THE UNIVERSITY OF CHICAGO

THE SEARCH FOR $K_L \to \pi^0\nu\overline{\nu}$ AND $K_L \to \pi^0\pi^0\gamma\gamma$ IN THE KOTO EXPERIMENT

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

BY

JOSEPH REDEKER

CHICAGO, ILLINOIS
AUGUST 2025

Copyright © 2025 by Joseph Redeker

All Rights Reserved

March on. Do not tarry. To go forward is to move toward perfection. March on, and fear not the thorns, or the sharp stones on life's path. -Khalil Gibran

ACKNOWLEDGMENTS

First and foremost I would like to thank my wife, Julia, for staying by my side through the hardest parts of my graduate studies. I could not have done this without the love and stability she puts into my life. I also want to thank my family, my parents and my sister, for always encouraging me to aim higher and not sell myself short.

This dissertation could not have been completed without the advice and help from my advisor and teacher, Yau Wah. He was always supportive of my work and encouraged me to explore my own ideas. I would also like to thank Chieh (Jay) Lin for being a great mentor for me as I worked through my graduate studies and being a good friend. The long lunch discussions on physics and politics was always a nice reprieve from work. As I was getting started in the group as an undergraduate, I would not have made it without the help of Yu-Chen Tung. Thank you for putting up with all of my questions. I am also grateful to Mario for and all the work and discussions on the DAQ and Level 3.

I am grateful to Koji Shiomi, Hajime Nanjo, and Tadashi Nomura for all of their help throughout the years on both analysis and the simulation library. I cannot quantify the amount your comments, suggestions, and advice I received helped me throughout my studies.

Lastly, I would like to thank my cat Daisy for her company while I am hard at work. Her companionship makes the hardest days a little easier.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iv

LIST OF FIGURES viii

LIST OF TABLES xiv

ABSTRACT xv

1 INTRODUCTION 1
- 1.1 Kaons and Strangeness 2
- 1.2 Neutral Kaon Mixing and CP-Violation 3

2 THE KOTO EXPERIMENT 11
- 2.1 Physics Motivation 11
- 2.2 $K_L \to \pi^0\nu\overline{\nu}$ and $K_L \to \pi^0\pi^0\gamma\gamma$ in the Standard Model 12
  - 2.2.1 Beyond Standard Model: $K_L \to \pi^0\pi^0 X$, where $X \to \gamma\gamma$ 14
- 2.3 KOTO Beamline and Detector 15
- 2.4 Experimental Principle 19
- 2.5 KOTO DAQ-Trigger System 21
  - 2.5.1 Stage-1 Trigger 22
  - 2.5.2 Stage-2 Trigger 23
- 2.6 Data Transfer 25
  - 2.6.1 Layer 1: ADC Crates 27
  - 2.6.2 Layer 2: OFC-I Crates 28
  - 2.6.3 Layer 3: OFC-II 30
- 2.7 Layer 4: The Dell Nodes and Late Stage Trigger 32
- 2.8 Error Monitoring 37

3 DATA COLLECTION PROCEDURE 38
- 3.1 Run Types 41
- 3.2 Run History 45

4 MONTE CARLO SIMULATION AND EVENT BUILDING 47
- 4.1 K_L Beamline Simulation 47
  - 4.1.1 GEANT3 simulation 48
  - 4.1.2 Empirical K_L simulation 49
- 4.2 GEANT4 and Detector Response Simulation 50
- 4.3 Accidental Overlay 53

5 EVENT RECONSTRUCTION . . . 55
5.1 Photon Energy and Hit Position Reconstruction . . . 55
5.2 Offline Clustering Algorithm . . . 57
5.3 $K_L \to \pi^0\nu\overline{\nu}$ Reconstruction . . . 58
5.4 $K_L \to \pi^0\pi^0\gamma\gamma$ Reconstruction . . . 59

6 DATA AND MONTE CARLO CONSISTENCY . . . 65
6.1 $K_L \to 3\pi^0$: Data vs. Monte Carlo . . . 65
6.2 Messy CSI Channels . . . 68
6.3 Systematic Error . . . 70
6.3.1 Veto Uncertainties . . . 70
6.3.2 Kinematic Uncertainties . . . 71
6.3.3 χ^2 Uncertainties . . . 71
6.3.4 Geometrical Uncertainties . . . 72
6.3.5 Trigger Uncertainties . . . 74
6.4 K_L Flux Calculation . . . 79

7 ANALYSIS OF $K_L \to \pi^0\pi^0\gamma\gamma$ AND $K_L \to \pi^0\pi^0X$ . . . 82
7.1 Simulation of $K_L \to \pi^0\pi^0\gamma\gamma$ . . . 82
7.1.1 $K_L \to \pi^0\pi^0X$ Simulation . . . 86
7.2 Measurement Principle . . . 86
7.3 $K_L \to 3\pi^0$ Background . . . 87
7.4 Event Selection Strategy . . . 89
7.5 Cut Optimization Strategy . . . 93
7.6 Control Region Study . . . 97
7.7 Single Event Sensitivity . . . 100
7.7.1 Systematic Uncertainties: A_{sig} . . . 101
7.7.2 Signal Acceptance and SES . . . 106
7.8 Results . . . 108
7.9 Discussion . . . 113

8 ANALYSIS OF $K_L \to \pi^0\nu\overline{\nu}$ . . . 118
8.1 Single Event Sensitivity . . . 119
8.2 History: 2016-2018 Data . . . 121
8.2.1 $K^\pm$ Background Rejection . . . 122
8.2.2 Halo $K_L \to 2\gamma$ Background Rejection . . . 129
8.3 Background Analysis: 2021 . . . 135
8.3.1 $K_L \to 2\pi^0$ Background . . . 136
8.3.2 Upstream π^0 . . . 142
8.3.3 Background Table: Summary . . . 142
8.3.4 Results . . . 143

9 CONCLUSION AND DISCUSSION 146
9.1 Interpretation of the Results 146
9.2 Potential Improvements 150

A LIST OF CUT CONDITIONS 152
A.1 Veto 152
A.2 Kinematic 153
A.3 Additional 155

B TABLE OF $K_L \to \pi^0\pi^0 X$ RESULTS 157

BIBLIOGRAPHY 158

LIST OF FIGURES

1.1 The two second order Feynman diagrams for neutral kaon transitions. 3

2.1 An overhead view of the experimental site. 11
2.2 The CKM triangle, and the relationship of its sides and height to notable kaon decays. 12
2.3 Dominant tree diagrams for the decay $K_L \to \pi^0\nu\overline{\nu}$ 13
2.4 Feynman Diagram of the CP-violating contribution to $K_L^0 \to 2\pi^0 l^- l^+$ from $K_L \to \pi^0\pi^0\gamma\gamma$ via internal $\gamma\gamma \to l^- l^+$ 14
2.5 The layout of the KOTO beamline. 16
2.6 A cutaway view of the KOTO detector. The K_L beam arrives from the left. . . 17
2.7 A representation of the front of the KOTO CSI, with the edge vetoes CC03 and Outer Edge Veto (OEV) [43]. 18
2.8 On the left is an individual MB module while on the right is a diagram of all modules assembled in the vacuum tank. The edges of each individual module point away from the beam center [52]. 19
2.9 A simple schematic showing the K_L beamline with an example $K_L \to \pi^0\nu\overline{\nu}$ decay. 20
2.10 $K_L \to \pi^0\nu\overline{\nu}$ Monte Carlo simulation using GEANT4. The red box signifies the signal region. 20
2.11 An overview of the KOTO trigger system. 21
2.12 A diagram of the daisy chain strategy to send ADC data for trigger decisions to the local CDT. 23
2.13 A simple example of the turn counting strategy to determine the number of clusters. 24
2.14 This table shows all different types of 2x2 sections, and the equivalent number of turns 25
2.15 Overview of the KOTO data transfer system from the ADC to Dell servers located off the experimental area. 26
2.16 The format for the ADC headers. They consist of six 16 bit words that contain the timestamp information, spill number, event number, cluster number, and the crate and slot number of the ADC 28
2.17 Strategy to output the data at the OFC-I stage. Even numbered events go to RAM-A, while odd numbered events go to RAM-B. 29
2.18 This plot shows the event loss due to OFC-I buffer full errors at different rates for several tests. 30
2.19 The event builder logic in the OFC-2. RX in the figure represents the remainder (X) modulo ten. 31
2.20 Control logic between the event builder and the 40 Gbps interface 32
2.21 Data capture from the Network Interface Card (NIC) to the RAM. 33
2.22 An example waveform from ADC data, with 64 samples. The first sample is the result from the online pedestal calculation. 34
2.23 A representation of the compression algorithm [27] 36

2.24 The total data reduction based on data taken in the Spring of 2024 [27]. The trigger type represents the mode in which the specific trigger conditions are designed for. 36

3.1 A schematic of the timing between the LIVE and Spill signals. A full LIVE or Spill cycle is 4.2 seconds in the most recent data taking runs. 38
3.2 A schematic of the timing between the rising edge of the LIVE and Spill signals. The optical link check, local timing alignment, and global timing alignment are done between the LIVE rising edge and the Spill rising edge. 39
3.3 An example of the waveform generated when an ADC receives the delta trigger. 40
3.4 The percentage distribution of the type of data taking runs during Run87 41
3.5 Cosmic muon trigger table from Run87. 42
3.6 The detector layout when the Z0 aluminum target is lowered. It sits upstream of the barrel counters. 43
3.7 The trigger table for physics runs taken during Run87. The most relevant triggers are highlighted as follows: $K_L \to \pi^0\nu\overline{\nu}$ trigger (red), normalization trigger (black), $K^{\pm}$ dedicated trigger (blue), $K_L \to 2\pi^0\gamma\gamma$ trigger (green) 44
3.8 The run history in terms of beam intensity and protons on target (POT) up to the year 2024. Data already analyzed for $K_L \to \pi^0\nu\overline{\nu}$ is labeled with a check mark. 45

4.1 The result of beam x-y position at the entrance of the KOTO detector based on GEANT3 simulation [44] 48
4.2 Relative amount of different particles at the entrance of the KOTO detector. Figure curtesy of Chieh Lin. [33] 49
4.3 Empirical K_L momentum spectrum obtained from engineering runs taken in 2012 analyzing $K_L \to \pi^+\pi^-$ using drift chambers. The y-axis is in arbitrary units and the x-axis is in GeV. 50
4.4 A visual representation of the GEANT4 particle tracking and generation from a 2 GeV electron shower in liquid Argon [50]. 51
4.5 The fitting of the attenuation function for a specific channel in the main barrel from data taken in 2013 and fitted by Takahiko Masuda in an internal meeting. 52
4.6 An example of an overlapped waveform. The high energy accidental hit overlaps a lower energy on time hit to mask the real timing. 54

5.1 An example of the linear fit done to evaluate the calibration constant for a CSI channel (module ID == 19). The linear fit is shown in red, and uses small peak heights for the best result. 56
5.2 Diagram explaining the offline clustering algorithm. The blue blocks represent a cluster of hit channels with energy deposit above 3 MeV. The red crystal is isolated, because its neighbor has energy deposit less than 3 MeV. The green colored crystals have a time measurement outside the timing window. 58
5.3 The improvement in $M_{\gamma\gamma}$ resolution after implementing the constrained fit algorithm. Shown in blue is the $K_L \to 3\pi^0$ simulation, and shown in red is the $K_L \to \pi^0\pi^0 X$, where $X \to \gamma\gamma$, simulation with an X mass of 200 MeV. 64

6.1 The comparison of $K_L \to 3\pi^0$ data and Monte Carlo. All kinematic cuts, except for the one shown in the histogram, are applied. Each histogram is normalized to one. 66
6.2 The comparison between exclusive veto efficiency in data (blue) and $K_L \to 3\pi^0$ (red) and for $K_L \to \pi^0\pi^0\gamma\gamma$ (green). 68
6.3 The hit position of γ_5 and γ_6 after $K_L \to \pi^0\pi^0\gamma\gamma$ reconstruction. The χ^2_{fit} with the $K_L \to \pi^0\pi^0\gamma\gamma$ assumption is small, while χ^2_{fit} with the $K_L \to 3\pi^0$ assumption is large. 69
6.4 The exclusive acceptance of the kinematic cuts used in the K_L flux calculation. 71
6.5 The exclusive acceptance of the χ^2 cuts used in the K_L flux calculation. 72
6.6 Comparison of the K_L momentum distribution before and after geometrical cuts requiring all six photons from the $K_L \to 3\pi^0$ decay incident on the CSI. 74
6.7 Distribution of Δ over 50,000 iterations of new parameters. The boundary shows the $\pm 1\sigma$ region and corresponds to the systematic error result. 75
6.8 The efficiency as a function of offline energy calculation. The vertical red line corresponds to the online cut value of 650 MeV. 76
6.9 The CDTNum distribution for data and Monte Carlo after all cuts 78

7.1 $M_{\gamma\gamma}$ Monte Carlo distribution. A cut requiring $|M_{\gamma\gamma} - m_{\pi^0}| < 10$ is applied to avoid the pole at the pion mass. 85
7.2 The KOTO simulation result overlaid on the Kambor-Funck distributions [24]. The cut value corresponds to the width of the $M_{\gamma\gamma}$ precut around the pion mass. 85
7.3 Green represents the barrel veto detectors (IB, CBAR, and FBAR) and the NCC. An example $K_L \to \pi^0\pi^0 X$ signal decay with all six final state photons incident in the CSI. 86
7.4 Overlapping histograms showing $\chi^2_{fit}(3\pi^0)$ and $\chi^2_{fit}(2\pi^0\gamma\gamma)$ results of $K_L \to 3\pi^0$ and $K_L \to \pi^0\pi^0 X$ Monte Carlo. The vertical red line indicates an example cut of $\chi^2_{fit}(3\pi^0) > 115$ and $\chi^2_{fit}(2\pi^0\gamma\gamma) < 10$. An X mass of 190 MeV is assumed in this Monte Carlo. 87
7.5 Background distribution with an example of the signal $K_L \to \pi^0\pi^0 X$ Monte Carlo with X mass of 190 MeV overlaid underneath. The blue box represents the low mass region which is treated as a control region, and the red box is the signal region. 88
7.6 The MinWFourierLH variable for well reconstructed $K_L \to 3\pi^0$ events. An event is well reconstructed if $\chi^2_{fit}(3\pi^0) < 1$. The vertical red line represents the cut value which attains 90% acceptance. 90
7.7 The MinWFourierLH rejection factor in the high mass region requiring a large $\chi^2_{fit}(3\pi^9)$ value. The error bars represent the statistical uncertainty in the rejection factor. 91
7.8 The MinCSDDLVal variable for background and signal Monte Carlo. The background is $K_L \to 3\pi^0$ events with $\chi^2_{fit}(3\pi^0) > 100$, and the signal is $K_L \to \pi^0\pi^0 X$ Monte Carlo with an X mass of 220 MeV. 92
7.9 The comparison between data and Monte Carlo under loose and tight cut conditions. 92

7.10 The Gaussian fits which determine the ROI for these analyses. 94
7.11 A example comparison between Data and $K_L \to 3\pi^0$ Monte Carlo in the low mass region under loose cut conditions. 98
7.12 These plots show the correlation between the background level scale factor calculation and the two most important cuts in the analyis, $\chi^2_{fit}(3\pi^0)$ and MinCSDDLVal. The error bars represent the statistical uncertainty in the ratio. 99
7.13 The background level scale factors for each analyses. The point labeled at an X mass of 240 MeV represents the scale factor for the $K_L \to \pi^0\pi^0\gamma\gamma$ analysis. . . . 99
7.14 The efficiency of WFourierLH > 0.19 cut condition as a function of photon energy in $K_L \to 3\pi^0$ decays. The error bars represent the statistical uncertainty. 101
7.15 The exclusive acceptance for MinWFourierLH, MinCSDDLVal and $\chi^2_{fit}(2\pi^0\gamma\gamma)$. 102
7.16 The Δ distribution for different resolution parameters, assuming $\chi^2_{fit}(2\pi^0\gamma\gamma) < 3$ The two vertical red lines represent the $\pm 1\sigma$ region. 104
7.17 The Δ distribution for different resolution parameters assuming $\chi^2_{fit}(3\pi^0) > 115$. The two vertical red lines represent the $\pm 1\sigma$ region. 105
7.18 The signal acceptance as a function of X mass. 106
7.19 The single event sensitivity as a function of X mass. 107
7.20 The results for each analysis. The $K_L \to 3\pi^0$ and $K_L \to \pi^0\pi^0\gamma\gamma$ Monte Carlo are included, and the expected number for each are shown in red and blue respectively. The signal region is denoted by the red box in the high mass region, and the expectation is denoted by the result labeled "SR". The expectation in the high mass region outside the signal region is labeled "Outer". 108
7.21 The upper limits set using the Feldman-Cousins approach. Due to the two events observed in the $M_X = 180$ MeV analysis, there is a non-zero lower limit on the signal rate at 90% C.L. 111
7.22 Comparison of Monte Carlo kinematic distribution and the two leftover events from the analysis with $M_X = 180$ MeV. Loose cut conditions are applied to observe the distributions. 113
7.23 Comparison of Monte Carlo kinematic distribution and the two leftover events from the analysis with $M_X = 180$ MeV. Loose cut conditions are applied to observe the distributions. All cuts except MinWFourierLH are applied with loose cuts, $\chi^2_{fit}(3\pi^0) > 50$, $\chi^2_{fit}(2\pi^0\gamma\gamma) < 10$, and MinCSDDLVal > 0.5. 115

8.1 The reconstructed K_L mass after the $K_L \to 2\pi^0$ selection criteria for the SES calculation [4]. 120
8.2 The results from the analysis of data taken in 2016-2018 [2]. The dotted region represents the signal region. 122
8.3 A potential background event from $K^\pm \to \pi^0 e^\pm \nu$. The Upstream Charged Veto (UCV) labeled here was not installed until 2020. Sizes are not to scale. 123
8.4 An image of the UCV installed before data taking in 2021 to reduce the background contamination from $K^\pm$. 123
8.5 The reconstructed kinematic variables for $K_L \to \pi^0\nu\overline{\nu}$ and $K^\pm$ background with an assumed kinematic constraints on the $e^\pm$. 124

8.6 The two distributions for both $K^{\pm} \to \pi^0 e^{\pm} \nu$ background and $K_L \to \pi^0 \nu\overline{\nu}$ signal Monte Carlo. Note that the signal tends to lie on the downstream side of the signal region. . . . 126
8.7 The model performance on $K^{\pm}$ background and signal Monte Carlo. . . . 127
8.8 The model performance on D1-off $K^{\pm}$ background in data and Monte Carlo. Only the blinded region was shown in these plots, with the number of events shown in red. . . . 128
8.9 A potential background event from $K_L \to 2\gamma$. Sizes are not to scale. . . . 129
8.10 A diagram showing the two assumptions made for the discrimination of Halo $K_L \to 2\gamma$, via the likelihood ratio. Plot generated by Yuya Noichi in an internal collaboration meeting. . . . 130
8.11 The results of the likelihood study on background and signal Monte Carlo. Plot generated by Yuya Noichi in an internal collaboration meeting. . . . 130
8.12 The results of the TMVA considering reconstructed variables. Plot generated by Yuya Noichi in an internal collaboration meeting . . . 131
8.13 The kinematic distributions for Halo $K_L \to 2\gamma$ background and signal Monte Carlo. 132
8.14 The two distributions for both halo $K_L \to 2\gamma$ background and $K_L \to \pi^0 \nu\overline{\nu}$ signal Monte Carlo. . . . 134
8.15 The model performance on Halo $K_L \to 2\gamma$ background and signal Monte Carlo. 135
8.16 A potential background event from $K_L \to 2\pi^0$. Sizes are not to scale. The red X near the γ represents a "missed" γ. . . . 136
8.17 The kinematic distributions for $K_L \to 2\pi^0$ background and signal Monte Carlo. 138
8.18 The input parameters importance in the model output. . . . 139
8.19 The comparison between signal and background Monte Carlos for the two most important input variables. These are compared to the leftover events from the full simulation Monte Carlo done in KEKCC and OSG. . . . 140
8.20 The two distributions for both $K_L \to 2\pi^0$ background and $K_L \to \pi^0 \nu\overline{\nu}$ signal Monte Carlo. . . . 140
8.21 The model performance for the $K_L \to 2\pi^0$ background and $K_L \to \pi^0 \nu\overline{\nu}$ signal Monte Carlo. . . . 141
8.22 A potential background event from an upstream π^0 generated from an interaction in the NCC. Sizes are not to scale. . . . 142
8.23 The final results of the 2021 data analysis. Red numbers indicate the expectation in that region from background Monte Carlo simulation. Black numbers represent the number observed in data. . . . 144
8.24 The final upper limit at 90% C.L. results for the analysis of $K_L \to \pi^0 X^0$ where X^0 is invisible. A mass range of 1-260 MeV is studied. . . . 145

9.1 The final upper limit at 90% C.L. results for the analysis of $K_L \to \pi^0 \pi^0 X$. The bars represent the respective signal regions for each X mass analysis, and the black dots show the invariant mass of events observed within their signal region. The continuous line represents the weighted upper limit calculation. . . . 147
9.2 The mean number observed assumed for the toy MC study as a function of test mass x_m. . . . 149

9.3 The upper limit result overlayed with the toy MC result. The one sigma bound on the averaged upper limit result is shown in the shaded red. 150
9.4 The likelihood upper limit result overlaid with the toy MC result. 151

LIST OF TABLES

6.1 The summary of systematic errors associated with the K_L yield calculation. The total is the quadratic sum of each of the sources. 78

6.2 This table contains all the cut values and variables used in the K_L yield calculation for the $K_L \to \pi^0\pi^0\gamma\gamma$ analysis. In addition to those listed here are "quality" cuts that remove events that are not the correct trigger type, or if there was a DAQ error during the spill. 81

7.1 This table contains all the fit parameters, the signal region (ROI) for each of the analyses, and the optimized cut values. All units are in MeV in the first four columns. The last column is the cut values obtained for the MinCSDDLVal cut. 97

7.2 This table contains all the the background level scale factor for all the analyses. ΔF is determined by the statistical error in the data to $K_L \to 3\pi^0$ Monte Carlo ratio. 100

7.3 This table contains the systematic error based on the $\chi^2_{fit}(3\pi^0)$ cut condition. In general, the larger the cut value then the larger the systematic uncertainty. . . . 105

7.4 This table contains the signal acceptances and SES for each analysis. The error is dominated by systematics, in particular related to $\chi^2_{fit}(3\pi^0)$, and not Monte Carlo statistics. 107

7.5 This table contains the 90% C.L. limits on the branching ratio for each analysis, and the number observed in data. 112

7.6 This table contains information regarding the three leftover events. Row one contains the event observed in two analyses. 113

8.1 The summary of systematic errors associated with the SES calculation. The total is the quadratic sum of each of the sources. 120

8.2 This table contains the background estimation in the 2016-2018 signal region. Rows marked with ^{a} represent background studied after opening the signal region. 121

8.3 This table contains the number and type of $K_L \to 2\pi^0$ background observed in simulation. 137

8.4 This table contains the background estimation in the 2021 signal region. The ordering is from largest to smallest, and backgrounds limited to upper limits are not included [4]. 143

A.1 List of Veto Detectors and their corresponding energy thresholds and timing window widths. Timing windows are set to prevent masked pulses being missed, consider possible time of flight hits, and the detector timing resolutions. 153

A.2 The modified veto conditions used in the $K_L \to \pi^0\nu\overline{\nu}$ analysis 154

B.1 The result of the weighted upper limit calculation for 50 test masses across the high mass region. 157

ABSTRACT

The KOTO experiment at J-PARC aims to study CP-symmetry breaking in kaon decays. In particular, KOTO searches for the rare decay, $K_L \to \pi^0\nu\overline{\nu}$, which is sometimes called the "Golden Mode". The mode is direct CP-violating, and with a small theoretical uncertainty, this search is highly sensitive to new physics. During the analysis of 2021 data, no events were observed in the signal region, which allowed KOTO to set the best upper limit on BR($K_L \to \pi^0\nu\overline{\nu}$) to date at $< 2.01 \times 10^{-9}$ (90% C.L.). This dissertation will include details on the analysis of $K_L \to \pi^0\nu\overline{\nu}$ in 2021 data.

Due to the high sensitivity required to study the $K_L \to \pi^0\nu\overline{\nu}$ decay, KOTO is in a unique position to study rare K_L interactions, such as sub-GeV quark coupling to dark matter. One avenue to study this is the mode $K_L \to \pi^0\pi^0 X$, where $X \to \gamma\gamma$. This mode was studied in the E391a experiment at KEK in the X mass region 194.3-219.3 MeV, with the best upper limit on the branching ratio set with an X mass of 214.3 MeV at $< 2.4 \times 10^{-7}$ [48]. In the analysis of 2021 data, the scope is increased to include X mass in the range 160-220 MeV, and the first search for the rare decay $K_L \to \pi^0\pi^0\gamma\gamma$ which is described by ChPT [24]. Two events were observed in the signal region near an X mass of 177 MeV, and no events were observed for the analysis of $K_L \to \pi^0\pi^0\gamma\gamma$, setting a world first upper limit of BR($K_L \to \pi^0\pi^0\gamma\gamma$) $< 9.77 \times 10^{-7}$ at 90% C.L.

CHAPTER 1

INTRODUCTION

One of the most difficult and glaringly human existential questions we may ask ourselves is: why are we here? From the perspective of a physicist, the search for that answer begins at the origin of time itself. At the beginning of time, matter and antimatter were created. If the laws of the universe treated both types of matter equivalently, then equal amounts of matter and antimatter would be uniformly distributed throughout space. As the universe expanded and cooled, we would expect all matter and anti-matter to annihilate. Yet, here we are in a universe which is predominantly matter.

The clear explanation is that the laws of the universe do not treat matter and anti-matter equally, but instead have a preference towards the creation of matter, called Baryogenesis [49]. For Baryogenesis to occur, it must satisfy the Sakharov conditions [42]:

- Baryon number violation
- Violation of charge and charge-parity symmetry (CP-violation)
- Interactions out of thermal equilibrium

In order to explain this phenomenon, it must invoke CP-violating processes [28]. Thus, exploring the nature of CP-violation in kaons [14] may lead to a better understanding of the nature of our universe and its creation. Alternative theories that may explain both Baryogenesis and dark matter [47] consist of a favored production of a combination of mesons and dark anti-mesons [26] which would satisfy the Sakharov conditions. Potential models include the creation of dark matter via a decay channel through the kaon, with the end state signature of the dark particle being two photons [31]. The purpose of this thesis is to introduce and evaluate the kaon decay channels $K_L \to \pi^0 \nu\overline{\nu}$, and $K_L \to \pi^0\pi^0 X$ where X is a dark particle in the mass range 160 MeV - 220 MeV, and promptly decays $X \to \gamma\gamma$.

This mode is analogous to the rare decay, $K_L \to \pi^0\pi^0\gamma\gamma$ which is vital to the understanding Chiral Perturbation Theory (ChPT) [24]. This analysis is the first search for such a mode, and it may be observed in this experiment with a predicted branching ratio on the order of 10^{-8}. These modes are studied at the KOTO experiment at J-PARC.

1.1 Kaons and Strangeness

Strange particles which appeared to have the same mass, about half the mass of a proton, were observed both in 1947 by G.D. Rochester and C.C. Butler [41] and in 1949 by Rosemary Brown [11]. The confounding observation of this particle is that Rochester and Butler observed the decay into two pions, with parity +1, while Brown observed a decay into three pions, with party -1. This led to what was called the "tau-theta puzzle" [20]. Two particles, with different parity, were postulated to exist with the same long lifetime, spin, charge and mass.

This was not the only puzzle these particles created. The relatively long lifetime on the order of 10^{-10} s differed from other known pion-proton interactions which occurred on the order of 10^{-23} s. This was later explained by Abraham Pais and Murray Gell-Man by introducing a new conserved quantity in particle interactions called "strangeness" [18]. This quantity would only be conserved through the strong force, but not the weak force. Thus, the new particles' long lifetime was explained by needing to decay weakly, and the name "strange" became associated with these new particles.

Around the same time, others began to question the long held belief in unbroken parity symmetry. This led to an investigation by C.S. Wu [51], who discovered parity violation occurs in weak interactions. Ultimately, both of these discoveries and ideas led to the resolution of the tau-theta puzzle. The two new particles were actually the same particle, characterized by its strangeness, and would later be called the charged K-meson.

This finding led to the discovery of four light particles which all have non-zero strangeness,

the K-mesons, also called kaons. There are four combinations of quarks which make up a kaon: K^+ ($\overline{s}u$), K^- ($\overline{u}s$), K^0 ($\overline{s}d$), and $\overline{K}^0$ ($\overline{d}s$). Note that the neutral kaon is not its own antiparticle.

1.2 Neutral Kaon Mixing and CP-Violation

It is crucial to first understand the concept of neutral kaon mixing in order to put into context the discovery and importance of CP-violation in kaon decays. Because K^0 ($\overline{s}d$), and $\overline{K}^0$ ($\overline{d}s$) are two distinct strange eigenstates differing by two units of strangeness, they may transition from one to the other via the weak force as shown in Figure 1.1.

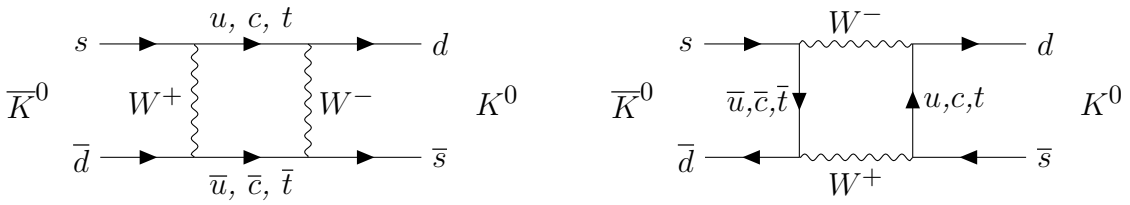


Figure 1.1: The two second order Feynman diagrams for neutral kaon transitions.

However, it may be noted that these strange eigenstates are *not* CP eigenstates. If we would like to consider a "physical" particle, one with definite mass and lifetime, then perhaps a more physical interpretation of the neutral kaon would be in terms of its weak eigenstates, as decays occur through the weak interaction [25]. Assuming CP is conserved in the weak interaction, then the CP eigenstates are equivalent to the weak eigenstates and can be defined simply as,

$$\begin{aligned} |K_1^0\rangle &= \frac{1}{\sqrt{2}}(|K_0\rangle + |\overline{K}_0\rangle) \\ |K_2^0\rangle &= \frac{1}{\sqrt{2}}(|K_0\rangle - |\overline{K}_0\rangle) \end{aligned} \tag{1.1}$$

Where the eigenvalues are +1 and -1 respectively,

$$\begin{aligned} CP|K_1^0\rangle &= \frac{1}{\sqrt{2}}(CP|K_0\rangle + CP|\overline{K}_0\rangle) \\ &= \frac{1}{\sqrt{2}}(|\overline{K}_0\rangle + |K_0\rangle) = |K_1^0\rangle \\ CP|K_2^0\rangle &= \frac{1}{\sqrt{2}}(|\overline{K}_0\rangle - |K_0\rangle) = -|K_2^0\rangle \end{aligned} \tag{1.2}$$

It is clear that since K_1^0 is CP-even, it will decay into only two pions and never three, due to the fact that pions have negative parity, while the opposite is true for K_2^0. Then the primary, non-leptonic decays of K_1^0 and K_2^0 are,

$$\begin{aligned} K_1^0 &\to \pi^0\pi^0 \qquad \& \qquad K_1^0 \to \pi^+\pi^- \\ K_2^0 &\to \pi^0\pi^0\pi^0 \qquad \& \qquad K_2^0 \to \pi^+\pi^-\pi^0 \end{aligned} \tag{1.3}$$

K_1^0 should have a definite lifetime that is much shorter than K_2^0 since the allowed phase space in the two-pion decay is larger. Physically, this means that the process of kaon decays may be better understood by studying the weak eigenstates.

Derivation of the Weak Eigenstates

The weak eigenstates of the neutral kaon are not obtained trivially. Of course, if the CP-symmetry of the weak force is assumed then the CP-eigenstates are equivalent to the weak eigenstates. However, if this assumption is withheld, an alternative form of the neutral kaon is obtained. These are typically labeled K_L and K_S, called "K-long" and "K-short" for their respective lifetimes. The derivation of the weak eigenstates is evaluated rigorously in this section, and is heavily inspired by the study of Mark Thomson's textbook on particle physics [46].

To start, it is useful to observe the time-dependent waveform of the strong eigenstate. For a K^0 that does not mix, this waveform is given by,

$$|K^0(t)\rangle = e^{-\Gamma t/2 - iM_K t}|K^0\rangle \tag{1.4}$$

where Γ is the lifetime of the unstable kaon, and M_K is the mass. Therefore, $|K^0(t)\rangle$ must satisfy,

$$i\frac{\delta}{\delta t}|K^0(t)\rangle = (m - \frac{i}{2}\Gamma)|K^0(t)\rangle \tag{1.5}$$

and the effective Hamiltonian is then,

$$H|K^0(t)\rangle = (m - \frac{i}{2}\Gamma)|K^0(t)\rangle \tag{1.6}$$

However, this is actually a mixing system, where a K^0 will gain a $\overline{K}^0$ component over time. This results in the time evolution of a neutral kaon system which takes the form,

$$|K(t)\rangle = a(t)|K^0\rangle + b(t)|\overline{K}^0\rangle \tag{1.7}$$

Following a similar strategy as above, the effective Hamiltonian is derived from the coupled equations,

$$\begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{21} - \frac{i}{2}\Gamma_{21} & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix} \begin{pmatrix} a(t)|K^0\rangle \\ b(t)|\overline{K}^0\rangle \end{pmatrix} = i\frac{\delta}{\delta t} \begin{pmatrix} a(t)|K^0\rangle \\ b(t)|\overline{K}^0\rangle \end{pmatrix} \tag{1.8}$$

and takes the form,

$$\hat{\mathbf{H}} = \hat{\mathbf{M}} - \frac{i}{2}\hat{\mathbf{\Gamma}} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \tag{1.9}$$

$\hat{\mathbf{\Gamma}}$ may be obtained simply by Fermi's Golden rule at lowest order, where $\hat{\mathbf{H}}_W$ represents the weak Hamiltonian.

$$\Gamma = 2\pi \sum_s \langle K(t)|\hat{\mathbf{H}}_W|s\rangle\langle s|\hat{\mathbf{H}}_W|K(t)\rangle\rho_s \tag{1.10}$$

By writing $K(t)$ in terms of K^0 and $\overline{K}^0$ there are four terms which each represent the terms in the Γ operator. Each of these Γ's contain terms of $|K^0\rangle$ and $|\overline{K}^0\rangle$, and for a final state s, they are represented by,

$$\Gamma_{11} = |a(t)|^2|\langle s|\hat{\mathbf{H}}_W|K^0\rangle|^2 \tag{1.11}$$

$$\Gamma_{22} = |b(t)|^2|\langle s|\hat{\mathbf{H}}_W|\overline{K}^0\rangle|^2 \tag{1.12}$$

$$\Gamma_{12} = a(t)b^*(t)\langle\overline{K}^0|\hat{\mathbf{H}}_W|s\rangle\langle s|\hat{\mathbf{H}}_W|K^0\rangle \tag{1.13}$$

$$\Gamma_{12} = a^*(t)b(t)\langle K^0|\hat{\mathbf{H}}_W|s\rangle\langle s|\hat{\mathbf{H}}_W|\overline{K}^0\rangle \tag{1.14}$$

Notice that $\Gamma_{21} = \Gamma_{12}^*$, and the Γ matrix is Hermitian. The mass matrix is slightly more complicated, as the masses are obtained from not only the masses of the quarks, but also the potential energy of the system. The potential energy contains interaction terms from the strong interaction, EM interaction, and weak interactions in the system. For example, the mass terms for K_0 and $\overline{K}_0$ are,

$$M_{11} = M_{\overline{s}} + M_d + \langle K^0|\hat{\mathbf{H}}_{QCD} + \hat{\mathbf{H}}_W + \hat{\mathbf{H}}_{EM}|K^0\rangle + O(G_F^2) + ... \tag{1.15}$$

$$M_{22} = M_s + M_{\overline{d}} + \langle\overline{K}^0|\hat{\mathbf{H}}_{QCD} + \hat{\mathbf{H}}_W + \hat{\mathbf{H}}_{EM}|\overline{K}^0\rangle + O(G_F^2) + ... \tag{1.16}$$

The mass interference terms are obtained from interactions in the mixing diagrams in Fig. 1.1. The only surviving terms must be at least 1st order, since there are two weak interaction points in the mixing diagram. Thus, to obtain the mass interference terms there must be a

sum over intermediate states s. Explicitly these terms are,

$$M_{12} = \sum_s \frac{\langle \overline{K}^0 | \hat{\mathbf{H}}_W | s \rangle \langle s | \hat{\mathbf{H}}_W | K^0 \rangle}{E_s - M} \tag{1.17}$$

$$M_{21} = \sum_s \frac{\langle K^0 | \hat{\mathbf{H}}_W | s \rangle \langle s | \hat{\mathbf{H}}_W | \overline{K}^0 \rangle}{E_s - M} \tag{1.18}$$

Again, notice that $M_{12} = M_{21}$ and the mass matrix is Hermitian, which implies that the Hamiltonian is Hermitian. Finally, if CPT symmetry is assumed for Standard Model interactions, then the masses and lifetime of the strong eigenstates K^0 and $\overline{K}^0$ must be the same. This allows us to write the Hamiltonian as,

$$\hat{\mathbf{H}} = \begin{pmatrix} M & M_{12} \\ M_{12}^* & M \end{pmatrix} - \frac{i}{2} \begin{pmatrix} \Gamma & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma \end{pmatrix} \tag{1.19}$$

Solving the eigenvalue equation by setting $\det(\hat{\mathbf{H}} - \lambda \hat{\mathbf{I}}) = 0$ gives two solutions for λ.

$$\lambda_\pm = M - \frac{i}{2}\Gamma \pm \left[(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)(M_{12} - \frac{i}{2}\Gamma_{12}) \right]^{\frac{1}{2}} = 0 \tag{1.20}$$

If M_{12} and Γ_{12} are real, then substituting these values back into the eigenvector equation for the Hamiltonian yields,

$$\vec{\lambda}_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad , \qquad \vec{\lambda}_- = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{1.21}$$

The weak eigenstates can be obtained by using the eigenvectors to construct the normalized

change of basis matrix.

$$\begin{pmatrix} |K_S\rangle \\ |K_L\rangle \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} |K^0\rangle \\ |\overline{K}^0\rangle \end{pmatrix} \tag{1.22}$$

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} |K^0\rangle + |\overline{K}^0\rangle \\ |K^0\rangle - |\overline{K}^0\rangle \end{pmatrix} \tag{1.23}$$

These are exactly K_1 and K_2 that were derived earlier in Eq. 1.1. However, here the only assumption was that M_{12} and Γ_{12} were real. The connection between CP being an exact symmetry of the weak force and M_{12} and Γ_{12} being real is related to the elements of the Cabibbo–Kobayashi–Maskawa (CKM) matrix. This unitary matrix is a representation of transition probability between quark flavour states via the weak interaction. *If* the elements of the CKM matrix are real, then M_{12} and Γ_{12} must also be real which implies the weak eigenstates are equivalent to the CP eigenstates.

$$\text{CP symmetry in the weak force} \iff \text{CKM matrix is real} \iff M_{12} \text{ and } \Gamma_{12} \text{ are real} \tag{1.24}$$

If the CKM matrix has imaginary terms, then the forms for the weak eigenstates must be adjusted. With a bit of algebra, the eigenvector equation yields,

$$\begin{pmatrix} M - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M^*_{12} - \frac{i}{2}\Gamma^*_{12} & M - \frac{i}{2}\Gamma \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix} = M - \frac{i}{2}\Gamma + \left[(M^*_{12} - \frac{i}{2}\Gamma^*_{12})(M - \frac{i}{2}\Gamma) \right]^{1/2} \begin{pmatrix} v \\ w \end{pmatrix} \tag{1.25}$$

$$\begin{pmatrix} (M - \frac{i}{2}\Gamma)v + (M_{12} - \frac{i}{2}\Gamma_{12})w \\ (M^*_{12} - \frac{i}{2}\Gamma^*_{12})v + (M - \frac{i}{2}\Gamma)w \end{pmatrix} = M - \frac{i}{2}\Gamma + \left[(M^*_{12} - \frac{i}{2}\Gamma^*_{12})(M - \frac{i}{2}\Gamma) \right]^{1/2} \begin{pmatrix} v \\ w \end{pmatrix} \tag{1.26}$$

The top and bottom elements give two coupled equations, and after some simplification the ratio of v and w takes the form,

$$\frac{v}{w} = \pm \left[\frac{M_{12} - \frac{i}{2}\Gamma_{12}}{M_{12}^* - \frac{i}{2}\Gamma_{12}^*} \right]^{1/2} \tag{1.27}$$

Calling this ratio ξ, the normalized eigenvectors are then,

$$\vec{\lambda}_+ = \frac{1}{\sqrt{1+|\xi|^2}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} \qquad , \qquad \vec{\lambda}_- = \frac{1}{\sqrt{1+|\xi|^2}} \begin{pmatrix} 1 \\ -\xi \end{pmatrix} \tag{1.28}$$

and the changing basis from the strong eigenstates to the weak eigenstates gives,

$$\begin{aligned} |K_S\rangle &= \frac{1}{\sqrt{1+|\xi|^2}}(|K^0\rangle + \xi|\overline{K}^0\rangle) \\ |K_L\rangle &= \frac{1}{\sqrt{1+|\xi|^2}}(|K^0\rangle - \xi|\overline{K}^0\rangle) \end{aligned} \tag{1.29}$$

If the imaginary terms in the CKM matrix are small relative to their real parts, then $\xi \approx 1$. In that case, it can be useful to write $\xi = \frac{1-\epsilon}{1+\epsilon}$, and to write the time dependence of the weak eigenstates in terms of the CP eigenstates.

$$\begin{aligned} |K_S\rangle &= \frac{1}{\sqrt{1+|\epsilon|^2}}(|K_1\rangle + \epsilon|K_2\rangle)e^{-i\lambda_S t} \\ |K_L\rangle &= \frac{1}{\sqrt{1+|\epsilon|^2}}(|K_1\rangle - \epsilon|K_2\rangle)e^{-i\lambda_L t} \end{aligned} \tag{1.30}$$

Each K_S and K_L are the physical particles which both have distinct lifetimes and masses, so λ_S and λ_L are then,

$$\lambda_{S/L} = \left(M \mp \frac{|\Delta M|}{2} \right) - \frac{i}{2}\left(\Gamma \pm \frac{|\Delta\Gamma|}{2} \right) \tag{1.31}$$

If ϵ is non-zero, then there is some CP-violating part in the weak eigenstates. This is

called "indirect" CP-violation, and arises due to small differences in $K^0 \leftrightarrow \overline{K}^0$ and $\overline{K}^0 \leftrightarrow K^0$ mixing. The magnitude of the indirect CP-violation parameter is currently estimated to be,

$$|\epsilon| = (2.228 \pm 0.011) \times 10^{-3} \tag{1.32}$$

which is obtained from the Particle Data Group (PDG) who evaluate central values based on a fit of all experimental measurements of a parameter [36]. "Direct" CP-violation is also possible and can be characterized, for example, by the $|K_2\rangle$ (CP-odd) component of the K_L directly decaying into two pions (CP-even). The magnitude of this effect is called ϵ' and is characterized by the ratio of the branching fractions shown in Eq. 1.33,

$$\epsilon' = \frac{\Gamma(K_2 \to \pi\pi)}{\Gamma(K_2 \to \pi\pi\pi)} \tag{1.33}$$

and the PDG calculated value for the relative magnitude compared to ϵ is,

$$\mathbf{Re}\left(\frac{\epsilon'}{\epsilon}\right) = (1.66 \pm 0.23) \times 10^{-3} \tag{1.34}$$

which is around 600 times smaller than ϵ [36]. However, the current CP-violation observed in the kaon (and B-meson) sector is relatively small, and cannot explain the matter and antimatter asymmetry. Because of this, CP-violating physics is still an area of interest for many in the physics community. This interest led to the development of the KOTO experiment at JPARC which aims to study rare kaon decays that are direct CP-violating.

CHAPTER 2

THE KOTO EXPERIMENT

The KOTO experiment is an international collaboration that currently consists of more than 10 universities and national labs in Japan, Taiwan, South Korea, and the United States. KOTO itself is an experiment located in J-PARC, which is situated in Tokai-mura, a small village in Japan 2.5 hours by car northeast of Tokyo.

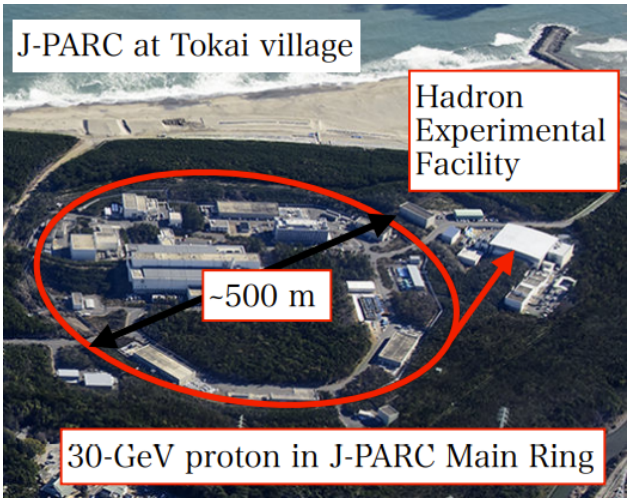


Figure 2.1: An overhead view of the experimental site.

JPARC specializes in its high intensity 30 GeV proton beam. This proton beam is extracted from the main ring and directed towards the Hadron Experimental Facility which is home to the KOTO target and detector, as seen in Fig. 2.2.

2.1 Physics Motivation

KOTO's main experimental goal is to measure the branching ratio of the ultra rare, Flavour Changing Neutral Current (FCNC) decay $K_L \to \pi^0\nu\overline{\nu}$, which is sometimes dramatically called the "Golden Mode". It obtained this moniker because: (1) It is direct CP-violating, and is directly proportional to the height of the CKM triangle, whose area represents the amount of CP-violation in weak decays and which can be constructed from the unitarity of

the CKM matrix. (2) It also has a very small theoretical uncertainty of around 2% in the SM estimation of the branching ratio [34]. Therefore, this mode is particularly sensitive to new physics.

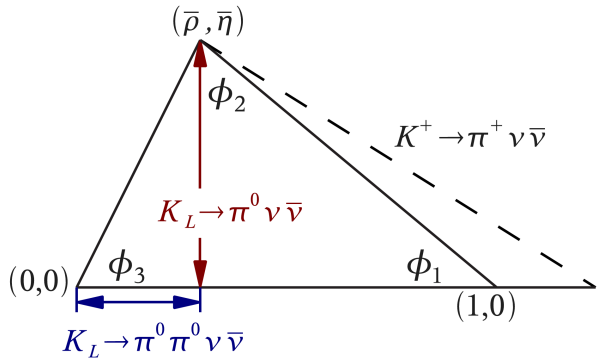


Figure 2.2: The CKM triangle, and the relationship of its sides and height to notable kaon decays.

It only takes a small difference in the measurement of the branching ratio compared to the Standard Model (SM) estimation to be a statistically significant discovery. KOTO has published experimental upper limits on this mode, with the previous best direct limit being set by KOTO in 2015 and the best limit being obtained in the analysis of 2021 data detailed in this thesis. However, between KOTO's best upper limits on $K_L \to \pi^0\nu\bar{\nu}$ to the SM branching ratio, there are still two orders of magnitude sensitivity space to search for new physics. That is the main motivation of the KOTO experiment.

2.2 $K_L \to \pi^0\nu\bar{\nu}$ and $K_L \to \pi^0\pi^0\gamma\gamma$ in the Standard Model

As mentioned in the previous section, the decay $K_L \to \pi^0\nu\bar{\nu}$ is a FCNC process which changes the strange quark to a down quark. This transition is dominated by short distance, $\Delta S = 1$ transitions. The long distance $\Delta S = 2$ transitions which arise from neutral kaon mixing, turning the $K_L \to K_S$, are much smaller because of the Glashow–Iliopoulos–Maiani (GIM) Mechanism [12]. The direct CP violating transition is thus dominant, and is expressed by the penguin and box diagrams shown in Fig. 2.3

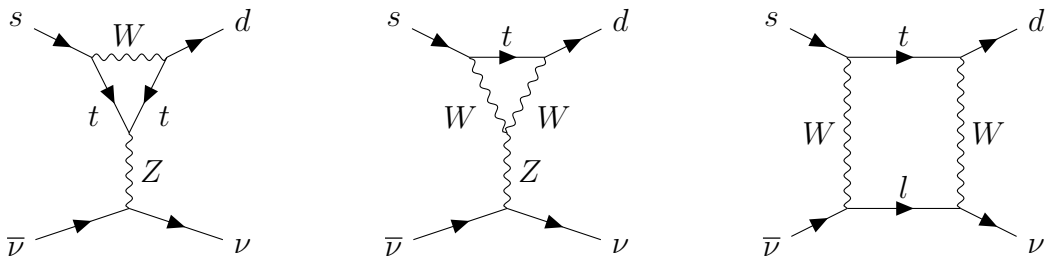


Figure 2.3: Dominant tree diagrams for the decay $K_L \to \pi^0 \nu\overline{\nu}$

Notably, the top quark plays a dominant role in the decay because the heavier mass leads to a smaller suppression factor on the order $\left(\sim \frac{m_q^2}{M_W^2}\right)$ [12]. As a result, and because the process is dominated by short interactions, the branching ratio is very nearly proportional to the height of the CKM triangle, η, through the relation shown in Eq. 2.3

$$BR(K_L \to \pi^0 \nu\overline{\nu}) \lessapprox BR(K^0 \to \pi^0 \nu\overline{\nu}) - BR(\overline{K}^0 \to \pi^0 \nu\overline{\nu}) \tag{2.1}$$

$$\lessapprox V_{ts} V_{td}^* - V_{ts}^* V_{td} \tag{2.2}$$

$$\propto \eta \tag{2.3}$$

This means that a direct measurement of the branching ratio of $K_L \to \pi^0 \nu\overline{\nu}$ is a test of the unitarity of the CKM matrix, and the main theoretical uncertainties arise from the CKM matrix elements. These parameters have been measured to relatively good uncertainty, which sets Standard Model uncertainty on the branching ratio of $K_L \to \pi^0 \nu\overline{\nu}$ on the order of 2% [13].

$$BR(K_L \to \pi^0 \nu\overline{\nu})_{SM} = 2.94 \pm 0.15 \times 10^{-11} \tag{2.4}$$

The radiative decay $K_L \to \pi^0 \pi^0 \gamma\gamma$ is another in a list of interesting K_L which may provide insight on Quantum Chromodynamics (QCD) low energy physics. There is a long history of studying such radiative decays in kaons, with interesting results. One such mode that is similar to $K_L \to \pi^0 \pi^0 \gamma\gamma$ is $K_L \to \pi^0 \gamma\gamma$, which was measured in KTeV [6] and NA48 [29] with

the current measurement uncertainty from the Particle Data Group (PDG) of around 3% [36]. This measurement was motivated by the need to understand the CP-conserving component of the rare decays of the form $K_L \to \pi^0 l^- l^+$, where the two photons from $K_L \to \pi^0 \gamma\gamma$ scatter to form a lepton pair. The theoretical estimation of the branching ratio was determined by ChPT, an effective field theory which uses the approximate chiral symmetry of QCD to study the low energy (<1 GeV) dynamics. Ultimately, this measurement had a strong disagreement with ChPT at $O(p^4)$ that led to a deeper understanding of interference between terms at $O(p^4)$ and $O(p^6)$ [22]. A similar story plays out between $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to 2\pi^0 l^- l^+$ which were studied in KTeV [8] [5]. The branching ratio of the latter is fundamentally affected by the process $K_L \to 2\pi^0\gamma^*\gamma^*$, this time CP-violating, shown in Fig. 2.4

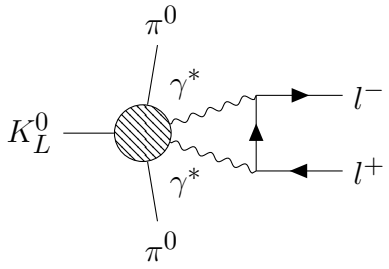


Figure 2.4: Feynman Diagram of the CP-violating contribution to $K_L^0 \to 2\pi^0 l^- l^+$ from $K_L \to \pi^0\pi^0\gamma\gamma$ via internal $\gamma\gamma \to l^- l^+$

The number of free parameters increases significantly between 2nd and 3rd order ChPT [22], and it has been shown in the study of $K_L \to \pi^0\gamma\gamma$ that the previously thought negligible $O(p^6)$ terms are not insignificant. The only current theoretical prediction of the branching ratio of $K_L \to \pi^0\pi^0\gamma\gamma$ is evaluated at $O(p^4)$ only [24], so it is possible to observe deviation from the theoretical prediction, and further test ChPT to third order.

2.2.1 Beyond Standard Model: $K_L \to \pi^0\pi^0 X$, where $X \to \gamma\gamma$

Analogous to $K_L \to \pi^0\pi^0\gamma\gamma$ is the mode $K_L \to \pi^0\pi^0 X$ where $X \to \gamma\gamma$. This mode was studied in KTeV [5] where $X \to \mu\mu$, and E391A [48]. In addition, there are many dark

matter models in the sub-GeV range which are gaining theoretical popularity. Some of these are models positing that such dark matter may couple strongly with quarks and still match observational limits on relic density [31]. Nonetheless, there is still a significant observational gap in the MeV range that KOTO is sensitive to. In particular, there are Minimal Flavour Violation (MFL) theoretical scenarios which posit that the three body kaon decays may be more strongly coupled to dark matter compared to two body decays [10]. Thus, the couplings to $K_L \to \pi^0\pi^0 X$ and $K_L \to \pi^+\pi^- X$ may be significantly larger than the coupling to well studied modes, $K_L \to \pi^0 X$ [4] [6] and $K^+ \to \pi^+ X$ [16]. For instance, if only the coupling to the W boson is non-zero, ($\sim c_{WW}(\Lambda) \neq 0$), then the ratio may be ≈ 1600 [10]. Additionally, the $K_L \to \pi^0\pi^0 X$ search is the weakest bound by far on the coupling [48], and further improvements in the limit will constrain NP models.

2.3 KOTO Beamline and Detector

The KOTO beamline is approximately 21 meters from the position of the fixed gold target, to the entrance of the KOTO detector. The 30 GeV beam arrives at the target in "spills" where beam is incident for two seconds and off for two seconds. This beam is extracted from the main ring via "slow" extraction using a third order resonance of betaron oscillations [35]. A critical component of the proton beam is related to its flatness with respect to time. If many of the on-spill protons have a high frequency component, then data acquisition efficiency will be lower. This flatness is typically expressed in terms of the duty factor,

$$F = \frac{\left(\int_{t_i}^{t_f} I(t)dt\right)^2}{\int_{t_i}^{t_f} I^2(t)} \tag{2.5}$$

where t_i and t_f represent the start and end time of the spill, while $I(t)$ represents the beam intensity on the target as a function of time over that spill. If the beam is perfectly flat, then this factor is exactly one. In 2021 data taking the duty factor was around 78% with an

average beam power at 64kW [35].

The KOTO beamline is comprised of three main components and can be seen in Fig.2.5. Nearest to the target is a photon absorber, which is followed by two collimators. The two collimators are meant to control the K_L beam momentum to be near the beam axis. In between the collimators is a sweeping magnet to keep charged particles from entering the KOTO detector decay volume. The long beamline is crucial to ensure that no K_S or other

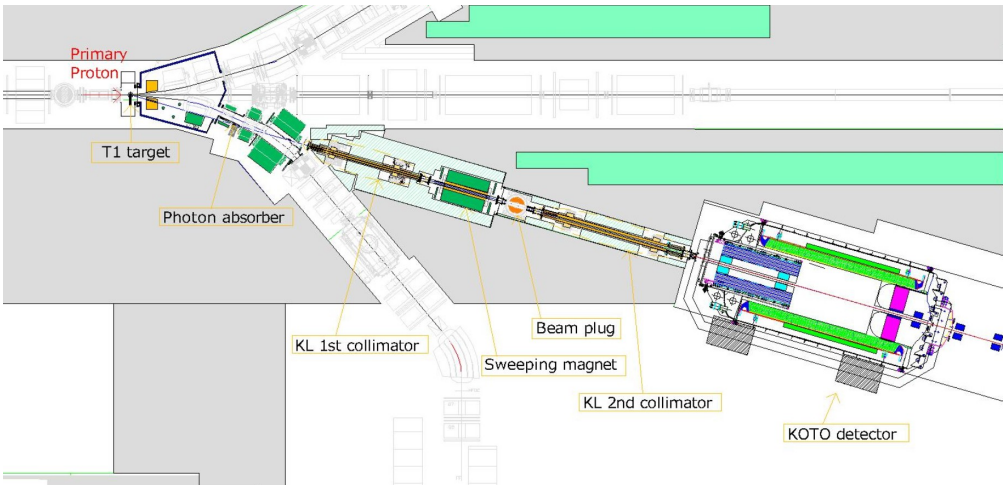


Figure 2.5: The layout of the KOTO beamline.

short lived particles make it into the decay region. With the long beamline and the sweeping magnet, the majority of particles entering the detector are K_L, n, or γ.

The KOTO detector is shown in Fig. 2.6. The main calorimeter consists of undoped Cesium Iodide (CsI) crystals, called the "CSI". The 2716 crystals are arranged in a "fish-eye" geometry, shown in Fig. 2.7 with the inner 1.2×1.2 meter region consisting of smaller crystals. The smaller (larger) crystals have dimensions $25 \times 25 \times 500$ millimeters ($50 \times 50 \times 500$ millimeters). The 500 mm depth corresponds to around 27 electromagnetic radiation lengths, so punch through is negligible.

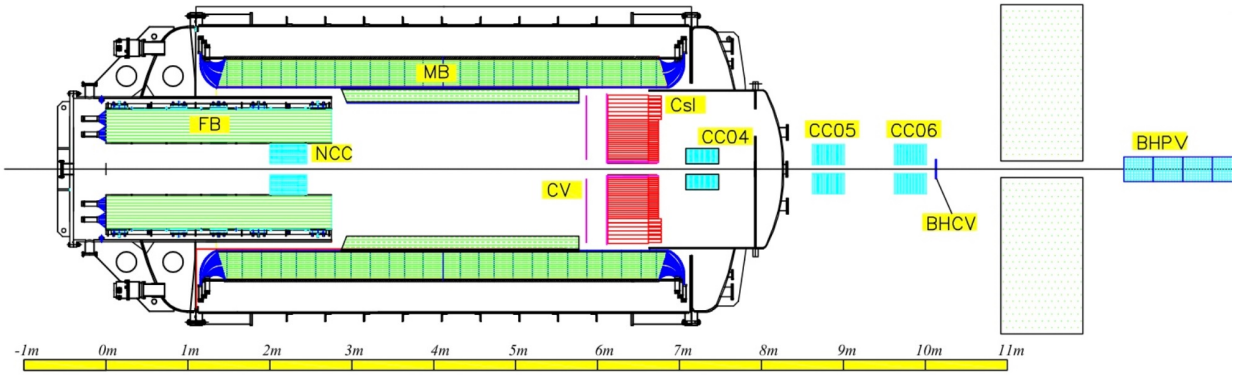


Figure 2.6: A cutaway view of the KOTO detector. The K_L beam arrives from the left.

Because of this geometry, a single photon hit will spread its energy out among many crystals. This forms a "cluster" of hits on the CSI which is used for the reconstruction of more advanced analysis variables. The energy (small and large crystal), timing, and position resolution are given by [43],

$$\frac{\sigma_E(\text{small})}{E} = (0.66 \pm 0.12 \pm 0.52)\% \oplus \frac{(1.81 \pm 0.04 \pm 0.02)\%}{\sqrt{E[\text{GeV}]}} \tag{2.6}$$

$$\frac{\sigma_E(\text{large})}{E} = (1.71 \pm 0.11 \pm 0.14)\% \oplus \frac{(1.31 \pm 0.10 \pm 0.01)\%}{\sqrt{E[\text{GeV}]}} \tag{2.7}$$

$$\sigma_t[\text{ns}] = \frac{(0.102 \pm 0.003)}{\sqrt{E[\text{GeV}]}} \oplus \frac{(0.004 \pm 0.001)}{E[\text{GeV}]} \oplus (0.11 \pm 0.02) \tag{2.8}$$

$$\sigma_{pos} = (1.99 \pm 0.23)\ \text{mm} \oplus \frac{(4.01 \pm 0.40)\ \text{mm}}{\sqrt{E[\text{GeV}]}} \tag{2.9}$$

Besides the calorimeter, the KOTO decay volume is fully surrounded by veto detectors. The largest being the barrel vetoes, called the main barrel (MB), inner barrel (IB), and front barrel (FB). These are lead-scintillator style detectors with the layers parallel to the beam and both end read-out. On the inside surface of the barrel vetoes, a thin strip of plastic scintillator is implemented to increase detection efficiency of charged particles. The

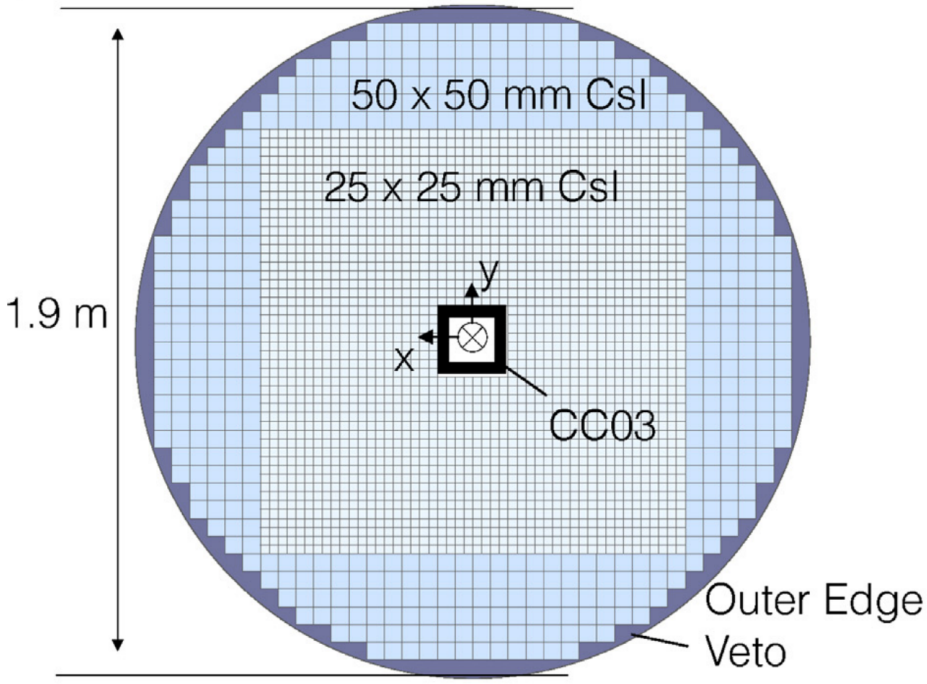


Figure 2.7: A representation of the front of the KOTO CSI, with the edge vetoes CC03 and Outer Edge Veto (OEV) [43].

individual modules which make up the barrel vetoes are constructed such that the edge does not point towards the beam center to avoid decreasing the efficiency, with a main barrel module shown in Fig. 2.8.

Other detectors include the Neutron Collar Counter (NCC) which sits at the upstream edge inside the front barrel, and CC0X detectors around the edge of the beam. These detectors are made from CsI crystals similar to those used in CSI. The charged veto (CV) detector is a thin two-layer plastic scintillator type detector which is placed in front of CSI to discriminate charged electromagnetic showers from neutral ones. Lastly, there are the detectors which lie in the beam-hole. Downstream of the CSI are the beam-hole charged veto (BHCV), which is an array of wire chambers, and the beam-hole photon veto (BHPV), a series of lead-aerogel style detectors.

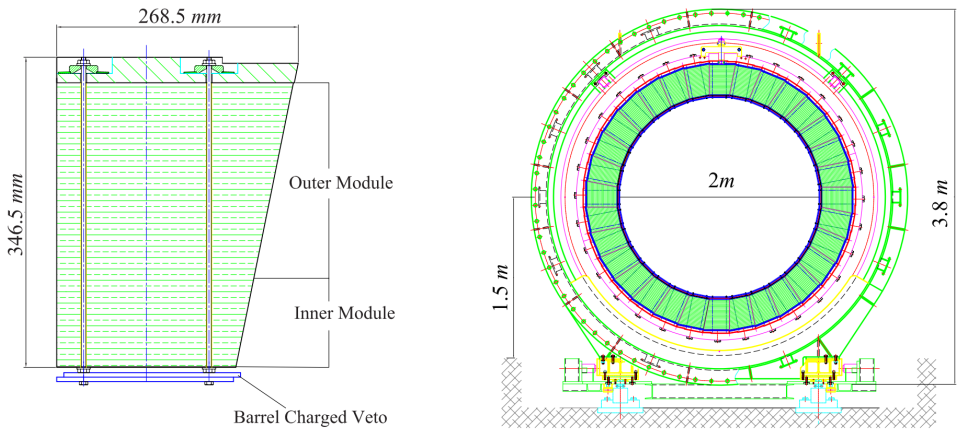


Figure 2.8: On the left is an individual MB module while on the right is a diagram of all modules assembled in the vacuum tank. The edges of each individual module point away from the beam center [52].

2.4 Experimental Principle

KOTO searches for an ultra-rare decay, $K_L \to \pi^0\nu\overline{\nu}$, which requires extensive understanding of potential background sources, such as those that come from neutrons or other K_L decays. Neutron background rejection requires a deep understanding of how hadronic showers in the CSI differ from EM showers. The latter requires a specific signal decay signature which KOTO can use to discriminate $K_L \to \pi^0\nu\overline{\nu}$ decays from other more common K_L decays. For instance, because $K_L \to \pi^0\nu\overline{\nu}$ has two neutrinos in the final state which go unmeasured, the two photons from the pion decay exhibit a high, off-beam axis momentum. If this fact is paired with an extremely efficient hermetic veto system, an event with two clusters that have high transverse momentum and no other hits in the veto detectors is unique to $K_L \to \pi^0\nu\overline{\nu}$. An example decay is shown in Fig. 2.9.

The pion reconstructed vertex, and hence transverse momentum, is reconstructed from the two photon hits in the CSI from Eq. 2.10.

$$M^2_{\pi^0} = 2E_1E_2(1 - \cos\theta) \tag{2.10}$$

E_1 and E_2 represent the reconstructed energies of the two photons from the energy deposit in the CSI, while θ represents the opening angle between the two photons. This allows the calculation of the decay vertex if the pion mass is assumed.

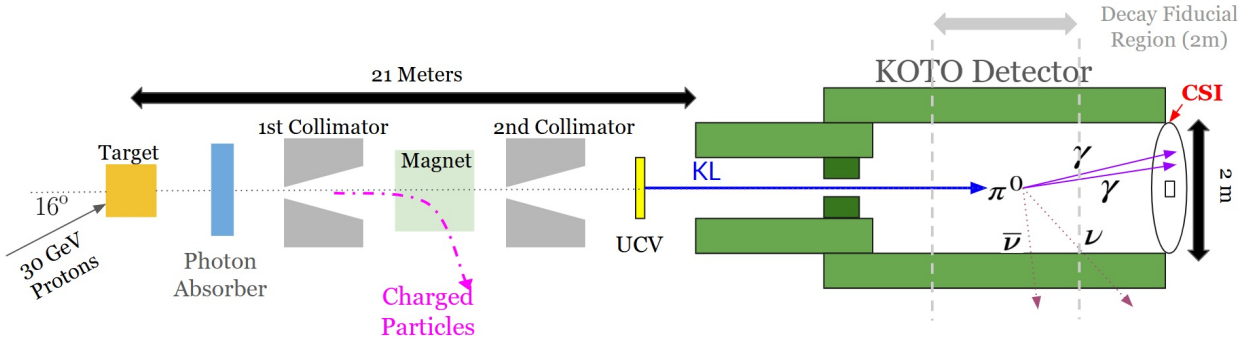


Figure 2.9: A simple schematic showing the K_L beamline with an example $K_L \to \pi^0\nu\overline{\nu}$ decay.

The signal region is defined as the pion reconstructed vertex on the beam axis versus the pion transverse momentum as shown in Fig.2.10.

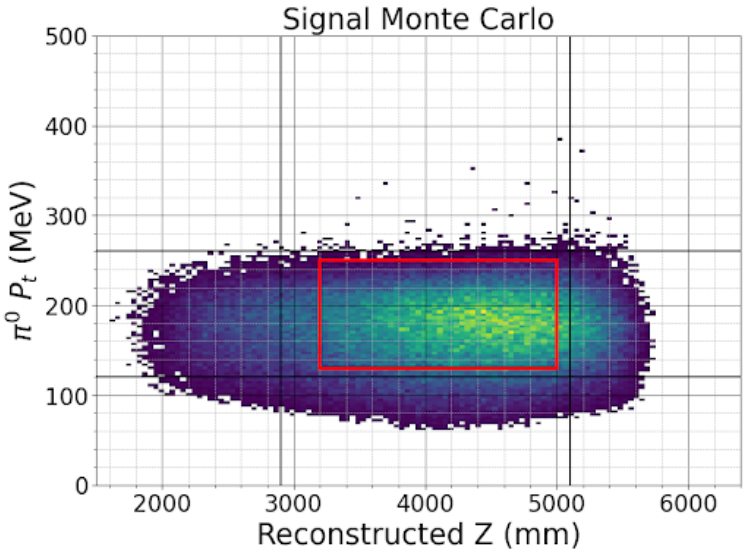


Figure 2.10: $K_L \to \pi^0\nu\overline{\nu}$ Monte Carlo simulation using GEANT4. The red box signifies the signal region.

2.5 KOTO DAQ-Trigger System

There are four main levels to the KOTO trigger system shown in Fig. 2.11: (1) gather the data, (2) extract the local information, (3) synthesize the global result, (4) make a trigger decision. All of this needs to be done in real time, clock edge by clock edge, with calculations that should be done as quickly as possible to minimize system deadtime since the ADC pipeline depth is only 5.2 μs. This must be done while handling nearly 4000 channels. At the lowest layer, there are 288 Chicago made analog to digital converters (ADC) housed in VME-64 crates. These turn the analog data into digital data. These ADCs then send information to a Clock Distribution (CDT) module which will concatenate the data from that VME-64 crate and send it to an Optical Fiber Center (OFC). It will synthesize that information and extract some complex information, such as detector energy thresholds or the number of clusters in the CsI. Once it is finished, the special OFC modules will send the complex trigger information to the Top-CDT, which will issue the trigger signal to all ADC's if the proper conditions are met.

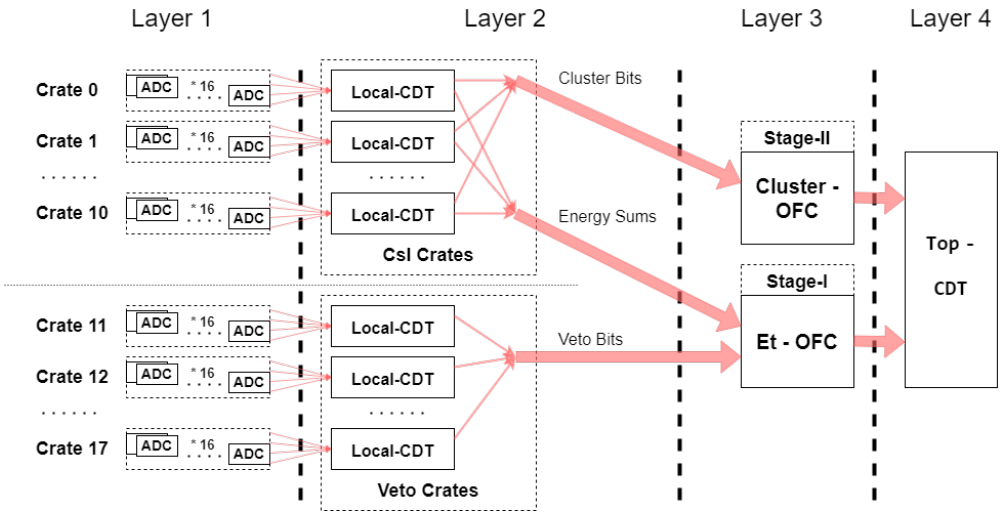


Figure 2.11: An overview of the KOTO trigger system.

2.5.1 Stage-1 Trigger

The Stage-1 Trigger consists of the CSI calorimeter on-time energy sum over all 2716 channels, as well as the on-time energy deposit in veto detectors. If the total energy deposited in the CSI exceeds the required threshold while veto detector energies are below their respective thresholds, this event will pass Stage 1.

To calculate the entire CSI E_t, the deposit energy of each individual channel is calculated and later summed up by the Et-OFC shown in Fig. 2.11. The channel energy is evaluated by taking advantage of the linearity between pulse height and deposit energy. Given calibration constants for each channel obtained using cosmic rays, the channel energy is simply calculated from,

$$E_i = c_0 P_h + c_1 \tag{2.11}$$

where c_0 and c_1 are the constants obtained from calibration, and P_h is the pulse height. The height is taken relative to the online pedestal value which is the mode of the pedestal during spill off. The local energy sums will be evaluated at the local CDT which then sends the result to the Et-OFC. The final CSI E_t and Veto energy sums are calculated there and then sent to the TopCDT. Based on user-defined thresholds, the result passes or fails the Stage-1 trigger. All trigger calculations are on a strict 5.2μs timer, and to decrease system deadtime, a daisy chain strategy is used to concatenate the energy data sent from each ADC. A diagram of the strategy is shown in Fig. 2.12. There are only four SFP inputs at the local CDT which allows for a 4x daisy chain. Since there are 16 ADCs per crate, the ADC trigger data is concatenated four times down the daisy chain so each SFP input into the Local CDT receives the data from four ADCs.

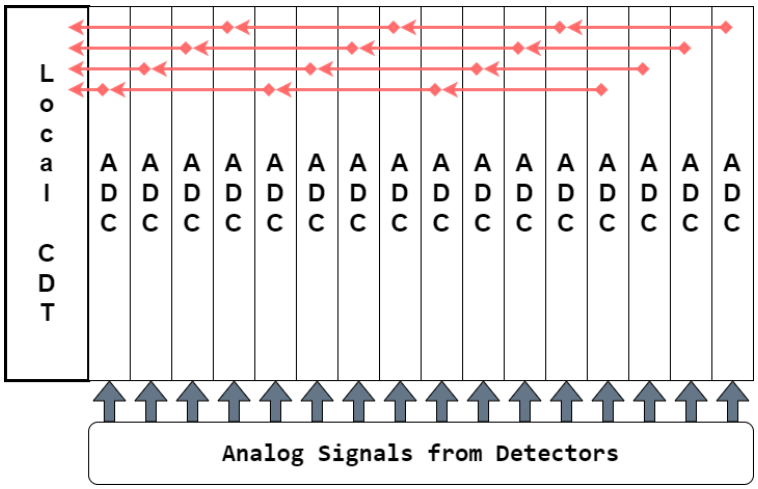


Figure 2.12: A diagram of the daisy chain strategy to send ADC data for trigger decisions to the local CDT.

2.5.2 Stage-2 Trigger

The Stage-2 trigger relies on the number of clusters in the CSI calorimeter. The calculation is only initiated after an event passes the Stage-1 trigger. Similar to Stage-1, the ADC Stage-2 trigger data is sent via daisy chain to each Local CDT which is then sent to the Clustering OFC module, as in Fig. 2.11. This trigger data consists of 16 "hit" bits for all ADCs with CSI channels. A channel has a "hit" if the energy deposit in the CSI channel is above threshold. This threshold is optimized for Stage-2 trigger efficiency of $K_L \to \pi^0\nu\overline{\nu}$ events. For small crystals, this threshold is 22 MeV while for large crystals the threshold is 44 MeV.

Once the "hit" map arrives at the Clustering OFC, the cluster count has to be done cleverly to avoid having a large deadtime. The basic strategy is tied to a topological fact about 2D objects-the number of clusters can be obtained via Eq. 2.12.

$$N_{clusters} = \frac{N_{convex} - N_{concave}}{4} \tag{2.12}$$

N_{convex} and $N_{concave}$ are the number of convex and concave turns respectively. To

demonstrate why this is the case, a simple example is shown in Fig. 2.13. The idea is very simple; there must be four 90 degree turns to make a shape. If there is a backwards turn of 90 degrees then an extra turn forward 90 degrees is needed.

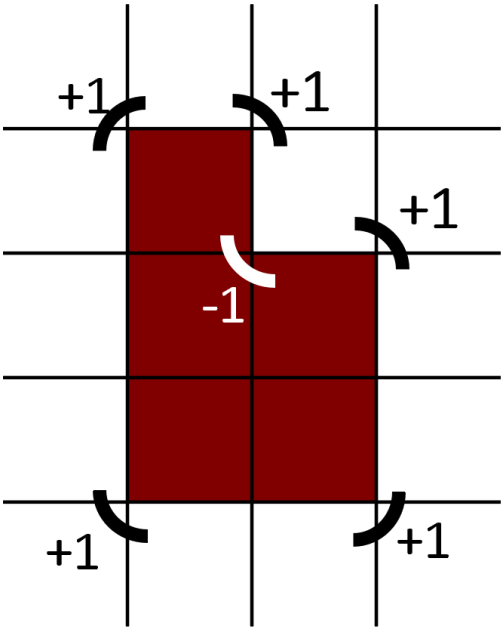


Figure 2.13: A simple example of the turn counting strategy to determine the number of clusters.

Counting the total turns gives four, and corresponds to a single cluster. The trick to this calculation is that it can be done in parallel in an FPGA by breaking up the hit map into many 2x2 sections and evaluating the number of turns for all of these sections at once. The number of turns in a given 2x2 section is shown in Fig. 2.14. As shown in Fig. 2.7, there are crystals of different sizes. The large ones, which have a cross sectional area of 5×5cm, are their own "2x2" section. The small crystals, which have one quarter the cross sectional area, each make up one corner of a 2x2 section. Afterwards, the total number of turns and hence the number of clusters is obtained. The system deadtime is limited primarily by the data transmission which takes 18, 125 MHz clock cycles, while amazingly, this algorithm calculates the number of clusters in two clocks, leading to a system downtime of only 0.16μs, and a corresponding loss rate of less than 1%.

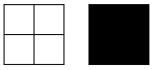
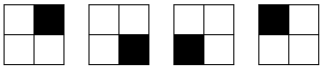
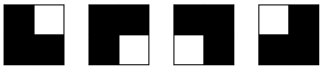
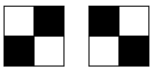
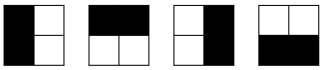
Cell Patterns	Physical Meanings	Pattern Values
	No corner	0
	One convex corner	+1
	One concave corner	-1
	Two convex corners	+2
	No corner	0

Figure 2.14: This table shows all different types of 2x2 sections, and the equivalent number of turns

2.6 Data Transfer

Once an event has passed both of the L1 stages, the Top-CDT issues a trigger to all ADCs to send the data. However, the ADCs do not send the entire pipeline for data storage. Instead, only the 64 digitized samples around the event pulses are sent. In this case, it is necessary that all ADCs send the data properly in a format aligned for data analysis. This alignment process is initiated by the Top-CDT at the start of a spill by sending an alignment trigger to all ADCs. Once an ADC receives this trigger, it self generates a "delta" event, which is a one sample wide pulse with a height of 10,000 counts. As this pulse travels through the trigger system, the relative delay of these delta events in clocks is measured. Once the delay time in clocks is known and configured for each ADC, then the correct 64 sample section is read out.

A general overview of the data transfer system is shown in Fig. 2.15. The system design is a pyramid style architecture which has major benefits. The first and most critical, is that it allows for complex late trigger decisions using GPU. This is because the event building can be done naturally in Layer 3 before the data reaches the Dell servers. Once the data arrives

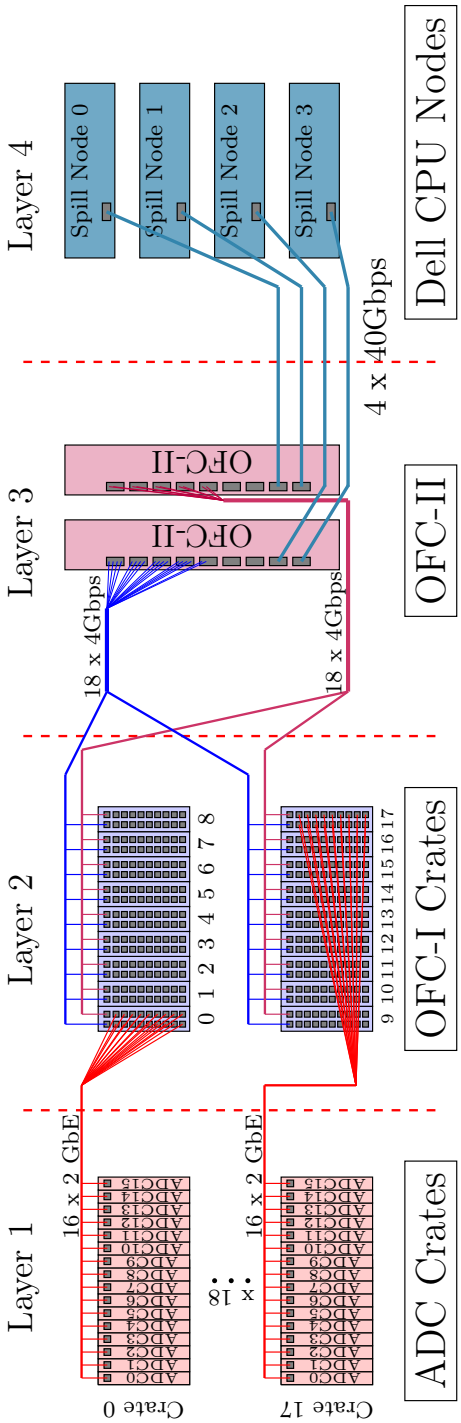


Figure 2.15: Overview of the KOTO data transfer system from the ADC to Dell servers located off the experimental area.

at the servers, it is already arranged in a way that is ideal for data analysis in software. Second, it is robust and sustainable. For instance, this strategy allows for easy expansion of the DAQ, with extra I/O capabilities in Layer 3 enabling more ADC crates if more physics motivated readouts are necessary. This concept may be expanded on in the second phase of the KOTO II experiment, which will require at least double the number of readout channels and more complex trigger decisions.

2.6.1 Layer 1: ADC Crates

The first layer consists of 288 14-bit ADCs. This transfer is done via a SFP cable and dell transceivers capable of 2 Gbps data transmission. The data transfer for the ADC is handled by a TLK chip made by Texas Instruments on the board. This special chip is required because there are no dedicated transceiver pins on the ADC's Stratix II FPGA. Each ADC needs to transfer 64, 14 bit samples from the FPGA pipeline for each accepted event. A clock frequency of 125 MHz corresponds to a payload per clock of 16 bits. However, the samples are only 14 bits which allows the space for two "header" bits. These header bits allow each module throughout the data transfer to determine what type of bit word is received at any given clock. There are three types of words: (1) header words [11], (2) energy words [10], and footer words [01]. The header words contain important event information that is used online and offline for event monitoring, with a detailed description given in Fig. 2.16

These headers are loaded in front of the energy words, and are used for recognizing and monitoring an incoming event at later layers of data transfer. Including header and footer words, each ADC will send out 6 (header) + 64 (energy) + 1 (footer) 16 bit words per event. This payload is sent to a corresponding OFC-I port, based on its crate and slot position.

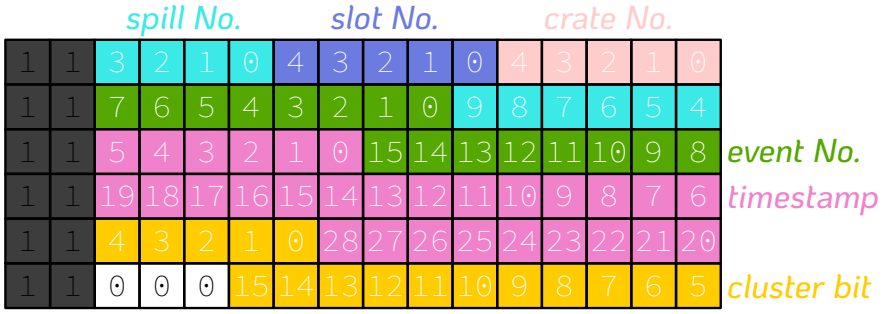


Figure 2.16: The format for the ADC headers. They consist of six 16 bit words that contain the timestamp information, spill number, event number, cluster number, and the crate and slot number of the ADC

2.6.2 Layer 2: OFC-I Crates

The OFC-I is a Chicago made module with an Arria V FPGA that consists of eighteen SFP ports capable of 5 Gbps transmission. Due to 10/8b encoding, the payload rate is only 4 Gbps. Sixteen of these ports are used to receive the data from a full VME crate of ADCs, which run at 2 Gbps as mentioned in the previous section. The two other SFP transceiver ports are used to output the data at 4 Gbps to the next layer of data transfer. The OFC-I has the role of concatenating data from each crate. For instance, all of the ADCs in crate 0 will send the data to the same OFC-I, which will then concatenate all the data from the crate including the headers and footers, and then send it to the next layer.

The entire event is kept together so that complex trigger decisions may be made later. To achieve this, the OFC-I employs a ping-pong strategy as shown in Fig. 2.17. Each even numbered event will be sent to a different designated on-chip memory location than an odd numbered event. Later, once the OFC-I has received all of the data from each of the 16 ADC's, the data is read from the memory sequentially at 4 Gbps, or 32 bits per clock cycle. This means that the OFC-I is the natural bottleneck of the data transfer. The incoming rate may be as high as 16×2 Gbps $= 32$ Gbps, while the maximum output rate is only 8 Gbps.

Even with this bottleneck in OFC-I, the typical rates during data taking fall far below the limit set by OFC-I. During data taking runs in January 2025 with the highest beam

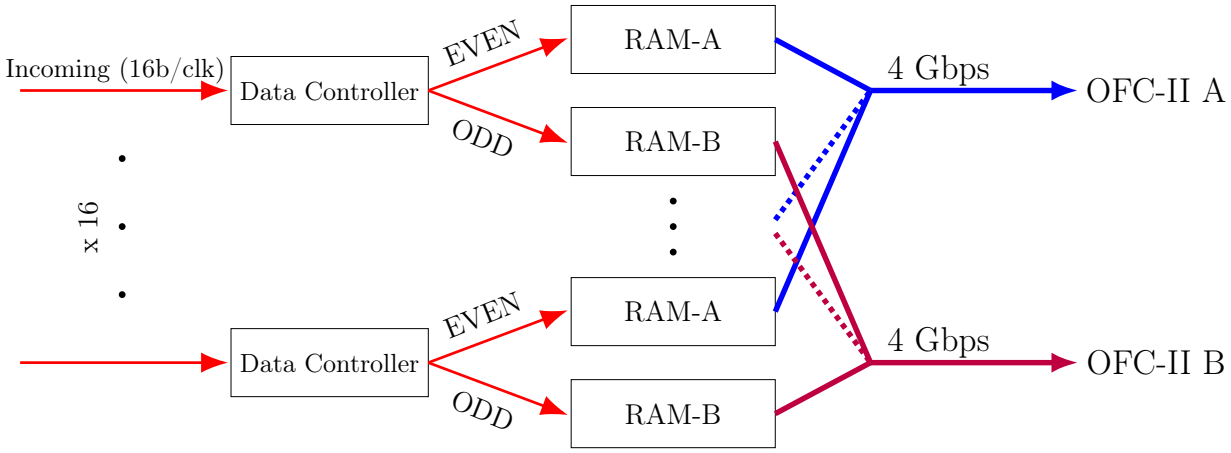


Figure 2.17: Strategy to output the data at the OFC-I stage. Even numbered events go to RAM-A, while odd numbered events go to RAM-B.

intensity to date at around 80 kW, and with a full trigger table, the number of events per second averaged 10,000. The results on the loss caused by OFC-I busy or buffer full during data taking in 2023 and in preliminary tests are shown in Fig. 2.18.

There are different results based on the test, so it is vital to describe the difference between tests in detail. The best performing curve represents a test done by optimizing the buffer size in the OFC-1. By increasing the buffer size, the OFC-1 is more robust against intermittent high rates caused by the non-uniform beam structure. In this case, most of the OFC-I memory is used for the buffer instead of runtime debugging logic such as a "spy memory" which can be read out in real time using the VME interface. The buffer size used during data taking is represented in the rest of the tests. The small scale test represents the tests done in the University of Chicago lab, with only a single crate of modules, while the full system test was done in JPARC. The rate in the full system test was achieved using concatenated timestamps selected randomly from several spills to mimic the beam structure for high intensity runs. The performance is nearly identical between these two tests, yet the performance of the system is significantly worse when done with real data. This difference is most probably caused by the method to test high rates in data which affects the trigger

timing structure, leading to intermittent high rates to which the OFC-I small buffer size is especially susceptible. Nonetheless, this test provides a worst case scenario for the loss percentage of the data transfer to CPUs.

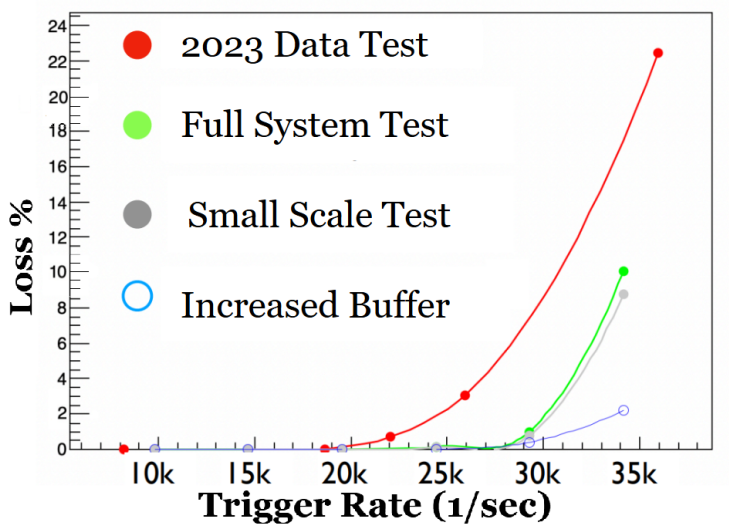


Figure 2.18: This plot shows the event loss due to OFC-I buffer full errors at different rates for several tests.

2.6.3 Layer 3: OFC-II

The OFC-2 is another homemade module from the University of Chicago that consists of 9 QSFP ports and houses a Stratix X FPGA. This FPGA has dedicated transceiver pins capable of up to 100 Gbps data transmission [32]. This module also contains 4 GB of high bandwidth DRAM memory (HBM) on chip. This HBM is capable of incredible read and write speeds of 800 Gbps. This provides the capability of further expansion of late stage trigger decisions or the calculation of complex variables within the OFC-2 module itself.

Five of the QSFP ports are used for handling the 18 inputs coming from each of the OFC-I, and the SFP $\rightarrow$ QSFP conversion is done using "octopus fibers" which combine 4 SFP cables into one QSFP cable. Since each OFC-I outputs two 4 Gbps lines, there are two OFC-2s to handle the inputs, labeled "OFC-2 A" and "OFC-2 B" as in Fig. 2.17. Each

OFC-2 will receive the data from an entire event, and do the event building before sending it via 40 Gbps Ethernet to the Dell servers. The event building logic is described in Fig. 2.19.

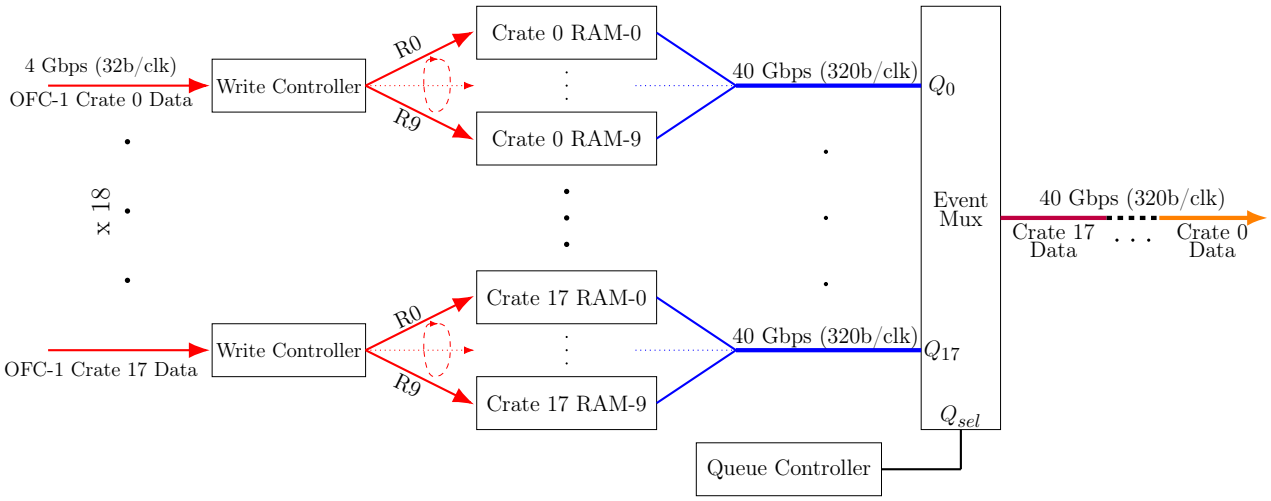


Figure 2.19: The event builder logic in the OFC-2. RX in the figure represents the remainder (X) modulo ten.

Since there are two 40 Gbps outputs in each OFC-2, a ping-pong strategy similar to that described in Fig. 2.17 is implemented before the event building. Once the data reaches the event building stage, it is directed into 10 separate, 32-bit wide RAMs depending on the order of its arrival. Each input has 10 designated RAMs, where the RAM the word is sent to is decided by the word number modulo 10. Later, once all 18 inputs deposit the data for a full event, then each of these 10 RAMs will be read in parallel to build a 320 bit bus for the data transfer. Each set of RAMs are read sequentially such that each crates data arrives to the Event Multiplexer at the correct time to be sent out in order.

The internal 320 bit bus needs to be handled correctly when interfacing with the 40 Gbps Ethernet IP. This is done by a system of control logic between event building and the 40 Gbps Ethernet IP interface. The details are shown in Fig. 2.20.

The "Event_waiting" bit is asserted once the number of events written into the memory becomes larger than the number read from the memory. The number written is incremented only when each of the 18 inputs fully writes the data into the RAM. The number of events

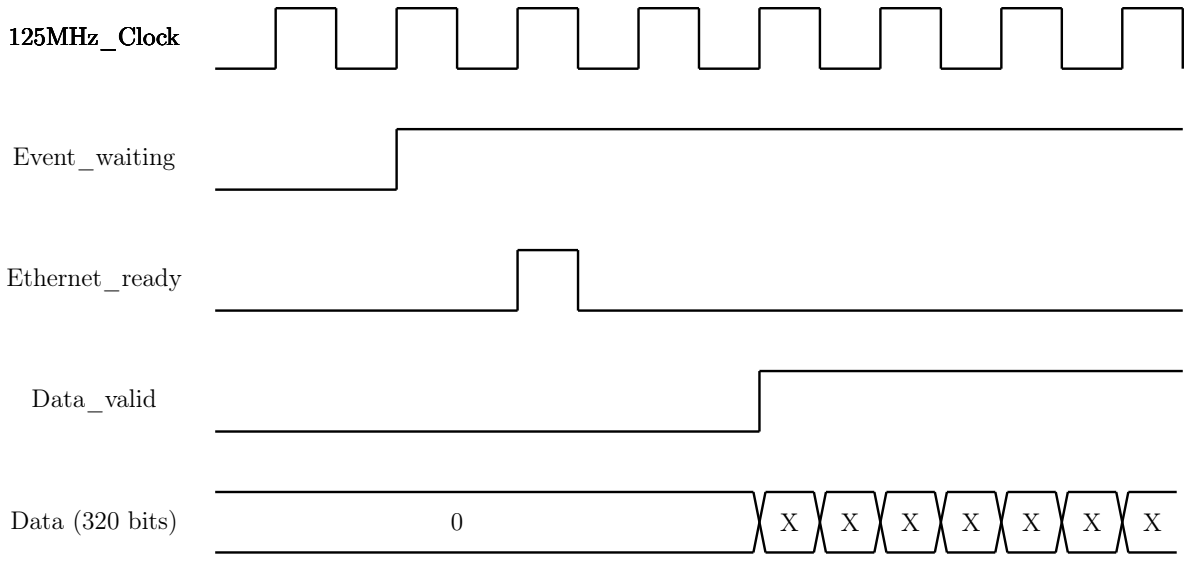


Figure 2.20: Control logic between the event builder and the 40 Gbps interface

read is only incremented once an event is fully read out. The Ethernet interface monitors the "Event_waiting" bit, and responds with a ready pulse based on two conditions: the previous data has been sent out and the state machine handling the interface logic is in the correct state to receive the data, and if the Ethernet IP reports stable TX/RX lanes. Once the ready pulse is received, the data is aligned to complete the event building and the first bits are sent out two clocks later.

2.7 Layer 4: The Dell Nodes and Late Stage Trigger

There are two type of nodes that are present in Layer 4, worker nodes and storage nodes. The four worker nodes are typically called spill nodes. They perform the 40 Gbps data capture from the OFC-2, data compression for transfer to KEKCC, and complex analysis for late stage trigger decisions. The two storage nodes are called disk nodes, and they collect the data from all four spill nodes after data processing, and store it on hard disk drives (HDD) capable of 96 TB of storage each.

The worker nodes must first perform the data capture at 40Gbps from the OFC-2. This

rate may be sustained for upwards of two seconds during the spill on cycle. In KOTO, this capture is done using a software package called "Netmap" [40], which allows for full control and management of the ring buffers in the kernel. The basic strategy to perform the data capture and transfer incoming data from the Network Interface Card (NIC) to the RAM is shown in Fig. 2.21 [27].

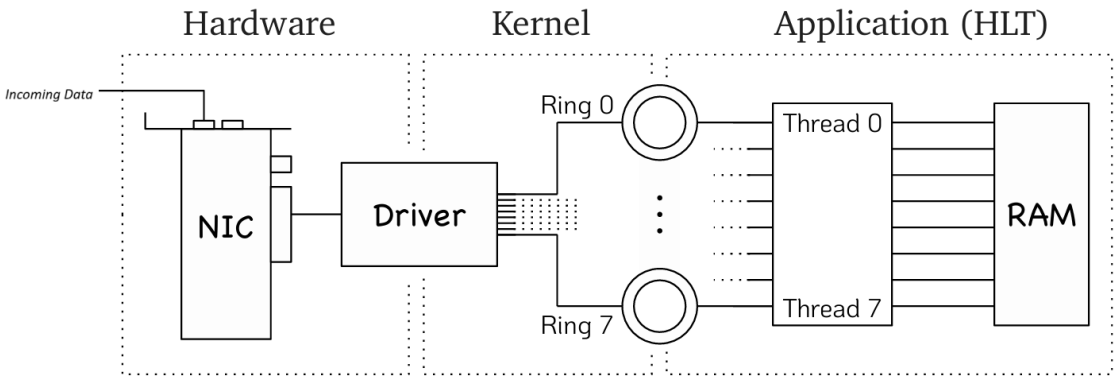


Figure 2.21: Data capture from the Network Interface Card (NIC) to the RAM.

The NIC hardware handles the data capture and the Dell driver interfaces the hardware with eight separate ring buffers. Internally with Netmap, these ring buffers are handled with eight CPU threads which write the data to the RAM. These threads handle and monitor the status of these ring buffers, and disperse the data in the RAM using non-uniform memory access (NUMA) techniques to improve write speed.

Once the event reaches the RAM, the spill nodes first calculate the number and location of clusters in the CsI using the event data. The energy deposit of each CsI crystal is calculated using a similar strategy as Eq. 2.11, except instead of pulse height, the integrated area under the pulse is used.

$$E_i = \frac{I_{pulse} \cdot \mathrm{MeVC}_i}{L_i \cdot G_i} \tag{2.13}$$

E_i represents the channel i, I_{pulse} is the integrated ADC pulse, and the calibration constant MeVC_i, is obtained preliminarily from cosmic ray data. The factor of G_i and L_i are the gain and linearity correction factors that are evaluated offline for more precision. The pulse is integrated relative to the online pedestal value, which is obtained during the spill off cycle and stored in the first sample of the waveform. An example ADC pulse with the online pedestal highlighted is shown in Fig. 2.22.

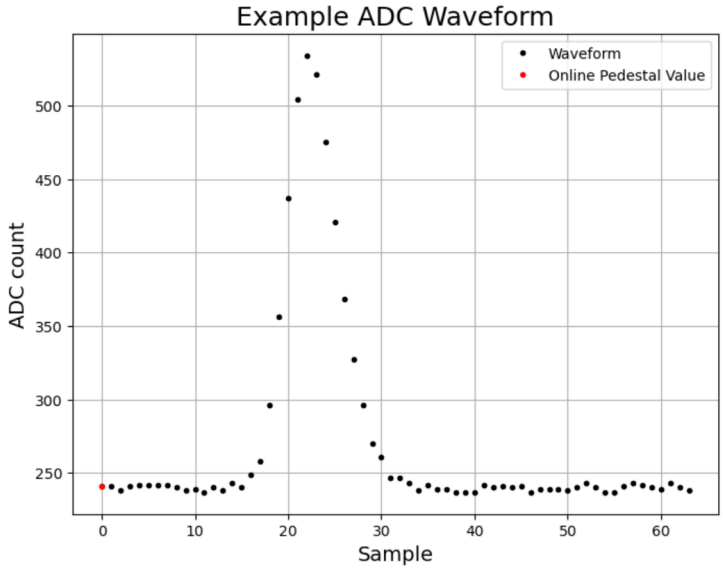


Figure 2.22: An example waveform from ADC data, with 64 samples. The first sample is the result from the online pedestal calculation.

Once the energy from each channel for a given event is obtained, a clustering algorithm is performed. This clustering algorithm is an adaption of the CLUE algorithm from CMS [3], with the basic strategy outlined as follows:

- Assign weights to each crystal based on the reconstructed energy.
- Find the closest higher energy neighbor.
- Repeat 2. until there is no close neighbor with higher energy.
- If the resulting crystal is above a set energy threshold, then it indicates a cluster, and this cluster is called the "cluster seed".

- Use the cluster seeds to obtain cluster energies and positions in the CsI.

The maximum crystal distance to be a "neighbor" and the energy thresholds are optimized using Monte Carlo results for $K_L \to \pi^0\nu\overline{\nu}$, $K_L \to 3\pi^0$, and $K_L \to \pi^0 e^+ e^-$ to produce the correct number of clusters and energies. So far, the spill nodes apply loose late stage trigger decisions, as the current DAQ has no bottleneck and can easily handle the current rates. In the future, if there is the physics need, then the late stage trigger can be a powerful tool. The current application of the late stage trigger is to reject low energy events, as well as events with a hit in the CsI that are too close to the edge for certain trigger types.

In addition to late stage trigger, the spill nodes also handle pedestal suppression and data compression. This is essential for keeping disk usage in KEKCC low, and for keeping the necessary transfer rate to KEKCC less than 5 Gbps. This limit is based on the shared 10 Gbps cable that connects the Hadron Hall to KEKCC.

Only a few dozen channels out of nearly 4000 will have a pulse recorded during a given event. To avoid writing the pedestal result for most of those channels, a pedestal suppression algorithm is used. If the given energy in a channel is less than 1 MeV, which means it is essentially noise, then the samples are set to zero to maximize the space saving factor from data compression. This threshold gives an online inefficiency of less than 0.1% [27].

The data compression is done simply. For a typical waveform, there 14 bits for all 64 samples. However, to represent the full range of the waveform, only the number of bits necessary to write $S_{max} - S_{min}$ per sample is needed. To be completely lossless, the compression algorithm also stores the minimum value. A representation of the algorithm is shown in Fig. 2.23.

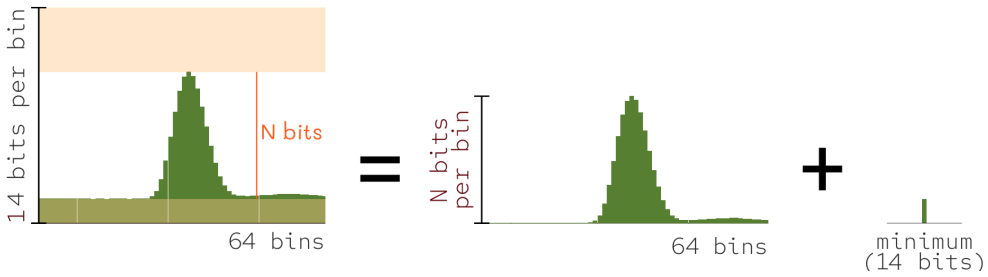


Figure 2.23: A representation of the compression algorithm [27]

The average compression is around a factor of three, and the total data size reduction obtained in in the spill nodes is shown in Fig 2.24.

Trigger	HLT-input rate (Spring 2024 physics runs)	Rate after event selection	rate after compression and ped. suppression
$K_L \to \pi^0\nu\overline{\nu}$	1.5 k/spill	Unchanged	/ 3.2
$K_L \to 3\pi^0$ (6 clus.)	2.0 k/spill	Unchanged	/ 5.3
$K^+ \to \pi^+\pi^0$	5.7 k/spill	/ 1.25	/ 5.4
$K_L \to 3\pi^0$ (5 clus.)	4.2 k/spill	/ 1.30	/ 5.3
$K_L \to \pi^0 e^+ e^-$	2.4 k/spill	/ 1.20	/ 5.3
Others	1.9 k/spill	Unchanged	/ 4.1
Total	**17.7 k/spill (20.0 Gbps)**	**17.2 Gbps**	**3.6 Gbps**

Figure 2.24: The total data reduction based on data taken in the Spring of 2024 [27]. The trigger type represents the mode in which the specific trigger conditions are designed for.

The trigger types in Fig. 2.24 represent the specific targeted K_L decays for a given trigger. For instance, the $K_L \to \pi^0\nu\overline{\nu}$ trigger represents two clusters in the CSI and no on-time hits in the veto detectors. On the other hand, $K_L \to \pi^0 e^+ e^-$ requires four hits in the CSI plus two hits in CV. In this case, the CV is taken out of veto and must see a hit. This flexibility in trigger allows KOTO to study important physics such as the K^+ flux, and $K_L \to \pi^0 e^+ e^-$. More details will be explored in the following chapter on the data taking procedures.

2.8 Error Monitoring

It is critical to have real time error monitoring in a DAQ system of this size. This monitoring is implemented in three ways:

1. Error bit transfer from Local Module → Local-CDT → Top-CDT for real time trigger management.

2. Spill-by-spill read out of on board statistics and status bits using VME protocol.

3. Layer 4 statistical analysis using a graphical users interface (GUI).

Each level of monitoring is used for separate purposes. The first, which is the error bit transfer for trigger management, is fundamental to data integrity. If there is an error reported in the system, the Top-CDT is able to withhold sending triggers. This immediately stops the DAQ if there is an error. The spill-by-spill read out allows for on-site experts to evaluate the type of errors that may have occurred on the order of seconds so they may be solved as soon as possible. The last method, which is the statistical analysis of incoming data at the Dell servers, allows for more detailed assessments of the type of error and loss associated with it. This analysis is done on the order of minutes, and error monitoring will be further discussed in the following chapter.

CHAPTER 3

DATA COLLECTION PROCEDURE

The data collection procedure requires precise coordination with the accelerator group. The trigger and data transfer system must be prepared for the moment protons hit the fixed target in order to maximize up time. This timing is provided by an accelerator group NIM module that provides the spill on cycle. When this line is up, it represents the period of protons on target, and the entire DAQ must be ready at that point. This is done internally using the LIVE signal transmitted by the Top-CDT via LVDS cables. The Top-CDT receives the result of the accelerator NIM module, then generates and distributes the LIVE signal appropriately to all modules in the system. The length and start of the LIVE signal may be modified by the DAQ group. A diagram showing the timing is given in Fig. 3.1.

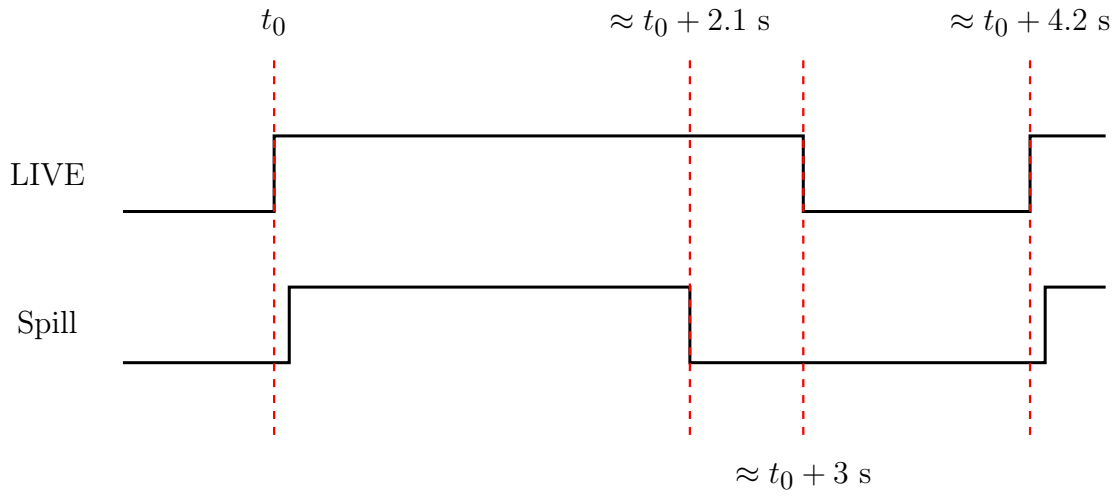


Figure 3.1: A schematic of the timing between the LIVE and Spill signals. A full LIVE or Spill cycle is 4.2 seconds in the most recent data taking runs.

The LIVE signal that is distributed to all modules signifies the start of the data taking procedure. At the falling edge of the LIVE off signal, each module will reset relevant statistics and registers, and wait in standby mode. The LIVE signal is set to arrive at the modules before the protons hit the target to check the healthiness of the optical links between all

modules, to perform the local timing alignment, and to perform the global alignment. If the timing and the optical links are not verified, real time errors are reported to the monitoring group, and the Top-CDT pauses the trigger until the next spill cycle. The time between the rising edge of the LIVE signal and the rising edge of the Spill signal is approximately 100 ms as shown in Fig. 3.2

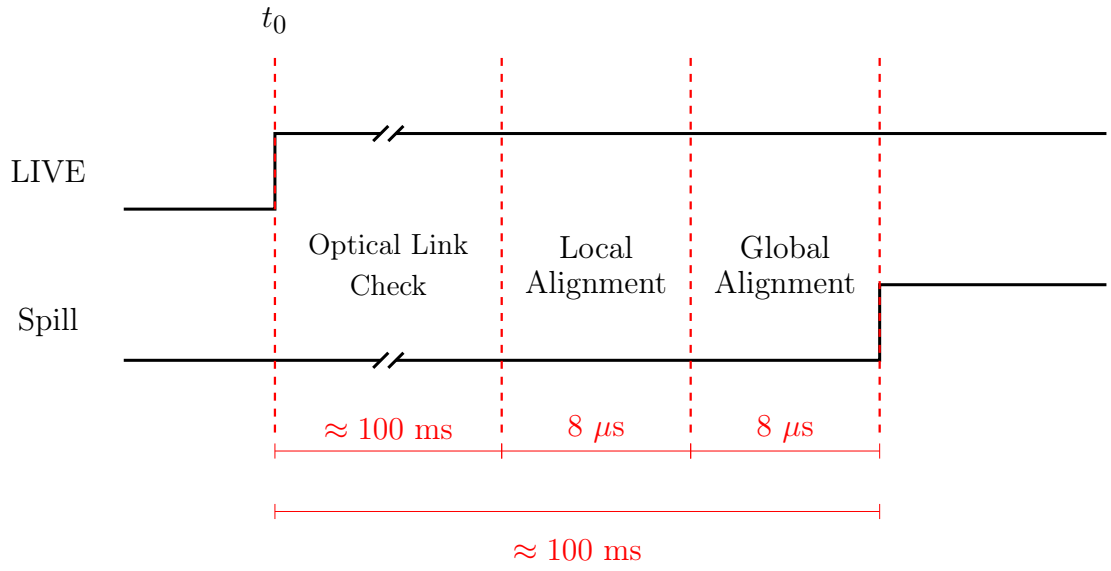


Figure 3.2: A schematic of the timing between the rising edge of the LIVE and Spill signals. The optical link check, local timing alignment, and global timing alignment are done between the LIVE rising edge and the Spill rising edge.

The optical link test is performed at each optical link between all modules. This test begins at the rising edge of the LIVE signal, and the sender outputs consecutive counters. The receiver continuously checks that the input increments every clock cycle over the course of 2000 clock cycles. If this check does not succeed within the 100 ms interval, an optical link error is reported to the Top-CDT by the receiving module.

The local alignment is performed approximately 100 ms after the rising edge of the LIVE signal if there is no optical link error reported. The alignment is done with a special alignment trigger sent out by the Top-CDT. The ADCs will then send specific alignment data through the trigger system. The time difference between when the alignment data is received and

when the alignment trigger is received gives the latency. This latency is evaluated at the crate level to ensure the trigger data for the entire crate of 16 ADCs is aligned before being sent to the Et-OFC and the Clustering-OFC to evaluate the trigger variables. If any module does not receive the alignment data, then an error is reported to the Top-CDT and no triggers are issued during that spill cycle.

The global alignment is performed after the local alignment to align the calorimeter timing with the veto timing. The time required for the E_t evaluation for both cases may be different. The global alignment is initiated in the Top-CDT, which will issue a different special trigger called a "delta" trigger. When an ADC receives this trigger, it replaces a portion of the internal pipeline with a delta function, as shown in Fig. 3.3.

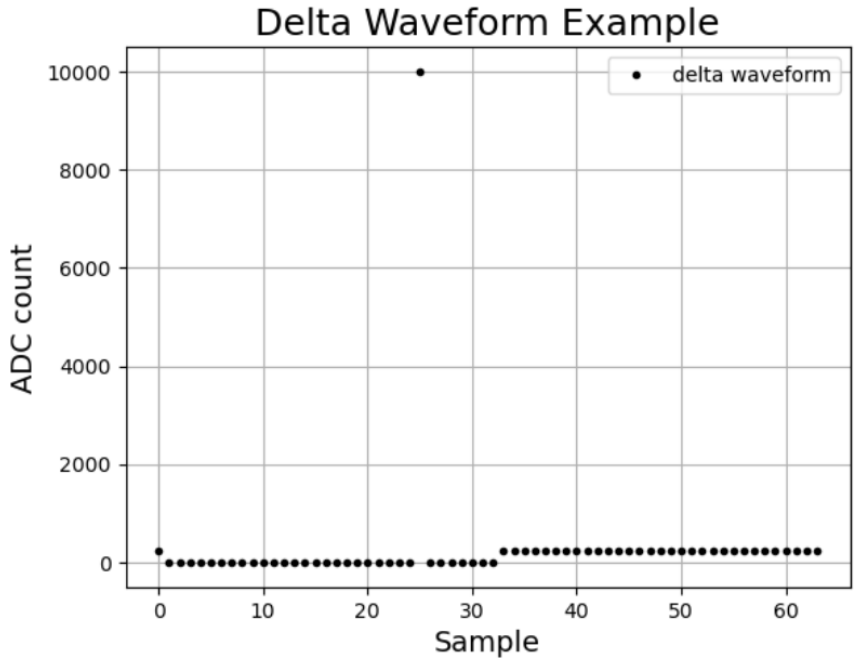


Figure 3.3: An example of the waveform generated when an ADC receives the delta trigger.

The first sample is the online pedestal followed by 0s, with one sample at 10000 ADC counts. The trigger data for this waveform is sent through the trigger evaluation, and the Top-CDT should receive a delta function from both the calorimeter E_t result and the veto result because the E_t will pass for only one clock. The time difference between these results

is the alignment result.

3.1 Run Types

KOTO utilizes the DAQ to perform many different types of data taking. These fall into three main categories,

- Calibration Runs
- Dedicated Background Study Runs
- Physics Runs

The majority of the beam time is dedicated to Physics runs, and most of the calibration data can be obtained without the beam. The full distribution of run types from data taking in the Spring of 2021 (Run87), which is the main focus of this analysis, is shown in Fig. 3.4.

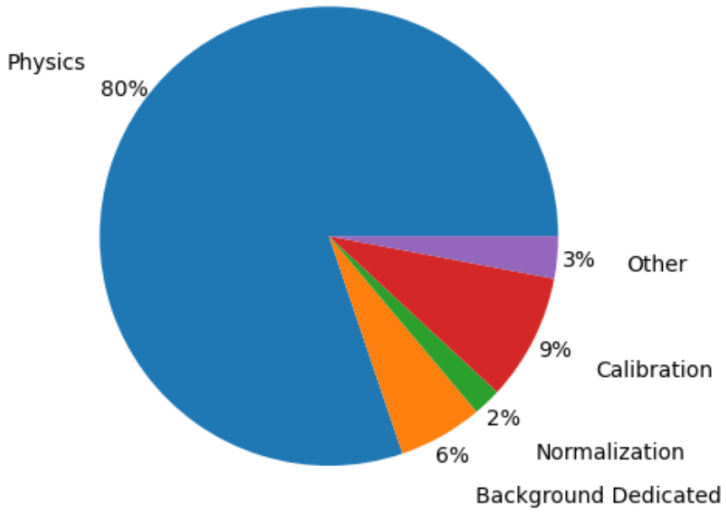


Figure 3.4: The percentage distribution of the type of data taking runs during Run87

Calibration Runs

The purpose of these runs is to perform timing and energy calibration for the calorimeter and veto detectors. The majority of these runs require a large amount of charged particles traveling through them to achieve the calibration result. For the most part, this can be done without the beam by taking cosmic ray data. By using cosmic muons over the course of a week prior to the scheduled beam time, the working analysis group can calculate preliminary calibration constants used in the trigger. This data can be simply collected by requiring hits in the VETO or calorimeter while there is no beam. An example of the real time generated trigger table for a cosmic muon data taking run is shown in Fig. 3.5.

On the left hand side is the description of the trigger, which shows if a detector is treated as a VETO ("!") or if a hit is required. In this case, no detectors are treated as a veto, and events with hits in the CSI, NCC, IB, CC04, CC05, or CC06 were triggered on. The number passing the trigger is shown in the second to last column.

#Trigger for each Trigger Type

Trigger Description	CDT Number	Pedestal Supp.	Prescale Factor(A)	Prescale Factor(B)	Lv1 Scaled(A)	Lv1 wCDTcls Scaled(B)	Raw/ Accepted
Trig0:[CsIEt]	-	-	1	1	698	541	1.3
Trig1:[NCC]	-	-	1	1	18,360	1,635	11.2
Trig2:[IB]	-	ON	1	1	14,932	2,626	5.7
Trig3:[CC04 | CC05 | CC06]	-	-	1	1	3,361	556	6.0

Figure 3.5: Cosmic muon trigger table from Run87.

However, there are many other types of runs dedicated to calibration. For instance, some detectors near the beam axis do not receive enough cosmic muons in the beam direction for calibration. Thus, a small percentage of the data taking time is necessary for beam muon runs. These runs are done by inserting a beam plug between the upstream collimators to block out most particles except muons. In addition, LEDs and lasers, which are made to flash simultaneously, are used during special runs for the timing calibration of special detectors in the beam downstream of the CSI. Lastly, an aluminum target is lowered inside the decay

fiducial region in order to generate a large number of pions from a particular z-position. These pions will promptly decay to two photons, and knowing the precise decay vertex at the aluminum target allows a more precise calibration of the CSI calorimeter.

Dedicated Background Study Runs

In Run87, there were two background types which required dedicated data taking runs. The first was related to beam neutrons. Since neutrons are long lived and neutral, they make it into the detector volume abundantly and represent a significant source of background if they interact with the KOTO calorimeter. Thus, dedicated runs to produce copious neutrons are needed in order to study these interactions and the rejection efficiency using hardware and analysis techniques. This is done by inserting a thin (3mm) aluminum target in the beam that scatters and creates beam neutrons. Since the aluminum target is lowered before the detector, contamination from pion generation and decay is much lower compared to the calibration runs described in the above section. A not-to-scale layout of this situation is shown in Fig. 3.6. Besides the neutron, a major background was discovered during the

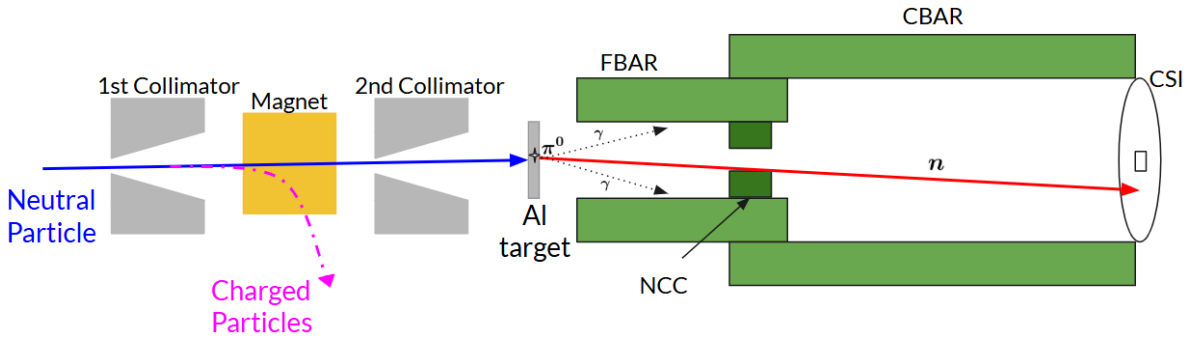


Figure 3.6: The detector layout when the Z0 aluminum target is lowered. It sits upstream of the barrel counters.

analysis of the 2016-2018 data taking runs. This background is induced by $K^{\pm}$ contamination in the beamline due to interactions with the 2nd collimator. The Upstream Charged Veto (UCV) was introduced in 2021 to VETO events with on-time charged particles upstream. To

study the rejection efficiency of this detector, special runs were done with the D1 magnet, shown in Fig. 3.6, turned off, allowing charged particles, in particular $K^{\pm}$, to enter the decay region. The detection efficiency can be found by pairing on-time hits in the UCV with reconstructed $K^{\pm} \to \pi^{\pm}\pi^0$ decays.

Physics Runs

The majority of the data taking time is done with a focus towards $K_L \to \pi^0\nu\overline{\nu}$. However, other triggers are utilized during these runs while minimizing DAQ dead-time and loss. A pre-scale can be applied to any trigger, meaning that only one event is triggered out of the scale factor number of events. For example, if there is a prescale of 30, only one event out of 30 is taken. The purpose of the prescale is to include triggers that have too high of a rate and would overwhelm the DAQ system. An example of the trigger table is shown in Fig. 3.7. There can be up to eight trigger conditions specified from the Top-CDT. For physics

#Trigger for each Trigger Type

Trigger Description	CDT Number	Pedestal Supp.	Prescale Factor(A)	Prescale Factor(B)	Lv1 Scaled(A)	Lv1 wCDTcls Scaled(B)	Raw/ Accepted
Trig0:[CsIEt !(CV | CC03 | CC04 | CC05 | CC06 | ncc | mb | ib)]	2	-	1	1	12,460	1,386	9.0
Trig1:[CsIEt !(CV | CC03 | CC04 | CC05 | CC06 | ncc | mb | ib)]	-	-	30	1	416	412	1.0
Trig2:[CsIEt]	-	-	900	1	455	446	1.0
Trig3:[CsIEt & CV !(FCVHitsA | CC03 | CC04 | CC05 | CC06 | NCC | MB | IB | RCVHitsA)]	3	ON	1	1	21,836	3,819	5.7
Trig4:[CsIEt & CV !(CC03 | NCC | MB | IB)]	3	-	1	30	112,979	761	148.5
Trig5:[CsIEt !(CV | CC03 | CC04 | CC05 | CC06 | ncc | mb | ib)]	6	ON	1	1	12,460	1,585	7.9
Trig6:[CsIEt OnlyOffSpill]	-	-	1	1	232	208	1.1
Trig7:[CsIEt & CV !(CC03 | CC04 | CC05 | CC06 | ncc | mb | ib)]	6 7	ON	1	1	101,015	1,131	89.3

Figure 3.7: The trigger table for physics runs taken during Run87. The most relevant triggers are highlighted as follows: $K_L \to \pi^0\nu\overline{\nu}$ trigger (red), normalization trigger (black), $K^{\pm}$ dedicated trigger (blue), $K_L \to 2\pi^0\gamma\gamma$ trigger (green)

runs, these trigger conditions are typically all used. Trig0, the $K_L \to \pi^0\nu\overline{\nu}$ trigger, requires two clusters in the CSI and no prescale. This is to maximize the number of signal obtained. The normalization trigger (Trig1) is the same as Trig0 with no requirement on the number of clusters. The purpose is to gather the K_L decays used in normalization such as $K_L \to 3\pi^0$, $K_L \to 2\pi^0$, and $K_L \to 2\gamma$. $K^{\pm}$ dedicated triggers (Trig3 and Trig4) have a CV detector hit requirement, the opposite of VETO, while also needing three clusters in the CSI. This is

to calculate the $K^{\pm}$ flux during physics runs using $K^{\pm} \to \pi^{\pm}\pi^0$ decays. Lastly, the trigger used to study $K_L \to \pi^0\pi^0\gamma\gamma$ (Trig5) is the same as $K_L \to \pi^0\nu\overline{\nu}$ but with a six cluster requirement. Other triggers are used for minimum bias data collection (Trig2) to study the online veto, and the CLOCK/TMON trigger (Trig6) used to study accidental rates. The last trigger was used for the novel study of $K_L \to \pi^0 e^+e^-e^+e^-$, and the precision measurement of π^0 dalitz decays in $K_L \to 3\pi^0$.

3.2 Run History

KOTO began taking data in 2013 and is still taking data to this day. The run history is shown in Fig. 3.8. The data taken in 2013, 2015, 2016-2018, and 2021 has been analyzed

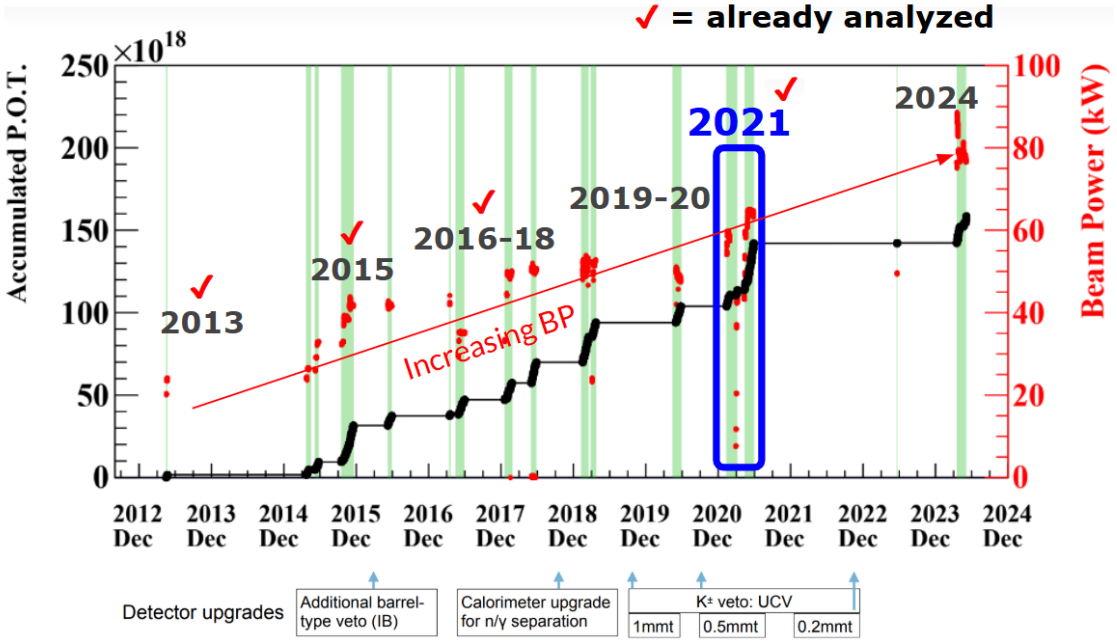


Figure 3.8: The run history in terms of beam intensity and protons on target (POT) up to the year 2024. Data already analyzed for $K_L \to \pi^0\nu\overline{\nu}$ is labeled with a check mark.

and published for $K_L \to \pi^0\nu\overline{\nu}$ [15] [1] [2] [4]. In 2016-2018, three candidate events were observed leading to the discovery of new background sources, in particular those related to $K^{\pm}$ decays. As a result, the upstream charged veto was developed and installed before

data taking in 2021, but was still absent during 2019-2020 (besides the prototype installed in 2020). Therefore, the $K^{\pm}$ contamination is still high for those runs, and 2021 data was analyzed first. In 2021, the average beam power was approximately 60kW, which is more than a 20% increase from the average of 50kW from 2016-2018. In the most recent data taking runs in 2025, the average beam intensity has been over 80kW. The total POT from 2021 is slightly larger than that from 2016-2018, making it the largest dataset KOTO has analyzed. This data is the focus of this dissertation.

CHAPTER 4

MONTE CARLO SIMULATION AND EVENT BUILDING

Given the detector material properties and the physics process interaction cross section, it is possible to fully simulate btoh the kaon decays and detector responses in KOTO. This simulation is crucial to understand the signal acceptance, background level and rejection, and background mechanisms. By carefully examining the real performance of the KOTO detectors, and simulating these behaviors, the KOTO analysis group can use simulation to evaluate the efficacy of the Monte Carlo in reproducing the real situation in data. On the other hand, discrepancies between data and Monte Carlo imply that something is missing in the analysis group's understanding of the process.

The KOTO group uses the GEANT4 toolkit [9] [7] to simulate particle interactions in the detector, and GEANT3 for the entire beamline simulation. From the particle interactions, the detector response and event reconstruction is simulated within the KOTO analysis library that is written in C/C++. The details are described in the following subsections.

4.1 K_L Beamline Simulation

The K_L beamline simulation is done in two parts. The first part is the Monte Carlo using the GEANT3 toolkit, which was used to determine the beamline parameters during the construction of the KOTO beamline and currently for the estimations of particle flux entering the K_L detector region. This is particularly important in preliminary estimations of $K^{\pm}$ entering the decay region. The second method uses data driven techniques during 2012 data taking runs to obtain K_L momentum spectrum and beam structure, to use for KOTO analysis using GEANT4 simulation.

4.1.1 GEANT3 simulation

The beam x-y position and K_L momentum as it enters the decay region is incredibly important. As the principle measurement of $K_L \to \pi^0\nu\overline{\nu}$ relies on missing momentum off the beam axis, it is critical that KOTO minimizes the off-axis momentum of the K_L. The beamline simulation using GEANT3 was used to create a tight beam profile, and the beam x-y position is shown in Fig. 4.1. The vast majority of particles entering the decay region

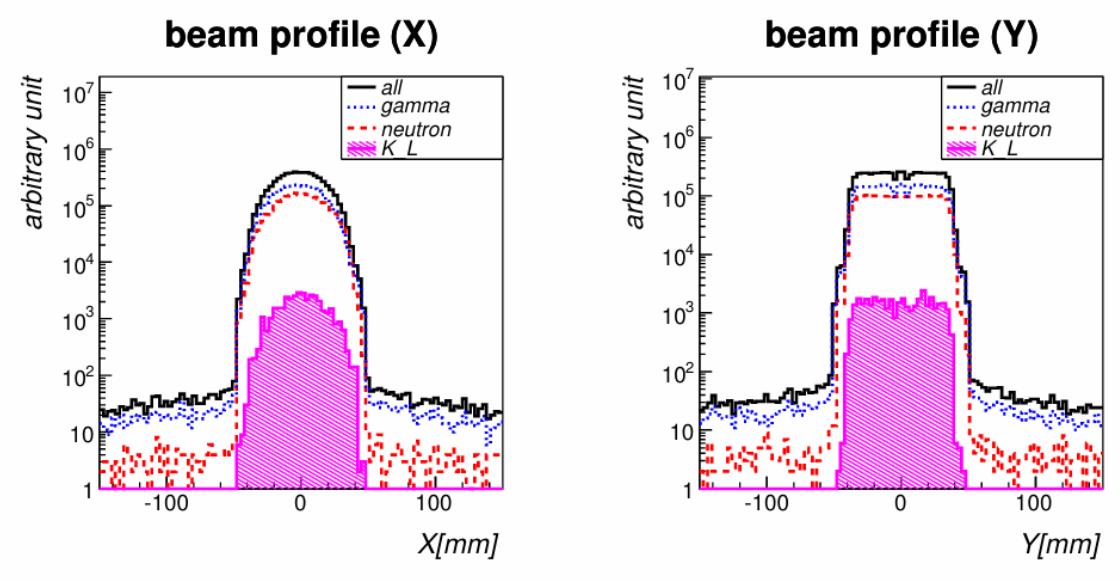


Figure 4.1: The result of beam x-y position at the entrance of the KOTO detector based on GEANT3 simulation [44]

consists of the particles shown in Fig. 4.1, which are the long lived neutral particles that propagate past the D1 magnet and the full 21 meters from the target. However, interactions in the second collimator can create multiple secondary particles such as $K^{\pm}$. The relative proportion of these particles is shown in Fig. 4.2, and notably the ratio of K_L to $K^{\pm}$ is on the order of 10^{-5}.

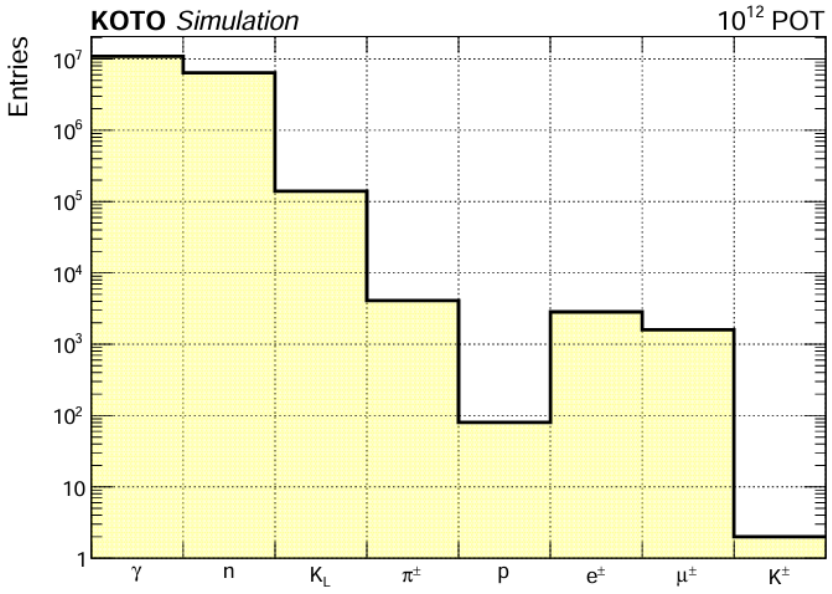


Figure 4.2: Relative amount of different particles at the entrance of the KOTO detector. Figure curtesy of Chieh Lin. [33]

4.1.2 Empirical K_L simulation

Before official data taking targeting $K_L \to \pi^0\nu\overline{\nu}$, there were engineering runs in 2012 in JPARC to study the beamline and detector performance. One of the purposes of these runs was to measure the K_L beam spectrum and momentum empirically. The K_L modes used in this analysis were $K_L \to \pi^+\pi^-\pi^0$ and $K_L \to \pi^+\pi^-$. The momentum of the charged particles was measured using a spectrometer (drift chamber), while the photon momentum was reconstructed from measurements in the CsI [44]. From this information, the K_L momentum can be reconstructed, and by mapping back using the geometrical acceptance from Monte Carlo, the original K_L momentum distribution was obtained.

This distribution is fitted by an asymmetric gaussian with four free parameters. The form is shown in Eq. 4.1

$$f(p) = \exp\left(-\frac{(p-\mu)^2}{\sigma(p)}\right) \tag{4.1}$$

The function $\sigma(p)$, is defined in Eq. 4.2

$$\sigma(p) = a \cdot (1 - (b + cp)(p - \mu)) \tag{4.2}$$

The fitted parameters were obtained as, $\mu = 1.42$, $a = 0.81$, $b = -0.301$, and $c = 0.017$. Thus, the resulting K_L momentum spectrum is peaked at around 1.4 GeV, and is shown in Fig. 4.3.

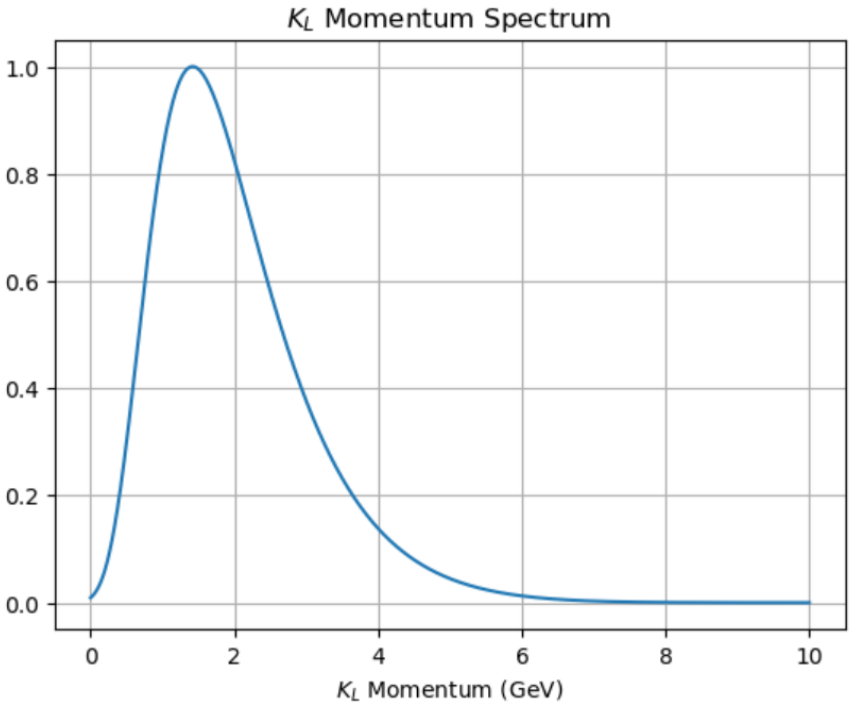


Figure 4.3: Empirical K_L momentum spectrum obtained from engineering runs taken in 2012 analyzing $K_L \to \pi^+\pi^-$ using drift chambers. The y-axis is in arbitrary units and the x-axis is in GeV.

4.2 GEANT4 and Detector Response Simulation

The GEANT4 simulation relies on detailed geometrical descriptions of the KOTO detector, such as the volume of space, and the material of a detector. Using these descriptions, and data derived interaction cross sections, the GEANT4 toolkit allows for step-by-step

particle tracking and generation. For instance, Fig. 4.4 shows the GEANT4 tracking of a 2 GeV electron incident on a liquid Argon volume denoted by the blue box. The energy deposited and particle production is calculated at each step for the electromagnetic shower, and this information along with the timing and position are stored in the simulation files. This information is then used to simulate the detector response. For instance, the deposit

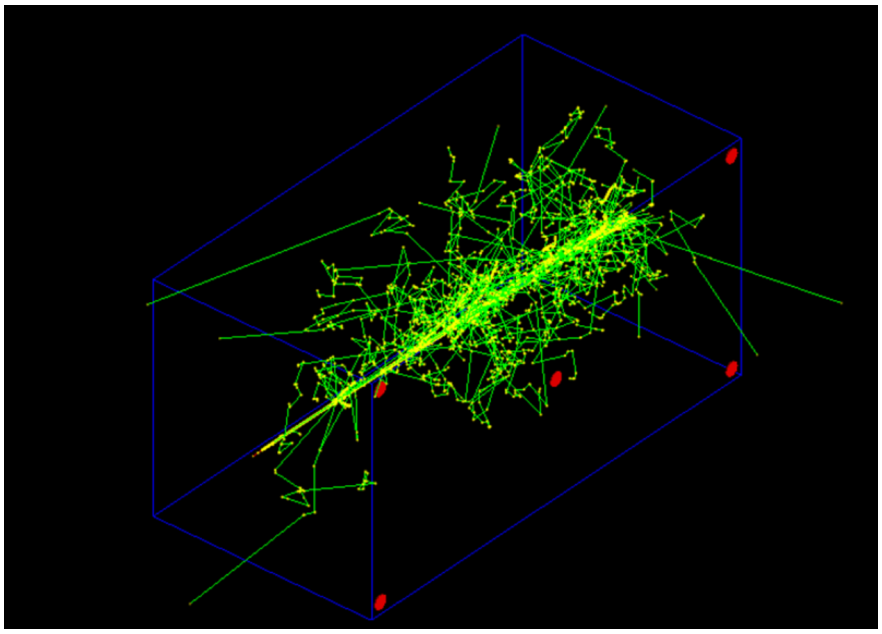


Figure 4.4: A visual representation of the GEANT4 particle tracking and generation from a 2 GeV electron shower in liquid Argon [50].

energy at a specific time and position manifests itself in terms of scintillation light, where the number of generated photons is given by the Poisson distribution, and depends on the light yield for a given material. Furthermore, the fraction of scintillation light which reaches the photomultiplier tube (PMT), and the timing of its arrival depends on the location. The timing and attenuation can also be modeled and simulated, as well as any non-uniformity of the material, but in reality the KOTO group does a fit with real data to obtain the energy and timing from simulation. For example, the simulated energy measurement from the interactions in a detector, like the CSI can be modeled by,

$$E_{sim} = \frac{C}{Y_{light}} \cdot \text{Pois}(A(z) \cdot E_{dep} \cdot Y_{light}) \tag{4.3}$$

where Pois() is the Poisson distribution, Y_{light} is the light yield, $A(z)$ is the attenuation function that is fitted with data, E_{dep} is the energy deposited in the detector, and C is a fitted constant. The attenuation function is particularly important for the barrel detectors which have readouts potentially quite far from the point of interaction. The function depends on the detector, but for the main barrel this attenuation function is given in Eq. 4.4.

$$A(z) = \begin{cases} \frac{1}{2}\exp(\frac{z_c - z}{\Lambda + \alpha(z - z_c)}) & \text{upstream readout} \\ \frac{1}{2}\exp(\frac{z - z_c}{\Lambda - \alpha(z - z_c)}) & \text{downstream readout} \end{cases} \tag{4.4}$$

The z_c is the z-position center of the detector, while Λ and α are fitted parameters. An example of the fitted attenuation function from data taken in 2013 is shown in Fig. 4.5

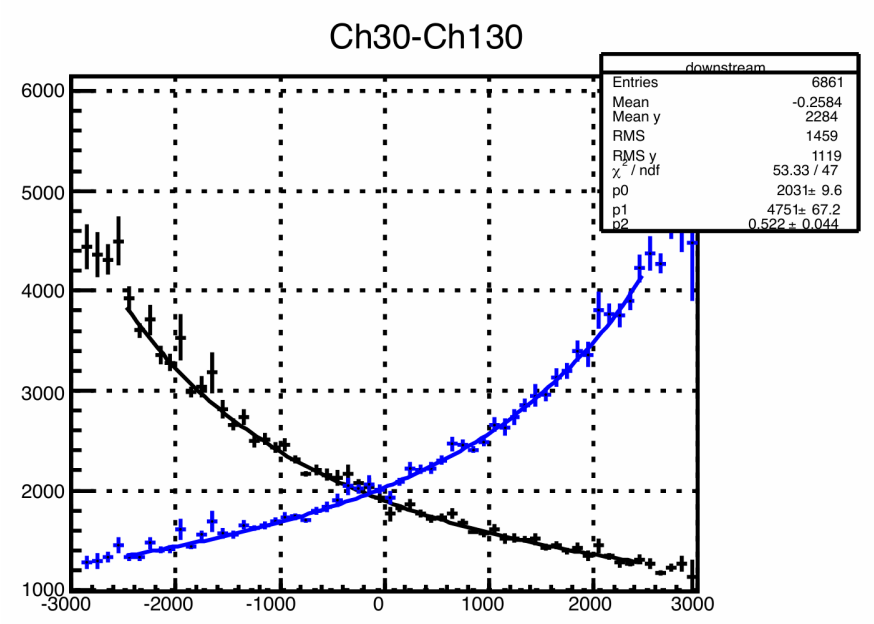


Figure 4.5: The fitting of the attenuation function for a specific channel in the main barrel from data taken in 2013 and fitted by Takahiko Masuda in an internal meeting.

Similar to the energy, the timing resolution can be simulated. This is also affected by

many factors such as non-uniformity in the detector material. The width of the timing resolution is modeled by Eq. 4.5, depending on the energy deposited in the detector.

$$\sigma_T(E_{dep}) = \frac{c_1}{E_{dep}} \oplus \frac{c_2}{\sqrt{E_{dep}}} \oplus c_3 \tag{4.5}$$

The parameters c_1, c_2, and c_3 are obtained from fitting, and the simulated pulse timing is smeared by a gaussian with a width given by σ_T.

Lastly, the pulse shape is simulated using a library of pulses taken from real data. This library is obtained for each detector, at a range of different energies to build the template library. Given a deposit energy and timing using Eq. 4.3 and Eq. 4.5, the discrete pulse samples may be generated from the template using a cubic spline interpolation. This is done because the true peak of the curve may not be represented in a specific sample in the 125 MHz ADC pulse.

4.3 Accidental Overlay

In real data, there are numerous interactions, typically caused by a stray neutron or gamma, in the KOTO veto and calorimeter which may occur at the same time as a real kaon decay interaction. When this happens on time in a veto detector, then the event is rejected, or worse, accidental pulses may mask on-time hits in the veto by overlapping the pulse. An example can be seen in Fig. 4.6

Thus, it is crucial to understand and simulate these types of interactions in the detector. The simulation of accidental hits is done by overlaying the result from real data. To pick out typical accidental events in data, KOTO uses a special trigger called "TMON" which stands for "target monitor". This monitor tracks the proton cycle in the accelerator which is highly correlated to the accidental hits. The monitor consists of a plastic scintillator counter set behind the primary beamline at 50°. Whenever a hit is measured in the target monitor, a

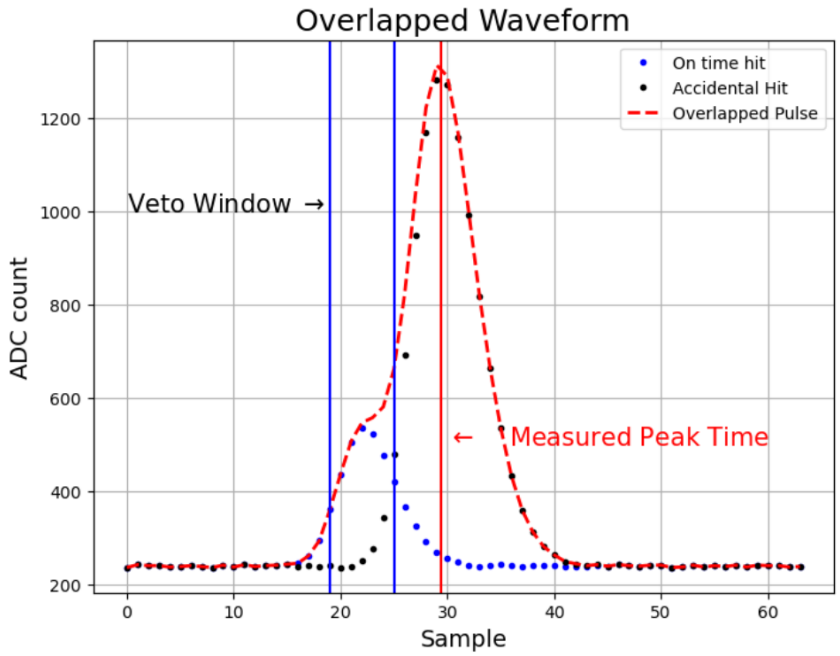


Figure 4.6: An example of an overlapped waveform. The high energy accidental hit overlaps a lower energy on time hit to mask the real timing.

corresponding external trigger is issued to the Top-CDT to gather data. The trigger timing is accounted for based on the path travel difference between the trigger monitor and KOTO detector.

This data is used to build a library of accidental hits, which is then overplayed onto the simulated pulse data. Using this method, the KOTO analysis group can study the loss due to accidental hits, and the probability of overlapped pulses leading to timing mismeasurement.

CHAPTER 5

EVENT RECONSTRUCTION

The event reconstruction corresponds to the calculation of complex analysis variables from the simulated (measured) waveforms in simulation (data). For instance, detailed calculations of the reconstructed energy and position of clusters in the CSI allow the evaluation of the decay vertex position and momentum of the K_L. In addition, variables used for the discrimination of the background and signal are evaluated.

5.1 Photon Energy and Hit Position Reconstruction

The first step in calculating these complex variables is the photon energy and hit position reconstruction. A high energy photon will typically have energy deposited at a specific time in multiple channels via an electromagnetic shower. The information that must be decoded is present in the waveform pulse in these hit channels. The waveform pulse consists of an area, which is proportional to the energy deposit, and a hit time, which is the peak position of the pulse. Each channel may have variable gain, and so the conversion from pulse area to energy requires calibration for each channel. To evaluate the waveform pulse area, the pedestal is subtracted. Instead of using the online pedestal, the offline pedestal calculation uses the average of the first ten samples, or the average of the last ten samples, depending on which has the lowest standard deviation.

The pulse area is evaluated via the expression shown in Eq. 5.1. The sample value is represented by s_i, the pedestal value is P, and the calibration constant is C and is calculated channel by channel. The factor of $L(H_{peak})$ corresponds to the non-linearity factor as a function of peak height. This is a second order effect, where at large pulse areas or peak heights, the energy diverges from proportionality. This factor is evaluated based on the pulse height.

$$A = \sum_{i=0}^{n} (s_i - P) \cdot \frac{C}{L(H_{peak})} \tag{5.1}$$

For most detectors, the value n ranges over the entire 64 sample waveform, while for some such as the beam hole detectors, the summation runs only in a small window around the peak sample. The constant C is obtained first during data taking using cosmic rays and later modified by a scale factor during DVU calibration data taking runs. An example of the data fitting for C from cosmic rays is shown in Fig. 5.1.

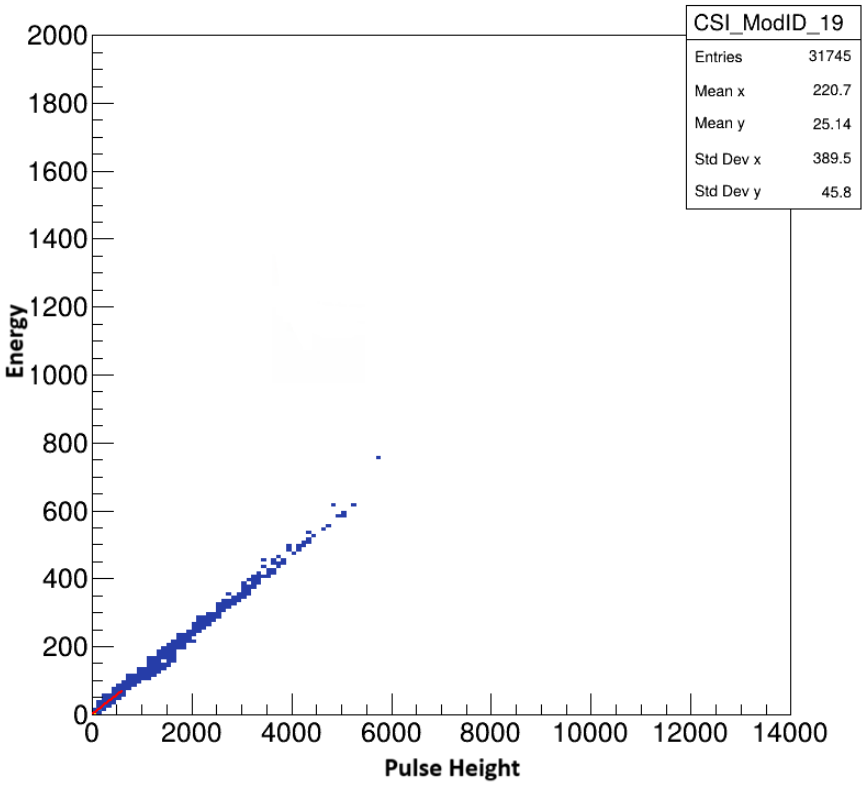


Figure 5.1: An example of the linear fit done to evaluate the calibration constant for a CSI channel (module ID == 19). The linear fit is shown in red, and uses small peak heights for the best result.

The timing for most detectors is evaluated by fitting a parabola to the peak sample and its two neighbors. One downside of this method is its sensitivity to pedestal fluctuations, so the

5 sample moving average is used instead to minimize the effect caused by these fluctuations.

5.2 Offline Clustering Algorithm

Once the energy and timing of each channel in the calorimeter are measured, the individual clusters are defined, counted, and have their respective energies and positions calculated. The offline algorithm is similar to that described in Section 2.7. Each CsI crystal with an energy deposit larger than 3 MeV is taken as a cluster seed. Centered at each cluster seed, a 14cm x 14cm square range is searched for additional seeds. If they are found, they are attributed to the same cluster and the search continues from this seed. This search range is decided based on two small crystal (2.5cm x 2.5cm) widths. The 14cm x 14cm area is chosen, instead of 10cm x 10cm, because the search range for a large crystal (5cm x 5cm) on the small-large crystal boundary encompasses two small crystal lengths, and one large crystal length. If a cluster seed is isolated from others, it is not labeled as a cluster and is instead labeled as an isolated crystal hit. These hits are typical of hadronic interactions in the CsI from neutrons, and if it is on-time with other clusters then this event is vetoed. Conversely, if a cluster seed does not fall within the correct timing window, it is regarded as an accidental. An simple diagram is shown in Fig. 5.2 showing the algorithm logic.

The timing of the clusters is determined by the average weighted time for each channel, where the weight is the timing uncertainty derived from the prototype CsI study [43]. This uncertainty is a function of energy and is shown in Eq. 2.8. Thus, the timing of a cluster is determined by,

$$T = \frac{\sum_{i=0}^{n} T_i/(\sigma_t(E_i))^2}{1/(\sigma_t(E_i))^2} \tag{5.2}$$

T_i is the channel time, and i runs to n where n is the number of channels in a given cluster. σ_t is a function of E_i, the energy deposit in the channel. If a crystal is more than

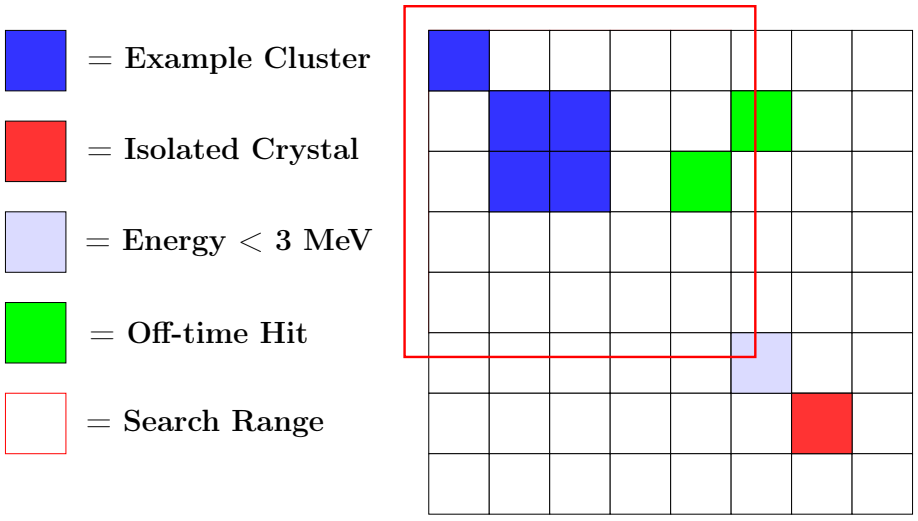


Figure 5.2: Diagram explaining the offline clustering algorithm. The blue blocks represent a cluster of hit channels with energy deposit above 3 MeV. The red crystal is isolated, because its neighbor has energy deposit less than 3 MeV. The green colored crystals have a time measurement outside the timing window.

$5\sigma_t$ away from the cluster timing, it is considered an accidental hit and removed from the cluster energy, timing, and position calculations. The cluster energy is simply the deposit energy sum over crystals in that cluster, while the x and y positions are the energy weighted positions, as shown in Eq. 5.3 and 5.4.

$$X = \frac{\sum_{i=0}^{n} E_i \cdot x_i}{\sum_{i=0}^{n} E_i} \tag{5.3}$$

$$Y = \frac{\sum_{i=0}^{n} E_i \cdot y_i}{\sum_{i=0}^{n} E_i} \tag{5.4}$$

5.3 $K_L \to \pi^0 \nu\overline{\nu}$ Reconstruction

This mode is the principle mode of study in KOTO. There are two neutrinos in the final state which escape undetected. This means that the K_L momentum cannot fully be reconstructed. However, the decay z-vertex can be determined assuming a decay on the beam axis. The

pion decays promptly into two photons, and to good approximation the K_L decay vertex is the π^0 decay vertex. Assuming the pion mass and conservation of momentum, the opening angle can be determined via Eq. 5.5.

$$M_{\pi^0}^2 = 2E_1E_2(1-\cos\theta) \tag{5.5}$$

From the opening angle, the hit positions on the CsI geometry can be used to calculate the decay vertex along the beam axis. In principle, there are four possible vertex solutions for a given opening angle, but nonphysical solutions are abandoned, such as those with a decay vertex behind the calorimeter. If there are two reasonable solutions, then the event is removed. Once the vertex is found, the momenta of the pion is calculated from the reconstructed cluster energies and positions in the CSI.

5.4 $K_L \to \pi^0\pi^0\gamma\gamma$ Reconstruction

The $K_L \to \pi^0\pi^0\gamma\gamma$ reconstruction is fundamentally different from the $K_L \to \pi^0\nu\overline{\nu}$ case, and is novel in KOTO. Here, there are no missing particles in the final state, and there are multiple combinations of two photons to pair with the two pions. For instance, given six photons in the final state the number of possible orderings are given in Eq. 5.6.

$$N_{pairings} = \frac{\binom{6}{2}\cdot\binom{4}{2}}{2} = 45 \tag{5.6}$$

The factor of two comes from the redundancy of pairing. As an example, pairing photons 1 & 2 is the same as pairing 2 & 1. The others come from the combinatorics of first choosing which two photons did not come from a π^0 (6 choose 2), and then the number of combinations of pairing the four leftover photons into two pairs (4 choose 2).

Because of all the different combinations, mispairings are a major issue both for reconstructing the signal, and discriminating from the background $K_L \to 3\pi^0$. To protect against

mispairings, the reconstruction of $K_L \to \pi^0\pi^0\gamma\gamma$ is done using a constraint fit strategy. This is the first analysis in KOTO to utilize this method of reconstruction. This method is well suited to this analysis, because there are no missing particles in the final state of six photons. Thus, all kinematics of the event should be well reconstructed. In total, there are five constraints applied in the reconstruction which are listed below.

1. K_L mass constraint: The invariant mass of the six photons should be the K_L mass.
2. π^0 mass constraint: The invariant mass of (γ_1, γ_2) should be exactly the pion mass.
3. 2nd π^0 mass constraint: The invariant mass of (γ_3, γ_4) should be exactly the pion mass.
4. Momentum conservation in x-direction: The K_L momentum in the x direction should lie collinear with the line starting from the fixed target to the x location of the center of energy on the CsI.
5. Momentum conservation in y-direction: The K_L momentum in the y direction should lie collinear with the line starting from the fixed target to the y location of energy on the CsI.

The process for this reconstruction is done in two stages. First, the reconstructed photon hit positions and energies are constructed as described in Section 5.2, and the pion decay vertex is obtained via the strategy described in Section 5.3 for each of the two pions. The K_L decay vertex is then the average of these two results, and the kinematics of the pions and kaon are reconstructed under this assumption.

The second stage of the reconstruction utilizes the constraints to modify the observables, such as kaon and pion momentum, and photon hit position and energy, to best fit the constraints. Knowing the position and energy resolution in the calorimeter, the "goodness" of the constrained fit can be quantified in a variable χ^2. The pairing which returns the smallest χ^2 value is chosen as the best fit for the reconstruction.

In details, the constrained fit uses the mathematical concept of Lagrange Multipliers to optimize the fitting result. For instance, if there is a multivariate function $f(x_1, ..., x_n)$, then it is simple to find the extremums by finding the point $(X_1, ..., X_n)$ such that,

$$\begin{aligned} \frac{\partial f(X_1, ..., X_n)}{\partial X_1} &= 0 \\ &\vdots \\ \frac{\partial f(X_1, ..., X_n)}{\partial X_n} &= 0 \end{aligned} \tag{5.7}$$

However, if this extremum is further required to have the point $(X_1, ..., X_n)$ solve some constraint function such that $H(X_1, ..., X_n) = 0$, then in addition to solving Equations listed in 5.7, the solution must additionally solve the following equations.

$$\nabla_{x_1,...,x_n} f(X_1, ..., X_n) = \lambda \nabla_{x_1,...,x_n} H(X_1, ..., X_n) \tag{5.8}$$

$$H(X_1, ..., X_n) = 0 \tag{5.9}$$

Or similarly, if a new function $F(x_1, ..., x_n)$ is defined such that,

$$F(x_1, ..., x_n, \lambda) = f(x_1, ..., x_n) + \lambda H(x_1, ..., x_n) \tag{5.10}$$

Then Equations 5.8 and 5.9 are satisfied if,

$$\nabla_{x_1,...,x_n,\lambda} F(X_1, ..., X_n, \lambda) = 0 \tag{5.11}$$

In the case of the reconstruction of $K_L \rightarrow \pi^0\pi^0\gamma\gamma$, the function that is minimized determines the χ^2, since ideally the reconstructed observables are close to the point that satisfies the constraints. This is a function of the photon x, y hit positions on the CsI, the

photon energies. Since the uncertainties on these observables are known, the minimizing function is defined in Eq 5.12.

$$\chi^2(x_1, y_1, E_1, ..., x_6, y_6, E_6) = \sum_{i=1}^{6} \frac{(x_i - x_{i,m})^2}{\sigma_{x_i}^2} + \sum_{i=1}^{6} \frac{(y_i - y_{i,m})^2}{\sigma_{y_i}^2} + \sum_{i=1}^{6} \frac{(E_i - E_{i,m})^2}{\sigma_{E_i}^2} \quad (5.12)$$

In the above equation, $x_{i,m}, y_{i,m}$ and $E_{i,m}$ represent the measured reconstructed position or energy of the photon hit. This may also be represented as the multiplication of matrices,

$$\chi^2 = (\alpha - \alpha_m)^T V^{-1} (\alpha - \alpha_m) \quad (5.13)$$

Where the matrices are defined as,

$$\alpha = \begin{pmatrix} x_1 \\ y_1 \\ E_1 \\ \vdots \\ x_6 \\ y_6 \\ E_6 \end{pmatrix}; \quad \alpha_m = \begin{pmatrix} x_{1,m} \\ y_{1,m} \\ E_{1,m} \\ \vdots \\ x_{6,m} \\ y_{6,m} \\ E_{6,m} \end{pmatrix}; \quad \text{and} \quad V = \begin{pmatrix} \sigma_{x_1}^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_{y_1}^2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_{E_1}^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & & \sigma_{E_6}^2 \end{pmatrix} \quad (5.14)$$

The solution, α, must also satisfy the constraints, which will be labeled H_i. As mentioned

previously, there are five constraints listed below.

$$H_1 = \left(\sum_{i=1}^{6} E_i\right)^2 - \left(\sum_{i=1}^{6} p_i\right)^2 - M_{K_L}^2 \tag{5.15}$$

$$H_2 = (E_{\gamma_1} + E_{\gamma_2})^2 - (p_{\gamma_1} + p_{\gamma_2})^2 - M_{\pi^0}^2 = 0 \tag{5.16}$$

$$H_3 = (E_{\gamma_3} + E_{\gamma_4})^2 - (p_{\gamma_3} + p_{\gamma_4})^2 - M_{\pi^0}^2 = 0 \tag{5.17}$$

$$H_4 = \frac{v_x}{L_{v_z}} - \frac{CoE_x}{L_{CSI}} = 0 \tag{5.18}$$

$$H_5 = \frac{v_y}{L_{v_z}} - \frac{CoE_y}{L_{CSI}} = 0 \tag{5.19}$$

Here, (v_x, v_y, v_z) correspond to the decay vertex position and L_{v_z} and L_{CSI} represent the distance from the target to the decay z position and CSI respectively. Thus, $H(\alpha) = 0$ where $H(\alpha) = (H_1(\alpha), ...H_5(\alpha)) = (0, ...0)$. Or, if the solution α is expanded around the measured result, then to first order,

$$H(\alpha) \approx H(\alpha_m) + \frac{\partial H(\alpha_m)}{\partial \alpha}(\alpha - \alpha_m) \tag{5.20}$$

If further matrices are defined as,

$$D = \begin{pmatrix} \frac{\partial H_1}{\partial \alpha_1} & \cdots & \frac{\partial H_1}{\partial \alpha_5} \\ \vdots & \ddots & \vdots \\ \frac{\partial H_5}{\partial \alpha_1} & \cdots & \frac{\partial H_5}{\partial \alpha_5} \end{pmatrix}; \quad \text{and} \quad h_m = \begin{pmatrix} H_1(\alpha_m) \\ \vdots \\ H_5(\alpha_m) \end{pmatrix} \tag{5.21}$$

then after this renaming, Eq. 5.20 simply becomes,

$$H(\alpha) \approx H(\alpha_m) + \frac{\partial H(\alpha_m)}{\partial \alpha}(\alpha - \alpha_m) = h_m + D(\alpha - \alpha_m) \tag{5.22}$$

Then the modified minimizing function becomes,

$$\chi^2 = (\alpha - \alpha_m)^T V^{-1} (\alpha - \alpha_m) + \lambda^T (h_m + D(\alpha - \alpha_m)) \tag{5.23}$$

following the strategy from Eq. 5.10. The Lagrange multipliers, λ, is a 5-vector and can be obtained as,

$$\lambda = (DVD^T)^{-1}(h_m + D(\alpha - \alpha_m)) \tag{5.24}$$

The minimizing solution is then given by,

$$\alpha = \alpha_m - VD^T\lambda \tag{5.25}$$

The values obtained from this calculation replace the original values obtained in the reconstruction. The quality of the fit is then the second term of our minimizing function,

$$\chi^2_{fit} = \lambda^T(h_m - D(\alpha - \alpha_m)) \tag{5.26}$$

This is the χ^2 which will be referenced in the analysis sections later in this dissertation. For all 45 combinations, this calculation is done and the reconstruction chosen is the one with the smallest χ^2_{fit} value. The difference in performance is shown in Fig. 5.3.

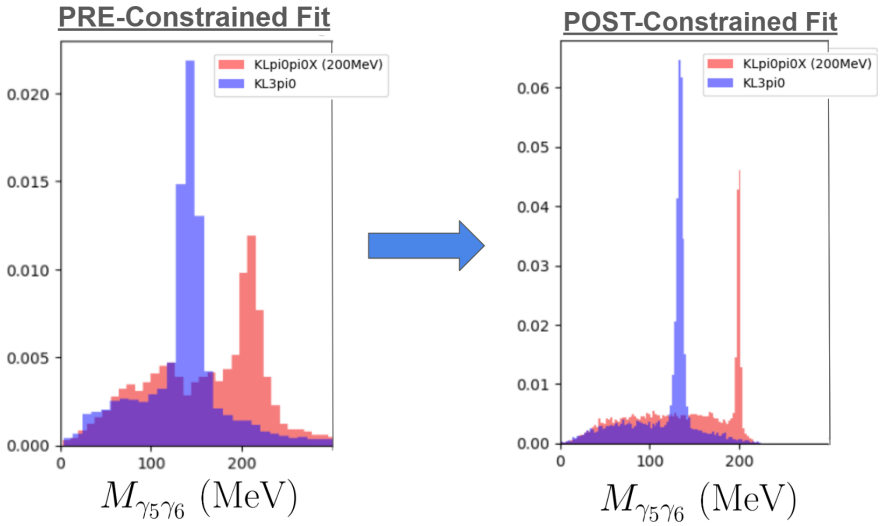


Figure 5.3: The improvement in $M_{\gamma\gamma}$ resolution after implementing the constrained fit algorithm. Shown in blue is the $K_L \to 3\pi^0$ simulation, and shown in red is the $K_L \to \pi^0\pi^0 X$, where $X \to \gamma\gamma$, simulation with an X mass of 200 MeV.

CHAPTER 6

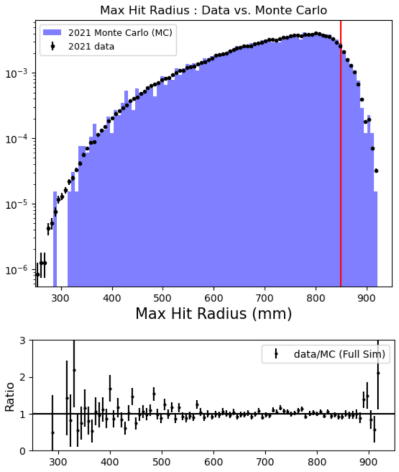
DATA AND MONTE CARLO CONSISTENCY

The Monte Carlo simulation and data collection strategies have been discussed in the previous sections. It is crucial to validate the Monte Carlo samples by comparing with data and curating the event selection process. The primary method employed in this dissertation uses the $K_L \to 3\pi^0$ simulation and data. The purpose of using $K_L \to 3\pi^0$ is three-fold. First, with the BR($K_L \to 3\pi^0$) $\approx 20\%$, it dominates the data with six clusters deposited in the CSI calorimeter. This means it is easy to create data samples that are nearly 100% $K_L \to 3\pi^0$. Second, it is highly constrained because there are no missing particles in the final state, and it has three pions that can be reconstructed. This means that, like $K_L \to \pi^0\pi^0\gamma\gamma$, a constrained fit can be applied to optimize the reconstruction. In the case of $K_L \to 3\pi^0$, there is an additional constraint added to account for the additional pion. Thirdly, the $K_L \to 3\pi^0$ is the dominant background source for $K_L \to \pi^0\pi^0\gamma\gamma$. With this in mind, it is crucial to study the behavior of the background and test the limits of the Monte Carlo in reproducing the data.

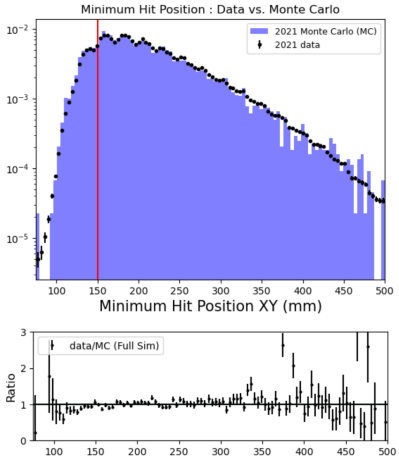
6.1 $K_L \to 3\pi^0$: Data vs. Monte Carlo

There are different types of selection criteria for the signal, commonly called "cuts" in the KOTO analysis. There are quality cuts, that ensure the quality of the data. These cuts are typically related to the selection of the trigger type and removal of events with runtime DAQ errors. The second type are veto cuts which are related to the veto energy and timing cut thresholds for each veto detector and are quite standardized between different analyses. More details are shown in Appendix A. The last are kinematic cuts. This section compares $K_L \to 3\pi^0$ data and Monte Carlo for the kinematic cuts applied in the analysis of $K_L \to \pi^0\pi^0\gamma\gamma$ to confirm the reproducibility of the kinematic cut performance in Monte Carlo simulation.

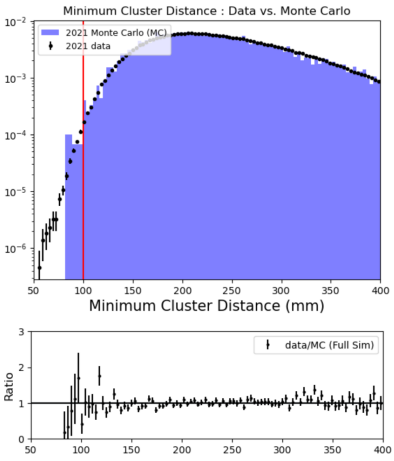
The comparison can be found in Fig. 6.1, where the red vertical line corresponds to the cut value used in the $K_L \to \pi^0\pi^0\gamma\gamma$ analysis.



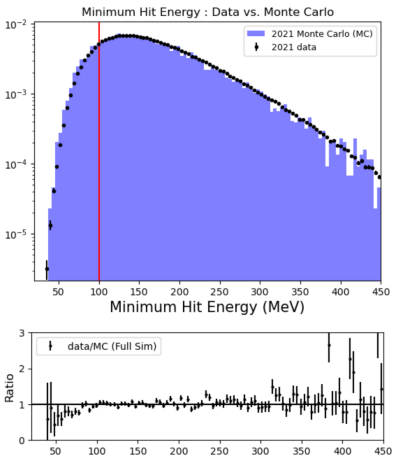
(a) The maximum photon hit radius in the CSI. The x-axis is in mm.



(b) The minimum photon hit X-Y in the CSI. The x-axis is in mm.

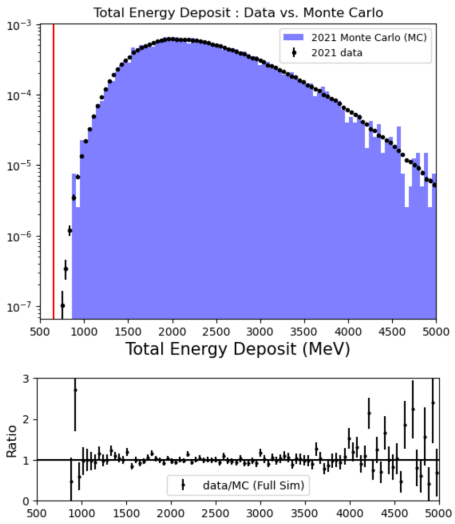


(c) The minimum of all cluster distances from each other in the CSI. The x-axis is in mm.

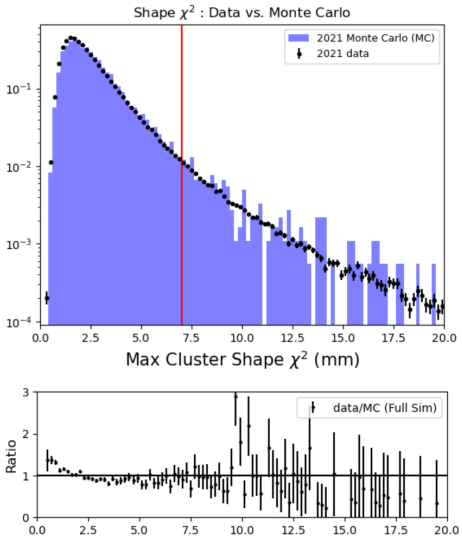


(d) The minimum reconstructed photon energy in the event. The x-axis is in MeV.

Figure 6.1: The comparison of $K_L \to 3\pi^0$ data and Monte Carlo. All kinematic cuts, except for the one shown in the histogram, are applied. Each histogram is normalized to one.



(e) The total deposit energy in the CSI. The x-axis is in MeV.



(f) The maximum cluster shape χ^2. This variable compares the cluster shape to a library of photon clusters to determine how "photon-like" a cluster is.

Figure 6.1: Continued

Additionally, veto and quality cuts are applied in Fig. 6.1. In general, inside the cut thresholds, data and Monte Carlo are consistent. In addition to the cuts above, three other kinematic cuts are applied in the analysis of $K_L \to \pi^0\pi^0\gamma\gamma$: the χ^2_{fit} with the $K_L \to 3\pi^0$ assumption, the χ^2_{fit} with the $K_L \to \pi^0\pi^0\gamma\gamma$ assumption, and the cluster shape discriminator using Deep Learning (CSDDL). These kinematic cuts are vital and are discussed in more detail in the following sections. For the purposes of this section, the selection used for the $K_L \to 3\pi^0$ data in these figures have $\chi^2_{fit} < 3$ to ensure these events are well reconstructed as $K_L \to 3\pi^0$.

To study the veto, a similar strategy is employed as in Fig. 6.1. The exclusive veto efficiency is compared between $K_L \to 3\pi^0$ data, $K_L \to 3\pi^0$ background Monte Carlo and $K_L \to \pi^0\pi^0\gamma\gamma$ Monte Carlo. The exclusive veto efficiency is defined in Eq. 6.1.

$$E_{veto,i} = \frac{N_{all}}{N_{all,i}} \tag{6.1}$$

$E_{veto,i}$ is defined as the veto efficiency of veto detector "i". N_{all} is the number of events after all veto and quality cuts, while $N_{all,i}$ is the number of events after all veto and quality cuts except the veto detector "i". The results of this study is shown in Fig. 6.2.

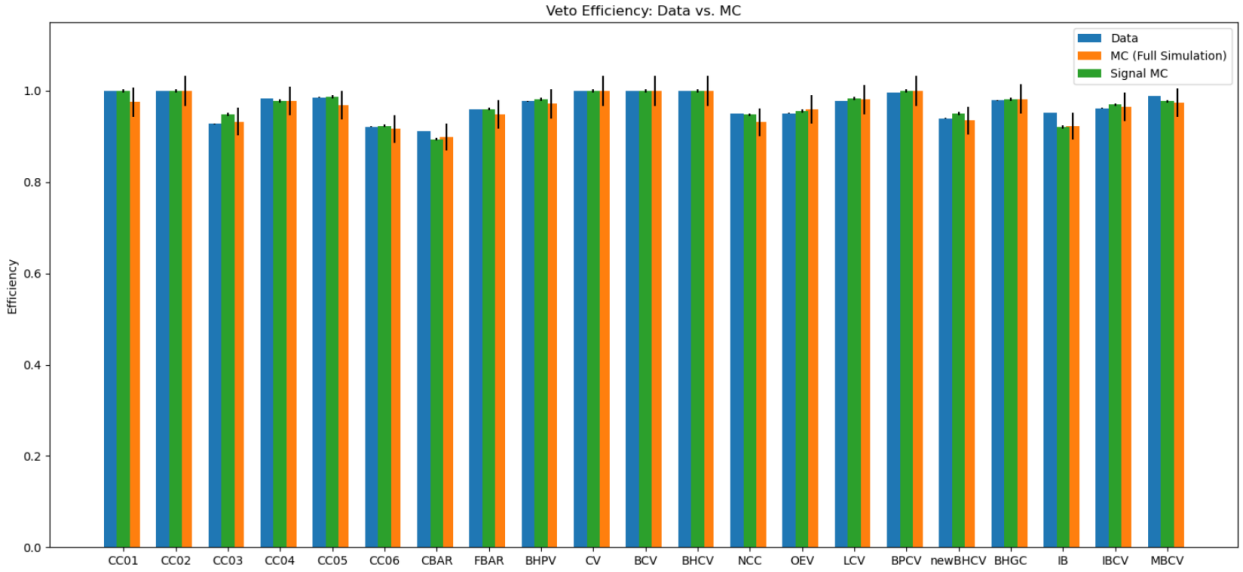


Figure 6.2: The comparison between exclusive veto efficiency in data (blue) and $K_L \to 3\pi^0$ (red) and for $K_L \to \pi^0\pi^0\gamma\gamma$ (green).

6.2 Messy CSI Channels

The analysis of $K_L \to \pi^0\pi^0\gamma\gamma$ is particularly sensitive to the energy resolution in the CSI calorimeter. This is due to the fact that $K_L \to 3\pi^0$ will start to resemble or signal if there is an energy mismeasurement that causes a photon mis-pairing to occur. If a mis-pairing occurs in the $K_L \to 3\pi^0$ decay, it can push the reconstructed decay vertex of the K_L downstream and push the invariant mass, $M_{\gamma_5\gamma_6}$, above the pion mass and into the signal region.

In the data taking runs in 2021, some problematic channels in the CSI calorimeter that were not well simulated in Monte Carlo were discovered. Photon hits near these channels were more likely to cause an energy mismeasurement, and events with hits near this CSI position were rejected for the analysis of $K_L \to \pi^0\pi^0\gamma\gamma$. This added CSI fiducial cut required

no hits in the region shown in Eq. 6.2.

$$\begin{aligned} 210 \text{ mm} < \text{HitX} < 260 \text{ mm} \\ 100 \text{ mm} < \text{HitY} < 135 \text{ mm} \end{aligned} \tag{6.2}$$

These messy channels were discovered by plotting the γ_5 and γ_6 reconstructed hit positions in data which had a large $K_L \to 3\pi^0$ constrained fit χ^2_{fit}. The plot of this analysis is shown in Fig. 6.3.

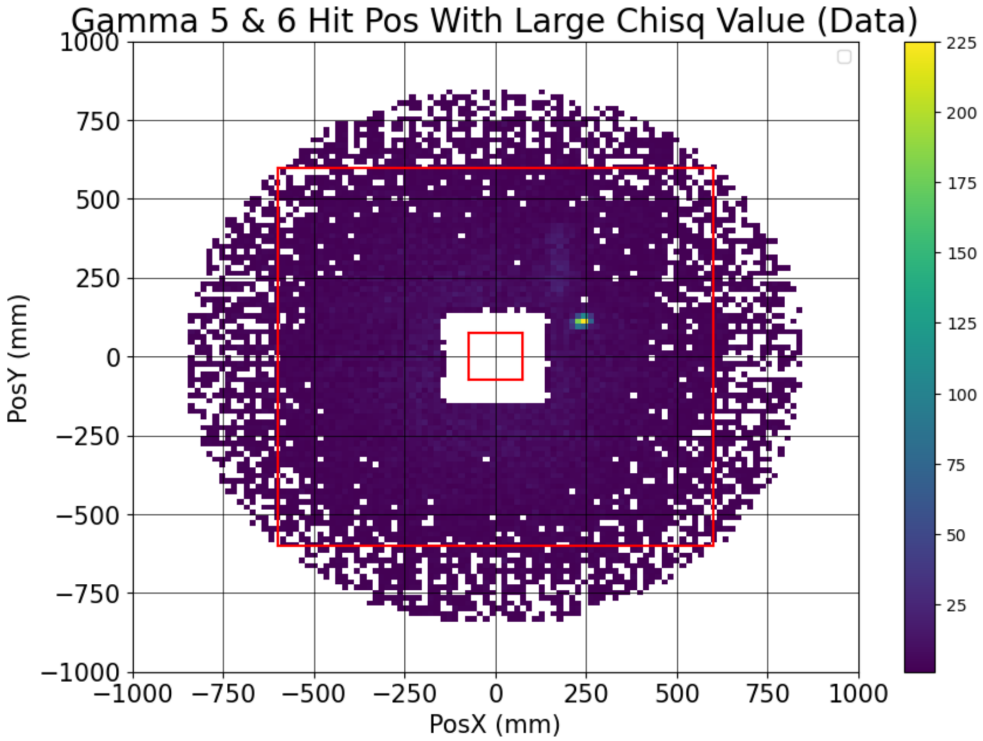


Figure 6.3: The hit position of γ_5 and γ_6 after $K_L \to \pi^0\pi^0\gamma\gamma$ reconstruction. The χ^2_{fit} with the $K_L \to \pi^0\pi^0\gamma\gamma$ assumption is small, while χ^2_{fit} with the $K_L \to 3\pi^0$ assumption is large.

The hot spot can be seen around the x-y position (235,120). Thus, these messy channels are disproportionally likely to cause a $K_L \to 3\pi^0$ background event, which motivates the cut in Eq. 6.2.

6.3 Systematic Error

Before evaluating the K_L flux in Section 6.4, the systematic error must be examined. This error stems from the discrepancies between data and Monte Carlo, and its impact on all calculations done using a simulated result. In reference to the K_L flux calculation, it requires the $K_L \to 3\pi^0$ acceptance estimated from Monte Carlo because this mode is used for the normalization.

6.3.1 Veto Uncertainties

The systematic errors associated with veto cuts rely entirely on the veto timing windows and energies being well simulated in Monte Carlo. The barrel counters are the most difficult to simulate well, and are the cause of the majority of systematic error, as can be seen in Fig. 6.2. The error is obtained by comparing the exclusive acceptance, A_i, for data and $K_L \to 3\pi^0$ Monte Carlo, where i represents the veto cut not applied. The formula for the exclusive acceptance is shown in Eq. 6.3.

$$A_i = \frac{\#\text{ of events after all cuts}}{\#\text{ of events after all cuts except the i-th cut}} \tag{6.3}$$

The systematic error for a particular veto cut is represented by Eq. 6.4.

$$\Delta_i = \frac{A_{i,mc} - A_{i,data}}{A_{i,data}} \tag{6.4}$$

Then in total, for each veto cut, the systematic error is then given by Eq. 6.5, where Δ_i is quadratically summed.

$$\Delta_{tot} = \sqrt{\Delta_i^2} = 5.5\% \tag{6.5}$$

6.3.2 Kinematic Uncertainties

A similar strategy is used for the kinematic cuts used in the K_L yield estimation. The exclusive cut values are compared between $K_L \to 3\pi^0$ data and $K_L \to 3\pi^0$ Monte Carlo in Fig. 6.4.

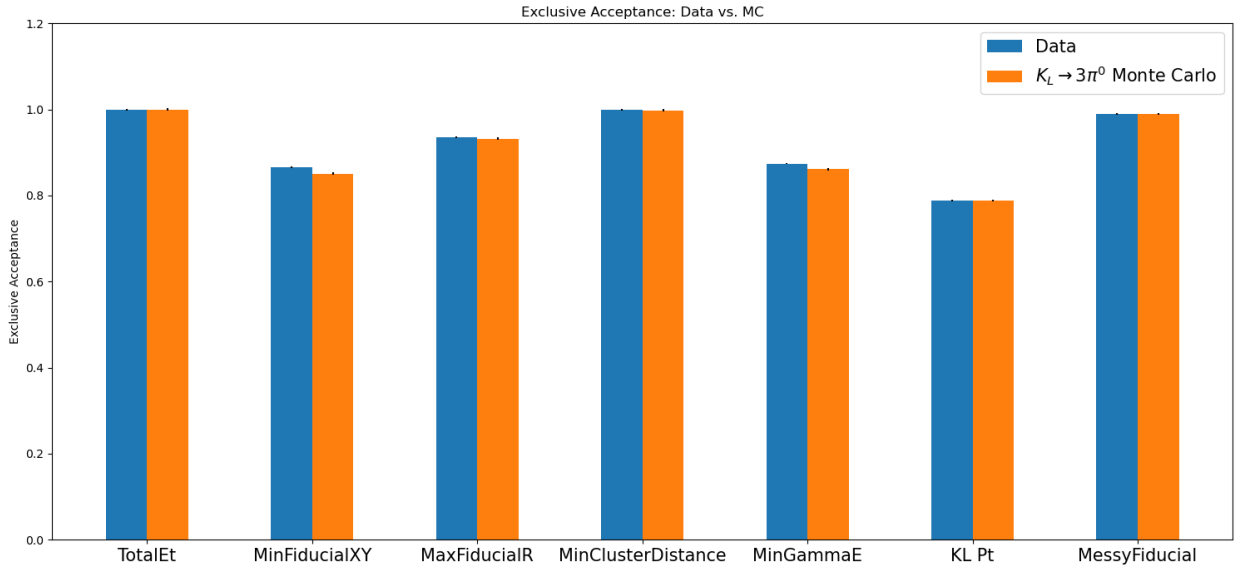


Figure 6.4: The exclusive acceptance of the kinematic cuts used in the K_L flux calculation.

following the strategy used in Eq. 6.4 and Eq. 6.5, the systematic error is given by,

$$\Delta_{tot} = \sqrt{\Delta_i^2} = 2.3\% \tag{6.6}$$

6.3.3 χ^2 Uncertainties

Following the same strategy as the previous sections, the systematic uncertainties related to the calculation of the flux using χ^2 related cuts are evaluated. These cuts are related to the cluster shape χ^2, and the two χ^2_{fit} variables obtained during reconstruction. The cluster shape χ^2 is a variable that compares the reconstructed cluster energy and timing distribution with a photon cluster library. Based on the energy and position of that cluster, this variable is compared against the library of clusters at that energy and position and a

χ^2 is constructed from the deviation from the expected result. This χ^2 is primarily used for $n-\gamma$ discrimination, or minimum ionization interactions, such as from a charged pion. The exclusive acceptance is shown in Fig. 6.5.

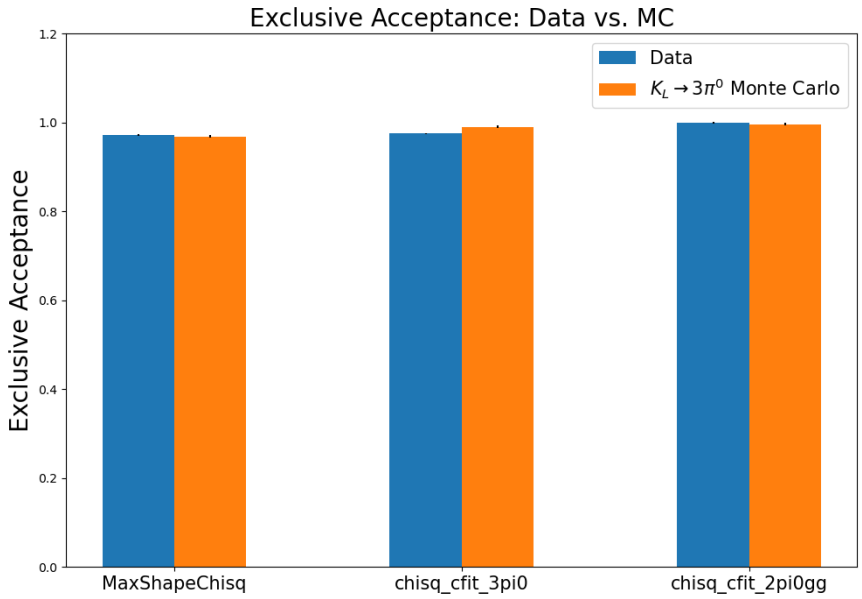


Figure 6.5: The exclusive acceptance of the χ^2 cuts used in the K_L flux calculation.

From this, the total systematic uncertainty is shown in Eq. 6.7.

$$\Delta_{tot} = \sqrt{\Delta_i^2} = 1.7\% \tag{6.7}$$

6.3.4 Geometrical Uncertainties

The geometrical acceptance of $K_L \to 3\pi^0$, or in other words, the probability of all final state particles hitting the calorimeter, is a by-product of the K_L momentum spectrum and detector configuration. If the K_L momentum spectrum in data differs from the model to simulate K_L's in Monte Carlo, then it will systematically affect the acceptance and flux calculation. To determine the size of this effect, the covariance matrix for the four parameters, (μ, a, b, c) is used from the K_L momentum measurement done in the engineering runs in 2012. This matrix is shown in Eq. 6.8.

$$V = \begin{pmatrix} 0.00144 & 0.00063 & 0.00133 & -0.00017 \\ 0.00063 & 0.00103 & 0.00127 & -0.00013 \\ 0.00133 & 0.00127 & 0.00216 & -0.00026 \\ -0.00017 & -0.00013 & -0.00026 & 0.000036 \end{pmatrix} \tag{6.8}$$

The off-diagonal elements of the covariance matrix represent the correlation between these parameters, and the variance in their values is simply given by the diagonal elements. Thus, by modifying the momentum parameters by random sampling and multiplying by the covariance matrix after Cholesky Decomposition [30], different momentum spectra are produced based on the uncertainties from the measurement.

To avoid high statistics Monte Carlo for each of these momentum spectra, a weighted acceptance method is used. Using the original K_L momentum spectrum, the geometrical acceptance is evaluated as a function of K_L momentum. An example of the K_L momentum spectrum before and after requiring all six photons incident in the CSI is shown in Fig. 6.6. This weight is then applied to the new K_L momentum function generated from the modified parameters.

The modified acceptance using the weighting method is shown in Eq. 6.9, where A_{mod} is the modified acceptance under the new parameters, $A_i(p)$ is the acceptance as a function of K_L momentum under the initial parameters, and $f_{mod}(p)$ is the modified K_L momentum distributions respectively.

$$A_{mod} = \frac{\int_0^\infty A_i(p) f_{mod}(p) dp}{\int_0^\infty f_{mod}(p) dp} \tag{6.9}$$

Doing this calculation for multiple new parameters gives many different A_{mod}'s, and plotting the relative difference for each new parameter,

$$\Delta = \frac{A_{mod} - A_i}{A_i} \tag{6.10}$$

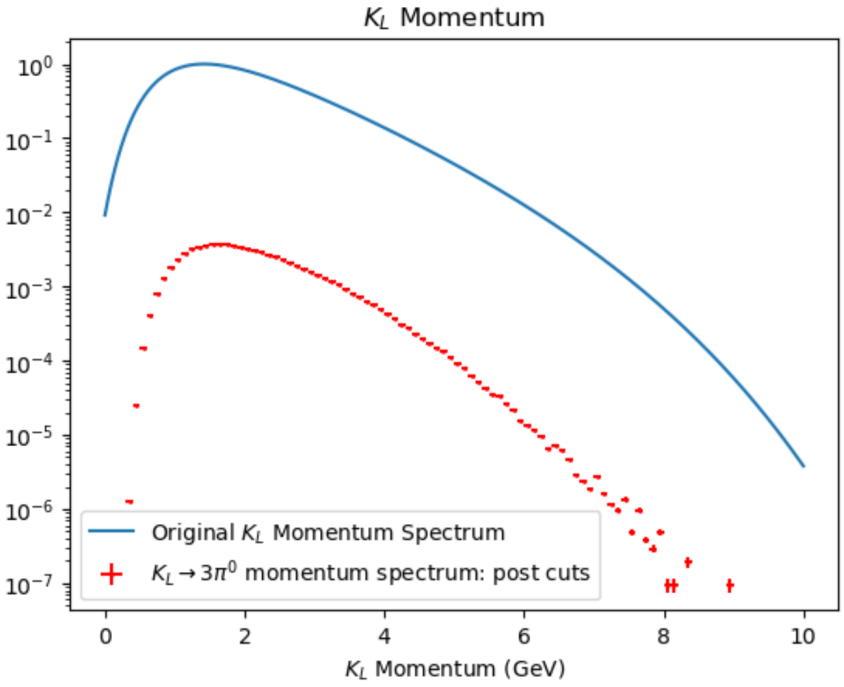


Figure 6.6: Comparison of the K_L momentum distribution before and after geometrical cuts requiring all six photons from the $K_L \to 3\pi^0$ decay incident on the CSI.

gives the result shown in Fig. 6.7. The systematic error is quoted at the 1σ boundary, or 1.2%.

6.3.5 Trigger Uncertainties

The trigger uncertainties are those associated with the offline cuts applied to the data under the assumption that they are not impacted by the online trigger. For instance, the online clustering requirement is six, but the algorithm is significantly different than that used offline to determine the cluster number with a more sophisticated strategy. If there is unexpected loss due to this trigger decision, it contributes to the systematic error. To study the online trigger effect, minimum bias trigger is used.

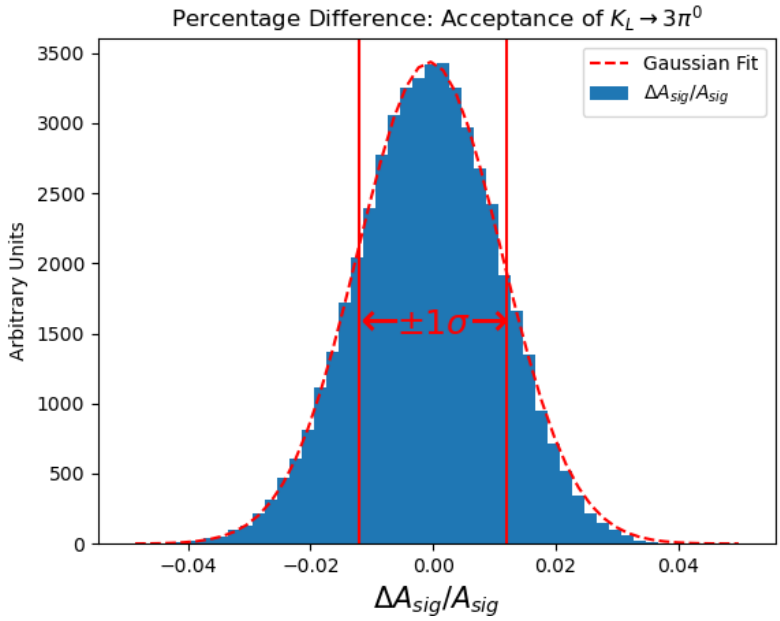


Figure 6.7: Distribution of Δ over 50,000 iterations of new parameters. The boundary shows the $\pm 1\sigma$ region and corresponds to the systematic error result.

Level 1 Trigger: E_t trigger

The Level 1 trigger E_t trigger requires the energy deposit in the CSI to be larger than 550 MeV. Offline, the cut threshold is 650 MeV. The online value is calculated using peak heights and calibration data using cosmic rays, whereas the offline data further improves the calibration using aluminum target runs, and calculates the pulse area. The uncertainty is obtained by a special data-taking run with a low online E_t threshold at 300 MeV. The efficiency is shown in Eq. 6.11.

$$Eff = \frac{\#\text{ of events passing online Et}}{\text{Total }\#\text{ of events}} \tag{6.11}$$

This efficiency is calculated for multiple ranges of the offline energy. It is then plotted as a function of the offline energy calculation and then fitted to the function shown in Eq. 6.12.

$$f(E_{off}) = a + \frac{b}{1 + \exp\left(\frac{(E_{off} - \mu)}{\sigma}\right)} \tag{6.12}$$

The fit result and efficiency are plotted in Fig. 6.8.

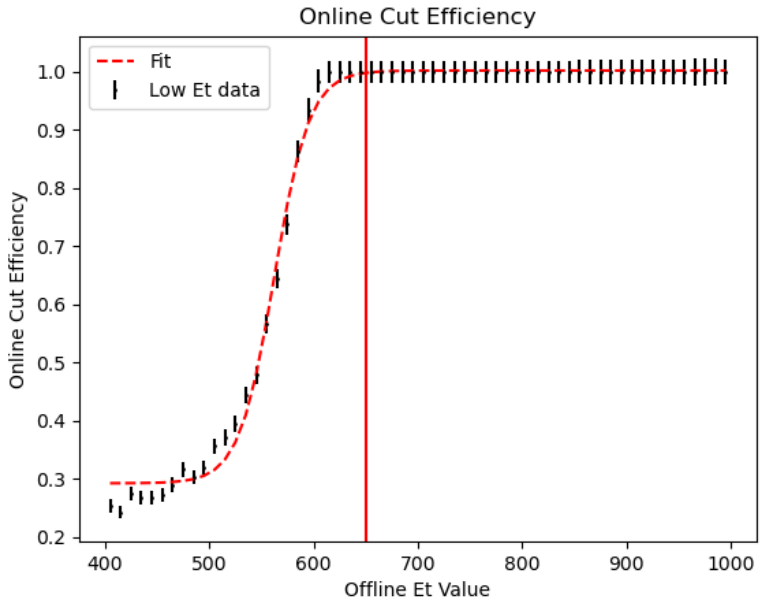


Figure 6.8: The efficiency as a function of offline energy calculation. The vertical red line corresponds to the online cut value of 650 MeV.

The fit results are shown in Eq. 6.13.

$$\begin{aligned} \mu &= 562.4 \pm 0.9 \\ \sigma &= 16.9 \pm 0.7 \end{aligned} \tag{6.13}$$

μ is more than 5σ away from the offline cut threshold. The loss expected from online E_t is therefore expected to be negligible.

Level 1 Trigger: Veto

The Level 1 trigger decision requires no hits in the veto detectors. Similar to the previous section, the online veto energies are evaluated using pulse heights and tight online timing windows. Offline, the pulse areas and wider windows are typically used with tighter energy requirements. To estimate the systematic loss due to the Level 1 veto, minimum bias trigger is used because it removes the veto requirements in physics trigger. The loss is determined with and without the online veto, shown in Eq. 6.14.

$$Rej = 1 - \frac{\#\text{ of events after online veto cut}}{\#\text{ of events with all cuts except online veto}} \tag{6.14}$$

This loss was found to be 1.0%.

Level 2 Trigger: Clustering

The systematic uncertainty of the Level 2 trigger is obtained by evaluating the difference in the online cluster count (CDTNum) calculation of $K_L \to 3\pi^0$ data and $K_L \to 3\pi^0$ Monte Carlo. The data is obtained through the Minimum Bias trigger, which does not apply the Level 2 trigger decision. From that data, the same $K_L \to 3\pi^0$ selection as that in the normalization is used, except for the online cluster count. The CDTNum distribution is shown in Fig. 6.9.

The systematic uncertainty is the difference in acceptance after requiring CDTNum==6. This uncertainty is shown in Eq. 6.15.

$$\Delta = \frac{A_{data} - A_{mc}}{A_{data}} = 4.0\% \tag{6.15}$$

Combining this result with the previous two sections by summing quadratically, the total systematic error associated with the online trigger decision is 4.1%, dominated by the Level 2 clustering calculation.

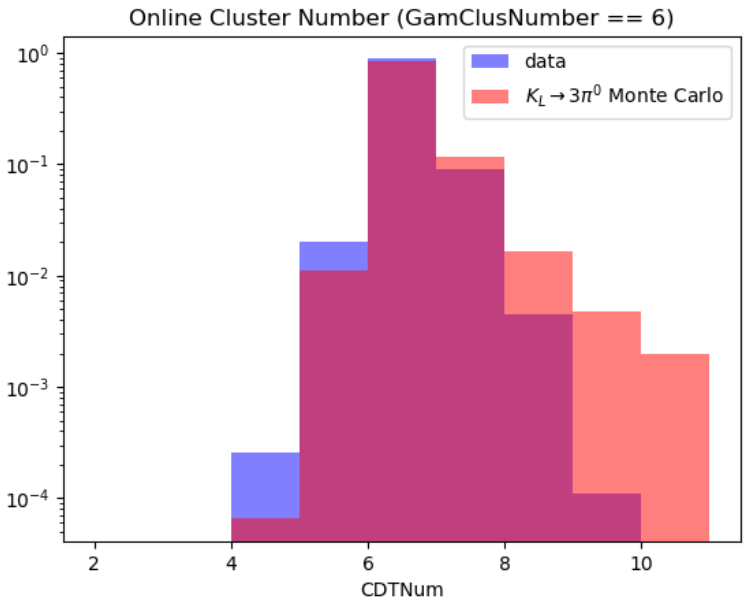


Figure 6.9: The CDTNum distribution for data and Monte Carlo after all cuts

Summary of Systematic Errors

The dominant uncertainties stem from the online cluster number calculation and from the offline veto cut discrepancy between data and Monte Carlo. The situation is summarized in Table 6.1.

Source	Uncertainty
Veto Cuts	5.1%
Kinematic Cuts	2.3%
χ^2 Cuts	1.7%
K_L Momentum	1.2%
Trigger	4.1%
Total	**7.2%**

Table 6.1: The summary of systematic errors associated with the K_L yield calculation. The total is the quadratic sum of each of the sources.

6.4 K_L Flux Calculation

Good agreement is observed between data and Monte Carlo for both background and signal. With this good agreement, the evaluation of the K_L yield and flux may be calculated using the Monte Carlo to determine the efficiency. The yield is calculated simply using $K_L \to 3\pi^0$ for the analysis of $K_L \to \pi^0\pi^0\gamma\gamma$. However, in the analysis of $K_L \to \pi^0\nu\overline{\nu}$, the yield is estimated from the normalization channel $K_L \to 2\pi^0$. In principle, these values should be the same. The yield is obtained via Eq. 6.16

$$Y = \frac{N_{3\pi^0}}{\mathrm{BR}(K_L \to 3\pi^0) \cdot A_{3\pi^0}} \tag{6.16}$$

Where $N_{3\pi^0}$ is the number of $K_L \to 3\pi^0$ observed in data, and $A_{3\pi^0}$ is the acceptance of the $K_L \to 3\pi^0$ mode determined in the Monte Carlo simulation. This is evaluated using Eq. 6.17.

$$A_{sig} = \frac{N_{sig}}{N_{gen}} \tag{6.17}$$

The acceptance, A_{sig}, is determined by the number of signal after all cuts, N_{sig}, divided by the number of signal Monte Carlo generated, N_{gen}. The generated Monte Carlo originates at the exit of the second collimator. This acceptance and yield calculation gives the K_L yield at the entrance of the detector. The yield itself is directly proportional to the K_L flux, where the conversion factor is the inverse number of protons on target (POT). Once the yield is obtained, the flux is given by Eq. 6.18.

$$F = \frac{Y}{POT} \tag{6.18}$$

In the analysis of $K_L \to \pi^0\pi^0\gamma\gamma$, the $N_{3\pi^0}$ after all quality, veto, and kinematic cuts was given by Eq. 6.19.

$$N_{data} = (2.201 \pm 0.001_{stat} \pm 0.159_{sys}) \times 10^7 \tag{6.19}$$

The acceptance is given in Eq. 6.20.

$$A_{3\pi^0} = (2.93 \pm 0.02_{stat} \pm 0.21_{sys}) \times 10^{-5} \tag{6.20}$$

This gives a yield using Eq. 6.16 as,

$$Y = (4.27 \pm 0.03_{stat} \pm 0.31_{sys}) \times 10^{12} \tag{6.21}$$

The complete set of cuts used in this calculation with explanations is shown in Table 6.2. The systematic errors are discussed later in this section. Note that this result differs from the yield from the published analysis of $K_L \to \pi^0\nu\overline{\nu}$ [4]. This difference results from the face that for the analysis of $K_L \to \pi^0\pi^0\gamma\gamma$, only a partial set of the data taking runs is analyzed. In particular, the set only contains runs with a beam intensity larger than 60 kW.

Cut Type	Cut Condition	Explanation
Veto	MyVetoCondition == 0	Default veto detector energy thresholds and timing windows (see Appendix A).
Kinematic	$E_{\gamma,min} > 100$ MeV	Minimum photon energy requirement.
Kinematic	$D_{min} > 100$ mm	Minimum distance between clusters must be large to prevent overlap.
Kinematic	$E_{tot} > 650$ MeV	Minimum total event energy requirement.
Kinematic	K_L $P_t < 8$ MeV	K_L must have low off-beam axis momentum.
Kinematic	$Min(X, Y) > 150$ mm	Requires all clusters to be outside 150 $\times$ 150 mm square centered at the beam hole.
Kinematic	$Max(R) < 850$ mm	Requires all clusters to have a hit radius far from the outer CSI edge.
Kinematic	210mm $<$ HitX $<$ 260mm and 100mm $<$ HitY $<$ 135mm	All photon hits must be outside the region affected by messy CSI channels.
χ^2	$\chi^2_{shape} < 7$	All clusters must be photon-like.
χ^2	$\chi^2_{fit}(3\pi^0) < 10$	Constrained Fit χ^2 must be small ($K_L \to 3\pi^0$) assumption.
χ^2	$\chi^2_{fit}(2\pi^0\gamma\gamma) < 10$	Constrained Fit χ^2 must be small ($K_L \to \pi^0\pi^0\gamma\gamma$) assumption.

Table 6.2: This table contains all the cut values and variables used in the K_L yield calculation for the $K_L \to \pi^0\pi^0\gamma\gamma$ analysis. In addition to those listed here are "quality" cuts that remove events that are not the correct trigger type, or if there was a DAQ error during the spill.

CHAPTER 7

ANALYSIS OF $K_L \to \pi^0\pi^0\gamma\gamma$ AND $K_L \to \pi^0\pi^0 X$

The primary focus of this dissertation is the novel analysis of $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\pi^0 X$. One major benefit of this analysis is the limited number of background sources. However, the primary background source $K_L \to 3\pi^0$ has a large branching ratio around 20%. Pushing the boundary of discovery for $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\pi^0 X$ thus requires strong rejection methods and detailed simulation strategies developed for these modes.

7.1 Simulation of $K_L \to \pi^0\pi^0\gamma\gamma$

The mode $K_L \to \pi^0\pi^0\gamma\gamma$ is a radiative K_L decay, which may generally be understood through the framework of ChPT, which is useful for "low energy" QCD interactions. In the case of $K_L \to \pi^0\pi^0\gamma\gamma$, the branching fraction and phase space can be "unambiguously" determined through ChPT [24]. Because the indirect CP-violating term, ε, is small, the amplitude of the decay $K_L \to \pi^0\pi^0\gamma\gamma$ can be roughly approximated as in Eq. 7.1.

$$A(K_L \to \pi^0\pi^0\gamma\gamma) = A(K_2 \to \pi^0\pi^0\gamma\gamma) + \varepsilon A(K_1 \to \pi^0\pi^0\gamma\gamma) \approx A(K_2 \to \pi^0\pi^0\gamma\gamma) \quad (7.1)$$

The leading order (p^4) contributions to the amplitude are shown in Eq. 7.2 and all parameters are explained below.

$$A(K_L \to \pi^0\pi^0\gamma\gamma) \approx A(K_2 \to \pi^0\pi^0\gamma\gamma) = i\frac{e^2 g}{3(4\pi^2)F^3} M_K^2 \epsilon_\mu(q_1)\epsilon_\nu(q_2) \times \left(\frac{3}{m_\pi^2 - 2z - i m_\pi \Gamma_\pi} + \frac{1}{M_\eta^2 - 2z}\right) \cdot T_3^{\mu\nu} \quad (7.2)$$

Where $z = q_1 \cdot q_2$, and q_1, q_2 are the radiative photons 4-momentum respectively.

$$F = 93.3 \text{ MeV}, \qquad g = 3.14 \times 10^{-7}$$
$$m_\pi = 134.9768 \text{ MeV/c}, \qquad M_K = 497.611 \text{ MeV/c}$$
$$M_\eta = 547.862 \text{ MeV/c}, \qquad T_3^{\mu\nu} = \varepsilon^{\mu\nu\rho\sigma} q_{1\rho} q_{2\sigma} \tag{7.3}$$

Clearly z is proportional to the invariant mass of the two radiative photons. Thus, if the invariant mass is plotted, there will be a pole at the pion mass, evidenced by the first parenthetical term in Eq. 7.2. The flat phase space assumption is therefore not realistic, and to understand the acceptance of $K_L \to \pi^0\pi^0\gamma\gamma$, the ChPT theory simulation using the amplitude shown here is necessary.

It is useful to consider this decay as a 3-body decay, with the two photons system behaving like a third particle with invariant mass proportional to z. The partial decay rate can broadly be thought of in two terms, shown in Eq. 7.4.

$$d\Gamma \propto |M|^2 \cdot dA \tag{7.4}$$

$|M|^2$ of course is the square of the amplitude shown in Eq. 7.2, while dA is the phase space area. Knowing that the phase space area is proportional to its corresponding Dalitz plot [19], the decay rate can be further simplified as,

$$d\Gamma \propto |M|^2 \cdot dA_{dalitz} \tag{7.5}$$

The square of the amplitude is shown in Eq. 7.6 after some algebraic simplification.

$$|M|^2 \propto z^2 \cdot \frac{(3\phi_\eta\phi_\pi + \phi_\pi^2 + m_\pi^2\Gamma_\pi^2)^2 + (m_\pi\Gamma_\pi)^2}{\phi_\eta^2(\phi_\pi^2 + m_\pi^2\Gamma_\pi^2)^2} \tag{7.6}$$

Here, $\phi_\eta = M_\eta^2 - 2z$ and $\phi_\pi = m_\pi^2 - 2z$. The area of a symmetric Dalitz plot can be

obtained via the integral [21] in Eq. 7.7 below,

$$\textbf{Area} \ \propto \int_{s_2}^{s_3} \frac{\sqrt{(s-s_1)(s-s_2)(s_3-s)(s_4-s)}}{s} \tag{7.7}$$

where,

$$\begin{aligned} s_1 &= (m_\pi - m_\pi)^2 = 0 \\ s_2 &= (m_\pi + m_\pi)^2 \\ s_3 &= (M_{K_L} - M_{\gamma\gamma})^2 \\ s_4 &= (M_{K_L} + M_{\gamma\gamma})^2 \end{aligned} \tag{7.8}$$

$$\tag{7.9}$$

The decay rate depends only on z, and by choosing the invariant mass of the two photons, the other two pions can be generated randomly assuming a typical three body decay. From the phase space area and the amplitude found in Eq. 7.6, the invariant mass distribution may be reproduced by first choosing an invariant mass uniformly in the range $(0, M_{K_L} - 2m_{\pi^0})$ and then using the Inverse Cumulative Distribution Function (ICDF) from the $d\Gamma$ calculation. The result is shown in Fig. 7.1.

This strategy was implemented into the KOTO simulation library, and the kinematics were compared in Fig. 7.2 using the KOTO simulation library and that of the Kambor-Funck analysis.

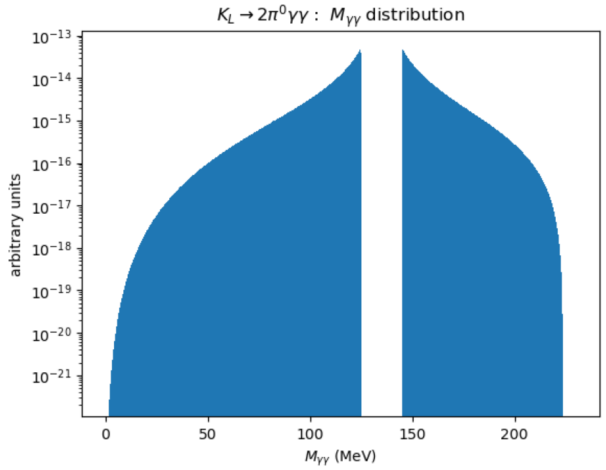
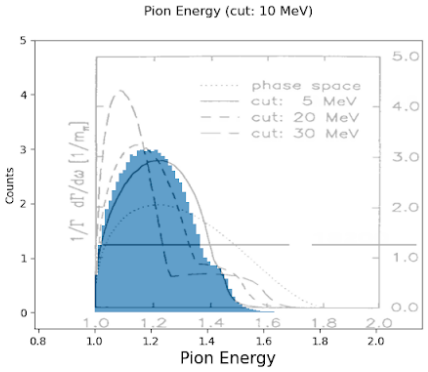
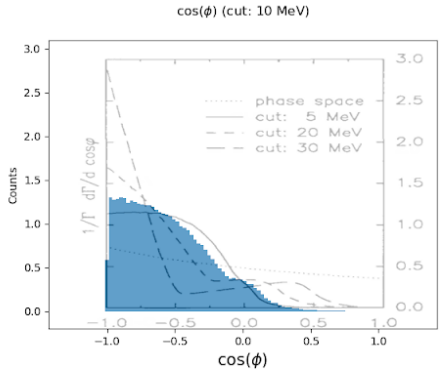


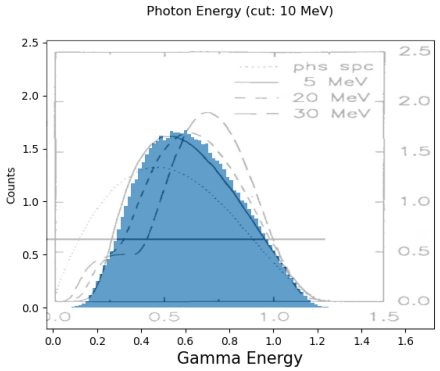
Figure 7.1: $M_{\gamma\gamma}$ Monte Carlo distribution. A cut requiring $|M_{\gamma\gamma} - m_{\pi^0}| < 10$ is applied to avoid the pole at the pion mass.



(a) The pion energy normalized to pion mass



(b) The opening angle between the two photons.



(c) The energy distributions of the photons, normalized to pion mass.

Figure 7.2: The KOTO simulation result overlaid on the Kambor-Funck distributions [24]. The cut value corresponds to the width of the $M_{\gamma\gamma}$ precut around the pion mass.

7.1.1 $K_L \to \pi^0\pi^0 X$ Simulation

The simulation of $K_L \to \pi^0\pi^0 X$ is done by implementing the decay directly in GEANT4 with defined X mass. The X particle is set to promptly decay into two photons with 100% decay probability and have a negligible lifetime. The K_L decay is treated as a uniform 3-body phase space decay. A range of different X masses is generated in 5 MeV intervals from 160-220 MeV. Masses below the pion mass were not studied in this analysis.

7.2 Measurement Principle

The measurement principle is simply six photon hits in the CSI that can be well reconstructed to form a K_L as in Section 5.4. The two pions in $K_L \to \pi^0\pi^0\gamma\gamma$ will promptly decay via, $\pi^0 \to 2\gamma$ 98.8% of the time [36], while for $K_L \to \pi^0\pi^0 X$, X is also assumed to promptly decay into two photons. This final state is remarkably similar to the $K_L \to 3\pi^0$ decay. Thus, the primary purpose of this analysis is to differentiate the two photons from X or the $K_L \to \pi^0\pi^0\gamma\gamma$ decay from a pion decay. An example diagram of $K_L \to \pi^0\pi^0 X$ signal interaction in the detector volume is shown in Fig. 7.3.

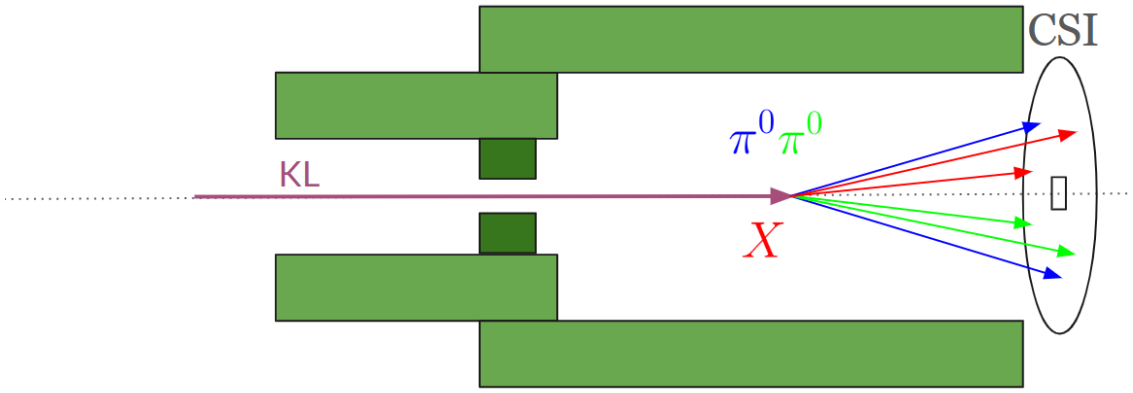


Figure 7.3: Green represents the barrel veto detectors (IB, CBAR, and FBAR) and the NCC. An example $K_L \to \pi^0\pi^0 X$ signal decay with all six final state photons incident in the CSI.

If a signal decay occurs in the detector region and all six photons are well measured,

the $\chi^2_{fit}(2\pi^0\gamma\gamma)$ should be small. However, if the $3\pi^0$ constraint is assumed, $\chi^2_{fit}(3\pi^0)$ should be large. This is due to the fact the X mass is assumed to be different than the π^0 mass. Similarly, the invariant mass $M_{\gamma_5\gamma_6}$ from the $K_L \to \pi^0\pi^0\gamma\gamma$ decay should also be different from the π^0 mass. The reconstructed result should thus not match the constraints well, leading to a large $\chi^2_{fit}(3\pi^0)$ value. To emphasize this difference, the $\chi^2_{fit}(3\pi^0)$ and $\chi^2_{fit}(2\pi^0\gamma\gamma)$ variables are plotted for $K_L \to 3\pi^0$ and $K_L \to \pi^0\pi^0X$ Monte Carlo in Fig. 7.4.

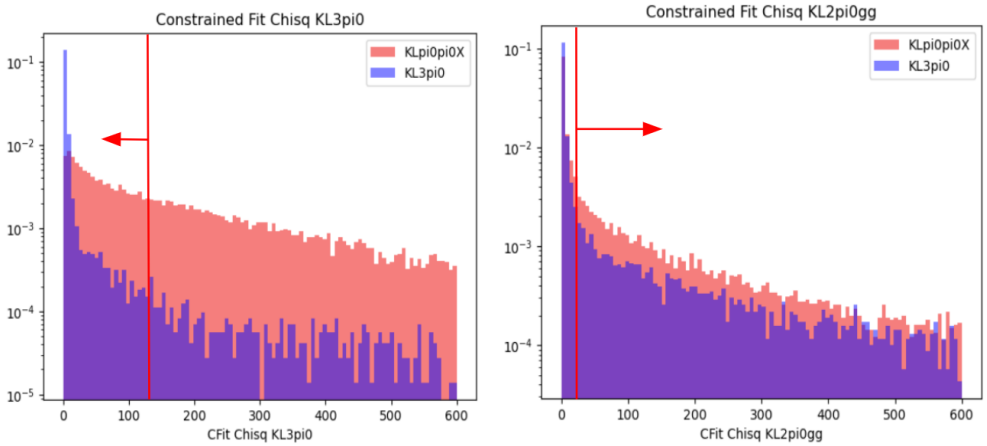


Figure 7.4: Overlapping histograms showing $\chi^2_{fit}(3\pi^0)$ and $\chi^2_{fit}(2\pi^0\gamma\gamma)$ results of $K_L \to 3\pi^0$ and $K_L \to \pi^0\pi^0X$ Monte Carlo. The vertical red line indicates an example cut of $\chi^2_{fit}(3\pi^0) > 115$ and $\chi^2_{fit}(2\pi^0\gamma\gamma) < 10$. An X mass of 190 MeV is assumed in this Monte Carlo.

The difference between the signal and background in Fig. 7.4 is profound, and provides the basis for background discrimination in this analysis.

7.3 $K_L \to 3\pi^0$ Background

The branching ratio of $K_L \to 3\pi^0$ is $\approx 20\%$. Being the primary background in a rare decay search like $K_L \to \pi^0\pi^0\gamma\gamma$ or $K_L \to \pi^0\pi^0X$ is extremely challenging and requires precise knowledge on the performance of the calorimeter. As mentioned in Sec. 5.4, there are 45 different pairing combinations for the six photons coming from three pions, with only one of those combinations being the correct one that leads to the true vertex reconstruction. Due

to the energy resolution of the CSI, it is possible that if the energy is slightly mismeasured, a wrong combination may lead to a "better" reconstruction result. This probability increases the more a photon is mismeasured in the calorimeter, and could result in a much different reconstruction result. An example of the $K_L \to 3\pi^0$ background tail is shown in Fig. 7.5.

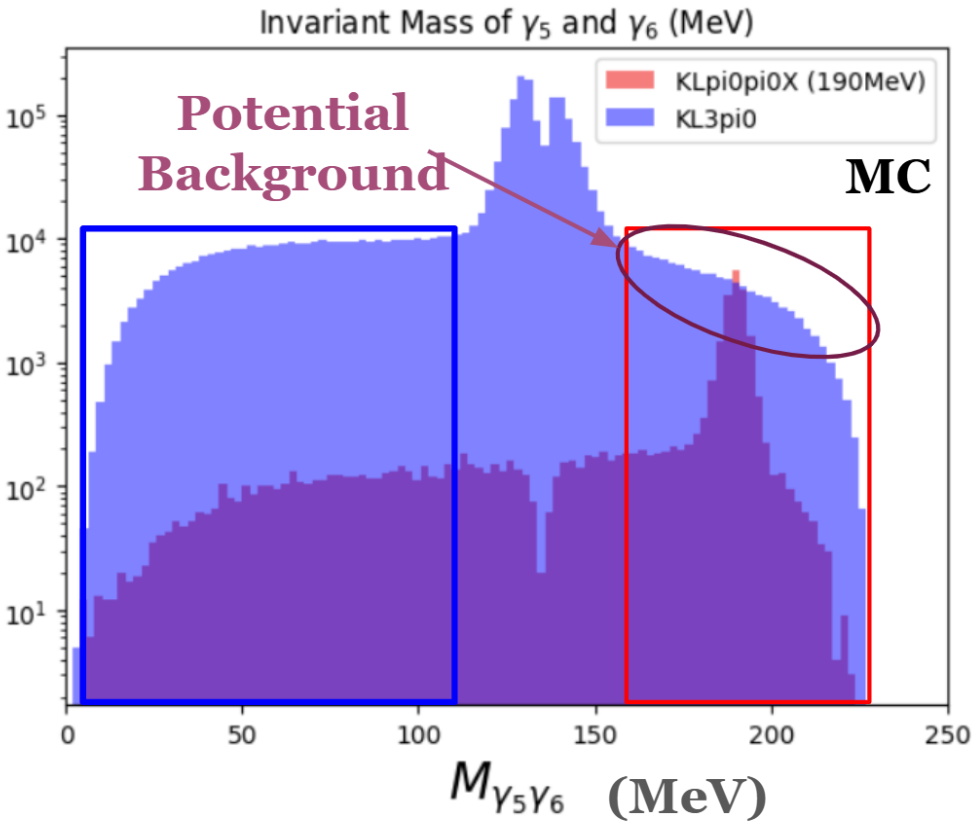


Figure 7.5: Background distribution with an example of the signal $K_L \to \pi^0\pi^0 X$ Monte Carlo with X mass of 190 MeV overlaid underneath. The blue box represents the low mass region which is treated as a control region, and the red box is the signal region.

Notably, the high mass tail is much smaller than the low mass tail. And even so, due to the large branching ratio of the $K_L \to 3\pi^0$ decay, the high mass tail may completely cover any signal unless background rejection methods are implemented. There are a few strange features in Fig. 7.5. The most prominent being the double peak around the pion mass. This double peak occurs because of the reconstruction strategy outlined in Sec. 5.4. The constraints pick out the four photons which best match the pion mass constraints and leave

the last two photons untouched. Thus, by definition, the invariant mass of the last pair of photons must deviate from the pion mass more than the other two pairs, leaving an inverse peak at the pion mass.

In another interesting observation, the low mass tail is larger than the high mass tail. This revolves around the fact that energy mismeasurement is strongly biased towards missing energy. There are many mechanisms for missing energy in the reconstruction, such as energy leakage and backsplash, or via the photonuclear effect. If energy loss occurs for one photon, the two other pairs will reconstruct the pion mass well and give the correct K_L vertex. However, the last pair will have some energy missing and a smaller invariant mass leading to a large low mass tail. In order to make it into the high mass tail, either energy needs to be added into the system, such as an accidental hit that is on-time with the other clusters, or an energy mismeasurement that leads to a mispairing. The mispairing causes a vertex misreconstruction that could push the invariant mass of the last pair either up or down.

7.4 Event Selection Strategy

There are many different types of signals in this analysis, $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\pi^0 X$, where the X mass ranges from 160-220 MeV. In principle, each analysis may require a separate event selection process with different cut thresholds. As an example, the $K_L \to 3\pi^0$ background level is lower near the kinematic boundary of $M_{\gamma\gamma} \approx 220$ MeV, and so looser cuts may still keep the background at acceptable levels.

The general kinematic and veto cuts and their basic descriptions and purpose, which are applied to each of these analyses are described in Table 6.2. These cuts are applied to ensure quality data. The cut conditions are the same as those listed in Table 6.2 except for the χ^2_{fit} variables listed in the last two rows which have dynamic cut thresholds that depend on the analysis. In addition, the following cuts are applied and their purposes are explained.

- MinWFourierLH < 0.19 : – This variable was developed for the discrimination of neu-

tron background in the $K_L \to \pi^0\nu\overline{\nu}$ analysis. It is a likelihood method that compares the pulse shapes in a shower to a library of photon and neutron pulse shapes taken from data. Neutron interactions in the CSI lead to longer pulse tails, and this difference can be seen in the low frequency amplitudes by taking the Fourier Transform. This variable is powerful in this analysis due to the photonuclear effect, which is the dominate background source, impacting the pulse shape.

The signal acceptance of this cut is done using $K_L \to 3\pi^0$ data which are well reconstructed. Monte Carlo is not used, as this method relies heavily on the pulse shape which is not perfectly simulated for hadronic interactions. The plot shown in Fig. 7.6, has the cut value that leads to 90% acceptance of well reconstructed events.

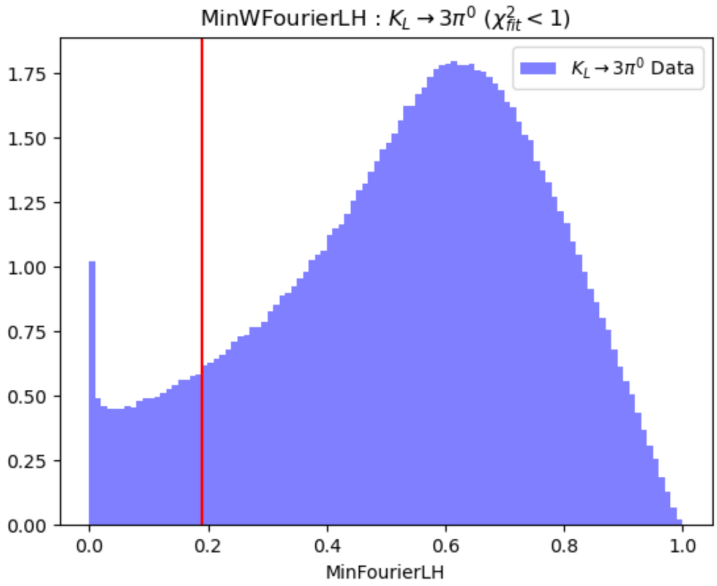


Figure 7.6: The MinWFourierLH variable for well reconstructed $K_L \to 3\pi^0$ events. An event is well reconstructed if $\chi^2_{fit}(3\pi^0) < 1$. The vertical red line represents the cut value which attains 90% acceptance.

While 90% signal acceptance is expected, the background rejection in the high mass region is large though this cut was not created specifically for this analysis. The performance of this cut is evaluated using data in the high mass region under loose cut conditions. This is

primarily misreconstructed $K_L \to 3\pi^0$ data, and 40% rejection is expected, as shown in Fig. 7.7.

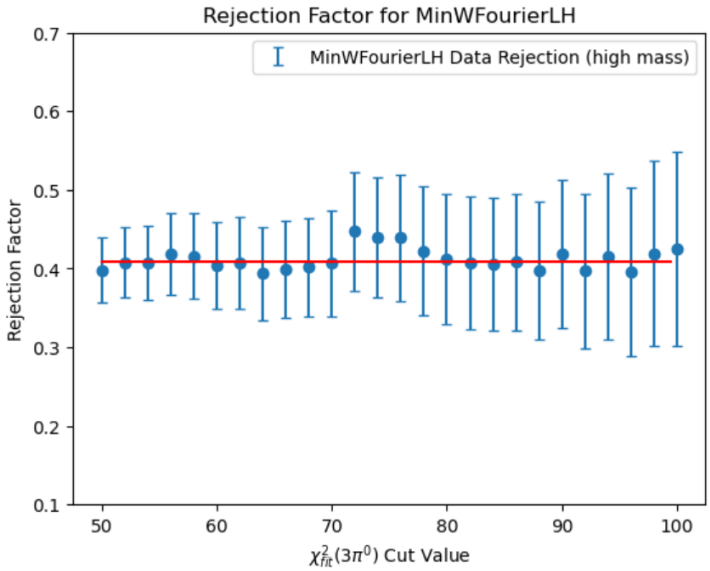


Figure 7.7: The MinWFourierLH rejection factor in the high mass region requiring a large $\chi^2_{fit}(3\pi^9)$ value. The error bars represent the statistical uncertainty in the rejection factor.

- MinCSDDLVal (dynamic) : – The MinCSDDLVal variable is obtained using a deep learning convolutional neural network. This network was trained to discriminate the cluster shape of hadronic and electromagnetic showers. This variable is important to the analysis, as the photonuclear effect also impacts the cluster shape.

MinCSDDLVal is treated as a dynamic cut, as it is such a powerful rejection tool in this analysis, but tight cut values lead to lower acceptance. Thus, analyses which require tighter cut thresholds to keep background levels low may prefer a tighter MinCSDDLVal cut value. To visualize the power of this cut, the difference between $K_L \to \pi^0\pi^0 X$ signal and $K_L \to 3\pi^0$ background is compared in Fig. 7.8.

However, understanding the true performance requires a comparison between data and Monte Carlo. This comparison is shown in Fig. 7.9 for different $\chi^2_{fit}(3\pi^0)$ cut values.

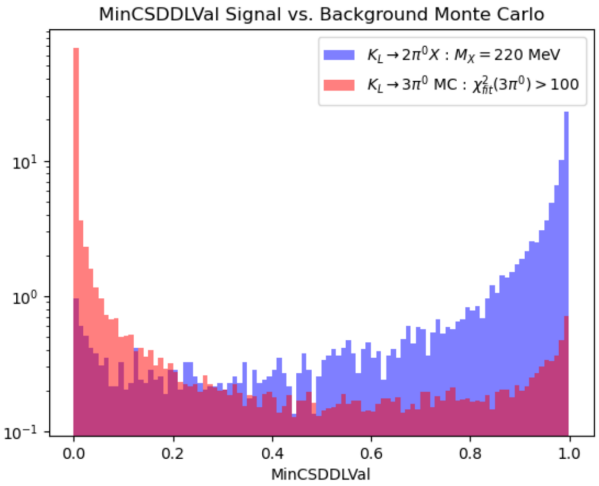
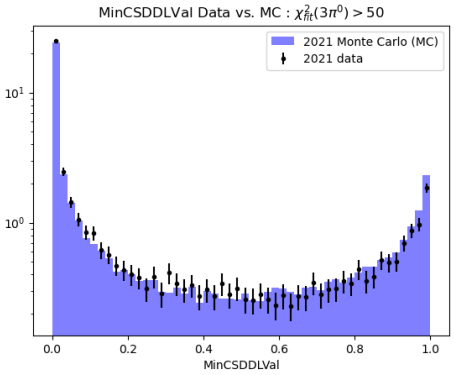
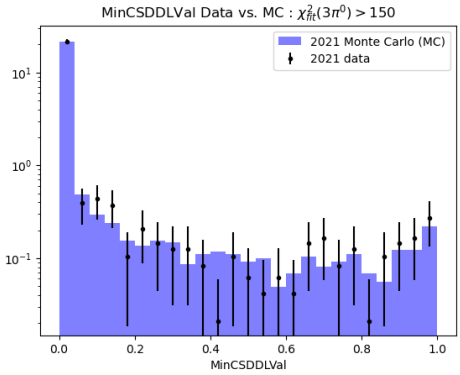


Figure 7.8: The MinCSDDLVal variable for background and signal Monte Carlo. The background is $K_L \to 3\pi^0$ events with $\chi^2_{fit}(3\pi^0) > 100$, and the signal is $K_L \to \pi^0\pi^0 X$ Monte Carlo with an X mass of 220 MeV.



(a) MinCSDDLVal with with loose $\chi^2_{fit}(3\pi^0) > 50$ cut.



(b) MinCSDDLVal with with loose $\chi^2_{fit}(3\pi^0) > 150$ cut.

Figure 7.9: The comparison between data and Monte Carlo under loose and tight cut conditions.

- $\underline{\Delta T_{vertex} < 1 \text{ ns}:}$ – This is a timing quality cut, requiring the first and second best reconstructions have similar K_L decay times from time of flight calculations. In the analysis of $K_L \to \pi^0\nu\overline{\nu}$, this cut rejects neutron events that may have large timing differences. For this analysis, it may be large if there is a misreconstruction or accidental hit and the K_L vertex position is off. Applying this cut can remove problematic events.

- $\underline{\text{AverageClusterTime} < 15 \text{ ns}:}$ – This is another timing quality cut, which requires the average reconstructed cluster times to fall within a tighter timing window. This cut suppresses events with clusters of varying times, such as those with accidental hits.

7.5 Cut Optimization Strategy

To optimize the cut selection, a $K_L \to 3\pi^0$ Monte Carlo dataset that corresponds to approximately five times the K_L from 2021 data taking. This dataset is used for the selection of cut values for the 14 different analyses, including $K_L \to \pi^0\pi^0\gamma\gamma$ and the thirteen X masses ranging from 160-220 MeV for $K_L \to \pi^0\pi^0 X$. The cut optimization focused on the three most important variables in these analyses, $\chi^2_{fit}(3\pi^0)$, $\chi^2_{fit}(2\pi^0\gamma\gamma)$, and MinCSDDLVal.

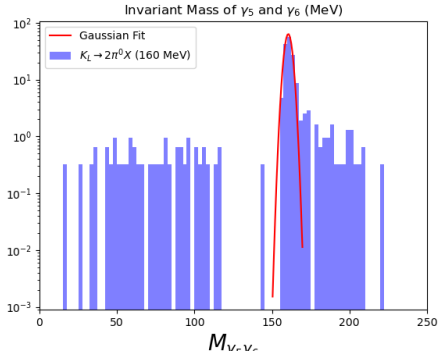
A three dimensional grid search was done to test each set of cuts and evaluate them based on a Figure of Merit (FOM) strategy. A FOM strategy optimizes the potential upper limit result by maximizing a function $f(N_{sig}, N_{bkg})$. Three different FOM were attempted, shown in Eq. 7.10. The same result was obtained using each of the methods.

N_{sig}, and N_{bkg} were calculated within the region of interest (ROI) for each analysis. For the X mass analyses, the ROI is found by applying a gaussian fit to the signal Monte Carlo, and defined by $\mu_{fit} \pm 2\sigma_{git}$. These fits are shown in Fig. 7.10. The signal is shown after a loose $\chi^2_{fit}(3\pi^0) > 100$ cut.

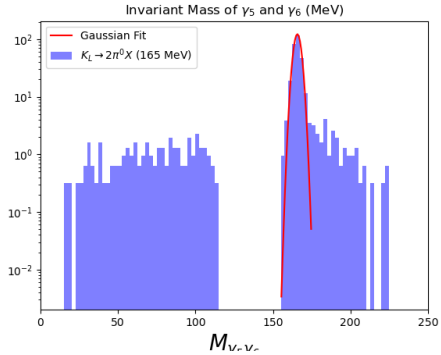
$$f_1 N_{sig}, N_{bkg}) = \sqrt{N_{sig} + N_{bkg}} - \sqrt{N_{bkg}}$$

$$f_2(N_{sig}, N_{bkg}) = \frac{N_{sig}}{\sqrt{N_{sig} + N_{bkg}}}$$

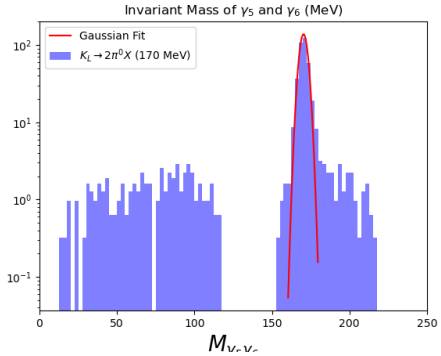
$$f_3(N_{sig}, N_{bkg}) = \sqrt{2\ln(Q)}\,, \qquad Q = e^{-N_{sig}} \left(1 + \frac{N_{sig}}{N_{bkg}}\right)^{N_{sig}+N_{bkg}} \tag{7.10}$$



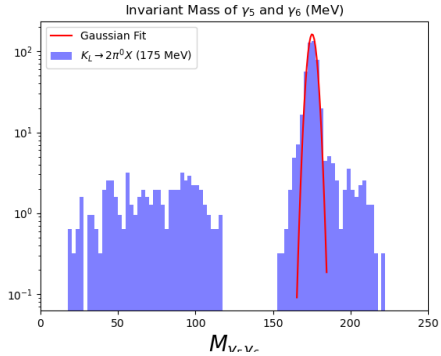
(a) Gaussian fit for X mass of 160 MeV.



(b) Gaussian fit for X mass of 165 MeV.

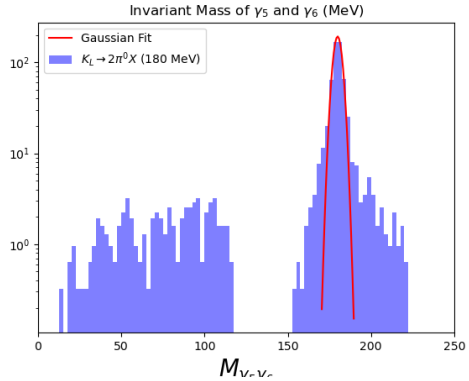


(c) Gaussian fit for X mass of 170 MeV.

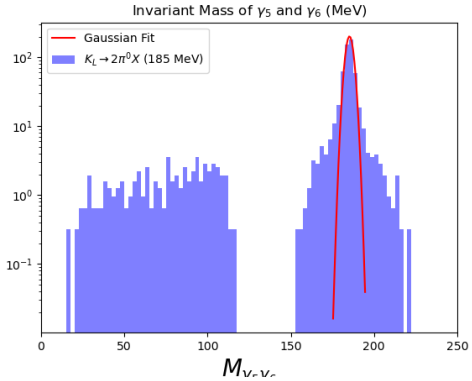


(d) Gaussian fit for X mass of 175 MeV.

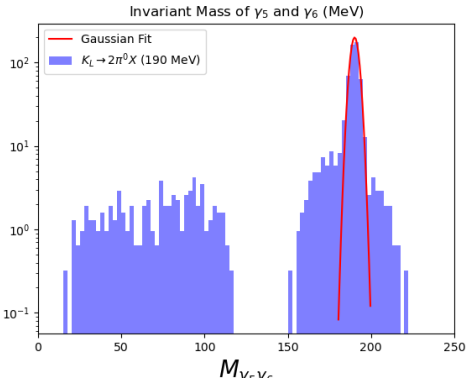
Figure 7.10: The Gaussian fits which determine the ROI for these analyses.



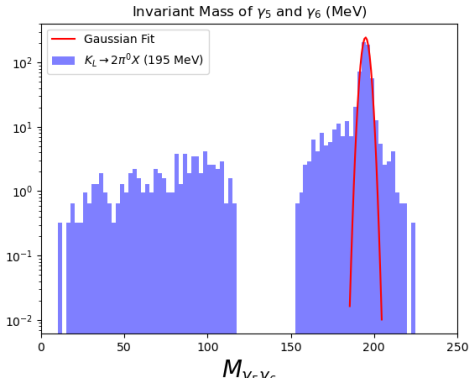
(e) Gaussian fit for X mass of 180 MeV.



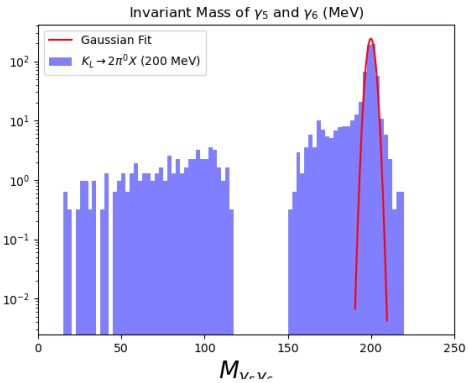
(f) Gaussian fit for X mass of 185 MeV.



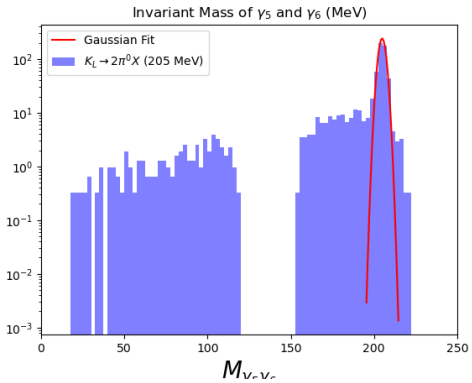
(g) Gaussian fit for X mass of 190 MeV.



(h) Gaussian fit for X mass of 195 MeV.

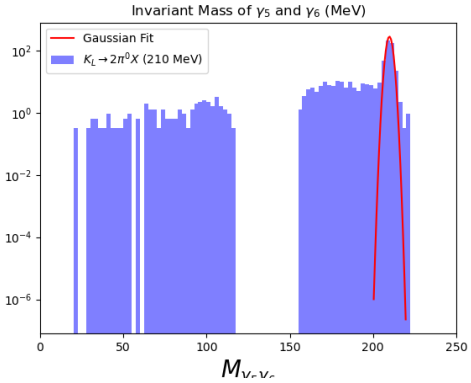


(i) Gaussian fit for X mass of 200 MeV.

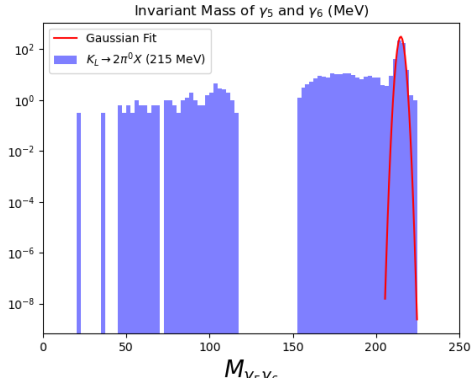


(j) Gaussian fit for X mass of 205 MeV.

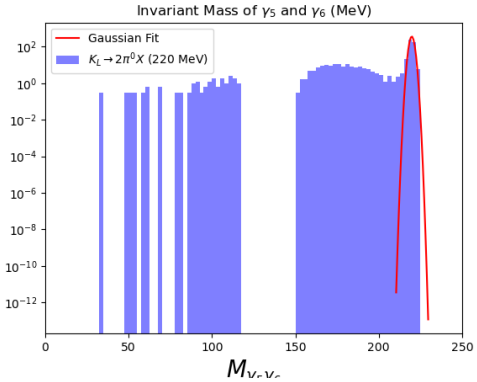
Figure 7.10: Continued



(k) Gaussian fit for X mass of 210 MeV.



(l) Gaussian fit for X mass of 215 MeV.



(m) Gaussian fit for X mass of 220 MeV.

Figure 7.10: Continued

Notably, the peak width tends to decrease at higher X mass. The gap around the pion mass with no events is caused by the cut requiring $\chi^2_{fit}(3\pi^0) > 100$. Table 7.1 shows the fitted parameters and signal regions. To determine the number of signal observed for a given analysis, $M_{\gamma\gamma}$ must fall within the ROI. If the number of background in the ROI was zero for a given set of cuts, the signal acceptance was taken as the FOM to optimize. In all cases, after this FOM strategy, there were no Monte Carlo events observed in the ROIs. The optimal cut values are shown in Table 7.1 with the fit parameters.

Analysis	μ_{fit}	σ_{fit}	Signal Region (MeV)	$\chi^2_{fit}(3\pi^0)$	$\chi^2_{fit}(2\pi^0)$	MinCSDDLVal
$M_X = 160$	160.48	2.15	$M_{\gamma\gamma} \in (156.18, 164.78)$	178	9	0.6
$M_X = 165$	165.66	2.23	$M_{\gamma\gamma} \in (161.20, 170.12)$	146	5	0.7
$M_X = 170$	170.34	2.46	$M_{\gamma\gamma} \in (165.42, 175.26)$	146	5	0.7
$M_X = 175$	175.26	2.52	$M_{\gamma\gamma} \in (170.22, 180.30)$	143	3	0.68
$M_X = 180$	179.96	2.54	$M_{\gamma\gamma} \in (174.88, 185.04)$	137	3	0.82
$M_X = 185$	185.21	2.30	$M_{\gamma\gamma} \in (180.61, 189.21)$	160	6	0.82
$M_X = 190$	190.13	2.46	$M_{\gamma\gamma} \in (185.21, 195.05)$	188	9	0.8
$M_X = 195$	194.88	2.19	$M_{\gamma\gamma} \in (190.50, 199.26)$	159	3	0.6
$M_X = 200$	199.92	2.05	$M_{\gamma\gamma} \in (195.82, 204.02)$	159	5	0.6
$M_X = 205$	204.85	1.92	$M_{\gamma\gamma} \in (200.99, 208.71)$	144	6	0.6
$M_X = 210$	209.86	1.51	$M_{\gamma\gamma} \in (206.84, 212.88)$	115	6	0.69
$M_X = 215$	214.84	1.36	$M_{\gamma\gamma} \in (212.12, 217.56)$	127	9	0.6
$M_X = 220$	219.77	1.17	$M_{\gamma\gamma} \in (217.43, 222.11)$	166	9	0.6
$K_L \to \pi^0\pi^0\gamma\gamma$	-	-	$M_{\gamma\gamma} \in (160, M_{K_L} - 2M_{\pi^0})$	178	3	0.78

Table 7.1: This table contains all the fit parameters, the signal region (ROI) for each of the analyses, and the optimized cut values. All units are in MeV in the first four columns. The last column is the cut values obtained for the MinCSDDLVal cut.

7.6 Control Region Study

The low mass region, below the pion mass, is crucial for understanding the agreement between data and Monte Carlo. There may be many causes that affect the performance of simulation. The photonuclear effect is not well studied in different materials for example, and a low energy model is used to estimate the cross sections in our detector [50] that rely on data measuring the cross section for other materials, and not CsI. If the photonuclear cross section in the CsI tends to be larger than expected in KOTO's energy range, Monte Carlo may underestimate the background levels.

Thus, a background level scale factor may be determined by studying the low mass region after tight cut selections. This factor is defined in Eq. 7.11, as

$$F = \frac{\#\text{ observed in data}}{\#\text{ left in MC}} \tag{7.11}$$

An example of the discrepancy observed in the low mass region can be seen in Fig. 7.11. In this figure, the comparison is done after applying all cuts except for MinWFourierLH. The cut values used for $\chi^2_{fit}(3\pi^0)$, $\chi^2_{fit}(2\pi^0\gamma\gamma)$, and MinCSDDLVal were 100, 5, and 0.1 respectively.

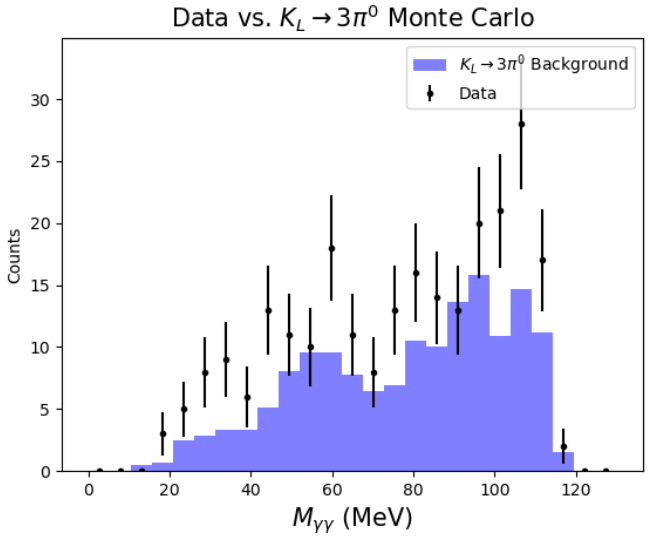
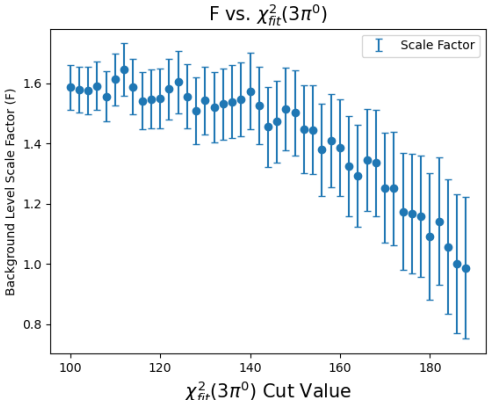


Figure 7.11: A example comparison between Data and $K_L \to 3\pi^0$ Monte Carlo in the low mass region under loose cut conditions.

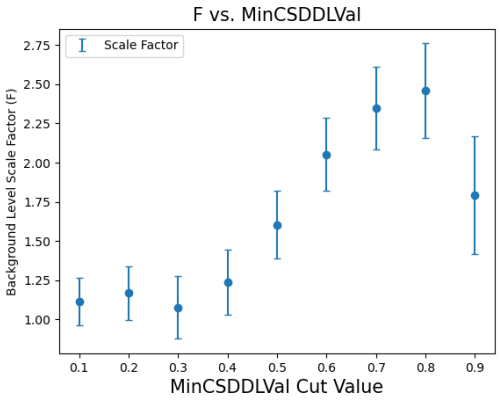
The difference between data and Monte Carlo is significant. Thus, it is crucial to study the correlation between the scale factor shown in Eq. 7.11 and $\chi^2_{fit}(3\pi^0)$ and MinCSDDLVal, the two most powerful cuts used in the analysis.

The correlation to the scale factor was observed and examples are shown in Fig. 7.12. All cuts are applied except for MinWFourierLH.

Since there is a strong correlation between these cuts and the scale factor, the background expectation based on Monte Carlo in each of the signal ROIs is scaled by this factor, which



(a) Correlation between the scale factor and $\chi^2_{fit}(3\pi^0)$.



(b) Correlation between the scale factor and MinCSDDLVal

Figure 7.12: These plots show the correlation between the background level scale factor calculation and the two most important cuts in the analyis, $\chi^2_{fit}(3\pi^0)$ and MinCSDDLVal. The error bars represent the statistical uncertainty in the ratio.

depends on the set of cuts applied. This factor is calculated for each set of cuts, and shown in Fig. 7.13. In general, if the analysis has a tight cut value for MinCSDDLVal, then the background level scale factor is larger.

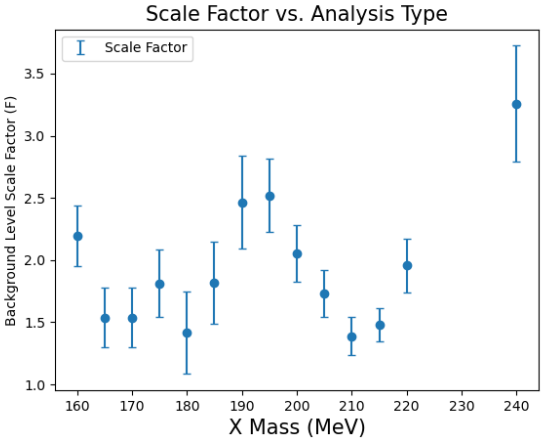


Figure 7.13: The background level scale factors for each analyses. The point labeled at an X mass of 240 MeV represents the scale factor for the $K_L \to \pi^0\pi^0\gamma\gamma$ analysis.

The results are further shown in the Table 7.2,

Analysis	$F \pm \Delta F$
$M_X = 160$	2.2 ± 0.2
$M_X = 165$	1.5 ± 0.2
$M_X = 170$	1.5 ± 0.2
$M_X = 175$	1.8 ± 0.3
$M_X = 180$	1.4 ± 0.3
$M_X = 185$	1.8 ± 0.3
$M_X = 190$	2.5 ± 0.4
$M_X = 195$	2.5 ± 0.3
$M_X = 200$	2.0 ± 0.2
$M_X = 205$	1.7 ± 0.2
$M_X = 210$	1.39 ± 0.15
$M_X = 215$	1.48 ± 0.13
$M_X = 220$	2.0 ± 0.2
$K_L \to \pi^0\pi^0\gamma\gamma$	3.2 ± 0.5

Table 7.2: This table contains all the the background level scale factor for all the analyses. ΔF is determined by the statistical error in the data to $K_L \to 3\pi^0$ Monte Carlo ratio.

7.7 Single Event Sensitivity

The Single Event Sensitivity (SES) is determined by the number of K_L and the acceptance of the signal and represents the central value of the branching ratio if one signal event is observed. The formula is shown in Eq. 7.12.

$$\text{SES} = \frac{1}{A_{sig} \cdot N_{K_L}} \tag{7.12}$$

The calculation of N_{K_L} is discussed in Section 6.4. It requires understanding the statistical and systematic uncertainties of the acceptance of $K_L \to 3\pi^0$ under a specific cut condition. In the analysis of $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\pi^0 X$, the cut conditions are slightly different for the following cuts: MinWFourierLH, MinCSDDLVal, $\chi^2_{fit}(3\pi^0)$, and $\chi^2_{fit}(2\pi^0)$. Thus, the systematic uncertainties related to these cut conditions must be understood.

7.7.1 Systematic Uncertainties: A_{sig}

This section follows the strategies outlined before in Section 6.3. The main difference arises from the study of MinWFourierLH variable. This variable relies on the pulse shape, which is not well simulated in Monte Carlo for hadronic interactions. Therefore, the acceptance weighted method is applied, similar to the geometrical acceptance systematic uncertainty calculation. The efficiency of the MinWFourierLH cut condition is evaluated using $K_L \to 3\pi^0$ data as a function of photon energy, shown in Fig. 7.14.

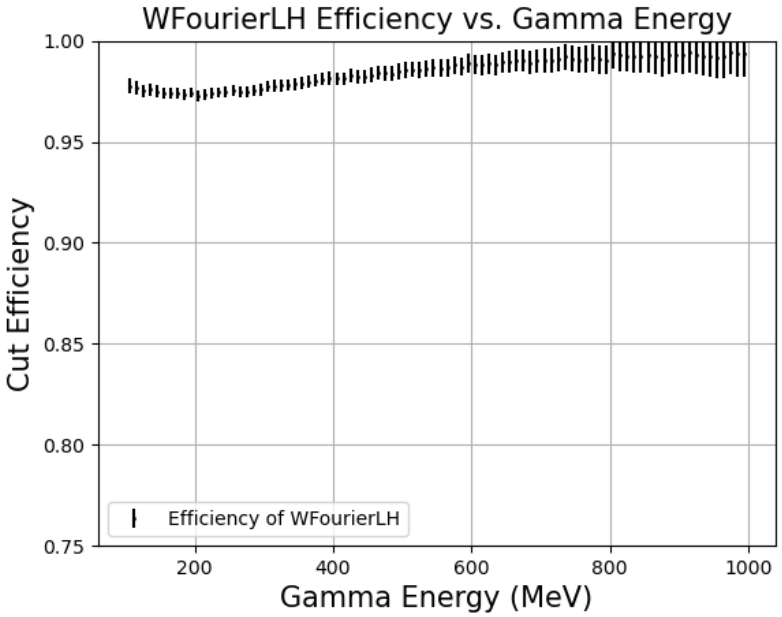


Figure 7.14: The efficiency of WFourierLH > 0.19 cut condition as a function of photon energy in $K_L \to 3\pi^0$ decays. The error bars represent the statistical uncertainty.

Notably, the efficiency drops at lower photon energies. This is likely caused by small pulses being more susceptible to pedestal fluctuations, affecting the pulse shape. To compare the acceptance in $K_L \to 3\pi^0$ data, which is found by directly applying the cut condition, to the acceptance of the $K_L \to 3\pi^0$ Monte Carlo, this efficiency function is used to determine the probability of rejection. A random number $r \in (0, 1)$ is generated uniformly for each photon in the decay. Based on the energy of that photon and the efficiency at the energy bin in data, the event is rejected if the random number is larger than the efficiency. The result after evaluating the acceptance using these strategies is shown in Eq. 7.13.

$$\Delta = \frac{|A_{data} - A_{MC}|}{A_{data}} = 0.7\% \tag{7.13}$$

For the other variables, MinCSDDLVal and $\chi^2_{fit}(2\pi^0\gamma\gamma)$, the strategy follows directly from the exclusive acceptance comparison between $K_L \to 3\pi^0$ data and $K_L \to 3\pi^0$ Monte Carlo which is described in Section 6.3. The only difference is that the cut conditions are tighter. The result is shown in Fig. 7.15.

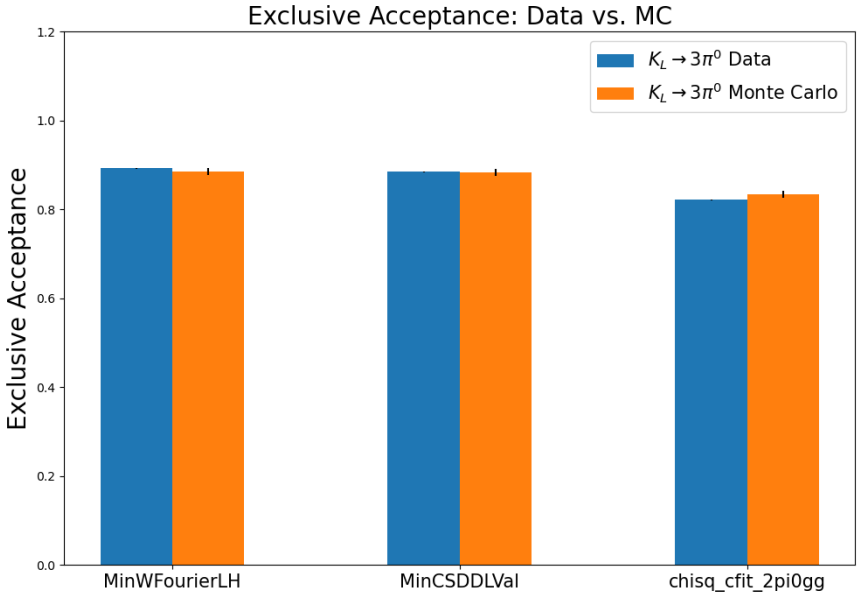


Figure 7.15: The exclusive acceptance for MinWFourierLH, MinCSDDLVal and $\chi^2_{fit}(2\pi^0\gamma\gamma)$.

The final systematic error result for these new cut conditions is shown in Eq. 7.14.

$$\Delta_{tot} = \sqrt{\sum_i \Delta_i^2} = 1.7\% \tag{7.14}$$

Lastly, the systematic error study for $\chi^2_{fit}(3\pi^0)$ is considered. Because this variable is explicitly constrained for $K_L \to 3\pi^0$, the same strategy of comparing $K_L \to 3\pi^0$ data and $K_L \to 3\pi^0$ Monte Carlo cannot be used like for $\chi^2_{fit}(2\pi^0\gamma\gamma)$. Instead, the strategy follows similarly to that done for the systematic uncertainty related to the geometrical acceptance.

The evaluation of the χ^2_{fit} relies on the measured resolution of the CSI calorimeter, which is listed in Eq. 2.6 - 2.8. If the true resolutions are different than their measured values used in the Monte Carlo simulation, it will impact the χ^2_{fit} result, as can be seen in Eq 5.12. Simultaneously, the systematic error result is already obtained for $\chi^2_{fit}(2\pi^0\gamma\gamma)$ at 1.5%. The strategy goes as follows:

1. Select a random number, r, from a gaussian distribution centered at zero, and with a width given by $\sqrt{2} \cdot \delta\sigma$, which is the error propagation term to χ^2_{fit}.
2. Scale the χ^2_{fit} by this random number such that $\chi^2_{fit,new} = (1 + r)\chi^2_{fit}$.
3. Evaluate the exclusive acceptance using $\chi^2_{fit,new}$ and χ^2_{fit}.
4. Determine the acceptance difference ratio, $\Delta = \frac{A_{ini} - A_{new}}{A_{ini}}$ for each χ^2_{fit}.
5. Repeat the process 100,000 times with different random numbers.
6. Plot the different Δs, and fit a Gaussian function to the two results. The systematic error is the $\pm 1\sigma$ result from the fitting.
7. Confirm the $\chi^2_{fit}(2\pi^0)$ result matches the expectation from the $K_L \to 3\pi^0$ data and Monte Carlo study.

The optimal Gaussian width for the random number generator was determined to be about 2.8%, which corresponds to an averaged 2% error in the CSI resolution values. The fitting for $\chi^2_{fit}(2\pi^0\gamma\gamma)$ is shown in Fig. 7.16. The width matches expectation at 1.5% systematic error obtained from the previous result.

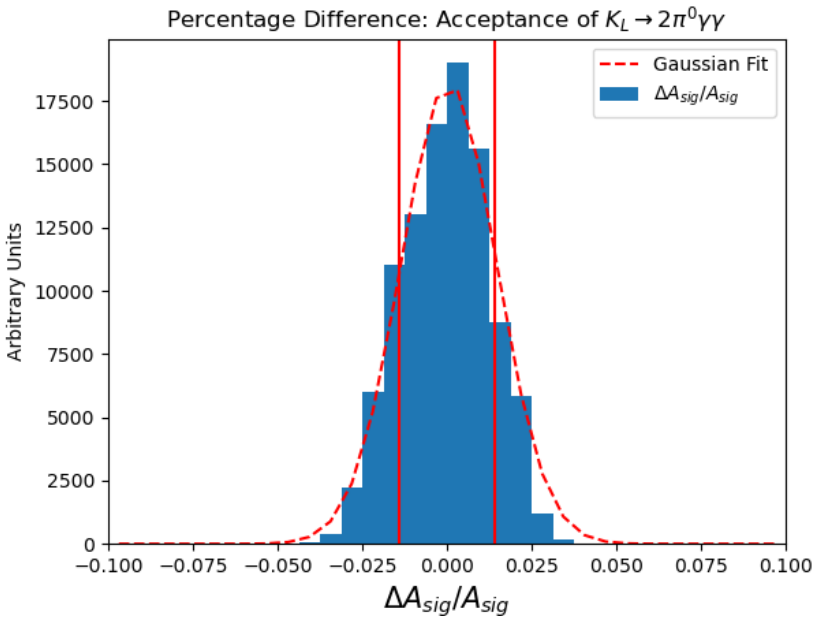


Figure 7.16: The Δ distribution for different resolution parameters, assuming $\chi^2_{fit}(2\pi^0\gamma\gamma) < 3$ The two vertical red lines represent the $\pm 1\sigma$ region.

There are multiple different cut values for $\chi^2_{fit}(3\pi^0)$ ranging from 115-188. Each of these values was tested to see if there was a dependence of cut value on the systematic uncertainty. The results are outlined in Table 7.3. The systematic error is larger than that of the $\chi^2_{fit}(2\pi^0\gamma\gamma) < 3$ cut condition because large χ^2 values are impacted significantly more by a percentage error. An example distribution of Δs is shown in Fig. 7.17.

Cut Condition	Systematic Uncertainty
$\chi^2_{fit} < 115$	5.0%
$\chi^2_{fit} < 127$	6.4%
$\chi^2_{fit} < 137$	5.7%
$\chi^2_{fit} < 143$	6.3%
$\chi^2_{fit} < 144$	6.3%
$\chi^2_{fit} < 146$	7.0%
$\chi^2_{fit} < 159$	7.2%
$\chi^2_{fit} < 160$	7.1%
$\chi^2_{fit} < 166$	6.1%
$\chi^2_{fit} < 178$	8.8%
$\chi^2_{fit} < 188$	6.6%

Table 7.3: This table contains the systematic error based on the $\chi^2_{fit}(3\pi^0)$ cut condition. In general, the larger the cut value then the larger the systematic uncertainty.

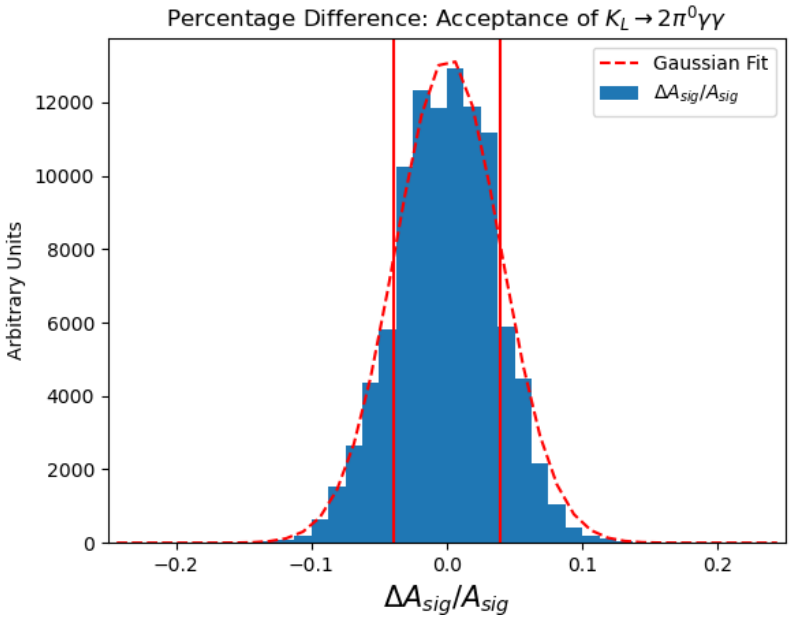


Figure 7.17: The Δ distribution for different resolution parameters assuming $\chi^2_{fit}(3\pi^0) > 115$. The two vertical red lines represent the $\pm 1\sigma$ region.

7.7.2 Signal Acceptance and SES

The signal acceptance can now be evaluated from the Monte Carlo after applying their respective cut conditions. The result is shown in Fig. 7.18, with the $K_L \to \pi^0\pi^0\gamma\gamma$ result shown by the horizontal line. The shaded area around it is the $\pm 1\sigma$ result.

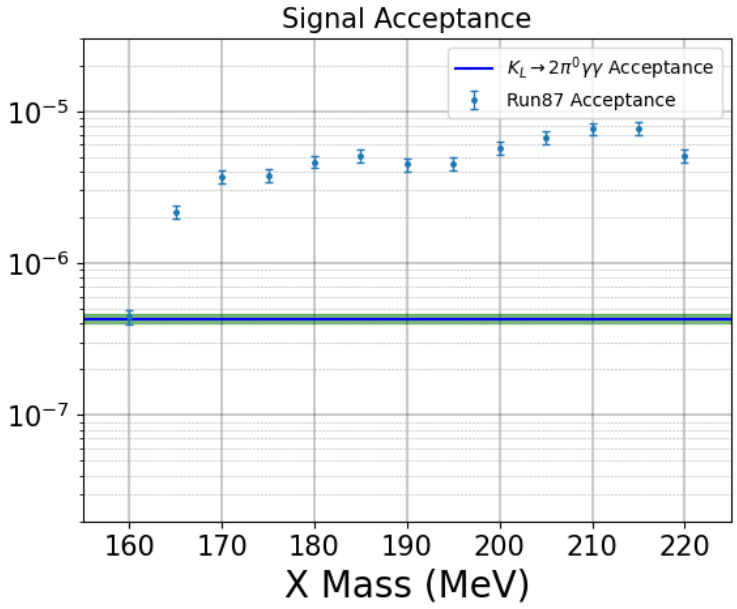


Figure 7.18: The signal acceptance as a function of X mass.

In general, the closer to the pion mass, the worse the signal acceptance. Two factors account for this trend. First, the background levels are larger towards the pion mass, and second the signal efficiency after applying the $\chi^2_{fit}(3\pi^0)$ cut condition is low because the X mass is closer to the pion mass. Directly scaling to the SES by using Eq. 7.12 gives Fig. 7.19. The summary of the acceptance and SES results are shown in Table 7.4.

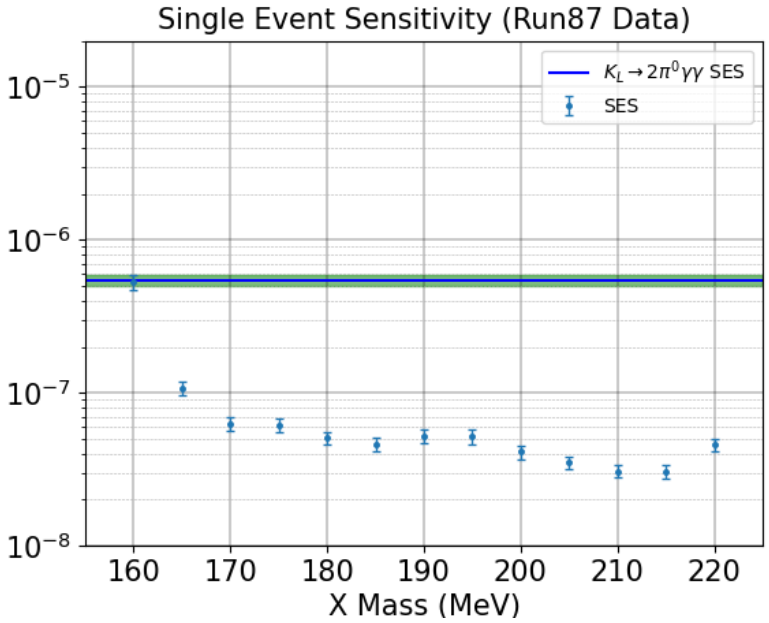


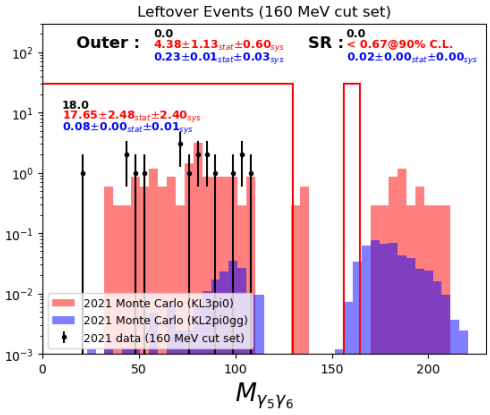
Figure 7.19: The single event sensitivity as a function of X mass.

Analysis	Acceptance ($\times 10^{-6}$)	SES ($\times 10^{-8}$)
$M_X = 160$	0.444 ± 0.051	52.8 ± 6.0
$M_X = 165$	2.17 ± 0.22	10.8 ± 1.1
$M_X = 170$	3.71 ± 0.37	6.32 ± 0.64
$M_X = 175$	3.79 ± 0.36	6.18 ± 0.60
$M_X = 180$	4.63 ± 0.43	5.06 ± 0.47
$M_X = 185$	5.10 ± 0.52	4.60 ± 0.47
$M_X = 190$	4.47 ± 0.44	5.25 ± 0.52
$M_X = 195$	4.53 ± 0.46	5.18 ± 0.53
$M_X = 200$	5.70 ± 0.58	4.11 ± 0.42
$M_X = 205$	6.70 ± 0.64	3.50 ± 0.33
$M_X = 210$	7.61 ± 0.67	3.08 ± 0.27
$M_X = 215$	7.66 ± 0.74	3.06 ± 0.30
$M_X = 220$	5.10 ± 0.48	4.60 ± 0.44
$K_L \to \pi^0\pi^0\gamma\gamma$	0.431 ± 0.041	54.4 ± 5.2

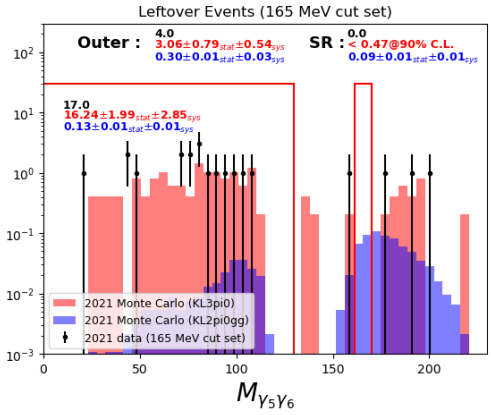
Table 7.4: This table contains the signal acceptances and SES for each analysis. The error is dominated by systematics, in particular related to $\chi^2_{fit}(3\pi^0)$, and not Monte Carlo statistics.

7.8 Results

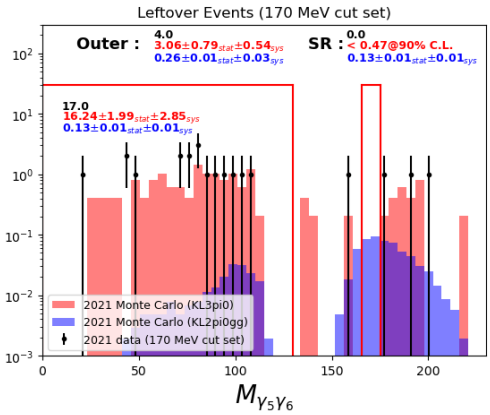
The results are shown for each analysis, based on the cut conditions shown in Tables 6.2 and 7.1. The data result is compared to the $K_L \to 3\pi^0$ Monte Carlo, and to $K_L \to \pi^0\pi^0\gamma\gamma$ Monte Carlo assuming a branching ratio of 8.4×10^{-8} [24].



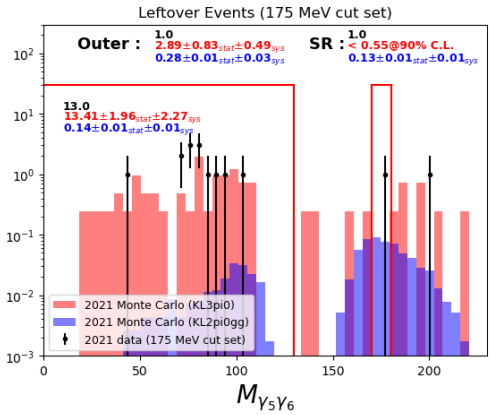
(a) The result for X mass of 160 MeV.



(b) The result for X mass of 165 MeV.

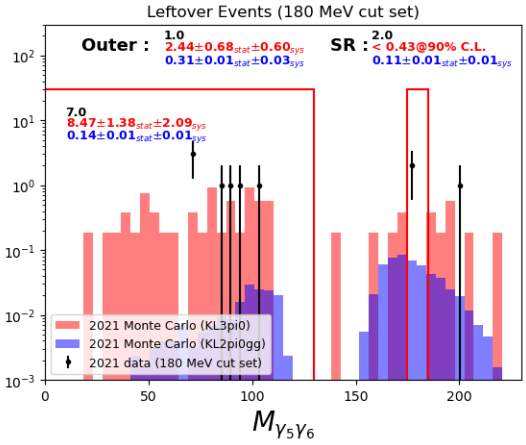


(c) The result for X mass of 170 MeV.

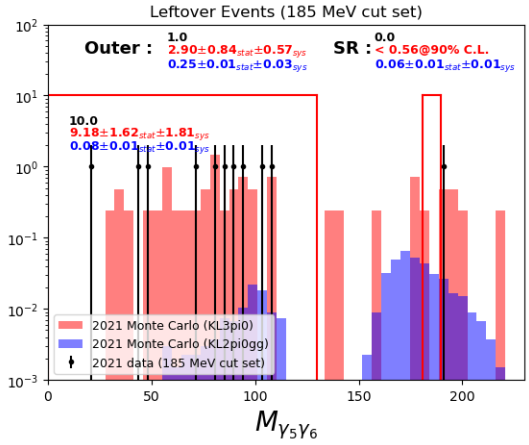


(d) The result for X mass of 175 MeV.

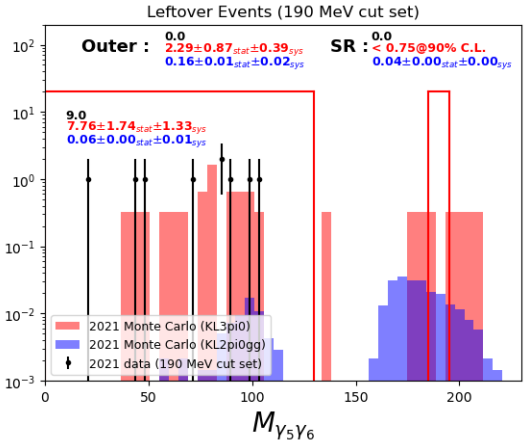
Figure 7.20: The results for each analysis. The $K_L \to 3\pi^0$ and $K_L \to \pi^0\pi^0\gamma\gamma$ Monte Carlo are included, and the expected number for each are shown in red and blue respectively. The signal region is denoted by the red box in the high mass region, and the expectation is denoted by the result labeled "SR". The expectation in the high mass region outside the signal region is labeled "Outer".



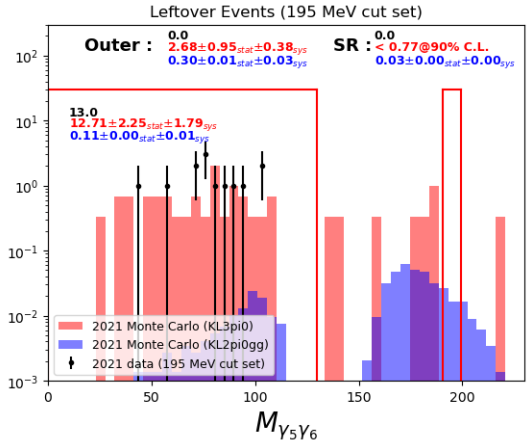
(e) The result for X mass of 180 MeV.



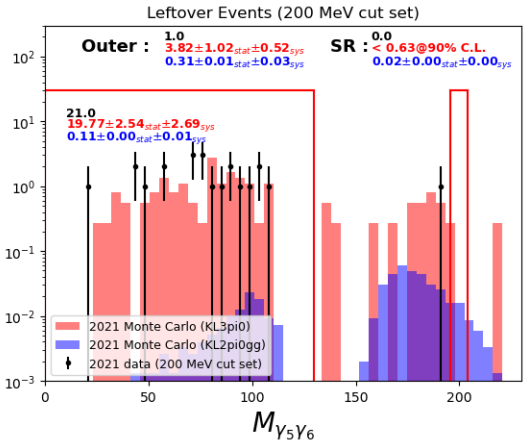
(f) The result for X mass of 185 MeV.



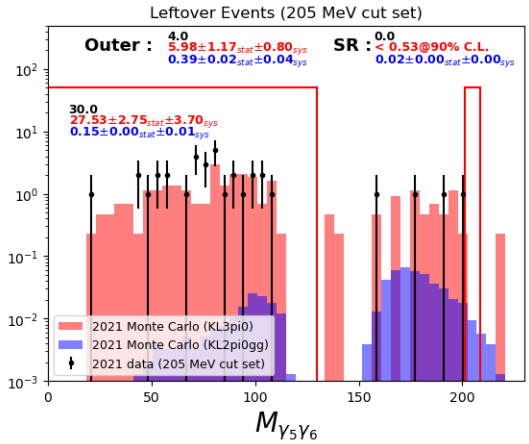
(g) The result for X mass of 190 MeV.



(h) The result for X mass of 195 MeV.

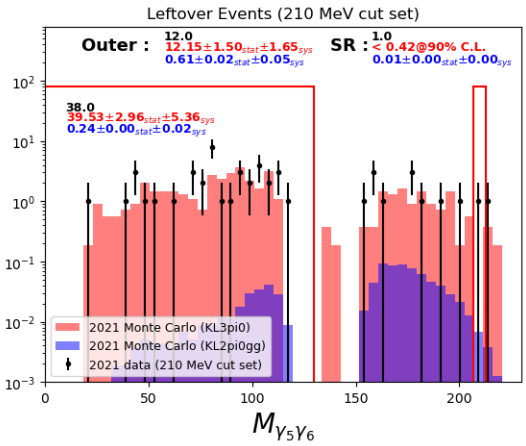


(i) The result for X mass of 200 MeV.

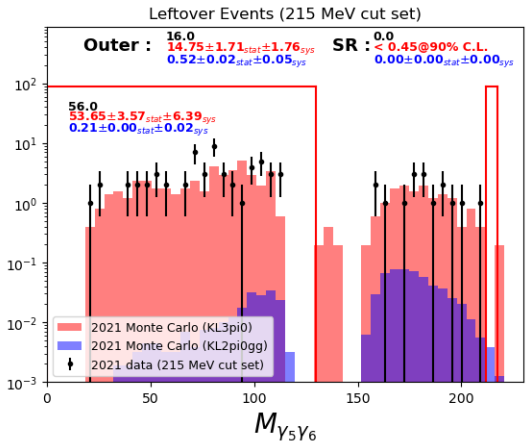


(j) The result for X mass of 205 MeV.

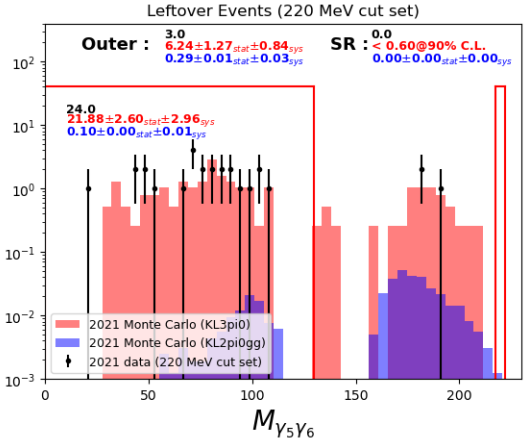
Figure 7.20: Continued



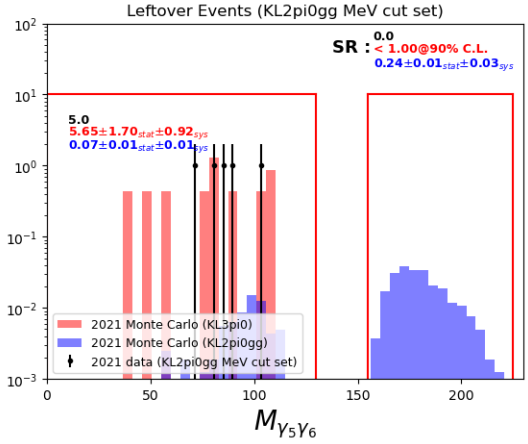
(k) The result for X mass of 210 MeV.



(l) The result for X mass of 215 MeV.



(m) The result for X mass of 220 MeV.



(n) The result for $K_L \to \pi^0\pi^0\gamma\gamma$.

Figure 7.20: Continued

From these results, the upper limits are constructed using the Feldman-Cousins approach [23], where the mean background was assumed to be,

$$N_{bkg} = UL_{3\pi^0}/2.3 + N_{2\pi^0\gamma\gamma} \tag{7.15}$$

where $UL_{3\pi^0}$ is the 90% C.L. upper limit on the $K_L \to 3\pi^0$ background in the ROI, and $N_{2\pi^0\gamma\gamma}$ is the number of expected $K_L \to \pi^0\pi^0\gamma\gamma$ in the ROI based on the theoretical branching ratio. The final results are shown in Fig. 7.21 and outlined in detail in Table 7.5. In Fig. 7.21, the KOTO results are compared to the E391A results, which saw up to an order of magnitude improvement in sensitivity in comparable regions and new results near the pion mass.

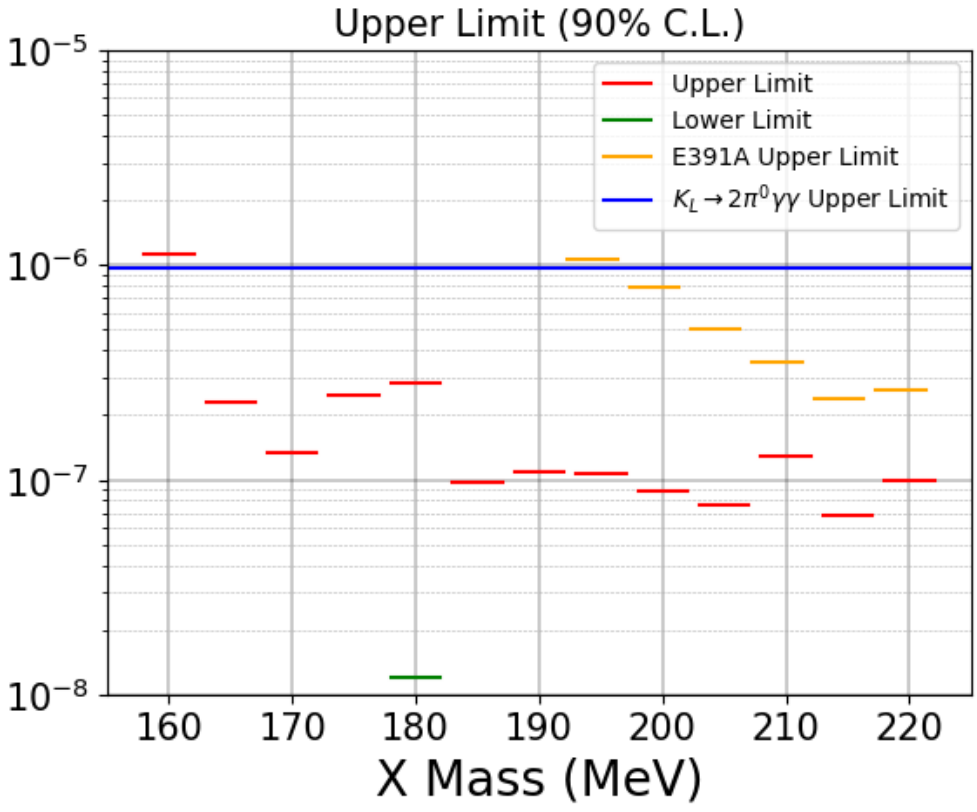


Figure 7.21: The upper limits set using the Feldman-Cousins approach. Due to the two events observed in the $M_X = 180$ MeV analysis, there is a non-zero lower limit on the signal rate at 90% C.L.

Analysis	N_{obs}	Result @90% C.L. ($\times 10^{-7}$)
$M_X = 160$ MeV	0	$\mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 11.2$
$M_X = 165$ MeV	0	$\mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 2.31$
$M_X = 170$ MeV	0	$\mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 1.33$
$M_X = 175$ MeV	1	$\mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 2.47$
$M_X = 180$ MeV	2	$0.12 < \mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 2.84$
$M_X = 185$ MeV	0	$\mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 0.980$
$M_X = 190$ MeV	0	$\mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 1.09$
$M_X = 195$ MeV	0	$\mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 1.07$
$M_X = 200$ MeV	0	$\mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 0.880$
$M_X = 205$ MeV	0	$\mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 0.765$
$M_X = 210$ MeV	1	$\mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 1.28$
$M_X = 215$ MeV	0	$\mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 0.685$
$M_X = 220$ MeV	0	$\mathrm{BR}(K_L \to \pi^0\pi^0 X, X \to \gamma\gamma) < 1.00$
$K_L \to \pi^0\pi^0\gamma\gamma$	0	$\mathrm{BR}(K_L \to \pi^0\pi^0\gamma\gamma) < 9.77$

Table 7.5: This table contains the 90% C.L. limits on the branching ratio for each analysis, and the number observed in data.

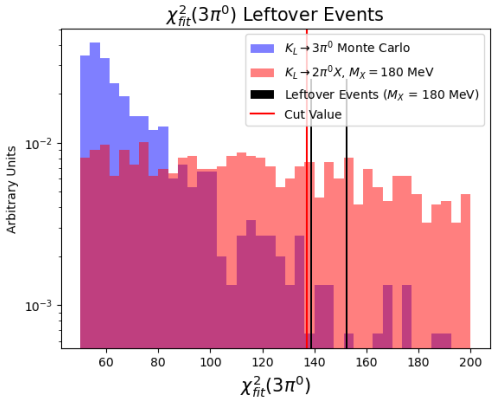
7.9 Discussion

The sum of $N_{obs} = 4$ from Table 7.4. However, there are only three unique events. One of the events observed in the $M_X = 180$ MeV analysis is also the event observed in the $M_X = 175$ MeV analysis. The two events observed near 180 MeV are contained within the same histogram bin, and have similar invariant masses. Details on these events are shown in Table 7.6.

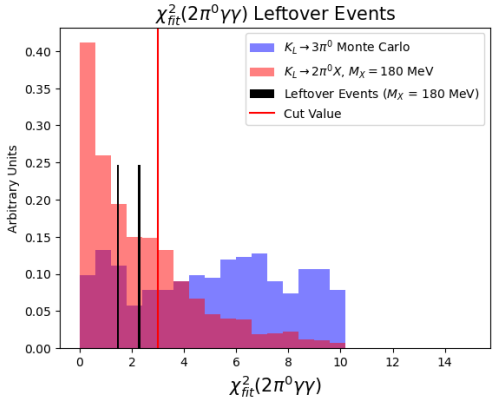
$M_{\gamma\gamma}$ (MeV)	$\chi^2_{fit}(3\pi^0)$	$\chi^2_{fit}(2\pi^0\gamma\gamma)$	MinCSDDLVal
176.674	152.0	1.47	0.90
178.522	138.8	2.27	0.96
211.344	132.1	4.35	0.89

Table 7.6: This table contains information regarding the three leftover events. Row one contains the event observed in two analyses.

The kinematics of the leftover events in the analysis with $M_X = 180$ MeV are shown in Fig. 7.22, overlaid with the distributions of $K_L \to 3\pi^0$ and $K_L \to \pi^0\pi^0 X$ Monte Carlo under loose cut conditions. The details are listed in the caption.

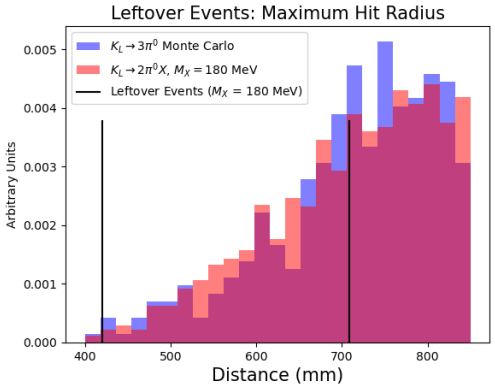


(a) The $\chi^2_{fit}(3\pi^0)$ for the leftover events compared with Monte Carlo.

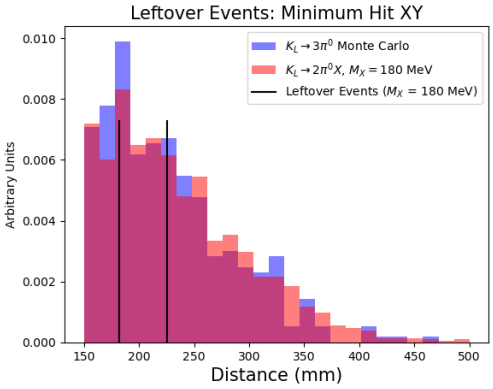


(b) The $\chi^2_{fit}(2\pi^0\gamma\gamma)$ for the leftover events compared with Monte Carlo.

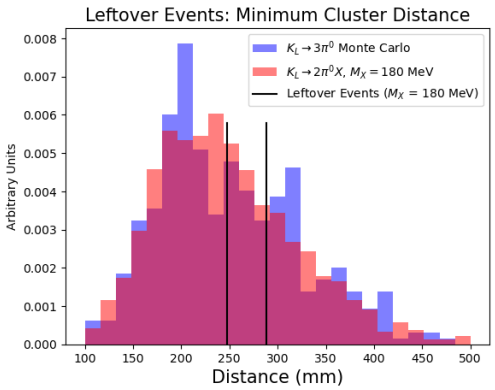
Figure 7.22: Comparison of Monte Carlo kinematic distribution and the two leftover events from the analysis with $M_X = 180$ MeV. Loose cut conditions are applied to observe the distributions.



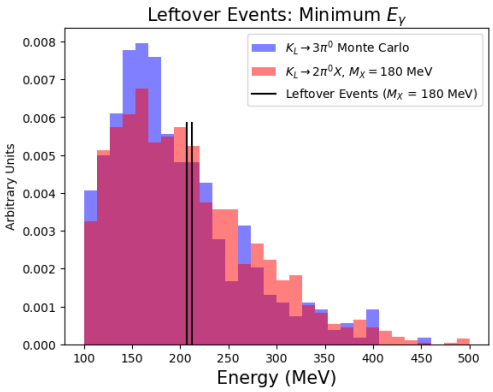
(c) The maximum photon hit radius in the CSI. The x-axis is in mm.



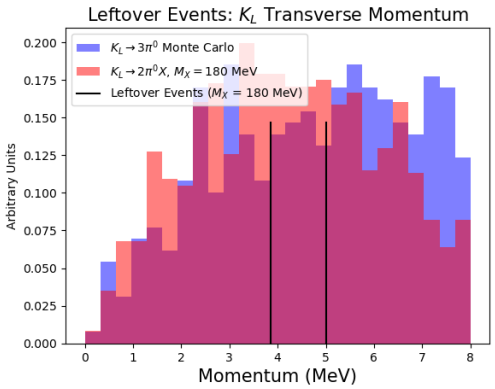
(d) The minimum photon hit X-Y in the CSI. The x-axis is in mm.



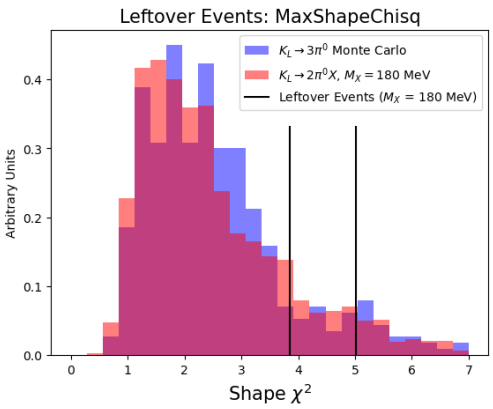
(e) The minimum of all cluster distances from each other in the CSI. The x-axis is in mm.



(f) The minimum reconstructed photon energy in the event. The x-axis is in MeV.



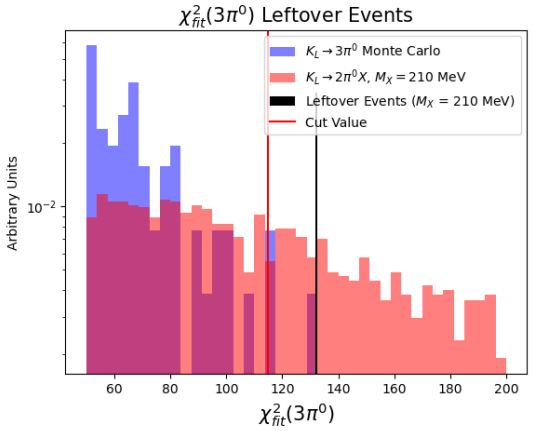
(g) The K_L transverse momentum.



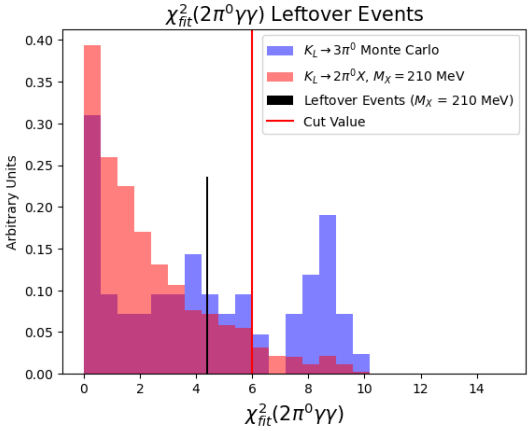
(h) The Shape χ^2 of the leftover events.

Figure 7.22: Continued

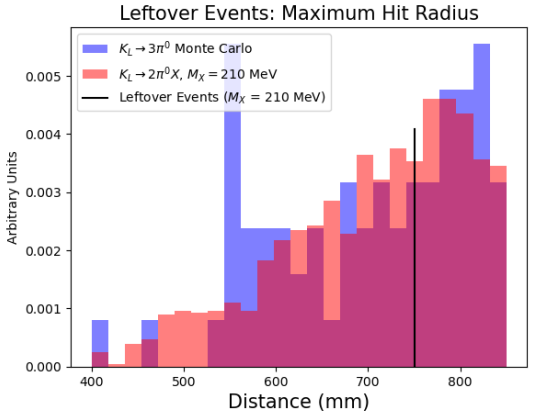
The notable points of interest is that one of the events has a very low maximum hit radius, and the two events have a larger than normal Shape χ^2. However, the MinCSDDLVal is well within the cut values for these events. The kinematics for the one event observed in the $M_X = 210$ MeV analysis are shown in Fig. 7.9.



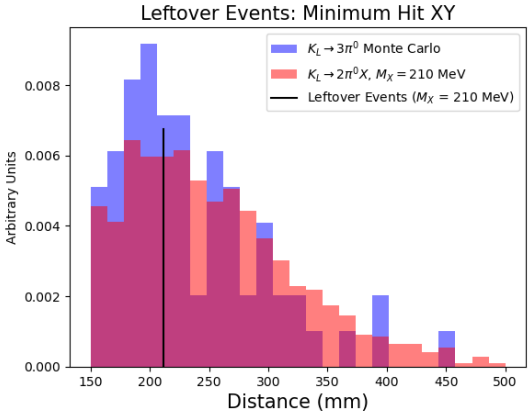
(a) The $\chi^2_{fit}(3\pi^0)$ for the leftover events compared with Monte Carlo.



(b) The $\chi^2_{fit}(2\pi^0\gamma\gamma)$ for the leftover events compared with Monte Carlo.

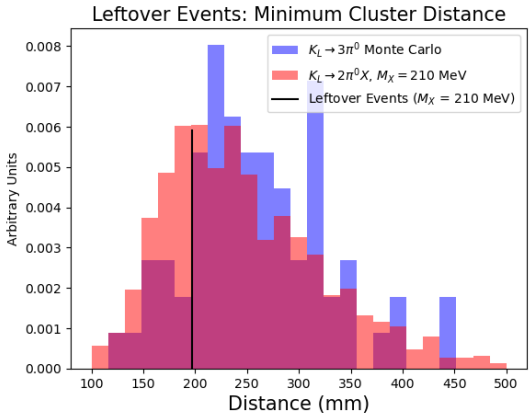


(c) The maximum photon hit radius in the CSI. The x-axis is in mm.

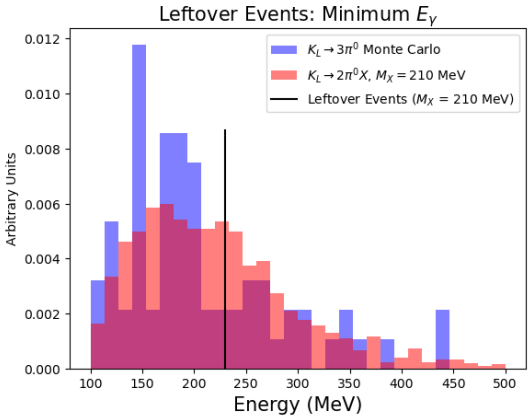


(d) The minimum photon hit X-Y in the CSI. The x-axis is in mm.

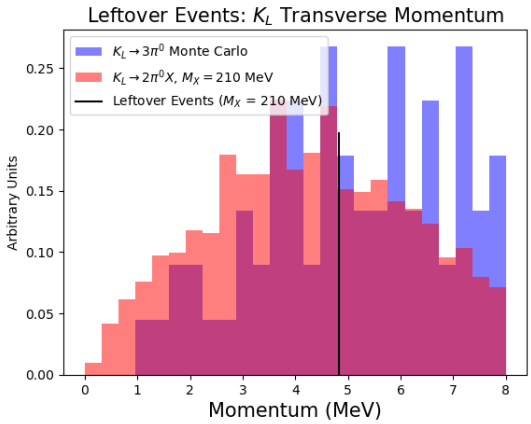
Figure 7.23: Comparison of Monte Carlo kinematic distribution and the two leftover events from the analysis with $M_X = 180$ MeV. Loose cut conditions are applied to observe the distributions. All cuts except MinWFourierLH are applied with loose cuts, $\chi^2_{fit}(3\pi^0) > 50$, $\chi^2_{fit}(2\pi^0\gamma\gamma) < 10$, and MinCSDDLVal > 0.5.



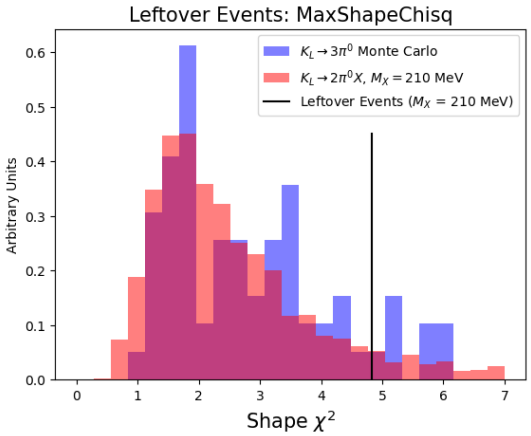
(e) The minimum of all cluster distances from each other in the CSI. The x-axis is in mm.



(f) The minimum reconstructed photon energy in the event. The x-axis is in MeV.



(g) The K_L transverse momentum.



(h) The Shape χ^2 of the leftover events.

Figure 7.23: Continued

There are a few discrepancies between the result in data and the expectation from Monte Carlo simulation. The largest discrepancy occurs in the analysis with $M_X = 160$ MeV. The background expectation outside of the signal region was $4.38 \pm 1.13_{stat} + 0.60_{sys}$, while the number observed was 0. Using Poisson statistics, and averaging after sampling around the mean based on the statistical and systematic error, the probability of this observation is around 3%.

Applying the same reasoning to the analysis with $M_X = 180$ MeV and assuming there is no signal and the mean background for $K_L \to 3\pi^0$ is at the 90% C.L. upper limit, the

probability to observe two events including $K_L \to \pi^0\pi^0\gamma\gamma$ is about 8.5%. Ultimately, this analysis relies on high statistics Monte Carlo, and yet the background estimation is set with an upper limit within the signal region. To provide detailed insights on this result, at least a factor of two times more $K_L \to 3\pi^0$ Monte Carlo statistics is necessary, as it will provide a better understanding of the background level inside and outside of the signal regions.

The interpretation of these results is not trivial, and work is ongoing. More discussion on this topic may be found in Section 9.1.

CHAPTER 8

ANALYSIS OF $K_L \to \pi^0\nu\overline{\nu}$

The primary purpose of the KOTO experiment is geared toward this analysis, which differs from the the $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\pi^0 X, X \to \gamma\gamma$ search in fundamental ways. There are missing particles in this search, being the two neutrinos, which means the signal region cannot be defined in terms of an invariant mass. Instead, the signal region is a square region in the π^0 p_T vs. Z_{vtx} plot. Additionally, this is a two cluster analysis which makes the number of potential background sources significantly higher.

However, like the analysis $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\pi^0 X, X \to \gamma\gamma$, this is a blind analysis. In the previous case, the entire high mass region was blinded prior to "opening" the box for each analysis. For $K_L \to \pi^0\nu\overline{\nu}$, the blind region is defined in Eq. 8.1.

$$120 \text{ MeV} \leq \pi^0 \, p_T \leq 250 \text{ MeV}$$

$$2900 \text{ mm} \leq Z_{vtx} \leq 5100 \text{ mm} \tag{8.1}$$

The process to determine the result follows the same strategy as the previous analyses.

1. Determine the SES.
2. Evaluate the agreement between data and the Monte Carlo simulation.
3. Estimate the background contamination inside the signal region using Monte Carlo simulation.
4. Confirm the result by studying control regions.
5. Open the blinded region and determine the result.

8.1 Single Event Sensitivity

The total amount of data taken with $K_L \to \pi^0\nu\overline{\nu}$ physics trigger corresponds to 3.3×10^{19} protons on target (POT). This data was taken with a range of beam intensities from 10 kW to 64 kW. Unlike the calculation in the previous analysis, the K_L yield is evaluated for three normalization modes, $K_L \to \gamma\gamma$, $K_L \to 2\pi^0$, and $K_L \to 3\pi^0$. However, the final SES was obtained using $K_L \to 2\pi^0$ data. The acceptance of $K_L \to 2\pi^0$ data and the $K_L \to \pi^0\nu\overline{\nu}$ signal was estimated using Monte Carlo simulation. These values were $A_{norm} = 3.44 \times 10^{-3}$ and $A_{sig} = 4.75 \times 10^{-3}$. The total number of $K_L \to 2\pi^0$ observed in data was $N_{obs} = 6.72 \times 10^5$. From that value, the SES was then calculated using Eq. 8.2.

$$SES = \frac{A_{norm} \cdot \mathrm{BR}(K_L \to 2\pi^0)}{A_{sig} \cdot N_{norm}} = (9.33 \pm 0.06_{stat} \pm 0.84_{sys}) \times 10^{-10} \tag{8.2}$$

Note that the uncertainty in the SES was dominated by systematic uncertainty. The summary of systematic error was not calculated here for the purpose of this dissertation, and is instead cited from the 2021 paper. However, the calculation follows the same strategies outlined in Section 6.3.5, which typically rely on the exclusive acceptance difference, $\Delta = \frac{A_{data} - A_{sig}}{A_{data}}$. In this case, A_{sig} represents the exclusive acceptance of $K_L \to \pi^0\nu\overline{\nu}$ signal from Monte Carlo simulation, and A_{data} corresponds to the exclusive acceptance of $K_L \to 2\pi^0$ data. Ultimately, the systematic error encodes the extent to which data and Monte Carlo results differ. The reconstructed K_L mass is a litmus test on the efficacy of the reconstruction, and is shown in Fig. 8.1, while the results of the systematic error are shown in Table 8.1.

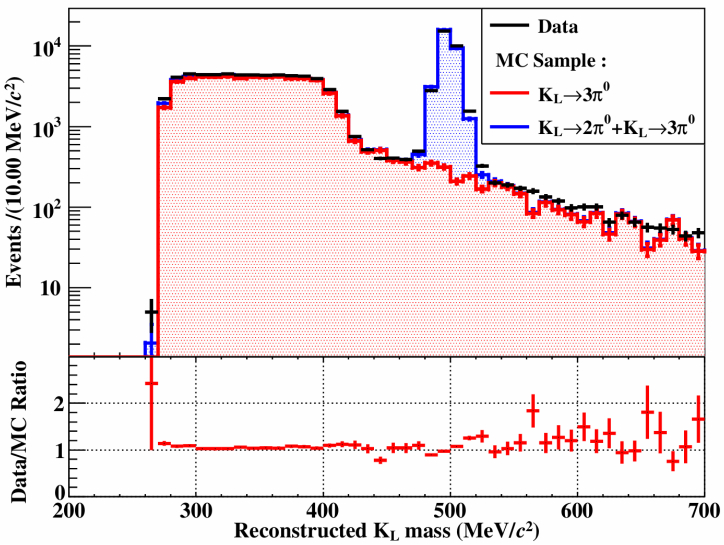


Figure 8.1: The reconstructed K_L mass after the $K_L \to 2\pi^0$ selection criteria for the SES calculation [4].

Source	Uncertainty
Veto Cuts	3.4%
Kinematic Cuts for $K_L \to 2\pi^0$	4.6%
Shape Related Cuts	4.2%
Normalization Mode Inconsistency	3.5%
Kinematic Cuts for $K_L \to \pi^0\nu\overline{\nu}$	3.5%
K_L Momentum Spectrum	0.98%
Trigger	1.91%
Photon Selection Cuts	0.79%
$K_L \to 2\pi^0$ Branching Fraction	0.69%
Δ T Cut	0.19%
Total	**9.0%**

Table 8.1: The summary of systematic errors associated with the SES calculation. The total is the quadratic sum of each of the sources.

Interestingly, the systematic error for the veto and online trigger are less than that observed in the analysis of $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\pi^0 X, X \to \gamma\gamma$ because of the difference in final state particles incident on the CSI. For instance, the online clustering algorithm was impacted by a few channels having the cables swapped. This impacted the accuracy of the online clustering algorithm, which indirectly affected the analyses with six incident clusters instead of two. Likewise, the veto systematic uncertainty is dominated by the inconsistency in exclusive acceptance for the Inner Barrel Veto. This veto is impacted the most by splash back, which may not be well simulated and proportionally impacts data with more clusters in the CSI.

8.2 History: 2016-2018 Data

To understand the work and progress made in the KOTO analysis for 2021 data, a quick look into the history of work done in 2016-2018 is necessary. This analysis was plagued by a large background estimation, with the details shown in Table 8.2.

Source	Background Level
$K_L \to 3\pi^0$	0.01 ± 0.01
$K_L \to 2\gamma$ (beam halo)	0.26 ± 0.07^a
Other K_L decays	0.05 ± 0.05
$K^\pm$ decays	0.87 ± 0.25^a
Neutron induced	0.017 ± 0.002
CV η	0.03 ± 0.01
NCC π^0	0.03 ± 0.03
Total	1.22 ± 0.26

Table 8.2: This table contains the background estimation in the 2016-2018 signal region. Rows marked with ^{a} represent background studied after opening the signal region.

The total background estimation was 1.22 ± 0.26, and after opening the blinded region, three events were observed. Using Poisson statistics, the probability of this fluctuation, accounting for the quoted error and assuming the three events are induced by the background studied, is around 10.5%. The result after opening the box is shown in Fig. 8.2.

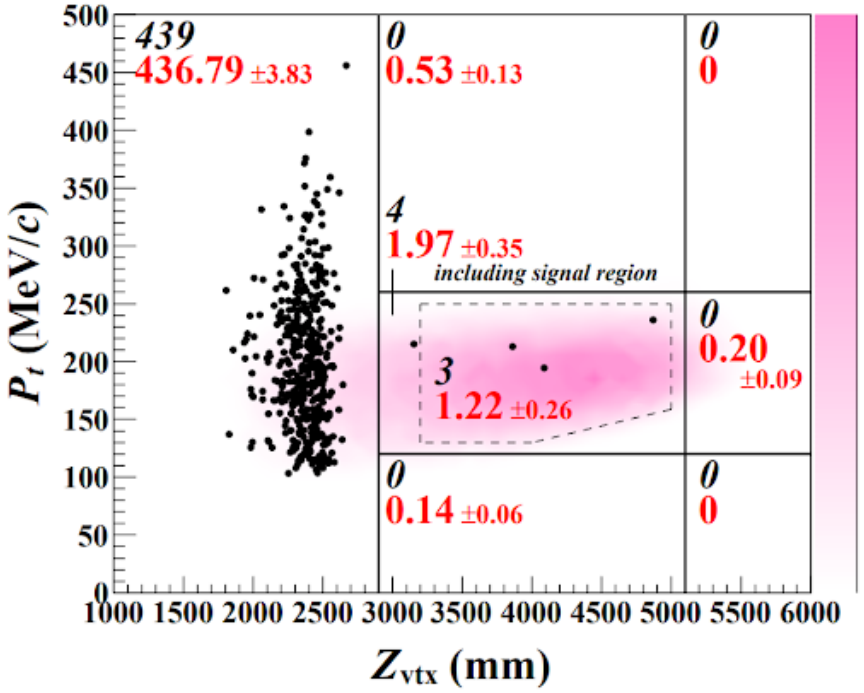


Figure 8.2: The results from the analysis of data taken in 2016-2018 [2]. The dotted region represents the signal region.

8.2.1 $K^{\pm}$ Background Rejection

In principle, there should be no charged particles entering the decay fiducial region due to the sweeping magnet inserted between the two collimators in the beamline. However, after further study it was found that the $K^{\pm}$ contamination was not negligible and was caused primarily by neutral particle interactions in the second collimator which may have charged particles as byproducts. The source of this background is predominantly from the $K^{\pm} \to \pi^0 e^{\pm} \nu$ decays. Notice that the final state is extremely similar to the decay $K_L \to \pi^0 \nu \bar{\nu}$, except with a neutrino exchanged with an $e^{\pm}$. If this $e^{\pm}$ is missed, either by inefficiency due

to it being low energy, or if it hits some dead material by traveling backwards, it is difficult to discern this decay from real signal. A diagram displaying a typical background event is shown in Fig. 8.3.

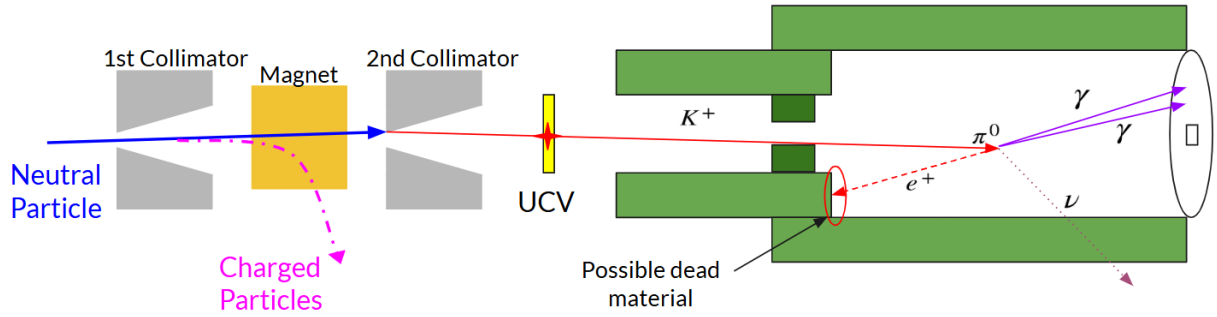


Figure 8.3: A potential background event from $K^{\pm} \to \pi^0 e^{\pm} \nu$. The Upstream Charged Veto (UCV) labeled here was not installed until 2020. Sizes are not to scale.

To combat this background, there are two potential approaches. The first is to install a plastic scintillator detector in the upstream area of the beamline. This detector, called the Upstream Charged Veto (UCV), sits between the second collimator and the entrance to the Front Barrel. An image of the UCV installed prior to the data taking runs in 2021 is shown in Fig. 8.4.

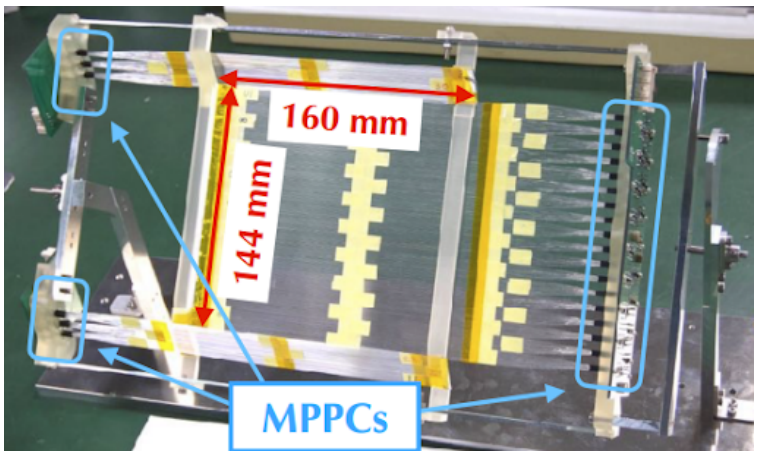
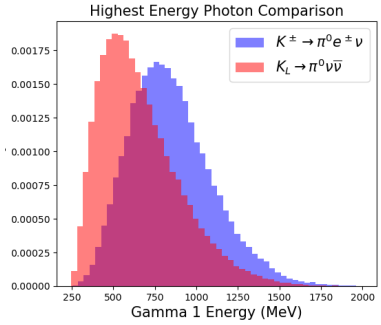


Figure 8.4: An image of the UCV installed before data taking in 2021 to reduce the background contamination from $K^{\pm}$.

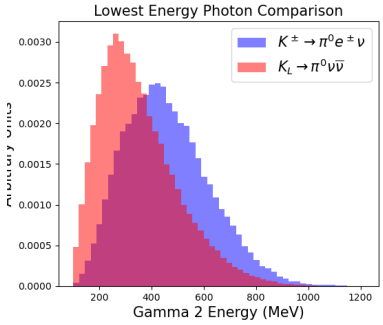
This detector is a 0.5mm thickness plastic scintillator with MPPC readout. The $K^{\pm}$

rejection was evaluated with a trigger to select $K^{\pm} \to \pi^{\pm}\pi^0$ events. The rejection factor was found to be $\approx \frac{1}{13}$.

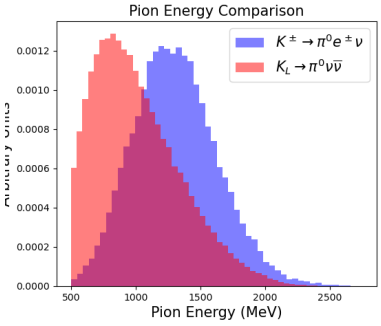
For the majority of data taken in 2019-2020 there is no UCV, and instead a new method of background rejection is needed which uses the kinematics of the typical $K^{\pm}$ background events. While the final state is remarkably similar to $K_L \to \pi^0\nu\overline{\nu}$, there are restrictions on the missing $e^{\pm}$. If it is high energy in the lab frame, the probability of detection is much higher. Thus, by requiring the $e^{\pm}$ to travel upstream or enter the in veto detectors with a very low energy (< 5 MeV), the backwards $e^{\pm}$ will affect the π^0 kinematics, which boost it downstream with higher energy in the lab frame. This can be seen by comparing Monte Carlo results requiring these strict kinematic constraints on the background in Fig. 8.5.



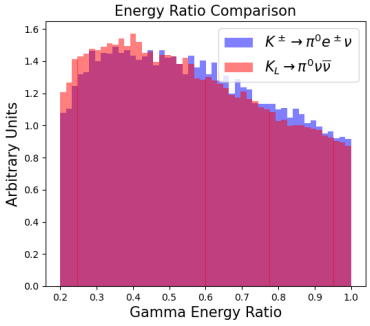
(a) The high energy photon comparison.



(b) The low energy photon comparison.

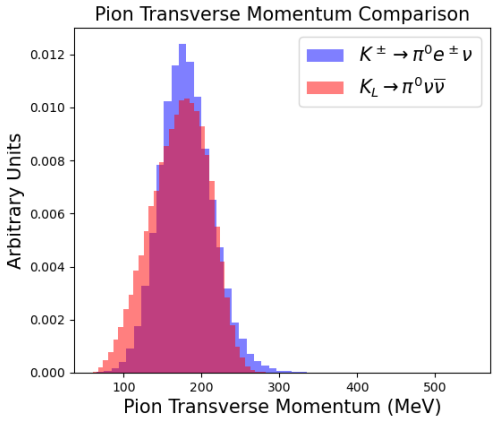


(c) The reconstructed pion energy comparison.

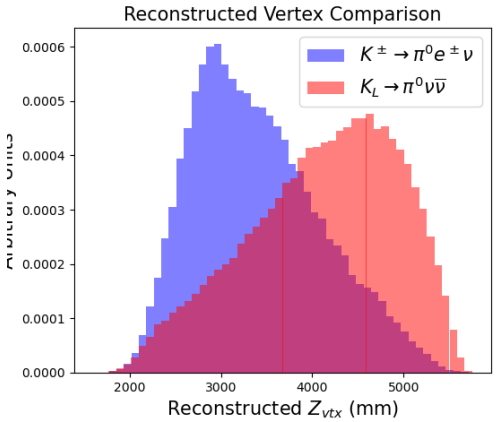


(d) The photon energy ratio comparison.

Figure 8.5: The reconstructed kinematic variables for $K_L \to \pi^0\nu\overline{\nu}$ and $K^{\pm}$ background with an assumed kinematic constraints on the $e^{\pm}$.



(e) The pion transverse momentum comparison between $K\pm$ background and signal.

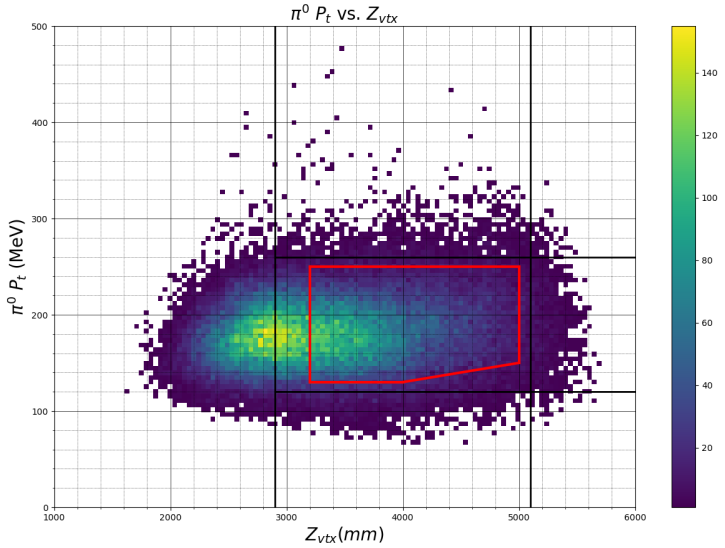


(f) The reconstructed Z_{vtx} position comparison between $K\pm$ background and signal.

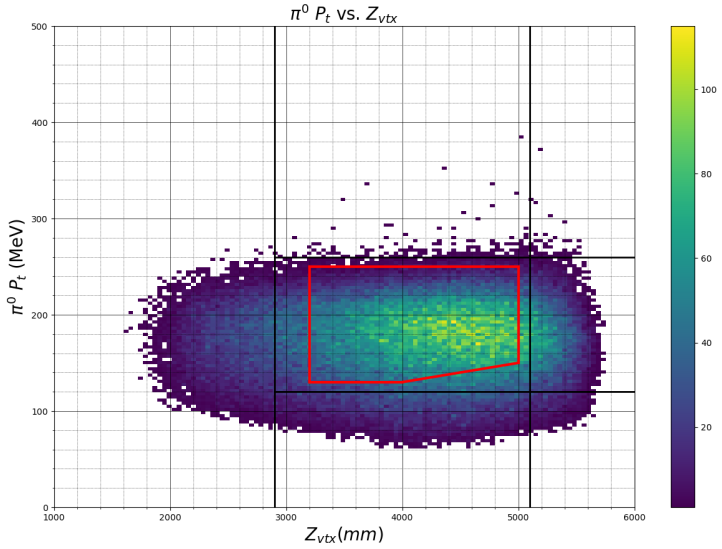
Figure 8.5: Continued

The boost in pion energy for the background also impacts other kinematics in the event, such as the reconstructed Z_{vtx}. This is due to the fact that a higher pion energy means the off-axis angle for the two photons tends to be smaller. Thus, the acceptance of decays farther from the CSI surface is larger, which can be seen more clearly when observing the 2 dimensional plots of π^0 P_t vs. Z_{vtx} for the background and signal in Fig. 8.6.

Due to the kinematic constraints on the background source, it is clear that there are kinematic correlations which may determine the $K^{\pm}$ background vs. the $K_L \to \pi^0 \nu\overline{\nu}$ signal. To optimize the rejection and signal acceptance, a simple MLP was trained with 2 hidden layers and 50 neurons for each layer. The input parameters consist of the photon and pion energies, the photon energy ratio, the cluster positions and distance, the projection angle on the CSI, the pion transverse momentum and reconstructed Z_{vtx}. The model was trained over 200 epochs, with a batch size of 1000. The inputs were normalized, with the activation layer defined by the hyperbolic tangent function. As this is a binary classification problem, the loss function was defined as the binary cross entropy, with the metric being the mean squared error. The model performance is shown in the following two plots in Fig. 8.7.

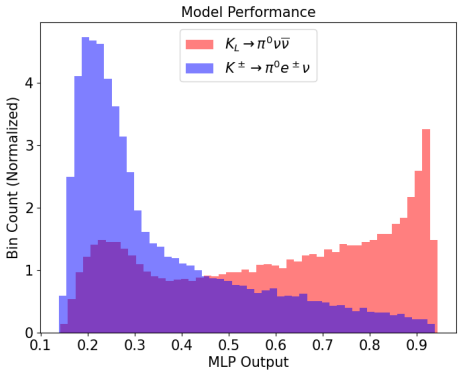


(a) The P_t vs. Z distribution for $K^{\pm} \to \pi^0 e^{\pm} \nu$ background Monte Carlo. The signal region is in red.

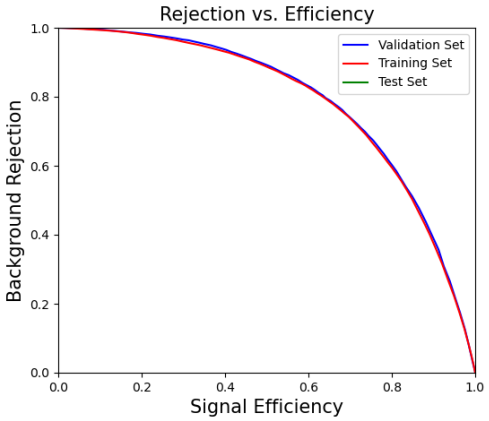


(b) The P_t vs. Z distribution for the $K_L \to \pi^0 \nu \overline{\nu}$ signal Monte Carlo. The signal region is in red.

Figure 8.6: The two distributions for both $K^{\pm} \to \pi^0 e^{\pm} \nu$ background and $K_L \to \pi^0 \nu \overline{\nu}$ signal Monte Carlo. Note that the signal tends to lie on the downstream side of the signal region.



(a) The model predictions for the training set, with values close to one being "signal-like" and values close to zero being "background-like".

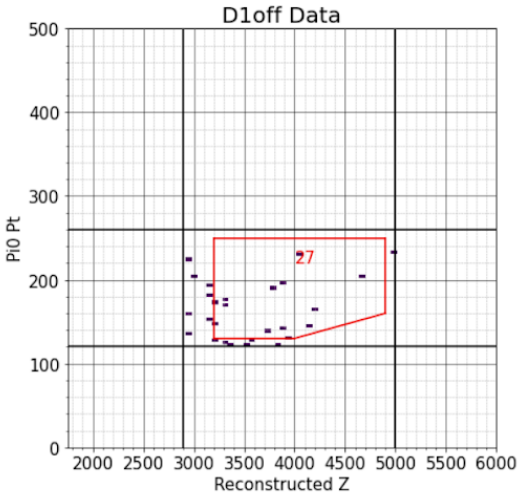


(b) The background rejection vs. signal efficiency at different cut neural network cut values. The perfomance is compared against the validation set and the test set.

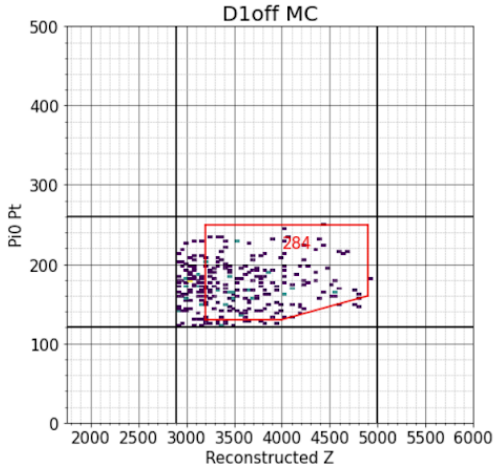
Figure 8.7: The model performance on $K^\pm$ background and signal Monte Carlo.

The performance is more efficient than a pure Z_{vtx} cut, but applying it would require a steep price in signal acceptance. However, it is a powerful tool for the $K^\pm$ background rejection for the data taken without UCV. Before that, it is crucial to confirm this model will perform well while applying it to real data. This check was done by taking data with the D1 magnet, which is the magnet that sits between the two collimators, off. By applying the typical $K_L \to \pi^0 \nu\bar{\nu}$ cuts to this data, the majority of the leftover events will be $K^\pm \to \pi^0 e^\pm \nu$ background. At the same time, it can be compared to a background Monte Carlo done with the D1 off, and confirm the agreement between the two. The results of this analysis are shown below in Fig. 8.8.

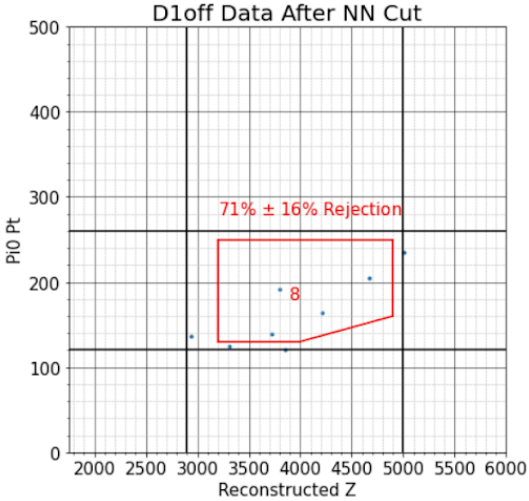
The agreement is good between data and Monte Carlo, while the statistics are low. This cut value corresponds to approximately a 40% loss in signal, which is not insignificant. More discussions must be done later as the analysis of data taken without UCV becomes more mature in applying this cut condition.



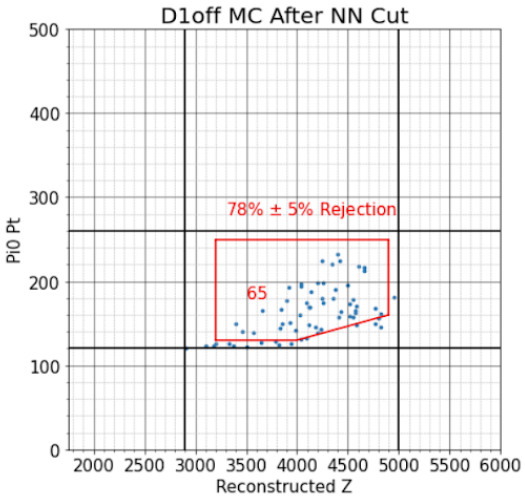
(a) The D1-off data pre-neural network cut.



(b) The D1-off Monte Carlo pre-neural network cut.



(c) The D1-off data post-neural network cut.



(d) The D1-off Monte Carlo post-neural network cut.

Figure 8.8: The model performance on D1-off $K^{\pm}$ background in data and Monte Carlo. Only the blinded region was shown in these plots, with the number of events shown in red.

8.2.2 *Halo $K_L \to 2\gamma$ Background Rejection*

The mechanism for this background is similar to the previous section. Typically, a $K_L \to 2\gamma$ decay is easily rejected by symmetry arguments. If it decays along the beam axis, the two clusters should be on opposite sides of the CSI and can rejected by the projection angle cut which requires the projected angle of the two clusters on the CSI to be smaller than 150°. However, if the decay occurs off the beam axis and close to the CSI surface, then it may be improperly reconstructed on the beam axis as a π^0. An example of this background mechanism is shown in Fig. 8.9.

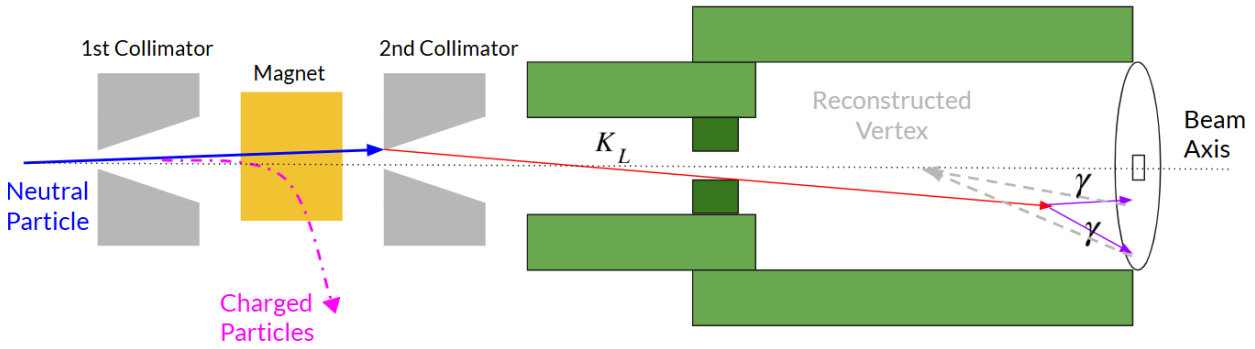


Figure 8.9: A potential background event from $K_L \to 2\gamma$. Sizes are not to scale.

An off-axis decay can occur when neutral particles scatter off of the second collimator and make it through to the decay fiducial region. Like the previous case, the kinematics of this event should be discernible from signal. Furthermore, the cluster shape should be different given that the true incident gamma angle is different from the reconstructed angle. By taking advantage of this difference, the event is reconstructed under two assumptions. The first is the background assumption, where it is a K_L decay off the beam axis originating from the second collimator and pointing towards the center of energy on the CSI. The second is the standard $K_L \to \pi^0\nu\overline{\nu}$ decay reconstruction. The cluster shapes are then compared, and a likelihood ratio value is evaluated and is defined in Eq. 8.3, with a diagram explaining the method in Fig. 8.10.

$$LR = \frac{\text{PDF(on-axis)}}{\text{PDF(on-axis)} + \text{PDF(off-axis)}} \tag{8.3}$$

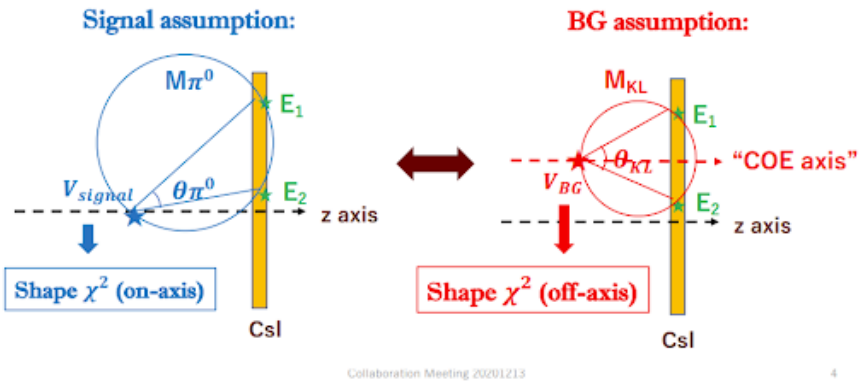


Figure 8.10: A diagram showing the two assumptions made for the discrimination of Halo $K_L \to 2\gamma$, via the likelihood ratio. Plot generated by Yuya Noichi in an internal collaboration meeting.

The results for signal and background Monte Carlo are shown in Fig. 8.11.

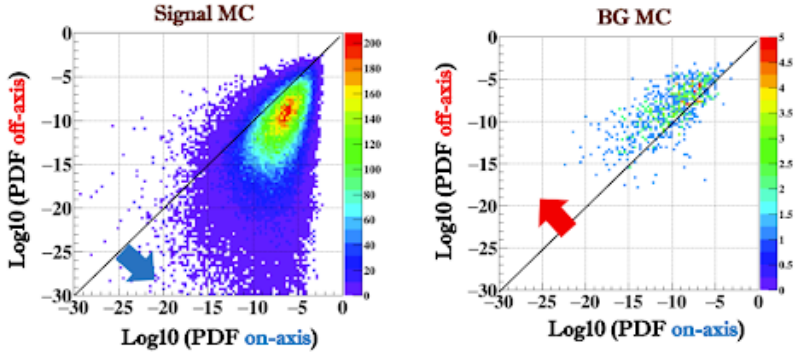


Figure 8.11: The results of the likelihood study on background and signal Monte Carlo. Plot generated by Yuya Noichi in an internal collaboration meeting.

Using this liklihood ratio, and a TMVA which also accounts for the reconstructed variables, the background rejection is around a factor of $\frac{1}{8}$, while keeping 94% signal acceptance. The result of the TMVA is shown in Fig. 8.12.

With the experience from the charged kaon background I also attempted to improve this result with the same strategy that was implemented for $K^{\pm} \to \pi^0 e^{\pm} \nu$, while also including

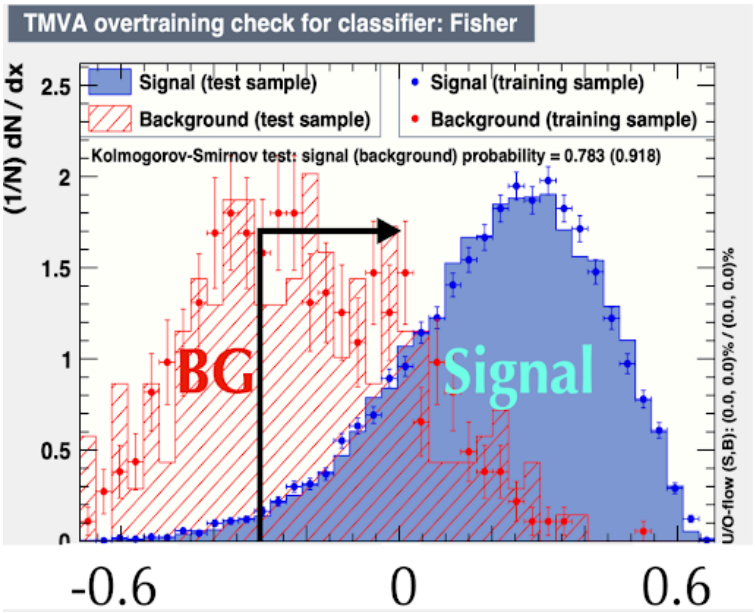
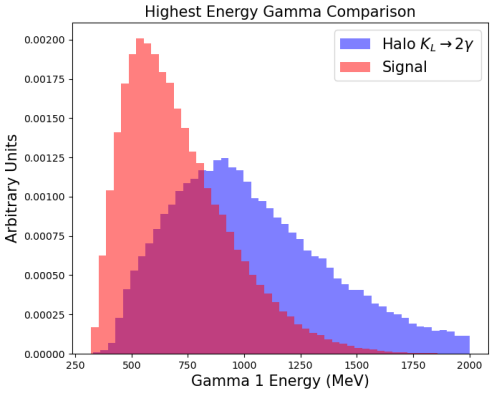
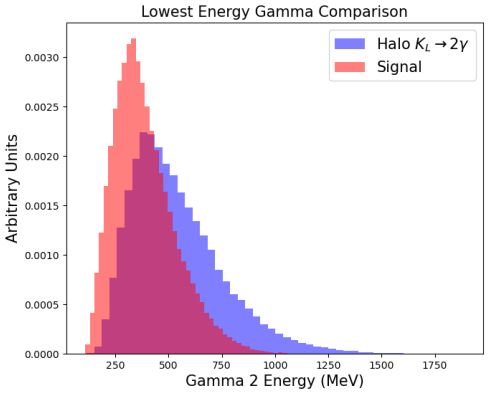


Figure 8.12: The results of the TMVA considering reconstructed variables. Plot generated by Yuya Noichi in an internal collaboration meeting

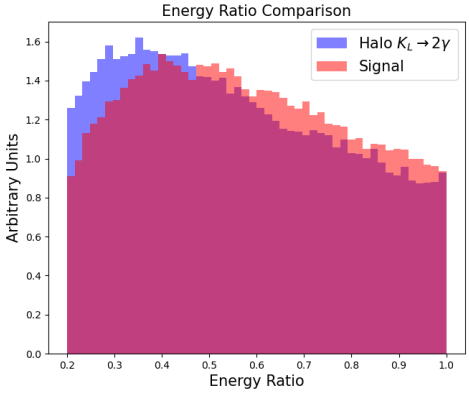
the cluster shape discriminating likelihood ratio variable described in Eq. 8.3. In addition, the previous analysis was limited by Monte Carlo statistics, while neural networks typically perform much better with larger datasets to train on. For my study, event recycling methods were used to increase the Halo K_L events obtained from the beamline simulation. These methods consisted of boosting the K_L daughter particles back into the rest frame of the K_L, performing a random, uniform rotation, and then boosting back, which provided a much larger seed file of Halo K_L Monte Carlo seed to perform the simulation. The comparison of kinematic variables used for training is shown in Fig. 8.13.



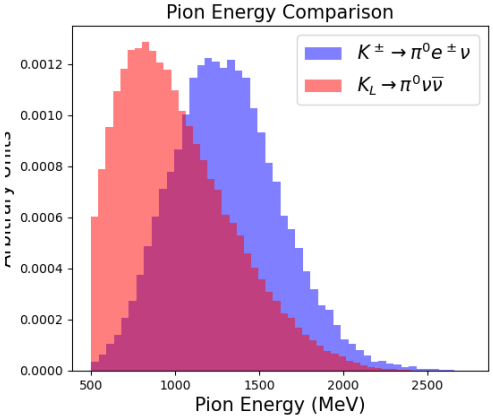
(a) The highest energy photon comparison between Halo $K_L \to 2\gamma$ and $K_L \to \pi^0\nu\overline{\nu}$ signal.



(b) The lowest energy photon comparison between Halo $K_L \to 2\gamma$ and $K_L \to \pi^0\nu\overline{\nu}$ signal.

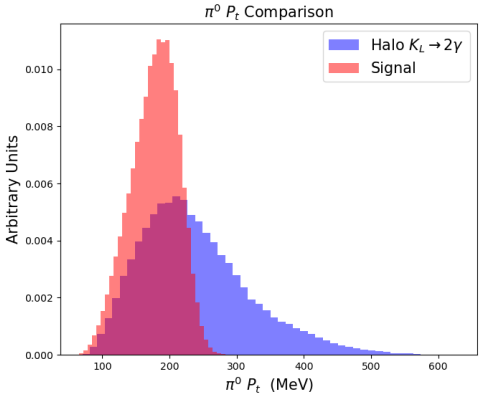


(c) The energy ratio comparison between Halo $K_L \to 2\gamma$ and $K_L \to \pi^0\nu\overline{\nu}$ signal.

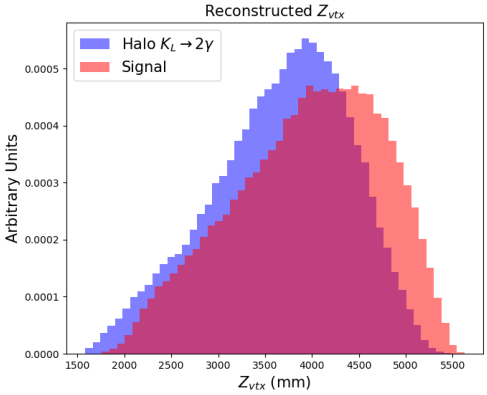


(d) The pion energy comparison between Halo $K_L \to 2\gamma$ and $K_L \to \pi^0\nu\overline{\nu}$ signal.

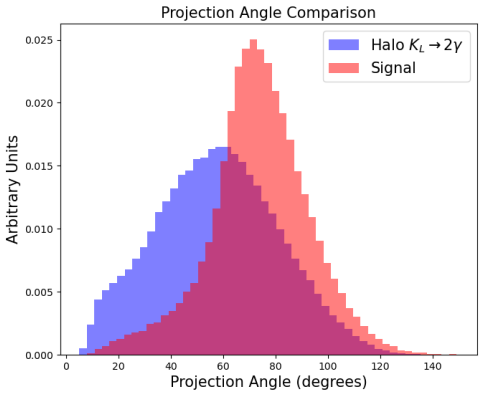
Figure 8.13: The kinematic distributions for Halo $K_L \to 2\gamma$ background and signal Monte Carlo.



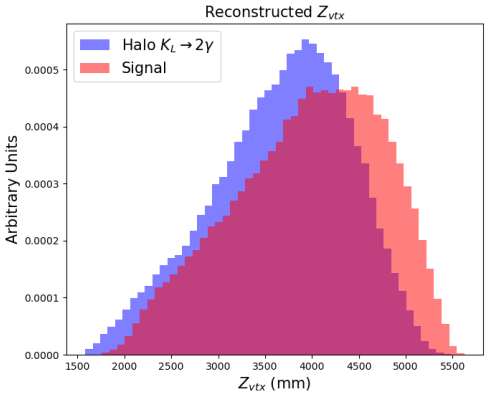
(e) The pion transverse momentum comparison between Halo $K_L \to 2\gamma$ and $K_L \to \pi^0\nu\overline{\nu}$ signal.



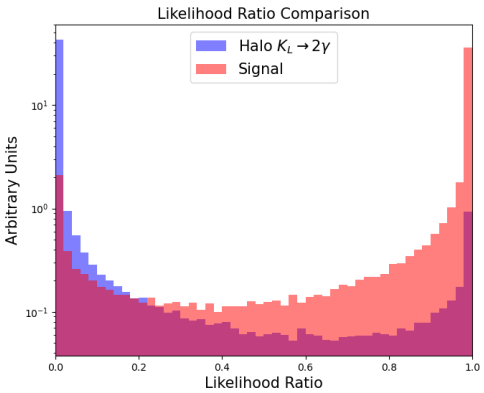
(f) The reconstructed vertex comparison between Halo $K_L \to 2\gamma$ and $K_L \to \pi^0\nu\overline{\nu}$ signal.



(g) The projection angle comparison between Halo $K_L \to 2\gamma$ and $K_L \to \pi^0\nu\overline{\nu}$ signal.



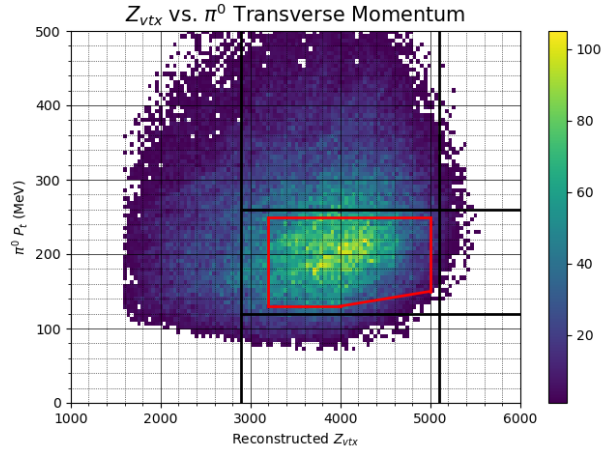
(h) The cluster distance comparison between Halo $K_L \to 2\gamma$ and $K_L \to \pi^0\nu\overline{\nu}$ signal.



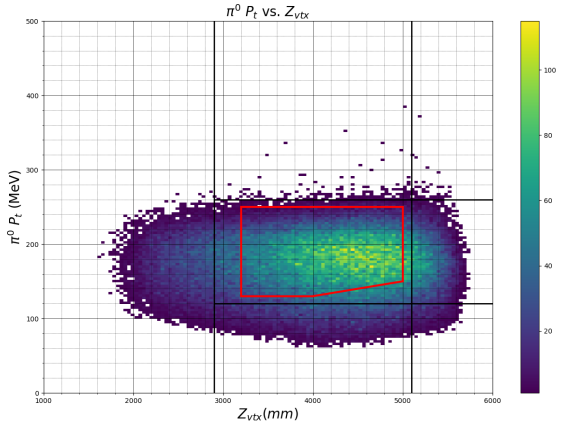
(i) The likelihood ratio comparison between Halo $K_L \to 2\gamma$ and $K_L \to \pi^0\nu\overline{\nu}$ signal.

Figure 8.13: Continued

A comparison of the 2 dimensional plots of Z_{vtx} vs. π^0 p_T is shown in Fig. 8.14. The same model hyperparameters and inputs to the model were selected as previously, except for the added likelihood variable. The normalization and training strategies were also the same. The result of the training is shown in Fig. 8.15.



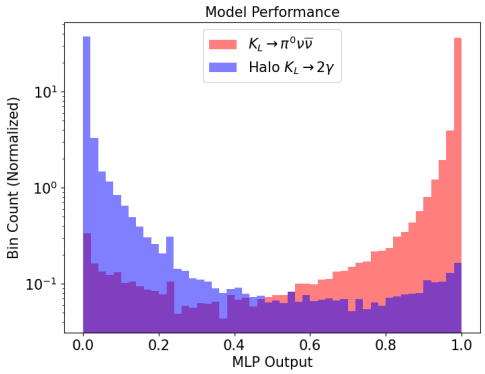
(a) The P_t vs. Z distribution for Halo $K_L \to 2\gamma$ background Monte Carlo. The signal region is in red.



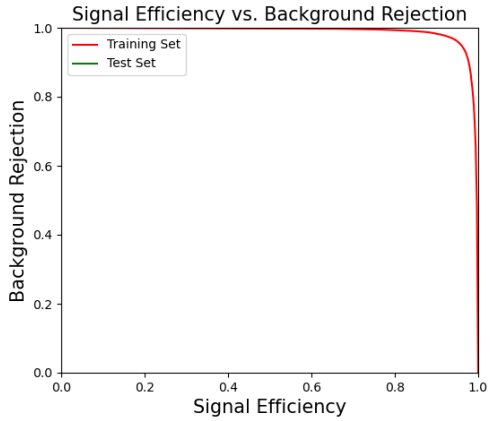
(b) The P_t vs. Z distribution for the $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo. The signal region is in red.

Figure 8.14: The two distributions for both halo $K_L \to 2\gamma$ background and $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo.

This corresponds to around a 98.5% background rejection in the signal region with a signal efficiency of 90%. The MLP is a bit more than two times more effective than the



(a) The model predictions for the training set, with values close to one being "signal-like" and values close to zero being "background-like".



(b) The background rejection vs. signal efficiency in the signal region at different cut neural network cut values. The performance is compared against the independent test set.

Figure 8.15: The model performance on Halo $K_L \to 2\gamma$ background and signal Monte Carlo.

likelihood ratio and TMVA strategy alone. Using these analyses and detector strategies, the KOTO group was able to significantly decrease the background estimation during the analysis of the 2021 data. In particular, the $K_L \to 2\pi^0$ and Upstream π^0 background become the dominant sources of background in this analysis. The previously troublesome background, $K^{\pm}$ and Halo $K_L \to 2\gamma$, are now relatively small, with additional strategies to decrease them further in the future.

8.3 Background Analysis: 2021

The estimated background for the analysis of 2021 data is shown in Table 8.4. As mentioned in the previous section, the Upstream π^0 and $K_L \to 2\pi^0$ background are now the largest. The background discrimination strategies I have applied to the background in the previous section have also been implemented on the $K_L \to 2\pi^0$ to reduce the background level.

8.3.1 $K_L \to 2\pi^0$ Background

This background came to the forefront after updating the GEANT4 version used in the simulation. The main difference was that the probability of photo-nuclear effect occurring increased in the new models. AS a result, the background level was estimated to be around 0.09. Before going into the details on the background rejection, the background cause is explained.

The $K_L \to 2\pi^0$ is a fundamental background source in this experiment. Its level is fundamentally tied to the veto inefficiency of the veto detectors. For instance, if the two photons from one of the pions is incident on the beam hole or barrel detectors, but not measured, the background is extremely similar to signal. An example of the type of background observed after simulation is shown in Fig. 8.16.

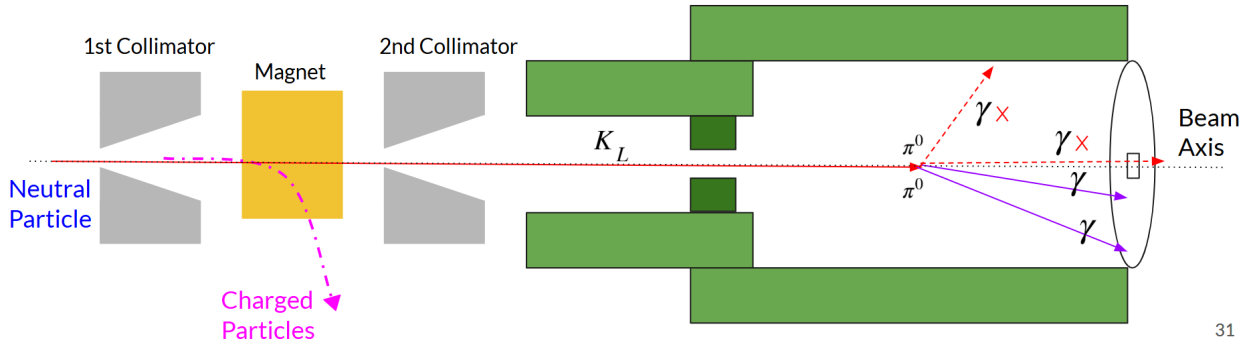


Figure 8.16: A potential background event from $K_L \to 2\pi^0$. Sizes are not to scale. The red X near the γ represents a "missed" γ.

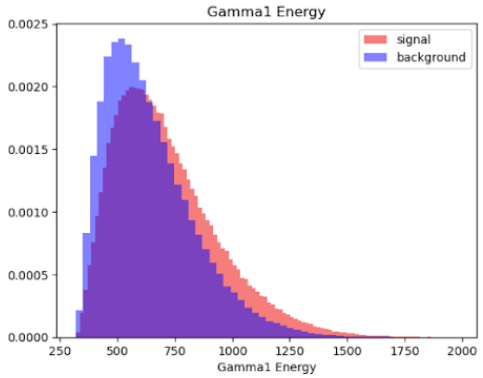
In total, the background can be separated into three main types, shown below in Table 8.3. The dominant background type occurs when both photons are missed after hitting the barrel vetoes. The type that accounts for the smallest percentage is the "beam hole" type. An event is labeled this way if at least one of the photons is missing down the beam hole, typically a high energy photon. Fusion type events account for $\approx 1/3$ of the total background source. These occur when two photons hit the CSI so close together that their respective

electromagnetic showers are fully or partially overlapped. This is particularly troublesome when it is impossible to tell if the cluster contains one or two photon hits, leading to a misreconstructed event. If it is paired with a low energy photon missed in the barrel vetoes, then it becomes a background source.

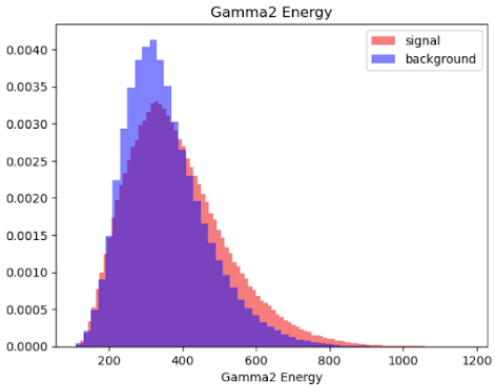
Type	N_{obs}
Fusion	6
Barrel	11
Beam hole	3

Table 8.3: This table contains the number and type of $K_L \to 2\pi^0$ background observed in simulation.

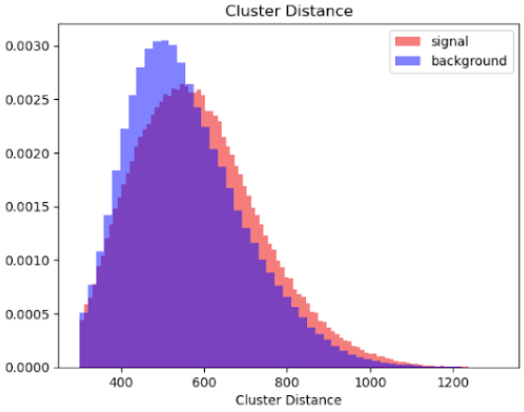
This corresponds to multiple months of $K_L \to 2\pi^0$ background simulation done in KEKCC as well as with the Open Science Grid (OSG). To generate a dataset large enough to train an MLP, selective strategies must be done to narrow down the number of events to simulate. Thus, fast simulation is done to observe the true hit positions and energies of each of the photons, and only events with a high likelihood of becoming background are chosen. For example, an interesting observation was that for the barrel and beam hole type background, the energy ratio of the two missing photons was large, typically over 50. Knowing the true hit positions of each photon, the energy ratio was also checked. If it was larger than 50 and had two photons incident in the barrel or beam hole, with one of the photons being low energy (<50 MeV), this event was selected as a potential background on which to be trained. Additionally, fusion events were selected by taking events where the true hit position of two photons was within 5 cm of each other, and the other photon was low energy and incident on the barrel vetoes. The ratio of these types of events in the training sample was controlled such that it matched the ratio observed in the full simulation data. The comparison between data and Monte Carlo is shown in Fig. 8.17.



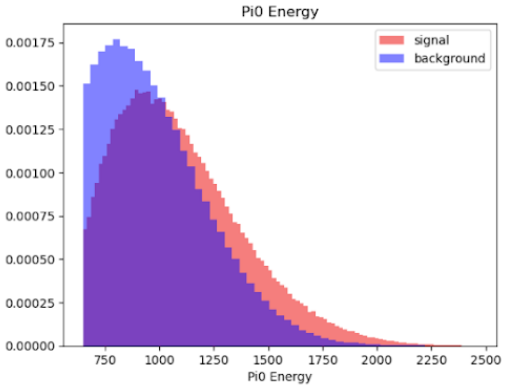
(a) The highest energy photon comparison between $K_L \to 2\pi^0$ and $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo.



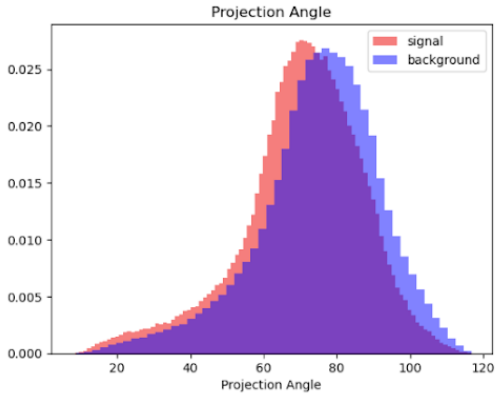
(b) The lowest energy photon comparison between $K_L \to 2\pi^0$ and $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo.



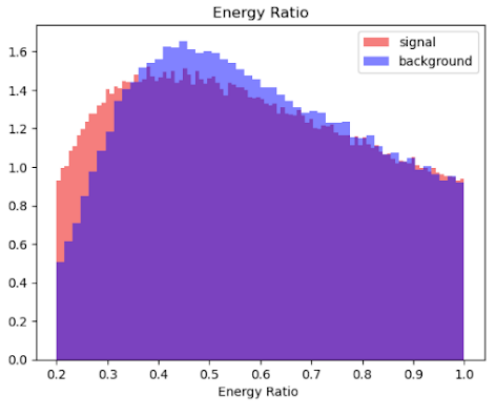
(c) The cluster distance comparison between $K_L \to 2\pi^0$ and $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo.



(d) The pion energy comparison between $K_L \to 2\pi^0$ and $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo.

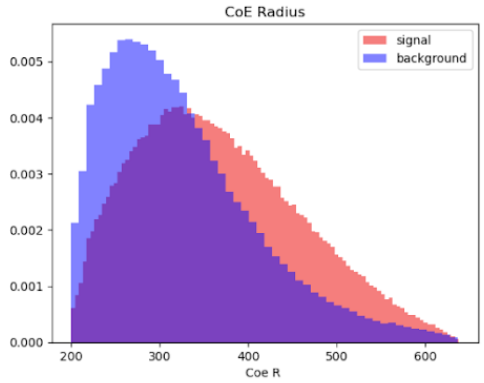


(e) The projection angle comparison between $K_L \to 2\pi^0$ and $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo.



(f) The photon energy ratio comparison between $K_L \to 2\pi^0$ and $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo.

Figure 8.17: The kinematic distributions for $K_L \to 2\pi^0$ background and signal Monte Carlo.



(g) The center of energy radius comparison between $K_L \to 2\pi^0$ and $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo.

Figure 8.17: Continued

The model training followed the previous analyses. The most important variables were obtained by evaluating the average impact on the model output, and are shown in Fig. 8.18.

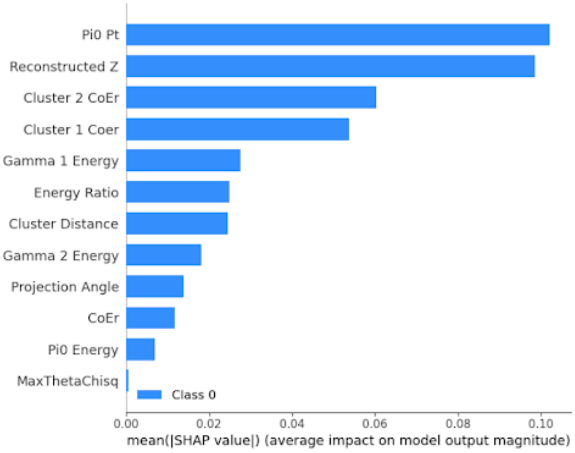
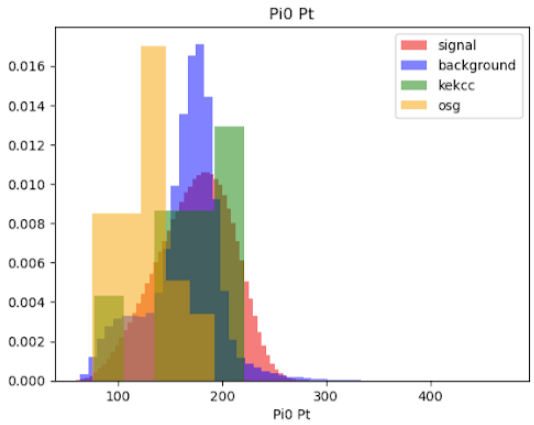


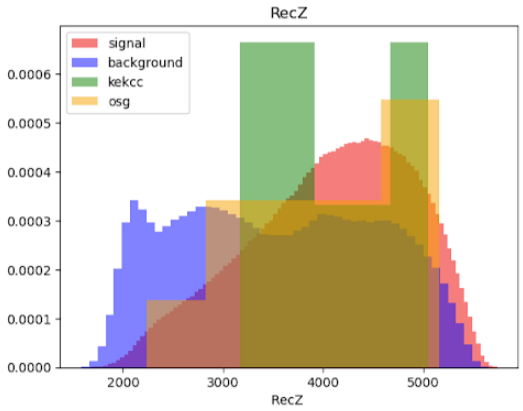
Figure 8.18: The input parameters importance in the model output.

By far, the two most important parameters are π^0 transverse momentum and the reconstructed Z_{vtx}, followed closely by the center of energy radius for the two clusters. These variables are compared between the background Monte Carlo data used for the training set, the signal Monte Carlo, and the full simulation result using OSG and KEKCC in Fig. 8.19.

Overall, the events leftover from the OSG full simulation tended to have a smaller π^0 transverse momentum than those observed at KEKCC, but low statistics likely play a role.



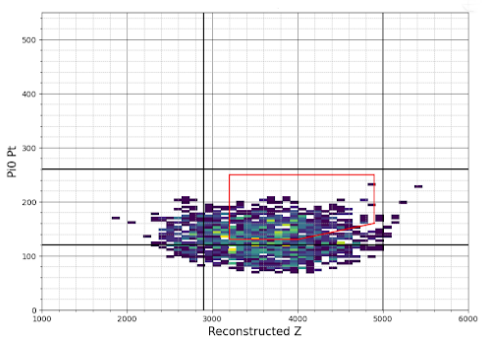
(a) The pion transverse momentum comparison.



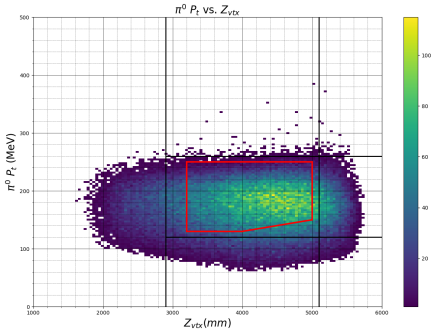
(b) The reconstructed Z_{vtx} comparison.

Figure 8.19: The comparison between signal and background Monte Carlos for the two most important input variables. These are compared to the leftover events from the full simulation Monte Carlo done in KEKCC and OSG.

The 2 dimensional plots for the signal and background are shown in Fig. 8.20.



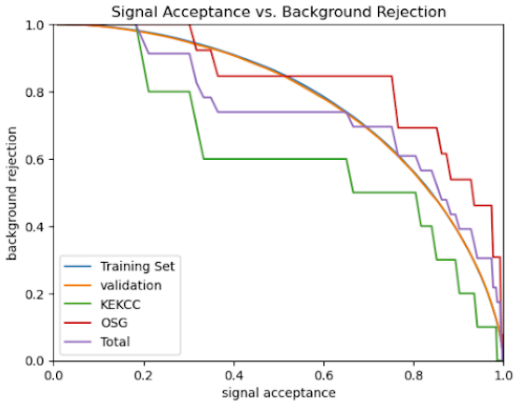
(a) The P_t vs. Z distribution for $K_L \to 2\pi^0$ background Monte Carlo used for the training.



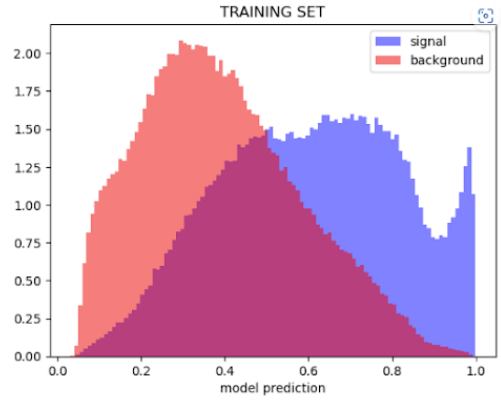
(b) The P_t vs. Z distribution for the $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo.

Figure 8.20: The two distributions for both $K_L \to 2\pi^0$ background and $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo.

The model performance is shown in Fig. 8.21, and was tested against the full simulation Monte Carlo done in KEKCC and OSG, although the statistics are low. Due to the fact the OSG Monte Carlo had remaining events that tended to have lower π^0 transverse momentum, the rejection was better, whereas the leftover events from the KEKCC Monte Carlo tended to have higher transverse momentum with poor rejection. By combining the two there was general agreement with the model performance of the training set. Note the interesting distribution for the neural network output for the signal, which has two peaks. This distribution likely resulted in part from the high transverse momentum events for the signal Monte Carlo, which are above the threshold for the highest transverse momentum observed in the background set. Thus, the model easily gives them a high score while struggling to discriminate signal with low transverse momentum.



(a) The signal efficiency vs. background rejection for the training, validation and test sets. It is compared to the low statistic full simulation results from KEKCC and OSG respectively.



(b) The model prediction distribution for the $K_L \to 2\pi^0$ background and $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo used in the training.

Figure 8.21: The model performance for the $K_L \to 2\pi^0$ background and $K_L \to \pi^0\nu\overline{\nu}$ signal Monte Carlo.

As the background level is still at a manageable level, a loose cut with 90% signal acceptance was applied in the 2021 data analysis with an estimated background rejection of 40%.

8.3.2 *Upstream π^0*

Currently, largest background in the analysis of 2021 data is the upstream π^0 background. This background is induced when there is an interaction in the NCC, by a beam halo neutron, which generates a π^0. If this π^0 promptly decays into two photons which are incident on the CSI, and the vertex is misreconstructed into the signal region, it may become a background source. In fact, many such events are observed upstream in the 2016-2018 data seen in Fig 8.2. This distribution of events has a tail that may leak into the signal region if the photon has an energy mis-measurement. Typically, this may be caused by the photo-nuclear effect in the CSI, the study of which is one motivation of the $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\pi^0 X, X \to \gamma\gamma$ analysis. An example of the nature of the background is shown in Fig. 8.22.

8.3.3 *Background Table: Summary*

The final summary of all backgrounds with a central value in the 2021 data analysis are shown in Table 8.4. The systematic and statistical errors are included. As for the background not discussed so far in this dissertation, further details may be found in Appendix A

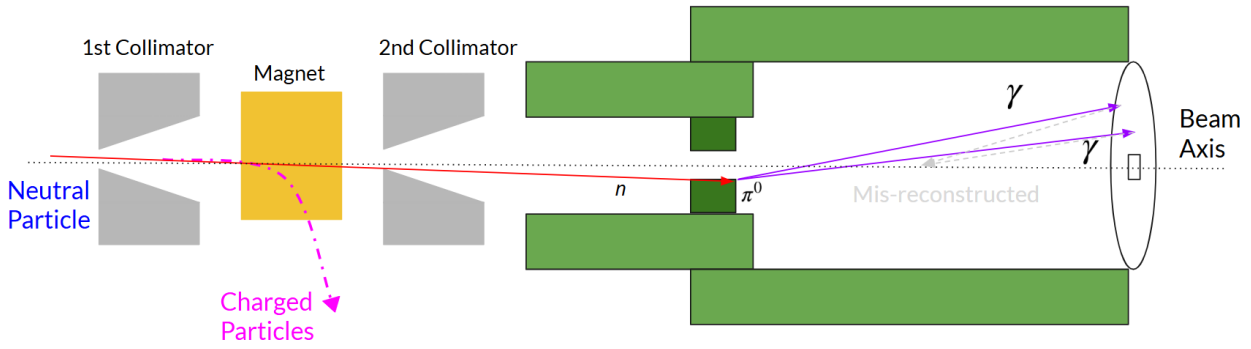


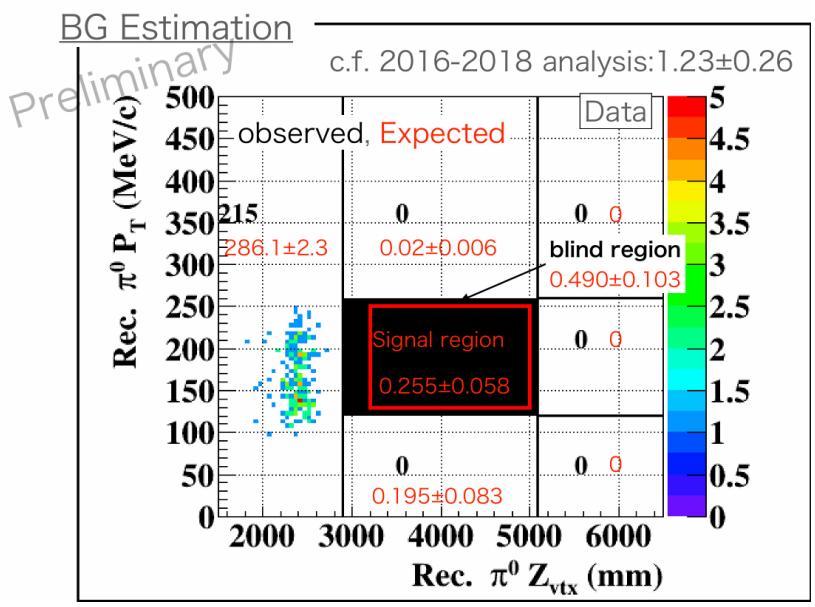
Figure 8.22: A potential background event from an upstream π^0 generated from an interaction in the NCC. Sizes are not to scale.

Source	Background Level
Upstream π^0	$0.060 \pm 0.046_{\text{stat}} \pm 0.007_{\text{sys}}$
$K_L \to 2\pi^0$	$0.059 \pm 0.022_{\text{stat}}{}^{+0.051}_{-0.060}{}_{\text{sys}}$
Scattered + Halo $K_L \to 2\gamma$	$0.045 \pm 0.010_{\text{sys}} \pm 0.006_{\text{sys}}$
$K^{\pm}$	$0.042 \pm 0.014_{\text{stat}}{}^{+0.004}_{-0.029}{}_{\text{sys}}$
Hadronic Cluster	$0.024 \pm 0.004_{\text{stat}} \pm 0.006_{\text{sys}}$
CV-η	$0.023 \pm 0.010_{\text{stat}} \pm 0.005_{\text{sys}}$
TOTAL	$0.253 \pm 0.055_{\text{stat}}{}^{+0.052}_{-0.067}{}_{\text{sys}}$

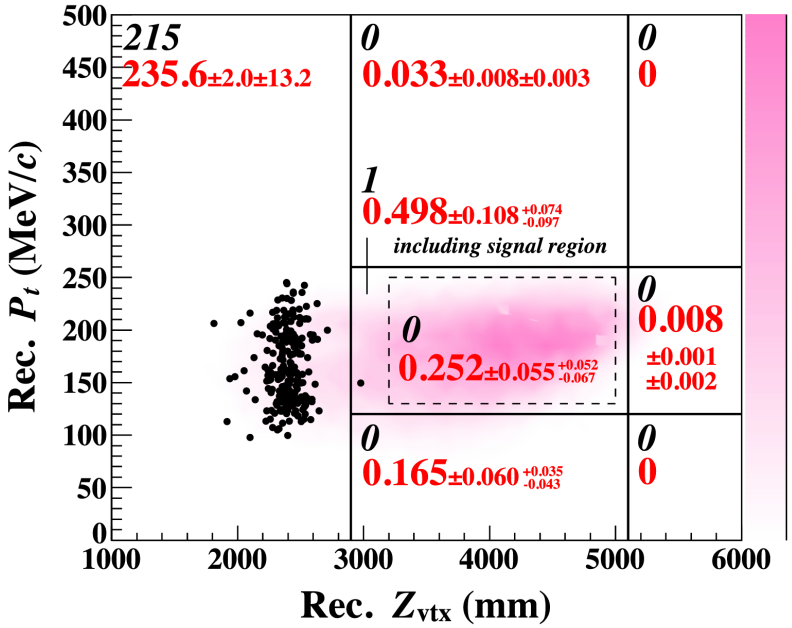
Table 8.4: This table contains the background estimation in the 2021 signal region. The ordering is from largest to smallest, and backgrounds limited to upper limits are not included [4].

8.3.4 Results

The preliminary results before opening the signal region, and the final results after further study of the NCC region and the open box result are shown in Fig. 8.23. The final cut conditions and cut descriptions are explained in detail in Appendix A. No events were observed in the signal region, which allowed KOTO to set the world best upper limit on the branching ratio of $K_L \to \pi^0\nu\overline{\nu}$. In addition, the total background expectation was a factor of five less than the result from data taken in 2016-2018 through the use of improved analysis techniques and new detector hardware. Using Poisson statistics and accounting for the uncertainty in the SES calculation, the results are shown in Fig. 8.23 [17] [4].



(a) The preliminary result of the 2021 data analysis situation before determining to open the box. Plot generated by Koji Shiomi in an internal meeting.



(b) The open box result for the $K_L \to \pi^0\nu\bar{\nu}$ analysis of 2021 data [4].

Figure 8.23: The final results of the 2021 data analysis. Red numbers indicate the expectation in that region from background Monte Carlo simulation. Black numbers represent the number observed in data.

$$\mathrm{BR}(K_L \to \pi^0 \nu\overline{\nu}) < 2.2 \times 10^{-9} \text{ (90\% C.L.)} \tag{8.4}$$

Furthermore, the study of $K_L \to \pi^0 X^0$ where X^0 is invisible is deeply tied to this result. By doing a Monte Carlo for a range of X^0 masses, and evaluating the signal efficiency based on the final cut conditions, an upper limit can be found for each X^0 mass. The result is shown in Fig. 8.24 [4].

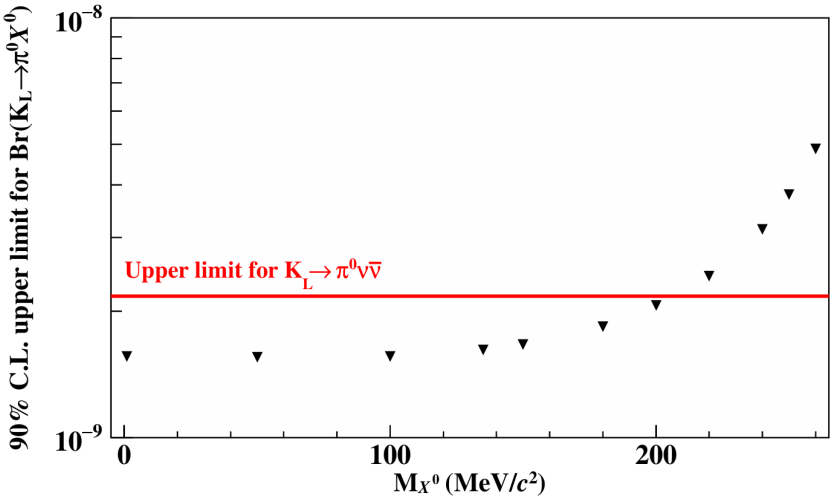


Figure 8.24: The final upper limit at 90% C.L. results for the analysis of $K_L \to \pi^0 X^0$ where X^0 is invisible. A mass range of 1-260 MeV is studied.

CHAPTER 9

CONCLUSION AND DISCUSSION

In this dissertation, I described the KOTO experiment at J-PARC which aims to study CP-symmetry breaking in kaon decays. In particular, the analysis of the rare decay, $K_L \to \pi^0\nu\overline{\nu}$, using 2021 data had no events observed in the signal region, which allowed KOTO to set the best upper limit on BR($K_L \to \pi^0\nu\overline{\nu}$) to date at $< 2.2 \times 10^{-9}$ (90% C.L.). Additionally, a novel analysis was performed on 2021 data on the modes $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\pi^0 X$ where $X \to \gamma\gamma$. Due to the improved resolution and analysis techniques from the study in E391A, the scope is increased to include X mass in the range 160-220 MeV, and the sensitivity is improved by close to an order of magnitude in some mass ranges. The analysis includes the first search for the rare decay $K_L \to \pi^0\pi^0\gamma\gamma$, which is described by ChPT. Two events were observed in the signal region near an X mass of 177 MeV, and no events were observed for the analysis of $K_L \to \pi^0\pi^0\gamma\gamma$, setting a world first upper limit of BR($K_L \to \pi^0\pi^0\gamma\gamma$) $< 9.77 \times 10^{-7}$ at 90% C.L.

9.1 Interpretation of the Results

The interpretation of the results for $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\nu\overline{\nu}$ are straightforward, as there are no events observed in the signal regions, which allows a clean upper limit on the branching ratio. For the analysis of $K_L \to \pi^0\pi^0 X$, the situation is more complicated due to the nature of the analysis strategy. Each X mass analysis is treated entirely separately and independently of others, however the results are not independent. There may be events which occur in two separate signal regions even though cut conditions are different, due to the overlap of the signal regions. This is shown in the modified results figure shown in Fig. 9.1. Of note, there is one event which is observed in the $M_X = 175$ MeV and $M_X = 180$ MeV analysis.

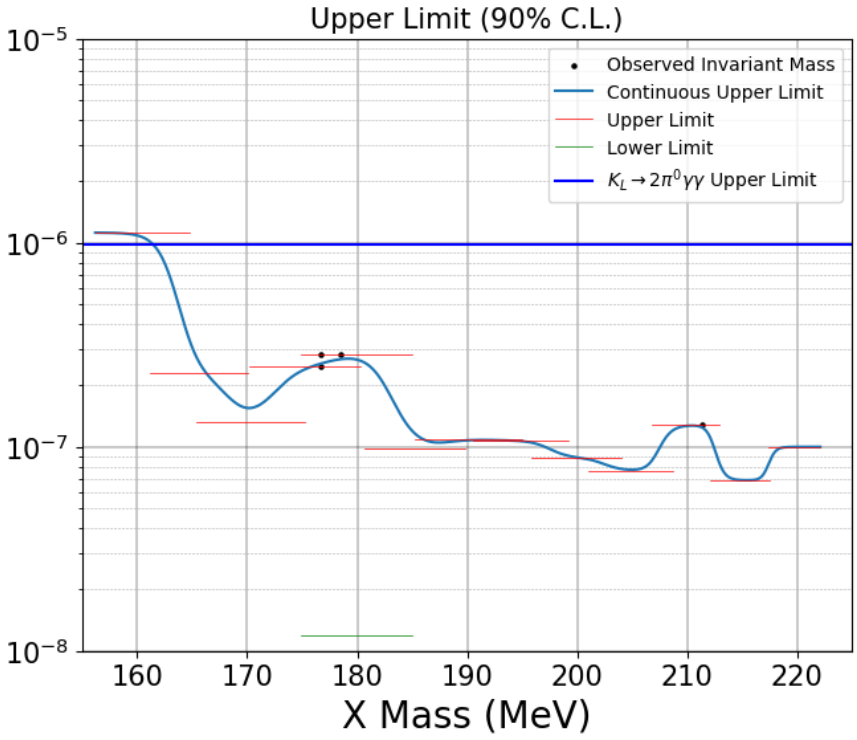


Figure 9.1: The final upper limit at 90% C.L. results for the analysis of $K_L \to \pi^0\pi^0 X$. The bars represent the respective signal regions for each X mass analysis, and the black dots show the invariant mass of events observed within their signal region. The continuous line represents the weighted upper limit calculation.

As discussed in Sec. 7.5, the benefits of this analysis strategy allow for the improvement in sensitivity throughout the mass region search. This is particularly important as the signal and background levels are low in these analyses. However, it makes the interpretation of the result difficult when there is event and signal region overlap. One way to overcome this issue of interpretation is to realize that each analysis impacts nearby analyses, and evaluate a weighted upper limit. For instance, a hypothetical $K_L \to \pi^0\pi^0 X$ decay, with $M_X = 180$ MeV, may have an imperfectly measured reconstructed invariant mass due to the energy resolution and nature of the cut conditions. Thus, the analysis result assuming an X mass of 175 MeV is not independent of the result assuming an X mass of 180 MeV. The amount of impact depends on the mass resolution, and in particular the gaussian fits obtained in Fig. 7.10. The weighted upper limits at test masses x_m, are calculated by the formula shown in

Eq. 9.1, where each X mass analysis is indexed by i.

$$UL(x_m) = \frac{\sum_i UL_i \cdot \mathrm{Gaus}_i(x_m)}{\sum_i \mathrm{Gaus}_i(x_m)} \tag{9.1}$$

UL_i and Gaus_i represent the upper limits and gaussian fit function for each X mass analysis respectively. The mean and and standard deviations for each analysis is shown in Table 7.1. This means that analysis results within 2-3 standard deviations of the result have impact on the upper limit. The weighted upper limits calculated at 50 test masses across the high mass region are shown in Appendix B.

Furthermore, this method can be tested with a toy MC result. Using the 13 analyses already completed, and produce an expected result after many trials. The steps are outlined below.

1. Choose approximately 100 test masses, x_m.

2. Determine the mass width of the signal at the test mass, x_m, via quadratic interpolation from the three nearest analysis results.

3. Generate 50,000 reconstructed masses from the gaussian mass width, and evalated the signal acceptance from quadratic interpolation from the three nearest analysis results.

4. Calculate the total acceptance, and hence the SES.

5. Obtain the Feldman Cousins variables, "mean background" and "number observed". The mean background is determined by quadratic interpolation of the background assumed in Eq. 7.15 for the three nearest analysis results. The number observed is determined by random sampling of a Poisson distribution with mean given by the sum of the mean background plus the height normalized test mass gaussian probability distribution function at the observed masses. The mean number obsered function is shown in Fig. 9.2.

6. Run 200 “experiments” with 200 observations and evaluate the averaged upper limit via Feldman Cousins [23].

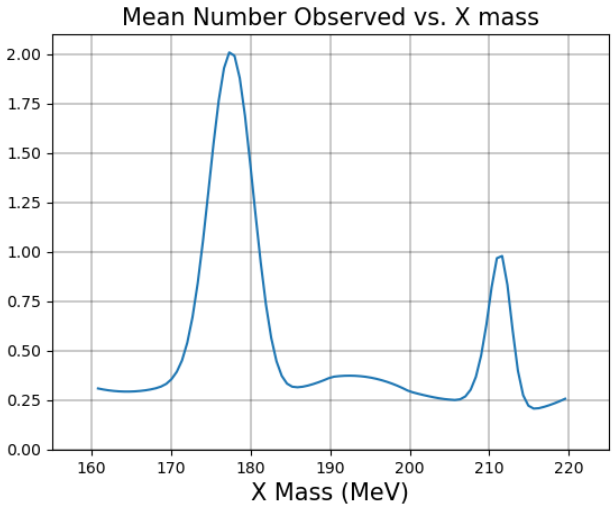


Figure 9.2: The mean number observed assumed for the toy MC study as a function of test mass x_m.

The height is normalized such that the mean number observed is two near the two observed events invariant masses, and one near the observed event at 211 MeV invariant mass. The averaged upper limit result is shown in Fig. 9.3.

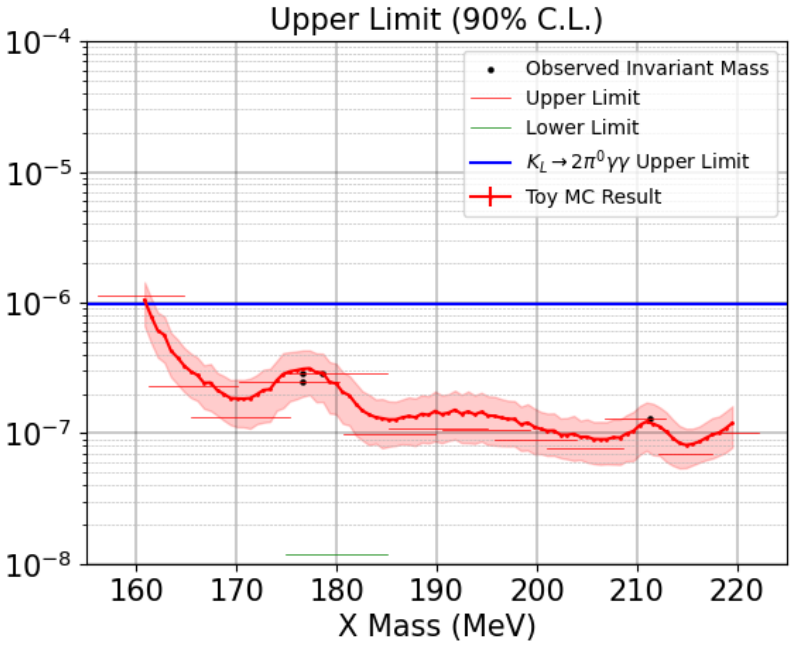


Figure 9.3: The upper limit result overlayed with the toy MC result. The one sigma bound on the averaged upper limit result is shown in the shaded red.

9.2 Potential Improvements

The analysis of $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\pi^0 X$ is severely limited by Monte Carlo statistics of $K_L \to 3\pi^0$ background. The current analysis only has 90% C.L. upper limits on the background in the signal region, which complicates the analysis of the results. Ongoing work includes doubling the statistics of $K_L \to 3\pi^0$ Monte Carlo such that the central value of the background can be obtained. This erases the ambiguity of the Feldman Cousins result which requires a central mean background.

The work of doubling statistics is expected to take around 1.5-2 months of continuous job submissions using the Open Science Grid (OSG) [39] [45] [37] [38]. This work is expected to finish before the publication of this result.

Additionally, the interpretation of the results is still in discussion. More study is ongoing using a frequentist method and a branching ratio scan. This method will not rely on

interpolation, and instead calculate the ratio of probabilities shown in Eq. 9.2.

$$r = \frac{P(n_{s+b} \leq n_{obs}|s+b)}{P(n_b \leq n_{obs}|b)} \tag{9.2}$$

Assuming different branching ratios of the signal, and different values of $s+b$, the signal assumption which results in a ratio closest to 10% gives the best upper limit at 90% C.L. This can be calculated at any test mass x_m, by calculating the average weighted ratio shown in Eq. 9.3

$$r_w = \frac{\sum_i \text{Gaus}_i(x_m) \cdot r_i}{\sum_i \text{Gaus}_i(x_m)} \tag{9.3}$$

The index i sums over the two nearest analysis results, and $\text{Gaus}_i(x_m)$ is the amplitude of the Gaussian fit at the test mass x_m for the analysis i. Thus, the weighted method has results which are most weighted towards the analysis closest to the test mass x_m. The preliminary result using this strategy is shown in Fig. 9.4, and due to low statistics on signal acceptance far from the signal mean, the result is plagued by high statistical uncertainty. The result is consistent with the previous method.

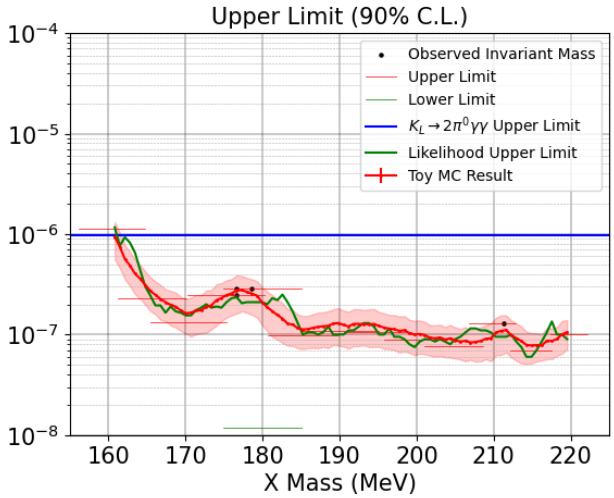


Figure 9.4: The likelihood upper limit result overlaid with the toy MC result.

APPENDIX A

LIST OF CUT CONDITIONS

A.1 Veto

The veto cut conditions used for the analysis of $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\pi^0 X$ are shown in the following table. In addition to those shown in A.1, there is also a "CSI Veto" to reject events with on-time single crystal hits. The timing window for these on-time hits is ± 10 ns. Typically, these are caused by real photons from the decay that hit the very edge of the CSI, but don't become a cluster under the analysis definition (and so pass the MaxFiducialR cut condition) or, photonuclear type interactions from a real hit that keeps the electromagnetic shower from developing. The veto cut conditions for these types of hits in the CSI depend on the distance from the closest cluster, and the energy deposit in the isolated crystal hit. The details are shown in Eq. A.1.

$$\begin{aligned} &E_{dep} > 10 \text{ MeV}, && \text{if } D < 200 \text{ mm} \\ &E_{dep} > \left(10 - \frac{7(D-200)}{400}\right) \text{ MeV}, && \text{if } 200 \le D \le 400 \text{ mm} \\ &E_{dep} > 3 \text{ MeV}, && \text{if } D > 400 \text{ mm} \end{aligned} \tag{A.1}$$

Furthermore, in the two cluster analysis of $K_L \to \pi^0\nu\overline{\nu}$, the veto conditions are further modified from Table A.1 in the case that masking may have occured. Three detectors CV, NCC, and FBAR are all susceptible to large acceptance loss due to accidental hits. This is due to their large volumes and wide timing windows to protect against masked pulses. However, in the offline analysis the probability of masking is checked by observing the pulse shape after applying the Fourier transform to the waveform. If this pulse is "masked-like" then the wide window is used instead of the window width shown in Table A.1. Otherwise,

Detector Name	Energy Threshold (MeV)	Timing Window Width (ns)
UCV	0.05	15
UCVLG	0.05	30
CC03	3	30
CC04	3	30
CC05	3	30
CC06	3	30
CBAR	2	60
FBAR	2	50
BHPV	2.5	15
CV	0.2	20
BCV	1	60
NCC	2	40
OEV	2	20
LCV	0.6	30
BPCV	1	24
BHCV	221×10^{-6}	25
BHGC	2.5	15
IB	2	60
IBCV	1	60
MBCV	1	60
DCV	0.2	60

Table A.1: List of Veto Detectors and their corresponding energy thresholds and timing window widths. Timing windows are set to prevent masked pulses being missed, consider possible time of flight hits, and the detector timing resolutions.

a shorter timing window is used to minimize accidental loss. The modified veto conditions are listed in Table A.2, including changes to IB and CBAR veto windows.

A.2 Kinematic

The kinematic cut conditions included in Table 6.2 were used for the analysis of $K_L \to \pi^0\pi^0\gamma\gamma$ and $K_L \to \pi^0\pi^0 X$. However, additional cuts were used in the analysis of $K_L \to \pi^0\nu\overline{\nu}$. These are discussed in the list below.

- **Min(ClusterRMS) $>$ 10:** The cluster root mean square (ClusterRMS) is a cluster shape cut, which encapsulates the "size" of the cluster. This cut is used to reject

Detector Name	Original Timing Window	Modified Timing Window (ns)
CV	80	20 (150 if FTT $\chi^2 > 12$)
FBAR	50	50 (170 if FTT $\chi^2 > 20$)
CBAR	60	50 (new timing algo.)
IB	60	45 ((new timing algo.)
NCC	40	40 (100 if FTT $\chi^2 > 40$)

Table A.2: The modified veto conditions used in the $K_L \to \pi^0 \nu\overline{\nu}$ analysis

small clusters which tend to be caused by hadronic interactions. The calculation of the ClusterRMS is shown in Eq. A.2, where the index i goes through the individual crystals within the cluster.

$$\text{ClusterRMS} = \sum_i \frac{E_i \sqrt{(x_i - x_{pos})^2 + (y_i - y_{pos})^2}}{E_{tot}} \tag{A.2}$$

- **CSICoER $>$ 200 mm:** The radius of the center of energy in the CSI must be larger than 200 mm. Background from fusion type events, such as those in $K_L \to 2\pi^0$ background, tend to have smaller center of energy radii since less off-axis energy is missing.

- **Min(GamClusterSize) $>$ 5:** Similar to the ClusterRMS cut, this is a cut based on the cluster size, and requires the minimum number of crystals in a cluster to be 5.

- **MaxGammaE $<$ 2000 MeV:** The maximum reconstructed gamma energy should be less than 2000 MeV. This is to reduce fusion and accidental type backgrounds.

- **Energy Ratio $<$ 5:** The ratio of the higher energy photon to the lower energy photon must be less than 5. This is to reduce the odd pairing or fusion background from $K_L \to 2\pi^0$, since the two photons from the same π^0 are unlikely to have such a high energy ratio.

- **Min($E \cdot \theta$) $>$ 2.5 GeV:** This cut is also performed to reduce the background caused

from $K_L \to 2\pi^0$, since mispairings lead to incorrect Z_{vtx} which would lead to the wrong incident angle. For $K_L \to 2\pi^0$ background, the distribution tends to be smaller than $K_L \to \pi^0\nu\overline{\nu}$ signal.

- π^0 **Kinematic:** This is a two dimensional cut based on the ratio π^0 transverse momentum to its beam direction momentum, and the reconstructed Z_{vtx}. This is important for the rejection of $K_L \to 2\pi^0$ background caused by mis-pairings or fusion. In general, the false π^0 tends to have a small ratio, P_T/P_Z, with a dependence on the reconstructed Z_{vtx}. The conditions are shown in Eq. A.4.

$$\begin{aligned} &P_T/P_Z > 0.1, && \text{if } Z_{vtx} < 4 \text{ m} \\ &P_T/P_Z > 0.1 + 0.05 \cdot (Z_{vtx} - 4), && \text{if } 4 \leq Z_{vtx} \leq 5 \text{ m} \\ &P_T/P_Z < 0.2 + 0.07 \cdot (Z_{vtx} - 3), && \text{if } 3 \leq Z_{vtx} \leq 5 \text{ m} \end{aligned} \tag{A.3}$$

(A.4)

- **Projection Angle:** The projection angle is defined as the angle between the two clusters on the CSI projected onto the CSI plane. It is required to be smaller than 150° to remove $K_L \to 2\gamma$ background. Since the photons are back to back, $K_L \to 2\gamma$ decays with small K_L transverse momentum should have a projection angle close to 180°.

A.3 Additional

The following additional cuts were applied to the $K_L \to \pi^0\nu\overline{\nu}$ analysis, but not the analyis of $K_L \to \pi^0\pi^0 X$ and $K_L \to \pi^0\pi^0\gamma\gamma$. The $K_L \to 2\pi^0$ NN cut is described in Section 8.3.1, and the Halo $K_L \to 2\gamma$ cuts were described in Section 8.2.2.

- **EtaCSD**: The Eta Cluster Shape Discriminator (EtaCSD) is a convolutional neural network model training on background caused by $\eta \to 2\gamma$ decays which may occur due to accidental interactions in the CV. If the two photons are mis-reconstructed into the decay region under the pion mass and beam line decay assumptions, then it may look like a signal decay.

- **Front End Timing**: MPPC readout installed on the front end of the CSI allow for the timing of a cluster to be evaluated from both sides of the CSI. Due to time of flight, neutron interactions which tend to occur in the middle or back end of the CSI, will have a different relative timing difference than electromagnetic clusters which begin at the front of the CSI. By requiring the timing difference between the rear and front timing to be large, neutron type background may be reduced.

APPENDIX B

TABLE OF $K_L \to \pi^0\pi^0 X$ RESULTS

Mass (MeV)	Upper Limit Result (90% C.L.)
156.18	1.12e-06
157.5	1.12e-06
158.82	1.11e-06
160.14	1.08e-06
161.46	9.78e-07
162.79	7.27e-07
164.11	4.29e-07
165.43	2.72e-07
166.75	2.14e-07
168.07	1.81e-07
169.39	1.59e-07
170.71	1.57e-07
172.03	1.75e-07
173.36	2.05e-07
174.68	2.31e-07
176.0	2.49e-07
177.32	2.61e-07
178.64	2.69e-07
179.96	2.68e-07
181.28	2.45e-07
182.6	1.96e-07
183.93	1.45e-07
185.25	1.16e-07
186.57	1.06e-07
187.89	1.05e-07
189.21	1.07e-07
190.53	1.08e-07
191.85	1.08e-07
193.17	1.08e-07
194.5	1.07e-07
195.82	1.05e-07
197.14	9.96e-08
198.46	9.29e-08
199.78	8.91e-08
201.1	8.67e-08
202.42	8.25e-08
203.74	7.85e-08
205.07	7.73e-08
206.39	8.11e-08
207.71	1.04e-07
209.03	1.23e-07
210.35	1.27e-07
211.67	1.21e-07
212.99	8.2e-08
214.31	6.93e-08
215.64	6.86e-08
216.96	7.35e-08
218.28	9.73e-08
219.6	9.99e-08
220.92	1e-07

Table B.1: The result of the weighted upper limit calculation for 50 test masses across the high mass region.

REFERENCES

[1] Search for $K_L \to \pi^0\nu\overline{\nu}$ and $K_L \to \pi^0 X^0$ decays at the j-parc koto experiment. *Phys. Rev. Lett.*, 122:021802, Jan 2019.

[2] First search for $K_L \to \pi^0\gamma$. *Phys. Rev. D*, 102:051103, Sep 2020.

[3] CLUE: a clustering algorithm for current and future experiments. 2022.

[4] Search for the $K_L \to \pi^0\nu\overline{\nu}$ decay at the j-parc koto experiment. *Phys. Rev. Lett.*, 134:081802, Feb 2025.

[5] E. Abouzaid, M. Arenton, and et al. Search for the rare decays $K_L \to \pi^0\pi^0\mu^+\mu^-$ and $K_L \to \pi^0\pi^0 X^0 \to \pi^0\pi^0\mu^+\mu^-$. *Phys. Rev. Lett.*, 107:201803, Nov 2011.

[6] E. Abouzaid et al. Final Results from the KTeV Experiment on the Decay $K_L \to \pi^0\gamma\gamma$. *Phys. Rev. D*, 77:112004, 2008.

[7] S. Agostinelli, J. Allison, and et al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[8] A. Alavi-Harati, T. Alexopoulos, and et al. Search for the $K_L \to \pi^0\pi^0 e^+e^-$ decay in the ktev experiment. *Phys. Rev. Lett.*, 89:211801, Oct 2002.

[9] J. Allison, K. Amako, and et al. Recent developments in geant4. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 835:186–225, 2016.

[10] Reuven Balkin, Stefania Gori, and Christiane Scherb. Probing cp and flavor violation in neutral kaon decays with alps, 2025.

[11] R. Brown, U. Camerini, P.H. Fowler, H. Muirhead, and C.F. Powell. Observations with electron sensitive plates exposed to cosmic radiation. *Nature*, 163:82, 1949.

[12] Gerhard Buchalla and Gino Isidori. The cp-conserving contribution to $k_l \to \pi^0 \nu\overline{\nu}$ in the standard model. *Physics Letters B*, 440(1):170–178, 1998.

[13] Andrzej J. Buras and Elena Venturini. Searching for new physics in rare k and b decays without $|v_{cb}|$ and $|v_{ub}|$ uncertainties, 2022.

[14] J. H. Christenson, J. W. Cronin, V. L. Fitch, and et al. Evidence for the 2pi decay of the k02 meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[15] J-PARC KOTO Collaboration. A new search for the kl→0^{-} and kl→0x0 decays. *Progress of Theoretical and Experimental Physics*, 2017(2):021C01, 02 2017.

[16] E. Cortina Gil and et al. Measurement of the k+→+ decay. *Physics Letters B*, 850:138513, 2024.

[17] Robert D. Cousins and Virgil L. Highland. Incorporating systematic uncertainties into an upper limit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 320(1):331–335, 1992.

[18] Robert P. Crease. *Abraham Pais 1918-2000*. National Academy of Sciences, 2011.

[19] R.H. Dalitz. Cxii. on the analysis of -meson data and the nature of the -meson. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(357):1068–1080, 1953.

[20] R.H. Dalitz. The $\tau - \theta$ puzzle. *AIP Conf.Proc.*, 300:141–158, 1994.

[21] A I Davydychev and R Delbourgo. Explicitly symmetrical treatment of three-body phase space. *Journal of Physics A: Mathematical and General*, 37(17):4871, apr 2004.

[22] Gerhard Ecker. Status of chiral perturbation theory. *Frascati Phys. Ser.*, 7:347–362, 1997.

[23] Gary J. Feldman and Robert D. Cousins. Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D*, 57:3873–3889, Apr 1998.

[24] R. Funck and J. Kambor. The decays k(l,s) —> pi0 pi0 gamma gamma and k(l) —> pi0 pi0 gamma* in the effective chiral lagrangian approach. *Nucl.Phys.B*, 396:53–80, 1993.

[25] M. Gell-Mann and A. Pais. Behavior of neutral particles under charge conjugation. *Phys. Rev*, 97(1387), 1955.

[26] Elor Gilly, Escudero Miguel, and Ann E. Nelson. Baryogenesis and dark matter from mesons. *Phys. Rev. D*, 99(035031), 2019.

[27] Mario Gonzalez. The Data-Acquisition System and new GPU-based HLT of the KOTO Experiment. *PoS*, ICHEP2024:1009, 2025.

[28] Baldes Iason, Nicole F. Bell, Alexander Millar, Kalliopi Petraki, and Raymond R. Volkas. The role of cp violating scatterings in baryogenesis - case study of the neutron portal. *JCAP*, (1411), 2014.

[29] Lydia Iconomidou-Fayard. Results on CP violation from the NA48 experiment at CERN. *Int. J. Mod. Phys. A*, 17:3012–3025, 2002.

[30] Aravindh Krishnamoorthy and Deepak Menon. Matrix inversion using cholesky decomposition, 2013.

[31] Jason Kumar. Indirect detection of sub-gev dark matter coupling to quarks. *Phys. Rev. D*, 98(116009), 2018.

[32] David Lewis, Gordon Chiu, and et al. The stratix™ 10 highly pipelined fpga architecture. pages 159–168, 02 2016.

[33] Chieh Lin. *Study of KL0 →0 and KL0 →0 with the Cluster-Finding Trigger at KOTO*. PhD thesis, National Taiwan University, 2021.

[34] Laurence S. Littenberg. Cp-violating decay $K_L^0 \to \pi^0 \nu\nu^-$. *Phys. Rev. D*, 39:3322–3324, Jun 1989.

[35] R. Muto. Status and outlook on slow extraction operation at j-parc main ring. In *Proc. IPAC'24*, number 15 in IPAC'24 - 15th International Particle Accelerator Conference, pages 1905–1910. JACoW Publishing, Geneva, Switzerland, 05 2024.

[36] Navas, S., et al. Review of particle physics. *Phys. Rev. D*, 110(3):030001, 2024.

[37] OSG. Ospool, 2006.

[38] OSG. Open science data federation, 2015.

[39] Ruth Pordes, Don Petravick, Bill Kramer, Doug Olson, Miron Livny, Alain Roy, Paul Avery, Kent Blackburn, Torre Wenaus, Frank Würthwein, Ian Foster, Rob Gardner, Mike Wilde, Alan Blatecky, John McGee, and Rob Quick. The open science grid. In *J. Phys. Conf. Ser.*, volume 78 of *78*, page 012057, 2007.

[40] Luigi Rizzo. Netmap: a novel framework for fast packet i/o. In *Proceedings of the 2012 USENIX Conference on Annual Technical Conference*, USENIX ATC'12, page 9, USA, 2012. USENIX Association.

[41] G.D. ROCHESTER and C.C. BUTLER. Evidence for the existence of new unstable elementary particles. *Nature*, 160:855–857, 1947.

[42] A.D. Sakharov. Violation of cp invariance, c asymmetry, and baryon asymmetry of the universe. *Pisma Zh.Eksp.Teor.Fiz.*, 5:32–35, 1967.

[43] K. Sato, J. W. Lee, and et al. Csi calorimeter for the j-parc koto experiment. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 982, dec 2020. Publisher Copyright: © 2020 Elsevier B.V.

[44] Kazufumi Sato. *Measurement of the CsI calorimeter performance and KL momentum spectrum for the J-PARC KOTO experiment.* PhD thesis, Osaka University, 2015.

[45] Igor Sfiligoi, Daniel C Bradley, Burt Holzman, Parag Mhashilkar, Sanjay Padhi, and Frank Wurthwein. The pilot way to grid resources using glideinwms. In *2009 WRI World Congress on Computer Science and Information Engineering*, volume 2 of *2*, pages 428–432, 2009.

[46] Mark Thomson. *Modern Particle Physics.* Cambridge University Press, 2013.

[47] Virginia Trimble. Existence and nature of dark matter in the universe. *Annual Review of Astronomy and Astrophysics*, 25:425–472, 1987.

[48] Y. C. Tung, Y. B. Hsiung, and et al. Search for a light pseudoscalar particle in the decay $K_L^0 \to \pi^0\pi^0 x$. *Phys. Rev. Lett.*, 102:051802, Feb 2009.

[49] Michael S. Turner. Big band baryosynthesis and grand unification. *AIP Conf. Proc.*, 72:224–243, 1981.

[50] Hans Wenzel, Jun Soon Yung, and Krzysztof Genser. Advanced example cats.

[51] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105(1413), 1957.

[52] W. Wah Yau, Taku Yamanaka, Tadashi Nomura, Hajime Nanjo, and et al. Proposal for kl $\to$ 0 experiment at j-parc, 2006.