# Lecture Recording

\* Note: These lectures will be recorded and posted onto the IMPRS website

- Dear participants,
- We will record all lectures on "Making sense of data: introduction to statistics for gravitational wave astronomy", including possible Q&A after the presentation, and we will make the recordings publicly available on the IMPRS lecture website at:
  - https://imprs-gw-lectures.aei.mpg.de/2021-making-sense-of-data/
- By participating in this Zoom meeting, you are giving your explicit consent to the recording of the lecture and the publication of the recording on the course website.

#### Making sense of data: introduction to statistics for gravitational wave astronomy Lecture 8: Examples of Bayesian statistics in GW data analysis

AEI IMPRS Lecture Course Jonathan Gair jgair@aei.mpg.de



#### Parameter Estimation in LIGO

## **LIGO Parameter Estimation**

- LIGO parameter estimation uses Bayesian methods. Results are quoted as posterior distributions, or posterior median values and credible intervals.
- We will show results from the first LVC catalogue, GWTC-1, for illustration. More recent results can be found in GWTC-2 (arxiv:2010.14527) and GWTC-3 (arxiv:2111.03606).

Event	$m_1/M_{\odot}$	$m_2/M_{\odot}$	$\mathcal{M}/M_{\odot}$	Xeff	$M_f/M_{\odot}$	$a_f$	$E_{\rm rad}/(M_\odot c^2)$	$\ell_{\rm peak}/({\rm erg}~{\rm s}^{-1})$	$d_L/{\rm Mpc}$	Z.	$\Delta\Omega/deg^2$
GW150914	$35.6^{+4.7}_{-3.1}$	$30.6^{+3.0}_{-4.4}$	$28.6^{+1.7}_{-1.5}$	$-0.01\substack{+0.12\\-0.13}$	$63.1_{-3.0}^{+3.4}$	$0.69\substack{+0.05 \\ -0.04}$	$3.1_{-0.4}^{+0.4}$	$3.6^{+0.4}_{-0.4}  imes 10^{56}$	$440^{+150}_{-170}$	$0.09\substack{+0.03 \\ -0.03}$	182
GW151012	$23.2^{+14.9}_{-5.5}$	$13.6^{+4.1}_{-4.8}$	$15.2^{+2.1}_{-1.2}$	$0.05\substack{+0.31 \\ -0.20}$	$35.6^{+10.8}_{-3.8}$	$0.67\substack{+0.13 \\ -0.11}$	$1.6^{+0.6}_{-0.5}$	$3.2^{+0.8}_{-1.7} \times 10^{56}$	$1080^{+550}_{-490}$	$0.21\substack{+0.09 \\ -0.09}$	1523
GW151226	$13.7^{+8.8}_{-3.2}$	$7.7^{+2.2}_{-2.5}$	$8.9^{+0.3}_{-0.3}$	$0.18\substack{+0.20 \\ -0.12}$	$20.5^{+6.4}_{-1.5}$	$0.74\substack{+0.07 \\ -0.05}$	$1.0\substack{+0.1\\-0.2}$	$3.4^{+0.7}_{-1.7}\times10^{56}$	$450^{+180}_{-190}$	$0.09\substack{+0.04 \\ -0.04}$	1033
GW170104	$30.8^{+7.3}_{-5.6}$	$20.0^{+4.9}_{-4.6}$	$21.4^{+2.2}_{-1.8}$	$-0.04\substack{+0.17\\-0.21}$	$48.9\substack{+5.1\\-4.0}$	$0.66\substack{+0.08\\-0.11}$	$2.2^{+0.5}_{-0.5}$	$3.3^{+0.6}_{-1.0}\times10^{56}$	$990_{-430}^{+440}$	$0.20\substack{+0.08 \\ -0.08}$	921
GW170608	$11.0^{+5.5}_{-1.7}$	$7.6^{+1.4}_{-2.2}$	$7.9^{+0.2}_{-0.2}$	$0.03\substack{+0.19 \\ -0.07}$	$17.8^{+3.4}_{-0.7}$	$0.69\substack{+0.04 \\ -0.04}$	$0.9\substack{+0.0\\-0.1}$	$3.5^{+0.4}_{-1.3}\times10^{56}$	$320^{+120}_{-110}$	$0.07\substack{+0.02 \\ -0.02}$	392
GW170729	$50.2^{+16.2}_{-10.2}$	$34.0^{+9.1}_{-10.1}$	$35.4_{-4.8}^{+6.5}$	$0.37\substack{+0.21 \\ -0.25}$	$79.5^{+14.7}_{-10.2}$	$0.81\substack{+0.07 \\ -0.13}$	$4.8^{+1.7}_{-1.7}$	$4.2^{+0.9}_{-1.5}\times10^{56}$	$2840^{+1400}_{-1360}$	$0.49\substack{+0.19 \\ -0.21}$	1041
GW170809	$35.0^{+8.3}_{-5.9}$	$23.8\substack{+5.1\\-5.2}$	$24.9^{+2.1}_{-1.7}$	$0.08\substack{+0.17 \\ -0.17}$	$56.3^{+5.2}_{-3.8}$	$0.70\substack{+0.08 \\ -0.09}$	$2.7\substack{+0.6 \\ -0.6}$	$3.5^{+0.6}_{-0.9}\times10^{56}$	$1030^{+320}_{-390}$	$0.20\substack{+0.05 \\ -0.07}$	308
GW170814	$30.6^{+5.6}_{-3.0}$	$25.2\substack{+2.8\\-4.0}$	$24.1^{+1.4}_{-1.1}$	$0.07\substack{+0.12 \\ -0.12}$	$53.2^{+3.2}_{-2.4}$	$0.72\substack{+0.07 \\ -0.05}$	$2.7\substack{+0.4 \\ -0.3}$	$3.7^{+0.4}_{-0.5}\times10^{56}$	$600^{+150}_{-220}$	$0.12\substack{+0.03 \\ -0.04}$	87
GW170817	$1.46\substack{+0.12 \\ -0.10}$	$1.27\substack{+0.09 \\ -0.09}$	$1.186\substack{+0.001\\-0.001}$	$0.00\substack{+0.02\\-0.01}$	$\leq 2.8$	$\leq 0.89$	$\geq 0.04$	$\geq 0.1 \times 10^{56}$	$40^{+7}_{-15}$	$0.01\substack{+0.00\\-0.00}$	16
GW170818	$35.4_{-4.7}^{+7.5}$	$26.7^{+4.3}_{-5.2}$	$26.5^{+2.1}_{-1.7}$	$-0.09\substack{+0.18\\-0.21}$	$59.4_{-3.8}^{+4.9}$	$0.67\substack{+0.07 \\ -0.08}$	$2.7\substack{+0.5 \\ -0.5}$	$3.4^{+0.5}_{-0.7}\times10^{56}$	$1060^{+420}_{-380}$	$0.21\substack{+0.07 \\ -0.07}$	39
GW170823	$39.5^{+11.2}_{-6.7}$	$29.0_{-7.8}^{+6.7}$	$29.2_{-3.6}^{+4.6}$	$0.09\substack{+0.22 \\ -0.26}$	$65.4^{+10.1}_{-7.4}$	$0.72\substack{+0.09 \\ -0.12}$	$3.3^{+1.0}_{-0.9}$	$3.6^{+0.7}_{-1.1} \times 10^{56}$	$1940^{+970}_{-900}$	$0.35\substack{+0.15 \\ -0.15}$	1666

LVC GWTC-1 (2019)

## LIGO PE codes

- In O1 and O2, LIGO parameter estimation used the LALInference code. This includes two separate algorithms
  - *LALInferenceMCMC*: A Markov Chain Monte Carlo code based on the Metropolis-Hastings algorithm. Proposal distributions are tuned to features of the likelihood expected for CBC inspirals.
  - \* *LALInferenceNest*: A bespoke nested sampling algorithm. New live points are drawn by evolving mini-MCMC chains until an independent point is obtained.
- During O3 a new software package, *Bilby*, was introduced (also with a parallel implementation, *parallel bilby*). The sampling algorithms in *Bilby* are not bespoke.
  Instead it uses freely available packages such as *dynesty*. LALInference was also used in O3, alongside *Bilby*. It is anticipated that Bilby will be the primary inference package for O4 onwards.







LVC GWTC-1 (2019)



LVC GWTC-1 (2019)





 Bayesian methods are also used for tests of general relativity, e.g., to place bounds on pPN deviations in the observed waveform.



LVC TGR with GWTC-2 (2020)

## **Population Inference**

# **Population Inference**

- \* LIGO employs Bayesian hierarchical models to constrain the parameters of the astrophysical population from which the sources are drawn.
- Examples
  - Cosmological parameter inference: estimation of the Hubble constant or other cosmological parameters from sets of events (see lecture 6).
  - Rate estimation: estimation of the rate of mergers of different types occurring in the Universe, and its evolution with redshift.
  - Source population properties: inference on the distribution of masses and spins of black holes etc.

## Rate Estimation

- Alternative approach to rate estimation simultaneously model foreground and background distributions and try to measure rates.
- Data is a set of observed statistic values, *x<sub>i</sub>*, e.g., max template SNR, evidence etc.
  Each event has an (unknown) flag, *f<sub>i</sub>*, labelling it as either foreground (*f<sub>i</sub>* =1) or background (*f<sub>i</sub>* =0).

$$d = \{x_i | i = 1, \dots, N\}$$

\* Foreground and background events are Poisson distributed with rates

$$\frac{\mathrm{d}N_f}{\mathrm{d}x} = R_f \hat{f}(x,\theta_f) \qquad \qquad \frac{\mathrm{d}N_b}{\mathrm{d}x} = R_b \hat{b}(x,\theta_b)$$

\* and corresponding cumulative distributions  $\hat{F}(x, \theta_f)$ ,  $\hat{B}(x, \theta_b)$ .

#### Rate Estimation



## Rate estimation: GW150914

\* Rate estimation requires foreground distribution. This was complicated for GW150914 by the presence of another, lower significance event, LVT151012.



#### Rate estimation: GWTC-1



## Rate estimation: GWTC-3

 With more events, can simultaneously constrain rate and more sophisticated population models.

	BNS	NSBH	BBH	NS-Gap	BBH-gap	Full	
	$m_1 \in [1, 2.5] M_{\odot}$	$m_1 \in [2.5, 50] M_{\odot}$	$m_1 \in [2.5, 100] M_{\odot}$	$m_1 \in [2.5, 5] M_{\odot}$	$m_1 \in [2.5, 100] M_{\odot}$	$m_1 \in [1, 100] M_{\odot}$	
	$m_2 \in [1, 2.5] M_{\odot}$	$m_2 \in [1, 2.5] M_{\odot}$	$m_2 \in [2.5, 100] M_{\odot}$	$m_2 \in [1, 2.5] M_{\odot}$	$m_2 \in [2.5, 5] M_{\odot}$	$m_2 \in [1, 100] M_{\odot}$	
PDB (pair)	$960^{+1700}_{-700}$	$59^{+81}_{-38}$	$25^{+10}_{-7}$	$41^{+69}_{-30}$	$9.3^{+19.0}_{-7.6}$	$1100^{+1700}_{-750}$	
PDB (ind)	$250^{+640}_{-200}$	$170_{-89}^{+150}$	$22^{+9}_{-6}$	$29^{+55}_{-23}$	$10^{+15}_{-8}$	$470_{-300}^{+830}$	
MS	$470^{+1400}_{-410}$	$57^{+120}_{-42}$	$42^{+88}_{-20}$	$3.7^{+20}_{-3.4}$	$0.17\substack{+56 \\ -0.16}$	$650^{+1600}_{-460}$	
BGP	$99^{+260}_{-86}$	$32^{+62}_{-25}$	$33^{+16}_{-10}$	$2.1^{+33}_{-2.1}$	$5.1^{+12}_{-4.0}$	$180^{+260}_{-110}$	
Merged	13 - 1900	7.4 - 320	16-130	0.029 - 84	0.01 - 56	71 - 2200	
	$m_1 \in [5, 20] M_{\odot}$		$m_1 \in [20, 50] M_{\odot}$	$m_1$	All BBH		
	$m_2 \in [5, 20] M_{\odot}$		$m_2 \in [5, 50] M_{\odot}$	$m_2$			
PP	23.4	$4^{+12.9}_{-8.6}$	$4.5^{+1.8}_{-1.3}$		$28.1^{+14.8}_{-10.0}$		
BGP	20.0	$)^{+10.0}_{-8.0}$	$6.4^{+3.0}_{-2.1}$		$33.0^{+16.0}_{-10.0}$		
$\mathrm{FM}$	21.1	$1^{+10.7}_{-8.3}$	$4.1^{+2.0}_{-1.4}$		$26.0^{+11.5}_{-8.7}$		
PS	27	+12 -9.4	$3.6^{+1.5}_{-1.1}$		$32^{+14}_{-9.6}$		
Merged	12.8	8 - 40	0.098 - 6.3		17.3 - 45		

- Infer mass distribution of black holes by using hierarchical models with different forms for the mass prior.
- Simplest model is a power law

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

- LIGO analysis of GWTC-1 used two variants
  - A: minimum mass fixed to 5 solar masses, flat in mass ratio;
  - B: all parameters allowed to vary.
- \* The LVC GWTC-1 analysis also used a *power law+peak* model designed to identify an excess of black holes at the edge of the pair-instability supernova mass gap

$$p(m_1|\theta) = \left[ (1 - \lambda_m) A(\theta) m_1^{-\alpha} \Theta(m_{\max} - m_1) + \lambda_m B(\theta) \exp\left(-\frac{(m_1 - \mu_m)^2}{2\sigma_m^2}\right) \right] S(m_1, m_{\min}, \delta m)$$
$$p(q|m_1, \theta) = C(m_1, \theta) q^{\beta_q} S(m_2, m_{\min}, \delta m).$$







- Using the 76 events that comprise GWTC-3 it was possible (and necessary) to fit a wider variety of models (arxiv:2111.03634). These included one additional parametric model (power-law + dip + peak), and three non-parametric models.
  - power-law + spline: modify the truncated power-law mass function by an arbitrary perturbation modelled as a cubic spline

 $p(m_1|\alpha, m_{\min}, m_{\max}, \delta_m, \{f_i\}) = k \ p(m_1|\alpha, m_{\min}, m_{\max}, \delta_m) \exp[f(m_1|\{f_i\})]$ 

flexible mixtures: model the distribution as a sum of separable components.
 Model the distribution over masses and spins simultaneously.

 $p(\mathcal{M}, q, s_{1z}, s_{2z}|\theta) = \sum_{i=1}^{N} w_i G(\mathcal{M}|\mu_i^{\mathcal{M}}, \sigma_i^{\mathcal{M}}) G(s_{1z}|\mu_i^{sz}, \sigma_i^{sz}) G(s_{2z}|\mu_i^{sz}, \sigma_i^{sz}) P(q|\alpha_i^q, q_i^{\min}, 1)$ 

- **binned Gaussian process**: construct bins in the mass-mass plane, regard the rates in each bin as unknown parameters, which are connected through the specification of a Gaussian process mean and covariance.



LVC BBH properties from GWTC-3 (2021)



LVC BBH properties from GWTC-3 (2021)

## Population inference: spins

\* Describe spin magnitude distribution using a Beta distribution (support is in the desired range [0,1]).  $\alpha^{\alpha_a - 1}(1 - \alpha)\beta_a - 1$ 

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



## Population inference: spins

p(

 Can also use a non-parametric model where the fraction of spin magnitudes in different bins are the hyperparameters, e.g., 3-bin model.

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



## Population inference: spins

Probe binary formation mechanisms by constraining spin distribution as a combination of an isotropic component, and a preferentially aligned component.  $p(\cos t_1, \cos t_2 | \sigma_1, \sigma_2, \zeta) = \frac{(1-\zeta)}{4}$ • Gaussian (G):  $\zeta = 1$ .  $+\frac{2\zeta}{\pi}\prod_{i\in\{1,2\}}\frac{\exp\left(-(1-\cos t_i)^2/(2\sigma_i^2)\right)}{\sigma_i \operatorname{erf}(\sqrt{2}/\sigma_i)}$ • Mixture (M):  $0 \le \zeta \le 1$ . 2.0 GWTC-2 1.2Mixture GWTC-3 Gaussian 1.01.5Isotropic  $b(\cos\theta) = \frac{1}{2} \int \frac{1}{2} d\theta$ 0.8 $p(\cos t_1)$ 0.40.50.20.0 + -1.0 $0.0 \ -1.0$ 0.5 0.0-0.51.0 -0.50.00.51.0 $\cos t_1$  $\cos\theta$ LVC BBH properties from GWTC-1 (2019) LVC BBH properties from GWTC-3 (2021)

#### $\cos\theta$ .

#### Population inference: mass-spin correlations

 Probe correlations between parameters by fitting joint distributions, or allowing model parameters to depend on other parameters. In the analysis of GWTC-3, the LVK explored the variation of the spin distribution with mass ratio.



# Population inference: rate evolution

 Constrain evolution of the rate of BBH mergers by including redshift dependence in the rate model

$$\frac{\mathrm{d}R}{\mathrm{d}\xi} \left( z|\theta \right) = R_0 p(\xi|\theta) (1+z)^{\lambda}$$

 Generate combined constraints on the rate evolution exponent and the parameters of the simple power-law mass function (model A).



# Population inference: rate evolution

\* GWTC-3 shows some evidence for rate evolution, with the rate higher in the past.



- Bayesian inference is also used for model selection based on the evidence ratio or Bayes factor.
- \* Example applications to LIGO include
  - test for presence / absence of gravitational radiation after the merger of a binary neutron star;
  - test of GW polarisation tensor versus scale or vector polarisations.

#### LVC, Phys. Rev. X 9 011001 (2019)

We determine the relative evidence for two models: that the on-source data are described by Gaussian noise only, or by Gaussian noise plus a GW signal as described in Refs. [151,152]. We find that the Gaussian noise model is strongly preferred, with a Bayes factor (evidence ratio) of 256.79 over the signal model. This result is consistent with both prompt collapse to a BH and with a postmerger signal that is too weak to be measurable with our current sensitivity. We further characterize the absence of a

#### LVC, Phys. Rev. D **100** 104036 (2019)

The Bayes factors (marginalized likelihood ratios) obtained in this case are  $12 \pm 3$  for tensor vs vector and  $407 \pm 100$  for tensor vs scalar, where the error corresponds to the uncertainty due to discrete sampling in the evidence computations. These values are comparable to those from GW170814, for which the latest recalibrated and cleaned data (cf. Sec. II) yield Bayes factors of  $30 \pm 4$  and  $220 \pm 27$  for tensor vs vector and scalar respectively.<sup>22</sup> Values from these binary black holes are many orders of magnitude weaker than those obtained from GW170817, where we benefited from the precise sky-localization provided by an electromagnetic counterpart [8].

 Tests of no-hair property: use evidence to quantify evidence for ringdown modes different to those predicted by general relativity.





Brito, Buonanno, Raymond (2019)

\* Another example: *Probing dynamical gravity with the polarisation of continuous gravitational waves, Isi et al.* (2017) *Phys. Rev. D* **96** 042001. (not yet a LIGO search).



#### Source reconstruction



- \* The *BayesWave* pipeline uses a Bayesian non-parametric approach to reconstruct noise and signal components from the data.
- \* The smooth noise PSD component is modelled using a cubic spline.
- \* Lines in the instrumental noise are modelled using Lorentzian functions.

$$p(x; b, m) = \frac{1}{\pi} \frac{b}{(x-m)^2 + b^2}$$

- The remaining components of the data are modelled using *wavelets*, which resolve time series at particular times and frequencies. *BayesWave* uses the Morley-Gabor basis.
- \* There is a coherent wavelet component for sources and incoherent components to represent glitches.



sky location, orientation, etc.

morphological params.





# Rapid localisation





- Bayesian techniques are also used to obtain rapid sky localisation of GW transients to send triggers to astronomers for EM follow-up.
- Bayestar employs the autocorrelation likelihood (likelihood evaluated at MLE parameter values)

$$\exp\left[-\frac{1}{2}\sum_{i}\rho_{i}^{2}+\sum_{i}\rho_{i}\Re\left\{e^{-i\gamma_{i}}z_{i}^{*}(\tau_{i})\right\}\right]$$

- Rapid marginalisation over parameters other than sky location achieved via integral approximation and look-up tables.
- \* Result is a sky map probability density.





## Bayestar



#### LALInference Burst

 LALInference Burst is another tool for rapid source localisation. The signal is modelled as a sine-Gaussian, coherent between detectors.

$$h_{x}(t) = \sin(\alpha) \frac{h_{rss}}{\sqrt{Q(1 - \cos(2\phi_o) e^{-Q^2})/4f_o\sqrt{\pi}}}$$
  
× sin (2\pi f\_o(t - t\_o) + \phi\_o) e^{-(t - t\_o)^2/\tau^2}

$$h_{+}(t) = \cos(\alpha) \frac{h_{rss}}{\sqrt{Q(1 + \cos(2\phi_o) e^{-Q^2})/4 f_o \sqrt{\pi}}} \\ \times \cos(2\pi f_o(t - t_o) + \phi_o) e^{-(t - t_o)^2/\tau^2}.$$

- Figure of merit (left) is sky area searched before true sky location identified.
- Bayestar and LIB are not true Bayesian algorithms since model of data generation process is approximated.



# oLIB

- There is also an online version of LIB, that mixes frequentist and Bayesian techniques.
- LIB is used to compute Bayes factors for the signal versus noise hypothesis (BSN) and for coherent versus incoherent triggers across detectors (BCI).
- oLIB, along with CWB, were the first algorithms to detect GW150914, as they were the online online algorithms running at the start of O1.

