Searching for Gravitational Waves from Compact Binary Coalescences and Stochastic Backgrounds in the LIGO–Virgo Detector Network

Thesis by
Liting Xiao (肖鹂葶)

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2023
Defended October 13, 2022
ACKNOWLEDGEMENTS

"Cogito, ergo sum."
– René Descartes

First and foremost, I would like to thank my advisor, Prof. Alan J. Weinstein, for his support and guidance throughout the years. I could not have finished this tough journey without his mentorship. It was the hardest six years of my life, also the most formative and life-changing years. Alan, thank you for not giving up on me. Thank you for your patience and constant nudging over the six years. Thank you for truly caring. Thank you for encouraging me to explore topics that interest me.

From my earliest LIGO adventures as a SURF student in summer 2014, I would like to thank Tjonnie Li, Surabhi Sachdev, Kent Blackburn, and Alan for a lot of hand-holding in helping me understand the LIGO science that I worked on and in mentoring me to discover LIGO data analysis. From my later PhD years, I would like to thank Derek Davis for being the coolest PyCBC buddy in the US. Jonah Kanner, for being so passionate about open science and for my first experience as an instructor in GW Open Data Workshop. Mahlet Shiferaw, Phoebe McClincy, Sierra Garza, for being such awesome SURF students who made the summer more fun and helped me understand the science better by co-mentoring. I would also like to thank Arianna Renzini for being the most amazing collaborator on the most interesting stochastic background project that I have always been curious about.

I am also inspired by the current and past members of the Caltech LIGO Data Analysis Group: Tom Callister, Maximiliano Isi, Thomas J Massinger, Jess McIver, Alex Urban, Colm Talbot, Michael Coughlin, Shreya Anand, Ka-Lok (Rico) Lo, Ka-Yue (Alvin) Li. I am deeply grateful to my PyCBC collaborators on the other side of the Atlantic: Tito Dal Canton, Gareth Davies, Ian Harry, Thomas Dent, Bhooshan Gadre. Your passion, hard work, and dedication set the best examples that I can always look up to.

I am in gratitude for my time as a LSC Fellow at the LIGO Livingston Observatory in the summer of 2018. Special thanks to Arnaud Pele, Guillermo Valdes, Carl Blair, Stuart Aston, Brian O’Riley, Joe Betzweiser, Keith Thorne. I am fortunate to have worked on the commissioning of the LIGO Livingston detector towards LIGO–Virgo’s third observing run in that summer for a holistic view of LIGO.
I am in debt to my parents for their love and support from the other side of the planet. Sorry for my stubbornness and inconsideration and sorry for the quarrels that we had. I hope one day you will understand me. And I am sorry I could not be back home often enough throughout the years of my college and graduate studies for about 11 years, and possibly more to come after graduation.

I sincerely thank all my friends for the great conversations and the unforgettable time spent together over the years. I thank my crew of oldest childhood friends for always being my biggest supporters. I have also made the most wonderful “new” friends in LA and OC. Thank you all for being there when I needed ears, advice, and company. You have been a big part of my graduate life and have made it more colorful and meaningful. And I look forward to the years to come with you all.

Last but not least, I want to thank my wife-to-be, Xinghui Yan (闫星会). With you, it’s always astonishment and joy. You are my greater than 5σ detection, with a false alarm rate less than 1 event per Hubble time. I have long been a lone star roaming in the Universe, until you swept by and captured me. We fall into a cosmic dance as we spiral towards each other, and everything else just becomes the background, insignificant. Our lives may be ordinary, but we can shake space and time and outshine all the stars in the Universe with our moments together. I’ve never been more certain of anything in my life.

I am a member of the LIGO Laboratory. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY-0757058. I carried out most of the work presented here as a member of the LIGO Scientific Collaboration and benefited from interaction with many of its members. However, except where otherwise indicated, the descriptions and results presented in this thesis are my own and not necessarily those of the Collaboration.

I would like to conclude my acknowledgements with an ancient Chinese proverb. The quest for truth is long and ceaseless. And I will always be seeking the truth.

路漫漫其修远兮，吾将上下而求索。

The end of this chapter of my life is the beginning of the next.
ABSTRACT

Gravitational waves (GWs) are ripples in spacetime generated by accelerating masses, carrying away information about the underlying processes. There are four main astrophysical sources detectable in the sensitive band of the LIGO–VIRGO–KAGRA (LVK) GW detector network: compact binary coalescences, burst sources, continuous waves and stochastic gravitational-wave backgrounds. This thesis focuses on the detection methods of two of these categories, coalescing compact binaries and stochastic backgrounds, and their search results across LIGO–Virgo’s first three observing runs spanning from 2015 to 2020.

Compact binary coalescences of black holes and/or neutron stars are the only type of GW sources detected so far in the LVK frequency band. Such binary systems lose orbital energy via GW emission and are compact enough to merge within the age of the Universe. PyCBC is a matched-filter, all-sky pipeline for GW signals from compact binary mergers using a bank of modeled gravitational waveform templates. We describe the methods employed in PyCBC and present the developmental updates both in its archival and low-latency configurations for LIGO–Virgo’s third observing run. Using PyCBC to analyze the data from LIGO–Virgo’s first three observing runs, we summarize our results of the searches in gravitational-wave transient catalogs and characterize some exceptional events.

A stochastic gravitational-wave background consists of a large number of weak, independent and uncorrelated events of astrophysical or cosmological origin. The GW power on the sky is assumed to contain anisotropies on top of an isotropic component, i.e., the angular monopole. Complementary to the LVK searches, we develop an efficient analysis pipeline to compute the maximum-likelihood anisotropic sky maps in stochastic backgrounds directly in the sky pixel domain using data folded over one sidereal day. We invert the full pixel-pixel correlation matrix in map-making of the GW sky, up to an optimal eigenmode cutoff decided systematically using simulations. In addition to modeled mapping, we implement a model-independent method to probe spectral shapes of stochastic backgrounds. Using data from LIGO–Virgo’s first three observing runs, we obtain upper limits on anisotropies as well as the isotropic monopole as a limiting case, consistent with the LVK results. We also set constraints on the spectral shape of the stochastic background using this novel model-independent method.


detector in preparation for O3, resulting in sensitivity improvements. DOI: 10.1103/physrevd.102.062003. URL: https://doi.org/10.1103/physrevd.102.062003.


## CONTENTS

Acknowledgements ................................................................. iii
Abstract .................................................................................. v
Published Content and Contributions ....................................... vi
Contents .................................................................................. vii
List of Figures ........................................................................... x
List of Tables ............................................................................ xviii

Chapter I: Introduction: The Dawn of Gravitational-wave Science ........................................................................ 1
  1.1 A Brief History of the Hunt for Gravitational Waves .......... 1
  1.2 Outline of Chapters ......................................................... 3

Chapter II: Introduction to Gravitational Waves and LIGO ..................................................................................... 5
  2.1 Gravitational Waves in General Relativity ....................... 5
  2.2 Sources of Gravitational Waves ....................................... 13
  2.3 Advanced LIGO Detectors .............................................. 16
  2.4 Overview of LIGO Science .............................................. 20

Chapter III: Compact Binary Coalescences ................................................. 22
  3.1 Astrophysical Processes ................................................ 23
  3.2 Gravitational Waveforms .............................................. 26

Chapter IV: The PyCBC Search Pipeline for Gravitational Waves from Compact Binary Coalescences .............................. 31
  4.1 Matched-filter Search Pipeline ....................................... 32
  4.2 PyCBC Live: A Real-time Search .................................... 45

Chapter V: Results of the PyCBC Search across LIGO–Virgo’s First Three Observing Run .................................................. 47
  5.1 Gravitational-wave Transient Catalogs ........................... 47
  5.2 Exceptional Gravitational-wave Events ........................... 50
  5.3 Conclusions ................................................................... 55

Chapter VI: Stochastic Gravitational-wave Backgrounds ................................................................. 56
  6.1 Properties of Stochastic Backgrounds ............................. 57
  6.2 Sources of Stochastic Backgrounds ................................. 60
  6.3 Detection Methods ........................................................ 64
  6.4 Current Observational Constraints from LIGO–VIRGO–KAGRA . 69

Chapter VII: Model-independent Search for Anisotropies in Stochastic Gravitational-wave Backgrounds ...................... 71
  7.1 SGWB Mapping ............................................................ 72
  7.2 Spectral Shape: Model-independent Approach ................ 84
  7.3 Simulations .................................................................... 85

Chapter VIII: Upper Limits on Anisotropic Backgrounds through LIGO–Virgo’s First Three Observing Runs .................. 90
  8.1 Spectral-model-dependent, Broadband Limits .................. 90
8.2 Spectral-model-independent, Narrowband Limits ................. 93
8.3 Discussion and Future Work .......................................... 95
Chapter IX: Advanced LIGO Detector Commissioning for LIGO–Virgo’s
Third Observing Run ...................................................... 98
  9.1 Implementation of a Real-time Kalman Filter in the LIGO Thermal
      Compensation System .............................................. 98
  9.2 Calibration of Subsystem Cavity Lengths .......................... 108
Chapter X: Summary and Outlook ...................................... 111
  10.1 Summary of the Previous Chapters ............................ 111
  10.2 Outlook .......................................................... 112
Bibliography .................................................................. 114
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The effect of a monochromatic gravitational wave with angular frequency $\omega = 2\pi / T$ propagating perpendicular to the plane of a circular ring of test particles. The lower panel shows the distortion caused by the + and $\times$ polarizations, respectively, in a local inertial frame [36].</td>
<td>12</td>
</tr>
<tr>
<td>2.2</td>
<td>Sources associated with each gravitational wave frequency band and the corresponding types of detectors required for detection [38]. Credit: NASA / J. I. Thorpe.</td>
<td>14</td>
</tr>
<tr>
<td>2.3</td>
<td>A simplified optical layout of an Advanced LIGO detector planned for the fourth observing run (O4). The laser beam from the pre-stabilized laser is stabilized and cleaned in terms of spatial profile, polarization, jitter and frequency noise in the input mode cleaner. The power-recycling mirror increases the beam power to improve the shot noise sensing limit. The cleaned, enhanced beam enters the Fabry-Pérot cavities of the main interferometer. After bouncing back and forth $\sim 300$ times, the beam transmits through the beam-splitter to the signal-recycling mirror, which improves the frequency response. The output Faraday isolator prevents back-reflected light from re-entering the interferometer. The squeezer produces quantum squeezed-vacuum to reduce broadband quantum shot noise. The filter cavity planned for O4 will enable frequency-dependent squeezing. The output mode cleaner cleans spatial and frequency components of the beam before the photodiodes measure the differential arm length [70].</td>
<td>17</td>
</tr>
<tr>
<td>2.4</td>
<td>Strain sensitivities as a function of frequency of the Advanced LIGO Hanford detector in O1 through O3. Only the O3 spectrum includes squeezed light injection. The Advanced LIGO design curve (with 125 W input laser power but no squeezing) and the A+ design sensitivity (with 12 dB squeezing and a factor of two lower coatings thermal noise) are also shown [70].</td>
<td>20</td>
</tr>
</tbody>
</table>
3.1 The inspiral, merger and ringdown stages of the GW signal from a non-spinning BBH merger as a function of time. The different approximation schemes and their range of validity are indicated. Wavy lines illustrate the regime close to merger where analytical methods have to be bridged by numerical relativity [144]. ........................................ 27

3.2 Example GR-modeled gravitational waveforms of BBHs in the time-domain (left) and frequency-domain (right). All waveforms are generated using the SEOBNRv4_ROM [148] waveform approximant. On the top, we show equal mass ratio BBHs of different total masses. In the middle are BBHs of different mass ratios, where the mass of the secondary is $10M_\odot$. On the bottom, $10M_\odot-10M_\odot$ BBHs with different $\chi_{\text{eff}}$ are presented. We see that the higher the total mass and the smaller the aligned spin, the shorter the waveforms above a $f_{\text{min}}$ determined by the detector sensitive band. ........................................ 29

4.1 Antenna patterns of a LIGO GW detector in the small-antenna limit ($L \ll c/f_{\text{GW}}$). The interferometer beam-splitter is located at the center of each pattern and the black lines represent the arms of the interferometer. GW sensitivity of the detector in a certain direction is measured by the distance of the plot surface to the center. At $f_{\text{GW}} = 0$, the left panel is for $+$ polarization, the middle panel is for $\times$ polarization, and the right panel is for unpolarized GWs [196]...... 33

4.2 A flowchart indicating the different steps of the PyCBC pipeline. ... 34

4.3 The template bank used in O3, plotted in the $m_1-m_2$ plane. Each blue dot represents a waveform template. Note that the templates are chosen in a space higher than two dimensions for aligned spinning systems, causing apparent overlap in the $m_1-m_2$ plane. ............... 36

4.4 Example of gating. The 5-second stretch of data is taken from LIGO’s second observing run, injected with two glitches. The threshold of gating is $50\sigma$ in whitened strain. Only one glitch that surpasses this threshold is detected and gated. ......................... 37

4.5 Single-detector reweighted SNRs from LIGO Hanford for BNS injections, BBH injections and noise. For injected GW signals of BBHs and BNSs, the reweighted SNR of the multi-detector matching template is close to the maximum SNR of all triggered templates, while for noise they can be quite different. ......................... 44
4.6 Two-detector template coherence statistic $\alpha$ for BNS injections, BBH injections and noise. With a threshold applied, we can weight down noise events using $\alpha$ with barely any compromise of sensitivity to GW signals. The improvement is especially significant for high-mass BBH events.

5.1 Masses of BHs and NSs in the stellar graveyard detected through GWs and EM waves. This plot contains all GW events detected through GWTC-3 with $p_{\text{astro}} > 0.5$ [247]. Credit: LVK / Aaron Geller / Northwestern University.

5.2 The number of confidently identified CBC candidates with a probability of astrophysical origin $p_{\text{astro}} > 0.5$ versus the detector network’s effective surveyed BNS time–volume [228]. The colored bands mark the different observing runs. The solid black line indicates the cumulative number of probable candidates. The blue line, dark blue band and light blue band are the median, 50% confidence interval and 90% confidence interval for a Poisson distribution fit to the number of candidates at the end of O3b [6].

5.3 Histogram of reweighted SNRs for single-detector triggers in the BNS region. The red curve is for LIGO Livingston (LLO) background events from O1 and O2, the blue curve for LIGO Hanford (LHO) from O1 and O2, and the yellow curve for LLO from the first 50 days of O3. GW190425 is louder than all the background events. Triggers for GW170817 from LHO and LLO are also shown for reference.
6.1 The landscape of potential SGWB signals and various current observational constraints across the frequency spectrum, taken from [69]. The light blue curve shows the prediction for a SGWB from a slow-roll inflationary scalar field [303]. The pink curve shows a hypothetical SGWB from interactions of cosmic strings [304]. The brown curve shows a SGWB from inspiralling supermassive BBHs predicted from the NANOGrav 12.5-year data set [280]. The two grey curves show SGWBs produced by first-order phase transitions at the electroweak scale (∼ 200 GeV) and the QCD scale (∼ 200 MeV), respectively [305]. The yellow curve shows a SGWB generated by compact binary coalescences, based on the mass distributions and merger rates inferred by LVK detections [292]. The dashed curves show various current observational constraints at 95% confidence level. The dotted curve shows the integrated constraint from measurements of $N_{\text{eff}}$ [306], the effective number of events of SGWB signals in a co-moving volume. 61

6.2 Energy density spectra predicted for contributions from BBHs, BNSs and NSBHs to the SGWB and for the combined SGWB due to CBCs [26]. The right panel also plots the 2σ power-law integrated sensitivity curves of Advanced LIGO and Advanced Virgo for current and future configurations. Shaded bands encompass 90% credible level. 62

6.3 Normalized overlap reduction functions (ORFs) $\gamma(f)$ for constituting baselines HL, HV and LV in the HLV detector network in the small antenna limit ($L \ll c/f_{\text{GW}}$). The ORFs show that the correlation of a detector pair to an unpolarized SGWB falls off with frequency rapidly. 66

6.4 Time evolution of the antenna pattern map for unpolarized waves at $f_{\text{GW}} = 0$ in the HLV network in the small-antenna limit. Each subplot shows the antenna pattern at a different hour of a day, in intervals of 3 hours. A full cycle corresponds to one sidereal day. 67
7.1 Point spread functions $\Gamma_{\eta\eta'}$ for the HLV detector network in its O3 sensitivity, assuming a flat strain spectrum for point sources. Black dots denote point sources at particular chosen sky location values $\Theta_{\eta}$. Due to finite coverage of the sky, signals from point sources spread to other locations on the sky. The point spread function is translationally invariant along the right ascension of a source, given data over a full sidereal day. The last subplot shows the interplay of point spread functions for all above carefully placed point sources on a cross on the sky. These use the HEALPix pixelization scheme with $N_{\text{side}} = 16, N_{\text{pix}} = 3072$.

7.2 Fisher matrices in the pixel domain with HEALPix ordering, with HEALPix parameter $N_{\text{nside}} = 16$, for a total of $N_{\text{pix}} = 12N_{\text{nside}}^2 = 3072$ pixels, and corresponding eigenvalues from a singular value decomposition (SVD). From top to bottom, power laws of spectral indices $0, 2/3$ and 3 are represented for the HLV network between 20 and 1726 Hz in its O3 sensitivity. On the left, Fisher matrices contain null directions, leading to very low SVD eigenvalues shown on the right. On the right, we see Fisher matrices of the three-detector HLV network are better conditioned than those of the constituting baselines, HL, HV and LV. Grey dashed lines point to the optimal minimum eigenvalue cutoffs and numbers of eigenmodes kept in each case.

7.3 Comparison of Fisher matrix eigenvalues between O1, O2, O3 and combined O1+O2+O3 analyses for spectral indices of 0, 2/3 and 3. In all cases, O3 contributes the most sensitivity to the combined O1+O2+O3 analysis.

7.4 The left panel shows the condition numbers (i.e., eigenvalue thresholds) and resulting normalized residual sums of squares for power laws of spectral indices 0, 2/3 and 3. We select the threshold returning the least residual sum of squares in a monopole injected simulation for each spectral index. The right panel illustrates the comparison between the Fisher matrix eigenvalue distributions for the HLV network for different values of $\alpha$. The horizontal dashed lines show the optimal thresholds determined via monopole simulations.
7.5 Top to bottom: input, clean, SNR, normalized residual maps and residual histograms from simulations described in Section 7.3. From left to right: monopole ($\alpha = 0$), Gaussian random field ($\alpha = 2/3$), 30 random points ($\alpha = 3$). For all simulations, the pixel with the maximum residual is at the level of a few percent of the injected signal. We have verified that the residuals are Gaussian distributed with norm 1. In the monopole reconstruction, the SNR map presents a characteristic horizontal band due to the shape the ORF traces on the sky over 1 day; this may be also noticed in Fig. 8.4. In the Gaussian field case, the injected map has patches of zero power, and is thus more subject to poor estimation due to noise fluctuations than the other cases shown. In the case of 30 random points, the SNR map presents a residual of the point-spread function with negative values as it is the result of the matrix operation in Eq. 7.22, which can give rise to negative fluctuations where the pixel power is very low.

7.6 Example of recovery of 30 point sources for $\alpha = 0$. The smearing of the point spread is not completely deconvolved due to the singularity of the Fisher matrix in this case. The negative power in $\hat{P}$ is due to noise fluctuations.

7.7 Reconstructed normalized energy densities and spectral dependencies in 30 adaptive frequency bands for maps of a monopole (blue), a Gaussian random field (orange), and 30 random point sources (green) for a non-power-law spectral shape. On the left: $H(f)$; on the right: $\Omega(f)$, as defined in Eqs. (6.15) and (7.1). The spectral shapes are well reconstructed for maps of extended sources (note that the recoveries almost overlap in the plots), whereas for point sources reconstruction is imperfect in particular in the lower frequency bands. See the text for details.

7.8 Left to right are input map for the spectral-model-independent run in 30 adaptive frequency bands and reconstructed clean maps for bands [20, 93.46875] Hz and [1092.71875, 1726] Hz.

8.1 Clean maps, SNR maps, $\Omega_{GW}$ upper limit maps, noise maps and SNR histograms for results integrated over a broadband of frequencies between 20 and 1726 Hz at a reference frequency of $f_{\text{ref}} = 25$ Hz using data from LIGO–Virgo’s first three observing runs. From left to right are for spectral indices 0, 2/3 and 3.
8.2 95% upper limits on the angular power spectrum $C_\ell$ of the SGWB for power laws of $\alpha = 0, 2/3$ and 3 at a reference frequency $f_{\text{ref}} = 25$ using data from LIGO–Virgo’s first three observing runs. We have noted the outlier for $\ell = 6$ in the $\alpha = 3$ case: this is currently under investigation and is believed to be due to a noise fluctuation which makes the point value of $C_6$ negative. .................................................. 93

8.3 95% upper limits on the energy densities in distinct frequency bands used in the spectral-model-independent spectral analysis. We show our results in 10 and 20 bands. The method assumes a scale invariant spectral shape in each spectral band. The results are consistent with noise dominated estimates. .................................................. 94

8.4 Clean maps, SNR maps and noise maps for three frequency bands representative of low, mid and high frequencies in the spectral-model-independent search. In this analysis of the combined O1+O2+O3 data, the search range is divided into 10 bands. From left to right, the plots shown are for frequencies between 20 and 133.125 Hz, between 270.5 and 324.21875 Hz and between 765.8125 and 1726 Hz, and $N_{\text{pix}} = 192, 768, 3072$ respectively. .................................................. 96

9.1 The Advanced LIGO interferometer in Livingston, LA, USA [376]. Credit: Caltech / MIT / LIGO Lab. .................................................. 98

9.2 A schematic of the Advanced LIGO Thermal Compensation System. The Hartmann Wavefront Sensor measures the test mass thermal aberrations. The radiative ring heater and the CO$_2$ laser add heat to minimize thermal gradients in the test masses. The ring heater is installed around all test masses while the CO$_2$ laser actuates only on the compensation plates of the input test masses [1]. .................................................. 99

9.3 The recursive Kalman filter (KF) algorithm, initialized with estimates from the previous time step, consists of the prediction and the correction phases. The KF first provides the $a$ priori estimate based on the previous state in the prediction phase, then updates its estimate using the current measurement in the $a$ posteriori phase. .................................................. 102
9.4 Performance of the Kalman filter (KF) on archival Hartmann Wave-front Sensor (HWS) measurements in the LIGO Livingston Thermal Compensation System. Differential heat was added onto the test mass at the beginning of measurements, and the HWS measured the optical aberration caused by it in terms of the spherical power. The KF estimates are more precise than the HWS measurements.

9.5 Comparison of the effectiveness of the Kalman filter (top panel) and the Bayesian-weighted Kalman filter (bottom panel) on an example 6-hour stretch of archival HWS measurements. The uphill climbs are due to controlled heat added onto the test mass for test mass characterization and TCS testing. The slow downhills trend are due to natural cooling of the test mass, and the steep downhill trends are because of the TCS corrections. The Kalman filter is prone to outliers, while the Bayesian-weighted Kalman filter detects and treats outliers, although imperfectly.

9.6 Optical layout of the Advanced LIGO detectors with seismic isolation and suspensions for O4 [38].
LIST OF TABLES

<table>
<thead>
<tr>
<th>Number</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>6.1</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>6.2</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>7.1</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>8.1</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>8.2</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>8.3</td>
<td>95</td>
</tr>
</tbody>
</table>
9.1 Best fitted model parameters of the state space representations and
of the analytical models for the spherical power of thermo-optical
aberrations. The parameters were obtained by fitting a double ex-
ponential model to the detector measurement data using the Least
Absolute Residuals method in [383]. For the goodness-of-fit statis-
tics for all components, the sums of squares due to error are $O(10^{-6})$,
the root mean squared errors are $O(10^{-4})$, and R-squared and adjusted
R-squared are 1.
Chapter 1

INTRODUCTION: THE DAWN OF GRAVITATIONAL-WAVE SCIENCE

In 1609, Galileo Galilei pointed his “spyglass” (telescope) up to observe the sky and started the field of modern electromagnetic-wave astronomy. At last, humankind was no longer “blind” to the cosmos. Over the course of four centuries, discoveries and surprises flooded in with each opening of frequency bands across the electromagnetic (EM) spectrum.

On September 14, 2015, the Advanced Laser Interferometer Gravitational-wave Observatory (LIGO) detectors [1] made the first direct detection of gravitational waves (GWs), GW150914, coming from the merger of two black holes [2]. GW150914 marked the beginning of gravitational-wave astronomy. At last, humankind was no longer “deaf” to the cosmos. An entire new window parallel to the EM spectrum onto the Universe has opened and was anticipated to trigger a huge wave of insights previously inaccessible and, more excitingly, to reveal surprises.

Merely 2 years later, on August 17, 2017, the Advanced LIGO and Advanced Virgo [3] detectors made the first detection of GWs from two colliding neutrons stars, GW170817 [4]. Subsequently, this event was seen in EM waves across the whole EM spectrum in the following seconds to tens of days by observatories around the globe [5]. The joint GW-EM observations of GW170817 ushered a new era of multi-messenger astronomy.

Ever since then, until the end of LIGO–Virgo’s third observing run (O3), the LIGO Scientific, VIRGO and KAGRA (LVK) collaboration has cataloged dozens more GW signals full of excitement, bringing the total number of GW events to 90 [6]. Among which there are GW190521 [7], where one or both of the progenitor black holes were in the pair instability mass gap [8], resulting in a $142M_\odot$ final black hole, the first detection of an intermediate-mass black hole [9], and GW200105 and GW200115 [10], two neutron star-black hole binaries.

1.1 A Brief History of the Hunt for Gravitational Waves

Gravitational waves were first proposed by Henri Poincaré [11] in 1905, drawing an analogy from electromagnetism. In 1916, Albert Einstein predicted the existence
of GWs [12] after making various approximations and assumptions to his general theory of relativity (GR) [13]. Several theorists, including Einstein himself, doubted this solution to the field equations. Arthur Eddington proved in 1922 that two of the three types of GWs found by Einstein were spurious waves, mathematical artifacts of the coordinate system Einstein used [14]. Einstein then came to believe GWs do not exist, but he eventually became convinced of the existence of GWs after decades. Even then, he was skeptical whether GWs could ever be discovered since they interact with matter so weakly, and if detectable, whether they could be scientifically valuable.

The effect of a GW passing can be described, in the simplest way, as stretching and squeezing spacetime between objects when a GW passes through. Thus in principle, by measuring distances between carefully placed objects, one can detect passing of GWs. But in reality, as we shall discuss more in detail later, this effect is so minuscule that it took thousands of scientists decades [15] to detect these tiny ripples in the fabric of spacetime from some of the most cataclysmic astrophysical processes in the Universe.

In the 1960s, Joseph Weber at the University of Maryland built the first GW detectors – Weber bars to observe resonant vibrations induced in large aluminum cylinders. Weber continuously claimed to have detected GWs in 1969 and onward. However, Weber’s claim was discredited because his proclaimed detection results were not reproducible by other teams.

In 1974, Russell A. Hulse and Joseph H. Taylor Jr. discovered the first binary pulsar system [16] and its orbital decay matched the dissipation of energy in the form of GWs predicted by GR [17–19]. The indirect evidence of GWs boosted morale amongst then disappointed GW enthusiasts, and the search efforts for GWs redoubled. Hulse and Taylor jointly won the Nobel Prize in Physics 1993 "for the discovery of a new type of pulsar, a discovery that has opened up new possibilities for the study of gravitation." [20]

Also starting in the 1960s, the idea of interferometric GW detection was pursued independently by a number of researchers. On the theoretical front, Kip Throne led a group at Caltech to study the theory of GWs and their astrophysical sources and later on noise in GW interferometers in the 1970s. Meanwhile on the experimental side, first prototypes of interferometric GW detectors were built independently by Robert Forward and Rainer Weiss. Later, more sensitive instruments were constructed by Rainer Weiss at MIT and Ronald Drever and Stan Whitcomb at Caltech in the 1970s.
and 1980s, which eventually led to the building of the LIGO project.

In 1988, the LIGO project received funding from the US National Science Foundation (NSF). Barry Barish of Caltech was appointed principal investigator in 1994 and the construction of the two observatories in Hanford, WA and in Livingston, LA broke ground. The construction completed in 1997 and two organizational institutions were formed: LIGO Laboratory, responsible for operations and R&D of the LIGO detectors and test facilities, and LIGO Scientific Collaboration (LSC), responsible for technical and scientific research and data analysis in LIGO.

The LIGO project was conceived in two stages, Initial LIGO and Advanced LIGO. The Initial LIGO configuration searched for GWs from 2002 to 2010 but returned null results [21]. The upgrade to Advanced LIGO took place from 2010 to 2014, and both detectors went into engineering mode in February 2015. Shortly before the first official observing run of Advanced LIGO in September 2015, nature surprised us with a loud and unambiguous chirp, GW150914. And as it turned out, with many more GW signals to follow. The era of gravitational-wave astronomy has only just begun and is beginning to deliver its promises.

The Nobel Prize in Physics 2017 was awarded to Rainer Weiss, Kip S. Thorne and Barry C. Barish “for decisive contributions to the LIGO detector and the observation of gravitational waves.” [22]

1.2 Outline of Chapters

This thesis focuses on the detection methods of two of the four main astrophysical sources of GWs, compact binary coalescences and stochastic gravitational-wave backgrounds. There also contains miscellaneous work done on Advanced LIGO detector commissioning for O3.

Chapter 2 formulates GWs in GR, introduces different categories of astrophysical GW sources, paints the basic idea of the LIGO detectors and highlights the vast areas of science LIGO makes connections to.

Chapter 3 briefly describes compact binary coalescences, focusing on their formation channels and gravitational waveforms predicted by GR. Chapter 4 outlines the frequency-domain matched filtering based PyCBC pipeline to search for GW signals from coalescing binary mergers and reports some improvement efforts towards LIGO–Virgo’s third observing run. Chapter 5 discusses the GW transient catalogs and some exceptional detections in LIGO–Virgo’s first three observing runs and their astrophysical lessons.
Chapter 6 presents foundational knowledge of stochastic gravitational-wave backgrounds such as their properties, sources and detection methods. Chapter 7 details a maximum-likelihood mapping method in the pixel domain to search for anisotropies in stochastic backgrounds, along with a novel model-independent technique to probe their spectral profiles. Chapter 8 summarizes the results of running both modeled and unmodeled anisotropic searches using LIGO–Virgo’s data from the first three observing runs.

Chapter 9 compiles my Advanced LIGO detector commissioning work for O3, done in the summer of 2018 as a LSC fellow at the LIGO Livingston Observatory.

Chapter 10 concludes this thesis and provides an outlook on the future of gravitational-wave science.
Chapter 2

INTRODUCTION TO GRAVITATIONAL WAVES AND LIGO

Static gravitational fields, such as ones created by resting masses, are described as curvature of spacetime in GR – "gravity". Matter and energy curve spacetime, and in return, curvature of spacetime causes matter and energy to move along geodesics. Dynamic gravitational fields, such as ones created by accelerating masses, generate ripples in spacetime in GR – "gravitational waves". GWs are messengers of changing gravitational fields, carrying away information about underlying events.

There are numerous astrophysical sources of GWs, with ground-based detectors concentrating on coalescing binaries, burst sources, continuous waves and stochastic gravitational-wave backgrounds. The Advanced LIGO detectors are ground-based GW interferometers with the most sensitive frequency range from roughly 20 Hz to 2000 Hz [23, 24].

As a GW passes by objects, an observer on Earth would detect that distances between objects increase and decrease periodically, known as the effect of strain. GWs are extremely weak, with a typical strain signal on the order of $10^{-21}$, corresponding to $10^{-18}$ m length change in Advanced LIGO detectors’ 4 km interferometer arms. This is 1000 times smaller than the proton radius of $10^{-15}$ m.

With GWs, we probe fundamental physics by testing gravity theories and studying physical processes in energy scales unachievable in laboratories [25]. We extract physical properties of the most puzzling stellar objects and processes via GW observations and infer their formation and evolution history via rates and populations studies [26]. We also use the standard siren method [27, 28] to study the expansion history of the Universe [29].

2.1 Gravitational Waves in General Relativity

Einstein’s theory of special relativity (SR) [30] in 1905 combines three-dimensional space and time into four-dimensional spacetime and deals with physics in inertial reference frames, i.e., in flat spacetime or in the absence of gravity. Physical laws are invariant under the Lorentz transformation in SR.

Einstein’s GR [13] in 1915 relaxes this constraint and introduces non-inertial frames. Any general transformation is permissible as long as physical laws are consistent
under it. And by the strong equivalence principle, we can equate inertial acceleration with gravitation. Hence, GR becomes a theory to study gravitation.

GWs are a direct consequence of GR [12]. In this section, we follow [31–37] to formulate GWs in GR. Throughout the derivation, we use the standard 4-vector notation $x^\mu \equiv (x^0, x^1, x^2, x^3)$, where the index $\mu$ takes values $\mu = 0, 1, 2, 3$. Here we set $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$. We also imply the use of Einstein summation convention for concise notation.

**Geometry of Curved Spaces**

To characterize a curved space, one cannot specify a global Cartesian coordinate system. In fact, multiple coordinate systems are needed, each on a sufficiently flat patch of the curved space. We thus carry out our analysis in these locally flat patches and borrow the machinery from the flat space case.

The spacetime of GR is a four-dimensional Lorentzian manifold on which a metric with signature $(−, +, +, +)$ is defined. The metric $g$ is a covariant, symmetric, second-rank tensor, which encapsulates the geometric and causal structure of spacetime. For a locally flat region in a curved space, the metric tensor determines the invariant squared interval between infinitesimally adjacent points $x^\mu$ and $x^\mu + dx^\mu$,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \quad (2.1)$$

The metric tensor also connects the covariant and contravariant components of other tensors by lowering and raising indices:

$$g_{\mu\nu} A^\nu = A_\mu, \quad g^{\nu\gamma} A_\nu = A^\gamma, \quad g_{\mu\nu} g^{\nu\gamma} = \delta_\mu^\gamma, \quad (2.2)$$

where $\delta_\mu^\gamma$ is the Kronecker delta.

To parallel transport a contravariant tensor $A^\mu$ from a point $P(x^\kappa)$ to a nearby point $Q(x^\kappa + dx^\kappa)$ in a manifold, we first choose a locally Cartesian coordinate system around $P$ and transform the tensor in general coordinates $A^\mu$ to Cartesian coordinates $A'^\mu$. Then we exploit the constancy of Cartesian coordinates, $\delta A'^\mu = 0$, and transform $A'^\mu$ at $Q$ back to the general coordinates to obtain $A^\mu + \delta A^\mu$ at $Q$.

Following this procedure, we obtain

$$\delta A^\mu = -\Gamma^\mu_{\nu\kappa} A^\nu dx^\kappa, \quad (2.3)$$

where the Christoffel symbols $\Gamma^\mu_{\nu\kappa}$ are given by

$$\Gamma^\mu_{\nu\kappa} = \frac{\partial^2 x^\mu}{\partial x^\alpha \partial x^\beta} \frac{\partial x^\alpha}{\partial x^\nu} \frac{\partial x^\beta}{\partial x^\kappa}. \quad (2.4)$$
The Christoffel symbols $\Gamma^\nu_{\mu\kappa}$ are the affine connection coefficients, which essentially are first derivatives of the metric describing how the local coordinate bases change from point to point.

In terms of the metric, the Christoffel symbols are

$$\Gamma^\mu_{\nu\kappa} = \frac{1}{2} \left( g^\mu_{\nu,\kappa} + g^\mu_{\kappa,\nu} - g^\mu_{\nu\kappa} \right),$$  \hspace{1cm} (2.5)

$$\Gamma^\nu_{\kappa\kappa} = \frac{1}{2} g^\mu_{\nu} \left( g_{\gamma\nu,\kappa} + g_{\gamma\kappa,\nu} - g_{\gamma\nu\kappa} \right),$$  \hspace{1cm} (2.6)

where $g_{\mu\nu,\kappa} := \partial g_{\mu\nu}/\partial x^\kappa$ and respectively.

We assume the contravariant components $A^\mu$ are differentiable by $x^\kappa$ as many times as required. Now we define the covariant derivative of $A^\mu$ to be

$$A^\mu_{;\kappa} := A^\mu_{,\kappa} + \Gamma^\mu_{\nu\kappa} A^\nu,$$  \hspace{1cm} (2.7)

where $A^\mu_{,\kappa} := \partial A^\mu/\partial x^\kappa$. The covariant differential is then $DA^\mu := A^\mu_{;\kappa} \, dx^\kappa$.

Similarly for a covariant tensor $B_\mu$, the covariant derivative and differential are

$$B_{\mu;\kappa} := B_{\mu,\kappa} - \Gamma^\nu_{\mu\kappa} B_\nu,$$  \hspace{1cm} (2.8)

$$DB_\mu = B_{\mu;\kappa} \, dx^\kappa.$$  \hspace{1cm} (2.9)

Curvature is an intrinsic property of space. Nonzero curvature can be inferred if a vector does not return to itself after parallel transporting it over a closed loop. We use the Riemann-Christoffel curvature tensor to describe curvature,

$$R^\mu_{\alpha\beta\gamma} := \Gamma^\mu_{\alpha\gamma,\beta} - \Gamma^\mu_{\alpha\beta,\gamma} + \Gamma^\mu_{\nu\gamma} \Gamma^\nu_{\alpha\beta} - \Gamma^\mu_{\nu\beta} \Gamma^\nu_{\alpha\gamma}.$$  \hspace{1cm} (10.10)

A space is flat if the Riemann tensor is zero everywhere, whereas it is curved if the Riemann tensor is nonzero somewhere.

One important symmetry property of the Riemann tensor is the Bianchi identity,

$$R_{\mu\alpha\beta\gamma:;\sigma} = \frac{1}{3} \left( R_{\mu\alpha\beta\gamma;\sigma} + R_{\mu\alpha\gamma\sigma;\beta} + R_{\mu\gamma\sigma\beta;\gamma} \right) = 0.$$  \hspace{1cm} (2.11)

By contracting the Riemann tensor on the first and the third indices, we attain the Ricci tensor, which is symmetric,

$$R_{\mu\nu} := R_{\mu\nu}^\gamma = 2 \left( \Gamma^\gamma_{\mu\nu,\gamma} - \Gamma^\gamma_{\sigma\nu} \Gamma^\sigma_{\mu\gamma} \right).$$  \hspace{1cm} (2.12)
The Ricci scalar, a scalar measure of curvature, is given by further contracting the Ricci tensor,
\[ R := R_\mu^\mu = g^{\mu\nu} R_{\mu\nu} = 2 g^{\mu\nu} (\Gamma^\gamma_{\mu\nu,\gamma} - \Gamma^\gamma_{\sigma\nu} \Gamma^\sigma_{\gamma\mu}). \] (2.13)

The Einstein tensor is then constructed via the Ricci tensor and the Ricci scalar as
\[ G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}, \] (2.14)
of which the vanishing divergence \( G^\mu_{\nu;\mu} = 0 \) can be proved using the Bianchi identity in Eq. (2.11).

**Einstein Field Equations**

Einstein field equations relate curvature of spacetime with energy and momentum distribution within it and are written as
\[
G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu},
\] (2.15)
where the right hand side contains the stress-energy tensor \( T_{\mu\nu} \) and the left hand side embodies curvature using the Einstein tensor \( G_{\mu\nu} \). The stress-energy tensor and the Einstein tensor are both divergence-free, which means GR preserves the conservation of energy and momentum.

Einstein field equations are second-order, coupled, nonlinear partial differential equations of the metric. Because of the symmetry of the Einstein tensor, we have ten independent equations out of sixteen. The Bianchi identity, or the vanishing divergence of the Einstein tensor, places four constraints on the ten equations, further bringing down the number from ten to six. But this is compensated by four degrees of gauge freedom. Therefore, the problem is well-posed mathematically.

In practice, the field equations are often intractable and exact solutions require strict symmetry conditions that are astrophysically unrealistic. Nevertheless, for exact solutions, we have the Schwarzschild metric for a spherically-symmetric, uncharged, non-rotating black hole, the Reissner-Nordström metric for a charged, non-rotating black hole, the Kerr metric for an uncharged, rotating black hole and the Kerr-Newman metric for a charged, rotating black hole. In cosmology, the Friedmann-Lemaître-Robertson-Walker (FLRW) metric describes the cosmological evolution of a homogeneous, isotropic, expanding, path-connected universe under a uniform distribution of matter and energy.
In the weak field, slow motion limit, Einstein’s equations reduce to Gauss’s law for Newtonian gravity,

$$\nabla^2 \phi = 4\pi G \rho,$$

(2.16)

where \(\phi\) is the Newtonian gravitational potential, \(G\) is the gravitational constant and \(\rho\) is the mass density.

**Linearized Gravity**

Since we expect the waves coming from distant astrophysical sources to be weak, we work out the solutions to Einstein field equations in the weak field regime. When the spacetime metric \(g_{\mu\nu}\) only slightly deviates from the flat Minkowski metric \(\eta_{\mu\nu}\),

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

(2.17)

linearized gravity is adequate to approximate GR. Here, the Minkowski metric is defined as \(\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)\), and \(||h_{\mu\nu}|| \ll 1\) constrains the gravitational field to be weak. Our coordinate system is therefore approximately Cartesian.

To linearize the Einstein tensor, let us start with the Riemann tensor in Eq. (2.10),

$$R_{\alpha\mu\beta\nu} = \frac{1}{2}(g_{\alpha\gamma,\mu\beta} + g_{\mu\beta,\alpha\gamma} - g_{\alpha\beta,\mu\gamma} - g_{\mu\gamma,\alpha\beta}) + \Gamma^2 \text{ terms}. \quad (2.18)$$

Note that \(\eta_{\mu\nu}\) are constants and the derivatives are thus zero. Also, in the weak field limit, \(\Gamma \sim O(h)\). Hence, to first order,

$$R_{\alpha\mu\beta\nu} \simeq \frac{1}{2}(h_{\alpha\gamma,\mu\beta} + h_{\mu\beta,\alpha\gamma} - h_{\alpha\beta,\mu\gamma} - h_{\mu\gamma,\alpha\beta}). \quad (2.19)$$

To proceed, we define the trace reverse of \(h_{\mu\nu}\) to be

$$\bar{h}_{\mu\nu} := h_{\mu\nu} - \frac{1}{2} h \eta_{\mu\nu},$$

(2.20)

where \(h = \eta^{\mu\nu} h_{\mu\nu}\) is the trace of \(h_{\mu\nu}\). Note that in the weak field approximation, we raise and lower indices by \(\eta_{\mu\nu}\).

To simplify the algebra, we exploit gauge freedom and employ the *Lorenz gauge*, also known as the harmonic gauge,

$$\bar{h}_{\mu\nu}^\gamma = h_{\mu\nu}^\gamma - \frac{1}{2} h_{,\mu} = 0. \quad (2.21)$$

In the Lorenz gauge, the Ricci tensor \(R_{\mu\nu}\) becomes

$$R_{\mu\nu} = \eta^{\alpha\beta} R_{\alpha\mu\beta\nu}$$

$$= \frac{1}{2}(h_{\nu,\mu\beta}^\alpha + h_{\mu,\alpha\gamma}^\nu - h_{,\alpha\beta} h_{\mu\gamma,\alpha\beta})$$

$$= -\frac{1}{2} \eta^{\alpha\beta} h_{\mu\nu,\alpha\beta} = -\frac{1}{2} \Box h_{\mu\nu}, \quad (2.22)$$
where $\eta^{\alpha\beta} h_{\mu\nu,\alpha\beta}$ is the d'Alembertian of $h_{\mu\nu}$, written as $\Box h_{\mu\nu}$ in our approximate Cartesian coordinate system.

Then the Ricci scalar $R$ becomes
\begin{equation}
R = \eta^{\mu\nu} R_{\mu\nu} = -\frac{1}{2} \Box h. \quad (2.23)
\end{equation}

Combining Eqs. (2.22) and (2.23), the Einstein tensor $G_{\mu\nu}$ is now given by
\begin{equation}
G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} R = -\frac{1}{2} \Box h_{\mu\nu}. \quad (2.24)
\end{equation}

Finally, in the weak field limit Eq. (2.17), Einstein field equations take the form
\begin{equation}
\Box h_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu}. \quad (2.25)
\end{equation}

**Plane Wave Solutions**

Now we consider the simplest solutions to Einstein field equations, the plane wave solutions. Since astrophysical sources are extremely far from the detectors, the plane wave solutions are accurate enough approximations of the propagation of GWs.

The linearized Einstein’s equations in vacuum are obtained by setting the right hand side of Eq. (2.25) to 0,
\begin{equation}
\Box h_{\mu\nu} = 0. \quad (2.26)
\end{equation}

Eq. (2.26) is the wave equation, so we consider complex wave solutions of the form,
\begin{equation}
\tilde{h}_{\mu\nu} = A_{\mu\nu} e^{ik_\alpha x^\alpha}, \quad (2.27)
\end{equation}
where $A_{\mu\nu}$ is a constant symmetric matrix and $k_\alpha$ is the wave vector indicating the direction of the wave in the spacetime.

Plugging Eq. (2.27) into Eq. (2.26) gives the dispersion relation
\begin{equation}
\Box \tilde{h}_{\mu\nu} = \eta^{\alpha\beta} k_\alpha k_\beta \tilde{h}_{\mu\nu} = 0
\end{equation}
\begin{equation}
\implies k_\alpha k^\alpha = 0. \quad (2.28)
\end{equation}

This implies that the wave vector $k_\alpha$ is null, meaning that GWs travel at $c$, the speed of light, in GR.

Furthermore, imposing the Lorentz gauge given by Eq. (2.21) on the wave solutions Eq. (2.27) yields
\begin{equation}
\tilde{h}_{\mu\nu}^\nu = 0 \implies A_{\mu\nu} k^\nu = 0, \quad (2.29)
\end{equation}
which means the wave vector $k^\nu$ is orthogonal to the matrix $A_{\mu\nu}$. Therefore, GWs are transverse to the direction of propagation in GR.
Transverse Traceless Gauge

In addition to the Lorenz gauge, we can further constrain the metric perturbation $h_{\mu\nu}$ by imposing the *transverse traceless* (TT) gauge. The extra conditions are

$$A^\mu_\mu = 0 \quad \text{(traceless),} \quad (2.30)$$
$$A_{\mu\nu} U^\nu = 0 \quad \text{(transverse),} \quad (2.31)$$

where $U^\nu$ is any fixed timelike vector.

By going to the TT gauge along with the Lorenz gauge, the metric perturbation $h_{\mu\nu}$ is purely spatial and traceless. Originally, $h_{\mu\nu}$ has ten independent components. Fixing the Lorenz gauge imposes four constraints. The TT gauge further introduces four constraints. The metric perturbation $h_{\mu\nu}$ now has only two independent components, which are the two intrinsic polarizations of $h_{\mu\nu}$. The $A_{\mu\nu}$ matrix is denoted by $A_{\mu\nu}^{TT}$ and the metric perturbation $h_{\mu\nu}$ by $h_{\mu\nu}^{TT}$ in this gauge.

Polarizations

In the TT gauge, if we choose our coordinate system such that the wave travels in the $+z$ direction, we can then write

$$A_{\mu\nu}^{TT} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A_+ & A_\times & 0 \\ 0 & A_\times & -A_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (2.32)$$

and the corresponding metric components are $h_+ = A_+ e^{-ik_\alpha^+}$ and $h_\times = A_\times e^{-ik_\alpha^\times}$.

We now define the *plus* ($+$) and *cross* ($\times$) polarization basis tensors corresponding to each polarization,

$$e^{+}_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad e^{\times}_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (2.33)$$

In terms of these basis tensors, the general wave is a linear combination of the two polarizations and can be written as

$$h_{\mu\nu}^{TT} = h_+ e^{+}_{\mu\nu} + h_\times e^{\times}_{\mu\nu}. \quad (2.34)$$
The effect of a monochromatic gravitational wave with angular frequency $\omega = 2\pi/T$ propagating perpendicular to the plane of a circular ring of test particles. The lower panel shows the distortion caused by the $+$ and $\times$ polarizations, respectively, in a local inertial frame [36].

**Effect of Gravitational Waves**

Consider a circular ring of test particles initially at rest in the plane of this page, the $x$-$y$ plane. When a GW traveling along the $z$ direction is incident on this ring, the passing wave would distort the spacetime the ring lives in. Subsequently, the ring is stretched and squeezed periodically by the wave as illustrated in Fig. 2.1 for a plus or a cross polarized GW, respectively. For the plus polarization, these fluctuations are along the $x$ and $y$ axes, while for the cross polarization, they are long the $y = x$ and $y = -x$ lines. The effect of the passage of a GW is thus to cause a tidal deformation of objects, transverse to the direction of the wave propagation.

**Generation of Gravitational Waves**

The discussion so far has been on the propagation of GWs, setting the source term to zero. Now we return to the linearized Einstein field equations in Eq. (2.25) to address how GWs are generated by accelerating masses.

The solutions to Eq. (2.25) can be attained by the radiative Green’s functions for
inhomogeneous wave equation. Let the observer’s position be at \(x\) and the source position be at \(x'\). Then the distance between the observer and the source is \(|x - x'|\). The solutions are then given by

\[
\tilde{h}_{\mu\nu}(t, x) = \frac{4G}{c^4} \int d^3x' \frac{[T_{\mu\nu}'(t', x')]_{\text{ret}}}{|x - x'|},
\]

(2.35)

where \([T_{\mu\nu}'(t', x')]_{\text{ret}}\) is the retarded stress-energy tensor at the retarded time \(t' = t - |x - x'|/c\).

Assuming that the size of the source is much smaller than its distance to the observer \(|x'| \ll |x| = r\) and the source is moving at a non-relativistic speed \(v \ll c\), we can replace \(|x - x'|\) by \(r\) in Eq. (2.35),

\[
\tilde{h}_{\mu\nu}(t, x) = \frac{4G}{c^4 r} \int d^3x' T_{\mu\nu}(t - r/c, x').
\]

(2.36)

Utilizing the divergence-free and symmetric properties of the stress-energy tensor, we can further simplify the formula to

\[
\tilde{h}_{\mu\nu}(t, x) = \frac{2G}{c^4 r} \partial_0^2 \int d^3x' x'_\mu x'_\nu T^{00}(t - r/c, x').
\]

(2.37)

Here, \(T^{00} = \rho c^2\) is the mass-energy density and we define the second moment of mass-energy to be

\[
I_{\mu\nu}(t') = \int d^3x' x'_\mu x'_\nu \rho(t', x').
\]

(2.38)

Using these and then projecting the quantities to the TT gauge where \(\tilde{h}_{\mu\nu} = h_{\mu\nu}\), we finally arrive at the quadrupole formula,

\[
h_{\mu\nu}^{TT}(t, x) \approx \frac{2G}{c^4 r} \partial_0^2 I_{\mu\nu}^{TT}(t - r/c).\]

(2.39)

We remark that the GW amplitude falls with distance as \(1/r\) and the leading order emission is quadrupole gravitational radiation.

### 2.2 Sources of Gravitational Waves

From Eq. (2.39), any massive system with an accelerating quadrupole moment generates GWs. In this section, we go over the GW spectrum and summarize the main sources for ground-based detectors.
Gravitational-wave Spectrum

We expect different classes of GW sources spanning a wide range of frequencies with different types of detectors required for detection, as shown in Fig. 2.2.

For terrestrial detectors sensitive to frequencies ranging from few Hz to few kHz, stellar-mass compact binaries composed of neutron stars (NSs) or stellar-mass black holes (BHs) of masses \( \sim 5-100M_\odot \) merge in this “audible band” [39]. The merger frequency occurs at the closest approach of the two bodies, the Schwarzschild radius of the system \( a_{\text{min}} = 2GM/c^2 \), where \( M \) is the total mass. Kepler’s third law relates the orbital period \( P \) with the total mass \( M \) and the orbital separation \( a \), \( P^2 \propto M^{-1}a^3 \). Since the GW frequency \( f \) is twice the orbital frequency \( \omega = 2\pi/P \), we have \( f \propto M^{1/2}a^{-3/2} \). At merger, \( f_{\text{merger}} \propto M^{1/2}a_{\text{min}}^{-3/2} \propto M^{-1} \). Hence, the more massive the binary, the smaller the final merger frequency. Rapidly spinning pulsars with tiny asymmetries radiate weak but continuous waves that also fall into this band. Additionally, asymmetric supernovae emit bursts of GWs in the few hundred Hz range. And finally, stochastic backgrounds consisted of unresolved sources or primordial GWs are also an ambitious target in this band.

For space-based interferometers, the targets are GW sources in the mHz and deci-Hz range. For one, supermassive black holes (SMBHs) [40] of masses \( 10^6M_\odot \) to \( 10^9M_\odot \) merge in the mHz range. Another promising source is extreme mass
ratio inspirals (EMRIs) [41], where a stellar BH or a NS spirals into a SMBH. Moreover, space-borne GW detectors can monitor early inspirals of stellar-mass binaries weeks to years before the binaries enter the sensitive band of ground-based detectors, allowing for early alerts. They also provide a window to observe GW backgrounds. Current efforts include the planned Laser Interferometer Space Antenna (LISA) [42] joint mission of ESA and NASA and the TianQin [43] space mission of China both launching in the 2030s, as well as the proposed Deci-hertz Interferometer Gravitational wave Observatory (DECIGO) [44, 45].

In the nHz range, merging SMBHs and GW backgrounds are the primary targets for pulsar timing arrays (PTAs) [46, 47]. Pulsing radio beams from magnetic poles of spinning neutron stars sweep by the Earth in precise time intervals. One can track an array of pulsars to monitor jitters of beam arrival times and then correlate these jitters to detect passage of GWs. The Square Kilometre Array (SKA) [48] and the Five-hundred-meter Aperture Spherical radio Telescope (FAST) [49] are planned radio telescopes to follow PTAs for extreme tests of GR.

We reach the lowest frequencies from $10^{-17}$ Hz to $10^{-15}$ Hz with the B-mode polarization anisotropy [50–52] of the Cosmic Microwave Background (CMB) [53, 54]. The CMB B-mode may contain the imprint of a primordial GW background, providing a test for inflationary models. Planck, WMAP, BICEP/Keck Array are some of the CMB probes setting constraints on primordial GWs [55].

Now we briefly describe the four main types of sources for ground-based detectors and refer interested readers to consult [34, 35] for a comprehensive discussion.

**Compact Binary Coalescences**

Compact binary coalescences (CBCs) are comprised of binary black hole (BBH), binary neutron star (BNS) and neutron star-black hole (NSBH) mergers within the frequency band of ground-based detectors. When a binary coalesces, it loses energy in the form of GWs, leading to reduction in its orbital distance. As the orbit shrinks, the orbital period is shorter and the system emits stronger waves and loses energy at an even faster rate. This runaway process endows the signal with a chirp-like characteristic, increasing both in frequency and in amplitude as it draws closer to merger.

CBCs have been the only detected group up till LIGO–Virgo’s third observing run [6]. BHs are not directly observable in EM waves. GWs, instead, are messengers of their existence and a great tool to study their properties and dynamics [56]. On
the other hand, NSs are still very much an enigma. Just like how the first GW signal from a BNS, GW170817, enabled us to learn so much about NSs, GWs will continue guiding us to solve the mystery of them.

We will come back to CBCs in Chapter 3 for a more detailed introduction.

**Bursts**

Unless there is a complete symmetry in the explosion of a supernova, it will generate GWs. Such GWs are not well modeled, because supernovae are not well understood. Asymmetric core-collapse supernovae can leave burst-like signals in the LIGO band [57]. Instabilities in NSs can also be a source of burst signals [58, 59]. In addition, GWs associated with gamma-ray bursts [60, 61], fast radio bursts [62], magnetar flares [63], soft gamma repeaters [64], cosmic string cusps [65], etc., are also searched for in the LVK collaboration. And more excitingly, there might be sources presently unknown.

**Non-axisymmetric Pulsars**

Pulsars, or spinning NSs, are symmetric to a high degree for their high densities and high angular velocities [58, 59]. Perfectly symmetric objects have no accelerating mass quadrupole and thus do not generate GWs. In case there exist tiny deformations on a pulsar, it will emit GWs categorized as continuous waves with fairly constant amplitudes and frequencies [66]. Although continuous waves are comparably weaker than other types, we can integrate them over longer periods to accumulate detectable signals [67, 68].

**Stochastic Gravitational-wave Backgrounds**

GWs not localized in either time or frequency constitute background signals. We expect background signals of two origins, astrophysical and cosmological [69]. The astrophysical background is a collection of individually unresolved sources, while the cosmological background is the relic gravitational radiation from the early Universe. We defer to Chapter 6 for a more in-depth introduction.

**2.3 Advanced LIGO Detectors**

**Detector Design**

The Advanced LIGO detectors are squeezing enhanced, dual-recycled, Fabry-Pérot, Michelson interferometers consisted of two 4 km arms [70]. A Michelson interferometer splits a coherent laser beam into two equal power beams along its two
Figure 2.3: A simplified optical layout of an Advanced LIGO detector planned for the fourth observing run (O4). The laser beam from the pre-stabilized laser is stabilized and cleaned in terms of spatial profile, polarization, jitter and frequency noise in the input mode cleaner. The power-recycling mirror increases the beam power to improve the shot noise sensing limit. The cleaned, enhanced beam enters the Fabry-Pérot cavities of the main interferometer. After bouncing back and forth ~300 times, the beam transmits through the beam-splitter to the signal-recycling mirror, which improves the frequency response. The output Faraday isolator prevents back-reflected light from re-entering the interferometer. The squeezer produces quantum squeezed-vacuum to reduce broadband quantum shot noise. The filter cavity planned for O4 will enable frequency-dependent squeezing. The output mode cleaner cleans spatial and frequency components of the beam before the photodiodes measure the differential arm length [70].
orthogonal equal length arms with a beam-splitter. A passing GW would stretch one arm and squeeze the other alternately, creating differences in light travel paths. The two beams are then reflected by test masses at the end of each arm and recombined at the beam-splitter to form an interference pattern observed on photodiodes at the detection port.

In the long wavelength limit $\lambda_{GW} \gg L$, where the GW wavelength $\lambda_{GW}$ is much larger than the length $L$ of interferometer arms, the GW amplitude $h$ is approximately given by the fractional change in the differential length $\triangle L$ of interferometer arms,

$$h = |h_{x} e^{+}_{\mu \nu} + h_{x} e^{x}_{\mu \nu}| \sim \frac{\triangle L}{L}.$$  \hspace{1cm} (2.40)

Most systems in the detectors are designed to amplify this minute length change or to mitigate noise faking or masking GWs. The laser beams are enclosed in vacuum tunnels to eliminate light scattering. The Fabry-Pérot optical cavities in each arm effectively increase the arm length a few hundred times. Power recycling improves the sensitivity by an order of magnitude by increasing the beam power [71]. Signal recycling is used to broaden the detector bandwidth [72]. Squeezing reduces the inherent uncertainty in photon arrival time measurements [73].

A schematic of an Advanced LIGO detector is presented in Fig. 2.3.

**Noise Sources**

In low frequencies ($\sim 10$ Hz), the ground-based GW detectors are limited by seismic noise. The LIGO detectors use suspended pendulums up to four stages to isolate seismic noise. In a suspended pendulum, movements reaching the suspension point are reduced by a factor of $1/f^2$, for frequencies much larger than the pendulum’s resonant frequency. In addition to passive isolation, ground movements sensed by seismometers and accelerometers are fed back to actively reduce seismic noise.

The mid frequencies ($\sim 100$ Hz) of the sensitive band are dominated by thermal noise caused by microscopic fluctuations of atoms in the mirrors and their suspensions. The thermal energy $k_B T/2$ per degree of freedom, where $k_B$ is the Boltzmann constant and $T$ is the temperature in Kelvin, excites different oscillation modes in the mirrors and their suspensions. Thermal noise is combated by choosing fused silica to be the material of mirrors and suspensions. Fused silica has the property of having very low loss of energy from its normal modes of vibration (high quality factor), thereby confining thermal noise to narrow frequency bands.
Quantum noise prevails at high frequencies (above \( \sim 200 \) Hz) of the sensitive band. Quantum noise is due to shot noise arising from Poisson fluctuations in photon arrival time at the detection port and photon radiation pressure arising from laser power variations. Based on the Heisenberg uncertainty principle, increasing laser power (not indefinitely) and utilizing power recycling diminish shot noise from photon counting. Squeezing further reduces noise in the phase quadrature at the expense of larger amplitude fluctuations, allowing for more accurate photon arrival time measurements.

In Fig. 2.4, we show representative GW sensitivity spectra of the Advanced LIGO Hanford detector in the first three observing runs (O1, O2 and O3), as well as the Advanced LIGO design curve planned for the fourth observing run (O4) and the target curve of the future upgrade, A+, for the fifth observing run (O5). The three aforementioned noise sources essentially determine the sensitivity curve while there exist many more other noise sources.

For more details on the detectors, we refer interested readers to [1, 70, 74–76].

**A Global Detector Network**

We need multiple detectors to confidently detect and locate GW sources. In the United States, there are two such GW interferometers located in remote areas: one in Hanford, WA, the other in Livingston, LA. These detectors are 3000 km apart, so signals can arrive at the two detectors with delays up to 10 ms.

Laser interferometric GW detectors operate as antennas, sensitive to large portions of the sky but unable to locate sources in the sky. A global network of detectors is needed for accurate sky localization of sources [76]. Currently, active GW interferometers include Advanced LIGO Hanford, Advanced LIGO Livingston, Advanced Virgo in Italy and KAGRA [77] in Japan. A fifth detector, LIGO-India [78], is in the planning phase and is expected to join the network by the end of 2020s.

There are also proposed third-generation ground-based detectors, Einstein Telescope [79, 80] in Europe and Cosmic Explorer [81–83] in the United States, which will have the sensitivity to detect all binary mergers in the Universe.
Figure 2.4: Strain sensitivities as a function of frequency of the Advanced LIGO Hanford detector in O1 through O3. Only the O3 spectrum includes squeezed light injection. The Advanced LIGO design curve (with 125 W input laser power but no squeezing) and the A+ design sensitivity (with 12 dB squeezing and a factor of two lower coatings thermal noise) are also shown [70].

2.4 Overview of LIGO Science

Fundamental Physics

So far, GR has passed all experimental and observational tests [84] in linear or mildly nonlinear regimes. However, GR’s inconsistency with quantum mechanics suggests that it may be a low-energy limit of a more general theory amenable to quantization. GR makes specific predictions about the nature of GWs and thus GWs provide a unique opportunity to test GR in a strongly nonlinear and highly relativistic domain.

GWs can be used to test strong-field predictions of GR by looking for deviations from numerical relativity modeling of the late inspiral stages of binary mergers [85]. Spectra of quasi-normal modes of black holes inferred by ringdown measurements of binary mergers are also a strong-field gravity test [86]. Moreover, any variations of the speed of GWs [87], Lorentz and parity violations [88, 89] and more polarization states than just the tensor polarization [90] may eliminate GR and place
constraints on alternative theories of gravity [25]. Beyond testing GR, detection of a GW background may reveal physical processes during phase transitions at different energy scales in the early Universe [91].

**Astrophysics**

From observed GWs, we can extract physical parameters of exotic stellar objects and cataclysmic events such as neutron stars, black holes and supernovae [92]. For compact binaries, physical observables such as binary masses, spins, orientations, etc., may shed light on their poorly understood formation channels [93]. We can learn about the underlying populations of sources and rates of events from an ensemble of detected individual events [26]. For neutron stars, tidal deformation encodes the key to resolve the dense nuclear equation of state [94]. Detection of GWs from galactic supernovae contribute directly to understanding the core-collapse supernova mechanism [95]. And detection of an astrophysical GW background may constrain the star formation history [96, 97].

**Cosmology**

We can make cosmological measurements, such as the Hubble constant, using detected GWs. A GW source is a “standard siren” [27, 28] of known loudness, allowing for a measurement of the luminosity distance to the source. Combined with a redshift measurement, either from an observation of an EM counterpart or from statistical calculations by overlaying galaxy catalogs onto source localization, we can probe the expansion history of the Universe [29].
COMPACT BINARY COALESCENCES

Compact binary coalescences of BHs and/or NSs are the most prominent GW sources in the frequency band of ground-based detectors. Such binary systems lose orbital energy via GW emission and are compact enough to merge within the age of the Universe. Up till the end of O3, the LVK collaboration has registered 90 compact binary mergers covering all flavors (BBHs, BNSs and NSBHs) detectable in the LVK band in the third Gravitational-wave Transient Catalog (GWTC-3) [6]. The population is expected to increase significantly in upcoming observing runs with improved sensitivities of existing detectors and addition of new ground-based interferometers [76].

Compared to the fast-growing population of detected CBCs, the theoretical framework for the formation and evolution mechanisms of merging compact binaries is still highly uncertain because of uncertainties and degeneracies of the astrophysical models involved. Two main formation channels are typically considered in the literature, isolated field binaries [98–105] and dynamical interactions [105–115]. More recently, a hybrid scenario of the two has also been suggested [116–119]. With direct evidence of their existence and their GW (and possibly EM) signatures, we are presented with an unprecedented opportunity to study the physical mechanisms of CBCs and their progenitor stars.

To detect GW signals from the inspiral, merger and ringdown phases of CBCs, the LVK collaboration employs both modeled searches and unmodelled or burst searches. CBCs are the most extensively modeled target sources of GWs with precise expected gravitational waveforms computed based on GR, thanks to perturbative approximations [120] and numerical relativity [121]. Modeled searches are thus able to use matched filtering as the (nearly) optimal technique to detect these weak signals buried in detector noise. On the other hand, burst searches assume minimal information about waveforms and look for generic transient signals in detector data.

In this chapter, we review the two main formation channels for compact binaries following [122] and briefly discuss aspects of GR-modeled gravitational waveforms for CBCs. We defer introduction to the matched filtering technique and PyCBC pipeline specific discussions to Chapters 4 and 5.
3.1 Astrophysical Processes

In theory, two main formation channels have been proposed to form compact binaries. One is the isolated binary channel [98–105], where two progenitor stars are gravitationally bound at birth, evolve in isolation into a compact binary towards the end of their life and merge relatively unperturbed in a diffuse stellar environment. And the other is the dynamical channel [105–115], where two compact remnants draw close to each other and are bound after several gravitational interactions with each other and other objects in a dense stellar environment. More recently, a hybrid scenario, where dynamical interactions eject a compact binary from its cluster to merge alone in the field, has been suggested [116–119]. Other scenarios that have been proposed include formation within the accretion disks of active galactic nuclei (AGNs) [123, 124], chemically homogeneous evolution [125–127], hierarchical mergers [128], and more [129].

Binary Evolution

It has been shown that more massive stellar objects tend to form stellar systems of higher multiplicity [130], suggesting most BH and NS progenitors are members of binaries, triples, and even quadruple systems. The timescale of a binary coalescence due to GW emission is given by [131]

$$t_{GW} = \frac{5}{256} \frac{c^4 a^4}{G^3 m_1 m_2 (m_1 + m_2)} ,$$

where $m_1$ is the mass of the primary star, $m_2$ is the mass of the secondary and $a$ is the semimajor axis of the binary. For example, for a compact binary consisting of two $10M_\odot$ BHs to merge within a Hubble time, the separation is thus smaller than $\sim 0.1$ AU. However, at solar metallicity, the stellar progenitor of a $10M_\odot$ BH has a radius of $\approx 0.18$ AU. Therefore, the stellar progenitors of such a system must have been born at a wider separation and have been brought closer together by other mechanisms (e.g., common envelope, tidal dissipation, natal kicks, etc) [129].

Common Envelope – A way to transfer mass from a star to its companion is through Roche lobe overflow. The Roche lobe defines a region of a star within which orbiting material is gravitationally bound to the star. Approximately teardrop-shaped, the Roche lobe sets the boundary of the critical gravitational potential. In a binary, the Roche radius of the critical gravitational potential of the primary star approximately writes [132]

$$R_{L,1} = a \frac{0.49 q^{2/3}}{0.6 q^{2/3} + \ln(1 + q^{1/3})} ,$$

where $q = m_2 / m_1$. 

where \( q = m_1/m_2 \geq 1 \) is the mass ratio of the primary to the secondary. The two Roche lobes in a binary are connected at the first Lagrangian point \( L_1 \) of the system.

The common envelope (CE)\(^{[98, 133]} \) evolution is a process that can shrink a binary separation considerably \(^{[103, 104]} \). At first, both stars lie within their Roche lobes. When a star overfills its Roche lobe, mass transfer from the donor to the accretor kicks off. Unstable mass transfer further shrinks both the binary orbit and the sizes of the Roche lobes, causing an accelerating Roche lobe overflow and hence an accelerating mass transfer. At a certain point, this runaway process of dynamically unstable mass transfer exceeds the accretion ability of the companion. Excess material engulfs the accretor, and eventually forms a common envelope encompassing both stars. With a shift of its center of mass, the gaseous envelope exerts a drag force onto the secondary and outsets an inspiral of the binary. During the inspiral, orbital energy (as heat) and angular momentum are transferred to the envelope and consequently heats up and expands the envelope. Towards the end of CE, either the envelope is expelled, leaving behind a short-period binary or the two cores are tidally disrupted and merge into a single star.

**Stellar Tides** – A star in a close stellar binary (or a compact star in a compact binary such as a white dwarf or a NS) experiences the tidal field of its companion and generates a gravitational quadruple moment due to tidal deformations. This quadruple moment causes loss of the binary’s orbital energy, with the precise dissipation process and rate depending on the stellar structure \(^{[134, 135]} \). Tidal dissipation tends to circularize eccentric binaries and spin up stars in close binaries. Tidal spin-up drives synchronization of both magnitude and direction of the component spins with the angular momentum of the system. Another consequence of tides in close stellar binaries is rotational mixing \(^{[136, 137]} \) of the stellar interior leading to chemical homogeneous evolution. Chemically-homogeneous stars skip the giant phase and thus remain compact to evolve close to each other without merging, which is a potential pathway to form aligned-spin, massive (\( > 20M_\odot \)), nearly equal mass ratio (\( q \approx 1 \)) BBHs \(^{[125–127]} \).

**Natal Kicks** – Supernovae can change binary orbits by imparting natal kicks, either tightening or loosening binaries. There are two main physical mechanisms leading to natal kicks. The first kind is the sudden mass loss during a supernova explosion. Impulsive mass ejection decreases the gravitational potential and thus loosens or even unbinds a binary. For a binary of total mass \( M \) and semimajor axis \( a_0 \) losing a
mass $\Delta m$ instantaneously, the semimajor axis $a_1$ after the explosion is \[3.3\]

$$\frac{a_1}{a_0} = \frac{M - \Delta m}{M - 2 \Delta m a_0 r},$$

where $r$ is the distance between the two bodies at the time of the explosion ($r \equiv a_0$ for a circular binary). The second kind is the velocity kick caused by asymmetries in the supernova ejecta. Depending on its direction with respect to the orbital velocity, the velocity kick can tighten or loosen binaries and change the eccentricity of orbits. Strong velocity kicks can result in significantly misaligned or anti-aligned spins in isolated binaries \[139\]. Natal kicks can also cast out compact objects and binaries from their birth star clusters.

**Dynamical Interactions**

A star cluster is a relatively dense region of stars bound by gravity consisting of a high-density core and a low-density halo surrounding the core. In a cluster, heavier stars tend to sink towards the core and travel at a slower pace while lighter ones flow into the halo (or even escape the cluster) and travel faster. This core collapse process increases the density of the core as well as the probability of gravitational encounters at the core. Cores of star clusters are thus an ideal environment for dynamical interactions among stars \[140\].

In a star cluster, a binary is considered hard if its gravitational binding energy $E_{\text{bin}} = G m_1 m_2/(2a)$ is higher than the average kinetic energy $E_k = \bar{m} v_\infty^2/2$ of its neighboring stars, and soft if lower. Stellar dynamics induced by the core collapse of a cluster harden hard binaries, shrinking their semimajor axes, and soften soft binaries, widening their semimajor axes. Moreover, dynamical interactions can trigger orbital eccentricities of binaries. High eccentricities greatly shorten the binary coalescence timescale, since for eccentricities close to 1, the coalescence time now scales as $t_{GW} \propto (1 - e^2)^{7/2}$, where $e$ is the binary eccentricity \[141\].

Binary hardening is arguably one of the most important processes to form BBHs. BHs concentrated at the core of a star cluster repeatedly go through three-body encounters during the cluster core collapse. This process ends until one of the three outcomes realizes: (i) the core collapse is reversed and the dynamical interaction rate drops; (ii) the binary is ejected from the cluster due to its high recoil velocity; (iii) the binary merges due to GW emission.

During three-body encounters, one member of the binary may be swapped with the initially free body, usually with the lightest body among the three being swapped out
of the binary. Therefore, dynamical encounters prefer forming BBHs with higher masses and equal mass ratios. Dynamically constructed binaries also do not have correlated spins. Moreover, most of such binaries are circularized via GW emission by the time they reach the LVK band, washing out any residual eccentricity at $> 10 \text{ Hz}$ [142]. Only those with extreme eccentricities may leave some of their eccentricities above 10 Hz.

Other processes of stellar dynamics leading to GW mergers include gravitational interactions in small-N systems such as triples, quadruples and higher hierarchical systems [128] and within AGN disks [123, 124].

### 3.2 Gravitational Waveforms

**Inspiral, Merger and Ringdown**

GW signals from CBCs can be typically split into three phases: inspiral, merger and ringdown (IMR), as illustrated in Fig. 3.1 for a non-spinning BBH merger. As the orbit shrinks, the GW frequency $f_{\text{GW}} = 2f_{\text{orb}}$ increases with time as well as the GW amplitude. Such a signal with a monotonically increasing frequency as a function of time is called a *chirp*.

Since no analytic solutions to the Einstein’s equations for CBCs are known, gravitational waveforms from the IMR of CBCs are constructed by combining perturbative calculations [120] in GR with numerical relativity simulations [121]. The early inspiral phase can be well approximated by a post-Newtonian expansion in powers of $v/c$, where $v$ is the binary orbital velocity. The late inspiral and merger phases are calculated through numerical relativity solving the full Einstein field equations. The final ringdown phase is computed using the BH perturbation theory.

**Inspiral** – The inspiral phase is when the two compact objects in a binary are relatively far apart, with a separation larger than its innermost stable circular orbit (ISCO). In general, the higher the total mass and the smaller the aligned spin of the system, the larger the ISCO. Consider a test particle orbiting a BH of mass $m$ [143]:

(i) if the BH is a Schwarzschild BH, the ISCO is at $r_{\text{ISCO}} = 6Gm/c^2$; (ii) if the BH is a Kerr BH with aligned spins, the ISCO is at $r_{\text{ISCO}} = Gm/c^2$; (iii) if the BH is a Kerr BH with anti-aligned spins, the ISCO is at $r_{\text{ISCO}} = 9Gm/c^2$. The inspiral of a CBC can last over millions of years before heading into the highly relativistic, highly nonlinear regime of the merger.

**Merger** – The merger phase is an intermediate stage joining the inspiral and ringdown, when the two compact objects are close enough to collide and merge into a
Figure 3.1: The inspiral, merger and ringdown stages of the GW signal from a non-spinning BBH merger as a function of time. The different approximation schemes and their range of validity are indicated. Wavy lines illustrate the regime close to merger where analytical methods have to be bridged by numerical relativity [144].

single BH. The strongest GWs in the coalescence are emitted in this phase. Since this process is highly nonlinear ($v/c$ is approaching 1), perturbative approximations are known to not converge and numerical relativity comes to the rescue, albeit computationally expensive. The merger only lasts a fraction of a second.

Ringdown – The ringdown phase defines the time after forming a single unperturbed rotating BH. This is a period when the gravitational waveform decays exponentially. The GW signal during the ringdown resembles an oscillating function with an exponentially damped amplitude. Here again, duration is a fraction of a second.

Parameter Space
In the framework of GR, to fully describe a gravitational waveform from a CBC requires at least 15 parameters listed in Table 3.1, with 8 source intrinsic parameters (masses and spins) and 7 location and orientation dependent extrinsic parameters. Note that we consider only quasi-circular binaries since most eccentricities are washed out by the time GW signals reach the LVK band as explained earlier. We do not consider matter effects in BNS and NSBH mergers as well since they have negligible contributions to overall search sensitivities, but they are particularly important in constraining the equation of state of NSs and other astrophysical mechanisms and hence are considered in parameter estimation after detections.

Gravitational waveforms expected from CBCs based on GR can be well modeled by combinations of the intrinsic binary component parameters. To first order, the phase evolution of signals from the inspiral of a binary is determined by the chirp
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_1, m_2$</td>
<td>component masses</td>
</tr>
<tr>
<td>$\vec{S}_1, \vec{S}_2$</td>
<td>component spin angular momenta</td>
</tr>
<tr>
<td>$\alpha, \delta$</td>
<td>sky position: right ascension, declination</td>
</tr>
<tr>
<td>$i$</td>
<td>inclination of the binary orbital plane w.r.t. the line of sight</td>
</tr>
<tr>
<td>$\psi$</td>
<td>polarization direction w.r.t. detector arms</td>
</tr>
<tr>
<td>$D$</td>
<td>luminosity distance</td>
</tr>
<tr>
<td>$\phi_c$</td>
<td>coalescence phase</td>
</tr>
<tr>
<td>$t_c$</td>
<td>coalescence time</td>
</tr>
</tbody>
</table>

Table 3.1: The minimal parameter space of gravitational waveforms from CBCs in the framework of GR. At least 15 parameters are needed to fully describe a CBC waveform. Masses and spins are source intrinsic parameters and others are location and orientation dependent extrinsic parameters.

mass of the system [141, 145]

$$M = \frac{(m_1m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \equiv \eta^{3/5}M, \quad (3.4)$$

where $\eta = m_1m_2/(m_1 + m_2)^2$ is the symmetric mass ratio and $M = m_1 + m_2$ is the total mass of the binary. In the merger and ringdown stages, we usually use the total mass $M$ and the mass ratio $q = m_2/m_1 \leq 1$ to describe a binary system.

Using the spin angular momentum $\vec{S}_i$ for component $i = \{1, 2\}$, we define the dimensionless component spin vector to be

$$\vec{\chi}_i = \frac{c\vec{S}_i}{Gm_i^2}, \quad (3.5)$$

which ranges in magnitude from 0 (non-spinning) to 1 (Kerr limit) for BHs. The two component spins are then combined to construct an effective inspiral spin of the binary via [146, 147]

$$\chi_{\text{eff}} = \frac{(m_1\vec{\chi}_1 + m_2\vec{\chi}_2) \cdot \hat{L}_N}{M}, \quad (3.6)$$

where $\hat{L}_N$ is a unit vector along the orbital angular momentum of the binary. $\chi_{\text{eff}}$ has the meaning of the mass-weighted projection of component spins parallel to the orbital angular momentum. A positive value indicates a net aligned spin with the orbital angular momentum while a negative value indicates a net anti-aligned spin.
Figure 3.2: Example GR-modeled gravitational waveforms of BBHs in the time-domain (left) and frequency-domain (right). All waveforms are generated using the SEOBNRv4_ROM [148] waveform approximant. On the top, we show equal mass ratio BBHs of different total masses. In the middle are BBHs of different mass ratios, where the mass of the secondary is $10M_\odot$. On the bottom, $10M_\odot$–$10M_\odot$ BBHs with different $\chi_{\text{eff}}$ are presented. We see that the higher the total mass and the smaller the aligned spin, the shorter the waveforms above a $f_{\text{min}}$ determined by the detector sensitive band.
We also define a dimensionless quantity to describe the mass-weighted in-plane spin component contributing to spin precession, the effective precession spin,

\[ \chi_p = \max \left( \chi_{1,\perp}, \frac{q(4q + 3)}{4 + 3q} \chi_{2,\perp} \right), \tag{3.7} \]

where \( \chi_{l,\perp} \) is the component of spin perpendicular to the orbital angular momentum \( \hat{L}_N \). \( \chi_p \) ranges from 0 (no precession) to 1 (maximal precession).

\( \chi_{\text{eff}} \) is the combination of the six spin components that is best measured from the GW waveform; and \( \chi_p \) is next. For finite signal-to-noise ratio signals, the other four degrees of freedom of spins are less well-constrained. However, we can (crudely) measure the spin tilt angles at which the spin axes of BHs are inclined with respect to the orbital angular momentum of the binary. Measuring the spin, especially \( \chi_p \), is our most powerful tool for distinguishing between the common evolution and dynamical formation channels.

We illustrate the effect on waveforms of varying the total mass \( M \), mass ratio \( q \) and the effective inspiral spin \( \chi_{\text{eff}} \) in Fig. 3.2. LIGO and Virgo have negligible sensitivity below 20 Hz, so we use waveforms beginning roughly there. The waveforms we use are tapered, ramping up from 0 at \( f \sim 10 \) Hz to full strength at \( f \leq 20 \) Hz. In Fig. 3.2, we see that the higher the total mass and the smaller the aligned spin, the shorter the waveforms above \( f_{\text{min}} = 20 \) Hz, corresponding to larger ISCOs.

**Waveform Models**

There are a number of waveform families developed over the years both in the time domain and in the frequency domain. Waveform families are mainly classified into three groups, depending on the techniques used: the IMR phenomenological model [146, 149, 150], the Effective-one-body (EOB) approach [151–153] and the numerical relativity surrogate model [154–156]. Details of various waveform generation techniques and waveform families are beyond the scope of this thesis, and we refer readers to cited references on those topics.
THE PYCBC SEARCH PIPELINE FOR GRAVITATIONAL WAVES FROM COMPACT BINARY COALESCENCES

To identify candidate GW signals and assess their significance relative to noise, we can either filter the detector data using expected gravitational waveforms to find matches or search for generic transient signals with minimal waveform assumptions [157]. In the case of CBCs, where gravitational waveforms are well-modeled using GR, matched filtering [158] is the optimal technique to search for their GW signals buried in Gaussian-stationary (white) detector noise [159–163].

Matched filtering correlates known waveform templates with detector data to uncover the presence of templates in the data stream. However, the detection problem is complicated by the facts that: (i) we do not a priori know the parameters of incoming GW signals; (ii) detector data are by no means Gaussian or stationary [1, 164]. To account for all expected GW signals from CBCs, we construct reservoirs of waveform templates, template banks, discretely sampled from the parameter space of CBCs. To deal with non-Gaussian, non-stationary excursions, glitches, in the data stream, we discard data segments containing loud glitches detrimental to search sensitivities [165] and design a number of signal consistency tests [166, 167].

The detection pipelines for CBC searches in the LVK collaboration include PyCBC [168–175], GstLAL [176–181], MBTA [182, 183], SPIIR [184–186] and cWB [187–191]. Other than SPIIR, which operates only in low latency, others have both online and offline configurations. Of the five pipelines, cWB is a generic burst search whereas PyCBC, GstLAL, MBTA and SPIIR are template-based searches. Currently in template-based searches, we restrict waveform templates to the dominant GW harmonic of aligned spinning systems where extrinsic parameters only affect the overall amplitude and phase of the signals. This choice reduces the parameter space of the search and simplifies the maximization over extrinsic parameters, despite some sacrifice of sensitivities by neglecting higher order modes and for systems with precession, eccentricities, etc. To further improve search sensitivities, there are ongoing efforts towards including precession [192, 193] and higher order modes [194, 195].

PyCBC is a matched-filter, all-sky pipeline for GW signals from compact binary
mergers using a bank of pre-calculated gravitational waveform templates. In this chapter, we describe the entire PyCBC workflow in Section 4.1, with the author’s research efforts discussed in the R&D subsection and Section 4.2. The methods presented are common to both the archival and PyCBC Live (real-time) configurations, and we highlight the difference between the two when relevant in Section 4.2.

4.1 Matched-filter Search Pipeline

The observed data $s(t)$ in GW detectors can be modeled as a sum of detector noise $n(t)$ and GW-induced strain $h'(t)$,

$$s(t) = n(t) + h'(t).$$

(4.1)

The strain $h'(t)$ observed at detectors is a linear combination of the two GW polarizations $h_+(t)$ and $h_\times(t)$ [162],

$$h'(t) = F^+(t; \Theta)h_+(t) + F^\times(t; \Theta)h_\times(t),$$

(4.2)

where the antenna patterns $F_+(t; \Theta)$ and $F_\times(t; \Theta)$ depend on the source polarization vector $\Theta$ with respect to the detector arms. The detector response functions for $P = \{+, \times\}$ plane polarized waves are computed as

$$F^P(t; \Theta) = \frac{1}{2}[X^\mu(t)X^\nu(t) - Y^\mu(t)Y^\nu(t)]e^P_{\mu\nu}(\Theta),$$

(4.3)

where $X(t)$ and $Y(t)$ are unit vectors pointing along the interferometer arms, and $e^P_{\mu\nu}(\Theta)$ is the polarization basis tensor. Vectors $X(t)$ and $Y(t)$ in Eq. (4.3) are all time-dependent due to the Earth’s rotational motion, giving rise to periodic changes of antenna patterns. The beam pattern for an unpolarized wave is then given by

$$F(t; \Theta) = (|F^+(t; \Theta)|^2 + |F^\times(t; \Theta)|^2)^{1/2}.$$

(4.4)

Fig. 4.1 displays the antenna patterns of a ground-based Michelson GW interferometer with arms oriented 90 degrees apart for $+$ polarized waves, $\times$ polarized waves and unpolarized waves at $f_{\text{GW}} = 0$ in the small-antenna limit ($L \ll c/f_{\text{GW}}$). The current ground-based GW detector network is sensitive to $f_{\text{GW}}$ up to $\sim 4000$ Hz, safely in the small-antenna limit; but not for space-borne detectors like LISA.

In the presence of Gaussian-stationary noise only, the statistical properties of the detector noise $n(t)$ are well-characterized by the one-sided power spectral density (PSD) written as

$$P_n(f) = \lim_{T \to \infty} \frac{2}{T} \left| \int_{-T/2}^{T/2} n(t)e^{-it2\pi ft} \, dt \right|^2.$$  

(4.5)
Figure 4.1: Antenna patterns of a LIGO GW detector in the small-antenna limit \((L \ll c/f_{\text{GW}})\). The interferometer beam-splitter is located at the center of each pattern and the black lines represent the arms of the interferometer. GW sensitivity of the detector in a certain direction is measured by the distance of the plot surface to the center. At \(f_{\text{GW}} = 0\), the left panel is for + polarization, the middle panel is for \(\times \) polarization, and the right panel is for unpolarized GWs [196].

Or equivalently,

\[
\frac{1}{2} P_n(f) \delta(f - f') = \langle \hat{n}^* (f') \hat{n}(f) \rangle. \tag{4.6}
\]

Here, \(\delta(f - f')\) is the Dirac delta function, the angle bracket denotes averaging over many noise realizations, and \(\hat{n}(f)\) is the Fourier transform of the detector noise time series \(n(t)\). The Fourier transform of a time series \(\psi(t)\) is defined to be

\[
\tilde{\psi}(f) = \int_{-\infty}^{\infty} \psi(t) e^{-i2\pi ft} \, dt. \tag{4.7}
\]

In Fig. 2.4, we show the one-sided amplitude spectral density (ASD), the square root of the one-sided PSD, for Advanced LIGO in its first three observing runs and at design sensitivity.

In this straightforward case, the matched filter is optimal in that it maximizes the output signal-to-noise ratio (SNR), which suffices as a detection statistic [159]. In reality, additional to Gaussian-stationary noise from fundamental processes such as seismic noise, thermal noise and quantum noise, there exists non-Gaussian, non-stationary transient noise of environmental and instrumental origin. To deal with these artifacts, data quality vetoes are applied to periods of poor detector performance before pipeline searches and most importantly, the detection statistic has been generalized in numerous ways beyond the simple SNR in order to be more robust against data quality issues.

The offline PyCBC workflow begins when we have data-quality vetted, calibrated strain data [197, 198]. Before filtering, we compute the harmonic mean of the
PSDs of detector data, averaged over an observing duration and all detectors in the search network. The average PSD is used to generate a waveform template bank covering the CBC parameter space. The pipeline detects and removes loud noise transients that survive data-quality investigations. For each detector, the data are then matched filtered, producing a SNR time series for every template in the bank. The resulting SNR time series is thresholded and clustered in time and in template bank to generate a list of single-detector triggers. Triggers are then subject to two $\chi^2$ tests, distinguishing signals from noise based on their morphology in the time-frequency space. Single-detector triggers in the detector network passing multi-detector coincidence and coherence checks are labeled GW candidate events.

To gauge the statistical significance of candidates, a noise background is computed empirically via numerous time shifts. Using the noise background long with a signal model, the candidates are subsequently ranked by the false-alarm rate (FAR), and associated false-alarm probability (FAP) and inverse FAR (IFAR). The entire workflow is summarized in Fig. 4.2. All procedures are described in detail below.
In LIGO–Virgo’s second and third observing runs (O2 and O3), for offline analyses, we apply the PyCBC search in two regimes, PyCBC-broad and PyCBC-BBH analyses. The PyCBC-broad analysis searches for all three types of CBCs, whereas the PyCBC-BBH analysis focuses on BBH coalescences with a total mass range of $[10M_\odot, 500M_\odot]$. The PyCBC-BBH analysis is out of the scope of this thesis and we will confine our discussion to PyCBC-broad.

**Template Bank**

Since we do not know the signal parameters beforehand, we filter the detector data against a template bank, a collection of expected GW signals covering the intrinsic parameter space of interest. Currently, we only use templates of the dominant GW harmonic of CBCs with aligned spins, constraining the parameter space to four dimensions of component masses and component aligned spins.

Before filtering, we calculate a template bank to be used across detectors and over an entire observing period. The density of templates across the bank is determined by the PSD of the detectors. For each independent detector, to obtain a representative PSD averaged over time, we first measure the median noise PSD every 2048 s over an observing period and compute the harmonic mean of the PSDs by

$$P_n^{\text{harmonic}}(f) = \frac{N}{\sum_{i=1}^{N} 1/P_n^i(f)},$$

where $N$ is the number of noise PSDs in an observing period. We then compute the final harmonic mean PSD using resulting PSDs from each detector in the network via Eq. 4.8. The template bank needs to be regenerated when detector sensitivities have changed significantly. By using a single template bank across time and detectors, we have reduced the computational cost considerably and simplified signal coincidence tests in downstream stages.

In O3, the PyCBC-broad analysis uses a template bank with $\sim 430,000$ templates covering a parameter space of total masses $[2M_\odot, 500M_\odot]$ redshifted to the detector frame [199]. The minimum component mass is $1M_\odot$. For templates with a total mass below $4M_\odot$, we use the TaylorF2 [147] waveform approximant, which utilizes a frequency-domain, post-Newtonian, inspiral-only formulation [200]. For systems with a total mass above $4M_\odot$, we adopt the SEOBNRv4_ROM [148] waveform approximant, modeled using the effective-one-body (EOB) formalism [153] tuned to numerical relativity for the merger and ringdown, with computationally efficient reduced-order modeling techniques [201]. Templates are placed using an optimized
Figure 4.3: The template bank used in O3, plotted in the $m_1-m_2$ plane. Each blue dot represents a waveform template. Note that the templates are chosen in a space higher than two dimensions for aligned spinning systems, causing apparent overlap in the $m_1-m_2$ plane.

hybrid geometric–random method \cite{202, 203} with a minimum duration cutoff of 0.15 s \cite{204}. The loss in matched-filter SNR due to the discrete placements of templates is less than 3%.

In Fig. 4.3, we plot the resulting O3 template bank in the $m_1-m_2$ plane, in which each dot represents a waveform template. The boundary at mass ratio $q = 1$ has a lower density because the waveform templates are more reliable. The high-mass boundary is determined by the minimum requirement on waveform duration. The gaps and structure are artifacts caused by the waveform placement algorithm, with no compromise in the loss of SNR being less than 3%. Lower-mass systems such as BNSs and NSBHs, with more GW cycles in band, require more closely-spaced templates than higher-mass BBHs, hence an apparent lower density of templates in the upper-right corner of the plot.

**Gating**

Short-duration loud glitches can produce high SNR triggers, severely impairing the search sensitivity by incurring dead time due to trigger clustering and ringing of the matched filter \cite{205}. Specifically, when a loud glitch is present, the long-standing impulse response of the filters in the SNR computation for a given template causes the generation of several high SNR triggers over a timescale much longer than the
Figure 4.4: Example of gating. The 5-second stretch of data is taken from LIGO’s second observing run, injected with two glitches. The threshold of gating is $50\sigma$ in whitened strain. Only one glitch that surpasses this threshold is detected and gated.

glitch itself. These spurious high SNR triggers can dominate over possible GW signals and increase the noise background of the search.

We implement a gating algorithm, which windows out the GW strain data for a short duration centered on the detected glitch prior to matched filtering [171]. The strain time series is first Fourier transformed and whitened via

$$\tilde{s}_{\text{white}}(f) = \frac{\tilde{s}(f)}{\sqrt{P_n(f)/2}}.$$  \hspace{1cm} (4.9)

After the inverse Fourier transform of $\tilde{s}_{\text{white}}(f)$, loud excursions of the whitened strain time series are identified as glitches. We then multiply the data with the complement of a symmetric Tukey window centered on the detected glitch time. The window is configured to have central zeros and symmetric smooth taper on both sides. The gating procedure applied to two example injected loud glitches is shown in Fig. 4.4.

Currently we fix a window duration for the algorithm, longer than the glitch itself but shorter than 1 s [164]. Going forward, a fast way to estimate the duration of a glitch is beneficial to be included in the gating algorithm. Another method to deal with glitches is by “inpainting” the identified bad data segments such that the twice-whitened data are zero at bad times [206].

**Matched Filtering**

Since the gravitational waveforms of CBCs are well modeled, the pipeline employs matched filtering [158] to search for these signals embedded in detector noise. A
matched filter correlates a known waveform template with the data and outputs the SNR,
\[
\rho(t) \equiv \langle s(t), \hat{h}(t) \rangle = \int_{-\infty}^{\infty} \frac{\tilde{s}(f)\tilde{h}^*(f) + \tilde{s}^*(f)\tilde{h}(f)}{P_n(f)} e^{i2\pi ft} df
\]
\[
= 4\text{Re} \int_0^{\infty} \frac{\tilde{s}(f)\tilde{h}^*(f)}{P_n(f)} e^{i2\pi ft} df,
\]
where the angle bracket \(\langle \cdot, \cdot \rangle\) denotes the inner product between the data \(s(t)\) and the normalized waveform template \(\hat{h}(t)\) [207]. In Eq. (4.10), \(\tilde{s}(f)\) and \(\hat{h}(f)\) are the Fourier transforms of the strain time series \(s(t)\) and the normalized template time series \(\hat{h}(t) = h(t)/\sqrt{\langle h(t), h(t) \rangle}\), respectively.

The real strain signal \(h'(t)\) in Eq. (4.1) is determined by a number of intrinsic and extrinsic parameters (see Table 3.1) we do not know \textit{a priori}. To simplify the search, currently we only consider the dominant GW mode of binaries with aligned spins. For such systems, extrinsic parameters such as sky location and orientation have effects only on the overall amplitude and phase of the waveform. To maximize over the coalescence phase \(\phi_c\), we decompose the waveform template \(\hat{h}(t)\) into two orthogonal phases \(h_{\cos}(t)\) and \(h_{\sin}(t)\) and construct the squared SNR by [168]
\[
\rho^2(t) = \rho_{\cos}^2 + \rho_{\sin}^2 = \frac{\langle s, h_{\cos} \rangle^2}{\langle h_{\cos}, h_{\cos} \rangle} + \frac{\langle s, h_{\sin} \rangle^2}{\langle h_{\sin}, h_{\sin} \rangle} = \frac{\langle s, h_{\cos} \rangle^2 + \langle s, h_{\sin} \rangle^2}{\langle h_{\cos}, h_{\cos} \rangle}.
\]

The two orthogonal phases of a template are connected by \(\tilde{h}_{\sin}(f) = i\tilde{h}_{\cos}(f)\). To maximize over the coalescence time \(t_c\), the pipeline records a list of times when the SNR exceeds a chosen threshold value for each template. The amplitude of the template after maximizing over \(\phi_c\) and \(t_c\) gives a measure of the effective distance of the source [168], a function of all the other extrinsic parameters,
\[
D_{\text{eff}} = D \left[ F_+^2 \left( \frac{1 + \cos^2 i}{2} \right)^2 + F_\times^2 \cos^2 i \right]^{-1/2}.
\]

We maximize over intrinsic parameters by computing the matched-filter SNR using Eq. (4.11) and thresholding for each template in the template bank in each detector. A time-clustering algorithm is then applied, keeping local maxima within a preset window of the SNR time series. The final product of this stage is a list of \textit{triggers} for each detector when a GW signal may exist.
Signal Consistency Tests

Detector data contain non-Gaussianity and non-stationarity, hence we need additional methods to distinguish between signals and noise for single-detector triggers, such as the two chi-squared signal consistency tests in PyCBC. Both of the $\chi^2$ tests are purposed to determine whether there exists a statistically significant difference between what is expected in the matching template and what is measured in the detector data.

The first $\chi^2$ discriminator compares the time-frequency distribution of power in the data with that of the matching template from the earlier matched filtering stage [169]. Based on the parameters of the template, the algorithm decides adaptively on the number of frequency bins to split the template. Lower-mass systems stay longer (more GW cycles) in the frequency band of ground-based GW interferometers, and thus we are able to choose more bins for a more effective test. After selecting the number of bins $p$, the frequency bins are chosen such that each bin contributes the same amount of power to the total SNR for a perfectly matching signal. For each bin, we then calculate the matched-filter SNR $\rho_i$ for the data. If the data contains a real signal, we expect $\rho_i$ to be $1/p$ of the total SNR. A $\chi^2$ statistic is constructed to measure the residual power in each frequency bin by [169]

$$
\chi^2 = p \sum_{i=1}^{p} \left[ \left( \frac{\rho_{\cos,i}^2}{p} - \rho_{\cos}^2 \right)^2 + \left( \frac{\rho_{\sin,i}^2}{p} - \rho_{\sin}^2 \right)^2 \right],
$$

where $\rho_{\cos}^2$ and $\rho_{\sin}^2$ are the squared SNRs of the two orthogonal phases of the matched filter in Eq. (4.11).

For a trigger of certain matched-filter SNR, when the reduced chi-squared $\chi^2_r = \chi^2 / (2p - 2)$ is large, the trigger is more likely to be noise; when $\chi^2_r$ is close to unity, the trigger is more likely to be a signal. We suppress noise triggers by re-weighting the matched-filter SNR in Eq. (4.11) via [208]

$$
\hat{\rho} = \begin{cases} 
\rho / \left[ (1 + (\chi^2_r)^3) / 2 \right]^{1/6}, & \text{if } \chi^2_r > 1, \\
\rho, & \text{if } \chi^2_r \leq 1.
\end{cases}
$$

We also include a high-frequency sine-Gaussian $\chi^2$ discriminator for short-duration transient noise, “blip glitches”, in the ranking of single-detector triggers. Most compact binary mergers do not have significant signal power at frequencies higher than the ringdown of the dominant harmonic. Exploiting this fact, we place $N$ sine-Gaussian tiles and measure the excess power at the time of peak signal amplitude
and at various frequencies above the ringdown. A sine-Gaussian reduced $\chi^2$ statistic is constructed by computing the sum of the squared SNRs in each individual sine-Gaussian tile [173]

$$\chi^2_{r,sg} = \frac{1}{2N} \sum_i^N \langle \tilde{s}(f), \tilde{g}_i(f; f_0, t_0, Q) \rangle^2,$$

where in the time domain each sine-Gaussian tile $g_i$ is defined as

$$g(t; f_0, t_0, \phi_0, Q) = A \exp \left( -\frac{4\pi f_0^2 (t - t_0)^2}{Q^2} \right) \cos(2\pi f_0 t + \phi_0).$$

Here, $f_0$ and $t_0$ are the central frequency and time of the sine-Gaussian, respectively, $Q$ is the quality factor set to 20 currently, $A$ is the amplitude set to 1, and $\phi_0$ is the phase. The sine-Gaussian $\chi^2$ discriminator is most effective at preventing glitches from triggering high-mass templates with final frequencies of $\sim$100 Hz or less, increasing the search sensitivity to high-mass BBHs.

Finally, we combine $\chi^2_{r,sg}$ with the reweighted SNR $\hat{\rho}$ defined in Eq. (4.14) to obtain the final single-detector reweighted SNR $\hat{\rho}_{sg}$ by [173]

$$\hat{\rho}_{sg} = \begin{cases} \hat{\rho}(\chi^2_{r,sg}/4)^{-1/2}, & \text{if } \chi^2_{r,sg} > 4, \\ \hat{\rho}, & \text{if } \chi^2_{r,sg} \leq 4. \end{cases}$$

**Multi-detector Coincidence**

GW detection pipelines such as PyCBC greatly benefit from requiring coincidences from multiple detectors, rejecting the vast majority of uncorrelated noise triggers [171]. Triggers of astrophysical origin in different detectors fall within a short time window of each other, given by time-of-flight between detectors and timing measurement uncertainty. We also require coincident triggers to be seen with the same waveform template. In the analysis, coincidences for all combinations of detectors in the network are constructed as a first step. Multi-detector coincidences are formed if the two-detector coincidence test is passed for each pair of involved detectors [174]. Coincident triggers that survive the time and parameter coincidence tests are labeled *candidate events*.

To compare GW candidates in different detector combinations systematically, a ranking statistic that reflects the degree of belief for a candidate’s astrophysical origin and signal consistency was developed in [174]. The Neyman-Pearson optimal detection statistic for triggered searches forms a likelihood ratio of the signal and
noise event rate densities \[209],

\[
\Lambda(\tilde{k}) = \frac{\hat{r}_S(\tilde{k})}{r_N(\tilde{k})}.
\]

(4.18)

In Eq. (4.18), \(r_N(\tilde{k})\) defines the noise event rate density and \(r_S = \mu_S \hat{r}_S(\tilde{k})\) is the signal event rate density, where \(\mu_S\) is the binary coalescence rate per co-moving volume per time and \(\hat{r}_S(\tilde{k})\) is the signal recovery rate depending on detector orientation and sensitivities. The binary coalescence rate \(\mu_S\) is assumed to be constant and does not affect the likelihood ratio \(\Lambda(\tilde{k})\). This assumption remains valid for ground-based detectors with their detection horizon within a redshift of \(z \leq 1\). Note that out to higher redshifts, this rate can be varying such that the constancy assumption breaks down for GW detectors like LISA.

Parameters \(\tilde{k}\) are used to collectively describe a GW candidate of unknown origin, which include the single-detector reweighted SNR \(\hat{\rho}_{sg:a} = \hat{\rho}_a\) (to simplify the notation) and template sensitivity \(\sigma_a\) for each participating detector \(a\); the intrinsic parameters \(\tilde{\theta}\) of the matching template; the amplitude ratio \(A_{ab}\), time difference \(\delta t_{ab}\) and phase difference \(\delta \phi_{ab}\) between each distinct detector pair \(a-b\) in the network,

\[
\tilde{k} = \{[\hat{\rho}_a, \sigma_a], \tilde{\theta}, [A_{ab}, \delta t_{ab}, \delta \phi_{ab}]\}.
\]

(4.19)

**Noise Model** – The coincident noise rate density is estimated by fitting the distribution of triggers above a reweighted SNR threshold in the individual detectors and combining the trigger rates for different detector combinations.

For detector \(a\), for template \(i\) with parameters \(\tilde{\theta}\), we model the distribution of triggers with respect to \(\hat{\rho}\) above a threshold \(\hat{\rho}_{th}\) as a decreasing exponential,

\[
r_{N:a;i}(\hat{\rho}; \tilde{\theta}) = \mu(\tilde{\theta}) p(\hat{\rho} | \tilde{\theta}, N),
\]

(4.20)

where

\[
p(\hat{\rho} | \tilde{\theta}, N) = \begin{cases} 
\alpha(\tilde{\theta}) \exp[-\alpha(\tilde{\theta})(\hat{\rho} - \hat{\rho}_{th})], & \hat{\rho} > \hat{\rho}_{th}, \\
0, & \hat{\rho} \leq \hat{\rho}_{th}.
\end{cases}
\]

(4.21)

Here, \(\mu(\tilde{\theta})\) is the number of triggers for the template above threshold and \(\alpha(\tilde{\theta})\) is the exponential decay rate. Before maximum-likelihood (ML) fitting, a fixed, small number of the loudest triggers from each detector are removed to mitigate possible contamination due to signals. The ML estimator for the exponential decay rate is found to be

\[
\hat{\alpha}_{ML} = (\hat{\rho} - \hat{\rho}_{th})^{-1},
\]

(4.22)
where $\tilde{\rho}$ is the mean value of $\hat{\rho}$ and the variance of the fit is the inverse of the number of triggers above threshold for the template.

The rate of coincident noise events is found by multiplying single-detector noise trigger rates $r_{N:a;\hat{\theta}}(\hat{\rho}; \tilde{\theta})$ for each detector $a$ and the size of the allowed time window for coincidences $A_N$,

$$r_{N;\hat{\theta}}(\hat{\rho}; \tilde{\theta}) = A_N \prod_a r_{N;a;\hat{\theta}}(\hat{\rho}; \tilde{\theta}). \quad (4.23)$$

**Signal Model** – The probability densities over the extrinsic parameter space $\tilde{\Omega} = \{A_{ab}, \delta t_{ab}, \delta \phi_{ab}\}$ for signal and noise coincidences are expected to be quite different. Noise coincidences should be uniformly distributed over the extrinsic parameter space, while the probability densities for signal coincidences for different detector combinations were estimated using Monte Carlo simulations given the detector locations in [172, 174].

The template sensitivity $\sigma_t$ acts as a proxy to the luminosity distance of a detector at which a standard CBC is recovered with a given expected SNR. For a network, the network sensitivity is determined by the least sensitive detector with $\sigma_{\text{min},i}$. Assuming a homogeneous distribution of sources in the nearby Universe, the expected signal recovery rate for a detector combination is proportional to $\sigma_{\text{min},i}^3$, which is further normalized by the median network sensitivity for HL coincidences $\tilde{\sigma}_{HL,i}^3$.

Combining the above considerations, the time-dependent signal event rate density for a given detector combination is proportional to

$$r_{S;i} \propto \frac{p(\tilde{\Omega}|S)}{p(\tilde{\Omega}|N)} \frac{\sigma_{\text{min},i}^3}{\tilde{\sigma}_{HL,i}^3}. \quad (4.24)$$

**Final Ranking Statistic** – Eq. (4.18) ensures the ranking statistic is comparable across different detector combinations. In practice, we use the logarithm of the ratio of signal and noise rate densities as the ranking statistic. The final ranking statistic is

$$R = \log \tilde{r}_S(\tilde{\kappa}) - \log r_N(\tilde{\kappa})$$

$$= -\log A_{N;\{d\}} - \sum_d \log r_{N;d;\hat{\theta}} + \log p(\tilde{\Omega}|S)$$

$$- \log p(\tilde{\Omega}|N) + 3(\log \sigma_{\text{min},i} - \log \tilde{\sigma}_{HL,i}). \quad (4.25)$$

**Candidate Significance**

Since we cannot shield the detectors from GW sources and no theoretical model of the detector noise exists, we empirically estimate the background distribution
of noise coincidences [167]. A collection of noise coincidences is fabricated by combining time-shifted single-detector triggers. In a detector network, only one detector is selected for time-shifting while all others are fixed, with the least sensitive detector kept in the fixed sub-network [174]. Before the time-shifted analyses, coincident triggers from confident signals with FAR below a specified threshold in any combination are removed [210]. For a given detector combination, we only use triggers from times when all participating detectors are observing. The time shift is chosen to be larger than the GW travel time between the farthest sites so that the coincidence is not physically permissible. To test the null hypothesis, we compute the ranking statistic for all time-shifted noise coincidences and obtain a distribution of the ranking statistic for noise.

Then for a zero-lag GW candidate event, we estimate its significance using the FAR statistic, the expected rate of coincident noise events with a higher ranking statistic. We count the number of noise coincidences ranked higher than the candidate event and divide the count by the effective length of time for constructing the background distribution. For week-long worth of data, FARs constructed as described can be as low as one per tens of millennia [174]. Formally, the FAR as a function of the ranking statistic threshold \( \Lambda^* \) is

\[
\text{FAR}(\Lambda^*) = \int r_N(\tilde{\kappa}) \Theta(\Lambda(\tilde{\kappa}) - \Lambda^*) \, d^n\tilde{\kappa}.
\]

We note though the most sensitive searches for GWs from CBCs rely on coincidence tests, there are certainly single-detector events of astrophysical origin. A new method for extrapolating the significance of single-detector candidates beyond the live-time of the analysis is described in [211].

**Research and Development**

This subsection contains unpublished development efforts towards improving the multi-detector coincidence checks for PyCBC. We revisit the requirement of coincident triggers being seen with the same waveform template. When a strand of data is matched filtered, it fires up more than one templates in the bank. Often when a trigger for this strand passes the threshold, a number of triggers also pass the threshold and go into the stage of multi-detector coincidence tests. In the current pipeline, we require only one template overlap among the triggers for different detectors. However, the geographies of triggered templates in the bank for GW signals and for noise are quite different. Real signals trigger templates roughly in the same
region of parameter space, while for noise the distribution is random all across the template bank.

To further distinguish signals from noise, we exploit the difference between their triggered template distributions. We show our work in the LIGO Hanford-LIGO Livingston (HL) two-detector case, noting that the method can be easily extended to a multi-detector GW network with possibly better separating power. Without using the entire distribution of triggered templates, which can be computationally expensive, we construct a two-detector template coherence statistic by

$$\alpha = \frac{(\hat{\rho}_{H}^2 + \hat{\rho}_{L}^2)^{1/2}}{(\hat{\rho}_{\text{max};H}^2 + \hat{\rho}_{\text{max};L}^2)^{1/2}}.$$  \hspace{1cm} (4.27)

Here, $\hat{\rho}_{H,L}$ are the single-detector reweighted SNRs for the matching template and $\hat{\rho}_{\text{max};H,L}$ are the maximum reweighted SNRs of all triggers for each detector. For GW signals, $\hat{\rho}_{H,L}$ and $\hat{\rho}_{\text{max};H,L}$ are close in values; for noise, they can be quite different, as shown in Fig. 4.5 for $\hat{\rho}_{H}$ and $\hat{\rho}_{\text{max};H}$ for BNS injections, BBH injections and background events from the first week of O3 data.

We test the differentiating power of our method using the first week of O3 data from LIGO Hanford and LIGO Livingston. The simplest way to utilize the template coherence statistic in the pipeline is by thresholding on it: if $\alpha \geq \alpha_{\text{th}}$, we do not change the ranking statistic of coincident triggers; if $\alpha < \alpha_{\text{th}}$, $\alpha$ can act as a weight and be multiplied into the ranking statistic directly. For a threshold of $\alpha_{\text{th}} = 0.9$ in Fig. 4.6, we weight down $\sim 15\%$ background events while only $\sim 1\%$ injected signals already with low IFARs are also weighted down. More complicated ways to include the template coherence statistic may achieve better improvements. Going forward, it
Figure 4.6: Two-detector template coherence statistic $\alpha$ for BNS injections, BBH injections and noise. With a threshold applied, we can weight down noise events using $\alpha$ with barely any compromise of sensitivity to GW signals. The improvement is especially significant for high-mass BBH events.

is beneficial to further develop, test and incorporate the template coherence statistic into the multi-detector detection statistic of PyCBC.

4.2 PyCBC Live: A Real-time Search

PyCBC Live [212, 213] is the real-time version of the PyCBC pipeline [171, 174, 175] to search for GWs from CBCs based on frequency-domain matched filtering. The purpose of GW low-latency searches is to generate rapid alerts from the data acquired by LIGO–Virgo in order to allow follow-up observation campaigns with EM and neutrino observatories around the globe [214]. Rapid identification and localization of GW candidates by search pipelines is uploaded to GraceDB [215] and announced on the Gamma-ray Coordinates Network (GCN) [216, 217] Notices and Circulars for follow-ups. Later re-analysis of the strain data can lead to some candidates being retracted and additional events being discovered, benefitting from improvements in calibration, data quality and data analysis from the entire observing run.

PyCBC Live was first introduced in [212] for LIGO’s O2. During O3, with new pipeline developments and the addition of Virgo, there are a number of new methods employed, which we summarize below and refer readers to [213] for details.

(a) Before O3, loud instrumental transients were not dealt with in PyCBC Live when streaming the data-quality investigated data. In early O3, this resulted in occasional gaps in the production of triggers from a given detector, lasting from a few seconds to several tens of seconds, depending on the glitch. During O3, the gating algorithm described earlier in PyCBC’s offline version was adopted in PyCBC Live. We used a threshold of $50\sigma$ on the absolute value of the whitened strain time series as a glitch.
detector. Each detected glitch was gated with a symmetric complemented Tukey window configured to have 0.5 s of central zeros and 0.25 s of smooth taper on both sides. This approach significantly reduced the impact of high SNR non-Gaussian transient noise with no visible impact on the latency of the analysis.

(b) The sine-Gaussian $\chi^2$ discriminator [173] described earlier was also implemented in PyCBC Live during O3.

(c) PyCBC Live in its O3 configuration included additional detectors for coincident searches. For each detector pair in the GW network, we independently perform the double-coincident analysis and a FAR is computed locally using the last 5 hr of data, tracking possible detector performance changes. If multiple coincidences are formed, the one with the lowest FAR is chosen and if tied, the one with the highest ranking statistic is chosen. For the remaining follow-up detectors, SNR time series and associated p-values are obtained using the template of the selected candidate.

(d) For single-detector candidates, extrapolation of the FAR beyond once per live time is used in order to generate low-latency alerts [211].

(e) In the later part of O3, source classification between the different possible types of CBCs is improved by using the chirp mass recovered by the pipeline as input and implicitly assuming a uniform density of candidate signals over the plane of component masses. This method constrains the chirp mass to be within a certain region in the plane of component masses. The probabilities of a candidate belonging to each source category are then proportional to the associated area of region.

(f) To produce rapid localization by BAYESTAR [218, 219], PyCBC Live needs to provide the template that maximizes the likelihood assuming Gaussian-stationary noise. However, in the pipeline run, the final selected template does not necessarily maximize the network matched-filter SNR under Gaussian-stationary noise assumption. Thus, a follow-up process to reanalyze a short stretch of data around the candidate and find the optimal template for Gaussian-stationary noise is launched after a candidate is reported to GraceDB. A new candidate with the maximized network matched-filter SNR is uploaded for localization after the process converges or a time-out of 400 s is reached.
Chapter 5

RESULTS OF THE PYCBC SEARCH ACROSS LIGO–VIRGO’S FIRST THREE OBSERVING RUN

We report results of the PyCBC searches as a part of the Gravitational-wave Transient Catalogs (GWTCs) published by the LVK collaboration to date, with discussions on some notable events. GWTC-1 [220] covers the first observing run (O1) of Advanced LIGO between September 12, 2015 and January 19, 2016 and its second observing run (O2) between November 30, 2016 and August 25, 2017, with Advanced Virgo joining on August 1, 2017. GWTC-1 contains 11 confident detections, with 10 BBHs [2, 221–226] and 1 BNS [4, 5, 227], and 14 marginal triggers.

GWTC-2 reports on 39 compact binary coalescences passing a FAR threshold of 2 per year [228] in the first part (O3a) of LIGO–Virgo’s third observing run (O3) between April 1, 2019 and October 1, 2019. GWTC-2.1 presents a deeper list of candidate events over O3a by including ones that have an astrophysical probability greater than 0.5 \( (p_{\text{astro}} > 0.5) \) [229], updating the total CBC count to 55. Shortly after a commissioning break between October 1, 2019 and November 1, 2019, LIGO and Virgo started the second half (O3b) of O3 between November 1, 2019 and March 27, 2020, with KAGRA also joining the global GW network. GWTC-3 identifies 35 candidates with \( p_{\text{astro}} > 0.5 \) consistent with GW signals from CBCs over O3b [6]. Across the first three observing runs, the LVK collaboration has observed 90 compact binary mergers of black holes and neutron stars, expanding the horizon of fundamental physics [25], astrophysics [26] and cosmology [29].

In addition to candidates reported by the LVK, other searches in strain data released to the public through the Gravitational Wave Open Science Center (GWOSC) [230–235] have identified a few more GW candidates [236–243].

5.1 Gravitational-wave Transient Catalogs

GWTC-1

GWTC-1 [220] is the first catalog detailing the GW signals detected during O1 and O2 by the LIGO Scientific and Virgo collaboration. Two matched-filter pipelines, PyCBC and GstLAL, and one minimally modeled cWB pipeline were employed for GWTC-1. PyCBC and cWB did not include Virgo data for analysis while GstLAL
analyzed Virgo data for August 2017. Note the PyCBC ranking statistic of GW candidates for the LIGO two-detector network was simpler at the time of GWTC-1 than described in Chapter 4. During O1, three BBHs were detected [223] while O2 saw the first detection of a BNS [4, 5, 227] and an additional seven detections of BBHs [224–226]. In addition, a list of 14 marginal event candidates with a FAR less than 1 per 30 days is provided in GWTC-1. Sensitivity of a GW interferometer is conventionally measured in terms of the inspiral range of a $1.4M_\odot + 1.4M_\odot$ BNS, which is the distance such a system can be detected with a SNR of 8. During O1 and O2, the BNS inspiral ranges span from 60 Mpc to 80 Mpc for LIGO Hanford and from 60 Mpc to 100 Mpc for LIGO Livingston. During the one-month of Virgo’s up-time, its BNS inspiral range was roughly at 25 Mpc.

**GWTC-2**

Thanks to several upgrades during the commissioning period prior to the start of O3a, the three detectors significantly improved their sensitivities, with median BNS inspiral ranges achieving 108 Mpc for LIGO Hanford, 135 Mpc for LIGO Livingston and 45 Mpc for Virgo [228]. The sensitivity increase drastically boosted the number of detections to 39 over O3a, compared to 11 across O1 and O2 [220]. Of the 39 new GW candidates reported by GWTC-2 over O3a, 26 were previously identified by low-latency searches and announced in near real-time through GCN Notices and Circulars while 13 were found by more in-depth offline reanalyses. GWTC-2 imposes a FAR threshold of 2 per year in each of the three offline search pipelines (cWB$^1$, GstLAL, PyCBC), with an expected noise contamination percentage of less than 10%. At this FAR threshold, the hyper volume surveyed by PyCBC is 0.296 Gpc$^3$ yr [228]. In GWTC-2, PyCBC did not scan the Virgo data because the three-detector searches were not integrated into the pipeline. Among the 39 reported candidates, 27 were detected by PyCBC, fewer than the number of 36 for GstLAL primarily due to the non-inclusion of Virgo data. When accounting for the difference in analyzed data, both GstLAL and PyCBC have detected a comparable amount of GW signals.

**GWTC-2.1**

In GWTC-2.1, we use the final version of the O3a strain data with improved calibration [23] and non-linear noise subtraction around 60 Hz [244, 245]. In addition to PyCBC and GstLAL, MBTA in its offline configuration was employed for the first

$^1$cWB did not detect any events that were not also seen by the two CBC pipelines; no unmodeled burst events have been confidently detected in O1 through O3.
time in GWTC-2.1. In this analysis, PyCBC was extended to search data from the three-detector LIGO–Virgo network, along with updates to the event ranking statistic [174] and the \( p_{\text{astro}} \) [246] calculation and a new method to estimate source class probability [213]. GWTC-2.1 imposes a looser FAR threshold of 2 per day, resulting in a deeper list of 1201 candidate events [229]. Astronomical investigations of subthreshold GW candidates may lead to multi-messenger discoveries, enhancing our understanding of such systems. Added with the requirement of having a probability of astrophysical origin greater than 0.5 (\( p_{\text{astro}} > 0.5 \)), a subset of 44 high-significance candidates were identified, of which 36 have been reported in GWTC-2.

**GWTC-3**

GWTC-3 is the most comprehensive catalog of GW events up to date, recording all 90 signals with \( p_{\text{astro}} > 0.5 \) found by the LVK collaboration across LIGO–Virgo’s first three observing runs [6]. Out of the 35 new significant events discovered during O3b, 18 were previously broadcast in low-latency through GCN Notices and Circulars and 17 were presented for the first time. The noise contamination rate is expected to be \( \sim 10–15\% \). Additionally, there are 1048 subthreshold candidates with a FAR less than 2 per day but that do not surpass the \( p_{\text{astro}} \) threshold. Significant candidates from O3b are comprised of BBHs and NSBHs, with none from BNSs.

**Figure 5.1**: Masses of BHs and NSs in the stellar graveyard detected through GWs and EM waves. This plot contains all GW events detected through GWTC-3 with \( p_{\text{astro}} > 0.5 \) [247]. Credit: LVK / Aaron Geller / Northwestern University.
Figure 5.2: The number of confidently identified CBC candidates with a probability of astrophysical origin $p_{\text{astro}} > 0.5$ versus the detector network’s effective surveyed BNS time-volume [228]. The colored bands mark the different observing runs. The solid black line indicates the cumulative number of probable candidates. The blue line, dark blue band and light blue band are the median, 50% confidence interval and 90% confidence interval for a Poisson distribution fit to the number of candidates at the end of O3b [6].

Masses of BHs and NSs detected through GWs and EM waves are plotted in Fig. 5.1, containing all GW events in GWTC-3 with $p_{\text{astro}} > 0.5$. Compared to O3a, the median BNS inspiral ranges are 115 Mpc for LIGO Hanford, 133 Mpc for LIGO Livingston and 51 Mpc for Virgo. Fig. 5.2 shows the increase in the number of candidates across observing runs. The PyCBC pipeline employed for GWTC-3 has been detailed in Chapter 4, with a new introduction of the use of graphics processing unit (GPU) cores or distributed computing through the Open Science Grid (OSG) [248, 249] for faster computation. Among the 35 events, 29 were found by one or both of the PyCBC-broad and PyCBC-BBH analyses. 21 events were found by two or more offline search pipelines (PyCBC, GstLAL, MBTA, cWB).

5.2 Exceptional Gravitational-wave Events

GW150914: First Direct Detection of Gravitational Waves

GW150914, observed on September 14, 2015, was the first direct detection of GWs and the first direct detection of a BBH merger to form a single BH [2]. The GW
signal was first identified by a generic transient search [250] and later recovered by PyCBC and GstLAL with a combined SNR of 24 and a FAR less than 1 per 203000 years\(^2\), equivalent to a greater than 5.1\(\sigma\) significance [221]. The reconstructed waveform matches the predictions of GR for a GW coming from the merger of BHs of \(36^{+5}_{-4} M_\odot\) and \(29^{+4}_{-4} M_\odot\). The final BH mass is \(62^{+4}_{-4} M_\odot\), with \(3^{+0.5}_{-0.5} M_\odot c^2\) radiated away in GWs. The masses of pre-merger BHs for GW150914 are more massive than previously expected from X-ray binary observations, implying relatively weak stellar winds from their progenitor stars and hence low metallicity [142]. This event marks the inauguration of GW astronomy.

**GW170814: First Three-Detector Observation**

GW170814 was the first signal observed by the two Advanced LIGO detectors and the Advanced Virgo detector coherently on August 14, 2017. The combined SNR for this event is 18.3, with individual detector SNRs of 9.7, 14.8 and 4.8 for LIGO Hanford, LIGO Livingston and Virgo, respectively. PyCBC estimated its FAR to be 1 per 27000 years and GstLAL estimated it to be 1 per 140000 years. The inclusion of Virgo in the GW detector network greatly improves its localization, containing the position to a 90\% credible area of 60 deg\(^2\). GW170814 places strong constraints on the polarization states of a GW being purely tensor over purely vector or purely scalar [226].

**GW170817: First Binary Neutron Star Merger**

GW170817 was the first BNS inspiral observed by the LIGO and Virgo detectors on August 17, 2017 [4]. The signal was initially detected by the low-latency GstLAL pipeline as a single-detector event with the LIGO Hanford detector. After suppressing the glitch at the LIGO Livingston detector around the coalescence time, the searches determined the SNRs to be 18.8, 26.4 and 2.0 for LIGO Hanford, LIGO Livingston and Virgo, respectively, with a network SNR of 32.4. PyCBC placed the FAR of this event at less than 1 per \(8.0 \times 10^4\) years, while GstLAL estimated its FAR to be 1 per \(1.1 \times 10^6\) years. The source was localized to a 90\% credible region of the sky 28 deg\(^2\) in area and had a luminosity distance of \(40^{+8}_{-14}\) Mpc. Independently, a short gamma ray burst (GRB 170817A) [251, 252] was reported by Fermi-GBM [253] and INTEGRAL [254] \(\sim 1.7\) s after the coalescence time. An extensive observing campaign across the entire EM spectrum was soon launched [5]. A bright optical

\(^2\)Since this event was much more significant than any other background event in the data analyzed, the FAR was provided as an upper bound.
counterpart AT 2017gfo [255] first discovered by the One-Meter, Two-Hemisphere (1M2H) team using the 1 m Swope Telescope at Las Campanas Observatory in Chile and later by multiple teams placed the binary in the shell elliptical galaxy NGC 4993. Signals from X-ray and radio emission were observed in the following tens of days. This is the first time GW and EM signals from a single source have been observed. The multi-messenger observations of GW170817 established that GW170817 was produced by a pair of coalescing NSs in NGC 4993, followed by a short GRB and a kilonova powered by the radioactive decay of r-process nuclei synthesized in the ejecta [5].

**GW190412: First Unequal Mass-ratio Coalescence**

The offline PyCBC search detected GW190412, observed on April 12, 2019 by all three detectors, with a network SNR of 17.4 and a FAR less than $1.1 \times 10^{-4}$. GW190412 is unique in the sense that it was the first merger detected with unquestionable significant unequal masses [256]. All preceding detections have mass ratios consistent with unity, while GW190412 has a mass ratio of $q = 0.28^{+0.12}_{-0.07}$ (90% credible interval). The component masses in the source frame are $m_1 = 30.1^{+4.6}_{-5.3} \, M_\odot$ and $m_2 = 8.3^{+1.6}_{-0.9} \, M_\odot$, consistent with the BH population inferred from the first two LIGO and Virgo observing runs [220]. However, unlike in GWTC-1, $q > 0.5$ is excluded with 99% probability.

GW signals can be decomposed into multiple moments of spherical harmonics in the source frame, of which the leading emission is quadrupolar [257–259]. In the basis of the $-2$ spin-weighted spherical harmonics $Y_{lm}(-2, \theta, \phi)$, the multipole decomposition of the strain is

$$h = \sum_{l \geq 2} \sum_{-l \leq m \leq l} \frac{h_{lm}(t, \vec{\theta})}{D_L} \, Y_{lm}(-2, \theta, \phi),$$

where $(\theta, \phi)$ are the polar and azimuthal angles specifying the direction of propagation from the source to the observer, $D_L$ is the luminosity distance from the observer, and $\vec{\theta}$ define the source properties such as masses and spins. Binary systems with more asymmetric masses are predicted to have stronger contributions to GW signals from higher order modes. GW190412 is the first signal carrying clearly detectable imprints of GW higher multipoles. For GW190412, the $h_{33}$ subdominant harmonic is found to be the strongest. Including higher order modes better constrains the inference on source inclination and luminosity distance. Though no conclusive formation channel has been identified for GW190412.
GW190425: Second Binary Neutron Star Merger

GW190425 was identified as a single-detector event in the LIGO Livingston detector on April 25, 2019 first by the low-latency GstLAL pipeline with a matched-filter SNR of 12.9 [260]. LIGO Hanford was offline at the time of the event and Virgo was not sensitive enough to generate a high-SNR trigger. PyCBC Live, MBTAonline and SPIIR later analyzed the strain data around GW190425, all having consistent match-filter SNRs, with 12.1, 12.9 and 12.0 for each pipeline, respectively. The well-constrained chirp mass was also measured to be consistent across all pipelines. Estimating the significance of single-detector events is challenging. Thus, in Fig. 5.3, we present the histogram made by the author for [260] of reweighted SNRs for background events from O1, O2 and the first 50 days of O3 in the BNS region of the parameter space from the PyCBC search results, compared with GW190425. Triggers from GW170817 are also shown for reference. GW190425 is louder than all background events.

GW190521: First Detection of Intermediate-mass Black Holes

GW190521 was observed to be a short-duration transient signal by the LIGO and Virgo detectors on May 21, 2019 [7]. PyCBC Live first reported this event with a network SNR of 14.5 and a FAR of 1 per 8 years. The unmodeled cWB search in its
IMBH configuration also identified GW190521, with a network SNR of 15.0 and a FAR of 1 per 28 years. GstLAL and SPIIR found consistent candidates with lower significance. The offline cWB analysis placed this event at a network SNR of 14.7 and a FAR of 1 per 4900 years. If GW190521 is consistent with a CBC, its component BHs have masses of $85^{+21}_{-14}M_\odot$ and $66^{+17}_{-18}M_\odot$, the largest progenitor masses observed to date. The primary BH falls within the pulsational pair-instability mass gap [8]. The mass of the remnant BH is measured to be $142^{+28}_{-16}M_\odot$, the first clear detection of an IMBH [9]. The radiated energy of $> 8M_\odot$ makes it the most energetic GW event ever observed.

**GW190814: Most Unequal Mass-ratio Event To Date**

GW190814 was observed on August 14, 2019 as a three-detector event, with SNRs of 21.6, 10.6 and 4.3 for LIGO Livingston, LIGO Hanford and Virgo, respectively, and a FAR of 1 per $1.3 \times 10^3$ years, as measured by the GstLAL pipeline [261]. PyCBC detected GW190521 with consistent SNRs and a FAR of 1 per 8.1 years. A compact binary merger with the most significantly unequal masses to date, its mass ratio was estimated to be $q = 0.112^{+0.008}_{-0.009}$. The primary component is conclusively a BH with mass $m_1 = 23.2^{+1.1}_{-1.0}M_\odot$, whereas the nature of the secondary with mass $m_2 = 2.59^{+0.08}_{-0.09}M_\odot$ is unclear. No evidence of tidal effects or spin-induced quadruple effects was found in the signal, and no EM counterpart has been identified. Current knowledge of the NS equation of state does not give a definitive conclusion on the nature of the secondary component. The secondary is either the lightest BH or the heaviest NS in a compact binary system. Just like GW190412, GW190814 also exhibits significant power in the $h_{33}$ subdominant harmonic.

**GW200105 and GW200115: First Neutron Star–Black Hole Mergers**

GW200105 and GW200115 were two NSBH mergers observed on January 5, 2020 and January 15, 2020, respectively, the first observational evidence for such systems [10]. For the single-detector event GW200105 in LIGO Livingston, PyCBC detected it with a SNR of 13.1$^3$, while GstLAL found it with a SNR of 13.9 and a FAR of 1 per 2.8 years. For GW200115, PyCBC detected it with a SNR of 10.8 and a FAR of 1 per $5.6 \times 10^4$ years, while GstLAL found it with a SNR of 11.6 and a FAR of 1 per $1 \times 10^5$ years. Their primaries with masses $8.9^{+1.2}_{-1.5}M_\odot$ and $5.7^{+1.8}_{-2.1}M_\odot$ are conclusively BHs, whereas the secondary masses of $1.9^{+0.3}_{-0.2}M_\odot$ and $1.5^{+0.7}_{-0.5}M_\odot$

$^3$PyCBC did not have the capability to assign a significance to single-detector triggers at the time of the event.
lie below the maximal mass of a NS with 89%–96% and 87%–98% probability, respectively. For either event, no evidence of tidal effects or spin-induced quadruple effects was found in the signal, and no EM counterpart has been identified. Hence, the nature of neither secondary component can be determined conclusively.

5.3 Conclusions

By the end of LIGO–Virgo’s O3, we have seen all three categories of CBCs (BBH, BNS and NSBH coalescences). We used very confident events with FAR less than 1 per 1000 years in GWTC-3 to find possible breakdowns of GR in nine different ways and found them all consistent with GR [25]. We measured the merger rates of all categories of GW events (informing the predicted stochastic signal, to be discussed in the next chapters) and are building the distributions of the mass and spin of compact objects in binaries, delving into the formation and evolution of the compact objects detectable with LIGO and Virgo [26]. Using 47 events in GWTC-3, we measured the local expansion rate of the Universe $H_0$ to be $68^{+12}_{-7}$ km s$^{-1}$ Mpc$^{-1}$ with BBH redshifted masses and $68^{+8}_{-6}$ km s$^{-1}$ Mpc$^{-1}$ with the galaxy catalog method [29]. Looking to the future, as we continue to observe more GW events, we expect to further expand the frontiers of fundamental physics, astrophysics, and cosmology.
Direct detections of GWs from Advanced LIGO [1], Advanced Virgo [3] and KAGRA [77] detectors so far have been traced back to point-like sources, which make up a tiny fraction of the GW sky. The bulk of unresolved signals associated with multiple point sources or extended sources combine incoherently to form backgrounds of GWs. A stochastic gravitational-wave background (SGWB) consists of a large number of independent and uncorrelated events which are typically individually weak, i.e., below the detection threshold of the detector. SGWBs can be categorized as either astrophysical, when produced by low-redshift, individually indistinguishable GW events [96, 97], or cosmological [262], as a result of high-energy events in the early Universe such as cosmic inflation [263–265], cosmic string networks [266–269], primordial black holes [270–272], and first order phase transitions [273–278].

Ground-based GW detectors are sensitive to SGWBs constrained between tens of Hz and a few hundred Hz. In other frequency bands, upper limits on SGWBs are set by the isotropy of the CMB [279] in the lowest frequencies [55], by timing residual analyses in millisecond pulsars in the nHz band [280], by normal modes of the Earth [281] and the Moon [282] in the mHz to Hz band, and loosely by primordial deuterium abundance from Big Bang Nucleosynthesis [283, 284] over a broad frequency range.

Background gravitational radiation is stochastic in the sense that it can only be characterized statistically, in terms of moments of its probability distribution. Stochastic GW signals can mimic shot noise, appearing as individual bursts in the timestream, or they can be described as Gaussian, where a multitude of signals overlap so that the central limit theorem applies. They may also exhibit popcorn-like features in the time domain, with partially overlapping signals but still far from the Gaussian regime [69]. To differentiate between the aforementioned sources of stochastic backgrounds, several subtraction or multi-fitting methods have been proposed [285–287], which leverage on the particular statistical nature of each signal contribution.

Studying SGWBs may potentially open a window onto $\sim 10^{-32}$ s (at a redshift $z > 10^{25}$) after the Big Bang. Our current knowledge of the early Universe mostly comes from the CMB [53, 54], the relic EM radiation from 380,000 years ($z \sim 1100$)
after the Big Bang during the epoch of recombination. Before recombination, the Universe was opaque to EM waves. Hence, GWs present a unique opportunity to probe the earliest moments of the Universe.

In this chapter, we examine some properties and sources of SGWBs, introduce the cross-correlation method [288–291] for two detectors, and summarize the current observational constraints set by the LVK collaboration [292, 293], following the treatments in [69, 294–297] and refer readers to them for more details.

6.1 Properties of Stochastic Backgrounds

Plane Wave Expansions

To describe a SGWB, the plane wave solution Eq. (2.27) can be expanded into a superposition of sinusoidal plane waves coming from all directions on the sky,

\[ h_{\mu\nu}(t, \vec{x}) = \int_{-\infty}^{\infty} df \int_{S^2} d\Theta \ h_{\mu\nu}(f, \Theta) e^{i2\pi f(t-\Omega \cdot x/c)}, \]

where we define a direction on the sky to be \( \Theta = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) \) in standard angular coordinates \((\theta, \phi)\) on the two-sphere, \( \theta \in [0, \pi], \phi \in [0, 2\pi) \).

For a SGWB, the metric perturbations \( h_{\mu\nu}(f, \Theta) \) in Eq. (6.1) are random variables of some probability distribution that determines the statistical properties of the background, assumed to be stationary (constant in time). We can further expand the metric perturbations \( h_{\mu\nu}(f, \Theta) \) using + and \( \times \) polarization basis tensors as in Eq. (2.34), with the frequency and direction dependence specified explicitly,

\[ h_{\mu\nu}(f, \Theta) = h_+(f, \Theta) e_{\mu\nu}^+(\Theta) + h_\times(f, \Theta) e_{\mu\nu}^\times(\Theta). \]

In terms of the angular unit vectors,

\[ e_\theta = (\cos \theta \cos \phi, \cos \theta \sin \phi, -\sin \theta), \]
\[ e_\phi = (-\sin \phi, \cos \phi, 0), \]

the polarization basis tensors are written as

\[ e^+ = e_\theta \otimes e_\theta - e_\phi \otimes e_\phi, \]
\[ e^\times = e_\theta \otimes e_\phi + e_\phi \otimes e_\theta. \]

Putting Eq. (6.2) back into Eq. (6.1) gives

\[ h_{\mu\nu}(t, \vec{x}) = \int_{-\infty}^{\infty} df \int_{S^2} d\Theta \sum_{p=+, \times} h_p(f, \Theta) e_{\mu\nu}^p(\Theta) e^{i2\pi f(t-\Omega \cdot x/c)}, \]

where the Fourier amplitudes \( h_p(f, \Theta) \) are complex functions satisfying \( h_p(-f, \Theta) = h_p^*(f, \Theta) \) due to the reality constraint.
Statistical Properties

Gaussian – For cosmological backgrounds, we typically can assume they originate from Gaussian random processes [262]. Even for the more recent astrophysical backgrounds [96, 97], by the central limit theorem, a collection of a vast number of independent random processes is Gaussian. This means the joint density function is a multivariate normal distribution.

Stationary – The stationarity assumption in our analysis arises naturally since both the period of waves and the period of observations are orders of magnitude smaller compared to the age of the Universe. It seems unlikely that the properties of a SGWB would vary over either of these time scales. This means its statistical properties only depend on time differences rather than absolute points in time.

Unpolarized – A SGWB is unpolarized in the sense that incident GWs have statistically equivalent + and × polarization content and thus have no preferred component.

In this chapter, we deal with Gaussian, stationary, unpolarized backgrounds, with uncorrelated radiation from different directions on the sky. In addition to the three aforementioned assumptions, we can further assume (without loss of generality) the stochastic background has zero mean

\[ \langle h_{\mu\nu}(t, \vec{x}) \rangle = 0 \quad \iff \quad \langle h_{P}(f, \Theta) \rangle = 0. \quad (6.8) \]

For such a zero-mean Gaussian SGWB, the quadratic moment uniquely determines its statistical properties. The specific form of the second moment depends on the source of the background.

Since the CMB is highly isotropic [279], of which the largest intrinsic deviation from isotropy is \( \sim 10^{-5} \) due to the non-uniform distribution of matter at last scattering [298, 299], it is reasonable to assume that a SGWB of cosmological origin is also isotropic to a high degree.

The quadratic expectation values for a Gaussian, stationary, unpolarized, isotropic background are

\[ \langle h_{P}(f, \Theta) h^{*}_{P'}(f', \Theta') \rangle = \frac{1}{16\pi} S_{h}(f) \delta_{PP'} \delta_{ff'} \delta_{\Theta\Theta'}. \quad (6.9) \]

The Kronecker delta function \( \delta_{PP'} \) ensures the background has no preferred polarization; \( \delta_{ff'} \) is because of our stationarity assumption; \( \delta_{\Theta\Theta'} \) is due to uncorrelatedness of different directions on the sky. We add the constant \( 1/(16\pi) \) in the expression
so that \( S_h(f) \) can be interpreted as the one-sided GW strain power spectral density function of units \( \text{strain}^2 \text{Hz}^{-1} \), integrated over the sky and polarizations.

A SGWB of astrophysical origin is anticipated to be \textit{anisotropic}, concentrated in certain patches of the sky, such as following the distribution of the closest galaxies to us \([96, 97]\). And the astrophysical background from the closer-by Universe can overwhelm the cosmological one, forming a foreground to the more sought-after distant background. Notwithstanding, we cannot rule out the possibility of an anisotropic cosmological background.

The most general form of the quadratic expectation values for a Gaussian, stationary, unpolarized, anisotropic background is

\[
\langle h_P(f, \Theta) h_P^*(f', \Theta') \rangle = \frac{1}{4} \mathcal{P}(f, \Theta) \delta_{PP'} \delta_{ff'} \delta_{\Theta \Theta'},
\]

(6.10)

where \( \mathcal{P}(f, \Theta) \) specifies the spectral and spatial power distribution in the SGWB and

\[
S_h(f) = \int_{S^2} d\Theta \mathcal{P}(f, \Theta).
\]

(6.11)

For more general backgrounds (e.g., polarized, non-Gaussian, correlated radiation), we refer readers to \([296]\).

**Energy Density Spectrum**

Analogous to the CMB, we characterize the spectral properties of a SGWB by describing its energy distribution in frequency. Specifically, we introduce a dimensionless quantity, the normalized GW energy density spectrum,

\[
\Omega_{GW}(f) \equiv \frac{1}{\rho_c} \frac{d\rho_{GW}}{d \log f},
\]

(6.12)

where \( \rho_{GW} \) is the GW energy density and \( \rho_c \) is the critical energy density required to close the Universe today,

\[
\rho_c = \frac{3H_0^2c^2}{8\pi G}.
\]

(6.13)

Here, \( c = 2.998 \times 10^8 \text{ m s}^{-1} \) is the speed of light and \( H_0 = 67.4 \text{ km s}^{-1} \text{ Mpc}^{-1} \) \([300]\) is the Hubble expansion rate of the current epoch (with some controversy on its measured value in the literature \([29, 300–302]\)). Conceptually, \( \Omega_{GW}(f)(df/f) \) is the ratio of the GW energy density to the total energy density required to close the Universe today in a small frequency interval from \( f \) to \( f + df \).
For an isotropic background, we can relate the GW energy density spectrum $\Omega_{GW}(f)$ and the strain power density spectrum $S_h(f)$ by

$$\Omega_{GW}(f) = \frac{2\pi^2}{3H_0^2} f^3 S_h(f). \tag{6.14}$$

Similarly for an anisotropic background,

$$\Omega_{GW}(f, \Theta) = \frac{2\pi^2}{3H_0^2} f^3 \mathcal{P}(f, \Theta). \tag{6.15}$$

The total energy density in GWs normalized by the critical energy density is thus

$$\Omega_{GW} = \int_0^\infty d(ln f) \Omega_{GW}(f) \quad \text{or} \quad \Omega_{GW} = \int_{S^2} d\Theta \int_0^\infty d(ln f) \Omega_{GW}(f, \Theta). \tag{6.16}$$

We can then compare $\Omega_{GW}$ with other forms of energy density in the Universe: $\Omega_R$, radiation that includes photons and relativistic neutrinos; $\Omega_M$, matter that includes baryons and cold dark matter; $\Omega_k$, spatial curvature; and $\Omega_\Lambda$, vacuum energy.

With the assumption that a SGWB is Gaussian, stationary and unpolarized, the statistical properties of the isotropic background are completely specified by its energy density spectrum $\Omega_{GW}(f)$.

### 6.2 Sources of Stochastic Backgrounds

Based on the underlying generation mechanism, we categorize SGWBs broadly into two families, astrophysical and cosmological. Astrophysical backgrounds arise from individually undetected signals of astrophysical origin [96, 97], whereas cosmological backgrounds form hypothetically from vacuum fluctuations in inflation [263–265], via emission by topological defects [267–269], and during first order phase transitions in the early Universe [273, 274]. In this section, we briefly review both astrophysical and cosmological sources for SGWBs. We summarize the landscape of potential stochastic signals and various current observational constraints across the frequency spectrum in Fig. 6.1 [69].

**Astrophysical Backgrounds**

Essentially any unresolved GW signals generated by astrophysical processes add up to form astrophysical backgrounds, with the most prominent contributor being compact binary coalescences (CBCs) [96, 97].

The energy density spectrum due to a particular source class $q$ is

$$\Omega_q(f) = \frac{f}{\rho_0} \int_0^{z_{\text{max}}} dz \frac{R_q(z) (\frac{dE}{df_z})_{f_z}}{(1 + z) H(z)}, \tag{6.17}$$
Figure 6.1: The landscape of potential SGWB signals and various current observational constraints across the frequency spectrum, taken from [69]. The light blue curve shows the prediction for a SGWB from a slow-roll inflationary scalar field [303]. The pink curve shows a hypothetical SGWB from interactions of cosmic strings [304]. The brown curve shows a SGWB from inspiralling supermassive BBHs predicted from the NANOGrav 12.5-year data set [280]. The two grey curves show SGWBs produced by first-order phase transitions at the electroweak scale (\( \sim 200 \text{ GeV} \)) and the QCD scale (\( \sim 200 \text{ MeV} \)), respectively [305]. The yellow curve shows a SGWB generated by compact binary coalescences, based on the mass distributions and merger rates inferred by LVK detections [292]. The dashed curves show various current observational constraints at 95% confidence level. The dotted curve shows the integrated constraint from measurements of \( N_{\text{eff}} \) [306], the effective number of events of SGWB signals in a co-moving volume.

where \( R_q(z) \) is the source-frame rate per co-moving volume of GW emission from sources of class \( q \) and \( H(z) = H_0 \sqrt{\Omega_M (1 + z)^3 + \Omega_\Lambda} \) is the Hubble parameter at redshift \( z \). The quantity \( (dE_s/df_s)|_{f_s} \) is the source-frame energy radiated by a single source, evaluated at the source frequency \( f_s = f (1 + z) \). And we average over the ensemble properties of the given class \( q \),

\[
\left\langle \frac{dE_s}{df_s} \right|_{f_s} \right|_q = \int d\phi \; p_q(\phi) \frac{dE_s}{df_s}(\phi),
\]

(6.18)

where \( p_q(\phi) \) is the probability distribution of source parameters \( \phi \) such as masses, spins, etc., across class \( q \).

With confirmed detections from all three classes of compact binaries (BBHs, BNSs,
<table>
<thead>
<tr>
<th>$q$</th>
<th>$R_q(z = 0.2)$ [Gpc$^{-3}$ yr$^{-1}$]</th>
<th>$\Omega_q(25$ Hz$)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBH</td>
<td>$17.9 - 44$</td>
<td>$5.0^{+1.4}_{-1.8} \times 10^{-10}$</td>
</tr>
<tr>
<td>BNS</td>
<td>$10 - 1700$</td>
<td>$0.6^{+1.7}_{-0.5} \times 10^{-10}$</td>
</tr>
<tr>
<td>NSBH</td>
<td>$7.8 - 140$</td>
<td>$0.9^{+2.2}_{-0.7} \times 10^{-10}$</td>
</tr>
</tbody>
</table>

Table 6.1: Estimated merger rates and SGWB energy densities at a reference frequency of 25 Hz through GWTC-3 for three classes of compact binaries: BBHs, BNSs, NSBHs [26].

Figure 6.2: Energy density spectra predicted for contributions from BBHs, BNSs and NSBHs to the SGWB and for the combined SGWB due to CBCs [26]. The right panel also plots the 2$\sigma$ power-law integrated sensitivity curves of Advanced LIGO and Advanced Virgo for current and future configurations. Shaded bands encompass 90% credible level.

NSBHs) through GWTC-3 [6, 220, 228, 229], we predict contributions from every source class of CBCs to the SGWB [26, 292, 307–310]. The merger rates and energy densities at a reference frequency of 25 Hz for three classes of CBCs are listed in Table 6.1. The combined SGWB due to CBCs is bounded to be $\Omega_{CBC}(25$ Hz$) = 6.9^{+3.0}_{-2.1} \times 10^{-10}$ through GWTC-3 [26]. The energy density spectra due to CBCs and the comparison with current and future detector sensitivities are plotted in Fig. 6.2. Although the estimated SGWB amplitude due to CBCs lies well below the current LIGO Hanford–LIGO Livingston–Virgo (HLV) network sensitivity, it may be accessible with future detector upgrades such as “A+”.

Other astrophysical sources include asymmetrical core collapse supernovae [311], asymmetrically rotating neutron stars [312–314], stellar core collapses [315, 316], hypothetical depleting boson clouds around BHs [317–319], etc. In expectation
of louder astrophysical SGWBs masking cosmological ones, several methods have been proposed for subtracting astrophysical foregrounds to reveal cosmological backgrounds [320–322].

**Cosmic Inflation**

The astounding isotropy of the CMB [53, 54, 279] poses several interesting questions about the early Universe including the horizon and flatness problems. The horizon problem refers to the isotropy in the CMB temperatures across the whole sky despite not being in causal contact to establish thermal equilibrium. The flatness problem concerns the seemingly fine-tuned density of matter and energy in the Universe, of which a slight departure from the critical value would lead to a runaway process resulting in an open universe or a “Big Crunch” instead of a flat universe.

Inflation [323, 324] is a theory that solves these problems within the hot Big Bang model [325]. The inflationary epoch is a period of exponential growth of space that lasted from $10^{-36}$ s to $10^{-32}$ s after the Big Bang. In this epoch, GWs are conjectured to be produced by tensor fluctuations in the inflationary scalar field [263–265] and by the decay of the scalar field into regular matter during pre-heating at the end of inflation [326–328].

**Cosmic Strings**

Cosmic strings are hypothetical classical one-dimensional topological defects in a hypothetical quantum field formed during a symmetry-breaking phase transition at the end of inflation [266, 329, 330]. Cosmic superstrings are the cosmologically extended quantum counterpart of cosmic strings produced at the end of brane inflation [266, 331]. When cosmic strings or cosmic superstrings intersect, they swap partners with probability of 1 [332] or less than 1 [331]. These processes form cusps and kinks [333], which create bursts of GWs [267–269]. The superposition of these GW bursts from cosmic strings or cosmic superstrings over the history of the Universe creates a SGWB [268, 269].

**First Order Phase Transitions**

Modifications to the Standard Model of particle physics theorize a first order electroweak phase transition, which has a discontinuity in the first derivative of the free energy [334]. During such a first order phase transition, many physical mechanisms may generate GWs [273, 274]. Bubbles of a new vacuum phase are created when the electroweak symmetry is spontaneously broken and GWs are generated when
vacuum bubbles collide [275, 276]. Shocks and sound waves in the plasma [278] and magnetohydrodynamic turbulence [277] may also produce GWs.

6.3 Detection Methods

As instrumental noise is itself stochastic, this sort of signal is not clearly distinguishable from noise in a single detector, in particular in the case where the signal is weak with respect to the noise and both are hard to model independently. However, even a weak stochastic background induces a correlated signal across multiple detectors. Even though detector outputs are dominated by noise rather than by signals from stochastic backgrounds, one can still detect the signals by correlating outputs of multiple GW detectors for sufficiently long observing time, provided each detector has uncorrelated noise.

In this section, we present the cross-correlation technique [288–290], taking into account the separation and orientations of two different detectors [291], to search for a Gaussian, stationary, unpolarized, isotropic background following [296].

Cross-correlation Technique

Consider the case of a baseline $I$ made up of two ground-based GW detectors 1, 2 each with data output

$$ s(t) = h(t) + n(t), \quad \text{(6.19)} $$

where $h(t)$ denotes the strain due to a SGWB and $n(t)$ denotes the detector noise. The basic idea of cross-correlation is to multiply together the outputs of the two independent detectors and integrate,

$$ \langle C^I(t) \rangle = \langle s_1(t)s_2(t) \rangle = \int_{-T/2}^{T/2} dt \; s_1(t)s_2(t), \quad \text{(6.20)} $$

where $T$ is the coincident observing time of both detectors.

When the background gravitational radiation is much smaller than the detector noise, $\langle |h(t)|^2 \rangle \ll \langle |n(t)|^2 \rangle$, we can write

$$ \langle C^I \rangle = \langle h_1 h_2 \rangle + \langle h_1 n_2 \rangle + \langle h_2 n_1 \rangle + \langle n_1 n_2 \rangle \approx \langle h_1 h_2 \rangle + \langle n_1 n_2 \rangle \approx \langle h_1 h_2 \rangle. \quad \text{(6.21)} $$

We drop terms $\langle h_1 n_2 \rangle$ and $\langle h_2 n_1 \rangle$ since the GW signals and the instrumental noise are uncorrelated. The angle brackets here refer to an ensemble averaging, which is taken in practice by averaging over time, as well as frequency, baselines, and all other available independent measurements of the signal.
When detector noise is uncorrelated within the baseline, the expectation value of the cross-correlation between the strain in detector 1, \( s_1 \), and the strain in detector 2, \( s_2 \), will be sensitive to the signal component only. \( \langle C^I \rangle \simeq \langle h_1 h_2 \rangle \) is just the variance (or power) of the stochastic GW signals. In current stochastic searches performed on LIGO–Virgo data, the noise is assumed to be fully independent between detectors, hence the cross-correlation of the data streams yields an optimal statistic for the stochastic signal. The latter is often referred to in the literature as an \textit{optimal filter} [295].

We do not consider correlated noise in our discussion. However, there exists a type of noise, Schumann magnetic resonances caused by the EM field of the Earth, which can mimic a correlated SGWB in the detectors. Several methods have been proposed to mitigate these effects in a GW detector network, including noise subtraction methods [335–338], the “GW Geodesy” method [339, 340], and spectral modeling [341].

\textbf{Overlap Reduction Functions}

In reality, GW detectors are not co-located or co-oriented, inducing nontrivial angular response functions. We thereby use an overlap reduction function (ORF) to take care of the non-parallel alignment and the separation time delay between two detectors, which depends entirely on the relative position and orientation of a detector pair [291]. This normalized function of frequency returns values less than or equal to 1, with unity achieved when the two detectors are co-aligned and coincident.

The normalized ORF takes the form

\[
\gamma^I(t; f) \equiv \gamma^I_n(t; f) = \frac{5}{8\pi} \sum_{p=\pm,\times} \int_{S^2} d\Theta \, F^P_1(t; \Theta) F^P_2(t; \Theta) e^{i2\pi f\Theta \cdot \Delta x(t)/c} e_\eta(\Theta),
\]

(6.22)

where \( \Theta \) is a unit vector, \( e_\eta(\Theta) \) is a set of basis functions on the two sphere, \( \Delta x(t) \) is the separation vector between the two detectors at time \( t \), and \( F^P(t; \Theta) \) is the detector antenna response to the \( P = \pm, \times \) polarization. Fig. 6.3 shows the normalized ORFs for LIGO Hanford–LIGO Livingston\(^1\) (HL), LIGO Hanford–Virgo (HV) and LIGO Livingston–Virgo (LV) baselines in the small-antenna limit (\( L \ll c/f_{GW} \)), up to 500 Hz. The normalized ORF is periodic with a period of a sidereal day because the detector antenna response is periodic. Fig. 6.4 gives the cyclic modulation of

\(^1\) HL and LV are designed to be as co-oriented as possible (modulo the curvature of the Earth), up to an overall 90° rotation of the detector arms; hence the ORF approaches \(-1\) at low frequencies.
Figure 6.3: Normalized overlap reduction functions (ORFs) $\gamma(f)$ for constituting baselines HL, HV and LV in the HLV detector network in the small-antenna limit ($L \ll c/f_{GW}$). The ORFs show that the correlation of a detector pair to an unpolarized SGWB falls off with frequency rapidly.

network antenna patterns defined as

$$F(t; \Theta) = \sum_{\text{all baselines}} F_1^+(t; \Theta) F_2^+(t; \Theta) + F_1^x(t; \Theta) F_2^x(t; \Theta)$$  \hspace{1cm} (6.23)$$

for unpolarized waves in the HLV network due to the Earth’s daily rotation.

**Optimal Filtering for Isotropic Backgrounds**

To illustrate the cross-correlation technique, we derive the (nearly) optimal filtering for a Gaussian, stationary, unpolarized, isotropic background using two interferometric GW detectors with uncorrelated noise. We will elaborate on the maximum-likelihood estimation for a Gaussian, stationary, unpolarized, anisotropic background in Ch. 7.

Start with the Fourier transform of the cross-correlated power defined in Eq. (6.20),

$$\langle \hat{C}^I(f) \rangle = \langle \tilde{s}_1(f) \tilde{s}_2^*(f) \rangle$$

$$\approx \langle \tilde{h}_1(f) \tilde{h}_2^*(f) \rangle$$

$$= \frac{T}{2} \gamma^I(f) S_h(f),$$  \hspace{1cm} (6.24)$$

where tilde denotes the Fourier transform of a function, asterisk represents the complex conjugate, $T$ is the observing time, $\gamma^I(f)$ is the ORF of the two detectors, and $S_h(f)$ is the GW strain power spectrum.
Figure 6.4: Time evolution of the antenna pattern map for unpolarized waves at $f_{GW} = 0$ in the HLV network in the small-antenna limit. Each subplot shows the antenna pattern at a different hour of a day, in intervals of 3 hours. A full cycle corresponds to one sidereal day.
In the weak signal limit, the covariance matrix reduces to

\[ N(f, f') \equiv \langle \tilde{C}^l(f)\tilde{C}^{l*}(f') \rangle - \langle \tilde{C}^l(f) \rangle \langle \tilde{C}^{l*}(f') \rangle \]

\[ \simeq \langle \tilde{n}_1(f)\tilde{n}_1^*(f') \rangle \langle \tilde{n}_2(f)\tilde{n}_2^*(f') \rangle \]

\[ = \frac{T}{4} P_{n_1}(f) P_{n_2}(f) \delta_{ff'}. \]  

(6.25)

Here, \( P_n(f) \) are the one-sided noise power spectral densities in the two detectors satisfying

\[ \langle \tilde{n}(f)\tilde{n}^*(f') \rangle = \frac{1}{2} P_n(f) \delta_{ff'}. \]  

(6.26)

In this approximation,

\[ \int_{-\infty}^{\infty} df' N^{-1}(f, f') \simeq \frac{4}{T} \frac{1}{P_{n_1}(f) P_{n_2}(f)}. \]  

(6.27)

The spectral dependence of a SGWB is crudely modeled as (an approximation to the spectrum in Fig. 6.2 for \( f \lesssim 100 \text{ Hz} \)) a power law with spectral index \( \alpha \),

\[ \Omega_{GW}(f) = \Omega_\alpha \left( \frac{f}{f_{\text{ref}}} \right)^\alpha. \]  

(6.28)

For spectral index \( \alpha \), \( \Omega_\alpha \) is the normalized GW energy density at a reference frequency \( f_{\text{ref}} \), usually chosen in the most sensitive region of the analysis.

Then, according to Eq. (6.14),

\[ S_h(f) = \frac{3H_0^2}{2\pi^2} \frac{\Omega_\alpha}{f_{\text{ref}}^3} \left( \frac{f}{f_{\text{ref}}} \right)^{\alpha-3} = \Omega_\alpha H_\alpha(f), \]  

(6.29)

with

\[ H_\alpha(f) = \frac{3H_0^2}{2\pi^2} \frac{1}{f_{\text{ref}}^3} \left( \frac{f}{f_{\text{ref}}} \right)^{\alpha-3}. \]  

(6.30)

Now with Eq. (6.24) and Eq. (6.29), we can construct an estimator for \( \Omega_\alpha \),

\[ \hat{\Omega}_\alpha = \Xi \int_{-\infty}^{\infty} df \frac{\gamma_{12}(f) H_\alpha(f)}{P_{n_1}(f) P_{n_2}(f)} \tilde{C}^l(f), \]  

(6.31)

where

\[ \Xi \equiv \left[ \frac{T}{2} \int_{-\infty}^{\infty} df \frac{\gamma_{12}^2(f) H_\alpha^2(f)}{P_{n_1}(f) P_{n_2}(f)} \right]^{-1}. \]  

(6.32)

The variance of the estimator is \( \sigma^2_{\hat{\Omega}_\alpha} = \Xi/2 \) and the expected signal-to-noise ratio is

\[ \rho_{\hat{\Omega}_\alpha} = \sqrt{T} \left[ \int_{-\infty}^{\infty} df \frac{\gamma_{12}^2 f S_h^2(f)}{P_{n_1}(f) P_{n_2}(f)} \right]^{1/2}. \]  

(6.33)
We highlight that Eq. (6.33) implies with sufficient observing time, in principle we can detect a SGWB regardless of the level of detector noise.

Finally, the optimal filter for a Gaussian, stationary, unpolarized, isotropic background is found to be

\[
\tilde{Q}^l(f) = \Xi_\ast \gamma^l(f) H_\alpha(f) \frac{P_{n_1}(f) P_{n_2}(f)}{\tilde{C}^l(f)}.
\]

(6.34)

6.4 Current Observational Constraints from LIGO–VIRGO–KAGRA

For both isotropic and anisotropic searches, three power laws are often considered: \(\alpha = 0\) corresponds to a flat energy density spectrum from cosmic strings [269] and slow-roll inflation [342]; \(\alpha = 2/3\) recoupts unresolvable inspirals from CBCs [96]; and \(\alpha = 3\) characterizes a fiducial flat strain power spectrum from supernovae [311]. In the absence of SGWB detections, we place upper limits on the normalized GW energy density spectrum as a function of \(\alpha\).

Isotropic Searches – The latest O3 analysis from the LVK collaboration, combined with upper limits from the earlier O1 and O2 runs, places constraints on power-law isotropic SGWBs at the 95% credible level to have amplitudes \(\Omega_\alpha\) at a reference frequency \(f_{\text{ref}} = 25\) Hz listed in Table 6.2: \(\Omega_0 \leq 5.8 \times 10^{-9}\), \(\Omega_{2/3} \leq 3.4 \times 10^{-9}\), and \(\Omega_3 \leq 3.9 \times 10^{-10}\) [292].

Anisotropic Searches – Using a combined dataset from the first three observing runs of all operational detectors (both LIGO detectors in all three runs and Virgo in O3), the LVK collaboration produced upper limit anisotropic sky maps of SGWBs [293]. Table 6.2 lists the 95% confidence-level upper limit ranges at a reference frequency \(f_{\text{ref}} = 25\) Hz on the GW energy density flux \(F_{\alpha,\Theta}\) [erg cm\(^{-2}\) s\(^{-1}\) Hz\(^{-1}\)] from a broadband radiometer search [343, 344] for point sources and the normalized GW energy density \(\Omega_\alpha(\Theta)\) [sr\(^{-1}\)] from a spherical harmonic decomposition search [345] for spatially-extended sources. We refer readers to [293] for the upper limit sky maps and more details.
Table 6.2: 95% confidence-level upper limits on the normalized GW energy density, normalized GW energy density spectrum and GW energy flux using combined data from LIGO–Virgo’s first three observing runs \([292, 293]\). Results are from an isotropic search, a spherical harmonic decomposition (SHD) search for spatially-extended sources and a broadband radiometer (BBR) search for point sources.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>Isotropic</th>
<th>SHD</th>
<th>BBR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5.8</td>
<td>3.2 – 9.3</td>
<td>1.7 – 7.6</td>
</tr>
<tr>
<td>2/3</td>
<td>3.4</td>
<td>2.4 – 9.3</td>
<td>0.85 – 4.1</td>
</tr>
<tr>
<td>3</td>
<td>0.39</td>
<td>0.57 – 3.4</td>
<td>0.013 – 0.11</td>
</tr>
</tbody>
</table>
MODEL-INDEPENDENT SEARCH FOR ANISOTROPIES IN STOCHASTIC GRAVITATIONAL-WAVE BACKGROUND

At cosmological scales, we assume the GW sky is isotropic based on the isotropy [279] of the CMB [53, 54], which traces the matter distribution in the Universe. However, at local scales, the nonuniform distribution of astrophysical GW sources mentioned in Section 6.2 may produce an anisotropic SGWB [96, 97]. Moreover, similar to the CMB dipole anisotropy [298, 299] (the kinematic interpretation of which is recently under contestation with high significance [346] by a quasar [347–349] study using the Wide-field Infrared Survey Explorer (WISE) [350], which indicates a possible deviation from the standard Friedman-Lemaître-Robertson-Walker cosmology), our peculiar motion with respect to the SGWB rest frame induces a recurring modulation affecting the dipole. It is thus fair to assume the SGWB power contains anisotropic components on top of an ensemble average isotropic value.

An approach to reconstruct the angular power distribution in an anisotropic SGWB is computing the maximum-likelihood map solutions using cross-correlated data [288–290] from a network of ground-based GW detectors. This is typically done assuming stationary Gaussian detector noise and a specific model for the spectral power distribution of the signal, and employing a weak-signal approximation [344, 345]. The latter implies we can ignore any signal contribution to the data auto-correlations, essentially allowing us to estimate the noise covariance from the data directly [351]. Mapping can be carried out in any set of basis functions on the sky, e.g., spherical harmonics of the SGWB power as adopted by the LVK collaboration [292, 293], or sky pixel indices as in [352–357].

Directional searches by the LVK collaboration [293, 358–360] include the broadband radiometer analysis (BBR) [343], the spherical harmonic decomposition (SHD) [345], the narrow band radiometer analysis (NBR) [343], and the all-sky, all-frequency analysis (ASAF) [361]. The BBR targets persistent point sources emitting GWs over a wide frequency range, whereas the SHD hunts for extended sources with smooth frequency spectra. The NBR inspects frequency spectra from specific locations on the sky, such as Scorpius X-1, SN 1987A and the Galactic Center, in
narrow frequency bands. The ASAF scans the sky in individual frequency bins, searching for excess GW power for each pixel-frequency pair. The BBR, SHD and NBR approaches integrate over frequencies employing a filter which includes a power-law model for the SGWB power spectrum, while the ASAF is a generic unmodeled search. Out of all of these, the SHD search is the only one that takes pixel-pixel correlations into account.

Complementary to the LVK searches, we present an efficient analysis pipeline built in Python to map anisotropies in SGWBs directly in the sky pixel domain using data folded over one sidereal day [362]. Our pipeline is tailored to folded data [363–365], which assumes the SGWB signal is stationary (i.e., is time-independent) and exploits the temporal symmetry of the Earth’s rotation to reduce the computation time by a factor of total observing days. In the pipeline, we use the HEALPix hierarchical pixelization scheme [366], in which the sky is discretized into equal area elements. We invert the “full” pixel–pixel correlation matrix in map-making of the GW sky, up to an optimal eigenmode cutoff decided systematically using simulations. In addition to modeled maximum-likelihood mapping, we implement a spectral-model-independent method to probe the spectral shape of a SGWB based on previous work in [355], now taking into account the deconvolution regularization problem systematically as a function of frequency. In both approaches, sky maps are converted from the pixel domain to the Fourier domain to place upper limits on the angular power spectrum, as well as the power spectrum of the monopole component of the background.

7.1 SGWB Mapping
We follow [296, 344, 345] to derive the maximum-likelihood map solutions, which are equivalent to the optimal filter Eq. (6.34) introduced in Chapter 6, for a Gaussian, stationary, unpolarized, anisotropic SGWB. In what follows, we adopt the celestial coordinate system with the origin at the center of the Earth and $e_z$ along the rotational axis of the Earth.

Strain Power Factorization
The SGWB is, in general, a function of both GW frequency and sky location, and is likely to have different frequency dependence from different sky locations. For the sake of simplicity, stochastic searches typically assume that the directionality and the spectral shape of the signal are independent, such that the GW strain power
\( \mathcal{P}(f, \Theta) \) in Eq. (6.9) may be factored into a spectral term and an angular term\(^1\),

\[ \mathcal{P}(f, \Theta) = H(f)\mathcal{P}(\Theta). \tag{7.1} \]

The spectral shape \( H(f) \) is usually modeled as a power law given by

\[ H_\alpha(f) = \left( \frac{f}{f_{\text{ref}}} \right)^{\alpha-3}, \tag{7.2} \]

where \( \alpha \) is the spectral index and \( f_{\text{ref}} \) is a reference frequency. This choice of model is well-motivated by many astrophysical and cosmological models [96, 262], however there are well-known spectral shapes outside this regime, e.g., the combined SGWB from CBCs at higher frequencies shown on the right of Fig. 6.2 [26]. The power-law assumption is a good approximation for the CBC SGWB at current detector sensitivities, but is expected to break down as sensitivity increases. For a direct comparison with the LVK results [293], we also set \( f_{\text{ref}} \) to 25 Hz.

\( \mathcal{P}(\Theta) \) in Eq. (7.1) is the angular power distribution that can be expanded in a set of chosen basis functions \( e_\eta(\Theta) \) on the two sphere,

\[ \mathcal{P}(\Theta) = \sum_\eta \mathcal{P}_\eta e_\eta(\Theta). \tag{7.3} \]

For a pixel basis, we write

\[ \mathcal{P}(\Theta) \equiv \mathcal{P}(\Theta_p) = \mathcal{P}_{p'} \delta(\Theta_p, \Theta_{p'}), \tag{7.4} \]

where \( \mathcal{P}_{p'} \) is the power of the signal in each pixel. For a spherical harmonic expansion,

\[ \mathcal{P}(\Theta) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \mathcal{P}_{lm} Y_{lm}(\Theta), \tag{7.5} \]

where \( \mathcal{P}_{lm} \) are the spherical harmonic coefficients of the signal and \( Y_{lm}(\Theta) \) are the spherical harmonic basis functions. Note that in general the units of sky power components may be different depending on the basis and conventions used. Here, we assume units of GW sky power are strain power per steradian.

**Cross-correlation Statistic**

In practice, it is usually more efficient to work with data divided into smaller time segments and transformed to the frequency domain, making use of the Fast Fourier

\(^1\)See Section 7.2 for a brief discussion of the validity of this assumption.
Transform (FFT) [367] algorithm and parallel processing. In the case we consider here, the data are split into segments of equal duration \( \tau \), where \( \tau \) is much bigger than the light travel time between the two detectors but small enough so that detector response functions do not change significantly over the interval.

The cross-spectral density (CSD) for a baseline \( I \) of two detectors evaluated in time segment labeled \( t \) and at frequency \( f \) is defined as

\[
C^I(t; f) = \frac{2}{\tau} \tilde{s}_1^*(t; f) \tilde{s}_2(t; f) \approx \frac{2}{\tau} \tilde{h}_1^*(t; f) \tilde{h}_2(t; f),
\]

(7.6)

where \( \tilde{s}(t; f) \) is the short-term Fourier transform (SFT) of \( s(t) \) of duration \( \tau \). For conventions used, please see [296]. Then, by Eq. (7.1) and the SFT of Eq. (6.7), the expectation value of \( C^I(t; f) \) is given by [296]

\[
\langle C^I(t; f) \rangle = \tau H(f) \sum_\eta \mathcal{P}_\eta \gamma^I_\eta(t; f),
\]

(7.7)

where \( \gamma^I_\eta(t; f) \) is the unnormalized ORF in Eq. (6.22), which describes the correlated sensitivity of the baseline to the sky and frequency modes of the signal.

In the pixel basis, \( \eta \to p \), so that the unnormalized ORF becomes

\[
\gamma^I_{p; tf} = \sum_{p=+} F^1_P(t; \Theta_p) F^P_2(t; \Theta_p) e^{i2\pi f \Theta_p \cdot \Delta x(t)/c},
\]

(7.8)

where \( F^P(t; \Theta_p) \) are detector response functions for \( P = \{+,-\} \) plane polarized waves, and \( \Theta_p \) is the general direction on the sky discretized in the pixel domain, i.e., it is the direction pointing to the center of the pixel \( p \). The ORF can be transformed to the spherical harmonic basis by

\[
\gamma^I_{lm; tf} = \int_{S^2} d\Theta_p \gamma^I_{p; tf} Y^*_{lm}(\Theta_p).
\]

(7.9)

Note that the normalization of this function on the whole sky is \( 5/(8\pi) \) [295].

Using compact notation with summation over directions \( \Theta \) on the sky implied, we put the signal model Eq. (7.7) in a general basis into matrix form

\[
\langle C^I(t; f) \rangle = K^I(t; f, \Theta) \mathcal{P}(\Theta) \Rightarrow \langle C^I_{i\eta} \rangle = K^I_{i\eta} \cdot \mathcal{P}_\eta,
\]

(7.10)

where

\[
K^I_{i\eta} \equiv \tau H(f) \gamma^I_\eta(t; f).
\]

(7.11)
The noise covariance matrix for the CSD is subsequently [296]

\[
N_{t,f,t',f'}^l \equiv \langle C_{t,f}^l C_{t',f'}^l \rangle - \langle C_{t,f}^l \rangle \langle C_{t',f'}^l \rangle \\
\approx \frac{\tau^2}{4} \delta_{tt'} \delta_{ff'} P_{n_1}(t; f) P_{n_2}(t; f),
\]
(7.12)

where the one-sided noise power spectrum \(P_n\) satisfies

\[
\langle \tilde{s}(t; f) \tilde{s}^*(t'; f') \rangle \simeq \langle \tilde{n}(t; f) \tilde{n}^*(t'; f') \rangle \\
= \frac{\tau}{2} \delta_{tt'} \delta_{ff'} P_n(t; f).
\]
(7.13)

To lighten the notation in remaining derivation, we drop superscripts for detector baselines and subscripts for function dependencies when there is no confusion.

**Likelihood Function**

We assume detector noise is Gaussian and stationary on the timescale \(\tau\), and further assume that the SGWB is Gaussian, unpolarized, and its spectral shape \(H(f)\) is known\(^2\).

The likelihood function for the cross-correlation statistic of a single baseline is then (using short-hand notation)

\[
\mathcal{L}(C|\mathcal{P}) \propto \prod_{t,f} \exp \left[ -\frac{1}{2} \chi^2(\mathcal{P}) \right],
\]
(7.14)

where, given the signal model in Eq. (7.10), the chi-squared statistic is

\[
\chi^2(\mathcal{P}) \equiv (C - \langle C \rangle)^\dagger N^{-1}(C - \langle C \rangle) \\
= (C - K \cdot \mathcal{P})^\dagger N^{-1}(C - K \cdot \mathcal{P}),
\]
(7.15)

where the dot product indicates a sum over spatial indices.

**Maximum-likelihood Maps**

Maximizing the likelihood function Eq. (7.14) with respect to \(\mathcal{P}\) is equivalent to minimizing the chi-squared statistic Eq. (7.15). By matrix differentiation, we derive the maximum-likelihood estimates of angular power spectrum, the *clean map*,

\[
\hat{\mathcal{P}} = (K^\dagger \cdot N^{-1} \cdot K)^{-1} \cdot (K^\dagger \cdot N^{-1} \cdot C) \equiv \Gamma^{-1} \cdot X \\
\Rightarrow \hat{\mathcal{P}}_\eta = \sum_{\eta'} \Gamma_{\eta\eta'}^{-1} X_{\eta'},
\]
(7.16)

\(^2\)In case of a non-Gaussian signal, we can expect the approach to be sub-optimal, as the likelihood used does not capture key features of the signal. In case of a polarized background, extra terms to the ORF must be considered [353].
where $X$ is the so-called dirty map, and $\Gamma$ is the Fisher information matrix. On the left we show computations in matrix form, and on the right we show the corresponding element-wise computations in a general basis.

The clean map statistic can be viewed as a directional extension of the optimal statistic in Eq. (6.34) derived in [295], and is thus robust to noise non-stationarity on time-scales longer than the analyzed time segment $\tau$, as it consists of an inverse noise-weighted average over segments.

The dirty map represents the GW sky seen through the beam matrix of the two detectors and is given by

$$X = K^\dagger \cdot N^{-1} \cdot C \quad \Rightarrow \quad X_\eta = \frac{4}{\tau} \sum_{t,f} \frac{H(f) \gamma_{\eta f}^{I*}}{P_{n_1}(t; f)P_{n_2}(t; f)} \tilde{s}_1(t; f) \tilde{s}_2(t; f). \quad (7.17)$$

The Fisher matrix, which can be interpreted as a point spread function, codifying how signals from point sources spread elsewhere due to finite coverage of the sky by a network of GW detectors, is

$$\Gamma = K^\dagger \cdot N^{-1} \cdot K \quad \Rightarrow \quad \Gamma_{\eta \eta'} = 4 \sum_{t,f} \frac{H^2(f)}{P_{n_1}(t; f)P_{n_2}(t; f)} \gamma_{\eta f}^{I*} \gamma_{\eta' f}^I. \quad (7.18)$$

In Fig. 7.1, we give plots of the point spread functions for the HLV detector network in its O3 sensitivity, assuming a flat strain spectrum $H_a(f) = 1$ (i.e., $\alpha = 3$) for point sources. Given data over a full sidereal day, the point spread function is translationally invariant along the right ascension of a source.

The above derivation for a baseline of two GW detectors is easily generalized to a multi-detector network. Assuming each baseline provides an independent measurement of the signal, it is sufficient to sum dirty maps and Fisher matrices over all baselines in the network

$$X = \sum_I X_I', \quad \Gamma = \sum_I \Gamma_I'. \quad (7.19)$$

to obtain the network clean map using Eq. (7.16).

In the weak signal limit, we can further show [345]

$$\langle X \cdot X^\dagger \rangle - \langle X \rangle \langle X^\dagger \rangle \approx \Gamma, \quad (7.20)$$

$$\langle \Phi \cdot \Phi^\dagger \rangle - \langle \Phi \rangle \langle \Phi^\dagger \rangle \approx \Gamma^{-1}. \quad (7.21)$$
Figure 7.1: Point spread functions $\Gamma_{\eta \eta'}$ for the HLV detector network in its O3 sensitivity, assuming a flat strain spectrum for point sources. Black dots denote point sources at particular chosen sky location values $\Theta_\eta$. Due to finite coverage of the sky, signals from point sources spread to other locations on the sky. The point spread function is translationally invariant along the right ascension of a source, given data over a full sidereal day. The last subplot shows the interplay of point spread functions for all above carefully placed point sources on a cross on the sky. These use the HEALPix pixelization scheme with $N_{\text{side}} = 16$, $N_{\text{pix}} = 3072$.

Therefore, $\Gamma$ is the covariance matrix for the dirty map $X$ and $\Gamma^{-1}$ is the covariance matrix for the clean map $\hat{P}$.

We can then define the SNR map to be the result of the matrix multiplication \cite{354}

$$\rho = \Gamma^{\frac{1}{2}} \cdot \hat{P} \quad \Rightarrow \quad \rho_\eta = \sum_{\eta'} \Gamma_{\eta \eta'}^{\frac{1}{2}} \hat{P}_{\eta'},$$

which takes off-diagonal elements of the Fisher matrix into account, and the noise standard deviation map to be

$$\sigma = \sqrt{\text{diag} \Gamma^{-1}} \quad \Rightarrow \quad \sigma_\eta = \Gamma_{\eta \eta}^{-\frac{1}{2}}.$$  

(7.23)

The noise map so defined is only sensitive to diagonal elements of the inverse Fisher matrix, ignoring all pixel-pixel correlations. However, correlations between different locations on the sky are nontrivial. The noise map is thereby only an approximation of the noise standard deviation of the estimator $\hat{P}$. In the case of a
singular Fisher matrix, the calculation of the SNR requires regularizing adjustments as described below.

The dirty maps and Fisher matrices may be calculated over broad frequency bands to improve detection statistics. However, this implies integrating over the spectral shape of the SGWB, $H(f)$, hence the resulting clean map Eq. (7.16) is a biased estimator of the angular power distribution, as we do not know $H(f)$ \textit{a priori}. The standard spectral-model approach is to assume a power-law spectral model $H_\alpha(f)$ as in Eq. (7.2) and estimate $\hat{P}$ for a set of $\alpha$ candidates. We consider here three possible spectral index values, in keeping with the LVK searches [293]: $\alpha = 0$, a flat energy density spectrum consistent with many cosmological models [262]; $\alpha = 2/3$, an astrophysical background dominated by CBCs [96]; and $\alpha = 3$, a generic flat strain spectrum [311]. Other approaches, such as the ASAF approach, solve for $\hat{P}$ in each frequency bin, and do not require a model for $H(f)$; however, in this case it is not possible to invert the full Fisher matrix, as it is prohibitively singular in a single frequency bin. This is the main motivation for integrating over broader frequency ranges when taking pixel-pixel correlations into account.

**Deconvolution Regularization**

To perform the deconvolution in Eq. (7.16), we need to invert the Fisher matrix which is typically singular due to the uneven sampling of the sky. In the absence of detections, current search methods employed by both the LVK collaboration and independent groups condition the Fisher matrix in an ad hoc way; specifically, the LVK has proceeded either by restricting only to diagonal elements and hence ignoring all pixel-pixel correlations as in the BBR search for point sources, or discarding the smaller 1/3 of eigenvalues and fixing a maximum multipole as in the SHD search for extended sources [293]. Other groups have instead chosen a fixed condition number for the Fisher matrix [354, 355]. It is clear that moving towards the detection era for SGWBs systematic ways to better regularize the Fisher matrix must be established, especially to claim confident detections.

The Fisher matrix is in general singular since there exist null directions the detector network is insensitive to [353, 368]. On the left of Fig. 7.2, we show Fisher matrices of the HLV network in its O3 sensitivity for different power laws, where in all cases null directions are evident as the blue areas of the matrices with values close to zero. To address this issue, we use a singular value decomposition (SVD) [369] of the Fisher matrix to rank contributing directions and discard eigenmodes which carry
Figure 7.2: Fisher matrices in the pixel domain with HEALPix ordering, with HEALPix parameter $N_{\text{side}} = 16$, for a total of $N_{\text{pix}}^3 = 12N_{\text{side}}^3 = 3072$ pixels, and corresponding eigenvalues from a singular value decomposition (SVD). From top to bottom, power laws of spectral indices 0, 2/3 and 3 are represented for the HLV network between 20 and 1726 Hz in its O3 sensitivity. On the left, Fisher matrices contain null directions, leading to very low SVD eigenvalues shown on the right. On the right, we see Fisher matrices of the three-detector HLV network are better conditioned than those of the constituting baselines, HL, HV and LV. Grey dashed lines point to the optimal minimum eigenvalue cutoffs and numbers of eigenmodes kept in each case.
little information [345].

The inherent condition number of the matrix, i.e., the ratio between the minimum and maximum eigenvalues, depends on the spectral shape. Including too many eigenmodes introduces unwanted noise whereas discarding too many eigenmodes sacrifices signals. The SVD technique allows to condition the matrix, i.e., impose an eigenvalue threshold such that all normalized eigenvalues that are smaller than the imposed condition number are discarded. This approach was previously explored systematically in [357].

We present a method to determine this threshold empirically via simulations [362]. For each spectral index $\alpha$, we impose the condition number returning the least residual sum of squares (RSS) from a respective high SNR monopole simulation,

$$\text{RSS} = (\mathcal{P}_{\text{inj}} - \hat{\mathcal{P}})^T (\mathcal{P}_{\text{inj}} - \hat{\mathcal{P}}),$$

(7.24)

where $\mathcal{P}_{\text{inj}}$ is the injected monopole. Monopole simulations are used for the calculation since we expect an intrinsic monopole irrespective of spectral shapes, and stronger than any higher multipoles [69]. The monopole simulations are generated using the HLV network configuration in its O3 sensitivity, since most sensitivity of the combined O1+O2+O3 runs comes from O3 as demonstrated in Fig. 7.3. The residuals computed for different condition numbers are illustrated in the left panel of Fig. 7.4. The optimal condition numbers with the smallest residuals for different power laws are listed in Table 7.1 and also indicated on the right subplots of Fig. 7.2. The percentages of eigenvalues kept using the optimal condition numbers are also shown in Table 7.1. Note these are quite different from the nominal value of $2/3$ in the LVK SHD searches, and depend strongly on the spectral index. The comparison between the Fisher matrix eigenvalues and the associated optimal condition numbers for the HLV network for different power laws is shown in the right panel of Fig. 7.4.

A GW detector network is diffraction-limited, i.e., the resolution and hence the point spread function inherently depends on the frequency of the source. Choosing a pixel basis with a higher resolution than the internal detector resolution at the relevant signal frequencies compromises SNRs of the deconvolved map. The angular resolution of a two-detector baseline is estimated by the diffraction limit [370]

$$\Delta \theta \simeq \frac{\lambda}{2D} = \frac{c}{2fD},$$

(7.25)

where $D$ is the baseline length. The expected angular resolution $\ell_{\text{max}} = \pi/\Delta \theta$ per frequency for our analyses can be derived from Eq. (7.25), with some technicalities.
The HL baseline length $D_{HL} = 3002$ km is used throughout the analyses for being the most sensitive baseline, so we expect this baseline to dominate the resolution.

Furthermore, for broadband analyses we expect each frequency to contribute as a function of overall signal spectral shape [293]. While we quote results at a fixed reference frequency in this case, chosen in line with previous results, this frequency does not determine our angular resolution.

On the other hand, for the banded approach described below, we quote results at the midpoint of each frequency band considered. For example, for our analyses between 20 and 1726 Hz, the frequency midpoint $f = 873$ Hz and $D \approx 3000$ km give $\Delta \theta = 0.0572$ rad. To cover the sky with $4\pi$ rad$^2$, we need $4\pi / (\Delta \theta)^2 = 3841$ pixels, comparable to the number of pixels $N_{\text{pix}} = 3072$ in sky maps with $N_{\text{side}} = 16$. This indicates that we are searching the data for point sources with an angular resolution.
Figure 7.4: The left panel shows the condition numbers (i.e., eigenvalue thresholds) and resulting normalized residual sums of squares for power laws of spectral indices 0, 2/3 and 3. We select the threshold returning the least residual sum of squares in a monopole injected simulation for each spectral index. The right panel illustrates the comparison between the Fisher matrix eigenvalue distributions for the HLV network for different values of $\alpha$. The horizontal dashed lines show the optimal thresholds determined via monopole simulations.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>Condition number</th>
<th>Percentage of eigenvalues kept</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$10^{-5.5}$</td>
<td>27.51%</td>
</tr>
<tr>
<td>2/3</td>
<td>$10^{-5.5}$</td>
<td>48.93%</td>
</tr>
<tr>
<td>3</td>
<td>$10^{-3.5}$</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 7.1: Optimal condition numbers and associated percentages of eigenvalues kept for Fisher matrices of the HLV network in its O3 sensitivity. Results are computed empirically via monopole injected simulations in a pixel basis of $N_{\text{pix}} = 3072$ pixels.

appropriate for a diffraction-limited measurement. As these are not broad-band integrated results, the reference frequency used here can give a reasonable estimate of the expected angular resolution in each band. Note that a recent study shows that the diffraction limit is not optimal to resolve sources [371], however we are most concerned here with maximizing the detection statistic, not the recovered resolution. In our pixel-basis approach, we use this limit as a lower bound on the number of pixels to use, so as to over-resolve the anisotropies. The upper bound on pixel number is set by Fisher matrix regularization, as described below.

Finally, adding more detectors to the network is a form of regularization, since it provides larger coverage of the sky. With ever-improving sensitivities of existing detectors and addition of new detectors (KAGRA [77] and LIGO India [78]) in
the future, Fisher matrices in modeled broadband searches will be much better conditioned so that specialized regularization techniques will become less important. On the other hand, however, the spectral-model-independent method described in Section 7.2 relies heavily on trustworthy regularization of all Fisher matrices in its narrowband searches.

**Multipole Moments**

Extended anisotropies are parameterized in multipole moments of the power on the sky, which are quantified by their spherical harmonic coefficients. We carry out our analysis in the pixel domain by choosing a pixel basis expansion as in Eq. (7.4). Hence, to obtain limits on anisotropies about the mean background, we convert from the pixel basis to the spherical harmonic basis.

We can construct estimators of spherical harmonic coefficients \( \hat{P}_{lm} \) for the GW sky directly using estimated angular power in pixels \( \hat{P}_\ell \) by

\[
\hat{P}_{lm} = (Y^\dagger \cdot \Gamma \cdot Y)^{-1} \cdot (Y^\dagger \cdot \Gamma \cdot \hat{P}_\ell),
\]

(7.26)

where \( Y = Y_{lm,p} \) is the spherical harmonic basis matrix. Noise in the Fourier domain can be computed as [355]

\[
\hat{N}_l = \frac{1}{1 + 2l} \sum_m \left| \sum_{pp'} Y_{lm,p} \Gamma_{pp'} Y_{lm,l'}^* \right|^2.
\]

(7.27)

Analogous to the approach in CMB experiments, we construct unbiased estimators of the squared angular power \( \hat{C}_l \) in the spherical harmonic basis by

\[
\hat{C}_l = \frac{1}{1 + 2l} \sum_m |\hat{P}_{lm}|^2 - \hat{N}_l.
\]

(7.28)

Assuming a spectral index \( \alpha \), our maximum-likelihood estimates \( \hat{P}_\eta \) of the GW angular power spectrum yield an estimate of the normalized GW energy density \( \Omega_{GW} \) at a reference frequency \( f_{\text{ref}} \), integrated over a broad band of frequencies. The normalized GW energy density at the reference frequency \( f_{\text{ref}} \) is calculated using the noise-weighted monopole value \( \hat{P}_{00} \) of the GW power across the sky estimated from the maps by Eq. (7.26),

\[
\Omega_\alpha = \Omega_{GW}(f_{\text{ref}}) = \frac{2\pi^2}{3H_0^2} f_{\text{ref}}^3 \hat{P}_{00}.
\]

(7.29)

Note the computation of \( \hat{P}_{00} \) includes a normalization by a factor of \( 5/(8\pi) \) due to the normalization of detector overlap functions defined in Eq. (6.22) [295]. The
GW energy density spectrum at arbitrary frequencies is then obtained by re-scaling
the frequency-integrated estimate of $\Omega_{GW}$ with its spectral shape,

$$\Omega_{GW}(f) = \Omega \left( \frac{f}{f_{\text{ref}}} \right)^\alpha .$$

(7.30)

### 7.2 Spectral Shape: Model-independent Approach

In Eq. (7.1), we assume the GW power on the sky can be factored into separate
directional and frequency components, and we further assume the spectral shape
is a power law of index $\alpha$ as in Eq. (7.2). Though these two simplifications are
motivated by many astrophysical and cosmological models [96, 262], they are not
exact and will eventually break down.

There are SGWBs with non power-law spectral shapes, e.g., see Figs. 6.1 and 6.2.
In low frequencies of Fig. 6.2, the SGWB due to CBCs is well modeled by a power
law of spectral index 2/3. However, in high frequencies, we expect a spectral
turnover determined by the redshift-dependent star formation rate and the average
total mass of BBHs [372, 373]. Measuring this turnover will thus allow us to probe
the average BBH total mass, the evolution of that quantity over cosmic time, and the
star formation history of the Universe. Moreover, there may even be backgrounds
with direction-dependent spectral emission, which the spectral-model search is not
optimal for.

A generic, spectral-model-independent approach thus allows us to probe the spectral
shape of the SGWB and potentially identify contributing sources and mechanisms.
Towards building a general, model-agnostic search for SGWBs, a first step is to
reduce the assumption of spectral shapes to a minimum while maintaining the GW
strain power factorization of Eq. (7.1).

### Adaptive Frequency Banding

To reconstruct the spectral dependence of a SGWB, we run map-making in distinct
frequency bands of adaptively chosen bandwidths. The number of bands is a
user input to the pipeline, which ideally is numerous enough to achieve a good
approximation of the spectral shape. Nevertheless, it competes with the conditioning
of the Fisher matrix in each band. Each band needs to be wide enough for the Fisher
matrix to be adequately well-behaved so as to allow inversion. With the number of
bands as input, the algorithm chooses frequency endpoints with each band containing
equal amount of noise-weighted strain power. Within each band, we then assume
a fixed, least-informative prior flat in energy density, $H_\alpha(f) \sim f^{-3}$, to run map-
making. The optimal condition number of the Fisher matrix for each band would ideally be determined independently using the method presented in Section 7.1 with a monopole simulation in that band. However, running a simulation for each band is computationally expensive. At the current sensitivity level, we choose to use the broadband optimal condition number in Table 7.1 for each spectral shape as a proxy. With the assumption of an angular-independent spectral shape and with the Fisher matrix properly conditioned, estimated GW energy densities in each frequency band trace out the strain power spectral dependence.

**Adaptive Pixelization**

In the spectral-model-independent method, a single angular resolution does not accommodate all frequency bands due to different diffraction limits estimated via Eq. (7.25) [362]. Fixing an angular resolution across all bands over-resolves lower frequencies and hence impairs the Fisher matrix conditioning, and under-resolves higher frequencies and hence loses attainable SNRs. We therefore independently estimate the expected angular resolution for each frequency band using Eq. (7.25), with $D = D_{\text{HL}}$ and $f$ to be the midpoint of the band [362]. We limit ourselves to the optimal resolutions within the HEALPix package [366] (i.e., choices of resolution $N_{\text{side}} = 2^n$), implying a coarse resolution variation over frequency bands, as described in Section 7.3. We choose the pixel resolution such that point sources are over-resolved: specifically, spread across four pixels. We leave the pixelization optimization as well as the exploration of alternative pixelization schemes which allow for finer resolution variations for future work.

**7.3 Simulations**

We demonstrate the maximum-likelihood mapping method in the pixel domain outlined in Section 7.1 in the spectral-model-dependent case by running our pipeline to recover various injected maps for power-law models of spectral indices 0, 2/3 and 3. We also illustrate the spectral-model-independent approach to probe spectral dependence as described in Section 7.2 via simulations using a realistic spectral shape from the population studies of GWTC-3 [26]. We use the present sensitivity from the HLV detectors to construct the simulations, released publicly in [374]. For all spectral models, we inject loud angular power distributions of monopoles, Gaussian random fields, and random point sources on the sky. The simulated input strain power is $h^2 \sim O(10^{-45})$ for the map-making verification, while we use $h^2 \sim O(10^{-40})$ for the spectral-model-independent method testing. These may be
considered very high SNR cases at present sensitivity, as may be observed in the SNR maps presented in Fig. 7.5.

The simulated data consist of sequential CSD frequency segments corresponding to time segments of $\tau = 96$ s over a sidereal day, same as the LIGO–Virgo folded dataset format [374], where changes in ORFs are negligible and the noise within each segment stays constant to a good approximation. We simulate CSDs for all three baselines in the HLV detector network in [20, 1726] Hz in both the spectral-model-dependent and independent cases. We then run the complete analysis pipeline to compute maximum-likelihood map solutions for comparisons with injected maps.

We generate simulated CSD time series via

$$C_{\text{sim}} = K \cdot P_{\text{inj}} \implies C_{\text{sim}}(t; f) = \sum_\Theta K(t; f, \Theta) P_{\text{inj}}(\Theta).$$

We also add simulated Gaussian noise to the CSD time series in Eq. (7.31). To verify the pipeline implementation, we use Eqs. (7.16) – (7.18) to compute $\hat{\mathcal{P}}$.

When computations are implemented correctly and the Fisher matrix is properly well-conditioned, the pipeline recovers injected maps,

$$\hat{\mathcal{P}} = (K^\dagger \cdot N^{-1} \cdot K)^{-1} \cdot (K^\dagger \cdot N^{-1} \cdot (K \cdot P_{\text{inj}}))$$

$$= K^{-1} \cdot N \cdot (K^\dagger)^{-1} \cdot K^\dagger \cdot N^{-1} \cdot K \cdot P_{\text{inj}} = P_{\text{inj}}.$$  

With expected SGWBs and associated spectral indices in mind, we demonstrate the map-making functionality in three cases: a monopole map with $\alpha = 0$; a Gaussian random field with $\alpha = 2/3$ and an maximum resolution $\ell_{\text{max}} = 8$; and a map of 30 random point sources with $\alpha = 3$. We run our searches in the frequency range of [20, 1726] Hz, similarly to the LVK broadband searches in O3 [293]. We choose a pixel basis of 3072 pixels, or equivalently $N_{\text{side}} = 16$ in the HEALPix scheme, with each pixel covering 13.4 deg$^2$.

In Fig. 7.5, we show input maps, reconstructed clean maps, SNR maps, normalized residual maps and residual histograms for all three cases. All injected maps are successfully recovered, with minimal residual maps. We have also verified all combinations of injected maps and spectral indices not shown in Fig. 7.5. Note that our mapping method in the pixel domain successfully recovers both extended sources as in the case of the monopole and Gaussian maps as well as the map with 30 random point sources, although with some caveats. The “point” sources are generated in the same resolution as the recovery map, meaning that each “point”
Figure 7.5: Top to bottom: input, clean, SNR, normalized residual maps and residual histograms from simulations described in Section 7.3. From left to right: monopole ($\alpha = 0$), Gaussian random field ($\alpha = 2/3$), 30 random points ($\alpha = 3$). For all simulations, the pixel with the maximum residual is at the level of a few percent of the injected signal. We have verified that the residuals are Gaussian distributed with norm 1. In the monopole reconstruction, the SNR map presents a characteristic horizontal band due to the shape the ORF traces on the sky over 1 day; this may be also noticed in Fig. 8.4. In the Gaussian field case, the injected map has patches of zero power, and is thus more subject to poor estimation due to noise fluctuations than the other cases shown. In the case of 30 random points, the SNR map presents a residual of the point-spread function with negative values as it is the result of the matrix operation in Eq. 7.22, which can give rise to negative fluctuations where the pixel power is very low.
Figure 7.6: Example of recovery of 30 point sources for $\alpha = 0$. The smearing of the point spread is not completely deconvolved due to the singularity of the Fisher matrix in this case. The negative power in $\hat{P}$ is due to noise fluctuations.

here spans 13.4 deg$^2$. For a more realistic check, point source simulations need to be generated at a finer resolution and recovered by coarser graining. Also, for $\alpha = 3$, the Fisher matrix is well-conditioned hence we do not need to apply conditioning in its inversion. Without the information loss, point sources for $\alpha = 3$ are well recovered whereas for other spectral indices the recoveries manifest leakage and loss of resolution. An example of point source recovery for $\alpha = 0$ is shown in Fig. 7.6, illustrating the “smearing” of the point source recovery. Limited by computational resources, we defer work on improving the resolution on the clean map to future work.

For spectral-model-independent narrowband searches, we show the pipeline’s capability to probe spectral shapes using injected maps of a monopole, a Gaussian random field and 30 random point sources. The pipeline runs map-making in 30 adaptively chosen frequency bands in the search range of [20, 1726] Hz for each simulation. In each band, we set $f$ to be the midpoint frequency, and adaptively produce sky maps of 192, 768, 3072 pixels, or equivalently $N_{\text{side}} = 4, 8, 16$ in the HEALPix scheme, with each pixel covering 214.9 deg$^2$, 53.7 deg$^2$, 13.4 deg$^2$ respectively. These choices allow for good regularization of the Fisher matrix, and allow us to aptly over-resolve anisotropies according to the diffraction limit. We plot the reconstructed spectral shapes and energy densities in each case, along with the target model in Fig. 7.7. In the first two cases, map monopoles in different bands collectively trace out the expected spectral shape. The recovery of the spectrum is harder in the case of random point sources: we find that the monopole is not well recovered at lower frequencies, while the spectrum emerges in the higher frequency bands. As may be observed in Fig. 7.8, the recovered maps at lower frequencies do not resolve the point sources, causing GW power leakage. More on this sort of effect is explained in [375]. This may also be due to a sub-optimal conditioning of the Fisher matrix, which can be explored by repeating simulations as described in
Figure 7.7: Reconstructed normalized energy densities and spectral dependencies in 30 adaptive frequency bands for maps of a monopole (blue), a Gaussian random field (orange), and 30 random point sources (green) for a non-power-law spectral shape. On the left: $H(f)$; on the right: $\Omega(f)$, as defined in Eqs. (6.15) and (7.1). The spectral shapes are well reconstructed for maps of extended sources (note that the recoveries almost overlap in the plots), whereas for point sources reconstruction is imperfect in particular in the lower frequency bands. See the text for details.

Figure 7.8: Left to right are input map for the spectral-model-independent run in 30 adaptive frequency bands and reconstructed clean maps for bands [20, 93.46875] Hz and [1092.71875, 1726] Hz.

Section 7.1 in each individual frequency band.
Chapter 8

UPPER LIMITS ON ANISOTROPIC BACKGROUNDS THROUGH LIGO–VIRGO’S FIRST THREE OBSERVING RUNS

We apply the methods outlined above to real data obtained by the LIGO and Virgo GW detectors. Our results clearly show no evidence for a signal, in agreement with the LVK results [293], hence we set upper limits on anisotropies as well as the isotropic monopole as a limiting case using the maximum-likelihood mapping method in the pixel domain described in Section 7.1. We also set constraints on the spectral shape of the SGWB using the spectral-model-independent method described in Section 7.2.

For the analyses, we use the publicly available folded datasets of the first three observing runs of Advanced LIGO and Advanced Virgo [374]. The strain time series is Fourier transformed and cross-correlated between each available pair of detectors in the network at the time of observing. The cross-correlated data from each pair are then folded over one sidereal day [363, 364], reducing the computation time for anisotropic searches by a factor of the number of total observing days. This makes the processing of stochastic searches feasible in any modern-day personal computer. For O1 and O2, cross-correlated data only exist for the HL baseline, while for O3, data from all three combinations, HL, HV and LV, are available. Each sidereal-day folded dataset is chunked into 898 segments, with each segment lasting $\tau = 96$ s.

We perform all our analyses in the frequency range between 20 and 1726 Hz at a resolution of 1/32 Hz, although 99% of sensitivity for isotropic broadband analyses comes from the frequency band between 20 and 300 Hz [292]. This is because, depending on the spectral shape of the signal and the regularity of the Fisher matrix, anisotropic searches are not limited by the same sky-integrated sensitivity as isotropic searches.

8.1 Spectral-model-dependent, Broadband Limits

For the spectral-model-dependent, broadband searches, we present the results using three spectral indices, $\alpha = 0, 2/3$ and 3, same as the LVK searches [293]. The entire range of frequencies is integrated into a single map for each case of $\alpha$. Combining
Figure 8.1: Clean maps, SNR maps, $\Omega_{GW}$ upper limit maps, noise maps and SNR histograms for results integrated over a broadband of frequencies between 20 and 1726 Hz at a reference frequency of $f_{\text{ref}} = 25$ Hz using data from LIGO–Virgo’s first three observing runs. From left to right are for spectral indices 0, 2/3 and 3.
We calculate the normalized GW energy density at a reference frequency of $f_{\text{ref}} = 25$ Hz for each spectral index and find these are consistent with 0, hence we set frequentist 95% confidence level upper limits summarized in Table 8.1. Our upper limits are consistent with the LVK isotropic search results [292]. From the SNR maps in Fig. 8.1 we find the maximum SNR values across the sky, reported in Table 8.1. These are well below a significant deviation from 0. To confirm this, we calculate p-values from the distributions of the SNR maps; these are also reported in Table 8.1. We thus conclude that we find no evidence of GW signals in either the monopole or anisotropies. Note that the SNR maps are Gaussian distributed with norms less than 1: the same behaviour is observed in the LVK collaboration work [293], and stems from the fact that the maps have been regularized. The p-values calculated here include this re-normalization.

We also show the upper limits on the angular power spectrum $C_\ell$’s of the SGWB obtained via Eq. (7.28) in Fig. 8.2. These are approximately consistent with the LVK anisotropic search results [293], given that regularisation is performed very differently, hence the spread over $\ell$ modes appears different in the two upper limits. Our choice of the maximum $\ell$ mode included here is dictated by our pixel resolution, jointly with the expected angular resolution of this style of search discussed in [293]. The relation between $\ell$ mode and number of pixels necessary to resolve it, expressed in terms of the HEALPix $N_{\text{side}}$ parameter, is roughly $\ell_{\text{max}} \sim 2N_{\text{side}}$. This would suggest going up to an $\ell_{\text{max}} = 32$ for our analysis. However, even in the most sensitive scenario ($\alpha = 3$), according to [293] we expect resolutions higher than

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>Max SNR (% p-value)</th>
<th>95% upper limit on $\Omega_\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.9 (6)</td>
<td>$7.3 \times 10^{-9}$</td>
</tr>
<tr>
<td>2/3</td>
<td>2.2 (3)</td>
<td>$5.1 \times 10^{-9}$</td>
</tr>
<tr>
<td>3</td>
<td>2.7 (1)</td>
<td>$5.1 \times 10^{-10}$</td>
</tr>
</tbody>
</table>

Table 8.1: Maximum pixel SNRs of the reconstructed broadband clean maps and 95% confidence level upper limits on the normalized GW energy density at a reference frequency of $f_{\text{ref}} = 25$ Hz for the HLV network using data from the first three observing runs.

O1, O2 and O3 analyses, we show the reconstructed clean maps computed via Eq. (7.16), SNR maps via Eq. (7.22) and noise maps via Eq. (7.23) in Fig. 8.1 for each spectral index $\alpha = 0, 2/3$ and 3 from left to right respectively. The condition number chosen for each index is listed in Table 7.1.

We calculate the normalized GW energy density at a reference frequency of $f_{\text{ref}} = 25$ Hz for each spectral index and find these are consistent with 0, hence we set frequentist 95% confidence level upper limits summarized in Table 8.1. Our upper limits are consistent with the LVK isotropic search results [292]. From the SNR maps in Fig. 8.1 we find the maximum SNR values across the sky, reported in Table 8.1. These are well below a significant deviation from 0. To confirm this, we calculate p-values from the distributions of the SNR maps; these are also reported in Table 8.1. We thus conclude that we find no evidence of GW signals in either the monopole or anisotropies. Note that the SNR maps are Gaussian distributed with norms less than 1: the same behaviour is observed in the LVK collaboration work [293], and stems from the fact that the maps have been regularized. The p-values calculated here include this re-normalization.

We also show the upper limits on the angular power spectrum $C_\ell$’s of the SGWB obtained via Eq. (7.28) in Fig. 8.2. These are approximately consistent with the LVK anisotropic search results [293], given that regularisation is performed very differently, hence the spread over $\ell$ modes appears different in the two upper limits. Our choice of the maximum $\ell$ mode included here is dictated by our pixel resolution, jointly with the expected angular resolution of this style of search discussed in [293]. The relation between $\ell$ mode and number of pixels necessary to resolve it, expressed in terms of the HEALPix $N_{\text{side}}$ parameter, is roughly $\ell_{\text{max}} \sim 2N_{\text{side}}$. This would suggest going up to an $\ell_{\text{max}} = 32$ for our analysis. However, even in the most sensitive scenario ($\alpha = 3$), according to [293] we expect resolutions higher than
Figure 8.2: 95% upper limits on the angular power spectrum $C_\ell$ of the SGWB for power laws of $\alpha = 0$, $2/3$ and $3$ at a reference frequency $f_{\text{ref}} = 25$ using data from LIGO–Virgo’s first three observing runs. We have noted the outlier for $\ell = 6$ in the $\alpha = 3$ case: this is currently under investigation and is believed to be due to a noise fluctuation which makes the point value of $C_\ell$ negative.

$\ell_{\text{max}} > 16$ to be unattainable, due to the shape of the LIGO and Virgo noise curves. Hence, we select $\ell_{\text{max}} = 16$.

8.2 Spectral-model-independent, Narrowband Limits

Using the spectral-model-independent method described in Section 7.2, we first divide the search range between 20 and 1726 Hz into 10 and 20 frequency bands with adaptively chosen endpoints. Since O3 achieves the best sensitivity out of the three observing runs and HL is the most sensitive out of the three baselines, the frequency endpoints are chosen such that each band contains the same amount of noise weighted strain power of the O3 HL data. O1 HL, O2 HL, O3 HV and O3 LV analyses then employ the same frequency banding as O3 HL.

The mapping method described in Section 7.1 is run on each band separately and the resulting upper limits on the GW energy density are plotted in Fig. 8.3. The energy densities in different bands collectively probe the spectral shape of the SGWB. The spectral shapes obtained in our analyses are consistent with a noise-dominated estimate with increasing power as a function of frequency, resembling the detector noise curve. We detail the frequency endpoints, angular resolution and upper limit in each band for the 10-band case in Table 8.2 and the 20-band case in Table 8.3.
Figure 8.3: 95% upper limits on the energy densities in distinct frequency bands used in the spectral-model-independent spectral analysis. We show our results in 10 and 20 bands. The method assumes a scale invariant spectral shape in each spectral band. The results are consistent with noise dominated estimates.

<table>
<thead>
<tr>
<th>Frequency band [Hz]</th>
<th>$N_{\text{pix}}$</th>
<th>$N_{\text{pix kept}}$</th>
<th>$A_{\text{pix}}$ [deg$^2$]</th>
<th>$\Omega_0$ upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>[20, 133.125)</td>
<td>192</td>
<td>192</td>
<td>214.9</td>
<td>$4.5 \times 10^{-10}$</td>
</tr>
<tr>
<td>[133.125, 180.6875)</td>
<td>192</td>
<td>192</td>
<td>214.9</td>
<td>$1.5 \times 10^{-8}$</td>
</tr>
<tr>
<td>[180.6875, 225.0)</td>
<td>768</td>
<td>767</td>
<td>53.7</td>
<td>$5.0 \times 10^{-7}$</td>
</tr>
<tr>
<td>[225.0, 270.5)</td>
<td>768</td>
<td>768</td>
<td>53.7</td>
<td>$6.4 \times 10^{-7}$</td>
</tr>
<tr>
<td>[270.5, 324.21875)</td>
<td>768</td>
<td>768</td>
<td>53.7</td>
<td>$3.0 \times 10^{-6}$</td>
</tr>
<tr>
<td>[324.21875, 381.28125)</td>
<td>768</td>
<td>768</td>
<td>53.7</td>
<td>$5.4 \times 10^{-6}$</td>
</tr>
<tr>
<td>[381.28125, 454.9375)</td>
<td>3072</td>
<td>2512</td>
<td>13.4</td>
<td>$4.5 \times 10^{-5}$</td>
</tr>
<tr>
<td>[454.9375, 578.0)</td>
<td>3072</td>
<td>3042</td>
<td>13.4</td>
<td>$1.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>[578.0, 765.8125)</td>
<td>3072</td>
<td>3072</td>
<td>13.4</td>
<td>$3.5 \times 10^{-4}$</td>
</tr>
<tr>
<td>[765.8125, 1726]</td>
<td>3072</td>
<td>3072</td>
<td>13.4</td>
<td>$5.4 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

Table 8.2: Model-independent search results of each band for 10 adaptively chosen frequency bands, including the number of pixels determined based on the diffraction limit, the number of pixels passing the condition number threshold $10^{-5.5}$ and the 95% upper limit on GW energy density. The results are for the combined O1+O2+O3 data.

We also show the clean maps, SNR maps and noise maps for three narrowband analyses in low, mid and high frequencies of the 10-band case in Fig. 8.4. The lowest frequency band is between 20 and 133.125 Hz; the mid band is between 270.5 and 324.21875 Hz; and the highest band is between 765.8125 and 1726 Hz. We note that the changes in the scale of structures are evident as frequencies increase and our method chooses the resolution of each band accordingly as demonstrated in Fig. 8.4.
<table>
<thead>
<tr>
<th>Frequency band [Hz]</th>
<th>$N_{\text{pix}}$</th>
<th>$N_{\text{pix kept}}$</th>
<th>$A_{\text{pix}}$ [deg$^2$]</th>
<th>$\Omega_0$ upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>[20, 104.65625]</td>
<td>192</td>
<td>192</td>
<td>214.9</td>
<td>$4.6 \times 10^{-10}$</td>
</tr>
<tr>
<td>[104.65625, 133.125)</td>
<td>192</td>
<td>192</td>
<td>214.9</td>
<td>$8.9 \times 10^{-9}$</td>
</tr>
<tr>
<td>[133.125, 157.375]</td>
<td>192</td>
<td>192</td>
<td>214.9</td>
<td>$1.8 \times 10^{-8}$</td>
</tr>
<tr>
<td>[157.375, 180.6875]</td>
<td>192</td>
<td>192</td>
<td>214.9</td>
<td>$4.2 \times 10^{-8}$</td>
</tr>
<tr>
<td>[180.6875, 203.09375]</td>
<td>768</td>
<td>701</td>
<td>53.7</td>
<td>$5.7 \times 10^{-7}$</td>
</tr>
<tr>
<td>[203.09375, 225.0]</td>
<td>768</td>
<td>739</td>
<td>53.7</td>
<td>$7.1 \times 10^{-7}$</td>
</tr>
<tr>
<td>[225.0, 247.21875]</td>
<td>768</td>
<td>751</td>
<td>53.7</td>
<td>$9.6 \times 10^{-7}$</td>
</tr>
<tr>
<td>[247.21875, 270.5]</td>
<td>768</td>
<td>753</td>
<td>53.7</td>
<td>$6.1 \times 10^{-7}$</td>
</tr>
<tr>
<td>[270.5, 296.5625]</td>
<td>768</td>
<td>767</td>
<td>53.7</td>
<td>$3.6 \times 10^{-6}$</td>
</tr>
<tr>
<td>[296.5625, 324.21875]</td>
<td>768</td>
<td>768</td>
<td>53.7</td>
<td>$3.9 \times 10^{-6}$</td>
</tr>
<tr>
<td>[324.21875, 351.28125]</td>
<td>768</td>
<td>763</td>
<td>53.7</td>
<td>$4.7 \times 10^{-6}$</td>
</tr>
<tr>
<td>[351.28125, 381.28125]</td>
<td>3072</td>
<td>1734</td>
<td>13.4</td>
<td>$6.7 \times 10^{-5}$</td>
</tr>
<tr>
<td>[381.28125, 415.53125]</td>
<td>3072</td>
<td>1963</td>
<td>13.4</td>
<td>$4.9 \times 10^{-5}$</td>
</tr>
<tr>
<td>[415.53125, 454.9375]</td>
<td>3072</td>
<td>2113</td>
<td>13.4</td>
<td>$8.6 \times 10^{-5}$</td>
</tr>
<tr>
<td>[454.9375, 514.5]</td>
<td>3072</td>
<td>2621</td>
<td>13.4</td>
<td>$1.1 \times 10^{-4}$</td>
</tr>
<tr>
<td>[514.5, 578.0]</td>
<td>3072</td>
<td>2887</td>
<td>13.4</td>
<td>$1.7 \times 10^{-4}$</td>
</tr>
<tr>
<td>[578.0, 656.21875]</td>
<td>3072</td>
<td>3072</td>
<td>13.4</td>
<td>$4.8 \times 10^{-4}$</td>
</tr>
<tr>
<td>[656.21875, 765.8125]</td>
<td>3072</td>
<td>3072</td>
<td>13.4</td>
<td>$3.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>[765.8125, 964.03125]</td>
<td>3072</td>
<td>3072</td>
<td>13.4</td>
<td>$6.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>[964.03125, 1726]</td>
<td>3072</td>
<td>3072</td>
<td>13.4</td>
<td>$1.1 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

Table 8.3: Model-independent search results of each band for 20 adaptively chosen frequency bands, including the number of pixels determined based on the diffraction limit, the number of pixels passing the condition number threshold $10^{-5.5}$ and the 95% upper limit on GW energy density. The results are for the combined O1+O2+O3 data.

### 8.3 Discussion and Future Work

In this work presented from Chapter 7 through Chapter 8, we have developed a maximum-likelihood mapping method in the pixel domain for the SGWB power on the sky, complimentary to the methods of the LVK collaboration [293]. In SGWB mapping, Fisher matrix regularization has long been an active area of research. We have presented an empirical method, albeit preliminary, to systematically regularize the Fisher matrix in mapping deconvolution via monopole simulations. In addition to modeled searches, we have introduced an improved spectral-model-independent, narrowband search method to probe the spectral shape of the SGWB, with adaptive frequency banding and adaptive pixelization techniques applied to each band. We have shown that this is a valid method to probe spectral shapes of anisotropic backgrounds, and may serve as a first step to characterize these signals which may
Figure 8.4: Clean maps, SNR maps and noise maps for three frequency bands representative of low, mid and high frequencies in the spectral-model-independent search. In this analysis of the combined O1+O2+O3 data, the search range is divided into 10 bands. From left to right, the plots shown are for frequencies between 20 and 133.125 Hz, between 270.5 and 324.21875 Hz and between 765.8125 and 1726 Hz, and $N_{\text{pix}} = 192, 768, 3072$ respectively.

then inspire a targeted search with a more refined model.

We have verified both the modeled and the unmodeled methods in various simulations and we apply both to LIGO–Virgo’s folded datasets from the first three observing runs. In the spectral-model-dependent, broadband searches, we do not find any excess signals on top of the detector noise. In the spectral-model-independent, narrowband searches, our obtained spectral shapes are consistent with noise dominated estimates. Our results are in agreement with what is found by the LVK, as summarized in Table 8.1.

In future work, we will improve the reliability of the Fisher matrix regularization method when applied to narrowband searches. The method is sub-optimal in
narrow bands as the Fisher matrix conditioning will be band-dependent. This is particularly evident in our simulations of point sources, which are very sensitive to Fisher regularization (see Figs. 7.7 and 7.8). Ideally, the condition number for each band is independently determined, while in this study we have used the broad-band condition numbers as an alternative. Furthermore, we can explore setting constraints on different parametric models of the SGWB spectral shape starting from our spectral-model-independent results. Finally, the ultimate goal of the spectral-model-independent method is to extend its capability to search for angular-dependent, frequency-dependent (most general) backgrounds.

In expectation of a first detection of SGWBs in the coming observing runs, we also plan to use the pipeline to probe interesting questions. For example, we aim to assess whether we should expect to detect the isotropic or anisotropic component of the SGWB first, assuming different observing scenarios and signal characteristics.
ADVANCED LIGO DETECTOR COMMISSIONING FOR LIGO–VIRGO’S THIRD OBSERVING RUN

In this chapter, I summarize my work at the LIGO Livingston Observatory in the summer of 2018. The LIGO Livingston interferometer is shown in Fig. 9.1. In Section 9.1, we explain the method and the results of the implementation of a real-time Kalman filter in the LIGO Thermal Compensation System. In Section 9.2, we discuss the calibration of cavity lengths for several subsystems of the LIGO Livingston detector. In addition to what is documented here, I was also involved in miscellaneous hands-on detector work, mainly on measurements, re-balancing, and re-tuning of interferometer parts.

Figure 9.1: The Advanced LIGO interferometer in Livingston, LA, USA [376]. Credit: Caltech / MIT / LIGO Lab.

9.1 Implementation of a Real-time Kalman Filter in the LIGO Thermal Compensation System

Advanced LIGO operates high power lasers and will store up to 750kW optical power in its Fabry-Perot arm cavities at design sensitivity [1]. Absorption of laser power in the test masses produces temperature gradients across the entire volume of the test masses and forms thermal lenses. The test masses are 40 kg right circular
Figure 9.2: A schematic of the Advanced LIGO Thermal Compensation System. The Hartmann Wavefront Sensor measures the test mass thermal aberrations. The radiative ring heater and the CO₂ laser add heat to minimize thermal gradients in the test masses. The ring heater is installed around all test masses while the CO₂ laser actuates only on the compensation plates of the input test masses [1].

cylinders of 34 cm diameter and 20 cm thickness made of fused silica and supported by four very thin silica fibers. The reflective surfaces are made of alternating layers of silicon-dioxide and tantalum pentoxide. In operation, the main laser beam passes through the bulk of the input test masses (ITMs), suffering otherwise-uncontrolled thermal lensing. Optical aberrations of the test masses caused by thermal lenses impair the sensitivity of the LIGO detectors and thus need to be compensated for.

The Thermal Compensation System (TCS), consisting of a circular ring heater (RH) actuator, a CO₂ laser projector, and a Hartmann Wavefront Sensor (HWS), is designed to mitigate, or control, optical aberrations by reducing temperature gradients in the test masses [377]. To apply distortion corrections, the TCS adds heat to create conjugate thermal lenses to the ones formed by the main beam heating. The HWS first registers optical aberrations. The RH then heats the outer edge of the corresponding test mass. For residual ITM deformation, the CO₂ laser projects annular or central shaped beams onto the compensation plate (CP) of the ITM. A schematic of the TCS is shown in Fig. 9.2.

The measurements from the HWS suffer from significant levels of noise. To provide better optical aberration estimates to be corrected by the TCS, we implement a real-time Kalman filter (KF) [378] for the noisy HWS measurements. The KF takes in three inputs: HWS aberration measurements, control input power from RHs,
CO₂ lasers and the main laser beam, and a state-space aberration model. Such implementation is motivated because the KF aberration estimates are more precise than the HWS measurements while being more accurate than the aberration model. Besides test mass distortion corrections, the KF might also yield unexpected results in the residuals of subtracting the KF estimates from the HWS measurements, such as identification of unknown contributions to aberrations from sources other than laser self-heating, RHs, and CO₂ lasers.

**Kalman Filter**

Kalman filtering, named after Rudolph E. Kálmán, is an often-used tool of stochastic estimation from noisy measurements using a simple form of feedback control [378]. The recursive algorithm deals with discrete-data linear systems, where multi-dimensional observables are linearly related to linearly evolving multi-dimensional system states with additive Gaussian noise.

At each time step, the KF first provides the *a priori* estimates of the current state of a system using the previous state, then updates its prediction with the new measurement from the current time step. It is an optimal predictor-corrector estimator in that it minimizes the *a posteriori* state error covariance when certain criteria are met. Even though conditions for optimality rarely occur, the technique and its extended variants perform well due to their simplicity and robustness.

Kalman filtering is ideal for low-latency, real-time processing problems because only information of the previous state of a system is needed to proceed instead of the full history. Common applications of the KF include motion tracking, autonomous navigation, and vehicle control. It has also been employed in GW searches, specifically in the estimation of the time-dependent violin modes of test mass suspensions using LIGO detector outputs [379, 380].

Mathematically, the KF formulation assumes a discrete-time process described by the linear stochastic difference equation,

\[ X_k = A_k X_{k-1} + B_k U_k + s_k, \]  

(9.1)

where \( X_k \) is the current state with covariance \( P_k \), \( A_k \) is the state-transition matrix applied to the previous state \( X_{k-1} \), \( B_k \) is the input-control matrix applied to the control signal \( U_k \), and \( s_k \) is the zero-mean Gaussian process noise with covariance \( Q_k \), i.e., \( s_k \sim N(0, Q_k) \).
The measurements are related to the system states by
\[ Z_k = C_k X_k + v_k, \] (9.2)
where \( Z_k \) is the current measurement, \( C_k \) is the observation model that maps the state space into the measurement space, and \( v_k \) is the zero-mean Gaussian measurement noise with covariance \( R_k \), i.e., \( v_k \sim N(0, R_k) \).

In general, the matrices \( A_k, B_k, C_k \) are time-dependent. In our application, we model these matrices all to be constant because the TCS is a slowly varying system that they are constant in the time scale of the KF. Thus going forward, we drop the subscripts and simply use \( A, B, C \).

In the prediction phase, at time step \( k \), we propagate the previous state \( X_{k-1} \) and the associated error covariance \( P_{k-1} \) ahead using quantities from the previous state,
\[ X_k^- = AX_{k-1} + BU_k, \] (9.3)
\[ P_k^- = AP_{k-1}A^T + Q_k. \] (9.4)

The residual between the current measurement \( Z_k \) and the \textit{a priori} measurement space estimate \( CX_k^- \) is then
\[ Y_k = Z_k - CX_k^-, \] (9.5)
with the covariance \( O_k \) of the residual being
\[ O_k = CP_k^-C^T + R_k. \] (9.6)

Using Eqs. (9.4) and (9.6), we define the optimal Kalman gain at time step \( k \) to be
\[ K_k = P_k^-C^TO_k^{-1}, \] (9.7)
which represents the weight given to the measurements and the current state estimate.

In the correction phase, we update the \textit{a priori} estimate Eqs. (9.3) and (9.4) using the measurement residual Eq. (9.5) and the Kalman gain Eq. (9.7) to obtain the \textit{a posteriori} estimate,
\[ X_k = X_k^- + K_k Y_k, \] (9.8)
\[ P_k = (I - K_k C)P_k^-, \] (9.9)
where \( I \) is the identity matrix.

The complete KF process is summarized in Fig. 9.3.
**Prediction:**

1. Propagate the state ahead:
   \[ X_k^- = AX_{k-1} + BU_k \]
2. Propagate the error covariance ahead:
   \[ P_k^- = AP_{k-1}A^T + Q_k. \]

**Correction:**

1. Compute the Kalman gain:
   \[ K_k = P_k^- C^T (CP_k^- C^T + R_k)^{-1} \]
2. Update the state with the measurement:
   \[ X_k = X_k^- + K_k (Z_k - CX_k^-) \]
3. Update the state error covariance:
   \[ P_k = (I - K_k C) P_k^- \]

Initial estimates for \(X_{k-1}\) and \(P_{k-1}\)

Figure 9.3: The recursive Kalman filter (KF) algorithm, initialized with estimates from the previous time step, consists of the prediction and the correction phases. The KF first provides the *a priori* estimate based on the previous state in the prediction phase, then updates its estimate using the current measurement in the *a posteriori* phase.

**Thermo-optical Aberration State-space Model**

A number of analytical models of the temperature fields and thermo-optical aberrations in the test masses have been developed [381, 382]. However, state-space representations of these models are too complicated to derive or are computationally expensive. To simplify the computation, the model presented here uses only one optical aberration parameter, the spherical power. Spherical power, denoted by \(S\), is the spherical component of the optical aberration \(\psi\), and is inversely proportional to the radius of curvature. The unit of spherical power is diopters (\(m^{-1}\)) [383].

In our implementation, we model the spherical power as the sum of two exponentials [384],

\[
S(t) = TV(t) \left[ \mu_1 (1 - e^{-t/\tau_1}) + \mu_2 (1 - e^{-t/\tau_2}) \right], \quad (9.10)
\]

where \(T\) is the transmission coefficient of the test masses, \(V(t)\) is the input power, \(\mu_1\) and \(\mu_2\) are scaling factors of unit \(m^{-1}W^{-1}\), \(\tau_1\) and \(\tau_2\) are time constants. We also make two important assumptions about the spherical power model. The transmission factor is high, i.e., \(T \to 1\). And for a unit input power, the saturation value when \(t \to \infty\) is 1, i.e., \(\lim_{t \to \infty} S(t) = \mu_1 + \mu_2 = 1\).

To find the state space representation of the spherical power, we first compute the unit step response by applying Laplace transform to \(S(t)\),

\[
S(s) = \mathcal{L}\left\{ \mu_1 (1 - e^{-t/\tau_1}) + \mu_2 (1 - e^{-t/\tau_2}) \right\} = \frac{sN_2 + N_1}{s^3 + s^2 D_2 + sD_1}, \quad (9.11)
\]
where \( D_1, D_2, N_1, N_2 \) are given by
\[
D_1 = \frac{1}{\tau_1 \tau_2}, \quad D_2 = \frac{1}{\tau_1} + \frac{1}{\tau_2}, \quad N_1 = \frac{\mu_1 + \mu_2}{\tau_1 \tau_2}, \quad N_2 = \frac{\mu_1}{\tau_1} + \frac{\mu_2}{\tau_2}.
\] (9.12)

The transfer function of the system is then
\[
H(s) = \frac{S(s)}{V(s)} = s \cdot S(s) = \frac{sN_2 + N_1}{s^2 + sD_2 + D_1}.
\] (9.13)

By applying the inverse Laplace transform to the transfer function
\[
\mathcal{L}^{-1}\{(s^2 + sD_2 + D_1)S(s) = (sN_2 + N_1)V(s)\},
\] (9.14)

we obtain the differential equation
\[
\ddot{S}(t) + D_2 \dot{S}(t) + D_1 S(t) = N_2 \dot{V}(t) + N_1 V(t),
\] (9.15)

where the over-dots denote time derivatives.

Finally, we obtain the state space representation that bears the form of Eq. (9.1),
\[
\begin{pmatrix}
S_k \\
\dot{S}_k
\end{pmatrix} = \begin{pmatrix}
1 & \Delta t \\
-D_1 \Delta t & 1 - D_2 \Delta t
\end{pmatrix} \begin{pmatrix}
S_{k-1} \\
\dot{S}_{k-1}
\end{pmatrix} + \begin{pmatrix}
0 & 0 \\
N_1 \Delta t & N_2 \Delta t
\end{pmatrix} \begin{pmatrix}
V_k \\
\dot{V}_k
\end{pmatrix},
\] (9.16)

where \( \Delta t \) is the sampling period of the KF.

In the notation of Eq. (9.1), the state-transition matrix \( A_k \) and the input-control matrix \( B_k \) are
\[
A_k = A = \begin{pmatrix}
1 & \Delta t \\
-D_1 \Delta t & 1 - D_2 \Delta t
\end{pmatrix}, \quad B_k = B = \begin{pmatrix}
0 & 0 \\
N_1 \Delta t & N_2 \Delta t
\end{pmatrix},
\] (9.17)

and the current state matrix \( X_k \) and the control signal \( U_k \) are
\[
X_k = \begin{pmatrix}
S_k \\
\dot{S}_k
\end{pmatrix}, \quad U_k = \begin{pmatrix}
V_k \\
\dot{V}_k
\end{pmatrix}.
\] (9.18)

The HWS measures the spherical power directly and thus the measurement model is
\[
Z_k = \begin{pmatrix}
1 & 0
\end{pmatrix} \begin{pmatrix}
S_k \\
\dot{S}_k
\end{pmatrix}.
\] (9.19)

In the notation of Eq. (9.2), the observation model \( C_k \) is
\[
C_k = C = \begin{pmatrix}
1 & 0
\end{pmatrix}.
\] (9.20)

The parameters in the state space representation are found by fitting the double exponential model Eq. (9.10) to the detector measurement data in [383] using the Least Absolute Residuals method, summarized in Table 9.1.
Table 9.1: Best fitted model parameters of the state space representations and of the analytical models for the spherical power of thermo-optical aberrations. The parameters were obtained by fitting a double exponential model to the detector measurement data using the Least Absolute Residuals method in [383]. For the goodness-of-fit statistics for all components, the sums of squares due to error are $O(10^{-6})$, the root mean squared errors are $O(10^{-4})$, and R-squared and adjusted R-squared are 1.

**Real-time Implementation of a Kalman Filter in the TCS**

With an aberration state space model and HWS measurements, we implement a real-time KF for better optical aberration estimates in the TCS. Since TCS is a slowly evolving system, we build a guardian node TCS_KAL coded in Python in the Guardian finite state machine [385] for the KF implementation instead of in a real “real-time” machine (for fast control with bandwidths 16 Hz or higher) coded in C.

We modify the L1TCS_SCS model with additional EPICS channels designated for TCS_KAL in each test mass:

```
L1:TCS-SIM_$TM$_SUB_DEFOCUS_KALMAN_APOST_PROC_STDDEV
L1:TCS-SIM_$TM$_SUB_DEFOCUS_KALMAN_APOST_MEAS_STDDEV
L1:TCS-SIM_$TM$_SUB_DEFOCUS_KALMAN_APOST_EST
L1:TCS-SIM_$TM$_SUB_DEFOCUS_KALMAN_APOST_RES
```

For a given aberration state space model, TCS_KAL reads in HWS measurements at 16 Hz, input powers of the main laser self-heating, RHs and CO$_2$ lasers, process noise and measurement noise from the witness EPICS channels. It outputs a posteriori KF estimates of optical aberrations in the test masses and the residuals between KF estimates and HWS measurements. All parameters of the state space model are
A Bayesian-weighted Outlier-robust Kalman Filter

The standard KF as described earlier is an optimal estimator for real-time linear dynamic systems. However, the standard KF is not robust against outliers as seen in Fig. 9.5. And this insufficiency will affect the performance of the KF implementation in the TCS, with ineffective aberration estimates around outliers and thus hinder further analysis on aberrations in the test masses.

We thus explore ideas of automatic outlier detection and removal to be implemented in the TCS Kalman filtering. We here outline a weighted least squares-like approach studied in details in [386]. The method introduces a Bayesian weight statistic to treat the weight of each data sample probabilistically. [386] deals with uncontrolled...
automatic systems, whereas in our application, we modify the method to include system controls.

Each data point $Z_k$ is attributed with a scalar weight $w_k$ such that the variance of $Z_k$ is weighted with $w_k$. The weights are modeled to be Gamma distributed random variables. The resulting priors are

\begin{align}
(Z_k | X_k, w_k) &\sim \text{Normal}(C X_k, R_k / w_k), \quad (9.21) \\
(X_k | X_{k-1}) &\sim \text{Normal}(A X_{k-1} + B U_k, Q_k), \quad (9.22) \\
w_k &\sim \text{Gamma}(a_{w_k}, b_{w_k}), \quad (9.23)
\end{align}

which are a generalization of Eqs. (9.1) and (9.2).

The problem is then treated as an Expectation-Minimization-like (EM) learning problem, where the Bayesian method maximizes the “incomplete” log likelihood $\log p(X_{1:N})$ with all other variables marginalized out. When assuming the model is time invariant, the final EM equations at time step $k$ are formulated as follows.

In the expectation phase, we first compute the weight $w_k$ of the data sample $Z_k$ at the current time step $k$ to be

\begin{equation}
\langle w_k \rangle = \frac{a_{w_k,0} + \frac{1}{2}}{b_{w_k,0} + \langle (Z_k - C X_k)^T R_k^{-1} (Z_k - C X_k) \rangle}, \quad (9.24)
\end{equation}

where $a_{w_k,0}$ and $b_{w_k,0}$ are prior scale parameters for the weight. $a_{w_k,0}$ and $b_{w_k,0}$ should be chosen such that $\langle w_k \rangle = 1$ with some confidence, which is to say most data samples are considered normal to start off.

Generalizing Eq. (9.9), we then compute the \textit{a posteriori} covariance $P_k$ of the state $X_k$ to be,

\begin{equation}
P_k = (\langle w_k \rangle C^T R_k^{-1} C + Q_k^{-1})^{-1}. \quad (9.25)
\end{equation}

Finally, generalizing Eq. (9.8), we calculate the \textit{a posteriori} mean $\langle X_k \rangle$ of the state $X_k$,

\begin{equation}
\langle X_k \rangle = P_k \left[ Q_k^{-1} (A \langle X_{k-1} \rangle + B U_k) + \langle w_k \rangle C^T R_k^{-1} Z \right]. \quad (9.26)
\end{equation}

In the minimization phase, the system’s process noise and measurement noise are updated.

The $m$th coefficient of the vector $R_k$, $r_{km}$, is given by

\begin{equation}
r_{km} = \frac{1}{k} \sum_{i=1}^{k} \langle w_i \rangle \left( (Z_{im} - C(m,:) X_i)^2 \right), \quad (9.27)
\end{equation}
where $C(m, :)$ is the $m$th row of the matrix $C$.

And the $n$th coefficient of the vector $Q_k$, $q_{kn}$, is defined as

$$ q_{kn} = \frac{1}{k} \sum_{i=1}^{k} \left( (X_{in} - A(n, :)X_{i-1} - B(n, :)U_{i-1})^2 \right), \quad (9.28) $$

where $A(n, :)$ is the $n$th row of the matrix $A$, and similarly for $B(n, :)$.

Since we deal with a streaming problem, where data samples arrive one at a time and not all past data are retrievable, we need to rewrite the minimization phase in incremental form:

$$ r_{km} = \frac{1}{k} \left[ \text{sum}_{km}^{w ZZ} - 2C(m, :)\left(\text{sum}_{km}^{w ZX}\right) 
+ \text{diag}\{C(m, :)\left(\text{sum}_{k}^{w XX^T}\right)C(m, :)^T\} \right], \quad (9.29) $$

$$ q_{kn} = \frac{1}{k} \left[ \text{sum}_{kn}^{X^2} - 2A(n, :)\left(\text{sum}_{kn}^{XX'}\right) - 2B(n, :)\left(\text{sum}_{kn}^{UU}\right) 
+ \text{diag}\{A(n, :)\left(\text{sum}_{k}^{XX'}\right)A(n, :)^T + 2A(n, :)\left(\text{sum}_{k}^{UU}\right)B(n, :)^T 
+ B(n, :)\left(\text{sum}_{k}^{UU}\right)B(n, :)^T\} \right]. \quad (9.30) $$

And the incremental sum statistics are:

$$ \text{sum}_{km}^{w ZZ} = \langle w_k \rangle Z_{km}^2 + \text{sum}_{k-1, m}^{w ZZ} \quad (9.31) $$

$$ \text{sum}_{km}^{w ZX} = \langle w_k \rangle Z_{km}X_k + \text{sum}_{k-1, m}^{w ZX} \quad (9.32) $$

$$ \text{sum}_{k}^{w XX^T} = \langle w_k \rangle \langle X_k X_k^T \rangle + \text{sum}_{k-1}^{w XX^T} \quad (9.33) $$

$$ \text{sum}_{kn}^{X^2} = \langle X_{kn}^2 \rangle + \text{sum}_{k-1, n}^{X^2} \quad (9.34) $$

$$ \text{sum}_{kn}^{XX'} = \langle X_{kn} \rangle \langle X_{k-1} \rangle^T + \text{sum}_{k-1, n}^{XX'} \quad (9.35) $$

$$ \text{sum}_{kn}^{UU} = \langle U_k \rangle^T + \text{sum}_{k-1, n}^{UU} \quad (9.36) $$

$$ \text{sum}_{k}^{XX'} = \langle X_{k-1} X_{k-1}^T \rangle + \text{sum}_{k-1}^{XX'} \quad (9.37) $$

$$ \text{sum}_{k}^{UU} = \langle U_k U_k^T \rangle + \text{sum}_{k-1}^{UU} \quad (9.38) $$

$$ \text{sum}_{k}^{X'U} = \langle X_k' U_k \rangle + \text{sum}_{k-1}^{X'U} \quad (9.39) $$

At the time of the project, the development of a robust KF in the TCS was still in early phase. One particular challenge was that the algorithm relied heavily on matrix manipulations in Python, which slowed down the process quite significantly compared to the standard KF. Also, robust KF is shown to be optimal in treating systems with time-variant state space models in [386], whereas the method could
Figure 9.5: Comparison of the effectiveness of the Kalman filter (top panel) and the Bayesian-weighted Kalman filter (bottom panel) on an example 6-hour stretch of archival HWS measurements. The uphill climbs are due to controlled heat added onto the test mass for test mass characterization and TCS testing. The slow downhills trend are due to natural cooling of the test mass, and the steep downhill trends are because of the TCS corrections. The Kalman filter is prone to outliers, while the Bayesian-weighted Kalman filter detects and treats outliers, although imperfectly. Not perfectly track and remove all outliers in our application. Furthermore, in our application, its implementation is very sensitive to some of the parameters.

Figure 9.5 compares the effectiveness of the robust KF and the standard KF on the same 6-hour stretch of archival HWS data set. Although not perfect, the robust KF is still an improvement compared to the standard KF in terms of outlier robustness. At the time of the project, we identified online code testing, performance boost of the robust KF, and exploration of other aberration state space models as future work [387].

9.2 Calibration of Subsystem Cavity Lengths

We calibrate the cavity lengths of four LIGO Livingston interferometer subsystems: Input Mode Cleaner (IMC), Michelson (MICH), Power Recycling Cavity (PRC), and Signal Recycling Cavity (SRC) [1] as shown in Fig. 9.6.
The triangular IMC cleans the spatial profile, polarization, jitter, and frequency noise of the prestabilized laser beam from the input port. MICH defines the pair of optical paths between the beam splitter and the ITMs. The PRC and SRC are lengths between power- and signal-recycling mirrors, the beam splitter, and the ITMs, of which the PRC increases the effective laser power while the SRC maintains a broad detector frequency response [1].

To lock the light through these paths in a resonant condition, we optimize the cavity lengths by control systems. Before we can achieve such control, accurate representations of cavity lengths must be established. We therefore utilize three channels to compute cavity lengths as well as displacements of intermediate mass stages and make sure they all return the quantities in interest accurately.

The first channel is the theoretical computations through suspension models and actuator signals of each mass stage from the Length Sensing and Control subsystem. For each of IMC, MICH, PRC, and SRC, contributions sum up linearly from each stage of suspension to arrive at final cavity lengths. The second one is the online calibration signals directly measuring cavity lengths. The third channel is by taking GS-13 seismometer measurements at suspension points and propagating down to final mass stages to obtain cavity lengths.
In principle, all three calculations should match well for each cavity length. However, during the pre-O3 commissioning of the Livingston detector, some cavity length computations of these four subsystems mismatched. We thus characterized the spectral performance of related filters and channels, and engaged updated filter functions subsequently in preparation of LIGO–Virgo’s O3. We refer readers to [389–391] for more details.
Chapter 10

SUMMARY AND OUTLOOK

10.1 Summary of the Previous Chapters

In 2015, the Advanced LIGO detectors [1] made the first direct detection of GWs from the merger of two massive black holes [2], GW150914, ushering the beginning of GW astronomy. In 2017, the Advanced LIGO and Advanced Virgo [3] detectors observed the first GW signals from two colliding neutron stars [4], GW170817. The subsequent extensive observing campaign across the electromagnetic spectrum marked a new era of multi-messenger astronomy [5]. Ever since then, until the end of LIGO–Virgo’s third observing run, the LIGO Scientific, VIRGO and KAGRA collaboration has cataloged dozens more GW signals from compact binary mergers, bringing the total number of GW events to 90 [6].

GWs carry signatures of their progenitor sources and generation mechanisms. They provide unique strong-field tests of gravity [25] and a unique way to probe the physics of matter at densities unattainable on Earth [94]. They open a window to study the most exotic astrophysical objects and the most violent events in the Universe. They help us learn about the evolution of stars and their populations [26]. They even offer a peek inside the earliest moments of the Big Bang and present a standard siren way to measure the expansion history of the Universe [29].

GWs can come from a variety of astrophysical sources, including coalescing compact binaries, continuous waves, burst sources and stochastic backgrounds. This thesis focuses on developing the detection methods for two out of the four mentioned: compact binary coalescences and stochastic gravitational-wave backgrounds.

We started with a brief history of the hunt for GWs in Chapter 1. In Chapter 2, to lay out some background knowledge, we briefly described GWs in terms of its formulation in GR and its main astrophysical sources. Leading into the detection aspect of GWs, we also introduced the Advanced LIGO detectors for measuring the stretching and squeezing of spacetime and the scientific values of such endeavors in fundamental physics, astrophysics and cosmology.

In Chapter 3, we expanded on compact binary coalescences, with a focus on their formation channels and gravitational waveforms. In Chapter 4, we followed through
the entire PyCBC search pipeline for GW signals from compact binary coalescences deeply buried in GW detector data. We also recounted the development efforts including improving the multi-detector coincidence checks and dealing with loud instrumental glitches towards the third observing run in PyCBC and in PyCBC Live. In Chapter 5, we compiled the PyCBC search results along with results from other pipelines in LIGO–Virgo’s first three observing runs into GW transient catalogs and characterized several notable events.

In Chapter 6, we presented the foundational knowledge on stochastic gravitational-wave backgrounds: statistical properties, sources, detection methods and the current most stringent upper limits on the stochastic backgrounds. In Chapter 7, we built a maximum-likelihood mapping pipeline in the pixel domain for the GW power on the sky, complimentary to the LVK search methods. To obtain a “good” deconvolution, regularization of the Fisher matrix must be addressed. We presented a novel method to empirically determine the optimal cutoff eigenvalue for its singular value decomposition via monopole simulations. In addition to the modeled maximum-likelihood map-making, we also implemented a model-independent method to probe the spectral shape of stochastic backgrounds. We verified both of our methods with various simulations. We applied our methods to the combined O1, O2 and O3 LIGO–Virgo data and detailed the search results in Chapter 8.

In Chapter 9, we went over my detector commissioning work at the LIGO Livingston Observatory in summer 2018, preceding the third observing run. Specifically, we developed a Kalman filter to provide better optical aberration estimates to be corrected by the Thermal Compensation System. We also characterized and calibrated the cavity lengths of several detector subsystems.

### 10.2 Outlook

For PyCBC developments towards the fourth observing run, we want to expand the search parameter space by including precessing waveforms. To improve the search sensitivity, we work on including data quality products in the ranking statistic. We will also incorporate a new method to compute the FAR for single-detector triggers [211]. For PyCBC Live, infrastructure and configuration improvements are underway. Low-latency early warnings for BNS mergers and \( p_{\text{astro}} \) functionalities will be also new additions.

For SGWB mapping developments, we want to expand the Fisher matrix regularization method to narrowband searches. We also want to extend the model-independent
method to search for angular-dependent, frequency-dependent (most general) backgrounds. In expectation of a first detection of SGWBs in the upcoming observing runs [26, 69, 307, 308], various interesting questions can be probed with the pipeline. For example, we want to simulate signals with anisotropies on top of an isotropic background and test whether the isotropic search or the anisotropic search is more sensitive.

GW astronomy is still in its infancy. Over the seven years after our first detection in 2015, we have already seen a number of exciting breakthroughs such as detections of exceptional events and joint GW-EM observations. The prospects are even brighter with the expanding global GW detector network [76]. During O1 and most of O2, the Advanced LIGO detectors were the only two observing GW detectors. At the end of O2 (August 2017), Advanced Virgo joined and contributed to the confirmation and localization of detections [220]. KAGRA joined the network in O3 and is performing upgrades to reach comparable sensitivity with the other detectors in the network [392]. A future LIGO in India is under construction and is expected to be operational by 2025 [78]. With further improvement in sensitivities of current detectors, expansion of the GW detector network and early warnings for pre-merger BNS detections [393–395], future observing runs offer great prospects for GW astronomy and multi-messenger astronomy, broadening our understanding of various aspects of GW science. Routine detections of BBHs forming a large population will allow us to probe the major formation channel for BBHs and distinguish between separate channels. Multi-messenger observations further facilitate the study of matter at extreme densities and provide a distinct way to measure the Hubble constant. Detections of GW backgrounds can give us insight into the physical mechanisms in the early Universe. There can also be surprises from burst sources or unexpected sources.
BIBLIOGRAPHY


[27] B. F. Schutz. “Determining the Hubble constant from gravitational wave observations”. In: Nature 323.6086 (Sept. 1986), pp. 310–311. ISSN: 1476-4687. DOI: 10.1038/323310a0. URL: https://doi.org/10.1038/323310a0.


[38] Imagine the universe! URL: https://imagine.gsfc.nasa.gov/observatories/satellite/lisa/.


[93] S. Vitale et al. “Use of Gravitational Waves to Probe the Formation Channels of Compact Binaries”. In: *Classical and Quantum Gravity* 34.3 (Jan. 2017), 03LT01. DOI: 10.1088/1361-6382/aa552e. URL: https://doi.org/10.1088/1361-6382/aa552e.

[94] L. Baiotti. “Gravitational Waves from Neutron Star Mergers and Their Relation to the Nuclear Equation of State”. In: *Progress in Particle and Nuclear


[125] Marchant, P. et al. “A new route towards merging massive black holes”. In: Astronomy and Astrophysics 588 (2016), A50. DOI: 10.1051/0004-6361/201628133. URL: https://doi.org/10.1051/0004-6361/201628133.


[130] M. Moe and R. D. Stefano. “Mind Your Ps and Qs: The Interrelation between Period (P) and Mass-ratio (Q) Distributions of Binary Stars”. In: The Astrophysical Journal Supplement Series 230.2 (June 2017), p. 15. DOI:


[362] L. Xiao, A. I. Renzini, and A. J. Weinstein. “Model-independent search for anisotropies in stochastic gravitational-wave backgrounds and application to LIGO-Virgo’s first three observing Runs”. In: (2023). L.X. participated in the conception of the project, solved and analyzed equations, prepared and analyzed data, and wrote the manuscript. DOI: 10.48550/ARXIV.2211.10010. URL: https://arxiv.org/abs/2211.10010.


