Detection and characterization of gravitational-wave bursts with multi-detector networks

Yi Shuen Christine Lee

ORCID: 0000-0002-8738-3299

May, 2025

A thesis submitted in complete fulfillment
of the requirements for the degree of

Doctor of Philosophy
Physics
The University of Melbourne

In loving memory of my grandfather, my aunt and my uncle.

Abstract

Transient sources of gravitational waves (GWs), lasting up to a few seconds, can be categorized into well-modeled compact binary mergers and minimally-modeled GW bursts. Possible sources of GW bursts include known phenomena, such as core-collapse supernovae (CCSNe) and gamma-ray bursts (GRBs), along with unclassified astrophysical events. In the advanced detector era, the LIGO-Virgo-KAGRA (LVK) collaboration has detected over 100 GW transients, and the number continues to grow at an average rate of one detection per day in the ongoing fourth observing run. Despite the large number of detections, no minimally-modeled GW bursts have been detected to date; compact binary mergers between black holes and neutron stars remain the only confirmed sources of GWs. Due to the complex and unpredictable nature of GW bursts, their detection and characterization require methods that make minimal assumptions about the signal source and morphology. The *BayesWave* algorithm is a well-established Bayesian burst analysis method within the LVK community. In anticipation of minimally-modeled GW burst detections as the global detector network expands, this thesis studies the burst detection and characterization performance of *BayesWave* with multi-detector networks.

Instrumental glitches are transient, non-Gaussian power spikes in the detector data, of non-astrophysical origins. Much like GW bursts, glitches typically last a few seconds and can occur across a broad range of frequencies. For this reason, glitches can potentially mask or mimic burst signals, resulting in false alarms. *BayesWave* employs a Bayesian framework to jointly characterize GW signals and glitches. The detection statistic is given by the log Bayes factor between the coherent signal model and the incoherent glitch model, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$; positive values of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ indicate a preference for the signal model, and vice versa. A previous study shows that $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ of astrophysical events increases as the detector network expand, but the study does not account for false alarms caused by glitches in real observing scenarios. In this thesis, the impact of glitches on the multi-detector burst detection performance of *BayesWave* is assessed, using data from the first half of the third observing run (O3a). *BayesWave*'s performance with the LIGO Hanford-Livingston (HL; two-detector) network is compared to that of the HL-Virgo (HLV; three-detector) network in terms of detection efficiency, measured using receiver operating characteristic (ROC) curves. Measurements of the O3a noise background show that the false alarm probability is, on average, an order of magnitude lower in HL compared to HLV. As a result, an analysis using phenomenological waveforms of binary black holes (BBH) as representative GW burst sources yields similar ROC curves with HL and HLV, indicating comparable overall detection efficiency, despite the higher $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ observed with HLV. A separate analysis finds that 18 GWs resembling compact binary mergers detected during O3 are also detected with comparable astrophysical significance in HL and HLV. Altogether, the study concludes that *BayesWave*'s burst detection performance does not improve with larger detector networks at O3 sensitivity, as the increased glitch rate undermines the advantage of increased $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$.

Our understanding of CCSNe explosion mechanisms and their GW signatures remains incomplete. Hydrodynamical instabilities, such as the standing accretion shock instability (SASI) and neutrino-driven convection, are theorized to be the key explosion drivers, and simulations suggest that they emit low-frequency ($\lesssim 250\,\mathrm{Hz}$) GWs. Therefore, the detection of low-frequency GWs, or lack thereof, is useful for constraining CCSNe explosion mechanisms. The *dedicated-frequency framework* developed in this thesis aims to enhance the detection of frequency-specific signatures, by performing band-passed *BayesWave* follow-up analyses on GW burst candidates detected by the complementary algorithm, coherent WaveBurst (cWB). The primary goal is to study whether low-frequency (LF) follow-up, limited to $\leq 256\,\mathrm{Hz}$, can detect GW signatures associated with CCSNe explosion mechanisms, under O3 observing conditions. The dataset consists of GWs extracted from five CCSN simulations, each exhibiting varying strengths of low-frequency emissions attributed to the SASI and neutrino-driven convection. The study finds that CCSNe with stronger low-frequency GW emissions are more readily detected by the LF follow-up analysis with *BayesWave*, and successful LF detections are useful for constraining CCSN explosion mechanisms. Another application of the dedicated-frequency framework is to enhance detection significance. An illustrative study is conducted on the loudest trigger during SN 2019fcn. Since the trigger exhibits minimal power below 256 Hz, it is followed up using a high-frequency (HF) *BayesWave* analysis, limited to $\geq 256\,\mathrm{Hz}$. The HF analysis reduces the trigger's false-alarm rate by approximately a factor of four compared to the initial full-band analysis.

According to Einstein's theory of general relativity, GWs consist of two tensor polarization modes, namely plus (+) and cross (×). *BayesWave* offers two tensor-polarized signal models: (i) the elliptical polarization model (E), where the + and × polarizations are exactly 90° out of phase, and (ii) the relaxed polarization model (R), which does not have this restriction. As the global detector network expands, GW polarizations can be measured with increasing accuracy. In light of future detector advancements, a multi-detector network analysis of *BayesWave*'s performance in characterizing elliptical and non-elliptical GW burst polarizations is presented, using design sensitivity curves of future upgrades of the LIGO, Virgo and KAGRA detectors. Non-precessing and precessing BBHs are used to represent elliptical and non-elliptical GW signals, respectively. The performance of E relative to R is evaluated by plotting the log Bayes factor $\ln \mathcal{B}_{R,E}$ as a function of the overlap $\mathcal{O}_{R,E}$, where $\mathcal{O}_{R,E} = 1$ indicates full agreement between E and R reconstructions, and $\mathcal{O}_{R,E} = 0$ indicates no agreement. For non-precessing or minimally-precessing signals, we find $\mathcal{O}_{R,E} > 0.98$ and $\ln \mathcal{B}_{R,E} < 0$, as E is preferred over R for its simplicity. In contrast, for strongly precessing signals, R provides more accurate signal characterization than E, yielding $\mathcal{O}_{R,E} < 0.98$ and $\ln \mathcal{B}_{R,E} > 0$. As the detector network expands, $\mathcal{O}_{R,E}$ decreases for strongly precessing signals, and $\ln \mathcal{B}_{R,E}$ becomes increasingly positive. The analysis is repeated for O3 BBH events, with all events yielding $\ln \mathcal{B}_{R,E} \leq 0$, except GW200129_065458 which is known to have anomalous properties, possibly precession or a coincident glitch. These findings suggest that $\ln \mathcal{B}_{R,E}$, together with $\mathcal{O}_{R,E}$, can indicate deviations from the strict elliptical polarization assumed by E. The non-precessing and precessing BBHs are also used to demonstrate how R measures the polarization content of GW bursts in terms of Stokes parameters. The measurement accuracy improves with larger detector networks, and for sufficiently large networks (three or more detectors), the accuracy is not affected by signal morphology.

Declaration

This page certifies that:

- This thesis contains only original work towards a Doctor of Philosophy, except where indicated in the preface.
- Due acknowledgement has been made in the text to all other material used.
- This thesis is fewer than 100 000 words in length, exclusive of tables, figures, bibliographies, and appendices.

Yi Shuen (Christine) Lee

Preface

This thesis is an original work of the author (Y. S. C. Lee) reporting research done alone or in collaboration with other authors. This section provides a chapter-by-chapter summary of the author's contributions and the publication status of all material.

Chapter 1 is a comprehensive literature review written by the author for this thesis. It is an original work of the author, with editing from A. Melatos and M. Millhouse, and has not been submitted for publication. All figures and tables are the work of the author, unless stated otherwise in the captions.

Chapter 2 is a comprehensive review of the methods and applications of the *BayesWave* algorithm [1, 2], which is the main focus of this thesis. This chapter contains reproduced and adapted materials from *Y. S. C. Lee, M. Millhouse, and A. Melatos*, Physical Review D 103, 062002 (2021) [3] and *Y. S. C. Lee, M. Millhouse, and A. Melatos*, Physical Review D 103, 082002 (2024) [4], both of which were written primarily by the author, with scientific input and editing from the co-authors. Reproduction and adaptations to the original texts are summarized as follows:

- Sections 2.1 is partially reproduced from Section II.A of [4].
- Section 2.2.1 is partially reproduced from Section II.A of [4].
- Sections 2.2.2-2.2.4 are reproduced verbatim from Sections II.B.1-II.B.3 of [4].
- Section 2.4.1 is partially reproduced from Section II.C of [4].
- Section 2.4.2 is adapted with substantial rewording from Section IV.A of [3].
- Section 2.5 is adapted with substantial rewording from Section IV.B, Section VI.A, Appendix B and Appendix C of [3].

The calculations, methods, and figures in Sections 2.4.2 and 2.5 originally appeared in the thesis submitted for the degree of Master of Science (Physics) at The University of Melbourne, awarded to the author in 2020. The adapted content is included in this thesis for completeness, and where relevant, the respective sources are acknowledged through in-text footnotes. Sections not listed above are original works of the author, with editing from A. Melatos and M. Millhouse, and have not been submitted for publication. All figures are the work of the author.

Chapter 3 is published as *Y. S. C. Lee, M. Millhouse, and A. Melatos*, Physical Review D 103, 082002 (2024) [4]. This work was written primarily by the author, with scientific input and editing from the co-authors. The coherent WaveBurst (cWB) team produced the background trigger list used in this study; the names of responsible team members are listed in the Acknowledgements section at the end of this chapter. Section II of Ref. [4] is excluded from this chapter, as it has been incorporated into Chapter 2 of this thesis. All figures and tables are the work of the author.

Chapter 4 is submitted to Physical Review D as *Y. S. C. Lee, M. J. Szczepańczyk, T. Mishra, M. Millhouse, and A. Melatos* (2025)[1]. This work was written primarily by the author, with scientific input and editing from the co-authors. M. J. Szczepańczyk collated and analyzed the GW waveforms for the core-collapse supernova models, and performed all necessary cWB analyses for the study with methodological input from T. Mishra. This chapter also includes Appendix 4.A1, which is an original work of the author, with editing from A. Melatos, and has not been submitted for publication. All figures and tables are the work of the author.

Chapter 5 is published as *Y. S. C. Lee, S. Doshi, M. Millhouse, and A. Melatos*, Physical Review D 111, 082002 (2025) [6]. This work was written primarily by the author, with scientific input and editing from the co-authors. S. Doshi conducted the analysis for the results shown in Table 5.1 and Figure 5.5. This chapter also includes Appendices 5.A3 and 5.A4, which are original works of the author, with editing from A. Melatos, and have not been submitted for publication. All figures and tables are the work of the author.

Chapter 6 summarizes the work in Chapters 3–5. It includes some exploratory future directions for the work in this thesis. It was written by the author, with editing from A. Melatos. All figures and tables are the work of the author.

The author contributed to two other publications during their PhD candidature, which are not included for examination in this thesis, but are listed below for completeness:

- *A. Mandlik et al.*, Monthly Notices of the Royal Astronomical Society, 532, 2644 (2024) [7] describes the upgrades to the Molonglo Observatory Synthesis Telescope (MOST) and its performance in the UTMOST-NS pulsar timing program, as well as in searches for single-pulse radio transients (e.g. fast radio bursts). The author worked with Chris Flynn to develop the MOST flux density calibration pipeline.

- *L. Dunn et al.*, Monthly Notices of the Royal Astronomical Society, 541, 1792 (2025) [8] presents results from the UTMOST-NS pulsar timing program. The author worked with Chris Flynn to develop the MOST flux density calibration pipeline.

The author also led the development of an undergraduate computational physics activity that introduces Bayesian inference and its applications in gravitational-wave astronomy, delivered as part of the subject Laboratory and Computational Physics 3 (PHYC30021) at the University of Melbourne, and funded by the Laby Foundation. This work is published as *Y. S. C. Lee, J. K. H. Thong, and Y. Bu*, Proceedings of The Australian Conference on Science and Mathematics Education (2023) [9] during the author's PhD candidature. This work is not included for examination in this thesis.

The author acknowledges funding support from the Australian Research Council (ARC) Centre of Excellence for Gravitational Wave Discovery (OzGrav) (project numbers CE170100004 and CE230100016), and from The University of Melbourne (Women in Physics Award and Jean Laby Women in Physics Travel Award). The author is also supported by a Melbourne Research Scholarship.

[1]Chapter 4 was submitted to Physical Review D before this thesis was examined, and thus does not include minor revisions and the additional appendix recommended by the reviewer in the published article: *Y. S. C. Lee, M. J. Szczepańczyk, T. Mishra, M. Millhouse, and A. Melatos*, Phys. Rev. D 112, 082006 (2025) [5].

Acknowledgments

Ten years ago, I took a leap of faith, turning down medical school offers to pursue physics. What began as a simple quest to indulge my curiosity serendipitously led me to this thesis. Mum and Dad, thank you profusely for your unwavering support and the countless sacrifices you have made to make all of this possible.

My PhD journey got off to a rocky start due to unexpected COVID-19 lockdowns and the loss of two dear family members. To Andrew, my supervisor, thank you for your support and understanding during these challenging times. Thank you also for your guidance over the years; your experience and wisdom have not only influenced my growth as a scientist and broadened my perspective, but have also cultivated my English literacy in ways that no formal education ever could. To Meg, my co-supervisor halfway across the world, thank you for laying the groundwork for my scientific interests, for your invaluable insights, and for being present at all my milestones, despite the inconvenient time zones. To Lloyd, the chair of my supervisory panel, thank you for your support and thoughtful counsel throughout my PhD. To Rachel, thank you for encouraging me to pursue astrophysics during my undergraduate days and for your generous mentorship along the way. To Harry, thank you for your patience and support as Head of School, especially in the final stages of my submission. To Marek and Tanmaya, thank you for your expertise and efforts in bringing our supernovae project to fruition. And to my colleagues at the School of Physics, especially those in the NS/GW group, thank you for the stimulating discussions, the corridor conversations, and for being such wonderful company overall — it is truly a pleasure to know you and work alongside you all.

On a more personal note, to my partner Cody, thank you for seeing me through all the ups and downs, and for helping me stay grounded throughout this journey. To my aunts, Cody's family, and close friends, thank you for your care and support over the years, especially Niken and Josh, who also shared this decade-long journey at university with me. Last but not least, I am grateful for my cat Iris, who is part menace, part distraction, but without question, the most endearing source of emotional support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1. A brief history of transient astronomy

Excerpted from Rumi's poem 'This body is a rose': "Each form you see has its unseen archetype. If the form is transient, its essence is eternal". Indeed, astronomical transients, despite their brevity, reveal the most violent and energetic processes to ever exist, and leave significant imprints in our understanding of the universe. Transient astronomy has a rich history that spans centuries. One of the earliest documented astronomical transients is the unprecedented appearance of a 'guest star', potentially a supernova (SN 185), observed by Chinese astronomers in the 2nd century [10]. In the 11th century, events like SN 1006, the brightest supernova recorded to date [11], and SN 1054, the supernova associated with the infamous Crab Pulsar [12, 13], were observed and documented by various civilizations. Centuries later came Tycho Brahe's supernova SN 1572 and Kepler's SN 1604 [14], leading up to the first telescopic observation of extragalactic supernovae in 1885 (SN 1885) [15]. Despite various misinterpretations of the transient origins at the time of observation, documentations of these events provide invaluable insights to modern astronomy, e.g. the frequency of galactic supernovae [16], and the characterization of Type Ia supernovae [17].

Over the last century, transient astronomy has progressed significantly alongside technological advancements. The observation of transients is no longer limited to those that are nearby (galactic) or visible to the naked eye. With the aid of ground- and space-based telescopes, astronomers are now observing transients on a daily basis, across all wavelengths of the electromagnetic (EM) spectrum. In addition to supernovae (SNe), other examples of transient sources detected to date through EM observations include, but are not limited to, gamma-ray bursts (GRBs) [18], active galactic nuclei (AGNs) [19], tidal disruption events (TDEs) [20] and fast radio bursts (FRBs) [21]. Altogether, EM observations, transient and perpetual, have significantly improved our understanding of the universe from our very own solar system to the cosmic microwave background. Nevertheless, EM astronomy is limited by light-matter interaction, i.e. EM waves are readily absorbed and/or scattered by materials like dust and gas. This results in the loss of information, especially when propagating across large cosmic distances.

On 14th of September 2015, the Laser Interferometer Gravitational-wave Observatory (LIGO) detectors in the USA observed the first merger between two black holes (GW150914) [22, 23] through the perturbations in the fabric of spacetime, i.e. 'gravita-

tional waves' (GWs). GW150914 is the first direct detection of gravitational waves, and also an astrophysical transient without EM counterparts. GWs are a unique astrophysical messenger because it propagates unobstructed through spacetime at the speed of light, and carries information of the emitting source. That is, GWs offer a new observational window to the universe, enabling the study of obscured astrophysical sources and source interiors that were previously inaccessible through electromagnetic observations. There exists a diverse array of GW-emitting sources, encompassing both transient and persistent phenomena, but we restrict our attention to the former in this thesis. Much like EM transients, GW transients are also produced by cataclysmic sources. In fact, EM and GW emissions sometimes come hand in hand, e.g. the binary neutron star merger GW170817 [24, 25], thus advancing multi-messenger astronomy. Compact binary coalescences (CBCs) like GW150914 and GW170817 are the most common sources for GW transients, so they are well-understood and robustly modeled. Other potential GW transient sources include, but are not limited to, SNe and GRBs. However, these sources are often challenging to model due to their inherent complexity and unpredictability. Although CBCs remain the only transient GW sources detected to date, the detection of other transient sources is highly-anticipated by the GW community. In light of this prospect, this thesis focuses on the methods of detecting and characterizing GW transients with minimum assumptions.

The rest of this chapter serves as an overview of the theory and background relevant to the topics explored in this thesis. In Section 1.2, we expand on the discussion of GW transient sources, and summarize the search efforts and detections made to date. In Section 1.3, we discuss the geometrical origins and properties of GWs. We also derive the physically-motivated solution to demonstrate how GWs are generated by sources such as those discussed in Section 1.2. In Section 1.4 we discuss the current designs, operational status and future plans of ground-based GW detectors, along with the adverse impacts of noise transients (glitches) in practical observing scenarios. In Section 1.5 we overview the methods for detecting well-modeled and minimally-modeled GW transients. Finally, in Section 1.6, we outline the structure for the remainder of this thesis.

1.2. Gravitational-wave transients

1.2.1. Gravitational-wave transient sources

Figure 1.1 illustrates the classification of GW sources based on their signal duration and the extent to which they can be modeled. The left-hand column displays the two types of GW transients: CBCs, which are well-modeled, and bursts, which are poorly constrained or unmodeled. These sources emit signals ranging from a few milliseconds to several weeks, and are useful for studying the behavior of matter and the nature of gravity in extreme conditions.

CBCs are by far the most common type of transient GW sources [26]. As its name suggests, a CBC consists of two compact objects, e.g. black holes (BH) and neutron stars (NS), spiraling toward each other and ultimately merging. The systematic orbital decay of CBCs produce the characteristic inspiral-merger-ringdown GW waveform that can be accurately modeled using general relativity [27].

There is also a wide variety of predicted transient GW sources, including but not lim-

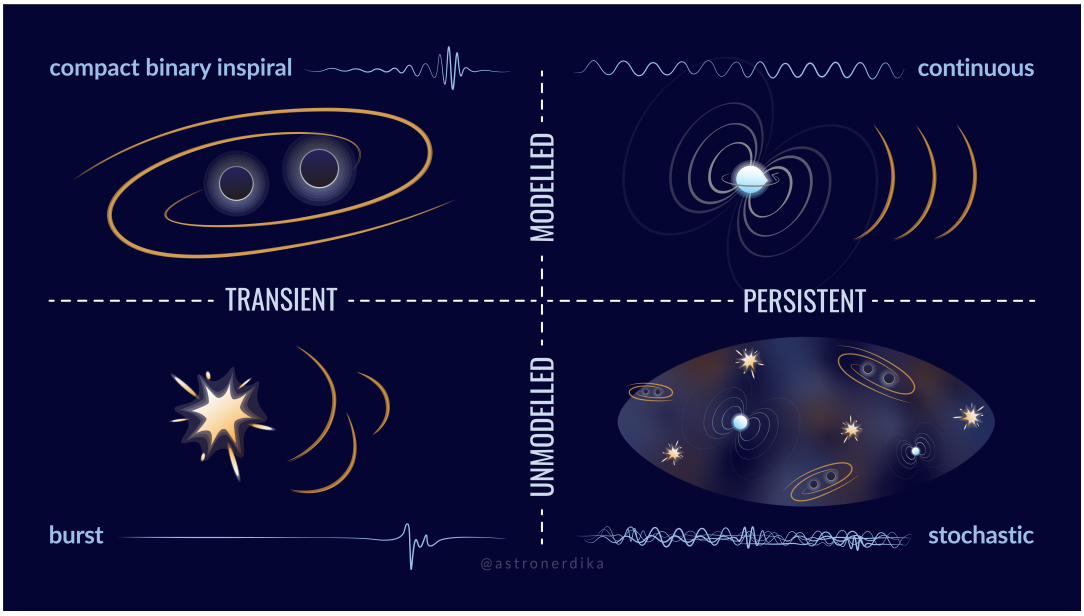

Figure 1.1: Types of GWs. The left-hand column shows transient GWs, with durations ranging from milliseconds to several weeks; their sources are discussed in Section 1.2.1. The right-hand column shows persistent GWs, with durations $\gtrsim 1$ yr; their sources are discussed in Section 1.2.3. [Graphics credit: Shanika Galaudage].

ited to: core-collapse supernovae (CCSNe) of massive stars [28], GRBs [29, 30], FRBs [31], binary neutron star (BNS) post-merger remnants [32], astrophysical glitches in isolated pulsars [33], and highly eccentric binary black holes (BBHs) [34]. However, the complex and non-deterministic nature of these sources makes it difficult to produce accurate models for the GW signals. These anticipated but poorly-constrained sources, along with any unclassified GW transients, are collectively referred to as (unmodeled) GW bursts.

1.2.2. Summary of gravitational-wave transient searches and detections

Searches for transient GW signals from well-modeled CBCs and generic burst sources require different techniques. The former relies on template-based matched filtering, while the latter requires minimally-modeled analyses which make few assumptions about the source origins. These techniques are discussed in further detail in Section 1.5. Here, we provide an overview of the efforts by the LIGO-Virgo-KAGRA (LVK) collaboration in searching for GW transients and the detections made to date.

The Advanced LIGO and Advanced Virgo detectors have completed three observing runs—O1, O2 and O3 —between 2015 and 2020. The CBC search results from these runs are presented in the Gravitational-Wave Transient Catalogs (GWTCs) [35–38]. A total of 93 CBCs have been detected by the end of O3, including the first detection GW159014. These detections predominantly consists of binary black hole (BBH) mergers, along with two confirmed binary neutron star (BNS) mergers and at least[1] two neutron star-black

[1]Detections such as GW190814 [39] confirm the presence of compact objects within the mass gap of $2.5M_\odot$–$5M_\odot$. The classification of the mass-gap object, and hence the type of CBC, remains ambiguous.

hole binary (NSBH) mergers. Every CBC detection refines our understanding of the underlying population of compact objects and the distribution of their properties, e.g. mass, spin, merger rates [40–42]. However, some events with exceptional properties offer additional astrophysical insights. We summarize these events and their implications in chronological order, from top to bottom, in Table 1.1. The last entry of Table 1.1, GW230529, is the first published detection from the ongoing fourth observing run O4 [43]. The number of detections in O4 has not been finalized, but the total number of observed GW events are expected to exceed 200 by the end of the observing run, currently scheduled for the 7th of October 2025.

The LVK collaboration has also conducted all-sky searches for GW bursts. These searches make minimal assumptions about the source properties, such as the source direction, the event epoch and the signal morphology. For each observing run (O1-O3), the searches for short-duration ($\lesssim 1$ s) and long-duration (2–500 s) bursts are carried out separately, using minimally-modeled burst analysis pipelines optimized for the respective duration. The short-duration searches spans a broader frequency range of $32\,\mathrm{Hz} \lesssim f \leq 4096\,\mathrm{Hz}$ [44–46], while the range of the long-duration searches are limited to $24\,\mathrm{Hz} \leq f \leq 2048\,\mathrm{Hz}$ [47–49]. None of the searches find statistically significant evidence for non-CBC GW bursts. However, they report the burst sensitivity of existing detectors by conducting searches using injected astrophysical and ad-hoc waveforms. The sensitivity of the most recent (O3) short-duration search peaks at 70 Hz. Using the 50% detection efficiency at false-alarm rate (FAR) 1 per 100 years as the detection benchmark, short-duration searches can detect galactic sources at 10 kpc that radiate energy as little as $\sim 10^{-10}\, M_\odot c^2$ [46]. The sensitivity of the long-duration search, on the other hand, peaks at ~150 Hz. With the same detection benchmark, but a less stringent FAR of 1 per 50 years, the O3 long-duration search can detect sources up to 10 kpc that emit a minimum energy of $\sim 10^{-9}\, M_\odot c^2$ [49]. For reference, the energy radiated by the BNS GW170817 is estimated to be $> 0.025\, M_\odot c^2$, and it was detected by the CBC searches with a FAR of 1 per 8×10^4 years, at a distance of ~40 Mpc [24].

1.2.3. Other gravitational-wave sources

The right-hand column of Figure 1.1 displays the two types of persistent GWs. In contrast to transients, persistent GW sources produce long-lasting signals that span last for years rather than seconds. Rapidly-rotating neutron stars with non-axisymmetric deformations are expected to produced quasi-monochromatic[2] and well-modeled continuous waves (CWs) signals [50, 51]. Persistent GWs can also arise from the superposition of GW signals from various cosmological and/or astrophysical processes across the universe, collectively known as the stochastic GW background [52]. Unlike CWs, this type of GW is poorly-modeled because its sources ranges from inherently stochastic (e.g. primordial GWs from inflation, cosmic string defects), to an assortment of deterministic but unresolved GW sources (e.g. supermassive BH binaries, stellar-mass CBCs, CCSNe) [53].

In the remainder of this thesis, we restrict our attention to short-duration GW transients lasting up to a few seconds. We do not consider long-duration ($\gtrsim 2$ s) transients or persistent GWs further.

[2] In this context, quasi-monochromatic refers to GWs whose frequency remains approximately constant over timescales of years.

LVK run	Event name	CBC type	Notes/Implications	Ref(s).
O1	GW150914	BBH	• First direct detection of GWs • Evidence for formation of heavy BHs ($\gtrsim 25M_\odot$) • First tests for general relativity and merger rate inference	[22, 23, 54–56]
O1	GW170817	BNS	• First BNS and first GW detection with EM counterparts	[24, 25]
O2	GW170104	BBH	• Evidence of black-hole spin misalignment with orbital angular momentum provide insights into BH formation channels	[57]
O2	GW170814	BBH	• First three-detector detection, enabling tests for GW polarizations	[58]
O3a	GW190412	BBH	• Asymmetric component masses ($30.1^{+4.6}_{-5.3}M_\odot$ and $8.3^{+1.6}_{-0.9}M_\odot$) • Evidence for higher-order multipoles	[59, 60]
O3a	GW190425	BNS	• Second BNS, but without EM counterpart • Heaviest observed BNS system	[61]
O3a	GW190521	BBH	• BBH with heaviest progenitor • Merger remnant is an intermediate-mass BH	[62]
O3a	GW190814	Ambiguous	• Merger between BH and a mass-gap object (heaviest NS or lightest BH) • Potentially a new class of CBC	[39]
O3b	GW200105 & GW200115	NSBH	• First two confirmed NSBH detections • Uneven component masses enable tidal disruption studies	[63]
O4a	GW230529	Ambiguous	• Merger between NS and a mass-gap object • Constrains mass-gap merger rate	[43]

Table 1.1: List of exceptional GW detections and their implications. The suffixes ‘a’ and ‘b’ after O3/O4 respectively indicate the first and second halves of the corresponding observing runs. The events above are selected for having dedicated LVK discovery papers.

1.3. Theory of gravitational-wave sources

We now discuss how GWs emerge from the framework of general relativity. First, in Section 1.3.1, we derive GWs from Einstein's field equations (EFEs) using the linearized theory. In Section 1.3.2, we discuss how the tensor GW polarizations arise as a solution to the EFEs in the linearized theory of general relativity. This provide the underlying theory for Chapter 5, where we explore the prospects of characterizing of tensor polarizations with GW bursts. In Section 1.3.2, we derive a physically motivated solution for GWs within the framework of linearized theory, which reveals their physical origins and lays the groundwork for the discussion of GW extraction from CCSNe in Chapter 4.

1.3.1. Linearized theory

The emergence of GW from general relativity finds its roots in the EFEs:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{1.1}$$

The left-hand side of Equation 1.1 comprises the space-time metric $g_{\mu\nu}$, plus its first and second derivatives through the Ricci tensor $R_{\mu\nu}$ and the Ricci scalar R. The right-hand side comprises the gravitational constant G, the speed of light c and the stress-energy tensor $T_{\mu\nu}$. Altogether, Equation 1.1 relates the local spacetime curvature, to the local density and flux of energy and momentum. It is challenging to obtain exact solutions to the EFEs due to their nonlinearity. Fortunately, however, GWs are produced by astrophysical sources (e.g. compact binary mergers and supernovae) where the gravitational fields are weak. Since the spacetime is flat in the absence of gravity, the metric for weak gravitational fields can be modeled as the Minkowski metric $\eta_{\mu\nu}$ plus a small perturbation term $h_{\mu\nu}$, i.e.

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \qquad \left|h_{\mu\nu}\right| \ll 1. \tag{1.2}$$

The weak-field solutions can then be approximated by expanding the EFEs to linear order of $h_{\mu\nu}$. We summarize the derivation below.

To linear order in $h_{\mu\nu}$, the Riemann curvature tensor is given by

$$R_{\alpha\mu\beta\nu} = \frac{1}{2}\left(\partial_\mu\partial_\beta h_{\alpha\nu} + \partial_\nu\partial_\alpha h_{\beta\mu} - \partial_\mu\partial_\nu h_{\alpha\beta} - \partial_\alpha\partial_\beta h_{\mu\nu}\right) \tag{1.3}$$

In the linearized theory, $\eta_{\mu\nu}$ is used to raise and lower indices. Thus the Ricci tensor, obtained by contracting the first and third indices, becomes

$$R_{\mu\nu} := R^{\beta}{}_{\mu\beta\nu} = \frac{1}{2}\left(\partial_\mu\partial^\beta h_{\beta\nu} + \partial_\nu\partial^\beta h_{\beta\mu} - \partial_\mu\partial_\nu h - \Box h_{\mu\nu}\right), \tag{1.4}$$

where h is the trace $h^{\mu}{}_{\mu} = \eta^{\mu\nu} h_{\mu\nu}$ and $\Box$ is the flat space D'Alembertian operator, i.e. $\partial^\mu\partial_\mu = \eta^{\mu\nu}\partial_\mu\partial_\nu$). Similarly, the Ricci scalar is given by

$$R = R^{\mu}{}_{\mu} = \partial^\sigma\partial^\gamma h_{\sigma\gamma} - \Box h. \tag{1.5}$$

Combining Equations 1.1, 1.4 and 1.5, and keeping only the linear order terms of $h_{\mu\nu}$, we obtain the set of linearized EFEs:

$$\partial_\mu\partial^\alpha h_{\alpha\nu} + \partial_\nu\partial^\alpha h_{\alpha\mu} - \partial_\mu\partial_\nu h - \Box h_{\mu\nu} - \eta_{\mu\nu}\left(\partial^\sigma\partial^\gamma h_{\sigma\gamma} - \Box h\right) = \frac{16\pi G}{c^4} T_{\mu\nu}. \tag{1.6}$$

Equation 1.6 can be simplified further using appropriate coordinate transformations. First, we show that linearity of Equation 1.2, and hence Equation 1.6, are preserved under the coordinate transformation $x^\mu \to x'^\mu = x^\mu + \xi^\mu$, if $\left|\partial_\mu \xi_\nu\right|$ is approximately the same order as $\left|h_{\mu\nu}\right| \ll 1$. Consequently, the metric as in Equation 1.2 transforms as [64]:

$$g_{\mu\nu} \to g'_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} - (\partial_\mu \xi_\nu + \partial_\nu \xi_\mu) + \mathcal{O}(h\partial\xi) + \mathcal{O}(\partial\xi\,\partial\xi). \tag{1.7}$$

Since h and $\partial\xi$ are small, the last two terms of $g'_{\mu\nu}$ in Equation 1.7 are negligible. Thus, the coordinate transform essentially resulted in the transformation of the metric perturbation:

$$h_{\mu\nu} \to h'_{\mu\nu} = h_{\mu\nu} - (\partial_\mu \xi_\nu + \partial_\nu \xi_\mu). \tag{1.8}$$

Under such transformation, the linearity of Equation 1.2 is preserved because $\left|\partial_\mu \xi_\nu\right| \ll 1$. Now we define the trace-reverse field

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h, \tag{1.9}$$

such that the inverse gives

$$h_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\bar{h} \tag{1.10}$$

with $h \equiv \eta^{\mu\nu} h_{\mu\nu} = -\bar{h}$. Since h is small, the trace-reverse field satisfies the transformation properties above. Substituting Equation 1.10 into Equation 1.6 results in the following simplification:

$$-\Box\bar{h}_{\mu\nu} + \partial_\mu \partial^\alpha \bar{h}_{\alpha\nu} + \partial_\nu \partial^\alpha \bar{h}_{\alpha\mu} - \eta_{\mu\nu}\partial^\sigma \partial^\gamma h_{\sigma\gamma} = \frac{16\pi G}{c^4} T_{\mu\nu}. \tag{1.11}$$

By gauge freedom, we can also choose the four gauge transform functions ξ_μ to satisfy the Lorentz gauge condition $\partial^\nu \bar{h}_{\mu\nu} = 0$. As a result, the last three terms of Equation 1.11 vanish, reducing it to a simple wave equation

$$\Box\bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu}. \tag{1.12}$$

1.3.2. Gravitational-wave polarizations

Since GWs propagate across large distances to reach the earth-bound detectors, it is more relevant to consider the far-field EFEs in vacuum, where $T_{\mu\nu} = 0$:

$$\Box\bar{h}_{\mu\nu} = 0. \tag{1.13}$$

Equation 1.13 implies that GWs in vacuum travel at the speed of light and follow a plane-wave solution. The solutions are derived as follows.

For the transformation $h_{\mu\nu} \to \bar{h}_{\mu\nu}$ as in Equation 1.8, gauge freedom allows us to choose any four ξ_μ that satisfy the Lorentz condition $\partial^\nu \bar{h}_{\mu\nu} = 0$ and thus Equation 1.13. For a GW propagating in the z-direction with angular frequency ω, there exists a space-like Lorentz gauge which reduces the plane-wave solution of the 4×4 perturbation

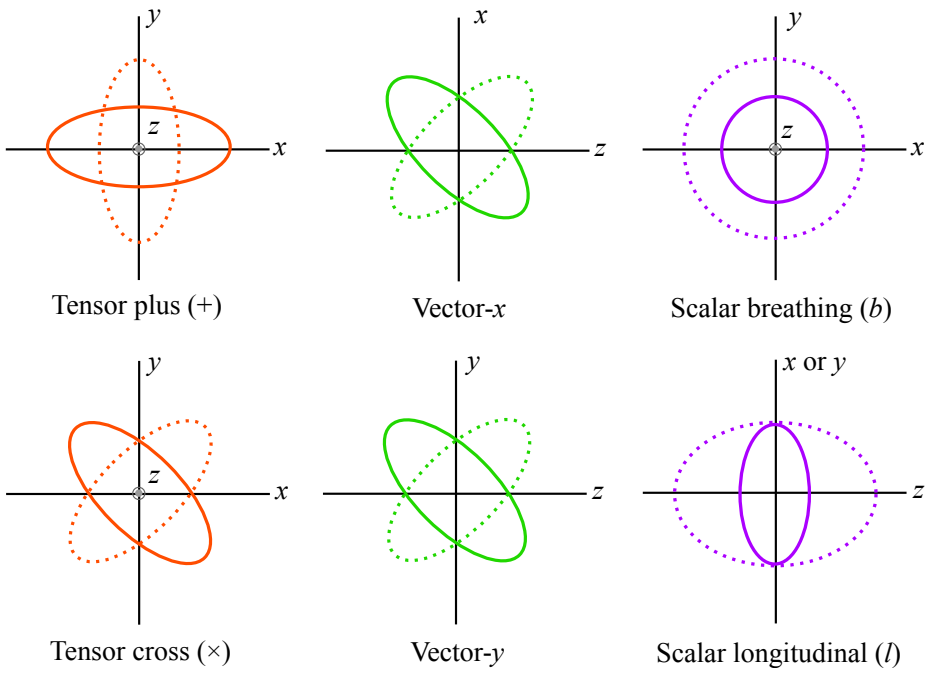

Figure 1.2: Effects of GW polarizations on a ring of free-falling particles, for a wave propagating in the z-direction (out of page). The plus (+) and cross (×) tensor polarizations are shown in orange, the vector-x and -y polarizations are shown in green, the scalar breathing (b) and longitudinal (l) polarizations are shown in purple.

metric tensor to:

$$\bar{h}_{\mu\nu}(t,z) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_b + h_+ & h_\times & h_x \\ 0 & h_\times & h_b - h_+ & h_y \\ 0 & h_x & h_y & h_l \end{pmatrix}_{\mu\nu} \cos\left[w\left(t - \frac{z}{c}\right)\right]. \tag{1.14}$$

Equation 1.14 represents the most general solution for GWs consistent with Einstein's equivalence principle, where gravitational interactions are fully described by the coupling of matter to a space-time metric tensor [65]. The six degrees of freedom in Equation 1.14, represented by the symbols h_P with subscripts $P \in \{+, \times, x, y, l, b\}$, correspond to the amplitudes of the linear polarization modes permitted by generic metric theories of gravity [66]. The subscripts + and × denote the tensor plus and cross polarization modes; x and y denote the vector-x and -y polarization modes; l and b denote the longitudinal and breathing scalar polarization modes. Figure 1.2 illustrates the effect of each polarization mode as the GW propagates through a ring of free-falling particles [67, 68].

In general relativity, the divergence of $T_{\mu\nu}$ is zero, consistent with the Bianchi identities, which ensure the conservation of energy and momentum. As a result, the metric perturbation $\bar{h}_{\mu\nu}$ must also be transverse and traceless. These requirements, together with the Lorentz conditions $\partial^\nu \bar{h}_{\mu\nu} = 0$, imply that $\bar{h}_{\mu\nu}$ must remain invariant under a further coordinate transformation $x^\mu \longrightarrow x'^\mu = x^\mu + \bar{\xi}^\mu$. This coordinate transformation

gives rise to the trace-reverse field transformation:

$$\bar{h}_{\mu\nu} \rightarrow \bar{h}'_{\mu\nu} = \bar{h}_{\mu\nu} - \bar{\xi}_{\mu\nu}, \tag{1.15}$$

where $\bar{\xi}_{\mu\nu} = \partial_\mu \bar{\xi}_\nu + \partial_\nu \bar{\xi}_\mu - \eta_{\mu\nu}\partial_\rho \bar{\xi}^\rho$. By gauge freedom, we can require that $\Box \bar{\xi}_\mu = 0$. Since the flat space D'Alembertian commutes with ∂_μ [69], setting $\Box \bar{\xi}_\mu = 0$ leads to $\Box \bar{\xi}_{\mu\nu} = 0$, i.e. Equation 1.13 is preserved under the above transformation. This additional transformation imposes four more constrains on $\bar{h}_{\mu\nu}$, in addition to $\partial^\nu \bar{h}_{\mu\nu} = 0$. Three of the transform functions $\bar{\xi}^i$ can be chosen so that $h_{0i} = 0$, for the spatial components $i \in \{1, 2, 3\}$. This simplifies the Lorentz condition of the time component ($\mu = 0$) to $\partial^0 h_{00} + \partial^i h_{0i} = \partial^0 h_{00} = 0$. Since h_{00} is time-independent and does not contribute to the emission of GWs, we can conveniently set $h_{00} = 0$ without loss of generality. With $h_{\mu 0} = 0$ for all μ, the overall Lorentz gauge condition reduces to $\partial^i \bar{h}_{ij} = 0$. This divergence-free condition on the spatial coordinates implies that the only non-zero spatial components of the metric perturbation are those transverse to the direction of propagation. The fourth transform function $\bar{\xi}^0$ is chosen such that the trace satisfies $h^\mu{}_\mu = h^i{}_i = 0$, and hence $\bar{h}_{\mu\nu} = h_{\mu\nu}$. Altogether these conditions define the transverse-traceless (TT) gauge. In this gauge, the metric perturbation for a GW propagating in the z-direction reduces to

$$h^{\mathrm{TT}}_{\mu\nu}(t, z) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}_{\mu\nu} \cos\left[w\left(t - \frac{z}{c}\right)\right]. \tag{1.16}$$

In other words, general relativity predicts that GWs are fully characterized by the two tensor polarization modes: + and ×. We present a method to measure the tensor polarization content of GW bursts in Chapter 5 of this thesis.

1.3.3. Generation of gravitational-waves

So far we have only discussed the geometrical origin and properties GWs. Nevertheless, the EFEs in Equation 1.12 are crucial for understanding the physical mechanisms that produce GWs. We discuss their implications below.

As noted earlier, the Bianchi identities guarantee the conservation of energy and momentum within the EFEs. Moreover, since the EFEs are linear in $h_{\mu\nu}$, they can be solved using a Green's function, $G(x, x')$ which satisfies

$$\Box_x G(x - x') = \delta^4(x - x'), \tag{1.17}$$

where x and x' can be treated physically as the positions of the observer and the emitting source respectively. By analogy to electromagnetism, the appropriate solution is the retarded Green's function:

$$G(x, x') = -\frac{1}{4\pi|\boldsymbol{x} - \boldsymbol{x}'|}\delta\left[c\left(t - \frac{|\boldsymbol{x} - \boldsymbol{x}'|}{c}\right) - ct'\right], \tag{1.18}$$

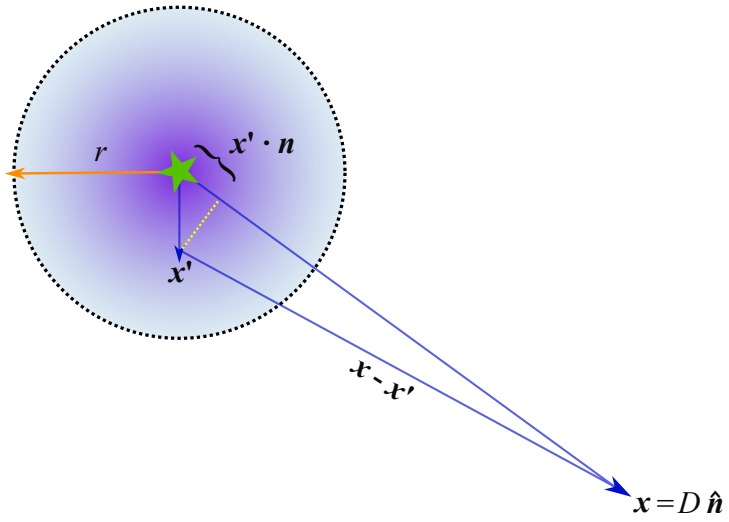

Figure 1.3: Illustration of the coordinate system which results in Equation 1.20. The green star indicates the center of the source.

because GWs also travel at a finite speed, i.e. the speed of light c. The resulting solution to Equation 1.12 is

$$\bar{h}_{\mu\nu}(t,\boldsymbol{x}) = \frac{4G}{c^4}\int d^4x' \frac{1}{|\boldsymbol{x}-\boldsymbol{x}'|} T_{\mu\nu}\left(t - \frac{|\boldsymbol{x}-\boldsymbol{x}'|}{c}, \boldsymbol{x}'\right). \tag{1.19}$$

Let r be the radius of the source, D be the distance to the source and the vector $\boldsymbol{n}$ be the direction at which the wave propagates (see Figure 1.3). Then for a faraway source, i.e. $D \gg r$, we can write

$$\left|\boldsymbol{x}-\boldsymbol{x}'\right| \simeq D - \boldsymbol{x}'\cdot\hat{\boldsymbol{n}}. \tag{1.20}$$

For non-relativistic sources[3], $T_{\mu\nu}$ is only non-vanishing inside the source, i.e. when $\boldsymbol{x}'\cdot\hat{\boldsymbol{n}} \lesssim r \ll D$. Thus, one can expand the stress-energy tensor as follows

$$T_{\mu\nu}\left(t - \frac{D}{c} + \frac{\boldsymbol{x}'\cdot\hat{\boldsymbol{n}}}{c}, \boldsymbol{x}'\right) = \left[T_{\mu\nu} + \frac{x'^i n^i}{c}\partial_t T_{\mu\nu} + \frac{1}{2c^2}\left(x'^i x'^j n^i n^j\right)\partial_t^2 T_{\mu\nu} + \cdots\right]_{\text{ret}}, \tag{1.21}$$

where the subscript 'ret' implies that tensor $T_{\mu\nu}$ and its derivatives are evaluated at the retarded time $t - D/c$. Substituting Equation 1.21 into 1.19, we obtain:

$$\bar{h}_{\mu\nu}(t,\boldsymbol{x}) = \frac{4G}{Dc^4}\int d^4x' T_{\mu\nu}\left(t - \frac{D}{c}, \boldsymbol{x}'\right), \tag{1.22}$$

where we have approximated $1/|\boldsymbol{x}-\boldsymbol{x}'| \approx 1/D$, keeping only the terms of order $\mathcal{O}(1/D)$. The energy-momentum conservation condition can be explicitly written as $\partial_0 T^{00} + \partial_i T^{i0} = 0$, where the indices 0 and i denote the time (t) and spatial (x) coordinates

[3] A source is non-relativistic if the typical velocity of particles within the source is $v \ll c$. The frequency of motion is given by $\omega \sim v/r$, so for non-relativistic sources, the radiated wavelengths $\lambda \sim r(c/v)$ are much larger than the size of the source, r.

respectively. Combining this with Gauss' theorem and noting that the surface integral of $T^{\mu\nu}$ vanishes at $D \gg r$, it can be shown that Equation 1.22 is equivalent to [70]

$$\bar{h}_{\mu\nu}(t, \boldsymbol{x}) = \frac{2G}{Dc^4} \ddot{M}_{\mu\nu}\left(t - \frac{D}{c}\right). \tag{1.23}$$

Here, $\ddot{M}_{\mu\nu}$ is second-order time derivative of the mass quadrupole moment tensor, viz.

$$\ddot{M}^{\mu\nu}\left(t - \frac{D}{c}, \boldsymbol{x}'\right) = \partial_0^2 \int d^3x' T^{00}\left(t - \frac{D}{c}, \boldsymbol{x}'\right) x^\mu x^\nu. \tag{1.24}$$

For non-relativistic sources in the weak-field limit, one can disregard contributions from the kinetic energy, potential energy, and other comparable terms of the particles constituting the source. As a result, T^{00}/c^2 reduces to the rest mass density and the quadrupole moment becomes proportional to the moment of inertia. The monopole and dipole moments do not contribute to the generation of GWs, because their respective time derivatives vanish as a result of energy and momentum conservation.

Equation 1.23 shows that GWs, to the lowest order, are produced by a time-varying moment of inertia. In other words, GW sources must entail asymmetries in their mass distributions. For a binary with equal component masses M, orbital separation R and orbital angular frequency $\omega \sim \sqrt{GM/4R^3}$, the second time-derivative of the quadrupole moment can be approximated as $\ddot{M}(t - D/c) \sim M\omega^2(2R^2)$. Using Equation 1.23 and assuming the source parameters of GW150914, i.e. $M \sim 30M_\odot$ and $D \sim 400$ Mpc, we can estimate strain amplitude as $h \sim 10^{-21}$, when R is 10 times the Schwarzschild radius. This simple exercise reveals that the space-time perturbations caused by GWs, upon their arrival at Earth, are exceedingly small. Therefore, detectable GW sources are restricted to compact objects with a large and rapidly varying moment of inertia, such as those discussed in Section 1.2.

1.4. Ground-based gravitational-wave detectors

The passage of GWs through space-time causes fluctuations in the relative distance between two fixed points. That is, the strain amplitude can be physically measured as

$$h = \frac{\Delta L}{L}, \tag{1.25}$$

where ΔL is the change in spatial distance between the two points and L is the reference distance when GWs are absent. As shown above, GW amplitudes are on the order of $h \sim 10^{-21}$, so physical detectors with $L \sim 1$ km must be capable of measuring ΔL with a precision on the order of 10^{-18} m. Achieving this level of precision in ground-based GW detectors requires advanced laser interferometry and noise mitigation techniques.

In Section 1.4.1, we outline the design and technologies employed by current ground-based interferometric GW detectors to achieve the sensitivity needed to detect the anticipated GW sources. In Section 1.4.2, we overview the existing GW observatories, the plans for future-generation detectors and how the expanding global detector network enhances GW detection and characterization. In Section 1.4.3 we discuss the types of unexpected transient noise artifacts (glitches) that have been observed in GW detectors, and how they affect transient GW detections.

1.4.1. *Detector design*

The current second-generation ground-based GW detectors are modified Michelson interferometers with kilometer-scale arms. Figure 1.4 shows a simplified setup of a modified Michelson interferometer. The recombination of laser beams from the orthogonal arms at the beam splitter produces time-evolving interference patterns, revealing changes in the differential arm length $\Delta L(t) = L_x(t) - L_y(t)$ caused by a passing GW. Conventionally, the detectors use a laser with wavelength of $\lambda_{\text{laser}} = 1064$ nm, due to its stability in terms of amplitude, frequency and beam geometry [23]. Therefore, the sensitivity of a standard Michelson interferometer with arm lengths $L \sim 1$ km can be estimated as $h \sim \lambda_{\text{laser}}/L \sim 10^{-9}$. This sensitivity is far from what is needed to detect the anticipated GWs with $h \sim 10^{-21}$.

Existing detectors implement several modifications to the Michelson setup to achieve desirable GW detection sensitivities. Below, we outline some of the key technologies:

(1) **Fabry-Peŕot cavity**. According to Equation 1.25, the strain sensitivity can be improved with increasing detector arm length L. While it is not feasible to construct an exceptionally large detector, it is possible to increase the effective arm length L_{eff} by increasing the optical paths of the laser. This is achieved using a Fabry-Peŕot resonant optical cavity enclosed between two partially reflective test masses, as illustrated in Figure 1.4. By allowing the light to traverse back and forth multiple times within the cavity before exiting, the power inside the cavity, as well as the effective optical path, is increased by up to a factor of 10^3, i.e. $L_{\text{eff}} \sim 1000$ km. The increase in L_{eff} is limited by the requirement that it must not exceed the wavelength $\lambda_{\text{GW}} \sim c/f_{\text{GW}}$ of the target GW with frequency f_{GW}, to ensure that the reference cavity length remains unaffected by the metric perturbation.

(2) **Power-recycling cavity.** The overall laser power can also be enhanced before entering the Fabry-Peŕot cavities. This additional resonant power buildup is achieved with a power-recycling cavity, formed using partially transmissive mirrors, and placed between the laser source and the beam splitter (see Figure 1.4). The LIGO power-recycling cavities as of O4 increases the circulating power by up to a factor of 50 [71]. The impact of increased power on sensitivity will be discussed shortly.

(3) **Signal-recycling cavity.** Laser light consists of photons that are subject to quantum fluctuations, that is, the arrival times of photons at the photodetector are inherently random. This randomness gives rise to shot noise, which cannot be eliminated but can be counteracted. One way to do so is by increasing the laser power (as described above), which reduces the impact of quantum fluctuations relative to fluctuations from a GW signal. Another way to counteract shot noise is by amplify the signal response, which increases sensitivity to $\Delta L(t)$. This is achieved by placing a signal-recycling cavity between the beam splitter output and the photodetector. The mirrors in the signal-recycling cavity can be adjusted to target specific frequency bands [72]. Currently, the detectors are tuned to be sensitive across a broadband frequency range of ∼20–5000 Hz [73].

The combined use of the recycling cavities, as described in (2) and (3), is known as dual recycling [74]. This is the state-of-the-art technology employed in the second-generation ground-based GW detectors.

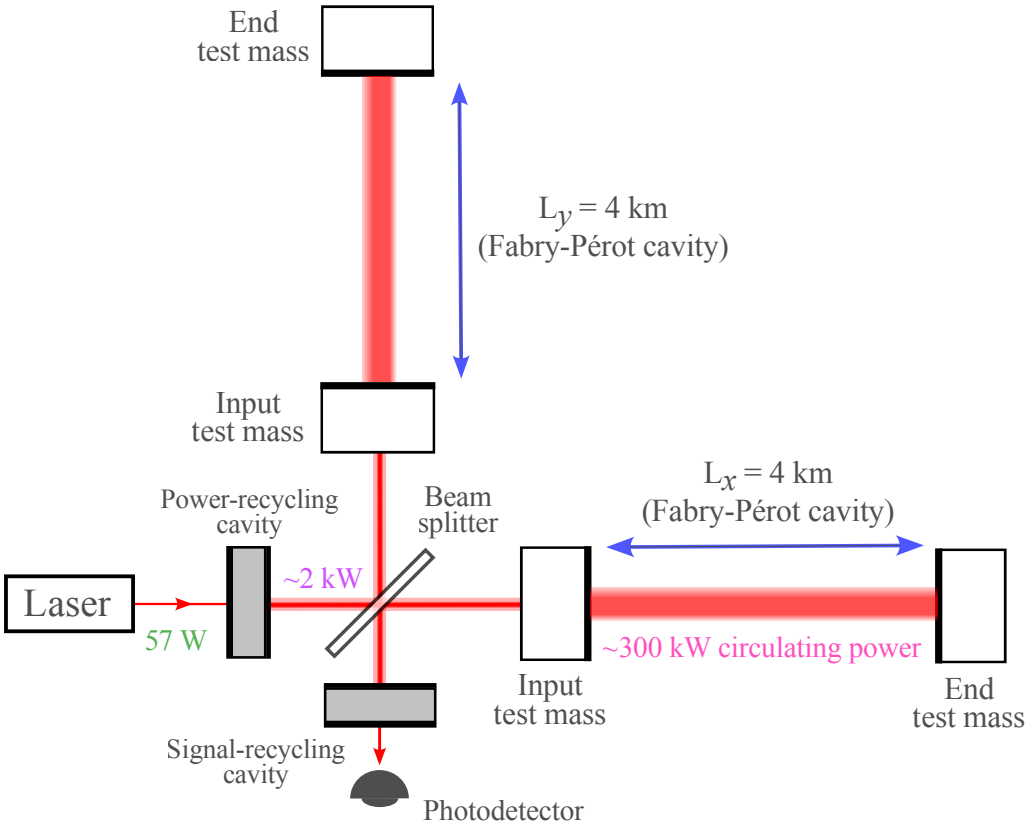

Figure 1.4: Simplified schematic of a modified Michelson interferometer with arm lengths $L_x = L_y = 4$ km (not to scale). This figure is adapted from Figure 3 of Ref. [23]; the laser input and cavity power values are updated to be consistent with the performance summary of the Advanced LIGO detectors in O4 [71].

We can now estimate the detector sensitivity with the above technologies implemented. First, we assume a Poisson-distributed photon arrival time. That is, if the photodetector is expected to collect N_{photons} over time τ, the fluctuations are given by $N_{\text{photons}}^{1/2}$. Therefore, the minimum detectable change in path length is given by [75]:

$$\Delta L \sim \frac{N_{\text{photons}}^{1/2}}{N_{\text{photons}}} \lambda_{\text{laser}}. \tag{1.26}$$

Secondly, we set $\tau \sim 1/f_{\text{GW}}$ so that N_{photons} depends on the frequency of the gravitational wave that we aim to detect; larger τ will cause the noise fluctuations to mask the GW signal. With this assumption, we can write

$$N_{\text{photons}} \sim \frac{P_{\text{cavity}}}{(h_{\text{planck}} c/\lambda_{\text{laser}})} \frac{1}{f_{\text{GW}}}, \tag{1.27}$$

where P_{cavity} denotes the circulating laser power inside the Fabry-Pérot cavities, and h_{planck} is the Planck constant (not to be confused with the GW strain h). With $\lambda_{\text{laser}} = 1064$ nm, $f_{\text{GW}} = 300$ Hz, and $P_{\text{cavity}} = 300$ kW as depicted in Figure 1.4, we obtain

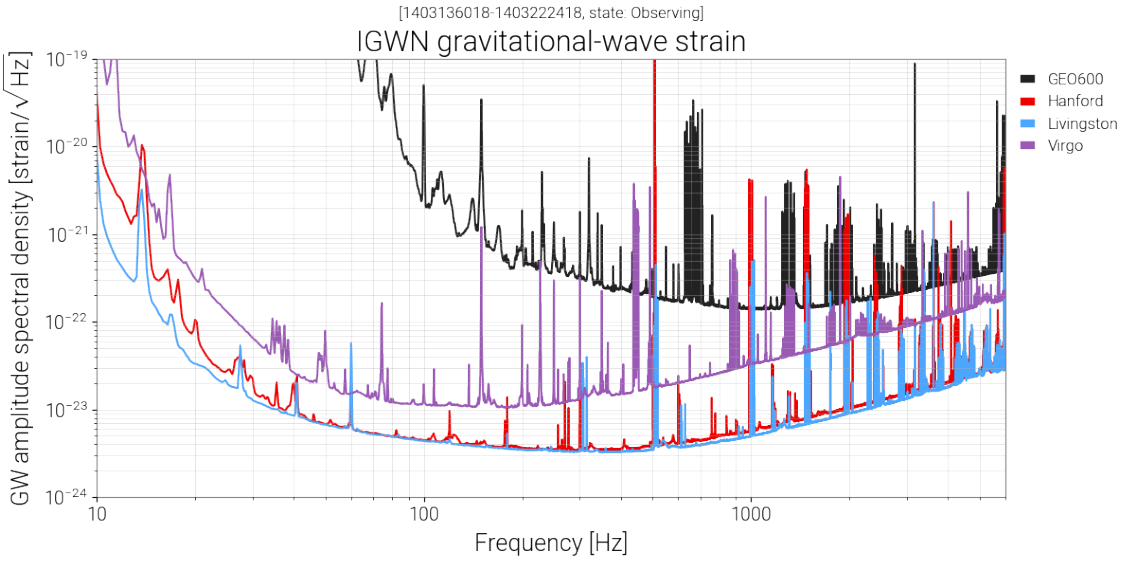

Figure 1.5: Representative O4b amplitude strain density (ASD) curves of the LIGO Hanford, LIGO Livingston, Virgo and GEO600 detectors. The ASD curves, measured as a function of frequency in units of $\mathrm{Hz}^{-1/2}$, display the daily temporal median sensitivity of each detector, calculated on the 23rd of June 2024 where all four detectors were active concurrently. These curves are obtained from the Gravitational Wave Open Science Center (GWOSC) [81].

$N_{\text{photons}} \sim 10^{21}$, and hence $\Delta L \sim 10^{-17}$ m. Given the effective optical path is also increased to $L_{\text{eff}} \sim 1000$ km, we achieve $h \sim \Delta L / L_{\text{eff}} \sim 10^{-23}$, which meets the required sensitivity for detecting GWs with $h \sim 10^{-21}$.

1.4.2. Global gravitational-wave observatories

There are currently five operational large-scale GW observatories around the world: the Advanced LIGO Hanford (H) and LIGO Livingston (L) detectors in the United States, each with 4 km arms [76]; Advanced Virgo (V) in Italy [77] and the Kamioka Gravitational Wave Detector (KAGRA, K) in Japan [78], each with 3 km arms; and GEO600 in Germany with 600 m arms [79]. The HLV network and GEO600 were active throughout O4b, which lasted 355 days, concluding on the 1st of April 2025; representative sensitivity curves of the active detectors during the run are shown in Figure 1.5. KAGRA has been offline due to the impact of the 2024 Noto earthquake and is undergoing commissioning. The third and final phase of O4 is scheduled to begin in June 2025, with plans for KAGRA to rejoin by the end of O4 ($\sim$ October 2025) [80].

The ground-based GW detector network is set to expand further in the coming decades, with the commissioning of LIGO-India [82] already underway. Third-generation detectors like Cosmic Explorer [83] in the United States and Einstein Telescope [84] in Europe have also been proposed. These next-generation detectors are expected to cover a similar frequency range to current detectors, but aim to improve sensitivity by up to an order of magnitude through longer arm lengths.

A larger and more sensitive global detector network has many advantages, including

but not limited to:

- **Increased signal-to-noise ratio (SNR) and sensitivity volume.** The overall network SNR for a network comprising $\mathcal{I}$ detectors is given by [85]:

$$\mathrm{SNR}^2_{\mathrm{net}} = \sum_{i=1}^{\mathcal{I}} \mathrm{SNR}^2_i. \tag{1.28}$$

  This can be approximated as $\mathrm{SNR}_{\mathrm{net}} \sim \sqrt{\mathcal{I}}\,\mathrm{SNR}_{\mathrm{IFO}}$, assuming each detector contributes the same $\mathrm{SNR}_{\mathrm{IFO}}$. For a single detector, the SNR scales with the source luminosity distance as $\mathrm{SNR}_{\mathrm{IFO}} \sim 1/D_L$ [86]. Consequently, for a network of $\mathcal{I}$ detectors, $\mathrm{SNR}_{\mathrm{net}} \sim \sqrt{\mathcal{I}}/D_L$, i.e. for the same D_L, the SNR is amplified by a factor of $\sqrt{\mathcal{I}}$. In other words, a network with $\mathcal{I}$ detectors can detect sources up to $\sqrt{\mathcal{I}}$ times farther than a single detector. Altogether, increasing $\mathcal{I}$ amplifies the SNR across all GW detections, and consequently the search volume $V \propto D_L^3$.

- **Increased sky coverage.** Assuming GWs follow the predictions of general relativity, a detector's unique antenna pattern functions, F_+ and $F_\times$, govern its sensitivity to the plus and cross polarization modes. The overall response of a single detector can then be expressed in the time domain as

$$h(t) = F_+(\alpha, \delta, \psi)h_+(t) + F_\times(\alpha, \delta, \psi)h_\times(t), \tag{1.29}$$

  where α, δ and ψ respectively denote the right ascension, declination and the polarization angle[4] relative to the detector axes. Equation 1.29 suggests that adding more detectors with different orientations and at various geographical location increases sky coverage, i.e. the network sensitivity broadens, covering a larger area of the sky [89].

- **Better characterization of GW polarizations.** Since each detector has its own response described by Equation 1.29, increasing the number of detectors creates additional simultaneous equations for the signal response. Provided the detectors are not co-aligned, increasing the number of detectors improves our ability to distinguish between GW polarizations [90].

- **Improved sky localization.** Sky localization is achieved by triangulating the time delays between multiple detectors. Therefore increasing the number of detectors and the spatial separation between detectors improves the accuracy of sky localization [90–92].

- **Increased duty cycle and detection confidence.** The duty cycle, during which at least two detectors are observing simultaneously, increases with a larger network. Consequently, the rate of coherent multi-detector observations increases, and the agreement between detectors strengthens the detection confidence [90].

However, in real observing scenarios, GW detectors are subject to unexpected noise transients (glitches). As more detectors are added, the frequency of glitches is also expected to increase, which could potentially offset some of the improvements described above. We discuss the properties and adverse impacts of glitches in the following section.

[4] The polarization angle ψ is the angle of rotation of between the GW polarization basis and the detector axes [87, 88].

1.4.3. *Glitches in gravitational-wave detectors*

As demonstrated in Section 1.4.1, GW detectors are sensitive to fractional changes in distance as small as 10^{-23}. Therefore, they are susceptible to a wide range of instrumental and environmental noise sources. Spectral analysis of the GW detector noise floor, such as the one shown in Figure 1.5, reveals that thermal noise, seismic noise and quantum noise are the primary long-term contributors to broadband frequency noise [73]. There are also persistent and elevated narrowband noise (spectral lines), which are typically of instrumental, environmental or calibration origin [93]. Most noise sources in the existing GW detectors are well-understood, and thus, can be accurately modeled and subtracted [94]. Nevertheless, transient noise with non-Gaussian spectral features can also arise from unforeseen sources, e.g. anthropogenic activities, weather changes, instrumental defects or malfunctions, and occasionally, unknown origins. These non-stationary and non-Gaussian transient noise artifacts are called glitches [95].

Figure 1.6 shows the time-frequency structure of several common glitch morphologies observed in the LIGO detectors. As illustrated, glitches manifest as short bursts of excess noise in the GW data stream, typically lasting less than a few seconds, and can span a wide range of frequencies, from narrowband to broadband. The transient and unpredictable nature of glitches pose challenges to the detection of GW transients, because they can mimic or mask real astrophysical signals, resulting in false-alarm or low confidence detections. Glitches in close proximity to or overlapping with GW signals can also bias parameter estimation, if they are not accurately modeled and subtracted, e.g. GW170817 on the characterization of tidal deformability [97] and GW200129_065458 on the inference of spin-precession [98].

Glitches are common occurrences throughout the three LVK observing runs, and various efforts have been put into the characterization and mitigation of glitches [99–102]. During O3a, the Advanced LIGO and Virgo detectors operated with improved sensitivities compared to O2, but the average glitch rate in the LIGO detectors increases from ∼0.1 per minute in O2 to ∼0.5 per minute in O3a [36]. Despite the higher glitch rate, the number of detections during the 6-month O3a run surpassed those of the 8-month O2 run by a factor 5. This observation raises a nontrivial question of how the rate of glitches affects the overall GW detection performance, and whether the impact differs between a smaller and larger detector networks. At first glance, one might assume that a larger detector network with improved sensitivity would naturally have higher detection efficiency. However, this is not necessarily true, as the higher glitch rates in larger detector networks may significantly lower the detection confidence, resulting in reduced detection efficiency. Since glitch rates and their impact on multi-detector network performances cannot be predicted using models, answering the above question require empirical studies. We explore this topic more thoroughly in Chapter 3.

1.5. Detecting gravitational-wave transients

Previously in Section 1.2.2, we noted that searches for well-modeled CBCs and poorly-constrained GW bursts require different analysis techniques. In Section 1.5.1, we discuss the matched-filtering approach for CBC searches, and in Section 1.5.2, we overview the minimally-modeled analyses used for detecting and characterizing GW bursts.

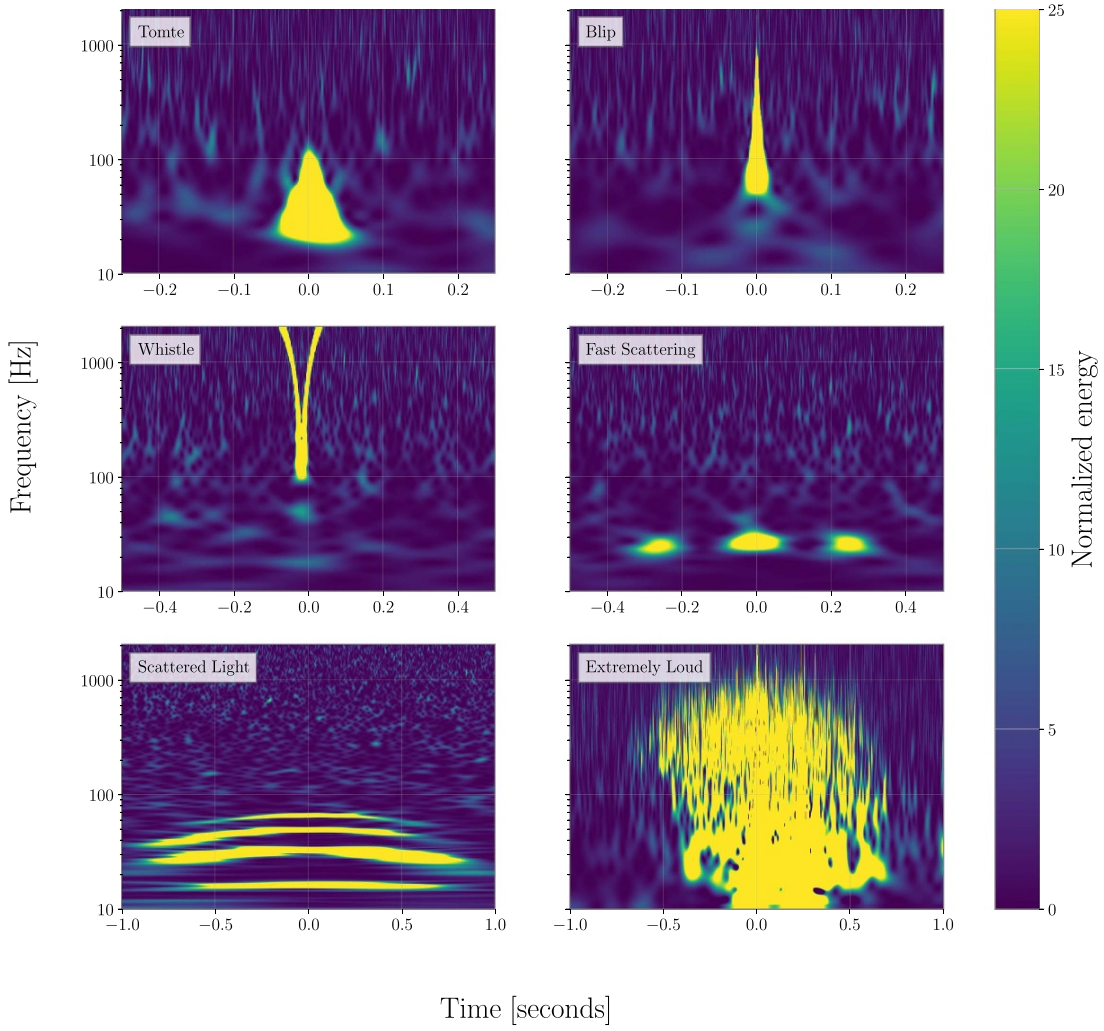

Figure 1.6: Time-frequency spectrograms of common glitch morphologies observed in the LIGO detectors. The horizontal axes, representing time in seconds, scales differently in each panel. The Tomte, Blip and Whistle glitches are short-duration ($\sim$0.1$-$0.25 s) glitches, while the Fast Scattering, Scattered Light and Extremely Loud glitches have comparatively longer durations ($\sim$1$-$2 s). Each of the short- and long-duration glitches also cover different frequency ranges. This figure is reproduced directly from Ref. [96].

1.5.1. *Matched filtering for well-modeled signals*

Matched filtering involves cross-correlating a set of predicted waveforms (templates) with the detector data, to identify the template that most closely matches (overlaps) with the data. We outline the methodology below.

In the absence of glitches, the Fourier-domain detector data $\tilde{d}(f)$ is given by the superposition of stationary Gaussian noise $\tilde{n}(f)$ and possibly a signal $\tilde{h}(f)$, viz.

$$\tilde{d}(f) = \tilde{n}(f) + \tilde{h}(f). \tag{1.30}$$

Since the Gaussian noise is uncorrelated across the spectrum, one can write:

$$\langle \tilde{n}(f)\tilde{n}^*(f')\rangle = \frac{1}{2}\delta(f-f')S_n(f), \tag{1.31}$$

where $\tilde{n}^*$ denotes the complex conjugate of $\tilde{n}$, $\langle \cdots \rangle$ denotes the ensemble average over many noise realizations, δ denotes the Dirac-delta function and $S_n(f)$ denotes the one-sided noise power spectral density (PSD). $S_n(f)$ is called the one-sided PSD because $\tilde{n}^*(f) = \tilde{n}(-f)$ for real detector output, implying that $S_n(f) = S_n(-f)$. That is, $S_n(f)$ is an even function across all Fourier frequencies, which also explains why the factor of $1/2$ in Equation 1.31 is necessary.

Now let us define a real-valued filter $K(t)$ with Fourier transform $\tilde{K}(f)$. If a signal $h(t)$ is present in the time-domain data, the convolution theorem allows us to write its filtered value as

$$S = \int_{-\infty}^{\infty} dt\, h(t) K(t) \tag{1.32}$$

$$= \int_{-\infty}^{\infty} df\, \tilde{h}(f) \tilde{K}^*(f), \tag{1.33}$$

where $\tilde{h}$ denotes the Fourier transform of h. Using the properties of noise as in Equation 1.31, we obtain

$$\langle N \rangle^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} df' df\, \langle \tilde{n}(f) \tilde{n}^*(f') \rangle \tilde{K}(f) \tilde{K}^*(f') \tag{1.34}$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} df \left| \tilde{K}(f) \right|^2 S_n(f). \tag{1.35}$$

The SNR is then given by

$$\mathrm{SNR}^2 = \frac{S^2}{\langle N \rangle^2}. \tag{1.36}$$

Our goal now is to find an optimal $K(t)$ that maximizes the SNR. The Cauchy-Schwarz inequality states that

$$\left| \int_{-\infty}^{\infty} df\, A(f) B(f) \right|^2 \leq \int_{-\infty}^{\infty} df |A(f)|^2 \int_{-\infty}^{\infty} df |B(f)|^2. \tag{1.37}$$

Since Equation 1.33 is equivalent to

$$S = \int_{-\infty}^{\infty} df\, \frac{\tilde{h}(f)}{\sqrt{S_n(f)}} \tilde{K}^*(f) \sqrt{S_n(f)}, \tag{1.38}$$

the SNR (Equation 1.36) can be maximized by expressing S^2 as the right-hand side of Cauchy-Schwarz inequality (Equation 1.37) with $A(f) = \tilde{h}(f)/\sqrt{S_n(f)}$ and $B(f) = \tilde{K}^*(f)\sqrt{S_n(f)}$. This results in [103]:

$$\mathrm{SNR}^2 = \frac{\int_{-\infty}^{\infty} df \frac{|\tilde{h}(f)|^2}{S_n(f)} \int_{-\infty}^{\infty} df \left| \tilde{K}(f) \right|^2 S_n(f)}{\frac{1}{2} \int_{-\infty}^{\infty} df \left| \tilde{K}(f) \right|^2 S_n(f)} \tag{1.39}$$

$$= 2 \int_{-\infty}^{\infty} df\, \frac{\tilde{h}(f) \tilde{h}^*(f)}{S_n(f)} \tag{1.40}$$

$$= 4\,\mathrm{Re} \int_{0}^{\infty} df\, \frac{\tilde{h}(f) \tilde{h}^*(f)}{S_n(f)} \tag{1.41}$$

$$= \left\langle \tilde{h} \middle| \tilde{h} \right\rangle. \tag{1.42}$$

Equation 1.41 reveals $\tilde{K}(f) = 4\tilde{h}(f)/S_n(f)$ as the filter that maximizes the SNR when extracting a signal embedded in Gaussian noise. Equation 1.41 is otherwise known as the noise-weighted inner product, and is more compactly represented as in Equation 1.42.

Matched-filtering aims to find a signal template $\tilde{h}_{\mathrm{T}}(f)$ that maximizes the SNR by correlating it with the observed data $\tilde{d}(f)$. In other words, it seeks to maximize the *overlap* between $\tilde{h}_{\mathrm{T}}$ and $\tilde{d}$, given by [104]

$$\mathcal{O} = \frac{\left\langle \tilde{h}_{\mathrm{T}} \middle| \tilde{d} \right\rangle}{\sqrt{\left\langle \tilde{h}_{\mathrm{T}} \middle| \tilde{h}_{\mathrm{T}} \right\rangle \left\langle \tilde{d} \middle| \tilde{d} \right\rangle}}, \tag{1.43}$$

where $\mathcal{O}$ is normalized to range from -1 (perfect anticorrelation) to 1 (perfect match). To achieve this, one needs a set of phenomenologically modeled templates (template bank) that spans a wide range of signal parameters. The spacing of signal parameters is optimized to maximize coverage of the parameter space while also keeping computational costs within reasonable limits [105, 106]. Therefore, matched-filtering is only suitable for detecting GW sources with well-constrained parameter spaces, such as CBCs. The matched-filter pipelines used for the CBC searches in the GWTCs [35–38] are: PyCBC [107, 108], the GStreamer LIGO Algorithm Library (GstLAL) [109, 110], the Multi-Band Template Analysis (MBTA) [111] and the Summed Parallel Infinite Impulse Response (SPIIR) pipeline [112, 113]. This thesis focuses on methods for detecting generic bursts, discussed in Section 1.5.2, so matched-filtering is not considered further.

1.5.2. Minimally-modeled analyses for generic bursts

The detection and characterization of GW bursts, where the signal models are poorly-constrained or unavailable, rely on analyses that make minimal assumptions about the signal source and morphology. However, minimally-modeled analyses are more susceptible to glitches than matched filtering due to their increased flexibility. Therefore, to effectively distinguish astrophysical signals from glitches, these analyses must cross-correlate data from multiple detectors. Additionally, to ensure consistent and reliable search results, the LVK all-sky searches for GW bursts employ multiple minimally-modeled analyses to perform independent searches. Since the analysis pipelines are sensitive to different regions of the parameter space, combining multiple pipelines also improves the search sensitivity across a wider range of signals. Although analysis pipelines are not optimized to target a particular type of burst source, they are optimized to detect signals of specific durations. The short-duration burst searches [44–46] are conducted using three algorithms: coherent WaveBurst (cWB) [114–116], omicron-LIB (oLIB) [117] and *BayesWave* [1, 2]. The long-duration burst searches [47–49] use the all-sky Stochastic Transient Analysis Multi-detector Pipeline (STAMP-AS) [118–120], the X-pipeline Spherical Radiometer (X-SphRad) [121, 122], as well as the long-duration configuration of cWB. Low-latency (real-time) analyses in O4 are conducted using cWB and MLy [123].

Minimally-modeled analyses are usually performed in the time-frequency plane, where regions with coherent excess power across the detector network are reconstructed using generic mathematical frames, e.g. wavelets, that can dynamically adapt to changes in time and frequency. Table 1.2 summarizes the variety of methods and detection statistics used by the above-mentioned LVK search pipelines, demonstrating the versatility of minimally-modeled analyses.

Pipeline	Burst duration	Method	Detection statistic
cWB [114–116]	Short & long	Adaptive signal characterization using multi-resolution wavelet transforms. There are currently two versions of cWB: (i) cWB-2G uses the Wilson-Daubechies-Meyer wavelet transform [124] (ii) cWB-XP uses the *WaveScan* transform based on Gabor wavelets [125]	All versions of cWB up to and including cWB-2G use the excess-power statistic. cWB-XP uses both excess-power and cross-power statistics.
MLy [123]	Short	Low-latency convolutional neural network (CNN) analysis with Pearson cross-correlation to distinguish correlated signals from uncorrelated glitches.	CNN score
oLIB [117]	Short	Hierarchical analyses to identify incoherent and coherent power. Signals and glitches are modeled using sine-Gaussian wavelets.	Joint likelihood ratio of Bayes factors.
BayesWave [1, 2]	Short	Independent reconstruction of signals and glitches using sine-Gaussian wavelets.	Signal-versus-glitch Bayes factor
STAMP-AS [118–120]	Long	Cross-power time-frequency map reconstruction. Triggers are determined by pixel clusters that exceed a nominal SNR threshold.	Network cross-power SNR
X-SphRad [121, 122]	Long	Frequency-domain cross-power analysis of spherical harmonics; evaluates power across various angular scales, corresponding to different degrees of homogeneous polynomials ($\ell \geq 0$).	Power ratio between $\ell > 0$ modes and $\ell = 0$

Table 1.2: Overview of the minimally-modeled analysis pipelines used in LVK burst searches. The columns, from left to right, display: (i) the name of the burst pipeline, (ii) the burst duration the pipeline is configured for, (iii) a summary of the pipeline method, and (iv) the detection statistic.

1.6. Thesis outline

This thesis studies the performance and applications of the minimally-modeled analysis algorithm, *BayesWave* [1, 2], in detecting and characterizing short-duration GW bursts. Henceforth, unless stated otherwise, any mention of GW bursts refers to short-duration bursts. The rest of this thesis is structured as follows.

In Chapter 2, we detail the methods of the *BayesWave* algorithm, analyze the behavior of its detection statistics, and provide a comprehensive literature review of *BayesWave*'s diverse applications.

In Chapter 3, we assess the impact of glitches in practical observing scenarios, on the multi-detector GW burst detection efficiency of *BayesWave*. We find a counterintuitive result, where the three-detector LIGO-Virgo network does not outperform the LIGO two-detector network in burst detection efficiency with *BayesWave* during O3. This is because the increased detection sensitivity of the three-detector network is offset by the higher glitch rate.

Low-frequency ($\lesssim 250$ Hz) GWs emitted by CCSNe are known to have implications for deducing the physical mechanisms driving their explosions. In Chapter 4, we propose a dedicated-frequency follow-up framework using a hierarchical analysis pipeline consisting of cWB [114–116] and *BayesWave*, to explore the detectability of low-frequency GW emissions under observing conditions similar to O3. We use the HL configuration in this study guided by the results from previous cWB studies and the *BayesWave* efficiency analysis in Chapter 2. We find that dedicated low-frequency follow-up analyses, limited to the range 32–256 Hz, can be used to constrain the explosion model of a CCSN candidate if a successful detection is made.

In Chapter 5, we shift our focus from current observing scenarios to explore the prospects of *BayesWave* in characterizing tensor GW burst polarizations using the global ground-based detector network, once the LIGO, Virgo and KAGRA reach their target sensitivities. Specifically, we compare the performance of *BayesWave*'s relaxed (R) polarization model with the default elliptical (E) polarization model, and find that R is preferred for the characterization of signals with pronounced non-elliptical polarizations. We also show that R can measure burst polarization content in terms of Stokes parameters, with improved accuracy as the detector network expands.

We conclude in Chapter 6 with summaries of the above works and discussions on future research directions.

Chapter 2

The *BayesWave* algorithm

This chapter contains reproduced text from Section II of [4]:

> *Y. S. C. Lee, M. Millhouse, and A. Melatos*,
> Physical Review D 109, 082002 (2024).

and calculations, methods, and results adapted with substantial rewording from [3]:

> *Y. S. C. Lee, M. Millhouse, and A. Melatos*,
> Physical Review D 103, 062002 (2021).

Parts of the adapted and reproduced text from [3] and [4] are combined with supplementary text and figures, to produce a coherent description of the *BayesWave* algorithm and its applications. The adaptations and reproductions are detailed in the preface. Where relevant, sources of adapted texts and materials are acknowledged through in-text footnotes. Notations from the original texts are updated to match the rest of this thesis. This chapter also contains original work of the author that has not been submitted for publication.

2.1. *BayesWave* Overview

The *BayesWave* algorithm is designed to adaptively reconstruct non-stationary and non-Gaussian transients in the data, using models with variable dimensions. As discussed in Section 1.5.2, *BayesWave* is suitable for analyzing generic GW bursts because it makes no *a priori* assumption about its source or waveform morphology. The algorithm is designed to target GW bursts with durations of up to a few seconds. A novel feature of *BayesWave* that sets it apart from other GW burst analysis algorithms is its ability to jointly characterize instrumental transients (glitches) and astrophysical signals. The name *BayesWave* expresses the two key concepts underpinning the algorithm: (i) waveform reconstruction using sine-Gaussian (also known as Morlet-Gabor) wavelets, and (ii) the implementation of Bayesian inference to discriminate signals from glitches.

This chapter outlines the fundamental principles and applications of the *BayesWave* algorithm. Section 2.2 describes the models for data reconstruction. Section 2.3 overviews the methods for sampling the model parameters. Section 2.4 discusses the role of the Bayes factor as *BayesWave*'s detection statistic. Section 2.5 discusses the scaling behavior of the Bayes factor. Finally, in Section 2.6 we review the applications of *BayesWave* in the existing literature.

2.2. Modeling the data

2.2.1. Sine-Gaussian Wavelets

For a given detector i, the data $d_i(t)$ at time t can consist of up to three components: the GW signal $h_i(t)$, which is bounded in t for burst sources; glitches $g_i(t)$, which are also bounded in t; and random detector noise $n_i(t)$, which is present continuously. That is, in general, the data can be expressed as $d_i(t) = h_i(t) + g_i(t) + n_i(t)$. *BayesWave* reconstructs the transient, non-Gaussian features, i.e. $h_i(t)$ and/or $g_i(t)$, by summing a set of sine-Gaussian wavelets. Figure 2.1 shows a visualized example of how the GW150914 time-domain signal is represented as a sum of five wavelets.

A single sine-Gaussian wavelet in the time domain takes the mathematical form

$$\Psi(t;\boldsymbol{\lambda}) = Ae^{-(t-t_0)^2/\tau^2}\cos\left[2\pi f_0(t-t_0)+\phi_0\right], \tag{2.1}$$

with $\tau = Q/(2\pi f_0)$ and $\boldsymbol{\lambda} = \{t_0, f_0, Q, A, \phi_0\}$. The symbols t_0, f_0, Q, A, ϕ_0 denote the central time, central frequency, quality factor, amplitude and phase offset of the wavelet respectively, collectively known as the intrinsic parameters. One reason for choosing sine-Gaussian wavelets over other basis frames[1] is due to its flexibility in the time-frequency domain, making them well-suited for modeling arbitrarily-shaped signals and glitches. The sine-Gaussian wavelet also has an analytic Fourier representation

$$\tilde{\Psi}(f;\boldsymbol{\lambda}) = \frac{\sqrt{\pi}A\tau}{2}e^{-\pi^2\tau^2(f-f_0)^2}\left\{e^{i[\phi_0+2\pi(f-f_0)t_0]} + e^{-i[\phi_0+2\pi(f+f_0)t_0]}e^{-Q^2f/f_0}\right\}, \tag{2.2}$$

so the *BayesWave* analysis can be performed entirely in the frequency domain. This is useful because, in the frequency domain, the time-delay operator simplifies to a phase shift (see Equation 2.3). Representations of bursts and glitches in the time-frequency domain can be achieved using a non-orthogonal sine-Gaussian wavelet frame, with the implementation of systematic priors. We discuss this further in Section 2.3.

The *BayesWave* algorithm offers four independent models for data reconstruction, namely (i) the GW signal plus Gaussian-noise model, $\mathcal{S}$, (ii) glitches plus Gaussian-noise model, $\mathcal{G}$ (iii) Gaussian-noise model, $\mathcal{N}$ and (iv) signal plus glitches plus Gaussian-noise model, or in short, the joint model $\mathcal{J}$. The properties of each model are described in the following subsections in the same order. While the applications of $\mathcal{J}$ are not explored in this thesis, its description is included here for completeness.

2.2.2. Signal model, $\mathcal{S}$

Recall that the five intrinsic parameters of a sine-Gaussian wavelet can represented with a single parameter vector $\boldsymbol{\lambda} = \{t_0, f_0, Q, A, \phi_0\}$. If a real GW signal is present in the data of a multi-detector network, we expect it to be coherent across all detectors in the network, albeit with different signal-to-noise ratio (SNR) and polarization per detector depending on the sensitivity and orientation of the detectors respectively. Therefore

[1] *BayesWave* reconstruction has been empirically shown to improve when chirplets with time-evolving frequency content are used instead of sine-Gaussian wavelets. However, using chirplets results in less effective signal and glitch separation through Bayes factor comparison [126].

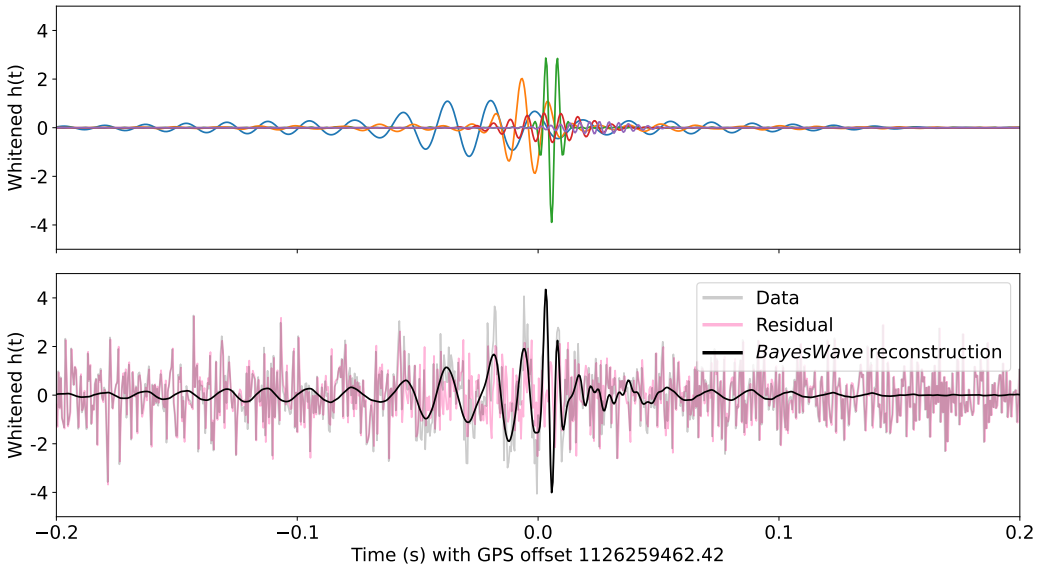

Figure 2.1: Visualization of *BayesWave* waveform reconstruction using sine-Gaussian wavelets. The top panel shows individual sine-Gaussian wavelets, each of different colour, used in the reconstruction of the first detection event GW150914 in the time-domain. The bottom panel shows the resulting *BayesWave* waveform (black) from summing the wavelets in the top panel, overlaid on the actual data (grey); the residuals (pink) are obtained by subtracting the waveform from the data. The vertical axes in both figures show the time-domain whitened strain amplitude, which is the inverse Fourier transform of the frequency-domain amplitude normalized by the Gaussian noise PSD.

when reconstructing the data using the signal model, the same wavelet parameters are used across all detectors in the network. The set of intrinsic parameters for the signal model ($\mathcal{S}$) is given by $\boldsymbol{\lambda}^{\mathcal{S}} = \boldsymbol{\lambda}_1 \cup \boldsymbol{\lambda}_2 \cdots \cup \boldsymbol{\lambda}_{N^{\mathcal{S}}}$, where $N^{\mathcal{S}}$ denotes the number of wavelets used in the signal reconstruction. These parameters are geocentric, meaning they are measured at a reference point located at the center of the Earth.

Since the signal models represent astrophysical GW signals, all $N^{\mathcal{S}}$ wavelets used in the reconstruction also share a set of extrinsic parameters $\boldsymbol{\Phi} = \{\alpha, \delta, \epsilon, \psi\}$. The symbols denote the right ascension, declination, ellipticity and polarization angle of the GW in order of appearance. The complete set of signal model parameters is then given by $\boldsymbol{\theta}^{\mathcal{S}} = \boldsymbol{\lambda}^{\mathcal{S}} \cup \boldsymbol{\Phi}$.

The geocentrically measured signal waveforms, parameterized by $\boldsymbol{\lambda}^{\mathcal{S}}$, can be projected onto the i-th detector using the detector's unique time-delay $\Delta t_i(\alpha, \delta)$, along with the antenna beam pattern response functions $F_i^+(\alpha, \delta, \psi)$ and $F_i^+(\alpha, \delta, \psi)$ of the plus (+) and cross (×) polarizations[2]. Mathematically we write [85, 88]

$$\tilde{h}_i(f; \boldsymbol{\lambda}^{\mathcal{S}}, \boldsymbol{\Phi}, N^{\mathcal{S}}) = \left(F_i^+ \tilde{h}_+ + F_i^\times \tilde{h}_\times\right) e^{2\pi i f \Delta t_i}, \tag{2.3}$$

where $\tilde{h}_p$ denotes the Fourier transform of the time domain geocentric GW signal, $h_p(t)$

[2]The antenna pattern functions F^+ and $F^\times$ are typically periodic functions of time due to the diurnal rotational motion of Earth [127]. However, the time dependence is omitted here with the assumption that the antenna patterns are constant over the short duration of GW burst. This assumption is conventional across all burst searches.

for polarization p. The version of *BayesWave* used in Chapters 3 and 4 assumes elliptical polarization such that the ellipticity parameter ϵ maps $\tilde{h}_+$ to the cross polarization $\tilde{h}_\times$ via

$$\tilde{h}_\times = \epsilon \tilde{h}_+ e^{i\pi/2}, \tag{2.4}$$

and $\tilde{h}_+$ is expressed as a linear combination of sine-Gaussian wavelets in the frequency domain as in Equation 2.2:

$$\tilde{h}_+(f) = \sum_{n=1}^{N^{\mathcal{S}}} \tilde{\Psi}(f; \boldsymbol{\lambda}_n). \tag{2.5}$$

Alternative signal polarization models offered by *BayesWave* are discussed in Chapter 5.

2.2.3. Glitch model, $\mathcal{G}$

Unlike GW signals, instrumental glitches and noise are uncorrelated across the detector network. Therefore the glitch model uses independent sets of wavelets to reconstruct glitches in each detector. Let $N^{\mathcal{G}_i}$ denote the number of wavelets and $\boldsymbol{\lambda}^{\mathcal{G}_i} = \boldsymbol{\lambda}^i_1 \cup \boldsymbol{\lambda}^i_2 \cdots \cup \boldsymbol{\lambda}^i_{N^{\mathcal{G}_i}}$ be the set of wavelet parameters used in the glitch model reconstruction of detector i. We can then write the glitch model for the i-th detector as

$$g(\boldsymbol{\lambda}^{\mathcal{G}_i}, N^{\mathcal{G}_i}) = \sum_{n=1}^{N^{\mathcal{G}_i}} \tilde{\Psi}(f; \boldsymbol{\lambda}^i_n). \tag{2.6}$$

Thus for a network with $\mathcal{I}$ detectors, the complete set of glitch model wavelet parameters is given by $\boldsymbol{\theta}^{\mathcal{G}} = \boldsymbol{\lambda}^{\mathcal{G}_1} \cup \boldsymbol{\lambda}^{\mathcal{G}_2} \cdots \cup \boldsymbol{\lambda}^{\mathcal{G}_\mathcal{I}}$. Note that the glitch model does not include extrinsic parameters for projecting the non-Gaussianity between detectors, because glitches are non-astrophysical and independent in each detector [85].

2.2.4. Gaussian-noise model, $\mathcal{N}$

In contrast to $\mathcal{S}$ and $\mathcal{G}$ which model non-Gaussian transient components of the detector data, *BayesWave* also offers a Gaussian-noise model $\mathcal{N}$ which treats the data as pure Gaussian noise. This model is used to characterize the stationary detector noise floors, i.e. PSDs[3]. Gaussian noise sources in the LIGO and Virgo detectors can be classified into three broad frequency bands: (i) seismic noise ($\sim$10 Hz), (ii) thermal noise ($\sim$10-200 Hz) and (iii) quantum (photon) shot noise ($\gtrsim$200 Hz). Moreover, various aspects of the detector apparatus including mirror suspensions, calibration lines, and the AC electrical supply are recurrent sources of high-power, narrow-band spectral lines. The $\mathcal{N}$ model uses the *BayesLine* algorithm [128], integrated within *BayesWave*, to collectively characterize these noise features in the detector PSDs. *BayesLine* employs cubic splines to model smooth broad-band noise and Lorentzians for narrow-band line-like features. The mathematical details of *BayesLine* are incidental to this thesis, a complete description can be found in Ref. [128].

The $\mathcal{S}$, $\mathcal{G}$ and $\mathcal{J}$ also incorporate $\mathcal{N}$ to model the PSDs alongside the wavelet models for non-Gaussian feature(s), and are therefore known as composite models [1].

[3]The PSD is related to the ASD as PSD= ASD^2

2.2.5. *Joint model, $\mathcal{J}$*

The $\mathcal{S}$ and $\mathcal{G}$ models introduced above are disjoint hypotheses where *BayesWave* assumes the detector data contains either signals or glitches but not both. When glitches occur near a candidate GW signal, analyses using signal-only ($\mathcal{S}$) model in *BayesWave* may include coherent excess power from the glitches across the network causing the inferred waveform to deviate from the actual signal. The glitch-only ($\mathcal{G}$) model, on the other hand, treat the power spikes in the detectors independently in order to remove excess noise from the individual detector data. Thus the disjoint models may infer both the nearby glitch and signal as a single non-astrophysical power spike or mistaken as a section of the signal as a glitch, resulting in removal of signal power.

The joint ($\mathcal{J}$) model reconstructs the data with a presumption that both a signal and glitches can occur simultaneously. It considers coherent features in the detectors as part of the signal and uncorrelated excess power in independent detectors as glitches. Altogether one can write a complete set of joint model parameters as $\boldsymbol{\theta}^{\mathcal{J}} = \{\boldsymbol{\theta}^{\mathcal{S}} \cup \boldsymbol{\theta}^{\mathcal{G}}\}$. The model is shown to improve the fidelity of *BayesWave*'s waveform reconstruction. In a test study [2], the joint model accurately recovered both the GW signal and instrumental glitch separately, regardless of the overlap in time. The $\mathcal{J}$ model is used by the LVK community to subtract glitches overlapping with a CBC signal [98, 102, 129]. In this context, where the signal source is known (e.g. CBC), the wavelet-based $\mathcal{S}$ model is replaced with phenomenological waveforms to improve detection sensitivity. We do not discuss $\mathcal{J}$ further in the remainder of this thesis.

2.3. Sampling of model parameters

Given the joint data $\boldsymbol{d} = \{d_1, \cdots, d_{\mathcal{I}}\}$ of a multi-detector network comprising detectors $1, \cdots, \mathcal{I}$, one can infer the posterior probability $p(\boldsymbol{\theta}^{\mathcal{M}}|\boldsymbol{d}, \mathcal{M})$ of a model $\mathcal{M}$ parameterized by $\boldsymbol{\theta}^{\mathcal{M}}$ via Bayes' theorem:

$$p(\boldsymbol{\theta}^{\mathcal{M}}|\boldsymbol{d}, \mathcal{M}) = \frac{p(\boldsymbol{\theta}^{\mathcal{M}}|\mathcal{M})p(\boldsymbol{d}|\boldsymbol{\theta}^{\mathcal{M}}, \mathcal{M})}{p(\boldsymbol{d}|\mathcal{M})}. \tag{2.7}$$

Here, $p(\boldsymbol{\theta}^{\mathcal{M}}|\mathcal{M})$ denotes the prior distribution of the model parameters $\boldsymbol{\theta}^{\mathcal{M}}$ for $\mathcal{M} \in \{\mathcal{S}, \mathcal{G}, \mathcal{N}\}$, $p(\boldsymbol{d}|\boldsymbol{\theta}^{\mathcal{M}}, \mathcal{M})$ denotes the likelihood function and $p(\boldsymbol{d}|\mathcal{M})$ denotes the model evidence, given by

$$p(\boldsymbol{d}|\mathcal{M}) = \int d\boldsymbol{\theta}^{\mathcal{M}} p(\boldsymbol{\theta}^{\mathcal{M}}|\mathcal{M})p(\boldsymbol{d}|\boldsymbol{\theta}^{\mathcal{M}}, \mathcal{M}). \tag{2.8}$$

In essence, the evidence is the likelihood of producing the data $\boldsymbol{d}$ from the hypothesis $\mathcal{M}$ marginalized over the parameter space of $\boldsymbol{\theta}^{\mathcal{M}}$, so it is otherwise known as the marginalized likelihood [85].

BayesWave uses a trans-dimensional Reversible-Jump Markov Chain Monte Carlo (RJMCMC) algorithm to sample the model parameters, which includes the number of wavelets (i.e. model complexity) and the wavelet parameters. The RJMCMC requires a likelihood function and model parameter priors. We overview these key components in the two following subsections, along with the techniques used to implement the RJMCMC in Section 2.3.3. We then discuss how the wavelet parameter posteriors are used to construct a waveform posterior that represents the data, in Section 2.3.4.

2.3.1. The likelihood function

In *BayesWave*, the likelihood function is given by [1, 130]

$$p(\boldsymbol{d}|\boldsymbol{\theta}^{\mathcal{M}}, \mathcal{M}) = \frac{1}{\det(\pi \boldsymbol{\mathcal{C}})} e^{-\tilde{\boldsymbol{r}}^{\mathrm{T}} \boldsymbol{\mathcal{C}}^{-1} \tilde{\boldsymbol{r}}}, \tag{2.9}$$

where $\tilde{\boldsymbol{r}} = \tilde{\boldsymbol{d}} - \tilde{\boldsymbol{h}} - \tilde{\boldsymbol{g}}$ is the Fourier domain residual upon subtracting the GW signal $\tilde{\boldsymbol{h}}$ and glitch $\tilde{\boldsymbol{g}}$ models from the data, $\tilde{\boldsymbol{r}}^{\mathrm{T}}$ is the conjugate transpose of $\tilde{\boldsymbol{r}}$, and $\boldsymbol{C}$ is the noise covariance matrix.

The likelihood function in Equation 2.9 is defined under the assumption that $\tilde{\boldsymbol{r}}$ is consistent with the Gaussian-noise $\tilde{\boldsymbol{n}}$ modeled by $\mathcal{N}$, where the noise is temporally stationary and uncorrelated across all frequencies. As a result, $\boldsymbol{\mathcal{C}}$ is a diagonal matrix, viz. for a single detector with PSD $S_n(f)$,

$$\boldsymbol{\mathcal{C}}_{k\ell} = S_n(f_k)\delta_{k,\ell}. \tag{2.10}$$

Here, $\delta_{k,\ell}$ denotes the Kronecker delta function and the subscripts k, ℓ denote the indices of the discrete frequency samples. For multi-detector networks, $\boldsymbol{\mathcal{C}}$ becomes a block diagonal matrix by construction, but the overall matrix remains diagonal because noise in the detectors are also expected to be independent from one another.

2.3.2. Model parameter priors

The model parameters include the number of wavelets, the intrinsic wavelet parameters and extrinsic parameters for the signal model. By default, the current version of *BayesWave* [2] uses an empirical prior for the number of wavelets, N.

$$p(N) = \frac{4\sqrt{3}N}{2\pi b^2 \left(3 + \frac{N}{b}\right)^4} \tag{2.11}$$

with $b = 2.9$. Equation 2.11 is derived using an analytic fit to the N-distribution of the 500 most significant background triggers from *BayesWave*'s all-sky burst analysis of O1 data [44]. Alternatively, if preferred, one can still opt for the flat N prior, as in the original *BayesWave* configuration [1]. The range of N is adjustable for the $\mathcal{S}$ and $\mathcal{G}$ models, but the Gaussian-noise model $\mathcal{N}$ insists that $N = 0$.

All intrinsic wavelet parameters, except for the wavelet amplitude A, have flat priors covering the range: t_0 within a user-selected time window, typically 1s and centered on the time of the candidate event; f_0 within the analysis frequency band, e.g. 20–1024 Hz; $Q \in \{0.1, 40\}$ by default, but adjustable by the user; and $\phi_0 \in \{0, 2\pi\}$. Similarly for the extrinsic parameters of $\mathcal{S}$: the source location prior is uniform across the sky, $\alpha \in \{0, 2\pi\}$ and $\sin\delta \in \{-1, 1\}$; the flat priors for polarization angle and ellipticity respectively cover the range $\phi \in \{0, \pi\}$ and $\epsilon \in \{0, 1\}$. If the source has been localized through other methods, such as EM observations, *BayesWave* provides the option to fix the sky location, which can help the sampler produce more accurate estimates of other parameters.

The wavelet amplitude prior is derived from a physically-motivated SNR prior, constrained by the SNR of real GW observations. The amplitude of a glitch wavelet in a

single detector with PSD $S_n(f_0)$ can be estimated from its SNR using

$$A^2 \simeq \frac{2\sqrt{2\pi} f_0 S_n(f_0)}{Q} \mathrm{SNR}^2. \tag{2.12}$$

Equation 2.12 also applies to a signal wavelet. However, in a network with $\mathcal{I} \geq 2$ detectors sharing the same signal wavelet, the detector-dependent $S_n(f_0)$ in Equation 2.12 is replaced by the network average:

$$\overline{S}_n(f_0) = \left(\sum_{i=1}^{\mathcal{I}} \frac{F_{+,i}^2 + \epsilon^2 F_{\times,i}^2}{S_{n,i}(f_0)} \right)^{-1}. \tag{2.13}$$

Here, $S_{n,i}(f_0)$ and $F_{p,i}$ respectively denote the PSD and antenna pattern functions for polarization p of the i-th detector. Observations of astrophysical signals and instrumental glitches have shown that low SNR events are more common than louder ones. However, one also has to consider that low amplitude, i.e. low SNR ($\lesssim 3$), wavelets have minimal impact on the likelihood and have a lower chance of being accepted into the sampling chain. Altogether, the prior should disfavor wavelets with either very high or very low SNRs. To achieve this, the SNR prior is constructed to peak at a nominal $\mathrm{SNR}_* = 5$, with the probability decreasing asymmetrically from the peak: the decrease is steeper at low SNRs (< 5) than at high SNRs (> 5). However, loud astrophysical signals are less common than loud instrumental glitches [44], so *BayesWave* uses separate SNR priors for the signal and glitch models. The prior distribution for the glitch model is given by [2]

$$p(\mathrm{SNR}|\mathcal{G}) = \frac{\mathrm{SNR}}{2\mathrm{SNR}_*^2 \left(1 + \frac{\mathrm{SNR}}{2\mathrm{SNR}_*}\right)^3}, \tag{2.14}$$

and the prior for the signal model is

$$p(\mathrm{SNR}|\mathcal{S}) = \frac{3\mathrm{SNR}}{4\mathrm{SNR}_*^2 \left(1 + \frac{\mathrm{SNR}}{4\mathrm{SNR}_*}\right)^5}. \tag{2.15}$$

Equation 2.15 (signal prior) falls off faster as SNR^{-4}, compared to SNR^{-2} in Equation 2.14 (glitch prior), reflecting the lower probability of astrophysical signals relative to instrumental glitches at high SNRs. Furthermore, the SNR^{-4} scaling in the signal model prior is consistent with the SNR distribution of astrophysical sources [86].

2.3.3. Posterior sampling with the RJMCMC algorithm

Using a transdimensional-RJMCMC algorithm, the number of wavelets used in *BayesWave* models along with the wavelet parameters are sampled on the fly to accommodate the intricacies of the data. While RJMCMC facilitates the detection of generic GW bursts without *a priori* knowledge of the waveform model, it is the most computationally demanding aspect of *BayesWave*. Run times for a single event can range from hours to days, depending on analysis settings e.g. number of sampler iterations and duration of the data being analyzed. Below, we outline the various techniques used by *BayesWave* to improve sampling efficiency.

Maximization during burn-in

The initial state of the sampling chain is arbitrarily defined. Thus the sampler requires a 'burn-in' period to reach a stationary state where it accurately reflects the target (posterior) distribution. Samples from the burn-in period are discarded to ensure that the final distribution is independent of the chain's initial state. In *BayesWave*, t_0, ϕ_0 and A for each wavelet is maximized during the burn-in period, to accelerate signal/glitch identification and convergence to the target distribution.

Parallel tempering

BayesWave also implements parallel tempering [131] to speed-up convergence. Using this approach, Markov chains at different temperatures T run in parallel. The temperature scales the likelihood function of each chain, viz. $p(\boldsymbol{d}|\mathcal{M}) \to p(\boldsymbol{d}|\mathcal{M})^{1/T}$ for model $\mathcal{M}$. This scaling suggests that hotter (high T) chains have flatter likelihood functions and are able to explore more of the prior volume, whereas colder chains are better for exploring localized peaks. A Metropolis-Hastings algorithm then is used to initiate state exchange between hot and cold chains. A successful exchange allows colder chains to escape local minima by adopting states discovered by hotter chains, while hotter chains can refine their exploration using states from colder chains. Altogether, parallel tempering ensures efficient and thorough exploration of the parameter space. In Section 2.4, we discuss how parallel tempering also facilitates the evaluation of the *BayesWave* model evidences.

Fisher matrix proposals for wavelet parameters

Another way of improving sampler efficiency is to ensure that the proposed (trial) parameters has a high acceptance rate. *BayesWave* leverages the Fisher Information Matrix (FIM) to construct custom proposal distributions for updating wavelet parameters while maximizing the acceptance rate. The FIM Γ of sine-Gaussian wavelets is known analytically; in terms of the five intrinsic parameters $\{t_0, f_0, Q, A, \phi_0\}$, it is expressed as follows:

$$\Gamma = \mathrm{SNR}^2 \begin{pmatrix} \frac{4\pi^2 f_0^2(1+Q^2)}{Q^2} & 0 & 0 & 0 & -2\pi f_0 \\ 0 & \frac{3+Q^2}{4f_0^2} & -\frac{3}{4Qf_0} & -\frac{1}{2f_0} & 0 \\ 0 & -\frac{3}{4Qf_0} & \frac{3}{4Q^2} & \frac{1}{2Q} & 0 \\ 0 & -\frac{1}{2f_0} & \frac{1}{2Q} & 1 & 0 \\ -2\pi f_0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.16}$$

By definition, the elements of the FIM for a wavelet Ψ are given by $\Gamma_{mn} \equiv \left(\frac{\partial\Psi}{\partial\lambda_m}|\frac{\partial\Psi}{\partial\lambda_n}\right)$. That is, the FIM encodes information on the local curvature of the wavelet parameter likelihood. One can therefore improve convergence by proposing jumps along the eigenvectors of the FIM, towards regions of higher likelihood [1]. In *BayesWave*, the

multivariate Gaussian proposals used to draw updates to the wavelet parameters are constructed using the FIMs:

$$q(\mathbf{y}|\mathbf{x}) = \frac{\det \Gamma}{(2\pi)^2} \exp\left(-\frac{1}{2}\Gamma_{ij}\Delta x^i \Delta x^j\right), \tag{2.17}$$

Here, $\Delta x^i = x^i - y^i$ denotes the displacement in the intrinsic parameter i, where x^i and y^i are the values before and after the update respectively.

Proposals for extrinsic parameters are also generated using FIMs. However, the signal model derivatives with respect to the extrinsic parameters cannot be expressed analytically due to degeneracies between the parameters and non-Gaussian, multimodal likelihood distributions. Thus, the FIMs are evaluated numerically using finite difference approximations [1]. With two-detector networks, sky locations are poorly constrained. Thus to accelerate convergence, proposals for sky location are further constrained within a degenerate ring which preserves the constant time-delay between the two detectors. As the detector network expands, the sky-ring proposal becomes less effective in facilitating convergence. However, this is not a major concern, because convergence improves naturally as the source location is better constrained with the addition of more detectors.

Trans-dimensional and within-dimension proposals

BayesWave also leverages domain knowledge to construct two custom proposals for sampling trans-dimensional moves (adding or removing a wavelet) and within-dimension moves (simultaneously adding and removing a wavelet): (i) a data-driven $\{t, f, Q\}$ (TFQ) proposal [2], and (ii) a proximity proposal [1]. We detail their formulation and motivation below.

The TFQ proposal guides the placement of wavelets in the $\{t, f, Q\}$ space. To construct the TFQ proposal, we first consider a constant-Q transform, which produces a map of the excess power relative to the average detector noise floor in the $\{t, f\}$ (TF) space, at a fixed Q [132, 133]. This TF excess-power map can be converted into a two-dimensional probability density function (PDF) by normalizing it to the total power. Similarly, by stacking constant-Q transforms at multiple Q values and applying an overall normalization, one constructs the TFQ proposal: a three-dimensional PDF from which the wavelet parameters (t_0, f_0, Q) are drawn. The TFQ proposal facilitates trans-dimensional moves by preferentially adding wavelets to regions with high power and removing wavelets from regions with low power in the TFQ space. For within-dimension moves, the new wavelet is also fairly drawn from the TFQ proposal to replace the removed wavelet.

Burst sources like CBCs and CCSNs, as well as instrumental glitches, are known to produce localized cluster(s) of power in the time-frequency plane. Therefore, compact clusters of wavelets are generally preferred over scattered wavelets. However, favoring clustered wavelet configurations could lead to over-clustering, i.e. placing unnecessary wavelets that overlap extensively in both time and frequency with others. To balance these considerations, *BayesWave* employs a proximity proposal that encourages placement of new wavelets near an existing cluster while minimizing their time-frequency overlap. To construct the proximity proposal, we first define the Gaussian-like PDF of a

wavelet in the time-frequency space:

$$p_{\text{Gaussian}}(t,f) = \frac{1}{2\pi\sigma_f\sigma_t} e^{-(\Delta f^2/2\sigma_f^2 + \Delta t^2/2\sigma_t^2)}, \tag{2.18}$$

where $\Delta f = f_0 - f$, $\Delta t = t_0 - t$, $\sigma_t = \tau$ and $\sigma_f = 1/\pi\tau$. To promote clustering, the probability of wavelet placement in the time-frequency plane should be high near existing wavelets, while to minimize overlap, it should be low in regions already occupied by other wavelets. This PDF is represented using a hollowed-out Gaussian:

$$p(t,f) = \frac{1}{2\pi\sigma_f\sigma_t(\mu^2-\nu^2)} \left(e^{-(\Delta f^2/2\sigma_f^2 + \Delta t^2/2\sigma_t^2)/\mu^2} - e^{-(\Delta f^2/2\sigma_f^2 + \Delta t^2/2\sigma_t^2)/\nu^2} \right), \tag{2.19}$$

where $\mu > \nu$ controls the width of the PDF, and ν controls the width of the hollowed-out region. Equation 2.19 is, however, only valid for a single wavelet. For a *BayesWave* model with N wavelets, the proximity proposal for the $(N+1)$-th wavelet is given by:

$$q_{\text{proxi}}(t,f) = \left[\frac{(1-\gamma)}{N} \sum_{n=1}^{N} p_n(t,f) \right] + \frac{\gamma}{V_{\text{TF}}}, \tag{2.20}$$

where $p_n(t,f)$ denotes Equation 2.19 for the n-th wavelet. The second term in Equation 2.20 is a uniform density component that allows for some degree of unbiased exploration across the entire time-frequency volume V_{TF}. The fractional contributions of the hollowed-out Gaussians and the uniform density component to the proximity proposal are quantified by $1-\gamma$ and γ respectively. The *BayesWave* proximity proposal adopts the values $\mu = 2$, $\nu = 0.5$ and $\gamma = 0.5$.

When adding or replacing a wavelet, the custom proposals only initialize the parameters $\{t_0, f_0, Q\}$; the remaining parameters, A and ϕ_0, are sampled from their respective priors. For the signal ($\mathcal{S}$) model, the new wavelet inherits the same extrinsic parameters as the existing wavelets. If the signal model does not have any wavelets yet, the extrinsic parameters are drawn from their respective prior distributions.

2.3.4. Waveform posteriors

So far, we have discussed how *BayesWave* samples the posterior distributions of the wavelet parameters. Ultimately, however, we are interested in the reconstructed waveform and its overall properties. For each sample in the posterior, *BayesWave* sums the wavelets with the corresponding parameters to produce a signal or glitch waveform. The signal waveform shown in Figure 2.1 is one of the many reconstructions from *BayesWave*'s waveform posterior. To visualize the waveform posterior, the top panel of Figure 2.2 shows eight arbitrarily chosen waveform samples, each with a different color, for the same event (i.e. GW150914). However, *BayesWave* wavelet parameters, and hence waveform posteriors, can have up to 10^6 samples. Thus, to summarize the overall distribution of waveform samples, *BayesWave* calculates the median and the 50% and 90% credible interval (C.I.) bounds at every time step[4] t_i of the waveforms $h(t_i)$. An example of this, also for GW150914, is shown the bottom panel of Figure 2.2.

[4]The interval of the time steps is determined by the user-selected sampling rate.

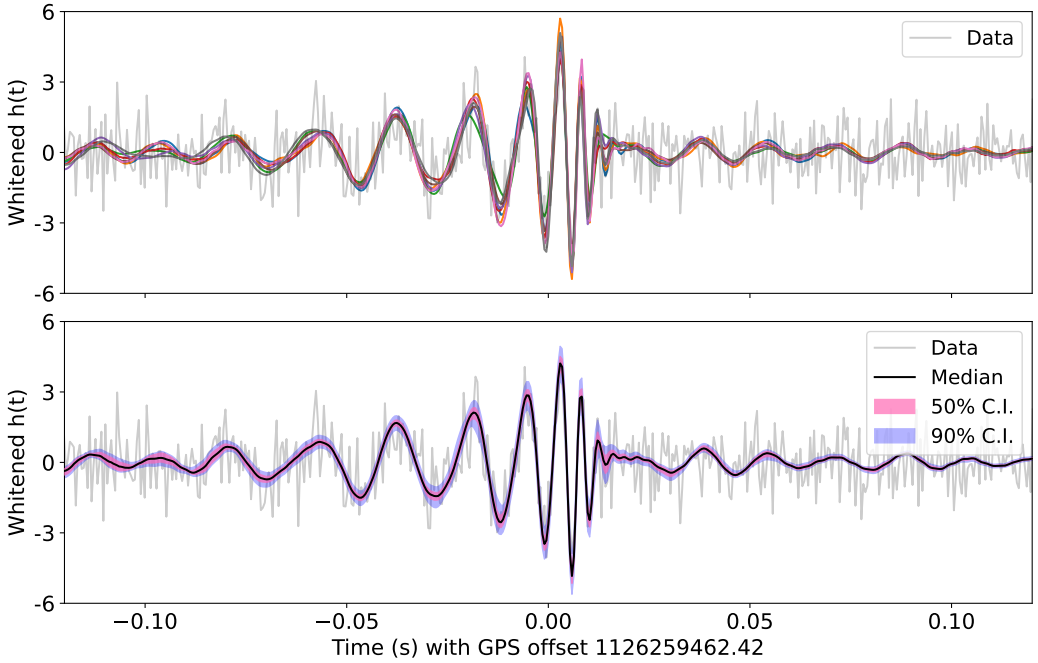

Figure 2.2: Visualization of the waveform posterior. The top panel shows eight time-domain waveform samples, each of different color, from the *BayesWave* signal waveform posterior of GW150914. The bottom panel shows the median (solid black curve), the 50% credible interval (C.I., blue shaded bands) and the 90% C.I. (pink shaded bands) of the same waveform posterior as in the top panel.

2.4. Bayesian model selection

2.4.1. *Bayes factor as detection statistic*

Given two models $\mathcal{M}_1$ and $\mathcal{M}_2$, one can measure which of the two is more strongly supported by the data $\boldsymbol{d}$, by computing the (posterior) odds ratio. This process is called model selection. According to Bayes' theorem, the odds ratio is given by

$$\frac{p(\mathcal{M}_1|\boldsymbol{d})}{p(\mathcal{M}_2|\boldsymbol{d})} = \frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)} \frac{p(\boldsymbol{d}|\mathcal{M}_1)}{p(\boldsymbol{d}|\mathcal{M}_2)}. \tag{2.21}$$

The right-hand side of Equation 2.21 comprises the prior odds between $\mathcal{M}_1$ and $\mathcal{M}_2$, and their evidence ratio

$$\mathcal{B}_{1,2}(\boldsymbol{d}) = \frac{p(\boldsymbol{d}|\mathcal{M}_1)}{p(\boldsymbol{d}|\mathcal{M}_2)}, \tag{2.22}$$

otherwise known as the Bayes factor. Assuming equal prior probability, i.e. $p(\mathcal{M}_1) = p(\mathcal{M}_2)$, one can perform model selection directly using the Bayes factor. $\mathcal{B}_{1,2}(\boldsymbol{d}) > 1$ suggests that $\mathcal{M}_1$ is more strongly supported by the data and vice versa.

BayesWave uses the log Bayes factor, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$, between the $\mathcal{S}$ and $\mathcal{G}$ models to quantify its confidence in distinguishing between GW signals and instrumental glitches. In other words, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ is the detection statistic of *BayesWave*. However, obtaining model

evidences directly from the integral in Equation 2.8 is computationally expensive, especially for complex and highly parameterized models. To overcome the computational challenges, *BayesWave* combines the parallel-tempered RJMCMC algorithm [131] and thermodynamic integration [134] to obtain the evidences required for the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ calculation. Implementations of these methods are detailed in Appendix 5.A1.

2.4.2. *Occam penalty* [5]

A unique property of the Bayes factor is that it inherently accounts for model complexity during model selection. When two competing models fit the data equally well, the Bayes factor penalizes the more complex model for potential overfitting. This behavior reflects Occam's razor, which posits that simpler explanations are generally preferred. The mathematical interpretation of this *Occam penalty* is detailed below.

Let $\boldsymbol{\theta}_{\rm MAP}$ denote the maximum a posteriori (MAP) value, i.e. the mode of the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{d},\mathcal{M})$. For strong detection candidates with SNR $\gtrsim 10$, the integrand of the model evidence in Equation 2.8 is expected to peak around $\boldsymbol{\theta}_{\rm MAP}$. One can then apply the Laplace-Fisher approximation to estimate the model evidence as [85]

$$p(\boldsymbol{d}|\mathcal{M}) \simeq p(\boldsymbol{d}|\boldsymbol{\theta}_{\rm MAP},\mathcal{M})p(\boldsymbol{\theta}_{\rm MAP}|\mathcal{M})(2\pi)^{D/2}\sqrt{\det C}. \tag{2.23}$$

In Equation 2.23, $p(\boldsymbol{d}|\boldsymbol{\theta}_{\rm MAP},\mathcal{M})$ is the MAP likelihood; $p(\boldsymbol{\theta}_{\rm MAP}|\mathcal{M})$ is the prior evaluated at the MAP parameter values; D is the model dimension; and $\det C$ is the determinant of the full covariance matrix C for the N wavelets used in waveform reconstruction[6]. If the parameter spaces of all N wavelets overlap minimally, the full covariance matrix becomes block-diagonal, and its determinant simplifies to

$$\det C = \prod_{n=1}^{N} \det C_n, \tag{2.24}$$

where C_n denotes the covariance matrix of a single wavelet. Consequently, C can be approximated as the inverse of the N-wavelet FIM $\overline{\Gamma}$, allowing $\det C$ to be replaced by $1/\det\overline{\Gamma}$ [135].

We now interpret the physical significance of the factors in Equation 2.23. Assuming uniform priors over all intrinsic wavelet parameters, the second factor in Equation 2.23 simplifies to $p(\boldsymbol{\lambda}_{\rm MAP}|\mathcal{M}) = 1/V_{\mathcal{M}}$, where $V_{\mathcal{M}}$ is the total intrinsic parameter space volume. By definition, $\det C$ measures the variance of the likelihood, and thus $\sqrt{\det C}$ quantifies the corresponding spread. As a result, the product of $\sqrt{\det C}$ and $(2\pi)^{D/2}$ measures the volume of the uncertainty ellipsoid, i.e. the posterior volume $\Delta V_{\mathcal{M}}$, for a given model $\mathcal{M}$ [53, 85, 136]. Collectively, the last three factors of Equation 2.23 represent the fraction of the prior volume occupied by the posterior $\Delta V_{\mathcal{M}}/V_{\mathcal{M}}$, and Equation 2.23 can be re-expressed as:

$$p(\boldsymbol{d}|\mathcal{M}) \simeq p(\boldsymbol{d}|\boldsymbol{\theta}_{\rm MAP},\mathcal{M})\frac{\Delta V_{\mathcal{M}}}{V_{\mathcal{M}}}. \tag{2.25}$$

[5] This section is adapted with substantial rewording from Section IV.A of Ref. [3], which originally appeared in the thesis submitted for the degree of Master of Science (Physics) at The University of Melbourne, awarded to the author in 2020.

[6] The covariance matrix C for the wavelet parameters is not to be confused with the noise covariance matrix $\boldsymbol{C}$ in Equation 2.9.

The fraction $\Delta V_{\mathcal{M}}/V_{\mathcal{M}}$ is known as the "Occam penalty factor" [3].

Combining Equations 2.22 and 2.25, the Bayes factor which compares the evidences between two models $\mathcal{M}_1$ and $\mathcal{M}_2$ is given by:

$$\mathcal{B}_{1,2}(\boldsymbol{d}) \simeq \Lambda_{1,2}(\boldsymbol{d}) \frac{\Delta V_1}{V_1} \frac{V_2}{\Delta V_2}, \tag{2.26}$$

where $\Lambda_{1,2}(\boldsymbol{d}) = p(\boldsymbol{d}|\boldsymbol{\theta}_{\mathrm{MAP},1})/p(\boldsymbol{d}|\boldsymbol{\theta}_{\mathrm{MAP},2})$ is the ratio of MAP likelihoods. Equation 2.26 shows that the Bayes factor is depends directly on both the likelihood ratio and the ratio of the Occam penalty factors. If two models occupy the same posterior volume, i.e. $\Delta V_1 = \Delta V_2$ and represent the data equally well, i.e. $\Lambda_{1,2}(\boldsymbol{d}) \approx 1$, the Occam factor suppresses the evidence for the model with the larger prior volume, effectively penalizing it for introducing unnecessary complexity in fitting the data. It is important to clarify that the Occam penalty is not intentionally imposed, but rather emerges naturally from the use of Bayes' Theorem.

To provide intuition on how the Occam penalty facilitates the separation of signals from glitches in *BayesWave*, we discuss the model selection problem heuristically as follows. First, we consider the dimensionality of the signal model ($\mathcal{S}$), where all detectors share the same intrinsic wavelet parameters, plus four extrinsic parameters. Each wavelet has five intrinsic parameters (t_0, f_0, Q, A, ϕ_0), so the dimensionality of $\mathcal{S}$ with N wavelets scales as

$$D_{\mathcal{S}} \sim 5N + 4. \tag{2.27}$$

The glitch model ($\mathcal{G}$), on the other hand, does not include extrinsic parameters. However, each detector has an independent glitch model, characterized by a unique set of intrinsic parameters. Suppose $\mathcal{S}$ and $\mathcal{G}$ use the same number of wavelets per detector, i.e. $N_i^{\mathcal{G}} = N^{\mathcal{S}} = N$ for all detectors i. Then for a network with $\mathcal{I}$ detectors, the dimensionality of $\mathcal{G}$ scales as [1]

$$D_{\mathcal{G}} \sim 5N\mathcal{I}. \tag{2.28}$$

For non-trivial events with $N \geq 1$, we observe that $D_{\mathcal{G}} > D_{\mathcal{S}}$ when $\mathcal{I} \geq 2$. In other words, for multi-detector networks, the glitch model generally occupies a larger parameter space volume than the signal model, i.e. $V_{\mathcal{G}} > V_{\mathcal{S}}$. If $\mathcal{S}$ and $\mathcal{G}$ fit the data equally well, viz. $\Lambda_{\mathcal{S},\mathcal{G}} \approx 1$ and $\Delta V_{\mathcal{S}} \approx \Delta V_{\mathcal{G}}$, then according to Occam's razor, the preference for $\mathcal{S}$ strengthens as $\mathcal{I}$ increases. This behavior is reflected in the Bayes factor:

$$\mathcal{B}_{\mathcal{S},\mathcal{G}}(\boldsymbol{d}) \simeq \Lambda_{B_{\mathcal{S},\mathcal{G}}}(\boldsymbol{d}) \frac{\Delta V_{\mathcal{S}}}{\Delta V_{\mathcal{G}}} \frac{V_{\mathcal{G}}}{V_{\mathcal{S}}} \tag{2.29}$$

where, if $V_{\mathcal{G}}$ increases more rapidly than $V_{\mathcal{S}}$ with increasing $\mathcal{I}$, then $\mathcal{B}_{\mathcal{S},\mathcal{G}}(\boldsymbol{d})$ also increases more rapidly in favor of $\mathcal{S}$. Looking ahead to Chapter 5, we use an analogous example to compare the preferences between two *BayesWave* signal polarization models.

2.5. Behavior of detection statistic

The detection statistic $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ is the fundamental quantifier of *BayesWave*'s performance in burst characterization. Ref. [85] derives an analytic expression for $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ to improve understanding of the factors that influence *BayesWave*'s performance in distinguishing

astrophysical signals from instrumental glitches. The analytic expression reveals that $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ primarily scales as

$$\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \propto N \ln(\mathrm{SNR}_{\mathrm{net}}), \tag{2.30}$$

with N denoting the number of wavelets used in the reconstruction, and $\mathrm{SNR}_{\mathrm{net}}$ denoting the overall network SNR, as given by Equation 1.28. The dependence of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ on N, which measures the waveform complexity in the time-frequency plane, distinguishes *BayesWave* from other unmodeled searches, whose detection statistics primarily scale with SNR. The scaling as in Equation 2.30 assumes a network comprising two GW detectors. However, the heuristic discussion in Section 2.4.2 suggests that size of the detector network also affects $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. In light of the expanding global detector network, Ref. [3] generalizes this scaling to accommodate an arbitrary number of detectors, $\mathcal{I}$. In Section 2.5.1, we rederive the analytic scaling of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ with $\mathcal{I}$, following the approach in Ref. [3]. In Section 2.5.2, we highlight the empirical results from Ref. [3], which are used to validate the analytic scaling.

2.5.1. *Analytic scaling* [7]

Ref. [85] applies the Laplace approximation, as shown in Equation2.23, to derive approximate expressions for the signal model evidence:

$$\ln p(\boldsymbol{d}|\mathcal{S}) \simeq \frac{\mathrm{SNR}^2_{\mathrm{net}}}{2} - \frac{5N^{\mathcal{S}}}{2} - N^{\mathcal{S}} \ln(V_\lambda) + \sum_{n=1}^{N^{\mathcal{S}}} \ln\left(\frac{\overline{Q}_n}{\mathrm{SNR}^5_{\mathrm{net},n}}\right) + \frac{D_\Phi}{2} + \ln \frac{\sqrt{\det C_\Phi}}{V_\Phi}, \tag{2.31}$$

and the glitch model evidence:

$$\ln p(\boldsymbol{d}|\mathcal{G}) \simeq \frac{\mathrm{SNR}^2_{\mathrm{net}}}{2} - \sum_{i=1}^{\mathcal{I}} \left[\frac{5N_i^{\mathcal{G}}}{2} + N_i^{\mathcal{G}} \ln(V_\lambda) - \sum_{n=1}^{N_i^{\mathcal{G}}} \ln\left(\frac{\overline{Q}_n}{\mathrm{SNR}^5_{i,n}}\right)\right]. \tag{2.32}$$

In Equations 2.31 and 2.32, $\overline{Q}_n \equiv (2\pi)^{5/2} \frac{\sqrt{2}Q_n}{\pi}$, and V_λ denotes the prior volume associated with the intrinsic wavelet parameters. The subscript n denotes the wavelet index within a wavelet model (either $\mathcal{S}$ or $\mathcal{G}$), while i denotes the detector index within $\mathcal{G}$. Accordingly, $\mathrm{SNR}_{\mathrm{net},n}$ in Equation 2.31 denotes the network SNR of the n-th wavelet in $\mathcal{S}$; $\mathrm{SNR}_{i,n}$ and $N_i^{\mathcal{G}}$ in Equation 2.32 respectively denote the SNR of the n-th wavelet and the total number of wavelets used in $\mathcal{G}_i$. In Equation 2.31, D_Φ=4, C_Φ, and V_Φ denote the dimension, covariance matrix, and the prior volume of extrinsic parameters respectively.

To derive the scaling of the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ with respect to the waveform properties and the number of detectors $\mathcal{I}$, we begin by simplifying Equations 2.31 and 2.32. The simplification is based on the following assumptions introduced in Ref. [85], which apply to loud candidates with SNR $\gtrsim 10$:

(1) The observed signal strength is comparable across all detectors, so the SNR for each detector i is given by:

$$\mathrm{SNR}_{i,n} \approx \frac{\mathrm{SNR}_{\mathrm{net},n}}{\sqrt{\mathcal{I}}}. \tag{2.33}$$

[7] This section is adapted with substantial rewording from Section IV.B and Appendix B of Ref. [3], which originally appeared in the thesis submitted for the degree of Master of Science (Physics) at The University of Melbourne, awarded to the author in 2020.

(2) For all detectors i, the glitch model $\mathcal{G}_i$ uses reconstruction parameters similar to those of the signal model $\mathcal{S}$. As a result, both models share the same number of wavelets per detector, i.e. $N_i^{\mathcal{G}} \approx N^{\mathcal{S}} \equiv N$, and their wavelets also have the same quality factor, viz. $Q_{i,n}^{\mathcal{G}} \approx Q_n^{\mathcal{S}} \equiv Q$.

(3) All wavelets have the same SNR, such that

$$\mathrm{SNR}_{\mathrm{net},n} \approx \frac{\mathrm{SNR}_{\mathrm{net}}}{\sqrt{N}}. \tag{2.34}$$

This assumption is supported by empirical observations [85].

(4) Given that $N_i^{\mathcal{G}}$ is the number of wavelets in the glitch model for a single detector i, Assumption (2) implies that the total number of wavelets in the glitch model across a network of $\mathcal{I}$ detectors can be approximated as $\mathcal{I}N_i^{\mathcal{G}} \approx \mathcal{I}N$.

Under these assumptions, Equations 2.31 and 2.32 respectively reduce to:

$$\ln p(\boldsymbol{d}|\mathcal{S}) \simeq \frac{\mathrm{SNR}_{\mathrm{net}}^2}{2} - N\left[\frac{5}{2} - \ln(V_\lambda) + \ln\left(\overline{Q}\right) - 5\ln\left(\frac{\mathrm{SNR}_{\mathrm{net}}}{\sqrt{N}}\right)\right] + 2 + \ln\frac{\sqrt{\det C_\Phi}}{V_\Phi}, \tag{2.35}$$

and

$$\ln p(\boldsymbol{d}|\mathcal{G}) \simeq \frac{\mathrm{SNR}_{\mathrm{net}}^2}{2} - \mathcal{I}N\left[\frac{5}{2} + \ln(V_\lambda) - \ln\left(\overline{Q}\right) + 5\ln\left(\frac{\mathrm{SNR}_{\mathrm{net}}}{\sqrt{N\mathcal{I}}}\right)\right], \tag{2.36}$$

with $\overline{Q} \equiv (2\pi)^{5/2}\frac{\sqrt{2}Q}{\pi}$.

By taking the difference between Equations 2.35 and 2.36, one obtains the analytic expression for the log Bayes factor, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \simeq \ln p(\boldsymbol{d}|\mathcal{S}) - \ln p(\boldsymbol{d}|\mathcal{G})$, for a network of $\mathcal{I}$ detectors:

$$\begin{aligned}\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \simeq N(\mathcal{I}-1)\left[\frac{5}{2} + \ln(V_\lambda) - \ln\left(\overline{Q}\right) + 5\ln\left(\frac{\mathrm{SNR}_{\mathrm{net}}}{\sqrt{N}}\right)\right] \\ - \frac{5}{2}\mathcal{I}N\ln(\mathcal{I}) + \left(2 + \ln\frac{\sqrt{\det C_\Phi}}{V_\Phi}\right).\end{aligned} \tag{2.37}$$

Equation 2.37 demonstrates that $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ scales primarily with the number of detectors $\mathcal{I}$, number of wavelets N (waveform complexity), and network SNR, viz.

$$\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \propto \mathcal{I}N\ln(\mathrm{SNR}_{\mathrm{net}}). \tag{2.38}$$

The additional scaling factor $\mathcal{I}$ distinguishes Equation 2.38 from Equation 2.30.

Equation 2.38 is similar to Equation 2.30 in that $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ increases with N, indicating a stronger preference for the signal model over the glitch model as the waveform complexity increases. However, Equation 2.38 further shows that, for a fixed SNR and N, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ increases with the number of detectors in the network. This behavior is consistent with the heuristic arguments in Section 2.4.2, where adding detectors increases the dimensionality of the glitch model more rapidly than that of the signal model, thereby strengthening the preference for the simpler signal model.

2.5.2. *Empirical validation*

BBH injections[8]

Ref. [3] uses simulated BBHs to empirically validate the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ scaling given by Equation 2.38. The analysis dataset comprises 150 non-spinning BBHs, modeled using the phenomenological waveform approximant `IMRPhenomD` [137, 138]. The BBHs are injected into Gaussian detector noise, colored by projected PSDs of detector configurations of different sizes, and are recovered using *BayesWave*. The detector configurations include the Hanford-Livingston (HL, two-detector), HL-Virgo (HLV, three-detector), and HL-KAGRA-Virgo (HLKV, four-detector) networks.

Figure 2.3 presents the results from this study. The top left panel shows $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ as a function of the injected $\mathrm{SNR}_{\mathrm{net}}$ for the HL, HLV and HLKV networks. As expected, all three networks show a clear trend where $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ increases with $\mathrm{SNR}_{\mathrm{net}}$. Additionally, for a fixed SNR, larger detector networks generally yield higher values of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$; as $\mathcal{I}$ increases from 2 (HL) to 4 (HLKV), the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ also approximately doubles across all values of $\mathrm{SNR}_{\mathrm{net}}$. This provides evidence that $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ scales directly with $\mathcal{I}$, in addition to $\mathrm{SNR}_{\mathrm{net}}$, as predicted by Equation 2.38.

Ref. [3] also explores the impact of $\mathrm{SNR}_{\mathrm{net}}$ and $\mathcal{I}$ on N, which, in turn, influences the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. Recall from Section 2.3.4 that, for a single event, *BayesWave* outputs a waveform posterior comprising waveforms constructed with different numbers of wavelets, N. The median N from the posterior serves as a representative measure of an event's waveform complexity. The top right panel of Figure 2.3 plots the median N as a function of $\mathrm{SNR}_{\mathrm{net}}$, showing that N increases systematically with $\mathrm{SNR}_{\mathrm{net}}$ across all three detector networks. This observation is expected because higher SNRs reveal more intricate and complex structures in the waveform, enabling *BayesWave* to place more wavelets with increased confidence. Since the same BBH injections are used across the three detector configurations, similar linear trends in N versus $\mathrm{SNR}_{\mathrm{net}}$ are observed. However, slight deviations from linearity are also noted at $\mathrm{SNR}_{\mathrm{net}} \lesssim 15$ and $N \sim 3$ across all configurations. This behavior is primarily due to the prior on N, which is asymmetric around its peak at $N = 3$ [2], as described by Equation 2.11. In summary, N is influenced by $\mathrm{SNR}_{\mathrm{net}}$ but not directly by $\mathcal{I}$.

The bottom panel of Figure 2.3 illustrates all components of Equation 2.38: the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ is plotted as a fuction of N for different $\mathcal{I}$, and the color bars indicate the injected $\mathrm{SNR}_{\mathrm{net}}$ of each injection in the corresponding detector configuration. Injections with $\mathrm{SNR}_{\mathrm{net}} \lesssim 15$ are typically recovered by *BayesWave* with $N \leq 3$ and $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \leq 50$, as low SNR injections have low-amplitude features that cannot be confidently reconstructed, so no obvious trends are observed in this regime. In contrast, injections with $\mathrm{SNR}_{\mathrm{net}} \gtrsim 15$ are recovered with $N > 3$, and as predicted by Equation 2.38, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ generally increases with N, for all three detector configurations. Moreover, for comparable values of N and SNRnet, the HLKV network typically yields higher $\ln \mathcal{BS},\mathcal{G}$ than HL or HLV, further emphasizing that $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ scales with the number of detectors $\mathcal{I}$.

[8] This section is adapted with substantial rewording from Section VI.A of Ref. [3], which originally appeared in the thesis submitted for the degree of Master of Science (Physics) at The University of Melbourne, awarded to the author in 2020.

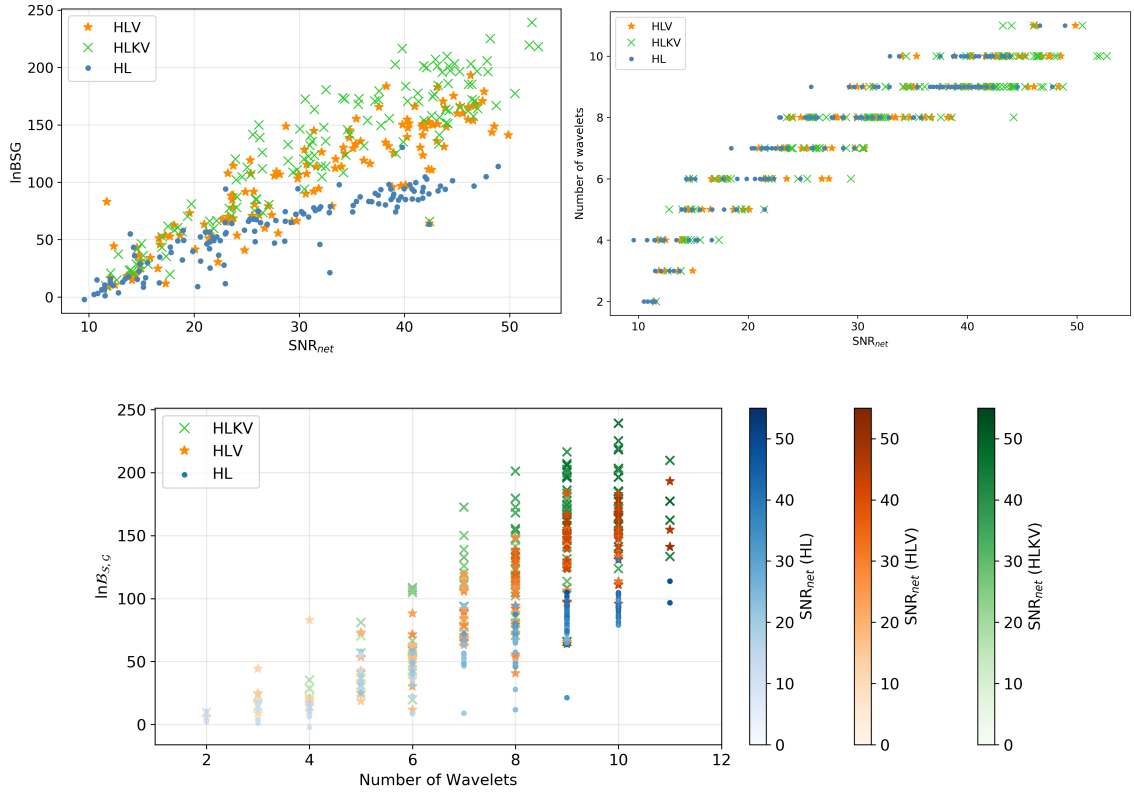

Figure 2.3: Empirical scaling of the log signal-to-glitch Bayes Factor $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$, based on BBH injections. The top left panel shows $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ versus the injected network signal-to-noise ratio, $\mathrm{SNR_{net}}$. Each data point represents a single BBH injection. The top right panel shows the median number of wavelets used in signal model reconstruction, N versus $\mathrm{SNR_{net}}$. The bottom panel shows $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ versus N, with the three color bars indicating the $\mathrm{SNR_{net}}$ of BBH injections in the corresponding detector networks. In the top panels, the horizontal axis corresponds three different network SNRs: (i) for the blue data points it corresponds to $\mathrm{SNR_{net}}$ of the HL network, (ii) for the orange data points it corresponds to $\mathrm{SNR_{net}}$ of the HLV network, (iii) for the green data points it corresponds to $\mathrm{SNR_{net}}$ of the HLKV network. Figures and caption reproduced from Ref. [3].

Sine-Gaussian wavelet injections[9]

Ref. [3] also uses sine-Gaussian wavelet signals to validate the analytic scaling of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. Since sine-Gaussian wavelets form the frame of *BayesWave* reconstructions, all sine-Gaussian signal injections are expected to require exactly one wavelet ($N = 1$) for reconstruction, regardless of SNR. Thus, this test focuses specifically on the scaling behavior of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ with respect to $\mathcal{I}$, without the influence of N.

The analysis dataset consists of 150 sine-Gaussian wavelets, injected as coherent signals into Gaussian detector noise of the HL, HLV and HLKV networks. All intrinsic wavelet parameters, except for A, are uniformly sampled from the following ranges: $t_0 \in [1.5, 2.5]$ s (with $t = 2$ s as the center of the *BayesWave* analysis window), $f_0 \in$

[9]This section is adapted with substantial rewording from Appendix C of Ref. [3], which originally appeared in the thesis submitted for the degree of Master of Science (Physics) at The University of Melbourne, awarded to the author in 2020.

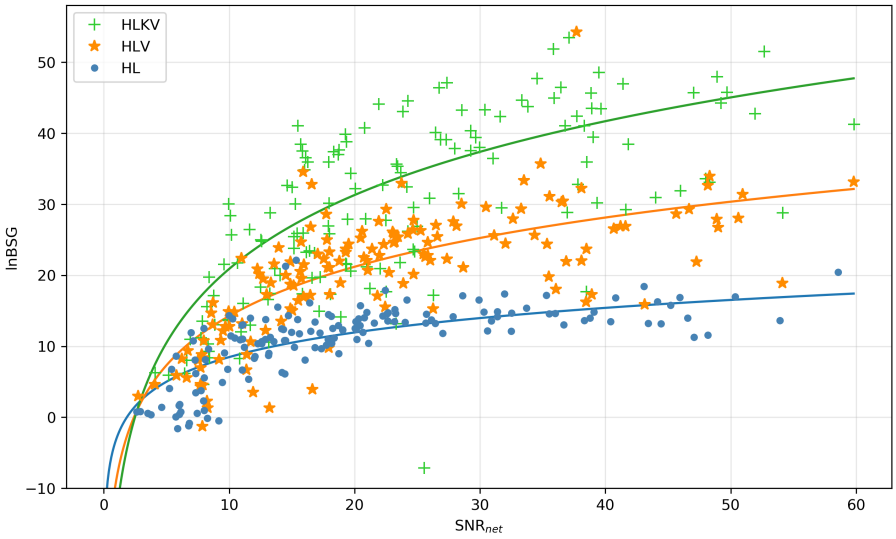

Figure 2.4: Sine-Gaussian wavelet injections, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ versus $\mathrm{SNR}_{\mathrm{net}}$. The solid lines with colors corresponding to the data symbols are analytic predictions of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ given by Equation 2.39. Figure and caption reproduced from Ref. [3].

$[32, 1000]$ Hz, $Q \in [0.1, 40]$ and $\phi_0 \in [0, 2\pi]$. The extrinsic signal parameters are also uniformly sampled from the ranges: $\alpha \in [0, 2\pi]$, $\sin\delta \in [-\pi/2, \pi/2]$, $\psi \in [0, 2\pi]$, and $\epsilon \in [-0.99, 0.99]$. The wavelet SNRs are uniformly sampled from the range SNR $\in [10, 50]$, and the amplitude A is calculated as per Equation 2.12.

Figure 2.4 shows $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ as a function $\mathrm{SNR}_{\mathrm{net}}$ for the sine-Gaussian wavelet injections across the HL, HLV, and HLKV networks. At fixed $\mathrm{SNR}_{\mathrm{net}}$, it is evident that $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ increases with $\mathcal{I}$, as predicted by Equation 2.38. Unlike BBH injections, where N depends on $\mathrm{SNR}_{\mathrm{net}}$, sine-Gaussian injections yield $N = 1$ for all $\mathrm{SNR}_{\mathrm{net}}$. Therefore, the differences in $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ between detector configurations at a fixed $\mathrm{SNR}_{\mathrm{net}}$ is entirely attributed to $\mathcal{I}$.

Ref. [3] also compares the analytic and empirical scaling of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ with $\mathcal{I}$ by fitting the generalized analytic expression:

$$\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \approx (\mathcal{I} - 1)[5 \ln \mathrm{SNR}_{\mathrm{net}} + a] + \frac{5}{2}\mathcal{I} \ln \mathcal{I} + b, \tag{2.39}$$

to the empirical data shown in Figure 2.4. Equation 2.39 is derived from Equation 2.37 by setting $N = 1$ and combining the constant terms into $a = \frac{5}{2} - \ln(V_\lambda) + \ln(Q)$ and $b = 2 + \ln \frac{\sqrt{\det C_\Phi}}{V_\Phi}$. The analytic fits as shown as solid curves in Figure 2.4, with colors corresponding to the respective detector network. The constants, $a = -10$ and $b = 4$, are determined by eye, and the same values are assumed for all fits. The solid curves are broadly consistent with the data, indicating general agreement between the analytic and empirical scalings of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ with $\mathcal{I}$. However, one should note that the solid curves represent networks with detectors of equal sensitivity, whereas the data is obtained from LVK networks comprising detectors with varying sensitivities. Therefore, the analytic fits should be regarded as rough approximations.

2.6. Applications of *BayesWave*

BayesWave is a versatile and well-established tool within the GW burst community, known for its model flexibility and its ability to distinguish signals from glitches by incorporating waveform complexity through $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. Below, we summarize the various applications of *BayesWave* in detecting and characterizing GW bursts and glitches, as reported in the existing literature:

(1) **All-sky search for short-duration GW bursts**. As discussed in Section 1.5.2, *BayesWave* is one of the three analysis pipelines used in the LVK all-sky short-duration GW burst searches [44–46]. However, *BayesWave* is computationally intensive due to the implementation of the RJMCMC sampling, so it is used as a follow-up to cWB. The hierarchical implementation of cWB and *BayesWave* has been shown to increase detection confidence of burst signals with non-trivial (i.e. non sine-Gaussian) time-frequency structure, with network SNR $\gtrsim 10$ [139].

(2) **Waveform consistency tests for CBC detections**. While matched-filtering is a robust method for detecting well-modeled CBCs, its effectiveness depends on the completeness of parameter space coverage by template banks. In other words, matched filtering is inherently limited by the availability of models, which reflect our current understanding of the parameter spaces. Minimally-modeled analyses like *BayesWave*, on the other hand, use flexible and generic frame functions that can capture signal properties that extend beyond the scope of the template bank, or deviate from the predictions of general relativity. Therefore, it is crucial to verify the consistency between the waveform reconstructions produced by the two methods. *BayesWave* is one of the algorithms used to perform waveform consistency tests for the detections reported in the GWTCs [35, 36, 38]. The consistency between *BayesWave* and template-based reconstructions is measured using the overlap, similar to that defined in Equation 1.43 [140].

(3) **Glitch subtraction**. *BayesWave* can simultaneously model a signal and glitch, making it an effective glitch subtraction tool. *BayesWave* subtracted the glitch overlapping with the BNS detection GW170817, while preserving the integrity of the underlying signal [97]. Since then, *BayesWave* has become one of the leading tools for glitch removal in the data-processing step before signal parameter estimation, in GWTC-2 and -3 [36, 38]. The performance and impact of *BayesWave* in glitch subtraction is also studied extensively in Refs. [98, 101, 102].

(4) **Spectral noise estimation**. As noted in Section 2.2.4, *BayesWave* can be used to model the noise PSD. In the GWTCs, *BayesWave* is used to model the PSDs for template-based parameter estimation of CBC detections [141].

(5) **Analyzing under-constrained burst sources**. In addition to all-sky searches for generic bursts, *BayesWave* has also been used to characterize specific types of unmodeled or under-constrained sources, and set upper limits on their detectability. Examples of these sources include BNS post-mergers [142–144], eccentric BBHs [145, 146], BBH post-merger echoes [147, 148], and CCSNe [46, 149, 150].

(6) **Model-independent parameter estimation**. *BayesWave* signal reconstruction using sine-Gaussian wavelets cannot be used to infer source-specific physical properties, e.g. masses and spins of CBCs. However, it can be used to estimate the sky location and the model-independent central moments of the signal, which includes the central time, duration, central frequency and bandwidth. The parameter estimation and sky localization performance of *BayesWave* for BBHs and generic burst signals, such as sine-Gaussians, Gaussians, and white noise bursts, are studied in Ref. [151]. The overall performance of *BayesWave* is shown to improve for signals that are more localized in the time-frequency plane [152]. Additionally, the impact of expanding detector networks on *BayesWave*'s sky localization performance is studied in Ref. [3].

(7) **Model-independent tests of general relativity**. *BayesWave* is used for the model-independent residual test of general relativity for GW150914 [56] defined below, as well as for other CBC candidates reported in the GWTCs [153–155]. This approach employs *BayesWave* to search for excess power in residual data after subtracting the best-fitting signal template predicted by general relativity. The presence of excess power, or lack thereof, quantifies the degree of agreement between the waveforms predicted by general relativity and the observed data.

(8) **Study of GW polarizations**. The signal model described in Section 2.2.2 comprises two tensor polarizations, consistent with the predictions general relativity. Nevertheless, *BayesWave*'s signal model can be generalized to include up to six polarizations, accounting for the scalar and tensor polarizations predicted by alternative theories of gravity. The generalized model is used in Ref. [156] to (i) study the ability to constrain the presence of scalar/vector polarization in addition to the tensor polarizations in GW detection candidates, and (ii) place upper limits on the detectability of scalar/vector polarizations.

In the remainder of this thesis, we study three different aspects of the performance and applications of *BayesWave*. In Chapter 3, we assess the multi-detector burst detection performance of *BayesWave* in the presence of glitches. In Chapter 4, we explore the possibility of using *BayesWave* to detect frequency-specific GW signatures and consequently constrain explosion models of CCSNe. In Chapter 5, we present *BayesWave* as a tool for characterizing GW burst polarizations.

Chapter 3

Impact of noise transients on gravitational-wave burst detection efficiency of the *BayesWave* pipeline with multi-detector networks

This chapter is a reproduction of [4]

> *Y. S. C. Lee, M. Millhouse, and A. Melatos*,
> Physical Review D 109, 082002 (2024).

reformatted with the following changes only:

- **Section II of Ref. [4] is excluded from this chapter as it has been adapted and incorporated into Chapter 2 of this thesis.**
- **The first two paragraphs of Section 3.7.2 are revised and expanded in response to suggestions from the thesis examiners.**
- **The text is styled to match the rest of this thesis.**
- **Notations are updated to match the rest of this thesis.**
- **Where necessary, bibliographic records are updated.**

3.1. Abstract

Detection confidence of the source-agnostic gravitational-wave burst search pipeline *BayesWave* is quantified by the log signal-versus-glitch Bayes factor, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. A recent study shows that $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ increases with the number of detectors. However, the increasing frequency of non-Gaussian noise transients (glitches) in expanded detector networks is not accounted for in the study. Glitches can mimic or mask burst signals resulting in false alarm detections, consequently reducing detection confidence. This paper presents an empirical study on the impact of false alarms on the overall performance of *BayesWave*, with expanded detector networks. The noise background of *BayesWave* for the Hanford-Livingston (HL, two-detector) and Hanford-Livingston-Virgo (HLV, three-detector) networks are measured using a set of non-astrophysical background triggers from the first half of Advanced LIGO and Advanced Virgo's Third Observing Run (O3a). Efficiency curves are constructed by combining $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ of simulated binary black hole

signals with the background measurements, to characterize *BayesWaves*'s detection efficiency as a function of the per-trigger false alarm probability. The HL and HLV network efficiency curves are shown to be similar. A separate analysis finds that detection significance of O3 gravitational-wave candidates as measured by *BayesWave* are also comparable for the HL and HLV networks. Consistent results from the two independent analyses suggests that the overall burst detection performance of *BayesWave* does not improve with the addition of Virgo at O3a sensitivity, because the increased false alarm probability offsets the advantage of higher $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$.

3.2. Introduction

The Advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) [76] detectors in Hanford, Washington and Livingston, Louisiana, USA have completed three observing runs O1, O2 and O3 between 2015 and 2020, two of which were joint observations with the Advanced Virgo detector in Cascina, Italy [77]. The Kamioka Gravitational Wave Detector (KAGRA) [78] located in Hida, Japan also came online towards the end of O3, conducting a joint observation (O3GK) [157] with the GEO600 [79] detector in Hannover, Germany. As of the three observing runs, around 90 candidate gravitational wave (GW) events were collectively observed and reported in the Gravitational-wave Transient Catalogs (GWTCs) [35, 36, 38]. In May 2023, the LIGO-Virgo-KAGRA (LVK) collaboration began the fourth observing run O4 with the two LIGO detectors. The Virgo and KAGRA detectors are also expected to join O4 at a later date.

GW events observed so far by the LVK detectors are compact binary coalescences (CBCs), namely the mergers of binary black holes (BBH), binary neutron stars and neutron star-black hole binaries. CBCs are transient GW events, otherwise known as GW burst sources. Aside from CBCs, we expect to observe GW bursts from other astrophysical sources including but not limited to core-collapse supernovae [158–160], pulsar glitches [161], magnetar bursts[1] [162, 163], nonlinear gravitational memory due to low-mass BBH mergers [164] and cosmic string cusps or kinks [165–167]. In addition, the possibility exists of GW bursts from astrophysical objects or processes that have not yet been discovered through electromagnetic observations. By their nature, GW waveforms of such novel signals are unclassified at present.

Traditionally, GW transient search pipelines use a matched filter [168–171] to compare the data to a bank of waveform templates obtained through various waveform modeling techniques [172–175]. Unlike CBCs, the waveforms of most prospective GW burst sources vary unpredictably from one event to the next and involve complicated physics beyond general relativity (e.g. hydrodynamics and neutrino transport). It is challenging to construct robust models with a few well-defined parameters which predict the waveforms, so template-based matched-filter searches for unmodeled GW bursts are impractical.

Several developed and emerging pipelines exist to perform source-agnostic GW burst searches [117, 121, 123], including but not limited to coherent WaveBurst (cWB) and *BayesWave*. The cWB [114–116, 176] burst search pipeline is used for offline analysis

[1] Magnetar bursts are short bursts ($\sim 0.1s$) of soft gamma-rays emitted by highly magnetized, isolated neutron stars. Their physical mechanism is unknown.

and online, low-latency generation of triggers for electromagnetic followups. Detection statistics of the cWB algorithm scales with the excess power in the time-frequency domain. *BayesWave* uses the transdimensional Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm which adjusts the model dimension in response to the data. For that reason, *BayesWave* is computationally intensive and is only used to follow-up potential GW candidates identified by other search pipelines. In the all-sky GW burst searches of the three Advanced LIGO and Advanced Virgo observing runs [44–46], cWB is used to analyze the full dataset and *BayesWave* is used to follow-up cWB triggers [1, 2, 136]. Previous studies have shown that hierarchical implementation of cWB and *BayesWave* enhances detection confidence [139].

As of O4, the LVK global network comprises of four large-scale detectors. With the commissioning of LIGO-India well under way [82], the network of GW detectors is expected to expand in the coming years. The expanding network of detectors with improved sensitivities increases the duty cycle, sky coverage and the accuracy of sky localization [90, 177]. However, having more detectors also increases the susceptibility of the network to transient non-astrophysical disturbances, as noted in O3 [36, 38, 46, 101]. These non-Gaussian instrumental noise transients, otherwise known as "glitches", appear as excess power in detector data and can mimic or mask unmodeled GW bursts. To enable high confidence detections with high astrophysical significance, glitches have to be identified and mitigated appropriately. Several efforts have been made to identify and characterize glitches by their origin and/or morphology [96, 99, 178–180]. Three common glitches in the LIGO-Virgo detectors are termed blip [181], whistle [182] and scattered light [183]. The whistle and scattered light glitches are of relatively longer duration ($\sim$ 0.7-2.0 s) and their origins are well-understood. Blip glitches, on the other hand, are transient power spikes which lasts for $\sim$ 0.1 s and spans a wide frequency band ($\sim 10^2$ Hz), typically of unknown origin. In cases where the glitch origin is unknown, further investigations are necessary before flagging a glitch and regressing it from the data to avoid overlooking astrophysical signals [101, 184–190].

The *BayesWave* algorithm enables the joint detection and characterization of GW burst and instrumental glitches, with no *a priori* assumptions of the source or morphology. Studies have been conducted to evaluate various aspects of *BayesWave*'s performance with multi-detector networks, including detection confidence, parameter estimation and waveform reconstruction [3, 85, 139, 140, 151]. In Ref. [3], the detection confidence of *BayesWave* with multi-detector networks is quantified using the algorithm's detection statistic: the log signal-to-glitch Bayes factor, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. The study showed analytically that increasing the number of detectors in a network has a positive impact on $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$, following derivations in Ref. [85]. The results are verified empirically with simulated BBH signals. While the outcome is promising, the study does not consider the increase in glitch rate in an expanded detector network, i.e. it only focuses on the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ of astrophysical events injected into simulated data in the absence of glitches. This paper generalizes Ref. [3], presenting a fuller analysis of *BayesWave*'s burst detection performance with expanded detector networks by accounting for the detector noise background using real detector data. For noise background measurements, we combine data from the first half of O3 (O3a) for the LIGO Hanford (H), LIGO Livingston (L) and Virgo (V) detectors, in particular the HL (two-detector) and HLV (three-detector) networks. We compare the overall performance of *BayesWave* between the HL and HLV

networks in O3a, noting that Virgo is less sensitive than HL; in contrast, the sensitivities of all three detectors may be comparable in future observing runs. The performance of *BayesWave* is evaluated by comparing the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ produced by astrophysical signals against the respective detector network backgrounds, using two independent injection sets. A set of simulated BBH signals is used to construct efficiency curves for characterizing *BayesWave*'s detection efficiency as a function of detection significance. To check for consistency, we analyze O3-like CBC signals to measure *BayesWave*'s detection significance of O3 GW candidates from GWTC-2 [36] and GWTC-3 [38].

The rest of this paper is organised as follows. In Section 3.3 we discuss the datasets used to study *BayesWave*'s performance: (i) HL and HLV background triggers for background measurements, (ii) simulated BBH injections and (iii) O3-like CBC waveform injections. In Section 3.4 we present *BayesWave*'s background measurements. In Section 3.5 we present the results for *BayesWave*'s efficiency analysis with the simulated BBH injections, and in Section 3.6 the significance measurements for the O3 GW candidates. We summarize our findings and discuss avenues for future work in Section 3.7.

3.3. *BayesWave* efficiency analysis

In standard GW searches, the astrophysical significance of a detection candidate is determined by the frequency of false alarms. False alarms are non-astrophysical events with detection statistics corresponding to that of GW candidates. To estimate the prevalence of false alarms, one can count the number of triggers produced by the detector background which does not contain astrophysical signals.

Ref. [3] assesses *BayesWave*'s detection confidence for expanded detector networks using only the detection statistic $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ produced by astrophysical events. However as the global detector network expands, the likelihood of instrumental glitches increases. The associated increase in false alarm detections reduces astrophysical significance of detections, thereby reducing detection confidence. Unmodeled burst searches (e.g. with *BayesWave*) place fewer constraints on the waveform morphology, and are therefore confounded more readily by glitches compared to modeled searches (e.g. with a matched filter) [191–195]. Since the significance of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ is influenced by false alarms, we present a more complete analysis of *BayesWave*'s performance with expanded detector networks by considering the impact of detector noise backgrounds on detection confidence.

We use detection efficiency $P_{\rm det}$ as a figure of merit to compare the overall performance of *BayesWave* between the HL (two-detector) and HLV (three-detector) networks. $P_{\rm det}$ is typically characterized as a function of detection significance by means of a receiver-operating-characteristic (ROC) curve[2], also known as an efficiency curve. In this study, we use the per-trigger false alarm probability $P_{\rm FA}$ as a measure of significance. We define $P_{\rm FA}$ as the probability that a trigger measured with a given detection statistic is a false alarm, and $P_{\rm det}$ as the probability of detecting an astrophysical event with a given significance. Higher $P_{\rm FA}$ indicates low astrophysical significance. Therefore higher $P_{\rm det}$ is achieved, if higher $P_{\rm FA}$ is tolerated. $P_{\rm FA}$ should not be confused with

[2]Typical ROC curves plot probability of detection (true positives) on the vertical axis and probability of false alarm (false positives) on the horizontal axis

the false alarm rate (FAR), which measures the number of false alarms per unit time. We further discuss the difference in Section 3.4, and explain why we use $P_{\rm FA}$ instead of FAR in our analysis.

In order to measure $P_{\rm FA}$, we need to understand the distribution of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ produced by the detector noise background. It is challenging to construct models that can accurately predict the noise background, so we empirically obtain the background distribution by applying *BayesWave* to triggers identified by cWB from time-shifted background data of the HL and HLV networks. The distribution of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ produced by the background triggers is then used to compute $P_{\rm FA}$. Using *BayesWave* to recover a population of injected signals, we obtain a distribution of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ for astrophysical events. Combining $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ of background triggers and astrophysical injections, we compute $P_{\rm det}$ as a function of $P_{\rm FA}$ to construct efficiency curves. We discuss the methods of constructing efficiency curves in greater detail in Section 3.5.1.

To study the impact of the noise background on *BayesWave*'s overall performance with expanded detector networks, we compare the efficiency curves between the HL and HLV networks for a synthesized population of BBHs. As a consistency check, we also analyze a set of O3-like CBC waveforms to measure and compare *BayesWave*'s detection significance of O3 GW detection events for HL and HLV. In the following sections, we detail the background and injection datasets for the analyses.

3.3.1. Background data

In GW data analysis, it is standard practice to use the time-shifting method to create pseudo-real detector datasets for noise background estimations [44–46, 196–198]. The time-shifting method introduces artificial time off-sets between the outputs of GW detectors operating in concert. The offsets are much larger than the coherence time (∼10 ms) of any real GW signals between the detectors, determined by the distance between the detectors and the GW propagation speed. As a result, coincident triggers in the time-shifted data cannot be astrophysical. By performing time-shifts repetitively on months worth of detector data, we obtain an artificially extended set of background data with effective livetimes[3] spanning thousands of years. We can use this to estimate $P_{\rm FA}$ by empirically measuring the fraction of background (i.e. noise-induced) triggers above a selected detection threshold in the time-shifted background [196].

In the Advanced LIGO and Advanced Virgo all-sky searches for short GW bursts [44–46], the cWB algorithm is used to analyze the full observational data. Due to the implementation of RJMCMC, *BayesWave* is computationally intensive. Thus, *BayesWave* is only used to follow-up subsets of cWB triggers. Although Ref. [139] has shown that the hierarchical implementation of cWB and *BayesWave* enhances detection confidence in all-sky burst searches, the aim of our study is to assess the independent burst detection performance of *BayesWave*. By convention, we use pre-existing trigger lists generated by the cWB pipeline [46, 176] to downselect triggers for *BayesWave* background measurements, but we do not make any claims on cWB's background and detection efficiency. We choose to use the trigger list for the first half of O3 (O3a), acquired from the cWB low-frequency (16–1024 Hz) all-sky analysis of the full time-shifted O3a background data. The analysis is conducted separately for the HL and HLV networks, on background

[3] The extended time interval obtained as a result of time-shifting is known as the effective livetime.

data obtained by applying time-shifts on 104.94-day (HL) and 75.19-day (HLV) segments of the real-time O3a detector data. The time-shifted background data accumulates 981 years and 573 years of effective livetimes for the HL and HLV networks respectively.

We select triggers by thresholding their cWB detection statistic, η_c, which scales with the network SNR of the signal present in the data [114, 115]. We arbitrarily nominate $\eta_{c,threshold} = 7$ as the significance threshold, in line with previous work [46, 139]. Triggers below the threshold are presumed to have insignificant impacts on detection efficiency and therefore excluded from the noise background measurement. cWB identifies 2×10^3 and 7×10^3 triggers[4] with $\eta_c > \eta_{c,threshold}$ in the HL and HLV background datasets respectively, but thousands of triggers are still too expensive to handle computationally in this paper. A straightforward approach is to increase $\eta_{c,threshold}$, but that would deliberately exclude low-SNR triggers from the background measurement. To avoid implementing a stricter $\eta_{c,threshold}$, we run *BayesWave* on a fraction (denoted by X) of randomly selected triggers from the full trigger list, all of which satisfy $\eta_c > \eta_{c,threshold} = 7$. We set $X = 0.45$ and $X = 0.15$ for the HL and HLV datasets respectively to deliver roughly equal numbers of triggers from the two networks. The reduced HL and HLV background datasets consist of 1008 and 1134 triggers respectively. We employ *BayesWave* to analyse the datasets to obtain $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ for each background trigger. The *BayesWave* analysis uses the same settings as the Advanced LIGO and Advanced Virgo O3 all-sky search for short GW bursts [46] (see Appendix 3.A1).

From the *BayesWave* analysis, we flag background triggers that are more consistent with the pure Gaussian-noise model, $\mathcal{N}$ than the composite signal plus Gaussian-noise model $\mathcal{S}$. By definition, the $\ln \mathcal{B}_{\mathcal{S},\mathcal{N}}$ error bars of these triggers encompass values less than or equal to zero, i.e.

$$\ln \mathcal{B}_{\mathcal{S},\mathcal{N}} - \Delta \ln \mathcal{B}_{\mathcal{S},\mathcal{N}} \leq 0 \tag{3.1}$$

where $\Delta \ln \mathcal{B}_{\mathcal{S},\mathcal{N}} = \sqrt{[\Delta \ln p(\boldsymbol{d}|\mathcal{S})]^2 + [\Delta \ln p(\boldsymbol{d}|\mathcal{N})]^2}$ is the width of the error bars and $\Delta \ln p(\boldsymbol{d}|\mathcal{M})$ denotes the uncertainty of the log evidence of model $\mathcal{M}$ [1]. For these Gaussian-noise-like triggers, the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ is meaningless as it serves to compare the evidences of models that characterize non-Gaussianity. Nevertheless, these triggers cannot be discarded from the background measurement as they satisfy $\eta_c > \eta_{c,threshold}$. Therefore we assign them with an arbitrarily low detection statistic, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = -500$, to indicate minimal astrophysical significance. A total of 268 (218) out of 1008 (1134) HL (HLV) triggers are assigned $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = -500$.

We present and discuss the background measurements in Section 3.4.

3.3.2. *Injections*

In addition to the background measurement, the detection statistic distribution for astrophysical signals is required to evaluate *BayesWave*'s burst detection performance. We inject waveforms of known morphology and recover them using *BayesWave* to empirically measure the distribution of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. Since CBCs are well-understood, we use them in this study to assess *BayesWave*'s independent performance with HL and HLV. We

[4] These counts include triggers from all three search bins used in the cWB O3a low-frequency burst analysis: LF1, LF2 and LF3. The bins are classified based on trigger morphologies. Classification details can be found in [46] and [199].

analyze two different source populations: Injection Set 1 (IS1), comprising phenomenological BBH waveforms with fixed component masses but uniformly distributed SNR and extrinsic parameters, and Injection Set 2 (IS2), comprising CBC waveforms with parameters that resemble real GW events from O3. The following two subsections describe the objective and properties of each injection dataset in order.

Phenomenological BBH waveforms (IS1)

IS1 consists of simulated BBH waveforms with a choice of parameter space encompassing the range detectable by the Advanced LIGO and Advanced Virgo detectors. The waveforms are added to temporally spread out segments of HL and HLV data across all of O3a to reflect practical observation intervals. We use IS1 to characterize *BayesWave*'s detection efficiency ($P_{\rm det}$) as a function of detection significance ($P_{\rm FA}$) via efficiency curves and compare the performance of *BayesWave* with the HL and HLV networks.

IS1 copies the injection set described in Section V of Ref. [3]. It consists of 1200 simulated BBH waveforms phenomenologically modeled using the `IMRPhenomD` [137, 138] approximant. The BBH sources are non-spinning, non-precessing and have equal component masses of $30M_{\odot}$. They also have uniformly distributed sky locations, inclinations and polarization angles. The distances are randomly sampled such that the signal amplitude is detectable in simulated HLV data with network signal-to-noise ratio within range $10 \leq \mathrm{SNR}_{\rm net} \leq 50$. We use the same injection dataset for both the HL and HLV networks; we simply exclude Virgo data in the HL analysis. By Equation 1.28, we expect $\mathrm{SNR}_{\rm net}$ of any given event to be lower in the HL network compared to HLV.

The analysis in Ref. [3] injects and recovers waveforms using projected (simulated) O4 detector data. However, *BayesWave*'s background measurements for HL and HLV in this study are carried out using O3a background triggers as discussed in Section 3.3.1. In order to measure the $P_{\rm det}$ as a function of $P_{\rm FA}$, detection statistics ($\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$) of the astrophysical signals must be compared with the background triggers of the same detector data. Thus we inject IS1 into arbitrarily selected segments of HL and HLV data throughout O3a. The O3a strain data is publicly available at the Gravitational Wave Open Science Center (GWOSC) [200, 201] and Figure 2 of Ref. [36] shows representative amplitude spectral densities of the detectors. As with the background, IS1 is analyzed using the same *BayesWave* settings as Ref. [46].

Events of IS1 are injected into O3a data with the same distances sampled from the simulated data. Since O3a data is noisier and has a different characteristic PSD compared to the simulated HLV data, the $\mathrm{SNR}_{\rm net}$ of IS1 events when injected into O3a data is lower than the referenced range $10 \leq \mathrm{SNR}_{\rm net} \leq 50$. In order to assess *BayesWave*'s performance under conditions relevant to practical searches, events below a designated detection threshold must be eliminated from the injection data. This is because they cannot serve as triggers by definition, in the same way that background triggers with $\eta_{\rm c} < \eta_{\rm c,threshold} = 7$ do not count as false alarms. Since this a designated search to assess the stand-alone efficiency of the *BayesWave* algorithm, independent of cWB, we set a nominal significance threshold of $\mathrm{SNR}_{\text{cut-off}} = 10$ for *BayesWave* viz. only injection events with $\mathrm{SNR}_{\rm net} \geq \mathrm{SNR}_{\text{cut-off}}$ in *both* the HL and HLV networks are adequately significant to be followed-up by *BayesWave* and included in the efficiency curve analysis. Out of 1200 injections, 412 non-detection events are filtered out from IS1, leaving 788 events going forward.

From the remaining 788 events, *BayesWave* identifies 157 (89) events consistent with Gaussian noise in the HL (HLV) network according to the $\ln \mathcal{B}_{\mathcal{S},\mathcal{N}}$ constraint defined in Section 3.3.1. These events are retained in the analysis dataset since they satisfy $\mathrm{SNR}_{\mathrm{net}} \geq 10$ but as with the background triggers, they are assigned $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = -500$ to indicate low detection significance.

To show the overall distribution of IS1 events, we plot $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ versus $\mathrm{SNR}_{\mathrm{net}}$ for the HL (blue circles) and HLV (orange stars) injections in Figure 3.1. The plot shows all but the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = -500$ events to focus on events with astrophysically relevant $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. Injections with comparable $\mathrm{SNR}_{\mathrm{net}}$ are evidently recovered with higher $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ in HLV compared to HL. This observation is consistent with Ref. [3] where $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ is analytically and empirically shown to increase primarily with $\mathcal{I}$.

Despite the astrophysical origin of IS1, *BayesWave* recovers two of the HL events with $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} < 0 \neq -500$ in Figure 3.1, suggesting that the evidence for the 'incoherent' glitch model ($\mathcal{G}$) is higher than for the 'coherent' signal model ($\mathcal{S}$). These events are also not consistent with Gaussian-noise i.e. they have $\ln \mathcal{B}_{\mathcal{S},\mathcal{N}} > 0$. This is because the injected signal power in the frequency domain is only marginally above the sensitivity threshold in one detector, and is approximately one order of magnitude lower in the other. Equation 2 shows that the sensitivity of each detector to different sky locations, at a given time, depends on the antenna pattern functions. Therefore the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} < 0$ recovery of the two HL injections, caused by the signal power imbalance across the detectors, is an inadvertent result of mismatched detector sensitivities to the randomly sampled sky locations at the time of injection. With additional coherent signal power from Virgo, the HLV-equivalents of these two events are recovered with $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \sim 10^1$. This argument also applies to IS2 injections, discussed in Section 3.3.2. We present the results of *BayesWave* detection efficiency analysis with IS1 in Section 3.5.

O3-like CBC waveforms (IS2)

To check for consistency with IS1, we measure *BayesWave*'s detection significance for real GW detection events in terms of P_{FA}, and compare the measurements between the HL and HLV networks. For this purpose, we implement *BayesWave* on IS2 consisting CBC waveforms resembling O3a and O3b GW events from GWTC-2 [36] and GWTC-3 [38] respectively, otherwise known as off-source injected waveforms.

In IS1 the BBH waveforms are sampled from a fixed parameter space and added to detector data spread out across all of O3a; in IS2 the off-source injections are sampled from the matched-filter source parameter posteriors for GW detection events and added into the background data around the event epoch. Off-source injections are used in the GWTCs to test the consistency[5] between matched-filter (template-based) CBC waveforms and minimally-modeled waveform reconstructions (e.g. cWB and *BayesWave*) [35, 36, 38].

[5]Consistency test are performed by comparing the on-source and off-source match. On-source waveforms are reconstructed directly from the event data. The match, defined by $\mathcal{O} = \langle h_1 \mid h_2 \rangle / \sqrt{\langle h_1 \mid h_1 \rangle \langle h_2 \mid h_2 \rangle}$, measures the overlap between two waveforms h_1 and h_2. $\langle \cdot \mid \cdot \rangle$ is the noise-weighted inner product [202]. On-source match compares the maximum likelihood waveform from template-based parameter estimation of the actual event with the point estimate from minimally-modeled reconstructions; off-source match compares the off-source injections with their respective reconstructions.

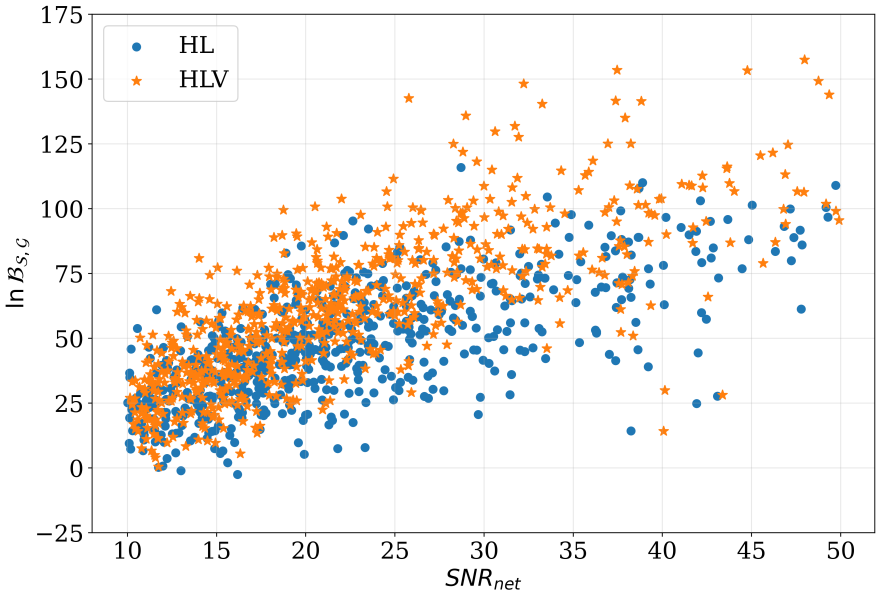

Figure 3.1: Log signal-to-glitch Bayes factor $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ versus network signal-to-noise ratio $\mathrm{SNR}_{\mathrm{net}}$ for IS1. The blue circles (orange stars) correspond to HL (HLV) network injections; each data point corresponds to a single injection. Gaussian-noise-like events with $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = -500$ are not shown.

IS2 comprises off-source injections of 22 independent GW events detected by the HLV network in O3. We summarize the relevant event properties in Table 3.1. All events listed in Table 3.1, except for GW200202_154313, are *BayesWave* waveform consistency test candidates [36, 38]. GW200202_154313 is excluded from the GWTC-3 consistency test due to low on-source match, but since the off-source injections for this event are available there is no reason to exclude it from IS2 for the assessment of *BayesWave*'s detection significance. A set of 200 off-source injections is available for each of the 22 GW events [36, 38]. We arbitrarily select 50 out of the 200 off-source injections for each GW event, totaling $22 \times 50 = 1100$ injections in IS2. Even though a fraction of the GW events are O3b detections, we inject all off-source events into segments of O3a HLV data to ensure comparability with the O3a noise background described in Section 3.3.1. The HL data are equivalent to the HLV data with Virgo removed.

As with IS1, only injections above the *BayesWave* significance threshold are retained in IS2. The last column in Table 3.1 shows the number of off-source injections that exceeds the significance threshold i.e. $\mathrm{SNR}_{\mathrm{net}} \geq \mathrm{SNR}_{\text{cut-off}} = 10$ for each GW event. There are four GW events with less than 25 off-source injections (i.e. $< 50\%$) satisfying the significance threshold, namely: GW190517_055101, GW190720_000836, GW190828_065509 and GW200219_094415. Since the $\mathrm{SNR}_{\mathrm{net}}$ of off-source injections for these four events are sampled from match-filter network SNR posteriors with medians $\lesssim 11$ (see Table 3.1), they are less likely to satisfy $\mathrm{SNR}_{\mathrm{net}} \geq 10$. Assessments of astrophysical significance for GW events with ≤ 25 off-source injections are unreliable due to insufficient P_{FA} measurements. Therefore, the four events listed above are excluded from the IS2 analysis. For the remaining 18 GW events, the numbers of injections shown in the last column of Table 3.1 include events that are more consistent with Gaussian noise than a GW signal according to *BayesWave*. As discussed in Section 3.3.1, these events are

LVK run	Event name	m_1 ($M_\odot$)	m_2 ($M_\odot$)	Network SNR†	# off-source injections in IS2
O3a	GW190408_181802	$24.6^{+5.1}_{-3.4}$	$18.4^{+3.3}_{-3.6}$	$15.3^{+0.2}_{-0.3}$	50
O3a	GW190412	$30.1^{+4.7}_{-5.1}$	$8.3^{+1.6}_{-0.9}$	$18.9^{+0.2}_{-0.3}$	48
O3a	GW190503_185404	$43.3^{+9.2}_{-8.1}$	$28.4^{+7.7}_{-8.0}$	$12.4^{+0.2}_{-0.3}$	47
O3a	GW190512_180714	$23.3^{+5.3}_{-5.8}$	$12.6^{+3.6}_{-2.5}$	$12.2^{+0.2}_{-0.4}$	40
O3a	GW190513_205428	$35.7^{+9.5}_{-9.2}$	$18.0^{+7.7}_{-4.1}$	$12.9^{+0.3}_{-0.4}$	49
O3a	GW190517_055101	$37.4^{+11.7}_{-7.6}$	$25.3^{+7.0}_{-7.3}$	$10.7^{+0.4}_{-0.6}$	21
O3a	GW190519_153544	$66.0^{+10.7}_{-12.0}$	$40.5^{+11.0}_{-11.1}$	$15.6^{+0.2}_{-0.3}$	48
O3a	GW190521	$95.3^{+28.7}_{-18.9}$	$69.0^{+22.7}_{-23.1}$	$14.2^{+0.3}_{-0.3}$	44
O3a	GW190602_175927	$69.1^{+15.7}_{-13.0}$	$47.8^{+14.3}_{-17.4}$	$12.8^{+0.2}_{-0.3}$	44
O3a	GW190706_222641	$67.0^{+14.6}_{-16.2}$	$38.2^{+14.6}_{-13.3}$	$12.6^{+0.2}_{-0.4}$	41
O3a	GW190720_000836	$13.4^{+6.7}_{-3.0}$	$7.8^{+2.3}_{-2.2}$	$11.0^{+0.3}_{-0.7}$	24
O3a	GW190727_060333	$38.0^{+9.5}_{-6.2}$	$29.4^{+7.1}_{-8.4}$	$11.9^{+0.3}_{-0.5}$	49
O3a	GW190728_064510	$12.3^{+7.2}_{-2.2}$	$8.1^{+1.7}_{-2.6}$	$13.0^{+0.2}_{-0.4}$	48
O3a	GW190828_063405	$32.1^{+5.8}_{-4.0}$	$26.2^{+4.6}_{-4.8}$	$16.2^{+0.2}_{-0.3}$	48
O3a	GW190828_065509	$24.1^{+7.0}_{-7.2}$	$10.2^{+3.6}_{-2.1}$	$10.0^{+0.3}_{-0.5}$	19
O3a	GW190915_235702	$35.3^{+9.5}_{-6.4}$	$24.4^{+5.6}_{-6.1}$	$13.6^{+0.2}_{-0.3}$	47
O3a	GW190924_021846	$8.9^{+7.0}_{-2.0}$	$5.0^{+1.4}_{-1.9}$	$11.5^{+0.3}_{-0.4}$	36
O3b	GW200129_065458	$34.5^{+9.9}_{-3.2}$	$28.9^{+3.4}_{-9.3}$	$26.8^{+0.2}_{-0.2}$	50
O3b	GW200202_154313	$10.1^{+3.5}_{-1.4}$	$7.3^{+1.1}_{-1.7}$	$10.8^{+0.2}_{-0.4}$	35
O3b	GW200219_094415	$37.5^{+10.1}_{-6.9}$	$27.9^{+7.4}_{-8.4}$	$10.7^{+0.3}_{-0.5}$	13
O3b	GW200224_222234	$40.0^{+6.9}_{-4.5}$	$32.5^{+5.0}_{-7.2}$	$20.0^{+0.2}_{-0.2}$	36
O3b	GW200311_115853	$34.2^{+6.4}_{-3.8}$	$27.7^{+4.1}_{-5.9}$	$17.8^{+0.2}_{-0.2}$	48

Table 3.1: List of O3 GW events used to generate the off-source injections of IS2. The columns from left to right show: (i) The LIGO-Virgo-KAGRA (LVK) observing run in which the event was detected, (ii) event name, (iii) primary component mass m_1, (iv) secondary component mass m_2, (v) HLV network matched-filter SNR† and (vi) number of off-source injections (out of 50) that satisfy $\mathrm{SNR}_{\mathrm{net}} \geq \mathrm{SNR}_{\text{cut-off}}$ and retained in IS2. Source parameter values displayed in the table are the median and the 90% symmetric credible intervals of the Bayesian posterior. Information in this table is copied directly from Table VI of GWTC-2 [36] (O3a events) and Table IV of GWTC-3 [38] (O3b events). †The network matched-filter SNR in this table is not to be confused with $\mathrm{SNR}_{\mathrm{net}}$ which denotes *injected* network SNR of IS1 and IS2 events.

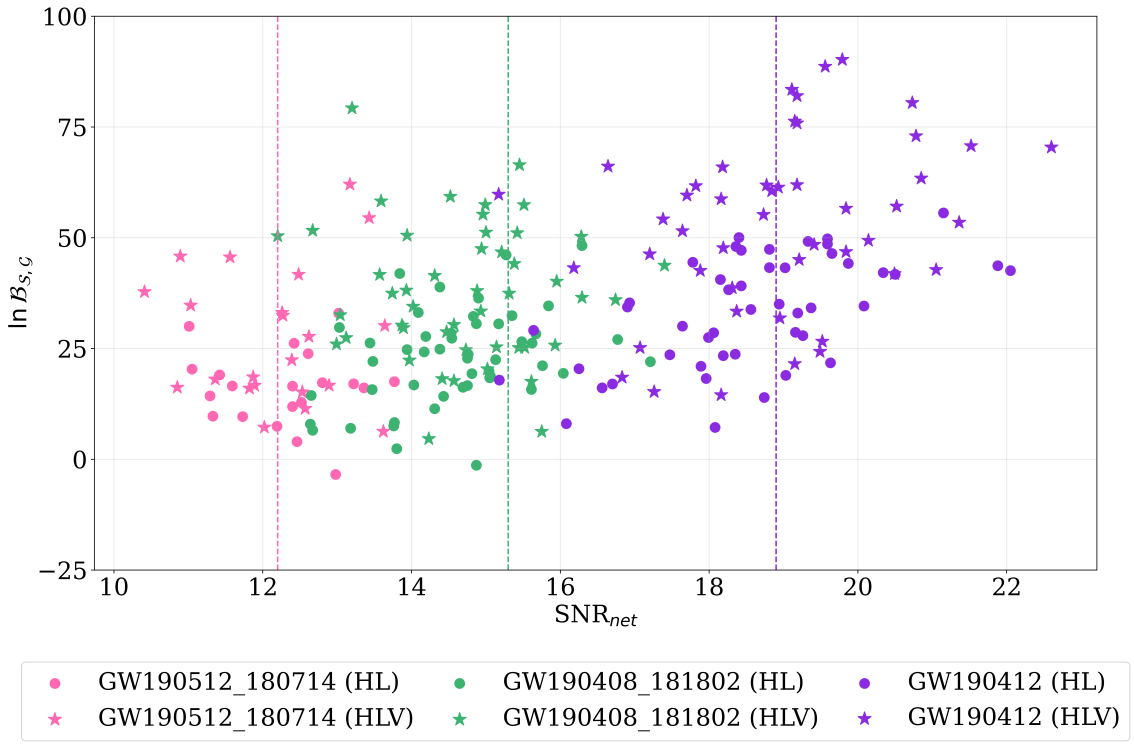

Figure 3.2: $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ versus $\mathrm{SNR}_{\mathrm{net}}$ for IS2 off-source injections of GW190512_180714 (pink), GW190408_181802 (green) and GW190412 (purple). The vertical dashed lines in the respective colors at $\mathrm{SNR}_{\mathrm{net}} = 12.2, 15.3, 18.9$ indicate the median HLV network match-filter SNRs of the GW events (from Table 3.1). The circles and stars correspond to HL and HLV injections respectively. Gaussian-noise-like events with $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = -500$ are not shown.

assigned $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = -500$ to indicate low significance.

Figure 3.2 shows the distribution of off-source injections in IS2 for each GW event in different colors. To avoid clutter, we show only three arbitrarily selected events with contrasting HLV network match-filter SNRs from Table 3.1, namely GW190512_180714 (pink), GW190408_181802 (green) and GW190412 (purple). Each circle (star) data point correspond to an individual HL (HLV) injections. For each GW event in Figure 3.2, the off-source injection $\mathrm{SNR}_{\mathrm{net}}$ are distributed within an approximate range of ± 5 from their respective median HLV network match-filter SNR, indicated by the vertical dashed lines in corresponding colors. The $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ also scales with $\mathcal{I}$, consistent with Ref. [3]. According to Table 3.1, the three events in Figure 3.2 also have comparable number of off-source injections in IS2. However, the number of injections for GW190512_180714 (pink) is visibly lower than the other two events, because the plot excludes Gaussian-noise-like events with $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = -500$. GW190512_180714 has the lowest network match-filter SNR of the three events, so its offsource injections in both the HL and HLV networks also have comparably lower $\mathrm{SNR}_{\mathrm{net}}$. Hence, the *BayesWave* evidences favours the Gaussian noise model more strongly than the signal model for a larger proportion ($\sim$50%) of GW190512_180714's offsource injections c.f. $\sim$0–5% for the other two events.

The comparison of *BayesWave*'s detection significance (P_{FA}) between the HL and HLV networks is presented in Section 3.6, for all 18 O3 GW events in IS2.

3.4. Background measurements

In this section, we discuss the suitability of using P_{FA} (as opposed to FAR) as a significance measure for the purpose of our analysis. We then present and discuss the noise background measurements. Using the dataset described in Section 3.3.1, we obtain the distribution of P_{FA} as a function of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$.

P_{FA} is the probability that a trigger of a given detection statistic ($\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$) is a false alarm i.e. non-astrophysical. In the context of hypothesis testing, P_{FA} represents the false positive rate (type I error) and is a dimensionless quantity by definition. In contrast, FAR measures the *temporal* frequency of false alarms producing a detection statistic value equal to or higher than a specified GW candidate event [22]. In other words, FAR is a time-average quantity which conflates *BayesWave*'s performance with engineering factors such as the detector glitch rate. As discussed in Section 3.3.1, *BayesWave* is not suitable for a full all-sky search of an observational dataset and is used instead to follow up triggers identified by other burst search pipelines like cWB. In this study, we measure *BayesWave*'s background for the HL and HLV networks using populations of background triggers arbitrarily downselected from the cWB all-sky analysis of the respective O3a time-shifted background data. Since P_{FA} is time-independent and marginalizes over the number of triggers analyzed, it relates directly to how *BayesWave* is used in this study. It is therefore more appropriate to compare *BayesWave*'s performance between the HL and HLV networks using P_{FA} as a measure of detection significance[6].

Figure 3.3 shows the HL and HLV network background as measured by *BayesWave* with the background trigger datasets described in Section 3.3.1. P_{FA}, plotted on the vertical axis, is computed as the fraction (i.e. per-trigger probability) of non-astrophysical triggers in the background exceeding the corresponding $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ on the horizontal axis. We restrict the plot to $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} > -20$, the range relevant to real astrophysical signals. Although not shown in Figure 3.3, triggers with $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} < -20$ are included in the denominators for computing P_{FA}, that is 1008 and 1134 respectively for HL and HLV. To estimate the uncertainties in our background measurements, we conventionally assume the detector noise background can be modeled as a Poisson process. The shaded regions show the 1 σ Poisson uncertainty region for HL and HLV in corresponding colors. In Appendix 3.A2, we show the implementation of P_{FA} in Poisson statistics as opposed to FAR, along with the derivation of the Poisson uncertainty regions.

The background measurements show that P_{FA} is higher for HLV than for HL at all $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ as the occurrence of background triggers increases with the number of detectors. As a result, events detected by the HLV network need to attain a higher $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ in order to achieve the same significance (P_{FA}) as the HL network. For example, to achieve $P_{\mathrm{FA}} = 0.1$, a HL event requires $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = 25.6$; c.f. $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = 43.1$ for HLV. Additionally, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ of the HLV background triggers are higher overall compared to HL. This is because the increased trigger frequency in HLV results in the increased likelihood of coincident triggers which more closely resemble coherent signals, and are therefore recovered with higher detection statistics by *BayesWave*. Furthermore, the misalignment of the Virgo

[6] P_{FA} should not be confused with the definition of false alarm probability, $\mathrm{FAP} = 1 - \exp(-\mathrm{T}_{\mathrm{obs}} \times \mathrm{FAR})$ used in other analysis pipelines e.g. PyCBC [108]. FAP is the probability of finding *one or more* noise background events with significance equal to or higher than FAR (of a candidate event) within an observation period T_{obs}.

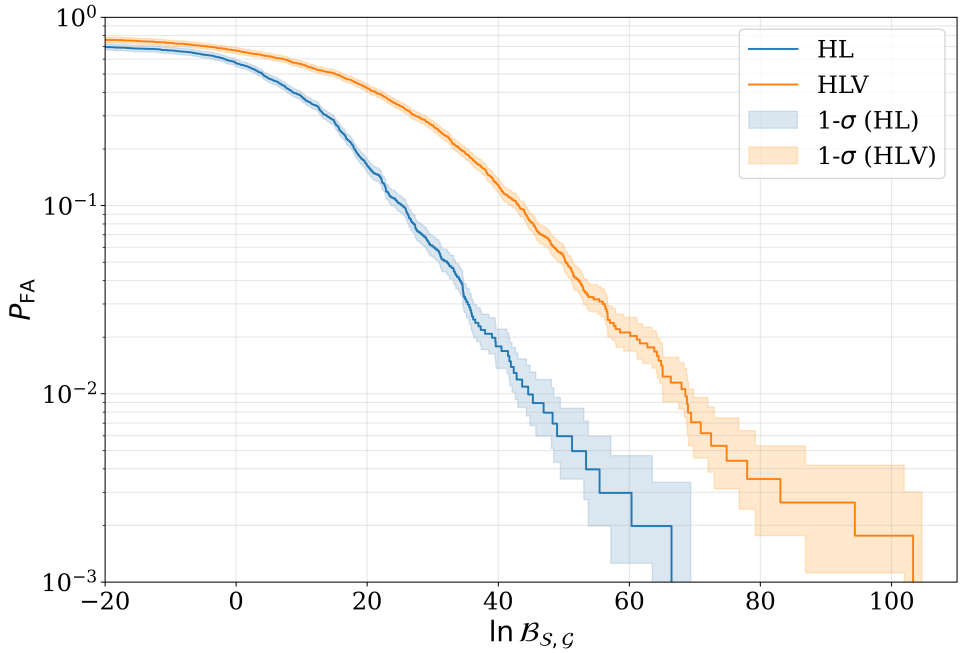

Figure 3.3: Background measurements for the *BayesWave* algorithm. The blue (orange) curve corresponds to the HL (HLV) background measured using the downselected O3a background triggers described in Section 3.3.1. The shaded bands show the 1-σ Poisson uncertainty regions for each network in corresponding colors.

detector senses a different signal polarization to the two co-aligned LIGO detectors, thus imposing a less stringent constraints on signal coherence. This reduces the efficiency of HLV in discriminating coincident glitches from signals.

3.5. *BayesWave* detection efficiency with BBH waveforms

3.5.1. Constructing efficiency curves

In Ref. [139], the performance of a hierarchical pipeline consisting of cWB and *BayesWave* is quantified using efficiency curves, which show the fraction of injected signal waveforms recovered above various significance thresholds. We use the same approach in this work to study the independent performance of *BayesWave*. Using IS1 described in Section 3.3.2, we construct efficiency curves for the HL and HLV networks by plotting $P_{\rm det}$ as a function of $P_{\rm FA}$.

As per Ref. [139], $P_{\rm det}$ is calculated as the fraction of astrophysical events recovered with detection statistic above a threshold. For *BayesWave*, this threshold is set by the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ corresponding to a user-selected significance i.e. $P_{\rm FA}$. As noted in Figure 3.3, the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ threshold is higher for HLV than for HL at a fixed $P_{\rm FA}$. The following example shows how $P_{\rm det}$ is computed for an arbitrary but representative choice $P_{\rm FA} = 0.2$.

Figure 3.4 shows histograms of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ for the HL (blue) and HLV (orange) O3a background triggers in the top panel, and for IS1 in the bottom panel. The $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ thresholds

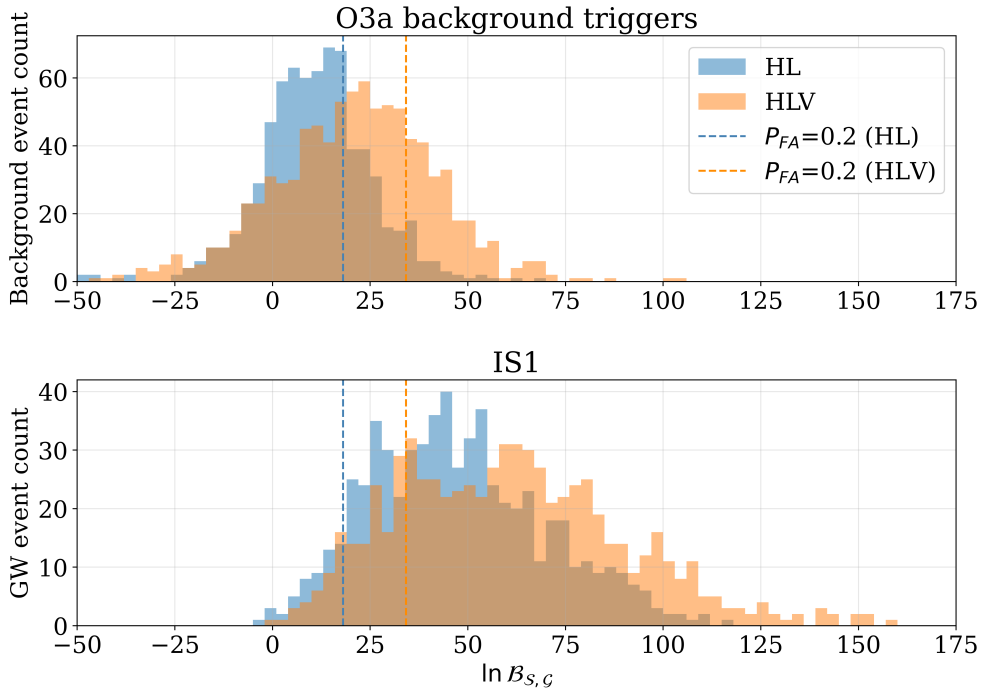

Figure 3.4: Worked example: computing one representative point on the efficiency curve for a significance threshold $P_{\mathrm{FA}} = 0.2$. The top panel shows the histogram of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ for the O3a background triggers described in Section 3.3.1. The bottom panel shows the histogram of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ for IS1. The HL and HLV network histograms are color-coded blue and orange respectively. In both panels, the vertical dashed lines at $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = 18.1$ (HL) and 34.2 (HLV) indicates the threshold for $P_{\mathrm{FA}} = 0.2$. The fraction of injections to the right of the thresholds in the bottom panel yields $P_{\mathrm{det}} = 0.74$ (HL) and 0.71 (HLV).

for $P_{\mathrm{FA}} = 0.2$ is set by the background triggers in the top panel. In both panels, we indicate the thersholds by the vertical dashed lines at 18.1 (HL, blue) and 34.2 (HLV, orange). With the HL and HLV thresholds established, we turn to the bottom panel of Figure 3.4 where we compute P_{det} as the fraction of IS1 injections detected by HL (HLV) greater than the threshold, i.e. to the right of the blue (orange) vertical line. We find $P_{\mathrm{det}} = 0.74$ and 0.71 for HL and HLV respectively. The procedure is repeated for P_{FA} in the range $0 \leq P_{\mathrm{FA}} \leq 1$ to construct the efficiency curves for HL and HLV.

3.5.2. Efficiency analysis

The efficiency curves of IS1 for the characterization of *BayesWave*'s overall burst detection efficiency is shown in Figure 3.5. The blue and orange curves correspond to the HL and HLV networks respectively. To indicate the error margins of P_{FA} from the background measurements, we carry over the 1-σ Poisson uncertainty regions onto the horizontal axis of the efficiency curves. From the background measurements, we also noted that the minimum $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ required to achieve a given significance reduces with increasing tolerance for P_{FA}. Therefore the efficiency curves show that P_{det} increases with P_{FA}, as more events in IS1 satisfy the reduced $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ threshold. The cluster of data points at $P_{\mathrm{det}} = P_{\mathrm{FA}} = 1$, disjointed from the rest of the efficiency curves, is an artifact

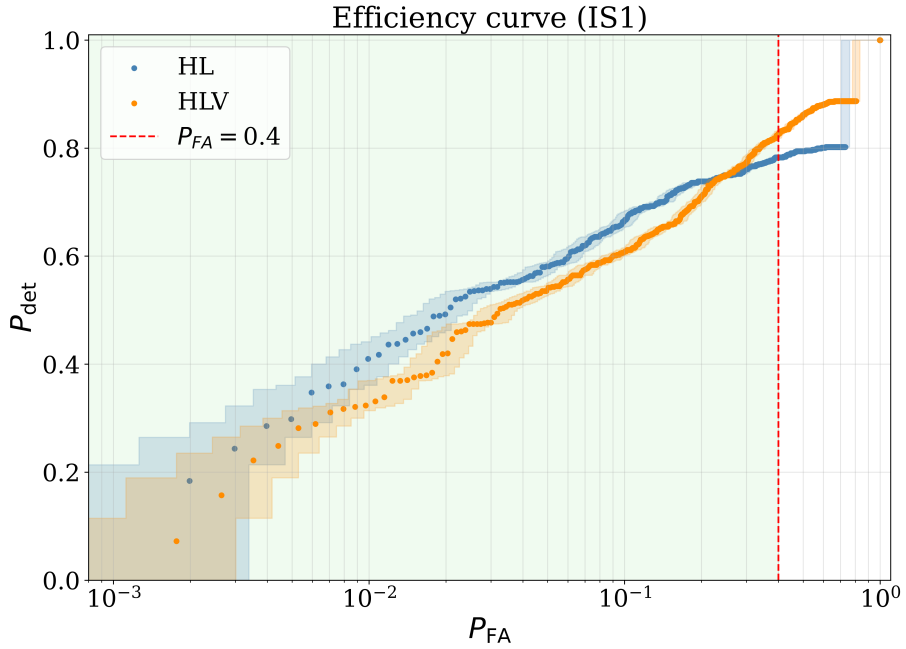

Figure 3.5: *BayesWave* efficiency curves constructed using IS1 for the HL (blue) and HLV (orange) networks. The shaded bands with matching colors are the 1-σ Poisson uncertainty regions for $P_{\rm FA}$, same as in Figure 3.3. The region where $P_{\rm FA} \leq 0.4$ is shaded green to indicate astrophysical relevance.

from assigning an arbitrarily low significance of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = -500$ to Gaussian-noise-like events. As discussed in Section 3.3.2, these events occupy 20% (11%) of the HL (HLV) IS1 injections. Therefore we observe a discrete jump in the fraction of recovered injections i.e. $P_{\rm det}$ of HL (HLV) from 0.80 (0.89) to 1. We also note a gap in the $P_{\rm FA}$ between the cluster of data points and the point before $P_{\rm FA} = 1$. This is because the second lowest $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ for the HL and HLV IS1 injections are of order -10^1 according to Figure 3.1.

In order to assess the overall detection efficiency of *BayesWave* with the HL and HLV networks, we focus on the region where $P_{\rm FA}$ is low enough to be astrophysically relevant. We arbitrarily define this region to be where $P_{\rm FA} \leq 0.4$ as indicated by the green shading. In this region, $P_{\rm det}$ of HL is generally higher than HLV, but the opposite is true for $P_{\rm FA} \gtrsim 0.25$. By quantifying the ratio between HL and HLV $P_{\rm det}$ for all data points in $P_{\rm FA} \leq 0.4$, we find that the HL network is only 1.02 times (i.e. 2%) more efficient in detecting IS1 injections than the HLV network on average. Hence, there are no significant differences in *BayesWave*'s overall detection efficiency with a two- or three-detector configuration.

To justify our findings, we show the event-wise comparison of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ between the HL and HLV networks for IS1 in Figure 3.6, color-coded according to the $\rm SNR_{net}$ of HLV[7]. The dashed diagonal line indicates where $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ is equal in both networks. For a specified detection significance ($P_{\rm FA}$), the plot can be divided into four quadrants by the corresponding $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ thresholds of the HL and HLV networks. Using $P_{\rm FA} = 0.2$ again as a representative example, we indicate the HL (HLV) threshold with a blue (orange)

[7]The HL and HLV network $\rm SNR_{net}$ are equally representative of the ensemble $\rm SNR_{net}$ of IS1 (see Figure 3.1). Thus we show only the HLV network $\rm SNR_{net}$ in Figure 3.6 to avoid clutter.

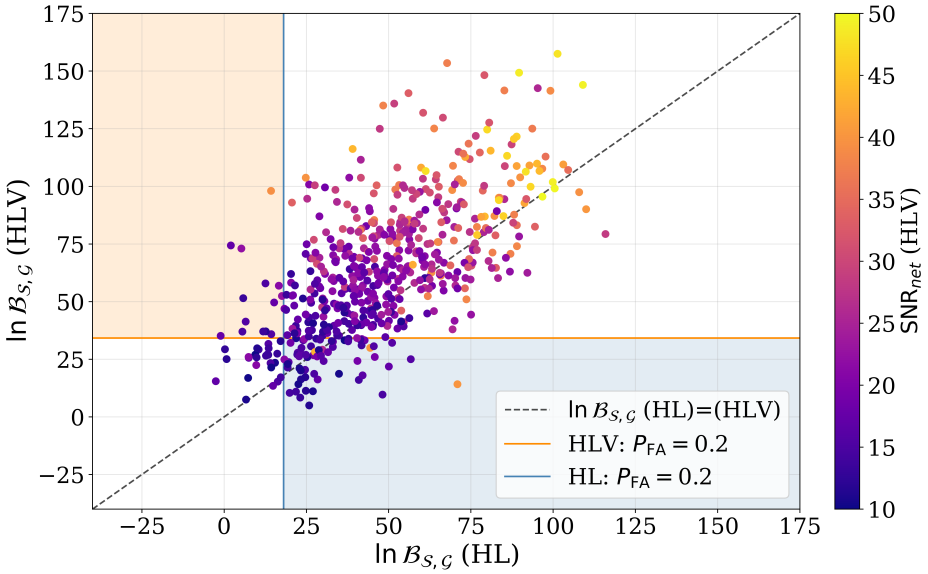

Figure 3.6: Log signal-to-glitch Bayes factor, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ of the HLV network versus the HL network for IS1. The color bar shows $\mathrm{SNR}_{\mathrm{net}}$ in HLV for each injection. The diagonal line indicates equal $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ for both networks. The dashed lines at $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}(\mathrm{HL}) = 18.1$ and $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}(\mathrm{HLV}) = 34.2$ indicate the thresholds for $P_{\mathrm{FA}} \leq 0.2$ with the respective networks. Gaussian-noise-like events with $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = -500$ are excluded in this plot.

solid line in Figure 3.6. The quadrants classify IS1 events based on their detectability. A successful detection in the HL (HLV) network is when the event $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ is higher than the detection threshold set by the blue (orange) line. By this definition, events in the top left quadrant (shaded orange) are detected by the HLV network only; the bottom right (shaded blue) by the HL network only; the top right by both networks and the bottom left by neither. We note that a fraction of events (in the blue shaded region) are only detected by HL despite having higher $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ in HLV. This is because a successful detection with the HLV network requires the increased $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ to satisfy a higher detection threshold to achieve the same significance as HL. In other words, the advantage of increased $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ in larger detector networks is offset by the higher detection thresholds due to the increased probability of false alarms in the background. This explains why the efficiency curves are comparable between the HL and HLV networks.

From Figure 3.6, we can also see that $\mathrm{SNR}_{\mathrm{net}}$ affects detectability. The top right quadrant contains events with overall higher $\mathrm{SNR}_{\mathrm{net}}$ compared to the other quadrants. That is, events with higher $\mathrm{SNR}_{\mathrm{net}}$ and hence higher $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ are more likely to be detected by both HL and HLV. The remaining IS1 events with lower $\mathrm{SNR}_{\mathrm{net}}$ are distributed across the other three quadrants where they fall short of at least one of the HL or HLV detectability thresholds, as indicated by the orange and blue dashed lines respectively. This is true for all P_{FA}. We discuss the cases where events are only detected by one of the two configurations. For events detected only by HLV (orange shaded region), it is straightforward to argue that adding Virgo increases the sensitivity of the network to the signal that are too low to be detected by HL. Consequently, this increases $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ and boosts the detection significance past the required threshold. For the less intuitive case where events are detected only by HL (blue shaded region), we need to justify for two

scenarios: (i) where $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ for HLV is higher than HL and (ii) vice versa. The former is discussed above. The latter suggests that the removal of Virgo boosts the signal evidence. This occurs when the addition of Virgo introduces non-Gaussian noise artifacts across the network which outweighs the sensitivity gain for the signal. These effects matter most for low $\mathrm{SNR}_{\mathrm{net}}$ injections.

In summary, the efficiency curves in Figure 3.5 show that *BayesWave*'s overall burst detection performance with the HL and HLV networks are comparable in the nominal astrophysically relevant range $P_{\mathrm{FA}} \leq 0.4$. This is because the noisier detector background of the HLV offsets the advantage of increased $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$, as revealed by the granular event analysis in Figure 3.6. Additionally we note that for low-$\mathrm{SNR}_{\mathrm{net}}$ injections at any given significance, adding an extra detector may tip them over or under the detection threshold unpredictably, due to a hard-to-quantify trade-off between the added noise and added sensitivity. High-$\mathrm{SNR}_{\mathrm{net}}$ injections, on the other hand, are more likely to be detected by both networks.

3.6. *BayesWave* detection significance of O3 GW events

The analysis with IS1 inferred that the overall detection efficiency of *BayesWave* is comparable between the HL (two-detector) network and HLV (three-detector) network. Using IS2 described in Section 3.3.2, we conduct a consistency test for the results of IS1 by comparing *BayesWave*'s detection significance of O3 GW events between the two network configurations.

The off-source injections in IS2 correspond to 18 independent O3 GW events. The final column of Table 3.1 shows the number of off-source waveforms available for each event. In order to measure the detection significance of these GW events according to *BayesWave*, we first quantify the significance for each off-source injection. This is done by comparing the recovered $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ in HL and HLV with the corresponding background measurements in Figure 3.3. To obtain a single-valued significance measurement for each GW event, we take the median[8] P_{FA} of the corresponding off-source waveforms. Figure 3.7 shows the median HLV P_{FA} versus that of HL. We use the interquartile range (IQR), that is the range encompassing the middle 50% of the off-source P_{FA} within each GW event, to represent the uncertainty in our measurements. The horizontal and vertical gray bars show the IQRs for the HL and HLV P_{FA} measurements respectively. We find that all data points are within close proximity of the diagonal line where P_{FA} is equal for HL and HLV. The size of the HL and HLV IQRs are also comparable. This suggests that *BayesWave*'s detection significance of O3 GW events are similar for both networks, further confirming that *BayesWave*'s burst detection performance with the HLV network does not exceed HL when the detector backgrounds are taken into account.

We also note that the median P_{FA} increases with decreasing $\mathrm{SNR}_{\mathrm{net}}$, because the colors of the points darkens as one moves from the bottom left to the top right of the plot. According to Figure 3.2, GW events with low network match filter SNR have low $\mathrm{SNR}_{\mathrm{net}}$ off-source injections that are generally more consistent with Gaussian noise.

[8] We show the median P_{FA} instead of the mean, because the median value excludes any biases introduced by the Gaussian-noise like events with $P_{\mathrm{FA}} = 1$, due to their arbitrarily low detection statistic $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = -500$.

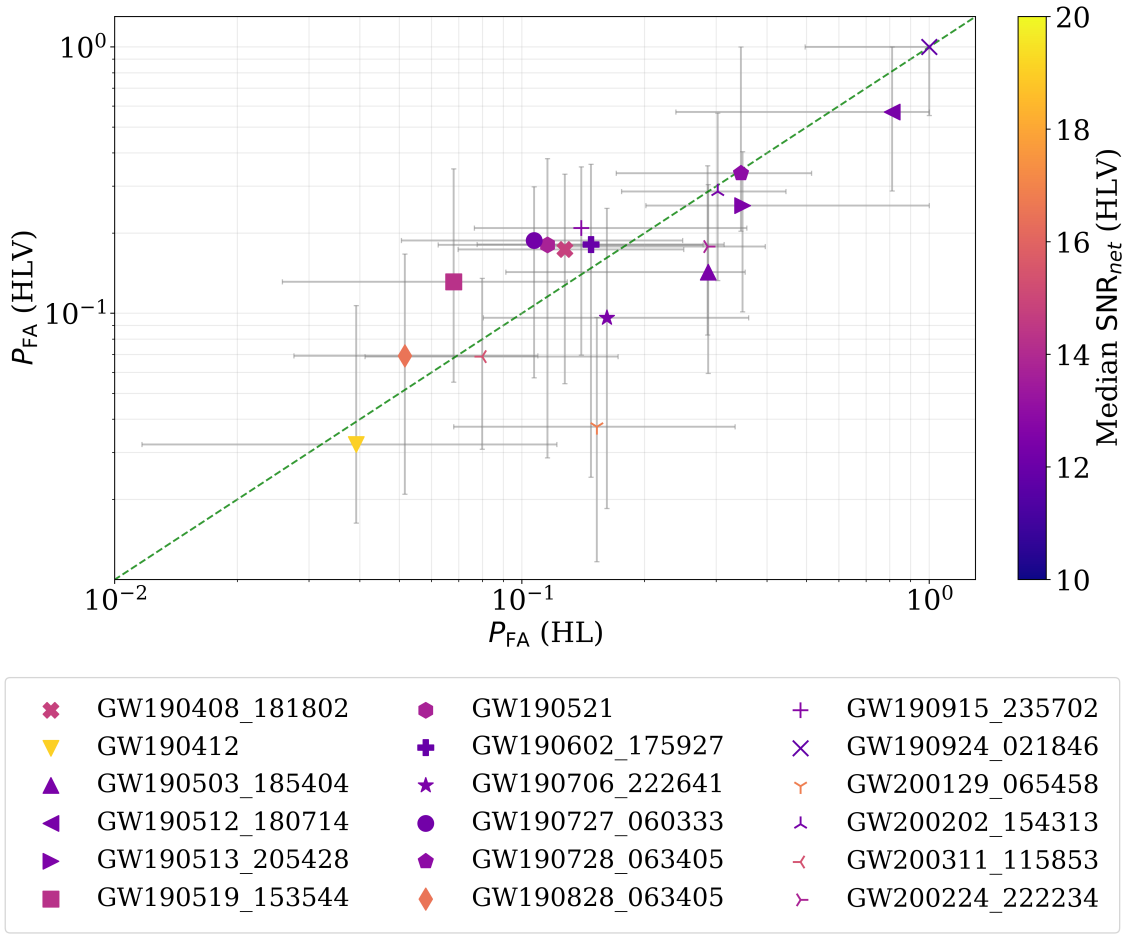

Figure 3.7: P_{FA} of the HLV network versus the HL network for O3 GW events in IS2. Each point represents a single GW event as shown in the legend and is color coded by the HLV $\mathrm{SNR}_{\mathrm{net}}$. The P_{FA} and $\mathrm{SNR}_{\mathrm{net}}$ shown are the medians of the off-source injections of the corresponding event; the horizontal (vertical) gray bars span the interquartile range of the HL (HLV) P_{FA} measurements. The diagonal line indicates equal P_{FA} for both networks.

Therefore the median P_{FA} of GW events in the top right of Figure 3.7 approaches unity. This observation is consistent with Figure 3.6, where events with low $\mathrm{SNR}_{\mathrm{net}}$ and hence low $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ are only detectable by both HL and HLV when higher P_{FA} are tolerated. Furthermore, the size of the IQRs are within the same order of magnitude as the median P_{FA} viz. the P_{FA} measurement uncertainties are larger for events with lower $\mathrm{SNR}_{\mathrm{net}}$ off-source injections (top right corner) compared to those with higher $\mathrm{SNR}_{\mathrm{net}}$ (bottom left corner). The wider IQRs suggest that the increased presence of Gaussian-noise-like injections not only reduces the astrophysical significance, but also increases the uncertainty in the significance measurements for GW events with low network match filer SNR.

Altogether IS2 shows that significance measurements with *BayesWave* is comparable for the HL and HLV networks, consistent with the findings of the IS1 efficiency curve analysis. We also find that P_{FA} and the uncertainty in its measurement increases with decreasing network match-filter SNR.

3.7. Conclusion and discussion

3.7.1. Summary of results

In practice, the source-agnostic *BayesWave* algorithm is used in conjunction with other search pipelines to enhance detection confidence of GW transients. In this work, however, we study the stand-alone performance of *BayesWave* with expanded detector networks. Detection confidence of *BayesWave* is assessed using the algorithm's detection statistics, the log signal-to-glitch Bayes factor $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$, which measures the extent of supporting evidence for the signal model over the glitch model. A previous study shows that $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ increases with increasing number of detectors, $\mathcal{I}$, in a network of GW interferometers [3]. However, the study did not account for the increase in glitch occurrence and the associated increase in false alarm detections, as more detectors are added to the network. This paper extends Ref. [3] with the goal of determining whether *BayesWave*'s overall burst detection performance is enhanced or reduced as $\mathcal{I}$ increases, when the detector noise background is taken into account. This is done by measuring the noise backgrounds produced by *BayesWave* and comparing the efficiency curves between the HL (two-detector) and HLV (three-detector) networks.

We obtain the noise backgrounds measurements of *BayesWave* for the HL and HLV networks by analyzing non-astrophysical triggers, downselected from the cWB analysis of the O3a time-slide background data. The background measurements show that per-trigger false alarm probability $P_{\rm FA}$ is higher in the HLV network than in HL, throughout the astrophysically relevant range $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \geq -20$. This is due to the increased likelihood of background triggers with an additional detector. We reiterate that the cWB algorithm is only used to downselect triggers for *BayesWave*'s background measurements, we do not investigate cWB's background and/or detection efficiency in this paper.

For the efficiency curve analysis, we implement *BayesWave* on a population of non-precessing and non-spinning phenomenological BBH waveforms (IS1) sampled from a parameter space detectable by the Advanced LIGO and Advanced Virgo detectors. IS1 is injected into segments of HL and HLV data spread out across all of O3a, to ensure comparability of the detection statistics with the background measurements. The efficiency curves plots detection efficiency, $P_{\rm det}$, of IS1 events as a function of the per-trigger false alarm probability, $P_{\rm FA}$, to characterize *BayesWave*'s performance over a range of significance thresholds. We find similar efficiency curves for the HL and HLV networks within a nominal significance range with plausible astrophysical implications i.e. $P_{\rm FA} \leq 0.4$. In other words, there are no major differences between *BayesWave*'s overall performance with HL and HLV. This counterintuitive finding is justified by event-wise comparison of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ between the HL and HLV IS1 injections in Figure 3.6. The plot reveals that the advantage of increasing $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ with $\mathcal{I}$ is offset by the increased $P_{\rm FA}$. Adding more detectors to the network increases the likelihood of noise events (i.e. false alarms). Therefore, events in larger detecter networks are required to satisfy higher detection thresholds to achieve the same significance as smaller networks. Additionally, the detectability of events by the HL and HLV networks at any given significance threshold ($P_{\rm FA}$) scales with $\rm SNR_{net}$. For events with low $\rm SNR_{net}$, the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ and hence $P_{\rm FA}$ in each detector network are more sensitive to subtle changes in detector noise variation. Therefore, the addition of Virgo can unpredictably tip an event over or under the HL or HLV signifi-

cance threshold.

To check for consistency with the efficiency analysis, we separately analyze a set of O3-like CBC waveforms (IS2), otherwise referred to as off-source injections. Parameters of off-source injections are sampled from the match-filter posteriors of 18 GW events from O3. We use $P_{\rm FA}$ to quantify *BayesWave*'s significance for each GW event. This is evaluated by comparing the $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ of their respective off-source injections with the O3a background measurements. The comparison of $P_{\rm FA}$ between HL and HLV reveals that *BayesWave* recovers all 18 events with similar significance from both networks. This result is consistent with the IS1 detection efficiency analysis.

Altogether, this study investigates the impact of glitches on the detection significance ($P_{\rm FA}$) and the overall performance of *BayesWave*, as a function of $\mathcal{I}$. From two independent analyses with IS1 and IS2, we conclude that there are no significant differences between *BayesWave*'s overall burst detection performance with the HL and HLV networks. Despite the improvement in detection statistic with the addition of Virgo, the associated increase in non-astrophysical background triggers raises the detection statistic threshold which the HLV network need to attain in order to achieve the same per-trigger $P_{\rm FA}$ as HL. Therefore the HLV configuration, despite having more detectors, does not have an advantage over HL in terms of detection efficiency. Our findings are consistent with previous studies [46, 203]. Although expanded detector networks improve accuracy of reconstruction and sky localization of the GW signal, Refs. [46] and [203] suggest that HL rejects glitches more efficiently compared to HLV and is therefore preferred in unmodeled burst searches to maximize detection efficiency. This is because HL comprises only of the co-aligned LIGO detectors with similar sensitivities to GW polarization components from all directions, therefore it poses more stringent constraints on signal coherence across the network. On top of that, the overall strain sensitivity of Virgo is lower than the two LIGO detectors in O3a, as shown in Figure 2 of Ref. [36]. This is a reason why the HLV network does not significantly outperform the HL network.

3.7.2. Future work

Although the comparison in this work shows that the HLV network does not surpass HL in burst-detection efficiency, it does not account for the longer duty cycles of larger networks. That is, the HLV network yields more data during which at least two detectors are simultaneously observing, compared to HL. With Virgo being less sensitive, the detection efficiency of the two-detector HV and LV networks is expected to be lower than that of HL. Nonetheless, future work could compare the joint performance of HLV's two-detector subsets (i.e. HL, HV, and LV) with the standalone performance of HL to fully leverage the advantages of expanded detector networks.

With the recently approved commissioning of LIGO-India with design sensitivity planned to match the LIGO detectors [82], it would be also worthwhile for prospective studies on *BayesWave*'s detection efficiency to compare HL with network configurations comprising three or more detectors of equal sensitivities. Furthermore, with improved detector sensitivity and the expansion of the global detector network, we anticipate an increase in correlated environmental glitches caused by Earth-scale events, such as lightning strikes [204]. Understanding BayesWave's overall burst detection performance in the presence of such glitches will be a valuable direction for future research.

In this work, we use trigger lists generated by the cWB algorithm from the LVK O3 all-sky burst search [46] to downselect triggers for *BayesWave* background measurements in this study. However, Ref. [203] conducted the same search using the cWB algorithm enhanced by machine-learning (ML) which shows improved overall search sensitivity compared to the standard cWB. We therefore suggest a complementary study to follow-up on whether the *BayesWave* background measurements can be improved if the triggers are downselected from the ML-enhanced cWB trigger list instead.

While *BayesWave* targets a broad range of unmodeled GW bursts, this study considers only CBC waveforms as they are the only source category detected in the LVK observing runs to date. One can generalize this study to alternative transient sources like supernovae and generic white noise bursts, but the analysis presented in this work is limited to comparing the overall trends of *BayesWave*'s independent performance between the HL and HLV networks. We did not study the sensitivity of *BayesWave* to specific types of burst signal because *BayesWave* is not used independently in practice, but rather to follow-up cWB triggers to enhance detection confidence. With promising outlooks for the ML-enhanced cWB [203] and O4 in progress, future work should consider assessing the joint performance of the ML-enhanced cWB algorithm with *BayesWave* for different types of burst sources as in Ref. [139].

Acknowledgements

This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation. Parts of this research were conducted by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), through project number CE170100004. The authors are grateful for computational resources provided by the LIGO Laboratory and supported by National Science Foundation Grants PHY-0757058 and PHY-0823459. Y. S. C. L. is supported by a Melbourne Research Scholarship and The University of Melbourne Women in Physics Award.

We thank the members of the cWB team for producing the background triggers that we used in our analysis: Shubhanshu Tiwari, Claudia Lazzaro, Marco Drago, Francesco Salemi, Gabriele Vedovato, Sergey Klimenko. We also thank Marek Szczepańczyk, Tyson Littenberg and Neil Cornish for their helpful comments.

3.A1. *BayesWave* configuration

To assist with reproducibility, we detail the *BayesWave* settings for the background measurements and injection analyses. The following settings are adapted from the *BayesWave* analysis used in the O3 all-sky burst search [46].

To down-select candidates from the cWB trigger list for *BayesWave*'s background measurements, we specify the significance threshold, $\eta_{\mathrm{c,threshold}} = 7$, as a first cut. We further reduce the dataset by keeping only a fraction of triggers satisfying $\eta_{\mathrm{c}} > \eta_{\mathrm{c,threshold}}$. This fraction is denoted by X in the main text.

For the signal injections, we use the O3a calibrated strain data for the LIGO Hanford (H1), LIGO Livingston (L1) and Virgo (V1) detectors [200]. The frame calibration

includes a noise subtraction procedure detailed in Ref. [188]. The Advanced LIGO (H1 and L1) noise subtraction targets noise from beam jitter, detector calibration lines and the main power grid line (at 60 Hz) [205]. For Advanced Virgo, we use low-latency (online) strain data which includes subtraction of frequency noise from the input laser, Michelson noise from displacement of the beam splitter mirrors, amplitude noise from auxiliary modulation and scattered light noise [206].

For all analyses, we set the low frequency cutoff at 20 Hz by convention [207]. The sampling rate is set at 2048 Hz to achieve a Nyquist frequency of 1024 Hz. For PSD estimation i.e. to construct the model $\mathcal{N}$, we employ the *BayesLine* algorithm. The *BayesWave* analysis segment length is set to 4 seconds, even though GW burst signals (especially CBCs) are typically shorter. This is to ensure that detector noise is relatively stationary in analysis segment for accurate prediction of the noise spectral density with *BayesLine*. Altogether, our search targets GW bursts signals with duration of milliseconds up to a few seconds, with frequencies in the 20-1024 Hz frequency band of Advanced LIGO and Advanced Virgo at O3a sensitivities.

3.A2. Poisson noise background

The Poisson process models a series of randomly occurring events where the average time between events are known, but not the exact time of arrival of each event. Events modeled as Poisson process are expected to have a probability mass function given by

$$P(n, \lambda) = \frac{\lambda^n \exp^{-\lambda}}{n!}. \tag{3.2}$$

Otherwise known as the Poisson distribution, Equation 3.2 measures the probability P of n number of events occurring within a population for a given rate parameter, $\lambda > 0$. In this context, 'population' refers to a group of events in a fixed temporal or spatial interval. By definition, λ is the expected number of events in a given population, independent of the type of interval specified i.e. it is dimensionless.

3.A2.1. $P_{\rm FA}$ vs. FAR in modelling Poisson noise

The noise background of the Advanced LIGO and Virgo detectors are modeled as a Poisson process in the standard LVK GW transient searches [35, 36, 38]. In modelling a Poisson noise background, $P_{\rm FA}$ and FAR play an analogous role of representing the rate of noise events, which directly influences the rate parameter λ. In the case of FAR where rate is measured in units of time, the time of observation $T_{\rm obs}$ is the interval required to obtain the expected number of noise events in the background, $\lambda = T_{\rm obs} \times \text{FAR}$. Conversely, $P_{\rm FA}$ measures the noise occurrence rate in units of events. Therefore $\lambda = N_{\rm obs} \times \text{FAR}$, where the interval is now given by the total number of events observed $N_{\rm obs}$. One can then show the relationship between $P_{\rm FA}$ and FAR as $P_{\rm FA} = (N_{\rm obs}/T_{\rm obs}) \times \text{FAR}$.

3.A2.2. Poisson uncertainty regions of $P_{\rm FA}$

Since the background triggers used for *BayesWave*'s background measurements in Section 3.4 are subsets of the cWB all-sky analysis of the full O3a time-shifted background

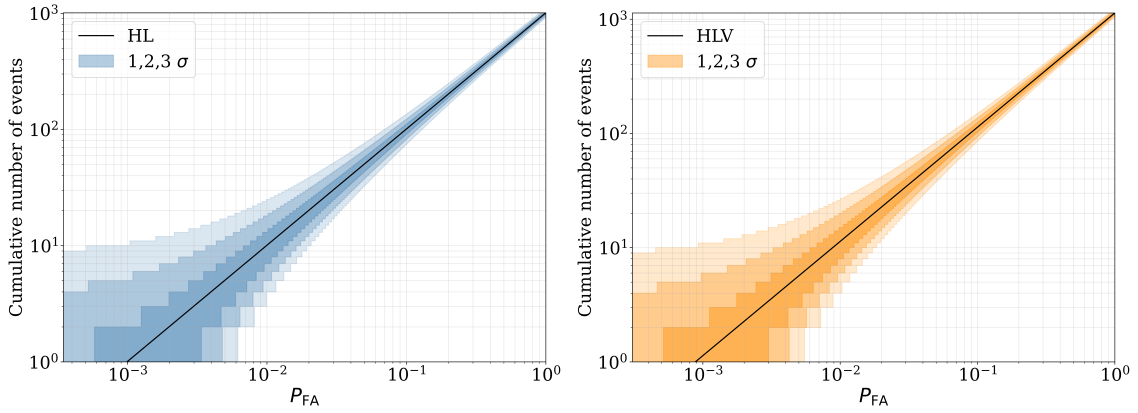

Figure 3.8: Cumulative number of events versus per-trigger false alarm probability, P_{FA} for the *BayesWave* background measurements of the HL (left panel, 1008 triggers) and HLV (right panel, 1134 triggers) network.

of the LIGO-Virgo network, we can thereby assume the triggers obey the Poisson distribution. Consequently, we can use the standard deviation (σ) of Equation 3.2 to represent the error margins of our P_{FA} measurements. We show the derivation as follows.

In the background measurements shown in Figure 3.3, P_{FA} (on the vertical axis) is computed as the fraction of background triggers recovered by *BayesWave* with $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ exceeding the corresponding threshold, $\ln \mathcal{B}^{*}_{\mathcal{S},\mathcal{G}}$ (on the horizontal axis) viz.

$$P_{\mathrm{FA}} = \frac{n(\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \geq \ln \mathcal{B}^{*}_{\mathcal{S},\mathcal{G}})}{n_{\mathrm{tot}}}, \tag{3.3}$$

where n_{tot} is the total number of triggers in the background dataset (the population). The numerator is essentially the expected occurrence of events exceeding $\ln \mathcal{B}^{*}_{\mathcal{S},\mathcal{G}}$, hence $\lambda = n(\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \geq \ln \mathcal{B}^{*}_{\mathcal{S},\mathcal{G}})$. One can then derive the 1-σ error margin for counting the number of events n exceeding $\ln \mathcal{B}^{*}_{\mathcal{S},\mathcal{G}}$ from the variance of the Poisson distribution:

$$\sigma = \sqrt{\sum_{n=1}^{\lambda} (n-\lambda)^2 P(n,\lambda)} = \sqrt{\lambda}. \tag{3.4}$$

Combining Equations 3.3 and 3.4, the 1 σ Poisson uncertainty region of P_{FA} for a given $\ln \mathcal{B}^{*}_{\mathcal{S},\mathcal{G}}$ is bounded by

$$\frac{\lambda - \sqrt{\lambda}}{n_{\mathrm{tot}}} \leq P_{\mathrm{FA}} \leq \frac{\lambda + \sqrt{\lambda}}{n_{\mathrm{tot}}}, \tag{3.5}$$

as indicated by the shaded regions in Figures 3.3 and 3.5.

To check for viability, we plot the cumulative number of triggers against P_{FA} in Figure 3.8 and the shaded regions show the 1-, 2- and 3-σ P_{FA} Poisson uncertainty regions. We compare our plots to the O3 backgrounds in Ref. [46] measured with inverse FAR. Even though we use a difference quantity to measure significance, the relative extent of the shaded regions are comparable. It is therefore appropriate to use the Poisson uncertainty described above as the error margins for our P_{FA} measurements.

Chapter 4

Dedicated-frequency analysis of gravitational-wave bursts from core-collapse supernovae with minimal assumptions

This chapter is a reproduction of

> *Y. S. C. Lee, M. Szczepańczyk, T. Mishra, M. Millhouse, and A. Melatos*, submitted to Physical Review D (2025)[1].

reformatted with the following changes only:

- **The text is styled to match the rest of this thesis.**
- **Appendix 4.A1 is added to discuss how GW signals are extracted from CCSNe using the quadrupole approximation. This appendix is not submitted for publication.**
- **Where relevant, brief descriptions of Appendix 4.A1 are added to Section 4.5 of the text.**

4.1. Abstract

Gravitational-wave (GW) emissions from core-collapse supernovae (CCSNe) provide insights into the internal processes leading up to their explosions. Theory predicts that CCSN explosions are driven by hydrodynamical instabilities like the standing accretion shock instability (SASI) or neutrino-driven convection, and simulations show that these mechanisms emit GWs at low frequencies ($\lesssim 0.25\,\mathrm{kHz}$). Thus the detection of low-frequency GWs, or lack thereof, is useful for constraining explosion mechanisms in CCSNe. This paper introduces the dedicated-frequency framework, which is designed to follow-up GW burst detections using bandpass analyses. The primary aim is to study whether low-frequency (LF) follow-up analyses, limited to $\leq 256\,\mathrm{Hz}$, constrain CCSN explosion models in practical observing scenarios. The analysis dataset comprises waveforms from five CCSN models with different strengths of low-frequency GW emissions induced by SASI and/or neutrino-driven convection, injected into the Advanced LIGO data from the Third Observing Run (O3). Eligible candidates for the LF follow-up must

[1] A version of this chapter, with minor changes not reflected in this thesis, is published as *Y. S. C. Lee, M. J. Szczepańczyk, T. Mishra, M. Millhouse, and A. Melatos*, Phys. Rev. D 112, 082006 (2025) [5].

satisfy a benchmark detection significance and are identified using the coherent WaveBurst (cWB) algorithm. The LF follow-up analyses are performed using the *BayesWave* algorithm. Both cWB and *BayesWave* make minimal assumptions about the signal's morphology. The results suggest that the successful detection of a CCSN in the LF follow-up analysis constrains its explosion mechanism. The dedicated-frequency framework also has other applications. As a demonstration, the loudest trigger from the SN 2019fcn supernova search is followed-up using a high-frequency (HF) analysis, limited to $\geq 256\,\mathrm{Hz}$. The trigger has negligible power below 256 Hz, and the HF analysis successfully enhances its detection significance.

4.2. Introduction

Core-collapse supernovae (CCSNe) are energetic explosions of massive stars ($\gtrsim 8M_{\odot}$) at the end of their lifetimes. CCSNe synthesize heavy elements during the explosion process, which are subsequently dispersed into the interstellar medium and inherited by the next generation of stars. CCSNe are also known to produce compact objects remnants like neutron stars and black holes; these remnants have been observed by the the Advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) [76] and Advanced Virgo [77] detectors. CCSNe play a vital role in stellar formation and evolution, which has led to extensive theoretical [208] and numerical [209–211] studies of their explosion mechanisms over the last few decades. It is believed that the initial collapse of the iron core produces an outward-bound hydrodynamic shock. Left to its own devices, however, the shock stalls within milliseconds due to energy dissipation, and fails to eject stellar material [210, 212]. This suggests the existence of a secondary process that revives the outward progress of the shock [213].

Constraining the explosion mechanisms of CCSNe requires accurate probing of the stellar interior, before and during the explosion. Electromagnetic signatures of the pre-explosion dynamics cannot penetrate the atmosphere of the progenitor. Neutrinos and gravitational-waves (GWs), on the other hand, propagate unscattered and unobstructed through the stellar atmosphere. Thus they carry information on the physical processes that drive the explosions. SN 1987A is the only CCSN observed to date accompanied by neutrino emissions [214–216], providing the first empirical evidence for a neutrino-driven explosion [217, 218]. Transient GW (burst) emissions due to asymmetrical motions in and around the compact stellar core are also expected during CCSNe [211, 219]. From the first observing run (O1) to the second half of the fourth observing run (O4b), the LIGO-Virgo-KAGRA (LVK) collaboration reported nearly 100 published GW detections [35–38] and over 200 candidates[2], but none are associated with CCSNe. The highly anticipated detection of GWs from CCSNe in the ongoing and upcoming observing runs incentivizes studies to improve GW burst analysis methods.

Physical processes occurring within CCSNe can be characterized by the frequency of their GW signatures [211]. Simulations have shown that the standing accretion shock instability (SASI) [220] and neutrino-driven convection [212] contribute to reviving stalled shock waves, ultimately driving CCSN explosions, and these processes generate low-

[2]The list of O4 GW candidates is available at: https://gracedb.ligo.org/superevents/public/O4/.

frequency signals ($\lesssim 0.25\,$kHz). Low-frequency GWs from CCSNe are therefore useful for constraining the explosion mechanisms. By dividing the frequency range into two bins above and below 0.25 kHz, Ref. [221] compares the low- and high-frequency reconstructions of CCSN GWs. The study finds that for CCSN models with stronger low-frequency emissions, such as SFHx [222] and mesa20 [223], the low-frequency reconstructions are more accurate than the high-frequency and full-band reconstructions. This suggests that bandpass analyses could be useful for enhancing detection sensitivity to frequency-specific GW signatures.

Dedicated-frequency studies to date, such as those in Ref. [221], assume Gaussian detector noise colored by simulated power spectral densities (PSDs). The results may differ for real observing scenarios, as the noise PSDs are not precisely known, and the data contain non-Gaussian noise transients (glitches), which are known to reduce detector sensitivity of GW bursts like CCSNe [38, 73]. The central goal of this paper is to explore dedicated-frequency analysis in realistic observing conditions. Standard GW burst analyses for CCSNe consider the frequency range $32-2048\,$Hz, and different classes of glitches occupy different regions of the time-frequency plane [96]. Hence one should ask whether excluding high frequencies from burst analyses can enhance sensitivity to low-frequency gravitational wave (GW) signatures, and vice versa. To address this question, we propose the *dedicated-frequency framework* which follows-up eligible detection candidates with bandpass analyses, e.g. the low-frequency (LF; $\leq 256\,$Hz) and high-frequency (HF; $\geq 256\,$Hz) follow-up analyses. By limiting the band, the outcomes of the LF (HF) analyses are not influenced by irrelevant HF (LF) glitches. We assess the astrophysical significance of detection candidates using empirical noise backgrounds drawn from real detector data.

The dedicated-frequency framework uses a hierarchical analysis pipeline comprising two independent algorithms: coherent WaveBurst (cWB) [114–116, 176, 224] and *BayesWave* [1, 2, 85], which are used by the LVK community to characterize generic GW bursts with minimal assumptions about the source and signal morphology [44–46]. The performances of cWB and *BayesWave* with CCSNe have been studied separately in Refs. [225] and [149] respectively. However, *BayesWave* is computationally intensive; it uses a reversible jump Markov Chain Monte Carlo (RJMCMC) algorithm to marginalize over its model dimensions. Therefore *BayesWave* is typically used to follow-up cWB triggers. The hierarchical implementation of cWB and *BayesWave* improves GW burst detection significance [139]. Here we present the methodology of the dedicated-frequency framework in the context of the hierarchical pipeline. As noted above, the detection of low-frequency GW signatures from the SASI and neutrino-driven convection, or lack thereof, places constraints on the explosion mechanism of a CCSN. Hence the primary objective of this paper is to study whether the detection of low-frequency GW signatures using the LF analyses can be used to constrain CCSN models. This study uses GW signals extracted from five distinct three-dimensional CCSN simulations, each with varying strengths of low-frequency GW emissions attributed to SASI and neutrino-driven convection. We also explore whether the HF analysis within the dedicated-frequency framework can enhance the detection significance of candidates with minimal low-frequency power. This study is conducted using the loudest trigger from SN 2019fcn. To ensure relevance to practical observing scenarios, both the LF and HF studies are carried out using real O3 data.

The rest of this paper is organized as follows. Section 4.3 discusses the physical processes within CCSNe, which produce GW signatures at different frequencies. Section 4.4 presents the dedicated-frequency framework in two parts: (i) an overview of the burst analysis pipelines in Section 4.4.1, and (ii) the applications and workflow in Section 4.4.2. Section 4.5 presents the methods and results of the LF study, demonstrating its application in constraining CCSN explosion mechanisms. Section 4.6 demonstrates an application of the HF follow-up using the loudest trigger of SN 2019fcn. We conclude in Section 4.7, with discussions of future work and broader applications of the dedicated-frequency framework.

4.3. Gravitational-wave signatures of core-collapse supernovae

CCSNe occur when stars with mass $8\,M_\odot \lesssim M \lesssim 100\,M_\odot$ enter the final stages of exoergic nuclear fusion [226]. If the iron core of a star exceeds the effective Chandrasekhar mass ($\sim 1.5\,M_\odot$) [227], the gravitational instability triggers core-collapse. The infall of stellar material compresses the core, forming a protoneutron star (PNS) [228, 229]. Once the core density exceeds nuclear density, further compression is no longer possible and a rebound occurs. The rebound process, otherwise known as core bounce, launches an outbound hydrodynamic shock wave [230]. When the shock encounters the still-collapsing outer core, it loses energy through the dissociation of heavy nuclei into nucleons and neutrinos [212, 231]. In other words, the outbound shock stalls and fails to produce an explosion. Therefore, a secondary shock-revival mechanism is required to trigger the explosion and this remains an active area of research [209, 213].

One way to study the physical mechanisms within CCSNe leading up to their explosion is to analyze the post-bounce GW emissions. The explosion mechanism is strongly influenced by progenitor rotation (see Ref. [213] and references therein). In this paper, we focus on slowly-rotating progenitors as they are expected to be the most common CCSN sources [232]. Their GW emissions exhibit two distinct spectral signatures: (i) a high-frequency component starting at $\sim 0.4\,\mathrm{kHz}$, which increases in frequency over time, and (ii) a low-frequency component confined to $\lesssim 0.25\,\mathrm{kHz}$. The two components are discussed in Sections 4.3.1 and 4.3.2 respectively.

4.3.1. High frequencies

Multi-dimensional numerical models have shown that high-frequency ($\gtrsim 0.4\,\mathrm{kHz}$) GW emissions predominantly originate within the PNS [233, 234]. The processes that give rise to such emissions include sustained Ledoux convection and convective overshoot (see Ref. [235] and references therein). Sustained Ledoux convection arises from persistent lepton gradients within the PNS [236]. Convective overshoot occurs when Ledoux convection extends into the stable outer layer [237]. Quadrupolar ($\ell = 2$) oscillations driven by intermittent aspherical accretion onto the PNS [235] also emit high-frequency GWs. The emission of high-frequency GWs typically commences $0.1-0.2\,\mathrm{s}$ post-bounce. In the early stages, the high-frequency emissions are driven by low-order ($n = 1, 2$) g-

modes[3]. As the PNS contracts and the equation of state (EOS) stiffens over time, the frequency of the GW signal increases. In the later stages, the high-frequency GW emissions are associated with the fundamental ($n = 0$) f-modes on the surface of the PNS [233, 238]. This evolutionary timeline varies based on the physical properties of the progenitor star (e.g. mass and rotation), as well as the PNS EOS. Altogether, the high-frequency GW emissions probe the evolution of the PNS structure and hence the EOS, but they do not explicitly reveal details about shock revival or the explosion itself.

4.3.2. *Low frequencies*

Two hydrodynamical instabilities have been identified as plausible shock-revival mechanisms: (i) the standing-accretion-shock instability (SASI) [220] and (ii) neutrino-driven convection [212, 239]. The SASI arises due to asymmetric radial density and velocity fluctuations, causing the initially spherical shock front to oscillate. The oscillation modulates the accretion flow, and the aspherical movement of matter results in the emission of low-frequency ($\lesssim 0.25\,\mathrm{kHz}$) GW [234, 235, 240]. GW emissions from SASI diminish upon the revival of the shock [238, 241]. The convection scenario, on the other hand, suggests the development of an entropy gradient as the outward shock weakens. The entropy gradient arises due to neutrino heating in the gain region[4], and the outflow of hot neutrinos revives the shock. The convective instabilities also modulate the accretion flow [242], but they produce weaker low-frequency GWs [235]. While the physical mechanisms behind the low-frequency emissions cannot be fully disentangled, the observation of low-frequency GW signatures offers insights into the pre-explosion processes and the timing of shock revival, which are useful for constraining the CCSN explosion mechanism.

In some slowly-rotating progenitors, Ledoux convection occurring shortly ($\lesssim 10\,\mathrm{ms}$) after core bounce also emits GWs. This short-lived convection, lasting $\lesssim 50\,\mathrm{ms}$, is otherwise known as prompt convection and is typically followed by a quiescent phase lasting $\sim 0.1\,\mathrm{s}$, with minimal GW emissions before the onset of the low- and high-frequency emissions described above. Figure 4.1 shows where the different GW signatures discussed above are located in the time-frequency plane. Simulations have showed that the peak GW frequency of prompt convection varies based on the EOS and numerical perturbation seeds, with values ranging across $0.1-0.7\,\mathrm{kHz}$ [219, 234, 238, 243]. Indeed, the GW emissions of the prompt convection may overlap with the low-frequency emissions from SASI and neutrino-driven convection. However, as depicted in Figure 4.1, the emission from prompt convection is temporally distinct and therefore distinguishable from other low-frequency emissions. Moreover, due to its brief duration, prompt convection contributes only a small fraction of the overall signal power [235, 238].

4.4. Dedicated-frequency framework

As discussed in Section 4.3, GW signatures at different frequencies probe different interior dynamics of CCSNe. Can one improve the characterization of frequency-specific

[3]Restoring forces of g-mode oscillations are exerted by buoyancy i.e. gravity.

[4]The gain region is where the energy gain through neutrino absorption exceeds the energy loss through neutrino emission [208].

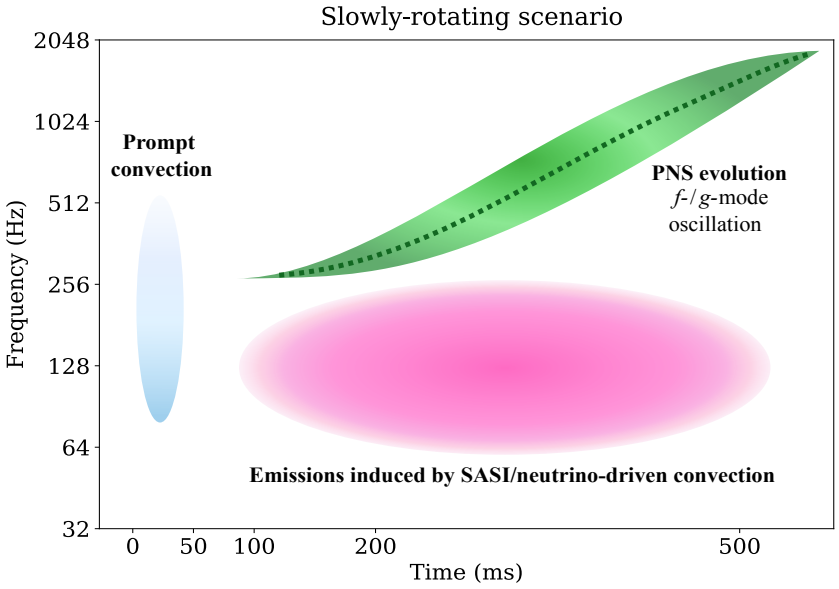

Figure 4.1: Schematic of GW emissions from CCSNe with slowly-rotating progenitors in the time-frequency plane. The scales on the axes are representative only. The actual durations and spectra of the GW signatures vary depending on the physical properties of the CCSN progenitor, and not all CCSNe will exhibit every signature shown.

GW signatures by ignoring contributions from irrelevant frequency ranges? Here, we introduce the dedicated-frequency framework, which is designed to follow-up GW burst detections with bandpass analyses. In Section 4.4.1, we overview the relevant burst analysis pipelines used in the framework. In Section 4.4.2 we discuss the applications and workflow of the dedicated-frequency analyses.

4.4.1. Hierarchical burst analysis pipeline

The LVK all-sky burst searches [44–46] implement multiple minimally-modeled burst analysis algorithms to ensure consistency. Two algorithms are coherent WaveBurst (cWB) and *BayesWave*. The hierarchical pipeline, in which *BayesWave* is used as a follow-up to cWB, enhances GW burst detection confidence [139], especially for complex waveforms (e.g. binary black holes and generic white noise bursts) with low signal-to-noise ratio (SNR).

The cWB algorithm uses the multi-resolution *WaveScan* transform [125] to compute excess power and cross-power statistics in the time-frequency data, and employs the constrained maximum likelihood formalism [115, 244] to reconstruct the signal waveforms and sky location. Each identified event is tagged with estimated summary statistics, describing the time-frequency structure, signal strength, and coherence across the multi-detector network. In order to limit the trigger production due to non-stationary detector noise, cWB implement vetoes related to some of the summary statistics to reduce excess background and false alarms. For example, one may discard triggers with network cross-correlation coefficient, $c_c = E_c/(E_c + E_n)$, below a nominal threshold ($= 0.5$ in this study), where E_c and E_n denote the coherent signal energy and residual noise energy respectively [115]. Events passing the summary-statistic vetoes are

ranked by cWB's main detection statistic η_c, which is effectively the coherent network SNR, defined by [224, 245]

$$\eta_c = \sqrt{\frac{E_c}{1 + \tilde{\chi}^2(\max\{1, \tilde{\chi}^2\} - 1)}}. \tag{4.1}$$

The reduced chi-squared statistic $\tilde{\chi}^2 = E_n/N_{DoF}$ in Equation 4.1 serves as an auxiliary glitch veto. For coherent signals, $\tilde{\chi}^2$ is approximately unity because the residual noise energy E_n is expected to obey the χ^2-distribution, with degrees of freedom N_{DoF} proportional to the number of *WaveScan* pixels used in the event reconstruction. Conventionally, events with $\tilde{\chi}^2 > 2.5$ are vetoed [160].

On top of the standard cWB vetoes, we use a machine-learning classification algorithm called eXtreme-Gradient Boost (XGBoost) to further distinguish GW signals and noise transients, i.e. to reduce false-alarm triggers [224, 245, 246]. The XGBoost model for generic burst searches is trained on a representative set of background noise events and stochastic white-noise-burst (WNB) signals[5] that do not correspond to any known GW sources [203]. Signal and noise classification with XGBoost is achieved by applying additional post-production vetoes on the cWB summary statistics, and the detection statistic reduces to

$$\eta_r = \eta_c W_{XGB}. \tag{4.2}$$

In Equation 4.2, W_{XGB} denotes the XGBoost penalty factor, which has a value between 0 (noise) and 1 (signal) [245]. Unless stated otherwise, all cWB analyses in this paper incorporate the XGBoost optimization and use η_r as the detection statistic.

BayesWave, on the other hand, takes a Bayesian approach when characterizing non-Gaussian features in the data. The algorithm reconstructs the data $\boldsymbol{d}$ using three independent models, namely the coherent signal plus Gaussian noise ($\mathcal{S}$) model, the incoherent glitch plus Gaussian noise ($\mathcal{G}$) model and the pure Gaussian noise model ($\mathcal{N}$). The $\mathcal{S}$ and $\mathcal{G}$ *BayesWave* models are constructed using sine-Gaussian wavelet frames [1]. A parallel-tempered Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm is used to sample the number of wavelets N and wavelet parameters for model $\mathcal{M} \in \{\mathcal{S}, \mathcal{G}, \mathcal{N}\}$. The associated Bayesian evidence $p(\boldsymbol{d}|\mathcal{M})$ is then computed using thermodynamic integration [136]. Model selection in *BayesWave* is conducted by comparing the Bayes factor between the models. The detection statistic of *BayesWave* is the log Bayes factor between $\mathcal{S}$ and $\mathcal{G}$,

$$\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = \ln p(\boldsymbol{d}|\mathcal{S}) - \ln p(\boldsymbol{d}|\mathcal{G}), \tag{4.3}$$

which scales not only with the network SNR, but also the model complexity quantified by N, and the number of detectors $\mathcal{I}$, viz. $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \sim \mathcal{O}(\mathcal{I}N \ln \mathrm{SNR}_{net})$ [3, 85].

Since the RJMCMC algorithm in *BayesWave* is computationally intensive, *BayesWave* is not used to analyze extended data segments, e.g. all-sky searches. Instead, *BayesWave* serves as a follow-up tool for analyzing targeted data segments flagged by other burst searches. In this work, we first run cWB to analyze the full dataset, and then use *BayesWave* to follow up on triggers that satisfy a nominal cWB detection threshold. The

[5] WNBs are band-limited and temporally localized signals with flat spectra that resemble white noise, but with power exceeding the average noise floor.

η_r in cWB compares signals against Gaussian noise, whereas the log Bayes factor $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ in *BayesWave* compares signals against glitches. This sequence of analyses is called the hierarchical pipeline and is commonly used in the LVK all-sky burst searches [44–46].

4.4.2. Applications and workflow

The dedicated-frequency framework is designed to be a follow-up tool, used exclusively for events that satisfy a benchmark significance criterion (e.g. FAR $\leq 1\,\mathrm{yr}^{-1}$). In other words, the dedicated-frequency framework does not redefine the standard LVK detection criteria; it follows up standard detections. In this framework, eligible candidates are identified using the standard (full-band) burst analysis which covers the band 32–2048 Hz. Once identified, they can be followed up with low-frequency (LF) and high-frequency (HF) analyses, below and above 256 Hz, respectively. The LF-HF boundary can be customized to study different sources and sceintific questions. For dedicated-frequency follow-ups of CCSNe, the LF and HF analyses distinguish SASI and neutrino-driven convection ($\lesssim 0.25$ kHz) from PNS oscillations ($\gtrsim 0.4$ kHz).

The dedicated-frequency framework implements the hierarchical pipeline as follows. cWB identifies triggers by clustering contiguous time-frequency pixels with coherent power exceeding the Gaussian noise floor; a trigger is recorded when the accumulated SNR of nearby pixel clusters surpasses a nominal threshold. A dedicated low-frequency search using cWB in its current configuration may not effectively detect SASI and neutrino-driven convection emissions, especially in broadband signals, due to insufficient power clustering in the low-frequency range. However, cWB has a low runtime and well-suited for processing extensive datasets. Therefore, we use the full-band cWB analysis to identify triggers, and then use *BayesWave* to follow-up on eligible (i.e. astrophysically-relevant) triggers with the dedicated-frequency (LF/HF) analyses. *BayesWave* is used for the dedicated-frequency follow-ups, because the successive application of cWB followed by *BayesWave* improves detection significance for signals with non-trivial time-frequency spectrograms [139].

In *BayesWave*, the coherent signal ($\mathcal{S}$) and incoherent glitch ($\mathcal{G}$) models are constructed by summing a set of continuous sine-Gaussian wavelets. A sine-Gaussian wavelet is intrinsically parameterized by its central frequency. Therefore, when restricting the band to LF or HF, we restrict the central frequency prior, not the overall wavelet spectrum. Consequently, some wavelet power during an LF analysis may leak into the HF band, and vice versa. For a model $\mathcal{M} \in \{\mathcal{S}, \mathcal{G}, \mathcal{N}\}$ parameterized by $\boldsymbol{\theta}^{\mathcal{M}}$, the Bayesian evidence is given by

$$p(\boldsymbol{d}|\mathcal{M}) = \int d\boldsymbol{\theta}^{\mathcal{M}} p(\boldsymbol{\theta}^{\mathcal{M}}|\mathcal{M}) p(\boldsymbol{d}|\boldsymbol{\theta}^{\mathcal{M}}, \mathcal{M}), \tag{4.4}$$

where $p(\boldsymbol{\theta}^{\mathcal{M}}|\mathcal{M})$ is the prior and $p(\boldsymbol{d}|\boldsymbol{\theta}^{\mathcal{M}}, \mathcal{M})$ is the likelihood. In *BayesWave*, the likelihood is calculated in the frequency domain (see Equation 4 in Ref. [85]), which is restricted to match the dedicated-frequency (LF/HF) band.

For real data containing glitches, the *BayesWave* dedicated-frequency follow-up analyses must be applied to both the background (instrumental noise) and foreground (potential GW) triggers. The background analysis provides false-alarm rate measurements for assessing the astrophysical significance of foreground detections. In addition to the

dedicated-frequency follow-ups, *BayesWave* also performs the full-band analysis on eligible events to provide significance estimates independent of cWB.

We demonstrate applications of the LF and HF follow-up analyses in Sections 4.5 and 4.6 respectively.

4.5. Constraining CCSN explosion mechanisms

The primary motivation of the dedicated-frequency framework is to constrain CCSN models for eligible detection candidates, in real observing scenarios. The detection of low-frequency ($\lesssim 0.25\,\mathrm{kHz}$) GW signatures, or lack thereof, can help select between CCSN models and explosion mechanisms. Furthermore, low-frequency GW emissions fall within the most sensitive frequency bands of existing interferometric GW detectors like LIGO and Virgo[6]. In this section, we apply the LF follow-up to simulated GW signals from five distinct CCSN models, featuring different amplitudes of low-frequency GWs induced by the SASI or neutrino-driven convection. We then assess whether CCSN models with prominent low-frequency GW signatures achieve higher detection efficiency with the LF follow-up, compared to those with little to no low-frequency emissions.

In Section 4.5.1, we detail the key features of the five CCSN models used in this study. The CCSN signals are injected into real O3 data. Noise background measurements are required to evaluate the significance of burst triggers. We present the background measurements in Section 4.5.2. In Section 4.5.3, we discuss the properties of the injected CCSN signals and how we ensure that they are detectable up to a nominal significance threshold in O3 data. The analysis results are presented and interpreted in Section 4.5.4.

4.5.1. CCSN models

We use the predicted GW waveforms of five three-dimensional CCSN models. The selected models are a subset of those used to test the LVK search sensitivity for GWs associated with SN 2023ixf [247], one of the closest CCSNe observed in the last decade. For this proof-of-principle study, we choose models that represent typical CCSNe, where the progenitors have solar metallicity and do not rotate [46, 209, 211].

- The SFHx model [222], otherwise known as s15, has a progenitor with zero-age main sequence (ZAMS) mass equal to $15\,M_\odot$ and a soft, i.e. low-pressure, EOS as described in Ref. [248]. The soft EOS results in more vigorous SASI activity, and hence strong low-frequency GW emissions (0.05–0.2 kHz). The high-frequency component of the signal reaches up to ∼1 kHz and is attributed to the PNS surface g-mode oscillations. This simulation is truncated ∼0.35 s after core bounce.

- The s25 model [238] has a progenitor with ZAMS mass equal to $25\,M_\odot$ and assumes the SFHo EOS [248], which differs slightly from SFHx in terms of the mass-radius relationship. The GW signal starts with distinct low-frequency GW emissions at ∼0.1 kHz associated with prompt convection[7]. This is followed by dis-

[6] See representative O3 noise curves in Figure 2 of Ref. [38]

[7] Figure 4 in Ref. [238] shows that the GW energy emitted $\lesssim 0.1$ s after core bounce constitutes $\lesssim 2\%$ of the total GW energy. That is, prompt convection accounts for only a small fraction of the total signal power.

tinct low-frequency emissions (0.05–0.2 kHz) associated with SASI, and high-frequency emissions (0.4–1 kHz) associated with the surface f- and g-modes of the PNS. The simulation is truncated ∼0.6 s after core bounce.

- The D15 model [249] has a progenitor with ZAMS mass equal to $15\,M_\odot$ and is simulated using the D-series CHIMERA code [250]. The GW signal is dominated by high-frequency emissions peaking at ∼1 kHz, largely due to Ledoux convection in the PNS. There are also secondary emissions below ∼0.25 kHz associated with the SASI and neutrino-driven turbulent convection. The explosion occurs ∼0.5 s after core bounce and the simulation is truncated at ∼0.75 s.

- The mesa20_pert model [223] has a progenitor with ZAMS mass equal to $20\,M_\odot$ and assumes the SFHo EOS [248]. Precollapse perturbations are introduced through an aspherical matter velocity field, leading to increased turbulence in the gain region. The GW signal is dominated by high-frequency components (0.3–1.2 kHz) associated with the PNS contraction, and is accompanied by low-frequency components (0.05–0.2 kHz) due to convection and the SASI. The simulation is truncated ∼0.52 s after core bounce.

- The s18 model [251] has a progenitor with ZAMS mass equal to $18\,M_\odot$ and is simulated using the neutrino hydrodynamics COCONUT-FMT code [252]. The GW signal peaks between 0.8–1 kHz due to g-mode PNS oscillations. There is minimal low-frequency emission associated with the SASI. The explosion occurs ∼0.25 s after core bounce and the simulation is truncated at 0.89 s.

The extraction of GW signals from CCSN simulations is computationally expensive, so the simulations of SFHx [222], s25 [238] and mesa20_pert [223] are truncated before the GW signal is fully developed, i.e. they exclude signals from the shock revival phase. Appendix 4.A1 details how GWs are extracted from quadrupole moments in CCSN simulations.

Different SASI signatures are observed for the selected models. The models are listed above in descending order of low-frequency power, i.e. SFHx waveforms generally have the strongest LF GW emissions, followed by s25 and so on. We use the continuous wavelet transform (CWT) [253] to quantify the LF power in each model. To visualize the process, we show the CWT time-frequency spectrograms of arbitrary selected sample waveforms from SFHx (left), s25 (middle) and s18 (right) in Figure 4.2. The white horizontal lines at 256 Hz divide the LF and HF components below and above the lines. The horizontal axis in Figure 4.2 shows the time, t, relative to the squared-strain-weighted central time,

$$t_0 = \frac{1}{h_{\rm rss}^2} \int_{-\infty}^{\infty} dt\, h(t)^2 t. \tag{4.5}$$

The root-sum-squared (rss) strain amplitude of the + and × polarizations,

$$h_{\rm rss}^2 = \int_{-\infty}^{\infty} dt\, [h_+^2(t) + h_\times^2(t)], \tag{4.6}$$

is proportional to the total energy of the GW signal, and serves as a normalization factor in Equation 4.5. The orange vertical lines in Figure 4.2 indicates $t = t_0$ of each waveform.

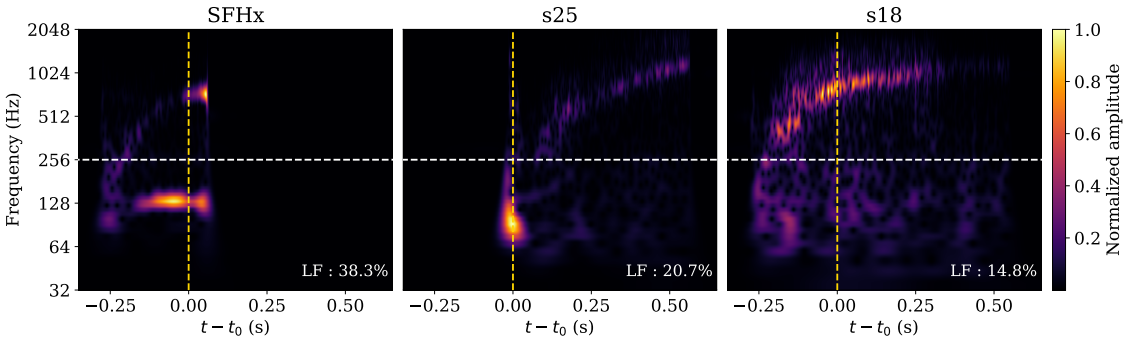

Figure 4.2: CWT time-frequency spectrograms of a sample GW signal for SFHx (left), s25 (middle) and s18 (right). The vertical axis shows the full-band analysis frequency range (32–2048 Hz) and the white horizontal lines at 256 Hz indicate the boundary between LF and HF components, as per the dedicated-frequency framework. The horizontal axis shows the time t relative to the central time t_0 of the signal, and the orange vertical lines indicate $t = t_0$. The horizontal scales are the same for all three plots, and the approximate duration of the SFHx, s25 and s18 signals are 350 ms, 625 ms and 900 ms respectively. The color bar represents the linearly scaled amplitude of the signals; the minimum and maximum amplitudes in each panel correspond to the values 0 and 1 respectively. The LF energy divided by the total signal energy is quoted in the bottom right corner of each plot.

By the definition in Equation 4.5, t_0 is the time at which the majority of the signal energy is concentrated. Visual inspection of Figure 4.2 reveals that the LF features of SFHx (left) and s25 (middle) are close to t_0, suggesting LF emissions for SFHx and s25 contribute considerably to the total signal power. For s18, on the other hand, t_0 roughly aligns with the peak emission at ∼1 kHz, suggesting HF emissions have a stronger influence. To quantify the extent of LF emissions, we compute the signal energy for frequency $f \leq 256$ Hz (i.e. below the white horizontal lines) divided by the total signal energy in the range $32\,\text{Hz} \leq f \leq 2048\,\text{Hz}$. For the particular waveforms plotted in Figure 4.2 for the models SFHx, s25 and s18, the LF contributions are 38.3%, 20.7% and 14.8% respectively. We repeat this calculation for 175 randomly realized waveforms per CCSN model. We find that, on average, the LF emissions contribute 36.6%, 19.4%, 18.4%, 16.2% and 14.9% to the overall signal energy of the SFHx, s25, D15, mesa20_pert and s18 models respectively.

4.5.2. Background measurements

To demonstrate the application of LF follow-ups in realistic observing scenarios, we inject GW signals from the five CCSN models into O3 data. Real detector data are susceptible to glitches and are therefore capable of producing false alarm triggers. In order to assess the significance of triggers produced by an analysis pipeline, one has to empirically measure the noise background, i.e. the rate of false alarms produced by the corresponding pipeline in the absence of GW signals. Trigger sensitivity varies unpredictably across different analysis pipelines. To address this, cWB first measures the background for all of O3, and *BayesWave* independently follows up on the astrophysically-relevant cWB triggers to evaluate its own background. The independent background measure-

ments with cWB and the *BayesWave* follow-up account for the individual strengths and shortcomings of each algorithm, thereby improving the reliability of their respective significance estimates. The background measurements are used to evaluate the significance of detection candidates in terms of the false alarm rate (FAR), which in turn is used to assess the eligibility of detection candidates for the dedicated-frequency follow-up.

We use the standard time-shift analysis to conduct the background measurements [196]. That is, we produce artificially extended detector background data by introducing temporal offsets, which are long enough to nullify any meaningful correlations between the outputs of two or more detectors. We choose to use the data from a two-detector configuration, comprising the LIGO Hanford (H) and LIGO Livingston (L) detectors. Previous works have shown that a three-detector configuration, comprising HL plus Virgo, does not outperform the HL-only network in terms of detection efficiency in O3 [4, 46, 203]. Altogether we accumulate 605 years of HL background by applying $\sim 2.4 \times 10^4$ time-shifts on 9.5 days of O3a data[8].

cWB background

The cWB background is measured by analyzing the entire time-shifted data set. The same XGBoost model is used for both the background and foreground cWB analyses. In this study, the model is trained using: (i) a randomly selected subset (70%) of the background noise events and (ii) four sets of WNB signals with central frequency, bandwidth and duration uniformly sampled from overlapping subsets within the ranges $[24, 2048]$ Hz, $[10, 800]$ Hz and $[0.1, 500]$ ms respectively. The parameters of the WNB training sets are chosen based on the expected time-frequency volume of GW bursts (see Appendix A of Ref. [203] for more details). The XGBoost training parameters for this study are similar to those in Ref. [247].

cWB does not perform the dedicated-frequency follow-ups; it is only used to identify eligible candidates via the full-band analysis. Therefore, only the full-band background measurement is necessary. This measurement is conducted using the remaining 30% (182 years) of data not used for the XGBoost model training, and is shown in the top panel of Figure 4.3. The FAR is calculated as the number of background cWB triggers exceeding the corresponding η_r, divided by the total background livetime of 182 years. The green horizontal line at FAR $= 1\,\mathrm{yr}^{-1}$ indicates the nominal detection threshold, below which a cWB trigger qualifies as a detection candidate in this study, corresponding to $\eta_r = 0.78$. That is, a cWB trigger must satisfy $\eta_r \geq 0.78$ to qualify as a LF follow-up candidate. The reason for choosing the FAR $= 1\,\mathrm{yr}^{-1}$ threshold is as follows. In the O3 all-sky burst search [46], an event is considered a significant detection for FAR $\leq 0.01\,\mathrm{yr}^{-1}$. However, Ref. [46] shows that typical non-rotating CCSNe with solar-metallicity, i.e. the models used in this study, are undetectable at such low FARs with O3 detector sensitivities. Therefore, to showcase the applications of the dedicated LF follow-up for anticipated CCSN detections with existing detector configurations, we arbitrarily increase the significance tolerance to FAR $\leq 1\,\mathrm{yr}^{-1}$.

[8] We use O3a data because its overall glitch rate is lower than in O3b [38]. More specifically, we leverage the higher burst detection sensitivity in O3a to produce more detection candidates eligible for dedicated-frequency follow-ups.

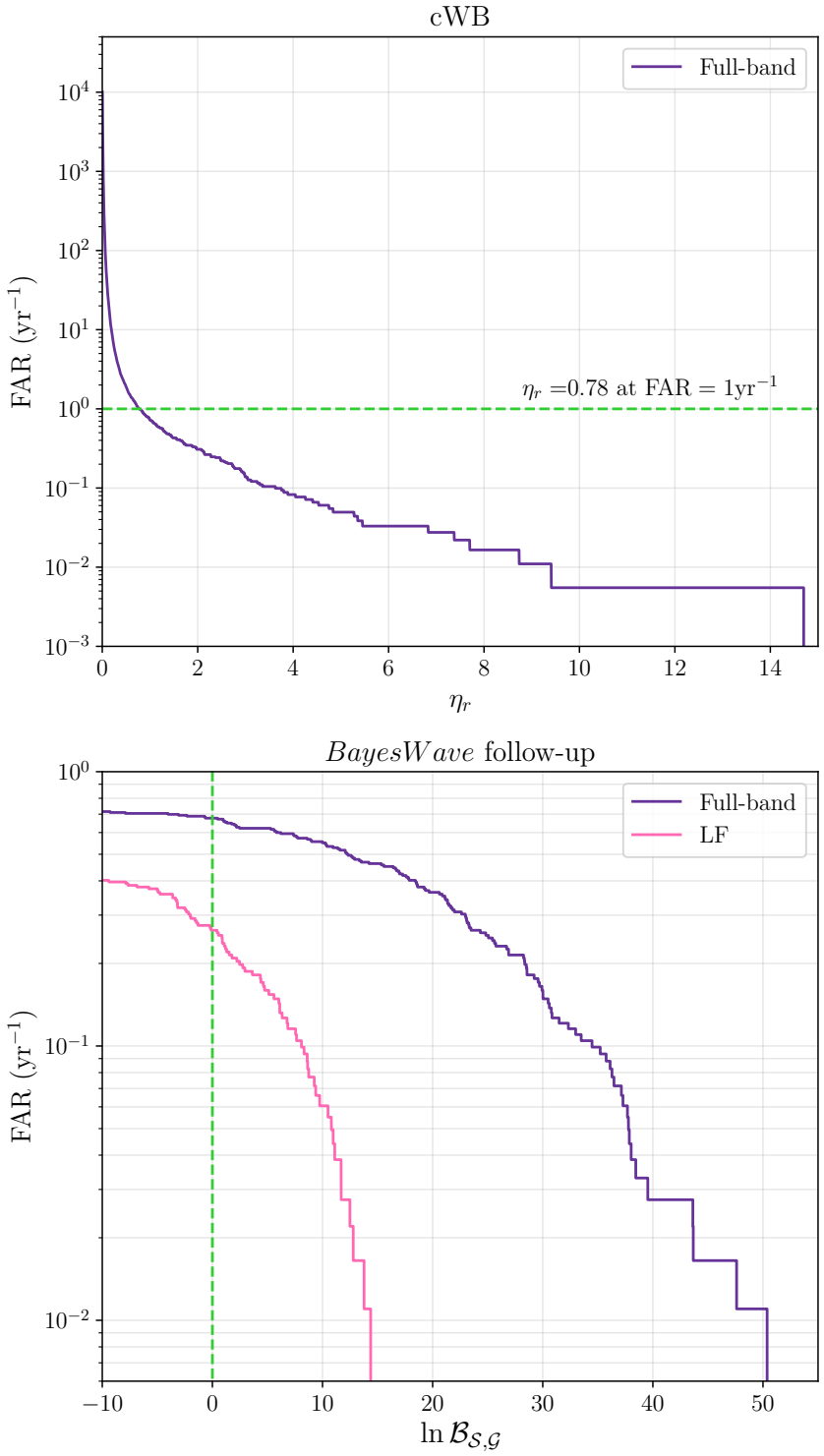

Figure 4.3: O3a background measurements. The top panel shows the FAR of the background triggers as a function of the detection statistic η_{r}, for the full-band cWB analysis. The bottom panel shows the same but for the *BayesWave* full-band (purple curve) and LF (pink curve) triggers, as a function of the detection statistic $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. The horizontal green line at FAR = 1 yr^{-1} (top panel) and the vertical green line at $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = 0$ (bottom panel) indicate the detection thresholds for cWB and *BayesWave* respectively.

BayesWave background

Eligible cWB triggers are followed up using the full-band and LF *BayesWave* analyses. The full-band analysis checks that the cWB trigger is also detected by *BayesWave*, while the LF analysis checks for low-frequency GW signatures, or the lack thereof. Therefore, we must conduct two separate background measurements for *BayesWave*, one using the full-band analysis and the other using LF. *BayesWave* is computationally expensive and therefore it is impractical to follow-up the entire time-shifted data (cf. cWB). Instead, *BayesWave*'s background is measured by following-up cWB background triggers above a nominal significance threshold. This threshold is the same as the dedicated-frequency follow-up criterion, FAR $= 1\,\mathrm{yr}^{-1}$, discussed above. Since triggers with FAR $> 1\,\mathrm{yr}^{-1}$ are not valid candidates for the LF follow-up, it is unnecessary to quantify their detection significance; they can be excluded from *BayesWave*'s background measurements, to conserve computational resources.

The *BayesWave* analyses are configured as follows. The analysis segment spans 4 seconds, centered on the event epoch, to provide a sufficiently long data segment for accurate noise PSD estimation. However, burst signals are typically shorter than a second. Therefore, the time window in which *BayesWave* is allowed to place signal and glitch wavelets is limited to 1 second, also centered on the event epoch. The sampling rate is twice the maximum frequency of the analysis band, e.g. the maximum frequency of the full-band analysis is 2048 Hz, so the sampling rate is 4096 Hz. This configuration applies to all *BayesWave* analyses in this paper, and is similar to that used in Ref. [149] for CCSNe analysis.

The bottom panel of Figure 4.3 shows the backgrounds measured by *BayesWave*; the purple and pink curves show the full-band and LF measurements respectively. The *BayesWave* FAR is calculated the same as with cWB, except that η_r is replaced by $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. An event must be more consistent with the signal model than the glitch model to qualify as an astrophysically-relevant *BayesWave* trigger, i.e. it must satisfy $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} > 0$. For reference, the green vertical line in the bottom panel of Figure 4.3 indicates where $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = 0$. In the full-band and LF background, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = 0$ correspond to FAR $= 0.68\,\mathrm{yr}^{-1}$ and $0.26\,\mathrm{yr}^{-1}$ respectively. It is expected that FAR $< 1\,\mathrm{yr}^{-1}$ for $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} > 0$, because only a subset of cWB triggers with FAR $\leq 1\,\mathrm{yr}^{-1}$ also qualify as *BayesWave* triggers. Henceforth we focus our discussion on the background measurements for $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} > 0$, as they represent the distribution of false alarms with astrophysically relevant detection statistics.

The LF *BayesWave* background (pink curve) is approximately an order of magnitude lower than the full-band background (purple curve) on average. There are two reasons for this. First, $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ scales with the SNR, as well as with the number of wavelets N, which reflects the model complexity. In the LF analysis, the HF contributions of the background triggers are disregarded, which reduces the trigger SNRs and the number of wavelets required for their *BayesWave* reconstruction; these reductions collectively result in lower $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. Second, for events with minimal LF power, disregarding the HF contributions can cause the data to become more consistent with the Gaussian noise model ($\mathcal{N}$) than with the non-Gaussian signal model ($\mathcal{S}$). That is, the log Bayes factor between the signal and Gaussian noise model ($\ln \mathcal{B}_{\mathcal{S},\mathcal{N}}$) includes values below zero

within its error bars $\Delta \ln \mathcal{B}_{\mathcal{S},\mathcal{N}}$, viz.

$$\ln \mathcal{B}_{\mathcal{S},\mathcal{N}} \pm \Delta \ln \mathcal{B}_{\mathcal{S},\mathcal{N}} \leq 0. \tag{4.7}$$

Events that satisfy Equation 4.7 are not astrophysically relevant and are therefore excluded from the FAR calculation, regardless of their $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. With this criterion in place, 17% of the cWB background triggers are excluded from the *BayesWave* full-band background measurements, cf. 49% for the *BayesWave* LF measurements. In other words, the LF analysis significantly reduces the noise background, which further explains why the LF background (pink curve) is lower compared to the full-band background (purple curve).

Overall, the background measurements show that a LF-analysis *BayesWave* trigger can achieve the same FAR as a full-band analysis trigger, with a lower $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$.

4.5.3. CCSN injection properties

As noted previously, only triggers with FAR $\leq 1\,\mathrm{yr}^{-1}$ qualify for the *BayesWave* LF follow-ups. Therefore we must ensure that the injected CCSNe also satisfy FAR $\leq 1\,\mathrm{yr}^{-1}$ in the cWB full-band analysis, to qualify for the LF follow-up study. Here, we discuss how to use cWB to compute the appropriate signal amplitudes for the CCSN injections.

The detectability of GW bursts at a given FAR is typically quantified by their detection efficiency, as a function of h_{rss}, defined in Equation 4.6. The detection efficiencies are evaluated empirically by injecting the same set of signals into detector noise at different h_{rss}, and then calculating the fraction of signals that are recovered by the full-band cWB analysis with FAR $\leq 1\,\mathrm{yr}^{-1}$. In the LF analysis, the CCSN signals are injected at an amplitude $h_{\mathrm{rss},50}$, corresponding to 50% detection efficiency. This choice aligns with the benchmark sensitivity used in standard burst searches and ensures that the CCSN signals are both detectable and eligible for LF follow-up. The model-specific $h_{\mathrm{rss},50}$ is evaluated as follows. First, we choose eight approximately uniform h_{rss} values from the range $5 \times 10^{-23}\,\mathrm{Hz}^{-1/2} \leq h_{\mathrm{rss}} \leq 4 \times 10^{-21}\,\mathrm{Hz}^{-1/2}$. Then, for each CCSN model, we inject $\sim$500 signals per h_{rss} value[9], and compute the corresponding detection efficiency. The detection efficiencies as a function of h_{rss} are plotted as discrete data points in Figure 4.4. The different colors indicate different CCSN models. The $h_{\mathrm{rss},50}$ for each CCSN model is obtained by least-squares fitting a cumulative log-normal distribution function of h_{rss}. Figure 4.4 shows the cumulative log-normal fits as dashed curves in colors corresponding to the data points they are fitting. The $h_{\mathrm{rss},50}$ value for each model is enclosed within the parentheses in the legend. These are the h_{rss} values at which we inject the waveforms for the LF follow-up study.

For an isotropically emitting source, the source distance r scales inversely with the h_{rss} as [254] (see derivation in Appendix 4.A1):

$$r^2 = \frac{G E_{\mathrm{GW}}}{\pi^2 c^3 h_{\mathrm{rss}}^2 f_0^2}. \tag{4.8}$$

[9]The number of injections varies across the CCSN models, ranging from 526 to 585 injections, depending on the number of available waveforms for each model. However, for each CCSN model, the number of injections per h_{rss} value is the same.

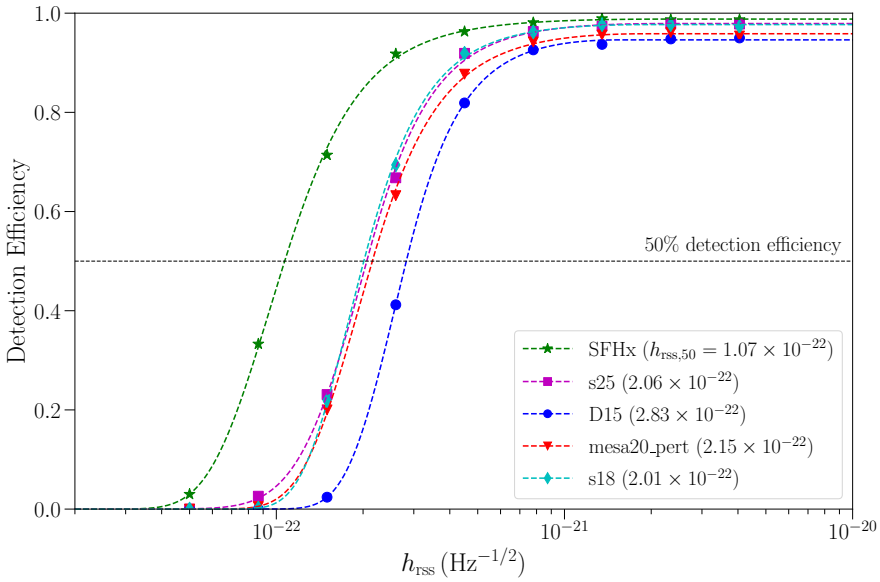

Figure 4.4: Detection efficiency for events with FAR $\leq 1\,\mathrm{yr}^{-1}$ versus signal amplitude $h_{\rm rss}$. Each point represents the empirically measured detection efficiency with ~500 injections. The colored points represent different CCSN models as indicated by the legend, and the dashed curves in corresponding colors show the least-square log-normal fit. The numbers in parentheses are $h_{\rm rss,50}$ (in units of $\mathrm{Hz}^{-1/2}$), i.e. the $h_{\rm rss}$ value that results in 50% detection efficiency, as indicated by the horizontal dashed line.

In Equation 4.8, G and c respectively denote the gravitational constant and the speed of light, and $E_{\rm GW}$ is the fixed total energy emitted by the GW source. The central frequency f_0 is defined as

$$f_0 = \frac{2}{h_{\rm rss}^2} \int_0^\infty df \left|\tilde{h}(f)\right|^2 f, \tag{4.9}$$

where $\tilde{h}(f) = \tilde{h}(-f)$ denotes the Fourier transform of the real-valued time-domain amplitude $h(t)$. Table 4.1 lists the $E_{\rm GW}$ and the peak frequency $f_{\rm peak}$ of the GW signal for the five CCSN models. Assuming that $f_{\rm peak}$ roughly approximates f_0, one can use Equation 4.8 to estimate the source distance r_{50} corresponding to $h_{\rm rss,50}$ for each CCSN model [203]. The $h_{\rm rss,50}$ from Figure 4.4 and the corresponding r_{50} are also listed in Table 4.1. Using 50% detection efficiency at FAR $\leq 1\,\mathrm{yr}^{-1}$ as the detection benchmark, we find that all five CCSN models in this study are detectable at distances no greater than that of the Galactic Center (8.5 kpc). Among the five models, SFHx has the highest detectability, reaching distances up to 7.8 kpc; mesa20_pert has the lowest detectability, reaching only 0.9 kpc.

The source distance estimates in Table 4.1 are broadly consistent with those from the O3 targeted CCSN search, as reported in Table III of Ref. [160]. Two possible reasons for the discrepancy in the distance estimates are: (i) our analysis adopts a stricter detection benchmark of FAR $\leq 1\,\mathrm{yr}^{-1}$, whereas Ref. [160] tolerates higher FARs corresponding to the loudest event associated with each CCSN; and (ii) our estimates are derived from an all-sky search, whereas Ref. [160] is based on targeted searches. Since Ref. [160] uses a less stringent detection benchmark, one would expect its r_{50} estimates to exceed those

Model	$E_{\rm GW}$ $[M_\odot c^2]$	$f_{\rm peak}$ [Hz]	$h_{\rm rss,50}$ [$\rm Hz^{-1/2}$]	r_{50} [kpc]
SFHx	1.1×10^{-9}	267	1.07×10^{-22}	7.8
s25	2.7×10^{-8}	1132	2.06×10^{-22}	4.9
D15	8.9×10^{-9}	1102	2.83×10^{-22}	2.1
mesa20_pert	9.4×10^{-10}	1103	2.15×10^{-22}	0.9
s18	1.6×10^{-8}	818	2.01×10^{-22}	5.3

Table 4.1: Estimating the source distance r_{50} corresponding to 50% detection efficiency. The columns from left to right list: (i) the CCSN models, (ii) the fixed total energy emitted by the source $E_{\rm GW}$, (iii) the peak frequency $f_{\rm peak}$ of the GW signal, (iv) the $h_{\rm rss,50}$ derived from Figure 4.4, and (v) the corresponding r_{50} calculated using Equation 4.8, assuming that $f_0 \approx f_{\rm peak}$.

in Table 4.1. This, however, is not always true; specifically, the r_{50} for s25 and s18 in Table 4.1 are further than all corresponding estimates in Ref. [160]. This comparison suggests that targeted searches do not necessarily provide better sensitivity than all-sky searches, justifying our use of an all-sky search to minimize bias in our results toward any particular sky location.

4.5.4. Results and discussion

We inject approximately 350 waveforms per CCSN model at the corresponding $h_{\rm rss,50}$ (as per Figure 4.4), into the same 9.5-day segment of O3a data that produced the background measurements in Figure 4.3. We perform the full-band cWB analysis on the injections to identify eligible candidates with FAR $\leq 1\,\rm yr^{-1}$ for the LF follow-up analysis. The lowest fraction of candidates (71% of 384 injections) is obtained for model s25, while the highest (80% of 378 injections) is for mesa20_pert. For each model, we arbitrarily select 175 of the eligible candidates, to ensure that all CCSN models have equal-sized datasets for a fair comparison.

The selected cWB candidates are followed-up using the *BayesWave* LF and full-band analyses. For each CCSN model and analysis, we calculate the detection efficiency, defined as the number of candidates recovered with $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \geq 0$ by *BayesWave*, divided by the total number of candidates (175). The results are plotted in Figure 4.5. As per the dedicated-frequency follow-up criteria, the detection efficiency for the cWB full-band analysis must be unity. Therefore, it is unsurprising that the *BayesWave* full-band analysis, i.e. the purple crosses in Figure 4.5, yields the same results. However, the detection efficiencies of the *BayesWave* LF follow-up analysis are non-trivial. We find that the LF detection efficiency decreases from left to right across the plot, i.e. as the LF power in the CCSN model decreases. We note, in particular, that the s18 model, which exhibit minimal SASI and has the lowest LF power (14.9% of the total power), has the lowest detection efficiency at 0.10. In contrast, the detection efficiency for SFHx, with 36.6% LF power and strong SASI emission, amounts to 0.97. That is, the detectability of SFHx with the LF follow-up is highly likely. However, for s25, D15 and mesa20_pert, which have

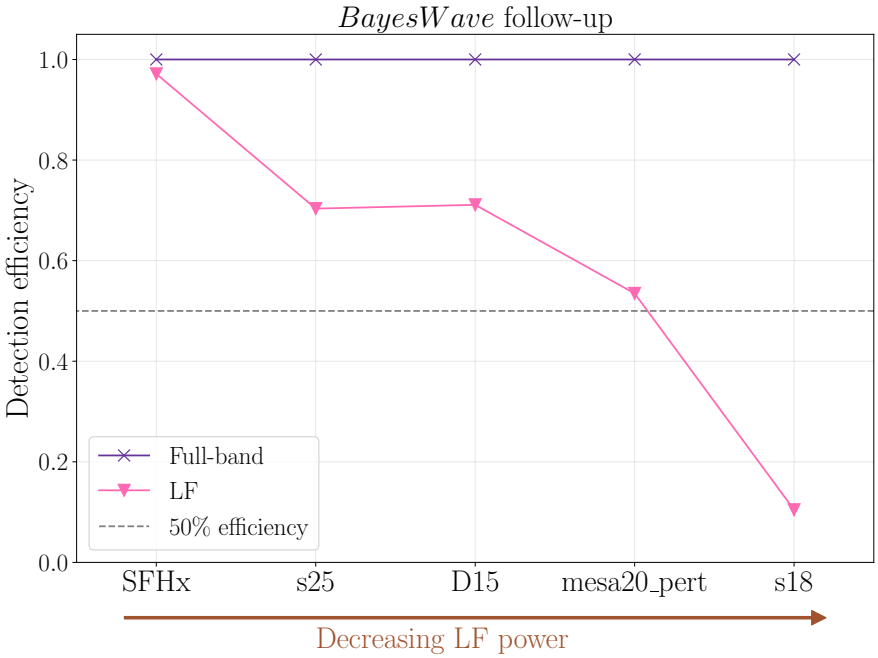

Figure 4.5: Detection efficiency for events with FAR $\leq 1\,\mathrm{yr}^{-1}$ for the five CCSN models, based on the *BayesWave* follow-up analyses. The pink triangles and purple crosses show the efficiencies for the LF and full-band analyses respectively. The CCSN models are displayed on the horizontal axis in order of decreasing LF power, spaced equally from left to right. The gray dashed line indicates 50% detection efficiency, for reference.

~10% less LF power than SFHx but still exhibit SASI emissions, the detection efficiency ranges from 0.5 to 0.75. That is, CCSN models with moderate SASI-related LF emission have a fair chance of being overlooked in LF follow-up analyses.

Detection efficiency can only be evaluated for injected signals, where both the number of injected and recovered signals are known. For detections in real data, the true population of signals is unknown, so the fraction of detected and missed signals cannot be determined. For this reason, detectability is typically quantified by the FAR. Accordingly, we present a FAR-based analysis of the CCSN injections to demonstrate how LF analysis results are interpreted for detections in real data. In Figure 4.6, the crosses (triangles) indicate the median FAR of the 175 injections per model, recovered by the *BayesWave* full-band (LF) analysis. The error bars bracket the interquartile ranges (IQRs). CCSN injections that satisfy Equation 4.7 are not successfully detected by *BayesWave*. However, these events cannot be excluded from the FAR analysis, as it would skew the IQRs and medians. Therefore, in Figure 4.6, the non-detections are assigned the maximum FAR=$1 \times 10^4\,\mathrm{yr}^{-1}$ derived from the cWB background, to reflect minimal astrophysical significance. The black (orange) horizontal dashed line represents the FAR threshold, below which an event is considered a successful detection for the full-band (LF) analysis; the FAR threshold corresponds to $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = 0$. For all five CCSN models, the median FARs (cross symbols) and the IQRs of the full-band analysis fall below the horizontal black dashed line. That is, $\geq 50\%$ of the waveforms per CCSN model satisfy the full-band detection threshold. This observation is consistent with the unit detection efficiency for the full-band analysis in Figure 4.5. The LF-analysis median FAR (triangle symbols), on

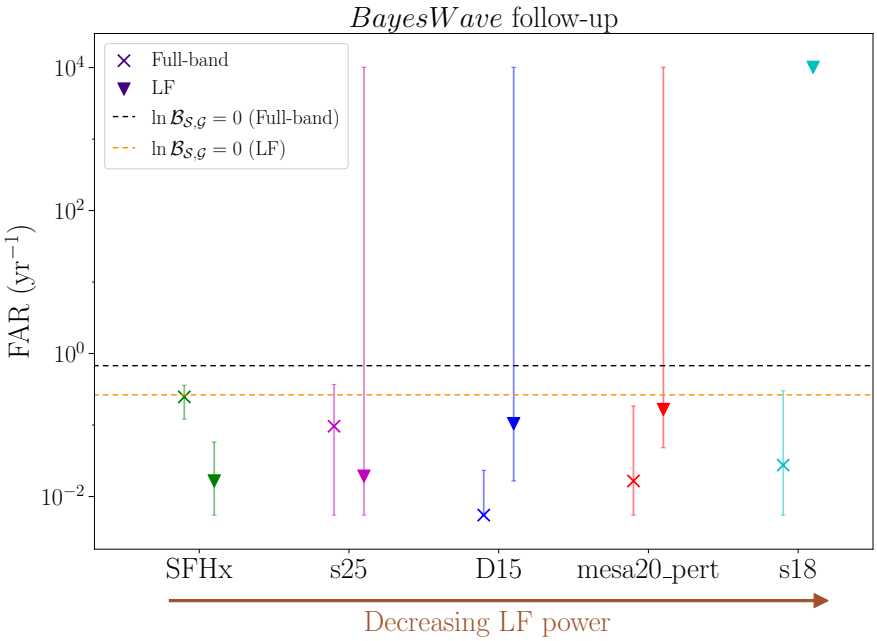

Figure 4.6: False alarm rate (FAR) of the CCSN injections produced by the *BayesWave* full-band and LF analyses. The crosses (triangles) show the median FAR of the 175 injections for each CCSN model, obtained using the full-band (LF) analysis. The error bars indicate the interquartile range (IQR) of the FARs. Non-detections, as defined by Equation 4.7, are assigned the maximum FAR ($1 \times 10^4\,\mathrm{yr}^{-1}$) of the cWB background. The LF error bars for the s18 model are invisible because all injections within the IQR are non-detections. The black (orange) horizontal dashed line at FAR $= 0.68\,\mathrm{yr}^{-1}$ ($0.26\,\mathrm{yr}^{-1}$) indicate the *BayesWave* detection threshold for the full-band (LF) analysis, below which an event is considered a detection. These thresholds correspond to $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = 0$ and are derived from the bottom panel of Figure 4.3. As with Figure 4.5, the CCSN models are displayed on the horizontal axis in order of decreasing LF power, equally spaced from left to right.

the other hand, increases from left to right. This suggests that CCSN models with less LF power are generally recovered with higher FAR by the LF analyses, leading to a lower corresponding detection efficiency, as noted in Figure 4.5. Overall, Figures 4.5 and 4.6 have similar implications for the LF follow-up performance, but Figure 4.6 presents the results in terms of a measurable quantity for detections in real data.

In summary, we find that the *BayesWave* LF follow-up analyses detect CCSN models with higher LF content at lower FAR. On the other hand, the LF follow-up may miss CCSNe with moderate LF content, so the absence of LF detection does not necessarily indicate the lack of LF content. Therefore, we conclude that a successful LF follow-up detection is useful for constraining the explosion model of a CCSN candidate, but the lack of a detection is inconclusive.

4.6. High-frequency analysis: loudest trigger during SN 2019fcn

In this section, we demonstrate another application of the dedicated-frequency framework: to follow-up a potential CCSN GW burst candidate with the HF analysis. We note in advance that the following illustrative study on the loudest trigger during SN 2019fcn does not yield astrophysical results. The aim, rather, is to demonstrate by way of a concrete example how to apply the HF analysis to astrophysically significant candidates produced by future GW burst searches.

4.6.1. Selection of a candidate

Ref. [160] presents cWB search results for GW bursts associated with optically observed CCSNe during O3. The CCSNe are selected based on (i) their optically inferred distances and (ii) whether there is sufficient GW data to generate a few years of time-shifted background for significance assessment. For each CCSN event, the cWB search is conducted over a methodically selected time interval, otherwise known as the on-source window (OSW), which is expected to contain the GW signal up to a specified probability (see Refs. [159] and [160] for details). Although no detections were reported from the searches, the OSW (i.e. non-time-shifted) analysis for each CCSN using cWB produces a list of triggers. The trigger with the lowest FAR is called the 'loudest trigger' and is considered the most plausible GW candidate for the associated CCSN. It is therefore interesting to apply the dedicated-frequency follow-up to one of the loudest triggers, to study whether the follow-ups can improve the detection significance in practical observing scenarios.

In order to choose an illustrative follow-up candidate, we follow up all of the loudest cWB triggers in Ref. [160] using the full-band *BayesWave* analysis. Valid candidates are those with positive $\ln \mathcal{B}_{\mathcal{S},\mathcal{N}}$ and $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. Among four valid candidates, the loudest trigger during SN 2019fcn has the highest $\ln \mathcal{B}_{\mathcal{S},\mathcal{N}}$ (14.7) and the second highest $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ (7.5). It also has the lowest cWB FAR ($22\ \mathrm{yr}^{-1}$), according to Table II in Ref. [160]; the FAR is evaluated using a background measured with time-shifted data from the 4.54-day OSW of SN 2019fcn. Altogether, this trigger is the best available O3 burst candidate for a dedicated-frequency follow-up, according to cWB and *BayesWave*. The cWB and *BayesWave* full-band analysis reconstructions of SN 2019fcn's loudest trigger, displayed in Figure 4.7, show that the signal predominantly comprises high-frequency power with central frequency $f_0 \sim 1\,\mathrm{kHz}$. With minimal low-frequency emission, this trigger is well-suited for the HF follow-up, which focuses on the band where the signal is expected and suppresses noise from irrelevant frequencies to enhance overall detectability. We reiterate that the trigger is not a realistic candidate due to its high ($\geq 1\ \mathrm{yr}^{-1}$) FAR. Therefore the results of the HF follow-up do not have any astrophysical implications. The aim is simply to demonstrate the methodology.

Figure 4.7 shows that the cWB reconstruction picks up signal power at $\sim$2 kHz which is absent from the *BayesWave* reconstruction. In contrast, the *BayesWave* reconstruction picks up signal power at $\sim$0.5 kHz, which cWB does not. This observation suggests that cWB and *BayesWave* are sensitive to different parts of the spectrum, but with only one event reconstruction, we are unable to reach a definitive conclusion. Analyzing more

events presents an interesting avenue for future studies.

4.6.2. Background measurements

The cWB background of SN 2019fcn is measured in Ref. [160]. Unlike in Section 4.5, the background measurement and analysis of SN 2019fcn in Ref. [160] is conducted without XGBoost and uses an older version of cWB, because the updated methods were not yet available. We refer the reader to Section IIIB of Ref. [160] for the full description of the ranking statistic η_c used in this section.

We aim primarily to follow-up the SN 2019fcn trigger using the *BayesWave* HF analysis, but for completeness we also follow-up using the *BayesWave* full-band analysis. Hence we measure two backgrounds. As before, the *BayesWave* background measurement follows up cWB background triggers above a nominal threshold. Here, the follow-up threshold for the full and HF *BayesWave* background measurements is $\eta_c = 6.7$, i.e. the value at which the cWB full-band analysis recovers the loudest trigger during SN 2019fcn, as reported in Ref. [160].

The purple and blue curves in Figure 4.8 show the background measured by *BayesWave* using the full-band and HF analyses respectively. Although the HF background is lower than the full-band background, the discrepancy is smaller than that observed between the LF and full-band backgrounds in Figure 4.3. This makes sense because the frequency range of the HF analysis (256–2048 Hz) overlaps more with the full-band analysis (32–2048 Hz) than the LF analysis (32–256 Hz). That is, the SNR of the HF triggers should be similar to the full-band analysis triggers, and the number of wavelets N used for reconstruction should likewise be similar, as should $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. The discrepancy could also partially be due to the different cWB versions used to measure the backgrounds for SN 2019fcn and Figure 4.3.

We use the backgrounds in Figure 4.8 to evaluate in Section 4.6.3 the *BayesWave* full-band and HF FARs for the loudest trigger during SN 2019fcn.

4.6.3. FAR analysis

Table 4.2 summarizes the detection statistics, η_c and $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$, recovered by the cWB and *BayesWave* analyses respectively, for the loudest trigger during SN 2019fcn. The cWB full-band analysis results are taken directly from Ref. [160]. The *BayesWave* FARs are evaluated by comparing the full-band and HF $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ values to the corresponding background measurements in Figure 4.8. First, we note that the *BayesWave* full-band analysis achieves a lower FAR ($6.4\,\mathrm{yr}^{-1}$) than cWB ($22.1\,\mathrm{yr}^{-1}$). This demonstrates how the hierarchical pipeline enhances detection significance, as previously reported in Ref. [139]. Second, the HF follow-up further reduces the FAR to $4.9\,\mathrm{yr}^{-1}$. This is because the HF follow-up yields higher $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = 8.6$, cf. $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = 7.5$ in the full-band analysis. According to Figure 4.8, higher $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ results in lower FAR. The increase in $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ with the HF follow-up is justified as follows. Since the SN 2019fcn trigger has minimal power below 256 Hz, the HF and full-band signal reconstructions are similar, i.e. they use approximately the same number of wavelets (median $N = 9$). The signal evidence depends on the volume of the parameter space, V and the subvolume that the model occupies, ΔV [85]; a higher signal evidence is achieved when $\Delta V/V$ is larger. If the HF signal

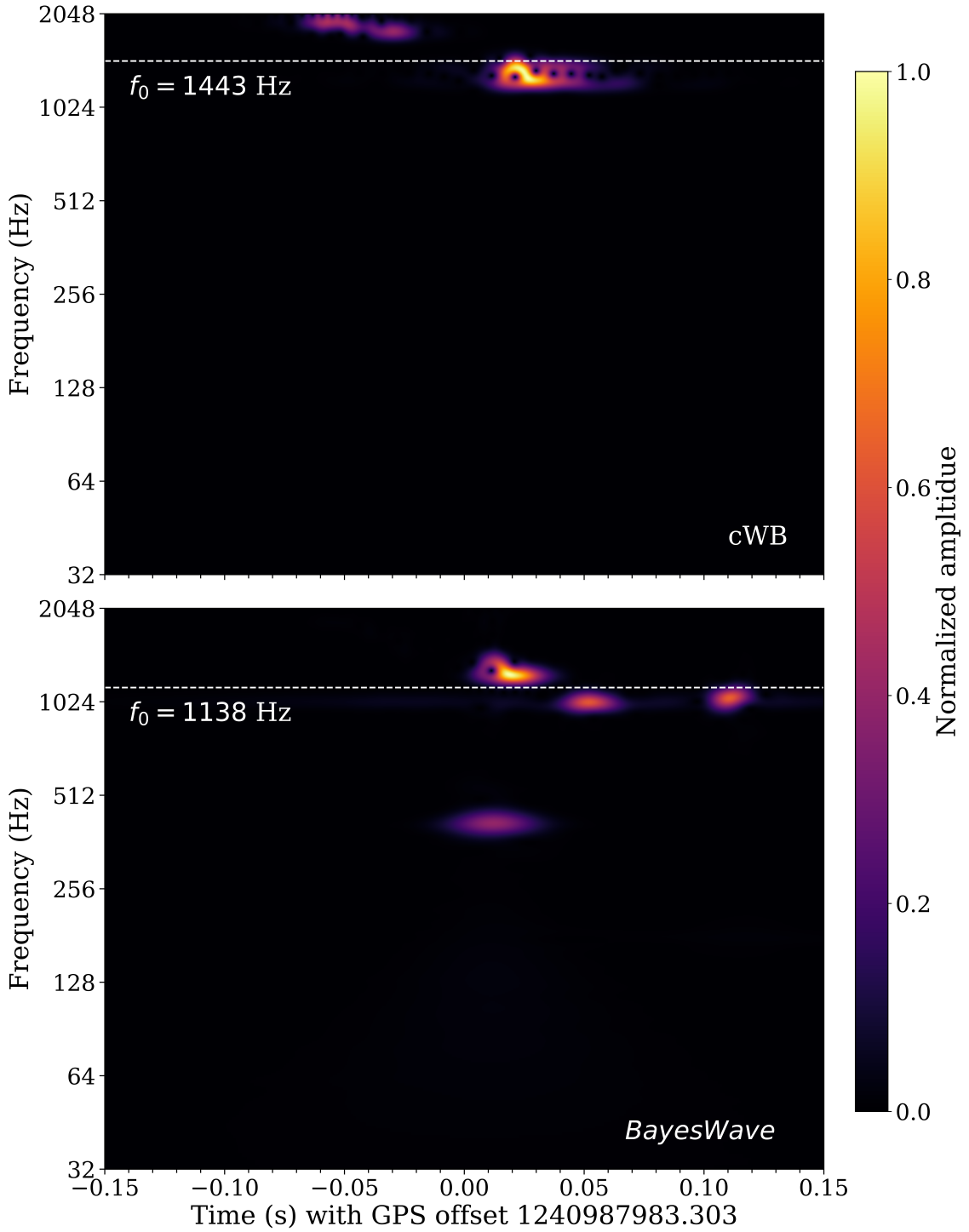

Figure 4.7: Whitened CWT time-frequency spectrogram of the loudest trigger during SN 2019fcn. The top panel shows the cWB full-band reconstruction. The horizontal axis indicates the time relative to the central GPS time of the trigger. The color bars show the normalized amplitude, defined the same way as in Figure 4.2. The horizontal line indicates $f_0 = 1443\,\mathrm{Hz}$. The bottom panel shows the same as the top panel but for *BayesWave*, with $f_0 = 1138\,\mathrm{Hz}$.

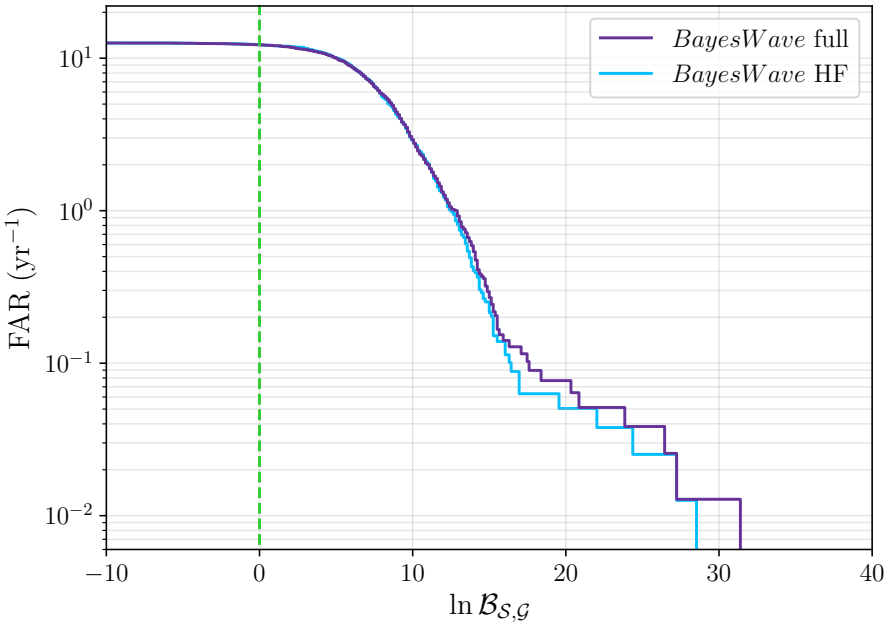

Figure 4.8: SN 2019fcn background measured using the *BayesWave* full-band (purple curve) and HF (blue curve) follow-up analyses. The vertical green line at $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} = 0$ indicates the *BayesWave* detection threshold.

reconstruction is similar to the full-band reconstruction, then $\Delta V_{\text{full}} \approx \Delta V_{\text{HF}}$. However, the HF analysis is restricted to a narrower bandwidth than the full-band analysis, so it follows that $V_{\text{HF}} < V_{\text{full}}$. As a result, the HF analysis yields higher signal evidence, and consequently higher $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$.

To summarize, Table 4.2 demonstrates that (i) the successive application of cWB and *BayesWave* reduces the FAR by a factor of 3.4, and (ii) the HF *BayesWave* follow-up, limited to $256\,\text{Hz} \leq f \leq 2048\,\text{Hz}$, further reduces the FAR for the loudest trigger during SN 2019fcn by a factor of 1.3.

We reiterate that the results in Table 4.2 have no astrophysical significance, because the loudest trigger during SN 2019fcn is not a real GW burst candidate. The goal in this paper is to demonstrate concretely how to apply HF follow-up to one representative event. This method is suitable for following up future detection candidates with marginal significance (e.g. FAR$\sim 1\,\text{yr}^{-1}$) and minimal low-frequency ($f \leq 256\,\text{Hz}$) power. One can also adapt the method by adjusting the LF-HF boundary to match other emission mechanisms.

4.7. Conclusions

GW signals from CCSNe contain spectral signatures which reflect the physical mechanisms that occur within the progenitor star immediately prior to its explosion. In particular, low-frequency gravitational wave signatures ($f \lesssim 250\,\text{Hz}$), can be used to detect hydrodynamical instabilities, such as the SASI and neutrino-driven convection, which are thought to play a crucial role in driving CCSN explosions. In this paper, we introduce the dedicated-frequency framework, designed to characterize spectral signatures

	Full-band			HF		
Pipeline	η_c	$\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$	FAR (yr^{-1})	η_c	$\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$	FAR (yr^{-1})
cWB	6.7	-	22.1	-	-	-
BayesWave	-	7.5	6.4	-	8.6	4.9

Table 4.2: Follow-up analysis output for the loudest trigger during SN 2019fcn. η_c and $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ are the detection statistics produced by cWB and *BayesWave* respectively. The cWB and *BayesWave* FARs are estimated independently using the corresponding backgrounds of the full-band and HF analyses; see Section 4.6.2.

of GW burst detection candidates. The framework uses a hierarchical detection pipeline comprising two minimally-modeled burst analysis algorithms: cWB, to identify eligible candidates based on their astrophysical significance, and *BayesWave*, to follow-up eligible burst candidates using bandpass analyses. We demonstrate two applications of the dedicated-frequency framework: (i) to detect GW signatures associated with the SASI and neutrino-driven convection in CCSNe using a low-frequency (LF) analysis, limited to the range $32\,\mathrm{Hz} \leq f \leq 256\,\mathrm{Hz}$; and (ii) to enhance detectability of the loudest trigger from SN 2019fcn using a high-frequency (HF) analysis, limited to the range $256\,\mathrm{Hz} \leq f \leq 2048\,\mathrm{Hz}$.

The LF study uses GW waveforms from five different CCSN models with typical (non-rotating, with solar metallicity) progenitors: SFHx, s25, D15, mesa20_pert and s18. These models range from the highest to the lowest LF power, respectively. The waveforms are injected into real O3a data, and the distribution of background noise triggers in the data, i.e. the false-alarm rate (FAR), is measured using a time-shift analysis. The backgrounds for the cWB full-band analysis, *BayesWave* full-band analysis and *BayesWave* LF analysis are evaluated separately to allow for an independent assessment of detection significance for each algorithm and analysis band. To qualify for the LF follow-up in this study, a CCSN injection must satisfy FAR $\leq 1\,\mathrm{yr}^{-1}$ according to the full-band cWB analysis. Hence, the waveforms for each CCSN model are injected at amplitudes corresponding to the 50% detection efficiency ($h_{\mathrm{rss},50}$) at FAR $\leq 1\,\mathrm{yr}^{-1}$. The value of $h_{\mathrm{rss},50}$ is determined empirically using the full-band cWB analysis. For each CCSN model, 175 injections are chosen arbitrarily from the full list of successful (FAR $\leq 1\,\mathrm{yr}^{-1}$) detection candidates to conduct the LF and full-band *BayesWave* follow-ups. Figure 4.5 shows that the LF detection efficiency increases, as the LF power increases. This is because the LF FAR reduces with increasing LF power, as shown Figure 4.6. The study shows that detecting CCSN models like s25, D15, and mesa20_pert with the LF analysis is not guaranteed, even if they exhibit SASI-related LF emission. That is, an unsuccessful detection with the LF follow-up does not imply an absence of LF emission. However, the converse holds true: a successful detection using the LF follow-up indicates the presence of LF emission and can therefore be used to constrain the CCSN explosion models for real detection candidates.

To demonstrate another application of the dedicated-frequency framework, we perform a HF follow-up analyses of the loudest trigger during SN 2019fcn [160]. The se-

lected trigger is not a real detection candidate because its cWB FAR does not satisfy the LVK detection benchmark (FAR $\leq 0.01\,\mathrm{yr}^{-1}$); it serves purely to demonstrate the method in practice, using the best available GW burst candidate from O3. The follow-up is conducted using the *BayesWave* full-band and HF analyses, with the noise background for each analysis measured separately. We find that the full-band *BayesWave* follow-up reduces the FAR to $6.4\,\mathrm{yr}^{-1}$, down from the initial cWB full-band analysis FAR of $22.1\,\mathrm{yr}^{-1}$. This finding reinforces the result from Ref. [139], that the successive application of cWB and *BayesWave* improves detection confidence. Since the SN 2019fcn trigger has minimal power in the range $f \leq 256\,\mathrm{Hz}$, limiting the HF follow-up analyses range to $f \geq 256\,\mathrm{Hz}$ reduces false-alarm triggers in the detector backgrounds. The HF *BayesWave* follow-up further reduces the FAR from $6.4\,\mathrm{yr}^{-1}$ to $4.6\,\mathrm{yr}^{-1}$. Altogether we find that the HF follow-up improves the detection of burst triggers with minimal low-frequency power. We also note that the *BayesWave* full-band analysis reconstruction of the SN 2019fcn trigger reveals non-negligible power at $f \sim 0.5\,\mathrm{kHz}$, a feature not observed in the corresponding cWB reconstruction. Conversely, cWB detects power at $f \sim 2\,\mathrm{kHz}$, yet *BayesWave* does not. This suggest that cWB and *BayesWave* may be sensitive to different frequency ranges. While our results do not provide sufficient evidence to confirm this claim, we recommend exploring the topic further in future work.

In conclusion, the *BayesWave* LF follow-ups within the dedicated-frequency framework are capable of identifying the presence of LF GW emissions in CCSN detections, thereby constraining CCSN explosion mechanisms. The HF follow-up for the loudest trigger during SN 2019fcn demonstrates that it is possible in principle to improve the detection significance of triggers with minimal LF power.

In this paper, cWB is not used for dedicated-frequency analyses for the reasons discussed in Section 4.4.2. However, it is an interesting avenue for future work to develop a framework which allows cWB to infer frequency-specific content of GW bursts independently of *BayesWave*, e.g. by analyzing subsets of the full-band reconstructions, instead of bandpass analyses. The dedicated-frequency framework is a versatile GW burst follow-up tool, which can improve detection confidence and characterization of the signal within a specific frequency range. The dedicated-frequency follow-ups can be tailored to any user-defined frequency range, and may therefore be tuned in the future to detect other types of burst signals with frequency-specific GW signatures, e.g. binary neutron-star post-merger remnants [32, 255] ($\gtrsim$1 kHz) and eccentric binary black-holes [256].

Acknowledgements

This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation (NSF). Parts of this research were conducted by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), through project numbers CE170100004 and CE230100016. The authors are grateful for computational resources provided by the LIGO Laboratory and supported by National Science Foundation Grants PHY-0757058 and PHY-0823459. Y. S. C. L. is supported by a Melbourne Research Scholarship. M. S. acknowledges Polish National Science Centre Grant No. UMO-2023/49/B/ST9/02777 and the Polish National Agency for Academic Exchange within Polish Returns Programme

Grant No. BPN/PPO/2023/1/00019. M. M. gratefully acknowledges support from the NSF through grant PHY-2409714.

The authors thank Chad Henshaw for the codebase that was used to produce the CWT time-frequency spectrograms presented in this paper. The authors also thank Anthony Mezzacappa, Jade Powell and Yanyan Zheng for their helpful comments.

4.A1. Extracting gravitational waves from CCSNe

In Section 1.3.3 of this thesis, we show that GWs are produced by accelerating matter with a time-varying quadrupole moment. Consequently, the mass quadrupole approximation $\ddot{M}^{\mu\nu}$ as shown in Equation 1.24 can be used to extract GWs from the CCSNe simulations. We outline the methods below.

The first step is to transform $\ddot{M}^{\mu\nu}$ into the TT-gauge for convenience. Recall that all time components of the metric perturbation in Equation 1.16 are set to zero. Therefore, the TT-gauge transformation can be written purely in the spatial coordinates, viz.

$$h_{ij}^{\mathrm{TT}} = \Lambda_{ij,kl} h_{kl}, \tag{4.10}$$

with spatial indices $i, j, k, l \in \{1, 2, 3\}$. For a wave propagating in the direction $\hat{\boldsymbol{n}}$,

$$\Lambda_{ij,kl}(\hat{\boldsymbol{n}}) = P_{ik} P_{jl} - \frac{1}{2} P_{ij} P_{kl}, \tag{4.11}$$

where $P_{ij}(\hat{\boldsymbol{n}}) = \delta_{ij} - n_i n_j$ is a symmetric and traverse tensor with trace $P_{ii} = 2$. By this definition, we can then write

$$\ddot{M}_{ij}^{\mathrm{TT}} = \Lambda_{ij,kl} \ddot{M}_{kl}. \tag{4.12}$$

For a GW propagating along $\hat{\boldsymbol{n}} = \hat{\boldsymbol{z}}$, one can combine Equation 4.12 and 1.23 to obtain:

$$h_{ij}^{\mathrm{TT}} = \frac{2G}{Dc^4} \begin{pmatrix} (\ddot{M}_{11} - \ddot{M}_{22})/2 & \ddot{M}_{12} & 0 \\ \ddot{M}_{21} & -(\ddot{M}_{11} - \ddot{M}_{22})/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij.} \tag{4.13}$$

Equating the spatial components of Equation 1.16 with those in Equation 4.13 yields:

$$h_{+} = \frac{G}{Dc^4} (\ddot{M}_{11} - \ddot{M}_{22}) \tag{4.14}$$

$$h_{\times} = \frac{2G}{Dc^4} \ddot{M}_{12}. \tag{4.15}$$

Equations 4.14 and 4.15 are used to extract the GWs from the CCSNe models in terms of their energy distribution. The energy flux of a GW through a unit surface dA is given by [69]

$$\frac{dE_{\mathrm{GW}}}{dA\,dt} = \frac{c^3}{32\pi G} \left\langle \dot{h}_{ij}^{\mathrm{TT}} \dot{h}_{ij}^{\mathrm{TT}} \right\rangle. \tag{4.16}$$

The angle brackets $\langle\cdot\rangle$ denote the temporal average over multiple oscillation periods of the GW. For GW bursts like CCSNe with duration T, the average is evaluated as follows:

$$\left\langle \dot{h}_{ij}^{\mathrm{TT}} \dot{h}_{ij}^{\mathrm{TT}} \right\rangle = \left\langle 2\left[\dot{h}_{+}^{2}(t) + \dot{h}_{\times}^{2}(t)\right]\right\rangle \tag{4.17}$$

$$= \frac{2}{T}\int_{-T/2}^{T/2} dt\left[\dot{h}_{+}^{2}(t) + \dot{h}_{\times}^{2}(t)\right]. \tag{4.18}$$

Let $dA = r^2 d\Omega$, where r is the distance to the source and Ω is the solid angle. Then substituting Equation 4.18 into Equation 4.16 yields

$$\frac{dE_{\mathrm{GW}}}{d\Omega\, dt} = \frac{c^3 r^2}{16\pi G}\frac{1}{T}\int_{-T/2}^{T/2} dt\left[\dot{h}_{+}^{2}(t) + \dot{h}_{\times}^{2}(t)\right] \tag{4.19}$$

Since the time-domain metric perturbation can be decomposed spectrally as

$$h_{ij}(t) = \int_{-\infty}^{\infty} df\, \tilde{h}_{ij}(f) e^{2\pi i f t}, \tag{4.20}$$

Equation 4.19 can be re-expressed as [254]

$$\frac{dE_{\mathrm{GW}}}{d\Omega\, dt} = \frac{\pi c^3 r^2}{4G}\frac{1}{T}\int_{-\infty}^{\infty} df \left|\tilde{h}(f)\right|^2 f^2, \tag{4.21}$$

with $\left|\tilde{h}(f)\right|^2 = \left|\tilde{h}_{+}(f)\right|^2 + \left|\tilde{h}_{\times}(f)\right|^2$. Integrating Equation 4.21 over the signal duration T and the solid angle Ω yiel ds the total GW energy:

$$E_{\mathrm{GW}} = \frac{\pi c^3 r^2}{4G}\int d\Omega \int_{-\infty}^{\infty} df \left|\tilde{h}(f)\right|^2 f^2 \tag{4.22}$$

For GWs with real amplitudes, i.e. $\tilde{h}(-f) = \tilde{h}^*(f)$, Equation 4.22 becomes

$$E_{\mathrm{GW}} = \frac{\pi c^3 r^2}{2G}\int d\Omega \int_{0}^{\infty} df \left|\tilde{h}(f)\right|^2 f^2. \tag{4.23}$$

Consequently, the GW energy spectrum is given by [257]

$$\frac{dE_{\mathrm{GW}}}{df} = \frac{\pi c^3 r^2}{8G} h_{\mathrm{char}}^2 \int d\Omega, \tag{4.24}$$

where h_{char} is the characteristic strain [103], defined such that $h_{\mathrm{char}}^2 = 4\left|\tilde{h}(f)\right|^2 f^2$.

In the case of an isotropically emitting source, Equation 4.24 simplifies to [225]

$$\frac{dE_{\mathrm{GW}}}{df} = \frac{\pi^2 c^3 r^2}{2G} h_{\mathrm{char}}^2. \tag{4.25}$$

If the isotropic source also emits a narrowband signal with central frequency f_0 (as defined in Equation 4.9), the total GW energy can be approximated as

$$E_{\mathrm{GW}} = \frac{\pi^2 c^3 r^2}{G} f_0^2 h_{\mathrm{rss}}^2. \tag{4.26}$$

Chapter 5

Multi-detector characterization of gravitational-wave burst tensor polarizations with the *BayesWave* algorithm

This chapter is a reproduction of [6]

> *Y. S. C. Lee, S. Doshi, M. Millhouse, and A. Melatos,*
> Physical Review D 111, 082002 (2025)

reformatted with the following changes only:

- **The text is styled to match the rest of this thesis.**
- **Table 5.1 is reformatted to display properly on the page.**
- **The last paragraph of Section 5.7 is revised and expanded in response to a suggestion from the thesis examiner.**
- **Appendix 5.A3 is added to present a supplementary study on how we ensure the convergence of *BayesWave*'s signal evidences for this study. This appendix is not submitted for publication.**
- **Appendix 5.A4 is added to discuss the theory of Stokes parameters, for completeness. This appendix is not submitted for publication.**
- **Brief descriptions of Appendices 5.A3 and 5.A4 are added to the relevant sections of the text.**

5.1. Abstract

Einstein's theory of general relativity predicts that gravitational waves (GWs) are tensor-polarized, with two modes of polarization: plus (h_+) and cross ($h_\times$). The unmodeled GW burst analysis pipeline, *BayesWave*, offers two tensor-polarized signal models: the elliptical polarization model (E) and the relaxed polarization model (R). Future expansion of the global GW detector network will enable more accurate studies of GW polarizations with GW bursts. Here a multi-detector analysis is conducted to compare the performance of E and R in characterizing elliptical and non-elliptical GW polarizations, using non-precessing and precessing binary black holes (BBHs) respectively as representative synthetic sources. It is found that both models reconstruct the elliptical non-precessing BBH signals accurately, but E has a higher Bayesian evidence than R as it is has fewer

model parameters. The same is true for precessing BBHs that are reconstructed equally well by both models. However, for some events with high precession and especially with three or more detectors, the reconstruction accuracy and evidence of R surpass E. The analysis is repeated for BBH events from the third LIGO-Virgo-KAGRA observing run, and the results show that E is preferred over R for existing detections. Although E is generally preferred for its simplicity, it insists on elliptical polarizations, whereas R can measure generic GW polarization content in terms of Stokes parameters. The accuracy of R in recovering polarization content improves as the detector network expands, and the performance is independent of the GW signal morphology.

5.2. Introduction

The global gravitational-wave (GW) detector network currently comprises four interferometers, namely the Hanford and Livingston Laser Interferometer Gravitational-Wave Observatory (LIGO) detectors in the USA [76], the Virgo detector in Italy [77] and the Kamioka Gravitational Wave Detector (KAGRA) in Japan [78]. With the commissioning of LIGO-India [82] in progress and the proposals to build next-generation ground-based interfrometers like the Cosmic Explorer [83] and Einstein Telescope [84], the global detector network will grow in the coming years. Benefits of a larger network include increased duty cycle, sky coverage, signal-to-noise ratio (SNR) and accuracy of source localization [90–92].

Larger networks also enable more accurate studies of GW polarization structure. General relativity (GR) predicts that GWs are a linear combination of two transverse-traceless tensor polarizations, namely plus (+) and cross (×). Metric theories beyond GR allow up to six independent polarizations, comprising additional vector modes (x and y) and scalar modes (b and l) [67, 68]. The overall amplitude of the metric perturbation in the time domain observed by a detector I is given by [156]

$$h_I(t) = \sum_P F_I^P(t,\psi,\Omega) h^P\left[t + \Delta t_I(\Omega)\right], \tag{5.1}$$

for $P \in \{+,\times,x,y,l,b\}$. In other words, the sensitivity of an individual GW detector to the amplitude h^P of polarization mode P is determined by a unique antenna pattern F^P, which is a function of time, the source orientation[1] ψ and the sky location $\Omega = \{\alpha,\delta\}$ expressed in terms of the right ascension (α) and declination (δ). $\Delta t_I(\Omega)$ is the arrival time delay between detector I and a nominal reference location.

Long-duration (continuous) GW signals can be used to measure polarizations. In theory, this can be done with a single detector by searching for the unique t dependence in $F_I^P(t,\psi,\Omega)$ [258]. Multiple detectors improve polarization detection [259]. Previous studies focus on persistent sources like a stochastic GW background [260–264] and spinning neutron stars with an asymmetric moment of inertia [265, 266]. To date, however, these signals have not been detected by ground-based interferometric detectors. In contrast, LIGO [76] and Virgo [77] have detected 90 compact binary coalescences (CBCs) up

[1] ψ is also known as the polarization angle. The standard LIGO-Virgo-KAGRA (LVK) definition of ψ can be found in Refs. [87, 88].

to the third observing run (O3) [35, 36, 38], and over 80 detection candidates from the ongoing fourth observing run (O4). Transient GW (burst) signals from CBCs typically last a few milliseconds to a few seconds, during which F_I^P in Equation 5.1 is roughly constant. Since h^P is the same for all detectors up to Δt_I, n detectors are required at minimum to break n polarization degeneracies if ψ and Ω are well-constrained [267]. However, differential-arm detectors like LIGO, Virgo and KAGRA cannot distinguish between the two scalar modes (l and b) because the antenna pattern functions are degenerate with respect to l and b, meaning that only five polarization modes can effectively be disentangled with current detector designs [268]. This is still enough in principle to distinguish between GR (tensor) and non-GR (non-tensor) polarizations.

This paper focuses on characterizing GR-consistent (tensor) polarization structures. We employ the *BayesWave* burst analysis algorithm [1, 2, 85] to detect and characterize the GW bursts with minimal assumptions. *BayesWave* provides two tensor-polarized signal models: the elliptical (E) and the relaxed (R) polarization models [1, 2, 85], represented as a sum of sine-Gaussian wavelets. E restricts the frequency-domain amplitudes to $\tilde{h}_\times = i\epsilon\tilde{h}_+$ with ellipticity parameter ϵ, whereas R reconstructs $\tilde{h}_+$ and $\tilde{h}_\times$ independently. In this work, we present an empirical study on the polarization characterization performance of E and R, for elliptical and non-elliptical GW burst signals. The Bayes factor between E and R is used to determine the model that best represents the data. According to GR, nonspinning binary black holes (BBHs) do not precess and emit elliptically polarized GWs [268]. Spinning CBCs, on the other hand, undergo precession if the component spin vectors orthogonal to the angular momentum vector are nonzero [269]. The angular momentum vector of precessing CBCs evolves with time, so the GW polarization is not strictly elliptical. Therefore we use simulated non-precessing and precessing BBHs to represent the elliptical and non-elliptical GW burst sources respectively. We also quantify how accurately R recovers GW polarization content of the non-precessing and precessing BBHs in terms of Stokes parameters. This is not possible with E because the overall polarization is described by a single ellipticity parameter ϵ. The size of a detector network impacts *BayesWave*'s signal characterization [3] and hence also the accuracy of measuring GW polarizations. For example, the Hanford (H) and Livingston (L) LIGO detectors are approximately coaligned, so their polarization measurements as in Equation 5.1 are not fully independent. As a result, the HL network provides limited information about the signal polarization structure. Thus we also compare *BayesWave*'s polarization characterization performance between the HL (two-detector), HL-Virgo (three-detector) and HLV-KAGRA (four-detector) networks.

The remaining sections of this paper are organized as follows. Section 5.3 summarizes the *BayesWave* algorithm and the signal models E and R. Section 5.4 details the non-precessing and precessing BBH injection sets used in this study. Section 5.5 discusses the aim and methods of comparing the burst characterization performance of E and R; the results are presented in Section 5.5.2 and the demonstration of methodology using real GW events from O3 is presented in Section 5.5.3. Section 5.6 details the methods to quantify the accuracy of R in recovering GW polarization; the multi-detector comparison for the non-precessing and precessing BBHs is presented in Section 5.6.3. We summarize our findings in Section 5.7.

5.3. BayesWave and Signal Models

BayesWave is a source-agnostic analysis algorithm used by the LVK community to jointly characterize instrumental glitches and generic GW bursts [35, 36, 38, 44–46]. The algorithm uses a sum of sine-Gaussian wavelets to reconstruct transient and non-Gaussian features in the data, without prior assumptions about their source. A sine-Gaussian wavelet is parameterised by its central time t_0, central frequency f_0, quality factor Q, amplitude $\mathcal{A}$ and phase ϕ_0, packaged into a vector $\boldsymbol{\lambda}$. It is expressed mathematically in the time domain as

$$\Psi(t;\boldsymbol{\lambda}) = \mathcal{A}e^{-(t-t_0)^2/\tau^2}\cos\left[2\pi f_0(t-t_0)+\phi_0\right], \tag{5.2}$$

with $\tau \equiv Q/(2\pi f_0)$. *BayesWave* uses a trans-dimensional Reversible Jump Markov Chain Monte Carlo (RJMCMC) sampler to marginalize over the wavelet parameters and the number of wavelets (i.e. model complexity) needed to fit the data. Although RJMCMC is computationally expensive, it has enough flexibility to characterize generic GW bursts without assuming a physically motivated waveform a priori.

BayesWave currently offers four independent phenomenological models: the Gaussian noise model ($\mathcal{N}$), the signal plus Gaussian noise model ($\mathcal{S}$), the glitch plus Gaussian noise model ($\mathcal{G}$), and the joint model ($\mathcal{S}$+$\mathcal{G}$) which simultaneously allows for a signal and a glitch. $\mathcal{S}$ assumes a coherent signal across all detectors in the network, whereas $\mathcal{G}$ is specific to each detector. Detailed descriptions of *BayesWave*'s models can be found in Refs. [2, 85]. The model that best describes the data is determined by calculating the Bayes factor, i.e. the ratio of Bayesian evidences. *BayesWave* uses thermodynamic integration to calculate the evidences (see Appendix 5.A1).

In this paper, we focus on comparing two GR-consistent signal models with different polarization restrictions: elliptical (E) and relaxed (R), with $\{E,R\} \in \mathcal{S}$. We describe these models in Sections 5.3.1 and 5.3.2 respectively.

5.3.1. *Elliptical polarization, E*

Equation 5.1 specifies the time-domain response of a GW detector I to a combination of polarization modes. In GR, the scalar and vector modes are omitted. As *BayesWave* calculates the likelihoods in the frequency domain, we rewrite the detector response as

$$\tilde{h}_I(f) = \left[F_I^{\times}(\psi,\Omega)\tilde{h}^{\times}(f) + F_I^{+}(\psi,\Omega)\tilde{h}^{+}(f)\right]e^{2\pi i f \Delta t_I(\Omega)}. \tag{5.3}$$

BayesWave analyzes burst signals much shorter than one day, so the diurnal and annual variations in $F_I^P(t,\psi,\Omega)$ are negligible, and we can write $F_I^P(t,\psi,\Omega) \approx F_I^P(\psi,\Omega)$. Equation 5.3 assumes that GWs travel at the speed of light and are not dispersed [156, 270]. Hence GW signals are coherent across all detectors and can be represented by a single projected waveform.

In order to disentangle $\tilde{h}_+(f)$ and $\tilde{h}_\times(f)$ from the measurements of $\tilde{h}_I(f)$ as in Equation 5.3, one needs to know ψ and Ω, which are extrinsic parameters of *BayesWave* signal model. In general, at least two detectors are required to break degeneracies between the two polarization modes, and additional detectors are required to constrain ψ and Ω. In the first observing run (O1) in 2015, only the two LIGO detectors were operating, whose

arms are approximately parallel, leading to polarization degeneracies. Therefore the first signal model implemented in *BayesWave* [1] assumes elliptical polarization (E), where $\tilde{h}^+$ is expressed as a sum of N wavelets and $\tilde{h}^\times$ is proportional to $\tilde{h}^+$ and out of phase by $\pi/2$ radians, viz.

$$\tilde{h}^+ = \sum_{n=1}^{N} \Psi\left[f; t_0^{(n)}, f_0^{(n)}, Q^{(n)}, \mathcal{A}^{(n)}, \phi_0^{(n)}\right], \tag{5.4}$$

$$\tilde{h}^\times = i\epsilon\tilde{h}^+. \tag{5.5}$$

The ellipticity parameter ϵ takes a value between zero (linear) and one (circular). Altogether E is parameterized by $5N$ intrinsic parameters from the wavelets plus the four extrinsic parameters ϵ, ψ, and $\Omega = (\alpha, \delta)$.

5.3.2. *Relaxed polarization, R*

The model E has achieved success in reconstructing GW signals like binary black holes (BBH) [140], binary neutron stars (BNS) post-merger emissions [143, 144], eccentric BBHs [145], and core-collapse supernovae (CCSNe) [149]. Most CBCs like BBHs and BNSs are elliptically polarized as in Equation 5.8, whereas generic bursts like CCSNe [271, 272] are not. Furthermore, E does not hold for CBCs with time-dependent polarization content, e.g. signals with spin-precession and/or higher-order spherical harmonic modes (beyond $\ell = |m| = 2$).

As Virgo and KAGRA join the global detector network, the task of separating the $+$ and $\times$ polarizations becomes more pressing. Each detector independently measures the combinations of polarizations as in Equation 5.3, so expanded detector networks enable better separation of the polarization modes. In response, *BayesWave* has been generalized to support the relaxed polarization model R [2]:

$$\tilde{h}^+ = \sum_{n=1}^{N} \Psi\left[f; t_0^{(n)}, f_0^{(n)}, Q^{(n)}, \mathcal{A}_+^{(n)}, \phi_{0,+}^{(n)}\right], \tag{5.6}$$

$$\tilde{h}^\times = \sum_{n=1}^{N} \Psi\left[f; t_0^{(n)}, f_0^{(n)}, Q^{(n)}, \mathcal{A}_\times^{(n)}, \phi_{0,\times}^{(n)}\right]. \tag{5.7}$$

The main difference between E and R is that $\tilde{h}^+$ and $\tilde{h}^\times$ are reconstructed separately in R without the restriction imposed by Equation 5.5. However, $\tilde{h}^+$ and $\tilde{h}^\times$ are not independent as they are both expressed as a sum of the same number (N) of wavelets, and the n-th plus and cross wavelets share the same central time $t_0^{(n)}$, central frequency $f_0^{(n)}$ and quality factor $Q^{(n)}$; only their amplitudes and phases differ. This arrangement is valid for two reasons. First, $\tilde{h}^+$ and $\tilde{h}^\times$ are generated by the same physical processes, so their time-frequency structure should be closely related [156]. Second, according to Equation 5.3, the time-frequency structure of a coherent signal should be independent of the detector; only the amplitude and phase are modified when projected from the reference location onto the detector. With R, one can also set $\psi = 0$ in Equation 5.3 without loss of generality[2]. Thus the overall dimensionality of R is given by the $7N$ intrinsic wavelet parameters $t_0^{(n)}, f_0^{(n)}, Q^{(n)}, \mathcal{A}_+^{(n)}, \phi_{0,+}^{(n)}, \mathcal{A}_\times^{(n)}, \phi_{0,\times}^{(n)}$ plus two extrinsic parameters $\Omega = (\alpha, \delta)$.

[2]This is because ψ and $\{\mathcal{A}_+^n, \mathcal{A}_\times^n\}$ are degenerate in Equation 5.3 i.e. the transformation

5.4. Injection datasets

The two key goals of this paper are (i) to test whether model E or model R reconstructs more accurately the signals from non-precessing (zero-spin) and precessing (nonzero-spin) BBHs, and (ii) to quantify how accurately R recovers generic BBH polarizations in terms of Stokes parameters. We address these goals quantitatively in Sections 5.5 and 5.6 by analyzing two separate sets of simulated injections, one comprising non-precessing BBHs and the other comprising precessing BBHs. We specify the two injection sets in this section. A previous study showed that *BayesWave*'s detection statistic, viz. the log signal-versus-glitch Bayes factor, scales with the number of detectors in the network [3]. Furthermore, the size of the detector network affects the distinguishability of polarization modes and hence the accuracy of polarization measurements [267, 270]. So we inject the BBHs into simulated noise of three detector configurations, namely Hanford-Livingston (HL, two detectors), Hanford-Livingston-Virgo (HLV, three detectors) and Hanford-Livingston-KAGRA-Virgo (HLKV, four detectors) to compare their performances in characterizing GW burst polarizations.

The non-precessing and precessing BBH injection sets are simulated using the `IMRPhenomXPHM` waveform approximant [273–276]. Each injection set consists of 200 waveforms with sky locations uniformly distributed on a sphere. The main difference between the two injection sets is the distributions of the dimensionless spin $|\boldsymbol{\chi}_{\mathcal{C}}|$, where the subscript $\mathcal{C} = 1, 2$ labels the primary and secondary black holes respectively. The effective precession spin parameter $0 \leq \chi_p \leq 1$ is conventionally used to characterize the degree of precession in BBHs [269]. Since χ_p depends on $|\boldsymbol{\chi}_{\mathcal{C}}|$ (see Appendix 5.A2), we set $|\boldsymbol{\chi}_{\mathcal{C}}| = 0$ to obtain $\chi_p = 0$ for all of the non-precessing BBH injections. For the precessing BBHs, $\boldsymbol{\chi}_{\mathcal{C}}$ and χ_p evolve with time. Although we cannot guarantee precession throughout the coalescence because the motion is too complicated to predict analytically, we can ensure that the system precesses initially. This is achieved by sampling uniformly in the domain $0.1 \leq |\boldsymbol{\chi}_{\mathcal{C}}| \leq 1$ to obtain $\chi_{p,\mathrm{init}} \neq 0$ initially at $f = 16$ Hz for all precessing BBH injections. Equation 5.38 shows that χ_p depends only on $\left|\boldsymbol{\chi}_{\mathcal{C},\perp}\right|$ i.e. spin components orthogonal to the angular momentum vector of the binary system. In other words, only part of $|\boldsymbol{\chi}_{\mathcal{C}}|$ contributes to χ_p. Therefore we impose a lower bound of 0.1 on $|\boldsymbol{\chi}_{\mathcal{C}}|$ to avoid obtaining injections with negligible $\left|\boldsymbol{\chi}_{\mathcal{C},\perp}\right|$, and hence $\chi_{p,\mathrm{init}}$, which are indistinguishable from non-precessing BBHs.

In this work, we use the non-precessing (precessing) BBHs injections to compare the performance between R and E for elliptical (non-elliptical) GW signals. Therefore we need to ensure that the injection sets fulfill their respective polarization requirements. The polarization of non-precessing CBCs is quantified by the ellipticity $\epsilon = |h_\times(t)|/|h_+(t)|$, which depends on the inclination angle ι between the line of sight and the normal to the orbital plane [64], according to

$$\frac{|h_\times(t)|}{|h_+(t)|} = \frac{2\cos\iota}{1+\cos^2\iota}. \tag{5.8}$$

Therefore, in theory, all inclination angles except $\iota = 90^\circ$ produce elliptically-polarized

$\{F^+(\psi,\Omega), F^\times(\psi,\Omega)\} \mapsto \{F^+(\psi',\Omega), F^\times(\psi',\Omega)\}$ for $\psi' = \psi + \Delta\psi$ can be fully described by $\{\mathcal{A}_+^{(n)}, \mathcal{A}_\times^{(n)}\} \mapsto \{\mathcal{A}_+^{(n)'}, \mathcal{A}_\times^{(n)'}\}$ (see Appendix A of Ref. [259] for more details).

non-precessing BBHs that are valid for our study. However, the source orientation affects the radiated power, e.g. $\iota = 90°$ (edge-on) results in minimum power, and $\iota = 0°$ or $180°$ (face-on or face-off) results in maximum power [90]. This, in turn, can affect the distinguishability of the GW polarization modes, so we select ι with caution. For non-precessing BBHs, ι is fixed throughout the inspiral. In contrast, for precessing BBHs, ι evolves throughout the inspiral and the polarizations are not related as in Equation 5.8. Since $|\chi_{\mathcal{C}}|$ of the precessing BBHs is uniformly sampled, some injections are expected to undergo strong precession, leading to significant deviations from the initial ι, while others will exhibit minimal precession, maintaining an inclination close to the initial ι. Although only the strongly precessing BBHs are useful for comparing the E and R performance between non-precessing and precessing signals, the minimally precessing BBHs can be used to assess the consistency of E and R compared to the non-precessing BBHs, provided their ι values are similar. For this reason, we nominally set $\iota = 45°$ for all non-precessing BBHs and initialize all precessing BBH inspirals with $\iota = 45°$ at $f = 16$ Hz to facilitate the consistency check. The initial ι does not affect the strongly precessing BBHs, as they are expected to deviate from it regardless.

To compare R versus E fairly, we keep all other parameters fixed across both injection sets. Asymmetric masses are known to amplify precession-related GR effects [277]. Hence we set the primary and secondary black hole masses to $m_1 = 40M_\odot$ and $m_2 = 8M_\odot$ respectively, to achieve a mass ratio of $q = m_1/m_2 = 5$ for all injections. The source distances are chosen to obtain an injected network signal-to-noise ratio $\mathrm{SNR}_{\mathrm{net}} = 50$ in simulated HLV data. The same distance distribution is used for the HL and HLKV injections, i.e. the sources have the same inherent loudness, but $\mathrm{SNR}_{\mathrm{net}}$ is adjusted by removing Virgo contributions and adding KAGRA contributions respectively. As a result, the HLKV injections have the highest $\mathrm{SNR}_{\mathrm{net}}$, followed by HLV and then HL. All non-precessing and precessing BBHs are injected into Gaussian noise colored by the sensitivity curves shown in Figure 5.1. In this paper, with an eye to the future, we use the same sensitivity curves for Virgo and KAGRA. We also note that $\mathrm{SNR}_{\mathrm{net}}(\mathrm{HLV}) = 50$ is a few factors higher than the BBH detections reported in the Gravitational-wave Transient Catalogs (GWTCs), where the HLV network matched-filter SNRs is typically ≤ 25 [35, 36, 38]. Therefore the studies presented in this paper are a proof of concept. We do not use the results to make claims about the polarization properties of existing detections.

For each injection set, we perform two separate *BayesWave* analyses with the signal models E and R. To ensure comparability between the two analyses, both analyses are configured the same. Firstly we assume fixed detector noise curves as in Figure 5.1 for all the analyses. By convention, we set the lower bound of the analysis frequency range to 20 Hz, to avoid the high-amplitude noise floor <20 Hz [94]. We choose a sampling rate of 2048 Hz, where the Nyquist frequency sets the upper analysis frequency bound at 1024 Hz. The overall analysis frequency range is therefore 20-1024 Hz which is sufficient for the detection of BBHs. The analyzed segments are set to 4 seconds, which is longer than typical CBC signals (∼1 second), to prevent truncating detectable segments of the BBH signals.

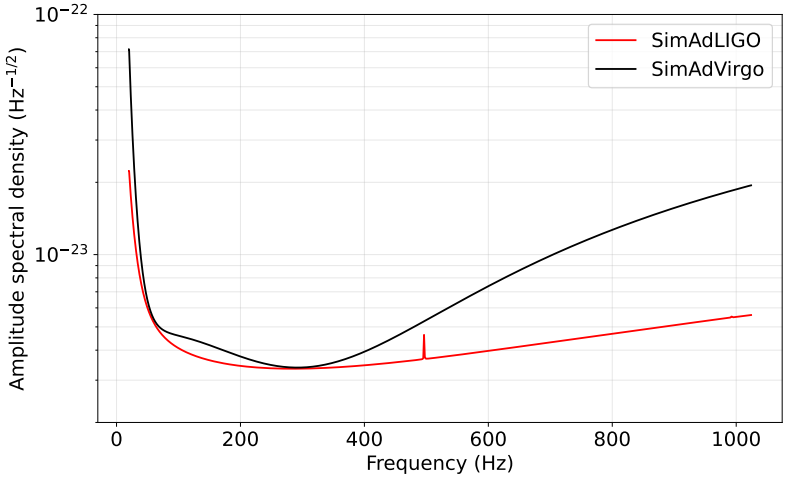

Figure 5.1: Simulated amplitude spectral density (ASD) curves of the Advanced LIGO (SimAdLIGO, red) and Virgo (SimAdVirgo, black) detectors from the LVK Algorithm Library [278]. SimAdLIGO represents the LIGO Hanford (H) and Livingston (L) detectors, whereas SimAdVirgo represents Virgo (V) and KAGRA (K) in this study.

5.5. R versus E

Figure 1 of Ref. [2] qualitatively demonstrates that R reconstructs BBH signals with spin-induced precession more accurately than E, assuming a zero-noise realization. The figure shows the injected and reconstructed time-domain signal amplitudes for one non-precessing and two precessing BBHs. Both E and R perform comparably when reconstructing the non-precessing BBH signal. On the other hand, the amplitudes of the precessing BBHs are modulated, and E can reproduce the modulations to some extent, although they are represented better by R. However, better reconstruction does not necessarily imply that R is supported more strongly by the data, once the Occam penalty in the Bayesian evidence is taken into account. Furthermore, the true signal morphology and hence the accuracy of reconstruction are unknown when analyzing real, astronomical data.

In this work, we use the Bayes factor between R and E ($\ln \mathcal{B}_{R,E}$) to study whether there is a preferred polarization model for different signal morphologies. If there is, can $\ln \mathcal{B}_{R,E}$ be used to distinguish between elliptical and non-elliptical polarizations? We use the non-precessing (elliptical) and precessing (non-elliptical) BBH injections described in Section 5.4 as representative case studies. In Section 5.5.1, we define and justify the metrics used to compare E and R. The results involving synthetic data are presented in Section 5.5.2. We also demonstrate how one can compute the relevant comparison metrics for real O3 events in Section 5.5.3.

5.5.1. *Comparison metrics*

It is convenient to compute the log Bayes factor (i.e. the log evidence ratio) between R and E,

$$\ln \mathcal{B}_{R,E} = \ln p(\mathbf{d}|R) - \ln p(\mathbf{d}|E), \tag{5.9}$$

as a basis for model selection. $\ln \mathcal{B}_{R,E} > 0$ indicates that R is more strongly supported by the data $\mathbf{d}$ than E, and vice versa. The comparison via Equation 5.9 is valid when the E and R analyses use the same noise power spectral density (PSD). That is, we assume the PSD is known and is the same in each analysis. In Appendix 5.A3, we conduct a supplementary study to identify an appropriate *BayesWave* sampler setting that ensures reliable model evidence calculations, to achieve accurate comparisons of E and R.

Our goal is to study how $\ln \mathcal{B}_{R,E}$ varies as the discrepancy increases between the signal waveforms reconstructed by E and R. We apply the standard LVK definition of the network overlap between the signal waveforms h^R and h^E reconstructed by R and E respectively [140]:

$$\mathcal{O}_{R,E} \equiv \frac{\langle h^R | h^E \rangle}{\sqrt{\langle h^R | h^R \rangle \langle h^E | h^E \rangle}}. \tag{5.10}$$

For a network with $\mathcal{I}$ detectors, one has

$$\langle h^a | h^b \rangle = \sum_{i=1}^{\mathcal{I}} \langle \tilde{h}_i^a | \tilde{h}_i^b \rangle, \tag{5.11}$$

and the noise-weighted inner product between waveforms $\tilde{h}_i^a$ and $\tilde{h}_i^b$ of the i-th detector is given by

$$\langle \tilde{h}_i^a | \tilde{h}_i^b \rangle = 4\,\mathrm{Re} \int_0^\infty df\, \frac{\tilde{h}_i^a(f) \tilde{h}_i^{b*}(f)}{S_{n,i}(f)}. \tag{5.12}$$

In Equation 5.12, $\tilde{h}^*$ denotes the complex conjugate of $\tilde{h}$ and $S_{n,i}(f)$ is the one-sided[3] noise PSD of detector i. By definition, $\mathcal{O}_{R,E}$ takes values between -1 and 1, where $\mathcal{O}_{R,E} = 1$ indicates perfect agreement between the E and R waveforms, $\mathcal{O}_{R,E} = 0$ indicates no agreement and $\mathcal{O}_{R,E} = -1$ indicates anti-correlation.

For simulated injections, network overlaps between the injected and recovered signal waveforms ($\mathcal{O}_{\mathrm{inj},R}$ and $\mathcal{O}_{\mathrm{inj},E}$) are typically used to quantify the reconstruction accuracy of *BayesWave*'s signal models i.e. they are good performance indicators. Thus we take advantage of the simulated injections to supplement our understanding about how $\mathcal{O}_{\mathrm{inj},R}$ and $\mathcal{O}_{\mathrm{inj},E}$ correlate with $\mathcal{O}_{R,E}$, for different BBH waveforms. We achieve this by evaluating the overlap ratio $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E}$ as a function of $\mathcal{O}_{R,E}$, where $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E} > 1$ indicates that R recovers the injected signal more accurately than E. However, $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E}$ quantifies the relative reconstruction accuracy of the two models. Hence we also present the results for $\mathcal{O}_{\mathrm{inj},R}$ versus $\mathcal{O}_{\mathrm{inj},E}$, to compare the absolute reconstruction accuracy of each model.

To illustrate how the overlaps ($\mathcal{O}_{R,E}$, $\mathcal{O}_{\mathrm{inj},R}$, $\mathcal{O}_{\mathrm{inj},E}$) quantify agreement between waveforms, we show the time-domain signal reconstructions of one non-precessing BBH and three precessing BBHs with contrasting $\chi_{p,\mathrm{init}}$ in Figure 5.2. The purple and green shaded regions show the 90% credible interval of the E and R recovered waveforms respectively. Most of the purple and green shaded regions enclose the injected waveform (black solid curves), implying that both E and R reconstructions are comparable and consistent with the true waveform. This is also reflected in the high overlaps, i.e.

[3] It makes sense to work with the one-sided PSD $S_{n,i}(f)$, because the GW signal and detector output are real, which implies $S_{n,i}(-f) = S_{n,i}(f)$. This is also why the Fourier integral in Equation 5.12 only includes positive frequencies [103].

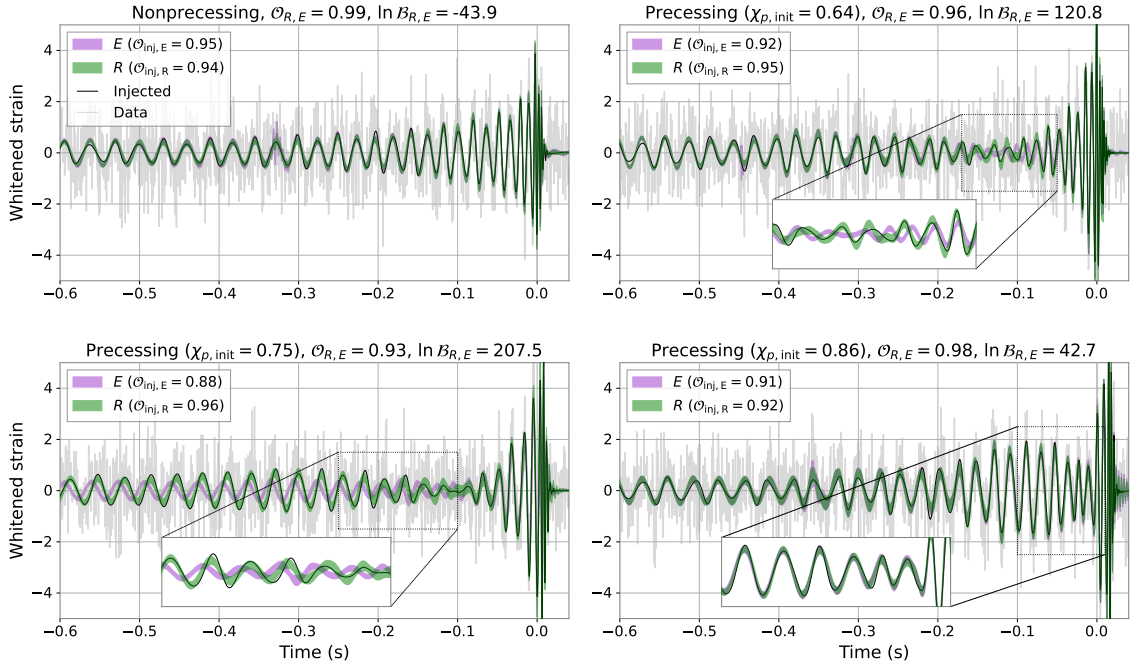

Figure 5.2: Injected and recovered time-domain signals for an arbitrarily chosen detector (LIGO Hanford). The top left panel shows the signal for a non-precessing BBH; the remaining three panels show signals for precessing BBHs with $\chi_{p,\mathrm{init}} = 0.64, 0.75$ and 0.86. The black solid curves indicate the injected signal and the grey curve indicate the detector data, which includes both the signal and noise. The 90% credible intervals for the waveforms recovered by E and R are indicated by the purple and green shaded bands respectively. $\mathcal{O}_{R,E}$ and $\ln \mathcal{B}_{R,E}$ for each event is shown at the top of the respective panels and the network overlaps $\mathcal{O}_{\mathrm{inj},R}$ and $\mathcal{O}_{\mathrm{inj},E}$ are shown in the legends. The overlaps displayed here are obtained from the HLKV analyses; these four events are injected into HLKV with $\mathrm{SNR}_{\mathrm{net}} \approx 65$.

$\mathcal{O}_{\mathrm{inj},E} \geq 0.88$ and $\mathcal{O}_{\mathrm{inj},R} \geq 0.90$ as quoted in the legends. However, the two precessing events with $\chi_{p,\mathrm{init}} = 0.64$ (top right) and $\chi_{p,\mathrm{init}} = 0.75$ (bottom left) in Figure 5.2 show qualitatively better reconstructions with R (green) then E (purple). The insets show that the improvements with R are especially prominent where the amplitudes are modulated. Quantitatively, we also obtain $\mathcal{O}_{\mathrm{inj},R} > \mathcal{O}_{\mathrm{inj},E}$ for these two events. Due to the discrepancies between R and E, $\mathcal{O}_{R,E}$ is also comparatively lower (0.96 and 0.93) than for the other two events (0.98 and 0.99). We also find that $\ln \mathcal{B}_{R,E}$ for these two events (120.8 and 207.5) exceeds the other two (−43.9 for non-precessing and 42.7 for precessing with $\chi_{p,\mathrm{init}} = 0.86$). Altogether $\mathcal{O}_{R,E}$ and $\ln \mathcal{B}_{R,E}$ suggest that the R model is strongly preferred over E for the two precessing events with $\chi_{p,\mathrm{init}} = 0.64$ and $\chi_{p,\mathrm{init}} = 0.75$. Notably, $\ln \mathcal{B}_{R,E}$ for the event with $\chi_{p,\mathrm{init}} = 0.86$ is lowest among the three precessing BBH examples, despite having the highest $\chi_{p,\mathrm{init}}$. Furthermore, its signal exhibits modulations like the other two precessing events, but the inset shows that the modulations are reconstructed equally well by E and R, yielding $\mathcal{O}_{\mathrm{inj},R} \approx \mathcal{O}_{\mathrm{inj},E}$.

To justify the above observations, we analyze the injected polarization content of the three precessing BBHs (plots omitted for brevity). We find that the $\chi_{p,\mathrm{init}} = 0.86$ event has higher circular polarization than the other two precessing events. Therefore, E char-

acterizes the $\chi_{p,\mathrm{init}} = 0.86$ event with an accuracy comparable to R. In summary, the above analysis implies two things. (1) Precessing BBHs evolve towards different inclination angles (and hence ellitpical polarizations) unpredictably, without a simple connection to $\chi_{p,\mathrm{init}}$. This is because it is possible for some BBHs with high initial precession to evolve into a stable configuration with slowly varying inclination before merging. Therefore, some BBHs with high $\chi_{p,\mathrm{init}}$ may emit more elliptical GW signals in the later stages of the inspiral. Conversely, BBHs with low $\chi_{p,\mathrm{init}}$ sometimes exhibit the opposite behavior. (2) Amplitude modulations in $h_+(t)$ and $h_\times(t)$ do not necessarily imply significant deviations from elliptical polarization. As a result, E is able to reconstruct some of the modulations, but R does better in general. These features agree with Figure 1 of Ref. [2].

5.5.2. *Non-precessing versus precessing BBHs*

Figure 5.3 presents the R versus E analyses for the non-precessing and precessing BBH injections, using the comparison metrics described above. The top row plots $\ln \mathcal{B}_{R,E}$ as a function of $\mathcal{O}_{R,E}$; the bottom row plots $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E}$ as a function of $\mathcal{O}_{R,E}$. All plots show results from the HL, HLV and HLKV analyses. The comparison of R and E through $\ln \mathcal{B}_{R,E}$ is meaningful, when the event is consistent with the signal model. Therefore Figure 5.3 only plots events with signal evidences greater than the glitch and/or Gaussian noise model in the analysis, i.e. events that satisfy $\ln \mathcal{B}_{R,\mathcal{G}} > 0$ and $\ln \mathcal{B}_{R,\mathcal{N}} > 0$, and similarly for E. Less than three of the 200 injections are removed from the analysis per detector configuration for failing to satisfy these conditions. This is expected, because all BBHs in this study are injected with high SNR, so noise- or glitch-like events are scarce. All the removed events have low SNR in one or more detectors, resulting in higher evidence for the incoherent $\mathcal{G}$ model than the coherent signal models R and E.

We first discuss the plots of $\ln \mathcal{B}_{R,E}$ versus $\mathcal{O}_{R,E}$ in the top row of Figure 5.3. In order to compare how the non-precessing and precessing BBHs are distributed, we segment both plots in the top row using black vertical and horizontal dashed lines placed at $\mathcal{O}_{R,E} = 0.98$ and $\ln \mathcal{B}_{R,E} = 0$ respectively. We set $\mathcal{O}_{R,E} = 0.98$ as the nominal threshold above which E and R models do equally well in fitting the data[4]. $\ln \mathcal{B}_{R,E} = 0$, by definition, is where the model evidences of E and R are equal i.e. neither model is preferred over the other from a Bayesian perspective. The top left plot in Figure 5.3 shows that non-precessing BBHs are generally found in the bottom right segment, with $\mathcal{O}_{R,E} > 0.98$ and $\ln \mathcal{B}_{R,E} < 0$. This is true for all three detector configurations. The evidence is higher for E than R because of the Occam penalty on unnecessary model dimensions of R. Previous studies have shown that for high SNR ($\gtrsim 10$) events, the Bayesian evidence of model $\mathcal{H}$ can be approximated by [3, 85]

$$p(\mathbf{d}|\mathcal{H}) \simeq p(\mathbf{d}|\boldsymbol{\theta}_{\mathrm{MAP}}, \mathcal{H}) \frac{\Delta V_{\mathcal{H}}}{V_{\mathcal{H}}}, \tag{5.13}$$

where $p(\mathbf{d}|\boldsymbol{\theta}_{\mathrm{MAP}}, \mathcal{H})$ is the likelihood evaluated at the maximum *a posteriori* (MAP) parameter value[5] $\boldsymbol{\theta}_{\mathrm{MAP}}$, and $\Delta V_{\mathcal{H}}$ and $V_{\mathcal{H}}$ are respectively the posterior and prior volumes

[4] We set the threshold as $\mathcal{O}_{R,E} = 0.98$ instead of $\mathcal{O}_{R,E} = 1$, because the model parameters of E and R differ, so it is unrealistic to expect their reconstructions to match exactly.

[5] $\boldsymbol{\theta} = \boldsymbol{\theta}_{\mathrm{MAP}}$ is the parameter value at which the posterior $p(\boldsymbol{\theta}|\mathbf{d}, \mathcal{H})$ peaks.

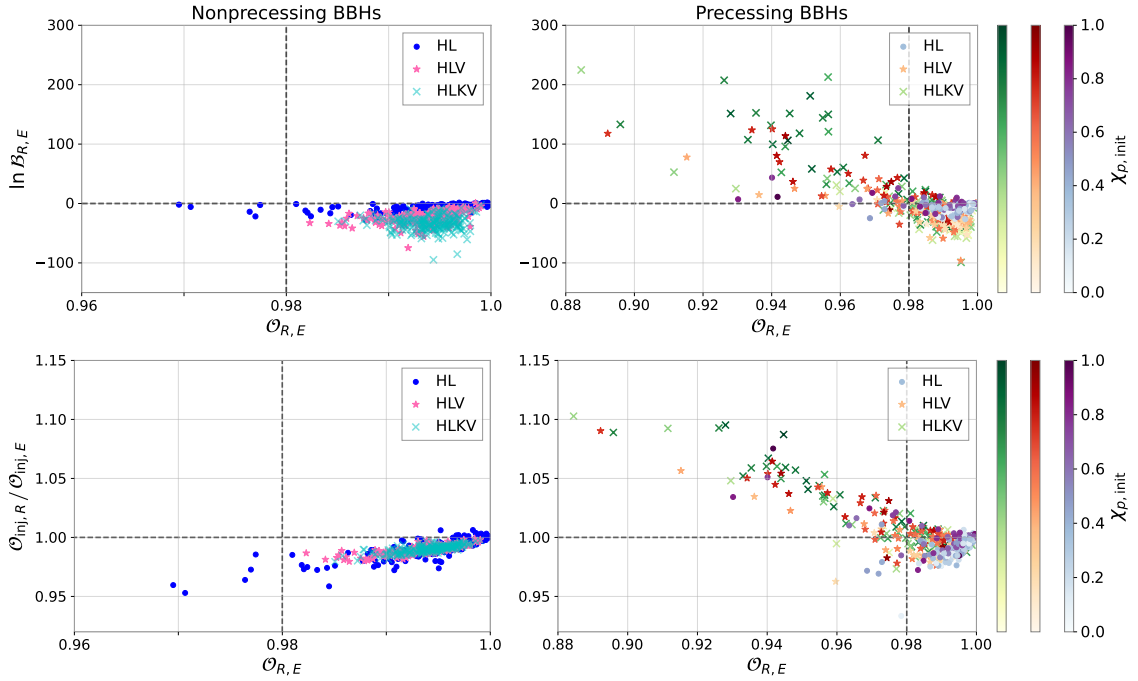

Figure 5.3: Performance of R versus E. The top row shows $\ln \mathcal{B}_{R,E}$ versus $\mathcal{O}_{R,E}$ and the bottom row shows $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E}$ versus $\mathcal{O}_{R,E}$. In each row, the left panel shows the properties of the non-precessing BBH injections for HL (dark blue circles), HLV (pink stars) and HLKV (light blue crosses). The right panel shows the properties of the precessing BBH injections for HL (purple circles), HLV (orange stars) and HLKV (green crosses). The horizontal scales are different for the left and right panels, but the same in the corresponding top and bottom panels. The three color bars per row apply to the right panels only; they indicate $\chi_{p,\mathrm{init}}$ of the precessing BBHs for HL, HLV, and HLKV in matching colors. All non-precessing binaries in the left panels have $\chi_p = 0$. The dashed lines in the top row at $\ln \mathcal{B}_{R,E} = 0$ (horizontal) and $\mathcal{O}_{R,E} = 0.98$ (vertical) divide the plot into four quadrants, as do the dashed lines at $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E} = 1$ (horizontal) and $\mathcal{O}_{R,E} = 0.98$ (vertical) in the bottom row.

of $\mathcal{H}$. The fraction $\Delta V_{\mathcal{H}}/V_{\mathcal{H}}$ is known as the Occam penalty because it suppresses the evidence of models with unnecessarily large prior volume. Equations 5.9 and 5.13 can be combined to give

$$\ln \mathcal{B}_{R,E} \simeq \ln \frac{p(\mathbf{d}|\boldsymbol{\theta}_{\mathrm{MAP}}, R)}{p(\mathbf{d}|\boldsymbol{\theta}_{\mathrm{MAP}}, E)} + \ln \frac{\Delta V_R}{\Delta V_E} + \ln \frac{V_E}{V_R}. \tag{5.14}$$

We find empirically $\mathcal{O}_{R,E} > 0.98 \approx 1$ for the non-precessing BBHs. The high $\mathcal{O}_{R,E}$ implies that the likelihoods of E and R are similar. As a result, we have $p(\mathbf{d}|\boldsymbol{\theta}_{\mathrm{MAP}}, R) \approx p(\mathbf{d}|\boldsymbol{\theta}_{\mathrm{MAP}}, E)$. Thus, the first term on the right-hand side of Equation 5.14 is approximately zero. Likewise the second term is approximately zero, because E and R have approximately equal posterior volumes. Therefore when the reconstructed waveforms of R and E agree closely, $\ln \mathcal{B}_{R,E}$ depends mainly on the ratio of prior volumes V_E/V_R. As discussed in Section 2.1, R has $7N+2$ parameters and E has $5N+4$. Since all *BayesWave* signal models require $N \geq 1$, R always has more parameters than E, implying $V_R > V_E$.

Hence for non-precessing BBHs, the simpler model E is preferred. There are no obvious correlations between $\ln \mathcal{B}_{R,E}$ and $\mathcal{O}_{R,E}$ for the non-precessing BBHs (top left panel of Figure 5.3). However, as the detector network expands, $\mathcal{O}_{R,E}$ generally shifts closer to unity and $\ln \mathcal{B}_{R,E}$ becomes more negative. The accuracy of E and R reconstructions, and hence the similarity of their reconstructions, increases with additional detectors. As a result, their posterior distributions also become more similar, and the first and second terms of Equation 5.14 become negligible. Hence $\ln \mathcal{B}_{R,E}$ depends more on $\ln(V_E/V_R)$ and becomes more negative.

The top right panel of Figure 5.3 shows $\ln \mathcal{B}_{R,E}$ versus $\mathcal{O}_{R,E}$ for the precessing BBHs. While most events reside in the bottom right quadrant as with the non-precessing BBHs, the top left quadrant ($\mathcal{O}_{R,E} < 0.98$, $\ln \mathcal{B}_{R,E} > 0$) is populated by 5%, 17% and 20% of the HL, HLV and HLKV precessing BBHs. In fact there is a visible trend, where $\ln \mathcal{B}_{R,E}$ increases, as $\mathcal{O}_{R,E}$ decreases, implying that the evidence of R for precessing BBHs increases when the reconstructions of E and R deviate further from one another. We also note that the color of the data points darken, as one moves from the bottom right to the top left quadrant, indicating that $\chi_{p,\mathrm{init}}$ increases. In fact, $\ln \mathcal{B}_{R,E} > 0$ is generally observed in events with $\chi_{p,\mathrm{init}} \gtrsim 0.5$. This is because BBHs with high initial precession are more likely to maintain precessional motion throughout the inspiral, leading to strong signal modulations that cannot be modeled by E, thereby strengthening the preference for R. However, as noted in Figure 5.2, deviations from elliptical polarization do not connect simply and predictably to $\chi_{p,\mathrm{init}}$, i.e. signals with high $\chi_{p,\mathrm{init}}$ can evolve into low-ellipticity signals, and vice versa. Therefore, we also observe some dark-colored (light-colored) data points in the bottom right (top left) quadrants, that do not follow the general $\chi_{p,\mathrm{init}}$ color trend. As quoted above, the percentage of events in the top left quadrant is highest for HLKV, followed by HLV then HL, which suggests that $\ln \mathcal{B}_{R,E} > 0$ is more likely with larger detector networks. We also note that $\ln \mathcal{B}_{R,E}$ increases with the number of detectors $\mathcal{I}$.

In the bottom row of Figure 5.3, we plot $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E}$ versus $\mathcal{O}_{R,E}$ to compare the reconstruction accuracy of E and R, as their inferred waveforms deviate. As in the top row, to compare the non-precessing and precessing BBHs, we divide the plots into four segments by placing vertical and horizontal black dashed-lines respectively at $\mathcal{O}_{R,E} = 0.98$ and $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E} = 1$, where E and R reconstruct the injected signal equally well. The bottom left plot shows that non-precessing BBHs generally fall within the bottom right quadrant, with $\mathcal{O}_{R,E} > 0.98$ and $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E} < 1$, i.e. non-precessing BBHs are reconstructed more accurately by E than R. This is because non-precessing BBHs are fundamentally elliptically polarized, so they are represented adequately by E. The polarization of R, on the other hand, is not restricted, which is disadvantageous when reconstructing fundamentally elliptical waveforms like non-precessing BBHs, especially when the SNR is low. However $\mathrm{SNR}_{\mathrm{net}}$ of the non-precessing BBHs used in this study is high ($\gtrsim 40$), so E and R perform similarly, i.e. we find $0.95 < (\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E}) < 1$. Furthermore we observe a trend in the bottom left plot of Figure 5.3, where $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E}$ decreases with $\mathcal{O}_{R,E}$. Decreasing $\mathcal{O}_{R,E}$ indicates that the waveform inferred by R becomes less elliptical, resulting in lower reconstruction accuracy for the non-precessing BBHs. We also find that $\mathcal{O}_{R,E}$ of HL is generally lower than HLV and HLKV, because HL has the lowest $\mathrm{SNR}_{\mathrm{net}}$ amongst the three configurations. This confirms that R deviates further from E as the SNR decreases. In the bottom right panel of Figure 5.3, we see that

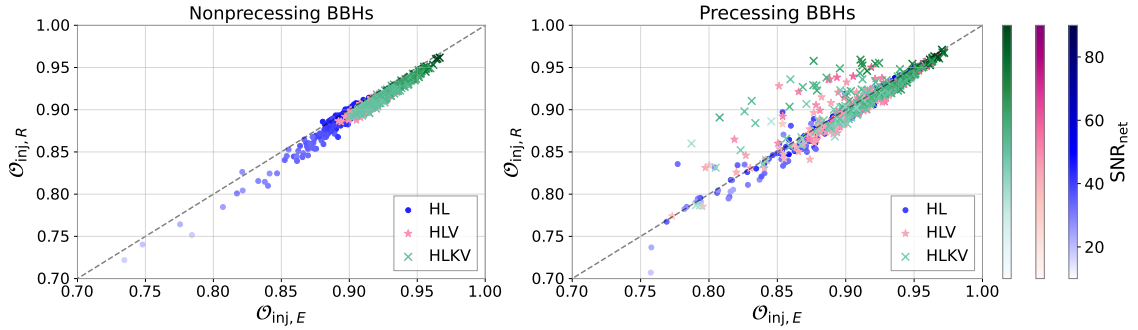

Figure 5.4: $\mathcal{O}_{\mathrm{inj},R}$ versus $\mathcal{O}_{\mathrm{inj},E}$ for the non-precessing (left panel) and precessing (right panel) BBHs. The blue circles, pink stars and green crosses indicate HL, HLV, and HLKV events respectively. The three color bars indicate $\mathrm{SNR}_{\mathrm{net}}$ per event for HL, HLV and HLKV in corresponding colors; this applies to both panels. The diagonal grey dashed lines in both panels indicate $\mathcal{O}_{\mathrm{inj},R} = \mathcal{O}_{\mathrm{inj},E}$.

$\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E}$ for the precessing BBHs follows similar trends as for the non-precessing BBHs for $\mathcal{O}_{R,E} > 0.98$. However, the trend changes for $\mathcal{O}_{R,E} < 0.98$, where we observe that $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E}$ increases, as $\mathcal{O}_{R,E}$ decreases. Note that the horizontal scales in the top and bottom rows of Figure 5.3 are the same. We infer that events with $\mathcal{O}_{R,E} < 0.98$ and $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E} > 1$ in the top left quadrant (bottom right panel) correspond to those with $\ln \mathcal{B}_{R,E} > 0$ in the top right panel. In other words, when R is favored by model selection, the reconstruction accuracy of R is also higher than E. This is expected for the reasons discussed previously, that is, R can reconstruct precession-induced amplitude modulations of precessing BBHs better than E.

Figure 5.4 shows the event-wise comparison of $\mathcal{O}_{\mathrm{inj},R}$ versus $\mathcal{O}_{\mathrm{inj},E}$ for the non-precessing (left panel) and precessing (right panel) BBHs. The information presented in this plot complements the bottom row in Figure 5.3 by explicitly showing the reconstruction accuracy of each signal model, i.e. $\mathcal{O}_{\mathrm{inj},R}$ and $\mathcal{O}_{\mathrm{inj},E}$, instead of their ratio $\mathcal{O}_{\mathrm{inj},R}/\mathcal{O}_{\mathrm{inj},E}$. Figure 5.4 shows that $\mathcal{O}_{\mathrm{inj},R}$ and $\mathcal{O}_{\mathrm{inj},E}$ exceed 0.7 individually for both the non-precessing and precessing BBHs, in all three detector configurations. That is, both E and R are able to reconstruct the precessing and non-precessing BBHs signals reliably in general; but as discussed above, one model can outperform the other depending on the signal morphology. We find that the HLKV (green) events are reconstructed most accurately with $\mathcal{O}_{\mathrm{inj},R}$ and $\mathcal{O}_{\mathrm{inj},E}$ generally exceeding 0.85, followed by HLV (exceeding 0.80) and then HL (exceeding 0.70). This makes sense because HLKV events have the highest $\mathrm{SNR}_{\mathrm{net}}$ as indicated by the color bars. This observation, where overlap increases with $\mathrm{SNR}_{\mathrm{net}}$, is consistent with Refs. [3, 140, 151].

In summary, we compare the multi-detector network performance of the two *BayesWave* signal models, R and E, for elliptical and non-elliptical GW burst signals. We use non-precessing and precessing BBHs to represent the elliptical and non-elliptical signals respectively. Our analyses find that signal reconstructions by E and R are comparable for the non-precessing BBHs signals, so the simpler E model attracts a lower Occam penalty and is typically preferred over R. The same is true for most ($\gtrsim 80\%$) of the precessing BBHs because, even though the polarizations of all precessing BBHs are non-elliptical to some extent, most do not deviate far enough from elliptical polarizations to require R for accurate signal reconstruction. However, for precessing BBHs with more pronounced

non-elliptical polarizations, the R reconstruction diverges from E, with the former providing more accurate representations of the signal. In such cases, R is preferred over E. This is more commonly observed in BBHs whose precession amplitude is relatively large, and in larger detector networks. From these results, we conclude that: (i) $\ln \mathcal{B}_{R,E}$ and $\mathcal{O}_{R,E}$ together can provide indications for non-elliptical polarization, but are insufficient to distinguish between purely elliptical and slightly non-elliptical signals; and (ii) R can improve signal characterization of GW signals for non-elliptical polarizations, especially with expanded detector networks.

5.5.3. *Demonstration with O3 data*

We extend the $\ln \mathcal{B}_{R,E}$ versus $\mathcal{O}_{R,E}$ analyses to real GW events from O3. The analyses of simulated BBHs in Section 5.5.2 use Gaussian noise colored by fixed PSDs as in Figure 5.1. In contrast, PSDs for real data are not known a priori, and vary temporally. For the analysis of real events, we estimate the PSD using the *BayesLine* algorithm, which uses an RJMCMC to model the PSDs as a smooth Akima spline with Lorentzians to reconstruct the narrow-band spectral features [128, 279]. For each event, we use *BayesLine* to produce a posterior distribution of the on-source PSD, and use the median of that posterior distribution as the PSD used in BayesWave's likelihood calculation for both the R and E analyses. By using the same PSD estimate, we are able to directly compare evidences between the two models.

BayesWave signal reconstructions are known to improve with decreasing signal duration [140, 152]; short-duration signals occupy less time-frequency volume, so *BayesWave* produces more compact and faithful wavelet representations. Hence, for the R versus E analysis, we use a subset of O3 events that are sufficiently loud and short for *BayesWave* to produce reliable reconstructions. We list the events in Table 5.1. These events are also used for *BayesWave*'s waveform consistency tests previously in GWTC-2 (O3a) [36] and GWTC-3 (O3b) [38]. The O3 events in Table 5.1 also span a range of chirp masses $\mathcal{M} = (m_1 m_2)^{3/5}/(m_1+m_2)^{1/5}$, where m_1 and m_2 are the primary and secondary component masses. GW signals for low-$\mathcal{M}$ mergers last longer in the LVK detector frequency bands and the components merge at higher frequencies, so longer analysis segments and higher sampling rates are required to accurately reconstruct the time-frequency structure. The opposite is true for high-$\mathcal{M}$ events. Following the general procedure of waveform reconstructions in GWTC2 [36] and GWTC3 [38], we use the same settings for the segment length, sampling rate, and lower frequency bound used in the parameter estimation analysis presented in the catalogs [6].

The values of $\ln \mathcal{B}_{R,E}$ and $\mathcal{O}_{R,E}$ for each event are listed in Table 5.1. We also plot $\ln \mathcal{B}_{R,E}$ against $\mathcal{O}_{R,E}$ in Figure 5.5 to study their mutual correlation. We divide Figure 5.5 into the same four quadrants as in the top panels of Figure 5.3. We find that $\mathcal{O}_{R,E}$ for the O3 events is generally lower than for the simulated BBHs. This is because the SNRs of O3 events, as shown in the forth column of Table 5.1, are lower than for the simulated BBHs (SNR$\sim$50), resulting in larger discrepancies between E and R. Nevertheless, we find $\mathcal{O}_{R,E} \gtrsim 0.90$ for the O3 events, indicating that E and R reconstructions are still comparable. Figure 5.5 also shows that we have $\ln \mathcal{B}_{R,E} < 0$ for the O3 events, and there

[6] The parameter estimation configurations (as well as results) can be found via the Online Gravitational-Wave Transient Catalog: https://gwosc.org/eventapi/html/GWTC/

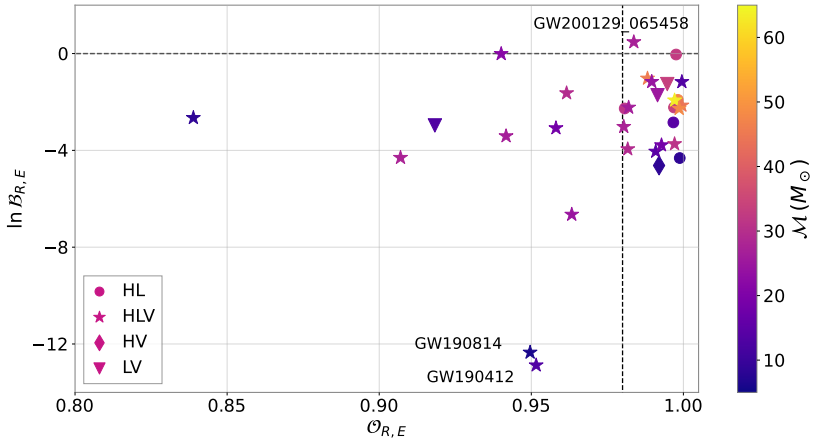

Figure 5.5: $\ln \mathcal{B}_{R,E}$ versus $\mathcal{O}_{R,E}$ for O3 events listed in Table 5.1. The different symbols indicate the detectors operating during the event, as shown in the legend, color coded by the chirp mass $\mathcal{M}$, in units of solar masses $M_\odot$ (see color bar). The horizontal axis is extended slightly beyond the maximum ($\mathcal{O}_{R,E} = 1$) to reduce symbol clutter at the right margin of the frame.

is no obvious trend relating $\ln \mathcal{B}_{R,E}$ and $\mathcal{O}_{R,E}$. Altogether, the $\ln \mathcal{B}_{R,E}$ versus $\mathcal{O}_{R,E}$ plot of the O3 events is similar to the non-precessing BBHs in the top left panel of Figure 5.3, viz. the O3 events are generally more consistent with the elliptically polarized signals characterized by E, compared to R. However, there are three events that stand out from the rest, namely GW200129_065458, GW190412 and GW190814, as indicated in Figure 5.5. We discuss them in further detail below.

GW200129_065458 is the only event in Table 5.1 with $\ln \mathcal{B}_{R,E} > 0$. Ref. [280] reports GW200129_065458 as the first CBC detection with spin-precession. However, the LIGO Livingston detector experienced an instrumental glitch around the time of detection. Ref. [98] shows, using *BayesWave*, that the uncertainties in glitch modelling reduces the evidence for spin-precession. Thus the presence of precession in GW200129_065458 remains ambiguous. The marginally positive Bayes factor ($\ln \mathcal{B}_{R,E} = 0.48 \pm 0.35$) is also insufficient to make claims about the preferred polarization of GW200129_065458. On the contrary, $\ln \mathcal{B}_{R,E} \approx -12$ for GW190412 and GW190814 are more negative than for other O3 events ($\ln \mathcal{B}_{R,E} \sim -5$) in Table 5.1. The color bar in Figure 5.5 indicates that both latter events have lower $\mathcal{M}$ than most of the other O3 events and therefore last longer. Table 5.1 shows that both E and R recover these events with high SNR (~ 20). Altogether E and R require more wavelets to characterize GW190412 and GW190814, compared to other shorter-duration O3 signals. Recall from Sections 5.3.1 and 5.3.2 that E and R have $D_E = 5N + 4$ and $D_R = 7N + 2$ model parameters respectively. Assuming equal N for both models, the difference in model dimensions between E and R, i.e. $D_R - D_E$, increases with N. Hence for GW190412 and GW190814, where N is large, we have $D_R \gg D_E$. Since the E and R reconstructions are also similar ($\mathcal{O}_{R,E} \approx 0.95$), the simpler model E is favoured over R. We also recognize that GW190814 and GW190412 show evidence of higher-order multipole moments [60]. The GW signal morphology in the presence of higher-order multipoles have been shown to deviate from quadrupole-only signals [281, 282]. The change in signal morphology may affect the reconstruction accuracy of E and R, and consequently the preferred model, an interesting avenue for future work.

LVK run	Event name	Detectors	Network SNR	$\mathcal{M}$ ($M_\odot$)	$\ln \mathcal{B}_{R,E}$	$\mathcal{O}_{R,E}$	Recovered SNR	
							E	R
O3a	GW190421_213856	HL	$10.7^{+0.2}_{-0.4}$	$31.4^{+6.0}_{-4.6}$	-2.27 ± 0.19	0.981	11.1	11.3
O3a	GW190521_074359	HL	$25.9^{+0.1}_{-0.2}$	$32.8^{+3.2}_{-2.8}$	-0.04 ± 0.20	0.998	25.8	25.8
O3a	GW190707_093326	HL	$13.1^{+0.2}_{-0.4}$	$8.4^{+0.6}_{-0.4}$	-4.32 ± 0.04	0.999	10.4	10.0
O3a	GW190408_181802	HLV	$14.6^{+0.2}_{-0.3}$	$18.5^{+1.9}_{-1.2}$	-3.08 ± 0.05	0.958	15.1	14.1
O3a	GW190412	HLV	$19.8^{+0.2}_{-0.3}$	$13.3^{+0.5}_{-0.5}$	-12.88 ± 0.06	0.952	19.1	18.2
O3a	GW190503_185404	HLV	$12.2^{+0.2}_{-0.4}$	$29.3^{+4.5}_{-4.4}$	-3.95 ± 0.04	0.982	12.0	11.4
O3a	GW190512_180714	HLV	$12.7^{+0.3}_{-0.4}$	$14.6^{+1.4}_{-0.9}$	-4.05 ± 0.04	0.991	11.2	11.2
O3a	GW190513_205428	HLV	$12.5^{+0.3}_{-0.4}$	$21.8^{+3.8}_{-2.2}$	-0.01 ± 0.05	0.940	12.8	12.9
O3a	GW190517_055101	HLV	$10.8^{+0.5}_{-0.6}$	$26.5^{+4.0}_{-4.2}$	-2.23 ± 0.04	0.982	10.0	9.7
O3a	GW190519_153544	HLV	$15.9^{+0.2}_{-0.3}$	$44.3^{+6.8}_{-7.5}$	-2.14 ± 0.13	1.000	15.4	15.6
O3a	GW190521	HLV	$14.3^{+0.5}_{-0.4}$	$63.3^{+19.6}_{-14.6}$	-1.93 ± 0.04	0.997	15.2	14.8
O3a	GW190602_175927	HLV	$13.2^{+0.2}_{-0.3}$	$48.0^{+9.5}_{-9.7}$	-2.28 ± 0.17	0.999	13.2	13.1
O3a	GW190706_222641	HLV	$13.4^{+0.2}_{-0.4}$	$45.6^{+13.0}_{-9.1}$	-1.03 ± 0.18	0.988	12.8	13.0
O3a	GW190727_060333	HLV	$11.7^{+0.2}_{-0.5}$	$29.4^{+4.6}_{-3.7}$	-1.63 ± 0.04	0.962	11.9	12.3
O3a	GW190728_064510	HLV	$13.1^{+0.3}_{-0.4}$	$8.6^{+0.6}_{-0.3}$	-2.65 ± 0.04	0.839	8.8	7.6
O3a	GW190814	HLV	$25.3^{+0.1}_{-0.2}$	$6.11^{+0.06}_{-0.05}$	-12.35 ± 0.32	0.950	19.5	19.2
O3a	GW190828_063405	HLV	$16.5^{+0.2}_{-0.3}$	$24.6^{+3.6}_{-2.0}$	-6.65 ± 0.05	0.963	17.0	15.6
O3a	GW190828_065509	HLV	$10.2^{+0.4}_{-0.5}$	$13.4^{+1.4}_{-1.0}$	-1.17 ± 0.12	0.999	6.5	6.3
O3a	GW190915_235702	HLV	$13.1^{+0.2}_{-0.3}$	$24.4^{+3.0}_{-2.3}$	-1.16 ± 0.04	0.990	12.9	13.6
O3a	GW190630_185205	LV	$16.4^{+0.2}_{-0.3}$	$25.1^{+2.2}_{-2.1}$	-1.72 ± 0.05	0.991	14.7	15.0
O3a	GW190708_232457	LV	$13.4^{+0.2}_{-0.3}$	$13.1^{+0.9}_{-0.6}$	-2.97 ± 0.21	0.918	11.8	10.7
O3a	GW190910_112807	LV	$14.5^{+0.2}_{-0.3}$	$33.5^{+4.2}_{-4.1}$	-1.26 ± 0.17	0.995	14.1	14.1
O3b	GW191109_010717	HL	$17.3^{+0.5}_{-0.5}$	$47.5^{+9.6}_{-7.5}$	-1.91 ± 0.05	0.998	17.5	17.5
O3b	GW191222_033537	HL	$12.5^{+0.2}_{-0.3}$	$33.8^{+7.1}_{-5.0}$	-2.23 ± 0.20	0.997	11.7	11.8
O3b	GW200225_060421	HL	$12.5^{+0.3}_{-0.4}$	$14.2^{+1.5}_{-1.4}$	-2.85 ± 0.19	0.997	12.6	12.4
O3b	GW191215_223052	HLV	$11.2^{+0.3}_{-0.4}$	$18.4^{+2.2}_{-1.7}$	-3.78 ± 0.04	0.993	9.9	9.6
O3b	GW200129_065458	HLV	$26.8^{+0.2}_{-0.2}$	$27.2^{+2.1}_{-2.3}$	0.48 ± 0.35	0.984	26.9	27.3
O3b	GW200208_130117	HLV	$10.8^{+0.3}_{-0.4}$	$27.7^{+3.7}_{-3.1}$	-3.02 ± 0.04	0.980	9.6	7.9
O3b	GW200219_094415	HLV	$10.7^{+0.3}_{-0.5}$	$27.6^{+5.6}_{-3.8}$	-4.30 ± 0.20	0.907	12.0	10.2
O3b	GW200224_222234	HLV	$20.0^{+0.2}_{-0.2}$	$31.1^{+3.3}_{-2.7}$	-3.74 ± 0.31	0.997	20.1	20.3
O3b	GW200311_115853	HLV	$17.8^{+0.2}_{-0.2}$	$26.6^{+2.4}_{-2.0}$	-3.41 ± 0.05	0.942	17.4	17.8
O3b	GW191216_213338	HV	$18.6^{+0.2}_{-0.2}$	$8.33^{+0.22}_{-0.19}$	-4.63 ± 0.39	0.992	17.7	16.4

Table 5.1: O3 GW events used in R versus E model selection. The columns from left to right quote: (i) the LVK observing run in which the event occurred, (ii) event name, (iii) detectors operating at time of detection, (iv) network matched-filter SNR, (v) chirp mass $\mathcal{M}$ in units of solar masses $M_\odot$, (v) log Bayes factor between R and E, $\ln \mathcal{B}_{R,E}$, (vi) network overlap between the R and E signal reconstructions, $\mathcal{O}_{R,E}$, and (vii) network SNR of the signals recovered by E and R. The network matched-filter SNR and $\mathcal{M}$ are copied directly from GWTC-2 [36] (O3a events) and GWTC-3 [38] (O3b events), and quote the median and the 90% symmetric credible intervals of the Bayesian posteriors, obtained from parameter estimation pipelines. $\ln \mathcal{B}_{R,E}$ and its error margins are calculated as per Equations 5.31 and 5.32. The network matched-filter SNR in this table is not to be confused with $\mathrm{SNR}_{\mathrm{net}}$ in the main text which denotes the injected network SNR for the simulated BBHs.

In summary, we perform model selection between E and R with O3 events. We find that E and R reconstructions of O3 events are comparable with $\mathcal{O}_{R,E} \gtrsim 0.90$. As with the non-precessing BBHs in Section 5.5.2, E is generally preferred over R, implying that the O3 GW signals are more consistent with elliptical polarizations than otherwise. The outlier GW200129_065458 reveals that $\ln \mathcal{B}_{R,E} > 0$ can indicate anomalous source properties, which justify further investigation with supplementary analysis pipelines.

5.6. Stokes parameters with R

The physical interpretation of ϵ in E (as in Equation 5.5) is unambiguous when the underlying signal is elliptically polarized and its polarization is constant with time. On the other hand, R can infer the polarization content of GW signals without constraints. In this section, we study how accurately R recovers polarization content for signals with different polarizations, using the non-precessing (elliptical) and precessing (non-elliptical) BBHs described in Section 5.4. We also study how the accuracy is affected by the size of detector networks using the HL, HLV and HLKV networks.

Ref. [2] explains how Stokes parameters describe the polarization content of GW signals. We define the fraction of linear ($F_{\rm L}$), circular ($F_{\rm C}$) and total ($F_{\rm T}$) polarizations relative to the total signal intensity in Section 5.6.1. In Section 5.6.2, we introduce a metric to quantify the accuracy between the injected and recovered $F_{\rm L}$, $F_{\rm C}$ and $F_{\rm T}$. We present the results for the non-precessing and precessing BBHs using the HL, HLV, and HLKV networks in Section 5.6.3.

5.6.1. Stokes parameters

The polarization content of a tensor-polarized GW is described by the four Stokes parameters [53, 283]

$$I = |\tilde{h}_+|^2 + |\tilde{h}_\times|^2 \tag{5.15}$$

$$Q = |\tilde{h}_+|^2 - |\tilde{h}_\times|^2 \tag{5.16}$$

$$U = \tilde{h}_+ \tilde{h}_\times^* + \tilde{h}_\times \tilde{h}_+^* \tag{5.17}$$

$$V = i(\tilde{h}_+ \tilde{h}_\times^* - \tilde{h}_\times \tilde{h}_+^*). \tag{5.18}$$

Equations 5.15 and 5.18 apply to monochromatic plane waves with linearly-independent complex amplitudes $\tilde{h}_+$ and $\tilde{h}_\times$. A GW signal, however, is typically modulated in amplitude and frequency, i.e. $\tilde{h}_+ = \tilde{h}_+(f)$ and $\tilde{h}_\times = \tilde{h}_\times(f)$, so its Stokes parameters decompose spectrally into functions $I(f)$, $Q(f)$, $U(f)$, and $V(f)$ of the Fourier frequency f. By definition, I is proportional to the total intensity; V describes circular polarization; Q and U together describe linear polarization. Therefore we can measure the polarization content of GWs through the fractional circular polarization

$$F_{\rm C}(f) = \frac{|V|}{I}, \tag{5.19}$$

fractional linear polarization

$$F_{\rm L}(f) = \frac{\sqrt{Q^2 + U^2}}{I}, \tag{5.20}$$

and the total polarization

$$F_{\rm T}(f) = \frac{\sqrt{Q^2 + U^2 + V^2}}{I}. \tag{5.21}$$

Detailed discussions of the theory underlying these quantities can be found in Appendix 5.A4. Going forward we refer to the above quantities collectively as the fractional polarizations $F_{\mathcal{P}}$, for $\mathcal{P} \in \{\mathrm{C}, \mathrm{L}, \mathrm{T}\}$. Stokes parameters are real numbers satisfying $I^2 \geq Q^2 + U^2 + V^2$, and $0 \leq F_{\mathcal{P}} \leq 1$.

5.6.2. *Accuracy metric: root mean squared residuals, $\mathcal{R}_{\rm RMS}$*

For E, the fractional polarizations $F_{\mathcal{P}}$ reduce to the following constants: $F_{\rm L} = (1 - \epsilon^2)/(1 + \epsilon^2)$, $F_{\rm C} = 2\epsilon/(1 + \epsilon^2)$ and $F_{\rm T} = 1$. For R, $F_{\mathcal{P}}$ are functions of frequency. We use the root mean squared (RMS) residuals between the injected and recovered $F_{\mathcal{P}}$, to quantify the measurement accuracy with R.

In *BayesWave*, the Fourier frequency f_i appearing as the argument of $F_{\mathcal{P}}(f_i)$ is discrete, with n intervals[7]. Therefore the root mean square (RMS) residuals between the injected (inj) and recovered (rec) $F_{\mathcal{P}}$ are given by

$$\mathcal{R}_{\rm RMS}(F_{\mathcal{P}}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\Big[F_{\mathcal{P},\rm rec}(f_i) - F_{\mathcal{P},\rm inj}(f_i)\Big]^2}, \tag{5.22}$$

with $0 \leq \mathcal{R}_{\rm RMS}(F_{\mathcal{P}}) \leq 1$. Lower $\mathcal{R}_{\rm RMS}(F_{\mathcal{P}})$ indicates better agreement between the injected and recovered $F_{\mathcal{P}}$, i.e. higher measurement accuracy, and vice versa. We demonstrate this in Figure 5.6, which shows the injected (colored dashed lines) and recovered (solid black line) $F_{\mathcal{P}}$ for a representative non-precessing (top panel) and a precessing (bottom panel) BBH from the injection sets described in Section 5.4. For each BBH, $F_{\rm L}$, $F_{\rm C}$ and $F_{\rm T}$ are plotted as functions of frequency in order, from top to bottom. To maximize the accuracy, i.e. to minimize $\mathcal{R}_{\rm RMS}(F_{\mathcal{P}})$, we only consider the frequency range $40\,\mathrm{Hz} \leq f \leq 400\,\mathrm{Hz}$ where the GW detector noise floor is lowest, as seen in Figure 5.1. The two contrasting examples in Figure 5.6 show that $F_{\mathcal{P}}$ recovered by HLKV (green) resembles more closely the injection (black) than HLV (red) and HL (purple). Therefore one expects $\mathcal{R}_{\rm RMS}(F_{\mathcal{P}})$ to be lowest for HLKV. Using $F_{\rm C}$ (middle row) of the precessing BBH (bottom panel) as an example, $\mathcal{R}_{\rm RMS}(F_{\rm C})$ of the HLKV, HLV and HL networks is given by 0.095, 0.321 and 0.810 respectively.

In Figure 5.6, we intentionally choose a precessing BBH with $\chi_{p,\rm init} = 0.77$, i.e. substantial precession, to showcase the difference in polarization content compared to a non-precessing BBHs, with $\chi_p = 0$ throughout the signal. For the non-precessing BBH (top panel), the $F_{\mathcal{P}}$ displays little or no frequency dependence. We find $F_{\rm C} > F_{\rm L}$ throughout the range $40\,\mathrm{Hz} \leq f \leq 400\,\mathrm{Hz}$, implying that the signal is more circularly polarized than linearly polarized. In contrast, for the precessing BBH example (bottom panel), $F_{\rm L}$ and $F_{\rm C}$ fluctuate with frequency, and the linear and circular polarizations dominate at different frequencies. We find $F_{\rm T} = 1$ for both the non-precessing and precessing examples, indicating that the signals are completely polarized at all f.

[7]For n evenly spaced samples, i.e. $\Delta f = f_{i+1} - f_i$ for all $i \in \{1, 2, \cdots n\}$, and for a specified frequency range $f_{\rm min} \leq f_i \leq f_{\rm max}$, one can write $n = (f_{\rm max} - f_{\rm min})/\Delta f$.

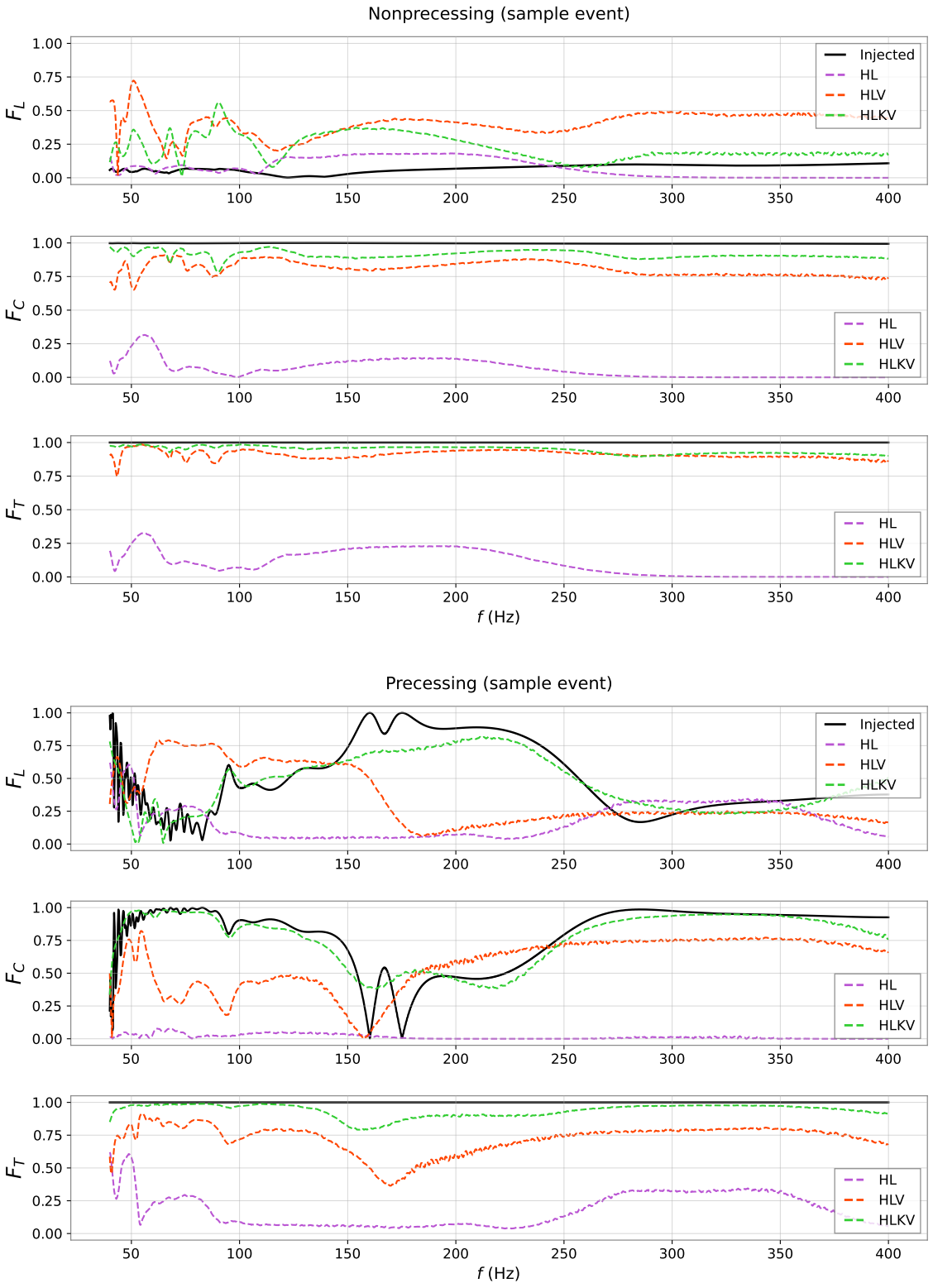

Figure 5.6: Recovering Stokes parameters. Top panel: injected and recovered fractional polarizations F_{L}, F_{C} and F_{T} (in order, from top to bottom) as a function of frequency $40\,\mathrm{Hz} \leq f \leq 400\,\mathrm{Hz}$, for a non-precessing BBH. Bottom panel: as for top panel, but for a precessing BBH with $\chi_{p,\mathrm{init}} = 0.77$. Injected $F_{\mathcal{P}}$ are graphed as black solid curves. Recovered $F_{\mathcal{P}}$ for HL, HLV, and HLKV are graphed as green, red and purple dashed curves respectively.

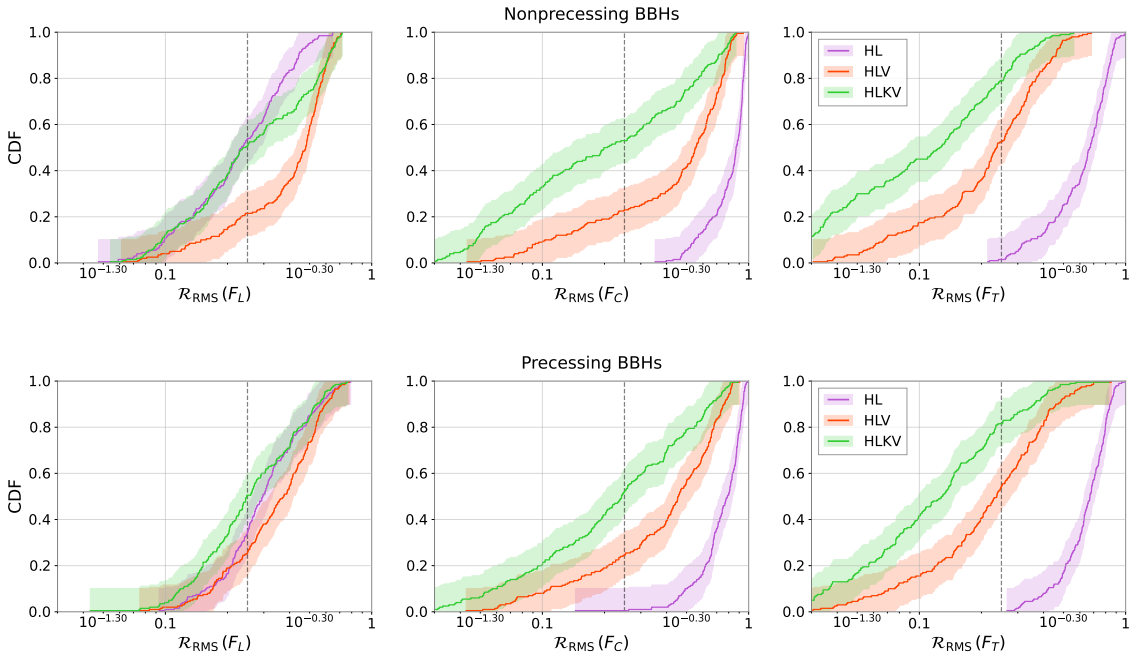

Figure 5.7: Overall perspective on accuracy: cumulative distribution functions (CDFs) of Stokes parameter root mean square residuals $\mathcal{R}_{\rm RMS}(F_{\mathcal{P}})$. The CDFs for $\mathcal{R}_{\rm RMS}(F_{\rm L})$, $\mathcal{R}_{\rm RMS}(F_{\rm C})$ and $\mathcal{R}_{\rm RMS}(F_{\rm T})$ are shown in the left, middle and right columns respectively. The top and bottom rows display the CDFs for non-precessing BBHs and precessing BBHs respectively. The CDFs for HL, HLV and HLKV are shown in purple, red and green respectively; the shaded region in corresponding colors brackets the 2-σ confidence interval (see Appendix 5.A5). The vertical dashed lines indicate $\mathcal{R}_{\rm RMS}(F_{\mathcal{P}}) = 0.25$. Higher curves imply higher measurement accuracy.

5.6.3. *Non-precessing and precessing BBH*

Figure 5.7 shows the cumulative distribution functions (CDFs) of $\mathcal{R}_{\rm RMS}(F_{\rm L})$, $\mathcal{R}_{\rm RMS}(F_{\rm C})$ and $\mathcal{R}_{\rm RMS}(F_{\rm T})$ in order from left to right. The top (bottom) row shows the CDFs for the non-precessing (precessing) BBHs, and CDFs for the HL, HLV and HLKV networks are shown in purple, red and green respectively. The shaded regions in corresponding colors indicate the 2-σ confidence intervals derived using the Dvoretzky-Kiefer-Wolfowitz inequality as described in Appendix 5.A5. We set a practical accuracy threshold of $\mathcal{R}_{\rm RMS}(F_{\mathcal{P}}) = 0.25$, and we only compare the CDFs for $\mathcal{R}_{\rm RMS}(F_{\mathcal{P}}) \leq 0.25$, i.e. to the left of the vertical dashed lines, in each panel of Figure 5.7. Higher CDF curves imply lower discrepancies between the injected and recovered $F_{\mathcal{P}}$, that is, $F_{\mathcal{P}}$ is recovered more accurately.

We first discuss the CDFs of $\mathcal{R}_{\rm RMS}(F_{\rm L})$, in the left column of Figure 5.7. For the non-precessing BBHs (top left panel), the HL CDF at $\mathcal{R}_{\rm RMS}(F_{\rm L}) \leq 0.25$ is 0.54. For precessing BBHs (bottom left panel), it is 0.34. In other words, the $F_{\rm L}$ measurement accuracy of HL is higher for non-precessing BBHs than for precessing BBHs. The contrast is due to the different linear polarization content, which we justify as follows. For a given polarization

fraction $F_{\mathcal{P}}$, we write its average across the frequency range $40\,\mathrm{Hz} \leq f_i \leq 400\,\mathrm{Hz}$ as

$$\overline{F_{\mathcal{P}}} = \frac{\sum_{i=1}^{n} F_{\mathcal{P},\mathrm{inj}}(f_i)}{n}. \tag{5.23}$$

The median $\overline{F_{\mathrm{L}}}$ for the non-precessing and precessing BBHs are 0.06 and 0.16 respectively. In Figure 5.6, one can see that $F_{\mathcal{P}}$ recovered by HL is generally close to zero, regardless of the injected $F_{\mathcal{P}}$. Since $\overline{F_{\mathrm{L}}}$ of the non-precessing BBHs is generally lower than the precessing BBHs, HL recovers F_{L} of non-precessing BBHs more accurately by coincidence. The HLV $\mathcal{R}_{\mathrm{RMS}}(F_{\mathrm{L}})$ CDFs, on the other hand, are quantitatively comparable between the non-precessing and precessing BBHs; at $\mathcal{R}_{\mathrm{RMS}}(F_{\mathrm{L}}) = 0.25$, the HLV CDFs are respectively 0.22 and 0.26 for the non-precessing and precessing BBHs. The same is true for HLKV, where the CDFs at $\mathcal{R}_{\mathrm{RMS}}(F_{\mathrm{L}}) = 0.25$ are the same (0.51) for both BBHs. These observations suggest that networks with three or more detectors measure low amplitude linear polarizations more consistently compared to a two-detector network, i.e. the measurement accuracy is not affected by signal morphology. Furthermore, the CDFs of HLKV (green) are higher than HLV (red) for all BBHs in the range $\mathcal{R}_{\mathrm{RMS}}(F_{\mathcal{P}}) \leq 0.25$, implying that larger detector networks recover F_{L} more accurately.

Next we discuss $\mathcal{R}_{\mathrm{RMS}}(F_{\mathrm{C}})$ in the middle column of Figure 5.7. For non-precessing BBHs (top middle panel), the CDF of HLKV is highest in the practically useful range $\mathcal{R}_{\mathrm{RMS}}(F_{\mathrm{C}}) \leq 0.25$, followed by HLV and then HL. The confidence intervals overlap marginally, suggesting that the differences in the CDFs are statistically significant. The same is true for precessing BBHs (bottom middle). At $\mathcal{R}_{\mathrm{RMS}}(F_{\mathrm{C}}) = 0.25$ the CDF of HL, HLV and HLKV is 0.0, 0.23, 0.53 respectively for the non-precessing BBHs and 0.01, 0.25 and 0.52 for the precessing BBHs. These values suggest that the CDFs of the non-precessing and precessing BBHs are quantitatively similar for the same detector networks. In other words, the polarization of the signal does not affect the measurement accuracy, if the same detector configuration is used. Additionally, the CDFs increase with the number of detectors, which suggests that larger detector networks recover F_{C} more accurately. We find that for HL, the $\mathcal{R}_{\mathrm{RMS}}(F_{\mathrm{C}})$ CDF is generally lower than $\mathcal{R}_{\mathrm{RMS}}(F_{\mathrm{L}})$, suggesting that HL recovers $\overline{F_{\mathrm{C}}}$ less accurately than $\overline{F_{\mathrm{L}}}$. The median $\overline{F_{\mathrm{C}}}$ of the non-precessing and precessing BBHs are 1.0 and 0.97 respectively, higher then $\overline{F_{\mathrm{L}}}$. This confirms that HL performs relatively poorly in recovering nontrivial fractional polarizations $F_{\mathcal{P}}$. The same arguments apply to $\mathcal{R}_{\mathrm{RMS}}(F_{\mathrm{T}})$, in the right column of Figure 5.7.

Altogether, we find that *BayesWave*'s signal model R recovers polarization content of GW bursts more accurately as the detector network expands. We also find that $F_{\mathcal{P}}$ of non-precessing and precessing BBHs are recovered with comparable accuracy using networks with three or more detectors, i.e. HLV and HLKV. This observation suggests that the underlying signal morphology (e.g. precessing or non-precessing) does not affect the performance of R in recovering polarization content, when the detector network is sufficiently large.

5.7. Conclusions

The *BayesWave* algorithm offers two tensor-polarized signal models: the elliptical polarization model E and the relaxed polarization model R. This paper studies two aspects

of their performance: (i) the signal characterization by R compared to E in Section 5.5, and (ii) the accuracy of R in measuring GW burst polarizations via Stokes parameters in Section 5.6. The expanding global detector network enables better signal characterization and hence more accurate measurements of GW polarizations, so we present multi-detector network comparisons between the HL (two-detector), HLV (three-detector) and HLKV (four-detector) networks to study how the network size impacts the performance of E and R.

Ref. [2] finds quanlitatively that R reconstructs GW burst signals with time-varying polarization content better than E, using precessing BBHs as an example. However, model selection does not depend entirely on reconstruction accuracy; R has more model parameters than E and incurs an Occam penalty in its Bayesian model evidence. In Section 5.5, we conduct a quantitative study of how the Bayes factor between R and E ($\ln \mathcal{B}_{R,E}$) is related to the discrepancy of reconstructed waveforms between R and E ($\mathcal{O}_{R,E}$). The analysis uses two sets of simulated BBHs, non-precessing and precessing, which produce elliptical and non-elliptical GW polarizations respectively. We find that E and R reconstruct GW signals from non-precessing BBHs comparably well and therefore E is preferred over R for its simplicity. We also note that $\mathcal{O}_{R,E}$ for non-precessing BBHs is generally higher with larger detector networks, because additional detectors enhances SNR and hence improves the reconstruction accuracy for both models. This results in a stronger preference for E in larger detector networks like HLV and HLKV. Similar trends apply for precessing BBH, but we find that R is preferred in some cases, especially for $\chi_{p,\mathrm{init}} \gtrsim 0.5$. High $\chi_{p,\mathrm{init}}$ indicates high initial precession, where the GW signals are more likely to possess time-varying (non-elliptical) polarization, whereupon they are better modeled by R. As a result, signal reconstructions of E and R generally show larger discrepancies for high $\chi_{p,\mathrm{init}}$ events, where E is less consistent with the data, and the preference for R increases. For the precessing BBHs, we also find that the preference for R increases with larger detector networks, suggesting that additional detectors enable R to characterize non-elliptical polarizations more accurately. Altogether the results imply empirically, that $\ln \mathcal{B}_{R,E}$ and $\mathcal{O}_{R,E}$ together indicate reliably when a GW signal deviates from elliptical polarization. However, the simulated BBHs are injected at unrealistically high SNR for the proof-of-concept study in Section 5.5.2. Therefore Section 5.5.3 repeats the study using 32 real GW events from O3. The study finds that E is generally preferred over R, suggesting perhaps that $\chi_{p,\mathrm{init}}$ is relatively low in most of the O3 events. We reiterate that $\ln \mathcal{B}_{R,E}$ and $\mathcal{O}_{R,E}$ cannot make definitive claims about the nature of the O3 events, but they can motivate follow-up analyses for events with potentially interesting properties.

Even though model selection generally prefers E over R, E cannot be used to measure the polarization content of generic GW bursts because it assumes elliptical polarization regardless of the true underlying polarization structure. R, on the other hand, is not restricted to a fixed polarization structure. In Section 5.6, we demonstrate how R measures generic polarization content, using the same non-precessing and precessing BBHs as above. We use the frequency-dependent fractional circular (F_{C}), fractional linear (F_{L}) and total (F_{T}) polarizations derived from the Stokes parameters to quantify polarization as a function of frequency f for polychromatic GW signals. For $F_{\mathcal{P}}$ with $\mathcal{P} \in \{\mathrm{C}, \mathrm{L}, \mathrm{T}\}$, we use the root mean squared residuals $\mathcal{R}_{\mathrm{RMS}}(F_{\mathcal{P}})$ to quantify the accuracy of R. As expected, we find that larger detector networks measure polarization more accurately;

the overall $\mathcal{R}_{\rm RMS}(F_{\mathcal{P}})$ CDFs of HLKV is higher than HLV by a factor of ~ 2. We also find the $\mathcal{R}_{\rm RMS}(F_{\mathcal{P}})$ CDFs for non-precessing and precessing BBHs are approximately equal, when one observes with either HLV or HLKV, i.e. the accuracy of polarization measurement is independent of signal morphology, when the detector network is sufficiently large. The inferior consistency and accuracy of the smaller HL network is attributed to degeneracies stemming from detector alignment. Avenues for future work include quantifying polarization measurements for other two-detector configurations, e.g. HV or LV, and comparing their performances to HLV and HLKV. We also recommend looking into reconstructions of unpolarized signals beyond BBHs, e.g. CCSNe and generic white noise bursts, to understand more fully the performance of R.

How do the results of this paper impact future tests of GR? Various analyses have been conducted to probe the polarization content of CBC signals by comparing GR models against alternative theories of gravity [66, 153–155, 284, 285]. Supplementary approaches are based on generalizing the parameterized post-Einsteinian [286–288] and null-stream frameworks[8] [292]. The null-stream approach does not rely on waveform templates and has been implemented as a consistency check when analyzing binary black hole (BBH) mergers in GWTC-2 and -3 [154, 155]. However, this method requires exact knowledge of the source location and forfeits phase information [154]. By further generalizing R to allow up to six independent polarization modes, *BayesWave* offers a model-independent framework to probe polarizations of GW bursts without the above limitations [156]. In principle, one can compare the Bayes factor between alternative-theory (beyond-GR) and GR polarization models. A first demonstration of this approach is presented in Ref. [56] for GW150914, where *BayesWave* is used to compare signal models with purely tensor versus purely scalar polarization; however, this represents a highly simplified and idealized comparison of purely GR versus purely non-GR polarization. Given the strong evidence for the existence of tensor modes [153–155], it is more informative to compare signal models with mixed polarizations (e.g. tensor plus scalar for beyond-GR) against pure-tensor (GR) models. Model evidences, however, are sensitive to the prior volume: one can arbitrarily increase the prior range of a model to decrease preference for the model, or vice versa. *BayesWave* generally uses flat priors for both the intrinsic and extrinsic model parameters [2]. Since R and E are both tensor-polarized, i.e. they both have exactly two polarization modes, they share similar parameters and hence similar priors. In contrast, if both the mixed-polarization and pure-tensor models use flat priors, the larger prior volume of the mixed-polarization models with more than two polarizations attracts a larger Occam penalty, biasing the model selection towards the pure-tensor models. As the global detector expands and more GWs are detected, the priors of nontensorial polarizations are expected to become better constrained. Therefore the methods presented in Section 5.5 of this paper can be adapted to distinguish between pure-tensor (GR) and mixed-polarizations (beyond-GR) models in the future.

[8] Null streams are linear combinations of detector outputs which are insensitive to tensor polarizations for a specified sky location. Null-stream fluctuations obey a chi-squared distribution, if the sky location is known exactly [289–291].

Acknowledgements

This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation. Parts of this research were conducted by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), through project numbers CE170100004 and CE230100016. The authors are grateful for computational resources provided by the LIGO Laboratory and supported by National Science Foundation Grants PHY-0757058 and PHY-0823459. Y. S. C. L. is supported by a Melbourne Research Scholarship and the Jean Laby Women in Physics Travel Award. M. M. gratefully acknowledges the NSF for support from Grant PHY-2110481. S. D. was supported by the Letson Summer Internship Award. The authors also thank Katerina Chatziioannou, Neil Cornish and Megan Arogeti for their helpful comments.

5.A1. Bayesian model evidence via thermodynamic integration

BayesWave uses parallel tempering [131] to improve convergence of the RJMCMC to the target (posterior) distribution. A parallel-tempered MCMC allows multiple Markov chains with different 'temperatures' (T) to run in parallel, while occasionally letting the chains exchange positions to improve coverage of the parameter space. Let $p(\mathbf{d}|\boldsymbol{\theta},\mathcal{H})$ denote the likelihood of the data $\mathbf{d}$ for a given model $\mathcal{H}$ parameterized by $\boldsymbol{\theta}$. The temperature modifies the likelihood according to $p(\mathbf{d}|\boldsymbol{\theta},\mathcal{H}) \mapsto p(\mathbf{d}|\boldsymbol{\theta},\mathcal{H})^{1/T}$, viz. high T chains survey more of the prior volume, whereas low T chains explore regions surrounding the posterior distribution. In addition to improving convergence, the chains can be used to directly calculate the Bayesian model evidence using thermodynamic integration [136]. The procedure is detailed below.

The model evidence is the likelihood of producing the data $\mathbf{d}$ given model $\mathcal{H}$ parameterized by $\boldsymbol{\theta}$, which by definition is the likelihood of $\mathcal{H}$ marginalized over the domain of $\boldsymbol{\theta}$, i.e.

$$p(\mathbf{d}|\mathcal{H}) = \int d\boldsymbol{\theta}\; p(\boldsymbol{\theta}|\mathcal{H}) p(\mathbf{d}|\boldsymbol{\theta},\mathcal{H}). \tag{5.24}$$

$p(\boldsymbol{\theta}|\mathcal{H})$ is the prior distribution of $\boldsymbol{\theta}$. By analogy, one can compute the evidence for $\mathcal{H}$ at a given temperature $T = 1/\beta$:

$$Z(\beta) = \int d\boldsymbol{\theta}\; p(\boldsymbol{\theta}|\mathcal{H}) p(\mathbf{d}|\boldsymbol{\theta},\mathcal{H})^{\beta}, \tag{5.25}$$

with $0 \leq \beta \leq 1$. $Z(\beta)$ is equivalent to a partition function in a physical system described by thermodynamic variables. Since the prior is independent of β, Equation 5.25 can be rewritten as

$$\frac{d}{d\beta} \ln Z(\beta) = \int d\boldsymbol{\theta} \left[\frac{p(\boldsymbol{\theta}|\mathcal{H}) p(\mathbf{d}|\boldsymbol{\theta},\mathcal{H})^{\beta}}{Z(\beta)} \right] \ln p(\mathbf{d}|\boldsymbol{\theta},\mathcal{H}), \tag{5.26}$$

using the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}\beta} \ln f(\beta) = \frac{1}{f(\beta)} \frac{\mathrm{d}f}{\mathrm{d}\beta}. \tag{5.27}$$

According to Bayes' Theorem, the terms enclosed within the square brackets in Equation 5.26 collectively represent the posterior probability $p_\beta(\boldsymbol{\theta}|\mathbf{d},\mathcal{H})$ of $\boldsymbol{\theta}$ for the chain at temperature $T = 1/\beta$. Therefore the right-hand side of Equation 5.26 is, by definition, the expectation value of the log likelihood $\langle \ln p(\mathbf{d}|\boldsymbol{\theta},\mathcal{H})\rangle_\beta$ for the corresponding β. This quantity can be evaluated directly from the RJMCMC chain [293]:

$$\langle \ln p(\mathbf{d}|\boldsymbol{\theta},\mathcal{H})\rangle_\beta = \frac{1}{\mu}\sum_{i=1}^{\mu} \ln p(\mathbf{d}|\boldsymbol{\theta}_{i,\beta},\mathcal{H}). \tag{5.28}$$

where μ denotes the number of samples in the chain after the burn-in period and $\{\boldsymbol{\theta}_{i,\beta}\}$ represents the set of samples for the $T = 1/\beta$, for $i \in \{1, 2, \cdots, \mu\}$.

The model evidence $p(\mathbf{d}|\mathcal{H})$ can be interpreted similarly to the Helmholtz free energy of a canonical physical ensemble [294], i.e. the logarithm of the partition function $\ln Z(\beta)$. This can be obtained by integrating Equation 5.26 over β:

$$\ln p(\mathbf{d}|\mathcal{H}) = \ln Z(\beta) \tag{5.29}$$

$$= \int_0^1 d\beta\ \langle \ln p(\mathbf{d}|\boldsymbol{\theta},\mathcal{H})\rangle_\beta . \tag{5.30}$$

Due to the finite number of parallel tempering chains, the integral 5.30 is evaluated as a discrete sum in practice and is subject to discretization error. Additionally, $\langle \ln p(\mathbf{d}|\boldsymbol{\theta},\mathcal{H})\rangle_\beta$ is estimated using the RJMCMC marginalization of model parameters and therefore contain statistical errors. It is, however, challenging to estimate these errors analytically. In *BayesWave*, the integral of Equation 5.30 is computed at each RJMCMC iteration, which produces a posterior distribution of $\ln p(\mathbf{d}|\mathcal{H})$. The expectation value $E[\ln p(\mathbf{d}|\mathcal{H})]$ and variance $V[\ln p(\mathbf{d}|\mathcal{H})]$ of the posterior provide the central value and error estimate of $\ln p(\mathbf{d}|\mathcal{H})$ respectively. That is, the log Bayes factor between two models $\mathcal{H}_1$ and $\mathcal{H}_2$ is given by

$$\ln \mathcal{B}_{1,2} = E[\ln p(\mathbf{d}|\mathcal{H}_1)] - E[\ln p(\mathbf{d}|\mathcal{H}_2)] \tag{5.31}$$

and has an error margin

$$\Delta \ln \mathcal{B}_{1,2} = \sqrt{V[\ln p(\mathbf{d}|\mathcal{H}_1)] + V[\ln p(\mathbf{d}|\mathcal{H}_2)]}. \tag{5.32}$$

5.A2. Precession spin parameter, χ_p

BBHs are generally characterized by eight intrinsic physical parameters: the component masses (m_1, m_2) and the spin angular momentum vectors ($\boldsymbol{S}_1$, $\boldsymbol{S}_2$) [40]. The indices 1 and 2 denote the primary and secondary black holes respectively. For brevity, we write $\boldsymbol{S}_\mathcal{C}$ for $\mathcal{C} = 1, 2$ when discussing properties that are relevant to both $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$. $\boldsymbol{S}_\mathcal{C}$ can be decomposed into spin components that are parallel ($\boldsymbol{S}_{\mathcal{C},\parallel}$) and orthogonal ($\boldsymbol{S}_{\mathcal{C},\perp}$) to the binary orbital angular momentum vector $\boldsymbol{L}$:

$$\boldsymbol{S}_\mathcal{C} = \boldsymbol{S}_{\mathcal{C},\parallel} + \boldsymbol{S}_{\mathcal{C},\perp}. \tag{5.33}$$

$\boldsymbol{S}_{\mathcal{C},\perp}$ is otherwise known as the in-plane spin and can be used to approximate the effective precession of a binary system through a single precession spin parameter χ_p [269]. We summarize its derivation as follows.

The leading order of the post-Newtonian precession equation is given by [295]

$$\dot{\boldsymbol{L}} = \frac{1}{r^2}(A_1\boldsymbol{S}_1 + A_1\boldsymbol{S}_2)\times\boldsymbol{L}, \tag{5.34}$$

with $A_1 = 2 + 3/(2q)$, $A_2 = 2 + 3q/2$, and $q = m_1/m_2 \geq 1$, where r denotes the orbital separation. Since one has $\boldsymbol{S}_{c,\parallel}\times\boldsymbol{L} = 0$ by definition, we can rewrite Equation 5.34 as

$$\dot{\boldsymbol{L}} = \frac{1}{r^2}(A_1\boldsymbol{S}_{1,\perp} + A_1\boldsymbol{S}_{2,\perp})\times\boldsymbol{L}, \tag{5.35}$$

to indicate explicitly that the time evolution of $\boldsymbol{L}$ in a precessing system is only driven by $\boldsymbol{S}_{\mathcal{C},\perp}$. In a precessing system, the angle between $\boldsymbol{S}_{\mathcal{C},\perp}$ and $\boldsymbol{L}$ as well as its magnitude $|\boldsymbol{S}_{\mathcal{C},\perp}| = S_{\mathcal{C},\perp}$ change with time. Figure 5 of Ref. [269] shows that the magnitude $|A_1\boldsymbol{S}_{1,\perp} + A_1\boldsymbol{S}_{2,\perp}|$ of Equation 5.35 oscillates about a mean value which is consistent over multiple precession cycles. The oscillation amplitudes is small, so precise modelling of this quantity is not necessary for accurate description of the waveform. Instead, the overall in-plane spin magnitudes in Equation 5.35 can be approximated by a single parameter

$$S_p = \frac{1}{2}\left[(A_1 S_{1,\perp} + A_2 S_{2,\perp}) + \left|A_1 S_{1,\perp} - A_2 S_{2,\perp}\right|\right] \tag{5.36}$$

$$= \max(A_1 S_{1,\perp}, A_2 S_{2,\perp}), \tag{5.37}$$

which effectively represents the mean oscillation amplitude of Equation 5.35. By definition, S_p is the average between the maximum and minimum in-plane spin contributions of $\boldsymbol{S}_{1,\perp}$ and $\boldsymbol{S}_{2,\perp}$, that is when they are parallel and anti-parallel respectively.

The spin angular momentum $\boldsymbol{S}_{\mathcal{C}}$ is conventionally referenced using its dimensionless counterpart $\boldsymbol{\chi}_{\mathcal{C}} = \boldsymbol{S}_{\mathcal{C}}/m_{\mathcal{C}}^2$, with $0 \leq |\boldsymbol{\chi}_{\mathcal{C}}| \leq 1$. By analogy one can define a dimensionless precession spin parameter

$$\chi_p = \frac{S_p}{A_1 m_1^2}. \tag{5.38}$$

Note that the primary black hole appears in the denominator of Equation 5.38. This is done because S_p typically reduces to $A_1 S_{1,\perp}$, because the primary black hole spin tends to dominate, as q increases [276]. For an interpretation relevant to similar-mass binaries, we refer the reader to Section III of Ref. [269] for further details.

5.A3. Statistical stability of *BayesWave* model evidences

By default *BayesWave* uses 20 chains for parallel tempering as described in Appendix 5.A1. Previous studies have shown that 20 chains are sufficient to achieve RJMCMC convergence, thereby yielding statistically stable estimates of the signal model evidence $\ln p(\boldsymbol{d}|\mathcal{S})$ for events with $\mathrm{SNR}_{\mathrm{net}} \sim 20$ [1]. However, this is not the case for the simulated BBH signals described in Section 5.4, with $\mathrm{SNR}_{\mathrm{net}}{\sim}50$. The high SNR of our injections increases the number of wavelets and the parameter space volume required to fit the data, so using 20 chains is inadequate for the convergence of the RJMCMC to the target distributions. Consequently, we obtain unreliable estimates of $\ln p(\boldsymbol{d}|\mathcal{S})$, with

large error margins $\Delta \ln p(\boldsymbol{d}|\mathcal{S}) \sim 10$. To determine appropriate convergence settings for our high-SNR injections, we experiment with various combination of *BayesWave* settings and compare their performances in evaluating $\ln p(\boldsymbol{d}|\mathcal{S})$. The settings are selected based on: (i) computational efficiency, and (ii) reliability, i.e. statistically stable estimates of $\ln p(\boldsymbol{d}|\mathcal{S})$ with minimal $\Delta \ln p(\boldsymbol{d}|\mathcal{S})$. We present the methods and results of this supplementary study below.

A conventional approach to test the convergence, and hence the reliability, of stochastic samplers is to repeat the analysis with multiple sampling seeds and ensure that the outcomes are consistent. In this study, we vary the seed of the RJMCMC sampler to assess the consistency of the $\ln p(\boldsymbol{d}|\mathcal{S})$ computed by *BayesWave*, using different analysis settings. The dataset consists of six randomly selected HLKV injections. Since HLKV is the largest detector configuration used in our polarization study, the injected $\mathrm{SNR}_{\mathrm{net}}$ is higher compared to HL and HLV. Therefore, if a setting ensures sampler convergence for HLKV injections, it should also guarantee convergence for HL and HLV injections. To conserve computational resources, we conduct the convergence study using only E and assume that the same convergence settings apply to R.

The number of chains and the number of RJMCMC iterations are two key factors that affect sampler convergence. Therefore, for each HLKV event, we perform the *BayesWave* analysis using five different chain/iteration settings[9]: (i) 50 chains/4-million iterations, (ii) 50 chains/8-million iterations, (iii) 60 chains/4-million iterations, (iv) 60 chains/8-million iterations and (v) 60 chains/10-million iterations. For each setting, we repeat the analysis with 11 different RJMCMC sampling seeds, taking integer values from 1234 to 1244, inclusive. In total, we perform $11 \times 5 = 55$ *BayesWave* analyses per event.

We present the results for all six events in Figure 5.8. In each event panel, we plot the signal evidence estimates for the E model ($\ln p(\boldsymbol{d}|E)$; top panel) and the corresponding error estimate ($\Delta \ln p(\boldsymbol{d}|E)$; bottom panel) as a function of the chain seed. Results for the five different *BayesWave* settings are distinguished by color. An indication of RJMCMC convergence is when the $\ln p(\boldsymbol{d}|E)$ is statistical stable, i.e. independent of the sampling seed. The vertical axis in the top panels shows the relative deviation from the median $\ln p(\boldsymbol{d}|E)$, calculated across the 55 *BayesWave* analyses of the corresponding event; the median value is displayed in the top-left corner of the plot. The 50-chain analyses (blue and purple) show frequent (2–3 per event) and large ($\gtrsim$20) deviations from the median $\ln p(\boldsymbol{d}|E)$, c.f. the 60-chain analyses (orange, red and green), generally yield uniform $\ln p(\boldsymbol{d}|E)$ across different seeds, with at most one noticeable deviation ($\lesssim$10) from the median per event. This suggests that $\ln p(\boldsymbol{d}|E)$ estimates with the 60-chain analyses are more statistically stable compared to the 50-chain analyses. Among the three 60-chain analyses, the 10-million iterations (green) analyses show the smallest the deviations from the median $\ln p(\boldsymbol{d}|E)$, followed by the 8-million iterations (red), and then the 4-million iterations (orange). While the deviations are minimized, the 8- and 10-million iteration analyses take approximately two and four times longer to run than the 4-million iteration analyses ($\sim$1 day), respectively. Therefore, we establish that the 60 chains/4-million iterations setting is the most efficient for obtaining statistically stable $\ln p(\boldsymbol{d}|E)$ estimates, within ± 5 of the expected median.

[9]The number of wavelets N, and thus the parameter space volumes of the signal models ($5N + 4$ for E and $7N + 4$ for R), scales linearly with the SNR. Therefore, we choose to experiment with 50 and 60 chains for injections with $\mathrm{SNR}_{\mathrm{net}} \sim 50$, c.f. 20 chains for $\mathrm{SNR}_{\mathrm{net}} \sim 20$.

Although we have identified a reasonable compromise, the occasional deviations in $\ln p(\boldsymbol{d}|E)$ estimates, i.e. non-convergence of the RJMCMC, must be addressed systematically. In Figure 5.8, the noticeable deviations of $\ln p(\boldsymbol{d}|E)$ from the median are associated with elevated $\Delta \ln p(\boldsymbol{d}|E)$. Furthermore, higher $\Delta \ln p(\boldsymbol{d}|E)$ generally leads to larger deviations in $\ln p(\boldsymbol{d}|E)$. Thus, $\Delta \ln p(\boldsymbol{d}|E)$ can be used as a measure of reliability. To demonstrate this in practice, we define a nominal threshold at five times the median $\Delta \ln p(\boldsymbol{d}|E)$ of the 60 chains/4-million iteration analyses across the 11 chain seeds, as indicated by the gray horizontal dashed lines in Figure 5.8. For the runs with $\Delta \ln p(\boldsymbol{d}|\mathcal{S})$ below the threshold, we observe little to no deviation from the median $\ln p(\boldsymbol{d}|E)$. The converse is true for runs with $\Delta \ln p(\boldsymbol{d}|E)$ above the threshold. Moreover, the thresholds are similar across the six events, suggesting that we can set a global $\Delta \ln p(\boldsymbol{d}|E)$ threshold applicable to all events, rather than one specific to each event. For each BBH injection set described in Section 5.4, we define two thresholds: one for the E analysis, and another for R. The thresholds are set to five times the median $\Delta \ln p(\boldsymbol{d}|\mathcal{S})$ across the 200 injections, for the corresponding analysis $\mathcal{S} \in \{E, R\}$. For events that exceed the threshold ($\sim$20 events per injection set, per detector configuration), we repeat the analysis with more iterations (8 million). All repeat analyses obtain reliable $\ln p(\boldsymbol{d}|E)$ estimates with this setting.

In summary, we find that using a combination of 60 chains and four million RJMCMC iterations is the most computationally efficient way to obtain reliable model evidences for $\mathrm{SNR}_{\mathrm{net}} \sim 50$ signals. We can also use the evidence error estimates of the overall event population as a nominal reliability benchmark.

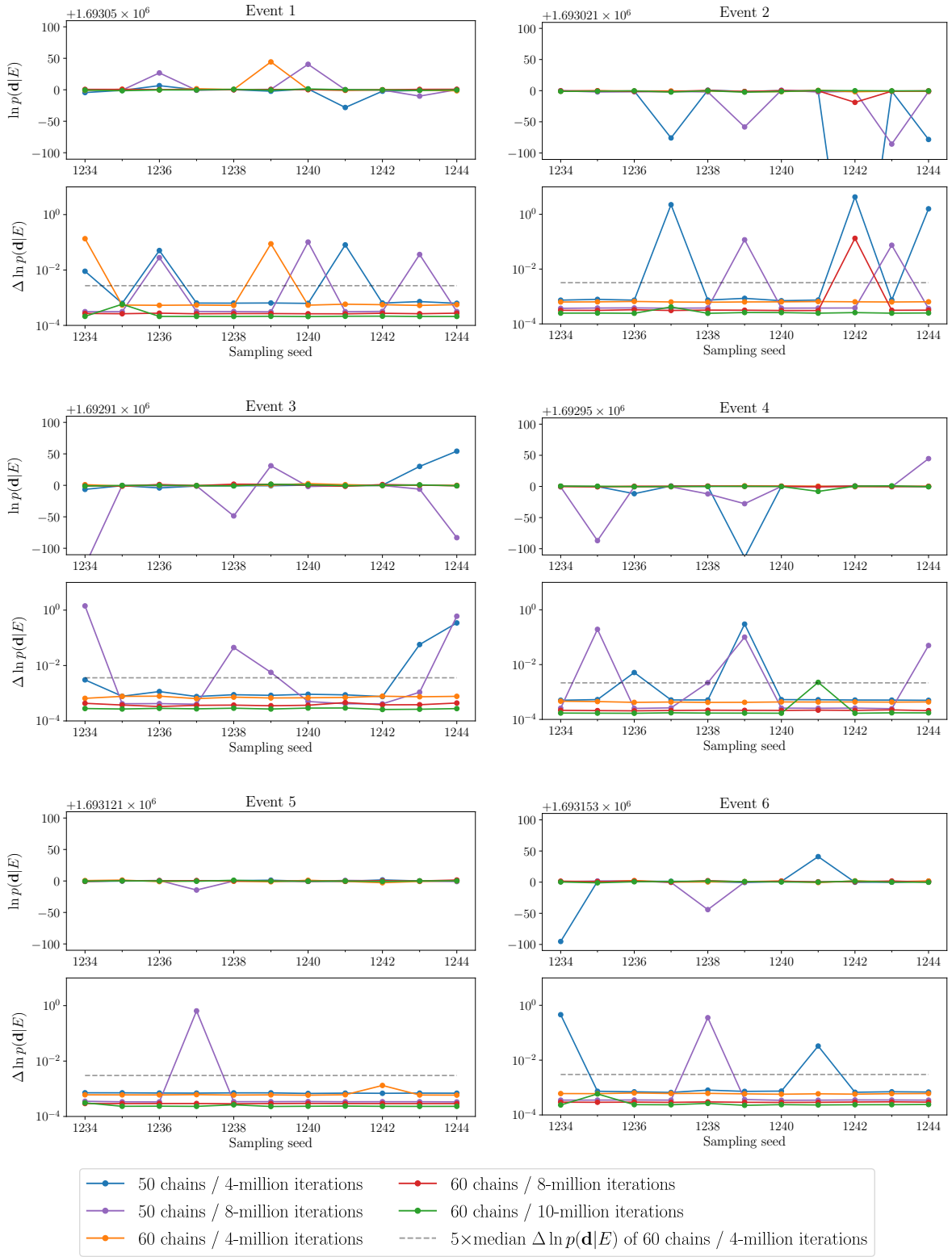

Figure 5.8: Statistical stability tests of *BayesWave*'s E model evidence estimate, using six randomly selected HLKV injections. In each event panel, the top plot shows the signal evidence, $\ln p(\boldsymbol{d}|E)$ as a function of the Markov chain seed used in the RJMCMC sampling; the bottom plot shows the same but for the corresponding error estimate, $\Delta \ln p(\boldsymbol{d}|E)$. The different colors represent the results from different run settings, as indicated by the legend. The gray dashed lines in the bottom plots indicate the $\Delta \ln p(\boldsymbol{d}|E)$ threshold above which the signal evidence is deemed unreliable. This threshold is placed arbitrarily at five times the median $\Delta \ln p(\boldsymbol{d}|E)$ of the 60 chains/4-million iteration analyses (orange).

5.A4. Theory of Stokes parameters

Stokes parameters are familiar in the context of electrodynamics [296]. The set of four quantities provide a convenient and complete description for the state of polarization in terms of the total wave intensity, fractional (degree of) polarization and the properties of the polarization ellipse. For completeness, we define the Stokes parameters from first principle for electromagnetic (EM) waves, and then generalize it to GW applications.

First, we consider two linearly independent and monochromatic EM plane waves with frequency ω, given by

$$\mathbf{E}_x = \boldsymbol{\epsilon}_x E_x e^{i\mathbf{k}\cdot\mathbf{x}-i\omega t} \tag{5.39}$$

$$\mathbf{E}_y = \boldsymbol{\epsilon}_y E_y e^{i\mathbf{k}\cdot\mathbf{x}-i\omega t}. \tag{5.40}$$

E_x and E_y are complex amplitudes, $\mathbf{k}$ is the wave vector, and $\boldsymbol{\epsilon}_x$ and $\boldsymbol{\epsilon}_y$ are linearly independent polarization vectors. One can then obtain a general homogeneous plane wave

$$\mathbf{E}(\mathbf{x}, t) = (\boldsymbol{\epsilon}_x E_x + \boldsymbol{\epsilon}_y E_y) e^{i\mathbf{k}\cdot\mathbf{x}-i\omega t} \tag{5.41}$$

propagating in the direction $\mathbf{k} = k\mathbf{n}$. The complex amplitudes E_x and E_y can be decomposed into:

$$E_x = A_x e^{ib_x} \tag{5.42}$$

$$E_y = A_y e^{ib_y}, \tag{5.43}$$

where A and b denote the magnitude and phase, respectively. When $b_x = b_y$, that is when E_x and E_y have the same phase, Equation 5.41 describes a linearly polarized wave, with the polarization vector forming an angle of $\tan^{-1}(E_y/E_x)$ with respect to $\boldsymbol{\epsilon}_x$. Conversely when $b_x \neq b_y$, the wave described by Equation 5.41 is elliptically polarized. In the special case where $A_x = A_y$ and $b_x = b_y \pm \pi/2$, the polarization becomes circular. Altogether, the polarization content of Equation 5.41 can be described fully by four real quantities $\{A_x, A_y, b_x, b_y\}$ and can thus be encoded as a 2×2 Hermitian polarization matrix:

$$J_{nm} = \mathcal{E}_n \mathcal{E}_m^*, \tag{5.44}$$

where $\mathcal{E}_\mu = |\boldsymbol{\epsilon}_\mu \cdot \mathbf{E}| = E_\mu e^{i\mathbf{k}\cdot\mathbf{x}-i\omega t}$ and $\mathcal{E}_\mu^*$ is the complex conjugate of $\mathcal{E}_\mu$. By the definition of Hermitian matrices, the polarization matrix can be expressed as a linear combination of the Pauli matrices[10] $\{\sigma_1, \sigma_2, \sigma_3\}$:

$$J_{nm} = \frac{1}{2} I \delta_{nm} + \frac{1}{2} (U\sigma_1 + V\sigma_2 + Q\sigma_3)_{nm} \tag{5.45}$$

$$= \frac{1}{2} \begin{pmatrix} I+Q & U-iV \\ U+iV & I-Q \end{pmatrix}, \tag{5.46}$$

[10] The Pauli matrices are given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

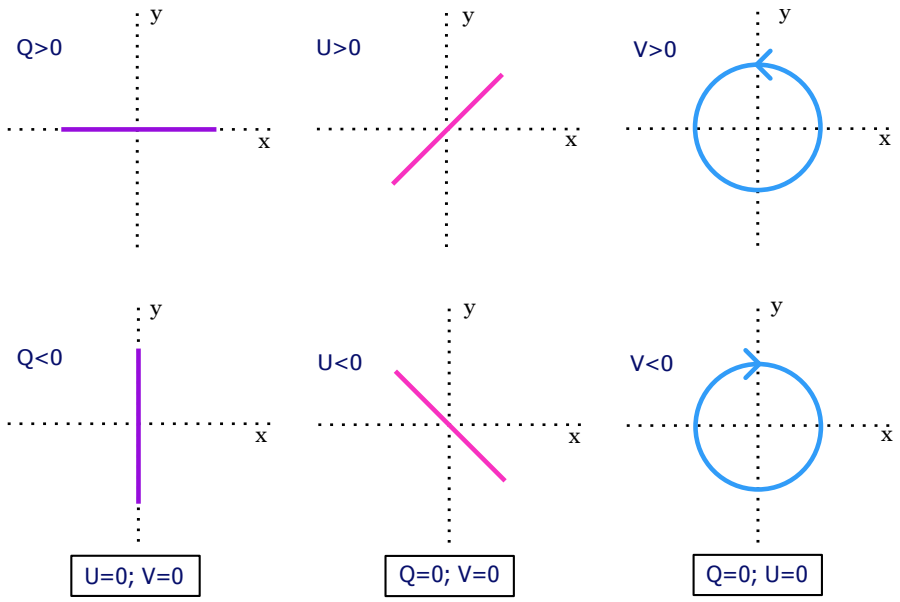

Figure 5.9: Visualizing polarization states; each locus represents the trajectory traced by the tip of the electric field vector across one wave period. The left column shows polarizations corresponding to $Q \neq 0$ and $U = V = 0$ in purple. The middle column shows polarizations corresponding to $U \neq 0$ and $Q = V = 0$ in pink. The right column shows polarizations corresponding to $V \neq 0$ and $U = V = 0$ in blue.

where δ_{nm} is the Kronecker delta. I, Q, U, V are real coefficients and are collectively referred to as the Stokes parameters: [283]

$$I = |\mathcal{E}_x|^2 + \left|\mathcal{E}_y\right|^2 = |E_x|^2 + \left|E_y\right|^2 \tag{5.47}$$

$$Q = |\mathcal{E}_x|^2 - \left|\mathcal{E}_y\right|^2 = |E_x|^2 - \left|E_y\right|^2 \tag{5.48}$$

$$U = 2\,\mathrm{Re}[\mathcal{E}_x^* \mathcal{E}_y] = E_x E_y^* + E_y E_x^* \tag{5.49}$$

$$V = 2\,\mathrm{Im}[\mathcal{E}_x^* \mathcal{E}_y] = i(E_x E_y^* - E_y E_x^*). \tag{5.50}$$

Alternatively, in terms of A and b:

$$I = A_x^2 + A_y^2 \tag{5.51}$$

$$Q = A_x^2 - A_y^2 \tag{5.52}$$

$$U = 2A_x A_y \cos\left(b_y - b_x\right) \tag{5.53}$$

$$V = 2A_x A_y \sin\left(b_y - b_x\right). \tag{5.54}$$

Physically, I is proportional to the intensity of Equation 5.41 and is always positive. In contrast, Q, U and V, can take both positive or negative values, with the signs conveying physical information about the polarization states. As discussed above, $\mathbf{E}$ is linearly polarized when $b_x = b_y$; in the special case where we also have $A_x = A_y$ ($A_x = -A_y$), $\mathbf{E}$ is linearly polarized at an angle of $\pi/4$ ($3\pi/4$) with respect to $\boldsymbol{\epsilon}_x$ i.e. the x axis, and the Stokes parameters are $U > 0$ ($U < 0$) and $Q = V = 0$. This is visualized in the top-middle

(bottom-middle) panel of Figure 5.9. When $A_x \neq 0$ and $A_y = 0$ ($A_y \neq 0$ and $A_x = 0$), **E** is linearly polarized along the x (y) axis, with $Q \neq 0$ and $U = V = 0$. This is visualized in the top-left (bottom-left) panel of Figure 5.9). With this, we establish that Q and U characterize linear polarizations. On the other hand, V describes circular polarization, which we justify as follows. When $A_x = A_y$ and $b_x = b_y \pm \pi/2$, we obtain $\mathcal{E}_x = \pm i\mathcal{E}_y$. That is, $\boldsymbol{E}$ traces a circular path in the plane transverse to the direction of propagation. The Stokes parameters for $b_x = b_y + \pi/2$ are $V > 0$ and $Q = U = 0$. An observer viewing the incoming wave face on will see an anticlockwise electric field as depicted in the top right panel of Figure 5.9. Conversely, for $b_x = b_y - \pi/2$, the electric field will rotate in a clockwise direction, with $V < 0$ and $Q = U = 0$, as shown in the bottom right panel of Figure 5.9. Equations 5.48-5.50 also imply that

$$I^2 \geq Q^2 + U^2 + V^2. \tag{5.55}$$

In other words, the total polarized intensity must not exceed the total intensity of the wave. If the wave is unpolarized, the Stokes parameters are $Q = U = V = 0$.

By analogy to EM waves, GW metric perturbations $\tilde{h}_{\mu\nu}$ in the frequency-domain can be written as a linear combination of polarization basis tensors. Tensor-polarized GWs have two basis tensors: $e^{+}_{\mu\nu}$ and $e^{\times}_{\mu\nu}$, hence the metric perturbation is given by

$$\tilde{h}_{\mu\nu}(f) = \tilde{h}_{+}(f)e^{+}_{\mu\nu} + \tilde{h}_{\times}(f)e^{\times}_{\mu\nu}. \tag{5.56}$$

Since the polarization amplitudes $\tilde{h}_{+}(f)$ and $\tilde{h}_{\times}(f)$ are analogous to the complex electric field amplitudes E_x and E_y, the Stokes parameters for tensor-polarized GW are given by:

$$I(f) = |\tilde{h}_{+}|^2 + |\tilde{h}_{\times}|^2 \tag{5.57}$$

$$Q(f) = |\tilde{h}_{+}|^2 - |\tilde{h}_{\times}|^2 \tag{5.58}$$

$$U(f) = \tilde{h}_{+}\tilde{h}^{*}_{\times} + \tilde{h}_{\times}\tilde{h}^{*}_{+} \tag{5.59}$$

$$V(f) = i(\tilde{h}_{+}\tilde{h}^{*}_{\times} - \tilde{h}_{\times}\tilde{h}^{*}_{+}), \tag{5.60}$$

where we write $\tilde{h}_{+} = \tilde{h}_{+}(f)$ and $\tilde{h}_{\times} = \tilde{h}_{\times}(f)$ for conciseness.

Stokes parameters in Equations 5.47-5.50 are defined for monochromatic plane waves. In general, EM and GW waves are polychromatic. Therefore the GW Stokes parameters, given by Equations 5.57-5.60, are functions of the GW frequency f. Consequently, the fractional linear polarization (as defined in the main text) is also a function of f:

$$F_{\mathrm{L}}(f) = \frac{\sqrt{Q(f)^2 + U(f)^2}}{I(f)}. \tag{5.61}$$

Similarly, the fractional circular polarization is given by

$$F_{\mathrm{C}}(f) = \frac{V(f)}{I(f)}; \tag{5.62}$$

and the total (degree of) polarization is

$$F_{\mathrm{T}}(f) = \frac{\sqrt{Q(f)^2 + U(f)^2 + V(f)^2}}{I(f)}. \tag{5.63}$$

5.A5. Confidence intervals of cumulative distribution functions

Dvoretzky-Kiefer-Wolfowitz (DKW) derived an inequality which can be used to compute confidence intervals for empirically determined cumulative distribution functions (CDF) [297]. For an empirical CDF, $P(x)$, with n data points, the interval which contains the true distribution $P_{\text{true}}(x)$ with a probability $\mathcal{L}$ is given by

$$P_n(x) - \varepsilon \leq P_{\text{true}}(x) \leq P_n(x) + \varepsilon, \tag{5.64}$$

with

$$\varepsilon = \sqrt{\frac{\ln\left(\frac{2}{1-\mathcal{L}}\right)}{2n}}. \tag{5.65}$$

The 2-σ confidence intervals in Figure 5.7 of the main text are calculated using $\mathcal{L} = 0.95$.

Chapter 6

Conclusions and outlook

This chapter summarizes the key results of the thesis and discuss potential directions for future research. Section 6.1 reviews the work presented in Chapter 3, in which we compare *BayesWave*'s GW burst detection efficiency between a two- and three-detector configuration in O3. Section 6.2 reviews the applications of the dedicated-frequency framework proposed in Chapter 4, with an emphasis on its potential to constrain explosion mechanisms with future CCSNe GW detections. Section 6.3 reviews the applications of *BayesWave* in characterizing GW burst polarizations, as demonstrated in Chapter 5.

6.1. Assessing burst detection performance of expanded detector networks in the presence of glitches

GW detectors are susceptible to a broad range of noise. While the sources of the persistent detector noise floors are generally well-identified and characterized, transient noise events (glitches) have been observed recurrently in the LIGO and Virgo detectors during the first three LVK observing runs, sometimes with unknown origins. Glitches challenge the detection and characterization of GW bursts, as their similar time-frequency structures can mimic or mask burst signals, leading to false alarms or missed detections.

BayesWave is a burst analysis algorithm used widely by the LVK community for burst searches and waveform consistency tests. The algorithm distinguishes between a coherent signal $\mathcal{S}$ and uncorrelated glitches $\mathcal{G}$ across the detector network, by comparing their Bayesian evidences via the Bayes factor $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$, which serves as the detection statistic. A novel feature of $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ is that it scales not only with the signal strength but also the time-frequency complexity and network coherence of the signal [85], setting *BayesWave* apart from other burst analysis algorithms, whose detection statistics depend mainly on signal strength. Consequently, *BayesWave* is expected to detect signals with increased confidence as the detector network expands [3]. However, during O3, the overall glitch rate in the LIGO detectors is approximately five times higher than in O2 [36]. While the O3 glitch rate in Virgo is lower than in O2, it remains comparable to that of the LIGO detectors. This raises the question of whether *BayesWave* continues to perform better with expanded detector networks at their current sensitivities, in the presence of glitches.

6.1.1. *Summary of results*

We address the above query in Chapter 3 of this thesis, by comparing *BayesWave*'s burst detection efficiency between a two- and three-detector network in O3. The two-detector network (HL) comprises the LIGO detectors in Hanford and Livingston, and the three-detector network (HLV) comprises HL plus Virgo. The detection efficiency $P_{\rm det}$ of the detector configurations is compared as a function of false-alarm probability $P_{\rm FA}$ using efficiency (ROC) curves. False alarms, by definition, are non-astrophysical triggers in the detector noise background with detection statistics comparable to that of astrophysical triggers. Since each burst analysis algorithm employs a unique detection statistic, they characterize false alarms differently and produce varying noise background measurements. The noise background cannot be modeled from first principles at present and must be measured empirically. The O3 background for *BayesWave* is measured by analyzing a subset of triggers identified by the coherent WaveBurst (cWB) algorithm in the time-shifted O3 HL and HLV data, which ensures the absence of coherent GW signals. The use of cWB is necessary due to *BayesWave*'s high computational cost, which makes it impractical to analyze the full 10^3-year span of time-shifted data. The background measurements show that the likelihood of false-alarm detections with *BayesWave* is generally an order of magnitude higher in the larger HLV network compared to HL. We then compare the ROC curves between HL and HLV by analyzing astrophysical BBH signals injected into the same segment of O3 data that produced the background measurements above. Although the $P_{\rm det}$ of HL is, on average, 2% higher than that of HLV in the astrophysically relevant range $P_{\rm FA} \leq 0.4$, the ROC curves for both detector configurations remain comparable overall. As a consistency check, we compare the detection significance of 18 O3-like CBC signals measured by HL and HLV, and find that they are also comparable. Altogether, our results suggest that the increased $P_{\rm FA}$ offsets the advantage of increased $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ in larger detector networks. Therefore, a larger network does not always achieve better detection efficiency than a smaller one. Our findings are consistent with other burst analyses [46, 203].

6.1.2. *A pilot study: detection efficiency dependence on SNR*

In Section 3.7.2, we proposed several potential avenues for future work, which include a comparison of *BayesWave*'s burst detection performance between HL and a three-detector configuration with equal sensitivity. The deployment of LIGO-India will facilitate this study. However, there is a considerable wait time before this can be explored, as the impact of glitches on detection efficiency can only be assessed using real observational data, given that glitches cannot be predicted in advance. Therefore, it is worthwhile to reassess and compare the efficiencies of the HL and HLV networks with improved sensitivities [71], once the data from the fourth LVK observing run becomes available, as well as in future runs. In the meantime, we can also consider studying the dependence of detection efficiency on SNR. We outline the motivation and suggested methods for this study below.

The detection efficiency study presented in Chapter 3 uses a set of BBHs with mixed SNRs ranging from ∼10 to ∼50. This dataset and the results are representative of the collective CBC population, where signals span a range of SNR values (∼5-30 in GWTC-3). Although HL performs comparably to HLV for this signal population for the mixed-

SNR population, are there specific SNR ranges where HLV is more sensitive than HL, or vice versa? If so, would detections be missed if we use only one detector configuration for the search?

As a first step towards answering this question, we conduct a pilot study using five injection sets, IS-a, IS-b, IS-c, IS-d, and IS-e, each consisting of 100 BBH signals injected at $\mathrm{SNR_{net}}\approx$10, 11, 12, 14 and 19 respectively. To minimize SNR variations due to sky location, all injected BBHs share the same sky location as GW170817. Additionally, to reduce the impact of the detectors' diurnal sensitivity cycle on the SNR, the BBHs are injected into a 6-hour segment of O3 data, where the network sensitivity remains relatively consistent. The left column of Figure 6.1 shows the $\mathrm{SNR_{net}}$ distributions of the injection sets, and the corresponding efficiency curves are shown in the right column. We find that the HLV efficiency curves for IS-a and IS-b are generally higher than those of HL in the astrophyically relevant region $P_{\mathrm{FA}} \leq 0.4$; the opposite is true for the remaining three higher SNR injection sets. These preliminary results suggest that HLV may be more efficient than HL at detecting lower SNR signals, and vice versa. A more comprehensive study, with careful consideration of sky locations and the physical properties of the BBHs, is needed to confirm this finding. Assuming that GW sources are uniformly distributed throughout the universe, their SNR priors are expected to scale as $\sim \mathrm{SNR}^{-4}$, i.e. low SNR signals are much more common than higher SNR signals [86]. Therefore, if HLV can detect low SNR signals more effectively than HL, we should assess the likelihood of missing low SNR signals with HL alone, and consider whether it is worthwhile to use both configurations in future searches to ensure they are not overlooked.

6.2. Constraining core-collapse supernova explosion models with gravitational-wave burst detections

CCSNe have been studied through EM observations for decades. However, these observations do not provide insight into the interior of the CCSNe before or during the explosion, as the pre-explosion EM emissions are unable to penetrate the outer layers of stellar material. The detection of GW bursts from CCSNe is highly anticipated within the astrophysics community, for it will enable studies of the CCSN interior, such as the PNS EOS and explosion mechanisms, for the first time. In Chapter 4 of this thesis, we focus on constraining CCSN explosion mechanisms through GWs. Various multi-dimensional hydrodynamic simulations have identified the SASI and neutrino-driven convection as potential drivers of CCSN explosions. These processes have also been shown to emit low-frequency GW signatures ($\lesssim$ 250 Hz). Therefore, CCSN explosion models can be constrained based on the detection of low-frequency GW emissions, or lack thereof. We propose the *dedicated-frequency framework*, a new approach inspired by Ref. [221], for detecting and characterizing frequency-specific GW signatures. Below, we summarize methods and applications of the framework.

6.2.1. Summary of results

The dedicated-frequency framework seeks to enhance the detection of frequency-specific GW signatures using follow-up analyses, limited to a user-selected frequency range.

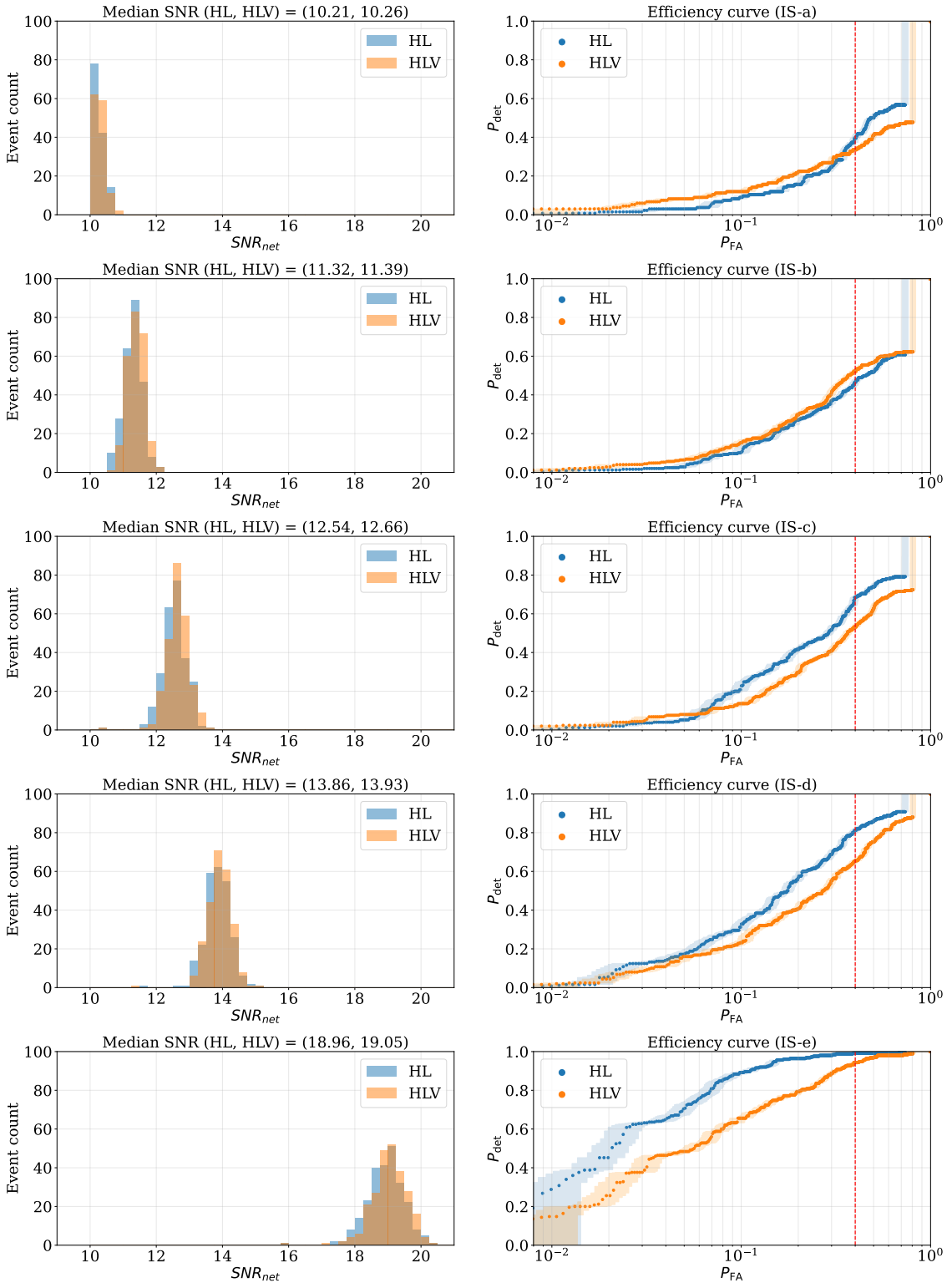

Figure 6.1: Detection efficiency dependence on SNR. The rows, from top to bottom, show the plots for IS-a, IS-b, IS-c, IS-d, and IS-e, arranged in increasing order of overall injected network SNR (SNR_{net}). The left column shows the SNR_{net} distributions of the HL (blue) and HLV (orange) injections, with the corresponding medians labeled at the top of each plot. The right column shows the efficiency curves, i.e. detection efficiency P_{det} as a function of false alarm probability P_{FA}. The region $P_{FA} \leq 0.4$ is shaded green to indicate astrophysical relevance.

The dedicated-frequency analyses are not burst detection tools; they serve as follow-ups to burst triggers that satisfy the standard GW detection benchmark. The framework adopts the hierarchical cWB plus *BayesWave* pipeline, which is also employed in the LVK all-sky short-duration burst searches [44–46] and has been shown to enhance detection confidence compared to independent cWB and BayesWave analyses. [139]. The cWB algorithm identifies suitable follow-up candidates, and *BayesWave* performs the dedicated-frequency (LF and/or HF) analyses as a follow-up. In this work, we demonstrate two applications of the dedicated-frequency follow-ups: (i) the LF analysis, limited to 32-256 Hz, for constraining CCSN explosion models through the detection of low-frequency GWs associated with SASI and/or neutrino-driven convection, and (ii) the HF analysis, limited to range 256-2048 Hz, for enhancing detection significance of burst triggers with minimal power in the low-frequency range.

The LF analysis dataset comprises GW signals from five non-rotating, solar-metallicity CCSN models: SFHx, s25, D15, mesa20_pert, and s18, listed in decreasing order of low-frequency ($\leq$ 256 Hz) power. The CCSN models are injected into the HL O3 data, rather than HLV, informed by results from previous cWB studies and the findings from Chapter 3. We find that the SFHx model, which exhibits the highest low-frequency power, achieves a 97% detection efficiency. That is, its detection in the LF analyses is almost guaranteed. In contrast, s18 only has a 10.5% detection efficiency with the LF analyses, due to the absence of low-frequency GWs induced by SASI/neutrino-driven convection. The s25, D15 and mesa20_pert models have $\sim$50-70% LF detection efficiency, suggesting that the LF analysis may fail to detect CCSN signals with moderate levels of SASI and neutrino-convection emissions. This result suggests that a non-detection with the LF analysis provides weak constraints on explosion models, as it is ambiguous whether the low-frequency GW emissions are missed or genuinely absent. In contrast, a successful LF detection confirms the presence of low-frequency GW emissions, making it a valuable tool for constraining CCSN explosion models.

The HF analysis of the dedicated-frequency framework can be used to enhance burst detection significance. We demonstrate this using the loudest trigger associated with SN 2019fcn, an optically-observed CCSN during O3. This trigger is a suitable candidate because it has minimal power below 256 Hz, allowing the HF analysis to reassess its significance without truncating any signal power. We show that the FAR of the trigger is reduced from 22.1 yr^{-1} with the full-band cWB analysis to 4.6 yr^{-1} with the *BayesWave* HF follow-up.

The dedicated-frequency framework can be extended in several directions. We discuss possible extensions in Section 6.2.2 and applications in Section 6.2.3.

6.2.2. Dedicated-frequency analyses with cWB

We initially propose to use both cWB and *BayesWave* for the dedicated-frequency follow-ups. However, we find that the cWB LF analyses are less efficient at detecting low-frequency GW signatures compared to *BayesWave* LF follow-ups. To illustrate this, we show the cWB results for the CCSN analysis in Figure 6.2, which can be compared with the *BayesWave* results in Chapter 4.

The left panel of Figure 6.2 shows the cWB background measurements. Similar to *BayesWave* (bottom panel of Figure 4.3), the LF background (pink) is lower than the

full-band background (purple) because noise events dominated by high-frequency ($f \geq$ 256 Hz) power are suppressed. However, the cWB full-band and LF backgrounds differ by a factor of 1.5 on average, c.f. *BayesWave*, in which the LF background is a factor of 10 lower than the full-band background on average. This is because the detection statistic η_r of cWB scales only with the network SNR, whereas $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$ of *BayesWave* scales with both the network SNR and the number of wavelets, i.e. model complexity. Therefore, limiting the analysis frequency range has less impact on η_r than on $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$. The right panel of Figure 6.2 shows the FAR retrieved by cWB, for the same injections as in Figure 4.6. The horizontal dashed lines at FAR= 1 yr^{-1} indicate the cWB detection threshold, below which an event qualifies as a detection and eligible for dedicated-frequency followups. Figure 4.6 shows that for the full-band cWB analysis, the median FAR (cross symbols) and the corresponding IQR[1] satisfy FAR < 1 yr^{-1}, confirming that the CCSN injections satisfy the criteria for dedicated-frequency follow-ups. In contrast, for the cWB LF analysis, the median FAR (triangle symbols) for all models, except SFHx, are above the FAR = 1 yr^{-1} threshold; the IQRs for all models encompass FAR values greater than 1 yr^{-1}. Moreover, the medians of D15 and mesa20_pert are at the maximum FAR, suggesting that >50% of the signals with moderate low-frequency emissions are missed by the cWB LF follow-up, c.f. <50% missed by *BayesWave* in Figure 4.6. Altogether, we find that LF analyses with cWB are generally less effective than with *BayesWave* at detecting low-frequency GW signatures for constraining CCSN models. Therefore, we present only the *BayesWave* results for the dedicated-frequency follow-ups in Chapter 4.

Nevertheless, we suggest exploring alternative approaches for frequency-specific GW characterization with cWB, as consistent results from two independent methods improve reliability of the findings. The cWB algorithm identifies triggers by clustering power in excess of the average detector noise fluctuations in the time-frequency domain. Since the data is bandpassed for dedicated-frequency analyses, excess-power clustering in the time-frequency domain is reduced, thereby diminishing cWB's signal characterization performance. Therefore, future work could consider maintaining the full-band cWB analysis to preserve the accuracy of signal reconstruction, and analyze specific bands of the reconstructed full-band signal thereafter. We also recommend establishing metrics to quantify cWB signal characterization performance within specific frequency bands, for example: (i) to quantify the accuracy of frequency-specific characterization, we can compare the injected and recovered power fractions in a specific frequency band relative to the full-band signal, and (ii) we can empirically determine the minimum required fraction of recovered power for a signal, to be considered a successful detection within the selected band. In principle, these metrics should also be applicable to the full-band analyses with *BayesWave*. Therefore, we suggest exploring this alternative approach with *BayesWave*, to assess its potential in complementing the dedicated-frequency analyses.

6.2.3. Non-CCSN bursts

So far, we have applied the dedicated-frequency framework in the LF (32-256 Hz) and HF (256-2048 Hz) ranges for characterizing GW signatures of CCSNe. However, there are

[1]The IQRs of the full-band cWB analysis are not visible in the right panel of Figure 4.6 because the lower and upper quartiles are exactly at the maximum FAR = $1 \times 10^4\,\mathrm{yr}^{-1}$, along with the median.

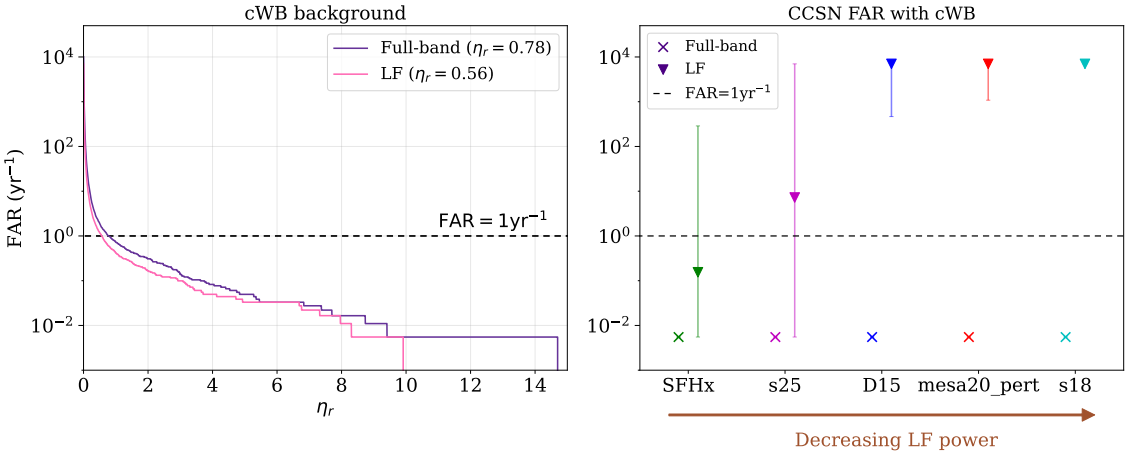

Figure 6.2: Dedicated-frequency analysis of the CCSN injections with cWB. The left panel shows the full-band (purple curve) and LF (pink curve) background measurements with cWB; the full-band cWB background is identical to that in Figure 4.3. The right panel shows the same as for Figure 4.6, but for cWB. The horizontal dashed lines in both panels indicate the cWB detection threshold at FAR=1 yr^{-1}. The η_r values corresponding to FAR=1 yr^{-1} are shown in the legend of the left panel.

other burst sources with frequency-specific signatures that can provide valuable insights for constraining the properties of these sources.

GWs emitted by BBHs with eccentric orbits can be decomposed into a series of harmonics. Ref. [256] shows that the frequency of a (2,2) multipole GW signal can be expressed as

$$f_{\mathrm{GW},k} = 2f_\phi + kf_r. \tag{6.1}$$

Here, f_ϕ is the azimuthal frequency, defined by the time taken for the binary to complete one orbit; $f_r (< f_\phi)$ is the radial frequency, defined by the time taken for the binary to return to apoapsis; and k denotes integer multiples. The GW signal power is dominated by the zeroth order mode with $f_{\mathrm{GW},0} = 2f_\phi$, which is analogous to a non-eccentric system where the GW frequency is twice the orbital frequency. Consequently, the eccentricity can be inferred using the amplitude ratio between the sub-dominant $k \neq 0$ modes and the dominant $k = 0$ mode, which increases with increasing eccentricity [256]. Since the sub-dominant and dominant modes are distinguishable by frequency, as shown in Equation 6.1, dedicated-frequency analyses which exclude the dominant-mode frequency ($f_{\mathrm{GW},0}$) could improve the characterization of the sub-dominant modes and facilitate eccentricity inference. This presents an interesting avenue for future work.

Another potential burst source for dedicated-frequency followups is the remnant of BNS mergers. If the remnant is a quasi-stable NS, it may produce a short-lived (~10-100 ms) GW signal with frequencies in the 1-4 kHz range, cause by transient non-axisymmetric deformations and quadrupole oscillations. Distinct spectral features in the post-merger GW signal correspond to different physical processes, and the characterization of these features helps constrain physical properties of the remnant. For example, the dominant quadrupolar oscillation of the NS remnant gives rise to the primary spectral peak, whose frequency is directly correlated with the radius of a non-rotating NS. This property, together with the mass, can be used to constrain the NS EOS [255]. If

the remnant is a BH, the GW spectra will feature a cut-off after the BH formation at ∼2-3 kHz, which can be used to infer the BNS mass ratio, and consequently the extent of tidal effects during the merger[2] [32]; for a fixed total mass, lower cut-off frequencies correspond to high mass ratio (unequal mass) systems, as stronger tidal effects cause the BH to form earlier during the inspiral. Dedicated-frequency analyses focused on the bands containing these spectral features, could improve their characterization and offer further insights into the physical properties of a BNS and its remnant.

6.3. Characterizing gravitational-wave burst polarizations with the global detector network

The expanding network of GW detectors facilitates the decomposition of polarization content in GW bursts, which in turn can offer valuable insights into the physical and geometric properties of the emitting source. In Chapter 5, we pivot from the detection aspect of *BayesWave*, and explore its prospects in characterizing tensor polarizations of GW bursts with multi-detector networks of different sizes. *BayesWave* offers two signal models —elliptical (E) and relaxed (R) —which are consistent with the tensor GW polarizations predicted by general relativity. This study focuses on two performance aspects of *BayesWave*: (i) its ability to identify deviations from E through Bayesian model selection between E and R, and (ii) the accuracy of R in reconstructing tensor GW polarizations in terms of Stokes parameters.

6.3.1. Summary of results

For this study, *BayesWave* analyzes two injection sets: one comprising non-precessing BBHs, which represent elliptical GW signals with fixed polarization content, and the other comprising precessing BBHs, which represent non-elliptical signals with time-varying polarization content. The signals are injected into Gaussian noise colored by the design sensitivity curves of future upgrades of the LIGO, Virgo and KAGRA detectors, and analyzed using both the E and R models.

In order to compare the performance between E and R, we evaluate the log Bayes factor $\ln \mathcal{B}_{R,E}$ and waveform overlap $\mathcal{O}_{R,E}$ between the R and E signal reconstructions. We find $\mathcal{O}_{R,E} > 0.98 \approx 1$ for non-precessing BBHs, indicating that the signals are reconstructed equally well by E and R. We also find $\ln \mathcal{B}_{R,E} < 0$, meaning that E is generally preferred as it has fewer parameters compared to R. The same holds true for minimally precessing BBHs, typically with initial precession spin parameter $\chi_{p,\mathrm{init}} \lesssim 0.5$. Precessing BBHs with high $\chi_{p,\mathrm{init}} \gtrsim 0.5$, in contrast, are more likely to exhibit time-varying (non-elliptical) polarization which is more accurately modeled by R then E. Consequently, $\ln \mathcal{B}_{R,E} > 0$, indicating a preference for R over E. The preference for

[2] Asymmetric matter movement in the disk surrounding the BH remnant produces a broadband peak around 5-6 kHz in the GW spectrum, whose width also reveals the extent of tidal effects during the BNS merger [32]. However, the amplitude of this peak is approximately an order of magnitude lower than that of the inspiral signal, and its frequency lies beyond the observing bands (<5 kHz) of current-generation detectors. Detection of this peak would require, at a minimum, kHz-band detectors like the Neutron Star Extreme Matter Observatory [298].

R strengthens as the detector network expands, as R can more accurately characterize the features of strongly precessing signals. The E versus R analysis is repeated for a subset of O3 GW events, selected based on their loudness and short signal duration to ensure reliable reconstructions by *BayesWave*. We find that E is preferred over R with $-12.88 \leq \ln \mathcal{B}_{\mathcal{S},\mathcal{G}} \leq -0.01$ for all O3 events except GW200129_065458, which remains under debate as to whether it is precessing. Although $\ln \mathcal{B}_{R,E} = 0.48 \pm 0.35$ for GW200129_065458 does not strongly favor R, it is the only event with positive $\ln \mathcal{B}_{R,E}$. Altogether, the results suggest that $\ln \mathcal{B}_{R,E}$ and $\mathcal{O}_{R,E}$ together can provide indications for exceptional source properties, such as potential deviations from elliptical polarizations, especially with larger detector networks.

Using the same BBH injection sets, we assess the accuracy of R in measuring the fractional circular (F_{C}), linear (F_{L}) and total (F_{T}) polarizations of GW bursts, as defined by Stokes' parameters. Using the root mean squared residual $\mathcal{R}_{\mathrm{RMS}}(F_{\mathcal{P}})$ for $\mathcal{P} \in \{\mathrm{C}, \mathrm{L}, \mathrm{T}\}$ to quantify the accuracy of measurement, we find that a four-detector network reduces $\mathcal{R}_{\mathrm{RMS}}(F_{\mathcal{P}})$ by approximately a factor of two compared to a three-detector network. In other words, larger detector networks yield more accurate polarization measurements, as expected. Additionally, we find that the accuracy of polarization measurements becomes independent of signal morphology when the detector network is large enough, i.e. with three or more detectors.

6.3.2. *Future burst polarization studies with BayesWave*

In Section 5.7, we discussed the prospects of using *BayesWave* to test general relativity through Bayesian model selection. This is only viable once the expanding detector networks impose tighter constraints on the parameter space of non-GR polarizations. Another possible extension is to leverage the *BayesWave* E and R models to enhance the detection and characterization of targeted burst sources through GW polarizations. This work is achievable at present for CCSNe, guided by GWs extracted from hydrodynamic simulations.

Ref. [240] shows that low-frequency ($\sim$100–200 Hz) GWs associated with SASI in the SFHx [222] CCSN model exhibit strong circular polarizations, V, as defined in Equation 5.60. The SNR of the full signal is measured in terms of the total intensity I as defined in Equation 5.57:

$$\mathrm{SNR}_I = \frac{I(h+n)}{\sqrt{\langle I(n) - \langle I(n) \rangle \rangle^2}}. \tag{6.2}$$

Here, $I(h+n)$ is the total intensity of the signal h plus Gaussian noise n, $I(n)$ is the total intensity of n alone, and $\langle x \rangle$ denotes the time average of quantity x over the duration spanned by $h+n$. By replacing I with V in Equation 6.2, one measures the SNR of the circularly polarized component. Ref. [240] shows that $\mathrm{SNR}_V > \mathrm{SNR}_I$, suggesting that the circular polarization component has a higher detectability than the full signal. As discussed in Chapter 4 of this thesis, the detection of SASI is useful for constraining CCSNe explosion models. It is therefore interesting to study the prospects of *BayesWave* in detecting and characterizing SASI through the circular GW polarizations.

To give a taste of what is possible, we perform an exploratory study by analyzing 75

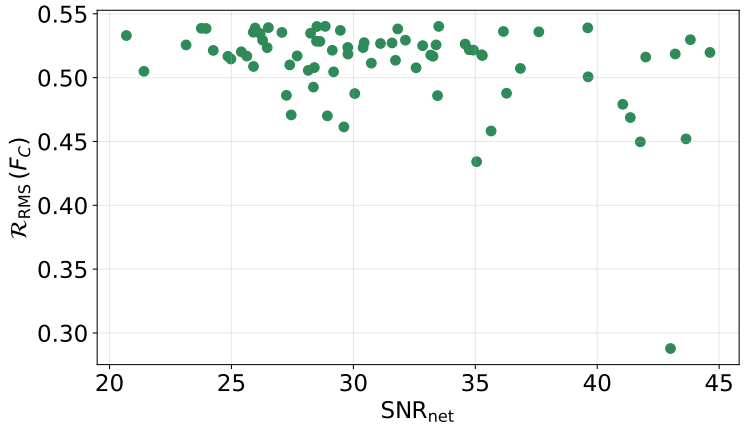

Figure 6.3: *BayesWave* measurements of SFHx circular polarization with the HLV network. The measurement accuracy, quantified by the RMS residual of the fractional circular polarization $\mathcal{R}_{\rm RMS}(F_{\rm C})$, is shown as a function of the injected network signal-to-noise ratio, $\rm SNR_{net}$. The $\mathcal{R}_{\rm RMS}(F_{\rm C})$ is evaluated for the 100–200 Hz band, which is relevant to the SASI.

SFHx waveforms injected into HLV data[3], assuming Gaussian and stationary noise. All waveforms are injected with $h_{\rm rss} = 1.07\times10^{-22}$ and share the same orientations relative to the GW detectors, meaning the intrinsic loudness and degree of circular polarization are identical across all injections. However, the injected network signal-to-noise ratio, $\rm SNR_{net}$, differ because the identical waveforms are injected at different times, and the network's sensitivity to the signal's fixed sky location varies throughout the day. Using the R model, we quantify the accuracy of circular polarization measurement through the RMS residual $\mathcal{R}_{\rm RMS}(F_{\rm C})$, as defined in Equation 5.22. Figure 6.3 plots $\mathcal{R}_{\rm RMS}(F_{\rm C})$ as a function of $\rm SNR_{net}$. We find that the residuals $\mathcal{R}_{\rm RMS}(F_{\rm C})$, evaluated for the SASI-related frequency range 100–200 Hz, are generally double the fiducial accuracy threshold $\mathcal{R}_{\rm RMS}(F_{\rm C}) = 0.25$ established in Section 5.6.3. Moreover, $\mathcal{R}_{\rm RMS}(F_{\rm C})$ does not decrease significantly with increasing $\rm SNR_{net}$. This suggests that, for SFHx signals with $\rm SNR_{net} < 45$, HLV network lacks the sensitivity needed for accurate recovery of the circular polarization content associated with SASI. Although higher SNR signals could be considered, it is more meaningful to focus on realistic SNR values. Therefore, future studies should consider including KAGRA and/or future detectors, with representative sensitivities as in Figure 6.4, to explore the potential for improving the accuracy of circular polarization recovery.

All SFHx injections are recovered with $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}} > 0$ and $\ln \mathcal{B}_{\mathcal{S},\mathcal{N}} > 0$, indicating that *BayesWave* successfully identified the injections as signals using both models $\mathcal{S} \in \{R, E\}$. Therefore, in addition to the circular polarization analysis, we can compare the performance of E and R in detecting and characterizing the full SFHx signal. The left panel of Figure 6.5 shows an event-wise comparison of the reconstruction accuracy between R and E, quantified by their overlaps with the injected waveform, $\mathcal{O}_{{\rm inj},R}$ and $\mathcal{O}_{{\rm inj},E}$. All events lie close to the locus $\mathcal{O}_{{\rm inj},E} = \mathcal{O}_{{\rm inj},R}$, which implies that R and E characterize the SFHx waveforms equally well. The color of the data points darkens as one moves from

[3]Since this is an exploratory study, we use HLV instead of HLKV to conserve computational resources.

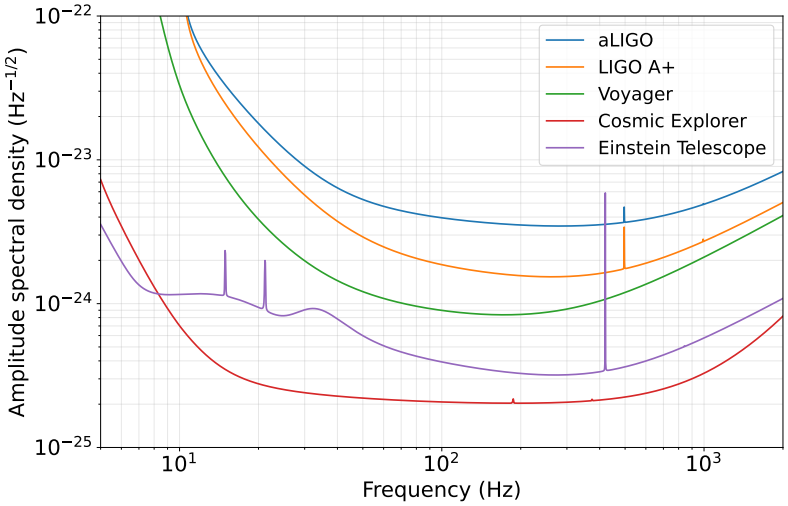

Figure 6.4: Projected sensitivities of future detectors: the upgraded LIGO A+ [299] and LIGO Voyager [300], the Cosmic Explorer [301] and Einstein Telescope [302]. The projected sensitivity of current Advanced LIGO (aLIGO) [303] detectors is included for reference.

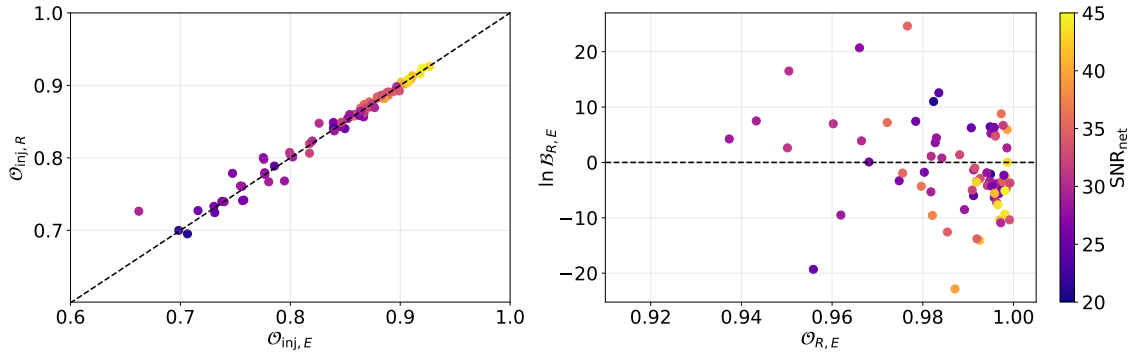

Figure 6.5: R vs E for SFHx injections. The left panel compares accuracy of waveform reconstructed between R and E, by plotting overlaps $\mathcal{O}_{\mathrm{inj},R}$ versus $\mathcal{O}_{\mathrm{inj},E}$; the black diagonal line represents the locus where $\mathcal{O}_{\mathrm{inj},R} = \mathcal{O}_{\mathrm{inj},E}$. The right panel shows $\ln \mathcal{B}_{R,E}$ versus $\mathcal{O}_{R,E}$; the black horizontal line corresponds to $\ln \mathcal{B}_{R,E} = 0$. The color bar indicates the SNR$_{\mathrm{net}}$ and applies to both panels.

the top right to the bottom left of the plot, indicating that reconstruction accuracy decreases with the $\mathrm{SNR}_{\mathrm{net}}$, as expected. The right panel of Figure 6.5 shows $\ln \mathcal{B}_{R,E}$ as a function of $\mathcal{O}_{R,E}$. We find that events with lower $\mathrm{SNR}_{\mathrm{net}}$ yield lower $\mathcal{O}_{R,E}$ due to the reduced reconstruction accuracy, which consequently results in larger deviations from $\ln \mathcal{B}_{R,E} = 0$. However, at a fixed $\mathcal{O}_{R,E}$, we observe roughly equal numbers of events above and below $\ln \mathcal{B}_{R,E} = 0$; this result makes sense because the SFHx injections are identically polarized, so there should be no preference bias towards a specific signal polarization model. The spread in $\ln \mathcal{B}_{R,E}$ at a fixed $\mathcal{O}_{R,E}$ could arise from: (i) the difference in $\mathrm{SNR}_{\mathrm{net}}$, and (ii) the change in detector orientation relative to the fixed GW source, caused by Earth's rotation, which affects the distinguishability of the polarization modes and hence the signal characterization. Altogether, we conclude that R and E are equally preferred for the SFHx injections. However, this result only holds for a particular orientation of the SFHx model. It would be valuable to follow up with an injection set that features uniformly distributed orientations, to determine whether E or R can improve the overall detectability of the SFHx waveforms.

Another possible follow-up study is to implement the dedicated-frequency framework introduced in Chapter 4, to assess whether *BayesWave*'s polarization characterization performance can be improved using bandpass analyses.

6.4. Final remarks

BayesWave is one of the leading burst analysis tools available to the GW community. This thesis contributes to understanding the performance and applications of *BayesWave* in burst detection and characterization. We conclude from Chapter 3 that larger detector networks do not necessarily improve the sensitivity of *BayesWave* to bursts due to the increased presence of glitches. Guided by the results of Chapter 3, we use the two-detector (HL) network in Chapter 4 to demonstrate how *BayesWave*, together with cWB, can serve as a tool for constraining CCSN explosion models. In Chapter 5, we show that larger detector networks are useful for characterizing GW burst polarizations. We are excited to see how our work will impact the future of GW burst analysis.

References

[1]N. J. Cornish and T. B. Littenberg, Classical and Quantum Gravity **32**, 135012 (2015).

[2]N. J. Cornish, T. B. Littenberg, B. Bécsy, K. Chatziioannou, J. A. Clark, S. Ghonge, and M. Millhouse, Physical Review D **103**, 044006 (2021).

[3]Y. S. C. Lee, M. Millhouse, and A. Melatos, Physical Review D **103**, 062002 (2021).

[4]Y. S. C. Lee, M. Millhouse, and A. Melatos, Physical Review D **109**, 082002 (2024).

[5]Y. S. C. Lee, M. J. Szczepańczyk, T. Mishra, M. Millhouse, and A. Melatos, Physical Review D **112**, 082006 (2025).

[6]Y. S. C. Lee, S. Doshi, M. Millhouse, and A. Melatos, Physical Review D **111**, 082002 (2025).

[7]A. Mandlik et al., Monthly Notices of the Royal Astronomical Society **532**, 2644–2656 (2024).

[8]L. Dunn, C. Flynn, M. Bailes, Y. S. C. Lee, G. Howitt, A. Melatos, V. Gupta, A. Mandlik, and A. Deller, Monthly Notices of the Royal Astronomical Society **541**, 1792–1815 (2025).

[9]Y. S. C. Lee, K. H. Thong, and Y. Bu, in Proceedings of The Australian Conference on Science and Mathematics Education (2023), p. 48.

[10]F.-Y. Zhao, R. G. Strom, and S.-Y. Jiang, Chinese Journal of Astronomy and Astrophysics **6**, 635–640 (2006).

[11]P. F. Winkler, G. Gupta, and K. S. Long, The Astrophysical Journal **585**, 324–335 (2003).

[12]D. H. Staelin and I. Reifenstein Edward C., Science **162**, 1481–1483 (1968).

[13]D. W. Pankenier, Journal of Astronomical History and Heritage **9**, 77–82 (2006).

[14]D. A. Green and F. R. Stephenson, "Historical Supernovae", in *Supernovae and gamma-ray bursters*, Vol. 598, edited by K. Weiler (2003), pp. 7–19.

[15]G. de Vaucouleurs and J. Corwin H. G., The Astrophysical Journal **295**, 287 (1985).

[16]G. A. Tammann, W. Loeffler, and A. Schroeder, Astrophysical Journal Supplement **92**, 487 (1994).

[17]P. Ruiz-Lapuente, The Astrophysical Journal **612**, 357–363 (2004).

[18]R. W. Klebesadel, I. B. Strong, and R. A. Olson, The Astrophysical Journal Letters **182**, L85 (1973).

[19]C. K. Seyfert, The Astrophysical Journal **97**, 28 (1943).

[20]S. Komossa and J. Greiner, Astronomy and Astrophysics **349**, L45–L48 (1999).

[21]D. R. Lorimer, M. Bailes, M. A. McLaughlin, D. J. Narkevic, and F. Crawford, Science **318**, 777 (2007).

[22]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review D **93**, 122003 (2016).

[23]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review Letters **116**, 061102 (2016).

[24]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review Letters **119**, 161101 (2017).

[25]B. P. Abbott et al. (LIGO Scientific Collaboration, Virgo Collaboration, Fermi, GBM, INTEGRAL, et al.), The Astrophysical Journal Letters **848**, L12 (2017).

[26]M. Spera, A. A. Trani, and M. Mencagli, Galaxies **10**, 76 (2022).

[27]A. Ghosh, N. K. Johnson-McDaniel, A. Ghosh, C. Kant Mishra, P. Ajith, W. Del Pozzo, C. P. L. Berry, A. B. Nielsen, and L. London, Classical and Quantum Gravity **35**, 014002 (2018).

[28]C. L. Fryer and K. C. B. New, Living Reviews in Relativity **6**, 2 (2003).

[29]E. Thrane et al., Physical Review D **83**, 083004 (2011).

[30]B. Giacomazzo, R. Perna, L. Rezzolla, E. Troja, and D. Lazzati, The Astrophysical Journal Letters **762**, L18 (2013).

[31]E. Platts, A. Weltman, A. Walters, S. P. Tendulkar, J. E. B. Gordin, and S. Kandhai, Physics Reports **821**, 1–27 (2019).

[32]K. Kiuchi, Y. Sekiguchi, M. Shibata, and K. Taniguchi, Physical Review D **80**, 064037 (2009).

[33]B. Haskell and D. I. Jones, Astroparticle Physics **157**, 102921 (2024).

[34]B. Vaishnav, I. Hinder, D. Shoemaker, and F. Herrmann, Classical and Quantum Gravity **26**, 204008 (2009).

[35]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review X **9**, 031040 (2019).

[36]R. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review X **11**, 021053 (2021).

[37]R. Abbott et al., Physical Review D **109**, 022001 (2024).

[38]R. Abbott et al. (LIGO Scientific, Virgo and KAGRA Collaboration), Physical Review X **13**, 041039 (2023).

[39]R. Abbott et al. (LIGO Scientific and Virgo Collaboration), The Astrophysical Journal Letters **896**, L44 (2020).

[40]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Astrophys. J. Lett. **882**, L24 (2019).

[41]R. Abbott et al. (LIGO Scientific and Virgo Collaboration), The Astrophysical Journal Letters **913**, L7 (2021).

[42]R. Abbott et al. (LIGO Scientific, Virgo and KAGRA Collaboration), Physical Review X **13**, 011048 (2023).

[43]A. G. Abac et al. (LIGO Scientific, Virgo and KAGRA Collaboration), The Astrophysical Journal Letters **970**, L34 (2024).

[44]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review D **95**, 042003 (2017).

[45]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review D **100**, 024017 (2019).

[46]R. Abbott et al. (LIGO Scientific, Virgo and KAGRA Collaboration), Physical Review D **104**, 122004 (2021).

[47]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Classical and Quantum Gravity **35**, 065009 (2018).

[48]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review D **99**, 104033 (2019).

[49]R. Abbott et al. (LIGO Scientific, Virgo and KAGRA Collaboration), Physical Review D **104**, 102001 (2021).

[50]P. D. Lasky, Publications of the Astronomical Society of Australia **32**, e034 (2015).

[51]M. Sieniawska and M. Bejger, Universe **5**, 217 (2019).

[52]A. I. Renzini, B. Goncharov, A. C. Jenkins, and P. M. Meyers, Galaxies **10**, 34 (2022).

[53]J. D. Romano and N. J. Cornish, Living Reviews in Relativity **20**, 2 (2017).

[54]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review Letters **116**, 241102 (2016).

[55]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), The Astrophysical Journal Letters **818**, L22 (2016).

[56]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review Letters **116**, 221101 (2016).

[57]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review Letters **118**, [Erratum: Phys.Rev.Lett. 121, 129901 (2018)], 221101 (2017).

[58]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review Letters **119**, 141101 (2017).

[59]R. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review D **102**, 043015 (2020).

[60]C. Hoy, C. Mills, and S. Fairhurst, Physical Review D **106**, 023019 (2022).

[61]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), The Astrophysical Journal Letters **892**, L3 (2020).

[62]R. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review Letters **125**, 101102 (2020).

[63]R. Abbott et al. (LIGO Scientific, Virgo and KAGRA Collaboration), The Astrophysical Journal Letters **915**, L5 (2021).

[64]B. Schutz, *A first course in general relativity*, 3rd ed. (Cambridge University Press, 2022).

[65]C. M. Will, *Theory and experiment in gravitational physics*, 2nd ed. (Cambridge University Press, 2018).

[66]M. Isi and A. J. Weinstein, arXiv e-prints, arXiv:1710.03794 (2017).

[67]D. M. Eardley, D. L. Lee, A. P. Lightman, R. V. Wagoner, and C. M. Will, Physical Review Letters **30**, 884–886 (1973).

[68]D. M. Eardley, D. L. Lee, and A. P. Lightman, Physical Review D **8**, 3308–3321 (1973).

[69]M. Maggiore, *Gravitational waves: volume 1: theory and experiments* (Oxford University Press, Oct. 2007).

[70]É. É. Flanagan and S. A. Hughes, New Journal of Physics **7**, 204 (2005).

[71]E. Capote et al., arXiv e-prints, arXiv:2411.14607 (2024).

[72]C. Bond, D. Brown, A. Freise, and K. A. Strain, Living Reviews in Relativity **19**, 3 (2016).

[73]A. Buikema et al. (LIGO Instrument Science), Physical Review D **102**, 062003 (2020).

[74]B. J. Meers, Physical Review D **38**, 2317–2326 (1988).

[75]J. D. E. Creighton and W. G. Anderson, "Gravitational-wave detectors", in *Gravitational-wave physics and astronomy* (John Wiley Sons, Ltd, 2011) Chap. 6, pp. 197–267.

[76]J. Aasi et al. (LIGO Scientific Collaboration), Classical and Quantum Gravity **32**, 074001 (2015).

[77]F. Acernese et al. (Virgo Collaboration), Classical and Quantum Gravity **32**, 024001 (2015).

[78]T. Akutsu et al. (KAGRA Collaboration), Progress of Theoretical and Experimental Physics **2021**, 05A101 (2021).

[79]K. L. Dooley et al., Classical and Quantum Gravity **33**, 075009 (2016).

[80]LIGO Scientific, Virgo and KAGRA Collaboration, *LIGO, Virgo and KAGRA observing run plans*, https://observing.docs.ligo.org/plan/, [Accessed: Apr 2025], 2025.

[81]LIGO Scientific, Virgo and KAGRA Collaboration, *IGWN gravitational-wave strain*, https://gwosc.org/detector_status/day/20240623/, [Accessed: Jan 2025], 2024.

[82]B. Iyer, T. Souradeep, C. S. Unnikrishnan, S. Dhurandhar, S. Raja, and A. Sengupta, *LIGO-India technical report*, https://dcc.ligo.org/LIGO-M1100296/public, 2011.

[83]B. P. Abbott et al. (LIGO Scientific Collaboration), Classical and Quantum Gravity **34**, 044001 (2017).

[84]B. Sathyaprakash et al., Classical and Quantum Gravity. **29**, edited by M. Hannam, P. Sutton, S. Hild, and C. van den Broeck, [Erratum: Class.Quant.Grav. 30, 079501 (2013)], 124013 (2012).

[85]T. B. Littenberg, J. B. Kanner, N. J. Cornish, and M. Millhouse, Physical Review D **94**, 044050 (2016).

[86]H.-Y. Chen and D. E. Holz, arXiv e-prints, arXiv:1409.0522 (2014).

[87]K. Cannon, J. Creighton, and T. Creighton (LIGO Scientific and Virgo Collaboration), *header `lalsimulation.h` (coordinate systems)*, [Accessed: May 2023].

[88]W. Anderson, P. Brady, D. Chin, J. Creighton, K. Riles, and J. Whelan, *Beam Pattern Response Functions and Times of Arrival for Earthbound Interferometer*, `https://dcc.ligo.org/LIGO-T010110/public`, 2002.

[89]A. Manzotti and A. Dietz, arXiv e-prints, arXiv:1202.4031 (2012).

[90]B. F. Schutz, Classical and Quantum Gravity **28**, 125023 (2011).

[91]B. P. Abbott et al. (LIGO Scientific, Virgo and KAGRA Collaboration), Living Reviews in Relativity **21**, 3 (2018).

[92]C. Pankow, E. A. Chase, S. Coughlin, M. Zevin, and V. Kalogera, The Astrophysical Journal Letters **854**, L25 (2018).

[93]P. B. Covas et al., Physical Review D **97**, 082002 (2018).

[94]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Classical and Quantum Gravity. **37**, 055002 (2020).

[95]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Classical and Quantum Gravity **33**, 134001 (2016).

[96]J. Glanzer et al., Classical and Quantum Gravity **40**, 065004 (2023).

[97]C. Pankow et al., Physical Review D **98**, 084016 (2018).

[98]E. Payne, S. Hourihane, J. Golomb, R. Udall, R. Udall, D. Davis, and K. Chatziioannou, Physical Review D **106**, 104017 (2022).

[99]M. Zevin, S. Coughlin, S. Bahaadini, E. Besler, et al., Classical and Quantum Gravity **34**, 064003 (2017).

[100]L. K. Nuttall, Philosophical Transactions of the Royal Society of London Series A **376**, 20170286 (2018).

[101]D. Davis, T. B. Littenberg, I. M. Romero-Shaw, M. Millhouse, J. McIver, F. Di Renzo, and G. Ashton, Classical and Quantum Gravity **39**, 245013 (2022).

[102]S. Hourihane, K. Chatziioannou, M. Wijngaarden, D. Davis, T. Littenberg, and N. Cornish, Physical Review D **106**, 042006 (2022).

[103]C. J. Moore, R. H. Cole, and C. P. L. Berry, Classical and Quantum Gravity. **32**, 015014 (2015).

[104]C. Cutler and É. E. Flanagan, Physical Review D **49**, 2658–2697 (1994).

[105]S. Babak, R. Balasubramanian, D. Churches, T. Cokelaer, and B. S. Sathyaprakash, Classical and Quantum Gravity **23**, 5477–5504 (2006).

[106]T. Cokelaer, Classical and Quantum Gravity **24**, 6227–6242 (2007).

[107]T. Dal Canton et al., Physical Review D **90**, 082004 (2014).

[108]S. A. Usman et al., Classical and Quantum Gravity **33**, 215004 (2016).

[109]B. Ewing et al., Physical Review D **109**, 042008 (2024).

[110]L. Tsukada et al., Physical Review D **108**, 043004 (2023).

[111]T. Adams, D. Buskulic, V. Germain, G. M. Guidi, F. Marion, M. Montani, B. Mours, F. Piergiovanni, and G. Wang, Classical and Quantum Gravity **33**, 175012 (2016).

[112]J. Luan, S. Hooper, L. Wen, and Y. Chen, Physical Review D **85**, 102002 (2012).

[113]Q. Chu et al., arXiv e-prints, arXiv:2011.06787 (2020).

[114]S. Klimenko, I. Yakushin, A. Mercer, and G. Mitselmakher, Classical and Quantum Gravity **25**, 114029 (2008).

[115]S. Klimenko et al., Physical Review D **93**, 042004 (2016).

[116]M. Drago et al., SoftwareX **14**, 100678 (2021).

[117]R. Lynch, S. Vitale, R. Essick, E. Katsavounidis, and F. Robinet, Physical Review D **95**, 104046 (2017).

[118]E. Thrane and M. Coughlin, Physical Review Letters **115**, 181102 (2015).

[119]B. P. Abbott et al., Physical Review D **93**, 042005 (2016).

[120]A. Macquet, M. A. Bizouard, N. Christensen, and M. Coughlin, Physical Review D **104**, 102005 (2021).

[121]P. J. Sutton et al., New Journal of Physics **12**, 053034 (2010).

[122]K. C. Cannon, Physical Review D **75**, 123003 (2007).

[123]V. Skliris, M. R. K. Norman, and P. J. Sutton, Physical Review D **110**, 104034 (2024).

[124]V. Necula, S. Klimenko, and G. Mitselmakher, in Journal of physics conference series, Vol. 363, Journal of Physics Conference Series (June 2012), p. 012032.

[125]S. Klimenko, arXiv e-prints, arXiv:2201.01096 (2022).

[126]M. Millhouse, N. J. Cornish, and T. Littenberg, Physical Review D **97**, 104057 (2018).

[127]P. Jaranowski, A. Królak, and B. F. Schutz, Physical Review D **58**, 063001 (1998).

[128]T. B. Littenberg and N. J. Cornish, Physical Review D **91**, 084034 (2015).

[129]S. Ghonge, J. Brandt, J. M. Sullivan, M. Millhouse, K. Chatziioannou, J. A. Clark, T. Littenberg, N. Cornish, S. Hourihane, and L. Cadonati, Physical Review D **110**, 122002 (2024).

[130]L. S. Finn, arXiv e-prints, gr–qc/9903107 (1999).

[131]R. H. Swendsen and J.-S. Wang, Physical Review Letters **57**, 2607–2609 (1986).

[132]S. Chatterji, L. Blackburn, G. Martin, and E. Katsavounidis, Classical and Quantum Gravity **21**, S1809–S1818 (2004).

[133]S. K. Chatterji, "The search for gravitational wave bursts in data from the second LIGO science run", PhD thesis (Massachusetts Institute of Technology, Jan. 2005).

[134]P. M. Goggans and C.-Y. Chan, eds., *Proceedings, 29th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (MaxEnt 2009): Oxford, Mississippi, July 5-10, 2009*, Vol. 1193 (2009), pp.1–430.

[135]M. Abt and W. J. Welch, Canadian Journal of Statistics **26**, 127–137 (1998).

[136]T. B. Littenberg and N. J. Cornish, Physical Review D **80**, 063007 (2009).

[137]S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza, and A. Bohé, Physical Review D **93**, 044006 (2016).

[138]S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Physical Review D **93**, 044007 (2016).

[139]J. B. Kanner, T. B. Littenberg, N. Cornish, M. Millhouse, E. Xhakaj, F. Salemi, M. Drago, G. Vedovato, and S. Klimenko, Physical Review D **93**, 022002 (2016).

[140]S. Ghonge, K. Chatziioannou, J. A. Clark, T. Littenberg, M. Millhouse, L. Cadonati, and N. Cornish, Physical Review D **102**, 064056 (2020).

[141]K. Chatziioannou, C.-J. Haster, T. B. Littenberg, W. M. Farr, S. Ghonge, M. Millhouse, J. A. Clark, and N. Cornish, Physical Review D **100**, 104004 (2019).

[142]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review X **9**, 011001 (2019).

[143]K. Chatziioannou, J. A. Clark, A. Bauswein, M. Millhouse, T. B. Littenberg, and N. Cornish, Physical Review D **96**, 124035 (2017).

[144]A. Torres-Rivas, K. Chatziioannou, A. Bauswein, and J. A. Clark, Physical Review D **99**, 044014 (2019).

[145]G. Dálya, P. Raffai, and B. Bécsy, Classical and Quantum Gravity. **38**, 065002 (2021).

[146]N. Gupte et al., arXiv e-prints, arXiv:2404.14286 (2024).

[147]K. W. Tsang, M. Rollier, A. Ghosh, A. Samajdar, M. Agathos, K. Chatziioannou, V. Cardoso, G. Khanna, and C. Van Den Broeck, Physical Review D **98**, 024023 (2018).

[148]K. W. Tsang, A. Ghosh, A. Samajdar, K. Chatziioannou, S. Mastrogiovanni, M. Agathos, and C. Van Den Broeck, Physical Review D **101**, 064012 (2020).

[149]N. Raza, J. McIver, G. Dálya, and P. Raffai, Physical Review D **106**, 063014 (2022).

[150]K. Gill, W. Wang, O. Valdez, M. Szczepanczyk, M. Zanolin, and S. Mukherjee, arXiv e-prints, arXiv:1802.07255 (2018).

[151]B. Bécsy, P. Raffai, N. J. Cornish, R. Essick, J. Kanner, E. Katsavounidis, T. B. Littenberg, M. Millhouse, and S. Vitale, The Astrophysical Journal **839**, 15 (2017).

[152]F. Pannarale, R. Macas, and P. J. Sutton, Classical and Quantum Gravity. **36**, 035011 (2019).

[153]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review D **100**, 104036 (2019).

[154]R. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review D **103**, 122002 (2021).

[155]R. Abbott et al. (LIGO Scientific, Virgo and KAGRA Collaboration), (2021).

[156]K. Chatziioannou, M. Isi, C.-J. Haster, and T. B. Littenberg, Physical Review D **104**, 044005 (2021).

[157]R. Abbott et al. (LIGO Scientific, Virgo and KAGRA Collaboration), Progress of Theoretical and Experimental Physics **2022**, 063F01 (2022).

[158]C. L. Fryer and K. C. B. New, Living Reviews in Relativity **14**, 1 (2011).

[159]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review D **101**, 084002 (2020).

[160]M. J. Szczepańczyk et al., Physical Review D **110**, 042007 (2024).

[161]E. Stopnitzky and S. Profumo, The Astrophysical Journal **787**, 114 (2014).

[162]S. Mereghetti, J. A. Pons, and A. Melatos, Space Science Reviews **191**, 315–338 (2015).

[163]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), The Astrophysical Journal **874**, 163 (2019).

[164]M. Ebersold and S. Tiwari, Physical Review D **101**, 104041 (2020).

[165]M. Sakellariadou, Physical Review D **42**, 354–360 (1990).

[166]T. Damour and A. Vilenkin, Physical Review D **71**, 063510 (2005).

[167]R. Abbott et al. (LIGO Scientific, Virgo and KAGRA Collaboration), Physical Review Letters **126**, 241102 (2021).

[168]B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton, Physical Review D **85**, 122006 (2012).

[169]D. A. Brown, I. Harry, A. Lundgren, and A. H. Nitz, Physical Review D **86**, 084017 (2012).

[170]P. Ajith, N. Fotopoulos, S. Privitera, A. Neunzert, N. Mazumder, and A. J. Weinstein, Physical Review D **89**, 084041 (2014).

[171]T. Dal Canton and I. W. Harry, arXiv e-prints, arXiv:1705.01845 (2017).

[172]A. Taracchini et al., Physical Review D **89**, 061502 (2014).

[173]L. Blanchet, Living Reviews in Relativity **17**, 2 (2014).

[174]S. Privitera, S. R. P. Mohapatra, P. Ajith, K. Cannon, N. Fotopoulos, M. A. Frei, C. Hanna, A. J. Weinstein, and J. T. Whelan, Physical Review D **89**, 024003 (2014).

[175]S. V. Dhurandhar and B. S. Sathyaprakash, Physical Review D **49**, 1707–1722 (1994).

[176]S. Klimenko et al., *cWB pipeline library: 6.4.0*, version cWB-6.4.0, Jan. 2021.

[177]B. P. Abbott et al. (LIGO Scientific, Virgo and KAGRA Collaboration), Living Reviews in Relativity **21**, 3 (2018).

[178]J. Powell, Classical and Quantum Gravity **35**, 155017 (2018).

[179]F. Robinet, N. Arnaud, N. Leroy, A. Lundgren, D. Macleod, and J. McIver, SoftwareX **12**, 100620 (2020).

[180]R. E. Colgan, K. R. Corley, Y. Lau, I. Bartos, J. N. Wright, Z. Márka, and S. Márka, Physical Review D **101**, 102003 (2020).

[181]M. Cabero et al., Classical and Quantum Gravity **36**, 155010 (2019).

[182]L. K. Nuttall et al., Classical and Quantum Gravity **32**, 245005 (2015).

[183]S. Soni, C. Austin, A. Effler, et al. (LIGO Scientific Collaboration), Classical and Quantum Gravity **38**, 025016 (2021).

[184]R. DeRosa, J. C. Driggers, D. Atkinson, H. Miao, V. Frolov, M. Landry, J. A. Giaime, and R. X. Adhikari, Classical and Quantum Gravity **29**, 215008 (2012).

[185]V. Tiwari et al., Classical and Quantum Gravity **32**, 165014 (2015).

[186]G. D. Meadors, K. Kawabe, and K. Riles, Classical and Quantum Gravity **31**, 105014 (2014).

[187]J. C. Driggers, S. Vitale, A. P. Lundgren, et al., Physical Review D **99**, 042001 (2019).

[188]G. Vajente, Y. Huang, M. Isi, J. C. Driggers, J. S. Kissel, M. J. Szczepańczyk, and S. Vitale, Physical Review D **101**, 042003 (2020).

[189]R. Ormiston, T. Nguyen, M. Coughlin, R. X. Adhikari, and E. Katsavounidis, Physical Review Research **2**, 033066 (2020).

[190]D. Davis and M. Walker, Galaxies **10**, 12 (2022).

[191]L. Blackburn, L. Cadonati, et al., Classical and Quantum Gravity **25**, 184004 (2008).

[192]B. P. Abbott, R. Abbott, et al., Physical Review D **80**, 102001 (2009).

[193]L. S. Collaboration, Reports on Progress in Physics **72**, 076901 (2009).

[194]J. Aasi, J. Abadie, et al., Classical and Quantum Gravity **29**, 155002 (2012).

[195]J. Abadie et al. (LIGO Scientific and Virgo Collaboration), Physical Review D **85**, 122007 (2012).

[196]M. Wąs, M.-A. Bizouard, V. Brisson, F. Cavalier, M. Davier, P. Hello, N. Leroy, F. Robinet, and M. Vavoulidis, Classical and Quantum Gravity **27**, 015005 (2010).

[197]J. Abadie, B. P. Abbott, et al., Physical Review D **81**, 102001 (2010).

[198]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review D **93**, 122004 (2016).

[199]D. Lopez, V. Gayathri, A. Pai, I. S. Heng, C. Messenger, and S. K. Gupta, Physical Review D **105**, 063024 (2022).

[200]R. Abbott et al. (LIGO Scientific, Virgo and KAGRA Collaboration), The Astrophysical Journal Supplement Series **267**, 29 (2023).

[201]LIGO Scientific and Virgo Collaboration, *LIGO Virgo strain data from observing run O3a*, `https://gwosc.org/O3/O3a/`, 2021.

[202]L. S. Finn, Physical Review D **46**, 5236–5249 (1992).

[203]M. J. Szczepańczyk et al., Physical Review D **107**, 062002 (2023).

[204]K. Janssens et al., Physical Review D **107**, 022004 (2023).

[205]D. Davis, T. Massinger, A. Lundgren, J. C. Driggers, A. L. Urban, and L. Nuttall, Classical and Quantum Gravity **36**, 055011 (2019).

[206]F. Acernese et al. (Virgo Collaboration), Classical and Quantum Gravity **39**, 045006 (2022).

[207]B. P. Abbott, R. Abbott, T. D. Abbott, et al., Classical and Quantum Gravity **37**, 055002 (2020).

[208]H. A. Bethe and J. R. Wilson, The Astrophysical Journal **295**, 14–23 (1985).

[209]H.-T. Janka, "Neutrino-Driven Explosions", in *Handbook of supernovae*, edited by A. W. Alsabti and P. Murdin (2017), p. 1095.

[210]B. Müller, Living Reviews in Computational Astrophysics **6**, 3 (2020).

[211]E. Abdikamalov, G. Pagliaroli, and D. Radice, arXiv e-prints, arXiv:2010.04356 (2020).

[212]H. A. Bethe, Review of Modern Physics **62**, 801–866 (1990).

[213]H.-T. Janka, Annual Review of Nuclear and Particle Science **62**, 407–451 (2012).

[214]K. Hirata et al., Physical Review Letters **58**, 1490–1493 (1987).

[215]R. M. Bionta et al., Physical Review Letters **58**, 1494–1496 (1987).

[216]E. N. Alexeyev, L. N. Alexeyeva, I. V. Krivosheina, and V. I. Volchenko, Physics Letters B **205**, 209–214 (1988).

[217]A. Burrows and J. M. Lattimer, The Astrophysical Journal **307**, 178 (1986).

[218]S. E. Woosley, The Astrophysical Journal **330**, 218 (1988).

[219]C. D. Ott, Classical and Quantum Gravity **26**, 063001 (2009).

[220]J. M. Blondin, A. Mezzacappa, and C. DeMarino, The Astrophysical Journal **584**, 971–980 (2003).

[221]M. Szczepańczyk and M. Zanolin, Galaxies **10**, 70 (2022).

[222]T. Kuroda, K. Kotake, and T. Takiwaki, The Astrophysical Journal Letters **829**, L14 (2016).

[223]E. P. O'Connor and S. M. Couch, The Astrophysical Journal **865**, 81 (2018).

[224]T. Mishra, S. Bhaumik, V. Gayathri, M. J. Szczepańczyk, I. Bartos, and S. Klimenko, Physical Review D **111**, 023054 (2025).

[225]M. J. Szczepańczyk et al., Physical Review D **104**, 102002 (2021).

[226]S. Woosley and H.-T. Janka, Nature Physics **1**, 147–154 (2005).

[227]S. E. Woosley, A. Heger, and T. A. Weaver, Reviews of Modern Physics **74**, 1015–1071 (2002).

[228]H. A. Bethe, G. E. Brown, J. Applegate, and J. M. Lattimer, Nuclear Physics A **324**, 487–533 (1979).

[229]E. Baron, J. Cooperstein, and S. Kahana, Physical Review Letters **55**, 126–129 (1985).

[230]S. A. Colgate and R. H. White, The Astrophysical Journal **143**, 626 (1966).

[231]A. Burrows, Reviews of Modern Physics **85**, 245–261 (2013).

[232]A. Heger, S. E. Woosley, and H. C. Spruit, The Astrophysical Journal **626**, 350–363 (2005).

[233]V. Morozova, D. Radice, A. Burrows, and D. Vartanyan, The Astrophysical Journal **861**, 10 (2018).

[234]A. Mezzacappa and M. Zanolin, arXiv e-prints, arXiv:2401.11635 (2024).

[235]H. Andresen, B. Müller, E. Müller, and H.-T. Janka, Monthly Notices of the Royal Astronomical Society **468**, 2032–2051 (2017).

[236]J. R. Wilson and R. W. Mayle, Physics Reports **163**, 63–77 (1988).

[237]J. W. Murphy, C. D. Ott, and A. Burrows, The Astrophysical Journal **707**, 1173–1190 (2009).

[238]D. Radice, V. Morozova, A. Burrows, D. Vartanyan, and H. Nagakura, The Astrophysical Journal Letters **876**, L9 (2019).

[239]A. Mezzacappa, A. C. Calder, S. W. Bruenn, J. M. Blondin, M. W. Guidry, M. R. Strayer, and A. S. Umar, The Astrophysical Journal **495**, 911–926 (1998).

[240]K. Hayama, T. Kuroda, K. Kotake, and T. Takiwaki, Monthly Notices of the Royal Astronomical Society **477**, L96–L100 (2018).

[241]J. Powell, B. Müller, and A. Heger, Monthly Notices of the Royal Astronomical Society **503**, 2108–2122 (2021).

[242]T. Foglizzo, L. Scheck, and H.-T. Janka, The Astrophysical Journal **652**, 1436–1450 (2006).

[243]A. Marek, H.-T. Janka, and E. Müller, Astronomy and Astrophysics **496**, 475–494 (2009).

[244]S. Klimenko, S. Mohanty, M. Rakhmanov, and G. Mitselmakher, Physical Review D **72**, 122002 (2005).

[245]T. Mishra, B. O'Brien, V. Gayathri, M. Szczepańczyk, S. Bhaumik, I. Bartos, and S. Klimenko, Physical Review D **104**, 023014 (2021).

[246]T. Mishra et al., Physical Review D **105**, 083018 (2022).

[247]A. G. Abac et al. (LIGO Scientific, Virgo and KAGRA Collaboration), arXiv e-prints, arXiv:2410.16565 (2024).

[248]A. W. Steiner, M. Hempel, and T. Fischer, The Astrophysical Journal **774**, 17 (2013).

[249]A. Mezzacappa et al., Physical Review D **107**, 043008 (2023).

[250]S. W. Bruenn et al., Astrophysical Journal, Supplement **248**, 11 (2020).

[251]J. Powell and B. Müller, Monthly Notices of the Royal Astronomical Society **487**, 1178–1190 (2019).

[252]B. Müller and H.-T. Janka, Monthly Notices of the Royal Astronomical Society **448**, 2141–2174 (2015).

[253]C. Henshaw, M. Arogeti, A. Heranval, and L. Cadonati, arXiv e-prints, arXiv:2402.16533 (2024).

[254]P. J. Sutton, arXiv e-prints, arXiv:1304.0210 (2013).

[255]J. A. Clark, A. Bauswein, N. Stergioulas, and D. Shoemaker, Classical and Quantum Gravity **33**, 085003 (2016).

[256]B. G. Patterson, S. M. Tomson, and S. Fairhurst, arXiv e-prints, arXiv:2411.04187 (2024).

[257]É. É. Flanagan and S. A. Hughes, Physical Review D **57**, 4535–4565 (1998).

[258]M. Isi, A. J. Weinstein, C. Mead, and M. Pitkin, Physical Review D **91**, 082002 (2015).

[259]M. Isi, M. Pitkin, and A. J. Weinstein, Physical Review D **96**, 042001 (2017).

[260]T. Callister, A. S. Biscoveanu, N. Christensen, M. Isi, A. Matas, O. Minazzoli, T. Regimbau, M. Sakellariadou, J. Tasson, and E. Thrane, Physical Review X **7**, 041058 (2017).

[261]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review Letters **120**, 201102 (2018).

[262]A. Nishizawa, A. Taruya, K. Hayama, S. Kawamura, and M.-a. Sakagami, Physical Review D **79**, 082002 (2009).

[263]T. Callister, A. S. Biscoveanu, N. Christensen, M. Isi, A. Matas, O. Minazzoli, T. Regimbau, M. Sakellariadou, J. Tasson, and E. Thrane, Physical Review X **7**, 041058 (2017).

[264]L. Tsukada, Physical Review D **108**, 124042 (2023).

[265]B. P. Abbott et al. (LIGO Scientific and Virgo Collaboration), Physical Review Letters **120**, 031104 (2018).

[266]B. P. Abbott et al. (Virgo Collaboration), Astrophys. J. **713**, 671–685 (2010).

[267]H. Takeda, A. Nishizawa, Y. Michimura, K. Nagano, K. Komori, M. Ando, and K. Hayama, Physical Review D **98**, 022008 (2018).

[268]M. Isi, Classical and Quantum Gravity. **40**, 203001 (2023).

[269]P. Schmidt, F. Ohme, and M. Hannam, Physical Review D **91**, 024043 (2015).

[270]M. Isi and A. J. Weinstein, arXiv e-prints, arXiv:1710.03794 (2017).

[271]A. Mezzacappa et al., Physical Review D **107**, 043008 (2023).

[272]D. Vartanyan, A. Burrows, T. Wang, M. S. B. Coleman, and C. J. White, Physical Review D **107**, 103015 (2023).

[273]G. Pratten, S. Husa, C. Garcia-Quiros, M. Colleoni, A. Ramos-Buades, H. Estelles, and R. Jaume, Physical Review D **102**, 064001 (2020).

[274]C. García-Quirós, M. Colleoni, S. Husa, H. Estellés, G. Pratten, A. Ramos-Buades, M. Mateu-Lucena, and R. Jaume, Physical Review D **102**, 064002 (2020).

[275]C. García-Quirós, S. Husa, M. Mateu-Lucena, and A. Borchers, Classical and Quantum Gravity. **38**, 015006 (2021).

[276]G. Pratten et al., Physical Review D **103**, 104056 (2021).

[277]G. Pratten, P. Schmidt, R. Buscicchio, and L. M. Thomas, Physical Review Research **2**, 043096 (2020).

[278]LIGO Scientific, Virgo and KAGRA Collaboration, *LVK Algorithm Library - LALSuite*, Free software (GPL), 2018.

[279]T. Gupta and N. J. Cornish, Physical Review D **109**, 064040 (2024).

[280]M. Hannam et al., Nature **610**, 652–655 (2022).

[281]L. London, S. Khan, E. Fauchon-Jones, C. García, M. Hannam, S. Husa, X. Jiménez-Forteza, C. Kalaghatgi, F. Ohme, and F. Pannarale, Physical Review Letters **120**, 161102 (2018).

[282]C. García-Quirós, M. Colleoni, S. Husa, H. Estellés, G. Pratten, A. Ramos-Buades, M. Mateu-Lucena, and R. Jaume, Physical Review D **102**, 064002 (2020).

[283]J. D. Jackson, *Classical Electrodynamics, 3rd Edition* (1998).

[284]H. Takeda, S. Morisaki, and A. Nishizawa, Physical Review D **103**, 064037 (2021).

[285]H. Takeda, S. Morisaki, and A. Nishizawa, Physical Review D **105**, 084019 (2022).

[286]N. Yunes and F. Pretorius, Physical Review D **80**, 122003 (2009).

[287]N. Cornish, L. Sampson, N. Yunes, and F. Pretorius, Physical Review D **84**, 062003 (2011).

[288]K. Chatziioannou, N. Yunes, and N. Cornish, Physical Review D **86**, 022004 (2012).

[289]K. Chatziioannou, N. Yunes, and N. Cornish, Physical Review D **95**, 129901 (2017).

[290]Y. Hagihara, N. Era, D. Iikawa, A. Nishizawa, and H. Asada, Physical Review D **100**, 064010 (2019).

[291]P. T. H. Pang, R. K. L. Lo, I. C. F. Wong, T. G. F. Li, and C. Van Den Broeck, Physical Review D **101**, 104055 (2020).

[292]Y. Gürsel and M. Tinto, Physical Review D **40**, 3884–3938 (1989).

[293]P. Gregory, *Bayesian logical data analysis for the physical sciences* (Cambridge University Press, USA, 2005).

[294]E. Aponte, Y. Yao, S. Raman, S. Frässle, J. Heinzle, W. Penny, and K. Stephan, English, Cognitive Neurodynamics **16**, 1–15 (2022).

[295]T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, Physical Review D **49**, 6274–6297 (1994).

[296]G. G. Stokes, "On the composition and resolution of streams of polarized light from different sources", in *Mathematical and physical papers*, Cambridge Library Collection - Mathematics (Cambridge University Press, 2009), pp. 233–258.

[297]A. Dvoretzky, J. Kiefer, and J. Wolfowitz, The Annals of Mathematical Statistics **27**, 642–669 (1956).

[298]K. Ackley et al., Publications of the Astronomical Society of Australia **37**, e047 (2020).

[299]L. Barsotti, L. McCuller, M. Evans, and P. Fritschel, *The A+ design curve*, https://dcc.ligo.org/LIGO-T1800042/public, [Accessed: Mar 2025], 2018.

[300]R. Adhikari et al., *LIGO Voyager Upgrade Concept*, https://docs.ligo.org/voyager/voyagerwhitepaper/main.pdf, [Accessed: Mar 2025], 2018.

[301]K. Kuns, P. Fulda, L. Barsotti, and M. Evans, *Cosmic Explorer Strain Sensitivity*, https://dcc.cosmicexplorer.org/cgi-bin/DocDB/ShowDocument?docid=T2000017, [Accessed: Mar 2025], 2023.

[302]S. Hild et al., Classical and Quantum Gravity **28**, 094013 (2011).

[303]M. Evans, R. Sturani, S. Vitale, and E. Hall, *Unofficial sensitivity curves (ASD) for aLIGO, Kagra, Virgo, Voyager, Cosmic Explorer, and Einstein Telescope*, https://dcc.ligo.org/ligo-t1500293/public, [Accessed: Mar 2025], 2020.