Hello everyone!

My name is Nadia Dimitrova and I am a senior in Aerospace Engineering at MIT. My project is, however, focused on machine learning and how it can be used to improve LIGO’s performance.

As many of you mentioned, LIGO is an interferometer. The instrument splits a light beam into two parts and recombines them after travelling different optical paths between mirrors – called test masses [CLICK]. Each separated beam enters a so-called Fabry Perot cavity [CLICK] and bounces between its two additional mirrors about 300 times before being merged with the beam from the other arm.

Passing gravitational waves emitted by massive stellar objects, including colliding neutron stars or black holes, cause ripples in space-time. And at the intersection between the perpendicular arms, a photo detector observes fluctuations in the induced differential arm length, or DARM signal. These fluctuations are due to the strain (or bending) of space-time that creates optical paths of different length. The signals over time can be analyzed to infer the nature of physical phenomena causing gravitational waves. [CLICK] Power and signal amplification is created through recycling optical systems, while ground motion is reduced by suspending all instruments.

In order to measure physical events below the limit of the detector (like higher mass binary black holes, early warnings of binary neutron star mergers, and cosmology with high red-shift sources), we focus on improving LIGO’s sensitivity. Any disturbance that limits the sensitivity is defined as noise. [CLICK] And noise is everywhere! Noise due to fundamental infrastructure limits, such as thermal and quantum fluctuations, is unsubtractable. The subtractable noise is caused by environmental perturbations and technical noise.

[CLICK]

If the laser beams are off-center on the suspended mirrors, the Fabry-Perot cavities create resonance effects as the beams bounce in between the two mirrors. Despite the elaborate seismic isolations through mechanical pendulums and springs, the residual motion of the end mass mirrors is large and a control system, shown here, is implemented to manage this resonance.

Starting from the left, a ground motion disturbance couples with the motion of a mirror. The motion is observed by sensors that transmit the information to a controller, which in turn commands actuators. The goal is to reduce undesired noises by compensating for their effects on the residual signal.

[CLICK]

The current regression method employed by aLIGO minimizes the squared error between the DARM readout channel and the predicted noises from the physical environmental monitor channels – all the sensors. Nonetheless, this Wiener filtering performs regression analysis that does not capture non-linearly coupled noises well. As seen in the figure, the advanced LIGO design exhibits a curbed performance in the [CLICK] sub-60 Hz frequency range due to many noise sources. Most of you talked about minimizing those today.

The problem is that feedback loops can be tricked to account for noise introduced by the sensors or actuators [CLICK] into the DARM main readout signal. Often this contamination is non-linear and hard to be removed by the implemented linear techniques.

[CLICK]

This project focused on one particular optical phenomenon caused by subtractable noise sources. Illustrated here, the angular rotation of the mirror can couple with the beam spot motion to create a cavity length signal that mimics a gravitational wave. For example, As seen in the figure, if the mirror is rotated by $\Delta \theta$ and the beam spot is displaced by $\Delta y$ from the rotational pivot, the fluctuation in length induced is $\Delta l$. This length fluctuation can be mistaken for a real physical phenomena, but is not filtered by the aLIGO Wiener filter that focuses on linearly coupled noises.

[CLICK]

This bilinear coupling means that for a beam that is no longer centered in the mirror's rotational pivot, beam spot offset and mirror rotation shown on the left and right respectively, affect the optical path linearly given that the other is held constant.

[CLICK]

This project focuses on modeling the spot motion from sensors, or witness channels.

[CLICK]

The analytical beam spot position on each mirror can be derived analytically as a function of the residual angular motions of the four test masses ETMX, ITMX, ETMY and ITMY, given the cavity’s geometry. Since these relations are purely geometric, the coupling should be instantaneous. Thus, we can expect that the spot motion can be modeled only by data witnessed at the current time step.

[CLICK]

Though there are many sensors in aLIGO's digital control and supervision architecture, even amounting to thousands of witness channels, we can expect the spot motion to be explained by only a few of those. [CLICK] From the physical properties of the system, there are a few witness channels which directly correlate to the spot motion.

The first three sets of channels are a part of the alignment sensing and control system (ASC). This is important in order to account for physical effects such as radiation pressure. Note that all channels have different DC offsets.

[CLICK] Also, the beam motion is affected by ground disturbances. The OPLEV channels are a part of the suspension system (SUS). [CLICK] To capture more data, the Internal Seismic Isolation platforms (ISI) measure the longitudinal motion at the top of the suspension chain.

[CLICK]

Before creating a model, it is helpful to understand what these channels actually look like in the time and in the frequency domain. Here, you can see a witness channel from a few of the groups mentioned above as it varies its output over time. [CLICK] Compared to the spot motion, or target data, these channels have a much higher fluctuation over time. The spot motion is directly measured via the ADS, and then the signal is coupled into the DARM main readout. We can also reconstruct the ADS channel from the main readout by inversing the signal processing.

[CLICK]

Why was it important to see these graphs? Well, as you can see, the measured spot motion varies on the scale of 5-20 seconds. This means that the frequency we will be interested in modelling will be in the 0.05-0.2 Hz band. To confirm this, I also plotted the linear coherency of the witnesses with the target spot motion and in that band, the coherency is up to .9, which is a strong linear relationship (think of the R squared values in excel).

[CLICK]

Naturally, the analysis continued by plotting the power spectra of the data. The peak around 0.12 Hz represents the motion over time-periods similar to that of the beam spot motion. We will focus on creating a model to mimic this signal.

[CLICK]

So how do we model the signal if the analytical coupling transfer function is not known? Well, machine learning neural networks perform that exact mapping of input to output data, without having to know much about the coupling in between.

[CLICK]

In our model, the input data undergoes pre-processing in order to improve numerical training. This includes a scaling method that reduces the variance of the input dataset to unity. The mean is preserved, since it represents a DC offset, which we mentioned might be important for some of the witness channels, and is thus a valuable information to the network.

[CLICK]

The magic in the middle of the network is simply a regression model. The simplest model, a dense model, takes the input data as separate blocks and connects their values, the circles, via layers with different weighs (or edge opacity). The network aims to minimize the difference between its output and some target output.

In our case, this implies that we can only combine data from the same time-step from different witnesses. Nonetheless, this should be able to capture the physical equations theorized.

[CLICK]

A convolutional layer can also be added, since it allows us to filter the inputs. This filtering can happen over time and thus prior inputs can help in the prediction of future signals.

[CLICK]

In order to test if the theory matches expectations, we can use not the ADS signal, but rather a reconstructed ADS signal. Starting from the main readout, we reverse the signal processing to obtain what the ADS signal should measure. If the same network architecture performs well on the reconstructed data but not in reality, there might be an effect that is overlooked in the system. Understanding this might pave the way toward improving the sensitivity of LIGO even further.

[CLICK]

Reconstructed spot position dense – can see the bump; the low frequency is captured better; loss plateaus around 1 (so the model is still not very good)

[CLICK]

Measured spot position dense – blue and orange don’t match so the network doesn’t capture the information relevant. The time-series are dominated by high-frequency components; loss plateaus around 1

[CLICK]

Reconstructed spot position conv – not much better than dense; as expected by theory, the dense linear model is sufficient to model the data; but probably needs better hyperparameter tuning and more training epochs and data; no need to use convolution yet

[CLICK]

Measured spot position conv – Similar to the previous real data model, this one does not perform well. This is likely due to the large range of features in the signals.

[CLICK]

A machine learning approach via nonlinear regression can be very useful when accounting for non-linear coupling effects since it does not care about the system that caused them, but simply by how the input affects the output.

Networks developed so far demonstrate an ability to model bilinear coupling given perfect witnesses of beam spot motion and angular motion. This report focused on modeling the beam spot motion in order to help with the reduction of the bilinearly coupled noise it induces.

[CLICK]

Imperfect real-world data leads to poor performance. A loss function that focuses on minimizing the error in the micro-seismic band, or in other words, deconstructing the power spectrum and minimizing the distance between the desired and predicted spectral densities in the band.

In order to emphasize these lower frequencies when training the model, we can also whiten the input data. This will increase power in the low frequency range which will make for an easier numerical training. Hyper-parameter optimization can be developed to best capture the desired dependencies and further reduce noise in the seismic band. If these improvements can create a model that predicts the spot motion well, the results can be incorporated in order to remove the bilinear noise coupling.

[CLICK]

I would like to thank my mentor Hang Yu for the constant support and guidance this summer with my project, to Rana Adhikari for always ensuring that I understand the physical meaning and motivation behind my work, and to Zachary Mark for introducing me to Neptune and its amazing organization.

Thank you for the attention. Are there any questions?