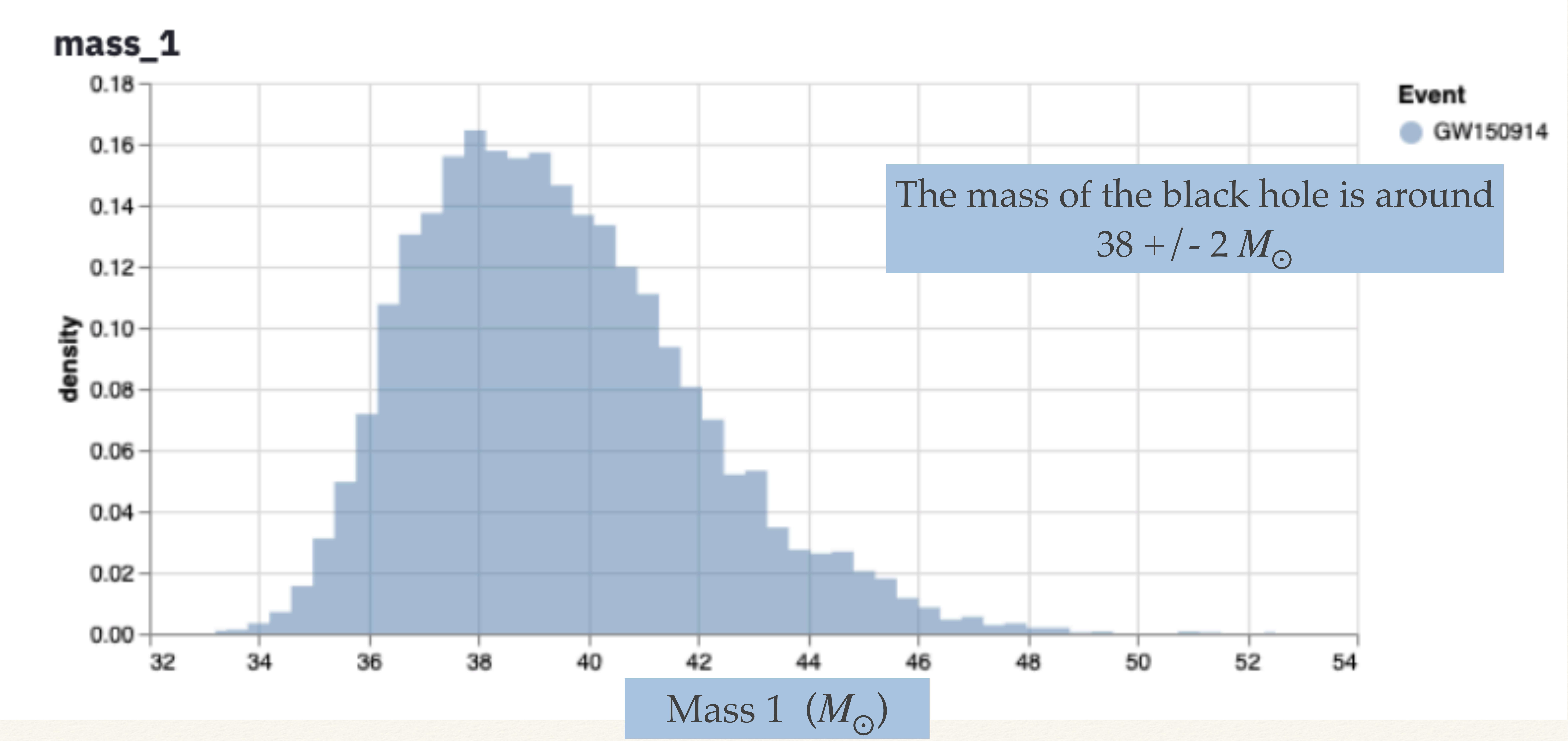
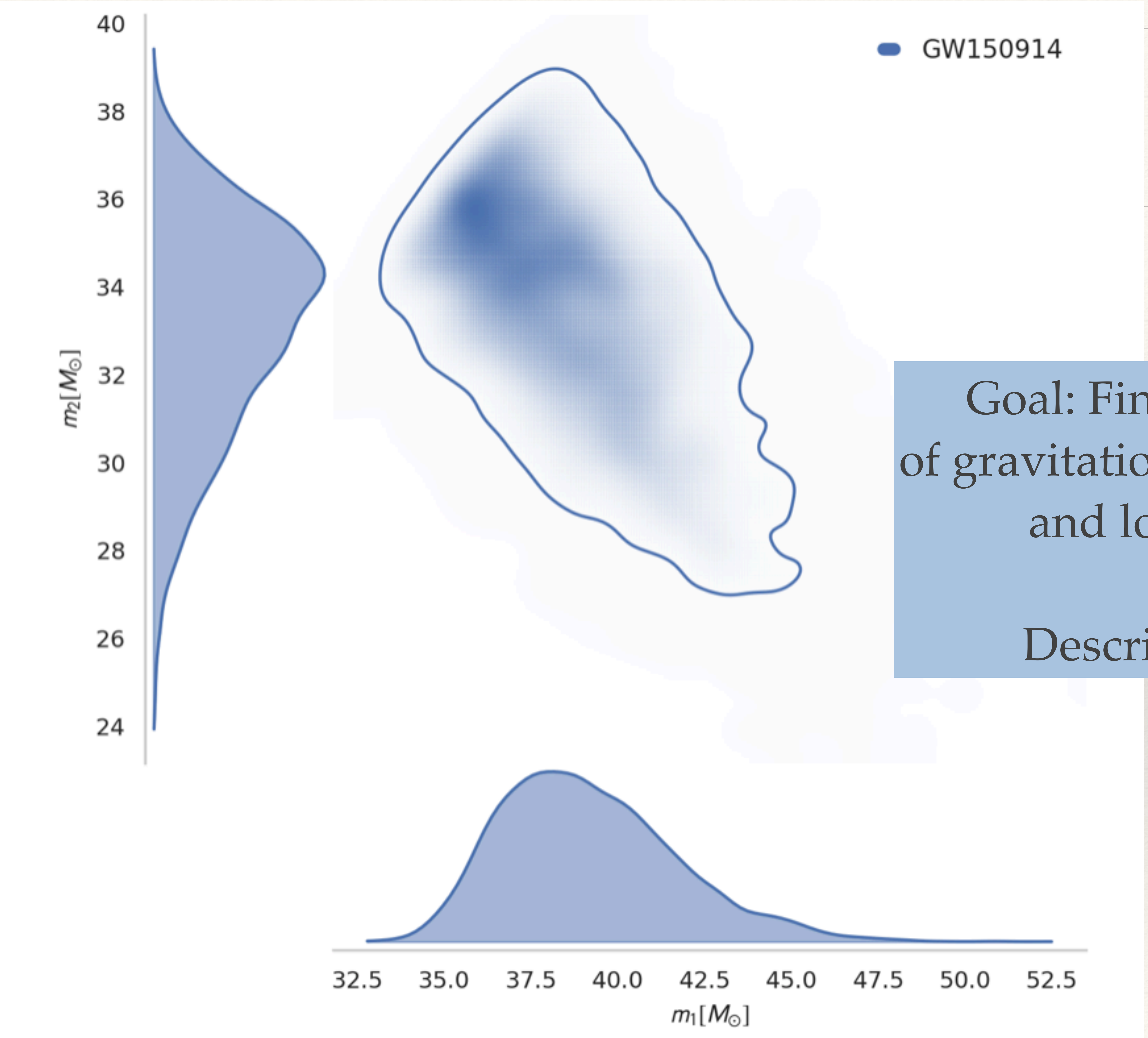
An upside-down guide to LIGO parameter estimation

Jonah Kanner
August, 2021

Some slides adopted from Katerina Chatziioannou, Open Data Workshop #1
DCC: G1800247

The end


mass_1
density
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
32
34
36
38
40
42
44
46
48
50
52
54
Event
GW150914
The mass of the black hole is around 38 +/- 2 $M_\odot$
Mass 1 ($M_\odot$)



The end

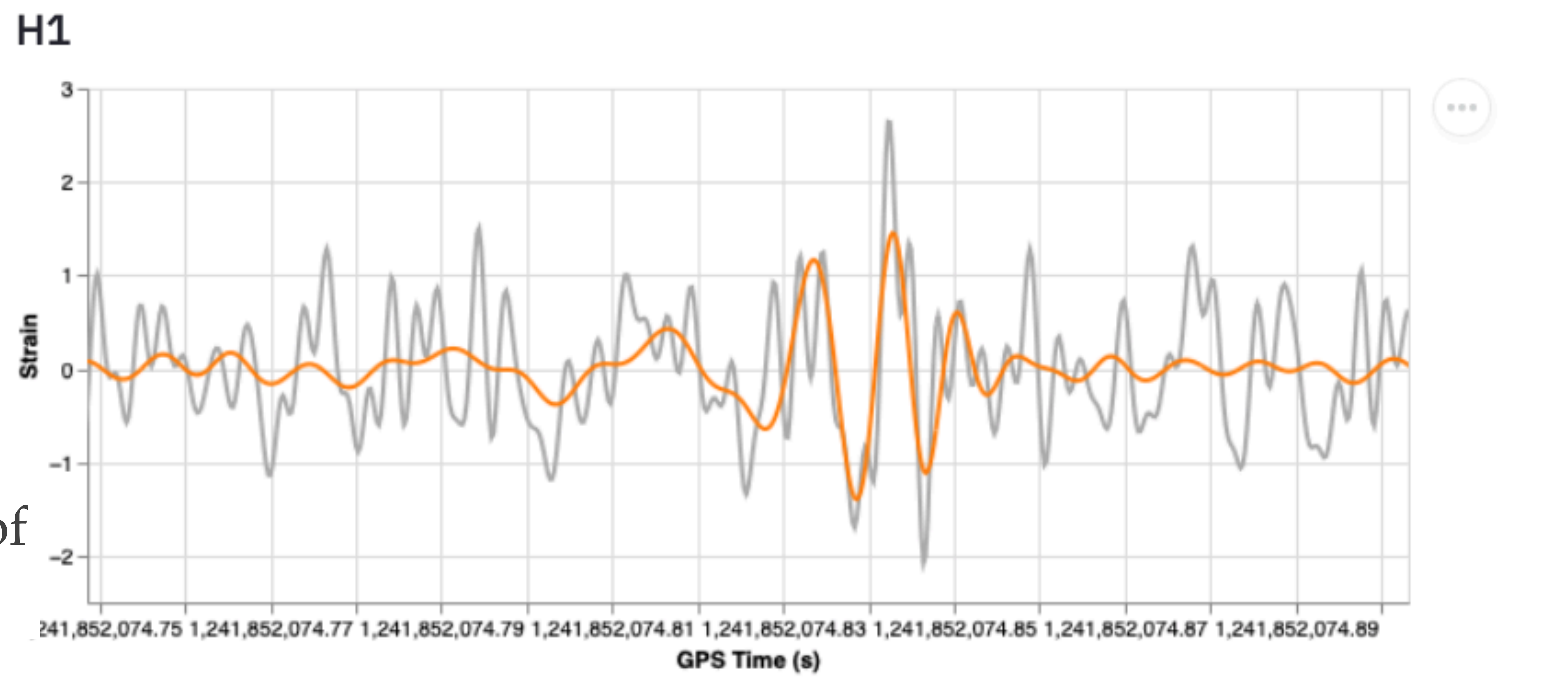
Goal: Find the "parameters" of the sources of gravitational waves, such as the masses, spins, and locations of merging black holes

Describe as a probability distribution

Find parameters by fitting waveforms

m1, m2, s1, s2, RA, DEC, iota, …

Waveform

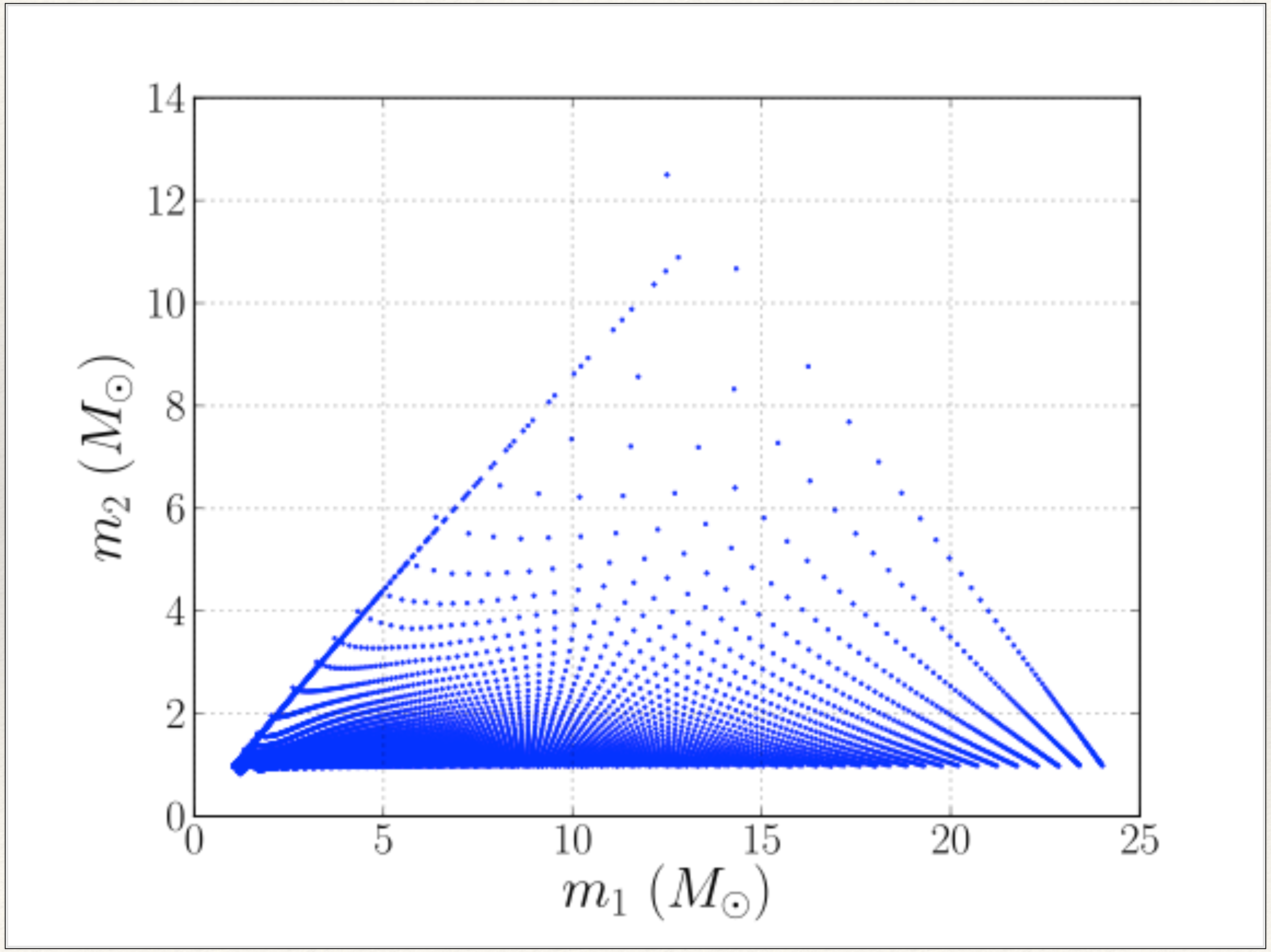


By checking how well a waveform fits the data, we get a measurement of how likely is this set of parameters

How many waveforms?

- For searches, we use 1 detector at a time, so we can ignore lots of parameters
- Searches typically use a 2 or 3 dimensional parameter space, with m1, m2, and maybe a spin parameter.

2-D “Parameter Space”



For a search, we have a 2-D parameters space

If we want to cover the parameter space in 1% steps, maybe we need 100^2 ~ 10,000 waveforms

A modern computer can perform 10^9 calculations per second, so with enough computers, we can compute the SNR for all of these templates

PE uses a 15 dimensional parameter space

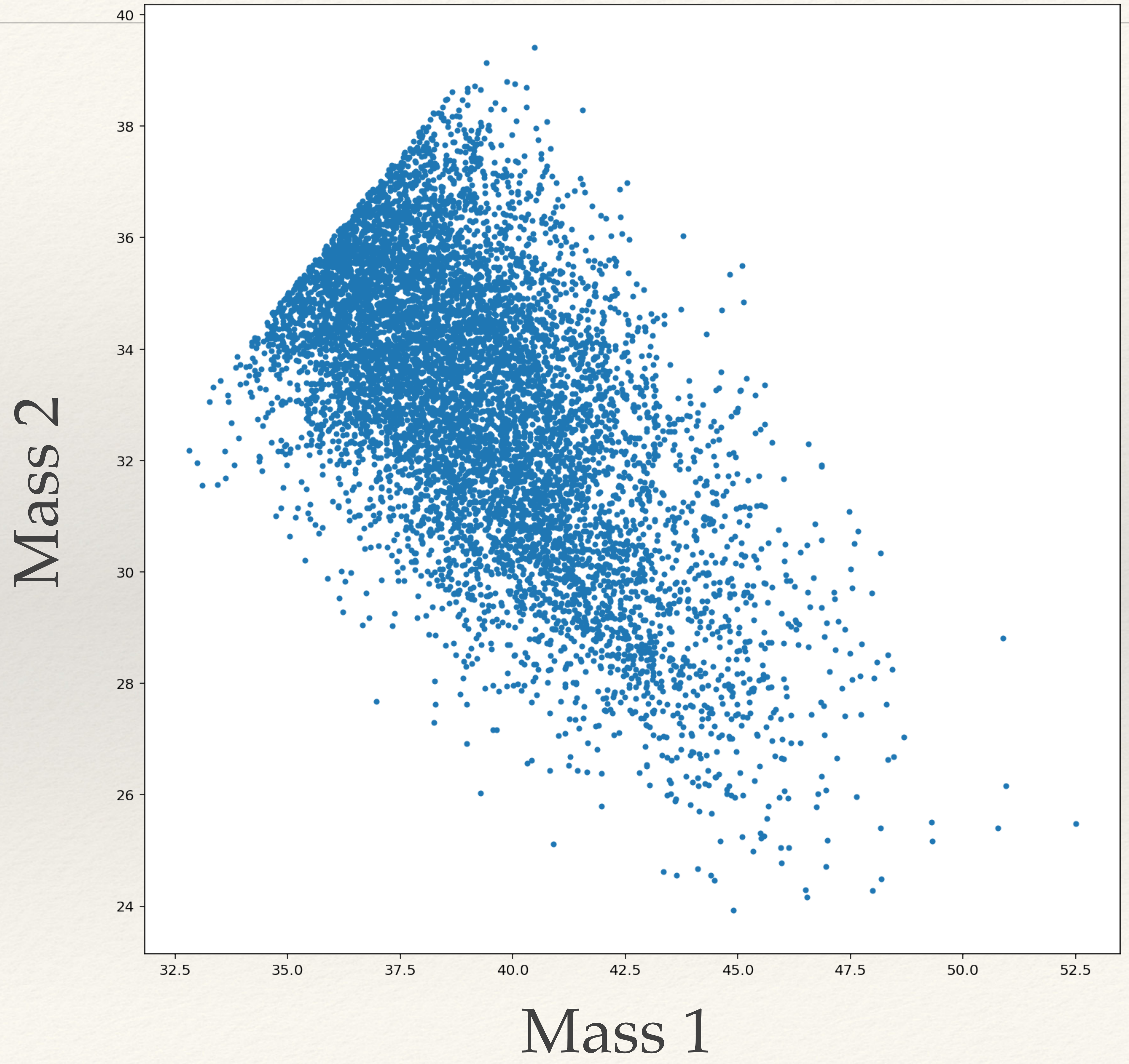
- For parameter estimation, we want to evaluate all the parameters
- This includes lots of position parameters that can be ignored for the search (e.g. the source location in the sky)
- Then, we have a 15 dimensional parameter space. If we want a grid with 1% spacing, that would be 100^15 = 10^30 waveforms (!)
- At modern computing speeds, it would longer than the age of the universe to make all these calculations. So, using a grid is not possible

Bayesian Samplers to the rescue

- "Samplers:"

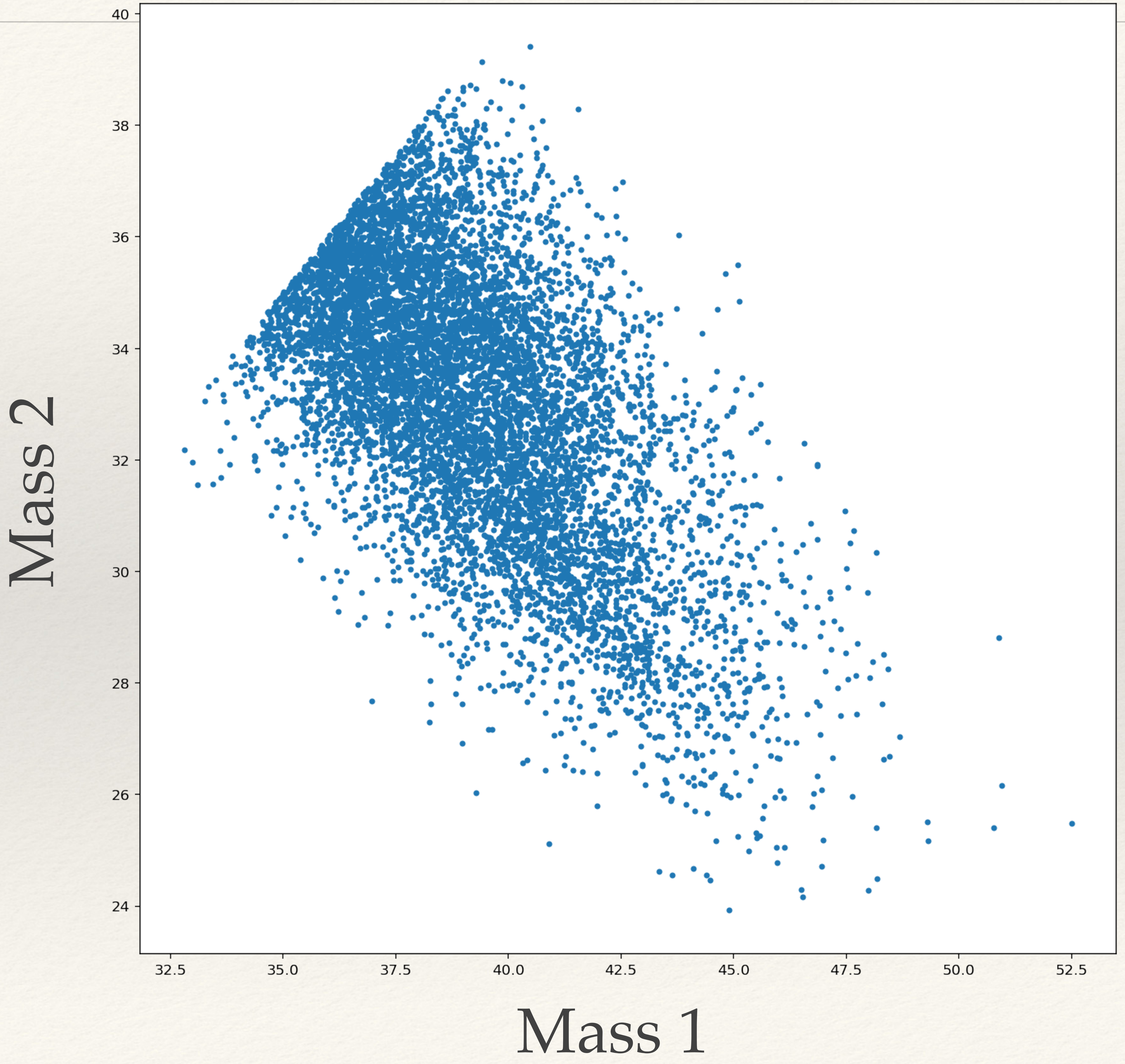
1. Select a point in parameter space
2. Check the waveform fit (likelihood)
3. Record the values
4. Repeat

Samplers are designed to be more likely to select points with a better fit



Samplers to the rescue

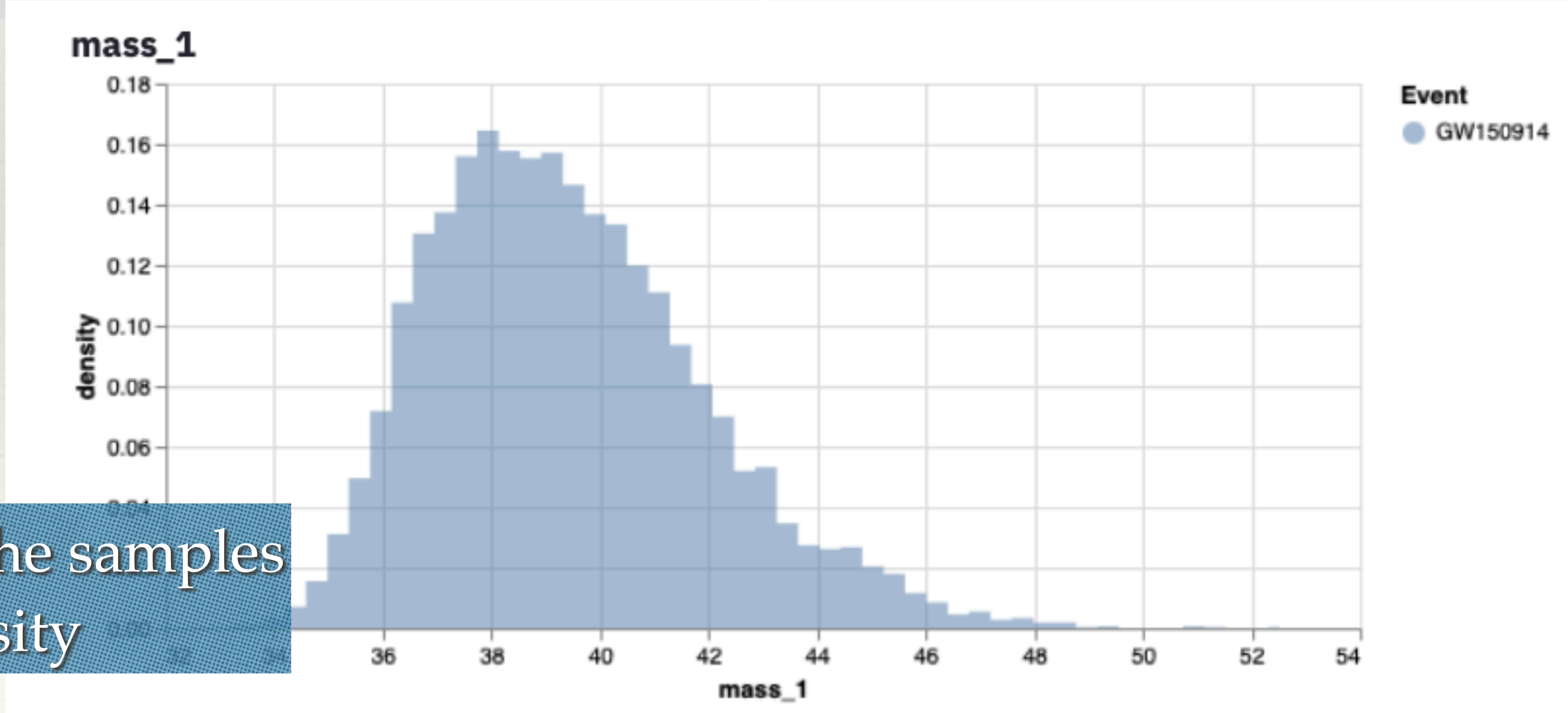
- There are a few different classes of sampling algorithms
- Go by names like “Markov Chain Monte Carlo” and “Nested Sampling”



Samplers are designed to be more likely to select points with a better fit

A bucket of samples

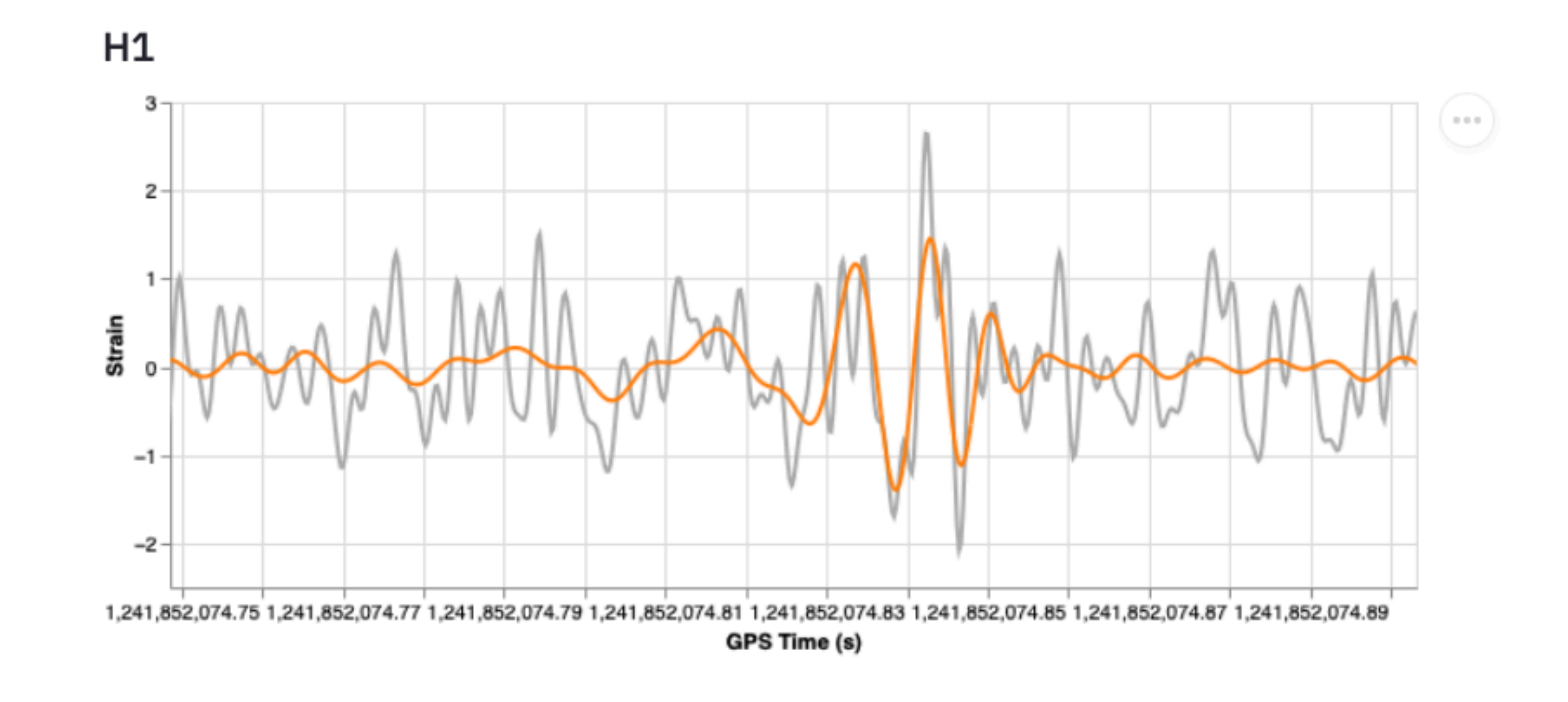
m1_detector_frame_Msun	m2_detector_frame_Msun	spin1	spin2	costilt1	costilt2
39.037380	37.044563	0.417147	0.867740	-0.280624	0.403853
34.620096	34.184416	0.125709	0.260679	-0.757349	-0.312285
37.894343	33.970520	0.581047	0.926893	0.649781	-0.510843
36.412973	35.684463	0.235808	0.094391	0.116578	-0.720505
39.477251	31.645008	0.511521	0.868009	-0.438237	0.269333



After running sampler, histogram the samples to calculate probability density

Likelihoods

- At each step in sampler, need to evaluate how well the waveform fits the data
- Do this with a quantity called the "Likelihood"
- The likelihood asks:
  "Assuming these are the right parameters, what is the probability of getting this particular data"



$$p(d\,|\,h')$$

"The probability of the data, given this waveform"

Calculating the likelihood

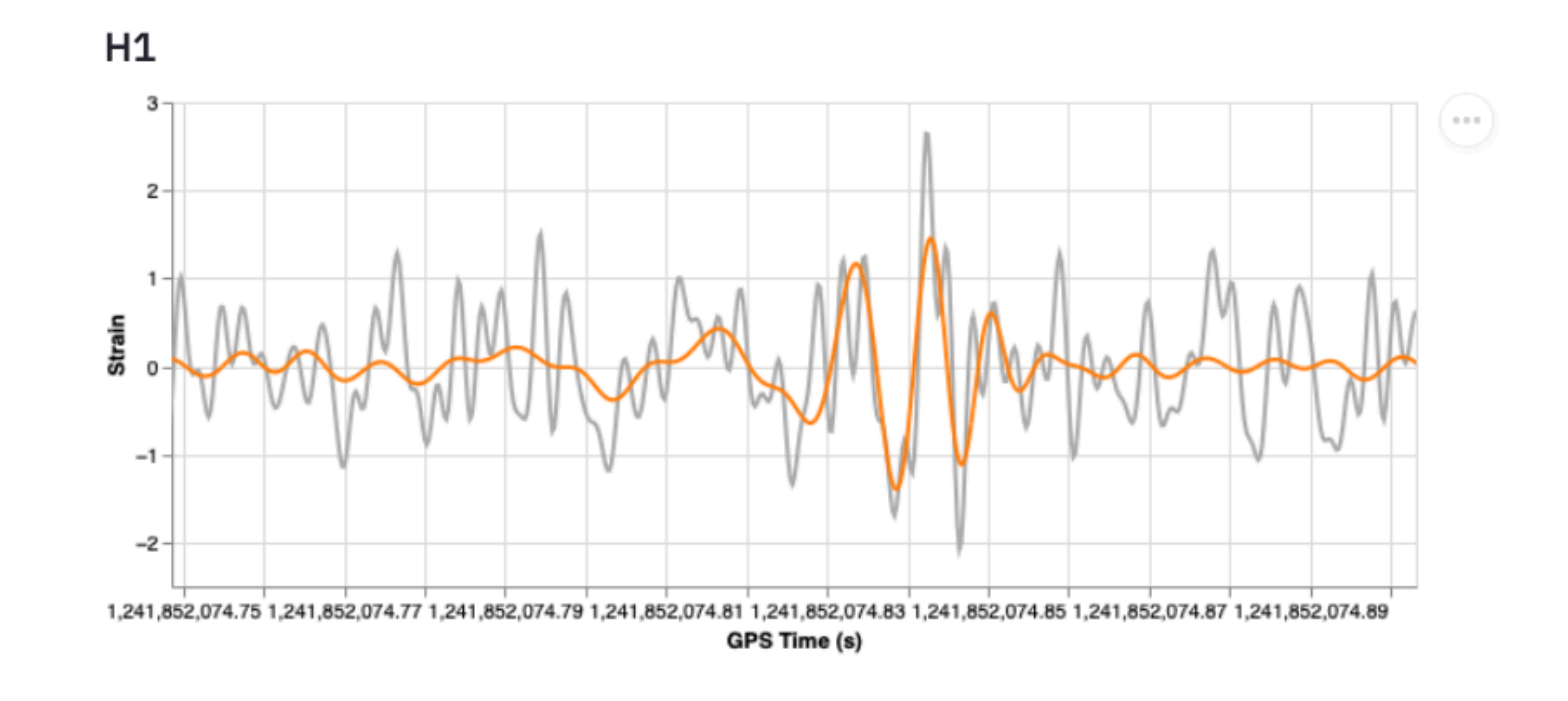
$$d = h' + n$$

data = signal + noise

$$n = d - h'$$

noise = data - signal

Call this the "residual" data, after subtracting the trial wavefom



Calculating the likelihood

By assuming the data are Gaussian, we know how likely we are to get a noise sample.

Noise much louder than the PSD value is very unlikely!

$$p(d\,|\,h') \propto e^{-\frac{1}{2}(n|n)} = e^{-\frac{1}{2}(d-h'|d-h')}$$

The Likelihood is the probability of getting the residual we see.

Waveforms that fit badly will tend to leave large residuals, and so will have a low likelihood

$$(a|b) = 4\int \frac{a(f)b^*(f)}{S_n(f)}df$$

Above, we used this definition of an "inner product" which is just a cross-correlation between the two values

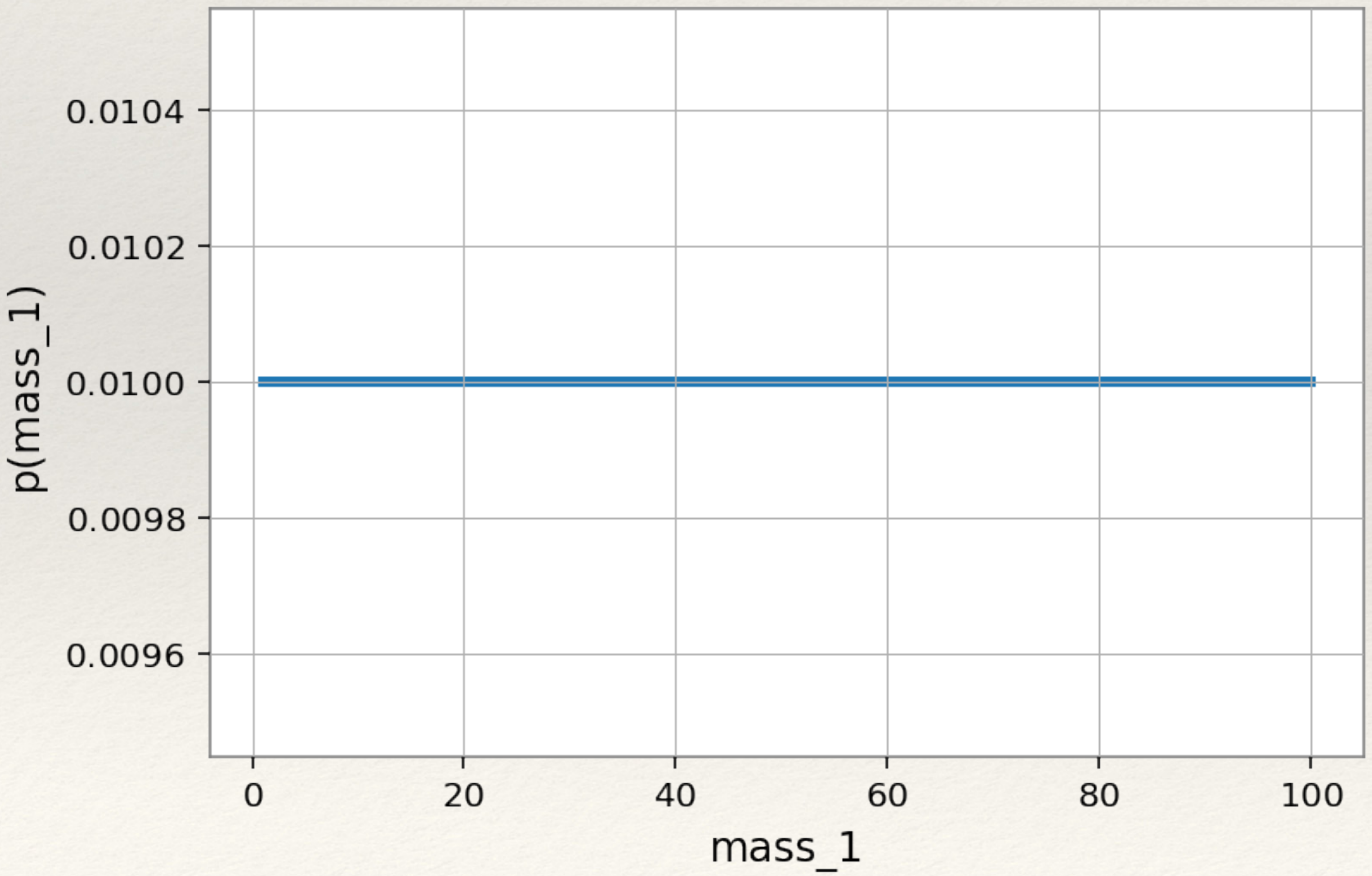
One more element: The prior

- For all of this to work, we need one more thing: “the prior”
- Prior distributions encode how likely we are to find each value of a parameter
- An important step in defining priors is to set the possible range of each parameter
  - For example, we might expect masses from [1, 100] $M_\odot$, and latitudes (DEC) from [-$\pi/2$, $\pi/2$]

One more element: The prior

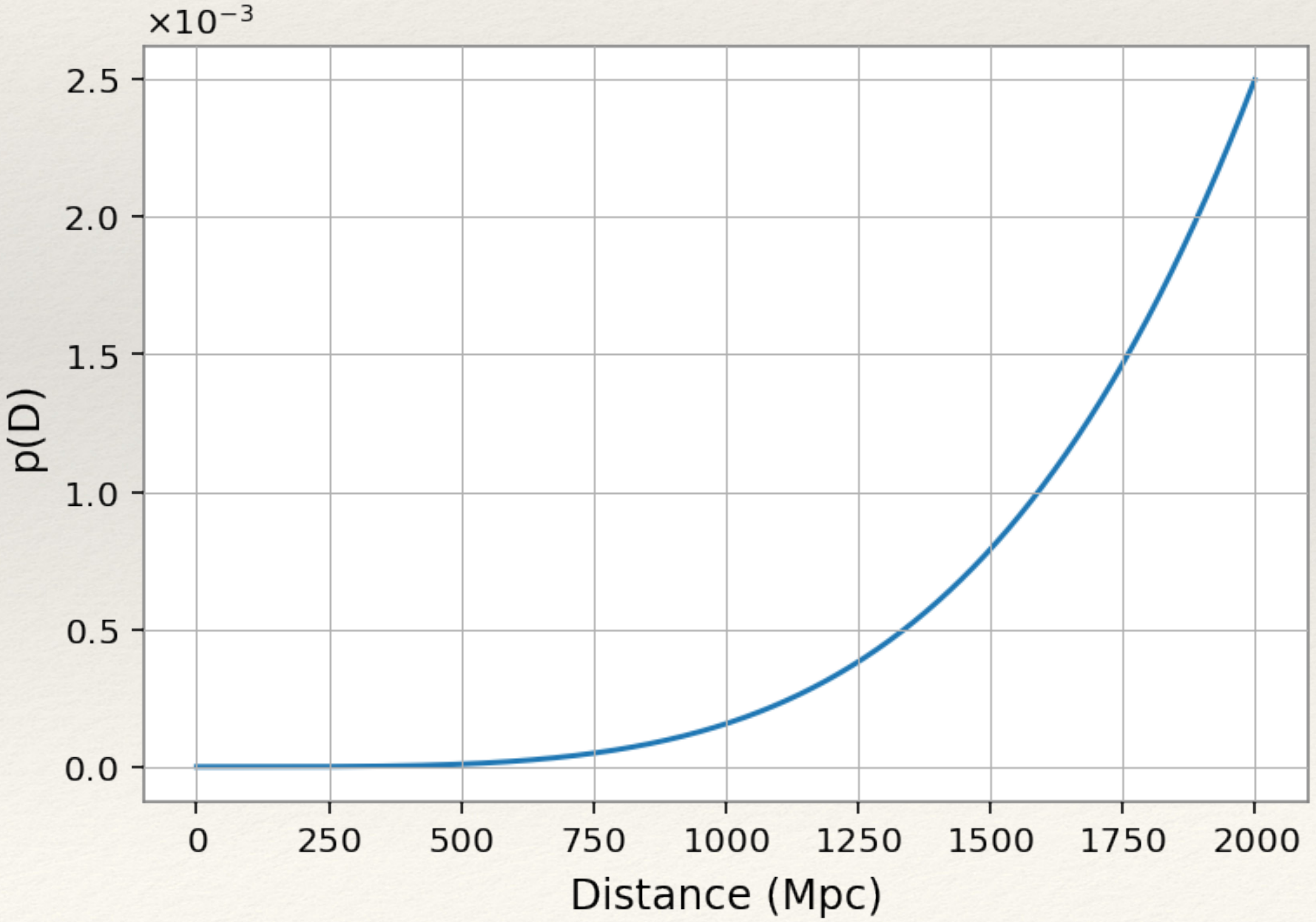
For mass, a simple prior would be to look for any mass from [1,100] $M_\odot$ with a flat distribution

$$p(m_1)$$



Expect BBH to be distributed uniformly in volume
So, a prior on distance would have more black holes further away

$$p(D)$$



Putting it together: Bayes' Theorem

$$p(h'|d) = \frac{p(d|h')p(h')}{p(d)}$$

Putting it together: Bayes' Theorem

$$p(h'|d) = \frac{p(d|h')p(h')}{p(d)}$$

Putting it together: Bayes' Theorem

$$p(h'|d) = \frac{p(d|h')p(h')}{p(d)}$$

Putting it together: Bayes' Theorem

"The Posterior"
Probability the source has these parameters, given the data

This is what we want to know!

"The Likelihood"
Probability of getting these data if this is the right parameter

"The Prior"
Guess about how likely is this parameter

$$p(h'|d) = \frac{p(d|h')p(h')}{p(d)}$$

"The Evidence"
With one model, just a normalization factor

PE: Right-side up

- Start with Bayes Theorem
    - Build a model
        - Set priors on all parameters
            - Run a sampler to explore parameter space
                - Collect a bucket of posterior samples
                    - Histogram the samples to measure Posterior distributions

PE: Right-side up

- Start with Bayes Theorem
- Build a model
- Set priors on all parameters
- Run a sampler to explore parameter space
- Collect a bucket of posterior samples
- Histogram the samples to measure Posterior distributions

$$p(h'|d) = \frac{p(d|h')p(h')}{p(d)}$$

PE: Right-side up

Start with Bayes Theorem

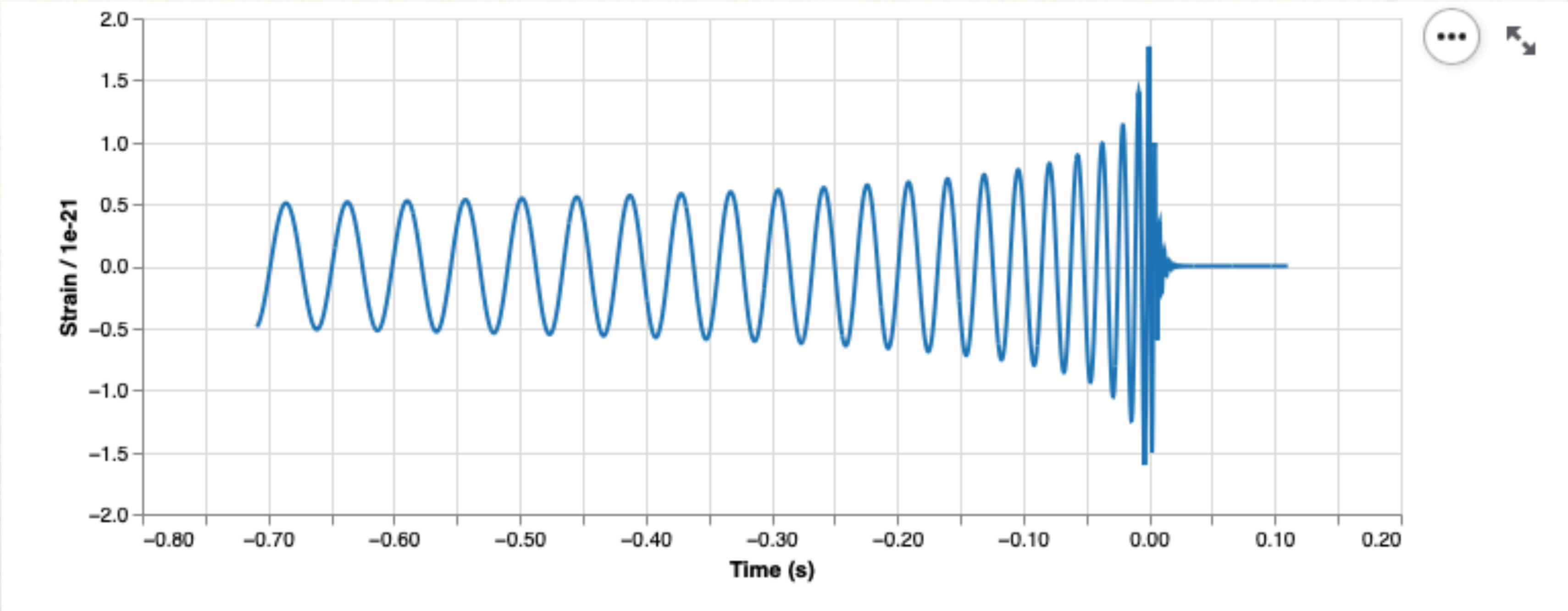
Build a model

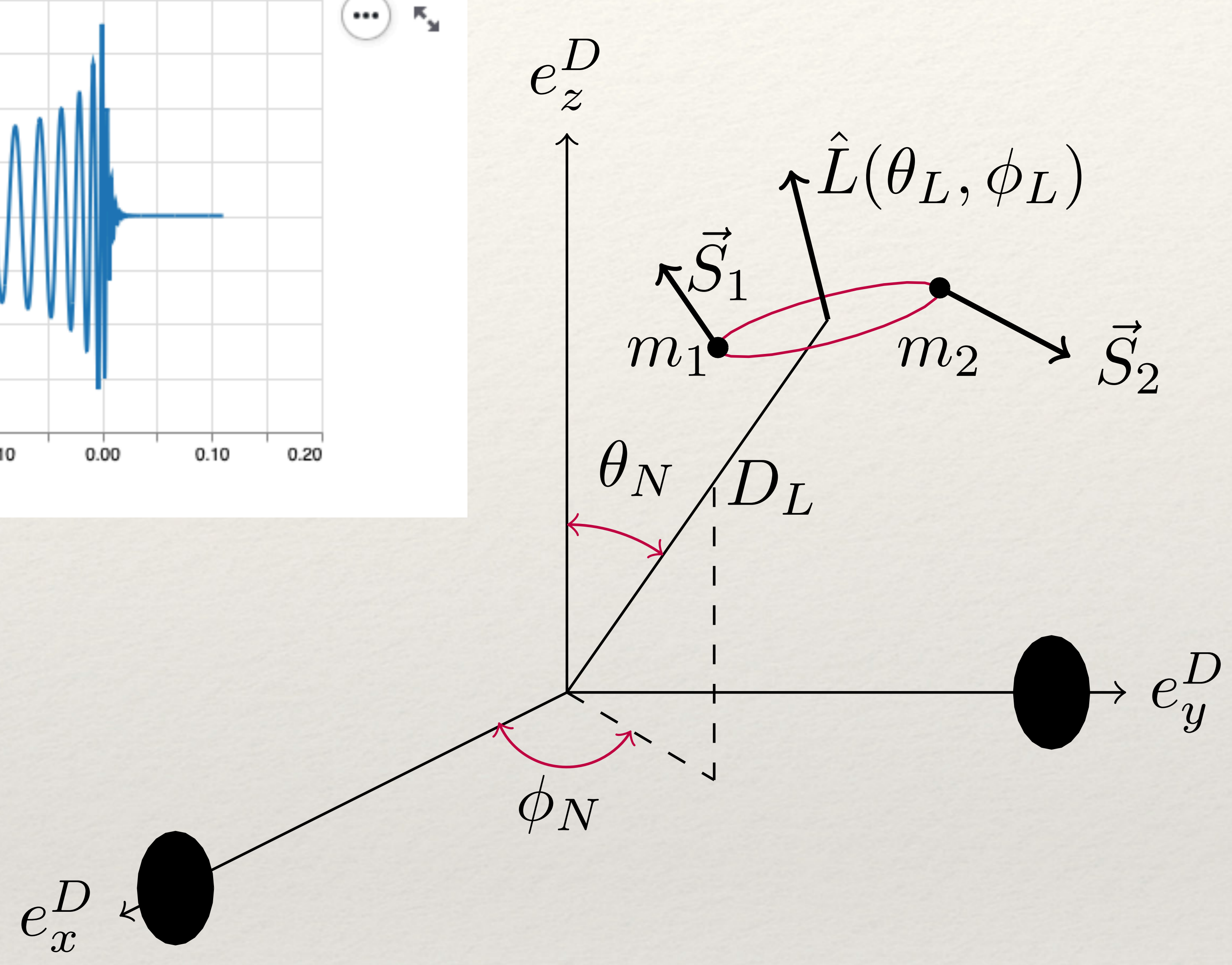
Set priors on all parameters

Run a sampler to explore parameter space

Collect a bucket of posterior samples

Histogram the samples to measure Posterior distributions





PE: Right-side up

Start with Bayes Theorem

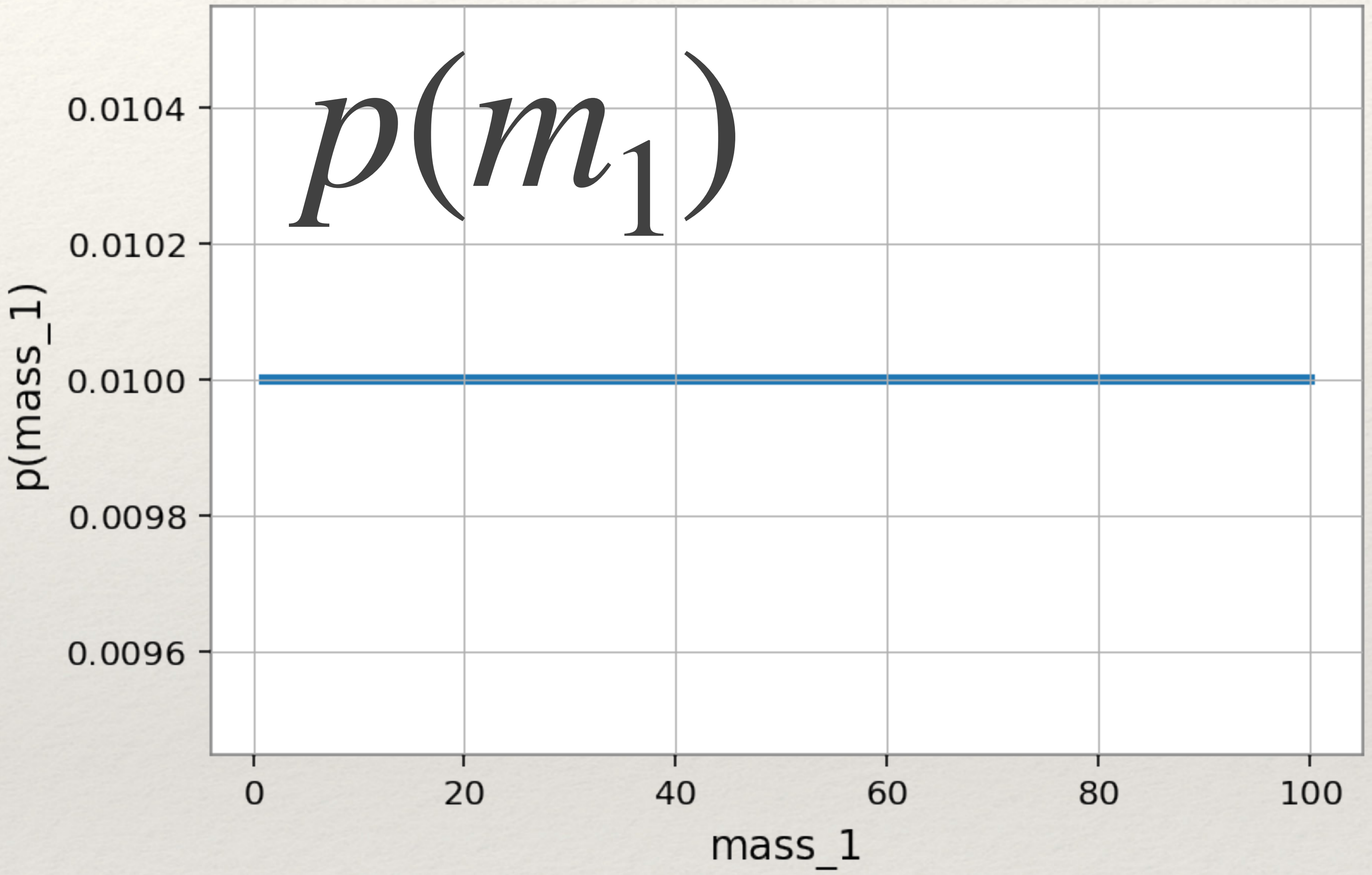
Build a model

Set priors on all parameters

Run a sampler to explore parameter space

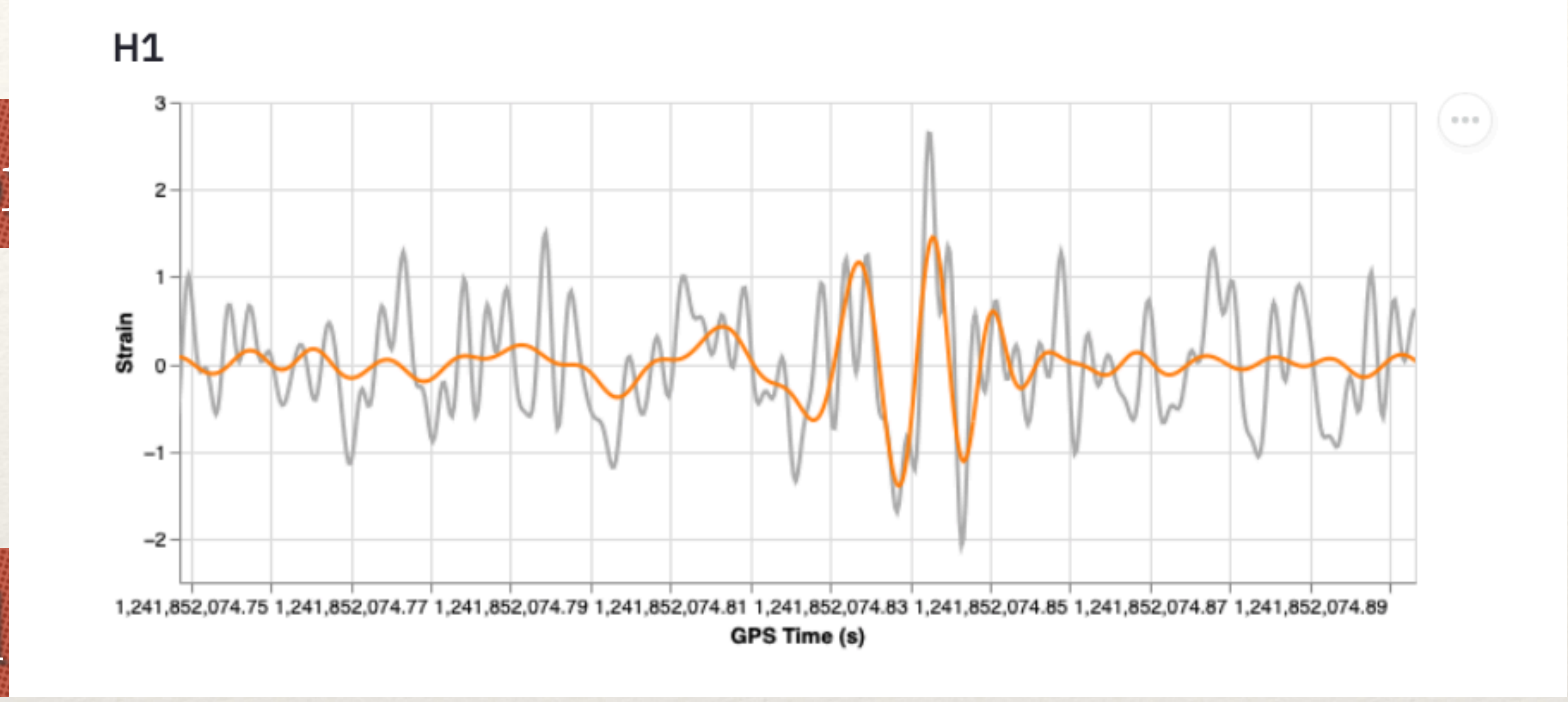
Collect a bucket of posterior samples

Histogram the samples to measure Posterior distributions



PE: Right-side up

Start with Bayes Theo

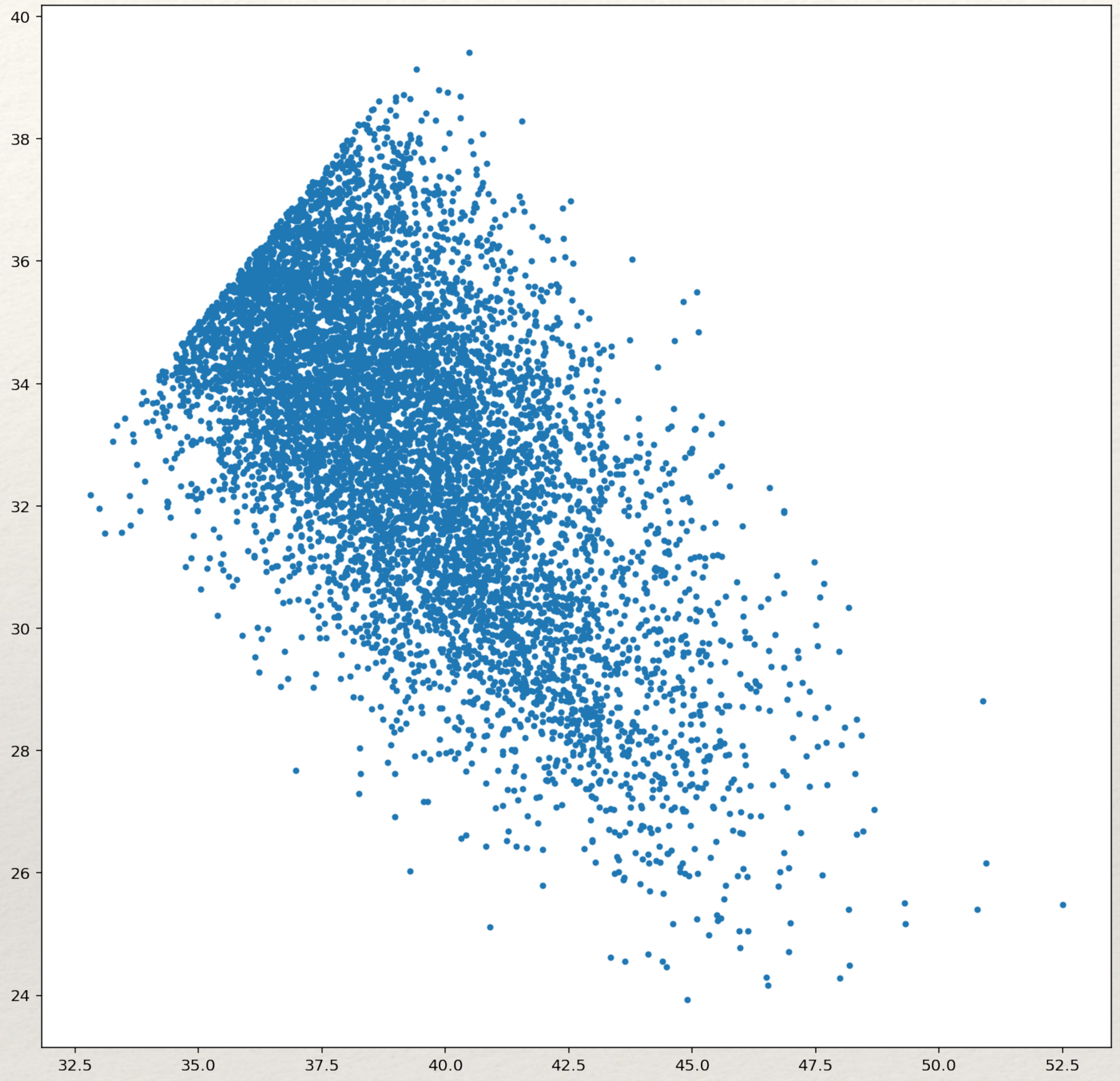


Build a model

Set priors on all

Run a sampler to explore parameter space

Collect a bucket of posterior samples



Histogram the samples to measure Posterior distributions

PE: Right-side up

Start with Bayes Theorem

Build a model

Set priors on all parameters

m1_detector_frame_Msun	m2_detector_frame_Msun	spin1	spin2	costilt1	costilt2
39.037380	37.044563	0.417147	0.867740	-0.280624	0.403853
34.620096	34.184416	0.125709	0.260679	-0.757349	-0.312285
37.894343	33.970520	0.581047	0.926893	0.649781	-0.510843
36.412973	35.684463	0.235808	0.094391	0.116578	-0.720505
39.477251	31.645008	0.511521	0.868009	-0.438237	0.269333

Run a sampler to explore parameter space

Collect a bucket of posterior samples

Histogram the samples to measure Posterior distributions

PE: Right-side

Start with Bayes Theorem

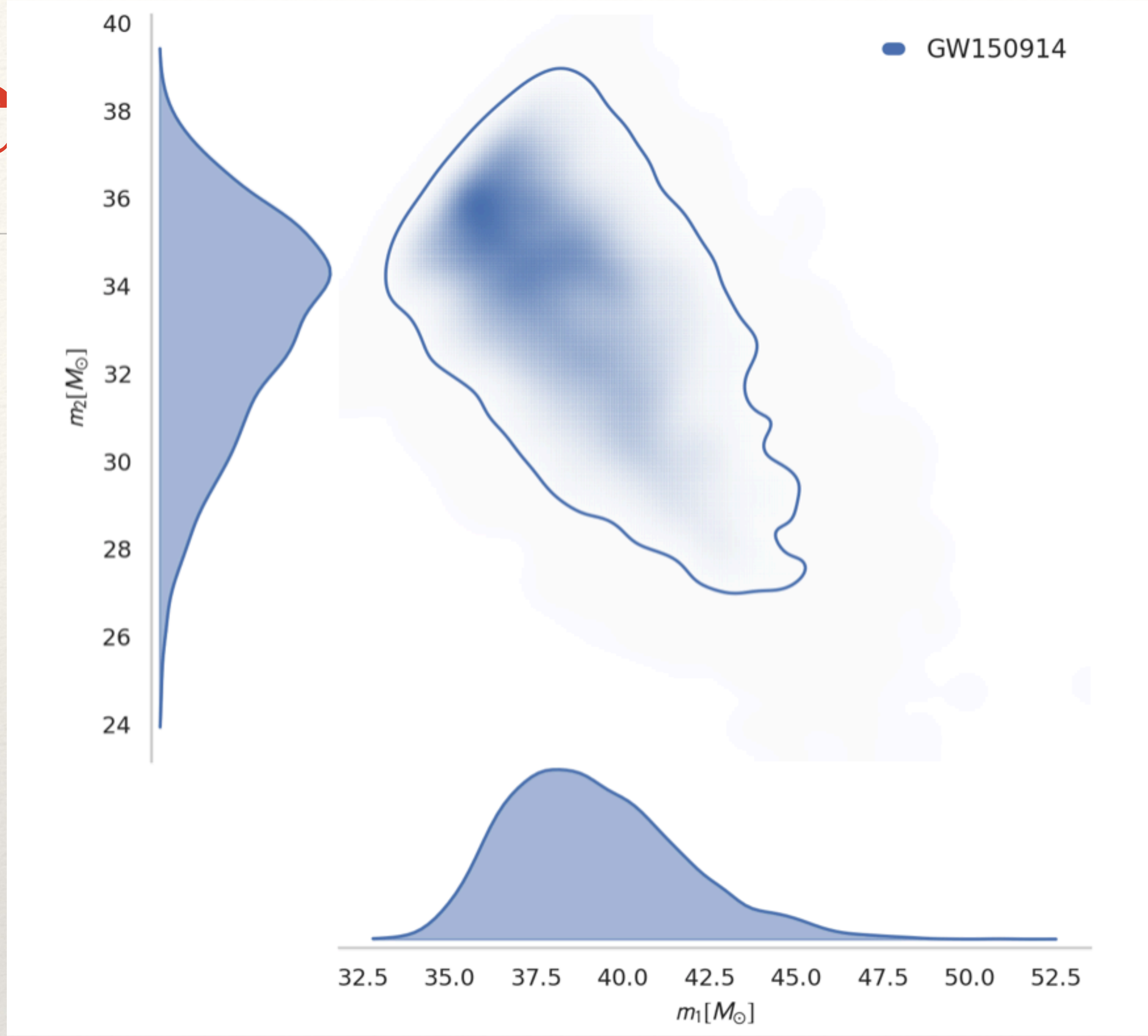
Build a model

Set priors on all parameters

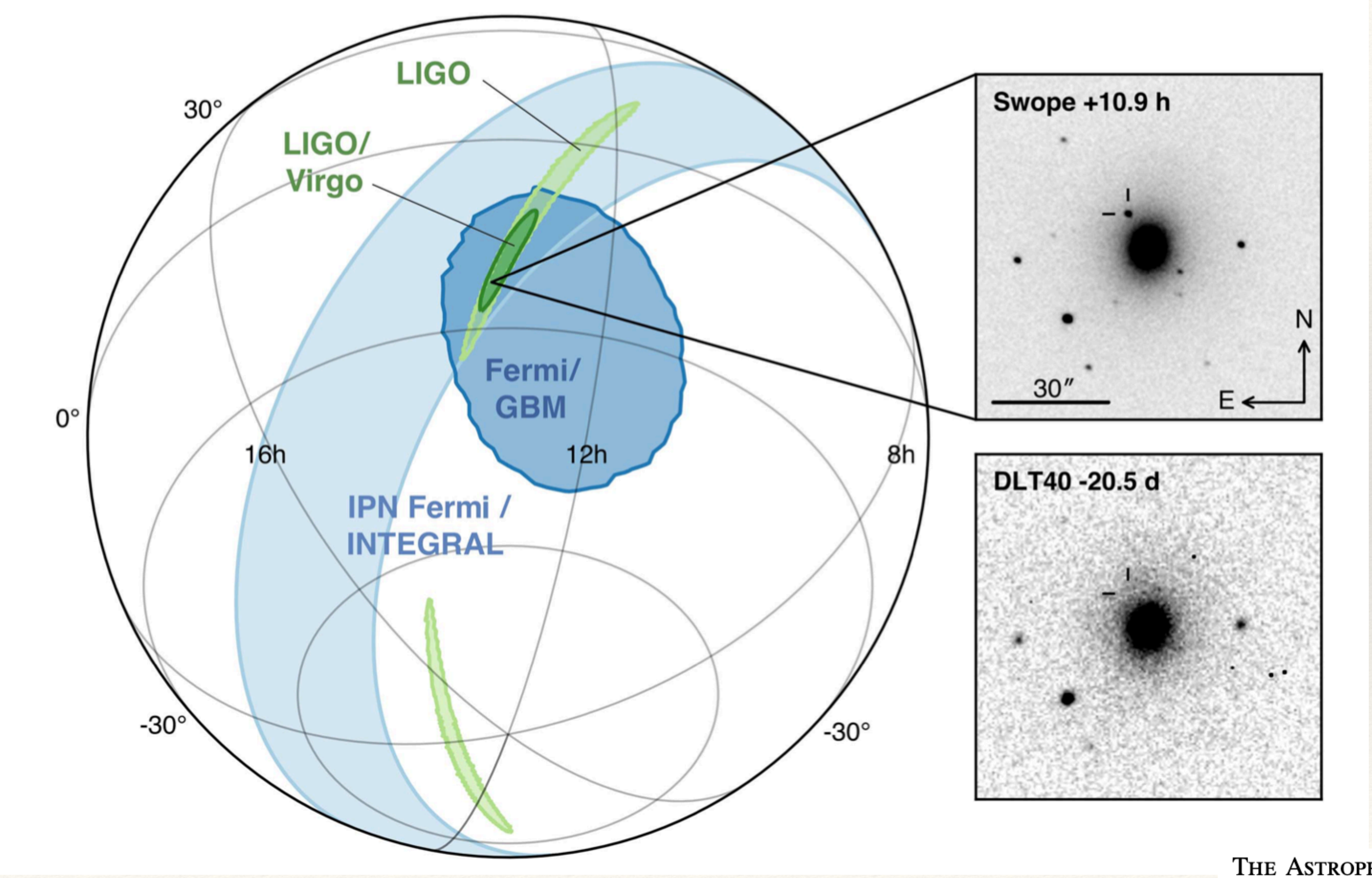
Run a sampler to explore parameter space

Collect a bucket of posterior samples

Histogram the samples to measure Posterior distributions



Key parameters: Skymaps

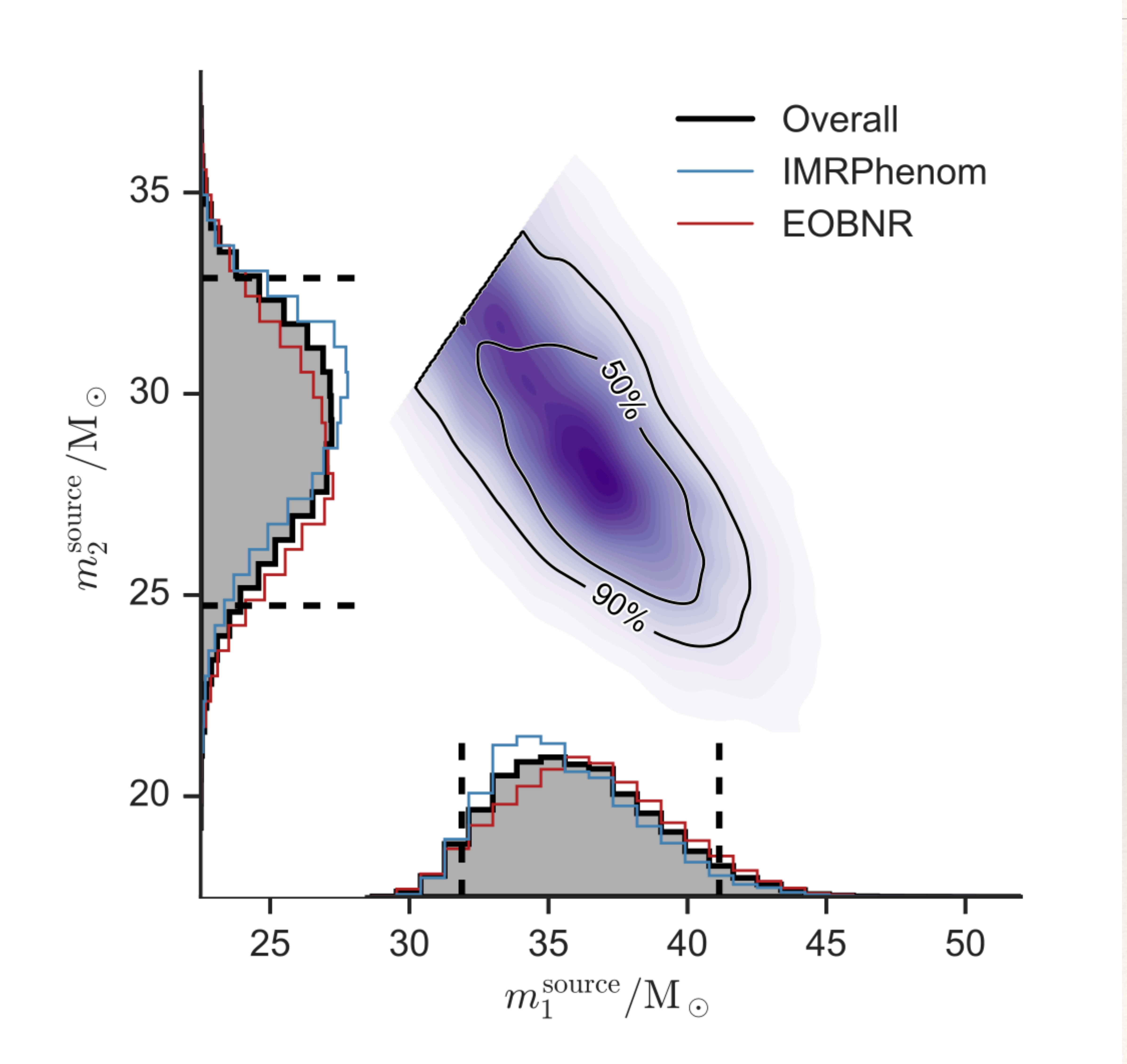


The Declination (latitude)
and
Right Ascension
(longitude)
of a source tell us where to find it

Measured mainly by time-of-flight between detectors

The Astrophysical Journal Letters, 848:L12 (59pp), 2017 October 20

Mass

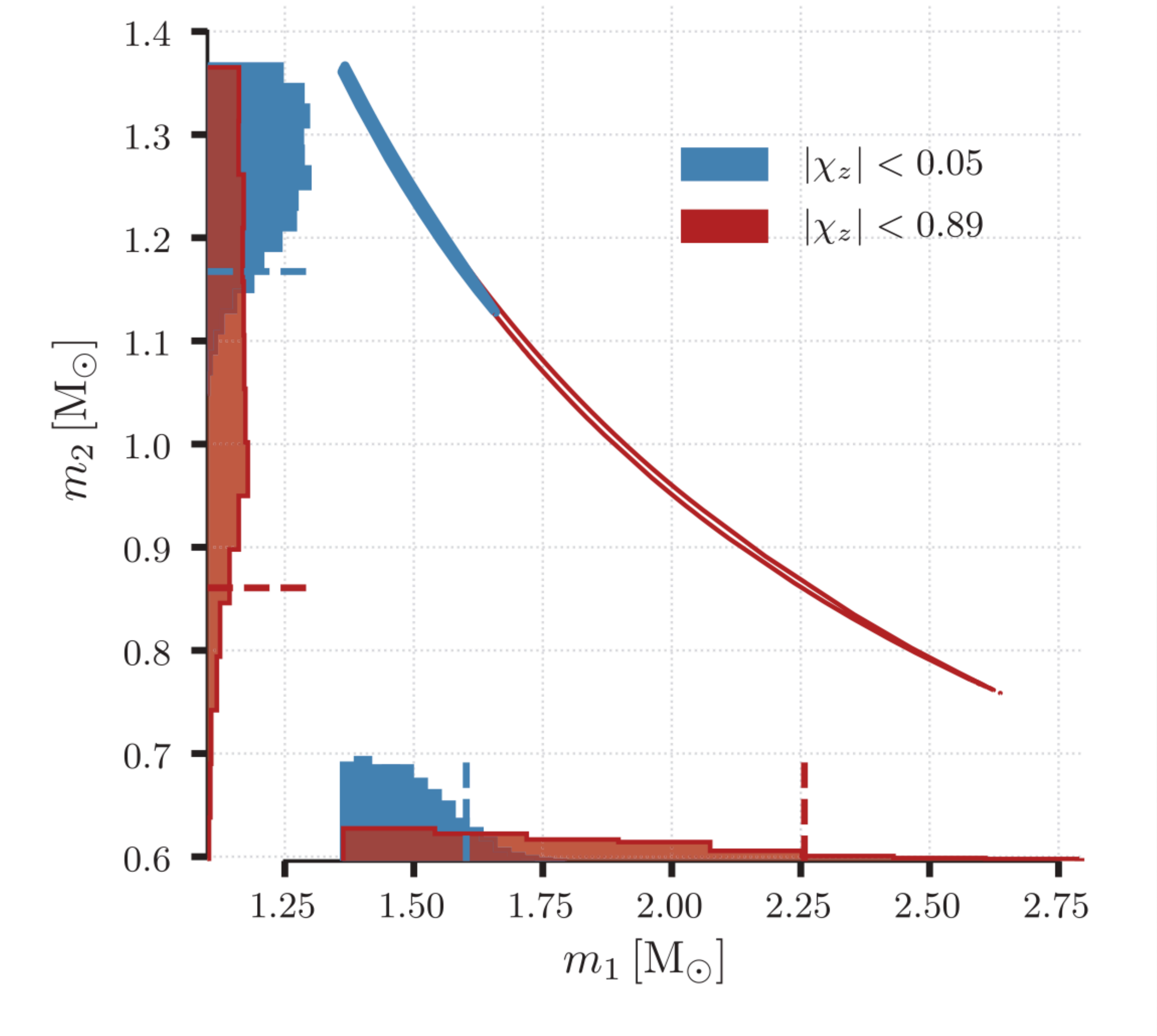


Mass posteriors for GW150914

The size of an object gives clues to its origins and history

Measured mainly by frequency evolution

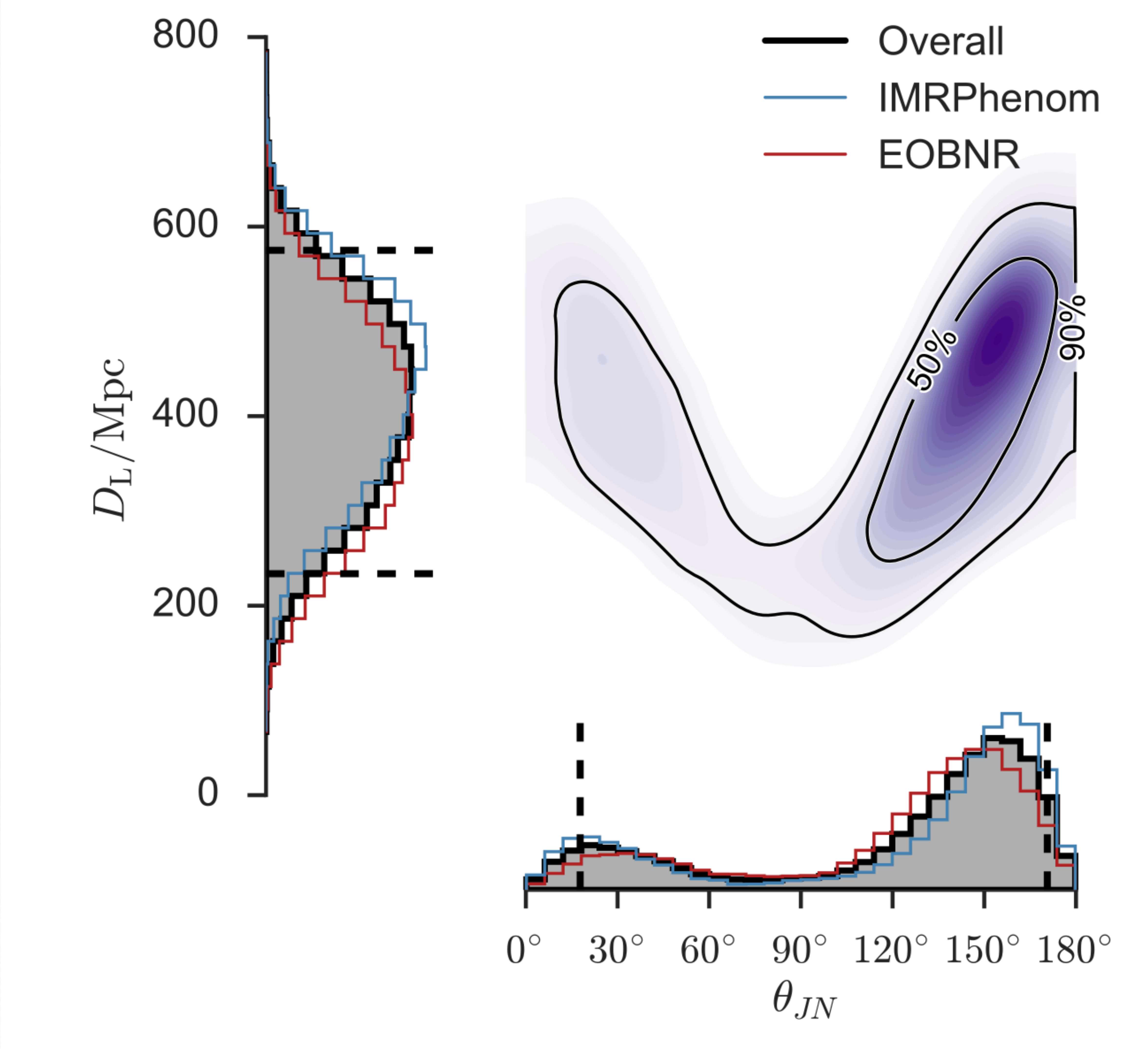
Degeneracies



Sometimes, the measurement of one parameter depends on the measurement of another parameter.

Here, the chirp mass is measured very well, but it can be difficult to measure individual masses

Degeneracy: Distance and inclination angle



The distance and inclination angle of a source both impact the amplitude of the signal

Summary

- Parameter Estimation is used to learn the masses, spins, and position of sources
- Uses a Bayesian statistical framework and a sampler (e.g. MCMC) to explore a large parameter space
- Posterior samples represent the probability of different parameter values