LIGO-P1800233-v2

Characterization and Reduction of Noise in Gravitational Wave Detectors

by

Nikhil Mukund K

A thesis submitted for the degree of

Doctor of Philosophy

(in Physics)

Submitted to the

Jawaharlal Nehru University
New Mehrauli Road, Munirka, Delhi 110067, India

Supervisor: **Prof. Sanjit Mitra**

Inter-University Centre for Astronomy and Astrophysics
Post Bag 4, Ganeshkhind, Pune 411007, India

July 2018

Certificate

This is to certify that the thesis entitled "**Characterization and Reduction of Noise in Gravitational Wave Detectors**" submitted by Mr. Nikhil Mukund K for the award of the degree of Doctor of Philosophy of Jawaharlal Nehru University, New Delhi is his original work. This has not been published or submitted to any other University for any other Degree or Diploma.

Prof. Somak Raychaudhury	**Prof. Sanjit Mitra**
(Director of IUCAA)	(Thesis Supervisor)

Declaration by the Candidate

I, hereby declare that the work reported in this thesis is entirely original. This thesis is composed independently by me at the Inter-University Centre for Astronomy and Astrophysics, Pune under the supervision of Prof. Sanjit Mitra. I further declare that the subject matter presented in the thesis has not previously formed the basis for the award of any degree, diploma, membership, associateship, fellowship, or any other similar title of any University or Institution.

Prof. Sanjit Mitra
(Thesis Supervisor)

Nikhil Mukund K
(Ph.D. Candidate)

Abstract

Advanced gravitational wave (GW) interferometers like LIGO and Virgo have reached a sensitivity necessary to detect the coalescence of compact objects like binary black holes and neutron stars. Six confirmed detections, which includes the historic discovery of binary blackhole merger (GW150914) and the inspiral signal from binary neutron stars (GW170817) which also had an electromagnetic counterpart, have provided tremendous insights into the nature of Einstein's theory of gravity and ushered an era of multi-messenger astronomy. Further improvement in sensitivity along with higher uptime will respectively help in probing the similar sources at higher redshifts and at the same time increase the chance of detecting GWs from the burst, continuous and stochastic sources. The current GW detectors are sensitive to astrophysical signals in the frequency range of 25 Hz to a few KHz. Making the detectors sensitive at lower frequencies is essential, where compact binaries spend a significant portion of their orbital evolution. The sensitivity in those frequencies is generally limited by a number of noise sources, one dominant contributor being the seismic waves. Their effect span a wide range of frequencies (30 mHz to 30 Hz) and lead to loss of interferometer lock, misalignment of optics, enhanced levels of scattering noise and local density perturbations leading to Newtonian Noise (NN), a.k.a. Gravity gradient noise. NN in particular, due to its direct coupling to the test masses, cannot be shielded using any kind of vibration isolation and is predicted to be a limiting noise for ground-based detectors in the frequency range 10 - 20 Hz. This thesis explores multiple aspects of these noise sources and discusses various strategies to minimize their effect on both the day-to-day detector operation as well as on GW strain data. This involves designing & placement of optimal sensor arrays for NN cancellation, development of early earthquake warning systems based on accurate ground motion prediction and techniques for system identification using static and adaptive filters. We also look into non-astrophysical transients or glitches which trigger false alarms in various data analysis pipelines lowering the significance of true detections. In this thesis, we access the impact of such transients on the detector, use automated techniques to classify them, obtain their relative distribution in strain & auxiliary channels and probe the reason of origin for some of those glitch classes. Finally, we describe the various machine learning tools that were developed to aid detector characterization and commissioning efforts at the sites.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my guide, Prof. Sanjit Mitra for his constant support and encouragement. Working with you towards my Ph.D. was such a delightful experience!

I am grateful to Prof. Sanjeev Dhurandhar, Prof. Tarun Souradeep & Prof. Sukanta Bose for their insightful comments and suggestions. Special thanks to Prof. Somak Raychaudhury, Prof. Ajit Kembhavi, Prof. Varun Sahni, Prof. Kandaswamy Subramanian, Prof. Thanu Padmanabhan, Prof. A. N. Ramaprakash, Prof. Dipankar Bhattacharya and Prof. R. Srianand for providing support at different stages of my Ph.D.

I am also grateful to Prof. Bala Iyer for being a constant source of inspiration and also thank all the members of the IndIGO Consortium specially Prof. Unnikrishnan, Prof. Ajith Parameswaran, Prof. K. G. Arun, Prof. Archana Pai, Prof. Rajesh Nayak, Prof. Anand Sengupta, Prof. Achamveedu Gopakumar, Prof. Surendra Somala and Arnab Dasgupta.

Very special gratitude to Prof. Rana Adhikari for all his critical suggestions and guidance during my visit to Caltech. Sincere thanks to Koji Arai and Gautam Venugopalan for all their help during my visit to the 40M. I extend my warm gratitude to all my research collaborators especially Jan Harms, Michael Coughlin, Sheelu Abraham and Prof. Ninan Sajeeth Philip.

I take this opportunity to thank the LIGO Livingston commissioning team especially Brian O'Reilly, Valera Frolov, Gabriela González, Shivaraj Kandhasamy, Arnaud Pele, Marie Kasprzack, Anamaria Effler, Adam Mullavey, Denis Martynov, Carl Blair, Stuart Aston, Ryan DeRosa and many others. Many thanks to the LIGO Working groups such as Control Systems, Detector Characterization, Seismic Isolation & Suspension. I also extend my thanks to the members of LIGO Scientific Collaboration especially to Brian Lantz and Dennis Coyne.

Special mention to all the members of the wonderful IUCAA family especially Javed, Kabir, Shabbir, Labani, Sumanta, Shakti, Anirban, Bhooshan, Ruchika, Suman, Swagat, Avyarthana, Prathik, Bera, Jishnu, Sreenath, Jithesh, Vivek, Savithri, Reju, Joe, Biju, Avinash, Nithin, Rinku, Biny, Nayana, Athul, Charles, Karthik, Sujatha, Suvodip, Sabyasachi, Jayanti, Sarah, Yogesh, Ajay, Santosh, Kaustubh, Kaushal and many others not in this list. My sincere thanks

to Roby Cheriyanadu for introducing me to the wonderful world of gravitational waves.

I am grateful to IUCAA administrative staff, specially Niranjan V. Abhyankar, Mukund S. Sahasrabudhe, Balaji V. Sawant, Santosh N. Khadilkar, Rajesh D. Pardeshi, V. Chellathurai, Senith S. Samuel, Swati D. Kakade, Deepika M. Susainathan, Rahul S Gaikwad for their timely support and cooperation.

I am thankful to the Council of Scientific & Industrial Research (CSIR) for the Junior and Senior Research Fellowship support during the tenure of my Ph.D. I would like to acknowledge the support from Navajbai Ratan Tata Trust (NRTT) for supporting my trip to Caltech and LIGO Livingston. Many thanks to the HPC facilities at IUCAA that were utilized for most of my analysis.

I would also like to acknowledge the constant support received from all my friends particularly Vishnu, Sachin, Anin, Sreyas, Vijin, Tony, Arjun, and Harinarayanan.

I would also like to extend my gratitude to all my teachers and gurus especially Maya D Nair, Rema Narayanan, Unnikrishnan Sir, Satheesh Sir, Sreelatha Ma'am, Prof. Raghu Chathanathodi, Prof. Ravi Varma, Prof. Pramod Dominic any many others.

Finally, I would like to thank my parents, grandmother, Madhu mama, Gopika, Gopal and all my family members for their warmth and support.

Contents

List of Figures

List of Tables

List of Publications of Nikhil Mukund K (Limited Author)

[1] **N Mukund**, B O'Reilly, S Somala, and S Mitra. Effect of induced seismicity on advanced gravitational wave interferometers. *In preparation towards submission to Phys. Rev. D*.

[2] **N Mukund**, M Coughlin, J Harms et al. Prediction of surface wave velocities with historical seismic data. *In preparation towards submission to Classical and Quantum Gravity*.

[3] **N Mukund**, S Thakur, S Abraham, A. K. Aniyan, S Mitra, N S. Philip, K Vaghmare, and D. P. Acharjya. An information retrieval and recommendation system for astronomical observatories. *The Astrophysical Journal Supplement Series*, 235(1):22, 2018.

[4] S Biscans, J Warner, R Mittleman, C Buchanan, M Coughlin, M Evans, H Gabbard, J Harms, B Lantz, **N Mukund**, A Pele, C Pezerat, P Picart, H Radkins, and T Shaffer. Control strategy to limit duty cycle impact of earthquakes on the ligo gravitational-wave detectors. *Classical and Quantum Gravity*, 35(5):055004, 2018.

[5] **N Mukund**, S Abraham, S Kandhasamy, S Mitra, and N S Philip. Transient classification in ligo data using difference boosting neural network. *Phys. Rev. D*, 95:104059, May 2017.

[6] M Coughlin, P Earle, J Harms, S Biscans, C Buchanan, E Coughlin, F Donovan, J Fee, H Gabbard, M Guy, **N Mukund**, and M Perry. Limiting the effects of earthquakes on gravitational-wave interferometers. *Classical and Quantum Gravity*, 34(4):044004, 2017.

[7] M Coughlin, **N Mukund**, J Harms, J Driggers, R Adhikari, and S Mitra. Towards a first design of a newtonian-noise cancellation system for advanced ligo. *Classical Quantum Gravity*, 33(24):244001, 2016.

List of Publications of Nikhil Mukund K (LSC Collaboration)

[1] Benjamin P. Abbott et al. Search for Tensor, Vector, and Scalar Polarizations in the Stochastic Gravitational-Wave Background. *Phys. Rev. Lett.*, 120(20):201102, 2018.

[2] Benjamin P. Abbott et al. Full Band All-sky Search for Periodic Gravitational Waves in the O1 LIGO Data. *Phys. Rev.*, D97(10):102003, 2018.

[3] B. P. et al. Abbott. Constraints on cosmic strings using data from the first advanced ligo observing run. *Phys. Rev. D*, 97:102002, May 2018.

[4] P. B. et al. Covas. Identification and mitigation of narrow spectral artifacts that degrade searches for persistent gravitational waves in the first two observing runs of advanced ligo. *Phys. Rev. D*, 97:082002, Apr 2018.

[5] Benjamin P. Abbott et al. GW170817: Implications for the Stochastic Gravitational-Wave Background from Compact Binary Coalescences. *Phys. Rev. Lett.*, 120(9):091101, 2018.

[6] Benjamin P. Abbott et al. All-sky search for long-duration gravitational wave transients in the first Advanced LIGO observing run. *Class. Quant. Grav.*, 35(6):065009, 2018.

[7] B P Abbott et al. Effects of data quality vetoes on a search for compact binary coalescences in Advanced LIGO's first observing run. *Class. Quant. Grav.*, 35(6):065010, 2018.

[8] Benjamin P. Abbott et al. First search for nontensorial gravitational waves from known pulsars. *Phys. Rev. Lett.*, 120(3):031104, 2018.

[9] Benjamin P. Abbott et al. First narrow-band search for continuous gravitational waves from known pulsars in advanced detector data. *Phys. Rev.*, D96(12):122006, 2017. [Erratum: Phys. Rev.D97,no.12,129903(2018)].

[10] B.. P.. Abbott et al. GW170608: Observation of a 19-solar-mass Binary Black Hole Coalescence. *Astrophys. J.*, 851(2):L35, 2017.

[11] Benjamin P. Abbott et al. First low-frequency Einstein@Home all-sky search for continuous gravitational waves in Advanced LIGO data. *Phys. Rev.*, D96(12):122004, 2017.

[12] B. P. Abbott et al. Search for Post-merger Gravitational Waves from the Remnant of the Binary Neutron Star Merger GW170817. *Astrophys. J.*, 851(1):L16, 2017.

[13] B. P. Abbott et al. On the Progenitor of Binary Neutron Star Merger GW170817. *Astrophys. J.*, 850(2):L40, 2017.

[14] B. P. Abbott et al. Estimating the Contribution of Dynamical Ejecta in the Kilonova Associated with GW170817. *Astrophys. J.*, 850(2):L39, 2017.

[15] A. Albert et al. Search for High-energy Neutrinos from Binary Neutron Star Merger GW170817 with ANTARES, IceCube, and the Pierre Auger Observatory. *Astrophys. J.*, 850(2):L35, 2017.

[16] B. P. Abbott et al. A gravitational-wave standard siren measurement of the Hubble constant. *Nature*, 551(7678):85–88, 2017.

[17] B. P. Abbott et al. Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A. *Astrophys. J.*, 848(2):L13, 2017.

[18] B. P. Abbott et al. Multi-messenger Observations of a Binary Neutron Star Merger. *Astrophys. J.*, 848(2):L12, 2017.

[19] B. P. et al. Abbott. Gw170817: Observation of gravitational waves from a binary neutron star inspiral. *Phys. Rev. Lett.*, 119:161101, Oct 2017.

[20] B. P. Abbott et al. GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence. *Phys. Rev. Lett.*, 119(14):141101, 2017.

[21] B. P. Abbott et al. Upper Limits on Gravitational Waves from Scorpius X-1 from a Model-Based Cross-Correlation Search in Advanced LIGO Data. *Astrophys. J.*, 847(1):47, 2017.

[22] Benjamin P. Abbott et al. All-sky Search for Periodic Gravitational Waves in the O1 LIGO Data. *Phys. Rev.*, D96(6):062002, 2017.

[23] A. Albert et al. Search for High-energy Neutrinos from Gravitational Wave Event GW151226 and Candidate LVT151012 with ANTARES and IceCube. *Phys. Rev.*, D96(2):022005, 2017.

[24] Benjamin P. Abbott et al. Search for intermediate mass black hole binaries in the first observing run of Advanced LIGO. *Phys. Rev.*, D96(2):022001, 2017.

[25] Benjamin P. Abbott et al. Search for gravitational waves from Scorpius X-1 in the first Advanced LIGO observing run with a hidden Markov model. *Phys. Rev.*, D95(12):122003, 2017.

[26] Benjamin P. Abbott et al. GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2. *Phys. Rev. Lett.*, 118(22):221101, 2017.

[27] B. P. Abbott et al. Search for Gravitational Waves Associated with Gamma-Ray Bursts During the First Advanced LIGO Observing Run and Implications for the Origin of GRB 150906B. *Astrophys. J.*, 841(2):89, 2017.

[28] Thomas D. Abbott et al. Search for continuous gravitational waves from neutron stars in globular cluster NGC 6544. *Phys. Rev.*, D95(8):082005, 2017.

[29] Benjamin P. Abbott et al. Effects of waveform model systematics on the interpretation of GW150914. *Class. Quant. Grav.*, 34(10):104002, 2017.

[30] Benjamin P. Abbott et al. First search for gravitational waves from known pulsars with Advanced LIGO. *Astrophys. J.*, 839(1):12, 2017. [Erratum: Astrophys. J.851,no.1,71(2017)].

[31] Benjamin P. Abbott et al. Upper Limits on the Stochastic Gravitational-Wave Background from Advanced LIGO's First Observing Run. *Phys. Rev. Lett.*, 118(12):121101, 2017. [Erratum: Phys. Rev. Lett.119,no.2,029901(2017)].

[32] Benjamin P. Abbott et al. Directional Limits on Persistent Gravitational Waves from Advanced LIGO's First Observing Run. *Phys. Rev. Lett.*, 118(12):121102, 2017.

[33] Benjamin P. Abbott et al. All-sky search for short gravitational-wave bursts in the first Advanced LIGO run. *Phys. Rev.*, D95(4):042003, 2017.

[34] Benjamin P. Abbott et al. Exploring the Sensitivity of Next Generation Gravitational Wave Detectors. *Class. Quant. Grav.*, 34(4):044001, 2017.

[35] Benjamin P. Abbott et al. The basic physics of the binary black hole merger GW150914. *Annalen Phys.*, 529(1-2):1600209, 2017.

[36] B. P. Abbott et al. Supplement: The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914. *Astrophys. J. Suppl.*, 227(2):14, 2016.

[37] B. P. Abbott et al. The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914. *Astrophys. J.*, 833(1):L1, 2016.

[38] Benjamin P. Abbott et al. Upper Limits on the Rates of Binary Neutron Star and Neutron Star–black Hole Mergers From Advanced Ligo's First Observing run. *Astrophys. J.*, 832(2):L21, 2016.

[39] Benjamin P. Abbott et al. Results of the deepest all-sky survey for continuous gravitational waves on LIGO S6 data running on the Einstein@Home volunteer distributed computing project. *Phys. Rev.*, D94(10):102002, 2016.

[40] B. P. Abbott et al. Binary Black Hole Mergers in the first Advanced LIGO Observing Run. *Phys. Rev.*, X6(4):041015, 2016.

[41] Thomas D. Abbott et al. Improved analysis of GW150914 using a fully spin-precessing waveform Model. *Phys. Rev.*, X6(4):041014, 2016.

[42] B. P. Abbott et al. Directly comparing GW150914 with numerical solutions of Einstein's equations for binary black hole coalescence. *Phys. Rev.*, D94(6):064035, 2016.

[43] B. P. Abbott et al. Comprehensive all-sky search for periodic gravitational waves in the sixth science run LIGO data. *Phys. Rev.*, D94(4):042002, 2016.

[44] B. P. Abbott et al. Supplement: Localization and broadband follow-up of the gravitational-wave transient GW150914. *Astrophys. J. Suppl.*, 225(1):8, 2016.

[45] B. P. Abbott et al. Localization and broadband follow-up of the gravitational-wave transient GW150914. *Astrophys. J.*, 826(1):L13, 2016.

[46] S. Adrian-Martinez et al. High-energy Neutrino follow-up search of Gravitational Wave Event GW150914 with ANTARES and IceCube. *Phys. Rev.*, D93(12):122010, 2016.

[47] B. P. Abbott et al. GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence. *Phys. Rev. Lett.*, 116(24):241103, 2016.

[48] B. P. Abbott et al. Properties of the Binary Black Hole Merger GW150914. *Phys. Rev. Lett.*, 116(24):241102, 2016.

[49] B. P. Abbott et al. Observing gravitational-wave transient GW150914 with minimal assumptions. *Phys. Rev.*, D93(12):122004, 2016. [Addendum: Phys. Rev.D94,no.6,069903(2016)].

[50] B. P. Abbott et al. GW150914: First results from the search for binary black hole coalescence with Advanced LIGO. *Phys. Rev.*, D93(12):122003, 2016.

[51] B. P. Abbott et al. Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914. *Class. Quant. Grav.*, 33(13):134001, 2016.

[52] B. P. Abbott et al. GW150914: Implications for the stochastic gravitational wave background from binary black holes. *Phys. Rev. Lett.*, 116(13):131102, 2016.

[53] B. P. Abbott et al. GW150914: The Advanced LIGO Detectors in the Era of First Discoveries. *Phys. Rev. Lett.*, 116(13):131103, 2016.

[54] B. P. Abbott et al. Astrophysical Implications of the Binary Black-Hole Merger GW150914. *Astrophys. J.*, 818(2):L22, 2016.

[55] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016.

[56] Benjamin P. Abbott et al. Prospects for Observing and Localizing Gravitational-Wave Transients with Advanced LIGO, Advanced Virgo and KAGRA. *Living Rev. Rel.*, 21:3, 2018. [Living Rev. Rel.19,1(2016)].

This thesis is dedicated to the loving memory of my aunt

Sathi Ramachandran

and to the celestial musician

Mandolin U Shrinivas.

Chapter 1

Introduction

1.1 Need for Gravitational Wave Detectors

Einstein's general theory of relativity provides a geometric picture of gravitation describes it as a manifestation of the curvature of spacetime. It reconciles the special theory of relativity with Newton's laws of gravity in a simple yet elegant manner. The Newtonian theory describes gravity as a force whose effect is instantaneous, and this violates the condition that there exists a maximal invariant speed for the propagation of interactions. Several alternative theories have been proposed to explain the phenomenon of gravity. The basic conditions for any such theory to be successful is that they must be complete, self-consistent, relativistic and must reduce to Newtonian theory under certain limits. Experimental evidence is crucial to validate the correctness of any such theory. Since 1916 several tests have been proposed to test relativity which includes the classical as well as the modern tests. The first two tests of general relativity were the comparison between prediction and observation for the perihelion advance of the planet Mercury (Einstein, 1916) and for the deflection of light near the Sun. In 1959, measurements of the gravitational redshift provided an additional test. All the above tests were proposed by Einstein, and these and subsequent measurements are known as the classical tests of general relativity. Measurement of the Shapiro time delay is sometimes called the "fourth test of general relativity." The need for modern tests of general relativity arose partly because there were other relativistic classical field theories (metric theories) like Brans-Dicke theory which were also in agreement with some of the previous experiments. Consequently, modern tests of gravity were proposed which comprised of laser interferometry, pulsar timing, CMB polarization studies, etc. The detection of gravitational wave (GWs) would be a direct verification of GR and not merely a higher order correction to the Newtonian theory. Apart from confirming GR, their detection will add a new window to observe the universe and hence will truly be a new candidate to the multi-messenger astronomy.

In fact, GWs are not only opening a new observational band but a new spectrum of bands containing information complementary to their electromagnetic counterparts. They will carry information about the internal structure of objects like neutron stars and black holes. GWs can help in determining the value of Hubble constant, constrain the neutron star equation of state, provide details of the central engine of gamma-ray burst and shed light on the early universe & structure formation.

Projects like Advanced LIGO, Advanced VIRGO, KAGRA, GEO600 and the future ones like LIGO-India, Cosmic Explorer, LISA and Einstein Telescope will push the current technologies to the extreme in an attempt to verify the unique predictions made by the general theory of relativity. Advanced LIGO has already reached a sensitivity needed to detect GWs and made historic detections of binary blackhole merger (Abbott, 2016c) and binary neutron star inspiral (Abbott, B. P. et al., 2017c). Joint detection with Advanced Virgo of binary black hole coalescence has also been achieved (Abbott, B. P. et al., 2017b), which made it possible for the first time to learn more about the GW polarizations enabling phenomenological tests of gravity.

Achieving the desired sensitivity necessary to observe these astrophysical signals requires careful understanding of the various noise sources that limit the sensitivity of the instrument at various frequency bands and development of techniques to mitigate their effects. For example, improving the performance at the lower frequencies below 30 Hz would make it possible for the inspiralling compact binaries to spent more cycles within the detection band, thus increasing the signal to noise ratio (SNR) and decreasing the uncertainty in the inferred parameters. In this thesis, we leverage various optimization and data analysis techniques to better understand and tackle some of the sensitivity limiting noise sources. We show how techniques from the field of machine learning (ML) can lead to significant automation and create better models to predict and mitigate the impact of various noise sources on the interferometer.

1.2 Linearized Einstein's Field Equations

Einstein's field equations connecting the spacetime curvature and the mass energy distribution is given by (using geometrized units where G = c = 1)

$$G_{\mu\nu} = 8\pi T_{\mu\nu}\,, \tag{1.1}$$

where $G_{\mu\nu}$ is the Einstein tensor and $T_{\mu\nu}$ is the stress-energy tensor. $G_{\mu\nu}$ is a compact notation which can be written in terms of Ricci tensor ($R_{\mu\nu}$) and Ricci scalar (R),

$$G_{uv} = R_{uv} - \frac{1}{2} R g_{\mu\nu}\,. \tag{1.2}$$

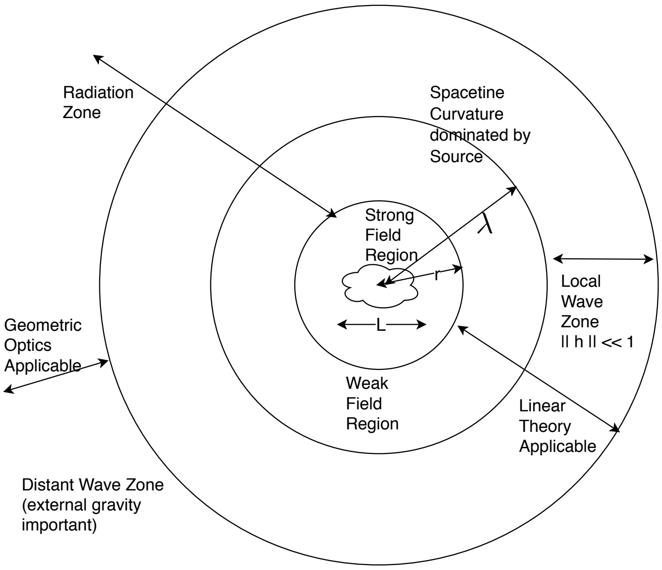


FIGURE 1.1: Gravitational radiation field zones.

When the gravitational field is weak, we can linearize Einstein's field equations around flat spacetime. The metric ($g_{\mu\nu}$) can then be written as a sum of unperturbed flat spacetime metric ($\eta_{\mu\nu}$) with a perturbation term ($h_{\mu\nu}$) in a gauge where $|h_{\mu\nu}| << 1$. Fig. 1.1 shows the various radiation field zones and indicates where the predictions from linearized theory are valid. The trace reversed metric perturbation, $\bar{h}_{\mu\nu}$, is defined as,

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} h\, \eta_{\mu\nu} \qquad \Longrightarrow \quad h = h^{\alpha}_{\alpha} = h_{\alpha\beta}\eta^{\alpha\beta}\,. \tag{1.3}$$

In terms of $\bar{h}_{\mu\nu}$, the Einstein's field equations linearized about flat spacetime becomes,

$$2G_{\mu\nu} = -\bar{h}^{,\alpha}_{\mu\nu,\alpha} - \eta_{\mu\nu}\bar{h}^{\alpha\beta}_{\alpha\beta,} + \bar{h}^{\alpha}_{\mu\alpha,\nu} + \bar{h}^{\alpha}_{\mu\alpha,\mu} = 16\pi T_{\mu\nu}\,. \tag{1.4}$$

There is a striking analogy (Peters and Mathews, 1963) between the linearized Einstein's field equations and Maxwell's electromagnetic field equations,

$$-\Box A_{\mu} = 4\pi J_{\mu}\,, \tag{1.5}$$

where A^{μ} is the Four-current and J^{μ} is the Four-current. Now for gravity in weak field regime, the self gravity of the source can be neglected. Introducing a gravitational Lorenz gauge $\bar{h}'^{\beta}_{\alpha\beta} = 0$ and commuting the derivatives, results in the following equation,

$$-\Box\bar{h}_{\mu\nu} = 16\pi T_{\mu\nu}\,, \tag{1.6}$$

whose solutions describe propagating gravitational waves.

To the lowest-order gravitational waves emitted by a system is given by the quadrupole formula,

$$h_{jk} = \frac{2}{r}\frac{d^2\mathcal{I}_{jk}}{dt^2}\,, \tag{1.7}$$

where $\mathcal{I}_{jk}$ is the second moment of the mass distribution given by

$$\mathcal{I}_{jk} = \int T^{00}x_j x_k dx^3\,. \tag{1.8}$$

In vacuum, Eq. (1.6) becomes $\Box h_{\mu\nu} = 0$, which has the standard plane wave solutions. In the transverse-traceless (TT) gauge the solutions could be written as,

$$h_{\mu\nu}(z,t) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -h_+ & h_\times & 0 \\ 0 & h_\times & h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = h_+\mathbf{e}_+ + h_\times\mathbf{e}_\times\,. \tag{1.9}$$

where $\mathbf{e}_+$ and $\mathbf{e}_\times$ are the two polarizations tensors.

When plane polarized GWs are incident on a ring of particles, they undergo the area-preserving deformation, as shown in Fig. 1.2. Since GWs produce strain in spacetime, the change in dimension is proportional the size of the body. This is why large Michelson interferometers are suitable for the detection of GWs. However, the challenge in detection is enormous because the GWs received on the earth produce extremely tiny strain on terrestrial bodies, as we discuss in the following section.

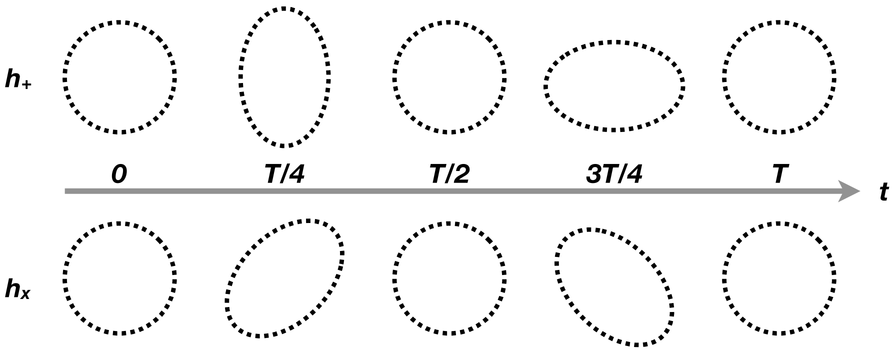


FIGURE 1.2: The effect of GWs (exaggerated) of period T on a ring of particles.

1.3 Sources of Gravitational Waves

In the following section, we discuss the astrophysical origin, typical strength, frequency range of various types of GW sources that are of relevance to the current generation detectors.

1.3.1 Inspiralling Compact Binaries

These are the most promising sources for the current generation for ground and space-based based detectors. The varying quadrupole moment associated with a binary system (NS-NS/BH-BH/NS-BH) leads to the emission to GWs which results in orbital shrinkage. Such orbital decay was first reported by Hulse-Taylor for the PSR 1913+16 (binary neutron star system consisting of a pulsar) which showed the almost perfect alignment with the predictions from the general theory of relativity. The gravitational waveforms from these coalescing binaries can be accurately modeled using analytical and numerical methods making the detection process easier using techniques like matched filtering(Dhurandhar and Sathyaprakash, 1994). In general, the waveform has three distinct phases (inspiral, merger and ringdown) each of which is modeled using a specific technique. The inspiral stage is accurately modeled using Post-Newtonian (PN) theory which is applicable when the gravitational field is weak, and the motion of matter is slow compared to the speed of light. The method involves expanding Einstein's field equations as a power series in terms of deviations from the flat Minkowski metric. But the scheme runs into problems in strong field regime encountered during the merger phase of binaries where the most accurate solutions are obtained using methods of numerical relativity. The technique is powerful enough to obtain the exact solution of coupled nonlinear PDEs but is computationally very expensive. The post-merger ringdown signal from the final blackhole is obtained using blackhole perturbation theory. Typical blackhole binary merger will have an amplitude

$$h_{eff} \sim 3 \times 10^{-22} \left(\frac{500\ Mpc}{r}\right) \left(\frac{\eta}{0.25}\right)^{1/2} \left(\frac{M}{50\ M_{\odot}}\right)^{5/6} \left(\frac{100\ Hz}{f}\right)^{1/6} . \tag{1.10}$$

The first-ever detection of such black hole merger happened on 14th September 2015 jointly by the Advanced LIGO detectors situated at Livingston and Hanford, US (Abbott, 2016c). Similarly, the first binary neutron star inspiral signal was detected on 17th August 2017 (Abbott, B. P. et al., 2017c) whose EM counterpart was picked up by a number of satellites and observatories operating in x-ray, optical and radio frequencies. Detection of a large number of compact binary coalescences can provide answers to a wide variety of unsolved issues in astrophysics and cosmology. It can provide insights about the preferred formation route of binary blackhole binaries which could either have an isolated evolution in galactic fields or could be the result of dynamical capture in dense cluster environments. The masses of the observed blackhole binaries can provide information about the metallicity of its neighborhood environment. Multi-messenger observation using GWs and EM signals for binary neutron star system makes it possible to probe the properties of neutron star matter, heavy element formation in the universe and provide a new independent way to estimate the local expansion rate of the universe (Hubble constant).

1.3.2 Burst Sources

A star is in a state of equilibrium if the inward gravitational force towards the core is balanced by the outward radiation pressure arising from the thermonuclear reactions. This equilibrium is broken once the star runs out of its nuclear fuel leading to a stellar collapse. The associated rapid release of gravitational potential energy can cause the outer layers to heat up which may finally lead to a supernova explosion or a long duration Gamma Ray Burst (GRB). If the collapse is non-spherical, it could lead to GW emission that can take away 10^{-5} *to* 10^{-7} of the total mass-energy. The associated GW emission can roughly have an amplitude as given below

$$h \sim 6 \times 10^{-21} \left(\frac{E}{10^{-7}\, M_\odot}\right)^{1/2} \left(\frac{1\; ms}{T}\right)^{1/2} \left(\frac{1\; kHz}{f}\right) \left(\frac{10\; kpc}{r}\right). \tag{1.11}$$

Although the detectable event rates for these sources could be low, observation of these sources can provide valuable information regarding the nuclear physics such as the equation of state at the core. Because of the lack of accurate waveforms, the best way to detect such burst sources would be using unmodeled searches that look for coincident and coherent excess energy across multiple detectors (Klimenko and Mitselmakher, 2004; Klimenko et al., 2008).

1.3.3 Continuous Sources

Pulsars are the most promising candidates for continuous GW signals. Pulsars are highly magnetized and spherically symmetric neutron stars (NS) formed from the spherical collapse of massive stars that emit highly regular pulses. It can be shown that a slight asymmetry in the form of small bumps on the surface of such rotating neutron stars can cause quadrupole moment variation sufficient to generate detectable gravitational waves at twice the rotational frequency. If a NS of radius R has a mass M with a bump of mass m (fractional asymmetry $\epsilon = (I_{XX} - I_{YY})/I_{ZZ} \sim m/M$, where I is the moment of inertia tensor), the order of magnitude estimate of resulting GW signal is given by,

$$h \sim Mass \times \frac{1}{r} v^2_{nonspherical} = \frac{4}{5}(2\pi m f R)^2 \frac{\epsilon M}{r} , \tag{1.12}$$

and the gravitational wave luminosity

$$L_{GW} \sim \frac{15}{125}(2\pi f)^6 \epsilon^2 M^2 R^4 . \tag{1.13}$$

GW emission takes away the NS rotational energy resulting in a spin down whose time scale is given by

$$t_{spin-down} \sim \frac{1/5 M v^2}{L_{GW}} \sim \epsilon^{-2} f^{-2} \left(\frac{M}{R}\right)^{-1} v^{-3} . \tag{1.14}$$

It needs to be noted that while continuous GWs have not been observed yet, the lack of detection allows us to place upper limits on the asymmetry parameter ϵ leading to scientifically important inferences. Multi-messenger astronomy also has to play an important role here. EM observations provides the spin-down rate, $M.R,\ v,\ \&\ f$ while L_{GW} is probed by GW observations thus constraining ϵ.

1.3.4 Stochastic Background

A Stochastic GW Background (SGWB) can be created from astrophysical as well as primordial sources. The former is made up of random superpositions of GW signals from unresolved astrophysical sources (e.g., distant compact binaries) while the latter could have contributions from early universe events, like cosmic inflation. For ground based observatories, cross-correlation of the data streams between two detectors is the best way to detect stochastic signals from the instrumental background noise. If we make the following assumption about the stochastic background namely (1) Stationary (2) Gaussian (3) Isotropic (4) Unpolarized then we can uniquely characterize it using its power spectral density $H(f)$, defined through the relation

$$< \tilde{h}^*_A(f, \hat{\mathbf{n}}) \tilde{h}_{A'}(f', \hat{\mathbf{n}}') >= \delta(f - f') \, \frac{\delta^2(\hat{\mathbf{n}}, \hat{\mathbf{n}}')}{4\pi} \, \delta A A' \, \frac{1}{2} \, H(f) . \tag{1.15}$$

There is currently a lot of interest in pursuing multi-wavelength study of the stochastic background and understand the dominant contributors to the spectrum. Detection of primordial stochastic GW background (SGWB) would let us probe the universe far beyond what is possible from electromagnetic and neutrino based observations. It will be a smoking gun test of cosmic inflation, an important pillar of the Big Bang model. Advanced LIGO aims to reach a sensitivity close to $\Omega_{gw} \sim 10^{-9}$ at 100 Hz which is still far way from the inflationary signal. Looking for evidence in the cosmic microwave radiation (CMB)'s tensor modes is one way to probe them. GWs leave their imprint in CMB neither through density perturbations nor temperature fluctuations but through polarization signal detectable as curl in the Stokes map. Post-inflation physics like the signature of an electroweak signal will be red-shifted to milliHertz which falls within the LISA frequency band and while the signals from QCD phase transition could be seen in Pulsar Timing Array (PTA) band consisting of periods of few years.

1.4 Noise Sources

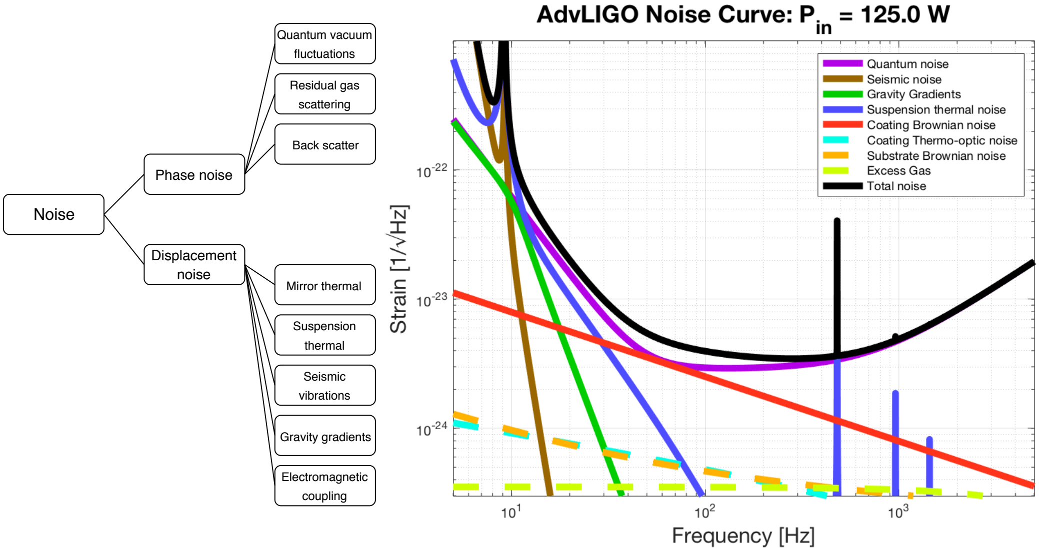


FIGURE 1.3: Fundamental noise sources that limit the sensitivity of advanced GW detectors. The noise budget is constructed using the Gravitational Wave Interferometer Noise Calculator (GWINC) code.

The fundamental noise sources that limit the sensitivity of GW detectors are listed in 1.3 along with their frequency dependent contribution of the Advanced LIGO noise budget as calculated by the GWINC code. It is essential to keep the test mass displacement less than the amplitude of the expected GW signal.

It is also necessary to read out the test mass position to a very high degree of accuracy keeping the back action noise to a minimum. Below we briefly describe some of these noise sources. Detailed discussion about the Newtonian and seismic noise is provided in Chapter 2,5,3,4.

1.4.1 Quantum Vacuum Fluctuations

Arrival times of photons on the photodetector follow a Poissonian distribution which results in fluctuations in the photocurrent and contributes as shot noise. For a simple Michelson interferometer of length L and optical power P in the interferometer arms, the stain equivalent of photon shot noise is given by (Saulson, 1994)

$$h_{sn}(f) = \frac{1}{L}\sqrt{\frac{\hbar c\lambda}{2\pi P}} \,. \tag{1.16}$$

As the photons hit the test masses, there is an associated momentum transfer and the previously mentioned Poissonian nature leads to force fluctuations resulting in photon radiation pressure noise, which has the following amplitude spectral density,(Saulson, 1994)

$$h_{rp}(f) = \frac{1}{mf^2L}\sqrt{\frac{\hbar P}{2\pi^3 c\lambda}} \,. \tag{1.17}$$

The total optical readout noise would be quadrature sum of shot noise and radiation pressure noise, given by,

$$h_{readout}(F) = \sqrt{h_{sn}^2(f) + h_{rp}^2(f)} \,. \tag{1.18}$$

For a given set of design parameters, we would have at every frequency, an optimal power which matches the contributions from the shot noise and the radiation pressure noise. Tracing out this optimum value would lead to a line called Standard Quantum Limit (SQL),

$$S_x(f) = \frac{2\hbar}{m(2\pi f)^2} \,, \tag{1.19}$$

which represents the lower bound quantum noise spectra envelope in classical interferometers. It was shown by Caves (1980) that a complete and consistent quantum mechanical description of optical readout noise requires the consideration of vacuum fluctuations of the electromagnetic field that enters the interferometer through the anti-symmetric port. In this picture, the readout noise is due to the beat between these vacuum fluctuations and the light fields from the stabilized laser.

1.4.2 Back Scatter

Scatter noise is thought to be a contributor to the mysterious excess noise seen from 40 Hz to 100 Hz (Ottaway, Fritschel, and Waldman, 2012a). A simple model for scattering caused due to the micro-roughness of the optics is given in Fig. 1.4. Here x_{AC} is the high-frequency beam tube motion along the direction of the beam and x_G is the low-frequency drift in optical path caused by changes in beam alignment. This is usually enhanced during periods of excess ground motion. The nonlinear nature of this noise is evident from Eq. 1.20 which leads to scattering up-conversions in the frequency range 40 Hz to 250 Hz. For example, Fig. 1.5 shows the side bands from bilinear coupling seen in differential arm motion (DARM) when the auxiliary optics table (HAM) is driven at 203 Hz using PZT shakers and the signal-recycling mirror is driven at a very low frequency to mimic the effects of excess ground motion.

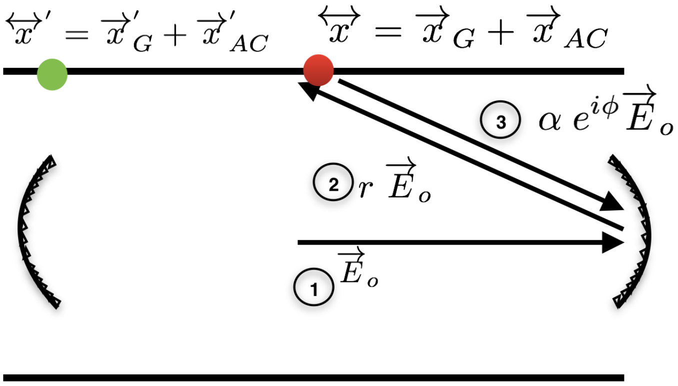


FIGURE 1.4: Back scatter caused due to the micro-roughness of the optics. The red dot is the scatterer and the green dot gives the position of the sensor which picks up the beam tube motion.

The relative intensity noise (RIN) is given by (Accadia, 2010a),

$$\begin{aligned} RIN &= A\sin(\frac{4\pi}{\lambda}x)\ , \\ &= A\sin(\frac{4\pi}{\lambda}[x_G + x_{AC}])\ , \\ &= A\sin(\frac{4\pi}{\lambda}[x_G + B\ sin(2\pi f_o t)])\ . \end{aligned} \tag{1.20}$$

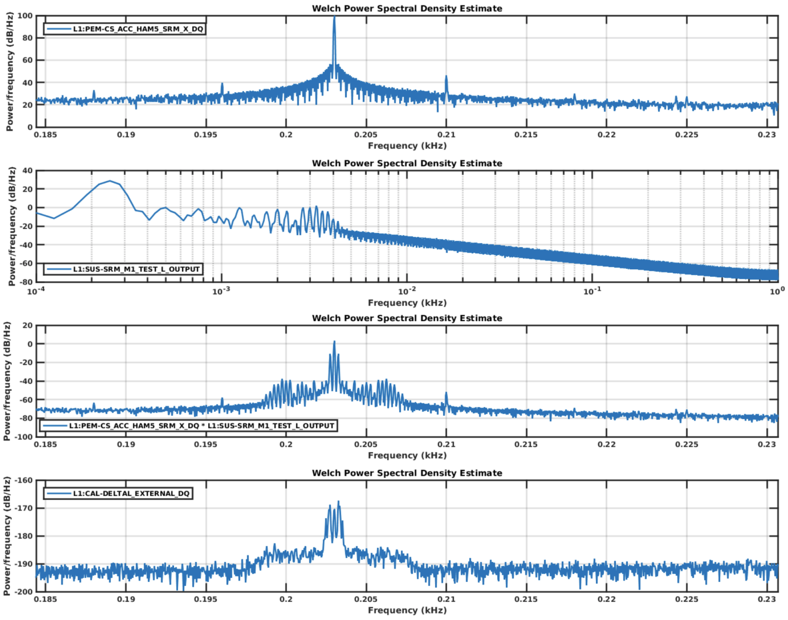


FIGURE 1.5: Modeling the bilinear coupling seen in the GW strain channel.

1.4.3 Thermal noise

Thermal noise (TN) in general has contributions from the substrate, coating materials, and suspension fibers. Intrinsic mechanical losses and temperature fluctuation induced thermoelastic effects are the cause of TN for the first two while the pendulum mode and the violin modes account for the suspension fiber TN. Much of the TN analysis is based on the results from the Fluctuation-Dissipation theorem developed by Callen et al. in 1951. This work depicted the connection between the excitation of a system (fluctuation) and the fictional loss (dissipation) for a linear system in thermal equilibrium.

According to the theorem, the thermal noise power spectral density of a system having a temperature T and an admittance $Z(\omega) = F(w)/v(\omega)$ is given by (Saulson, 1990),

$$S_x(\omega) = \frac{4k_B T}{\omega^2}\mathcal{R}[Z^{-1}(\omega)]\ . \tag{1.21}$$

Mirror thermal noise

The thermal fluctuations leads to mechanical resonances which gets dissipated through internal and external modes of damping. This internal dissipation within an anelastic material is related to the delay between the applied stress(σ) and the corresponding strain(ϵ) that is developed. The applied stress and the corresponding strain are given by

$$\begin{aligned} \sigma &= \sigma_o e^{i\omega t} \;, \\ \epsilon &= \epsilon_o e^{i(\omega t - \phi)} \;. \end{aligned} \tag{1.22}$$

This leads to a frequency dependent complex modulus of elasticity($M = \sigma/\epsilon$),

$$M = \frac{\sigma}{\epsilon} = M_1(\omega) + iM_2(\omega) \;. \tag{1.23}$$

The loss angle (ϕ) which tells us how much the strain is lagging behind the strain is given by

$$\tan\phi = \frac{M_2}{M_1} \;. \tag{1.24}$$

The anelastic response of a material to an applied force results in internal dissipation which can be represented by a modified form of Hooke's Law. The internal source of dissipation arising from Coating Brownian TN is expected to limit the performance of Advanced LIGO detectors in the frequency range 50 Hz to 100 Hz.

The displacement noise power spectral density arising from coating TN has the following form

$$S_x(f,T) \approx \frac{2k_B T}{\pi^2 f}\frac{d}{w^2 Y}\phi\left(\frac{Y^{'}}{Y} + \frac{Y}{Y^{\prime}}\right) \;. \tag{1.25}$$

Here T, d, w, ϕ respectively represent temperature, coating thickness, laser beam radius and coating mechanical loss. Coatings are currently made of alternating layers of dielectric materials (amorphous glassy oxides) with different refractive indexes. Typically a combination of SiO2 (Silica) and TiO2:: Ta2O5 (Titania-doped Tantala) are used to achieve the desired levels of reflectance while still maintaining a low level of optical loss.

Suspension thermal noise

The mechanical loss in the test mass pendulum suspensions leads to suspension thermal noise (Saulson, 1990; González, 2000; Blair et al., 2012) and these can arise from either the suspension pendulum modes or the suspension violin modes. The loss angle ϕ_{mat} for the suspension material is the sum of bulk, thermoelastic and surface components. Since the lossless gravitational restoring force is much larger than the lossy elastic restoring force, suspension loss angle tends to have a value which is higher than the effective pendulum loss ϕ_{pend}.

$$\phi_{pend} = \phi_{mat}\frac{E_{el}}{E_{grav}+E_{el}} \approx \phi_{mat}\frac{E_{el}}{E_{grav}} \quad (E_{el} << E_{grav})\ . \tag{1.26}$$

We can also express suspension loss in terms of number of suspension elements (N), suspended mass (m), suspension length (L), tension in suspension elements (T_o), Young's modulus of suspension and the cross-section moment of inertia of the suspension element (I) as follows (Blair et al., 2012),

$$\phi_{pend}(\omega) = \phi\omega_{mat}\frac{N\sqrt{T_o Y_o I}}{2mgL}\ . \tag{1.27}$$

1.4.4 Electromagnetic Coupling

One major contributor to EM coupling is the coupling between stray RF fields and the control system electronics used for the pre-stabilized laser (PSL) (The LIGO Scientific Collaboration and Virgo Collaboration, 2016). Laser frequency stabilization is achieved within the PSL by locking it to the length of the mode cleaner cavity. Due to its electronics noise, the laser follows the cavity at high frequencies while lower seismic isolation below at frequencies below the suspension resonance results in cavity following the laser at lower frequencies. The mode cleaner is locked by the Pound-Hall-Drever (PDH) signal which is in turn used by the voltage controlled oscillator (VCO) to the generate proportional RF signal. This signal drives an acousto-optic modulator which corrects for the laser frequency fluctuations. Presence of non-RF connectors, radiation from BALUN transformers, cable without RF ferrite choke, etc. could lead to an excess of stray RF fields within the LVEA (Laser Vacuum Equipment Area). These fields have a frequency close to the VCO signal ($80 \pm 5MHz$ and their beat signal can produce a low-frequency signal in kilohertz which show up as whistle glitches in the strain data.

Ambient magnetic fields follow a $1/f$ behavior which has dominant contributions from AC power line at 60 Hz and higher harmonics in the US. Schumann resonance arising from traveling waves in the Earth's surface-ionosphere cavity has also been detected at 5-50 Hz. They can potentially couple in a coherent fashion to multiple GW detectors through actuation magnets, and their removal would require post-offline Wiener filtering.

1.5 Characterization & Mitigation of Noise Sources

The first major task is to identify the fundamental and technical noise sources that can get into the instrument. For those noise sources that are not specific to GW detectors, instead of reinventing the wheel it would be wise first to try the strategies (or their modified versions) that have previously proven to be successful. This could apply to noises like stray light control, electronic noise, vibration isolation, etc. GW detectors deploy a combination of passive and active systems leveraging on the best aspects of both the schemes. The first step in mitigation techniques would be to try out a passive solution. Example of this would be the baffles for reducing stray light, tuned mass dampers for mechanical vibration isolation, pendulum suspensions, etc. One advantage of passive systems is that they usually have no moving parts and needs little maintenance. But their performance is limited at lower frequencies (below few Hertz). As compared to active isolators, the settling time for passive systems is larger because of their lower values related to Eigen frequency and damping. Active isolation would require a sensor, controller and an actuator that would respectively measure the vibration, generate the appropriate control signal and cancel the vibration by physically moving the system to counteract the effect of the coupling noise. Active systems typically have less settling time and provide good suppression at lower frequencies. But they are sensitive to sensor & actuator noise and requires intelligent algorithms to design the controller. Active systems can further be classified into adaptive or static based on whether the filter coefficients are updated based on the nature of the error signal. Adaptive controls can in principle take into account variations in the plant as well as the coupling functions and change the filter coefficients on the fly. Both static Wiener filter and its adaptive variants have been applied to varying success within LIGO (Driggers et al., 2012). Often the bottleneck is the hardware implementation part where one is constrained to use a limited number of filter coefficients, and filter update equations need to be recalculated fast enough so that the phase delay between the noise stream and the actuator is kept to a minimum.

Post-offline subtraction is the scheme where a pre-trained filter to remove correlated artifacts from the data. Wiener filtering is one such method whose application leads to an improvement of 20 % in the sensitivity of Hanford observatory during the second observation run. Newtonian noise cancellation would also be carried out post-offline through the construction of an optimal Wiener filter. In this thesis, we show how the noise characterization and reduction efforts can benefit a lot from the use of intelligent algorithms. Multi-parameter optimization, especially when the cost function has a large number of local minima, has always been a challenging problem. In Chapter 2 we show a modified version of the Simulated Annealing algorithm, which worked successfully for optimal array placement required for Newtonian noise subtraction. We show the effectiveness of the same algorithm in the system identification through the optimal placement of zeros and poles in Chapter 5.

Machine learning (ML) has previously been successfully applied to a wide variety of problems in various disciplines in science and technology. In Astronomy and Astrophysics, they are used to perform tasks such as automated classification of galaxies & stellar spectra, using deep convolutional neural networks, removal of correlated noise from radio antennae data using Gaussian process regression, telescope image reconstruction using generative adversarial networks, etc. ML can be broadly classified into three main types, Supervised learning, Unsupervised learning and Reinforcement learning. When we have labeled data, supervised learning can be used to train a model which then can be applied to newer datasets. If there is a huge number of training samples and enough computational resources, using deep neural net classifiers would often lead to improved prediction accuracy. These techniques are seen to be useful in image and speech recognition tasks. In Chapter 7 we explore wavelet-based feature extraction along with various kinds of neural networks and use them to classify the transients seen in LIGO's strain and auxiliary channels.

When labeled data is not available, if we can extract significant features from the data, we can still get an idea about the underlying populations within the dataset using unsupervised clustering algorithms. They are extensively used in pattern recognition, recommendation systems, and outlier detection. We make use of unsupervised learning called DBSCAN in Chapter 6 to probe the unusual behavior in LIGO data channels seconds before the lockloss and identify the nature of the lockloss event.

In Chapter 8 we use a natural language processing based information retrieval system where a shallow layer neural net is used to perform semantic learning followed by a K-Means clustering in the word vector space to identify the relevant archival data similar to user query.

The above two techniques can lead to artificial specific intelligence as they can only get good at a certain task. Reinforcement learning (RL) fills this gap and can be used to emulate artificial generic intelligence. RL is starting to find applications in control systems where the algorithms are required to make decisions which can have consequences, and the optimal strategy would be to take those decisions that maximize the reward. This technique could find future applications within GW detectors in areas such as vibration isolation, cavity locking, etc.

This thesis has been assigned LIGO document number LIGO-P1800233. The views and opinions presented in this work are those of the author and does not reflect the scientific opinion of the LIGO Scientific Collaboration (LSC).

Chapter 2

Newtonian Noise

2.1 Introduction

The recent discovery of gravitational waves has increased the importance of improving the low-frequency sensitivity of gravitational wave detectors. The Newtonian noise (NN) or the Gravity Gradient Noise (GGN) can arise from perturbations in density (either ground, atmospheric or moving objects) which lead to fluctuations in the local 'Newtonian' gravitational field around the vicinity of the interferometer.

There could be many routes to this as described in Fig. 2.1.

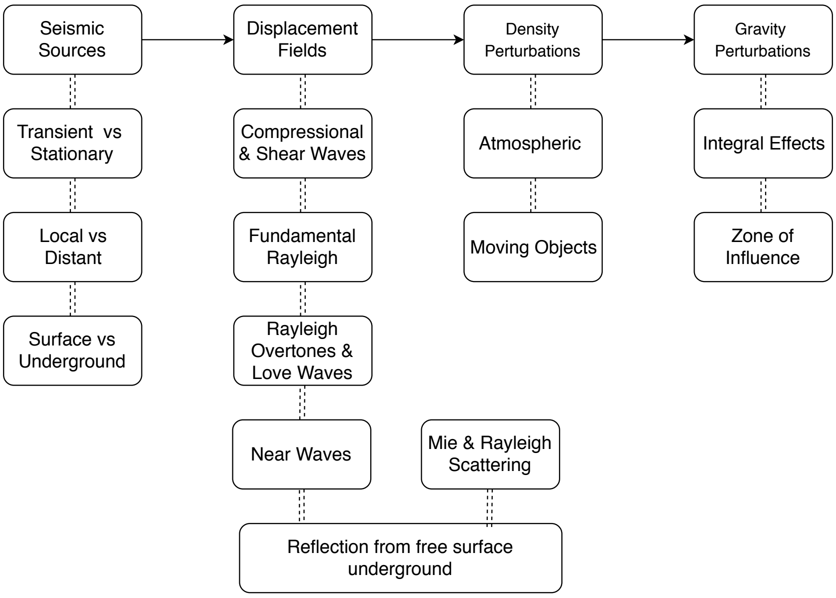


FIGURE 2.1: Various sources of Newtonian Noise

Seismic Newtonian Noise (NN) is anticipated to emerge because of gravitational coupling between seismic ground movement and the test masses, bypassing the whole seismic isolation framework (Saulson, 1984a; Hughes and Thorne, 1998a; Creighton, 2008) and is thought to be a limiting noise source for the advanced detectors between 10 – 20 Hz. One way to circumvent this noise especially for future detectors is to build detectors in sites with very low levels of seismic noise (Beker et al., 2011; Harms et al., 2010; Punturo et al., 2010). But for the second generation detectors, the strategy could be to employ post-offline feed-forward noise cancellation using arrays of seismic sensors measuring the local seismic fields (Cella, 2000).

Previously Driggers et al. (Driggers, Harms, and Adhikari, 2012) explored schemes to mitigate NN by estimating and subtracting it from the interferometer data and found that optimizing the shape of seismic arrays maximized the efficiency of the subtraction. They explored causal and acausal implementations of noise subtraction using simulated seismic data and achieved subtraction of about a factor of 10.

In this work, we focus on achieving the best possible array configurations using a given set of seismic sensors that would provide the maximal noise subtraction at a given frequency of interest. The array optimization technique developed takes into account the nature of seismic fields properties such as levels of homogeneity, isotropy, scattering and generate an appropriate array solution for each case. We also make use of the results from seismic array measurements carried out the LIGO Hanford end-station and found the optimal solution. Along with the Rayleigh plane-wave model used in the previous analysis, we use three more models for which we optimize arrays. Some of the parameters in these models such as the seismic speeds are inferred from the previously mentioned LHO array data.

One question is whether all degrees of freedom of the seismic field can be monitored. We explore the contributions from surface vs. body waves in the analysis below. Another potential issue is the scattering of seismic waves, which can cause difficulty in characterizing the seismic field structure completely. It is also possible that the scattered waves have higher wave numbers compared to the freely propagating surface waves, and therefore a higher density seismic array would be required to monitor the field accurately. Scattering will likely only be a problem if it creates large differences in seismic waveforms over short distances near the test mass, which is unlikely due to a relatively uniform media in the ground near the test masses. The presence of a large number of local minima encountered while finding the optimal array solutions necessitates the need for high-dimensional samplers to compute numerical solutions of the NN equations.

The structure of this chapter is as follows. We review the NN formalism and the models used in this analysis in section 2.2. In section 2.3, we describe how we perform the array optimization using parallel cascaded simulated annealing. Discussion about the inference of values of parameters used in our seismic field models using the previous measurements from an array of accelerometers at the LHO site is done in section 2.4. In section 2.5, we provide the results of analytic and global optimizer for seismic reference fields. Finally we summarize the conclusions in section 2.6. The results presented in this chapter is based on the publication M. Coughlin, N. Mukund, et al. 2016 Class. Quantum Grav. 33 244001.

2.2 Newtonian Noise Models

2.2.1 NN Estimate based on Seismic Spectrum

The following section describes the transfer function, $T(f)$ that describes the effect of seismic induced ground motion ($\tilde{W}(f)$) on the differential test mass motion ($\tilde{x}(f)$) at the GW detectors. For a mass fluctuation $\Delta M(t) = M(t) - < M(t) >$ in a region of size $\lambda/2$ where $\lambda = v_s/f$ is the acoustic wavelength, the gravitational acceleration produced at the test mass is given by

$$\frac{\overrightarrow{\mathbf{F}}(t)}{m} = \frac{G\Delta\mathbf{M}(t)\hat{\mathbf{r}}}{r^2} \,. \tag{2.1}$$

In the Fourier domain, acceleration along the x-direction,

$$\frac{F_x}{m} = \frac{G\Delta\mathbf{M}(\omega_o)\cos\theta}{r^2} \,. \tag{2.2}$$

The left hand side of this equation can be written using the transfer function of a spring mass damper system with a resonant frequency ω_o and a decay time $\tau = \frac{2Q}{\omega_o}$. This leads to

$$\frac{\tilde{x}(\omega)}{F_x} = \frac{1}{m}\frac{1}{-\omega^2 + i\frac{\omega_o}{2Q}w + \omega^2} \,. \tag{2.3}$$

Substituting the above expression in Eq. 2.2 and taking the square modulus, we obtain

$$\left[(\omega^2 - \omega_o^2)^2 + \frac{\omega^2}{\tau^2}\right] x(\omega)^2 = G^2|\Delta\mathbf{M}|^2\frac{\cos^2(\theta)}{r^4} \,. \tag{2.4}$$

Eq. 2.4 gives the contribution from a single fluctuating region with size of order $\lambda/2$. If we assume the each such fluctuating region is independent of each other, we can then sum over such regions and approximate it as an integral by introducing an inner cutoff radius, $r_{min} = \lambda/4$, where $\lambda = v_s/f$ is the acoustic wavelength. This leads to

$$\left[(\omega^2 - \omega_o^2)^2 + \frac{\omega^2}{\tau^2}\right] x(\omega)^2 = \frac{4G^2}{3\pi^3 v_s^4}\omega^4 |\Delta\mathbf{M}(\omega)|^2 \,. \tag{2.5}$$

For $\lambda << L$, the disturbance at each of the test mass can be considered to be uncorrelated and so can be added in quadrature. Hence the total noise contribution to the difference in separation of two pairs of test masses (ignoring the effects of local sources) would lead to a factor of four increase as compared to Eq. 2.5,

$$\left[(\omega^2 - \omega_o^2)^2 + \frac{\omega^2}{\tau^2}\right] x(\omega)^2 = \frac{16G^2}{3\pi^3 v_s^4}\omega^4 |\Delta\mathbf{M}(\omega)|^2 \tag{2.6}$$

$|\Delta\mathbf{M}(\omega)|$ can be rewritten in terms of $|\Delta\mathbf{X}(\omega)|$ as

$$|\Delta\mathbf{M}(\omega)| = \frac{\sqrt{\pi}}{4}\rho_e \lambda^2 |\Delta\mathbf{X}(\omega)| = \pi^{\frac{5}{2}} r^2 \rho_e \frac{|\Delta\mathbf{X}(\omega)|}{\omega^2} \,. \tag{2.7}$$

Thus Eq. 2.4 now takes form

$$\left[(\omega^2 - \omega_o^2)^2 + \frac{\omega^2}{\tau^2}\right] x(\omega)^2 = \frac{16\pi^2}{3} G^2 \rho_e^2 |\Delta\mathbf{X}(\omega)|^2 \,. \tag{2.8}$$

For typical decay time of $10^8\,s$ and resonance frequency ω_o around 1 Hz, we obtain

$$\omega^4 \, x(\omega)^2 = \frac{16\pi^2}{3} G^2 \rho_e^2 |\Delta\mathbf{X}(\omega)|^2 \,. \tag{2.9}$$

This leads to the following expression for the Newtonian noise transfer function

$$\frac{\tilde{x}(f)}{\tilde{X}(f)} = \frac{4\pi\rho_e G}{(2\pi f)^2}\beta(f) \,, \tag{2.10}$$

where $\beta(f) = 1/\sqrt{3} =: \beta_{\text{Saulson}}$, is called the reduced transfer function. While β_{Saulson} provides a good approximation to $\beta(f)$, its more precise estimate involves accounting for the Rayleigh and Love modes found in the seismic spectrum. Detailed analysis about the range of values for β is given in Hughes and Thorne (1998b), where they show that β in general can vary between 0.15 – 1.4.

The NN transfer function can be rewritten as $T(f) \equiv \frac{\tilde{x}(f)}{\tilde{W}(f)}$ where $\tilde{W}(f)$ is the rms displacement averaged over the horizontal and vertical directions, given by

$$W(f) = \sqrt{\frac{2\tilde{X}^2(f) + \tilde{Z}^2(f)}{3}}\,. \tag{2.11}$$

Here $\tilde{X}(f)$ and $\tilde{Z}(f)$ in units of $m/\sqrt{Hz}$ respectively are the square root of the spectral densities of earth's horizontal and vertical surface displacements. We model $\tilde{W}(f)$ using the standard LIGO seismic spectrum,

$$\tilde{W}(f) = \begin{cases} 1\times 10^{-9}\frac{m}{\sqrt{Hz}} & \text{at } 1 < f < 10\text{ Hz} \\ 1\times 10^{-9}\frac{m}{\sqrt{Hz}}\left(\frac{10Hz}{f}\right)^2 & \text{at } f \geq 10\text{ Hz}\,. \end{cases}$$

Fig. 2.2 shows the seismic NN strain noise plotted assuming $\rho_e = 1.8g/cm^3$, $0.15 < \beta < 1.4$ and a typical direction averaged seismic displacement model spectra observed at the LIGO sites.

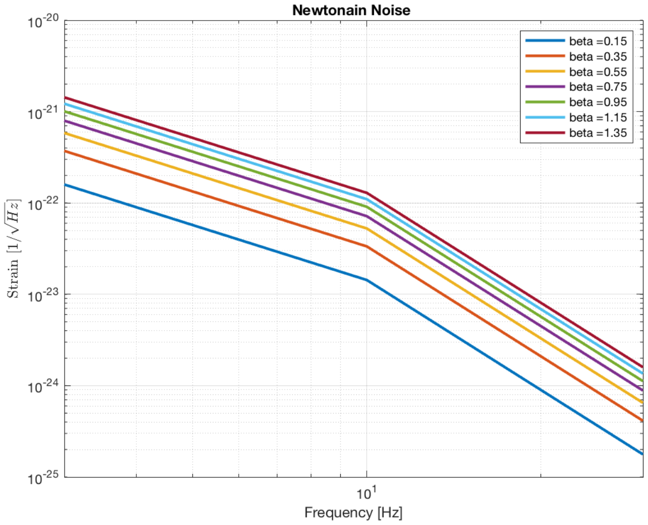


FIGURE 2.2: Newtonian noise modeled using Eq. 2.10

In the case of weak ground displacement, $\overrightarrow{\xi}(\overrightarrow{r}, t)$, the test-mass acceleration due to NN (for arbitrary seismic fields) can be estimated by the integral over the entire ground medium resulting in

$$\delta\overrightarrow{a}_{NN}(\overrightarrow{r_0}, t) = G\int dV \frac{\rho_0(\overrightarrow{r})}{\|\overrightarrow{r} - \overrightarrow{r_0}\|^3}\left(\overrightarrow{\xi}(\overrightarrow{r}, t) - 3(\overrightarrow{e}_r \cdot \overrightarrow{\xi}(\overrightarrow{r}, t))\overrightarrow{e}_r\right)\,. \tag{2.12}$$

Here, ρ_0 is the density of the ground, G is the gravitational constant, r_0 is the position of the test mass, r is the location in the ground, and $\overrightarrow{e}_r$ is the unit vector pointing from r_0 to r. We expect surface waves to be the dominant contribution to NN, due to their larger amplitudes than body waves. We test this expectation with seismic array measurements at LHO.

In the presence of stationary noise, the optimal array solution for NN subtraction turns out to be the Wiener filter and for was first explored by Beker et al. (2011) and then extended by Driggers, Harms, and Adhikari (2012). It was shown that the average subtraction residual, as a function of the correlation between sensors and the test mass as a function of seismometer locations, can be written as

$$R = 1 - \frac{\overrightarrow{C}^{T}_{SN} \cdot (C_{SS})^{-1} \cdot \overrightarrow{C}_{SN}}{C_{NN}} \ . \tag{2.13}$$

Here, C_{SN} is the cross-correlation vector between seismometers and the NN acceleration of the test mass, C_{SS} is the cross-correlation matrix between seismometers, and C_{NN} is the NN variance. Given a fixed number of seismometers, the optimal array is found by changing seismometer locations and minimizing $\sqrt{R}$. R represents the subtraction performance of any linear filter.

If we consider the signal (X_I) from the $\mathcal{I}th$ seismic sensor to be a linear combination of the displacement signal (s_I) plus some sensor noise (σ_I) and the GW sensitive channel (Y) to be made of the NN signal (N) plus some something which is uncorrelated with the seismic signal then we have

$$\begin{aligned} X_I &= s_I + \sigma_I \ , \\ Y &= H + N \ . \end{aligned} \tag{2.14}$$

We would then be required to generate a pseudo-witness channel, Y_s, by linearly combining each sensor such that its sum with the GW channel should have reduced power at each frequency bin as compared to the original. The total sum is given by

$$Y_s(\omega) = Y_\omega + \int d\omega' \sum_I \alpha_I(\omega, \omega') X_I(\omega') \ , \tag{2.15}$$

where $X_I(\omega)$, Y_ω are respectively the Fourier transforms of $X_I(t)$ and $Y(t)$. Minimization of the power spectrum $S(Y_s, Y_s) =< Y_s(\omega)^* Y_s(\omega') >$ with respect to the functions $\alpha_I(\omega)$ leads to the linear filter solution

$$\alpha_I(\omega', \omega) = -\sum_K \int d\omega'' \mathcal{K}^{-1}_{IK}(\omega, \omega'') < X_K(\omega'')^* Y(\omega') > \ , \tag{2.16}$$

where the kernel $\mathcal{K}_{IK}(\omega, \omega') \equiv < X_K(\omega^* X_J(\omega')) >$.

If seismic noise is stationary, then we have,

$$\begin{aligned} < X_I(\omega)^* Y(\omega^{'}) > &= 2\pi\delta(\omega-\omega^{'})C_{SN_I}(\omega)\,, \\ < X_I(\omega)^* X_J(\omega^{'}) > &= 2\pi\delta(\omega-\omega^{'})[C_{SS_{IJ}}(\omega)+C_{\Sigma\Sigma_{IJ}}(\omega)]\,, \\ < Y(\omega)^* Y(\omega^{'}) > &= 2\pi\delta(\omega-\omega^{'})[C_{NN}(\omega)+C_{HH}(\omega)]\,. \end{aligned} \tag{2.17}$$

This leads to the following optimal Wiener filter coefficients

$$\alpha_I(\omega,\omega^{'}) = -\delta(\omega-\omega')[C_{SS}(\omega)+C_{\Sigma\Sigma}(\omega)]^{-1}_{IJ}[C_{SN}(\omega)]_J\,. \tag{2.18}$$

We now explore the variables that contribute to equation 2.13. In its most generic version, the terms C_{NN} and C_{SN} can be given by

$$\begin{aligned} &C_{NN}(\delta a_x;\overrightarrow{\rho}_{\rm tm},\omega) = \\ &(2\pi G\rho_0\gamma(\nu))^2\int\int\frac{d^2k}{(2\pi)^2}\frac{d^2k'}{(2\pi)^2}S(\xi_z;\overrightarrow{k}_\rho,\overrightarrow{k}'_\rho,\omega)\frac{k_x}{k_\rho}\frac{k'_x}{k'_\rho}e^{-hk_\rho}e^{-hk'_\rho}e^{i\overrightarrow{\rho}_{\rm tm}\cdot(\overrightarrow{k}_\rho-\overrightarrow{k}'_\rho)}\,, \end{aligned} \tag{2.19}$$

and

$$\begin{aligned} &C_{SN}(\delta a_x;\overrightarrow{\rho}_{\rm tm},\overrightarrow{\rho},\omega) = \\ &-2\pi i G\rho_0\gamma(\nu)\int\int\frac{d^2k}{(2\pi)^2}\frac{d^2k'}{(2\pi)^2}S(\xi_z;\overrightarrow{k}_\rho,\overrightarrow{k}'_\rho,\omega)\frac{k_x}{k_\rho}e^{-hk_\rho}e^{i(\overrightarrow{\rho}_{\rm tm}\cdot\overrightarrow{k}_\rho-\overrightarrow{\rho}\cdot\overrightarrow{k}'_\rho)}\,. \end{aligned} \tag{2.20}$$

Solving these integrals, we obtain

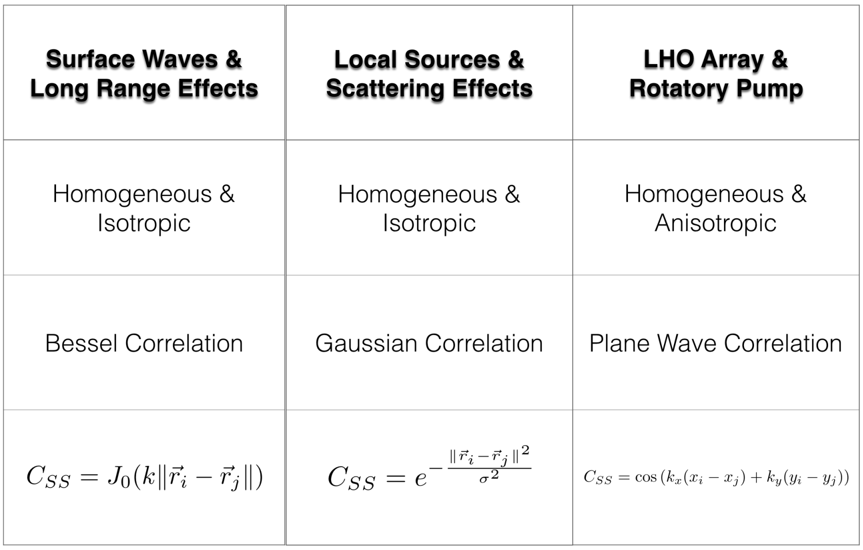
$$\begin{aligned} C_{NN}(\delta a_x;\overrightarrow{\rho}_{\rm tm},\omega) &= (2\pi G\rho_0\gamma(\nu))^2\int\int d^2\rho d^2\rho' C(\xi_z;\overrightarrow{\rho},\overrightarrow{\rho}',\omega)\times \\ &\frac{(x-x_{\rm tm})(x'-x_{\rm tm})}{(h^2+(x-x_{\rm tm})^2+(y-y_{\rm tm})^2)^{3/2}(h^2+(x'-x_{\rm tm})^2+(y'-y_{\rm tm})^2)^{3/2}}\,. \end{aligned} \tag{2.21}$$

and

$$\begin{aligned} &C_{SN}(\delta a_x;\overrightarrow{\rho}_{\rm tm},\overrightarrow{\rho},\omega) = \\ &G\rho_0\gamma(\nu)\int\int d^2\rho' C(\xi_z;\overrightarrow{\rho}_{\rm tm},\overrightarrow{\rho}',\omega)\frac{x-x'}{(h^2+(x-x')^2+(y-y')^2)^{3/2}}\,. \end{aligned} \tag{2.22}$$

We solve these integrals for a few cases, including the isotropic case, where C_{SS} takes the form of a Bessel function, a case where C_{SS} is a Gaussian, a case where C_{SS} is a plane-wave model, and a case where C_{SS} is a double harmonic

model. In the Rayleigh wave case, $C_{SS} = J_0(2\pi\|\overrightarrow{r}_i - \overrightarrow{r}_j\|/\lambda)$. In the Gaussian case, $C_{SS} = e^{-\frac{\|\overrightarrow{r}_i - \overrightarrow{r}_j\|^2}{\sigma^2}}$. In the Plane-wave case, $C_{SS} = \cos\left(2\pi(\frac{x_i - x_j}{a} + \frac{y_i - y_j}{b})\right)$. In the double harmonic case, where we allow for a wavenumber orthogonal to the dominant wave-vector,
$C_{SS} = \cos\left(2\pi(\frac{x_i - x_j}{a} + \frac{y_i - y_j}{b})\right) \cos\left(\frac{2\pi}{c}(-\frac{x_i - x_j}{a} + \frac{y_i - y_j}{b})\right)$. It is an isotropic, homogeneous correlation, but it is not the correlation coming from an isotropic, homogeneous Rayleigh-wave field. The Gaussian correlation function implies the presence of local seismic sources, which can also include scattering. For the homogeneous, isotropic Gaussian C_{SS}, the distribution of local sources must be homogeneous and isotropic as well. In the plane-wave case, as we will see, the measured correlations are best approximated by a plane-wave, and therefore we will use this case to understand and compare to the measured correlations. Table 2.1 shows the various scenarios that are encountered, the associated seismic field pattern and the corresponding the correlation models.



Surface Waves & Long Range Effects	**Local Sources & Scattering Effects**	**LHO Array & Rotatory Pump**
Homogeneous & Isotropic	Homogeneous & Isotropic	Homogeneous & Anisotropic
Bessel Correlation	Gaussian Correlation	Plane Wave Correlation
$C_{SS} = J_0(k\|\vec{r}_i - \vec{r}_j\|)$	$C_{SS} = e^{-\frac{\|\vec{r}_i - \vec{r}_j\|^2}{\sigma^2}}$	$C_{SS} = \cos\left(k_x(x_i - x_j) + k_y(y_i - y_j)\right)$

TABLE 2.1: Various possible scenarios of seismic field correlations.

2.3 Array Optimization

The primary task of this work is to find the locations of the seismometer array ($\overrightarrow{r}_i$) via minimization of Eq. 2.13 for a given seismic field correlation model (Table 2.1). Finding analytic solutions is seen to be difficult if we use more than six sensors. Also, the nature of the correlation functions as described above leads to a cost function that has numerous local minima. Numerical solutions using algorithms designed for high-dimensional problems could be the best strategy. What we require is a robust yet scalable global optimizer than could generate the desired optimal arrays in the least amount of time.

In the following, we use three numerical optimization algorithms. These work by optimizing all seismometer locations simultaneously. We expect the optimal array to be symmetric in the isotropic case, which we discuss below. Our previous attempts using Nested Sampling and Particle Swarm Optimization (PSO) did not quite yield configurations that were close to global minima and was not computationally efficient even for the six sensor case. In this work, we explore the Simulated Annealing (SA) algorithm, widely used in combinatorial optimization problems and constructed a version of it that could hit optimal solution with minimal time. SA is a probabilistic method (see Fig. 2.3) for mining a cost function consisting of numerous local minima whose name draws inspiration from a metallurgic annealing process where heating followed by slow cooling results in achieving minimum energy configuration in a crystal lattice. The algorithm starts as a random search in the parameter space and gradually moves towards a greedy search. As the number of iterations increases, it evokes a time-dependent jump proposal that is determined by the temperature of the chains. In our implementation, we perform parallel analyses at each iteration on different computing cores and selects the best configuration for further use in further iterations as shown in Fig. 2.4.

For optimizing the array locations, the algorithm is as follows. We use either the analytic or measured seismometer correlations, denoted as C_{SS}. Arrays with N seismometers are then generated by the algorithms. The C_{SS} (and thereby the C_{SN}) is then computed using the expressions given above for the various cases. This process is repeated until the algorithm's stopping criterion is reached, and after that the array with the smallest residual is taken to be the optimal array.

We make some assumptions in the following analysis. We assume that the seismic noise is stationary. We also assume that a signal-to-noise ratio of 100 is achievable with the seismic sensors. We choose for the test mass to be suspended 1.5 m above ground, which is approximately the height of the LIGO test masses. In our simulation, we use a 100 m x 100 m surface area with a test mass at its center. This is conservative, as this is larger than the area from which interesting NN contributions are expected (Harms et al., 2009).

2.4 LHO Array

In Fig. 2.6, we use the measured seismometer correlations carried out using the seismic array deployed at Hanford (2012) to inform the speed and wavelength

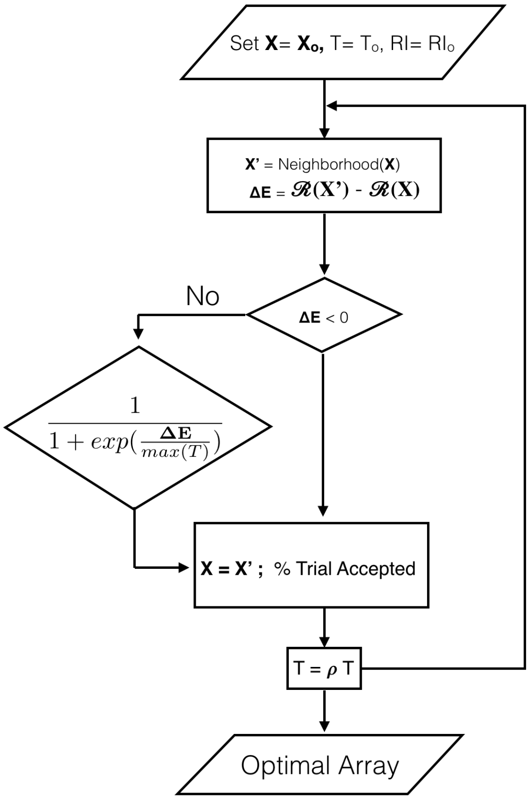


FIGURE 2.3: Stimulated Annealing algorithm

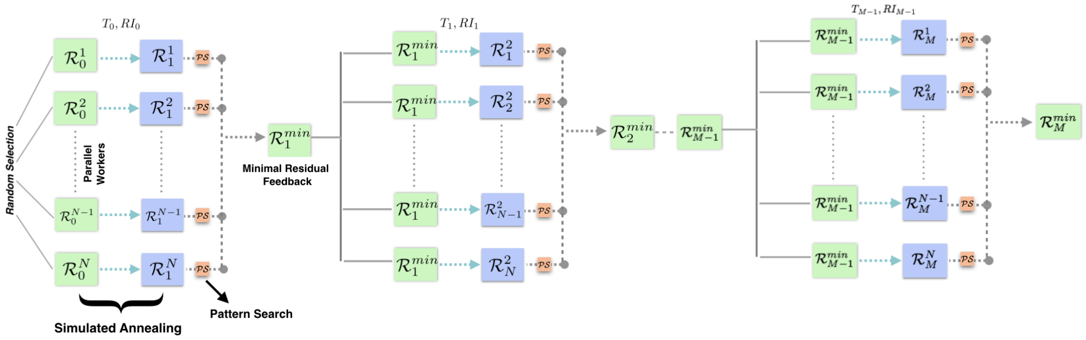


FIGURE 2.4: Parallel Cascaded Stimulated Annealing

of the seismic waves observed at the cite. A correlation analysis of the array data shows that the speed of sound for seismic waves in the ground is approximately 300 m/s, shown in Fig. 2.5. We find that at 15 Hz, the wavelength of the seismic waves is $\lambda = 20\,\mathrm{m}$. Using the relation, $\sigma = \lambda/\pi$, $\sigma = 6.5\,\mathrm{m}$. Bessel and Gaussian correlation models used in our analysis make use of these parameters to figure out the optimal array. As for the plane-wave model, the best fit parameters of $a = -42.3\,\mathrm{m}$, $b = 25.0\,\mathrm{m}$, and $c = 1.6$.

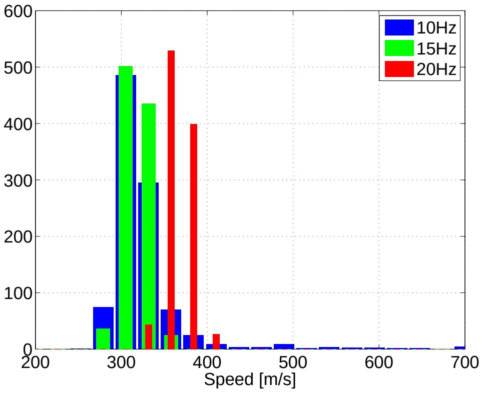


FIGURE 2.5: Histograms of seismic speed measurements using LHO array data at 10, 15, and 20 Hz.

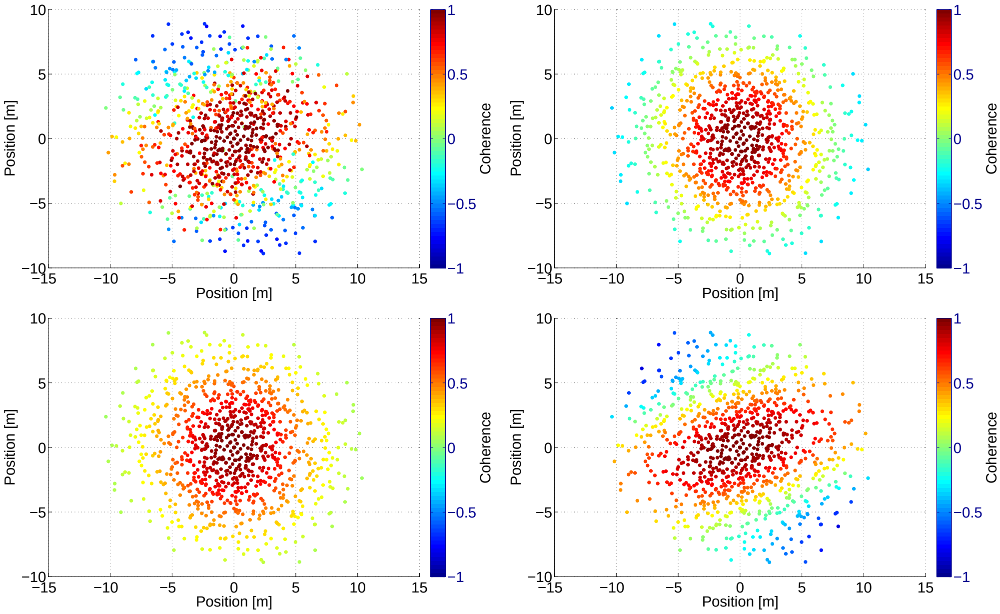


FIGURE 2.6: On the top left is the measured correlation function $C(\xi_z;\rho',\omega)$ for the LIGO Hanford site at 15 Hz. On the top right, bottom left, and bottom right respectively is the Bessel, Gaussian, and plane-wave models for the seismic fields.

2.5 Optimization Results

We start with thee simplest case of a single test mass affected by homogeneous and isotropic fields. The subtraction results using eight sensors are shown in Fig. 2.7. This is the ideal situation as we rule out the presence of isolated sources like a rotatory pump that can lead to directional dependence. We calculated the array solutions when there is slight anisotropy in the Rayleigh fields but found that the optimal solution looks almost identical to the Bessel function correlation.

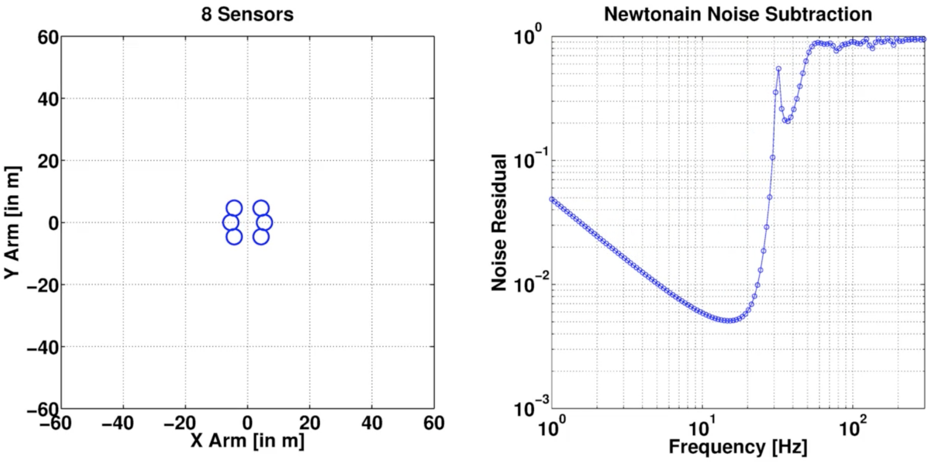


FIGURE 2.7: Optimized array geometry for homogeneous and isotropic fields. Subtraction of over a factor of 600 obtained at the target frequency of 15 Hz using 8 sensors (only six sensors are visible since two of the sensors are stacked on top of other sensors. This indicates the need for sensors with higher sensitivity).

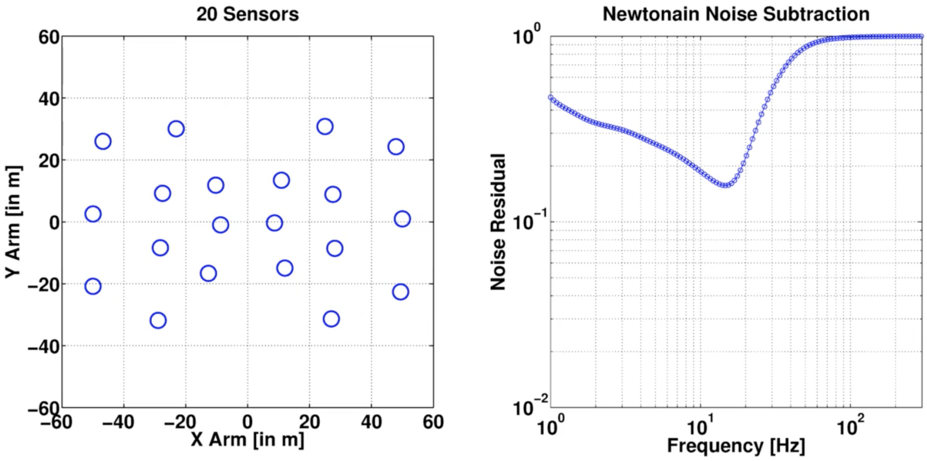


FIGURE 2.8: Optimized array geometry when there is significant scattering. Subtraction close to a factor of ten obtained at the target frequency of 15 Hz using 20 sensors.

Subtraction results for a configuration of 20 seismometers in the homogeneous and isotropic field case are presented in Fig. 2.9; As discussed above, this analysis is performed for a $\lambda = 20\,\mathrm{m}$, corresponding to seismic speeds of $c = 300\,\mathrm{m/s}$. In Fig. 2.11 we present the results after the minimization of Eq. 2.13 both analytically and using simulated annealing. For an array of 20 seismometers, the best achievable subtraction at 15 Hz is 265. The key feature of these arrays is that they are symmetric. In the isotropic case, the optimal configuration is simply a circle of seismometers at a distance from the test mass

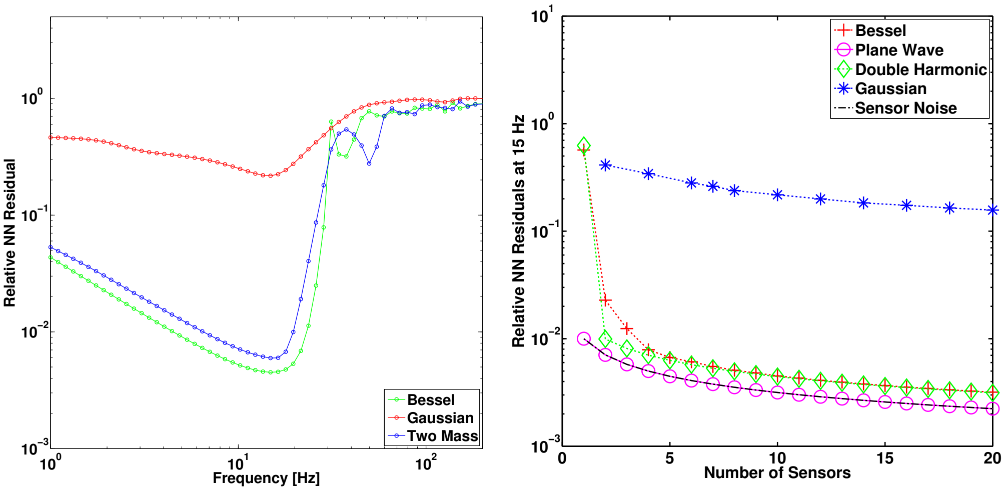


FIGURE 2.9: Subtraction residual vs. frequency for the optimal array for the Gaussian and Bessel cases with 20 seismometers. On the right is the noise reduction at 15 Hz as a function of the number of sensors for the Gaussian and Bessel cases. A reduction of approximately a factor of 5 is achieved when using 20 sensors for Gaussian, 265 for Bessel, 315 for the double plane-wave model, 432 for the plane-wave, and 447 for the theoretical SNR $\times \sqrt{N}$ subtraction.

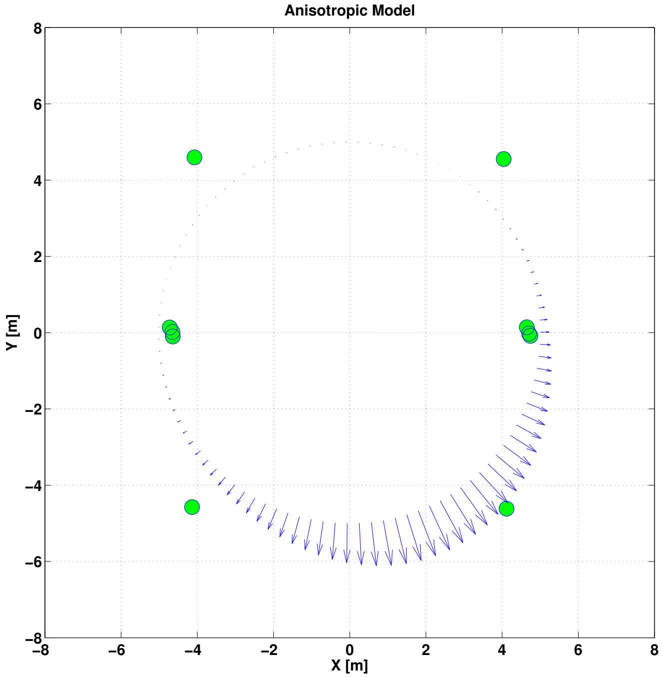


FIGURE 2.10: Effect of anisotropy Rayleigh fields on array optimization (10 sensors). Results are very similar to the isotropic case.

which is proportional to the wavelength of interest. The shape of the subtraction residuals can be understood as follows. In the seismic case, at low frequency, which corresponds to a longer wavelength, the seismometers are effectively closer together. As frequency decreases, the seismic signals become more or less identical. In this case, perfect cancellation, up to the sensor limit, can be expected for all wavelengths that are much larger than the distance between the seismometers. On the other hand, for NN cancellation, a seismometer near the test mass has a vanishing correlation with the NN (for a Rayleigh field). As frequency decreases, which effectively moves the seismometer closer to the test mass, the correlation with the test mass is lost for that sensor, which increases residuals. At high frequency, the correlation goes to 0, and therefore the residuals increase. For this reason, we need higher density (shorter distances between seismometers) for seismic cancellation than for NN cancellation. We find three algorithms arrive at the same result, presented in figure 2.11.

The second result presented is for a configuration of seismometers in the Gaussian case. This analysis is performed for a $\sigma = 6.4$, corresponding to seismic speeds of $c = 300\,\mathrm{m/s}$. For an array of 20 seismometers, the best achievable subtraction at 15 Hz is 4.5. As explained above, this analysis is useful as a direct comparison with the measured seismic results from the seismic array. These seismic arrays are also symmetric, but the array is not at a constant radial distance for this situation. The subtraction residuals have a slightly different shape and magnitude than in the isotropic case. In general, the residuals are much higher than that for the Rayleigh field case. This is due to the significant complication from local sources or seismic scattering. They also flatten at low frequency and rise more slowly at high frequency.

The third result we present is for the most optimistic seismic field, that of a plane-wave, with a theoretical minimum of SNR $\times \sqrt{N}$ subtraction (1 - $\frac{1}{1+SNR^2}$). In this case, the subtraction is a factor of 432, which is about 10% higher than the sensor noise limit. The optimal arrays for this scenario are not symmetric, and there is no single optimal array, but instead, only optimal lines on which the seismometers lie.

The fourth and final result presented is for the measured correlations given by the modified plane-wave approximation for the seismic array. The analysis is performed in the same way as above, but using fits to the measured correlations between the LHO array seismometers. For an array of 20 seismometers, the best achievable subtraction at 15 Hz is a factor of 316. The best array is not identical to, but in general, has similar features to that of the array above. Deviations from the plane-wave approximation are to be expected due to the anisotropies of the seismic field located at the detectors. The short-coming to this analysis is that the array coverage, which is approximately 10 m, is significantly smaller than is required to see the correlations go to 0 at a far distance. A future dedicated array at LHO will significantly expand the spatial coverage, and therefore this should be improved.

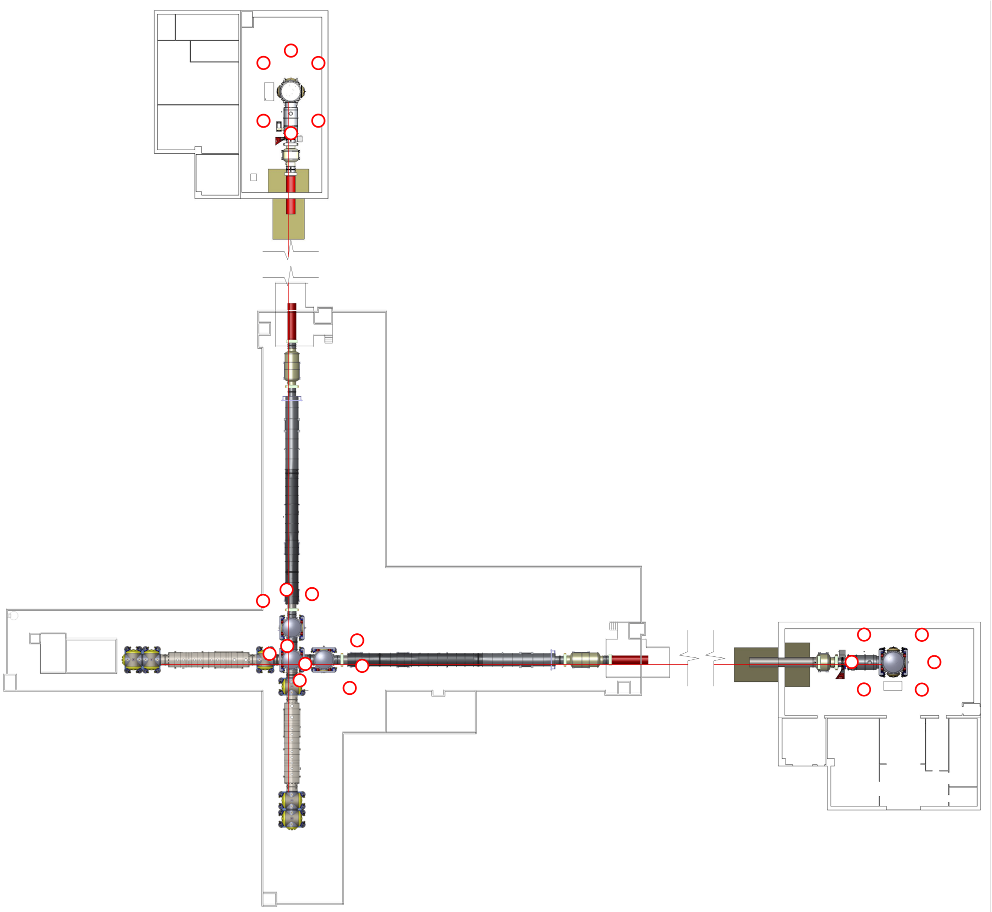


FIGURE 2.11: An example arrangement of seismometers at both end stations and the corner stations. The end station placement is based on the Bessel model discussed in the text surrounding the test mass. The corner station arrangement is based on the Bessel model surrounding both of the corner station test masses.

We present an application of these results in Fig. 2.11. In this analysis, we adopt a Bessel model for both the end stations and the corner station at LHO. The end station analysis is identical to that of the above. The corner station is slightly more complicated, due to the presence of multiple test masses. In this case, we optimize the array for NN subtraction at both of the test masses at the corner station. This changes the shape of the optimal array for the corner station. But it clearly shows that we only need to monitor the seismic fields with the center station to achieve the desired levels of subtraction for homogeneous yet not perfectly isotropic fields which is the case for Advanced LIGO detectors.

2.6 Conclusion

In this work, we have investigated optimal seismic arrays for performing NN subtraction at the LIGO Hanford site. We have made a series of assumptions about the seismic field, from isotropic and homogeneous to non-isotropic and inhomogeneous, and computed the optimal seismic arrays for these situations. Measurements from a seismic array installed at the LIGO Hanford site allowed us to move beyond assumptions of an isotropic and homogeneous field. We show that we can use such an array to subtract NN noise from the gravitational-wave detectors should they become limited by this source.

Although subtraction residuals do not depend significantly on array configurations and instead on the number of seismometers and the ground covered by the array, we benefit greatly from a cross-check between the array derived by the analytic solution in the isotropic and homogeneous case and the numerical solution derived in the non-isotropic and inhomogeneous case. Our numerical solutions benefit from measured correlations, which include contributions from not only surface waves, but body waves and scattered waves as well. This should allow for NN subtraction by factors of 10 or more.

The accuracy and scalability of the developed techniques would enable us to deploy the appropriate array that could be used for post-offline NN subtraction thus lowering the noise floor for second-generation gravitational-wave detectors. Further improvements in these algorithms can also make them beneficial for the third generation detectors. In future, we would like to explore the effectiveness of such optimized arrays in the presence of varying levels of non-stationarity noise.

Chapter 3

Earthquake Warning Based on Ground Motion Prediction

3.1 Introduction

With the advent of gravitational wave (GW) astronomy, it is essential to maximize the duty cycle of second-generation gravitational-wave detectors such as the Laser Interferometer Gravitational-wave Observatory (LIGO) (J Aasi et al., 2015), Virgo (F Acernese et al, 2015), and GEO600 (Grote H. for the LIGO Scientific Collaboration, 2010) detectors. Any increase in duty cycle increases the sensitivity of GW searches, including the observations of binary black hole mergers (Abbott, B. P. et al, 2016; Abbott, B. P. et al., 2016; Abbott, B. P. et al., 2017a; Abbott, B. P. et al., 2017b) and binary neutron stars (Abbott, B. P. et al., 2017c). GWs from these induce small displacements in the detectors, which are designed to be free from environmental disturbances and limited only by processes of fundamental physics. These detectors are subject to non-Gaussian noise transients due to either internal behavior of the instrument or interactions between the detector and its environment (The LIGO Scientific Collaboration and Virgo Collaboration, 2016). To minimize the effect of the environment, the LIGO detectors contain 200,000 auxiliary channels which are designed to monitor both the behavior of the instrument and the environmental conditions. A subset of these sensors is physical environmental monitor sensors dedicated to monitoring the environment, including seismometers, magnetometers, microphones, and many others. The Advanced LIGO (J Aasi et al., 2015) and the Advanced Virgo (F Acernese et al, 2015) have in particular driven the development of both seismic (Beker et al., 2016) and rotation (Venkateswara et al., 2014) sensors. Among these, seismometers are useful for measuring any source of ground motion that can couple into the interferometers. One source of ground motion that destabilizes the detectors are earthquakes (Coughlin et al., 2015; Coughlin et al., 2017b), despite seismic isolation systems designed to minimize such effects (Abbott and Adhikari, 2002; Stochino et al., 2009; Matichard et al., 2015). The surface waves, typically the highest amplitude component from earthquakes in the frequency band of interest with the longest duration, adversely affect the detectors. This occurs by making it impossible to keep the detectors operable or induce higher frequency noise by up-converting low-frequency optical motion.

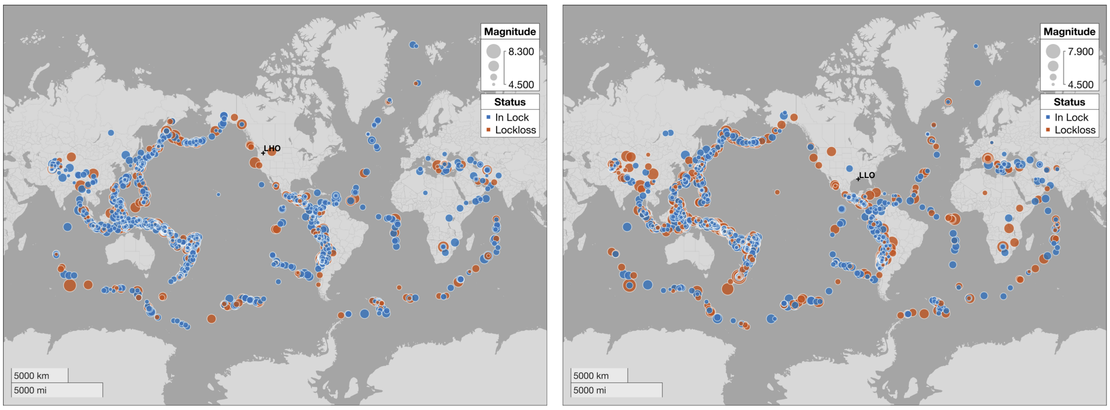


FIGURE 3.1: Impact of earthquakes happening worldwide on LIGO Interferometers at Hanford and Livingston during the first and second observation period. Points marked in red indicate the instances when the resulting ground motion caused the interferometer to go out of lock.

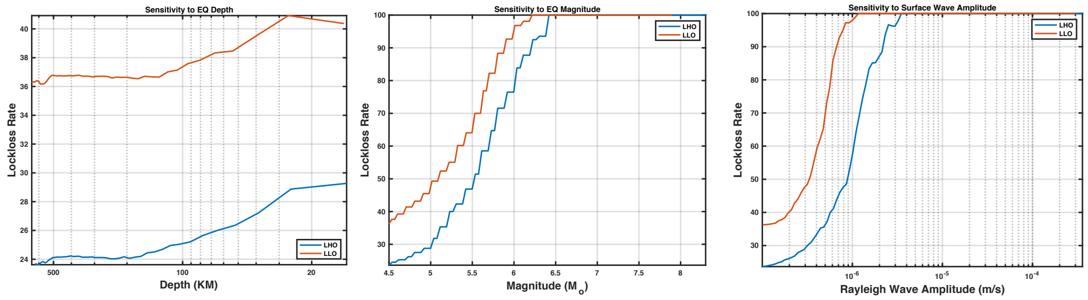


FIGURE 3.2: Plots showing lockloss rate associated with earthquake magnitude and Rayleigh wave amplitude at both the LIGO detectors.

Earthquake early warning (EEW) is a burgeoning field dedicated to the rapid detection and characterization of earthquakes as well as the dissemination of that information to people and infrastructure in their path (RM. Allen, 2012; Kuyuk, H.S. and RM. Allen, 2013; Strauss and Allen, 2016). Worldwide, many seismic and geodetic (GPS) sensor arrays exist that produce rapid earthquake information products, from magnitude and location estimates to regional centroid moment tensors (CMTs) and advanced slip inversions. With wide-ranging public warning systems in Mexico and Japan and smaller-scale systems in many other countries, warnings from seconds to minutes are now available to reduce the impact of earthquakes on society (Strauss and Allen, 2016). The short warning times arise out of the physical processes that drive the earthquake rupture, where the warning is given by seismometers measuring P-waves ($\approx 8\,\mathrm{km/s}$) and S-waves ($\approx 4\,\mathrm{km/s}$). Reliability of these estimates are one of the most important aspects of EEW systems, and their improvements generally rely on increasing the number of stations involved in the warning decisions as well as increasing alarm thresholds on ground motion, both seeking to limit the number of false positives (Kuyuk et al., 2015). Both of these strategies come at the cost of decreasing the warning time. As these systems minimize the time required to calculate the source parameters of earthquakes (i.e., their location and magnitude), it becomes important to predict the ground motion that the earthquakes will cause as a function of location and distance with high accuracy.

This chapter is organized as follows: Section 3.2 talks about the sensitivity of the GW interferometers to earthquakes and the previous attempts to model them. In Section 3.3 we describe how the seismic data was obtained from the site as well as the IRIS seismic array. Section 3.4 describes the deployed regression and clustering techniques and their requirements. Finally, the relative performance of various prediction algorithms along with the ability to guess the state of the interferometer are covered in Section 3.5.

3.2 Impact Assessment

Fig. 3.1 depicts the distribution of global seismic events and their respective effect on the state of the GW interferometers during LIGO’s first and second observation run. The blue circles (scaled as per the magnitude) represent scenarios where the ground motion was high enough to cause instabilities leading to a state of lockloss. Sensitivities to parameters such as magnitude and surface wave amplitude are shown in Fig. 3.2 where the steepness of the curve indicates higher sensitivity to the respective parameter. As expected LLO is seen to be more vulnerable to ground shaking which could be attributed to its local geology and soil properties. The main goal of LIGO/Virgo EEW methods would be to generate reliable relations between earthquake source parameters and ground motion metrics. Examples in the time domain include peak ground acceleration, peak ground velocity, and peak ground displacement, while in the frequency domain there are spectral accelerations, velocities, and displacements as well as predominant periods (Douglas, 2003). Early estimates of magnitudes tend to underestimate the energy released due to the non-instantaneous pattern of slip. For this reason, the early estimates of the ground velocity amplitudes are often not as accurate as the later values. The effects of these errors are particularly pronounced for larger earthquakes, where the estimates of the fault lengths become more important. Thus, these larger earthquakes tend to have their amplitudes under-predicted. The loss of performance that results from use of the rapid estimates is acceptable to be used as rapid warnings. Melgar et al. (2015) showed that real-time GPS waveforms could rapidly determine the magnitude within the first minute of rupture and in many cases before rupture is complete.

In the initial stage, we used advances in early earthquake warning to develop a low-latency earthquake early warning client named *Seismon* (Coughlin et al., 2017b). This system uses a real-time event messaging system of the U.S. Geological Survey (USGS) to mitigate the effects of teleseismic events on ground-based gravitational-wave detectors. Using information about the earthquake source characteristics such as magnitude, depth, and distance, ground velocity induced by the earthquakes were predicted. In the initial version of the algorithm, we used an empirical fit to an equation derived to account for physical effects. As per this equation, the amplitude of the surface waves, $\mathrm{Rf}_{\mathrm{amp}}$, is given by

$$\mathrm{Rf}_{\mathrm{amp}} = M\frac{\mathrm{a}}{f_c^b}\frac{e^{-2\pi h f_{\mathrm{c}}/c}}{r^d}\,, \tag{3.1}$$

where $f_{\mathrm{c}} = 10^{2.3-M/2}$ is the corner frequency of the earthquake, M is the magnitude of the earthquake,a is a scaling factor ,h is the depth, r is the distance to the detectors, and c is the speed of the surface-waves, all in SI units. A Metropolis Hastings Multi-Chain, Monte Carlo algorithm implementing adaptive simulated annealing, which statistically guarantees to obtain solutions close to the global minima (Kirkpatrick, Gelatt, and Vecchi, 1983; Ingber, 2000) was used to optimize the parameters a, b, c and d. Optimization was carried out by minimizing the difference between the prediction $\mathrm{Rf}_{\mathrm{amp}}$ and the set of historical data samples of ground motion measurements made at the sites. This algorithm is the same as the one used in the optimization of seismometer arrays for gravity gradient noise cancellation in gravitational-wave detectors, and a thorough explanation can be found in Chapter 2. The best-fit parameters so obtained for each of the observatories is given in Table. 3.1.

This equation succeeded in predicting peak ground velocity such that 90%

Detector	a	b	c	d
LHO	0.16	1.31	4672.83	0.83
LLO	0.16	1.31	4672.83	0.81
VIRGO	1.60	0.89	4992.70	0.83
GEO	8.65	1.92	324.52	1.40

TABLE 3.1: Best-fit parameters to the peak velocities seen at the interferometers to Eq. 3.1.

of events had a measured ground velocity within a factor of 5 of the predicted value. But there were a few downsides to this empirical fit. First of all, while it was derived with physical effects in mind, it was predominantly an empirical construction. It was also found that the parameters in the model were quite degenerate, which meant that parameters derived to be physically meaningful quantities showed significant differences from site to site which were unlikely to be very different. Finally, to be useful to the detectors, there is a goal of a factor of 2 in relative error in the ground velocity predictions to make accurate assessments of whether the gravitational-wave detectors will be affected, which is much smaller than the factor of 5 scatter seen.

One of the key aspects of the system is the ground velocity predictions, $\mathrm{Rf}_{\mathrm{amp}}$, for each site. These predictions have two purposes. First of all, they provide a meaningful metric which on-site-staff at the detectors can use to plan the response to the incoming earthquake. The response could be in the form of switching seismic isolation loops to steer the interferometer to a more robust configuration keeping it locked although with a lesser sensitivity (Biscans et al., 2018). The predictions also serve as inputs to the algorithms which make lock-loss predictions, which we will describe in the following Chapter. Any such information about upcoming downtime can be utilized to perform opportunistic maintenance to rectify problems typically scheduled for weekly maintenance periods.

3.3 Methods: Seismic data.

The first part of our analysis uses data obtained from the GW sites. For each earthquake event from the archival database, we take the vertical component of broadband (velocity) data that is filtered using an acausal 0.1 Hz low-pass Butterworth filter. The data is calibrated into ground velocity using a constant V to m/s value appropriate for each seismometer. Time-series are chosen to encompass the P-wave arrival to surface waves calculated assuming a (very conservative) seismic velocity of 2 km/s. As the seismometers located at the end

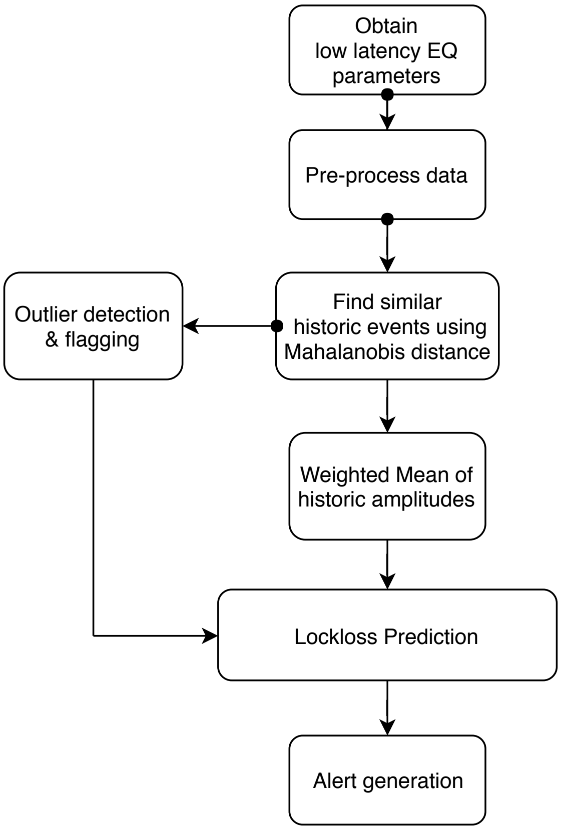


FIGURE 3.3: Process flowchart depicting the low latency earthquake warning pipeline.

and center stations observed similar values for the relevant frequency band, we use the center station sensor for the rest of our analysis.

We also perform an analysis of seismic time series that were made available through IRIS (*US Array Website*), covering the last ten years. These stations have time-series with response between 10 mHz to 10 Hz. The noise for these instruments is determined by a variety of sources including anthropogenic and atmospheric disturbances, earthquakes and ocean waves (Bonnefoy-Claudet, Cotton, and Bard, 2006). Seismic noise models are developed using global seismometer arrays (Peterson, 1993; Berger, Davis, and Ekström, 2004; McNamara et al., 2009). We use stations across 51 US states to explore the effects of a variety of different sites, which can have noise spectra that have significant variation, due to location aspects such as topography and proximity to urban settlements. One source present across the world is the oceanic microseism around 0.3 Hz that dominate seismic ground spectra everywhere on the Earth (Haubrich, Munk, and Snodgrass, 1963; Toksöz and Lacoss, 1968; Cessaro, 1994; Friedrich, Krüger, and Klinge, 1998). IRIS contains data for some stations as far back as the early 1970's, of which we systematically downloaded and processed data from almost all stations from January 2005 to May 2017 consisting of around 21412 earthquakes. Stations are supplied with Nanometrics T240, Streckeisen STS-1/STS-2, Güralp CMG-3T and Geotech KS-54000 broadband seismometers. The magnitudes range from 6.0 to 9.2, chosen to cover the range of earthquake magnitudes likely to significantly effect the gravitational-wave detectors.

3.4 Methods: MLA description

Machine learning (ML) has recently become an important aspect of EEW and seismology in general. For example, The *MyShake* EEW system uses artificial neural networks to differentiate earthquake and human motions, with 98% of earthquake records within 10 km correctly identified, and only 7% of people-induced transients is appearing to be earthquakes to the algorithm (Kong et al., 2016). They have been used to differentiate earthquakes from other seismic transients (Kuyuk et al., 2011; Kortström, Uski, and Tiira, 2016; Perol, Gharbi, and Denolle, 2017), discriminate between deep and shallow microearthquakes (Mousavi et al., 2016) and to add to undersampled or missing traces (Jia and Ma, 2017). Also, they have been used to make full-wave tomography images (Diersen et al., 2011).

The idea of this analysis is to compare historical ground velocity measurements to predictions made using different machine learning algorithm techniques. The inputs to the algorithm are the earthquake magnitude, latitude, longitude, distance, depth, and earthquake azimuth relative to the detector. The target output is the measured ground velocity. This improves on the analytical equation in a few ways. First of all, by switching to ML, we eliminate the dependence on a functional form. Second, it trivially includes more parameters, such as latitude, longitude, and earthquake azimuth relative to the detector above and beyond the initial analytical formalism.

In particular, we compare the efficiency as the two different machine learning approaches: regression and clustering. Within regression, we evaluate the performance of the Tensorflow implementation of deep neural networks (DNN) (Abadi et al., 2016), stacked ensemble regressors (Wolpert, 1992; Laan, Polley, and Hubbard, 2007) and Gaussian Process Regression (GPR) (Rasmussen and Williams, 2006), while in clustering we use a Mahalanobis distance (Mahalanobis, 1936) based similarity search to make the predictions. The parameters that enter the predictions are M, the magnitude of the earthquake, h, the depth, r, the distance to the detectors, θ and ϕ, the latitude and longitude, and α, the earthquake azimuth relative to the detector. All of these variables are available in low latency from the USGS. On longer timescales, the earthquake slip inversion, strike, rake, and dip, and the moment tensor values, M_{rt}, M_{tp}, M_{rp}, M_{tt}, M_{rr}, and M_{pp} are also available, serving as additional parameters which could be help the prediction process. The target variables correspond to peak ground velocities measured using seismometers.

The performance of each algorithm is accessed using real and simulated datasets. For each dataset, 80% is used for training with the remaining 20% used for predictions. As for simulated data, new samples are generated from each of the original datasets by creating a Gaussian jitter distribution centered around the parameter value followed by a random draw of samples from these distributions. Artificially adding noise (or jitter) to the predictor and response variables in a controlled fashion helps to improve the learning and prevent early stopping. The presence of noise enhances the ability of the MLA to better learn and generalize to the underlying smooth, non-linear function. Selective boosting is done to minimize the imbalances in the dataset using the Synthetic Minority Oversampling TEchnique (SMOTE). Prediction results are based on the unseen test data which account for 20% of the total data samples.

The deep neural network (DNN) that we employ to carry out the nonlinear regression has a topology inspired from generalized regression neural networks (Specht, 1991), but we back-propagate the errors and update the weights by training it through several epochs. DNNs, in general, require larger data sets to learn the underlying function without over-fitting the data and we observed them be sensitive to the network architecture and the activation functions. We use a sequential network with nine dense layers with exponential linear unit activation and the first-order gradient-based Adam optimizer (Kingma and Ba, 2014). Results from using this network are given in Fig. 3.4.

Recently stacked ensemble regressors have gained much prominence and are consistently outperforming others in several datasets hosted at Kaggle (*Kaggle Website*). The first level consists of a set of base learners who are individually trained and cross-validated. Their predictions form input to second level meta-learner regressor which is further trained to generate the final ensemble prediction. Such systems are theoretically guaranteed to present the optimal learners in the asymptotic sense. Success with DNN and ensemble techniques crucially depends on the number of training data and is sensitive to the hyper-parameters. As for the Gaussian Process Regression or kriging, the results are seen to be better than the above two when we use squared exponential kernel with prediction based on block coordinate descent. The GPR hyper-parameters

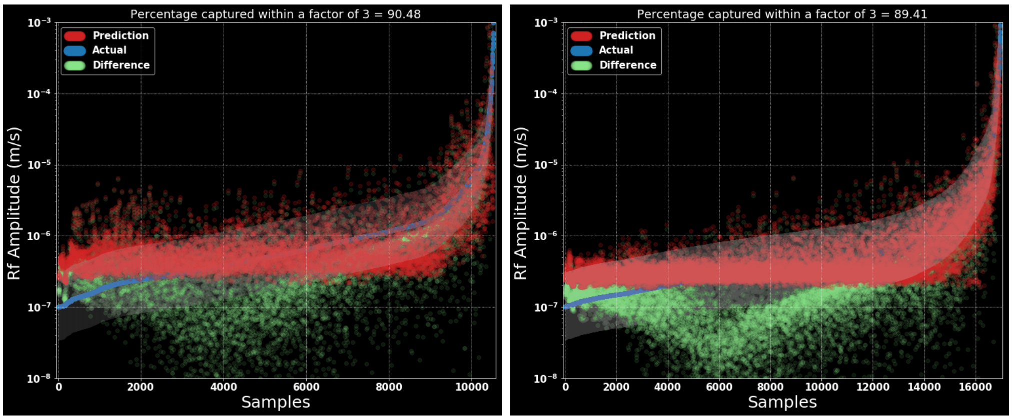


FIGURE 3.4: Performance of Deep Neural Network in predicting the peak ground motion based on low latency earthquake parameters.

are optimized using Bayesian optimization. But the method scales as O(n^3) resulting in high memory requirements and training time for large data sets.

Mahalanobis distance is the multi-dimensional generalization of z-score which tells how many sigmas is the data away from the mean distribution. It is observed to be a very robust technique as it takes into account the covariances between the variables. In Fig. 3.5 we show that results from using our clustering technique which makes use of this metric to find the closest matching earthquakes that happened in the past. This scheme naturally lets one identify outlier earthquakes with no similar events in the archival data.

Among the various MLAs, we choose clustering-based prediction for EEW pipeline, and the process flowchart is given in Fig. 3.3. In addition to having the best prediction accuracy, it has the following advantages. Firstly, as there is no training involved, the need for hyper-parameter tuning is eliminated. As we are constantly monitoring the seismic data and appending the earthquake database, with time we expect a decrease in prediction error along with a reduction in the number of outliers.

3.5 Results

Fig. 3.5 shows the prediction results from applying Mahalanobis based clustering on the simulated and real earthquake data. We obtain an accuracy of above 90% in both the cases. This performance improvement in prediction accuracy from a factor of 5 to 2.5 can be attributed to increased availability of data, the inclusion of more earthquake parameters and the usage of robust algorithms. Prediction results are similar when the algorithm is applied to earthquakes observed at the Virgo Observatory 3.6. As for the outliers, they do

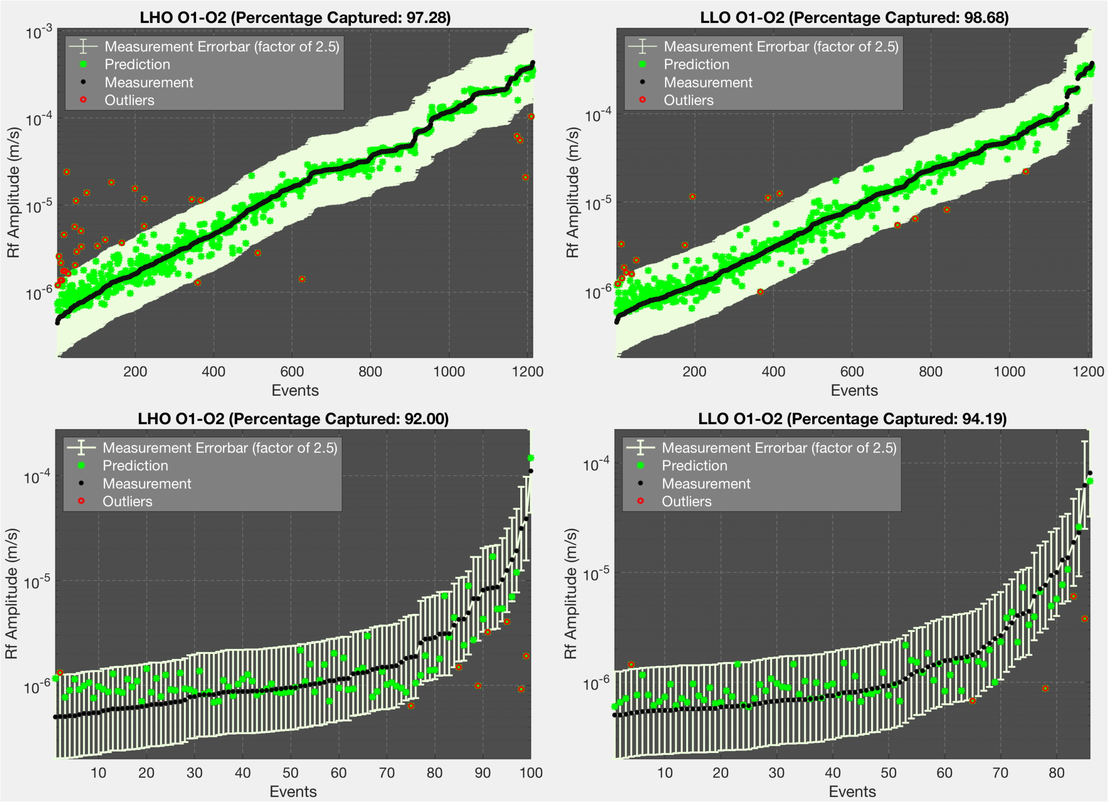


FIGURE 3.5: Estimate of peak velocities seen during O1-O2 at the interferometers (LHO, LLO) using Mahalanobis distance-based clustering. Results on simulated and real data are respectively shown in the top and bottom rows. The events have been ordered by their measured peak ground velocity (in gray), and a yellow error bar corresponds to a factor of 2.5 within the predicted value. More than 90% of the events are within a factor of 2.5 of the predicted value.

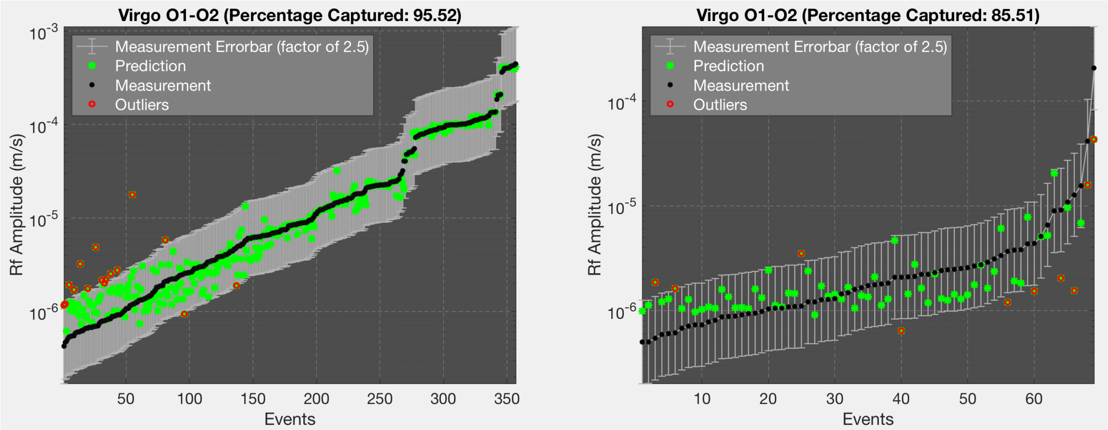


FIGURE 3.6: Fit of peak velocities seen during O1-O2 at the Virgo Interferometer

not seem to show any specific dependence concerning the input parameters. One reason could be that their parameter combination is rather uncommon, so the predicted amplitude is averaged across not so similar events. The general trend among outliers seems to be that the higher amplitude events are underestimated, and the lower amplitude events are over-estimated. Such outliers should decrease as we gather more training data. The possibility of further improvements using moment tensor solutions is left to future work.

TABLE 3.2: Rf amplitude prediction performance of different ML algorithms for simulated (upper right) and real (lower left) data

	Deep Neural Nets	Stacked Ensemble	GPR	Clustering
LIGO Livingston	89% 85%	93% 89%	94% 87%	98% 94%
LIGO Hanford	86% 84%	91% 88%	92% 89%	97% 92%

We also demonstrate the resourcefulness of the above scheme by making predictions across the United States using the data recorded by the IRIS network. For ground motion recorded in each state, we perform the similarity search using archival data from the same state as well as the other states and compare the prediction accuracy as shown in Fig. 3.7. For each row representing a state, we show how well we can make ground velocity predictions based on archival events from itself as well as other states. For at-least 24 states, the accuracy is seen to be above 90% when we use its training data for predicting the new events. The variation seen in predictability along the diagonal might be due to the differences in local geology across the US. This observation of unpredictability could be beneficial for future site selection surveys looking for suitable locations for next-generation interferometers. High values seen along several of the off-diagonal terms are due to the nearly identical response to teleseismic events could mean the corresponding similarity within their geological properties such as shear velocity profiles, elasticity, and local soil density, etc. Future state-wise seismic modeling could use this information to the data with the ones from similar states especially if the original dataset is sparse. The performance shows that the archival data-based prediction scheme can be extended beyond the realm of gravitational-wave detector sites for hazard-based early warning alerts.

3.6 Conclusions

In conclusion, we have used MLAs to predict peak ground velocities from teleseismic earthquakes. The estimated ground velocity is used to forecast the potential effect of earthquakes on gravitational-wave detectors and issue near

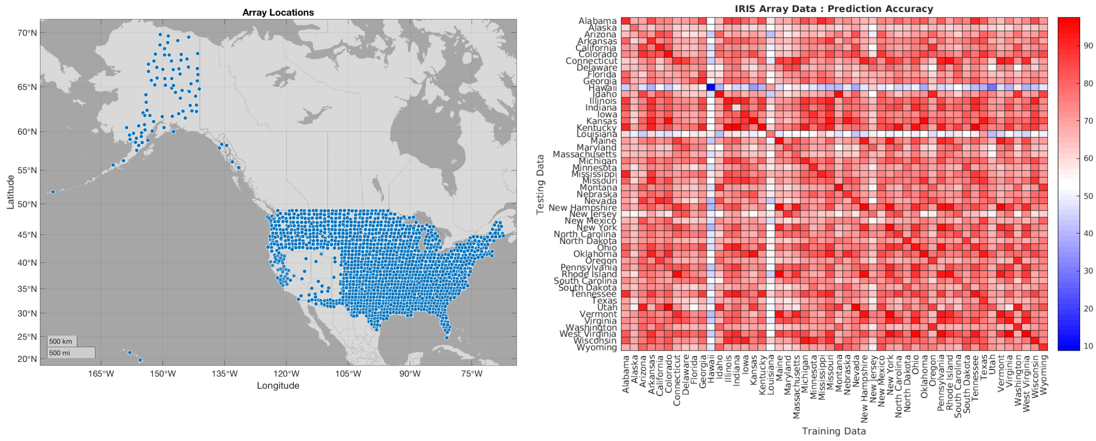


FIGURE 3.7: Figure to the left shows the locations of the seismometer array used to collect the earthquake data. Heatmap on the right shows the ground motion prediction accuracy within each state making use of archival data from the same as well as every other state.

real-time alerts at the site. Alert system based on this scheme has been implemented at the Advanced GW observatories and would be used in the near future to switch seismic filters at very low frequencies. Given the significant interest in accurate ground velocity predictions for EEW systems in general, we believe the techniques here would be beneficial beyond the gravitational-wave community. While we focus on the prediction of peak ground velocity here, in future we would explore the possibility of using historical seismometer data along with CMT parameters to predict radiation pattern associated with the fault rupture.

Chapter 4

Impact of Induced Seismicity on Advanced Gravitational Wave Interferometers

4.1 Introduction

Terrestrial gravitational wave observatories like Advanced LIGO and Advanced Virgo have reached a sensitivity necessary to detect compact binary coalescence from binary black holes and neutron stars (Abbott, B. P. et al, 2016; Abbott, B. P. et al., 2016; Abbott, B. P. et al., 2017a; Abbott, B. P. et al., 2017c; Abbott, B. P. et al., 2017b) and aims to improve it further to probe the universe at higher redshifts. This process requires to correctly understand the limiting noise sources at the lower frequency (below 30 Hz) and develop strategies to effectively tackle them. Low-frequency ground vibrations significantly affect the duty cycle and the quality of the strain data used for GW analysis. The passive and active isolation installed at Advanced LIGO has been designed to provide seismic noise suppression at frequencies above 1 Hz (Matichard et al., 2015). Typically there is no effective isolation in earthquake frequency band (0.01 Hz to 0.03 Hz) which implies that the interferometer goes out of lock whenever the ground motion from earthquakes crosses a threshold value (typically around few microns per second). They are among the major contributors to detector downtime and the resulting recovery period can vary from a few minutes to hours. Low latency early earthquake warning (EEW) warning systems described in Chapter 3 which are installed at the GW detectors focus on high magnitude teleseismic events. The alert issued based on information received from USGS seismological network has a latency of fifteen minutes and includes information about the time of arrival and the expected ground motion from the Rayleigh waves. Such alerts are used to either prepare the instrument to face the impact as well as to unlock the interferometer to minimize the effects from excessive ground motion. For the first time, we discuss the impact of induced low magnitude earthquakes($M_w < 5$) of regional origin on current and next-generation GW interferometers.

The chapter is organized as follows, section 4.2 will discuss induced seismicity and the sudden increase in regional earthquakes in the conterminous US. Section 4.3 will focus specifically on the Oklahoma region and locklosses caused by events from this region. In section 4.4, we describe the results from simulated events from the Tuscaloosa Marine Shale (TMS) region and look into the proximity effects. Finally in section 4.5, we briefly discuss how this new source of noise could be detrimental to the future generation of terrestrial detectors and possible strategies to minimize the impact.

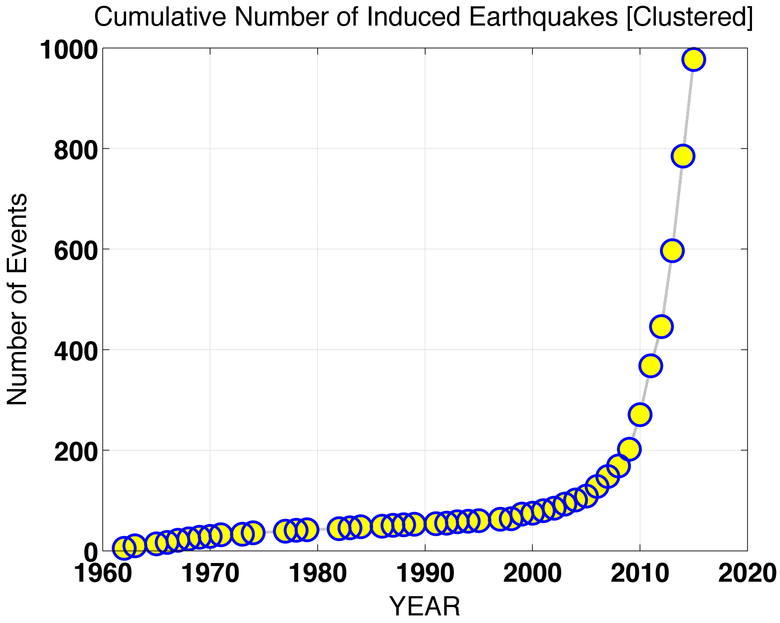


FIGURE 4.1: Cumulative distribution of induced earthquakes in conterminous US. The exponential increase since 2010 strongly correlates with the increased activities related to hydraulic fracking.

4.2 Induced Seismicity

Induced seismicity refers to low magnitude earthquake activity caused due to human activities that stresses the fault structures resulting in a rupture. Earthquakes happen when there is a slip in the fault resulting in a release of stored stress energy. Rupture is initiated when the effective shear stress exceeds the frictional resistance, and the energy gets radiated away as seismic waves. Fault rupture occurs when the shear stress between the foot-wall and the hanging wall exceeds the effective normal stress according to,

$$\tau_c = \tau_o + \mu(\sigma_n - P\ I)\,, \tag{4.1}$$

where τ_c: critical stress, τ_o: shear stress, μ: coefficient of friction, σ_n: normal stress, P: Pore pressure and I: unit tensor.

Last several years have seen an unprecedented increase in the number of induced earthquakes in the conterminous US from regions that were previously seismically quiet (See Fig. 4.1). Oklahoma region is the most striking example where the rise in low magnitude earthquakes has been alarmingly high since 2010. This phenomenon is strongly correlated with the rise in fracking activities within the state as part of the booming oil industry. Hydraulic fracking is the process of horizontal drilling followed by injection of liquid at high pressure. This process results in excess of wastewater which is hazardous in general and so needs to be disposed of by re-injecting it back to crust through class II wells. Injection fluid can directly get in contact with the fault structure resulting in an increase in pore pressure, or it could lead to a change the loading conditions on the fault leading to a tremor.

Proximity to the site implies even minor quakes from these regions could be a potential cause of lockloss at the interferometers. The proximity of the hypo-center $\approx 800\ km$ results in induced events that can hamper the day-to-day operation of the detector.

Factors that determine induced seismicity is still an active area of research (Weingarten et al., 2015) but in general, they include proximity to the number of class II injection wells, the rate of injection, wellhead injection pressure, proximity to the crystalline basin and cumulative injection volume. The triggered low magnitude earthquakes usually do not result in structural damage as the induced motion is of orders of tens of micrometers but is sufficient to disrupt the functioning of a gravitational wave interferometer.

4.3 Oklahoma Events

Oklahoma has been experiencing a series of earthquakes of magnitude above 4 in the past few years4.2, and several previous studies have associated these with wastewater injection activities in the state. Table 4.1 list the 31 earthquakes that occurred in Oklahoma region from 2013 to 2018. We describe results from analysis of these events using data from three STS broadband seismometers (ETMX, ETMY, ITMY) located respectively near both the end arms of the interferometer and at the corner station. The signal is analyzed in various frequency bands: 30 mHz - 100 mHz (earthquake band), 100 mHz - 300 mHz (micro-seismic band) and 300 mHz - 1Hz (ground motion) to get a better understanding of these shallow regional earthquakes. Results from the measurement of peak ground motion and earthquake duration are shown in figure 4.6.

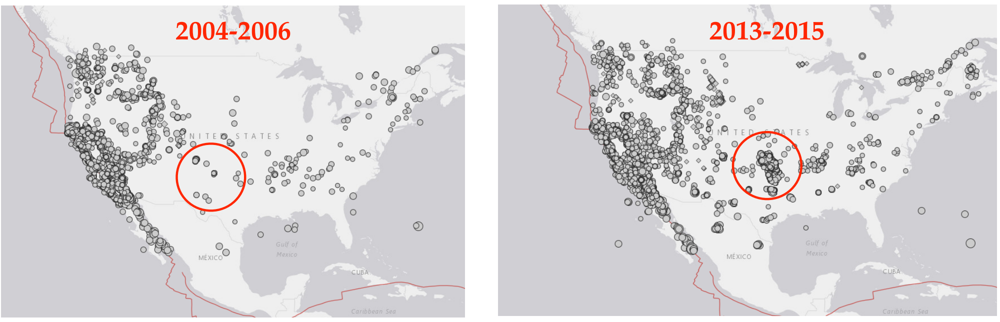


FIGURE 4.2: Plot comparing the earthquakes in Oklahoma during 2004-06 and 2013-15.

Earthquake band at the Livingston observatory sees a median ground velocity of order a micron per sec with the highest amplitude along the X direction, followed by Y and Z direction. Interestingly, the micro-seismic and ground motion band shows even higher levels of motion, making these events different from the typical teleseismic events observed at LIGO. There is significant motion along the horizontal axis which is undesirable as it could lead to misalignment of the optics potentially increasing the time to lock the cavity again after an earthquake-induced lockloss. We see a motion of micro radian for the suspension optical levers in pitch and yaw degrees of freedom. If the test mass mirrors swing with high amplitude, they are also susceptible to stray charge deposition from the suspension cage. Increase in seismic motion could couple in a nonlinear fashion to the gravitational wave strain channel leading and amplify scattering noise in 50 to 200 Hz. Such unwanted effects make it difficult to attain the desired sensitivity required for gravitational wave detection.

In Fig. 4.3, we compare a 4.3 magnitude induced earthquake from Oklahoma with a magnitude seven teleseismic event from Japan. The time-frequency plot clearly shows the difference between the two both in frequency content and in duration.

Magnitude 5.0 induced earthquake that occurred at Oklahoma on 7th November 2016 01:44:24 UTC lead to a loss of interferometric lock at the Livingston observatory (See Fig. 4.4). The latency of early warning system currently installed at the site is around 20 minutes which is greater than the arrival time of Rayleigh waves from regional sites like Oklahoma. The event led to a Rayleigh wave amplitude of 11 microns/s and saturated most of the suspension models and possibly the seismometers. The lockloss happened within 5 minutes of the

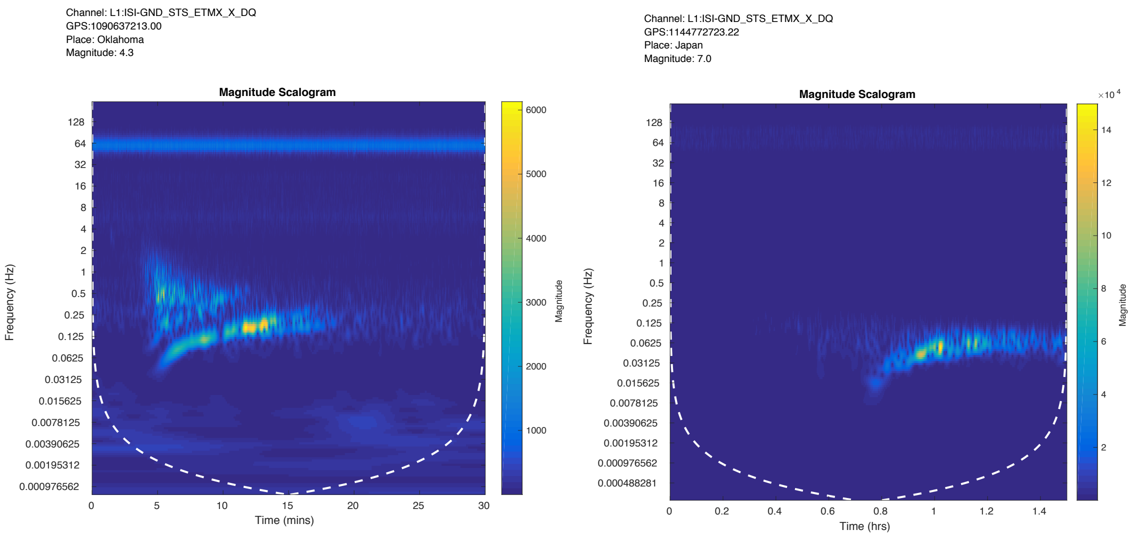


FIGURE 4.3: The plot compares the time-frequency content along the horizontal direction for induced and normal earthquakes. The left plot is caused by a 4.3 magnitude induced earthquake at Oklahoma while the right plot is for seven mag earthquake in Japan. Shallowness in depth and nearness to site causes the Oklahoma event to produce significant motion at higher frequencies.

event which was less than the time it took to show up as an alert in USGS and subsequently in Seismon (LIGO's Earthquake Warning Application,(Coughlin et al., 2017a)). There was a delay of 3 hrs before the interferometer could be successfully locked again although it usually takes around 30-45 minutes to recover from such a quake. In fact, our studies show that any earthquake from Oklahoma with a magnitude above 4.2 and occurring at a depth around 5 km has the potential to cause lockloss at the site. In Fig: 4.5 we show how an earthquake drives the interferometer to a state of lockloss. It shows the behavior of various subsystems (including Alignment Sensing & Control [**ASC**], Internal Seismic Isolation [**ISI**] and Length Sensing & Control [**LSC**]) seconds prior to the lockloss. The excess ground motion triggers instabilities in the control signals which either tend to drive the digital to analog converters beyond their actuation limit or causes short duration transients or glitches finally leading to loss of resonance in the optical cavities.

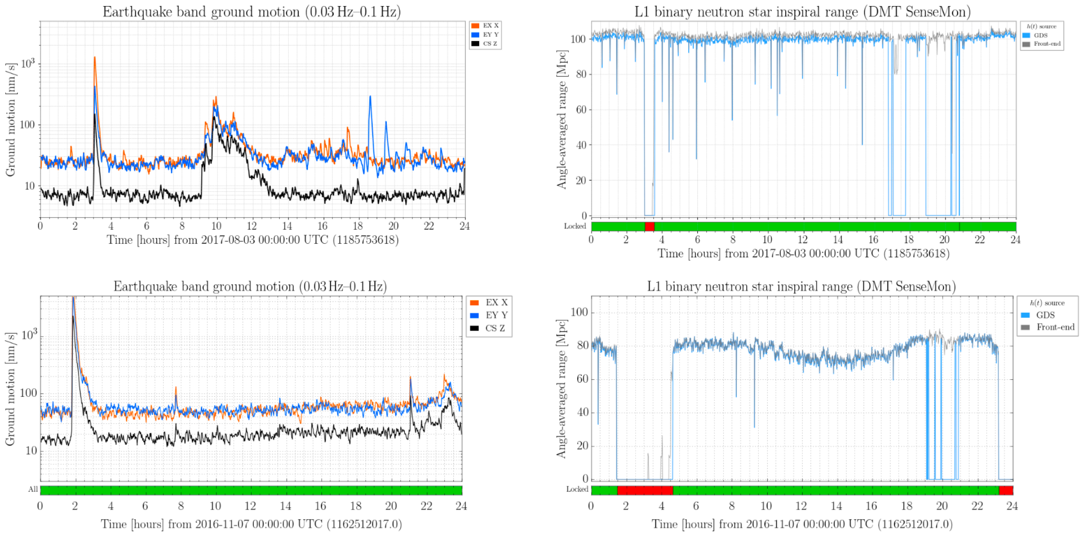


FIGURE 4.4: Effect of Oklahoma earthquakes on the LIGO Livingston detector. The first lockloss caused by a 4.2 magnitude earthquake leads to downtime of thirty minutes. The effect is more severe with the second one of magnitude five where the instrument is seen to take more than three hours to be back in operation.

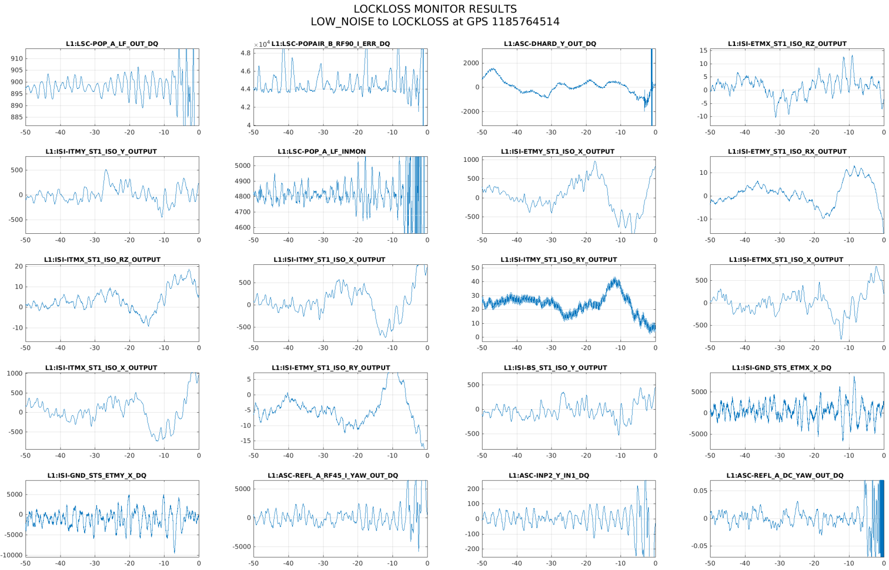


FIGURE 4.5: Effect of 4.2 magnitude Oklahoma earthquake on various LIGO subsystems. Growing oscillations are visible in all the corner and end station vibration isolation control loops tens of seconds prior to lockloss.

4.3.1 Newtonian Noise from Oklahoma earthquakes

As shown previously, the nearby induced events tend to have significant energy content in the frequencies above the earthquake band, with few of them

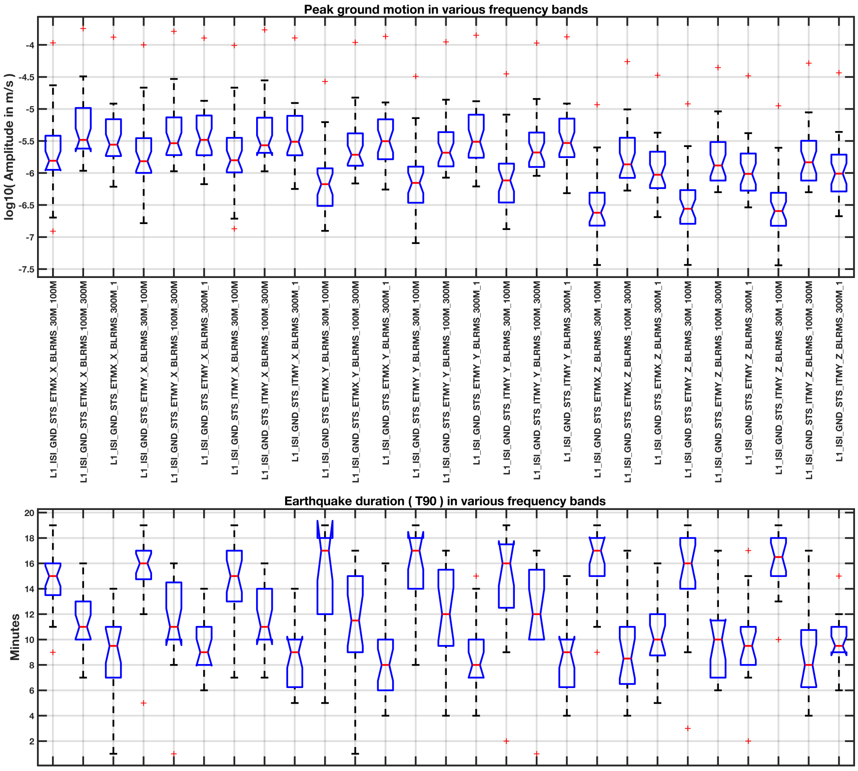


FIGURE 4.6: Box plots showing the distribution of peak ground motion and duration of earthquake as seen in multiple seismometer at different frequency bands.

even showing up to 10 Hz - 30 Hz band it is likely to also have an impact on the Newtonian noise budget. We estimated the level of NN produced by the Oklahoma earthquakes such as those given in Table 4.1. In Chapter 2, we assumed an NN model which is dominated by Rayleigh waves. Fig. 4.7 gives the seismic spectral histogram for these events along with the expected NN. We did not find any significant increase compared to what we expect from typical seismic waves without such earthquake content. This is expected, as our earthquake sample was dominated by events with a magnitude close to 4.3 which do not have much energy content in the NN band of interest 10 Hz - 20 Hz. It needs to be noted that we have very little information about the event rate and parameters of microquakes and events with magnitude less than 2.7 which are hardly picked up by seismic arrays. If there is a background of such events occurring fairly regularly, but randomly, it could then lead to a stochastic NN background that could already be affecting the noise floor and enhancing the effects of scattering noise, etc.

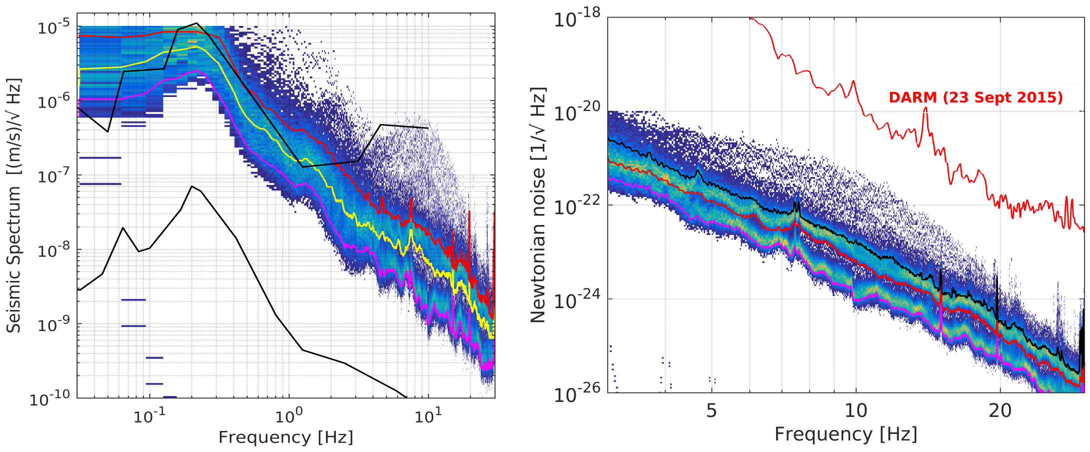


FIGURE 4.7: Seismic spectral histogram of Oklahoma events along with the corresponding Newtonian noise contribution.

4.4 Modeling Expected Ground Motion

Robust modeling is necessary to foresee the effects of possible future events, insights from which incorporated for future site selection studies. In this section, we estimate the expected ground motion via two different methods: (1) Analytical equations and (2) Spectral finite element modeling. We use the region around St. Tammany Parish, Louisiana, located within 60 miles from LLO. Fracking activities are likely to start in this region shortly. It is part of the Tuscaloosa Marine Shale (TMS) region with oil deposits trapped at 11,000-15,000 feet and is geologically similar to the Eagle Ford formation. Fig. 4.8 shows the TMS region along with the active fault structure in the state of Louisiana.

4.4.1 Analytic Model

Peak Rayleigh wave amplitude can be approximately modeled as a function in earthquake magnitude, depth and distance (Coughlin et al., 2017b) as follows

$$Rf_{amp} = M\frac{a}{f_c^b}\frac{e^{-2\pi h f_c/c}}{r^d} . \tag{4.2}$$

Here the free parameters a,b,c,d are best-fit parameters which make the model predict 90 % of the observed motion within a factor of five. For this simulation, we use a = 0.16, b = 1.31, c = 4672.83 and d = 0.81. As for magnitude and depth, we assume a distribution similar to the average US induced seismicity data available since 1967. Within the TMS region, we choose regions with high resistivities (of order seven ohm-m and higher), as they have a higher chance of being exploited for oil. We randomly sample the parameter distribution and estimate the peak amplitude motion using Eq. 4.2. Results are shown in Fig. 4.9,

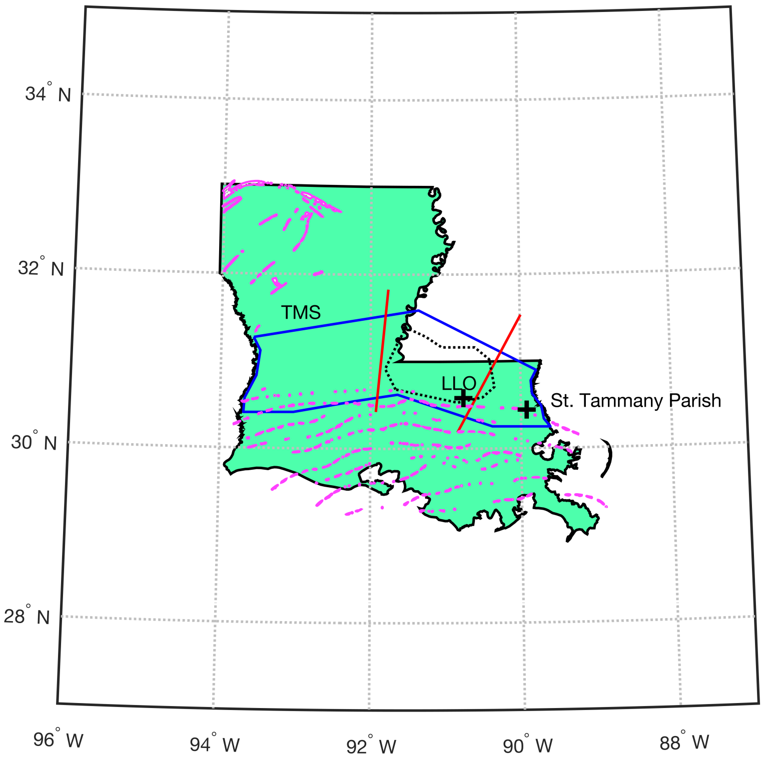


FIGURE 4.8: Map of Louisiana overlaid with known active faults. Tuscaloosa Marine Shale region enclosed within the blue polygon. The region within the TMS having high resistivity is indicated by the dotted line. High resistivity is indicative of the presence of hydrocarbons making the region ideal for fracking.

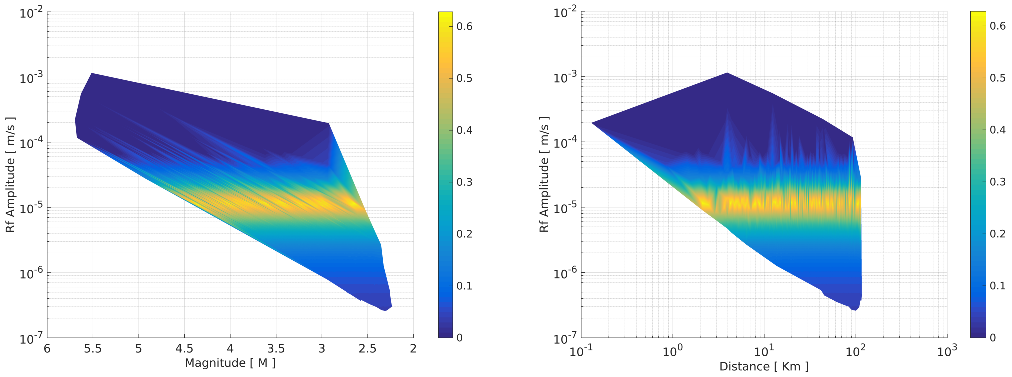


FIGURE 4.9: Simulation of expected ground motion from induced Earthquakes from the Tuscaloosa Marine Shale region. Color bar gives the probability of occurrence.

where we observe more than 50% of the events to cause motion of order tens of microns per second. The proximity to the Livingston site would lead to microquakes of magnitude even 2.7 to be detrimental to the operation of the detector.

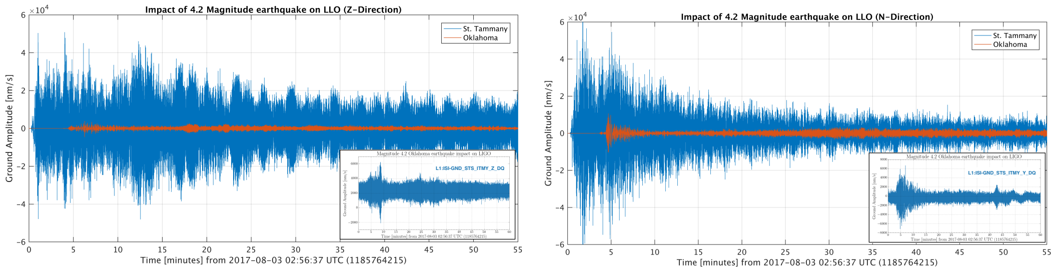


FIGURE 4.10: Expected ground motion along vertical and horizontal direction from a 4.2 magnitude earthquake triggered at St. Tammany Parish (Tuscaloosa Marine Shale) and Oklahoma region modeled using SPECFEM3D. The inset plots show the measured ground motion for a similar earthquake that happened on Aug 3, 2018 at Edmond, Oklahoma.

4.4.2 Spectral Element Model

Next, we study the impact using the forward simulation of the earthquake by spectral finite element method through the SPECFEM3D software (Komatitsch and Vilotte, 1998; Tromp, Komatitsch, and Liu, 2008). It uses spectral element method, a special case of the Galerkin method, to simulate seismic wave propagation. The simulation consists of mesh generation, database generation and finally running the solver.

We recreate the 4.2 Oklahoma earthquake occurred on 3rd August 2018 at Edmond Oklahoma (*M 4.2 - 6km ENE of Edmond, Oklahoma*) and compared the results with the actual seismometer readings obtained from the site. We then compare the results by shifting the fault rupture location to St. Tammany Parish region, a potential fracking site within the TMS region located around 60 miles from LLO. For the forward simulation, we require centroid moment tensor (CMT) values that describe the fault rupture obtained from the United States Geological Survey (USGS) in addition to the location, depth, and local geological information. Mesh of depth 60 km is used with absorbing boundary conditions. We assume the compressional (V_p) and shear velocity (V_s) to be uniform from the source to destination and to follow a vertical trend typical of what is observed in Louisiana. We use results from Gaite et al. (2015) to get the shear wave velocity (V_s) till depths of up to 30 km. V_p is calculated using Castagna's Relation which provides a model for regions near the Gulf Of Mexico as

$$V_p = 1.13\ V_s + 1360\ . \tag{4.3}$$

Our model uses results from (Castagna and Backus, 1993) for the variation of shale density as a function of depth. Fig. 4.10 shows the results of our simulation. We observe the close similarity between the simulated seismogram and the seismometer reading both regarding arrival time and the peak amplitude. It is evident that a similar event from the TMS region would hit the site in less than a minute with an amplitude of 60 microns/second, an order of magnitude higher than the original Oklahoma event. Such events triggered by the exploration and exploitation of Tuscaloosa Marine Shale would be detrimental to the operation of LIGO Livingston.

4.5 Impact on future GW detectors

The recent 4.6 M earthquake that occurred in the Gulf of Mexico on 2018-05-06 16:47:09 UTC registered a band limited RMS well above 6 microns along the horizontal direction in 1 Hz - 10 Hz and was strongly witnessed in the 10 Hz - 30 Hz band (See Fig. 4.11). Although the event is not linked with induced seismicity, it provides an idea about how geographical nearness and shallowness in depth can affect the frequency bands that are of astrophysical significant for the next generation detectors. This observation strengthens the need for future earthquake isolation designs to also consider frequency bands above traditional 30 mHz -100 mHz range.

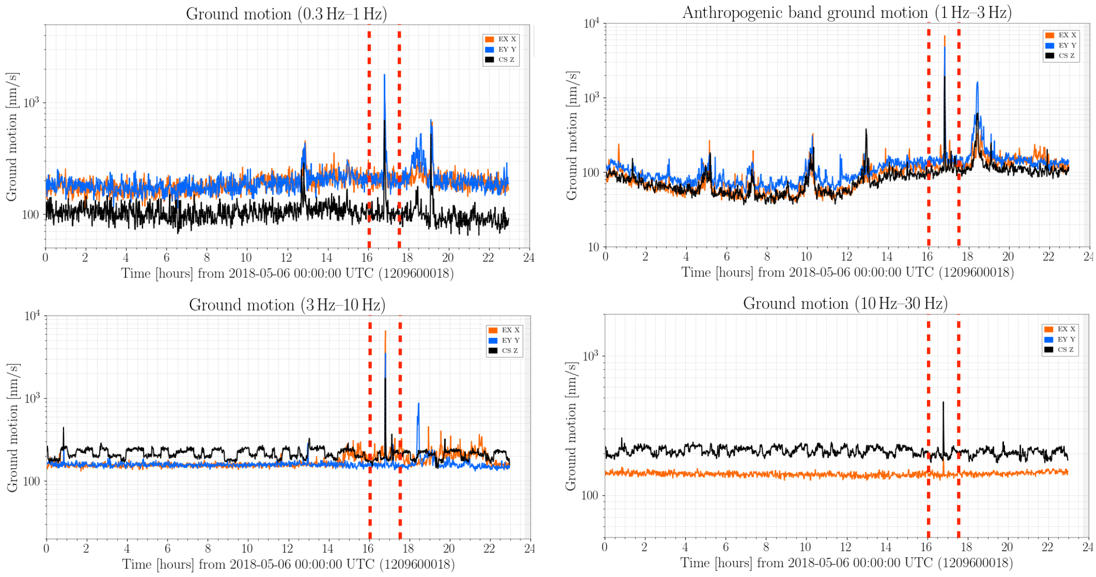


FIGURE 4.11: Observation of high frequency content in the 4.6M earthquake that occurred on 2018-05-06 16:47:09 UTC at the Gulf of Mexico (182km SSE of Buras-Triumph, Louisiana at a depth of 10km).

Several seismic isolation control loops are designed to provide noise suppression in the band of interest (typically above 1 Hz) and often has gain peeking in lower frequencies, most of these could be attributed to sub-optimal loop design. This makes the instrument susceptible to noise re-injection during periods of elevated ground motion. The above sections show that proximity to regions with a high rate of induced seismicity could be detrimental in maintaining steady lock for current and future generation detectors. Future underground observatory like Einstein telescope (Abernathy et al., 2011) is likely to built at a depth of 100 m - 200 m to avoid the effects of surface Rayleigh waves. We repeated the spectral finite element simulations described in section 4.4.2 and measured the ground motion at a depth of 150 m and found that the corresponding suppression in vibration as compared to the surface is negligible. Thus it is advisable to access level of induced seismicity prior to site selection for the next generation terrestrial interferometers.

As the period between the earthquake and its impact is of the order eight minutes, warning systems using gravity gradiometers could be useful to issue low latency alerts at the site. Previous works (Harms et al., 2015; Montagner et al., 2016) have shown the possibility of detecting transient changes in gravity within 10sec after the fault rupture. Such extremely sensitive gravity strainmeters will not be limited by seismic propagation speed and may be able to work independently of the seismometers to inform the site commissioners well in advance of the arrival of the P waves.

A swarm of low magnitude induced earthquakes can lead to a persistent seismic background which in general would be less dominant to the other competing low frequency noises. But the Rayleigh waves from these could cause density perturbations around the vicinity of the detector and contribute to the already existing gravity gradient or Newtonian (NN) noise (Saulson, 1984b; Hughes and Thorne, 1998b). NN is a dominant sensitivity limiting noise at frequencies below 20 Hz and require a post off line subtraction via Wiener filtering using an optimal array of seismic sensors (Harms, 2015; Coughlin et al., 2016). Microquakes from the induced seismicity “hot spots” could alter the seismic correlation patterns thus affecting the overall subtraction efficiency. In future, we would like to study the NN background from a swarm of low magnitude nearby events and study their contribution to the advanced LIGO noise budget.

It has been demonstrated that restrictions on drilling newer wells near areas of reported seismic activity and curbing the operations of existing injection wells lead to decrease in induced earthquakes. For example the state of Arkansas saw an upsurge of earthquakes in 2009 which dropped since 2011 due to prompt legislative measures. Rate of wastewater fluid injection often correlates with the number of observed earthquakes as is seen from the Rocky

Mountain Arsenal data (Evans, 1966; Hsieh and Bredehoeft, 1981). Controlling the amount of injected fluid could be one strategy to minimize the triggering. Restrictions imposed at Oklahoma on the rate of fluid injections since May 28, 2016 led to a decrease in the number of earthquakes and is clearly seen in Fig. 4.12 which shows the cumulative number of earthquakes with magnitude greater than three that occurred in the state.

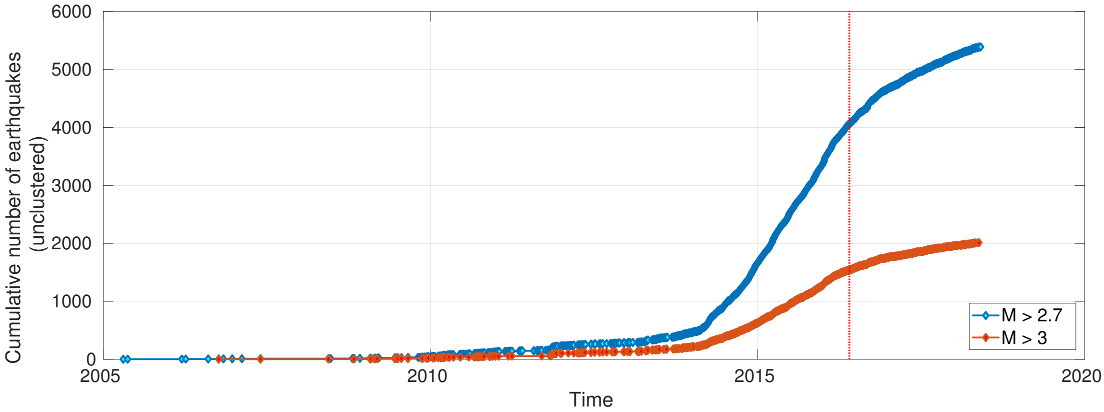


FIGURE 4.12: The cumulative number of earthquakes with magnitude greater than three that happened in Oklahoma. Regulations in place since May 28, 2016, is seen to decrease the frequency of these events.

4.6 Conclusions

In this chapter, we have discussed the effects of induced seismicity on the gravitational wave interferometers. In particular, we accessed the impact of events from Oklahoma and showed that due to the geographical nearness and the shallowness of the hypo-center, earthquakes of magnitude as small as 4.2 from these area could lead to lockloss scenarios at the Livingston detector. The closeness of the source also results in a contribution to the spectral content beyond the earthquake band which can affect the seismic control loops deployed at the interferometer.

We further estimated the possible effects to LIGO Livingston detector from the Tuscaloosa Marine Shale formation which is a potential candidate for future oil extraction in the conterminous US. Proximity to the site implies even minor quakes from these regions could be a potential cause of lockloss at the interferometers. The latency of early warning system currently installed at the site is around 20 minutes which will not be sufficient to handle events described in this here. Alternative schemes based on gravity gradiometers need to be explored for prompt early detection and warning. A better understanding of the

	time	place	mag	latitude	longitude	depth
0	2018-04-09T10:22:20.050Z	27km WSW of Perry, Oklahoma	4.5	36.2165	-97.5820	3.890
1	2018-04-07T12:16:03.300Z	20km W of Perry, Oklahoma	4.6	36.2896	-97.5166	5.791
2	2018-03-05T03:40:59.200Z	15km NE of Enid, Oklahoma	4.2	36.4833	-97.7491	8.229
3	2018-03-04T23:17:17.400Z	15km NE of Enid, Oklahoma	4.2	36.4780	-97.7411	8.103
4	2017-09-08T02:26:23.700Z	12km SSE of Medford, Oklahoma	4.3	36.6996	-97.6830	6.073
5	2017-08-03T02:56:37.300Z	6km ENE of Edmond, Oklahoma	4.2	35.6720	-97.4047	5.031
6	2017-07-14T13:47:35.300Z	12km N of Stroud, Oklahoma	4.2	35.8591	-96.6833	6.813
7	2017-05-13T08:32:37.300Z	32km NW of Fairview, Oklahoma	4.2	36.4782	-98.7365	5.367
8	2016-11-07T01:44:24.500Z	3km W of Cushing, Oklahoma	5.0	35.9907	-96.8030	4.430
9	2016-11-02T04:26:54.000Z	12km ESE of Pawnee, Oklahoma	4.4	36.3047	-96.6657	4.349
10	2016-09-03T12:02:44.400Z	14km NW of Pawnee, Oklahoma	5.8	36.4251	-96.9291	5.557
11	2016-07-17T04:17:58.900Z	20km W of Perry, Oklahoma	4.2	36.2841	-97.5138	4.787
12	2016-07-09T02:04:27.400Z	33km NW of Fairview, Oklahoma	4.4	36.4638	-98.7584	7.242
13	2016-07-08T22:29:38.000Z	33km NW of Fairview, Oklahoma	4.2	36.4749	-98.7459	6.356
14	2016-07-08T21:31:57.600Z	32km NW of Fairview, Oklahoma	4.2	36.4765	-98.7387	7.315
15	2016-04-07T22:27:30.400Z	1km E of Luther, Oklahoma	4.2	35.6619	-97.1741	6.106
16	2016-03-29T04:53:01.200Z	4km NNE of Crescent, Oklahoma	4.2	35.9900	-97.5773	5.183
17	2016-02-13T17:07:06.290Z	31km NW of Fairview, Oklahoma	5.1	36.4898	-98.7090	8.310
18	2016-01-07T08:37:11.100Z	32km NW of Fairview, Oklahoma	4.4	36.4754	-98.7342	6.647
19	2016-01-07T04:27:57.600Z	33km NW of Fairview, Oklahoma	4.7	36.4955	-98.7254	4.058
20	2016-01-07T04:27:27.900Z	33km NW of Fairview, Oklahoma	4.4	36.4860	-98.7412	7.093
21	2016-01-01T11:39:39.800Z	6km ENE of Edmond, Oklahoma	4.2	35.6688	-97.4065	5.825
22	2015-12-29T11:39:19.200Z	6km ENE of Edmond, Oklahoma	4.3	35.6654	-97.4054	6.532
23	2015-11-30T09:49:12.800Z	26km E of Cherokee, Oklahoma	4.7	36.7509	-98.0561	5.629
24	2015-11-23T21:17:46.500Z	11km NE of Cherokee, Oklahoma	4.4	36.8382	-98.2762	5.028
25	2015-11-19T07:42:12.000Z	13km SW of Cherokee, Oklahoma	4.7	36.6602	-98.4594	5.914
26	2015-11-15T09:45:31.300Z	33km NW of Fairview, Oklahoma	4.3	36.4696	-98.7549	5.109
27	2015-10-10T22:03:05.300Z	3km W of Cushing, Oklahoma	4.3	35.9860	-96.8032	3.274
28	2015-10-10T09:20:43.000Z	20km WSW of Medford, Oklahoma	4.4	36.7187	-97.9311	5.629
29	2015-07-27T18:12:15.400Z	4km NNE of Crescent, Oklahoma	4.5	35.9889	-97.5717	5.000
30	2015-07-20T20:19:03.400Z	13km NE of Cherokee, Oklahoma	4.4	36.8420	-98.2593	4.079

TABLE 4.1: List of earthquakes that occurred in Oklahoma with a magnitude of at least 4.2

underlying fault structure and accurate data on the distribution of class II injection wells along with the rate of injection will be essential to minimize the impact induced seismicity on the ground-based interferometers. Accessing the level of induced seismicity near the site will be beneficial for operation of next generation ground based and underground gravitational wave observatories.

Chapter 5

System Identification for Feedforward Noise Cancellation

5.1 Introduction

Control system design in GW interferometers is mostly based on measured transfer functions which often tend to be corrupted by noise. A significant amount of time is spent on developing a model based on these measurements which are further used to design the optimal controller. In this chapter, we present methods used for system identification of linear time-invariant systems and possible ways of their automation.

In section 5.3 we describe the technique of vector fitting, a powerful method that generates stable infinite impulse response (IIR) filters which often provide good approximations to the measured transfer functions. Sections 5.4 & 5.5 describe characteristics of feedback (FB) & feedforward (FF) filters and show how carefully implemented FF filters can carry out noise suppression in seismic platforms that host the LIGO test masses. The resourcefulness of FIR Wiener filtering in removing the persistent noise as well as transient glitches is demonstrated in section 5.6. Effects of removing correlation on the matched filter analysis are provided in 5.8. Plant dynamics are often susceptible to environmental factors like temperature. Hence filters that can track such variation and modify the control signal often perform superior to their static counterparts. Benefits of using the adaptive version for tracking variations in impulse response in shown in section 5.9.

5.2 Characterizing Linear Time Invariant Systems

A system which does the following mapping $y(t) = \mathcal{H}x(t)$ is called a linear time invariant (LTI) system if it satisfies the following relations:

- Linearity,

$$\mathcal{H}[\alpha x1(t) + \beta x2(t)] = \alpha \mathcal{H}[x1(t)] + \beta \mathcal{H}[x2(t)] \ . \tag{5.1}$$

- Time Invariance,

$$y(t) = \mathcal{H}[x(t)] \implies y(t-\tau) = \mathcal{H}[x(t-\tau)] \ . \tag{5.2}$$

We can completely characterize an LTI system through its impulse response $h(t)$ in time domain or through the transfer function $H(\omega)$ in the frequency domain. $H(\omega)$ and $h(t)$ are related to each other through the Fourier transform,

$$H(\omega) = \mathcal{F}[h(t)] = \int_{-\infty}^{+\infty} h(t) e^{-i\omega\tau} d\tau \ . \tag{5.3}$$

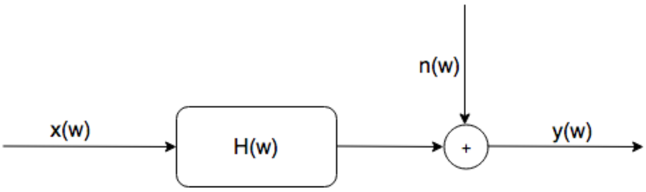


FIGURE 5.1: Linear time invariant (LTI) system subject to an excitation signal.

As shown in Fig. 5.1, if we excite a linear system, by sending a signal $x(\omega)$ and then corresponding response $y(\omega)$ in the presence of noise $n(\omega)$ is given by

$$y(\omega) = x(\omega) H(\overrightarrow{\theta}; \omega) + n(\omega) \ , \tag{5.4}$$

where the transfer function, $H(\overrightarrow{\theta}; \omega)$ is calculated as

$$H(\overrightarrow{\theta}; \omega) = \frac{S_{xy}(\omega)}{S_{xx}(\omega)} \ . \tag{5.5}$$

Here S_{xy} and S_{xx} are respectively the input-output cross-power spectral density and the input power spectral density. In terms of coherence C_{xy},

$$H(\overrightarrow{\theta}; \omega) = \frac{S_{xy}(\omega)}{S_{xx}(\omega)} = \frac{< \tilde{x}(\omega)\tilde{y}(\omega)^* >}{< \tilde{x}(\omega)\tilde{x}(\omega)^* >} = C_{xy} \sqrt{\frac{< \tilde{y}(\omega)\tilde{y}(\omega)^* >}{< \tilde{x}(\omega)\tilde{x}(\omega)^* >}} \ , \tag{5.6}$$

where

$$C_{xy} = \frac{< \tilde{x}(\omega)\tilde{y}(\omega)^* >}{\sqrt{< \tilde{x}(\omega)\tilde{x}(\omega)^* >}\sqrt{< \tilde{y}(\omega)\tilde{y}(\omega)^* >}} \ . \tag{5.7}$$

5.3 Fitting the measured transfer functions

TF Measurements are carried out by sending excitations ($\tilde{x}(\omega)$) either as swept sine or band-limited white noise through an actuator and recording the system response as measured by a sensor. The amplitude of the actuator signal is often chosen such that the output signals have high coherence with respect to the input. One methodology adopted is to fit the measurement with a rational function approximation. The linear model so created can be used to fine tune the controller either through software-in-loop simulations or via hardware-in-loop simulations. $H(s)$ can be approximated as a ratio of two polynomials,

$$H(s) \approx \frac{a_0 + a_1 s + a_2 s^2 + ... + a_N s^N}{b_0 + b_1 s + b_2 s^2 + ... + b_N s^N} \ . \tag{5.8}$$

It can also be written as a rational function,

$$H(s) \approx \sum_{n=1}^{N} \frac{c_n}{s - a_n} + d + sh \ . \tag{5.9}$$

The process of finding the optimal coefficients of the above rational function is done through the process of vector fitting (Gustavsen and Semlyen, 1999; Gustavsen, 2006). It has to be noted that the Eq. 5.9 cannot be solved as a linear problem as the unknowns a_n occur in the denominator. Vector Fitting (VF) tackles this by solving it hierarchically in two stages, the first one being pole identification followed by the identification of residues. We start with an initial set of poles $\tilde{a}_n$ and use an unknown function $\sigma(s)$ so that the problem now becomes,

$$\begin{bmatrix} \sigma(s)H(s) \\ \sigma(s) \end{bmatrix} \approx \begin{bmatrix} \sum\limits_{n=1}^{N} \frac{c_n}{s-\tilde{a}_n} + d + sh \\ \sum\limits_{n=1}^{N} \frac{\tilde{c}_n}{s-\tilde{a}_n} + 1 \end{bmatrix} \ . \tag{5.10}$$

The above equation can be transformed into the following form which now becomes linear in its unknowns $c_n, d, h, \tilde{c}_n$. This leads to

$$\left(\sum_{n=1}^{N} \frac{c_n}{s - \tilde{a}_n} + d + sh \right) - \left(\sum_{n=1}^{N} \frac{\tilde{c}_n}{s - \tilde{a}_n} \right) H(s) \approx H(s) \ , \tag{5.11}$$

which has the form,

$$A_k x = b_k \ , \tag{5.12}$$

where

$$A_k = \left[\frac{1}{s_k-\tilde{a}_1} \quad \dots \quad \frac{1}{s_k-\tilde{a}_N} \quad 1 \quad s_k \frac{-f(s_k)}{s_k-\tilde{a}_1} \quad \dots \quad \frac{-f(s_k)}{s_k-\tilde{a}_N},\right], \tag{5.13}$$

$$x = \left[c_1 \quad \dots \quad c_N \quad d \quad h \quad \tilde{c}_1 \quad \dots \quad \tilde{c}_N\right], \tag{5.14}$$

$$b_k = f(s_k)\,. \tag{5.15}$$

For complex poles, the consecutive terms that occur in the partial fractions for the initial poles $\tilde{a}$ and the residues c are constructed of complex conjugate pairs. Rewriting the transfer function, we get

$$H(s) = \frac{\left(\sum\limits_{n=1}^{N} \frac{c_n}{s-\tilde{a}_n} + d + sh\right)}{\left(\sum\limits_{n=1}^{N} \frac{\tilde{c}_n}{s-\tilde{a}_n} + 1\right)}, \tag{5.16}$$

which in fractional form becomes,

$$H(s) = \frac{h\frac{\prod\limits_{n=1}^{N+1}(s-z_n)}{\prod\limits_{n=1}^{N}(s-\tilde{a}_n)}}{\frac{\prod\limits_{n=1}^{N}(s-\tilde{z}_n)}{\prod\limits_{n=1}^{N}(s-\tilde{a}_n)}} = \frac{h\prod\limits_{n=1}^{N+1}(s-z_n)}{\prod\limits_{n=1}^{N}(s-\tilde{z}_n)}\,. \tag{5.17}$$

The poles a_n of $H(s)$ are given by the zeros $\tilde{z}_n$ which are obtained as the eigenvalues of the matrix $A - b\tilde{c}^T$. The residues of $H(s)$ are finally obtained by solving Eq. 5.9 making use of a_n obtained in the previous step. The accuracy of vector fitting is seen to be sensitive to the initial location of the poles, and the use of real poles is seen to cause ill-conditioning within the linear problem. So the recommended strategy would be to use a set of complex conjugate starting poles who's imaginary part linearly span the interested frequency range with their real part taking similar values but suppressed by a factor of hundred.

Based on the above mentioned fitting procedure, we developed an application called 'TFestimate' [1] to automate the process of transfer function based system identification (See Fig. 5.2) . One issue with the measurements is that they are often noisy which confuses the fitting routine leading it to use more poles and zeros. To tackle this, the application lets the user smooth the measurements before fitting. This is achieved either through convolution with Savitzky–Golay filter (Savitzky and Golay, 1964) or using cubic spline smoothing.

---

[1] Github link to TFestimate App: `https://github.com/Nikhil-Mukund/TFestimate`

Input Data Type: ● [Freq Real_TF Imag_TF] ○ [Freq Complex_TF] ○ FRD Model

Chose TF: isolated_one_over_plant.txt

Min Poles 6   Max Poles 11

Frequency Weighting Filter: ● Passbands (Hz) 0.01,5

Result_folder_name TF_Results   Max Iter 100

Init Method svf   Options: iv, svf, gpmf, n4sid, auto

Search Method lsqnonlin   Options: gn, gna, lm, lsqnonlin, grad, auto

Smooth 10   Extreme Smooth 0.9999900

Fig Name test

FIGURE 5.2: TFestimate: A MATLAB GUI Application for fitting measured transfer functions.

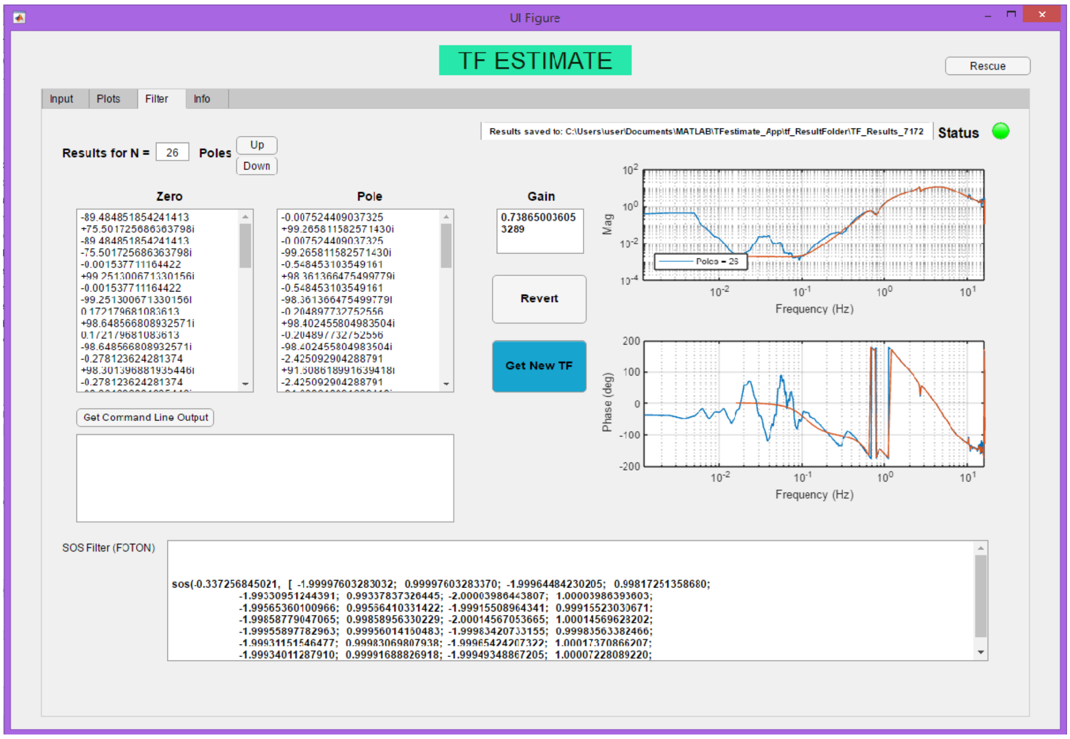
The large dynamic range across the given frequency bandwidth can be tackled by weighting the cost function using the inverse of the transfer function or by focusing only on the desired range and ignoring the rest of frequency band.

The result of fitting an inverse plant transfer function is shown in Fig. 5.3. The application ensures stability by forcing all the poles to be on the left half plane. It generates models in both zero-pole-gain (ZPK) format and as second-order-section (SOS) filter format.

5.4 Seismic Control: Feedback vs Feedforward

Typical ground vibration at $10Hz$ is of order few $nm/\sqrt{Hz}$ while for sensing GWs we need to push this to $10^{-19}m/\sqrt{Hz}$. This is achieved in advanced GW detectors using a combination of active and passive vibration isolation techniques. Passive ways include the use of multiple stage pendulums, hydraulic external pre-isolators or mass dampers tuned to certain resonant modes or through the use of materials like Viton. Active methods involve the use of feedback and feedforward control loops to achieve desired levels of isolation.

Fig. 5.4 compares the pathways for feedback (FB) and feedforward (FF) control systems. In FB, sensor is placed on the platform and is good for set point tracking. Sensors while picking up ground noise also add their noise ($10^{-13}\ m\ to\ 10^{-14}\ m$) which implies that sensor noise would be re-injected into



Results saved to: C:\Users\user\Documents\MATLAB\TFestimate_App\tf_ResultFolder\TF_Results_7172

Status

Results for N = 26 Poles

Zero

```
-89.484851854241413
+75.501725686363798i
-89.484851854241413
-75.501725686363798i
-0.001537711164422
+99.251300671330156i
-0.001537711164422
-99.251300671330156i
0.172179681083613
+98.648566808932571i
0.172179681083613
-98.648566808932571i
-0.278123624281374
+98.301396881935446i
-0.278123624281374
```

Pole

```
-0.007524409037325
+99.265811582571430i
-0.007524409037325
-99.265811582571430i
-0.548453103549161
+98.361366475499779i
-0.548453103549161
-98.361366475499779i
-0.204897732752556
+98.402455804983504i
-0.204897732752556
-98.402455804983504i
-2.425092904288791
+91.508618991639418i
-2.425092904288791
```

Gain

```
0.73865003605
3289
```

SOS Filter (FOTON)

```
sos(-0.337256845021,  [ -1.99997603283032;  0.99997603283370;  -1.99964484230205;  0.99817251358680;
                -1.99330951244391;  0.99337837326445; -2.00003986443807;  1.00003986393603;
                -1.99565360100966;  0.99566410331422; -1.99915508964341;  0.99915523030671;
                -1.99858779047065;  0.99858956330229; -2.00014567053665;  1.00014569628202;
                -1.99955897782963;  0.99956014150483; -1.99983420733155;  0.99983563382466;
                -1.99931151546477;  0.99983069807938; -1.99965424207322;  1.00017370866207;
                -1.99934011287910;  0.99991688826918; -1.99949348867205;  1.00007228089220;
```

FIGURE 5.3: Results for fitting inverse plant transfer function using the TFestimate App. The application is asked to provide a good fit using only right half plane poles in the frequency range 0.1 Hz to 10 Hz.

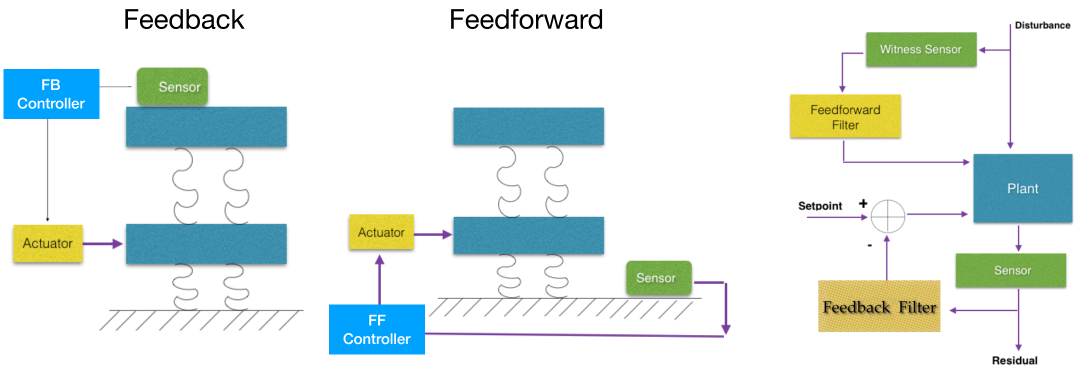
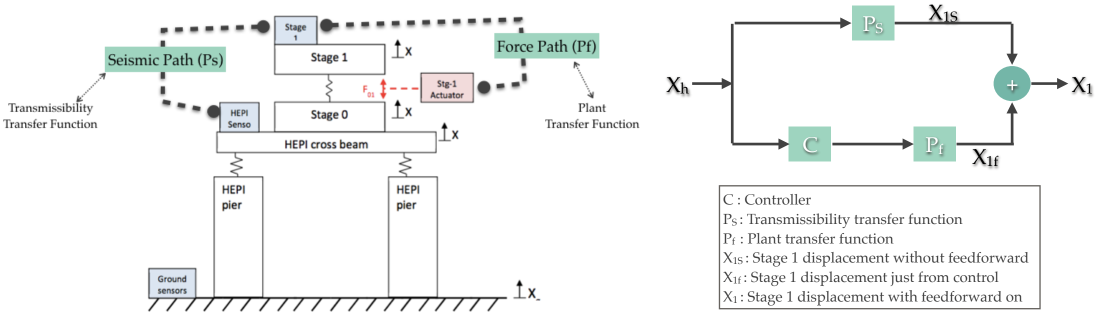


FIGURE 5.4: Schematic of feedback and feedforward control strategies. Feedback is suited for setpoint tracking while feedforward is used from noise rejection.

the system. But the variation in temperature and shifts in mechanical resonances does not affect the loops. FB, in general, does not require an accurate system model but is not very effective for multi-order systems which have long time delays. In FF, the sensor is attached to the ground hence ground noise affects the actuator in a causal manner with a comparatively lower time lag.

Sensor noise is added before the passive seismic isolation. The downside of FF is that if the cancellation to be achieved is a factor of 100 then the system has to be modeled within 1 percent error. While modeling the system to within 10 percent is often easy, getting to around 1 percent is quite difficult. When measurements of the transfer function of the system are made, the actual statistical uncertainty on each data point is $\sigma = \sqrt{\frac{1-C}{2CN}}$ where C is the coherence and N is the number of averages. Quite a lot of measurement excitations or measurement time is required for this purpose which naturally makes it more challenging to implement compared to the feedback method.



(i) (ii)

FIGURE 5.5: (i) Seismic isolation stack consisting of hydraulic external pre-isolator and the actively controlled platforms. (ii) Feedforward pathway with the optimal controller, C given as $-\frac{P_s}{P_f}$.

5.5 Implementation of Feedforward Controls at LLO

Schematic of seismic stack used for isolating test mass and auxiliary optics is shown in Fig. 5.5. The optimal feedforward controller C can be obtain as follows,

$$\begin{aligned} X_1 &= X_{1S} + X_{1f} \rightarrow 0 \; (optimal\ controller) \\ &\implies X_h \cdot P_s + X_h \cdot C \cdot P_f = 0 \\ &\implies C = -\frac{P_s}{P_f} \end{aligned} \tag{5.18}$$

As described previously this technique requires precise information and modeling of the plant and the transmissibility transfer function which is quite often a challenging task.

5.5.1 Feedforward at HAM2 & HAM3 Chambers

The Horizontal Access Module (HAM) chambers HAM2 and HAM3, host the input mode cleaners as well as the power recycling cavity mirrors. Fig. 5.6 shows the results from the implementation of feedforward controller at these chambers. We use the signal from the HEPI L4C seismic sensor to reduce Internal Seismic Isolation (ISI) stage 1 table motion in X, Y, Z degrees of freedom (Mukund, 2017a). The target was to reduce the peaks seen in 5-30 Hz which showed coherence with ground motion. As described in Eq. 5.18 the feedforward transfer function is estimated as -T/P (cts/nm/s) where Transmissibility, T (nm/s /nm/s) is measured between HEPI L4Cs and Stage 1 GS13 and Plant, P (nm/c/cts) is measured between the FF excitation channel and Stage 1 GS13. Fitting is carried out using the method of rational fitting as described in Section 5.3. We performed data preconditioning via median filtering and cubic spline smoothing before fitting the noisy transfer function. Plant and Transmissibility measurements were done with the chamber in an isolated state. Prominent peak seen at 16.5 Hz in both the chambers is seen to be reduced by a factor of 5 in Z and Y dofs. While peak at 29.5 Hz in X dof is reduced by a factor of 2. Overall noise in the frequency range 5 to 20 Hz got suppressed by a factor of 4-10 in both the chambers. We could also not find any noise injection from these new FF filters at frequencies outside the target of interest in all the dofs.

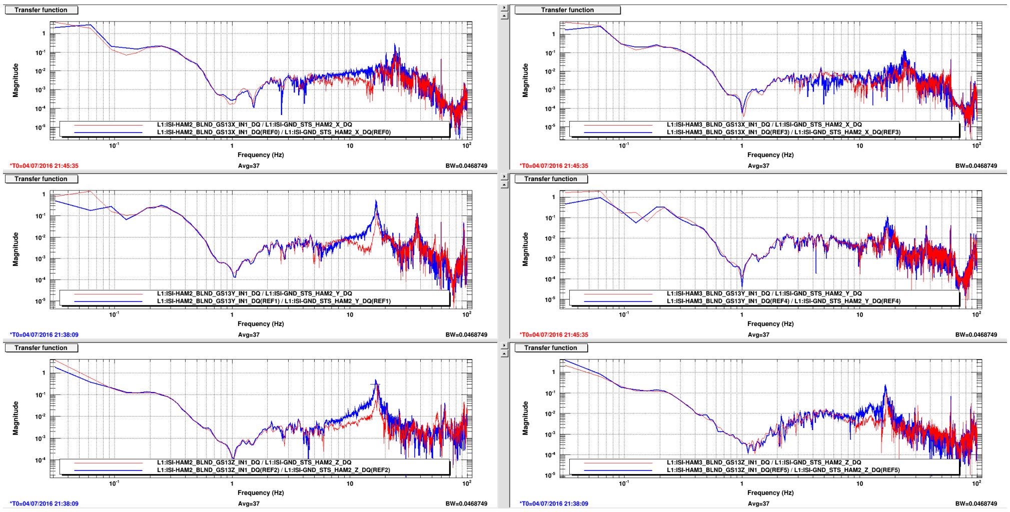


FIGURE 5.6: Each column respectively show how the LLO HAM 2 & 3 stage 1 platform noise (blue) gets suppressed (red) along the X,Y,Z directions in the frequency range 5 to 20 Hz when the feedforward controllers are engaged.

5.5.2 Feedforward at BSC End Station

The test mass optics and the beam splitter are isolated from ground motion through the Basic Symmetric Chamber (BSC). Here we used similar techniques as discussed in the previous section to suppress the vertical motion of the Stage 2 platform at both the end stations at LLO using the respective sensor signal from their Stage 1 L4C. The black line in Fig. 5.7 shows the seismic ground motion while the brown & blue curves respectively show the Stage 2 motion before and after the filter is engaged. We observe a reduction in seismic transmission to stage 2 in frequency 10 Hz to 20 Hz by a factor of $\sim$ 10 in both the chambers. Coherence between the two stages, often used as a measure of linear coupling, is also seen to decrease in the frequency 1 - 15 Hz. Results from similar implementation at the corner station are shown in Fig. 5.8.

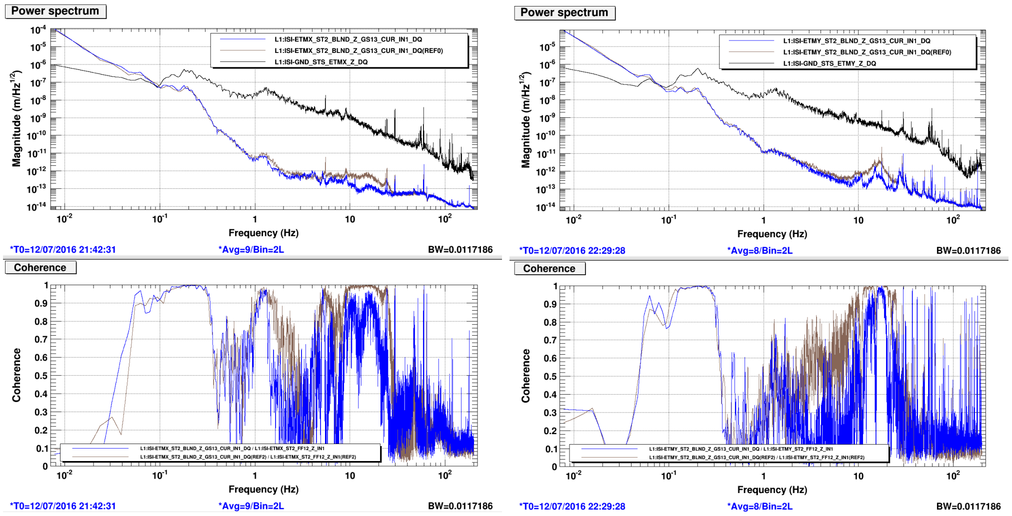


FIGURE 5.7: Top row shows how the BSC end chamber stage 2 platform motion (brown) caused by the ground vibration (black) is further suppressed (blue) by a factor by of 10 in frequency range 10 - 20 Hz when the FF filters are engaged. The bottom row shows the corresponding decrease in coherence between motion sensors kept on stage 1 and stage 2.

5.5.3 Length to Angle (L2A) Decoupling

Cross-coupling either from length to angle (L2A) or from angle to length (A2L) is routinely observed whenever there are suspended optics. For example, if the laser beam is off-centered, then any angular motion (jitter) can couple to changes in an optical path leading to displacement noise. Or, if there is any imbalance in the length control signal applied to the test mass coil drivers, then it could cause angular motion in either pitch or yaw degrees of freedom. Proper identification and modeling of length to angle transfer function can be used to

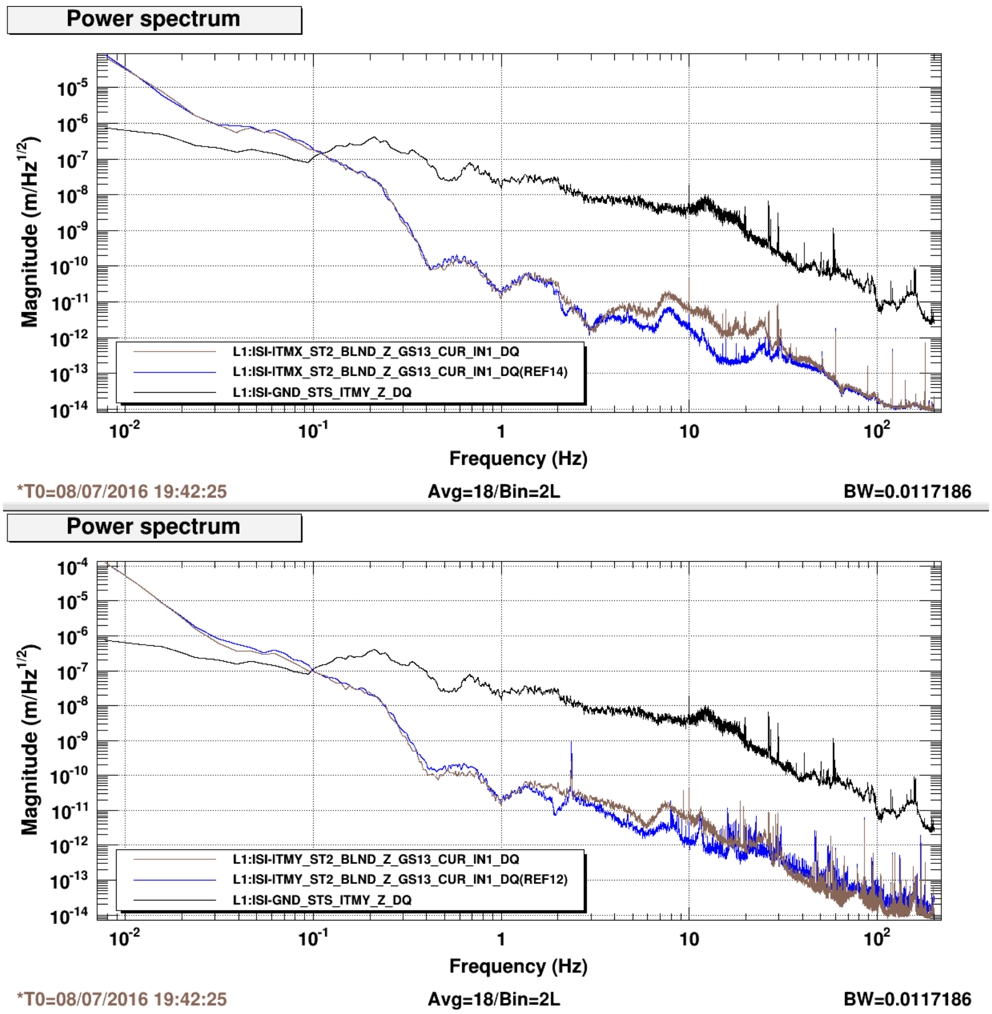
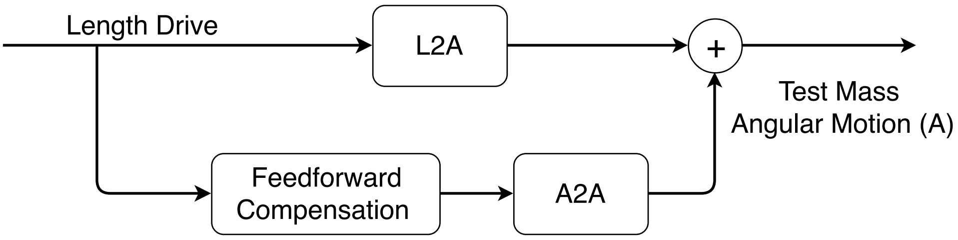


FIGURE 5.8: Noise suppression at the corner station BSC chamber using a feedforward controller. We see an improvement in the platform motion (brown) for the input test mass optics X & Y further improved (blue) in the range 1 Hz - 15 Hz.

decouple this noise via feedforward cancellation. Fig. 5.9 shows the length to pitch (L2P) noise suppression power recycling cavity mirror where the system identification is carried out using the previously described TFestimate application. Accurate fitting leads to a factor of ten suppression in the range 0.01 Hz to 1 Hz (Pele, 2017).

5.6 Wiener Filtering

Wiener filter obtained by solving the ***Wiener-Hopf Equation*** can yield the optimal filter coefficients which will minimize the mean squared error (MSE) between the target and the witness channels. It is an FIR filter and is free from instability issues caused by unstable poles as in IIR filters. They were recently used to remove certain linear coupling from the GW strain channel that leads to



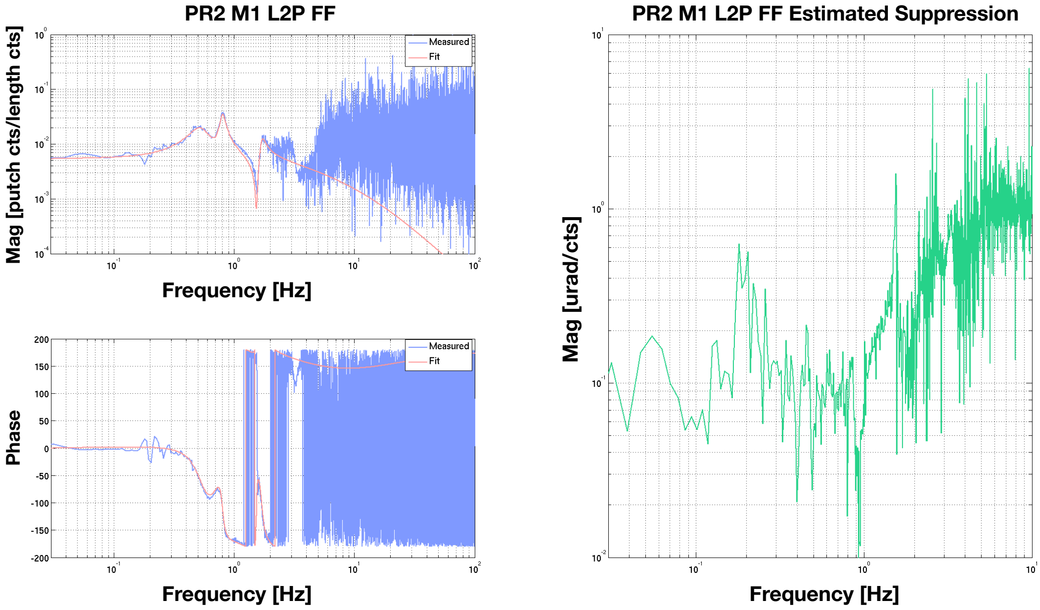


FIGURE 5.9: We show an instance when the system identification needs to be carried out on a noisy transfer function. The best fit model is used in the implementation of Length to Angle decoupling filter for the LLO Power Recycling Mirror.

significant improvement in the observation range by around 25 Mpc. If d is the desired signal to be cleaned and $y = \overline{W}^T x$ is the linear filter output constructed using the witness channel x and filter coefficients $\overline{W}$, then the minimization of MSE leads to,

$$\begin{aligned} J_n \equiv E[e^2(n)] &= E[\ (d(n) - y(n))^2\ ] && (5.19) \\ &= E[\ (d(n) - \overline{W}^T \overline{x})^2\ ] && (5.20) \\ &= E[d^2(n)] - 2\overline{W}^T \overline{P} + \overline{W}^T R \overline{W}\,, && (5.21) \end{aligned}$$

where $\overline{P}$ is the cross correlation matrix between witness and target channels. R is the auto-correlation matrix for witness channels. Minimizing J with respect to the filter coefficients $\overline{W}$ lead to the optimal Wiener filter solution,

$$\frac{dJ_n}{d\omega_i} = \frac{d\epsilon^2}{d\omega_i} = 2[R\overline{W} - \overline{P}], \tag{5.22}$$

$$\overline{W} = R^{-1}\overline{P}. \tag{5.23}$$

Below we demonstrate two scenarios where Wiener filter is used to clean the LIGO auxiliary channel data. We use the input test mass suspension position monitor at the Hanford corner station as our target channel while the seismometer used for physical environment monitoring at the LVEA (laser and vacuum equipment area) is used as the witness channels. To check for performance on newer data sets the same filter is applied to thirty-minute chunks of data measured since 07:00:00 UTC 10-3-2010 (twenty days later). The same procedure is done for the adaptive filter where the initial coefficients are set using the previously calculated Wiener coefficients. Seismic spectra seem to vary over timescales of 30 min as seen in the fluctuations of ITMX SUS-Position signal. Wiener filter remains stable enough to provide noise reduction especially in the lower frequency band (see Fig. 5.10). Later in the section, we will show how an adaptive algorithm outperforms static Wiener and provides better tracking of fluctuations.

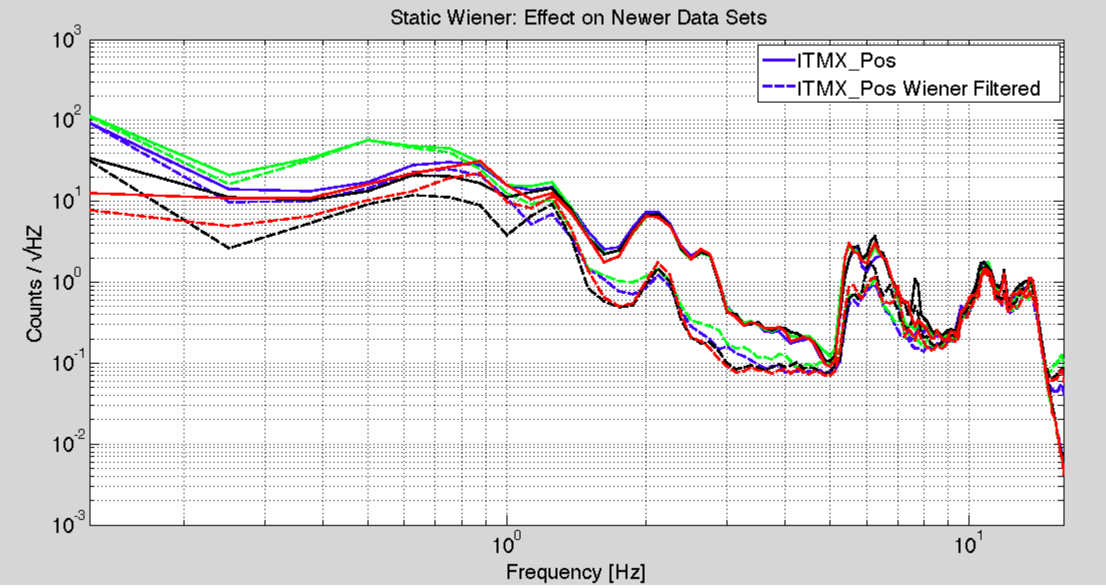


FIGURE 5.10: Application of pre-trained Wiener filter for noise suppression in suspension point motion. The performance is accessed by applying the filter to thirty minute chunks of data obtained few days later.

Now we demonstrate the effectiveness of the algorithm on data corrupted with glitches. Such glitches can often confuse the controllers and can drive the actuator to saturation. Our target signal is the horizontal motion sensed by the L4C sensors mounted on the Y-end station HEPI external pre-isolator with witness channel being the PEM seismometer located close to the isolation platform. Through this post-offline filter construction, we demonstrate that the overall signal amplitude can be minimized along with the suppression of the above-mentioned glitches. What we need is an online version of the algorithm for real-time cancellation. Previously this would require us to construct an IIR equivalent of the same which often degrades the performance due to challenges in stable IIR filter design. Recently LIGO has added the provision to directly implement the FIR version thus making Wiener filtering more practical.

It has to be kept in mind that the matrix inversion required in Eq. 5.23 results in issues related to ill-conditioning whenever the witness channels are highly correlated. It is also based on the strict assumption that the noise is stationary Gaussian which is not always true. It is also very likely that optimal solution might not even exist for MISO (Multiple Output Single Output) systems. And if the transfer function does change with time (say due to temperature fluctuations), the filter will not be able to track those variations. This motivates the need for an adaptive algorithm, which starts with some arbitrary filter coefficient values and after a brief training period settles down to the optimal Wiener solution.

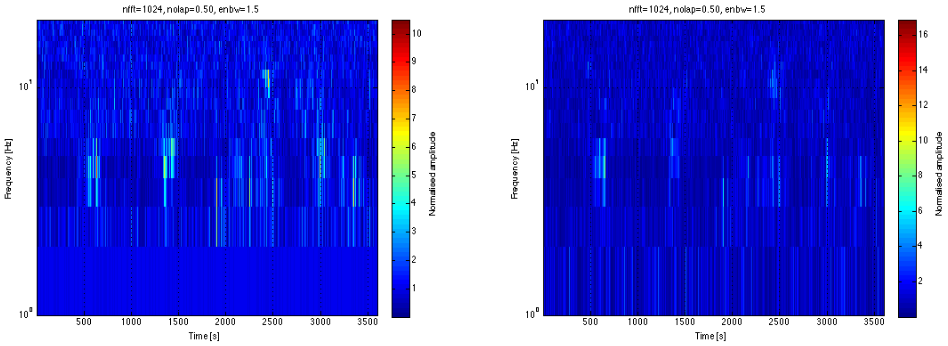


FIGURE 5.11: Demonstration of seismic glitch removal from the HEPI external pre-isolator platforms using the Wiener filtering technique.

5.7 Adaptive Filtering

In adaptive filtering, the impulse response of the filter can be adjusted to remove the correlated noise signal. No previous knowledge of signal and noise is required for these filters. Their filter coefficients vary and adjust themselves through an adaptive algorithm. After a brief initial learning period, the filter coefficients settle down to the optimal Wiener solution effectively reducing the noise content. This dynamic nature makes them extremely important for real-time filtering.

5.7.1 Adaptive Noise Cancellation

Adaptive Noise cancellation is a form of filtering which involves estimating the noise signal by filtering the reference input signal and then subtracting it from the primary input which contains both the original signal and the correlated noise. The basic strategy involved here is to minimize the total output power, this reduces total noise which in turn maximizes the signal to noise ratio of the output signal. Noise reduction in this manner does not distort the original signal and removes only the uncorrelated noise.

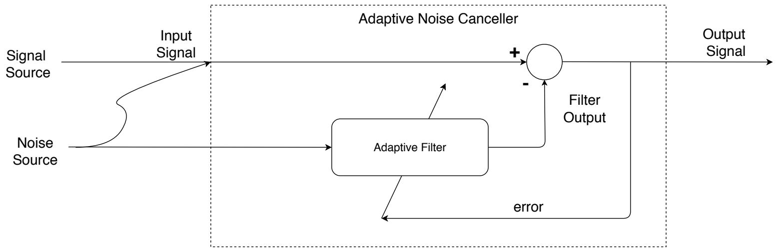


FIGURE 5.12: Adaptive Noise Cancellation Schematic

It is clear from the block diagram that the adaptive noise canceler has two inputs: primary and reference. Primary consists of the input signal which is of our interest and the reference signal is a noise signal which is uncorrelated with input signal but correlated with the noise signal that goes into the primary. The filtered reference signal is subtracted from the primary to get an output or error signal which is an estimation of the actual signal. This error signal is fed back to the adaptive filter which correspondingly modifies the filter coefficients.

$$\hat{s} = (s + n) - \hat{n}\,,$$

$$min\ E[\hat{s}^2] = min\ E[s^2] + min\ E[(n - \hat{n})^2]\,. \tag{5.24}$$

The key idea is that by minimizing the total output power, the output noise power gets minimized and thus output SNR gets maximized.

5.7.2 From Steepest Descent to Least Mean Square

The Wiener filter previously discussed is optimal only if proper statistics of the underlying signals are available (Haykin, 2008). This method may not always suitable in GW detectors since exact knowledge about the noise characteristics is often difficult to obtain, and also the no,is also varies in time. Another strategy would be to use an iterative algorithm that would start with some arbitrary value of filter coefficients and would then gradually move towards the optimal tap weights. The main aim involved is to reduce the cost function. A good candidate for this purpose is the steepest descent algorithm. Here one starts with an arbitrary point on the performance surface and proceeds in the direction which minimizes the cost function. This is along the steepest descent slope of the performance surface at that point. Again this process is continued for some finite number of iterations till filter coefficients approach the optimal Wiener solution. Here ω_n is the previous filter coefficient, ω_{n+1} is the updated coefficient, ϵ^2 represents error, and μ is called the learning rate parameter. The algorithm that we actually used is a variant of the steepest descent algorithm called the least mean square algorithm. It is obtained by having some approximations to the previous algorithm which is explained below. Expectation value of the square of the error signal is used as a figure of merit for obtaining the Wiener filter. Difference between the target signal and the signal calculated from the filtered witness channels gives the error signal. The filter update equation in this case is given by

$$\omega_{n+1} = \omega_n - \mu[R\overline{W} - \overline{P}] \,. \tag{5.25}$$

The cross-correlation vector $\overline{P}$,

$$P = E[\bar{x}(n)d(n)] \,, \tag{5.26}$$

can be written as

$$P = \frac{\bar{x}(n)d(n) + \bar{x}(n-1)d(n-1) + ... + \bar{x}(n-N)d(n-N)}{N+1}] \simeq \bar{x}(n)d(n) \,. \tag{5.27}$$

Similarly, the autocorrelation matrix R is given by

$$R = \bar{x}(n)\bar{x}^T(n) \,. \tag{5.28}$$

Substituting these in the above expression of steepest descent algorithm we get the least mean square (LMS) algorithm,

$$\begin{aligned} \bar{\omega}_{n+1} &= \bar{\omega}_n + \mu[P - R\bar{W}] \,, \\ &= \bar{\omega}_n + \mu\bar{x}(n)[d(n) - \bar{\omega}_T(n)\bar{x}(n)] \,, \\ &= \bar{\omega}_n + \mu\bar{x}(n)e(n) \,. \end{aligned} \tag{5.29}$$

The LMS algorithm, proposed by Widrow and Hoff in 1959, is a stochastic gradient algorithm (Haykin, 2008) and is still one of the most extensively used adaptive filtering algorithms. Although it may seem simple computationally, its mathematical analysis is quite complicated due to its stochastic and nonlinear nature. In general, all the filter coefficients are initially set equal to zero and the filter output $y(n)$ is given by $\omega^T x(n)$. The error signal i.e. output estimate is given by $e(n) = d(n) - y(n)$. The algorithm does not converge to exact Weiner solution in the absolute sense but after finite number of iterations the mean DC value of the filter coefficients converge to the optimal Wiener solution.

Advantages of LMS Algorithm as compared to Wiener filter

- Simplicity, Robustness, Computationally less expensive
- No matrix inversion is required
- No need to measure correlation functions
- Tracks changes in signal statistics efficiently and continuously adapt the filter

Stability Criterion Since LMS algorithm involves feedback, stability can be an issue. Stability criterion can be stated by imposing that the mean square error term J_n attains a constant value as the number of iterations, n approaches infinity. An algorithm that satisfies this is said to be stable in the mean square and this, in turn, constrains the step-size parameter which is related to the spectral content of its tap inputs,

$$0 < \mu < \frac{2}{\lambda_{max}}, \tag{5.30}$$

λ_{max} is the largest Eigenvalue of the correlation matrix $\bar{R}$.

We now repeat the previous simulation of Wiener filter based suspension point noise suppression (see Fig. 5.10) but this time using the ANC algorithm. It is seen to outperform the static Wiener and provides better tracking of fluctuations as is evident from noise cancellation achieved around frequencies 1.25 Hz and lower.

5.8 Effect of Noise Cancellation on Matched Filtering

Glitches are short duration transients which are not of astrophysical origin and a certain class of which can be modeled as Sine-Gaussian (SG) waveforms with

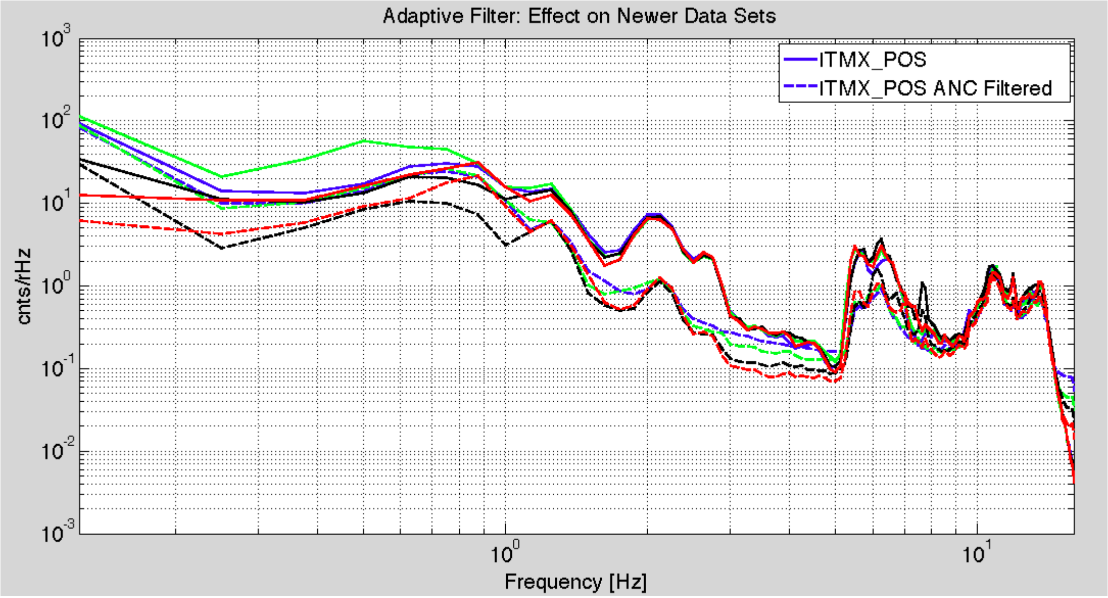


FIGURE 5.13: Application of adaptive filter for noise suppression in suspension point motion. The performance is accessed by applying the filter to thirty minute chunks of data obtained few days later.

a low-quality factor. The technique used to search for signals from compact binaries is matched filtering which involves cross-correlating the detector strain data with the waveform template and thus forming a signal to noise estimate for that waveform. Whenever SNR exceeds a given threshold, it leads to a trigger which is then checked for coincident occurrence in other detectors and then further analysis is carried out. Glitches turned out to be a major issue in first generation detectors since many leads to high SNR triggers which increased the false alarm rate that loss of confidence in detecting actual inspiral signals. The origin of glitches can be numerous like electromagnetic fluctuations, weather variations, seismic events, control system related issues, scattered light, etc. and they are discussed in detail in Chapter 7. In the time domain, SG transients can be modeled as,

$$s(t) = A\, exp(-\frac{(t-t_o)^2}{\tau^2})\cos(2\pi f_o t + \phi_o)\,, \tag{5.31}$$

where A gives the overall amplitude, t_o and f_o, the location of the sine-Gaussian in time and frequency, ϕ_o is the oscillation phase at t = 0 and $\tau = Q/2\pi f_o$. Here Q is the dimensionless quality factor.

Here we show a simple example to demonstrate how such noise cancellation techniques mentioned above could help the matched filtering analysis. For a signal buried in stationary Gaussian noise, matched filtering will be the optimal filtering technique. If the data stream is s(t) = h(t) + n(t) where h(t) is the signal and n(t) is the detector noise with one-sided power spectral density, $S_n(f)$ defined as $< \tilde{n}(f)\tilde{n}^*(f') >= \frac{1}{2}S_n(|f|)\delta(f-f')$ then the matched filter output that computes the arrival times accurately, is given by (Allen et al., 2012),

$$x(t) = 4\, Re \int_0^\infty \frac{\tilde{s}(f)\, \tilde{h}^*_{template}(f)}{S_n(f)}\, e^{2\pi i f t}\, df\,. \tag{5.32}$$

The phase information can be captured by calculating the quadrature sum of the matched filter output in phase ($\phi_o = 0$) and an orthogonal phase ($\phi = \pi/2$). The complex matched filter output z(t) provides the answer to this issue,

$$z(t) = x(t) + iy(t) = 4 \int_0^\infty \frac{\tilde{s}(f)\, \tilde{h}^*_{template}(f)}{S_n(f)}\, e^{2\pi i f t}\, df\,. \tag{5.33}$$

$|z(t)|^2$ gives the quadrature sum of two orthogonal matched filters and $y(t)$ gives the matched filter output when the template used is given by: $\tilde{h}_{template,\phi_0\rightarrow\phi_o-\pi/2}(f) = \tilde{h}_{template}(f)e^{i\pi/2} = i\tilde{h}_{template}(f)$.

As per convention, the waveform templates are generated for systems at an effective distance of $D_{eff} = 1Mpc$. Before computing SNR, a normalization constant for the template is calculated:

$$\sigma_m^2 = 4 \int_0^\infty \frac{|\tilde{h}_{1Mpc,m(f)}|^2}{S_n(f)} df\,. \tag{5.34}$$

When the signal is just stationary and random Gaussian noise, $< x_m(t) >=< y_m(t) >= 0$ and $< x_m^2(t) >=< y_m^2(t) >= \sigma_m^2$. The amplitude of the signal to noise ratio of the quadrature matched filter is given as :

$$\rho_m(t) = \frac{|z_m(t)|}{\sigma_m}\,. \tag{5.35}$$

Our idea was to see how stable the LMS active noise cancellation codes were to data having sine-Gaussian glitches. The simulation included injecting sinusoidal signal and a gravitational wave signal (injected at t = 95 s) into seismic data with glitches injected at random and some even within the signal. Matched filtering was carried out on the 120s long data set to see for triggers before and after active noise cancellation.

Results, shown in Fig. 5.14 and Fig. 5.15, indicate that there is a significant reduction in SNR triggers after active filtering, and the GW signal is satisfactorily recovered with almost similar SNR as before. The significance of this process is made even more clear when we inject very high amplitude glitches 5-6 times the background noise. Before noise cancellation, matched filtering

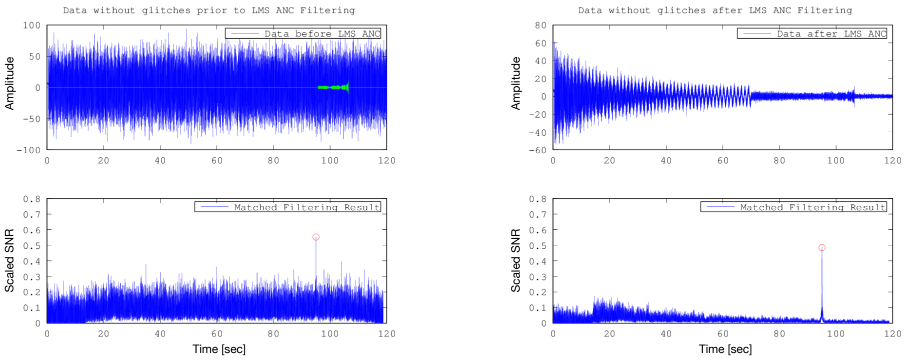


FIGURE 5.14: Results on detector strain without glitches before after least mean square adaptive noise cancellation

yields many high SNR events that can be misinterpreted as gravitational wave signal exactly where glitches are injected. These false triggers are removed, and the GW signal is recovered after LMS ANC.

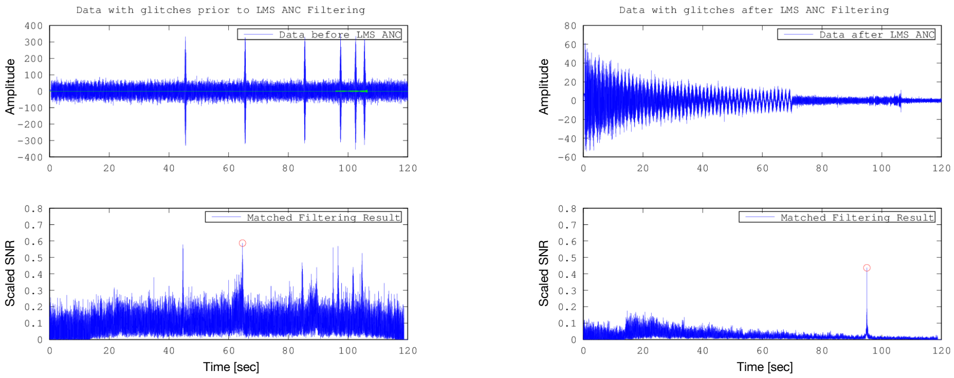


FIGURE 5.15: Results on detector strain with glitches before after least mean square adaptive noise cancellation

5.9 ANC based on Sparseness of Impulse Response

Adaptive Noise Cancellation that was previously implemented at LIGO (Driggers et al., 2012) relies on NLMS algorithm gives equal weightage to all filter coefficients. This and other traditional linear update based algorithms do not take into account the sparseness of the impulse response that they try to identify. The sparseness measure (ζ) is given by

$$\zeta(n) = \frac{L}{L - \sqrt{L}} \left[1 - \frac{\|h(n)\|_1}{\sqrt{L}\|h(n)\|_2} \right],$$

where L is the filter length and h(n) is the impulse response. Sparse impulse response will have a large fraction of its energy contained within a short duration implying that most components will have close to zero magnitudes with a very few having significantly higher magnitude. NLMS widely used in signal processing due to its simplicity and robustness. But its shown to have slow convergence and tracking capabilities compared to more advanced ANC algorithms.

One strategy (named PNLMS) can be used to give proportionate adaptive weightage to filter coefficients meaning that higher magnitude coefficients will receive higher increment leading to an increased convergence of that coefficient. Although PNLMS gives better convergence for sparse systems, the performance is worse than NLMS for non-sparse impulse responses.

5.9.1 Tracking the Variation in Impulse Response

IPNLMS (Improved Proportionate NLMS Algorithm) gives a better balance between NLMS and PNLMS by acting like NLMS when impulse response is not sparse and giving proportionate weightage for sparse systems (Huang, Benesty, and Chen, 2006). It is always found to track and converge better than NLMS and PNLMS irrespective of the sparseness involved. The filter update equation for IPNLMS is given as

$$\begin{aligned}\overline{W}(n) &= \overline{W}(n-1) + \mu(k) g_{ip,l}(k-1)\overline{x}(k-1)e(k)\\ \mu(k) &= \frac{\alpha}{\sum_{j=0}^{L-1} x^2(n-j) g_{ip,j}(n-1) + \delta_{IPNLMS}} \quad .\\ g_{ip,l}(n-1) &= \frac{1-\beta}{2L} + (1+\beta)\frac{|\hat{h}_l(n-1)|}{2\|\hat{h}(n-1)\|_l + \epsilon}\end{aligned} \tag{5.36}$$

where $0 < \alpha < 1$, $-1 \leq \beta < 1$, $\delta_{IPNLMS} \propto \sigma_x^2 \frac{1-\beta}{2L}$ and $\epsilon > 0$.

Simulations were performed on a data set of 20 seconds long where sparseness measure ζ of the input impulse response was changed from 0.85 to 0.53 after the first 10 seconds as shown in Fig. 5.16. The resulting tracking and convergence capabilities of all the three algorithms are plotted given in Fig. Here we assumed that filter order to be 512 for both input and simulated filter. The impulse response of the dynamical system can be sensitive factors like temperature which necessitates the need for the ANC algorithm to track impulse response variations in short time scales. Our simulations show that IPNLMS is better equipped to track those variations compared to the traditionally used

NLMS algorithm. Also note that most acoustic impulse responses are sparse in general. So IPNLMS is better suited for channels like LIGO PEM microphone monitors. The performance is estimated in terms of Normalized Misalignment given by

$$\mathit{Normalized\ Misalignment\ (in\ dB)} = 20\, log_{10} \frac{\|\mathbf{h} - \hat{\mathbf{h}}(k)\|_2}{\|\mathbf{h}\|_2}, \tag{5.37}$$

where $\mathbf{h}$ is the actual impulse response and $\hat{\mathbf{h}}(k)$ is the estimated one at k^{th} instant of time. Note that using a filter of length smaller than the actual impulse

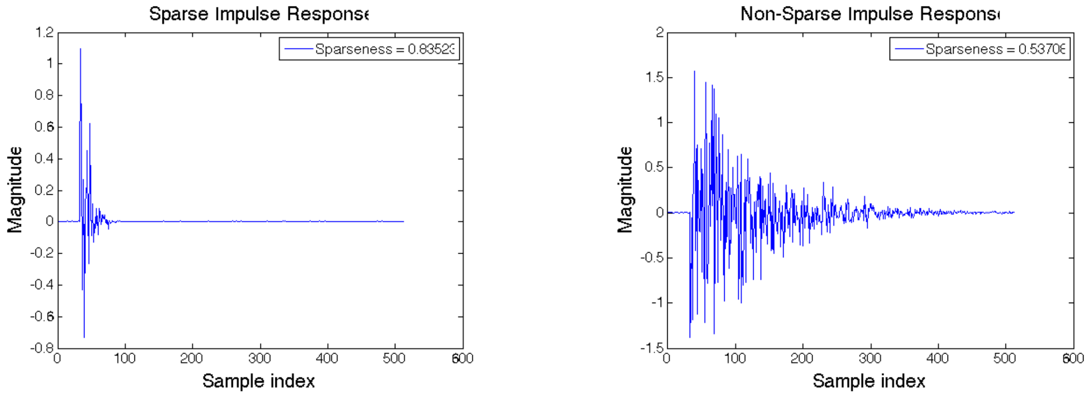


FIGURE 5.16: Impulse response with high and low levels of sparseness.

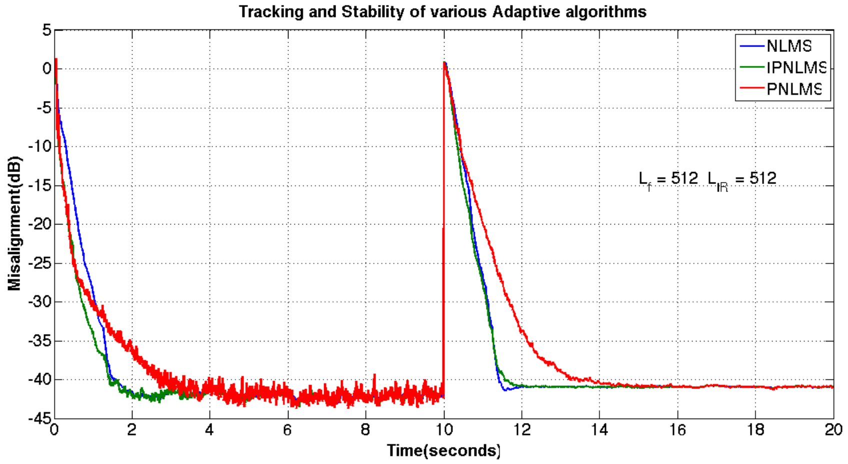


FIGURE 5.17: Comparison of various adaptive filtering algorithms regarding their effectiveness to capture variation in the impulse response. Best performance is provided by Improved Proportionate Normalized LMS algorithm.

response leads to higher Mean Square Error (MSE), but the initial convergence rate is considerably higher for low filter length case.

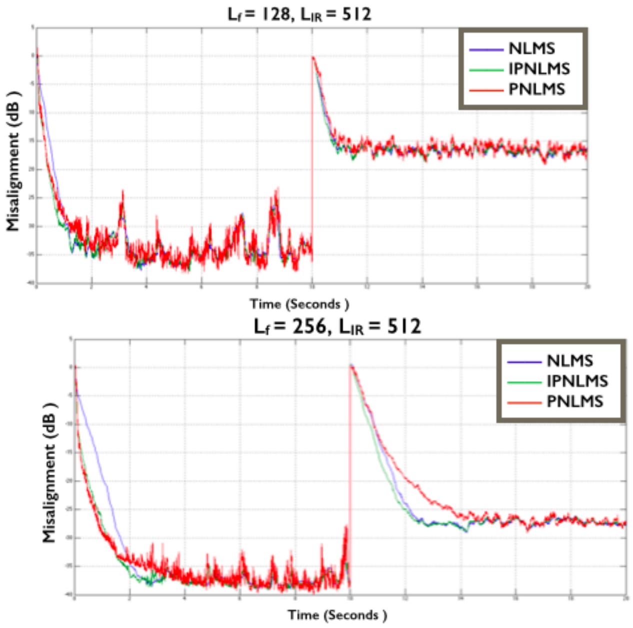
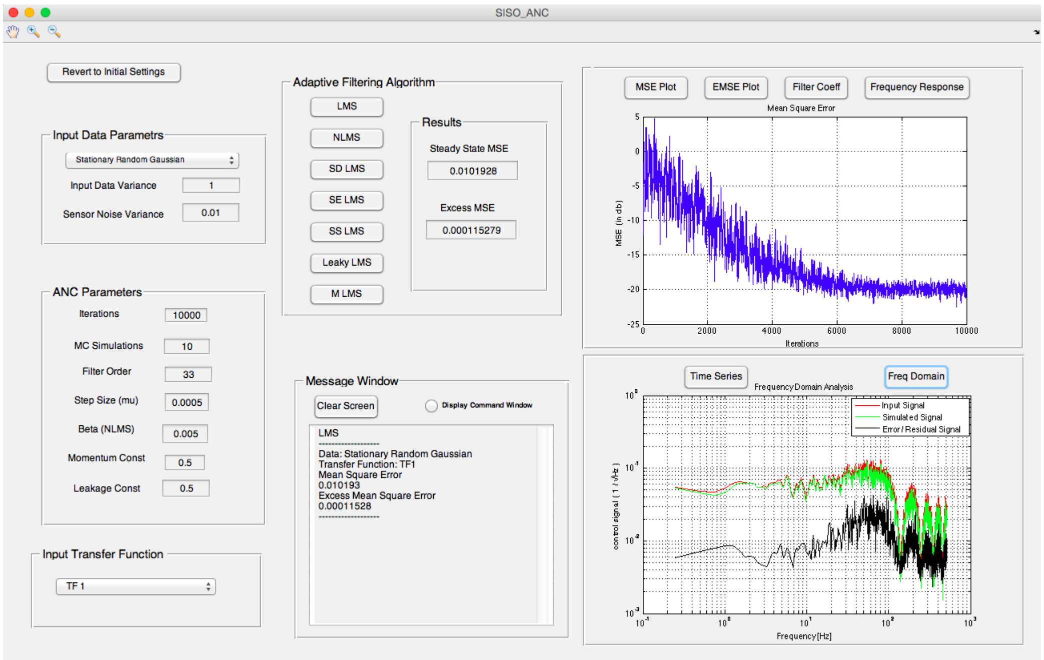


FIGURE 5.18: Comparison of the ability to track the variations in impulse response as function of filter coefficients.

Fig. 5.19 shows a MATLAB Graphical User Interface (GUI) that we developed to access the performance of various adaptive algorithms. The GUI in its current version can be used for Single Input Single Output (SISO) systems and can take into account sensor noise within the channel. We studied the convergence and tracking ability of various LMS algorithms such as Normalized LMS and Leaky LMS along with hardware friendly algorithms like signed LMS. With simulated data, we could reach up to sensor noise level using an appropriate filter order and an optimal learning rate.

5.10 Conclusions and Future Work

The techniques involved in system identification based on measured transfer functions were discussed, and their application was demonstrated through feedforward noise suppression in Advanced LIGO seismic isolation platforms. Automated ways to model the system dynamics can save time and lead to the desired level of control in the desired frequency band without causing out of band noise re-injection. We did a comparative study of static Wiener and adaptive filtering algorithms through its application to simulated data as well as the data obtain from LIGO seismic sensor channels. Further, we accessed the impact of adaptive filters on matched filter analysis especially when the input data had been significantly corrupted by correlated noise sources. In future, we would like to study the hardware feasibility of the various adaptive algorithms



Input Data Parametrs
Stationary Random Gaussian
Input Data Variance: 1
Sensor Noise Variance: 0.01

ANC Parameters
Iterations: 10000
MC Simulations: 10
Filter Order: 33
Step Size (mu): 0.0005
Beta (NLMS): 0.005
Momentum Const: 0.5
Leakage Const: 0.5

Input Transfer Function
TF 1

Results
Steady State MSE: 0.0101928
Excess MSE: 0.000115279

```
LMS
-------------------
Data: Stationary Random Gaussian
Transfer Function: TF1
Mean Square Error
0.010193
Excess Mean Square Error
0.00011528
-------------------
```

FIGURE 5.19: MATLAB GUI developed to test the performance of various adaptive filtering algorithms.

using field programmable gate arrays (FPGAs) and compare the performance of IIR and FIR filters for real-time applications.

Chapter 6

Lockloss Monitoring and Prediction for Advanced LIGO

6.1 Locking Acquisition

Although in general, a Fabry-Perot cavity can be locked to a stabilized laser using the well-known Pound-Drever-Hall (PDH) technique, for Advanced LIGO which is a Dual-Recycled Fabry-Perot Michelson Interferometer, cavity resonance is achieved through a more elaborate hierarchical control strategy. The main error signals needed for this control are typically obtained through the reflection signals from the full interferometer, the anti-symmetric port and the reflection signal from the beam splitter. It is required to bring all the five degrees of freedom (common mode, differential mode, Michelson length, power recycling length, and signal recycling length) to a linear operating point. One major issue comes from the fact that these errors signals have a very narrow linear regime and using them outside of this limit would lead to misleading results. Part of this nonlinearity could be associated with the inclusion of recycling cavities within the Fabry-Perot Michelson interferometer. Initial step involved in locking is to lock the arms cavities using the end station green laser (532 nm) through the arm length stabilization system. The frequency of the end station lasers is offset such that the cavities are non-resonant for the main infrared laser (1064 nm). The dual recycled Michelson interferometer (DRMI) is then locked to the main laser using the third harmonic (3f) technique. 3f signal is generated mainly due to the beat between the RF sidebands and its second harmonic both of which unlike the carrier are non-resonant within the arm cavities. Once the DRMI is locked, the previously introduced frequency offset is decreased to zero so that the main laser is resonant in the arm cavities finally leading to a full lock of the interferometer.

Before locking, it is also necessary to stabilize the sub-systems from various kinds of noises and instabilities. This includes frequency & intensity stabilization of laser, active stabilization of seismic platforms, violin mode damping of suspension, beam spot centering, minimizing cross-coupling with the alignment and length sensing to name a few. The presence of such multitude of control loops running within the interferometer also makes it susceptible to instabilities or glitches that can occur within any of these sub-systems.

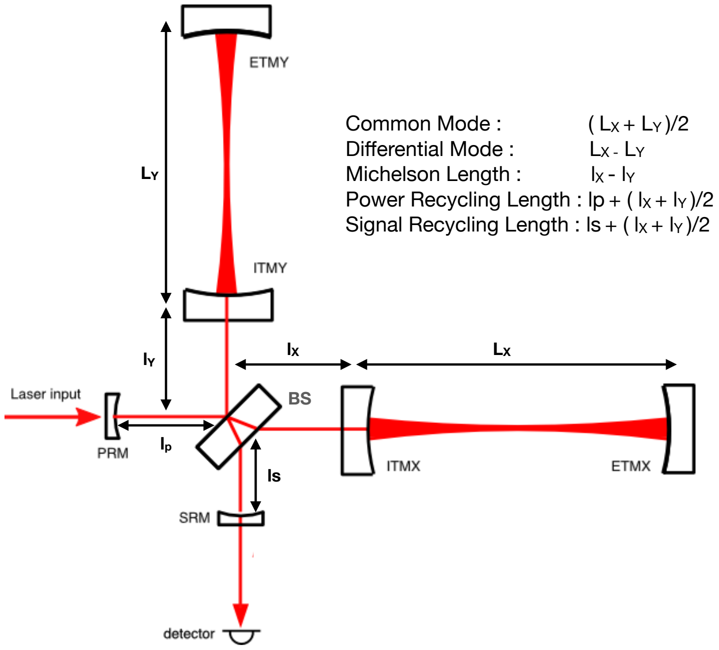


FIGURE 6.1: Simplified optical layout of the interferometer showing the five degrees of freedom.

6.2 Probing lockloss

Often the desired controller is designed by hand such that desired suppression is obtained in the band of interest, but often this also leads to undesired out-of-band gain peaking. Hence when such unforeseen excess noise is injected into the loop, it drives the system into instability leading to a loss of power within the arm cavities resulting in a state of lockloss. It is undesirable, as typically it takes tens of minutes to regain lock thus losing coincident observation time. The inter-relatedness of the various subsystems and availability of tens of thousands of data channels to probe the issue makes lockloss debugging a computationally intensive process. We are also faced with the problem of decoupling the cause and effect when probing through the different LIGO data channels.

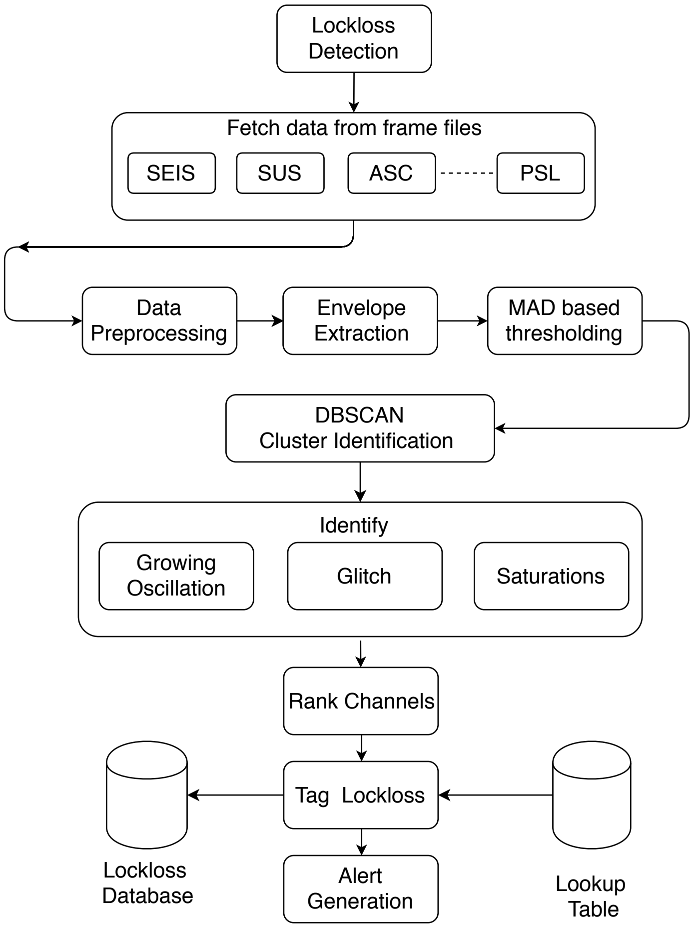


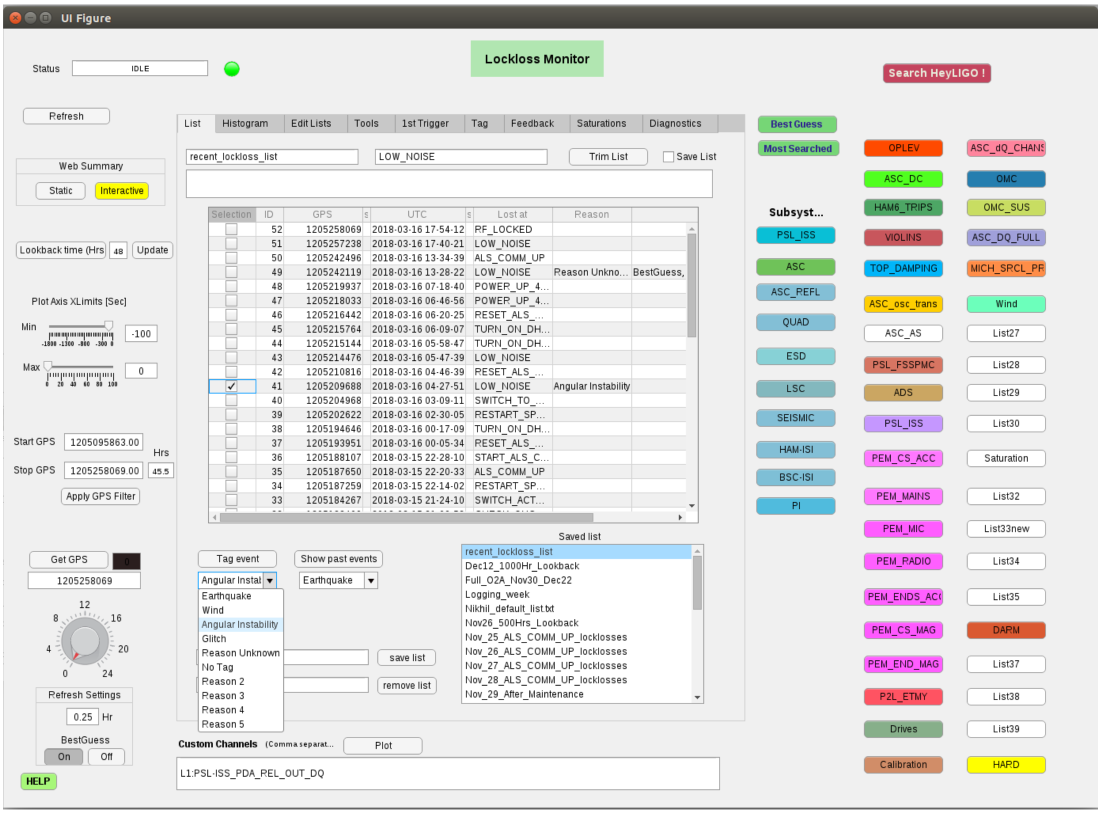
FIGURE 6.2: Algorithm used for real-time tagging of lockloss events.

In this chapter, we present the development of an automated near real-time lockloss monitor that uses techniques from machine learning and attempts to diagnose the cause of lockloss by looking across the various subsystem channels. This low latency initial information provided by the tool could be further used by the commissioners & subsystem engineers to probe the origin of the lockloss.

We now describe the basic algorithm behind the auto-tagging feature of lockloss monitor given in Fig. 6.2. The interface constantly checks for new lockloss events by searching the Guardian logs. Guardian (Rollins, 2016) is a distributed state machine supervision control software that continuously monitors the states of various subsystems within the interferometer. Once we detect a lockloss, the exact time of the event is obtained by checking for a sudden drop in cavity power from a pickup photo-diode. We then load 100 seconds of data before the lockloss for hundreds of subsystem channels that are carefully selected based on inputs from the commissioners and subsystem engineers. Channel-specific preprocessing is carried out in the form of decimation, and bandpass/high-pass filtering to respectively decrease the data volume and to boost any transients. After this step, we look for spurious events in each data channel by looking for regions in the data that exceed a threshold (usually few factors of median absolute deviation above or below the median). DBSCAN (Ester et al., 1996) algorithm is then utilized to identify clusters along with outliers in an automated fashion. Each channel is ranked based on the first occurrence of such clusters. We then try to look for patterns within this ranked channel-list that point towards a specific issue within a subsystem. A comparison is done using a lookup table that is created based on past observations of such similar locklosses and based on the similarity between the sets of channels a tag is issued. Fig. 6.6 and Fig. 6.7 show the results from analyzing two different kinds of locklosses. The first one triggered by a saturation happening with the electrostatic drive actuator and the second one due to instabilities occurring with the length and alignment control signals.

6.2.1 DBSCAN

Here we elaborate on DBSCAN, the core algorithm used in this analysis and motivate why it was chosen. Machine learning algorithms always encounter difficulties when dealing with outliers. And this is especially true for classification based algorithms such as deep neural nets as they have been mostly trained on datasets that belong to specific categories. Even they encounter data that have features different from those used in training or having characteristics in between say, two classes, on most occasions they fail to identify its uniqueness and tries to classify it into the pre-learned groups with different levels of confidence. Also for a complex dynamical system like Advanced LIGO, it is impractical to identify all possible scenarios of unusual channel behavior and perform training. DBSCAN, which stands for Density-Based Spatial Clustering of Applications with Noise, does not require the user to provide in advance the number of expected clusters in the data and so is different from traditional clustering algorithms such as K-Means clustering (Hartigan and Wong, 1979). The minimum number of points in the cluster (MinPts) and the maximum distance between two samples for them to be considered as part of the same cluster (EPS) are the two input parameters that need to be provided to identify the clusters in the data. If a region has more than or equal to the specified number of points within a maximum distance provided by the user, it is labeled as a unique cluster.



Status: IDLE

Lookback time (Hrs): 48

Min: -100
Max: 0

Start GPS: 1205095863.00
Stop GPS: 1205258069.00
Hrs: 45.5

Get GPS: 1205258069

Refresh Settings: 0.25 Hr

recent_lockloss_list
LOW_NOISE

Selection	ID	GPS	s	UTC	s	Lost at	Reason	
☐	52	1205258069		2018-03-16 17-54-12		RF_LOCKED		
☐	51	1205257238		2018-03-16 17-40-21		LOW_NOISE		
☐	50	1205242496		2018-03-16 13-34-39		ALS_COMM_UP		
☐	49	1205242119		2018-03-16 13-28-22		LOW_NOISE	Reason Unkno…	BestGuess,
☐	48	1205219937		2018-03-16 07-18-40		POWER_UP_4…		
☐	47	1205218033		2018-03-16 06-46-56		POWER_UP_4…		
☐	46	1205216442		2018-03-16 06-20-25		RESET_ALS_…		
☐	45	1205215764		2018-03-16 06-09-07		TURN_ON_DH…		
☐	44	1205215144		2018-03-16 05-58-47		TURN_ON_DH…		
☐	43	1205214476		2018-03-16 05-47-39		LOW_NOISE		
☐	42	1205210816		2018-03-16 04-46-39		RESET_ALS_…		
☑	41	1205209688		2018-03-16 04-27-51		LOW_NOISE	Angular Instability	
☐	40	1205204968		2018-03-16 03-09-11		SWITCH_TO_…		
☐	39	1205202622		2018-03-16 02-30-05		RESTART_SP…		
☐	38	1205194646		2018-03-16 00-17-09		TURN_ON_DH…		
☐	37	1205193951		2018-03-16 00-05-34		RESET_ALS_…		
☐	36	1205188107		2018-03-15 22-28-10		START_ALS_C…		
☐	35	1205187650		2018-03-15 22-20-33		ALS_COMM_UP		
☐	34	1205187259		2018-03-15 22-14-02		RESTART_SP…		
☐	33	1205184267		2018-03-15 21-24-10		SWITCH_ACT…		

Angular Instal
- Earthquake
- Wind
- Angular Instability
- Glitch
- Reason Unknown
- No Tag
- Reason 2
- Reason 3
- Reason 4
- Reason 5

Earthquake

Saved list:
- recent_lockloss_list
- Dec12_1000Hr_Lookback
- Full_O2A_Nov30_Dec22
- Logging_week
- Nikhil_default_list.txt
- Nov26_500Hrs_Lookback
- Nov_25_ALS_COMM_UP_locklosses
- Nov_26_ALS_COMM_UP_locklosses
- Nov_27_ALS_COMM_UP_locklosses
- Nov_28_ALS_COMM_UP_locklosses
- Nov_29_After_Maintenance

Custom Channels: L1:PSL-ISS_PDA_PEL_OUT_DQ

FIGURE 6.3: Lockloss Monitor GUI installed at LLO. Application continuously monitoring the state of the instrument and in the event of a lockloss, initiates an ML-based analysis looking for abnormal behavior in the various subsystem channels few tens of seconds before the lockloss. It then sorts the list of channels so detected and issues a tag indicating the plausible reason for the lockloss. Interactive lockloss results plots are generated for each event which can be accessed from the web by anyone with valid LIGO credentials. It can also look for saturations and can be used as a lockloss debugging tool for each of the individual subsystems.

6.2.2 Tracking the user search patterns

In this section, we describe an alternative approach to debug the lockloss events. It turns out that the subject expertize or specific domain knowledge is something that would be hard for any ML algorithm to emulate at least with the present technologies. But we can always learn from how an experienced person probes a lockloss event. Often his or her experience based intuition and subsequent search for features in different channels results in finding the right reason for the lockloss event. To take advantage of this, we added a provision where users can do investigation with the lockloss monitor application and if successful add their conclusions for the given lockloss. The patterns associated with the event are analyzed based on inputs from each user and for any new lockloss event weighted average results based on previous user feedback can

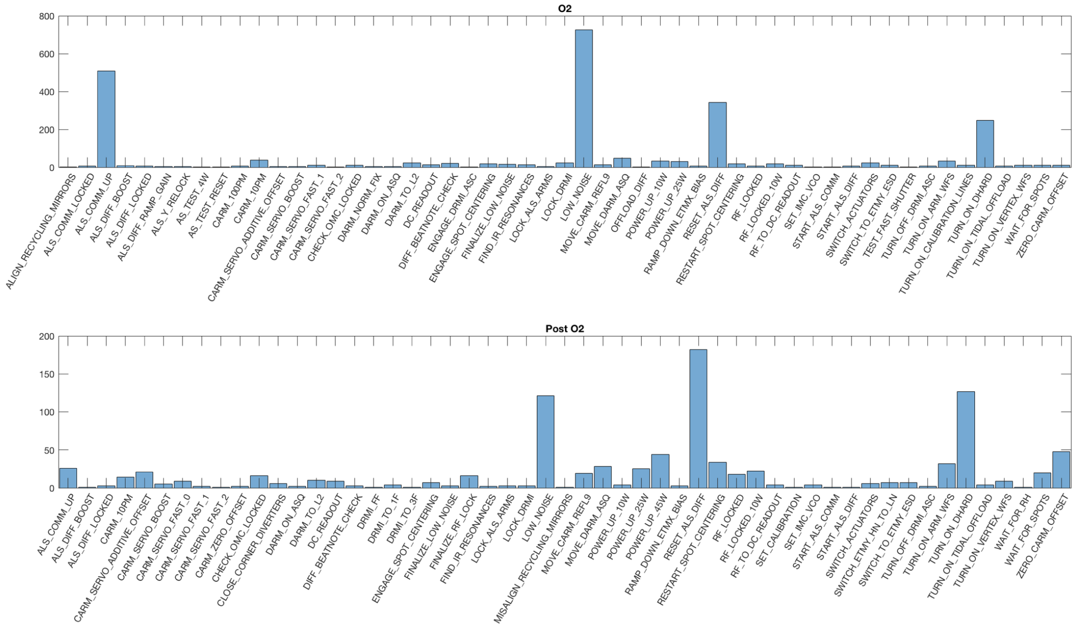


FIGURE 6.4: Comparison of the frequency of locklosses that happened at various stages in the locking acquisition.

be utilized to come up with a list of channels that behaved oddly during the event.

6.3 Analysis of O2 Unknown LLO Locklosses

Fig. 6.4 shows the distribution of LLO locklosses across multiple stages of locking procedure. During the science run period (O2), the goal would be to maintain a long stretch of a lock with minimal configuration changes so that we obtain good quality coincident data. Large variations in the PSD is also undesirable as it can effect the parameter estimation analysis. Lockloss often occurs at multiple other intermediate stages also where various configuration changes are carried out in an attempt to improve the sensitivity.

Results from analyzing 76 unknown locklosses that happened during O2 are reported below[1](Mukund, 2017b). These events were identified by scanning through operator logbook entries written during the period and also using the data collected by the Lockloss App. We analyze 100 seconds of data prior to lockloss and generated the ranked list of channels. Table 6.2 lists all

---

[1] Lockloss plots for all the unknown O2 locklosses can be accessed at: *https://llocds.ligo-la.caltech.edu/data/nikhil.mukund/LOCKLOSS_App_SpecificAnalysis/Plots/Unknowns/*

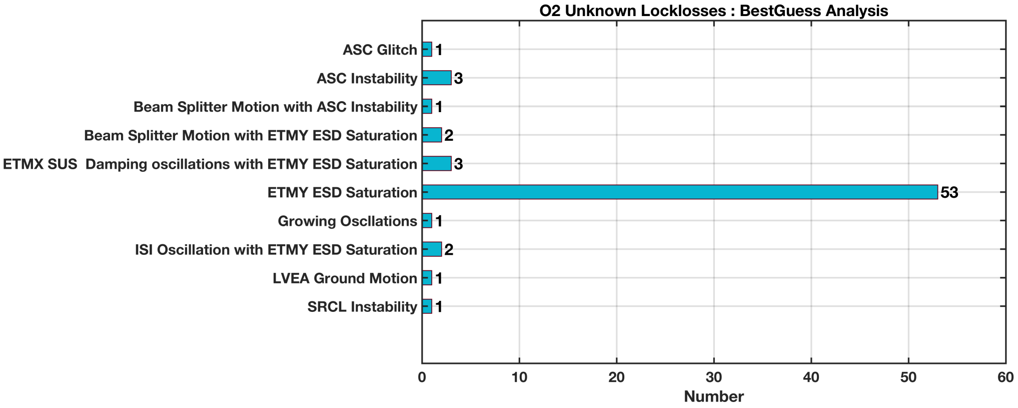


FIGURE 6.5: Distribution of unknown locklosses observed at LLO during O2

these events along with their probable cause. Majority of them show association with ETMY ESD saturation, growing oscillations in ASC, LVEA ground motion and oscillations in BS.

Often to minimize the noise arising while converting control signals from digital to analog domain (DAC noise), data is digitally low passed around a few Hertz, and then the filter is inverted in the analog domain. If the low pass filter has a suppression factor, then we would have to compensate for it through a higher gain analog filter. So if a high-frequency transient event occurs in the DAC, say due to bit quantization error, then it could get amplified to a value that is high enough to saturate the electrostatic drive (ESD) actuator. As for the unknown O2 locklosses, we found 60 of these 76 to be triggered by such ETMY ESD saturation. Of these, around 23 events show the presence of a strong glitch in various channels 0.1-1.5 seconds before lockloss. The transient is primarily seen in the channel used in ETMY L2/L3 stage actuation. It is also witnessed by ETMY SUS ESD monitoring channels at around the same time. This signal could be the one that triggers the ESD saturation at a later stage leading to the lockloss.

6.4 Predicting earthquake related locklosses

The main benefit of ground velocity predictions for gravitational-wave detectors described in Chapter 3 is to inform predictions of whether an earthquake will cause the loss of data for the detector. The Rayleigh component within the seismic wave is the most damaging for precision measurements and often generates sustain the ground motion of few tens of microns forcing the interferometer to lose lock as shown in Fig. 6.9. The exact mechanism that triggers

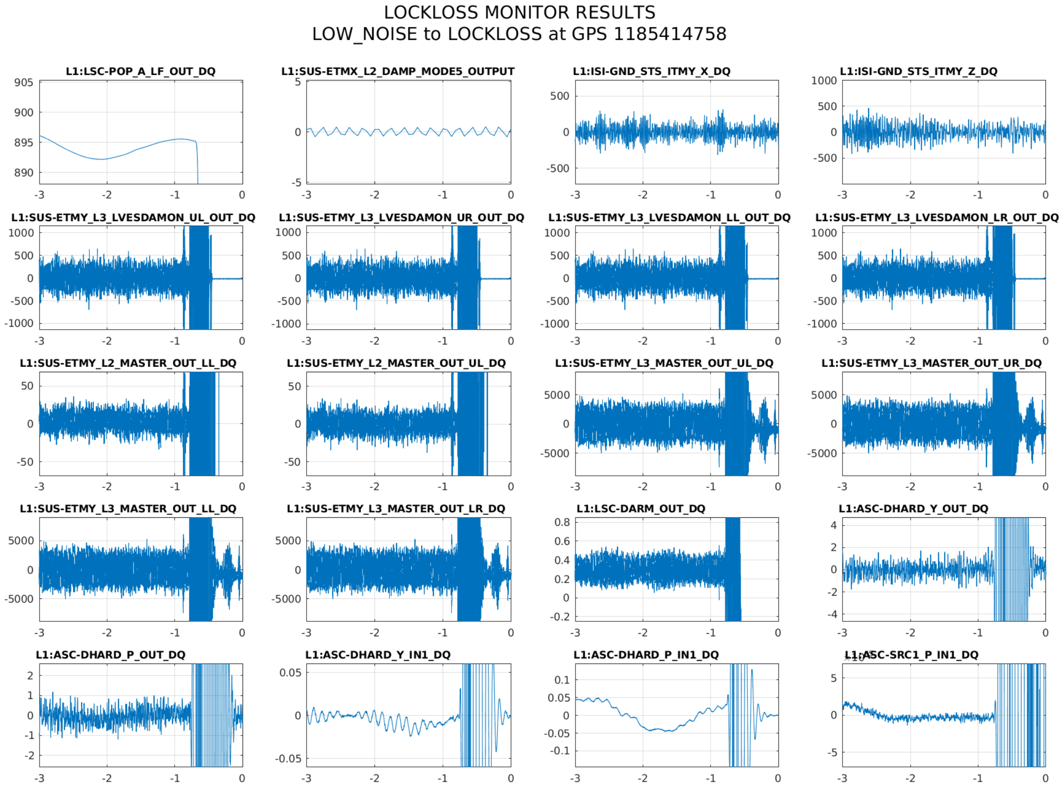


FIGURE 6.6: Lock loss due to ESD saturation. Presence of a strong glitch few milliseconds prior to the saturation is clearly seen in the ETMY suspension L2 stage channels. These plots are automatically generated by the lockloss monitor with a latency as low as five minutes.

lockloss during periods of excess ground motion is still not fully understood, but techniques such as the one presented here provide certain clues to how it affects the interferometer. Signals seen in the various control channels seem to be within the actuation range, so we would expect the system to stay locked. We have observed a ringing in the suspension violin modes during some of the earthquake related locklosses. But in general, it is preceded by growing instabilities in the control channels related to internal seismic isolation (ISI) and alignment sensing & control (ASC) subsystems as shown in Fig. 6.8. We have previously developed techniques for preventing earthquakes from causing the loss of data taking if advanced notice is given (Biscans et al., 2018). In the following section, we will use a machine learning algorithm to develop a lockloss prediction model as well. We use the same set of inputs to the algorithm as in the ground velocity prediction case, but also include the ground velocity predictions themselves as inputs. To generate the target variable, we take times when the gravitational-wave detectors lost the ability to take data during an earthquake and assign a value of 1, and a 0 otherwise. We use the same

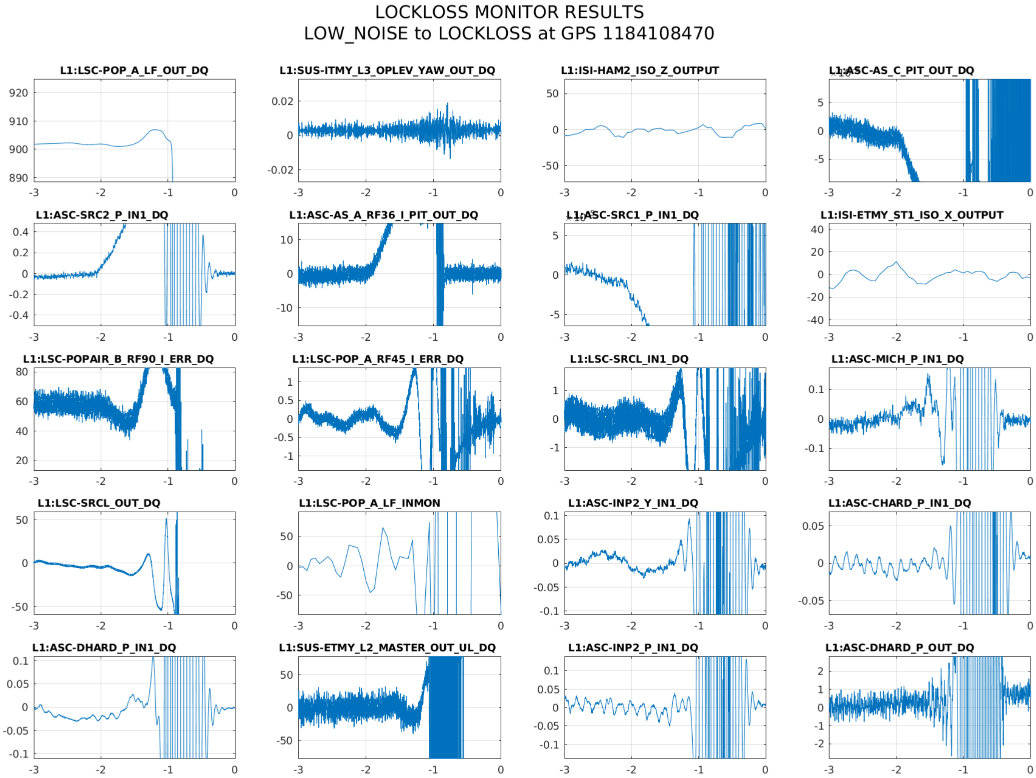


FIGURE 6.7: Lock loss due to instabilities in length and alignment control signals. Lockloss Monitor is trained to issue to an "Angular Instability" flag for these kindof locklosses.

clustering based algorithm as applied for the ground velocity predictions and allow for outliers detection. Acknowledging that there is a trade-off between false-alarm probability and efficiency standard, We are able to make predictions for the inliers with an accuracy above 92 %, keeping the associated false-alarm probability to be less than 10 % (see Table 6.1). Traditionally the seismic isolation at LIGO was designed to provide isolation above 100 mHz and had a significant gain at the lower frequencies making it vulnerable to earthquakes. This happens due to limitations in our filter design, which is often constructed by hand and leads to noise injection outside the band of interest. Since regaining interferometric lock takes tens of minutes, its preferable to stay locked albeit at a lower sensitivity. Thus if we have prior information about the earthquake and corresponding prediction about an impending lockloss, it would be more desirable to switch to a configuration that provides enough freedom for common mode motion but at the same time enforces the best possible suppression for the local differential motion.

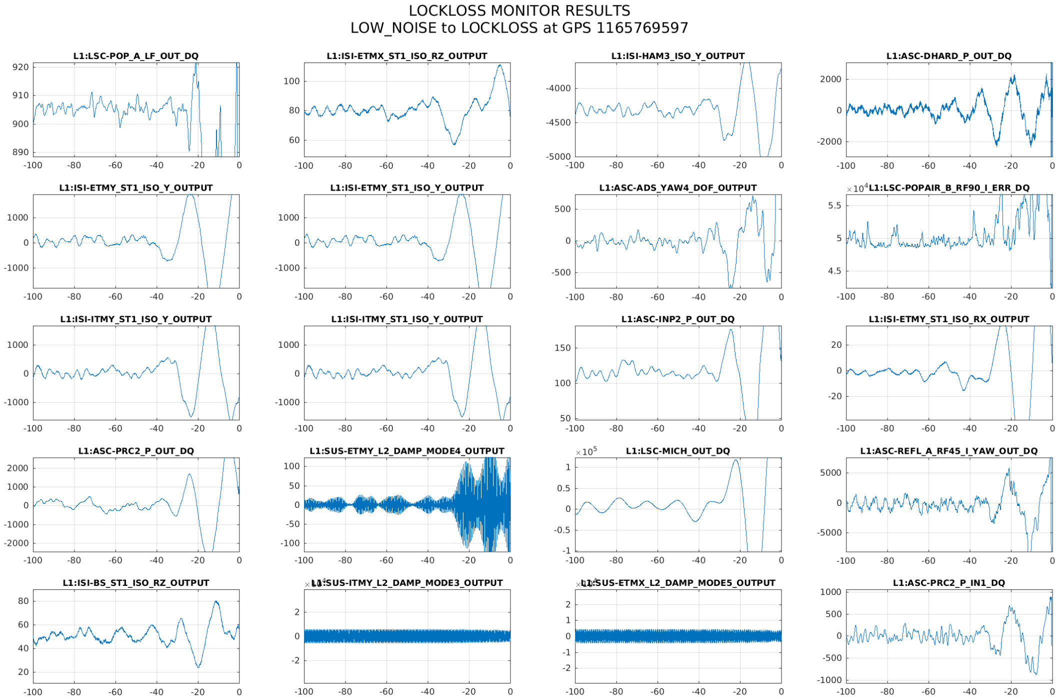


FIGURE 6.8: Lock loss due to excessive ground motion caused by an earthquake. Ringing up of suspension violin modes can be seen from the second plot on the third row.

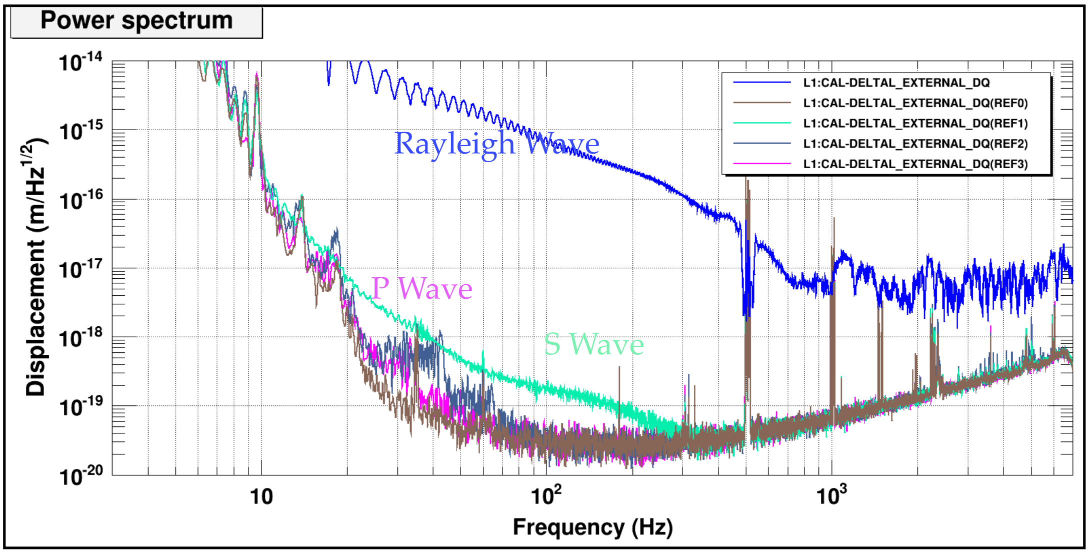


FIGURE 6.9: The effect of P, S and R seismic waves on a locked interferometer. Surface Rayleigh generates sustained ground motion measuring few tens of microns thus forcing the interferometer to lose lock.

GPS	Reason	Comments	Related Channels
1182989555	ASC Instability	Oscillations in ASC MICH	L1:ASC-ADS_PIT4_DOF_OUTOUT, L1:ASC-REFL_A_RF45_I_YAW_OUT_DQ

1182992674	ETMY ESD Saturation	Glitch seen in ETMY L3 LVESDAMON 0.5 seconds prior to lockloss	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ
1183000957	ETMY ESD Saturation	Glitch seen in ETMY L3 LVESDAMON 0.2 seconds prior to lockloss	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ
1183093432	ETMY ESD Saturation	Glitch seen in ETMY L3 LVESDAMON 0.3 seconds prior to lockloss	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ
1183194813	ETMY ESD Saturation	Growing oscillation in ETMX SUS L2 DAMP, Glitch in ETMY SUS L2	L1:SUS-ETMX_L2_DAMP_MODE5_OUTPUT, L1:SUS-ETMY_L2_MASTER_OUT_DQ
1183273720	ETMY ESD Saturation	Growing oscillation in ETMX SUS L2 DAMP MODE5	L1:SUS-ETMX_L2_DAMP_MODE5_OUTPUT
1183418522	ETMY ESD Saturation	No glitch seen prior to saturation	
1183491395	ETMY ESD Saturation	No glitch seen prior to saturation	
1183495156	ETMY ESD Saturation	No glitch seen prior to saturation	
1183581056	ETMY ESD Saturation	Big glitch seen in ETMY {L3,L2} LVESDAMON 0.3 seconds prior to lockloss	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ
1183587880	ETMY ESD Saturation	No glitch seen prior to saturation	
1183616183	Beam Splitter Motion with ETMY ESD Saturation	Oscillation in Beam Splitter Stage 1	L1:ISO-BS_ST1_{X,Y,RX,RY,RZ}_OUT_DQ
1183618655	ETMY ESD Saturation	Growing oscillation in ITMY SUS Damping channels	L1:SUS-ITMY_L2_DAMP_MODE3_OUTPUT
1183654850	ETMY ESD Saturation	Glitch seen in ETMY L3 1.5 seconds prior to lockloss	L1:SUS-ETMY_L3_MASTER_OUT_{UL,UR,LL,LR}_DQ
1183712934	ETMY ESD Saturation	Growing oscillation in ITMY SUS Damping channels	L1:SUS-ITMY_L2_DAMP_MODE3_OUTPUT
1183910338	SRCL Instability	Signal Recycling Cavity Length signal abruptly changes	L1:LSC-SRCL_IN1_DQ
1184108470	ASC Instability	Growing Oscillation in ASC {MICH,CHARD, DHARD}	L1:ASC-CHARD_P_IN1_DQ, L1:LSC-DHARD_P_OUT_DQ
1184137869	ETMY ESD Saturation	No glitch seen prior to saturation	
1184222703	ETMY ESD Saturation	Big glitch seen in ETMY {L3,L2} 1 seconds prior to lockloss	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:SUS-ETMY_L2_MASTER_OUT_{UL,UR,LL,LR}_DQ
1184243998	ETMY ESD Saturation	Small glitch seen in ETMY {L3,L2} 1.5 seconds prior to lockloss	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:SUS-ETMY_L2_MASTER_OUT_{UL,UR,LL,LR}_DQ
1184362931	ETMY ESD Saturation	Big glitch seen in ETMY L2 0.4 seconds prior to lockloss	L1:SUS-ETMY_L2_MASTER_OUT_{UL,UR,LL,LR}_DQ
1184475187	ETMY ESD Saturation	Big glitch seen in ETMY L2 0.7 seconds prior to lockloss	L1:SUS-ETMY_L2_MASTER_OUT_{UL,UR,LL,LR}_DQ
1184558105	ETMY ESD Saturation	Glitch seen in ETMY L3 LVESDAMON 0.7 seconds prior to lockloss	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ
1184608567	ETMY ESD Saturation	Big glitch seen in ETMY {L3,L2} 0.5 seconds prior to lockloss	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:SUS-ETMY_L2_MASTER_OUT_{UL,UR,LL,LR}_DQ
1184612981	ETMY ESD Saturation	No glitch seen prior to saturation	
1184734470	ETMY ESD Saturation	Multiple glitches seen in ETMY {L3,L2} 0.5v seconds prior to lockloss	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:SUS-ETMY_L2_MASTER_OUT_{UL,UR,LL,LR}_DQ
1184753518	ETMY ESD Saturation	Small glitch seen in SUS ETMY L3 1.5 seconds prior to lockloss	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ
1184796479	ETMY ESD Saturation	No glitch seen prior to saturation	
1184816444	ETMY ESD Saturation	No glitch seen prior to saturation	
1184914222	ETMY ESD Saturation	Big glitch seen in ETMY L3 LVESDAMON 0.15 seconds prior to lockloss	
1184980235	ETMY ESD Saturation	No glitch seen	
1185011198	ETMY ESD Saturation	Small glitch seen in SUS ETMY L3 1 seconds prior to lockloss	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ
1185027836	ASC Instability	Transient in ASC {CHARD, MICH}	L1:ASC-SRCl_{P,Y}_IN1_DQ, L1:ASC-CHARD_{P,Y}_IN1_DQ
1185047720	Growing Oscillations	Multiple subsystems show growing oscillations ∼5 seconds prior to lockloss	L1:ASC-ADS_PIT3_DOF_OUTOUT, L1:LSC-SRCL_IN1_DQ, L1:LSC-POP_A_RF45_I_ERR_DQ
1185078066	Beam Splitter Motion with ASC Instability	Glitch (∼1.3s) in Stage 1 Beam Splitter followed by ASC oscillations	L1:ISI-BS-ST1_ISO_{X,Y,RX,RY}_OUTPUT
1185209259	ETMY ESD Saturation	No glitch seen prior to saturation	
185212384	ETMY ESD Saturation	No glitch seen prior to saturation	
185326281	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD,SRCl} oscillates around saturation	

1185414758	ETMY ESD Saturation	Big glitch seen in ETMY L3 LVESDMON 0.1 seconds prior to lockloss, ASC {DHARD,SRCl} oscillates around saturation	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:ASC-DHARD_P_OUT_DQ
1185454502	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD,SRCl} oscillates around saturation	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:ASC-DHARD_P_OUT_DQ, L1:ASC-SRCl_P_IN1_DQ
1185493492	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD,SRCl} oscillates around saturation	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:ASC-DHARD_P_OUT_DQ, L1:ASC-SRCl_P_IN1_DQ
1185509875	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD,SRCl} oscillates around saturation	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:ASC-DHARD_P_OUT_DQ, L1:ASC-SRCl_P_IN1_DQ
1185632047	LVEA Ground Motion	Growing oscillation in ISI channels close to ITMX, ITMY, HAM2, HAM3	L1:ISI-ITMX_ST1_ISO_RX_OUTPUT, L1:IS1-HAM2_ISO_X_OUTPUT
1185686039	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD,SRCl} oscillates around saturation	L1:ASC-DHARD_P_OUT_DQ, L1:ASC-SRCl_P_IN1_DQ
1185692847	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD} oscillates around saturation	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:ASC-DHARD_P_OUT_DQ
1185703719	HAM Chamber Motion, ESD Saturation	HAM 4 & 6 seem to be moving, ASC {DHARD, SRCl} oscillates around saturation	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:ASC-DHARD_P_OUT_DQ
1185714098	ETMY ESD Saturation	Big glitch seen in ETMY L3 LVESDAMON 0.5 seconds prior to lockloss, ASC {DHARD,SRCl} oscillates around saturation	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:ASC-DHARD_P_OUT_DQ, L1:ASC-SRCl_P_IN1_DQ
1185725684	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD} oscillates around saturation	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:ASC-DHARD_P_OUT_DQ
1185852420	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD} oscillates around saturation	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:ASC-DHARD_P_OUT_DQ
1185877302	ASC Instability	ASC {MICH, DHARD, CHARD} signals starts oscillations prior to lockloss	L1:ASC-MICH_Y_IN1_DQ, L1:ASC-CHARD_P_IN1_DQ
1185909701	ETMY ESD Saturation	Big glitch seen in ETMY L3 LVESDAMON 0.1 seconds prior to lockloss, ASC {DHARD,SRCl} oscillates around saturation	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:ASC-DHARD_P_OUT_DQ
1185936851	ETMY ESD Saturation	Small glitch seen in ETMY L3 LVESDAMON 0.5 seconds prior to lockloss, ASC {DHARD} oscillates around saturation	L1:SUS-ETMY_L3_LVESDAMON_{UL,UR,LL,LR}_OUT_DQ, L1:ASC-DHARD_P_OUT_DQ
1185961205	ETMY ESD Saturation	Big glitch seen in ETMY L3 0.3 seconds prior to lockloss, ASC {DHARD,SRCl} oscillates around saturation	L1:SUS-ETMY_L3_MASTER_OUT_{UL,UR,LL,LR}_DQ, L1:ASC-DHARD_P_OUT_DQ
1185999079	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD} oscillates around saturation	L1:SUS-ETMY_L3_MASTER_OUT_{UL,UR,LL,LR}_DQ, L1:ASC-DHARD_P_OUT_DQ
1186054462	ETMY ESD Saturation	Small glitch seen in ETMY L3 1 second prior to lockloss, ASC {DHARD} oscillates around saturation	L1:SUS-ETMY_L3_MASTER_OUT_{UL,UR,LL,LR}_DQ, L1:ASC-DHARD_P_OUT_DQ
1186061168	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD} oscillates around saturation	L1:SUS-ETMY_L3_MASTER_OUT_{UL,UR,LL,LR}_DQ, L1:ASC-DHARD_P_OUT_DQ
1186095025	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD} oscillates around saturation	L1:SUS-ETMY_L3_MASTER_OUT_{UL,UR,LL,LR}_DQ, L1:ASC-DHARD_P_OUT_DQ
1186107561	ISI Oscillation with ETMY ESD Saturation	Growing oscillation seen in ETMY Stage 1 Isolation loop	L1:ISI-ETMY_ST1_ISO_Y_OUTPUT
1186119410	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD} oscillates around saturation	L1:SUS-ETMY_L3_MASTER_OUT_{UL,UR,LL,LR}_DQ, L1:ASC-DHARD_P_OUT_DQ
1186201223	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD} oscillates around saturation	L1:SUS-ETMY_L3_MASTER_OUT_{UL,UR,LL,LR}_DQ, L1:ASC-DHARD_P_OUT_DQ
1186304059	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD} oscillates around saturation	L1:SUS-ETMY_L3_MASTER_OUT_{UL,UR,LL,LR}_DQ, L1:ASC-DHARD_P_OUT_DQ
1186435959	ISI Oscillation with ETMY ESD Saturation	Growing oscillation/transient seen in HAM2 ISI RY Isolation loop	L1:ISI-HAM2_ISO_RY_OUTPUT, L1:SUS-ITMY_OPLEV_YAW_OUT_DQ
1186460396	ETMY ESD Saturation	No glitch seen prior to saturation, ASC {DHARD} oscillates around saturation	L1:SUS-ETMY_L3_MASTER_OUT_{UL,UR,LL,LR}_DQ, L1:ASC-DHARD_P_OUT_DQ
1186490407	ETMY ESD Saturation	No glitch seen prior to saturation, Suspension damping modes	L1:SUS-ITMY_L2_DAMP_MODE4_OUTPUT
1186561213	ETMY ESD Saturation	No glitch seen prior to saturation, Suspension damping modes	L1:SUS-ETMX_L2_DAMP_MODE5_OUTPUT
1186621179	ETMY ESD Saturation	No glitch seen prior to saturation, Suspension damping modes	L1:SUS-ETMX_L2_DAMP_MODE5_OUTPUT
1186668200	ETMY ESD Saturation	No glitch seen prior to saturation, Suspension damping modes	L1:SUS-ETMX_L2_DAMP_MODE6_OUTPUT
1186676278	Beam Splitter Motion with ETMY ESD Saturation	Oscillation in Stage 1 Beam Splitter isolation loop	L1:ISI-BS_ST1_ISO_RY_OUTPUT, L1:ISI-ETMY_ST1_ISO_Z_OUTPUT

LHO	In Lock	Lockloss
True Positives	39	10
False Positives	1	3
Precision	0.97	0.77
Sensitivity	0.93	0.91
Specificity	0.91	0.93

LLO	In Lock	Lockloss
True Positives	49	16
False Positives	3	2
Precision	0.94	0.89
Sensitivity	0.96	0.84
Specificity	0.84	0.96

TABLE 6.1: Performance analysis of lockloss prediction models for LHO and LLO. Each of them respectively has an accuracy of 92% and 93%.

6.5 Conclusions

In this chapter, we have discussed locking the interferometer and the need for constant monitoring and diagnosis of events leading to a lockloss. We show how we incorporate techniques from machine learning to develop an automated tool that uses a clustering based algorithm to track and find the channels that behaved unusually few seconds before the lockloss event. We carry out the analysis of several lockloss events that happened during O2 whose reasons were unknown and throw some light on their possible origin. We further use historic lockloss data collected during periods of high seismic activity and use pattern recognition tools to predict the probability of an impending lockloss based on initial earthquake parameters obtained from the USGS alert system. This information could be used in future to switch the seismic filters so that the interferometer would continue to be in a locked state thus increasing the coincident observation time.

Chapter 7

Transient Classification in Strain and Auxiliary Channels

7.1 Introduction

Detection of short duration GW signals in LIGO data requires reliable identification and removal of noise transients produced by a variety of non-astrophysical sources (Abbott, 2016c; Abbott, 2016b). Noise transients present in the data reduces the reliability of a GW detection by increasing its false alarm probability. Mitigation of noise transients is a major challenge in searches for GW, especially for short duration events, like high mass blackhole binaries, where the signal can be easily mimicked by non-astrophysical transients of varied origin. These often have a waveform morphology close to that of the targeted signal, thus making the differentiation even more difficult (Abbott, 2016a).

With the advent of big data analysis, machine learning has emerged as a useful tool to handle huge volumes of data and to interpret meaningful results from them. In the past few decades, machine learning algorithms such as artificial neural network (ANN) (Adams and Woolley, 1994; Biswas, 2013), support vector machines (Cortes and Vapnik, 1995; Chang and Lin, 2011), random forest (Breiman, 2001), Gaussian mixture model (Powell et al., 2015) etc. found many applications in astronomy and occasionally, have been used for the study of noise artifacts in GW analyses. Since the visual inspection of individual events and their classification is time-consuming and prone to errors, machine learning methods are more effective and reliable for the detection of hidden signatures of astrophysical GW in the data.

We present a hybrid classifier that combines features from supervised and unsupervised machine learning algorithms to do the transient classification. Our classifier performs an unsupervised hierarchical clustering of the incoming data to identify possible groups and a supervised Bayesian (Bayes, 1763) classifier to do the final classification. The classifier code uses features extracted from wavelet analysis of the data in a fast and efficient manner using GPU and MPI parallelization techniques, whereby, making it a good candidate for real-time burst trigger classification and detector characterization. When used to predict the class labels for an input data, the classifier ranks the most likely classes each with an associated probability (confidence level) that may be used to set a threshold to discard unreliable predictions. This multiple-class prediction is useful to identify borderline examples in the feature space. In our study, the classifier was first tested on simulated data consisting of astrophysical bursts along with commonly observed instrumental glitches and then on the LIGO's sixth science run burst hardware injections. Targeted searches for specific glitch types seen in Advanced LIGO's first observation data were also carried out, and the results are reported. Recent methods like deep learning (Krizhevsky, Sutskever, and Hinton, 2012; Zevin et al., 2017) using convolutional neural networks require a large number of training data and are computationally expensive. The fact that we can represent the transient classes with minimal features and fewer training data samples makes our method less susceptible to such issues and speeds up the training process, making it suitable for real-time applications. The results presented in this chapter is based on the publication: Nikhil Mukund et al. Phys. Rev. D 95, 104059.

7.2 Transient Events in GW Data

Table 7.1 lists the transients used in our analysis. Standard searches for compact binary coalescences use matched filtering as the base algorithm (Sathyaprakash and Dhurandhar, 1991), while the burst searches primarily look for excess power in the data along with the time coincidence to trigger a detection (Klimenko and Mitselmakher, 2004; Klimenko et al., 2008). Both these searches are followed by extensive sanity checks, where the auxiliary channels insensitive to astrophysical signals are inspected to rule out possible terrestrial coupling (Abbott, 2016a). Auxiliary channels are often in the thousands, and their coupling with the GW strain sensitive channel is seen to fluctuate in time due to the dynamic nature of the instrument. This often makes the auxiliary channel veto procedure a daunting task. Incorporating a machine learning based veto procedure to identify well-known classes of non-astrophysical transients can help discern the trigger right at the strain channel and thus reduce false alarms.

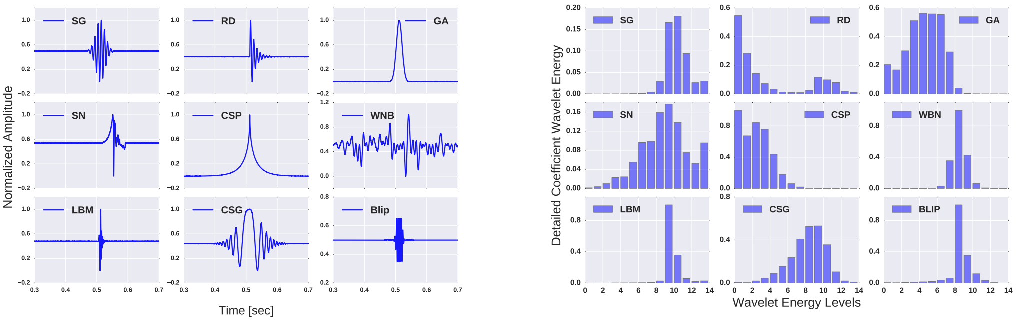


FIGURE 7.1: Left panel depicts typical transient events (SNR set to 50 for better visualization). Wavelet energy median distribution for simulated data (SNR varied from 8 to 100) shown in the right panel

TABLE 7.1: Details of transients used. A=astrophysical, NA=non-astrophysical, O1: Advanced LIGO 1st science run, S6=LIGO 6th science run, Sim=simulated

Transient	Symbol	Type	Search type
Sine Gaussian (Abbott et al., 2004; Canton et al., 2014)	SG	NA	S6, Sim
Ring down (Vishveshwara, 1970)	RD	A	S6, Sim
Gaussian (Abbott et al., 2004; Abadie, 2012)	GA	NA	S6, Sim
Supernova (Zwerger and Mueller, 1997; Burrows et al., 2006)	SN	A	S6, Sim
Cusp (Damour and Vilenkin, 2000)	CSP	A	S6, Sim
White burst noise (Abadie, 2012)	WNB	NA	S6, Sim
Black hole merger (Baker, Campanelli, and Lousto, 2002; Baker et al., 2002)	LBM	A	S6, Sim
Chirping sine Gaussian (Bose et al., 2016)	CSG	NA	Sim
Blip (Abbott, 2016a)	Blip	NA	Sim
Scattering (Abbott, 2016a; Accadia, 2010a; Ottaway, Fritschel, and Waldman, 2012a)	SCT	NA	O1
Type A (low frequency)	A	NA	O1
Type B	B	NA	O1
Type C (blip, Fig. 7.7) (*Adam Day, 2016*)	C	NA	O1
Type D	D	NA	O1
Type E (koi fish) (*Adam Day, 2016*)	E	NA	O1
Type F (needle)	F	NA	O1
Type G	G	NA	O1
Lightning (Abbott, 2016a)	LGN	NA	Targeted

7.3 Feature Extraction

Raw time series is preprocessed by applying a whitening transformation, which enhances the short duration features seen in the data. Transient signals occurring in power systems and neuro-magnetic brain responses have structural and temporal similarities with the glitch signals found in LIGO data streams. Wavelet-based feature extraction for classifying these transients are detailed in Bhagat et al. (2009) and Jayasree, Devaraj, and Sukanesh (2009). A wavelet is a function having a smooth oscillatory pattern which vanishes near the ends (Bhagat et al., 2009). With its desirable qualities like good localization in time and frequency domain, it seems to be a natural choice for extracting information from transient signals. A discrete wavelet transform results in sparser signal representation consisting of a reduced feature set, but still preserves information necessary to differentiate among the classes.

The mother wavelet is defined as

$$\psi_{a,b}(t) = |a|^{-1/2}\psi(\frac{t-b}{a}), \tag{7.1}$$

where a,b $\in$ R and a $\neq$ 0 is scaled and time shifted to form the wavelet family.

The orthonormal basis of Hilbert space $L^2(R)$ consisting of finite-energy signals is obtained by discretization of scale and translation parameters $a_j = 2^{-j}$ and $b_{j,k} = 2^{-j}k$ giving the family wavelet as

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \ with \ j,k \in Z \tag{7.2}$$

The all resolution level wavelet decomposition of the signal has the form,

$$S(t) = \sum_{j=-N}^{-1} \sum_{k} C_j(k)\psi j,k(t) = \sum_{j=-N}^{-1} r_j(t)\,, \tag{7.3}$$

where N = log2 (signal length). The energy E_j at each resolution level is computed as

$$E_j = \|r_j\|^2 = \sum_k |C_j(k)|^2\,. \tag{7.4}$$

The relative wavelet energy at each resolution level

$$p_j = \frac{E_j}{E_{tot}}, \quad \text{where} \quad E_{tot} = \sum_{j<0} E_j\,. \tag{7.5}$$

Wavelet entropy S_{WE} which encodes the degree of disorder in a signal can be written as

$$S_{WE} = -\sum_{j<0} p_j \, ln[p_j]\,. \tag{7.6}$$

Here we carry out similar N-level one-dimensional wavelet decomposition (MATLAB and Wavelet Toolbox Release 2013a, 2013) using an appropriately chosen mother wavelet. Feature extraction for simulated and LIGO O1 data is done using a Daubechies 2 (`db2`) wavelet, while for other search cases, a discrete Meyer (`dmey`) wavelet is used. We use $N = 12$ and 14, respectively, for data sampled at 4 and 16 kHz. Energy in the detail levels and wavelet entropy are then computed and are normalized to unity. Also, the kurtosis of the whitened signal is also used as a distinguishing feature. These features along with the class labels form the input for our Bayesian classifier. Figure 7.1 shows typical transients and their detail coefficient wavelet energy.

7.4 Discrete vs Continuous Wavelet transform

The scale parameter described in Eq. 7.1 is the key parameter that distinguishes between discrete (DWT), and continuous wavelet transforms (CWT). The former uses coarse discretization in terms of powers of 2, $2^j, j = 1, 2, 3, ...$ while the latter discretizes more finely using fractional powers of 2, $2^{j/v}, j = 1, 2, 3....$ DWT becomes the transform of choice when we need sparse representation of our signals and becomes useful for de-noising applications. The orthonormal nature of DWT transform has to lead to its wide spread applications in statistical signal processing and can provide a perfect reconstruction of the original signal from its approximation and detail coefficients. On the other side, DWT is susceptible to shift in the input signal. To tackle this in our analysis, we used the shift-invariant non-decimated discrete wavelet transform. CWT on the other side despite being computationally more expensive provides a higher-resolution signal analysis. They are shifting invariant but in general, do not provide perfect reconstruction, unlike DWT. It is more suited for localization of transients in time and frequency domain. Time-frequency maps (scalograms) constructed using CWT can be used for deep learning based image classification tasks.

7.5 Classifier

Machine learning involves techniques which allow systems to automatically learn and improve prediction accuracies by exploring their past experiences on data. It mimics human decision-making ability by discovering the relationships between the variables of a system from a given set of examples that have both the variables and the observed outcomes. Here we use a hybrid classifier, a supervised Bayesian (Bayes, 1763) one called difference boosting neural network (DBNN) (Philip and Joseph, 2000;) to classify the burst signals.

The DBNN can impose conditional independence on data without a loss of accuracy in the posterior computations. It does this by associating a threshold window with each of the attribute values of the examples (Philip et al., 2002). The network is designed to work with discrete value input features, while GW data features are continuous. A simple method to deal with continuous feature value is to recast it into a suitable number of bins. There are no fixed criteria for the number of bins each feature may take. It might be argued that smaller the bin size, conditions can be imposed with better accuracy. However, in most practical situations, the optimal bin size is close to the square root of N, where N represents the number of discrete values present in the data for that variable.

Once the bins are defined, for each feature bin and the given classes, the allowed ranges for all the remaining features are registered.

The DBNN, being a supervised neural network, requires a training data to configure the network before it can be used for classification of unseen data. The learning takes place by highlighting the difference between the features in two or more classes (Philip et al., 2002) by using Bayesian probability as its central rule for decision making. The confidence in a prediction (Philip, 2010) is the value of the posterior Bayesian probability for a given set of input features.

The working of DBNN can be divided into three units: Bayesian probability estimator, gradient descent boosting algorithm, and a discriminant function estimator (Philip et al., 2002). The network starts with a flat prior for all the classes $P(C_k) = 1/N$, preventing the training from being biased to any specific prior distribution. The first unit in DBNN (executed by option 0 in the implementation) computes Bayesian probability and the threshold function for each of the training sample by constructing a grid for each class with columns representing the attributes and rows their respective values. The bin location for each attribute value is decided such that the full range of values can be uniformly covered by the set number of bins for that attribute across the classes. Initially, the content in attribute bins is all set to one. The training examples are taken one by one, and the bin corresponding to each attribute value for its class is incremented by one. This sampled data is used to compute the likelihood for an attribute value to favor a class, $P(U_m|C_k)$, as the ratio of occurrences (counts) in its bin for the class C_K to the total counts in all k classes for the same bin number that U_m holds for that attribute. The classifier also makes notes for each attribute value and its class, the allowed maximum and minimum values taken by the remaining attributes in the entire training sample. This information is used to negate the possibility that the value of one feature may favor multiple classes unless all other features also have values in the same range across the classes.

Though we started with a flat prior, to compute the Bayesian probability, we need to estimate the actual prior. In the Bayesian framework, prior has no special meaning. It is a weighted bias (belief) about the probable outcome of an experiment based on experiences in the past. In the second unit (executed by option 1 in the implementation), the DBNN estimates prior based on its experience with the given training data. The DBNN does not make any change in the prior for correctly classified examples. In the case of failed examples, it attaches an additional weight to the attributes so that, it may also get correctly classified. To avoid random fluctuations due to the introduction of arbitrary priors, this is done by modifying the flat prior incrementally by $\Delta W_m = \alpha(1 - \frac{P_k}{P_{k^*}})$ through a set of repeated rounds on the training data until the example gets correctly classified. That is, until $P(U|C_k) = \Pi_m P(U_m|C)$ goes to a maximum for the true class represented by the data. Here P_k and P_{k^*}, respectively, represent the calculated Bayesian probability for the true class and the wrongly estimated class, and α is a fraction called the learning rate (Philip, 2010). Since the ratio of the probabilities is taken, this is much like the way humans arrive at their decisions based on their cumulative experiences in the past. This process is called training, and after training, the estimated likelihoods and prior are saved for future use. The assumption during the training process is that a representative training data is available that has suitable examples to represent all the variants in the target space.

The third unit (executed by options 2 and 3 in the implementation) computes the discriminant function. According to Bayesian theorem, the updated belief or the posterior is the product of the prior and the evidence normalized over all possibilities. This can be written as

$$P(C_k|U) = \frac{\prod_m (P(U_m|C_k)W_m}{\sum_k \prod_m (P(U_m|C_k)W_m}, \tag{7.7}$$

where W_m represent the prior weight vector.

DBNN has been successfully applied to many astronomical problems such as star-galaxy classification (Philip et al., 2002), classification of point sources such as quasars, stars, and unresolved galaxies (Abraham et al., 2012), transient classification (Philip, 2012) to indicate a few.

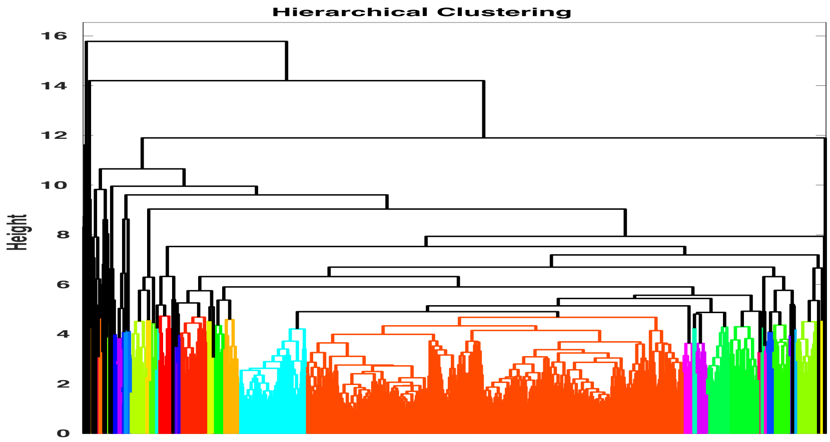


FIGURE 7.2: Dendrogram showing hierarchical clustering of 1000 transient triggers identified in O1 Data from Hanford observatory by the `Omicron` algorithm (Robinet, 2015). The transient morphology changes progressively from left to right

As for the case of all supervised networks, the accuracy of the predictions depends on the initial class selection and quality of the training data sets. When encountering real instrument data where it is difficult to know beforehand the actual groups present, running an unsupervised classifier before the Wavelet-DBNN classifier was seen to vastly improve the results. This step becomes more relevant for targeted searches looking for a particular transient class where unsupervised learning can yield insights into contamination from other glitch classes. Prior information about other glitches with very similar morphology can be made use of by the network to learn to differentiate between them whereby improving the accuracy.

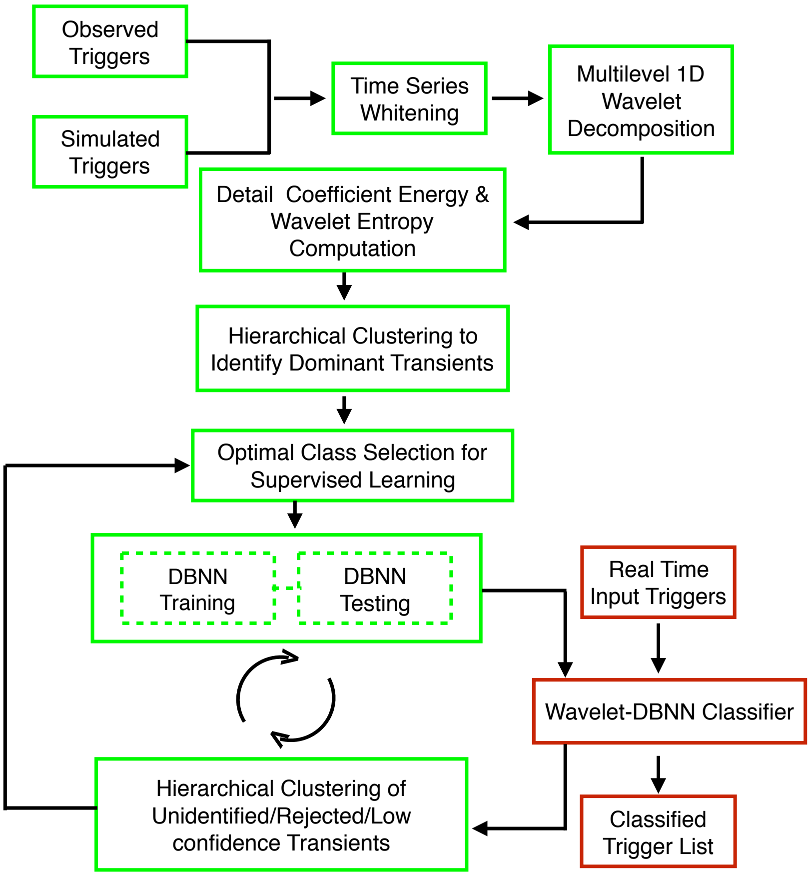
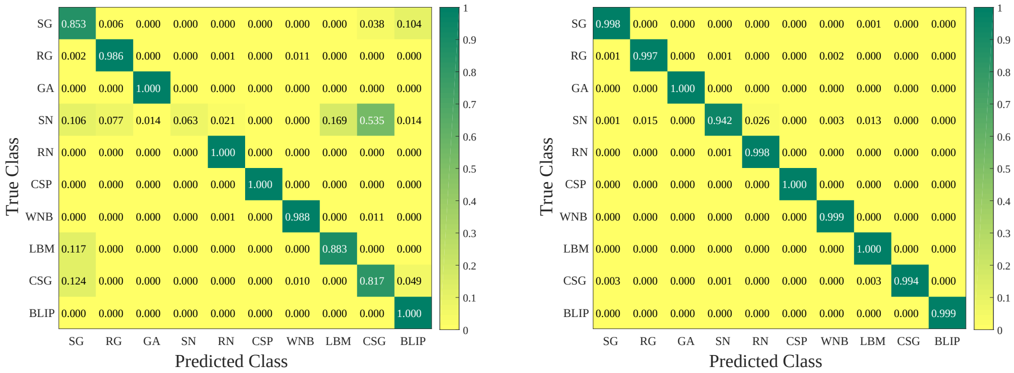


FIGURE 7.3: Block diagram of the proposed hybrid classifier.

We run an unsupervised classifier using Hierarchical clustering on the data to get an idea about the possible transient groups currently present in the data and their respective distribution (see Fig. 7.2). A classifier trained this way is observed to outperform the other scenarios where class selection is done either by visual inspection or by using predefined classes.

We employ a bottom up agglomerate clustering where the pairwise distance is calculated using the Mahalanobis distance measure (Mahalanobis, 1936). The



	SG	RG	GA	SN	RN	CSP	WNB	LBM	CSG	BLIP
SG	0.853	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.038	0.104
RG	0.002	0.986	0.000	0.000	0.001	0.000	0.011	0.000	0.000	0.000
GA	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SN	0.106	0.077	0.014	0.063	0.021	0.000	0.000	0.169	0.535	0.014
RN	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
CSP	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
WNB	0.000	0.000	0.000	0.000	0.001	0.000	0.988	0.000	0.011	0.000
LBM	0.117	0.000	0.000	0.000	0.000	0.000	0.000	0.883	0.000	0.000
CSG	0.124	0.000	0.000	0.000	0.000	0.000	0.010	0.000	0.817	0.049
BLIP	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000

	SG	RG	GA	SN	RN	CSP	WNB	LBM	CSG	BLIP
SG	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000
RG	0.001	0.997	0.000	0.001	0.000	0.000	0.002	0.000	0.000	0.000
GA	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SN	0.001	0.015	0.000	0.942	0.026	0.000	0.003	0.013	0.000	0.000
RN	0.000	0.000	0.000	0.001	0.998	0.000	0.000	0.000	0.000	0.000
CSP	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
WNB	0.000	0.000	0.000	0.000	0.000	0.000	0.999	0.000	0.000	0.000
LBM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
CSG	0.003	0.000	0.000	0.001	0.000	0.000	0.000	0.003	0.994	0.000
BLIP	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.999

FIGURE 7.4: Confusion matrix for simulated data: Results from traditional SVM (left) and DBNN (right) classifiers. The closer a diagonal element is to unity, the better is the classification for the corresponding type. Accuracy of our method is thus evident.

criterion for estimating the linkage between the clusters is based on the average distance between pairs of signals among the clusters, weighted by the numbers of elements in each cluster. Cluster linkage at each level of dendrogram is calculated recursively whose value for a given pair of clusters is given by

$$d(r,s) = \sqrt{\frac{2n_r n_s}{(n_r + n_s)}} \|\tilde{x}_r - \tilde{x}_s\|_2 \,. \tag{7.8}$$

The optimal distance measure used for linkage and the original mother wavelet used for decomposition are both selected based on the value of the cophenetic correlation coefficient, c, (Sneath and Sokal, 1973) with a value close to unity being ideal,

$$c = \frac{\sum\limits_{i<j} (Y_{ij} - y)(Z_{ij} - z)}{\sqrt{\sum\limits_{i<j} (Y_{ij} - y)^2 \sum\limits_{i<j} (Z_{ij} - z)^2}} \,. \tag{7.9}$$

Y_{ij} is the pairwise distance between parametrized waveforms while Z_{ij} is their linkage distance. y, z, respectively, represent the average value of the corresponding distance measures.

Optimal leaf ordering of the resulting dendrogram is achieved by maximizing the sum of similarities between adjacent leaves (Bar-Joseph, Gifford, and Jaakkola, 2001). This step is carried out to identify the relationship between the various clusters and to locate possible subgroups. For example, in Fig. 7.2, transients at both ends are least related to each other. The schematic of the hybrid classifier useful for real-time transient classification is shown in Fig. 7.3.

7.6 Simulated Data

A simulated data set was created consisting of 49845 transients from ten classes (refer to Table 7.1) whose SNR is varied uniformly between 8 and 100 using Gaussian white noise addition. Signals in each class are generated for different values of parameters sampled from a wide range. Details of bursts used in the simulation are given below (t_o is set to 0.5 Sec).

(i) Gaussian (GN)

These broadband non-astrophysical signals are modeled as simple Gaussians with a duration parameter τ taking values 0.0005, 0.001, 0.0025, 0.005, 0.0075, 0.01, 0.02, and 0.05 (Abbott et al., 2004),

$$s(t) = \exp(-\frac{(t-t_o)^2}{\tau^2})\,.$$

(ii) Sine-Gaussian (SG)

SG models a non-astrophysical glitch which produces significant triggers in matched filtering analysis for coalescing compact binaries (Canton et al., 2014). τ is set to $2/f_o$ with a central frequency (f_o) logarithmically spanning from 100 Hz to 2000 Hz,

$$s(t) = \exp(-\frac{(t-t_o)^2}{\tau^2})\ \sin(2\pi f_o(t-t_o))\,.$$

(iii) Ringdown (RG)

RG signals have a longer duration but a shorter bandwidth and are modeled as damped sinusoids. They are produced from quasi-normal modes of a final black hole formed from coalescing compact binaries (Vishveshwara, 1970). Here we set $\tau = 4/f_o$ with f_o similar to that of sine-Gaussian data set,

$$s(t) = \begin{cases} \exp(-\frac{(t-t_o)}{\tau})\ \cos(2\pi f_o(t-t_o)) & \text{if } t \geq t_o \\ 0 & \text{if } t < t_o\,. \end{cases}$$

(iv) Chirping Sine Gaussian (CSG)

CSG is similar to SGs but with an additional chirping parameter (Bose et al., 2016). This signal closely models the whistle glitches frequently seen in LIGO detector data. The equation below gives the waveform model where each of the parameter is varied as follows: f_o:{5,100}, α:{10,100} and

τ:{0.001,0.025} ,

$$s(t) = \frac{\exp\left(\frac{-(1-1i\alpha)\ (t-t_o)^2}{(4\tau^2)} + 2\pi i(t-t_o)f_o\right)}{(2\pi\tau^2)^{\frac{1}{4}}}.$$

(v) Supernova (SN)

Zwerger-Mueller waveforms (Zwerger and Mueller, 1997), one of the Supernova waveforms, are produced by an axisymmetric core collapse of supernovae. These are obtained by hydrodynamical simulations of the stellar core collapse by varying the initial conditions like adiabatic index, spin, and differential rotation profile. We incorporate 78 models (with varying SNR) consisting of a simple analytic equation of state. We also make use of Ott-Burrows supernova waveforms (Burrows et al., 2006) in our analysis.

(vi) Cusp (CSP)

Symmetry breaking phase transitions in the early Universe could generate cosmic strings (Damour and Vilenkin, 2000) with a cusp-like signal, $h(f) = A(f)f^{-4/3}$. Such waveforms are simulated with an exponential roll off after a cutoff frequency f_o which is varied from 50 Hz to 2000 Hz.

(vii) White Noise Bursts (WBN)

WBN has in general very complex time-frequency morphology. Their spectra are white in the specified band and zero outside (Abadie, 2012). Here we construct a set of burst signals which have central frequency spanning from 50 to 300 Hz, bandwidth 50 to 150 Hz and duration 0.1 to 0.4 seconds.

(viii) Black Hole Merger (LBM)

These waveforms capture the coalescence radiation emitted from a merger of binary black hole systems using Lazarus approach (Baker, Campanelli, and Lousto, 2002). Analytic approximation (Baker et al., 2002) is used to construct time domain templates to replicate the merger scenarios. We considered black hole binaries with a chirp mass in the range $\{20, 50\}$ and cosine of inclination angle varied between zero and one.

(ix) Blip (Blip)

Blips are observed frequently in both LIGO detectors, but their origin is not well-understood (*Adam Day, 2016*). Hardware injections carried out at the observatories sometimes hit the saturation limit of the actuator resulting in signals which look similar to blips. Hence we simulate them by clipping sine-Gaussians at a few percent levels around the mean amplitude.

While sampling the parameters, care was taken to ensure that the signals within a class are significantly different. Table 7.2 shows the performance of a Wavelet-DBNN classifier. The total number of samples, size of the training set, true positives (TP), false positives (FP), precision, sensitivity, and specificity are reported (see (Powers, 2007) for terms definition). The resulting confusion matrix is shown on the right panel of Fig, 7.4.

TABLE 7.2: Simulated transient signals.

Name	Total	Train.	TP	FP	Preci.	Sensi.	Speci.
SG	5000	552	4991	22	0.99	0.99	1.00
RG	5000	311	4984	16	0.99	0.99	1.00
GA	5000	155	5000	0	1.00	1.00	1.00
SN	745	313	702	14	0.98	0.94	1.00
RN	5000	114	4992	19	0.99	0.99	1.00
CSP	5000	14	5000	0	1.00	1.00	1.00
WNB	10000	421	9994	15	0.99	0.99	1.00
LBM	5000	586	4999	31	0.99	1.00	0.99
CSG	5000	378	4969	0	1.00	0.99	1.00
BLIP	4100	153	4097	0	1.00	0.99	1.00

Our Positive Condition is defined as the classifier predicting that the transient is belonging to a particular class. Now that can be true or false.

$$True\ Positive\ Rate\ (Sensitivity) = \frac{\Sigma Predicted\ positive\ condition\ actually\ true}{\Sigma All\ actually\ positive\ conditions} \tag{7.10}$$

$$False\ Positive\ Rate = \frac{\Sigma Predicted\ positive\ condition\ actually\ false}{\Sigma All\ actually\ negative\ conditions} \tag{7.11}$$

$$False\ Negative\ Rate = \frac{\Sigma Predicted\ negative\ condition\ actually\ false}{\Sigma All\ actually\ positive\ conditions} \tag{7.12}$$

$$True\ Negative\ Rate\ (Sensitivity) = \frac{\Sigma Predicted\ negative\ condition\ actually\ true}{\Sigma All\ actually\ negative\ conditions} \tag{7.13}$$

For comparison with a standard classifier, we use publicly available support vector machine (SVM) implementation `LIBSVM` (Chang and Lin, 2011; Eaton et al., 2015) on the same wavelet decomposed parameter sets. Figure 7.4 clearly shows how our Wavelet-DBNN classifier outperforms the traditional classifier. The stark difference is observed for Supernova signals where the SVM shows a very high misclassification, most likely due to the limited number of data samples and the inherent diversity in their morphology.

7.7 S6 hardware Injections

To check the performance of our classifier on real data, we use the classifier on the LIGO strain data obtained from the sixth science run (Aasi, 2015). We apply our classifier to six different classes of hardware injected short duration transient signals as given in Table 7.3. These are traditionally performed to access the efficiency of various astrophysical search pipelines both regarding detection and parameter estimation. Another motivation is to check the low latency alert system that is in place to provide rapid alerts to the astronomers outside the LIGO collaboration to perform follow-up observations in various electromagnetic bands. The strain data is whitened to better identify the transients and then down sampled to 4096 Hz. 1634 transients with SNR greater than ten are used in the analysis. Table 7.3 gives the results after classification.

TABLE 7.3: S6 hardware injections.

Name	Total	Train.	TP	FP	Preci.	Sensi.	Speci
SG	1476	69	1476	10	0.99	1.00	0.94
RG	36	25	33	0	1.00	0.92	1.00
GA	46	38	44	8	0.85	0.96	1.00
SN	41	34	33	4	1.00	0.86	1.00
CSP	28	27	24	0	1.00	0.86	1.00
WNB	29	27	24	0	1.00	0.83	1.00

7.8 Targeted Search: LIGO Strain Channel

Detector characterization studies revealed several kinds of non-astrophysical transients in the Advanced LIGO's strain data during its first observation run (Abbott, 2016a). Identification of these transients and establishment of their non-astrophysical origin were crucial for the detection of GW signal (Abbott, 2016c; Abbott, 2016b; Abbott, 2016a). For those known classes, if we could automate their detection using machine learning methods, it will reduce the noise background in the astrophysical GW searches. We used LIGO Hanford Observatory (LHO) strain data (September 18 to January 12) consisting of 28354 transient triggers (Robinet, 2015) with SNR ranging from 8 to 100 and having a maximum frequency of 2096 Hz. Our classifier is then used to search for events which look similar to the major transient classes (see Table 7.1) evident from the initial hierarchical clustering. Several of these could potentially limit generic burst searches, in particular, cosmic cusps and supernova events. One second whitened data around the trigger is used for feature extraction. The training set consisted of minimal samples ranging from five to ten per class. Fig. 7.5 shows the distribution of classified transients with a similar morphology to the training set.

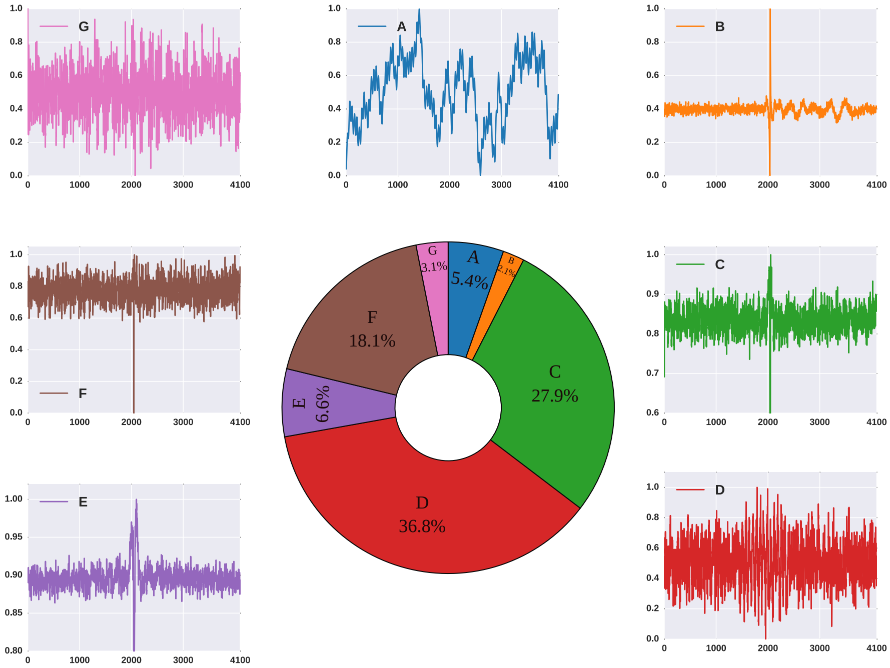


FIGURE 7.5: Donut chart shows the distribution of O1 transients classified by DBNN-Wavelet classifier into the major classes identified by hierarchical clustering. Representative transients from each class (sampled at 4096 Hz) are also shown.

We also analyzed with LIGO Livingston data for triggers in SNR range 10 to 100 and obtained comparable results. Classifier found a strong presence of glitches caused due to previously proposed scattering mechanism (Accadia, 2010a; Ottaway, Fritschel, and Waldman, 2012a). We observe coincidence (within one second) between some of the classified scattering glitches and the triggers seen in LIGO's auxiliary angular length sensing channels. These auxiliary channels carry information about the motion of signal recycling cavity optics. The coupling was seen to occur predominantly from the pitch and yaw degrees of freedom with a respective contribution of 17.3% and 12.5% with 50% of the glitches coincidentally seen in both the channels. Scattering happens when an off-axis beam gets reflected from the beam tube and recombines with the main beam. These morphology based identification coupled with coincident analysis would help one to narrow down to the region mostly likely to cause the transients and also help in applying appropriate data quality vetoes.

7.9 Targeted Search: LIGO Auxiliary Channels

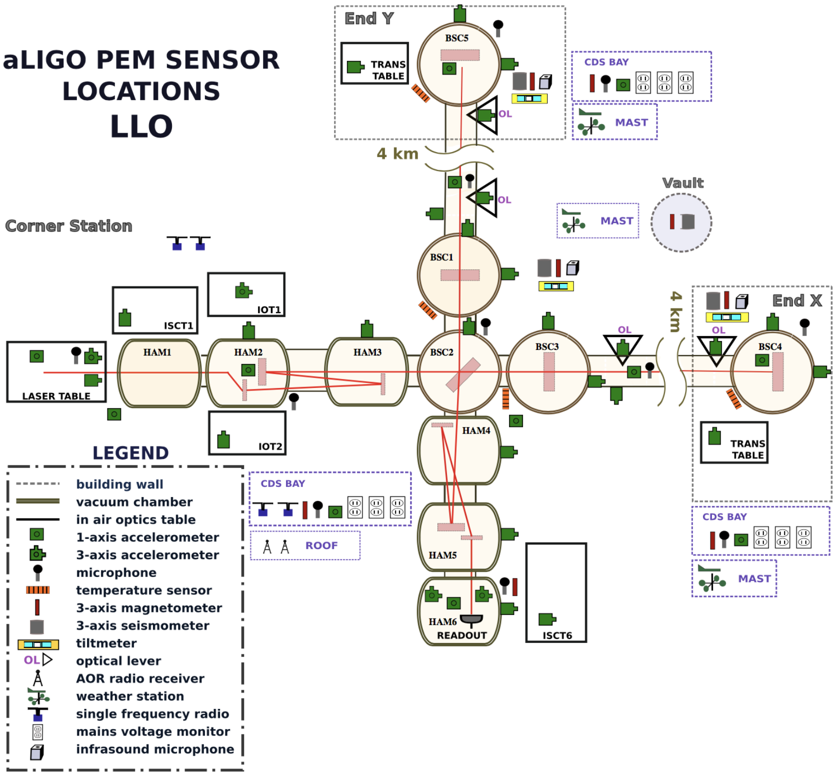


FIGURE 7.6: Physical Environment Monitoring (PEM) channels used at LIGO Livingston. *http://pem.ligo.org/channelinfo/index.php*

Apart from the channel that carries information about the gravitational waves GW detectors uses numerous auxiliary channels to gather information about the environmental conditions about the instrument (See Fig. 7.6). These can carry information about the ambient temperature, stray EM & magnetic fields, ground motion, acoustic vibrations, to name a few. Data from many of these (ex:seismic) are used to control the isolation platforms while some others are used to veto spurious transients seen in strain channel based on their temporal coincidence. Severe weather conditions can affect both the detectors and, if not properly vetoed, can be misinterpreted as a true signal. Variation in the ambient magnetic field during lightning and a thunderstorm around LIGO can affect sensors and actuators present in multistage suspension systems that isolate and control the LIGO test mass. They are seen in magnetometers with a very distinct time-frequency morphology (Fig. 7.7, right panel). These also induce currents in the beam tube and are detected by on-site clamp meters. Here we apply our classifier to separate out lightning events from the other transients seen in the magnetometer data. We use LIGO Livingston Observatory (LLO) Y-arm magnetometer Omicron triggers generated from 16:00:00 to 23:00:00 UTC of December 16, 2015. Hierarchical clustering on the first 30 minutes of data generates the training set, which is fed to the supervised classifier that performs the final targeted search. Triggers with SNR 15 to 1000 and frequency 1 to 1024 Hz are used for the analysis. 42 out of 689 such triggers are identified to be caused by lightning. Similar search carried out in LLO X-arm magnetometer data for the same period identifies 45 lightning triggers. Our results are consistent with the local weather data, which reported lightning activity during the same period. The number of misclassification in these cases turned out to be only one and six respectively.

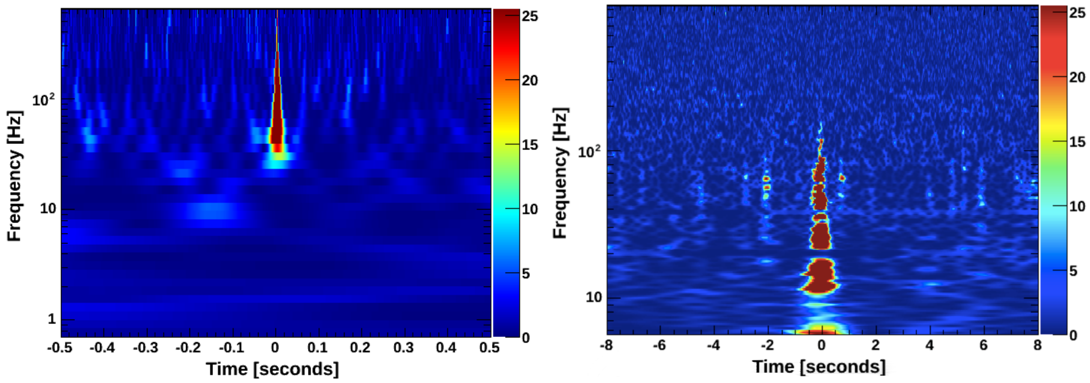


FIGURE 7.7: OmegaScan: Type C glitch in LHO strain channel (left) and lightning glitch in LLO magnetometer (right). Plots generated through LIGOdv-web (*LigoDV-Web*).

7.10 Development of an Improved Classifier

After the successful application of our wavelet-DBNN classifier on O1 data, we improved upon our classifier before its application on O2 data. We observed that there are subclasses within the glitch classes that were previously considered hence required a more refined classifier for more accurate prediction. This also implied that we learn the finer details of each transient using better feature extraction techniques. The previous classifier, although probabilistic, had the disadvantage that it tried to classify every trigger into either of the training classes, thus reducing the chance of detecting a true outlier. An ideal classifier should generate a very low probability for every trained class if it encounters an event which has a previously unseen morphology. Another point to consider was the fact that each classifier would have its short comings, so better strategy to maximize the predictions was to combine the results from more than one classifier each trained independently from the another.

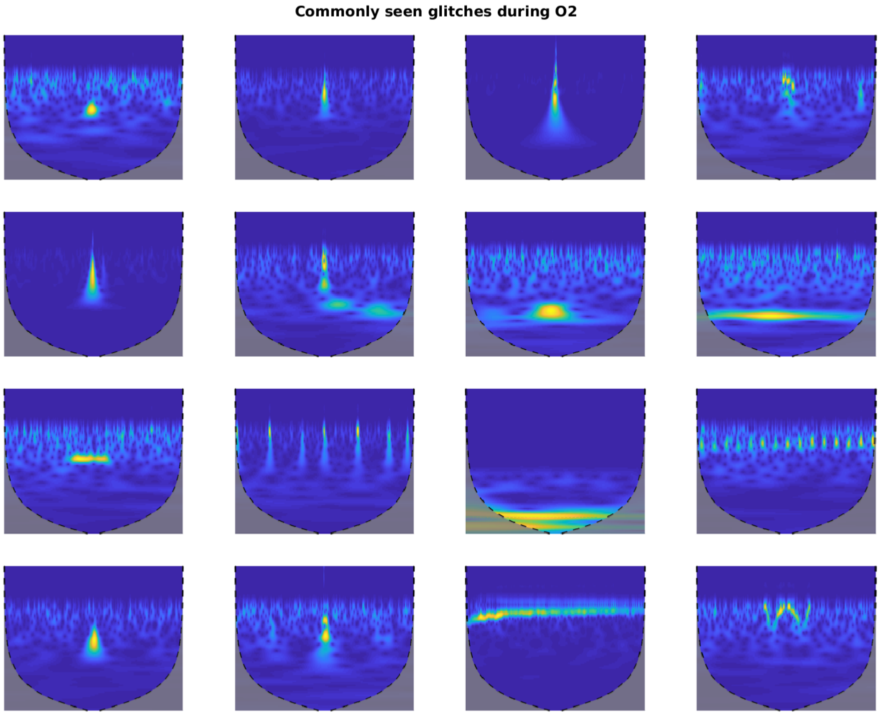


FIGURE 7.8: Continuous wavelet transform based image scalogram of commonly observed glitches during the O2 science run. Each image is made using one second of whitened strain data and consists of frequency information till 512 Hz. Names of the classes (left to right) are as follows: air compressor, blip, extremely loud, helix, koi fish, light modulation, low-frequency burst, low-frequency lines, power line, repeating blips, scattered light, scratchy, tomte, violin mode, wandering line and whistle.

7.10.1 Deep transfer learning

If one needs to train a Deep Neural Net (DNN) from scratch, it typically requires a very large data set consisting of millions of images and very long training periods. Usually, such large data sets are never available or are hard to generate. The alternative approach would be to use a network which has previously trained on a very large dataset, but consisting of totally unrelated images. The initial layers of the DNN serve the purpose of extracting features of the image but at different scales (similar to convolution using different types of filters) while the last layers connect these learned features to the categories of objects available for prediction. So if we retrain the initial layers and replace the last layers and train them on the dataset of our interest, we would still be able to extract the needed features and force the classifier to make predictions on objects that are of interest to us. We use AlexNet classifier and change the last three layers and retrain the network using the training data set of transient scalograms.

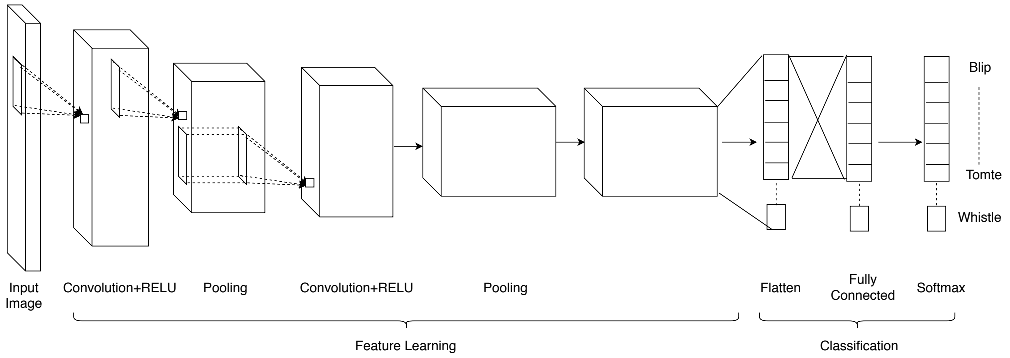


FIGURE 7.9: Deep transfer learning using AlexNet

7.10.2 Multi-Class SVM based on Wavelet Packet feature extraction

The other parallel classifier we use is based on direct feature extraction from the whitened time series. In our last work, we had used DWT and had computed the relative energy at 14 levels. As we are now encountering more classes, it was necessary to switch to a more detailed feature learning technique for which we make use of wavelet packets. In this technique at each level we further decompose the detail coefficients into approximation and detail coefficients (see Fig. 7.10) leading to superior frequency resolution as compared to the original DWT method (Laine and Fan, 1993; Ta, 1994; Walczak, Van Den Bogaert, and Massart, 1996). It is clear that the median wavelet packet distribution is shown in 7.11 offers much more scope for separation among the various transient classes as compared to the histogram provided in Fig. 7.1.

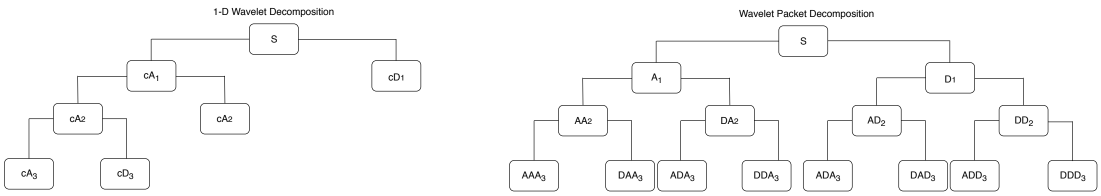


FIGURE 7.10: Comparison of 1-D wavelet decomposition to wavelet packet decomposition at level three.

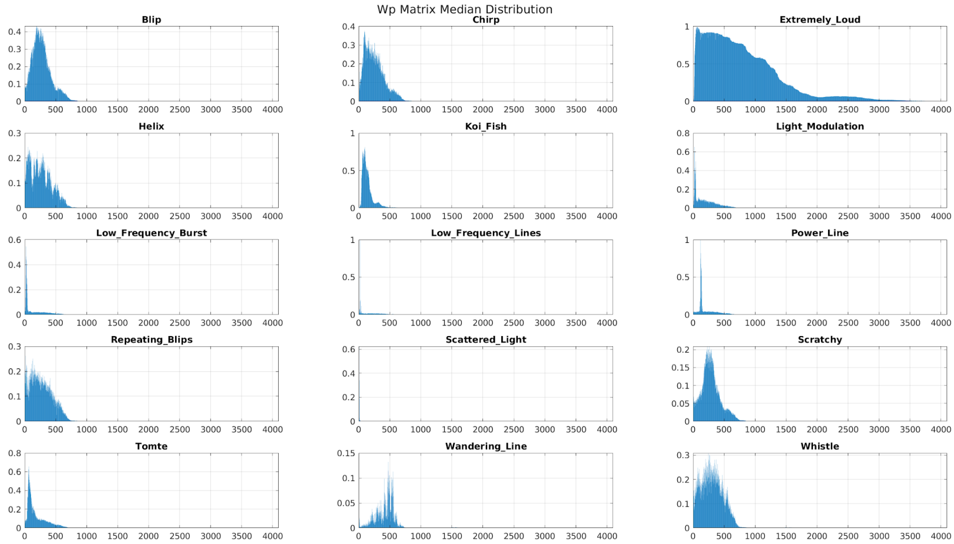


FIGURE 7.11: Wavelet packet feature extraction

7.10.3 Ensemble Classifier

We combine the prediction probabilities from each of the above classifiers and supply them as input to the decision tree based ensemble classifier. With this step, we ensure that the inability for certain method to distinguish between certain classes is decreased thus improving the reliability of our predictions. The results from applying the ensemble classifier to omicron triggers found in Advanced LIGO second observation run data is given in Fig. 7.12. [1] Astrophysical signal search pipelines looking for high mass BBH signals can make use of the prediction probabilities to issue morphology based veto for the triggers seen in the data. This could in future be used either as an alternative or in addition to the traditional chi-square based veto.

---

[1] Probabilistic classification results for each trigger can also be accessed at the following web link: `https://ldas-jobs.gw.iucaa.in/~nikhil.mukund/WebInterface/`.

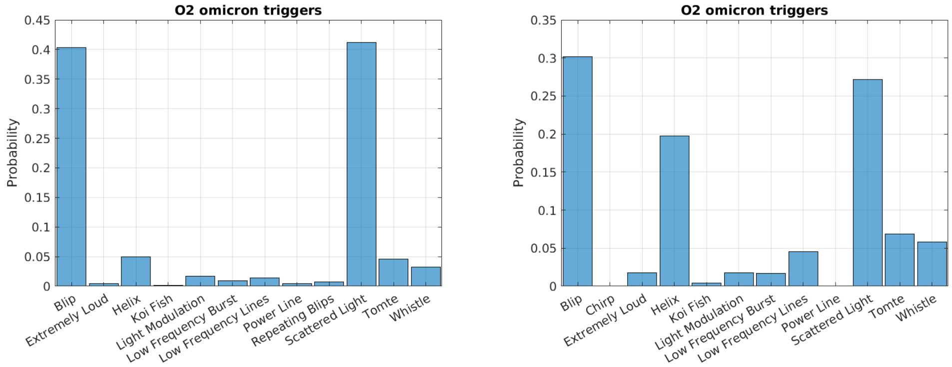


FIGURE 7.12: Results from ensemble classification of omicron triggers seen in GW strain channel during Advanced LIGO's second observation run at both the LIGO detectors. The ensemble learner combines inputs from deep learning classifier based on transfer learning and traditional feature extraction based SVM classifier.

7.11 Conclusions

We have convincingly demonstrated the resourcefulness of machine learning in detector characterization and burst signal analysis in LIGO like complex instruments. We showed that an effective feature extraction technique, in conjunction with an efficient classifier, can be used to classify a variety of transients in a practical situation involving real data. In the initial analysis, we used relative wavelet energy, wavelet entropy, and kurtosis as a possible parameter set for classifier input. This, coupled with a difference boosting neural network, was very accurate in discerning between classes with slightly different morphology and possibly different physical origin. The usefulness of the method was shown in our analysis where we could do an accurate targeted search for a specific glitch using minimal training sets. The parameter set used here can be expanded to include other features which can aid the classification even when the corresponding values are unavailable for other classes. The special construction of the classifier makes sure that it does not suffer from the curse of dimensionality, unlike most neural network classifiers. Hence the feature set can be expanded in future without causing much computational overhead. Combining class information along with multichannel coincidence analysis will help to narrow down to the cause for a particular kind of transient present in the data. If there is a good enough reason to believe that the trigger is non-astrophysical, then glitch based vetoes can be applied to those times. This would lower background triggers in search pipelines thus enhancing confidence in the true detections. We plan to develop such a data quality vector which can be used to directly veto low latency triggers produced by search pipelines looking for astrophysical signals. We further improved the analysis by improving both the feature extraction technique as well as the classifier accuracy. We introduced a way to combine the predictions from multiple classifiers through the use of an ensemble classifier. We finally use it to on the Advanced LIGO's second observation data and find out the distribution of the dominant glitch groups present in the strain data.

Chapter 8

Information Retrieval & Recommendation System for GW Detectors

8.1 Introduction

Large science projects, especially astronomical observatories, have plenty of data about operations, scheduling, maintenance, and general observational activities all logged in text form which is often unstructured. Over the years, these logbook entries will accumulate, covering most of the aspects regarding multiple instruments within the observatory. It has to be noted that although the key technologies are changing at a rapid phase, the fundamental principles involved in construction and maintenance at these places are getting altered less rapidly. This necessitates the need for keeping a record of activities carried out over the years which would be useful for prompt diagnostics. Projects like SKA, TMT, LIGO, SALT, and JWST also require extensive internal coordination. These typically are a collaboration consisting of hundreds to thousands of scientists whose research can span areas like instrument fabrication, installation, commissioning, characterization, maintenance, upgrade, data analysis, and parameter estimation. Often the project time span is spread across few decades and thus generate information whose volume and complexity cannot be handled effectively by traditional search-engine-backed information processing tools. On the positive side, analysis of such big data volumes can yield powerful insights into the inherent trends and fluctuations within the concerned project.

Data mining in the big data framework often encounters difficulty in both extracting the relevant information from the data and in coming up with meaningful interpretations in a highly coherent fashion (Fan, Han, and Liu, 2014; Khan et al., 2014; Wu et al., 2014). Due to the lack of standardization, the logged data often come in a format that is not suitable to be stored in relational databases of coherent hierarchy (Stephens et al., 2015). The methods in which data are stored and associated with different entities also pose a challenge in mining the required information from it. For example, in a gravitational wave observatory, there is a core science dataset with plenty of ancillary data on the observation and a variety of other auxiliary datasets collected from various sensors and actuators. Ancillary data includes information in the form of technical documents, images, drawing, and measurements.

Even though data mining methods like association analysis, clustering, and other machine-learning techniques exist, the presentation of unstructured data into these algorithms and inference generation is observed to be a non-trivial task (Han, Kamber, and Pei, 2011). The generation of insights from big data with recommendation systems based on learning from unstructured text data (Pazzani and Billsus, 2007) tackles these challenges (LaValle et al., 2011; Hu et al., 2014). Descriptive recommendations and information retrieval (Sigurbjörnsson and Zwol, 2008; Gretzel et al., 2004) have recently gained popularity and have been applied to areas like travel recommendation systems (Gretzel et al., 2004) and content personalization systems (Liang, Lai, and Ku, 2006). Besides commercial applications, text-summarization-based content recommendation (Hassan et al., 2009) is an exciting area that has a high level of applicability to different areas of science and research (Miner, Elder IV, and Hill, 2012; Kerzendorf, 2017). Unlike conventional rank-based search systems, these do not perform topical modeling and rank topics of recurring interest (Zoghbi, Vulić, and Moens, 2013). Topical modeling is usually done for retrieving information from a single website with multiple topics. The challenge is when different topics on a single site may be weakly linked to each other (Cointet and Roth, 2010). While there could be already known relations among different entities, the process of data mining and better data representation can reveal the latent unanticipated linkages among different topical entities (Behrens and Bassu, 2006).

In this work, we make use of the open source logbook data from the Laser Interferometric Gravitational Observatory (The LIGO Scientific Collaboration, 2015) to depict the resourcefulness of knowledge rediscovery backed by natural language-processing (NLP, Ricci, Rokach, and Shapira, 2011). Additionally, such efforts will efficiently disseminate technical knowledge to a wider audience and will help the ongoing efforts to build upcoming detectors, such as LIGO-India, by even helping to foresee possible challenges during the design phase. This is a novel approach to observational astronomy, and the developed software is made available to the public through a web application named **Hey**

LIGO[1]. We also show the application of descriptive content-based recommendations to compare common issues among multiple observatories. These methods are scalable and will be very useful in the event of upcoming projects like the Square Kilometre Array (SKA) and the LIGO-India detector.

This chapter is organized as follows. In Section 8.2 we describe the methodology adopted to convert raw data into useful and representable information. Section 8.3 provides the details of the data used in our analysis. The features of the recommendation system are outlined in Section 8.4. Finally, in Section 8.5 we apply the scheme to various gravitational wave observatories around the world and discuss the results obtained in Section 8.6.

8.2 Contextual learning of unstructured Data

Structured data are highly organized and usually reside in a relational database schema. Unstructured data refers to information that does not follow the traditional database scheme. For example, e-mails, web pages, business documents, and FAQs are some examples of unstructured data. They include text and even multimedia content. Therefore, processing such information is usually a tedious task.

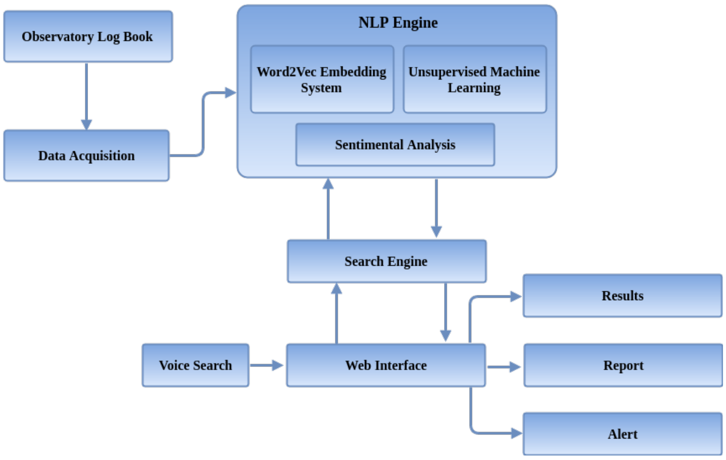


FIGURE 8.1: Schematic depicting the information retrieval and recommendation system.

This section briefly describes the manner in which unstructured textual data are acquired, processed, and finally given a structure. Moreover, it also enumerates the various steps that are involved in the development of a machine-learning model and are used to differentiate between the available textual data

---

[1] `heyligo.gw.iucaa.in`

points contextually. Finally, the model is used to perform clustering over all the textual data, thus adding a structure for the ease of access. Fig. 8.1 shows the schematic representation of our web interface that has been used to implement the scheme.

The unstructured dataset that we use is in the form of textual web pages. These pages have an identical HTML structure and certain attributes defined for every data point. Due to this common structure and the open source nature of the web pages, it is possible to write a script that can extract each attribute from the HTML source code and organize the complete information into a dataframe. A dataframe is a tabular structure with columns as attributes and rows as individual datapoints. The first part of our algorithm perfroms data acquisition using the python package 'Beautiful Soup' (Richardson, 2017) to retrieve information from web pages by searching through all the new posts and related data. It saves them into relevant files for future utilization.

Once the data sets are stored locally, non-essential attributes are removed, textual timestamps are converted to system timestamps, duplicate data points are removed/combined, and the resulting data are passed onto the text-processing unit. A vocabulary for our data is generated by converting the unstructured data into stem words. For that, we have removed all special characters and punctuations such as $!, @, \#, \$, \%, *, \&, {}', {}'', (,)$ etc. All non-English words and other HTML tags, and URLs, are also excluded from the data. The text is then tokenized (Huang et al., 2007) by splitting the strings of text into a list of words called tokens. To reduce the redundancy in the vocabulary, we convert the related token forms and their derivatives to a common base stem by a process known as 'stemming' (Frakes, 1992).

We convert textual data into vectors that can be easily handled by the computer (Li et al., 2015) using the process of 'embedding'. There exist various embedding algorithms like One-Hot Encoding (Harris and Harris, 2012) and Term Frequency Inverse Document Frequency (TFIDF) (Leskovec, Rajaraman, and Ullman, 2014). Not all methods can capture the contextual differences between the words. However, a recent breakthrough in the field of NLP incorporates neural networks that can learn the vector values for each word by iterating over the text multiple times using a gradient-based algorithm (Mikolov et al., 2013b; Mikolov et al., 2013a). Bengio et al. (2003) have coined the term word embeddings with a neural language model to train them with the model's parameter.

One of the commonly used tools for converting words into vectors is "Word2vec," described in Mikolov et al. (2013b). Word2vec has a single hidden layer, fully connected neural network that takes a large text corpus as input and produces a higher dimensional vector for each unique word in the corpus. Words that share common contexts in the corpus are located close to each other in the vector space. Word2vec models do not consider word order and can capture semantic information between words in a very efficient way (Ling et al., 2015). With the help of Word2vec embeddings, a computer can differentiate between words of different types. Word2vec implements two computationally less expensive models known as Continuous Bag of Words (CBOW) and a Skip-gram model (Mikolov et al., 2013b) to learn word embeddings. The representation of a corpus of text or an entire document in the form of a list of words (Multiset) is referred to as Bag of Words representation (Markov and Larose, 2007). The algorithm essentially tries to predict the target based on a set of context words (Mikolov et al., 2013b; Mikolov et al., 2013a).

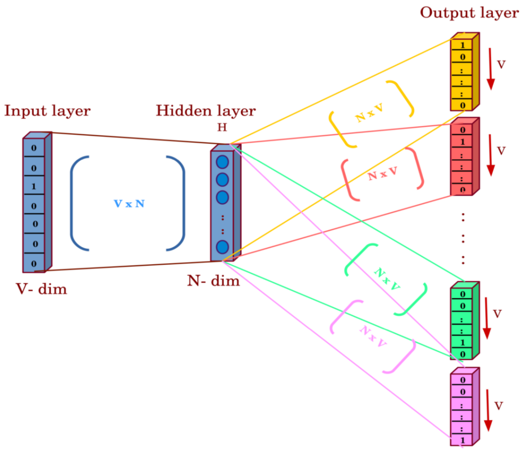


FIGURE 8.2: Skip-gram model for creating word vectors using neural networks. The model predicts contextual words after being given a target word. A linear transformation is used to project the input layer (one hot encoding) to the hidden layer via a $\mathbf{V} \times \mathbf{N}$ matrix. The hidden layer consisting of nonlinear activation functions is connected to the output layer via $\mathbf{N} \times \mathbf{V}$ matrix.

The model that we have used in this work is the Skip-gram model, whose architecture is shown in Figure 8.2. This representation is similar to the CBOW model, but instead of predicting the target word, it predicts the context words based on a given target word (Mikolov et al., 2013a). The model maximizes the probability for classification of a word based on another word in the same sentence (Mikolov et al., 2013b). Thus, the vector representation is capable of capturing the semantic meanings of the words from a sequence of training words $w1, w2, ..., wT$ and their contexts c. The algorithm can be briefed as follows. First, the words are applied as input to a log-linear classan ifier whose objective is to maximize the average log probability given by

$$\mathcal{L} = \frac{1}{T}\sum_{t=1}^{T} \sum_{-c\leq j\leq c, j\neq 0} \log\, P(w_{t+j}|w_t)\;. \tag{8.1}$$

Larger values of c can result in a higher accuracy but require more training time (Mikolov et al., 2013a). To obtain the output probability, $P(w_o|w_i)$, the model estimates a matrix that maps the embeddings into a V-dimensional vector O_{w_i}. Thus, the probability of predicting the word w_o given the word w_i is defined using the softmax function

$$P(w_o|w_i) = \frac{\exp\left(O_{w_i}(w_o)\right)}{\sum_{w\epsilon V}\exp\left(O_{w_i}(w)\right)}\;,$$

where V is the number of words in the vocabulary (Mikolov et al., 2013a; Ling et al., 2015). But this formulation is computationally intensive for larger vocabularies. This problem is alleviated in Word2vec by using the hierarchical softmax function (Morin and Bengio, 2005) or with a negative sampling approach (Goldberg and Levy, 2014).

After embedding all words, every data point is represented as the average of all the word vectors of the words present in it. A Nearest Neighbor Algorithm (Andoni and Indyk, 2006) is then used to cluster these data point vectors to respective clusters efficiently. The optimal number of clusters is estimated iteratively until it is observed that the accuracy has peaked, which in our case was found to be $1/5^{th}$ of the vocabulary of our model. We used the Python implementation of $Scikit - learn$ (Pedregosa et al., 2012) package for applying the nearest neighbor algorithm.

Even after NLP classification, we find that quite a few relevant posts are left out unobserved. So we added one more layer of processing by analyzing the overall emotional content of the reports. We used the AFINN lexicon (Nielsen, 2011) consisting of a collection of 2477 words each with an associated integer value ranging between -5 to +5 representing a transition from a negative to positive sentiment. Modifying the word valence and appending the lexicon with technical words that better represent the associated sentiments was found to produce better results. For example, the LIGO specific application associates terms like 'lockloss' and 'scatter noise' with negative sentiment, while 'new filter installed' would be associated with a positive sentiment.

Our implementation for LIGO is designed so that the users can query for information through a web interface. The stem words in the query are identified, and the vector is projected into the previously modeled word vector space. The nearest neighbor model retrieves the top neighbors for the query vector, which are then displayed as search results on the web interface. Search results are then filtered to check for the presence of the query words in either the title or in its content to weed out false positives. In Fig. 8.3, we have shown a simple search query displayed on the web interface. The different features incorporated in the web interface are described in Section 8.4.

8.3 Gravitational Wave Observatories

The complex nature of GW interferometers experiment requires scientists from multiple domains of expertise to work together and share information. Rigorous commissioning and characterization efforts have been carried over two decades to reach the current level of sensitivity. Efforts to enhance the detector uptime led to increased coincident observation, improving the likelihood of the detection of astrophysical signals. LIGO (The LIGO Scientific Collaboration, 2015), Virgo (Acernese, Agathos, and Agatsuma, 2015), GEO600 (Dooley, Leong, and Adams, 2016), and KAGRA (Aso et al., 2013) archive most of their day-to-day site activities using web interfaces known as logbooks. These may range from installation activities to noise hunting, and mitigation works carried out during the lifespan of the observatory. Although there are site-specific issues, they often encounter problems of a similar nature, and it has often been the case that employing solutions that worked at the other sites turned out to be useful. Also, it is not uncommon to see previously fixed issues to reappear at a later time on timescales varying from a few months to years. This happens due to recurring environmental fluctuations and configuration changes in the detector. Since the current GW detectors aim at coincident detection of events, the joint uptimes of the instruments are crucial. This is more significant because the probability of detection scales linearly with observation time and cubically with the sensitivity of the instrument.

Information extraction and processing of logbook information, as envisaged here, is expected to help with making better decisions about detector maintenance. For example, identifying the subsystems that could get affected during instrument upgrade will be vital for scheduling and coordinating tasks among the sub-groups involved. Similarly, long-term tracking of an issue can be carried out to see if the various overhauling attempts indeed lead to an improvement in performance, which correlates with a lesser number of related posts.

Usually, logbooks are accessible to the public, and valid credentials are required to create or edit them. Additional supporting materials like measurement figures, photographs, sensor data, codes, etc., can be uploaded with these entries. It is also possible to add comments and carry out further discussion on any of the logbook entries. It is mandatory for all the entries to have a title, section, task description, and author details. Details on retrieved information are given below in Table 8.1

Observatory	Logbook Entries	Contributors	Time span	Dictionary Size	Clusters
LIGO Livingston[2]	24351	261	2010-2017	2273	455
LIGO Hanford[3]	24968	237	2010-2017	2713	543
Virgo[4]	34592	660	2010-2017	5026	1005

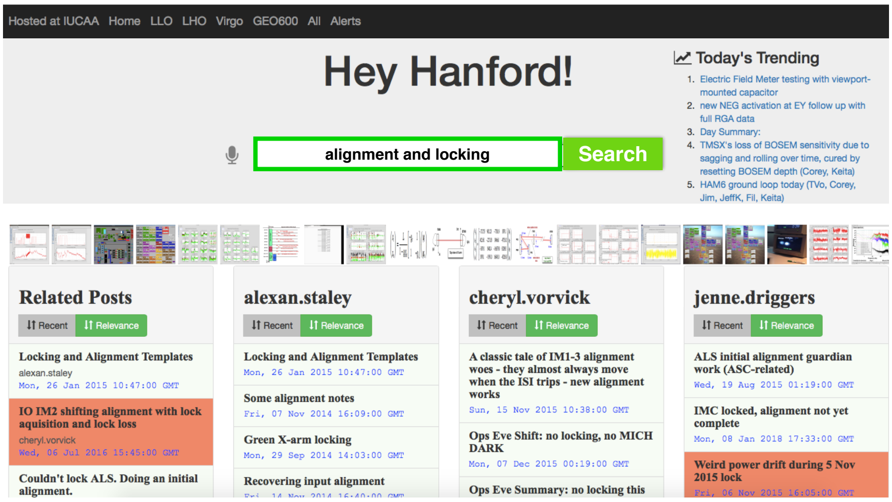
TABLE 8.1: Logbook Details Retrieved from different GW Observatories.

8.4 Hey LIGO Functionalities

We have developed and deployed an open-access NLP-based web application named **Hey LIGO** to support the commissioning and characterization efforts at the GW observatories. It relies on the logbook data recorded since 2010 by scientists specializing in different aspects of the detector. Every query is answered by matching it with most relevant logbook entries sorted as per their closeness to the query term in the word vector space. We further analyze the sentiment of the post and color code so that green indicates a positive outcome and red corresponds to something undesirable in the context of activities carried out at the detector. An image retrieval facility displays the thumbnail of the figures attached to the sorted data, simplifying the knowledge discovery process. Contextual data visualization across multiple detectors is carried out as shown in Fig. 8.7 and Fig. 8.6. This feature lets the user compare and see the trends in the searched keyword across different observatories.

An automatic check for new data entries is done periodically so that the NLP models are regularly updated. We track the volume of discussions happening on various topics and identify and rank the trending issues on a daily basis. Scientists involved with the project will mostly be interested in getting notified about specific problems that correlate with their domain of expertise, so the application only issues alerts to registered participants with matching interests. This targeted delivery will remove clutter and will ensure proper dissemination of information to the concerned people.

Code development is a tedious procedure wherein a significant amount of time is spent on readability and re-usability to benefit a wider research community. Our application makes better use of this idea by auto-detecting and notifying the user about the presence of codes in the searched content. We believe this feature would simplify the procedure involved in result reproduction and its consequent independent verification.



Hey Hanford!

alignment and locking

Today's Trending

1. Electric Field Meter testing with viewport-mounted capacitor
2. new NEG activation at EY follow up with full RGA data
3. Day Summary:
4. TMSX's loss of BOSEM sensitivity due to sagging and rolling over time, cured by resetting BOSEM depth (Corey, Keita)
5. HAM6 ground loop today (TVo, Corey, Jim, JeffK, Fil, Keita)

Related Posts

Locking and Alignment Templates
alexan.staley
Mon, 26 Jan 2015 10:47:00 GMT

IO IM2 shifting alignment with lock aquisition and lock loss
cheryl.vorvick
Wed, 06 Jul 2016 15:45:00 GMT

Couldn't lock ALS. Doing an initial alignment.

alexan.staley

Locking and Alignment Templates
Mon, 26 Jan 2015 10:47:00 GMT

Some alignment notes
Fri, 07 Nov 2014 16:09:00 GMT

Green X-arm locking
Mon, 29 Sep 2014 14:03:00 GMT

Recovering input alignment

cheryl.vorvick

A classic tale of IM1-3 alignment woes - they almost always move when the ISI trips - new alignment works
Sun, 15 Nov 2015 10:38:00 GMT

Ops Eve Shift: no locking, no MICH DARK
Mon, 07 Dec 2015 00:19:00 GMT

Ops Eve Summary: no locking this

jenne.driggers

ALS initial alignment guardian work (ASC-related)
Wed, 19 Aug 2015 01:19:00 GMT

IMC locked, alignment not yet complete
Mon, 08 Jan 2018 17:33:00 GMT

Weird power drift during 5 Nov 2015 lock
Fri, 06 Nov 2015 16:05:00 GMT

FIGURE 8.3: Screenshot of the Hey LIGO Web Interface. Search results are color tagged based on their overall sentiment. Trending posts are identified based on the associated metadata consisting of comments and discussions made within the LIGO community.

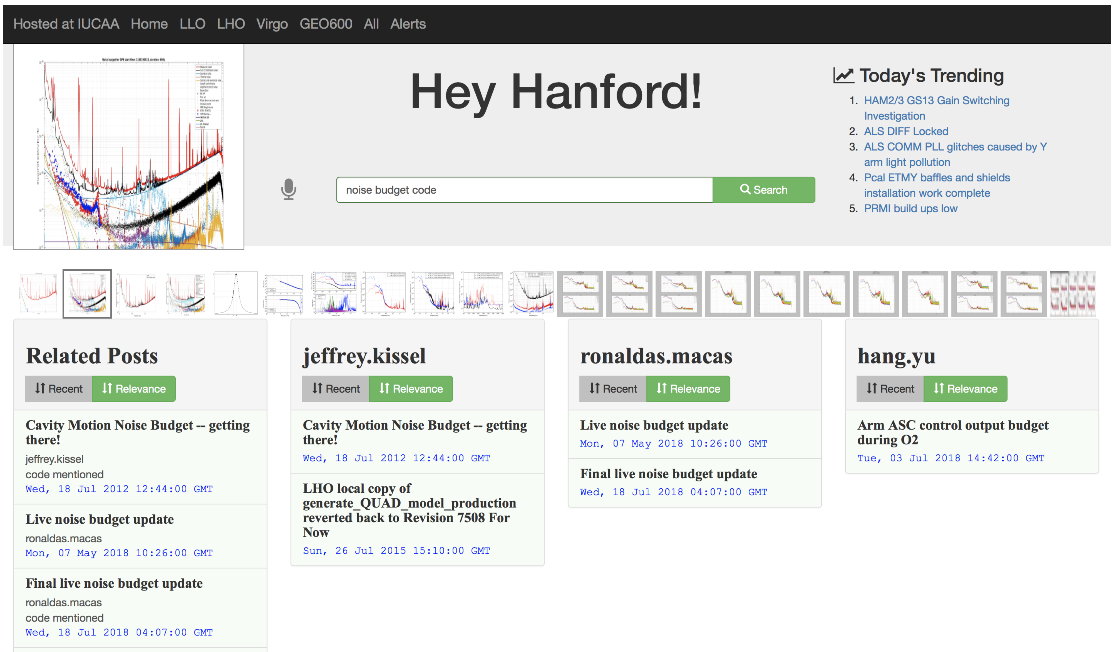
To check the capability of our application, we analyze six months of LIGO Livingston Laboratory (LLO) logbook data (2017 January 1 to June 30) and compare the NLP results with the actual entries. Table 8.2 shows the recovery performance for a specific set of randomly chosen keywords. In most cases, the false alarms occur at the tail end of the search results, which represent the neighbors of least relevance in the k-NN search. These can be removed either by setting a threshold on the similarity distance measure or by post-filtering the results by additionally comparing the content of each post. Currently, we have implemented the process of post-filtering to remove such posts in the final web application. In the future, we are planning to incorporate a mechanism that will make use of feedback received from the users and utilize it to improve the accuracy in retrieving relevant posts.

8.5 Inferring from logbook entries

Once the relevant logbook entries are identified using the techniques mentioned above, their associated metadata can be utilized to obtain several kinds of quantitative information about the topic of interest.

TABLE 8.2: Prominent LLO Logbook Keywords from 2017 January 1 to June 30

Keyword	Log book Entries	Posts Retrieved by NLP Code	
		Total	Relevant
Lock Loss	108	108	89
Earthquake	83	94	80
Charge Measurement	62	65	58
Guardian	55	65	55
Optical Lever	63	61	48
Calibration Lines	55	52	45



Today's Trending

1. HAM2/3 GS13 Gain Switching Investigation
2. ALS DIFF Locked
3. ALS COMM PLL glitches caused by Y arm light pollution
4. Pcal ETMY baffles and shields installation work complete
5. PRMI build ups low

Related Posts

Cavity Motion Noise Budget -- getting there!
jeffrey.kissel
code mentioned
Wed, 18 Jul 2012 12:44:00 GMT

Live noise budget update
ronaldas.macas
Mon, 07 May 2018 10:26:00 GMT

Final live noise budget update
ronaldas.macas
code mentioned
Wed, 18 Jul 2018 04:07:00 GMT

jeffrey.kissel

Cavity Motion Noise Budget -- getting there!
Wed, 18 Jul 2012 12:44:00 GMT

LHO local copy of generate_QUAD_model_production reverted back to Revision 7508 For Now
Sun, 26 Jul 2015 15:10:00 GMT

ronaldas.macas

Live noise budget update
Mon, 07 May 2018 10:26:00 GMT

Final live noise budget update
Wed, 18 Jul 2018 04:07:00 GMT

hang.yu

Arm ASC control output budget during O2
Tue, 03 Jul 2018 14:42:00 GMT

FIGURE 8.4: HeyLIGO search results related to Hanford noise budget. The code automatically identifies the presence of codes and notifies the user.

8.5.1 Trends within Detectors

We briefly compare the trends obtained for a few test search queries and discuss the observed patterns. Although of similar configuration, the effect of various noises on each detector can be of a different nature. Variation in instrumental behavior and environmental effects due to geographical location can also influence the efficiency of implemented mitigation measures. Figuring out such details may positively speed up the commissioning activities of future detectors like LIGO-India.

1. *Installation*. The first plot in Fig. 8.7 shows the trends in posts related to installation work at each of the observatories. Activities picked up momentum in 2010 at LIGO and continued until mid of 2014, after which testing and commissioning tasks started. Advanced Virgo seems to have begun such activities in 2014 and carried them on until the end of 2016.

2. *Jitter Noise*. Jitter noise arising out of laser pointing fluctuations (Martynov, Hall, and Abbott, 2016) is sensitive to cavity alignments and angular mirror motions. It has been partly caused by the pre-stabilized laser (PSL) periscope motion induced by a chiller water flow around PSL's high power oscillator. Various efforts to understand its possible origin and subsequent attempts to subtract it from the data stream are reflected in the increased number of logbook entries at the LIGO Hanford Observatory (LHO) relative to other sites. Commissioners performed online feedforward noise subtraction using auxiliary witness channels that reduced the coupling significantly (Sigg, 2016; Vajente, 2016; Driggers, 2017).

3. *Scattering Noise in LHO and LLO*. Noise from scattered light is one of the factors that limit the sensitivity in the frequency bin from 50 to 200 Hz (see Fig. 8.5), especially during periods of high microseism. An off-axis beam-scattered laser beam could hit a reflecting surface like a camera mirror mount or beam tube and re-enter the cavity. Nonlinear features are seen in the data when this beam picks up resonances from reflecting surfaces, which then get upconverted or phase-modulated by low-frequency seismic-like motion. Its effect at LLO is more pronounced as compared to LHO, as the former is vulnerable to microseismic activity (Ottaway, Fritschel, and Waldman, 2012b).

   Fig. 8.5 shows the effect of acoustic excitation on the 82 Hz peak seen in gravitational wave differential arm motion (DARM) data. The acoustic injections carried out at the LIGO Y-end station are reconstructed using the scatter noise model S(t) (Accadia, 2010b) given by

   $$S(t) = A\sin(\frac{4\,\pi}{\lambda}(n\,Y_{rms}(t) + Y_{ac}(t)))$$

   where $Y_{rms}(t)$ is the ground motion and $Y_{ac}(t) = B\sin(2\pi f_o t)$ is the chamber motion, with (A,n,B) being the tunable parameters. The model parameters using are fine-tuned using pattern search. The scatter noise projection to DARM from ambient motion is obtained by scaling down the

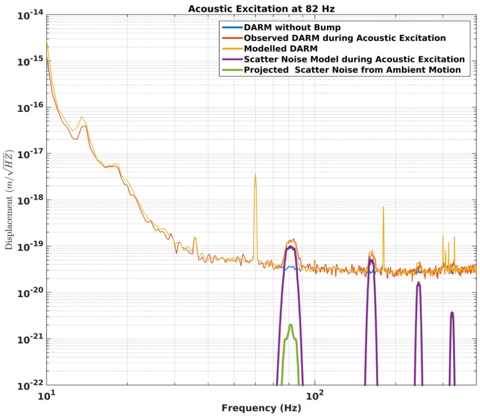


FIGURE 8.5: Effects of scattered light observed at LIGO Livingston. Noise gets amplified and upconverted during periods of high microseism and limits the sensitivity range of the GW detectors.

chamber motion based on the accelerometer signal before and after injection.

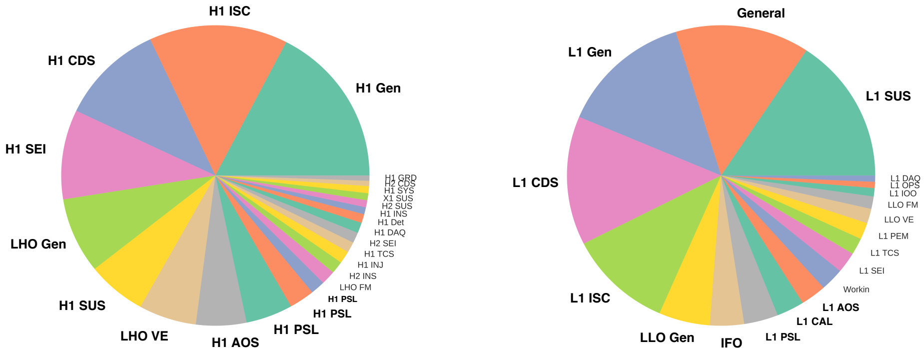


FIGURE 8.6: Pie diagram showing how various LIGO subsystems contribute to the various non-astrophysical glitches seen in LIGO. The acronyms in the diagram are defined as follows: **ISC**: Interferometer Sensing and Controls; **CAL**: Calibration; **AOS**: Auxiliary Optics Support; **SUS**: Suspension; **VE**: Vacuum Electronics; **SEI**: Seismic External Isolation; **CDS**: Control and Design System.

4. *Non-astrophysical Transients* . Glitches often show up in the strain data, leading to false alarms in the various search pipelines that look for astrophysical signals (Abbott, Abbott, and Abbott, 2016). They are often witnessed in auxiliary sensor channels while a few of them have been reported to

cause loss of lock of the interferometer (Mukund, 2017b). The report generation feature of our application provides the following glitch distribution (Fig. 8.6) across multiple subsystems based on their tags in the data. The general operation of all the three detectors has been affected by such transients ever since their beginning of operation (see Fig. 8.7). It is interesting to note the subtle variations in the noise sources between LHO and LLO. The origin of many of them has been studied and reported in the logbooks while a vast majority are still not well understood.

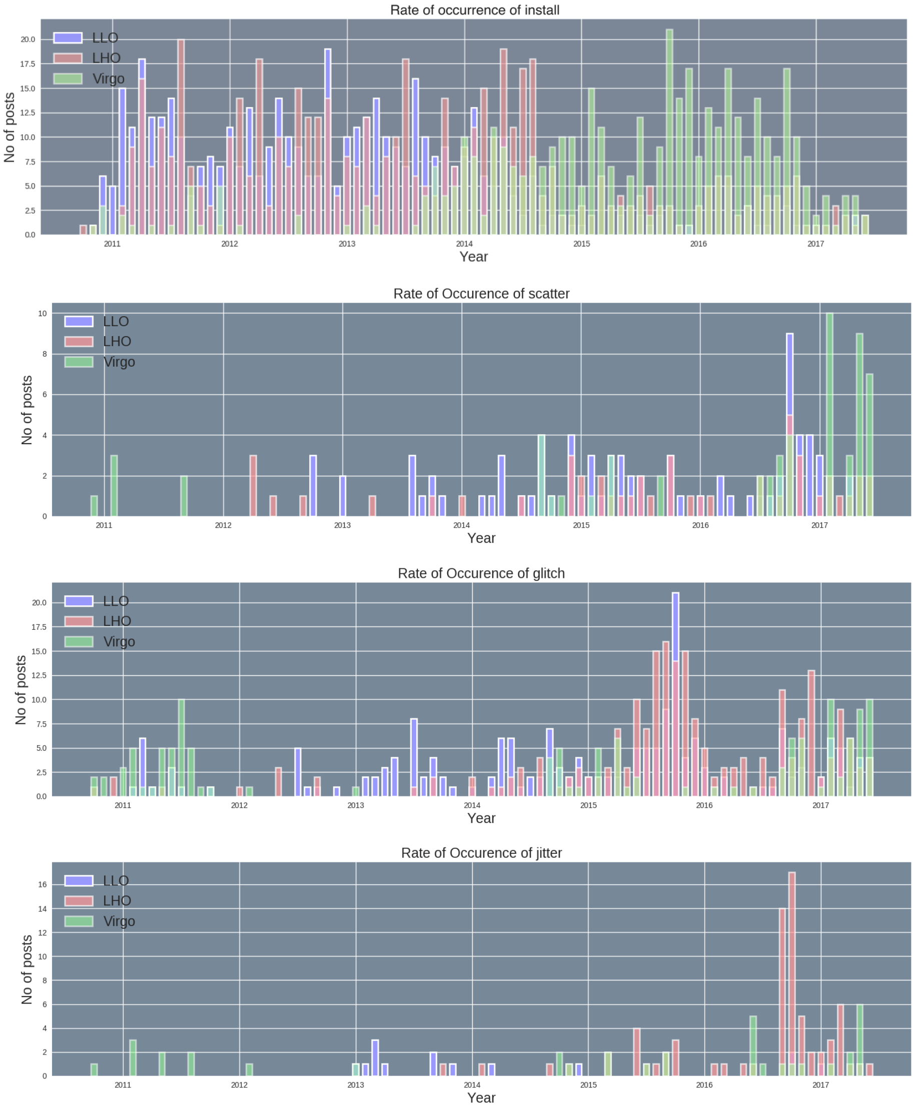


FIGURE 8.7: Rate of occurrence of different keywords in the logbooks of multiple detectors as a function of time.

8.5.2 Visualizing an Observatory as a Complex Network

The behavior of an observatory and the elements that lead to changes in the system behavior can be studied through its representation as a complex network. Complexity is expressed through nodes and links within the network. Here, the nodes can be either subsystems or specific instruments or even subgroups within the observatory and edges between them provide the probability of each one of them being connected to the other as inferred from the logbook entries. We first create a dictionary of subsystem keywords and for each one, find the frequency of their joint occurrence with everyone else. This information is then used to form the adjacency matrix whose diagonal elements are all zero and the off-diagonal value representing the linkage is given by the ratio of joint occurrence frequency divided by the total occurrence of the keyword. The adjacency matrix being non-symmetric leads to a directed graph. The number of incident edges determines the node size while the edge width is given by the associated connection probability. To better aid visualization, we adopt Force Atlas 2 layout (Jacomy et al., 2014) with repulsion being approximated using Barnes Hut optimization (Barnes and Hut, 1986) which is well suited for larger graphs. The interconnectedness information within the observatory revealed through these networks may help in identifying the critical nodes in the system, making it easier to identify the vulnerable connections. These representations could be useful during large-scale repair and maintenance as they reveal the other subsystems that can get affected in the process.

In Fig. 8.8 we show the network connection for a few prominent nodes of the Virgo observatory. It differs from real-world networks regarding its degree distribution (degree refers to the number of edges connected to each node). Sparse networks are characterized by a degree distribution that takes the form of a power law and is commonly seen in biological networks and computer networks (Barabási and Pósfai, 2016). For the case of Virgo network, this distribution deviates from such a power law, indicating a dense connection between the nodes. Further research is needed to analyze the network and study the instrument's robustness to random sub-system failures.

8.6 Discussions and Conclusion

We have demonstrated how information retrieval and recommendation systems could be useful for LIGO-like astronomical observatories. Compared with conventional searches associated with the existing sites, our web application incorporates an NLP-based information retrieval system that can also perform

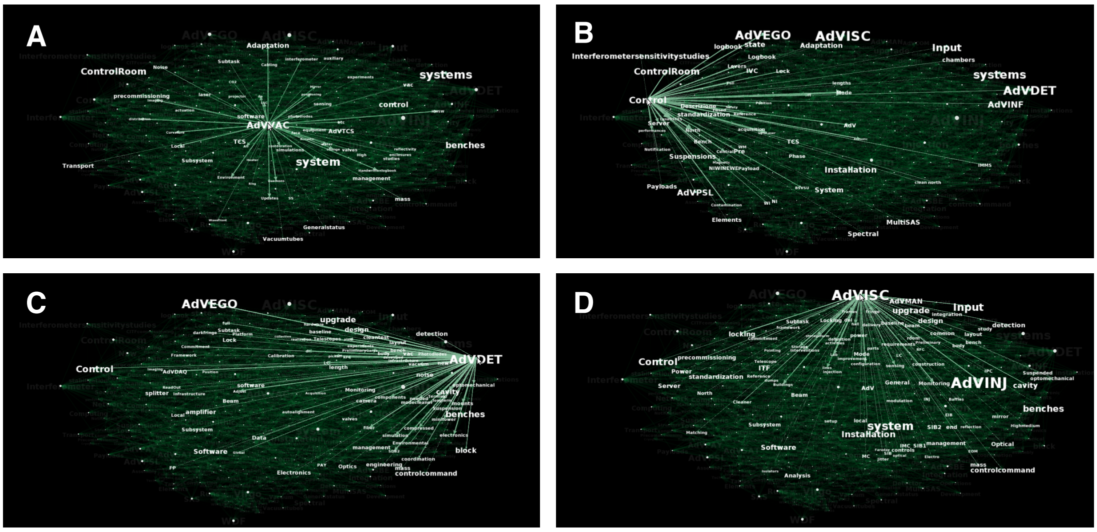


FIGURE 8.8: Network plot of the Virgo detector highlighting the inter-connections between various subsystems. The subplots highlight the directed graph for: (A) vacuum, (B) control system, (C) photodetector, and (D) interferometric sensing and control.

visualization of the user-queried data. Involving a wider science community in big science projects can alleviate some of the issues related to lack of sufficient human resources within the project. The developed interface identifies the major issues based on the discussions done within the LIGO community and recognizes the trending issues. It is plausible that someone involved with the project has already seen and solved these issues. Hence, proper dissemination of information will help technical experts within the collaboration to contribute, leading to overall performance improvement for the instrument.

Coordinated efforts are being undertaken worldwide to carry out electromagnetic follow-up searches looking for counterparts to coalescing binaries sources (Abbott, Abbott, and Abbott, 2016). During the instance of GW candidate event alerts, astronomers may be able to take advantage of our application and know more about the instrument.

Future improvements in the application would be to include capabilities wherein an identified issue will be provided with possible fixes making use of information from past attempts that fixed an identical issue. This would require text abstraction and summarization, which are quite challenging when the data have ample numbers of technical terms. Efforts to add other GW detectors like GEO600 and KAGRA are currently in progress and will enhance the effectiveness of our application.

This kind of system has many potential applications following the commissioning and running of large science projects like the SKA and future LIGO observatories. In this project, the data source was more unstructured and had few tags related to the status of different activities. At present, institutions like SKA South Africa, which is in charge of building the MeerKAT [5] telescope, one of the precursors to SKA, uses a more structured systems like JIRA [6] for issue tracking and log keeping[7]. Scaling our present system to such databases can improve the efficiency of topical modeling. This also enables auto-updates of the learning database as more and more information is logged into the system finally making it robust.

The feature above will become advantageous as the organization includes more participants. The availability of such systems will make the re-usability of information much easier and efficient. This will also help with resolving instrument issues more quickly and more efficiently. Enhanced analytics of key components and recurring issues can help improve the fault tolerance of different subsystems and could provide insights into how to modify them for better performance.

---

[5] `http://www.ska.ac.za/science-engineering/meerkat/`

[6] `https://www.atlassian.com/software/jira`

[7] `https://indico.skatelescope.org/event/402/material/1/6.pdf`

Chapter 9

Conclusions

In this thesis, we analyzed the impact of a certain subset of noise sources for current and next-generation detectors and studied various schemes suitable to minimize their impact on the instrument and its data. We combined techniques from signal processing, global optimization, data science and machine learning to device appropriate strategies which were implemented on real and simulated data. Several applications were developed and installed at the site to help the commissioners and subsystem engineers working on improving the sensitivity and robustness of the instrument.

Chapter 2 focused on optimal sensor placements suitable for Newtonian Noise (NN) subtraction in a variety of seismic field correlations. The optimal sensor array geometry was obtained for each possible seismic field correlation that varied from homogeneous and isotropic seismic field to one that is dominated by scattering. Based on the seismic field correlations obtained from the Hanford site, we showed that sensor noise limited NN subtraction using Weiner filtering is possible using a few seismic surface displacement monitors deployed inside the interferometer building. The global optimizer based on simulated annealing developed for this project is found to be useful in any generic scenarios where the cost function has multiple local minima, and simultaneous minimization of few tens of parameters are needed. Mitigation of NN will improve the sensitivity of the detectors in the lower frequencies where the binaries spend most of their time before merger while in the detector's sensitive frequency band. This will help us in estimating the intrinsic binary parameters more precisely, which will, in turn, enable a much stronger test of General Relativity, a fundamental goal of Modern Physics.

In Chapter 3 we presented the need for early earthquake warning systems and how accurate predictions about the expected ground motion can be achieved using the past earthquake measurements done at the site. Efforts are ongoing to effectively utilize this information and switch controls enabling the detector to stay locked during periods of excess ground motion. If successful, this would increase the coincident observation time crucial for the detection confidence and would lead to decreased sky localization error necessary for electromagnetic (EM) follow up.

We identified a potentially new form of seismic noise arising from hydraulic fracking-related waste water injection and reported its impact in Chapter 4. The difference in triggering mechanism of these induced earthquakes causes even low magnitude earthquakes to throw the interferometer out of lock. We also report sustained seismic activity in these areas for long periods of time even after the stoppage of fracking activities. This study suggests the need for ongoing site selection surveys for the next generation detectors to also access level of induced seismicity as a precautionary measure so that the continuous operation of the interferometer would not be hampered. It also motivates the need for very low latency warning systems that would be required to make filter configuration changes well in advance of the incoming seismic waves.

On the fly system identification would be crucial when the plant characteristics are subject to external fluctuations, and routine manual transfer functions would become impractical as it would decrease the duty cycle leading to decreased coincident observation time. Incorporating the concepts from machine learning into hardware is one area which we would like to explore in the future. Custom hardware (e.g., FPGA) for carrying out Wiener filter and its adaptive versions may eliminate the need for post-offline cleaning leading to rapid parameter estimation of detected astrophysical sources. GW instrument is already at the cutting edge of research in many areas, so improving the sensitivity any further would require the use of state of the art technology in both software and hardware end. But there is still enough scope for improvement in areas like control systems where we are still in the classical control regime and have not adopted robust modern control strategies. The techniques explored in this thesis can be easily integrated into the controllers designed for complex dynamical systems.

Steps involved in the development of an automated lockloss monitoring tool installed at LIGO Livingston is explained in Chapter 6 whose resourcefulness is elucidated through the analysis of several unknown lockloss happened during Advanced LIGO second observation run. To ease the lockloss debugging procedure, we introduced an automated tagging feature which identifies the probable cause of lockloss within five minutes based on nature of channels that get tried prior to the lockloss event. We further analyze the sensitivity of the GW detector to different earthquake parameters and show that it is possible to alert the site staff about an impending lockloss based on analysis of previous such earthquake events. This valuable information can be utilized by the site staff to take the necessary steps to minimize the impact of the seismic event.

Presence of non-astrophysical transients has been a long-standing problem for the compact binary coalescence (CBC) and burst data analysis algorithms. In Chapter 7 we discussed a morphology-based tool developed using machine learning and demonstrated its robustness by classifying the triggers found in Advanced LIGO's first observation data. The distinct morphology of these triggers is due to the difference in their origin, hence by understanding the relative population of these events we can infer about the various environmental couplings existing within the interferometer. Further development has lead to a probabilistic version of the classifier whose outputs can now be incorporated into various transient search pipelines and can be used either in conjunction with or as an alternative to the traditional chi-square based veto procedure. This machine learning based method can be incorporated into the existing template based matched filtering and unmodeled burst signal search pipelines to effectively veto the non-astrophysical transients from the strain data. Reducing the background would increase the confidence of true coincident detections and would be an effective tool to characterize the single detector events. Thus, these techniques can help us in detecting many more events, which would otherwise be discarded as low-significance triggers.

Rapid dissemination of information is needed to speed up the commissioning, characterization and maintenance activities happening at the working at the detector site. The need for such a system also becomes relevant in the context of projects like LIGO-India which needs to be completed in a short timescale without compromise in sensitivity. Using ideas based on data mining and natural language processing as described in Chapter 8 we were able to utilize decades of logbook entries covering various aspects of the instrument and create a knowledge engine equipped to answer queries about the functioning of the instrument. Analyzing the entire interferometer as a complex network led us to reveal the subtle interconnection between various multiple subsystems which could provide insights into the possible effects of certain critical subsystem failures onto others. The utility of the system based on feedback received from the users has to lead us to expand the knowledge engine to include the logbook information that is separately maintained for each subsystem like the ones related to controls, instrumentation and data analysis. The generic nature of the technique makes it useful for other astronomical projects such SKA, TMT which also involves integration of multiple subsystems similar to GW observatories.

Through this thesis, we have explored different techniques to characterize multiple noise sources affecting the GW detectors and devised ways to tackle them in an automated fashion. We leveraged on optimization and machine learning algorithms to create applications that improve the robustness and quality of the collected data. Improved duty cycle and the availability of cleaned data would be of utmost importance to make frequent high confidence signal detections from several kinds of astrophysical and cosmological GW sources. Finally, we hope that the experience gained and the tools developed in this thesis will directly benefit the current, upcoming and next-generation GW detectors, especially the LIGO-India project.

Bibliography

http://www.iucaa.ernet.in/~nspp/dbnn.html.

Aasi, J. et al. (2015). “Characterization of the LIGO detectors during their sixth science run”. In: *Classical Quantum Gravity* 32.11, p. 115012. URL: http://stacks.iop.org/0264-9381/32/i=11/a=115012.

Abadi, Martín et al. (2016). “TensorFlow: A System for Large-scale Machine Learning”. In: *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation*. OSDI’16. USENIX Association, pp. 265–283. ISBN: 978-1-931971-33-1. URL: http://dl.acm.org/citation.cfm?id=3026877.3026899.

Abadie, J. et al. (June 2012). “All-sky search for gravitational-wave bursts in the second joint LIGO-Virgo run”. In: *Phys. Rev. D* 85 (12), p. 122007. DOI: 10.1103/PhysRevD.85.122007. URL: http://link.aps.org/doi/10.1103/PhysRevD.85.122007.

Abbott, B et al. (2004). “First upper limits from LIGO on gravitational wave bursts”. In: *Phys. Rev. D* 69.10, p. 102001. URL: https://link.aps.org/doi/10.1103/PhysRevD.69.102001.

Abbott, B. P., R. Abbott, T. D. Abbott, et al. (2016). “Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914”. In: *Classical and Quantum Gravity* 33.13, p. 134001. URL: http://stacks.iop.org/0264-9381/33/i=13/a=134001.

Abbott, B. P., R. Abbott, T. D. Abbott, et al. (2016). “Localization and Broadband Follow-up of the Gravitational-wave Transient GW150914”. In: *The Astrophysical Journal Letters* 826.1, p. L13. URL: http://stacks.iop.org/2041-8205/826/i=1/a=L13.

Abbott, B. P. et al. (2016a). “Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914”. In: *Classical Quantum Gravity* 33.13, p. 134001. URL: http://stacks.iop.org/0264-9381/33/i=13/a=134001.

— (June 2016b). “GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence”. In: *Phys. Rev. Lett.* 116 (24), p. 241103. DOI: 10.1103/PhysRevLett.116.241103. URL: http://link.aps.org/doi/10.1103/PhysRevLett.116.241103.

Abbott, B. P. et al. (Feb. 2016c). "Observation of Gravitational Waves from a Binary Black Hole Merger". In: *Phys. Rev. Lett.* 116 (6), p. 061102. DOI: 10.1103/PhysRevLett.116.061102. URL: http://link.aps.org/doi/10.1103/PhysRevLett.116.061102.

Abbott, R, R Adhikari, et al. (2002). "Seismic isolation for Advanced LIGO". In: *Classical and Quantum Gravity* 19.7, p. 1591. URL: http://stacks.iop.org/0264-9381/19/i=7/a=349.

Abbott, B. P. et al. (2016). "GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence". In: *Phys. Rev. Lett.* 116 (24), p. 241103. DOI: 10.1103/PhysRevLett.116.241103. URL: http://link.aps.org/doi/10.1103/PhysRevLett.116.241103.

Abbott, B. P. et al (2016). "Observation of Gravitational Waves from a Binary Black Hole Merger". In: *Phys. Rev. Lett.* 116 (6), p. 061102. DOI: 10.1103/PhysRevLett.116.061102. URL: http://link.aps.org/doi/10.1103/PhysRevLett.116.061102.

Abbott, B. P. et al. (2017a). "GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2". In: *Phys. Rev. Lett.* 118 (22), p. 221101. DOI: 10.1103/PhysRevLett.118.221101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.118.221101.

Abbott, B. P. et al. (2017b). "GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence". In: *Phys. Rev. Lett.* 119 (14), p. 141101. DOI: 10.1103/PhysRevLett.119.141101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.141101.

— (2017c). "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral". In: *Phys. Rev. Lett.* 119 (16), p. 161101. DOI: 10.1103/PhysRevLett.119.161101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.161101.

Abernathy, Matt et al. (2011). "Einstein gravitational wave Telescope conceptual design study". In:

Abraham, S. et al. (Jan. 2012). "A photometric catalogue of quasars and other point sources in the Sloan Digital Sky Survey". In: *MNRAS* 419, pp. 80–94. DOI: 10.1111/j.1365-2966.2011.19674.x. arXiv: 1011.2173 [astro-ph.IM].

Accadia, T et al. (2010a). "Noise from scattered light in Virgo's second science run data". In: *Classical Quantum Gravity* 27.19, p. 194011. URL: http://stacks.iop.org/0264-9381/27/i=19/a=194011.

— (2010b). "Noise from scattered light in Virgo's second science run data". In: *Classical and Quantum Gravity* 27.19, p. 194011. URL: http://stacks.iop.org/0264-9381/27/i=19/a=194011.

Acernese, F., M. Agathos, K. Agatsuma, et al. (2015). “Advanced Virgo: a second-generation interferometric gravitational wave detector”. In: *Classical and Quantum Gravity* 32.2, p. 024001. URL: `http://stacks.iop.org/0264-9381/32/i=2/a=024001`.

Adam Day, 2016. `https://cqgplus.com/2016/06/06/how-do-we-know-ligo-detected-gravitational-waves/`.

Adams, A. and A. Woolley (1994). “Hubble classification of galaxies using neural networks”. In: *Vistas Astron.* 38, pp. 273–280. DOI: `10.1016/0083-6656(94)90037-X`.

Allen, Bruce et al. (2012). “FINDCHIRP: An algorithm for detection of gravitational waves from inspiraling compact binaries”. In: *Phys. Rev. D* 85 (12), p. 122006. DOI: `10.1103/PhysRevD.85.122006`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.85.122006`.

Andoni, Alexandr and Piotr Indyk (2006). “Near-Optimal Hashing Algorithms for Approximate Nearest Neighbor in High Dimensions”. In: *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*. FOCS ’06. Washington, DC, USA: IEEE Computer Society, pp. 459–468. ISBN: 0-7695-2720-5. DOI: `10.1109/FOCS.2006.49`. URL: `http://dx.doi.org/10.1109/FOCS.2006.49`.

Areeda, J and J Smith. *LigoDV-Web*. `https://ldvw.ligo.caltech.edu/ldvw/view`.

Aso, Yoichi et al. (Aug. 2013). “Interferometer design of the KAGRA gravitational wave detector”. In: *Phys. Rev. D* 88 (4), p. 043007. DOI: `10.1103/PhysRevD.88.043007`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.88.043007`.

Baker, J. et al. (2002). “Modeling gravitational radiation from coalescing binary black holes”. In: *Phys. Rev. D* 65 (12), p. 124012. DOI: `10.1103/PhysRevD.65.124012`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.65.124012`.

Baker, John, Manuela Campanelli, and Carlos O. Lousto (2002). “The Lazarus project: A pragmatic approach to binary black hole evolutions”. In: *Phys. Rev. D* 65 (4), p. 044001. DOI: `10.1103/PhysRevD.65.044001`. URL: `http://link.aps.org/doi/10.1103/PhysRevD.65.044001`.

Bar-Joseph, Ziv, David K. Gifford, and Tommi S. Jaakkola (2001). “Fast optimal leaf ordering for hierarchical clustering”. In: *Bioinformatics* 17.suppl 1, S22–S29. DOI: `10.1093/bioinformatics/17.suppl_1.S22`. URL: `http://bioinformatics.oxfordjournals.org/content/17/suppl_1/S22.abstract`.

Barabási, Albert-László and Márton Pósfai (2016). *Network science*. Cambridge: Cambridge University Press. ISBN: 9781107076266 1107076269. URL: http://barabasi.com/networksciencebook/.

Barnes, J. E. and P. Hut (1986). "A hierarchical O(n-log-n) force calculation algorithm". In: *Nature* 324, p. 446. URL: http://dx.doi.org/10.1038/324446a0.

Bayes, Thomas (1763). "A Letter from the Late Reverend Mr. Thomas Bayes, F. R. S. to John Canton, M. A. and F. R. S." In: *Phil. Trans.* 53, pp. 269–271. DOI: 10.1098/rstl.1763.0044. URL: http://rstl.royalsocietypublishing.org/content/53/269.short.

Behrens, Clifford A and Devasis Bassu (Dec. 2006). *Information retrieval and text mining using distributed latent semantic indexing*. US Patent 7,152,065. URL: https://patents.google.com/patent/US7152065.

Beker, Mark et al. (2016). "Innovations in seismic sensors driven by the search for gravitational waves". In: *The Leading Edge* 35.7, pp. 590–593. DOI: 10.1190/tle35070590.1.

Beker, MG et al. (2011). "Improving the sensitivity of future GW observatories in the 1–10 Hz band: Newtonian and seismic noise". In: *General Relativity and Gravitation* 43.2, pp. 623–656.

Bengio, Yoshua et al. (Mar. 2003). "A Neural Probabilistic Language Model". In: *J. Mach. Learn. Res.* 3, pp. 1137–1155. ISSN: 1532-4435. URL: http://dl.acm.org/citation.cfm?id=944919.944966.

Berger, J., P. Davis, and G. Ekström (2004). "Ambient Earth Noise: A survey of the Global Seismographic Network". In: *Journal of Geophysical Research* 109, B11307.

Bhagat, M. et al. (2009). "Investigating neuromagnetic brain responses against chromatic flickering stimuli by wavelet entropies". In: *PloS one* 4.9, e7173. URL: https://doi.org/10.1371/journal.pone.0007173.

Biscans, Sebastien et al. (2018). "Control strategy to limit duty cycle impact of earthquakes on the LIGO gravitational-wave detectors." In: *Classical and Quantum Gravity*. URL: http://iopscience.iop.org/10.1088/1361-6382/aaa4aa.

Biswas, R. et al. (Sept. 2013). "Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data". In: 88.6, 062003, p. 062003. DOI: 10.1103/PhysRevD.88.062003.

Blair, David G et al. (2012). *Advanced gravitational wave detectors*. Cambridge University Press.

Bonnefoy-Claudet, S., F. Cotton, and P.-Y. Bard (2006). "The nature of noise wavefield and its applications for site effects studies; A literature review". In: *Earth-Science Rev.* 79, p. 205.

Bose, Sukanta et al. (2016). "Towards mitigating the effect of sine-Gaussian noise transients on searches for gravitational waves from compact binary coalescences". In: *Phys. Rev. D* 94 (12), p. 122004. DOI: `10.1103/PhysRevD.94.122004`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.94.122004`.

Breiman, Leo (2001). "Random Forests". In: *Machine Learning* 45.1, pp. 5–32. ISSN: 1573-0565. DOI: `10.1023/A:1010933404324`. URL: `http://dx.doi.org/10.1023/A:1010933404324`.

Burrows, A. et al. (2006). "A new mechanism for core-collapse supernova explosions". In: *Astrophys. J.* 640, pp. 878–890. DOI: `10.1086/500174`.

Canton, T Dal et al. (2014). "Effect of sine-Gaussian glitches on searches for binary coalescence". In: *Classical Quantum Gravity* 31.1, p. 015016. URL: `http://stacks.iop.org/0264-9381/31/i=1/a=015016`.

Castagna, John P and MM Backus (1993). "AVO analysis—Tutorial and review". In: *Offset-dependent reflectivity: Theory and practice of AVO analysis: SEG Investigations in Geophysics* 8, pp. 3–36.

Caves, C. M. (1980). "Quantum-mechanical noise in an interferometer". In: *Phys. Rev. D* 23, p. 1693.

Cella, G (2000). "Off-line Subtraction of Seismic Newtonian Noise". In: *Recent Developments in General Relativity*. Springer, pp. 495–503.

Cessaro, R. K. (1994). "Sources of Primary and Secondary Microseisms". In: *BSSA* 84, p. 142.

Chang, Chih-Chung and Chih-Jen Lin (2011). "LIBSVM: a library for support vector machines". In: *ACM Transactions on Intelligent Systems and Technology* 2.3, p. 1. URL: `http://doi.acm.org/10.1145/1961189.1961199`.

Cointet, Jean-Philippe and Camille Roth (2010). "Local Networks, Local Topics: Structural and Semantic Proximity in Blogspace." In: *ICWSM*.

Cortes, Corinna and Vladimir Vapnik (1995). "Support-Vector Networks". In: *Machine Learning* 20.3, p. 273. DOI: `10.1007/BF00994018`. URL: `http://dblp.uni-trier.de/rec/bib/journals/ml/CortesV95`.

Coughlin, M et al. (2017a). "Limiting the effects of earthquakes on gravitational-wave interferometers". In: *Classical and Quantum Gravity* 34.4, p. 044004. URL: `http://stacks.iop.org/0264-9381/34/i=4/a=044004`.

Coughlin, Michael et al. (2015). "Real-time earthquake warning for astronomical observatories". English. In: *Experimental Astronomy* 39.2, pp. 387–404. ISSN: 0922-6435. DOI: `10.1007/s10686-015-9454-1`. URL: `http://dx.doi.org/10.1007/s10686-015-9454-1`.

Coughlin, Michael et al. (2016). "Towards a first design of a Newtonian-noise cancellation system for Advanced LIGO". In: *Classical and Quantum Gravity* 33.24, p. 244001.

Coughlin, Michael et al. (2017b). “Limiting the effects of earthquakes on gravitational-wave interferometers”. In: *Classical and Quantum Gravity* 34.4, p. 044004. URL: http://stacks.iop.org/0264-9381/34/i=4/a=044004.

Creighton, T. (2008). “Tumbleweeds and airborne gravitational noise sources for LIGO”. In: *Class. Quantum Grav.* 25, p. 125011.

Damour, Thibault and Alexander Vilenkin (2000). “Gravitational Wave Bursts from Cosmic Strings”. In: *Phys. Rev. Lett.* 85 (18), pp. 3761–3764. DOI: 10.1103/PhysRevLett.85.3761. URL: http://link.aps.org/doi/10.1103/PhysRevLett.85.3761.

Dhurandhar, S. V. and B. S. Sathyaprakash (1994). “Choice of filters for the detection of gravitational waves from coalescing binaries. II. Detection in colored noise”. In: *Phys. Rev. D* 49, p. 1707.

Diersen, Steve et al. (2011). “Classification of Seismic Windows Using Artificial Neural Networks”. In: *Procedia Computer Science* 4.Supplement C. Proceedings of the International Conference on Computational Science, ICCS 2011, pp. 1572 –1581. ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2011.04.170. URL: http://www.sciencedirect.com/science/article/pii/S1877050911002286.

Dooley, K L, J R Leong, T Adams, et al. (2016). “GEO 600 and the GEO-HF upgrade program: successes and challenges”. In: *Classical and Quantum Gravity* 33.7, p. 075009. URL: http://stacks.iop.org/0264-9381/33/i=7/a=075009.

Douglas, J. (Apr. 2003). “Earthquake ground motion estimation using strong-motion records: a review of equations for the estimation of peak ground acceleration and response spectral ordinates”. In: *Earth Science Reviews* 61, pp. 43–104. DOI: 10.1016/S0012-8252(02)00112-5.

Driggers, J.C. et al. (2012). “Active noise cancellation in a suspended interferometer”. In: *Review of Scientific Instruments* 83.2, pp. 024501–024501.

Driggers, Jenne (2017). *LHO alog 34631*. https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=34631.

Driggers, Jennifer C., Jan Harms, and Rana X. Adhikari (2012). “Subtraction of Newtonian noise using optimized sensor arrays”. In: *Phys. Rev. D* 86 (10), p. 102001. DOI: 10.1103/PhysRevD.86.102001. URL: http://link.aps.org/doi/10.1103/PhysRevD.86.102001.

Eaton, John W. et al. (2015). *GNU Octave Version 4.0.0 Manual: A High-Level Interactive Language for Numerical Computations*. URL: http://www.gnu.org/software/octave/doc/interpreter.

Einstein, Albert (1916). “The foundation of the general theory of relativity”. In: *Annalen Phys.* 14, pp. 769–822.

Ester, Martin et al. (1996). "A density-based algorithm for discovering clusters in large spatial databases with noise." In: *Kdd*. Vol. 96. 34, pp. 226–231.

Evans, D.M. (Jan. 1966). "Denver area earthquakes and the Rocky Mountain Arsenal disposal well". In: *Mt. Geol.; (United States)*.

F Acernese et al (2015). "Advanced Virgo: a second-generation interferometric gravitational wave detector". In: *Classical and Quantum Gravity* 32.2, p. 024001. URL: `http://stacks.iop.org/0264-9381/32/i=2/a=024001`.

Fan, Jianqing, Fang Han, and Han Liu (2014). "Challenges of Big Data analysis". In: *National Science Review* 1.2, pp. 293–314. DOI: `10.1093/nsr/nwt032`. eprint: `/oup/backfile/content_public/journal/nsr/1/2/10.1093_nsr_nwt032/3/nwt032.pdf`. URL: `http://dx.doi.org/10.1093/nsr/nwt032`.

Frakes, W. B. (1992). In: *Information Retrieval*. Ed. by William B. Frakes and Ricardo Baeza-Yates. Upper Saddle River, NJ, USA: Prentice-Hall, Inc. Chap. Stemming Algorithms, pp. 131–160. ISBN: 0-13-463837-9. URL: `http://dl.acm.org/citation.cfm?id=129687.129695`.

Friedrich, A., F. Krüger, and K. Klinge (1998). "Oceangenerated microseismic noise located with the Gräfenberg array". In: *J. Seism.* 2, p. 47.

Gaite, B et al. (2015). "A 3-D shear velocity model of the southern North American and Caribbean plates from ambient noise and earthquake tomography". In: *Solid Earth* 6.1, p. 271.

Goldberg, Y. and O. Levy (Feb. 2014). "word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method". In: *ArXiv e-prints*. arXiv: `1402.3722 [cs.CL]`. URL: `https://arxiv.org/abs/1402.3722`.

González, Gabriela (2000). "Suspensions thermal noise in the LIGO gravitational wave detector". In: *Classical and Quantum Gravity* 17.21, p. 4409.

Gretzel, Ulrike et al. (2004). "Tell me who you are and I will tell you where to go: Use of travel personalities in destination recommendation systems". In: *Information Technology & Tourism* 7.1, pp. 3–12. DOI: `doi:10.3727/1098305042781129`. URL: `https://www.ingentaconnect.com/content/cog/itt/2004/00000007/00000001/art00001`.

Grote H. for the LIGO Scientific Collaboration (2010). "The GEO 600 status". In: *Class. Quantum Grav.* 27, p. 084003.

Gustavsen, Bjorn (2006). "Relaxed vector fitting algorithm for rational approximation of frequency domain responses". In: *Signal Propagation on Interconnects, 2006. IEEE Workshop on*. IEEE, pp. 97–100.

Gustavsen, Bjorn and Adam Semlyen (1999). "Rational approximation of frequency domain responses by vector fitting". In: *IEEE Transactions on power delivery* 14.3, pp. 1052–1061.

Han, Jiawei, Micheline Kamber, and Jian Pei (2011). *Data Mining: Concepts and Techniques*. 3rd. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. ISBN: 0123814790, 9780123814791. URL: `https://www.sciencedirect.com/science/book/9780123814791`.

Harms, J. et al. (2009). “Simulation of underground gravity gradients from stochastic seismic fields”. In: *Phys. Rev. D* 80, p. 122001.

Harms, J et al. (2015). “Transient gravity perturbations induced by earthquake rupture”. In: *Geophysical Journal International* 201.3, pp. 1416–1425.

Harms, Jan (2015). “Terrestrial Gravity Fluctuations”. In: *Living Reviews in Relativity* 18.1, p. 3. ISSN: 1433-8351. DOI: `10.1007/lrr-2015-3`. URL: `https://doi.org/10.1007/lrr-2015-3`.

Harms, Jan et al. (2010). “Characterization of the seismic environment at the Sanford Underground Laboratory, South Dakota”. In: *Classical and Quantum Gravity* 27.22, p. 225011.

Harris, David and Sarah Harris (2012). *Digital design and computer architecture*. Elsevier. URL: `https://textbooks.elsevier.com/web/product_details.aspx?isbn=9780123944245`.

Hartigan, J. A. and M. A. Wong (1979). “A k-means clustering algorithm”. In: *JSTOR: Applied Statistics* 28.1, pp. 100–108.

Hassan, Ahmed et al. (2009). “Content based recommendation and summarization in the blogosphere”. In: *Ann Arbor* 1001, p. 48109. URL: `https://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/203`.

Haubrich, R. A., W. H. Munk, and F. E. Snodgrass (1963). “Comparative spectra of microseisms and swell”. In: *BSSA* 53, p. 27.

Haykin, Simon S (2008). *Adaptive filter theory*. Pearson Education India.

Hsieh, Paul A and John D Bredehoeft (1981). “A reservoir analysis of the Denver earthquakes: A case of induced seismicity”. In: *Journal of Geophysical Research: Solid Earth* 86.B2, pp. 903–920.

Hu, Han et al. (2014). “Toward scalable systems for big data analytics: A technology tutorial”. In: *IEEE Access* 2, pp. 652–687. URL: `http://ieeexplore.ieee.org/document/6842585/`.

Huang, Chu-Ren et al. (2007). “Rethinking Chinese Word Segmentation: Tokenization, Character Classification, or Wordbreak Identification”. In: *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*. ACL ’07. Prague, Czech Republic: Association for Computational Linguistics, pp. 69–72. URL: `http://dl.acm.org/citation.cfm?id=1557769.1557791`.

Huang, Yiteng, Jacob Benesty, and Jingdong Chen (2006). *Acoustic MIMO signal processing*. Springer Science & Business Media.

Hughes, S. A. and K. S. Thorne (1998a). "Seismic gravity-gradient noise in interferometric gravitational-wave detectors". In: *Phys. Rev. D* 58, p. 122002.

Hughes, Scott A and Kip S Thorne (1998b). "Seismic gravity-gradient noise in interferometric gravitational-wave detectors". In: *Physical Review D* 58.12, p. 122002.

Ingber, Lester (2000). "Adaptive simulated annealing (ASA): Lessons learned". In: *CoRR* cs.MS/0001018. URL: `http://arxiv.org/abs/cs.MS/0001018`.

J Aasi et al. (2015). "Advanced LIGO". In: *Classical and Quantum Gravity* 32.7, p. 074001. URL: `http://stacks.iop.org/0264-9381/32/i=7/a=074001`.

Jacomy, Mathieu et al. (June 2014). "ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software". In: *PLOS ONE* 9.6, pp. 1–12. DOI: `10.1371/journal.pone.0098679`. URL: `https://doi.org/10.1371/journal.pone.0098679`.

Jayasree, T, D Devaraj, and R Sukanesh (2009). "Classification of Transients using WaveletBased Entropy and Radial Basis NeuralNetworks". In: *Int. J. Comp. Electrical Engineering* 1.5, p. 590.

Jia, Yongna and Jianwei Ma (2017). "What can machine learning do for seismic data processing? An interpolation application". In: *GEOPHYSICS* 82.3, pp. V163–V177. DOI: `10.1190/geo2016-0300.1`.

Kaggle Website. `https://www.kaggle.com`. Accessed: 2018-06-15.

Kerzendorf, W. E. (May 2017). "Knowledge discovery through text-based similarity searches for astronomy literature". In: *ArXiv e-prints*. arXiv: `1705.05840 [cs.DL]`. URL: `https://arxiv.org/abs/1705.05840`.

Khan, Nawsher et al. (2014). "Big data: survey, technologies, opportunities, and challenges". In: *The Scientific World Journal* 2014. URL: `http://dx.doi.org/10.1155/2014/712826`.

Kingma, Diederik P and Jimmy Ba (2014). "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980*.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi (1983). "Optimization by Simulated Annealing". In: *Science* 220.4598, pp. 671–680. DOI: `10.1126/science.220.4598.671`. URL: `http://science.sciencemag.org/content/220/4598/671`.

Klimenko, Sergey and Guenakh Mitselmakher (2004). "A wavelet method for detection of gravitational wave bursts". In: *Classical Quantum Gravity* 21.20, S1819. URL: `http://stacks.iop.org/0264-9381/21/i=20/a=025`.

Klimenko, Sergey et al. (2008). "A coherent method for detection of gravitational wave bursts". In: *Classical Quantum Gravity* 25.11, p. 114029. URL: `http://stacks.iop.org/0264-9381/25/i=11/a=114029`.

Komatitsch, D. and J. P. Vilotte (1998). “The spectral-element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures”. In: *bssa* 88.2, pp. 368–392.

Kong, Qingkai et al. (2016). “MyShake: A smartphone seismic network for earthquake early warning and beyond”. In: *Science Advances* 2.2. DOI: `10.1126/sciadv.1501055`. eprint: `http://advances.sciencemag.org/content/2/2/e1501055.full.pdf`. URL: `http://advances.sciencemag.org/content/2/2/e1501055`.

Kortström, Jari, Marja Uski, and Timo Tiira (2016). “Automatic classification of seismic events within a regional seismograph network”. In: *Computers & Geosciences* 87.Supplement C, pp. 22 –30. ISSN: 0098-3004. DOI: `https://doi.org/10.1016/j.cageo.2015.11.006`. URL: `http://www.sciencedirect.com/science/article/pii/S0098300415300832`.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton (2012). “ImageNet Classification with Deep Convolutional Neural Networks”. In: *Proceedings of the 25th International Conference on Neural Information Processing Systems*. NIPS’12. Lake Tahoe, Nevada: Curran Associates Inc., p. 1097. URL: `http://dl.acm.org/citation.cfm?id=2999134.2999257`.

Kuyuk, H. S. et al. (Jan. 2011). “An unsupervised learning algorithm: application to the discrimination of seismic events and quarry blasts in the vicinity of Istanbul”. In: *Natural Hazards and Earth System Sciences* 11, pp. 93–100. DOI: `10.5194/nhess-11-93-2011`.

Kuyuk, H. S. et al. (2015). “Automatic earthquake confirmation for early warning system”. In: *Geophysical Research Letters* 42.13. 2015GL063881, pp. 5266–5273. ISSN: 1944-8007. DOI: `10.1002/2015GL063881`. URL: `http://dx.doi.org/10.1002/2015GL063881`.

Kuyuk, H.S. and RM. Allen (2013). “A global approach to provide magnitude estimates for earthquake early warning”. In: *Geophysical Research Letters* 40.

Laan, Mark J Van der, Eric C Polley, and Alan E Hubbard (2007). “Super learner”. In: *Statistical applications in genetics and molecular biology* 6.1.

Laine, A. and J. Fan (Nov. 1993). “Texture Classification by Wavelet Packet Signatures”. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 15.11, pp. 1186–1191. ISSN: 0162-8828. DOI: `10.1109/34.244679`. URL: `http://dx.doi.org/10.1109/34.244679`.

LaValle, Steve et al. (2011). “Big data, analytics and the path from insights to value”. In: *MIT sloan management review* 52.2, p. 21. URL: `https://sloanreview.mit.edu/article/big-data-analytics-and-the-path-from-insights-to-value/`.

Leskovec, Jure, Anand Rajaraman, and Jeffrey David Ullman (2014). *Mining of massive datasets*. Cambridge University Press. URL: `https://www.cambridge.org/core/books/mining-of-massive-datasets/C1B37BA2CBB8361B94FDD1C6F4E47922`.

Li, Yitan et al. (2015). "Word Embedding Revisited: A New Representation Learning and Explicit Matrix Factorization Perspective". In: *Proceedings of the 24th International Conference on Artificial Intelligence*. IJCAI'15. Buenos Aires, Argentina: AAAI Press, pp. 3650–3656. ISBN: 978-1-57735-738-4. URL: `http://dl.acm.org/citation.cfm?id=2832747.2832758`.

Liang, Ting-Peng, Hung-Jen Lai, and Yi-Cheng Ku (2006). "Personalized content recommendation and user satisfaction: Theoretical synthesis and empirical findings". In: *Journal of Management Information Systems* 23.3, pp. 45–70. URL: `http://www.jstor.org/stable/40398855`.

Ling, Wang et al. (2015). "Two/Too Simple Adaptations of Word2Vec for Syntax Problems." In: *HLT-NAACL*, pp. 1299–1304. URL: `http://www.aclweb.org/anthology/N15-1142`.

Mahalanobis, Prasanta Chandra (1936). "On the generalized distance in statistics". In: *Proc. Natl. Inst. Sci. India* 2, p. 49.

Markov, Zdravko and Daniel T Larose (2007). *Data mining the Web: uncovering patterns in Web content, structure, and usage*. John Wiley & Sons. URL: `http://onlinelibrary.wiley.com/book/10.1002/0470108096`.

Martynov, D. V., E. D. Hall, B. P. Abbott, et al. (June 2016). "Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy". In: *Phys. Rev. D* 93 (11), p. 112004. DOI: `10.1103/PhysRevD.93.112004`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.93.112004`.

Matichard, F et al. (2015). "Seismic isolation of Advanced LIGO: Review of strategy, instrumentation and performance". In: *Classical and Quantum Gravity* 32.18, p. 185003. URL: `http://stacks.iop.org/0264-9381/32/i=18/a=185003`.

MATLAB and Wavelet Toolbox Release 2013a (2013). *Version 8.1.0.604*. Natick, Massachusetts: The MathWorks Inc.

McNamara, D. E. et al. (2009). "A Method to Establish Seismic Noise Baselines for Automated Station Assessment". In: *Seism. Res. Lett.* 80, p. 628.

Melgar, Diego et al. (2015). "Earthquake magnitude calculation without saturation from the scaling of peak ground displacement". In: *Geophysical Research Letters* 42.13. 2015GL064278, pp. 5197–5205. ISSN: 1944-8007. DOI: `10.1002/2015GL064278`. URL: `http://dx.doi.org/10.1002/2015GL064278`.

Mikolov, T. et al. (Oct. 2013a). "Distributed Representations of Words and Phrases and their Compositionality". In: *ArXiv e-prints*. arXiv: `1310.4546 [cs.CL]`. URL: `https://arxiv.org/abs/1310.4546`.

Mikolov, T. et al. (Jan. 2013b). “Efficient Estimation of Word Representations in Vector Space”. In: *ArXiv e-prints*. arXiv: `1301.3781 [cs.CL]`. URL: `https://arxiv.org/abs/1301.3781`.

Miner, Gary, John Elder IV, and Thomas Hill (2012). *Practical text mining and statistical analysis for non-structured text data applications*. Academic Press. URL: `https://www.sciencedirect.com/science/book/9780123869791`.

Montagner, Jean-Paul et al. (2016). “Prompt gravity signal induced by the 2011 Tohoku-Oki earthquake”. In: *Nature communications* 7, p. 13349.

Morin, Frederic and Yoshua Bengio (2005). “Hierarchical Probabilistic Neural Network Language Model.” In: *Aistats*. Vol. 5, pp. 246–252. URL: `http://www.iro.umontreal.ca/~lisa/pointeurs/hierarchical-nnlm-aistats05.pdf`.

Mousavi, S. Mostafa et al. (2016). “Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression”. In: *Geophysical Journal International* 207.1, pp. 29–46.

Mukund, Nikhil (2017a). *LLO alog 26850*. `https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=26850`.

— (2017b). *LLO alog 36195*. `https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=36195`.

Nielsen, F. Å. (2011). *AFINN*. Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby. URL: `http://www2.imm.dtu.dk/pubdb/p.php?6010`.

Ottaway, David J, Peter Fritschel, and Samuel J Waldman (2012a). “Impact of upconverted scattered light on advanced interferometric gravitational wave detectors”. In: *Opt. express* 20.8, p. 8329. URL: `http://www.opticsexpress.org/abstract.cfm?URI=oe-20-8-8329`.

Ottaway, David J, Peter Fritschel, and Samuel J. Waldman (Apr. 2012b). “Impact of upconverted scattered light on advanced interferometric gravitational wave detectors”. In: *Opt. Express* 20.8, pp. 8329–8336. DOI: `10.1364/OE.20.008329`. URL: `http://www.opticsexpress.org/abstract.cfm?URI=oe-20-8-8329`.

Pazzani, Michael J. and Daniel Billsus (2007). *The Adaptive Web*. Berlin, Heidelberg: Springer-Verlag. Chap. Content-based Recommendation Systems, pp. 325–341. ISBN: 978-3-540-72078-2. URL: `http://dl.acm.org/citation.cfm?id=1768197.1768209`.

Pedregosa, F. et al. (Jan. 2012). “Scikit-learn: Machine Learning in Python”. In: *ArXiv e-prints*. arXiv: `1201.0490 [cs.LG]`. URL: `https://arxiv.org/abs/1201.0490`.

Pele, Arnaud (2017). *LLO alog 32532*. `https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=32532`.

Perol, T., M. Gharbi, and M. Denolle (Feb. 2017). “Convolutional Neural Network for Earthquake Detection and Location”. In: *ArXiv e-prints*. arXiv: `1702.02073 [physics.geo-ph]`.

Peters, PC and Jon Mathews (1963). “Gravitational radiation from point masses in a Keplerian orbit”. In: *Physical Review* 131.1, p. 435.

Peterson, J. (1993). “Observation and modeling of seismic background noise”. In: *Open-file report* 93-322.

Philip, N. S. et al. (Apr. 2002). “A difference boosting neural network for automated star-galaxy classification”. In: *A&A* 385, pp. 1119–1126. DOI: `10.1051/0004-6361:20020219`. eprint: `astro-ph/0202127`.

Philip, N. S. et al. (2012). “Classification by boosting differences in input vectors: an application to datasets from astronomy”. In: *Astronomical Society of India Conference Series*. Ed. by P. Prugniel and H. P. Singh. Vol. 6. Bulletin of the Astronomical Society of India, India, p. 151. URL: `http://adsabs.harvard.edu/abs/2012ASInC...6..151P`.

Philip, Ninan Sajeeth (2010). “A learning algorithm based on primary school teaching wisdom”. In: *Paladyn* 1.3, p. 160. URL: `http://dx.doi.org/10.2478/s13230-011-0002-z`.

Philip, Ninan Sajeeth and K. Babu Joseph (Dec. 2000). “Boosting the Differences: A Fast Bayesian Classifier Neural Network”. In: *Intell. Data Anal.* 4.6, pp. 463–473. ISSN: 1088-467X. URL: `http://dl.acm.org/citation.cfm?id=1294154.1294155`.

Powell, J. et al. (Nov. 2015). “Classification methods for noise transients in advanced gravitational-wave detectors”. In: *Classical Quantum Gravity* 32.21, 215012, p. 215012. DOI: `10.1088/0264-9381/32/21/215012`.

Powers, David Martin Ward (2007). *Evaluation: from precision, recall and F-factor to ROC, informedness, markedness and correlation*.

Punturo, M et al. (2010). “The Einstein Telescope: a third-generation gravitational wave observatory”. In: *Classical and Quantum Gravity* 27.19, p. 194002.

Rasmussen, Carl Edward and Christopher KI Williams (2006). “Gaussian processes for machine learning. 2006”. In: *The MIT Press, Cambridge, MA, USA* 38, pp. 715–719.

Ricci, Francesco, Lior Rokach, and Bracha Shapira (2011). *Introduction to recommender systems handbook*. Springer. URL: `http://www.springer.com/us/book/9780387858203`.

Richardson, Leonard (2017). *Beautiful Soup*. `https://www.crummy.com/software/BeautifulSoup/`.

RM. Allen (2012). “Transforming Earthquake Detection?” In: *Science* 335, pp. 297–298.

Robinet, F. (2015). *Omicron: An Algorithm to Detect and Characterize Transient Events in Gravitational-Wave Detectors*. Tech. rep. URL: `https://tds.ego-gw.it/ql/?c=10651`.

Rollins, Jameson Graef (2016). “Distributed state machine supervision for long-baseline gravitational-wave detectors”. In: *Review of Scientific Instruments* 87.9, p. 094502.

Sathyaprakash, B. S. and S. V. Dhurandhar (1991). In: *Phys. Rev. D* 44 (12), pp. 3819–3834. DOI: `10.1103/PhysRevD.44.3819`. URL: `http://link.aps.org/doi/10.1103/PhysRevD.44.3819`.

Saulson, P. R. (1984a). “Terrestrial gravitational noise on a gravitational wave antenna”. In: *Phys. Rev. D* 30, p. 732.

Saulson, Peter R (1984b). “Terrestrial gravitational noise on a gravitational wave antenna”. In: *Physical Review D* 30.4, p. 732.

Saulson, Peter R. (1990). “Thermal noise in mechanical experiments”. In: *Phys. Rev. D* 42 (8), pp. 2437–2445. DOI: `10.1103/PhysRevD.42.2437`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.42.2437`.

Saulson, Peter R (1994). *Fundamentals of interferometric gravitational wave detectors*. World Scientific.

Savitzky, Abraham and Marcel JE Golay (1964). “Smoothing and differentiation of data by simplified least squares procedures.” In: *Analytical chemistry* 36.8, pp. 1627–1639.

Sigg, Daniel (2016). *LHO alog 30412*. `https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=30412`.

Sigurbjörnsson, Börkur and Roelof van Zwol (2008). “Flickr Tag Recommendation Based on Collective Knowledge”. In: *Proceedings of the 17th International Conference on World Wide Web*. WWW ’08. Beijing, China: ACM, pp. 327–336. ISBN: 978-1-60558-085-2. DOI: `10.1145/1367497.1367542`. URL: `http://doi.acm.org/10.1145/1367497.1367542`.

Sneath, P.H.A. and R.R. Sokal (1973). *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. A Series of Books in Biology. W. H. Freeman, San Francisco. ISBN: 9780716706977. URL: `https://books.google.co.in/books?id=iWWcQgAACAAJ`.

Specht, Donald F (1991). “A general regression neural network”. In: *IEEE transactions on neural networks* 2.6, pp. 568–576.

Stephens, Zachary D et al. (2015). “Big data: astronomical or genomical?” In: *PLoS Biol* 13.7, e1002195. URL: `https://doi.org/10.1371/journal.pbio.1002195`.

Stochino, Alberto et al. (2009). “The Seismic Attenuation System (SAS) for the Advanced LIGO gravitational wave interferometric detectors”. In: *Nuclear*

Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 598.3, pp. 737 –753. ISSN: 0168-9002.

Strauss, Jennifer A. and Richard M. Allen (2016). "Benefits and Costs of Earthquake Early Warning". In: *Seismological Research Letters* 87.3, pp. 765–772. ISSN: 0895-0695. DOI: 10.1785/0220150149. eprint: http://srl.geoscienceworld.org/content/87/3/765.full.pdf. URL: http://srl.geoscienceworld.org/content/87/3/765.

Ta, Nhi P (1994). "A wavelet packet approach to radio signal classification". In: *Time-Frequency and Time-Scale Analysis, 1994., Proceedings of the IEEE-SP International Symposium on*. IEEE, pp. 508–511.

The LIGO Scientific Collaboration (2015). "Advanced LIGO". In: *Classical and Quantum Gravity* 32.7, p. 074001. URL: http://stacks.iop.org/0264-9381/32/i=7/a=074001.

The LIGO Scientific Collaboration and Virgo Collaboration (2016). "Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914". In: *Classical and Quantum Gravity* 33.13, p. 134001. URL: http://stacks.iop.org/0264-9381/33/i=13/a=134001.

Toksöz, M. N. and R. T. Lacoss (1968). "Microseisms: Mode Structure and Sources". In: *Science* 159, p. 872.

Tromp, Jeroen, Dimitri Komatitsch, and Qinya Liu (2008). "Spectral-element and adjoint methods in seismology". In: *Communications in Computational Physics* 3.1, pp. 1–32.

US Array Website. http://www.usarray.org. Accessed: 2018-03-15.

USGS. *M 4.2 - 6km ENE of Edmond, Oklahoma*. URL: https://earthquake.usgs.gov/earthquakes/eventpage/us2000a3y4#moment-tensor.

Vajente, Gabriele (2016). *LHO alog 30473*. https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=30473.

Venkateswara, Krishna et al. (2014). "A high-precision mechanical absolute-rotation sensor". In: *Review of Scientific Instruments* 85.1, p. 015005. DOI: 10.1063/1.4862816.

Vishveshwara, CV (1970). "Scattering of gravitational radiation by a Schwarzschild black-hole". In: *Nature* 227.5261, pp. 936–938. URL: http://dx.doi.org/10.1038/227936a0.

Walczak, Beata, Bas Van Den Bogaert, and Desire Luc Massart (1996). "Application of wavelet packet transform in pattern recognition of near-IR data". In: *Analytical Chemistry* 68.10, pp. 1742–1747.

Weingarten, Matthew et al. (2015). "High-rate injection is associated with the increase in US mid-continent seismicity". In: *Science* 348.6241, pp. 1336–1340.

Wolpert, David H (1992). “Stacked generalization”. In: *Neural networks* 5.2, pp. 241–259.

Wu, Xindong et al. (2014). “Data mining with big data”. In: *ieee transactions on knowledge and data engineering* 26.1, pp. 97–107. URL: `https://doi.org/10.1109/TKDE.2013.109`.

Zevin, M et al. (2017). “Gravity Spy: integrating advanced LIGO detector characterization, machine learning, and citizen science”. In: *Classical Quantum Gravity* 34.6, p. 064003. URL: `http://stacks.iop.org/0264-9381/34/i=6/a=064003`.

Zoghbi, Susana, Ivan Vulić, and Marie-Francine Moens (2013). “Are Words Enough?: A Study on Text-based Representations and Retrieval Models for Linking Pins to Online Shops”. In: *Proceedings of the 2013 International Workshop on Mining Unstructured Big Data Using Natural Language Processing*. UnstructureNLP ’13. San Francisco, California, USA: ACM, pp. 45–52. ISBN: 978-1-4503-2415-1. DOI: `10.1145/2513549.2513557`. URL: `http://doi.acm.org/10.1145/2513549.2513557`.

Zwerger, Thomas and Ewald Mueller (1997). “Dynamics and gravitational wave signature of axisymmetric rotational core collapse.” In: *Astronomy and Astrophysics* 320, pp. 209–227. URL: `http://adsabs.harvard.edu/abs/1997A%26A...320..209Z`.