

ENVIRONMENTAL NOISE IN GRAVITATIONAL WAVE DETECTORS AND
THE SEARCH FOR GRAVITATIONAL WAVE SIGNALS ASSOCIATED WITH
GAMMA-RAY BURSTS DURING LIGO'S THIRD OBSERVING RUN

by

PHILIPPE DAI-QUANG NGUYEN

A DISSERTATION

Presented to the Department of Physics
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

July 2022

© 2022 Philippe Dai-Quang Nguyen
All rights reserved.

DISSERTATION ABSTRACT

Philippe Dai-Quang Nguyen

Doctor of Philosophy

Department of Physics

July 2022

Title: ENVIRONMENTAL NOISE IN GRAVITATIONAL WAVE DETECTORS
AND THE SEARCH FOR GRAVITATIONAL WAVE SIGNALS ASSOCIATED
WITH GAMMA-RAY BURSTS DURING LIGO'S THIRD OBSERVING RUN

My abstract goes here.

CURRICULUM VITAE

NAME OF AUTHOR: Philippe Dai-Quang Nguyen

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR, USA
California Polytechnic State University, San Luis Obispo, CA, USA

DEGREES AWARDED:

Doctor of Philosophy, Physics, 2022, University of Oregon
Bachelor of Science, Physics, 2016, Cal Poly San Luis Obispo

AREAS OF SPECIAL INTEREST:

Gravitational Wave Astrophysics

PROFESSIONAL EXPERIENCE:

Research Assistant, Department of Physics, University of Oregon, 2017-present
Teaching Assistant, Department of Physics, University of Oregon, 2016-2019

GRANTS, AWARDS AND HONORS:

PUBLICATIONS:

Nguyen, P., et al. (2021). Environmental noise in Advanced LIGO detectors. *Classical Quantum Gravity*, 38(14), 145001.

Davis, D., et al. (2021). LIGO detector characterization in the second and third observing runs. *Classical Quantum Gravity*, 38(13), 135014.

Fiori I., Effler A., Nguyen P., Paoletti F., Schofield R.M.S., Tringali M.C.
(2021) Environmental Noise in Gravitational-Wave Interferometers.

In: Bambi C., Katsanevas S., Kokkotas K.D. (eds) Handbook of
Gravitational Wave Astronomy. Springer, Singapore.

Abbott, R., et al. (2022). Search for Gravitational Waves Associated
with Gamma-Ray Bursts Detected by Fermi and Swift during the
LIGO–Virgo Run O3b. *The Astrophysical Journal*, 928(2), 186.

ACKNOWLEDGEMENTS

Acknowledgments go here.

TABLE OF CONTENTS

| Chapter | Page |
|--|------|
| I. INTRODUCTION | 1 |
| II. GRAVITATIONAL WAVE ASTRONOMY | 4 |
| 2.1. Gravitational waves | 4 |
| 2.2. Sources of gravitational waves | 6 |
| 2.3. Multi-messenger astronomy | 8 |
| 2.4. Gravitational wave detectors | 9 |
| III. METHODS FOR STUDYING ENVIRONMENTAL NOISE | 12 |
| 3.1. Sources of environmental noise | 13 |
| 3.2. The PEM sensor array | 13 |
| 3.2.1. Monitoring the monitors with ligocam | 15 |
| 3.3. Environmental noise injections | 17 |
| 3.3.1. Vibrational injections | 19 |
| 3.3.1.1. Beating-shakers technique | 20 |
| 3.3.1.2. Impulse injections | 21 |
| 3.3.2. Magnetic injections | 22 |
| 3.4. Coupling functions | 23 |
| 3.4.1. Single coupling site, sensor, and injection | 23 |
| 3.4.2. Multiple coupling sites, sensors, and injections | 25 |
| 3.4.3. Solving the coupling equations | 26 |
| 3.5. Uncertainties and limitations of coupling functions | 32 |
| 3.5.1. Comparison to transfer functions | 32 |
| 3.5.2. Assumptions about coupling mechanisms | 32 |

| Chapter | | Page |
|---|----|------|
| 3.5.3. Hardware limitations | 33 | |
| 3.6. The pemcoupling package | 37 | |
| 3.6.1. Processing steps | 38 | |
| 3.6.2. Data products | 39 | |
| IV. STUDIES OF ENVIRONMENTAL NOISE DURING O3 | 41 | |
| 4.1. Evaluation of coupling functions | 41 | |
| 4.2. Vibrational noise studies during O3 | 42 | |
| 4.2.1. Scattered light at the HAM5/6 septum | 42 | |
| 4.2.2. Search for the source of a 48-Hz peak | 45 | |
| 4.2.3. Input beam jitter | 46 | |
| 4.3. Magnetic noise studies during O3 | 47 | |
| 4.4. Validation of gravitational wave event candidates | 51 | |
| 4.4.1. Automated validation of O3 events | 54 | |
| 4.4.2. Event validation in O4 | 55 | |
| 4.4.2.1. Coupling function tuning | 58 | |
| 4.4.2.2. Grouping nearby sensors | 60 | |
| V. GRAVITATIONAL WAVES ASSOCIATED WITH GAMMA-RAY BURSTS | 61 | |
| 5.1. Gamma-ray bursts | 61 | |
| 5.2. GW searches | 62 | |
| 5.2.1. X-Pipeline | 65 | |
| 5.3. O3b search for GWs associated with GRBs | 68 | |
| 5.3.1. GRB sample | 70 | |
| 5.3.2. Results of the O3b search | 70 | |
| 5.3.3. Model exclusion | 73 | |
| 5.3.4. Noise effects in the generic transient search | 75 | |

| Chapter | Page |
|----------------------------|------|
| REFERENCES CITED | 79 |

LIST OF FIGURES

| Figure | Page |
|--|------|
| 1. The PEM system layout at LLO during O3, as seen on the PEM public website. | 14 |
| 2. Standard locations for vibration and magnetic injections at the LHO corner station (left), Y end station (top right), and X end stations (bottom right). | 17 |
| 3. Injection equipment photos | 19 |
| 4. Comparison of the old small-coil comb magnetic field injections with the new large-coil broadband injections. | 22 |
| 5. Diagram of a trivial coupling function measurement in the trivial case. | 26 |
| 6. Expanding Figure 5 to more than one sensor and coupling site. | 27 |
| 7. Example of a broadband acoustic noise injection and measurement of a single-injection coupling function | 30 |
| 8. Ambient noise level for the LHO HAM6 Y-axis accelerometer | 31 |
| 9. Single-injection coupling functions (upper limits not shown) for the HAM5 Y-axis accelerometer for three different shaker injection locations (on top of HAM5, on top of HAM6, and on the HAM5 chamber door). | 35 |
| 10. Single-injection coupling functions (upper limits not shown) for various magnetometers and various injections at both observatories. | 36 |
| 11. Ambient estimate of vibrational noise levels at LHO (top) and LLO (bottom). | 43 |
| 12. Time series (left) and spectrograms (right) of a vibrational impulse injection produced at the output arm of the LHO detector. | 45 |
| 13. Spectrograms of DARM and various accelerometers near the input arm and beam splitter showing a beating-shakers injection at 48 Hz. | 46 |

| Figure | Page |
|---|------|
| 14. LHO DARM spectrum before and after mitigation of the 48-Hz peak. | 47 |
| 15. Improvement in jitter coupling at LHO (left) and LLO (right) between the start and end of O3. | 47 |
| 16. Ambient estimate of magnetic noise levels at LHO (top) and LLO (bottom). | 48 |
| 17. LHO magnetic coupling measured by the last five wall-mounted coil injections performed in O3. | 49 |
| 18. Time-lines of magnetic coupling changes relative to the start of the observing run. | 50 |
| 19. Weekly trends in frequency (top) and amplitude (top) of peaks in the magnetic coupling functions. | 51 |
| 20. Probability spectrogram of an ETMY accelerometer at LLO and a constant-Q transform of the GW strain channel | 57 |
| 21. Coupling function tuning and the resulting probability spectrogram | 59 |
| 22. Cumulative distribution of p-values for the loudest on-source events of the O3b X-pipeline analyses. | 71 |
| 23. Cumulative distributions of O3b exclusion distances for 150-Hz sine-Gaussian and ADI-A waveforms. | 72 |
| 24. Histogram of redshift measurements for GRBs detected by Swift BAT. | 74 |
| 25. Cumulative distributions of O3a and O3b exclusion distances for the ADI-C (top) and 300-Hz sine-Gaussian (bottom) waveforms. | 77 |
| 26. Cumulative distributions of CSG-300 Hz exclusion distances (top) and the detector network antenna factors (bottom) for GRBs before and after the implementation of RC tracking. | 78 |

LIST OF TABLES

| Table | Page |
|--|------|
| 1. Specifications for injection equipment. | 18 |
| 2. Column descriptions for the single-injection coupling function output of the pemcoupling package. | 40 |
| 3. Parameters and properties for accretion disk instability waveform injections. | 68 |
| 4. Median 90% exclusion distances (D_{90}) for the generic transient search during O3b. | 73 |
| 5. Exclusion confidence for each injected waveform model. | 75 |
| 6. Relative increase in median D_{90} for each x-pipeline simulated waveform. | 76 |

CHAPTER I

INTRODUCTION

On April 1, 2019, the Laser Interferometer Gravitational-Wave Observatory (LIGO) collaboration ushered forth a new phase of gravitational wave (GW) astronomy as it began its third observing run, which ran in two six-month stages and concluded March 27, 2020. Together with their European counterpart Virgo, the two LIGO detectors would detect a total of 74 new GW signals throughout the third observing run (O3) (Abbott et al., 2021a, 2021b), over three times the detection rate in the first two runs. In the first observing run (O1) and the second observing run (O2), LIGO had already made the first detection of GWs from the inspiral of binary black hole (BBH) and binary neutron star (BNS) systems (GW150914 and GW170817, respectively) (Abbott et al., 2016b, 2017). The detections made in O3 include an additional detection of a BNS merger (B. P. Abbott et al., 2020), the first two detections of neutron star-black hole (NSBH) mergers (Abbott et al., 2021a), the first clear detection of an intermediate-mass black hole forming from a BBH merger (R. Abbott et al., 2020), and a number of other BBH mergers that have expanded and challenged our understanding of black hole populations.

The dramatic increase in detection rate could not have been achieved without the myriad upgrades made to the LIGO and Virgo interferometers themselves (Buikema et al., 2020). These upgrades range from changes in the laser system to replacing core interferometer optics to mitigation of disruptive external signals. Studying the behavior of the detectors is crucial to finding new ways to improve their sensitivity and stability. Detector characterization involves deploying a wide array of data analysis tools and experimental tests to understand how noise

originating from within and outside of the detectors couples into the GW data stream (Davis et al., 2019, 2021).

Despite our best efforts, unwanted noise signals still affect the detector in a number of ways. Short-duration transient signals, called *glitches*, impact analysis pipelines searching for GWs. Thus to keep up with the high event detection rates in O3 many analyses necessary in the validation of GW event candidates have been automated, with more sophisticated methods being developed for future observing runs.

This dissertation describes my contributions to GW astronomy along multiple avenues. First, Chapter II introduces gravitational waves and the anatomy of a gravitational-wave interferometer. Chapter III discusses methods used to characterize environmental noise, unwanted signals originating from outside the interferometer. Chapter IV presents the results of those methods during O3, an overview of several noise investigations in which those methods have played a crucial role, and the implementation of an automated algorithm for vetting GW detections.

The detection of a GW signal coincident with a short gamma-ray burst (GRB) originating from the BNS merger GW170817 was a breakthrough moment for the astronomical community, shedding light on the mysterious properties of GRBs. Even as the LIGO and Virgo detectors improve, GW170817 was a fortunate discovery considering its incredibly close proximity (40 Mpc). Targeted searches for GWs associated with GRBs allow much more sensitive searches for potential joint observations, and are necessary for expanding our ability to make joint detections at greater distances. Chapter V describes the connection between GWs

and GRBs, the search for GW-GRB associations during the third observing run, and considerations for future analyses.

CHAPTER II

GRAVITATIONAL WAVE ASTRONOMY

Add a little history about GW astro. This chapter gives an overview of the theoretical background necessary for understanding the emission of GWs (Section 2.1), known and expected sources of GW radiation (Section 2.2), and how a GW interferometer works (Section 2.4).

2.1 Gravitational waves

To develop theory of GWs, we begin with the Einstein field equations

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}$$

where $T_{\mu\nu}$ is the stress-energy tensor and $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$ is the Einstein tensor, comprised of the Ricci tensor $R_{\mu\nu}$, Ricci scalar R , and spacetime metric $g_{\mu\nu}$. We can approximate the solution as the flat Minkowski metric $\eta_{\mu\nu}$ plus a perturbation $h_{\mu\nu}$: $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. Assuming the perturbation is small $|h_{\mu\nu}| \ll 1$, we can construct the linearized Ricci tensor

$$R_{\mu\nu} = \frac{1}{2}(\partial_\sigma \partial_\mu h_\nu^\sigma + (\mu \leftrightarrow \nu) - \partial_\mu \partial_\nu h - \square h_{\mu\nu})$$

where $h = \eta^{\mu\nu} h_{\mu\nu}$ is the trace of $h_{\mu\nu}$, and the linearized Ricci scalar

$$R = \eta_{\mu\nu} R^{\mu\nu} = \partial_\mu \partial_\nu h - \square h_{\mu\nu}$$

This yields the linearized Einstein tensor

$$G_{\mu\nu} = \frac{1}{2} (\partial_\mu \partial_\sigma \bar{h}_\nu^\rho + (\mu \leftrightarrow \nu) - \eta^{\sigma\rho} \partial_\sigma \partial_\rho \bar{h}_{\mu\nu} - \eta_{\mu\nu} \partial_\sigma \partial_\rho \bar{h}^{\sigma\rho})$$

where $\bar{h}_{\mu\nu} := h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$ is the trace-reversed metric perturbation.

We can simplify these terms further by choosing the appropriate gauge. The Lorenz, or harmonic, gauge allows us to drop the terms containing the divergence of $\bar{h}_{\mu\nu}$. In this gauge we fix $\partial_\mu \bar{h}^{\mu\sigma} = 0$, reducing the linearized Einstein field equations to simply

$$-\square \bar{h}_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \quad (2.1)$$

In the Newtonian (slowly-varying) limit, the D'Alembertian operator becomes a spatial Laplace operator, and it can be shown that the trace-reversed perturbation reduces to the Newtonian gravitational potential Φ and the stress-energy tensor reduces to just the mass density, recovering the Poisson equation for Newtonian gravity: $\nabla^2 \Phi = 4\pi G\rho$.

In vacuum, the field equation is simply $\square \bar{h}_{\mu\nu} = 0$, the solution to which is a GW propagating at the speed of light: $\bar{h}_{\mu\nu} = A_{\mu\nu} \cos(k_\sigma x^\sigma - \phi_{\mu\nu})$, where $A_{\mu\nu}$ and $\phi_{\mu\nu}$ are the amplitude and phase of the wave.

In the slowly-varying case the gauge condition can be further restricted by making the metric perturbation purely spatial, ($h_{00} = h_{0i} = 0$) and traceless ($h = h_i^i = 0$). In this transverse-traceless (TT) gauge, we write the metric perturbation as $h_{\mu\nu}^{TT}$ (no overline necessary because trace-reversal has no effect on the traceless perturbation). The gravitational wave described by $h_{\mu\nu}^{TT}$ is transverse to its direction of propagation. These gauge conditions cause all components to vanish except two independent ones. For a monochromatic plane wave propagating

in the z direction, these two components are

$$h_{11}^{TT} = -h_{22}^{TT} = h_+(t-z)$$

$$h_{12}^{TT} = h_{21}^{TT} = h_\times(t-z)$$

and are called the plus and cross polarizations, respectively. The effect of these polarizations on an array of test particles is a stretching and compressing of the distances between the particles in the xy plane.

The general solution to (2.1) is analogous to the problem of solving for the vector potential given a source current, where here we solve for the metric perturbation given a stress-energy tensor describing the gravitational wave source. The result is an integral of the stress-energy tensor:

$$\bar{h}_{\mu\nu}(t, \vec{x}) = \frac{4G}{c^4} \int \frac{T_{\mu\nu}(t - \|\vec{x} - \vec{x}'\|/c, \vec{x}')}{\|\vec{x} - \vec{x}'\|} d^3x'$$

If the source is slowly moving and the gravitational wave energy is small, the solution can be simplified to

$$h_{ij}^{TT}(t, \vec{x}) \simeq \frac{2G}{c^4 r} \ddot{I}_{ij}^{TT}(t - r/c) \quad (2.2)$$

where I_{ij}^{TT} is the transverse-traceless projections of the quadrupole moment tensor of the source, $I_{ij}(t) = \int x_i x_j \rho(t, \vec{x}) d^3x$.

2.2 Sources of gravitational waves

Eq. 2.2 shows that any system whose quadrupole moment has a non-vanishing second derivative can generate GWs, which requires some non-spherically

symmetric motion of masses. The frequency of the waves is determined by the motion, usually some rotational frequency, of the masses. Generating strong waves requires high masses and speeds, so it is reasonable to look for GWs from astrophysical sources.

The simplest example is a binary system of massive, compact objects, such as neutron stars (NSs) or black holes (BHs). For most of their lifetime, the binary generates continuous GWs at a more-or-less steady frequency twice the orbital frequency of the components. Over time, the orbit decays due to the loss of energy to GW emission, causing the frequency and amplitude of the emission to increase. The time evolution scales with the component masses, so highly massive binaries, e.g. BBH systems, generate gravitational waves that sweep up in frequency and amplitude more quickly than less massive ones, e.g. BNS and NSBH systems. These compact binary coalescences (CBCs) creates a distinct GW signature, characterized by a relatively short-duration ($\lesssim 1$ s for BBHs, tens to hundreds of seconds for NSBHs and BNSs), high frequency ($\sim 10\text{-}1000$ Hz) chirp. These violent merger events happen very frequently, making them the best candidate for detecting gravitational waves with existing detectors.

Continuous GWs generated from binary systems may range from very low frequency (nanoHertz-range) waves from supermassive BH binaries, to milliHertz waves from stellar-mass galactic binaries, but in higher frequency bands the best candidates are isolated, rapidly rotating NSs (Riles, 2017). If such an NS exhibits any asymmetry with respect to its rotational axis, it generates GWs at twice its rotational frequency.

Burst sources of GWs are short duration events not generated by binary mergers; their time evolution is too difficult to model due to their unpredictable

or poorly understood dynamical behavior. Core-collapse supernovae are the most promising source to be detected, but their GW emission is still expected to be too weak for detecting events outside the galactic neighborhood. Other potential sources include magnetar flares and cosmic string cusps.

The superposition of all GWs forms a stochastic GW background analogous to the cosmic microwave background (CMB) (Christensen, 2018). This background is comprised of stellar-mass binary BH and NS mergers at frequencies currently observable by GW detectors, but at lower frequencies galactic white dwarf binaries and supermassive BH mergers would also contribute to the stochastic background. At cosmological distances, relic gravitational waves from the very early universe could be detectable via their effect on the polarization of the CMB radiation.

Thus far, 90 gravitational wave signals originating from CBCs have been detected by the LIGO-Virgo detector network (Abbott et al., 2021b). These include two BNS (GW170817 and GW190425) and two NSBH (GW200105 and GW200115) mergers. The rest are associated with BBH mergers. There has yet to be a detection of a GW signal with a significant probability of having a non-CBC origin.

2.3 Multi-messenger astronomy

In addition to being the first ever detection of a merger between two neutron stars, GW170817 ushered forth a new era of astronomy, making history as the first astronomical event observed by both gravitational waves and electromagnetic waves. Introduce what can be learned (basically summarize the intro to GRB section but also mention other targeted searches)

2.4 Gravitational wave detectors

Detection of gravitational waves requires measuring the transverse stretching and compressing of space. The earliest attempt at this was done through resonant mass detectors, solid, vibrationally isolated cylinders tuned to a particular frequency that could be used to detect the effect of gravitational waves on the length of the cylinders. These proved incapable of reaching the required sensitivity for detecting even the strongest gravitational waves in the frequency band they were designed for (~ 1 kHz).

The current era of GW detection is dominated by laser interferometers inspired by the simple Michelson interferometer. There are currently three observatories in operation: LIGO, consisting of LIGO Hanford Observatory (LHO) in Washington and LIGO Livingston Observatory (LLO) in Louisiana, and the Virgo observatory in Italy. Additional detectors in Japan (Kagra) and India (LIGO India) are under construction, and projects for next-generation detectors (Einstein Telescope, Cosmic Explorer, and the Laser Interferometer Space Antenna (LISA)) are on the horizon.

A GW interferometer is configured so that the anti-symmetric output port (the output not leading back to the laser source) observes no signal due to destructive interference. If one interferometer arm is elongated relative to the other, the phase shift between the signals from both arms results in some constructive interference; thus the amplitude of a gravitational wave passing through the plane of the detector arms is converted to an amplitude in laser light measured at the output port. Sensitivity scales with interferometer arm length, but the incredible length required to detect gravitational waves from even the loudest sources (100s of km in order to observe black hole binary mergers) makes this infeasible. However,

the effective arm length can be increased by making the arms Fabry-Perot cavities, so that light travels many (~ 300) times back and forth in each arm before leaving through the beam splitter.

Probably talk about a bunch of instrument noise sources here, e.g. shot noise and thermal noise.

Other methods that have been implemented to dramatically improve the sensitivity current-generation of Advanced LIGO (aLIGO) and Advanced Virgo include increasing laser power or injecting squeezed light to reduce the effects of photon shot noise; implementing a power recycling cavity at the symmetric output that sends the constructively-interfering light leaving in that direction back into the interferometer; implementing a signal recycling cavity at the anti-symmetric output to tune the most sensitive frequency band of the detector; larger mirrors to combat radiation pressure on the mirrors due to high laser power (Creighton & Anderson, 2011).

The two LIGO detectors began O1 on September 12, 2015, and made the first detection of a GW from a BBH merger on September 14, 2015 (Abbott et al., 2016b). This was followed by two more BBH detections before the end of the run on January 16, 2016 (Abbott et al., 2019). O2 began on November 30, 2016 after a period of detector upgrades and ended on August 25, 2017. During that run, in addition to several more BBH detections, the LIGO network (with the addition of Virgo towards the end of the run) observed the first BNS merger on August 17, 2017 (Abbott et al., 2017). O3, which spanned April 1, 2019 to March 27, 2020, came after another round of major improvements in the performance of the detectors (Buikema et al., 2020) and the full inclusion of the Virgo detector

in the GW network. Throughout O3, LIGO and Virgo observed a total of 74 GW events (Abbott et al., 2021a, 2021b).

CHAPTER III

METHODS FOR STUDYING ENVIRONMENTAL NOISE

Environmental noise refers to any signal originating from outside the structure of a GW detector that can impact the detector's sensitivity or disrupt its ability to achieve and maintain lock (Effler et al., 2015; Nguyen et al., 2021). The effects of environmental noise on detector sensitivity can range from persistent excess noise in the GW strain data, to short-duration transient signals, or *glitches*.

One goal of studying environmental noise is to aid in the validation of GW events. Due to the sophisticated nature of the search pipelines used to detect gravitational waves in the LIGO data, environmental glitches are highly unlikely to fully account for a GW event candidate. That said, glitches capable of influencing analyses occur frequently at both observatories.

Unlike instrumental noise, environmental noise can potentially be correlated between different detectors, i.e. stemming from a common source as opposed to stemming from chance coincidence. Such correlated noise is not accounted for in the estimation of false-alarm probabilities, which is done by time-shifting background data from each LIGO detector to produce long stretches of coincident background.

Environmental noise is particularly important in searches for unmodeled sources of gravitational waves, as these look for excess power without the use of waveform templates. Even for highly significant CBC events, contamination of the strain data can be detrimental to parameter estimation analyses that infer source properties from the morphology of the event. Thus it is important to have a quantitative solution for identifying and evaluating the impact of environmental transients when they occur coincide with candidate events.

The second goal is to improve the sensitivity and performance of the detector by localizing noise sources and coupling mechanisms. Once tracked down, they can be mitigated by eliminating the noise sources, attenuating the propagation of the signal, or modifying the detector itself.

In this chapter I introduce sources of environmental noise (Section 3.1), describe the hardware used to monitor those noise sources (Section 3.2) and to reproduce them experimentally (Section 3.3). I present a method for quantifying environmental coupling (Section 3.4), then I discuss the limitations (Section 3.5) and software implementation (Section 3.6) of the method.

3.1 Sources of environmental noise

The environment can influence a GW detector through physical contact (via vibrations or temperature fluctuations), electromagnetic waves, static electric and magnetic fields, and possibly high-energy radiation.

Vibrations can affect the data by directly moving the test masses. They can also modulate the paths of light that is scattered off of surfaces inside the interferometer, or introduce fluctuations in the alignment of the input laser beam.

3.2 The PEM sensor array

Understanding environmental influences on the detectors requires comprehensive monitoring of its physical surroundings. This is done through the physical environmental monitoring (PEM) system of auxiliary sensors (Figure 1), which consists of accelerometers for high-frequency vibrations (between tens to thousands of Hertz), seismometers for low-frequency vibrations (up to tens of Hertz), microphones, magnetometers, voltage monitors that measure the

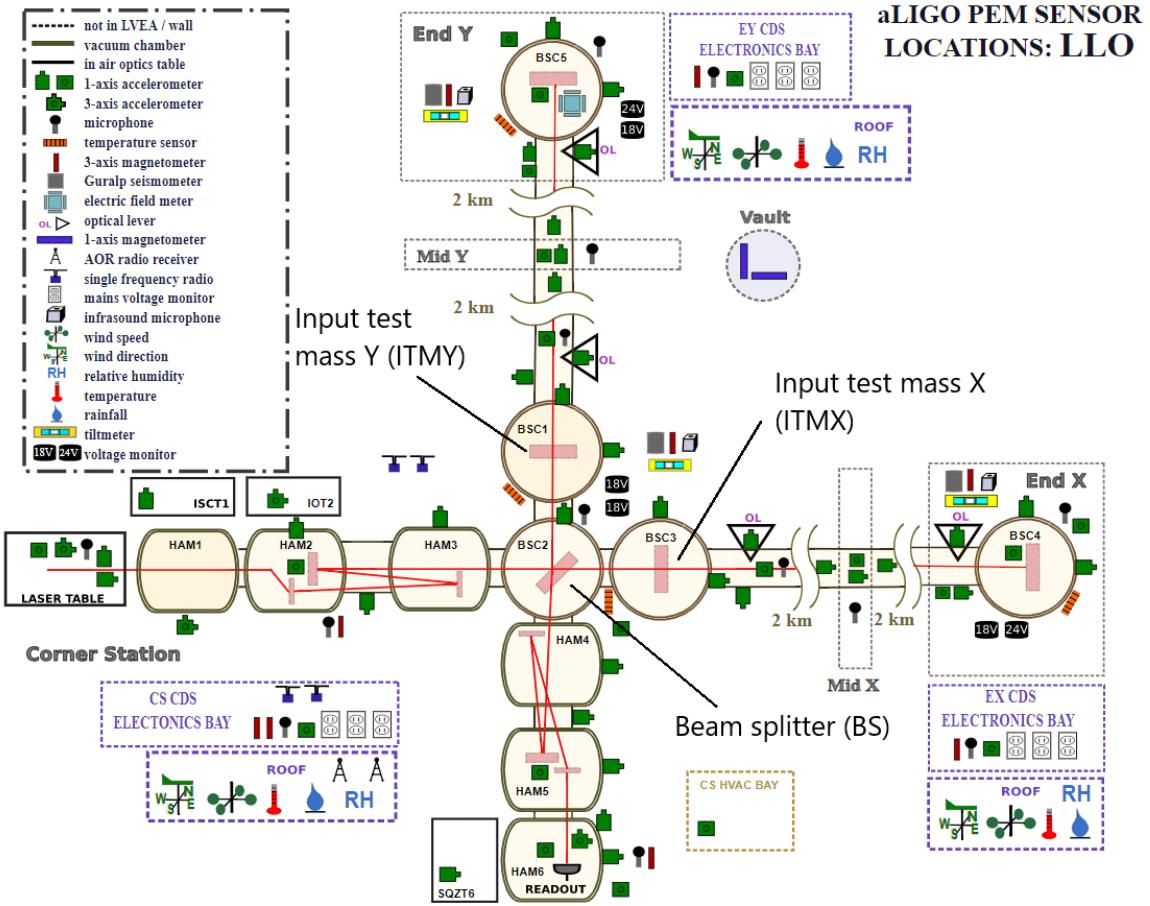


Figure 1. The PEM system layout at LLO during O3, as seen on the PEM public website.

voltage of electric power supplied to the detector sites, radio-frequency (RF) receivers, a cosmic-ray detector for high-energy particles, and wind, temperature and humidity sensors. Detailed information on PEM sensors, including example background spectra and calibration data, can be found on the PEM website, PEM.LIGO.org (R. Schofield, Effler, Nguyen, et al., 2021).

Most sensors produce an analog signal that is converted to a digital signal via an analog-to-digital converter (ADC), then processed by a data acquisition system and saved into frame files. The frame file are the primary data format for GW detector data (Blackburn et al., 2019) and are remotely accessible via

the LIGO grid computing clusters located at the observatories as well as various LIGO-affiliated institutions. From here on, I will use *sensor* to refer to the physical monitoring device and *channel* to refer to the data stream as read from frame files.

3.2.1 Monitoring the monitors with `ligocam`

Maintaining an extensive, ever-expanding array of auxiliary sensors requires a scalable solution for monitoring their behavior. Frequent changes to the sensors, as well as of interferometer hardware, result in changes in ambient spectral characteristics. Worse yet, sensors, their power supplies, their ADCs, and the data acquisition system that process their signals all have the potential to malfunction, leaving blind spots in our ability to detect signals that may affect the GW strain data.

The LIGO channel activity monitor, or `ligocam`, is a program that checks for various signs of unusual sensor behavior using a number of spectral cues (Taluker, Nguyen, et al., 2021). The goal of `ligocam` is to produce human-readable summaries of sensor behavior for the entire PEM network at an observatory, and send email alerts to experts when significant malfunctions are identified.

Leveraging the LIGO data grid computing clusters, `ligocam` parallelizes its analysis by splitting the full channel list (over 100 channels), processing only five per job. This allows it to run hourly, scheduled by a `cron` job, outputting a summary page and transmitting email alerts (if applicable) within minutes. A `cron` job is scheduled at each of the LIGO observatories. Although such expediency is not necessary for fixing malfunctioning hardware, it is important to have up-to-date

channel status when validating recently detected GW event candidates, as discussed in Section 4.4.

The program reads in time series data using the `gwpy` library (Macleod et al., 2021) and converts each time series to an amplitude spectral density (ASD). When run for the first time, `ligocam` saves the current ASDs as reference ASDs. On each subsequent run at a later time t , the new ASD $X_t(f)$ is compared to the reference $\bar{X}_{t-1}(f)$ to determine if there is anomalous behavior present. If $X_t(f)$, shows no anomalous behavior relative to $\bar{X}_{t-1}(f)$, then the reference is updated as an exponential average of all past non-anomalous ASDs:

$$\bar{X}_t(f) = \begin{cases} X_0(f), & t = 0 \\ \alpha X_t(f) + (1 - \alpha) \cdot \bar{X}_{t-1}(f), & t > 0 \end{cases} \quad (3.1)$$

where α^{-1} is the averaging decay length. If `ligocam` is run hourly with $\alpha^{-1} = 1/6$, for instance, then $\bar{X}_t(f)$ is roughly the average of the past six hours of channel data. If any anomalous behavior is detected, it is reported on the output HTML page and (if sufficiently egregious) via email to the relevant parties, and $\bar{X}_t(f)$ is not updated.

Anomalies are identified and reported in a number of ways. Noise far above or below the reference is a sign of changes to hardware or infrastructure in the vicinity of the sensor, which may not adversely affect coverage of the area but could adversely affect noise estimates as described in Section 4.4. If $X_t(f)$ is many orders of magnitude below the reference but still above the electronic noise background of the ADC, the sensor is considered to be faulty. Some sensors, especially magnetometers, can have background noise levels low enough to be near the ADC noise floor. For this reason, checks for magnetometers are performed specifically on

the 60 Hz mains signal, while other sensors (accelerometers, microphones, etc.) are analyzed for broadband variations. If the signal is even weaker than the electronic noise background, then the data acquisition system is blamed for the failure.

3.3 Environmental noise injections

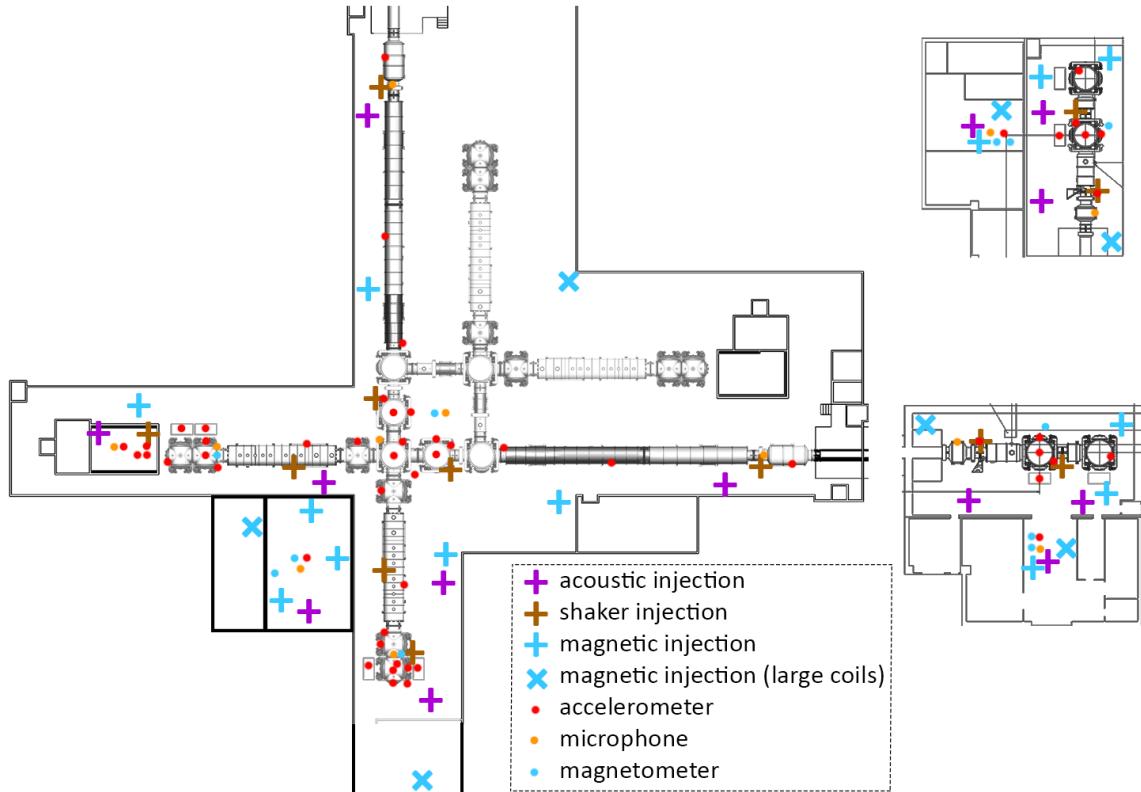


Figure 2. Standard locations for vibration and magnetic injections at the LHO corner station (left), Y end station (top right), and X end stations (bottom right).

The effect of environmental influences on the sensitivity of a GW detector can be studied by making noise *injections*. These are signals produced by human-operated sources with the intention of replicating environmental disturbances with sufficient amplitude to produce excess noise in the differential arm length (DARM) spectrum. The amplitude of the excess, combined with measurements of the input signal, can be used to quantify the coupling behavior (Section 3.4). The

most common examples are acoustic injections, generated using speakers, seismic injections generated by vibrational shakers, and magnetic field injections generated by electrical current loops.

At each observatory we inject from 13 locations with acoustic injections, about 12 with shaking injections, and 15 with small-coil magnetic injections, with 7 large-coil magnetic injection locations planned for the fourth observing run (O4) (as explained in Section 3.3.2). The number and locations of shaker injections vary between injection campaigns. For all injection types, multiple injections are made at each location in order to focus on different frequency bands. Additionally, impulse injections (not shown) are made at locations where vibrational injections have revealed strong coupling sites.

Table 1. Specifications for injection equipment.

| Equipment | Injection type |
|--|----------------|
| Custom enclosure with two 14-in. speakers | Acoustic |
| Various smaller speakers | Acoustic |
| APS 113 Electro-Seis®Long Stroke Shaker (APS Dynamics, 2014) | Vibrational |
| Piezosystem® (Piezosystems, n.d.) shaker with custom reaction mass | Vibrational |
| Brüel & Kjær® (Brüel & Kjær, 2021) EM shaker with custom reaction mass | Vibrational |
| 1 m diameter copper coil (100 turns) | Magnetic |
| 3 x 3 m and 5 x 5 m coils (80-100 turns) | Magnetic |

Injection locations are chosen to best mimic disturbances from outside the detector (Figure 2). To do so we choose them to be as far from the detector and environmental sensors as possible, but we are usually limited by the size of the detector sites themselves (some injections can be made from outside). Time dedicated to these tests has to be balanced against other instrumental work and

observing time, which leads to a trade-off between measurement uncertainty and coverage. We perform injections from as many locations as time allows in order to maximize coverage of potential coupling sites. Increased time allocation toward environmental studies in recent years has allowed for a significant increase in the number of injection locations. Table 1 summarizes the current equipment used and Figure 3 shows photos of some of the equipment.



Figure 3. Injection equipment photos. From left to right: wall-mounted magnetic field injection coil; 14-in. speakers; APS 113 shaker connected to the door of a vacuum chamber by a rigid fiberglass rod; modified Piezosystem shaker clamped to an electronics rack; modified B&K shaker clamped to a beam tube support.

3.3.1 Vibrational injections

Acoustic injections are produced by large speakers. For the corner stations, a pair of speakers mounted on a vibrationally isolated cart (to minimize ground-based vibrational signals) are used. Typically, the injection signal is white noise band-passed between 20-2000 Hz, with narrower bands being used for special follow-up of particular coupling sites.

Seismic injections at low frequency (up to tens of Hertz) during Initial LIGO (iLIGO) were performed with small electromagnetic and piezoelectric shakers and a weighted cart. A large shaker has been used since the beginning of

noise studies for O3. The large shaker can impart up to 133 N of sine force and a peak-to-peak displacement of 158 mm (APS Dynamics, 2014), compared to the electromagnetic shaker which imparts up to 45 N of force and a displacement of 8 mm (Brüel & Kjær, 2021). While smaller shakers can be directly clamped to the interferometer supports, a rigid fiberglass rod is used to connect the large shaker to the interferometer. This has an added benefit of being better able to adjust the direction of the actuation by angling the rod and shaker accordingly.

Two new injection techniques have been developed for localizing vibration coupling sites connected to the vacuum enclosure, such as locations on the vacuum enclosure that reflect scattered light. The techniques rely on the slow propagation speeds (hundreds of meters per second) of vibrations on the steel vacuum enclosure walls or, for acoustic injections, in air. These two techniques were essential in the localization of coupling sites in O3, as discussed in Section 4.2.

3.3.1.1 Beating-shakers technique.

The beating-shakers technique is narrow-band, and involves vibrating the vacuum enclosure at two slightly different frequencies, each injected from a shaker or a speaker at a different location (e.g. a shaker at one location injects a sine wave at frequency f and a shaker at the other location injections at frequency $f + 0.01 \text{ Hz}$). The two injections are adjusted in amplitude to produce strong beats in the GW channel.

Because the injection locations are different, the relative phase of the two injected signals varies with location on the vacuum enclosure. As a result, the phase of the beat envelope varies with position, and different sites experience maximum chamber wall motion at different times. The sites with accelerometer signals that have the same beat envelope phase as the DARM signal are candidates

for the scattering sites on the vacuum enclosure walls. Other sensors that are not near the coupling site may also match the phase by chance, but these false positives can be rejected by varying the locations of the shakers.

3.3.1.2 Impulse injections.

The second injection technique, which is broad band, involves propagation delays in impulse injections. Impulse injections are performed by striking the vacuum enclosure directly with enough force to produce a transient in the GW channel and in nearby accelerometers. The vibrational impulse propagates through the structure of the vacuum enclosure, arriving at different accelerometers and coupling sites at different times.

We can distinguish these arrival times because the propagation velocity is much slower than in solid material, and is only roughly 300 m/s in our case. Using time series plots, the arrival time of the impulse in the GW channel is compared to the arrival time of the impulse in multiple accelerometers. The accelerometers that show the same arrival time as in the GW channel are more likely to be near a coupling site than those that observe the impulse much earlier or later. Again, varying the location of the injection eliminates sensors that match the detector time-of-arrival by chance but are actually far from the coupling site.

An additional consistency check is that the coupling of accelerometers near the coupling site will vary less between different impulse locations than that of accelerometers far from the coupling site. Finally, if the accelerometer is at the coupling site, the impulse in the GW channel will have a resonance structure that is similar to the resonance structure of the accelerometer signal, which can be judged from spectrograms.

3.3.2 Magnetic injections

Improvements have also been made to the magnetic field injection equipment. In order to generate fields strong enough to couple into the GW channel using the 1 m magnetic field coils built during iLIGO, we must focus the power of the coil into narrow bands and combs instead of injecting broadband signals. This was sufficient in iLIGO when strong magnetic coupling occurred primarily through permanent magnets. However, due to the removal of permanent magnets from the test masses, coupling from those sources has decreased and cables and connectors have become the dominant coupling sites above about 80 Hz, introducing more structure to the coupling functions and requiring stronger injections.

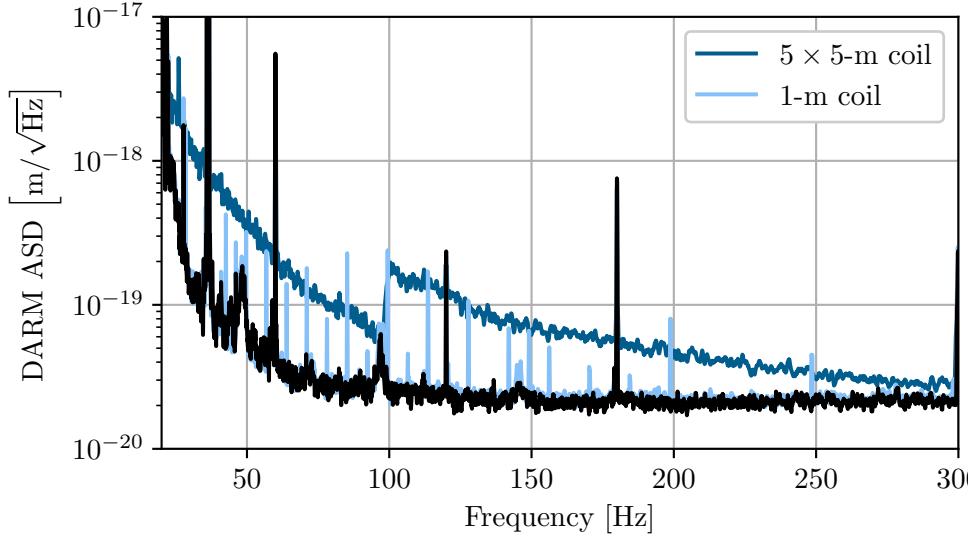


Figure 4. Comparison of the old small-coil comb magnetic field injections with the new large-coil broadband injections.

To achieve high-amplitude broadband magnetic injections, seven wall-mounted coils, each one a 3 m x 3 m or 5 m x 5 m square of 80-100 turns, are being installed at each site; three at the corner station and two at each end

station. These coils are fixed in place and can be operated remotely, allowing for weekly injections to monitor variations in magnetic coupling caused by changes to electronics. Figure 4 compares the old and new magnetic injections. Some coils were installed and operated at the sites during O3 (discussed in Section 4.3); the project will be completed by the start of O4.

3.4 Coupling functions

Goal here is to give much more mathematical rigor to coupling function model than was provided in the relatively terse CF section of the PEM paper. Still trying different notation conventions. Hard to choose a convention for the many variables and indices that isn't confusing, since there are many similar but not identical quantities.

3.4.1 Single coupling site, sensor, and injection

Suppose there exists exactly one coupling site, i.e. one location at which incident environmental signals result in excess noise in the GW strain data. Suppose also that a sensor is placed at the location of the coupling site, and a noise injection is performed that produces a signal observable by the sensor and the interferometer readout. The coupling mechanism can be modeled in the frequency domain as a linear system:

$$h(f) = C(f)x(f), \quad (3.2)$$

where $h(f)$ is the ASD of the detector (strain) response, x is the ASD of the injection signal as measured by the sensor, and $C(f)$ is the *coupling function*, which represents the amplitude of gravitational wave strain noise per unit amplitude in

the sensor. By convention, the strain is typically converted to DARM, in meters, in which case $C(f)$ represents test mass displacement per unit of sensor amplitude. If the injection signal is an acoustic signal and the sensor is a microphone measuring amplitude in Pa, for instance, then the acoustic coupling function is in units of m/Pa.

In both the witness sensor and the detector, some ambient background noise is always present whether or not an injection is produced. Let $h_{\text{bkg}}(f)$ and $h_{\text{inj}}(f)$ be the ASDs of the detector during a period of background noise and during the injection, respectively. Likewise let $x_{\text{bkg}}(f)$ and $x_{\text{inj}}(f)$ be the background and injection ASDs of the sensor. Since the noise adds linearly in the power spectral domain, the actual signal in each is the difference between the injection-time and background-time power spectral densities (PSDs):

$$[h(f)]^2 = [h_{\text{bkg}}(f)]^2 - [h_{\text{bkg}}(f)]^2 \quad (3.3)$$

$$[x(f)]^2 = [x_{\text{bkg}}(f)]^2 - [x_{\text{bkg}}(f)]^2 \quad (3.4)$$

Combining eqs. (3.2)–(3.4), we can measure the coupling function from the background and injection ASDs (Kruk & Schofield, 2016; Nguyen, 2020):

$$C(f) = \sqrt{\frac{[h_{\text{inj}}(f)]^2 - [h_{\text{bkg}}(f)]^2}{[x_{\text{inj}}(f)]^2 - [x_{\text{bkg}}(f)]^2}}. \quad (3.5)$$

The value of a coupling function at a single frequency bin is referred to as a *coupling factor*.

3.4.2 Multiple coupling sites, sensors, and injections

Suppose now there are multiple coupling sites, and a sensor is placed at the location of each site. The detector response to an environmental signal now becomes a linear combination of the sensor signals and their sensor-specific coupling functions:

$$h(f) = \sum_{j=1}^m C_j(f)x_j(f), \quad (3.6)$$

Solving for the coupling function now would require producing multiple injections instead of just one, resulting in a system of n equations with m unknown coupling functions, where n and m are the numbers of injections and sensors, respectively:

$$h_i(f) = \sum_{j=1}^m C_j(f)x_{ij}(f). \quad (3.7)$$

Here $h_i(f)$ is the detector response during injection i , $x_{ij}(f)$ is the amplitude measured by sensor j during injection i , and $C_j(f)$ is the coupling function of sensor j . If $n = m$, eq. (3.7) could be solved to determine the coupling functions of all sensors.

Thus far it has been assumed that the witness sensors are placed precisely at the locations of the coupling mechanisms, but such perfect placement is not realistically feasible given that there are an unknown number of coupling sites at unknown locations. A sensor, even if it is near a coupling site, only measures the injection amplitude at its own location, not at the coupling location. Therefore, when using real-world sensors, eq. (3.4) is not exact, so eq. (3.5) does not exactly describe the coupling at the coupling site. Nevertheless, as explained above, sensors are distributed in order to maximize coverage of coupling sites and this has been

sufficient for producing reliable coupling functions for all sensors, as discussed further in Section 3.5. As seen in Figure 5, the sensor coupling function $C_j(f)$ is a good approximation to the true coupling $C(f)$ if the injection distance from the sensor r_{ij} is much greater than the distance between the coupling site and sensor. This allows us to estimate excess noise in the GW channel $h_i(f)$ without directly measuring the coupling actuation $A_i(f)$.

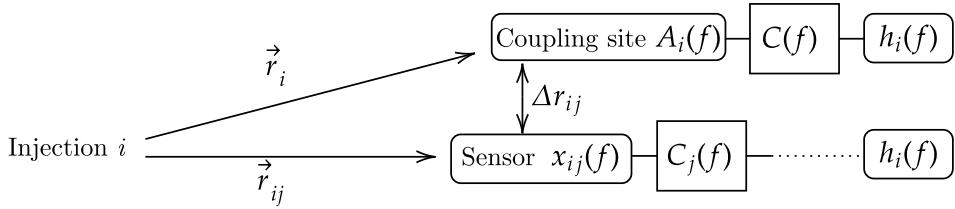


Figure 5. Diagram of a coupling function measurement in the trivial case. An injection i produces a signal $A_i(f)$ at a coupling site, which cannot be measured directly. If the injection is performed sufficiently far, then $\Delta r_{ij} \ll |\vec{r}_{ij}|$, so $C_j(f) \approx C(f)$.

In Figure 6 we can see how complex the experiment system becomes when there are multiple coupling sites of unknown location. If there are two coupling sites A and B , at least two injections i_1 and i_2 are required to distinguish between the coupling mechanisms. The GW channel response to injection i_1 is therefore the sum of the contributions of the actuation signals $A_{i_1}(f)$ and $B_{i_1}(f)$, and likewise for injection i_2 . Again, at best we merely approximate the actuation signals $A_i(f)$ and $B_i(f)$ with sufficiently close sensors, although we do not know *a priori* which sensors are near coupling sites.

3.4.3 Solving the coupling equations

One hurdle remains in attempting to solve eq. (3.7). In practice, typically $n < m$ due to logistical constraints on the number of injections one could

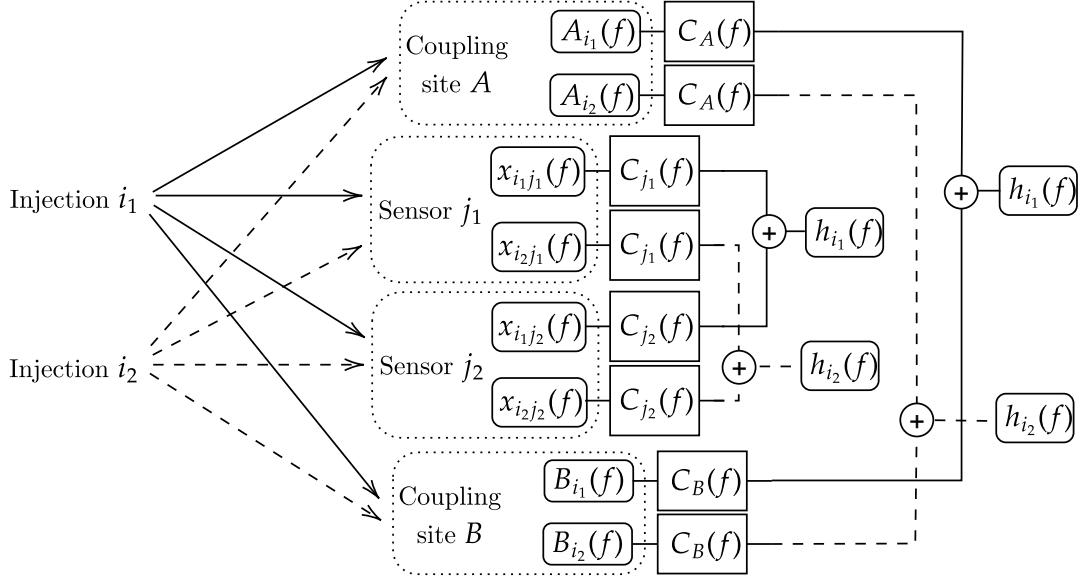


Figure 6. Expanding Figure 5 to two injections i_1 and i_2 , two sensors j_1 and j_2 , and two coupling sites A and B . Each injection response $h_i(f)$ is itself a sum across contributions from A and B . The goal is to determine the sensor couplings $C_j(f)$ that reproduce $h_i(f)$ using the sensor measurements $x_{ij}(f)$ since the true coupling actuations $A_i(f)$ and $B_i(f)$ are not known.

perform during a realistic time window, which makes the system of equations underdetermined. A straight-forward least-squares regression is therefore not always feasible. Below are two approximation methods for determining $C_j(f)$ for all sensors.

Nearest-sensors approximation. One method of forcing $n = m$ is reducing the number of sensors in each equation, by asserting $x_{ij}(f) = 0$ for sensors that are sufficiently far from the source of injection i . This can be done by ordering the sensors by distance from the injection source and applying the assertion to the $m - n$ farthest sensors. Issues can arise if there are sensors that are never near enough to any injection source, causing them to zeroed out for all injections; this requires that injections be distributed such that each sensor is near enough to at least one injection.

Nearest-injection approximation. Instead of solving eq. (3.7) in full, one can approximate $C_j(f)$ for each sensor independently of other sensors. Given a sensor j , eq. (3.5) can be repurposed by replacing x with x_{ij} and h with h_i to compute a single-injection “coupling function” $\mathcal{C}_{ij}(f)$ for each injection:

$$\mathcal{C}_{ij}(f) = \sqrt{\frac{[h_{i,\text{inj}}(f)]^2 - [h_{i,\text{bkg}}(f)]^2}{[x_{ij,\text{inj}}(f)]^2 - [x_{ij,\text{bkg}}(f)]^2}}. \quad (3.8)$$

The closer an injection is to a sensor i , the more accurately $\mathcal{C}_{ij}(f)$ approximates $C_j(f)$, since the detector response would be dominated by coupling near sensor j . Therefore one can construct the sensor coupling function by choosing at each frequency bin the coupling factor corresponding to the nearest injection, determined by the highest sensor amplitude (using the assumption that injection amplitudes are equivalent). That is, for a frequency f_k and a set of injections \mathcal{I} , one can measure the sensor amplitudes $\{x_{ij}(f_k) \mid i \in \mathcal{I}\}$, compute the single-injection coupling functions $\{\mathcal{C}_{ij}(f_k) \mid i \in \mathcal{I}\}$, and construct the approximate sensor coupling function

$$\tilde{C}_j(f_k) := \mathcal{C}_{lj}(f_k) \text{ where } l = \underset{i \in \mathcal{I}}{\operatorname{argmax}} (x_{ij}(f_k)). \quad (3.9)$$

If the distribution of injection locations provides sufficient coverage of sensor locations, then $\tilde{C}_j(f) \approx C_j(f)$. Shortcomings of this assumption are discussed in Section 3.5.

Figure 7 provides an example of a single-injection coupling function measurement for a pre-stabilized laser (PSL) acoustic injection. Figure 8 shows an estimated ambient noise based on an accelerometer coupling function constructed from five single-injection coupling functions. For simplicity only five injections

were used to produce this example, however in practice the number of injections performed near a sensor can be much higher.

Due to hardware limitations it can be possible for an injection signal to be strong enough to produce excess noise in a sensor ASD but not in the GW detector ASD. For frequency bins where this is the case, an upper limit on $\mathcal{C}_{ij}(f_k)$ can be established by assuming, as a worst-case scenario, that all of the detector noise at that frequency is produced by the coupling alone

$$\mathcal{C}_{ij,\text{UL}}(f) = \frac{h_{i,\text{bkg}}(f)}{\sqrt{[x_{ij,\text{inj}}(f)]^2 - [x_{ij,\text{bkg}}(f)]^2}}. \quad (3.10)$$

The larger the injection amplitude, the better this upper limit can be constrained. The boundaries between measurements, upper limits, and null results are established by two ASD ratio thresholds: a sensor threshold and a detector threshold. Let $r_x := x_{ij,\text{inj}}(f)/x_{ij,\text{bkg}}(f)$ and $r_h := h_{i,\text{inj}}(f)/h_{i,\text{bkg}}(f)$ represent the injection signal-to-noise ratios if the sensor and GW detector ASDs, respectively. If $r_x \geq t_x$ and $r_h \geq t_h$, where t_x is the sensor threshold and t_h is the detector threshold, then a measurement is computed via eq. (3.8). Otherwise, if $r_x \geq t_x$ but $r_h < t_h$, then eq. (3.10) is used to place an upper limit on the coupling. If $r_x < t_x$ and $r_h < t_h$, then neither a measurement nor upper limit is computed. The null hypothesis is thus assumed:

$$\mathcal{C}_{ij,\text{null}}(f) = \frac{h_{i,\text{bkg}}(f)}{x_{i,\text{bkg}}(f)}. \quad (3.11)$$

The values of t_x and t_h are determined based on typical level of random fluctuations observed in the spectra, but often values of $t_x = 10$ and $t_h = 2$ are used for most types of sensors and injections. The higher choice of t_x is due to the

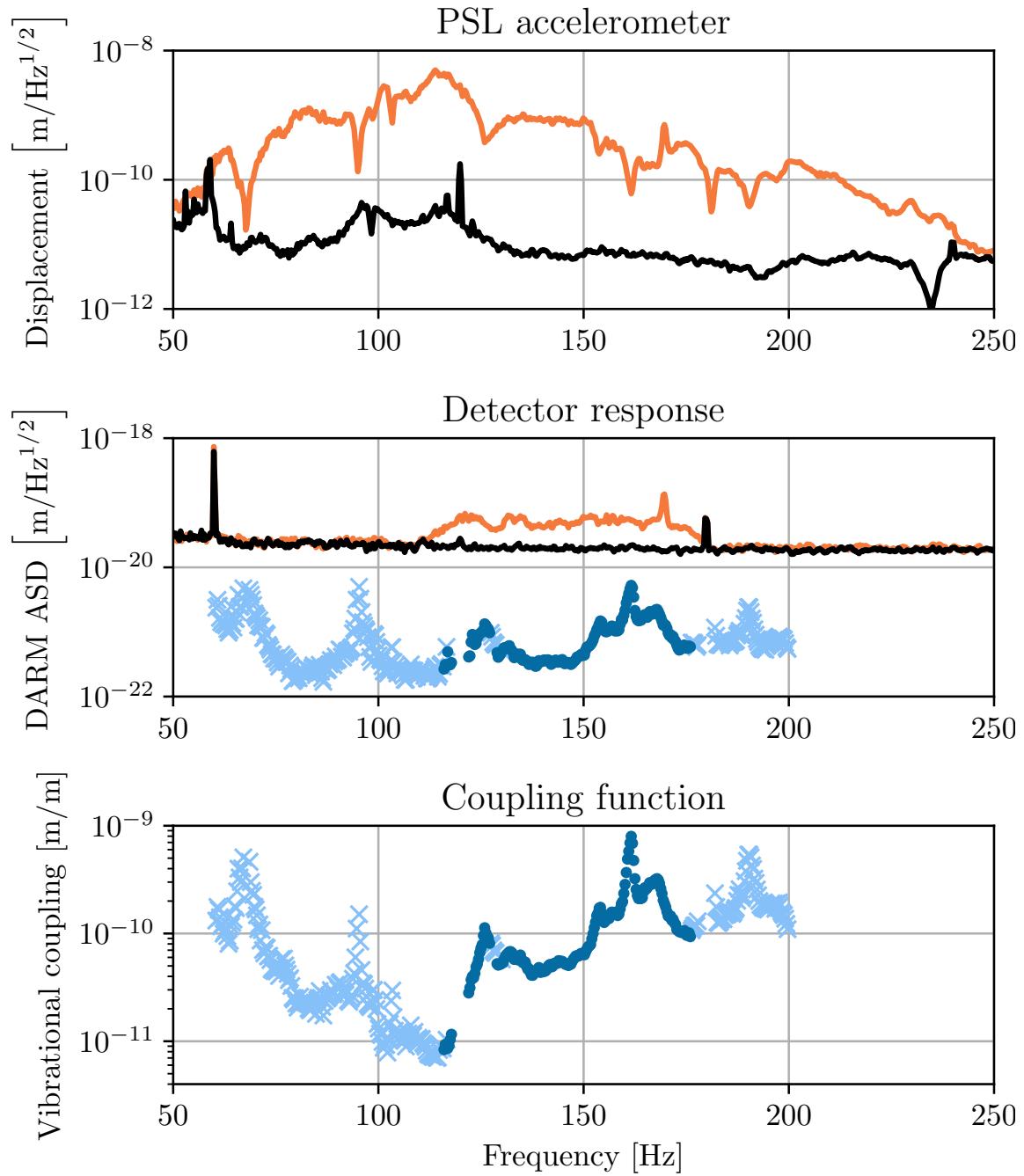


Figure 7. Example of a broadband acoustic noise injection and measurement of a single-injection coupling function. Top: displacement of an accelerometer in the PSL room during background time (black) and injection time (orange). Middle: DARM during background time (black) and injection time (orange). Estimated ambient levels for the accelerometer are shown as dark blue dots, with upper limits shown as light blue crosses. Bottom: single-injection coupling function used to produce estimated ambient above.

environmental sensors being much more sensitive to random fluctuations in the ambient noise level than the interferometer is.

The coupling function as approximated in eq. (3.9) is used for comparing coupling between different sensor locations and producing estimates of interferometer noise levels, e.g. as part of event validation (see Section 4.4). References to a sensor’s coupling function will hereafter refer to this approximate quantity. Figure 8 provides an example of an estimated ambient for an accelerometer on the HAM6 vacuum chamber (which houses the interferometer output optics). The PEM website provides coupling functions for all accelerometers, microphones, and magnetometers produced from the most recent campaign of injections (R. Schofield et al., 2021).

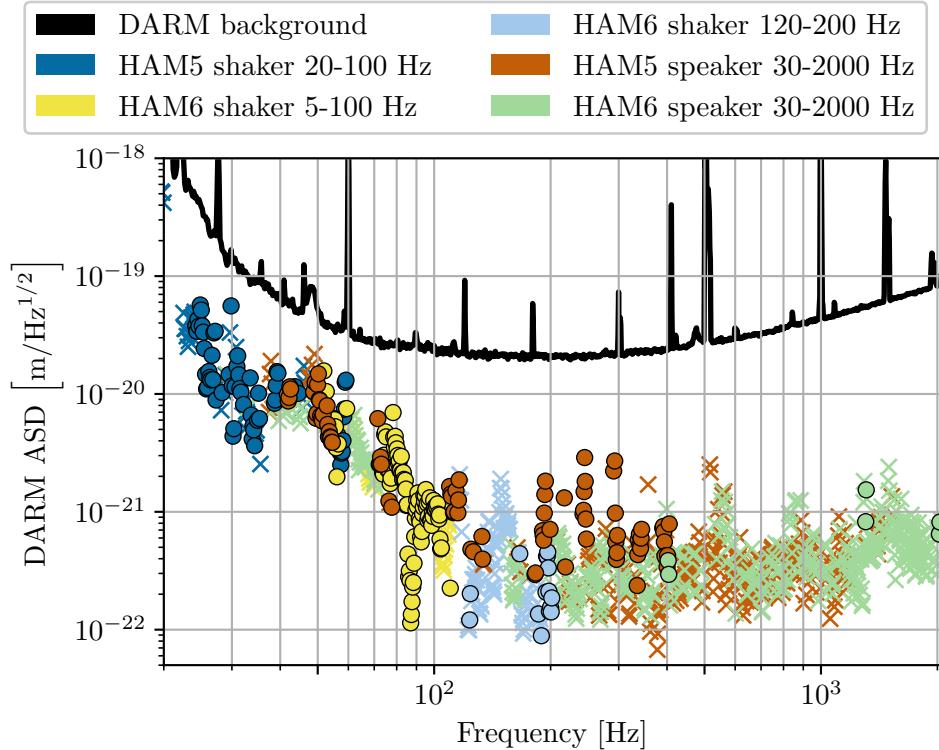


Figure 8. Ambient noise level for the LHO HAM6 Y-axis accelerometer estimated from a composite coupling function, using acoustic and seismic injections near the output arm.

3.5 Uncertainties and limitations of coupling functions

3.5.1 Comparison to transfer functions

Environmental coupling is characterized using coupling functions instead of transfer functions because perfect coherence is not assumed in the system. Low coherence can arise either due to non-linearity in the coupling or due to the spacing between the sensor and coupling site. On a superficial level, a coupling function lacks a phase response component, representing only the magnitude response in the system. Coupling functions also differ fundamentally from transfer functions in the sense that they do not assume the input signal to be the true actuation signal, but rather merely a witness of the actuation, while the actuation is in fact occurring at the location of the true coupling site.

3.5.2 Assumptions about coupling mechanisms

Equation (3.5) relies on two assumptions about the coupling mechanism. First, the coupling is assumed to be linear in amplitude, e.g. doubling the amplitude of the injection would double the amplitude of the GW detector response. This is confirmed when performing injections by repeating them with different amplitudes and ensuring that the detector response scales proportionally with the injection amplitude. Second, the coupling function ignores any up- or down-conversion of the signal between the sensor and the GW detector. Such non-linear coupling can be very significant for scattering noise and bilinear coupling, but is not accounted for in the estimates of linear coupling. One way to check for non-linear coupling is by sweeping single frequency injections over time and searching for off-frequency responses in the GW detector. Frequency changes

from non-linear coupling can be an issue in broadband injections where up- or down-converted noise in the interferometer readout appears in the injection band, resulting in artificially higher estimates at those frequencies. We split broadband injections into smaller frequency bands to avoid this effect when necessary. One approach for quantifying non-linear coupling is presented in Washimi et al. (2020).

3.5.3 Hardware limitations

Injection amplitudes. To measure coupling, we inject signals large enough to produce a response in the detector, but the maximum amplitude of injections is limited by the sensitive range of the environmental sensors (saturation produces an overestimate of coupling). This effectively limits how far below the detector noise background we can probe for coupling or establish upper limits.

Recall that eq. (3.9) was based on assuming that injection amplitudes are equivalent. However, this assumption is ambitious: since injections vary with distance to sensors, the amplitudes used have to be adjusted to achieve a large signals in the sensors and in the GW channel. This means that the highest-amplitude injection measured by a sensor is not necessarily the nearest injection to that sensor. If a further injection was performed using a much larger amplitude, its measured amplitude can trump that of a nearer injection, leading to the algorithm choose a more distant injection source location when determining \tilde{C}_j . To prevent this issue, once C_{ij} is computed for all sensors for a single injection, an additional sensor threshold is applied that is a fraction of the highest amplitude observed by all sensors. This threshold is used only to demote measurements to upper limits. By doing so, this injection-dependent threshold vetoes measurements produced by sensors that are far from the injection. For the O3 injections, the threshold was set

at one third, meaning that a coupling measurement is demoted to an upper limit if the sensor measuring it observes the injection at less than a third of the amplitude observed by the sensor closest to the injection.

Need to talk about this a lot more. Injection-depended thresholding reflects a fundamental issue with the distribution of sensors and injections. In principle if you have many one-to-one correspondence b/w sensors and injections, you would just have each injection produce measurements in its nearest sensor, which is effectively what you get by setting this threshold to one. But in absence of that ideal injection distribution you have to lower the threshold by some amount and it's not obvious how to operationalize this. Nevertheless some even as an ad hoc thresholding scheme it works very well for preventing inaccurate coupling functions from misattributing the nearest injection to a sensor.

I need to emphasize that this is basically a way of normalizing injections.

Uncertainty due to injection locations. As mentioned above, the model in eq. (3.7) relies on the assumption that the environment is monitored at the coupling site. The density of sensors is not great enough for this to be strictly true, especially if the source of the environmental signal is closer to the coupling site than the sensor is. The detector response to an injection depends on the distance between the injection and coupling site, whereas the sensor response depends on the distance between the injection and sensor. Varying the injection location therefore varies the relative scaling of the numerator and denominator of eq. (3.8), affecting the measurement of $\mathcal{C}_{ij}(f)$ and subsequently the sensor coupling function via eq. (3.9). Therefore, a finite spacing of sensors leads to some degree of uncertainty in the coupling functions. This uncertainty also propagates to projected noise levels in the GW channel using these coupling functions.

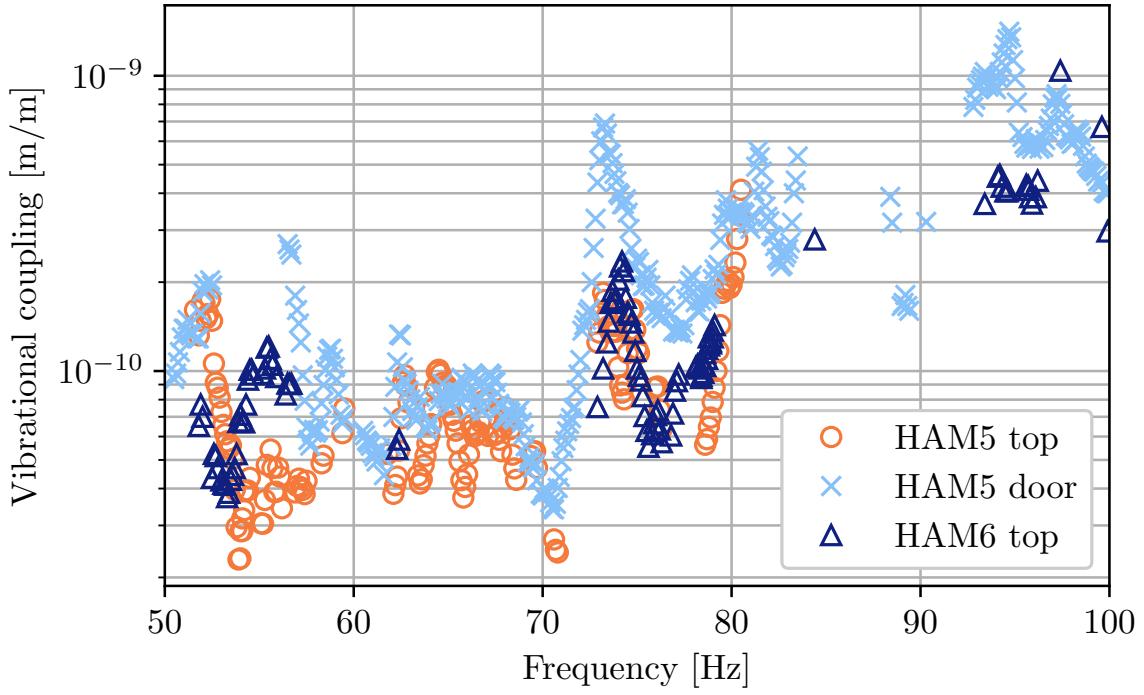


Figure 9. Single-injection coupling functions (upper limits not shown) for the HAM5 Y-axis accelerometer for three different shaker injection locations (on top of HAM5, on top of HAM6, and on the HAM5 chamber door).

Since the uncertainty manifests as a multiplicative scaling of $\mathcal{C}_{ij}(f)$, it can be described by computing a geometric standard deviation of $\mathcal{C}_{ij}(f)$ for a single sensor over a range of injection locations, at each frequency bin. Figure 9 shows single-injection coupling functions for an accelerometer measured from shaker injections produced from three locations (the distribution of injection locations is discussed in Section 3.3. Since the injection locations are close enough to the accelerometer, it can be assumed that the variance is primarily due to finite spacing effect. Averaged across all frequency bins, the geometric standard deviation between injection locations is 1.4, i.e. coupling functions measured from vibrational injections can be expected to vary by a factor of 1.4 when measured by different injection locations.

Figure 10. Single-injection coupling functions (upper limits not shown) for various magnetometers and various injections at both observatories.

A similar study was performed combining geometric standard deviations for various magnetometers at both observatories. There are fewer magnetic injection locations to use for the comparison, but since coupling can be measured at each station (a corner station and both end stations) at each of the two LIGO observatories, there are twelve magnetometers that can be used, each with two or more injections nearby. The result of this study is that magnetic coupling measurements and noise projections vary by a factor of 1.7. This is slightly greater than that of vibrational measurements, since the lower number of magnetometers means that the distances between coupling sites and sensors is greater, amplifying the finite spacing effect.

For both vibrational and magnetic coupling, these estimated uncertainties are acceptable given that conclusions made from coupling functions are often more qualitative than strictly quantitative, i.e. identifying and localizing coupling mechanisms is more important than precise estimates of the detector response. That said, more precise noise estimates may become important for quantifying the impact of environmental transients on GW event candidates, as discussed in Section 4.4.

Nodal artifacts from acoustic injections. In the case of acoustic injections, the uncertainty in a coupling function can be exacerbated when nodes and anti-nodes in the acoustic signal coincide with the location of a sensor but not a coupling site. This results in peaks and troughs in the sensor spectrum at frequencies that have a node or anti-node at the sensor location, respectively. These artifacts can impact any sensor, but are more noticeable in microphone spectra

than accelerometer spectra, possibly because the stiffness of the vacuum enclosure results in effectively averaging over a larger area; in microphones, the peak-to-trough ratio is typically a factor of a few. The peaks and troughs are present in the sensor but not in the detector spectrum, because the sensor monitors a single point whereas the coupling to the interferometer is spread across a large enough area for the effects of nodes and anti-nodes to average out. Consequently, this effect imprints troughs and peaks onto the coupling function.

The artifacts can be smoothed out of the spectra by applying a moving average over $x_{ij,\text{inj}}(f)$ before computing $\mathcal{C}_{ij}(f)$. The moving average window must be on the scale of a few Hz since this is typically the scale of the peak-to-peak distances. On the other hand, smoothing of spectra can also result in less accurate coupling measurements when narrow mechanical resonances are present, so the window must balance the smoothing of artifacts against this disadvantage. For accelerometer spectra, analyzing injections with various smoothing parameters show that a logarithmically-scaled window which is **XX** Hz wide at 100 Hz and **XX** Hz wide at 1000 Hz best satisfy these constraints. Since microphones are much more sensitive to nodal artifacts while being less sensitive to narrow mechanical resonances (they would have to be strong enough to produce audible signals), their spectra can be smoothed much more aggressively: a logarithmically-scaled window is used which is 15 Hz wide at 100 Hz and 150 Hz wide at 1000 Hz.

3.6 The **pemcoupling** package

This section covers the technical details of the `pemcoupling` python package (Nguyen, 2020), which includes command-line tools for processing large

numbers of injections and producing single-injection coupling functions, coupling functions, and multi-channel summary coupling functions.

The package uses the `gwpv` library for fetching raw time series data and producing ASDs of the GW strain channel and auxiliary channels from user-provided background and injection times.

3.6.1 Processing steps

This subsection is quite messy. The confusion comes from talking about different sensors at different steps. I think I could try splitting it by sensor type, like list all steps for accelerometers and microphones, then all steps for magnetometers.

A number of pre-processing steps are performed to condition the data for analysis. First, the auxiliary channel time series are examined for evidence of saturation. Injections can cause saturation in two ways: actuator saturation due to over-driven amplifiers, or sensor saturation due to hitting the maximum amplitude the sensor can record.

During magnetic injections, the injection amplitude typically has to exceed background by a few orders of magnitude to produce noise in the GW channel, so comb magnetic injections are the most likely to saturate at the actuator. They are also capable of saturating the sensors if performed too close (such as in the electronics racks). In either case, intermodulation distortion occurs, generating peaks at sums and differences of the comb line frequencies while also reducing signal power in the injection lines. If the saturation is occurring at the actuator, the measured coupling is still accurate, but upper limits are higher due to the reduced signal power. On the other hand, sensor saturation results in inaccurate

coupling measurements, because the line amplitudes no longer reflect the true magnetic signal at the sensor location.

Accelerometers can saturate in the presence of broadband vibrational injections. This results in artificial broadband noise in the amplitude spectra. The consequence is measuring a lower coupling function at frequencies where the saturation artifact dominates.

For these reasons, coupling functions are only measured for sensors that do not exceed 32,000 counts, as this is the sign of sensor saturation.

Once the time series are converted to ASDs, triaxial magnetometer channels are combined in quadrature. Each magnetometer measures the x , y , and z components of its local magnetic field independently. An artificial channel is produced to represent the absolute magnitude of the field.

All PEM channels have accompanying calibration measurements (R. Schofield et al., 2021). Calibration for microphones and magnetometers is simply a constant conversion of Pa/count or T/count, respectively. Accelerometer ASDs are converted to acceleration (m/s^2), then to displacement (m) by dividing each bin by $(2\pi f)^2$ where f is the bin frequency.

For acoustic injections, spectral smoothing is applied at this point, in order to suppress nodal artifacts as explained in Section 3.5.3. Finally, the single-injection coupling functions $\mathcal{C}_{ij}(f)$ and upper limits $\mathcal{C}_{ul}(f)$ are measured as in eq. (3.8).

Paragraph or two on the CF calculation step.

3.6.2 Data products

For each $\mathcal{C}_{ij}(f)$, the data are saved in the following forms:

| column | description |
|---------------|--|
| frequency | bin center frequency [Hz] |
| factor | coupling factor in [m/calibrated sensor unit] |
| factor_counts | coupling factor in [m/ADC count] |
| flag | “Measured”, “Upper Limit”, “Thresholds not met”, or “No data” |
| sensINJ | sensor amplitude at injection time [calibrated sensor unit/ $\text{Hz}^{1/2}$] |
| sensBG | sensor amplitude at background time [calibrated sensor unit/ $\text{Hz}^{1/2}$] |
| darmINJ | GW channel amplitude at injection time [m/ $\text{Hz}^{1/2}$] |
| darmBG | GW channel amplitude at background time [m/ $\text{Hz}^{1/2}$] |

Table 2. Column descriptions for the single-injection coupling function output of the `pemcoupling` package.

1. comma-separated text file consisting of coupling measurements, flags, and raw spectra
2. plot of the raw coupling function (units of meters per analog-to-digital counts)
3. plot of the coupling function in physical units (meters per calibrated sensor unit, e.g. Tesla for magnetometers)
4. figure containing two subplots: one showing the background and injection spectra of the auxiliary sensor, and one showing the background and injection spectra of the GW strain data and the estimated environmental noise projection.

Post-processing, to aggregate single-injection coupling functions into coupling functions, and produce site-wide coupling plots, is done via additional commands, `pemcoupling-composite` and `pemcoupling-summary`.

Expand on these final steps a bit.

CHAPTER IV

STUDIES OF ENVIRONMENTAL NOISE DURING O3

Introduce how observing/engineering runs work, O3 schedule, etc.

O3 was preceded by an engineering run, a month-long period during which the interferometer is operational at low-noise levels but not observing GW events, to provide time for detector commissioning and noise studies.

4.1 Evaluation of coupling functions

Pretty much just copied from PEM paper. Goal is to include most or all of the referenced calculations and figures here.

Before discussing results of injection studies and coupling function measurements in O3, it is necessary that we take a moment to assess how well they estimate excess noise in the interferometer when the noise source is known to be the result of environmental coupling. Of course, one could do this by applying coupling functions to a set of “test” injections, but this would require performing more injections, which as mentioned before is typically not feasible. Furthermore, since the intention of the coupling functions is to estimate the impact of noise signals originating from a broad variety of sources, it makes sense to use non-laboratory signals when evaluating coupling estimates. Therefore, we use noise events not produced by injection equipment to evaluate how accurately the coupling functions recover the actual excess noise observed in the GW strain data.

Thunderstorms are known to produce short-duration transients in the strain data at tens of Hz. At LLO, coupling functions for several accelerometers at the Y end station, where vibrational coupling was the highest, are capable of

estimating the amplitude of multiple noise transients to within a factor two during a particularly loud thunderstorm (Nguyen & Schofield, 2019).

Helicopter flyovers can produce narrow-band features up to tens of seconds long. Coupling functions of various sensors at both observatories can predict the amplitudes of lines produced by multiple helicopter flyovers during O3 to within a factor of two in most cases (Nguyen & Schofield, 2020).

Long-duration noise due to vibrations from rain and the building heating, ventilation, and air conditioning (HVAC) is also well characterized by coupling functions at LHO (Banagiri, Covas, & Schofield, 2019; R. M. S. Schofield, Nguyen, Banagiri, Merfeld, & Effler, 2019).

4.2 Vibrational noise studies during O3

Figure 11 shows the ambient contribution of vibrational noise during O3, produced by combining the highest coupling factors among accelerometers and microphones measured from an injection campaign at the beginning of O3. At the end of O3, the vibration noise background at both observatories was dominated by input beam jitter above 100 Hz (discussed in Section 4.2.1). At LHO, the dominant coupling region below 100 Hz was the output arm. At LLO, the dominant coupling regions were the Y-end in the 40-60 Hz band and the output arm in the 60-100 Hz band.

4.2.1 Scattered light at the HAM5/6 septum

A major source of detector noise and reduced sensitivity to GWs is the scattering of light from the beam spot on a test mass or other optic to surfaces that are moving relative to the optic, like vacuum chamber walls. A very small

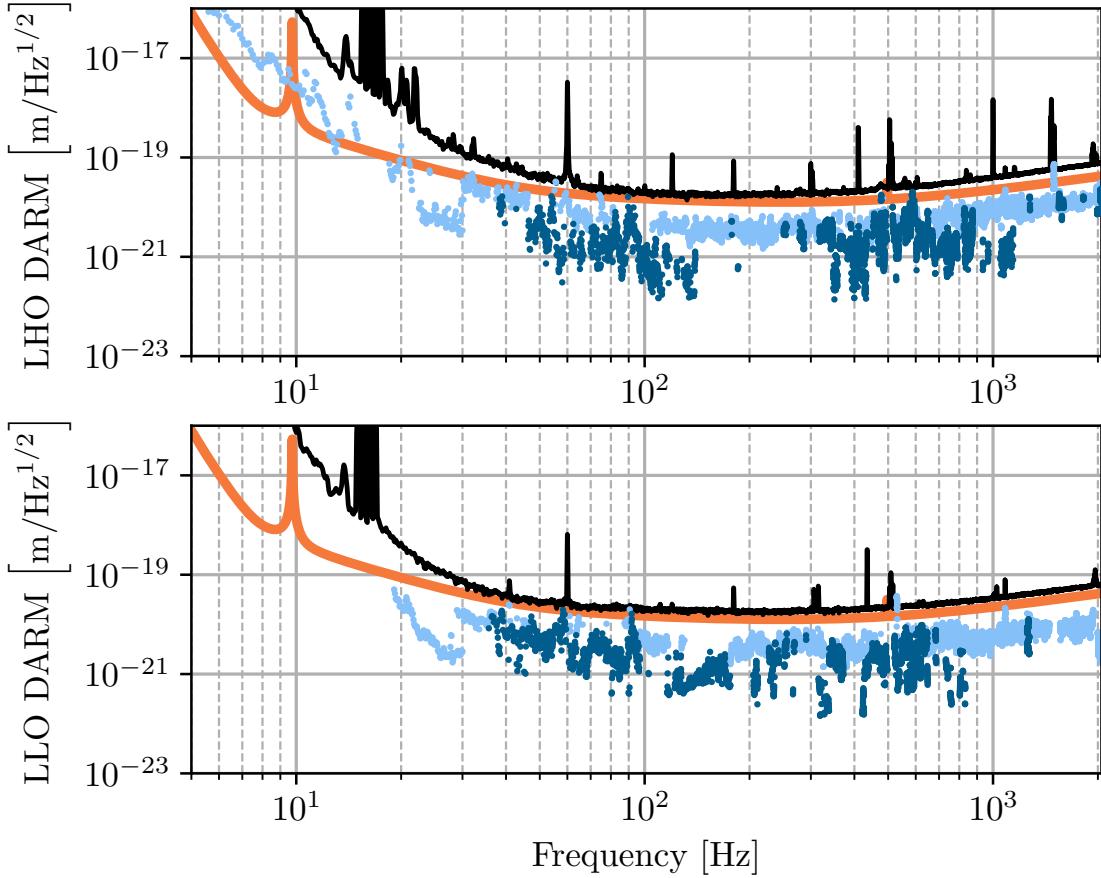


Figure 11. Ambient estimate of vibrational noise levels at LHO (top) and LLO (bottom).

fraction of the light reaching the moving surface is reflected to the originating or another beam spot, where it scatters back into the main interferometer beam. As the distance to the moving surface changes, the phase of the returning light changes relative to the main beam, producing fluctuations in the amplitude of the beam, that, at 1 part in 10^{20} can be on the scale of those produced by gravitational waves. In addition to this sensitivity to recombined scattered light, the scattering noise is problematic because of non-linear coupling when the path length modulation becomes comparable to the wavelength of the light, producing noise at harmonics of modulation frequencies (Soni et al., 2020).

At LHO, investigations throughout O3 showed that scattering noise produces noise near the detector noise background in the frequency range of 38–100 Hz. The sensors with the highest ambient projections in this band were accelerometers located on the HAM 5 and HAM 6 vacuum chambers, which contain the GW channel readout, as well as the optics of the output mode cleaner, squeezed light system, and signal recycling cavity. The coupling was excited most strongly by injections around the output arm, but acoustic noise produced as far as the Y-arm manifold \approx 50 m away produced excess noise in DARM. Spectra from the nearest injections also show non-linear coupling behavior ([Expand on this](#)). Shaker injections were also performed throughout the corner station, confirming that the dominant coupling site was in the HAM 5 and 6 area. This evidence suggested that the coupling was caused by scattered light within HAM 5 or HAM 6. A more thorough investigation was required to localize the exact scattering surface.

Figure 12 shows time series and spectrograms of a single impulse injection performed in the horizontal access module (HAM) 5/6 area. The plots show a single impulse injection signal in the GW channel and in various output optics accelerometers. Multiple sensors observe an impulse time-of-arrival matching that of the GW channel, but repeating the injection from various other locations rules out sensors that do not match it consistently across multiple injections. In this case the septum (separating the HAM5 and HAM6 chambers) accelerometer signal matches the DARM signal most consistently (other injections not shown for brevity).

Spectrograms of the same impulse injection for DARM and the three sensors with the closest matching time-of-arrival to that of DARM reveals similarities between the frequency structure of the septum accelerometer signal and that of

DARM. This provides further support that the septum is the dominant coupling site in the output arm.

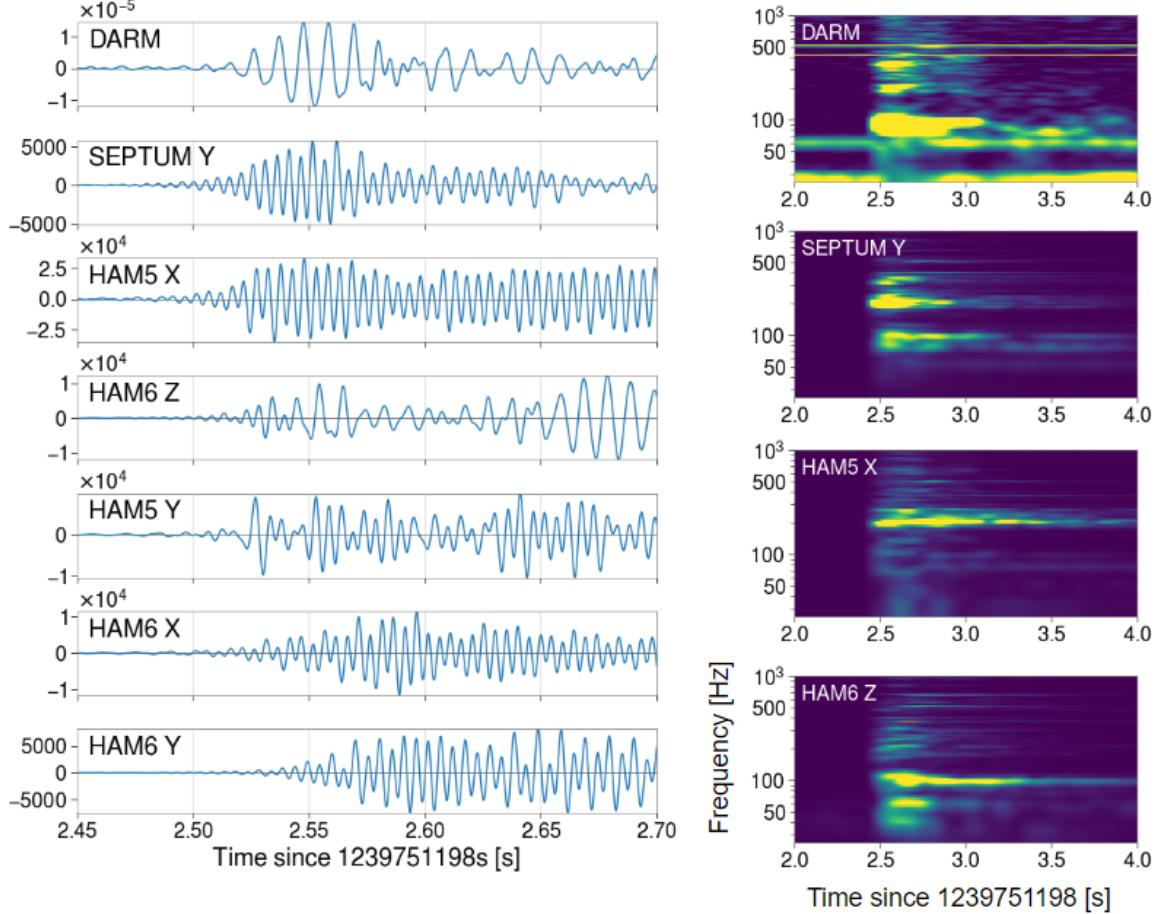


Figure 12. Time series (left) and spectrograms (right) of a vibrational impulse injection produced at the output arm of the LHO detector.

4.2.2 Search for the source of a 48-Hz peak

Figure 13 shows spectrograms of a beating-shakers injection used to localize the coupling site responsible for a 48-Hz noise peak in the DARM spectrum. The shakers were injecting at 48 and 48.01 Hz. The Y-axes of the spectrograms are centered along at 48 Hz and show the combined signal in each sensor modulating at the beat frequency (0.01 Hz). This set of spectrograms suggests that the

accelerometers on the input test mass (ITM) chambers and the Y-axis HAM2 accelerometer are likely not close to the true coupling location, since the beat envelopes are the furthest offset from the beat envelope in the DARM response.

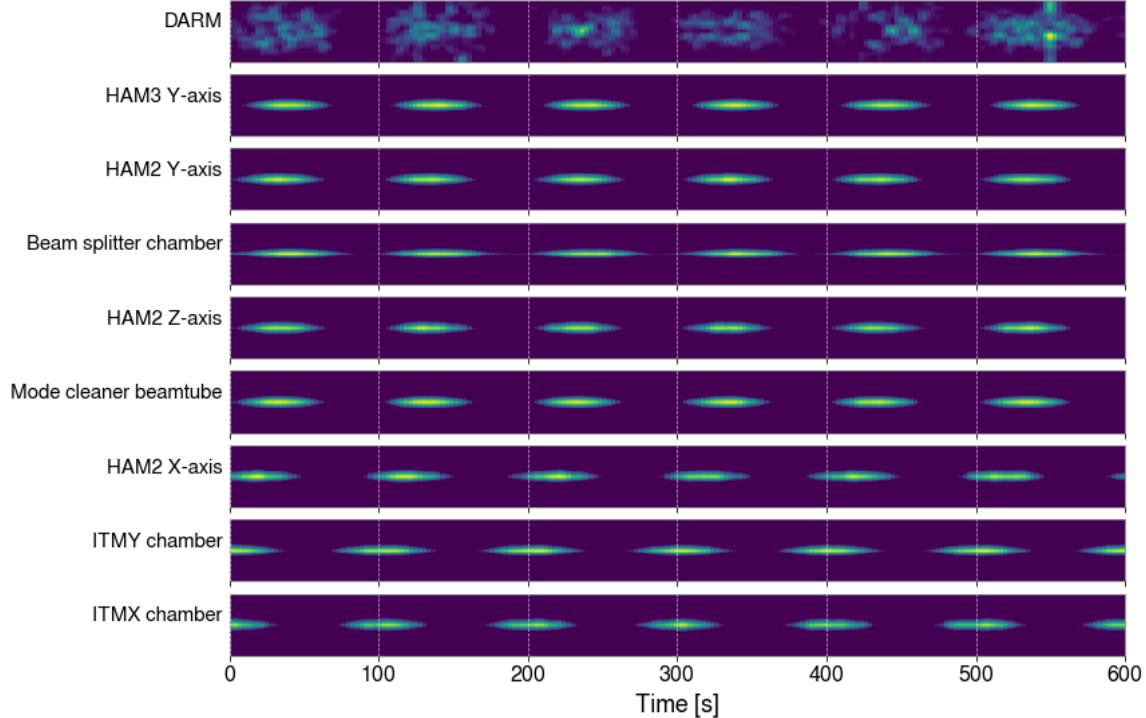


Figure 13. Spectrograms of DARM and various accelerometers near the input arm and beam splitter showing a beating-shakers injection at 48 Hz.

Multiple other injections were made (not shown here) with varying shaker locations in order to rule out other sensors until the most likely candidate remaining was the HAM3 Y-axis accelerometer. Black glass was used to block scattered light at this location and the peak was eliminated for the second half of the O3 observation run. Figure 14 shows the noise reduction in the GW channel after the light scattering was mitigated.

4.2.3 Input beam jitter

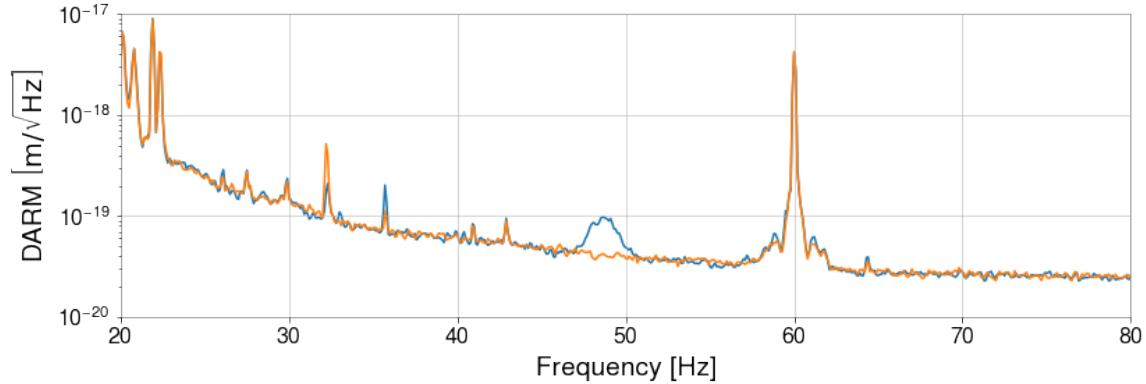


Figure 14. LHO DARM spectrum before and after mitigation of the 48-Hz peak.

Figure 15. Improvement in jitter coupling at LHO (left) and LLO (right) between the start and end of O3.

Short discussion, summarize measurements and improvement. Will make sure to discuss impact of test mass defect on jitter noise.

4.3 Magnetic noise studies during O3

Magnetic injections early in aLIGO suggested that coupling to permanent magnets in the suspension system could prevent LIGO from reaching design sensitivity in the 10-20 Hz regions (R. M. S. Schofield, 2013). While the test mass actuator is electrostatic and not magnetic (as in iLIGO), a number of permanent magnets were used in the suspensions, including for actuation in the first three of the four levels of the isolation chain and for eddy current damping. The greatest number of permanent magnets were in the eddy current damping arrays and these were removed. Nevertheless, ambient fields are still predicted to produce noise at greater than one-tenth of the design sensitivity in the 10-20 Hz band (Figure 16), and may need to be further addressed as we reach design sensitivity in the 10 Hz region.

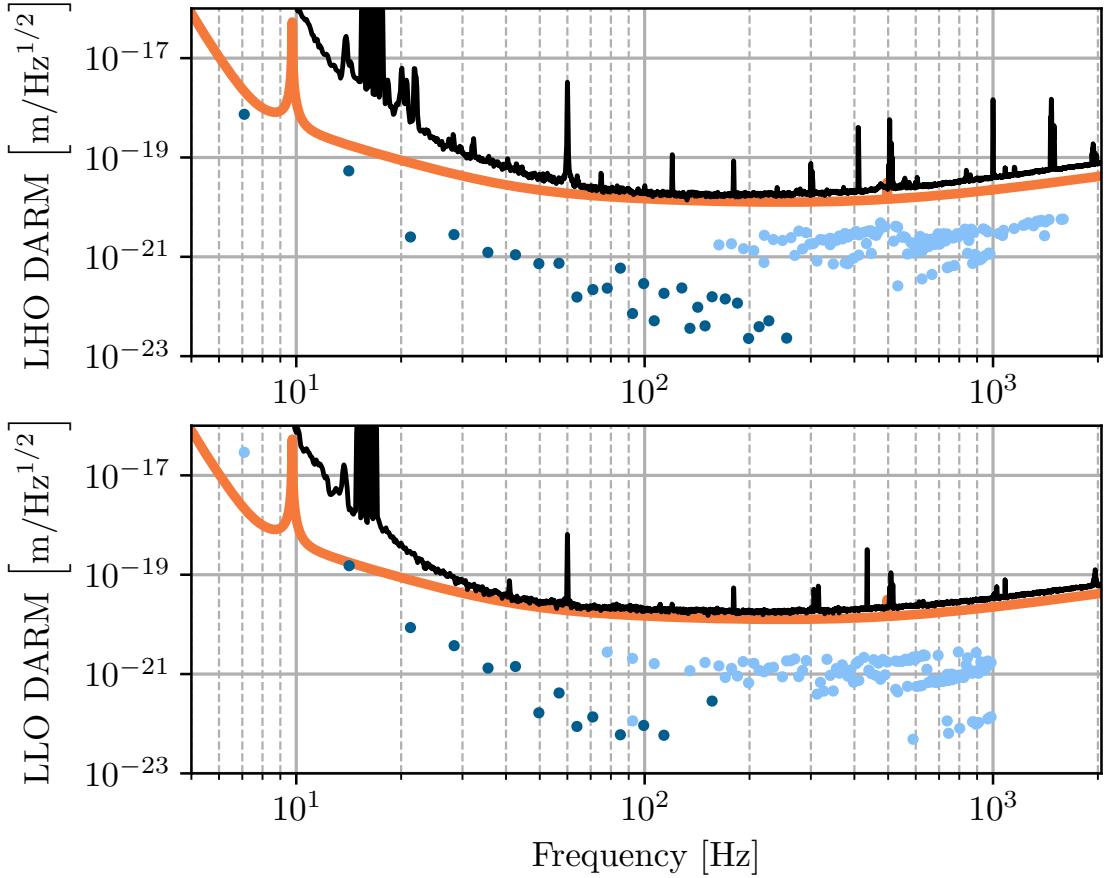


Figure 16. Ambient estimate of magnetic noise levels at LHO (top) and LLO (bottom).

At higher frequencies, generally above about 30 Hz, the dominant magnetic coupling appears to be through induction of currents in cables and at connectors, mainly to actuator cabling and other cabling in the control system. Mitigation of coupling to cables and connectors has required a continuing program of monitoring coupling since cables are often disconnected and reconnected during runs as electronics are replaced for problems or upgrades. This program consists of making weekly, broadband magnetic field injections using the large wall-mounted coils described in Section 3.3.2. Since the injections are scheduled for every Tuesday morning before the routine detector maintenance period, the interferometer may

or may not be in a locked state at the time, so an injections were not always performed.

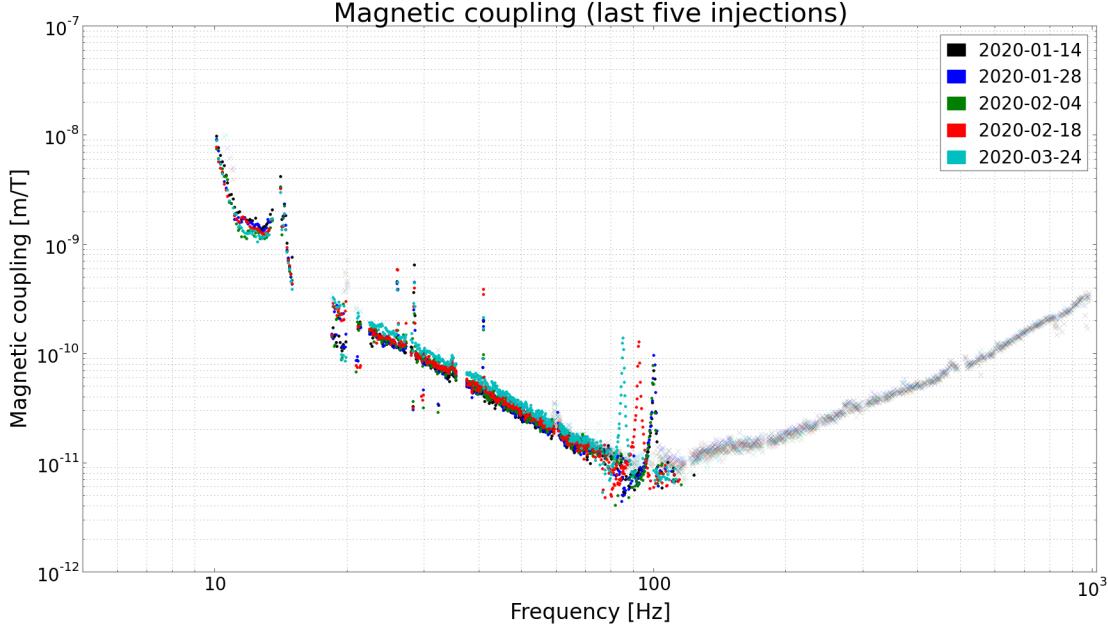


Figure 17. LHO magnetic coupling measured by the last five wall-mounted coil injections performed in O3.

Figures 17, 18, and 19 are automatically generated every week by the code that analyzes the injections. Figure 17 shows magnetic coupling functions measured over the last five weeks for which a broadband injection was performed. Changes can be seen in both the broadband and narrowband structure of the coupling function. Just within these five weeks, the level of broadband coupling varied by as much as a factor of about 1.5. Since the injection is produced from a same location and at the same amplitude every time, uncertainties due to the injection source as discussed in Section 3.5 do not account for these variations. Broadband changes

over specific frequency ranges tracked over the full course of O3 (Figure 18) show significant fluctuations throughout the run.

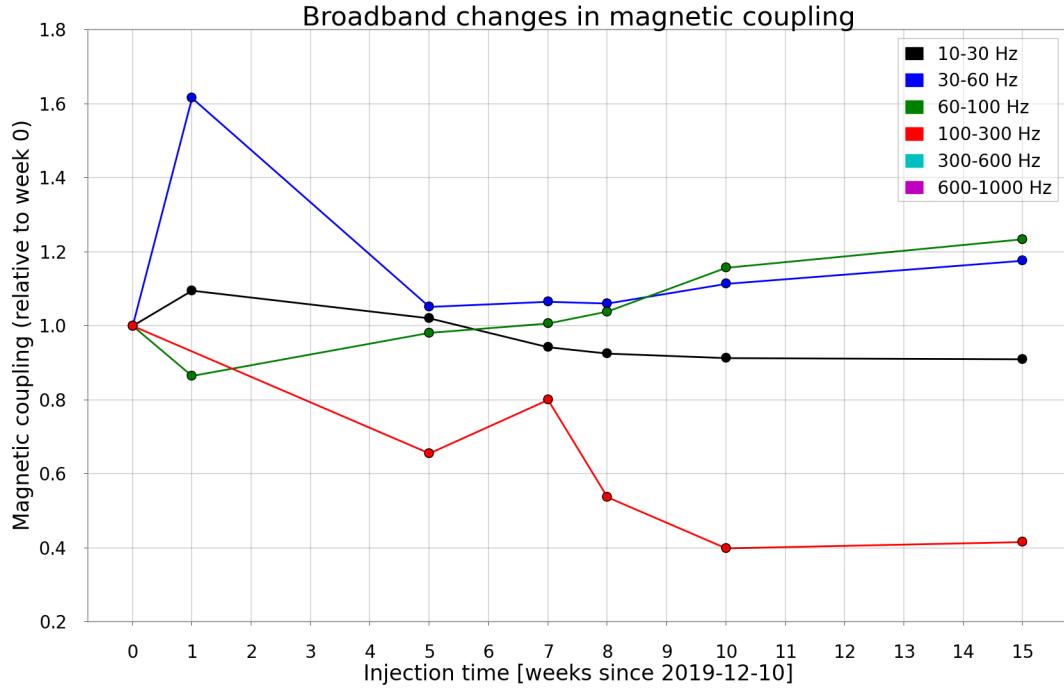


Figure 18. Time-lines of magnetic coupling changes relative to the start of the observing run.

Furthermore, a large peak in the coupling is seen migrating between 90 and 110 Hz. This is precisely the type of feature often missed by comb injections but will be routinely discovered by broadband injections in future observing runs. Figure 19 shows frequency and amplitude fluctuations in coupling peaks such as this. Although the coupling of the $\tilde{100}$ Hz peak is still well below the level that would produce excess noise in the GW channel, its presence gives reason to be concerned that similar peaks could arise in the future that do couple significantly. These weekly injections would help in identifying when the coupling changed, so instrumentalists can deduce what changes to the electronics may have affected it.

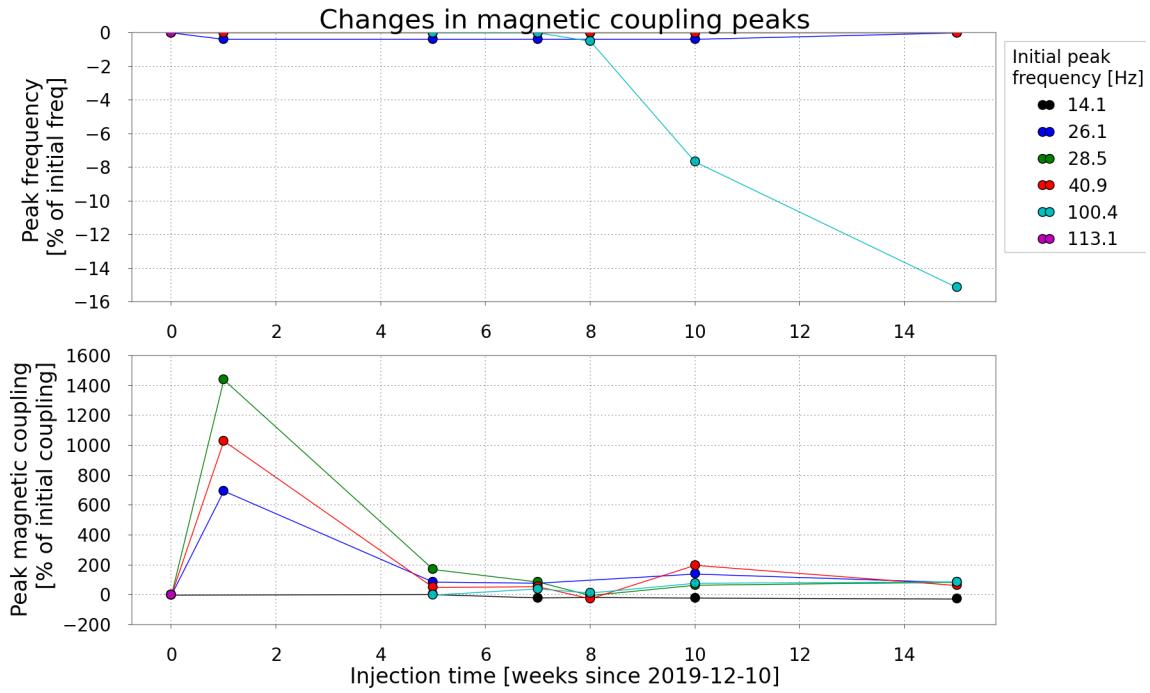


Figure 19. Weekly trends in frequency (top) and amplitude (bottom) of peaks in the magnetic coupling functions.

4.4 Validation of gravitational wave event candidates

In addition to investigating sources of environmental influences, knowledge acquired from environmental studies contributes to the vetting of GW event candidates. Analysis pipelines search the strain data for astrophysical signals. They are categorized into modeled searches for binary mergers that match the data to template waveforms (e.g. GstLAL (Cannon et al., 2012) and PyCBC (Usman et al., 2016)) and unmodeled searches that identify excess energy coherent between multiple detectors (e.g. cWB (Klimenko, Yakushin, Mercer, & Mitselmakher, 2008), oLIB (Lynch, Vitale, Essick, Katsavounidis, & Robinet, 2017), and BW (Cornish & Littenberg, 2015)).

Contamination of the GW data can occur through any of the means discussed in previous sections. Environmental noise has the potential to be correlated between detectors by stemming from a common source, such as through electromagnetic signals from distant sources or glitches in GPS-correlated electronics. The analysis pipelines estimate the false-alarm probabilities for GW events based on the background rate of randomly coincident events in the detector network. They generate background events by time-shifting the data stream of one detector relative to another by time steps much longer than the light travel time between detectors and longer than the duration of GW signals. This method does not account for the possibility of transients being correlated between the detectors due to a common environmental source.

Environmental noise is also particularly relevant to unmodeled searches. Unlike template-based methods, these searches make minimal assumptions about the signal waveform and rely more heavily on signal correlation between sites.

The first observation of a GW occurred on 14 Sept 2015 (Abbott et al., 2016b). The event, a short-duration binary black hole merger designated GW150914, required a number of follow-up investigations to find potential noise sources around the time of the event (Abbott et al., 2016a). This included an examination of the status of all PEM sensors and any significant signals they observed for possible contamination of the GW signal (R. M. S. Schofield, Roma, et al., 2018). A few of the PEM sensors were not working, but because of redundancy, coverage was sufficient.

Comparisons between Q-transform spectrograms (Chatterji, Blackburn, Martin, & Katsavounidis, 2004) of all coincident events in environmental sensors to the time-frequency path of the event revealed that no environmental signals

had paths similar to the event candidate. Q-transforms produce a quality-factor-optimized logarithmic tiling of the time-frequency space, making them useful for visualizing transients. The signal-to-noise ratios (SNRs) of the matching signals were also compared to that of the event, showing that even if there were overlapping time-frequency paths, none of the environmental signals were large enough to influence the strain data at the SNR level of the event, based on multiplying the environmental signals by their respective sensor coupling functions.

The validation process for novel events such as GW150914 also includes redundant checks for global sources of environmental noise. We use a dedicated cosmic ray detector located below an input test mass at LHO to examine any association of cosmic ray showers to excess noise in DARM. We also check external observatories for coronal mass ejections, solar radio signals, geomagnetic signals, and radio-frequency (RF) signals in the detection band as well as higher frequencies.

There was specific concern over a co-incident extremely-high current (504 kA) lightning strike over Burkina Faso, prompting additional studies of the effects of lightning on the interferometer (R. M. S. Schofield, 2018). Investigations of similar strikes found no effect on the strain data and investigations of closer strikes confirmed that the magnetometers were much more sensitive to lightning strikes than the interferometer was. In conclusion there was no reason to veto the first detection based on environmental disturbances.

Subsequent detections throughout O1 and O2 employed a similar procedure; however the development of the method described in Section 3.4 for producing coupling functions for all sensors expedited the process. This was especially important for examining environmental noise during GW170817, the first long-

duration event detected by LIGO (Abbott et al., 2017; R. M. S. Schofield, Nguyen, et al., 2018). The longer duration of this event (75 s) unsurprisingly overlapped with many environmental signals. Based on the coupling functions for those sensors, several of these environmental events were loud enough (estimated DARM signals of up to SNR 4) to have contributed to the interferometer readout, but not enough to account for the GW signal. Furthermore, none of them had a time-frequency morphology that correlated with any features in the candidate signal.

4.4.1 Automated validation of O3 events

Since the start of O3, most of the procedure described above has been automated in order to handle the increase in detection rate. The automated vetting is performed by the `pemcheck` routine, which is a part of the Data Quality Report (DQR). When an event is detected by the astrophysical search pipelines, a DQR is initiated and assembles a plethora of tasks for assessing the data quality at each observatory during the time of the event. Among these tasks, an omega scan pipeline (Chatterji et al., 2004; Davis et al., 2021) is used to search for transient noise in all PEM sensors in the time window spanning the event candidate. It does so by producing a Q-transform for each sensor and reporting those in which there is a transient signal with a false-alarm rate below 10^{-3} Hz. The omega scan also reports the frequency and amplitude of the most significant tile for each sensor. The `pemcheck` in O3 used the output of the omega scan to estimate each sensor’s potential affect on the data quality of the detector. The coupling function of each sensor was interpolated at the peak frequency and multiplied by the peak amplitude, producing an estimated DARM amplitude.

Sensors whose estimated contribution exceed one tenth of the DARM background level were flagged for human input, requiring a comparison of the environmental signal morphology to that of the event candidate. If there was sufficient signal overlap, reviewers may advise that analysts perform some noise removal in the data, such as by gating or filtering out the appropriate time or frequency range, before performing further follow up analyses. The event could be retracted, if gating or filtering out the environmental contribution would reduce the signal-to-noise ratio of the candidate to a level no longer consistent with a GW detection.

During O3, no candidates were retracted on the basis of the environmental coupling check alone. Some human input was still required for all of the **XX** events reported in (Abbott et al., 2021a), although little to no signal overlap of environmental transients was seen.

4.4.2 Event validation in O4

With the GW detection rate expected to increase in O4, a more sophisticated and streamlined vetting routine is necessary. The DQR in O4 will report a p-value for each vetting task, including the `pemcheck` task. In the case of environmental noise vetting, this p-value represents the null-hypothesis probability, i.e. the probability that environmental disturbances are not contaminating the GW event signal. This probability is defined based on the uncertainty of the coupling functions.

Suppose we have measured the coupling for a sensor at a single frequency bin, $C(f_k)$. As we know systematic uncertainties result in coupling measurements to be log-normally distributed (see Section 3.5.3), we can describe the true coupling

with a log-normal distribution with mean $\ln C(f_k)$ and standard deviation $\ln 2$ (corresponding to a factor two uncertainty). If only an upper limit is available, then the coupling must be equal to or less than the upper limit, so it can be described by a uniform probability distribution between 0 and the upper limit. Thus multiplying these distributions by the sensor ambient we get a probability distribution for the projected noise level $h_p(f_k)$ in the GW channel in terms of $\mu = C(f_k) \cdot X(f_k)$:

$$P(h_p) = \frac{1}{h_p \sqrt{2\pi} (\ln 2)^2} \exp \left[-\frac{\ln h_p - (\ln \mu)^2}{2(\ln 2)^2} \right] \quad (4.1)$$

for measurements and

$$P(h_p) = \begin{cases} \frac{1}{CX} & \text{if } x < CX \\ 0 & \text{otherwise} \end{cases} \quad (4.2)$$

for upper limits.

We can write the probability that the projected noise actually exceeds some GW channel background level $h_{\text{bkg}}(f_k)$ based on the corresponding cumulative distribution functions:

$$P(h_p > h_{\text{bkg}}) = \begin{cases} 1 - \frac{1}{2} \left[\operatorname{erf} \left(\frac{\ln h_p - \ln \mu}{\sqrt{2} \ln 2} \right) \right] & \text{(measurement)} \\ 1 - \min(1, \frac{h_p}{\mu}), & \text{(upper limit)} \end{cases} \quad (4.3)$$

This is computed for every time-frequency pixel of a spectrogram, producing a probability spectrogram image that shows how likely there is to be excess noise within each pixel. We can then search for the highest probability within pixels that overlap a GW transient in order to report the probability that environmental

noise in the vicinity of the PEM sensor is contaminating the signal of interest. This algorithm is run on every PEM sensor with an available coupling function and the highest probability among all sensors is reported.

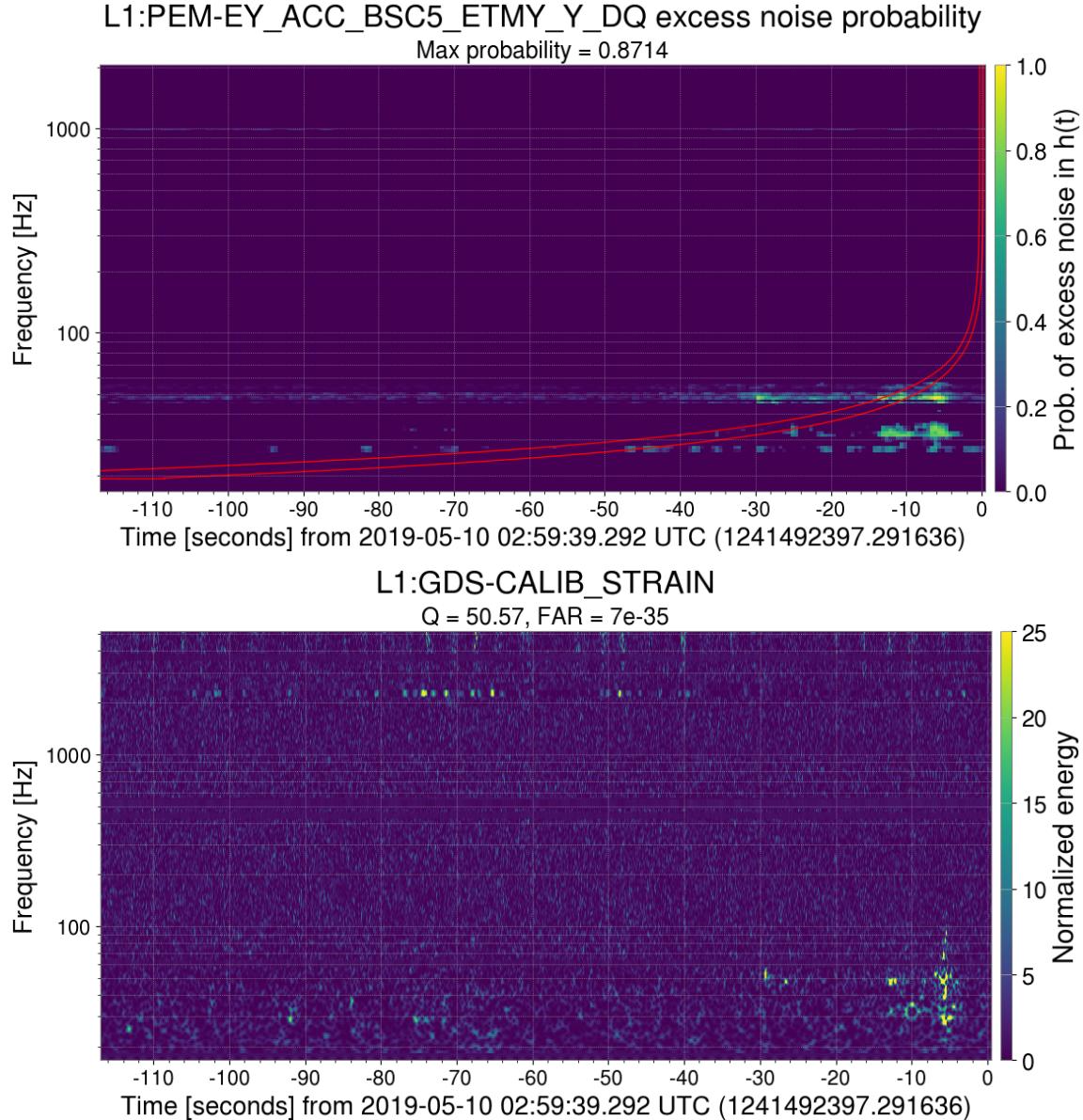


Figure 20. Probability spectrogram of an ETMY accelerometer at LLO (top) and a constant-Q transform of the GW strain channel (bottom). The red lines show the time-frequency path of GW event candidate S190510g.

Figure 20 provides an example of a probability spectrogram output by pemcheck for an O3 GW event candidate, S190510g. The candidate signal coincides with scattering noise produced at the LLO Y-end station by a thunderstorm. The coupling projection predicts a peak probability of 0.87, meaning there is an 87% probability that noise is present in the interferometer overlapping the time-frequency window of the GW event.

There are still limitations to how well this method can predict the presence of excess noise in the GW channel, discussed below.

4.4.2.1 *Coupling function tuning.* Coupling function projections often overestimate the GW background level for a number of reasons. The most common causes are all linked to the fact that coupling functions can only be measured during extended periods of detector maintenance, usually at the start of an observing run, leaving them vulnerable to becoming outdated as noise sources are introduced, changed, or mitigated. If a noise source is introduced near a sensor, such as through the installation of new hardware, the sensor ambient can increase dramatically whether or not the noise source actually couples to the GW channel. If the new noise does not couple, then the projection overestimates the detector amplitude, leading to a spurious claim that a GW candidate signal may be affected by environmental noise. Conversely, if an existing noise source is removed or its coupling mechanism mitigated, then the GW background drops, potentially below the level of estimated ambient noise projections.

For these reasons, coupling functions must be tuned to account for overestimated noise in the time around an event candidate. This is done by simply treating background noise in a long stretch of time before the candidate as if it were a new injection. Rather than re-measuring the coupling function, we can

check for frequencies where projections estimate ambient coupling to exceed the GW background, dividing the coupling function by the ratio by which it is overestimating. The result is a coupling function that at most projects noise at background levels.

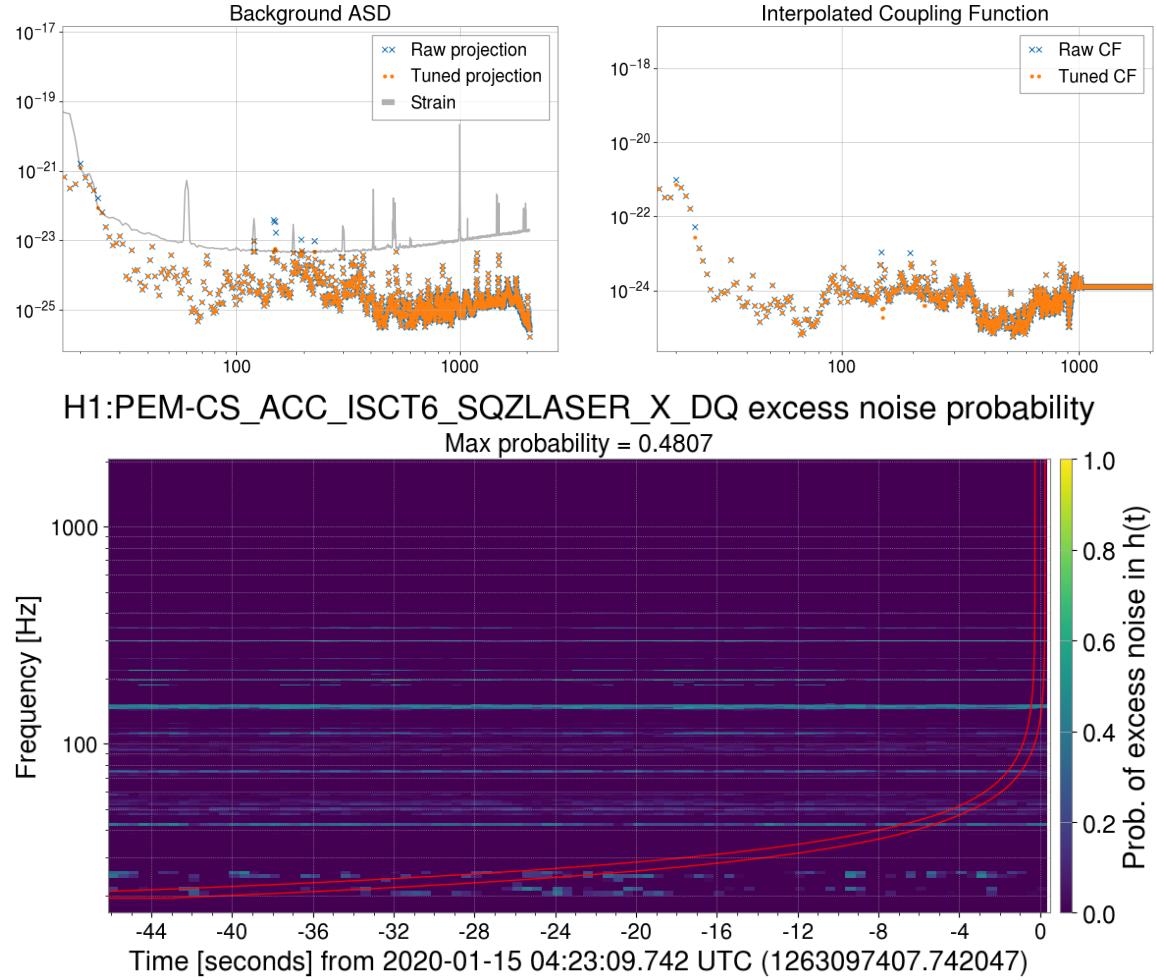


Figure 21. Coupling function tuning (top) and the resulting probability spectrogram (bottom).

Figure 21 shows an egregious case of projected noise for an accelerometer on the squeezed light optics table overestimating the strain background amplitude. The tuning reduces this to match the strain background, and the resulting

probability spectrogram does not estimate a significant probability that excess noise is present.

4.4.2.2 *Grouping nearby sensors.* Even after tuning coupling functions, projections can still be overestimated in the presence of short duration (≤ 1 s) transients that are not likely to couple based on physical reasons. Rack magnetometers (placed on the metal racks that hold the various control systems, data acquisitions systems, and power supplies in the electronics rooms) observe the highest rates of localized short-duration transients, routinely picking up magnetic fields from changes in nearby currents. When these glitches coincide with a GW event candidate they predict a probability of excess noise usually above 90%. Although the coincident rate is low for typical (BBH) events, a long-duration GW signal such as a BNS will typically overlap with at least a few electronics glitches.

Since coupling functions are only intended to be used for noise sources distant from the sensors, projections from sensors placed close to each other ($\lesssim 1$ m) should be highly correlated if the input noise is environmental and not local to the individual sensors. To incorporate this knowledge, `pemcheck` groups together magnetometers in each electronics room; their probability spectrograms are stacked and a pixel-wise minimum is computed to generate a combined group spectrogram. This suppresses signals that project above the strain background in one magnetometer but not the others, so the projected excess noise probabilities are much lower than any of the peak probabilities reported by the individual sensors.

CHAPTER V

GRAVITATIONAL WAVES ASSOCIATED WITH GAMMA-RAY BURSTS

5.1 Gamma-ray bursts

GRBs are short, energetic bursts of gamma rays in the MeV range, first discovered in 1967 (Klebesadel, Strong, & Olson, 1973). Observations throughout the following decades revealed that GRBs could be classified based on duration and spectral hardness (Kouveliotou et al., 1993). This classification has become the most used to describe GRBs: long-soft bursts last $\gtrsim 2$ s and have soft emission spectra, i.e. lacking in higher energy photons, while short-hard bursts last $\lesssim 2$ s and had harder emission spectra. It is believed that the ultra-relativistic jets required to produce GRBs come from either black holes (Woosley, 1993) or magnetars (Dai & Lu, 1998).

A multitude of models have been proposed throughout the decades to explain the origins of GRBs. [Go through a few, reference e.g. review papers.](#)

Photometry and spectroscopy data provide evidence that long GRBs originate from core-collapse supernovae (CCSNe), whereas short GRBs are believed to be associated with compact binary mergers, such as the BNS merger GW170817 (Abbott et al., 2017).

In the case of long GRBs, some models predict the emission of GWs as a result of asymmetries in the core-collapse phase. Such GWs would be short, lasting less than a second. Extreme emission models predict a wide variety of signals, often longer in duration. Matter surrounding the remnant of a CCSN forms an accretion disk, in which turbulent behavior can arise. For instance, instabilities in the accretion disk can lead to the formation of a clump of matter which can then migrate inwards, shedding angular momentum in the form of gravitational waves.

5.2 GW searches

In 2017, the first binary neutron star merger GW170817 was detected by advanced LIGO and Virgo, immediately accompanied by the detection of a relatively low-luminosity gamma-ray burst GRB170817A by Fermi Gamma-ray Burst Monitor (GBM) two seconds later (Abbott et al., 2017). This prompted a global effort to find an optical, UV, and infrared counterpart that would make up the signature of a kilonova. The localization provided by the joint detection was sufficient to locate a counterpart near NGC 4993.

GW parameter estimation suffers from a degeneracy between distance and orbital plane inclination; increasing the distance of the merger and orienting its orbital plane would both result in a lower GW amplitude. This degeneracy can be broken if either one could be measured externally. If the joint detection localization were good enough to determine a host galaxy, this would greatly help resolve the distance, however this will more likely require observation of an optical or UV counterpart due to the poor localization provided by current GRB and GW detectors. In the case of GW170817, the host galaxy whose redshift was known was used to determine the distance, which then allowed for a more precise measurement of the inclination angle. A separate analysis using the distance measured from the GW detection combined with the known redshift made the first joint GW-EM measurement of the Hubble constant, albeit with very large uncertainty due to the small sample size (Abbott et al., 2017).

[Discuss background and disagreement among Hubble constant measurements.](#)

One of the many unanswered questions surrounding GRBs pertains to their jet profile, the luminosity as a function of viewing angle (the angle between the

observer and the symmetric axis of the jet; the profile is assumed to be axially symmetric and independent of distance). When information about the jet profile is required, e.g. for making rate estimates for GRB detections, the profile is typically modeled as a top-hat (uniform within some opening angle and dropping sharply beyond it) for simplicity, but the true profile may be different.

Determining the viewing angle θ_{obs} of a GRB is essential for distinguishing between different jet profile models, but it relies on the ability to observe an afterglow emission. These emissions, ranging from radio to X-rays, follow the prompt emission of γ -rays and are generated by the interaction between the relativistic outflow and the surrounding medium. The observation of these afterglows was crucial in improving localization of the GRB sources and provided valuable information about the energy scale of the jet, as well as measurements of θ_{obs} . The latter came from observing the signature jet break in the afterglow light curve, resulting from the lateral spreading of jet material as it expands; the opening angle can be determined from the timing of the jet break. However, few jet breaks have been observed for short GRBs, so existing observations do not place tight constraints on opening angle (Biscoveanu, Thrane, & Vitale, 2020).

Joint GW-GRB observations can provide much more information on jet properties (Farah et al., 2020; Mogushi, Cavaglià, & Siellez, 2019). GRB170817 was orders of magnitude less energetic than most short GRBs, so it likely would have been ignored in the absence of a GW coincidence. Its low luminosity immediately ruled out an on-axis top-hat jet. An off-axis top-hat jet was considered unlikely as well because the narrow opening angle predicted for top-hat jets based on theory and past GRB measurements only allowed for $\theta_{\text{obs}} \lesssim 10$ deg. More evidence against an off-axis top-hat model arose when the bright afterglow expected to

emerge after \sim 1 day was not observed. These observations instead favored a wide-angled, structured jet model for GRB170817. A structured jet model may refer to any luminosity function that decreases gradually with θ_{obs} rather than abruptly, e.g. a Gaussian or power-law with uniform center. One mechanism that would explain such a model is a cocoon emission, in which the relativistic jet interacts dissipatively with the surrounding merger ejecta, depositing its energy into a cocoon fireball that results in a structured jet (Abbott et al., 2017).

Discuss other lessons learned from GW170817 kilonova.

Motivate using a targeted search: justify w/ astrophysics and better sensitivity.

The LIGO-Virgo collaboration searches for gravitational waves associated with GRBs detected by Fermi GBM and Swift Burst Alert Telescope (BAT) using two analyses: a template-based matched-filter search using the `pyGRB` pipeline and a generic transient search using `x-pipeline`. The `pyGRB` pipeline is used for analyzing short and ambiguous GRBs, while `x-pipeline` is used for all GRBs, short, ambiguous or long. The GRBs sample for a LIGO-Virgo observing run is collected from the Fermi and Swift GRB catalogs, and the best sky localization and timing information is used. The GRB classifications are determined based on T_{90} (and their associated δT_{90}), the time interval over which 90% of the total background-subtracted photon counts are observed by the reporting GRB detector. GRBs are considered short if $T_{90} + |\delta T_{90}| < 2$ s, long if $T_{90} - |\delta T_{90}| > 4$ s, and ambiguous if they lie in between. In particular, it does not account for short GRBs that are followed by periods of extended emission (Norris & Bonnell, 2006; M. H. P. M. van Putten, Lee, Della Valle, Amati, & Levinson, 2014), for which measures of T_{90} may substantially exceed these thresholds. For more robust

classification one must also consider spectral properties, most commonly the spectral hardness or peak energy of the event, but since our sample consists of observations from multiple observatories with different spectral sensitivities we do not employ such quantities when organizing our sample.

5.2.1 X-Pipeline

One of the analyses searching for GWs associated with GRBs uses the generic transient search library `x-pipeline`. This pipeline analyzes GW strain data from multiple observatories around the time of a GRB. This is an unmodeled method for finding GWs, as it does not rely on template-based matched-filtering. Instead, given a GRB event, `x-pipeline` searches for excess power coherent between LIGO-Virgo detectors and consistent with the sky localization and time window of the GRB. The search time window starts 600 s before the GRB trigger time and ends at 60 s after trigger time, or at T_{90} if $T_{90} > 60$ s. This is sufficient to cover the time delay between GW emission from a progenitor and any GRB prompt emission (Aloy, Müller, Ibáñez, Martí, & MacFadyen, 2000; Burlon, Ghirlanda, Ghisellini, Greiner, & Celotti, 2009; Burlon et al., 2008; Koshut et al., 1995; Lazzati, 2005; Lazzati, Morsony, & Begelman, 2009; MacFadyen, Woosley, & Heger, 2001; Vedrenne & Atteia, 2009; Wang & Mészáros, 2007; Zhang, Woosley, & MacFadyen, 2003). While some GW emissions, such as from CCSN, are expected to reach frequencies up to a few kilohertz (Radice, Morozova, Burrows, Vartanyan, & Nagakura, 2019), we restrict our search frequency range to the most sensitive band of the GW detectors, 20–500 Hz, since detecting such signals above a few hundred hertz requires extremely high GW energies (Abbott et al., 2019) and expanding the frequency range would also significantly increase the computational cost. To

constrain the sky location of the GRB event, the search is using a grid based on the best localization known either from Fermi GBM or Swift BAT.

`X-pipeline` produces time–frequency maps of the GW data coherently combined between the detectors. These maps give access to the temporal evolution of the spectral properties of the signal and enable the pipeline to search for clusters of pixels containing excess energy. The pipeline assigns each cluster a detection statistic based on energy and ranks them accordingly. A coherent consistency test, based on correlations between data in different detectors, then vetoes clusters that are associated with noise transients. The surviving cluster with the largest ranking statistic is the best candidate for a GW detection, and the search quantifies its significance as the probability of the event being produced by the background alone. This is determined by comparing the SNR of the trigger within the 660 s on-source window to the distribution of the SNRs of the loudest triggers in the 660 s off-source windows. As a requirement, the off-source data consist of at least $\tilde{1.5}$ hr of coincident data from at least two detectors around the time of a GRB. This is small enough to select data where the detectors should be in a similar state of operation as during the GRB on-source window, and large enough so that probability estimates using artificial time-shifting of the data are at the sub-percent level.

We quantify the sensitivity of the generic transient search by injecting simulated signals into off-source data. For each waveform family injected we determine the largest significance of any surviving cluster associated with the injections. We compute the percentage of injections that have a significance higher than the best event candidate and look for the amplitude at which this percentage is above 90%, which sets the upper limit. We include O3b calibration

errors (Acernese et al., 2022; Sun et al., 2021) by jittering the amplitude and arrival time according to a Gaussian distribution representative of the calibration uncertainties. As with the modeled search, these injection sets allow us to calculate the 90% exclusion distance, D_{90} , for each injection waveform. These D_{90} estimates represent the distance within which our null result is 90% likely to exclude the existence of a GW signal caused by the emission mechanism for that waveform.

We choose simulated waveforms to cover the search parameter space of three distinct sets of circularly polarized GW waveforms: BNS and NSBH binary inspiral signals, stellar collapse signals, and disk instability signals.

Circular sine-Gaussian (CSG) injections represent GW emission from stellar collapses defined in Equation (1) of Abbott et al. (2017) with a Q factor of 9 and varying center frequency of 70, 100, 150, and 300 Hz. In all cases, we assume an optimistic emission of energy in GWs of $E_{\text{GW}} = 10^{-2}M_{\odot}c^2$. This set of waveforms covers a number of emission models for short-duration GWs occurring during the collapse of the star.

Binary inspiral injections are characterized by a Gaussian distribution centered at $1.4 M_{\odot}$, with a width of $0.2 M_{\odot}$ for an NS in a BNS, and with a width of $0.4 M_{\odot}$ for an NS in an NSBH. The distribution for GWs emitted by BNS mergers addresses the case of short GRB events as in Abbott et al. (2017) and adopted in the pyGRB search.

Long-duration Accretion disk instability (ADI) injections are used to represent GWs produced by instabilities in the magnetically suspended torus around a rapidly spinning BH (M. H. van Putten, 2001; M. H. van Putten et al., 2004). The model for these waveforms is parameterized by the mass M and dimensionless spin parameter χ of the central BH, and the fraction ϵ of the

accretion disk mass (which is fixed at $1.5 M_{\odot}$) that forms clumps. The parameters used to generate the five families of ADI signals are shown in Table 3, along with the duration, frequency, and E_{GW} of each waveform.

Table 3. Parameters and properties for accretion disk instability waveform injections.

| Waveform Label | $M (M_{\odot})$ | χ | ϵ | Duration (s) | Frequency (Hz) | $E_{\text{GW}} (M_{\odot}c^2)$ |
|----------------|-----------------|--------|------------|--------------|----------------|--------------------------------|
| ADI-A | 5 | 0.30 | 0.050 | 39 | 135–166 | 0.02 |
| ADI-B | 10 | 0.95 | 0.200 | 9 | 110–209 | 0.22 |
| ADI-C | 10 | 0.95 | 0.040 | 236 | 130–251 | 0.25 |
| ADI-D | 3 | 0.70 | 0.035 | 142 | 119–173 | 0.02 |
| ADI-E | 8 | 0.99 | 0.065 | 76 | 111–234 | 0.17 |

5.3 O3b search for GWs associated with GRBs

Since O3 was split into two halves, O3a and O3b, the LIGO-Virgo GRB search was also split in two. Changes in the detector during the in-between commissioning phase resulted in minor improvements to the sensitivity of the LIGO detectors, and a major change was made to x-pipeline to deal with noise issues.

In the O3a search, the sensitivity to long-duration (≥ 10 s) signals was often limited by loud background noise transients known as glitches (Davis et al. 2021). While X-Pipeline’s coherent consistency tests easily veto these glitches, many long-duration simulated signals would overlap such a glitch by chance. In these cases the simulated signal and glitch would be clustered together and subsequently vetoed together. To address this problem, an “autogating” procedure was implemented for O3b. For each detector, the total energy in the whitened data stream is computed over a 1 s window. If this total fluctuates by more than 50 standard deviations above the median value, then the data is zeroed out over the

interval where the threshold is exceeded. A 1 s inverse Tukey window is applied at each end of the zeroed interval to transition smoothly between the whitened and zeroed data. To minimize the possibility of a loud GW transient triggering a gate, the procedure cancels a gate if there is a simultaneous energy excursion above 10 standard deviations in any other detector. The threshold of 50 standard deviations is low enough to gate the most problematic loud glitches, while being high enough that the only GWs zeroed out by the gate would have been detectable by all-sky searches. Empirically we find that this procedure is effective at reducing the impact of loud glitches without affecting the sensitivity to low-amplitude GW signals.

In the O3b search, autogating was performed on more than half of the GRBs analyzed by `x-pipeline`. Even though loud injections zeroed out by the glitch would have been detectable by all-sky searches, we ran `x-pipeline` on each GRB without autogating first, and performed an autogated rerun only if the closed box results suggest that autogating would improve the sensitivity of the search. Running `x-pipeline` itself produces a “closed-box” results page, which summarizes the results of the search on a “dummy” on-source window chosen from among the off-source windows. This allows us to tune the parameters of the search, the injections, or the pre- and postprocessing to produce the most sensitive closed-box results before actually analyzing the on-source data (the “open-box” analysis). The glitches targeted by autogating affect the ability of `x-pipeline` to detect long-duration signals, so we can determine if a closed-box analysis would perform better with autogating applied by looking for a clear plateau or dip in the detection efficiency curves of the long-duration ADI waveforms. If such a problem does exist, then the closed-box analysis is rerun with autogating, and the box with better efficiency curves (usually the autogated version) is opened. Otherwise, if

autogating is not likely to affect the results, then we open the box for the non-autogated analysis.

5.3.1 GRB sample

The full sample of GRBs occurring in O3b consists of seven short GRBs, 12 ambiguous GRBs, and 89 long GRBs. Of these, only two have known redshifts: GRB 191221B ($z = 1.148$) (Kuin & Swift/UVOT Team, 2019; Vielfaure et al., 2019) and GRB 200205B ($z = 1.465$) (Vielfaure & Stargate Collaboration, 2020). Since x-pipeline searches for coincident excess power, we perform the generic transient search for GRBs where at least two of the three LIGO-Virgo detectors were active. This leads to 86 GRBs to analyze and is also compatible with the network observing time of at least two detectors (85.3

5.3.2 Results of the O3b search

We rank each candidate by calculating a p-value, the probability of an event or a louder one in the on-source data, given the background distribution, under the null hypothesis. The p-value is calculated by counting the fraction of background trials that contain an event with a greater signal-to-noise ratio than that of the loudest on-source event. Figure 22 shows the distribution of p-values for the 86 GRBs analyzed by x-pipeline. In this plot, a significant event would appear at a much lower p-value in the lower left corner of the plots, and be outside (to the left) of the 90% confidence region. The plot shows that the p-value distribution is consistent with the background. The lowest reported p-value found during O3b for the generic transient search was 7.95×10^{-3} (GRB 200224B). Although this p-value is very small, it is not unexpected given the high number of GRBs analyzed.

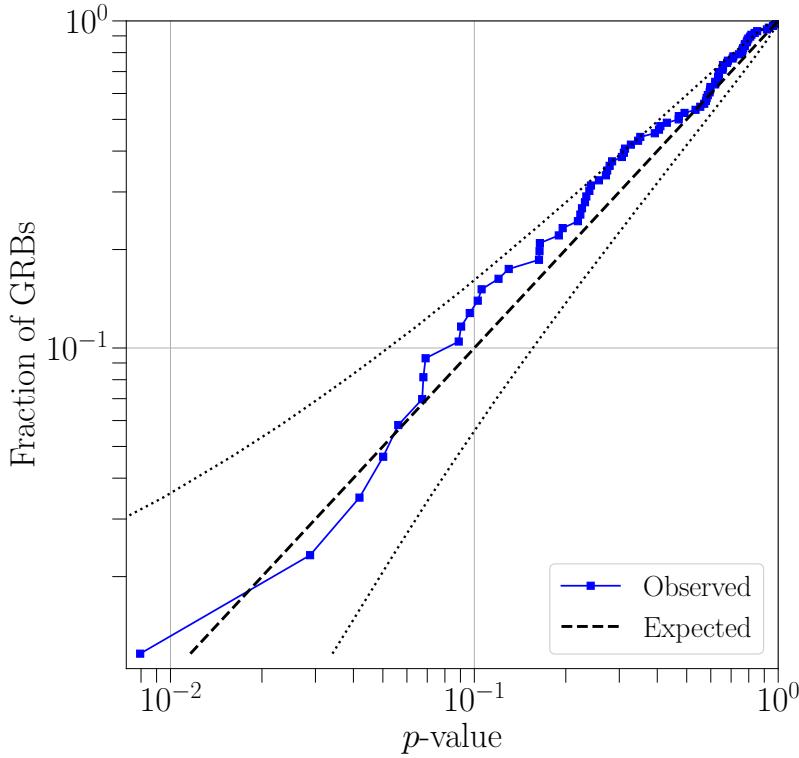


Figure 22. Cumulative distribution of p-values for the loudest on-source events of the O3b X-pipeline analyses. The dashed line indicates an expected uniform distribution of p-values under a no-signal hypothesis, with the corresponding 90% band as the dotted lines.

Given that no loud GW signals are observed coincident with any of the GRBs in this search, we perform a weighted binomial test to determine the probability of observing our set of p-values assuming a uniform background distribution (Abadie et al., 2012). A small probability would suggest that there may be a population of subthreshold GW signals that our search did not identify. This type of weighted binomial test uses the lowest re-weighted p-values from the searches. For the generic transient search, the test gives a probability of 0.76. The same test carried out in O3a returned a probability of 0.30 (Abbott et al., 2021b). In O2 (removing GW170817/GRB 170817A) and the first observing run of Advanced LIGO and Advanced Virgo (O1) the probabilities were 0.75 and 0.75,

respectively (Abbott et al., 2017; B. P. Abbott et al., 2019). As in these previous analyses, the probabilities obtained in O3b suggest that no weak GWs can be attributed to the population of GRBs.

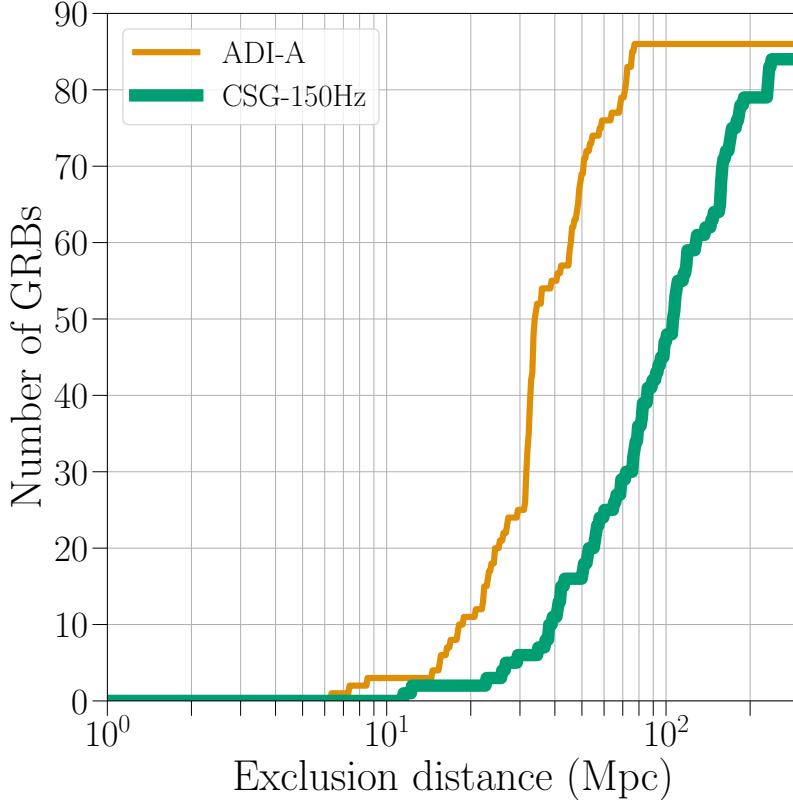


Figure 23. Cumulative distributions of O3b exclusion distances for 150-Hz sine-Gaussian and ADI-A waveforms.

We derive a 90% confidence level lower limit on the distance for each of the 86 GRBs analyzed with the generic transient search, based on the different emission models. Figure 23 shows the distribution of D_{90} values for the ADI-A model and for a CSG with central frequency of 150 Hz. The limits reported depend on the sensitivity of the instruments in the network, which change with time and sky localization of the GRB events. We marginalize these limits over errors introduced by detector calibration. Table 4 reports the median D_{90} , for the set of GRBs for

the different signals. The limits vary by nearly an order of magnitude due to the variety of signals used in our analysis. On average, the median values for the O3b generic transient search are about 50% greater than those reported in O3a (Abbott et al., 2021b).

Table 4. Median 90% exclusion distances (D_{90}) for the generic transient search during O3b.

| | CSG 70 Hz | CSG 100 Hz | CSG 150 Hz | CSG 300 Hz |
|----------------|-----------|------------|------------|------------|
| D_{90} [Mpc] | 166 | 126 | 92 | 42 |
| | ADI-A | ADI-B | ADI-C | ADI-D |
| D_{90} [Mpc] | 34 | 140 | 54 | 22 |
| | ADI-E | | | |
| D_{90} [Mpc] | | | 52 | |

We can primarily attribute this improvement to the use of autogating in O3b: the increase in exclusion distances is highest (up to a factor of 2) for the longest-duration waveforms, which are most impacted by the glitches removed by autogating. The exclusion distances for the shorter-duration CSG waveforms, which are not expected to be affected by autogating, increased by about 30% on average. This is more than could be accounted for by chance differences in the LIGO–Virgo antenna factors between the two samples. Rather, the increase is likely due to improvements in the performance of the detectors themselves, such as through the reduction of noise caused by scattered light in the LIGO detectors Soni et al. (2020) or the improvement in sensitivity of the Virgo detector Davis et al. (2021).

5.3.3 Model exclusion

Although x-pipeline is an entirely unmodeled GW search pipeline, given the substantial improvement in the exclusion distances in O3b, we may be approaching the point at which extreme GW emission models can be excluded

using out null results and the sensitivity estimates for injected waveforms. Using detection efficiency curves of both the O3a and O3b searches, we can compute for each model an exclusion confidence using the method described in Kalmus, Zanolin, and Klimenko (2013) for supernova searches (Abbott et al., 2019). The model exclusion probability given N targeted GRBs is

$$P_{\text{excl}} = 1 - \prod_{i=1}^N [1 - \varepsilon_i(d_i)] \quad (5.1)$$

where $\varepsilon_i(d_i)$ is the detection efficiency at the source distance of d_i . Of the 86 GRBs analyzed in the O3b search by `x-pipeline`, only one, GRB 191221B has a redshift measurement (GRB 200205B did not occur at a time when at least two detectors were active and was therefore not included). For the remaining GRBs, we have no distance measurement, but instead we can sample d_i from the distribution of redshifts measured by Swift BAT.

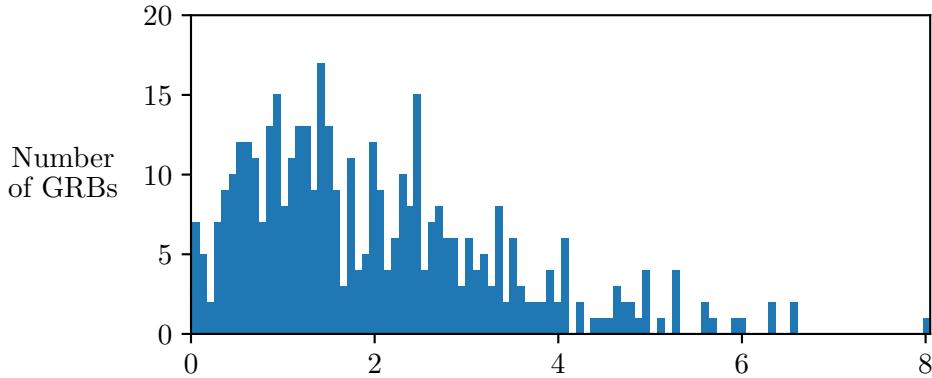


Figure 24. Histogram of redshift measurements for GRBs detected by Swift BAT.

The Swift GRB archive has redshift measurements for 411 triggers *Swift GRB Archive* (n.d.). Figure 24 shows the distribution of those measurements, from which we sample distances for the 86 GRBs analyzed, using inverse transform

sampling, and compute exclusion probability for each model. This is repeated 1,000 times and P_{excl} is averaged for each model across all trials.

Table 5. Exclusion confidence for each injected waveform model.

| | CSG 70 Hz | CSG 100 Hz | CSG 150 Hz | CSG 300 Hz |
|-------------------|-------------|-------------|------------|------------|
| P_{excl} | 0.74 | 0.63 | 0.44 | 0.12 |
| | ADI-A | ADI-B | ADI-C | ADI-D |
| P_{excl} | 0.04 | 0.61 | 0.15 | 0.00 |
| | ADI-E | | | |
| P_{excl} | | 0.18 | | |

The averages are presented in Table 5. The model with the highest exclusion probability is the 70-Hz sine-Gaussian model. Assuming that the Swift redshift distribution is representative of the GRBs analyzed by x-pipeline, it is more likely than not that we can rule out the very optimistic energies simulated by the low-frequency 70-Hz and 100-Hz sine-Gaussian injections ($E_{\text{GW}} = 10^{-2} M_{\odot} c^2$) based on the lack of GW-GRB joint detections at O3 sensitivity.

If search sensitivity continues to improve in O4, we may well be able to begin ruling out the most extreme emission models. The exclusion probability for the ADI-B model is 0.61, making it the only unfavorable ADI model so far. This accretion disk instability model is simulated for a central BH mass of $M = 10 M_{\odot}$, a dimensionless spin parameter of $\chi = 0.95$, and a very large clump mass of $\epsilon = 0.2$ times that the central BH. It is the most extreme of the ADI models, and it would not be surprising to see it being ruled out in future searches.

5.3.4 Noise effects in the generic transient search

Comparing the exclusion distances of the O3a and O3b search sheds some light on the impact of detector and search pipeline improvements on the sensitivity

of the x-pipeline analysis. Table 6 presents the improvement in the median D_{90} for each waveform model as a fraction of the O3a value. Given the large number of GRBs analyzed, it would be unlikely for the difference to be explained by chance improvement in antenna response (i.e. due to more GRBs lining up with the sky region where the GW detector network is most sensitive). x-pipeline reports the antenna response of the GW detectors for each GRB analyzed. These can be summed up over the active detectors for each GRB and averaged over the full GRB sample. The result is a 9.2% improvement in antenna response, which is not negligible but can only account for a small portion of the differences in exclusion distances.

Table 6. Relative increase in median D_{90} for each x-pipeline simulated waveform.

| | CSG 70 Hz | CSG 100 Hz | CSG 150 Hz | CSG 300 Hz |
|-----------------------------------|-----------|------------|------------|------------|
| $\Delta D_{90}/D_{90,\text{O3a}}$ | 0.12 | 0.19 | 0.25 | 0.49 |
| | ADI-A | ADI-B | ADI-C | ADI-D |
| $\Delta D_{90}/D_{90,\text{O3a}}$ | 0.43 | 0.14 | 0.85 | 0.92 |
| | ADI-E | | | |
| $\Delta D_{90}/D_{90,\text{O3a}}$ | 0.55 | | | |

For the longest duration ADI-C, ADI-D, and ADI-E waveforms, we see the greatest improvements, over 50% higher than in O3a (Figure 25, top). For short-duration waveforms, the highest-frequency 300-Hz sine-Gaussian injections show the greatest change (Figure 25, bottom). These are all well above the change expected from a chance difference in antenna response.

Considering that longer-duration waveforms see greater improvement, the most obvious likely contributor is the implementation of autogating. However, autogating cannot account for the changes in the D_{90} of short-duration models. These can only be explained by better GW detector sensitivity and/or reduced

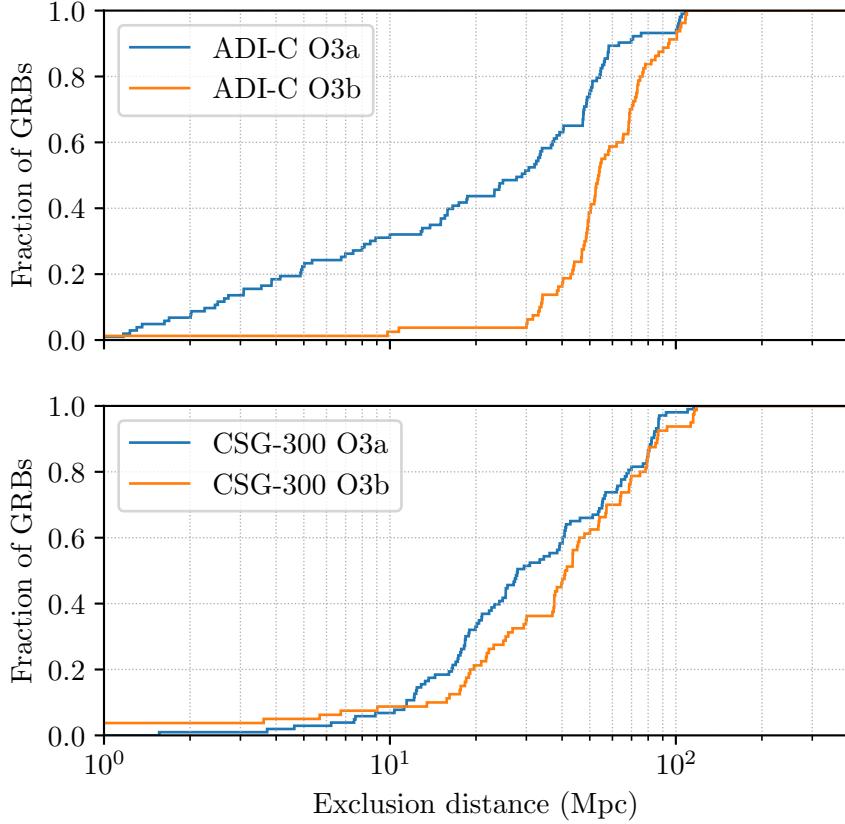


Figure 25. Cumulative distributions of O3a and O3b exclusion distances for the ADI-C (top) and 300-Hz sine-Gaussian (bottom) waveforms. These show the greatest improvement between O3a and O3b among their respective waveform sets.

rates of short-duration glitches that overlap the injected signal. Such glitches cause the entire injection to be vetoed, so it is not recovered by the search pipeline. In particular, glitches caused by low-frequency scattering noise (such as those discussed in Section 4.2) can pepper a stretch of time with short-duration glitches, overlapping injections made at multiple points. A method for reducing the occurrence of these glitches, reaction chain (RC) tracking, was implemented at for O3b, on Jan 7, 2020 at LLO and Jan 15, 2020 (Soni et al., 2020).

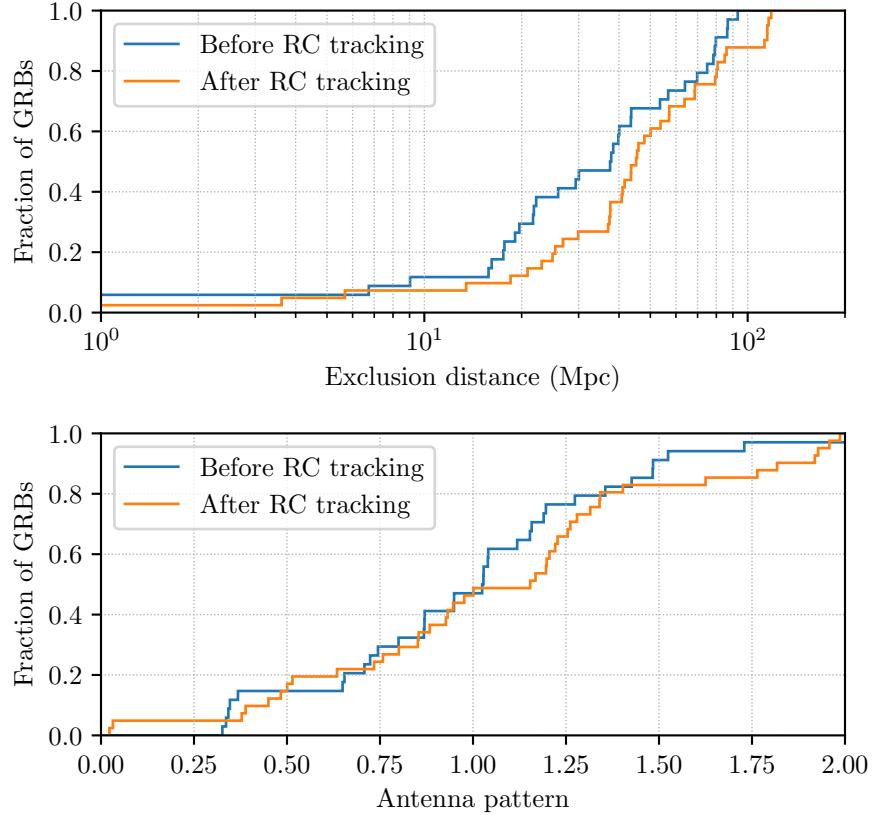


Figure 26. Cumulative distributions of CSG-300 Hz exclusion distances (top) and the detector network antenna factors (bottom) for GRBs before and after the implementation of RC tracking.

Figure 26 gives evidence for the effect of glitch mitigation on x-pipeline sensitivity. Although there is clearly a bias towards higher antenna responses after RC tracking, they do not entirely account for the huge substantial increase in exclusion distances. In summary, noise mitigation is a crucial part of improving the sensitivity of the generic transient search. As discussed above, there may plenty that can be inferred even in the absence of detections. Better detector sensitivities in O4, along with the development of new glitch subtraction methods, could allow more confident exclusion of GW emission models associated with GRBs.

REFERENCES CITED

- Abadie, J., et al. (2012, November). Search for Gravitational Waves Associated with Gamma-Ray Bursts during LIGO Science Run 6 and Virgo Science Runs 2 and 3. *ApJ*, *760*(1), 12. doi: 10.1088/0004-637X/760/1/12
- Abbott, B. P., Abbott, R., Abbott, T. D., Acernese, F., Ackley, K., Adams, C., ... et al. (2017, Oct). Gravitational waves and gamma-rays from a binary neutron star merger: Gw170817 and grb 170817a. *ApJ*, *848*(2), L13. Retrieved from <http://dx.doi.org/10.3847/2041-8213/aa920c> doi: 10.3847/2041-8213/aa920c
- Abbott, B. P., et al. (2016a, Jun). Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914. *Class. Quantum Grav.*, *33*(13), 134001.
- Abbott, B. P., et al. (2016b, Feb). Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, *116*, 061102.
- Abbott, B. P., et al. (2017, November). A gravitational-wave standard siren measurement of the Hubble constant. *Nature*, *551*(7678), 85-88. doi: 10.1038/nature24471
- Abbott, B. P., et al. (2017, Oct). GW170817: Observation of gravitational waves from a binary neutron star inspiral. *Phys. Rev. Lett.*, *119*, 161101.
- Abbott, B. P., et al. (2017, June). Search for Gravitational Waves Associated with Gamma-Ray Bursts during the First Advanced LIGO Observing Run and Implications for the Origin of GRB 150906B. *ApJ*, *841*(2), 89. doi: 10.3847/1538-4357/aa6c47
- Abbott, B. P., et al. (2019, July). All-sky search for short gravitational-wave bursts in the second Advanced LIGO and Advanced Virgo run. *Phys. Rev. D*, *100*(2), 024017. doi: 10.1103/PhysRevD.100.024017
- Abbott, B. P., et al. (2019, Sep). GWTC-1: A gravitational-wave transient catalog of compact binary mergers observed by LIGO and Virgo during the first and second observing runs. *Phys. Rev. X*, *9*, 031040. Retrieved from <https://link.aps.org/doi/10.1103/PhysRevX.9.031040> doi: 10.1103/PhysRevX.9.031040
- Abbott, B. P., et al. (2019, November). Search for Gravitational-wave Signals Associated with Gamma-Ray Bursts during the Second Observing Run of Advanced LIGO and Advanced Virgo. *ApJ*, *886*(1), 75. doi: 10.3847/1538-4357/ab4b48

- Abbott, B. P., et al. (2020, March). GW190425: Observation of a Compact Binary Coalescence with Total Mass $\sim 3.4 M_{\odot}$. *ApJ*, *892*(1), L3. doi: 10.3847/2041-8213/ab75f5
- Abbott, R., et al. (2020, September). GW190521: A Binary Black Hole Merger with a Total Mass of $150 M_{\odot}$. *Phys. Rev. Lett.*, *125*(10), 101102. doi: 10.1103/PhysRevLett.125.101102
- Abbott, R., et al. (2021a, Jun). Gwtc-2: Compact binary coalescences observed by ligo and virgo during the first half of the third observing run. *Phys. Rev. X*, *11*, 021053. Retrieved from <https://link.aps.org/doi/10.1103/PhysRevX.11.021053> doi: 10.1103/PhysRevX.11.021053
- Abbott, R., et al. (2021b). GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run.
- Abbott, R., et al. (2021a, July). Observation of Gravitational Waves from Two Neutron Star-Black Hole Coalescences. *ApJ*, *915*(1), L5. doi: 10.3847/2041-8213/ac082e
- Abbott, R., et al. (2021b, July). Search for Gravitational Waves Associated with Gamma-Ray Bursts Detected by Fermi and Swift during the LIGO-Virgo Run O3a. *ApJ*, *915*(2), 86. doi: 10.3847/1538-4357/abee15
- Acernese, F., et al. (2022, February). Calibration of advanced Virgo and reconstruction of the detector strain $h(t)$ during the observing run O3. *Classical and Quantum Gravity*, *39*(4), 045006. doi: 10.1088/1361-6382/ac3c8e
- Aloy, M. A., Müller, E., Ibáñez, J. M., Martí, J. M., & MacFadyen, A. (2000, March). Relativistic Jets from Collapsars. *ApJ*, *531*(2), L119-L122. doi: 10.1086/312537
- APS Dynamics. (2014, Mar). *APS 113 ELECTRO-SEIS Long Stroke Shaker*. https://www.apsdynamics.com/en/products/details/vibration-exciter/aps-113-ab.html?file=files/spektra_relaunch/inhalte/produktgruppen/datenblaetter/APS_113-AB_AB-HF_Data_Sheet_en.pdf.
- Banagiri, S., Covas, P., & Schofield, R. M. S. (2019, May). *PEM with Rain using HAM6 Vac accelerometer*. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=49495>. LIGO.
- Biscoveanu, S., Thrane, E., & Vitale, S. (2020, April). Constraining Short Gamma-Ray Burst Jet Properties with Gravitational Waves and Gamma-Rays. *ApJ*, *893*(1), 38. doi: 10.3847/1538-4357/ab7eaf

- Blackburn, K., et al. (2019, May). *Specification of a common data frame format for interferometric gravitational wave detectors (igwd)*.
<https://dcc.ligo.org/LIGO-T970130>. LIGO Document Control Center.
- Brüel & Kjær. (2021, Jun). *Type 4809 portable shaker/vibration exciter*. <https://www.bksv.com/-/media/literature/Product-Data/bp0231.ashx>.
- Buikema, A., et al. (2020, Sep). Sensitivity and performance of the Advanced LIGO detectors in the third observing run. *Phys. Rev. D*, 102, 062003.
- Burlon, D., Ghirlanda, G., Ghisellini, G., Greiner, J., & Celotti, A. (2009, October). Time resolved spectral behavior of bright BATSE precursors. *A&A*, 505(2), 569-575. doi: 10.1051/0004-6361/200912662
- Burlon, D., Ghirlanda, G., Ghisellini, G., Lazzati, D., Nava, L., Nardini, M., & Celotti, A. (2008, September). Precursors in Swift Gamma Ray Bursts with Redshift. *ApJ*, 685(1), L19. doi: 10.1086/592350
- Cannon, K., et al. (2012, March). Toward early-warning detection of gravitational waves from compact binary coalescence. *Astrophys. J.*, 748(2), 136.
- Chatterji, S., Blackburn, L., Martin, G., & Katsavounidis, E. (2004, sep). Multiresolution techniques for the detection of gravitational-wave bursts. *Class. Quantum Grav.*, 21(20), S1809–S1818.
- Christensen, N. (2018, Nov). Stochastic gravitational wave backgrounds. *Reports on Progress in Physics*, 82(1), 016903. Retrieved from <https://doi.org/10.1088/1361-6633/aae6b5> doi: 10.1088/1361-6633/aae6b5
- Cornish, N. J., & Littenberg, T. B. (2015, June). Bayeswave: Bayesian inference for gravitational wave bursts and instrument glitches. *Class. Quantum Grav.*, 32(13), 135012.
- Creighton, J. D. E., & Anderson, W. G. (2011). *Gravitational-wave physics and astronomy*. Wiley-VCH.
- Dai, Z. G., & Lu, T. (1998, May). Gamma-ray burst afterglows and evolution of postburst fireballs with energy injection from strongly magnetic millisecond pulsars. *A&A*, 333, L87-L90.
- Davis, D., Massinger, T., Lundgren, A., Driggers, J. C., Urban, A. L., & Nuttall, L. (2019, February). Improving the sensitivity of Advanced LIGO using noise subtraction. *Class. Quantum Grav.*, 36(5), 055011.

- Davis, D., et al. (2021, Apr). LIGO detector characterization in the second and third observing runs. *Class. Quantum Grav.*, *38*(13), 135014. Retrieved from <https://doi.org/10.1088/1361-6382/abfd85> doi: 10.1088/1361-6382/abfd85
- Effler, A., Schofield, R. M. S., Frolov, V. V., González, G., Kawabe, K., Smith, J. R., ... McCarthy, R. (2015, Jan). Environmental influences on the LIGO gravitational wave detectors during the 6th science run. *Class. Quantum Grav.*, *32*(3), 035017. doi: 10.1088/0264-9381/32/3/035017
- Farah, A., et al. (2020, June). Counting on Short Gamma-Ray Bursts: Gravitational-Wave Constraints of Jet Geometry. *ApJ*, *895*(2), 108. doi: 10.3847/1538-4357/ab8d26
- Kalmus, P., Zanolin, M., & Klimenko, S. (2013). *Excluding source models with multiple astrophysical observations*.
- Klebesadel, R. W., Strong, I. B., & Olson, R. A. (1973, June). Observations of Gamma-Ray Bursts of Cosmic Origin. *ApJ*, *182*, L85. doi: 10.1086/181225
- Klimenko, S., Yakushin, I., Mercer, A., & Mitselmakher, G. (2008, May). A coherent method for detection of gravitational wave bursts. *Class. Quantum Grav.*, *25*(11), 114029.
- Koshut, T. M., Kouveliotou, C., Paciesas, W. S., van Paradijs, J., Pendleton, G. N., Briggs, M. S., ... Meegan, C. A. (1995, October). Gamma-Ray Burst Precursor Activity as Observed with BATSE. *ApJ*, *452*, 145. doi: 10.1086/176286
- Kouveliotou, C., Meegan, C. A., Fishman, G. J., Bhat, N. P., Briggs, M. S., Koshut, T. M., ... Pendleton, G. N. (1993, August). Identification of Two Classes of Gamma-Ray Bursts. *ApJ*, *413*, L101. doi: 10.1086/186969
- Kruk, J., & Schofield, R. M. S. (2016, Aug). *Environmental Monitoring: Coupling Function Calculator*. <https://dcc.ligo.org/LIGO-T1600387/public>. LIGO Document Control Center.
- Kuin, N. P. M., & Swift/UVOT Team. (2019, December). GRB 191221B: Swift/UVOT redshift. *GRB Coordinates Network*, *26538*, 1.
- Lazzati, D. (2005, February). Precursor activity in bright, long BATSE gamma-ray bursts. *MNRAS*, *357*(2), 722-731. doi: 10.1111/j.1365-2966.2005.08687.x
- Lazzati, D., Morsony, B. J., & Begelman, M. C. (2009, July). Very High Efficiency Photospheric Emission in Long-Duration γ -Ray Bursts. *ApJ*, *700*(1), L47-L50. doi: 10.1088/0004-637X/700/1/L47

- Lynch, R., Vitale, S., Essick, R., Katsavounidis, E., & Robinet, F. (2017, May). Information-theoretic approach to the gravitational-wave burst detection problem. *Phys. Rev. D*, *95*, 104046.
- MacFadyen, A. I., Woosley, S. E., & Heger, A. (2001, March). Supernovae, Jets, and Collapsars. *ApJ*, *550*(1), 410-425. doi: 10.1086/319698
- Macleod, D., Urban, A. L., Coughlin, S., Massinger, T., Pitkin, M., Rngeorge, ... Badger, T. G. (2021). *gwpy/gwpy: Gwpy 2.1.3*. Zenodo. Retrieved from <https://zenodo.org/record/5788702> doi: 10.5281/ZENODO.5788702
- Mogushi, K., Cavaglià, M., & Siellez, K. (2019, Jul). Jet geometry and rate estimate of coincident gamma-ray burst and gravitational-wave observations. *ApJ*, *880*(1), 55. Retrieved from <http://dx.doi.org/10.3847/1538-4357/ab1f76> doi: 10.3847/1538-4357/ab1f76
- Nguyen, P. (2020, Apr). *PEM Coupling Function Tools*. <https://git.ligo.org/pem/pemcoupling>. GitLab.
- Nguyen, P., et al. (2021, jun). Environmental noise in Advanced LIGO detectors. *Class. Quantum Grav.*, *38*(14), 145001. Retrieved from <https://doi.org/10.1088/1361-6382/ac011a> doi: 10.1088/1361-6382/ac011a
- Nguyen, P., & Schofield, R. M. S. (2019, May). *Coupling estimates from PEM injections correctly predict LLO DARM glitch near S190510g from thunder-driven vibration at EY*. <https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=46025>. LIGO.
- Nguyen, P., & Schofield, R. M. S. (2020, Jul). *Coupling estimates of helicopter flyovers at both sites during O3*. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=56285>. LIGO.
- Norris, J. P., & Bonnell, J. T. (2006, May). Short Gamma-Ray Bursts with Extended Emission. *ApJ*, *643*(1), 266-275. doi: 10.1086/502796
- Piezosystems. (n.d.). *Piezo nanopositioning devices, actuators & controllers*. <https://www.piezosystem.com/>.
- Radice, D., Morozova, V., Burrows, A., Vartanyan, D., & Nagakura, H. (2019, May). Characterizing the Gravitational Wave Signal from Core-collapse Supernovae. *ApJ*, *876*(1), L9. doi: 10.3847/2041-8213/ab191a

- Riles, K. (2017, Dec). Recent searches for continuous gravitational waves. *Modern Physics Letters A*, 32(39), 1730035. Retrieved from
<http://dx.doi.org/10.1142/S021773231730035X> doi:
10.1142/s021773231730035x
- Schofield, R., Effler, A., Nguyen, P., et al. (2021). *PEM central*.
<http://pem.ligo.org>.
- Schofield, R. M. S. (2013, Mar). *Investigations of magnetic coupling to the quads*.
<https://dcc.ligo.org/LIGO-G1300300/public>. LIGO Document Control Center.
- Schofield, R. M. S. (2018, Nov). *Lightning, LIGO, and GW150914*.
<https://dcc.ligo.org/LIGO-T1800506/public>. LIGO Document Control Center.
- Schofield, R. M. S., Nguyen, P., Banagiri, S., Merfeld, K., & Effler, A. (2019, May). *HVAC noise in DARM roughly consistent with prediction for septum from PEM injections*.
<https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=48912>. LIGO.
- Schofield, R. M. S., Nguyen, P., Roma, V., Palamos, J., Schale, P., Effler, A., ... Frey, R. (2018, Nov). *PEM vetting report for GW170817*.
<https://dcc.ligo.org/LIGO-T1800508/public>. LIGO Document Control Center.
- Schofield, R. M. S., Roma, V., Palamos, J., Hardwick, T., Effler, A., Frey, R., & Talukder, D. (2018, Nov). *PEM vetting report for GW150914*.
<https://dcc.ligo.org/LIGO-T1800505/public>. LIGO Document Control Center.
- Soni, S., et al. (2020, Dec). Reducing scattered light in LIGO's third observing run. *Class. Quantum Grav.*, 38(02), 025016.
- Sun, L., Goetz, E., Kissel, J. S., Betzwieser, J., Karki, S., Bhattacharjee, D., ... Urban, A. (2021, June). Characterization of systematic error in Advanced LIGO calibration in the second half of O3. *arXiv e-prints*, arXiv:2107.00129.
- Swift grb archive*. (n.d.). NASA. Retrieved from
https://swift.gsfc.nasa.gov/archive/grb_table/
- Taluker, D., Nguyen, P., et al. (2021, Dec). *LIGO Channel Activity Monitor*.
<https://git.ligo.org/pem/ligocam>. GitLab.

- Usman, S., et al. (2016, October). The PyCBC search for gravitational waves from compact binary coalescence. *Class. Quantum Grav.*, *33*(21), 215004.
- van Putten, M. H. (2001, August). Proposed Source of Gravitational Radiation from a Torus around a Black Hole. *Phys. Rev. Lett.*, *87*(9), 091101. doi: 10.1103/PhysRevLett.87.091101
- van Putten, M. H., Levinson, A., Lee, H. K., Regimbau, T., Punturo, M., & Harry, G. M. (2004, February). Gravitational radiation from gamma-ray burst-supernovae as observational opportunities for LIGO and VIRGO. *Phys. Rev. D*, *69*(4), 044007. doi: 10.1103/PhysRevD.69.044007
- van Putten, M. H. P. M., Lee, G. M., Della Valle, M., Amati, L., & Levinson, A. (2014, October). On the origin of short GRBs with extended emission and long GRBs without associated SN. *MNRAS*, *444*, L58-L62. doi: 10.1093/mnrasl/slu113
- Vedrenne, G., & Atteia, J.-L. (2009). *Gamma-Ray Bursts*. doi: 10.1007/978-3-540-39088-6
- Vielfaure, J. B., Arabsalmani, M., Heintz, K. E., Tanvir, N. R., de Ugarte Postigo, A., Malesani, D. B., ... Stargate Collaboration (2019, December). GRB 191221B: VLT/X-shooter redshift. *GRB Coordinates Network*, *26553*, 1.
- Vielfaure, J. B., & Stargate Collaboration. (2020, February). GRB 200205B: VLT/X-shooter redshift. *GRB Coordinates Network*, *26998*, 1.
- Wang, X.-Y., & Mészáros, P. (2007, December). GRB Precursors in the Fallback Collapsar Scenario. *ApJ*, *670*(2), 1247-1253. doi: 10.1086/522820
- Washimi, T., Yokozawa, T., Tanaka, T., Itoh, Y., Kume, J., & Yokoyama, J. (2020). Method for environmental noise estimation via injection tests for ground-based gravitational wave detectors.
- Woosley, S. E. (1993, March). Gamma-Ray Bursts from Stellar Mass Accretion Disks around Black Holes. *ApJ*, *405*, 273. doi: 10.1086/172359
- Zhang, W., Woosley, S. E., & MacFadyen, A. I. (2003, March). Relativistic Jets in Collapsars. *ApJ*, *586*(1), 356-371. doi: 10.1086/367609