

# School of Physics and Astronomy



Improved Methods for the Detection of Gravitational  
Waves associated with Gamma-Ray Bursts

Iain Dorrington

Submitted for the degree of Doctor of Philosophy  
School of Physics and Astronomy  
Cardiff University

16th May 2019

# Summary of thesis

In this thesis we will see two targeted searches for gravitational wave (GW) signals associated with gamma-ray bursts (GRBs) and the work that is ongoing to make these two search pipelines faster and more sensitive. The first of these is PyGRB, a matched filter search that follows up short GRB detections. The second is X-pipeline, a burst search for both long and short GRBs.

We will begin with a chapter on GW background, where we will show that GWs are a consequence of General Relativity. We will look at some known and potential sources of GWs, and look into the basics of how GWs are detected.

We then summarise the current state of GRB science. We will look at the history of GRB astronomy, starting with their accidental discovery and ending with a summary of the key scientific findings that came from the first joint detection of a GW signal associated with a GRB. The second half of this chapter focuses on multi-messenger astronomy, how it is done and what it can teach us about the universe.

Chapter 4 is on PyGRB. This chapter begins with a look at the theory behind how the pipeline works, and then goes into the detail of how the pipeline works in practice. This chapter ends with a summary of the results from the PyGRB search of the latest observing run O2, including the detection of GW 170817 and the implications of this for the astrophysical rates of neutron star mergers.

We then look at the PyGRB redevelopment work that is currently ongoing. The code used in O2 is now dated and slow by comparison to work that is going on with other matched filter searches. We will look at the work that has already been done to rewrite the PyGRB pipeline to use the optimised code of the PyCBC software package. We will look at the speed improvements this brings, learn why it is scientifically important that the code runs quickly, and also see some of the new tools that integrating into PyCBC has made available for the GRB search.

We end by looking at X-pipeline. We will start by understanding the theory behind how X-pipeline works and looking at the results from the X-pipeline analysis of GRBs in O2. The second half of this chapter looks at how machine learning can be used to improve the sensitivity and speed of X-pipeline. We end this chapter with a discussion of future work that can be done on this pipeline, and some of the issues that arise from using machine learning.

# Declaration of authorship

- **DECLARATION:**

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed: ..... (candidate) Date: .....

- **STATEMENT 1:**

This thesis is being submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (PhD).

Signed: ..... (candidate) Date: .....

- **STATEMENT 2:**

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references.

Signed: ..... (candidate) Date: .....

- **STATEMENT 3**

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed: ..... (candidate) Date: .....

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Introduction to Gravitational Wave Astronomy</b>	<b>4</b>
2.1	Gravitational Waves . . . . .	4
2.1.1	Linearised Gravity . . . . .	4
2.1.2	Gauge Transformation . . . . .	5
2.1.3	Physical Effects of Gravitational Waves . . . . .	6
2.2	Gravitational Wave Generation . . . . .	8
2.3	Gravitational Wave Sources . . . . .	10
2.4	Gravitational Wave Detectors . . . . .	12
2.4.1	The Michelson Interferometer . . . . .	13
2.4.2	Interferometer Antenna Pattern . . . . .	17
2.4.3	Interferometer Noise . . . . .	19
2.4.4	The Global Network of Interferometers . . . . .	22
<b>3</b>	<b>Gamma-Ray Bursts</b>	<b>23</b>
3.1	The History of Gamma-ray Burst Astronomy . . . . .	23
3.1.1	Cold War Tension and an Unexpected Discovery . . . . .	23
3.1.2	BATSE and the Galactic/Extra-galactic Controversy . . . . .	25
3.1.3	The Long and Short of Gamma-ray Bursts . . . . .	27
3.1.4	BeppoSAX and the First Afterglows . . . . .	28
3.1.5	The Fireball Model . . . . .	30
3.1.6	Jets . . . . .	31
3.1.7	The Long GRB-Supernova Connection . . . . .	33
3.1.8	The Short GRB-Compact Binary Connection . . . . .	33
3.2	Gravitational Waves and GRB 170817A . . . . .	36
3.2.1	Initial Observation and Followup . . . . .	36
3.2.2	Kilonova Observation . . . . .	38
3.2.3	Structured Jets . . . . .	38
3.2.4	Other Scientific Results . . . . .	44
3.3	GRB Progenitors . . . . .	48
3.3.1	Compact Binary Coalescence . . . . .	48
3.3.2	Core Collapse . . . . .	49
3.4	GRB Detectors . . . . .	50
3.4.1	Swift . . . . .	50
3.4.2	Fermi . . . . .	50
3.4.3	The InterPlanetary Network . . . . .	50
3.5	GRB Gravitational Wave Search Strategies . . . . .	51
3.5.1	Triggered and Untriggered Searches . . . . .	51

3.5.2	Modeled and Burst Searches . . . . .	52
3.6	GRB Astrophysics with Gravitational Waves . . . . .	52
<b>4</b>	<b>A Targeted Search for Gravitational Waves associated with Short GRBs</b>	<b>55</b>
4.1	Introduction . . . . .	55
4.2	Coherent Matched Filtering . . . . .	56
4.2.1	Binary Coalescence Waveform . . . . .	56
4.2.2	Coherent SNR . . . . .	58
4.2.3	Comparison to Coincident Search . . . . .	60
4.2.4	Null SNR . . . . .	62
4.2.5	Searching for Face on Signals . . . . .	63
4.2.6	Coherent $\chi^2$ Tests . . . . .	64
4.3	The PyGRB Workflow . . . . .	68
4.3.1	Background Noise Rejection . . . . .	68
4.3.2	Reweighted SNR . . . . .	69
4.3.3	Event Significance . . . . .	70
4.3.4	Searching over a Sky Patch . . . . .	72
4.3.5	Calculating Search Sensitivity . . . . .	73
4.3.6	The PyGRB Workflow . . . . .	74
4.4	O2 PyGRB Search . . . . .	76
4.4.1	GRB sample . . . . .	76
4.4.2	Results . . . . .	76
4.4.3	Rates . . . . .	78
<b>5</b>	<b>The Future of PyGRB</b>	<b>81</b>
5.1	Reasons to Rewrite the Pipeline . . . . .	81
5.2	Workflow Modifications . . . . .	83
5.2.1	Consistency Checks . . . . .	83
5.2.2	Detection Statistic . . . . .	84
5.2.3	The new Coherent Matched Filtering Executable . . . . .	85
5.3	Search Performance . . . . .	86
5.3.1	Testing the Detection Statistic . . . . .	87
5.3.2	Computational Cost . . . . .	96
5.4	Future Plans . . . . .	98
<b>6</b>	<b>A Search for Unmodelled Gravitational Wave Signals using Machine Learning</b>	<b>100</b>
6.1	X-pipeline . . . . .	100
6.1.1	Burst Search Background . . . . .	103
6.1.2	Standard Likelihood . . . . .	104
6.1.3	Null Energy . . . . .	106
6.1.4	Incoherent Energy and Background Rejection . . . . .	106
6.1.5	Tuning and Trigger Significance . . . . .	107
6.1.6	O2 X-pipeline Search . . . . .	108
6.2	Machine Learning . . . . .	110
6.2.1	Supervised Machine Learning . . . . .	110
6.2.2	Boosted Decision Trees . . . . .	113
6.2.3	Training Data Preprocessing . . . . .	116
6.2.4	Optimisation and Validation . . . . .	117

6.3	Effectiveness of XTMVA . . . . .	123
6.3.1	Analysis Setup . . . . .	123
6.3.2	Sensitivity and Speed Comparison . . . . .	126
6.3.3	XTMVA Search Results . . . . .	129
6.4	Discussion and Future Work . . . . .	131
6.4.1	Low p-value Triggers . . . . .	131
6.4.2	Long Injections . . . . .	131
6.4.3	Machine Learning Developments . . . . .	134
6.4.4	Comparison to other Machine Learning Approaches . . . . .	137
<b>7</b>	<b>Conclusions</b>	<b>138</b>
<b>Appendices</b>		<b>141</b>
.1	General Relativity . . . . .	141
.2	PyGRB O2 Result Tables . . . . .	141
.3	MVA Results Tables . . . . .	144

# List of Figures

2.1	<b>Plus and Cross Polarisation.</b> The effect of the plus and cross polarisation of a GW on a ring of test particles. [27]	8
2.2	<b>The First Direct Detection of a Gravitational Wave.</b> The top panel shows the GW waveform with an inset cartoon of the state of the binary at four different phases of the binary inspiral. The bottom panel shows the Keplerian separation distance in units of Schwarzschild radii ( $R_s = \frac{2GM}{c^2}$ where $M$ is the total mass of the system). Also shown is the relative velocity of the black holes. [5]	13
2.3	Two identical schematics of a Michelson interferometer, each with a different naming convention. [10]	14
2.4	This schematic Michelson Interferometer shows our labelling convention for the different Electric fields. [10]	15
2.5	This plot shows how the detector output varies with the change in length of the interferometer arms. [10]	17
2.6	<b>Interferometer Antenna Pattern.</b> Here we see the antenna response on an interferometer with arms aligned with the x and y axes. The detector is most sensitive to GWs coming from perpendicular to the detector plane, while GWs approaching from within the detector plane but at an angle off-set from the arms by $\pi/4$ radians will be in a null of the detector. [33]	19
2.7	<b>ASD for LIGO Livingston Observatory.</b> Plotted in blue is the ASD for the LIGO Livingston Observatory as it was on the 20th August 2017. In grey we see the Gravitational Wave Interferometer Noise Curve (GWINC), a theoretical model of all the noise in the detector, as well as the O2 projected noise curve in orange.	20
2.8	<b>Noise Budget for LIGO Livingston Observatory.</b> Here we have the noise budget for the LIGO Livingston Observatory on 20th August 2017. The noise sources in LIGO detectors are described in detail in [26] and [3].	21
3.1	<b>BATSE gamma-ray light curves.</b> [31]	24
3.2	<b>BATSE GRB Fluence.</b> This plot shows the fluence (given by the colour of each point) and the sky position of each GRB detected by the BATSE mission. [29]	26

3.3	<b>LogN-logP for BATSEVPO.</b> Here we plot the log of the number of GRBs against the log of the peak flux. The sample includes GRBs detected by BATSE and by PVO. The energy range for BATSE was 50-300keV, and the energy range for PVO was 100-2000keV. For uniformly distributed GRBs, we expect this plot to have a gradient of $-3/2$ . The expected gradient is observed for high energy GRBs but not at lower energies. This suggests a limited distance to which GRBs can be observed. <a href="https://www.researchgate.net/figure/Distribution-log-N-log-P-for-a-combined-set-of-BATSEPV0-data-The-distributions-ma">cite https://www.researchgate.net/figure/Distribution-log-N-log-P-for-a-combined-set-of-BATSEPV0-data-The-distributions-ma</a> <a href="https://www.nature.com/articles/366040a0.pdf">fig5_242389649</a> <a href="https://www.nature.com/articles/366040a0.pdf">cite https://www.nature.com/articles/366040a0.pdf</a> .	27
3.4	<b>T90 vs the spectral hardness ratio.</b> Here we plot the $T_{90}$ values and the spectral hardness ratio for the BATSE GRBs. The top panel shows a histogram of the T90 data, which clearly has two populations of GRBs, short and long. The main plot shows T90 against spectral hardness, which makes the two populations even more clear and shows that short GRBs have harder spectra than long GRBs. Those GRBs with the greatest ratio of energy in the X-ray to gamma ray band, generally those with a peak energy of less than 15 keV, are called <i>X-ray flashes</i> (XRF). Those with comparable energy in the gamma-ray and X-ray band are called <i>X-ray rich</i> (XRR) GRBs. All other GRBs are simply called GRBs. These different classes of GRBs are marked on the plot. Also shown is the 2 second dividing line between short and long GRBs. <a href="https://arxiv.org/abs/1209.5029">cite What are Gamma-ray Bursts? Joshua Bloom</a>	29
3.5	<b>Break in spectrum due to jetting.</b> Here we see the optical light curves for the afterglow GRB 990510. A break in the spectrum is visible after approximately one day. <a href="https://iopscience.iop.org/article/10.1086/312282/fulltext/">cite https://iopscience.iop.org/article/10.1086/312282/fulltext/</a>	32
3.6	<b>GRB170817A and GW170817.</b> Here we see that the a coherent combination of the Hanford and Livingston strain data from GW170817 in the bottom panel. The top two panels shows the Fermi GRM curve in the 10-50keV and the 50-300keV range respectively. The INTEGRAL/SPI-ACS data is shown in the third plot. The background estimate for each detector is indicated by the red line. Note that the GRB was detected 1.7 seconds after the GW signal was detected. We can also see that Fermi detected a longer, softer signal in the 10-50 keV range, that lasted for a few seconds after the triggering pulse. <a href="https://arxiv.org/abs/1710.05833">cite Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A</a>	37

3.7	<b>Glitch in the LIGO Livingston Observatory.</b> The top panel shows a time frequency map for the Livingston observatory data at the detection time of GW 170817. A glitch is clearly visible approximately 1.5 seconds before the end of the signal. Despite this the signal is still clearly visible. The bottom plot shows the raw strain data from the Livingston observatory. This data is bandpassed between 30 Hz and 2 kHz to emphasise the sensitive range of the detector. The grey curve (and right axis) shows the inverse Tukey window used to smoothly zero out the data around the glitch before the rapid reanalysis of the data. The blue curve shows the waveform model used to subtract the glitch from the data before measurements of the source's properties were made. <b>cite GW170817 observation from BNS</b>	39
3.8	<b>GW 170817 Detection.</b> Here we see time frequency maps of the LIGO Hanford and Livingston observatories, and the Virgo observatory at the detection time of GW 170817. This data has been whitened and independently observable noise sources have been subtracted, including a glitch in the Livingston data. The non-detection by Virgo significantly reduced the amount of the sky that the signal could have originated from. <b>cite GW170817 observation from BNS</b>	40
3.9	<b>Sky map for GW170817/GRB170817A.</b> Here we see the 90% confidence sky localisation for the LIGO and Virgo collaborations in green, the GBM 90% localisation in purple, and the annulus formed by Fermi and INTEGRAL timing information in grey. <b>cite Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A</b>	41
3.10	<b>NGC 4993.</b> Image of NGC 4993 taken in 1992 by the Anglo-Australian Observatory (left) and August 18th 2017 by the Las Cumbres Observatory (right). <b>cite <a href="https://www.nature.com/articles/nature24291/figures">https://www.nature.com/articles/nature24291/figures</a></b>	41
3.11	<b>Brightness/Luminosity against redshift.</b> Here we see the distribution of $E_{\text{iso}}$ and $L_{\text{iso}}$ against redshift for every GBM-detected GRB with a measured refshift. For GRBs with power law spectra, marked with a downward pointing arrow, this is taken to be an upper limit. This is because the spectra must have curvature, and so extrapolating a power law leads to an overestimation. The green dashed line shows the approximate detection threshold for the GBM. These plots show that GRB 170817A was much dimmer than any other detected GRB. <b>[cite GRB BNS paper]</b>	42
3.12	<b>Jet Structure Scenarios.</b> Three different scenarios that could explain the low luminosity of GRB 170817A. The first scenario is that a Top-hat jet was viewed off-axis. The second is that the jet is structured, with photons emitted further from the axis being lower energy and fewer in number, and viewed relatively far from the axis. The third scenario is that a uniform jet has a surrounding cocoon that emits lower energy photons, and it was these lower energy photons that were detected. <b>cite GRB BNS paper</b>	43

3.13	<b>Structured Jet.</b> Left panel: A pseudocolour density image of the simulation used to compute the afterglow curves [citation in afterglow paper]. The low density core of the jet is the blue region near the middle. The orange and green region around the core is the slow moving wings. Top right panel: Here we see the 3 GHz flux detected from different parts of the structured jet as time progresses. The angle is relative jet axis, so the blue curve is the core of the jet, the orange curve is the fast wings of the jet, the orange curve is the material moving along the line of sight (an angle of about $33^\circ$ in this case), and the pink and brown curves correspond to large angles, that do not contribute much to the observed flux. Bottom right panel: The distribution of energy as a function of angular separation from the jet. cite GRB afterglow paper paper . . . . .	45
3.14	<b>Jet Model Comparison.</b> Here we see a comparison of the best fit for the structured jet, Top-hat jet seen off-axis, and isotropic models. The afterglow's measured flux density at 3 GHz is shown by the blue symbols (though the fits were performed with multiwavelength data). The inset shows the best fit isotropic energy and Lorentz factor for each model as a function of viewing angle. The arrows show the position of the observer for the structured and Top-hat jet models. cite GRB afterglow paper paper . . . . .	46
4.1	<b>Null Statistic Cut.</b> Here we plot the coherent SNR against the null SNR. The blue crosses are background triggers. The red pluses are signal injections. The black line is the veto line, with all triggers in the shaded region above the line being discarded. The green line indicates the expected SNR for optimally oriented injections. The magenta line shows the one sigma error on the green line. . . . .	69
4.2	<b>PyGRB Sky Grid.</b> Here we see an example of a full search grid used by PyGRB, indicated by the blue dots, and a reduced sky grid parsed by PyGRB in the case of a two detector search using the Hanford and Livingston detectors, the empty circles labeled 'parsed'. The parsed circles do not form a line due to the parsing routine, but this has no effect on analysis.[40] . . . . .	73
4.3	<b>PyGRB Workflow.</b> The workflow starts in two parallel branches, one that runs the injections jobs, and one that analyses the background and onsource. . . . .	75
4.4	<b>P-value for each GRB.</b> This is the p-value distribution for the 41 GRBs other than GRB 170817A. The GRBs with no trigger in the onsource window have upper and lower limits on the p-value. The upper limit is a p-value of 1. The lower limit is the fraction of offsource trials that also had no trigger. The distribution lies within the $2\sigma$ range, shown by the upper and lower dotted lines. . . . .	77
4.5	<b>Cumulative exclusion distance.</b> This is the cumulative 90% exclusion distance for every GRB analysed by PyGRB except GRB170817A. The 90% exclusion distance is the distance at which 90% of injected simulated signals are recovered with a greater coherent SNR than the loudest trigger in the onsource. . . . .	78

4.6	<b>Cumulative Rate of BNS and short GRB Events.</b> The magenta lines show the 90% confidence bounds for joint GRB-GW events as a function of redshift. This was calculated using the 41 non-detections and single detection by PyGRB during O2. The black line and the grey region shows the estimated BNS merger rate $1210^{+3230}_{-1040}$ . In green is shown the estimated Fermi detection rate and its 90% confidence region[21]. The measured redshifts of every short GRB apart from GRB 170817A are shown in brown. The gold sample refers to those GRBs that were localised to near a host galaxy, making the redshift measurement more reliable than short GRBs measured more distant from a host galaxy. Our results are compatible with both the Fermi-GBM observed rate and the predicted BNS merger rate.[11] . . . . .	80
5.1	<b>Coherent and Reweighted SNR vs Time for GRB 170817A.</b> The top panel shows the coherent SNR vs time for GRB 170817A. The GW is clearly visible, as are some smaller peaks that are due to noise. The bottom panel shows the null and $\chi^2$ reweighted SNR. We can see that GW is still very prominent, but the peaks in coherent SNR that were due to noise have been downweighted. . . . .	88
5.2	<b>Coherent and Reweighted SNR vs Time for GRB 170112A.</b> The top panel shows the coherent SNR vs time for GRB 170112A. There is no GW, but several glitches are clearly visible. The bottom panel shows the $\chi^2$ reweighted SNR. We can see that the glitches have been downweighted to be no more significant than the background. . .	88
5.3	<b>Null SNR vs Reweighted SNR for GRB 170817A.</b> Here we plot the null SNR against the coherent SNR (left) and the reweighted SNR (right) for GRB 170817A. Only triggers with a null SNR above 4.5 are reweighted by the null SNR (with the other triggers being reweighted only by their $\chi^2$ values), and we can see that these triggers have been downweighted more than the triggers with a low null SNR. The GW is clearly visible on the right of both plots. . . . .	89
5.4	<b>Reweighted SNR vs Network <math>\chi^2</math> for GRB 170817A.</b> Here we plots the network $\chi^2$ against the coherent SNR (left) and reweighted SNR (right) for GRB 170817A. We can see that triggers with the higher network $\chi^2$ of a trigger, the more it is downweighted. The GW is clearly visible on the right of both plots. . . . .	90
5.5	<b>Coherent SNR vs Network <math>\chi^2</math> for GRB 170112A.</b> Here we plot the network $\chi^2$ against the coherent SNR (left) and reweighted SNR (right) for GRB 1710112A. The data around this GRB contained several glitches, which are apparent from the triggers with a high network $\chi^2$ and coherent SNR. As we can see, these triggers are downweighted appropriately, such that the reweighted SNR contains no peaks. . .	91

5.6 Loudest Event per 6-second Trial for GRB 170817A.	Here we see the peak coherent SNR (blue line) and reweighted SNR (orange line) in each 6-second trial for the GRB 170817A analysis, with the stars indicating the GW. Again, we can see that the on-source has a much higher reweighted SNR than any of the off-source trials. Also note that the tail of events with a high coherent SNR of about 7-9, seen in figure ??, does not appear in this plot. This shows that the reweighting is lowering the significance of glitches, as it is supposed to do. <code>fix caption</code>	92
5.7 Loudest Event per 6-second Trial for GRB 170112A.	Here we see the peak coherent SNR (blue line) and reweighted SNR (orange line) in each 6-second trial for the GRB 170112A analysis. The peak coherent and reweighted SNR in the on-source trial are indicated by the red and blue stars respectively. We see that the on-source results are consistent with background. We can also see that reweighting the SNR removed the long tail of high SNR glitches for this analysis.	93
5.8 Injection Distance against Time.	170817 run	93
5.9 Injection Distance against Time.	This is the distance (MPc) vs time plot for the BNS injection run in the PyGRB O2 analysis of GRB 170817A. Blue crosses indicate that the injection was found and was more significant than event in the background data. Red crosses indicate that a trigger was found that was coincident with the injection, but it was vetoed. Black crosses indicate that no trigger was found that was coincident with the trigger. Coloured circles indicate that the injection was found but was not louder than all of the background, and in this case the circle colour indicates the FAP of the trigger. We can see that nearby injections are almost always found, and typically with a low FAP. More distant injections tend to be vetoed, missed completely, or have a relatively high FAP. It is at a distance of about 200 Mpc that injections start to be missed.	94
5.10 Injection Distance against Chirp Mass.	170817 run	95
5.11 Injection Distance against Chirp Mass.	This plot uses the same conventions and injection set as figure 5.9, but shows the distance (Mpc) against the chirp mass of the signal injections. We can see that, in general, the higher the chirp mass, the lower FAP of the injection, as we would expect.	95
6.1 X-pipeline Time-Frequency Map	This figure shows a time-frequency map from X-pipeline for a $1.4 - 10 M_{\odot}$ NSBH merger using simulated background from the two Hanford detectors. The top figure shows a coherent signal stream called the <i>standard likelihood</i> $E_{\text{SL}}$ and the bottom figure shows the top 1% of pixels. [cite xp paper]	101

6.2	<b>X-pipeline Background Rejection Test.</b> This figure shows an example of X-pipeline background rejection. The axes show two of the statistics that X-pipeline calculates. Specifically, the x-axis shows the coherent null energy and the y-axis shows the incoherent null energy (see section 6.1.4 for more details on these statistics). The red squares show simulated GW signals, and the crosses show background triggers. The colour bar shows the base 10 logarithm of the significance of each trigger. The injection amplitude plotted is chosen such that approximately 90% of injections will survive the cut. Hence, the cut eliminates most of the noise but only a few signals. [ <a href="#">cite X-pipeline paper</a> ]	102
6.3	<b>Cumulative Distribution of p-values.</b> Here we plotted the p-values for every GRB analysed by X-pipeline in O2 apart from GRB 170817A. Also plotted is the expected distribution and the $2\sigma$ deviation. The results are consistent with the no-signal hypothesis. [ <a href="#">cite O2 GRB paper</a> ]	109
6.4	<b>Cumulative Distribution of Exclusion Distance.</b> Here we plotted the 90% exclusion distance for every GRB analysed by X-pipeline in O2 apart from GRB 170817A. This is the distance to which 90% of injections can be recovered with a significance greater than the loudest event in the on-source. [ <a href="#">cite O2 GRB paper</a> ]	110
6.5	<b>Schematic Decision Tree.</b> To determine if a trigger is a signal or noise event the tree makes a series of cuts on the attributes $x[N]$ . If the inequality in a node is true, then the next node is the branch to the left. Otherwise the next node is the one to the right. The properties of the tree, such as the number of layers it has, are set by the user (see section 6.2.4).	114
6.6	<b>Visualising the Classifier.</b> In the top plot you can see the value for $\log(E_{\text{null}})$ and $\log(I_{\text{null}})$ for all the signal and background training data used to build the classifier. We chose one of these events at random (indicated by the star) and varied $E_{\text{null}}$ and $I_{\text{null}}$ to see how it changed the MVA score, indicated by the colour in the bottom plot. As we can see, increasing $I_{\text{null}}$ and decreasing $E_{\text{null}}$ leads to the event being more likely to be classed as a signal. This is akin to the X-pipeline cut shown in figure 6.2.	115
6.7	<b>Removing WNB and Cusp Waveforms from Training Set.</b> Here we plot the percentage change in 50% upper limit injection scale per waveform after removing WNB and Cusp waveforms from the training set. Negative values indicate that the sensitivity improved after the change. We see that the sensitivity to most waveforms drops, but by less than 1%. As we use a few hundred injections at each injection scale, this is not a statistically significant result. This shows that the MVA is able to detect GW morphologies that it has not been trained on.	120

6.8 <b>Effect of Hyperparameter Optimisation.</b> Here we see the effects of optimisation on the 50% upper limit injection scale. Lower values indicate a more sensitive search. The top panel shows the absolute values and the bottom panel shows the percentage change. The benefits of optimising the hyperparameters is no more than a $\sim 3\%$ improvement in sensitivity when compared to the default settings of the TMVA boosted decision tree classifier. It is also interesting to note that the three waveforms that have their sensitivity drop after optimisation (adi-a, adi-c, and adi-d) are all long waveforms. We will discuss the problem with these waveforms in section 6.4. . . . .	124
6.9 <b>MVA Improvement.</b> Here we see the effects of using the MVA on the 50% upper limit injection scale for the same GRB that was used for optimisation. The top panel shows the absolute values and the bottom panel shows the percentage change. We can see that the MVA outperforms X-pipeline on every waveform. As this was the GRB used to optimise the hyperparameters, it cannot be guaranteed that these results will hold for other GRB analyses. . . . .	125
6.10 <b>X-pipeline and XTMVA ADI-a 50% Injection Scale Upper Limit by GRB.</b> Here we plot the sensitivity to the ADI-a waveform of both X-pipeline and XTMVA. The lower injection scales for XTMVA show that XTMVA is more sensitive than X-pipeline to the ADI-a waveform. Also, note the lower variation in injection scale between GRBs for XTMVA, suggesting that XTMVA is more stable than X-pipeline. . . . .	126
6.11 <b>X-pipeline and XTMVA CSG 50% Injection Scale Upper Limit by GRB.</b> Here we plot the sensitivity to the 150 Hz circular sine-Gaussian waveform of both X-pipeline and XTMVA. The lower injection scales for XTMVA show that XTMVA is more sensitive than X-pipeline to the CSG waveform. Again, note the lower variation in injection scale between GRBs for XTMVA, suggesting that XTMVA is more stable than X-pipeline. . . . .	127
6.12 <b>Median 50% Injection Scale Upper Limit by Waveform.</b> Here we plot the median sensitivity to each waveform of both X-pipeline and XTMVA. Overall, XTMVA is more sensitive, especially to shorter waveforms such as sine-Gaussians. Apart from ADI-a, the MVA is worse than X-pipeline for long waveforms, though the difference is small. If the MVA could use the long injection code that X-pipeline uses, it is reasonable to expect that the MVA would outperform X-pipeline for long waveforms as well. . . . .	128

6.13 <b>XTMVA p-values.</b> Here we have plotted the p-values for 13 of the GRBs analysed with the MVA. The blue triangles indicate the p-value reported by the MVA, the black dotted lines show the expected distribution and a $2\sigma$ deviation. Two GRBs that were analysed were left out from this plot: GRB 170817A as it had a known GW counterpart and E264930 as it was used to tune the hyperparameters. The analysis shows some bias towards low p-values. In particular, two out of the 13 analysed GRBs have a p-value of $\sim 1\%$ . This can be compared to figure 6.3 which shows the X-pipeline p-values for O2. In particular, X-pipeline did not have the same bias towards low p-values that XTMVA does. This needs further investigation. . . . .	129
6.14 <b>Cumulative Distribution of Exclusion Distance.</b> Here we plotted the XTMVA 90% exclusion distance for the 13 GRBs in the results set. This is the distance to which 90% of injections can be recovered with a significance greater than the loudest event in the on-source. . . . .	130
6.15 <b>Generalised Clustering Sensitivity Change.</b> Here we see the change in 50% injection scale upper limit for XTMVA with and without using generalised clustering. A lower value indicates a more sensitive search. The sensitivity of XTMVA is significantly improved for long waveforms such as ADIs, BNS, and NSBH when using generalised clustering. There is, however, a slight reduction in sensitivity to short waveforms. . . . .	133
6.16 <b>Detection Efficiency Curve without Generalised Clustering.</b> This is the detection efficiency curve for an XTMVA analysis without generalised clustering. The x-axis shows the root-sum-square amplitude of the injected waveforms and the y-axis shows the fraction of injections detected. This plot shows that at low amplitude no injections are found, while for very loud injections there is almost a 100% detection efficiency. This is the typical, expected behaviour. . . . .	135
6.17 <b>Detection Efficiency Curve with Generalised Clustering.</b> This is the detection efficiency curve for an XTMVA analysis using generalised clustering. The x-axis shows the root-sum-square amplitude of the injected waveforms and the y-axis shows the fraction of injections detected. We can see that some very loud injections are being missed, despite the fact that close to 100% of some lower energy injection sets are being found. . . . .	136
6.18 <b>Time-Frequency Box Size.</b> Here we have a histogram of the time-frequency box size of triggers in the signal training set for an analysis with generalised clustering and without. We can see that the default run not only has a lot more triggers overall, but it has more triggers with a large time-frequency box as well. . . . .	136

# List of Tables

5.1	<b>Time to Complete Single Matched Filter job.</b> Here we see how long a single matched filter job in the GRB 170817A analysis took to complete with different $\chi^2$ tests being used. The new code does not use the bank or auto $\chi^2$ tests. We can see that there is a substantial speed up of the analysis. . . . .	97
6.1	<b>Example MVA training data.</b> Each event has a label and several attributes. The training sets we actually use have up to 15 attributes and thousands of events. The labels are various coherent and incoherent measures of energy, as well as trigger properties such as the bandwidth and duration. A full list is given in table 6.2.1. . . . .	111
6.2	<b>MVA Attributes.</b> The attributes used by the machine learning classifier. . . . .	112
6.3	<b>Training Waveforms.</b> Here we list the waveforms used by XTMVA for both training and testing, together with a short description. . . . .	119
1	<b>Median 50% Injection Scale Upper Limits.</b> Here we list the median 50% injection scale upper limit values for X-pipeline and the MVA. . . . .	144
2	<b>MVA 50% Injection Scale Upper Limit.</b> Here we list the 50% injection scale upper limit for each waveform type tested and for each GRB analysed. . . . .	145
3	<b>X-pipeline 50% Injection Scale Upper Limit.</b> Here we list the 50% injection scale upper limit for each waveform type tested and for each GRB analysed. . . . .	146

“[T]he Listeners are trying to work out precisely what it was that the Creator said when He made the universe... There were certain problems caused by the fact that they didn’t hear only the subtle echoes of the first words, but every other sound made on the Disc. In order to recognise the sound of the Words, they had to learn to recognise all the other noises. This called for a certain talent, and a novice was only accepted for training if he could distinguish by sound alone, at a distance of a thousand yards, which side a dropped coin landed. He wasn’t actually accepted into the order until he could tell what colour it was.”

*Terry Pratchett*

For my parents, not that you'll understand any of it.



# Collaborative Work

The following sections of this thesis contain work that has been co-authored with others, or that builds upon software developed by others and previously published:

- Chapter 4 discusses the PyGRB search pipeline. Previous publications describing developments to this pipeline and search results can be found in **cite harry macleod williamson and O1 GRB paper**.
- Chapter 4 ends with a discussion of the PyGRB search results from the most recent observing run. These results can be found in **cite O2 GRB paper**. I contributed to this paper compiling the required information on the Gamma-ray bursts (GRBs), analysing eight of the 42 GRBs in the PyGRB analysis sample, and co-authoring the paper.
- Chapter 5 discusses development work on PyGRB. I led this work, but with the input of many collaborators, especially from members of the Cardiff University Gravitational Wave group and the GRB working group of the LIGO-Virgo collaboration.
- Chapter 6 discusses development work on X-pipeline . X-pipeline has been previously written about in **cite sutton and was**.
- The X-pipeline search results from the most recent observing run, used in chapter 6, were previously published in **cite O2 GRB paper**, which I co-authored.
- The MVA search for GWs described in chapter 6 has been previously described in **cite Adams paper** and the software used is described in **cite TMVA user guide**.



# Chapter 1

## Introduction

In 1915, Albert Einstein published a new theory of gravity: the general theory of relativity **cite gr paper**. This new theory was indistinguishable from Newtonian gravity in most situations, but in the presence of strong gravitational fields the two theories made different predictions. For example, Einstein predicted that the sun would gravitationally lens light by twice the amount predicted using Newtonian arguments. An experiment was soon performed by Arthur Eddington and collaborators to measure the amount of gravitational lensing by the sun. This experiment looked for stars that should have been behind the sun during the 1919 solar eclipse **cite eddington**. Some of these stars were made visible, and from these the amount of lensing was measured and confirmed the prediction of general relativity. The results of this experiment and others meant general relativity was quickly accepted by the scientific community. However, there was one prediction of general relativity that remained elusive for decades to come: *gravitational waves* (GWs).

General relativity was a field theory of gravity, and bore a strong resemblance to Maxwell's theory of electromagnetism **cite maxwell paper**. It was reasoned by some that this meant there must be GWs produced by accelerating mass just as there are electromagnetic waves produced by accelerating charge. Others thought this was stretching the analogy to electromagnetism too far, and pointed to key differences between gravitation and electromagnetism **cite kannefeck**. For example, there is no negative mass to act as an analog to the negative charge of electromagnetism.<sup>1</sup> After much controversy, the consensus built that GWs do exist and experiments were conceived that had the potential to detect GWs. After yet more controversy, with the false claim of a detection by Joseph Weber **cite weber paper**, the first experimental evidence for the existence of GWs came in 1975 from observations of a binary star system. According to general relativity, a binary star system should slowly inspiral due to energy lost through GW emission. Hulse and Taylor observed a neutron star binary with precisely the orbital decay predicted by general relativity

---

<sup>1</sup>It can be shown that this is one of the reasons why GWs are much more difficult to detect than electromagnetic waves.

---

### cite hulse taylor.

A few years prior to this, Rainer Weiss set out a new vision for GW astronomy. Weiss proposed building a Michelson interferometer to detect GWs directly **cite weiss paper**. This led to founding of the Laser Interferometer Gravitational-wave Observatory (LIGO) in the early 1980s, which set out to build a pair of 4 km laser interferometers. The new detector network would be able to detect GWs emitted by compact binary systems, such as black hole and neutron star inspirals, and potentially also from nearby supernovas.

Around the same time that GW astronomy was just starting to get off the ground, another kind of transient astronomy was hitting its stride. Gamma-ray Bursts (GRBs) are bright short flashes of gamma rays that originate at extra-galactic distances. First detected in the 1960s, the cause of GRBs remained a mystery for decades. With the 1991 launch of the Burst and Transient Source Experiment (BATSE), a space based GRB detector, there was finally a large sample of GRB light curves. This confirmed that there were two populations of GRBs depending on the duration of the burst, called long and short GRBs. Theoretical work done in the early 1990s predicted that GRBs could be produced by compact objects such as neutron stars and black holes, possibly via supernova that creates the compact object. By the end of the decade, astronomical observations confirmed that long GRBs were linked to supernova explosions. Then, in the early 2000s, observations strongly suggested that short GRBs were produced by binary neutron star mergers.

The fact that GRB and GW astronomy were both now targeting the same extreme events made collaboration natural. There are several strategies used to search for coincident GRB/GW detections. The first approach is to look for coincident detections between GW and GRB triggers. The second approach uses the fact that GW detectors are sensitive to most of the sky by making followup observations of GW triggers with electromagnetic (EM) telescopes. The final method is to use advanced data analysis techniques to followup GRB detections with a GW search, which is the focus of this thesis.

The LIGO detectors first came online in 2002 until 2010, but did not detect any GWs. In 2010 the LIGO detectors were shut down for upgrades. In September 2015 the detectors came back online, and on the 14th of September 2015, the first GW detection was made [5]. The component masses of this binary were  $\sim 30M_{\odot}$ , too large to be neutron stars, and so not a candidate for EM followup observations. Then, on the 17th of August 2017, a short GRB was detected in coincidence with a low mass binary GW signal **cite gw170817 paper**. This was the first joint GW/GRB detection, and confirmed that at least some binary neutron star mergers produced short GRBs. This detection led to discoveries about GRB astrophysics, constrained the rates of binary neutron star mergers, gave a new method to measure the Hubble constant, and many other discoveries.<sup>2</sup>

---

<sup>2</sup>See section 3.2 for more details.

It is in this context that this thesis looks at improved methods for following up GRB observations. We will see two search pipelines. The first is looks for a binary inspiral signal associated with short GRBs, while the second is an unmodelled search, meaning it is sensitive to any waveform morphology, and looks for GWs associated with both long and short GRBs. We will also look at the results of the most recent searches for GWs associated with GRBs. We begin with an introduction to GW astronomy, followed by a more detailed overview of GRB astronomy than was given above.

## Chapter 2

# Introduction to Gravitational Wave Astronomy

In this chapter we will run through the basics of gravitational wave (GW) astronomy. We will start by showing that GWs are a natural consequence of general relativity and show that binary star systems emit GWs. We will then look at GW detectors, focusing on the noise in the detector and how the sensitivity changes over the sky, as these factors are important for data analysis.

## 2.1 Gravitational Waves

General Relativity shows that spacetime can curve and move. From this, it may seem obvious that waves can travel though spacetime, but this is not a trivial fact. In this section, we will show that gravitational waves do exist. We will do this by considering a perturbation of the flat Minkowski metric. We will then calculate the Ricci tensor and Ricci scalar for the perturbed metric and, with a clever choice of gauge transformation, see that these yield a wave solution to Einstein's equations. This section is based mostly on the derivations in [41] and [12].

### 2.1.1 Linearised Gravity

Let  $g_{\mu\nu}$  be the metric given by adding a small perturbation  $h_{\mu\nu}$  to the Minkowski metric  $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$ . We write this metric as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad |h_{\mu\nu}| \ll 1 . \quad (2.1)$$

To first order in  $h$ , we calculate the Ricci tensor (3)

$$R_{\mu\nu} = \eta^{\alpha\beta} R_{\alpha\mu\beta\nu} = \frac{1}{2} (\partial_\mu \partial^\alpha h_{\alpha\nu} + \partial_\nu \partial_\alpha h_\mu^\alpha - \square h_{\mu\nu} - \partial_\mu \partial_\nu h) \quad (2.2)$$

and the Ricci scalar (4)

$$R = \eta^{\mu\nu} R_{\mu\nu} = \partial_\mu \partial_\alpha h^{\mu\alpha} - \square h, \quad (2.3)$$

where  $\square = \partial_\mu \partial^\mu$  is the d'Alembertian operator and  $h = h_\mu^\mu$  is the trace of  $h$ . From these, we find the resulting Einstein tensor is also linear in  $h$

$$\begin{aligned} G_{\mu\nu} &= R_{\mu\nu} - \frac{1}{2} R \eta_{\mu\nu} \\ &= \frac{1}{2} (\partial_\sigma \partial_\nu h_\mu^\sigma + \partial_\sigma \partial_\mu h_\nu^\sigma - \partial_\mu \partial_\nu h - \square h \mu - \eta_{\mu\nu} \partial_\rho \partial_\lambda h^{\rho\lambda} + \eta_{\mu\nu} \square h) / fs \end{aligned} \quad (2.4)$$

### 2.1.2 Gauge Transformation

The metric given in 2.1 does not uniquely define a coordinate system, as there may be other coordinates where the metric is given by the Minkowski metric plus some other perturbation. This gives us the freedom to choose some coordinates that make the linear Einstein tensor simpler. To find these coordinates, we start with the following gauge transformation

$$x^{\mu'} = x^\mu + \chi^\mu(x), \quad \partial_\mu \chi^\nu \ll 1. \quad (2.5)$$

This gives us

$$\frac{\partial x^\mu}{\partial x^{\alpha'}} = \delta_\alpha^\mu - \partial_\alpha \chi^\mu + \mathcal{O}(|\partial \chi|^2). \quad (2.6)$$

Applying these results to our metric (2.1) we find

$$g_{\alpha'\beta'} = \frac{\partial x^\mu}{\partial x^{\alpha'}} \frac{\partial x^\nu}{\partial x^{\beta'}} g_{\mu\nu} = g_{\alpha\beta} - \partial_\alpha \chi_\beta - \partial_\beta \chi_\alpha. \quad (2.7)$$

Subtracting the Minkowski metric from each side, we find

$$h_{\alpha'\beta'} = h_{\alpha\beta} - \partial_\alpha \chi_\beta - \partial_\beta \chi_\alpha. \quad (2.8)$$

We have some freedom in choosing our  $\chi$ , so we impose the *harmonic gauge condition*

$$\partial_\mu h_\nu^\mu = \frac{1}{2} \partial_\nu h, \quad (2.9)$$

where  $h = h_\lambda^\lambda$ . We can always choose a  $\chi$  such that this is true. To see this, first note that  $\partial'_\mu = \partial_\mu - (\partial_\mu \chi^\lambda) \partial_\lambda$ . From this we find

$$(\partial'_\mu h''_\nu^\mu - \frac{1}{2} \partial'_\nu h') \approx (\partial_\mu h_\nu^\mu - \frac{1}{2} \partial_\nu h) - \square \chi_\nu. \quad (2.10)$$

Thus, if we are given an  $h$  such that 2.9 is not true, we can choose a  $\chi$  such that

$$\square \chi_\nu = (\partial_\mu h_\nu^\mu - \frac{1}{2} \partial_\nu h). \quad (2.11)$$

Using 2.9, we can simplify the Ricci tensor and scalar

$$R_{\mu\nu} = -\frac{1}{2}\square h_{\mu\nu} \quad (2.12)$$

$$R = -\frac{1}{2}\square h . \quad (2.13)$$

Using these in the linearised Einstein equations (2.4) gives us

$$\square h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\partial^2 h = -\frac{16\pi G}{c^4}T_{\mu\nu} . \quad (2.14)$$

Alternatively, we can use 6 to write this as

$$\square h_{\mu\nu} = -16\pi G \left( T_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}T \right) . \quad (2.15)$$

In a vacuum, the right hand side of this equation becomes zero, and we can recognise as the relativistic wave equation

$$\square h_{\mu\nu} = 0 . \quad (2.16)$$

### 2.1.3 Physical Effects of Gravitational Waves

The plane wave solution for the vacuum wave equation is

$$h_{\mu\nu}(x) = \epsilon_{\mu\nu}e^{ik_\alpha x^\alpha} \quad (2.17)$$

where  $\epsilon_{\mu\nu}$ , the polarisation tensor for the gravitational wave, is symmetric and constant, and  $k^\alpha$  is the 4-wave-vector given by  $k^\alpha = (\omega, \vec{k})$ . Substituting this into the vacuum wave equation, we find

$$k_\alpha k^\alpha = -\omega^2 + |\vec{k}|^2 = 0 . \quad (2.18)$$

Hence, gravitational waves travel at the speed of light.

The polarisation tensor must satisfy the wave equation and the harmonic gauge condition. Putting the wave solution 2.17 into the harmonic gauge condition 2.9, we find that gravitational waves are transverse

$$k_\mu \epsilon_\nu^\mu = \frac{1}{2}\partial_\nu \epsilon \quad (2.19)$$

where  $\epsilon = \epsilon_\lambda^\lambda$ .

We can impose further gauge conditions as long as the harmonic gauge (and hence 2.19) is not violated. As any transformation with  $\square\chi = 0$  will satisfy 2.11, and hence harmonic gauge condition, we express  $\chi$  as

$$\chi_\nu = X_\nu e^{ikx} . \quad (2.20)$$

Using 2.20 and 2.17 in the transformation equation 2.8, we find the transformation equation for the polarisation tensor

$$\epsilon'_{\mu\nu} = \epsilon_{\mu\nu} - ik_\mu X_\nu - ik_\nu X_\mu . \quad (2.21)$$

Taking the trace of this, we find

$$\epsilon'_\mu^\mu = \epsilon_\mu^\mu - 2ik^\mu X_\mu . \quad (2.22)$$

Thus, we can impose the further gauge condition that the polarisation matrix be traceless by choosing coordinates such that  $2ik^\mu X_0 = \epsilon_\mu^\mu - 2ik^i X_i$ . Using this in 2.19, we find

$$k^\mu \epsilon_{\mu\nu} = 0 . \quad (2.23)$$

We can fix the other elements of  $X_\mu$  by setting  $\epsilon_{i0} = 0$  for  $i = 1, 2, 3$ . We do this using 2.21 to find

$$\epsilon'_{i0} = \epsilon_{i0} - ik_i X_0 - ik_0 X_i . \quad (2.24)$$

Now we see that by choosing the  $X_i$  such that  $ik_0 X_i = ik_i X_0 - \epsilon_{i0}$ , we have  $\epsilon'_{i0} = 0$ . The gauge defined by the above conditions is called the *transverse-traceless* (TT) gauge.

To make things even simpler, assume the wave is traveling in the z-direction, that is  $\vec{k} = (0, 0, \omega)$ . The transverse condition 2.23 gives us  $\epsilon_{0\nu} = \epsilon_{3\nu}$ . This gives us  $\epsilon_{00} = 0$  and  $\epsilon_{3i} = 0$ , as we already have  $\epsilon_{30}$ . The traceless condition then becomes  $\epsilon_{11} + \epsilon_{22} = 0$ . Using the symmetry of the polarisation matrix to fill in the rest of the matrix, we find the following solution to 2.17

$$h_{\mu\nu}(x) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \epsilon_+ & \epsilon_\times & 0 \\ 0 & \epsilon_\times & -\epsilon_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} e^{(i\omega(z-t))} .$$

We can see that the polarisation matrix is fully described by two real numbers  $\epsilon_+$  and  $\epsilon_\times$ . We call these the plus and cross polarisation of the gravitational wave.

To understand what these two polarisations mean physically, consider two particles separated by  $\varphi^\mu = x_1^\mu - x_2^\mu = (0, \vec{\varphi})$ . The proper distance between these two points is given by

$$\Delta L = \sqrt{g_{\mu\nu} dx^\mu dx^\nu} = \sqrt{\vec{\varphi}^2 + h_{ij} \varphi^i \varphi^j} = |\vec{\varphi}| \left( 1 + \frac{1}{2} h_{ij} \varphi^i \varphi^j / \vec{\varphi}^2 + \mathcal{O}(h^2) \right) . \quad (2.25)$$

From this we see that the fractional change in distance between the two particles is

$$\frac{1}{2} h_{ij} \hat{\varphi}^i \hat{\varphi}^j = \frac{1}{2} (\epsilon_+ [(\hat{\varphi}^1)^2 + (\hat{\varphi}^2)^2] + 2\epsilon_\times \hat{\varphi}^1 \hat{\varphi}^2) e^{(i\omega(z-t))} \quad (2.26)$$

where  $\hat{\varphi}^i = \varphi^i / |\varphi|$ . From this formula, it becomes clear (with some thought) that the plus polarisation causes the x-axis and y-axis separation of particles to increase and then decrease  $180^\circ$  out of phase with each other. The cross polarisation has the same effect, but offset by  $45^\circ$ . This is shown more clearly in figure 2.1.

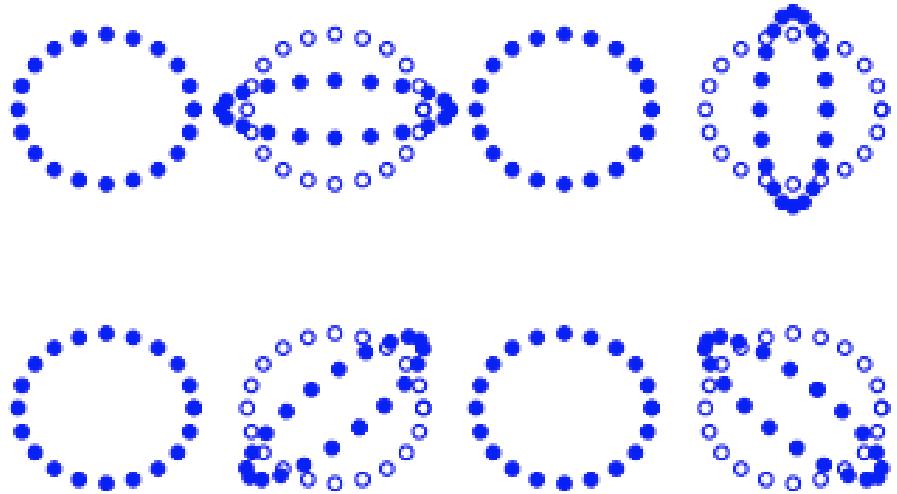


Figure 2.1: **Plus and Cross Polarisation.** The effect of the plus and cross polarisation of a GW on a ring of test particles. [27]

## 2.2 Gravitational Wave Generation

We will now show how GWs are emitted by a source. We start by defining the trace reversed perturbation

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h \quad (2.27)$$

we can simplify the linearised Einstein equation 2.14 to

$$\square\bar{h}_{\mu\nu} = -16\pi GT_{\mu\nu} \quad (2.28)$$

where we have now set  $c = 1$ . We can solve this equation by defining a Green's function such that

$$\square\mathcal{G}(x) = \delta^{(4)}(x) . \quad (2.29)$$

The physical interpretation of this solution is that  $\mathcal{G}(x)$  is a wave originating from a point source at the origin. The solution to the linearised Einstein equations can then be written as

$$\bar{h}_{\mu\nu}(x) = -16\pi G \int d^y y \mathcal{G}(x - y) T_{\mu\nu}(y) . \quad (2.30)$$

This Green's function is the same as used to model electromagnetic waves, which means we can get the form of  $\mathcal{G}$  for free

$$\mathcal{G}(t, \vec{x}) = -\frac{\theta(t)\delta(t-r)}{4\pi r} \quad (2.31)$$

where  $r = |\vec{x}|$  is the distance from the source, and  $\theta(t) = 1$  for  $t > 0$  and is zero otherwise. The interpretation of this solution is fairly simple,  $\theta$  ensures that the wave travels forwards in time,  $\delta(t-r)$  ensures the wave travels at the speed of light, and the final factor  $\frac{-1}{4\pi r}$  ensures that the second derivative of the Green's function is a delta function at the origin as  $\nabla^2(\frac{-1}{4\pi r}) = \delta^{(3)}(\vec{x})$ . Plugging 2.31 into 2.30, we obtain

$$\bar{h}_{\mu\nu}(t, \vec{x}) = 4G \int d^4y \frac{\theta(t-y^0)\delta(t-y^0 - |\vec{x} - \vec{y}|)T_{\mu\nu}(y)}{|\vec{x} - \vec{y}|} . \quad (2.32)$$

Using the delta function to integrate over  $y^0$  we obtain

$$\bar{h}_{\mu\nu}(t, \vec{x}) = 4G \int d^3y \frac{T_{\mu\nu}(t - |\vec{x} - \vec{y}|, \vec{y})}{|\vec{x} - \vec{y}|} . \quad (2.33)$$

Let us consider the case of an isolated source of GWs that is far away, so we can approximate  $|\vec{x} - \vec{y}| = r$  with  $r$  the distance to the source. Then we have

$$\bar{h}_{\mu\nu}(t, \vec{x}) = \frac{4G}{r} \int d^3y T_{\mu\nu}(t-r, \vec{y}) . \quad (2.34)$$

This formula is not very helpful for determining what a GW signal will look like, as it is difficult to interpret a stress-energy tensor. To make things simpler, we will now show how to write this solution in terms of just the energy density term of the stress energy tensor  $T_{00}$ , which is easy to understand. To do this, first note that the harmonic gauge condition gives us  $\partial^\mu \bar{h}_{\mu\nu} = 0$ , from which we get

$$\partial_0 \bar{h}^{0\nu} = -\partial_i \bar{h}^{0i} \quad (2.35)$$

which can be used to determine the timelike components of  $\bar{h}$  from the spacelike components. We therefore restrict our attention to finding  $\bar{h}_{ij}$ . Next, we insert a term  $\partial_k y^i = \delta_k^i$  into the above integrand, and then use integration by parts to obtain

$$\int d^3y T_{\mu\nu} = \int \partial_k(y^i T^{kj}) d^3y - \int y^i (\partial_k T^{kj}) d^3y . \quad (2.36)$$

The first term vanishes as it is a surface term and we have an isolated source. For the second term, we start by using the conservation law  $\partial_\mu T^{\mu\nu}$  to obtain

$$\partial_0 T^{0\nu} = -\partial_k T^{k\nu} \quad (2.37)$$

with  $\nu = j$  to obtain

$$\int d^3y T_{\mu\nu} = \int y^i \partial_0 T^{0j} d^3y . \quad (2.38)$$

The next step<sup>1</sup> is to use the symmetry of the stress-energy tensor to write

$$\int d^3y T_{\mu\nu} = \frac{1}{2} \int y^i \partial_0 T^{0j} + y^j \partial_0 T^{0i} d^3y . \quad (2.39)$$

We can then repeat the trick used to obtain 2.36, i.e. multiply both terms in the integrand by  $\partial_k y^i = \delta_k^i$  and use integration by parts. After some algebra, everything should cancel nicely to leave us with

$$\int d^3y T_{\mu\nu} = \frac{1}{2} \int \partial_k (y^i y^j \partial_0 T^{0k}) - y^i y^j \partial_0 \partial_k T^{0k} d^3y . \quad (2.40)$$

The first term vanishes as it is a surface term. The second term can again be simplified using 2.38 to give

$$\int d^3y T_{\mu\nu} = \frac{1}{2} \int y^i y^j (\partial_0)^2 T^{00} d^3y \quad (2.41)$$

which is purely in terms of the energy-density term of the stress-energy tensor, as we set out to show. Using this result in equation 2.34 gives us the formula for a GW signal

$$\bar{h} = \frac{2G}{r} \int y^i y^j (\partial_0)^2 T^{00} d^3y . \quad (2.42)$$

We then define the *quadrupole moment tensor* as

$$I^{ij}(t) = \int y^i y^j (\partial_0)^2 T^{00} d^3y \quad (2.43)$$

to obtain

$$\bar{h}(t, \vec{x}) = \frac{2G}{r} \frac{d^2 I^{ij}}{dt^2}(t_r) \quad (2.44)$$

where  $t_r = t - r$  is the retarded time.

## 2.3 Gravitational Wave Sources

We will now see how to use the quadrupole formula to calculate the morphology of a GW produced by a binary star system. We will assume the stars are orbiting in the  $x^1 - x^2$  plane at a distance  $R$  from the center of mass of the system, and that both stars are of mass  $M$ . To simplify our calculations, we will use the Newtonian approximation of the orbits. In this case we have

$$\frac{GM}{(2R^2)} = \frac{Mv^2}{R} \quad (2.45)$$

---

<sup>1</sup>This step seems arbitrary now but is necessary.

from which we find

$$v = \left( \frac{GM}{4R} \right)^{\frac{1}{2}} . \quad (2.46)$$

The orbital period of this system is

$$T = \frac{2\pi R}{v} \quad (2.47)$$

which gives us the angular frequency

$$\Omega = \frac{2\pi}{T} = \left( \frac{GM}{4R^3} \right)^{\frac{1}{2}} . \quad (2.48)$$

The path of the stars, which we label  $a$  and  $b$ , can then be written

$$x_a^1 = -x_b^1 = -R \cos \Omega t , \quad x_a^2 = -x_b^2 = -R \sin \Omega t . \quad (2.49)$$

The energy density of the system is then given by

$$T^{00}(t, \vec{x}) = M\delta(x^3)[\delta(x^1 - R \cos \Omega t)\delta(x^2 - R \sin \Omega t) + \delta(x^1 + R \cos \Omega t)\delta(x^2 + R \sin \Omega t)] . \quad (2.50)$$

Using this in the quadrupole moment tensor formula 2.43, we find

$$I_{11} = 2MR^2 \cos^2 \Omega t = MR^2(1 + \cos 2\Omega t) \quad (2.51)$$

$$I_{22} = 2MR^2 \sin^2 \Omega t = MR^2(1 - \cos 2\Omega t) \quad (2.52)$$

$$I_{12} = I_{21} = 2MR^2(\cos \Omega t)(\sin \Omega t) = MR^2(\sin 2\Omega t) \quad (2.53)$$

$$I_{33} = 0 . \quad (2.54)$$

From these, the metric perturbation is easily constructed from 2.44

$$\bar{h}_{ij}(x) = \frac{8GM}{r}\Omega^2 R^2 \begin{pmatrix} -\cos 2\Omega t_r & -\sin 2\Omega t_r & 0 \\ -\sin 2\Omega t_r & \cos 2\Omega t_r & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

From this we can see that GWs from a binary star system would produce a GW with an oscillating plus and cross polarisation. We should note that this derivation assumes no energy loss from the GWs. In fact, GWs do radiate energy and this causes the stars to inspiral, a feature not captured by our Newtonian approximation. As the stars inspiral, the angular frequency of the system increases. This causes the frequency of the GW to increase as well as the amplitude<sup>2</sup>. This is what creates the characteristic *chirp* of the GW from a binary inspiral and merger. Even without

---

<sup>2</sup>Though the inspiral also causes the separation between the black holes to shrink, which acts to lower the amplitude. In reality, the increase in orbital frequency wins out and the amplitude increases near the end of the inspiral.

taking into account the inspiral, this simple model illustrates many properties that are useful for GW astronomy. First is the  $1/r$  dependence on the size of the metric perturbation, which can be used to estimate the distance to the source. Another is the fact that the GW frequency is twice the orbital frequency of the system, so the measured GW frequency gives us the orbital frequency of the system. Also, the metric perturbation is larger for systems with high mass and high orbital frequency. High mass stars are therefore good candidates, but they also have a large radius, meaning that the stars in the binary will begin to merge while the binary still has a low angular frequency. Thus the best GW emitters are compact object binaries, such as neutron stars and black holes, as they are high mass but also have a small radius. We call these systems *compact binary coalescence* (CBC) systems.

On the 14th of September 2015, the first direct detection of a GW came from a CBC system [5]. In figure 2.2] we see the waveform that was detected, with a cartoon of the evolution of the inspiral. The bottom panel also shows how the velocity and separation of the black holes evolves. The signal is just as described above, an sine wave with increasing frequency and amplitude as the black holes radiate energy through GW emission and inspiral. Since this initial detection, many CBC signals have already been detected [[cite catalog paper](#)], mostly from binary black hole systems, but also one binary neutron star system, which will be discussed in detail in the next chapter. In chapter 4 we will see a search pipeline designed specifically to detect GWs from these CBC systems.

We end this section by noting that any system with a large amount of asymmetric rapidly accelerating mass could potentially be detected using GWs. There are searches ongoing for GWs from supernovas, gamma-ray bursts, non-axisymmetric pulsars, and black hole accretion disk instabilities. At the time of writing, the only GWs that have been detected are from CBC systems.

## 2.4 Gravitational Wave Detectors

As shown in section 2.1.3, GWs can be detected by measuring changes in length between two points, and this change in length is proportional to the distance between the two points and the amplitude of the GW. As astrophysical sources are expected to produce GWs with amplitudes of  $\sim 10^{-21}$ , we must choose two points that are as far apart as practical, at least on the order of kilometers, if the effect is to be measurable. Even then, the change in distance between the two points will still be only a fraction of the wavelength of light. To measure this change, we use a *Michelson interferometer* [[cite R.Weiss 1972](#)]. In this section, we will begin with a basic mathematical description of how a Michelson interferometer works. We will then move on to discuss the aspects of the GW detectors that are important for data analysis, namely the detector noise and directional sensitivity.

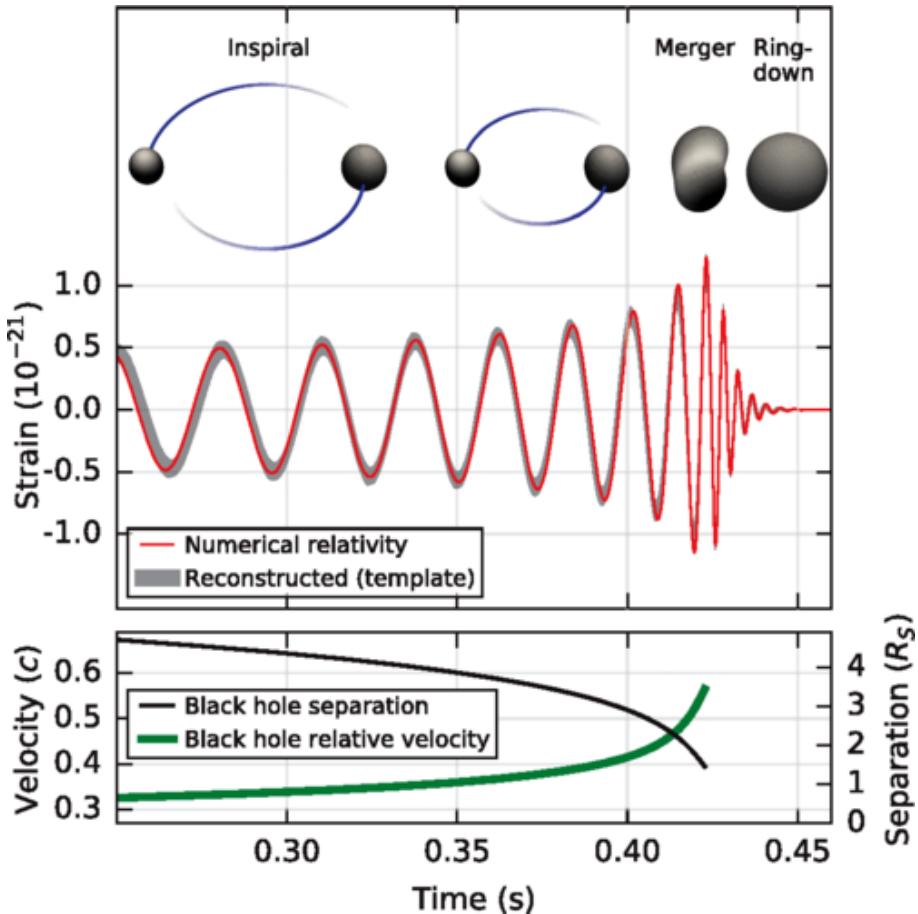


Figure 2.2: **The First Direct Detection of a Gravitational Wave.** The top panel shows the GW waveform with an inset cartoon of the state of the binary at four different phases of the binary inspiral. The bottom panel shows the Keplerian separation distance in units of Schwarzschild radii ( $R_s = \frac{2GM}{c^2}$  where  $M$  is the total mass of the system). Also shown is the relative velocity of the black holes. [5]

#### 2.4.1 The Michelson Interferometer

We will start with an overview of how the interferometers work before going into the mathematical details. Figure 2.3 is a schematic of a Michelson interferometer<sup>3</sup>. There is a laser in the West port which is aimed at the beam splitter in the middle. The beam splitter is a partially reflective mirror which reflects half the photons into the North port and transmits half into the East port. Both interferometer arms are the same length<sup>4</sup>. The light that is reflected into the North arm picks up a phase shift such that it is completely out of phase with the light transmitted into the East port. At the end of the North and East ports are mirrors that reflect the light back to the beam splitter. The light is then transmitted/reflected by the beam

<sup>3</sup>The left and right schematics are the same but each use a different naming convention. Both conventions are useful.

<sup>4</sup>Actually the important thing is that the two light rays have opposite phase in each arm when they return to the beam splitter.

splitter into either the South or West arms. If the North and East arms have the same length, then the light that from each arm will be in phase in the West port<sup>5</sup> but out of phase in the South port<sup>6</sup>. This means that if the arms are the same length then all the light goes into the West port. However, if the arm lengths differ by some small fraction of the wavelength of the laser light (possibly because of a gravitational wave), then some of the light will travel down the South port and be detected. This is the basic idea behind an interferometer.

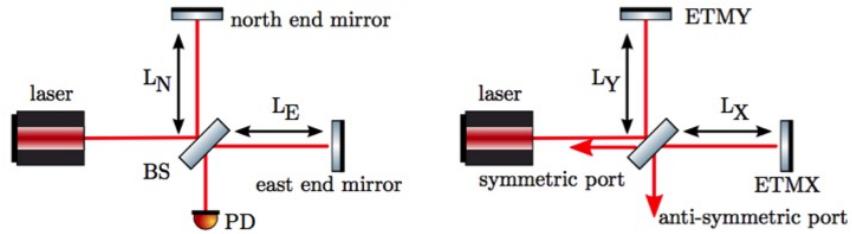


Figure 2.3: Two identical schematics of a Michelson interferometer, each with a different naming convention. [10]

The key feature of an interferometer is that the laser light is out of phase at the dark port. How does this happen? Suppose that the laser light is initially described by the electric field  $E_0$ . Each time the light encounters the beam splitter or a mirror, it will pick up a change in phase. In figure 2.4 we show the naming convention we will use for the different electric fields in the interferometer. The *magnitude of reflection*  $r$  and the *magnitude of transmission*  $t$  describe the reflectivity of a mirror. We will assume the beam splitter and mirrors are lossless, so  $t^2 = 1 - r^2$ . Once the light reaches the beam splitter it will either be reflected into the North arm and pick up a phase shift  $\varphi_r$ , or it will be transmitted into the East arm, in which case it is phase shifted by  $\varphi_t$ . These electric fields are respectively denoted by

$$E_1 = rE_0 \exp(i\varphi_r) \quad (2.55)$$

$$E_2 = tE_0 \exp(i\varphi_t) . \quad (2.56)$$

We denote the phase accumulated in the vertical and horizontal arms as  $\Phi_1$  and  $\Phi_2$  respectively. Thus the electric fields  $E_3$  and  $E_4$  are given by

$$E_3 = rE_0 \exp(i(\varphi_r + \Phi_1)) \quad (2.57)$$

$$E_4 = tE_0 \exp(i(\varphi_t + \Phi_2)) . \quad (2.58)$$

After going through the beam splitter one more time the light rays pick up another phase shift. We will not assume that the phase shift is the same on the front and

<sup>5</sup>Also known as the *Symmetric port*

<sup>6</sup>Also known as the *anti-symmetric port* or the *dark port*.

back of the mirror, so this time the light shifts by  $\varphi_{r'}$ . Hence the two light rays  $E_5$  and  $E_6$  are given by

$$E_5 = E_0 \left( r^2 \exp(i(2\varphi_r + \Phi_1)) + t^2 \exp(i(2\varphi_t + \Phi_2)) \right) \quad (2.59)$$

$$E_6 = rtE_0 \left( \exp(i(\varphi_r + \varphi_t + \Phi_1)) + \exp(i(\varphi_{r'} + \varphi_t + \Phi_2)) \right). \quad (2.60)$$

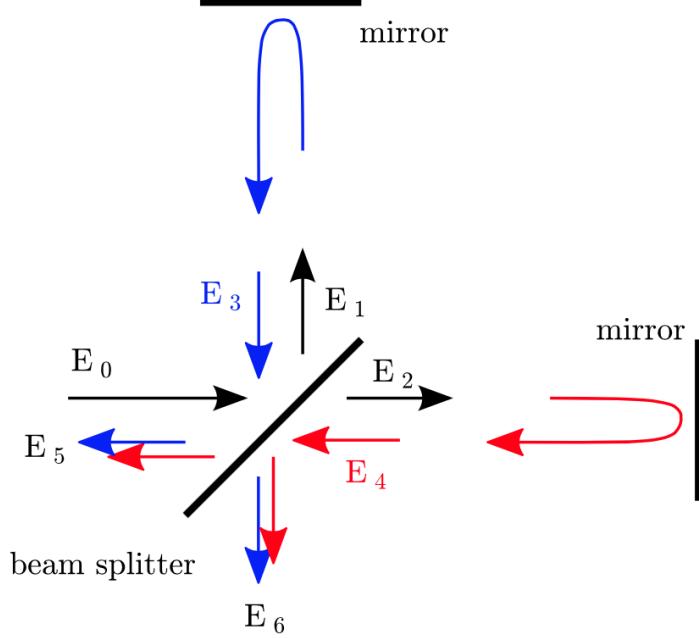


Figure 2.4: This schematic Michelson Interferometer shows our labelling convention for the different Electric fields. [10]

These equations give us the phase at every step of the lights path back out of the detector, but we can simplify things. First we define

$$\alpha_+ = \varphi_r + \varphi_t + \frac{1}{2}(\Phi_1 + \Phi_2) \quad (2.61)$$

$$\alpha_- = \varphi_r - \varphi_t + \frac{1}{2}(\Phi_1 - \Phi_2) \quad (2.62)$$

which gives us

$$E_5 = E_0 e^{i\alpha_+} (r^2 e^{i\alpha_-} + t^2 e^{-i\alpha_-}). \quad (2.63)$$

Similarly, if we define

$$\beta_+ = \varphi_t + \frac{1}{2}(\varphi_r + \varphi_{r'} + \Phi_1 + \Phi_2) \quad (2.64)$$

$$\beta_- = \frac{1}{2}(\varphi_r - \varphi_{r'} + \Phi_1 - \Phi_2) \quad (2.65)$$

which gives us

$$E_6 = rt E_0 e^{i\beta_+} (e^{i\beta_-} + e^{-i\beta_-}) . \quad (2.66)$$

For the case of a 50:50 beam splitter, we have  $r = t = 1/\sqrt{2}$ . Thus we can write

$$E_5 = E_0 e^{i\alpha_+} \cos(\alpha_-) \quad (2.67)$$

$$E_6 = E_0 e^{i\beta_+} \cos(\beta_-) . \quad (2.68)$$

Conservation of energy means we must have  $|E_0|^2 = |E_5|^2 + |E_6|^2$ , hence

$$\cos^2(\alpha_-) + \cos^2(\beta_-) = 1 . \quad (2.69)$$

Using the identity  $\cos^2(x) = \cos(2x) + 1$  we can write this condition as

$$\cos(2\alpha_-) = \cos(2\beta_-) . \quad (2.70)$$

Hence we have  $2\alpha_- = 2\beta_- + \pi(2N + 1)$ , where  $B \in \mathbb{Z}$ , which we can rearrange to obtain

$$\frac{1}{2}(\varphi_r + \varphi_{r'}) - \varphi_t = (2N + 1)\frac{\pi}{2} . \quad (2.71)$$

In most cases we do not know or need to know what the exact phase factors are and can choose any such that 2.71 is satisfied. We will use  $\varphi_r = \varphi_{r'} = 0$  and  $\varphi_t = \pi/2$ , so that the electric field in the south port, given by 2.60, becomes

$$E_6 = E_0 \frac{i}{2} (e^{i\Phi_1} + e^{i\Phi_2}) = E_0 \frac{i}{2} (e^{2ikL_N} + e^{2ikL_E}) \quad (2.72)$$

where  $L_N$  and  $L_E$  are the lengths of the North and East interferometer arms,  $k$  is the wavenumber of the laser light, and we have used the fact that the phase shift obtained by the light traveling down the interferometer arms and back again is  $\Phi_1 = 2kL_N$  and  $\Phi_2 = 2kL_E$ . We can simplify these expressions by writing them in terms of the average and differential arm lengths

$$\bar{L} = \frac{L_N + L_E}{2}, \quad \Delta L = L_N - L_E . \quad (2.73)$$

From this we obtain

$$2L_N = 2\bar{L} + \Delta L, \quad 2L_E = 2\bar{L} - \Delta L . \quad (2.74)$$

Thus, the electric field in the South port is given by

$$E_6 = E_0 \frac{i}{2} e^{2ik\bar{L}} (e^{ik\Delta L} + e^{-ik\Delta L}) = iE_0 e^{2ik\bar{L}} \cos(ik\Delta L) . \quad (2.75)$$

The photodetector in the South port produces a signal proportional to  $E_6 E_6^*$ , which

gives

$$S = E_6 E_6^* = P_0 \cos^2(k\Delta L) = P_0 \cos^2(2\pi\Delta L/\lambda) \quad (2.76)$$

where  $P_0 = E_0 E_0^*$ . In figure 2.5 we see how the output power varies with a change in the differential arm length. We can see that the output signal varies from zero up to the input power with a period of  $\Delta L/\lambda = 0.5$ . At the point  $\Delta L/\lambda = 0.25$ , there is no light entering the south port. We call this point the *Dark Fringe*. As long as  $\Delta L/\lambda < 0.25$ , i.e. the displacement caused by a GW is sufficiently small, then the amount of power at the dark port is proportional to the differential arm length, giving us a way to detect a passing GW.

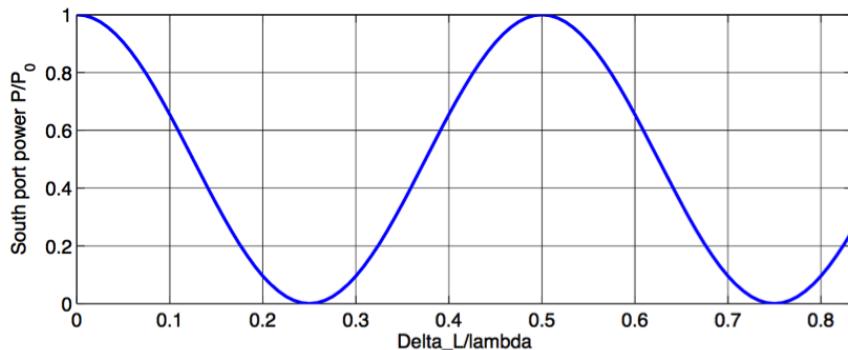


Figure 2.5: This plot shows how the detector output varies with the change in length of the interferometer arms. [10]

### 2.4.2 Interferometer Antenna Pattern

The response of an interferometer to a passing GW depends on the direction that the GW is approaching from. In this section we will see how interferometer sensitivity varies over the sky. The interferometer output depends on the phase shift induced by a change in the differential arm length. If we have an interferometer with arms of length  $L$  and aligned with the x and y axes, then equation 2.25 gives us

$$\Delta L_x = L(1 + \frac{1}{2}h_{xx})) + \mathcal{O}(h^2) \quad (2.77)$$

to be the change in length of the x arm of interferometer arm induced by a passing GW. Similarly, the change in length of the y-arm is given by

$$\Delta L_y = L(1 + \frac{1}{2}h_{yy})) + \mathcal{O}(h^2) . \quad (2.78)$$

Dropping the higher order terms, the differential arm length is then given by

$$\Delta L_x - \Delta L_y = \frac{1}{2}L(h_{xx} - h_{yy}) . \quad (2.79)$$

So far we have been working in the detector frame. To understand how the interferometer responds to GWs coming from different positions in the sky, we need to understand how  $h_{xx}$  and  $h_{yy}$  change for GWs coming from different sky positions. To do this, we consider a GW approaching the detector from an arbitrary direction and introduce a new frame of reference  $(x', y', z')$  such that the incoming GW is traveling along the  $z'$  axis. We relate the detector frame  $(x, y, z)$  to the GW frame using the polar angles  $\theta$  and  $\phi$ . The GW polarisation matrix is defined in the GW frame to be

$$h'_{ij} = \begin{pmatrix} h_+ & h_\times & 0 \\ h_\times & -h_+ & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

The rotation that brings the GW frame into the detector frame is rotation through  $\theta$  around the y-axis and a rotation through  $\phi$  around the z-axis. This gives us

$$\mathcal{R} = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} .$$

We can then calculate  $h_{xx}$  and  $h_{yy}$  using the formula

$$h_{ij} = \mathcal{R}_{ik}\mathcal{R}_{jl}h'_{kl} \quad (2.80)$$

to find

$$h_{xx} = h_+(\cos^2 \theta \cos^2 \phi - \sin^2 \phi) + 2h_\times \cos \theta \sin \phi \cos \phi \quad (2.81)$$

$$h_{yy} = h_+(\cos^2 \theta \sin^2 \phi - \cos^2 \phi) - 2h_\times \cos \theta \sin \phi \cos \phi . \quad (2.82)$$

Plugging these values into 2.79, and setting  $L = 1$  for convenience, gives us

$$\frac{1}{2}(h_{xx} - h_{yy}) = \frac{1}{2}h_+(1 + \cos^2 \theta) \cos 2\phi + h_\times \cos \theta \sin 2\phi . \quad (2.83)$$

Thus, the detector response to a GW  $s$  can be written as

$$s = F_+(\theta, \phi)h_+ + F_\times(\theta, \phi)h_\times \quad (2.84)$$

where  $F_{+, \times}$  are the antenna response to the plus and cross polarisations of the GW, given by

$$F_+(\theta, \phi) = \frac{1}{2}(1 + \cos^2 \theta) \cos 2\phi \quad (2.85)$$

$$F_\times(\theta, \phi) = \cos \theta \sin 2\phi . \quad (2.86)$$

From these we see that the sensitivity of an interferometer varies depending on the sky position of the source of the GW. In particular, we can see that interferometers

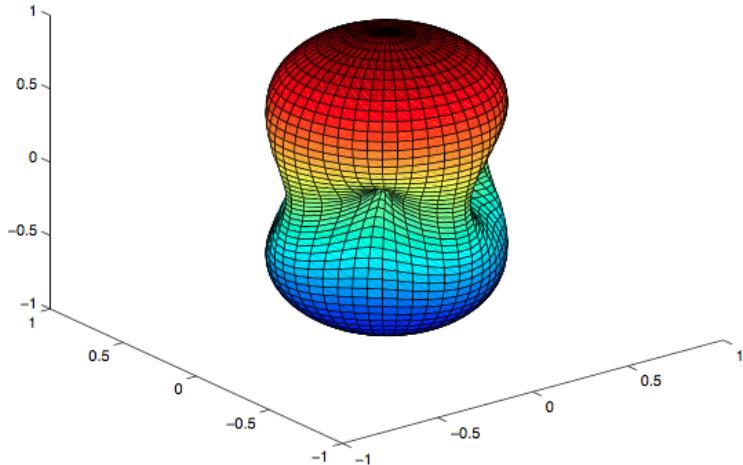


Figure 2.6: **Interferometer Antenna Pattern.** Here we see the antenna response on an interferometer with arms aligned with the x and y axes. The detector is most sensitive to GWs coming from perpendicular to the detector plane, while GWs approaching from within the detector plane but at an angle off-set from the arms by  $\pi/4$  radians will be in a null of the detector. [33]

have blind spots, known as the *null* of the detector. For example, the plus polarisation of a GW coming from an angle  $\phi = \frac{\pi}{4}$  will not be detected. In figure 2.6, we can see how the sensitivity of an interferometer with arms aligned with the x and y axes varies over the sky. In later chapters, we will see how the variation in sensitivity of each detector to different parts of the sky can be used to our advantage, and make a more sensitive GW search.

### 2.4.3 Interferometer Noise

The stationary noise in the detector is measured using the *power spectral density* (PSD). Let  $s = s(\tau)$  be the detector output, the PSD is defined as the Fourier transform of the autocorrelation function of  $s$

$$\frac{1}{2}S_n(f) = \int_{-\infty}^{\infty} s * s(\tau) e^{i2\pi f \tau} d\tau . \quad (2.87)$$

The factor of  $\frac{1}{2}$  is a convention. Taking the square root of the PSD gives us the *Amplitude Spectral Density* (ASD). The PSD and ASD are measures of the amount of time variation of a given frequency that occurs in the detector output. The PSD is measured in units of  $\text{Hz}^{-1}$ , which can be interpreted as the amount of variation in each frequency bin. Similarly, the units of the ASD are  $\text{Hz}^{-\frac{1}{2}}$ . Frequencies at which the ASD is small indicate frequencies where the detector is relatively quiet, while large values indicate a lot of motion in the detector. This can be seen in figure

2.7, which shows the ASD for the LIGO Livingston observatory<sup>7</sup> on 20th August 2017. The ASD is lowest from around 30 Hz to 1000 Hz, and so the detector is most sensitive to GW signals in this frequency range. In figure 2.8 we plot all the known sources of noise in the Livingston detector. From this plot you can see that below the sensitive frequency range of the detector, the noise rapidly rises due to low frequency seismic noise, the alignment and angular sensing and control systems (ASC), and the thermal noise of the mirror suspensions (thermal Susp) and coating (thermal coating). Above the sensitive frequency range it is quantum noise that dominates, where quantum noise is given by the sum of radiation pressure and shot noise. You can also see that at around 20-50 Hz, the sum of noises is much lower than the measured differential arm length (DARM). This indicated that the limiting factor of the detector sensitivity in this frequency range is unknown. The spikes in the ASD have multiple sources, including mirror suspension, calibration lines, and interference due to mains electricity. For a more detailed discussion of LIGO detector noise, see [26] [3].

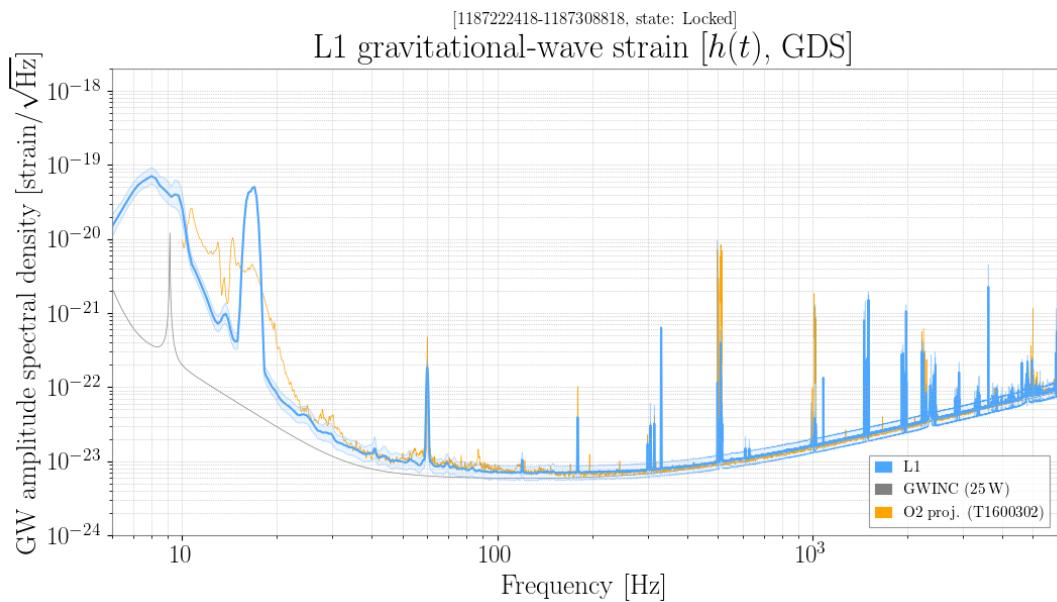


Figure 2.7: **ASD for LIGO Livingston Observatory.** Plotted in blue is the ASD for the LIGO Livingston Observatory as it was on the 20th August 2017. In grey we see the Gravitational Wave Interferometer Noise Curve (GWINC), a theoretical model of all the noise in the detector, as well as the O2 projected noise curve in orange.

<sup>7</sup>A current generation interferometer, see 2.4.4 for more details.

[1187222418-1187308818, state: Ready]

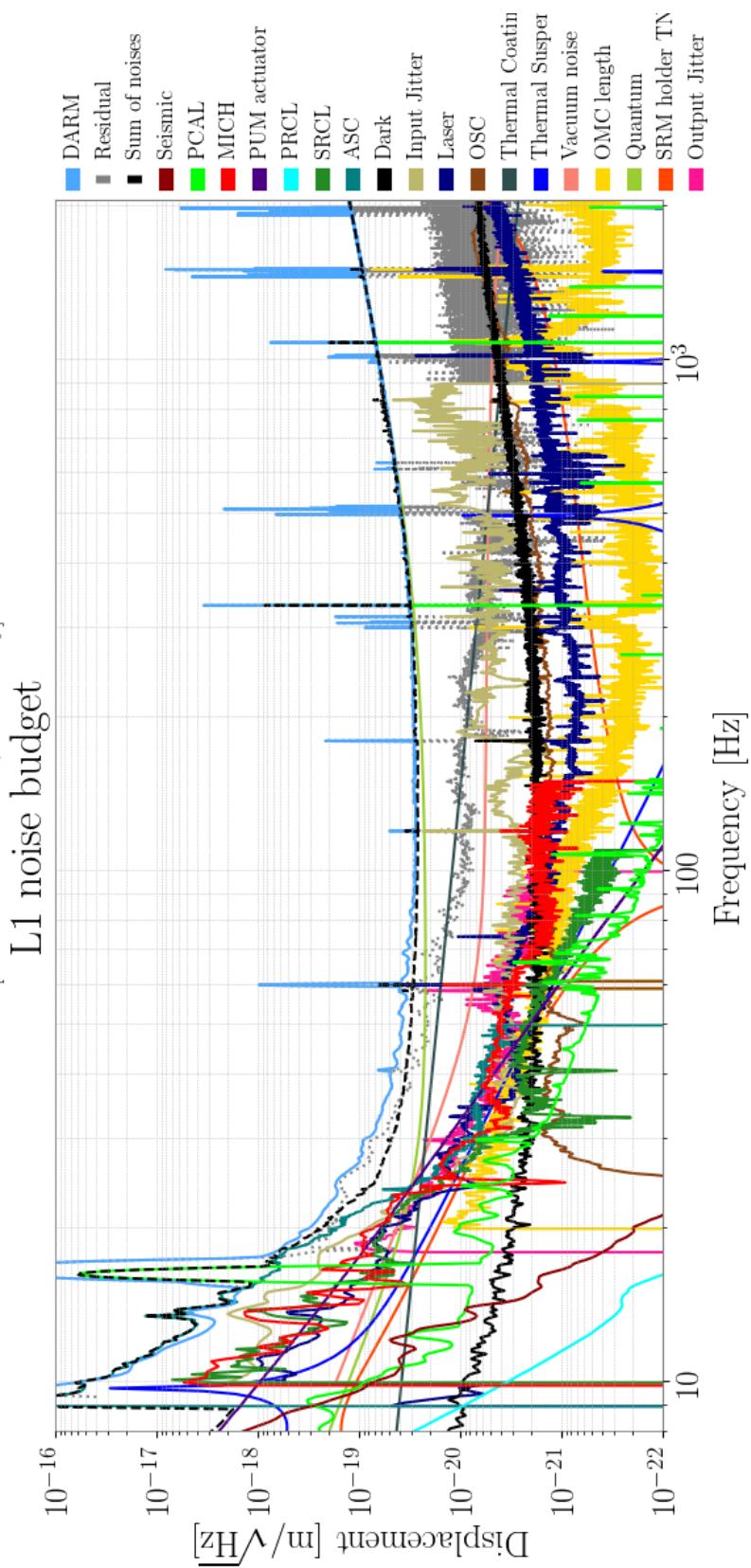


Figure 2.8: **Noise Budget for LIGO Livingston Observatory.** Here we have the noise budget for the LIGO Livingston Observatory on 20th August 2017. The noise sources in LIGO detectors are described in detail in [26] and [3].

Interferometer noise is not stationary. Occasionally there will be noise transients that appear in the data. These can be caused by many different factors. Common causes of glitches include acoustic noise from the environment or saturation overflows, where too much light shines on a photodetector. Stationary noise is relatively easy for GW searches to handle, but glitches require more care as they can easily be mistaken for a GW signal, as we will see in later chapters.

#### 2.4.4 The Global Network of Interferometers

In this section we will quickly summarise the current generation of detectors, including those that should be part of the global network in the near future.

**LIGO** The Laser Interferometer Gravitational-Wave Observatory (LIGO), currently has two detectors, one at Hanford, Washington, and the other at Livingston, Louisiana. The light travel time between the detectors is about 10 ms, reducing the chance of correlated noise and allowing for some triangulation to determine the sky position of a source. The LIGO detectors have 4 km long arms. The arms contain Fabry-Perot cavities, mirrors at each end of the arms that reflect the light back and forth many times. This effectively increases the length of the arms to  $L_{\text{eff}} \sim 1120$  km. Another LIGO detector is being constructed in India, which should be operational by 2024. This detector will be identical to the other two detectors, but by being built far from the other detectors in the network, will significantly increase the ability of the network to localise GW sources using triangulation.

**Virgo** Virgo is a 3 km interferometer, similar but not identical to the LIGO detectors. It is based in Cascina, Italy. Most of the results mentioned in this thesis used the LIGO-Virgo network.

**GEO600** A 600 m interferometer near Starstedt in Germany. Due to the short detector arms compared to the LIGO and Virgo interferometers, GEO600 is typically not used when analysing GW network data, but has been vital for the development of new technology for the LIGO and Virgo detectors.

**Kagra** The Kamioka Gravitational Wave Detector (KAGRA) is a Japanese interferometer being built entirely underground, with 3 km long arms, and will have cryogenic mirrors. It is hoped that KAGRA will be operational by the end of 2019.

# Chapter 3

## Gamma-Ray Bursts

*Gamma-ray Bursts* (GRBs) are exceptionally energetic flashes of gamma rays. They can last from just a few milliseconds up to several hours and have highly variable luminosity curves (see figure 3.1). They are detected at a rate of about one per day, are uniformly distributed over the sky, and are the most electromagnetically energetic objects in the universe. The short duration and huge energy emitted by GRBs suggests a violent origin, making them strong candidates for gravitational wave (GW) emission.

In this chapter we discuss GRB astrophysics. We will begin with some historical perspective, starting with the accidental first detection of a GRB in 1963 and continuing to the first GW detection associated with a GRB in 2017. We will see that GRBs can be classified as short-hard or long-soft, depending on their duration and spectral hardness. We will see the evidence that long GRBs are caused by core collapse supernova, and that short GRBs are caused by compact binary mergers involving at least one neutron star. We will discuss the physical processes that could be powering GRBs and what GW astronomy can teach us about these processes. We end this chapter with a discussion of current GRB detectors and search strategies for GW emission associated with GRBs.

### 3.1 The History of Gamma-ray Burst Astronomy

In this section we will discuss the key discoveries of GRB astronomy in their historical context. These discoveries are the motivation of GW searches associated with GRBs, and were used to develop the GW searches discussed in chapters 4 and 6.

#### 3.1.1 Cold War Tension and an Unexpected Discovery

The partial nuclear test ban treaty, agreed between the USA and the Soviet Union in 1963, banned atmospheric, underwater, and outer space nuclear weapon tests. This created a technical challenge: how to enforce the ban? Seismic sensors could be used for on-Earth tests, but would not work for outer space tests. The solution was to look

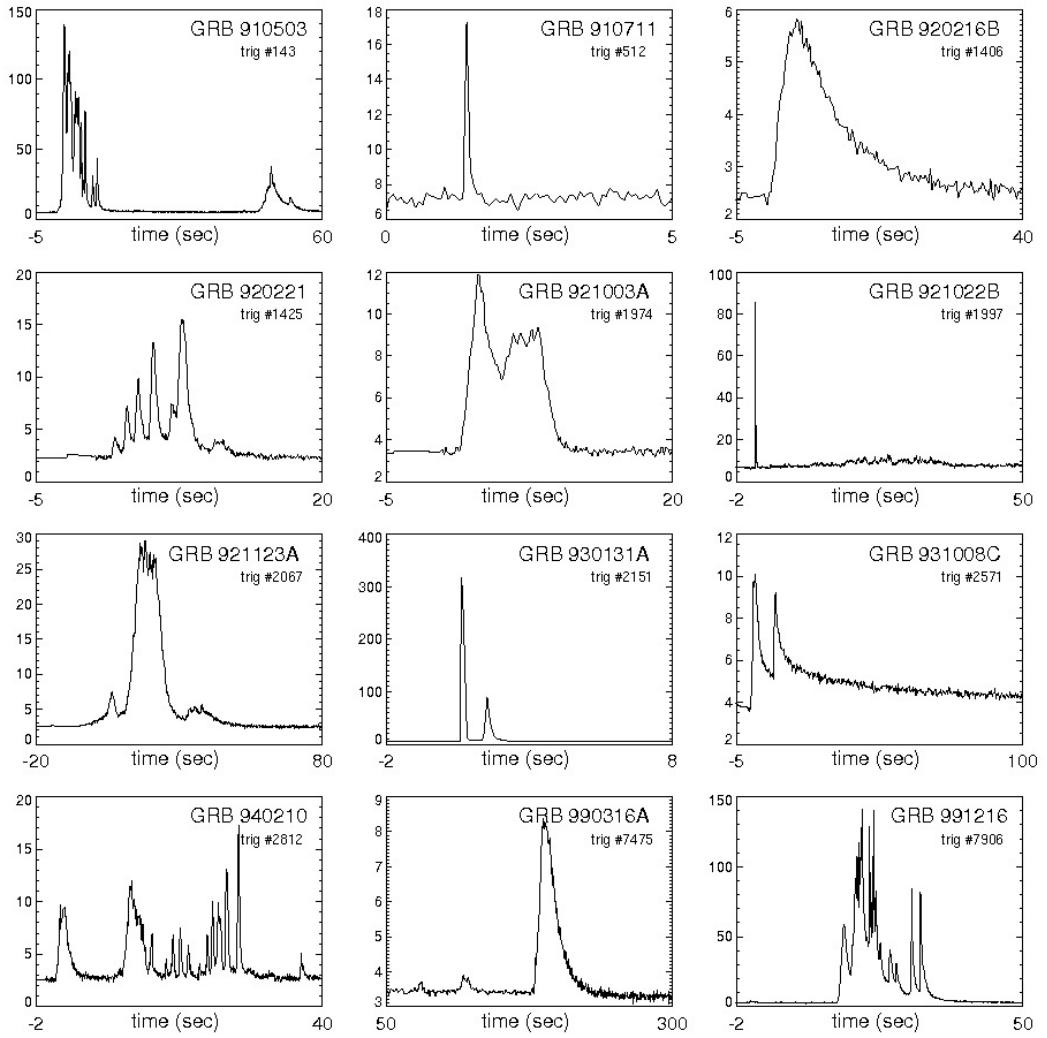


Figure 3.1: **BATSE gamma-ray light curves.** [31]

for the flash of gamma-rays produced in the first milliseconds of a nuclear explosion. Thus the Vela and Kosmos gamma ray detecting satellites were produced by the USA and Soviet Union respectively. These satellites contained only rudimentary gamma ray detectors, and each individual satellite was not capable of localisation. Some localisation was possible using time delay and Earth blocking information.

These satellites started to detect brief bursts of gamma rays, which were first reported in 1973. These events did not look like those expected from a nuclear test, and did not seem to be coming from the Earth or any nearby astronomical bodies such as the moon. It appeared a new, high-energy astronomical phenomenon had been discovered.

This phenomena, called *Gamma-ray Bursts* (GRBs), would appear and fade away in as little as a milliseconds, and could be as brighter than the rest of the gamma ray sky combined. The brevity of these events placed constraints on the size of the source, as the crossing time for a region cannot be less than the light travel time. Thus a 1ms GRB must have a source smaller than 300km across. This limits the potential candidates down to compact objects, such as neutron stars and black holes, or to small regions of larger objects, such as the cores of massive stars. Another important feature of GRBs is that the bursts do not repeat. This suggests a source that is destroyed when the GRB is produced. More measurements were needed to narrow down the number of possible sources for GRBs.

### 3.1.2 BATSE and the Galactic/Extra-galactic Controversy

Most early GRB detectors could not localise well. It was known that GRBs were not coming from any planets in the solar system or from the galactic center, but otherwise the location of GRB sources was a mystery. In particular, it was not clear whether GRBs were coming from galactic or extra-galactic sources. Answering this question was an important step towards identifying the origin of GRBs, as an extra-galactic source would require far greater energy than a galactic source to produce a GRB of equal brightness.

Better sky localisation would help to answer this question. If GRBs mostly occur on the galactic plane, then they are probably of galactic origin. If they were clustered around nearby galaxies, then GRBs probably come from those galaxies. If, however, they are distributed isotropically on the sky, then it would be likely that GRBs form at cosmological distances.<sup>1</sup>

Improved sky localisation was achieved in 1991, with the launch of the Burst and Transient Source Experiment (BATSE) on the Compton Gamma-ray Observatory. With BATSE, it became possible to determine the sky position of a GRB to within

---

<sup>1</sup>It should be noted that there are other ways that GRBs could be isotropically distributed. If they are only detectable to a few hundred parsecs, they would be entirely within the disc of the galactic plane and so would also be isotropically distributed. Alternatively, as neutron stars receive a ‘kick’ during their formation, they could form a corona around the galaxy.

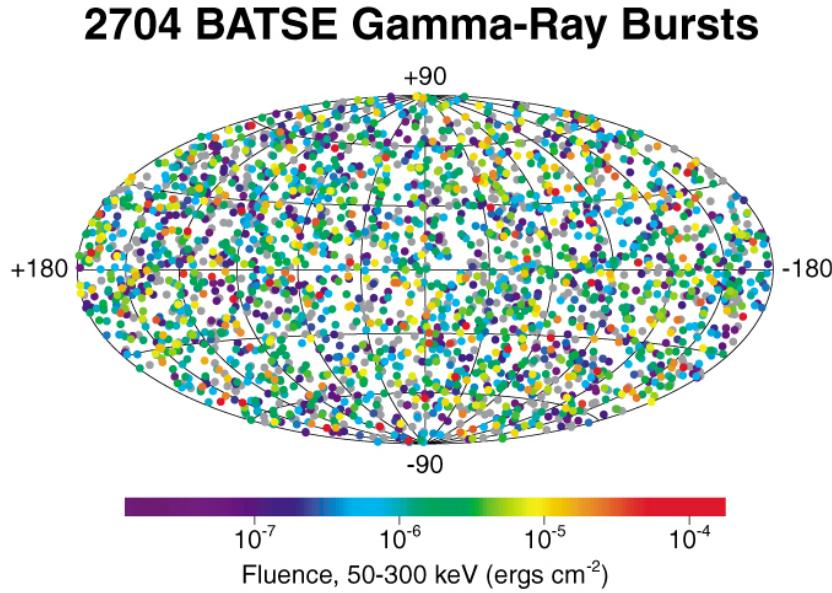


Figure 3.2: **BATSE GRB Fluence.** This plot shows the fluence (given by the colour of each point) and the sky position of each GRB detected by the BATSE mission. [29]

half a degree. During its mission, BATSE detected approximately 2700 GRBs and determined the sky position of a large number of these. For the first time their was a statistically significant sample of well localised GRBs. In figure 3.2 we can see the sky location of every GRB detected by BATSE, together with the fluence<sup>2</sup> of that GRB. This plot shows that GRBs are isotropically distributed over the sky. This evidence strongly suggests GRBs are of cosmological origin.

The BATSE data also provided evidence that GRBs were uniformly distributed and that we were seeing a limited horizon, beyond which GRBs became much harder to detect. This evidence came in the form of a  $\log N - \log P$  distribution, where  $N$  is the number of detected GRBs and  $P$  is the peak flux<sup>3</sup>. If GRBs are uniformly distributed in space then the number of GRBs out to a given distance increases as the cube of that distance. However, peak flux from the GRBs would decrease with the inverse square of the distance. Hence, plotting  $\log N$  against  $\log P$ , we expect to find a gradient of approximately  $-3/2$ . Any short fall from this expected distribution suggests that we have reached a horizon for detectable GRBs. In figure 3.3 we see the  $\log N - \log P$  distribution for a combined set of GRBs detected by BATSE and the Pioneer Venus Orbiter (PVO).<sup>4</sup>. The gradient for high energy GRBs is the expected  $-3/2$  but there is a shortfall at low energies, suggesting there is a

<sup>2</sup>Fluence is the time integral of the flux. Essentially a measure of the total detected energy.

<sup>3</sup>Some times the  $\log N - \log S$  distribution is preferred, with  $S$  being the given flux. The essence of the plots is the same.

<sup>4</sup> The PVO was less sensitive than BATSE, but it operated for 10 years and so observed a fairly large number of GRBs

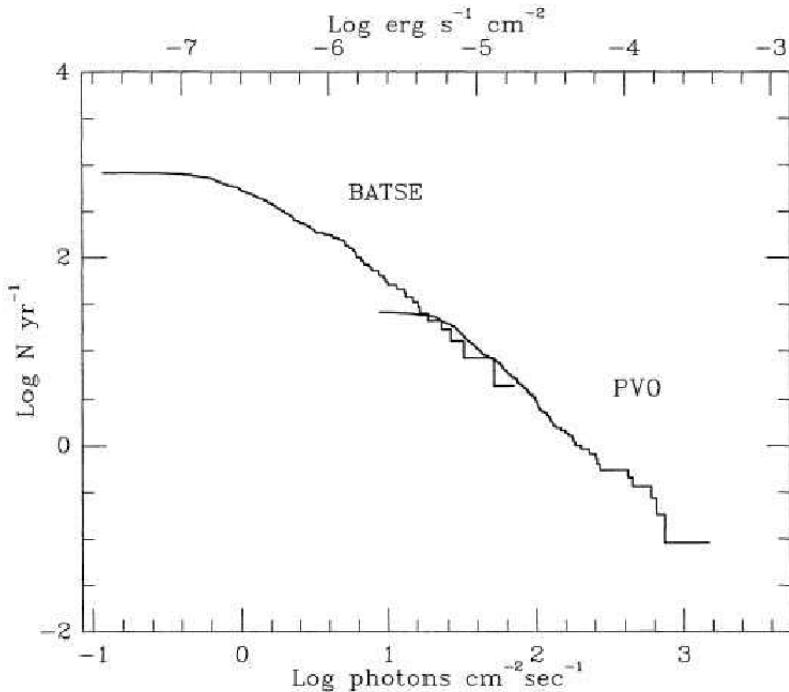


Figure 3.3: **LogN-logP for BATSEVPO.** Here we plot the log of the number of GRBs against the log of the peak flux. The sample includes GRBs detected by BATSE and by PVO. The energy range for BATSE was 50-300keV, and the energy range for PVO was 100-2000keV. For uniformly distributed GRBs, we expect this plot to have a gradient of  $-3/2$ . The expected gradient is observed for high energy GRBs but not at lower energies. This suggests a limited distance to which GRBs can be observed. cite [https://www.researchgate.net/figure/Distribution-log-N-log-P-for-a-combined-set-of-BATSEPV0-data-The-distributions-match\\_fig5\\_242389649](https://www.researchgate.net/figure/Distribution-log-N-log-P-for-a-combined-set-of-BATSEPV0-data-The-distributions-match_fig5_242389649) cite <https://www.nature.com/articles/366040a0.pdf>

limited distance to which GRBs can be viewed. Unfortunately this data does not inform us of where that horizon is. It could be within our galaxy, or it could be the horizon of the universe.

### 3.1.3 The Long and Short of Gamma-ray Bursts

Another important piece of evidence into the origins of GRBs came from their duration. As every burst has different properties<sup>5</sup> the duration of a GRB is not trivially defined. The measure most commonly used is the  $T_{90}$ , the time over which 90% of the total fluence is recorded.<sup>6</sup>

The top panel of figure 3.4 shows the number of bursts with a given  $T_{90}$  for the BATSE data. This plot shows that there are two populations of GRBs, the first with a  $T_{90}$  value of about 0.5s and the second with a  $T_{90}$  of about 30s. It is also

<sup>5</sup>For example, some bursts have multiple flares.

<sup>6</sup>Another common measure is the similarly defined  $T_{50}$ .

clear from this data that the longer population of GRBs are detected far more often. Plotting the spectral hardness of the BATSE GRBs against  $T_{90}$ , as has been done in figure 3.4, we see that the shorter GRBs also have a harder spectrum than the longer GRBs. This means short GRBs emit more high energy photons than long GRBs. Because of these two properties of the two populations, they are known as *short-hard* and *long-soft* GRBs. It is common to use the criteria that short GRBs are those that are less than 2s, and long GRBs are longer than 4s, with those inbetween being called *intermediate* GRBs.

It should be mentioned that while  $T_{90}$  is a very useful tool, it is instrument dependent. This is because more sensitive instruments will track GRBs for longer, and bursts have different durations in different energy bands. Also, the  $T_{90}$  is measured in the detector frame, and not the rest frame of the burst, which would make a GRB at a redshift of  $z$  appear a factor of  $(1+z)$  longer. For these reasons, using the  $T_{90}$  to classify short/long GRBs should only be considered approximate.

As a final note on GRB durations, it should be mentioned that there is some evidence that there may be intermediate and ultra-long populations of GRBs. It is an open question whether these GRBs represent new populations of GRBs or are part of the long and short GRB populations. As these are still contentious, we do not consider them further.

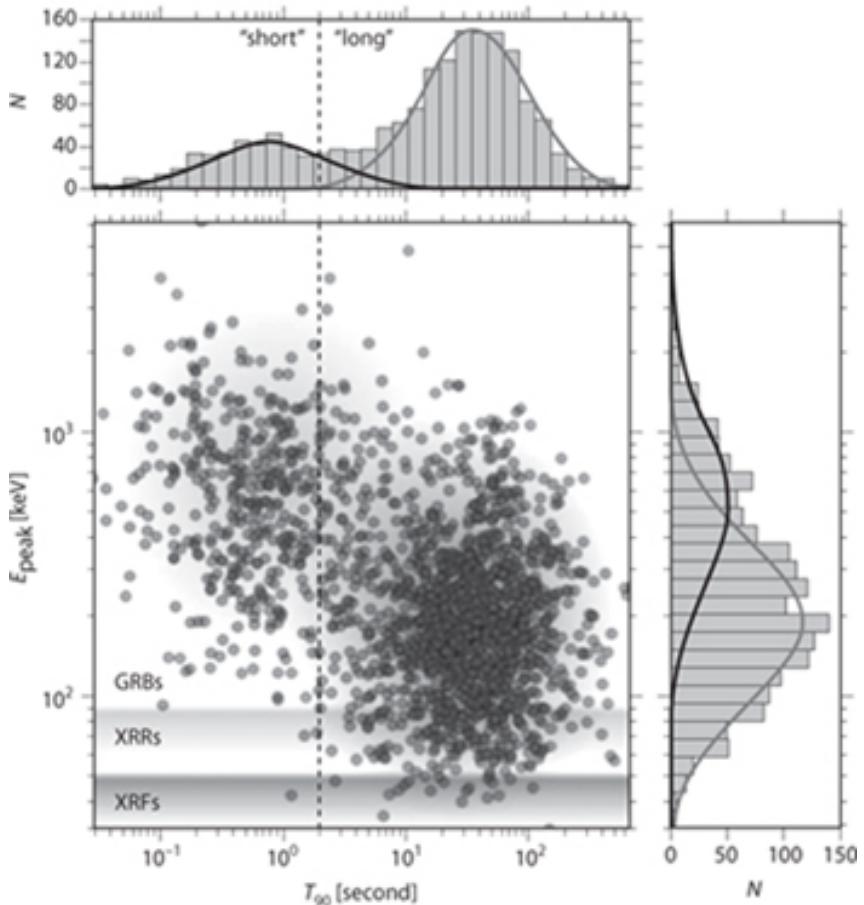
### 3.1.4 BeppoSAX and the First Afterglows

Important clues had been found into the origin of GRBs, but still no one had found any trace of a GRB after the prompt emission (the initial flash of gamma-rays). The search area provided by GRB detectors at the time were too large for ground based telescopes to have a realistic chance of finding the source of the GRB, though attempts were made. This changed with the launch of the BeppoSAX satellite in 1996. BeppoSAX was able to localise to within a few arcminutes, much better than the half a degree BATSE was capable of.<sup>7</sup> It also had a Narrow-Field x-ray Instrument (NFI) to search for X-ray counterparts to GRBs.

On February 28th 1997 a GRB was detected by BeppoSAX that was localised well enough for the satellite to aim the NFI at it. A new X-ray source was detected that faded away slowly over the next few days. The lower energy emission detectable in the days after the prompt emission is called the GRB *afterglow*, and this was what BeppoSAX had detected. The localisation was accurate enough to allow ground based telescopes to find the optical counterpart to this GRB as well. The optical images showed what some argued was a distant galaxy and others argued was a galactic nebula. On May 8th 1997, a GRB was detected by BeppoSAX and an

---

<sup>7</sup>It is true that the InterPlanetary Network, a network of GRB detectors placed on various spacecraft throughout the solar system (see section 3.4), could localise better than BATSE. The problem was that it took too long find the sky position of the GRB using this method, and any trace of the GRB had disappeared before astronomers could find it.



**Figure 3.4: T90 vs the spectral hardness ratio.** Here we plot the  $T_{90}$  values and the spectral hardness ratio for the BATSE GRBs. The top panel shows a histogram of the  $T_{90}$  data, which clearly has two populations of GRBs, short and long. The main plot shows  $T_{90}$  against spectral hardness, which makes the two populations even more clear and shows that short GRBs have harder spectra than long GRBs. Those GRBs with the greatest ratio of energy in the X-ray to gamma ray band, generally those with a peak energy of less than 15 keV, are called *X-ray flashes* (XRF). Those with comparable energy in the gamma-ray and X-ray band are called *X-ray rich* (XRR) GRBs. All other GRBs are simply called GRBs. These different classes of GRBs are marked on the plot. Also shown is the 2 second dividing line between short and long GRBs. **cite What are Gamma-ray Bursts? Joshua Bloom**

optical counterpart quickly found. A spectrum was obtained for the counterpart which showed iron and magnesium absorption lines that had been significantly redshifted. This showed that the GRB must have occurred at a distance greater than 5 Gpc and that the light passed through some gaseous cloud in a distant galaxy on its way to Earth. With this observation there could be no doubt that GRBs were originating at cosmological distances, as had been suspected.

### 3.1.5 The Fireball Model

The *fireball shock model* was developed in the 1990s. It attempted to describe the physical processes that cause GRBs without making many assumptions about the source of energy for the GRB, the *central engine*. At this time the evidence was building that GRBs originated at cosmological distances. At such great distances the inferred energy emitted at the source is far higher, as much as a solar mass if emitted isotropically. Light-travel time arguments showed that GRB sources must also be small, at most hundreds of kilometers across (see section 3.1.1). Realising that all GRBs are small and highly energetic was the starting point of the fireball shock model.

The next step was to notice that photons detected from GRBs are often above the pair production threshold<sup>8</sup> ( $2 \times m_e c^2 \approx 1\text{MeV}$ ), and so should have created electron-positron pairs rather than gamma-rays. This problem is solved by assuming that the energy released drives a relativistic expansion from the source. In this case the energy of the photons in their rest frame is inversely proportional to the bulk Lorentz factor of the outflow, i.e. the energy in the rest frame of the photons is much lower. It also causes photons to bunch up just ahead of the relativistic matter that is emitting the photons, causing it to seem more energetic to an observer. Accounting for these factors, it can be shown that the Lorentz factor must be in the hundreds for a typical GRB. This is as compared to a Lorentz factor of approximately 1.001 for a fast supernovae, and corresponds to GRB velocities of at least 99.9% of the speed of light.

Combining the evidence for the high speed of the ejected matter with the small size of the emitting object means that a large amount of energy must be rapidly dumped into a small volume before the GRB is emitted. Most of the energy is in the form of photons, so this is called the *radiation dominated* phase. This soup of photons with a small amount of matter is called the *fireball*. The fireball expands and the Lorentz factor grows with it. As it expands, the energy is absorbed by the matter (protons and electrons) in the form of kinetic energy. This is the *matter dominated* phase. Most of the particles are moving in the same direction now, with little random motion, i.e. the fireball has become cold.

A mechanism is needed to turn this kinetic energy into a GRB. The simplest

---

<sup>8</sup>Although for GRB 170717A, this was not case. See section 3.2 for more details

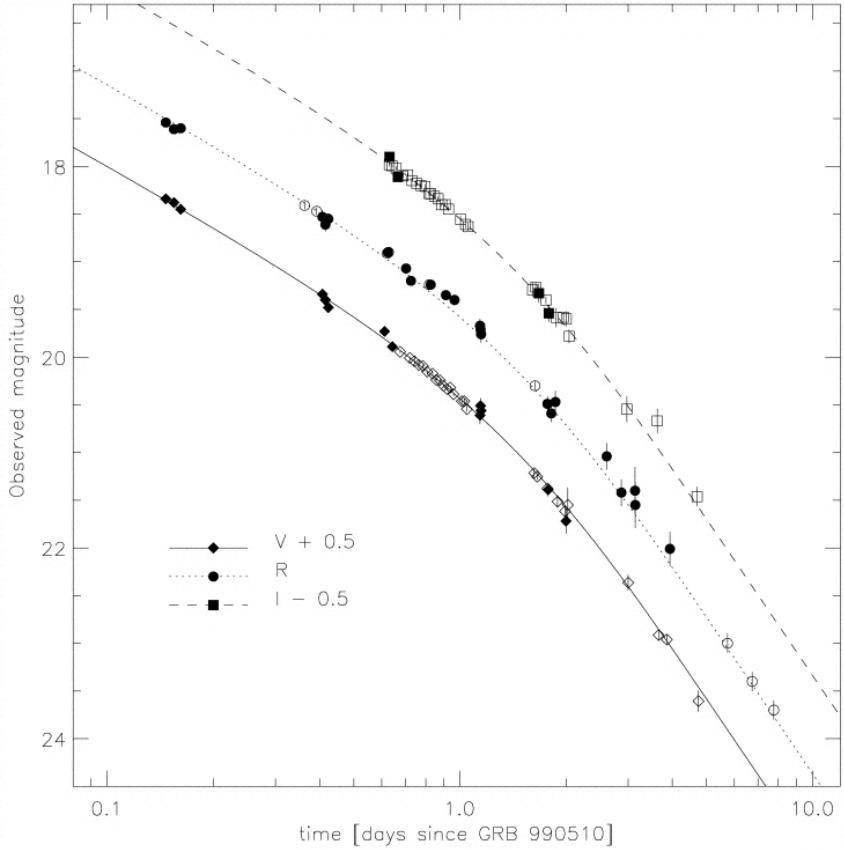
way to do this is to have the matter collide with slower moving matter surrounding the system. The matter is too sparse for direct collisions to happen enough to form a GRB. Instead, it is thought that magnetic fields near the edge of the fireball can cause the matter to slow down and radiate its energy. When the fireball interacts with the surrounding matter, the rapid change in temperature, pressure, and density travel through the medium faster than the medium can react. Like the sonic boom of a supersonic jet, this causes *shocks* in the surrounding medium. It is these shocks that are thought to be the source of the GRB. Precisely how these shocks power a GRB is not known. *Fermi acceleration* possibly plays a role. This is where charged electrons enter the shock and are reflected back by magnetic fields, increasing their kinetic energy. After several iterations of this, the electron can travel even faster than the shock. The magnetic fields could then cause the electrons path to curve, causing them to emit energy as synchotron radiation. Alternatively they might interact with photons, imparting their considerable energy onto the photon to create a high energy gamma-ray. This process is called *inverse Compton scattering*. There is still a lot to learn about the processes that power the GRB.

There are also two hypotheses as to what the slower moving matter that the fireball collides with could be. The first is simply material around the star, the *circumburst medium* (CBM). This is called an *external shock scenario*. The other theory is that multiple shells of material are emitted, and the shocks are created when faster moving shells catch up with slower moving shells. This is called the *internal shock scenario*. The internal shock method is generally favoured for the prompt emission. This is because GRB light curves are highly variable, with some showing multiple peaks (see figure 3.1). This is easily explained by the multiple shocks of the internal shock scenario, but more difficult to explain for external shocks. External shocks are thought to be the cause of the GRB afterglow.

### 3.1.6 Jets

The well-founded assumptions made by the fireball model, that GRB progenitors are small and dense, naturally lead to the prime suspects for powering GRBs being compact objects, such as black holes and neutron stars. If we assume that compact objects do play a role in generating GRBs, then it is not much of a jump to think that GRBs may be jetted. Jets are ubiquitous with both black holes and neutron stars: supermassive black holes and solar mass black holes have been observed to have jets, and highly magnetised, rotating neutron stars, called *pulsars*, are also known to produce jets. Assuming that GRBs are jetted substantially lowers the inferred energy emitted by the source, which has implications for the physical mechanism that causes GRBs. It is therefore important to determine whether GRBs are jetted or not.

What would be the observable effects of GRB jetting? The fireball model sug-



**Figure 3.5: Break in spectrum due to jetting.** Here we see the optical light curves for the afterglow GRB 990510. A break in the spectrum is visible after approximately one day. [cite <https://iopscience.iop.org/article/10.1086/312282/fulltext/>](https://iopscience.iop.org/article/10.1086/312282/fulltext/)

gests that the ejected matter will have a very high Lorentz factor. Special relativity tells us that objects moving with high Lorentz factor  $\Gamma$  emit most of their energy within an angle  $\theta \approx 2/\Gamma$  radians of the direction of travel. This is called *relativistic beaming*. If the matter is jetted and the beamed angle  $\theta$  is less than the opening angle of the jet, then there will be no observable difference between isotropic emission and jetted emission, as the edge of the jet will not be visible. However, once the matter reaches the interstellar medium it will begin to slow down, reducing the amount of beaming and increasing the angle to which the matter can be observed. Once the beaming angle has increased to a point where the edge of the jet is visible, the amount of energy being radiated towards the observer will suddenly drop, and this will happen across the whole electromagnetic spectrum at the same time. This *break* in the spectrum is observable in the afterglow of many GRBs. For example, figure 3.5 shows the clearly visible break in the afterglow spectrum of GRB 990510 after approximately one day. It is now widely accepted that short GRBs have opening angles of  $\sim 30^\circ$  and long GRBs have opening angles of  $\sim 5^\circ$ .

### 3.1.7 The Long GRB-Supernova Connection

As more afterglows were found, trends began to appear. Long GRBs tended to occur directly on host galaxies, not randomly in space.<sup>9</sup> Spectroscopy of these galaxies often found the presence of emission lines excited by star formation. In fact, no long GRB has been found in a non-starforming galaxy. The host galaxies also tended to be relatively faint. Low luminosity (i.e. low stellar mass) suggests galaxies have a low metallicity [cite someone], and spectrographic studies confirmed this. Star forming galaxies with low metallicity, such as those from which long GRBs originate, are also where type 1b and 1c supernovae are observed. Type 1b and 1c supernovae result from the core collapse of massive stars that have shed off their outer layers of hydrogen. In the case of type 1c supernovae, they have also shed most or all of their helium as well. That long GRBs originate from star forming galaxies with low metallicity raises the possibility that long GRBs are connected to the massive stars that can only exist in these kinds of galaxies. In particular, it is thought that a type 1b or 1c supernova could power a GRB.

The strongest evidence of a supernova connection came with the detection of GRB 980425. This GRB was exceptionally close; at just  $z = 0.0085$  it is still the closest GRB ever detected. Followup observations found a rising supernova, SN 1998bw. It was a very bright supernova, about ten times brighter than normal. It also showed no hydrogen or helium emission lines, making this a type 1c supernova. It seemed highly unlikely that the coincident GRB and supernova were unrelated. Questions remained though, because this GRB was exceptionally faint given its distance. Was it typical of other long GRBs?

After GRB 980425, searches were undertaken for other bright supernova afterglows associated with long GRBs. Not only were many found, but they were shown to have similar spectra to GRB 980425. In particular, GRB 030329 was another nearby GRB, at redshift  $z = 0.17$ , which had very detailed followup. It was 1000 times brighter than GRB 980425, but showed the same spectral features. With this discovery, the consensus grew that long GRBs were caused by type Ic supernova.

### 3.1.8 The Short GRB-Compact Binary Connection

Much had been learned about long GRBs by studying their afterglows, but no afterglow had been found for short GRBs. This was because short GRBs fade rapidly, reducing the amount of time for follow up observations. Rapid follow up was needed, and so the *Swift* satellite (see section 3.4) was launched in 2004. Swift was able to autonomously followup a GRB with X-ray, UV, and optical measurements within minutes, just what was needed for short GRBs.

The first detected short GRB afterglow was that of GRB 050509B. It was a

---

<sup>9</sup>We will see in section 3.1.8 that this is unlike short GRBs, which can originate quite far from any host galaxy.

faint X-ray afterglow that faded quickly, but it was localised well enough for ground based telescopes to determine the host galaxy, which had a redshift of  $z = 0.225$ . The afterglow and host galaxy were very different to those of long GRBs:

- The host galaxy was a massive elliptical galaxy which showed no evidence of star forming.
- The host galaxy was relatively nearby.
- The GRB took place far away from the galactic core.
- There was also no evidence of a supernova.

With more short GRB afterglow detections it became clear that these are all typical properties of short GRBs. Short GRBs can occur in any type of galaxy, unlike long GRBs. They show no evidence of being caused by a supernova. On average, short GRBs occur much closer than long GRBs<sup>10</sup>, though this could be partly due to selection effects.<sup>11</sup> Also short GRBs occur further from the center of their host galaxy than long GRBs, with an average offset of 4.5kpc and 1.5kpc respectively. Sometimes they are so far from any galaxy that determining the host is impossible; these are called *hostless* GRBs.

The observed properties of short GRBs can be explained naturally if short GRBs are caused by the merger of a neutron star with either another neutron star or a black hole. In this case the two objects would form in a binary system and slowly inspiral due to the loss of energy through gravitational wave emission, until they finally merge and emit the GRB. For neutron star - black hole (NS-BH) mergers, the black hole must be relatively low mass (less than  $\sim 10M_{\odot}$ )[\[cite paper\]](#). This is because if the black hole has a high mass then the neutron star will be swallowed whole by the black hole, and there will be no material to produce a GRB. For a low mass black hole the neutron star will be tidally disrupted before merger, i.e. the black hole will pull apart the neutron star, providing matter that can then power a GRB. The merger time for a binary scales with  $a^4$ , where  $a$  is the initial separation. This means a small change in the initial separation can lead to a very different merger time.<sup>12</sup> This explains their presence in elliptical galaxies that have long since stopped star forming, as well as in younger, star-forming galaxies.

It is also expected that these binary systems would receive a *kick* in their formation, which explains why so many short GRBs occur far from the center of their host galaxy. There appear to be two mechanisms through which the binary could receive a kick. The first is due to the supernovas that formed each component of the binary.

---

<sup>10</sup>The average redshift of a short GRB is about  $z = 0.5$ , while less than 10% of long GRBs have a redshift less than one. [\[cite https://arxiv.org/abs/1702.03338\]](#)

<sup>11</sup>Short GRBs tend to be dimmer and so not observable at greater distances. Also, GRBs at high redshifts will appear longer in the detector frame.

<sup>12</sup>In fact, some galactic binaries have merger times that exceed the age of the universe.

As the supernova will cause a large amount of matter to become unbounded from the system at a particular point in the orbit, conservation of momentum forces the binary system to recoil. This cannot be the only mechanism that causes a kick to neutron stars though, as pulsar observations show that even lone neutron stars receive kicks from the supernova that forms them, though the mechanism that causes this kick is not known.

The binary inspiral theory also explains why short GRBs are short. The duration of a GRB depends on how long it takes the gamma-rays to break through any surrounding material and how long the central engine remains active. For a core collapse supernova there can be a lot of material around the central engine, and the amount of time the central engine is active could also be highly variable. This large amount of variability explains why long GRBs can last from seconds to hours. A binary inspiral is expected to clear out the surrounding system of any matter long before merger. The only source of matter around a binary merger is from the stars themselves. And as the GRB emitted at or immediately after the merger of the compact objects (see section 3.3), the amount of time the central engine is active is not very variable. Putting this together we conclude that binary mergers cannot produce long GRBs, as the central engine is not active for very long and there is not a lot of matter for the GRB to break through.

The observation of supernova remnants coincident with long GRBs demonstrated the long GRB - supernova connection. One might expect to find an analogous remnant for short GRBs. But what would the remnant of a neutron star merger look like? Would it even be observable? It is useful to first consider the circumstances and processes that cause supernovas, and then contrast them with those of a neutron star merger. Supernova ejecta emits light due to the radioactive decay of *s*-process elements. The s-process refers to when an atom captures a neutron, which makes the atom unstable, and then  $\beta$ -decays<sup>13</sup> to a heavier element. This is how the iron group elements are produced and it is also what makes supernovas radiate light.

The s-process occurs only if the unstable atom decays before it captures another neutron. This is unlikely in neutron rich, high temperature environments, such as a neutron star merger. Instead the atoms will capture multiple neutrons before they  $\beta$ -decay. This is called the *r*-process and it creates elements with a high atomic number, such as gold and iodine. The light emitted by the decay of the r-process elements after a neutron star merger is called a *kilonova*. The r-process elements are optically opaque, making kilonovas peak in the infrared (IR) and be about 10-100 times fainter than supernova.

It is difficult to detect IR using ground based telescopes, due to atmospheric distortion and a bright sky. Despite this difficulty, a kilonova was eventually observed in conjunction with GRB 130603B, a short GRB. The Hubble space telescope observed the event fade over 10 days until it was undetectable in optical light, but still

---

<sup>13</sup> $\beta$ -decay is when a neutron decays into a proton, electron, and an anti-neutrino.

clearly visible in IR. The IR light was brighter than would be expected by simply extrapolating from the afterglow. This additional IR component to the light was the kilonova. This was the strongest evidence yet that neutron star mergers were the progenitors of short GRBs.

As discussed in chapter 2, compact binary coalescences (CBCs) are known to be strong emitters of GWs. This made short GRBs a promising target for GW astronomy, as a GW signal of a neutron star binary inspiral detected in coincidence with a short GRB would be the smoking-gun that neutron star mergers can produce short GRBs.

## 3.2 Gravitational Waves and GRB 170817A

On the 17th of August 2017, the Fermi Gamma-ray Burst Monitor (see section 3.4) observed a faint short GRB just two seconds after the LIGO and Virgo observatories detected a GW signal from a binary neutron star merger with a network SNR of 32.4 (see figure 3.6) [**cite BNS detection paper**]. Comparing the sky localisation of the GW detector network and the Fermi GRB detector, the source of the GRB was localised to a small sky patch. Ground based observatories scanned the sky patch for a remnant, and quickly identified a bright object near the galaxy NGC 4993 [**cite detection papers**]. This object had not been there when the same galaxy had been previously observed (see figure 3.10), and this galaxy was at the right distance as determined from the GW signal. Further study revealed the new object to be a rapidly evolving kilonova, confirming the theory that short GRBs are caused by neutron star mergers.

In this section we will discuss in detail the coincident detection of GW 170817 with GRB 170817A. We will start by discussing the initial GW detection, focusing on the data analysis used to determine the sky location and properties of the source. We will then discuss the results of the EM followup that found the afterglow of the GRB and showed that it was a kilonova. We will then end with a discussion of late time EM observations, which showed that GRB 170817A had a structured jet which was seen off-axis, explaining the relatively faint prompt emission.

### 3.2.1 Initial Observation and Followup

The initial detection of GW 170817 was made only by the LIGO Hanford observatory, despite the fact that both LIGO detectors and the Virgo detector were in observing mode. A quick investigation found that the Livingston detector data had not been included by the low-latency search due to a glitch approximately 1.5 seconds before the coalescence phase of the signal. Glitches such as this happen at a rate of once every few hours in the LIGO detectors. These glitches are not temporally correlated between the two LIGO detectors and their source is unknown.

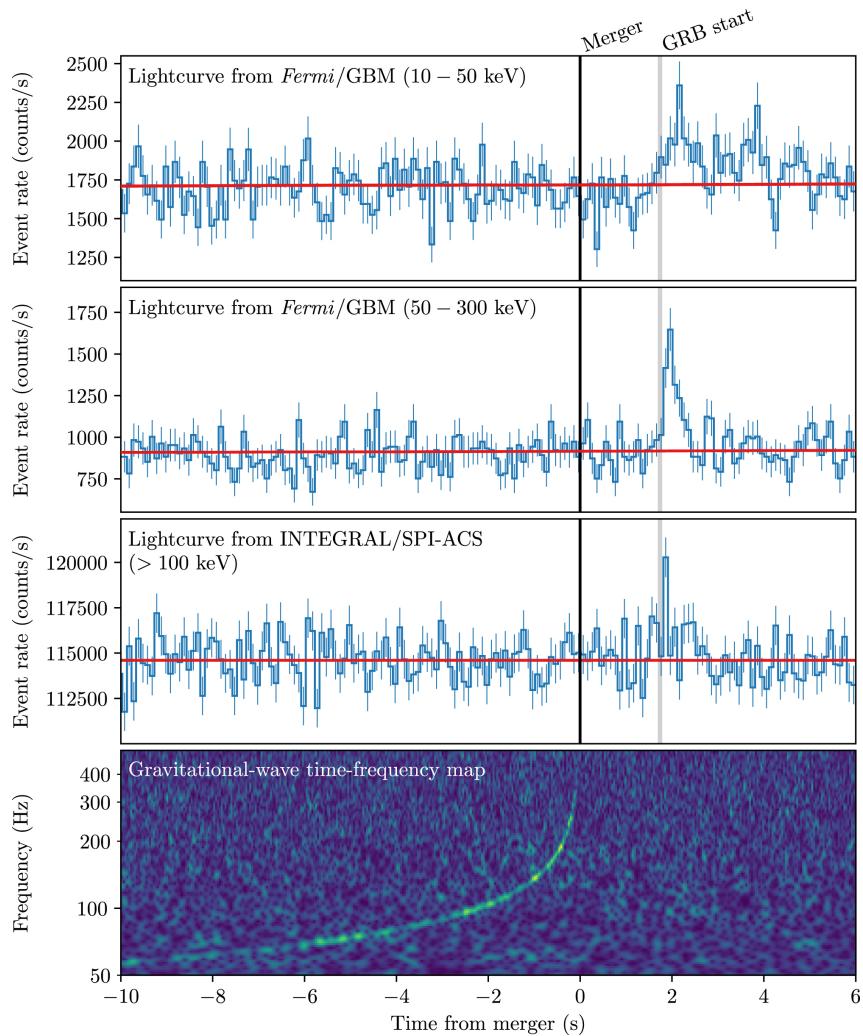


Figure 3.6: **GRB170817A and GW170817.** Here we see that the a coherent combination of the Hanford and Livingston strain data from GW170817 in the bottom panel. The top two panels shows the Fermi GRM curve in the 10-50keV and the 50-300keV range respectively. The INTEGRAL/SPI-ACS data is shown in the third plot. The background estimate for each detector is indicated by the red line. Note that the GRB was detected 1.7 seconds after the GW signal was detected. We can also see that Fermi detected a longer, softer signal in the 10-50 keV range, that lasted for a few seconds after the triggering pulse. [cite Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A](#)

Despite the glitch, the GW signal is clearly visible in the data (see the top panel of figure 3.7).

Due to the coincidence with GRB 170817A, a rapid reanalysis of the data was performed, with the glitch removed from the data using an inverse Tukey window (see the bottom panel of figure 3.7). Removing a glitch like this lowers the reported SNR of a matched filter search compared to if there was no glitch, but allows the trigger to pass signal consistency tests<sup>14</sup>. The data from the LIGO and Virgo detectors, with the Livingston glitch and some other noise removed, is shown in figure 3.8. The GW signal, clearly visible in the LIGO detectors, had an SNR of 18.8 in Hanford and 26.4 in Livingston, but an SNR of just 2.0 in Virgo.

The high SNR in the LIGO detectors compared to Virgo suggested that the GW originated from a part of the sky that Virgo was not sensitive to at the time of the trigger (see section 2.4.2). The source of the GW was localised to within 31 deg<sup>2</sup> using the time delay between the LIGO detectors and the fact that the source originated from a null of the Virgo detector<sup>15</sup> [**cite BNS detection paper**] and the sky localisation determined using GW data was consistent with the data determined by GRB detectors (see figure 3.9). GW data also determined the distance to the source was approximately 40 Mpc, close enough that relatively complete galaxy catalogues exist. Using this information, ground based observatories began scanning the sky patch for an afterglow. Less than 11 hours after the initial trigger, the Swope telescope in Chile, followed by five other observatories [**cite kilonova paper?**], found a new object on the edge of galaxy NGC 4993 (see figure 3.10).

### 3.2.2 Kilonova Observation

This new object was studied intensely over the following days. The transient was initially observed to be a rapidly fading blue transient, which had faded completely within 48 hours. The spectrum showed no supernova-like absorption lines, ruling out the possibility that this transient was caused by a supernova. Over the next  $\sim$ 10 days the spectrum grew redder, and observations by ESO-VLT/X-shooter showed evidence of the decay of r-process elements. This all strongly indicates that the transient was a kilonova. X-ray emission was detected 9 days after the merger, and radio emission after 16 days. This delayed radio emission was predicted from neutron star merger models as the ejecta interacts with the interstellar medium.

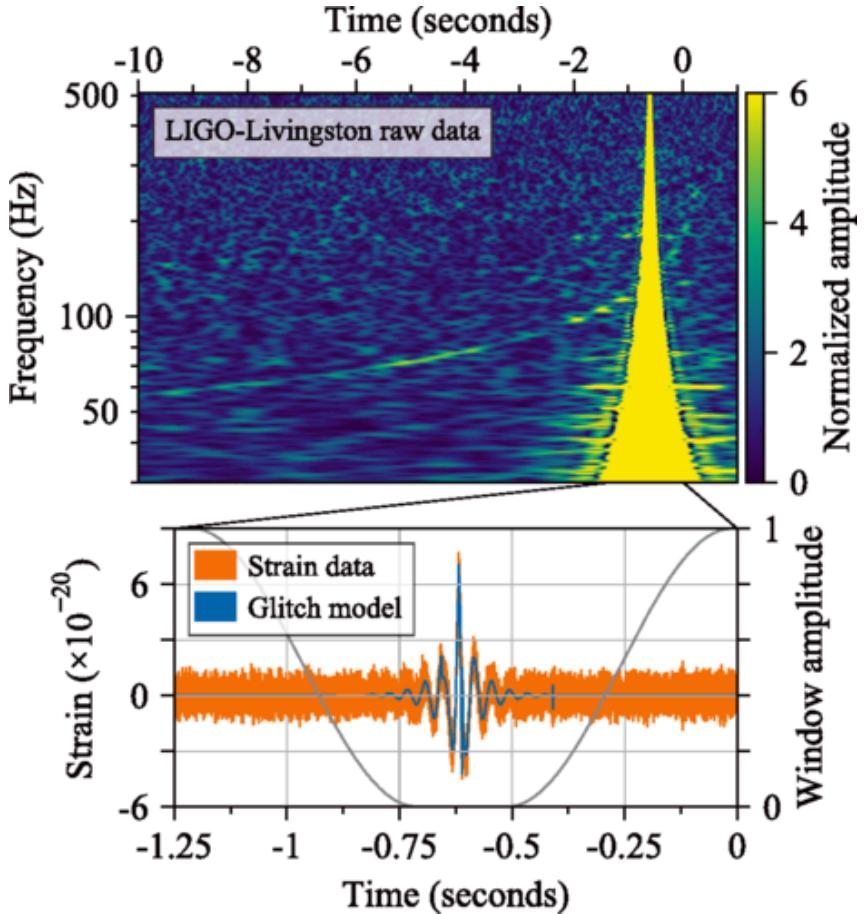
### 3.2.3 Structured Jets

GRB 170817A was significantly fainter than any other detected GRB (see figure 3.11). In fact, the GRB showed no evidence for photons with energy  $> 511$  keV,

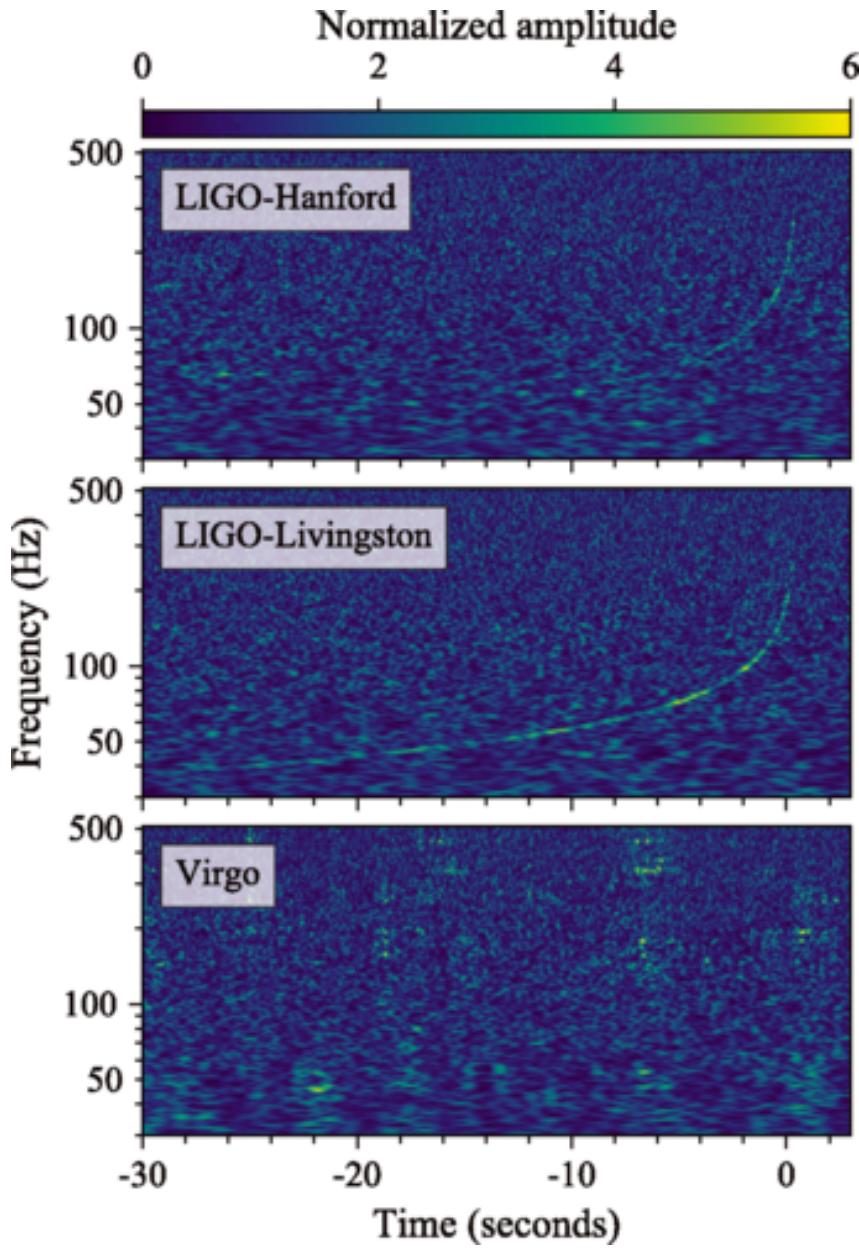
---

<sup>14</sup>In fact, some offline analyses automatically gate out glitches, see section ?? and [**cite someone**]

<sup>15</sup>For parameter estimation analysis of the signal, the glitch was modelled and subtracted from the data (see the bottom panel of figure 3.7), which reduced the sky patch to an area of 28 deg<sup>2</sup>.



**Figure 3.7: Glitch in the LIGO Livingston Observatory.** The top panel shows a time frequency map for the Livingston observatory data at the detection time of GW 170817. A glitch is clearly visible approximately 1.5 seconds before the end of the signal. Despite this the signal is still clearly visible. The bottom plot shows the raw strain data from the Livingston observatory. This data is bandpassed between 30 Hz and 2 kHz to emphasise the sensitive range of the detector. The grey curve (and right axis) shows the inverse Tukey window used to smoothly zero out the data around the glitch before the rapid reanalysis of the data. The blue curve shows the waveform model used to subtract the glitch from the data before measurements of the source's properties were made. [cite GW170817 observation from BNS](#)



**Figure 3.8: GW 170817 Detection.** Here we see time frequency maps of the LIGO Hanford and Livingston observatories, and the Virgo observatory at the detection time of GW 170817. This data has been whitened and independently observable noise sources have been subtracted, including a glitch in the Livingston data. The non-detection by Virgo significantly reduced the amount of the sky that the signal could have originated from. **cite GW170817 observation from BNS**

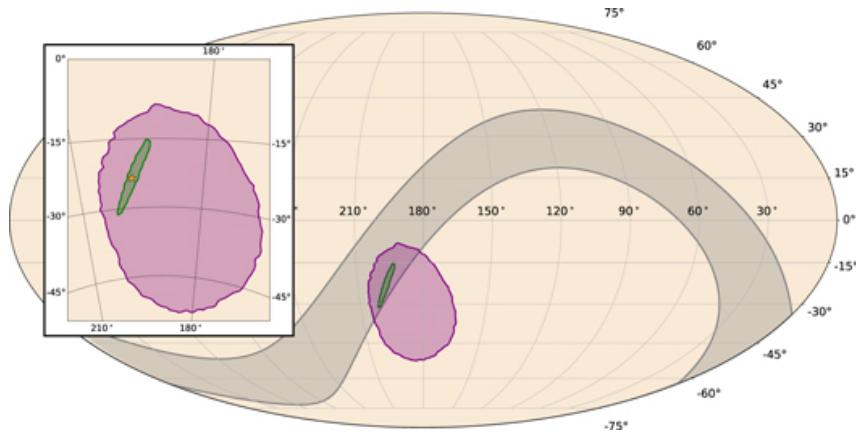


Figure 3.9: **Sky map for GW170817/GRB170817A.** Here we see the 90% confidence sky localisation for the LIGO and Virgo collaborations in green, the GBM 90% localisation in purple, and the annulus formed by Fermi and INTEGRAL timing information in grey. [cite Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A](#)

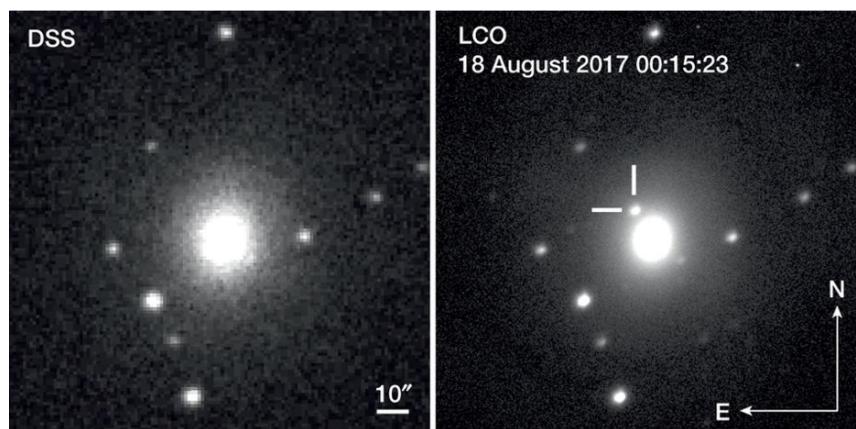
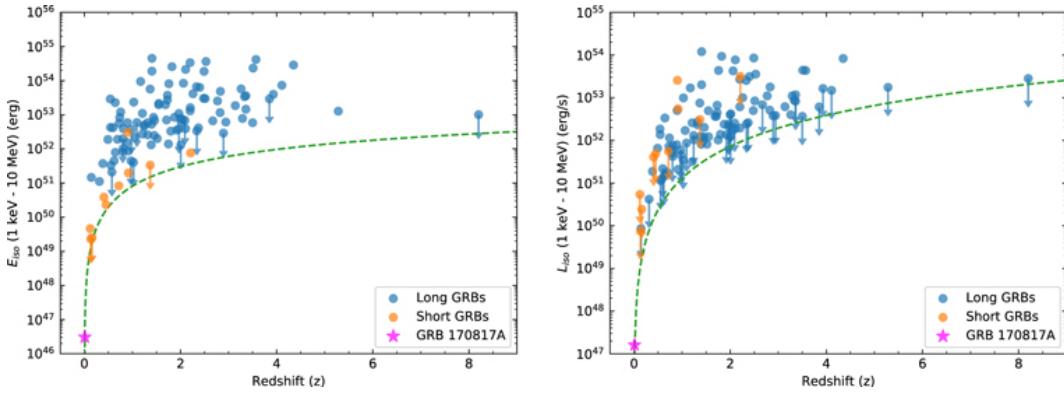


Figure 3.10: **NGC 4993.** Image of NGC 4993 taken in 1992 by the Anglo-Australian Observatory (left) and August 18th 2017 by the Las Cumbres Observatory (right). [cite <https://www.nature.com/articles/nature24291/figures/2>](https://www.nature.com/articles/nature24291/figures/2)

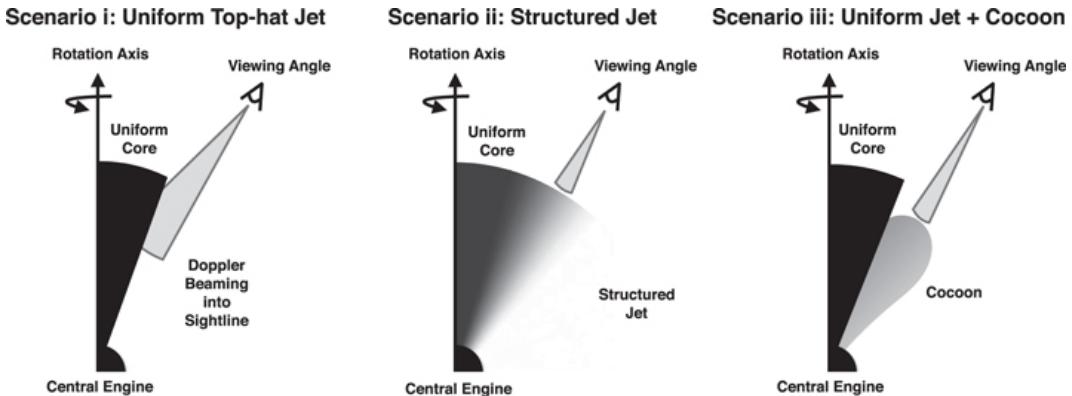


**Figure 3.11: Brightness/Luminosity against redshift.** Here we see the distribution of  $E_{\text{iso}}$  and  $L_{\text{iso}}$  against redshift for every GBM-detected GRB with a measured redshift. For GRBs with power law spectra, marked with a downward pointing arrow, this is taken to be an upper limit. This is because the spectra must have curvature, and so extrapolating a power law leads to an overestimation. The green dashed line shows the approximate detection threshold for the GBM. These plots show that GRB 170817A was much dimmer than any other detected GRB. [cite GRB BNS paper]

i.e. above the pair production threshold, meaning that the matter ejected from this GRB need not have been traveling at relativistic velocities<sup>16</sup>. There are several factors that affect GRB brightness, such as the angle the GRB is observed at and the intrinsic properties of the jet. The simplest model of the jet is the *top-hat jet*, which has a uniform core that terminates sharply at some angle from the GRB axis. It is possible that GRB 170817A was a top-hat jet viewed off-axis, making it appear dimmer. Another possibility is that the GRB had a *structured jet*, meaning that it gradually becomes less energetic as the angle from the axis increases. It is also possible that the GRB jet had a *cocoon* created by the relativistic jet shocking the non-relativistic matter surrounding the jet. These three possibilities are shown in figure 3.12. It could also be that GRB 170817A was produced by a new mechanism that is not observable at greater distances as it is intrinsically dim. It could also be that GRB 170817A is part of a population of subluminous GRBs that can only be detected if they occur unusually close or that the GRB was not jetted at all, and is a mildly relativistic, isotropic fireball. In this section we will consider each of these possibilities and compare each model to the observed spectral evolution of GRB 170817A, from the prompt emission to the late afterglow observations. We will see that the best explanation for the faint prompt emission of GRB 170817A is that it had a structured jet seen at a relatively large viewing angle.

We begin by using some simple arguments to show some of these models are

<sup>16</sup>Recall from section 3.1.5 that GRBs were assumed to accelerate matter to relativistic velocities as this would explain how GRBs could emit photons above the pair production threshold. The fact these high energy photons were not detected means that material ejected from this GRB need not travel at relativistic velocity



**Figure 3.12: Jet Structure Scenarios.** Three different scenarios that could explain the low luminosity of GRB 170817A. The first scenario is that a Top-hat jet was viewed off-axis. The second is that the jet is structured, with photons emitted further from the axis being lower energy and fewer in number, and viewed relatively far from the axis. The third scenario is that a uniform jet has a surrounding cocoon that emits lower energy photons, and it was these lower energy photons that were detected. [cite GRB BNS paper](#)

unlikely. The observed gamma-ray properties of GRB 170817A are similar to other short GRBs, making it unlikely that the prompt emission was caused by a different mechanism to other short GRBs. It is also unlikely that GRB 170817A represents the first detected member of a subluminous population as this would mean short GRB luminosities covers six orders of magnitude, which is difficult to conceive given the small range of physically possible neutron star masses. Though it should be noted that a wider range of intrinsic luminosities is possible if we assume that some short GRBs are produced by NS-BH mergers, or that the magnetic field strength of GRB progenitors is highly variable and significantly affects the intrinsic luminosity [[cite BNS GRB paper](#)].

This leaves only those models that focus on the jet structure and viewing angle of the prompt emission. In [[cite afterglow paper](#)] they compare the late time afterglow observations of GRB 170817A to the predictions made assuming three different jet models: A top-hat jet seen off-axis, a structured jet with a cocoon, and a mildly relativistic, isotropic fireball. The expected afterglow for a structured jet GRB is very different to the off-axis top-hat and the isotropic fireball. For a structured jet, the initial afterglow emission will be due material traveling down the line of sight. After a few days, this material will interact with the interstellar medium, causing it to decelerate and emit light. Over the next months, material ejected at an increasing large angle from the line of sight will become observable due to interactions with the interstellar medium, causing the afterglow to appear brighter. Eventually, on a timescale of months or years, the jet will become observable. At this point the afterglow will have reached peak luminosity and will start

to fade. This is qualitatively different to the top-hat and isotropic fireball case, in which all ejected material has approximately the same energy and so the afterglow rises more rapidly and fades more slowly, unlike the afterglow of a structured jet GRB which will rise slowly (see top right panel of figure 3.13). Note that if the jet is observed on axis, then the afterglow of a structured jet would be indistinguishable from a top-hat jet as the afterglow would peak quickly and then fade in both cases. In [cite afterglow paper] the authors use Markov Chain Monte Carlo to fit the structured jet, top-hat, and isotropic fireball model to the spectra of GRB 170817A. The 3 GHz light curve for the best fit of each model is plotted in figure 3.14. We can see that the structured jet model fits the data much better than either the top-hat or isotropic fireball model. The best fit model viewing angle is  $\theta = 33_{-2.5}^{+4}$ , which is comparable to the LIGO measurement of  $\leq 28^\circ$ . If this model of the structured jet is correct,<sup>17</sup> then approximately one in every 20 BNS systems detected with GWs should have a GRB counterpart [cite afterglow paper].

The structured jet model also has the advantage of being a natural consequence of binary neutron star mergers. As the jet shocks the slow moving material in the surrounding area, a cocoon of high pressure, subrelativistic matter will form. This cocoon creates a sheering force on the jet which creates a jet with a highly relativistic core, surrounded by lighter and slower moving material, with mildly relativistic wings at larger angles (see left panel of figure 3.13).

### 3.2.4 Other Scientific Results

We have so far focused on the physics of the jet and the progenitor of the GRB as these both effect our interpretation of the coincident GRB/GW signal. Now that we are confident that GRB 170817A resulted from a binary neutron star merger, we can use it as a probe for other physics.

#### Speed of Gravity

One result that can be derived immediately from the coincident detection is the speed of gravity  $v_{\text{GW}}$ . General relativity predicts that GWs will travel at the speed of light  $v_{\text{EM}}$ . We can now test this prediction by using the measured time delay between GW and GRB detection and different assumptions of time delay between the merger and GRB emission. The greater the distance to the GRB, the less the uncertainty in time delay. For this reason we will assume the GRB happened at a distance  $D = 26$  Mpc, the lower bound of the 90% credible distance determined from the GW signal. The delay between the peak of the GW signal and the start of the start of the GRB signal is  $1.74 \pm 0.05$  seconds. We can find an upper limit on the speed of gravity by assuming that the first photons were emitted at merger time.

---

<sup>17</sup>The best fit structured jet model has a  $\chi^2$  of 69 for 56 degrees of freedom and a probability of p=11%.

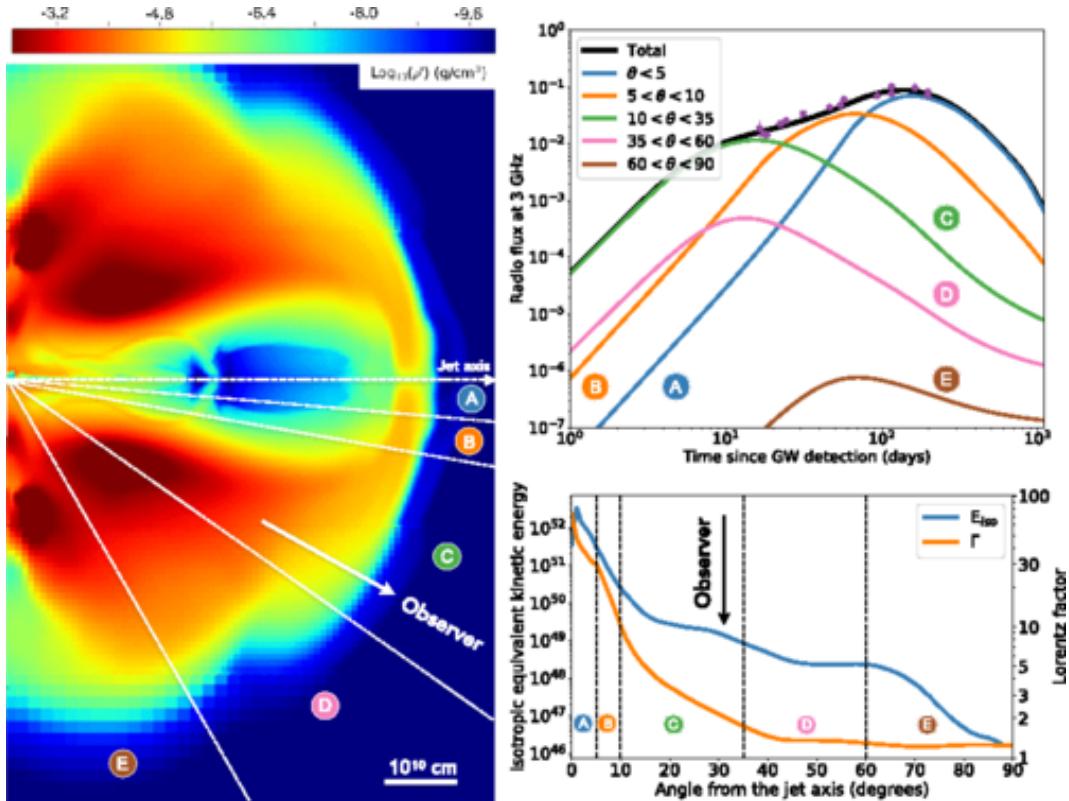


Figure 3.13: **Structured Jet.** Left panel: A pseudocolour density image of the simulation used to compute the afterglow curves [citation in afterglow paper]. The low density core of the jet is the blue region near the middle. The orange and green region around the core is the slow moving wings. Top right panel: Here we see the 3 GHz flux detected from different parts of the structured jet as time progresses. The angle is relative jet axis, so the blue curve is the core of the jet, the orange curve is the fast wings of the jet, the orange curve is the material moving along the line of sight (an angle of about  $33^\circ$  in this case), and the pink and brown curves correspond to large angles, that do not contribute much to the observed flux. Bottom right panel: The distribution of energy as a function of angular separation from the jet. cite GRB afterglow paper paper

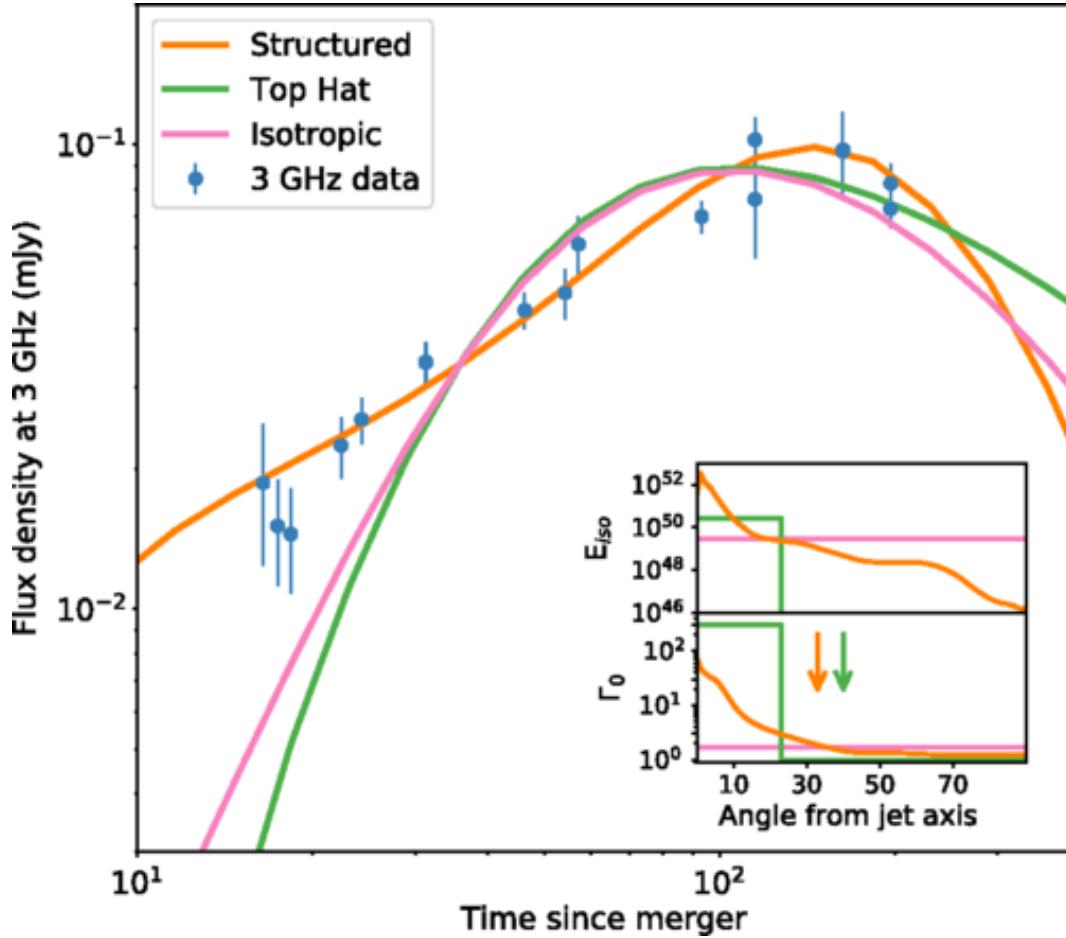


Figure 3.14: **Jet Model Comparison.** Here we see a comparison of the best fit for the structured jet, Top-hat jet seen off-axis, and isotropic models. The afterglow's measured flux density at 3 GHz is shown by the blue symbols (though the fits were performed with multiwavelength data). The inset shows the best fit isotropic energy and Lorentz factor for each model as a function of viewing angle. The arrows show the position of the observer for the structured and Top-hat jet models. [cite GRB afterglow paper](#)

This means the entire time delay is due to GWs traveling faster than light. Using the upper limit for the time delay, this gives us  $\delta t = 1.74 + 0.05$ . For a lower limit we can assume a significant delay between the merger and the first photons being emitted. It can be shown that the duration of a GRB is approximately the time delay between the merger and the GRB emission [cite GRB BNS paper]. It is therefore expected that the time delay between merger and emission is also  $\sim 2$  seconds. For a lower bound, we conservatively assume a time delay of  $\Delta t = 10 - 1.74 + 0.05 = 8.31$  seconds. As  $\Delta\nu/\nu_{\text{GW}} = -\nu_{\text{EM}}\Delta t/D$ , where  $\Delta\nu = \nu_{\text{GW}} - \nu_{\text{EM}}$ , we find

$$-3 \times 10^{-15} \leq \frac{\Delta\nu}{\nu_{\text{EM}}} \leq +7 \times 10^{-16}. \quad (3.1)$$

As this range includes zero, it is in agreement with general relativity. The largest source of uncertainty in this calculation is the time delay between GW and EM emission. As this does not depend on the distance to the source, more joint detections will allow this value to be constrained. This will allow for more accurate measurements of the speed of gravity and rule out more exotic EM emissions mechanisms, some of which predict time delays much greater than the 10 second bound we used in the above calculation.

## Hubble Constant

Another scientific result to come from GW 170817 is a new measurement of the Hubble constant. Using the distance inferred from GW data, and the redshift measurement of the associated galaxy NGC 4993, it is possible to calculate the Hubble constant directly, instead of using a cosmic distance ladder.<sup>18</sup> Using this method, the Hubble constant was measured to be  $H_0 = 70^{+12}_{-8} \text{ km s}^{-1}\text{Mpc}^{-1}$ . This is consistent with measurements made using both Planck data and standard candles.

## Rates

This detection can also be used to predict rates of BNS mergers and joint GW/GRB mergers. We will cover this in more detail in section 4.4, where we consider not only the detection of GRB 170817A and GW 170817, but also the 41 non-detections for other GRBs for which there was GW data available. There we will see that the expected rate of BNS mergers is 1-30 per year, with about 0.07-1.80 joint GRB-GW detections per year for the 2019-20 observing run.

---

<sup>18</sup>The cosmic distance ladder describes the fact that various different techniques are used to measure distances depending on the distance scale being measured. Each technique is useful for a certain range of distances, and the different techniques overlap in some distance range, allowing them to be calibrated to one another.

### 3.3 GRB Progenitors

In section 3.1 we discussed the evidence for short GRBs being produced by neutron star mergers and long GRBs being produced by core collapse supernova. In both of these cases, a large amount of energy is displaced rapidly from a small region. In section 3.1.5 we discussed the fireball model, which shows that these circumstances cause shocks and relativistic beaming which explain the high energy and variability of GRBs. But the fireball model only assumes that the *central engine*, the source of energy for the GRB, is small and highly energetic. There are many phenomena that satisfy this criteria. In this section we discuss some of these possibilities. We will see under which circumstances neutron star mergers and core collapse supernovae can power short and long GRBs respectively.

#### 3.3.1 Compact Binary Coalescence

As mentioned previously, short-hard GRBs are thought to be powered by the CBC involving at least one neutron star, with a neutron star or black hole as the other component of the binary. With the detection of GW170817 in conjunction with GRB170817A, we now know this is the case for at least some short-hard GRBs. In this section we will discuss the immediate aftermath of a NS merger and the different GRB and GW signals that each could be produced.

**NSBH Mergers** The simplest case is that of an NSBH merger where the BH mass is significantly greater than that of the NS. If the BH to NS mass ratio is greater than 5:1 then the innermost stable circular orbit (ISCO)<sup>19</sup> is greater than the tidal disruption radius<sup>20</sup>. This results in the NS being swallowed whole by the BH and leaving no accretion disk. The accretion disk is believed to be essential in powering a GRB after a NS merger, and so this case is not expected to produce a GRB. If the NSBH has a relatively low mass ratio, then the neutron star will be tidally disrupted before merging with the black hole. This will create a massive accretion disk and is expected to produce a GRB.

**BNS Mergers** For BNS mergers we must consider both the mass ratio and the total mass of the system. This is because it is possible for the merging neutron stars to either form a black hole or a *hypermassive neutron star* depending on the total mass of the binary system. The simplest of these is for a system with approximately equal mass ratio and a high total mass, such that a black hole can be produced immediately after merger. This requires the total mass to be above about  $2.9M_{\odot}$  [cite], but the exact value depends on the EOS of neutron stars. In this case, no accretion disk will be produced and so no GRB is expected.

---

<sup>19</sup>The ISCO is the smallest stable circular orbit for a test particle around a BH.

<sup>20</sup>The radius around a BH where the BH's tidal forces pull apart an in-falling star.

For an unequal mass ratio, the lighter neutron star is tidally disrupted by the larger neutron star. Matter from the lighter neutron star then accretes onto the more massive star, causing the more massive star to collapse into a black hole. This can potentially leave a massive accretion disk which could power a GRB.

If the total mass is less than about  $3M_{\odot}$  [cite] then the merger is not expected to immediately form a black hole but to instead form a hypermassive neutron star. This is a neutron star that is supported by differential rotation and thermal pressure. The merger will produce an accretion disk which could power a GRB. It is also possible that the rapidly rotating hypermassive neutron star will be ellipsoidal in shape, making a powerful emitter of GWs. These GWs would be detectable with aLIGO up to about 20Mpc [cite]. Over time, the hypermassive neutron star will lose angular momentum due to GW emission and magnetic force, its rotation will become more uniform due to magnetic forces and viscosity, and it will radiate away its heat. These factors cause the hypermassive neutron star to eventually collapse to a black hole. This can happen on a timescale as short as milliseconds [cite someone] up to seconds or minutes[citation in BNS GRB paper].

### 3.3.2 Core Collapse

Long GRBs are known to be caused by core collapse supernova. In this section we will discuss some of the possible central engines for long GRBs, focusing on those that could produce a GW signal strong enough to be detected with aLIGO.

**Rotational Instabilities of Proto-Neutron Stars** The cores of stars with initial masses in the range  $10M_{\odot} \lesssim M \lesssim 25M_{\odot}$  are expected to collapse to rapidly rotating *protoneutron stars*; stars which are cooling and contracting to form a neutron star. These protoneutron stars could exhibit non-axisymmetric deformations due to, for example, their rapid rotation driving hydrodynamic instabilities. The rapid rotation of the star could then drive significant GW emission, potentially detectable to 10 Mpc with current detectors and even further if the protoneutron star accretes supernova material.

**Non-axisymmetric Instabilities of Accretion Disks** For stars with initial mass greater than  $\sim 30M_{\odot}$ , the core collapse of the star will form a central black hole surrounded by an accretion disk. If the accretion disk has sufficiently high angular momentum and non-axisymmetric instabilities, then it can produce detectable GWs with a waveform similar to that of a low mass binary merger. For a stellar mass black hole with a clump of matter in its accretion disk of approximately  $0.1M_{\odot}$ , an accretion disk instability can potentially be detected out to 100Mpc with aLIGO [cite something]. An accretion disk instability can be caused by, for example, a high angular momentum such that the accretion disk is not gravitationally stable.

## 3.4 GRB Detectors

Most GRB detections have been come from the Swift and Fermi space telescopes, or a network of satellites called the Interplanetary Network. In this chapter we discuss the characteristics of these three missions, focusing on the aspects that affect a followup search for gravitational waves: sky coverage, source localisation, and sensitive energy range.

### 3.4.1 Swift

The Swift satellite is named for its ability to autonomously repoint itself towards a GRB within 90 seconds. It has three instruments: The first is the *Burst Alert Telescope* (BAT), the primary tool for GRB detection. It has a large field of view of approximately 2 steradians, and is highly sensitive in the 15-150 keV energy range. The BAT can localise GRBs to within 1-4 arcminutes. The other two instruments are the *X-ray Telescope* (XRT) and the *UV/Optical telescope* (UVOT). These are used for followup observations of the GRB afterglow. They can also reduce the sky-error; XRT can localise to 3-5 arcseconds and UVOT to 0.5 arcseconds **cite** [https://swift.gsfc.nasa.gov/about\\_swift/](https://swift.gsfc.nasa.gov/about_swift/). The UVOT is also used to determine the redshift of the host galaxy from which the GRB originated. To date, Swift has detected over 1000 GRBs, with approximately 10% of these having  $T_{90} < 2\text{s}$  **cite** [https://swift.gsfc.nasa.gov/archive/grb\\_table/stats/](https://swift.gsfc.nasa.gov/archive/grb_table/stats/)

### 3.4.2 Fermi

The Fermi Satellite [**cite** <https://www.nasa.gov/content/fermi/overview>] has two instruments: The first is the *Large Area Telescope* (LAT), and the second is the *Gamma-ray Burst Monitor* (GBM). The LAT is sensitive to higher energy photons (in the range 30 MeV-300 GeV), has sky coverage of 2 steradians, and can localise to within 1 arcminute of accuracy. The GBM is sensitive to lower energy photons (in the range 8keV-30MeV), is sensitive to a greater sky area than LAT (9.5 steradians), but cannot localise as well LAT (typically 10s or 100s of square degrees).

### 3.4.3 The InterPlanetary Network

The interplanetary network (IPN) [**cite something**] is a network of GRB detectors on spacecraft that are in low Earth orbit, eccentric Earth orbit, traveling to other planets, or orbiting other planets already. None of the spacecraft in the IPN are dedicated GRB detectors; they are individually unable to determine the sky position of a GRB and are not as sensitive as dedicated GRB detectors. But the network as a whole acts as an all-time all-sky GRB monitor. And as the spacecraft are great distances apart, the network can triangulate sky position to within several arcminutes. The accuracy is greatly dependent on the number and position of the satellites

that detect the GRB, with satellites at greater distances significantly improving the sky location error. A list of past and current IPN satellites can be found here [[cite](https://heasarc.gsfc.nasa.gov/W3Browse/all/ipngrb.html) <https://heasarc.gsfc.nasa.gov/W3Browse/all/ipngrb.html>].

## 3.5 GRB Gravitational Wave Search Strategies

In chapters 4 and 6, we discuss in detail two search pipelines for the CBC and burst search respectively. It is useful before discussing these searches to give an overview of the general strategy when searching for GWs associated with GRBs. This will give context to later discussions on the details of the searches, and also to the following section on the astrophysics that can be learned from detections of GWs from GRBs.

### 3.5.1 Triggered and Untriggered Searches

We can classify GW searches as being either *triggered* or *untriggered*. Untriggered searches will search all of the sky for all of the time when there is detector data. These all-sky all-time searches can be further divided into groups depending on the amount of time they take to run. Some searches, for example [[cite PyCBC live](#), [gstlal](#), [CWB papers](#)], are untriggered and low latency. They aim to find GW triggers and their sky position within minutes, allowing astronomers to followup the trigger. This is what happened for GW 170817, which was detected in low latency, in coincidence with a Fermi GRB, and followed up by ground and space based observatories. High latency untriggered searches, such as [[cite pycbc](#), [gstlal papers](#)], work on much longer timescales, as long as months, but aim to achieve very high sensitivity with very low false alarm rates. These pipelines are more sensitive and have lower false alarm rates than the low latency searches.

Triggered searches use sky location and time information from other messengers, such as GRBs. These searches have the advantage of only needing to analyse a limited amount of data (as the trigger time is known) and being able to use information gleaned from the other messengers to restrict the search. For example, both long and short GRBs are suspected to emit their jets along the axis of angular momentum, which can be used to infer than the GRB is circularly polarised (see section 4.2.5).

Results from both triggered and untriggered searches can be used in *Multimessenger* searches [[cite subthreshold paper](#)]. These searches attempt to combine subthreshold triggers from multiple different messengers to make a confident detection. For example, a core collapse supernova could produce a long GRB, a GW signal, and neutrinos, but be too distant for any search using just one of these messengers to make a confident detection.

### 3.5.2 Modeled and Burst Searches

Searches for GWs can also be classified as either *modeled* or *burst*. Modeled searches have theoretical models of the waveforms they are searching for. For example, CBC signals can be modeled by numerical relativity and analytic methods [**cite something**]. These waveforms can then be used to build a more sensitive search, such as a matched filter search (see sections **list matched filter sections**). In chapter 4 we will look at PyGRB, a targeted, modeled search for GWs associated with GRBs.

Burst searches use minimal assumptions about waveform morphology, instead relying on measures of coherence between detector data streams. This is in general less sensitive than if a waveform was known and a matched filter search could be carried out, but there are no such waveform models for supernovas/long GRBs. In chapter 6 we will discuss a burst search called X-pipeline.

## 3.6 GRB Astrophysics with Gravitational Waves

We have discussed what GRBs are, what causes GRBs, and seen that they are good emitters of GWs. We end this chapter with a discussion of how GW astronomy will add to our understanding of GRBs.

**Short GRB Progenitors** With the coincident detection of GRB 170817A and GW 170817, it is now known that at least some short GRBs are produced by neutron star mergers. As more detections are made it will become easier to determine whether this is the only source of short GRBs or if other mechanisms exist. Also, by comparing the electromagnetic counterpart to the knowledge gleaned from the GW signal about the binary, such as whether the system is an NSBH or a BNS, we can learn more about the central engine of the GRB.

**Long GRB Progenitors** There is a lot of evidence that at least some long GRBs are powered by core collapse supernova. If this is the case, then it should be possible to detect GWs in coincidence with a nearby long GRB. The GW signal would provide clues as to the evolution of the core collapse. If there is no GW detection, then it is possible to constrain the dynamics of the core collapse. Supernovae searches performed by LIGO at the moment only analyse supernovae that occur at a distance of less than 20 Mpc, and have so far seen no GW signal. This null result has been used to rule out some of the more extreme GW emission models for supernovae. [**cite O2 supernova paper**]

**Populations** There are many unknowns about the populations of short GRBs and compact binary systems. As a network of GW detectors is essentially sensitive to the whole sky, it will be possible to better understand the population of compact binary

systems within the well defined horizon of the detectors network. As GW signals also allow the direct inference of distance, it will even be possible to understand how the population of binary systems changes with distance up to the horizon of the detector network.<sup>21</sup> With the detection of GW 170817, the local BNS merger rate is measured to be  $1210^{+3230}_{-1040}$  [[cite catalog paper](#)]. This would correspond to a joint GRB/GW detection rate of between 0.07 and 1.80 events per year for the 2019/20 LIGO/Virgo observing run [[cite O2 GRB paper](#)].

**GRB Distance and Luminosity** For CBC GW signals where the inclination is approximately known, such as those associated with short GRBs, the distance to the source can be determined directly from the GW signal and without using a cosmological distance ladder. This allows independent verification of distances, as was done with GRB 170817A and GW 170817. Once the distance is known, the luminosity of the GRB can be reconstructed. This can also be used to as an alternative measure of Hubble's constant, again, as was done for GW 170817 [[cite paper](#)].

**Jet Structure** By comparing the rate of joint GW/GRB detections of BNS and NSBH mergers to the number detected through GWs alone, it will be possible to determine the opening angle of short GRBs. For the structured jet discussed in section 3.2, it is estimated that one in twenty BNS GW detections will have a GRB counterpart [[cite afterglow paper](#)]. For strong GW signals, it will also be possible to measure the polarisation of the GW. If the GW is face on, then it will be circularly polarised, but if it is edge on then it will be elliptically polarised. By measuring the polarisation of a large number of short GRBs and studying their prompt emission, it will be possible to fully describe the angular structure of short GRB jets.

**Neutron Star Equation of State** The equation of state (EoS) of matter at neutron star densities is poorly constrained. There are many ways that studying the GW signal of neutron star binaries can be used to determine the EoS. For example, a GW signal for a system with a neutron star of mass  $\sim 2M_{\odot}$  suggests a stiff EoS as a soft EoS could not support such a massive star. Another way of constraining the EoS of neutron stars is to look at the orbital frequency of a neutron star binary when tidal disruption occurs, as this depends on the radius of the neutron star. Using this radius with the mass of the neutron star (as determined by the inspiral waveform), we can determine the density of the neutron star. Tidal deformation can also affect the inspiral part of the waveform [[cite same paper as bartos paper](#)]. It has been shown [[cite Read et al](#)] that the advanced LIGO detectors can determine the radius of a neutron star to within  $\sim 1\text{km}$  for a source at 100 Mpc. It has also been theorised that the core of neutron stars may be made of strange quark matter.

---

<sup>21</sup>Due to the limited horizon of aLIGO and the low rate of neutron star mergers, this will have to wait for third generation GW detectors.

If this is true, then it will have a significantly different GW signal for both CBC systems [**cite same paper as bartos paper**] and neutron star instabilities after a core collapse supernova. Thus, GW astronomy can determine if neutron stars contain quark matter cores.

## Chapter 4

# A Targeted Search for Gravitational Waves associated with Short GRBs

### 4.1 Introduction

Compact binary coalescence (CBC) events are strong emitters of gravitational waves (GWs), and many searches exist to search for signals from these systems [34] [32] [28]. These searches take theoretical waveforms [23] and matched filter them against the strain data from GW detectors. It is known that the progenitor of at least some short GRBs are compact binaries, either binary neutron star (BNS) or low mass ratio neutron star - black hole (NSBH) systems. While the search pipelines mentioned above can detect the signals associated with short GRBs, they do not use any information from GRB detectors. By using this data, we can make a more sensitive search specifically designed to look for CBC signals associated with short GRBs. This is what PyGRB was developed to do.

There are several ways EM information can improve a GW search. The first is that the time and sky position of the short GRB are known from GRB detectors. This speeds up the analysis as we only use data around the GRB trigger time and only analyse a small patch of the sky. The sky position information can also be used to make a more powerful detection statistic, which we will see in section 4.2.2. We can also assume that the binary system that produced the GRB was face on, as GRBs are emitted perpendicular to the orbital plane. We will see in section 4.2.5 that this can also be used to improve the detection statistic.

In this chapter we describe PyGRB [19] [40], a targeted matched filter search for short GRBs. PyGRB is integrated into the PyCBC data analysis software [8] [7] [30] [15], which is primarily an all-time, all-sky, matched filter search for CBC signals. We will begin by discussing the theory behind the statistics used by PyGRB and show that it is more sensitive to GRB signals than the all-sky, all-time search.

In section 4.3 we will see how this theory is applied to make a functioning search pipeline. We end this chapter with a discussion on the findings of PyGRB in the second advanced LIGO and advanced Virgo observing run.

## 4.2 Coherent Matched Filtering

In this section we will discuss the theory behind PyGRB; a targeted, matched filter search for GWs associated with short GRBs. The detection statistic of PyGRB is based on the *coherent SNR*, which uses the sky position and time of the GRB with the antenna response function and PSD of each detector in the network. Using these, it is possible to calculate the amount of signal power expected in each detector and factor this into the detection statistic. For example, if a GRB was localised to a point in the sky directly above the most sensitive interferometer in the network, then this detector would be expected to have the highest SNR of all detectors in the network. If one of the other detectors in the network has a higher SNR, then it is less likely that this trigger is a GW. The coherent SNR provides a natural way to use the expected ratio of SNR in each detector into the detection statistic.

This qualitative description of coherence is made rigorous in the following sections 4.2.1 and 4.2.2. In section 4.2.3, we compare the coherent SNR to the *coincident SNR*, network detection statistic for the all-sky search. In section 4.2.4 we discuss the *null SNR*, another coherent statistic that measures the consistency of the signal across the detectors in the network. Short GRBs are expected to be emitted perpendicular to the orbital plane of the neutron star merger and with an opening angle of up to  $30^\circ$  (see sections 3.1.6 and 3.2.3). In section 4.2.5 we will see how this information can be used to build an even more sensitive detection statistic. We then end this section by discussing how to reject non-Gaussianities in the data by checking the consistency of the signal with the template that found the trigger.

### 4.2.1 Binary Coalescence Waveform

PyGRB searches for GWs from binaries with circular orbits and aligned spin components. These waveforms depend on 11 parameters: The two component masses, the component spins, the sky location  $(\theta, \phi)$ , the distance  $D$ , the coalescence time  $t_0$ , the inclination  $\iota$ , the polarisation angle  $\psi$ , and the coalescence phase  $\phi_0$ . We can reduce to nine parameters for the GRB search as the sky position is known.<sup>1</sup> Of the remaining parameters, the distance, binary inclination, polarisation, and coalescence phase affect the phase and amplitude of the waveform and not the signal morphology. This can be seen by writing the waveform can be in the following form

$$h_+(t) = \mathcal{A}^1 h_0(t) + \mathcal{A}^3 h_{\pi/2}(t) \quad (4.1)$$

---

<sup>1</sup>We will discuss how to deal with uncertainty in the sky position in section 4.3.4.

$$h_{\times}(t) = \mathcal{A}^2 h_0(t) + \mathcal{A}^4 h_{\pi/2}(t) \quad (4.2)$$

where  $h_0(t)$  and  $h_{\pi/2}(t)$  describe the two phases of the waveform, are usually assumed to be orthogonal, and depend only on the component masses and spins. The amplitudes  $\mathcal{A}^i$  are given by

$$\mathcal{A}^1 = \frac{D_0}{D} \left( \frac{(1 + \cos^2 \iota)}{2} \cos 2\phi_0 \cos 2\psi - \cos \iota \sin 2\phi_0 \sin 2\psi \right) \quad (4.3)$$

$$\mathcal{A}^2 = \frac{D_0}{D} \left( \frac{(1 + \cos^2 \iota)}{2} \cos 2\phi_0 \sin 2\psi + \cos \iota \sin 2\phi_0 \cos 2\psi \right) \quad (4.4)$$

$$\mathcal{A}^3 = -\frac{D_0}{D} \left( \frac{(1 + \cos^2 \iota)}{2} \sin 2\phi_0 \cos 2\psi + \cos \iota \cos 2\phi_0 \sin 2\psi \right) \quad (4.5)$$

$$\mathcal{A}^4 = \frac{D_0}{D} \left( -\frac{(1 + \cos^2 \iota)}{2} \sin 2\phi_0 \sin 2\psi + \cos \iota \cos 2\phi_0 \cos 2\psi \right) \quad (4.6)$$

where  $D_0$  is an arbitrary reference distance used to normalise the waveforms [19] [14]. For any  $\mathcal{A}^\mu$  we can invert equations (4.3)-(4.6) to obtain the physical parameters up to a reflection symmetry of the system.

The response of GW detector  $X$  to a GW  $h_{+, \times}$  is given by

$$h^X(t) = F_+(\theta^X, \phi^X, \chi^X) h_+(t^X) + F_\times(\theta^X, \phi^X, \chi^X) h_\times(t^X) \quad (4.7)$$

where  $F_{+, \times}^X$  is the antenna response of detector  $X$  to the plus and cross polarisation of the GW, the angles  $\theta^X$  and  $\phi^X$  give the sky position of the source relative to the detector, the polarisation angle between the detector frame to the radiation frame is given by  $\chi^X$ , the coalescence time at the Earth's centre is  $t$ , and the coalescence time in detector  $X$  is given by  $t^X$ . Combining 4.7 with 4.1 and 4.2 we can write the detector response in terms of the  $A^\mu$  as

$$h^X(t) = \mathcal{A}^\mu h_\mu^X(t) \quad (4.8)$$

where the  $h_\mu^X$  are given by

$$h_1^X = F_+^X h_0(t^X) \quad (4.9)$$

$$h_2^X = F_\times^X h_0(t^X) \quad (4.10)$$

$$h_3^X = F_+^X h_{\pi/2}(t^X) \quad (4.11)$$

$$h_4^X = F_\times^X h_{\pi/2}(t^X) . \quad (4.12)$$

### 4.2.2 Coherent SNR

In this section we will derive the *coherent SNR*, a detection statistic for a multidector matched filter search.<sup>2</sup> We begin the derivation by finding the likelihood that a trigger is a GW in terms of  $\mathcal{A}^\mu$  (assuming the detector noise is Gaussian). We then maximise the likelihood with respect to  $\mathcal{A}^\mu$ . This reduces the parameter space from nine dimensions to five, reduces the noise background, and improves the detection statistic. It also allows us to speed up the analysis by reducing the parameter space that needs to be searched over, meaning we can use a smaller template bank.

We begin by describing the output  $s^X(t)$  of detector  $X$ . The detector data is the sum of the antenna response to a GW, given by (4.7), with the noise in the detector  $n^X(t)$

$$s^X(t) = n^X(t) + h^X(t) . \quad (4.13)$$

The noise power spectral density (PSD)  $S_h^X$  for detector  $X$  is defined by

$$\langle \tilde{n}^X(f)[\tilde{n}^X(f')]^* \rangle = \delta(f - f') S_h^X(f) \quad (4.14)$$

where the angle brackets denote the time average of the noise and tildes indicate that the function has been Fourier transformed. The matched filter between the GW waveform<sup>3</sup>  $h$  and detector data is given by the inner product

$$(s^X|h) = 4\text{Re} \int_0^\infty \frac{\tilde{s}^X(f) \cdot [\tilde{h}(f)]^*}{S_h^X(f)} e^{2\pi ift} df . \quad (4.15)$$

The probability of obtaining detector data  $s^X$  given the presence of GW  $h$  is denoted  $P(s^X|h)$ . This is equivalent to the probability of a sample of detector noise being equal to the measured detector data minus the detector response to  $h$ , given by (4.7). If the detector noise is Gaussian, then the probability is

$$P(s^X|h) = \frac{1}{2\pi} e^{-(s^X-h^X|s^X-h^X)/2} . \quad (4.16)$$

The likelihood ratio is the probability of obtaining detector data  $s^X$  when the GW signal  $h$  is present, divided by the probability of obtaining the data  $s^X$  in the absence of a GW, which we denote  $P(s^X|0)$ . This gives us

$$\Lambda(h) = \frac{P(s^X|h)}{P(s^X|0)} = \frac{e^{-(s^X-h^X|s^X-h^X)/2}}{e^{-(s^X|s^X)/2}} . \quad (4.17)$$

---

<sup>2</sup>Note that this is not the final detection statistic for PyGRB, which is a modified version of the coherent SNR. See sections 4.3.2 and 4.2.5.

<sup>3</sup>Careful with notation here. The complex GW waveform is denoted  $h$ , and the detector response to the GW is denoted by  $h^X$ .

For convenience, we will use the log-likelihood

$$\log \Lambda = (s^X | h^X) - \frac{1}{2}(h^X | h^X) . \quad (4.18)$$

Using 4.8, we can rewrite the log-likelihood in terms of  $\mathcal{A}^\mu$

$$\log \Lambda = \mathcal{A}^\mu(s^X | h_\mu^X) + \frac{1}{2}\mathcal{A}^\mu(h_\mu^X | h_\nu^X)\mathcal{A}^\nu . \quad (4.19)$$

The log-likelihood defined above is a measure of the probability a trigger that has been found in a single detector is a GW. For a coherent search we need a likelihood measure that takes into consideration every detector in the network. To do this, we must first define the multidetector inner product, we can then follow the same steps as in the single detector case. Assuming that the noise in each detector is uncorrelated, the probability of obtaining detector data  $s^X$  in detector  $X$  and  $s^Y$  in detector  $Y$  will be the product of measuring  $s^X$  and  $s^Y$  independently. From equation 4.16, we see that this is equivalent to summing the inner products from each detector. We therefore define the multidetector inner product to be the sum of individual detector inner products for the  $d$  detectors in the network

$$(\mathbf{a}|\mathbf{b}) = \sum_{X=1}^d (a^X | b^X) . \quad (4.20)$$

The multidetector log-likelihood then becomes

$$\log \Lambda = (\mathbf{s}|\mathbf{h}) - \frac{1}{2}(\mathbf{h}|\mathbf{h}) \quad (4.21)$$

where  $\mathbf{h}$  is the vector of detector responses to the GW  $h^X$  and  $\mathbf{s}$  is the vector of detector outputs  $s^X$ . In terms of  $\mathcal{A}^\mu$ , the multidetector log-likelihood function is

$$\ln \Lambda = \left[ \mathcal{A}^\mu(\mathbf{s}|\mathbf{h}_\mu) - \frac{1}{2}\mathcal{A}^\mu\mathcal{M}_{\mu\nu}\mathcal{A}^\nu \right] \quad (4.22)$$

where  $\mathcal{M}_{\mu\nu}$  is the symmetric matrix given by

$$\mathcal{M}_{\mu\nu} = (\mathbf{h}_\mu | \mathbf{h}_\nu) . \quad (4.23)$$

We want to find the template that maximises the log-likelihood. The values for  $\mathcal{A}^\mu$  for which the log-likelihood 4.22 is maximal are given by

$$\hat{\mathcal{A}}^\mu = \mathcal{M}^{\mu\nu}(\mathbf{s}|\mathbf{h}_\nu) \quad (4.24)$$

where  $\mathcal{M}^{\mu\nu}$  is the inverse of  $\mathcal{M}_{\mu\nu}$ . The *coherent SNR*  $\rho_{\text{coh}}$  is then defined by

$$\rho_{\text{coh}}^2 = 2 \ln \Lambda|_{\max} = (\mathbf{s}|\mathbf{h}_\mu)\mathcal{M}^{\mu\nu}(\mathbf{s}|\mathbf{h}_\nu) . \quad (4.25)$$

We can simplify the matrix  $\mathcal{M}$ , and make it easily invertible, by noting that as CBC signals spend a large number of cycles in the sensitive frequency range of the detector, the 0 and  $\frac{\pi}{2}$  phases of the waveform are approximately orthogonal. The slow frequency evolution means that the two phases have roughly equal amplitude. Hence we find

$$(h_0^X | h_{\pi/2}^X) \approx 0, \quad (4.26)$$

$$(h_{\pi/2}^X | h_{\pi/2}^X) \approx (h_0^X | h_0^X) = (\sigma^X)^2. \quad (4.27)$$

Using these, we see that  $\mathcal{M}$  simplifies to

$$\mathcal{M} = \begin{bmatrix} A & C & 0 & 0 \\ C & B & 0 & 0 \\ 0 & 0 & A & C \\ 0 & 0 & C & B \end{bmatrix}$$

with

$$A = \sum_X (\sigma^X F_+^X)^2 \quad (4.28)$$

$$B = \sum_X (\sigma^X F_\times^X)^2 \quad (4.29)$$

$$C = \sum_X (\sigma^X F_+^X)(\sigma^X F_\times^X). \quad (4.30)$$

With the  $\mathcal{A}^\mu$  terms maximised over, we have reduced the number waveform parameters from nine to five, leaving only the masses, spins, and coalescence time to be searched over.

### 4.2.3 Comparison to Coincident Search

In this section we will put the coherent SNR into a form such that it is more easily compared to the *coincident SNR*, the detection statistic currently used by all-sky matched filter searches. The coincident SNR is defined to be the quadrature sum of the matched filter of the individual detectors

$$\rho_{\text{coinc}}^2 = \sum_{X,Y} \sum_{i=0,\pi/2} \left( s^X \left| \frac{h_i}{\sigma^X} \right. \right) [\delta^{XY}] \left( s^Y \left| \frac{h_i}{\sigma^Y} \right. \right). \quad (4.31)$$

We can make 4.25 more easily comparable to 4.31 by making  $\mathcal{M}$  diagonal. We do this by rotating the detector frame to the *Dominant Polarization Frame*. In this frame, the plus and cross antenna response functions for each detector are orthogonal, making  $C = 0$  and  $\mathcal{M}$  diagonal. As we included polarisation angles between the equatorial and radiation frames  $\chi$  in 4.7 and the radiation and source frames  $\psi$  in  $\mathcal{A}^\mu$ , we can rotate our network frame without placing further constraints on the system. In particular, we can rotate through an angle  $\chi^{\text{DP}}$  such that  $C = 0$ .

This gives us the following antenna response functions

$$F_+^{\text{DP},X} = F_+^X \cos 2\chi^{\text{DP}} + F_\times^X \sin 2\chi^{\text{DP}} \quad (4.32)$$

$$F_\times^{\text{DP},X} = -F_+^X \sin 2\chi^{\text{DP}} + F_\times^X \cos 2\chi^{\text{DP}}. \quad (4.33)$$

Using these antenna response functions in (4.30) and solving for  $\chi^{\text{DP}}$ , we find

$$\chi^{\text{DP}} = \frac{1}{4} \arctan \left( \frac{2 \sum_X (\sigma^X F_+^X)(\sigma^X F_\times^X)}{\sum_X [(\sigma^X F_+^X)^2 - (\sigma^X F_\times^X)^2]} \right). \quad (4.34)$$

This does not uniquely define the dominant polarisation frame, so we also require the network to be more sensitive to the plus polarisation than the cross polarisation

$$|F_+^{\text{DP},X}| \geq |F_\times^{\text{DP},X}|. \quad (4.35)$$

In what follows, we assume we are in the dominant polarisation frame and drop the DP superscript. Inverting  $\mathcal{M}$  and using (4.25), we see the coherent SNR in the dominant polarisation frame is

$$\rho_{\text{coh}}^2 = \frac{(\mathbf{s}|\mathbf{F}_+ \mathbf{h}_0)^2 + (\mathbf{s}|\mathbf{F}_+ \mathbf{h}_{\pi/2})^2}{(\mathbf{F}_+ \mathbf{h}_0 |\mathbf{F}_+ \mathbf{h}_0)^2} + \frac{(\mathbf{s}|\mathbf{F}_\times \mathbf{h}_0)^2 + (\mathbf{s}|\mathbf{F}_\times \mathbf{h}_{\pi/2})^2}{(\mathbf{F}_\times \mathbf{h}_0 |\mathbf{F}_\times \mathbf{h}_0)^2}. \quad (4.36)$$

We can rewrite this as

$$\rho_{\text{coh}}^2 = \sum_{X,Y} \sum_{i=0,\pi/2} \left( s^X \left| \frac{h_i}{\sigma^X} \right. \right) [f_+^X f_+^Y + f_\times^X f_\times^Y] \left( s^Y \left| \frac{h_i}{\sigma^Y} \right. \right) \quad (4.37)$$

where we define the orthonormal vectors

$$f_{+, \times}^X = \frac{\sigma^X F_{+, \times}^X}{\sqrt{\sum_Y (\sigma^Y F_{+, \times}^Y)^2}}. \quad (4.38)$$

We can now compare the coherent SNR as given in (4.37) to the coincident SNR given by (4.31). The coincident SNR is the quadrature sum of all the energy in each detector. In the space spanned by individual detector SNRs, where each trigger is represented by a point in this space, it is the distance from the origin to the trigger. It is the total energy of that trigger.

The coherent SNR is a projection of the total energy of a trigger onto the subspace spanned by  $f_+$  and  $f_\times$ . This subspace is the space consistent with a gravitational wave signal from the given sky location and with the PSD of the detectors at the given time, so we call it the *signal space*. Orthogonal to this space is the space inconsistent with a signal. Noise contributes energy to all components of the coincident SNR, and so projecting it onto the signal space will reduce its magnitude. The energy due to a GW signal lays entirely in the signal plane, so in the absence of

noise the projection will not change the magnitude of that trigger at all. Thus, in the case of a GW detection, the signal contribution to the coherent SNR and coincident SNR is the same, but the noise contribution to the coherent SNR is reduced. Thus, for a GW detection we expect  $\rho_{\text{coh}} \lesssim \rho_{\text{coinc}}$ .

It is also interesting think of the coherent and coincident SNR in terms of the number of degrees of freedom of the noise contribution. The coincident SNR has noise contributions from the phase and amplitude measurements in each detector, resulting in  $2N$  degrees of freedom, where  $N$  is the number of detectors. From (4.36) we can see that the coherent SNR has four degrees of freedom: the  $0$  and  $\pi/2$  phases of the plus and cross polarisation amplitudes of the gravitational wave. In the case where we have a non-degenerate (i.e. sensitive to both plus and cross polarisations) two detector network, both the coincident and coherent SNR have four degrees of freedom. In this case, the coherent and coincident SNRs are identical as

$$f_+^X f_+^Y + f_\times^X f_\times^Y = \delta^{XY} . \quad (4.39)$$

If we increase the number of detectors in the network, then the number of degrees of freedom of the coincident SNR will also increase, but remains constant for the coherent SNR. This gives the coherent analysis a sensitivity advantage.

#### 4.2.4 Null SNR

Gravitational waves have two polarisations. Therefore we can fully describe the signal with two non-aligned detectors. Adding additional detectors to our network allows us to construct additional data streams that have the GW signal removed. We can easily construct such a data stream by subtracting the coherent SNR (the energy consistent with a GW signal) from the coincident SNR (the total energy measured by the detectors). Any energy left from this is associated with noise, and we call this the *null SNR*<sup>4</sup>

$$\rho_N^2 = \rho_{\text{coinc}}^2 - \rho_{\text{coh}}^2 = \sum_{X,Y} \sum_{i=0,\pi/2} \left( s^X \left| \frac{h_i}{\sigma^X} \right. \right) [N^{XY}] \left( s^Y \left| \frac{h_i}{\sigma^Y} \right. \right) \quad (4.40)$$

where

$$N^{XY} = \delta^{XY} - f_+^X f_+^Y - f_\times^X f_\times^Y . \quad (4.41)$$

Glitches are typically uncorrelated between detectors and so not confined to the signal space. For this reason, we expect  $\rho_{\text{coinc}}^2 \gg \rho_{\text{coh}}^2$  for loud glitches and thus the null SNR is high. For a GW signal, most of the energy will lay in the signal plane. Therefore  $\rho_{\text{coinc}}^2 \simeq \rho_{\text{coh}}^2$  and the null SNR will be close to zero. Thus, applying a threshold on the null SNR gives us an additional background noise rejection test.

---

<sup>4</sup>The null SNR is similar to *null streams*, which are used by burst searches (see section 6.1). The null stream in a burst search is made to remove a GW signal of any morphology, unlike the null SNR which is template specific.

### 4.2.5 Searching for Face on Signals

Short GRB jets are emitted perpendicular to the orbital plane and with an opening angle  $< 30^\circ$ , as discussed in section 3.2. In a similar way to how the coherent SNR defined in section 4.2.2 accounts for the sky position of the GRB and the PSD of each detector, we can construct a new coherent SNR that takes into account that GWs associated with GRBs have an orbital inclination of  $\iota \sim 0$  or  $\iota \sim \pi$  with respect to the observer.

We can see from equations (4.3) to (4.6) that the GW amplitude terms depend linearly on  $\cos \iota$  and  $(1 + \cos^2 \iota)/2$ . These two terms are equal to within 1% for  $0^\circ \leq \iota < 30^\circ$ . Thus, for  $\iota \sim 0$ , we can use the approximation  $\cos \iota \approx (1 + \cos^2 \iota)/2$ . For  $\iota \sim \pi$  the two terms are approximately equal in magnitude, but with opposite signs.

Using the above approximation and defining

$$\tilde{D} = \frac{D}{\cos \iota} \text{ and } \chi_{l,r} = \phi_0 \pm \psi, \quad (4.42)$$

the GW amplitudes (4.3) to (4.6) with  $\iota \sim 0$  become

$$\mathcal{A}^1 \approx \mathcal{A}^4 \approx -\frac{D_0}{\tilde{D}} \cos 2\chi_l \equiv \mathcal{B}_1 \quad (4.43)$$

$$\mathcal{A}^2 \approx -\mathcal{A}^3 \approx \frac{D_0}{\tilde{D}} \sin 2\chi_l \equiv \mathcal{B}_2. \quad (4.44)$$

We see that now there are only two amplitude factors,  $\mathcal{B}_1$  and  $\mathcal{B}_2$ , and that the GW is circularly polarised. Similar results can be derived for the  $\iota \sim \pi$  case.

Substituting these into the equation for the log-likelihood (4.22) and working in the dominant polarisation frame<sup>5</sup>, we find

$$\ln \Lambda = \mathcal{B}_1(\mathbf{s}|\mathbf{F}_+ \mathbf{h}_0 + \mathbf{F}_\times \mathbf{h}_{\pi/2}) + \mathcal{B}_2(\mathbf{s}|\mathbf{F}_\times \mathbf{h}_0 - \mathbf{F}_+ \mathbf{h}_{\pi/2}) \quad (4.45)$$

$$-\frac{1}{2}[\mathcal{B}_1^2 + \mathcal{B}_2^2][(\mathbf{F}_+ \mathbf{h}_0|\mathbf{F}_+ \mathbf{h}_0) + (\mathbf{F}_\times \mathbf{h}_0|\mathbf{F}_\times \mathbf{h}_0)]. \quad (4.46)$$

We now want to maximise the log-likelihood, just as we did when deriving the coherent SNR in section 4.2.2. Defining

$$\alpha = (\mathbf{s}|\mathbf{F}_+ \mathbf{h}_0 + \mathbf{F}_\times \mathbf{h}_{\pi/2}) \quad (4.47)$$

$$\beta = (\mathbf{s}|\mathbf{F}_\times \mathbf{h}_0 - \mathbf{F}_+ \mathbf{h}_{\pi/2}) \quad (4.48)$$

and maximising the log-likelihood with respect to  $\mathcal{B}_1$  and  $\mathcal{B}_2$ , we find

$$\rho_{\text{coh}}^2 = 2 \ln \Lambda_{\text{max}} = \frac{\alpha^2 + \beta^2}{(\mathbf{F}_+ \mathbf{h}_0|\mathbf{F}_+ \mathbf{h}_0) + (\mathbf{F}_\times \mathbf{h}_0|\mathbf{F}_\times \mathbf{h}_0)}. \quad (4.49)$$

---

<sup>5</sup>Recall that  $(h_{\pi/2}^X|h_{\pi/2}^X) \approx (h_0^X|h_0^X)$  for CBC signals, see section 4.2.2.

In the  $\iota \sim \pi$  case, the coherent SNR takes the same form, but with

$$\alpha = (\mathbf{s}|\mathbf{F}_+ \mathbf{h}_0 - \mathbf{F}_\times \mathbf{h}_{\pi/2}) \quad (4.50)$$

$$\beta = (\mathbf{s}|\mathbf{F}_\times \mathbf{h}_0 + \mathbf{F}_+ \mathbf{h}_{\pi/2}) . \quad (4.51)$$

The circularly polarised coherent SNR allows for a 3% increase in range for a given FAP when compared to the generic coherent SNR [40]. This corresponds to a 10% increase in the rate of GW signals detected. Also, as the circularly polarised coherent SNR only has two degrees of freedom, we can construct an additional null statistic and further improve background noise rejection. As well as this, we can now calculate a null SNR even when analysing data from just two detectors.

#### 4.2.6 Coherent $\chi^2$ Tests

There are many instrumental and environmental sources of non-stationary, non-Gaussian noise in GW detectors. These frequently take the form of a large SNR spike in the data, which we call a *glitch*. The coherent statistics discussed in the preceding sections will downweight these glitches, but loud glitches will still have a high coherent SNR and so can appear to be a GW. We need a way to distinguish between glitches and GWs. We do this using *signal consistency tests*, which checks that triggers match the waveform of the template. In this section we introduce three signal consistency tests used by PyGRB.

#### Signal Consistency Tests

We can describe the detector output as the sum of a part that is proportional to the template  $h(t)$ , a part  $g(t)$  that is orthogonal to a GW template, and Gaussian noise  $n(t)$ , so that

$$s(t) = n(t) + Ah(t) + Bg(t) \quad (4.52)$$

where  $A$  and  $B$  are amplitude factors and we have normalised the the  $h$  and  $g$  components, so that

$$(g|h) = 0, \quad (g|g) = 1, \quad (h|h) = 1 . \quad (4.53)$$

Matched filtering this data will return a high SNR if  $B$  is large, regardless of whether there is a GW signal in the data or not. To distinguish between a glitch and a signal, note that the large SNR due to a glitch is *not* because it matches the template well, but because it has an intrinsically larger amplitude. This means that loud glitches should ring off against a wide variety of different templates in the template bank, unlike a true GW signal where we would expect the SNR to be high only for those templates that have similar morphology to the GW signal.

Motivated by this, we introduce a set of  $N$  templates  $T^i$  that are orthonormal to the template in question

$$(h|T^i) = 0, \quad (T^i|T^j) = \delta^{ij} . \quad (4.54)$$

We now sum the squares of the matched filter of the data  $s$  with the templates  $T^i$ . When the noise is stationary (i.e. when  $B = 0$ ), the orthonormality condition means that

$$\chi^2 \equiv \sum_{i=1}^N (T^i|s)^2 = \sum_{i=1}^N (T^i|n)^2 \quad (4.55)$$

whether a GW signal is present or not. Each term in the summation is the square of a Gaussian distributed with a mean of zero and variance of one, making this  $\chi^2$  distributed with  $N$  degrees of freedom. The mean and variance are therefore

$$\langle \chi^2 \rangle = N, \quad \text{Var}(\chi^2) = 2N . \quad (4.56)$$

Now suppose that there is a glitch in the data (i.e.  $B \neq 0$ ). In this case, we find

$$\chi^2 = \sum_{i=1}^N [(T^i|n) + B(T^i|g)]^2 . \quad (4.57)$$

This follows a non-central  $\chi^2$  distribution with  $N$  degrees of freedom and non-centrality parameter  $\lambda = B^2(T^i|g)^2$ . This has mean

$$\langle \chi^2 \rangle = N + B^2 \sum_{i=1}^N (T^i|g)^2 \quad (4.58)$$

and variance

$$\text{Var}(\chi^2) = 2N + 4B^2 \sum_{i=1}^N (T^i|g)^2 . \quad (4.59)$$

Now we can see how to use  $\chi^2$  to distinguish GW signals from glitches. Calculate  $\chi^2$  and divide by the number of degrees of freedom. For a GW signal this should be close to one, but for a glitch it will be greater than one by an amount that is proportional to the square of the amplitude of the glitch. Similarly, the variance of  $\chi^2$  per degree of freedom for a GW signal will be approximately two, whereas for a glitch it will scale with the square of the glitch amplitude. Thus we can apply a threshold on the value of  $\chi^2$  per degree of freedom, with triggers exceeding this value being considered glitches.

It is not trivial to find a set of orthonormal waveforms such as the  $T^i$  above. A set of waveforms  $t^i$  can be made into a set of waveforms orthogonal to  $h$  using the formula

$$T^i = \frac{t^i - (t^i|h)h}{\sqrt{1 - (t^i|h)^2}} \quad (4.60)$$

but in general these  $T^i$  will not be orthonormal to each other. Using these waveforms, the formula for  $\chi^2$  is no longer a sum of independent Gaussian variables but a sum of correlated Gaussian variables. For this distribution the mean is unchanged, but the variance becomes<sup>6</sup>

$$\text{Var}(\chi^2) = 2N + 2 \sum_{i \neq j} (T^i | T^j)^2 . \quad (4.61)$$

This does not cause significant problems though as  $\chi^2$  thresholds are set empirically.<sup>7</sup>

We can extend the idea of the  $\chi^2$  test to a multidetector search. We start with the four waveform components of  $\mathbf{h}_\mu$ , as described in section 4.2.2. We saw in section 4.2.3 that in the dominant polarisation frame, these components satisfy the equation

$$(\mathbf{h}_\mu | \mathbf{h}_\nu) = \mathcal{M}_{\mu\nu} = \text{diag}(A, B, A, B) \quad (4.62)$$

and as such are orthogonal, but not normalised. We can normalise them to obtain

$$(\hat{\mathbf{h}}_\mu | \hat{\mathbf{h}}_\nu) = \delta_{\mu\nu} . \quad (4.63)$$

We then take a set of normalised test waveforms  $\hat{t}_\mu^i$  and construct waveforms that are orthogonal to  $\hat{\mathbf{h}}_\mu$  using

$$T_\mu^i = \frac{\hat{t}_\mu^i - \sum_\nu (\hat{\mathbf{t}}_\mu^i | \hat{\mathbf{h}}_\nu) \hat{h}_\nu}{\sqrt{1 - \sum_\sigma (\hat{\mathbf{t}}_\sigma^i | \hat{\mathbf{h}}_\sigma)^2}} . \quad (4.64)$$

The multidetector  $\chi^2$  is then defined as

$$\chi^2 = \sum_{\mu=1}^4 \sum_{i=1}^N (\mathbf{T}_\mu^i | \mathbf{s})^2 \quad (4.65)$$

if we assume the templates used are orthonormal

$$(\mathbf{T}_\mu^i | \mathbf{T}_\nu^j) = \delta^{ij} \delta_{\mu\nu} . \quad (4.66)$$

This has a mean of  $4N$  and a variance of  $8N$ . The orthonormality condition is not in general true. Accounting for this, the mean is still  $4N$  but the variance becomes

$$\text{Var}(\chi^2) = 8N + 2 \sum_{i \neq j} \sum_{\mu \neq \nu} [(\mathbf{T}_\mu^i | \mathbf{T}_\nu^j)^2] \quad (4.67)$$

We have seen how  $\chi^2$  tests are expected to work, but there are several different

---

<sup>6</sup>The general formula for the variance of the sum of correlated Gaussian variables  $X_i$  is  $\text{Var}(\sum_i X_i) = \sum_i \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j)$  where  $\text{Cov}(X_i, X_j)$  is the covariance matrix of  $X_i$  and  $X_j$ .

<sup>7</sup>In fact, we could generate orthonormal waveforms, but it is not worth the computational cost.

methods for obtaining the waveforms  $\hat{\mathbf{t}}_\mu^i$ . In the rest of this section we will see the three different kinds of  $\chi^2$  test that PyGRB uses.

### The Coherent Bank $\chi^2$ Test

A glitch will cause a high SNR across many templates in a template bank. This is different to a GW, which will have a high SNR only for templates with similar parameters to the signal. By selecting a set of templates from across the template bank to be our test waveforms  $t^i$ , we can use the  $\chi^2$  test described above. The set of waveforms remains the same for every template  $h$  in the template bank. This test is most effective when  $\mathbf{T}_\mu^i$  are close to orthogonal, so templates are selected from across the mass space. This is sufficient to make the deviation from a true  $\chi^2$  test negligible. The single detector version of this tests was first described in [18].

### The Coherent Autocorrelation $\chi^2$ Test

Alternatively the  $t^i$  can be selected by applying timeshifts to the original template  $h$  [18]. The problem with this test is that the  $t^i$  are in general far from orthonormal, leading to a distribution that does not closely resemble a  $\chi^2$  distribution. As thresholds are set empirically, this can still be used as a consistency check, but it has a large variance.

### The Coherent Power $\chi^2$ Test

The power  $\chi^2$  test divides the triggering template  $h$  into  $N$  non-overlapping frequency windows such that each window has the same expected SNR [6]. This is then compared to the SNR in each window as measured from the data. For a real signal, each window would have an SNR  $\rho/N$ , where  $\rho$  is the SNR for the whole template. The expected SNR  $\rho^i$  is subtracted from the measured SNR in each window. In the presence of purely Gaussian noise this will be Gaussian distributed with mean zero and unit variance. Thus, we can obtain a  $\chi^2$  distributed variable with  $(N - 1)$  degrees of freedom as follows

$$\chi^2 = N \sum_{i=1}^N (\rho^i - \rho/N)^2 . \quad (4.68)$$

The coherent power  $\chi^2$  test is a natural extension of this. We define the SNR contribution to the  $i$ -th window to be

$$\rho_\mu^i = \frac{(\mathbf{s}|\mathbf{h}_\mu^i)}{\sqrt{(\mathbf{h}_\mu|\mathbf{h}_\mu)}} . \quad (4.69)$$

The  $\chi^2$  test is then

$$\chi^2 = N \sum_{i=1}^N \sum_{\mu=1}^4 (\rho_{\mu}^i - \rho_{\mu}/N)^2. \quad (4.70)$$

In Gaussian noise this will be the  $\chi^2$  distributed with  $4(N - 1)$  degrees of freedom.

## 4.3 The PyGRB Workflow

Now we have seen the theory behind a coherent matched filtering search, we will use what we have learned to make a targeted GW search pipeline. We will see how PyGRB combines the coherent SNR, null SNR, and coherent  $\chi^2$  values to make a detection statistic and how this detection statistic is used to estimate the false alarm probability (FAP) of triggers. We will see how PyGRB deals with uncertainty in the sky position of a GRB. In section 4.3.5 we will see how PyGRB estimates its sensitivity, which is important for providing upper limits on the rates of neutron star mergers. We end this section by bringing all these parts together and looking at the full workflow.

### 4.3.1 Background Noise Rejection

We have seen several statistics that can indicate whether a trigger is a GW signal or noise. We apply cuts on these statistics, discarding any triggers that fail the cuts. We can also speed up the analysis by discarding triggers before calculating computationally expensive statistics, such as the coherent SNR and the signal consistency tests. In this section we will discuss the PyGRB workflow as it was in the most recent LIGO observing run O2, focusing on how triggers are discarded in order to speed up the analysis.

The analysis begins with matched filtering the individual detector data. Using the known sky location of the GRB and the fact that gravitational waves travel at the speed of light, we timeshift the data from each detector so that a GW associated with the GRB will be coincident in each detector. A list of triggers is formed for each detector by keeping only times when the individual detector SNR is greater than four, the rest of the data is discarded. We then discard any trigger that is not coincident in at least two detectors in the network.

The coincident SNR is then computationally cheap to calculate, and any trigger with a coincident SNR below six is discarded. We then calculate the coherent SNR for the remaining triggers and apply the same threshold of six to the coherent SNR. Once we have the coincident and coherent SNR, the null SNR is cheap to calculate. We cut triggers with a high null SNR, as it is in general higher for glitches than for GWs. However, differences between the actual GW waveform and the template used for the matched filter, as well as inaccuracies in timing and sky position, can all lead to some fraction of the energy of a GW contributing to the null SNR. In

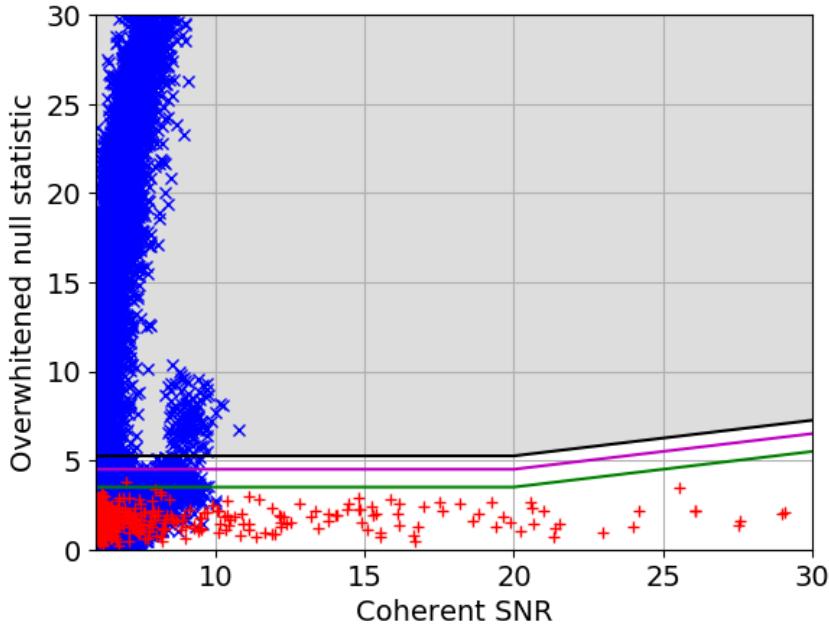


Figure 4.1: **Null Statistic Cut.** Here we plot the coherent SNR against the null SNR. The blue crosses are background triggers. The red pluses are signal injections. The black line is the veto line, with all triggers in the shaded region above the line being discarded. The green line indicates the expected SNR for optimally oriented injections. The magenta line shows the one sigma error on the green line.

practice, this could mean a loud GW has a high coherent SNR *and* a high null SNR. For this reason, we increase the null SNR threshold with the coherent SNR for loud triggers but keep the null SNR threshold fixed for quiet triggers, as can be seen in fig.4.1. To be precise, we keep triggers meeting either of the following criteria:

$$\rho_N \leq 5.25 \text{ and } \rho_{\text{coh}} \leq 20 \quad (4.71)$$

$$\rho_N \leq \frac{\rho_{\text{coh}}}{5} + 5.25 \text{ and } \rho_{\text{coh}} > 20 . \quad (4.72)$$

The  $\chi^2$  tests (described in 4.2.6) are computationally expensive to calculate, so we calculate these only on those triggers that survive all the other tests. The power  $\chi^2$  test in particular is very expensive, and so is only calculated on triggers that pass the bank and auto-correlation  $\chi^2$  tests.

### 4.3.2 Reweighted SNR

Once we have a final list of triggers, we need a way to rank the likelihood that they are a real GW signal. This could be done using the coherent SNR, but this is easily skewed by glitches, even with the above mentioned cuts. To deal with this problem we down-weight the coherent SNR of the remaining triggers according to their coherent power  $\chi^2$  values and their null SNR values. Thus, those triggers

that pass the cuts but nevertheless are inconsistent with a GW end up with a lower *reweighted SNR*. The reweighting happens first according to the  $\chi^2$  values

$$\rho_{\chi^2} = \begin{cases} \rho_{\text{coh}} & \text{if } \chi^2 \leq n_{\text{dof}} \\ \frac{\rho_{\text{coh}}}{\left(\frac{1}{2}\left[1+\left(\frac{\chi^2}{n_{\text{dof}}}\right)^3\right]\right)^{1/6}} & \text{if } \chi^2 > n_{\text{dof}} \end{cases} \quad (4.73)$$

and then they are reweighted again according to their null SNR,

$$\rho_{\text{rw}} = \begin{cases} \rho_{\chi^2} & \text{if } \rho_{\text{null}} \leq 4.25 \\ \frac{\rho_{\chi^2}}{\rho_{\text{null}} - 3.25} & \text{if } \rho_{\text{null}} > 4.25 \end{cases} \quad (4.74)$$

Note that this reweighting only down-weights triggers with a high  $\chi^2$  or a high null SNR. The reweighting should not significantly alter the network SNR of a GW signal, but will reduce the SNR of noise transients. The reweighted SNR is the detection statistic used by PyGRB.

### 4.3.3 Event Significance

We use signal consistency checks to remove glitches from the trigger list, and we use the reweighted SNR to rank the surviving triggers in order of likelihood of being a real GW signal. The final step is to determine the significance of the surviving triggers, i.e. how often will we find glitches with a given reweighted SNR that have survived the signal consistency checks. We do this by calculating the false alarm probability (FAP) for each trigger. The FAP depends on the data quality at the time being analysed, as poor data quality can lead to more high SNR glitches which makes finding a high SNR trigger less significant. For this reason, we must estimate the rate of high SNR glitches at the time of the GRB in order to determine the significance of a trigger. In this section we outline the process used to calculate event significance.

As detector noise is not stationary, we calculate the FAP by using the data around the time of the GRB. If a GW is there, then it is assumed to be in the window starting five seconds before and ending one second after the GRB trigger time. This window would allow the detection of GWs from most GW models associated with short GRBs. We call this 6s window the *on-source* window. The loudest<sup>8</sup> event, with the loudest template, in the on-source is taken as our trigger.

To evaluate the p-value of this trigger, we analyse approximately one hour of data around the on-source<sup>9</sup>, called the *off-source* data. This data is assumed to not contain a GW signal. The off-source is divided up into as many 6s segments as

<sup>8</sup>Here *loudest* means highest reweighted SNR.

<sup>9</sup>The amount of data used before and after the on-source varies depending on the data available. For example, sometimes a GRB will happen less than an hour before (after) an interferometer loses (acquires) lock.

possible, so that if we have  $T_{\text{off}}$  seconds of data then we have  $N = T_{\text{off}}/T_{\text{on}}$  segments, where  $T_{\text{on}}$  is the length of the on-source. These segments are called the *off-source-trials*. Using the same criteria as the for the on-source, we find the loudest trigger in each off-source-trial. As the off-source is assumed to only contain noise, the FAP of the on-source trigger is the fraction of off-source trials that have a louder trigger than the on-source, i.e. the probability that noise would produce a trigger with a reweighted SNR louder than the on-source.

This process alone is only capable of achieving a FAP of  $10^{-3}$  as there are not enough off-source trials to claim a lower FAP. This is low enough to reject a signal hypothesis, but not low enough to claim a detection, for which we require a FAP of less than  $10^{-5}$  [40], which corresponds to one false alarm in  $\sim 2000$  years if we assume one short GRB per week. One way to get around this would be to use a longer off-source, but as detector noise is not stationary, it is possible that this extra off-source data would have different properties to the time of the GRB. It would also significantly increase the computational time required to complete the analysis, as well as the amount of detector data required to analyse a GRB, meaning fewer GRBs are analysed.

To find a lower FAP, the off-source data from each detector is time-shifted relative to the other detectors in the network and the data is then reanalysed. The time-shifts are longer than the light travel time between the detectors in the network and also longer than the duration of a typical glitch (i.e. less than one second), so that triggers in the time shifted data are unlikely to appear coherently.

There are two types of time shifting, called *short slides* and *long slides*. The two types of time slide is arise naturally from the fact that PyGRB analyses data in segments, typically of 128s. Time shifting the data within a segment is called *short slides*. For example, in an HLV analysis the Hanford data is not shifted, the Livingston data is shifted in increments of 6s, and the Virgo data is shifted in increments of 12 seconds. This does not require the data to be matched filtered again, as the time shift happens after filtering. This means only the network statistics must be recomputed, making short slides computationally cheap but only able to demonstrate a FAP as low as  $10^{-4}$ . *Long slides* refers to when the segments themselves are shifted relative to one another. This requires the matched filtering to be redone, as the matched filter time series is not saved when analysing the following segment. This makes long slides much more computationally expensive than short slides. However, short slides can be done on top of the long slides, so that just 10 long slides are needed to show that a trigger has a  $\text{FAP} < 10^{-5}$ .

It should be noted that our treatment of time slides assumes that each combination of time shifted data is statistically independent of the others. This is not true, as each combination of time shifted data is still the same data being analysed. For this reason, the FAP calculated from time slides is a lower limit on the true FAP [38] [39].

#### 4.3.4 Searching over a Sky Patch

The sky position of GRBs cannot be determined exactly. The uncertainty in sky position depends on which satellites detected the GRB<sup>10</sup>. The BAT instrument on Swift can localise to within 1-4 arcminutes[9]. This is a small enough uncertainty that the GW search can use a single sky point[2]. Other GRB detectors cannot provide as precise localisation. The GBM on the Fermi satellite often provides roughly circular  $3\sigma$  confidence region with a radius of several degrees[17]. The IPN uses triangulation to determine sky position, so depending on the number of the satellites that make the detection and how spatially separated they are, the sky localisation can vary from under a square degree to thousands of square degrees.

It is important to have approximately the correct sky position for a coherent search. If the sky position is incorrect, then the detector antenna patterns will be incorrect for the GRB, reducing the search sensitivity. More crucially, an incorrect sky position will cause the calculated time delays between detectors in the network to be wrong, making it impossible to correctly find coincident triggers between detectors. To overcome these issues, PyGRB searches over a grid of points that cover the uncertainty region provided by GRB detectors[40][37].

The grid of points is constructed by filling the uncertainty region with concentric circles of points separated by  $\delta a$ , such that each ring has  $2\pi n/\delta a$  points with  $n = 0$  being the central point of the patch. These concentric circles extend to cover the  $3\sigma$  confidence region. The value of  $\delta a$  depends on the timing uncertainty we are willing to accept. In practice, we use a timing uncertainty of  $\delta t = 0.5\text{ms}$ , which will limit the lost SNR to  $< 5\%$ [40]. Searching over the sky grid is done after the computationally dominant step of matched filtering, so that searching over the sky grid does not significantly slow down the analysis.

In the two detector case, it is possible to search a reduced number of sky points without losing sensitivity. This is because there is a ring of sky points that give the same time delay between the two detectors. Searching different points on this ring would not change the time of arrival of a GW to the detectors, though it would change the antenna response factors. However, as the antenna response functions drop out of the formula for the coherent SNR (4.37) in the two detector case,<sup>11</sup> searching over different points on those sky rings would give exactly the same results. Thus, in the two detector case, we can analyse a greatly reduced set of sky points. Unfortunately, when we limit our search to looking for circularly polarised GWs, as discussed in section 4.2.5, then the antenna patterns again affect the detection statistic, and we must search over the sky grid.

---

<sup>10</sup>See section 3.4 for details on GRB detectors.

<sup>11</sup>This is because in the two detector case, both the coherent and coincident SNR have four degrees of freedom. Thus, any observed amplitude and phase is consistent with an astrophysical signal.

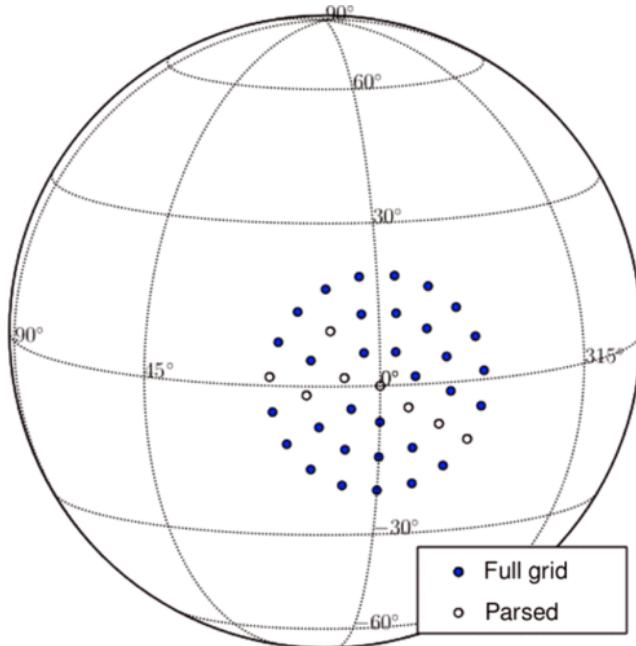


Figure 4.2: **PyGRB Sky Grid.** Here we see an example of a full search grid used by PyGRB, indicated by the blue dots, and a reduced sky grid parsed by PyGRB in the case of a two detector search using the Hanford and Livingston detectors, the empty circles labeled ‘parsed’. The parsed circles do not form a line due to the parsing routine, but this has no effect the on analysis.[40]

#### 4.3.5 Calculating Search Sensitivity

It is useful to estimate the sensitivity to GWs around the time of the analysed GRB. In the case that no GW is detected, this can be used to provide a lower limit on the distance to the source.

To do this, we inject simulated signals into the off-source data and see if PyGRB can find them. The injected signals are CBC waveforms drawn from three populations: BNS mergers, NS-BH systems that have spins aligned with the angular momentum, and NS-BH systems with spins generically aligned. The injections are drawn from an astrophysically motivated distribution of distances, component masses and spins, and binary inclination. The NS mass distribution is a normal distribution with mean  $1.4M_{\odot}$  and standard deviation of  $0.2M_{\odot}$ [24][42] restricted to the  $1M_{\odot} - 3M_{\odot}$  range, with the upper limit set by theoretical considerations [22]. The NS spins magnitude is restricted to be  $\leq 0.4$ , the fastest observed pulsar spin [20]. The BH masses are drawn from a normal distribution of mean  $10M_{\odot}$  with standard deviation of  $6M_{\odot}$ , restricted to the range  $3M_{\odot} - 15M_{\odot}$ . The BH spin magnitudes are  $\leq 0.98$ , motivated by X-ray observations[43]. The injections are limited to have binary inclinations of  $0^{\circ} - 30^{\circ}$  or  $150^{\circ} - 180^{\circ}$ , as the opening angle for a CBC powered GRB is expected to be less than  $30^{\circ}$  [25].

We quantify the sensitivity of the search using the 50% and 90% exclusion dis-

tance. This is the distance at which 50% and 90% of the injected signals are recovered with a greater reweighted SNR than the most significant on-source trigger.

### 4.3.6 The PyGRB Workflow

We conclude this section by seeing how all the different elements we have seen come together in the PyGRB workflow. Figure 4.3 shows the PyGRB workflow. The workflow starts with two parallel branches: one for the injection runs (right branch) and the other for the background and on-source analysis (left branch). We will discuss the non-injection branch first.

To speed up the analysis, the matched filter jobs are run in parallel. In order to do this, the template bank is first split into many smaller template banks (typically  $\sim 1500$ ). These are used by the *inspiral* jobs, which matched-filter the data, perform the consistency checks, and calculate the network SNR statistics. The results of each inspiral job are then compiled together by *trig combiner* and binned into off-source, on-source, or off-trial triggers.<sup>12</sup> Another file is produced which contains all triggers found (i.e. both in the on-source and in the off-source). Each of these is then *clustered*, meaning only the loudest trigger with the loudest template in any 0.1 second window is kept. The clustered output is then used to calculate detection efficiencies and make plots before going into the results webpage.

The injection jobs are similar but have a few extra steps. First the injections must be generated. At the moment this uses the same code (called *inspinj*) as the all-sky search for black hole binaries. As we are only looking for binaries that could theoretically produce a GRB, we remove any injections that have the wrong masses or orientations.<sup>13</sup> The injections are then split into multiple jobs and combined with different parts of the split bank to be analysed by the inspiral code. These jobs can be run quickly by only searching the data with those templates that are expected to find the injections, rather than using the full template bank. For this reason, the bank is only split into two for the injection runs. The *inj finder* code analyses the output of the injection inspiral jobs to determine which injections were found and which were not. The *inj combiner* job combines injections of the same waveform type from different injection campaigns into a single file. The inj combiner output is used to calculate injection detection efficiencies and make plots, which are then used in the results webpage.

---

<sup>12</sup>The off-trials are six second chunks of data which are each treated as a dummy on-source to assess whether the pipeline is behaving nominally. Only if these off-trials behave sensibly will we look at the on-source results.

<sup>13</sup>The scripts that remove unsuitable injections are not included in the flowchart as they do not add much to the understanding of the workflow. In future, a script that produces only the appropriate injections will be written to replace the ad hoc system we currently use.

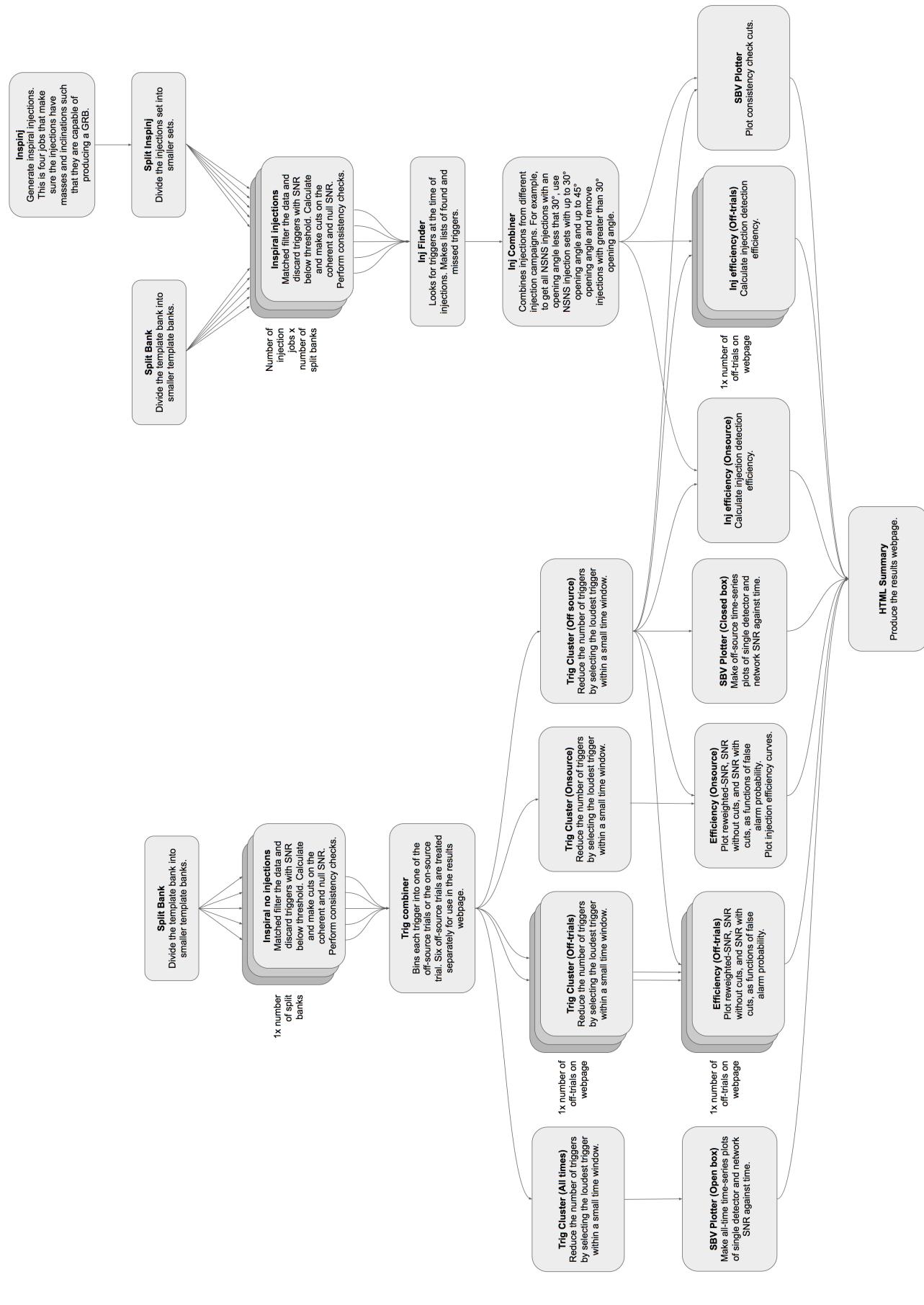


Figure 4.3: **PyGRB Workflow**. The workflow starts in two parallel branches, one that runs the injections jobs, and one that analyses the background and onsource.

## 4.4 O2 PyGRB Search

From November 2016 to August 2017, the Advanced LIGO detectors undertook their second observing run, O2, with the Advanced Virgo detector joining on August 1st 2017. PyGRB was used to search for GW signals associated with short and ambiguous duration GRBs detected during O2. The results of the search were reported in [11], but we will outline the key findings here. We begin with a brief discussion of the GRB sample that was analysed before moving on to the results of the search.

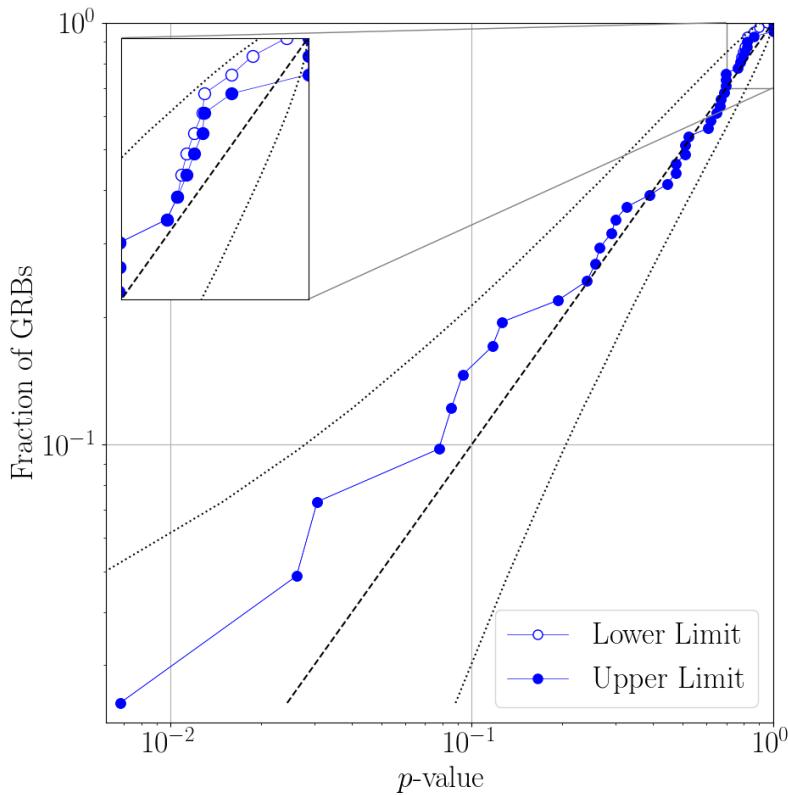
### 4.4.1 GRB sample

The GRBs analysed were those detected by Swift BAT, Fermi GBM, or the IPN (see section 3.4 for more details on these). There were in total 242 bursts detected by Swift and Fermi, and 52 detected by the IPN, with many GRBs appearing in both catalogues. Only short GRBs ( $T_{90} < 2s$ ) and ambiguous GRBs ( $2s < T_{90} < 4s$ ) were analysed using PyGRB, as long GRBs are not expected to have a CBC progenitor (as discussed in chapter 3). At least 1664s of data is required for PyGRB to correctly estimate the background. Any detector that does not have this much data available around the time of a GRB is not used for the analysis, and if no detectors have sufficient data available then PyGRB will not analyse that GRB. Removing the GRBs for which there was insufficient data left 42 short/ambiguous GRBs that could be analysed by PyGRB.

### 4.4.2 Results

The analysis found one GW signal, GW 170817. It was associated with GRB 170817A and had been previously reported [4][13]. The p-value for this event is  $\leq 9.38 \times 10^{-6}$  and the coherent SNR is 31.26[11]. No other GWs were detected in conjunction with any other GRBs. Apart from GW 170817, there were six GRBs with a p-value of less than 0.1. These candidate events had further data quality checks, which found no clear noise source that could explain the triggers. They were then analysed using lalInference[35], a coherent Bayesian parameter estimation pipeline, to determine if there could be a subthreshold signal in the data or if it was more likely to be background noise. No evidence was found of a subthreshold signal.

In figure 4.4 we have plotted the p-value for the remaining 41 GRBs analysed after the removal GW 170817 from the sample. For GRBs with no trigger in the on-source window, we provide upper and lower limits on the p-value. The upper limit is a p-value of 1. The lower limit is the fraction of off-source trials that also had no trigger. The distribution lays within the  $2\sigma$  range, shown by the upper and lower dotted lines. The population consistency with a no-signal hypothesis was calculated using the weighted binomial test outlined in [1]. This test considers the lowest 5%



**Figure 4.4: P-value for each GRB.** This is the p-value distribution for the 41 GRBs other than GRB 170817A. The GRBs with no trigger in the onsource window have upper and lower limits on the p-value. The upper limit is a p-value of 1. The lower limit is the fraction of offsource trials that also had no trigger. The distribution lays within the  $2\sigma$  range, shown by the upper and lower dotted lines.

of p-values in the population, weighted by their prior probability of detection based on the time and sky position of the GRB. This analysis did not include GW 170817 as it is a confirmed detection. The combined p-value for the search is 0.3. Thus we find no significant evidence for a population of subthreshold signals.

As GRB 170817A is known to have originated in the galaxy NGC 4993[13], the distance to this GRB is approximately 43Mpc. For the other GRBs in our sample we calculated the 90% exclusion distance. This is the distance at which 90% of the signal injections are recovered with a greater coherent SNR than the loudest trigger in the on-source. The 90% exclusion distances are plotted in figure 4.5. The median 90% exclusion distance for the BNS waveform is 80 Mpc, for an NS-BH waveform with generic spin is 105 Mpc, and for NS-BH with aligned spin is 144 Mpc. These values are slightly lower than in O1, which were 90 Mpc, 139 Mpc, and 150 Mpc respectively[16], though this seems to just be due to a statistical fluctuation due to the small sample size in O1. There is a summary table of the information about each GRB analysed by PyGRB in appendix .2. This table includes the 90% exclusion

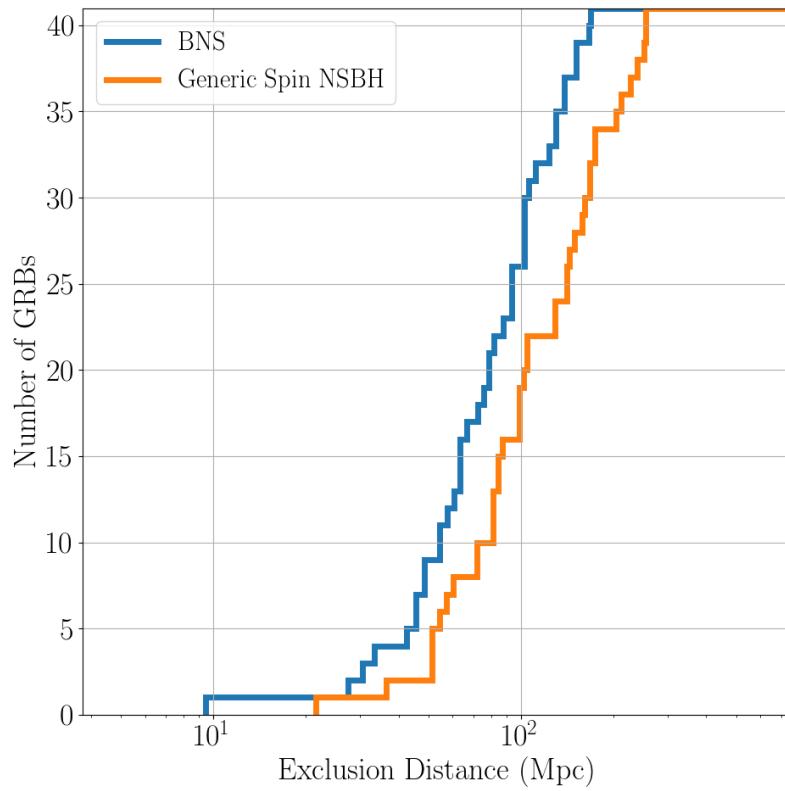


Figure 4.5: **Cumulative exclusion distance.** This is the cumulative 90% exclusion distance for every GRB analysed by PyGRB except GRB170817A. The 90% exclusion distance is the distance at which 90% of injected simulated signals are recovered with a greater coherent SNR than the loudest trigger in the onsource.

distance for each waveform type for each GRB.

#### 4.4.3 Rates

During O2 we detected just one GW signal from a sample of 42 short GRBs. For the non-detections we placed lower limits on the distance to the source. Using this information it is possible to estimate the rate at which BNS mergers produce GRBs luminous enough to be observed. This depends on the intrinsic luminosity of the GRB, as well as the jet structure and opening angle. While there is strong evidence that GRB jets are structured, we will treat all differences in measured luminosity as though they resulted from an intrinsically lower luminosity. This is acceptable for the purposes of modeling the potential for joint GW-GRB detections as we only need to know the number of faint GRB detections, not the reason why they are faint. As such, we use the luminosity function from [13], which extends the model

in [36] to have a second break at low luminosity. This function is given by

$$\phi_o(L_{\text{iso}}) = \begin{cases} \left(\frac{L_{\text{iso}}}{L_{**}}\right)^{-\gamma_L} \left(\frac{L_{**}}{L_*}\right)^{-\alpha_L} & L_{\text{iso}} < L_{**} \\ \left(\frac{L_{\text{iso}}}{L_*}\right)^{-\alpha_L} & L_{**} < L_{\text{iso}} < L_* , \\ \left(\frac{L_{\text{iso}}}{L_*}\right)^{-\beta_L} & L_{\text{iso}} > L_* \end{cases} \quad (4.75)$$

where  $L_{\text{iso}}$  is the isotropic luminosity of the GRB, and the parameters  $L_* = 2 \times 10^{52}$  ergs $^{-1}$ ,  $L_{**} = 5 \times 10^{49}$  ergs $^{-1}$ ,  $\alpha_L = 1$  and  $\beta_L = 2$  are used to fit the observed GRB redshift distribution, and  $\gamma_L$  is the free parameter associated with the second break. To fit this model to the data, the detectability threshold for Fermi-GBM used was 2 photons cm $^{-2}$ s $^{-1}$  for the 64 ms peak photon flux in the 50-300 keV band and the short GRB spectrum was modeled as a band function with  $E_{\text{peak}} = 800$  keV,  $\alpha_{\text{Band}} = 0.5$ ,  $\beta_{\text{Band}} = -2.25$ . The redshift distribution obtained by this model was then normalised to match the Fermi-GBM detection rate of 40 short GRBs per year. The free parameter  $\gamma_L$  was constrained using Monte Carlo sampling to calculate the probability of finding the results obtained during O2 for a given value of  $\gamma_L$ . This yields 90% confidence bounds on  $\gamma_L$  of [0.04, 0.98]. In figure 4.6 the luminosity functions corresponding to these bounds are plotted. For comparison, the estimated BNS merger rate[13] and the cumulative Fermi detection rate[21] are plotted, as well as the redshift for every short GRB for which a measurement exists apart from GRB 170817A. A subset of the measured redshifts, called the *gold sample*, are those for which a confident association with a host galaxy was made, which reduces the chance that the redshift measurement was made for the wrong galaxy.

As can be seen, the rates calculated using the PyGRB results are consistent with both the estimated BNS merger rate and the Fermi observed rate. With more GW detections of BNS systems, it will become possible to determine the ratio of BNS mergers that have produce detectable short GRBs, which in turn has implications for GRB jet structure and opening angle. Also, if short GRBs are detected within the LIGO-Virgo horizon that do not have a corresponding GW signal, then that would suggest that not all short GRBs have BNS progenitors. With the model described above, the 2019-20 LIGO-Virgo observing run is predicted to make 1-30 BNS detections, and 0.07-1.80 joint GW-GRB detections[11].

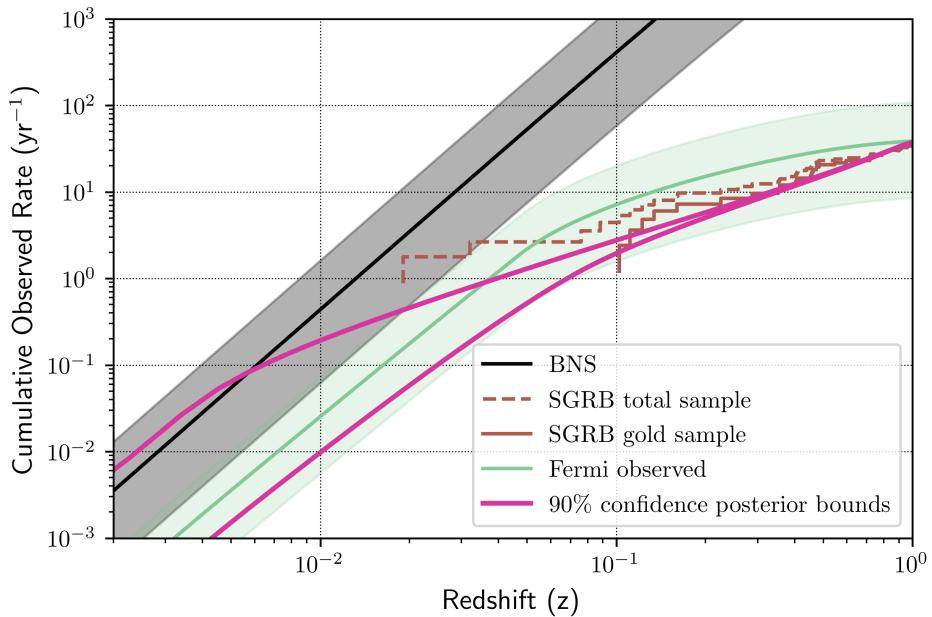


Figure 4.6: **Cumulative Rate of BNS and short GRB Events.** The magenta lines show the 90% confidence bounds for joint GRB-GW events as a function of redshift. This was calculated using the 41 non-detections and single detection by PyGRB during O2. The black line and the grey region shows the estimated BNS merger rate  $1210^{+3230}_{-1040}$ . In green is shown the estimated Fermi detection rate and its 90% confidence region[21]. The measured redshifts of every short GRB apart from GRB 170817A are shown in brown. The gold sample refers to those GRBs that were localised to near a host galaxy, making the redshift measurement more reliable than short GRBs measured more distant from a host galaxy. Our results are compatible with both the Fermi-GBM observed rate and the predicted BNS merger rate.[11]

# Chapter 5

## The Future of PyGRB

In the previous chapter we discussed the PyGRB pipeline as it was used in the second LIGO observing run O2. The code is now about 10 years old, and in that time new software tools have become available that could speed up and improve the analysis. In particular, there is the PyCBC software package [**cite pycbc paper**], which has highly optimised tools for performing a matched filter search. For this reason, PyGRB is being rewritten and fully implemented in the PyCBC framework. In this chapter we will begin by discussing the reasons why this change was necessary, then we will see the progress that has already been made in rewriting PyGRB and the improvements they provide. We will then end this chapter by looking forward at the changes that need to happen for the PyGRB rewrite to be used in a GW search.

### 5.1 Reasons to Rewrite the Pipeline

The primary reason for rewriting the pipeline is to make use of PyCBC’s faster and more efficient matched filtering software. Developments in the all-sky search have drastically sped up matched filtering in recent years, but these developments are currently unavailable to PyGRB. Speeding up the analysis will not only reduce the computational cost of the analysis, but allow new science to be done with PyGRB. The medium latency PyGRB search aims to provide a preliminary analysis of a GRB within a day. As the kilonova that follows a short GRB fades on a timescale of hours, this delay can mean missing vital data about the early evolution of the kilonova.<sup>1</sup> With the new code, we expect to be able to perform a preliminary analysis (with minimal injections and no timeslides) within a couple of hours. Eventually we could take advantage of the low-latency PyCBC search, *PyCBC live* **cite pycbc live paper**, to reduce this time to minutes or even seconds.

Slow matched filtering also limits the size of the sky patch that can be analysed in a reasonable amount of time. There were four short/ambiguous GRBs in O2 that

---

<sup>1</sup>For example, the top right panel of figure 3.13 shows that observations in the first day after a short GRB can be vital for determining the jet structure.

were too poorly localised for PyGRB to analyse in a reasonable amount of time.<sup>2</sup> With faster code, these could have been analysed. Similarly, slow matched filtering limits the number of timeslides that can be performed in a reasonable amount of time. This limits how accurately the FAP of a trigger can be measured.

There are two bottlenecks in the PyGRB workflow. The first is the matched filtering, where using the PyCBC matched filtering engine can lead to significant speed up. The other is in post-processing, which can be sped up significantly by using a modern file format.<sup>3</sup> The post-processing also frequently exceeds the memory allowance on computing clusters, which requires the job to be manually halted and rerun with the post-processing split into many smaller jobs. As PyGRB deals with far fewer triggers (and so should use less memory) than the all-sky search, this should not be a problem. Using tools in PyCBC will fix this.

Faster code is not the only reason for integrating PyGRB into PyCBC. It will also make new tools that were developed for the all-sky search available to the targeted search, as we will see in section 5.2. Integrating into PyCBC will also mean that any new tools developed in PyCBC will become immediately available to PyGRB, increasing the rate of development of PyGRB. Similarly, any techniques developed for the targeted search can easily be adopted by the all-sky search. For example, coherent follow-up, using PyGRB tools, could be carried out on the sky patch around GW triggers found using the all-sky search.

The final reason to integrate PyGRB into PyCBC is to make the code more user friendly and less cumbersome. This will speed up development and make it easier for new users to become proficient in using PyGRB. One way we will do this is by using the tools developed by PyCBC for making the results webpages. The PyGRB results webpages are static and contain little information about how to reproduce the results they show. The PyCBC webpages are far more user friendly, containing interactive tables and the command line arguments required to replicate a the plots. These tools will, for example, speed up diagnosing missed injections.

Another example of how integrating into PyCBC will make PyGRB simpler is the PyGRB implementation of autogating. Autogating is a method to prevent very loud glitches contaminating the matched filter time series. PyCBC does this by using a time-frequency excess power burst search **cite** <https://arxiv.org/pdf/1602.03839.pdf> **and refs there in** to find times of excess strain in the data.<sup>4</sup> These times are then zeroed out using an inverse Tukey window, as seen in figure 3.7. While autogating was used by PyGRB during O2, its implementation was ad hoc. A script was made that used PyCBC tools to autogate the data and write out the gated frame files,

---

<sup>2</sup>These were all IPN detections, which can have poor localisation depending on which satellites detect the GRB (see section 3.4).

<sup>3</sup>The old code used xml files, which are significantly slower for our purposes than the hdf5 files used by PyCBC.

<sup>4</sup>To be precise, the threshold for gating was set at a burst SNR of 100 during O2, higher than any expected GW signal. For example, GW 150914 had a burst SNR of 10.

which are then used by the rest of the PyGRB workflow. In future, the autogating will be handled automatically by the matched filtering executable, just as it is with PyCBC. This will not have a significant effect on analysis time, but is an example of how integrating into PyCBC will simplify the workflow.

## 5.2 Workflow Modifications

We have discussed the reasons why rewriting PyGRB was necessary. But the rewrite was not simply a case of carrying out the same analysis with a faster matched filtering engine. The new analysis is qualitatively different from the one described in chapter 4. The most significant change is to the signal consistency tests. Not only can we use new tools developed in PyCBC to make these tests more powerful, but we have chosen to move away from using coherent  $\chi^2$  tests in favour of the single detector tests. We will discuss these changes in 5.2.1. The change in  $\chi^2$  tests then necessitates the development of a new detection statistic, as the old one uses the coherent  $\chi^2$  tests. This is discussed in 5.2.2. We will begin with the most significant change to the PyGRB pipeline, the new coherent matched filtering executable.

### 5.2.1 Consistency Checks

Most of the changes made to the PyGRB pipeline were motivated by either speeding up the analysis or using new data analysis tools that were not previously available. One of the changes we made that does not fall into either of these two categories is abandoning the coherent  $\chi^2$  tests.<sup>5</sup> Apart from being computationally expensive, it was also clear that these tests were performing no better than the single detector  $\chi^2$  tests.<sup>6</sup>

The problem with coherent  $\chi^2$  tests is that glitches are a single detector phenomenon. Coherent matched filtering works well because we have a network of detectors that should all detect the same signal when a GW is present. In this case, the coherent matched filter provides a natural way to downweight coincident signals that are not consistent in phase or amplitude with the sky location and time of the GRB. Contrast this with a glitch, which a burst of energy in one detector. In this case, combining the data streams from multiple detectors washes out the effect of the glitch, making it harder to detect, as the signal consistency tests are supposed to do. This is clear from the formula for the coherent  $\chi^2$  tests, give by equation 4.65 but repeated here for convenience;

$$\chi^2 = \sum_{\mu=1}^4 \sum_{i=1}^N (\mathbf{T}_\mu^i | \mathbf{s})^2 \quad (5.1)$$

---

<sup>5</sup>Described in section 4.2.6.

<sup>6</sup>Also described in section 4.2.6.

where the  $\mathbf{T}_\mu^i$  are the  $N$  multidetector waveforms orthogonal to  $\mathbf{h}$  and  $\mathbf{s}$  is the vector of detector data streams. The sum implicit in the multidetector inner product makes the contribution of any one detector more difficult to detect, making a test for a single detector glitch less effective. Ultimately, this manifests itself in a higher variance for the  $\chi^2$ , given by equation 4.67 but repeated here for convenience;

$$\text{Var}(\chi^2) = 8N + 2 \sum_{i \neq j} \sum_{\mu \neq \nu} [(\mathbf{T}_\mu^i | \mathbf{T}_\nu^j)^2] \quad (5.2)$$

Compared to the single detector  $\chi^2$  test, given by 4.61. The variance of the multi-detector  $\chi^2$  test is  $6N$  larger from the first term, as well as the sum over  $\mu$  and the implicit sum over detectors in the inner product. As the variance is much higher in the coherent  $\chi^2$ , it is inherently more difficult to place an appropriate threshold to remove glitches. For these reasons, we have decided to drop the use of the coherent  $\chi^2$  tests from the PyGRB analysis, in favour of the single detector  $\chi^2$  tests used by the PyCBC all-sky search.

Not only have we moved to single detector  $\chi^2$  tests, but PyCBC has made improvements to the  $\chi^2$  test that can now be incorporated into the GRB search. The power  $\chi^2$  test described in 4.2.6 divided the template into a fixed number of bins, which should each contain an equal amount of signal power for a real GW. This test can be improved by using the parameters of the template to determine the number of bins. In particular, templates with a higher peak frequency have more cycles in the sensitive band of the detectors and so using more  $\chi^2$  bins can make the test more powerful. Conversely, a template with a low peak frequency will have fewer cycles in the sensitive band and so needs to use fewer bins, otherwise some of the bins will have an SNR comparable to the background noise.

PyGRB currently uses 16  $\chi^2$  bins, regardless of template parameters. The new code uses a variable number of bins  $p$  according to the formula

$$p = \lfloor 0.72 f_{\text{peak}}^{\frac{2}{3}} \rfloor \quad (5.3)$$

where  $f_{\text{peak}}$  is the peak frequency of the waveform, measured in Hertz. This formula is the same as the current PyCBC analysis, and was determined empirically. **cite GW150914 paper, though this has a different formula, look for more up-to-date citation**

### 5.2.2 Detection Statistic

As we no longer use the coherent  $\chi^2$  statistic, we also cannot use the reweighted SNR described in section 4.3.2. Instead we define a new network  $\chi^2$  statistic, and use this to define a new reweighted SNR. This statistic needs to be able to take into account both the signal consistency and the SNR of the trigger in each detector. If we only used the signal consistency, then a high  $\chi^2$  value in one detector would be

enough to veto a trigger, even if that detector only had a low SNR. In the worst case, a GW that occurred in the null of a detector (like GW 170817 did) will have a low SNR in that detector, but could also have a high  $\chi^2$ . The GW would then be vetoed despite having a large SNR and low  $\chi^2$  in the remaining detectors in the network. To get around this problem, we reweight the  $\chi^2$  contribution from each detector by the SNR from that detector before summing over the  $\chi^2$  value from each detector. The resulting network  $\chi^2$  is then used to reweight the coherent SNR, in a similar way as the PyCBC all-sky search reweights the single detector SNRs with the single detector  $\chi^2$ .

To be precise, the new network  $\chi^2$  statistic for an  $N$  detector network, where the SNR and  $\chi^2$  per degree of freedom values for each individual detector are given by  $\rho_i$  and  $\chi_i^2$  respectively, is given by

$$\chi_N^2 = \frac{1}{\rho_{\text{coinc}}^2} \sum_{i=0}^N \chi_i^2 \rho_i^2 . \quad (5.4)$$

From this, we calculate the new reweighted SNR to be

$$\rho_{\text{rw}} = \frac{\rho_{\text{network}}}{(1 + (\chi_N^2)^a)^b} \quad (5.5)$$

with  $a = 3$  and  $b = 1/6$  by default. Here  $\rho_{\text{network}}$  denotes a network SNR statistic, which for a network of three or more detectors will be the null reweighted coherent SNR (discussed in 4.3.2) and in the one or two detector case will be the coincident SNR (as there is no null stream). In the next section, we will see how well the network  $\chi^2$  and the new detection statistic perform in a search.

### 5.2.3 The new Coherent Matched Filtering Executable

The most significant change to the PyGRB workflow is the new matched filtering executable. This has been written entirely within the PyCBC framework. In this section we outline the algorithm the new matched filtering executable goes through. As coherent statistics and signal consistency tests are computationally expensive, the philosophy behind writing the coherent matched filter code has been to apply the computationally cheapest tests first, and only perform the more expensive tests on those triggers that survive all the previous tests.

With this in mind, let us start at the beginning. Using the time and sky location of the GRB being analysed, the antenna pattern for each interferometer is calculated. The coherent matched filtering code then reads in data from every detector in the network. The first template from the bank is then matched filtered against each the data from each detector. A threshold, usually an SNR of 4, is applied to the SNR time series from each detector. Any sample in the SNR time series that passes this threshold in any detector is called a *trigger*. The SNR time series from each detector

is then time shifted, using the GRB sky position, so that a GW would be coincident in all of them. At this stage, any trigger that is not coincident in at least two of the detectors is cut.<sup>7</sup> For each remaining trigger, the SNR from each detector is added in quadrature<sup>8</sup> to produce a the *coincident SNR*. A threshold is then applied, such that any trigger with a coincident SNR below 6 is cut.

All of the cuts made so far were computationally cheap, because now the algorithm gets more expensive. If this analysis uses three or more detectors, then we now calculate the coherent SNR. We first use the antenna pattern and the template to calculate the inverse of the matrix 4.2.2. We then use this to calculate the coherent SNR, following equation 4.25. We then apply another threshold such that any trigger with a coherent SNR below 6 is cut. It is then easy to calculate the null SNR by finding the quadrature difference between the coincident SNR and coherent SNR for the remaining triggers. We apply a threshold on the null SNR that increases with the coherent SNR of the trigger, as described in section 4.3.1.

By this point, it is very rare to have more than perhaps  $\sim 10$  triggers left. For these remaining triggers we calculate the  $\chi^2$  tests, and then use these to calculate the network  $\chi^2$ , as described in section 5.2.1. The  $\chi^2$  values are computationally very expensive to calculate, but few triggers survive the previous cuts so it is not too much of a burden. The network  $\chi^2$  is then used reweight the coherent SNR, as described in 5.2.2. The triggers are then reweighted again by the null SNR, as in section 4.3.2. No cut is applied to the reweighted SNR in the executable, as there are no more statistics left to calculate and so no improvement in analysis time to be gained by further cuts at this point.<sup>9</sup> The remaining triggers are then *clustered*, meaning that only the most significant trigger in any 0.1 second window is kept, and then stored in memory. The code then moves onto the next template in the template bank. Once all templates have been analysed, an output file containing all the information about each trigger is written.

### 5.3 Search Performance

In this section we see how the new code performs. We will start by looking at the results of some example analyses in order to convince ourselves that the code is behaving sensibly. Then we will look at how much faster the new code is than the old.

As the new code cannot yet search over a sky patch, we chose two well localised GRBs for the test. The first was GRB 170817A, this has an associated GW, and as an afterglow was located which provides very accurate sky localisation. This was

<sup>7</sup>This is a very powerful cut. For almost every template, this will leave no triggers remaining and we can move on to the next template.

<sup>8</sup>Note that this includes the SNR contribution from any detector that was below the single detector SNR threshold.

<sup>9</sup>Although more cuts are made on the triggers in post-processing.

also a three detector analysis, allowing us to check the coherent and null SNR are behaving correctly. The second GRB was GRB 170112A, which was a well localised swift GRB with Livingston and Hanford detector data available at the trigger time.<sup>10</sup> The data around the trigger time of this GRB had some noise transients, that can be used to test the effectiveness of the network  $\chi^2$  test and the new detection statistic.

### 5.3.1 Testing the Detection Statistic

In this section we will look at the new detection statistic, to see how well it performs in a GRB analysis. We will start by looking at the SNR time series for the two GRBs analysed. Figure 5.1 shows the coherent SNR (top panel) and the reweighted SNR (bottom panel) timeseries for GRB 170817A. The plots show the triggers for every template in the template bank,<sup>11</sup> clustered such that only the most significant trigger in each 0.1 second window is kept. In the top panel we can see the coherent SNR time series looks clean, and GW 170817 is clearly visible at the correct time with a coherent SNR of 29.22, which is comparable to the coherent SNR of 31.26 PyGRB obtained in O2.<sup>12</sup> The bottom panel shows the reweighted SNR time series. We can see that some of the noise transients that peak the coherent SNR have been downweighted to have a comparable significance to background noise, but the GW is still clearly visible. In figure 5.2 we plot the coherent (top panel) and reweighted (bottom panel) SNR time series for GRB 170112A. From the coherent SNR time series we can see that there are several significant glitches. The reweighted SNR time series shows that these glitches were downweighted to be no more significant than background noise.

From the time series plots, it seem that the reweighted SNR is behaving as expected. In figure 5.3 we have plotted the null SNR against the coherent SNR (left) and reweighted SNR (right) for GRB 170817A<sup>13</sup>, in order ensure that triggers are being correctly reweighted by the null SNR. The null SNR shows no significant deviations from the background level, with a peak of just 5.22. Only a few points are above the 4.25 threshold required to reweight a trigger by the null SNR, see equation 4.74 but the triggers that exceed this threshold are being appropriately downweighted.

Similarly, in figures 5.4 and 5.5, we plot the coherent SNR (left) and reweighted SNR (right) against the network  $\chi^2$  for GRB 170817A and GRB 170112A respectively. Figure 5.4 shows that the GW triggers with higher network  $\chi^2$  are downweighted more. The GW, visible as the rightmost point in each plot, is downweighted

---

<sup>10</sup>GRB 170112A was not a single sky point analysis when analysed by PyGRB in O2, but did have a small sky error. The O2 PyGRB analysis searched over three sky points.

<sup>11</sup>There are about 190000 templates in the template bank.

<sup>12</sup>Some differences in SNR are to be expected as the search is using different PSD options to the O2 PyGRB search.

<sup>13</sup>As GRB 170112A was a two detector analysis, there was no null SNR. Hence this plot is only made for GRB 170817A.

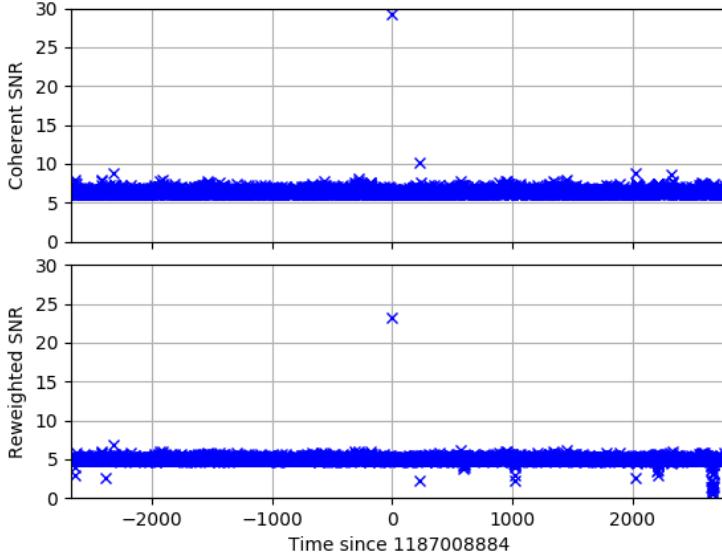


Figure 5.1: **Coherent and Reweighted SNR vs Time for GRB 170817A.** The top panel shows the coherent SNR vs time for GRB 170817A. The GW is clearly visible, as are some smaller peaks that are due to noise. The bottom panel shows the null and  $\chi^2$  reweighted SNR. We can see that GW is still very prominent, but the peaks in coherent SNR that were due to noise have been downweighted.

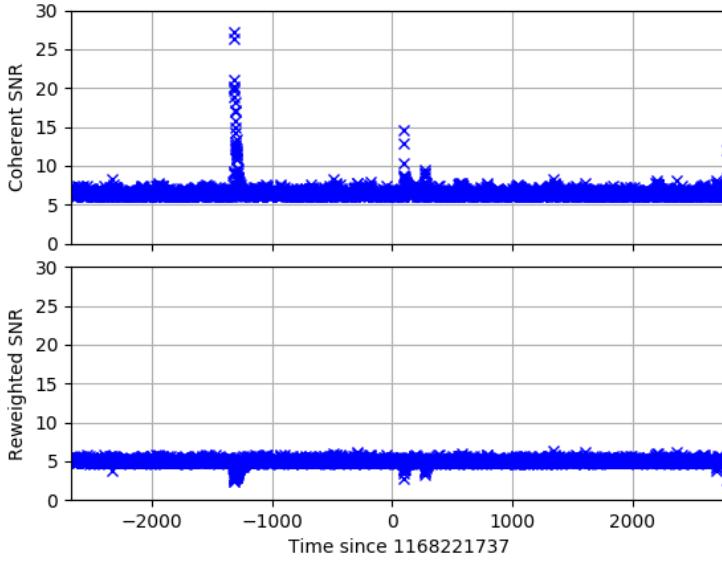
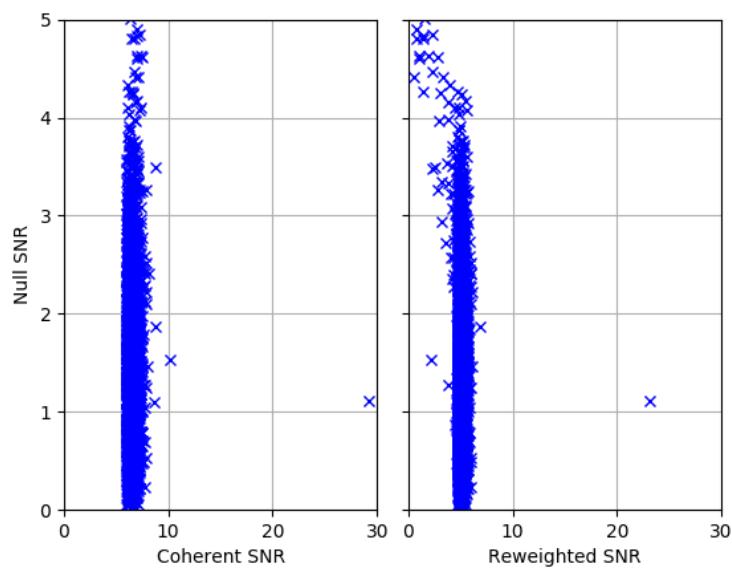


Figure 5.2: **Coherent and Reweighted SNR vs Time for GRB 170112A.** The top panel shows the coherent SNR vs time for GRB 170112A. There is no GW, but several glitches are clearly visible. The bottom panel shows the  $\chi^2$  reweighted SNR. We can see that the glitches have been downweighted to be no more significant than the background.



**Figure 5.3: Null SNR vs Reweighted SNR for GRB 170817A.** Here we plot the null SNR against the coherent SNR (left) and the reweighted SNR (right) for GRB 170817A. Only triggers with a null SNR above 4.5 are reweighted by the null SNR (with the other triggers being reweighted only by their  $\chi^2$  values), and we can see that these triggers have been downweighted more than the triggers with a low null SNR. The GW is clearly visible on the right of both plots.

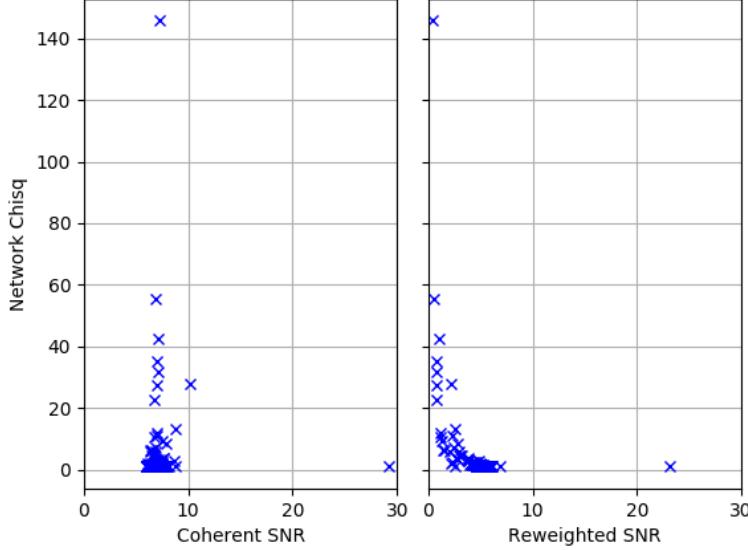
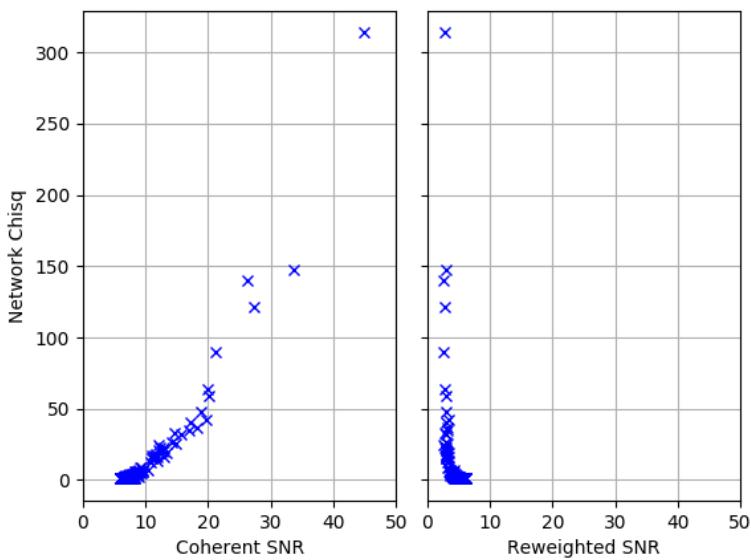


Figure 5.4: **Reweighted SNR vs Network  $\chi^2$  for GRB 170817A.** Here we plots the network  $\chi^2$  against the coherent SNR (left) and reweighted SNR (right) for GRB 170817A. We can see that triggers with the higher network  $\chi^2$  of a trigger, the more it is downweighted. The GW is clearly visible on the right of both plots.

but still far more significant than any background event. In figure 5.5, we see that the glitches in the data around GRB 170112A have a high coherent SNR and a high network  $\chi^2$ , but are appropriately downweighted so that these loud glitches are no more significant than background events.

We also plotted the peak coherent and reweighted SNR in each 6-second off-source trial, as well as the on-source. The results for GRB 170817A can be seen in figure 5.6. As we do not have timeslides, there are not many off-source trials (892 in total), but we can clearly see that the on-source trial, marked by the blue and red stars, are substantially louder than the background for both the coherent and reweighted SNR. Note that the coherent SNR has a small tail of higher significance events at around a coherent SNR of 7-9, which does not appear in the reweighted SNR. In figure 5.7 we make the analogous plot for GRB 170112A. As the data contained several glitches, this plot shows a long tail of events with a very high coherent SNR, which the reweighted SNR does not have. We can also see that the coherent SNR and reweighted SNR in the on-source window, indicated by the red and blue stars respectively, are consistent with background. This plot shows the importance of the reweighting. Suppose there had been a GW in the on-source of GRB 170112A that had a coherent SNR of  $\sim 30$  and a reweighted SNR of  $\sim 23$ , as GW 170817 did. In this case, the GW would *not* have had a higher coherent SNR than all of the off-source trials, making a detection claim based on the coherent SNR



**Figure 5.5: Coherent SNR vs Network  $\chi^2$  for GRB 170112A.** Here we plot the network  $\chi^2$  against the coherent SNR (left) and reweighted SNR (right) for GRB 170112A. The data around this GRB contained several glitches, which are apparent from the triggers with a high network  $\chi^2$  and coherent SNR. As we can see, these triggers are downweighted appropriately, such that the reweighted SNR contains no peaks.

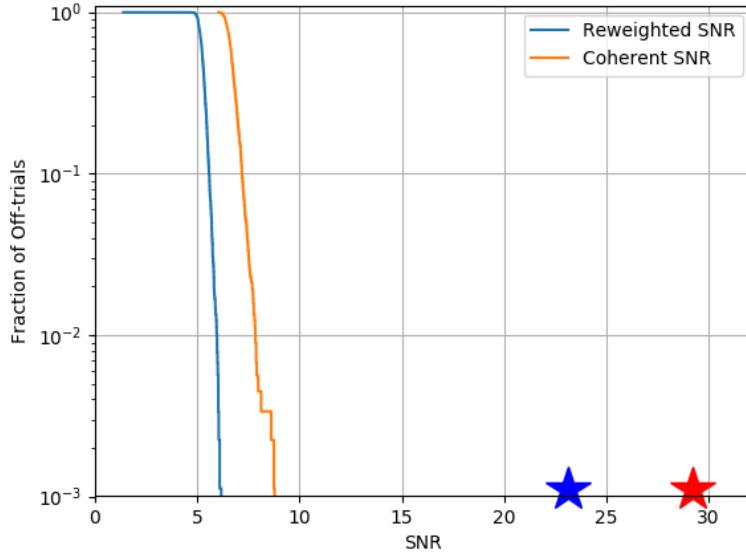


Figure 5.6: **Loudest Event per 6-second Trial for GRB 170817A.** Here we see the peak coherent SNR (blue line) and reweighted SNR (orange line) in each 6-second trial for the GRB 170817A analysis, with the stars indicating the GW. Again, we can see that the on-source has a much higher reweighted SNR than any of the off-source trials. Also note that the tail of events with a high coherent SNR of about 7-9, seen in figure ??, does not appear in this plot. This shows that the reweighting is lowering the significance of glitches, as it is supposed to do. **fix caption**

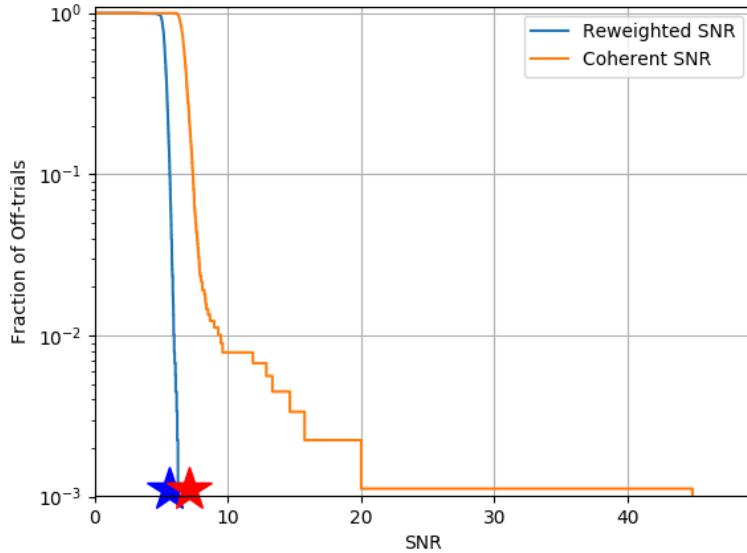
impossible. But using the reweighted SNR it would have been far more significant than anything in the background, which, with more off-source trials, could allow us to claim a detection.

We also injected BNS waveforms with an opening angle of up to  $30^\circ$  into the off-source data of GRB 170817A to see how well they were recovered. In figure 5.8 we have plotted the distance to the injection against time. From this we can see that the BNS waveforms are detectable up to about 200 Mpc. In figure 5.9, we show an analogous plot from the O2 PyGRB analysis of GRB 170817A. We can see that the old code could also detect these waveforms up to a distance of about 200 Mpc.<sup>14</sup>

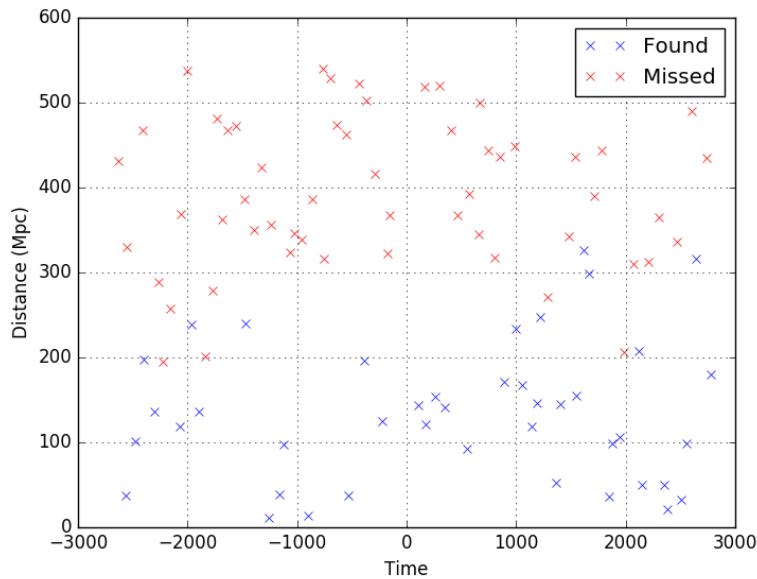
In figure 5.10 we plot the chirp mass of the injection against the distance. Comparing to the analogous plot from O2 PyGRB search, shown in figure 5.11, we see that the results are comparable, with an average range of about 200 Mpc, increasing slightly for higher chirp mass systems.

---

<sup>14</sup>The new code cannot yet provide a FAP for an injection as the old code can.



**Figure 5.7: Loudest Event per 6-second Trial for GRB 170112A.** Here we see the peak coherent SNR (blue line) and reweighted SNR (orange line) in each 6-second trial for the GRB 170112A analysis. The peak coherent and reweighted SNR in the on-source trial are indicated by the red and blue stars respectively. We see that the on-source results are consistent with background. We can also see that reweighting the SNR removed the long tail of high SNR glitches for this analysis.



**Figure 5.8: Injection Distance against Time.** 170817 run

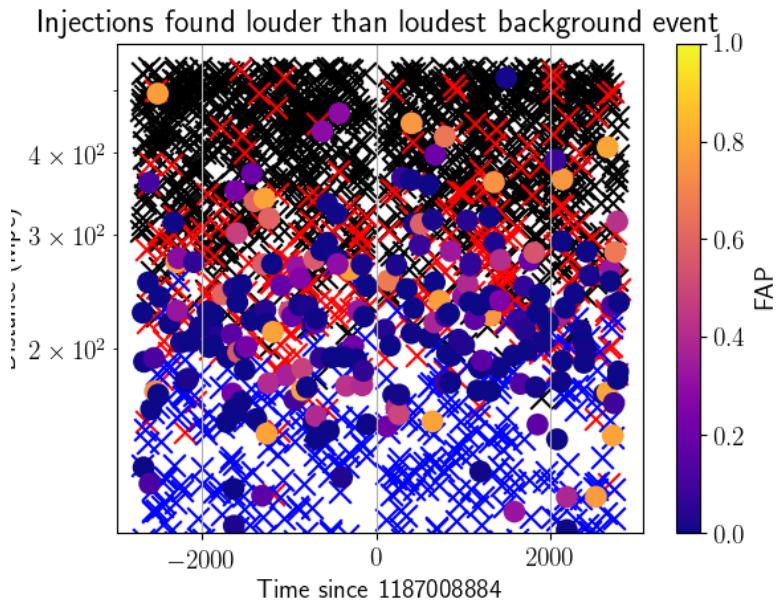


Figure 5.9: **Injection Distance against Time.** This is the distance (MPc) vs time plot for the BNS injection run in the PyGRB O2 analysis of GRB 170817A. Blue crosses indicate that the injection was found and was more significant than event in the background data. Red crosses indicate that a trigger was found that was coincident with the injection, but it was vetoed. Black crosses indicate that no trigger was found that was coincident with the trigger. Coloured circles indicate that the injection was found but was not louder than all of the background, and in this case the circle colour indicates the FAP of the trigger. We can see that nearby injections are almost always found, and typically with a low FAP. More distant injections tend to be vetoed, missed completely, or have a relatively high FAP. It is at a distance of about 200 Mpc that injections start to be missed.

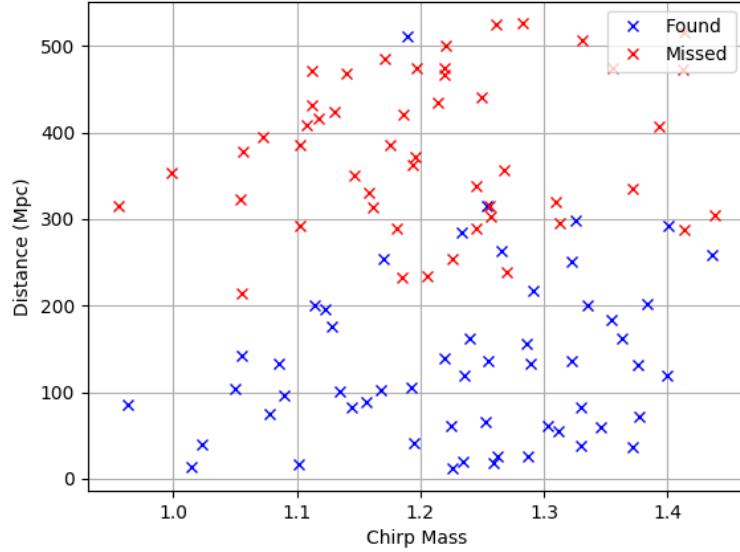


Figure 5.10: **Injection Distance against Chirp Mass.** 170817 run

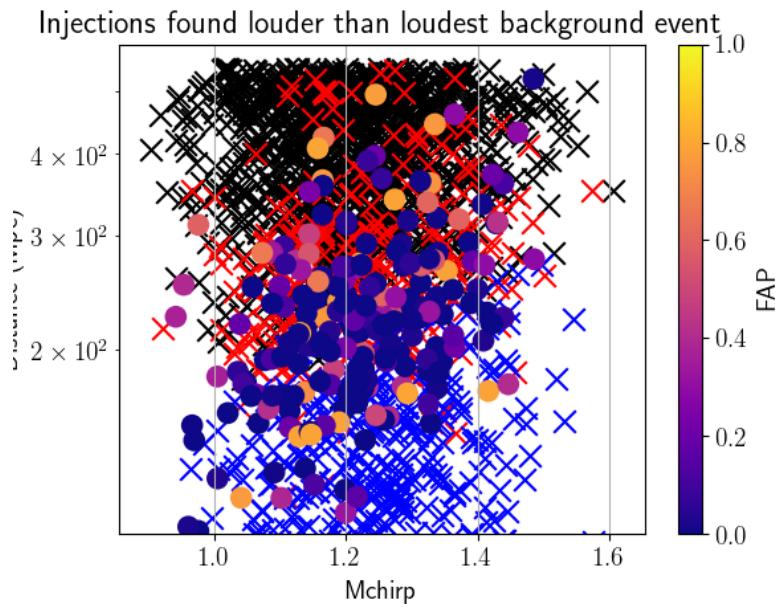


Figure 5.11: **Injection Distance against Chirp Mass.** This plot uses the same conventions and injection set as figure 5.9, but shows the distance (Mpc) against the chirp mass of the signal injections. We can see that, in general, the higher the chirp mass, the lower FAP of the injection, as we would expect.

### 5.3.2 Computational Cost

As mentioned previously, one of the primary motivations for integrating PyGRB into PyCBC was to speed up the analysis. To see the effect of the changes described above, we compare the analysis time of the new executables with the corresponding executable from the old code. The new matched filtering executable does not yet have all the functionality the old code has, and so any comparison of analysis time should be considered preliminary.

For testing, we ran as close as possible to an identical analysis with the old and new code. This meant disabling the features of the old analysis that the new code does not have. Features that were disabled include the face-on coherent analysis described in section 4.2.5, the bank and auto  $\chi^2$  tests, and timeslides. As the power  $\chi^2$  test is different in the old and new codes, we ran three tests: One with no  $\chi^2$  tests, one with only the power  $\chi^2$  test, and one with all the  $\chi^2$  tests. The analysis with no  $\chi^2$  tests will indicate the speed up in matched filtering. The test with just the power  $\chi^2$  test is a comparison between the current best analysis the new code can perform, and the old code equivalent. This test is not a fair comparison for the old code, however, as it performs the bank and auto  $\chi^2$  tests before the power  $\chi^2$  test as the coherent power  $\chi^2$  test is the most computationally expensive. We therefore also compare an analysis where the old code uses all the  $\chi^2$  tests it has available, even though the new code only uses the power  $\chi^2$  test.

We will compare the CPU runtime for several of the executables that have been rewritten. The most significant of these is the matched filtering job. This is by a large margin the most computationally expensive part of the analysis. Analysing a single GRB involves matched filtering over 1 hour of data with approximately 192000 templates. To speed this analysis up, the template bank is split into 1500 *split banks*, which are then analysed individually. There are then also the *injection jobs*, where GW signals are added to the data in order to measure the sensitivity of the analysis and ensure the pipeline is behaving correctly. In O2 there were seven different sets of injection waveforms used.<sup>15</sup> Each of these seven injection sets is split into 520 smaller injection sets, in order to speed up analysis and because the injections need to be spread throughout the data so that they do not skew the PSD. Just as with the non-injection runs, the injection jobs are sped up by splitting the template bank. However, for the injection jobs, the template bank is only split into two. The reason this is possible without being prohibitively slow is that PyGRB (both the old and new version) will only filter the data with templates that match the injections. This drastically speeds up analysis. Bringing all these factors together, we see that for a standard GRB analysis in O2, of the approximately 9000 individual jobs run, 8780 are matched filtering jobs. Of these

---

<sup>15</sup>These include one BNS set, and NSBH sets with and without precession, with and without aligned spin, and two different NSBH models **cite SEOB and STT2 paper**.

matched filtering jobs, 7280 are injection jobs and 1500 are non-injection jobs.

In either case, these matched filtering jobs can take hours. With the set up described above the jobs do not take as long as in a production run, as timeslides etc. are turned off, but they take long enough to test performance. In table 5.3.2 we see the CPU time for the old code and the new code to complete a single matched filter job, with a three detector network, with no injections, and filtering 128 templates. The left-most column indicates which  $\chi^2$  tests were used. As we can see, the new code is significantly faster than the old. Also, the use of  $\chi^2$  tests has almost no impact on the analysis time of the new code, in fact it improved the analysis time slightly. That the  $\chi^2$  tests have little impact on analysis time is not very surprising as the PyCBC  $\chi^2$  tests are highly optimised as it is a computational bottleneck for the all-sky search. It is not such a problem for PyGRB as the  $\chi^2$  tests are only calculated on triggers that have survived single detector SNR cuts, coincidence checks, and network SNR cuts. This leaves relatively few triggers for which the  $\chi^2$  test is calculated. From these results, we can conclude that the matched filtering alone is now more than twice as fast as the old code. When  $\chi^2$  tests are included in the analysis, the new code is almost five times faster than the old code. For a comparison of the matched filter job speed with injections, we only compared the analysis time with all  $\chi^2$  tests. In this case old code took 7781 seconds and the new code took 2132, more than three times faster.

It is also instructive to compare the computation time to the PyCBC all-sky matched filter job. These jobs matched filter data from just a single detector, as opposed to the three detector analysis described above. Hence, we should expect the matched filtering to take a third of the time of the new PyGRB code. The rest of the time difference between the two matched filter jobs can be accounted for by the coherent calculations performed by PyGRB. Analysing data from a single detector with 129 templates and no signal consistency tests takes the PyCBC all-sky code 2m 47s. This suggests that the coherent calculations take about a quarter of the total run time of the new PyGRB matched filtering executable.

$\chi^2$ tests used	Old code	New Code
None	24m 20s	12m 7s
Power	77m 18s	11m 28s
Power, auto, and bank	58m 37s	n/a

Table 5.1: **Time to Complete Single Matched Filter job.** Here we see how long a single matched filter job in the GRB 170817A analysis took to complete with different  $\chi^2$  tests being used. The new code does not use the bank or auto  $\chi^2$  tests. We can see that there is a substantial speed up of the analysis.

While the limiting factor on the analysis time is the matched filter jobs,<sup>16</sup> rewriting the post-processing jobs has also led to a useful speed up. In particular, the

<sup>16</sup>Both with the old and new code.

trigger combining job, which combines the results from all the different split bank matched filter jobs into one large results file, is now much faster. It previously had to be split into many smaller jobs to prevent it from exceeding the memory allowance of the computer cluster. In total, the trigger combining stage required 20 jobs of about 200 seconds each and one job of about 80 seconds. With the new code, this can be done with one job in about 70 seconds. A similar story applies to the trigger clustering stage, which used to be 20 jobs of about 60 seconds each followed by one job of about 30 seconds. This is now a single 4 second long job. It is expected that once the PyGRB rewrite is complete, the post-processing will be a negligible amount of the total run time.

## 5.4 Future Plans

We have shown that the work to integrate PyGRB into PyCBC has reached a significant milestone in being able to analyse a GRB and achieve a reasonable FAP of  $\approx 1/1000$ . We have also shown that it is significantly faster than the old code and described the new tools that are available for the PyGRB search. We end this chapter by looking at what there is left to be done to make the new pipeline as sensitive as the old one, and what can be done after this to make PyGRB even more sensitive.

In order for the new code to achieve the same confidence in a GRB trigger as the old code does, we must implement timeslides. Without timeslides, we can only achieve a FAP of about 1/1000. With timeslides implemented, a FAP of  $10^{-5}$  will be achievable. It should be noted that the relative speed increase seen with the new code is likely to be lower when using timeslides. This is because short timeslides do not require the matched-filtering step to be repeated, as mentioned in section 4.3.3, and this is the step where most of the speed increase was achieved. Long time slides do require the data to be matched filtered again, so some speed increase will be seen there.

The final thing required to make the rewrite as sensitive as the old code is to incorporate the fact that GRB GW signals are expected to be face on, as described in section 4.2.5. This led to a 3% increase in the sensitivity of the old code, and we expect the same improvement with the new code.

These changes will make the rewrite as sensitive as the old code, but we also need to be able to search over a sky patch. Some Swift GRBs are localised well enough that PyGRB can search a single sky point, but the majority require multiple sky points to be analysed. This means that the new code currently cannot analyse the majority of short GRBs. Work is currently ongoing to add this functionality, and we expect it to be incorporated into the new code in the near future.

Once the above functionality has been added to the new PyGRB pipeline, it will be faster and more sensitive than the old code. We can then turn our attention to

new methods to improve the pipeline, and the new science that can be done. At this point, it will be possible to analyse a GRB on a timescale of hours, significantly increasing the amount of science that can be done with a positive detection. We could also use PyGRB to followup well localised triggers from the all-sky search.<sup>17</sup> This could be built into the all-sky pipeline, as a hierarchical search, where the all-sky is used to determine candidates and their sky location for coherent followup. With more detectors, such as KAGRA and LIGO India, the case for a hierarchical search strengthens as localisation improves with the number of detectors and so does the power of coherent analysis.

We can also start to use some more of the PyCBC tools that have so far been unavailable to PyGRB. For example, PyCBC can downweight triggers based on properties of the template that found the trigger **cite**. This works because certain templates have been found to be more likely to produce a high SNR trigger when filtered on detector noise than other templates. There are also new  $\chi^2$  tests that use sine-Gaussian waveforms in order to better reject certain kinds of short duration glitches that are particularly common in the LIGO detectors. **cite**

There are also plans to make an astrophysically motivated detection statistic. This means that triggers that astrophysical models predict are less common will be downweighted, while sources that are expected to be common will not be. This is a natural next step for a coherent search, as the coherent SNR determines whether a trigger is consistent with the state of the detector network at the time of the trigger, but does not determine if the trigger is consistent with the current best models of the GW source populations.

---

<sup>17</sup>As with a network of three or more detectors, the localisation can be better than the localisation of some GRBs.

# Chapter 6

## A Search for Unmodelled Gravitational Wave Signals using Machine Learning

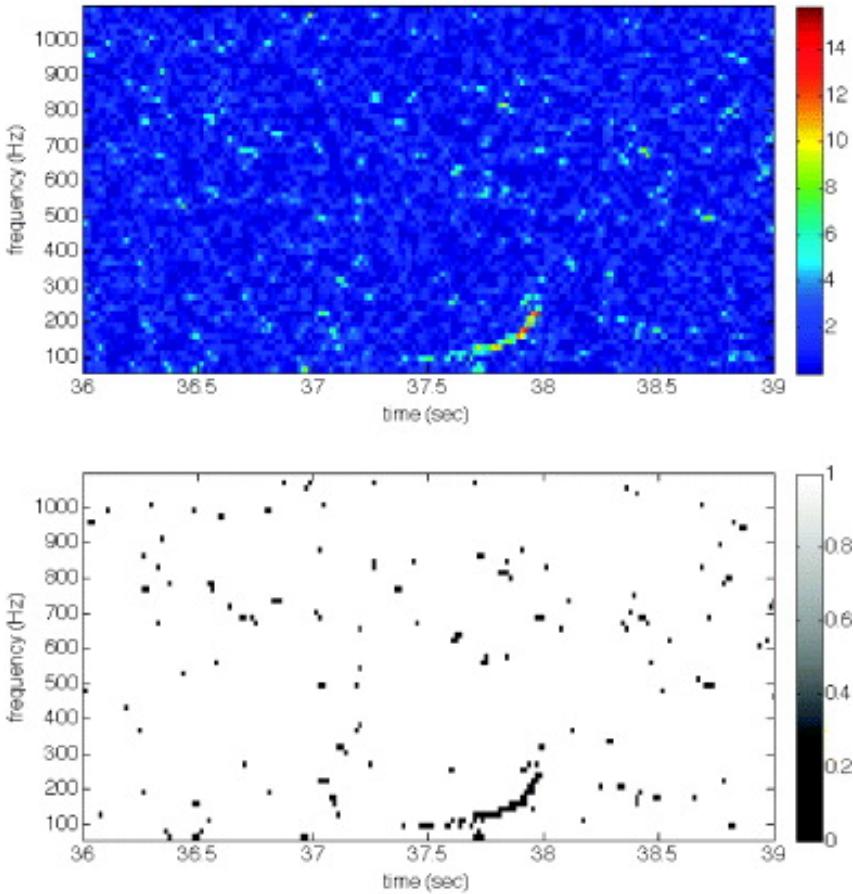
In previous chapters we focused on searching for gravitational wave (GW) signals from binary mergers using a matched-filter search. Matched filter searches are very sensitive, but they require theoretical waveforms to have been produced in advance. For this reason, matched filtering is not appropriate for some of the most interesting potential sources of GWs, e.g. core collapse supernova, as the GW morphology is not known. For this reason, it is important to develop unmodelled searches as well. Unmodelled searches, otherwise known as *burst* searches, look for coherence between the data streams of multiple interferometers. We will consider the case of a burst search where the sky position of the candidate source is known. This allows us to calculate the relative time of arrival and the relative signal power in each detector.

In this chapter, we will first discuss how an existing targeted burst search, called *X-pipeline*, searches for gravitational wave bursts (GWBs). We will then see how we can improve this pipeline by using *multivariate analysis* (MVA) to rank candidate GW events.

### 6.1 X-pipeline

In this section we provide a brief summary of X-pipeline , a targeted search for GWBs. For full details, see **cite sutton was**. It uses the sky location of an astrophysical event, such as a GRB, to coherently combine the data streams of each detector in the network. There are two types of coherent data stream made by X-pipeline :

- *Signal streams*, which increase the power of a GW signals relative to noise.
- *Null streams*, which reduce the power of true GW signals relative to noise.

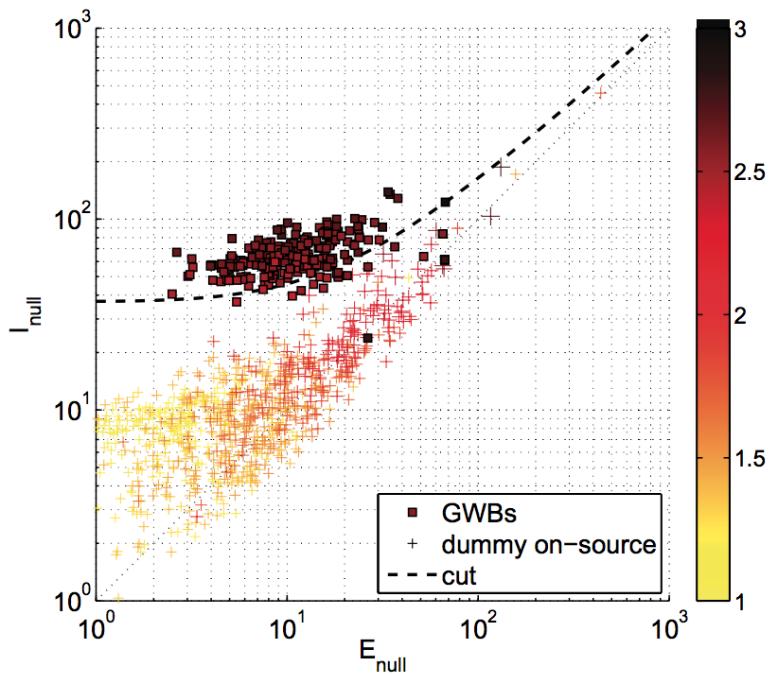


**Figure 6.1: X-pipeline Time-Frequency Map** This figure shows a time-frequency map from X-pipeline for a  $1.4 - 10M_{\odot}$  NSBH merger using simulated background from the two Hanford detectors. The top figure shows a coherent signal stream called the *standard likelihood*  $E_{\text{SL}}$  and the bottom figure shows the top 1% of pixels. [cite xp paper]

X-pipeline makes time-frequency maps from these coherent data streams. A group of neighbouring high energy pixels in a time-frequency map of a signal stream is called a *trigger* (see figure 6.1). X-pipeline cuts triggers based on correlations of the various signal and null data streams (see figure 6.2). The position of these cuts is set to give the best detection efficiency at a fixed false alarm rate, determined using a subset of the triggers that are used for tuning.<sup>1</sup>

We will begin by deriving the *standard likelihood*, a signal stream used by X-pipeline, and a corresponding null stream. Then we will see how X-pipeline uses these with incoherent data streams to reject background triggers and measure the significance of the surviving triggers. We will end this section by looking at the X-pipeline search results from the most recent observing run.

<sup>1</sup>For the MVA pipeline, there are no cuts. Instead the machine learning algorithm ranks triggers just that background triggers have a low score and GWs have a high score, as we will see in 6.2.



**Figure 6.2: X-pipeline Background Rejection Test.** This figure shows an example of X-pipeline background rejection. The axes show two of the statistics that X-pipeline calculates. Specifically, the x-axis shows the coherent null energy and the y-axis shows the incoherent null energy (see section 6.1.4 for more details on these statistics). The red squares show simulated GW signals, and the crosses show background triggers. The colour bar shows the base 10 logarithm of the significance of each trigger. The injection amplitude plotted is chosen such that approximately 90% of injections will survive the cut. Hence, the cut eliminates most of the noise but only a few signals. [ [cite X-pipeline paper](#) ]

### 6.1.1 Burst Search Background

Suppose we have a network of  $D$  detectors. A GW, described by  $h_+(t)$  and  $h_\times(t)$ , passes through the Earth from direction  $\hat{\Omega}$ . We describe the sensitivity of detector  $\alpha \in \{1, \dots, D\}$  to the plus and cross polarisations using the *antenna response functions*, denoted  $F_\alpha^+(\hat{\Omega})$  and  $F_\alpha^\times(\hat{\Omega})$ . The position of detector  $\alpha$  is denoted by  $\mathbf{r}_\alpha$  and  $n_\alpha(t)$  is the noise in this detector. The detector output  $d_\alpha(t)$  is then given by

$$d_\alpha(t + \Delta t_\alpha(\hat{\Omega})) = F_\alpha^+(\hat{\Omega})h_+(t) + F_\alpha^\times(\hat{\Omega})h_\times(t) + n_\alpha(t + \Delta t_\alpha(\hat{\Omega})). \quad (6.1)$$

Here  $\Delta t_\alpha$  is the time taken for the GW to reach the detector from some arbitrary reference point<sup>2</sup>  $\mathbf{r}_0$

$$\Delta t_\alpha(\hat{\Omega}) = \frac{1}{c}(\mathbf{r}_0 - \mathbf{r}_\alpha) \cdot \hat{\Omega}. \quad (6.2)$$

From now on we will suppress explicit mention of the reference point  $\mathbf{r}_0$  or the time delay  $\Delta t_\alpha$  on the understanding that detector outputs need to be time-shifted by an appropriate amount.

In reality, detector outputs are not continuous but sampled discretely. The discrete Fourier transform  $\tilde{x}[k]$  of the time series  $x[j]$ , and its inverse, are given by

$$\tilde{x}[k] = \sum_{j=0}^{N-1} x[j]e^{-i2\pi jk/N}, \quad x[j] = \frac{1}{N} \sum_{k=0}^{N-1} \tilde{x}[k]e^{i2\pi jk/N}. \quad (6.3)$$

For sampling rate  $f_s$  and  $N$  data points in the time domain, we convert continuous to discrete notation by using

$$x(t) \rightarrow x[j] \quad (6.4)$$

$$\tilde{x}(f) \rightarrow f_s^{-1}\tilde{x}[k] \quad (6.5)$$

$$\int dt \rightarrow f_s^{-1} \sum_j \quad (6.6)$$

$$\int df \rightarrow f_s N^{-1} \sum_k \quad (6.7)$$

$$\delta(t - t') \rightarrow f_s \delta_{jj'} \quad (6.8)$$

$$\delta(f - f') \rightarrow N f_s^{-1} \delta_{kk'} . \quad (6.9)$$

For example, the one-sided noise spectral density for a detector with noise  $n(t)$  can be written in continuous form as<sup>3</sup>

$$\langle \tilde{n}_\alpha^*(f) \tilde{n}_\beta(f') \rangle = \delta_{\alpha\beta} \delta(f - f') \frac{1}{2} S_n(f) \quad (6.10)$$

<sup>2</sup>The center of the Earth is a fairly intuitive choice for a worldwide detector network, but in practice it is more computationally efficient to use one of the detectors as the reference point.

<sup>3</sup>Note that this is the same as 4.14 up to a factor of 1/2.

where the angle brackets indicate an average over the noise. In the discrete notation listed above, this becomes

$$\langle \tilde{n}_\alpha^*[k] \tilde{n}_\beta[k'] \rangle = \frac{N}{2} \delta_{\alpha\beta} \delta_{kk'} S_\alpha[k] . \quad (6.11)$$

We will be working with the *noise-spectrum-weighted* quantities, defined by

$$\tilde{d}_{w\alpha}[k] = \frac{\tilde{d}_\alpha[k]}{\sqrt{\frac{N}{2} S_\alpha[k]}} \quad (6.12)$$

$$\tilde{n}_{w\alpha}[k] = \frac{\tilde{n}_\alpha[k]}{\sqrt{\frac{N}{2} S_\alpha[k]}} \quad (6.13)$$

$$F_{w\alpha}^{+,\times}(\hat{\Omega}, k) = \frac{F_\alpha^{+,\times}(\hat{\Omega})}{\sqrt{\frac{N}{2} S_\alpha[k]}} \quad (6.14)$$

The normalisation of the whitened data is

$$\langle \tilde{n}_{w\alpha}^*[k] \tilde{n}_{w\beta}[k'] \rangle = \delta_{\alpha\beta} \delta_{kk'} . \quad (6.15)$$

In what follows we will only use the whitened detector data, noise, and antenna patterns, and drop the subscript  $w$  for clarity. In vector notation, we can write 6.1 as

$$\tilde{\mathbf{d}} = \mathbf{F}\tilde{\mathbf{h}} + \tilde{\mathbf{n}} \quad (6.16)$$

where  $\mathbf{F} = [\mathbf{F}^+ \ \mathbf{F}^\times]$  and  $\tilde{\mathbf{h}} = [\tilde{h}_+ \ \tilde{h}_\times]^T$ .

### 6.1.2 Standard Likelihood

With the notation and conventions laid out in the previous section, we now derive the *standard likelihood*,<sup>4</sup> a coherent signal stream. We start by supposing a GW  $\tilde{\mathbf{h}}$  passes through a detector from a given direction. Assuming the detector noise is Gaussian, the probability of attaining whitened detector output  $\tilde{\mathbf{d}}$  in one time-frequency pixel is given by

$$P(\tilde{\mathbf{d}}|\tilde{\mathbf{h}}) = \frac{1}{(2\pi)^{D/2}} \exp \left[ -\frac{1}{2} \left| \tilde{\mathbf{d}} - \mathbf{F}\tilde{\mathbf{h}} \right|^2 \right] . \quad (6.17)$$

For a set  $\{\tilde{\mathbf{d}}\}$  of  $N_p$  time-frequency pixels, labeled by  $k$ , we have

$$P(\{\tilde{\mathbf{d}}\}|\{\tilde{\mathbf{h}}\}) = \frac{1}{(2\pi)^{N_p D/2}} \exp \left[ -\frac{1}{2} \sum_k \left| \tilde{\mathbf{d}}[k] - \mathbf{F}[k]\tilde{\mathbf{h}}[k] \right|^2 \right] . \quad (6.18)$$

---

<sup>4</sup>This derivation has a lot in common with the derivation of the coherent SNR in section 4.2.2. The primary difference being that here we consider the likelihood of obtaining a certain collection of time-frequency pixels, whereas before we considered the likelihood of measuring a certain SNR for a template.

By comparing this value to the probability that the detector produces this output in the absence of any GW signal, we can calculate a likelihood of the signal being a real GW signal.

The *Likelihood Ratio*  $L$  is the log of the probability that the detector network will have output  $\tilde{\mathbf{d}}$  in the presence of GW  $\tilde{\mathbf{h}}$  divided by the probability of obtaining the same output in the absence of a gravitational wave ( $\tilde{\mathbf{h}} = 0$ )

$$L = \ln \frac{P(\{\tilde{\mathbf{d}}\}|\{\tilde{\mathbf{h}}\})}{P(\{\tilde{\mathbf{d}}\}|0)} = \frac{1}{2} \sum_k \left[ |\tilde{\mathbf{d}}|^2 - |\tilde{\mathbf{d}} - \mathbf{F}\tilde{\mathbf{h}}|^2 \right]. \quad (6.19)$$

For the above analysis to be applied, we would need to know the waveform  $\tilde{\mathbf{h}}$  in advance. For GRBs and other unmodelled searches, this is not possible. One way to handle this problem is to fit the waveform in each time-frequency pixel to the data in such a way as to maximise the likelihood ratio. Hence we have

$$0 = \frac{\partial L}{\partial \tilde{\mathbf{h}}} \Big|_{\tilde{\mathbf{h}}=\tilde{\mathbf{h}}_{\max}}. \quad (6.20)$$

Solving this, we find

$$\tilde{\mathbf{h}}_{\max} = (\mathbf{F}^\dagger \mathbf{F})^{-1} \mathbf{F}^\dagger \tilde{\mathbf{d}} \quad (6.21)$$

where the superscript dagger  $^\dagger$  denotes the conjugate transpose.

Calculating the likelihood ratio for  $\tilde{\mathbf{h}}_{\max}$  gives us the *Standard Likelihood*<sup>5</sup>  $E_{\text{SL}}$

$$E_{\text{SL}} = 2L(\tilde{\mathbf{h}}_{\max}) = \sum_k \tilde{\mathbf{d}}^\dagger \mathbf{P}^{\text{GW}} \tilde{\mathbf{d}} \quad (6.22)$$

where

$$\mathbf{P}^{\text{GW}} \equiv \mathbf{F}(\mathbf{F}^\dagger \mathbf{F})^{-1} \mathbf{F}^\dagger. \quad (6.23)$$

We can see from equation 6.1 that the contribution made to the data output by a passing GW from fixed sky location is restricted to the subspace spanned by  $\mathbf{F}_+$  and  $\mathbf{F}_\times$ . Therefore the energy in this subspace is the energy that is consistent with a GW from a given sky location. We can show that  $\mathbf{P}^{\text{GW}}$  is a projection operator, projecting the data into this same subspace. The standard likelihood maximises the energy in this subspace, and so is the maximum energy contained in the whitened data that is consistent with a GW from the given sky location.<sup>6</sup> This is an example of the coherent signal streams that X-pipeline uses.

---

<sup>5</sup>Also known as the *maximum likelihood*.

<sup>6</sup>In practice, some of this energy will be due to noise. We say that the energy due to noise is *inconsistent* with a GW signal. The rest of the energy must be due to the GW, so we say it is *consistent* with the GW.

### 6.1.3 Null Energy

We can use the standard likelihood to construct a null stream. We start with the total energy in the data, given by

$$E_{\text{tot}} = \sum_k |\tilde{\mathbf{d}}|^2 . \quad (6.24)$$

This is an incoherent statistic as it contains auto-correlation terms but no cross-correlation terms, i.e. each detector is treated individually. If we subtract the standard likelihood from the total energy, we obtain the *null energy*<sup>7</sup>

$$E_{\text{null}} \equiv E_{\text{tot}} - E_{\text{SL}} = \sum_k \tilde{\mathbf{d}}^\dagger \mathbf{P}^{\text{null}} \tilde{\mathbf{d}} . \quad (6.25)$$

This is the energy that is inconsistent with a GW from given sky location, and must therefore be associated with noise. This is the minimum amount of energy in the whitened data that is inconsistent with the GW. It is an example of the null streams used by X-pipeline.

This shows one of the key advantages of coherent analysis. If we analysed our data incoherently, we would be working with just the total energy. By using coherent methods, we can project the whitened data into the subspace spanned by  $\mathbf{F}^+$  and  $\mathbf{F}^\times$ , removing some fraction of the noise without removing much of the signal energy.<sup>8</sup> The drawback is that if the sky position is not known in advance, then the analysis needs to be repeated for a set of directions that span the entire sky ( $\gtrsim 10^3$  directions), each with different antenna response functions  $\mathbf{F}^+$  and  $\mathbf{F}^\times$ . This will slow down analysis and increase the false alarm probability (FAP).

### 6.1.4 Incoherent Energy and Background Rejection

The diagonal elements of (6.25) are auto-correlation terms, and the other elements are cross-correlation terms. The auto-correlation part of the null energy is called the *incoherent null energy*, and denoted by

$$I_{\text{null}} = \sum_k \sum_\alpha P_{\alpha\alpha}^{\text{null}} |\tilde{d}_\alpha|^2 . \quad (6.26)$$

Background triggers are typically not correlated between the different detectors of the network, so the cross-correlation terms are small relative to the auto-correlation terms. This means that for glitches, we have

$$E_{\text{null}} \approx I_{\text{null}} . \quad (6.27)$$

---

<sup>7</sup>This is analogous to the null SNR described in section 4.2.4.

<sup>8</sup>Again, this is analogous to the coherent SNR and the discussion in section 4.2.3.

Compare this to the case of a GW signal. This will be correlated between the detectors. By construction, the energy due to the GW does not appear in the null stream. Therefore, in the presence of a strong GW, the incoherent energy is much larger than the null energy

$$E_{\text{null}} \ll I_{\text{null}} . \quad (6.28)$$

Using (6.27) and (6.28), we see that the ratio of  $E_{\text{null}}$  and  $I_{\text{null}}$  is very different in the case of a glitch as opposed to a GW signal. We can use this to make the following cut to remove noise triggers from our sample

$$I_{\text{null}}/E_{\text{null}} > C \quad (6.29)$$

for some constant  $C > 1$ . This test does not work as well for small amplitude glitches, where the statistical fluctuations can lead to  $E_{\text{null}}$  being smaller than  $I_{\text{null}}$ . For this reason X-pipeline varies  $C$  with  $I_{\text{null}}$ , as can be seen in figure 6.2. The precise position of the cut is set to maximise performance on a set of injections and background noise triggers.

We have seen how X-pipeline uses the coherent and incoherent null stream for background rejection. X-pipeline uses this same process for other measures of coherent energy, i.e. it uses the coherent energy to define an incoherent energy and then make a cut. Specifically, it uses the coherent energy for that assumes either a plus, cross, or circularly polarised GW, which are described in detail in [cite chateraaji was and sutton](#). In section 6.2 we will see how these statistics can instead be used by a machine learning algorithm to make more intelligent cuts, using arbitrary combinations of coherent and incoherent statistics, as well as other statistics.

### 6.1.5 Tuning and Trigger Significance

In this section we will describe how the coherent statistics are tuned and how X-pipeline calculates the significance of a trigger. X-pipeline searches for GWs in a window  $[-600, +60]$  seconds around the trigger time of a typical GRB, which we call the *on-source* data. For GRBs with a  $T_{90}$  value greater than 60 seconds, the on-source is extended to  $[-600, +T_{90}]$  seconds around the trigger time. To tune the cuts, GW signal injections are placed into the on-source data. The cuts that are chosen are the ones that minimise the amplitude of the GW injections for a fixed detection percentage and at a given false alarm rate. For example, the user could decide that a FAP of 1% is acceptable, and that the cuts should be chosen to minimise the injection scale at which 90% of the injections are detected.

In order to measure the significance of a trigger, approximately 1.5 hours of data either side of the on-source window is also analysed, called the *off-source* data. The off-source data is split into trials of equal length to the on-source window. The FAP is calculated as the fraction of off-source trials that have a higher value for

the detection statistic than the loudest trigger in the on-source. X-pipeline creates several ranking statistics, such as a Bayesian-inspired modification to the standard likelihood **cite was** and a statistic based on power law distributed noise **cite S6 GRB paper**. The ranking statistic used as the detection statistic is the one that gives the best sensitivity at the user specified FAP. To create more trials and hence better measure the FAP, timeslides are applied to the off-source to simulate more data.<sup>9</sup>

### 6.1.6 O2 X-pipeline Search

The Advanced LIGO second observing run started in November 2016 and ended in August 2017, with Advanced Virgo joining for the last month. During this time there were 242 GRBs reported by the Swift and Fermi GRB observatories, and 52 reported by the InterPlanetary Network (IPN), with many GRBs appearing in both lists. X-pipeline requires at least 660 seconds of coincident data between any two detectors in the network. This resulted in 98 GRBs being analysed by X-pipeline during O2. Data from Virgo was only used if it improved the sensitivity of the closed box analysis.

The injections used for this analysis were circular sine-gaussian (CSG) and accretion disk instability (ADI) [**citations in O2 GRB paper**] waveforms. The CSG injections are ad hoc waveforms, not based on any astrophysical model, chosen to have a total radiated energy of  $E_{\text{GW}} = 10^{-2} M_{\odot} c^2$  and a Q-factor of nine.<sup>10</sup> The ADI waveforms model the GW signal from instabilities in a magnetically suspended torus around a rapidly spinning black hole. These two waveforms are chosen as they have very different morphology and duration, with CSG waveforms lasting 0.01 – 1 second and ADI waveforms lasting for 10 – 250 seconds.

Of the 98 GRBs analysed, the only statistically significant signal was that of GRB 170817A, which was recovered with a FAP of  $3.1 \times 10^{-4}$ . The FAP of the 97 other GRBs analysed is plotted in figure 6.3. A weighted binomial test of the 5% most significant GRBs analysed by X-pipeline in O2 gave a combined p-value of 0.75, indicating that there is no evidence for a sub-threshold population of weak GRB triggers.

Plotted in 6.4 is the 90% confidence exclusion distance for each analysed GRB apart from GRB 170817A. This is the distance at which 90% of injected signals have a higher ranking statistic than the most significant event in the on-source. This can be used to rule out a burst event occurring within the exclusion distance for each GRB, and also places upper-limits on the amount of energy that GRBs emit as GWs. The CSG waveforms are not astrophysical and so have no well defined distance. The

---

<sup>9</sup>This is the same approach taken by PyGRB, as discussed in section 4.3.3. X-pipeline also uses long and short timeslides as well.

<sup>10</sup>The Q-factor corresponds to the number of cycles contained within some finite region of the sine-gaussian envelope.

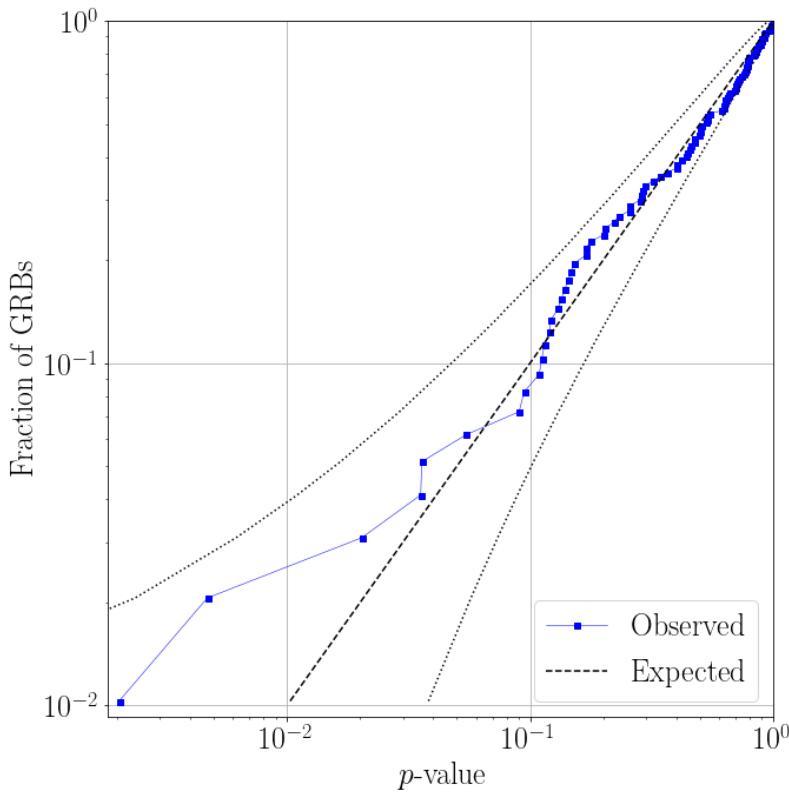


Figure 6.3: **Cumulative Distribution of p-values.** Here we plotted the p-values for every GRB analysed by X-pipeline in O2 apart from GRB 170817A. Also plotted is the expected distribution and the  $2\sigma$  deviation. The results are consistent with the no-signal hypothesis. [cite O2 GRB paper]

reported distance to the source of a CSG GW with energy  $E_{\text{GW}}$  and central frequency  $f_0$  is given by the formula cite <https://arxiv.org/pdf/1304.0210.pdf>

$$d^2 = \frac{5}{2} \frac{G}{\pi^2 c^3} \frac{E_{\text{GW}}}{f_0^2 h_{90}^2} \quad (6.30)$$

where  $h_{90}$  is the root-sum-square amplitude  $h_{\text{rss}}$  of the trigger at which 90% of injections are detected. The  $h_{\text{rss}}$  is a standard measure of amplitude used in burst searches and is given by

$$h_{\text{rss}} = \sqrt{\int dt (h_+^2(t) + h_x^2(t))} . \quad (6.31)$$

The exclusion distance can also be used as a measure of search sensitivity to burst signals. Comparison with the X-pipeline results from O1 [cite O1 paper] shows that the sensitivity is largely unchanged.

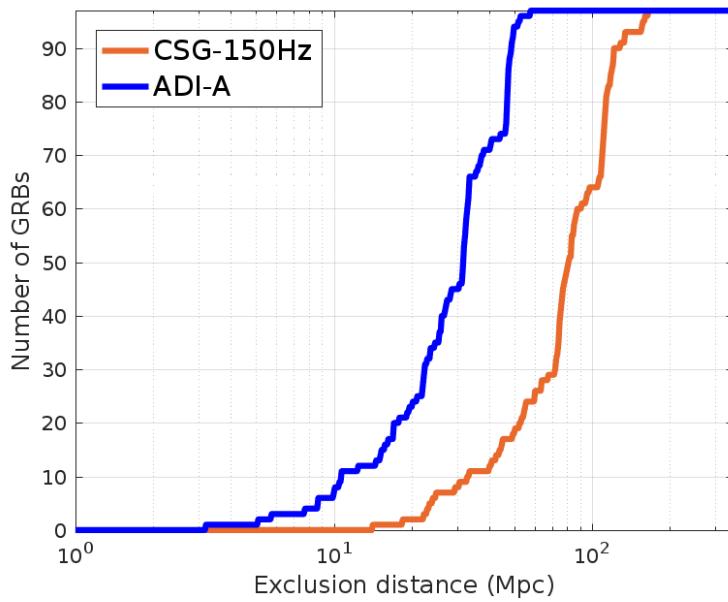


Figure 6.4: **Cumulative Distribution of Exclusion Distance.** Here we plotted the 90% exclusion distance for every GRB analysed by X-pipeline in O2 apart from GRB 170817A. This is the distance to which 90% of injections can be recovered with a significance greater than the loudest event in the on-source. [cite **O2 GRB paper**]

## 6.2 Machine Learning

We have seen how X-pipeline makes cuts on coherent statistics to distinguish between noise and GW signals. X-pipeline uses pairs of coherent and incoherent energy for background rejection, as in (6.29). These cuts are effective and well motivated<sup>11</sup> but they cannot explore the full parameter space, looking for patterns using all of the statistics at our disposal. In this section we discuss how to use machine learning to achieve this. The software we use is the Toolkit for Multivariate Analysis (TMVA) package in the ROOT data analysis framework. We call the pipeline that uses machine learning on X-pipeline triggers XTMVA. For details on the TMVA software package, see **cite TMVA user guide**. This pipeline has been previously reported on<sup>12</sup> in **cite toms paper**.

### 6.2.1 Supervised Machine Learning

The type of Machine learning we use is called *supervised machine learning*. Supervised machine learning algorithms are trained on data which has already been classified. They can then be shown a new, not-classified data point and determine the appropriate classification. Supervised machine learning requires data to be in

<sup>11</sup>As we saw in section 6.1.4

<sup>12</sup>Though they did not call the pipeline XTMVA.

Label	loghbayesiancirc	standard	circenergy	circinc
Background	12.3128	58.3523	44.7196	24.9015
Background	12.0349	67.5344	41.7045	22.4848
Signal	18.2145	59.8136	53.3320	22.0601
Signal	43.7113	123.9194	118.9774	43.9234
Signal	6422.1467	14426.9124	14167.2933	4991.7876

Table 6.1: **Example MVA training data.** Each event has a label and several attributes. The training sets we actually use have up to 15 attributes and thousands of events. The labels are various coherent and incoherent measures of energy, as well as trigger properties such as the bandwidth and duration. A full list is given in table 6.2.1.

a particular format (see table 6.1). It is a list of *events*, each with a *label* and corresponding *attributes*. The machine learning algorithm builds a *classifier* that can determine the label of an event when given the event’s attributes. In the case of our GW search, the events are the triggers returned by X-pipeline (see section 6.1). The labels are *signal* or *background*, and the attributes are the values of the signal and null data streams for those triggers and some statistics describing the time-frequency properties of the trigger. A full list of the attributes used, together with a short description of each attribute can be seen in table 6.2. The signal triggers are generated by injecting signal waveforms into the data. Background triggers are triggers that do not coincide with an injected signal and are typically formed from timeslides and using data outside the on-source window.

The signal and background trigger sets are each divided into two subsets: The *training set* and the *testing set*. The training set is used to build the classifier, while the testing set is used to measure the accuracy of the trained classifier. If the classifier performs much better on the training set than on the testing set, then the classifier is *overtrained*. This means that the classifier has learned the properties of precisely the signals and noise in the training set, rather than learning the general properties of signals and noise. When an overtrained classifier is used on a different data set with different noise properties (such as the testing set), it will perform poorly.

Statistic	Description
<code>loghbayesiancirc</code>	A likelihood ratio based on Bayesian methods, for the hypothesis of a circularly polarised GW vs Gaussian noise [cite was]
$E_{\text{SL}}$	The maximum energy in the whitened data that is consistent with a GW from a given sky location.
$E_{\text{null}}$	The minimum amount of energy in the whitened data that is inconsistent with a GW from a given sky location. Given by $E_{\text{SL}} - E_{\max}$ .
$I_{\text{null}}$	The sum of the autocorrelation terms of $E_{\text{null}}$ .
$E_{\text{circ}}$	The maximum energy in the whitened data consistent with a circularly polarised GW from a given sky location.
$I_{\text{circ}}$	The sum of the autocorrelation terms of the $E_{\text{circ}}$ .
$E_{\text{circnull}}$	The energy in the whitened data that is consistent with a GW but inconsistent with a circularly polarised GW from a given sky location. Given by $E_{\text{tot}} - E_{\text{circ}}$ .
$I_{\text{circnull}}$	The sum of the autocorrelation terms of $E_{\text{circnull}}$ .
$E_{\text{H1}}$	The energy in the H1 interferometer.
$E_{\text{L1}}$	The energy in the L1 interferometer.
$E_{\text{V1}}$	The energy in the V1 interferometer.
Number of Pixels	The number of pixels in the cluster.
Power law	The standard likelihood defined for the case where the noise is power law distributed instead of Gaussian distributed.
Duration	Time duration of the trigger in seconds.
Bandwidth	The frequency range spanned by the trigger in Hertz.
Sky Position	The sky position of the trigger.

Table 6.2: **MVA Attributes.** The attributes used by the machine learning classifier.

### 6.2.2 Boosted Decision Trees

A *Decision Tree* is a simple type of classifier. It is a flowchart of true/false statements about a trigger's attributes to determine the correct label. For example, consider the decision tree shown in figure 6.5. Here the attributes are labeled as the components of a vector  $x$ . We start at the top node and work downwards. If the statement in the node is true then we follow the branch to the left and if not then we follow the branch to the right. We then consider the statement at the end of whichever branch we follow. We continue this process until we reach a *leaf node*, which has no branches and contains a final classification for the trigger.

We can improve the performance of the classifier by using an *ensemble* (or *forest*) of trees. This means training multiple distinct trees, with each tree independently classifying the trigger. The final classification of each trigger is a normalised sum of the outputs of each tree, with +1 corresponding to signal, and -1 corresponding to background. This leads to different regions of the parameter space having different MVA scores, as can be seen in fig.6.6. The higher the score, the more likely an event is to be a signal.

To train a decision tree we must determine the best variable and cut for each decision node, as well as the correct label for each leaf node. Each of these values is set by brute force: trying every possible cut in some discrete range to get the best performance on the training events. Ensemble methods work best when each classifier in the ensemble is independent of the others, so training every tree on the same events is not going to give optimal results. For this reason each tree is trained on some subset of the training set. We could pick events at random to form these subsets, but a more powerful method is to use *Adaptive Boosting*.

When using adaptive boosting, each event in the training set is given a probability that it will be selected to train the next tree. Initially each event in the training set has the same probability of being selected. Then, after each tree is trained, the probabilities are updated such that misclassified events are more likely to be selected to train the next tree. In this way, the ensemble become gradually more effective at classifying those events that are most difficult to classify. To be precise, all the *misclassified* events have their probability of being selected for training the next tree multiplied by a *boost weight* before being renormalised. The boost weight is given by **cite tmva user guide**

$$\alpha = \frac{1 - \text{err}}{\text{err}} \quad (6.32)$$

where err is the *misclassification rate*, the fraction of events misclassified by the *previous* tree. Note that the boost weight is greater than one for  $\text{err} < 1/2$ , and that as the error gets smaller, the boost weight increases<sup>13</sup>. The effect of this is that each new tree is more likely to be trained on the events that are most difficult

---

<sup>13</sup>The misclassification rate is always less than half when there are only two labels as otherwise the algorithms would simply swap the labels around.

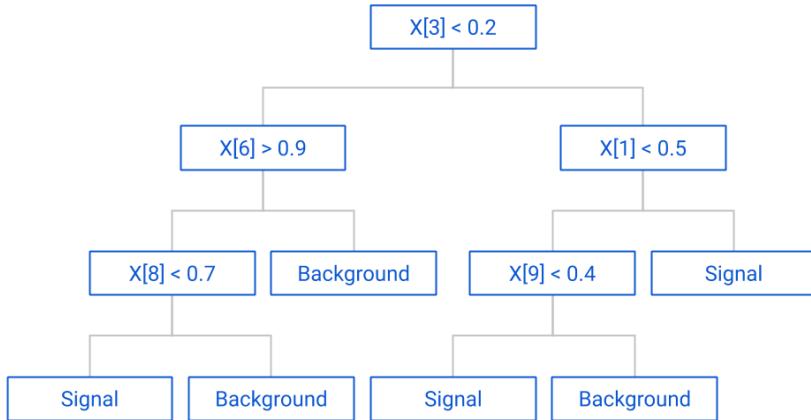


Figure 6.5: **Schematic Decision Tree.** To determine if a trigger is a signal or noise event the tree makes a series of cuts on the attributes  $x[N]$ . If the inequality in a node is true, then the next node is the branch to the left. Otherwise the next node is the one to the right. The properties of the tree, such as the number of layers it has, are set by the user (see section 6.2.4).

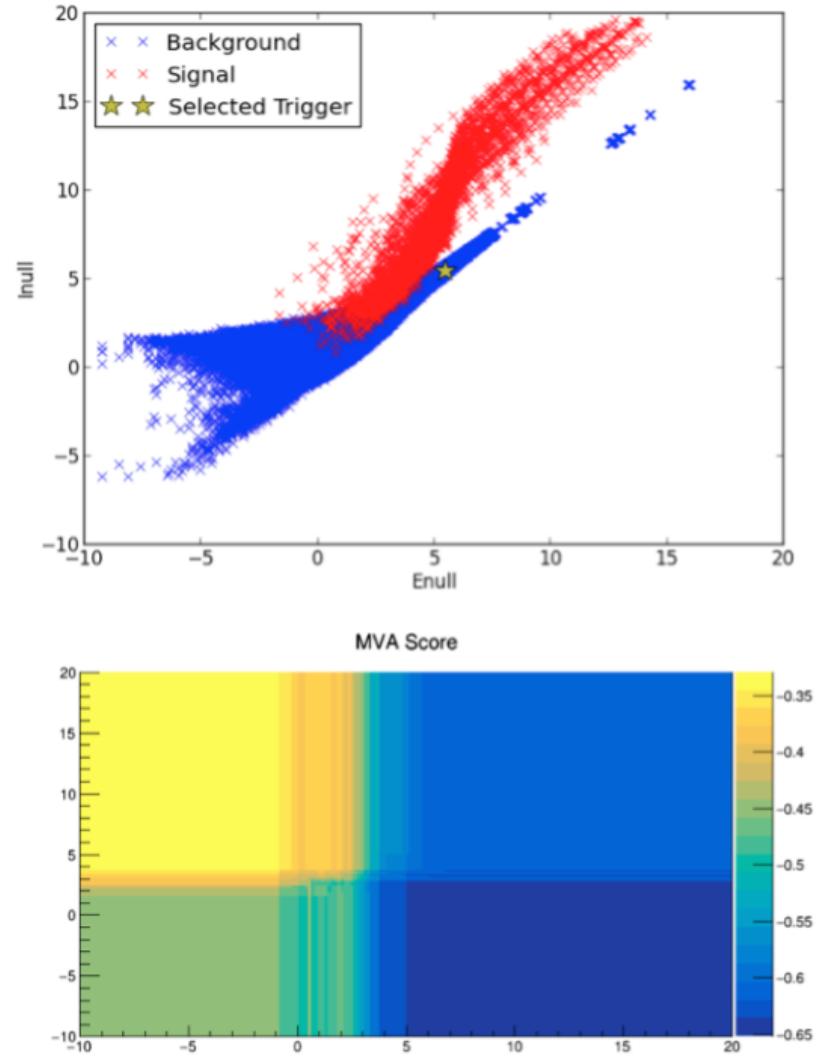
to classify. For example, if the previous tree only misclassified one event, then the misclassification error will be very small and the boost weight will be very large. This boost weight will then only be applied to the one misclassified event, making it almost certain that this event will be selected to train the next tree.

It can be useful to slow down the rate at which the probabilities are updated. To understand why, suppose that a difficult-to-classify event is, by chance, correctly classified by the first few trees. It will now have a low probability of being selected to train future trees. If we could slow down the rate at which the probabilities are updated, then only those events that are consistently classified correctly should end up with a low probability of being selected for training the next tree. This is done by adding an exponent  $\beta < 1$  to the boost weight, so that  $\alpha \rightarrow \alpha^\beta$ . The smaller  $\beta$  is, the slower the probabilities are updated. For this reason, we call  $\beta$  the *learning rate*.

The ensemble output is also changed, so that it is weighted rather than being a simple sum. If the output of the  $i$ th tree is given by  $h_i(\mathbf{x})$ , with  $\mathbf{x}$  being the event attributes, then the ensemble output is given by **cite tmva user guide**

$$H(\mathbf{x}) = \frac{1}{N} \sum_{i=0}^N \ln(\alpha_i) h_i(\mathbf{x}) \quad (6.33)$$

where  $N$  is the number of trees in the ensemble. In this way, greater significance is given to trees with lower misclassification rates. It is common practice when using a forest of decision trees such as this to say that values of  $H$  below 0.5 are background and those above 0.5 are signal. We determine classification in a slightly different



**Figure 6.6: Visualising the Classifier.** In the top plot you can see the value for  $\log(\text{Enull})$  and  $\log(\text{Inull})$  for all the signal and background training data used to build the classifier. We chose one of these events at random (indicated by the star) and varied  $\text{Enull}$  and  $\text{Inull}$  to see how it changed the MVA score, indicated by the colour in the bottom plot. As we can see, increasing  $\text{Inull}$  and decreasing  $\text{Enull}$  leads to the event being more likely to be classed as a signal. This is akin to the X-pipeline cut shown in figure 6.2.

way. The trained classifier is used on the off-source trials and it returns a value of  $H$  for each trial. We then calculate a FAP by determining the fraction of off-source trials that had a larger value of  $H$  than the on-source trial. In this way, small values of  $H$  are interpreted as more likely to be background events while larger values are more likely to be a signal, but a value of  $H$  on its own is not enough to determine if a trigger is a GW signal. Instead it is used to calculate a FAP for the trigger.

### 6.2.3 Training Data Preprocessing

Extra data preprocessing is required when training the classifier. This is because the MVA will learn what a signal looks like based on the training data, and so a small contamination can cause the classifier to fail completely. For example, suppose a small amplitude signal injection is added close to a large amplitude glitch in the data. This trigger will be labeled as a signal, due to the injection, but the properties of the trigger will resemble a glitch, as the glitch has a much larger amplitude than the signal injection. In this case, the injection is essentially a background trigger that has been mislabeled as a signal. This reduces the ability of the classifier to detect real signals by making the properties of signals harder for the algorithm to learn. Even worse, it can lead to background triggers being misclassified as a GW signal.

For these reasons it is important to make sure that signal injections do not overlap in time or frequency with background triggers. To prevent this from happening we remove any injected signals that coincide with noise, a process we call *cleaning* the data. This process starts with finding all of the triggers in the data for the smallest signal injection scale. The injected waveforms are too small to be detected so all of these triggers must be background. We then increase the injection amplitude and look for triggers again. Any triggers that overlap in time or frequency with the noise triggers are then identified and removed from the from the signal set.

We must also not include injections in our signal training set that are too small to be detected, as this would again be including triggers labeled as signals but with the properties of a noise trigger. For this reason we apply a threshold on the amplitude of the signal set, so any injection below that amplitude is removed. The level of the threshold has to be set experimentally. If the threshold is set too low then we will increase the chance of a false positive and hurt our sensitivity to real signals, but if it is too high then we will limit the classifiers ability to detect low amplitude signals.

As XTMVA is to be used for the unmodelled search, for which the physical waveforms are uncertain, it is also important to ensure that the search can find waveforms that are not in the training set. There are several tools that we use to achieve this. The first is to limit the amount of information the classifier is given about the waveform morphology. The classifier cannot know the precise morphology of any waveform because the only attributes that the classifier trains on are the duration,

band-width, and number of pixels in the trigger, as well as the various measurements of the coherent and incoherent energy between the detectors in the network. Thus the classifier has very limited information of the waveform morphology.

There is still a possibility that the classifier will become too specialised to the waveforms in the training set, as certain waveform morphologies may have particular characteristics that become apparent in the parameters that the MVA is given. For this reason we also trained the classifier on a variety of different waveforms. The training set includes long and short waveforms, and a variety of different morphologies. Some of the signals are astrophysically motivated, such as compact binary coalescence signals, while some are artificial, such as the white noise burst.

The final tool we use to ensure the classifier is sensitive to generic waveforms is to test the classifier on waveforms that are not included in the training set. If the classifier can find waveforms not in the training set, then we can be reasonably confident that it is sensitive to generic waveforms. We also try removing certain waveforms from the training set to test the robustness of the classifier. This will lead to a drop in sensitivity for that waveform, but if the drop in sensitivity is small, then we can be confident that the classifier is robust. Experiments with removing certain waveforms from the training set showed that the sensitivity did not change by more than a few percent.

It was also found to be important that the injection runs for the MVA were performed on the off-source data. This is different to how X-pipeline is typically tuned, where the injections are performed in the on-source data. It was found that using injections in the on-source data can lead to false positive results. This is probably because some of the injections will be masked by noise in the detector data, when the MVA classifier analyses the on-source, it will then see this noise and mistakenly identify it as a GW.

#### 6.2.4 Optimisation and Validation

With any machine learning algorithm there are *hyperparameters* that must be tuned. These are the parameters of the classifier itself, such as the number of trees in the ensemble, or the maximum depth of the trees. We optimise the hyperparameters by repeatedly running an example analysis with different hyperparameters, trying to improve the sensitivity of the search. We measured the sensitivity of the search by looking at a *dummy on-source*, the off-source trial that had the 90th percentile loudest trigger. We then calculate the 50% and 90% injection scale upper limit, a measure of injection amplitude at which 50% or 90% of the injections can be recovered with a detection statistic higher than in the dummy on-source. The smaller the 50% injection scale upper limit is, the more sensitive the search is to GW signals.

Setting our hyperparameters using the testing set can cause *data leakage*. Data leakage is when data from outside the training set is used to build a classifier. As we

tune our classifier on the testing set, it is possible that we will implicitly tune our classifier to work well *only* on this training and testing set.<sup>14</sup> To avoid this, once we have tuned the hyperparameters on a single GRB, we test the classifier on several other GRBs. If the performance drops significantly on these other GRB analyses, then we have had data leakage and we need to retune our classifier. This process of testing on previously unused data is called *validation*. If there is evidence of data leakage, then we must retune our classifier and validate it again, this time using a different (previously unused) GRB so as to avoid data leakage from the validation GRB.

Optimisation is a somewhat cyclical process, as once we have changed one hyperparameter, we must then go back and test that other parameters do not now need changing. But optimisation is not pure guess work, and in the rest of this section we will see the strategy we used to optimise our classifier. We first discuss optimising the training data, as the other hyperparameters will not have much of an effect if the training data is of poor quality. As mentioned in section 6.2.3, there are several choices to make regarding what data is used for training, such as setting the amplitude threshold. We then discuss optimising the BDT classifier itself.

## Training Set

Many of the hyperparameters mentioned in section 6.2.3 had to be optimised. Consider the amplitude threshold applied to triggers before they are included in the training set. We applied the threshold to the loghbayesian statistic (see table 6.1). Experimentation showed that allowing triggers with a loghbayesian value below 20 led to the misclassification of background triggers as signal. Setting a value higher than this slowly reduced the sensitivity of the search.

We also optimised the waveforms that are included in the training set. In particular we were interested in how sensitive the MVA is to waveforms not in the training set. What we found was that the MVA was fairly robust to having waveform types removed from the training set. For example, the final classifier we used was trained on a set of waveforms that included white noise burst (WNB) signals of about 0.1 seconds and about 10 seconds as well as cusp waveforms. Removing these from the training set surprisingly had almost no effect on the ability of the MVA to detect WNB signals, cusp signals, or any other waveform, as can be seen in figure 6.7. Similar results were found when removing other waveforms. This suggests that by only giving the MVA the coherent statistics calculated by X-pipeline, we have managed to build a classifier that is agnostic to waveform morphology, as required for a burst search. Table 6.2.4 details all the waveforms used for training and testing.

---

<sup>14</sup>This is very similar to overtraining discussed in section 6.2.

Waveform	Description
ADI	Accretion disk instability waveforms. Essentially, these model a ‘clump’ forming in the accretion disk, that spins down the black hole via GW emission. We use five different ADI waveforms, labeled a-e. They vary in the masses of the black hole, accretion disk, and the clump, as well as the spin of the black hole. Their durations range from 10 seconds to 250 seconds, hence they are all considered long waveforms for our purposes. <a href="#">cite van putten</a>
BNS	Inspiral and merger of a binary neutron star system. Typically has a duration of about 10 seconds.
NSBH	Inspiral and merger of a neutron star-black hole system. Typical duration of a few seconds.
SGC	Circularly polarised sine-Gaussian. Ad hoc waveforms, not based on any astrophysical source. We use four different SGC waveforms, which differ in their peak frequency. They are less than one second long.
SGE	Elliptically polarised sine-Gaussian waveform. Like SGC waveforms, this is not astrophysically motivated.
WNB	White noise burst. This waveform is random noise. We use two different durations of WNB, short 0.1 second bursts, and long 10 second bursts.
Cusp	Short duration waveform, predicted to be emitted by cosmic strings.

Table 6.3: **Training Waveforms.** Here we list the waveforms used by XTMVA for both training and testing, together with a short description.

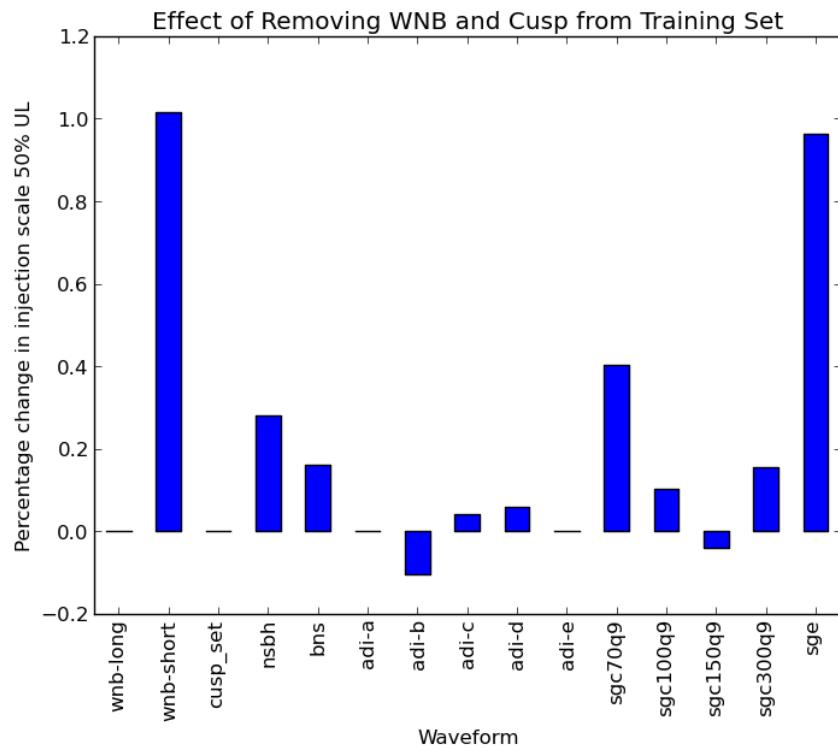


Figure 6.7: **Removing WNB and Cusp Waveforms from Training Set.** Here we plot the percentage change in 50% upper limit injection scale per waveform after removing WNB and Cusp waveforms from the training set. Negative values indicate that the sensitivity improved after the change. We see that the sensitivity to most waveforms drops, but by less than 1%. As we use a few hundred injections at each injection scale, this is not a statistically significant result. This shows that the MVA is able to detect GW morphologies that it has not been trained on.

## BDT Parameters

There are many hyperparameters that need to be set for a BDT analysis. In this section we discuss some of these hyperparameters and the method we used to optimise them for our analysis.

We began by setting values for  $NTrees$ , the number of trees in the ensemble, and the learning rate, discussed in section 6.2.2. These two parameters are set first to ensure the machine learning algorithm will converge in a reasonable amount of time, even if the results are not very sensitive. Setting the learning rate too low causes the classifier to take longer to converge. Setting the learning rate too high can cause the classifier to never converge. Similarly, using too many trees takes too long for the classifier to finish training, but too few and the training will terminate before the algorithm has converged. Setting these first ensures that we have a classifier that gives sensible results in a reasonable amount of time. While optimising, we set the learning rate slightly high and the number of trees slightly low, allow the analysis to complete quickly while we optimise our other parameters. Once the other parameters are set, we again optimise  $NTrees$  and the learning rate, increasing the number of trees and decreasing the learning rate to ensure the algorithm reaches its optimum performance, even if it increases the time taken for training.

We now tune the tree-specific parameters. Unlike the number of trees and the learning rate, which are primarily tuned to ensure the classifier will converge in a reasonable time, these parameters are set to ensure that the classifier is accurate but does not overtrain. Overtraining can happen when the trees are allowed to make cuts that are too fine, carving out regions of parameter space around anomalous events in the training set rather than finding cuts that generalise beyond the training data. The way to prevent this is to limit how fine the cuts made by the decision trees are allowed to be, while allowing cuts that are fine enough to pick out the general features of signal and background events in our data. There are several hyperparameters we can set to do this, which must all be tuned.

The first of these is the maximum depth of the trees, which is the maximum number of cuts a tree can make before it reaches a leaf node. Each cut divides the parameter space into ever smaller regions which it labels as background or signal. Setting the maximum number of cuts too low will therefore cause the classifier to be too coarse in dividing up the parameter-space, resulting in poor accuracy. Increasing the maximum depth allows the classifier to pick out smaller features in the parameter-space. If the maximum depth is too high then the classifier will overtrain; dividing the parameter-space into precisely the regions that work for the training set and losing generality. As we are using adaptive boosting, it is recommended **cite TMVA userguide** to use trees with fewer cuts. For this reason we tried values from 2-16 and recorded the effect this had on the sensitivity of the search. We found that for our problem a maximum depth of 8 was optimal.

A related hyperparameter is the minimum number of events that we allow in a leaf node. If we allow the training algorithm to have any number of events in a leaf node, then it will occasionally find cuts that results in a small number of events in one or more of the leaf nodes. Just as with the trees that were too deep, this can cause the classifier to carve out regions of parameter space around anomalous events in the training set rather than finding cuts that can generalise beyond the training set. Conversely, setting the minimum number of events allowed in the leaf nodes to be too high does not allow the classifier to pick out the key features in the data. We used a grid search over the values 100-1600 for the minimum number of events per leaf node and found the optimal value to be 400.<sup>15</sup>

The final hyperparameter we set was the number of cuts that the training algorithm scans over to find the best cut. When the classifier is training it searches for the best way to cut the parameter space into a signal subspace and a background subspace. To do this we can try every possible cut on every parameter. This can be very slow and lead to overtraining. To speed up our analysis and reduce the chance of overtraining, we can choose the number of cuts to try on each parameter. For example, we may decide to use 10 cuts for each parameter. In this case the algorithm will tune the cut on a parameter which ranges from 0-100 by trying cuts at 0,10,20,... and selecting the best of these cuts. To tune the number of cuts we wanted to use, we tested values in the range 10-160 as well as allowing the classifier to try every possible cut. We found that 80 cuts was optimal for our purposes.

## Result of Optimisation

Optimising the hyperparameters had relatively little effect on the overall sensitivity of the classifier. The results of the optimisation are shown in figure 6.8. There you can see that most of the optimisation made the classifier  $\sim 3\%$  more sensitive to certain waveforms than the TMVA boosted decision tree default settings, though at the expense of about a 1% drop in some others. With more time it might be possible to see greater improvements, but our findings are that as long as the hyperparameters are within some sensible range<sup>16</sup>, the performance of the MVA is fairly constant. The same cannot be said for work done on optimising the training set. Without appropriately set thresholds, the MVA will regularly make false detections, making the pipeline unusable for a GW search. In figure 6.9 we plot a comparison of the MVA against a standard X-pipeline analysis of the same GRB. You can see that the MVA is outperforming X-pipeline on every waveform. As this was the GRB used for optimisation, this results probably overestimate the improvement that the MVA will bring. In the next section we will analyse a selection of other GRBs from O2 to

---

<sup>15</sup>It should be noted that this value should scale with the size of the training data. So if we increase the number of events in the training data by a factor of 10 then we need to increase the minimum number of events allowed in each leaf node by a factor of 10 as well.

<sup>16</sup>For example, using only trees of depth two hurts sensitivity, but depth 3 and above leads to only modest improvements.

make a fairer comparison of the X-pipeline and XTMVA.

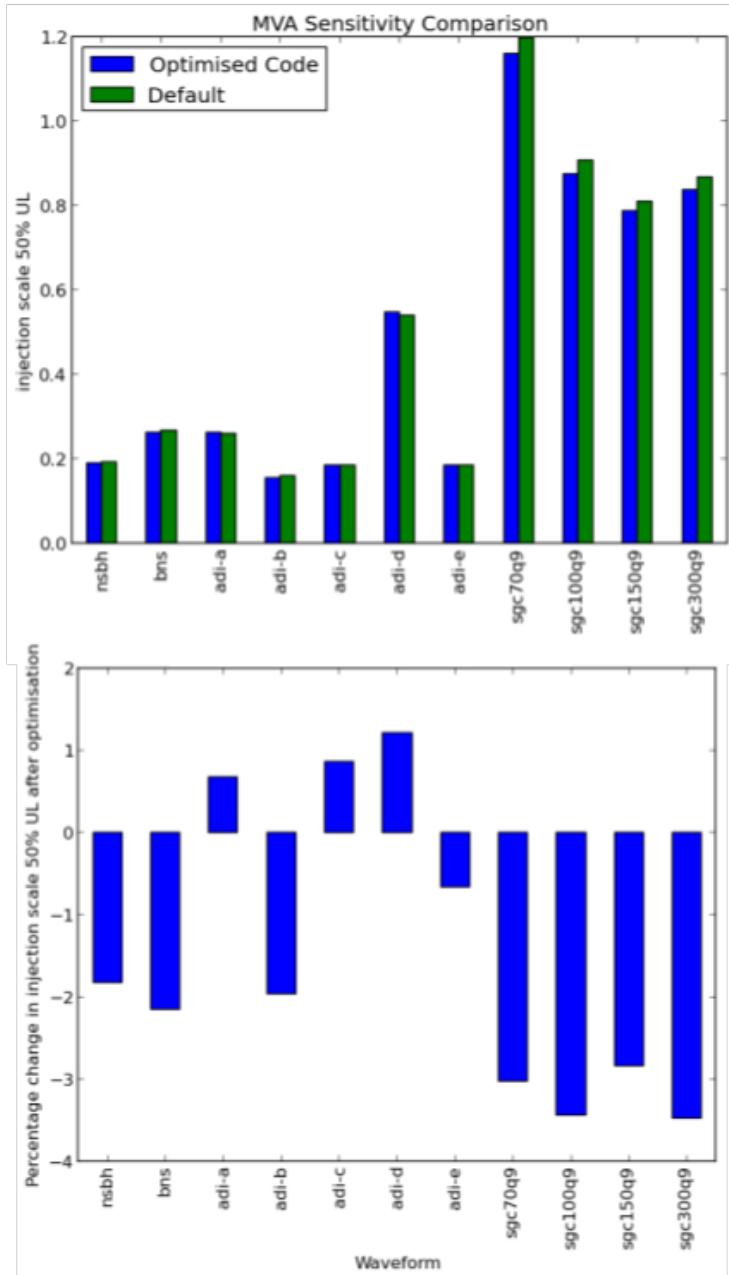
## 6.3 Effectiveness of XTMVA

In this section we will compare XTMVA to X-pipeline. To do this, we ran XTMVA on a selection of 15 GRBs from O2. Of these GRBs, one was used to optimise the hyperparameters of the classifier, and so should not be used for comparison to X-pipeline. The other was GW 170817, which is treated differently to the other GRBs analysed as it is a confirmed detection. The remaining 13 GRBs make up the results set. We will see that XTMVA has better sensitivity than X-pipeline, but also has some pathologies that make X-pipeline a more robust search at the current time. We end this section by discussing the implications of this research, and how we should proceed.

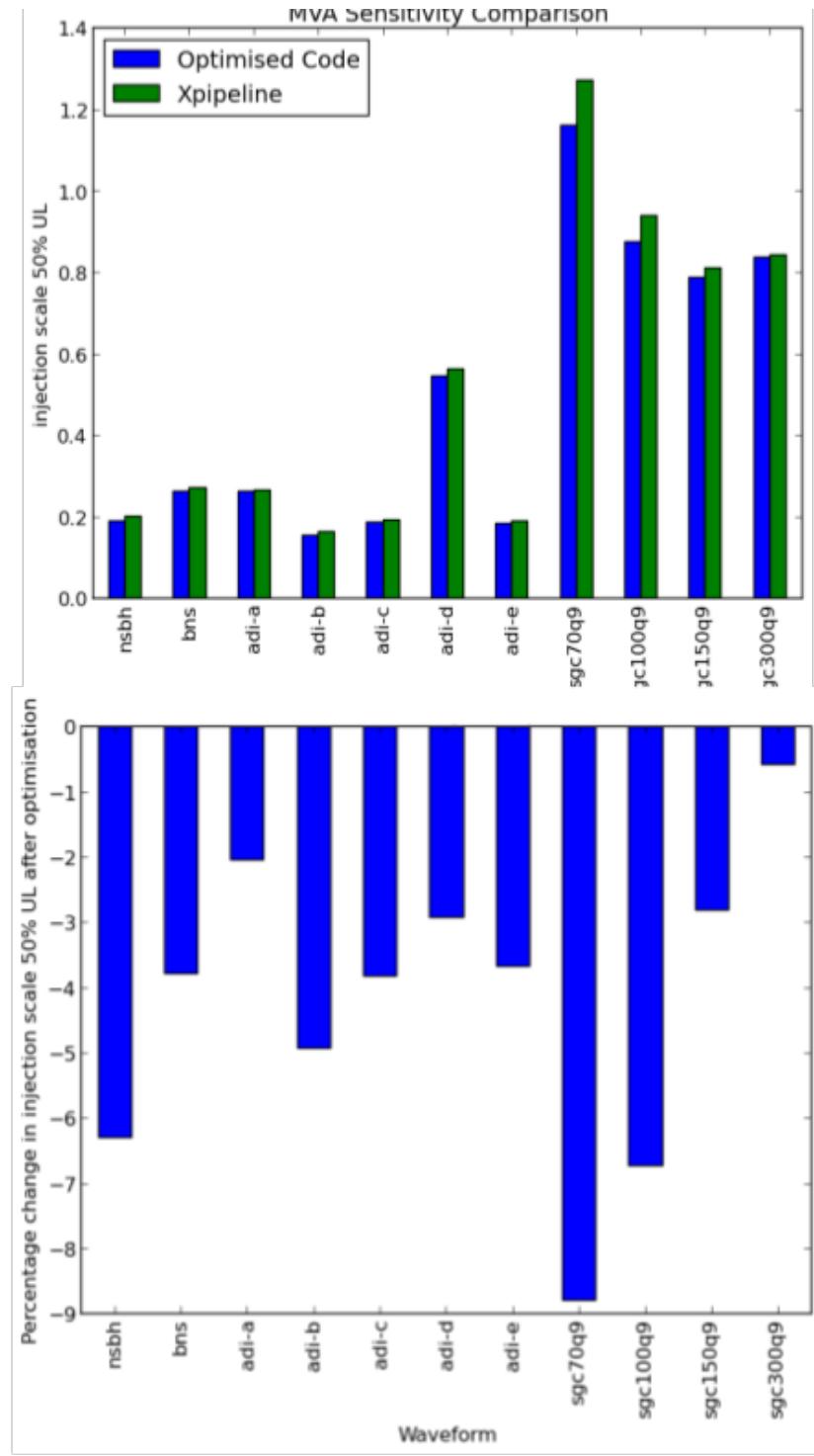
### 6.3.1 Analysis Setup

We compared the speed and sensitivity of XTMVA to X-pipeline on a subset of GRBs analysed in O2. When performing the GRB analyses with XTMVA, we tried as much as possible to keep the parameters of the analysis the same to make comparison easier. However, this was not always possible and it is important to clarify some of the differences between the X-pipeline and XTMVA analyses that could bias the results.

The first is that XTMVA used off-source injections, unlike X-pipeline which uses on-source injections. As mentioned in section 6.2.3, this is required for the MVA or it will make false detections. Using off-source injections forced another change upon the analysis. The O2 X-pipeline analysis used code that made the recovery of long injections easier. X-pipeline analyses data in chunks of 256 seconds. For long injections, especially those over  $\sim 100$  seconds long, this can lead to injections being spread over multiple segments. For this reason, X-pipeline will analyse two neighbouring segments of 256 seconds if a long injection is near the boundary between these chunks. This can increase the signal power of long injections and prevent them being broken up into multiple smaller signals. For purely technical reasons, the code that allows multiple chunks of data to be analysed is not compatible with off-source injections. Hence, the MVA is at a disadvantage when trying to find long injections. For short waveforms, it is very unlikely they would be injected near the boundary between two chunks of data and so the effect described above is negligible. Apart from these two changes (i.e. using off-source injections and not using the long-injection code) the MVA analysis was identical to the X-pipeline analysis used in O2.



**Figure 6.8: Effect of Hyperparameter Optimisation.** Here we see the effects of optimisation on the 50% upper limit injection scale. Lower values indicate a more sensitive search. The top panel shows the absolute values and the bottom panel shows the percentage change. The benefits of optimising the hyperparameters is no more than a  $\sim 3\%$  improvement in sensitivity when compared to the default settings of the TMVA boosted decision tree classifier. It is also interesting to note that the three waveforms that have their sensitivity drop after optimisation (adi-a, adi-c, and adi-d) are all long waveforms. We will discuss the problem with these waveforms in section 6.4.



**Figure 6.9: MVA Improvement.** Here we see the effects of using the MVA on the 50% upper limit injection scale for the same GRB that was used for optimisation. The top panel shows the absolute values and the bottom panel shows the percentage change. We can see that the MVA outperforms X-pipeline on every waveform. As this was the GRB used to optimise the hyperparameters, it cannot be guaranteed that these results will hold for other GRB analyses.

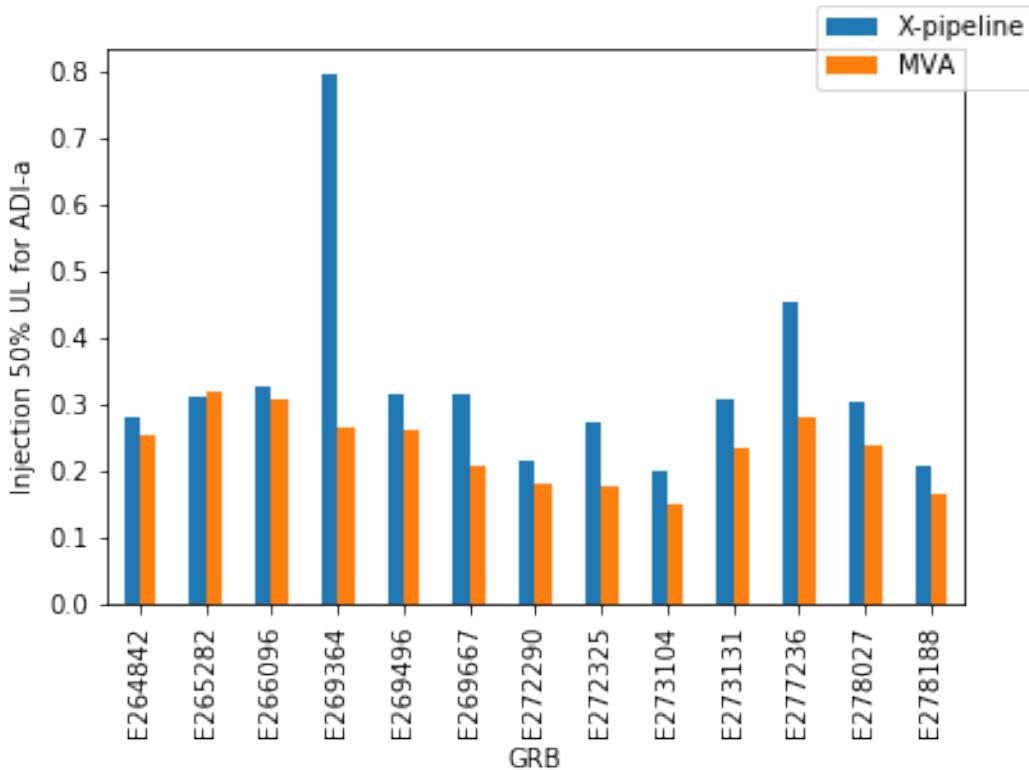


Figure 6.10: **X-pipeline and XTMVA ADI-a 50% Injection Scale Upper Limit by GRB.** Here we plot the sensitivity to the ADI-a waveform of both X-pipeline and XTMVA. The lower injection scales for XTMVA show that XTMVA is more sensitive than X-pipeline to the ADI-a waveform. Also, note the lower variation in injection scale between GRBs for XTMVA, suggesting that XTMVA is more stable than X-pipeline.

### 6.3.2 Sensitivity and Speed Comparison

To measure the sensitivity of XTMVA we used the same measure as in section 6.2.4, the 50% injection scale upper limit. In figures 6.10 and 6.11, we compare the sensitivity of XTMVA and X-pipeline on one of the accretion disk instability waveforms (ADI-a) and the 150 Hz circular sine Gaussian waveform (CSG) for the 13 GRBs in the results set. These plots show several interesting features. The MVA analysis is usually more sensitive to ADI-a waveforms than X-pipeline despite not benefiting from the long injection code.<sup>17</sup> The MVA also consistently outperforms X-pipeline for the CSG waveform. These plots also show that the MVA is more stable than X-pipeline, which has outliers for both waveforms, unlike the MVA.

In figure 6.12, we plot the median 50% injection scale upper limit for each waveform across the GRBs analysed. We can see that there is a small improvement when

<sup>17</sup> ADI-a waveforms are 39 seconds long and the analysis segments are about 256 seconds long. Therefore, we expect approximately 15% of them to intersect a segment boundary and therefore lose some signal power.

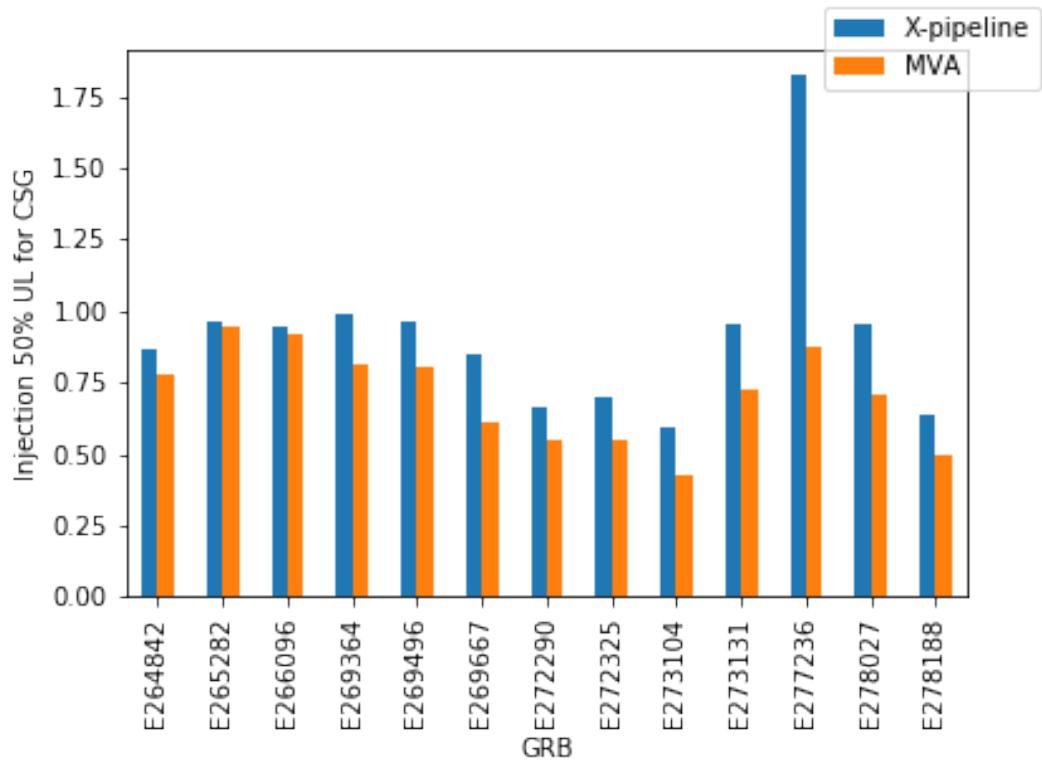


Figure 6.11: **X-pipeline and XTMVA CSG 50% Injection Scale Upper Limit by GRB.** Here we plot the sensitivity to the 150 Hz circular sine-Gaussian waveform of both X-pipeline and XTMVA. The lower injection scales for XTMVA show that XTMVA is more sensitive than X-pipeline to the CSG waveform. Again, note the lower variation in injection scale between GRBs for XTMVA, suggesting that XTMVA is more stable than X-pipeline.

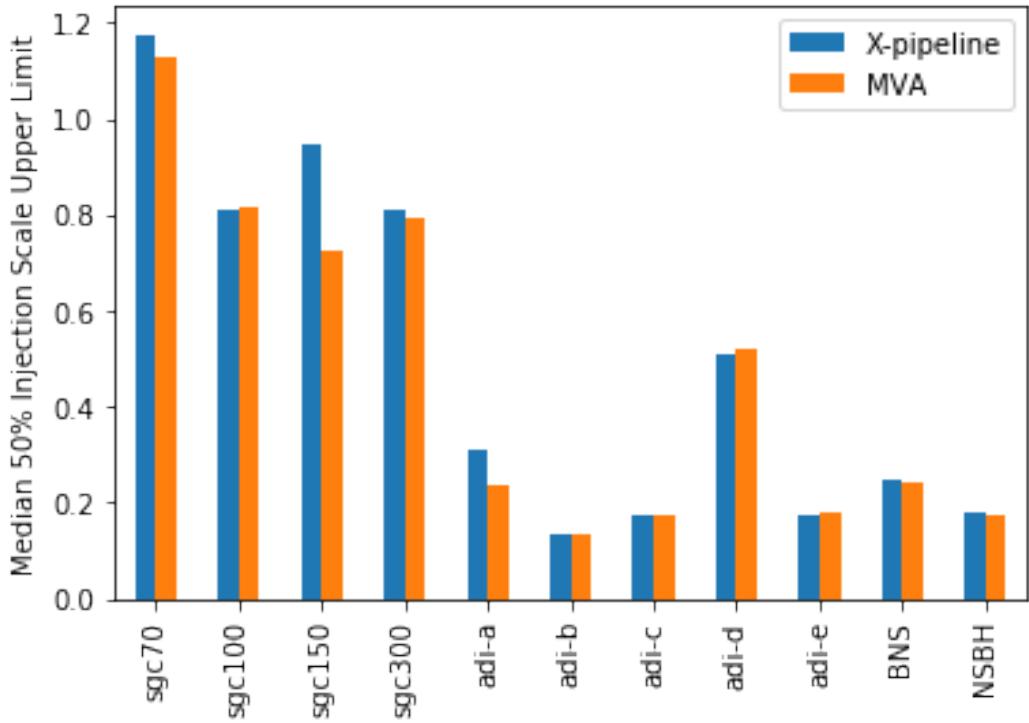


Figure 6.12: **Median 50% Injection Scale Upper Limit by Waveform.** Here we plot the median sensitivity to each waveform of both X-pipeline and XTMVA. Overall, XTMVA is more sensitive, especially to shorter waveforms such as sine-Gaussians. Apart from ADI-a, the MVA is worse than X-pipeline for long waveforms, though the difference is small. If the MVA could use the long injection code that X-pipeline uses, it is reasonable to expect that the MVA would outperform X-pipeline for long waveforms as well.

using the MVA. For most long waveforms, in particular the ADIs and the BNS waveforms, the difference is small (except for ADI-a). For most of the short waveforms, in particular the sine gaussian waveforms, the MVA displays a noticeable improvement (with the exception of the 100 Hz circular sine gaussian waveform). See appendix .3 for the full breakdown of sensitivity by waveform and GRB analysed.

As well as the sensitivity improvement, XTMVA is also uses less CPU time than X-pipeline. In particular, training an MVA classifier is much faster than calculating the optimal coherent cut with X-pipeline. The post-processing stage for a two detector analysis with X-pipeline takes approximately 75 CPU hours, while for XTMVA it takes about 18. Though it should be noted that the initial processing stage, where coherent statistics are calculated and triggers found, is still the most computationally intensive part of the process, taking about 650 hours. This part of the pipeline is identical between XTMVA and X-pipeline.

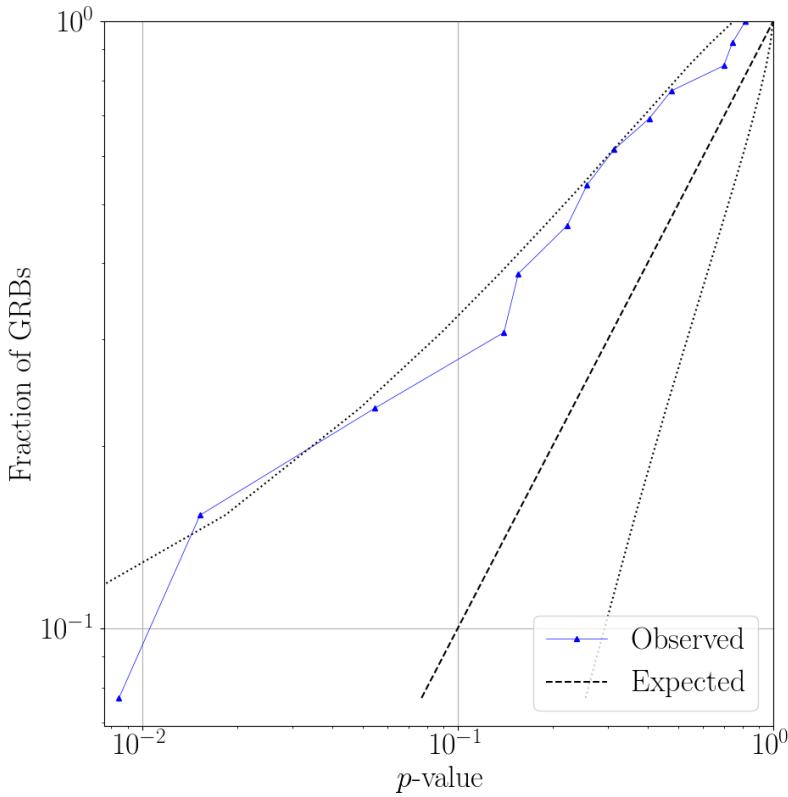


Figure 6.13: **XTMVA p-values.** Here we have plotted the p-values for 13 of the GRBs analysed with the MVA. The blue triangles indicate the p-value reported by the MVA, the black dotted lines show the expected distribution and a  $2\sigma$  deviation. Two GRBs that were analysed were left out from this plot: GRB 170817A as it had a known GW counterpart and E264930 as it was used to tune the hyperparameters. The analysis shows some bias towards low p-values. In particular, two out of the 13 analysed GRBs have a p-value of  $\sim 1\%$ . This can be compared to figure 6.3 which shows the X-pipeline p-values for O2. In particular, X-pipeline did not have the same bias towards low p-values that XTMVA does. This needs further investigation.

### 6.3.3 XTMVA Search Results

We analysed the 13 GRBs of the results set in the same manner as was done for the X-pipeline results in O2 (see section 6.1.6). In figure 6.13 shows the p-values for each GRB, which can be compared to 6.3. We see that the MVA has a bias towards low p-values. In particular, there are two p-values of about  $\sim 1\%$ . Further investigation shows that these two low p-value events were ranked by X-pipeline to be the most significant triggers in the on-source window before vetoes were applied, but failed the coherent cuts.

In figure 6.14 we replicated figure 6.4, the 90% exclusion distance plot. While we do not have as many GRBs in our sample as the X-pipeline O2 analysis did, we can see that the exclusion distance for ADI-a is higher than for the X-pipeline O2

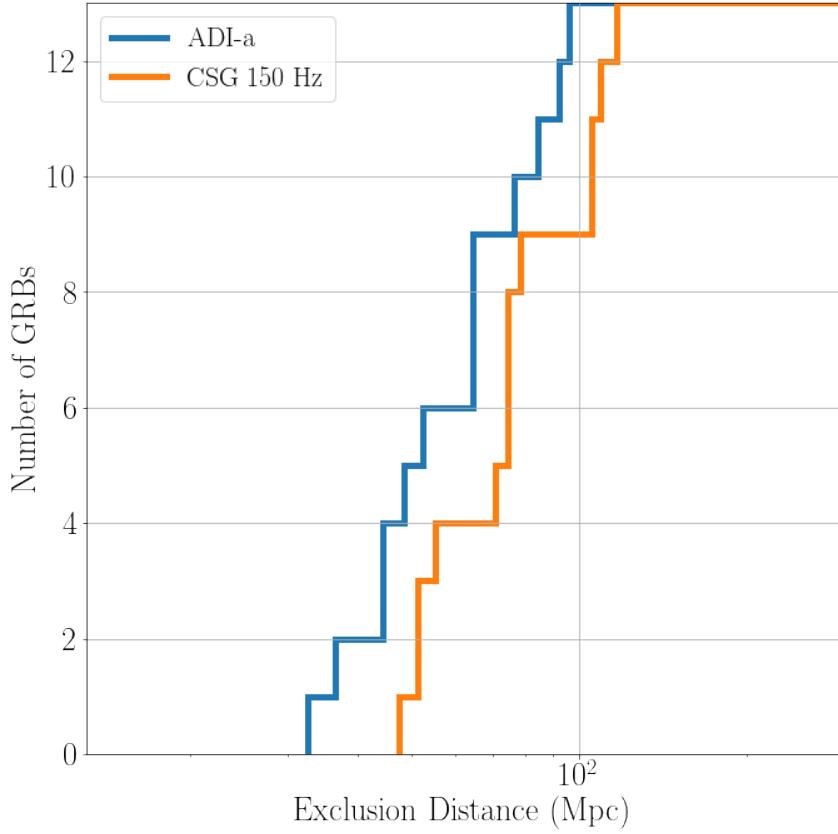


Figure 6.14: **Cumulative Distribution of Exclusion Distance.** Here we plotted the XTMVA 90% exclusion distance for the 13 GRBs in the results set. This is the distance to which 90% of injections can be recovered with a significance greater than the loudest event in the on-source.

analysis. The XTMVA analysis also does not have the long tail of low exclusion distance analyses that the X-pipeline analysis does. This is consistent with the results shown in figures 6.11 and 6.10, which show that for some GRBs the X-pipeline analysis has much lower sensitivity, while the MVA is far more consistent between analyses.

We also analysed GRB 170817A with XTMVA. This was the only known GW we analysed with XTMVA. The p-value of the most significant on-source trigger was  $4.095 \times 10^{-4}$ , a factor of twenty lower than the most significant trigger in the 13 other GRBs analysed and comparable to the p-value of  $3.1 \times 10^{-4}$  measured by X-pipeline. This p-value corresponds to just one off-source trial having a more significant event, and is therefore at the limit of how low a p-value we can achieve with the number of timeslides we were using. The on-source for this event also had two other low p-value events, which time-frequency data show are different parts of the inspiral signal. Had the energy from these different sections been considered one trigger, then the significance would have been even greater. We will discuss how this might be achieved in section 6.4.

## 6.4 Discussion and Future Work

In this chapter, we have shown that machine learning shows promise as a method for improving GW searches in the future. Building upon the work in [cite Toms paper](#), we showed that XTMVA is capable of detecting GWs, and that it is both computationally cheaper and more sensitive than X-pipeline. We tested XTMVA on a relatively large sample of real GRBs, and compared the results to a production run with X-pipeline and found the results to be promising. With that said, the analysis also has some problems that have not been previously reported, as well as some obvious areas for development. In this section we look at some of the problems with XTMVA, potential ways to fix these problems, and other areas of future development. We then end with a brief discussion of other work using machine learning for GW searches.

### 6.4.1 Low p-value Triggers

The most notable problem with XTMVA is the high number of low FAP triggers. While the results shown in section 6.3.3 would not suggest any detection other than GW 170817, it had two  $\sim 1\%$  FAP triggers in an analysis of 13 GRBs. Before XTMVA can be used in a search, we must be sure that it will not return a false positive. This means analysing more GRBs to see if the problem persists. If it does, then we must learn why XTMVA is ranking highly some glitches that fail the standard X-pipeline cuts.

### 6.4.2 Long Injections

We must also work to make XTMVA work with the long injection code. This is a problem for X-pipeline as well, as it will soon move to using off-source injections just as XTMVA does, and when it does it will also not be able to use the long injection code. This is a purely technical task and once fixed, XTMVA should see an appreciable improvement in sensitivity to longer waveforms, especially ADIs. At this point, XTMVA may run in to another problem we have noticed. The cleaning code<sup>18</sup> removes triggers that overlap in time or frequency with background triggers. This is much more likely to happen for longer signals, and we have noticed that for this reason, the signal training set ends up with relatively few long injections. We experimented with other cleaning methods to try to fix this. One was to only clean small injection scales, allowing loud injections to pass into the training set uncleaned. The logic was that a loud signal should still be detectable in the presence of some noise, and so cleaning was not as important. This approach did not work as it allowed too much noise into the training set which then led to more low FAP triggers.

---

<sup>18</sup>Discussed in section 6.2.3.

Another cleaning method was also tried. This required that the injection and noise time-frequency boxes overlap by at least some user-defined percentage to be removed from the training set. Overlaps of 50% and 90% were tested but both yielded the similar results as the original cleaning code. There are two likely reasons for this. The first is because glitches will typically be bigger in time-frequency space than most injections, especially in the frequency dimension. The other reason is because even when injections are long, they will usually be detected as many short triggers rather than one large trigger. This means that although the injection is long, the trigger that is reported will probably only be a fraction of the whole waveform, which then makes it more likely that the overlap between the glitch and the injection will be large enough for the trigger to be vetoed.

This brings us to another area of development for XTMVA, *Generalised Clustering*. This is a tool developed for X-pipeline<sup>19</sup> that changes the way triggers are defined. By default, triggers are groups of neighbouring pixels in the time-frequency maps, such as in figure 6.1. Specifically, for pixels to be clustered together, they must share an edge or a corner. Generalised clustering allows these pixels to be separated by a user-specified number of pixels. The effect of this is to prevent one large GW signal being broken up into many smaller triggers. This is what happened with the analysis of GW 170817 mentioned in section 6.3.3, where the MVA reported three high significance triggers in a small time-frequency window. The downside to using generalised clustering is that it can cause noise triggers to be grouped together as well, boosting the power of noise. This then causes our sensitivity to short triggers to be reduced slightly. Experimenting with generalised clustering in X-pipeline suggests that the benefits outweigh the costs, with an improvement of between 7% and 50% in detectable amplitude for long inspiral and ADI waveforms at the cost of a decrease of 3% to 9% for short sine-Gaussian waveforms.

Experiments were done to see if generalised clustering would improve the sensitivity of XTMVA. We found that the 50% injection scale upper limit for different waveforms saw roughly the same changes as X-pipeline, with improvements for long waveforms but not for short waveforms (see figure 6.15). However, the detection efficiency for very large injections was relatively low, meaning that the loudest signals were being missed. To see the problem, we have plotted a detection efficiency curves for the NSBH injection set of an XTMVA analysis without generalised clustering in figure 6.16 and with generalised clustering in figure 6.17. It is clearly visible from the plots that using generalised clustering causes high energy injections to be missed. The problem seems to only effect longer waveforms, such as ADIs and inspirals, while the shorter waveforms, such as sine-gaussians, do not see a fall in their detection efficiency at high amplitude.

The problem seems to be the cleaning algorithm again. Generalised clustering makes both signal and background triggers louder and longer, which increases the

---

<sup>19</sup>Though X-pipeline can use generalised clustering already, it has not yet been used in a search.

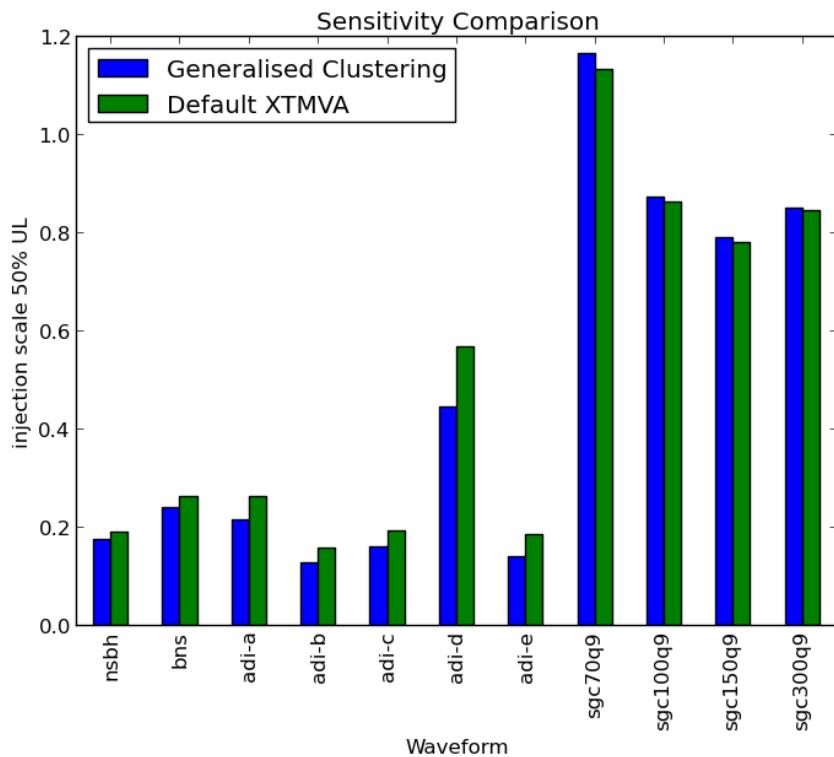


Figure 6.15: **Generalised Clustering Sensitivity Change.** Here we see the change in 50% injection scale upper limit for XTMVA with and without using generalised clustering. A lower value indicates a more sensitive search. The sensitivity of XTMVA is significantly improved for long waveforms such as ADIs, BNS, and NSBH when using generalised clustering. There is, however, a slight reduction in sensitivity to short waveforms.

chance that the injection will be coincident in time or frequency with a background trigger. This will then cause more injections to be removed by the cleaning code. To test this we looked at the time-frequency box size<sup>20</sup> of the triggers in the signal training set. The generalised clustering has a larger median trigger box size than the default analysis, as we would expect. However, the 90th percentile of box size is larger for the default analysis, suggesting that the longest triggers are not making it into the training set. In figure 6.18 we can see a histogram of the time-frequency box size for the triggers in the training set of the default and generalised clustering runs of XTMVA. We can see that the default analysis has more triggers with a large box size, supporting our hypothesis that the cleaning is removing long waveforms from the training set. The figure also shows that many shorter injections are also not making it into the training set.<sup>21</sup> This does not affect the detection efficiency of short waveforms by much as there are still many short waveforms making it into the training set. As the number of waveforms making it into the training set falls quickly with time-frequency box size in both the default and generalised clustering analyses, losing a relatively small number of long injections can have a significant detrimental effect on performance at large amplitudes, which is what we believe is happening.

For this reason, generalised clustering does not seem to be compatible with the MVA as it exists at the moment. Given the significant improvements it brings to X-pipeline, it is unlikely that XTMVA will be able to outperform X-pipeline until we can integrate generalised clustering into XTMVA. To fix this requires a rethink of how we clean the signal training set.

### 6.4.3 Machine Learning Developments

There are now many more machine learning packages available than when work on XTMVA first began. These new MVA packages have many tools that would make XTMVA more transparent and run faster. Some of these packages have large communities that could be used to speed up development. Work is also ongoing to convert X-pipeline to python [`cite scotty https://x-pipeline.github.io/`], at which point it will become much easier to use the many python based machine learning packages that exist. For these reasons, it seems worthwhile to stop using TMVA. Future work using MVA on X-pipeline triggers should move to a modern software package with a large community of developer and users. This will likely increase the potential sensitivity of the pipeline and speed up development. This will not involve drastic changes to the XTMVA infrastructure, as most of the pre-processing code (cleaning, thresholding, building training and testing sets etc.) will still be needed, only the machine learning engine and post-processing scripts will need to be edited.

---

<sup>20</sup>This is the bandwidth multiplied by the duration of the trigger.

<sup>21</sup>In fact, using generalised clustering leads to 20% fewer triggers in the training set.

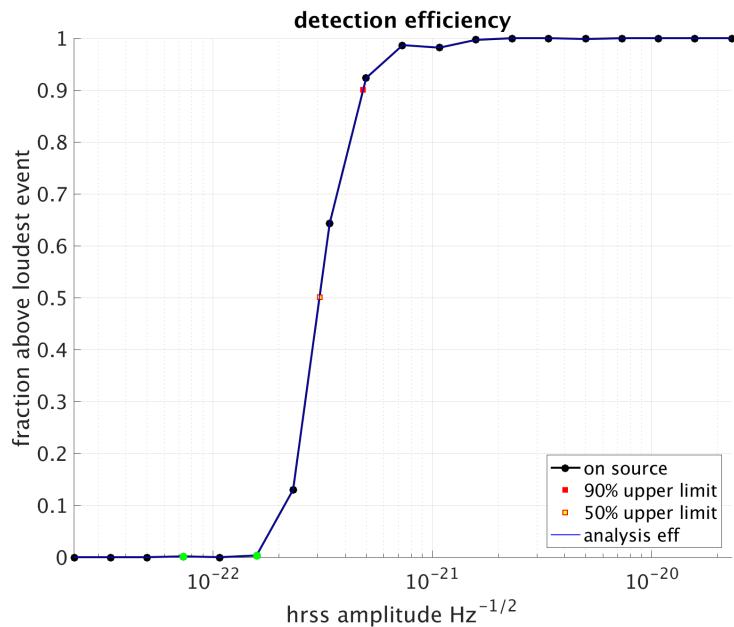


Figure 6.16: **Detection Efficiency Curve without Generalised Clustering.** This is the detection efficiency curve for an XTMVA analysis without generalised clustering. The x-axis shows the root-sum-square amplitude of the injected waveforms and the y-axis shows the fraction of injections detected. This plot shows that at low amplitude no injections are found, while for very loud injections there is almost a 100% detection efficiency. This is the typical, expected behaviour.

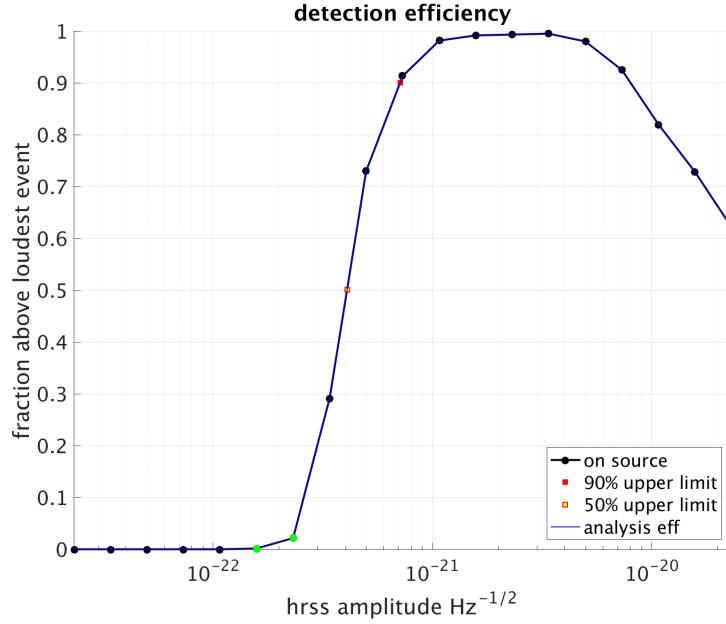


Figure 6.17: **Detection Efficiency Curve with Generalised Clustering.** This is the detection efficiency curve for an XTMVA analysis using generalised clustering. The x-axis shows the root-sum-square amplitude of the injected waveforms and the y-axis shows the fraction of injections detected. We can see that some very loud injections are being missed, despite the fact that close to 100% of some lower energy injection sets are being found.

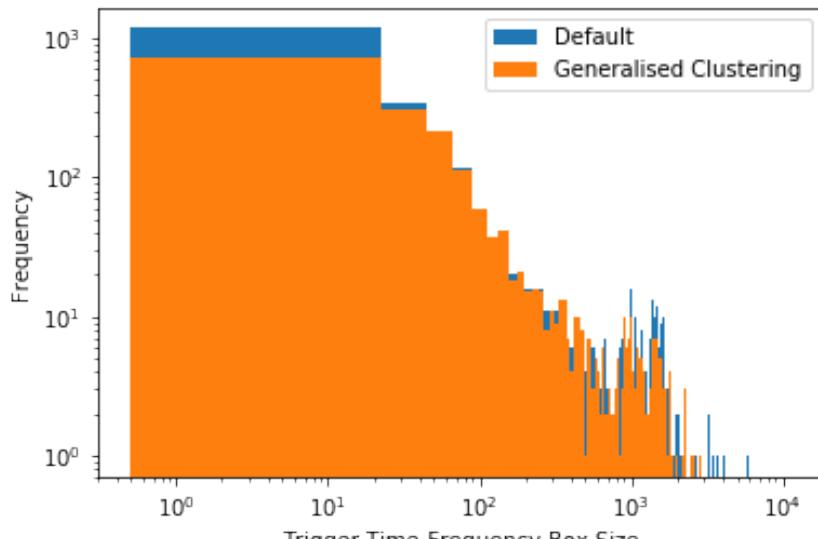


Figure 6.18: **Time-Frequency Box Size.** Here we have a histogram of the time-frequency box size of triggers in the signal training set for an analysis with generalised clustering and without. We can see that the default run not only has a lot more triggers overall, but it has more triggers with a large time-frequency box as well.

#### 6.4.4 Comparison to other Machine Learning Approaches

We end with a brief note on some of the other machine learning approaches for GW searches, and in particular on the differences with our search. Convolutional Neural Networks (CNNs) and other deep-learning tools have received a lot of attention in recent years, and have developed rapidly. These tools have been shown to be especially useful for image classification problems, and so it is perhaps natural that many different groups are trying to use CNNs to analyse time-frequency data to search for GWs **cite lots of people**. In essence, these CNNs are tasked with learning how to coherently combine data-streams, recognise an inspiral waveform, or invent some better technique GW detection. This is quite different from the method we have described in this chapter. We calculate statistics that we already know are useful for finding GWs, a process called *feature engineering*, and then task the classifier with finding more intelligent ways of combining these statistics. The advantages of the CNNs is that they are faster, as feature engineering can take time, and they are less restricted, which can allow them to find unexpected ways to solve the task they are given. The advantage of feature engineering is that it makes it easier to train a classifier, as it does not need to learn how to, for example, coherently combine data streams. Both approaches show a lot of promise for the future of GW astronomy.

# Chapter 7

## Conclusions

*“Equipped with his five senses, man explores the universe around him  
and calls the adventure **Science**. ”*

- Edwin Hubble

This thesis has covered many areas of joint GRB/GW astronomy. After summarising the current state of both GW and GRB astronomy, we discussed two current targeted GW searches for the followup of GRB observations. The first of these was the PyGRB targeted matched filter search for short GRB followup. We showed how PyGRB works and discussed the results of the most recent PyGRB search. We saw that PyGRB detected GWs associated with GRB 170817A, and used this with the non-detection from the 41 other GRBs analysed to make predictions on rates of joint detections in the future.

We discussed how PyGRB is being developed to make it faster, more sensitive, and more flexible for future development. The new code is able to analyse a GRB in half the CPU time of the old code, and can use new developments in the all-sky search to further improve speed and sensitivity. The shorter analysis time improves the scientific usefulness of a detection. It also allows poorly localised GRBs to be analysed, unlike the current search which cannot analyse poorly localised GRBs in a reasonable amount of time. Faster analysis also allows for more time slides, which allows for more accurate measurements of the false alarm rate. The new code can currently only analyse a single sky point, and cannot use time slides. Work is ongoing to add this functionality. Once this is done, the new pipeline will be ready to carry out a search.

We also saw the X-pipeline burst search for GWs associated with both long and short GRBs. After summarising the key features of the pipeline and looking at its performance in O2, we looked at how XTMVA attempted to improve the sensitivity of X-pipeline using machine learning. We saw that XTMVA had improved sensitivity and reduced the analysis time compared to X-pipeline. We discussed some of the

problems that come with using machine learning, as well as how these problems have been, or can be, overcome. Despite significant progress, it does not seem like XTMVA will be in a position to replace X-pipeline in the near future. This is partly for technical reasons, such as the need to rewrite and update the code to use modern software, and partly for reasons that are intrinsic to the approach we took to machine learning, such as the problem training a machine learning classifier on long waveforms without polluting the signal training set with noise. For these reasons, it is sensible to use X-pipeline with the other improvements, such as generalised clustering, that have been developed alongside the in parallel to XTMVA. That said, there is currently a lot of ongoing work to try and using machine learning for a GW search, and many of the problems we faced will likely affect these searches too. Hopefully they will learn from our experience.

With the detection of GWs associated with GRB 170817A, a new era of multi-messenger astronomy has begun. With the Kagra and LIGO India detectors coming online soon, the future of GW/GRB astronomy is promising. The same techniques can also be used to followup triggers from other messengers, such as neutrino signals, opening up yet more windows into the universe. Within hours of a trigger from anyone of these messengers, followup analysis will be carried out using the others. In the not-too-distant future, we will have a EM, GW, and neutrino observatories acting as different tools for the same searches.

Equipped with these three senses, we explore the universe.

# Appendices

---

## .1 General Relativity

Christoffel Symbol

$$\Gamma_{\mu\nu}^\lambda = \frac{1}{2}g^{\lambda\rho}[\partial_\nu g_{\mu\rho} + \partial_\mu g_{\nu\rho} - \partial_\rho g_{\mu\nu}] \quad (1)$$

Riemann Curvature Tensor

$$R_{\lambda\alpha\beta}^\mu = \partial_\alpha \Gamma_{\lambda\beta}^\mu - \partial_\beta \Gamma_{\lambda\alpha}^\mu + \Gamma_{\nu\alpha}^\mu \Gamma_{\lambda\beta}^\nu - \Gamma_{\nu\beta}^\mu \Gamma_{\lambda\alpha}^\nu \quad (2)$$

Ricci Tensor

$$R_{\mu\nu} = g^{\alpha\beta} R_{\alpha\mu\beta\nu} = R_{\mu\beta\nu}^\beta \quad (3)$$

Ricci Scalar

$$R = g^{\alpha\beta} R_{\alpha\beta} = R_\beta^\beta \quad (4)$$

The Einstein Equations

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \quad (5)$$

Alternative form of the Einstein equations

$$R_{\mu\nu} = \frac{8\pi G}{c^4} \left( T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu} \right) \quad (6)$$

## .2 PyGRB O2 Result Tables

GRB	UTC time	ra	dec	satellite	network	BNS	NSBH generic spin	NSBH aligned spin
161217128	03:03:46	14 <sup>h</sup> 26 <sup>m</sup> 31 <sup>s</sup>	51°58'	Fermi	H1L1	65	85	122
170112A	02:01:59	01 <sup>h</sup> 00 <sup>m</sup> 55 <sup>s</sup>	-17°13'	Swift	H1L1	83	106	144
170121133	03:10:52	16 <sup>h</sup> 07 <sup>m</sup> 58 <sup>s</sup>	13°48'	Fermi	H1L1	96	142	172
170125022	00:31:14	17 <sup>h</sup> 36 <sup>m</sup> 34 <sup>s</sup>	28°34'	Fermi	H1	46	52	57
170127067	01:35:48	22 <sup>h</sup> 37 <sup>m</sup> 19 <sup>s</sup>	-63°56'	Fermi	H1L1	76	129	141
170203486	11:40:26	16 <sup>h</sup> 20 <sup>m</sup> 22 <sup>s</sup>	-00°30'	Fermi	H1L1	66	99	119
170219002	00:03:07	03 <sup>h</sup> 39 <sup>m</sup> 22 <sup>s</sup>	50°04'	Fermi	H1L1	171	251	304
170302166	03:58:24	10 <sup>h</sup> 17 <sup>m</sup> 00 <sup>s</sup>	29°23'	Fermi	H1L1	107	175	206
170305256	06:09:07	02 <sup>h</sup> 34 <sup>m</sup> 38 <sup>s</sup>	12°05'	Fermi	L1	48	73	82
170325331	07:56:58	08 <sup>h</sup> 29 <sup>m</sup> 55 <sup>s</sup>	20°31'	Fermi	H1L1	73	88	125
170428A	09:13:42	22 <sup>h</sup> 00 <sup>m</sup> 12 <sup>s</sup>	26°54'	Swift	H1L1	105	167	178
170506169	04:02:48	07 <sup>h</sup> 29 <sup>m</sup> 02 <sup>s</sup>	51°52'	Fermi	H1L1	103	174	149
170614505	12:06:39	20 <sup>h</sup> 43 <sup>m</sup> 58 <sup>s</sup>	-37°54'	Fermi	H1	9	22	0
170709334	08:00:24	20 <sup>h</sup> 40 <sup>m</sup> 10 <sup>s</sup>	02°12'	Fermi	L1	139	228	255
170726249	05:58:15	11 <sup>h</sup> 05 <sup>m</sup> 41 <sup>s</sup>	-34°00'	Fermi	H1L1	124	152	207
170802638	15:18:25	03 <sup>h</sup> 29 <sup>m</sup> 12 <sup>s</sup>	-39°12'	Fermi	H1L1V1	45	62	72
170805B	14:18:49	09 <sup>h</sup> 42 <sup>m</sup> 31 <sup>s</sup>	69°54'	IPN	H1L1V1	132	163	218
170808065	01:34:09	00 <sup>h</sup> 13 <sup>m</sup> 12 <sup>s</sup>	62°18'	Fermi	L1V1	58	83	87
170817908	21:47:34	05 <sup>h</sup> 32 <sup>m</sup> 07 <sup>s</sup>	50°04'	Fermi	H1V1	35	51	63
161210524	12:33:54	18 <sup>h</sup> 52 <sup>m</sup> 29 <sup>s</sup>	63°03'	Fermi	H1L1	61	72	112
161212652	15:38:59	01 <sup>h</sup> 39 <sup>m</sup> 36 <sup>s</sup>	68°12'	Fermi	H1	49	59	60

GRB	UTC time	ra	dec	satellite	network	BNS	NSBH generic spin	NSBH aligned spin
170111815	19:34:01	18 <sup>h</sup> 03 <sup>m</sup> 31 <sup>s</sup>	63°42'	Fermi	H1	95	160	198
170121067	01:36:54	00 <sup>h</sup> 12 <sup>m</sup> 07 <sup>s</sup>	-75°37'	Fermi	H1L1	79	105	144
170124528	12:40:29	00 <sup>h</sup> 43 <sup>m</sup> 24 <sup>s</sup>	11°01'	Fermi	H1	65	101	116
170125102	02:27:10	23 <sup>h</sup> 57 <sup>m</sup> 38 <sup>s</sup>	-38°13'	Fermi	H1	30	39	63
170127B	15:13:29	01 <sup>h</sup> 19 <sup>m</sup> 58 <sup>s</sup>	-30°20'	Swift	H1	113	169	197
170206A	10:51:58	14 <sup>h</sup> 12 <sup>m</sup> 43 <sup>s</sup>	12°34'	IPN	H1L1	151	254	264
170222A	05:00:59	19 <sup>h</sup> 31 <sup>m</sup> 53 <sup>s</sup>	28°04'	IPN	H1L1	80	86	112
170304003	00:04:26	22 <sup>h</sup> 02 <sup>m</sup> 00 <sup>s</sup>	-73°45'	Fermi	H1L1	105	143	178
170318B	15:27:53	18 <sup>h</sup> 57 <sup>m</sup> 10 <sup>s</sup>	06°19'	Swift	H1L1	152	254	281
170403583	13:59:18	17 <sup>h</sup> 48 <sup>m</sup> 19 <sup>s</sup>	14°31'	Fermi	H1L1	166	240	261
170430204	04:54:20	01 <sup>h</sup> 35 <sup>m</sup> 26 <sup>s</sup>	30°07'	Fermi	H1	32	54	81
170604603	14:28:05	22 <sup>h</sup> 41 <sup>m</sup> 36 <sup>s</sup>	40°42'	Fermi	L1	131	204	237
170708046	01:06:11	22 <sup>h</sup> 13 <sup>m</sup> 00 <sup>s</sup>	25°37'	Fermi	L1	57	105	103
170723882	21:10:18	14 <sup>h</sup> 10 <sup>m</sup> 19 <sup>s</sup>	39°49'	Fermi	H1L1	95	83	179
170728A	06:53:29	03 <sup>h</sup> 55 <sup>m</sup> 36 <sup>s</sup>	12°09'	Swift	H1L1	89	129	163
170803172	04:07:16	05 <sup>h</sup> 06 <sup>m</sup> 00 <sup>s</sup>	24°00'	Fermi	H1L1V1	56	83	105
170803B	22:00:32	00 <sup>h</sup> 56 <sup>m</sup> 53 <sup>s</sup>	06°34'	IPN	L1	140	215	234
170805A	14:38:11	17 <sup>h</sup> 55 <sup>m</sup> 12 <sup>s</sup>	-23°30'	IPN	H1L1V1	69	100	114
170817A	12:41:06	13 <sup>h</sup> 09 <sup>m</sup> 36 <sup>s</sup>	-23°24'	Fermi	H1L1V1	0	0	41
170818137	03:17:20	19 <sup>h</sup> 48 <sup>m</sup> 53 <sup>s</sup>	06°21'	Fermi	H1L1	103	146	169
170816599	14:23:04	23 <sup>h</sup> 25 <sup>m</sup> 36 <sup>s</sup>	19°06'	Fermi	H1V1	46	56	73

### 3 MVA Results Tables

Waveform	X-pipeline	MVA
sgc70	1.18	1.13
sgc100	0.81	0.81
sgc150	0.95	0.72
sgc300	0.81	0.79
adi-a	0.31	0.24
adi-b	0.14	0.13
adi-c	0.18	0.17
adi-d	0.51	0.52
adi-e	0.17	0.18
BNS	0.25	0.24
NSBH	0.18	0.18

Table 1: **Median 50% Injection Scale Upper Limits.** Here we list the median 50% injection scale upper limit values for X-pipeline and the MVA.

GRB	sgc70	sgc100	sgc150	sgc300	adi-a	adi-b	adi-c	adi-d	adi-e	BNS	NSBH
E264842	1.14	0.88	0.78	0.87	0.26	0.16	0.18	0.52	0.18	0.26	0.18
E265282	1.36	1.04	0.95	1.06	0.32	0.18	0.23	0.64	0.24	0.32	0.23
E266096	1.40	1.08	0.92	1.06	0.31	0.18	0.24	0.63	0.23	0.32	0.24
E269364	1.17	0.88	0.82	0.85	0.27	0.17	0.19	0.56	0.19	0.27	0.20
E269496	1.28	0.87	0.80	0.84	0.26	0.16	0.19	0.56	0.19	0.27	0.19
E269667	1.10	0.76	0.61	0.67	0.21	0.13	0.16	0.43	0.15	0.22	0.16
E272290	0.83	0.59	0.55	0.58	0.18	0.11	0.13	0.37	0.13	0.18	0.13
E272325	0.84	0.63	0.55	0.59	0.18	0.11	0.13	0.38	0.13	0.18	0.14
E273104	0.65	0.51	0.42	0.51	0.15	0.08	0.11	0.30	0.11	0.15	0.11
E273131	1.12	0.81	0.72	0.79	0.23	0.13	0.17	0.48	0.18	0.24	0.17
E277236	1.33	1.09	0.87	0.98	0.28	0.18	0.22	0.63	0.23	0.29	0.22
E278027	1.13	0.81	0.71	0.79	0.24	0.13	0.17	0.52	0.18	0.24	0.18
E278188	0.72	0.54	0.49	0.54	0.17	0.09	0.12	0.33	0.12	0.17	0.12

Table 2: **MVA 50% Injection Scale Upper Limit.** Here we list the 50% injection scale upper limit for each waveform type tested and for each GRB analysed.

GRB	sgc70	sgc100	sgc150	sgc300	adi-a	adi-b	adi-c	adi-d	adi-e	BNS	NSBH
E264842	1.36	0.97	0.86	0.94	0.28	0.17	0.21	0.64	0.20	0.28	0.20
E265282	1.42	1.11	0.97	1.07	0.31	0.19	0.23	0.64	0.23	0.33	0.24
E266096	1.61	1.11	0.95	1.06	0.33	0.19	0.25	0.66	0.24	0.32	0.25
E269364	1.21	0.93	0.99	0.84	0.80	0.17	0.20	0.57	0.19	0.27	0.19
E269496	1.37	0.89	0.96	0.83	0.31	0.16	0.19	0.56	0.20	0.26	0.19
E269667	1.11	0.73	0.85	0.65	0.31	0.12	0.16	0.45	0.15	0.20	0.15
E272290	0.86	0.58	0.66	0.58	0.22	0.11	0.13	0.38	0.13	0.18	0.13
E272325	0.85	0.64	0.70	0.61	0.28	0.11	0.13	0.39	0.13	0.18	0.14
E273104	0.67	0.49	0.60	0.48	0.20	0.08	0.11	0.30	0.11	0.15	0.10
E273131	1.09	0.81	0.95	0.81	0.31	0.13	0.17	0.49	0.17	0.25	0.18
E277236	1.81	1.41	1.82	1.34	0.45	0.21	0.26	0.80	0.26	0.38	0.29
E278027	1.18	0.80	0.95	0.79	0.30	0.14	0.18	0.51	0.17	0.24	0.18
E278188	0.73	0.55	0.64	0.54	0.21	0.09	0.12	0.33	0.12	0.17	0.12

Table 3: **X-pipeline 50% Injection Scale Upper Limit.** Here we list the 50% injection scale upper limit for each waveform type tested and for each GRB analysed.

# Bibliography

- [1] J. Abadie et al. Search for gravitational waves associated with gamma-ray bursts during LIGO science run 6 and Virgo science runs 2 and 3. *Astrophys. J.*, 760:12, 2012.
- [2] B. P. Abbott, The LIGO Scientific Collaboration, and Virgo Collaboration. Prospects for observing and localizing gravitational-wave transients with advanced ligo and advanced virgo. *Living Reviews in Relativity*, 19(1):1, Feb 2016.
- [3] B. P. et al Abbott. Gw150914: The advanced ligo detectors in the era of first discoveries. *Phys. Rev. Lett.*, 116:131103, Mar 2016.
- [4] B. P. et al. Abbott. Gw170817: Observation of gravitational waves from a binary neutron star inspiral. *Phys. Rev. Lett.*, 119:161101, Oct 2017.
- [5] B. P. et all Abbott. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, Feb 2016.
- [6] Bruce Allen.  $\chi^2$  time-frequency discriminator for gravitational wave detection. *Phys. Rev. D*, 71:062001, Mar 2005.
- [7] Bruce Allen.  $\chi^2$  time-frequency discriminator for gravitational wave detection. *Phys. Rev.*, D71:062001, 2005.
- [8] Bruce Allen, Warren G. Anderson, Patrick R. Brady, Duncan A. Brown, and Jolien D. E. Creighton. FINDCHIRP: An Algorithm for detection of gravitational waves from inspiraling compact binaries. *Phys. Rev.*, D85:122006, 2012.
- [9] Scott D. et al. Barthelmy. The burst alert telescope (bat) on the swift midex mission. *Space Science Reviews*, 120(3):143–164, Oct 2005.
- [10] C. Bond, D. Brown, A. Freise, and K. Strain. Interferometer techniques for gravitational-wave detection, 2009. <https://arxiv.org/abs/0909.3661v3>.
- [11] Francesco Carotenuto and Badri Krishnan. Search for gravitational-wave signals associated with gamma-ray bursts during the second observing run of Advanced LIGO and Advanced Virgo. 2019.

- [12] Sean Carroll. *Spacetime and geometry : an introduction to general relativity*. Pearson, Harlow, Essex, 2014.
- [13] LIGO Scientific Collaboration, Fermi Gamma-ray Burst Monitor Virgo Collaboration, and INTEGRAL. Gravitational waves and gamma-rays from a binary neutron star merger: GW170817 and GRB 170817a. *The Astrophysical Journal*, 848(2):L13, oct 2017.
- [14] Neil J Cornish and Edward K Porter. The search for massive black hole binaries with LISA. *Classical and Quantum Gravity*, 24(23):5729–5755, nov 2007.
- [15] Tito Dal Canton et al. Implementing a search for aligned-spin neutron star-black hole systems with advanced ground based gravitational wave detectors. *Phys. Rev.*, D90(8):082004, 2014.
- [16] B. P. Abbott et al. Search for gravitational waves associated with gamma-ray bursts during the first advanced LIGO observing run and implications for the origin of GRB 150906b. *The Astrophysical Journal*, 841(2):89, may 2017.
- [17] Charles Meegan et al. The Fermi Gamma-Ray Burst Monitor. *The Astrophysical Journal*, 702(1):791–804, aug 2009.
- [18] Chad Hanna. Searching for gravitational waves from binary systems in non-stationary data. *PhD thesis, Louisiana State University*, 2008.
- [19] I. W. Harry and S. Fairhurst. Targeted coherent search for gravitational waves from compact binary coalescences. *Phys. Rev. D*, 83:084002, Apr 2011.
- [20] Jason W. T. Hessels, Scott M. Ransom, Ingrid H. Stairs, Paulo C. C. Freire, Victoria M. Kaspi, and Fernando Camilo. A radio pulsar spinning at 716 hz. *Science*, 311(5769):1901–1904, 2006.
- [21] E. J. Howell, K. Ackley, A. Rowlinson, and D. Coward. Joint gravitational wave – gamma-ray burst detection rates in the aftermath of GW170817. 2018.
- [22] Vassiliki Kalogera and Gordon Baym. The maximum mass of a neutron star. *The Astrophysical Journal*, 470(1):L61–L64, oct 1996.
- [23] Sebastian Khan, Sascha Husa, Mark Hannam, Frank Ohme, Michael Prrer, Xisco Jimnez Forteza, and Alejandro Boh. Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. *Phys. Rev.*, D93(4):044007, 2016.
- [24] Blent Kiziltan, Athanasios Kottas, Maria De Yoreo, and Stephen E. Thorsett. THE NEUTRON STAR MASS DISTRIBUTION. *The Astrophysical Journal*, 778(1):66, nov 2013.

- [25] Davide Lazzati, Rosalba Perna, Brian J. Morsony, Diego Lpez-Cmara, Matteo Cantiello, Riccardo Ciolfi, Bruno Giacomazzo, and Jared C. Workman. Late time afterglow observations reveal a collimated relativistic jet in the ejecta of the binary neutron star merger GW170817. *Phys. Rev. Lett.*, 120(24):241103, 2018.
- [26] D. V. et al Martynov. Sensitivity of the advanced ligo detectors at the beginning of gravitational wave astronomy. *Phys. Rev. D*, 93:112004, Jun 2016.
- [27] Dave McKechnie. On the use of higher order waveforms in the search for gravitational waves emitted by compact binary coalescences. *PhD thesis, Cardiff University*, 2010.
- [28] Cody et al. Messick. Analysis framework for the prompt discovery of compact binary mergers in gravitational-wave data. *Phys. Rev. D*, 95:042001, Feb 2017.
- [29] NASA. The burst and transient source experiment. <https://heasarc.gsfc.nasa.gov/docs/cgro/batse/>. Accessed: 2019-07-18.
- [30] Alexander H. Nitz, Thomas Dent, Tito Dal Canton, Stephen Fairhurst, and Duncan A. Brown. Detecting binary compact-object mergers with gravitational waves: Understanding and Improving the sensitivity of the PyCBC search. *Astrophys. J.*, 849(2):118, 2017.
- [31] Asaf Pe'er. Physics of Gamma-Ray Bursts Prompt Emission. *Adv. Astron.*, 2015:907321, 2015.
- [32] Surabhi Sachdev et al. The GstLAL Search Analysis Methods for Compact Binary Mergers in Advanced LIGO's Second and Advanced Virgo's First Observing Runs. 2019.
- [33] B.S. Sathyaprakash and B.F. Schutz. Physics, astrophysics and cosmology with gravitational waves. *Living Reviews in Relativity*, 12(2), 2009.
- [34] Samantha A. Usman et al. The PyCBC search for gravitational waves from compact binary coalescence. *Class. Quant. Grav.*, 33(21):215004, 2016.
- [35] John Veitch and Alberto Vecchio. A Bayesian approach to the follow-up of candidate gravitational wave signals. *Phys. Rev.*, D78:022001, 2008.
- [36] David Wanderman and Tsvi Piran. The rate, luminosity function and time delay of non-Collapsar short GRBs. *Monthly Notices of the Royal Astronomical Society*, 448(4):3026–3037, 03 2015.
- [37] M Was, P Sutton, Gareth Jones, and Isabel Leonor. Performance of an externally triggered gravitational-wave burst search. *Phys. Rev. D*, 86, 01 2012.

- [38] Michal Was, Marie-Anne Bizouard, Violette Brisson, Fabien Cavalier, Michel Davier, Patrice Hello, Nicolas Leroy, Florent Robinet, and Miltiadis Vavoulidis. On the background estimation by time slides in a network of gravitational wave detectors. *Classical and Quantum Gravity*, 27(1):015005, dec 2009.
- [39] Michal Was, Marie-Anne Bizouard, Violette Brisson, Fabien Cavalier, Michel Davier, Patrice Hello, Nicolas Leroy, Florent Robinet, and Miltiadis Vavoulidis. Limitations of the time slide method of background estimation. *Classical and Quantum Gravity*, 27(19):194014, sep 2010.
- [40] A. R. Williamson, C. Biwer, S. Fairhurst, I. W. Harry, E. Macdonald, D. Macleod, and V. Predoi. Improved methods for detecting gravitational waves associated with short gamma-ray bursts. *Phys. Rev.*, D90(12):122004, 2014.
- [41] A Zee. *Einstein gravity in a nutshell*. Princeton University Press, Princeton, 2013.
- [42] Feryal zel and Paulo Freire. Masses, radii, and the equation of state of neutron stars. *Annual Review of Astronomy and Astrophysics*, 54(1):401–440, 2016.
- [43] Feryal zel, Dimitrios Psaltis, Ramesh Narayan, and Jeffrey E. McClintock. THE BLACK HOLE MASS DISTRIBUTION IN THE GALAXY. *The Astrophysical Journal*, 725(2):1918–1927, dec 2010.