LASER INTERFEROMETER GRAVITATIONAL WAVE OBSERVATORY
- LIGO -
CALIFORNIA INSTITUTE OF TECHNOLOGY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Technical Note **LIGO-T1900526-v1-** July 2019

Extending the Reach of Gravitational-wave Detectors with Machine Learning Interim Report 2

Morgan Nanez

Distribution of this document:

LIGO Scientific Collaboration

Draft

California Institute of Technology
LIGO Project, MS 18-34
Pasadena, CA 91125
Phone (626) 395-2129
Fax (626) 304-9834
E-mail: info@ligo.caltech.edu

Massachusetts Institute of Technology
LIGO Project, Room NW22-295
Cambridge, MA 02139
Phone (617) 253-4824
Fax (617) 253-7014
E-mail: info@ligo.mit.edu

LIGO Hanford Observatory
Route 10, Mile Marker 2
Richland, WA 99352
Phone (509) 372-8106
Fax (509) 372-8137
E-mail: info@ligo.caltech.edu

LIGO Livingston Observatory
19100 LIGO Lane
Livingston, LA 70754
Phone (225) 686-3100
Fax (225) 686-7189
E-mail: info@ligo.caltech.edu

http://www.ligo.caltech.edu/

Abstract

This report presents the idea of using current machine learning techniques and algorithms to reduce the overall noise floor of the LIGO detectors. There will be a hard emphasis on techniques that analyze time series data, such as utilizing long short-term memory and nonlinear regression algorithms. While other sources of noises in the detectors are outlined in the proposal, there will be a focus on using machine learning algorithms to hone in on noise sources coming from physical attributes of the instrument itself. The goal is to increase the sensitivity of the detectors by subtracting linear and non-linear noise coupling mechanisms using machine learning.

1 Introduction

In recent years LIGO has made strides in the discovery of gravitational waves from stellar mass black holes and neutron star mergers. However, there are still many more waves viable for detection below the current surface of noise. With the application of machine learning algorithms to the gravitational-wave detector data and auxiliary channels on-site, there is a possibility to subtract the noise in the time-series due to instrumental artifacts. By reducing the current noise floor there will be greater sensitivity in the instrument, leading to a greater rate of detection. add about the physical aspects of ligo (how it works etc

2 Noise Sources

For this project, we want to focus on the what happens within the interferometer and why it causes noise. Within the instrument itself, there are two main types of noise: displacement and sensing noise. [8]. Displacement noise causes real motion of the test masses, four suspended mirrors that form the Fabre-Perot arm cavities, or their surfaces. Sensing noises limit the ability of the instrument to measure test mass motion.

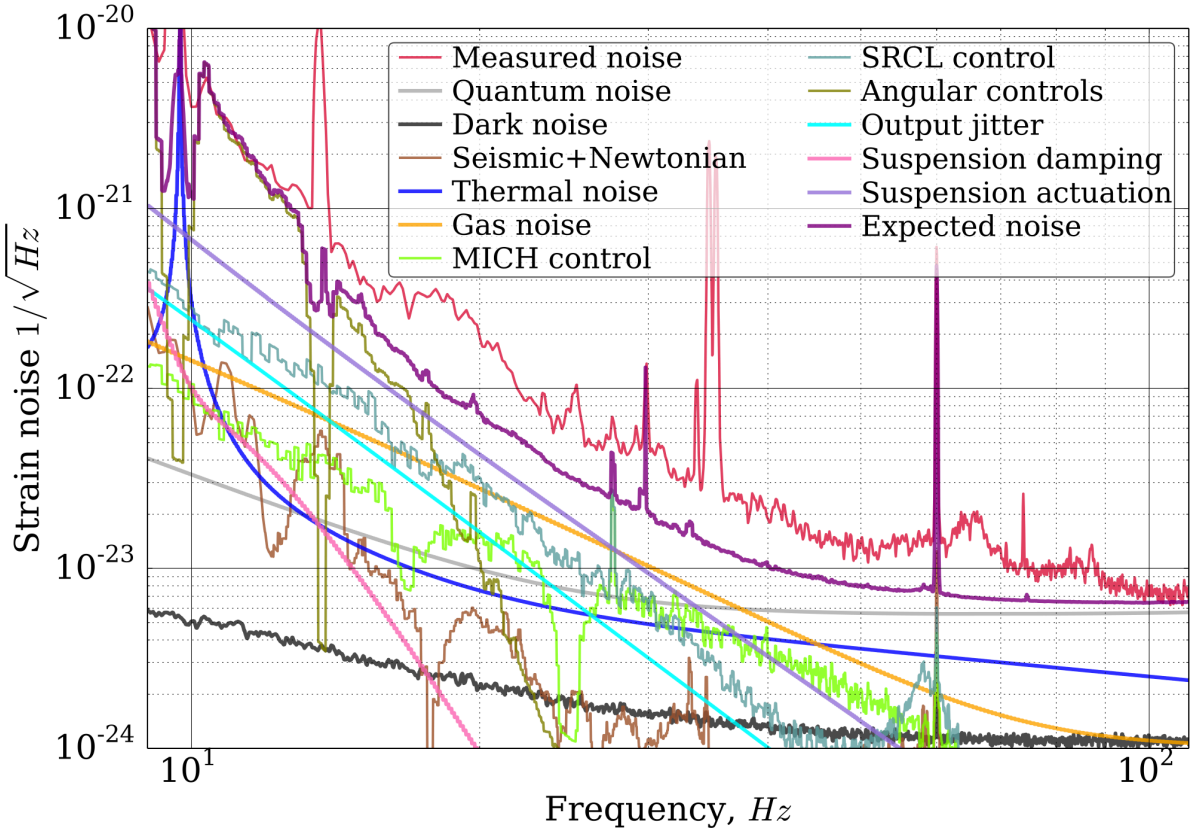

Figure 1: An example of source noises that combine to form the sensitivity of the detectors. This noise budget is for the Livingston detector at low frequencies. Source: [8]

The coupling of each noise source to the gravitational wave channel at a certain frequency f is estimated using the following equation: $L(f) = L_0 h(f) = T(f) \times N(f)$. Where $N(f)$ is the noise spectrum measured by an auxiliary (witness) sensor and $T(f)$ is the transfer function from the sensor to the gravitational wave channel.

Ultimately we want to be able to subtract these sources. When picking which witness channels to subtraction from the data in order to get rid of certain noise sources, they must be determined incapable of subtracting potential gravitational-wave signals. This can be tested by inserting simulated signals into a test data set and then processing them accordingly. While I will discuss known noise sources, there are still sources contributing heavily to the noise floor that have not been identified. It is important to explore and focus on the broadband noise as that limits the instrument sensitivity greatly and is harder to remove.

2.1 Fundamental Noise

Fundamental noises determine the ultimate design sensitivity of the interferometer. This noise cannot be reduced without upgrades to the instrument itself, such as installing a new laser. Multiple sources of thermal noise are present in the interferometer. Thermal vibrations of the suspension

fibers, in the form of suspensions thermal noise, causes motion of the test masses. Thermal fluctuations of the optical coatings cause coating Brownian noise. This was reduced by optimizing the thickness of the coatings. Thermal noise also arises in the substrates of the test masses [2, 11] Another fundamental noise is the quantum noise, which is propelled by the fluctuations of the optical vacuum field entering the interferometer through the antisymmetric port [4, 5] . By disturbing the optical field resonating in the arm cavities, the vacuum fluctuations creates a displacement noise by exerting a pressure force that causes physical motion of the test masses. The noise seen in the differential arm channel is given by:

$$L(f) = \frac{2}{cM\pi^2 f^2}(hvGP_{\text{arm}})^{1/2}K(f) \tag{1}$$

$$L(f) = \frac{1.38\times10^{-17}}{f^2}(\frac{P_{\text{arm}}}{100\text{kW}})^{1/2}K(f)\frac{\text{m}}{\sqrt{\text{Hz}}} \tag{2}$$

where h is Planck constant and P_{arm} is the power circulating in the arm cavities. This creates a fundamental limit to the detector sensitivity below 40Hz.[7]

The vacuum fluctuations entering through the port also introduce shot noise in the gravitational wave channel. [6] The total shot noise can be written as:

$$L(f) = \frac{\lambda}{4\pi G_{\text{arm}}}(\frac{2hvG_{source}}{G_{\text{prc}}P_{\text{in}}\nu})^{1/2}\frac{1}{K(f)} \tag{3}$$

$$L(f) = 2\times10^{-20}(\frac{100\text{kW}}{P_{\text{arm}}\nu})^{1/2}\frac{1}{K(f)}\frac{m}{\sqrt{\text{Hz}}} \tag{4}$$

where ν is the power that is transmitted to the photodiodes, 0.75. Shot noise level is independent of the differential arm offset for small offsets ΔL ¡ 100pm. Shot noise limits the design sensitivity above 40 Hz, and current sensitivity. above 100 Hz.

2.2 Technical Noise

Technical noise stems from intricacies in the instrument and can be reduced once the cause is carefully studied. A source of technical noise is the charging noise on the test masses. Ideally, the only charge on the test masses would be the one accumulated due the electrostatic actuation however there can be residual charge left from cleaning and protection of the optics. The surface of the test masses also loose charge due to the UV photons generated by the nearby ion pump used in the vacuum system. Two coupling mechanisms arise between charging noise and the gravitational wave channel. The first one comes from time and the second one comes from voltage fluctuations of the various pieces of grounded metal in the vicinity of the test masses. Voltage noise creates fluctuations in the electric field E and a applies a force F_{ch} on the test mass, following the equation: $F_{\text{ch}} = \int E\sigma dS$ where the integral is computed over both the front and back surfaces of the test mass. The total noise coupling above 10 Hz can be estimated using the equation

$$L(f) = \frac{F_{\rm ch}}{M(2\pi f)^{1/2}} \approx \frac{10^{-16}}{f^2} \frac{\sigma}{10^{-11}{\rm C/cm^2}} \frac{\rm m}{\sqrt{\rm Hz}} \tag{5}$$

By discharging the test masses, the coupling of voltage fluctuations on the ground plane to the gravitational wave signal was reduced by a factor of 10 to 100. Charge from the front surface can be removed with the use of ion gun [3]

Due to the nature of the interferometer, laser amplitude and frequency noise arises as well. Intensity and frequency stabilization servos actively suppress the laser noise along with a passive filter applied to the laser beam as it enters the main interferonmeter. While a largely portion of the laser frequency noise is cancelled at the antisymmetric port, there is still residual frequency noise couples into the gravitational wave channel.

To stabilize the interferonmeter optical response, active control of the mirror angular degrees of freedom is important. However, noise in the associated auxiliary degrees of freedom will couple to the gravitational wave channel at some level. Any residual fluctuation of the Michelson length $N_{\rm mich}$ couples to the transmitted power of the output mode cleaner, where the gravitational wave channel is transduced. The coupling mechanism can be described as:

$$L(f) = \frac{1}{G_{\rm arm}} N_{\rm mich}(f) \tag{6}$$

where $G_{\rm arm}$ is arm cavity build up. The coupling coefficient weakly depends on the differential arm offset and alignment. Residual fluctuations of the signal recycling cavity length also couple to gravitational wave channel, due to ΔL, the different arm offset, through radiational pressure force extereted on the test masses by. the resonating optical fields. Between the frequencies 10 to 70 Hz, you can model the differential arm noise $L(f)$ due to signal recycling cavity longitudinal noise $N_{\rm srcl}$ using:

$$L(f) = \frac{0.16}{f^2} \frac{\Delta L}{10{\rm pm}} N_{\rm srcl}(f) \tag{7}$$

A nonlinear coupling appears due to low-frequenecy modulation of the ΔL, arising form unsuppressed this word is right angular motion of the interferometer mirrors. At higher frequencies, above 70 Hz, the coupling signal depends on the mode matching between the signal recycling cavity and the arm cavity. Any residual angular motion of the test masses $N_{\rm ang}$ couples to the gravitational wave channel geometrically due to beam mis-centering d on the mirror, given by the equation $L(f) = d \times N_{\rm ang}(f)$. To mitigate the linear coupling of auxiliary degrees of freedom to the gravitational wave channel a realtime feed-forward cancellation technique is utilized. Witness signals are reshaped accordingly using time-domain filters and the cancellation signals are applued directly to the test masses. This reduces the amount of noise coming from the auxiliary degrees of freedom in the range 10-150Hz.

[8]. An instrumental cause of small fluctuations in the data is the jittering of the pre-stabilized laser. [9] The jittering was introduced when upgrades were added to the subsystem. A high powered oscillator was added to the system in an effort to increase laser power. However, the high powered oscillator required continuous heat dissipation via water cooling. Vibrations originated from the

water flow, introducing jitter into the beam angle and size, resulting in noise. Other efforts, such as adding sensors that measured radial beam distortions and thermal compensation systems were used to mitigate the noise from the jittering. However, the jitter remained causing noise throughout the second LIGO observation.

2.3 Environmental Noise

Environmental noise is the noise that comes from the environment, such as seismic motion, acoustic and magnetic noises. The vacuum chambers and the arm tubes are not isolated from the ground seismic or the ambient acoustic noises [8]. This motion can couple to the gravitational wave channel through scattered light. When the laser light hits the optical components, some light scatters out of the main beam and then scattered back into the main beam from the walls and other components. This is called backscattering. The backscattered light modulates the main beam in both phase and amplitude, introducing noise into the gravitational wave channel. For example, in the vacuum chambers, the coupling of scattered light noise is modulated by low frequenecy motion of the scattering surfaces, and therefore note linear. We can use the following equation to make a projection of scattered light noise to the gravitational wave channel.

$$L(f) = N_{\text{amb}}(f)\frac{L_{\text{ecx}}(f)}{N_{\text{ecx}}(f)} \tag{8}$$

where L_{ecx} and N_{ecx} are the spectra of the gravitational wave channel and of the back scattering element motion.

One source of environmental noise is seismic noise. Below 10Hz, residual seismic motion causes significant displacement noise. On average, at 10Hz the ground moves $10^{-9}\text{m}/\sqrt{\text{Hz}}$, which is 10 orders of magnitude larger than LIGO's target sensitivity. To address this, seismic noise is filtered using a combination of passive and active stages. The test masses are suspended from quadruple pendulums, providing isolation as $1/f^8$ in the detection bandwidth. The pendulums are then mounted on multistage active platforms [1]. By using very-low inertial sensors, the required isolation is achieved in the detection band and at lower frequencies.

The fluctuations of the local gravity fields around the test masses also couple to the gravitational wave channel as force noise, also known as gravity gradient noise.[10] The coupling to the differential arm length displacement is given by the following equations:

$$L(f) = 2\frac{N_{\text{grav}}(f)}{(2\pi f)^2} \tag{9}$$

$$N_{\text{grav}}(f) = \beta G_\rho N_{\text{sei}}(f) \tag{10}$$

where N_{grav} is the fluctuation of the local gravity field projected on the arm cavity axis, the factor of 2 accounts for the incoherent sum of noises from the four test masses, G is the gravitational constant,ρ 1800 kg m3is the ground density near the mirror, β 10 is a geometric factor, and N_{sei} is the seismic motion near the test mass. This is a limiting noise source typically in the frequencey range 10-20Hz

3 Previous Regression Attempts

Previously, linear noise was regressed by directly accepting the data from the physical environmental channels as the witness channels, while bilinear noise is handled by constructing witness channels from two or more PEM and or simulated channels. One of the more promising past regression methods is based on the construction of the Wiener-Kolmogorov (WK) filters, which has been preformed on seismic noise subtraction and extended to multi-channel analysis [12]. On a similar note, the Wiener method has been used to subtract noise as well. This method determines how to manipulate the data from a witness channel so that when it is subtracted from the gravitational wave channel the mean-square-error of the gravitational wave channel is minimized [13]. This method was used for low frequency seismic noise.

4 Mock Data

Our neural network must have some data to train, test and validate on. Because we are developing the neural network, we do not start with real data from the interferometer, as that is very noisy and hard to pinpoint which features break the network. As a result, we develop mock data to train the network on. Currently, we have five models that each cover different characteristic of the previously mentioned noise. Generally mock data is generated by inserted random gaussian noise

4.1 Current Models

The resonance model uses the previous inputs as well as a resonant frequency and a quality factor to produce the subtraction target and witness signals. The bilinear model attempts to represent data similar to the beam-spot motion and mirror angularly control signal. The beam-spot motion is caused by microseism, or very small earthquakes. Lastly, the scatter model attempts to mimic one witness moving slowly and one acoustically active witness coupling together. While all these model have their own strengths, they all fail to accurately mimic the real data being outputted by the interferometer.

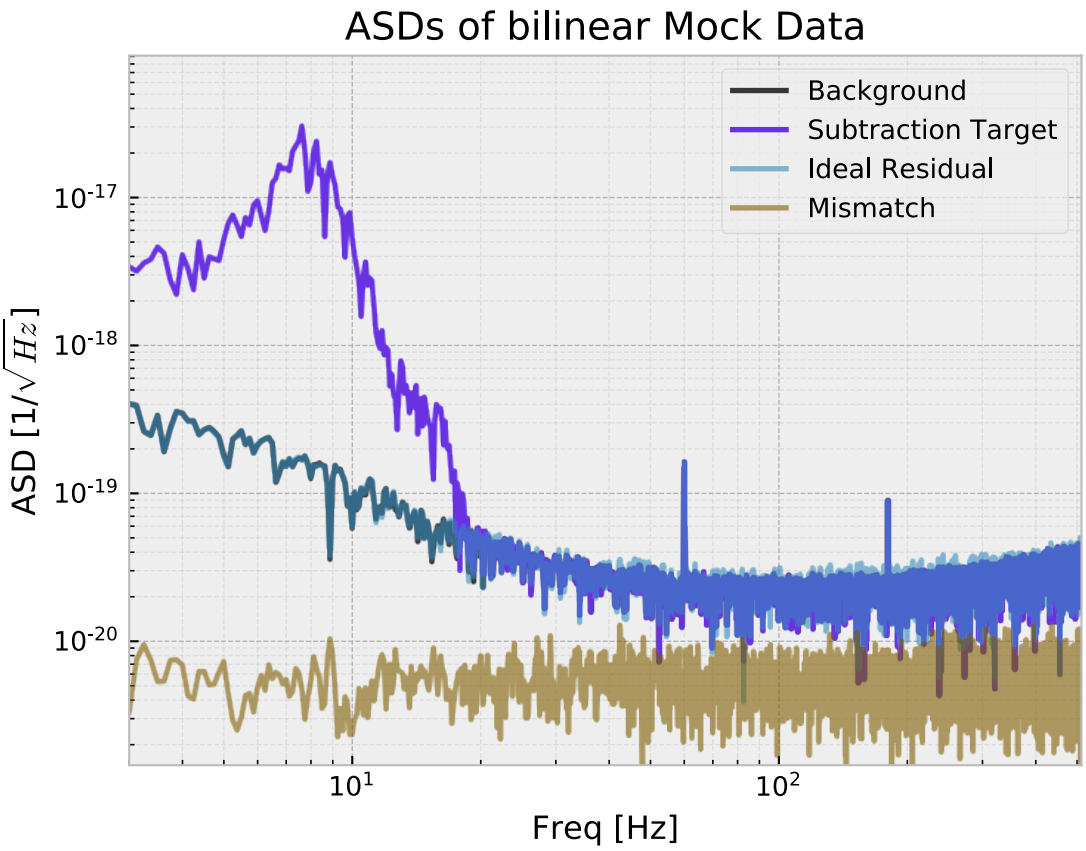

Figure 2: Spectral density of the data being generated by the bilinear filter

While we have not mimicked the real data very accurately, one of the harder features to remove from the data are the irregular spikes and cusps that are present. To mimic these features we are using a filter called ellip.. It is quite simple at the moment but it is just a preliminary filter to test and develop the neural network on. Right now, it is a simple bandstop filter but also contains the types of peaks and troughs we are aiming at removing. The idea is that if we test the neural network on this model, it will be easier to build up the model and adjust the network accordingly. This is simply a first step in the right direction and the ellip model does not contain all the factors needed to make an accurate data set. The ellip model is appealing because it time dependent. In order to create a good neural network capable of mimicking this filter, it should in theory need to look at previous inputs in order to correctly predict the output of the current input.

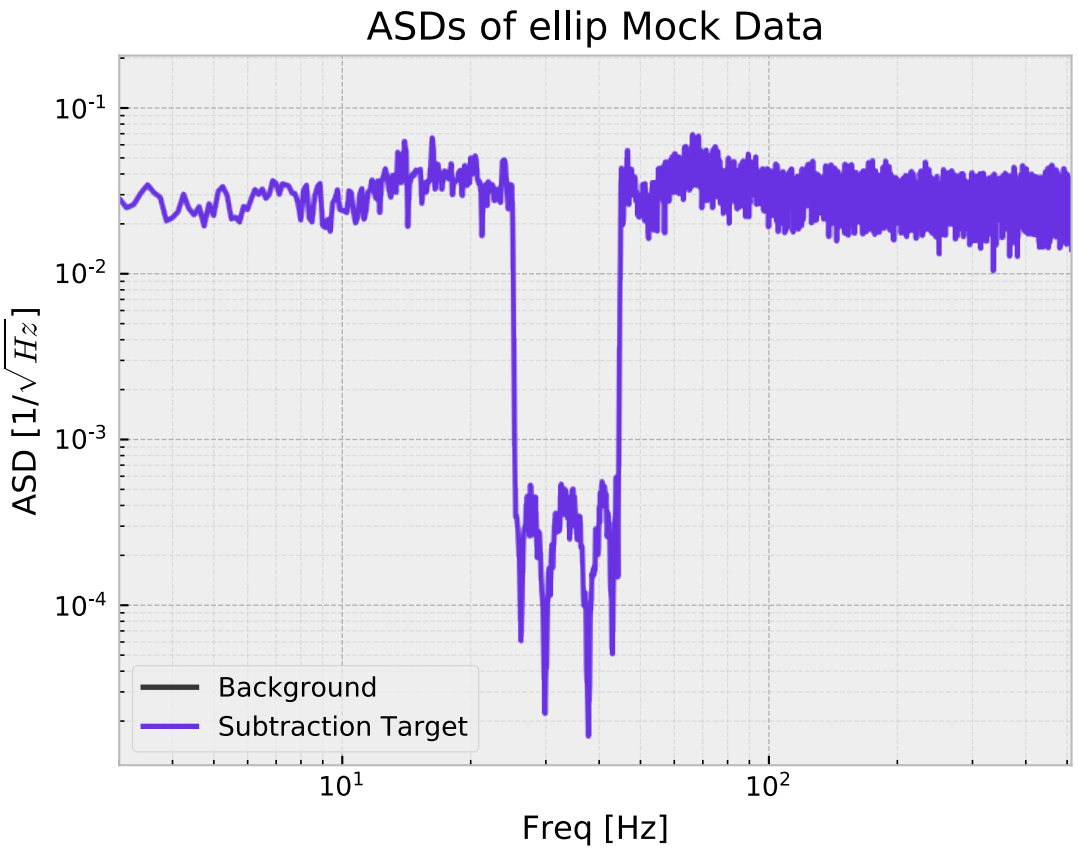

Figure 3: Spectral density of the data being generated by the ellip filter

Throughout the rest of the summer we hope to improve it and combine the previously mentioned models into one model that will accurately represent the real data from the instrument.

5 Machine Learning

There are two types of noise which contribute to the overall sensitivity of the detector. Non-removable noise sources, such as thermal noise, determine the underlying sensitivity of the detector. These noise sources can only be removed by improving the design of the detector itself. Other noise sources, known as removable noises, such as seismic noise, can be removed by monitoring the witness channels. Witness channels are the channels that monitor the physical environment around the detector. Machine learning algorithms coupled with data from the witness channels can be used on removable noise sources while keeping the signals we want to detect intact. By running the algorithms on past data, noise regression algorithms will be able to predict future noise sources and subtract them from the incoming data. It is also important to define neural networks as a set of algorithms that are designed to recognize patterns. These networks are made up of multiple layers that do different computations on the passing inputs, reducing the outputted data. These layers are identified as the input layer, the hidden layer where computations are preformed, and the output

layer. Because the data we are looking at is sequential in nature, we consider the outputted data to be time series data.

Time series data is defined to be taken sequentially. Time series forecasting is the process of predicting future events based on previous data. The most important components of time series are the trend, cycle, seasonality and error. This type of forecasting will be used on the time series data outputted by the detector to analyze the witness channels and the environmental factors contributing to removable noise sources.

We also want to be able to see how the data evolves and reacts to the data around it. For this reason, we want to deal with recurrent neural networks or convolutional neural networks.

Because the data we are dealing with it highly complicated, I plan on trying nonlinear regression algorithms. Nonlinear regression is used to find nonlinear relations between sets of data. It is ideal for the task at hand because it can estimate models with arbitrary relationships between independent and dependent variables.

6 Noise Couplings

While the fundamentals sources of noise determine the baseline LIGO sensitivity, many other sources and disturbances are coupled into the detector sub-systems, degrading the overall detector sensitivity. LIGO has thousands of physical environment monitors collecting data that can help characterize the coupling between the gravitational wave channel and the environment. Most of the couplings are linear

7 Neural Networks

Neural networks, broadly, are a set of algorithms that are designed to recognize patterns. The term deep learning is used to describe neural networks that are stacked, in other words, networks that have several layers. Each layer is made of nodes. A node is where the computation happens, combining inputs from the data with a set of weights, that either amplify or dampen that input, assigning significance to the input in regards to the algorithm the neural network is trying to learn. The output from the all input-weight calculations are summed up and then passed through the node's activation function. The activation function determines to what extent that signal should be passed through the network. If the signal passes through, the node is considered to be activated.

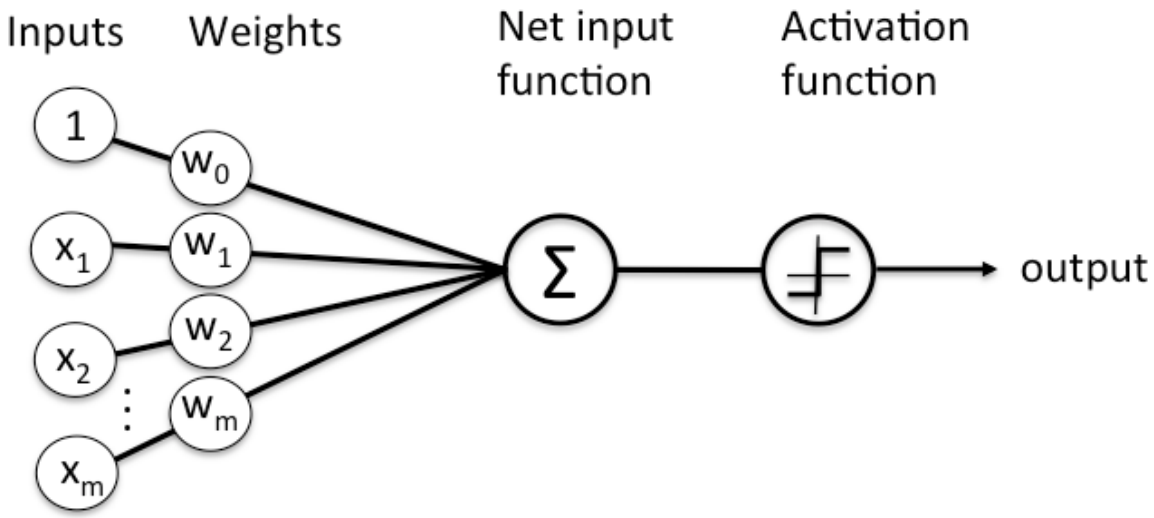

Figure 4: An example of what one node could look like. Source:https://skymind.com/wiki/neural-network

The goal of using a neural network is to get to a point of minimal error as fast as possible. We want to develop neural networks that are able to mimic the filtering of both linear and non-linear couplings. In doing so, we utilize a feedforward neural network. The input enters the naive. network and then a guess of the output is produced. The difference between the estimated output and the actual output is our error. Adjustments are then made in order to minimize the error and repeats the process again.

Our goal is to be able to apply different networks to the different auxiliary channels. As discussed above, different noise sources have different couplings and require different types of transformers to subtract the noise. I will discuss the three modern architectures we explored.

7.1 Regression

Linear regression is a machine learning algorithm based on supervised learning. Linear regression preforms the task to predict a dependent variable value y based on a given independent variable x.

$$Y_1 = bx_1 + a \tag{11}$$

Where Y is the estimated output, X is the input, the coefficient b is the scale factor to input X and a is the bias coefficient. This model can made more complicated by simply adding more inputs and scaling factor.

The process in which we optimize the values of the coefficients by iteratively minimizing the error of the model on the training data is called gradient descent. Essentially, you start with random values for each coefficient then after each iteration the sum of the squared errors are calculated for each input-output pair values. The coefficients are updated in the direction toward minimizing the error. This is repeated until a minimum sum squared error is achieved.

7.2 Multilayer Perceptrons

This is the classical type of neural network. These are simple computational units that have weighted input signals and produce and output signal using an activation function. An activation function is a simple mapping of summed weighted input to the output of the neuron, essentially governing the threshold at which the neuron is activated and the strength of the output signal. This has proved to work very well on linear couplings.

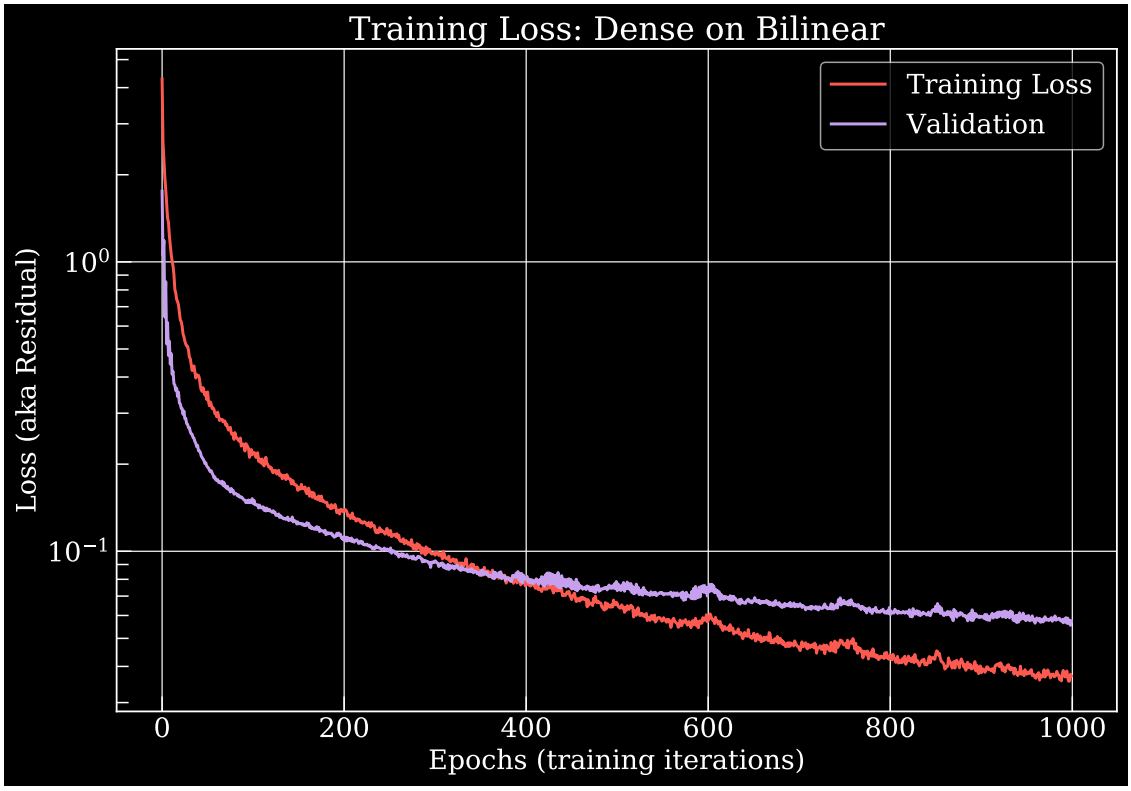

Figure 5: Loss using Bilinear mock data and a simple dense network

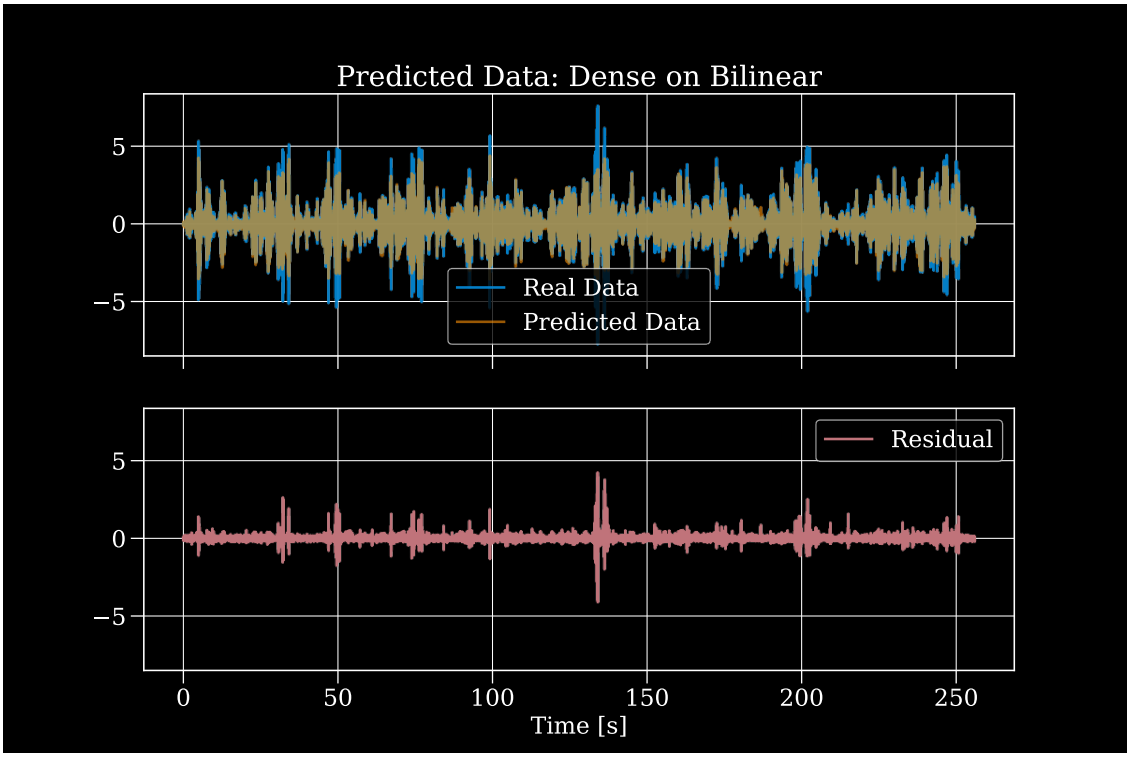

Figure 6: The output predicted data from the same neural network

7.3 Convolutional Neural Networks

These networks are unique in that that preserve the spatial structure of the problem by learning internal feature representations using small chunks of input data. This preservation allows the network to be used on data that has a spatial relationship. The convolutional layers are comprised o filters and feature maps. Filters are the equivalent of the classical neuron, also having weights and an output value. The layers also make use of a nonlinear transfer function as part of activation. This allows us to apply this type of network to the nonlinear noise channels.

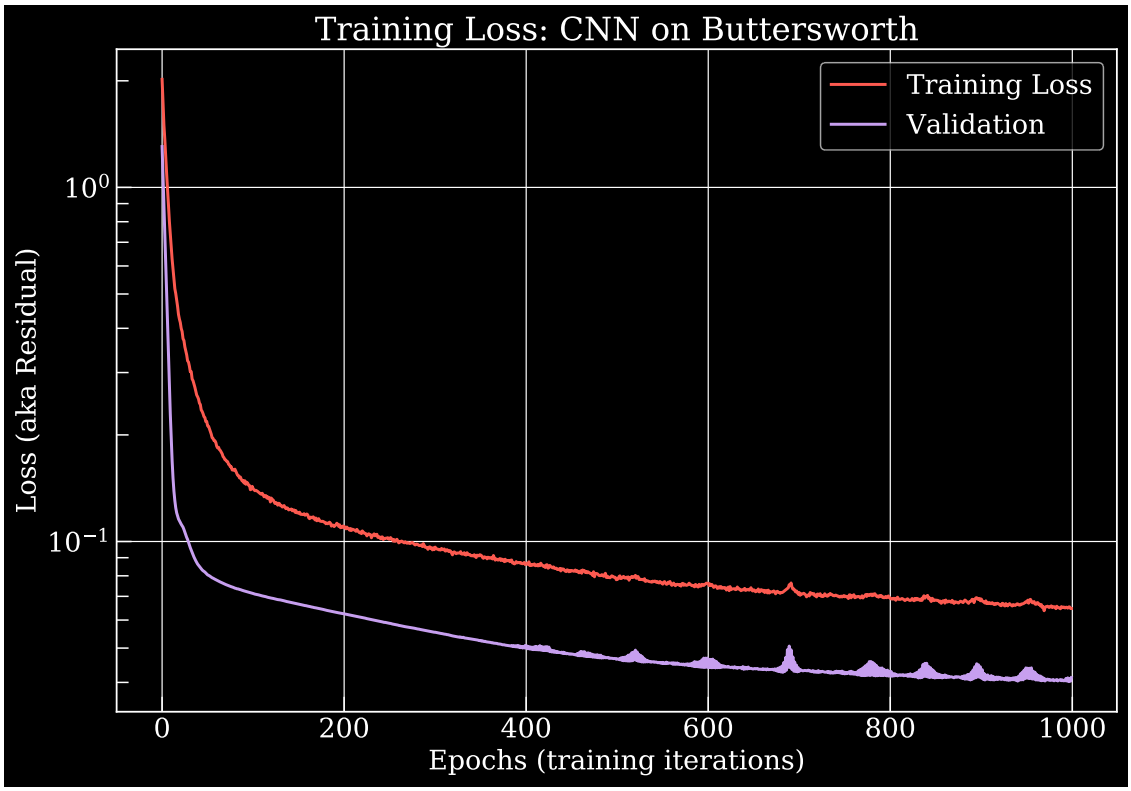

Figure 7: Loss using buttersworth mock data and a simple convolutional neural network

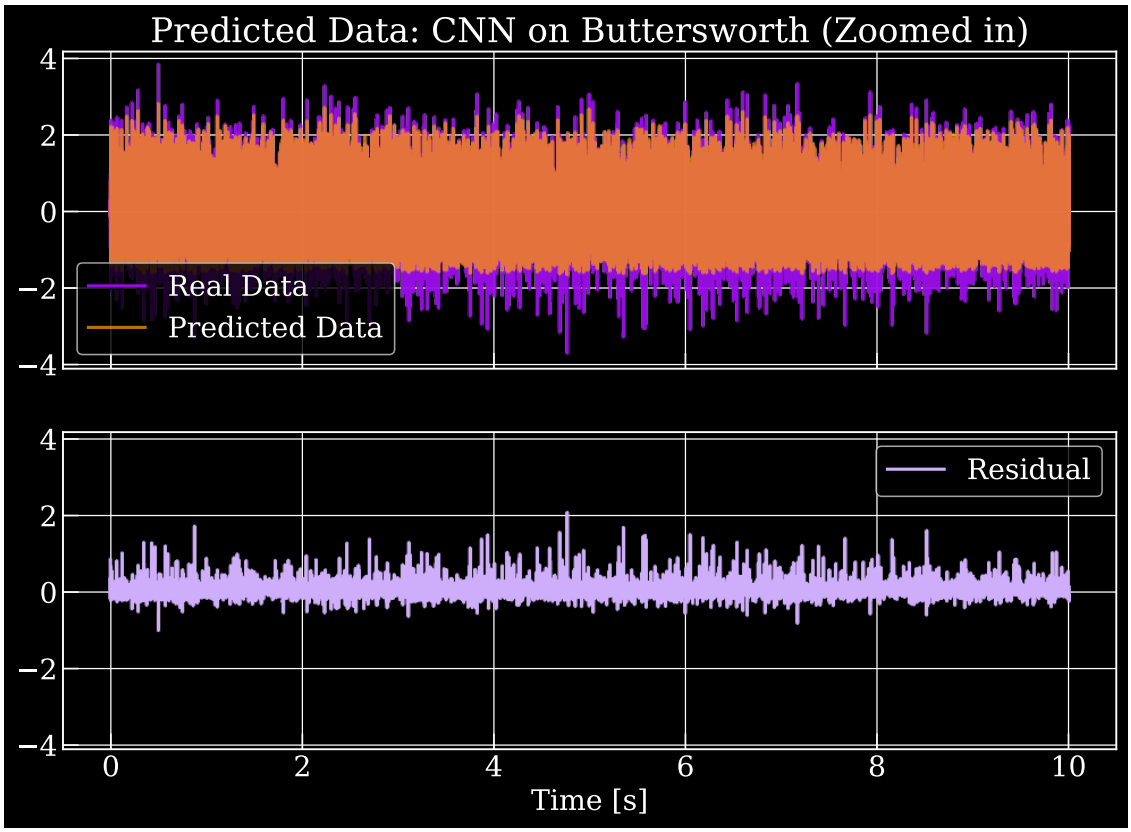

Figure 8: The output predicted data from the convolutional neural network

7.4 Recurrent Neural Networks

Recurrent neural networks are similar to multilayer perceptron networks, but have the addition of loops. The output of the network may feedback as an input to the network. These connections add memory to the network and allow it to learn. broader abstractions from the input sequences. There are many types of recurrent neural networks, but we have been focusing on using the long short-term memory model. It allows errors to flow backwards which leads to the network learning patterns that require memory of events that happened at earlier time steps. These are a type of recurrent neural network that process sequential inputs, taking into consideration past inputs to analyze the current input. This process is appealing because it can capture long-term dependencies in the data. Some of the removable noise sources, like seismic waves, will take several seconds to get to the witness channels, so having the ability to take in longer inputs is needed.

References

[1] S M Aston, M A Barton, A S Bell, N Beveridge, B Bland, A J Brummitt, G Cagnoli, C A Cantley, L Carbone, A V Cumming, L Cunningham, R M Cutler, R J S Greenhalgh, G D Hammond, K Haughian, T M Hayler, A Heptonstall, J Heefner, D Hoyland, J Hough, R Jones, J S Kissel, R Kumar, N A Lockerbie, D Lodhia, I W Martin, P G Murray, J O'Dell, M V Plissi, S Reid, J Romie, N A Robertson, S Rowan, B Shapiro, C C Speake, K A Strain,

K V Tokmakov, C Torrie, A A van Veggel, A Vecchio, and I Wilmut. Update on quadruple suspension design for advanced LIGO. *Classical and Quantum Gravity*, 29(23):235004, oct 2012.

[2] V.B. Braginsky, M.L. Gorodetsky, and S.P. Vyatchanin. Thermodynamical fluctuations and photo-thermal shot noise in gravitational wave antennae. *Physics Letters A*, 264(1):1 – 10, 1999.

[3] P Campsie, L Cunningham, M Hendry, J Hough, S Reid, S Rowan, and G D Hammond. Charge mitigation techniques using glow and corona discharges for advanced gravitational wave detectors. *Classical and Quantum Gravity*, 28(21):215016, sep 2011.

[4] Carlton M. Caves. Quantum-mechanical radiation-pressure fluctuations in an interferometer. *Phys. Rev. Lett.*, 45:75–79, Jul 1980.

[5] Carlton M. Caves. Quantum-mechanical noise in an interferometer. *Phys. Rev. D*, 23:1693–1708, Apr 1981.

[6] LIGO Scientific Collaboration. Enhanced sensitivity of the ligo gravitational wave detector by using squeezed states of light. *Nature Photonics*, 7:613 EP –, 07 2013.

[7] LIGO Scientific Collaboration. Advanced LIGO. *Classical and Quantum Gravity*, 32(7):074001, mar 2015.

[8] LIGO Scientific Collaboration. Sensitivity of the advanced ligo detectors at the beginning of gravitational wave astronomy. *Phys. Rev. D*, 93:112004, Jun 2016.

[9] D. Davis, T. Massinger, A. Lundgren, J. C. Driggers, A. L. Urban, and L. Nuttall. Improving the sensitivity of Advanced LIGO using noise subtraction. *Classical and Quantum Gravity*, 36(5):055011, March 2019.

[10] Jennifer C. Driggers, Jan Harms, and Rana X. Adhikari. Subtraction of newtonian noise using optimized sensor arrays. *Phys. Rev. D*, 86:102001, Nov 2012.

[11] Yuk Tung Liu and Kip S. Thorne. Thermoelastic noise and homogeneous thermal noise in finite sized gravitational-wave test masses. *Phys. Rev. D*, 62:122002, Nov 2000.

[12] V Tiwari, M Drago, V Frolov, S Klimenko, G Mitselmakher, V Necula, G Prodi, V Re, F Salemi, G Vedovato, and I Yakushin. Regression of environmental noise in LIGO data. *Classical and Quantum Gravity*, 32(16):165014, jul 2015.

[13] Norbert Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications*. The MIT Press, 1949.