Precision and Efficiency: An Automated System for Laser Alignment in Optical Experiments

Jerry Liu

Under the direction of

Yuqin Sophia Duan
EECS Ph.D
MIT

Qiushi Gu
Postdoctoral Research Associate
MIT

Research Science Institute
August 1, 2023

Abstract

The precise alignment of light sources with detectors or other sources poses a challenge in numerous optics experiments. For instance, interferometry experiments often demand light rays to be nearly parallel to observe interference effects. Furthermore, the modern growth of quantum optics, which utilizes very low levels of light such as single photons, emphasizes the importance of perfectly aligned light sources. This study specifically focuses on aligning a laser beam into a fiber optic cable to minimize power losses and maximize coupling efficiency. Currently, the alignment of light sources relies on the manual and time-consuming process known as "beam-walking," which involves using one's hands to fine-tune two mirrors to guide a light beam to a specific location and angle. Not only is a precise manual beam-walk very time-consuming, but, in the case of coupling efficiency, it is also susceptible to local maxima trapping or overshooting the true global maximum due to the imprecision of the human hand. In this study, we present an automated setup using stepper motors to precisely align and couple a laser beam into a fiber optic cable to maximize coupling efficiency. The setup utilizes a modified gradient descent algorithm with a variable learning rate to accurately converge to the maximum, with additional stochastic elements to prevent local maxima trapping. We show that such an automated setup achieves a $4-5\%$ increase in coupling efficiency in a $4-5$ hour time frame for a single-axis optimization. Moreover, with the integration of other degrees of freedom, as well as highly precise industrial motors and measuring devices, the efficiency and time consumption can be improved further. Ultimately, the automation of the laser alignment process has tremendous implications, as it will allow for greater precision and efficiency in various optical experiments.

Summary

Laser alignment plays a critical role in various optical experiments, from ensuring parallel light beams in interferometry to detecting single photons in quantum optics. This study specifically focuses on aligning and coupling a laser beam into a fiber optic cable to minimize power losses, which improves signal transmission and enhances optical communication. Here, we present an automated system to perform such an alignment, achieving a $4-5\%$ increase in coupling efficiency in a $4-5$ hour time frame for a single-axis optimization.

1 Introduction

The alignment of laser beams plays a crucial role in numerous optics experiments and has various applications, such as making laser beams collinear in interferometry [1, 2] and minimizing power losses when coupling a laser beam into an optical fiber [3, 4]. Furthermore, with the modern growth of quantum optics, which utilizes experiments involving extremely low light levels such as single-photon generation [5, 6, 7], the alignment of light sources into detectors is becoming even more significant. This study specifically focuses on aligning a laser beam into a fiber optic cable to minimize power losses and maximize coupling efficiency. Maximizing efficiency is important because it improves signal transmission, enabling more accurate and sensitive measurements in optical experiments and communications [8, 9].

Traditionally, the alignment of lasers is done through a manual process known as "beam-walking," which uses two adjustable mirrors to guide a laser to a specific point at a specific angle (refer to Figure 1) [10]. However, this process is laborious and time-consuming, and since it misses many possible mirror angles, it does not necessarily find the optimal alignment.

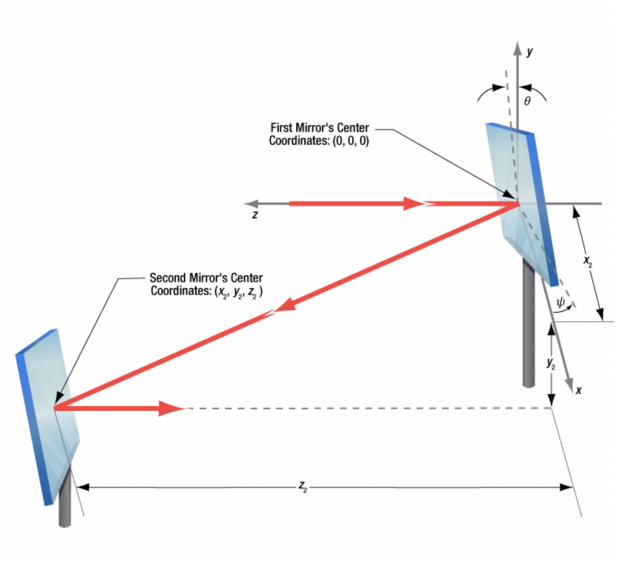

Figure 1: Beam-Walk: Two mirrors each with tip and tilt adjustors (indicated by θ and ψ on the first mirror) are used to deflect an incoming laser beam (red line) [11].

At the same time, machine learning (ML) is becoming more prominent in optical physics, and in the past decade, ML has assisted in various experiments from computation imaging and super-resolution [12, 13] to optical communication and networking [14, 15]. However, little research has been done on the applications of machine learning to the automation of beam-walking, despite the potential to achieve much more accurate laser alignments.

In this study, we propose an automated system, using cheap and open-source control electronics and off-the-shelf optomechanical components, utilizing ML techniques to align and couple a laser beam into a fiber optic cable while maximizing coupling efficiency. With results from this automated system, we can make comparisons with other methods of beam-walking—for instance, manual beam-walking or physics-informed models such as ray-tracing. Such comparisons are crucial because the ability to significantly increase coupling efficiency while maintaining a reasonable duration of automation will have tremendous implications on the field of optics, as it will facilitate improved transmission of light power and lead to substantial time savings in various optical experiments.

2 Background

2.1 Mirror Adjustment

Of particular importance in this paper is mirror adjustment within the beam-walking setup. As shown in Figure 2, the two knobs, referred to as the *pitch* and *yaw,* control the mirror's angle to the $y-$ and $x-$axis, respectively. In other words, this means that the *pitch* adjusts the reflected beam in the vertical direction and the *yaw* adjusts it in the horizontal direction, which makes it easy to direct the reflected beam to a specific location.

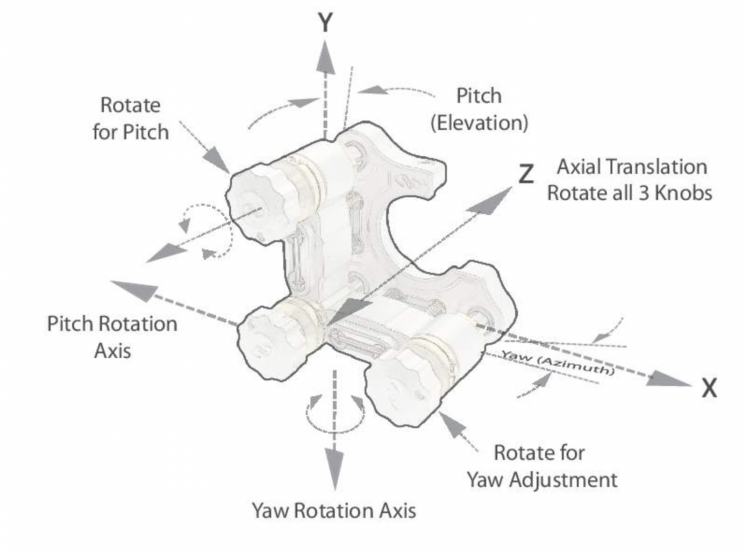

Figure 2: Diagram of the knobs and corresponding adjustments on a mirror [16].

Because the beam-walking setup contains two mirrors, there are four degrees of freedom (one for each knob). Realistically, there is coupling between the pitch and yaw axes, so the degrees of freedom are not independent.

2.2 Stepper Motors

To automate the adjustment of the mirrors, we connect each mirror knob to a stepper motor. The physical setup of the motors is shown in Figure 3.

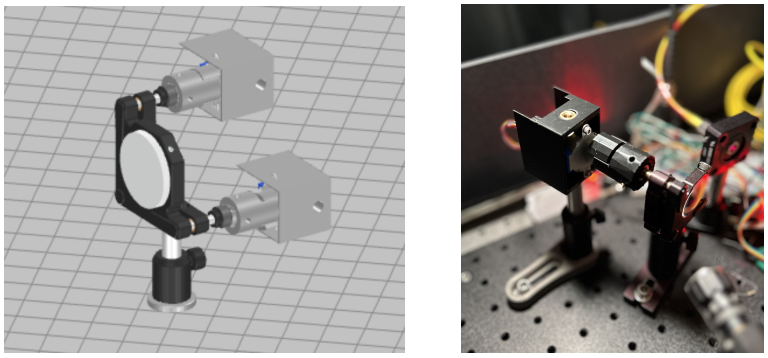

Figure 3: Physical setup of the motors: 3D Optix Drawing (left) and actual setup (right).

The smallest degree step that our motor can turn is 0.1766°. For the rest of the study, the data uses *steps* to characterize the angular deviation of a motor and mirror knob. Issues with the stepper motors, such as hysteresis, are discussed in Section 6.

2.3 Initial Alignment and Automated Setup

The automated method that we will present to maximize coupling efficiency requires first that the mirrors are (a) manually coarsely aligned and (b) manually beam-walked. Note, however, that the beam-walk in this case does not need to be very precise and is therefore much quicker than a traditional, precise manual beam-walk. In fact, the only reason why a manual beam-walk is necessary is to register a non-zero power reading from the detector fiber optic cable (see Figure 4). Furthermore, this beam-walk only needs to be performed once. Even with potential mirror and thermal drifts over time [17, 18], the deviations would be so small that simply running the automated algorithm again would realign the setup.

To begin a coarse alignment, position two laser sources, coupled into fiber optic cables and collimated, i.e. the light rays are parallel and not diverging, by lenses with two mirrors (similar to Figure 1) so that the laser paths are nearly coincident. Then, a coarse alignment can be performed by slightly adjusting the mirror angles using the knobs described in Section 2.1. See [11] for a coarse alignment algorithm.

To perform the manual beam-walk, replace one of the laser sources with a detector to measure the power output. The setup is similar to Figure 4, but without the stepper motors attached to the mirror knobs. In our setup and in the rest of this study, "Mirror 1" refers to the mirror closer to the detector. Refer to Appendix A for a beam-walk algorithm.

Finally, Figure 4 shows a diagram of the automated setup. Note that in the physical system, there would actually be four stepper motors, one for each mirror knob.

Additionally, see Figures 5 and 6 for details about the motors and voltage sensor.

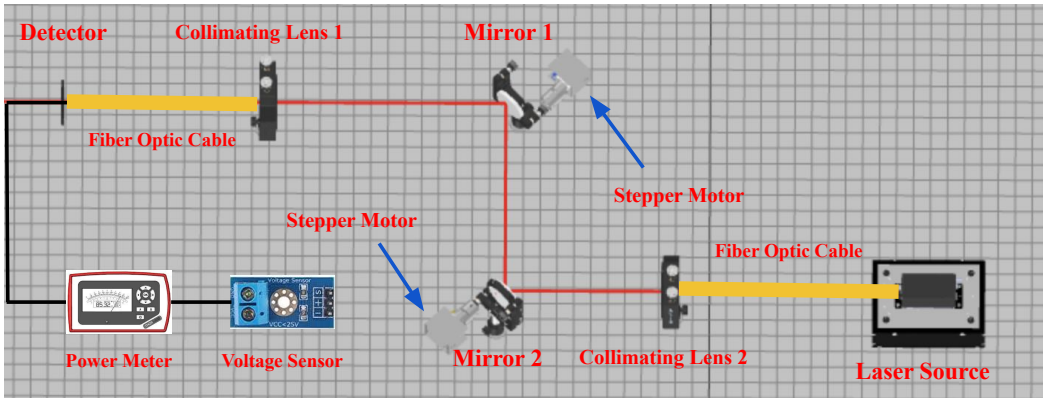

Figure 4: Automated Setup: Laser beam (red line) "walked" into a detector by using automated stepper motors attached to the mirror knobs. The detector and fiber optic cable are connected to a power meter to allow for coupling efficiency calculation.

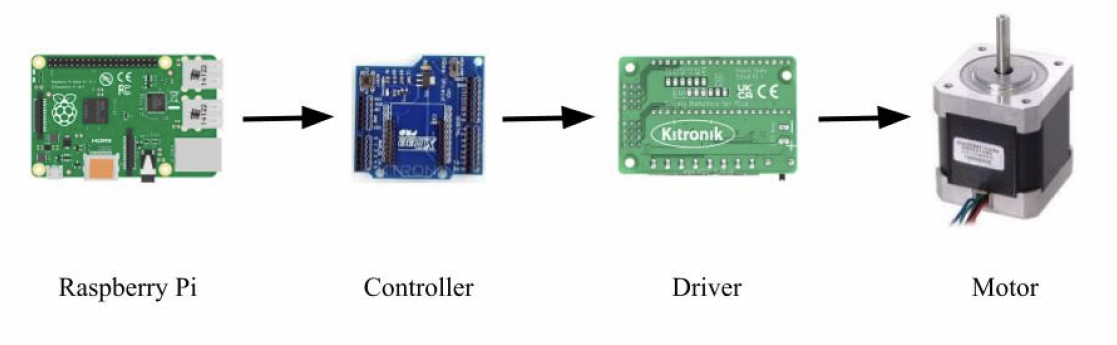

Figure 5: Motors are controlled by a Raspberry Pi, controller, and driver connected in series.

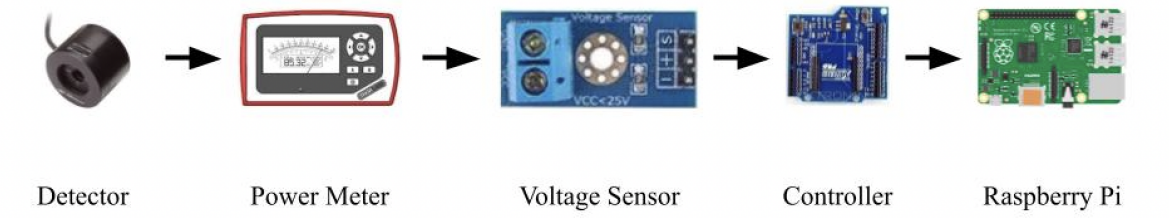

Figure 6: Automatic voltage measurement: Sensor draws voltage from the power meter and sends information to the Raspberry Pi via the controller.

3 Methods

3.1 Single-Knob Optimization

Our first goal is to automate a single-knob optimization, i.e. automate the turning of a single mirror knob to maximize coupling efficiency. It is traditionally known that the coupling efficiency, or alternatively the output power, follows a Gaussian distribution with respect to the angular displacement of a single mirror knob [19].

We can also verify this experimentally using our automated setup by sweeping through angles of Mirror 1's pitch and plotting the measured voltages. The result is shown below.

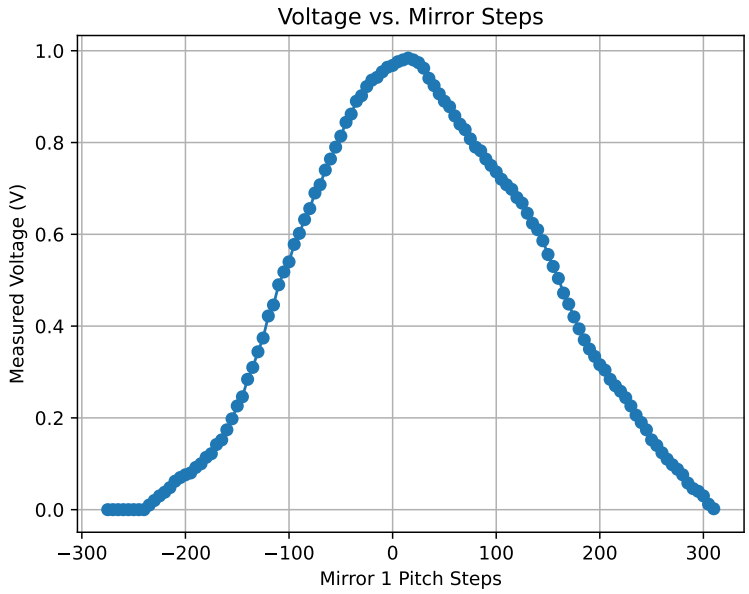

Figure 7: Graph of measured voltage vs. Mirror 1 Pitch Steps through brute force sweep.

Our experimentally measured data matches relatively well to the theoretically expected Gaussian distribution. More importantly, the data demonstrates that the graph of Measured Voltage vs. Mirror 1 Pitch Step for the single mirror knob is a single-peak maximum, suggesting that we can use a modified ML gradient descent algorithm [20] to locate and measure

this single maximum. Refer to Appendix B for a general algorithm outline.

Figure 8 shows the results of running our gradient descent algorithm on our setup from two different starting angles. Note that the background Gaussian distributions are slightly different because two different sweeps were run to make sure that coupling in other axes did not affect our algorithm. It is evident that our algorithm properly converges near the maximum of the Gaussian, with small deviations (which are explained in Appendix B).

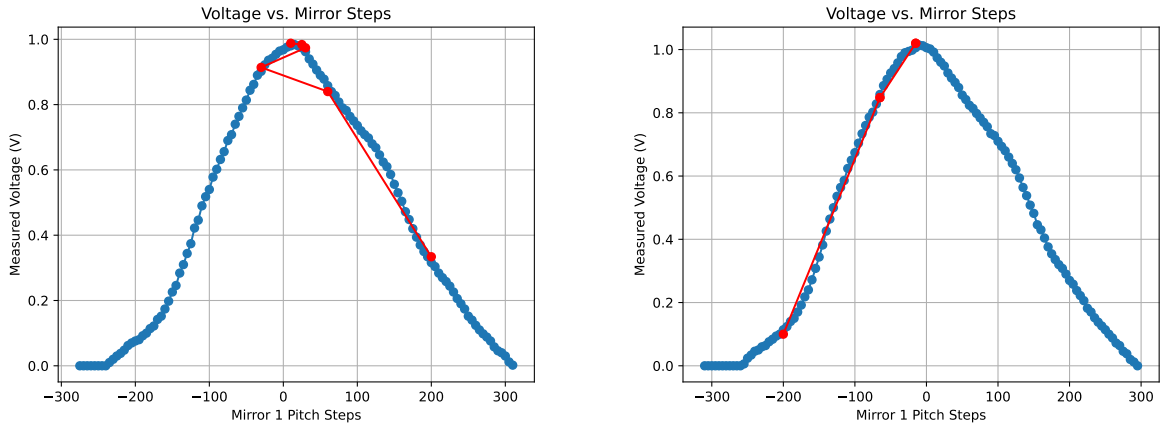

Figure 8: Gradient descent algorithm attempting to find the maximum measured voltage. The blue points represent the Gaussian distribution found via a sweep, and the red points and lines represent consecutive steps made by our algorithm. Initial conditions: $\theta_{\text{start}} = 200$ steps, $\alpha = 100000$ (left); $\theta_{\text{start}} = -200$ steps, $\alpha = 100000$ (right).

3.2 Single-Axis Optimization

Now we move on to a single-axis optimization, i.e. using two motors connected to the "pitch" mirror knobs to optimize the y-axis coupling. First, we find the general shape of the 3D plot of Measured Voltage vs. Motor 1 and 2 Pitch Angles.

To determine the shape theoretically, we use the Zemax ray tracing software, which is the industry-standard modeling system for optical experiments [21, 22, 23]. In the software, we input a general beam-walking setup that contains several specifics from our actual setup,

such as the collimating lenses used. Note that the goal of using Zemax is solely to obtain the general shape of a 2D map; in particular, the setup and map do not need to be extremely precise. See Appendix C for the Zemax setup used.

To get the 2D map, we take "slices" for Mirror 1 Pitch Steps and sweep over a large range of Mirror 2 Pitch Steps. Then, we graph it against the coupling efficiency calculated by the Zemax POP Function [24, 25]. Note that the coupling efficiency is proportional to the measured output voltage, since the input laser voltage and power is assumed to be constant. The results are shown in Figure 9 below.

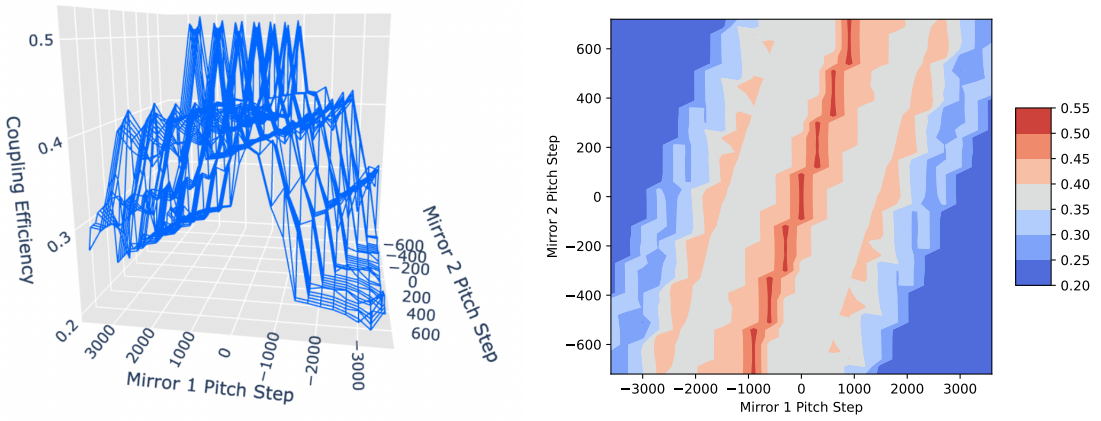

Figure 9: Zemax (theoretical) plots of Coupling Efficiency vs. Mirror 1 and 2 Pitch Angles by brute force sweep: 3D meshplot (left) and colormap (right).

First, observe that the cross-sections of the 3D meshplot are jagged peaks. This does not match with what we would physically expect; since we are sweeping through the mirror angles continuously, we should see a smooth 3D surface. In fact, the jaggedness arises inherently from the amount of sampling done by the Zemax software, which is limited by our computing power. For this reason, we cannot draw much information from these cross-sections.

Another feature that is particularly prominent is the linear nature of the peaks of these cross sections. This feature is seen clearly in the colormap in Figure 9 by looking at the points with maximum voltage (the red points). The points appear to fit a linear relationship, with

several discontinuities due to the slices made in the Mirror 1 Pitch Step (consecutive slices are taken 300 steps apart). From this, we may hypothesize that, experimentally, the 3D map has voltage peaks that lie roughly on a line.

To verify this feature experimentally, we again take slices for fixed Mirror 1 angles and sweep over Mirror 2 angles. To do this efficiently, for each slice, we maintain a sliding window of "interesting" Mirror 2 angles i.e. angles which give us a non-zero reading. Then, as we move through the slices, we only have to scan angles within the vicinity of the sliding window. Doing this, we obtain the 3D plots shown below.

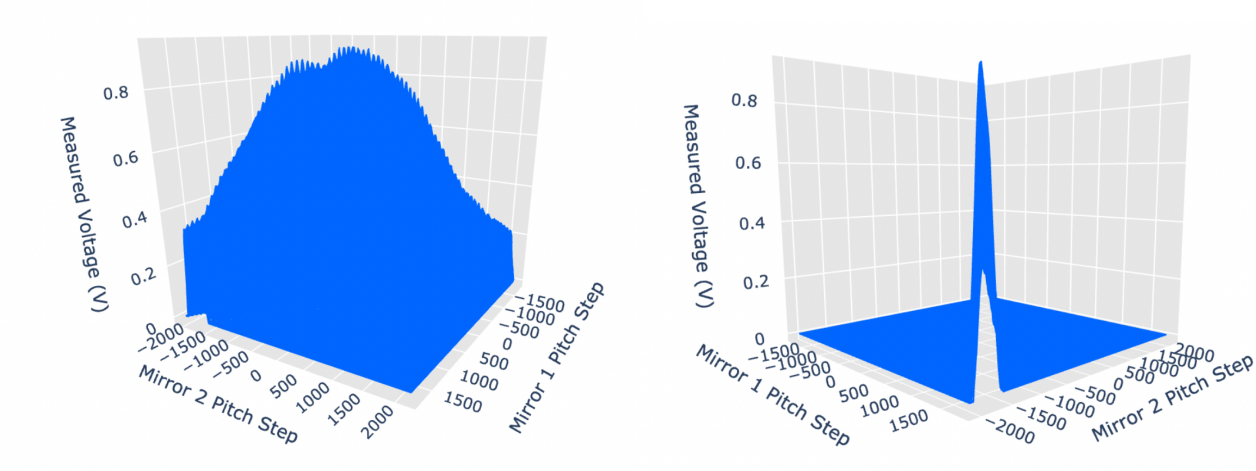

Figure 10: 3D mapping of Measured Voltage vs. Mirror 1 and 2 Pitch Steps.

Then, creating a colormap as well as a linear plot of the Mirror 1 and 2 Pitch Steps where the voltages are maximum, we obtain the plots shown in Figure 11.

The plots suggest that, generally, the voltage peaks do indeed lie on a line, which matches our theoretical expectation. However, in the linear plot, there are sharp discontinuities around Mirror 1 Pitch $= -1450$ steps and 1700 steps. But under closer inspection, these discontinuities are due to lower-order optical effects and can therefore be ignored, as their maxima are negligible compared to the global maximum. Refer to Appendix D for more details.

Thus, ignoring these discontinuities, we can calculate the primary gradient of the voltage peaks. This is shown by a red line (ranging from -1450 to 1650 steps) in the linear plot in

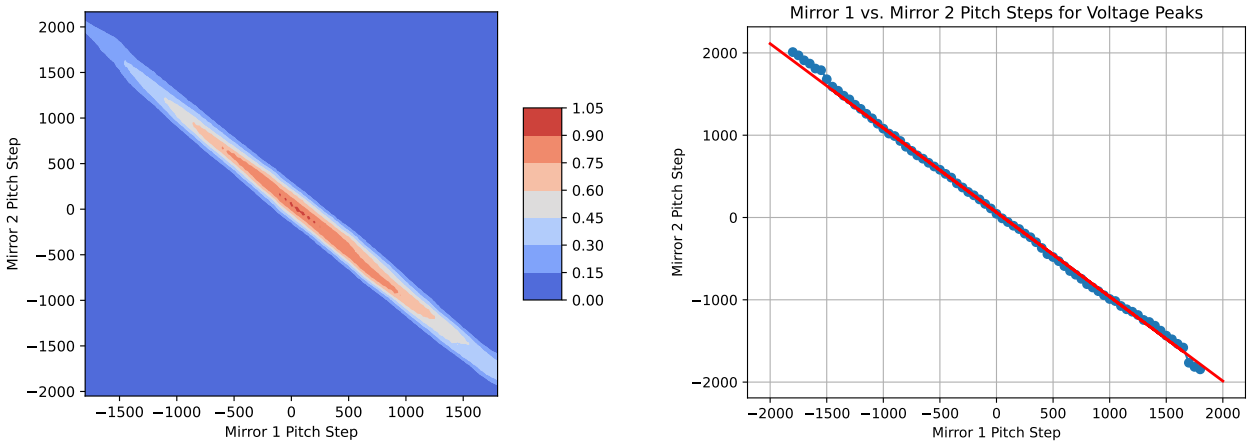

Figure 11: Colormap of measured voltages (left) and linear plot between Mirror 1 and Mirror 2 Pitch Steps for voltage peaks (right).

Figure 11. The equation of the least-squares regression line is

$$\text{Mirror 2 Pitch Step} = -1.024 \cdot (\text{Mirror 1 Pitch Step}) + 62.39 \tag{1}$$

The linear nature of the voltage peaks allows us to effectively reduce this 2D optimization problem into one dimension. In particular, we can use a similar 1D gradient descent as before, but instead always move along the peaks by using the measured slope.

For the sake of implementation, we approximate the slope (-1.024) to be -1, i.e. for every change in Mirror 1's Pitch we change Mirror 2's Pitch by an equal but opposite value. Then, every time we travel along this slope, we perform a small sweep of Mirror 2's Pitch to locate the true peak (since there is some error in the slope).

Additionally, since there are some local maxima along the envelope of the voltage peaks on the primary gradient (see Figure 10), we need to incorporate stochastic elements into our algorithm to prevent local maxima trapping. To do this, we randomly choose five starting points (within -1200 and 1200 Mirror 1 Pitch steps of the initial position, which will typically be the position of the initial hand alignment). Then, we perform our gradient descent from

each point and take the maximum at the end. To further increase the rate of convergence to the global maximum, the user can input more starting points.

Finally, see Appendix B for an appropriate learning rate. The results of running the algorithm for two different starting alignments are shown in the following section.

4 Results

Here, we show the results of running our gradient descent algorithm on two separate starting alignments. For each of the alignments, the gradient descents starting from different random Mirror 1 Pitch Steps are depicted in different colors. Furthermore, once our algorithm finds where the global maximum is roughly located, it performs a small 2D sweep in the vicinity to find the true maximum.

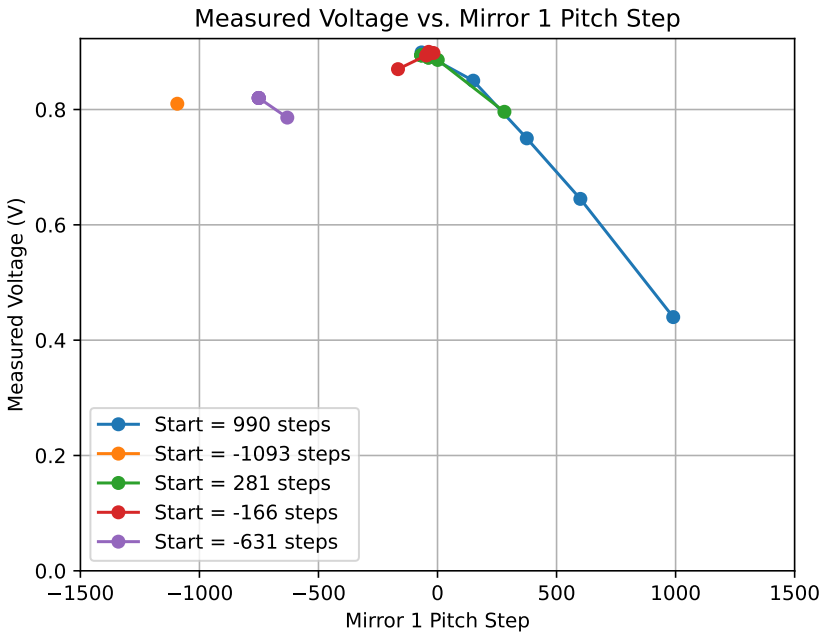

Figure 12: Trial 1: Initial hand alignment gives power output 77.1 μW. The maximal convergence point from the algorithm is at approximately Mirror 1 Pitch = -67 steps. Automated system's power output: 80.3 μW, Time of duration: 4 hours.

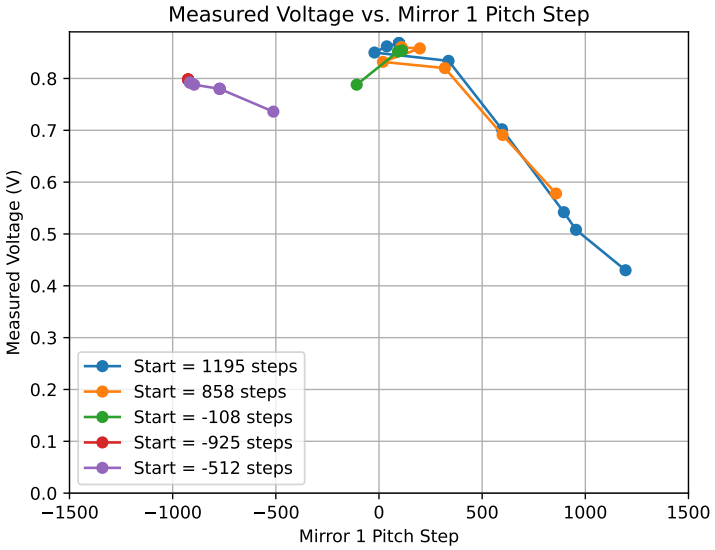

Figure 13: Trial 2: Initial hand alignment gives power output 74.6 μW. The maximal convergence point from the algorithm is at approximately Mirror 1 Pitch = 97 steps. Automated system's power output: 78.3 μW, Time of duration: 5 hours.

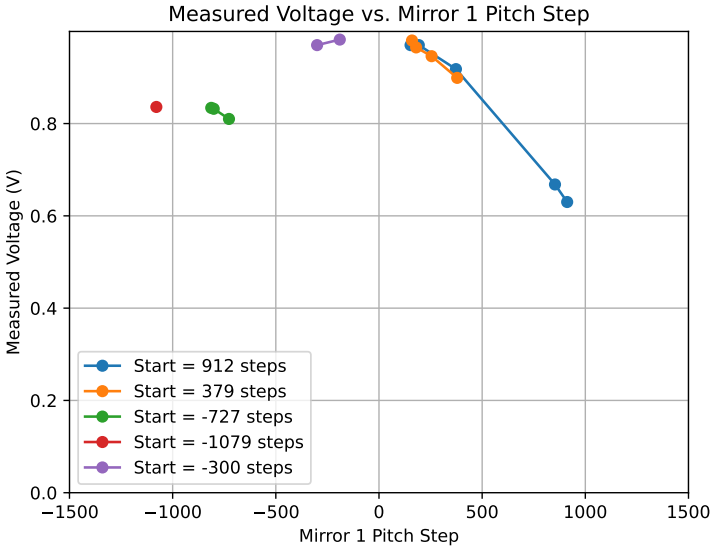

Figure 14: Trial 3: Initial hand alignment gives power output 80.1 μW. The maximal convergence point from the algorithm is at approximately Mirror 1 Pitch = -190 steps. Automated system's power output: 84.0 μW, Time of duration: 4 hours.

We can summarize our results in the following table:

Trial #	P_{Manual} (μW)	$P_{\text{Automated}}$ (μW)	% Gain	Time of Automation
1	77.1	80.3	4.15%	4 hrs
2	74.6	78.3	4.96%	5 hrs
3	80.1	84.0	4.87 %	4 hrs

In the table, the % Gain is calculated by

$$\% \text{ Gain} = \frac{P_{\text{Automated}} - P_{\text{Manual}}}{P_{\text{Manual}}} \tag{2}$$

which represents the percent increase in the power output. Since the laser source also has a constant power input, % Gain also represents the percent increase in the coupling efficiency.

5 Conclusion

The results suggest that, on average, our automated system's single-axis optimization improves a manual beam-walk's coupling efficiency by around $4-5\%$ in a time of automation of around $4-5$ hours. Although these results seem like an insignificant improvement, note that a manual beam-walk **should** perform relatively well in a single-axis optimization because many of the possible angles are still swept through. However, when additional degrees of freedom are added (e.g. the yaw axis), a manual beam-walk will begin missing many more positions, in which case the automated system will take over.

More importantly, the results demonstrate the repeatability and the validity of our gradient descent algorithm. As shown in Figures 12 to 14, each individual gradient descent seems to properly reach a local maximum. Furthermore, the random choice of starting points prevents our system from getting stuck at a single local maximum. The validity and repeatability of our algorithm has broad implications, as similar algorithms can be applied when more degrees of freedom are considered. This is discussed further in the next section.

6 Discussion and Future Work

In this study, we analyzed the applicability and benefits of an automated beam-walking alignment system for a single-axis optimization. For future research, it is necessary to also automate the other axis (the yaw) and perform an optimization on the four degrees of freedom. There are two potential methods of doing this:

1. Repeat the same optimization process as explained in Section 3.2. Do this process independently for the pitch and yaw axes, until the two maximums converge.

2. Still use the same optimization process as explained in Section 3.2, but do not do it independently for the pitch and yaw axes. Instead, since we can still walk along the voltage peaks for each axis, we can reduce this four-dimension optimization to a 2D gradient descent. This guarantees the global maximum, and the algorithm does not need to know anything about the physical system.

The first method is easier to implement and can be easily extended from this study. On the other hand, the second method explained above has more general applications. In particular, with additional degrees of freedom (e.g. the focal lengths of the collimating lenses, distances between mirrors), we can add more dimensions to the gradient descent.

Another part of the study that requires further research is the use of more precise equipment, e.g. industrial motors. The stepper motors that were used in this study were very basic 3D-printed CAD designs, and were thus subject to issues such as inaccuracy and hysteresis. For example, due to the imperfect alignment of the motor with the mirror knob as well as the gaps between the two joints, an additional angular "boost" was necessary to realign the motor and mirror and fill in the gap every time the motor changed directions. This led to slightly inaccurate motor turns, and additionally made it difficult to return the motor back to its original position. In our study, this issue was mitigated by constantly performing small

sweeps in the vicinity of an angle to find the true maximum. However, with more precise motors, this step can be bypassed completely, making the system more time-efficient. Additionally, more precise motors would be able to turn by smaller amounts, making the system less prone to overshooting the maximum coupling efficiency.

Another prominent issue was the inherent noise in the detector, power meter, and voltage sensor. Due to this noise, we were forced to select relatively large values of $\delta\theta$ to calculate the gradient at each point (refer to Appendix B) to ensure that calculated gradients were significant and not due to fluctuations in the readings. This, in turn, made the error of the location of the global maximum (which is on the order of $\delta\theta$) relatively large as well. Therefore, with more precise measuring equipment with less noise, we would be able to choose a smaller $\delta\theta$ and hence reduce this error.

Thus, for further research, the use of more accurate and precise equipment has the potential to significantly increase the coupling efficiency while lowering the time of automation.

7 Acknowledgments

From the MIT Quantum Photonics Group, I would like to thank my mentors, Yuqin Sophia Duan and Qiushi Gu, for guiding me through my project, helping me brainstorm new ideas, and always being patient with me. I would also like to thank the principal investigator, Professor Dirk Englund, for continually pushing for and sparking my interest in automation in various fields of research as well as keeping the Quantum Photonics Group going.

I would also like to thank my various teachers and tutors who helped me reach this point. Thank you to my RSI tutor, Agnes Robang, for giving me feedback and talking to me about my project. Thank you to my physics teacher at Amador Valley High School, Mrs. Barnett Dreyfuss, for being such a great mentor and supporter throughout my high school years.

Thank you also to the people who morally supported me through RSI: my friends, Alex

Li and Luke Huang; my counselor, Justin Ricketts; and my family.

Finally, thank you to the Massachusetts Institute of Technology (MIT), the Center for Excellence in Education (CEE), and sponsors for making this experience possible and giving me new insights into the marvelous world of research.

References

[1] M. M. Colavita. Adverse effects in dual-feed interferometry. *New Astronomy Reviews*, 53(11-12):344–352, 2009.

[2] M.-S. Hartig, S. Schuster, G. Heinzel, and G. Wanner. Non-geometric tilt-to-length coupling in precision interferometry: mechanisms and analytical descriptions. *Journal of Optics*, 25(5):055601, 2023.

[3] H. Kuniyil and K. Durak. Efficient coupling of down-converted photon pairs into single mode fiber. *Optics Communications*, 493:127038, 2021.

[4] A. A. Chesworth, R. K. Rannow, O. Ruiz, M. DeRemer, J. Leite, A. Martinez, and D. Guenther. Novel fiber fused lens for advanced optical communication systems. In *Terahertz, RF, Millimeter, and Submillimeter-Wave Technology and Applications IX*, volume 9747, pages 259–264. SPIE, 2016.

[5] G. Buller and R. Collins. Single-photon generation and detection. *Measurement Science and Technology*, 21(1):012002, 2009.

[6] C. Santori, D. Fattal, J. Vuckovic, G. S. Solomon, and Y. Yamamoto. Single-photon generation with inas quantum dots. *New Journal of Physics*, 6(1):89, 2004.

[7] S. Takeuchi. Recent progress in single-photon and entangled-photon generation and applications. *Japanese Journal of Applied Physics*, 53(3):030101, 2014.

[8] Q. He, Z. Zhao, X. Ye, C. Luo, D. Zhang, S. Wang, and X. Xu. Optical coupling efficiency of a coupler with double-combined collimating lenses and thermally expanded core fibers. *Micromachines*, 13(2):324, 2022.

[9] A. Suzuki, Y. Wakazono, S. Suzuki, M. Tamura, H. Masuda, T. Ishikawa, Y. Hashimoto, T. Suzuki, K. Kikuchi, H. Nakagawa, et al. High optical coupling efficiency using 45°-ended fibre for low-height and low-cost optical interconnect modules. *Electronics Letters*, 44(12):724–725, 2008.

[10] S. Agha and D. Minkin. Understanding 'walking the beam'. *Stony Brook Laser Teaching Center*, 2007.

[11] ThorLabs. Two steering mirrors can be used to walk a laser beam to a new path. https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=14221.

[12] T. Liu, K. De Haan, Y. Rivenson, Z. Wei, X. Zeng, Y. Zhang, and A. Ozcan. Deep learning-based super-resolution in coherent imaging systems. *Scientific reports*, 9(1):3926, 2019.

[13] G. Barbastathis, A. Ozcan, and G. Situ. On the use of deep learning for computational imaging. *Optica*, 6(8):921–943, 2019.

[14] D. Zibar, M. Piels, R. Jones, and C. G. Schäeffer. Machine learning techniques in optical communication. *Journal of Lightwave Technology*, 34(6):1442–1452, 2015.

[15] Y. Xie, Y. Wang, S. Kandeepan, and K. Wang. Machine learning applications for short reach optical communication. In *Photonics*, volume 9, page 30. MDPI, 2022.

[16] Newport. Optical mirror mount guide. https://www.newport.com/g/optical-mirror-mount-guide.

[17] A. Haber, J. E. Draganov, K. Heesh, J. Cadena, and M. Krainak. Modeling, experimental validation, and model order reduction of mirror thermal dynamics. *Optics Express*, 29(15):24508–24524, 2021.

[18] T. Le, A. Savchenkov, N. Yu, L. Maleki, and W. Steier. Optical resonant sensors: a method to reduce the effect of thermal drift. *Applied optics*, 48(3):458–463, 2009.

[19] M. Fadhali, J. Zainal, Y. Munajat, J. Ali, R. A. Rahman, et al. Mode matching for efficient laser diode to single mode fiber coupling. In *International Workshop and Conference on Photonics and Nanotechnology 2007*, volume 6793, pages 104–111. SPIE, 2008.

[20] S. Ruder. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*, 2016.

[21] N. Coluccelli. Nonsequential modeling of laser diode stacks using zemax: simulation, optimization, and experimental validation. *Applied optics*, 49(22):4237–4245, 2010.

[22] P. Wissmann, S. B. Oh, and G. Barbastathis. Simulation and optimization of volume holographic imaging systems in zemax®. *Optics express*, 16(10):7516–7524, 2008.

[23] A. Ahmed, M. T. Riaz, A. Raza, A. Zaib, M. A. Akbar, and M. B. Sarwar. Modeling and simulation of office desk illumination using zemax. In *2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, pages 1–6. IEEE, 2019.

[24] Y.-H. Chen, A. Morales, F. D. Gomez, T. Robertson, H.-H. Cheng, H. Chen, S. Lin, and K. Johnson. Design fiber-to-waveguide coupling for photonic integrated circuits. In *Optical Interconnects XXIII*, volume 12427, pages 71–74. SPIE, 2023.

[25] H. Wu, Z.-y. Yu, X.-S. Qian, B. Chen, Y.-q. Lu, et al. The efficiency improvement of a bidirectional optical subassembly. In *2009 15th Asia-Pacific Conference on Communications*, pages 597–600. IEEE, 2009.

[26] H. Liu, Q. Fu, L. Du, T. Zhang, G. Yu, S. Han, and D. Zhang. Learning rate perturbation: A generic plugin of learning rate schedule towards flatter local minima. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pages 4234–4238, 2022.

[27] K. Nakamura, B. Derbel, K.-J. Won, and B.-W. Hong. Learning-rate annealing methods for deep neural networks. *Electronics*, 10(16):2029, 2021.

[28] ThorLabs. F260fc-b - 633 nm, f = 15.15 mm, na = 0.16 fc/pc fiber collimation pkg. https://www.thorlabs.com/thorproduct.cfm?partnumber=F260FC-B.

[29] ThorLabs. A397tm-a - f = 11.00 mm, na = 0.30, wd = 8.44 mm, mounted aspheric lens, arc: 350 - 700 nm. https://www.thorlabs.com/thorproduct.cfm?partnumber=A397TM-A.

[30] P. O. Bayguinov, D. M. Oakley, C.-C. Shih, D. J. Geanon, M. S. Joens, and J. A. Fitzpatrick. Modern laser scanning confocal microscopy. *Current protocols in cytometry*, 85(1):e39, 2018.

[31] M. Toyoshima. Maximum fiber coupling efficiency and optimum beam size in the presence of random angular jitter for free-space laser systems and their applications. *JOSA A*, 23(9):2246–2250, 2006.

A Manual Beam-Walk Algorithm

Here, we describe a traditional algorithm to perform a relatively time-efficient manual beam-walk. We use a setup similar to Figure 4, without the stepper motors attached to the mirror knobs. Also, we attach an oscilloscope to the detector to see the output voltage with respect to time. Note that this algorithm applies for each mirror axis independently i.e. we perform it separately for the pitch and yaw knobs.

1. Rotate Mirror 2's knob in one direction until the oscilloscope reads a voltage of approximately η of its original value. The value of η is specified below.

2. Attempt to "compensate" this drop by rotating Mirror 1's knob. The rotation of Mirror 1's knob should result in a relatively Gaussian distribution, so rotate the knob in the direction that increases the voltage until it reaches a maximum.

3. - If the new maximum is **less** than the original voltage, the original Mirror 2 rotation direction was either incorrect or too far. In this case, repeat Steps 1 and 2 but rotate Mirror 2's knob in the other direction until the voltage is approximately η' ($\eta' > \eta$) of its original value, and compensate with Mirror 1 again. From now on, $\eta = \eta'$. The value(s) of η' are specified below.
   - Otherwise, the original Mirror 2 direction was correct. Repeat steps 1 and 2.

4. Repeat Steps 1 to 3 until $\eta \approx 0.75$.

Initially, $\eta \approx 0.25$. For the sake of time efficiency, η increases from 0.25 to 0.50 to 0.75, after which the algorithm ends. Perform the algorithm separately on the pitch and yaw axes, and repeat until both axes optimizations converge to the same value. This is necessary because, in a realistic setup, there may be coupling between the two axes.

B Gradient Descent Algorithm

In a classical gradient descent, the step size that we take at a point is proportional to the derivative at that point. To be more precise, the step size follows the equation

$$\Delta x = \alpha \cdot \frac{df}{dx} \tag{3}$$

where α is the learning rate. However, using a fixed learning rate has its downsides. In particular, a small learning rate accurately converges to the minimum or maximum but takes many iterations to do so. On the other hand, a large learning rate may result in the algorithm overshooting the minimum or maximum. For this reason, traditional gradient descent algorithms in machine learning often run multiple trials to determine the optimal learning rate [26], which can be inefficient.

To reduce the time of automation, we use a process similar to "learning rate annealing" [27], which utilizes a dynamic learning rate. In particular, if the algorithm overshoots the maximum, it returns to the previous step and reduces the learning rate by a factor of 2.

Refer to Figure 15 for a general algorithm outline. The algorithm allows the user to specify the initial learning rate, maximum number of iterations, and tolerance. The tolerance establishes a condition for stopping: if the percent increase of the measured value is less than the tolerance, then the simulation stops. By allowing the user to specify these parameters, the automated system allows for a trade-off between coupling efficiency and time of automation. For instance, for a greater coupling efficiency, the user could input a smaller learning rate, higher maximum number of iterations, and a lower tolerance.

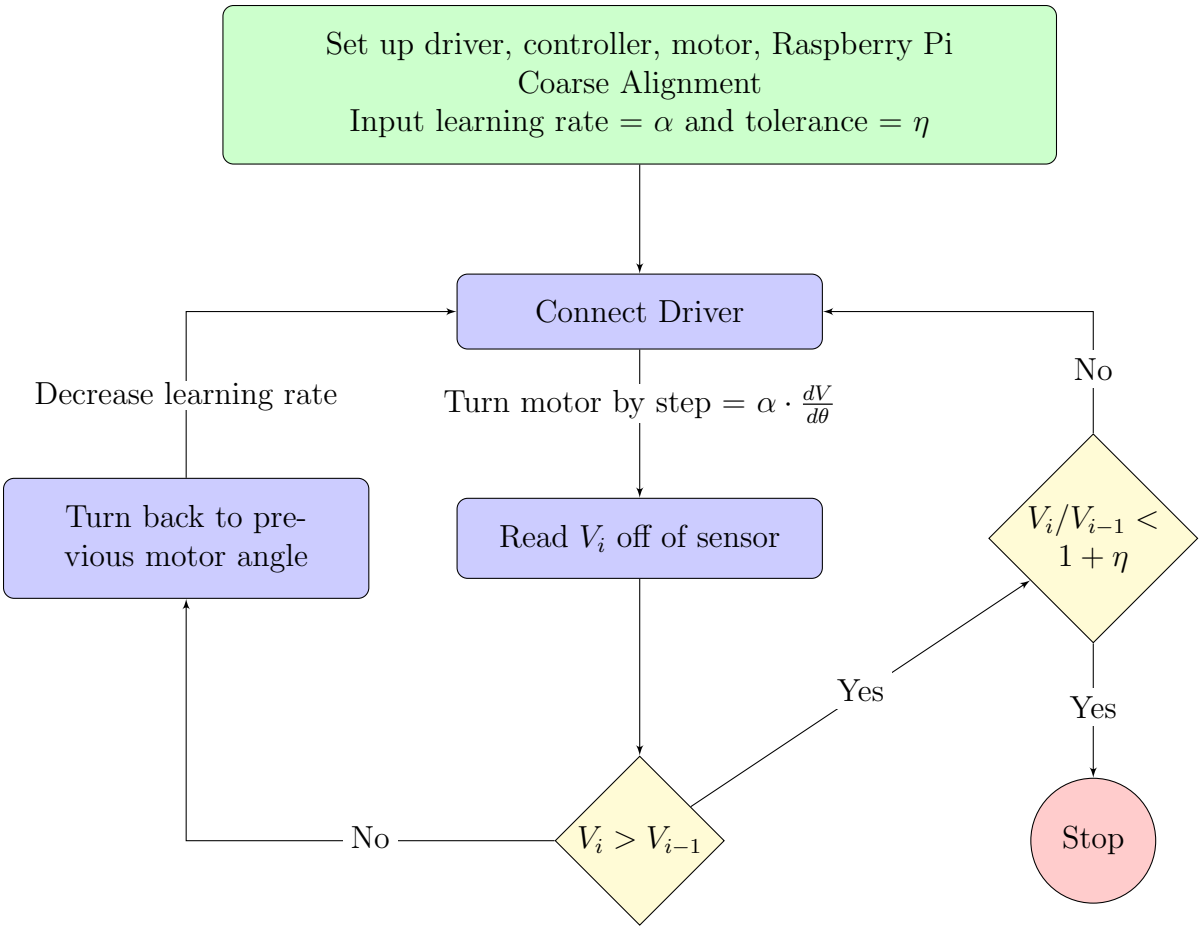

Figure 15: Automated gradient descent algorithm with dynamic learning rate. V_i refers to the measured voltage during the i-th iteration of the algorithm.

Of particular importance in Figure 15 is the calculation of $\frac{dV}{d\theta}$ at each step. Since it is impossible to move the motors by an infinitesimal angle and measure an infinitesimal change in voltage, we must use the fact that

$$\frac{dV}{d\theta} \approx \frac{V(\theta + \delta\theta) - V(\theta)}{\delta\theta} \tag{4}$$

for sufficiently small $\delta\theta$. Here, $\delta\theta$ is limited by the precision of the power voltage sensor. For less precise equipment, $\delta\theta$ will need to be larger in order to detect an accurate change in the voltage. Therefore, the algorithm also allows the user to input a gradient step $\delta\theta$. However,

this also means that the error of our algorithm in finding the angle with maximum coupling efficiency is on the order of $\delta\theta$. However, with small $\delta\theta$, this error can be easily mitigated by performing a small, brute force sweep at the end of our algorithm.

In the single-knob case, we use $\delta\theta = 25$ steps $\approx 4.4°$ to be able to detect a change in the voltage readings. Note that this also gives us an estimate on what the starting learning rate should be. An order of magnitude estimate tells us that

$$\Delta x \sim 100 \text{ steps}, \;\; \Delta V \sim 0.1 \text{ V}, \;\; \delta\theta \sim 25 \text{ steps} \implies \;\; \alpha \sim \frac{\Delta x}{\frac{\Delta V}{\delta\theta}} \sim 25000 \tag{5}$$

In the single-axis (double-knob) case, we use $\delta\theta = 50$ steps $\approx 8.8°$. Note that this is because, along the primary gradient in the 3D map, the envelope containing the voltage peaks (Figure 10) is much wider than the Gaussian distribution in the single-knob case (Figure 7), so we need a slightly larger $\delta\theta$ to detect a change. At the same time, the detected change in voltage will also decrease. Another estimate gives us

$$\Delta x \sim 100 \text{ steps}, \;\; \Delta V \sim 0.01 \text{ V}, \;\; \delta\theta \sim 50 \text{ steps} \implies \;\; \alpha \sim \frac{\Delta x}{\frac{\Delta V}{\delta\theta}} \sim 500000 \tag{6}$$

C Zemax Setup

Since we only want to find a general shape of the 3D mapping between Voltage and Mirror 1 and 2 Pitch Angles, we use a more general Zemax setup with several specifics from our physical setup. The Zemax setup is shown in Figure 16.

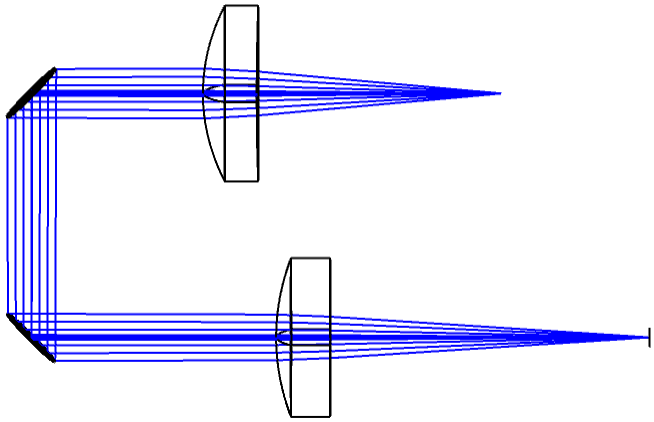

Figure 16: Basic beam-walking setup in Zemax. Note that the point at the top is the laser source, and the small line at the bottom is the detector. The distance between Lens 1 and the detector is optimized using the focus optimization function.

The collimating lenses are specific to our physcal setup. Lens 1 (the bottom lens) is ThorLabs F260FC-B [28], and Lens 2 (the top lens) is ThorLabs A397TM-A [29].

To sweep over all mirror angles, we continuously change the `TiltDecenterData` for both of the mirrors. To calculate the coupling efficiency, we use the code:

```
TheSystem.MFE.GetOperandValue(ZOSAPIEditors.MFE.MeritOperandType.POPD)
```

D Lower Order Optical Effects

Here, we propose an explanation for the discontinuities in the linear plot in Figure 11.

We focus first on the kink at Mirror 1 Pitch ≈ 1700 steps. Zooming into the 3D mapping (Figure 10) at around this region, we obtain the plot shown in Figure 17.

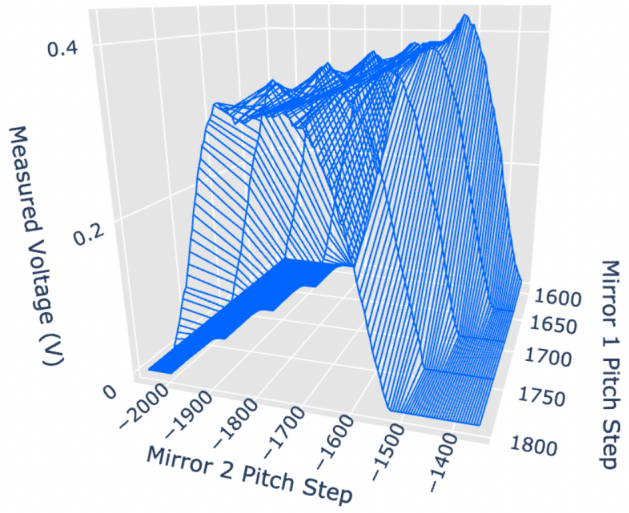

Figure 17: Zoom in of 3D mapping at around Mirror 1 Pitch ≈ 1700 steps.

As seen above, the discontinuity is due to the emergence of a second ridge that "overtakes" the original ridge. This results in an additional angle that the Mirror 2 pitch must turn before reaching the maximum, which in turn results in a discontinuity in the linear plot in Figure 11.

The second ridge can be a result of various lower-order optical effects, one of which is the airy disk effect [30, 31]. In particular, the intensity distribution of the laser is not a perfect Gaussian, but instead has additional maxima further from the center of the beam. Therefore, as the angular deviations of the mirrors increase, the intensity of the airy disk maxima may overtake the intensity of the original Gaussian.