8.4 Model selection

Bayesian methods are also applied to model selection using the LIGO/Virgo observations, through the evaluation of **evidence ratios** or **Bayes factors** for pairs of alternative hypotheses for the data. Some examples of applications to gravitational wave data are

- Test for the presence of a signal in the data after the end of the merger of the two neutron stars in GW170817. Such a signal might be evidence that the merger product was a hypermassive neutron star rather than a black hole. For GW170817 the Bayes factor for the noise model over the signal model was 256.79 (Abbott, B.P., et al., *Phys. Rev. X* **9** 011001 (2019)), providing strong evidence that no such signal was present.
- Test of the polarisation state of gravitational waves. Possible models are that the gravitational waves have tensor polarisation, as expected in GR, or have scalar polarisation or vector polarisation. The analysis of GW170818 gave Bayes factors of 12 for tensor versus vector polarisation and 407 for tensor versus scalar, while the analysis of GW170814 gave Bayes' factors of 30 and 220 respectively (Abbott, B.P., et al., *Phys. Rev. D* **100** 104036 (2019)).
- Tests of the no-hair property of the remnant black hole formed in a merger, by comparing the properties of the observed ringdown radiation to that predicted by GR (Brito, Buonanno and Raymond, *Phys. Rev. D* **98**, 084038 (2018)).

- Probing alternative theories of gravity. For example, looking for evidence for dynamical gravity with the polarisation of continuous gravitational waves (Isi et al., *Phys. Rev. D* **96**, 042001 (2017)).

## 8.5 Source reconstruction

Although Bayesian inference relies on the existence of models, it is also possible to use these methods to recover “unmodelled” sources. One such implementation is the BAYESWAVE algorithm. The method works by modelling the noise and signals in the data from the various detectors as a superposition of simple components. BAYESWAVE represents the noise as a combination of a smooth PSD component, described by a cubic spline, lines represented by Lorentzians and glitches modelled by wavelets. Signals in the data are also modelled by wavelets, but with parameters that are common across the detectors, as opposed to the noise components which are independent in different detectors. Wavelets are simple functions that are compact in both time and frequency. We will encounter these again in the non-parametric regression section of this course. There are many different wavelet families, but the wavelets used in BAYESWAVE are known as the Morley-Gabor basis.

BAYESWAVE fits its model using reversible jump MCMC. The reversible jump element is required to add or remove wavelet or line components, as the number of these required is not known a priori. Further details on the BAYESWAVE algorithm can be found in

- Cornish, N.J., and Littenberg, T.B., *Class. Quantum Grav.* **32**, 135012 (2015).
- Littenberg, T.B., and Cornish, N.J., *Phys. Rev. D* **91**, 084034 (2015).

BAYESWAVE is used in LIGO analyses for PSD estimation, glitch removal and for non-parametric waveform reconstruction. The good agreement between the BAYESWAVE reconstructed waveform and the best fit model found by parameter estimation for GW150914 (see Figure 38) provided extra support to the fact that this was a true signal.

## 8.6 Rapid localisation

Since the start of the O1 observing run, LIGO/Virgo have been sending out triggers to facilitate follow-up of gravitational wave events by electromagnetic telescopes. To avoid delays to these alerts, it is necessary to rapidly estimate the sky location of the triggers so that astronomers know where to point their telescopes. Bayesian techniques are also used for this purpose. Full Bayesian parameter estimation is not possible in low-latency, so the rapid localisation algorithms are not truly Bayesian, but make approximations in evaluating the posterior that allow it to be computed quickly.

The BAYESTAR algorithm replaces the full likelihood by the autocorrelation likelihood, which is the likelihood evaluated at the maximum likelihood parameter values, as returned by the online search algorithms. This autocorrelation likelihood takes the form

$$\exp \left[ -\frac{1}{2} \sum_i \rho_i^2 + \sum_i \rho_i \Re \left\{ e^{-i\gamma_i} z_i^*(\tau_i) \right\} \right]$$

where  $\rho_i$  denotes the signal to noise ratio in detector  $i$ ,  $\gamma_i$  and  $\tau_i$  are the phase and time of arrival of the trigger in detector  $i$  and  $z_i(t)$  is the time-series of the matched filter overlap in detector  $i$ . The marginalisation of this integral over all parameters except sky location is

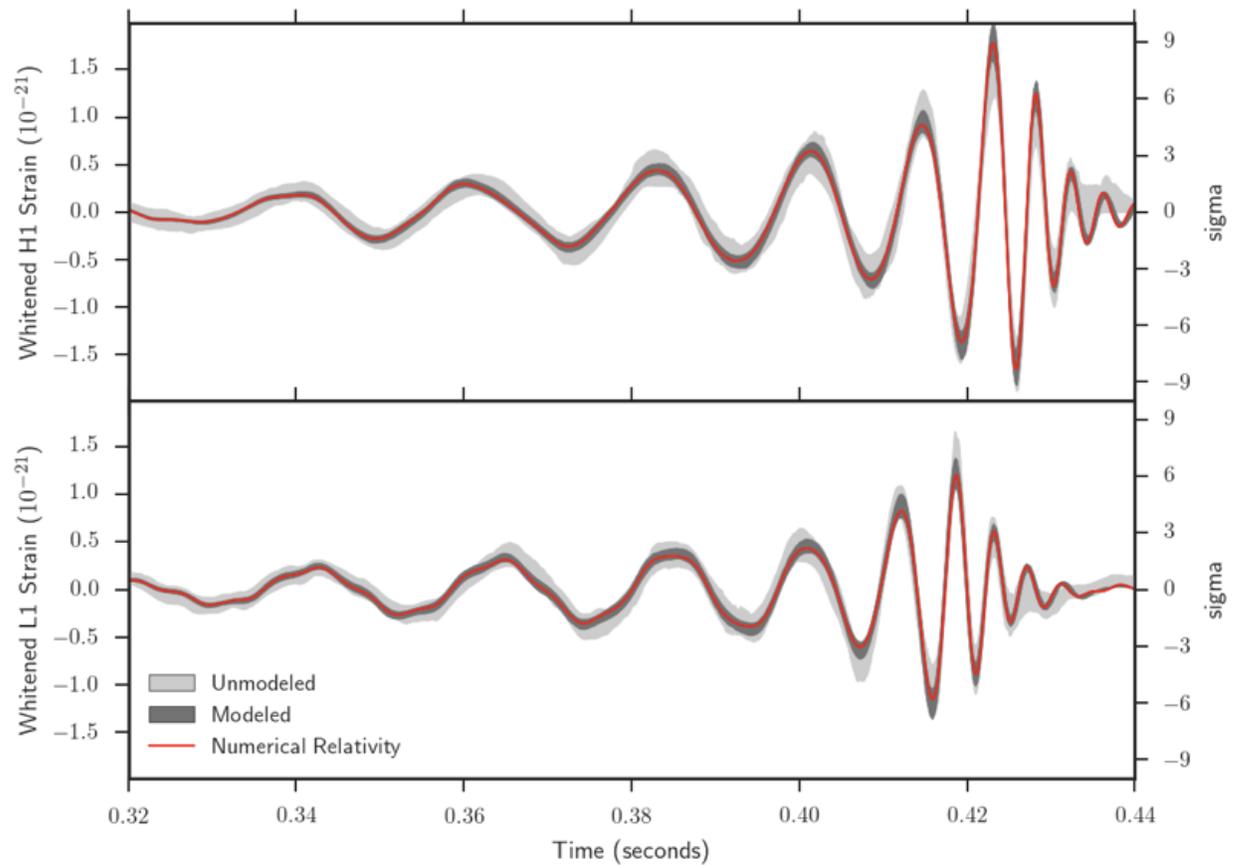


Figure 38: BAYESWAVE reconstruction of GW150914 (labelled “unmodelled”), compared to the waveform corresponding to the maximum a posteriori parameters obtained by parameter estimation (labelled “modelled”) and a numerical relativity waveform with consistent parameters. Figure reproduced from Abbott, B.P., et al., *Phys. Rev. Lett.* **116**, 061102 (2016).

accelerated using approximations to the marginalisation integrals and by employing look-up tables. The result of running the algorithm is a sky map probability density, i.e., a weighting of pixels on the sky by their relative probability of being the true location of the observed transient.

More details on the BAYESTAR algorithm can be found in

- Singer, L., and Price, L., *Phys. Rev. D* **93**, 024013 (2016).

Another rapid localisation algorithm used in LIGO is LALINFERENCEBURST or LIB. In this case, computational savings in the model are obtained by representing an arbitrary signal as a single sine-Gaussian

$$h_+(t) = \cos(\alpha) \frac{h_{\text{rss}}}{\sqrt{Q(1 + \cos(2\phi_0)e^{-Q^2})/4f_0\sqrt{\pi}}} \sin(2\pi f_0(t - t_0) + \phi_0)e^{-(t-t_0)^2/\tau^2}.$$

While this simple model cannot accurately describe all signals, it does represent the relative amplitudes of the signal in different detectors correctly and that is enough to obtain reasonable sky-localisation accuracies.

There is also an online version of LIB, called oLIB, that uses Bayesian evidences computed by LIB to assess triggers identified in a time-frequency analysis. The evidences for the triggers being noise versus signal and being coherent in different detectors versus incoherent are used to identify potentially interesting candidate events for follow-up. oLIB was running at the time of GW150914 and, along with CWB, was the first algorithm to identify this signal in the data.

More details on the LALINFERENCEBURST algorithm and on oLIB, can be found in

- Essick, R., Vitale, S., Katsavounidis, E., Vedovato, G., and Klimenko, S., *Astrophys. J.* **800**, 81 (2015).
- Lynch, R., Vitale, S., Essick, R., Katsavounidis, E., and Robinet, F., *Phys. Rev. D* **95**, 104046 (2017).

## 8.7 LISA parameter estimation

Bayesian methods have also been used in the context of data analysis development for LISA, mostly in the framework of the sequence of Mock LISA Data Challenges (MLDCs) that took place between 2006 and 2010. Bayesian techniques, with some frequentist simplifications such as the use of the  $\mathcal{F}$ -statistic, were used not only to characterise the identified sources, but also to search for sources in the data set. A variety of techniques were employed, including Markov Chain Monte Carlo algorithms, genetic algorithms and nested sampling. These methods were successfully able to find and characterise sources in the sample data sets, although these were somewhat simplified, containing only Gaussian instrumental noise with known PSD and a reduced number of astrophysical sources. In Figure 39 we show a table of parameter measurement precisions of supermassive black hole mergers for all submissions to the third round of the MLDC. The final two columns of the table show the fitting factor, i.e., overlap, of the submitted entry with the true source in each of the two independent LISA data channels,  $A$  and  $E$ .

The use of Bayesian techniques for searches as well as parameter estimation in the LISA context is motivated by the nature of the data. In the LIGO/Virgo context, most sources are

source (SNR <sub>true</sub> )	group	$\Delta M_c/M_c$ $\times 10^{-5}$	$\Delta \eta/\eta$ $\times 10^{-4}$	$\Delta t_c$ (sec)	$\Delta \text{sky}$ (deg)	$\Delta a_1$ $\times 10^{-3}$	$\Delta a_2$ $\times 10^{-3}$	$\Delta D/D$ $\times 10^{-2}$	SNR	FF <sub>A</sub>	FF <sub>E</sub>
<b>MBH-1</b> (1670.58)	AEI	2.4	6.1	62.9	11.6	7.6	47.4	8.0	1657.71	0.9936	0.9914
	CambAEI	3.4	40.7	24.8	2.0	8.5	79.6	0.7	1657.19	0.9925	0.9917
	MTAPC	24.8	41.2	619.2	171.0	13.3	28.7	4.0	1669.97	0.9996	0.9997
	JPL	40.5	186.6	23.0	26.9	39.4	66.1	6.9	1664.87	0.9972	0.9981
	GSFC	1904.0	593.2	183.9	82.5	5.7	124.3	94.9	267.04	0.1827	0.1426
<b>MBH-3</b> (847.61)	AEI	9.0	5.2	100.8	175.9	6.2	18.6	2.7	846.96	0.9995	0.9989
	CambAEI	13.5	57.4	138.9	179.0	21.3	7.2	1.5	847.04	0.9993	0.9993
	MTAPC	333.0	234.1	615.7	80.2	71.6	177.2	16.1	842.96	0.9943	0.9945
	JPL	153.0	51.4	356.8	11.2	187.7	414.9	2.7	835.73	0.9826	0.9898
	GSFC	8168.4	2489.9	3276.9	77.9	316.3	69.9	95.6	218.05	0.2815	0.2314
<b>MBH-4</b> (160.05)	AEI	4.5	75.2	31.4	0.1	47.1	173.6	9.1	160.05	0.9989	0.9994
	CambAEI	3.2	171.9	30.7	0.2	52.9	346.1	21.6	160.02	0.9991	0.9992
	MTAPC	48.6	2861.0	5.8	7.3	33.1	321.1	33.0	149.98	0.8766	0.9352
	JPL	302.6	262.0	289.3	4.0	47.6	184.5	28.3	158.34	0.8895	0.9925
	GSFC	831.3	1589.2	1597.6	94.4	59.8	566.7	95.4	-45.53	-0.1725	-0.2937
<b>MBH-2</b> (18.95)	AEI	1114.1	952.2	38160.8	171.1	331.7	409.0	15.3	20.54	0.9399	0.9469
	CambAEI	88.7	386.6	6139.7	172.4	210.8	130.7	24.4	20.36	0.9592	0.9697
	MTAPC	128.6	45.8	16612.0	8.9	321.4	242.4	13.1	20.27	0.9228	0.9260
	JPL	287.0	597.7	11015.7	11.8	375.3	146.3	9.9	18.69	0.9661	0.9709
<b>MBH-6</b> (12.82)	AEI	1042.3	1235.6	82343.2	2.1	258.2	191.6	26.0	13.69	0.9288	0.9293
	CambAEI	5253.2	1598.8	953108.0	158.3	350.8	215.4	29.4	10.17	0.4018	0.4399
	MTAPC	56608.7	296.7	180458.8	119.7	369.2	297.6	25.1	11.34	-0.0004	0.0016

Figure 39: Summary of the fractional errors in the recovery of parameters of the supermassive black hole binary mergers in the third MLDC data challenge. The final two columns, labelled FF<sub>A</sub> and FF<sub>E</sub>, give the overlap (or “fitting factor”) of the waveform corresponding to the recovered parameters with the true injected waveform. Each row represents a separate entry from one of the groups responding to the challenge. Table reproduced from Babak, S., et al., *Class. Quantum Grav.* **27**, 084009 (2010).

type <sup>1</sup>	$\nu$ (mHz)	$\mu/M_\odot$	$M/M_\odot$	$e_0$	$\theta_S$	$\varphi_S$	$\lambda$	$a/M^2$	SNR
True	0.1920421	10.296	9517952	0.21438	1.018	4.910	0.4394	0.69816	120.5
Found	0.1920437	10.288	9520796	0.21411	1.027	4.932	0.4384	0.69823	118.1
True	0.34227777	9.771	5215577	0.20791	1.211	4.6826	1.4358	0.63796	132.9
Found	0.34227742	9.769	5214091	0.20818	1.172	4.6822	1.4364	0.63804	132.8
True	0.3425731	9.697	5219668	0.19927	0.589	0.710	0.9282	0.53326	79.5
Found	0.3425712	9.694	5216925	0.19979	0.573	0.713	0.9298	0.53337	79.7
True	0.8514396	10.105	955795	0.45058	2.551	0.979	1.6707	0.62514	101.6
Found	0.8514390	10.106	955544	0.45053	2.565	1.012	1.6719	0.62534	96.0
True	0.8321840	9.790	1033413	0.42691	2.680	1.088	2.3196	0.65829	55.3
Found	0.8321846	9.787	1034208	0.42701	2.687	1.053	2.3153	0.65770	55.6
Blind									
True	0.1674472	10.131	10397935	0.25240	2.985	4.894	1.2056	0.65101	52.0
Found	0.1674462	10.111	10375301	0.25419	3.023	4.857	1.2097	0.65148	51.7
True	0.9997627	9.7478	975650	0.360970	1.453	4.95326	0.5110	0.65005	122.9
Found	0.9997626	9.7479	975610	0.360966	1.422	4.95339	0.5113	0.65007	116.0

Figure 40: Maximum a posteriori parameter values (labelled “Found”) recovered for all five EMRIs in the MLDC data set 1B (upper rows) and two additional random chosen sources. These are compared to the “True” parameters which were used to generate the injected signals. Table reproduced from Babak, S., Gair, J.R., and Porter, E.K., *Class. Quantum Grav.* **26**, 135004 (2009).

of short duration relative to the time between signals, and so it is necessary to efficiently sift through large amounts of data to find candidate sources of interest. In the LISA context, the source duration is comparable to the length of the data stream and so the entire data stream is relevant for the analysis of all sources. It is natural therefore to find and characterise sources simultaneously.

While the MLDCs demonstrated the effectiveness of the use of Bayesian methods to find and characterise most source types, several open questions remain, in particular related to the impact of non-stationary noise and instrumental artefacts such as gaps, the full extent of source confusion and the detection and characterisation of extreme-mass-ratio inspirals (EMRIs). While the EMRI sources in the MLDC data sets were successfully characterised under simplified assumptions (see Figure 41), the likelihood for an EMRI is very complicated, with many secondary maxima in parameter space. The successful algorithms relied on knowledge of the structure of the likelihood surface, which was specific to the simplified model of the EMRI employed in the MLDC, and the fact that all identified secondaries were generated by the same EMRI signal. While the structure of the likelihood surface can probably be learned for more accurate waveform models, the correct grouping of secondary modes will be much more challenging for real LISA data which could contain many hundreds of EMRIs.

Nested sampling has also been used in the context of LISA data analysis. In fact, the first application of the MULTINEST nested sampling algorithm in a gravitational wave context was to the characterisation of supermassive black hole mergers in LISA data (Feroz, F., Gair,

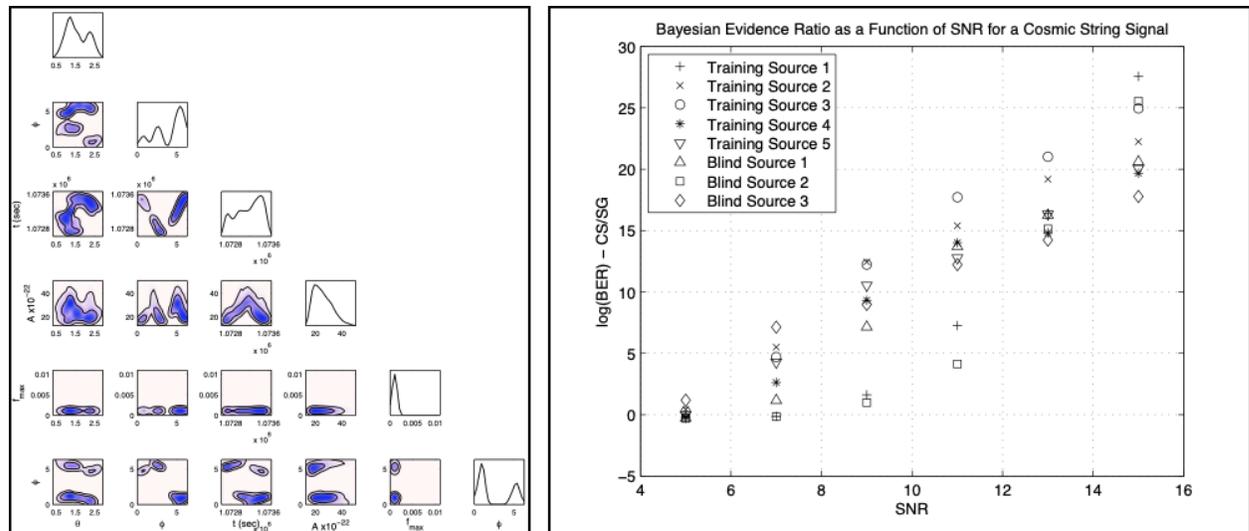


Figure 41: Left panel: posterior on the parameters characterising one of the cosmic string cusp gravitational wave bursts in the MLDC round 3 cosmic string data set. Right panel: evidence ratio in favour of the true (cosmic string cusp) model versus an alternative (sine-Gaussian) model for the burst, as a function of the burst signal-to-noise ratio. Figures reproduced from Feroz, F., Gair, J.R., Graff, P., Hobson, M.P., and Lasenby, A., *Class. Quantum Grav.* **27**, 075010 (2010).

J.R., Hobson, M.P., and Porter, E.K., *Class. Quantum Grav.* **26**, 215003). MULTINEST was also used to find and characterise supermassive black hole mergers and gravitational wave bursts from cosmic string cusps in MLDC data. In the latter case, the computed Bayesian evidences were used to test the hypothesis that the burst signals were consistent with a cosmic string cusp as opposed to a generic sine-Gaussian burst model (see Figure ?? and Feroz, F., Gair, J.R., Graff, P., Hobson, M.P., and Lasenby, A., *Class. Quantum Grav.* **27**, 075010 (2010)).

Further details on LISA data analysis can be found in the MLDC papers, and references therein:

- Arnaud, K.A., et al. *The Mock LISA Data Challenges: An overview*, *AIP Conf. Proc.* **873**, 619 (2006).
- Arnaud, K.A., et al., *A How-To for the Mock LISA Data Challenges*, *AIP Conf. Proc.* **873**, 625 (2006).
- Arnaud, K.A., et al., *Report on the first round of the Mock LISA Data Challenges*, *Class. Quantum Grav.* **24**, S529 (2007).
- Arnaud, K.A., et al., *An overview of the second round of the Mock LISA Data Challenges*, *Class. Quantum Grav.* **24**, S551 (2007).
- Babak, S., et al., *Report on the second Mock LISA Data Challenge*, *Class. Quantum Grav.* **25**, 114037 (2008).
- Babak, S., et al., *The Mock LISA Data Challenges: from Challenge 1B to Challenge 3*, *Class. Quantum Grav.* **25**, 184026 (2008).

- Babak, S., et al., *The Mock LISA Data Challenges: from Challenge 3 to Challenge 4*, *Class. Quantum Grav.* **27**, 084009 (2010).