Astrophysical inference and transient gravitational wave astronomy

Greg Ashton

Monash Centre for Astrophysics

Monash University

OzGrav

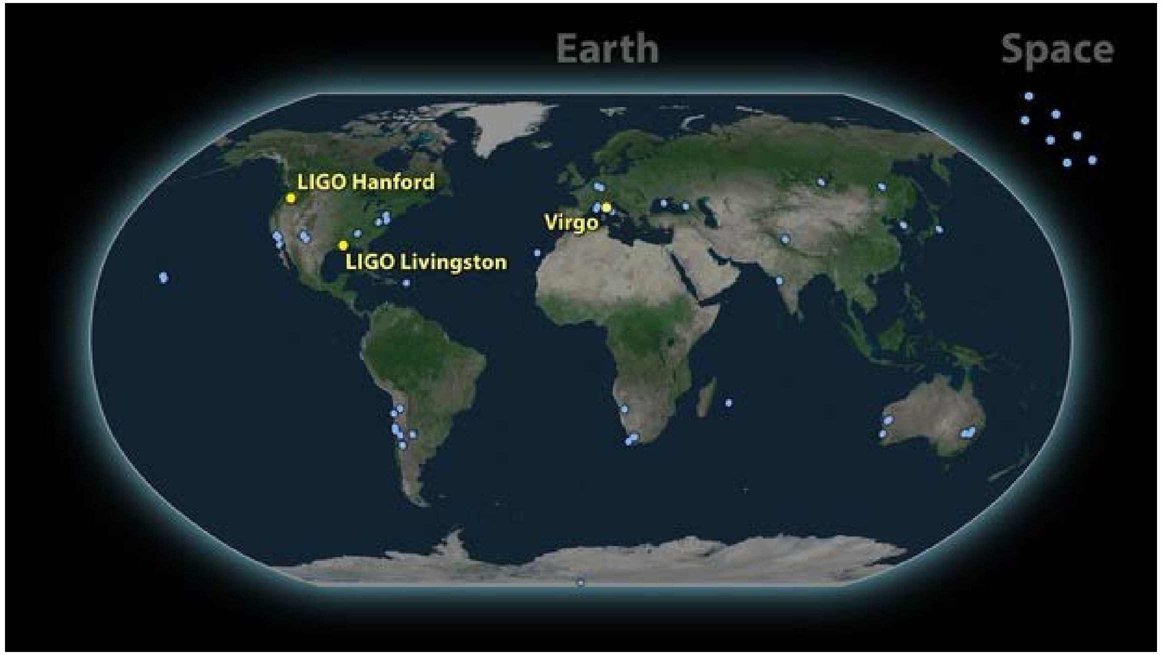
Talk overview

1. Transient gravitational wave astronomy
2. The tools of inference
3. Better understanding individual events
4. Better understanding a population of events
5. Inference as a detection tool
6. Introducing Bilby: the “Bayesian Inference Library”

Transient gravitational wave astronomy

Gravitational wave astronomy

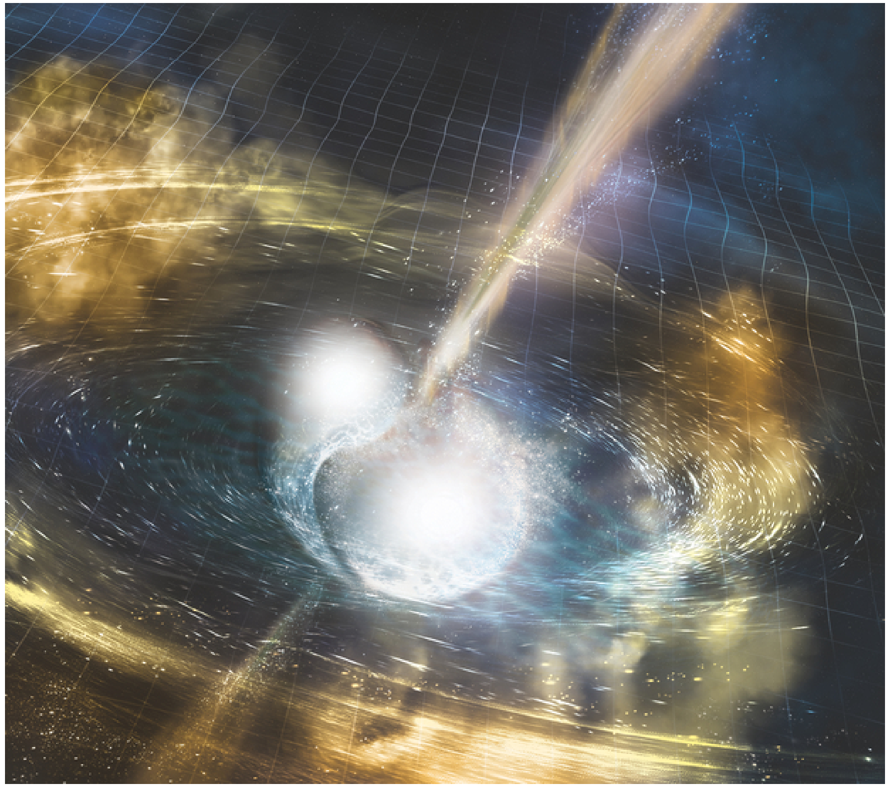
- Gravitational wave astronomy has already enjoyed much success
- Three ground-based interferometers
- Working in tandem with multi-messenger probes:
  - A multitude of telescopes spanning the electromagnetic spectrum
  - Neutrino detectors



[LIGO-Virgo]

Transient gravitational waves

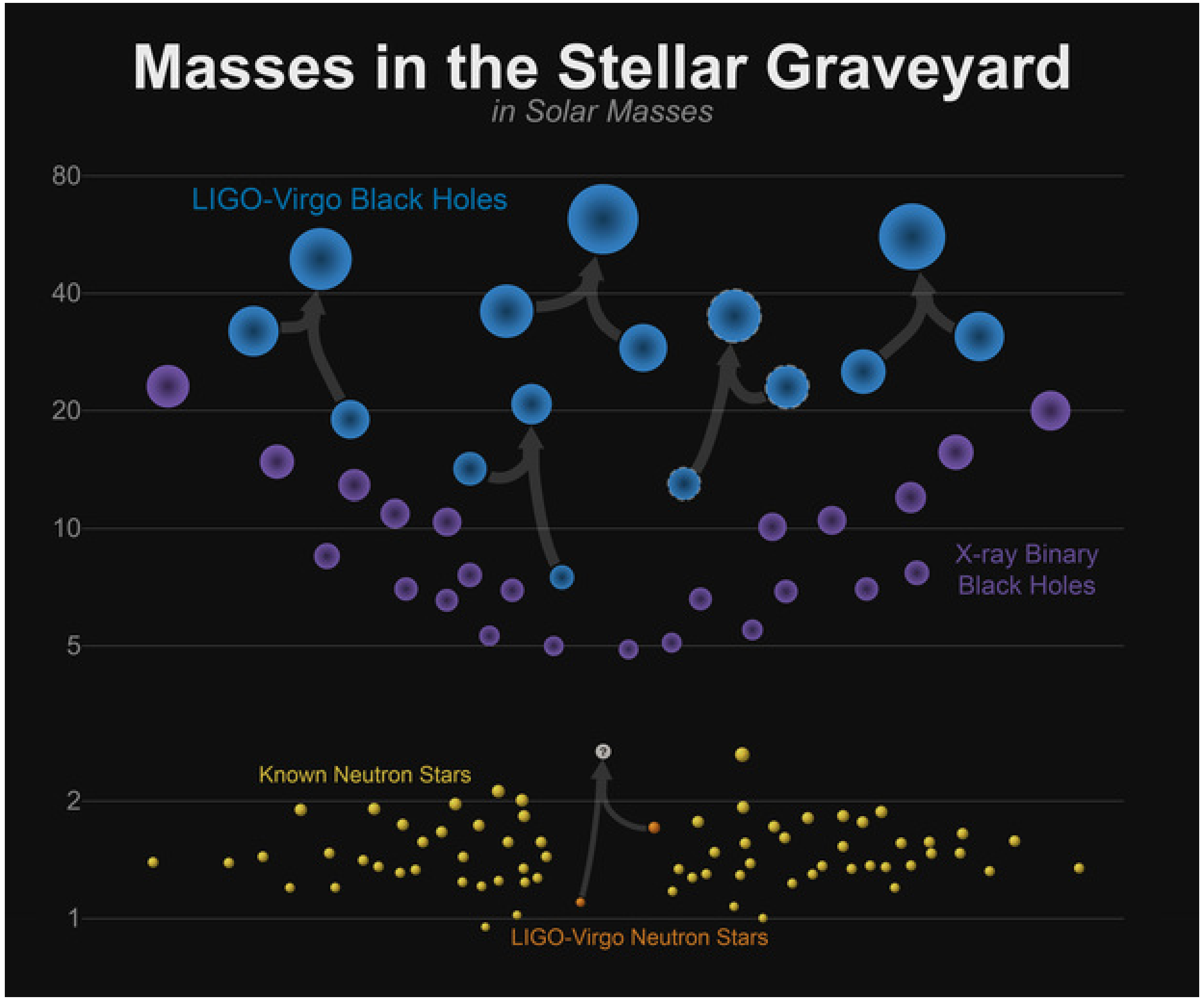
- Caused by catastrophic events
- Primarily considering CBC or *Compact Binary Coalescence*
- Typically "observable" for a fractions of a second up to a few tens of seconds
- Extra galactic in origin



NSF/LIGO/Sonoma State University/A. Simonnet

The view so far

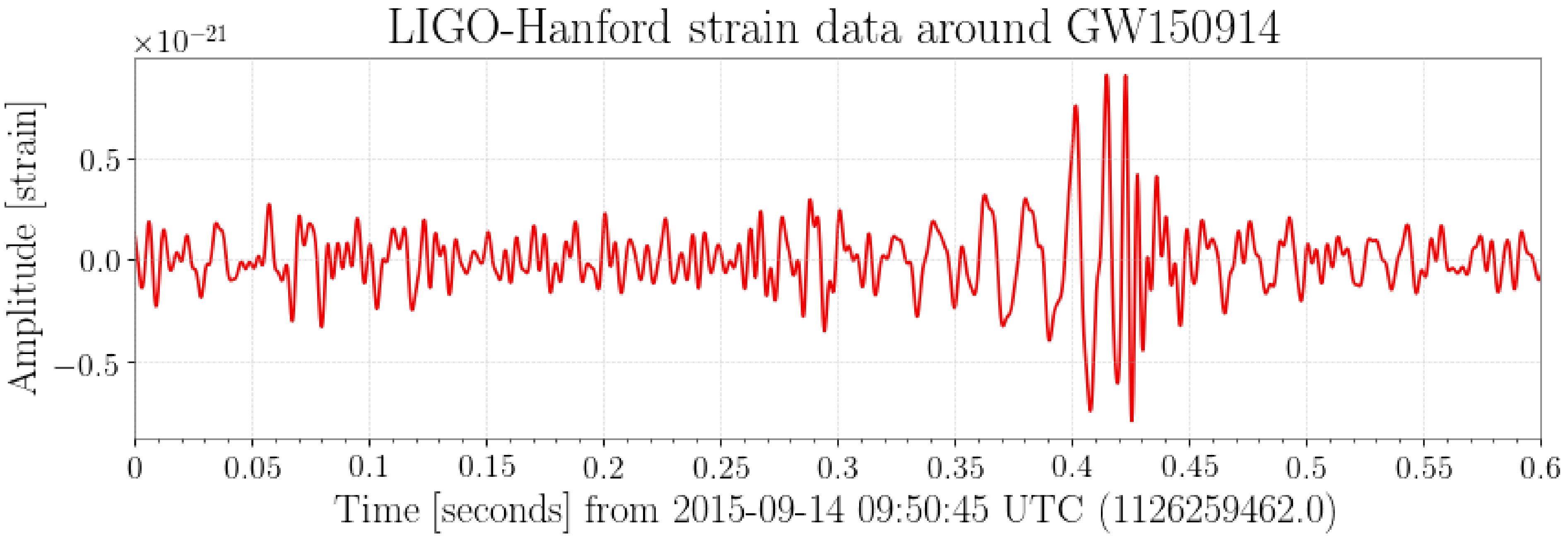
- After 2 observing runs
- 5 binary black hole mergers
- 1 binary neutron star merger:
  - Joint detection with Fermi-GBM/Integral
  - Subsequent detections across the spectrum
- 1 “LVT” (LIGO/Virgo trigger)



Credit: [LIGO-Virgo/Frank Elavsky/Northwestern University]

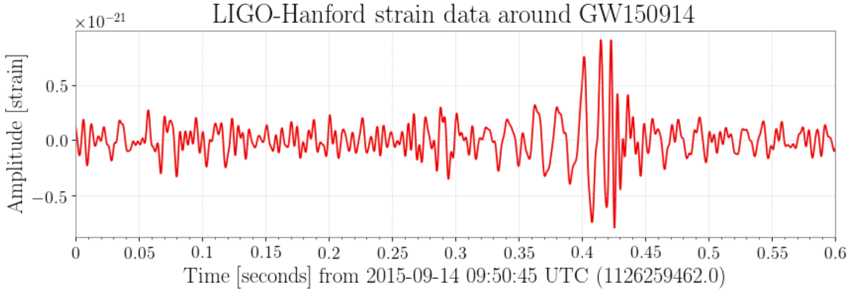
The tools of inference

Getting physics out of the interferometer



Source: [gwpy documentation]

Getting physics out of the interferometer



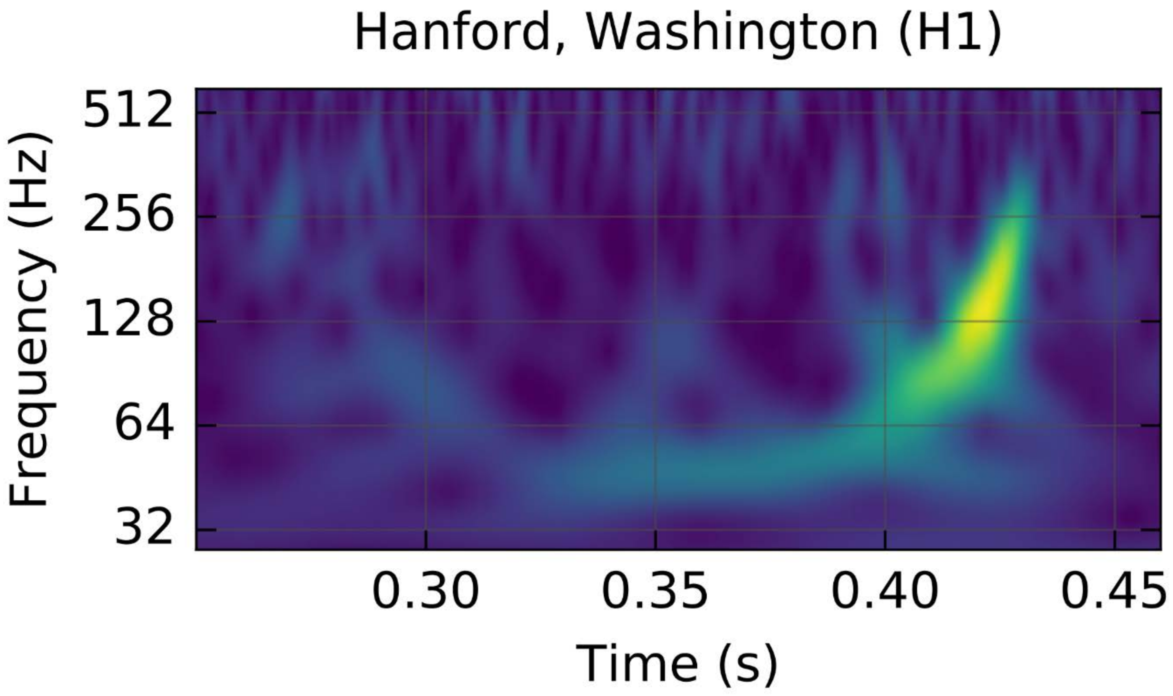
???

TABLE I. Source parameters for GW150914. We report median values with 90% credible intervals that include statistical errors, and systematic errors from averaging the results of different waveform models. Masses are given in the source frame; to convert to the detector frame multiply by $(1 + z)$ [90]. The source redshift assumes standard cosmology [91].

Primary black hole mass	$36^{+5}_{-4} M_\odot$
Secondary black hole mass	$29^{+4}_{-4} M_\odot$
Final black hole mass	$62^{+4}_{-4} M_\odot$
Final black hole spin	$0.67^{+0.05}_{-0.07}$
Luminosity distance	410^{+160}_{-180} Mpc
Source redshift z	$0.09^{+0.03}_{-0.04}$

Abbott et al (2016), PRL 116

From basic physics



Abbott et al (2016), PRL 116

- Abott et al (2016) Ann. Phys. 529
- Use time-frequency plot and peak amplitude
- Can obtain estimates such as

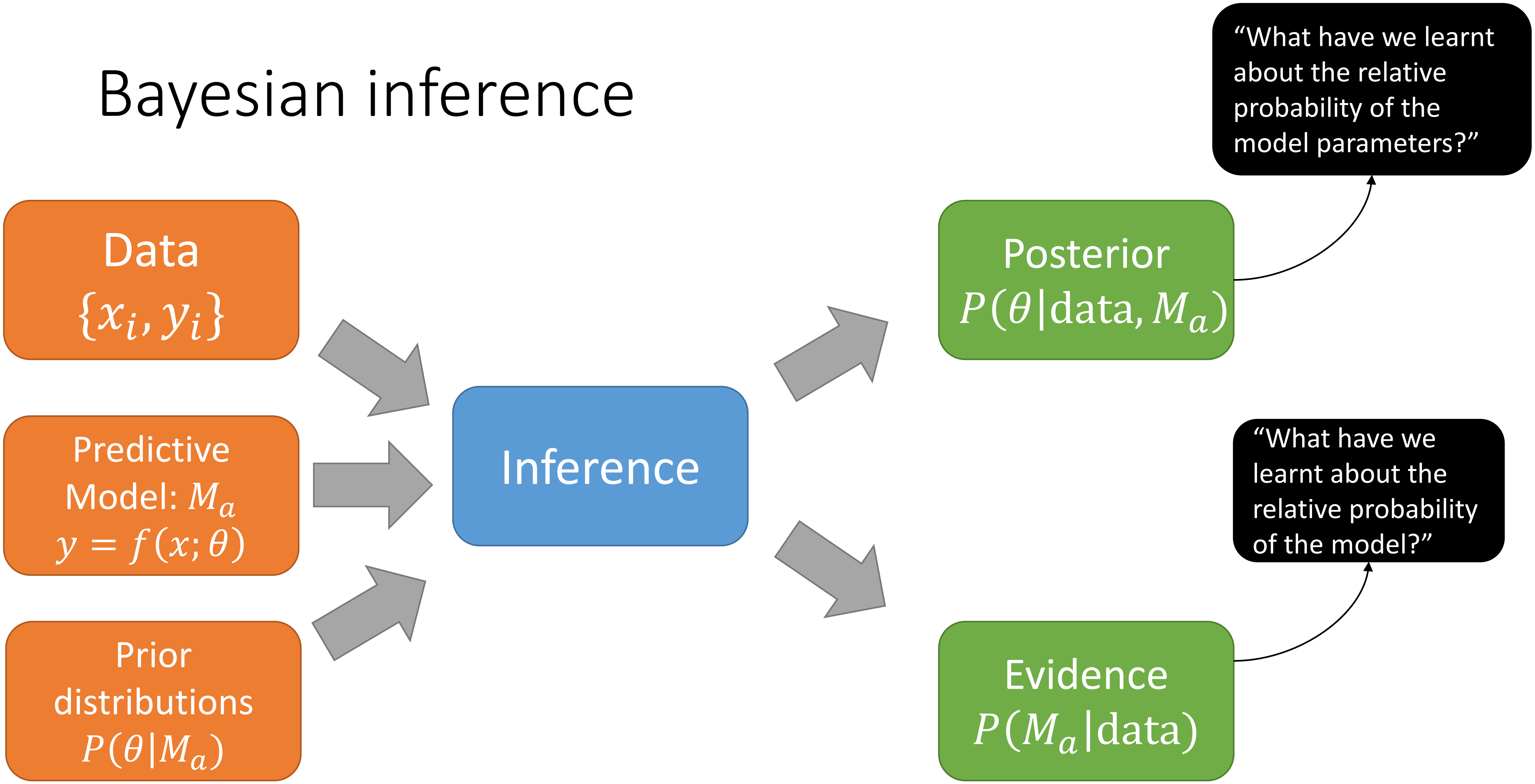
$$d_L \sim 300 \text{ Mpc}$$

- Not particularly accurate and difficult to estimate uncertainties

Bayesian inference

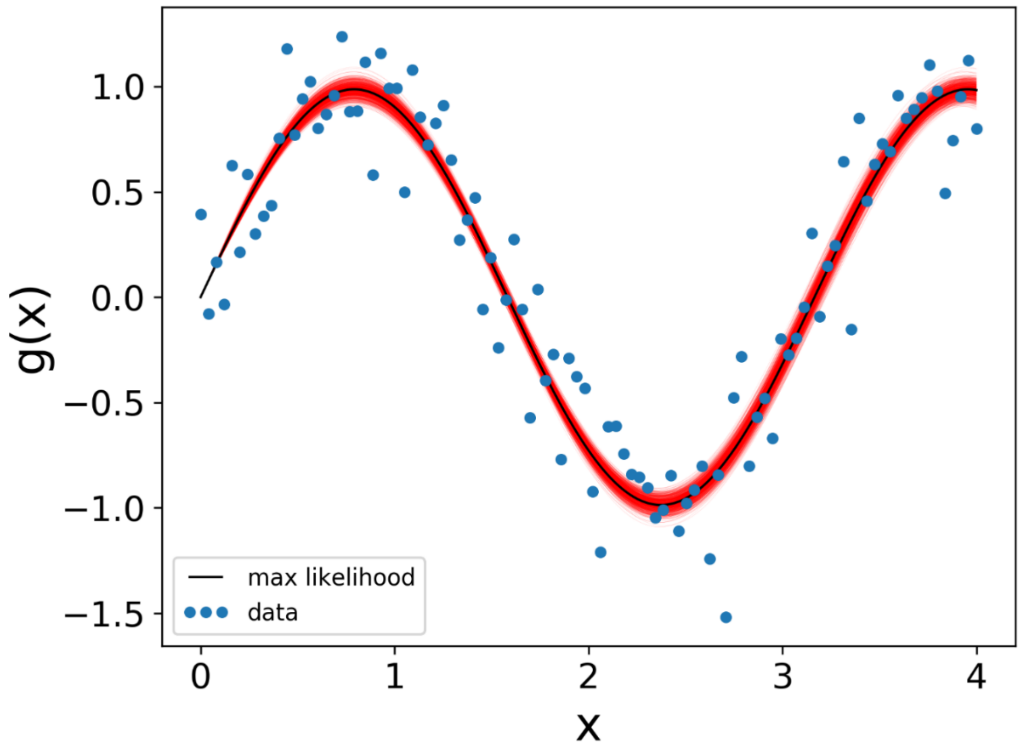
- Since we have uncertainties, we need to use probabilistic language
- Bayes rule: $P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$
- At its core, this is a statement of conditional probabilities
- But it provides tool set to:
    - Understand how well a given model fits the data
    - Compare between models

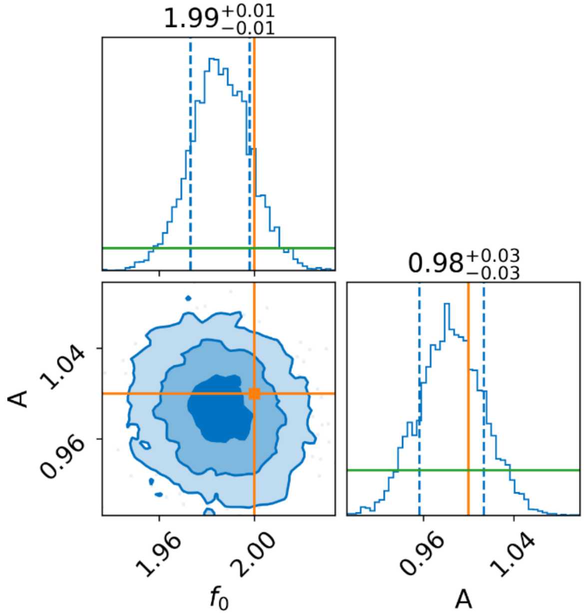
Bayesian inference



Simple example

- Model: $g(x; A, f_0) = A\sin(f\ x)$
- Data: simulate with some fixed values
- Priors:
  - $A \sim \mathrm{Uniform}(A_{\mathrm{min}}, A_{\mathrm{max}})$
  - $f_0 \sim \mathrm{Uniform}(f_{0\,\mathrm{min}}, f_{0\,\mathrm{max}})$
- Output:
  - Posteriors (corner plots)
  - Posterior predictive plots
  - Evidence (compare to other models)





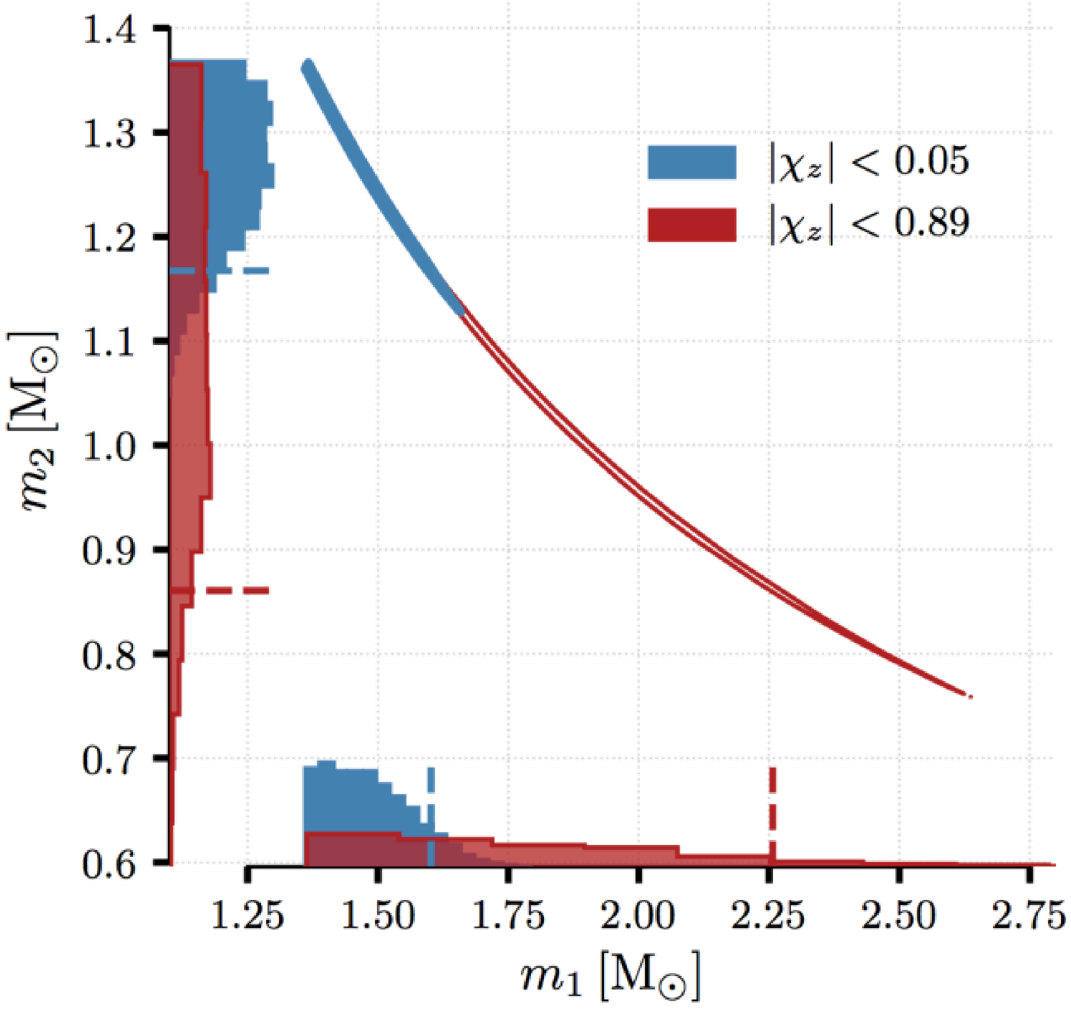
Inference for gravitational waves

- Fundamentally no different, but…
- Model of gravitational waveform
  - Typically has $\sim$15 parameters
  - Phenomological waveform tuned to numerical relativity
  - Slow to evaluate (typically $\sim$0.01 seconds/evaluation)
- Must also model the response of the detectors
  - E.g., frequency-dependent sensitivity
  - For long-lived signals this includes Doppler modulations

Better understanding individual events

Inference for individual events

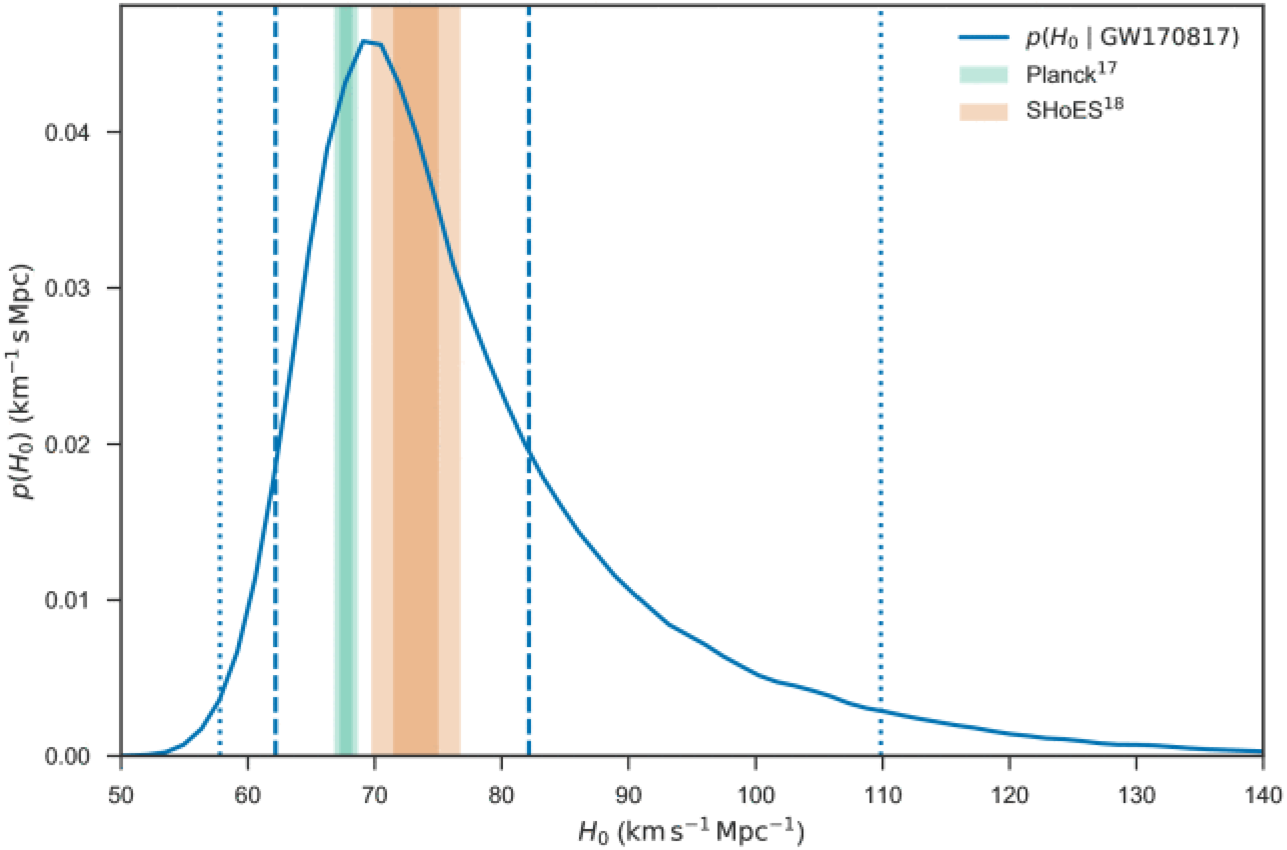
- GW170817 (the binary neutron star system)
- Masses clearly support that this is a pair of binary neutron stars
- Some wiggle room for a neutron star + black hole (but ruled out by other observations)



Abbott et al (2017) PRL 119

Inference for individual events

- GW170817 (the binary neutron star system)
- EM follow-up enabled a host-galaxy identification
- Giving an estimate of the distance
- Provides independent constraint on the Hubble constant



[Abbott et al (2017) PRL 119]()

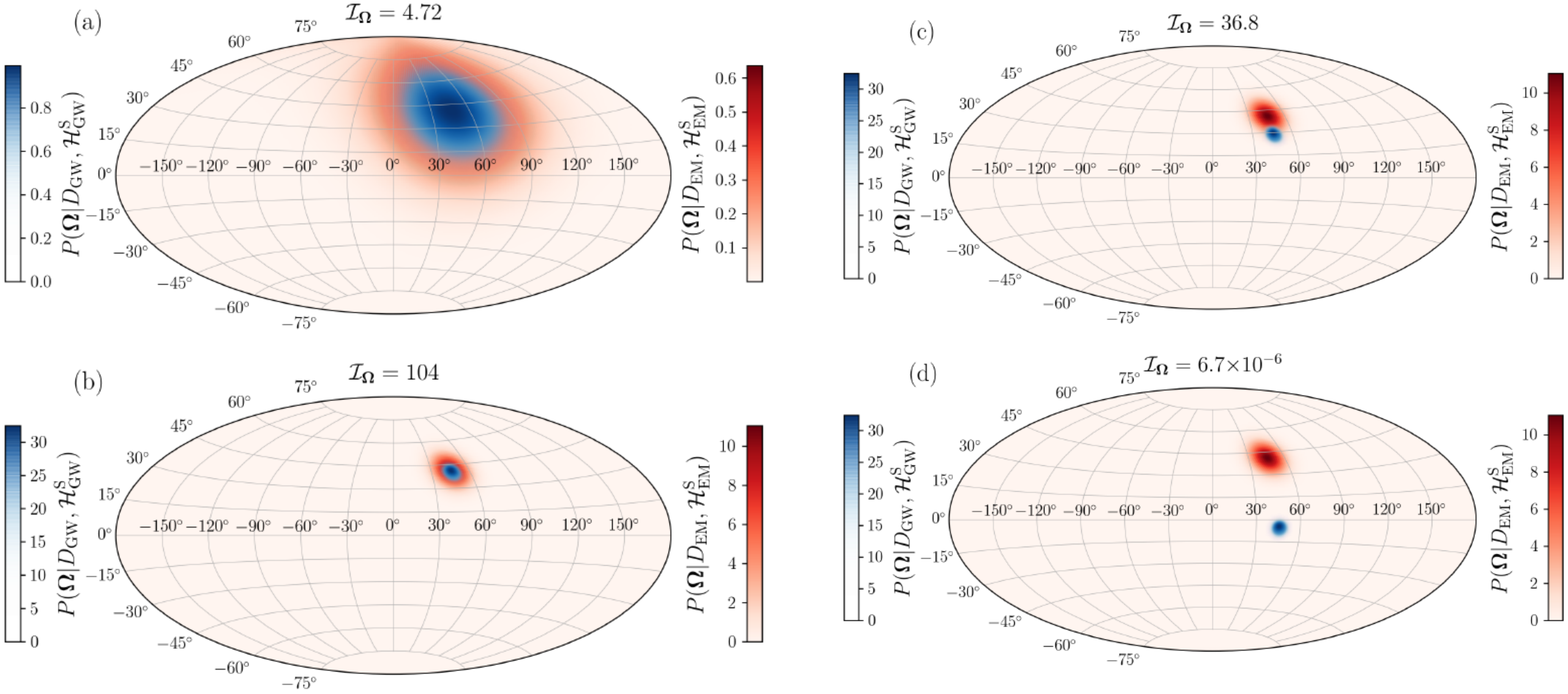
Joint inference between distinct data sets

- Given two data sets A and B both with "events" detected in them
- How do we decide if the events share a common origin?
- Generally, we look for common-parameters θ
- Ashton et al. (2018) APJ 860:

$$\frac{P(\mathrm{A}|\mathrm{Common})}{P(\mathrm{A}|\mathrm{Noise})} \times \frac{P(\mathrm{B}|\mathrm{Common})}{P(\mathrm{B}|\mathrm{Noise})} \times \int \frac{P(\theta|A)P(\theta|B)}{P(\theta)} d\theta = \frac{P(\mathrm{A,B}|\mathrm{Common})}{P(\mathrm{A,\,B}|\mathrm{Noise})}$$

Evidence for event in d_A × Evidence for event in d_B × Posterior overlap I_θ = Evidence for a common event

Example: sky-overlap



- Application to GW170817 and GRB 170817A favours a common signal by a factor of $\sim 10^6$
- Agreement with the original analysis of [Abbott et al (2017) APJL 848]

Better understanding a population of events

Hierarchical modelling

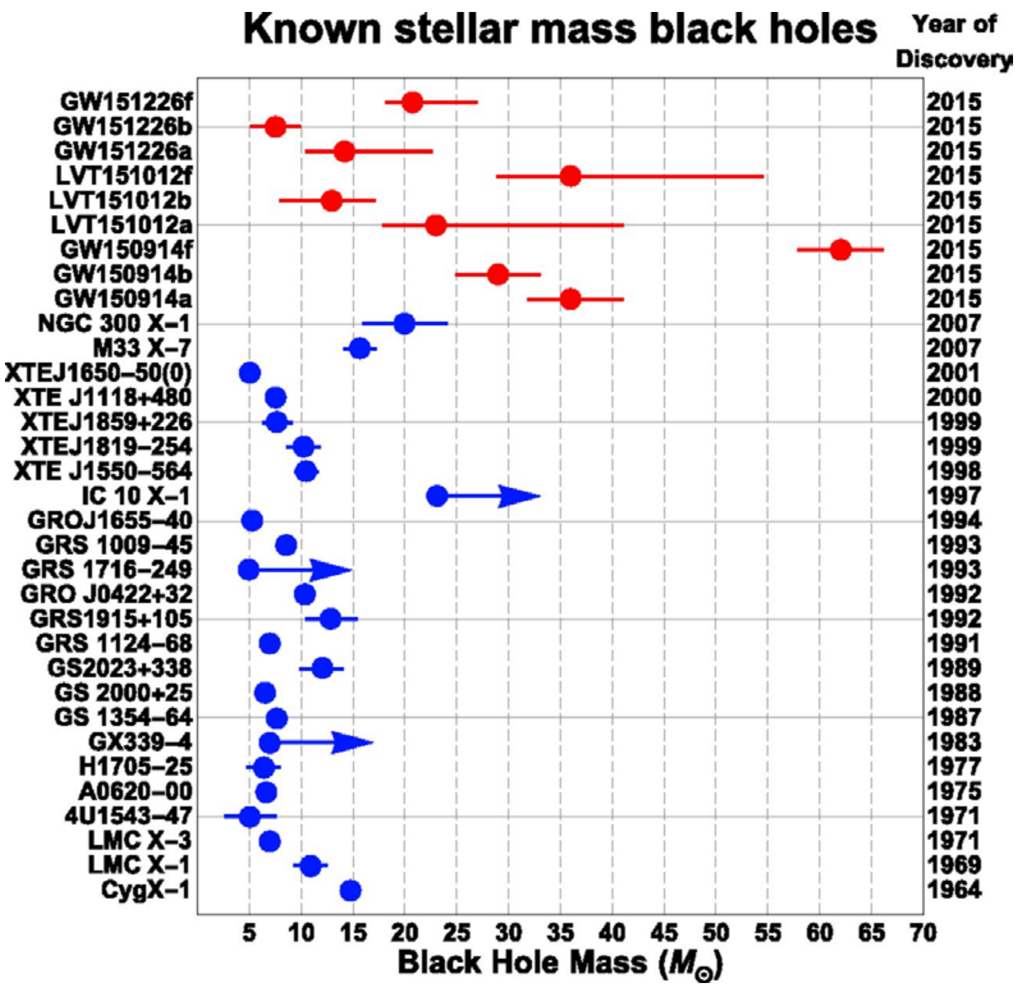
Hierarchical models

<u>Individual inference</u>

- Given $d_1, d_2, d_3 \ldots, d_N$ independent data sets
- Infer model parameters: θ for each data set :$P(\theta|d_i)$

<u>Hyperparameter inference</u>

- Infer hyperparameters Λ
- Hyperparameters model population-level effects
- Requires a predictive model



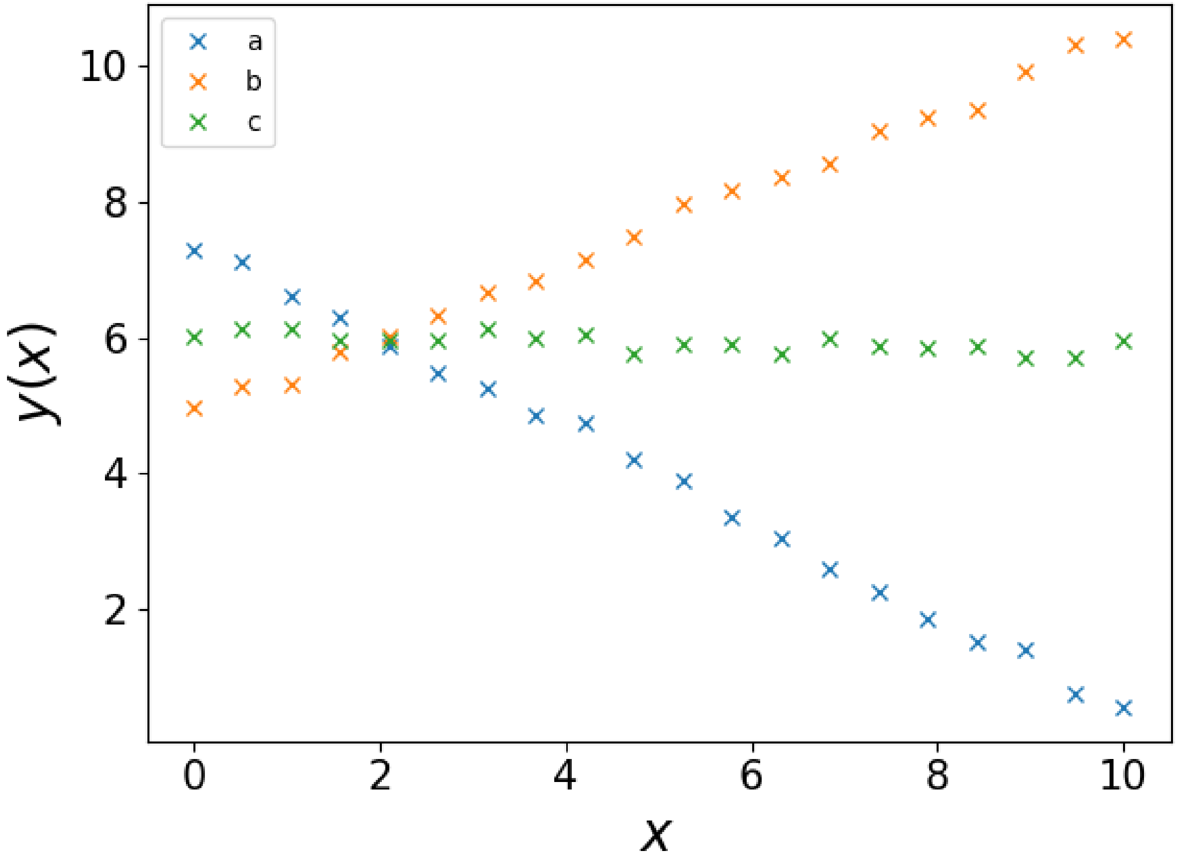
[Kalogera, Lazzarini (2017)](#)

Hierarchical modelling: toy example

- Three data sets: a, b, and c
- For the ith individual data set:

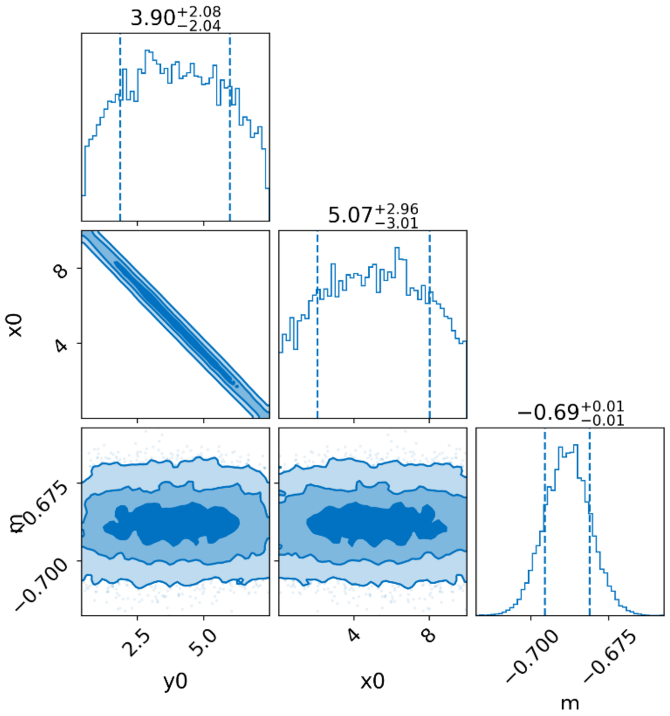
$$y(x) = y_0 + m_i\,(x - x_0)$$

- Where y_0, x_0 are population-parameters while m_i is not
- How can we infer x_0 and y_0 ?
- Note: x_0 and y_0 are degenerate

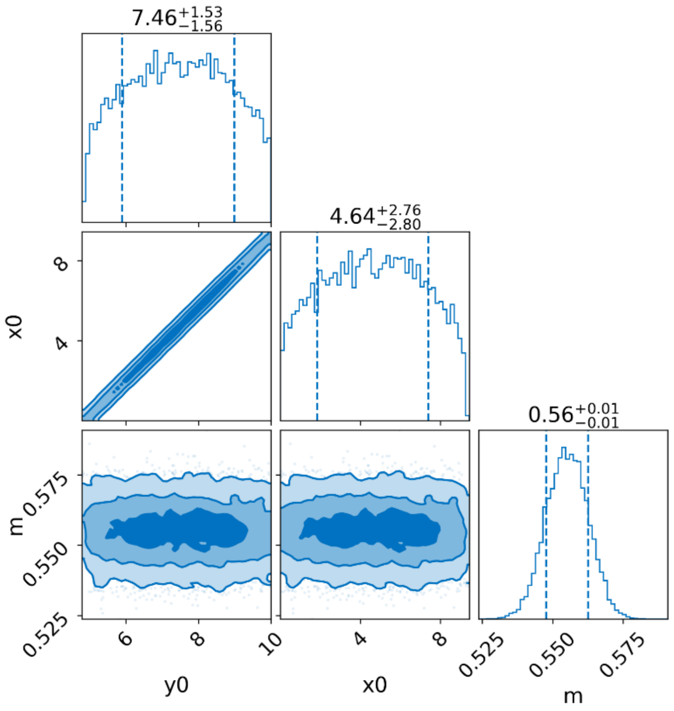


Individual posteriors

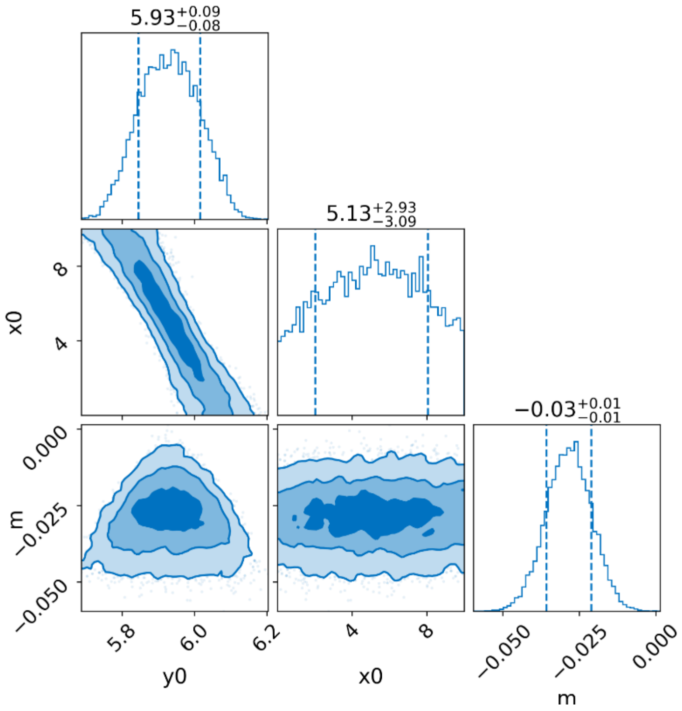
For an individual event, the gradient is well measured but x_0 and y_0 are not



Posterior for data A



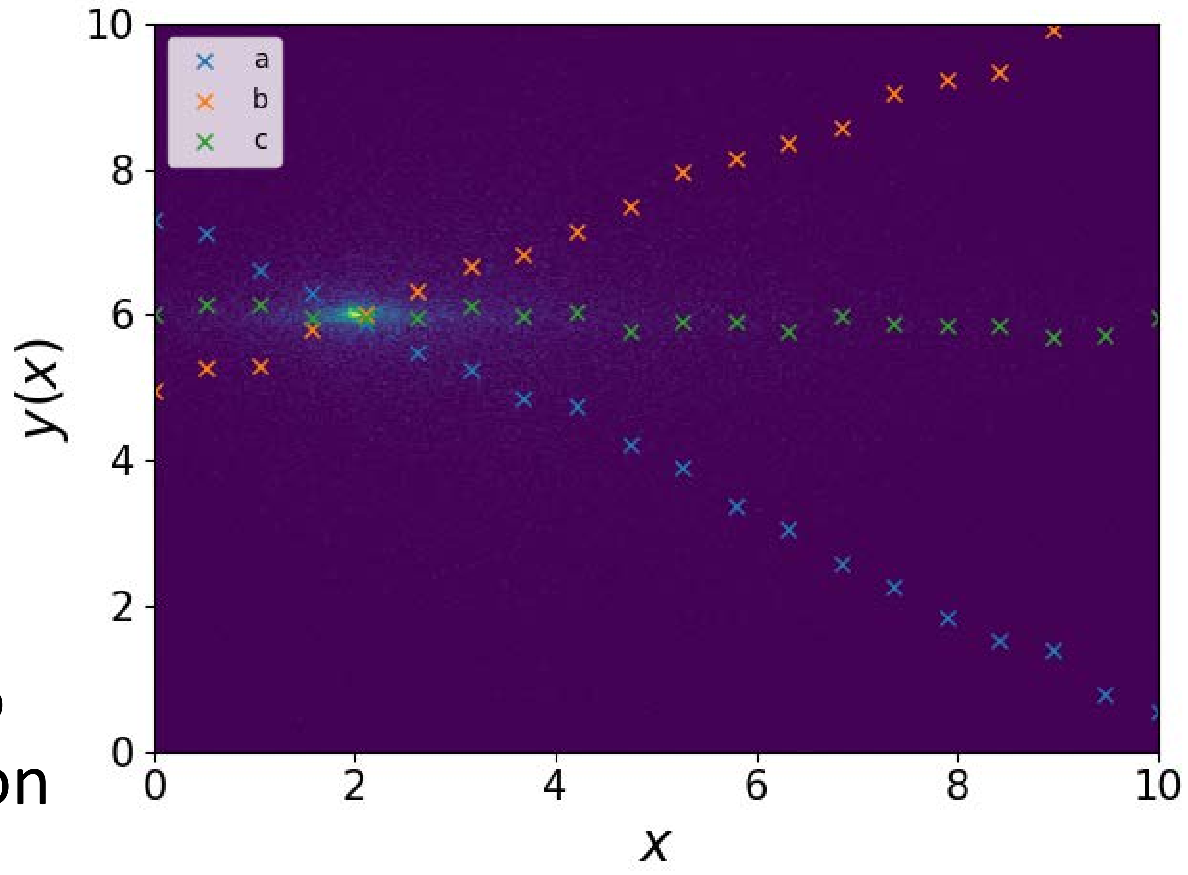
Posterior for data B



Posterior for data C

Hyperparameter inference

- We can model the population behaviour as follows:
- Let:
  - $x_0 \sim \mathrm{Normal}(\mu_{x_0}, \sigma_{x_0})$
  - $y_0 \sim \mathrm{Normal}(\mu_{y_0}, \sigma_{y_0})$
- Then infer the posterior on the hyperparameters: $\mu_{x_0}, \sigma_{x_0}, \mu_{y_0}, \sigma_{y_0}$
- Finally look at the predicted position for x_0, y_0



Applying hierarchical models to gravitational wave transients

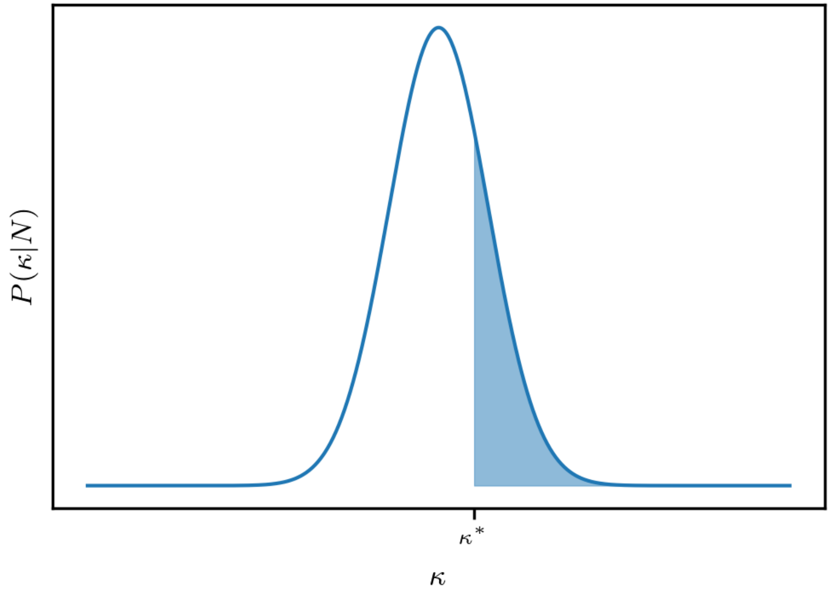
- Use the observed events to determine population properties
  - E.g., Talbot & Thrane (2017) PRD 96
  - Determine formation channels and identify sub-populations
- Constrain formation rates and test cosmology
- Search for an astrophysical background of events
  - Smith & Thrane (2018) PRX 8
  - Time to detect the background reduced from 40 months (using traditional cross-correlation) to 1 day of design-sensitivity data

Inference as a detection tool

Inference for detection

- Detection methods currently rely on frequentist method:
  - Define background distribution (time-slides)
  - Compute a p-value (probability of event or louder give background)
  - Small p-value ⇒ unlikely to be background
- This has some issues
  - Calculation of background is difficult/expensive
  - Saturation, e.g., Was et al (2009) arXiv:0906.2120
  - How to incorporate this in a Bayesian framework

$$p = \int_{\kappa *}^{\infty} P(\kappa |\, N)\, d\kappa$$



$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

???

Time-slides

Hyperparameterised odds

- In a normal setting, we ask

"what is the evidence for a signal in the i$^{\text{th}}$ segment?"

- Amounts to calculating

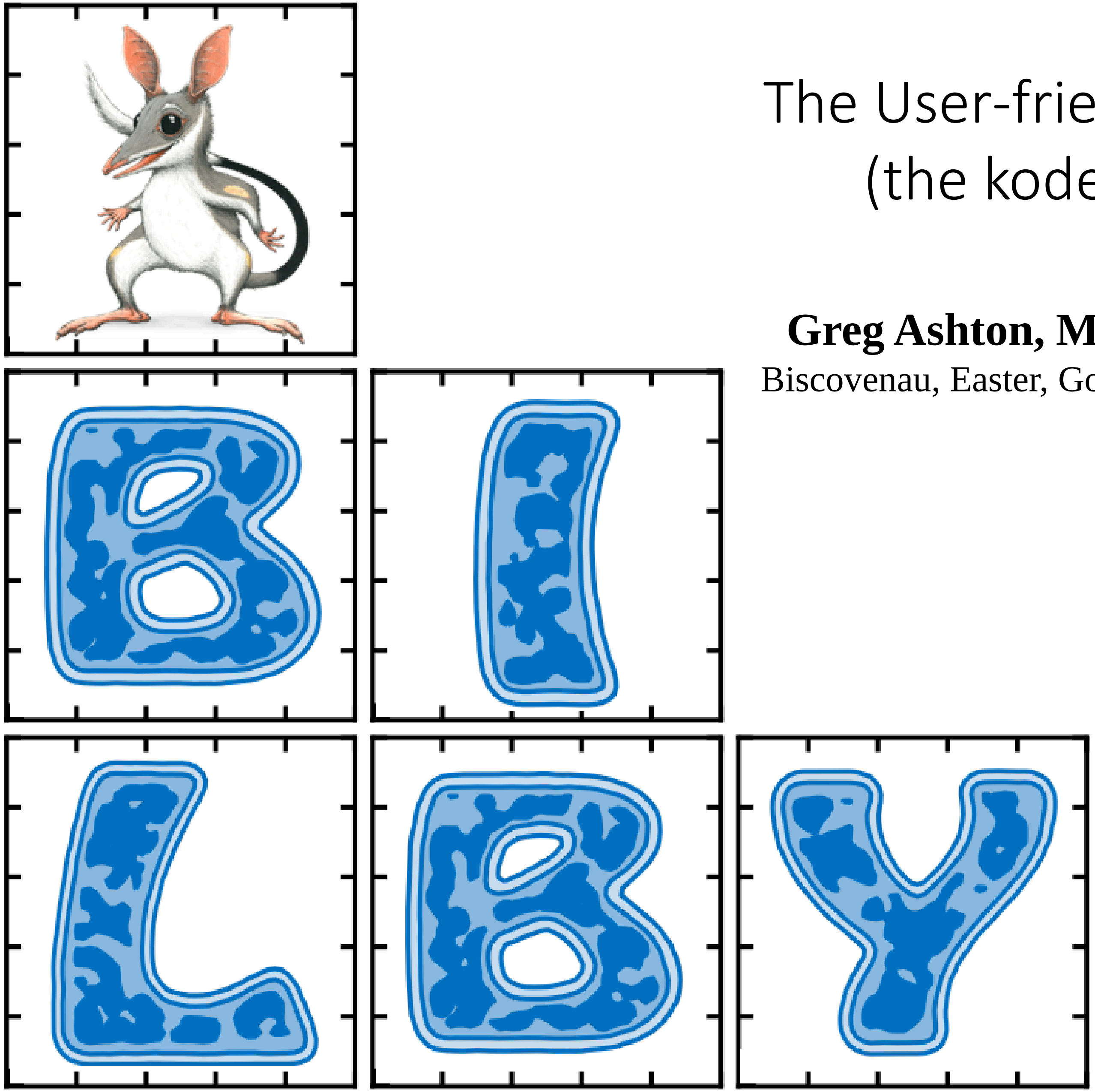
$$O_{S_i/N_i}(d_i) = \frac{P(S_i \mid d_i)}{P(N_i \mid d_i)} = \frac{P(d_i|S_i)P(S_i)}{P(d_i|N_i)P(N_i)}$$

- Instead, we want to ask

"what is the evidence for a signal in the i$^{\text{th}}$ segment, given **all** the data?"

$$O_{S_i/N_i}(\boldsymbol{d}) = \frac{P(S_i \mid \boldsymbol{d})}{P(N_i \mid \boldsymbol{d})} = \frac{\int d\Lambda\, P(d_i|S_i,\Lambda)P(S_i|\Lambda)\, P(\Lambda|\boldsymbol{d}_{i\neq k})}{\int d\Lambda\, P(d_i|N_i,\Lambda)P(N_i|\Lambda)\, P(\Lambda|\boldsymbol{d}_{i\neq k})}$$

Ashton, Smith & Thrane (in prep.)



The User-friendly Bayesian inference library (the kode formerly known as tupak)

Greg Ashton, Moritz Hübner, Paul Lasky, Colm Talbot
Biscovenau, Easter, Goncharov, Lower, Payne, Powell, Sarin, Smith, Thrane

- General code for inference
- Capable of handling arbitrary data/models
- Multiple samplers
- gw-specific and hyperparameter modules
- www.monash.docs.ligo.org/bilby/

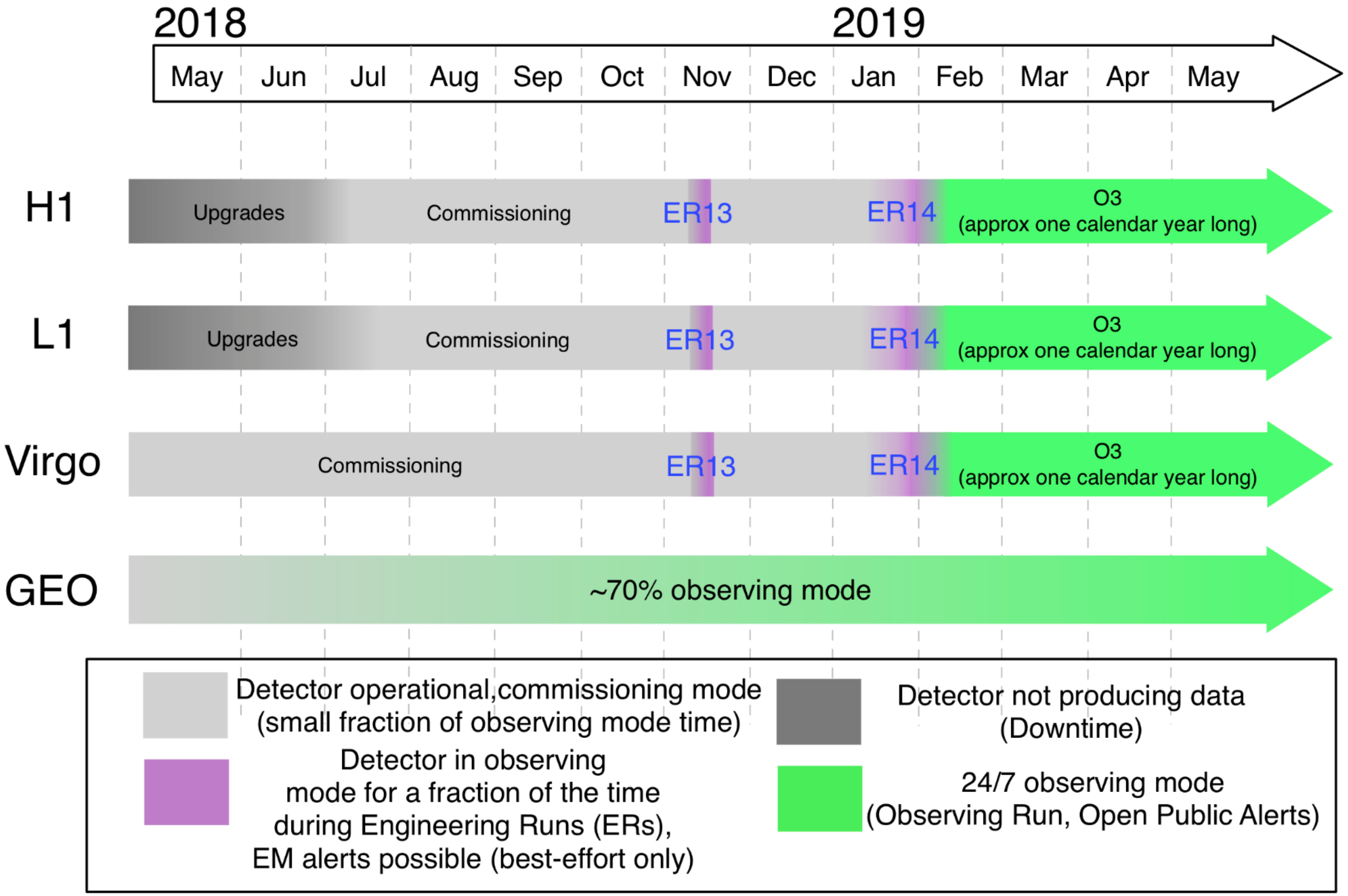
Overview

- Inference provides a universal tool for astrophysics
- Can help us understand individual events
- Hierarchical modelling can be used to
  - Understand populations
  - Detect backgrounds of events
  - Detect events
- Bilby: an OzGrav effort to build the next generation of inference tools

LIGO-VIRGO Joint Run Planning Committee

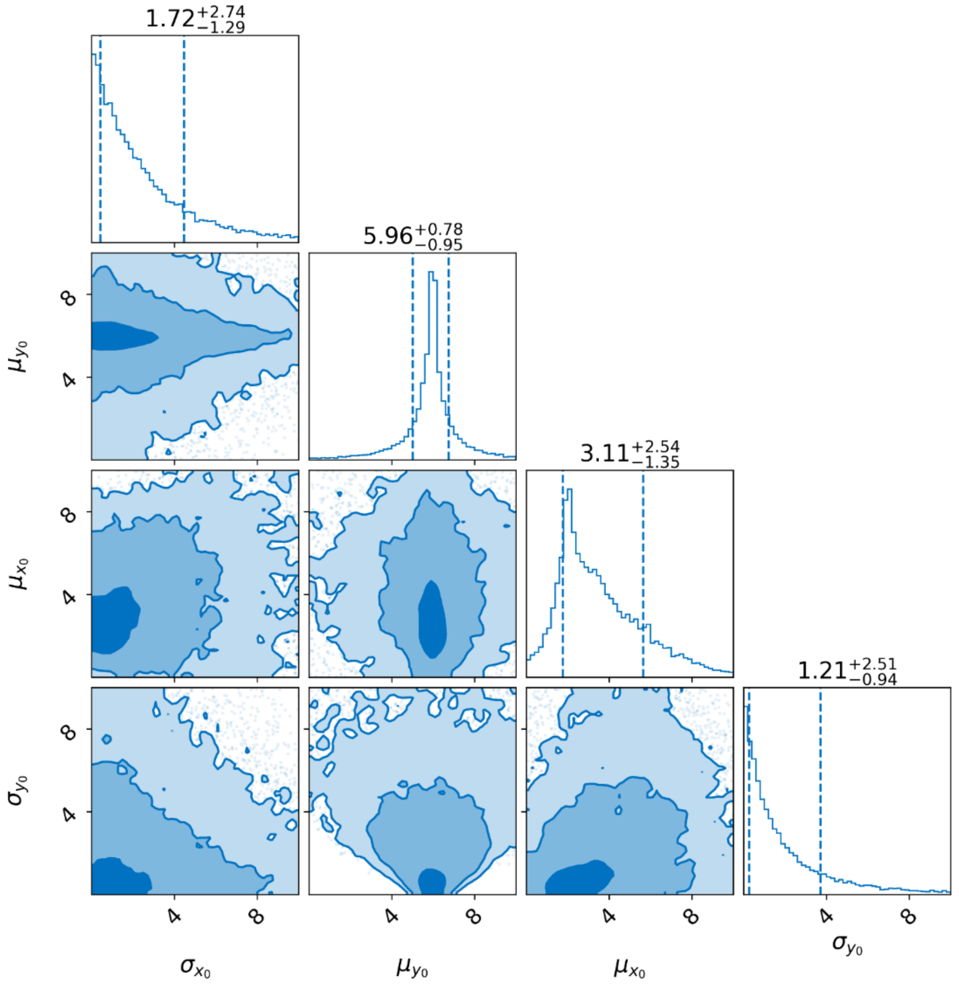
Working schedule for O3

(Public document G1801056-v2, based on G1800889-v5)



Credit: LIGO-Virgo

Posterior for hyperparameter example



Background and significance