Summer Research Project

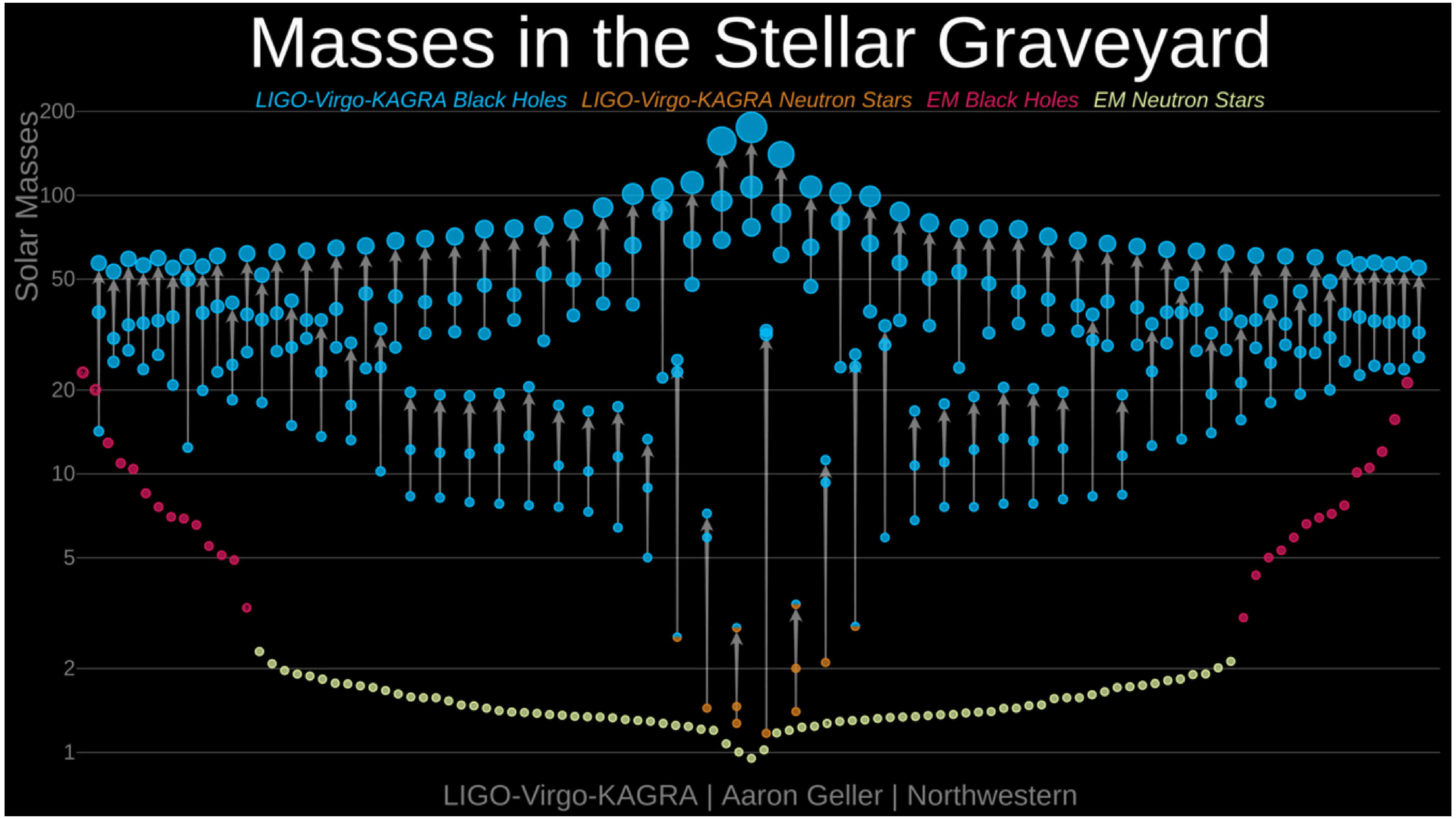
Black Hole Population Modelling from Gravitational Wave Data

Jupiter Stevenson

01/07 - 26/09

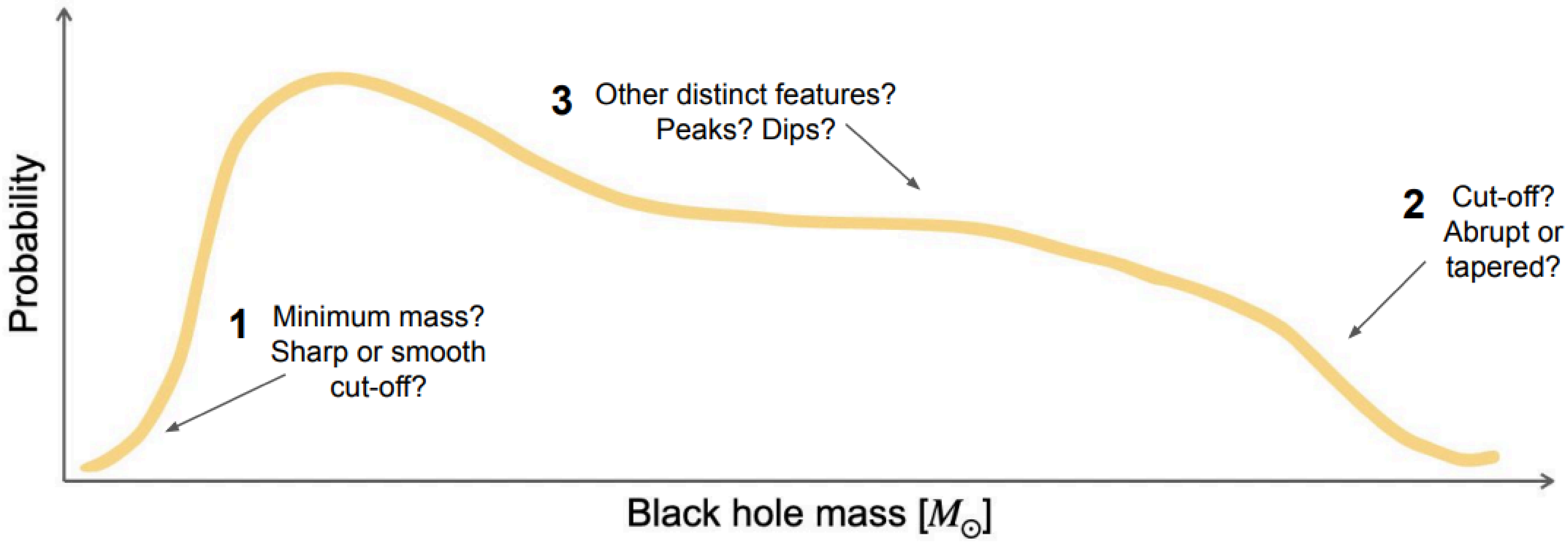
The Gravitational Wave Transient Catalogue

- Consists of publicly released data from LVK collaboration
- There are currently 182 events on the GWOSC events list, of which 93 are confident



GWTC (gwosc.org)
LIGO/Virgo Masses in the Stellar Graveyard

Mass Distribution Function



Analysis of the MDF furthers understanding of the evolution of massive stars

- Is there a mass gap between the highest mass NSs and the lowest mass BHs, and why?
- Where is the upper limit for BH formation?
- What features of stellar evolution lead to peaks and dips?
- How do second or third generation black holes factor into this distribution, and how can we identify them?

LIGO Webinar: GWTC-2 Populations

Bayesian Theory

$$P(M|D) = \frac{P(D|M)P(M)}{P(D)}$$

$$\text{Posterior} = \frac{\text{Likelihood x Prior}}{\text{Evidence}}$$

Posterior- P(M|D)
Prior- P(M)
Likelihood- P(D|M)
Evidence- P(D)

$$BF_{AB} = \frac{P(D|M_A)}{P(D|M_B)}$$

$$\text{Bayes Factor} = \frac{\text{Evidence of Model A}}{\text{Evidence of Model B}}$$

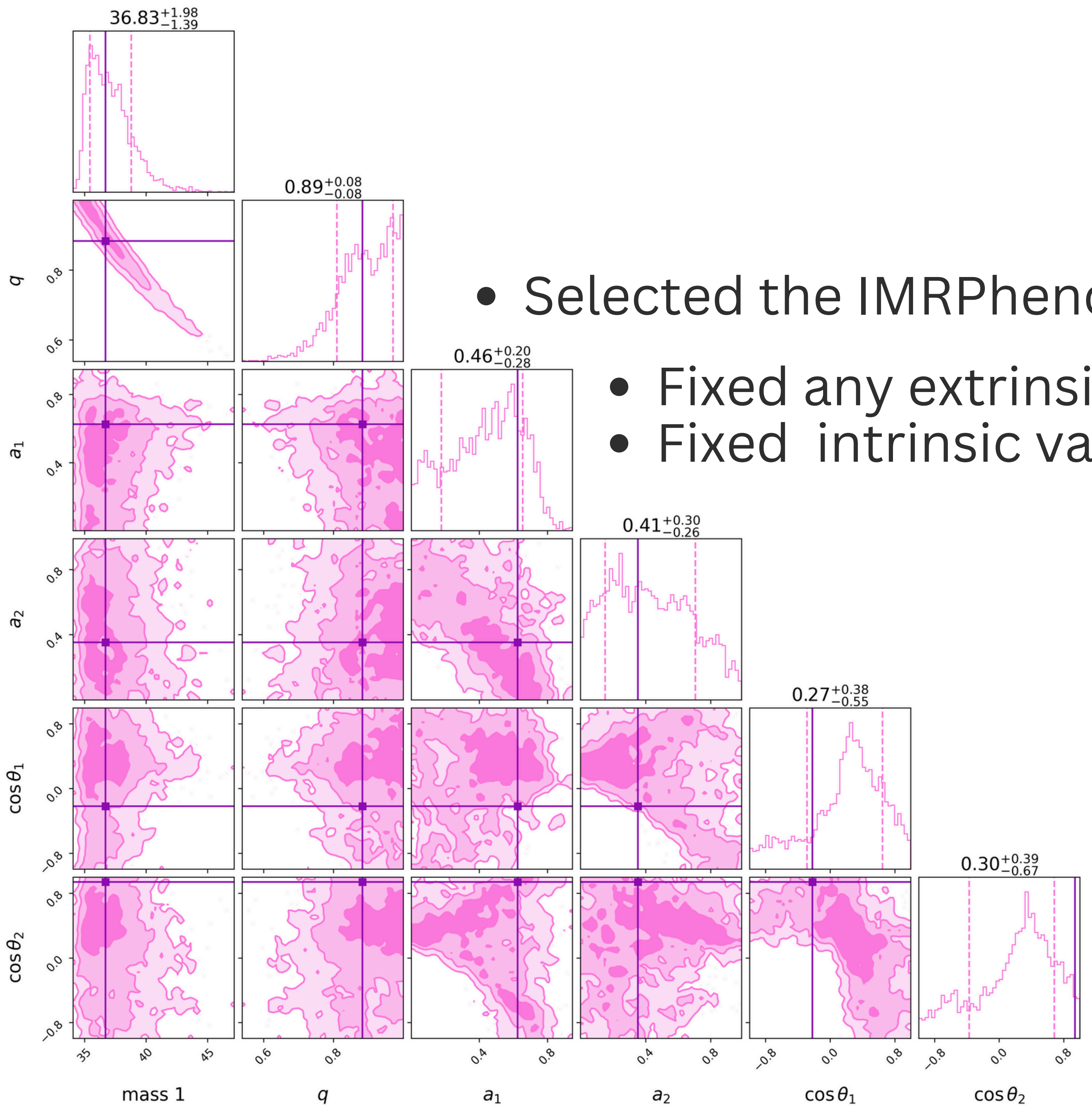
Why bother with Bayensian Theory?

The Bayes factor allows for easy model comparison, however it does not indicate if any given model is the best possible model

Parameter Estimation for a Single Event

Parameters	Desc	Intrinsic or Extrinsic?	Priors
M (solar mass)	Chirp Mass	Intrinsic	U(5,100)
q	Mass Ratio	Intrinsic	U(0.125, 0.99)
Xi	Spin	Intrinsic	U(0, 0.99)
ai	Spin Precession	Intrinsic	U(0, 0.99)
cos(θi)	cosine of tilt	Intrinsic	U(-1, 1)
t (seconds)	geocentric time	Extrinsic	U(t-0.1, t+0.1)
φ12	Spin Orientation	Extrinsic	U(0, 2π)
φjl	Momentum Alignment	Extrinsic	U(0, 2π)
ψ	Polarisation Factor	Extrinsic	U(0, π)
ra (radians)	right ascension	Extrinsic	U(0, 2π)
dec (radians)	declination	Extrinsic	Cosine
d (Mpc)	Luminosity Distance	Extrinsic	U(1e2, 5e3)
θjn	Inclination	Extrinsic	Sine
phase	phase	Extrinsic	U(0, 2π)

Parameter Estimation for a Single Event

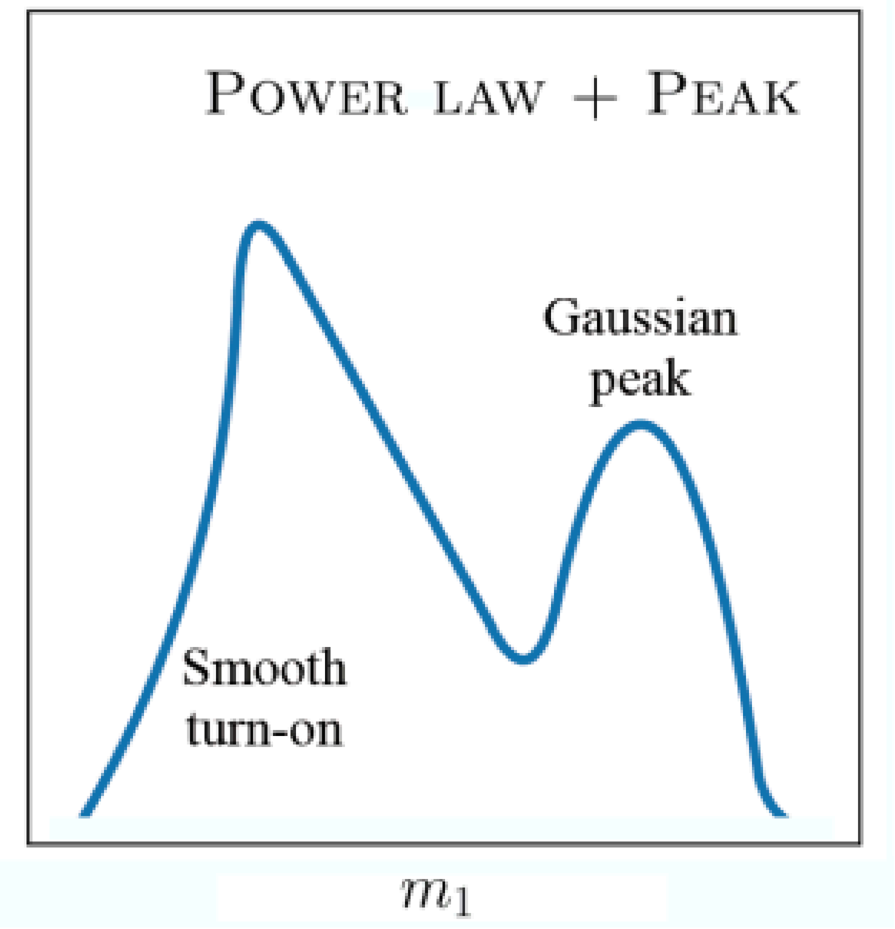


- Selected the IMRPhenomXPHM waveform approximant
  - Fixed any extrinsic values for individual events
  - Fixed intrinsic values relating to spin

Parameter Estimation for a Population of Events

Different mass functions are definied by different parameters- I selected a Power Law + Peak model which has 7 parameters to estimate for the primary mass distribution

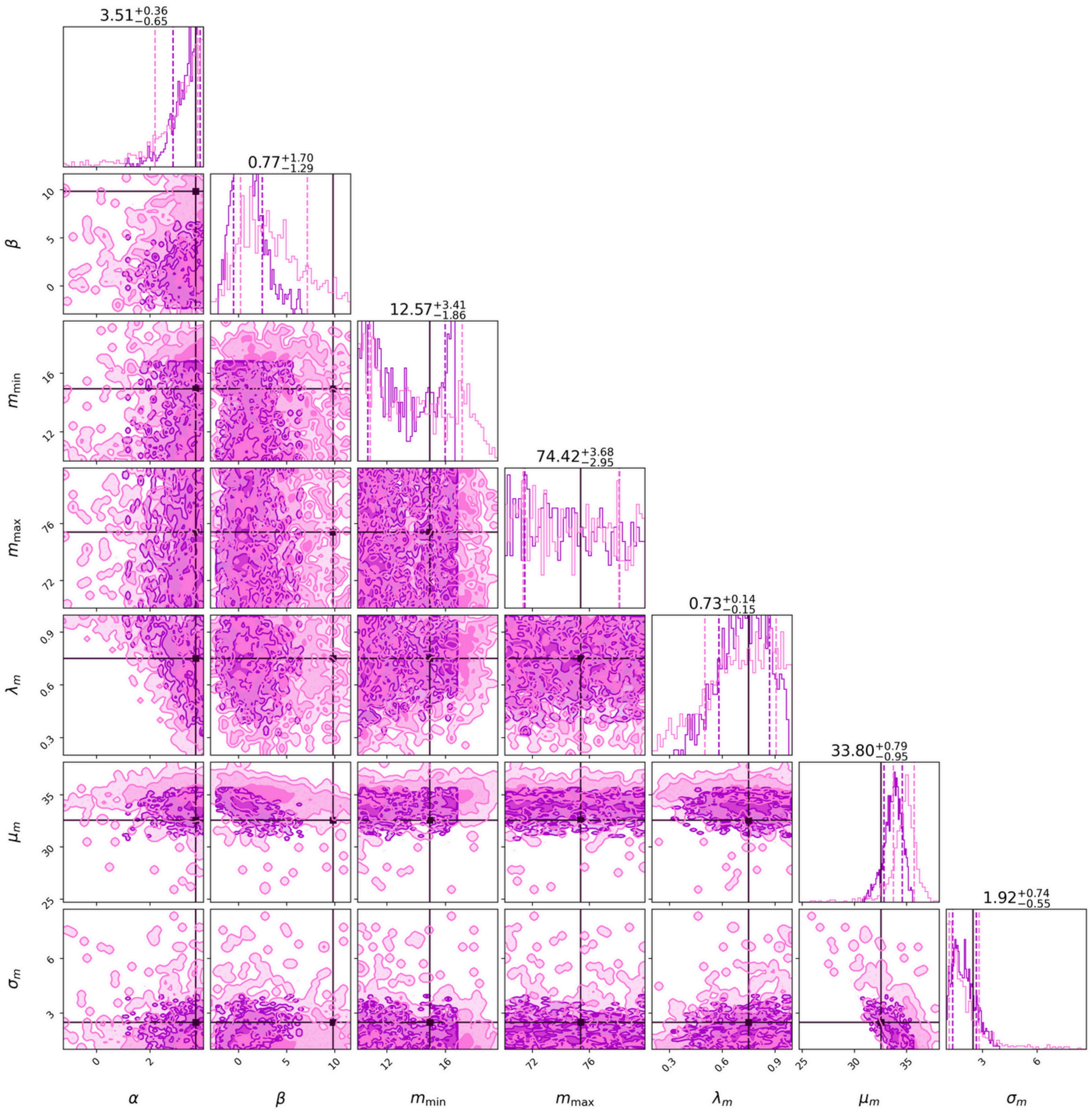
Parameters	Desc	Priors
α	spectral index for the power law component	U(-2, 4)
mmin	minimum mass for the power law component	U(10, 20)
mmax	maximum mass for the power law component	U(70, 80)
λ	fraction of systems in the gaussian component	U(0, 1)
μm	mean of the gaussian component	U(10, 50)
σm	width of the gaussian component	U(1, 5)
δm	range of mass tapering at the minimum	U(0, 10)



This model is justified as the peak arises as large mass (100-130 Solar masses) stars undergo mass loss as they go supernova (pulsational pair instability supernovae), causing a build-up of similar, lower mass black holes than what would be expected if the stars didn't undergo mass loss

Abbott et al 2021 ApJL 913 L7
Abbott et al 2019 ApJL 882 L24

Parameter Estimation for a Population of Events



My Simulations

Set a MDF, thereby defining a “Universe” for the run with fixed parameters

Draw mass 1, q, a1, a2, cos(θ1), cos(θ2) and redshift for an event from the prior distribution

Inject the waveform signal for the event into the noisy interferometers

Check if detector SNR > 12 and network SNR > 8 (i.e. if the event can be resolved from noise)

If so, continue to through to sampling

If not, reject the event and return to drawing events

Sample for the most likely parameters for the individual event

Repeat until a sufficient number of events have been sampled, then collate all the sampler outputs into a single file

Set the output file as the posteriors for MDF parameter estimation and calculate the likelihood

Samples parameters for the MDF to identify the most likely

Compare to the initial MDF for the Universe

Parameter	Set Value
α	3.71
mmin	14.927
mmax	75.3890
λ	0.75
μm	32.56
σm	2.49
δm	2

My Simulations

Applying the SNR threshold has have implications for the MDF:

- Lower mass events are less likely to make it past the cut
- Same for more distant events
- There is therefore an inherent bias towards louder events, which are either closer, higher mass or both (i.e. Malmquist bias)

This is to simulate which events would actually be flagged as significant by GW pipelines, and so which would continue to further analysis

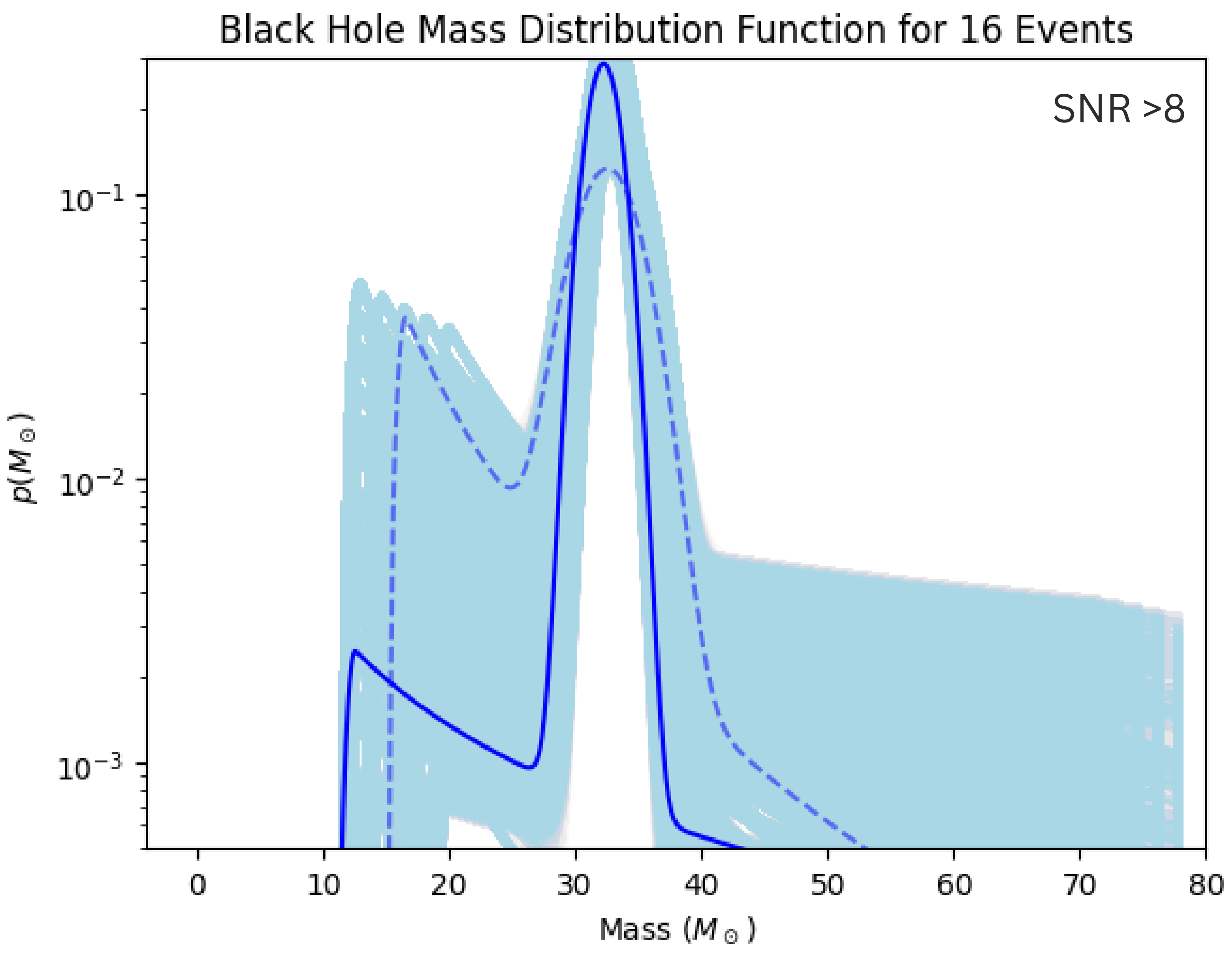
Ideal conditions are assumed within the simulations- the fixed extrinsic parameters are fixed to maximise their SNR (i.e. the gravitational waves are perfectly aligned with the IFO). This isn’t very realistic, but makes drawing conclusions easier

Some Meta-Data

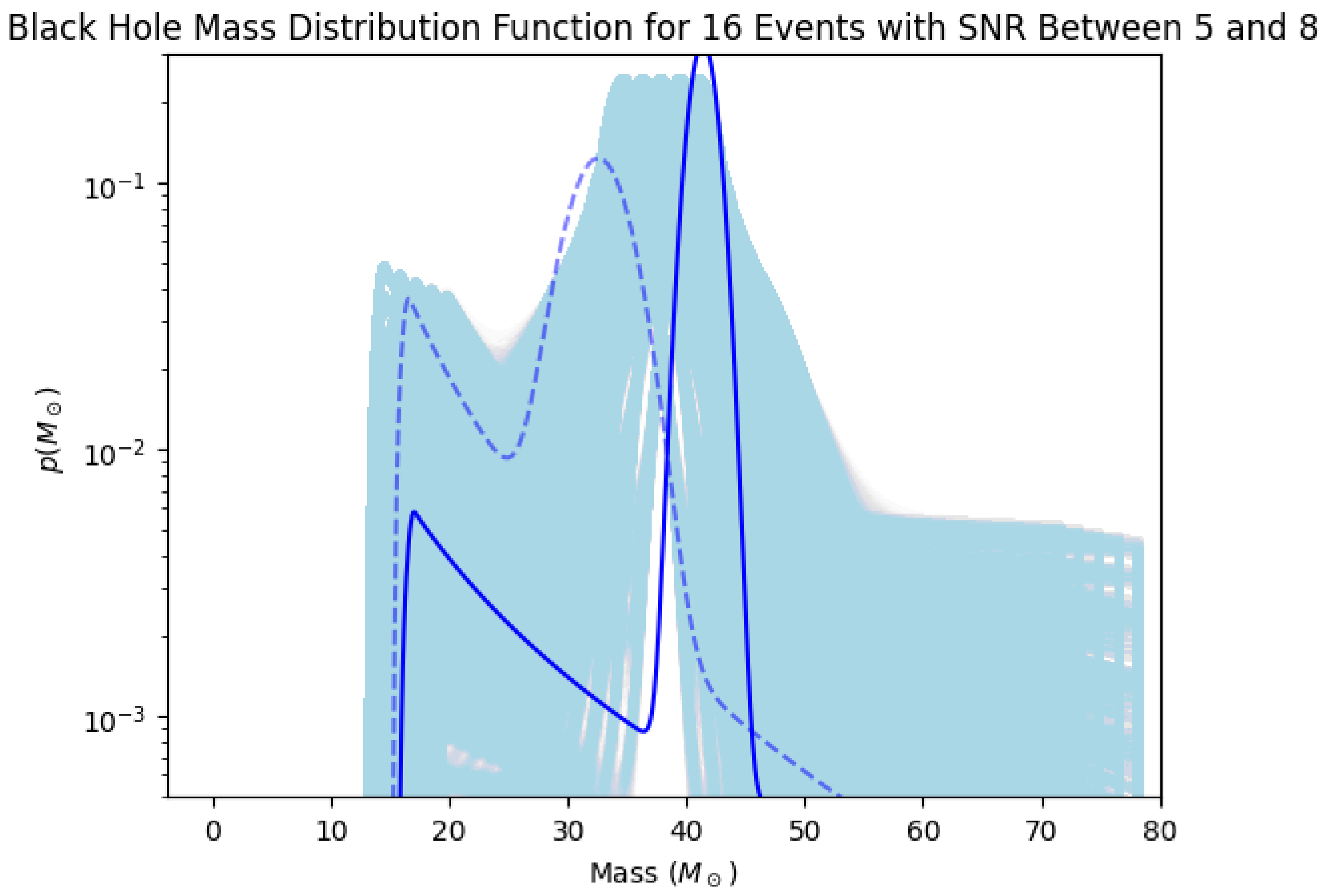
- Number of files generated
  - 17822 Singular Events Generated
  - 642 Events then sent through parameter estimation
  - 146 Events in the final “proper” data set
  - 128 Total Unique Events fed into the population parameter estimation
  - 99.3% of initial events are discarded

- Volume of data generated
  - 15.2GB data generated in total from the first event log to the final plot
  - Of this, 5.2 MB corresponds to data used in the generation of the final results
  - This is an efficiency of 34.2% - pretty good!
  - Disregarding files from failed runs, the maximum efficiency for data used/data created is 99.94%- even better!

Results

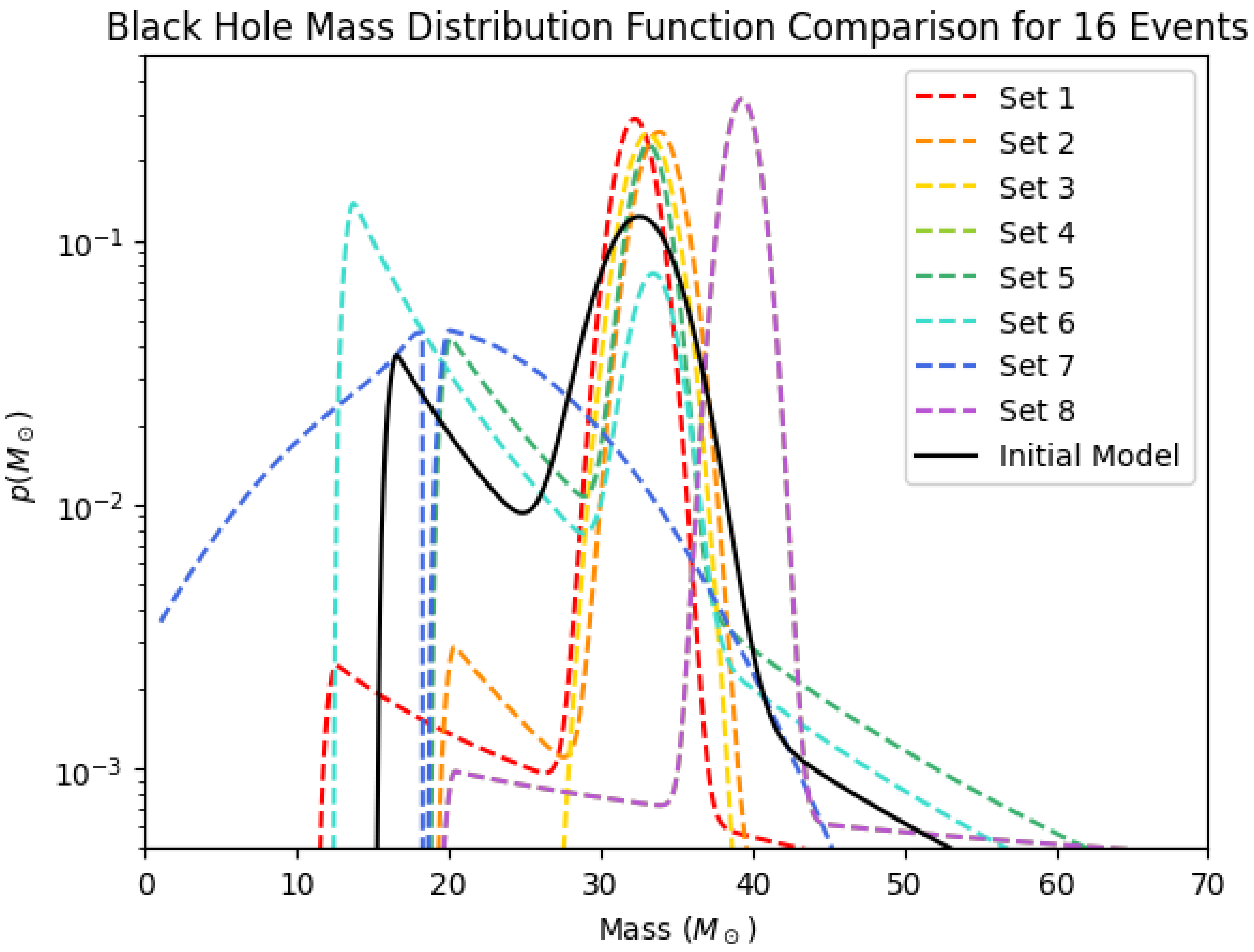


Black Hole Mass Distribution Function for 16 Events
SNR >8
10^{-1}
10^{-2}
10^{-3}
$p(M_\odot)$
0
10
20
30
40
50
60
70
80
Mass ($M_\odot$)



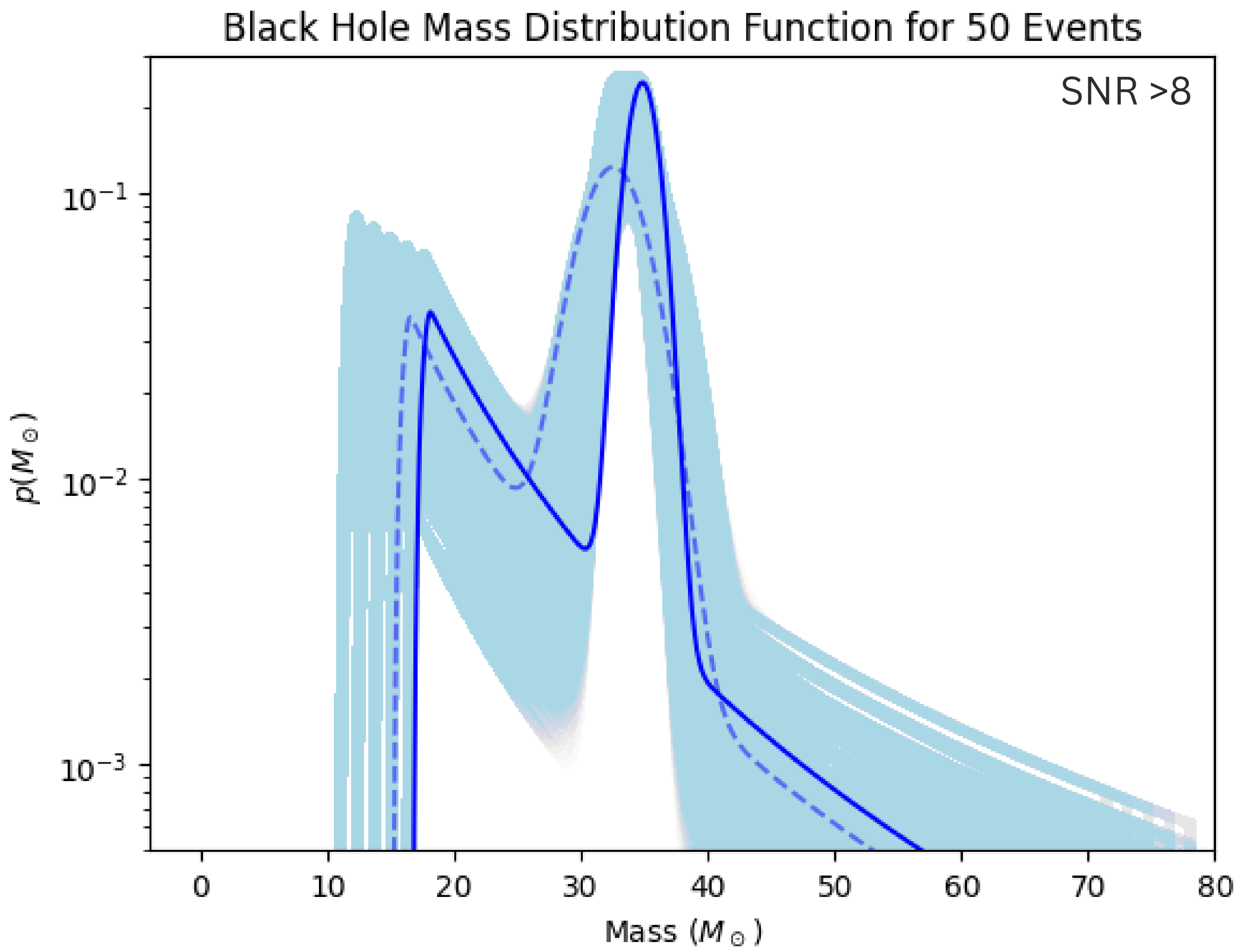
Black Hole Mass Distribution Function for 16 Events with SNR Between 5 and 8
10^{-1}
10^{-2}
10^{-3}
$p(M_\odot)$
0
10
20
30
40
50
60
70
80
Mass ($M_\odot$)

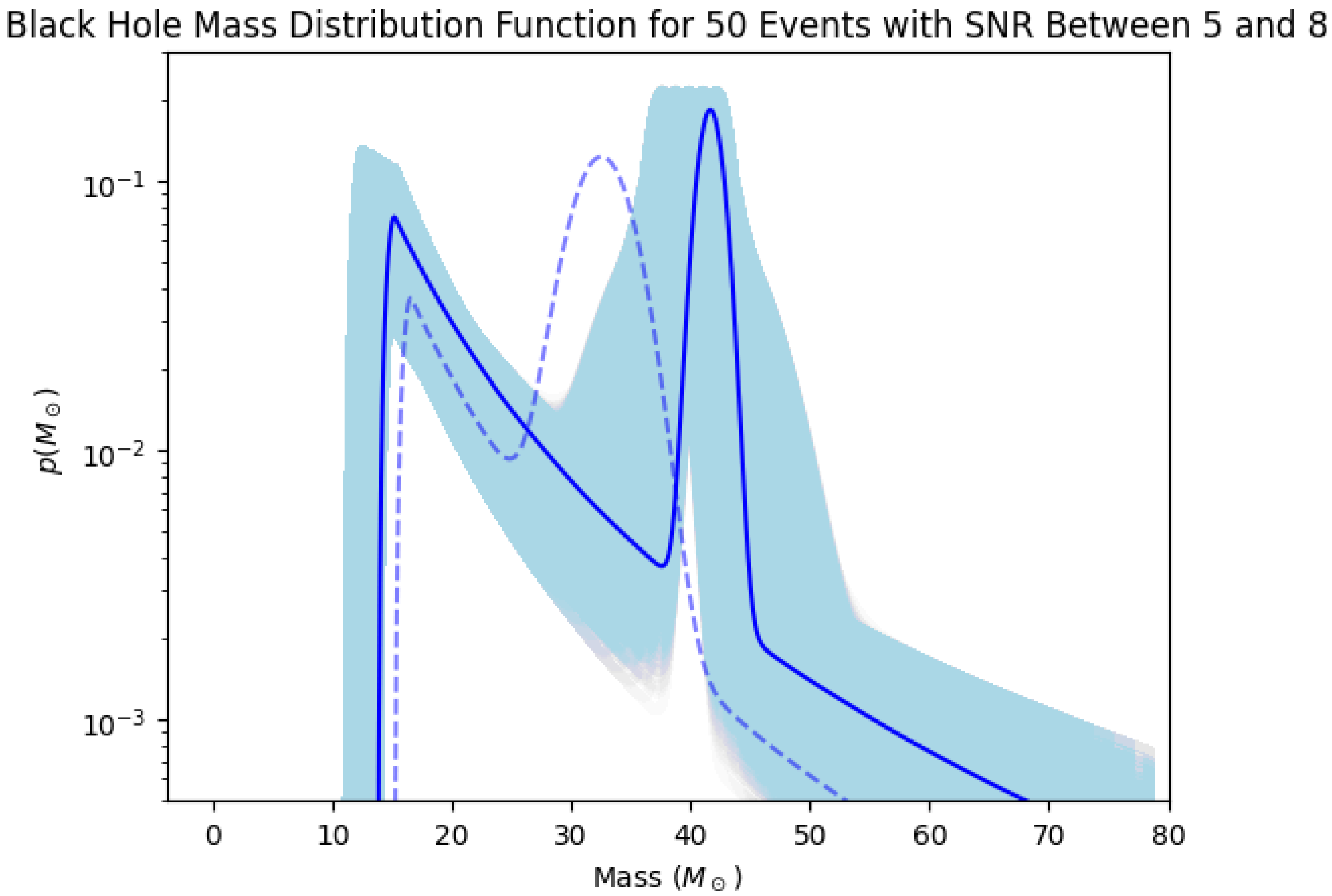
Results



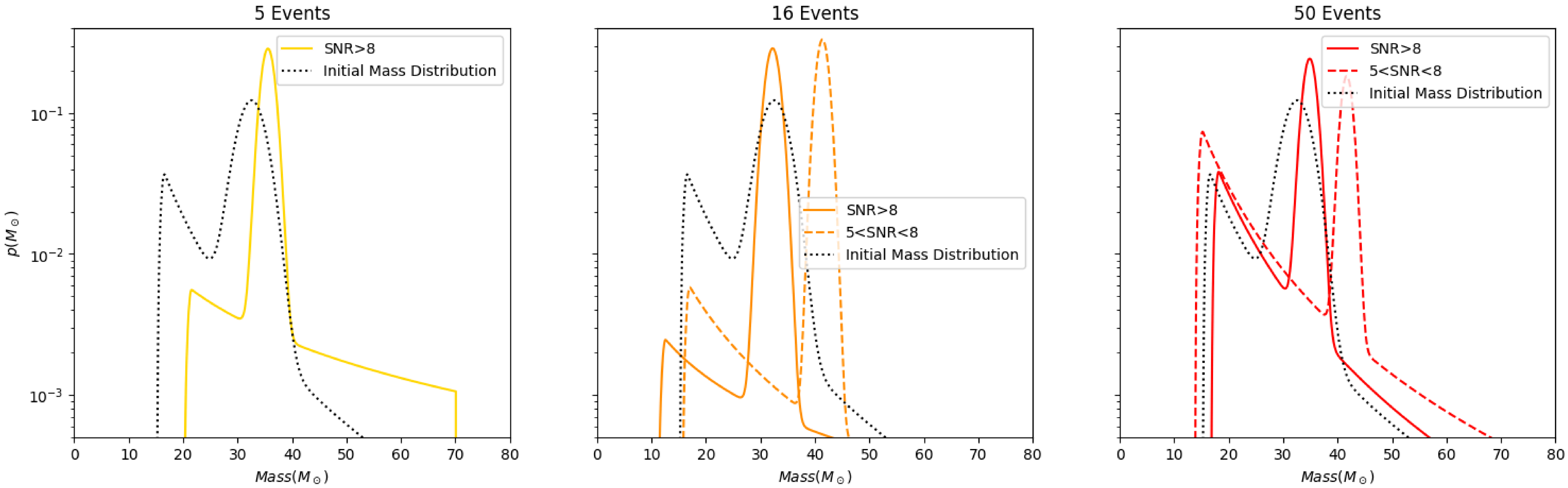
Black Hole Mass Distribution Function Comparison for 16 Events
$p(M_\odot)$
10^{-1}
10^{-2}
10^{-3}
0
10
20
30
40
50
60
70
Mass ($M_\odot$)
Set 1
Set 2
Set 3
Set 4
Set 5
Set 6
Set 7
Set 8
Initial Model

Results


Black Hole Mass Distribution Function for 50 Events
SNR >8
$p(M_\odot)$
10^{-1}
10^{-2}
10^{-3}
0
10
20
30
40
50
60
70
80
Mass ($M_\odot$)


Black Hole Mass Distribution Function for 50 Events with SNR Between 5 and 8
$p(M_\odot)$
10^{-1}
10^{-2}
10^{-3}
0
10
20
30
40
50
60
70
80
Mass ($M_\odot$)

Results



5 Events
SNR>8
Initial Mass Distribution
$p(M_\odot)$
10^{-1}
10^{-2}
10^{-3}
0 10 20 30 40 50 60 70 80
$Mass(M_\odot)$
16 Events
SNR>8
5<SNR<8
Initial Mass Distribution
0 10 20 30 40 50 60 70 80
$Mass(M_\odot)$
50 Events
SNR>8
5<SNR<8
Initial Mass Distribution
0 10 20 30 40 50 60 70 80
$Mass(M_\odot)$

Further Research

Model Comparison

- Comparing Different MDFs
- Comparing how different waveform approximations affect the MDF

Running parameter estimation on real GWTC data

- Comparing with Simulated

Estimating parameters for larger data sets

- Better Data output
- Conclusions could be drawn about the accuracy of MDFs derived from the real GWTC data

Further Exploration of Bias

- How much does the Malmquist bias caused by the SNR cut (applied due to the IFO sensitivity) affect the MDF?
- Can this be accurately adjusted for?

Introducing Spin

Thank you

Any Questions?