Measuring Light Distribution of LED Sources for Hyper-Kamiokande Detector Calibration

by

Nicholas A. Booth
B.Sc., Colorado School of Mines, 2018
M.Sc., Colorado School of Mines, 2019

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Physics and Astronomy

Measuring Light Distribution of LED Sources for Hyper-Kamiokande Detector Calibration

by

Nicholas A. Booth
B.Sc., Colorado School of Mines, 2018
M.Sc., Colorado School of Mines, 2019

Supervisory Committee

Dr. D. Karlen, Co-Supervisor
(Department of Physics and Astronomy)

Dr. M. Hartz, Co-Supervisor
(Department of Physics and Astronomy)

ABSTRACT

The Hyper-Kamiokande experiment uses water-Cherenkov detectors to study neutrino oscillation and charge conjugation parity (CP) violations with high precision. To reduce systematic errors, multi-photomultiplier tube modules (mPMTs) that comprise the Cherenkov photon detectors will include LED light sources for calibration purposes. This report describes how the LEDs are incorporated into mPMT design, how the distribution of LED light is measured, and discusses how the LEDs will be used for the experiment calibration upon detector completion.

Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

I would like to thank:

My partner Ms. Kaira Murphy and my cat Ajani for providing love and support throughout my degree.

Professors Mark Hartz and Dean Karlen for their mentoring, support, understanding, and patience.

My parents, David and Sue Booth, without whom none of this would have been possible.

The Natural Science and Engineering Research Council of Canada for supporting my research.

The government of Canada for permitting me to study in this incredible country.

Saruman believes it is only a great power that can hold evil in check. But that is not what I have found. I've found it is the small things, everyday deeds of ordinary folk that keeps the darkness at bay. Simple acts of kindness and love.

- Gandalf

Peter Jackson

DEDICATION

I dedicate this paper to Professor Patrick Kohl of the Colorado School of Mines, who helped guide and foster my love of physics. Thank you for inspiring me to pursue my dreams, and I hope to follow in your footsteps and inspire others in the same way.

Chapter 1

Introduction

1.1 Motivation

The Hyper-Kamiokande (HyperK) Water Cherenkov detector will be buried and sealed underground, with photomultiplier tubes (PMTs) lining the interior of the water tank to measure Cherenkov radiation. In order to create radiation to calibrate the PMTs, light sources must be included within the tank before installation. For HyperK and the Intermediate Water Cherenkov Detector (IWCD), this is achieved by incorporating LEDs that emit into the tank via a light pipe within the multi-PMT modules (mPMTs) housing for the detectors. Photogrammetry calibration will extract a 3 dimensional model from photographs taken by multiple cameras, and so 6 beacon LEDs emit from fixed locations around the edge of the mPMT lid. These beacons can be turned on or off so that mPMTs across the tank can measure and subsequently locate the source of the emitted light. For timing calibration, a pulsed LED is mounted to the centre of the mPMT lid and emits a set number of photons over a known length of time. This pulse is then measured by various mPMTs throughout the tank, and the arrival times will be used to calibrate the timing offsets from the passive electronics.

This paper details the research and measurement of a selection of light pipes, the process by which the measurements are converted to a probability density function modeling the emission, and a recommendation from the available options with justification.

1.2 Claims

This thesis presents an approach of measuring light from a distance using a camera and an intermediate target plane (referred to as the "target"). The light pipe shines towards the target such that the pattern is visible by the camera on the target's opposite side. The camera then measures the intensity of this pattern, assigning a brightness value to each pixel of the photograph. The measured intensity profiles are fed into an algorithm that calculates the target's emitted intensity from the camera's measured intensity and converts the pixel images into a light intensity distribution in the polar coordinates θ and ϕ. To obtain multiple trials for a given light pipe, the emitter is placed at distances further and further away from the target plane, while maintaining the assumption that the emitter can be approximated as a point source.

The decided-upon light pipe must meet the following criteria, as required by the hardware, photogrammetry, and timing calibrations.

1. A 10^6 photon pulse from the LED source is distributed such that counts are kept within the PMT dynamic range to prevent over-saturation;

2. The pulsed LEDs must be visible by as many PMTs as possible to calibrate timing offsets caused by passive electronics;

3. The distributed intensity must be as smooth as possible to determine the position of the emission from the measured counts.

To test for these conditions, the light distributions are used to generate simulated photons inside the detector tank and display the resulting intensity distribution on the inner wall of the tank where the mPMTs will be located.

From the available light pipe options, this thesis determines that there is no significant difference within the purview of this analysis. While there are logistical constraints to consider, all options had similar light pulse area, saturation levels, angular distribution, and intensity distribution for the IWCD.

1.3 Outline of Thesis

Chapter 2 describes the history of neutrino science, the purpose of the Hyper-Kamiokande experiment, the goals it hopes to achieve, and the calibration requirements to minimize systematic uncertainty.

Chapter 3 details the process by which the light distribution is calculated from camera images, including the geometry of the set up, the calibration process to relate pixels to physical space, and the algorithm to extract angular distribution data from a single data trial image.

Chapter 4 lists the measurement procedure with associated corrections and provides examples of the raw data collected. The chapter continues by explaining how the raw image data is converted to a histogram, presents combined histograms of all distance trials, and generates the probability density functions from the combined histograms.

Chapter 5 includes the simulation of the IWCD tank, the Monte Carlo used to generate photon trajectories, and the resulting heatmap and intensity distribution for each light pipe. The chapter concludes by recommending the Bivar SMFLP12.0 for use in HyperK.

Chapter 6 summarizes the claims and results of the dissertation. It also provides context for how the algorithm will be used for various calibrations in multiple facets of the HyperK experiment.

Chapter 2

Neutrino Oscillation and Hyper-Kamiokande

Over the 20th century, the definition of an elementary particle changed drastically. As detection technology and techniques advanced, particles previously thought to be fundamental building blocks of matter were found to be made up of smaller particles themselves. The Standard Model, developed over the latter half of the century, classifies elementary particles depending on their interactions with other particles. Quarks are affected by all three major forces - the strong force, the weak force, and electromagnetism. Leptons, however, do not interact with the strong force. Charged leptons (like electrons) interact with electromagnetism, but uncharged leptons (like neutrinos) only interact with other matter via the weak force. As a result, neutrinos are particularly elusive, since most neutrinos never interact with the weak force in a sufficient capacity to be observed. Therefore, to observe a neutrino, the opportunities for interaction must be incredibly frequent. The Hyper-Kamiokande detector, with its stronger beam and larger size, intends to observe neutrino interactions at a rate 20 times higher than previous Kamiokande experiments. A larger sample size will decrease the statistical error in the analysis of the data, allowing researchers to more accurately study phenomenon such as neutrino oscillations and CP violations that could help explain the matter-antimatter imbalance in the universe.

2.1 Neutrino Oscillations

Many neutrinos that pass through Earth are products of reactinos in the sun or in the upper atmosphere. Solar neutrinos make up a majority of the flux through the Earth, and this flux was initially calculated from the Standard Solar Model (SSM). In the 1960s, as detector technology advanced, the Homestake experiment [2] made the first measurement of electron neutrinos emitted from the core of the Sun, but only detected a fraction of the number of neutrinos the SSM predicted. This prompted decades of investigation, bringing both the experimental process of Homestake and the SSM in general into question, but none of the proposed problems were found to cause the discrepancy. The issue lied in the complexity of neutrinos - not only was the atmospheric component of the neutrino flux being ignored, but also the Sun only *emits* electron neutrinos. There are two other flavours of neutrinos, and it is possible for a neutrino to oscillate between the three possibilities.

The flavour of a neutrino is defined by the flavour of the associated charged lepton involved in its production. For example, a muon neutrino ν_μ is emitted when a π^+ particle decays into a muon μ^+ (Eqn 2.1), whereas an electron neutrino ν_e is emitted (with a muon antineutrino $\overline{\nu}_\mu$) when a muon decays into an electron e^+ (Eqn 2.2) [5].

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \tag{2.1}$$

$$\mu^+ \rightarrow e^+ + \overline{\nu}_\mu + \nu_e \tag{2.2}$$

Each neutrino flavour is a different mix of neutrino mass eigenstates. Following its production, the mixture changes due to the mass differences. As a result, the eventual interaction of the neutrino may produce a charged lepton that differs in flavour from the charged lepton involved in the production of the neutrino. From a simplified viewpoint, it would appear that the flavour of the neutrino changed, and the process is referred to as neutrino oscillation.

Discovering this experimentally took almost 50 years. The original Kamiokande detector, built in the 1980s to search for proton decay, saw a deficit between measured atmospheric neutrinos and predicted values [4]. Super-Kamiokane (Super-K) was built in the 1990s to further explore the discrepancies found by Kamiokande, specifically probing the dependency of the neutrino deficit as a function of zenith angle [5]. The experiment found that the inconsistencies was consistent with oscillations between mass states $5\text{x}10^{-4} < \Delta m^2 < 6\text{x}10^{-3} eV^2$. Working under this new

assumption of neutrino oscillation, the Sudbury Neutrino Observatory (SNO) utilized three different detection techniques to study solar neutrinos depending on their flavour [1]. For an energy threshold of 5 MeV, the neutral current method not only measured a total solar flux of $5.09\text{x}10^6\text{cm}^{-2}\text{s}^{-1}$, which agreed with solar models, but also the electron neutrino component ($\phi_e = 1.76\text{x}10^6\text{cm}^{-2}\text{s}^{-1}$) and non-electron neutrino component ($\phi_{\mu\tau} = 3.41\text{x}10^6\text{cm}^{-2}\text{s}^{-1}$) of the total flux. The discrepancies of the Homestake experiment had been answered; the chlorine detector used in Homestake could only detect electron neutrinos and could not observe the other components of the solar flux.

2.2 The Hyper-Kamiokande Experiment

To continue probing the nature of neutrino oscillations, the HyperK detector was approved in 2020 as the successor to Super-K and the associated T2K ("Tokai to Kamioka") experiment. Figure 2.1 shows the concept for a new water Cherenkov detector with 8 times the fiducial area and 2.5 times the beam intensity that will increase the rate of data collection twenty-fold as compared to T2K. More observed events will shrink the statistical uncertainty of the detector, allowing the experiment to measure imbalances between neutrino and antineutrino probabilities, known as Charge-Parity (CP) violations.

This measurement requires control over whether neutrinos or antineutrinos are entering the tank. A particle beam is created at the J-PARC accelerator, 295 km from the HyperK detector, and passed through a magnet horn to filter out the undesired pion charge [3]. It is ambiguous to classify mesons like a pion as matter or antimatter, since mesons consist of both a quark and an antiquark, so electric charge is used to specify the type of neutrino that will be produced. Equation 2.1 shows the decay of a positive pion into a positive (antimatter) muon and a muon neutrino (matter). The inverse is also true - a negative pion will decay into a matter muon and a muon antineutrino (antimatter) (Eqn 2.3). Depending on the direction of the current in the magnetic horn, the undesired charge will diverge from the particle beam, allowing a stream of neutrinos or antineutrinos to be directed towards the detector. Comparing the rate of events for each beam will provide insight into the conservation of CP symmetry.

$$\pi^- \rightarrow \mu^- + \overline{\nu}_\mu \tag{2.3}$$

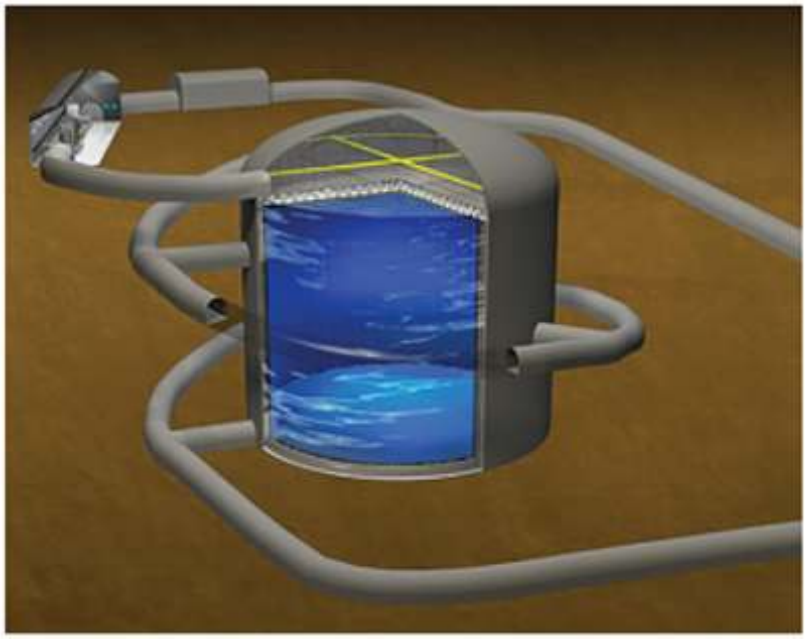


Figure 2.1: Concept Model of Hyper-Kamiokande Tank

However, this comparison requires a strong understanding of neutrino production and detection. CP violations could potentially manifest at extremely small orders of magnitude, so HyperK aims to constrain the uncertainty of the detector to the smallest orders of magnitude possible for the data to be meaningful. The detector will reduce statistical uncertainty by vastly increasing the expected number of measured events through the increased tank size and more powerful beam. Reducing systematic uncertainty, on the other hand, requires an incredibly well-tuned and well-understood model of the detector response to neutrinos. New state-of-the-art photomultiplier tubes (PMTs), designed specifically for the HyperK tank, will line the entire interior surface of the tank and record the intensity and pulse length of photons from the Cherenkov light showers produced by particle interactions. With this information and the known locations of the triggered PMTs, the detector can reconstruct the interaction vertex and other properties of neutrino events.

2.3 Detector Hardware and Calibration

Because the PMTs will be submerged in water within the HyperK tank, they must be sealed within a housing unit to prevent water contamination. Each unit, known as multiphotomultiplier tubes (mPMTs), protects nineteen PMT detectors mounted

on a plastic spherical cap beneath a domed acrylic lid (Figure 2.3). The detectors are powered and controlled by an electronics board mounted to the steel base of the mPMT, with cabling exiting the unit via the base into the non-submerged support structure of the tank to communicate with the primary control computer.

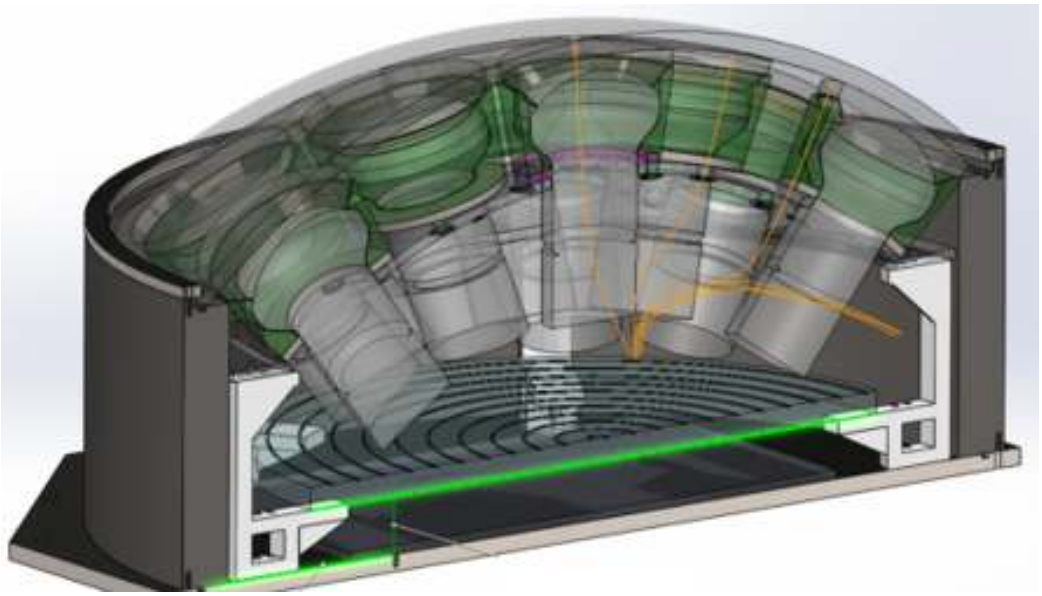


Figure 2.2: Cutaway View of multiPhotomultiplier Tube Interior. Example Light pipe placement highlighted in orange.

To measure mPMT orientations and estimate timing offsets, calibration LEDs are included inside the sealed housing unit, drawing power and commands from the mPMT electronics board. These LEDs are mounted directly onto the board at the base of the unit, and the light is transported from the board to the acrylic dome by pipes of glass and shielding, called light guides or light pipes. Figure 2.2 shows these pipes as orange lines interwoven through the PMTs, and Figure 2.3 shows the terminations of these pipes where the light will be emitted into the tank. Each mPMT will have 7 calibration LEDs - 6 beacon LEDs (represented in red in Figure 2.3) that have binary on/off states and a pulsed LED (represent in gray in Figure 2.3) that emits a known, timed light pulse. The photons from these LEDs will be captured by other PMTs and separate, isolated cameras in the tank to calibrate the photogrammetry and timing offsets.

Figure 2.3: Birdseye View of multiPhotomultiplier Tube. Red dots denote planned beacon light pipes, Black dots denote omitted light pipes, Grey dot denotes pulsed light pipe.

The purpose of this thesis is to quantify the light distribution that is output of the various light pipes available, simulate the projection of light onto the other PMTs in the tank, and recommend a specific light guide for use in the HyperK mPMT modules based on the simulation results.

2.4 Light Pipes to be Studied

Three possible light pipes are under consideration, an example of which is shown in Figure 2.4. These options were initially selected based upon availability and cost, but the angular distribution of each must be measured before a final decision can be made. Distinctions between each pipe is shown in Table 2.1. Fibre Thickness refers to the diameter of the optical fibre inside the light pipe, while Housing Thickness

refers to the total diameter of the light pipe, shielding included.

Label	Manufacturer	Product Number	Fibre Thickness	Housing Thickness
Pipe A	Bivar Inc.	SMFLP12.0	1 mm	2.20 mm
Pipe B	Bivar Inc.	SMTV01-B11	2 mm	3 mm
Pipe C	Visual Communications Co.	FLX_12	1.5 mm	2.20 mm

Table 2.1: Identification and Thickness of Light Pipe Options

The calibration discussed in Section 2.3 will be most successful when as many mPMTs as possible see the same calibration light. Chapter 4 explains the procedure by which each pipe's angular distribution is determined and presents the results for all options. Chapter 5 details the light distribution and intensity profile of each light pipe, determined by a simulation of the IWCD tank, and refers a single pipe for use.

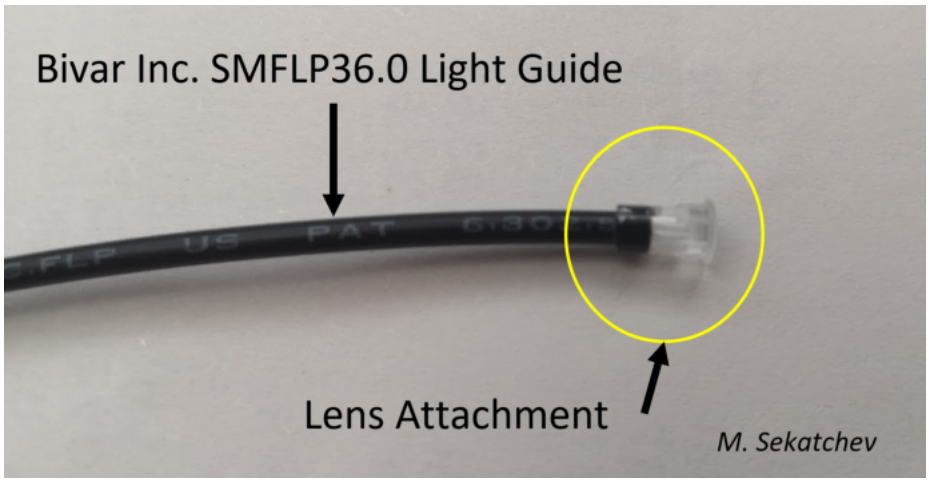


Figure 2.4: Example Light Pipe SMFLP36.0 (36" variant of Pipe A)

Chapter 3

Calibration and Comparison of Potential Lightpipes

The first step of any calibration process is to define the input values that will be used to standardize what the instrument observes. The calibration LEDs, both beacon and pulsed, emit photons according to a certain angular distribution unique to the hardware in question. If well defined, this distribution can then be used to predict what an mPMT will observe under the same conditions. Not only does this provide a baseline for individual instrument calibrations, but the distributions of various lightpipe options will be compared to determine which part under investigation best fits the needs of HyperK.

Unfortunately, the required information is not provided by the manufacturers of the optical fibres, and so the model must be built from scratch. The solution provided below is capable of extracting the angular distribution of a light source from images of the light shining on a translucent 2D surface. From this, the model can be expressed in Cartesian or cylindrical coordinates depending on the requirements of the experiment.

3.1 Cartesian Projections of Cylindrical Light

A photon emitted from a lightpipe has two relevant directional quantities - the azimuthal angle ϕ and the polar angle θ. Together, these two angles define the direction of an emission vector of arbitrary length (Figure 3.1). By placing an infinite target plane orthogonal to the z-axis at some set distance d away from the light source, an

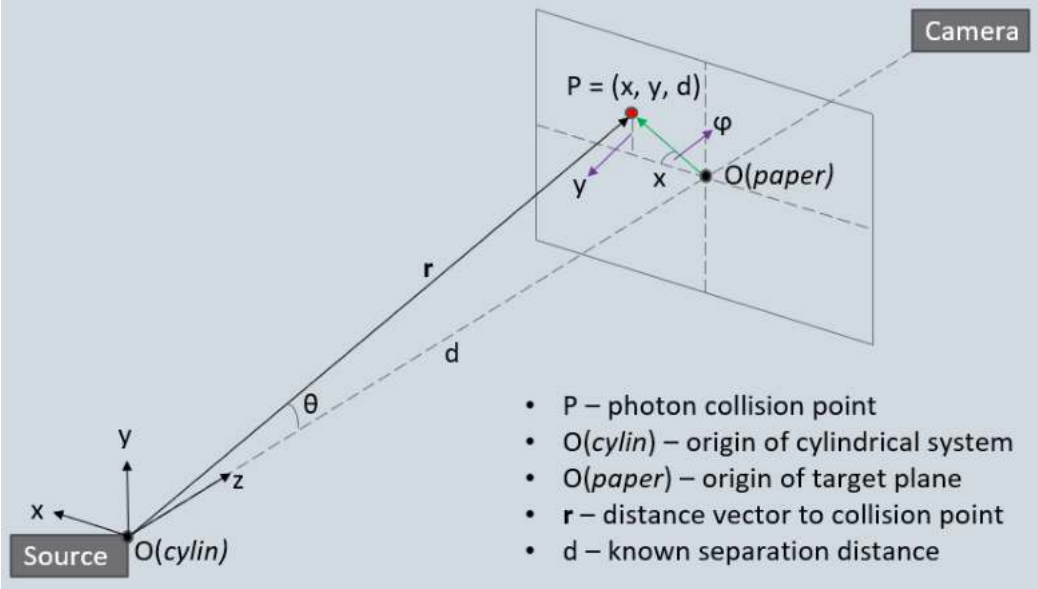


Figure 3.1: Line Drawing of Experimental Setup

infinitesimal slice of the light distribution can be observed. This flat projection of the emitted light allows the coordinates of a collision point P to be defined in terms of both the target plane and the emission source. As the same point is being described by each, a direct relationship can be drawn between the Cartesian coordinates (defined by the plane's origin O_{paper}) and the cylindrical coordinates (defined by the source origin O_{cylin}).

From this plan, the actual experimental procedure was developed across several months, improving accuracy and consistency over time. An initial proof of concept

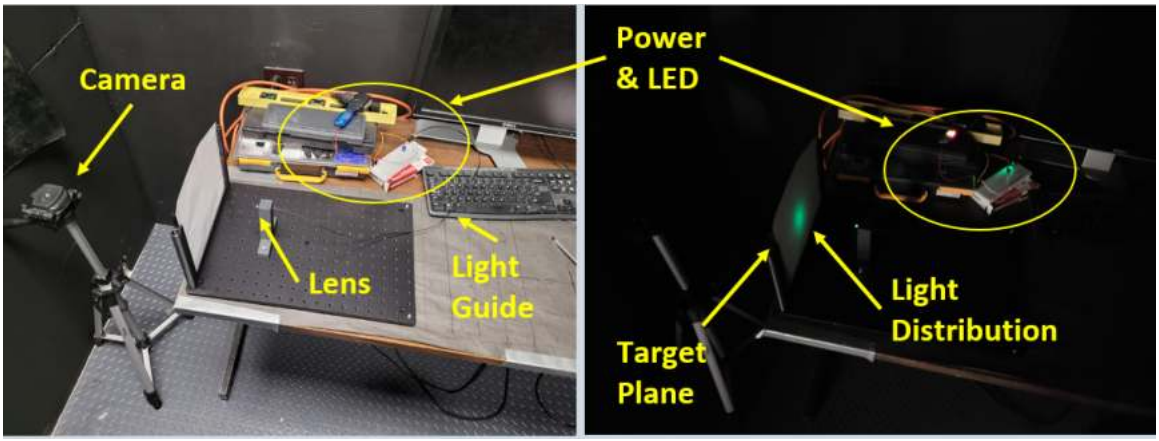


Figure 3.2: Experimental Setup and Example Data Trial

setup was developed using easily accessible supplies, but parts were slowly upgraded to 3D printed parts as the exact dimensions for the measurements were determined from the proof of concept procedure and data collection. The final set up, shown in Figure 3.2, mounted a frame for the target plane and a stand for the light guide onto a 12" by 9" optical table to ensure consistent distance trials for each iteration of the experiment.

The fundamental assumption while collecting and analyzing data is that the angular distribution of the light source is independent of the z axis. While the size of the beam spot widens as distance d is increased, the z coordinate of a collision P defines only where the photon *stops*, not how it is emitted. Therefore, a normalized distribution obtained from a data trial with distance d should be nearly identical to a normalized distribution obtained from a distance of $2d$ or any other distance.

Collecting images at several known distances between the light source and the target plane will model the light density across multiple independent data trials. This enables experimental error to be carried over when a single normalized density function for the source in question is extracted by combining the data trials.

3.1.1 Capturing the Beam Spot

Previously, an attempt was made to estimate the beam angle of a lightpipe by shining the LED directly into the camera aperture to directly count photons per pixel. However, a typical camera aperture is not large enough to effectively capture a light distribution at distances relevant to HyperK. As discussed in Chapter 2, the detectors exposed to this calibration light will be across the tank, far away from the emitter, where the light source can be approximated as a point source. Because the camera's aperture is not able to capture the beam spot by itself, the target plane captures the light pulse on a much larger 2 dimensional surface which is subsequently imaged by the camera.

Inserting a new element between the source and the measurement tool requires its own assumptions, however. The plane is considered to be uniformly diffuse; light from any angle is scattered isotropically towards the camera, maintaining the intensity profile of the light source. In addition, because the light is not being shined directly on the photosensors of the camera, a lens is needed for the camera to image the target plane. This lens distorts the light from the target plane to focus the image onto the photosensors, a process that must be reversed during data analysis in order

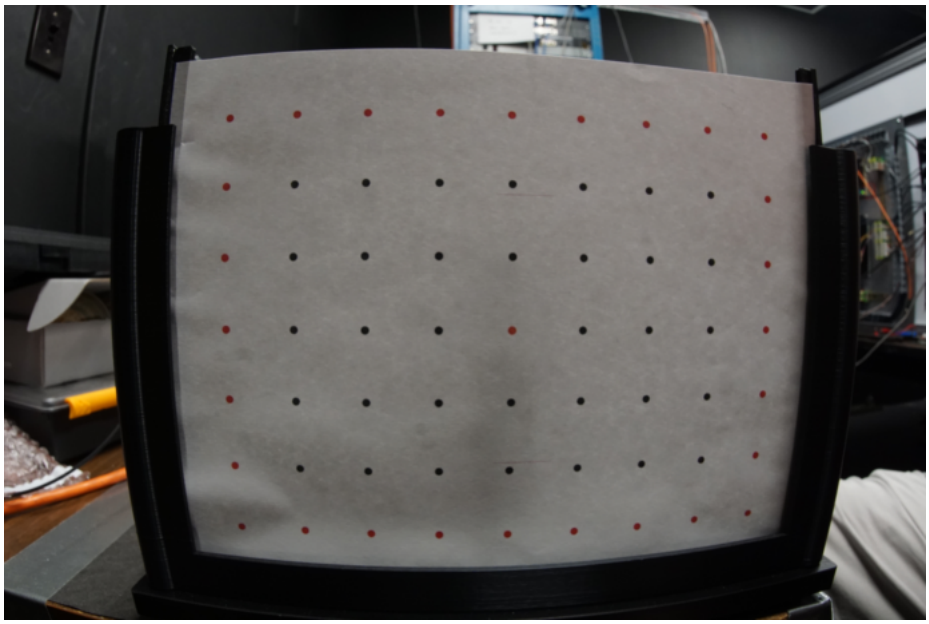


Figure 3.3: Calibration Screen

to measure the correct location of a photon on the target plane.

3.1.2 Defining the Pixel Coordinate System

The relevant information collected by an image is the Cartesian coordinates of any given point on the target plane. Combined with the known separation distance d (Figure 3.1), these Cartesian coordinates can be easily converted to ϕ and θ via trigonometry. However, cameras do not collect photons with a typical x-y Cartesian grid, but rather with an i-j pixel grid defined by the camera's hardware. Therefore, in order to express a given location in terms of ϕ and θ, the relationship between pixel number i and j and Cartesian coordinate x and y must be fully defined for the entire target plane.

Figure 3.3 is an example of a calibration screen used to define this relationship. The grid of uniformly spaced dots represents the Cartesian coordinate system, with each dot located at a known x-y position from the central red dot that defines the origin. When imaged by the camera, the light from each of these dots is distorted by the attached lens and then captured in a specific pixel of the photoreceptor. Because of the distortion, the relationship between pixel and Cartesian cannot be linear, and so a robust calibration process is required to calculate the conversion.

The grid in Figure 3.3 is calibrated by finding the pixel coordinates of a dot's

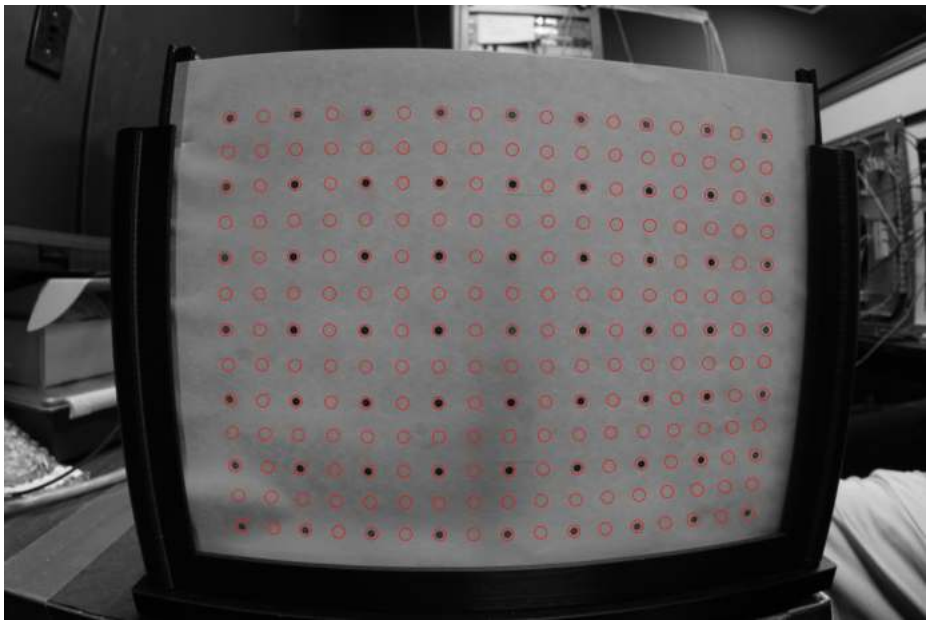


Figure 3.4: Fully Calibrated Pixel Grid. Red dots denote xy positions in between calibration points.

centre, thereby associating a known x-y pair with a known i-j pair and vice versa. This process is done three times:

1. by hand by the user to define the limits of the calibration area;
2. a rough calibration using these limits to place x-y coordinates close to their respective pixel coordinates;
3. a full calibration of the area including the space between the dots.

Each iteration defines a new set of initial conditions, becoming more and more accurate with each pass. First, it is necessary to brute force some x-y and i-j pairings by hand, so that the software has an initial guess to the space in question. The four corners and the origin are chosen for this purpose for convenience. Next, the algorithm locates all circles in the image, including many false positives outside the calibration area. For each Cartesian coordinate pair, the rough calibrator estimates a location based on a linear distance from the origin (defined in the previous iteration) and pairs the nearest found circle to the estimated pixel location. This not only generates a list of points for the next iteration to use as an initial guess, but also eliminates all false positives. Finally, the full calibrator inputs the grid of known x-y / i-j pairings into

a numerical interpolater that generates conversion matrices bounded by the red dots in Figure 3.1.

The result is an extension of the Cartesian grid by placing circles at the locations in between the defined dots, as demonstrated in Figure 3.4. With the generated conversion matrices, the algorithm is able to convert any given x-y Cartesian coordinate pair to its respective i-j pixel coordinate and vice versa.

3.1.3 Applying the Pixel-Cartesian Conversion

Now that each pixel of the image can be translated to an x-y coordinate relative to the centre of the target plane, each pixel can be assigned a (d, θ, ϕ) coordinate, where d is the separation distance between the source and the target plane. The origin defined in Section 3.1.2 is purposefully positioned to be on-axis with the centre of the light source and the centre of the camera lens, creating a fixed z-axis for the cylindrical geometry. As depicted in Figure 3.1, any point P with known Cartesian coordinates is easily defined in cylindrical coordinates through trigonometry explicitly laid out in Section 3.2. In practice, each pixel in the data image has an associated Cartesian coordinate and a number of associated photon counts. By iterating through each pixel in the calibrated area of the data image, these photon counts can be grouped by the θ and ϕ value of their respective pixel to create a distribution histogram.

3.2 The Algorithm

The software developed for this project was written in Python 3.9 and intended for integration with the HyperK code base. The user supplies the necessary images (calibration, background, and data), along with relevant measured quantities such as the distance between the camera and the target plane, the uniform spacing of the calibration screen, and most importantly the initial calibration coordinates of the four corners of the grid of dots. For each experimental trial, the software calibrates the pixel grid, removes background contributions, generates a distribution histogram for each distance tested, and then plots each histogram on top of each other to confirm the fundamental assumption that the angular distributions do not depend on d.

3.2.1 Background Contributions

Before data images can be analyzed, background contributions must be removed from the number of photon counts in each pixel. While great pains were taken to ensure a darkroom for optical research, it is impossible to completely remove all sources of light, especially when other instruments are in the room and have status lights of their own. Even if perfect darkness could be guaranteed, the camera very likely also records its own dark noise pedestal that must be passed to the algorithm. After taking the calibration image but before the light source is turned on, an image of the target plane in darkness is taken to establish the pedestal light contribution of the room (Figure 4.4 in Section 4.1). The algorithm's first action after opening the given set of data images is to convert the provided calibration image to greyscale (see Section 3.2.3) and subsequently save the background intensity within the dataset's Calibration object. This intensity can then be accessed later during data analysis, when the counts in each pixel need to be subtracted from the respective pixel in each data image.

3.2.2 The Calibration Process

The calibration screen depicted in Figure 3.3 provides specific locations that guide the calibration process. Each dot has a spacing of one inch from the other surrounding dots, and so each dot has an easily defined Cartesian coordinate from the centre origin dot. The algorithm's initial goal is to define the pixel coordinates of each dot's centre; by doing so, these pixel coordinates can be used to interpolate into the space between

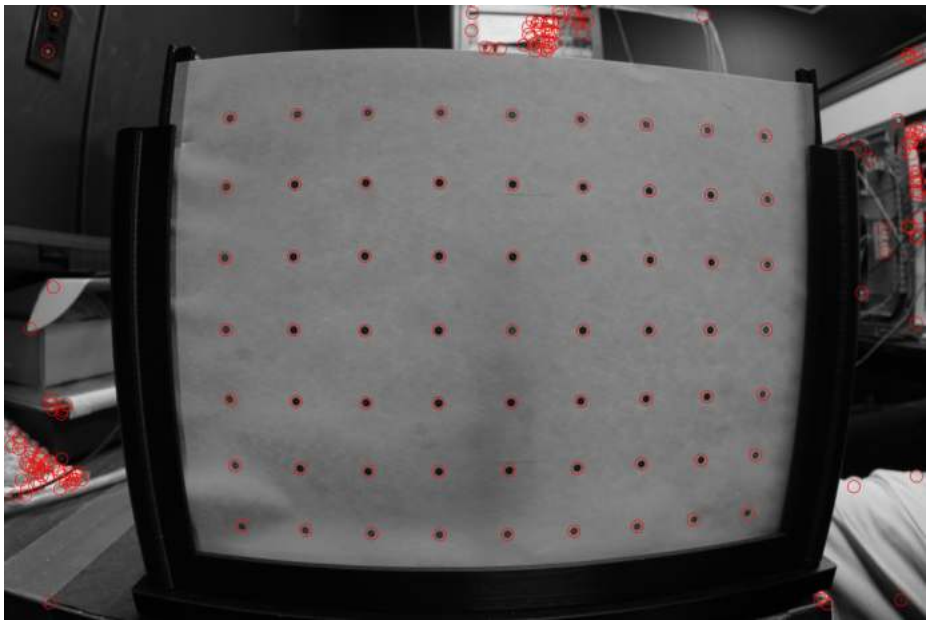


Figure 3.5: All Circles found by the HoughCircles Python Method

the dots.

To do this, the algorithm employs a Python method known as HoughCircles, which scours a given image for similarly-coloured pixels in a circular shape. Since the parameters for "circular" must include all the calibration dots, the method finds plenty of other circular entities as well. Figure 3.5 shows how HoughCircles not only located all the calibration dots, but also several screw heads, the end of a PVC pipe, and several dials on the instrument in the back. Before the algorithm can begin assigning pixel coordinates to the calibration dots, these false positives have to be removed from the analysis by defining the region to be calibrated.

Unfortunately, the algorithm cannot do this by itself. To find the boundary circles, HoughCircles would be required, which will find false positives, which will then need the boundary circles again that are still unknown. This degeneracy is avoided by requiring an independent initial guess from the user. For this research, the calibration images were opened in a basic image editing software (Microsoft Paint) to find the approximate pixel of the centres of the corner dots and origin. For each, the algorithm looks to the given approximate location and assigns the closest HoughCircle to the respective Cartesian location among $(\pm 4, \pm 3)$ and $(0, 0)$. Figure 3.6 illustrates the result of this step, highlighting the circles defined as the corners and origin.

However, the algorithm still needs to assign Cartesian coordinates to the remain-

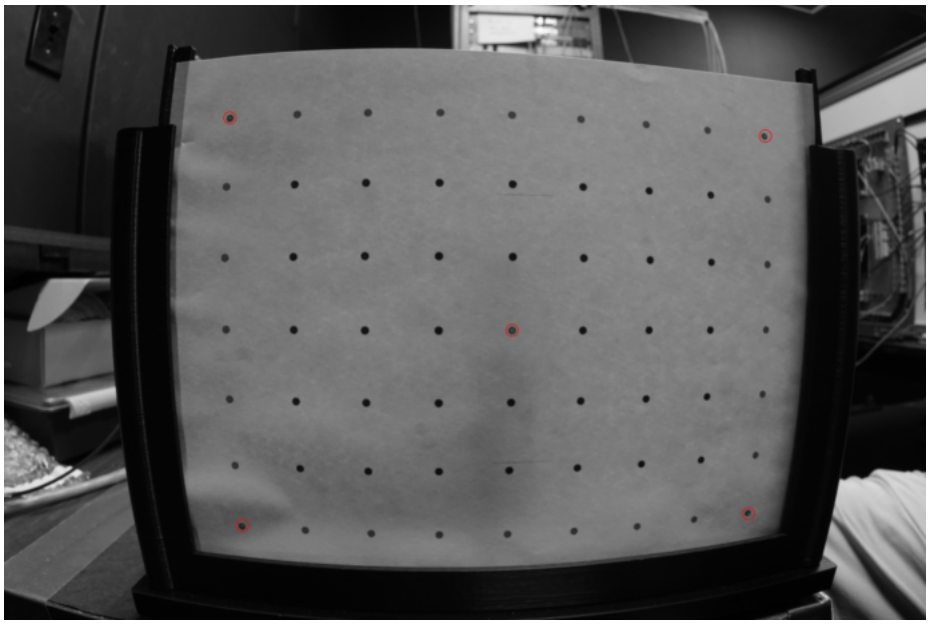
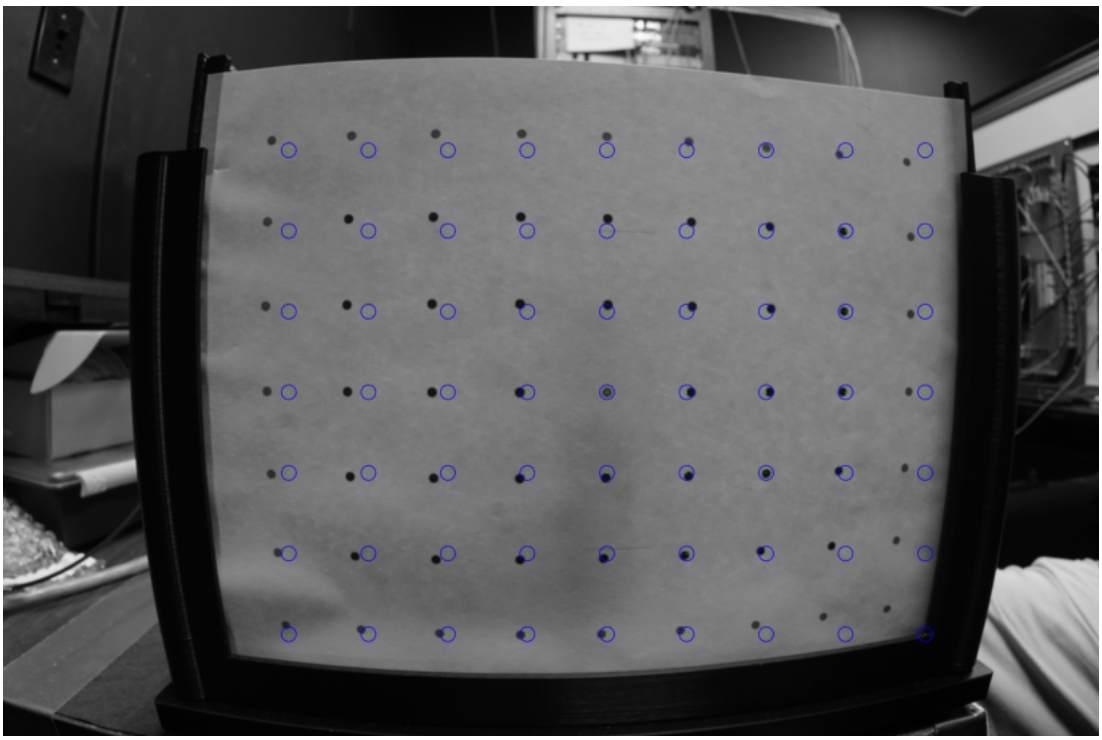


Figure 3.6: Calibration Area defined by Initial Guess - 4 corners and the origin

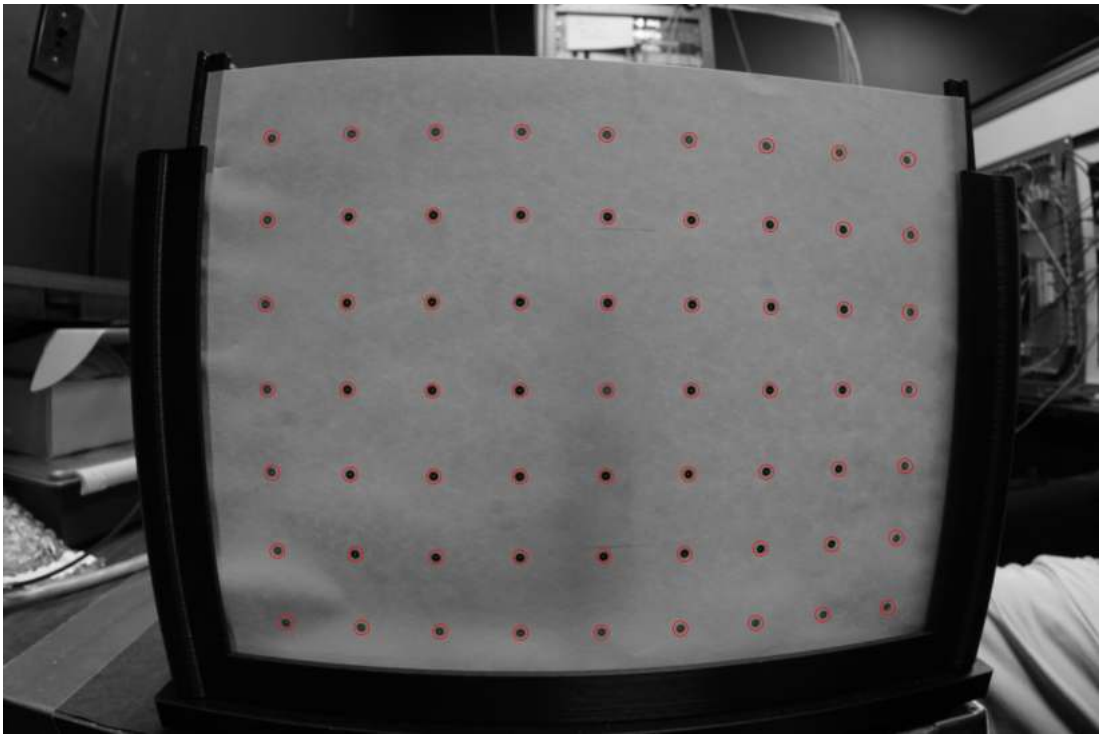
ing calibration dots. Asking for user inputs for these dots defeats the point of having a calibration algorithm, so these dots must be defined from the information already at hand. Because each dot is equally spaced from the others, a simple numerical interpolation is needed to estimate the location of an interior calibration dot. While the imaged $(-4, 0)$ dot is distorted from its actual location four inches to the left of the origin, it is safe to assume the HoughCircle associated with the $(-4, 0)$ dot is the closest to where an undistorted image would expect it to be. As both $(-4, -3)$ and $(-4, 3)$ have known pixel coordinates, the algorithm finds the midpoint of the distance between, finds the nearest HoughCircle, and assigns said circle the respective Cartesian coordinates $(-4, 0)$. This process is then repeated for every other calibration dot.

Figure 3.7 illustrates these steps, as well as the distortion present in the image that might not be clear to the naked eye. Each of the blue circles in Figure 3.7a shows where the Cartesian coordinates are expected to be, given the defined origin and a simple pixel-to-inch ratio. As the dots move further away from the origin, the distortion grows stronger, with significant white space between the far corners and their respective estimates. This implies that the distortion is radial, which is in agreement with the spherical lens used during data collection.

Once the algorithm has all the pixel locations for all the known Cartesian coor-



(a) Grid Location of Dots from simple Linear Interpolation



(b) Calibration Dots aligned with corresponding HoughCircles

Figure 3.7: Preparing for Full Calibration

dinates (Figure 3.7b), a much more accurate interpolation of the entire calibration area is found using SciPy's LinearNDInterpolator. This Python method numerically interpolates between a given array of defined points in a two or greater dimensional space. However, the method is only capable of interpolating a single quantity over the region bounded by the given points. In order to convert freely between pixel and Cartesian, this interpolator is ran four times - x and y each in terms of (i, j), and i and j each in terms of (x, y). The output of each is saved within the dataset's Calibrator object, and intermediary circles are added between the calibration dots to demonstrate the successful interpolation (see Figure 3.4 in Section 3.1.2).

3.2.3 Projections and Data Analysis

Before the analysis can begin, the data image sample must be prepared by the algorithm into the exact form the Analyzer object expects. Because the algorithm only analyzes images one at a time to prevent memory issues, the process listed below is repeated each time a new data image is provided.

1. User input specifies the set of data images and the specific image to be analyzed;
2. Open saved Calibrator object with background pedestal;
3. Open image from specified data set;
4. Convert data image to greyscale and subtract background;
5. Pass filtered data image into Analyzer object.

The data image in question, known as the trial image, must be converted to greyscale for the same reason as the background image in Section 3.2.1 - each pixel needs a single intensity value, not a combination of RGB colour information. Greyscale normalizes and blends the colours of a pixel into a number between 0 and 255, which is proportional to the number of photons that hit said pixel. This transformation is shown in Figures 4.5 and 4.6 in the following chapter.

In addition to the filtered trial image, the Analyzer object requires a calibrator (opened from the saved Calibrator object) and the physical dimensions (in centimetres) of the calibrated space. The Analyzer object uses the provided calibrator and physical dimensions to define the limits of the calibrated region within a copy of the trial image. Pixels inside the calibration area are assigned the filtered counts from

their respective trial image pixel, while the pixels outside the area are set to NaN ("Not a Number") to exclude them from the image data.

Projections in Pixel and Cartesian Space

Unfortunately, the rectangular target plane caused a portion of the collected data to be unusable - beyond a certain polar angle θ_{limit}, rings of constant ϕ are not fully included on the plane. In simpler terms, the collected data is limited by the largest full circle that can be drawn on the target plane (the yellow circle in Figure 3.8). To account for this, the active region is constrained by the "cos(θ) limit", shown by the red lines in Figure 3.8, and only includes the polar angle range that is covered at all azimuthal angles.

Importantly, this limit depends on the centre location of the light pulse, which is not exactly at the origin of the calibrated region. While great pains were taken to ensure the light source, target plane, and camera were nearly on axis, it is unreasonable to assume perfection, and so the centre of the light pulse must be determined independently. To do this, the light pulse is projected onto the horizontal and vertical axes, and a normal Gaussian (Equation 3.1) is fit to these projections.

$$f(x) = a \exp\left(-\frac{(x-b)^2}{2c^2}\right) \tag{3.1}$$

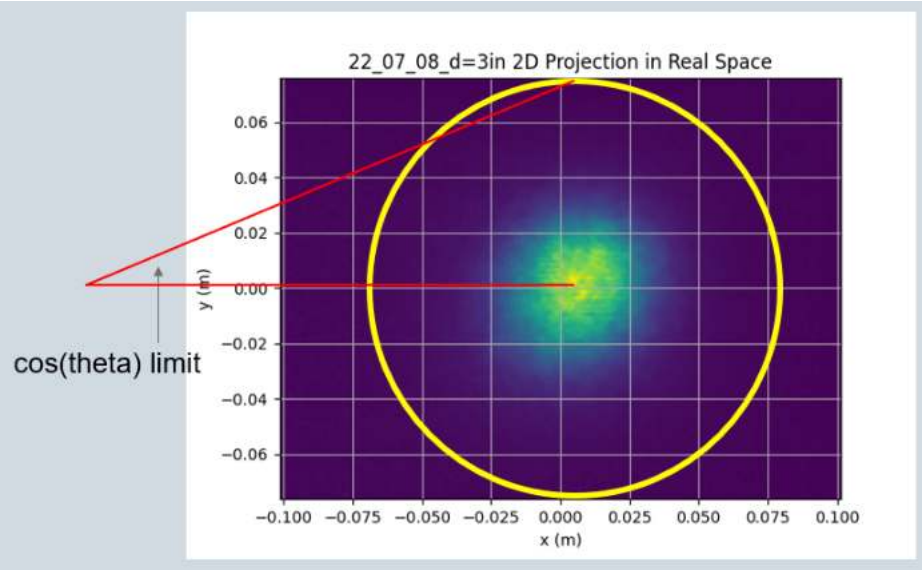


Figure 3.8: Calibrated Data Image in Cartesian Space

The assumption of a normal Gaussian is most valid around the centre of the projection, so the fit is limited to the region defined by the Full Width at Half Maxiumum (FWHM). The algorithm isolates the FWHM by first finding the maximum number of counts on the curve and then iterating from the left and right extremes until a count of more than half of said maximum is found.

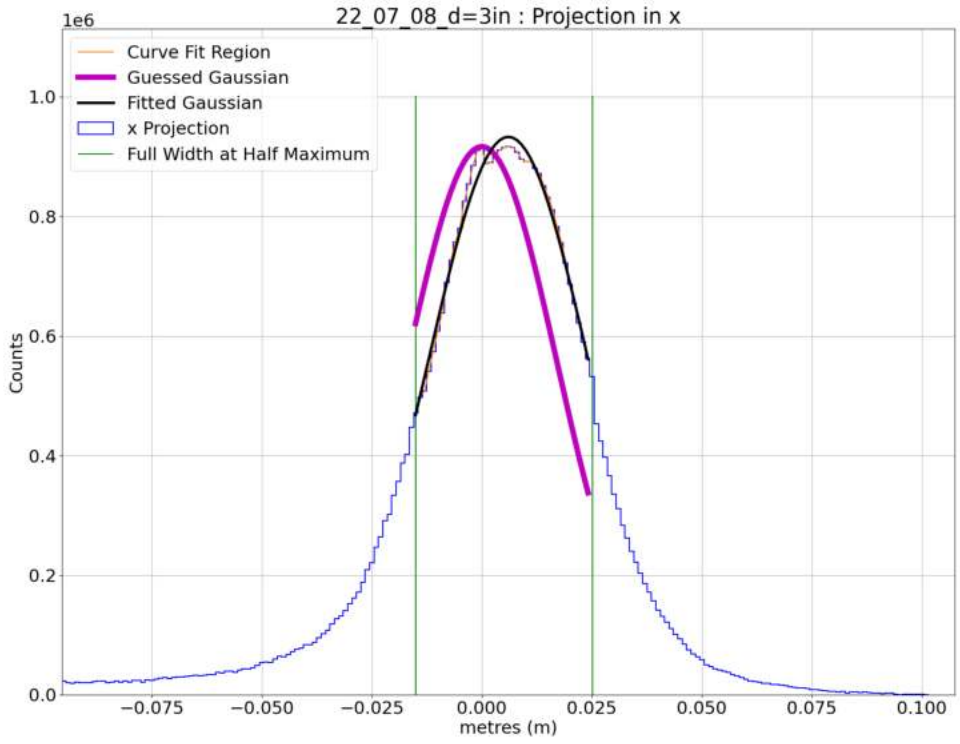


Figure 3.9: Projection of Light Distribution onto Horizontal Axis

The algorithm employs Python's curvefit() method to fit the projection to Equation 3.1, but an initial guess is required for the method to begin the numerical approximation. An assumption of the constants in Equation 3.1 - a, b, and c - must therefore be supplied from the projection curve. For the amplitude a, the previously defined maximum value is used, as it is assumed to be located at the peak of the curve. Parameter b represents the expectation value of the distribution, or the horizontal location of the amplitude a. Using the previous assumption that the maximum value is the amplitude, this location is simply be the midpoint of the defined FWHM region. Finally, the standard deviation σ (represented as parameter c) is found by dividing

the FWHM by 2.355, a known property of Gaussian distributions. These estimated parameters are then fed into the curvefit() method, alongside the projection data, to ascertain the fitted value for b, the centre of the Gaussian peak of the projection.

Each step of this Gaussian fit is plotted for an example trial image's horizontal projection in Figure 3.9. The vertical green lines are the FWHM boundaries found by iterating from the left and right, respectively. The blue curve shows the data excluded from the fit, while the data in orange is what is included. The purple line shows the Gaussian as defined by the initial guess parameters, while the black line shows the fitted Gaussian using the parameters output by curvefit().

Figure 3.10 combines the horizontal, vertical, and two-dimensional projections to illustrate how the Gaussian peaks translate to the centre of the light pulse. As a

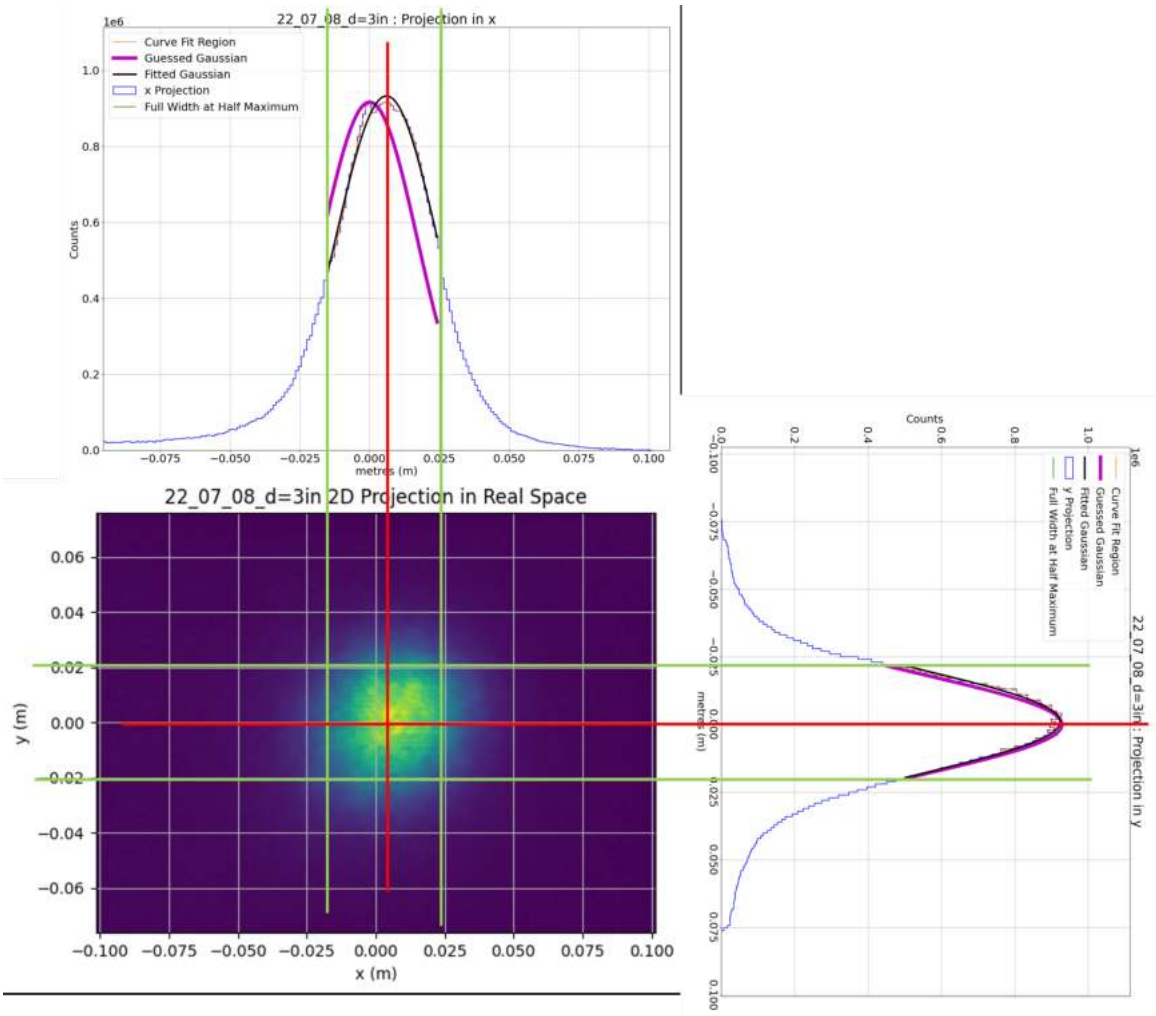


Figure 3.10: Horizontal and Vertical Projections of $d = 3$in Trial Image

troubleshooting measure, the fit procedure is done for ij as well as xy to test the calibrator. If the xy conversion of the ij coordinate pair is not within tolerance of the xy coordinate pair, the algorithm displays a message to alert the user.

Extraction of Azimuthal and Polar Distributions

With the centre of the light pulse associated with a Cartesian coordinate pair, the coordinate system of the target plane (centred on the origin of the target plane) can be replaced with a coordinate system of the light pulse (centred on the found Cartesian coordinates). The image's pixel matrix is then iterated through once more, calculating the $cos(\theta)$ and ϕ values for each pixel.

The relevant vector is the $\mathbf{r}$ distance vector in Figure 3.1, originating at source $(0, 0, 0)$ and impacting the target plane at (x, y, d). Now that x and y are known in addition to d for each pixel, this Cartesian coordinate pair for the collision point can be expressed in cylindrical coordinates (r, θ, z).

$$\begin{aligned} r &= \sqrt{x^2 + y^2} \\ \phi &= tan^{-1}\left(\frac{y}{x}\right) \end{aligned} \tag{3.2}$$

$$\begin{aligned} |\mathbf{r}_{source}| &= \sqrt{r^2 + d^2} \\ cos(\theta) &= \frac{d}{|\mathbf{r}_{source}|} \end{aligned} \tag{3.3}$$

While z remains unchanged, the x and y components are translated to r and ϕ via Equation 3.2. Finally, the azimuthal angle is calculated from the magnitude of the $\mathbf{r}$ distance vector and the known separation distance d (Equation 3.3).

Chapter 4

Measurements

As referenced in Section 3.1.3, the experimental setup was developed across several months, improving accuracy and consistency over time. The light pipe in question is connected to an LED and mounted a set distance away from the target plane. The light distribution is then captured by the camera on the opposite side of the target plane before the mount is moved to a new set distance further away. A visual of the final setup used is provided again here for convenience, and each aspect of the design is discussed in length below.

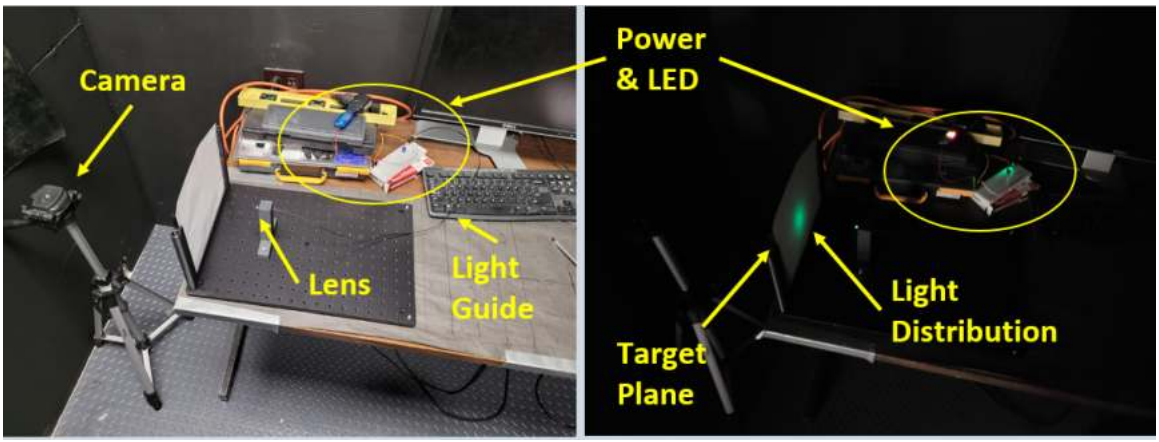


Figure 4.1: Experimental Setup and Example Data Trial

The camera is mounted upon a tripod to lock the axis normal to the camera lens to the axis normal to the light guide's lens. The tripod used allows for independent movement and locking in each of up/down, yaw rotation, and pitch rotation. Because the central consistent z-axis in Figure 3.1 is a fundamental assumption for the analysis geometry, the camera must be aligned by hand so that the camera is on-axis with the

emitter. While the camera height typically remained unchanged between trials, the optical board and its related parts had to be disassembled and stored away when data was not being taken. Although the edge of the table was used to make the position of the optical board as consistent as possible (Figure 4.1), it is unrealistic to expect separate measurements to have exactly the same positioning. Because of this chance the light guide had moved slightly, the camera was re-aligned to be on-axis with the emitter before any images were taken, including calibration and background images.

While the location of the camera must be realigned each time, this set up takes advantage of an optical board to hold the target plane in the same location relative to the lens stand for every trial. A U-shaped housing bracket is mounted to one end of the board via 2 vertical, cylindrical slots that slide onto optical posts. The U-shape enables efficient swapping of target planes between a calibration screen and a blank sheet of paper for data collection while maintaining the plane's origin on the aligned camera/emitter z-axis.

To minimize buckling of the translucent target plane, the paper was glued to another U-shaped bracket that could slide in and out of the above housing frame. While a full rectangular bracket would minimize the buckling even further, the Prusa Mini printers available could not print such a design without structural failure. Since the Prusa Mini printing beds could not fit the full frame laid flat, the part had to have been printed standing vertical. A compromise of a U-shape allowed the printers to construct a full bracket on the printing stage without the part collapsing under its own weight. If a 3D printer with a large enough bed is available, a full rectangular frame would be an upgrade to the experimental set up.

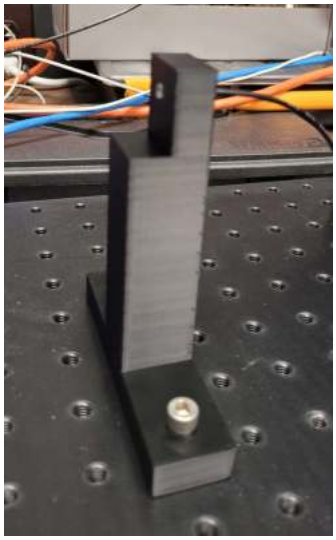


Figure 4.2: Lens Mount

The light output, the lens at the end of the light pipe, must be locked into only a single degree of freedom - on a rail forwards/backwards from the target plane. Once the target plane frame was constructed, the light output needed to be held at the height of the centre of the target plane. Because this stand only needed to support the small lens, the design incorporated holes on either side of the lens to interface directly with the optical board via screws. The light guide is threaded through a bore hole on the mount at the exact height of the centre of the target plane (Figure 4.2)

Figure 4.3: Collimator

and out towards the LED light source. The light pipe manufacturers did document a recommended bore size to thread the pipe through, and this recommended size allowed the lens to rest flush with the plastic and level with the optical table, as the pipe was kept straight and supported through the bore hole to ensure light left the lens as parallel to the table as possible.

Finally, the light source of the experiment is a single green LED in circuit with a voltage source and a variable resistor to control the brightness/power of the LED. A wooden cylindrical collimater with a hole at each end is secured onto the LED on one side and the light pipe on the other (see Figure 4.3). This collimator prevents light bleed from the LED into the measurement, as well as forces the light to enter the pipe at an acceptable angle for total internal refraction.

4.1 Measurement Procedure

Before anything else is done, it is imperative to configure the camera settings to have a preset aperture and exposure time to take an image in darkness. The required settings vary from camera to camera, and it is recommend to check the manuals and online resources to confirm the correct presets. In addition, the camera's focal length must be set and unchanged for both the darkness preset and the normal, fully lit camera mode for calibration images. Changing any aspect of the camera's configuration during a measurement can have catastrophic effects on the data collection, as everything except the plane-source separation distance must be kept as constant as possible.

After all the tools were correctly calibrated and configured, the first step of the data collection procedure was to assemble the setup in Figure 4.1 while the darkroom was otherwise unoccupied. Once the optical board was fully assembled, the camera was placed upon the tripod and aligned by eye, using the back display of the camera to position the target plane origin at the exact centre of the image.

The calibration screen is first inserted into the frame to obtain a calibration image while the lights were still on. An example of this calibration image is shown in Figure 3.3. The calibration screen is then replaced with a blank collection screen and the room lights are turned off to obtain a background image. Importantly, before an image is taken, the LED is turned on to ensure no bleeding from the source to the target plane, as such bleed could not be filtered out during analysis. After confirming

Figure 4.4: Example Background Image

no light bleed, the LED is turned off and a background image is taken. A background corresponding to the example in Figure 3.3 is featured in Figure 4.4.

After a background pedestal is recorded, the lights are turned back on and the light pipe was fed through the lens mount bore and connected to the LED circuit. The lens was gently pushed flush against the plastic, and all lens mount connections were hand-tightened to ensure no movement during data collection. The lights were turned off, the LED was turned on, and a data image of the resultant light distribution were taken. Two images were collected in order to not only protect against possible image corruption during data transfer to the algorithm, but also to confirm that the experimental setup did not move in between images. Once the images are collected, the lights were turned back on, the lens stand was unscrewed from the board, moved to the next trial distance, re-secured to the board, and new data images are taken. This continues until all trials are completed, for distances equal to 1, 3, 5, 7, and 10 inches. After all data images have been collected, the LED is turned off and a final background image is taken to compare with the initial background image to again protect against data corruption and ensure no camera motion during the measurement.

Finally, the experiment setup was disassembled and stored so that the darkroom could be used for other experiments. If multiple light pipes were being tested in a ses-

sion, the full assembly was reset back to the background image for each measurement to maximize experimental integrity.

4.1.1 Camera-to-Plane Correction

While the distance between the camera and the target plane is not relevant to the analytical geometry, each pixel on the target plane is a slightly different distance away the camera relative to the central origin pixel. As a result, the number of photons the camera sees in a pixel is only equal to the number of photons impacting said pixel for the exact centre, as all other pixels have a slight spread the causes some photons to miss the camera.

$$\text{relative efficiency} = \frac{d_{camera}^2}{d_{camera}^2 + r^2} \tag{4.1}$$

The relative efficiency in equation 4.1 is a ratio of distance vectors, the numerator representing the distance from the camera to the target plane's origin and the denominator representing the corrected distance to the camera from a specific pixel. r is calculated the same as in Equation 3.2 from a pixel's x and y coordinates, but d_{camera} must be obtained during the experimental procedure. After the equipment has been set up but before the calibration image is taken, the distance from the fixed and locked camera lens to the centre of the secured target plane is measured with a tape measure and recorded in inches, to remain consistent with the other distances measured. All unit conversions, as well as all calculations of Equation 4.1, are done by the Analysis object discussed in Section 3.2.3.

4.2 Results of Measurements

Images of light distributions is the important result of the experimental procedure, but not useful for any calibration purpose in image form. After being fully analyzed as explained in Section 3.2, the relevant information from the data images is presented as histograms. A $cos(\theta)$ and a ϕ histogram are created for each individual trial image and plotted overlaid with each other, so that the histograms can be compared qualitatively to confirm consistency across multiple distances. Finally, all trials for each light pipe are combined into a single distribution and normalized into a probability density function.

4.2.1 Example Results for Pipe A, $d = 3in$ Trial

Figures 3.8, 3.9, and 3.10 use the $d = 3in$ trial data from Pipe A (see Table 2.1) to demonstrate the Analysis algorithm, so the results of this trial, both image and histogram results, will be discussed here. Because the optical board spacing was in imperial, images will be referred to based upon the imperial separation distance, although all calculations are done in metric. The results from a single trial presented below are not representative of every measurement, but the procedure to obtain distribution histograms from an image is the same for all collected data.

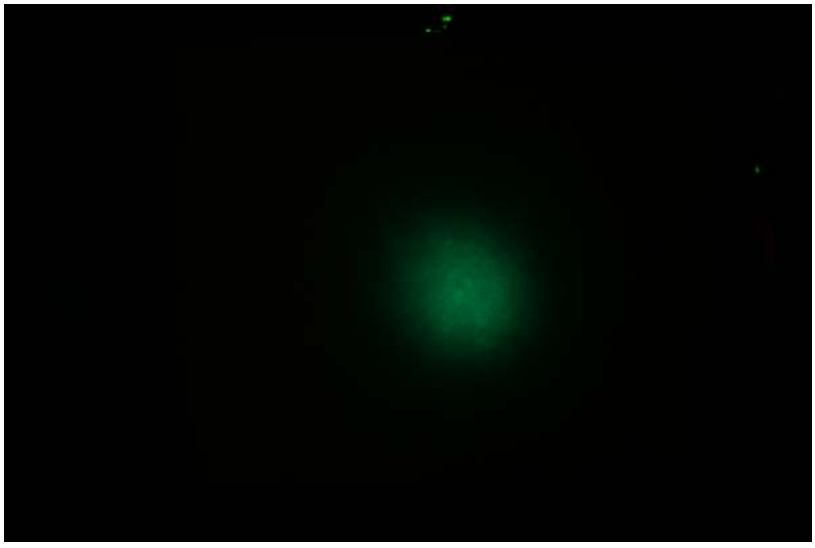


Figure 4.5: Data Image from Pipe A, $d = 3in$ Trial

Example of Image Results

Figure 4.5 shows a image result from the $d = 3in$ experimental trial, and it is readily apparent why the algorithm must first prepare a given image for analysis. First, light pollution from other instruments is visible at the top centre of Figure 4.5; this must be filtered out or will create false positives in the analysis. Second, the centre of the light distribution appears offset from the middle of the image, and so must be found independently of the target plane's pixel matrix. Section 3.2.3 discusses the procedure to remove these issues at length, but the results of these steps are shown here.

Figure 4.6 shows Figure 4.5 after the algorithm has converted the image to greyscale and the background has been removed. As confirmation, the light pollution at the top of the image is no longer present. Figure 4.7 shows the result of the calibration

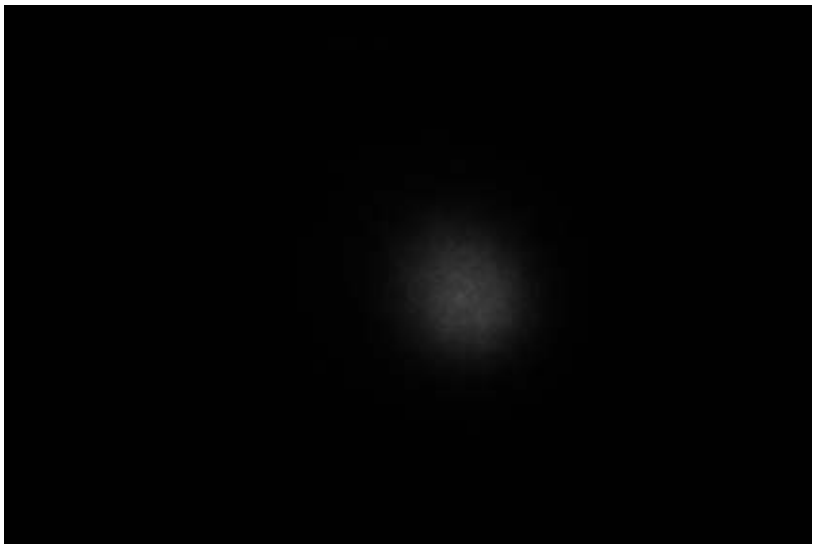


Figure 4.6: Filtered Image from Pipe A, $d = 3in$ Trial

in Section 3.1.2. While the exact centre of the light distribution is determined by the processes in Section 3.2.3, it is clearly evident that the centre of the light distribution is located much closer to the target plane's origin than Figure 4.5 suggests.

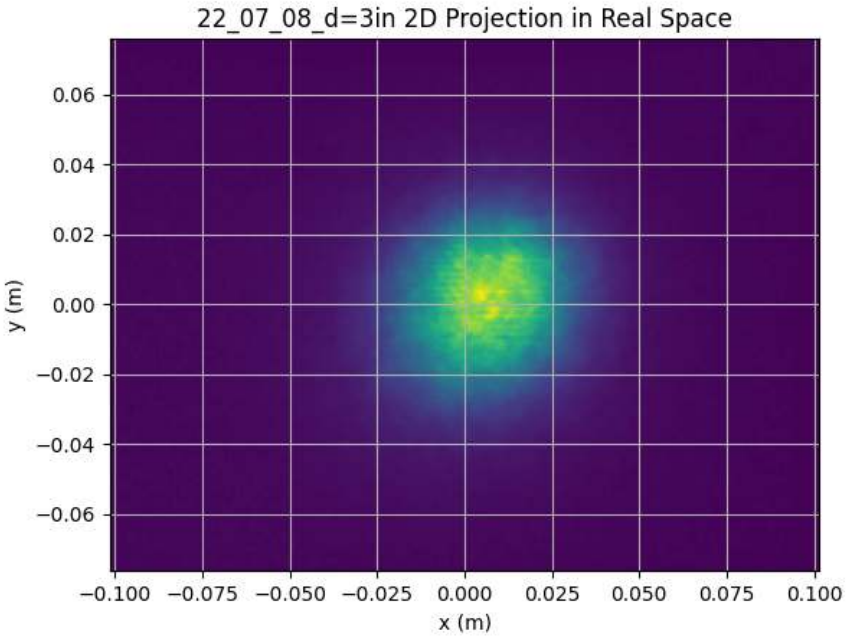


Figure 4.7: Reconstructed Projection of Pipe A, $d = 3in$ Results

Angular Distributions

Section 3.2.3 concludes with the trigonometry required to extract angular distributions from a calibrated and centred image such as Figure 4.7. Each pixel has an associated $cos(\theta)$, ϕ, and counts value, which are then plotted into $cos(\theta)$ and ϕ histograms, examples of which are shown in Figures 4.8 and 4.9 for Pipe A, $d = 3in$ data.

The polar $cos(\theta)$ distribution (Figure 4.8 is expected to descend rapidly from $cos(\theta) = 1$, the centre of the distribution. The red line marks the $cos(\theta)$ limit discussed in Section 3.2.3, where the data abruptly ends. This is another important reason to combine histograms - the $cos(\theta)$ limit moves closer to $cos(\theta) = 1$ the further away the source is to the target plane. The histogram for the $d = 1in$ trial is therefore more accurate for low $cos(\theta)$ values, as more data for such values are recorded, while the data from $d = 10in$ is predominately near $cos(\theta) = 1$.

The azimuthal ϕ distribution (Figure 4.9), while interesting to plot, is not impor-

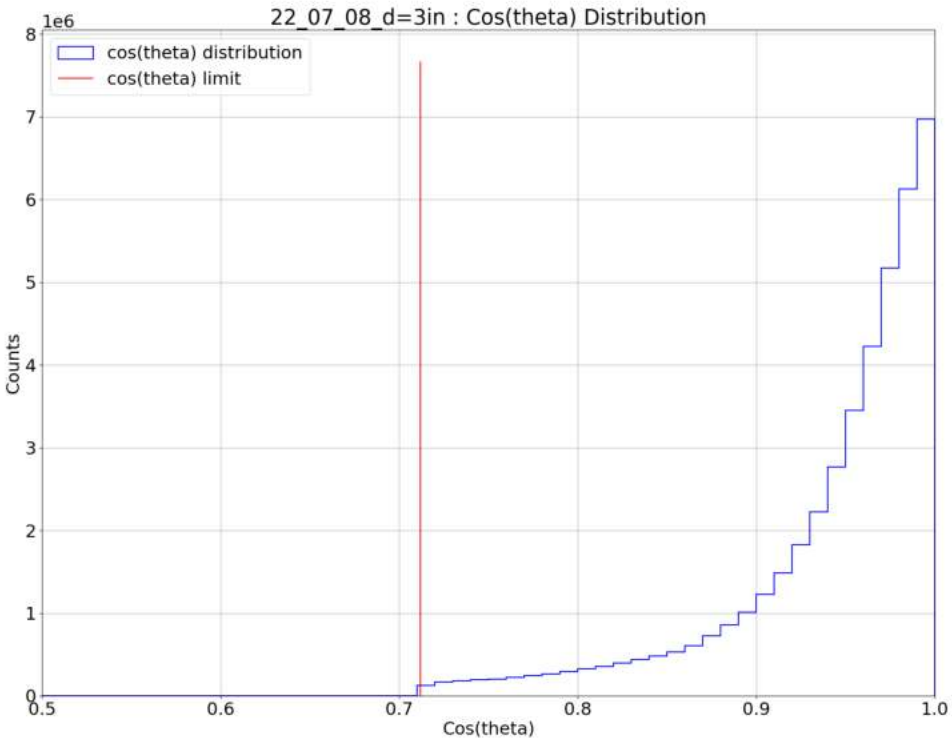


Figure 4.8: $cos(\theta)$ Histogram of Pipe A, $d = 3in$ Distribution

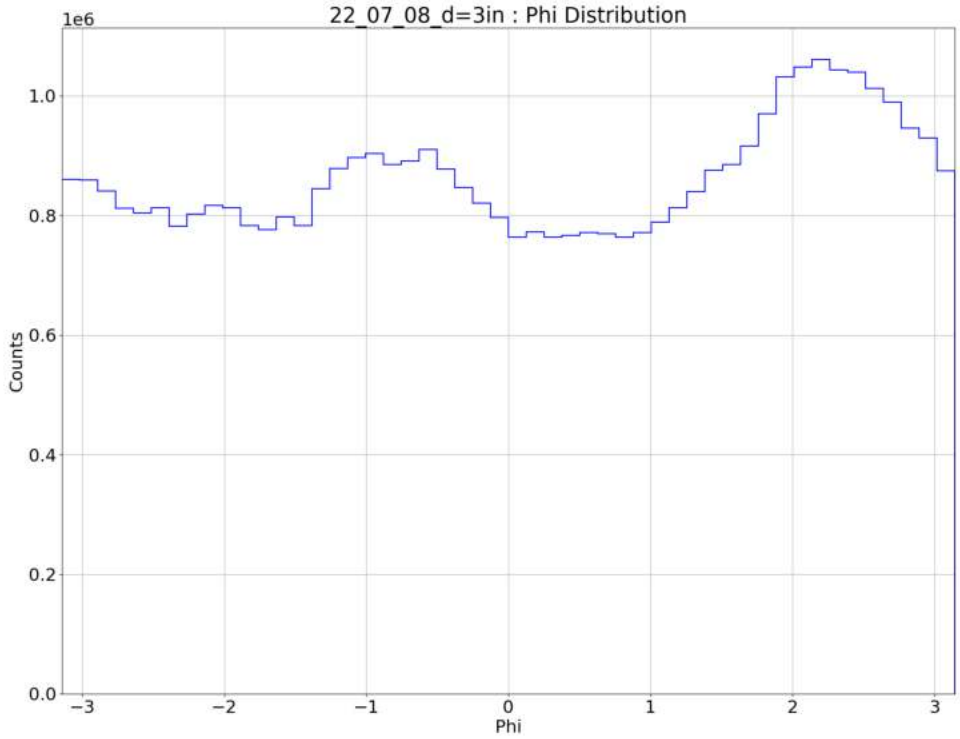


Figure 4.9: ϕ Histogram of Pipe A, $d = 3in$ Distribution

tant to the overall purpose of this research. While the polar distribution describes shape and size of the light pulse, the azimuthal distribution merely describes the rotation asymmetry of the lens. Although each ϕ histogram is expected to be uniform, deviations from a uniform distribution suggest imperfections in the lens itself, rather than speaking to the capabilities of the light pipe. The goal of this project is to determine the best light pipe for exposure inside the detector tank, and so the shape and size of the light pulse is the desired result, not the polar distribution. However, it is important to collect, save, and plot this data to confirm the fundamental assumption of the measurements - the angular distribution is independent of the separation distance between source and target plane.

4.2.2 Combined Results

Alone, the previous histograms do not communicate much relevant information. It is the comparison between distance trials of a single pipe that begets the combined

distribution needed for the simulations in Chapter 5. As such, it is no longer beneficial to use a single example, and so the combined distributions for each light pipe presented in Table 2.1 will be discussed in length.

Normalized ϕ Distributions

While the exact details of the polar histograms are not necessary for the tank calibration, it is still important to plot the histograms for each trial together to qualitatively confirm consistency across measurement trials. The fundamental assumption of the project is that the behaviour of the lens is consistent across all distances, and Figure 4.10 confirms that this is indeed the case for Pipe A.

Pipes B and C, however, appear much less consistent than Pipe A. While it is still safe to assume consistency over various distances, the overlapping plots vary wildly in the oscillations around a constant uniform average. While Figure 4.10 shows each distribution oscillating together, both Figure 4.11 and 4.12 have peaks and troughs

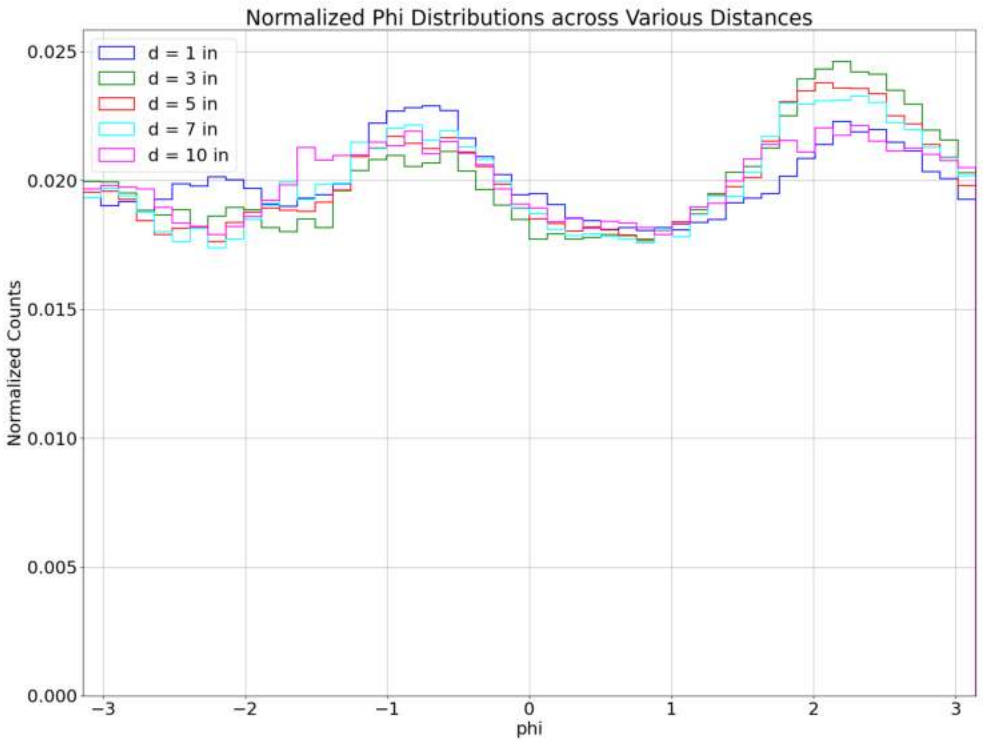


Figure 4.10: Normalized ϕ Distribution for Pipe A over Various Distances

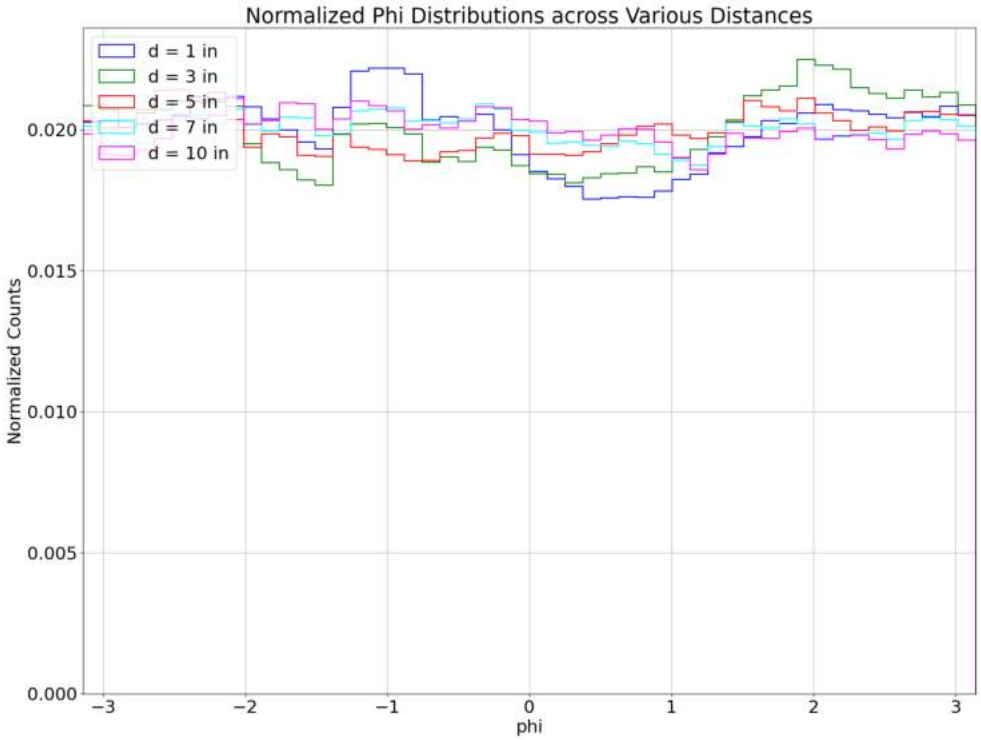


Figure 4.11: Normalized ϕ Distribution for Pipe B over Various Distances

that are not shared amongst all distance trials. Both Pipes B and C appear to have inconsistencies around $\pm\frac{\pi}{2}$, where one data trial peaks while another troughs. These inconsistencies are attributed to imperfections in the attached lens, as the rotational position of the light pipe was not fixed during data collection. It is easily possible that the light pipe and lens rotated within the stand during motion, resulting in the inconsistencies for Pipes B and C. While not a condemnation of these two options, it is important to note that the data suggestions Pipe A has superior rotational symmetry in its lens. Whether this is a product-wide tendency or specific to the parts measured is unknown, and a potential area of further research.

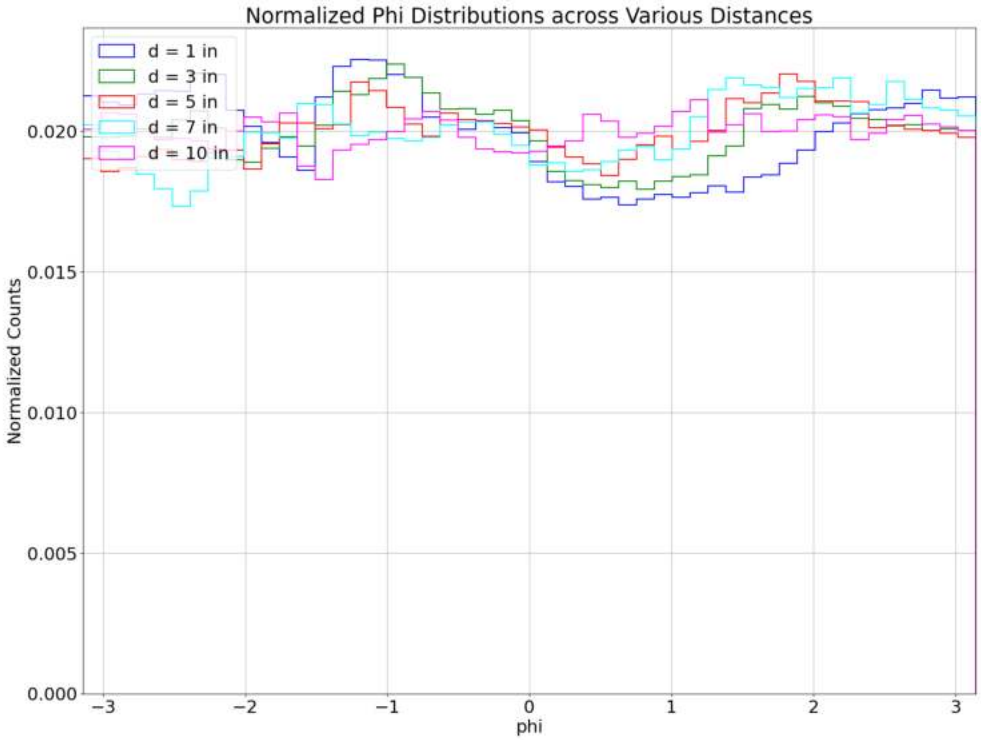


Figure 4.12: Normalized ϕ Distribution for Pipe C over Various Distances

Standardized $cos(\theta)$ Distributions

The histograms in Figures 4.10, 4.11, and 4.12 are not the raw distributions obtained like in Figure 4.9, as the total number of photons detected varies with the separation distance. In order to compare like to like, the integral of the distribution was set to one so that each bin represented a fraction of the total, rather than an arbitrary counts value. This normalization only scales the ϕ distributions vertically, and so the affect is only noticeable in the scale of the vertical axis.

However, the maximum number of counts has a dramatic effect on the $cos(\theta)$ distributions, particularly for higher $cos(\theta)$ values where the intensity of the beam spot is greatest. To account for this, each $cos(\theta)$ distribution was standardized around a single bin value, specifically the $cos(\theta) = 0.95$ bin. Each bin is then evaluated as a percentage of this bin's counts, so it is important to visualize the behaviour of the curve both below and above the chosen bin.

Figure 4.13 plots the standardized polar histograms for Pipe A on a logarithmic

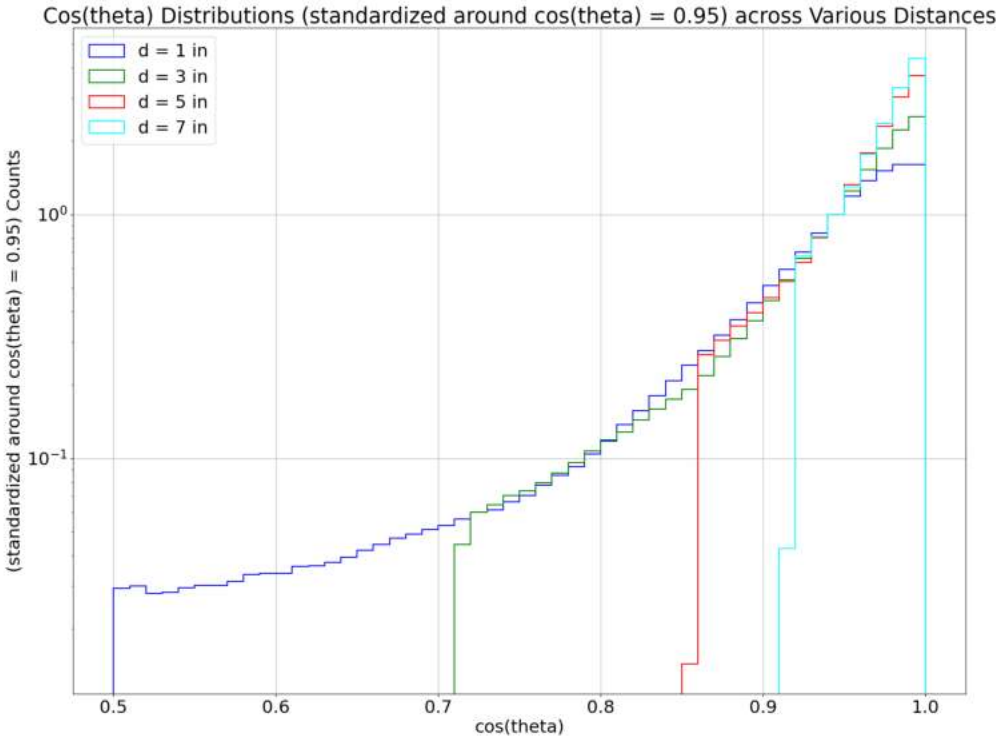


Figure 4.13: Standardized $cos(\theta)$ Log Distribution for Pipe A over Various Distances

vertical axis. A linear vertical axis hides the behaviour of the distributions for low $cos(\theta)$ values, as the photon count for $cos(\theta) = 1.00$ are several orders of magnitude higher than the photon count for pixels on the edges of the light pulse. A logarithmic axis, however, shifts the focus of the plot from only the most full bins to all bins, so the behaviour at lower $cos(\theta)$ is not excluded from evaluation. As mentioned in Section 4.2.1, $d = 1in$ data is most accurate for low $cos(\theta)$ and $d = 10in$ is most accurate for high $cos(\theta)$. This is evident in Figure 4.13 around $cos(\theta) = 1.00$, where the $d = 1in$ distribution terminates at a much lower value than any of the others. Despite this expected inconsistency, this histogram clearly depicts consistent behaviour across all distances for the measured $cos(\theta)$ spectrum.

Pipes B and C, on the other hand, exhibit far less consistency. Applying the same standardization and logarithmic axis as Pipe A, the distributions do not converge nearly as well either above or below the chosen $cos(\theta) = 0.95$ bin. Pipe B in Figure 4.14 has modest agreement for $\cos(\theta) > 0.9$, but the behaviour at low to mid

$cos(\theta)$ varies noticeably between distance trials. Pipe C (Figure 4.15) is even worse, exhibiting similar behaviour to Pipe B for low $cos(\theta)$ values but descending as $cos(\theta)$ approaches 1. While it is expected for $d = 1in$ to be inaccurate at higher $cos(\theta)$, here the $d = 1in$ trial behaves in the exact opposite way as the other trials; continuing to increase while the others drop off slightly.

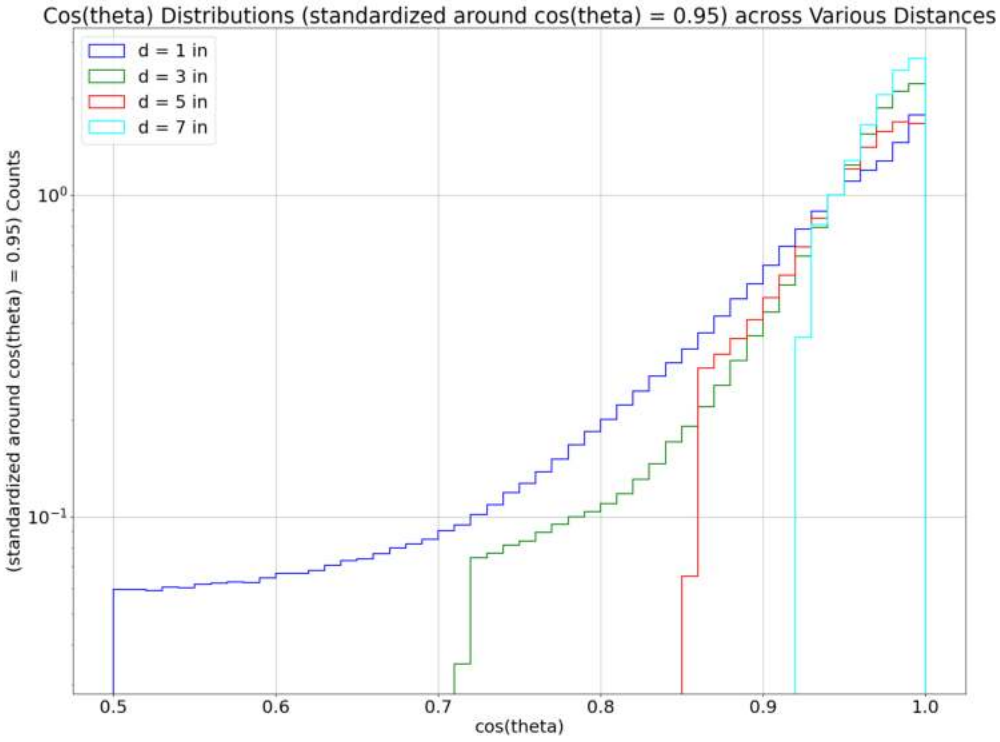


Figure 4.14: Standardized $cos(\theta)$ Log Distribution for Pipe B over Various Distances

Probability Density Functions

The combined $cos(\theta)$ plots above provide enough information to create a comprehensive single histogram object for the lens, independent of separation distance. For Pipe A, the $d = 1$ trial data is consistent with all the other trials up to the 0.90 bin, it can

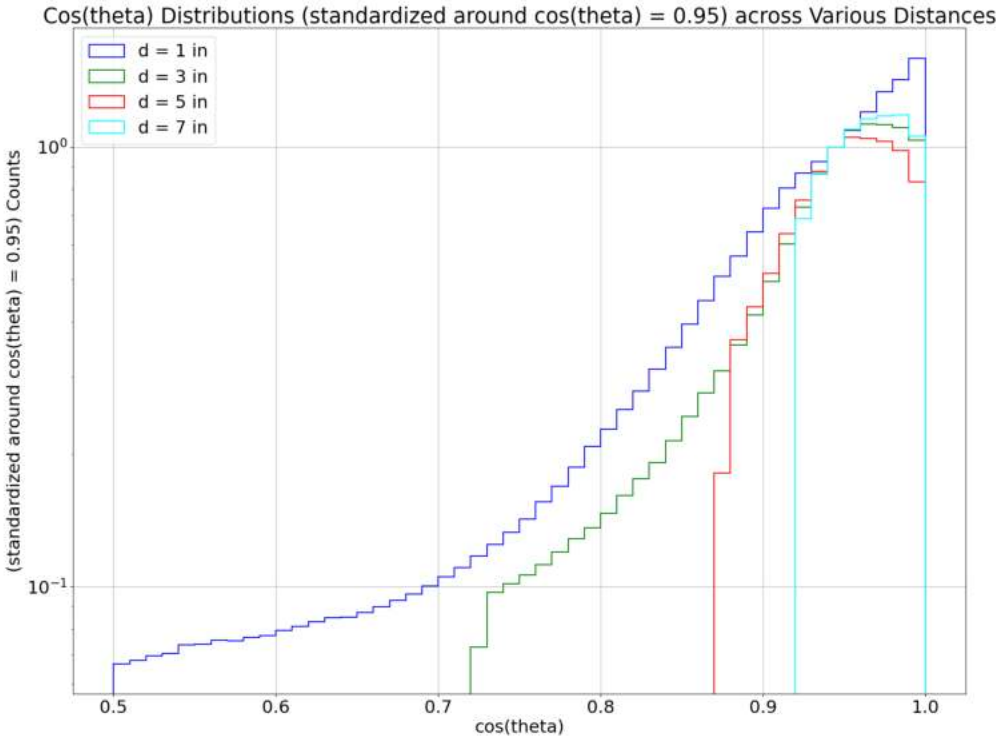


Figure 4.15: Standardized $cos(\theta)$ Log Distribution for Pipe C over Various Distances

be assumed that it would be consistent in shape with the other trials for $cos(\theta) < 0.75$ (if data could be collected for that region at higher separation distances). For Pipes B and C, although trial $d = 1$ is positioned slightly above the other trials, the shape of the curve below $cos(\theta) < 0.90$ remains consistent. As the final density function will be normalized independently, this discrepancy in magnitude is ignored. To complete the desired range for the PDFs, the $d = 5in$ distribution is consistent with higher distance trials for high $cos(\theta)$ values for each pipe, and so can be used for the $cos(\theta)$ values where the $d = 1in$ falters. Normalizing these two distributions and coupling the first full $d = 5in$ bin to the same $d = 1in$ bin creates a continuous, normalized probability density function, shown in Figures 4.16, 4.17, and 4.18. Defining these probability density functions is the final measurement of the project, and the LED calibration lights can be simulated inside a detector to evaluate the preferred light pipe.

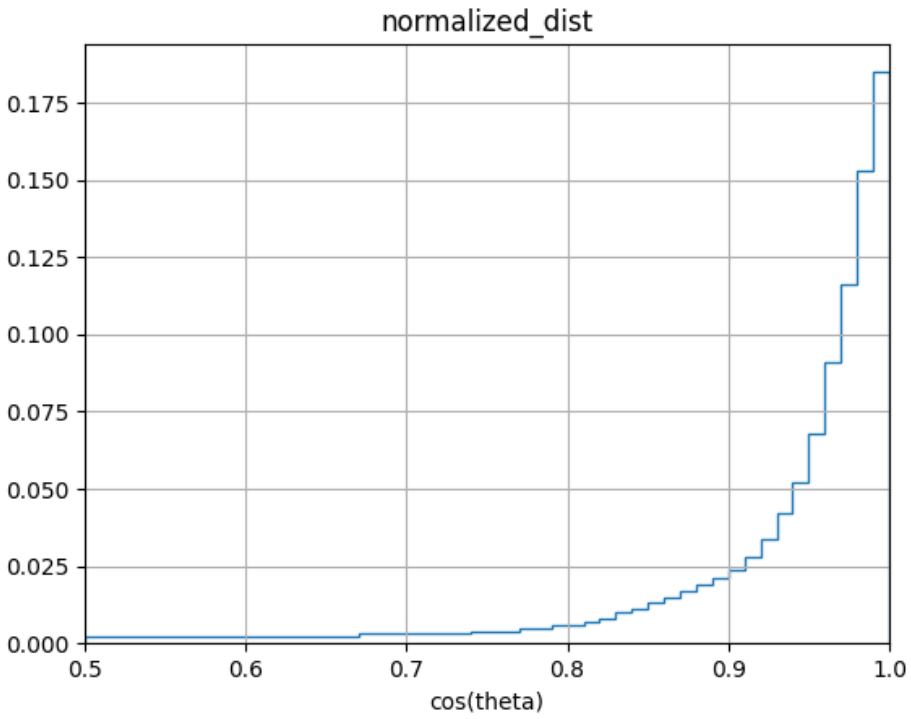


Figure 4.16: $cos(\theta)$ Probability Density Function for Pipe A

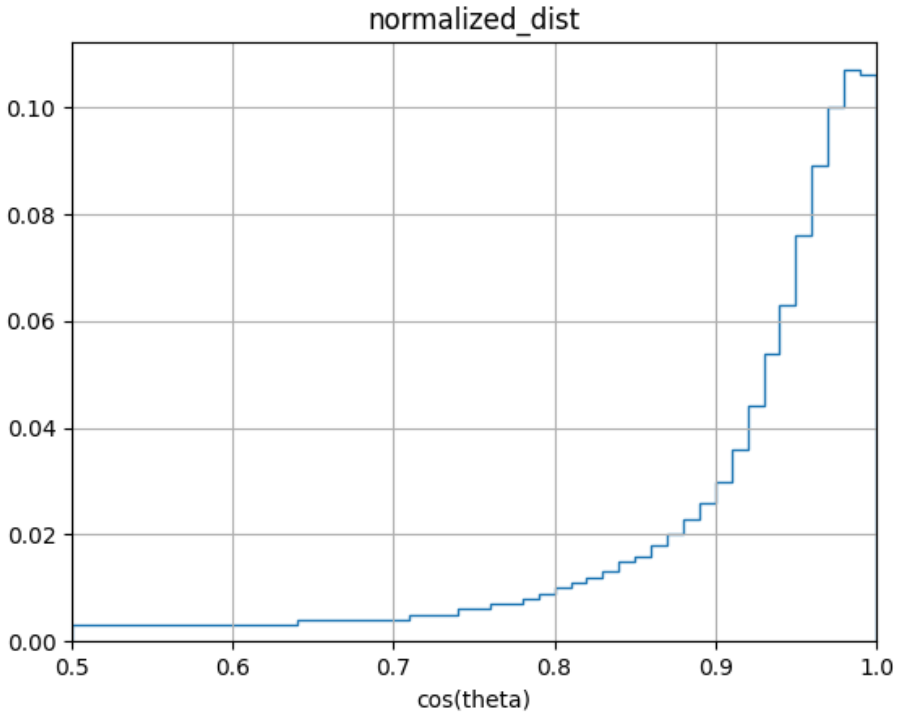


Figure 4.17: $cos(\theta)$ Probability Density Function for Pipe B

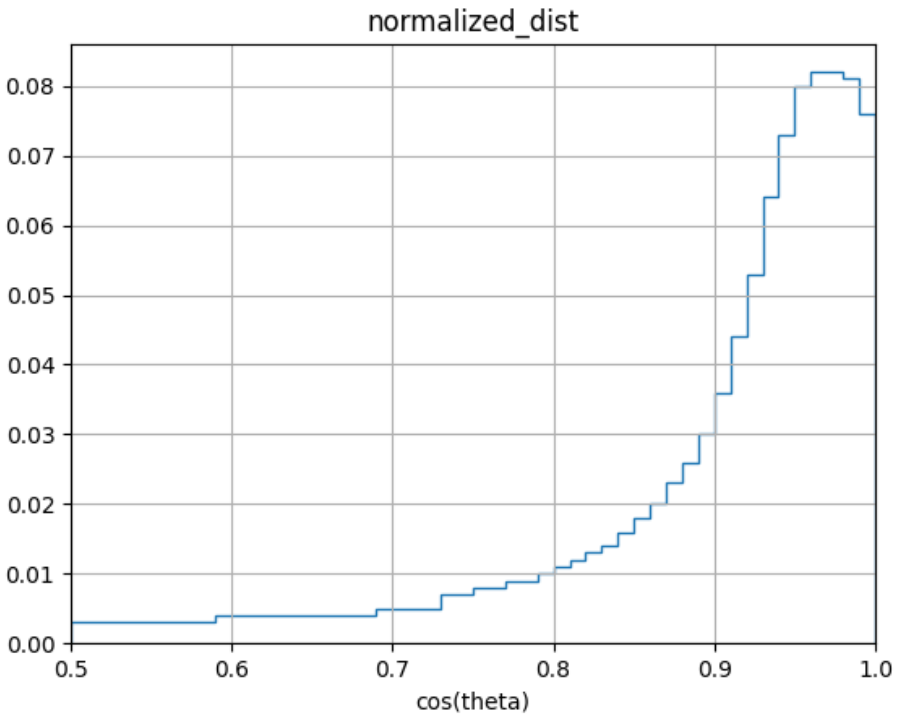


Figure 4.18: $cos(\theta)$ Probability Density Function for Pipe C

Chapter 5

Evaluation, Analysis and Comparisons

Now that the probability density function (pdf) for each light pipe has been measured, it is necessary to observe their respective effects in the detector calibration procedure. Figures 4.16, 4.17, and 4.18 merely dictate the likelihood of a photon to be emitted at a specific $cos(\theta)$ from the lens of the light pipe. In order to evaluate the option best suited for the calibration described in Section 2.3, a list of photon trajectories for each pipe is created via a Monte Carlo of its respective pdf and then fired into a simulation of the IWCD tank.

5.1 Monte Carlo Method

To generate an arbitrary number of photons trajectories, a random $cos(\theta)$ and ϕ must be generated from their respective probability density functions. For ϕ, a uniform distribution between 0 and 2π is assumed, as the imperfections in ϕ discussed in Section 4.2.2 are sufficiently small to be ignored at this stage. For $cos(\theta)$, however, the simulation employs an Acceptance-Rejection method to generate random values according to the given pdf. In brief, the two dimensions of the pdf ($cos(\theta)$ bin and probability density) are assigned uniform random distributions between the dimension's minimum and maximum. A random value is then generated for each dimension, for example 0.08 in the vertical and 0.75 in the horizontal. The random vertical value is then compared directly to the value of the PDF at the random horizontal value and accepted if the random value is lower than the PDF value. For the above example,

the PDF value at 0.75 on the horizontal is much less than 0.02 for each light pipe, and so a 0.08 on the vertical will be rejected regardless of the simulation. This ensures values for a given bin are generated at the frequency the pdf dictates, although processing time can become a concern if more than a billion photons are generated.

5.2 Tank Simulation

After generating a hundred million (10^8) photon trajectories, the final impact location of each photon can be determined given the dimensions of the tank and a fixed location of the source. While mPMTs will line the entirety of the tank interior, the simulated mPMT source is assumed to be fixed on the tank wall, centred vertically between the lid and the base.

5.2.1 Snell's Law

Before a photon can travel through the tank, it must first exit the mPMT. After leaving the light pipe, a photon passes from air (inside the sealed mPMT) through the plastic dome cover and into the water of the tank. Each of these substances has its own index of refraction, and so the refraction of the light as it passes through each interface must be accounted for.

$$n_1 sin(\theta_1) = n_2 sin(\theta_2) = n_3 sin(\theta_3) \tag{5.1}$$

Fortunately, from the symmetry of Snell's Law (Eq. 5.1), the intermediary plastic dome can be ignored, and only the refraction from air into water is relevant. The index of refraction for light and water are well established ($n_{air} = 1.0003$, $n_{water} = 1.33$), and the incoming angle is equivalent to the polar angle of the photon, so calculating the refraction is rudimentary within the simulation.

5.2.2 Tank Geometry

To determine the impact coordinates on the side of the tank, the photon trajectories have to be converted from the cylindrical coordinate system of the mPMT source to the cylindrical coordinate system of the tank itself. Following the coordinate system defined in Figure 3.1, the z-axis is defined as the normal axis of the mPMT, from which the azimuthal angle is measured. However, the cylindrical tank defines the

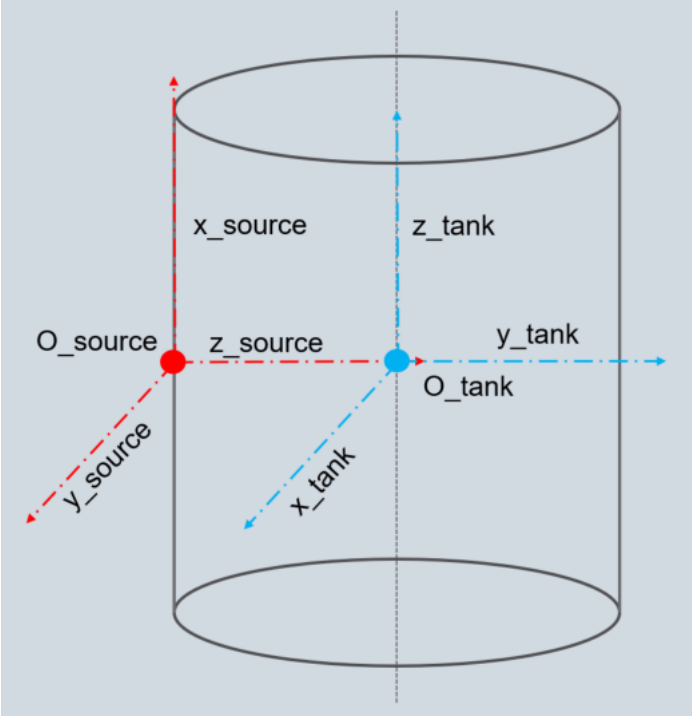


Figure 5.1: Line Drawing of Tank Simulation

z-axis between the lid and the base, a 90-degree rotation from the established mPMT source. To rotate the trajectory into the tank's coordinates, both systems are broken down into their Cartesian unit vectors (see Figure 5.1), and the rotation between them is defined in Equations 5.2. The generated (θ, ϕ) trajectory is reduced to the unit vectors of the source, rotated according to Eq. 5.2, then expanded back into the cylindrical coordinates of the tank.

$$\begin{aligned} \hat{x}_{tank} &= \hat{y}_{source} \\ \hat{y}_{tank} &= \hat{z}_{source} \\ \hat{z}_{tank} &= \hat{x}_{source} \end{aligned} \tag{5.2}$$

Because both Cartesian and cylindrical require the vertical (base to lid) component, the height of the collision point is defined first. If it is within the vertical bounds of the tank, the photon collides with the tank wall at the known radius of the tank. The linear path of the photon can then be projected into the XY plane as a polynomial $y = mx + b$, where the slope m is the ratio of known unit vectors $\hat{y}/\hat{x}$. This polynomial, the equation of a circle (Eq. 5.3), and the known radius of

the tank create a system of equations the simulation then solves for the Cartesian collision coordinates for the photon. α is used as the azimuthal tank coordinate in order to reduce confusion with the measured angles θ and ϕ.

$$\begin{aligned} x^2 + y^2 &= r^2 \\ tan(\alpha) &= \frac{y}{x} \end{aligned} \tag{5.3}$$

If the collision point is outside the vertical bounds of the tank, the impact will be on either the lid or base, depending on the sign of the trajectory, and determining the coordinates of impact is slightly more complicated. While the photon is blocked from hitting the wall by the lid/base, the path it takes will not change. The XY polynomial is created and solved as before to determine the location of the phantom collision, but then the simulation uses similar triangles to determine the Cartesian location of the impact on the lid/base from the known limiting z.

Finally, the Cartesian coordinates (x, y, z) of the collision are converted to the cylindrical coordinates (r, α, z) using the equations of a circle (Eq. 5.3). The impact coordinates are then stored in a data array according to the surface the photon collides with (i.e. lid, wall, or base).

5.2.3 Units and Weights

Currently, each entry in the data array represents the raw number of photons that would impact the respective "pixel" of the tank. However, this is a misleading statistic due to the circular shape of the base and lid. Pixels close to the centre of a circle will be much much smaller than not only those further from the centre, but also the pixels on the wall. To account for this, the raw count of photons in a pixel is divided by the area of that pixel, calculated from the known WTCE Tank dimensions. This converts the data array into counts per unit area, which can be compared directly between the three surfaces.

It is important that the area of each pixel is saved in a separate data structure of the tank as this conversion is done. When the intensity distributions are plotted in Section 5.3, the histograms need to be weighted according their respective areas. Without any weights, the probability of choosing an arbitrary pixel is proportional to the size of the pixels, which will inaccurately show a higher intensity for larger pixels. Including the pixel areas as weights for each pixel counts per unit area correctly makes

choosing a random pixel a uniform distribution, allowing again for direct comparisons between the three surfaces.

5.2.4 Visualization of Results

After each pixel has a counts per unit area, the simulation meshes each of the arrays into a two-dimensional data structure according to the shape of the surface. While the

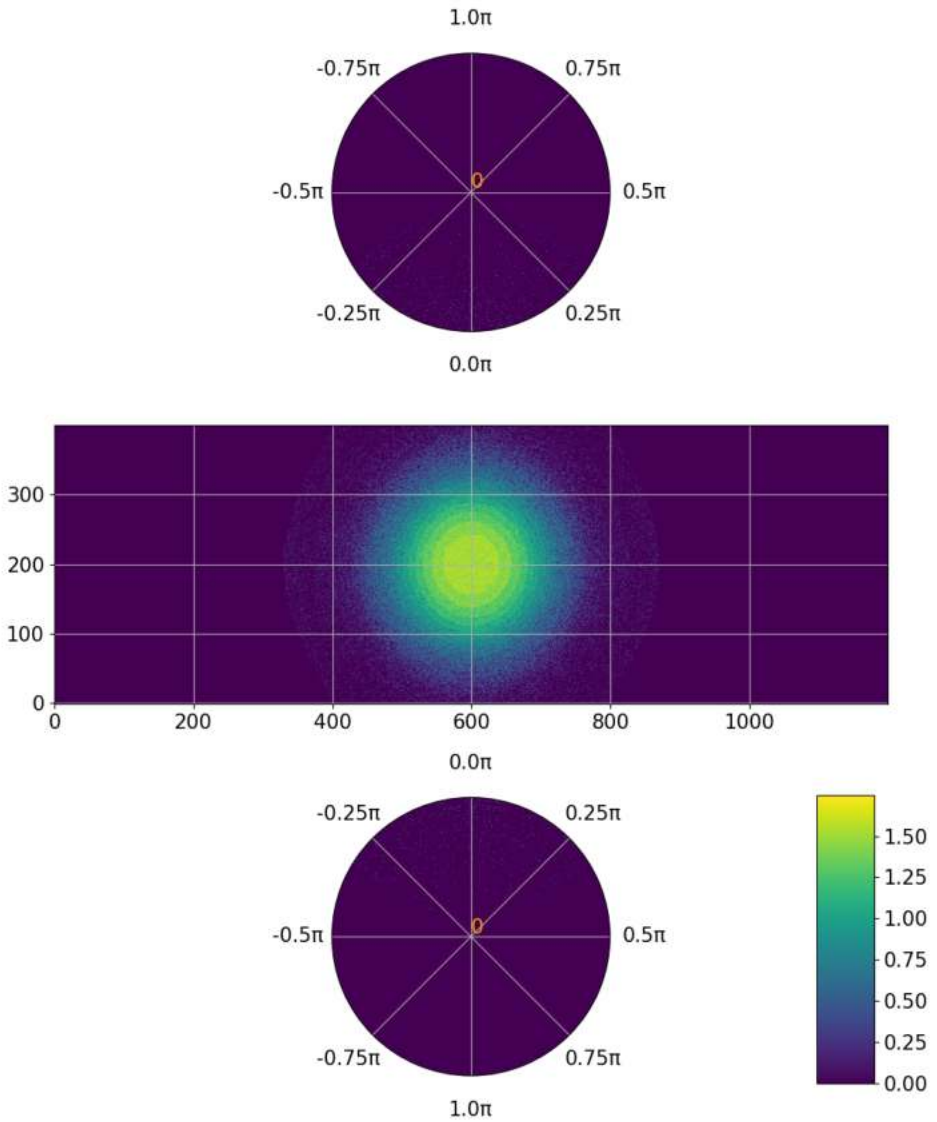


Figure 5.2: Tank Simulation for Pipe A using Logarithmic Scale. Lid/Base - Radians along circle exterior, Radius (cm) in orange within circle. Wall - Height (cm) on the vertical and normalized number of PMTs on the horizontal.

tank wall is a simple rectangle of (z, α), the lid and base are circles of (r, α). To display the full simulation of the tank, the algorithm uses Python method np.meshgrid() to create a plottable heatmap.

A logarithmic scale is used once again to highlight the behaviour at low $cos(\theta)$, where several orders of magnitude fewer photons will collide. This also provides a clear assessment for one of the requirements for calibration - over- or under-saturation

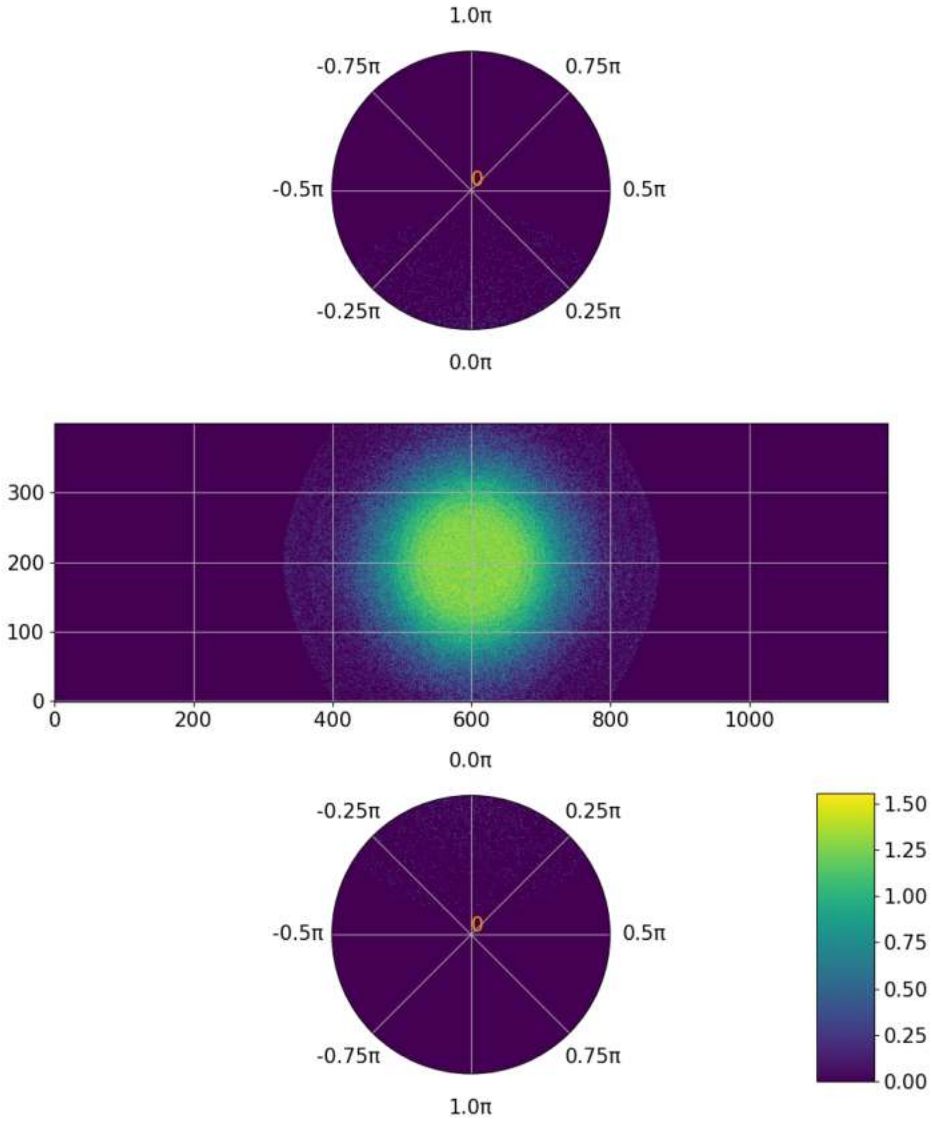


Figure 5.3: Tank Simulation for Pipe B using Logarithmic Scale. Lid/Base - Radians along circle exterior, Radius (cm) in orange within circle. Wall - Height (cm) on the vertical and normalized number of PMTs on the horizontal.

of the PMTs. The detectors have a dynamic range of 100, meaning that the difference between the minimum signal of a single photoelectron and complete saturation is a factor of 100, or 2 in log scale. To ensure all detectors that lie within the exposure range of the light pulse see measurable data, the relative intensity from the centre of the pulse to the edge of the heatmap must be within 2 orders of magnitude. Looking at Figure 5.2, this is indeed the case, as the intensity colour bar is capped around

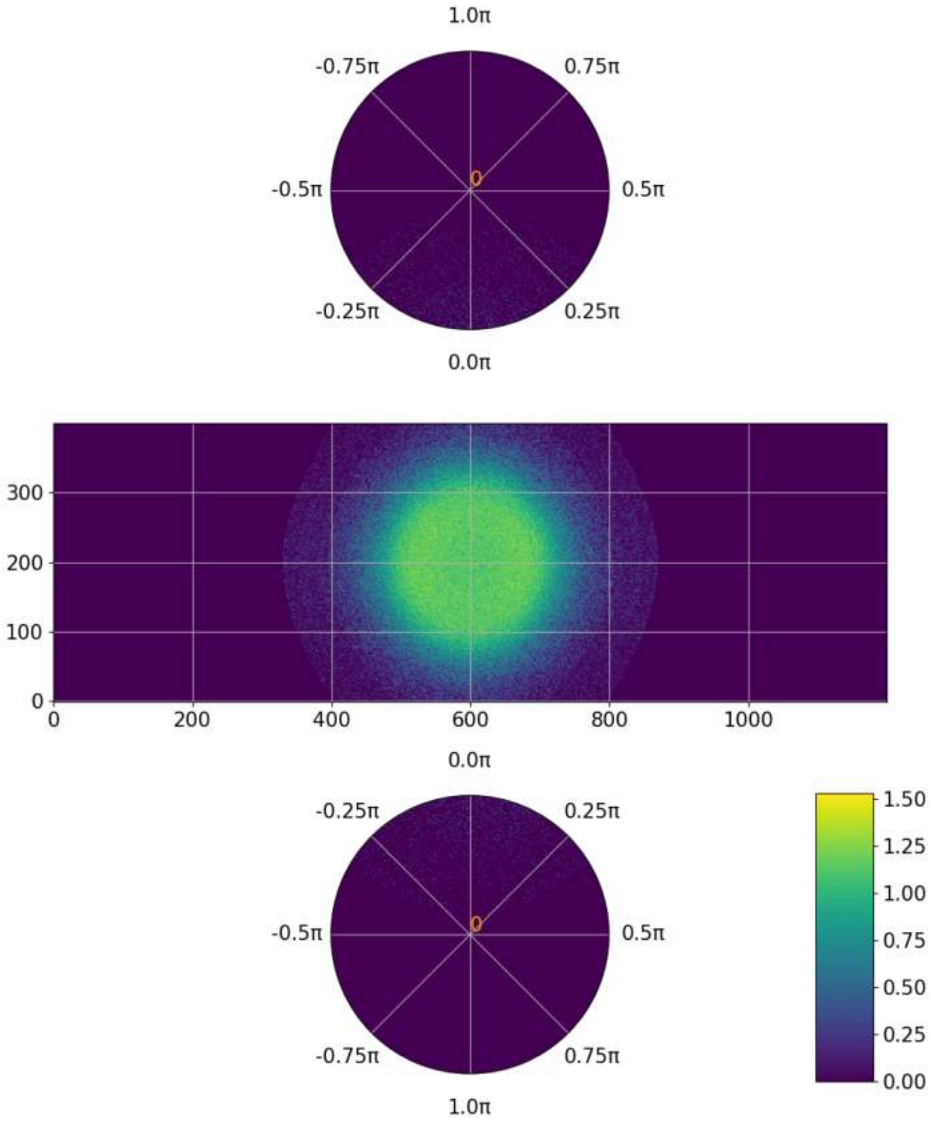


Figure 5.4: Tank Simulation for Pipe C using Logarithmic Scale. Lid/Base - Radians along circle exterior, Radius (cm) in orange within circle. Wall - Height (cm) on the vertical and normalized number of PMTs on the horizontal.

1.75. Figures 5.3 and 5.4 display similar behaviour, capping at 1.50 each, confirming that all three options fulfill the saturation requirement for calibration. As for total area covered by the light pulse, each light pipe results in a similar spread over the majority of the opposite side of the tank, from base to lid. Any variations in this area is insignificant in comparison to the total area covered, and again there is no major distinction between the options to justify a recommendation of a specific pipe.

5.3 Intensity Distribution

With all three pipe options being relatively interchangeable within the saturation and light area requirements, it falls to the intensity distribution to provide a clear justification for one option over the others. To reiterate Section 2.3, the calibration will be most successful when as many mPMTs see the same light distribution as possible. In terms of measurable quantities, as many mPMTs as possible need to see similar intensities of light so they can be compared directly during calibration.

Examining Figures 5.5 through 5.7, all three pipes have similar properties - two clear and distinct peaks accompanied by a continuous smear of decreasing probability. However, the peaks in the lower bins are deceiving, as the first peak at 0.0 is from

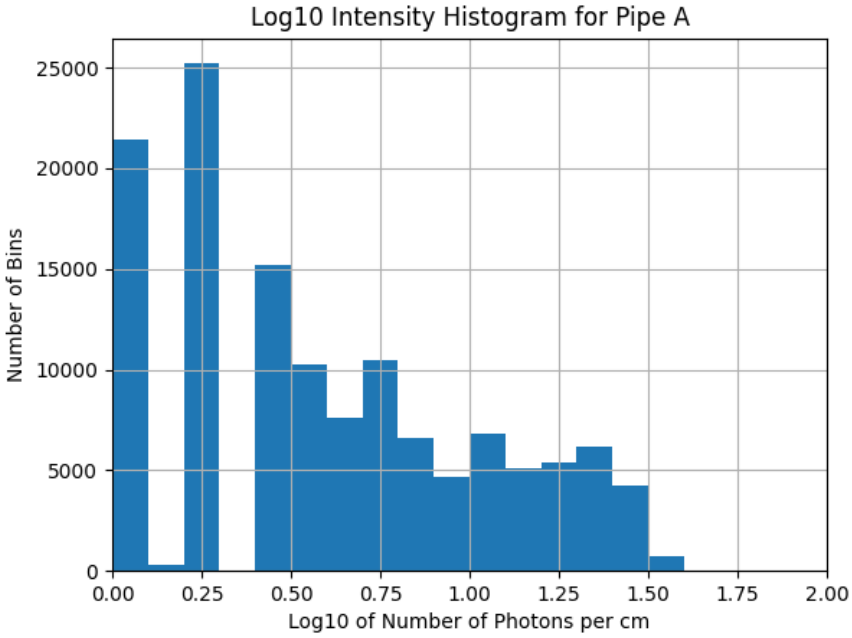


Figure 5.5: Intensity Distribution for Pipe A using Logarithmic Scale

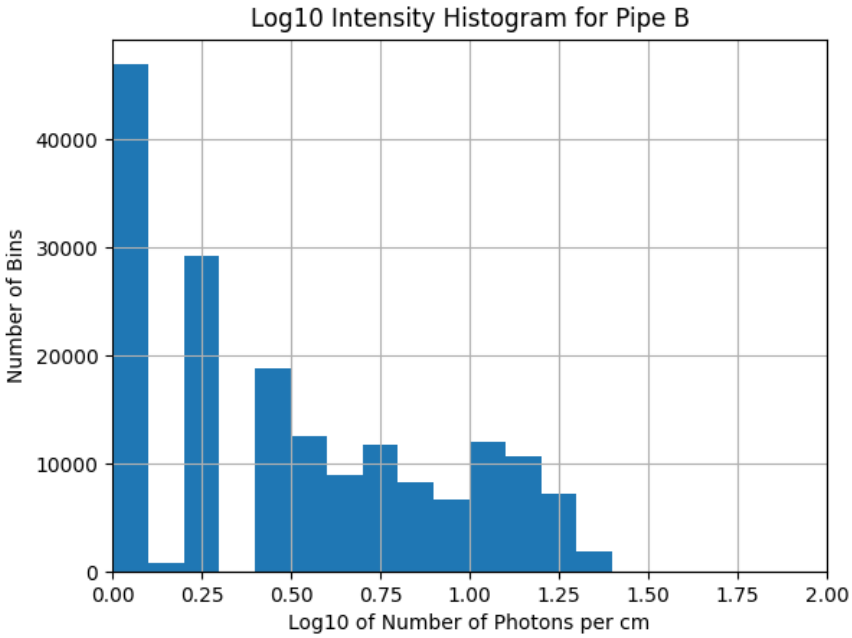


Figure 5.6: Intensity Distribution for Pipe B using Logarithmic Scale

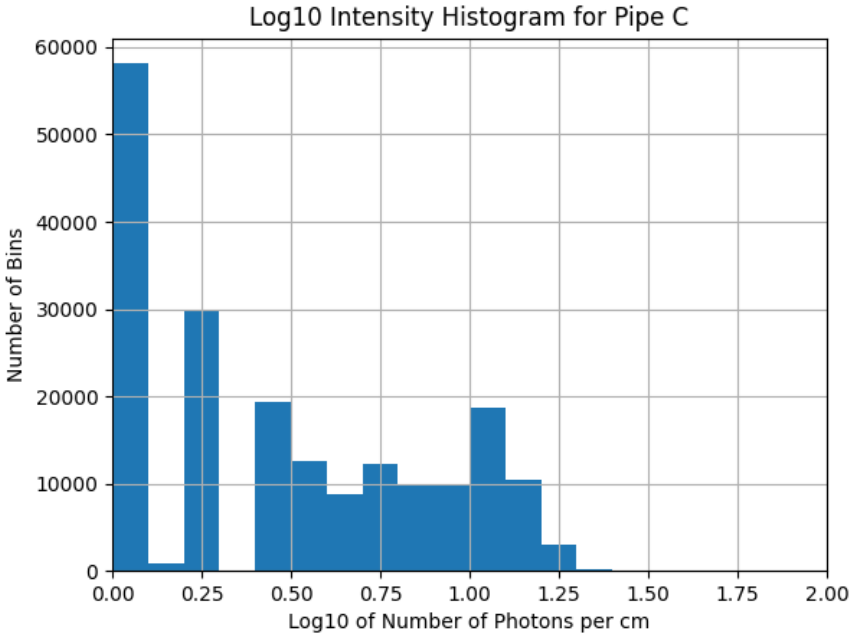


Figure 5.7: Intensity Distribution for Pipe C using Logarithmic Scale

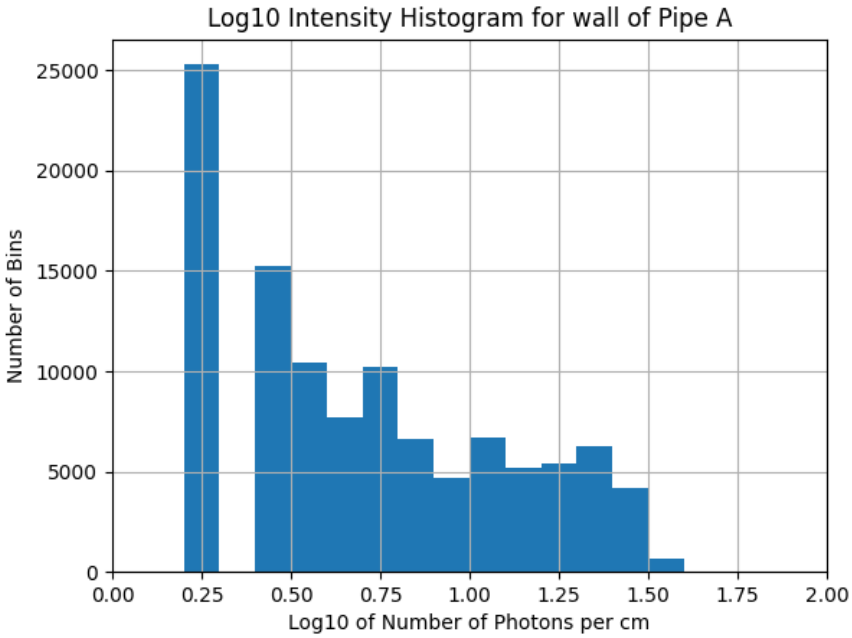


Figure 5.8: Intensity Distribution for Wall of Pipe A (Log Scale)

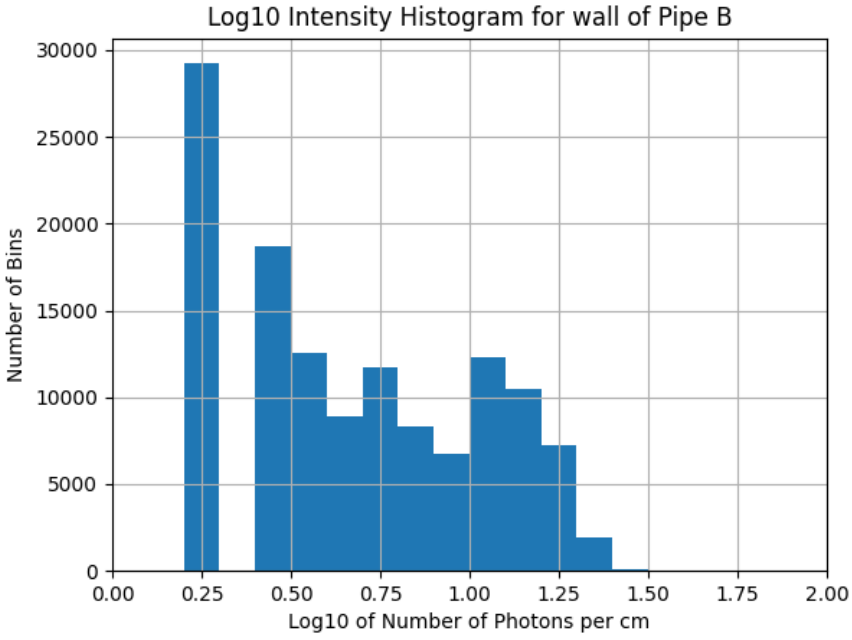


Figure 5.9: Intensity Distribution for Wall of Pipe B (Log Scale)

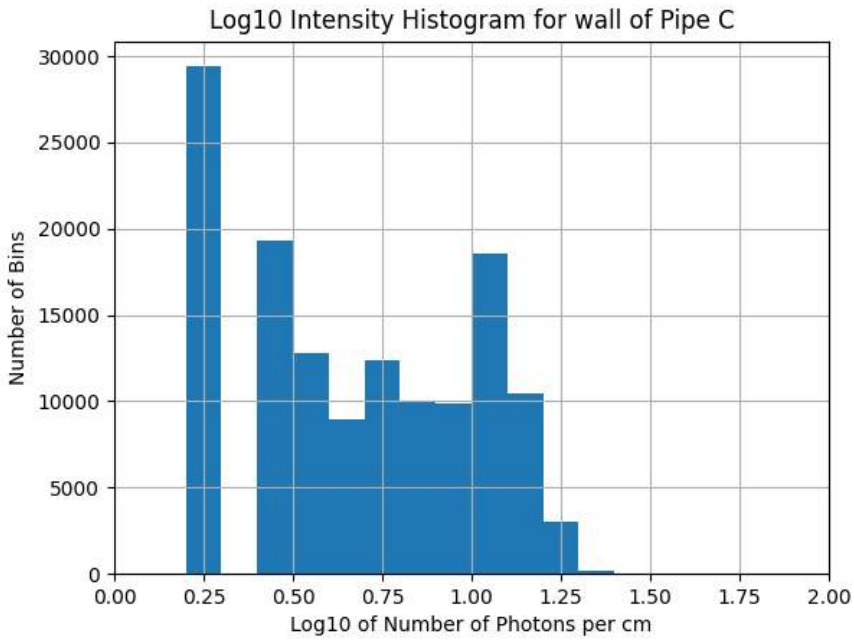


Figure 5.10: Intensity Distribution for Wall of Pipe C (Log Scale)

the lid and base regions and the second peak, and indeed the rest of the distribution, is from the wall. Figures 5.5 through 5.7 display histograms for the entire cylinder, followed by Figures 5.8 through 5.10presents the extracted wall region only. This refinement has a drastic affect on the vertical scale of each histogram. Removing the massive amounts of 0.0 counts from the endcap regions in the first trio causes the vertical axes of the second trio to be functionally identical.

While Figures 5.5 through 5.7 suggest a meaningful difference in scale between Pipe A and Pipes B and C, Figures 5.8, 5.9, and 5.10 all have a 0.25 peak within an order of magnitude (between 2.5e4 and 3.0e4 counts). While the three plots are far from identical, there is no noticeable or predictable pattern in their differences.

5.4 Light Pipe Recommendation

All three options available meet the saturation requirement and cover a sufficiently large area of the tank, and the intensity of each is distributed over the wall region of tank in a similar fashion. It is therefore the recommendation of this thesis and its associated research to focus any further efforts on the logistics of each option, such as cost and transportation, to select the light pipe best suited for the HyperK project.

Chapter 6

Conclusions

The research detailed in this thesis presents a new method of modeling angular distribution in polar coordinates through a collection of two dimensional images from a camera. While this process can be adapted to many different uses, the intended use is to quantify the intensity distribution of light pipes that will emit calibration light into Water Cherenkov detectors of the Hyper-Kamiokande experiment. The algorithm created for this dissertation calculates the distributions in both θ and ϕ, confirms the distributions are independent of the position of the source relative to the location of photon impact, and condenses the distributions into normalized probability density functions for emission angle.

While a more robust and detailed simulation will be developed in the future, the generated probability density functions were fed into a rudimentary simulation of the IWCD tank to visualize the expected light emissions. This confirmed each light pipe covers a similar area of the tank interior and keeps each mPMT exposed within the dynamic range to avoid saturation. In addition, the intensity distributions of each pipe's emitted light are very similar. After ignoring contributions from the endcap regions, Figures 5.8 through 5.10 are too alike to justify one option over the others, and so this thesis recommends the HyperK team to consult other limiting factors to determine which light pipe to use.

As HyperK progresses towards data collection, this measurement procedure and algorithm will be incorporated by many other aspects of the project to help calibrate photogrammetry and timing light sources, such as the LEDs built into the structure of the tank outside the mPMT housings. This algorithm was purposefully developed to be independent of the exact specification of this research, so that it can be adapted and used by any project that requires light calibration in the future.

Bibliography

[1] Q. R. Ahmad et al. Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, 2002.

[2] B. T. Cleveland, Timothy Daily, Raymond Davis, Jr., James R. Distel, Kenneth Lande, C. K. Lee, Paul S. Wildenhain, and Jack Ullman. Measurement of the solar electron neutrino flux with the Homestake chlorine detector. *Astrophys. J.*, 496:505–526, 1998.

[3] Francesca Di Lodovico. The Hyper-Kamiokande Experiment. *J. Phys. Conf. Ser.*, 888(1):012020, 2017.

[4] Y. Fukuda et al. Atmospheric muon-neutrino / electron-neutrino ratio in the multiGeV energy range. *Phys. Lett. B*, 335:237–245, 1994.

[5] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, 1998.