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

Technical Note	**LIGO-T2200177–**	2022/10/05

Fisher Information Analysis for Emissivity Estimation

Hiya Gada

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

The Laser Interferometer Gravitational-Wave Observatory (LIGO) measures gravitational waves and is one of the pioneering instruments which help detect black holes and neutron star mergers. The instrument, or more accurately, the suspended mirrors used, is highly susceptible to vibration noise. The third-generation upgrade, LIGO Voyager, is planned to improve the sensitivity by an additional factor of two and halve the low-frequency cutoff to 10 Hz by reducing quantum radiation pressure and shot noise, mirror thermal noise, mirror suspension thermal noise, and Newtonian gravity noise. The thermal vibrations (or thermal noise) are nullified by installing a cryogenic cooling facility (Figure 1) which radiatively cools the silicon test masses to 123 K.

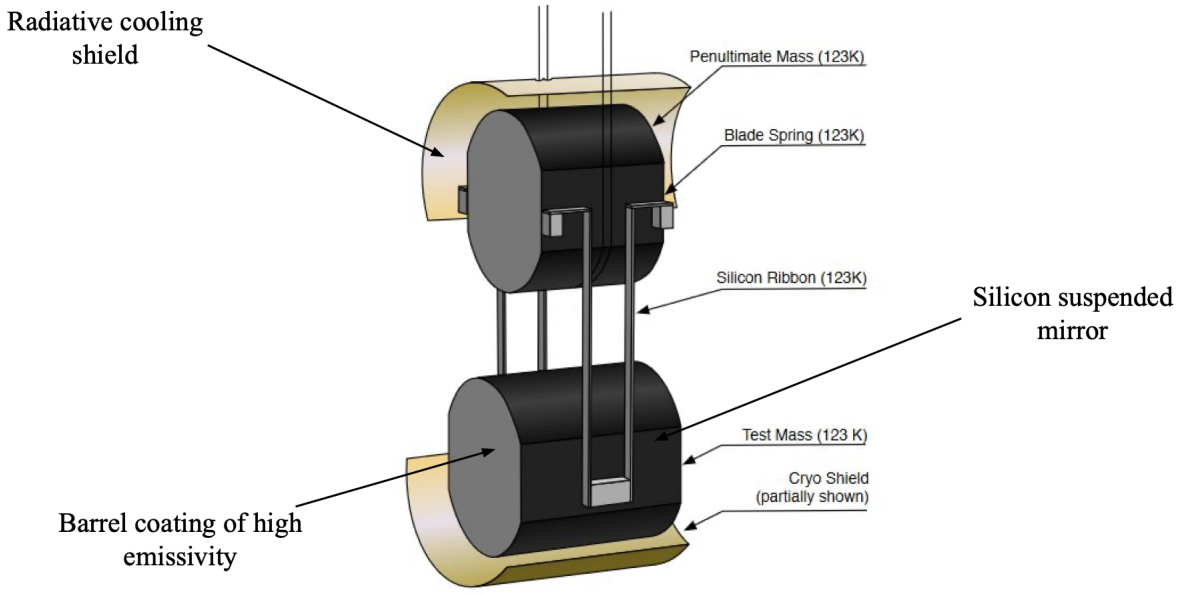

Figure 1: Cryogenic cooling of suspended mirrors in the to-be Voyager upgrade

1.1 Background

Constancio et al. [1] theorized that the silicon test masses would require a high thermal emissivity coating to increase the radiative coupling to its cold environment and effectively dissipate the absorbed laser power. To this end, we wish to determine the emissivities of various black coatings as a function of temperature and subsequently use the best emissivity material for the Mariner (Voyager prototype) upgrade at the Caltech 40m Lab. This is done by obtaining cool-down curves in a cryostat designed specifically for the purpose of emissivity measurement. Using a simplified heat transfer model, the emissivity and corresponding propagated uncertainty is extracted from the cool-down temperature data.

1.2 Motivation

Running the experiment and obtaining the cool-down data is expensive and time-consuming, with a time constant of several days. It becomes infeasible to run the experiments multiple times and find the expected value of the emissivity. We thus run simulations prior to the experiment and find the optimal experimental configuration and excitation, which gives us emissivity with the least uncertainty, using Fisher Information Matrix analysis. The optimal configuration and excitation input obtained will then be used in the experiment to get a close to accurate measurement of emissivity.

The lab's ongoing work includes making design changes to the cryostat, which would minimize heat leaks into the system. It would allow the test mass to cool down to 123 K quickly and reduce the thermal noise injected into the system to get a less uncertain cool-down output. My project would complement this effort by theoretically determining which design parameters contribute the most to the uncertainty in emissivity and even suggest changes in their values for future design upgrades of the cryostat. It would also corroborate the design changes already made and recommend what optimal excitation should be given to make the system robust to noise. Further the same optimal experimental configuration and excitation obtained as a result of this project can be used for emissivity tests of many key coating materials for the LIGO Voyager upgrade.

2 The Cryostat

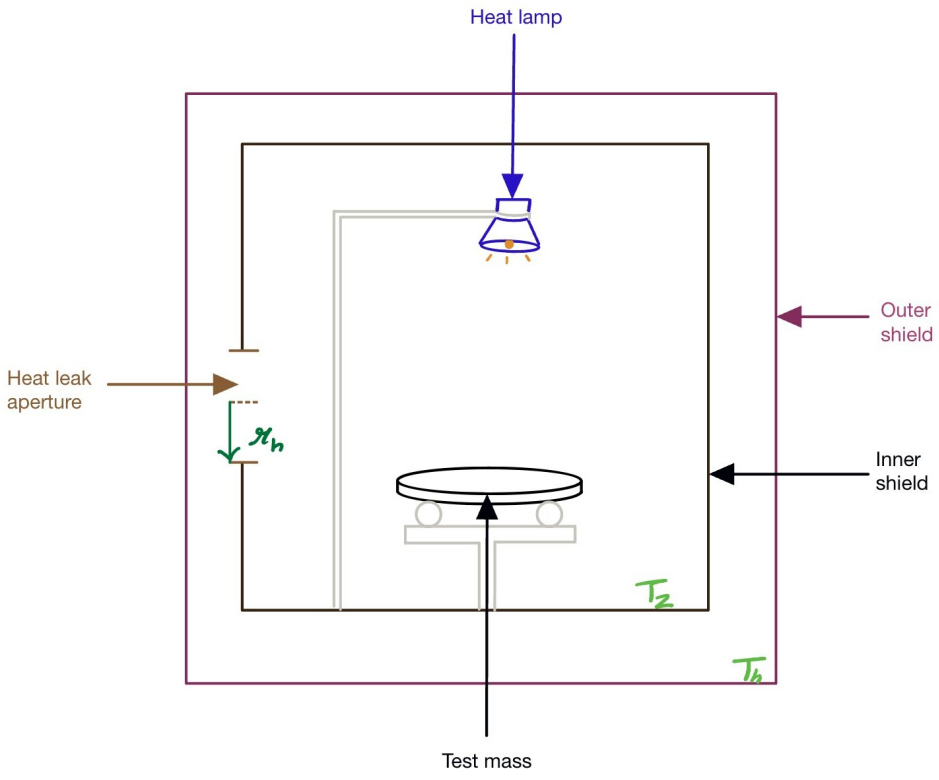

Figure 2: The cryostat

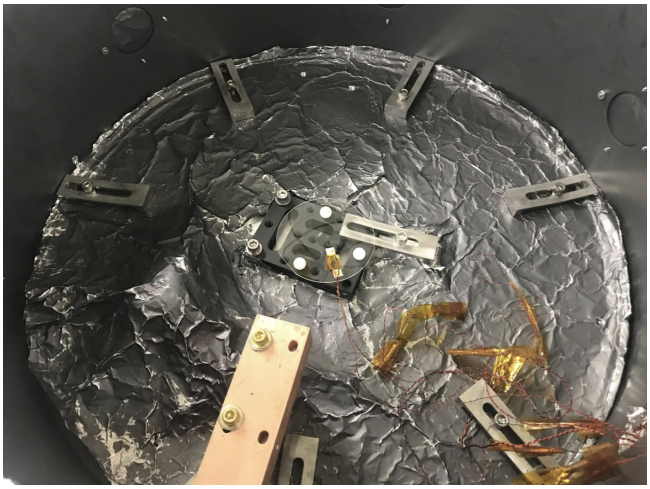

Figure 3: A look into the cryostat experimental set-up

In our latest experiment, we use a thin wafer and not a bulk mass because it is easier to mass produce wafers for numerous experiments at a low cost. Figure 2 shows a simplified system diagram of the cryostat. A mechanical setup (see figure 3) is designed to hold the wafer in place. We also add a heater (a heating lamp) to give the necessary optimal control input from our Fisher Matrix analysis. It is also useful to heat the set-up after an experiment back to room temperature. This helps reduces the experiment run-time. The notation used in this report is as follows,

T_1	Test mass temperature
T_2	Inner shield temperature
T_h	Outer shield temperature
ϵ_1	Test mass emissivity
ϵ_2	Inner shield emissivity
ϵ_h	Outer shield emissivity
A_h	Total area of heat leak from the inner shield to the outer shield
r_h	Effective radius of the heat leak area
$F_{1\to h}$	Geometric view factor from the test mass to the heat leak area
A_1	Test mass surface area
A_2	Inner shield surface area
C_p	Specific heat of test mass
m	Mass of silicon test mass

2.1 Heat Model

Using radiative heat transfer equations as formulated in [2] and keeping geometric view factors in mind, we get the following simplified model of our system,

$$mC_p\frac{dT_1}{dt} = \sigma A_1\left[\frac{T_2^4 - T_1^4}{\frac{1}{\epsilon_1}+\frac{A_1}{A_2}\left(\frac{1}{\epsilon_2}-1\right)} + \frac{T_h^4 - T_1^4}{\frac{1}{\epsilon_1}+\frac{A_1}{A_h}\left(\frac{1}{\epsilon_h}-1\right)+\frac{1}{F_{1\to h}}-1}\right], \tag{1}$$

where we assume that that the geometric factor from the test mass to the inner shield $F_{1\to 2} \approx 1$. The first term corresponds to the cooling term by the cold inner shield, while the second term corresponds to the heating term due to the heat leaks.

For our preliminary analysis we will consider only the steady state region of the cool-down curve. Thus, the ϵ_1, ϵ_2, ϵ_h, T_2 and T_h can be taken as constant. Moreover, it is experimentally observed that T_h realises steady state around 200 K almost every time. T_1 is the output, and $\bar{\theta} = \{\epsilon_1, \epsilon_2, \epsilon_h, r_h\}$ denotes the parameter vector for our system.

Note that the parameters A_h and $F_{1\to h}$ are combined in one parameter that is r_h. It gives us the radius of the effective circular hole causing the heat leak. Knowing r_h one can find out the corresponding A_h and $F_{1\to h}$.

3 Optimal System Identification

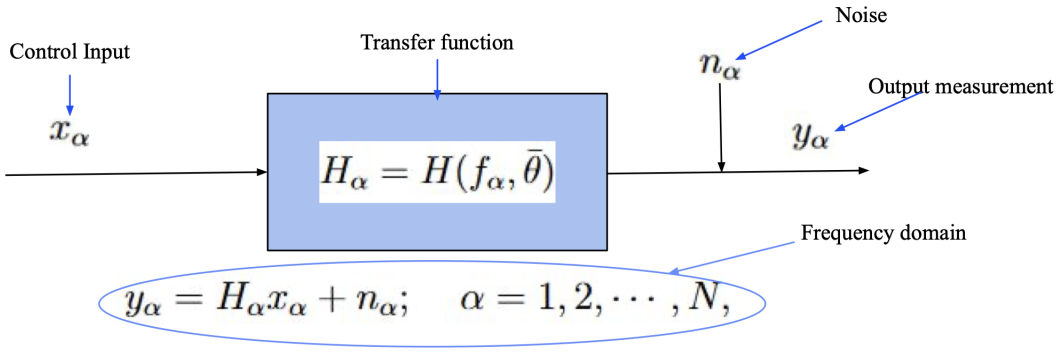

Figure 4: Frequency domain system with noise

3.1 Approach

The main question we want to answer how should our input signal x_α look like to determine the transfer function poles and zeros with least uncertainty and robust to noise n_α? We choose to do a frequency domain analysis as the equations become algebraic for linear systems. Once we obtain the optimal input or excitation we can back-calculate the transfer function H_α of the system and determine the poles and zeros. Thus, we have identify the system parameters.

3.2 Fisher Information Matrix

Assuming that the measurement noise as Gaussian, the Fisher Matrix is constructed from the likelihood function with the objective of maximising the curvature (or minimising uncertainty in output) as derived in [3]. The output measurement,

$$y_\alpha = H_\alpha x_\alpha + n_\alpha; \quad \alpha = 1, 2, \cdots, N,$$

where H_α is the system model, x_α is the input and n_α is the noise corresponding to the α^{th} frequency. The Fisher Matrix is given by,

$$\begin{aligned}
\mathcal{F}_{ij} &= -\left.\frac{\partial^2[\ln \mathcal{L}]}{\partial\theta_i\theta_j}\right|_{\bar{\theta}}, \\
&= \sum_\alpha \frac{1}{|n_\alpha|^2}\,\mathrm{Re}\left[\frac{\partial \hat{y}_\alpha^*}{\partial\theta_i}\frac{\partial \hat{y}_\alpha}{\partial\theta_j}\right]\Bigg|_{\bar{\theta}}, \\
&= \sum_\alpha \frac{1}{\sigma_{H_\alpha}^2}\,\mathrm{Re}\left[\frac{\partial \hat{H}_\alpha^*}{\partial\theta_i}\frac{\partial \hat{H}_\alpha}{\partial\theta_j}\right]\Bigg|_{\bar{\theta}},
\end{aligned}$$

where $\hat{H}_\alpha$ denotes the estimate of the system model, which is dependant on the parameter prior $\bar{\theta}$ we choose. Here $\sigma_{H_\alpha} = \frac{|n_\alpha|}{|x_\alpha|}$, depending both on noise and excitation amplitudes. The Cramer-Rao bound gives a lower limit on the covariance matrix $\mathcal{C}$ and relates it to the Fisher Matrix. For unbiased parameters,

$$\mathcal{C} \geq \mathcal{F}^{-1},$$

where the inequality is understood to be element-wise. Our objective is to select those parameters and input which maximize the Fisher information (and thus minimize variance-covariance). We do this by either maximising the determinant of the Fisher Matrix or minimising the variance of ϵ_1 as that is the parameter of major interest.

4 Optimal Control Input

4.1 Cryostat Transfer function

Consider additional power input by a lamp shining on our test mass and modifying equation (1) we get,

$$mC_p\frac{dT_1}{dt} = \sigma A_1\left[\frac{T_2^4 - T_1^4}{\frac{1}{\epsilon_1} + \frac{A_1}{A_2}\left(\frac{1}{\epsilon_2} - 1\right)} + \frac{T_h^4 - T_1^4}{\frac{1}{\epsilon_1} + \frac{A_1}{A_h}\left(\frac{1}{\epsilon_h} - 1\right) + \frac{1}{F_{1\to h}} - 1}\right] + P(t), \tag{2}$$

where $P(t)$ is the heat power input as a function of time. We will linearize this system about the steady state equilibrium and make a bode plot (figure 5).

The transfer function of the linearized system is calculated as the Laplace transform of the ratio of the output $T_1(t)$ to the heat input $P(t)$,

$$H_{T_1}(f_\alpha) = \mathcal{L}\left(\frac{T_1(t)}{P(t)}\right) \tag{3}$$

$$= \frac{1}{C_0 f_\alpha + 4(C_1 + C_2)T_1^{\circ 3}} \tag{4}$$

where, f_α denotes the frequency variable, T_1° is the equilibrium temperature of the test mass and the constants C_0, C_1 and C_2 are given below,

$$C_0 = \frac{\sigma A_1}{m C_p}$$

$$C_1 = \left(\frac{1}{\epsilon_1} + \frac{A_1}{A_2}\left(\frac{1}{\epsilon_2} - 1\right)\right)^{-1}$$

$$C_2 = \left(\frac{1}{\epsilon_1} + \frac{A_1}{A_2}\left(\frac{1}{\epsilon_h} - 1\right)\frac{1}{F_{1\to h}} - 1\right)^{-1}$$

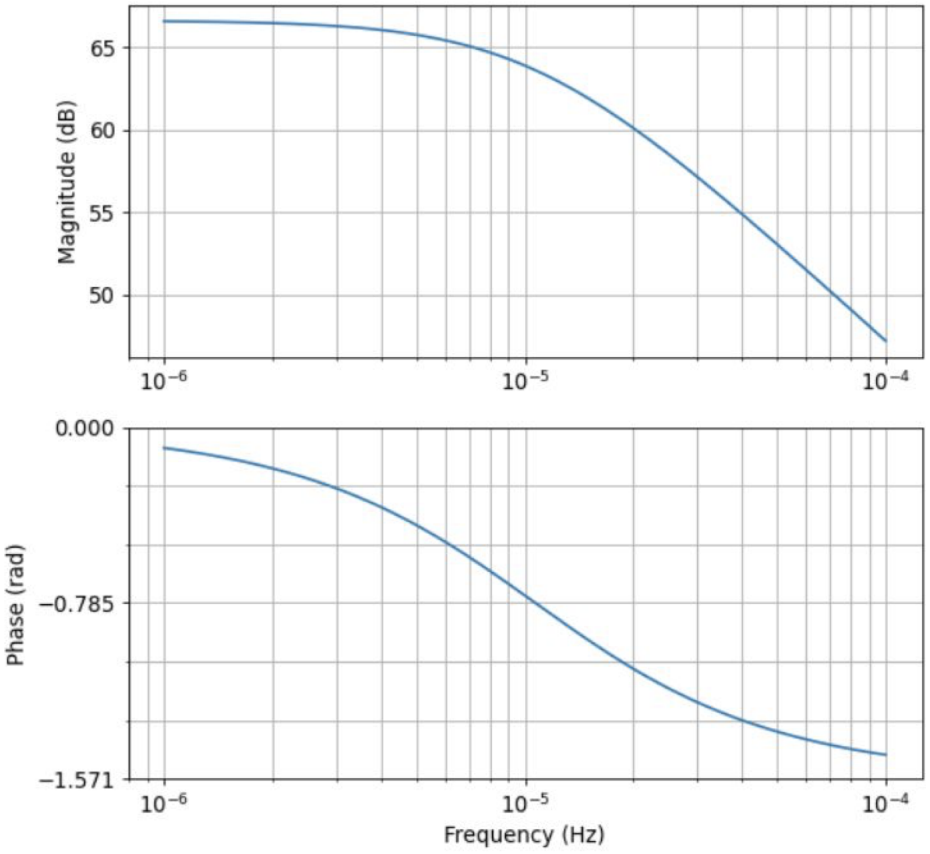

Figure 5: Bode plot of the heat model transfer function

As we can see the cross-over frequency is very low, around 10^{-5} Hz. This means that the time constant of our system is $\approx$ 27 hours. Any frequency above 10^{-5} Hz will be damped.

We will have our optimal frequency search region around this region. Intuitively, we can reason that it is not possible to extract information faster than the time constant of the system.

4.2 Frequency Domain Analysis of Linearized System

[3] provides a framework for finding optimal excitation using Fisher Matrix for a two parameter linear system. We started out with finding one optimal frequency, as done in [3], and then proceeded to use two frequency input signal. We use all four parameters to build a 4×4 Fisher Matrix for our linearized system given by equation (2).

4.2.1 One frequency input signal

Refer figure 6 which shows the variation of the determinant of the Fisher Matrix vs the frequency of the input signal. We see that the maxima occurs around optimal frequency of the order 10^{-5} Hz. This observation is consistent with the pole frequency of the transfer function of our linearized system (3), as seen in the bode plot (figure 5). In this analysis we assume white noise (or constant noise amplitude for all frequencies). The plot was generated by running an algorithm which calculated each element of the Fisher Matrix.

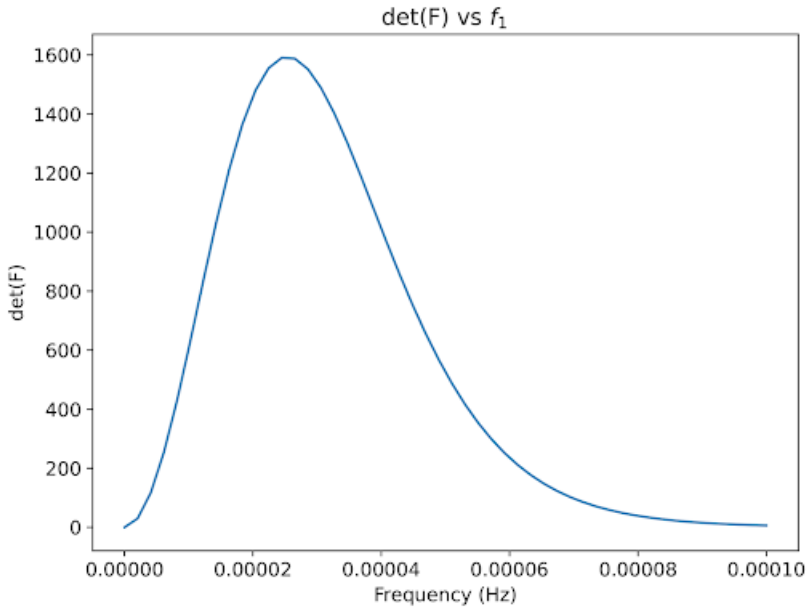

Figure 6: $\det(F)$ vs frequency f_1 for a one frequency input signal

4.2.2 Two frequency input signal

Let us see what would happen if we make our signal more complex by adding another frequency. We can plot the determinant contours on a graph and see at what combination of frequencies the maxima occurs. These frequencies would be our optimum frequencies which give minimum uncertainty in parameters. We observe in our result (figure 7) that the

optimum frequencies are degenerate and of the order 10^{-5} Hz. Thus, we obtained optimum frequencies are close to the time constant of the system, as predicted.

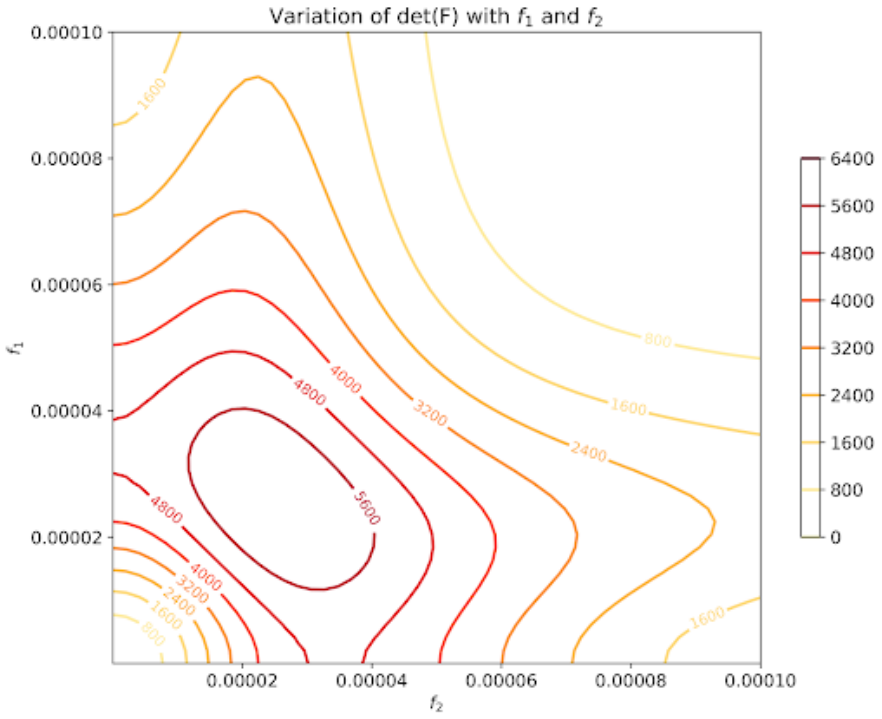

Figure 7: Contours of det (F) with both frequencies on the axes

A few points to note:

- Degeneracy occurs because we assigned the same signal and noise amplitudes to both the frequencies. This means we assigned the same σ_α to both the frequencies. Assigning a different σ_α would give us a non degenerate result.
- We assumed our noise to be white noise. Degeneracy is lost if we assume frequency dependant noise.
- Assigning different power or amplitudes to the signal of the two frequencies will also give a non-degenerate optimum result.

We can conclude that considering two input frequencies in the Fisher analysis gives us no new information (as compared to one input frequency) unless we consider that the system noise is frequency dependent. Hence, determining this frequency dependence is an important consideration for future work.

It is evident that there is a problem here. Practically, assigning signal amplitudes or deciding σ_α for the various frequencies is time intensive. There are many combinations of amplitudes possible for the signal and the noise for each frequency. Moreover, giving a higher magnitude of amplitude to a particular frequency automatically gives it more importance and can end up being the optimal frequency. The logic would become circular. Therefore, we want a method that can give us both the amplitude (or power) and the optimal frequency together, while maximizing the Fisher Information Matrix.

5 Power Spectrum Optimization

This method involves distributing some fixed total power among discrete frequencies within a particular range. Finding optimal frequencies by maximizing the determinant of the Fisher Matrix involves choosing the excitation amplitude of the signal beforehand (when assigning the standard deviation for that frequency). The power spectrum optimization algorithm introduced in [4] gives us information about what the optimum power distribution should be. The main advantage of this algorithm is that it can constrain the total power and also find the optimal excitation amplitude for a range of frequencies.

5.1 Dispersion Function

Dispersion function $\nu(\chi, \Omega_k)$, as defined in [4], for a given input power spectrum $\chi(\Omega) = (|U(1)|^2, \cdots, |U(F)|^2)$, with $\sum_{k=1}^{F} |U(k)|^2 = \mathfrak{P}$ is,

$$\nu(\chi, \Omega_k) = \text{trace}([\mathcal{F}(\chi)]^{-1} fi(\Omega_k)),$$

with $\mathcal{F}(\chi)$ the information matrix resulting from the design $\chi(\Omega)$, $fi(\Omega_k)$ the information matrix corresponding to a single frequency input with a normalized power spectrum $|U(k)|^2 = \mathfrak{P}$, and Ω_k the frequency.

To intuitively understand the meaning of this function, we will look at the one parameter case. The dispersion function then can be explained as a scalar measure of the covariance for the given input power spectrum with respect to the covariance for a single frequency input with normalized (concentrated) power spectrum.

One of the properties of the dispersion function is that the maximum of the dispersion function over the frequency grid is larger than or equal to the number of parameters. Thus our aim is to minimize the dispersion function and make it converge to the number of parameters.

In a way it can be understood how bad the covariance of our system is, at the estimated parameters, in comparison to the covariance when the power is concentrated at one frequency. It encourages us to find a power distribution which minimizes the covariance instead of concentrating the power at one frequency (which may or may not be one of the optimal frequencies).

Another way of expressing the dispersion function, as done in [4],

$$\nu(\chi, \Omega_k) = \frac{2\sigma_G^2(\Omega_k, \bar{\theta})\mathfrak{P}}{\sigma_U^2(k)|G(\Omega_k)|^2 + \sigma_Y^2(k) - 2\,\text{Re}(\sigma_{YU}^2(k)\bar{G}(\Omega_k))}.$$

This can be interpreted as the ratio of the variance of the system frequency response, calculated with the estimated parameters, to the noise power of the measurements at the frequency Ω_k. This makes sense as we want to minimize the variance of the frequency response (which encodes the system parameters) with respect to a particular noise power.

5.2 Algorithm

- **Step 1:** Select a set of frequencies $\mathbb{F} = \{\Omega_1, \cdots, \Omega_F\}$ within our band of interest. The input power is distributed equally amongst these F frequencies (called design χ_0).
- **Step 2:** Set $i = i + 1$ in the numerical algorithm and calculate $\nu(\chi_i, \Omega_k)$ for $k = 1, \cdots, F$.
- **Step 3:** Let n_θ be the number of parameters. Compose new design,

$$\chi_{i+1}(\Omega_k) = \chi_i(\Omega_k)\frac{\nu(\chi_i, \Omega_k)}{n_\theta} \quad \text{for } k = 1, \cdots, F.$$

- **Step 4:** If $\max(\nu(\chi_i, \Omega_k) - n_\theta) < \epsilon$ with a sufficiently small ϵ and $\Omega_k \in \mathbb{F}$, then the optimal design is found. Otherwise, we return to Step 2.

5.3 Results

We analyse the evolution of the power spectrum optimization algorithm for our linearized system given by equation (2). For iterations 1, 2, 5, 10 and 100 we obtain the results shown in figure 8. It can be observed with each iteration, the power spectrum becomes more and more concentrated around a maxima. After 100 iterations about 5W of the power is assigned to a frequency of the order 10^{-5} Hz. This result also matches our initial Fisher Matrix optimization analysis. We see the advantage of this method, as it spits out the optimal power distribution for the amplitudes of a range of frequency input signals.

6 Optimal Experimental Configuration

We try to see the trend of $\text{var}(\epsilon_1)$ or the variance of ϵ_1 and $\det(F)$ or the determinant of the Fisher Matrix with respect to various parameters (figure 9). The frequency used in this analysis is the optimal frequency obtained in the single frequency Fisher Matrix analysis.

We see that the variance of ϵ_1 does not change much with parameters ϵ_2 and ϵ_h. However, it does change a lot with r_h, which is the radius of the heat leak aperture. The variance increases with an increase in the radius, which is quite intuitive.

The trends for the determinant of the Fisher Matrix are hard to explain intuitively. More work can be done in validating these results. However, we do see the determinant attaining a maxima with parameters ϵ_1, ϵ_2 and ϵ_h being equal to 1. This means that information is maximized when the system components have maximum radiative coupling. We also note the determinant is max for the smallest value of r_h. We can thus say, that heat leaks add more variation in our system and make us less certain about the parameter estimates.

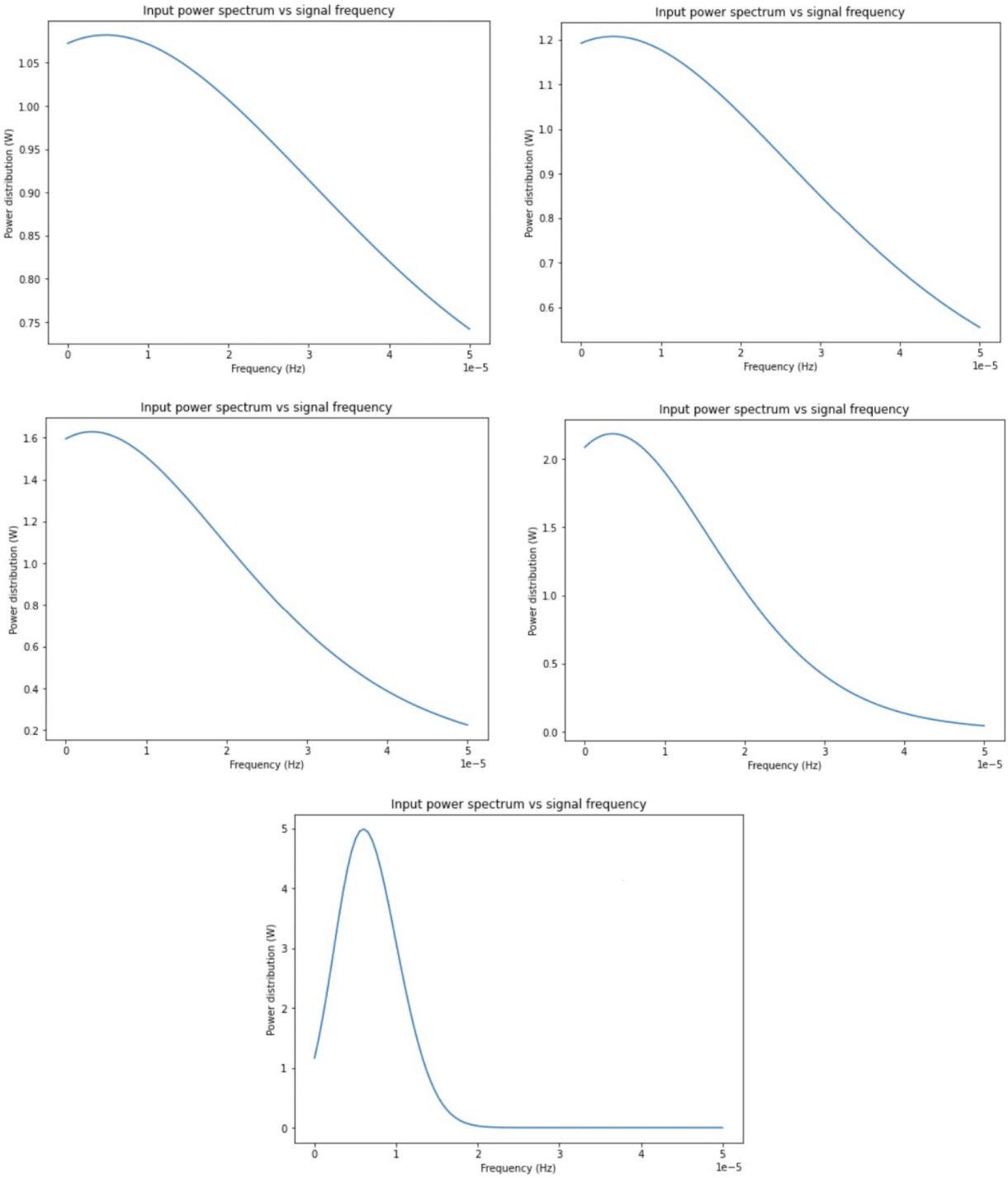

Figure 8: Power spectrum optimization for 1, 2, 5, 10 and 100 iterations respectively (from left to right)

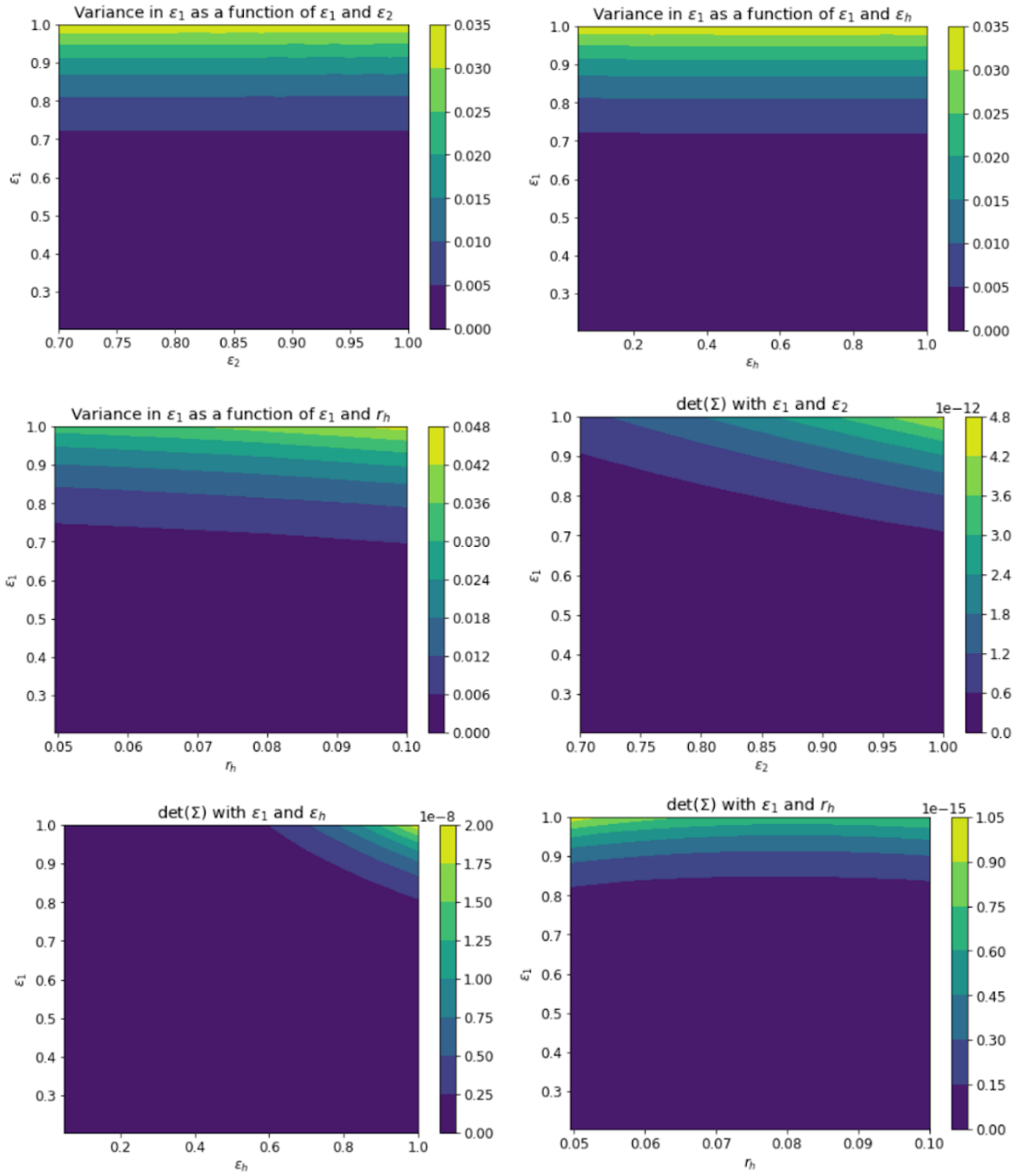

Figure 9: Variation of $var(\epsilon_1)$ and $\det(F)$ with various parameters

7 Future work

Some ideas that can be worked on in the future are the following:

1. Understanding the mathematics of the Power Spectrum Optimization algorithm. Answering questions like: What is being minimized? Why are the answers not exactly equal to the Optimal Control Frequency analysis?
2. Intuitively explaining trends observed in Optimal Experimental Configuration.
3. Estimating or modelling the noise in our system from various sources (ambient fluctuations of temperature, measurement noise, etc.).
4. Time-domain analysis of our original nonlinear system without assuming steady state conditions. As frequency domain is only for LTI systems, the evaluated optimal frequencies are still impractical.
5. A different problem can be posed: How do we get maximum information out of the system within a fixed amount of time?

8 Acknowledgments

I would like to thank my mentors Prof. Rana Adhikari, Christopher Wipf, and Radhika Bhatt for their constant guidance and support. I would also like to extend this gratitude towards the LIGO SURF program and National Science Foundation (NSF) for giving me the opportunity to work alongside scientists at LIGO on such a large scale project.

References

[1]M. Constancio Jr, R. X. Adhikari, O. D. Aguiar, K. Arai, A. Markowitz, M. A. Okada, and C. C. Wipf, "Silicon emissivity as a function of temperature", International Journal of Heat and Mass Transfer **157**, 119863 (2020).

[2]Y. A. Çengel, *Heat transfer: a practical approach*, McGraw-Hill series in mechanical engineering (McGraw-Hill, 2003) Chap. 12.

[3]E. D. Hall, "Fisher matrix methods for transfer function measurement", ligo (2015).

[4]R. Pintelon and J. Schoukens, *System identification: a frequency domain approach* (John Wiley & Sons, 2012).

[5]J. Tellinghuisen, "Statistical error propagation", The Journal of Physical Chemistry A **105**, 3917–3921 (2001).

[6]D. Wittman, "Fisher matrix for beginners", Technical report, UC Davis.

[7]R. X. Adhikari, K. Arai, A. F. Brooks, C. Wipf, O. Aguiar, P. Altin, B. Barr, L. Barsotti, R. Bassiri, A. Bell, G. Billingsley, R. Birney, D. Blair, E. Bonilla, J. Briggs, D. D. Brown, R. Byer, H. Cao, M. Constancio, S. Cooper, T. Corbitt, D. Coyne, A. Cumming, E. Daw, R. deRosa, G. Eddolls, J. Eichholz, M. Evans, M. Fejer, E. C. Ferreira, A. Freise, V. V. Frolov, S. Gras, A. Green, H. Grote, E. Gustafson, E. D. Hall, G. Hammond, J. Harms, G. Harry, K. Haughian, D. Heinert, M. Heintze, F. Hellman, J. Hennig, M. Hennig, S. Hild, J. Hough, W. Johnson, B. Kamai, D. Kapasi, K. Komori, D. Koptsov, M. Korobko, W. Z. Korth, K. Kuns, B. Lantz, S. Leavey, F. Magana-Sandoval, G. Mansell, A. Markosyan, A. Markowitz, I. Martin, R. Martin, D. Martynov, D. E. McClelland, G. McGhee, T. McRae, J. Mills, V. Mitrofanov, M. Molina-Ruiz, C. Mow-Lowry, J. Munch, P. Murray, S. Ng, M. A. Okada, D. J. Ottaway, L. Prokhorov, V. Quetschke, S. Reid, D. Reitze, J. Richardson, R. Robie, I. Romero-Shaw, R. Route, S. Rowan, R. Schnabel, M. Schneewind, F. Seifert, D. Shaddock, B. Shapiro, D. Shoemaker, A. S. Silva, B. Slagmolen, J. Smith, N. Smith, J. Steinlechner, K. Strain, D. Taira, S. Tait, D. Tanner, Z. Tornasi, C. Torrie, M. V. Veggel, J. Vanheijningen, P. Veitch, A. Wade, G. Wallace, R. Ward, R. Weiss, P. Wessels, B. Willke, H. Yamamoto, M. J. Yap, and C. Zhao, "A cryogenic silicon interferometer for gravitational-wave detection", Classical and Quantum Gravity **37**, 165003 (2020).