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

Technical Note **LIGO-T2600198–** 2026/11/01

A System Identification Framework for Technical Noise Reduction in the Caltech 40m Prototype

Michael Padilla[1], Rana Adhikari[2], Nathan Holland[2], Christopher Wipf[2]

[1] *University of California Riverside,* [2] *California Institute of Technology*

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/

Contents

1 Introduction **2**

2 Caltech 40m Input Mode Cleaner **4**

2.1 Mechanical Model 5

2.2 Radiation Pressure 6

2.3 Signal Modeling 8

3 Implementation in LIGO Real-time System **10**

3.1 Local Mirror System Modeling 10

3.2 Global Feedback Modeling 12

3.3 Validation 14

3.3.1 Step Responses 15

3.3.2 OLTF Measurements 15

3.3.3 Radiation Pressure Measurements 17

4 System Identification **19**

4.1 Manual sysID 20

4.2 Optimal Multisine sysID 21

4.2.1 The Pintelon-Schoukens Algorithm 22

4.2.2 Implemented Optimal Multisine Algorithm 23

4.2.3 CDS Backend Testing 23

5 Conclusion **26**

5.1 Next Steps 27

6 Acknowledgments **28**

A Closed-loop Plant Estimation **28**

Abstract

To achieve the sensitivity targets of future gravitational wave detector upgrades, low-frequency technical control noise must be significantly reduced. We propose the development of an accurate digital twin model. This model will serve as the environment for fast offline training of a reinforcement learning-based control policy that can be implemented directly into the real system. This report details our development of a physics-based digital twin of the Caltech 40m Prototype Input Mode Cleaner — a triangular, suspended optic cavity. We implemented a gray-box model directly within LIGO's real-time control architecture. In doing so we constructed the complete signal path to accurately capture the dynamics of the digital controller and electronics alongside the physical suspended optics. We also demonstrate the implementation of an optimal system identification algorithm for measurements of the 40m system. This algorithm allows for fast and optimal measurements that will serve as the basis for further validation and fitting of the digital twin, enabling a high-fidelity training environment for effective control policy implementation onto the real system.

1 Introduction

Since 2015, LIGO has detected dozens of astrophysical transients, providing significant insights into objects such as black holes and neutron stars [1–4], while additionally allowing us to test prominent theories and models like strong-field general relativity [5, 6]. To further push the detection horizon of gravitational wave interferometers, upgrades to LIGO and next-generation detectors are envisioned to come online in the next decade, focusing on improvements in sensitivity and more efficient power build up [7, 8]. Present noise budget considerations are categorized as high frequency (> 300 Hz), mid frequency (50-300 Hz), or low frequency (< 50 Hz). High frequency noise is primarily dominated by quantum shot noise, which will be suppressed by the upgrade's higher power and squeezing targets. Mid frequency noise is addressed through changes to test-mass coatings which would lower coating thermal noise. At low frequencies, the primary consideration of these future upgrades involves changes to the test masses and suspensions to reduce radiation pressure and seismic noise. However, without a considerable reduction in low-frequency technical noise to take advantage of reduced fluctuations in the interferometer's degrees of freedom, these design considerations alone will not reach sensitivity targets [7]. This project directly addresses this contribution to future design considerations through the implementation of a modeling to reinforcement learning pipeline that trains a more optimal control policy for lower control noise.

Contemporary control systems used to lock or control the interferometer's degrees of freedom are required for sensitive operation and can stabilize otherwise unstable systems, but they also inject sensor noise back onto the system. This is a consequence of a fundamental tradeoff: a control loop requires sufficient gain over a target bandwidth, yet outside that band its output should be attenuated as quickly as possible to avoid re-injecting sensor readout noise. Linear controllers cannot escape this tradeoff between loop stability and how sharply their response rolls off, so some noise injection is unavoidable. Recent work on a new approach to solving this problem has shown that the implementation of reinforcement learning (RL) algorithms as non-linear controllers can reduce injected control noise by more

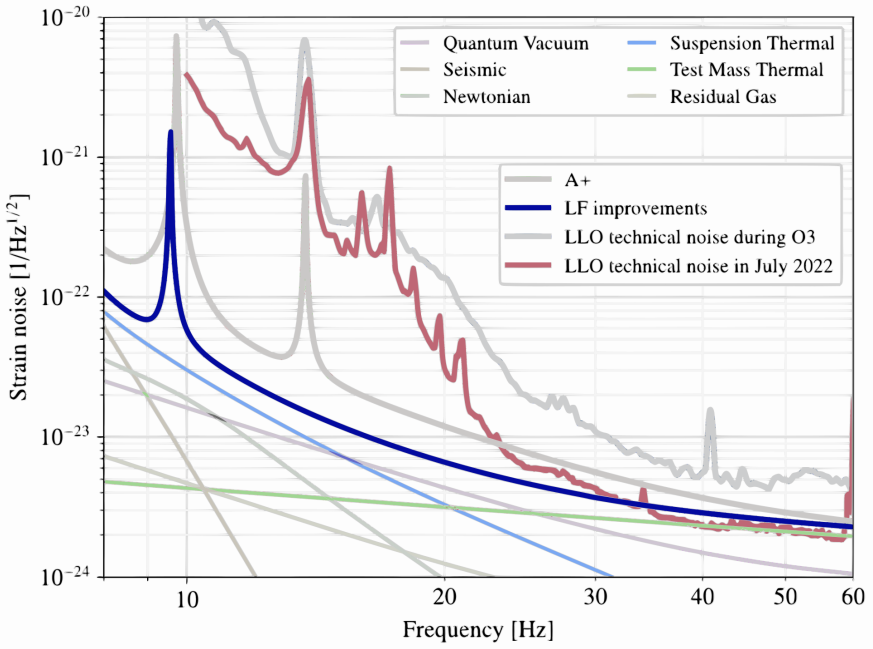


Figure 1: The dark blue curve is the theoretical sensitivity of LIGO following the improvements of the A# upgrade, including improvements to suspension and test mass noise sources (light blue and green). In contrast, technical noise measured at the Livingston site in 2022 is significantly greater than this fundamental noise floor. Figure modified from [7].

than a factor of 30 [9]. This was previously shown in the common hard pitch control loop in the LIGO Livingston facility. In order to create an effective RL model, its training environment must accurately model the behavior of the system it will control. Creating accurate models that reflect the behavior of real-world counterparts is especially difficult in the complex optical subsystems of the LIGO interferometer. In particular, this is because of the complex non-linear dynamics of these systems, which may be difficult to reproduce with a purely ground-up physics-based simulation. For this reason, this project explores an alternative approach to modeling, system identification, which will be explored to create accurate simulated models to facilitate a successful sim-to-real transfer of an RL agent policy.

In this report, we develop the modeling of the digital twin through derivation of radiation pressure in the linear-regime of the IMC. Additionally we port the modeled IMC physics into a Simulink model using the `rtsfreerun` tool with accurate modeling of signal processing electronics and digital filters. We detail system identification measurements taken to validate the `rtsfreerun` model, showing qualitative agreement between model and the real system. Finally, we show successful implementation of an optimal system identification algorithm on one of the 40m IMC angular control loops, demonstrating a method for automatic system identification to speed up future digital twin parameter fitting efforts.

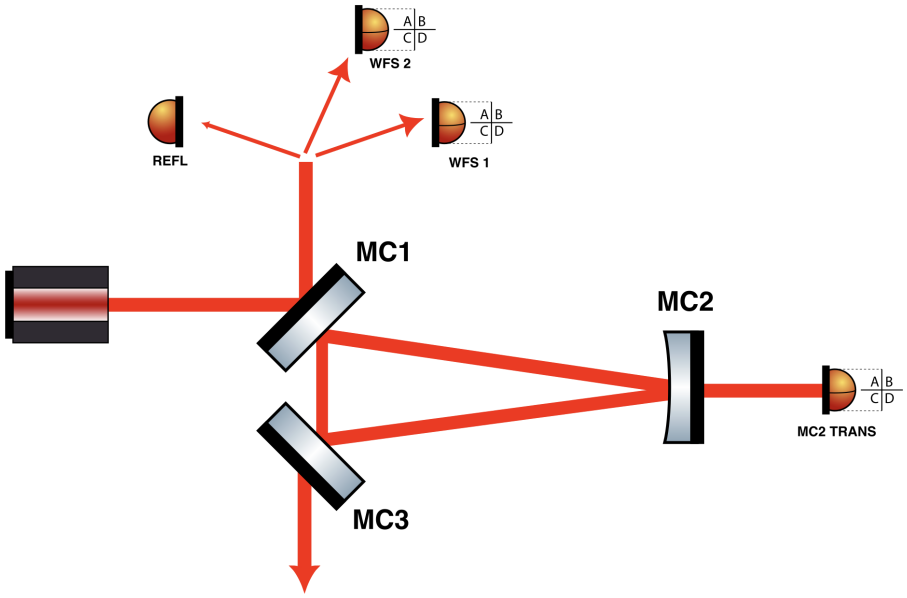


Figure 2: A simplified diagram of the 40m IMC and relevant sensors. The IMC is a triangular cavity composed of mirrors MC1, MC2, and MC3. The input laser enters through MC1 and exits through MC3. The reflection beam off of MC1 is sensed by the REFL PD and two quadrant photodetectors: WFS1 and WFS2. The transmission beam from MC2 is read by the MC2 TRANS quadrant photodetector.

2 Caltech 40m Input Mode Cleaner

The development and implementation of the modeling framework discussed above was focused on the input mode cleaner (IMC) in the Caltech 40m LIGO prototype. The IMC is a triangular suspended optic cavity in the input path of LIGO (see Figure 2). The main purpose of the IMC is for spatial mode cleaning, suppression of laser noise sources, and polarization filtering [10]. With various coupled dynamics and over a dozen controlled degrees of freedom, the IMC serves as a sufficiently complex system to develop the modeling and measurement framework that can be transferred to other LIGO subsystems. Our model of the IMC is limited to linear models of the suspension-optic systems, radiation pressure, and feedback loops. These are the primary physical mechanisms that influence the relevant degrees of freedom that an RL policy would control. Below we will detail the physical model used as the basis for digital twin development.

Each IMC mirror is held by a single pendulum suspension (see Figure 3). These suspensions are attached to suspension towers which are tightened to passive seismic isolation tables inside the vacuum system. The response of the mirror to direct actuation as well as background noise such as seismic motion is especially important to understand. This is because actuation and noise sources move the mirrors and affect cavity properties such as power gain and the cavity eigenmode. Thus, much of the physical modeling for this project is developed with the goal of accurate modeling of individual optic motions. There are three primary degrees of freedom for each mirror which we will consider: longitudinal (position), as well as pitch and yaw motion.

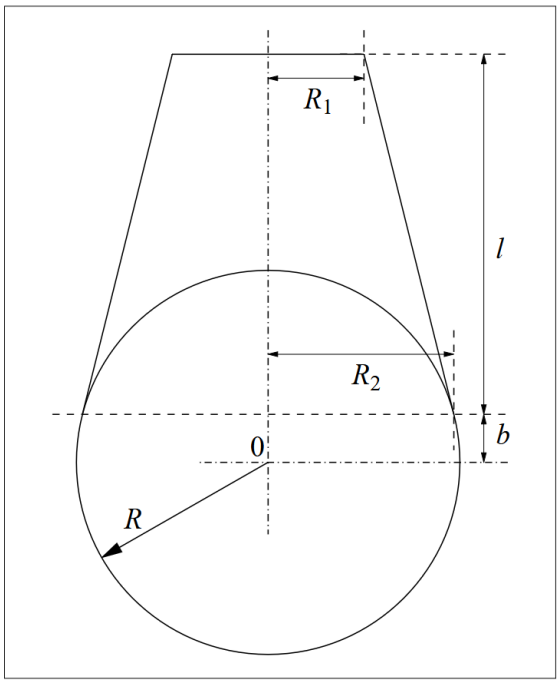


Figure 3: Technical drawing of a small optic suspension (applies to the 40m IMC mirrors) detailing relevant dimensions of the optic and suspension. IMC mirrors are hung by a single suspension. Figure sourced from [11].

Each mirror is controlled by five Optical Sensor and ElectroMagnetic actuators (OSEMs), four on each corner of the mirror face and one on the side. The four corner OSEMs actuate on the aforementioned three degrees of freedom, and so again these OSEMs will be of primary consideration for the modeling of electronics and control loops.

2.1 Mechanical Model

Our model of the optic-suspension motion of IMC mirrors is detailed in [11], which derives the dynamical equations and transfer functions of a single optic given four degrees of freedom: mirror position, pitch, and yaw, and suspension point longitudinal position. Approximating all motion to be small, the corresponding system is linear. From the derived Lagrangian of this system, the authors obtain three equations of motion of the mirror which correspond to its three degrees of freedom:

$$\ddot{x} + \gamma_x \dot{x} + \omega_x^2 x = \omega_x^2 (x_{sp} + b\theta) + f_x(t)$$

$$\ddot{\theta} + \gamma_\theta \dot{\theta} + \omega_\theta^2 \theta = \frac{\omega_\theta^2}{l+b}(x - x_{sp}) + f_\theta(t)$$

$$\ddot{\phi} + \gamma_\phi \dot{\phi} + \omega_\phi^2 \phi = \omega_\phi^2 \phi_{sp} + f_\phi(t)$$

Where x is the longitudinal position, θ is pitch, ϕ is yaw, and x_{sp} is the longitudinal position of the suspension point. It can be seen through these equations that position and pitch

motions are coupled. Additionally suspension point motion couples directly into the system introducing seismic noise into the system. For reference on the definition of system parameters, see [11]. Further detail of the derivation and analysis of the resulting mechanical model is discussed in [12].

2.2 Radiation Pressure

Radiation pressure within a cavity directly excites multi-mirror coupled angular modes and induces a longitudinal optical spring effect. This derivation focuses on the induced angular modes, as the optical spring effect is shown to be negligible for this system.

A beam displaced from the geometric center of a mirror induces a torque:

$$\tau = \frac{2Px}{c}$$

Where P is the circulating power, x is the beam position displacement from the center, and c is the speed of light. In order to accurately model the resulting torques on each mirror in the IMC, we need a model of the relationship between beam position displacements and the degrees of freedom of each mirror. Previously this was done in [13] by utilizing the model for a triangular cavity constructed in [14, 15]. However, the more recent analysis in [16] provides an accurate and more simple linear approximation of a triangular cavity, which has already been assumed in the model of WFS signals for angular control loops.

As detailed in [16], the beam displacements on the mirror plane are given by:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = -\frac{R}{\sqrt{2}} \begin{pmatrix} g & \sqrt{2} & g + \frac{2\ell}{R} \\ 1 & \sqrt{2} & 1 \\ g + \frac{2\ell}{R} & \sqrt{2} & g \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \sqrt{2}\xi & -\frac{\sqrt{2}\ell}{g} & -\sqrt{2}\xi \\ -\frac{\ell}{g} & \frac{L+\ell}{g} & -\frac{\ell}{g} \\ -\sqrt{2}\xi & -\frac{\sqrt{2}\ell}{g} & \sqrt{2}\xi \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix}$$

Where x is the horizontal displacement and y is the vertical displacement. θ is the pitch of the mirror (counter clockwise about the $+x$ axis of the mirror) and ϕ is the yaw of the mirror (counter clockwise about the $+y$ axis of the mirror). The remaining parameters are geometric. Parameters L and 2ℓ are the long side and short side cavity lengths respectively. R is the radius of curvature of mirror MC2. Finally g is the g-factor of the cavity defined as $g = 1 - \frac{L+\ell}{R}$ and ξ is defined as $\xi = L + \frac{\ell}{g}\left(1 - \frac{L}{R}\right)$.

We can write the resulting torques on the mirrors through the following vector equations:

$$\vec{\tau}_{pit} = I_{pit}\ddot{\vec{\theta}}$$

$$\vec{\tau}_{yaw} = I_{yaw}\ddot{\vec{\phi}}$$

Thus we get the following linear coupled differential equations:

$$\begin{pmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \\ \ddot{\theta}_3 \end{pmatrix} = \frac{PR\sqrt{2}}{cI_{pit}} \begin{pmatrix} \frac{g}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}}\left(g + \frac{2\ell}{R}\right) \\ 1 & \sqrt{2} & 1 \\ \frac{1}{\sqrt{2}}\left(g + \frac{2\ell}{R}\right) & 1 & \frac{g}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix}$$

$$\begin{pmatrix} \ddot{\phi}_1 \\ \ddot{\phi}_2 \\ \ddot{\phi}_3 \end{pmatrix} = \frac{2P}{cI_{yaw}} \begin{pmatrix} \xi & -\frac{\ell}{g} & -\xi \\ -\frac{\ell}{g} & \frac{L+\ell}{g} & -\frac{\ell}{g} \\ -\xi & -\frac{\ell}{g} & \xi \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix}$$

For the pitch differential equations an additional sign was accounted for due to the angle definition.

It is of note that these geometric parameters (ξ, g, L, ℓ) are dependent on the relative displacements of the mirrors. This would result in non-linear coupling of the position and yaw degrees of freedom in the model which have previously been completely decoupled. This coupling is additionally introduced through the circulating power's dependence on the relative displacements of the mirrors. However in the case of circulating power, when the cavity is locked on resonance this is a second order effect, so we can safely ignore the longitudinal optical spring effect that results from this displacement dependence. Additionally for the geometric parameters dependence, because we only expect the mirrors to move on the order of micrometers, the relative change in these values is roughly on the order of $10^{-6} - 10^{-7}$, thus this effect is similarly negligible.

Transforming the matrix into the eigenbasis and evaluating for the 40m IMC parameters, we find two negative eigenvalues and one positive eigenvalue for pitch motion indicating two stable and one unstable mode for the pitch degrees of freedom. We find one negative eigenvalue and two positive eigenvalues for the yaw motion indicating one stable and two unstable modes for the yaw degrees of freedom. For the 40m IMC with 500 W of resonating power, however, the effect of radiation pressure is small and not enough to change the stability of angular modes.

Using these radiation pressure induced torques, we can evaluate changes to the pitch and yaw transfer functions of all three mirrors in the IMC. Figure 4 shows the 3x3 matrix of transfer functions for all three mirror pitch and yaw degrees of freedom respectively.

We see that the incorporation of radiation pressure in the IMC results in coupling of separate mirror degrees of freedom as well as a shifting of resonance frequencies across all three mirrors.

The outputs of a numerical calculation of the shifts in the poles of the pitch and yaw degrees of freedom are shown in Table 1.

Table 1: Optomechanical pole frequencies and their shifts from the unperturbed (radiation-pressure-free) resonances for the pitch and yaw angular degrees of freedom, computed for the 40m IMC at 500 W circulating power.

DOF	**Pole**	**Frequency (Hz)**	**Shift (mHz)**
Pitch	1	0.7558	−17.2
	2	0.7771	+4.1
	3	0.7731	+0.1
Yaw	1	0.8560	+0.0
	2	0.8431	−12.9
	3	0.8337	−22.3

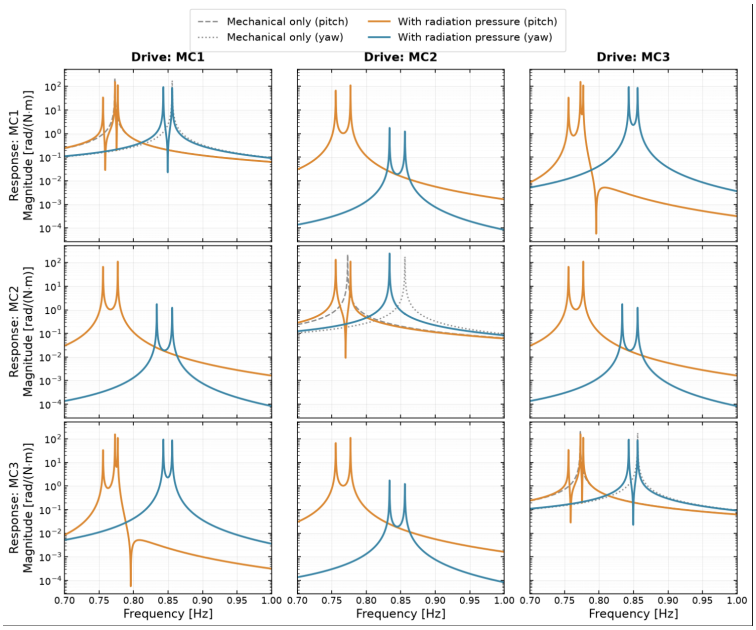


Figure 4: A visual of the radiation pressure modified plant transfer functions for the IMC structured by (input mirror, output mirror). This plot shows the introduction of mirror couplings and a lifting of resonance frequency degeneracy under the assumption of three identical optic-suspension systems. Note that although pitch and yaw are depicted on the same plots, pitch and yaw motions are fully decoupled in our model.

2.3 Signal Modeling

In addition to modeling of the underlying mechanical system, we model the processing of digital signals to complete the feedback loops. Each feedback loop is the series and parallel combination of various electronics and filters. We can categorize this modeling into: sensing and actuation, analog signal processing, analog-digital signal conversion, and digital signal processing. A simplified diagram of the signal path for these feedback loops is shown in Figure 5.

Sensing and actuation for the local damping loops is done with the OSEMs. OSEMs both sense and actuate each mirror: a flag on the mirror partially blocks an LED–photodiode beam, so the photocurrent measures the mirror's displacement over the flag's central linear range, while an adjacent coil drives a magnet on the mirror to apply a force. OSEM coils convert current to a magnetic field that pushes on these magnets. Voltage signals are converted into coil currents using the HAM-A coil driver. The conversion from input voltage to coil current is determined by the coil driver's transfer function and total impedance. From component values for the coil driver specified in [17], along with specific resistor component values specified in [18], we found the transconductance of the driver to be 11 mA/V for DC signals (Mechanical control loop signals of ∼1Hz can be approximated to the DC regime). Coil current to force requires a more detailed calculation which is not detailed in this report. However, document [17] reports a value of 0.0158 N/A for AOSEMs and 0.021 N/A for BOSEMs. Although the OSEMs in the 40m IMC are not BOSEMs, this conversion factor for BOSEMs is roughly assumed (0.02 N/A) for our initial baseline model of the system.

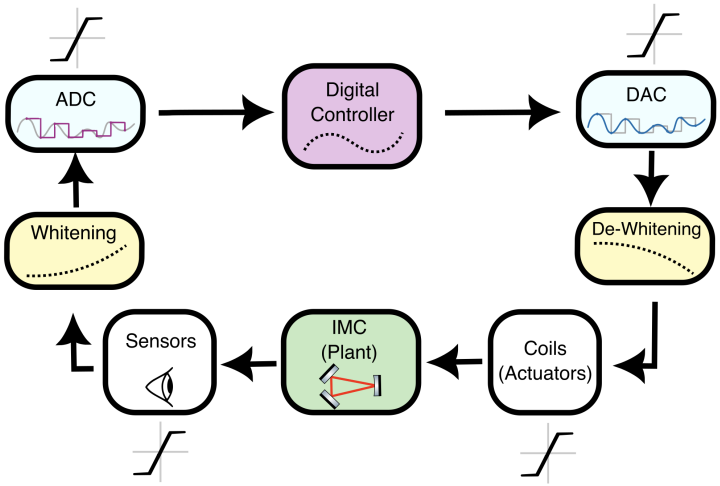


Figure 5: A simple diagram detailing the components of the signal path in the IMC system. Error signals begin as physical disturbances to the IMC degrees of freedom, converted to analog electronic signals with sensors. These signals pass through whitening filters and are digitized by the ADC, following which they are processed by digital filters. The resulting digital signals are converted back to analog before being passed to actuators which physically move the IMC mirrors. This signal geometry is general enough to describe all feedback loops for the IMC. As indicated by the diagram the ADC, DAC, sensor, and actuator components all have physical limits which can clip or distort signals, an important feature which our model handles.

Analog signals are processed using the whitening and de-whitening technique. The goal of whitening is to normalize signal power in each frequency bin such that frequencies of low signal are not overpowered by electronic noises. Whitening analog filters are used on the error signals measured by OSEMs, while de-whitening filters are used on analog feedback signals before being passed to the actuators. Each of these analog whitening filters is paired with a corresponding inverse digital filter to correct for whitening when the signal is digitized. For the IMC control loops, analog whitening filters are used while analog de-whitening filters are not being used.

To convert between analog and digital signals, analog-to-digital converters (ADC) and digital-to-analog converters (DAC) are used on the sensing and actuation sides respectively. The digitized signal is referred to in units of digital counts. The ADCs and DACs used for IMC feedback loops are 16-bit with a max voltage of ± 10 V. Therefore, the conversion from counts to voltage is $20\ \mathrm{V}/2^{16} \approx 3.05 \cdot 10^{-4}\ \mathrm{V/ct}$.

The IMC has 13 independent control loops, each controlling a separate degree of freedom (although not each of these degrees of freedom are independent). Nine of these control loops are damping loops which suppress the individual three degrees of freedom for each of the three mirrors. The remaining four control loops lock global, or cavity, degrees of freedom. These degrees of freedom include three alignment degrees of freedom and the round-trip length of the cavity. Each control loop is closed with a digital signal path which often contains multiple digital filters.

Table 2: Actuation conversion factors for the IMC control loops, from digital counts at the DAC to physical force on the mirror.

Stage	**Value**
DAC (counts → voltage)	3.05×10^{-4} V/ct
Coil driver (voltage → current)	10.9 mA/V
Coil + magnet (current → force)	0.02 N/A
Net (counts → force)	6.6×10^{-8} N/ct

3 Implementation in LIGO Real-time System

Effective sim-to-real transfer of a simulated control policy necessitates accurate modeling of the real environment it will be deployed on. In the case of the IMC, this is the digital controller. Digital feedback loops in LIGO are operated using the advLIGOrts tool [19], which compiles Simulink models into C code that is deployed to control computers for fast feedback (16384 Hz). This framework allows for easy manipulation of high level digital signal paths without a reduction in controller speed. The resulting RL control policy will be wrapped in this advLIGOrts tool, so we must be sure that the surrounding logic is being accurately modeled in simulated training. To do this, we utilize the `rtsfreerun` tool [20], recently developed by Christopher Wipf. The rtsfreerun tool wraps a compiled advLIGOrts model into a Pythonic object that can be simulated faster than real time. We additionally utilize the in development Model Builder tool [21] also by Christopher Wipf which allows for development of Simulink models with native support for custom CDS Simulink parts used in the real-time system. This section details the development and analysis of the rtsfreerun environment for the IMC.

The full IMC rtsfreerun model includes the model for the plant and the signal path for the 13 digital control loops. To construct this model, we built a Simulink model that contains plant dynamics, electronics, and digital logic. Figure 6 visualizes the high level structure for this modeling. We by convention chose to split the IMC model into 4 main subsystems: the three mirrors and global feedback. The mirror subsystems contain the mechanical plant model for each mirror as well as local damping loop signal processing. The global feedback subsystem contains signal modeling for global control loops (LSC and ASC loops) as well as global physical effects (radiation pressure).

3.1 Local Mirror System Modeling

The core of the local mirror subsystems is the model of the mechanics. This subsystem includes a plant for one of the mirrors, conversion matrices between the sensing and mechanical degree of freedom bases, and additional digital processing filters. We model the physical system using a Simulink state space block, which takes in state space matrices given by the state space formalism. We can obtain these matrices through a simple transformation of our coupled differential equations by defining a state vector x, an input vector u, and an output vector y. In the state space model, x represents the internal state of the system of interest, u represents a system input in the actuation basis, and y represents the observed state of

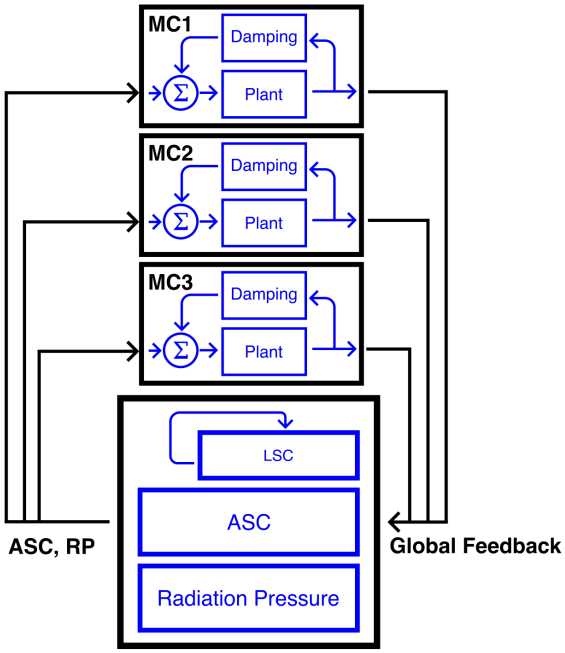


Figure 6: Simplified diagram of the global feedback topology used in the `rtsfreerun` model. Each suspension-optic system functioned as an independent system with in ports and out ports which feed into global control loops. ASC and radiation pressure signals are summed with damping loop controls for each mirror. The LSC loop is self-contained and does not affect the dynamics of the system.

the system. ABCD state space matrices are defined as follows:

$$\frac{d\vec{x}}{dt} = \boldsymbol{A}\vec{x} + \boldsymbol{B}\vec{u}$$
$$\vec{y} = \boldsymbol{C}\vec{x} + \boldsymbol{D}\vec{u}$$

The resulting ABCD matrices we obtain are Tustin discretized using the python-control library to convert these matrices into discrete time matrices with a sampling rate of 16384 Hz. We do this to model the discrete nature of sensing and actuation in closing these control loops.

We additionally determine the OSEM2DOF and DOF2OSEM matrices which allow us to convert between the degree of freedom basis of the mirror motion and corresponding OSEM sensor signals. These matrices are physical basis transformations. It is of note that these matrices can be collapsed into the state space matrices by redefining x and u, however we choose to keep this additional modeling layer for simulation analysis purposes. To obtain these matrices, we use the matrix U determined in the T000134 document and reorder components as well as change units such that the matrix converts physical distances in meters to OSEM signals in counts. U is defined as:

$$U = \frac{1}{4}\begin{pmatrix} \frac{1}{R_m} & -\frac{1}{R_m} & \frac{1}{R_m} & -\frac{1}{R_m} \\ 1 & 1 & 1 & 1 \\ \frac{1}{R_m} & \frac{1}{R_m} & -\frac{1}{R_m} & -\frac{1}{R_m} \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

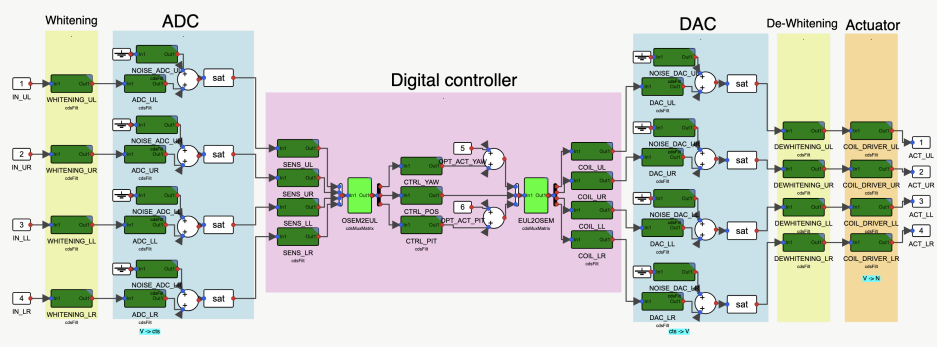


Figure 7: A screenshot of a subsystem of the Simulink model which includes the whitening to actuation section of the signal path for the damping loops on one of the IMC mirrors. Analog and digital filters are modeled using `cdsFilt` blocks and limits are modeled using saturation blocks. There are excitation points (also using `cdsFilt` blocks with a gain of 1) at the output of the ADC and DAC for modeling of electronics noise.

Where R_m is the effective lever arm of the magnets to the angular degrees of freedom. The last row of this matrix corresponds to a fourth degree of freedom, 'bend', which under a rigid-body assumption is always zero. This degree of freedom is added to square the matrix and make inversion easier; it is dropped after inversion. We define the OSEM2DOF matrix as UG, where G is the diagonal OSEM sensor gain matrix. The units of G are in V/m and V/rad, representing the conversion of mechanical degrees of freedom to OSEM voltages. The DOF2OSEM matrix is $U^{-1}A$, where A is the actuation gain matrix. This matrix is in units of N/V and $N \cdot m/V$ and contains the physics of the coil driver, coil, and magnets to convert voltage actuation signals into forces and torques on the mirrors.

The model also contains OSEM2EUL and EUL2OSEM matrices which are controller specified matrices used to convert ADC signals from the sensors into reconstructed mirror degree of freedom signals and back into DAC signals for OSEM coils. These matrix values were obtained through parsing a `.snap` file backup of EPICS [22] channel data.

Signal conditioning and loop shaping are processed within the real-time digital controller pipeline. This can be modeled through implementation of second-order section (SOS) filters used in the digital controller. These filters were extracted from the digital controller and stored in a `FOTON` readable file, a tool used to design and easily implement filters into control loops through specification of an array of bi-quad filter coefficients. EPICS channel data from the Control and Data System (CDS) is used to determine which filters are currently active in the system. Figure 7 shows a plot of the signal geometry for the damping loop control filters. These digital filters also contain freedom to include offsets and gains. We use the custom `cdsFilter` Simulink blocks which include these SOS filter modules as well as a gain and offset specification. EPICS channel data is similarly used to grab these values from the system and insert them into the model.

3.2 Global Feedback Modeling

The model of the global feedback loops branches from the local feedback loop models. Inputs to the global feedback loops branch from the outputs of the state space block and control

signals are summed with damping loop signals. This mirrors the geometry of these loops in the real system: they measure global degrees of freedom from the same plant but with different sensors, and their control signals are summed with local damping in the digital controller.

The global feedback subsystem consists of three main components: the LSC loops, the ASC loops, and radiation pressure. To implement radiation pressure, we take the outputs of the two angular degrees of freedom for all three mirrors and apply a matrix operation given by the coupled differential equations derived above. Because this is a purely physical effect, the output of these matrices (the differential change in the corresponding degrees of freedom) is routed directly to the input of the state space block and summed with digital signals. We note that this does not equate to a modified coupled state space model due to the additional time delay introduced along this path. This was done out of simplicity and in the interest of time to avoid rerouting the original decoupled three mirror Simulink model. Despite this implementation not being physically equivalent, we do not anticipate this to make a significant difference to the corresponding dynamics of the system. This is because the unit delay for our 16kHz system is 10^4 times smaller than the resonance frequencies of the system ($\sim$1Hz). Moving forward however the more robust method to implementing these effects would be to create a new state space block which couples all three of these mirrors together.

The LSC loop is particularly different compared to the ASC loops in that at our first order approximation it has no influence on the dynamics of the system. This is because the LSC loop only directly controls the frequency of the input laser to fix the optical path length of the cavity. Thus, the longitudinal positions of the mirrors are not directly actuated on by this loop. Because of this, although modeling of the LSC is important for completeness, it has been taken with less priority as its implementation is not important to accurately model the rest of the dynamics of the IMC. Physically the LSC loop is closed with an error signal from the mixing of a modulation signal and the REFL PD. Currently the `rtsfreerun` model does not include the electronics that generate this signal, instead the PDH equations are used to model the signal from the length de-tuning of the model plant. This loop topology is shown in Figure 8.

The longitudinal errors of all three mirrors are used to calculate a corresponding cavity length error. It is important to note that the coupling of the longitudinal degrees of freedom for each mirror to the length error is not symmetric across all three mirrors. This is because the angle between the normal of the mirror and the incident beam is dependent on the mirror. This means that a longitudinal error on MC2 has a coupling factor of ~ 2 to the length error, while this coupling factor is only $\sim \sqrt{2}$ for MC1 and MC3. The PDH equations are used to compute a corresponding PDH error signal slope, which is applied in our model as a simplified linear gain block with a saturation representing the limits of the linear PDH signal regime. This error signal then passes through the designed digital control filter and an additional filter block is used for the conversion from digital counts to a correction to round-trip cavity length in meters.

Measurements taken on the 40m indicate that the real analog LSC loop has a unity-gain frequency above 8.192 kHz. Because rtsfreerun operates at 16.384 kHz, with a Nyquist frequency of 8.192 kHz, the full analog LSC dynamics cannot be accurately modeled in the present simulation.

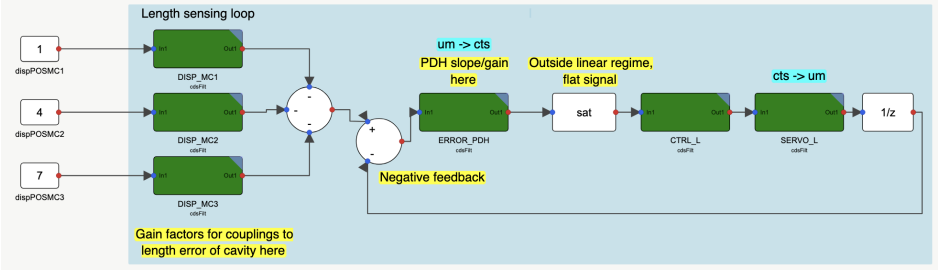


Figure 8: A screenshot of the LSC loop in the Simulink model. Displacements from each of the mirror are summed with different coupling factors to compute the round trip length error signal. This signal passes through a filter with a gain corresponding to the slope of the linear regime of the PDH error signal. A saturation block is used to specify the limits of this linearity assumption. Finally, the signal passes through a digital control filter, the output signal of which is fed back into the length error signal.

Finally, the ASC loops are based on the error signals of the Wavefront Sensors (WFSs) and the MC2 transmission sensor rather than OSEMs. WFSs are quadrant photodiodes sensing beam misalignment (tilt/translation) via RF heterodyne detection. We can model these sensors using the corresponding sensor matrices. These matrices were determined in the Tanioka analysis and were further restructured for the physical modeling purposes of this project in [12]. The two WFSs, one of which measures the near field limit of the REFL beam of the IMC and one which measures the far field limit, are able to sense two global degrees of freedom for both pitch and yaw motion of the IMC mirrors. This leaves us with one angular degree of freedom that cannot be sensed with these sensors. To fully resolve all alignment degrees of freedom, we utilize another sensor, the MC2 TRANS QPD. This gives us an additional angular degree of freedom for both pitch and yaw motions, giving us an invertible sensing matrix which is necessary for effective control. The corresponding error signals from all three sensors are fed into the digital controller subsystem which is a one-to-one model of the digital filters, as well as input and output matrices in the real-time system. The `rtsfreerun` model for the ASC loops is shown in Figure 9.

3.3 Validation

To evaluate the developed rtsfreerun model against the real counterpart, we conducted a series of simulated validation tests. One advantage of our approach to modeling is that due to the model's use of the same real time digital processing software as the real system, simulated excitations and measurements very closely replicate real measurements due to the one-to-one modeling of digital channels. This served as a particular benefit to the speed and simplicity of conducting simulated experiments. Our validation tests can be split into two categories: consistency checks and comparative tests. Consistency checks involved comparison of the dynamics of the rtsfreerun model to the previously developed simple python-control model. Comparative checks involved comparing a real and simulated measurement on the IMC.

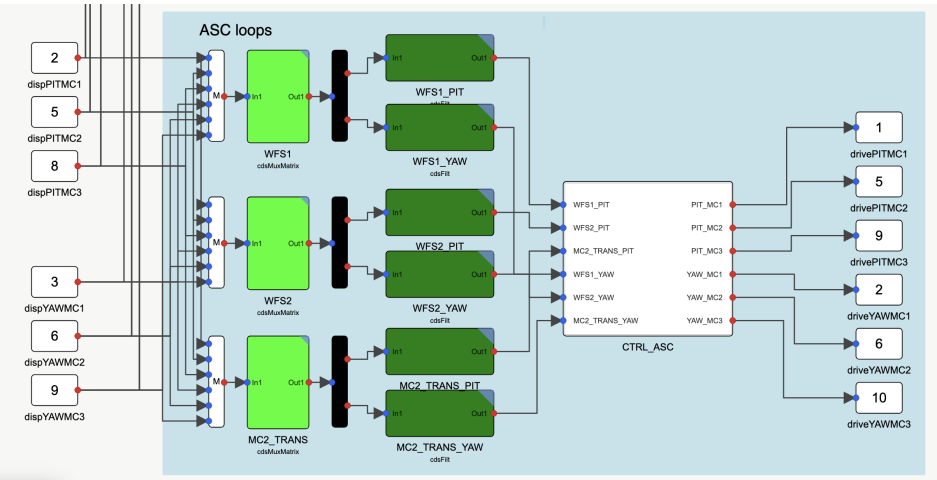


Figure 9: A screenshot of the ASC loops feedback in the Simulink model. Sensing matrices are used to convert angular DOFs to sensor signals. These signals are processed by a subsystem block which contains the relevant digital control filters.

3.3.1 Step Responses

We first compare our rtsfreerun model against the real system in the time domain. One way we can do this is through the step response, which is the time-domain response of the system to a step actuation to one of its inputs. The step response provides a simple way to characterize the system's time-domain dynamics, while being much easier to implement experimentally than an ideal impulse. For this comparison, we excited the system using the CDS and measured the corresponding responses of the mechanical degrees of freedom (position, pitch, yaw). The simplest approach to this measurement is to modify the offset value of the particular channel that we would like to excite in the digital controller. We chose to send in these excitations at the `{DOF}CAL` channels, since they serve as a digital reconstruction of the forces that will be applied to the mirrors in the mechanical DOF basis. We then in parallel measure the response at the `IN{DOF}` channels which are similarly a digital reconstruction of the errors in the DOF basis. We then replicate this same measurement in the rtsfreerun model. Figure 10 shows a comparison of these two measurements. We see a general agreement in ring-down behavior, with some discrepancies in the Q-factor and resonant frequency. The mismatch for the closed-loop Q-factor in particular implies a mismatch in the open-loop gain between the real system and the model. These mismatches are the main motivation for system identification, which will be discussed in the next section.

3.3.2 OLTF Measurements

In addition to time-domain comparisons, we also compare the frequency-domain behavior of our rtsfreerun model. One method of frequency-domain validation is through measurements of open-loop transfer functions (OLTFs). For a linear time invariant system, we can abstractly represent the feedback loop through a simple block diagram (see Figure 11) where block G is the forward path transfer function and block H is the feedback transfer function. The OLTF of a feedback loop is the transfer function of the corresponding feedback path, or $G(s)H(s)$. The OLTF is important because it is used to understand the stability of the loop

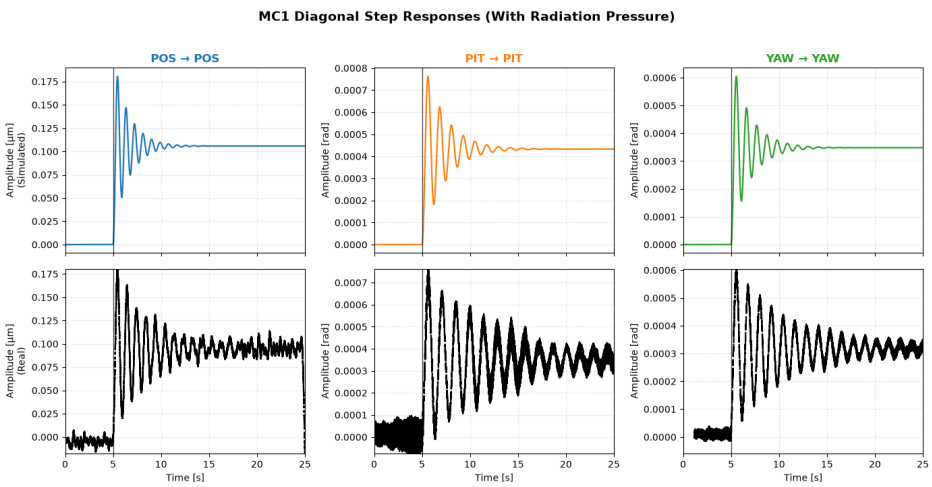


Figure 10: Step response comparison plots. The top row shows the time-domain response predicted by the `rtsfreerun` model, and the bottom row shows the real responses measured at the 40m. Columns are ordered position, pitch, and yaw. Measurements for MC1.

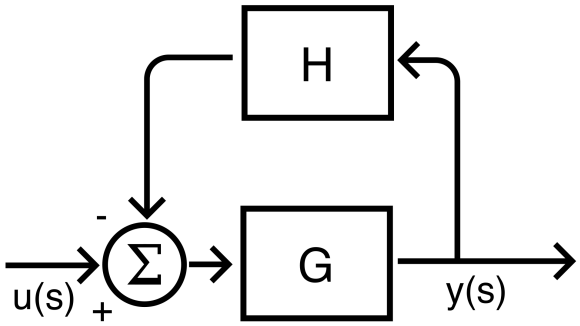


Figure 11: A simple diagram of a system with feedback. The block G represents the forward path and the block H represents the feedback path, summed with a flipped sign representing negative feedback.

and is related to the closed-loop transfer function, which describes how inputs are shaped when feedback is active.

As part of this validation we compare two OLTF measurements: the POS damping loop for MC1 and the WFS1 PIT ASC loop. The first of these measurements was completed by Arush prior to my arrival on site, while I worked directly on the second measurement. Thus, we will detail the process for this measurement further in section 4. Figure 12 and figure 13 show the comparison between model and measurement for the damping loop and ASC loop respectively. We see for the damping loop a strong agreement in both the magnitude and phase between our model and measurement in the high coherence band of the measurement. This is similarly the case for the WFS1 PIT loop, although we observe additional phase accumulation in the 1-3Hz band. This is not unexpected, as currently the model used as a reference in the rtsfreerun model has not incorporated all of the pre-processing filters that are used for WFS signals in the real system. A similar discrepancy I previously observed between the model and the measurement helped with identification of unmodeled low pass

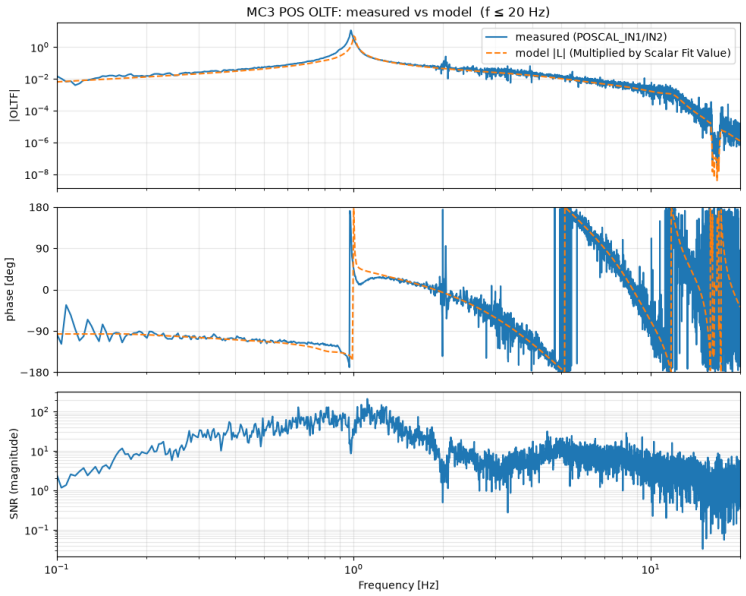


Figure 12: OLTF measurement of the MC3 position loop. The top two rows are a Bode plot comparison with orange curve being a magnitude scaled TF from the `rtsfreerun` model and the blue curve is the measured data. The bottom row is the log-scaled SNR for the measurement.

filters. We therefore suspect that this discrepancy is the result of a similarly unmodeled low-pass filter somewhere within the digital processing chain.

3.3.3 Radiation Pressure Measurements

The final validation test we conducted was to resolve the resonance splitting of the angular degrees of freedom of individual mirrors caused by radiation pressure. Resolving this splitting provides a direct validation of the radiation pressure model and its predicted mode structure. To do this, we took a fine measurement of the diagonal pitch to pitch transfer function for the MC1 plant. We focus on the MC1 pitch degree of freedom as the model predicts the largest and most resolvable splitting for this mirror and degree of freedom. The prediction from the modeling of radiation pressure suggests that this splitting should be resolvable with at least a 28mHz measurement resolution.

To isolate the plant dynamics from the surrounding control loops, we measure the transfer function from the total pitch actuation to the sensed pitch response. We send in an excitation u at the `PITCAL` channel and reference the sensed response y at the `SUSPIT` input to the total drive x read at `PITCAL_IN2`, which lies after where the excitation is summed with damping and ASC control signals. By computing individual cross spectral densities S_{yu} and S_{xu} and taking their ratio, we isolate the measurement of the plant transfer function P while removing all controller contributions and potential noise biases. For further justification on how this measurement isolates P, see Appendix A.

We expect the modeled nominal resonance frequency (∼0.773 Hz) to be offset from the true location, so we first performed a coarse white-noise excitation over 0.5–1 Hz to locate the res-

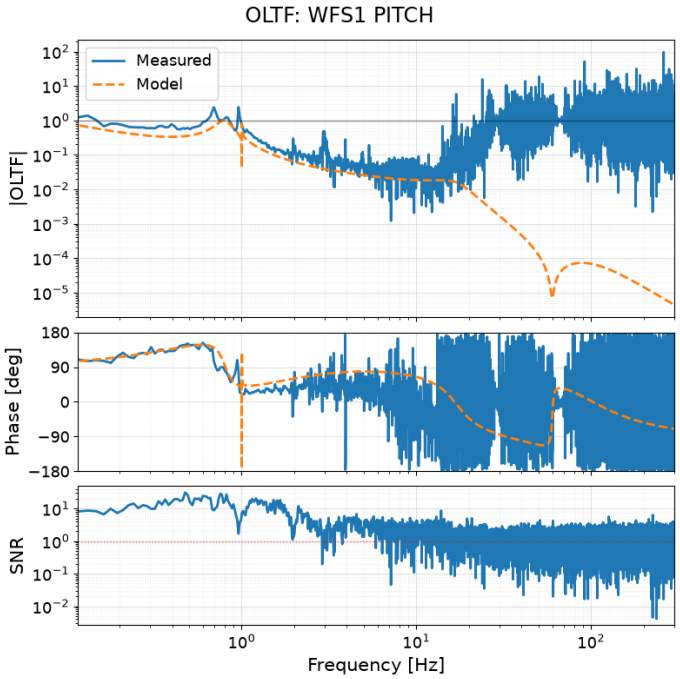


Figure 13: OLTF measurement of the WFS1 pitch loop. The top two rows are a Bode plot comparison with orange curve being a magnitude scaled TF from the `rtsfreerun` model and the blue curve is the measured data. The bottom row is the log-scaled SNR for the measurement.

onance, which we found at approximately 0.680 Hz. Rescaling our radiation pressure model with this resonance frequency, we expect spacings of 4.6, 19.5, 24.1 mHz for each respective pole pair. Following this a fine multisine excitation was used to resolve the resonance region using the excitation framework described in Section 4.2.2: five averaging segments of 750 s each, with excitation lines placed every 2 mHz across the band 0.656–0.720 Hz.

The frequency resolution, or spacing of the excitation lines, is set by the inverse of the segment duration; for our 750 s segments this is approximately 1.3 mHz, finer than the 2 mHz line spacing. The fundamental limit on distinguishing two nearby peaks, however, is set by the resonance linewidth $\Delta f \approx f/Q$. From the measurement a Q fit to the data could not constrain its value, meaning our 2 mHz resolution is too large to resolve the resonance. Thus we estimate $Q \gtrsim 340$ and a linewidth smaller than our frequency resolution. The model therefore predicts a three-peak measurable transfer function.

Figure 14 shows the measured MC1 pitch plant transfer function. We observe a single sharp resonance at approximately 0.680 Hz with high coherence (> 0.9) across the band. We see no resolved secondary peak, even at the $\sim$24 mHz scale at which poles are predicted to be shifted by. This indicates that one or more assumptions in the model is incorrect and the frequency shifts are smaller than expected. We identify two likely contributions. First, the predicted splitting scales linearly with the circulating power P. A circulating power lower than the value assumed in the model would compress all predicted shifts. Second, the approximately 13.7% discrepancy between the modeled nominal resonance (0.773 Hz) and the measured value (0.680 Hz) indicates that the pendulum and optical parameters of

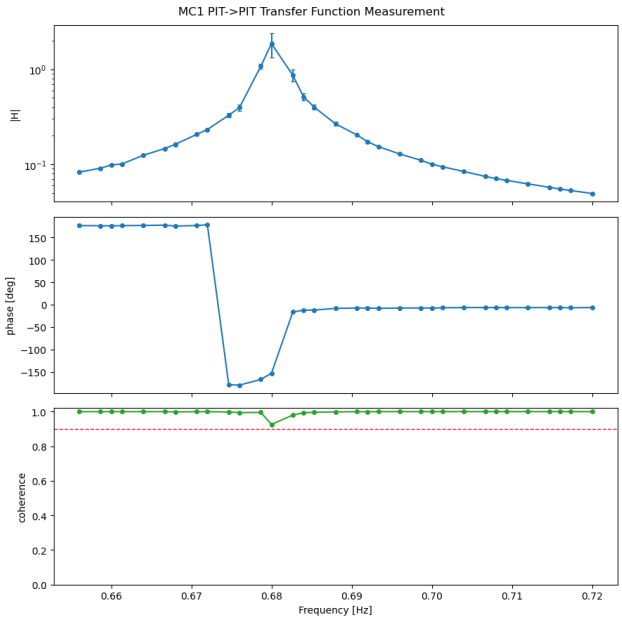


Figure 14: Plant TF measurement for the pitch degree of freedom on MC1. The top two panels are the measured Bode plot, and the bottom panel shows a consistent high coherence for the measurement.

the model are consequentially inaccurate. Since the predicted shifts depend on these same parameters, the true splitting may be smaller than predicted.

We will consider a possible discrepancy in circulating power during the measurement. Scaling the model by circulating power we find that the minimum power required to resolve a 2 mHz gap between peaks is $\sim$ 51 W, an order of magnitude lower than expected. The second of these possibilities is also considered. To do this, we run a Monte Carlo simulation with $N = 200,000$ to estimate the statistical likelihood of the inability to resolve peaks being a result of varied pendulum resonance frequencies or other system parameters. We assume a relative resonance frequency error of 13.4% corresponding to the measured discrepancy for MC1, a 10% error for geometric parameters (which in the case of lengths corresponds to at smallest the centimeter scale), and a prior of 10% error for circulating power. The results of this simulation show that not one case resulted in a minimum gap between resonances at less than 2 mHz, effectively ruling out this possibility.

4 System Identification

As the validation tests have shown, even though our existing model captures much of the characteristic behaviors of the real system, there still exist discrepancies between modeled and true system parameters. Thus, for an accurate digital twin, we must incorporate a method of model fitting. This is called system identification. System identification is the process of fitting a parameterized model of a system to measurements of its real counterpart. With the initial rtsfreerun model established, we moved to developing the process for taking measurements of the system for this fitting.

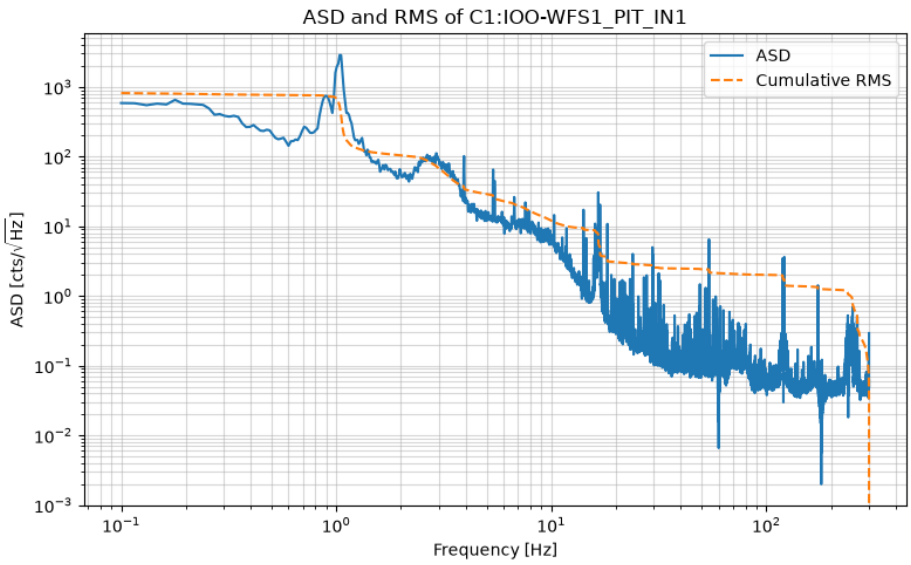


Figure 15: ASD measurement of the WFS1 PIT control channel with no excitation. The orange curve is the integrated RMS curve, a useful metric for determining the frequency dependent amplitude of the designed excitation.

SysID is most commonly done manually, meaning excitations and measurements are conducted by the decision of a specific person or group of people. The best way to measure systems ideally is to maximize the drive power to maximize the signal-to-noise ratio (SNR) of the measurement. For LIGO this presents two issues: particularly strong drives can push the system outside of its linear regime which our model assumes, and strong excitations have the risk of damaging the hardware in the system. Limited excitation strength can be compensated by increased measurement time, although this presents its own issues. Thus sysID measurements are typically slow iterative processes that strike a balance between SNR, measurement duration, and the linearity/safety limits of the system.

4.1 Manual sysID

The WFS1 PIT OLTF measurement mentioned in the validation section above is particularly illustrative of this sysID process. The first step involved a passive measurement of the channel we would like to excite, in this case being `C1:IOO-WFS1_PIT_IN1` as well as additional auxiliary channels for later comparative analysis. The ASD of this measurement is shown in Figure 15.

The RMS and spectral content of this passive measurement is used to shape the excitation, since an effective drive must sit just above the noise floor across the frequency band of interest (high enough to overcome noise, but no higher than necessary to avoid unnecessarily disturbing the system). We designed a shaping filter (an elliptic roll-off, a zpk filter, and two resgain stages) to match the measured RMS, and set the excitation amplitude accordingly.

The measurement then proceeded as an iterative sequence of excitations, with the amplitude raised in stages while monitoring the response. The first excitation established the correct

overall scale, and the second yielded a well-matched excitation shape but with a signal-to-noise ratio of only ∼1, motivating a further increase in amplitude. At the same time we steepened the elliptic low-pass roll-off to further suppress high-frequency drive, since additional low-pass filtering elsewhere in the loop already attenuates the system's response there, limiting the potential information gain. By the third excitation, the aligned cavity beam began to pulse and wobble visibly, indicating that we were approaching the limit of safe drive strength and should be careful not to push the excitation much further. The progression of shaping the drive to the noise floor, raising it in controlled steps, and stopping at the first physical sign of disturbance is precisely the slow, manual balancing of SNR against linearity and safety described above.

Throughout the measurement process we recorded not only the target WFS1 PIT channel but also the pitch and yaw channels of all three mirrors, the remaining WFS and MC2 TRANS channels, and the MC length channel. The mirror channels give a rough measure of how strongly the excitation drives the suspensions, the auxiliary WFS and TRANS channels probe coupling into the orthogonal ASC loops, and the length channel tests our earlier conclusion that the LSC loop is not important for modeling the system's angular dynamics. Figure 13 shows the resulting open-loop transfer function.

The measurement retains good coherence up to ∼1 Hz and loses it rapidly above. This is consistent with the loop rolling off: at high frequency the sensitivity approaches unity, indicating that the open-loop gain is approaching zero, so little additional information is available by probing further up in frequency. Our model likewise predicts that the ASC open-loop gain falls off rapidly above 10 Hz, so this band was not pursued. The model does, however, predict that the WFS loop gain rises above unity below ∼0.1 Hz (corresponding to a pole at DC), a region we have not yet measured. Extending the measurement to lower frequencies is therefore a natural next step for a higher-fidelity comparison against the ASC model.

This WFS1 PIT measurement illustrates both the value and the limitations of the manual approach: it produced a usable and coherent OLTF measurement, but required continuous human supervision and a conservative, iterative approach to drive strength that leaves room for improved measurement fidelity. This motivates an optimal input design approach to system identification, in which the excitation is shaped to extract the maximum information about the system parameters subject to fixed amplitude and duration constraints, rather than being tuned by hand. We call this method the Optimal Multisine method.

4.2 Optimal Multisine sysID

The slow nature of the manual sysID procedure, particularly for systems like the IMC defined by dozens of parameters, is motivation for the development of an algorithmic approach to sysID. The development of this algorithm logic was not within the scope of this project, however we will give an overview of the current algorithm such that the mathematical approach and application to the real system can be better appreciated and connected to completed work and existing problems.

4.2.1 The Pintelon-Schoukens Algorithm

The optimal multisine method implemented in this project is heavily based on the optimal excitation design algorithm described by Pintelon and Schoukens [23]. This algorithm is based on the Fisher information formalism which we will introduce here.

The Pintelon-Schoukens Algorithm derives itself from the simple goal: what is the optimal excitation $\chi(\Omega_k)$ (where Ω_k is the k-th frequency of a discrete excitation) to be injected into a system to minimize the uncertainty of a parameterized model estimate of the system. For continuous frequency spaces, this problem is particularly computationally expensive, and so a discrete frequency space (hence a multisine excitation) is used to find the mathematically optimal excitation. We choose to define this uncertainty as the determinant of the covariance matrix, which for a n_θ parameter system is an $n_\theta \times n_\theta$ matrix. The determinant of this matrix can be interpreted as the uncertainty volume in the system's parameter space. Thus, minimizing the determinant of the covariance matrix can be interpreted as constraining the volume of possible systems in the parameter space. To compute the covariance matrix, we can use the Cramér-Rao inequality to determine a lower bound. Assuming a sufficiently long measurement time, the lower bound CR is:

$$CR(\theta_0) = F_I^{-1}(\theta_0)$$

Where $F_I(\theta_0)$ is the Fisher information matrix. For a single-input single-output (SISO) system, we can approximate the Fisher information matrix as such [24]:

$$(F_I)_{ab} \simeq \sum_k \frac{P_{xx}}{P_{yy}} \cdot 2\mathrm{Re}\left[\left(\frac{\partial G(\Omega_k,\theta)}{\partial \theta_a}\right)^\dagger \left(\frac{\partial G(\Omega_k,\theta)}{\partial \theta_b}\right)\right]$$

Where $\frac{\partial G(\Omega_k,\theta)}{\partial \theta_a}$ is the component J_a of the Jacobian for G evaluated at Ω_k and $\frac{P_{xx}}{P_{yy}}$ is the weight for the corresponding frequency, or the ratio between the excitation power and the noise power at the frequency Ω_k. This expression relies on the assumptions of Gaussian noise, negligible input noise, and the high SNR/unbiased regime where $CR \approx F_I^{-1}$. Thus under these assumptions minimizing the determinant of the covariance matrix is equivalent to maximizing the determinant of the Fisher information matrix. In order to optimize our designed excitation $\chi(\Omega)$, we use the dispersion function $\nu(\chi, \Omega_k)$, which tells us how much the determinant of the Fisher information matrix will change for an increase in the excitation power at a particular frequency. This gives us a recursive formula for optimizing our excitation to maximize Fisher information:

$$\chi_{i+1}(\Omega_k) \propto \nu(\chi_i, \Omega_k)\chi_i(\Omega_k)$$

Where ν is directly related to F_I. More precisely, $\nu(\chi, \Omega_k) = \mathrm{tr}[F_I(\chi)^{-1} f_I(\Omega_k)]$ where $F_I(\chi)$ is the Fisher information matrix for the current excitation design and $f_I(\Omega_k)$ is the information contributed by a normalized excitation at frequency Ω_k, so that ν measures the relative value of allocating excitation power at that frequency. Subtly the above equation for F_I (and correspondingly ν) relies on the knowledge of $G(\Omega, \theta)$, which would of course be unknown in the context of the measurement. This underscores the importance of the prior and estimator. The prior model gives us an initial guess θ that allows us to guess the most optimal excitation, and following iterations update this guess using the estimator based on system measurements.

This is the mathematical basis for the algorithmic approach to system identification. With an iterative process of designing an excitation, estimating the model parameters θ, and using the dispersion function to design the next excitation, we have a fully algorithmic approach for system parameter estimation that is additionally mathematically optimal (under the specified assumptions, which we note below may not always hold).

4.2.2 Implemented Optimal Multisine Algorithm

The current algorithm code is quite dense and may be difficult to parse upon first pass, however its fundamental logic is relatively straightforward and easy to follow. Understanding the logic behind the algorithm allows clear understanding of how the above framework is incorporated into and how it interacts with the system. Below in Algorithm 1 we create a pseudo-code representation of the algorithm split into defined variables, a design function, and the high level iterative structure.

The algorithm can be understood as a simple conditional loop which references a method we call `DESIGN`. `DESIGN` requires two explicit arguments: the model prior or estimate θ and model uncertainty u. The core of the method runs the Pintelon–Schoukens algorithm for a fixed number of iterations n_i starting from an initial flat excitation spectrum. The final excitation is returned by the method. This method is iteratively called to give an updated excitation based on an updated model fit.

The Python implementation of this algorithm was developed in parallel to and independently of this work. Our contribution was to interface this implementation with the 40m real-time control system. We connected the design/measurement logic to the CDS excitation and readback channels and validated, to the extent possible within this project, that the pipeline takes correct measurements and correctly injects into the real system.

4.2.3 CDS Backend Testing

With the optimal algorithm implemented in Python, the next important step is to demonstrate its application on the real system. This must be taken with particular care, as any bug in the implementation that produces an unintended excitation risks damaging the hardware. For this reason, testing proceeded in stages of increasing code authority over the system: a fully simulated measurement, then injection into disconnected filters, and finally a live measurement on a real control loop. At each stage we validated correct behavior before granting the code more direct access to the physical system.

The first stage was a fully offline test, in which we ran a simulated injection and measurement for a simple two-pole transfer function. This exercised the excitation design, injection, and readback logic as well as the code's built in live dashboard tool without touching the system. We confirmed the expected ramp-on behavior and a clean multisine drive signal in the readback of the excitation and response channels.

The second stage tested the code's ability to connect to, act on, and cleanly detach from the real system, while isolating any faults from the physical interferometer. We first ran a read-only measurement on a single channel (`C1:SUS-SRM_SUSPIT_IN1_DQ`) to confirm the code could correctly connect to and read data from the system; the returned data was an

Algorithm 1 Optimal-multisine system identification (Pintelon–Schoukens)

Inputs:
θ_0 – prior model
u – prior model uncertainty (fractional)
n_a – number of averaged periods
n_i – inner (design) iterations
n_o – outer (measurement) iterations
δ – target fractional parameter uncertainty
P_{tot} – frequency integrated power budget

procedure DESIGN(θ, u)
  if $u > 0$ **then** ▷ *average over the plausible resonance band*
    form scaled models $\tilde{\theta}(u) = \{\theta$ with resonances scaled by $s : s \in [1-u, 1+u]\}$
    $\chi \leftarrow$ MEAN $\big($DESIGN$(\theta_k, 0)$ for $\theta_k \in \tilde{\theta}(u)\big)$
  else ▷ *dispersion function reweighting*
    initialize flat spectrum χ_0
    for $i = 0$ to $n_i - 1$ **do**
      $\chi_{i+1} \leftarrow \nu(\chi_i, \Omega_k)\,\chi_i$ ▷ *dispersion reweight*
      floor $\chi_{i+1}(\Omega_k)$; renormalize by P_{tot} ▷ *each bin has an excitation power floor*
    $\chi \leftarrow \chi_{n_i}$
  return χ

Initialize:
  passive measurement of readout noise P_{yy}
  $\chi_0 \leftarrow$ DESIGN(θ_0, u) ▷ *first pass is prior-robust*

for $j = 0$ to $n_o - 1$ **do**
  minimize crest-factor and ramp χ_j to waveform
  inject n_a periods; measure response Y, drive X ▷ *multiple periods to average out noise biases*
  $H \leftarrow \mathrm{mean}_a(Y)/\,\mathrm{mean}_a(X)$ ▷ *reference-based, cancels controller*
  accumulate H into a running estimate by inverse-variance weighting per bin
  fit θ_{j+1} to H
  compute Fisher information F_I
  $\sigma_{\theta_a} \leftarrow \sqrt{[F_I^{-1}]_{aa}}$
  if $\max_a \left(\dfrac{\sigma_{\theta_a}}{|\theta_a|}\right) < \delta$ **or** $j + 1 = n_o$ **then**
    stop ▷ *target or max iteration reached*
  else
    $\chi_{j+1} \leftarrow$ DESIGN$(\theta_{j+1}, 0)$

exact match to the same channel read independently through the ndscope tool, validating the readback path. We then performed an injection test on the (at the time) unused SRM POS damping loop, a signal path disconnected from the rest of the system. Injecting into an isolated filter ensures that any unexpected behavior, such as a larger than intended excitation, cannot propagate to the physical system. To emulate a realistic measurement, we implemented an SOS decomposition of our WFS1 PIT OLTF model into the SRM control filter modules, so that the injected excitation would be shaped by a known transfer function, and separately emulated readout noise drawn from the background spectrum of the `C1:IOO-WFS1_PIT_IN1` channel. This let us reproduce the full measurement of the WFS1 PIT loop on an isolated, safe filter path. We verified that the excitation was delivered to the correct channel, that the measurement correctly recovered the implemented filter, and that the code cleanly disconnected from the system after the excitation completed.

With the isolated-injection stage validated, the final step was a live measurement on the real WFS1 PIT loop. Because this grants the code direct authority over a physical loop, we first established well-defined limits on the excitation. The maximum RMS and peak amplitude were based on our earlier manual measurement of the same loop. This yielded a maximum RMS of 3190 counts and a maximum amplitude of 10000 counts.

For additional protection we imposed hard limits at two points. First, we capped the coil outputs of all mirrors at 75% of the DAC range (24576 counts) as a hard stop on actuation. Second, we set limits on the `C1:IOO-MC{NUM}_PIT` channels as a validation check on the excitation limits defined within the code itself, computed from the peak signal expected to reach each channel (the summed background and maximum expected excitation, propagated through the known gain factors between the excitation point and each channel). The initial values were 320, 500, and 830 counts for MC1, MC2, and MC3 respectively.

During the first live measurement we observed saturation in the `C1:IOO-MC{NUM}_PIT` channels, indicating these limits were being exceeded, which in turn produced an inaccurate transfer function through a spuriously high measured coherence at low frequencies. We attempted to abort early in the third iteration; however, the abort did not terminate cleanly (a segmentation fault, revealing an error in the termination handling), leaving a live excitation process that had to be killed manually. For the second measurement we raised the initial limits by a factor of two, which resolved the saturation on the first iteration but not on later iterations. Inspecting the control filter, we found that it enhanced higher frequencies, which explains the recurrence of saturation as later iterations trended toward higher frequencies (a bias discussed below). The final iteration of this measurement with the corresponding coherence is shown in Figure 16, the results of which are insightful to recognize the current failure mode of the algorithm.

Two behaviors observed during these tests are worth noting as open issues. First, the designed excitations tended to bias toward higher frequencies over successive iterations, in both the isolated-filter tests and the live measurement. We suspect this may be related to the high order prior model used to fit to the WFS loop as well as an implicit high SNR regime assumption of the algorithm not holding for the relatively low SNR measurements we have completed. This connects to the second observed behavior: measurements with non-negligible noise could diverge substantially from the prior even when a small prior uncertainty was specified. Resolving these behaviors, along with the parameter-fitting stage of

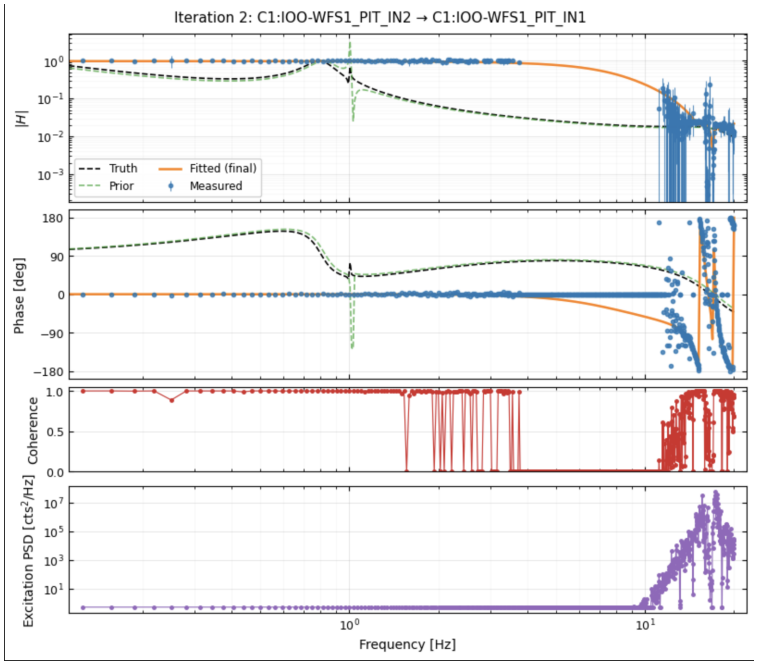


Figure 16: Final iteration of the final optimal sysID measurement taken at the 40m. We see a high coherence at low frequencies which appears to be an unintentional result of excitation signal clipping during the measurement process. The final panel shows how the algorithm tends to higher frequency measurements on later iterations.

the pipeline, remains the subject of ongoing work.

5 Conclusion

In this report we developed a physics-based digital twin of the Caltech 40m IMC and a system identification framework to measure and refine it against the real instrument. We derived the radiation-pressure-induced angular coupling of the three-mirror cavity in the linear regime, ported the full IMC physics of suspension dynamics, electronics, and digital control into an `rtsfreerun` model that shares the real-time software of the physical system. We validated this model against the 40m through three complementary measurements: step responses, open-loop transfer functions, and a fine radiation-pressure measurement of the MC1 pitch plant.

These validation tests establish both the fidelity and the limitations of the current model. The step-response and OLTF comparisons show close agreement in the high-coherence bands, while discrepancies such as a closed-loop Q-factor mismatch indicating an open-loop gain error point to where the model diverges from the real system and motivate a fitting stage. The radiation-pressure measurement result is null: no resonance splitting was resolved despite a measurement resolution and plant linewidth below the predicted mode separation. We showed this measurement to be a genuine property of the plant rather than an artifact of the active control loops, since referencing both the response and the total drive to the injected excitation cancels the controller contributions under the approximation of a SISO system.

The absence of the predicted splitting therefore likely points to inaccuracies in the assumed optomechanical parameters of the model.

To address these discrepancies systematically, we implemented and deployed an optimal-multisine system identification pipeline on the 40m real-time system. The core Pintelon–Schoukens algorithm and its Python implementation predate this work; our contribution was to interface this pipeline with the CDS excitation and readback channels and to validate, through a multi-staged process of increasing code authority, that it takes correct measurements and injects safely into the real instrument.

Several issues remain open. The parameter-fitting stage of the sysID pipeline is implemented but still is under development, so there has not yet been a verifiably accurate measurement for parameter fitting. The observed drift of the optimal excitation toward higher frequencies under low-SNR or a high order prior warrants further investigation before the method is applied broadly. The ultimate objective of training a reinforcement-learning control policy in the completed digital twin and transferring it to the real system remains future work that this framework is intended to enable. Together, these results deliver the modeling and measurement infrastructure on which the high-fidelity training environment depends.

5.1 Next Steps

Moving forward with the initial goals of the project, the next immediate step is an investigation into the current behavior of the implemented algorithm on noisy measurements. The current failure modes will have to be resolved before any trustworthy measurement of the system can be taken. Following the necessary corrections to the algorithm and validation, optimal sysID measurements will be taken to extract all the relevant parameters from the system. These parameters, along with their uncertainties, will be used to fit an ensemble of digital twin models (each with slightly tuned parameters dictated by measurement uncertainties). This ensemble of digital twin models will serve as the training environment for the reinforcement learning algorithm.

Along with advancements in the sysID algorithm, there is additional work to be done on the `rtsfreerun` model design. Currently many pre-processing and control filters for the ASC loops have not been fully implemented. Control filters have been added to the `rtsfreerun` model, however SOS filters and gain values have not been extracted from the system. The pre-processing filter blocks for WFS signals have not yet been modeled. The gain of the LSC loop and controller filter modules have also not been added to the model. Thus the current state of the `rtsfreerun` model is functioning damping filters and radiation pressure, with a few missing pieces for global ASC and LSC loops yet to be implemented.

After development of the digital twin has been completed, the final goal is to train a reinforcement learning control policy. This will be done using the Deep Loop Shaping RL method [9] developed to reduce the injected noise of the common-hard-pitch loop in LIGO. After training, this policy will be deployed on the real system and improvements to control noise will be evaluated. This concludes the project at the 40m, following which this approach can be adapted for various other systems in the LIGO facilities for improvements to technical noise for future upgrades.

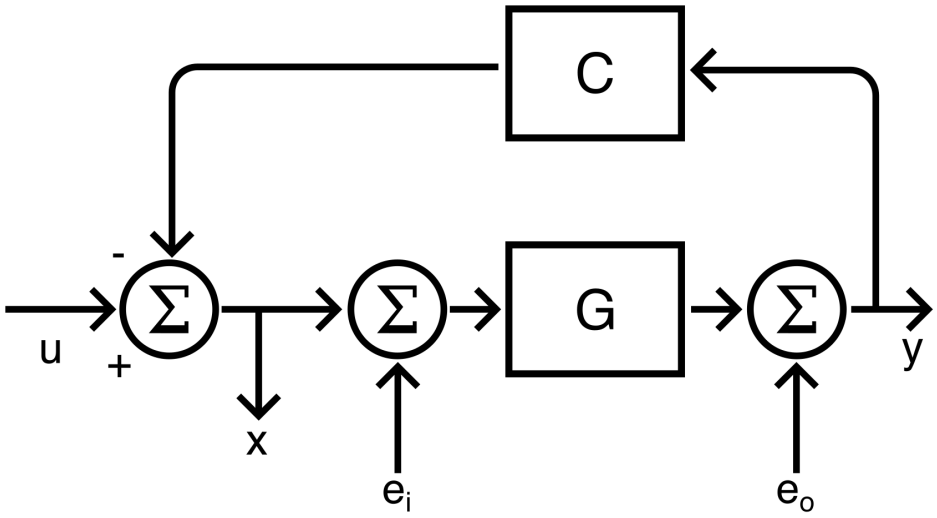


Figure 17: A simplified diagram of a feedback system with plant G and controller C. Inputs to the system are labeled u, the total drive (controller + inputs) is given by x, and the output of the plant is y. In this diagram input noise e_i and output noise e_o are also modeled to illustrate possible biases in specific measurement methods.

6 Acknowledgments

Thank you to Dr. Nathan Holland, Dr. Rana Adhikari, Dr. Christopher Wipf, and Arush Pimpalkar for all of your support, guidance, and insights throughout my work on this project. Thank you to the National Science Foundation, the California Institute of Technology Student-Faculty Programs, and LIGO lab for making my summer research possible. Additionally a big thank you to all of my fellow SURFs for making this summer an unforgettable experience.

A Closed-loop Plant Estimation

This section will justify the validity of the measurement discussed in 3.3.3 in analyzing the resonance splitting caused by radiation pressure in the IMC. Figure 17 illustrates a simplified diagram of the system this measurement is performed on.

Using basic loop algebra from the diagram, we can write two equations that relate u, x, and y:

$$y = G(x + e_i) + e_o$$

$$x = u - Cy$$

These can be rewritten into closed loop equations:

$$x = \frac{u - CGe_i - Ce_o}{1 + CG}$$

$$y = \frac{G(u + e_i) + e_o}{1 + CG}$$

These equations will be useful for the following calculations.

Because we are measuring a transfer function of a closed-loop system, we cannot simply compute a ratio of the output y to the input u because there are other control signals simultaneously passing through the loop. This can be avoided by instead computing the transfer function using the total drive input x, or the sum of the excitation signal u and controller signals. Under the assumption of no noise ($e_i = e_o = 0$), one valid way to measure the transfer function G therefore is through the ratio of cross spectral densities $\frac{S_{yx}}{S_{xx}}$. However this introduces potential measurement biases. To see why, we expand S_{yx}:

$$S_{yx} = \sum_N y(f)x^*(f)$$

$$S_{yx} = \sum_N Gx^*x + \sum_N (Ge_i + e_o)x^*$$

In this expression we can substitute the closed-loop equation for x described earlier:

$$S_{yx} = \sum_N Gx^*x + \sum_N \frac{1}{(1+CG)^*}(Ge_i + e_o)(u^* - C^*G^*e_i^* - C^*e_o^*)$$

The second term factors out into multiple cross spectra terms. We reasonably assume in this simple model that input and output noises are uncorrelated, and over an asymptotically long measurement both noise sources are uncorrelated with the excitation. Thus, the only terms that are left are noise power spectra terms:

$$S_{yx} = GS_{xx} - \frac{|G|^2C^*}{(1+CG)^*}S_{ii} - \frac{C^*}{(1+CG)^*}S_{oo}$$

The second and third power spectra terms correspond to biases in the resulting measurement. To avoid this, we instead compute the ratio of two cross-spectral densities $\frac{S_{yu}}{S_{xu}}$. To see why this eliminates the bias terms, we similarly expand:

$$S_{yu} = \sum_N [G(x + e_i) + e_o]u^*$$

$$S_{yu} = \sum_N Gxu^* + \sum_N Ge_iu^* + \sum_N e_ou^*$$

Where for the same assumptions as above the second and third terms go to zero. Thus, the ratio becomes:

$$\frac{S_{yu}}{S_{xu}} = \frac{\sum_N Gxu^*}{\sum_N xu^*} = G$$

We therefore correctly recover the plant transfer function we had intended to measure.

References

[1] LIGO Scientific Collaboration and Virgo Collaboration. "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs". In: *Phys. Rev. X* 9 (3 2019), p. 031040. DOI: 10.1103/PhysRevX.9.031040. URL: https://link.aps.org/doi/10.1103/PhysRevX.9.031040.

[2] LIGO Scientific Collaboration and Virgo Collaboration. "GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo during the First Half of the Third Observing Run". In: *Phys. Rev. X* 11 (2 2021), p. 021053. DOI: 10.1103/PhysRevX.11.021053. URL: https://link.aps.org/doi/10.1103/PhysRevX.11.021053.

[3] LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration. "GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo during the Second Part of the Third Observing Run". In: *Phys. Rev. X* 13 (4 2023), p. 041039. DOI: 10.1103/PhysRevX.13.041039. URL: https://link.aps.org/doi/10.1103/PhysRevX.13.041039.

[4] LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration. "GWTC-4.0: Updating the Gravitational-wave Transient Catalog with Observations from the First Part of the Fourth LIGO–Virgo–KAGRA Observing Run". In: *Astrophys. J. Lett.* 1004.2 (2026), p. L22. DOI: 10.3847/2041-8213/ae2c74. URL: https://doi.org/10.3847/2041-8213/ae2c74.

[5] B. P. Abbott et al. "Observation of GWs from a Binary BH Merger". In: *Phys. Rev. Lett.* 116 (6 2016), p. 061102. DOI: 10.1103/PhysRevLett.116.061102. URL: https://link.aps.org/doi/10.1103/PhysRevLett.116.061102.

[6] Vitor Cardoso and Paolo Pani. "Tests for the existence of black holes through gravitational wave echoes". In: *Nature Astronomy* 1.9 (Sept. 2017), pp. 586–591. ISSN: 2397-3366. DOI: 10.1038/s41550-017-0225-y. URL: https://doi.org/10.1038/s41550-017-0225-y.

[7] P. Fritschel et al. *LSC Post-O5 Study Group Report*. Tech. rep. LIGO-T2200287. LIGO Scientific Collaboration, 2022. URL: https://dcc.ligo.org/LIGO-T2200287/public.

[8] M. Evans et al. *A Horizon Study for Cosmic Explorer: Science, Observatories, and Community*. 2021. arXiv: 2109.09882 [astro-ph.IM]. URL: https://arxiv.org/abs/2109.09882.

[9] Jonas Buchli et al. "Improving cosmological reach of a gravitational wave observatory using Deep Loop Shaping". In: *Science* 389.6764 (2025), pp. 1012–1015. ISSN: 1095-9203. DOI: 10.1126/science.adw1291. URL: http://dx.doi.org/10.1126/science.adw1291.

[10] Chris L. Mueller et al. "The advanced LIGO input optics". In: *Review of Scientific Instruments* 87.1 (Jan. 2016), p. 014502. DOI: 10.1063/1.4936974. arXiv: 1601.05442 [physics.ins-det].

[11] M. Rakhmanov et al. *Dynamical Properties of LIGO Single Loop Suspended Mirrors*. Tech. rep. LIGO-T000134-00-D. Laser Interferometer Gravitational-Wave Observatory (LIGO), Aug. 2000.

[12] Arush Pimpalkar et al. *SURF Report - Arush Pimpalkar 2026*. LIGO Document T2600199-v1. LIGO Laboratory, May 2026. URL: https://dcc.ligo.org/LIGO-T2600199.

[13] Jenne Driggers. *LIGO SURF Report: Radiation Pressure and Triangular Cavities*. Tech. rep. LIGO-T060240-00. LIGO, 2006. URL: https://dcc.ligo.org/T060240.

[14] Yaron Hefetz, Nergis Mavalvala, and Daniel Sigg. "Principles of calculating alignment signals in complex resonant optical interferometers". In: *Journal of the Optical Society of America B* 14.7 (July 1997), pp. 1597–1606. DOI: 10.1364/JOSAB.14.001597.

[15] Daniel Sigg and Nergis Mavalvala. "Principles of calculating the dynamical response of misaligned complex resonant optical interferometers". In: *Journal of the Optical Society of America A* 17.9 (Sept. 2000), pp. 1642–1649. DOI: 10.1364/JOSAA.17.001642.

[16] Satoshi Tanioka et al. *Angular response of a triangular optical cavity analyzed by a linear approximation method*. 2020. arXiv: 2002.02703 [physics.optics]. URL: https://arxiv.org/abs/2002.02703.

[17] LIGO Scientific Collaboration. *HAM-A Coil Driver*. Tech. rep. LIGO-T1200264. LIGO, 2012. URL: https://dcc.ligo.org/T1200264.

[18] LIGO 40m Collaboration. *LIGO 40m Electronic Log 17860*. https://nodus.ligo.caltech.edu:8081/40m/17860. Caltech 40m Prototype.

[19] Rolf Bork et al. "advligorts: The Advanced LIGO real-time digital control and data acquisition system". In: *SoftwareX* 13 (2021), p. 100619. ISSN: 2352-7110. DOI: https://doi.org/10.1016/j.softx.2020.100619. URL: https://www.sciencedirect.com/science/article/pii/S2352711020303320.

[20] Christopher Wipf. *rtsfreerun*. https://git.ligo.org/christopher.wipf/rtsfreerun. LIGO GitLab repository, accessed 25 September 2026. 2022.

[21] Christopher Wipf. *model-builder*. https://git.ligo.org/christopher.wipf/model-builder. LIGO GitLab repository, accessed 9/25/26. 2026.

[22] EPICS Collaboration. *EPICS – Experimental Physics and Industrial Control System*. https://epics.anl.gov/. EPICS Base 7.0.10, accessed 25 September 2026.

[23] Rik Pintelon and Johan Schoukens. *System Identification: A Frequency Domain Approach*. 2nd. John Wiley & Sons, 2012. DOI: 10.1002/9781118287422. URL: https://onlinelibrary.wiley.com/doi/book/10.1002/9781118287422.

[24] Hang Yu. *Spectral Analysis and Measurements*. LIGO Technical Note LIGO-G2101503-v2. https://dcc.ligo.org/cgi-bin/private/DocDB/ShowDocument?.submit=Identifier&docid=G2101503&version=. Laser Interferometer Gravitational-Wave Observatory (LIGO), July 2021.