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

Technical Note	**LIGO-T2200196-v1**	2022/05/20

Developing Deep Learning Solutions for Lock Acquisition

Peter Ma — Mentor: Gabriele Vajente

Distribution of this document:

LIGO Scientific Collaboration

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 NW17-161
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 project will look into investigating and developing deep learning techniques to approach the problem of LIGO's lock acquisition. Specifically we will look into leveraging modern techniques in attention based learning to help estimate the state of the mirrors given optical signals from the Power Recycled Michelson configuration (PRMI). We will also look into the usage of deep reinforcement techniques and how one might craft a machine learning model that is agnostic to any kind of setup for various degrees of freedom. In exploring these techniques, should any approach prove successful, the impact would directly help improve LIGO's total operational time with an upper bound of improvement of 12%, which will help accelerate the rate at which gravitational wave events are detected.

1 Introduction

LIGO stands for the Laser Interferometer Gravitational-Wave Observatory and is an observatory built to detect the spatial dilations created by Gravitational Waves (GW). These dilations are small, on the order of $10^{-18} - 10^{-22}m$ [1]. To measure these small changes we use lasers, which are setup as a Michelson interferometer to help reveal to us the small dilations of space. LIGO works primarily by splitting a single beam in two which are sent down 2 orthogonal arms that then gets reflected back. The beam that's reflected back then gets combined where the lasers interfere optically. The brightness of the beam is an indication of how far off from being in-phase the beam is. This directly shows us the relative changes in path length of the two arms thus allowing us to measure small changes in space.

However to achieve the sensitivity needed to detect these small changes, we require multiple mirrors to amplify the signal, and to increase the power circulating through the interferometer shown in figure 1. One of these sets of mirrors is called the Fabry-Perot cavity [2]. These Fabry-Perot cavities are used to recirculate the photons inside the 4km-long arms. This is so that they can accumulate more change in phase for the same displacement of the mirror created by a GW. This is because reflecting the light between the cavity increases the path length light needs to travel. Thus this turns what is technically a 4km physical arm into a nearly 1120km large "interferometer" for GW detection.

Secondly, we have a cavity called the Power Recycling Cavity shown in figure 1. This is the stage before the beam of laser enters the beam splitter. The purpose of this cavity is that, once tuned to resonance, it puts more power through the entire interferometer [2]. We bounces light back and forth which turns a 50W laser into a much more powerful beam with a gain of 50-55.

Finally we have a set of Signal Recycling Mirrors [2] shown in figure 1. This is part of the last stages of the interferometer. Here we add a cavity that has the goal of amplifying the optical signal and increase the detector bandwitdh. We can tune the resonance to choose to amplify the signal amplitude at the price of further reducing the detector bandwidth or increase the detector bandwidth at the price of a dampening of the signal.

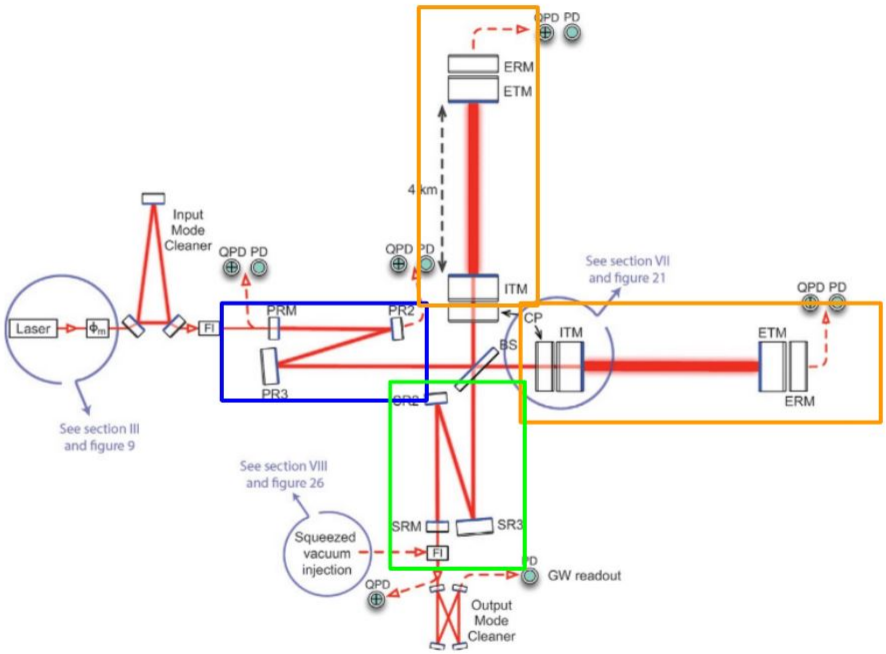

Figure 1: The blue highlighted boxes are the Power Recycled Mirrors, the orange boxes indicate the Fabry-Perot cavities and the green box indicates the signal recycling mirrors.[2]

For this project we will focus on a particular system called the Power Recycled Michelson configuration. Specifically we want to help improve putting the Dual Recycled Michelson (DRMI) in resonance, which is the area between the beam splitter and the two input test masses (ITM) as well as the two recycling cavities that aren't between the beam splitter and the two ITM. The purpose is to test our approaches to simpler setups before moving forward with various ideas. Details can be found on the figure 2.

Now, the issue with these mirrors and cavities is that these mirrors can move with seismic motions. To fix this we suspend them as these mirrors in a 4-stage pendulum above ground to passively dampen these motions. However, since we suspended them, their motion now has 6 degrees of freedom, 3 translations and 3 rotations. The pendulums allows us to reduce the motion above the pendulums resonances (greater than a few Hz). However these movements are large for motions at the resonances, on the order of a few microns which is far too large for our desired precision. We need a way to try and stabilize these motions of our mirrors. To simplify our problem let us restrict our attention to just longitudinal motion. This the motion along the direction in which the light travels and is the motion that affects our detectors the most. More specifically, there is a total of 7 mirrors. However, since we only care about the relative positions between a given pair of mirrors, we can reduce the number of degrees of freedom to the following: 1. the changes of the arm length due to GW 2. the changes between the beam splitter and the input test mass, ("short Michelson") 3. average length change of the arms, this is the Michelson interferometer (MICH) 4. the changes of the power recycling cavity length, (PRCL) 5. the changes to the signal recycling cavity length [2].

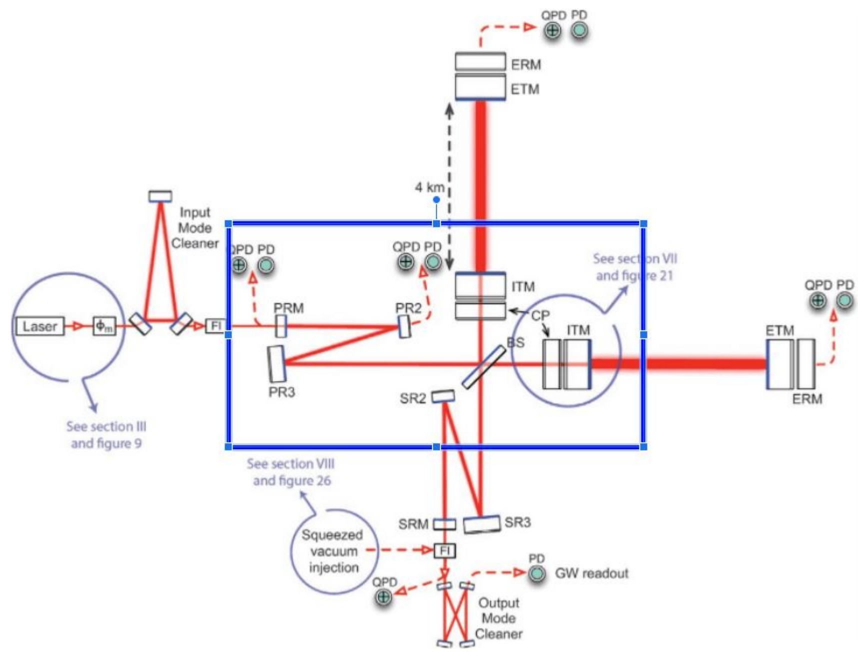

Figure 2: The blue highlighted region is the Power Recycled Michelson (PRMI). The mirrors labeled PRM stands for the power recycling mirrors and the ITM is the test mass and the BS is the beam splitter. The BS and ITM forms the short Michelson Interferometer and the PRM and the BS is the Power Recycled components.[2]

Despite this the issue with the movement of the mirrors, these motions are controllable as they are part of our suspension system. The regime when they are controllable is a linear region where we can easily retrieve the relative position of the mirrors and drive the motions close to 0. By 0 we mean the operational point, where all the resonance conditions are fulfilled to achieve power build up and high sensitivity of the instrument to GW. However, the mirrors cannot be controlled in all situations because in non-linear regions we don't know the relative positions generally.

That being said, what do these linear and non-linear regimes actually correspond to? These linear and non-linear region corresponds to the behaviour of the optical signals that we receive from each cavity. It turns out, we cannot retrieve information about the relative positions of the mirrors directly. Instead, we have signals from the lasers that are reflected within the cavities who's power corresponds to the relative position of the mirrors. The issue is that this relationship between position and the power of the signal is, in general, non-linear. There are regimes in which we know that the relationship between position and signal is linear. We can "detect" when we've achieved this is if the signals of the laser are of high power, and remains relatively constant. However, for practical cases we just require it to be high power. In this regime, the control problem is solved.

For completeness, these "optical signals" that are available to us comes from a technique called a variant of the Pound-Drever-Hall. What happens is that we inject a laser beam into the interferometer. The signal is frequency modulated such that the sidebands of the signal are not resonant with the Fabry Perot cavity. But instead we can tune the power recycling cavity such that it is resonant

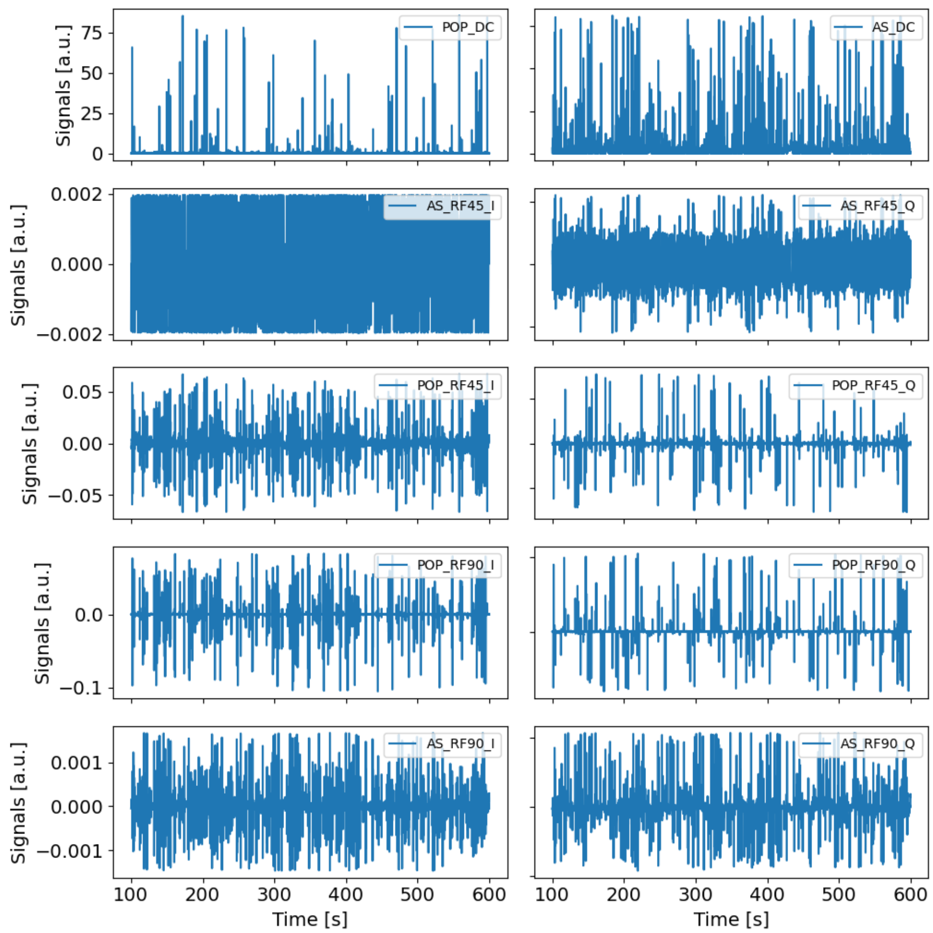

Figure 3: These 10 plots is an example of the 10 optical signals we would expect to receive. Notice each are labeled for the following ports ‘POP‘ - Power recycling cavity pick off and ‘AS‘ - stands for anti symmetrical port and ‘REFL‘ - stands for reflection. We have that ‘_ pow‘ = total power ‘45*‘ = demodulated at 45 MHz 90* = demodulated at 90 MHz. Then we have the (i) and the (q) which stands for each of the demodulated signals has two ”quadratures”. [2] The data shown are simulations of the optical signals we expect to receive.

with these sidebands. What this does is that this allows us to be sensitive to changes in the PRC and not in the Fabry Perot cavity which is subject to both GW and to various arm motions. We of course need to demodulate the signal, at the reflection port (REFL) or at the power recycling pick-off (POP) which provides signals that measure the power recycling cavity length (PRCL) etc. Specifics on what this signal actually looks like can be found in figure 3

In the end, what isn’t solved is the general case. To control these mirrors in general, we need to ”acquire the lock”. The whole purpose of lock acquisition is to develop a scheme in which we can drive the movement of the mirrors into a regime in which the relationship between signal and the relative positions have a linear behavior from a regime where the relationship is non-linear.

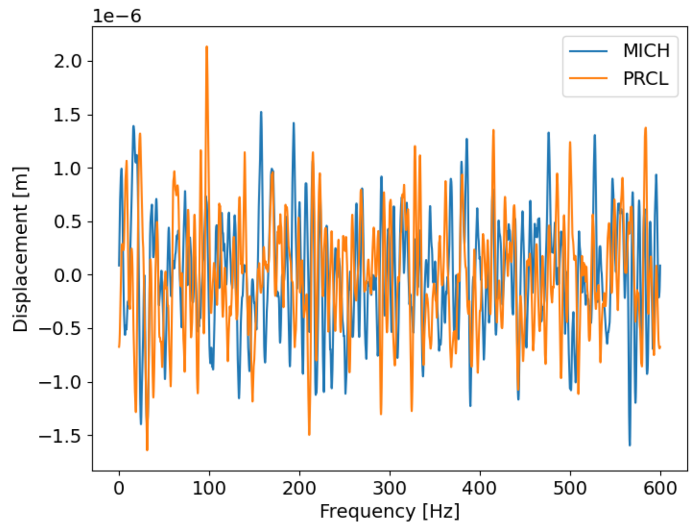

Figure 4: Here is an example of the position of simulated mirrors. This is the data we wish to recreate given the signals obtained from figure 3.

2 Problem Statement

The main problem we are trying to solve is acquiring the lock for these mirrors. As mentioned before, acquiring the lock when given the position of the mirrors is already solved! The problem is, in general, we don't know the state of the mirrors. The only data we receive are the optical signals that come from the interferometer (IFO), which is a complicated strongly nonlinear function, and the solutions are non-unique. This makes it a non linear control problem. What is even more difficult is that, attempts at linearlizing the problem works only a small fraction of the state space. The probability that the mirrors fall within the linear regime by chance is 10^{-9}. We can reduce the problem down to, **how can we develop a means of constructing the current state of the mirrors given this optical signal input?** Here is an example of what we want, we want to produce the relative motions of the mirrors, i.e the state of the mirrors shown in figure 4

Some requirements that we expect a solution to have:

· **1** Predict / infer the positions of the mirrors accurately

· **2** Run fast enough on real time data as to match the sampling rate

· **3** The predicted states must be a continuous mapping of signals to the states

These requirements suggests the end to build some kind of continuous non-linear state estimator. In this context, when we mean continuous, we mean that small changes in the data results in small changes to the outputs. A direct consequence of this constraint is that we expect the predictions to not include nonphysical jumps in the estimated positions. Example of these unphysical jumps

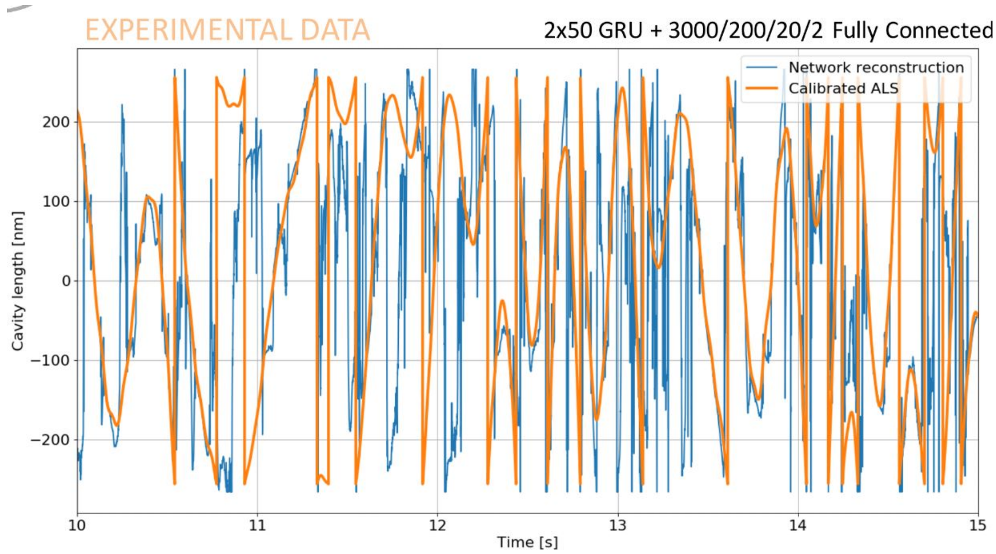

Figure 5: This is a sample of the state reconstruction performed by previous ML models. The jump of data points between -200nm to 200nm at short intervals of time are examples of these unphysical jumps that we wish to avoid. Note this isn't data from the PRMI but from the Fabry-Perot cavities.

created by previous ML based reconstructions of the mirrors states can be seen in figure 5. Thes are due to the periodic conditions and to having trained the deep neural network with wrapped positions

Additional elements to consider:

· **Memory** The model might need to have a mechanism or attention mechanism for the data input.

· **Online Learning** We might have to tune model in an online way

· **Small Models** Build small models and employ inductive biases in order to make things run faster with simpler solutions

As further explanation, there are a few additional reasons to leverage historical data when making inferences. Firstly, when building the model, the nature of the input data is that it is relatively small and ambiguous if taken out of the context from a history of past signals in past time steps. Thus if we were to build a model on just looking at individual samples it would be a more difficult challenge. Furthermore, we know before hand that the data is time series. We know the data represents a complicated nonlinear mapping of the positions of the mirrors of which follow some kind of dynamics i.e the positions evolve as a function of time. Thus it is to our advantage to construct a model that can exploit this assumption about the data when making an inference. We can exploit it either implicitly by feeding data that contains historical information or explicitly by "baking it into design" such that the architecture utilizes the temporal data in a sequential manner. An example of these sequential techniques are recurrent models.

3 The Status Quo

Currently the classical approach to this problem in the configurations for the Power Recycled Michelson (PRMI) and the Dual Recycled Michelson (DRMI) is to use linear combinations of these optical signals to reconstruct the mirror motions specifically when triggered by resonance crossing. This resonance crossing is when carrier or sideband power crosses a threshold. This is not ideal, firstly because there is a low probability that this might happen. Secondly the controllers have a short time to decelerate the mirrors to maintain this lock.

Previously, deep learning solutions to tackle this problem, researchers have developed a non-linear continuous state observer for the IFO system in an attempt to reconstruct these positions from the optical signals.

The approach uses Gated Recurrent Units (GRU) and deep learning techniques to tackle the non-linear nature of the error signals. This approach although relatively fruitful in demonstrating its ability in reconstructing these positions from simulations, still has problems. How do we to enforce the continuity of the output signals: some times they have large nonphysical jumps, see figure 5.

The root cause of the issue is that for any given signal, multiple positions can correspond to the same signal.The solutions are only unique up to a multiple of half the wavelength of the laser, in the case of the Fabry-Perot case. It is a bit more complicated in our setup but the essence of the problem is the same[1]. So if we want to devise an algorithm that inverts the problem, we are restricting the algorithm to just return the solutions from 0 to 1/2 wavelength instead. But, this would introduce sharp changes in actual positions because clearly any mirror can move beyond the restricted region, we just choose to shift all solutions between 0-1/2 wavelength solutions. To recover this nonphysical adjustment we apply something like the wrapping function, which detects this discontinuity and tries to adjust the solution such that the discontinuity is gone by effectively stitching the solutions back such that its continuous. The problem is that the current deep neural net fails because the solutions are smoother than what we can detect as being a discontinuity.

Furthermore, the models were relatively large and running them in real time might present itself as a challenge. We can see the architecture of the model in figure 6

These GRU's were trained on simulated data. Currently we have simulations that, when given the relative positions of the mirrors, it can simulate the signals we would typically expect from the signal readouts. This will form the basis of the training data for any kind of machine learning technique.

[1] https://gitlab.com/gabrielevajente/prmi-ml

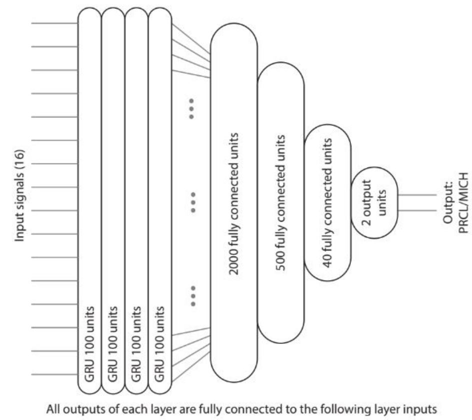

Figure 6: This is the GRU model architecture from previous attempts at producing a continuous state estimator.

4 Proposed Solutions

There are a few directions to approach this problem. One approach the problem in a encoder decoder model in a supervised learning setting. The second approach is to develop a similar model as the GRU except we use attention based recurrent networks instead of a simple gated recurrent network. Lastly, we can approach the problem to learn general controls in a reinforcement learning, online learning fashion. Here are the details of each approach.

4.1 Time Series Encoder / Decoder Method

Firstly we will make use of the given simulations from the mirrors. We will start off with a simple neural net that will take in time series data over a certain interval and we train it to recreate the positions of the mirrors. Then we tune the model to real data by placing it in an encoder decoder setup. This works by trying to output the positions of the mirrors during this time. We will then feed this back into the simulation and try to recreate the original signal in which we feed the data in. The errors created between the two are used to update the neural network. The benefit is that we can frame the problem into a simpler supervised learning encoder decoder format almost akin to variation of an autoencoder. These kinds of algorithms are well studied in literature and would be interesting to leverage these techniques. [3]

4.2 Attention Mechanism

Attention based models and Transformers are models that have taken the world of ML by storm [4][5]. The incredible ability to selectively tune focus on certain areas of data has rendered memory issues with previous recurrent neural networks an almost trivial problem. With this recent movement, attention mechanism could be a viable direction of investigation for state estimations. The goal is to replace the multiple GRU layers with with a multi-headed attention layer in order to decrease the size of the model. The hope is that this allows us to extend the scope [length of time] at which the model can take in data which can potentially produce more favourable results. Previous GRU's suffered from vanishing gradients when we backpropagated too far in time.

Furthermore, we can try to implement the same kind of regularization to ensure continuity as discussed in the previous online learning strategy.

4.3 Deep Deterministic Policy Gradient

If time allows we can consider the following:
Deep Deterministic Policy Gradients (DDPG) is a model-free off-policy algorithm for learning continous actions. It combines ideas from DPG (Deterministic Policy Gradient) and DQN (Deep Q-Network) [6]. It uses networks from DQN, and it is based on DPG, which can operate over continuous action spaces. We can try to change the supervised learning problem into an RL problem. We can attempt to transform the supervised approach into an RL problem by setting the action as the predicted state, and the state as the input signal.

Then we have an Actor. The Actor will be the model that predicts the position of the mirrors. The Critic could be the simulation that takes in the positions and recreates the supposed signals. The reward would the losses or in this case the reciprocal of the losses. This is different as the value of the action (the predicted state) depends not depend on the future state of the signal, but on the current state of the signal. The rest of the approach follows that of a DDPG model.

Another formulation is for the RL model to learn directly from interacting with the controllers. By controllers we mean the actual inputs that move the physical mirrors. So the approach will be bypassing creating the state of the mirrors by directly informing how the physical mechanisms of the mirrors should react. We develop the critic to evaluate how well the model is ”in the linear” regime. This can be gauged by the time spent in this linear regime. This way the model can learn how to act by playing the these controls. The benefit is that this would in theory make this ML model agnostic to any setup! With other approaches, we are limited to the simulations on various simplified setups. To develop a solution for the full problem, we'd need to then develop more complicated simulations for the entire setup, furthermore should any configuration change we'd have to adjust simulations to accompany this. Should a solution exist, the benefit with the RL model interacting with controllers is that no simulation is even required. This allow might allow us to generalize the solution better.

The reason we choose to approach it this way is because the action space is continuous. In other words, we can expect if the input states are continuous, and change very little between time steps, that the outputs would also change by small amounts. This might require us to abandon the approach where we compute positions in batches of time intervals, instead we predict positions sample by sample. By sample by sample, we mean that historical data is inputted but the output would be a single sample. Note we will still need historical data as to avoid trivial mappings of single samples to sample.

4.3.1 Online/Active Learning

We can also try approaching the various techniques with active learning specifically by updating the network live while making predictions[7].

The intention with approaching the problem with this direction is that it might not require us to hold memory of all the previous time series data since the model adapts and tunes the model on real data on the fly. The hope is that the model only needs to fit a small subset of the nonlinear relationship between the relative positions and the signals produced because its forced to fit to the data locally in time. Rather than building a model that fits to the data globally in the situation where we separated the training the execution by freezing the weights. Furthermore, regularization of nonphysical jumps is easier to control in this approach. We can penalize the model on the fly for when predictions are clearly nonphysical [at least we can try!].

5 Foreseeable Roadblocks

For the first RL approach, disregarding an explicit memory mechanism might be an issue. We might need to devise a means of passing information from previous snapshots in time series data. Furthermore in the online setting, controlling model might be difficult. If model evolves through time by updating the network on the fly, it's not always guaranteed it will behave in the way we intend it will. We might need to develop a series of checks and balances to guarantee the model doesn't produce unwieldy reconstructions, or have "backup" models.

Furthermore, the RL approach might also be difficult specifically in constructing a value function that instructs the policy on the kinds of outputs we should be expecting. Intuitively we would frame the value function as

Lastly, the approach using encoder decoders still lacks a very satisfying means of ensuring continuity. Furthermore, the solutions are not unique, and so building a one model fits all might not solve our problems in the way we intend it.

6 Timeline

This is the current running timeline for the project. The starting week would be in week of June 13th. Plan is to implement the online learning the approaches using JAX [8] and HAIKU [9] and the rest of the training for the Attention based RNN models either also JAX/HAIKU or KERAS

[10]. The reason for this choice is the flexibility in designing models and is the new trend of ML based frameworks.

· **Week 1** Review code base and re-implement previous GRU in KERAS framework [10]

· **Week 2** Benchmark KERAS framework (make sure they perform the way we expect it). Build environment for online learning simulation.

· **Week 3** Implement encoder decoder strategy

· **Week 4** Buffer week - in case results aren't as expected.

· **Week 5** Start attention RNN model approach

· **Week 6** Fully implement + tune the model

· **Week 7** Evaluate the performance

· **Week 8** Buffer week - if extra time look into using RL models

· **Week 9** Wrap up results and begin write up or presentation

· **Week 10** Discuss next steps and finish outstanding deliverables.

7 The Impact

There are two broad areas of impact for this kind of work should we be successful. Firstly, this will directly help LIGO operate at longer observational times. Currently, classical systems take approximately 12% of the time to acquire the lock and to drive the motions of the mirrors down to an operational regime, during which the detector cannot observe. Should we be successful in developing an effective locking acquisition scheme, we can work towards lowering the 12% standby time in hopes of increase in observing time. This is under the assumption our technique generalizes to LIGO's more complex configurations. We currently estimate, with back of envelope calculations, that we loose out on ~ 11 events, given a total 12% downtime and having discovered 90 events in the recent GWTC-3 catalogue release.

The second impact would be the demonstration of using ML based techniques in controlling state of the art equipment. The machine learning and reinforcement learning literature have been notoriously known for making improvements in simulated and game-like environments with few applications real world control problems. The reason why use cases of ML in real world control problems remain rare is mostly due to the simulation to real challenges (sim2real gap). Building such a successful model or approach would be an incredible demonstration of real world applications of the theoretical advancements in the field of ML and control theory to hard control problems.

8 Forward

Moving forward, the goal for this project is to experiment with these to ideas and to investigate the benefits and the draw backs of each approach. Future plans maybe to incorporate this kind of algorithm on table-top experiments or with the 40m prototype to test the performance of our approach in the real world.

References

[1] Bassan, M. (ed.) *Advanced Interferometers and the Search for Gravitational Waves* (Springer International Publishing, 2014). URL `https://doi.org/10.1007/978-3-319-03792-9`.

[2] Vajente, G., Gustafson, E. K. & Reitze, D. H. Precision interferometry for gravitational wave detection: Current status and future trends. In *Advances In Atomic, Molecular, and Optical Physics*, 75–148 (Elsevier, 2019). URL `https://doi.org/10.1016/bs.aamop.2019.04.002`.

[3] Peng, P., Jalali, S. & Yuan, X. Solving inverse problems via auto-encoders (2019). URL `https://arxiv.org/abs/1901.05045`.

[4] Niu, Z., Zhong, G. & Yu, H. A review on the attention mechanism of deep learning. *Neurocomputing* **452**, 48–62 (2021).

[5] Vaswani, A. *et al.* Attention is all you need (2017). URL `https://arxiv.org/abs/1706.03762`.

[6] Lillicrap, T. P. *et al.* Continuous control with deep reinforcement learning (2015). URL `https://arxiv.org/abs/1509.02971`.

[7] Si, J. & Wang, Y.-T. Online learning control by association and reinforcement. *IEEE Transactions on Neural Networks* **12**, 264–276 (2001).

[8] Bradbury, J. *et al.* JAX: composable transformations of Python+NumPy programs (2018). URL `http://github.com/google/jax`.

[9] Hennigan, T., Cai, T., Norman, T. & Babuschkin, I. Haiku: Sonnet for JAX (2020). URL `http://github.com/deepmind/dm-haiku`.

[10] Chollet, F. *et al.* Keras. `https://keras.io` (2015).