Search for the Rare Decay of the Neutral Kaon, $K_L^0 \to \pi^0 \nu\bar{\nu}$

by

Melissa A. Hutcheson

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2021

Doctoral Committee:

Professor Myron Campbell, Chair
Professor James Liu
Professor Sally Oey
Assistant Professor Joshua Spitz
Assistant Professor Liuyan Zhao


THE UNIVERSITY OF MICHIGAN
ARTES · SCIENTIA · VERITAS
1817

Melissa A. Hutcheson
melhutch@umich.edu
ORCID iD: 0000-0003-0306-9807

To all of my teachers, instructors, and professors, for encouraging me
and instilling in me a deep curiosity and a passion for learning,
and to Carl Sagan, for inspiring me to ponder the cosmos.

Acknowledgements

This thesis was accomplished only with the tremendous support of many people. Here, I would like to acknowledge their impact and offer my gratitude.

First and foremost, I would like to express my sincere appreciation and thanks to my advisor and mentor, Professor Myron Campbell, who accepted me as an REU student eight years ago and introduced me to KOTO and the world of high energy experimental physics. Since then, Myron has guided me on my journey to graduate school and throughout my graduate career, and in many ways, started me on the path to obtaining this degree. He has helped me grow from an unsure yet enthusiastic student to a confident and experienced researcher, he pushed me to tackle problems that no one had yet solved, and he helped me gain confidence in approaching problems and presenting my work. I appreciate and admire his patience, his overwhelming optimism, his thoughtful clarity, and his realistic approach to academic research and life in general. I will cherish all of the many bike rides in Japan, working on repairing equipment together, and his seemingly endless stream of trivia facts. Thank you, Myron, for everything.

I am deeply thankful for Dr. Monica Tecchio who I have also known now for eight years, and who has been a steady source of DAQ and analysis knowledge, as well as a close friend and colleague. While Monica wasn't my official advisor, she supported me every step of the way and often patiently walked me through the details of the DAQ system or the analysis tasks at hand. It has been a joy to work alongside her for so many years.

I also want to deeply thank Dr. Brian Beckford and Dr. Stephanie Su for their support and friendship throughout my graduate experience. I want to thank Brian for his excellent advice, for his constant support and encouragement, and for not allowing me to give up through all the ups and downs of graduate life. While he provided me with invaluable guidance in analysis and research, he also inspired me to continue to affect important and meaningful change and has given me a powerful perspective on how to do so. I want to thank Stephanie for being my fellow graduate student companion and for showing me the ropes as the senior graduate student. She was often my guide and helped me understand the DAQ system and analysis and was there for me through challenging times.

I will always appreciate my time and experiences with the above members of the "OG" Michigan KOTO group as well as with the many Michigan KOTO group members over the years– from beam time at J-PARC, to bike rides along the river and to the fish market, and group dinners at Monica's house. I enjoyed all of the hard work and fun times together and I am extremely grateful and proud to be a part of such a welcoming and inclusive group.

I must also express my sincere gratitude to all members of the KOTO collaboration. In particular I want to thank Chieh (Jay) Lin and Satoshi Shinohara for being my PhD comrades along this journey. Both Jay and Shinohara were very helpful in answering my many questions and I have enjoyed working with them during our PhD years. I want to express my sincere appreciation and thanks to Professor Taku Yamanaka for his support and guidance, as well as for hosting me at Osaka University to work on analysis. I would also like to graciously thank all of the members of the Osaka group for making my stay so welcoming. Additionally, I want to thank my KOTO colleagues Professor Koji Shiomi, Professor Tadashi Nomura, Professor Hajime Nanjo, Professor Nobuhiro Shimizu, Professor Yau Wah, Professor Yee Bob Hsiung, Dr. Katsushige Kotera, Dr. Kota Nakagiri, and Dr. Ichinori Kamiji for their

advice and discussions on research work at J-PARC and analysis. I am grateful to all members of the collaboration who worked on the 2016–2018 data collection and data analysis, as the effort and dedication required for this research is considerable. I am forever thankful to be a part of this experiment and team of collaborators and I appreciate the opportunities to participate in cutting-edge research, travel to different countries, and present my work and ideas.

I want to give a special thanks to Dr. Nhan Tran, who was my collaborating scientist and advisor at Fermilab for my SCGSR fellowship, as well as Dr. Lindsey Gray who also helped guide my machine learning research project while I was there. I appreciate all of their expertise and guidance on my project, as well as the tools and skills that I developed thanks to their support. I also want to thank Yongyi Wu for being a good roommate and someone to talk to while I was resident at Fermilab and I want to thank Dr. Bo Jayatilaka for his friendship during my time there, as well as for his continued support and friendship thereafter.

I am deeply thankful for all of the people who have supported me on this journey and have helped me on my path to achieving this goal. I want to thank my professors at Agnes Scott– Professor Chris De Pree, Professor Amy Lovell, Professor Nicole Ackerman, Professor Paul Wallace, and Professor Art Bowling, for encouraging me, guiding me, and giving me a strong physics foundation and creating such a fun and welcoming environment in which to learn physics. I want to thank my best friends and my companions for life, Kimberly Luong and Katie Butler, for being there for me every step of the way and offering their unconditional love and support. I want to thank my dear friend and college roommate, Lulu Liu, for being my unwavering Manchurian warrior and for fun times in Japan together during my graduate studies. I would like to thank Enrique Gamez, for being a supportive and understanding friend, especially in the first couple of years of grad school when I was struggling with coursework, and for always making me feel comfortable asking for help on problem

sets. I want to thank Dr. Kaleigh Fisher-Grant and Kyle Fisher-Grant for being such wonderful roommates when I first moved to Michigan and for being such supportive and down-to-earth friends throughout my PhD career. I also want to thank my current roommate, Akari Oya, for her close friendship and steadfast support, for listening to my struggles over writing this thesis, and for making a home with me in Michigan for the past few years.

Finally, I would like to thank my adopted godparents, John and Amy Ellison, who always encouraged me to pursue my interests in physics, astronomy, and science in general, and whose enthusiasm drove me to reach for ever loftier goals. I am forever grateful for my family for supporting my dreams and my education and for their love and endless encouragement. I would like to give a special thanks to my father for reading and providing feedback on every chapter of this thesis, and for always teaching me to think critically and use my resources. I want to thank my mother for instilling in me a love of learning and reading, and for her strong spirit and determination that inspired me to pursue challenges with confidence. I want to thank my sisters, Rebecca and Lucy, for their everlasting friendship and unquestioning support and understanding, for always being real with me, and for always being there. I am also thankful for my extended Graves family for their support and welcoming acceptance. From the bottom of my heart, I appreciate everyone who has helped me to walk my path.

Melissa A. Hutcheson

Ann Arbor, MI

March, 2021

Table of Contents

List of Figures

Figure

List of Tables

Table

List of Appendices

Appendix

Abstract

The KOTO experiment located at the J-PARC research facility in Tokai, Japan aims to to measure the branching ratio (BR) of the rare decay of the neutral kaon, $K_L^0 \to \pi^0\nu\bar{\nu}$. This decay has a Standard Model (SM) predicted BR of $(3.00 \pm 0.30) \times 10^{-11}$. While it is highly suppressed, it is an excellent probe to search for new physics beyond the SM as it directly violates charge-parity (CP) and has small theoretical uncertainties.

This thesis performs an analysis with data collected in 2016–2018 and reports on the $K_L^0 \to \pi^0\nu\bar{\nu}$ search. $K_L^0 \to \pi^0\nu\bar{\nu}$ events were identified by detecting two photons from the π^0 decay with a large transverse momentum and no other particles present. The data was analyzed using a blind analysis technique and Monte Carlo simulations to ensure the reproducibility of the data. The analysis included event reconstruction and selection, normalization, background estimations, and the estimation of systematic uncertainties.

The Single Event Sensitivity (SES) achieved with the 2016–2018 data was $(7.20 \pm 0.05_{\text{stat}} \pm 0.66_{\text{syst}}) \times 10^{-10}$ which is the highest sensitivity level ever reached for the $K_L^0 \to \pi^0\nu\bar{\nu}$ search. The expected background in the signal region was estimated to be 0.05 ± 0.02. After unblinding the data, three events were observed in the signal region and two new backgrounds from $K^{\pm}$ decays and scattered K_Ls were considered. Consequently, the background estimation in the signal region was updated to be 1.22 ± 0.26. Based on these results, an upper limit on the branching ratio of $K_L^0 \to \pi^0\nu\bar{\nu}$ was set to be 4.9×10^{-9} at the 90% Confidence Level.

Chapter 1

Introduction and Theory

Something strange happened at the beginning of the universe, and we can't quite explain it. It is thought that when our universe formed an estimated 13.8 billion years ago, there were equal amounts of matter and antimatter created immediately after the Big Bang. However, when matter and antimatter come into contact with one another, they completely annihilate and produce massive amounts of energy. Why then, was there some small amount of matter left over to make up the universe that we know today? This matter and antimatter asymmetry can be partially explained by a phenomenon called charge-parity (CP) violation which is sometimes seen in the weak interactions of particles. However, the amount of CP violation that we observe in weak decays doesn't account for the amount of matter that we see in the universe. This indicates that there may be decay processes that exhibit more CP violation than expected, or entirely new processes altogether that violate CP. Understanding CP symmetry-breaking is one of the central goals of particle physics, and one of the ways that we can investigate it is by studying the decay of the neutral kaon, $K_L^0 \to \pi^0\nu\bar{\nu}$. This chapter will introduce the necessary background of physics and the particles in question– kaons, and how they behave and what makes the $K_L^0 \to \pi^0\nu\bar{\nu}$ decay particularly interesting. We will explore extensions to the current theory and the history of experimental searches for $K_L^0 \to \pi^0\nu\bar{\nu}$. The KOTO experiment will be introduced and the overview and objective of this thesis will be explained.

1.1 The Standard Model of Physics

In the later half of the 20th century, particle physicists developed a theory that classified and described elementary particles and their interactions. This became known as the Standard Model (SM) of physics, and while it doesn't include gravity and some other key phenomenon[1], it does an excellent job of predicting what it does include. Figure 1.1 shows the fundamental particles of the SM. There are six types, known as "flavors", of quarks and six leptons, four gauge bosons or force carriers, and one scalar boson– the Higgs boson. Each SM particle has an associated antiparticle (denoted with a bar over the letter) with opposite physical charges but the same mass. Since some particles are not charged, they are their own antiparticle[2].

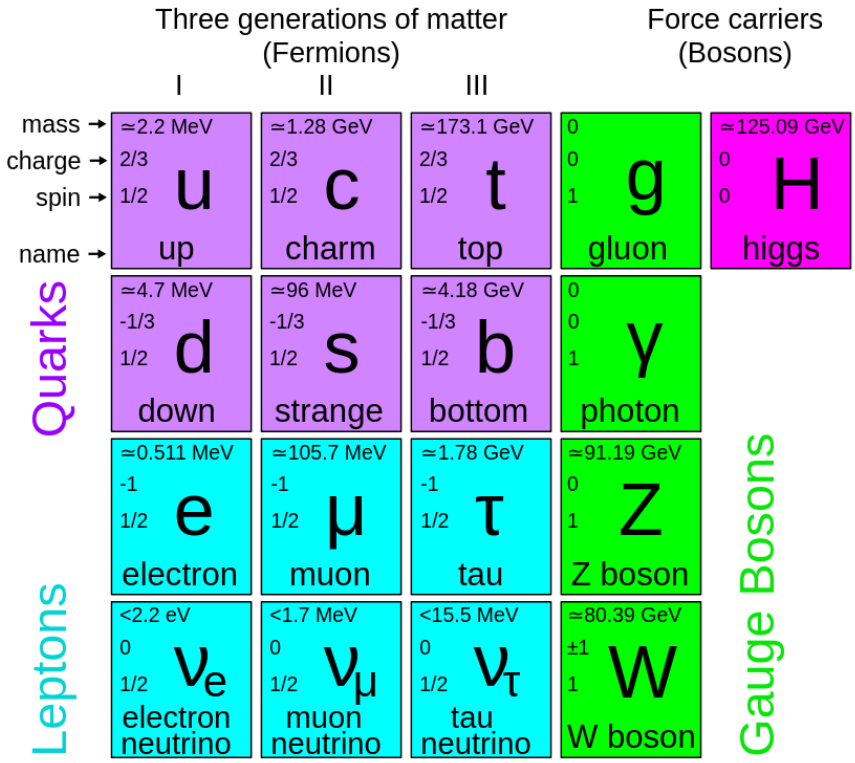


Figure 1.1: The Standard Model of particle physics. The particle masses listed in units of energy are assumed to be divided by c^2.

[1] Baryon asymmetry, dark matter, dark energy, neutrino oscillations

[2] The gluon, photon, Z boson, and Higgs boson are their own antiparticles.

The quarks can only exist in composite particles which are referred to as hadrons. Hadrons can be composed of three quarks, called baryons, or one quark and one antiquark, called mesons. The gauge bosons each correspond to a fundamental force and mediate the interactions between the other particles. The gluon mediates the strong force, the photon mediates the electromagnetic force, and the Z and W bosons mediate the weak force. Lastly, the Higgs boson, theoretically predicted in 1964 [1; 2], is responsible for giving particles mass and was the final SM particle that was experimentally confirmed by the ATLAS and CMS collaborations in 2012 [3; 4]. The experimental measurements of all of the elementary particles are consistent with the SM prediction.

1.2 CP Symmetry and Violation

There are three discrete symmetries that are fundamental in the Standard Model– charge conjugation (C), parity (P), and time reversal (T). A symmetry in physics means that the laws of physics should be the same under some transformation; if they are not, then that symmetry is broken, or violated. Charge conjugation transforms a particle to its antiparticle by changing the sign of all quantum charges. For example, an electron (e^-) is transformed under C to a positron (e^+). Parity changes a particle's "handedness" or "helicity", meaning it inverts a particle's spatial coordinates like the inversion that occurs in a mirror. Under P, the handedness of space is reversed, and a right-handed electron becomes a left-handed electron. Finally, time reversal symmetry says that any particle interaction is invariant under time reversal.

In the 1950s, it was discovered that both C and P symmetry are violated separately in weak interactions. After this time, it was thought that charge conjugation could be combined with parity to preserve a combined CP symmetry. For example, a left handed electron (e^-_L) is transformed under CP into a right-handed positron (e^+_R). While C and P are violated separately in many weak interactions, CP symmetry is

still preserved in most weak interaction processes. However, in 1964 physicists Cronin and Fitch discovered that CP is actually violated in neutral kaon decays [5]. Since then, CP violation has been observed in kaon and B meson weak decays, both directly and indirectly. It has also been experimentally observed in the B^0 meson system that T symmetry is violated [6]. While C, P, T, and the combined CP symmetry is broken, the combination of all three, CPT symmetry, still remains invariant and is recognized to be a fundamental property of the laws of physics.

The existence of CP violation indicates that the laws of physics are not the same for matter and antimatter. We know that CP violation exists and can measure how often it occurs, but there still may be some unknown CP violating processes that we haven't observed. Thus, understanding this phenomenon may help us understand why we live in a matter-dominated universe.

1.3 Kaon Phenomenology

Kaons are a type of meson containing an s or $\bar{s}$ quark. Kaons, like many composite particles, are unstable and after some period of time, decay through weak interactions into other particles. Since their discovery in 1947 [7], they have played a critical role in our understanding of the nature of particle interactions and the laws of physics. The charged kaons, K^+ and K^-, are composed of $u\bar{s}$ and $s\bar{u}$, respectively. The neutral kaon with an electric charge of zero has much more complex states, as illustrated in the following section.

1.3.1 $K^0 - \bar{K^0}$ Mixing

The neutral kaon, K^0, and its antiparticle, $\bar{K^0}$, can be represented by the mass eigenstates,

$$|K^0\rangle = |d\bar{s}\rangle \tag{1.1}$$

$$|\bar{K^0}\rangle = |s\bar{d}\rangle \tag{1.2}$$

These states are eigenstates of parity but not charge conjugation. If we operate on the mass eigenstates with the CP operator we see that neither of them are their weak eigenstates:

$$CP\,|K^0\rangle = -\,|\bar{K^0}\rangle \tag{1.3}$$

$$CP\,|\bar{K^0}\rangle = -\,|K^0\rangle \tag{1.4}$$

Thus, the normalized CP eigenstates are a combination of K^0 and $\bar{K^0}$ where,

$$|K_1\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\bar{K^0}\rangle) \tag{1.5}$$

$$|K_2\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\bar{K^0}\rangle) \tag{1.6}$$

and if we operate on K_1 and K_2 with the CP operator we can see that they are indeed eigenstates of CP:

$$CP\,|K_1\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\bar{K^0}\rangle) = |K_1\rangle \tag{1.7}$$

$$CP\,|K_2\rangle = -\frac{1}{\sqrt{2}}(|K^0\rangle + |\bar{K^0}\rangle) = -\,|K_2\rangle \tag{1.8}$$

Under this assumption, we expect the K_1 state to decay into two neutral pions ($2\pi^0$), and the K_2 state to decay into three neutral pions ($3\pi^0$), since the π^0 has a CP eigenvalue of -1 and the K_1 state is CP even and the K_2 state is CP odd. This can be seen in Equations 1.9 and 1.10.

$$\begin{array}{ccc} |K_1\rangle & \longrightarrow & 2\pi^0 \\ CP+1 & & CP+1 \end{array} \tag{1.9}$$

$$\begin{array}{ccc} |K_2\rangle & \longrightarrow & 3\pi^0 \\ CP-1 & & CP-1 \end{array} \tag{1.10}$$

Since the $3\pi^0$ decay has less decay phase space, we would expect K_2 to have a longer lifetime, and for K_1 to have a short lifetime as the $2\pi^0$ decay has more decay phase space. This prediction was made by Gell-Mann and Pais in 1954 [8] and these two states were actually observed at Brookhaven in 1956 [9]. The K_1 state was named K_S^0 (K-short) and the K_2 state was named K_L^0 (K-long) based on their lifetimes.

However, in 1964, Cronin and Fitch observed something strange. From a beam of K_L^0 particles they expected to detect only $3\pi^0$ decays, but instead observed that a small fraction (about one out of 500) decayed to $2\pi^0$ [5]. This was the first experimental evidence of CP symmetry violation, and this discovery won them the Nobel Prize in 1980. Thus, the weak eigenstates are actually,

$$|K_L^0\rangle = \frac{1}{\sqrt{1+|\epsilon|^2}}(|K_2\rangle + \epsilon\,|K_1\rangle) \tag{1.11}$$

$$|K_S^0\rangle = \frac{1}{\sqrt{1+|\epsilon|^2}}(|K_1\rangle - \epsilon\,|K_2\rangle) \tag{1.12}$$

where K_L^0 is mostly composed of the K_2 state but has a small amount of K_1 and vice versa for K_S^0. The average lifetime of K_S^0 is $(0.8953 \pm 0.0005) \times 10^{-10}$ seconds while the average lifetime of K_L^0 is $(5.116 \pm 0.020) \times 10^{-8}$ seconds [10]. The fact that the K_L^0 has a lifetime about 100 times longer than K_S^0 will be important later on. The value of ϵ indicates the amount of contamination from CP violation and has a value of [10]

$$|\epsilon| = (2.228 \pm 0.011) \times 10^{-3} \tag{1.13}$$

It should be noted that in this thesis I refer to the long-lived neutral kaon as K_L^0 and sometimes K_L. When I refer to "kaons" in general, I am referring to K_L^0. The short-lived neutral kaon will always be specified as K_S^0 or K_S.

1.4 The Cabibbo-Kobayashi-Maskawa Matrix

The origin of CP violation in the Standard Model comes from the Cabibbo-Kobayashi-Maskawa (CKM) matrix [11; 12]. The CKM matrix is a unitary matrix that relates the weak eigenstates (primed) to the mass eigenstates (unprimed) as shown in Equation 1.14 and describes the strength of flavor changing weak decays.

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = V_{CKM} \begin{bmatrix} d \\ s \\ b \end{bmatrix} \qquad \text{where} \qquad V_{CKM} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \tag{1.14}$$

In changing from one quark flavor, i, to another, j, the probability of that transition is proportional to $|V_{ij}^2|$. CP violation as well as quark mixing among the three generations are explained by the CKM mechanism. The CKM matrix is defined by four independent parameters and can be written in different parameterizations. The original parameterization [12] uses three angles (θ_1, θ_2, and θ_3) and one complex, CP violating phase (δ) which indicates the amount of CP violation. The Wolfenstein parameterization [13] has four parameters, λ, A, ρ, and η and can be written as

$$V_{CKM} = \begin{bmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & \end{bmatrix} + O(\lambda^4) \tag{1.15}$$

where the parameters are defined as

$$\lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}} \tag{1.16}$$

$$A\lambda^2 = \lambda \frac{|V_{cb}|}{|V_{us}|} \tag{1.17}$$

$$A\lambda^3(\rho + i\eta) = V_{ub}^* \tag{1.18}$$

The parameter η is the CP violating contribution and the magnitude of this value is a measure of CP violation in the SM. It is often convenient to define $\bar{\rho}$ and $\bar{\eta}$:

$$\bar{\rho} \sim \rho(1 - \lambda^2/2) \tag{1.19}$$

$$\bar{\eta} \sim \eta(1 - \lambda^2/2) \tag{1.20}$$

as we will see in the following section.

1.4.1 Unitarity Triangles

Because the CKM matrix is unitary, we can use this property as an important constraint to test the Standard Model. A unitary matrix, U, has the property

$$U^\dagger U = I \tag{1.21}$$

where I is the identity matrix and $U^\dagger$ indicates the conjugate transpose of U. If we apply this property to the CKM matrix we can write the unitary condition

$$V_{us}^* V_{ud} + V_{cs}^* V_{cd} + V_{ts}^* V_{td} = 0 \tag{1.22}$$

from which we can construct a triangle in the complex plane based on the magnitude of the terms, called a "unitarity triangle". If we normalize the bottom of the triangle to lie on the real axis and have a length of one, we can draw the unitarity triangle in the $\bar{\rho}$ – $\bar{\eta}$ plane as shown in Figure 1.2. It turns out, measurements on CP violating processes correspond to measurements of the unitarity triangle; for example, the length of the sides, the interior angles, and the height, all can be measured through some SM process. Therefore, a key test of the SM is to measure these processes and determine if the triangle is still unitary– if it's not, this would be a clear indication

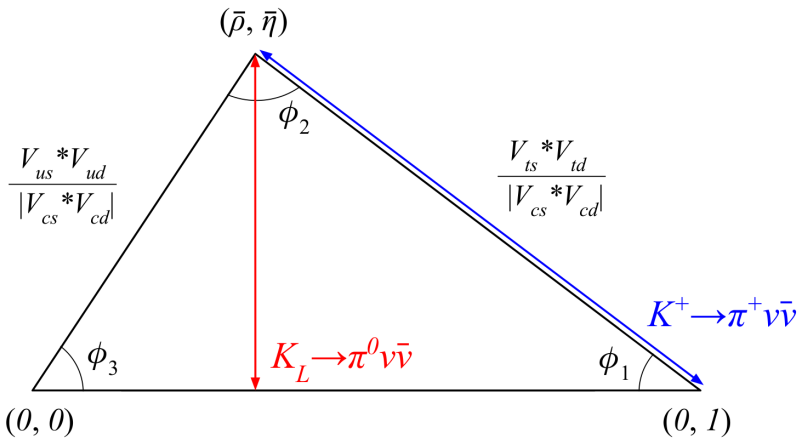


Figure 1.2: The kaon unitarity triangle in the $\bar{\rho}$ – $\bar{\eta}$ plane. The height of the unitarity triangle is directly related to the decay, $K_L^0 \to \pi^0 \nu\bar{\nu}$.

of physics beyond the Standard Model (BSM).

The triangle drawn in Figure 1.2 is the kaon unitarity triangle. Other triangles related to other processes, for example B physics, can be drawn. In particular, the CP violating parameter, $\bar{\eta}$, can be obtained directly by measuring the branching ratio (BR) of the decay $K_L^0 \to \pi^0 \nu\bar{\nu}$ which corresponds to the height of the unitarity triangle. The branching ratio for a decay is the fraction of time a particle decays to a particular final state. In other words, it is the fraction of particles which decay by a single decay mode with respect to the total number of particles that decay. $\bar{\eta}$ can also be determined indirectly from the decay $K^+ \to \pi^+ \nu\bar{\nu}$, and the two decays are closely related as we will see in Section 1.6. Thus, the $K_L^0 \to \pi^0 \nu\bar{\nu}$ decay plays an important role in testing CP violation in the SM.

1.5 The $K_L^0 \to \pi^0 \nu\bar{\nu}$ Decay

It is clear that studying CP violation is important to constraining the Standard Model parameters. In particular, CP violation can be studied in the quark or lepton sectors. Along with testing lepton universality in the lepton sector, there are four "golden" processes that are powerful tools for probing CP violation in the quark

sector: asymmetries in $B^0 \to J/\psi K_s^0$, the ratio of B_s to B_d mixing, and the decays $K^+ \to \pi^+ \nu\bar{\nu}$ and $K_L^0 \to \pi^0 \nu\bar{\nu}$ [14]. While these processes are extremely rare, they are referred to as "golden" because they are very clean processes and therefore well suited to investigate CP violation. The focus here is on $K_L^0 \to \pi^0 \nu\bar{\nu}$ mainly and $K^+ \to \pi^+ \nu\bar{\nu}$, though a short overview of the B physics processes and lepton universality are covered in Appendix B.

The $K_L^0 \to \pi^0 \nu\bar{\nu}$ decay is a Flavor Changing Neutral Current (FCNC) process that proceeds through second-order weak interactions and directly violates CP. FCNCs are forbidden at tree level in the SM and so this decay proceeds through a loop diagram as shown in Figure 1.3. The name FCNC comes from the fact that the flavor of the quark changes and the charge stays the same. In this case, a t quark mediates the decay of an s quark to a d quark. The flavor change can also be mediated through u or c, but the t quark dominates in the vertices because it has the largest mass.

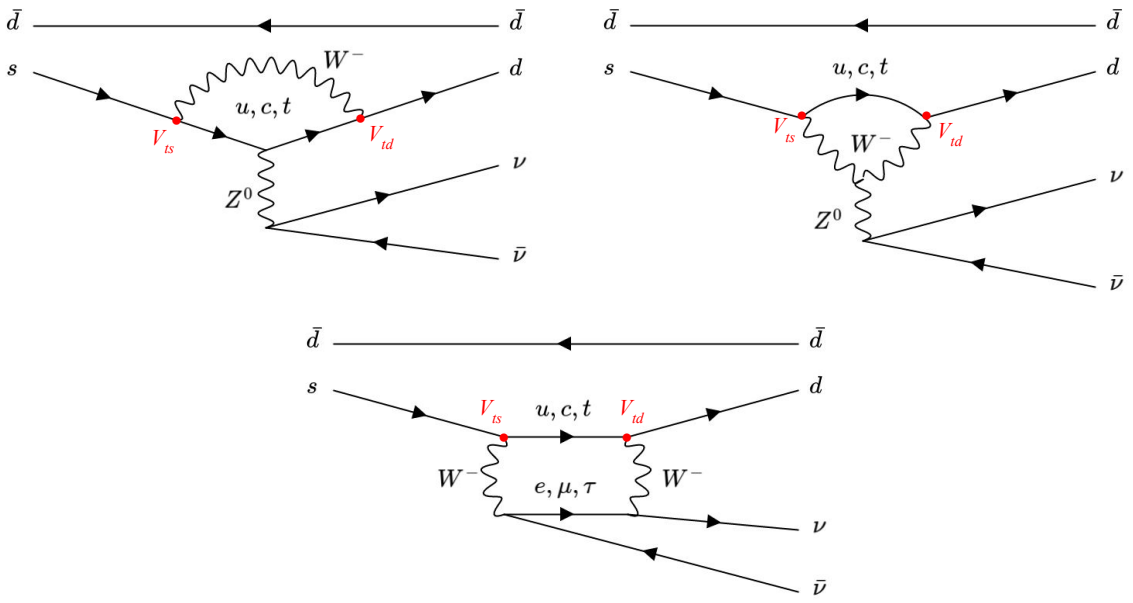


Figure 1.3: SM Feynman diagrams for the $K_L^0 \to \pi^0 \nu\bar{\nu}$ decay. The neutral kaon is represented as $\bar{d}s$ and the neutral pion is made of $\bar{d}d$. The $s \to d$ transition proceeds through a loop and is mediated by either a u, c, or (most often) t quark. The CKM matrix elements that contribute are shown in red. The top two diagrams are referred to as "penguin" diagrams and the bottom diagram is a box diagram.

Though $K_L^0 \to \pi^0 \nu\bar{\nu}$ is highly suppressed in the SM, the theoretical uncertainties on the SM predicted branching ratio are quite small and so it is an excellent candidate to search for BSM physics and investigate CP violation. If the experimentally measured branching ratio is higher than the predicted value, this would point to physics beyond the SM. Figure 1.4 shows a possible extension to the SM involving a charged Higgs and a BSM diagram in which an undiscovered particle, X, could both change the flavor and keep the current neutral without having to go through a loop.

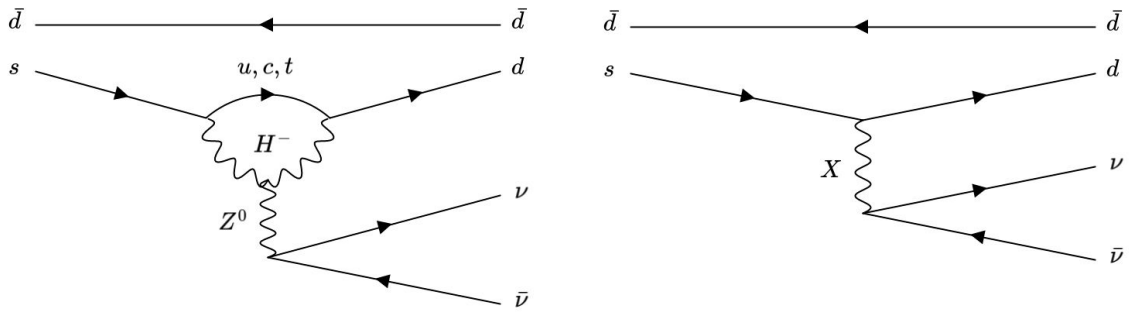


Figure 1.4: $K_L^0 \to \pi^0 \nu\bar{\nu}$ BSM diagrams. A possible extension to the SM in which a charged Higgs mediates the $s \to d$ transition (left) and a BSM diagram in which an unknown particle, X, mediates the decay without a loop (right).

The branching ratio of $K_L^0 \to \pi^0 \nu\bar{\nu}$ calculated from the SM theory [15] is given in Equation 1.23

$$\mathrm{BR}(K_L^0 \to \pi^0 \nu\bar{\nu}) = \kappa_L \left(\frac{\mathrm{Im}(V_{ts}^* V_{td})}{\lambda^5} X(x_t) \right)^2 \tag{1.23}$$

where x_t is the square of the ratio between the top quark mass and the W boson mass and can be expressed as

$$x_t = \frac{m_t^2}{M_W^2} \tag{1.24}$$

and $\lambda = |V_{us}|$. The Inami-Lim loop function, $X(x_t)$, factors in the loop effect and higher order quantum chromodynamics (QCD) corrections [16]. Finally, the factor κ_L includes other effects and is given as

$$\kappa_L = (2.231 \pm 0.013) \times 10^{-10} \left(\frac{\lambda}{0.225} \right)^8 \tag{1.25}$$

Because this decay mode involves only heavy particles such as the top quark and the W and Z bosons, the effects of long distance interactions from light quarks are negligible. Also, the K_L form factor does not depend on lattice QCD calculations which have large uncertainties and is instead obtained from K_L semi-leptonic decay rates which are well measured. Thus, the error on the BR calculation is dominated by the uncertainty in CKM parameters and even so is extremely small at about 1.2%. This makes $K_L^0 \to \pi^0 \nu\bar{\nu}$ theoretically very clean which is unique among other CP violating processes. Using the above equations, the SM predicted branching ratio is calculated to be [15]

$$\mathrm{BR}(K_L^0 \to \pi^0 \nu\bar{\nu}) = (3.00 \pm 0.30) \times 10^{-11} \tag{1.26}$$

This effectively means that this decay mode only happens once in roughly 33 billion neutral kaon decays. As rare as this processes is, measuring the branching ratio of $K_L^0 \to \pi^0 \nu\bar{\nu}$ would be a critical test of CP violation in the SM.

1.6 $K^+ \to \pi^+ \nu\bar{\nu}$ and the Grossman-Nir Bound

The charged kaon decay, $K^+ \to \pi^+ \nu\bar{\nu}$, is in many ways analogous to the neutral kaon decay, in that it proceeds through a similar FCNC process requiring a loop and is also highly suppressed in the SM. However, the charged kaon decay violates CP symmetry indirectly, while $K_L^0 \to \pi^0 \nu\bar{\nu}$ is directly CP violating. As seen in Section 1.4.1, the two decay processes are closely related. In fact, by using the branching ratio of the charged kaon decay, we can set an indirect limit on $\mathrm{BR}(K_L^0 \to \pi^0 \nu\bar{\nu})$. This model-independent upper bound on the branching ratio, called the Grossman-Nir bound [17], is derived from isospin symmetry arguments as

$$\mathrm{BR}(K_L^0 \to \pi^0 \nu\bar{\nu}) < 4.4 \times \mathrm{BR}(K^+ \to \pi^+ \nu\bar{\nu}) \tag{1.27}$$

Thus, the measured $\mathrm{BR}(K^+ \to \pi^+ \nu\bar{\nu})$ of $(1.7 \pm 1.1) \times 10^{-10}$ from the Brookhaven National Laboratory (BNL) E949 experiment [18] gives an upper limit on $\mathrm{BR}(K_L^0 \to \pi^0 \nu\bar{\nu})$ to be 1.5×10^{-9} at the 90% Confidence Level (CL). The NA62 experiment at CERN is a continuation of the E949 experiment and aims to measure $\mathrm{BR}(K^+ \to \pi^+ \nu\bar{\nu})$ which has a SM predicted BR of $(9.11 \pm 0.72) \times 10^{-11}$ [15]. Results from NA62 report one observed event which gives an upper limit of $\mathrm{BR}(K^+ \to \pi^+ \nu\bar{\nu}) < 14 \times 10^{-10}$ at the 95% CL [19]. These results are included in Figure 1.5 which shows the BR possibilities for both $K_L^0 \to \pi^0 \nu\bar{\nu}$ and $K^+ \to \pi^+ \nu\bar{\nu}$ for various theoretical models.

While the charged decay mode has been relatively well constrained, the neutral

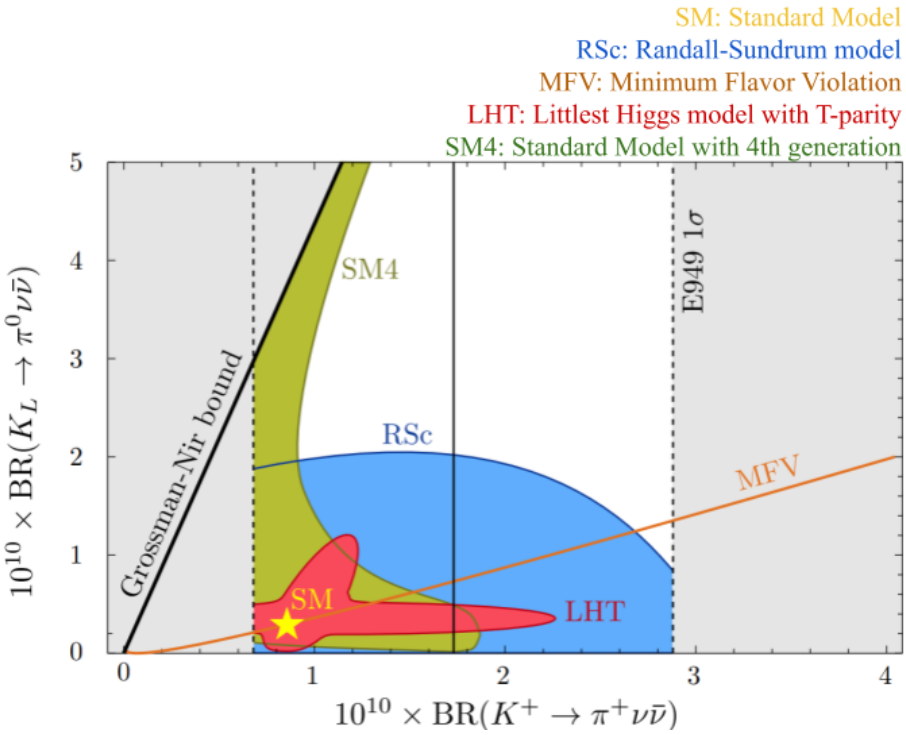


Figure 1.5: Branching ratios for $K_L^0 \to \pi^0 \nu\bar{\nu}$ and $K^+ \to \pi^+ \nu\bar{\nu}$ for various models. The grey area is ruled out experimentally by the E949 experiment (measurement shown in black with 1σ bounds) or by the Grossman-Nir bound. The SM prediction is indicated by the yellow star. Figure modified from [20].

decay mode still has a large range of possibilities for the branching ratio. A combination of both $K_L^0 \to \pi^0\nu\bar{\nu}$ and $K^+ \to \pi^+\nu\bar{\nu}$ measurements is necessary to developing a better understanding of the underlying CP violating processes.

1.7 Beyond SM Theories

The level of CP violation in the SM does not account for the amount of matter that is seen in the universe, which points to physics beyond the SM. There are several theoretical models that give a branching ratio for $K_L^0 \to \pi^0\nu\bar{\nu}$ that is larger than the SM prediction. Possible branching ratios for some of these models are shown in Figure 1.5. In particular, the Randall-Sundrum model with custodial symmetry (RSc) [21] and the Littlest Higgs model with T-parity (LHT) [22] both give a larger branching ratio for the K_L^0 and K^+ rare decays but the effects of this new physics on B mesons is expected to be small. In another model, Minimum Flavor Violation (MFV) [23], the deviation from the SM predicted branching ratio is relatively small. MFV like the SM, assumes only one CP violating phase in the CKM matrix. This indicates that if a large deviation from the SM is observed, new CP violating phases must exist. The SM with a 4th generation of quarks and leptons (SM4) would require two CP violating phases and a 4×4 flavor mixing matrix and the branching ratio of K_L^0 and K^+ would be affected [24]. Other models such as the Minimal Supersymmetric Model (MSSM) [25] and the Z' model [26] would also cause a sizeable deviation from the SM predicted branching ratios.

1.8 History of Experimental Searches and Previous Limits

The importance of the $K_L^0 \to \pi^0\nu\bar{\nu}$ decay is perhaps best illustrated by the search history (Figure 1.6). Since the late 1980s until now, many experiments have searched for $K_L^0 \to \pi^0\nu\bar{\nu}$ and set upper limits on the branching ratio. The first study was

performed by Littenberg in 1989 [27], in which he extracted a limit on the branching ratio for $K_L^0 \to \pi^0\nu\bar{\nu}$ based on the data from Cronin and Fitch's experiment with neutral kaons. Two experiments at Fermi National Accelerator Laboratory (Fermilab) set limits in 1992 [28] and 1994 [29] using the Dalitz decay of the π^0 ($\pi^0 \to e^+e^-\gamma$). The Dalitz decay occurs when one of the two photons from the π^0 decay converts to an electron positron pair and the e^+ and e^- can be detected. While it is easy to detect, this process only happens about 1% of the time.

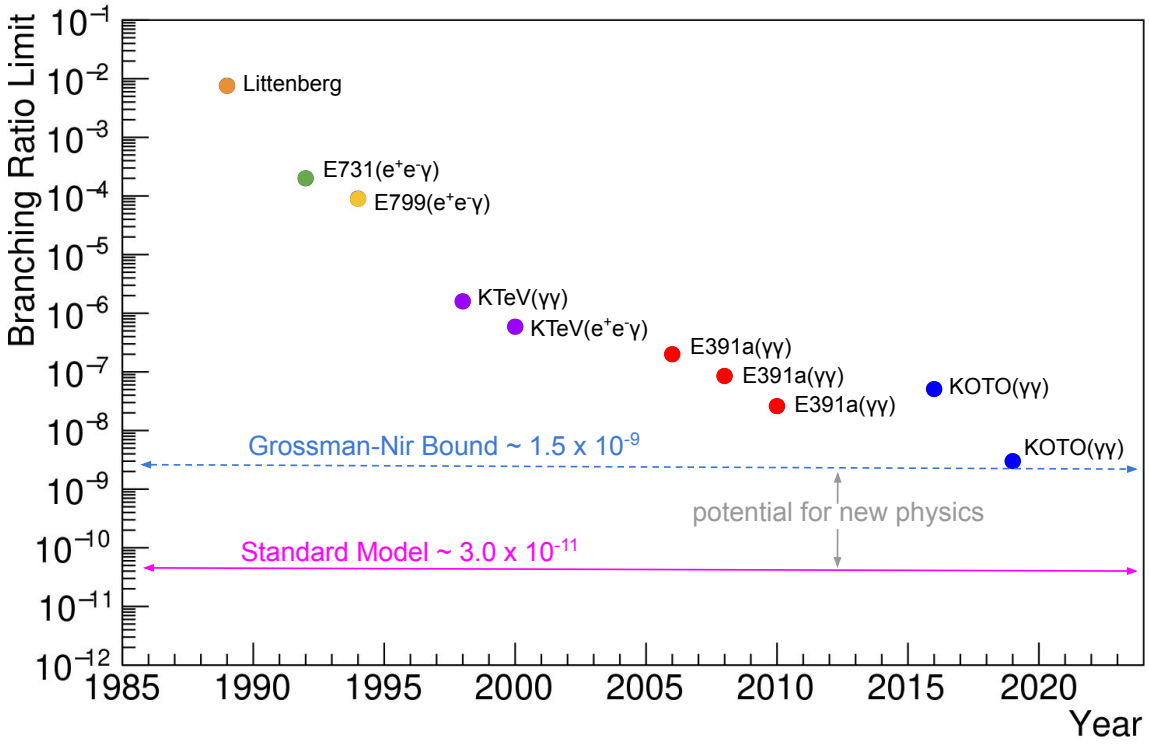


Figure 1.6: The search history of $K_L^0 \to \pi^0\nu\bar{\nu}$. The branching ratio upper limit is shown for various experiments.

The KTeV experiment at Fermilab was dedicated to measuring the direct CP violation parameter Re(ϵ'/ϵ) and various rare K_L decays. A limit on BR($K_L^0 \to \pi^0\nu\bar{\nu}$) was set in 1999 by observing the two photons from the π^0 decay [30]. Though $\pi^0 \to \gamma\gamma$ occurs 99% of the time, detecting a neutral final state is much more difficult as tracking detectors cannot be used. Despite this, the KTeV experiment was the first

to prove that detecting $\pi^0 \to \gamma\gamma$ was a viable method to search for $K_L^0 \to \pi^0\nu\bar{\nu}$. Another limit on BR($K_L^0 \to \pi^0\nu\bar{\nu}$) was set using the Dalitz decay in 2000 [31].

The first dedicated search for $K_L^0 \to \pi^0\nu\bar{\nu}$ was carried out by the E391a experiment at the High Energy Accelerator Research Organization (KEK) 12 GeV proton synchrotron in Tsukuba, Japan. E391a used an electromagnetic calorimeter and a hermetic veto system to detect the two photons from the π^0. The experiment set limits in 2006 [32] and 2008 [33], and the final experimental upper limit was set in 2010 as BR($K_L^0 \to \pi^0\nu\bar{\nu}$) $< 2.6 \times 10^{-8}$ at the 90% CL [34]. The KOTO experiment is the successor to E391a and is introduced in the following section. From data taken in 2013, KOTO set a limit of 5.1×10^{-8} at the 90% CL on BR($K_L^0 \to \pi^0\nu\bar{\nu}$) [35] and from data taken in 2015 KOTO set a new best experimental limit [36] of

$$\mathrm{BR}(K_L^0 \to \pi^0\nu\bar{\nu}) < 3.0 \times 10^{-9} \ \ (90\%\ \mathrm{CL}) \tag{1.28}$$

This result was published in 2019 and sits just above the Grossman-Nir bound as shown in Figure 1.6. Thus, the next measurement expects to push into the realm of potential new physics.

1.9 Introduction to the KOTO Experiment

The KOTO experiment was designed to search for $K_L^0 \to \pi^0\nu\bar{\nu}$ by observing the two photons from the pion decay and nothing else. The goal of KOTO is to reach the Standard Model sensitivity and observe the $K_L^0 \to \pi^0\nu\bar{\nu}$ decay and measure its branching ratio. KOTO, like other experiments searching for extremely rare processes, is a precision measurement experiment and is currently the only experiment in the world searching for $K_L^0 \to \pi^0\nu\bar{\nu}$. The experiment is located in Japan and the experimental method and setup is detailed in Chapter 2. KOTO is an international collaboration which brings together over 60 collaborators from 16 different institu-

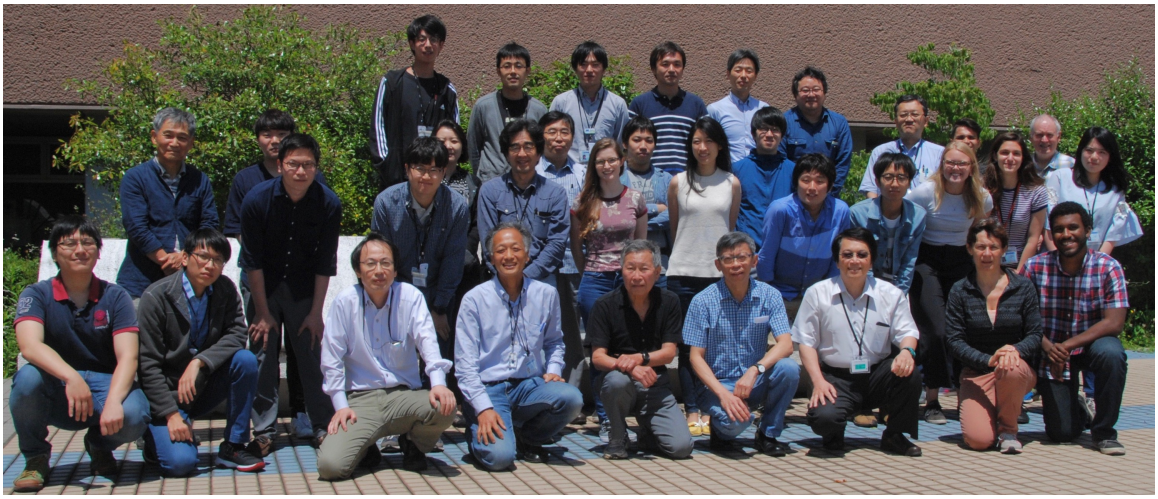


Figure 1.7: Members of the KOTO collaboration in June of 2018 (not all members present).

tions in Japan, the United States, Taiwan, South Korea, and Russia. A photo of the collaboration members in June of 2018 is shown in Figure 1.7.

1.10 Objective and Overview of This Thesis

The main purpose of this thesis is to report on the search for the $K_L^0 \to \pi^0\nu\bar{\nu}$ decay with data collected in 2016–2018. This involves several steps as follows:

1. Calculate the total number of kaons ($N_{K_L^0}$) collected
2. Estimate the Single Event Sensitivity (SES)
3. Estimate the total background contribution
4. Report a new branching ratio limit for $K_L^0 \to \pi^0\nu\bar{\nu}$

This chapter introduced the physics motivations and theory behind the search for $K_L^0 \to \pi^0\nu\bar{\nu}$. Chapter 2 contains details about the KOTO experiment's setup while Chapter 3 describes the Data Acquisition (DAQ) system used to collect the data. The 2016–2018 data collection and the performance of the DAQ system are discussed

in Chapter 4. Chapter 5 details the event reconstruction and selection as well as the the analysis strategy which includes Monte Carlo (MC) simulation methods, normalization analysis, and background estimations as described in Chapters 6, 7, and 8, respectively. The opening of the box and the final results are presented in Chapter 9. Lastly, Chapter 10 gives conclusions and a discussion of the results as well as future prospects for KOTO and the continued search for $K_L^0 \to \pi^0 \nu\bar{\nu}$. Preliminary work for a new background reduction method using machine learning is described in Appendix A and Appendix B contains a brief overview of B physics and lepton universality. A list of acronyms and definitions can be found in Appendix C.

Chapter 2

Experimental Method

The E14 experiment, also known as the KOTO experiment, is located at the Japan Proton Accelerator Research Complex (J-PARC) in Tokai Village, Ibaraki Prefecture, Japan and was designed to observe the decay, $K_L^0 \to \pi^0\nu\bar{\nu}$. The name "KOTO" comes from the neutral kaon and the experiment's location, $\mathbf{K^0}$ at **To**kai. The experiment was proposed in 2006 [37], a neutral K_L beam line was built in 2009, and KOTO took its first physics runs in 2013. This thesis focuses on data collected in 2016–2018, and the experimental setup during this time is described here. The concept of the experimental method is simple, though in practice it is extremely difficult. To observe the $K_L^0 \to \pi^0\nu\bar{\nu}$ decay, one must create over 33 billion neutral kaons, watch them all decay using the experiment's detector systems, and then select the one that decays into $\pi^0\nu\bar{\nu}$. It could be described as trying to find "a needle in a haystack", though a more accurate description might be trying to find "a needle in a needle-stack". The creation of the neutral kaons and the detectors that collect their decay products are described in this chapter.

2.1 Experimental Strategy

In order to identify the signal decay, we need a way to distinguish it from the various backgrounds. The signal decay occurs when a neutral kaon decays into a

neutral pion and two neutrinos– one neutrino and one anti-neutrino. The π^0 then immediately decays into two photons and the two neutrinos are not seen by the KOTO detectors as shown in Figure 2.1. The photons are captured by the main detector, an electromagnetic calorimeter composed of Cesium Iodide (CsI) crystals, and a hermetic veto system is in place to reject particles from other decays. The experimental difficulty in observing $K_L^0 \to \pi^0\nu\bar{\nu}$ is due to the lack of charged particles in the final state and due to the high efficiency required for the detection of extra particles from other decay modes, such as $K_L^0 \to 3\pi^0$, $K_L^0 \to 2\pi^0$, and $K_L^0 \to \pi^0\pi^+\pi^-$. The strategy of KOTO is to observe the two photons from the π^0 decay with no other particles present in the detectors, in which the two photons have a discernible, large, transverse momentum.

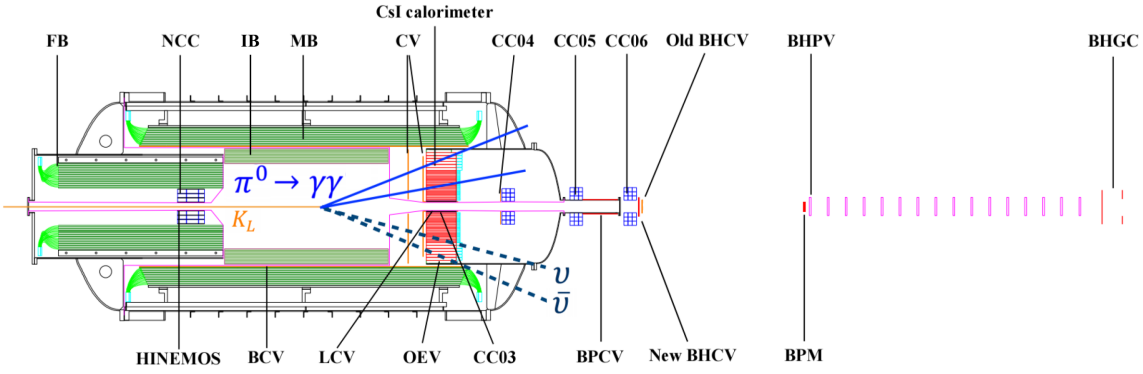


Figure 2.1: Schematic view of the KOTO detectors with the signal decay shown. The two photons are captured by the CsI calorimeter (red) and the two neutrinos are not seen by the detectors.

2.2 Experimental Setup

KOTO is what's called a "fixed target experiment" in particle physics, where particles such as protons or heavy nuclei are accelerated to high speeds and are then extracted from the main beam line to collide with a stationary target. These collisions produce a spray of many other particles, which can be filtered off to create a secondary

beam of particles. At the J-PARC facility, a beam of 30 GeV protons collides with a gold target and from these particle collisions, a secondary beam of neutral kaons is produced. In the KOTO experiment, we use a decay in-flight technique and observe the decay products from the kaons with the CsI calorimeter and various veto detectors. By using the average lifetime of the long-lived neutral kaon (around 5×10^{-8} seconds), we place the detectors in a region downstream from the target so that many of the kaons decay within the decay volume. The terms "downstream" and "upstream" are often used to describe the direction or location along the beam, in which downstream is in the direction of the beam flow and upstream is in the opposite direction. A small beam hole in the center of the calorimeter is needed to allow the kaons that did not decay to pass through the detector system with no interaction.

2.2.1 J-PARC Facility and Accelerator

The J-PARC facility is a high-intensity accelerator research facility in the Ibaraki Prefecture on the east coast of Japan that produces some of the highest intensity particle beams in the world. J-PARC uses three stages to accelerate protons to high energies and has three research facilities that utilize the high power proton beam. The first stage is a linear accelerator (Linac), the second stage is a 3 GeV Rapid Cycling Synchrotron (RCS), and the third stage is the main ring (MR). The research facilities are the Material and Life Science Experimental Facility (MLF), the Neutrino Experimental Facility (NU), and the Hadron Experimental Facility (HD or Hadron Hall). An aerial view of J-PARC is shown in Figure 2.2.

In the Linac, negative hydrogen ions, H^-, from a 50 keV ion source are accelerated up to 400 MeV. The accelerated H^- ions are stripped of their electrons with a stripper foil and the resultant protons are injected into the RCS [39]. In the RCS, the injected protons are accelerated up to 3 GeV at a repetition rate of 25 Hz in bunches of two [40]. The RCS transports 93−97% of the 3 GeV proton beam to the MLF and a

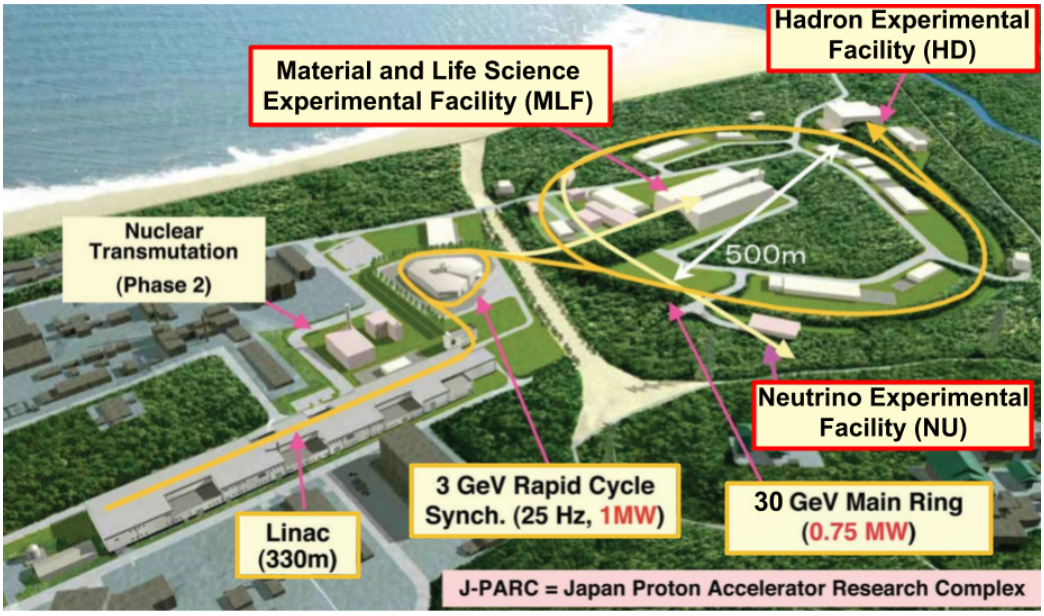


Figure 2.2: Aerial view of the J-PARC research facility. The three accelerator stages are shown in yellow and the research facilities are shown in red. Modified from [38].

small part of the beam from the RCS is extracted to the main ring. In this case, two bunches from the RCS are injected into the main ring four times and then those eight proton bunches are accelerated up to 30 GeV in under two seconds with repetition rates of $0.17 - 0.4$ Hz [41].

An important measure of the performance of the beam is the beam power, which can be defined as the beam intensity (number of particles per second) times the particle energy as shown in Equation 2.1.

$$\text{beam power} = dN/dt\ [\text{p}\mu\text{A}] \times \text{particle energy [GeV]} \tag{2.1}$$

In this equation dN/dt is the number of particles per second in particle micro amperes (pμA) in which 1 pμA $= 6.25 \times 10^{12}$ particles/sec. The beam power is typically measured in kilowatts (kW) and since the KOTO experiment aims to observe as many kaon decays as possible in order to observe $K_L^0 \to \pi^0\nu\bar{\nu}$, a high beam power is desirable. Since the energy of the protons is constant at 30 GeV, the beam intensity

is the variable factor, and more particles per second means more kaons produced that KOTO can collect. During the 2016–2018 data collection, the beam power varied from 31 kW to 51 kW with an average beam power of around 45 kW.

Once the proton bunches are in the main ring, they are extracted to the Neutrino Experimental Facility using fast extraction (FX) or to the Hadron Experimental Facility using slow extraction (SX) (Figure 2.3). The FX system delivers the beam to the NU using fast-kicker magnets which extract the eight circulating proton bunches all at once within 5μs. The SX system on the other hand delivers the beam to Hadron Hall over the course of two seconds. In SX mode, the beam is debunched by turning off the rf (radio frequency) power at 30 GeV and the protons are slowly extracted over a two second period using third-integer resonance extraction [41]. The purpose of using slow extraction is to reduce the instantaneous rate of particles coming into the detectors to avoid pileup of events. The slow extraction from the MR to Hadron Hall is detailed in the next section.

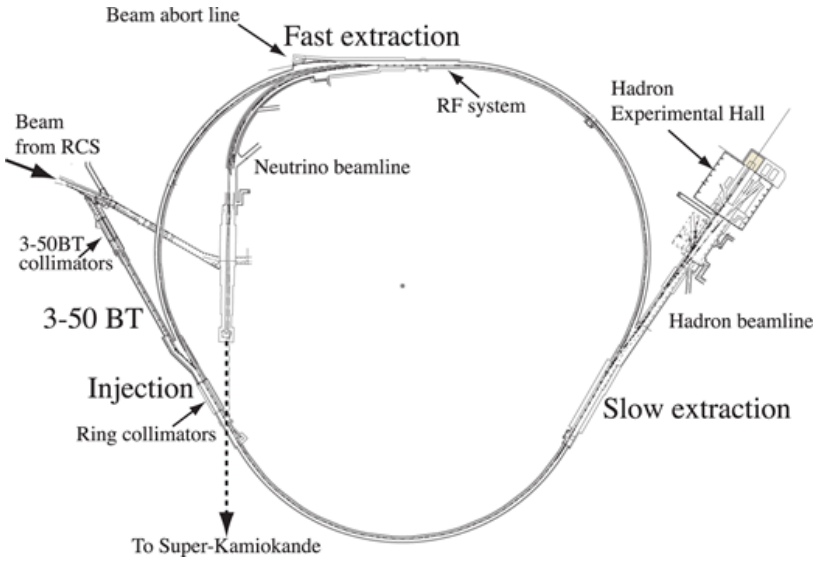


Figure 2.3: The J-PARC main ring showing the fast extraction and slow extraction systems. Fast extraction is used for the neutrino beam line and slow extraction is used for the experiments inside Hadron Hall. Image courtesy of [38].

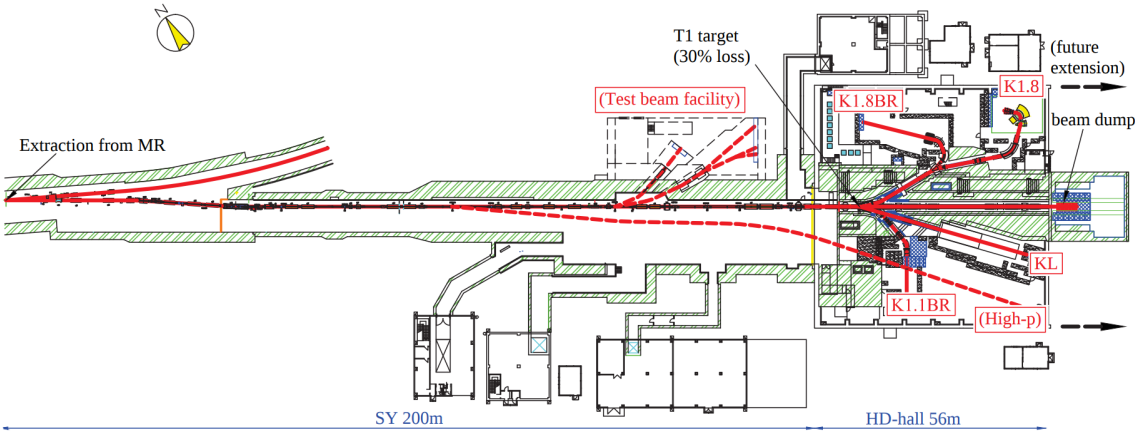


Figure 2.4: Schematic layout of the proton slow extraction from the MR to Hadron Hall. Image courtesy of [42].

2.2.2 Proton Extraction

The extracted protons are transported to the production target in Hadron Hall via the Switch Yard (SY), which is about 200m in length (Figure 2.4). During data collection in 2016–2018, the beam was extracted for approximately two seconds every 5.2-6.0 seconds. A shot of extracted protons in a single acceleration cycle is called a "spill". An important performance measure of the beam extraction is called the spill duty factor defined as,

$$\text{Spill Duty Factor} = \frac{\left(\int_0^T I(t)dt\right)^2}{T\int_0^T I^2(t)dt} \tag{2.2}$$

where $I(t)$ is the beam intensity and T is the extraction period or spill length. If the beam is extracted at a constant rate, the duty factor is one and the time structure of the extracted beam is flat. Thus, a higher duty factor corresponds to less time fluctuations in the beam which is desirable for data collection with the data acquisition system. The reported duty factor from the accelerator from 2016–2018 varied between 35%-55%, except in Run 74/75 in 2017 where there was an issue with the beam extraction (see Section 4.2) and the duty factor was occasionally as low as 16%.

2.2.3 Hadron Experimental Facility

The KOTO experiment is located in the Hadron Experimental Facility, or Hadron Hall, along with other nuclear and particle physics experiments (Figure 2.5). The hall itself is 60 m by 56 m. The protons that are extracted from the main ring are delivered to Hadron Hall and hit the T1 production target, creating a spray of particles including kaons, pions, and antiprotons that are used to form secondary beams for the various experiments. The T1 target used in the 2016–2018 runs is shown in Figure 2.6. It is composed of two gold bars each 66 mm long that sit atop a copper block with a water cooling system to prevent overheating. The proton beam is only incident on one of the gold bars at a time and the target can be moved horizontally with a mechanical driver so that switching between the gold bars is possible. About 50% of

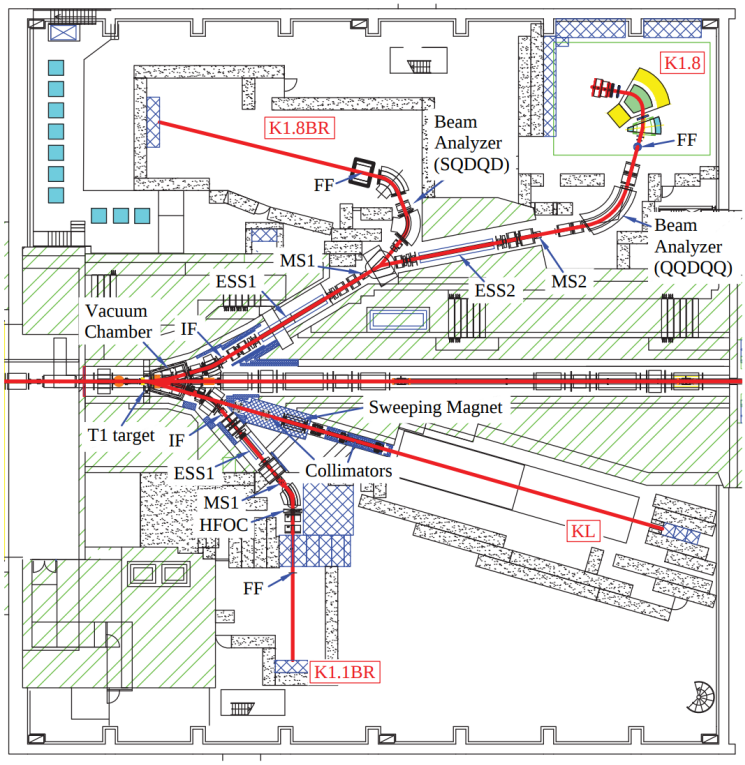


Figure 2.5: Schematic view of Hadron Hall. The extracted proton beam enters the hall from the left and hits the T1 target to produce secondary beams of particles that are used by the various experiments. The location of the KOTO experiment is indicated by the "KL" area. Image courtesy of [42].

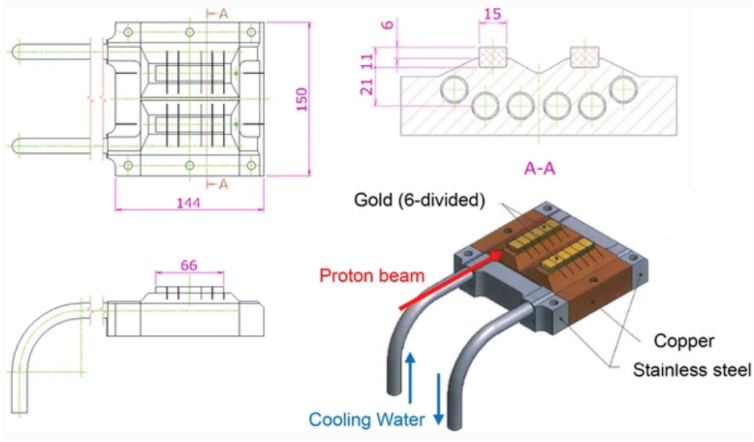


Figure 2.6: Structural drawings of the T1 target (in mm). Image courtesy of [43].

the protons interact in the target, and from these interactions two charged secondary beams (K1.8 and K1.8BR) and a neutral K_L beam are produced [43].

2.2.4 K_L Beam Line

At the T1 target, only about one out of every 10^5 particles created is a long-lived neutral kaon. Even so, the J-PARC facility produces the highest intensity neutral kaon beam in the world. The kaons that are produced are delivered from the T1 target to the KOTO detectors via the K_L beam line (Figure 2.7). The K_L beam line is 21 m long and is shifted 16° with respect to the primary proton beam line. This shift was optimized to reduce the number of high energy neutrons while keeping the kaon flux as high as possible. Another requirement is that the beam is highly collimated with a very small diameter because in order to fully reconstruct the π^0 decay position, we must use the beam axis as a constraint. This is achieved using collimators to produce a narrow "pencil beam" with square dimensions of 8×8 cm^2. A sweeping magnet removes charged particles in the beam, and because the long-lived neutral kaon has a lifetime about 100 times longer than most other particles, all short lived neutral particles decay by the time the beam reaches the KOTO detectors, leaving only K_L^0s, neutrons, and photons in the beam.

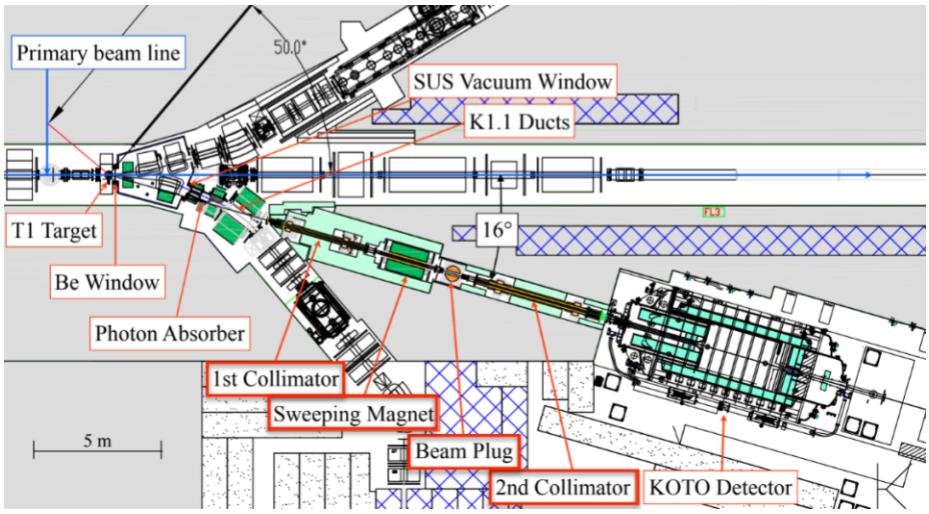


Figure 2.7: The K_L beam line in Hadron Hall. Image courtesy of [44].

The K_L beam line consists of a photon absorber, two collimators, a sweeping magnet, and a beam plug (Figure 2.8). The photon absorber is a 7 cm lead block placed in the beam to reduce photons produced at the T1 target. The two collimators are composed of iron with tungsten alloy at their edges. The first collimator is 4 m long and the second is 5 m long and they are placed 4.5 m away from each other along the beam line. The collimators are designed to reduce the beam diameter to 8×8 cm^2 at the end of the second collimator as well as to reduce "halo neutrons" (neutrons outside the beam width) that are produced by multiple scattering of the neutrons

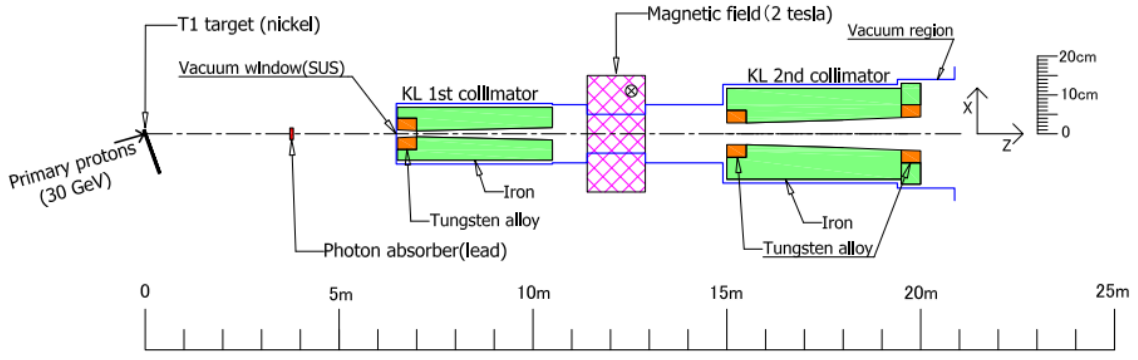


Figure 2.8: Schematic view of the K_L beam line components. Image courtesy of [45].

in the beam. These neutrons in the beam halo can produce a serious background because they interact with detector materials and fake a signal event. A 2 Tesla dipole magnet located between the collimators is used to sweep out charged particles from the beam, leaving only neutral particles. A beam plug placed after the sweeping magnet can be "closed" to prevent the beam from entering the KOTO detectors. We define the end of the beam line, or the downstream end of the second collimator, as the "beam exit".

It is important to know how many kaons are in the beam and the kaon momentum spectrum, as well as the amount of neutrons and photons in the beam that can contribute to background. In 2010 the K_L flux and momentum spectrum of the beam were measured by counting $K_L^0 \to \pi^0\pi^+\pi^-$ decays with a simple setup before the KOTO detectors were installed. The momentum spectrum was measured more accurately in 2012 using $K_L^0 \to \pi^0\pi^+\pi^-$ and $K_L^0 \to \pi^+\pi^-$ decays and gave similar results. From these measurements, the mean K_L momentum was determined to be 2 GeV/c with the peak at 1.4 GeV/c [46] and the distribution is shown in Figure 2.9.

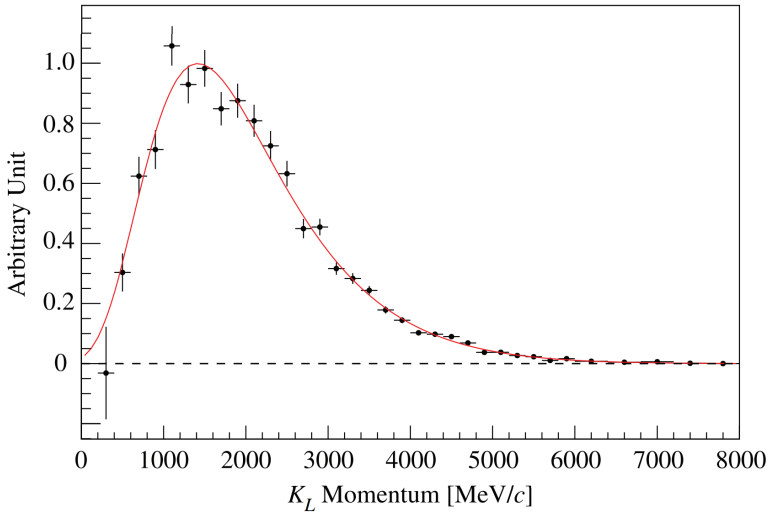


Figure 2.9: The K_L momentum spectrum at the beam exit measured in 2012 (black) with a Gaussian fit (red). The spectrum below 2000 MeV/c was measured with $K_L^0 \to \pi^+\pi^-$. The spectrum above 2000 MeV/c was measured with $K_L^0 \to \pi^0\pi^+\pi^-$ and $K_L^0 \to \pi^+\pi^-$. Image courtesy of [44].

Particles	Flux per 2×10^{14} POT
γ (> 1 MeV)	2.23×10^{9}
γ (> 50 MeV)	1.38×10^{7}
neutron (> 1 MeV)	1.40×10^{9}
neutron (> 1 GeV)	1.61×10^{8}
K_L	4.19×10^{7}

Table 2.1: Neutron and photon fluxes from beam line simulations [47]. The K_L flux was measured in 2012.

The K_L flux is defined in units of "2×10^{14} Protons on Target (POT)" which is the J-PARC design value for the number of Protons on Target per spill. The K_L flux measured in 2010 was 4.19×10^{7} per 2×10^{14} POT. The neutron flux and photon flux in the beam core were estimated using beam simulations and are summarized in Table 2.1. The ratio of halo neutrons to the beam core neutrons is well suppressed to the order of 10^{-5}. Figure 2.10 shows the simulated beam profile for kaons, neutrons, and photons.

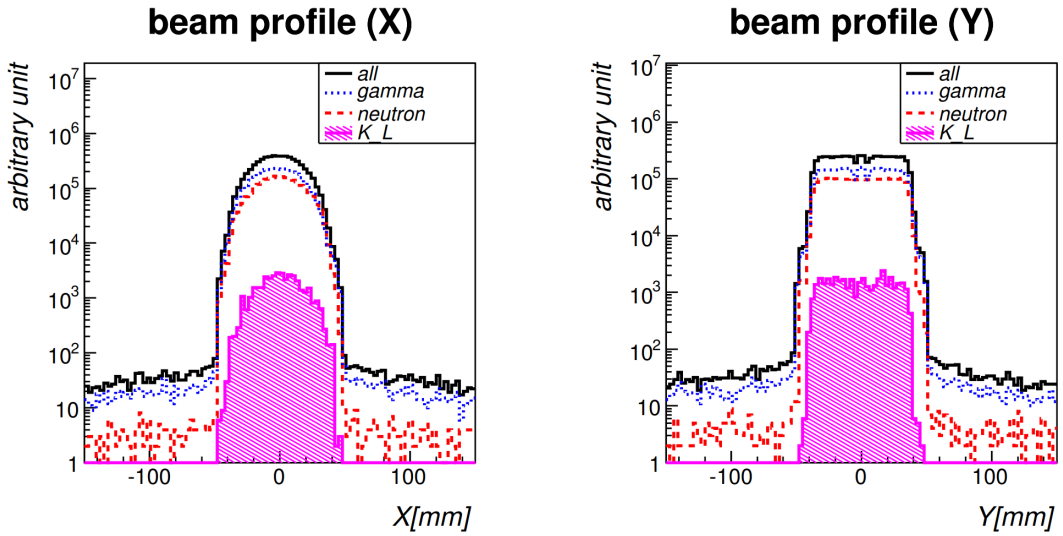


Figure 2.10: The beam profile of kaons, neutrons, and photons simulated using Geant3. Image courtesy of [46].

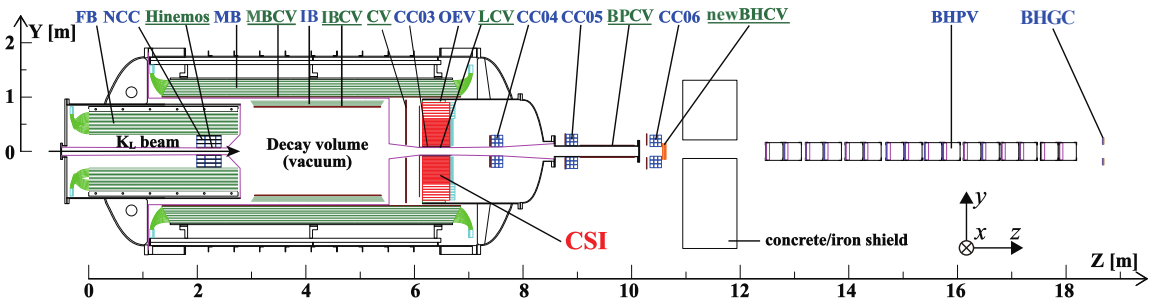


Figure 2.11: Schematic view of the KOTO detectors. The CsI calorimeter is shown in red. The detectors with green labels are charged particle veto counters and the detectors with blue labels are photon veto counters. The magenta lines indicate the vacuum membrane.

2.3 KOTO Detectors

The KOTO detectors are located at the end of the K_L beam line and collect the decay products from the kaons that decay inside the decay volume. Figure 2.11 shows a side view of the KOTO detectors used during the 2016–2018 data collection. The main detector is the CsI calorimeter (red) and the others are veto detectors in place to reject events with other particles present. Charged particle veto counters reject events with charged particles in them and photon veto counters reject neutral decays. It is essential that no particle (except for neutrinos) goes undetected in the KOTO experiment, so specific veto detectors are in place to catch any particles that might escape. A right-handed, Cartesian coordinate system is used, with the positive z direction pointing downstream where $z = 0$ starts at the upstream edge of the Front Barrel (FB). The x and y axes are in the horizontal and vertical directions, respectively. Most of the detector components are inside an external vacuum vessel which is shown in Figure 2.12. The decay volume is kept in vacuum at 10^{-5} Pa to suppress π^0 production by the interaction of neutrons in the beam with residual gas in the volume. Table 2.2 gives a summary of the different detectors and the following sections in Chapter 2 give details on each detector.

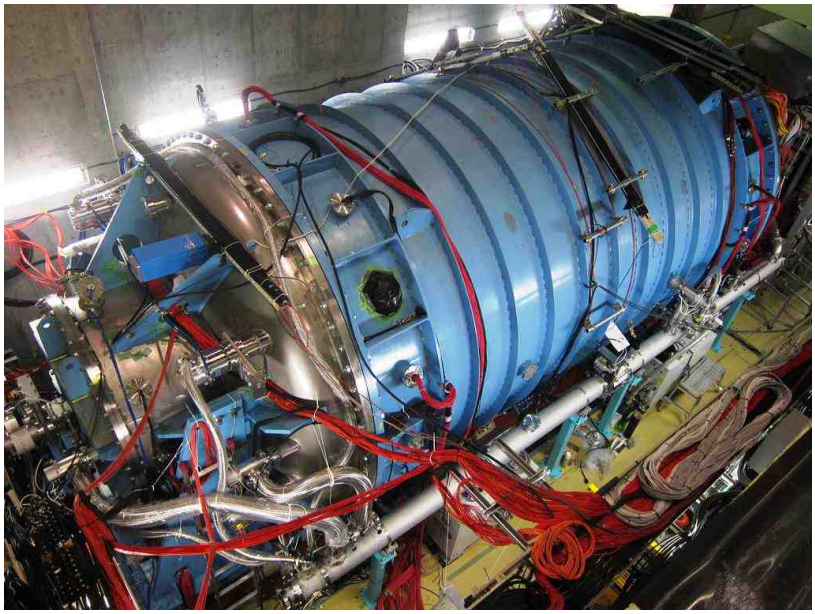


Figure 2.12: The external vacuum vessel houses most of the KOTO detectors.

Detector	Upstream surface z position (mm)	Module length (cm)	Readout	Material	Type
FB	0	275	single-end	scint.	photon veto
NCC	1990	44.5	single-end	cryst.	neutron veto
HINEMOS	1990	44.5	single-end	scint.	photon veto
MB	1355	550	dual-end	scint.	photon veto
MBCV	1355	550	single-end	scint.	charged veto
IB	2994.5	267.6	dual-end	scint.	photon veto
IBCV	2994.5	267.6	dual-end	scint.	charged veto
CV	5842	25.4	dual-end	scint.	charged veto
CsI	6168	50	single-end	sryst.	main photon detector
OEV	6168	42	single-end	scint.	photon veto
LCV	6168	57.5	single-end	scint.	charged veto
CC03	6168	50	single-end	cryst.	photon veto
CC04	7415	See Fig. 2.23	single-end	cryst. & scint.	photon & charged veto
CC05	8793	See Fig. 2.23	single-end	cryst. & scint.	photon & charged veto
CC06	10338	See Fig. 2.23	single-end	cryst. & scint.	photon & charged veto
BPCV	8793	100	single-end	scint.	charged veto
new BHCV	10571.7	2.79	single-end	wire	charged veto
BHPV	12077	33.2	dual-end	Cerenkov	photon veto
BHGC	18235	See Fig. 2.27	dual-end	Cerenkov	photon veto

Table 2.2: Summary of the different KOTO detectors. Detectors made from scintillator (scint.) and CsI crystals (cryst.) are listed as such in the table.

2.3.1 CsI Calorimeter

The CsI electromagnetic calorimeter (often simply called "the CsI") is made of 2716 undoped Cesium Iodide crystals stacked in a circular pattern and is shown in Figure 2.13. There are 2240 small crystals which have a cross-sectional area of 2.5×2.5 cm^2 and 476 large crystals with a cross-sectional area of 5×5 cm^2. The smaller crystals are stacked in the central 1.2×1.2 m^2 region around the beam hole and the larger crystals surround the peripheral region. A stainless steel cylinder with a diameter of 1.9 m holds the crystals and the 20×20 cm^2 beam hole at the center allows the other beam particles to pass through without interaction.

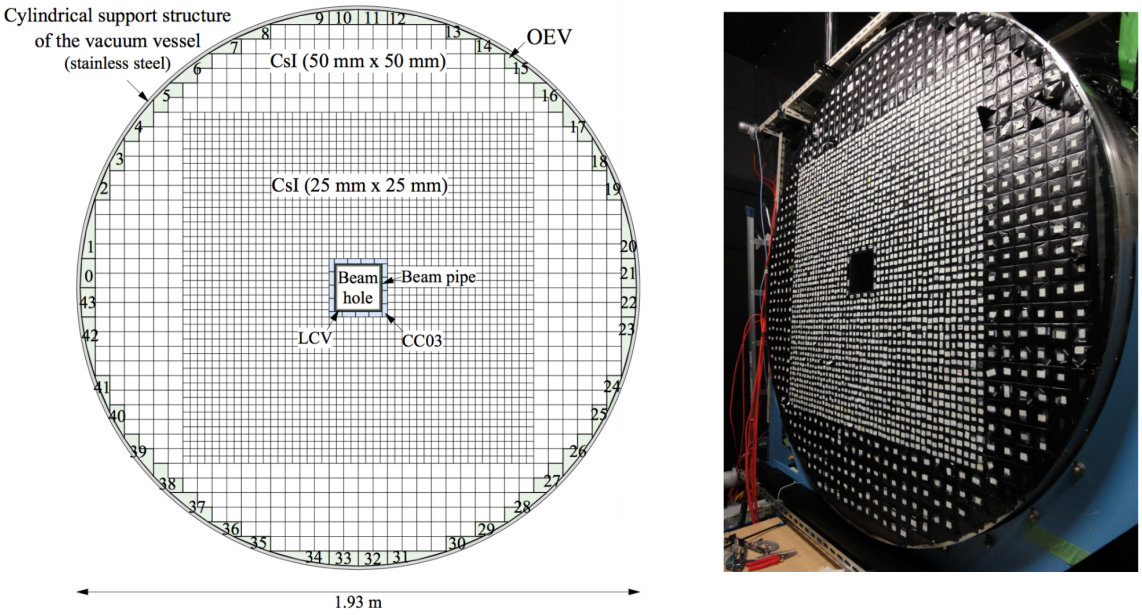


Figure 2.13: Schematic front view of the CsI calorimeter (left) and a picture of the detector during installation (right).

The CsI crystals are 50 cm long which corresponds to 27 radiation lengths (X_0) and the light transmitted through the crystals are read out by photomultiplier tubes (PMTs) mounted on the downstream end of each crystal.[1] By using such long crystals the problem of photon punch through and shower leakage is minimized. Each

[1]The CsI crystals and PMTs were recycled from the KTeV experiment [48].

crystal is also wrapped in a 13 μm layer of aluminized mylar so that they are optically isolated from each other. Two types of PMTs are used for the different sized crystals– Hamamatsu 3/4 inch PMTs (R5364) are used for the small crystals and Hamamatsu 1.5 inch PMTs (R5330) are used for the large crystals. A 4.6 mm scilicone cookie is placed between the crystal and the PMT to improve the efficiency of light transmission. The CsI calorimeter has an excellent energy resolution (σ_E) of $\sigma_E/E = 0.99\%/\sqrt{E_{\mathrm{GeV}}}$ where E_{GeV} is the incident energy of the electron in GeV. The timing resolution (σ_t) of the CsI calorimeter is $\sigma_t/E = 0.13/\sqrt{E_{\mathrm{GeV}}}$ ns and the position resolution (σ_d) is $\sigma_d/E \sim 2.5/\sqrt{E_{\mathrm{GeV}}}$ mm [49].

2.3.2 Front Barrel and Main Barrel

The Front Barrel (FB) and Main Barrel (MB or CBAR) are photon veto counters that surround the upstream region and decay volume, respectively. Both the FB and the MB are composed of trapezoidal modules made of layers of scintillator sandwiched between layers of lead. The FB has 16 modules that are 2.75 m long and the MB has 32 modules that are 5.5 m long. In both cases the modules are arranged to surround a cylindrical volume, as shown in Figures 2.14 and 2.15. The lead layers increase the probability of photon interaction in the module to create electromagnetic showers. When electrons are excited in the scintillator from an interaction, ultraviolet light is produced. The layers of scintillators have small grooves that hold wavelength shifting (WLS) fibers that convert the ultraviolet light to visible light which is collected by PMTs at the ends of the modules. The FB modules are read out by PMTs on the upstream end of the detector while the MB modules have dual-end readout with PMTs on both ends of the modules. The FB modules have a total radiation length of 16.5 X_0 while the MB modules have a total radiation length of 14.0 X_0.

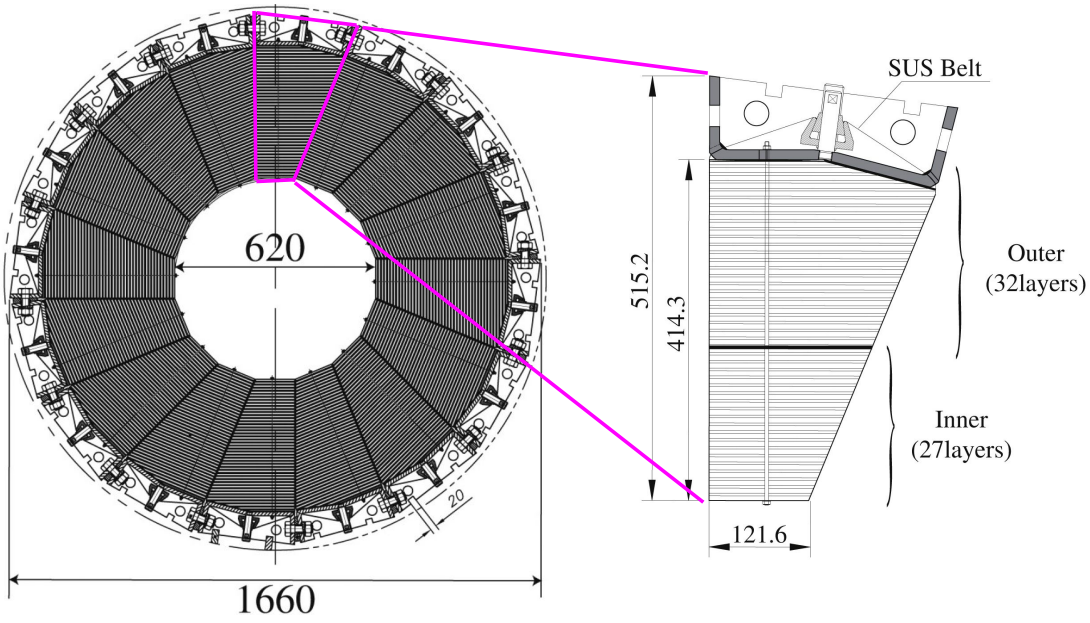


Figure 2.14: Cross-sectional view of the FB detector (left) and an individual module (right). Units are in mm. Figure modified from [50].

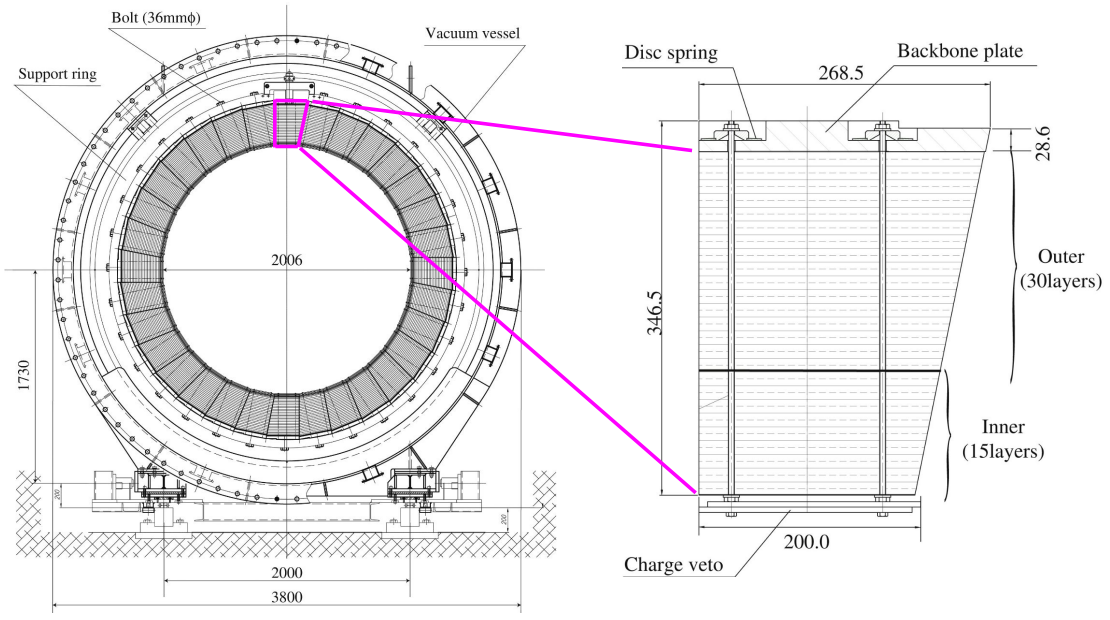


Figure 2.15: Cross-sectional view of the MB detector in the vacuum vessel (left) and an individual MB module (right). Units are in mm. Figure modified from [50].

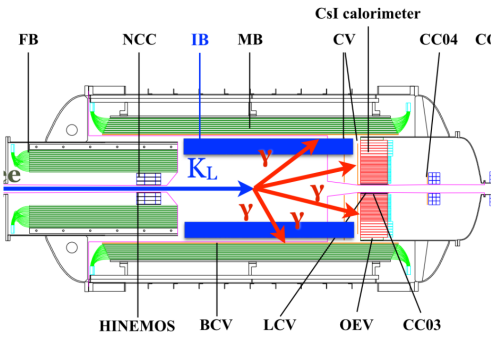


Figure 2.16: Side view of the KOTO detectors with the IB shown in blue. The $K_L^0 \to 2\pi^0$ background is also depicted. Image courtesy of [51].

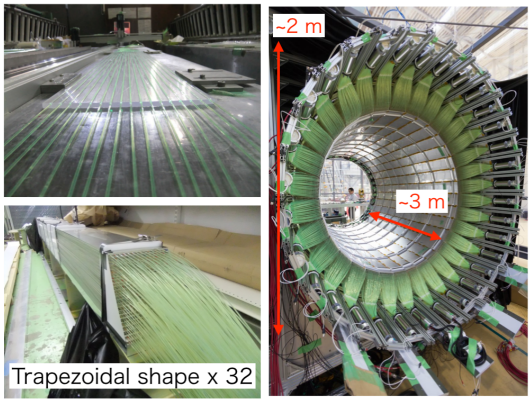


Figure 2.17: The WLS fibers in a scintillator layer (top left), a trapezoidal module (bottom left), and the IB before installation (right).

2.3.3 Inner Barrel

The Inner Barrel (IB) is an additional photon veto counter that was added before the 2016–2018 runs in order to increase the depth of detectors in the barrel region. It has a similar design as the FB and MB detectors. The installation of the IB added an additional five radiation lengths (X_0) to reduce the $K_L^0 \to 2\pi^0$ background, which was one of the major backgrounds in the 2015 data set. In this case, two of the four photons are detected by the CsI calorimeter while the other two photons are not seen in the barrel region due to the inefficiency in the MB (Figure 2.16).

Photos of the IB can be seen in Figure 2.17. The IB is composed of 32 trapezoidal modules of lead and scintillator with WLS fibers and has dual-end read out with PMTs on both ends of each module. The IB is 3 m long and has an inner diameter of 1.5 m. Based on Monte Carlo (MC) estimations, the IB reduces the amount of $K_L^0 \to 2\pi^0$ background by a factor of three [51].

2.3.4 Main Barrel Charged Veto and Inner Barrel Charged Veto

In order to detect charged particles in the barrel region, the Inner Barrel Charged Veto (IBCV) and Main Barrel Charged Veto (MBCV) are in place. The IBCV and

MBCV are made of 10 mm thick plastic scintillator modules with WLS fibers that surround the inner surface of the IBCV and MBCV. The IBCV consists of 32 scintillator modules arranged in a cylindrical shape and the signals from each module are read from both ends with PMTs. The 32 modules of the MBCV cover the barrel region not covered by the IBCV (see Figure 2.18) and are read out at the downstream end with 16 PMTs in which one PMT reads two MBCV modules.

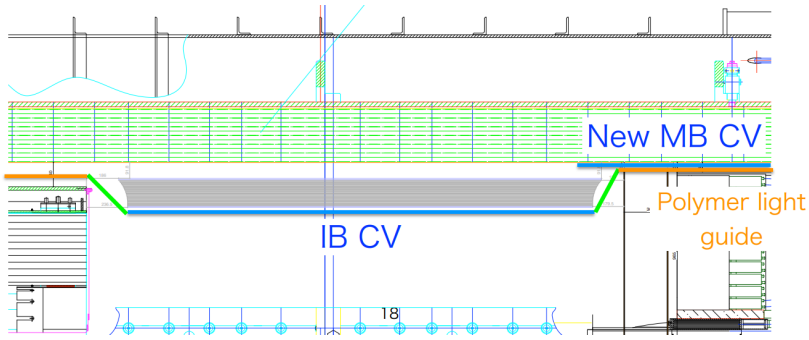


Figure 2.18: Schematic cross section of the IBCV and MBCV detectors (blue). The MBCV is shown as New MBCV in the figure.

2.3.5 Neutron Collar Counter and HINEMOS

The Neutron Collar Counter (NCC) is located inside the FB at the entrance of the decay volume and it is designed to reject K_L^0 decays that occur in the upstream region of the FB as well as to suppress the upstream-π^0 background which occurs when halo neutrons in the beam hit detector material upstream and produce π^0s. Another important purpose of the NCC is to measure the flux and spectrum of halo neutrons in the beam. The NCC is made of undoped CsI crystals and has 48 inner modules and eight additional outer modules. Each module is divided into three optically separate regions (front, middle, and rear) which is used to distinguish neutrons from other particles from K_L^0 decays for the halo neutron flux measurement. The inner modules are read out using WLS fibers, and the outer modules are read out with PMTs. The Horizontal Inner NCC Edge Mounted Scintillator (HINEMOS) is a charged particle

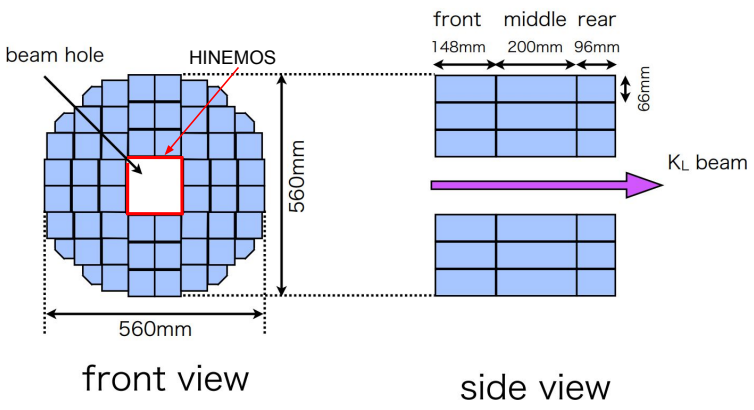


Figure 2.19: Schematic view of the NCC and HINEMOS. Figure modified from [52].

veto counter mounted on the inside of the NCC beam hole and is made of scintillator with WLS fibers. The NCC and HINEMOS detectors are shown in Figure 2.19.

2.3.6 Charged Veto

The Charged Veto (CV) is the main charged particle veto counter and is located in front of the CsI calorimeter. By identifying if the CV has a hit or not, we can identify if a particle that hit the CsI is neutral (no hit in CV) or charged (CV hit). The CV consists of two layers, one placed 5 cm upstream from the calorimeter and one placed 30 cm upstream from the calorimeter. Both layers are composed of plastic scintillator modules arranged in four quadrants (Figure 2.20) and the front and rear layers of the CV are arranged such that the x and y position of the hit can be uniquely identified. Each scintillator module is 3 mm thick and 69 mm wide and the front CV has 48 modules while the rear CV has 44 modules. The scintillation light is collected by WLS fibers in the modules and then read out via Multi-Pixel Photon Counters (MPPCs) attached to both ends of the CV modules [53].

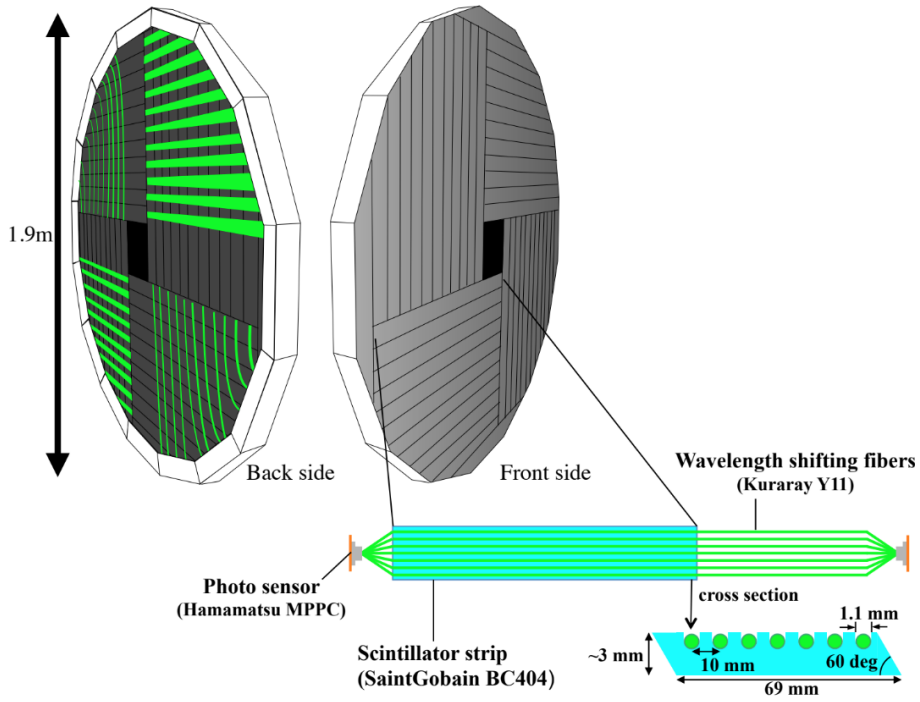


Figure 2.20: Schematic view of the CV detector. Figure courtesy of [54].

2.3.7 Collar Counter 3 and Liner Charged Veto

The Collar Counter 3 (CC03) is between the inside surface of the CsI crystals in the beam hole and the beam pipe. Its purpose is to veto K_L^0 decays that happen close to the CsI calorimeter. CC03 is made of undoped CsI crystals and contains 16 modules. The crystals have a cross section of 45.5×18 mm^2 and are 500 mm long and each crystal is read out with two PMTs at the downstream end. The Linear Charged Veto (LCV) is located on the inner surface of the beam pipe and it is designed to detect charged particles that hit inside the beam hole. The LCV is made of four plastic scintillator sheets with WLS fibers embedded. Both the LCV and CC03 are shown in Figure 2.21.

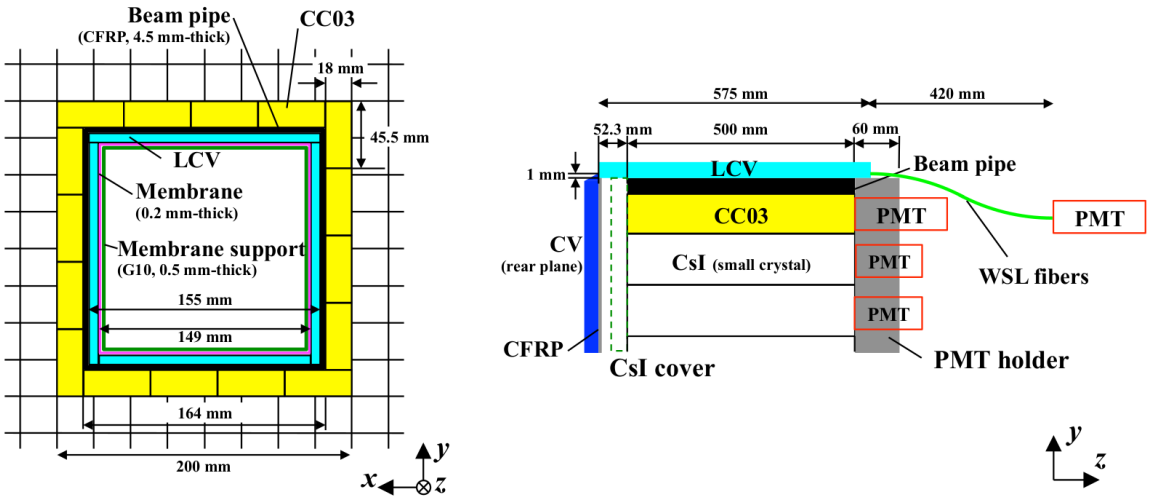


Figure 2.21: Schematic view of the LCV and CC03 looking downstream (left) and from the side (right). Figure courtesy of [54].

2.3.8 Outer Charged Veto

The purpose of the Outer Charged Veto (OEV) is to cover the gap between the edge of the CsI calorimeter and the cylindrical support structure of the external vacuum vessel (this can be seen in Figure 2.13) to ensure that no particles escape through this region. The OEV consists of 44 modules made of alternating layers of lead and scintillator. The scintillator sheets contain WLS fibers and are read out by PMTs. Figure 2.22 shows the OEV detector location and an OEV module.

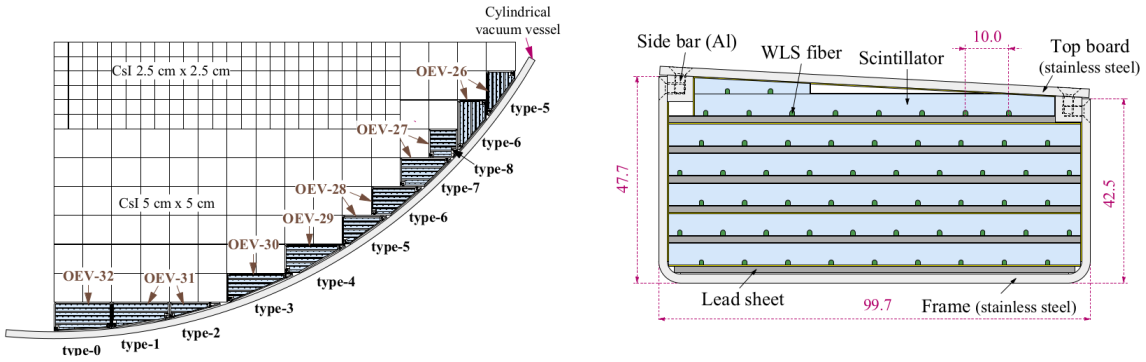


Figure 2.22: The OEV detector. The left figure shows a zoomed-in portion of Figure 2.13 highlighting the OEV location. The right figure shows an OEV module. Figure courtesy of [55].

2.3.9 Downstream Collar Counters

There are three other collar counters located downstream from the calorimeter in the beam pipe– CC04, CC05, and CC06. They are in place to catch photons and charged particles that escape through the beam hole in the calorimeter. The collar counters are constructed from undoped CsI crystals stacked around the beam pipe and there is a layer of plastic scintillator on the upstream side. CC04 is located inside the vacuum tank, while CC05 and CC06 are located outside. Figure 2.23 shows an illustration of the downstream collar counters.

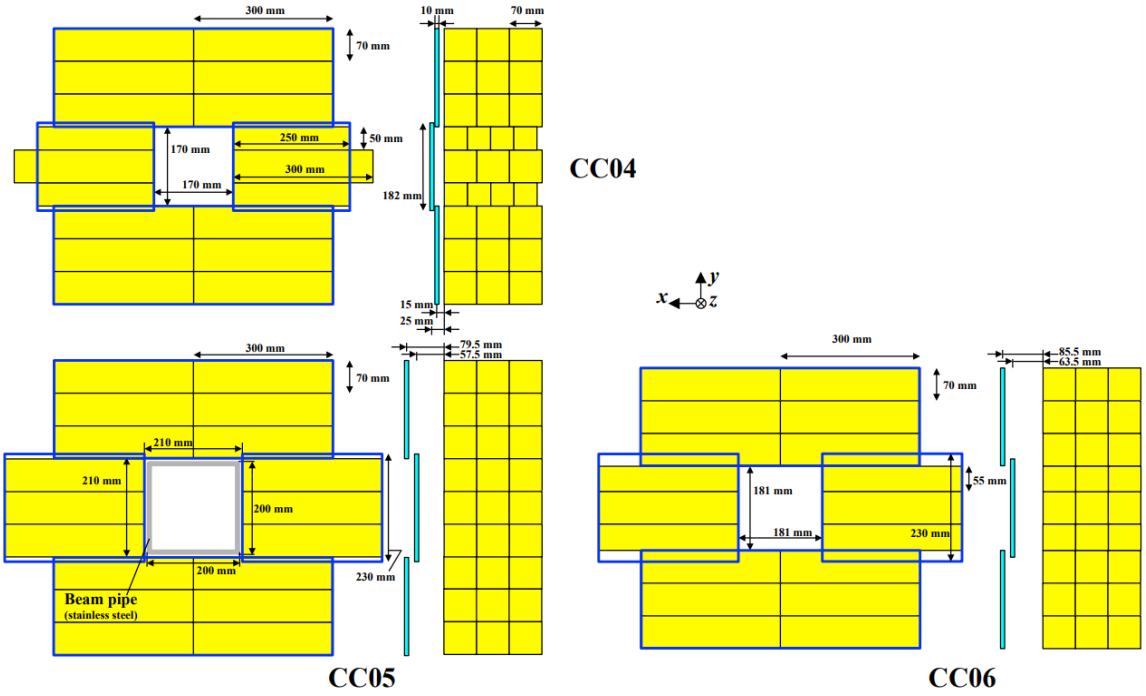


Figure 2.23: The downstream collar counters with a front view (left) and a side view (right). The yellow blocks are the CsI crystals and the blue lines and cyan lines indicate the scintillator modules. Figure courtesy of [47].

2.3.10 Beam Pipe Charged Veto

The Beam Pipe Charged Veto (BPCV) is located between the CC05 and CC06 detectors in the beam pipe and its purpose is to detect charged particles that escape

down the beam hole. In particular, the BPCV was specifically designed to detect charged pions from the $K_L^0 \to \pi^+\pi^-\pi^0$ background. In this case, the two charged pions escape down the beam hole and the π^0 is detected on the CsI with missing momentum which fakes a signal event. To mitigate this background the BPCV was installed after the first physics run in 2013. The BPCV is composed of four plastic scintillator modules that are 5 mm thick with WLS fibers embedded and is shown in Figure 2.24.

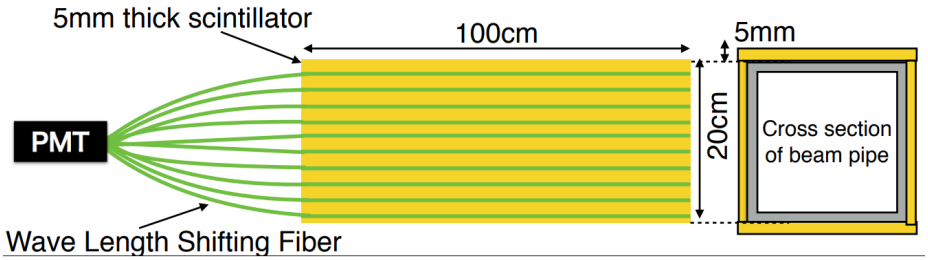


Figure 2.24: A BPCV detector module (left) and the BPCV in the beam pipe (right). Figure courtesy of [47].

2.3.11 Beam Hole Charged Veto

The new Beam Hole Charged Veto (new BHCV) is in place to detect charged particles in the beam in the downstream region. It is located behind the CC06 detector. Because this detector is in the beam with a high flux of K_Ls and neutrons, it must be insensitive to neutral particles. In order to achieve this, the new BHCV is a wire chamber with three layers and uses a mixture of n-Pentane and CF_4 gases as the drift gas. The new BHCV replaced the old BHCV which suffered from a high rate of accidental neutron hits. The new BHCV reduced the hit rate by a factor of three [56]. Figure 2.25 shows details of the new BHCV detector.

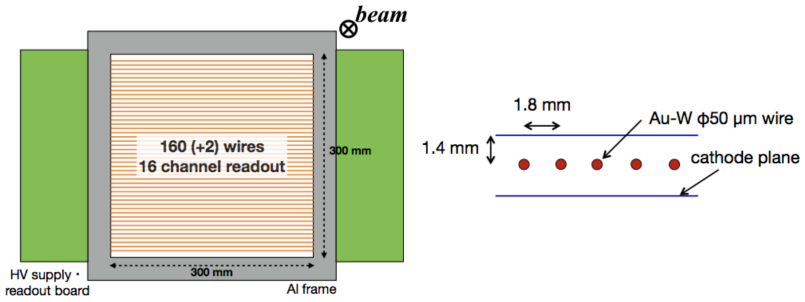


Figure 2.25: A new BHCV detector module (left) and its cell structure (right). Figure courtesy of [56].

2.3.12 Beam Hole Photon Veto

The Beam Hole Photon Veto (BHPV) is located downstream of the concrete/iron shielding and is a Cerenkov detector. Its purpose is to detect photons that escape through the beam hole and it is insensitive to neutrons in the beam. The BHPV consists of 16 lead-aerogel Cerenkov counters placed along the beam axis, which corresponds to a total radiation length of 6.2 X_0. Each BHPV module is composed of a lead converter, an aerogel radiator, light collecting mirrors, and has dual-end PMT readout. Figure 2.26 shows the BHPV detector and individual module structure.

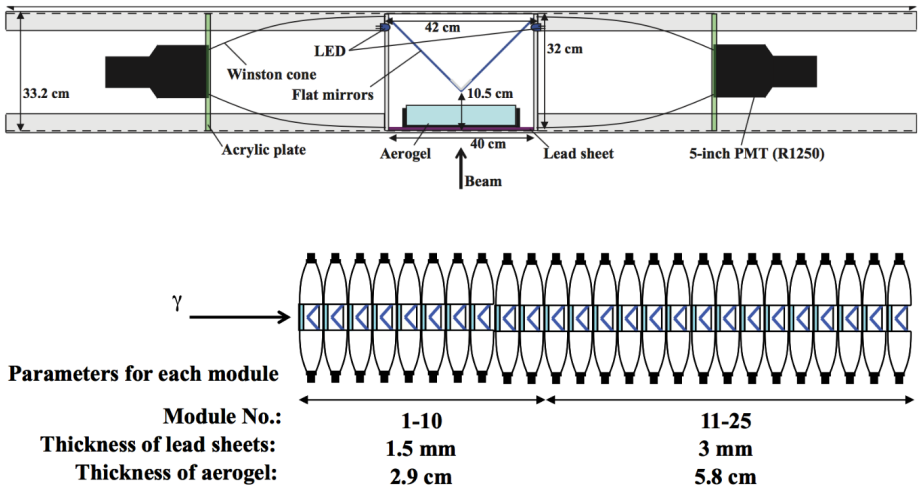


Figure 2.26: A BHPV module (top) and the full BHPV detector (bottom). Note that there are a total of 25 modules shown in the figure but only 16 modules were implemented in the 2016–2018 runs. Figure courtesy of [57].

2.3.13 Beam Hole Guard Counter

The Beam Hole Guard Counter (BHGC) is located at the very end of the KOTO detectors and is in place to detect photons that escape through the BHPV edge region (this can be seen in the top of Figure 2.27). The BHGC is composed of four modules of lead-acrylic Cerenkov counters in which each module is 500×120 mm^2. The lead plate converts the photon to charged particles and the charged particles create Cerenkov radiation in the acrylic plate and the signal is read out from both ends with PMTs. The BHGC is designed such that it is insensitive to neutrons in the beam [58]. Figure 2.27 illustrates the details of the BHGC detector.

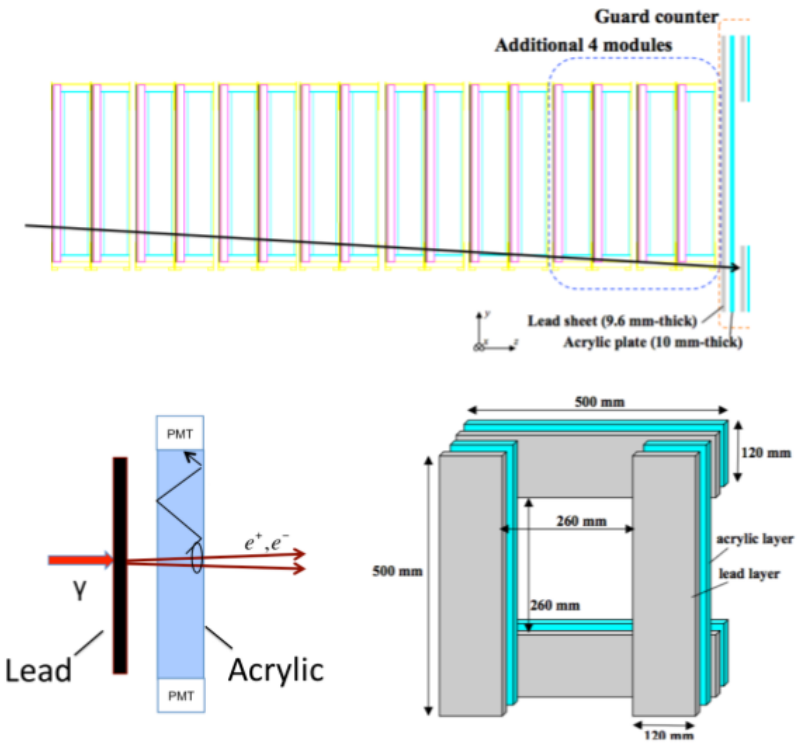


Figure 2.27: The BHGC detector details. Top: the BHGC is located behind the BHPV detector and is in place to catch the photons that go undetected through the BHPV edge regions. The bottom left shows a single BHGC module with a photon hit and the bottom right shows the configuration of the BHGC detector. Figure modified from [54].

2.3.14 Vacuum System

The final piece of the KOTO detector apparatus is the vacuum system. The decay region is evacuated to 10^{-5} Pa to prevent neutrons in the beam from interacting with residual gas in the decay volume and creating additional particles. The vacuum system consists of a high vacuum region kept around 5×10^{-5} Pa and a low vacuum region kept at less than one Pa. The low vacuum region acts as a buffer between the high vacuum region and the outside atmospheric pressure and the different regions are separated by thin films called "membranes". Figure 2.28 shows the vacuum system and the different regions.

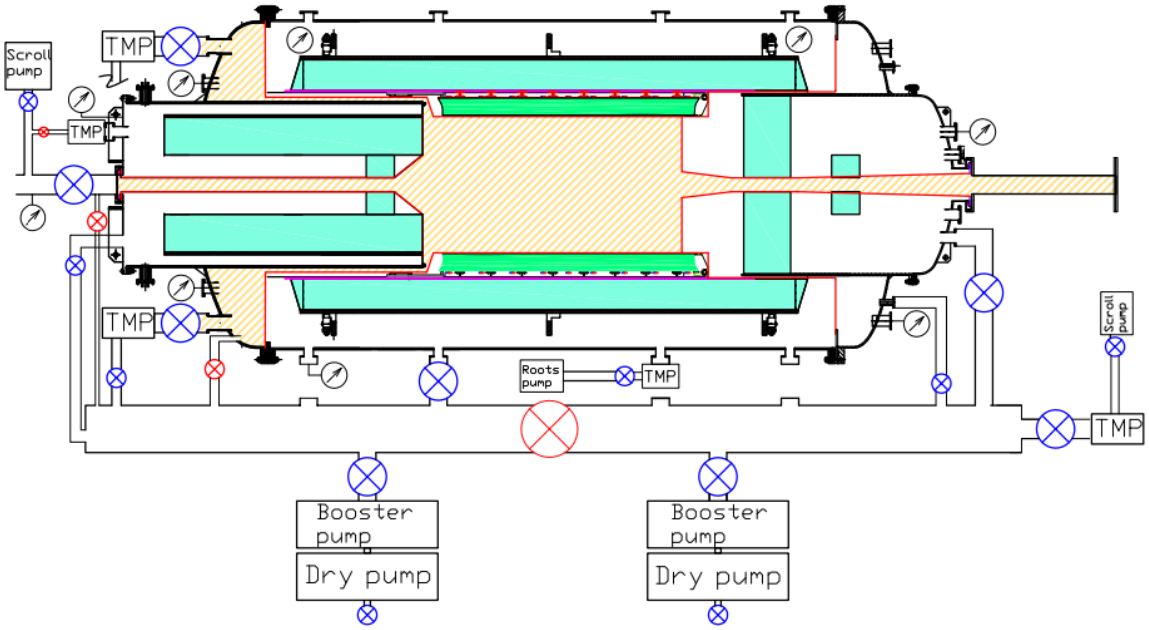


Figure 2.28: Schematic view of the vacuum system. The high vacuum region is indicated by the yellow hatched region and the low vacuum region is the surrounding hollow region within the external vacuum vessel. The detectors in cyan are in the low vacuum region. The red and blue crossed circles are closed and open valves, respectively. The circles with arrows indicate where vacuum monitors are installed.

Chapter 3

Data Acquisition System

In order to collect enough statistics to observe $K_L^0 \to \pi^0\nu\bar{\nu}$, we use the high intensity proton beam to produce the many billions of kaons needed for KOTO. However, processing the signals from thousands of detector channels during high beam intensity is a huge undertaking and requires a highly efficient system to accomplish this task. Because the rates at which we acquire data are so large, it is impossible to store all of the data and so we must selectively record events of interest by making trigger decisions on live data and discarding the rest. This is the main function of the data acquisition (DAQ) system– to collect the data from the detectors, process it, and cut it down to a reasonable amount so that it can be stored for later analysis.

To achieve this, the DAQ system was designed to sustain a high trigger rate with minimum dead time and efficiently and accurately select triggers. The DAQ system must also have good timing resolution to distinguish real events from accidental backgrounds that occur with high beam intensity. This is accomplished by using fast front-end digitization and three levels of triggers. The DAQ components and how they work are described in this chapter and the different system configurations that were used in the 2016–2018 data collection are described in Sections 3.2.2–3.2.5, as the DAQ was upgraded during this time. The performance of the DAQ system in the 2016–2018 runs is covered in Chapter 4.

3.1 Introduction to Triggering

To understand the DAQ system, one must first understand "triggers" or "triggering". In this case, the word "trigger" can mean the system that uses criteria to rapidly decide which events to keep, for example, the Level 1 trigger. It can also mean a single snapshot or instance of what is happening in the detectors– I refer to this as "a trigger". For example, in the KOTO DAQ system, every 8 ns we calculate the energy deposited in all the detectors and we issue a trigger if the energy in the CsI exceeds a certain threshold and there is no energy in the veto detectors. There are different types of triggers, depending on what criteria the DAQ system uses to determine a trigger and this is discussed in Chapter 4. In the KOTO DAQ system, we use Field Programmable Gate Arrays (FPGAs) on most of the electronics boards. FPGAs are integrated circuits that can be reprogrammed to implement different logic functions, which allows for flexible triggering and unlimited modifications to the system.

3.2 DAQ Overview and Upgrades

During the 2016–2018 data collection, we were constantly improving and upgrading the DAQ system to keep up with the high trigger rates that come with increasing beam power. From 2016 to 2018, changes to the system were gradually introduced and a new triggering system was fully functional by the final data collection period. The upgrades are briefly summarized below (Table 3.1) and the details of the different configurations in each run period are described in Sections 3.2.2–3.2.5. The components of the DAQ system are listed in Table 3.2 and the remainder of the chapter gives details on each DAQ component. The data collection period from 2016 to 2018 consisted of four run periods, Run 69 in 2016, Run 74/75 in 2017 (which are grouped together here because they were taken consecutively), and Runs 78 and 79 in 2018 (see Table 3.1). More details about the run periods are given in Chapter 4.

Run Period	Date	DAQ Upgrades
Run 69	6/2016	ADCs, Michigan L1, L2, and L3 triggers used
Run 74/75	4/2017	CDT trigger added and used as L1.5 and implemented software cut in L3
Run 78	1/2018	CDT performs extra online veto
Run 79	6/2018	CDT and OFC system replaces Michigan L1 and L2 trigger systems

Table 3.1: Summary of the upgrades to the DAQ system from 2016 to 2018. The run periods lasted anywhere from one to two months. Precise dates of operation can be found in Table 4.1.

In general, the KOTO DAQ system is composed of two types of front end analog-to-digital converter (ADC) modules and three levels of triggers. The Level 1 (L1) and Level 2 (L2) triggers are hardware triggers and Level 3 (L3) is a software trigger. The original L1 and L2 triggers were designed and built by the University of Michigan KOTO group and are referred to in this chapter as the "Michigan L1 and L2". The DAQ system in Run 69 consisted of the ADCs, the Michigan L1 and L2 triggers, and the L3 trigger[1]. The system is controlled by the MAster Control and TRIgger Supervisor (MACTRIS+) which distributes the clock and keeps track of DAQ diagnostics. In 2017, a new cluster counting trigger called the Clock Distribution and Trigger (CDT) was integrated into the system and we used the CDT as an additional trigger (Level 1.5) in Run 74/75. We also implemented a software trigger cut in the L3 system. In 2018, the DAQ system used in Run 78 was the same as in Run 74/75 except we added an extra online veto performed by the CDT trigger. In Run 79, the Optical Fiber Center (OFC) modules were integrated into the CDT trigger system and the OFC and CDT system replaced the Michigan L1 and L2 triggers which I refer to as the "new L1 and L2 triggers". The details of this new system are covered in Section 3.7. Table 3.2 describes the different DAQ components and in which periods they were used.

[1] Also built by the Michigan group

DAQ Component	Used In Run				Description
	69	74/75	78	79	
ADC modules	✓	✓	✓	✓	Data digitization and compression
Michigan L1 trigger	✓	✓	✓		Energy sum and veto cut
Michigan L2 trigger	✓	✓	✓		Center of Energy (COE) cut
L3 trigger	✓	✓	✓	✓	Event building (CDT 4 cluster veto cut added in Run 74/75)
MACTRIS+	✓	✓	✓	✓	System control, diagnostics, and clock
CDT trigger (L1.5)		✓	✓		Number of clusters cut (online veto added in Run 78; became new L2 trigger in Run 79)
new L1 trigger				✓	Energy sum and veto cut by OFC + CDT system
new L2 trigger				✓	Number of clusters cut by OFC + CDT system

Table 3.2: Summary of the DAQ components and the periods in which they were used. From 2016 to 2018 the CDT trigger system was gradually integrated into the DAQ until finally in Run 79, the CDT and OFC system became the new L1 and L2 triggers.

3.2.1 DAQ System Layout

Figure 3.1 show the location of the different DAQ components inside Hadron Hall and the KOTO counting room (the data collection/control room). The ADC crates sit inside the KOTO experimental area underneath the concrete radiation shielding blocks near the KOTO detectors. Thousands of Ethernet cables from the detectors are fed into the ADCs. The OFC and CDT system are also located within the radiation shielding and the ADCs are connected to the OFC and CDT system via optical fibers as well as Category-5 LVDS Ethernet cables. The full OFC and CDT system was not completed until Run 79. The Michigan L1 and L2 crates sit outside the radiation shielding blocks and the data and signals are sent through optical fibers that pass through the concrete shielding. The L3 trigger system is located in the KOTO counting room and Ethernet as well as optical fibers are used to deliver the data from L2 to L3. Once the data is processed in the L3 trigger, it is sent out to be permanently stored at the KEK Computer Cluster or KEKCC via the SINET4 (Science Information NETwork) network.

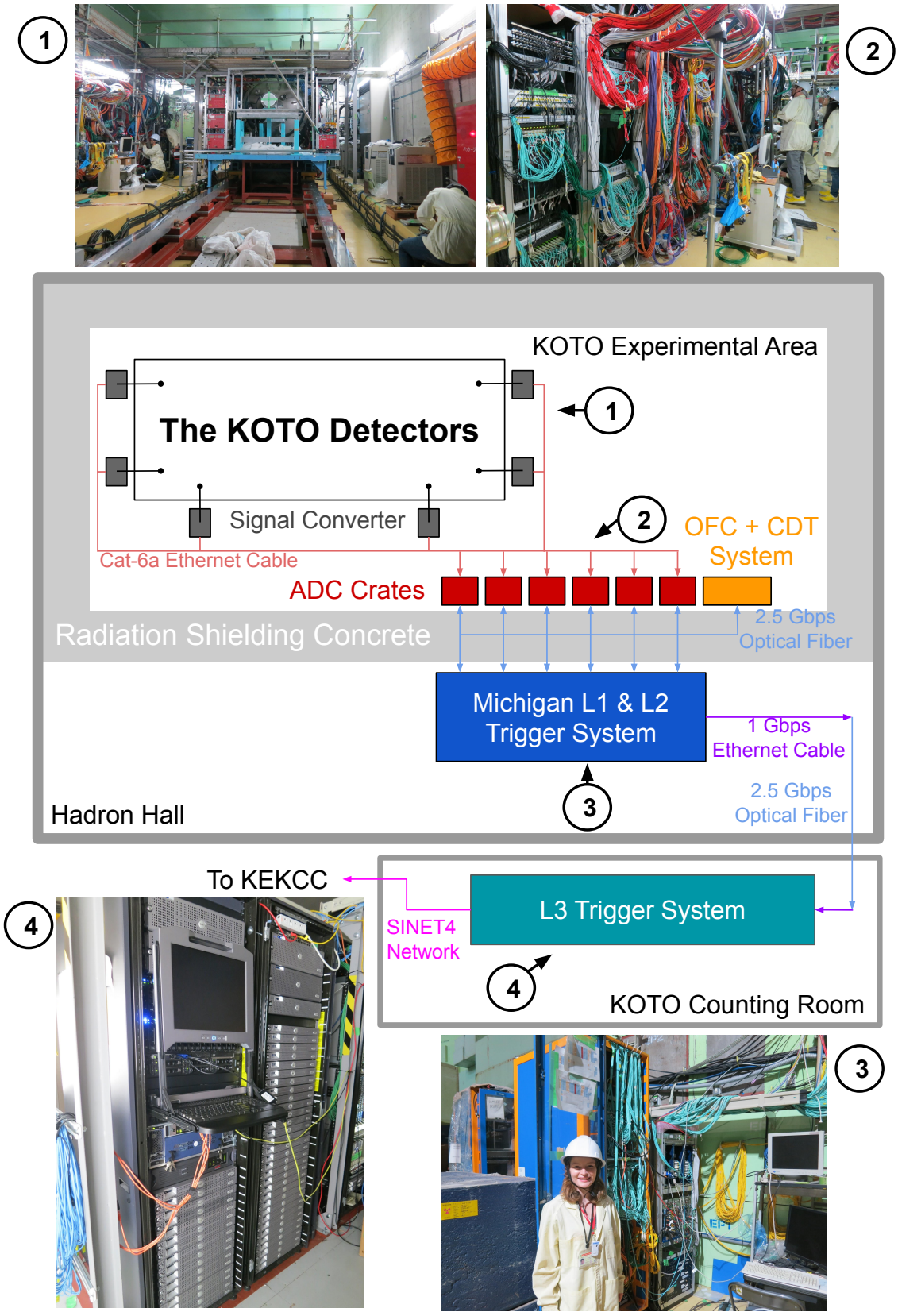


Figure 3.1: Layout of the DAQ system in Hadron Hall (not to scale).

3.2.2 Run 69 DAQ System

The DAQ system used in Run 69 is illustrated in Figure 3.2 and the data flow is described here. The analog waveforms from the detector PMTs come into the ADC modules via Category-6a Ethernet cables. The ADC modules convert the analog signals to a digital data packet and a lossless compression algorithm is used to compress the data (detailed in Section 3.9). The data is then simultaneously buffered in the ADC modules and another copy is sent out to the Michigan L1 trigger via 2.5 Gigabits per second (Gbps) optical fibers. The L1 trigger decision is made based on the total energy deposited in the CsI calorimeter and no activity in the veto detectors. This total energy sum and veto information is sent to MACTRIS+ and if these criteria are met, then MACTRIS+ sends a L1 trigger request to the ADC module where it meets the buffered data right before it falls off the end of the buffer.

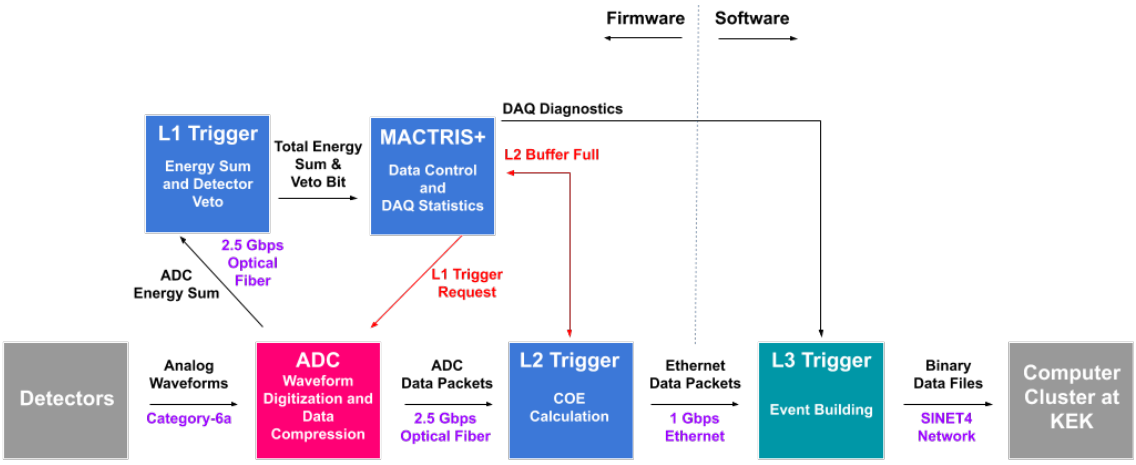


Figure 3.2: The Run 69 DAQ System and data flow. Black arrows and text indicates the data that is transferred between components and purple indicates the connections that allows the data passage. Red shows the system control signals that are sent.

All of the data that is met with a L1 request is sent from the ADC modules to the Michigan L2 trigger via 2.5 Gbps optical fibers. The Michigan L2 trigger makes a trigger decision based on the Center of Energy (COE) position on the CsI calorimeter and the data which makes the cut is sent to the L3 trigger via 1 Gbps Ethernet cabling.

The L3 software trigger receives fragments of events from L2 and uses Infiniband with MPI protocol to build complete events. It should be noted that in Run 69, the L3 was only used for event building and did not make any trigger decisions. After the events are built, the data is stored temporarily on disk arrays before being sent out to be permanently stored at the KEK Computer Cluster (KEKCC). Analysis and data processing that is done in the DAQ chain is referred to as "online" analysis and analysis performed on data that is stored on KEKCC is referred to as "offline" analysis.

3.2.3 Run 74/75 DAQ System

Before data collection in 2017, the DAQ system was upgraded with the addition of the CDT trigger as well as a software cut that was implemented in the L3 trigger. The CDT trigger calculates the number of clusters on the CsI which is used for versatile trigger selections. A "cluster" is defined as energy deposited in a group of CsI crystals from a particle hitting the calorimeter. Using a cluster counting trigger allows us to improve the efficiency of data collection for the signal decay (two clusters), increase the number of $2\pi^0$ and $3\pi^0$ events that are used for normalization analysis (four and six clusters), and eliminate events not used in analysis (odd number of clusters).

The data flow and connections of the DAQ system are very similar to Run 69, though the trigger decision flow was altered to accommodate the CDT. Figure 3.3 illustrates the DAQ chain with the CDT implemented. In this case, the data flow is the same as in Run 69 up until the total energy sum and veto information is sent to MACTRIS+, if the L1 trigger criteria is met, MACTRIS+ sends a L1 pre-accept signal to the ADC module and the ADC starts the clustering mask calculation. This cluster mask information is sent to the CDT trigger via Category-5 LVDS Ethernet cabling and the CDT calculates the number of clusters. The number of clusters is sent to MACTRIS+ and if the CDT trigger criteria are met then MACTRIS+ sends

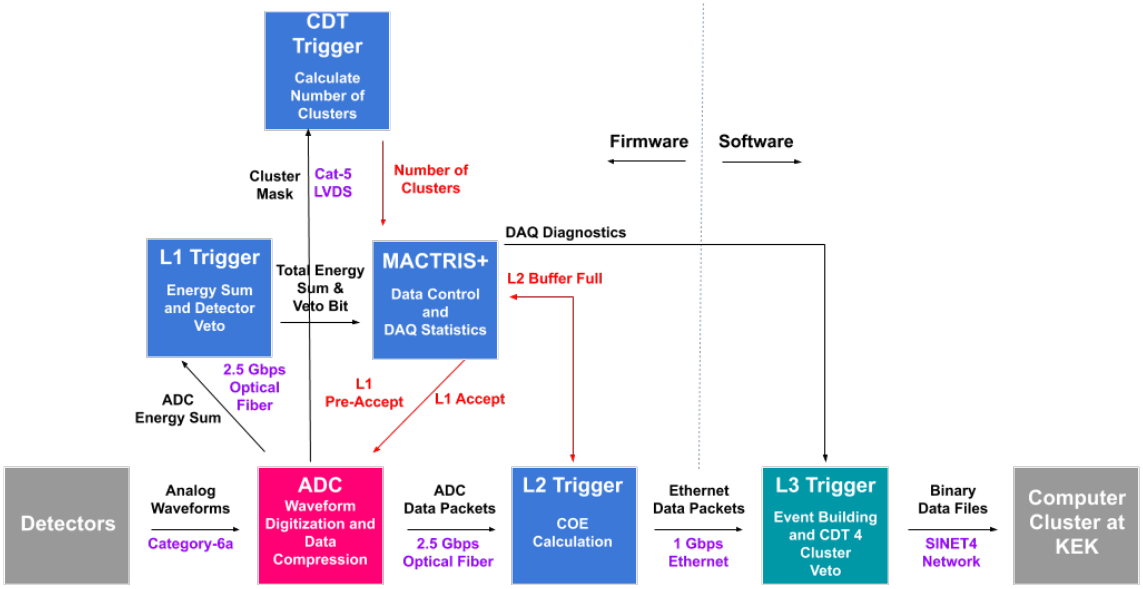


Figure 3.3: The Run 74/75 DAQ System and data flow. Note the addition of the CDT trigger and the L3 software cut. Black arrows and text indicates the data that is transferred between components and purple indicates the connections that allows the data passage. Red shows the system control signals that are sent.

a L1 accept signal to the ACD down the same line as the L1 pre-accept. These two signals are differentiated with a different length clock. As before, the L2 trigger decision is made and the data is passed to L3 which builds the events. In 2017 we also implemented a software cut in L3 on events with four clusters, which is discussed in Section 3.8.2.

3.2.4 Run 78 DAQ System

The DAQ system used in Run 78 is the same as the system used in Run 74/75 except for an upgrade to the CDT trigger. In anticipation of increased beam intensity, we added an extra criteria for the CDT to suppress the high trigger rate by also making a trigger decision on veto detector information as well as the number of clusters. The Michigan L1 trigger decision was still used and is based on the total energy in the CsI and the total energy in each of the veto detectors. We modified the CDT trigger to make an additional and tighter cut on the veto information by

looking at the individual channels of the veto detectors and factoring in the effect of energy calibration. In practice, if a L1 trigger accept signal was issued, the CDT trigger would perform the clustering calculation and also make a judgement on the veto energy information. If the CDT found any hits in the veto detectors that the Michigan L1 was unable to find, the Michigan L1 trigger decision was overridden. The rest of the DAQ system operated the same as before.

3.2.5 Run 79 DAQ System

Finally, in Run 79 we added the Optical Fiber Center (OFC) modules to the CDT trigger system and the OFC and CDT system became the new L1 and L2 triggers. The purpose of this upgrade was to reduce the dead time in the DAQ system, and increase the efficiency of collecting $K_L^0 \to \pi^0\nu\bar{\nu}$ events. Figure 3.4 shows the new data flow. The new L1 trigger is still based on the total energy deposited on the calorimeter and the veto detector information, but the OFC and CDT system is able to make a more accurate and efficient trigger selection. The energy sum and veto information is sent from the ADCs to the OFC and CDT system and the L1 trigger decision is made and sent back to the ADCs by the Top CDT, which controls the system. The Top CDT is responsible for issuing trigger signals to the rest of the system, while MACTRIS+ communicates with the Top CDT to secure the data flow between the different components and ensure there is no data package loss.

The new L2 trigger decision is based on the number of clusters on the calorimeter. The ADCs send partial cluster information to the OFC and CDT system and the number of clusters is calculated and sent to the Top CDT for the new L2 trigger decision. If the conditions are met, a trigger accept signal is sent back to the ADCs to meet the buffered data before it is passed on to the Michigan L2 trigger. In this case, the Michigan L2 trigger does not make any trigger decisions but the hardware remains in the system and the data is simply passed on to the L3 system for event

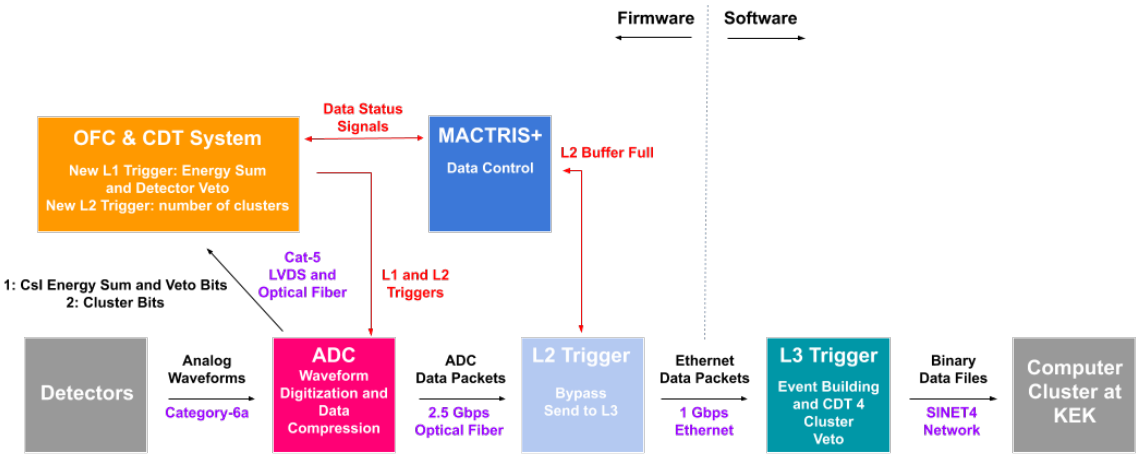


Figure 3.4: The Run 79 DAQ System and data flow. The CDT and OFC system replaces the Michigan L1 and L2 triggers. The new system works together with MACTRIS+ to ensure secure data passage. Black arrows and text indicates the data that is transferred between components and purple indicates the connections that allows the data passage. Red shows the system control signals that are sent.

building and software triggering. More details on the OFC and CDT system are covered in Section 3.7. The rest of the chapter details the different components of the DAQ system.

3.3 MACTRIS+ Control and Fanout System

MACTRIS+ is the brain of the DAQ system. Its purpose is to receive and send the signals from the other DAQ components to tell them how to behave and to make trigger decisions based on these signals. Along with the trigger signals (L1Pre-A, L1A, etc.), it also generates and sends out the 125 MHz (8 ns) system clock (CLOCK), handles DAQ diagnostics and packages and sends this information to the L3 trigger, and generates the logic gate for data collection (LIVE). There are two other flavors of MACTRIS+, the slave[2] MACTRIS+ and the COE MACTRIS+. The hardware on the MACTRIS+ board (Figure 3.5) is the same across the different flavors, but

[2]The use of "master" and "slave" in electrical engineering is antiquated and has racist undertones and so shouldn't be used for future naming schemes. The terms are only kept here for the sake of consistency with other documents. The terms "controller" and "responder" are more appropriate.

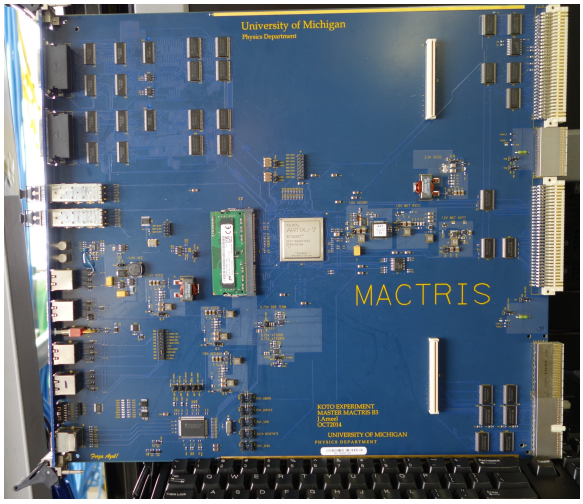


Figure 3.5: The MACTRIS+ control board.

the firmware on the FPGA[3] is programmed to do different tasks. The MACTRIS+ board sits in slot 12 of the Michigan L1 crate and the slave MACTRIS+ boards sit in slot 2 of the Michigan L1 and L2 crates. COE MACTRIS+ sits in slot 12 of the Michigan L2 crate. In this section, I refer only to MACTRIS+'s communication with the Michigan L1 and L2 triggers, not the new L1 and L2 system[4].

MACTRIS+ distributes the control signals to the ADCs via the fanout system, which consists of a master fanout board and two 9U VME crates of slave fanout boards. The signals are sent out via an RJ45 Ethernet port on the front panel of MACTRIS+ (Figure 3.6) and are received by the Ethernet port on the master fanout board (Figure 3.7) which creates 16 identical copies of the signals and distributes them to the slave fanout boards which in turn, create 16 more copies. The slave fanout boards then sends a single copy of the signal to each ADC board via Ethernet cables.

MACTRIS+ also delivers the system clock to the L1 and L2 boards via ribbon cables that connect to the slave MACTRIS+ boards in the L1 and L2 crates. The slave

[3]Xilinx XC5VFX30T

[4]In Run 79 the Top CDT took over the responsibility of issuing trigger signals, generating the clock, and logic gate signaling. MACTRIS+ worked with the Top CDT to secure the data handling.

Figure 3.6: The MACTRIS+ front panel. The system clock is distributed to the L1 and L2 slave MACTRIS+ boards via the ribbon cable ports. The DAQ statistics are output to the L3 trigger via one Ethernet port and another Ethernet port distributes the control signals to the master fanout board.

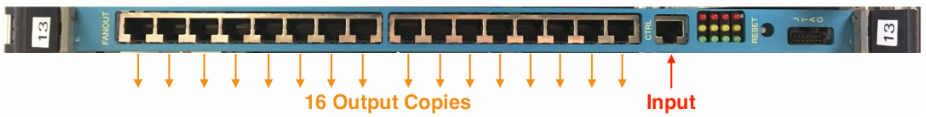


Figure 3.7: The fanout board front panel. The signals are input from MACTRIS+ via the Ethernet port on the right and 16 copies are output via Ethernet.

MACTRIS+ then delivers the clock to the L1 and L2 boards via the P2 backplane. In the L2 crate, the slave MACTRIS+ delivers the clock to COE MACTRIS+ through the front panel ribbon cable. The MACTRIS+ and the COE MACTRIS+ are in charge of the L1 and L2 trigger decisions, respectively. The L1 trigger is made every 8 ns and so no dead time is generated. The L2 trigger on the other hand, receives triggers at a non-fixed clock time and so dead time can be generated if triggers come too close together. For this reason, MACTRIS+ checks for available buffer space in the L2 trigger and if there is no space, MACTRIS+ will not send the L1 trigger signal to the ADCs to prevent the system from becoming overwhelmed. Though all triggers don't make it to the L3 system, we implemented a feature in MACTRIS+ that allows us to save diagnostic information about each L1 trigger regardless of whether it passed the L2 trigger[5]. This includes information like trigger type, timestamp, veto data, etc. that is used to track DAQ performance and troubleshoot errors in the system.

[5]This feature was used in all runs except for Run 79, in which the CDT and OFC system took over DAQ diagnostics.

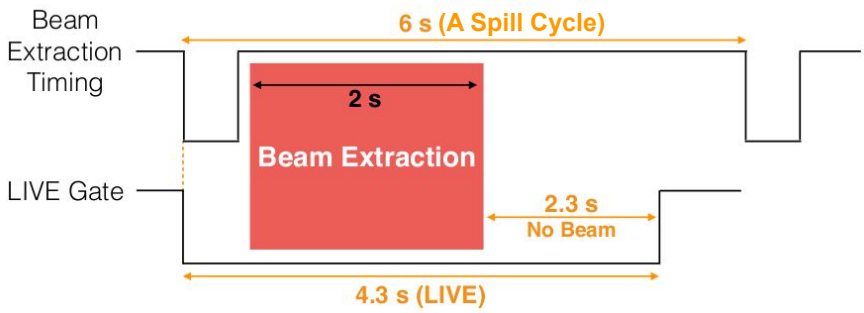


Figure 3.8: The data collection period during a six second spill cycle. The four second LIVE signal from MACTRIS+ tells the rest of the DAQ system to start processing triggers during the two second beam extraction.

Finally, in order to collect data during the beam extraction, MACTRIS+ generates a LIVE signal so that the DAQ system is ready to process triggers when the beam is extracted to Hadron Hall. During a six second spill cycle, the beam is extracted for two seconds (described in Section 2.2.2) and the KOTO DAQ system is LIVE for four seconds as shown in Figure 3.8. The LIVE signal is synchronized with the start of the beam extraction and we keep the system LIVE for four seconds to ensure that no part of the beam is missed. The time during a spill cycle where there is beam extracted is called "on-spill" and the time where there is no beam extracted is called "off-spill".

3.4 ADC Module

The output signals from the detector PMTs are analog waveforms, so the main purpose of the ADC modules (sometimes also called ADC boards) is to convert the analog signals to digital data so that high rate trigger decisions can be made. The analog signals from the PMTs are converted from 50 Ω single-ended signals to differential signals by a 100 Ω differential converter near the vacuum feedthrough for the detectors (this can be seen in Figure 3.1). Output signals from the CsI, CV, BHPV, new BHCV, BPCV, and BHGC already have differential signals so don't need to be

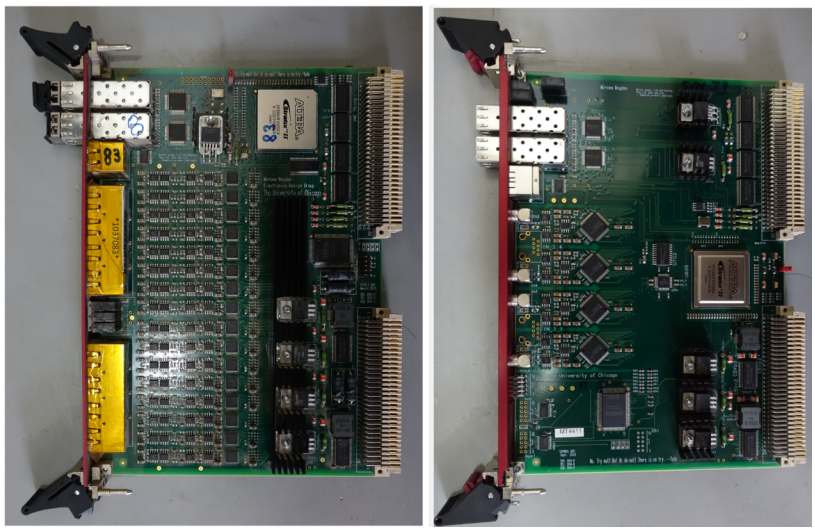


Figure 3.9: Pictures of the ADC modules. The left shows a photo of a 125 MHz ADC board which can take up to 16 detector channel inputs. The right shows a picture of a 500 MHz ADC board which can take up to four detector channel inputs.

converted. We use differential signaling to reduce the effects from noise and pulse widening while transmitting the signals along long lengths of cabling. The signals are transferred to the ADC modules via Category-6a Ethernet cables each with two shielded, twisted pairs. Due to the different detector locations, the cables from the detectors have different lengths and the signals arrive at the ADCs at different times. This timing disparity is accounted for in the FPGA firmware on the ADC boards.

We use two types of ADC modules to receive data from the detectors– 14-bit ADCs with a 125 MHz sampling rate and 12-bit ADCs with a 500 MHz sampling rate (Figure 3.9). Most of the detector signals are digitized with the 125 MHz ADCs but detectors that are close to the beam line have a higher event rate and so we use the 500 MHz ADCs for these detectors that require a faster sampling rate. In the 2016–2018 runs a total of 3864 detector channels were connected to the ADC modules. These boards were housed in 6U VME crates which can hold up to 16 ADC modules. In the 2016 runs, there were 17 VME crates and in 2017-2018 we added an additional VME crate to accommodate the CDT trigger. Table 3.3 lists the detector

channels and the corresponding number of ADC modules that were used.

Detector	Number of Channels	ADC Modules	
		125 MHz	500 MHz
FB	32	2	
NCC	204	13	
MB	128	8	
MBCV	16	1	
IB	64		16
IBCV	64	4	
CV	184	12	
CsI	2716	176	
OEV	44	3	
LCV	4	1	
CC03	32	2	
CC04	64	4	
CC05	60	4	
CC06	60	4	
BPCV	4	1	
new BHCV	48		12
BHPV	34		9
BHGC	8		2
old BHCV (not used)	8		2
old BCV (not used)	64	4	
Other	26	2	

Table 3.3: Detector channels and the corresponding ADC modules.

3.4.1 125 MHz ADC

The 125 MHz ADC modules digitize the waveform signals with a sampling rate of 8 ns, which means that we record the pulse height in ADC counts of the waveform every 8 ns as it comes into the ADCs. The 8 ns clock is provided by the MACTRIS+ system clock[6]. However, the leading edge of the PMT waveform is originally too fast for an 8 ns sampling, so before digitization we use a 10-pole Bessel filter to spread out and shape the pulse into a Gaussian waveform. After the waveform goes through the 10-pole filter it has a spread of about 64 ns. Then when we digitize the waveform a 1

[6]Provided by the Top CDT in Run 79

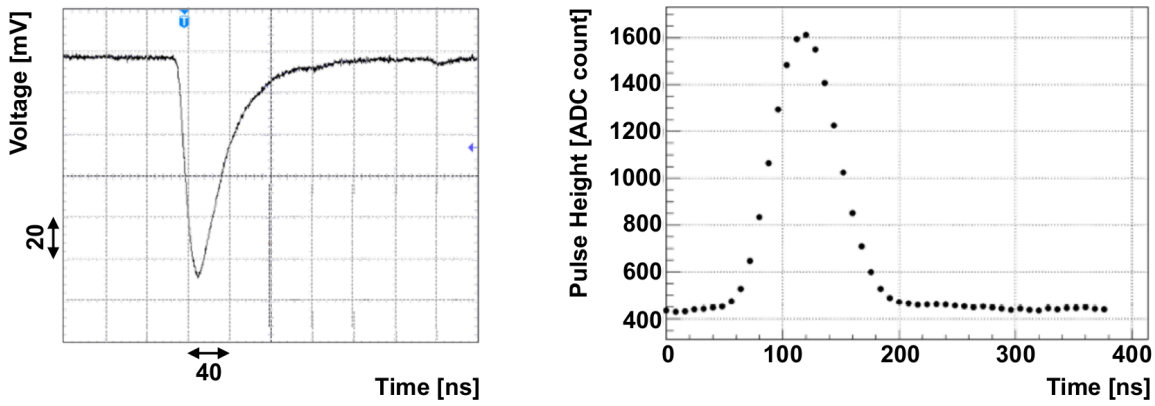


Figure 3.10: Digitization of an analog waveform. An analog pulse from a detector PMT is shown on an oscilloscope (left) and has a sharp leading edge. After shaping with the 10-pole filter, the waveform is digitized into 64 samples (right).

ns timing resolution can be achieved even with an 8 ns sampling rate. The resulting digitized waveform has 64 samples. This step to improve the timing resolution is essential, as a good timing resolution is required to distinguish overlapping events. Figure 3.10 shows the waveform before and after digitization in the 125 MHz ADCs.

The logic decisions described here are illustrated in Figure 3.11 and reflect the data and signal handling in all run periods except Run 79. Once the waveforms are digitized, the FPGA[7] on the ADC board synchronizes the signals from up to 16 channels and aligns them to the K_L beam pulse. The FPGA then carries out two processes in parallel. The energy sums of the waveforms are calculated and sent to the Michigan L1 trigger via 2.5 Gbps optical fibers for the trigger decision. At the same time, the data is packaged and compressed using a lossless compression algorithm (detailed in Section 3.9) and then buffered in a 512 clock-wide pipeline to wait for the L1 decision. If MACTRIS+ sends a signal back to the ADC, then the data is sent to the L2 trigger via 2.5 Gbps optical fibers, otherwise it is ignored. It should be noted that in Run 79, the ADC firmware was modified to fit the requirements of the new OFC and CDT system. These details are described in Section 3.7.

[7] Altera Stratix II EP2S60F1020

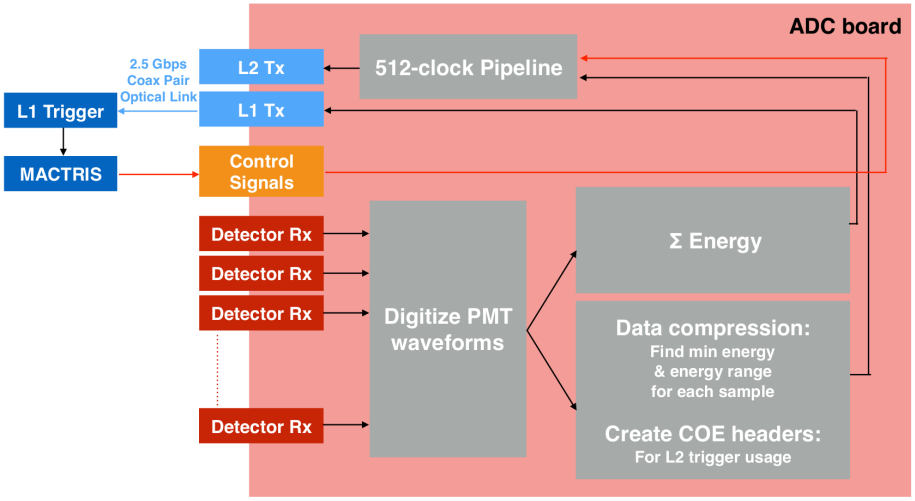


Figure 3.11: Schematic of the ADC module illustrating the firmware logic (excluding changes made for Run 79). The top port of the ADC module is configured to send the data packets to the L2 trigger while the port below it is configured to deliver the energy sum information to the L1 trigger. Figure courtesy of [54].

3.4.2 500 MHz ADC

The 500 MHz ADCs are used for detectors that have direct interaction with the beam and as a result, a higher event rate than detectors that are further away from the beam. By using a fast sampling rate of 2 ns we can capture the sharp edge of the waveform without needing any waveform shaping like is needed in the 125 MHz ADCs. Aside from the waveform shaping with the 10-pole filter, the data handling and logic in the 500 MHz ADCs is the same as in the 125 MHz ADCs.

3.5 Michigan Level 1 Trigger

The L1 trigger decision is made every 8 ns based on the total energy deposited on the CsI calorimeter and no energy in the veto detectors. Specifically, we require a minimum total energy (E_t) of 550 MeV deposited on the calorimeter and no energy in the NCC, CV, MB, IB and CC03 detectors. In this way, the L1 trigger removes events with charged particles and ensures that there is enough activity in the CsI to

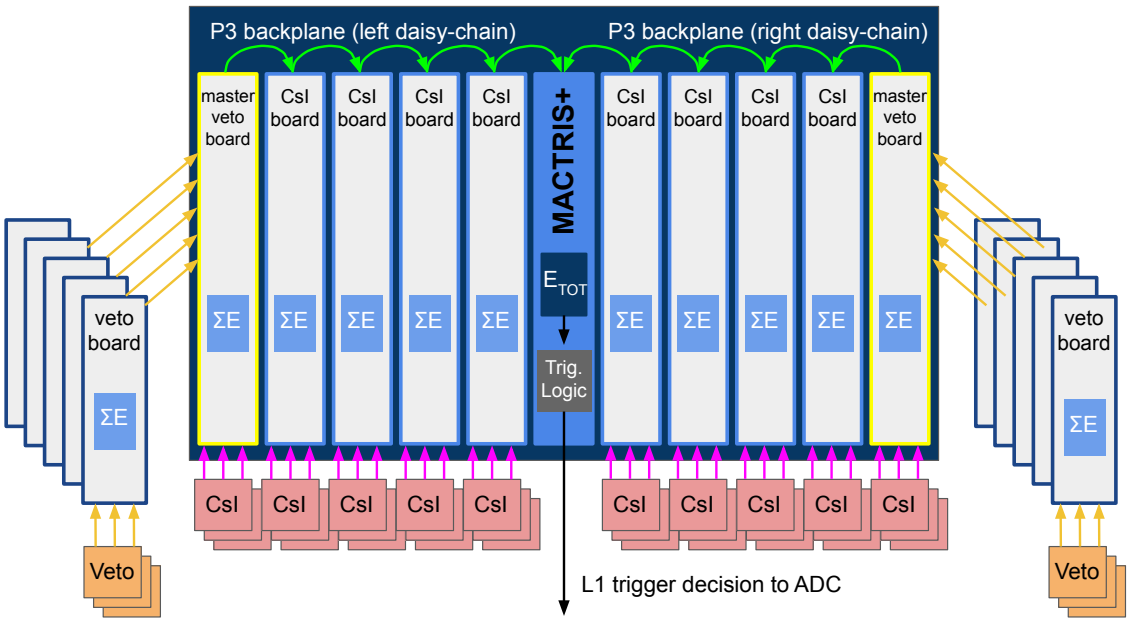


Figure 3.12: Schematic of the L1 trigger daisy-chain and information flow.

account for at least two photon hits.

There are three types of L1 trigger modules that have the same hardware but are programmed to do different jobs– the CsI trigger boards, the veto boards, and the master veto boards. The CsI trigger boards calculate the energy sum from the CsI calorimeter, the veto boards calculate the energy sum from the veto detectors, and the master veto boards receive energy sums from the CsI ADCs and the veto trigger decisions from the veto boards. All of the L1 trigger modules are housed in 9U VME crates connected with a P2 VME backplane and a custom made P3 VME backplane.

The L1 modules are equipped with a Virtex 5 Xilinx FPGA[8] to quickly calculate the energy sums from multiple optical fibers. Each L1 module receives the energy sums of the waveforms from up to 16 ADC boards and calculates the partial energy sum from the inputs. To get the total energy sum, we use the P3 backplane to connect the L1 boards in a daisy-chain and pass the energy sums from each board to the center of the crate where the MACTRIS+ board resides (Figure 3.12). Each

[8]XC5VFX70T

L1 board calculates the energy sum from its inputs and adds this to the energy sum passed from the neighboring board. The MACTRIS+ board sits in the middle slot of the VME crate and calculates the total energy from the left and right daisy-chain energy sums. The MACTRIS+ makes the trigger decision based on the CsI E_t and the veto detector energies and then sends a trigger signal to the ADCs if the conditions are met.

3.6 Michigan Level 2 Trigger

The L2 trigger was designed to select events that have a large transverse momentum, which is a unique characteristic of the signal decay. Because other K_L backgrounds decay exclusively into photons, an event where every photon hits the CsI will have a net transverse momentum near zero. In the case of the signal decay where the neutrinos are not seen by the KOTO detectors, this gives the π^0 and the resulting photons a significant amount of transverse momentum which can be seen by the hit positions on the CsI.

To make this distinction, the L2 trigger calculates the Center of Energy (COE) position on the face of the calorimeter for each event. The distance from the center of the calorimeter to the COE position is called the COE radius, which is determined by the hit position of the photons weighted by their energies as shown in Equation 3.1.

$$\text{COE radius} = \frac{\sqrt{\left(\sum_i E_i x_i\right)^2 + \left(\sum_i E_i y_i\right)^2}}{\sum_i E_i} \tag{3.1}$$

Here E_i is the energy deposited in the ith crystal and x_i and y_i are the x and y coordinates of the ith crystal. Figure 3.13 shows the COE radius distribution for several neutral backgrounds and the signal decay and it is easy to see that the COE

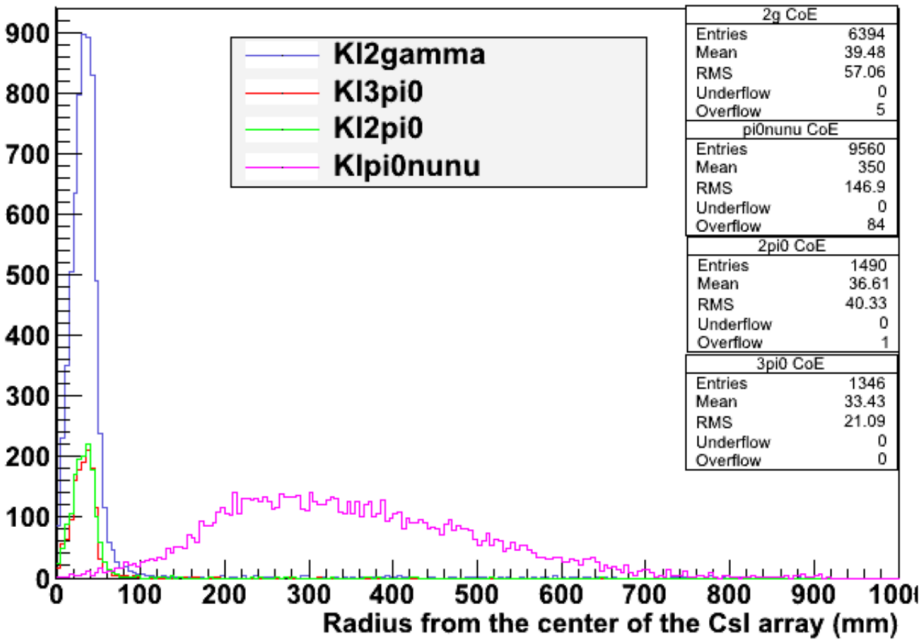


Figure 3.13: Center of Energy distribution for various K_L decays. Background decays such as $K_L^0 \to 3\pi^0$, $K_L^0 \to 2\pi^0$, and $K_L^0 \to 2\gamma$ have lower COE distributions than the signal decay, $K_L^0 \to \pi^0\nu\bar{\nu}$, because the signal decay has missing neutrinos that give the π^0 a large transverse momentum. The COE cut is made at 165 mm. Figure courtesy of [59].

radius is distinctly large for $K_L^0 \to \pi^0\nu\bar{\nu}$. After calculating the COE radius the L2 trigger selects only events that have a COE radius greater than 165 mm.

The L2 trigger boards have identical hardware to the L1 trigger boards though they are programmed to have a different functionality. Figure 3.14 shows the data flow of the L2 trigger. Each L2 board has two FPGAs, a Virtex 5 Xilinx FPGA[9] (V5) and a Virtex 4 Xilinx FPGA[10] (V4)[11]. There are 18 L2 boards housed in a 9U VME crate with P2 and P3 VME backplanes. The COE MACTRIS+ board sits in the center of the crate so that it can receive the COE calculations from each board and make the L2 trigger decision. The L2 boards receive data packets from up to 16 ADC boards via optical fibers and store the events in a First-In-First-Out (FIFO)

[9] XC5VFX70T
[10] XC4VFX12
[11] The L1 trigger modules have but do not use the V4 FPGA.

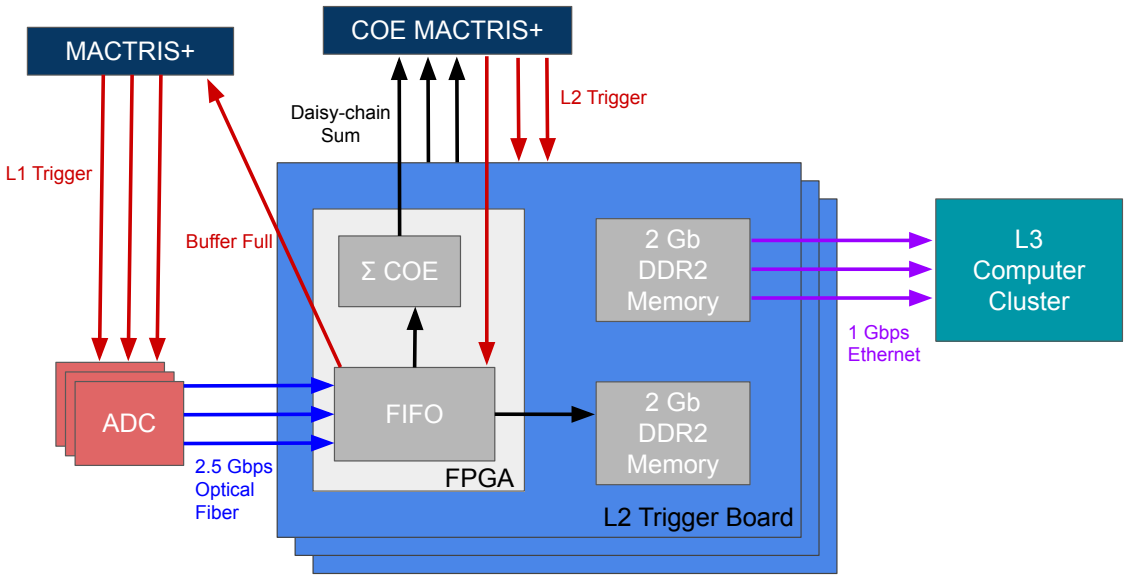


Figure 3.14: Schematic of the L2 trigger information flow.

buffer. The COE partial sum information is calculated sequentially for each event by the V5 FPGA and the information is passed from the L2 boards via the P3 backplane daisy-chain so that the COE MACTRIS+ can calculate the final COE radius. COE MACTRIS+ makes the trigger decision and issues a trigger accept signal to the L2 boards via the P2 VME backplane.

If a trigger accept signal is received by the L2 board, the events are read out from the FIFO to a 2 Gb DDR2 memory by the V4 FPGA to be stored until they can be transferred to the L3 computer cluster during the next spill cycle. There are two DDR2 memories on the L2 boards so that during each spill cycle we can write to one memory and read out to L3 with the other and prevent overwriting events. Each of the L2 boards can also communicate with MACTRIS+ in the event that there is not enough room in the FIFO to store incoming triggers. In this case, the L2 board will send a "buffer full" signal to MACTRIS+ to tell it to stop issuing L1 trigger signals until the buffer has room again. The capacity of the FIFO is 11 events and the lossless compression algorithm in the ADCs greatly reduces the data size so that the FIFO rarely becomes full.

3.7 OFC and CDT Trigger System

Before the last data collection period in 2018, a new L1 and L2 trigger system referred to as the Optical Fiber Center (OFC) and Clock Distribution and Trigger (CDT) system was completed, which replaced the original Michigan L1 and L2 triggers. The system was gradually introduced with the addition of the CDT trigger in 2017 and was finalized with the addition of the OFC modules in 2018. This upgrade allowed us to reduce DAQ dead time and improve the trigger efficiency. The performance of the new system is detailed in Chapter 4. The new L1 trigger is the same as before and is determined by the combination of the total energy deposited on the calorimeter and no energy in the veto detectors. The new L2 trigger is based on the number of clusters on the calorimeter.

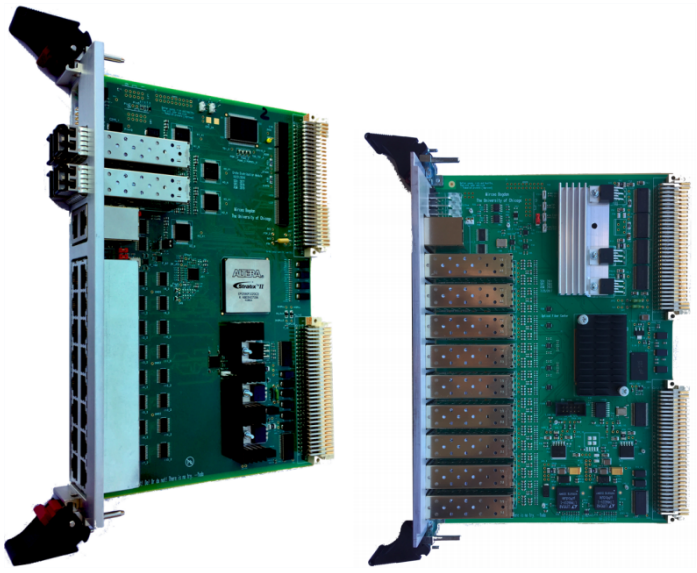


Figure 3.15: The CDT (left) and OFC (right) modules.

The new triggering system is composed of two different types of custom made modules, CDT modules and OFC modules (Figure 3.15) and both boards are equipped with FPGAs to complete different tasks. Most of the CDT modules act as Local CDTs, which collect the relevant information from the ADCs and pass it to either one

of two OFC modules. The Clustering OFC module collects the clustering information for the L2 trigger and the Energy/Veto OFC collects information for the L1 trigger. One CDT module acts as the Top CDT, which makes the final L1 and L2 trigger decisions and distributes other signals to the rest of the DAQ system (the Top CDT takes over some of the jobs of MACTRIS+). Figure 3.16 shows the data and signal flow of the OFC and CDT system and the details of the L1 and L2 triggers are described below.

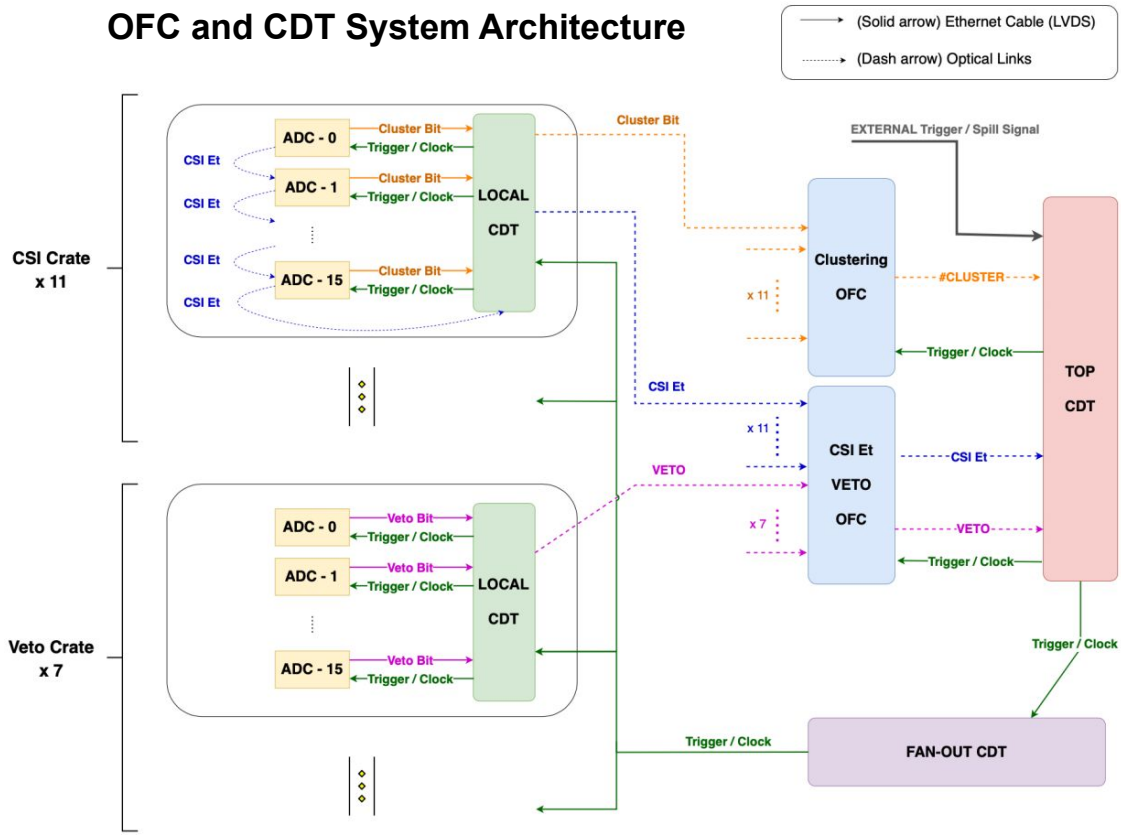


Figure 3.16: Schematic of the OFC and CDT system data and signal flow. Figure courtesy of [60].

3.7.1 New L1 Trigger

The new L1 trigger has the same criteria as the Michigan L1 trigger but is determined in a different way, with different hardware. The Top CDT distributes the 125 MHz clock to the rest of the system and the L1 trigger decision is made every 8 ns. There are 18 ADC crates that convert the analog signals from the detectors to digital data, 11 of which receive signals from the CsI calorimeter (CsI crates), and 7 of which receive signals from the veto detectors (veto crates). For the CsI crates, the partial energy sums from each ADC board are added via a daisy-chain using 2.5 Gbps optical fibers. For the veto crates, the veto bit is set HIGH when the energy exceeds the threshold in each module. The energy sum information and veto bits are sent to the Local CDT modules via 2.5 Gbps optical fibers and Category-5 LVDS Ethernet cables, respectively. Then 11 CsI E_t values and 7 veto bits are sent via optical links from the Local CDTs to the Energy/Veto OFC, which sums the CsI E_t values and obtains the final veto bits by taking a logical OR of the 7 veto inputs. The final energy sum and final veto bits are sent to the Top CDT via optical links, which makes the L1 trigger decision.

3.7.2 New L2 Trigger

If the L1 trigger conditions are met, the Top CDT distributes a cluster trigger signal via the fanout CDT to the Local CDTs, which distribute it to each of the ADC boards (Figure 3.16). When a cluster trigger is received by the CsI crates, a cluster bit is made for each channel based on the peak height of the waveform. These bits are sent to the Clustering OFC which uses a high-speed clustering algorithm to create a full map of the CsI calorimeter clusters inside the OFC module, an example of which can be seen in Figure 3.17.

The number of clusters is calculated and reported to the Top CDT via 2.5 Gbps optical fibers and the L2 trigger decision is made based on the number of clusters.

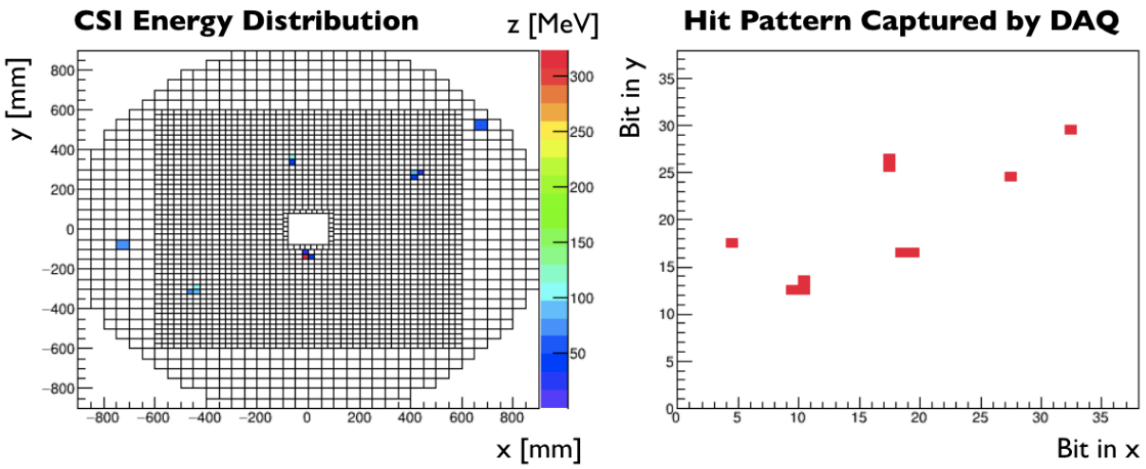


Figure 3.17: Energy deposition display of a six cluster event on the CsI calorimeter (left) and the corresponding CDT map with six clusters (right). Figure courtesy of [60].

In this way, we are able to collect events with two clusters (signal) as well as events with four and six clusters ($2\pi^0$ and $3\pi^0$) for normalization analysis. If the trigger conditions are satisfied, a trigger signal is sent to the ADCs and the data packages are written out from the ADCs to the Michigan L2 trigger (bypassed) and sent on to the L3 trigger system for event building.

3.8 Level 3 Trigger

Because the signals from the detectors come into the ADCs over multiple channels, the data for a single event is in pieces and must eventually be put back together to get a full picture of the event. Thus, the job of the L3 trigger is to collect these event fragments from L2, do event building and reconstruction, make online software cuts, and repackage the data for permanent storage on KEKCC.

The L3 software trigger is a computer cluster named "Banjo" that consists of 47 worker nodes (banjo01 − banjo47), and two head nodes (banjo and banjo−1) that supervise the worker nodes. Each node has two processors and eight CPU cores.

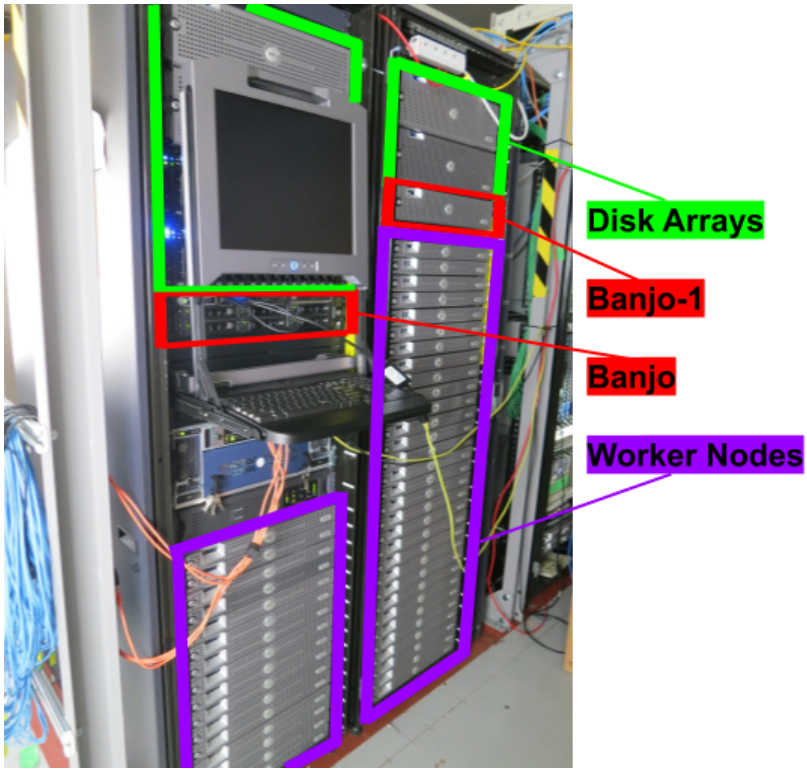


Figure 3.18: The Banjo computer cluster. The head nodes are shown in red, the worker nodes in purple, and the disk arrays in green.

The head nodes are Dell 2950 computers with 8 TB of storage[12] each and the worker nodes are Dell 1950 computers each with two 4 TB disk drives. The system also has eighteen stages of 16 TB RAID 5 disk arrays for temporary storage on the cluster. Figure 3.18 shows a photo of Banjo and the different components.

The worker nodes are split into two groups that accomplish different tasks– Type 1 nodes (banjo01 − banjo17) and Type 2 nodes (banjo18 − banjo47). The 17 Type 1 nodes receive event fragments from L2 and send the fragments to the Type 2 nodes. The 30 Type 2 nodes build the events, perform the online software cut, and compresses and package the data. The data is then saved to the node's local disk, later transferred to the cluster's disk arrays, and ultimately sent out to be permanently stored at the KEK computer cluster.

[12] 2 TB RAID 1 system disk and 6 TB RAID 5 local disks

3.8.1 Event Building and Packaging

When the Ethernet data packets are passed to L3, the events must be built, compressed, and written to disk before the next spill cycle starts. This means that L3 has less than six seconds to reconstruct and process the data. In order to achieve this high data processing rate, we use both the Ethernet network and the Message Passing Interface (MPI) network to create mesh communication so that there is direct communication between all nodes. We use a 48 port Ethernet switch to send control signals and a 48 port Infiniband switch to facilitate fast data transmission between nodes. Each node is connected to the Ethernet switch via 10 Gbps Ethernet cabling and to the Infiniband switch via 20 Gbps Infiniband cabling. We also use the multi-core capabilities of the nodes and the MPI protocol to run a multi-threading algorithm that does parallel processing of each task. We use the Real Time Red Hat Linux operating system for all nodes and the L3 software to do the event building, data processing, and packaging is written in C++.

The data flow and event reconstruction is described and depicted in Figure 3.19. Each Type 1 node has a direct connection to a L2 trigger board and receives event fragments using the 1 Gbps Ethernet via UDP protocol for fast data transferring. Type 1 nodes send event fragments to Type 2 nodes for event building using the Infiniband with the MPI network and TCP/IP protocol. Events are built using the the event ID and spill information contained in the Ethernet data packet header. Once the events are built the ADC-compressed data is decompressed to do trigger selections with further analysis, conduct checks of the data, and rearrange the data for offline analysis. The L3 uses little-endian as the data packaging style.

After this is done the data is recompressed using a lossless data compression algorithm to reduce the event size to save space on the disk storage. The data is packaged in a binary file and saved to the node's local disk, later transferred to the cluster's disk arrays, and ultimately copied and sent out to be permanently stored

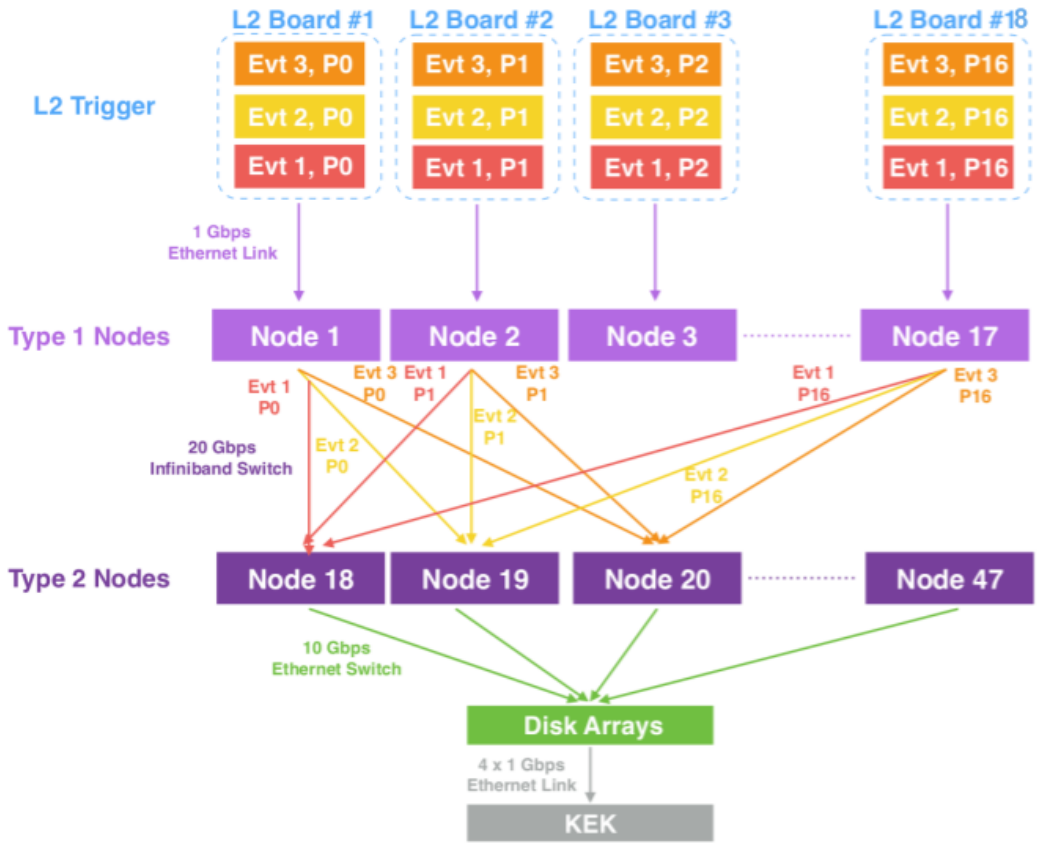


Figure 3.19: The L3 data flow. Figure modified from [54].

at the KEK computer cluster. We use the 4 Gbps SINET4 network[13] to transfer the data to KEKCC and remove the L3 local data after checking that the copy was transferred successfully to KEK.

3.8.2 L3 Software Cut

The L3 trigger software cut was developed and implemented in 2017. We determined through offline analysis studies that most events with four clusters cannot be reconstructed to the $K_L^0 \to 2\pi^0$ decay. For this reason, we designed an online cut for events with four clusters (determined by the CDT trigger) that checks the energy in the MB and IB and if the energy exceeds 10 MeV in any MB or IB module, the

[13]The Science Information NETwork (SINET) is a Japanese academic network for universities and research institutions.

event is rejected. From the offline analysis studies, this cut reduced this background by around 30%. The L3 software also carries out online quality checks of the data to ensure consistent data for offline analysis.

3.9 Data Compression

There are two purposes to the data compression in the KOTO DAQ system– firstly, the data is compressed in the ADCs to maximize data throughput with the available resources needed to process the data and secondly, the data is uncompressed and then recompressed in the L3 trigger to reduce the file size and save space on KEKCC. In order to keep the full information in the data, we use a bit-packing lossless compression algorithm which allows the original data to be perfectly reconstructed from the compressed data.

The waveform information that is received by the ADCs has 64 digital samplings that is compressed into data packets. The compression algorithm encodes the data into a 16-bit word with the last two bits reserved for header information. The compressed data is arranged by waveform samples in the ADCs and for each waveform the minimum and maximum energy values among the 64 samples are recorded. This gives the maximum number of bits needed to store the energy difference in each sample to the minimum energy and we store this energy difference for all 64 samples. The same algorithm is used in the L3 trigger, except the data packets are arranged by detector channels instead of waveform samples for easier offline analysis. The data compression is an essential part of the DAQ system as it reduces DAQ deadtime and allows us to collect and store as much data as possible so that we can collect enough statistics to observe $K_L^0 \to \pi^0 \nu \bar{\nu}$.

Chapter 4

Data Collection

The KOTO experiment collected its first physics data in 2013, though the data collection was stopped after 100 hours due to a radiation incident in Hadron Hall[1]. KOTO began running again in 2015 and since then has been collecting data periodically each year. This chapter will focus on the data collection from 2016 to 2018 which is the data set used in this thesis for the $K_L^0 \to \pi^0\nu\bar{\nu}$ search. In particular, it will cover the conditions during data collection, the types of experimental runs, trigger selections, and the performance of the DAQ system.

4.1 Overview of Data Collection

The J-PARC accelerator operates in either SX or FX, so the beam use is split between the two and KOTO takes data during the allocated SX beam time. The periods of data collection are split up in different experimental "Runs" which are dictated by the accelerator operation cycles. A Run typically lasts one to two months. Within an experimental Run the data collection is further divided into "runs" which are typically 500 spills. Since the beam extraction period (a spill cycle) is approximately six seconds, a run lasts about 50 minutes. Different trigger selections can be adjusted and troubleshooting can be carried out between runs.

[1] A report of the incident can be found at [61].

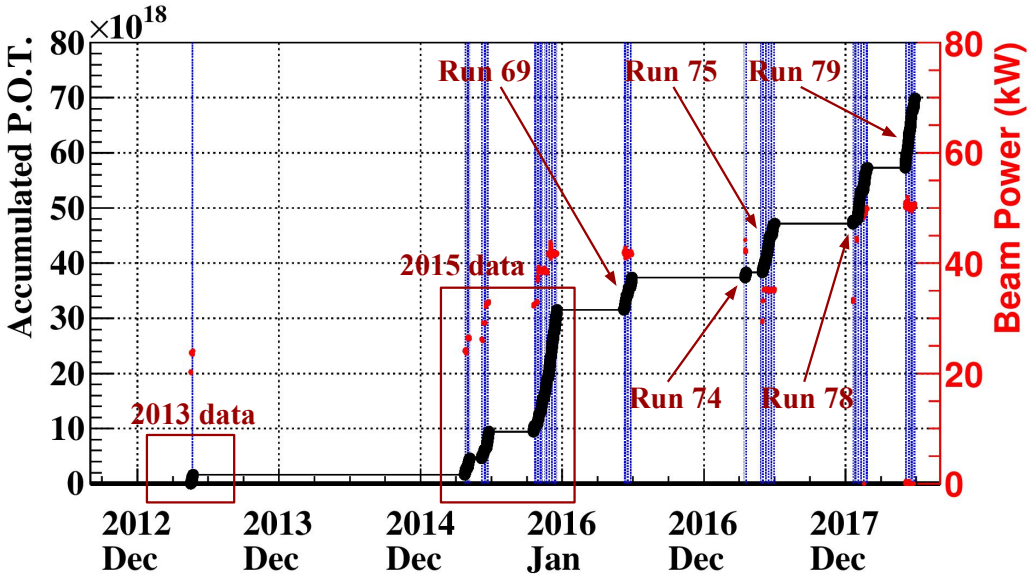


Figure 4.1: Accumulated POT (black) and beam power (red) over time from 2013 to 2018. The run periods taken during 2016–2018 are labeled.

Figure 4.1 shows the accumulated Protons on Target (POT) and the beam power from 2013 to 2018, which gives an indication of how much data we have collected since the first run in 2013. The 100 hours of data taken in 2013 was analyzed together as a data set (referred to as the 2013 run/analysis) and the data taken in 2015 was analyzed together (referred to as the 2015 runs/analysis). The data set used for this thesis includes data taken from 2016–2018 after the IB was installed and before the experiment was upgraded with the installation of double-ended readout MPPCs on the CsI calorimeter (details in Section 10.3.1). The different run periods in the 2016–2018 runs are shown in Figure 4.1 as Run 69, Run 74, Run 75, Run 78, and Run 79.

4.2 2016–2018 Data Collection

Table 4.1 lists the details of the run periods in 2016–2018. The IB was installed after the 2015 runs in preparation for Run 69 in June of 2016. The beam power in

Run	Dates	Beam Power (kW)	POT$\times 10^{18}$ Phys.	Phys. Corrected	Comments
69	5/28 ∼ 6/30/2016	42	5.5	3.1	Installed new IB
74	4/13 ∼ 4/24/2017	44	1.0	N/A	ESS trouble in MR, CDT added to DAQ
75	5/25 ∼ 7/2/2017	31, 35, 37.5	9.2	8.2	CDT trigger implemented
78	1/13 ∼ 2/26/2018	33, 44, 50	10.2	8.1	
79	6/7 ∼ 6/28/2018	51	12.6	11.2	new L1 and L2 triggers in DAQ

Table 4.1: Run periods in the 2016–2018 data set. The POT delivered to the T1 target in physics runs is indicated as "Phys.". "Phys. Corrected" indicates the POT in physics runs properly corrected based on the DAQ livetime and only using good runs and good spills.

Run 69 was 42 kW and due to an issue with the L2 trigger, we took physics and normalization triggers separately. During Run 74 in 2017, after about 1.5 weeks of data collection, the Electro-Static Septum 1 (ESS1) in the MR was damaged by an unstable proton bunch in the beam, causing the fine wire electrodes to break and short circuit when they touched the high voltage electrode. Because of this issue, the accelerator operation was stopped for about a month while the ESS1 was replaced with the ESS2, and the MR operated with one ESS instead of two[2]. Due to this limitation, the beam intensity was reduced and limited to a maximum of 37.5 kW. The majority of the data taken in Run 75 was with a beam power of 37.5 kW.

The ESS1 was repaired before Run 78 in January of 2018, and with a gradual increase in beam intensity, most of the data taken in Run 78 was at 50 kW beam power. Run 79 occurred in June of 2018 with a beam power of 51 kW, which is the maximum beam power for the T1 target due to cooling limitations. The repetition cycle for the beam extraction was 5.52 seconds in Runs 69, 74, and 75 and before Run 78 it was changed to 5.2 seconds and remained at 5.2 seconds for Run 79. The spill length varied from approximately 2.1 seconds to 1.8 seconds and often gradually decreased over the course of the run period. The POT collected for physics runs ("Phys.") is listed in Table 4.1 as well as the POT collected in physics runs

[2]More details can be found at [62].

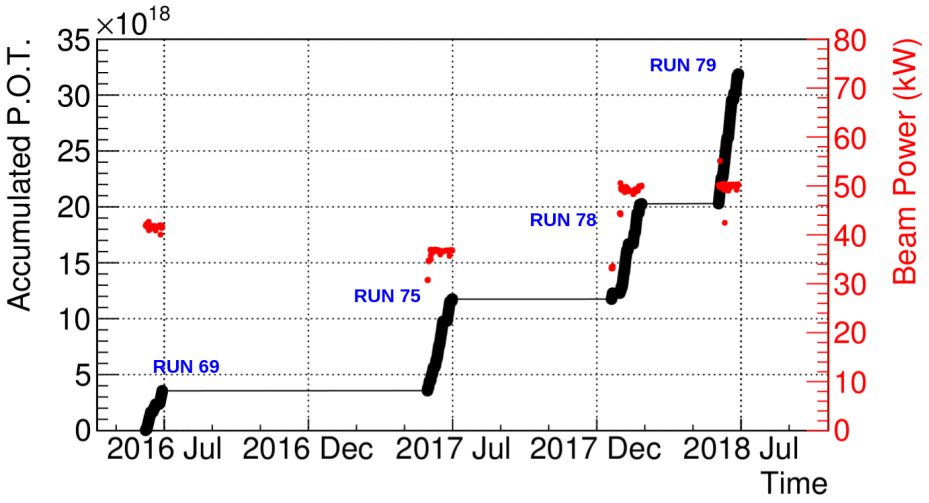


Figure 4.2: Accumulated POT (Phys. Corrected) (black) and beam power (red) for the 2016–2018 physics runs. Run 74 was not used in the normalization analysis so it is not shown here.

that has been corrected for loss factors ("Phys. Corrected"). Figure 4.2 shows the accumulated, corrected POT and beam power in physics runs for the 2016–2018 run periods. The total number of corrected POT that was collected during the 2016–2018 physics data collection was 3.05×10^{19}. It should be noted that because Run 74 is a very small data set, it was not included in the normalization analysis. Finally, there were two dead channels in the CsI calorimeter that were excluded in the analysis, channel 356 and 357, which can be seen in Figure 4.3. There were also two dead channels in the IB and IBCV detectors and their treatments are discussed in Section 5.3.4.3.

4.3 Trigger Selections

A mixture of several types of triggers were used to collect various events for the analysis. The primary trigger was the physics trigger, which was used to collect the $K_L^0 \to \pi^0\nu\bar{\nu}$ sample. We collected other trigger types simultaneously at lower prescale factors that are used for normalization analysis, checking accidental activity, and

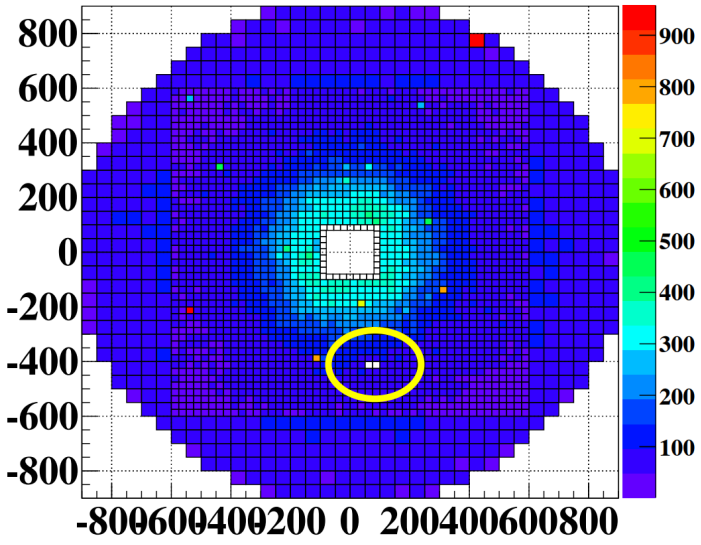


Figure 4.3: CsI dead channels 356 and 357 (circled in yellow) were excluded in the 2016–2018 analysis. The view of the calorimeter is from downstream.

tracking detector performance. The prescale factor was set for each trigger with the intention of collecting as many physics triggers as possible to maintain high statistics for $K_L^0 \to \pi^0\nu\bar{\nu}$ while also collecting a large enough sample of other triggers necessary for the analysis. The following sections detail the purposes and conditions for each trigger type.

4.3.1 Physics Trigger

The physics trigger was the main trigger designed to collect data sensitive to $K_L^0 \to \pi^0\nu\bar{\nu}$. The conditions of the physics trigger required a total energy (E_t) greater than 550 MeV deposited in the CsI calorimeter and no activity in the veto detectors which includes the NCC, CV, MB, IB and CC03 detectors (commonly called the "online veto"). The threshold for "no activity" in each of the veto detectors is shown in Table 4.2. The physics trigger also required a large transverse momentum which corresponds to a COE radius greater than 165 mm. In Run 75 we replaced the COE cut with the requirement of two, four, or six clusters on the calorimeter as calculated by the CDT trigger system. For events with four clusters, we required less

Veto Detector	ADC Count Threshold	Energy Threshold (Run 79)
NCC	$<$ 3000 ADC counts	14 MeV
CV	$<$ 7500 ADC counts	0.2 MeV
MB	$<$ 4300 ADC counts	14 MeV
IB	$<$ 1200 ADC counts	10 MeV
CC03	$<$ 5000 ADC counts	14 MeV
CC04-CC06	N/A	14 MeV

Table 4.2: Threshold of ADC counts/energy for each veto detector for the physics trigger. The ADC counts correspond to offline energies in MeV after applying a gain correction factor for each detector and were used in Runs 69-78. The new DAQ system (OFC+CDT) used the E_t sum in MeV and this was implemented in Run 79. The CC04-CC06 veto detectors were also used for the physics L1 trigger in Run 79.

than 10 MeV in all MB and IB modules, for the purpose of rejecting events that were tagged as four clusters, but had energy deposited in the MB or IB.

Thus, for Runs 75, 78, and 79, the physics trigger required the E_t sum, the online veto, two, four, or six clusters, and no activity in the IB or MB for events with four clusters. Though the $K_L^0 \to \pi^0\nu\bar{\nu}$ decay only has two photon clusters, we included four and six cluster events in the physics trigger in order to collect more $K_L^0 \to 3\pi^0$ and $K_L^0 \to 2\pi^0$ events to be used for the normalization analysis. The prescale factor for the physics trigger was one, since a high level of statistics is required for the $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis.

4.3.2 Normalization Trigger

The purpose of the normalization trigger was to collect $K_L^0 \to 3\pi^0$, $K_L^0 \to 2\pi^0$, and $K_L^0 \to 2\gamma$ events which are used in the normalization analysis (Chapter 7). In the case of Run 69, the normalization trigger was the same as the physics trigger, except the COE cut was not applied since there is no missing energy in $K_L^0 \to 3\pi^0$, $K_L^0 \to 2\pi^0$, or $K_L^0 \to 2\gamma$. In Run 75 and thereafter, the normalization trigger requirements remained the same, including only the E_t sum and the online veto, regardless of the number of clusters. To suppress the trigger rate the prescale factor for the normalization trigger was 30, meaning one out of every 30 triggers was a normalization trigger.

4.3.3 Minimum Bias Trigger

The minimum bias trigger required only that the energy deposited in the CsI was greater than 550 MeV. By excluding any information about the veto detectors, we were able to collect data free from any online veto trigger biases, and this is used to evaluate veto cut efficiencies. A prescale factor of 300 was applied for the minimum bias trigger to reduce the trigger rate.

4.3.4 Calibration Trigger

The purpose of the calibration trigger was to collect $K_L^0 \to 3\pi^0$ events for the energy calibration of the CsI calorimeter. There are no other K_L^0 decays that generate six photons, so $K_L^0 \to 3\pi^0$ events can be collected effectively by looking at the number of clusters. Before cluster counting with the CDT trigger system, the CsI was divided into twelve regions, as shown in Figure 4.4, which was used for rough cluster counting by region. The calibration trigger had the same requirements as the normalization trigger (E_t sum and online veto) and required hits in four or more regions with energy exceeding 120 MeV. Though cluster counting by the CDT trigger system was added in Run 75, the calibration trigger continued to use the region cluster counting method. A prescale factor of 10 was applied for this trigger.

4.3.5 External Triggers

There were five external triggers used in the 2016–2018 data collection– the cosmic ray trigger, clock trigger, laser trigger, LED trigger, and Target Monitor (TMon) trigger. The external triggers come from sources other than the K_L beam and they all had a prescale factor of one applied.

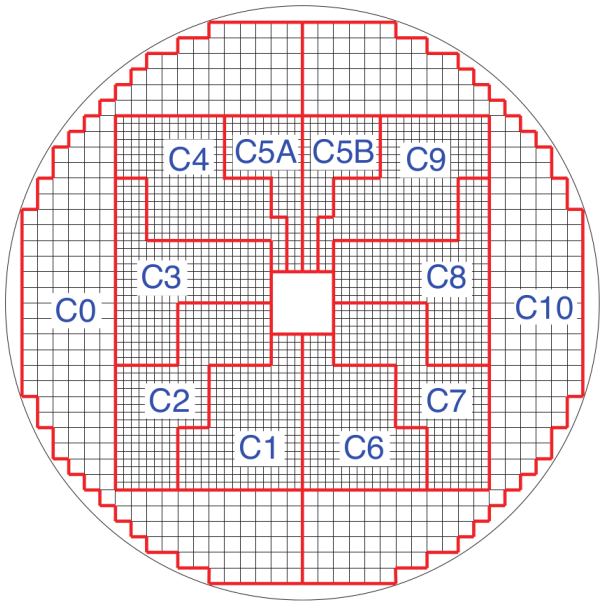


Figure 4.4: The CsI calorimeter divided into 12 sections for region cluster counting for the calibration trigger. Figure courtesy of [63].

4.3.5.1 Cosmic Ray Trigger

The purpose of the cosmic ray trigger was to continuously check the stability of the CsI, OEV, and NCC detectors. We collected cosmic ray data in the off-spill time during each spill cycle and the cosmic ray trigger was issued when the CsI, NCC, or OEV detected cosmic ray muons. The E_t thresholds for the CsI and the NCC were the same as in the physics trigger, and for the OEV we required two or more modules to have hits for a trigger to be issued.

4.3.5.2 Clock Trigger

The clock trigger, also known as the random trigger, was issued by a 10 Hz periodic clock in order to take random trigger data. This trigger was used to monitor counting rates and noise in detectors during a spill.

4.3.5.3 Laser Trigger

The laser trigger used light from a laser that was flashed at 5 Hz and triggers were issued synchronously with this timing. The purpose of the laser trigger was for the calorimeter PMT gain correction and channel timing alignment.

4.3.5.4 LED Trigger

The LED trigger was similar to the laser trigger in which LED lights were flashed at 10 Hz in the veto detectors. This trigger was used to monitor gain stability in the veto detector MPPCs and PMTs and adjust the timing offsets between veto detector modules.

4.3.5.5 Target Monitor Trigger

The Target Monitor (TMon) trigger was used to collect data with accidental hits related to the beam activity. The Target Monitor is located in Hadron Hall, 50° off the primary beam line and monitors the yield of secondary particles from the T1 target. By issuing triggers generated by the Target Monitor related to accidental activities in the beam, we could study the accidental activity and its effect on the physics trigger.

4.4 Special Runs

Most of the data that we collected were "physics runs" with the main purpose to collect high statistics for $K_L^0 \to \pi^0\nu\bar{\nu}$ and other events used in analysis. However, we also dedicated beam time to collecting other types of data used to calibrate detectors and study various backgrounds. These special runs collected during 2016–2018 were Aluminum (Al) target runs used primarily for studying the hadron-cluster background and muon runs used for calibration. Figure 4.5 shows a breakdown of the data and

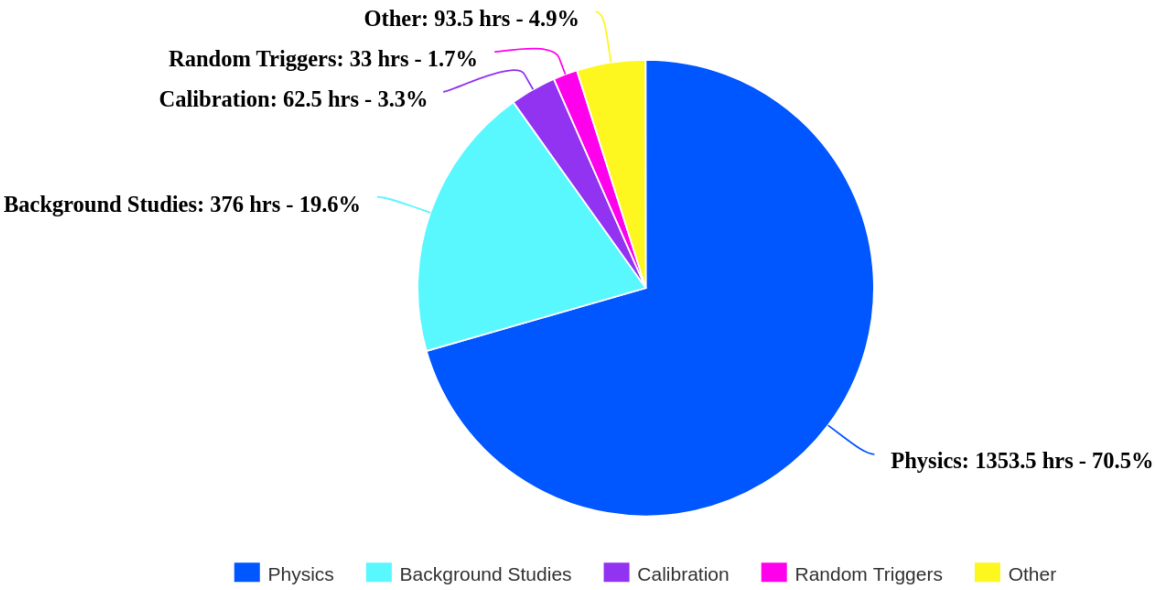


Figure 4.5: Amount of data collected during the 2016–2018 runs during beam time. "Physics" includes any normalization runs that were taken separately. "Background Studies" includes Al target runs, while "Calibration" includes muon runs and other runs taken to calibrate specific detectors. "Random Triggers" includes TMon and clock runs. It should be noted that many more hours of calibration data were taken before and after beam time and are not included in this chart.

the different types of runs in the 2016–2018 periods.

4.4.1 Aluminum Target Runs

In the 2016–2018 data collection, we took special runs in which an Aluminum plate, or "target", was inserted into the beam to scatter neutrons for background studies, as well as to use in studies that required known z positions. There were two types of Aluminum (Al) targets used, the Z0 Al target, and the Decay Volume Upstream (DVU) Al target. The Z0 Al target was located at the front of the FB around the Z=0 position and the DVU Al target was located at the upstream end of the decay volume (Figure 4.6 shows the positions of each). The Z0 Al target could be mechanically inserted into the beam using a motor system and the DVU Al target could be moved into the beam by rolling down a wire that it hung on (Figure 4.7). The specifications of each Al target are shown in Table 4.3.

	Z0 Al Target	DVU Al Target
Thickness	10 mm	5 mm
Size	80×80 mm^2	100 mm diameter
Z Position	-634 mm	2795 mm

Table 4.3: Specifications of the Z0 and DVU Al targets.

The Z0 Al target runs were used to collect samples of control data for the "hadron-cluster background", in which a scattered neutron hits the calorimeter and produces a hadronic shower and a neutron in the first hadronic shower scatters to produce a second hadronic shower. The Z0 Al target was inserted into the beam line to scatter neutrons for the Z0 Al target runs and was moved out of the way the rest of the time. More details on this background can be found in Section 8.4.1. We also took special runs with the DVU Al target inserted into the beam with the purpose of generating π^0s when a beam core neutron interacted with the target. The DVU Al target runs were mainly used for calibrating the CsI calorimeter with $\pi^0 \to 2\gamma$. Since the z generation position of the π^0 (the DVU Al target z position) is known, the π^0 mass can be reconstructed using Equation 5.18. Similarly, the energies of the photons can be obtained by assuming the π^0 mass. Thus, these data runs were used to correct the absolute energy scale of the energy calibration.

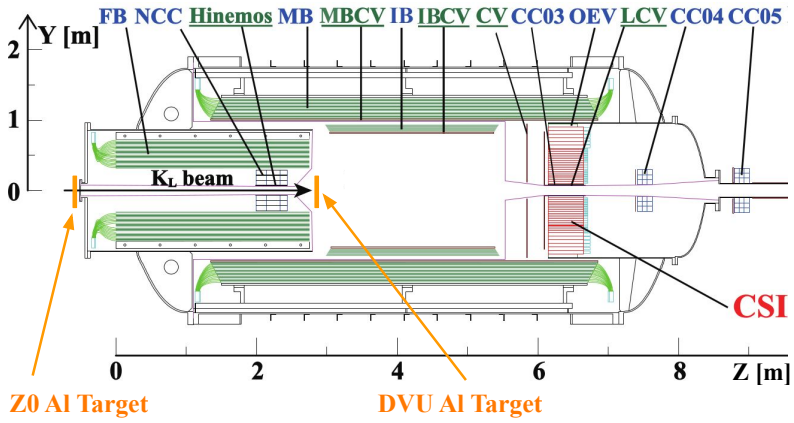


Figure 4.6: Locations of the Z0 and DVU Al targets in the KOTO detector system.

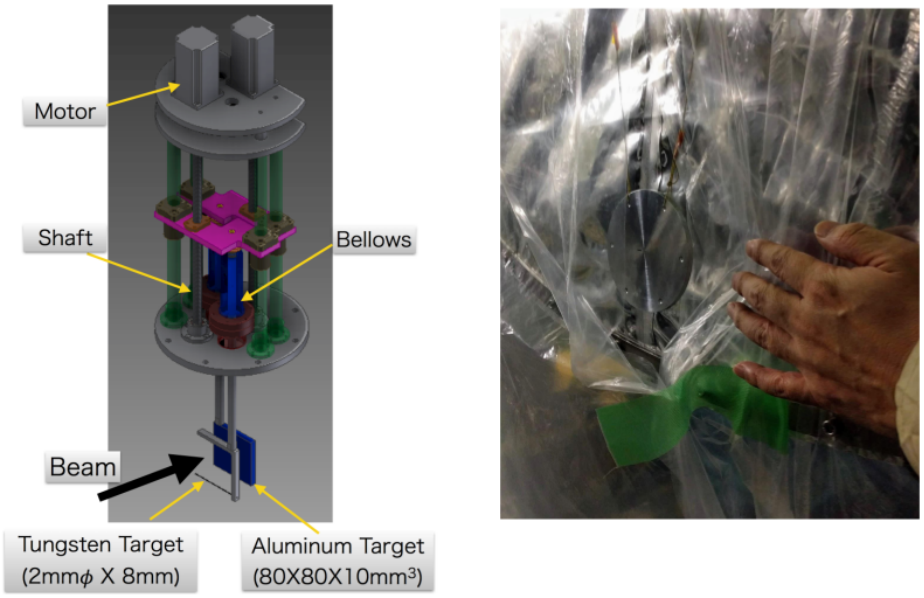


Figure 4.7: The Al targets used in the 2016–2018 data collection. A schematic drawing of the Z0 Al target and the motor system used to move it is shown on the left and a picture of the DVU Al target is shown on the right.

4.4.2 Muon Runs

There were two types of muon runs that we took during the 2016–2018 data collection: cosmic muon runs and beam muon runs. The purpose of the muon runs was for calibration of detectors and stability checks. We took cosmic muon runs before and after the accelerator runs as well as during any accelerator maintenance periods. Different trigger settings were used depending on the needs of the run and events were triggered by E_t sums in the CsI, NCC, OEV, MB, IB, BPCV, and downstream collar counters. The clock, laser, and LED external triggers were also used during these runs.

To calibrate detectors close to the beam hole that could not be calibrated in the cosmic muon runs, we took beam muon runs in which the beam plug (described in Section 2.2.4) was closed. By closing the beam plug and turning off the sweeping magnets, high momentum charged muons could penetrate through the beam plug and enter into the KOTO detector area. The beam muon runs were also used to

study how charged particles penetrate detectors close to the beam. In these runs events were triggered using a combination of E_t sums from the NCC, BHCV, and downstream collar counters.

4.5 DAQ Performance

The DAQ system was continuously upgraded from 2016 to 2018, and we tracked the effects of these changes through the performance of the system. One important criteria of the DAQ system that it sustain high trigger rates with minimum dead time, meaning that the system loses as little data as possible during data collection. Due to hardware limitations, stray errors during data passing, and corrupted data, the DAQ system does not operate at 100% efficiency. In order to accurately calculate how many kaons are coming into the KOTO detectors for the analysis, it's necessary to know how much data is lost through the inefficiency of the DAQ system. Thus, measuring the performance of the DAQ system over each run is an important step. It should be noted that since Run 74 was such a small data set, the DAQ performance was not calculated for this run period.

The DAQ system performance is quantified by what is called the "livetime ratio" or simply "livetime". The livetime ratio indicates the fraction of data that was passed to the next stage in the system, excluding the data that was intentionally removed by trigger cuts. The livetime for the L1, L2, and L3 triggers are covered in the next sections. Since some of the trigger systems were upgraded from run to run, the elements causing the dead time are different, but the livetime definition remains the same. That is, the ratio between the number of events coming out of the system and the number of events coming into the system. Some commonly used DAQ statistics are defined below:

- L1 Raw (L1Raw) – Raw trigger/event coming into the DAQ system

- L1 Gate (L1Gate) – The number of clocks that the L1 trigger requires to process events
- L1 Request (L1Req) – Event that is far enough apart from other triggers and passed the L1 trigger decision
- L1 Accept (L1A) – Event that passed the L1 trigger decision and L2 is not busy
- L1 Reject (L1Rej) – Event that was rejected by the L1 trigger decision
- L2 Accept (L2A) – Event that passed the L2 trigger decision
- L2 Reject (L2Rej) – Event that was rejected by the L2 trigger decision

4.5.1 L1 Livetime

The L1 livetime refers to the fraction of events that pass the L1 trigger, excluding those events that the L1 trigger intentionally rejects. Thus, the L1 livetime can be written as

$$\text{L1 Livetime Ratio} = \frac{\#\ \text{L1Req}}{\#\ \text{L1Raw}} \tag{4.1}$$

L1 Raw are events that are at least one clock[3] apart from each other that come into the DAQ system. Due to hardware limitations, the L1 trigger can only process events that are far enough apart from each other and so some events are lost, even though these events pass the L1 trigger decision. In Run 69 the L1 Gate was 10 clocks and in Runs 75 and 78 the L1 Gate was 100 clocks, due to the time the CDT required to process events. In Run 79 the L1 Gate was reduced to 20 clocks, which improved the L1 livetime, as seen in Figure 4.8.

[3]Each clock is 8 ns

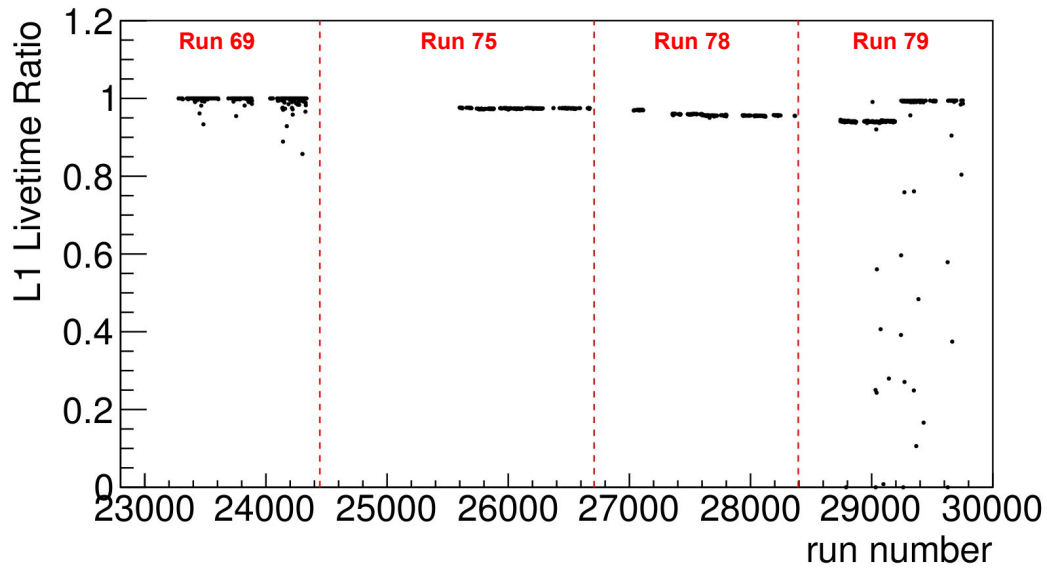


Figure 4.8: The L1 livetime ratio vs run number for run periods in 2016–2018. The L1 livetime considers the L1 Gate requirement for the L1 trigger.

4.5.2 L2 Livetime

Before the system was upgraded with the addition of the CDT, the largest bottleneck in the DAQ was the limited memory and buffer size in the L2 hardware. Because there was limited space in L2, some events that passed the L1 trigger couldn't be passed on to L2 because it was already busy processing other events. L1 Requests are events that passed the L1 trigger, while L1 Accepts are events that passed the L1 trigger without L2 being busy, and so the L2 livetime can be written as

$$\text{L2 Livetime Ratio} = \frac{\#\ \text{L1A}}{\#\ \text{L1Req}} \tag{4.2}$$

The addition of the CDT before the L2 trigger in Run 75 reduced the trigger rate coming into the L2 trigger, however the L1 Gate requirement of 100 clocks for the CDT moved most of the L2 dead time to the L1 trigger. This improvement in L2 livetime from Run 69 to Run 75 can be seen in Figure 4.9.

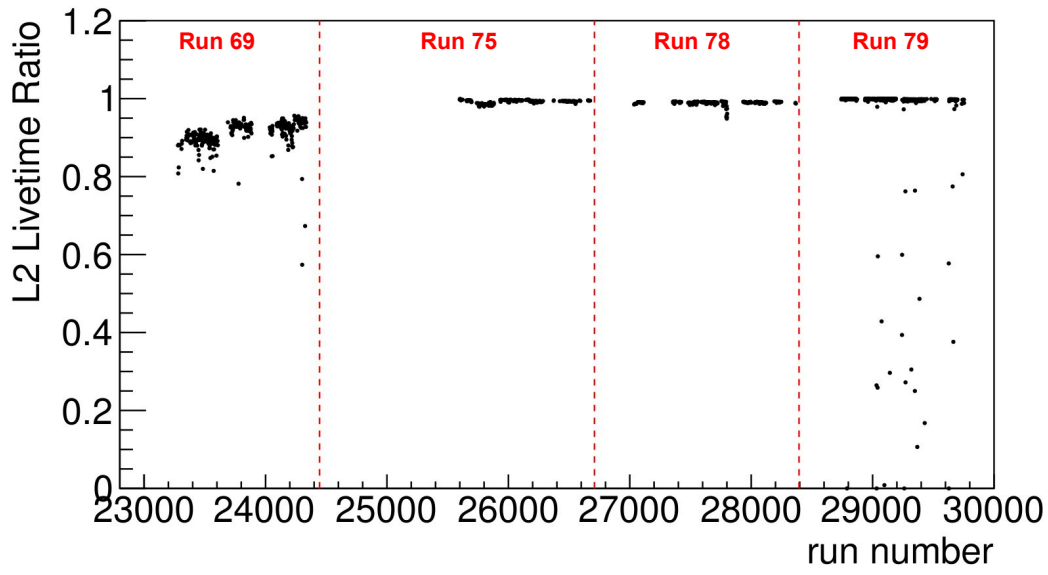


Figure 4.9: The L2 livetime ratio vs run number for run periods in 2016–2018. The largest contributor to the L2 livetime is the limited memory and buffer size in the L2 hardware.

4.5.3 L3 Livetime

The performance of the L3 trigger depends on several factors, including data passage, data quality, and the L3 event building performance. When event fragments are passed from the L2 trigger to L3, sometimes data packets are lost and so events cannot be completely built in L3 because of the missing information. Since the L3 trigger has a limited amount of time before the next spill cycle's data must be processed, some events may be unprocessed and contribute to the L3 dead time. The quality of the data may also be corrupted and this is considered in the L3 livetime calculation. Aside from the events intentionally removed by the L3 software cut, the L3 livetime is defined in Equation 4.3 as

$$\text{L3 Livetime Ratio} = \frac{\#\text{ events in offline analysis files}}{\#\text{ L2A}} \tag{4.3}$$

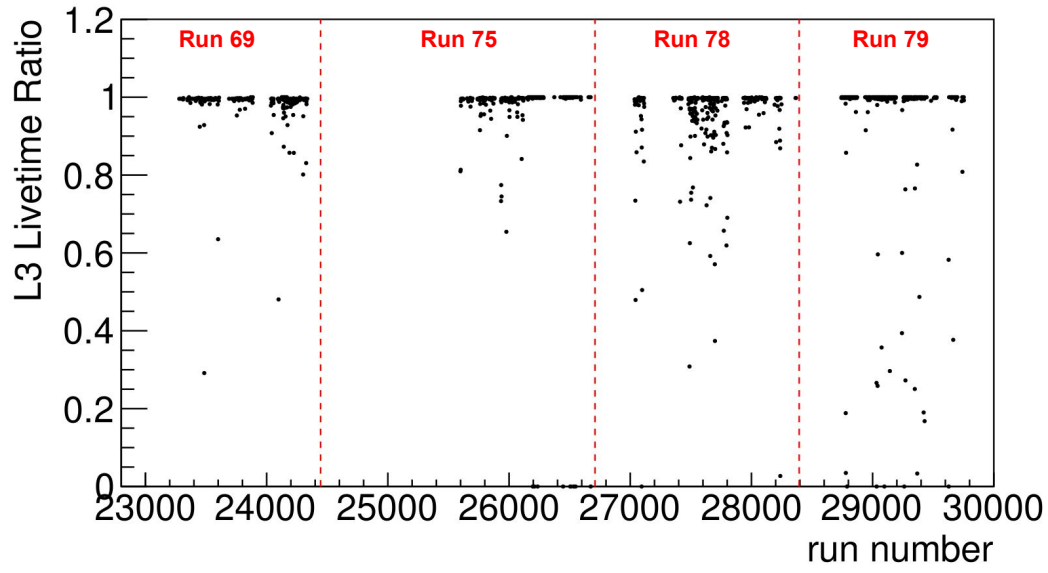


Figure 4.10: The L3 livetime ratio vs run number for run periods in 2016–2018. The L3 livetime depends on the data quality, event building performance, and the data passage from L2 to L3.

Since events can be lost in transit from the L2 trigger to the L3 trigger, the L3 livetime is defined as the faction of events that make it to the offline analysis files from the L2 trigger. The L3 livetime was consistent across all run periods in 2016–2018 as seen in Figure 4.10.

4.5.4 DAQ Livetime

To calculate the performance of the DAQ system as a whole, we simply multiply the L1, L2, and L3 livetime ratios. Thus, the overall DAQ livetime ratio is defined in Equation 4.4 and can be seen in Figure 4.11.

$$\text{DAQ Livetime Ratio} = \text{L1 Livetime} \times \text{L2 Livetime} \times \text{L3 Livetime} \tag{4.4}$$

The livetime from run period to run period is listed in Table 4.5.4. In Run 69 the livetime ratio was around 90% and by the end of Run 79, we were able to improve the DAQ livetime to around 99%. Even though the beam power increased overall

Run Period	Main Beam Power (kW)	DAQ Livetime Ratio
Run 69	44 kW	90%
Run 75	37.5 kW	97%
Run 78	50 kW	95%
Run 79	51 kW	99%

Table 4.4: The DAQ livetime ratio for each run period in the 2016–2018 data collection. Run 74 is excluded due to small statistics.

from 2016 to 2018, the livetime of the DAQ system also steadily increased with the DAQ upgrades, despite the higher trigger rates. The trigger rates seen in the new DAQ system (completed in 2018) with a beam power of 51 kW were as follows– the number of raw triggers based on E_t only was 480k/spill, after the L1 trigger selection the rate was suppressed to 22k/spill, and after the L2 trigger selection the rate was suppressed to 13k/spill. The new system dead time was measured to be 0.16 μs which was primarily due to the online cluster-finding algorithm. The efficiency of the new L2 trigger in collecting $K_L^0 \to \pi^0\nu\bar{\nu}$ events was estimated to be 99.6% based on MC simulations. More details on the performance of the new DAQ system can be found in [64].

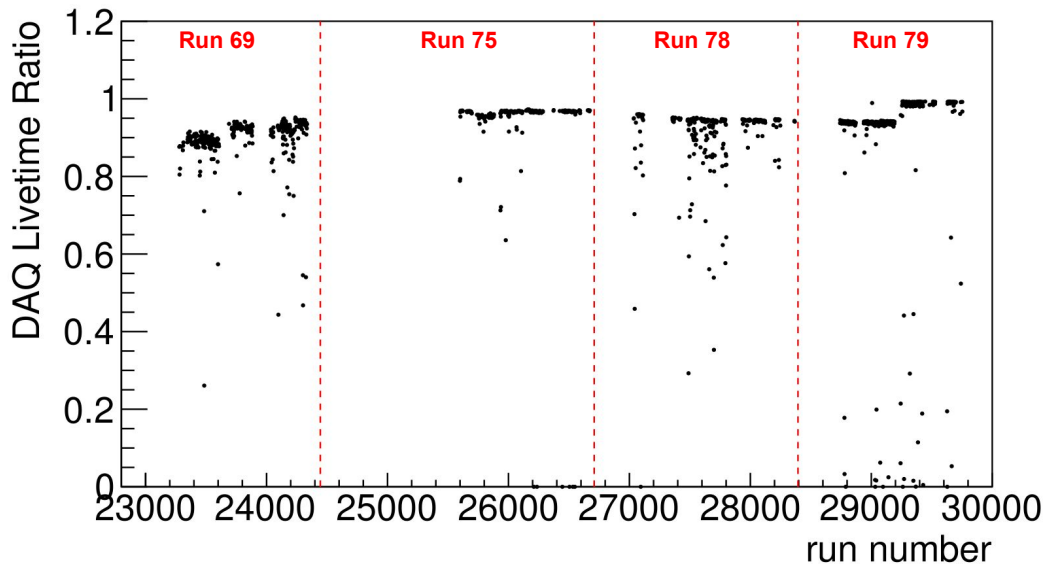


Figure 4.11: The overall DAQ livetime ratio vs run number for run periods in 2016–2018.

In summary, the KOTO experiment successfully collected 3.05×10^{19} POT of physics data from 2016–2018 which is 1.5 times more than what was collected in the 2015 runs. During this time we also improved the performance of the DAQ system from 90% livetime to 99% livetime despite increasing trigger rates due to higher beam power. After 2018, the CsI calorimeter was upgraded to have dual-ended readout and the details of these upgrades and further data collection are covered in Section 10.3. The KOTO experiment has continued to collect data each year since 2018, and will continue with data collection in the future. As the J-PARC accelerator continues to increase the beam power, improvements will be made to the DAQ to keep up with the increasing trigger rate and these plans are covered in Section 10.3.2.

Chapter 5

Event Reconstruction and Selection

After collecting the large amounts of data needed for the $K_L^0 \to \pi^0\nu\bar{\nu}$ search, the next task was to analyze the data. The KOTO data analysis involved three main steps: 1) event reconstruction and selection, 2) normalization analysis, and 3) background estimation and reduction. First, the physical properties of the events that deposit energy in the KOTO detectors were reconstructed from the waveform information saved in the raw data files. Once events were reconstructed, event selections were made and further analysis was carried out. In the KOTO analysis, we used Monte Carlo (MC) simulations to ensure that the data was reproducible and that the physics processes were well understood. The details of the Monte Carlo simulations are covered in Chapter 6. The normalization analysis (Chapter 7) was used to calculate the number of kaons produced and the background estimations (Chapter 8) were essential to making accurate cut selections to reduce background but retain signal events. This chapter will give an overview of the analysis methods and describe the event reconstruction and selection for the $K_L^0 \to \pi^0\nu\bar{\nu}$ search.

5.1 Overview of Analysis Strategy

The experimental strategy of KOTO (described in Section 2.1) is to observe the two photons from the π^0 decay on the calorimeter which have a large transverse

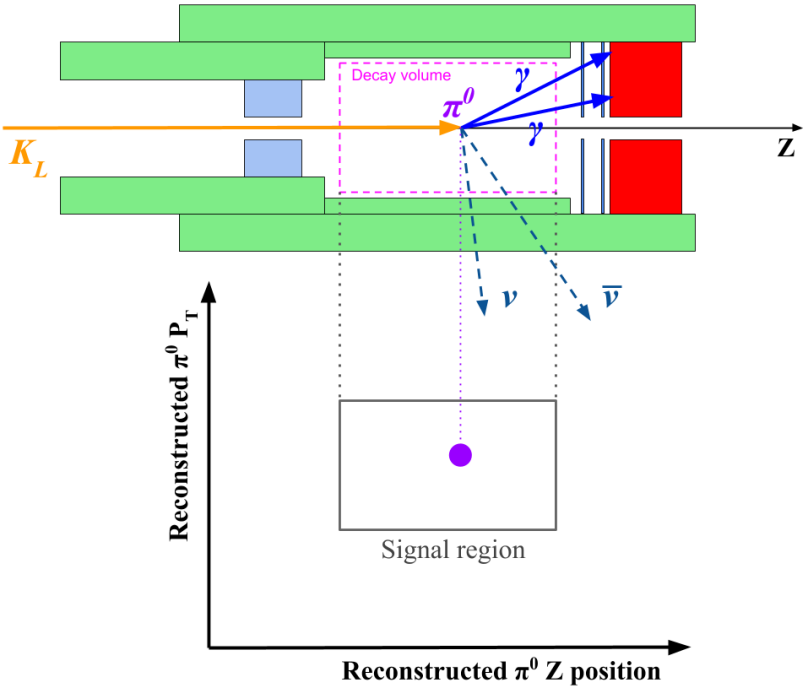


Figure 5.1: Schematic explanation of the signal identification in KOTO. The top of the figure shows a signal event inside the detectors in which the two photons from the π^0 are detected by the calorimeter (red). The π^0 z position and transverse momentum are reconstructed and the signal region lies in the P_T vs $Z_{\rm vtx}$ plane (bottom).

momentum with no other particles present in any detectors. The two neutrinos are not seen, but carry away momentum. The signal decay is identified by the properties of the π^0 in $K_L^0 \to \pi^0\nu\bar{\nu}$. By using the energies and the positions of the two photon hits on the calorimeter, the decay vertex along the z axis ($Z_{\rm vtx}$) and the transverse momentum (P_T) are reconstructed. The signal region is defined in the reconstructed π^0 P_T vs reconstructed π^0 $Z_{\rm vtx}$ plane as shown in Figure 5.1. This plot of P_T vs $Z_{\rm vtx}$ is often referred to as a "PtZ plot". To avoid any biases on selection criteria ("cuts"), a blind analysis method is used in which the signal region is covered with a black box until the final cuts have been made. Details of the signal region and blinded region are covered in Section 5.1.3. After the analysis is finalized, the unblinding of the signal region is called "opening the box" and the remaining events in the signal region are regarded as candidate events.

5.1.1 Background

Due to the extremely small branching ratio of $K_L^0 \to \pi^0\nu\bar{\nu}$, the principle challenge of the analysis is the reduction of background. Because of this, a thorough analysis of backgrounds is critical in order to observe the signal decay. The background estimations described in Chapter 8 are used to determine cut selections for the data and detailed studies are also done on how to reduce these backgrounds. In addition to estimating the background, preliminary work on a new background reduction method using machine learning to calculate the incident angles of the photon clusters on the CsI is detailed in Appendix A. In general, the backgrounds in the KOTO analysis are categorized into two groups, the K_L decay background and the neutron-induced background.

5.1.1.1 K_L Decay Background

The K_L decay background comes from other K_L decay modes. Table 5.1 shows the main decay modes of the K_L^0 and their branching ratios.

K_L Decay Mode	Branching Ratio
$K_L^0 \to \pi^\pm e^\mp \nu_e$	40.55 ± 0.11 %
$K_L^0 \to \pi^\pm \mu^\mp \nu_\mu$	27.04 ± 0.07 %
$K_L^0 \to 3\pi^0$	19.52 ± 0.12 %
$K_L^0 \to \pi^+\pi^-\pi^0$	12.54 ± 0.05 %
$K_L^0 \to \pi^+\pi^-$	$(1.967 \pm 0.010) \times 10^{-3}$
$K_L^0 \to 2\pi^0$	$(8.64 \pm 0.06) \times 10^{-4}$
$K_L^0 \to 2\gamma$	$(5.47 \pm 0.04) \times 10^{-4}$

Table 5.1: The main decay modes of K_L^0 and their branching ratios [10].

Although the decays $K_L^0 \to \pi^\pm e^\mp \nu_e$, $K_L^0 \to \pi^\pm \mu^\mp \nu_\mu$, and $K_L^0 \to \pi^+\pi^-$ can make two cluster hits in the CsI calorimeter, the CV detector in front of the CsI vetoes these events with charged particles in them. Therefore, K_L decays with π^0s in the final state are the main source of this background. In particular, the $K_L^0 \to 2\pi^0$ and the $K_L^0 \to \pi^+\pi^-\pi^0$ decays create a significant background because only two extra

particles can be used to veto these events. In the case of $K_L^0 \to 2\pi^0$, sometimes not all four photons hit the calorimeter. If this happens and two photons from the four photon final state are not detected by the veto detectors, this fakes a signal event. In the case of $K_L^0 \to \pi^+\pi^-\pi^0$, if the π^+ and π^- escape down the beam hole, this also creates a background, especially since the branching ratio for $K_L^0 \to \pi^+\pi^-\pi^0$ is so large. The $K_L^0 \to 2\gamma$ decay mode creates two clusters on the calorimeter like the signal decay, but can be rejected using kinematic cuts on the properties of the event, since the photons from $K_L^0 \to 2\gamma$ have no transverse momentum.

5.1.1.2 Neutron-Induced Background

The other source of background in KOTO is caused by halo neutrons in the beam. In this case, one of two things happen. Either halo neutrons hit detector material and produce particles that fake a signal event, or halo neutrons directly hit the calorimeter and produces hadronic showers that are mistaken for photons. In the latter case, a neutron hits the calorimeter and creates a primary hadronic shower and a neutron in the primary hadronic shower travels some distance inside the calorimeter and then produces a secondary hadronic shower. This is referred to as the "hadron-cluster background". The other neutron-induced backgrounds occur when a neutron hits detector material and produces either a π^0 or an η and they subsequently decay into two photons. More details about these backgrounds and their estimations are covered in Chapter 8.

The final piece in the analysis is the normalization, which is used to calculate the number of kaons at the beam exit. For the normalization studies we use three "normalization modes", $K_L^0 \to 3\pi^0$, $K_L^0 \to 2\pi^0$, and $K_L^0 \to 2\gamma$, to estimate the flux of K_L^0s into the KOTO detectors and calculate the total number of kaons. The normalization modes are also used to evaluate cut efficiencies and check the MC reproducibility of the data.

5.1.2 Branching Ratio and Single Event Sensitivity

There are three variables needed in order to estimate the branching ratio of $K_L^0 \to \pi^0\nu\bar{\nu}$ (Equation 5.1): the number of reconstructed $K_L^0 \to \pi^0\nu\bar{\nu}$ events (N_{signal}), the number of K_L^0 generated at the beam exit ($N_{K_L^0}$), as well as the acceptance of the signal mode (A_{signal}).

$$BR(K_L^0 \to \pi^0\nu\bar{\nu}) = \frac{N_{\text{signal}}}{N_{K_L^0} \times A_{\text{signal}}} \tag{5.1}$$

The number of signal events (N_{signal}) is obtained from the PtZ plot and the number of K_L^0s generated at the beam exit ($N_{K_L^0}$) is calculated using the normalization modes. The signal acceptance (A_{signal}) is estimated using simulations and includes the geometric acceptance of the detectors and the efficiencies of the cuts made on the data. Because the number of signal events is not known until the data is unblinded, the Single Event Sensitivity (SES) is defined as,

$$SES = \frac{1}{N_{K_L^0} \times A_{\text{signal}}} \tag{5.2}$$

and is a measure of how sensitive we are to observing a signal event. The smaller the SES, the higher the chance we are able to observe a signal event. Therefore, we want as many as possible K_L^0s generated and a high acceptance. If one or more signal events are observed, an estimation of BR($K_L^0 \to \pi^0\nu\bar{\nu}$) can be extracted from Equation 5.1, otherwise, an upper limit of the branching ratio can be calculated.

Most of the analysis steps are performed in parallel until the unblinding of the signal region. Figure 5.2 shows a picture of the analysis flow. The reconstruction of the energy and timing of the waveforms is performed first, and then the data goes through production where the detector calibration is done and the stability of the data

is evaluated. Further reconstruction of particles and clustering is then performed in order to check whether the calibration gives the expected results when compared with Monte Carlo. Several stages of production are made until the calibration is correct. Once this is done, the normalization analysis and background estimations are finalized in order to make the final cut selections and calculate the necessary quantities. Once the analysis is complete, the signal box is opened to reveal the result.

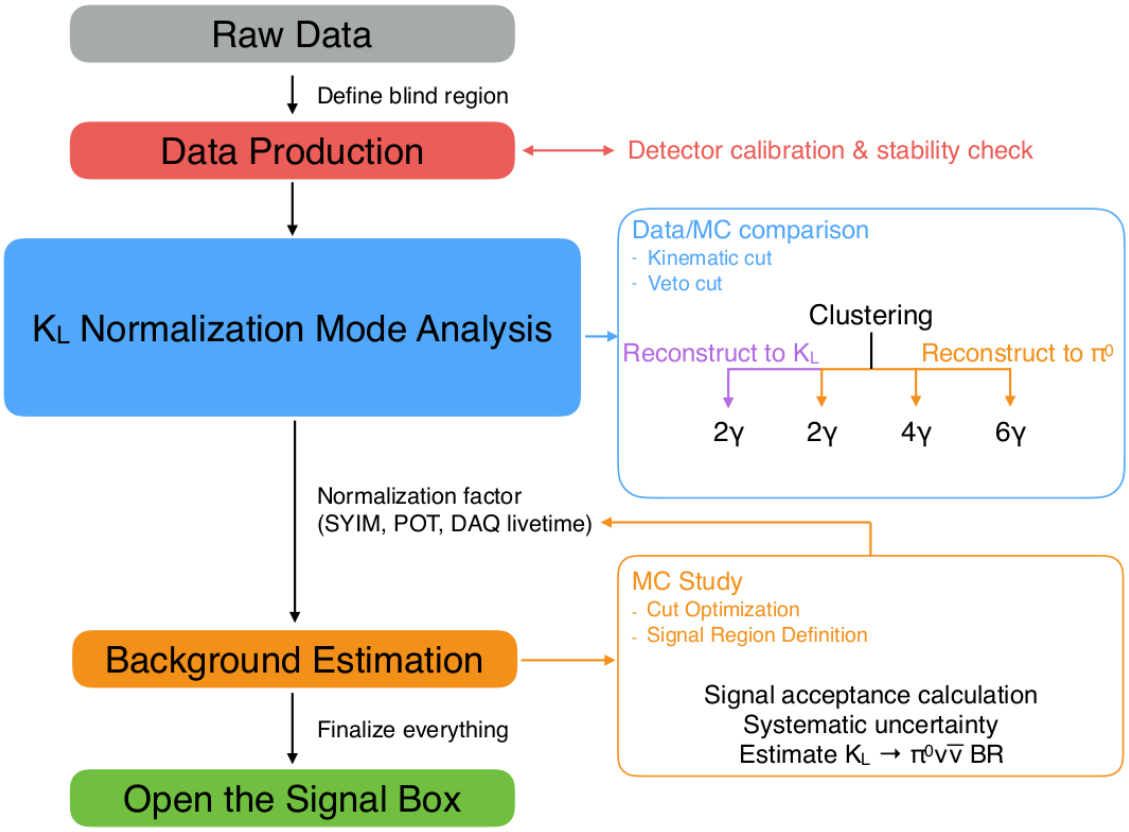


Figure 5.2: Flow chart of the KOTO analysis process. Raw data files are produced and the blind region is defined. The data goes through production for calibration and stability checks and clustering is performed in order to carry out the normalization analysis. The background estimation is done to optimize cuts and check the MC reproducibility of the data. Several stages of production are performed until this is achieved. Then the analysis is finalized and the signal box is opened. Figure courtesy of [54].

5.1.3 Signal Region and Blinded Region

For the analysis of the 2016–2018 data, the signal region was defined in order to maximize signal and minimize background and we used a blind analysis method to eliminate bias. The blinded region and signal region are shown in the left of Figure 5.3. We used MC simulations to estimate where the signal would be distributed in the P_T–Z plane as shown in the right of Figure 5.3 and also carried out detailed background estimations to understand where the backgrounds would be distributed. From these studies we defined the signal region as follows:

Signal Region

- 3200 mm $< Z^{\pi^0}_{\text{vtx}} <$ 5000 mm
- 130 MeV/c $< P^{\pi^0}_T <$ 250 MeV/c
- And a line drawn from 130 MeV/c to approximately 158 MeV/c across the $4000 < Z^{\pi^0}_{\text{vtx}} <$ 5000 mm range

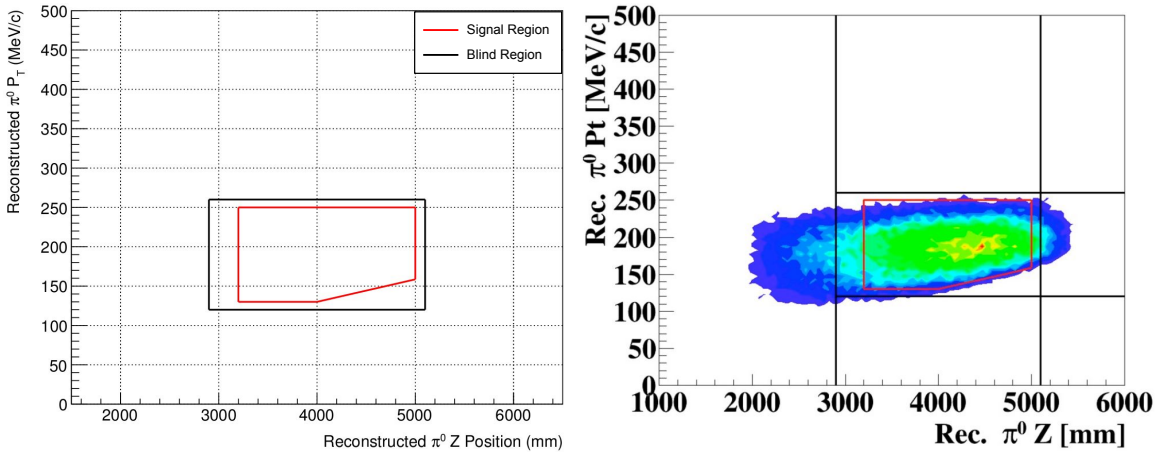


Figure 5.3: The signal region and the blinded region in the P_T–Z plane (left) and the MC distribution of $K^0_L \to \pi^0\nu\bar{\nu}$ events in the P_T–Z plane with the signal cut selections applied (right). In the right figure, the P_T–Z plane is split up into different control regions outside the blinded region in which to estimate the expected background contributions and compare it to the observed number of events.

The $Z^{\pi^0}_{\rm vtx}$ region was set in order to reduce the hadron-cluster background and other K_L decay backgrounds in the downstream region and the upstream π^0 background (a π^0 is generated at the NCC detector) in the upstream region as shown in the right of Figure 5.4. The lower bound of the signal box for the P_T range was set due to the $K^0_L \to \pi^+\pi^-\pi^0$ background, which occurs in the low P_T region. The diagonal cut across the bottom of the region as shown in Figure 5.3 was made in order to avoid this background. The upper bound in the P_T range was set using the signal distribution which has a distinctly higher P_T than the backgrounds as shown in the left of Figure 5.4.

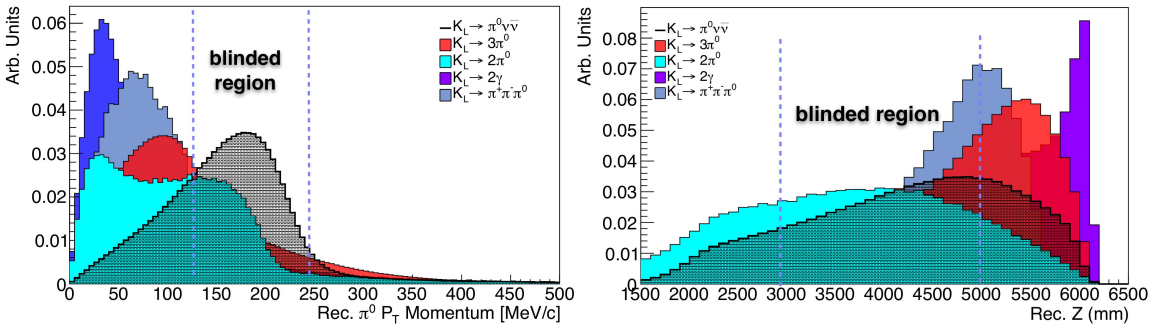


Figure 5.4: P_T and z distributions of the signal and various backgrounds without any cut selections that help determine the signal region. The left figure shows MC distributions in z of various K_L backgrounds and the signal (black). The right figure shows MC distributions in P_T for the signal (black) and various K_L backgrounds. Figure courtesy of [65].

The blind analysis method was used in order to reduce bias when optimizing cuts on the data. Since we regarded the events in the signal box as signal candidates, special treatment of any events in the signal region cannot occur if the region is blinded. In this way, we only use the region outside of the blinded region as a control to make informed cuts on the data and use Monte Carlo to estimate how much background resides inside the blinded region. The blinded region was defined as follows:

Blinded Region

- 29000 mm $< Z_{\text{vtx}}^{\pi^0} <$ 5100 mm
- 120 MeV/c $< P_T^{\pi^0} <$ 260 MeV/c

The blinded region stayed covered until the final cut selections were made. Events remaining in the signal region were regarded as signal candidates. The next section will cover the details of the event reconstruction. The event selections for the $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis are discussed in Section 5.3.

5.2 Event Reconstruction

The purpose of the event reconstruction was to go from raw waveform signals in the detectors to a full picture of the event, including the particles and their physical properties. Figure 5.5 gives an overview of the event reconstruction process. First, energy and timing information was extracted from the waveforms. From this information we identified clusters of energy on the CsI calorimeter to reconstruct photons (γ), the photon clusters were used to reconstruct pions (π^0), and then the pions were used to reconstruct kaons (K_L). Information in the veto detectors was also reconstructed and is used for veto decisions.

5.2.1 Energy and Timing Extraction from Waveform

When a CsI crystal or detector module was hit with a particle, the resulting signal was a waveform containing information about the energy that was deposited and the hit time. As described in Section 3.4, the analog waveforms were digitized by either 125 MHz or 500 MHz ADCs, which corresponds to a 512 ns event window. The following sections describe how the energy and timing information was obtained from detectors using each of these ADC modules.

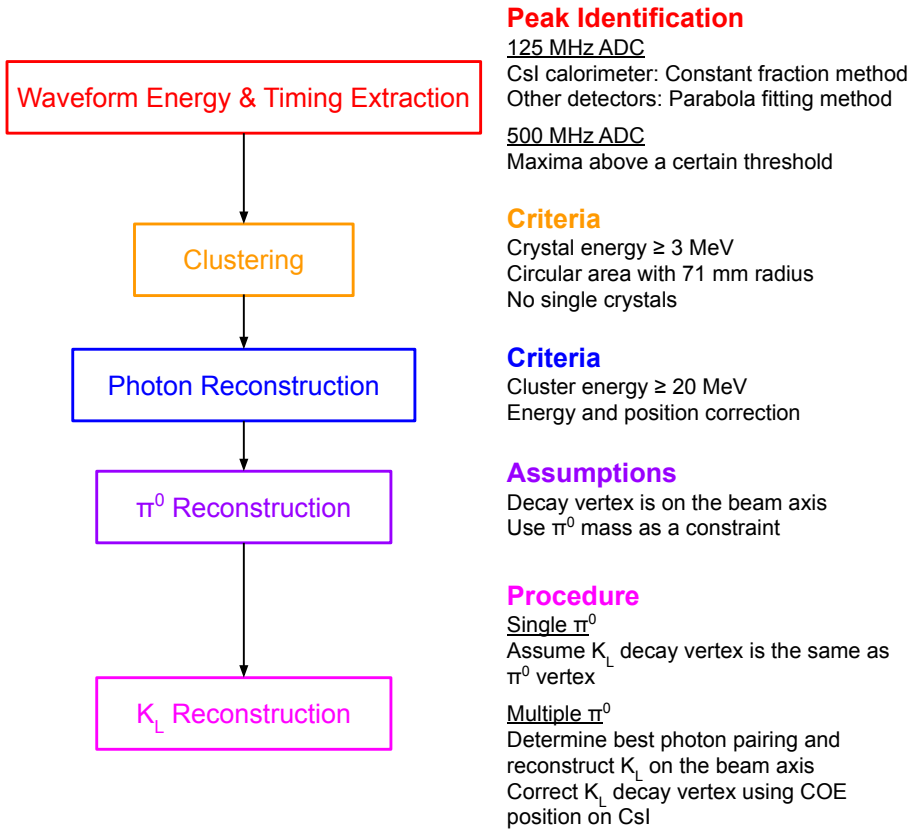


Figure 5.5: Overview and flow of the reconstruction procedure.

5.2.1.1 Detector Components with 125 MHz ADC Modules

The waveforms from the CsI calorimeter were digitized with the 125 MHz ADCs. Most of the other detectors also used the 125 MHz ADCs for waveform digitization, and a list of those that do can be found in Table 3.3. The waveforms collected by the 125 MHz ADCs were digitized in 64 samples with an 8 ns sampling rate, corresponding to a 512 ns event window. We expect a single pulse inside the event window, but in the case of multiple pulses, the pulse closest to the nominal time[1] was chosen for timing extraction. The "pedestal" was defined as the baseline of the waveform and was determined as the average of the first or last 10 samples. The average value with the smallest standard deviation was chosen as the pedestal value for that channel and the pedestal value was subtracted from each of the samples in the waveform.

[1]The nominal time was determined for each detector as the most probable timing of all channels.

Energy

The energy of the waveform was determined by integrating over all 64 samples after pedestal subtraction and then multiplying by a calibration constant. The calibration constant converted the ADC counts to energy in MeV. We determined the calibration constants separately for each detector channel by using the data from the cosmic muon runs, beam muon runs, and DVU Al target runs as described in Section 4.4.

Timing

Two different methods were used to obtain the timing information from the waveforms, the constant fraction method and the parabola fitting method. The constant fraction method was used for the CsI calorimeter because it gave a better timing resolution, which is important for the π^0 reconstruction. The parabola fitting method was used for all other detectors because it was better at distinguishing overlapping pulses in the case of pulse pileups, though this method has worse timing resolution. However, any timing shift due to overlapping pulses could lead to incorrect timing in the veto detectors, so we used the parabola fitting method to ensure that the veto timing was accurate. Since the calorimeter has a high granularity of CsI crystals, the pulse pileup rarely occurred in the CsI.

The constant fraction method calculates the timing based on the rising edge of the waveform pulse. In this case, the pulse time is defined when the rising edge of the pulse exceeds half of the maximum peak height. The parabola fitting method defines the timing based on the timing of the pulse peak. First the waveform is smoothed to reduce noise by taking a moving average. The peak of the waveform is then determined by fitting a parabolic function around the peak. In the event that there were multiple pulses in the event window, all pulses above a given threshold were fitted with the parabolic function and the one with timing closest to the nominal time was chosen to be the timing of the channel. Figure 5.6 shows examples of the

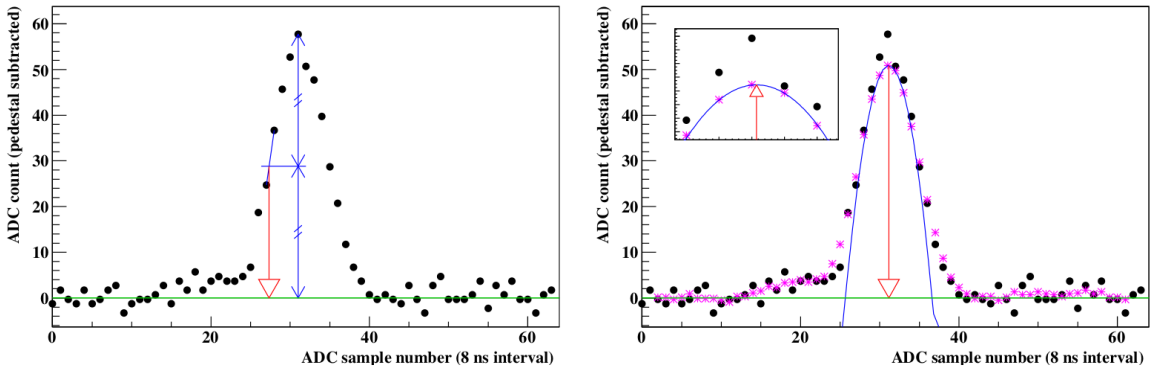


Figure 5.6: Waveform timing extraction with the constant fraction method (left) and the parabola fitting method (right). Black points are the waveform data, the green lines are the pedestal values, and the red arrows indicate the defined time. In the left figure, the blue line shows the peak of the waveform and the timing is determined by half of the peak height. In the right figure, the magenta points are the moving average of the waveform and the blue line is the fitted parabolic function using the three points around the peak. The timing of the waveform was determined by the location of the peak as shown in the zoomed panel. Figure courtesy of [47].

constant fraction method and the parabola fitting method for CsI channel waveform data. For both methods, a timing correction was applied for each channel using a timing calibration constant in order to correct for differences in timing offsets among different detector channels. The details of the timing calibration methods can be found in [47].

5.2.1.2 Detector Components with 500 MHz ADC Modules

The IB, new BHCV, BHPV, and the BHGC detectors used the 500 MHz ADCs for waveform digitization. In this case, the waveforms collected by the 500 MHz ADCs were digitized in 256 samples with a 2 ns sampling rate, corresponding to a 512 ns event window. These detectors that use the 500 MHz ADCs do so because they are close to the beam line and thus have a high counting rate requiring a faster sampling rate. We expect multiple pulses in the event window because of the high hit rate for these detectors. For this reason, every pulse in the event window was identified and we recorded the energy and timing information of each. Figure 5.7 shows an

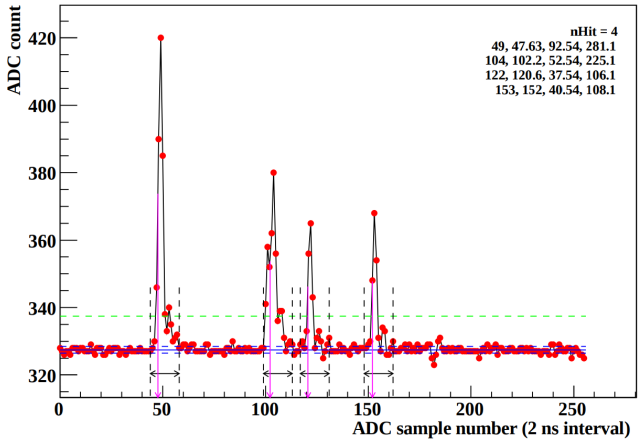


Figure 5.7: Example of energy and timing extraction from waveforms from the BHPV detector which has a 500 MHz sampling rate. The solid and dashed blue lines indicate the pedestal value and its 1σ region. The green dashed line shows the threshold for peak finding. The magenta arrows indicates the timing and the black arrows show the range for energy integration. Figure courtesy of [47].

example of multiple pulses in the BHPV detector. The pedestal value was defined by first fitting a Gaussian function to the waveform and then the most common value that appeared was set as the pedestal value. The pedestal value was then subtracted from each of the samples in the waveform. To identify pulses, we searched for local maximum points in the 256 samples above a detector specific threshold.

Energy

The energy of each pulse was calculated by integrating a range of samples around a pulse which was determined separately for each detector component. The calibration constant was used to convert the ADC counts to energy in MeV and was determined using the beam muon runs.

Timing

The timing for the waveforms digitized with the 500 MHz ADCs was determined using the constant fraction method and a timing calibration constant was used to correct the differences in timing among different detector channels.

5.2.2 Clustering

When a particle hits the CsI calorimeter, an electromagnetic shower is generated. Since the Molière radius is 3.57 cm for CsI [10] and the size of the (small) crystals are 2.5×2.5 cm^2, the shower spreads to multiple crystals. The group of crystals in which a particle deposits energy is called a "cluster" and the process of identifying clusters on the calorimeter is called "clustering". The procedure for clustering is outlined below and shown in Figure 5.8.

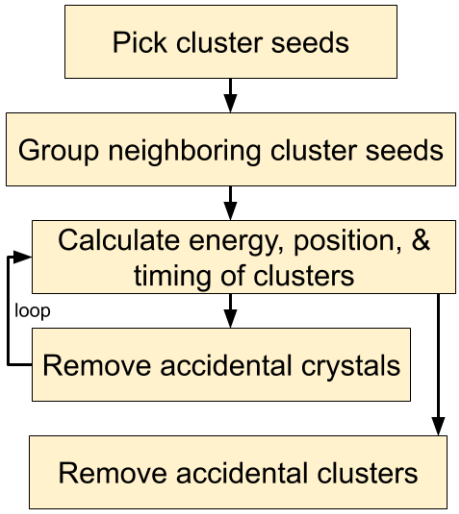


Figure 5.8: Flow chart of the clustering procedure. The crystals with deposited energy are grouped together into clusters, then the relevant quantities that describe the cluster are calculated. We used timing information of individual crystals to remove any crystals that do not belong to the cluster. The relevant quantities were calculated again and this process was iterated until all accidental crystals were removed. Finally, any clusters that did not meet the timing requirement for that event were removed.

First, crystals that had more than 3 MeV energy deposited and were within a 150 ns time window were selected as "cluster seeds". Cluster seeds with the highest energy were designated as core cluster seeds and then if any other cluster seeds were found within 71 mm of a core cluster seed, they were grouped together as a cluster as shown in Figure 5.9. Single cluster seeds were not considered clusters and are referred to as "isolated hit crystals" that were used to veto events, as described in Section

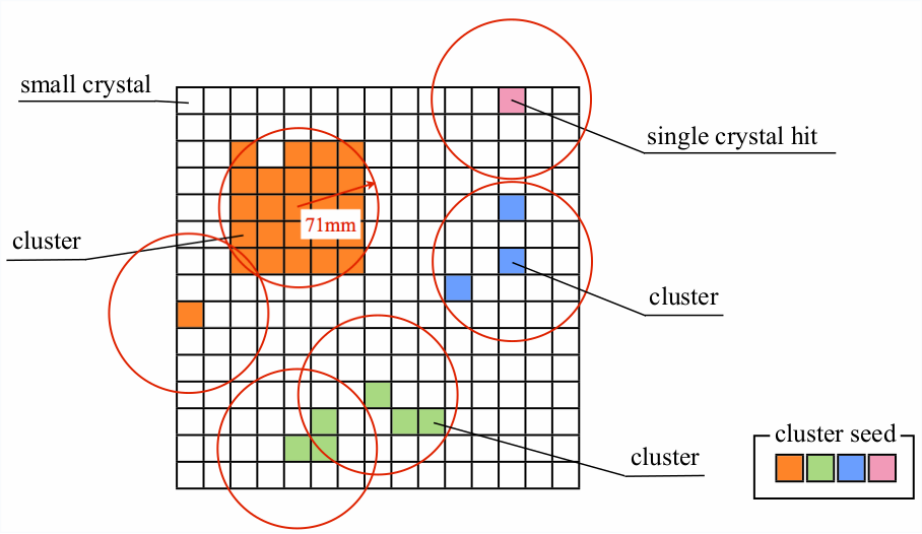


Figure 5.9: Examples of identifying clusters from cluster seeds. The different colors of the cluster seeds represent different clusters formed (orange, green, and blue) and the red circles represent the 71 mm radius. Since the pink cluster seed has no other cluster seeds around it, it is not classified as a cluster. Figure courtesy of [63].

5.3.4.1. After a cluster was defined, its energy (E_{cluster}), center of energy position in x and y (x_{cluster}, y_{cluster}), and timing (t_{cluster}) were calculated from the individual crystals in the cluster as shown in Equations 5.3, 5.4, 5.5, and 5.6.

$$E_{\text{cluster}} = \sum_{i=0}^{n} e_i \tag{5.3}$$

$$x_{\text{cluster}} = \frac{\sum_{i=0}^{n} e_i x_i}{\sum_{i=0}^{n} e_i} \tag{5.4}$$

$$y_{\text{cluster}} = \frac{\sum_{i=0}^{n} e_i y_i}{\sum_{i=0}^{n} e_i} \tag{5.5}$$

$$t_{\text{cluster}} = \frac{\sum_{i=0}^{n} t_i/\sigma_t^2}{\sum_{i=0}^{n} 1/\sigma_t^2} \tag{5.6}$$

where n is the total number of crystals a cluster contains and e_i, x_i, y_i, and t_i are the energy, x position, y position, and timing of the i-th crystal, respectively. The timing resolution of each crystal, σ_t, was measured in a past beam test [46] to be

$$\sigma_t(e_i) = \frac{5}{e_i} \oplus \frac{3.36}{\sqrt{e_i}} \oplus 0.13 \tag{5.7}$$

where σ_t is in ns, e_i is in MeV and $\oplus$ represents quadrature addition.

Once these quantities describing the cluster were calculated, the next step was to remove accidental crystal hits that were erroneously grouped into the cluster. This was done using the timing information of the crystal and the cluster timing. If any crystal had a hit timing outside of the range $t_{\text{cluster}} \pm 5\sigma(e)$, it was removed from the cluster and used for vetoing as described in Section 5.3.4.1. Figure 5.10 shows the timing difference between each crystal relative to the cluster timing as a function of the crystal energy. The two red lines represent the 5σ bounds, and any crystals outside of these bounds were removed from the cluster. This process was iterated until all of the crystals were contained within the 5σ allowed region.

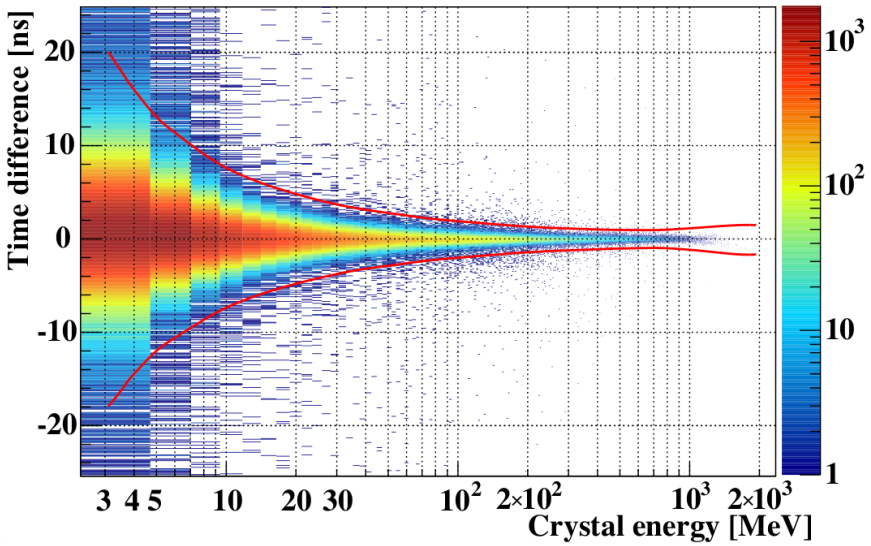


Figure 5.10: Distribution of crystal hit time relative to cluster hit time as a function of crystal energy. The red lines indicate the 5σ bound for t_{cluster} and the colors represent the number of events in arbitrary units. Figure courtesy of [63].

The last step of the clustering process was to remove any accidental clusters. For all clusters that hit the CsI, if the timing difference between the earliest and the latest cluster time was larger than 30 ns, the cluster with the largest time difference from the average timing of all the clusters was removed. This process was iterated until all of the clusters were contained within a 30 ns time window.

5.2.3 Photon Reconstruction

After clustering was completed, clusters that had energy larger than 20 MeV were defined as "photon (γ) clusters". The photon clusters were used to reconstruct the energy and hit position of the photon by making corrections to the photon cluster energy and position. An energy correction to the cluster itself is required because the energy of the cluster is usually smaller than the true energy of the photon due to shower leakage in the crystals as well as the finite energy threshold of 3 MeV used to select cluster seeds. The cluster energy was corrected using a correction map prepared by simulations.

5.2.3.1 Position and Energy Correction of Photon

Next, to obtain the energy of the photon and its position, a correction was made based on the approximate incident angle of the photon. Because the photons are neutral, there is no way to know for certain which direction they originate from. However, by using the reconstructed decay position of the π^0 or K_L where the photons were produced[2], a correction can be made. First, the π^0 or K_L was reconstructed to obtain the approximate decay vertex (described in Sections 5.2.4 and 5.2.5) and then the correction described below was applied. After applying the correction, the π^0 or K_L was reconstructed again.

Figure 5.11 shows a conceptual view of the position shift that occurs when a

[2]This is not necessarily true, as a number of backgrounds result from the assumptions we must make in the reconstruction.

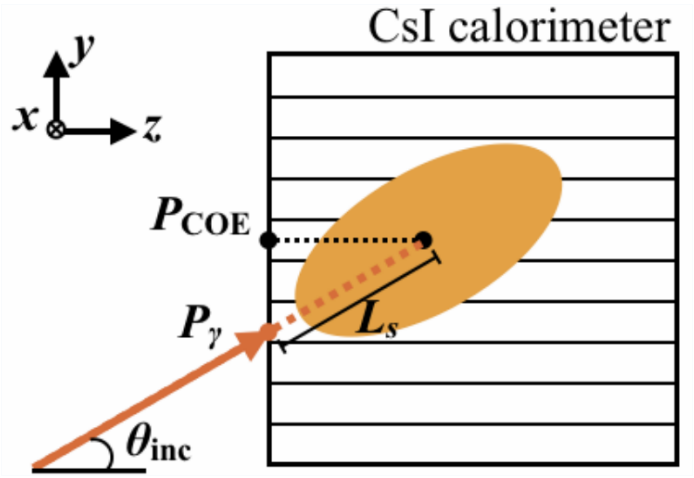


Figure 5.11: Conceptual view of true photon hit position (P_γ) compared to the COE position of the photon cluster (P_{COE}) due to the incident angle. Figure courtesy of [56].

photon hits the calorimeter at an angle. Since the photons are expected to originate from the beam line, they do not hit the calorimeter normal to the surface. Thus, the center of energy position (P_{COE}) of the photon cluster deviates from the true hit position of the photon (P_γ). P_γ is obtained as follows:

$$P_\gamma = P_{\mathrm{COE}} - L_s v \sin\theta_{\mathrm{inc}} \tag{5.8}$$

$$L_s/X_0 = p_0 + p_1 \ln(e_{\mathrm{clus}}) \tag{5.9}$$

where L_s is the shower length between the true hit position and the shower maximum position, $v = (\cos\phi_{\mathrm{inc}}, \sin\phi_{\mathrm{inc}})$ is the unit vector of the photon momentum in the $x-y$ plane, and θ_{inc} and ϕ_{inc} are the reconstructed polar and azimuthal angles of the photon momentum, respectively. $X_0 = 18.5$ mm is the radiation length of CsI, $p_0 = 6.490$ and $p_1 = 0.993$ are parameters obtained by simulations, and e_{clus} is the energy of the cluster in GeV. Finally, the energy of the photon was corrected using the correction map with incident angle information.

5.2.4 π^0 Reconstruction

The next step was to reconstruct π^0s assuming the photons came from the decay, $\pi^0 \to 2\gamma$. Again, since the photons are neutral, only the position and energy of the two photon hits on the calorimeter are known. In order to do a kinematic reconstruction of a two-body decay, we used the invariant mass of the π^0, $M_{\pi^0} = 134.9766$ MeV/c^2 [10], as a constraint as well as constrained the decay position to be on the beam axis (z axis). Using these constraints, the decay vertex z position of the π^0 ($Z^{\pi^0}_{\text{vtx}}$) can be calculated, and this can be used to calculate the transverse momentum of the π^0 (P_T) and the time at which the π^0 decayed, referred to here as the decay vertex time ($T^{\pi^0}_{\text{vtx}}$).

5.2.4.1 Decay Vertex Position

The geometry of the two-body decay, $\pi^0 \to 2\gamma$, is shown in Figure 5.12. The four momentum of the two photons (γ_1, γ_2) that hit the CsI can be written as

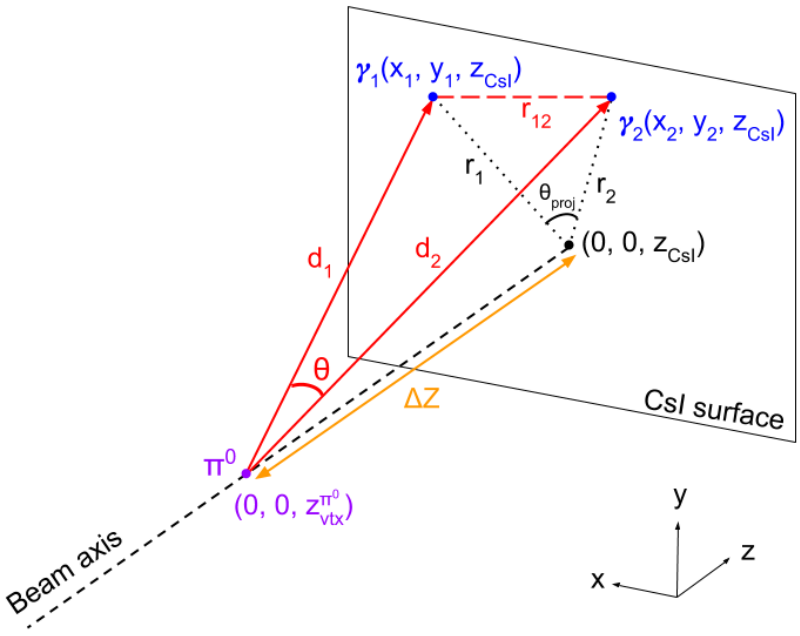


Figure 5.12: Schematic view of the π^0 reconstruction. The π^0 (purple) decays to two photons (blue). The red triangle is used in the reconstruction.

$(E_1, x_1, y_1, Z_{\mathrm{CsI}})$ and $(E_2, x_2, y_2, Z_{\mathrm{CsI}})$, and the four momentum of the π^0 is written as $(E, 0, 0, Z^{\pi^0}_{\mathrm{vtx}})$. The x and y momenta of the π^0 are both zero as we assume the π^0 decays on the beam axis. To calculate the decay vertex of the π^0, we used the geometry of the two-body decay, conservation of four momentum, and the law of cosines. The following geometrical relationships hold:

$$\Delta Z = Z_{\mathrm{CsI}} - Z^{\pi^0}_{\mathrm{vtx}} \tag{5.10}$$

$$r_1 = \sqrt{x_1^2 + y_1^2} \tag{5.11}$$

$$r_2 = \sqrt{x_2^2 + y_2^2} \tag{5.12}$$

$$d_1 = \sqrt{(\Delta Z)^2 + r_1^2} \tag{5.13}$$

$$d_2 = \sqrt{(\Delta Z)^2 + r_2^2} \tag{5.14}$$

$$r_{12}^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 \tag{5.15}$$

$$r_{12}^2 = d_1^2 + d_2^2 - d_1 d_2 \cos\theta \tag{5.16}$$

where the variables are shown in Figure 5.12. Conservation of four momentum for this two-body decay gives

$$M_{\pi^0}^2 = (E_1 + E_2)^2 - (\vec{p_1} + \vec{p_2})^2 \tag{5.17}$$

which can be expressed in terms of the opening angle, θ, between the two photons:

$$M_{\pi^0}^2 = 2E_1E_2(1 - \cos\theta) \tag{5.18}$$

Therefore, by assuming the π^0 invariant mass and constraining the decay to be on the beam axis, the opening angle between the two photons can be written as

$$\cos\theta = 1 - \frac{M_{\pi^0}^2}{2E_1E_2} \tag{5.19}$$

Then, using the law of cosines on the red triangle in Figure 5.12, a quadratic equation can be constructed for $(\Delta Z)^2$ as shown in Equation 5.20.

$$(1-\cos^2\theta)(\Delta Z)^4 + (2\vec{r_1}\cdot\vec{r_2} - (r_1^2+r_2^2)\cos^2\theta)(\Delta Z)^2 + (\vec{r_1}\cdot\vec{r_2})^2 - r_1^2r_2^2\cos^2\theta = 0 \tag{5.20}$$

$$\Delta Z = Z_{\text{CsI}} - Z_{\text{vtx}}^{\pi^0} \tag{5.21}$$

where $\vec{r_1}$ and $\vec{r_2}$ are the vectors for the photon hit positions on the calorimeter. When we solve for $Z_{\text{vtx}}^{\pi^0}$ using the two equations above and require it to be positive and real, there are two possible solutions. We choose the solution in which $Z < Z_{\text{CsI}}$ for $Z_{\text{vtx}}^{\pi^0}$.

5.2.4.2 Transverse Momentum

Once the π^0 reconstructed z vertex has been calculated, the transverse momentum of the π^0 can be calculated in units of MeV/c as

$$P_T = \left| \sum_{i=1,2} E_i \frac{\vec{r_i}}{\sqrt{r_i^2 + (\Delta Z)^2}} \right| \tag{5.22}$$

where E_i is the energy of the ith cluster, $\vec{r_i} = (x_i, y_i)$ and represents the two dimensional vector of each photon hit position, and $r_i = |\vec{r_i}|$.

5.2.4.3 Decay Vertex Time

The time at which the π^0 decayed can be calculated after the z vertex decay position is known. Since the arrival time of the photons to the surface of the calorimeter is known, we can compute the decay vertex time of the π^0 by using the positions of

the photon clusters on the CsI as well as the π^0 decay vertex position. The vertex time of each ith cluster is given by

$$T_{\text{vtx,i}} = t_i - \frac{\sqrt{r_i^2 + (\Delta Z)^2}}{c} \tag{5.23}$$

where t_i is the photon hit timing and c is the speed of light. Then, the π^0 decay vertex time is given by the weighted average of the ith photon cluster time for N number of clusters (Equation 5.24).

$$T_{\text{vtx}}^{\pi^0} = \sum_{i=1}^{N} \frac{T_{\text{vtx,i}}/\sigma_T^2(E_i)}{\sum_{j=1}^{N} 1/\sigma_T^2(E_j)} \tag{5.24}$$

where $\sigma_T^2(E)$ is the resolution for each vertex time as a function of energy and is given by Equation 5.25.

$$\sigma_T^2(E) = \frac{3.8}{\sqrt{e}} \oplus 0.19 \tag{5.25}$$

where the units for energy are in MeV and the units for time are in ns.

5.2.5 K_L^0 Reconstruction

After the π^0s were reconstructed, the next step was to reconstruct K_Ls. We reconstructed K_Ls for the normalization decay modes ($K_L^0 \to 3\pi^0$, $K_L^0 \to 2\pi^0$, and $K_L^0 \to 2\gamma$) and for the signal decay, $K_L^0 \to \pi^0\nu\bar{\nu}$. In this case, the K_Ls were reconstructed from either photons or π^0s. For the normalization modes, we expect $K_L^0 \to 3\pi^0$ to have six photons in the final state, $K_L^0 \to 2\pi^0$ to have four, and $K_L^0 \to 2\gamma$ to have two. Thus, we selected events that had a number of clusters that was either equal to or greater than the number of photons needed to reconstruct a particular decay mode. The clusters with the closest cluster timing were selected for

reconstruction and the others, if any, were used for vetoing, as described in Section 5.3.4.1.

5.2.5.1 Reconstruction of $K_L^0 \to 3\pi^0$, $K_L^0 \to 2\pi^0$, and $K_L^0 \to 2\gamma$

For $K_L^0 \to 2\gamma$ events, the K_L reconstruction procedure was the same as the $\pi^0 \to 2\gamma$ reconstruction in Section 5.2.4, except the π^0 mass was replaced with the invariant mass of the K_L, $M_{K_L} = 497.614$ MeV/c^2 [10]. To reconstruct the $K_L^0 \to 3\pi^0$ and $K_L^0 \to 2\pi^0$ decays, the correct pairing of photon clusters on the calorimeter to reconstruct the π^0s was necessary. Equation 5.26 gives the number of possible photon pairings

$$\text{Number of } \gamma \text{ pairings} = \prod_{i=1}^{N/2} \frac{\binom{2i}{2}}{(N/2)!} \tag{5.26}$$

where N is the total number of photon clusters and $\prod$ represents product notation. For $K_L^0 \to 2\pi^0$, there are three possible combinations of four photons forming two pairs, and for $K_L^0 \to 3\pi^0$, there are 15 possible combinations of six photons forming three pairs. For the K_L reconstruction, we assumed that the K_L decayed on the beam axis. Then, the reconstructed K_L decay vertex ($Z_{\text{vtx}}^{K_L}$) was calculated using the weighted average of each photon pairing combination as follows:

$$Z_{\text{vtx}}^{K_L} = \sum_{i=1}^{N/2} \frac{Z_{\text{vtx,i}}^{\pi^0}/\sigma_{\text{z,i}}^2}{\sum_{j=1}^{N/2} 1/\sigma_{\text{z,j}}^2} \tag{5.27}$$

where the reconstructed π^0 vertex ($Z_{\text{vtx,i}}^{\pi^0}$) and its position resolution ($\sigma_{\text{z,i}}$) were calculated for each ith π^0. To find the best pairing combination, the "pairing χ_z^2" was defined in Equation 5.28.

$$\chi_z^2 = \sum_{i=1}^{N/2} \frac{(Z_{\mathrm{vtx,i}}^{\pi^0} - Z_{\mathrm{vtx}}^{K_L})}{\sigma_{\mathrm{z,i}}^2} \tag{5.28}$$

Thus, the photon cluster pairing that had the smallest χ_z^2 was chosen as the best combination and its previously calculated $Z_{\mathrm{vtx}}^{K_L}$ was assigned to be the reconstructed K_L vertex position for the decay.

The final step in the K_L reconstruction for the normalization modes was a correction to the K_L decay vertex. It is not necessarily true that the K_L decayed on the beam axis, so after the initial calculation of $Z_{\mathrm{vtx}}^{K_L}$, a correction was made by using the x and y center of energy positions on the calorimeter and the creation point of the K_Ls, namely, the T1 target. A schematic drawing of this correction is shown in Figure 5.13 and the following equations are used to adjust the K_L decay vertex:

$$X_{\mathrm{COE}} = \frac{\sum_i^N E_i x_i}{\sum_i^N E_i} \tag{5.29}$$

$$Y_{\mathrm{COE}} = \frac{\sum_i^N E_i y_i}{\sum_i^N E_i} \tag{5.30}$$

$$X_{\mathrm{vtx}}^{K_L} = X_{\mathrm{COE}} \frac{Z_{\mathrm{vtx}}^{K_L} - Z_{\mathrm{target}}}{Z_{\mathrm{CsI}} - Z_{\mathrm{target}}} \tag{5.31}$$

$$Y_{\mathrm{vtx}}^{K_L} = Y_{\mathrm{COE}} \frac{Z_{\mathrm{vtx}}^{K_L} - Z_{\mathrm{target}}}{Z_{\mathrm{CsI}} - Z_{\mathrm{target}}} \tag{5.32}$$

where the center of energy positions on the calorimeter (X_{COE} and Y_{COE}) are calculated using the x and y positions of the ith photon (x_i and y_i), and the energy of

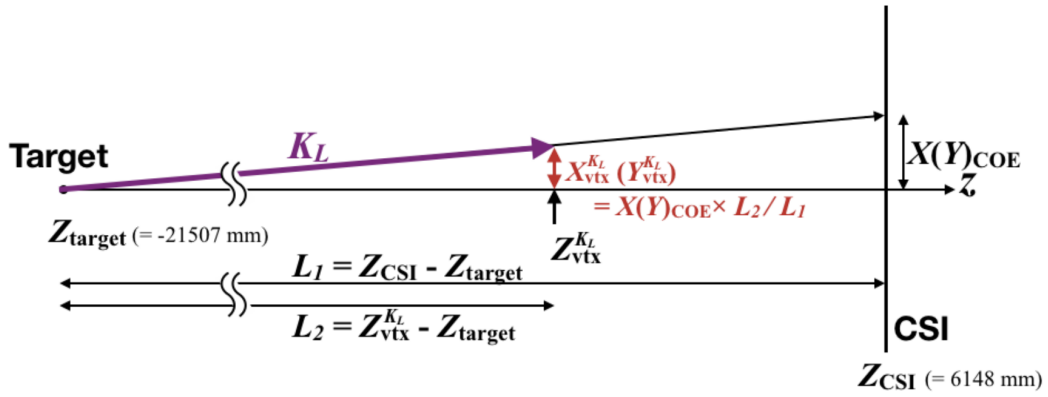


Figure 5.13: The correction to the K_L decay vertex. Figure courtesy of [56].

the ith photon (E_i). N is the number of photons. The K_L decay vertex in x and y are $X^{K_L}_{\text{vtx}}$ and $Y^{K_L}_{\text{vtx}}$, respectively, and the other variables are shown in Figure 5.13.

After determining the corrected K_L vertex position ($X^{K_L}_{\text{vtx}}$, $Y^{K_L}_{\text{vtx}}$, $Z^{K_L}_{\text{vtx}}$) for the $K^0_L \to 3\pi^0$ and $K^0_L \to 2\pi^0$ decays, we reconstructed the π^0 kinematic quantities again without using the π^0 mass as a constraint, and instead used the fixed K_L vertex position. Finally, the K_L was reconstructed using the four momentum of the π^0s instead of the K_L mass. Thus, distributions of of the reconstructed invariant mass for the π^0 and K_L can be obtained. After all the final corrections were made, the K_L decay vertex time was calculated using Equation 5.33 where the variables are the same as in Equation 5.24.

$$T^{K_L}_{\text{vtx}} = \sum_{i=1}^{N/2} \frac{T_{\text{vtx,i}}/\sigma^2_T(E_i)}{\sum_{j=1}^{N/2} 1/\sigma^2_T(E_j)} \tag{5.33}$$

5.2.5.2 Reconstruction of $K^0_L \to \pi^0\nu\bar{\nu}$

The signal decay reconstruction procedure is identical to that of the π^0, since $K^0_L \to \pi^0\nu\bar{\nu}$ only has one π^0 and the neutrinos go undetected. Since the lifetime of the π^0 is so short ($\sim 10^{-17}$ s), the distance it travels from the K_L is negligible, so we treat the K_L decay vertex position and time to be the same as that of the π^0.

5.3 $K_L^0 \to \pi^0\nu\bar{\nu}$ Event Selections

Once events were reconstructed, cuts were made on the physics data to remove background and identify $K_L^0 \to \pi^0\nu\bar{\nu}$ events. The criteria include data selections such as cuts to ensure good data quality, kinematic cuts that use information from the CsI calorimeter, veto cuts that use information from the veto detectors, and advanced cuts that use tools such as neural networks and deep learning in order to remove background. There are also some special cuts that deal with specifics such as removing trigger bias or dead channels. Some of the cut values were determined using information from the normalization analysis (Chapter 7) and background information (Chapter 8). Tables 5.2 and 5.3 at the end of Sections 5.3.3 and 5.3.4 include all of the cuts used for the 2016–2018 $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis.

5.3.1 Data Selection Cuts

Number of Clusters

To select $K_L^0 \to \pi^0\nu\bar{\nu}$ events, we chose events with two cluster hits on the CsI calorimeter.

Trigger Selection

We took different types of triggers in the physics runs, and so we selected only the physics trigger events (described in Section 4.3.1) intended to collect the $K_L^0 \to \pi^0\nu\bar{\nu}$ sample.

Data Quality

Since there were some issues during data collection or the data was sometimes corrupted, only events that were taken during good spills and good runs with no errors were selected. In this way, the data for the $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis was of good quality and excluded any data that was taken during unstable conditions.

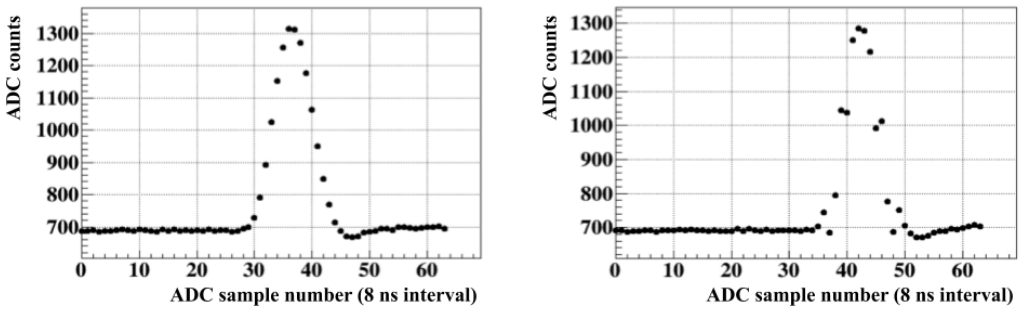


Figure 5.14: Example of a normal waveform (left) and a corrupted waveform (right). We used the second derivative value of the waveform to remove those that are abnormal.

Corrupted Waveform Removal

A small percentage of the waveforms in the 2016–2018 data were abnormal, as can be seen in the right of Figure 5.14 and so the strategy adopted was to remove them. Because of the abnormal nature, the corrupted waveforms tend to have larger second derivative values and so we used this feature to tag and remove the corrupted waveforms to ensure good data quality.

5.3.2 Trigger Bias Removal Cuts

These cuts eliminated the effect of the online trigger to prevent bias in the analysis.

Total γ Energy ($E_{tot,\gamma}$)

The sum of the energy of the two photons was required to be greater than or equal to 650 MeV to prevent trigger bias. Because of the lack of calibration information in the online trigger, the different gain performance of detector channels lead to a bias that was removed with this cut.

COE Radius

Similarly, the online COE radius cut introduced a bias as well, and so the offline COE radius was required to be greater than or equal to 200 mm.

Average Photon Cluster Time

In the event that two triggers occur within the same 64 sample window, both waveforms from two separate events would be overlapping. Because of this, the two events in the same window could be assigned the same timing which results in event duplication. To eliminate this double counting of an event, the average of the two cluster's timing was required to be within ±15 ns of the nominal trigger timing.

5.3.3 Kinematic Cuts

The kinematic cuts use the properties of the particles that hit the CsI calorimeter for event selection. They are split up into three categories– photon selection cuts that ensure good photon quality, background source cuts that distinguish $K_L^0 \to \pi^0\nu\bar{\nu}$ from other backgrounds with kinematic properties, and advanced cuts that use tools like machine learning to distinguish signal from background.

5.3.3.1 Photon Selection Cuts

Photon Energy (E_γ)

To ensure that only photons with good energy resolution were used in the reconstruction, we required the energy of each photon to be greater than or equal to 100 MeV and less than or equal to 2000 MeV.

Photon Cluster Position (CsI Fiducial)

To ensure that the electromagnetic showers were fully contained in the CsI calorimeter, any clusters that were too close to the outer or inner edges of the calorimeter were rejected. This cut is also referred to as the "CsI fiducial cut" as the photons were required to hit within the fiducial region of the CsI. Thus, the hit positions of each photon (x_γ, y_γ) were each required to be at least 150 mm away from the beam axis (0, 0) and the radius of the cluster hit position from the beam axis was required

to be no more than 850 mm.

Photon Cluster Distance

To prevent two clusters from fusing and being misidentified as a single cluster, we required the distance between two clusters to be greater than or equal to 300 mm.

Photon Cluster Distance from Dead Channels

We required the photon cluster hit position to be more than 53 mm away from any dead channels to prevent errors in the calculation of photon energies.

5.3.3.2 Background Source Cuts

Photon Projection Angle ($\theta_{\mathbf{proj},\gamma}$)

The opening angle of the two photon tracks projected onto the calorimeter surface ($x-y$ plane) is referred to as the projection angle and is shown in Figure 5.12 as θ_{proj}. We calculate the projection angle as follows:

$$\theta_{\text{proj}} = \cos^{-1}\left(\frac{\vec{r_1}\cdot\vec{r_2}}{|\vec{r_1}||\vec{r_2}|}\right) \tag{5.34}$$

The projection angle was required to be less than or equal to 150° in order to reduce the $K_L^0 \to 2\gamma$ background since the photons are back to back and the projection angle is 180° for the $K_L^0 \to 2\gamma$ decay.

$\mathbf{E_\gamma}$ Ratio

The energies of the two photons from $\pi^0 \to 2\gamma$ typically are not highly asymmetric. Therefore, we required the energy ratio of two photons (E_2/E_1) to be greater than or equal to 0.2 where $E_1 > E_2$. This cut reduces the mispairing of photons from the $K_L^0 \to 2\pi^0$ background, in which two photons are paired from different π^0s, referred to as the "odd pairing" (described in Section 8.3.2).

$\mathbf{E}_\gamma\ \theta_\gamma$

Similar to the energy ratio cut, this cut was used to reject the $K_L^0 \to 2\pi^0$ odd pairing background. In this case, the product of the photon energy and the angle between the reconstructed photon track and the beam axis was required to be greater than or equal to 2500 MeV·deg.

Photon Cluster Size

Cluster size is defined as the number of consecutive crystals contained in a cluster and we required the cluster size to be at least five crystals. This cut is effective in reducing hadronic clusters from neutrons because these clusters tend to have a smaller cluster size than photon clusters.

Photon Cluster RMS

Similar to cluster size, the root mean square (RMS) of the cluster is typically smaller for hadronic clusters compared to photon clusters. The cluster RMS is calculated using each crystal's energy and distance from the COE as shown in Equation 5.35.

$$\mathrm{RMS}_{\mathrm{clus}} = \sqrt{\sum_i \frac{e_i R_i^2}{\sum_j e_j}} \tag{5.35}$$

$$R_i^2 = (x_i - x_{\mathrm{COE},\gamma})^2 + (y_i - y_{\mathrm{COE},\gamma})^2 \tag{5.36}$$

$$x_{\mathrm{COE},\gamma} = \sum_i \frac{e_i x_i}{\sum_j e_j} \tag{5.37}$$

$$y_{\mathrm{COE},\gamma} = \sum_i \frac{e_i y_i}{\sum_j e_j} \tag{5.38}$$

where e_i represents the energy of the ith crystal, and x_i and y_i represent the x and y positions of the ith crystal, respectively. The distance between the ith crystal position and the COE position ($x_{\mathrm{COE},\gamma}$, $y_{\mathrm{COE},\gamma}$), R_i, is defined in Equation 5.36. For this cut, we require $\mathrm{RMS}_{\mathrm{clus}}$ to be greater than or equal to 10 mm.

π^0 Kinematic

In order to reject the "CV-η" background, resulting from a neutron hitting the CV detector and creating an η particle which decays into two photons, we set allowed regions in the $\mathrm{P}_{\mathrm{T}}^{\pi^0}/\mathrm{P}_z^{\pi^0}$–$Z_{\mathrm{vtx}}^{\pi^0}$ plane and the E_{π^0}–$Z_{\mathrm{vtx}}^{\pi^0}$ plane as shown in Figure 5.15. The variables $\mathrm{P}_z^{\pi^0}$ and E_{π^0} are the longitudinal π^0 momentum and π^0 energy, respectively.

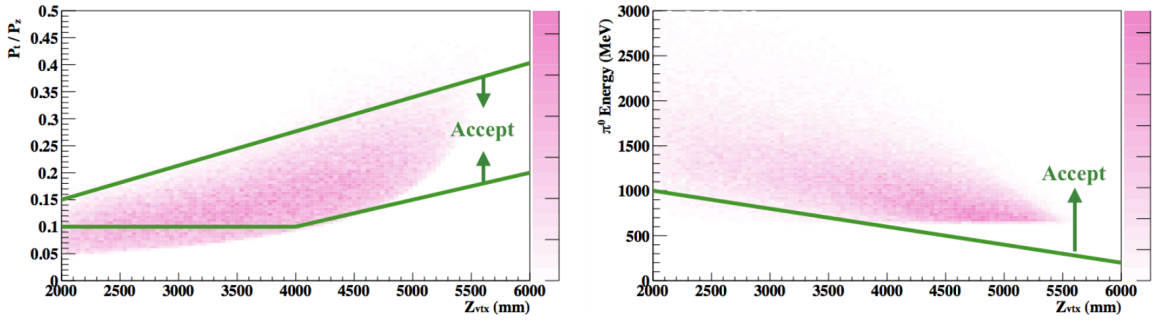


Figure 5.15: The accepted regions for the π^0 kinematic cut. The MC distribution of $K_L^0 \to \pi^0\nu\bar{\nu}$ events is shown in pink with the accepted regions bound by the green lines. Figure courtesy of [56].

Vertex Time Difference (ΔT_{vtx})

To ensure that the two photons originated from the same π^0, we required the vertex time difference between each photon and the reconstructed π^0 vertex time to be less than or equal to 1 ns. Equations 5.23 and 5.24 were used to calculate ΔT_{vtx} as follows:

$$\Delta T_{\mathrm{vtx}} = \left| T_{\mathrm{vtx,i}} - T_{\mathrm{vtx}}^{\pi^0} \right| \tag{5.39}$$

where the i represents the vertex time of the ith cluster. Also, in the reconstruction for $K_L^0 \to \pi^0\nu\bar{\nu}$ we assume the vertex time and position of the π^0 is the same as that of the K_L. Since any accidental hits or neutron hits would likely result in a difference of the reconstructed vertex time, this cut was effective at removing hadron-cluster background events.

5.3.3.3 Advanced Cuts

Neutron-Photon Cluster Shape Discrimination (CSDDL)

This cut was developed to reduce the hadron-cluster background (described in Section 5.1.1.2) and is referred to as the CSD Deep Learning (DL) cut. This method uses neural networks with deep learning to distinguish between neutron and photon clusters using cluster shape and kinematic information. The photon training sample used was $K_L^0 \to \pi^0\nu\bar{\nu}$ MC and the neutron training sample used was the Z0 Al target data (described in Section 4.4.1). The output of the network was a binary classification in which photons have a value of one and neutrons have a value of zero. A classification value greater than 0.985 for each cluster was required.

η Cluster Shape Discrimination (η CSD)

Similar to the CSDDL cut, the η CSD cut uses a neural network to distinguish between photons that come from $K_L^0 \to \pi^0\nu\bar{\nu}$ and photons that come from an η particle generated at the CV using cluster shape information. The signal training samples are $K_L^0 \to \pi^0\nu\bar{\nu}$ MC and the background training samples are CV-η MC in which an η particle originates from the CV detector. The output of the network is a binary classification in which signal photons have a value of one and background photons from the η particle have a value of zero. This cut required a classification value for each cluster greater than 0.91.

$K_L \to \pi^+\pi^-\pi^0$ Deep Learning ($\pi^+\pi^-\pi^0$ DL)

We used the $K_L \to \pi^+\pi^-\pi^0$ DL cut in order to reduce the $K_L \to \pi^+\pi^-\pi^0$ background in which the π^+ and π^- escape down the beam hole and the remaining π^0 produces two photons that hit the CsI. This cut uses a neural network with deep learning to distinguish between background and signal by using the features of both photons as well as the reconstructed π^0. The training sample for signal used was $K_L^0 \to \pi^0\nu\bar{\nu}$ MC and the background training sample used was $K_L \to \pi^+\pi^-\pi^0$ MC and the output of the network was a binary classification with signal being valued at one and background valued at zero. We required a classification value for each event to be greater than 0.922.

Reconstructed Polar Angle χ^2 (χ^2_θ)

To reduce the "CV-η" and "CV-π^0" backgrounds in which a neutron hits the CV and produces either an η or a π^0 which decays into two photons, we used the reconstructed polar angle of the cluster ($\theta_{\rm rec}$) and compared it to the output of a neural network that also calculated the polar angle ($\theta_{\rm NN}$) using features of the cluster. The neural network was trained with $K_L^0 \to \pi^0\nu\bar{\nu}$ MC. The consistency of the polar angle was quantified using the χ^2_θ value defined as

$$\chi^2_\theta = \frac{(\theta_{\rm rec} - \theta_{\rm NN})^2}{\sigma^2_{\theta_{\rm NN}}} \tag{5.40}$$

where $\sigma^2_{\theta_{\rm NN}}$ is the angle resolution in standard deviation of $\theta_{\rm NN}$. This cut required a χ^2_θ value less than 4.5 for each photon.

Fourier Pulse Shape Discrimination (FPSD)

This cut was designed to reduce neutron backgrounds by discriminating between pulses from electromagnetic showers (photons) and hadronic showers from neutrons.

A shower template was prepared from CsI waveform data and the pulses were compared to the template. We used $K_L^0 \to 3\pi^0$ data to prepare the template for the electromagnetic showers and used Z0 Al target data for the hadronic shower template. A Fast Fourier Transform (FTT) was performed on the waveforms and a variable called the pulse shape likelihood ratio characterized the waveform shape in the frequency domain. We required a value greater than 0.5 for the FPSD cut.

Cut Name	**Cut Requirement**
Data Selection Cuts	
Number of Clusters	$= 2$
Trigger Selection	= physics trigger
Data Quality	= good run, good spill, no error
Number of Corrupted Waveforms	$= 0$
Trigger Bias Removal Cuts	
$E_{tot,\gamma}$	≥ 650 MeV
COE Radius (offline)	≥ 200 mm
Average γ Cluster Time	$<$ nominal time $\pm$ 15 ns
Kinematic Cuts – *γ Selection Cuts*	
E_γ	$100 \text{ MeV} \leq E_\gamma \leq 2000 \text{ MeV}$
CsI Fiducial	$\lvert x_\gamma \rvert \geq 150$ mm, $\lvert y_\gamma \rvert \geq 150$ mm, $\sqrt{x_\gamma^2 + y_\gamma^2} \leq 850$ mm
γ Cluster Distance	≥ 300 mm
γ Cluster Distance from Dead Ch.	≥ 53 mm
Kinematic Cuts – *Background Source Cuts*	
$\theta_{proj,\gamma}$	$\leq 150°$
E_γ Ratio	≥ 0.2
E_γ θ_γ	≥ 2500 MeV·deg
γ Cluster Size	≥ 5
RMS_{clus}	≥ 10 mm
π^0 Kinematic	Accepted regions in Figure 5.15
ΔT_{vtx}	≤ 1 ns
Kinematic Cuts – *Advanced Cuts*	
CSDDL Value	> 0.985
η CSD Value	> 0.91
$\pi^+\pi^-\pi^0$ DL Value	> 0.922
χ^2_θ Value	< 4.5
FPSD Value	> 0.5

Table 5.2: The data selection, trigger bias removal, and kinematic cut requirements for the $K_L^0 \to \pi^0\nu\bar{\nu}$ event selection for the 2016–2018 data.

5.3.4 Veto Cuts

Since KOTO's experimental strategy is to observe two photon hits on the CsI and nothing else, we don't want any activity in any of the veto detectors. The criteria to veto events is based on the energy deposited in the veto counters, and also on the time any hits occur, since we don't want any extra hits in the veto detectors in coincidence with the two photon hits on the calorimeter. Energy thresholds and timing windows were set for each veto counter, some of which were based on background studies. Along with the veto detectors, information from the CsI calorimeter was also used to veto events. A full list of the veto cuts used in the 2016–2018 $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis is in Table 5.3.

5.3.4.1 CsI Calorimeter Veto Cuts

Isolated Hit Crystal

In Section 5.2.2, the process of clustering was described and it was noted that any isolated crystal was not considered a cluster. We used these isolated hit crystals to veto accidental hits not related to K_L events but that may have happened simultaneously. If the hit time of an isolated hit crystal was within $\pm$ 10 ns of the timing of its nearest photon cluster and the energy of the hit was larger than a given threshold, the event was rejected, otherwise the isolated hit crystal was ignored. The veto energy threshold, E_{isolated}, was determined using MC simulations and is given as a function of distance from the nearest cluster, d, as shown in Equation 5.41 and Figure 5.16.

$$E_{\text{isolated}} = \begin{cases} 10 \text{ MeV}, & (d \leq 200 \text{ mm}) \\ (13.5 - 0.0175d) \text{ MeV}, & (200 \text{ mm} < d < \dfrac{13.5 - T_{\text{sup}}}{0.0175} \text{ mm}) \\ T_{\text{sup}} \text{ MeV}, & (d \geq \dfrac{13.5 - T_{\text{sup}}}{0.0175} \text{ mm}) \end{cases} \tag{5.41}$$

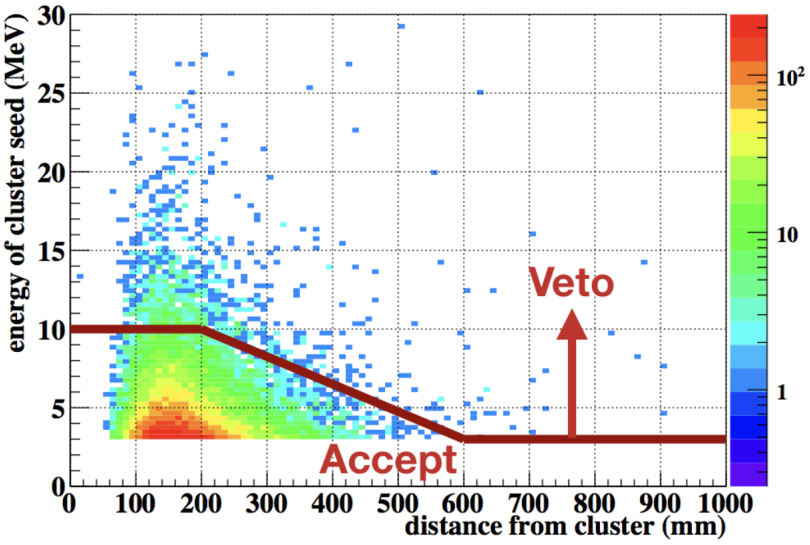


Figure 5.16: $K_L^0 \to \pi^0\nu\bar{\nu}$ MC simulation of isolated hit crystal energy as a function of distance from the nearest cluster with the isolated hit crystal veto threshold shown in red. In this figure the suppression threshold is 3 MeV. Any events above the red line are vetoed. Figure courtesy of [56].

where T_{sup} is the suppression threshold in MeV. The suppression threshold is determined individually for each run period and for each CsI channel and depends on the TMon data area Gaussian σ and the calibration constant of the channel. The threshold was set at a higher value for crystals with a smaller d in order to reduce the signal loss due to the isolated hit crystals that were not from accidental hits, but from gaps in the shower propagation.

Extra Cluster

Accidental activity in the decay volume can also cause extra clusters on the CsI, or extra clusters may be from $K_L^0 \to 3\pi^0$ or $K_L^0 \to 2\pi^0$ events. To reject these events with extra clusters, the reconstructed vertex time of these clusters was calculated and if it was within $\pm$ 10 ns of the event vertex time, these events were rejected. Since the cluster seed requirement was 3 MeV energy deposited and no single crystal was allowed, the minimum energy of an extra cluster was 6 MeV (two crystal cluster).

5.3.4.2 Veto Detector Cuts

The veto detector energy and timing extraction from waveform information was covered in Section 5.2.1. It should be noted that for the modules with single-end readout the energy and timing information of the detector channel was used as that of the module (module time and module energy). For the modules with double-sided readout, the energy and timing information was determined from both channels, and the reconstruction specifics for each dual-end readout veto detector can be found in [47].

To determine the veto timing of a detector, first the "module veto time" was determined for the different veto detector modules, which was based on the module hit time, the reconstructed vertex time, the geometry of the decay paths, and the time of flight. After the module veto time was calculated for each module, we took the final veto time to be the module veto time that was closest to the nominal time among all hits with energy larger than a certain threshold. The nominal time and energy threshold were determined separately for each detector. Finally, the veto energy was determined based on the module energy.

If a veto detector had a veto energy above a certain threshold and a veto time within the veto timing window, the event was rejected. The veto energy threshold and veto timing window were determined separately for each veto detector and are shown in Table 5.3. The wider the timing window, the larger the signal acceptance loss, but the narrower the timing window, the more difficult it is to distinguish between the actual pulse and an accidental pulse. Thus, in order to get the most out of the veto detectors, we used two veto timing windows for some detectors. The detectors with both narrow and wide timing windows were the FB, NCC, and CV. To determine which window to use, the waveform of the hit was investigated. Any hit with overlapping waveforms would benefit from a wider window while it would be better to use a narrow window for just a single pulse (see example of overlapped

waveforms in Figure 8.1). Thus, we extended the veto timing window to the wide window if the waveform had a double pulse.

To determine if there was a double pulse, we used the Fourier Transform Template (FTT) method. For the FTT method, we performed a Fourier transform on each waveform and compared it to a Fourier transform template of single pulse waveforms. The difference between the waveform and the template was evaluated using the χ^2 value and if the FTT χ^2 value was greater than a certain threshold, the wide timing window was used. For some veto detectors, the number of modules hit was also used in the veto decision, and these details are described below.

BHPV

The BHPV detector veto requirement depended on the number of modules hit. In particular, we required the hit coincidence (within a 10 ns window) of more than two consecutive modules in order to veto an event.

new BHCV

The veto requirements for the new BHCV detector were that more than one module had to be hit and the energy had to be above the veto threshold for an event to be rejected.

5.3.4.3 Dead Channel Treatment

There were two dead channels in the veto detectors, channel 23 in module 55 in the IB was dead (for Runs 74–79) and channel 127 in module 27 was dead in the IBCV. These modules required special treatment and so a study was performed to find the best veto threshold and timing window. As a result, the veto timing window was widened for these modules, as shown in Table 5.3.

Detector	Energy Threshold (MeV)	Time Window (ns)
CsI Veto Cuts		
CsI Isolated Hit Crystal	see Figure 5.16	nearest $t_{cluster} \pm 10$
CsI Extra Cluster	≤ 6	$\mid t_{cluster} - T_{vtx}^{K_L} \mid \geq 10$
Veto Detector Cuts		
CC03	≤ 3	32.2 ± 15.0
CC04 (Crystal)	≤ 3	4.31 ± 15.0
CC04 (Scintillator)	≤ 1	4.31 ± 15.0
CC05 (Crystal)	≤ 3	-24.3 ± 15.0
CC05 (Scintillator)	≤ 1	-24.3 ± 15.0
CC06 (Crystal)	≤ 3	-21.8 ± 15.0
CC06 (Scintillator)	≤ 1	-21.8 ± 15.0
FB (narrow)	≤ 1	$15.1 < t_{FB}^{veto} < 66$
FB (wide)	≤ 1	$-50.4 < t_{FB}^{veto} < 120.4$
NCC (narrow)	≤ 1	9.1 ± 20
NCC (wide)	≤ 1	9.1 ± 50
MB	≤ 1	$32.6 < t_{MB}^{veto} < 72.6$
IB	≤ 1	$-50.5 < t_{IB}^{veto} < -5.5$
CV (narrow)	≤ 0.2	54.1 ± 10
CV (wide)	≤ 0.2	54.1 ± 75
OEV	≤ 1	19.5 ± 10.0
LCV	≤ 0.6	14.9 ± 15.0
BPCV	≤ 1	21.4 ± 12.0
BHPV	# coin. modules hit ≤ 2	-76.2 ± 7.5
IBCV	≤ 0.5	16.9 ± 30.0
MBCV	≤ 0.5	19.4 ± 30.0
new BHCV	$\leq 2.21 \times 10^{-4}$ & # mod. hit ≤ 1	-93.5 ± 12.5
BHGC	≤ 2.5	-77.4 ± 7.5
HINEMOS	≤ 1	9.32 ± 20.0
Dead Channel Treatment		
IB Ch 55	≤ 1	-40.5 ± 80
IBCV Ch 27	≤ 0.5	$-33.1 < t_{IBCV}^{veto} < 116.9$

Table 5.3: The veto cut requirements for the $K_L^0 \to \pi^0\nu\bar{\nu}$ event selection for the 2016–2018 data. The veto timing window was arranged around the nominal time for a detector. The FB, NCC, and CV had both narrow and wide veto windows defined and the FTT method was used to determine which window to use. Events greater than the energy threshold and located inside the veto timing window were rejected.

Chapter 6

Monte Carlo Simulations

The Monte Carlo simulations, often abbreviated "MC", were an integral part of the $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis. We relied on these detailed simulations to understand how particles interacted in the KOTO detector system due to limited statistics in the data or limited knowledge about how decay processes were occurring. In this way, comparing the data to the Monte Carlo and determining if it was reproduced well gives confidence that the physics processes that occur in the MC were also occurring in the data. Since a blind analysis was used, the MC were also necessary to understand what the signal to background ratio was in the signal box before the unblinding of the data. In particular, the MC simulations were used to

- Carry out the normalization analysis (Chapter 7)
- Estimate the various backgrounds (Chapter 8)
- Evaluate the signal acceptance and the systematic uncertainties (Chapter 9)

The details of how the MC simulations were generated are covered in this chapter.

6.1 Overview

The Monte Carlo simulations were performed with version 9.5.2 of the Geant4 simulation toolkit [66; 67; 68], which is a package widely used in high energy physics

to simulate particle interactions in matter. Specifically, Geant4 uses Monte Carlo methods which rely on random sampling to obtain numerical results. Using Geant4, the KOTO detectors can be constructed in the simulation exactly as they are in the real experiment. The simulation proceeded in the following way– first, a beam of K_Ls with a known momentum spectrum was generated (Section 6.2). Then the decays of the incident K_Ls, their daughter particles, and their interactions in the detectors were simulated (Section 6.3). Finally, the simulated information was converted to the same format as that of the data while considering detector response (Section 6.5). At each stage of the simulation process, information such as energy, time, and position of each particle was recorded so we could understand exactly what was going on inside the detectors. We also factored in simulations of neutrons in the beam as described in Section 6.4.

6.1.1 Types of Simulation Methods

A few different kinds of simulation methods were used to generate MC for the analysis. Because producing a fully realistic simulation takes a significant amount of time and requires a large amount of disk space, we used several practical methods to generate MC more efficiently. The different MC methods are described below.

Full Simulation

The "full simulation" method recorded all information throughout the process and didn't skip or limit any steps. This method simulated all particles and interactions completely, and so it is the most realistic, but computationally expensive simulation method.

Fast Simulation

The "fast simulation" method limited the simulation to only simulate or compute the necessary quantities for a particular study in order to conserve resources. For

example, fast simulations were used for background studies on K_L decays that have large branching ratios in which the background mechanism occurs infrequently in the normal MC. In this case, we used the fast simulation to increase statistics of these particularly rare occurrences by requiring some features that were likely to produce the background process and only simulated and saved the events that satisfied these requirements. Limitations were placed on things like detector response, detector online thresholds, the order of particle track stacking, and the end points of particle tracks. Different requirements were made based on the needs of the analysis and this greatly reduced the time and disk storage space needed for the simulation.

Recycling Method

The purpose of the "recycling method" was to increase statistics. First, a sample of normal MC simulations were generated and events that satisfied some requirements were recorded as in the fast simulation method. The events that passed the requirements were then used as random seeds to generate more events with the same properties. Thus, utilizing the recycling method enhanced the MC statistics without consuming too many resources.

6.2 K_L Generation

Because it is extremely complicated and inefficient to simulate the generation of the K_Ls from protons incident on the T1 target, the K_Ls were generated at the beam exit. As described in Section 2.2.4 and shown in Figure 2.7, the beam exit was defined as the downstream end of the second collimator. For the generation we specified the momentum, position, and direction of the simulated K_L particles.

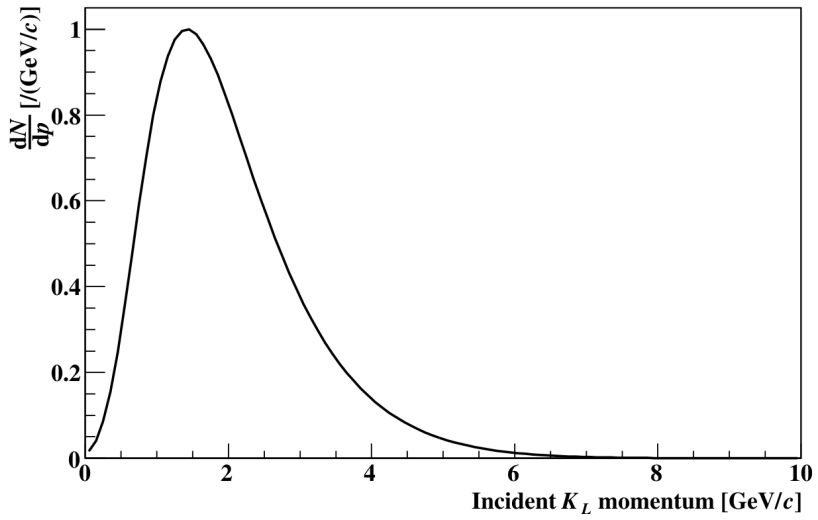


Figure 6.1: The K_L momentum spectrum used for the MC simulation. Figure courtesy of [47].

6.2.1 K_L Momentum Spectrum

The K_Ls were generated with a previously measured momentum spectrum that was obtained in 2012 by reconstructing $K_L^0 \to \pi^0\pi^+\pi^-$ and $K_L^0 \to \pi^+\pi^-$ decays using the CsI calorimeter and a spectrometer [46]. The resulting momentum distribution was fitted with the asymmetric Gaussian in Equation 6.1 and this function was used as the momentum spectrum of the simulated K_Ls.

$$\frac{dN}{dp} = \exp\left(-\frac{(p-\mu)^2}{2(\sigma_0(1-(A+Sp)(p-\mu)))^2}\right) \tag{6.1}$$

The variables and parameters in Equation 6.1 are defined as follows: dN/dp is the momentum distribution, p is the momentum of an incident K_L in GeV/c, μ is the peak momentum, σ_0 is the width of the distribution corresponding to the peak value μ, and A and S are asymmetric fit parameters for the Gaussian function. The values of μ, σ_0, A, and S are 1.420 GeV/c, 0.8102 GeV/c, -0.3014, and 0.01709 $(\text{GeV/c})^{-1}$, respectively. The momentum distribution is shown in Figure 6.1.

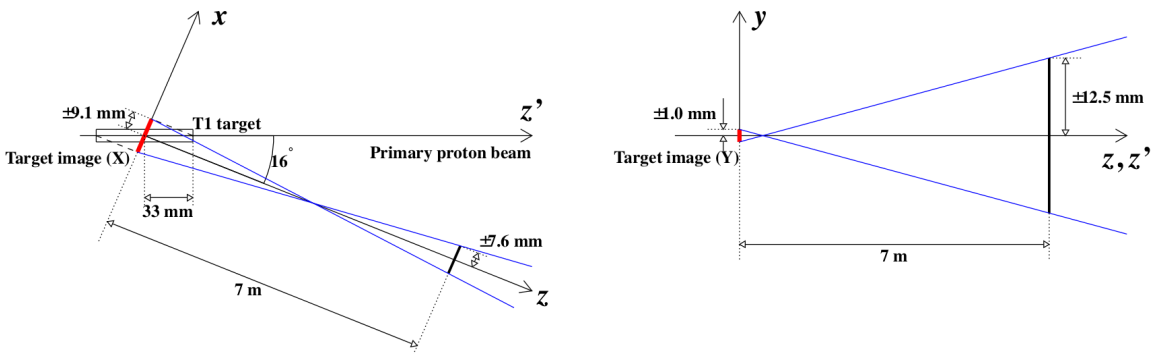


Figure 6.2: The illustrated criteria for generating K_Ls using the target image in the $x-z$ plane (left) and the $y-z$ plane (right). The target image is shown in red in each figure and the black bold lines at $z = 7$ m are the apertures of the collimator in x and y. The simulated K_Ls that pass the aperture points were used as incident K_Ls. Figure courtesy of [47].

6.2.2 K_L Incident Position and Direction

The position and direction distributions of the simulated K_Ls were determined using a simple beam optical simulation. This simulation randomly generated the K_L vertex position and its direction in a region referred to as the "target image" as shown in Figure 6.2. In this setup, z' is the direction of the primary proton beam and z is the direction of the K_L beam. The x axis is perpendicular to the z axis and lies in the z and z' plane and the y axis is perpendicular to this plane. The target image is the projection of the real target size onto the x and y plane and it has dimensions of $x = \pm 9.1$ mm and $y = \pm 1.0$ mm.

Once the K_L vertex was generated within the target image, the track was extrapolated in the downstream direction. If the track of the K_L passed the collimator at the aperture, the generated K_L information was used; if it did not, the generation process was repeated. The criteria for passing the collimator was determined using the narrowest point of the collimator, which was at $z = 7$ m from the center of the target in the x and y directions. The extrapolated x and y positions at this point were required to be within ± 7.6 mm and ± 12.5 mm, respectively.

6.3 K_L Decay and Particle Interaction with Detectors

After the K_Ls were generated, the next step was to simulate their decay and the decays of their daughter particles using Geant4. To accurately simulate the distribution of K_L decay vertices, the following probability function was used

$$P(L) = \exp\left(-\frac{L}{c\gamma\tau}\right) \tag{6.2}$$

where L is the length of the K_L flight path, c is the speed of light in a vacuum, γ is the Lorentz factor ($\gamma = 1/\sqrt{1-\beta^2}$), and τ is the mean lifetime of the K_L at rest. $\beta = v/c$ and v is the speed of the K_L which can be obtained from the K_L momentum. Interactions with materials such as air, vacuum windows, and detector materials along the flight path were considered and factored into the simulation.

If the K_L decayed, daughter particles were generated at the decay vertex and the behavior of each particle was simulated. The kinematics of the daughter particles were determined assuming the V–A interaction and the decay form factors were taken from [10]. To simplify the process, each decay mode of the K_L was simulated and analyzed separately and the branching ratio of each mode was factored in before obtaining any final results.

The interactions of particles with detector material were also simulated step by step in the Geant4 code. Interactions such as energy loss, scattering, generation of secondary particles, absorption, and particle decays were factored in by considering the cross section of each process. At each step in the simulation, five quantities were recorded: energy deposited (e_i), interaction time (t_i), and interaction position (x_i, y_i, z_i) where i indicates the step number. These quantities were used to obtain the visible energy and timing information in detector components as described in Section 6.5. The interaction processes were repeated for all generated particles until they lost

all of their kinetic energy or went outside of the decay volume. To simulate the often complicated hadronic interaction inside detector material, the QGSP_BERT physics list in Geant4 was used.

6.4 Neutron Generation

Along with kaons, we also simulated neutrons in the beam halo in order to accurately estimate the neutron-induced backgrounds as described in Section 5.1.1.2. In this case, neutrons with an energy of 1 GeV were generated at the position (1, 3.5, -1507) mm near the beam exit and propagated downstream. The neutrons that hit detectors (such as NCC or CV) and produced particles that could become a background were then used to increase statistics for the upstream π^0 background (Section 8.4.2), the CV-π^0 background (Section 8.4.3), and the CV-η background (Section 8.4.4. The simulation criteria and details of the estimation methods can be found in these background sections.

6.5 Detector Response

The quantities obtained from the Geant4 simulation such as particle interaction position, interaction time, and energy deposited can be thought of as the "ideal" or "true" information. However, in the real experiment, the detectors do not record the true information because of finite position resolutions, timing shifts due to light attenuation and propagation, and photon statistics effects. Thus, in order to make the simulations more realistic, these detector responses were also simulated.

In all of the KOTO detectors except for the BHPV and BHGC, which are Cerenkov detectors, there was a timing delay between when the particle actually hit the detector and when the scintillation light was received by the PMTs. This delay occurs because the scintillation light travels by multiple reflections inside the detector material as

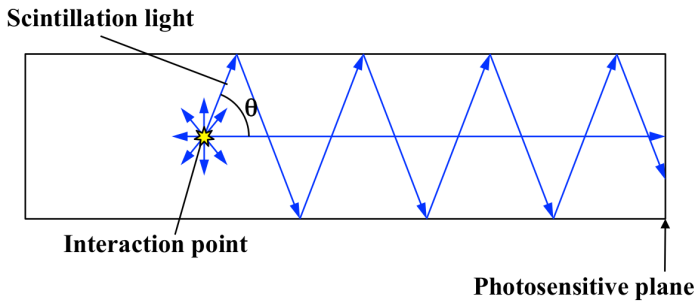


Figure 6.3: Example of light propagation inside a detector by multiple reflections. The delay between the hit time and detection time was corrected by using the interaction point position, the readout direction, and the effective speed of light calculated using the index of refraction and the emission angle, θ. Figure courtesy of [47].

illustrated in Figure 6.3. To correct for this effect, a delay was added to the simulated time of interaction based on the interaction position within the detector, the readout direction, and the "effective" speed of light in the scintillator which was measured for each detector. The effective speed accounts for the longer travel path due to multiple reflections by using the index of refraction and the emission angle, θ, as shown in Figure 6.3. For some detector components, an additional timing smearing effect was also applied using a Gaussian function. Since the particles in the Cerenkov detectors traveled faster than the speed of light inside the detector material, the effective speed of light correction was not necessary.

In addition to the timing correction, a correction was also applied to the energy information obtained from the simulation that was based on the hit position and took into account attenuation effects and the light collection efficiency of the detectors. The statistical fluctuation of the detected number of photons was also considered. In summary, the energy and timing response was considered for the CsI calorimeter, MB, IB, MBCV, IBCV, CV, FB, LCV, CC03, BHCV, and BPCV detectors. Only the timing response was considered for the OEV detector and there was no detector response implemented for the NCC, CC04, CC05, and CC06 detectors. The specific energy and timing correction details for each detector can be found in [47].

6.6 Waveform Simulations

Since the raw data from the detectors were analog waveforms that were digitized in 64 samples, we converted the simulated detector responses to waveforms with 64 samples so that the same analysis procedures could be applied for both data and MC. This allowed for the study of timing shifts due to overlapped waveforms caused by accidental activities in the beam. A waveform with 64 samples was created for each ith step in the simulation with energy and timing information corrected for detector response denoted as (e'_i, t'_i). The simulated waveform was formed by taking samples every 8 ns from the asymmetric Gaussian in Equation 6.3.

$$f(t) = \exp\left(-\frac{(t-\mu)^2}{(\sigma_0 + a(t-\mu)^2)^2}\right) \tag{6.3}$$

The variable t corresponds to the timing of the waveform sample in ns. The parameters σ_0 and a are the width and tail of the waveform, respectively, which were obtained by fitting the waveform of the real data channel by channel for each detector. The parameter μ is the peak timing of the waveform defined as

$$\mu = \mu_0 + t'_i - \delta t + t_{\text{offset}} \tag{6.4}$$

where μ_0 is the uncorrected peak timing, t'_i is the timing response of the ith simulation step, δt is the timing calibration constant, and t_{offset} is the timing offset adjustment to match the MC simulation timing to the timing of the data. Then the sum of the 64 corresponding values for $f(t)$ for step i was normalized to e'_i as shown in Equation 6.5, which represents the waveform collection, $\{a^i_n\}$.

$$a^i_j = e'_i \frac{f(8j)}{\sum_{k=0}^{63} f(8k)} \quad \text{where } (j = 0, 1, ..., 63) \tag{6.5}$$

Finally, the complete waveform information for each event was obtained by summing over all 64 samples in the waveform collection as shown in Equation 6.6.

$$A_j = \sum_i a_j^i \quad \text{where } (j = 0, 1, ..., 63) \tag{6.6}$$

6.7 Accidental Overlay

In the MC simulation, the K_Ls were generated and propagated downstream one at a time so that each recorded event contained a single K_L decay. However, pileup of multiple decays happens in the real world, especially with high beam power, and two or more K_L decays can occur very close to each other such that two K_L decays are recorded in the same event. This pileup of events can cause overlapped waveforms in which the original waveform is masked by another. Accidental activity in the beam can also cause this effect where the true hit is masked by other particles not related to the K_L decay. As a result, an overlapped pulse on top of the true pulse can cause a background if the measured time is shifted outside of the veto timing window. An example of this "masking background" can be seen in Figure 8.1 and is described in Section 8.2.

To reproduce this accidental activity in the MC simulation, we overlaid waveforms of accidental hits from the TMon trigger data (described in Section 4.3.5.5) onto the MC waveforms. The steps for this "accidental overlay" process are described as follows. First, a TMon trigger data file was randomly selected and then an event within the file was randomly chosen to begin the process. The waveforms in each MC event had the waveforms from an accidental event overlaid channel by channel. Because the process was random, some accidental events may have been used more than once for different MC events. For waveforms collected with the 125 MHz ADCs, the waveforms from the TMon data were added directly onto the generated MC

waveforms. For those waveforms collected with the 500 MHz ADCs, if the timing difference between the TMon data hit and the MC hit was less than 10 ns, the hit energies were merged and the timing was set to be the timing of the first hit. This procedure was used to reproduce the noise in each detector channel. After the accidental overlay was applied to the MC waveforms, the same procedures used for the data were used to calculate the waveform's energy and timing as described in Section 5.2.1.

Chapter 7

Normalization Analysis

The main purpose of the normalization analysis was to calculate the total number of kaons at the beam exit for the 2016–2018 data set as well as carrying out other important checks between data and MC. The total number of kaons, along with the signal acceptance, are used to calculate[1] the Single Event Sensitivity (SES) and subsequently the BR of $K_L^0 \to \pi^0\nu\bar{\nu}$, which are essential for the $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis. Three K_L decay modes, $K_L^0 \to 3\pi^0$, $K_L^0 \to 2\pi^0$, and $K_L^0 \to 2\gamma$, referred to as "normalization modes", were used for the analysis. We used these modes because they are easily reconstructed in the KOTO detectors and have high branching ratios relative to the signal decay. Specifically, the normalization modes were used to

- Calculate the total number of K_Ls at the beam exit
- Estimate the flux of K_Ls into the KOTO detectors
- Evaluate the kinematic and veto cut efficiencies
- Check the MC reproducibility of the data

This chapter details the steps of the 2016–2018 normalization analysis for each of the normalization modes and reports the final result on the total number of kaons.

[1] The calculation is described in Section 5.1.2.

7.1 Overview

The data collected with the normalization and minimum bias triggers described in Sections 4.3.2 and 4.3.3 were used for the normalization studies and we generated separate MC samples of the $K_L^0 \to 3\pi^0$, $K_L^0 \to 2\pi^0$, and $K_L^0 \to 2\gamma$ decay modes to compare to the data. The normalization MC samples had the accidental overlay (described in Section 6.7) applied and then the MC for each normalization mode went through the clustering and event reconstruction process described in Chapter 5. By generating the MC samples separately and using the event reconstruction process, we were able to determine which decay modes contributed when comparing the MC with the data and identify any "contamination" from other normalization modes. Other K_L decay modes such as $K_L^0 \to \pi^{\pm}e^{\mp}\nu_e$, $K_L^0 \to \pi^{\pm}\mu^{\mp}\nu_\mu$, and $K_L^0 \to \pi^+\pi^-\pi^0$ were also simulated and checked and were found to have small contributions to the normalization modes so they are not included in this analysis.

In order to purify the reconstructed data and MC, event selections for each normalization mode were made using a cut set similar to the $K_L^0 \to \pi^0\nu\bar{\nu}$ cut selection so that the event acceptance for the normalization analysis was as close as possible to the signal. The event selections used for the normalization analysis are covered in Section 7.2. After all of the cuts were applied, the data and MC were compared for each mode and the K_L flux and number of K_Ls were calculated as described in Sections 7.3 and 7.4, respectively.

Each step of the normalization analysis was done separately for each run period in which the accelerator conditions were different to account for different systematic uncertainties. In this case, the run periods for the 2016–2018 data described in Section 4.2 were further divided by beam power or run condition as shown in Table 7.1. The MC samples were also generated separately for each run period and the number of MC events generated for each period is also included in the table.

Run Period	Beam Power	# of MC Events	POT$\times 10^{18}$
Run 69 (0)	42 kW	2×10^8	1.780
Run 69 (1)	42 kW	2×10^8	1.335
Run 75 31 kW	31 kW	2×10^8	0.270
Run 75 35 kW	35 kW	2×10^8	0.582
Run 75 37.5 kW	37.5 kW	1×10^9	7.306
Run 78 33 kW	33 kW	2×10^8	0.509
Run 78 44 kW	44 kW	2×10^8	0.206
Run 78 50 kW	50 kW	1×10^9	7.359
Run 79 51 kW	51 kW	1×10^9	11.172

Table 7.1: The separate run periods for which the normalization analysis was done, along with beam power, the number of MC events used, and corrected physics POT for each run period. Run periods 75 37.5 kW, 78 50 kW, and 79 51 kW were the main run periods with large statistics so the number of MC events used for these periods was roughly an order of magnitude higher than the other periods. Run 69 was divided into two separate run periods before an amplifier for the FB was installed (0) and after the amplifier was installed (1).

7.2 Event Selection for the Normalization Analysis

The cuts applied on the data and MC for each normalization mode are described in this section. The kinematic cuts applied to each normalization mode were designed to maximized the purity of each mode and thus differed slightly from the kinematic cuts for the signal event selection. The veto cuts for the normalization analysis were the same as the veto cuts for the $K_L^0 \to \pi^0\nu\bar{\nu}$ event selection as described in Section 5.3.4 and Table 5.3. They were utilized in the normalization analysis to remove background due to accidental activity and other K_L decay modes in which all the photons do not hit the calorimeter. The kinematic cuts used in the normalization analysis are described below and summarized in Table 7.2.

7.2.1 Data Selection Cuts

The normalization analysis required two, four, or six clusters on the calorimeter and depending on the study, normalization or minimum bias trigger events were used. We required the same Data Quality and Corrupted Waveform Removal cuts

described in Section 5.3.1 to ensure the normalization data was of good quality and had no corrupted waveforms.

7.2.2 Trigger Bias Removal Cuts

In order to remove bias in the online trigger, we required the same Total γ Energy ($E_{tot,\gamma}$) and Average Photon Cluster Time cuts as described in Section 5.3.2. The COE Radius cut was not used because for the normalization and minimum bias triggers, the COE trigger cut was not required.

7.2.3 Kinematic Cuts

The kinematic cuts used for the normalization analysis are split into three categories– photon selection cuts that ensure good photon quality, cuts based on the properties of the reconstructed π^0s, and cuts based on the reconstructed K_L properties.

7.2.3.1 Photon Selection Cuts

The normalization analysis used the same Photon Energy (E_γ), Photon Cluster Position (CsI Fiducial), and Photon Cluster Distance cuts as described in Section 5.3.3.1 except the thresholds were different for E_γ and the Photon Cluster Distance. For the normalization modes we required E_γ to be at least 50 MeV and photon clusters to be at least 150 mm apart. The CsI Fiducial cut values were the same as the $K_L^0 \to \pi^0\nu\bar{\nu}$ event selection.

7.2.3.2 π^0 Cuts

Reconstructed π^0 Mass Difference (ΔM_{π^0})

For the $K_L^0 \to 3\pi^0$ and $K_L^0 \to 2\pi^0$ decays, the reconstructed π^0 mass was calculated assuming its decay vertex to be the K_L decay vertex. The reconstructed π^0 mass was required to be within 10 MeV/c^2 of the π^0 invariant mass for the $K_L^0 \to 3\pi^0$ decay

and 6 MeV/c^2 for the $K_L^0 \to 2\pi^0$ decay. This cut was applied in order to remove event reconstructions that used the wrong photon combinations for these decay modes.

Reconstructed π^0 Z Difference (ΔZ_{π^0})

The reconstructed π^0 z vertex was calculated for each π^0 and this cut required that the distance between the most upstream π^0 and most downstream π^0 was less than or equal to 400 mm. This helped eliminate events with bad pairing reconstructions in which the wrong photon combinations were used for the $K_L^0 \to 3\pi^0$ and $K_L^0 \to 2\pi^0$ decays.

7.2.3.3 K_L Cuts

Vertex Time Difference ($\Delta T_{\mathbf{vtx}}$)

This cut required that the timing difference between the reconstructed K_L time and the vertex time of each photon was less than or equal to 3 ns to ensure that two photons originated from the same π^0. The definition of this cut is exactly the same as the Vertex Time Difference described in Section 5.3.3.2, as the time of the π^0 vertex and the K_L vertex are the same for the signal decay.

Reconstructed K_L Mass Difference (ΔM_{K_L})

Similar to the Reconstructed π^0 Mass Difference cut, this cut required that the reconstructed K_L mass was within 15 MeV/c^2 of the K_L invariant mass for the $K_L^0 \to 3\pi^0$ and $K_L^0 \to 2\pi^0$ decays. Because the K_L mass cannot be calculated for the $K_L^0 \to 2\gamma$ decay and instead must be used as a constraint, there is no cut requirement for the $K_L^0 \to 2\gamma$ decay mode. This cut reduced $K_L^0 \to 3\pi^0$ and $K_L^0 \to 2\pi^0$ events with bad pairing reconstructions and the contamination of the $K_L^0 \to 3\pi^0$ mode in the $K_L^0 \to 2\pi^0$ analysis.

K_L Transverse Momentum (K_L P_T)

While the K_Ls ideally don't have any transverse momentum, we required the reconstructed K_L to have a P_T of no more than 50 MeV/c to ensure there were no missing particles.

K_L Reconstructed Z χ^2 (χ^2_{Z,K_L})

Equation 5.28 gives the definition of χ^2 for the reconstructed K_L z position which was required to be less than or equal to 20 for the $K^0_L \to 3\pi^0$ and $K^0_L \to 2\pi^0$ decays. This cut helped eliminate bad pairing combinations for these decays.

Reconstructed K_L Z Vertex ($Z^{K_L}_{\mathbf{vtx}}$)

The reconstructed K_L z vertex was required to be between 3000 and 5000 mm to ensure that the K_Ls used for the normalization analysis had no interaction with detector material outside of the decay volume.

K_L Beam Exit Position ($X^{K_L}_{\mathbf{Beam\ Exit}}$, $Y^{K_L}_{\mathbf{Beam\ Exit}}$)

For the $K^0_L \to 2\pi^0$ mode, the x and y positions of the K_Ls at the beam exit, ($X^{K_L}_{\text{Beam Exit}}$, $Y^{K_L}_{\text{Beam Exit}}$), were required to be within the beam core, which had a radius of 50 mm. The x and y beam exit positions of the K_Ls were calculated from the COE positions of the photons used in the reconstruction and then scaled to the z position of the beam line exit assuming that the T1 target is a point source fixed at $x = 0$, $y = 0$. Equations 7.1 and 7.2 give the x and y positions of the K_Ls at the beam exit.

$$X^{K_L}_{\text{Beam Exit}} = X_{\text{COE}} \frac{Z_{\text{exit}} - Z_{\text{CsI}}}{Z_{\text{target}} - Z_{\text{CsI}}} \tag{7.1}$$

$$Y^{K_L}_{\text{Beam Exit}} = Y_{\text{COE}} \frac{Z_{\text{exit}} - Z_{\text{CsI}}}{Z_{\text{target}} - Z_{\text{CsI}}} \tag{7.2}$$

where X_{COE} and Y_{COE} are defined in Equations 5.29 and 5.30, respectively, Z_{exit} is the z position of the beam line exit, Z_{CsI} is the z position of the upstream face of the CsI calorimeter, and Z_{target} is the z position of the T1 target. This cut was applied to reject $K_L^0 \to 2\pi^0$ events with missing energy.

Cut Name	$K_L^0 \to 3\pi^0$	$K_L^0 \to 2\pi^0$	$K_L^0 \to 2\gamma$
Data Selection Cuts			
**Number of Clusters	6	4	2
**Trigger Selection		norm. and min. bias triggers	
*Data Quality		good run, good spill, no errors	
*# Corrupt Waveforms		0	
Trigger Bias Removal Cuts			
*$E_{\text{tot},\gamma}$		≥ 650 MeV	
*Avg. γ Cluster Time		$<$ nominal time $\pm$ 15 ns	
Kinematic Cuts – *γ Selection Cuts*			
**E_γ		≥ 50 MeV	
*(CsI Fiducial)		$\lvert x_\gamma \rvert \geq 150$ mm, $\lvert y_\gamma \rvert \geq 150$ mm, $\sqrt{x_\gamma^2 + y_\gamma^2} \leq 850$ mm	
**γ Cluster Distance		≥ 150 mm	
Kinematic Cuts – *π^0 Cuts*			
ΔM_{π^0}	≤ 10 MeV/c^2	≤ 6 MeV/c^2	–
ΔZ_{π^0}	≤ 400 mm	≤ 400 mm	–
Kinematic Cuts – *K_L Cuts*			
**ΔT_{vtx}		≤ 3 ns	
ΔM_{K_L}	≤ 15 MeV/c^2	≤ 15 MeV/c^2	–
K_L P_T		≤ 50 MeV/c	
χ^2_{Z,K_L}	≤ 20	≤ 20	–
$Z^{K_L}_{\text{vtx}}$		3000 mm $\geq Z^{K_L}_{\text{vtx}} \leq 5000$ mm	
$(X^{K_L}_{\text{Beam Exit}}, Y^{K_L}_{\text{Beam Exit}})$	–	$\lvert X^{K_L}_{\text{Beam Exit}} \rvert, \lvert Y^{K_L}_{\text{Beam Exit}} \rvert \leq 50$ mm	–

Table 7.2: The data selection, trigger bias removal, and kinematic cut requirements for the 2016–2018 normalization analysis. The dash (–) indicates that the cut was not used for the corresponding normalization mode, otherwise the cut was applied for all normalization modes. Cuts with * indicate commonly used cuts for the $K_L^0 \to \pi^0\nu\bar{\nu}$ event selection with the same cut value and cuts with ** indicate commonly used cuts for the $K_L^0 \to \pi^0\nu\bar{\nu}$ event selection with a different cut value.

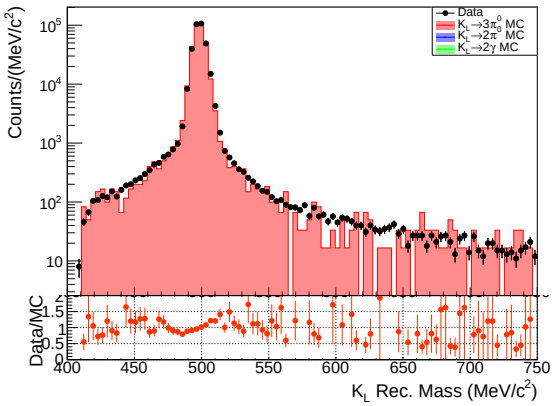
7.3 Normalization Modes

To ensure good agreement between the data and MC, the distributions of different kinematic quantities were checked for each normalization mode. For these distributions, all of the veto cuts and kinematic cuts were applied except for the kinematic cut on the corresponding kinematic variable being plotted. This section describes the normalization modes and shows the relevant kinematic distributions for each mode. All of the kinematic distributions shown were generated using normalization trigger data from Run 79 51 kW because this run period had the largest statistics. Each plot was normalized by area, meaning that the MC histograms were scaled to the area underneath the data histogram.

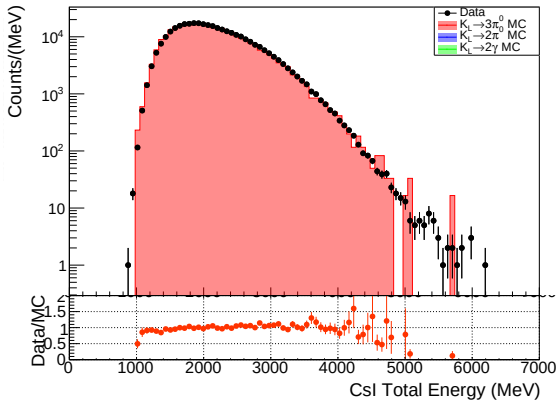
7.3.1 $K_L^0 \to 3\pi^0$

The $K_L^0 \to 3\pi^0$ decay is the cleanest normalization mode, due to the fact that it has six photons in the final state making it almost completely free from background contamination. This fact, along with its large branching ratio of 19.52% [10], makes $K_L^0 \to 3\pi^0$ an excellent decay mode for the normalization analysis. Figures 7.1 and 7.2 show the distributions for each of the kinematic variables used in the normalization event selection. The purity of the $K_L^0 \to 3\pi^0$ data was almost 100%, with contributions from $K_L^0 \to 2\pi^0$ and $K_L^0 \to 2\gamma$ being negligible.

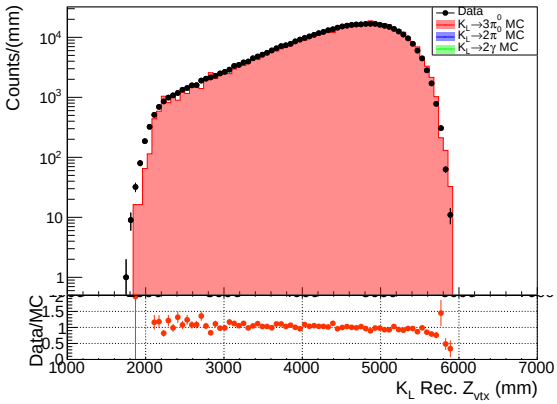
The reconstructed mass of the K_L as shown in Figure 7.1a has a sharp peak at the invariant K_L mass (497.614 MeV/c^2 [10]) and the data and MC are in good agreement at this point, though a slight shift of the peak can be seen, the source of which is not understood. The tails of the distribution were caused by the mispairing of photons when forming π^0s. The reconstructed decay vertex of the K_L (Figure 7.1c) is distributed from 2000 to 6000 mm, corresponding to the NCC detector and the front face of the CsI, respectively. The reconstructed P_T of the K_L is relatively small and results from the divergence of the K_L beam as shown in Figure 7.1d.



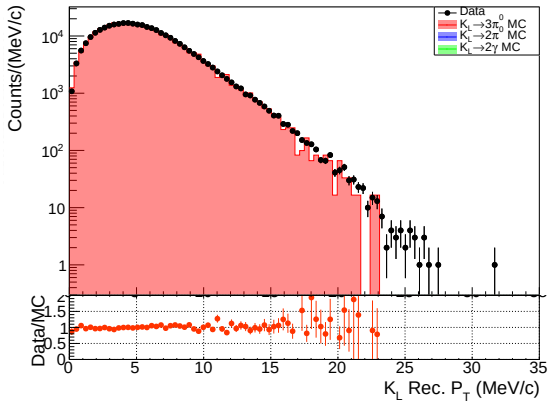
(a) K_L reconstructed mass



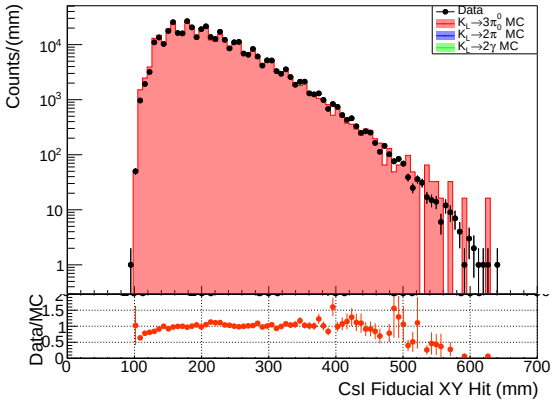
(b) Total energy deposited in CsI



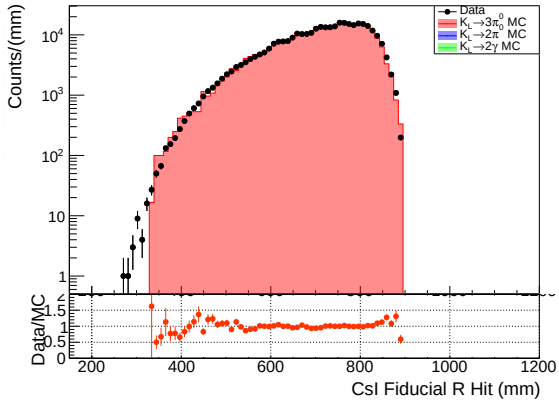
(c) K_L reconstructed z vertex



(d) K_L reconstructed P_T

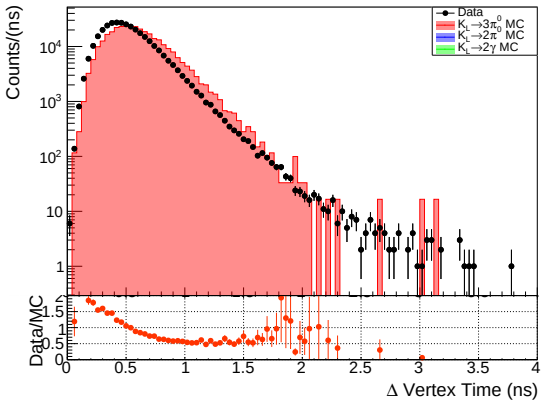


(e) Innermost photon cluster position (XY)

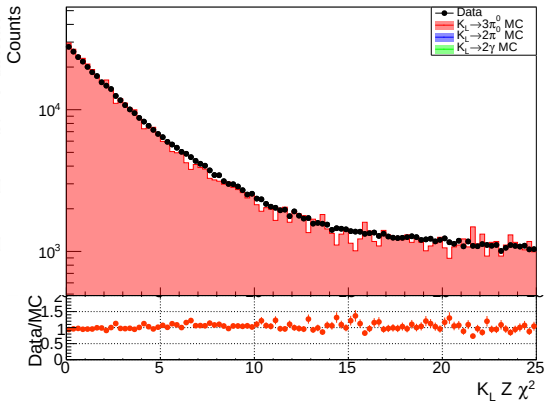


(f) Outermost photon cluster position (R)

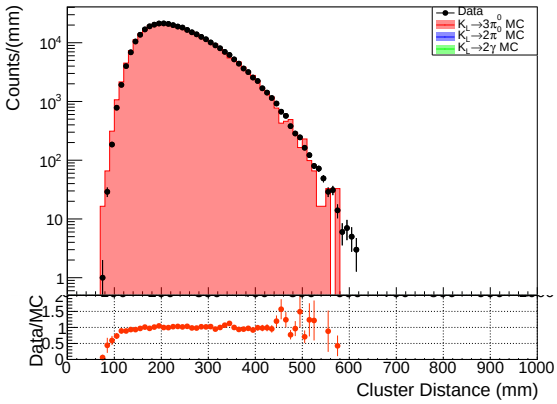
Figure 7.1: Kinematic distributions for reconstructed $K_L^0 \to 3\pi^0$ decays. Black points represent data and the colored histograms represent MC ($K_L^0 \to 3\pi^0$ = red, $K_L^0 \to 2\pi^0$ = blue, $K_L^0 \to 2\gamma$ = green). The data to MC ratio is shown in the bottom panel and the error bars represent statistical error.



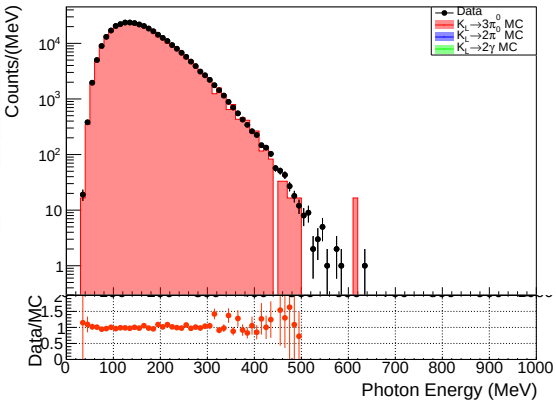
(a) Vertex time difference



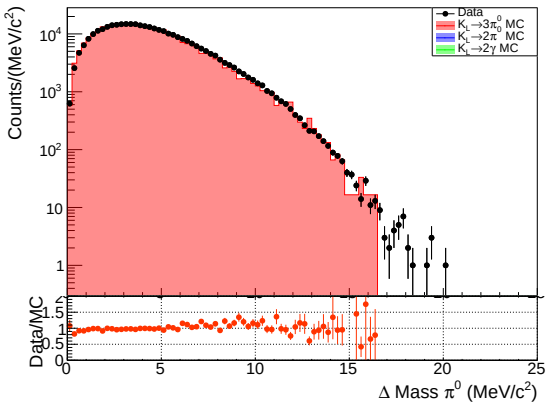
(b) K_L reconstructed z χ^2



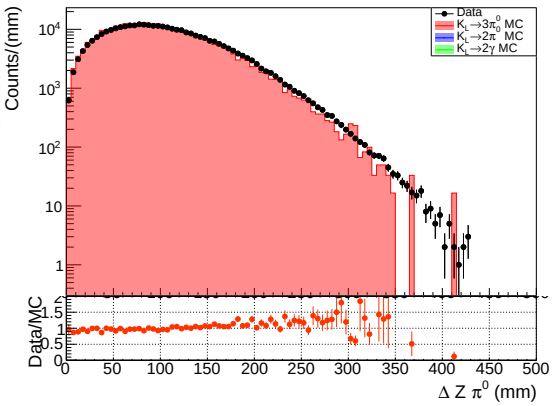
(c) Distance between two clusters



(d) Individual photon energy

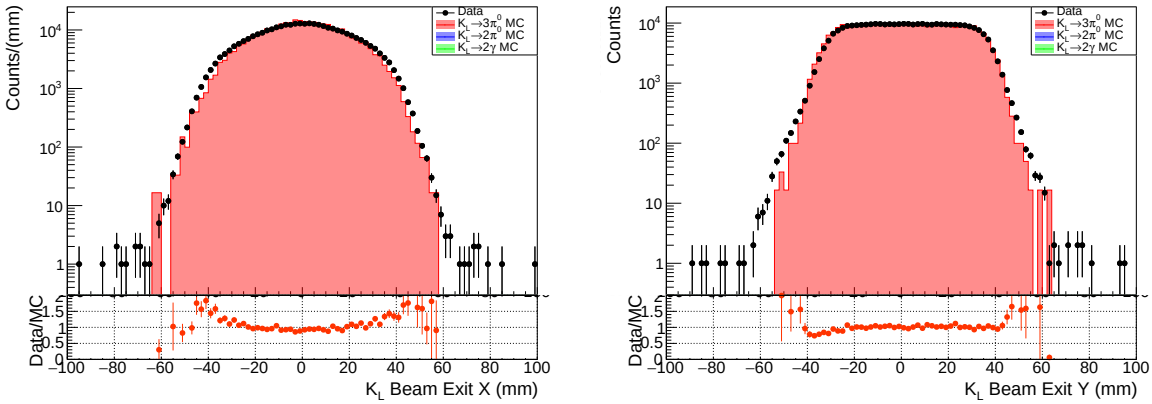


(e) Reconstructed π^0 mass difference



(f) Reconstructed π^0 z difference

Figure 7.2: Kinematic distributions for reconstructed $K_L^0 \to 3\pi^0$ decays. Black points represent data and the colored histograms represent MC ($K_L^0 \to 3\pi^0$ = red, $K_L^0 \to 2\pi^0$ = blue, $K_L^0 \to 2\gamma$ = green). The data to MC ratio is shown in the bottom panel and the error bars represent statistical error.



(a) K_L rec. x position at beam exit

(b) K_L rec. y position at beam exit

Figure 7.3: Kinematic distributions for reconstructed $K_L^0 \to 3\pi^0$ decays. Black points represent data and the colored histograms represent MC ($K_L^0 \to 3\pi^0$ = red, $K_L^0 \to 2\pi^0$ = blue, $K_L^0 \to 2\gamma$ = green). The data to MC ratio is shown in the bottom panel and the error bars represent statistical error.

Aside from the timing distribution (Figure 7.2a), there is reasonable agreement between the data and MC for the kinematic variables, which ensures that the simulation of the K_L beam's interaction with the detectors accurately describes what is seen in the data. The discrepancy in the vertex time difference comes from not fully understanding the timing response of the calorimeter and being able to model it effectively in the MC simulation. For this reason, a loose selection criteria of $\Delta T_{\text{vtx}} \leq 3$ ns was adopted to mitigate the effects of this difference. Finally, although a cut was not imposed on the x and y incident positions of the K_Ls at the beam exit, the distributions are shown in Figure 7.3. The difference in shape of the distributions comes from the geometry of the T1 target since the beam hits a larger spread in the x direction and effectively a point source in the y direction.[2]

7.3.2 $K_L^0 \to 2\pi^0$

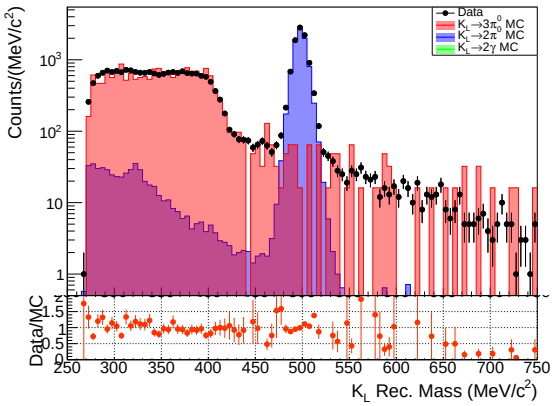
Though the $K_L^0 \to 2\pi^0$ decay has a branching ratio over 200 times smaller than the $K_L^0 \to 3\pi^0$ decay ($\sim 10^{-4}$), it plays an important role in the normalization analysis

[2] This can be seen in Figure 2.6.

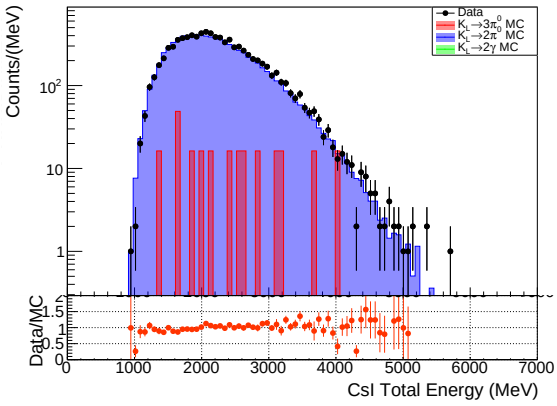
because it is the normalization mode most similar to the signal decay. Both $K_L^0 \to 2\pi^0$ and $K_L^0 \to \pi^0\nu\bar{\nu}$ have four final state particles and $K_L^0 \to 2\pi^0$ has a similar energy profile and momentum distribution to that of the signal decay. The largest background contribution for the $K_L^0 \to 2\pi^0$ mode is the $K_L^0 \to 3\pi^0$ decay if two of the six photons are missed and only four clusters are seen on the calorimeter or if cluster fusion occurs. Since the branching ratio of $K_L^0 \to 3\pi^0$ decay is two orders of magnitude higher than $K_L^0 \to 2\pi^0$ decay, the veto cuts were important for reducing background contamination. After applying all kinematic and veto cuts for the $K_L^0 \to 2\pi^0$ normalization event selection, the kinematic distributions are shown in Figures 7.4, 7.5, and 7.6.

The veto cuts were effective in removing $K_L^0 \to 3\pi^0$ events with missing photons, especially the cuts for the IB and MB detectors, and the $K_L^0 \to 3\pi^0$ events that remained were mostly due to the cluster fusion background. The $K_L^0 \to 3\pi^0$ contamination is best illustrated in the K_L reconstructed mass distribution (Figure 7.4a), but the K_L mass cut effectively removed this contribution and only the events in the final event selection region (482 MeV/c^2 $\leq M_{K_L} \leq$ 512 MeV/c^2) around the peak remained. Thus, the purity of the $K_L^0 \to 2\pi^0$ normalization sample was around 97%. The $K_L^0 \to 3\pi^0$ events that were left came from events that were reconstructed with four photons and had an invariant mass close to the K_L mass by happenstance.

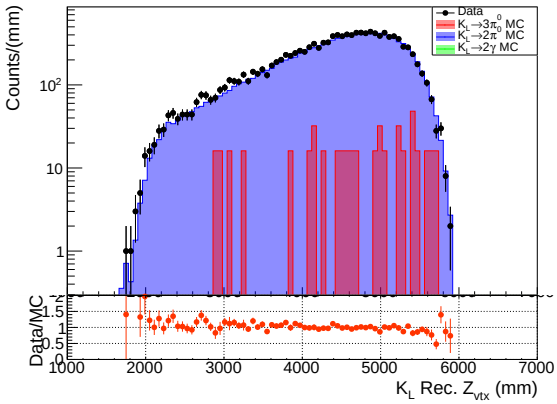
The other kinematic distributions are the same as those for $K_L^0 \to 3\pi^0$ and have similar trends, though the limited statistics in the $K_L^0 \to 2\pi^0$ data cause the distributions to have more statistical fluctuations than the $K_L^0 \to 3\pi^0$ distributions. The vertex time difference discrepancy between data and MC is also seen in the $K_L^0 \to 2\pi^0$ mode, but all other variables show good agreement between data and MC within the accepted cut regions. The distributions of the K_L x and y incident positions at the beam exit are shown in Figure 7.6 and have good agreement between data and MC within the accepted cut region of $\pm$50 mm.



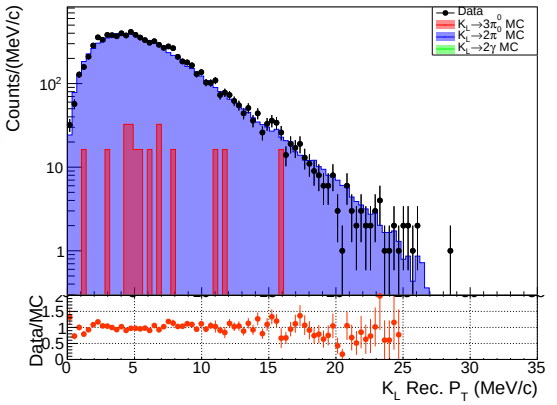
(a) K_L reconstructed mass



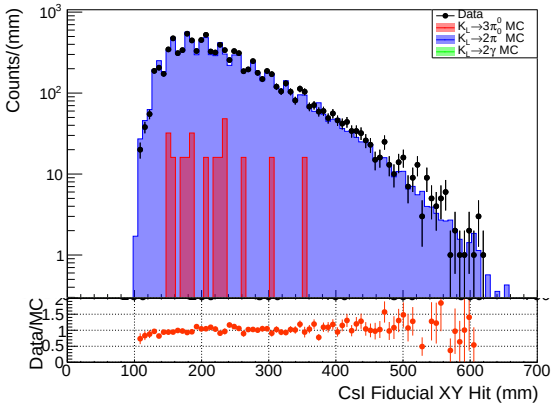
(b) Total energy deposited in CsI



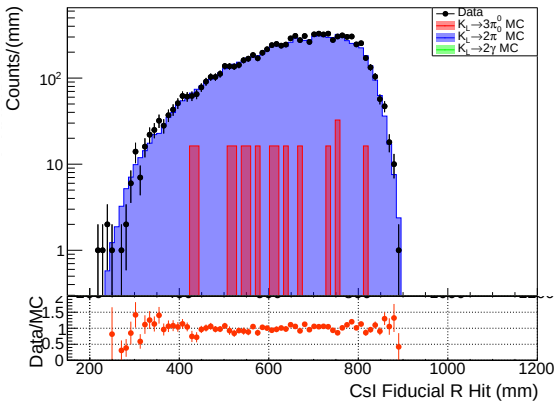
(c) K_L reconstructed z vertex



(d) K_L reconstructed P_T

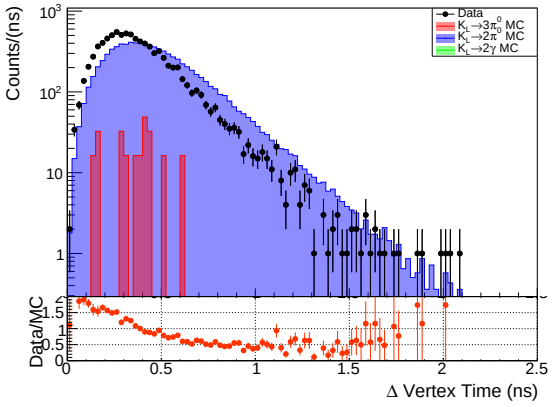


(e) Innermost photon cluster position (XY)

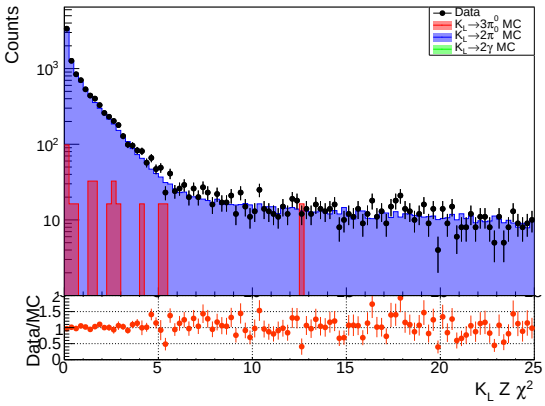


(f) Outermost photon cluster position (R)

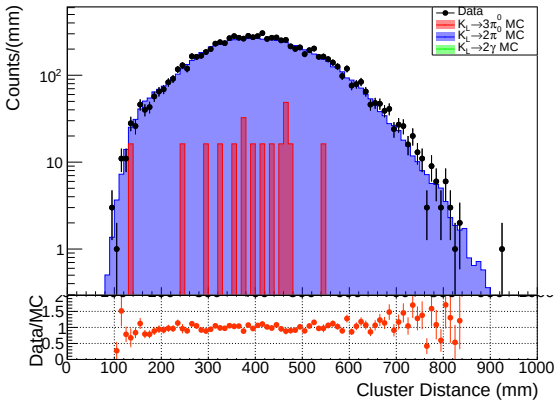
Figure 7.4: Kinematic distributions for reconstructed $K_L^0 \to 2\pi^0$ decays. Black points represent data and the colored histograms represent MC ($K_L^0 \to 3\pi^0$ = red, $K_L^0 \to 2\pi^0$ = blue, $K_L^0 \to 2\gamma$ = green). The data to MC ratio is shown in the bottom panel and the error bars represent statistical error.



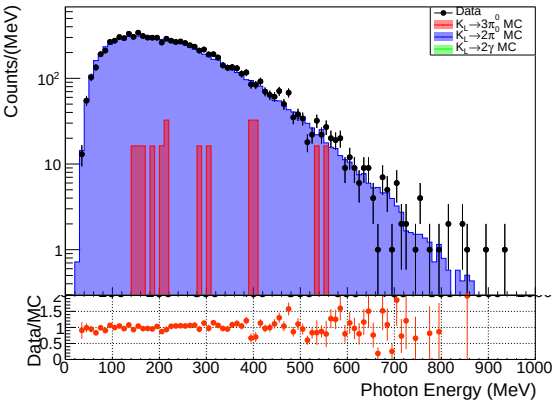
(a) Vertex time difference



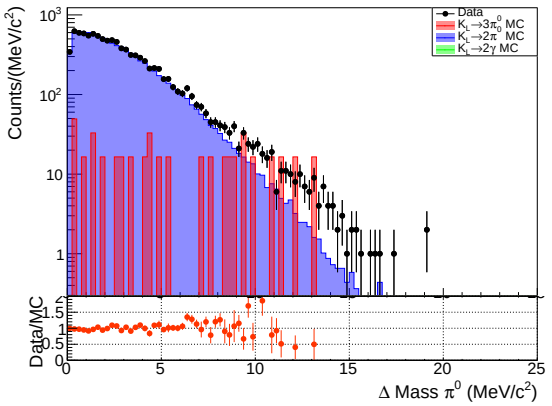
(b) K_L reconstructed z χ^2



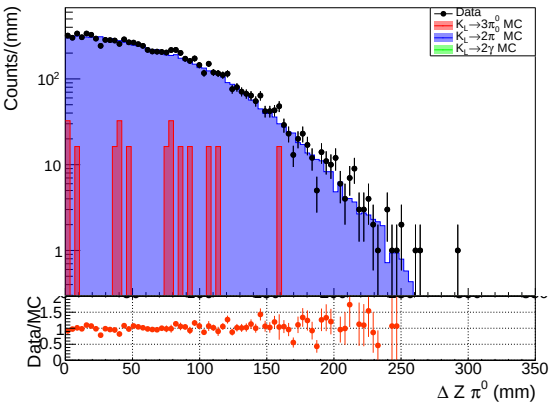
(c) Distance between two clusters



(d) Individual photon energy

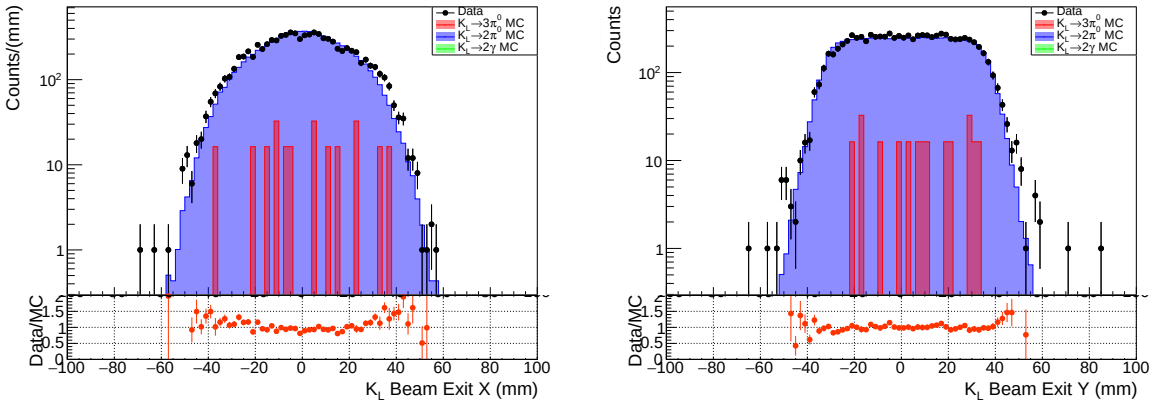


(e) Reconstructed π^0 mass difference



(f) Reconstructed π^0 z difference

Figure 7.5: Kinematic distributions for reconstructed $K_L^0 \to 2\pi^0$ decays. Black points represent data and the colored histograms represent MC ($K_L^0 \to 3\pi^0$ = red, $K_L^0 \to 2\pi^0$ = blue, $K_L^0 \to 2\gamma$ = green). The data to MC ratio is shown in the bottom panel and the error bars represent statistical error.



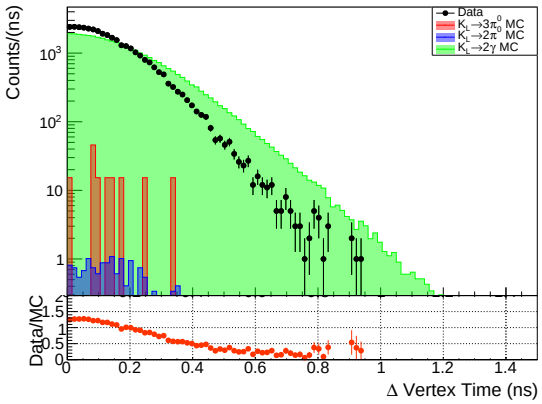
(a) K_L rec. x position at beam exit

(b) K_L rec. y position at beam exit

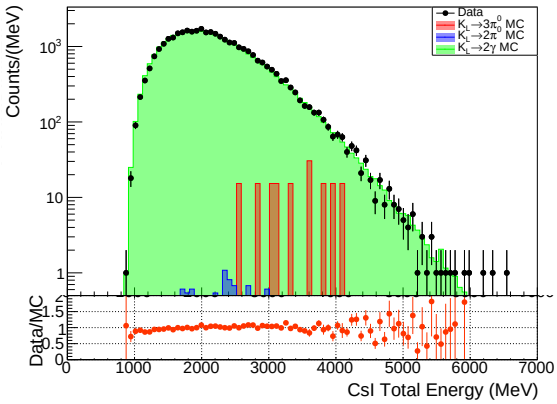
Figure 7.6: Kinematic distributions for reconstructed $K_L^0 \to 2\pi^0$ decays. Black points represent data and the colored histograms represent MC ($K_L^0 \to 3\pi^0$ = red, $K_L^0 \to 2\pi^0$ = blue, $K_L^0 \to 2\gamma$ = green). The data to MC ratio is shown in the bottom panel and the error bars represent statistical error.

7.3.3 $K_L^0 \to 2\gamma$

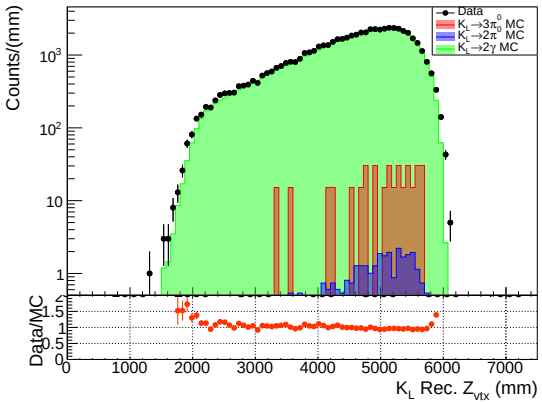
The final normalization mode used was the $K_L^0 \to 2\gamma$ decay which has an almost identical reconstruction procedure to the signal decay as both final states have two photons. Since the K_L mass was assumed in order to reconstruct this decay and there are no π^0s in the reconstruction, there are no kinematic distributions for these variables. Since there are only two photons in the final state, it is possible for the $K_L^0 \to 3\pi^0$ and $K_L^0 \to 2\pi^0$ modes to have background contributions if four or two photons are missed or the fusion of clusters occurs. Despite background contributions from these modes, the purity of the $K_L^0 \to 2\gamma$ normalization sample was still over 99% with the majority of the contamination coming from the $K_L^0 \to 3\pi^0$ mode because of its large branching ratio. After all of the kinematic and veto cuts were applied (except for the variable of interest), the kinematic distributions are shown in Figures 7.7 and 7.8. Overall, there is good agreement between the data and MC within the accepted cut regions, with the exception of the timing distribution as seen in Figure 7.7a which was also seen in the other modes.



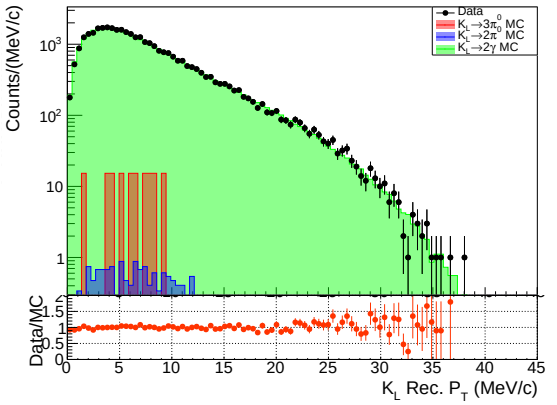
(a) Vertex time difference



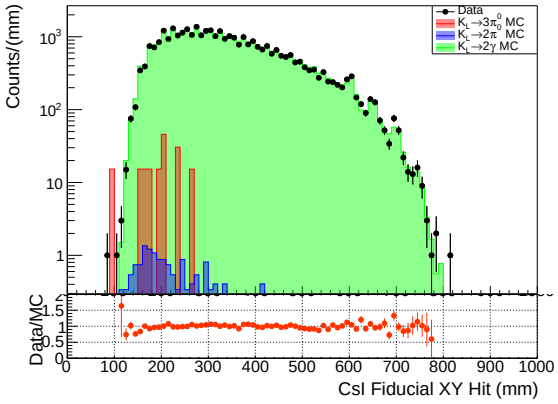
(b) Total energy deposited in CsI



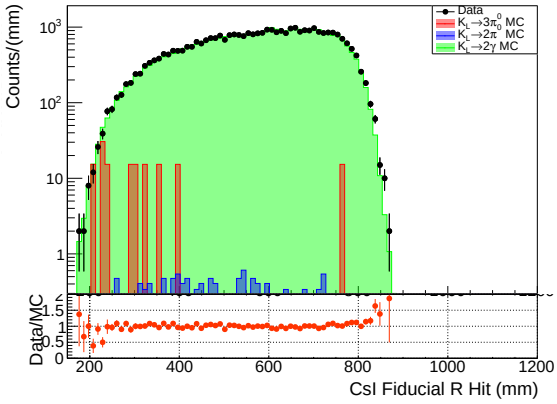
(c) K_L reconstructed z vertex



(d) K_L reconstructed P_T

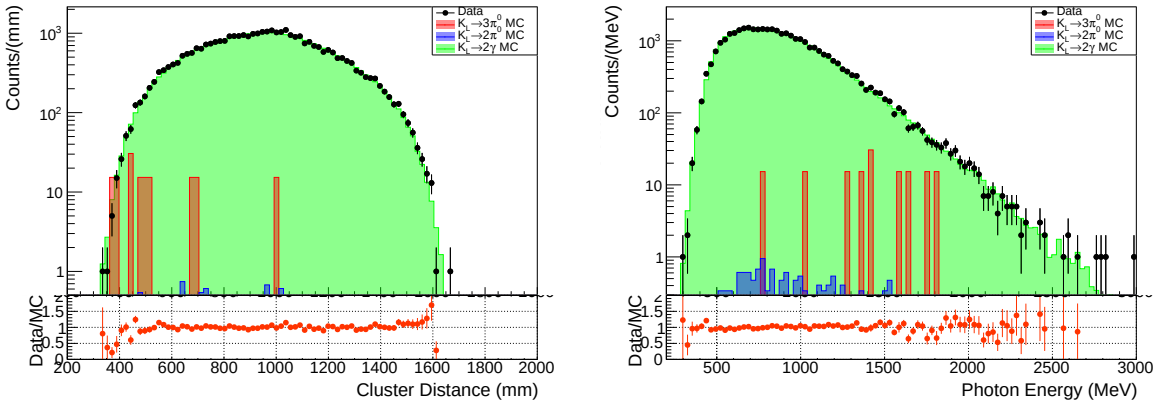


(e) Innermost photon cluster position (XY)



(f) Outermost photon cluster position (R)

Figure 7.7: Kinematic distributions for reconstructed $K_L^0 \to 2\gamma$ decays. Black points represent data and the colored histograms represent MC ($K_L^0 \to 3\pi^0$ = red, $K_L^0 \to 2\pi^0$ = blue, $K_L^0 \to 2\gamma$ = green). The data to MC ratio is shown in the bottom panel and the error bars represent statistical error.



(a) Distance between two clusters **(b)** Individual photon energy

Figure 7.8: Kinematic distributions for reconstructed $K_L^0 \to 2\gamma$ decays. Black points represent data and the colored histograms represent MC ($K_L^0 \to 3\pi^0$ = red, $K_L^0 \to 2\pi^0$ = blue, $K_L^0 \to 2\gamma$ = green). The data to MC ratio is shown in the bottom panel and the error bars represent statistical error.

7.4 Normalization Analysis Calculations and Results

In order to calculate the K_L flux, and subsequently the total number of kaons, several other quantities must first be calculated. As mentioned in Section 7.1, each step of the normalization analysis was performed separately for each run period. Thus, all the calculations described in this section were done for each period until being summed or averaged for the final result.

7.4.1 Acceptance

The same kinematic and veto cuts that were applied to the data for the normalization event selection were also applied to the MC simulation samples to check the MC reproducibility of the data, and also to calculate the acceptance, A. The acceptance is based on the MC and calculated separately for each mode as

$$A_{\text{mode}} = \frac{N_{\text{MC}}^{\text{accepted}}}{N_{\text{MC}}^{\text{generated}}} \tag{7.3}$$

where $N_{\mathrm{MC}}^{\mathrm{accepted}}$ and $N_{\mathrm{MC}}^{\mathrm{generated}}$ are the total number of MC events remaining after applying all cuts and the total number of MC events generated, respectively. Thus, the acceptance gives the probability that an event will remain after all cut selections have been applied. Since other decay modes can contribute as background in any normalization mode sample, the total acceptance, A_{total}, factored in the contributions from other modes and was defined as

$$A_{\mathrm{total}} = \sum_{\mathrm{mode}} A_{\mathrm{mode}}\ \mathrm{BR}_{\mathrm{mode}} \tag{7.4}$$

where $\mathrm{BR}_{\mathrm{mode}}$ is the branching ratio of the corresponding mode. The total acceptance for each run period and each normalization mode is shown in Figure 7.9. The acceptance varied from period to period, which was largely attributed to the change in beam power, and subsequently the loss from accidental activities in the beam. Thus, as the beam power increased so did the accidental loss, causing a lower acceptance. Since the acceptance is based purely on the MC simulation samples, agreement between the data and MC is important for an accurate calculation of the total number of K_Ls and the K_L flux.

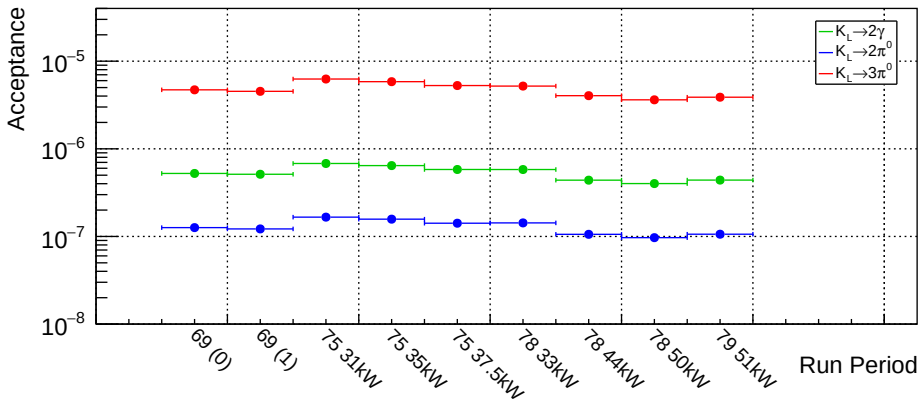


Figure 7.9: The total acceptance for each normalization mode over all run periods in 2016–2018. Vertical error bars represent statistical error, but the error is negligible.

7.4.2 Cut Efficiencies

The normalization studies were also used to compare the kinematic and veto cut efficiencies between data and MC. The efficiency (ϵ) of a cut, i, is defined in Equation 7.5 below.

$$\epsilon_i = \frac{\text{number of events w/ all cuts}}{\text{number of events w/ all cuts except } i_{th} \text{ cut}} \tag{7.5}$$

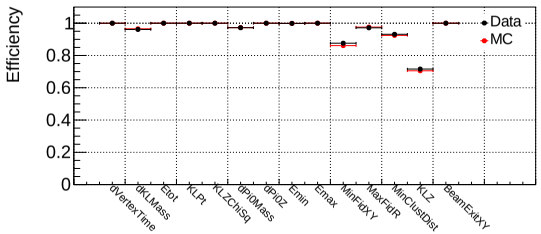
The efficiency of a particular cut reflects the cut's selection power– if the cut doesn't remove any events, then the efficiency is 1. The important part of evaluating the cut efficiencies was to compare the difference in efficiency between the data and MC, which quantitatively describes the systematic differences between data and MC detector responses, calibration constants, and energy corrections. These differences in the efficiencies for each cut were used in the estimation of systematic uncertainties for the $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis, as described in Section 9.1. The kinematic and veto cut efficiencies for each normalization mode for Run 79 51 kW are shown in Figure 7.10, as this run period had the largest statistics.

7.4.3 K_L Yield and Flux

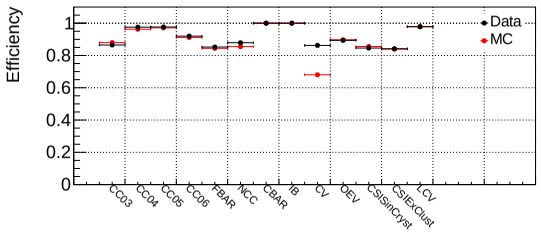
To calculate the flux of K_Ls into the KOTO detectors, the K_L yield is first defined in Equation 7.6,

$$K_L \text{ yield} = \frac{N_{\text{data}}}{A_{\text{total}}} \tag{7.6}$$

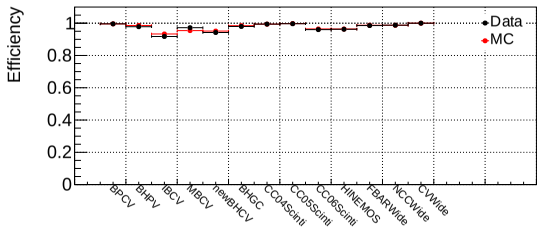
where N_{data} is the number of data events remaining after applying all cuts and A_{total} is the total acceptance defined in Equation 7.4. Then, to make this number comparable between different run periods, the K_L flux factors in the amount of data collected



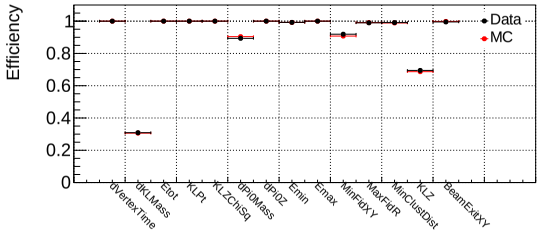
(a) $K_L^0 \to 3\pi^0$ kinematic cut efficiencies



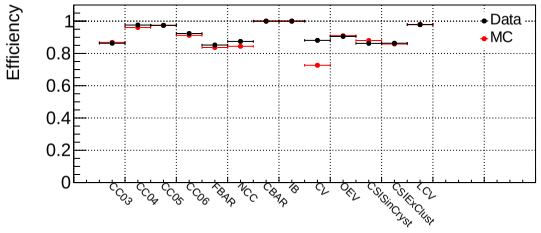
(b) $K_L^0 \to 3\pi^0$ veto cut efficiencies



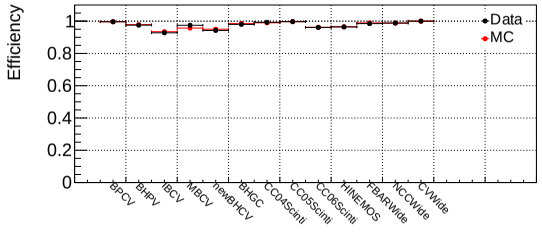
(c) $K_L^0 \to 3\pi^0$ veto cut efficiencies



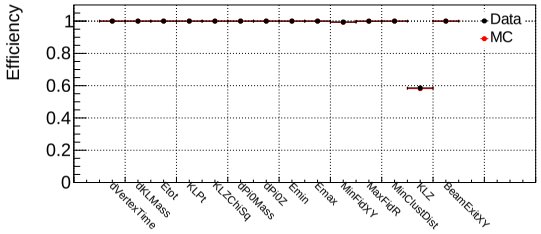
(d) $K_L^0 \to 2\pi^0$ kinematic cut efficiencies



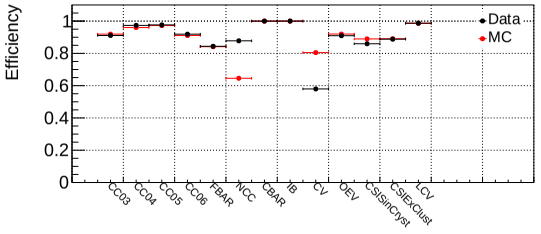
(e) $K_L^0 \to 2\pi^0$ veto cut efficiencies



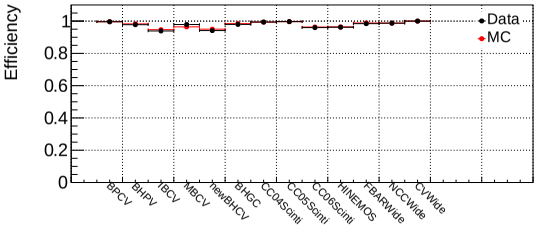
(f) $K_L^0 \to 2\pi^0$ veto cut efficiencies



(g) $K_L^0 \to 2\gamma$ kinematic cut efficiencies



(h) $K_L^0 \to 2\gamma$ veto cut efficiencies



(i) $K_L^0 \to 2\gamma$ veto cut efficiencies

Figure 7.10: The kinematic and veto cut efficiencies for each normalization mode in Run 79 51 kW. The cut efficiencies of the data and MC are shown in black and red, respectively. The kinematic cut efficiencies were evaluated using the normalization trigger data and the veto cut efficiencies were evaluated using the minimum bias trigger data. Vertical error bars represent statistical error, but the error is negligibly small.

each run period, by normalizing to the effective POT collected. Thus, the K_L flux is defined as

$$K_L \text{ flux} = \frac{K_L \text{ yield}}{\text{POT}_{\text{runs}}/\text{POT}_{\text{norm factor}}} \tag{7.7}$$

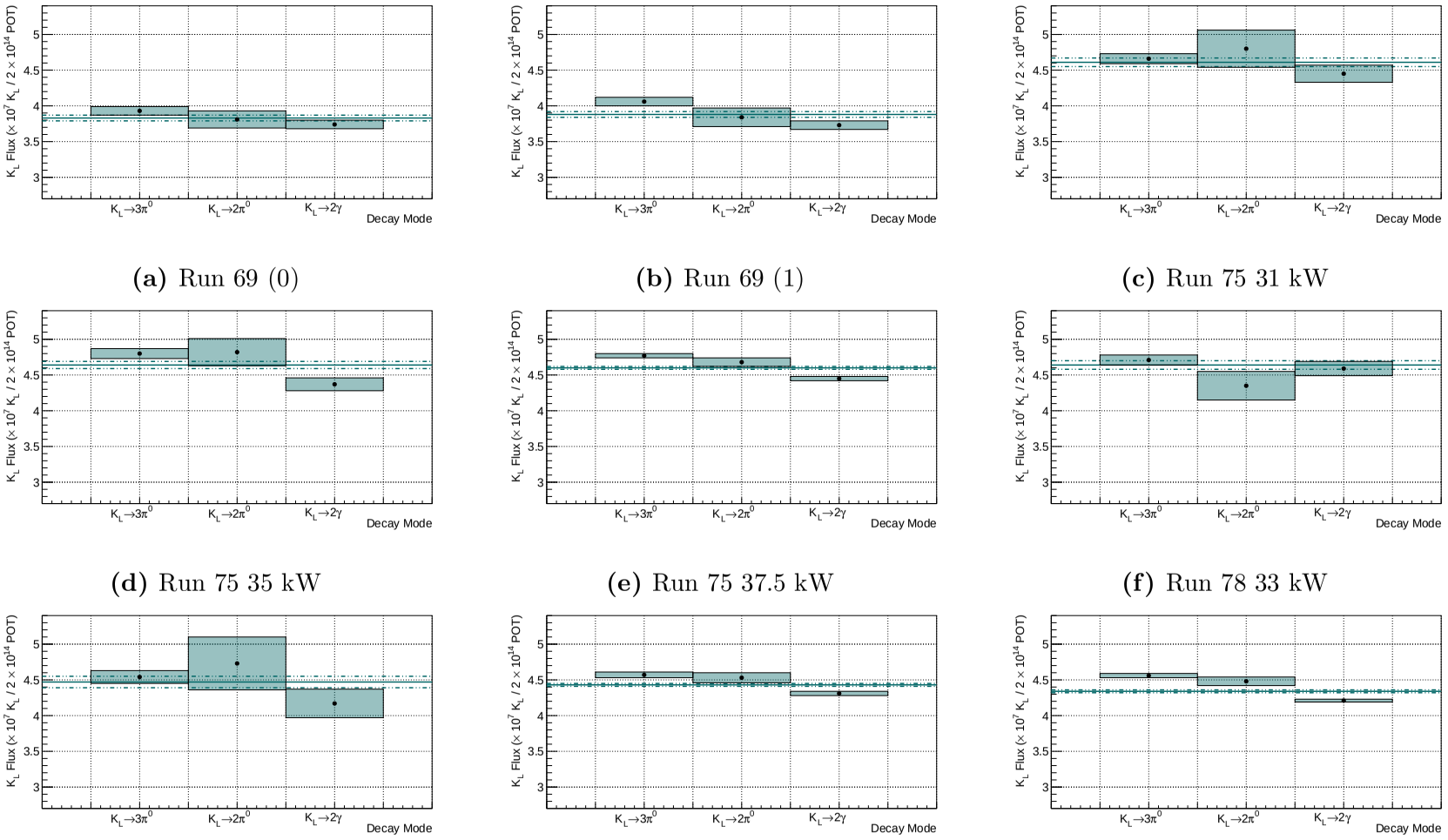
where POT_{runs} is the corrected physics POT divided by the prescale factor, as shown in Equation 7.8.

$$\text{POT}_{\text{runs}} = \frac{\text{POT}}{\text{prescale}} \tag{7.8}$$

The normalization trigger prescale was 30 and the minimum bias trigger prescale was 300 as described in Section 4.3. The K_L flux is defined in units of the POT normalization factor ($\text{POT}_{\text{norm factor}}$), which is the J-PARC design value for the number of POT per spill and is 2×10^{14}. The K_L flux for each normalization mode and each run period in 2016–2018 is listed in Table 7.3 and shown in Figure 7.11. The normalization trigger data was used in this analysis to calculate the K_L flux.

Run Period	K_L **Flux** ($\times$ **10**^{7} K_L**/2** $\times$ **10**^{14} **POT**)		
	$K_L^0 \to 3\pi^0$	$K_L^0 \to 2\pi^0$	$K_L^0 \to 2\gamma$
69 (0)	3.93 ± 0.06	3.81 ± 0.12	3.74 ± 0.06
69 (1)	4.06 ± 0.06	3.84 ± 0.13	3.73 ± 0.06
75 31 kW	4.66 ± 0.07	4.80 ± 0.26	4.45 ± 0.12
75 35 kW	4.80 ± 0.07	4.82 ± 0.19	4.37 ± 0.09
75 37.5 kW	4.77 ± 0.03	4.68 ± 0.06	4.46 ± 0.03
78 33 kW	4.71 ± 0.07	4.35 ± 0.20	4.59 ± 0.10
78 44 kW	4.54 ± 0.09	4.73 ± 0.37	4.17 ± 0.17
78 50 kW	4.57 ± 0.04	4.53 ± 0.07	4.31 ± 0.03
79 51 kW	4.56 ± 0.03	4.48 ± 0.06	4.21 ± 0.02

Table 7.3: The K_L flux for each normalization mode in each run period for the 2016–2018 data set. The error included is the statistical error.



(a) Run 69 (0) **(b)** Run 69 (1) **(c)** Run 75 31 kW

(d) Run 75 35 kW **(e)** Run 75 37.5 kW **(f)** Run 78 33 kW

(g) Run 78 44 kW **(h)** Run 78 50 kW **(i)** Run 79 51 kW

Figure 7.11: The K_L flux for each normalization mode for all run periods in 2016–2018. The black points represent the calculated flux and the colored boxes represent the statistical error. The solid line is the weighted average of the flux for the three modes with the dashed lines representing the statistical error on the weighted average.

Run Period	Number of K_L at Beam Exit		
	$K_L^0 \to 3\pi^0$	$K_L^0 \to 2\pi^0$	$K_L^0 \to 2\gamma$
69 (0)	$(3.50 \pm 0.05)\times10^{11}$	$(3.39 \pm 0.11)\times10^{11}$	$(3.33 \pm 0.05)\times10^{11}$
69 (1)	$(2.71 \pm 0.04)\times10^{11}$	$(2.56 \pm 0.08)\times10^{11}$	$(2.49 \pm 0.04)\times10^{11}$
75 31 kW	$(6.29 \pm 0.10)\times10^{10}$	$(6.48 \pm 0.35)\times10^{10}$	$(6.01 \pm 0.16)\times10^{10}$
75 35 kW	$(1.40 \pm 0.02)\times10^{11}$	$(1.40 \pm 0.06)\times10^{11}$	$(1.27 \pm 0.02)\times10^{11}$
75 37.5 kW	$(1.74 \pm 0.01)\times10^{12}$	$(1.71 \pm 0.02)\times10^{12}$	$(1.63 \pm 0.01)\times10^{12}$
78 33 kW	$(1.20 \pm 0.02)\times10^{11}$	$(1.11 \pm 0.05)\times10^{11}$	$(1.17 \pm 0.03)\times10^{11}$
78 44 kW	$(4.68 \pm 0.09)\times10^{10}$	$(4.88 \pm 0.38)\times10^{10}$	$(4.30 \pm 0.17)\times10^{10}$
78 50 kW	$(1.68 \pm 0.01)\times10^{12}$	$(1.67 \pm 0.02)\times10^{12}$	$(1.58 \pm 0.01)\times10^{12}$
79 51 kW	$(2.54 \pm 0.02)\times10^{12}$	$(2.50 \pm 0.03)\times10^{12}$	$(2.35 \pm 0.01)\times10^{12}$
Total	$\mathbf{(6.96 \pm 0.03)}\times10^{12}$	$\mathbf{(6.83 \pm 0.05)}\times10^{12}$	$\mathbf{(6.49 \pm 0.02)}\times10^{12}$

Table 7.4: The number of kaons at the beam exit calculated for each normalization mode in each run period for the 2016–2018 data set. The error included is the statistical error.

7.4.4 Total Number of Kaons

The final step in the normalization analysis was to calculate the total number of kaons at the beam exit. This number was obtained simply by multiplying the K_L yield by the prescale factor, as shown in Equation 7.9.

$$N_{K_L^0} = K_L \text{ yield} \times \text{prescale} \tag{7.9}$$

The number of kaons calculated for each normalization mode and each run period is shown in Table 7.4 and the totals for each mode are illustrated in Figure 7.12. It should be noted that among the three normalization modes, there is a discrepancy in the flux and $N_{K_L^0}$, in which the $K_L^0 \to 3\pi^0$ mode result is usually the largest and the $K_L^0 \to 2\gamma$ mode is typically the smallest, with the $K_L^0 \to 2\pi^0$ result residing between them (this can be seen in Figures 7.11 and 7.12). Ideally, since the branching ratios of the three modes have been taken into account, the K_L flux values and $N_{K_L^0}$ obtained should be consistent within statistical error in the same run period. However, there is about a 5–8% difference in the flux and $N_{K_L^0}$ across the three normalization modes, the exact source of which is presently unknown. This discrepancy between

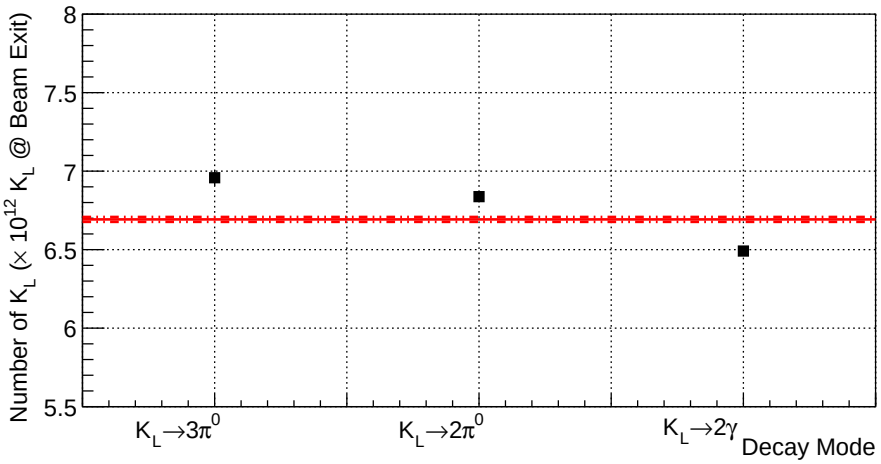


Figure 7.12: The total number of kaons at the beam exit for each normalization mode for the 2016–2018 data set. The solid red line is the weighted average of the total $N_{K_L^0}$ for the three modes with the dashed lines representing the statistical error on the weighted average. Vertical error bars on the black data points represent statistical error, but the error is negligibly small.

the normalization modes is accounted for as a source of systematic uncertainty in the evaluation of the Single Event Sensitivity, as described in Section 9.1.8.

After evaluating the consistency between the three normalization modes, the final value for the total number of kaons collected in the 2016–2018 data set was taken from the $K_L^0 \to 2\pi^0$ mode, as it has an energy profile and momentum distribution closest to the signal decay. Therefore, the total $N_{K_L^0}$ at the beam exit was $(6.83 \pm 0.05) \times 10^{12}$ for the 2016–2018 data set. The $K_L^0 \to 2\pi^0$ mode weighted average across all run periods for the K_L flux was $(4.47 \pm 0.03) \times 10^7$ $K_L/2 \times 10^{14}$ POT. This value of the total $N_{K_L^0}$ was used to calculate the final SES as reported in Chapter 9. The data is reasonably well reproduced by the MC simulations as seen in the kinematic distributions and cut efficiencies and the discrepancies are included in the evaluation of systematic uncertainties.

Chapter 8

Background Estimation

Because a blind analysis strategy was used, we estimated the expected number of background events in the blinded and signal regions using MC simulations or other reliable samples and compared these background estimations to data in the surrounding control regions. The background estimations were used to inform background reduction methods and optimize the final cut set such that the signal to background ratio was maximized. This chapter describes the background sources, their estimation methods, and the final estimated background contribution for each source.

8.1 Overview

The number of background events was estimated separately for each background source and for different regions in the P_T–Z plane. As described in Section 5.1.1, the different background sources are categorized into two groups, the K_L decay background and the neutron-induced background. The majority of the background estimations were done using MC simulations except for the hadron-cluster background in which we used a control sample of data collected using the Al target (described in Section 4.4.1). The details of the MC generation and data samples used are described in each background subsection. Background can also be caused by accidental activity in the beam where an accidental hit masks the waveform of a true hit and

Background Source	Number of Events Used
K_L Decay Background	
$K_L^0 \to \pi^+\pi^-\pi^0$	8.04×10^{13}
$K_L^0 \to 2\pi^0$	8.04×10^{10}
$K_L^0 \to 2\gamma$	7.395×10^{11}
$K_L^0 \to 3\pi^0$ (masking)	8.152×10^{13}
$K_L^0 \to \pi^\pm e^\mp \nu_e$ (masking)	8.156×10^{13}
Neutron-Induced Background	
Hadron-Cluster	9.09×10^{16}
Upstream π^0	4.9895×10^{11}
CV-π^0	7.1982×10^{12}
CV-η	7.1982×10^{13}

Table 8.1: The background sources considered in the $K_L^0 \to \pi^0\nu\bar{\nu}$ blind analysis and the number of events used for each background estimation.

the timing is incorrectly calculated. This "masking background" was also considered in the background contributions as described in Section 8.2.

The event selections for the $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis were applied to the background samples to evaluate the estimated background contribution for each source. Then, the number of remaining events in the simulation[1] were scaled to the amount of data collected and correction factors were applied to estimate the background contribution. Table 8.1 lists the number of events used to estimate each background source.

8.2 Masking Background

The masking background refers to when an accidental hit in the beam causes an overlapped pulse on a hit that should have been vetoed. This effect can be seen in Figure 8.1 where the measured time is shifted outside the veto window due to the overlapped pulse and the event is not vetoed and becomes a background. The FTT method described in Section 5.3.4.2 was used to discriminate overlapped pulses and reduce the masking background.

To study the effects of the masking background and estimate its contribution

[1] Or the Z0 Al target control sample

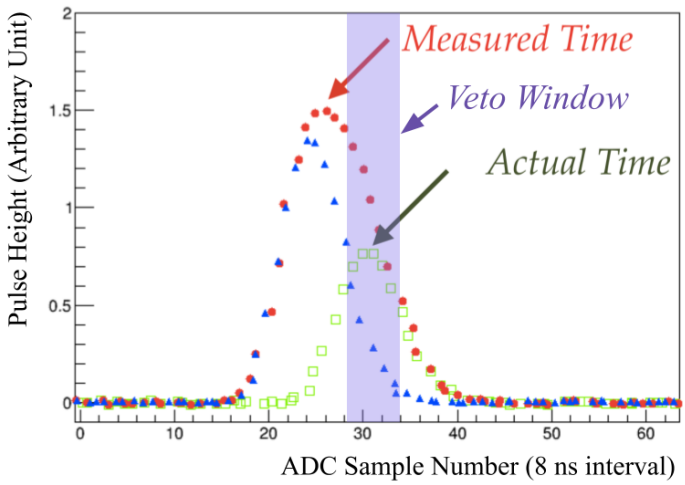


Figure 8.1: Example of the masking background in which an accidental hit (red) overlaps a true hit (green) and causes an incorrect calculation of the event time. Because the measured event time is outside of the veto window, this event is not vetoed and thus contributes to background.

we overlaid waveforms from the TMon trigger data[2] on the waveforms in the MC simulation. This process is also described in the "accidental overlay" Section 6.7. The accidental overlay was applied to all of the MC simulations for the background sources in order to more closely replicate the data. In particular, the masking background contributes in the $K_L^0 \to 3\pi^0$ and the $K_L^0 \to \pi^\pm e^\mp \nu_e$ backgrounds as described in Sections 8.3.4 and 8.3.5, respectively.

8.3 K_L Decay Background

The K_L background comes from other K_L decay modes as described in Section 5.1.1.1 and results primarily from veto detector inefficiencies. We used MC simulations to estimate the background contributions from $K_L^0 \to \pi^+\pi^-\pi^0$, $K_L^0 \to 2\pi^0$, $K_L^0 \to 3\pi^0$, $K_L^0 \to 2\gamma$, and $K_L^0 \to \pi^\pm e^\mp \nu_e$. For each K_L decay background, we scaled the remaining events in the simulation to the amount of data collected by multiplying by the normalization factor defined in Equation 8.1.

[2]Described in Section 4.3.5.5

$$\text{BG Norm Factor} = \frac{\text{Total \# Kaons} \times \text{BR(mode)}}{\text{\# MC Generated}} \tag{8.1}$$

where the total number of kaons at the beam exit was determined to be 6.83×10^{12} K_Ls as described in Section 7.4.4, BR(mode) is the branching ratio of the particular K_L decay, and the number of MC generated is the total number of MC events used for the background estimation. Any necessary correction factors were also applied as described in each background section.

8.3.1 $K_L^0 \to \pi^+\pi^-\pi^0$ Background

Mechanism

The $K_L^0 \to \pi^+\pi^-\pi^0$ decay becomes a background if the π^+ and π^- escape down the beam hole and are undetected by the downstream beam hole veto detectors and both the photons from the π^0 decay hit the calorimeter as shown in Figure 8.2. Since the branching ratio of this decay mode is large (∼12%) this background can contribute in a major way and so the CC05, CC06, BPCV, and new BHCV were important detectors for vetoing this background. The missing momentum carried away by the π^+ and π^- can cause the reconstructed π^0 from this decay to be pushed

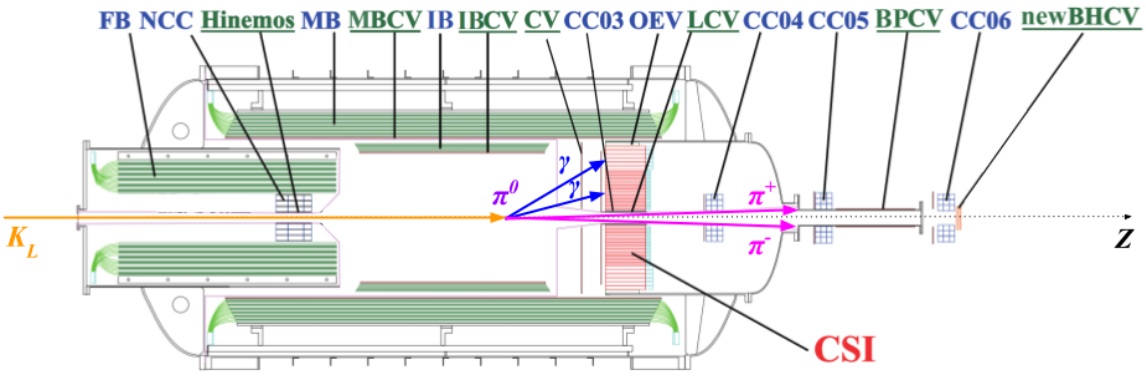


Figure 8.2: Example of the $K_L^0 \to \pi^+\pi^-\pi^0$ background mechanism.

into the bottom of the signal region. Although the maximum amount of transverse momentum the π^0 can have is 133 MeV/c, the reconstructed P_{T} can be larger due to the finite energy and position resolution of the calorimeter or if the K_L decays off the beam axis. The closer the K_L decays to the calorimeter, the higher the chance that the π^+ and π^- will escape down the beam hole due to the finite beam size. Thus, we cut the bottom right corner of the signal region to mitigate this background (this can be seen in Figure 8.3). The $K_L \to \pi^+\pi^-\pi^0$ Deep Learning cut was developed specifically to reduce this background.

Estimation Method

To estimate the contribution from the $K_L \to \pi^+\pi^-\pi^0$ background, the recycling method described in Section 6.1.1 was used to generate a MC sample of background events. First, the K_Ls were simulated to decay to $\pi^+\pi^-\pi^0$ and information from the daughter particles (π^+, π^-, and 2γ from the π^0) were obtained. Next, we selected events from the simulation in which both charged pions did not hit the calorimeter and had two photon hits on the CsI in the fiducial region of $|x|$ or $|y| > 130$ mm and $R < 900$ mm with a combined energy greater than 550 MeV. These events were then duplicated 100 times to increase the effectiveness of the background sample and input as seeds into the full simulation. The MC was scaled to the amount of data and correction factors for the loss introduced by using different event selections in the recycling method were applied. These included a correction factor for the CsI accidentals, a correction factor to account for backsplash and shower leakage, and a correction factor to account for cuts used in the seed selection but not the final cut set.

Result

The background distribution for $K_L \to \pi^+\pi^-\pi^0$ in the $\mathrm{P_T}$–Z plane can be seen in Figure 8.3 with (right) and without (left) the $K_L \to \pi^+\pi^-\pi^0$ Deep Learning cut.

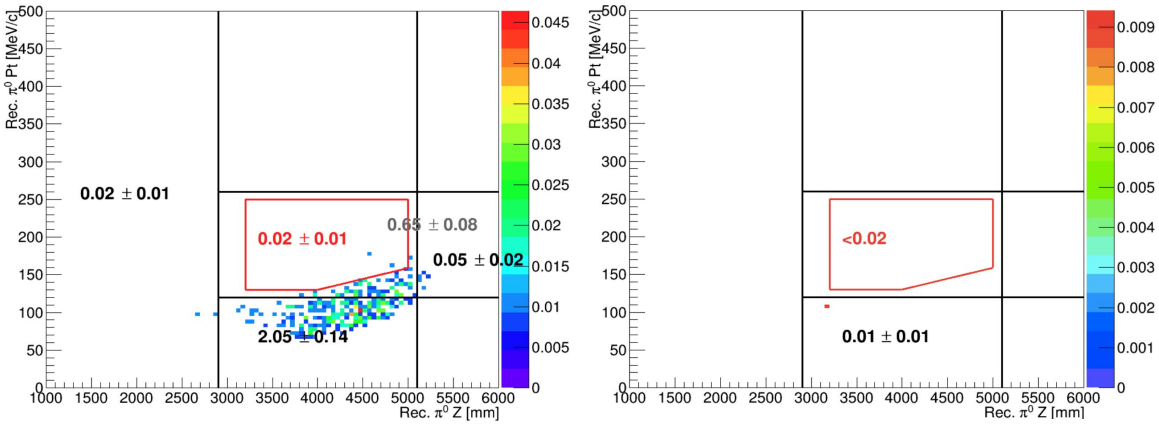


Figure 8.3: The estimated $K_L \to \pi^+\pi^-\pi^0$ background contribution in each region of the P_T–Z plane without (left) and with (right) the $K_L \to \pi^+\pi^-\pi^0$ DL cut at 90% signal acceptance. The values in black are the estimated number of events in the corresponding region, the values in grey are the estimated contribution inside the blinded region (including the signal region), and the values in red are the estimated contribution in the signal region. Regions without values have an estimated contribution of zero.

The background contribution from this mode was mainly concentrated in the low P_T region and the $K_L \to \pi^+\pi^-\pi^0$ DL cut effectively reduced this background in the region below the signal box while keeping the signal acceptance at 90%. Using the $K_L \to \pi^+\pi^-\pi^0$ DL cut, the background contribution in the signal region was estimated to be < 0.02 events at the 90% CL as shown in the right of Figure 8.3.

8.3.2 $K_L^0 \to 2\pi^0$ Background

Mechanism

The $K_L^0 \to 2\pi^0$ background can contribute in several ways, and because it only has two photons available for vetoing, detector inefficiencies can result in a large contribution from this background. The predominant mechanisms for the $K_L^0 \to 2\pi^0$ background are listed below.

- Photon Punch-Through

  If a photon is very high energy and there is not enough detector material along

its path, the photon can penetrate through the detector without any interaction. This "photon punch-through" occurred in the MB and IB detectors, though the installation of the IB greatly reduced the probability of punch-through and reduced the $K_L^0 \to 2\pi^0$ background by a factor of three. This inefficiency also occurred in the BHPV detector.

- Sampling Effect

  In the case of a low energy photon, the photon can be absorbed in the inactive layer of material in detectors with sandwich structure, such as the FB, IB, or MB. Thus, the photon is not detected and this detector inefficiency contributes to the background.

- Photo-Nuclear Interaction

  Photo-nuclear interaction happens when a high energy photon interacts with atomic nuclei in detector materials and the energy is absorbed by the nuclei. The absorbed energy excites the nuclei and the energy is emitted as hadrons such as neutrons or protons which are not detected. This effect contributes to detector inefficiencies and as a result, the $K_L^0 \to 2\pi^0$ background.

- Cluster Fusion

  In addition to missing photons, two photons can hit in the same spot on the calorimeter and fuse into a single cluster. This cluster fusion can contribute if three photons hit the calorimeter and two out of the three photons fuse into a single cluster while only one photon remains for vetoing. It can also contribute if four photons fuse into two clusters on the calorimeter which is misidentified as a two photon event.

- Odd Pairing

  The topology of the detected photons can affect the kinematic properties in the reconstruction of the π^0 and if the reconstruction is done incorrectly, can

contribute to background. "Even pairing" refers to the case when two photon clusters from the same π^0 are reconstructed, and "odd pairing" refers to the case when two photon clusters from different π^0s are paired together and reconstructed. If odd pairing happens, the π^0 is not reconstructed properly and can result in a background. The E_γ ratio cut and the E_γ θ_γ cut mentioned in Section 5.3.3.2 were designed to suppress the $K_L^0 \to 2\pi^0$ odd pairing background.

Estimation Method

To estimate the $K_L^0 \to 2\pi^0$ background, the fast simulation method was used. First $K_L^0 \to 2\pi^0$ decays were fully simulated and events were selected in which at least one of the photons hit the calorimeter in the fiducial region $|x|$ or $|y| > 100$ mm and $R < 900$ mm with an energy greater than 400 MeV. Then, these events were used as inputs to the full simulation to increase the $K_L^0 \to 2\pi^0$ background statistics. The MC was scaled to the amount of data collected and all of the cuts were applied.

Result

Figure 8.4 shows the $K_L^0 \to 2\pi^0$ background distribution in the P_T–Z plane for a loose cut condition[3] (left) and with the final cuts (right). While this background was mainly distributed in the lower P_T region of the signal box as shown on the left, the BHPV, BHGC, and IB detectors were effective in reducing the background in this region as well as the CsI Isolated Hit Crystal veto cut (described in Section 5.3.4.1) and the FTT veto method used to eliminate pulse pileups. With all of the final cuts, the estimated background contribution for the $K_L^0 \to 2\pi^0$ background in the signal region was < 0.18 events at the 90% CL.

[3]Including all kinematic cuts, all shape cuts to suppress the hadron-cluster background, and all veto cuts except for the CsI single crystal veto, BHPV veto, BHGC veto, and the wide veto window for CBAR and IB.

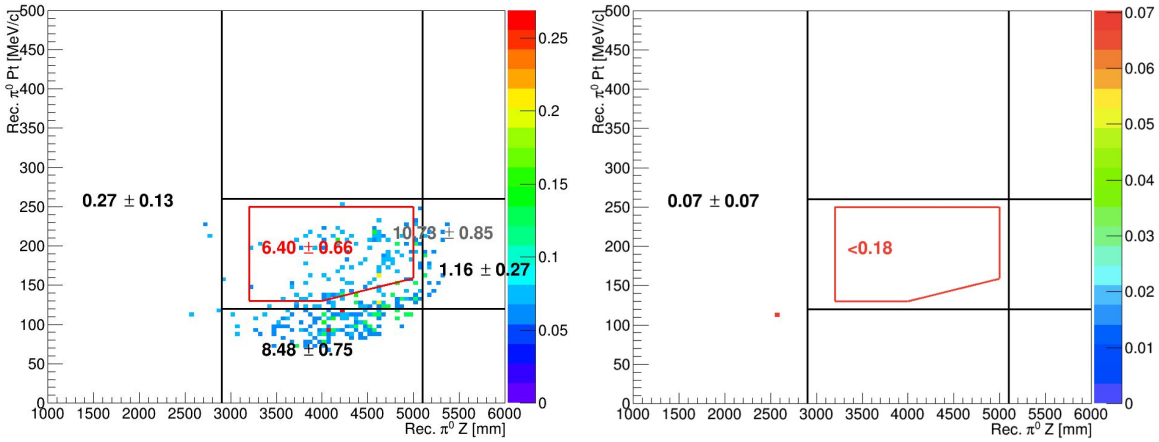


Figure 8.4: The estimated contribution in the P_T–Z plane for the $K_L^0 \to 2\pi^0$ background with a loose cut condition (left) and all of the cuts (right). The color of the number values follow the same convention as in Figure 8.3.

8.3.3 $K_L^0 \to 2\gamma$ Background

Mechanism

Though the $K_L^0 \to 2\gamma$ decay itself has no transverse momentum because the photons decay back to back, if the K_L itself has any P_T, the reconstructed π^0 from the two photons can become a background. For example, if the K_L scatters at the vacuum window, the off-axis $K_L^0 \to 2\gamma$ decay would produce two photons that appear to have a large P_T as illustrated in Figure 8.5. In addition, if the K_L decays before the decay volume and at least one of the photons hits an upstream detector, it can

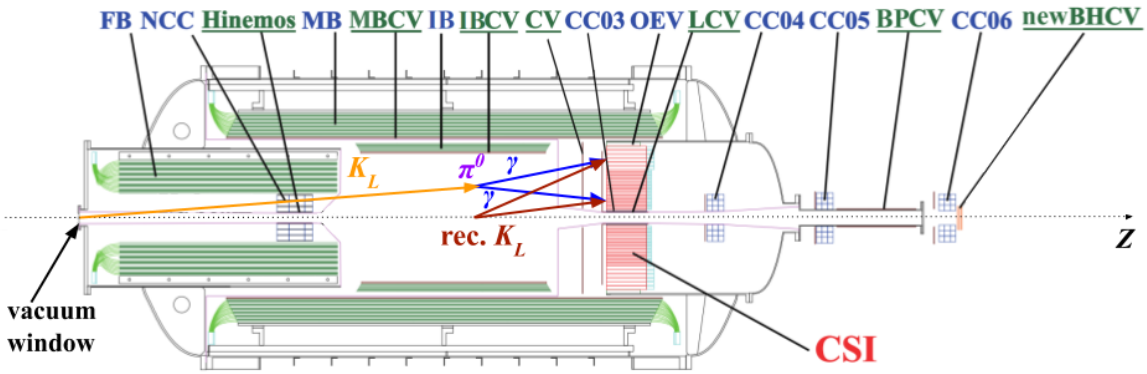


Figure 8.5: Example of the scattered $K_L^0 \to 2\gamma$ background mechanism.

produce particles that cause two cluster hits on the calorimeter, which would also cause a background.

Estimation Method

The fast simulation method was used to estimate the $K_L^0 \to 2\gamma$ background contribution. First, $K_L^0 \to 2\gamma$ decays were fully simulated at the beam exit and the simulation was stopped when the K_Ls reached the vacuum window. Then, only K_Ls with a transverse momentum greater than 20 MeV/c were selected and used as inputs to the full simulation so the background sample contained as many scattered K_Ls as possible. The MC was scaled to the data and the final cuts were applied.

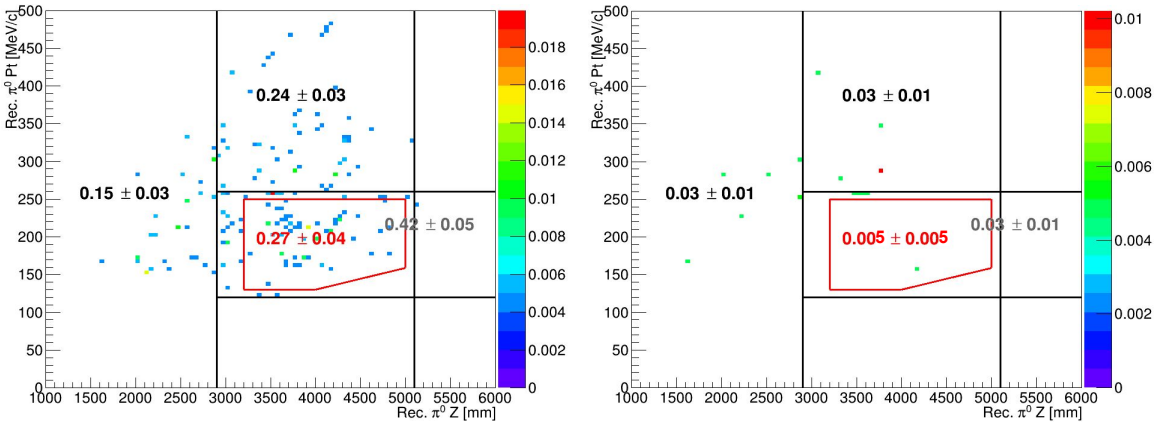


Figure 8.6: The estimated contribution in the P_T–Z plane for the $K_L^0 \to 2\gamma$ background without (left) and with (right) the advanced shape cuts. The color of the number values follow the same convention as in Figure 8.3.

Result

Figure 8.6 shows the $K_L^0 \to 2\gamma$ background distribution without (left) and with (right) the advanced shape cuts[4]. Though the shape cuts do not have much effect in removing scattered $K_L^0 \to 2\gamma$ background events, they are effective in removing other background contributions from the $K_L^0 \to 2\gamma$ decays as illustrated in the figure. The

[4]CSDDL, η CSD, and χ^2_θ

final background contribution in the signal region for $K_L^0 \to 2\gamma$ was estimated to be 0.005 ± 0.005 events.

8.3.4 $K_L^0 \to 3\pi^0$ Masking Background

Mechanism

The $K_L^0 \to 3\pi^0$ decay has six photons in the final state, so four photons can be used to veto the event. As a result, this background is highly suppressed by the veto cuts; however, it can contribute if there are multiple pulse pileups and the veto timing is incorrectly calculated. Veto detector inefficiencies also contribute to this background, and since the branching ratio of $K_L^0 \to 3\pi^0$ is large ($\sim$19%) the chance that overlapping pulses or detector inefficiencies occur is not negligible. This background is considered a "masking" background because of the high probability that accidental activities will mask photon hits in the veto detectors, cause a miscalculation of the veto timing, and result in a background.

Estimation Method

The recycling method described in Section 6.1.1 was used to create a background sample of $K_L^0 \to 3\pi^0$ masking events. First, events selected from the full $K_L^0 \to 3\pi^0$ MC simulation without the accidental overlay applied were used as seed samples. These seed samples then had different accidental waveforms from the TMon data overlaid 10,000 to 30,000 times for each seed sample to enhance the probability of masking. The MC was scaled to the amount of data collected and correction factors were applied to account for the loss due to CsI accidentals not accounted for in the seed samples and for cut differences among the seed samples and the final cut set.

Result

After applying all cuts, no remaining events were observed in the MC simulation and the estimated number of background events in the signal region for the $K_L^0 \to 3\pi^0$

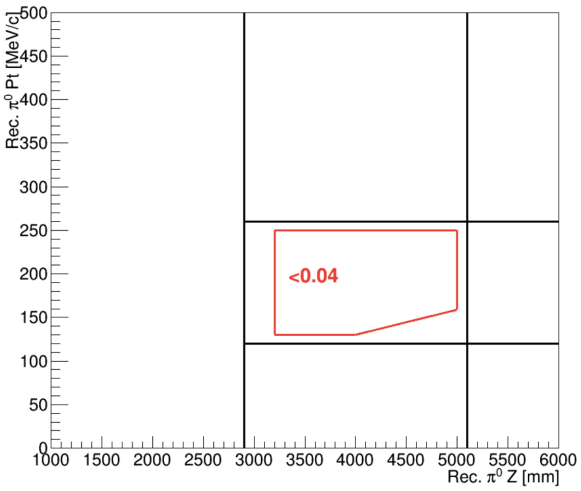


Figure 8.7: The estimated $K_L^0 \to 3\pi^0$ background contribution in the P_T–Z plane.

masking background was < 0.04 at the 90% CL as shown in Figure 8.7.

8.3.5 $Ke3$ Masking Background

Mechanism

The decay mode $K_L^0 \to \pi^\pm e^\mp \nu_e$ is often referred to as "$Ke3$" since it has three final state particles and one of them is an electron[5]. This decay can become a background if the CV detector fails to detect the two charged particles and they both hit the calorimeter and create two cluster hits. We also consider this decay mode a masking background because if there are any accidental hits in the CV detector that occur at the same time as this decay, the true veto hits could be masked by accidental pulses. Since this decay mode has a large branching ratio of about 40%, the background contribution must be considered.

Estimation Method

To estimate the $Ke3$ background we used a method similar to the estimation method used for the $K_L^0 \to 3\pi^0$ background. In this case we simulated full $K_L^0 \to \pi^\pm e^\mp \nu_e$ events without accidental overlay and selected events to be used as seeds. Dif-

[5]Similarly, the decay $K_L^0 \to \pi^\pm \mu^\mp \nu_\mu$ is called "$K\mu 3$".

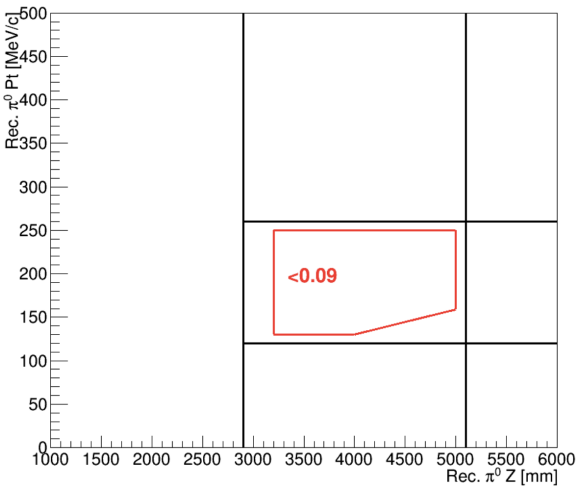


Figure 8.8: The estimated contribution in the P_T–Z plane for the $Ke3$ masking background.

ferent accidental waveforms were overlaid on each seed sample 10,000 to 30,000 times to enhance the probability of masking and increase statistics. The MC was scaled to the data and the same correction factors applied to the $K_L^0 \to 3\pi^0$ background were computed and applied to correct for the loss introduced by the $Ke3$ background estimation method.

Result

There were no remaining events observed in the MC simulation after all cuts were applied and the estimated number of $Ke3$ masking background events in the signal region was < 0.09 at the 90% CL as shown in Figure 8.8. It should also be noted that the $K\mu3$ background was estimated but it was estimated to be zero.

8.4 Neutron-Induced Background

As described in Section 5.1.1.2, one of the largest sources of background results from halo neutrons in the beam. In this case, halo neutrons can either hit the calorimeter directly and produce hadronic showers that are mistaken for electromagnetic show-

ers in the case of the hadron-cluster background, or halo neutrons can hit detector material and produce secondary particles like an η or π^0 that subsequently decay into two photons. For the hadron-cluster background we used a special control sample of scattered neutron data to estimate the background contribution and for the other background sources, MC simulations were used. The details of each background source are described below.

8.4.1 Hadron-Cluster Background

Mechanism

The hadron-cluster background is caused by a halo neutron directly hitting the calorimeter and producing a hadronic shower in which a neutron in the first shower travels some lateral distance inside the calorimeter and produces a secondary hadronic shower. These two distinct clusters on the calorimeter with no energy in the veto detectors are reconstructed as a π^0 and can produce a background as illustrated in Figure 8.9. Since they are both neutral, distinguishing between neutron and photon showers becomes a primary challenge. Discrimination methods using the shape of the cluster (CSDDL, Section 5.3.3.3) and the shape of the waveform pulse (FPSD, Section 5.3.3.3) were developed to differentiate neutron clusters from photon clusters.

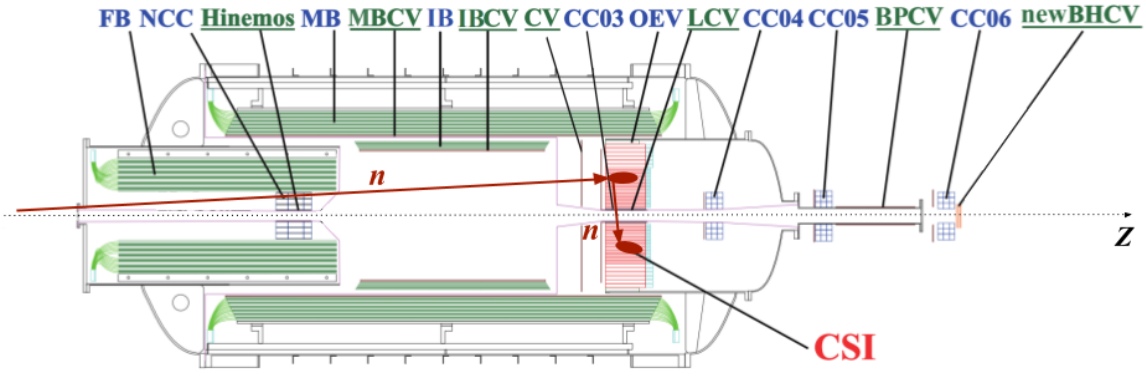


Figure 8.9: Mechanism for the hadron-cluster background in which a neutron directly hits the calorimeter and produces a primary and then secondary hadronic shower.

Estimation Method

To estimate the hadron-cluster background contribution, a data-driven approach was used. The Z0 Al target runs described in Section 4.4.1, were used to collect a control sample of scattered neutron data by inserting a thin Al plate into the beam. The neutrons scattered by the Al target were then used to estimate the background yield and the reduction power of different cuts. The reduction power is defined as the fraction of events that survive a cut, so a smaller reduction power makes the cut more effective at removing background. To ensure that the control sample from the Z0 Al target data matched the neutron distribution in the physics data, a cut set that excluded neutron related cuts was used, ensuring that most contributions from K_L decays were removed while neutron events remained. Figure 8.10 shows the distribution of physics data (left) and the Z0 Al target control sample (right). Aside from the events in the upstream region due to the upstream π^0 background, there is good agreement between the distributions and so the control sample can be used to estimate the neutron background in the physics data.

To calculate the number of events for the hadron-cluster background we scaled the

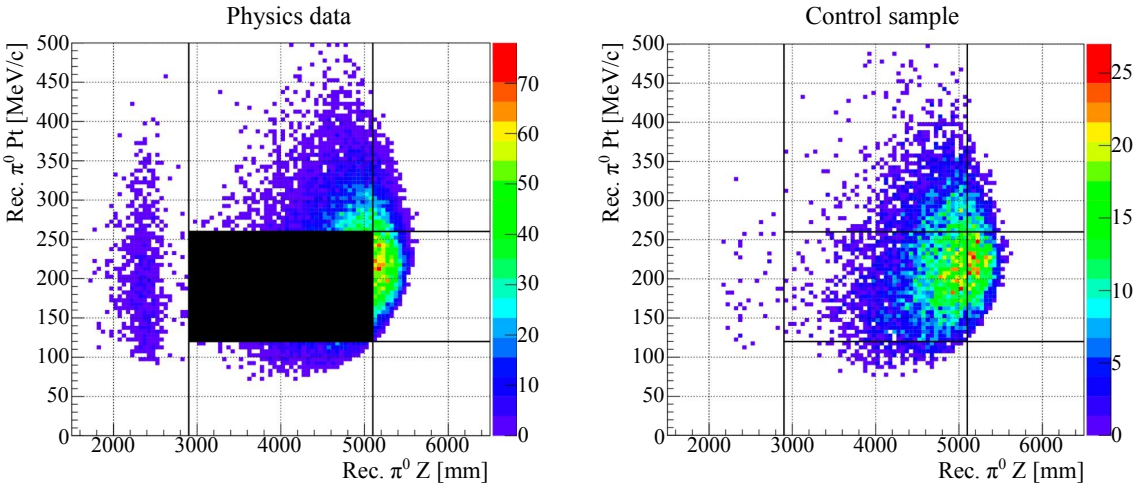


Figure 8.10: The distribution of neutron events in the physics data (left) and in the Z0 Al target control sample (right).

number of remaining events in the control sample after applying all cuts to the amount of physics data collected (the scaling was done for the region 2900 mm $< Z^{\pi^0}_{\rm vtx} <$ 6000 mm and 120 MeV/c $< P^{\pi^0}_T <$ 500 MeV/c excluding the blinded region) and then multiplied the scaled remaining events by the reduction power of the neutron cuts. The reduction power of the neutron cuts (CSDDL and FPSD) was determined by taking the number of events expected in the region 2900 mm $< Z^{\pi^0}_{\rm vtx} <$ 6000 mm and 120 MeV/c $< P^{\pi^0}_T <$ 500 MeV/c and dividing by the number of Z0 Al target events in that region without the neutron cuts.

Result

Using the method described above, the estimated background contribution in the signal region for the hadron-cluster background was estimated to be 0.017 ± 0.002 events. The distribution in the P_T–Z plane is shown in Figure 8.11.

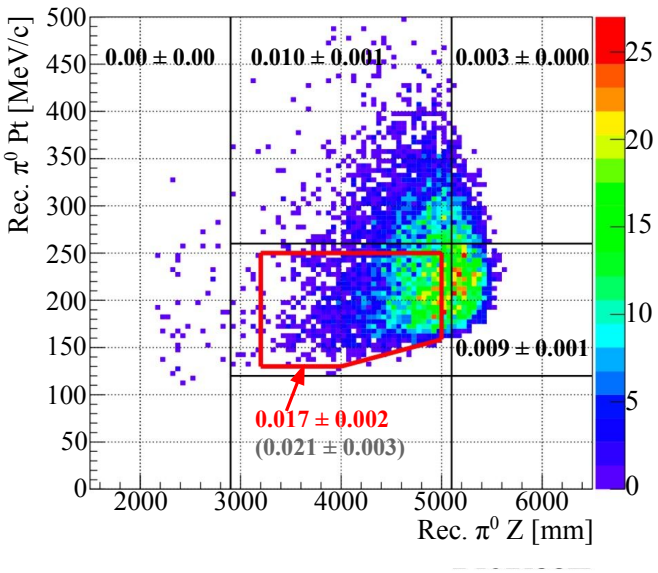


Figure 8.11: The estimated hadron-cluster background contribution in each region of the P_T–Z plane. The values in black are the estimated number of events in the corresponding region, the value in red is the estimated number of events in the signal region, and the grey value in parenthesis is the estimated contribution in the blinded region including the signal region. Regions without values have an estimated contribution of zero.

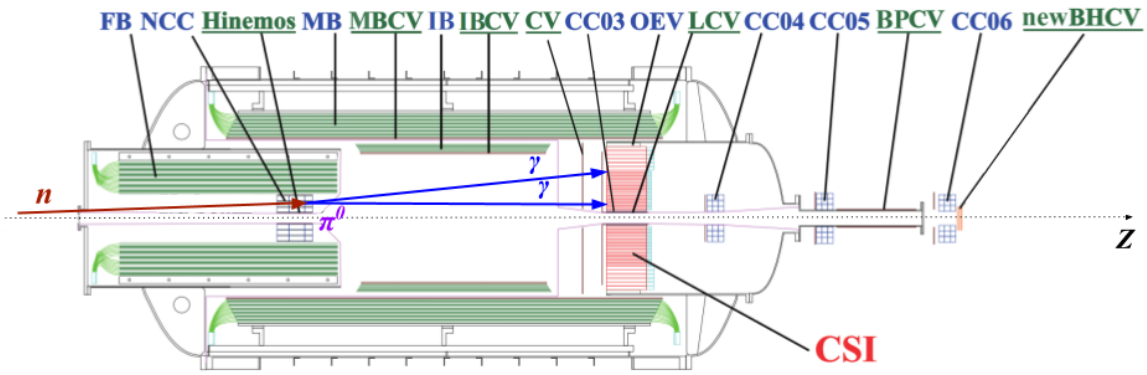


Figure 8.12: Mechanism for the upstream π^0 background.

8.4.2 Upstream π^0 Background

Mechanism

The upstream π^0 background occurs when a halo neutron hits the NCC detector and the interaction produces a π^0 which subsequently decays into two photons. A depiction of this background can be seen in Figure 8.12. Even though the z decay vertex is around 2500 mm, which is outside of the signal region, if the energy of the photons is mis-reconstructed to be smaller than the true photon energy due to photo-nuclear interactions, the reconstructed decay vertex can be shifted downstream into the signal region and cause a background.

Estimation Method

We used MC simulations and the recycling method described in Section 6.1.1 to estimate the background contribution from the upstream π^0 background. Halo neutrons were simulated at the beam exit and those events that interacted with the NCC detector and produced a π^0 were used as seeds. The π^0s were required to have a momentum greater than 100 MeV/c, and if they did the information was recycled 100 times to increase statistics. Next, the π^0 decay was simulated and the initial momentum of the photons was used to check whether the photons generated from the π^0 decay hit the calorimeter. At least one photon was required to have an energy

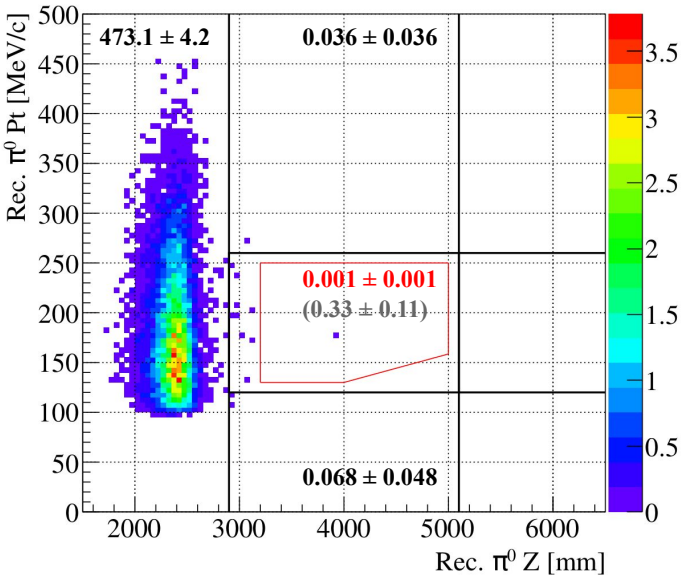


Figure 8.13: The estimated upstream π^0 background contribution in each region of the P_T–Z plane. The color of the number values follow the same convention as in Figure 8.11.

greater than 100 MeV and events that passed these criteria were used for the full simulation. As the upstream π^0 background is mainly distributed in the upstream region, the MC sample was normalized to the data using the region $Z^{\pi^0}_{\rm vtx} < 2900$ mm and a loose cut set that eliminated K_L decays and hadron-cluster events.

Result

After imposing all cuts, the number of upstream π^0 events in the signal region was estimated to be 0.001 ± 0.001 as shown in Figure 8.13. The estimated contribution in the upstream region was 473.1 ± 4.2 events which is consistent with the 462 events observed in this region within a few percent as shown in Figure 8.17.

8.4.3 CV-π^0 Background

Mechanism

A neutron in the beam halo can also hit the CV detector in front of the CsI and the interaction can produce $\pi^0 \to 2\gamma$ as depicted in Figure 8.14. Though the decay

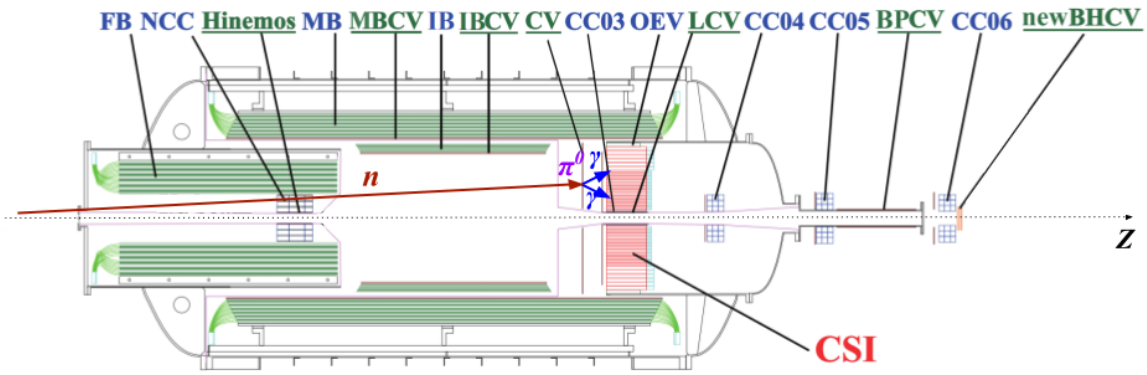


Figure 8.14: Mechanism for the CV-π^0 background in which a neutron hits the CV detector and produces a π^0 which subsequently decays to two photons.

vertex is outside of the signal region (Z_{vtx} ~5500–6000 mm), if a secondary particle associated with the π^0 production hits the calorimeter near one of the photon hits and increases the measured photon energy, the reconstructed decay vertex can be shifted upstream into the signal region. The Reconstructed Polar Angle χ^2 cut (Section 5.3.3.3) was developed to mitigate this background, as well as the CV-η background.

Estimation Method

To estimate the CV-π^0 background, the recycling method was used to generate MC samples with high statistics. To make the samples, halo neutrons were simulated from the beam exit and used as seeds. The halo neutrons that hit the CV without energy deposited on the upstream detectors were then used to make seeds of π^0s generated in the CV. At least one π^0 with an energy greater than 100 MeV had to be generated and the seeds that passed the momentum requirement were recycled 20 times. The π^0 decay was then simulated and the initial momentum of the photons was used to check whether the photons generated from the π^0 decay hit the calorimeter. At least one photon was required to have an energy greater than 100 MeV and events that passed these criteria were used for the full simulation. Finally, the simulation was scaled to the amount of data collected using the downstream region ($Z^{\pi^0}_{\text{vtx}} > 5100$ mm) and a loose cut set that eliminated K_L decays and hadron-cluster events.

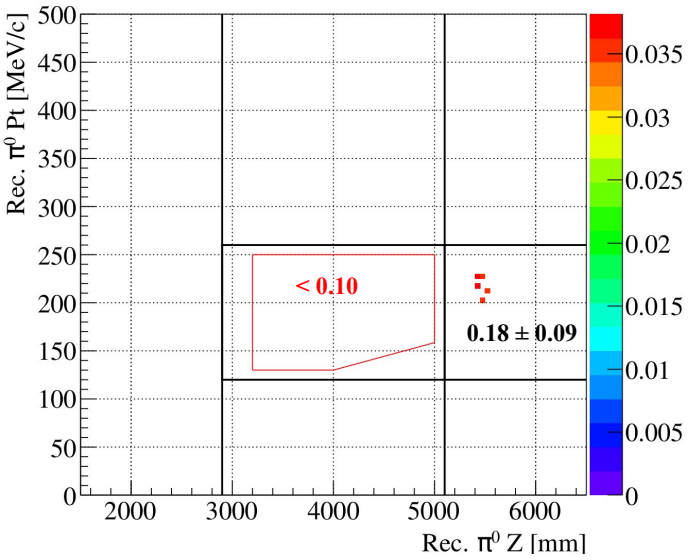


Figure 8.15: The estimated CV-π^0 background contribution in each region of the P_T–Z plane. The color of the number values follow the same convention as in Figure 8.11.

Result

Figure 8.15 shows the distribution of CV-π^0 events in the P_T–Z plane. After applying all cuts, there were no remaining events in the signal region and the estimated background contribution in the signal region was < 0.10 events at the 90% CL.

8.4.4 CV-η Background

Mechanism

The CV-η background has the same mechanism as the CV-π^0 background except in the interaction at the CV, an η particle is produced instead of a π^0. The η decays to two photons about 40% of the time and has an invariant mass of $M_\eta = 547.862$ MeV/c^2 [10]. When a π^0 is reconstructed with the two photon hits, the heavy mass of the η causes the reconstructed decay vertex to be pushed upstream into the signal region causing a background. The η-CSD cut described in Section 5.3.3.3 uses neural networks to distinguish this background from signal events.

Estimation Method

The estimation method for the CV-η background is very similar to the method used for the CV-π^0 background and the recycling method was used to increase statistics. Seeds of halo neutrons were generated in the same way and with the same criteria, and these halo neutron seeds were used to make seeds of η particles generated in the CV detector. At least one η with at least 100 MeV had to be generated and the seeds that met the criteria were recycled 200 times. The η decay was then simulated and both photons were required to hit the calorimeter with at least one photon having at least 100 MeV of energy deposited. Events that passed these requirements were then used as inputs to the full simulation. The same method used for the CV-π^0 background was used to scale MC to the amount of data collected.

Result

After imposing all cuts, the distribution of CV-η events is shown in Figure 8.16. The estimated contribution for the CV-η background in the signal region was 0.03 $\pm$ 0.01 events.

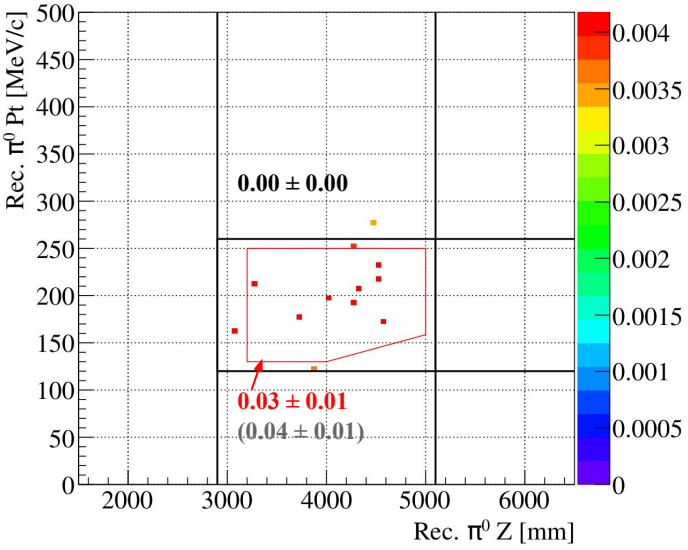


Figure 8.16: The estimated background for CV-η in each region of the P_T–Z plane. The color of the number values follow the same convention as in Figure 8.11.

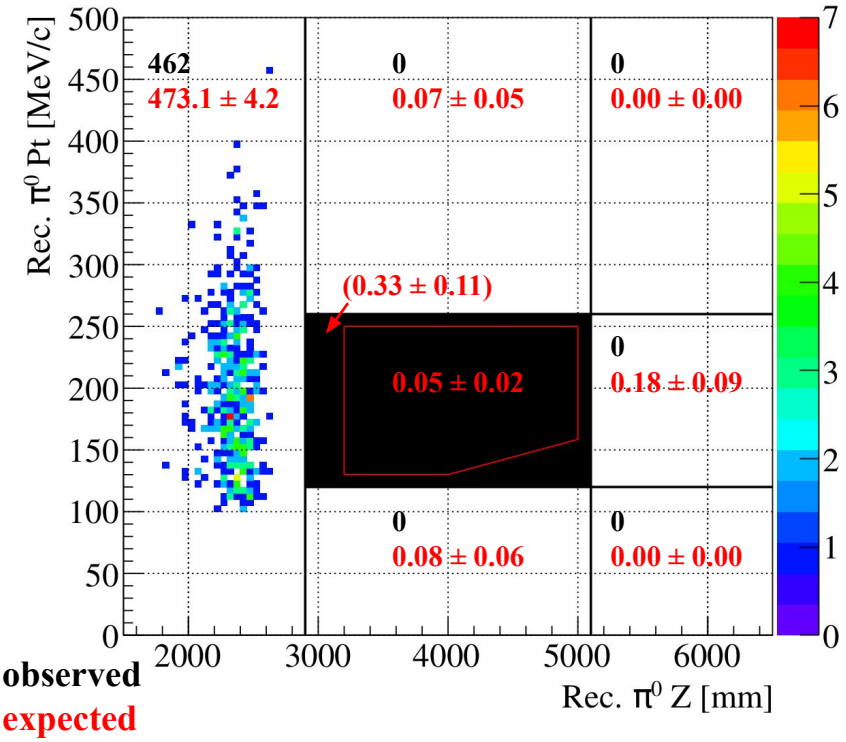


Figure 8.17: P_T–Z plot of the observed data with the estimated number of background events in each region for the 2016–2018 $K_L^0 \to \pi^0\nu\bar{\nu}$ blind analysis. The black number values indicate the observed number of events in that region and the red numbers indicate the estimated number of background events expected in that region. The red value in parenthesis is the estimated number of background events expected in the blinded region including the signal region.

8.5 Summary of Background Estimation

The background contributions were estimated for K_L decays and neutron-induced events using MC simulations and a control data sample for the hadron-cluster background. Figure 8.17 shows the P_T–Z plot with the data and the expected number of background events in each region after imposing all cuts and scaling the background estimations to the amount of data collected. The number of observed events is consistent with the background estimations in all of the sideband regions and there were no observed events in any of the sideband regions except for the upstream region containing upstream π^0 background events. The total number of background events estimated in the signal region was 0.05 ± 0.02.

Background Source	Estimated # of Background
K_L Decay Background	
$K_L^0 \to \pi^+\pi^-\pi^0$	< 0.02
$K_L^0 \to 2\pi^0$	< 0.18
$K_L^0 \to 2\gamma$	0.005 ± 0.005
$K_L^0 \to 3\pi^0$ (masking)	< 0.04
$K_L^0 \to \pi^\pm e^\mp \nu_e$ (masking)	< 0.09
Neutron-Induced Background	
Hadron-Cluster	0.017 ± 0.002
Upstream π^0	0.001 ± 0.001
CV-π^0	< 0.10
CV-η	0.03 ± 0.01
Total	0.05 ± 0.02

Table 8.2: Summary table of the estimated background contribution in the signal region for each background source. The largest sources that contribute to the total are the hadron-cluster background and the CV-η background. The backgrounds estimated with an upper limit were estimated at the 90% CL.

Table 8.2 lists the estimated background contribution in the signal region for each background source. The main contributions to the total estimated background of 0.05 ± 0.02 were from the hadron-cluster background and the CV-η background. The uncertainty on the background estimation comes from the statistical sizes of the MC samples and the Z0 Al target data sample.

Chapter 9

Results

After completing the event reconstruction and selection (Chapter 5), the normalization analysis (Chapter 7), and the background estimation (Chapter 8), the final steps in the analysis were to estimate the systematic uncertainties, calculate the Single Event Sensitivity, and "open the box" by unblinding the data in the signal region. As noted in Chapters 7 and 8, the normalization analysis yielded a total number of K_L at the beam exit of $(6.83 \pm 0.05) \times 10^{12}$ and the total number of estimated background events in the signal region was 0.05 ± 0.02. This chapter details the calculation of the systematic uncertainties, the Single Event Sensitivity, and the results and details of the unblinded data.

9.1 Systematic Uncertainties

As mentioned in Section 5.1.2, the branching ratio of $K_L^0 \to \pi^0 \nu\bar{\nu}$ is calculated as

$$BR(K_L^0 \to \pi^0 \nu\bar{\nu}) = \frac{N_{\text{signal}}}{N_{K_L^0} \times A_{\text{signal}}} \tag{9.1}$$

where N_{signal} is the number of signal events, $N_{K_L^0}$ is the total number of K_Ls at the beam exit, and A_{signal} is the signal acceptance. Correspondingly, the SES is the same

as the BR when $N_{\text{signal}} = 1$. As discussed in Chapter 7, the $K_L^0 \to 2\pi^0$ normalization mode was used to calculate $N_{K_L^0}$. Using the equations from Section 7.4 the SES can be written as

$$SES = \frac{1}{A_{\text{signal}}} \frac{A_{K_L \to 2\pi^0} \times BR(K_L \to 2\pi^0)}{N_{K_L \to 2\pi^0} \times \text{prescale}} \tag{9.2}$$

where $A_{K_L \to 2\pi^0}$ is the acceptance of the $K_L^0 \to 2\pi^0$ normalization mode, BR($K_L \to 2\pi^0$) is the branching ratio of $K_L^0 \to 2\pi^0$, $N_{K_L \to 2\pi^0}$ is the number of $K_L^0 \to 2\pi^0$ events remaining in the data after the normalization event selection, and the prescale factor is 30. By expressing the SES in this manner, systematic uncertainties that are common between $A_{K_L \to 2\pi^0}$ and A_{signal} can be cancelled, such as discrepancies between data and MC, DAQ efficiency, veto detector response, decay probability, and some of the kinematic cut selections. The acceptances can be written as

$$A_{K_L \to 2\pi^0} = A_{geom}^{K_L \to 2\pi^0} \times A_{trigger}^{K_L \to 2\pi^0} \times A_{\gamma}^{K_L \to 2\pi^0} \times A_{kine}^{K_L \to 2\pi^0} \times A_{veto}^{K_L \to 2\pi^0} \tag{9.3}$$

$$A_{sig} = A_{geom}^{sig} \times A_{trigger}^{sig} \times A_{\gamma}^{sig} \times A_{kine}^{sig} \times A_{shape}^{sig} \times A_{veto}^{sig} \tag{9.4}$$

where the subscripts refer to the type of acceptance– *geom*, *trigger*, γ, *kine*, *shape*, and *veto* represent the acceptance associated with detector geometry, trigger-related effects, photon selection cuts, kinematic cuts, shape-related cuts, and veto cuts. The acceptances were estimated using MC simulations or comparisons between data and MC. Each acceptance has a corresponding systematic uncertainty and these are described in the following subsections. The basic strategy to estimate the uncertainty on the acceptance is to prepare a control sample and compare the data and MC and calculate the difference between them for each cut and each run period. The weighted average over all run periods is taken and then the quadratic sum for each cut uncertainty is computed to get the total uncertainty. The other uncertainties

that contributed to the $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis were the uncertainty on the $K_L^0 \to 2\pi^0$ branching ratio and the uncertainty due to the inconsistency in the three normalization modes. The total uncertainty is summarized in Section 9.1.9.

9.1.1 Geometric Acceptance

The geometric acceptance was defined as the probability that a K_L entering the KOTO detectors decays within the region of $3000 < z < 4700$ mm and all photons produced in either the $K_L^0 \to \pi^0\nu\bar{\nu}$ or the $K_L^0 \to 2\pi^0$ decay hit the CsI calorimeter in the fiducial region[1]. The geometric acceptance was calculated using the fast simulation method described in Section 6.1.1 and the uncertainty comes from the K_L momentum spectrum. To estimate this uncertainty, the relative change in the acceptance ratio of $A_{geom}^{K_L\to 2\pi^0}/A_{geom}^{sig}$ when varying the parameters (μ, σ_0, A, S) in the K_L momentum spectrum in Equation 6.1 was evaluated. The distribution of the relative deviation of the acceptance ratio over 10^6 trials with various parameter settings is shown in Figure 9.1. The systematic uncertainty was defined as the 1σ range of the relative change and was estimated to be 0.98%. This is listed as the "K_L momentum spectrum" uncertainty in Table 9.1.

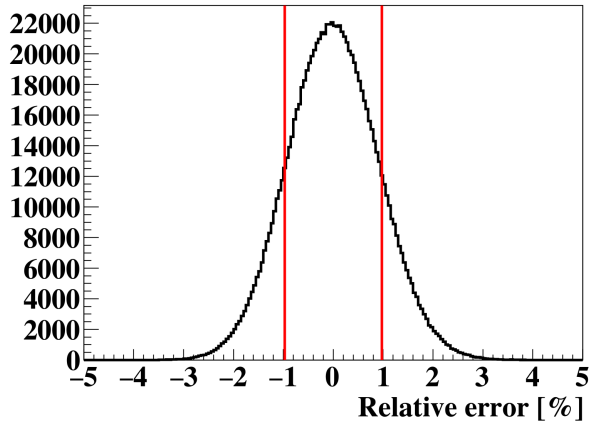


Figure 9.1: The relative deviation of the geometric acceptance ratio with various K_L momentum spectrum parameters. The red lines are the 1σ bounds.

[1]Defined in Section 5.3.3.1

9.1.2 Trigger-Related Effects

The efficiencies of the trigger-related cuts described in Section 5.3.2 affected the acceptance of the online trigger effect ($A_{trigger}$) and so systematic uncertainties were introduced. These uncertainties were evaluated separately for the CsI total energy (E_t) cut, the COE radius cut, the average photon cluster time cut, and the CDT cut and then the total trigger-related uncertainty was calculated as the combined quadrature sum from all of the sources. The E_t cut efficiency uncertainty was evaluated by studying the deviation in the efficiency function when the normalization area for the energy distribution was varied from 1000 to 1800 MeV. Because the E_t cut efficiency was almost 100%, the uncertainty was negligibly small ($< 0.1\%$) and so this systematic uncertainty was ignored.

The uncertainty on the COE cut efficiency was evaluated by comparing the data and MC of validation samples. The samples used were normalization trigger data and MC with two photon events outside of the signal region with a loose cut selection applied. The systematic uncertainty was calculated as the discrepancy between the ratio of the COE cut efficiencies between data and MC and was estimated to be 0.68%. Since the COE cut was only used for Runs 69 and part of Run 75, the uncertainty was scaled to the SES by multiplying by a factor of 0.27, resulting in a systematic uncertainty of 0.18%.

To evaluate the uncertainty introduced by the average photon cluster time cut, the deviation from the efficiency double ratio was calculated in Equation 9.5,

$$\epsilon \text{ ratio deviation} = \left[\left(\frac{\epsilon^{MC}_{K_L \to 2\pi^0}}{\epsilon^{data}_{K_L \to 2\pi^0}} \right) / \left(\frac{\epsilon^{MC}_{K_L \to 2\gamma}}{\epsilon^{data}_{K_L \to 2\gamma}} \right) \right] - 1 \tag{9.5}$$

where the efficiency, ϵ, is defined in Equation 7.5. Data and MC control samples of $K^0_L \to 2\pi^0$ and $K^0_L \to 2\gamma$ events were used to evaluate the deviation from the

efficiency ratio between data and MC with all cuts applied except for the average photon cluster time cut. The $K_L^0 \to 2\gamma$ events were used instead of $K_L^0 \to \pi^0\nu\bar{\nu}$ since they both have the same final state of two photons. The weighted average of the deviation from the efficiency ratio over all run periods was regarded as the systematic error and was estimated to be 0.18%.

The final trigger-related source of uncertainty was due to the CDT trigger cut efficiency. A control sample of minimum bias data and MC was used with a loose cut selection applied. The deviation from the efficiency ratio for data and MC was calculated in a similar way as the average photon cluster time cut and the data and MC were compared with and without the CDT trigger cut applied. The deviation was calculated for each run period and the weighted average of all run periods was calculated as 0.08% uncertainty. Since the CDT trigger was introduced in Run 75, this uncertainty was scaled to the SES by multiplying by a factor of 0.82, resulting in a systematic uncertainty of 0.07%. To get the final systematic uncertainty on the trigger-related effects, the quadrature sum was taken as

$$\sigma_{trigger} = \sigma_{CSIEt} \oplus \sigma_{COE} \oplus \sigma_{timing} \oplus \sigma_{CDT} \tag{9.6}$$

and the resulting systematic uncertainty was 0.26%.

9.1.3 Photon Selection Cuts

The uncertainty from the photon selection cuts (E_γ and CsI fiducial cuts described in Section 5.3.3.1) was estimated using a control sample of π^0s taken from $K_L^0 \to 2\pi^0$ data and $K_L^0 \to 2\pi^0$ MC with all cuts except for the photon selection cuts applied. The deviation (Δ_i) in acceptances between data and MC were compared using Equation 9.8 where the acceptance of the photon selection cut, A^i, has the same definition of the cut efficiency defined in Equation 7.5.

$$\Delta_i = (A^i_{MC} - A^i_{data})/A^i_{data} \tag{9.7}$$

The deviation was calculated for all run periods and the weighted average over all periods was taken as the uncertainty for each photon selection cut. Finally, the total systematic uncertainty of the photon selection cuts was determined by taking the quadrature sum of each photon selection cut uncertainty. This systematic uncertainty was estimated to be 0.57%.

9.1.4 Kinematic Cuts for $K^0_L \to \pi^0\nu\bar{\nu}$

The systematic uncertainty for the kinematic cuts for the signal[2] was estimated in the same way as the uncertainty on the photon selection cuts in Section 9.1.3. Control samples of $K^0_L \to 2\pi^0$ data and MC were used. Each pair of photons that formed a π^0 was used to reconstruct a π^0 assuming the decay vertex was on the beam axis as in the $K^0_L \to \pi^0\nu\bar{\nu}$ reconstruction method. The acceptances of each cut between data and MC were compared as in the above section, the weighted average over all run periods was taken, and the quadratic sum of all cut uncertainties was taken to be the total systematic uncertainty which was 2.9%.

9.1.5 Kinematic Cuts for $K^0_L \to 2\pi^0$

The $K^0_L \to 2\pi^0$ kinematic cuts[3] was different than the kinematic cuts on the signal, and so the uncertainty on the kinematic cuts for $K^0_L \to 2\pi^0$ were evaluated separately. The same method as described in Section 9.1.3 was used in which the difference of cut acceptances between data and MC were regarded as the systematic uncertainty. The total uncertainty was calculated to be 3.2%.

[2]Described in Section 5.3.3: γ cluster distance, γ cluster distance from dead ch., $\theta_{\mathrm{proj},\gamma}$, E_γ ratio, E_γ θ_γ, π^0 kinematic, offline COE, ΔT_{vtx}, π^0 rec. Z_{vtx}, π^0 rec. P_T

[3]Described in Section 7.2.3: γ cluster distance, ΔT_{vtx}, K_L rec. Z_{vtx}, K_L rec. P_T, $\Delta \mathrm{M}_{K_L}$, χ^2_{Z,K_L}, $\Delta \mathrm{M}_{\pi^0}$, $\Delta \mathrm{Z}_{\pi^0}$, K_L beam exit X,Y position

9.1.6 Shape-Related Cuts

The uncertainty on the shape-related cuts[4] was calculated using the same method as described in Section 9.1.4. The resulting total systematic uncertainty was 5.2%.

9.1.7 Veto Cuts

To evaluate the systematic uncertainty in the differences between the data and MC on the veto cuts, the same double ratio method used for the average photon cluster time uncertainty in Section 9.1.2 was used. Data and MC samples of $K_L^0 \to 2\pi^0$ and $K_L^0 \to 2\gamma$ events with kinematic cuts and a selection of loose veto cuts applied were used to evaluate the deviation of the efficiency double ratio as described in Equation 9.5. The weighted average over all periods was taken to be the systematic uncertainty for each veto cut and then the total systematic uncertainty was calculated by taking the quadratic sum of all veto cut uncertainties. The ensuing total systematic uncertainty for the veto cuts was 3.2%.

9.1.8 Other Sources of Uncertainty

The other sources of systematic uncertainty considered were the error on the branching ratio of the $K_L^0 \to 2\pi^0$ decay, which is 0.69% [10], and the inconsistency in the normalization analysis among the three normalization modes as discussed in Section 7.4.4. The difference seen in the K_L flux and the total number of K_L ($N_{K_L^0}$) between the three modes is not yet understood and so it was counted as a source of systematic uncertainty. This uncertainty was defined as the maximum difference between the SES calculated using the three normalization mode $N_{K_L^0}$ results and was estimated to be 5.2%.

[4]Described in Sections 5.3.3.2 and 5.3.3.3: γ cluster size, RMS_{clus}, CSDDL, η CSD, $\pi^+\pi^-\pi^0$ DL, FPSD, χ^2_θ

Source	Uncertainty [%]
K_L momentum spectrum	0.98
trigger-related effect	0.26
photon selection cuts	0.57
kinematic cuts for $K_L^0 \to \pi^0\nu\bar{\nu}$	2.9
kinematic cuts for $K_L^0 \to 2\pi^0$	3.2
shape-related cuts	5.2
veto cuts	3.2
$K_L^0 \to 2\pi^0$ branching ratio	0.69
normalization mode inconsistency	5.2
Total	9.2

Table 9.1: Summary of relative systematic uncertainties on the Single Event Sensitivity.

9.1.9 Total Uncertainty

The systematic uncertainties for the $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis are summarized in Table 9.1. The largest sources of uncertainty were due to the error on the shape-related cuts and the inconsistency between the three normalization modes. All of the sources of uncertainty were combined as their quadratic sum and the resulting total relative systematic uncertainty was 9.2%. It should also be noted that the total relative statistical uncertainty from $N_{K_L \to 2\pi^0}$ was 0.75%. Thus, the total relative uncertainty on the $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis was

$$\pm\,(1/\sqrt{N_{sig}} \times 100)\% \pm (0.75\%)_{\text{stat}} \pm (9.2\%)_{\text{syst}} \tag{9.8}$$

where the first term is the Poisson statistical uncertainty on the number of observed candidate events.

9.2 Single Event Sensitivity

The Single Event Sensitivity described in Section 5.1.2, represents how sensitive we are to observing a single event,

$$SES = \frac{1}{N_{K_L^0} \times A_{\text{signal}}} \tag{9.9}$$

where the variables are the same as in Equation 9.1. The $K_L^0 \to 2\pi^0$ normalization mode was used to calculate the total number of kaons at the beam exit which was $(6.83 \pm 0.05) \times 10^{12}$ and the signal acceptance was calculated to be 2.03×10^{-4}. The breakdown of the signal acceptance, number of K_L at the beam exit, and SES for each run period is listed in Table 9.2. Based on these values, the final SES for the 2016–2018 $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis was

$$SES = (7.20 \pm 0.05_{\text{stat}} \pm 0.66_{\text{syst}}) \times 10^{-10} \tag{9.10}$$

This is the highest level of experimental sensitivity for the $K_L^0 \to \pi^0\nu\bar{\nu}$ search to date, which improves on the previous level of sensitivity reached by KOTO (SES = $1.30 \pm 0.01_{\text{stat}} \pm 0.14_{\text{syst}} \times 10^{-9}$) [36] by a factor of 1.8. The systematic uncertainty

Run Period	$N_{K_L^0}$	$A_{\textbf{sig}}$	**SES**
Run 69 (0)	3.39×10^{11}	1.92×10^{-4}	1.53×10^{-8}
Run 69 (1)	2.56×10^{11}	1.88×10^{-4}	2.07×10^{-8}
Run 75 31 kW	6.48×10^{10}	2.85×10^{-4}	5.41×10^{-8}
Run 75 35 kW	1.40×10^{11}	2.71×10^{-4}	2.62×10^{-8}
Run 75 37.5 kW w/ COE w/o CDT	3.10×10^{11}	2.43×10^{-4}	1.33×10^{-8}
Run 75 37.5 kW w/ COE w/ CDT	5.31×10^{11}	2.43×10^{-4}	7.76×10^{-9}
Run 75 37.5 kW w/o COE w/ CDT	8.68×10^{11}	2.50×10^{-4}	4.61×10^{-9}
Run 78 33 kW	1.11×10^{11}	2.49×10^{-4}	3.63×10^{-8}
Run 78 44 kW	4.88×10^{10}	1.91×10^{-4}	1.07×10^{-7}
Run 78 50 kW	1.67×10^{12}	1.72×10^{-4}	3.49×10^{-9}
Run 79 51 kW	2.50×10^{12}	1.90×10^{-4}	2.11×10^{-9}
Total	6.83×10^{12}	2.03×10^{-4}	7.20×10^{-10}

Table 9.2: Number of K_L at the beam exit, signal acceptance, and SES for each run period in the 2016–2018 data set. Run 75 37.5 kW is further split into different periods in which the CDT trigger cut was replacing the COE trigger cut and different trigger conditions were used.

is greater than the statistical uncertainty and can be improved upon by identifying the discrepancy between the different normalization mode results, improving the agreement between data and MC for some of the advanced shape cuts that use deep learning, and increasing statistics.

9.3 Opening the Box

After all of the cuts were applied to the data, the normalization analysis was completed, the various backgrounds were estimated, and the SES was calculated, the signal box was opened to reveal what was in the blinded region. Figure 9.2 shows the PtZ plot for the unblinded 2016–2018 data set. The box was opened in August of 2019 and we observed four events in the signal region and one event in the upstream blinded region. After a careful rechecking of the analysis procedures and settings, we found an incorrect parameter setting for the nominal time in one

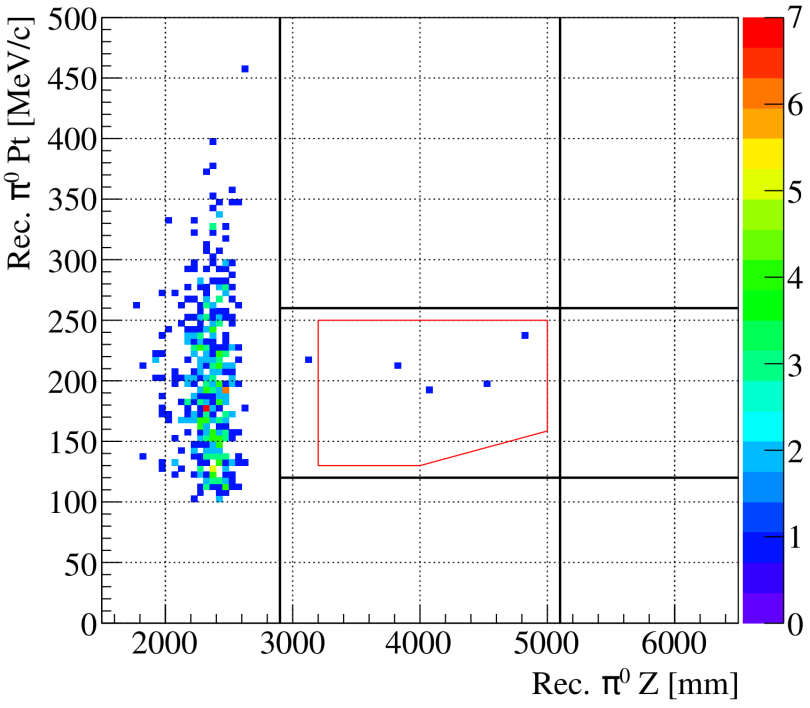


Figure 9.2: The PtZ plot after opening the box for the 2016–2018 $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis.

of the veto detectors that caused the second event from the right to be erroneously included. After making the correction to the nominal time and reprocessing the data with the same cut selection criteria, three of the original four candidate events in the signal region remained. The next section details the properties of the candidate events observed in the blinded region.

9.3.1 Properties of the Candidate Events

Since the background expected in the signal box before unblinding the data was estimated to be 0.05 ± 0.02 events, the candidate events and their properties were examined in close detail. The events in the blind region were assigned an ID based on the order in which the data was collected to help identify them, and this is shown in Figure 9.3 and Table 9.3. Event 2 was the event with the incorrect nominal time applied and so it is not a candidate event but it is included here to show the mistreatment of the waveform peak selection. Event 0 was found to have overlapped pulses in the NCC detector, suggesting that an accidental hit caused the veto timing to be calculated incorrectly causing a masking background, the details of which are covered in Section 9.3.1.1. Event 3 had hits in the FB just outside the veto window and is discussed in Section 9.3.1.4. Events 1 and (4) (outside the signal region) had no distinguishing features.

ID	**Run Period**	**run number**	**node**	**file**	**DST entry**	π^0 P_T **[MeV/c]**	π^0 $Z_{\mathbf{vtx}}$ **[mm]**
0	Run 69	23413	25	2	20074	194.36	4067.57
1	Run 69	24295	35	0	10527	235.97	4848.34
2	Run 75	26465	29	2	8462	197.86	4536.18
3	Run 79	29669	22	1	7164	213.08	3839.79
(4)	Run 79	29509	35	2	2330	215.19	3133.08

Table 9.3: The assigned ID and other identifying information for the observed candidate events including π^0 P_T and Z_{vtx}. The ID was assigned and sorted based on the run number. The node, file, and DST entry are other values to uniquely identify each event. The labeled candidate events can be seen in Figure 9.3.

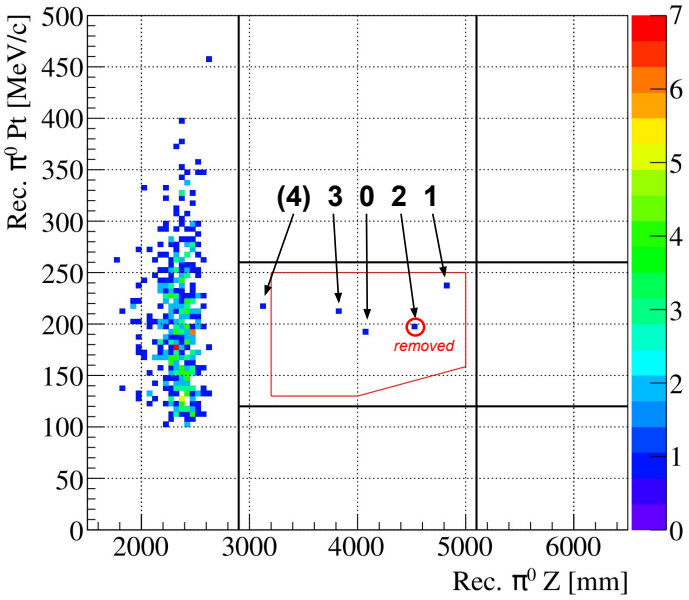


Figure 9.3: PtZ plot of the candidate events with their assigned IDs. Event 2 had an incorrect setting for the nominal time and so it was removed due to this mistake.

9.3.1.1 Event 0

Event 0 was collected in Run 69 and after a closer examination of the waveforms in the veto detectors, it was discovered to have overlapped pulses in the NCC detector which is located upstream and sits within the FB surrounding the beam hole[5]. The top left of Figure 9.4 shows a close up side-view of part of the NCC detector system, which is split up into three different modules– front, middle, and rear. Each module section of the NCC records waveform hits and then the "common" combination of all three detector modules was used to veto events. The right of Figure 9.4 shows the common waveform that was used for the veto decision and then the subsequent front, middle, and rear modules (below) with the waveforms recorded in these individual sections. In the front and middle modules of the NCC there were hits that were outside of the veto window time indicating that these hits were likely due to accidentals. However, in the rear module, an on-time hit occurred inside the veto window. Because the combination of the three modules was used in the veto decision, the accidental pulse

[5]See Section 2.3.5 for details on the NCC detector

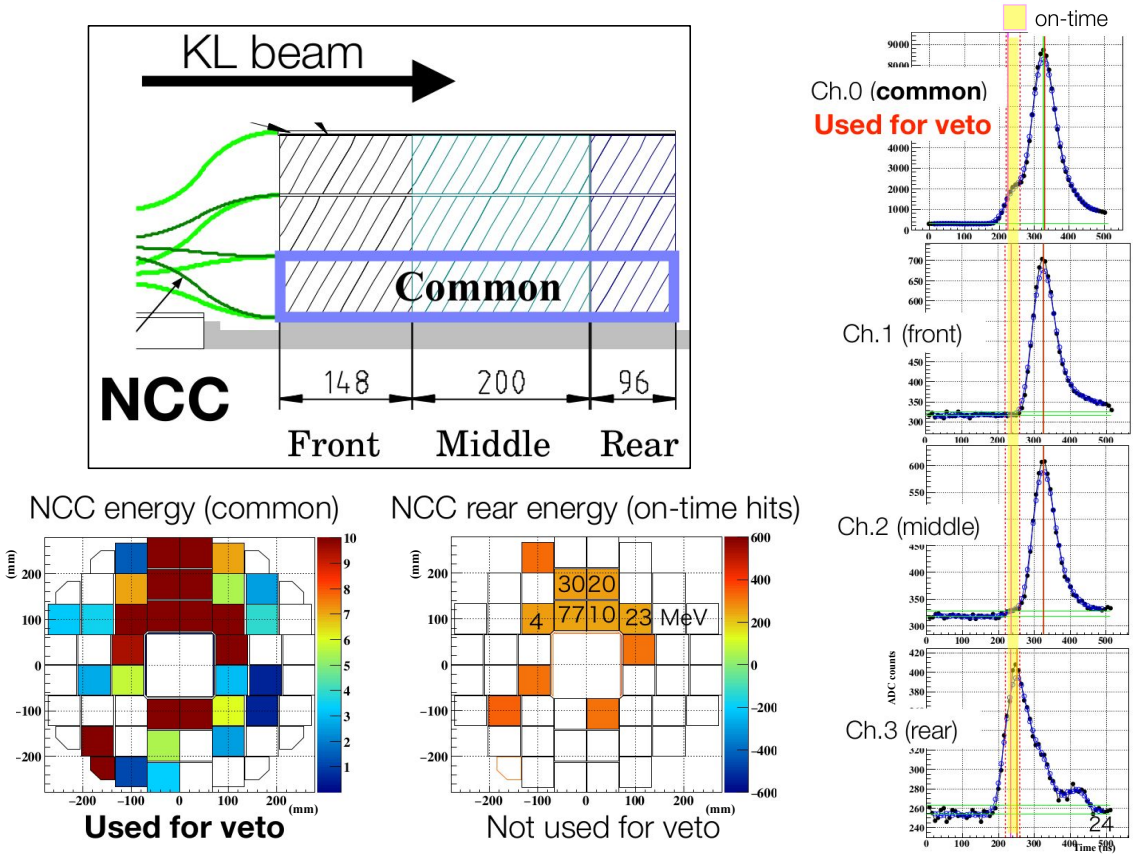


Figure 9.4: A detailed look at the properties of event 0 in which an on-time pulse in the rear NCC was masked by an accidental hit and not vetoed due to the combination of all modules being used for the veto decision. The top left shows the different modules (front, middle, and rear) of the NCC detector, the bottom left shows the NCC common energy vs the NCC rear energy, and on the right, an accidental hit outside the veto window in the front and middle modules masked the on-time hit in the rear module when they were combined into the common waveform used to veto.

in the front and middle modules masked the true hit in the rear module and this event could not be vetoed. The bottom left of Figure 9.4 shows the common energy distribution vs the rear energy distribution in the NCC, which had in-time hits with an energy range of 4 to 77 MeV. Due to waveform smoothing in obtaining the common waveform over all modules, the on-time peak disappeared and the FTT veto method was unable to identify a double pulse.

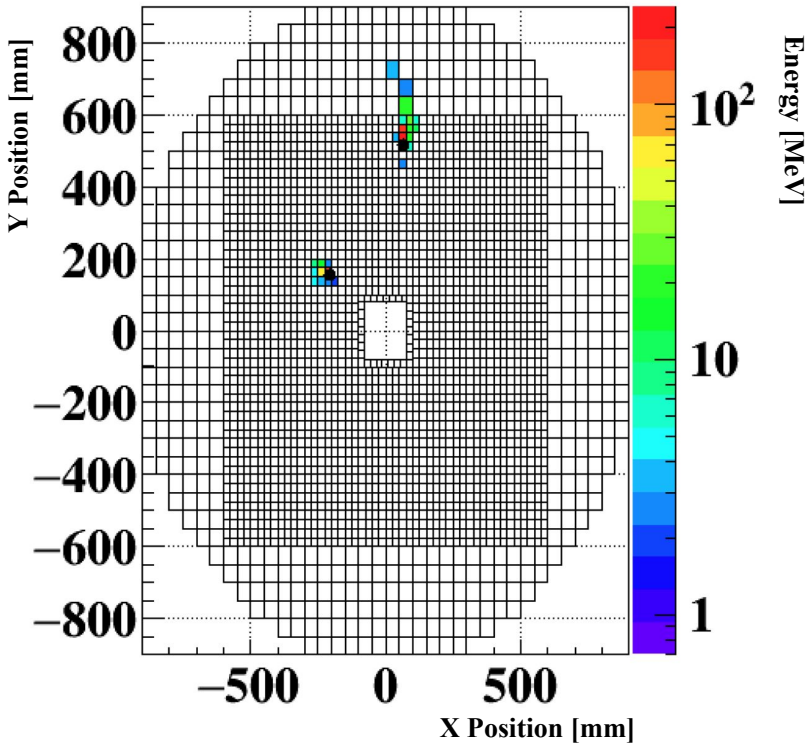


Figure 9.5: The energy deposited in the calorimeter for event 1. The black dots on each cluster indicate the reconstructed photon hit positions.

9.3.1.2 Event 1

Event 1 was collected in Run 69 and has no distinguishing features or energy deposited in any of the veto detectors. It is regarded as a candidate event. Figure 9.5 shows the energy deposited in the CsI calorimeter for event 1.

9.3.1.3 Event 2

Event 2 occurred in Run 75 and while it was one of the original four events in the signal box when the blinded region was revealed, a rechecking of the analysis parameters revealed that it should have been vetoed, as the nominal time was incorrectly set. The nominal time was a parameter determined for each veto detector as the most probable timing of all channels and was used to select waveform peaks. After checking for hits in the veto detectors for event 2, it was discovered that the

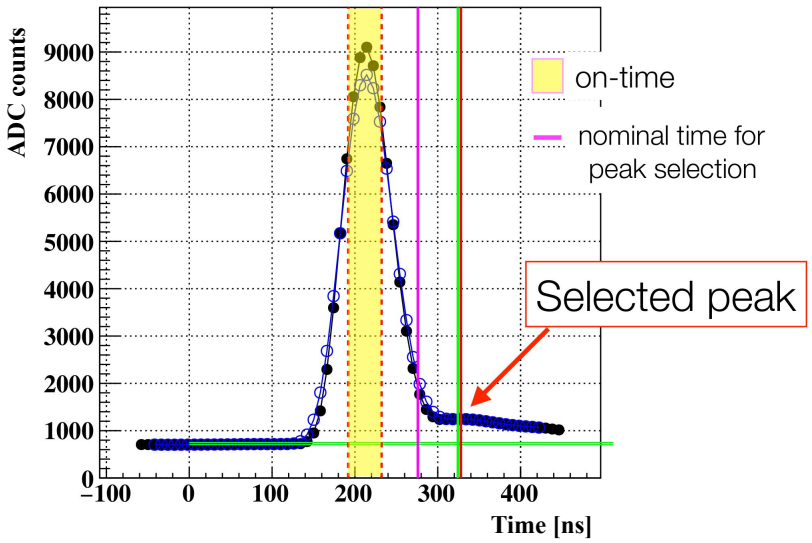


Figure 9.6: The mis-treatment of the peak selection due to an incorrect nominal time setting for event 2. The on-time waveform found in the HINEMOS detector was not used to veto because the selected peak (red/green line) was outside of the veto window. Had the correct nominal time (pink line) been used originally, this event would have been vetoed.

HINEMOS detector[6] had a hit within the veto window but since the nominal time was set incorrectly, this event was not vetoed as it should have been. Instead, a small peak after the true hit peak was selected as the on-time peak because it was closest to the erroneous nominal time. This can be seen in Figure 9.6. Setting the correct nominal time allowed this event to be vetoed. It should also be noted that this event also had energy hits in the CV detector, but they were below the veto threshold.

9.3.1.4 Event 3

Event 3 was collected in Run 79 and after checking for signals in the veto detectors, it was found to have a hit in the FB that occurred just before the veto window. Figure 9.7 shows the hit waveforms that were detected in channels 12 and 13 of the FB and the energy deposited in the FB. Because the hit timing occurs just before the veto window this event was not vetoed. Though this hit could have been due to accidentals,

[6]The inner scintillator of the NCC, described in Section 2.3.5

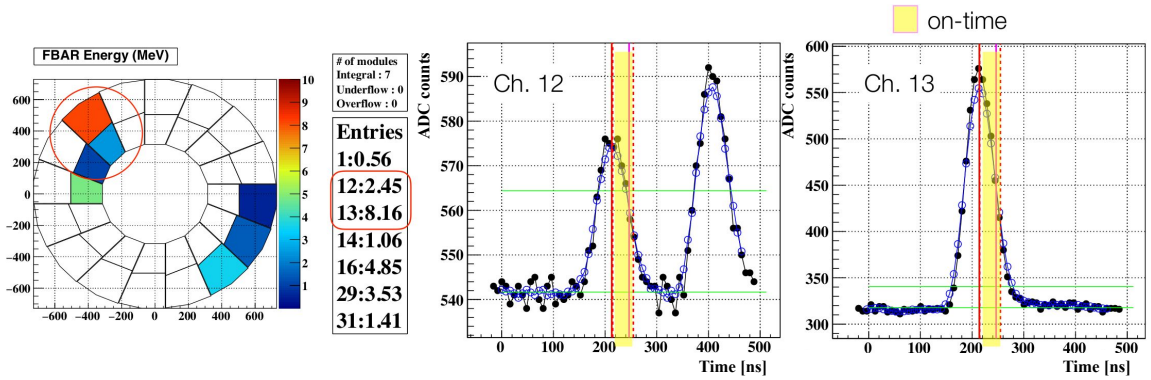


Figure 9.7: The FB energy deposits for event 3 (left) and the waveforms of the hit which occurred just before the veto window (right). The red line is the selected peak timing and the yellow area is the veto window.

we cannot be certain and so this event is regarded as a candidate event. The event display of the energy deposited in the calorimeter for this event is shown in Figure 9.8.

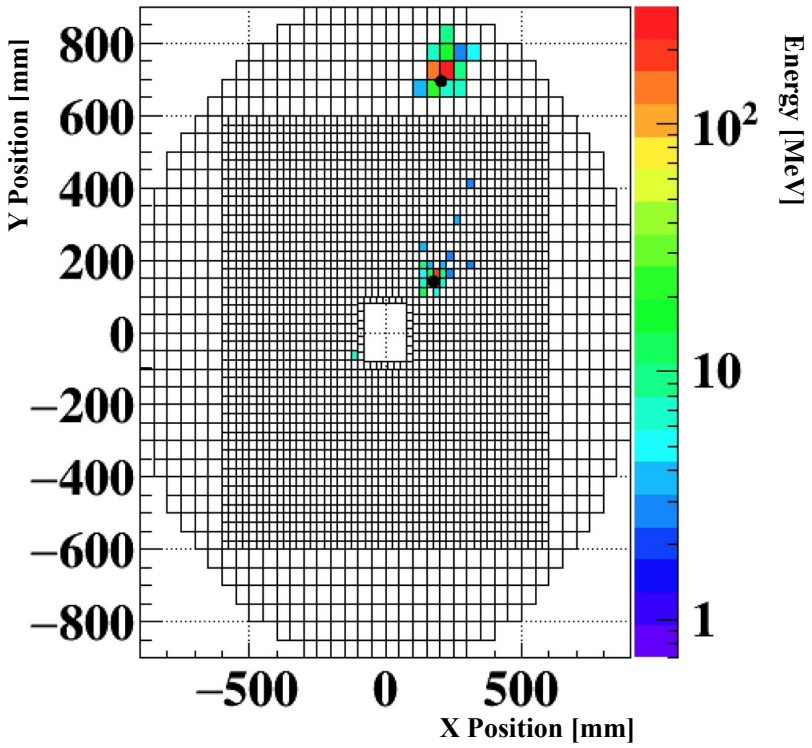


Figure 9.8: The energy deposited in the calorimeter for event 3. The black dots on each cluster indicate the reconstructed photon hit positions.

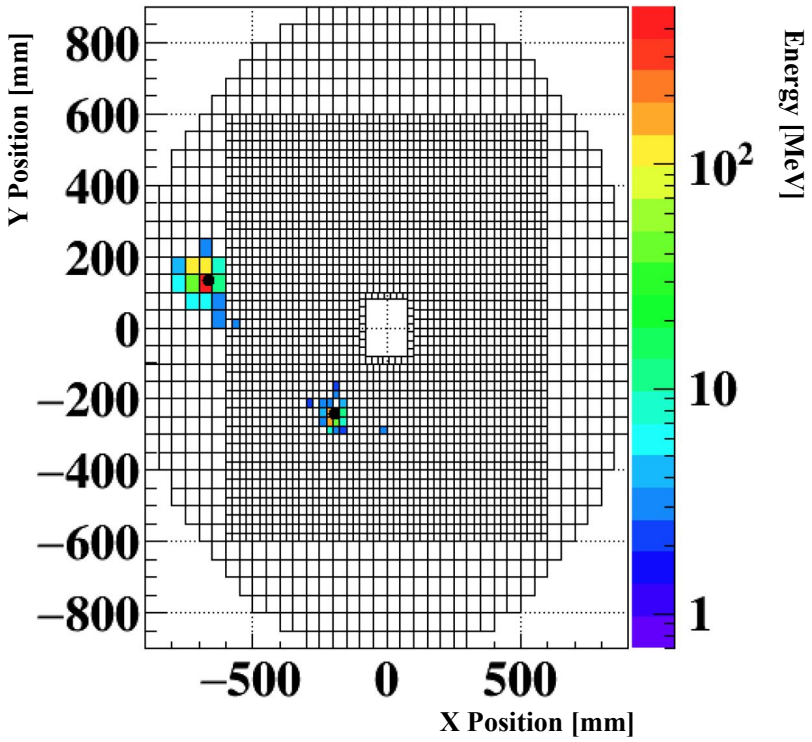


Figure 9.9: The energy deposited in the calorimeter for event (4). The black dots on each cluster indicate the reconstructed photon hit positions. This event was outside the signal region but within the blinded region.

9.3.1.5 Event (4)

Event (4) was collected in Run 79 and lies within the blinded region but outside the signal region. Because it is outside the signal region, it is not regarded as a candidate event, but its properties were studied and it was found to have no distinguishing features. Figure 9.9 shows the energy deposited in the calorimeter for this event.

9.4 Additional Background Studies

After examining the candidate events closely, the background estimations were re-visited and restudied. Since the original background estimation was 0.05 ± 0.02 and we observed three events in the signal region, we rechecked many of our background estimations and estimated contributions from other background sources that were not

previously considered. Table 8.2 in Chapter 8 shows the backgrounds that were estimated before opening the box and the following subsections describe the background studies done after examining the blind region. In particular, two main, new types of backgrounds were found and studied– a background resulting from $K^{\pm}$ decays which is called the $K^{\pm}$ background or the charged kaon background, (Section 9.4.1), and a background resulting from $K_L^0 \to 2\gamma$ decays that occur off the beam axis due to K_Ls in the beam halo which is called the halo K_L background (Section 9.4.2). Finally, Section 9.4.3 describes some previous background estimations that were reevaluated with larger statistics as well as background studies on other K_L decay modes that were expected to have small contributions. Table 9.5 at the end of this section lists the updated background estimations with the new studies included.

9.4.1 $K^{\pm}$ Background

The mechanism for the $K^{\pm}$ background is that a $K^{\pm}$ can be generated when a K_L in the beam hits the edge of the second collimator[7]. If the $K^{\pm}$ enters the KOTO detectors and decays via $K^{\pm} \to \pi^0 e^{\pm} \nu_e$ and the electron is emitted with low enough energy to escape detection by the veto counters, this decay can mimic a signal event. A depiction of this background mechanism is shown in Figure 9.10. Other $K^{\pm}$ decays (such as $K^{\pm} \to \pi^0 \mu^{\pm} \nu_{\mu}$ and $K^{\pm} \to \pi^0 \pi^{\pm}$) were also estimated, but $K^{\pm} \to \pi^0 e^{\pm} \nu_e$ is the most likely source of background because the kinematics of the π^0 in this decay are most similar to that of the π^0 in the signal decay.

9.4.1.1 $K^{\pm}$ Flux Measurement

To accurately estimate the charged kaon background, a dedicated data collection run (Run 85) was performed in June of 2020 to estimate the $K^{\pm}$ flux at the beam exit. This was done so that the simulated MC background distribution of various $K^{\pm}$

[7]Shown in Figures 2.7 and 2.8

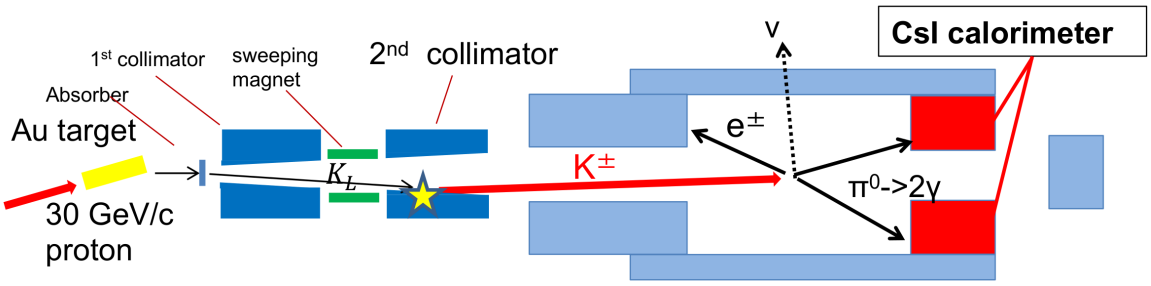


Figure 9.10: The $K^{\pm}$ background mechanism. A K_L in the beam hits the edge of the second collimator and a $K^{\pm}$ is generated. The $K^{\pm}$ enters the KOTO detectors and decays via $K^{\pm} \to \pi^0 e^{\pm} \nu_e$ and only the $e^{\pm}$ can be used for vetoing. If there is veto inefficiency or if the electron has low energy (particularly from moving backwards), then the π^0 decays into two photons and the missing momentum mimics a signal event.

decays could be normalized to the measured value of the $K^{\pm}$ flux, instead of relying soley on the MC simulation to estimate the $K^{\pm}$ contribution. A specialized trigger was used to collect $K^{\pm} \to \pi^0 \pi^{\pm}$ events (BR$\sim$ 20%) to evaluate the $K^{\pm}$ flux. This dedicated $\pi^0 \pi^{\pm}$ trigger selected events with three clusters on the CsI calorimeter, one coincidental hit in the CV from the $\pi^{\pm}$, and no other hits in any of the veto detectors.

To reconstruct $K^{\pm} \to \pi^0 \pi^{\pm}$ events, the CsI cluster closest to the extrapolated hit position in the CV was identified as the charged pion, and the other two clusters were identified as the two photons from the π^0 decay. The two photons were then used to reconstruct the decay vertex of the π^0 (assumed to be on the z axis) and since the π^0 travels a negligible distance before it decays, this is also the decay vertex of the $K^{\pm}$. From this, the transverse momentum of the π^0 was calculated ($P_T^{\pi^0}$) and the direction of the $\pi^{\pm}$ was obtained by using the decay vertex and the charged cluster position on the CsI. Then the absolute momentum of the $\pi^{\pm}$ ($|\vec{p}_{\pi^{\pm}}|$) was determined by assuming the transverse momentum of the $K^{\pm}$ was zero and using momentum conservation in the transverse plane as shown in Equation 9.11.

$$|\vec{p}_{\pi^{\pm}}| = \frac{P_T^{\pi^0}}{\sin \phi_{\pi^{\pm}}} \tag{9.11}$$

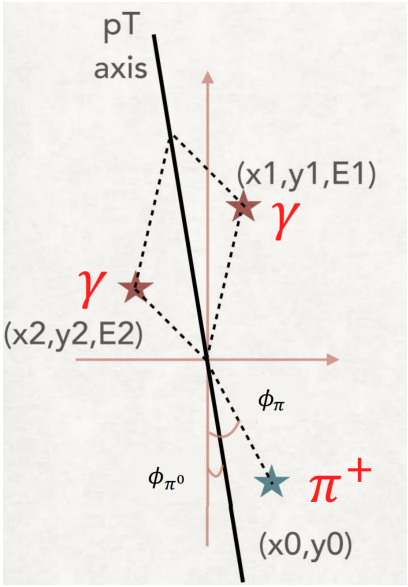


Figure 9.11: Using the hit positions of the particles in the $K^\pm \to \pi^0\pi^\pm$ reconstruction to calculate the momentum of the $\pi^\pm$.

The angle, $\phi_{\pi^\pm}$, is defined and shown as ϕ_π in Figure 9.11. Finally, after obtaining the four momentum for each photon (p_{γ_1}, p_{γ_2}) and the $\pi^\pm$ ($p_{\pi^\pm}$) from this information, the invariant mass of the $K^\pm$ was calculated in Equation 9.12.

$$M_{K^\pm} = \sqrt{(p_{\gamma_1} + p_{\gamma_2} + p_{\pi^\pm})^2} \tag{9.12}$$

To select $K^\pm \to \pi^0\pi^\pm$ events, we required the energy of the charged cluster ($E_{\pi^\pm}$) to be $200 < E_{\pi^\pm} < 400$ MeV in order to select a minimum-ionizing particle (MIP). We also required the reconstructed $K^\pm$ invariant mass to be $440 < M_{K^\pm} < 600$ MeV/c^2 and we required the same veto cuts as in the $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis except for the CV veto requirement. Figure 9.12 shows the reconstructed $K^\pm$ mass, the reconstructed π^0 z position, and the reconstructed $K^\pm$ momentum distributions for data and MC in which all of the cuts have been applied except for the cut on the variable being plotted. There is good agreement between the collected $K^\pm \to \pi^0\pi^\pm$ data and the MC simulation, and the $K^\pm \to \pi^0\pi^\pm$ events were able to be well distinguished from background events by using the reconstructed $K^\pm$ mass cut.

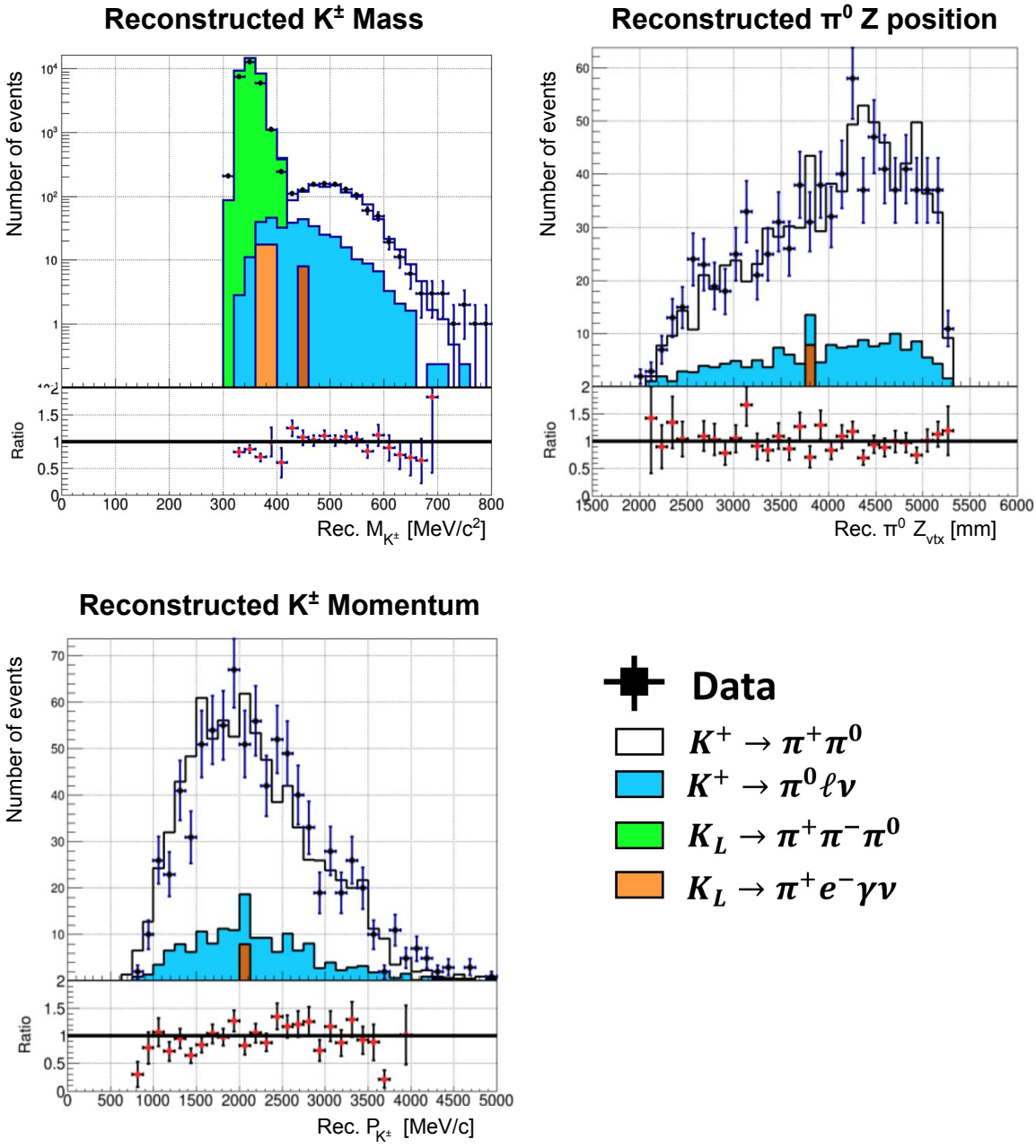


Figure 9.12: The reconstructed $K^{\pm}$ mass, the reconstructed π^0 z position, and the reconstructed $K^{\pm}$ momentum distributions for data and MC for the $K^{\pm} \to \pi^0\pi^{\pm}$ data collected in Run 85. The blue histogram represents $K^{\pm} \to \pi^0 e^{\pm}\nu_e$ and $K^{\pm} \to \pi^0\mu^{\pm}\nu_e$ decays. The data to MC ratio is shown in the bottom panel ("Ratio") and the error bars represent statistical error.

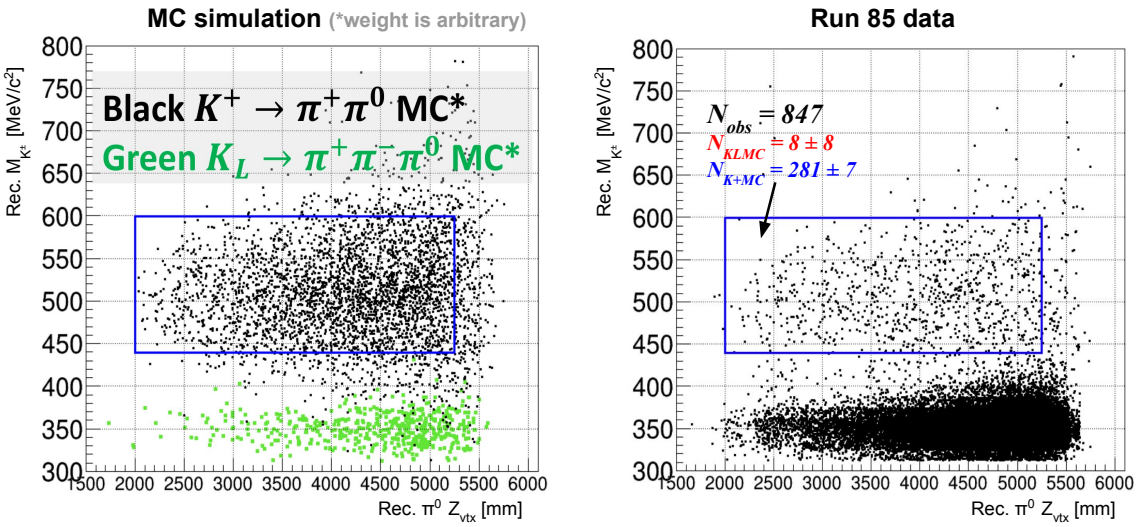


Figure 9.13: The reconstructed $K^\pm$ mass vs π^0 z vertex for Run 85 data (right) and MC simulation (left). The signal region to identify $K^\pm \to \pi^0\pi^\pm$ events is shown in blue and the data distribution was consistent with the MC simulation. There were 847 events observed in the signal region and this number was used to calculate the ratio of the $K^\pm$ to K_L flux.

To identify $K^\pm \to \pi^0\pi^\pm$ events and evaluate the K+ flux, a signal box was defined in the reconstructed $K^\pm$ mass vs π^0 z vertex position plane as shown in Figure 9.13. Events inside the box were used to calculate the $K^\pm$ flux as a ratio, comparing the $K^\pm$ flux with the K_L flux. Normalization data of $K_L^0 \to \pi^+\pi^-\pi^0$ decays were collected during Run 85 and the $K^\pm$ to K_L flux ratio was evaluated using a similar method as described in Section 7.4.3 which took into account the acceptance of the decays which was based on MC, the number of observed events after all cuts had been applied, the POT, and the prescale factor. Based on 847 $K^\pm \to \pi^0\pi^\pm$ candidate events as shown in the right of Figure 9.13, the ratio of the $K^\pm$ to K_L flux at the beam exit was measured to be

$$R_{K^\pm} = F_{K^\pm}/F_{K_L} = (2.6 \pm 0.1) \times 10^{-5} \tag{9.13}$$

where $F_{K^\pm}$ and F_{K_L} are the measured $K^\pm$ flux and K_L flux, respectively. The measured $K^\pm$ flux was around three times larger than what previous MC studies had predicted, which can be seen by comparing the number of observed events (N_{obs}) to

$K^{\pm}$ **Decay Mode**	**Estimated # of BG**
$K^{\pm} \to \pi^0 e^{\pm} \nu_e$	0.81 ± 0.13
$K^{\pm} \to \pi^0 \mu^{\pm} \nu_{\mu}$	0.02 ± 0.02
$K^{\pm} \to \pi^0 \pi^{\pm}$	0.004 ± 0.004
$K^{\pm} \to \mu^{\pm} \nu_{\mu}$	0
$K^{\pm} \to \pi^{\pm} \pi^{\pm} \pi^{\mp}$	0
$K^{\pm} \to \pi^0 \pi^0 \pi^{\pm}$	0
Total $K^{\pm}$ **BG**	0.84 ± 0.13

Table 9.4: The estimated number of background in the signal region for various $K^{\pm}$ decays. The decay modes $K^{\pm} \to \mu^{\pm} \nu_{\mu}$, $K^{\pm} \to \pi^{\pm} \pi^{\pm} \pi^{\mp}$, and $K^{\pm} \to \pi^0 \pi^0 \pi^{\pm}$ had no events remaining in MC after all of the cuts were applied, so these contributions are listed as zero.

the number of events predicted by the MC (N_{K+MC}) in Figure 9.13.

9.4.1.2 Background Estimation

To estimate the background contribution from $K^{\pm}$ decays, we used MC to simulate six major $K^{\pm}$ decays as shown in Table 9.4. The $K^{\pm} \to \pi^0 e^{\pm} \nu_e$ decay mode was expected to be the largest background contribution among these modes. The MC was scaled to the measured $K^{\pm}$ flux and all of the cuts for the 2016–2018 $K_L^0 \to \pi^0 \nu \bar{\nu}$ analysis were applied. The number of expected background in the signal region for each $K^{\pm}$ decay are shown in Table 9.4 and the PtZ plots of the $K^{\pm}$ background distributions are shown in Figure 9.14. The total number of background events from $K^{\pm}$ decays was estimated to be 0.84 ± 0.13, where 97.3% comes from $K^{\pm} \to \pi^0 e^{\pm} \nu_e$ decays, 2.3% comes from $K^{\pm} \to \pi^0 \mu^{\pm} \nu_{\mu}$ decays, and 0.4% comes from $K^{\pm} \to \pi^0 \pi^{\pm}$ decays. The other $K^{\pm}$ decay modes simulated had no events remaining in the MC after applying all cuts, and so their contribution was estimated to be zero.

Finally, the estimated $K^{\pm}$ background was corrected by factoring in the discrepancy in the acceptance between data and MC for the cuts used in the $K_L^0 \to \pi^0 \nu \bar{\nu}$ analysis against $K^{\pm}$ decays. To evaluate the $K^{\pm}$ background on the $K_L^0 \to \pi^0 \nu \bar{\nu}$ analysis, a data-driven method was used in which the $K^{\pm}$ yield was measured by

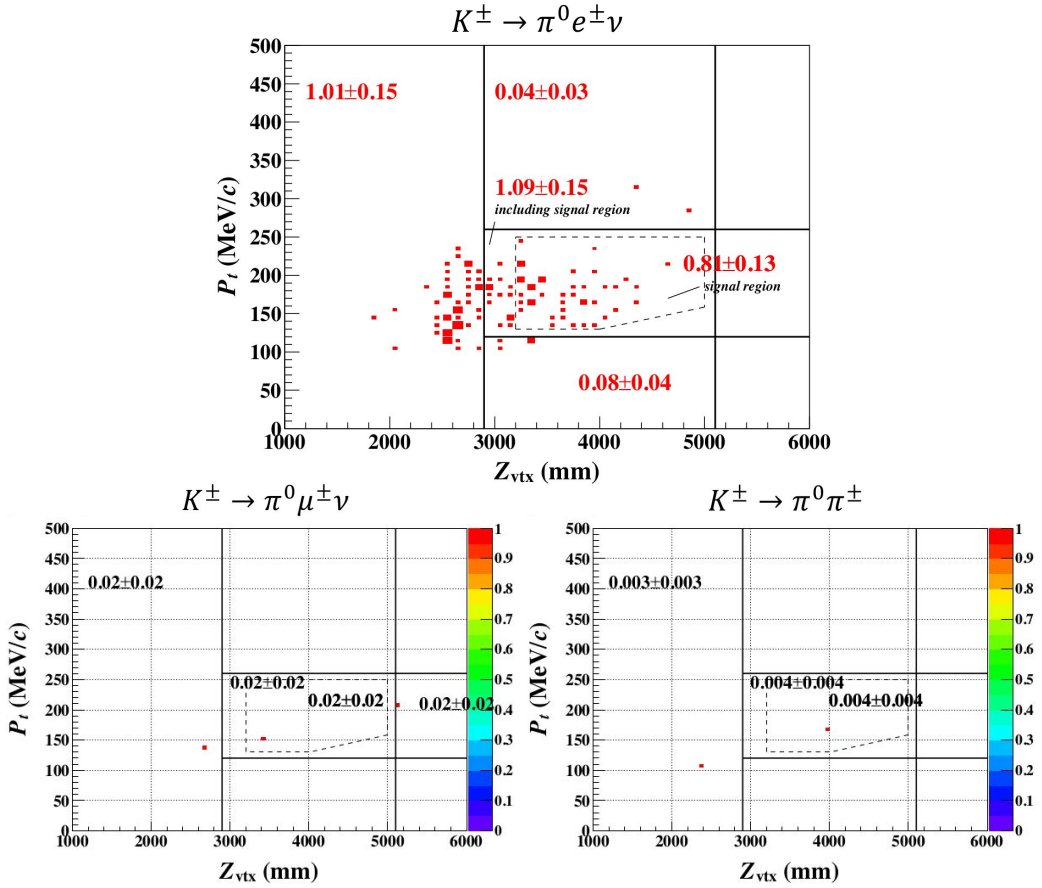


Figure 9.14: The estimated background contribution in each region of the P_T–Z plane for the $K^{\pm}$ decays. The top plot for $K^{\pm} \to \pi^0 e^{\pm} \nu_e$ has the largest contribution of 0.81 ± 0.13 events. Regions without values have an estimated contribution of zero.

collecting a control sample of $K^{\pm} \to \pi^0 \pi^{\pm}$ decays during Run 85. To increase the $K^{\pm}$ yield, the sweeping magnet in the KL beam line[8] was turned off and five hours of physics trigger data was collected. Along with the physics trigger, we simultaneously collected data with the $\pi^0 \pi^{\pm}$ trigger in the magnet-off configuration in order to normalize the $K^{\pm}$ yield. After imposing the $K_L^0 \to \pi^0 \nu \bar{\nu}$ selection criteria on the control sample, we observed 27 events in the blinded region which was consistent with the 26.0 ± 3.2 events expected in MC as shown in Figure 9.15. The ratio of these two numbers was taken to compare the $K^{\pm}$ acceptance of the data and the MC and was

[8]Described in Section 2.2.4 and shown in Figures 2.7 and 2.8

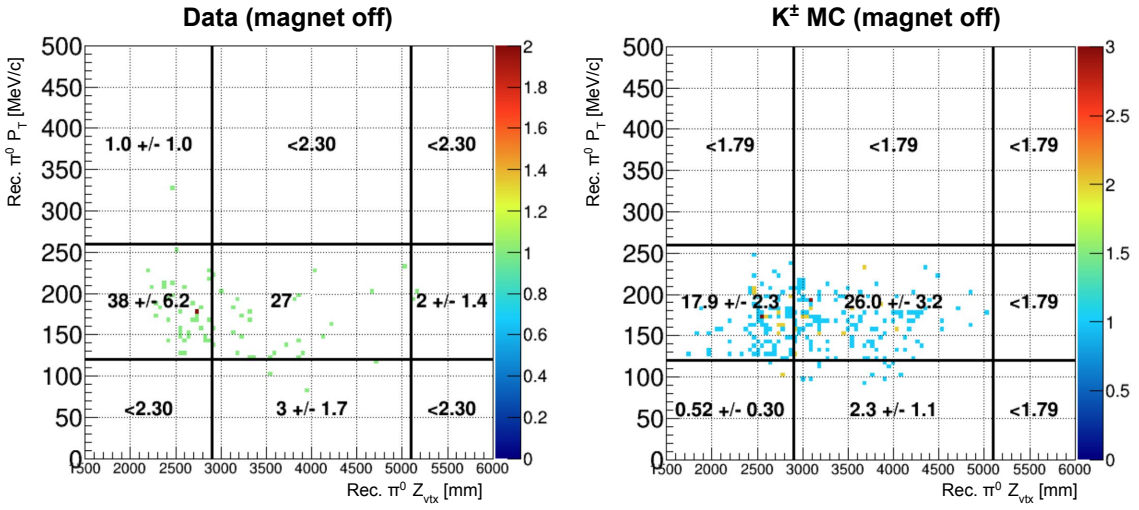


Figure 9.15: PtZ plots of the $K^{\pm}$ control sample data (left) and MC (right) with the sweeping magnet turned off and all cuts used in the $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis applied. The number of events in the blinded region for data (27) and MC (26.0 ± 3.2) were consistent.

determined to be

$$R_{A_{K^{\pm}}} = A_{\text{data}}^{K^{\pm}} / A_{\text{MC}}^{K^{\pm}} = 1.04 \pm 0.26 \tag{9.14}$$

where $A_{\text{data}}^{K^{\pm}}$ and $A_{\text{MC}}^{K^{\pm}}$ are the acceptances of the control sample data and MC, respectively. The uncertainty on this value comes from the $K^{\pm}$ spectrum difference between the magnet on and off configurations, as well as statistical uncertainties. By multiplying the total number of $K^{\pm}$ background (0.84 ± 0.13) by the acceptance ratio (1.04 ± 0.26), the corrected, total number of $K^{\pm}$ background was 0.87 ± 0.25. It should be noted that the $K^{\pm}$ background is the largest background contribution and because of this, a prototype of a new detector called the Upstream Charged Veto (UCV) was designed, built, and installed in Run 85 to test its performance on suppressing the $K^{\pm}$ background. More details on the UCV detector and other ways to reduce the $K^{\pm}$ background are covered in Section 10.3.1.

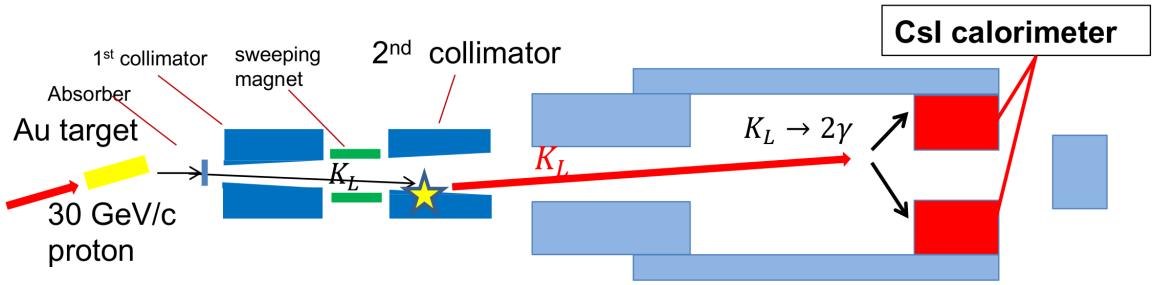


Figure 9.16: The halo K_L background mechanism occurs when a K_L in the beam scatters off the edge of the second collimator and decays via $K_L^0 \to 2\gamma$. Since the scattered K_L has a finite transverse momentum and no other final state particles to veto with, it can mimic a signal event.

9.4.2 Halo K_L Background

Another background that was not previously considered occurs when a K_L in the beam scatters off the second collimator and then that scattered K_L decays to two photons via $K_L^0 \to 2\gamma$. The mechanism for this background is depicted in Figure 9.16. This background is particularly difficult to reduce because while the branching ratio of $K_L^0 \to 2\gamma$ is only $\sim 5 \times 10^{-4}$, there are no extra particles in the final state with which to veto, so if the K_L has any transverse momentum at all, it can be mistaken for a signal event. It should be noted that the halo K_L background is different from the $K_L^0 \to 2\gamma$ background discussed in Section 8.3.3, in which a K_L in the beam scatters at the vacuum window.

The halo K_L background was calculated by estimating the halo K_L yield using a sample of $K_L^0 \to 3\pi^0$ events that was collected in the 2016–2018 run periods. The $K_L^0 \to 3\pi^0$ events were selected to have large COE radius values (indicating they came from halo K_Ls) and the ratio between data and MC for this sample was evaluated. The data/MC ratio was then used to scale the estimated number of halo K_L background events in MC to the collected halo K_L data by multiplying by this factor. After this correction, the halo K_L background in the signal region was estimated to be 0.26 ± 0.07 events as shown in Figure 9.17.

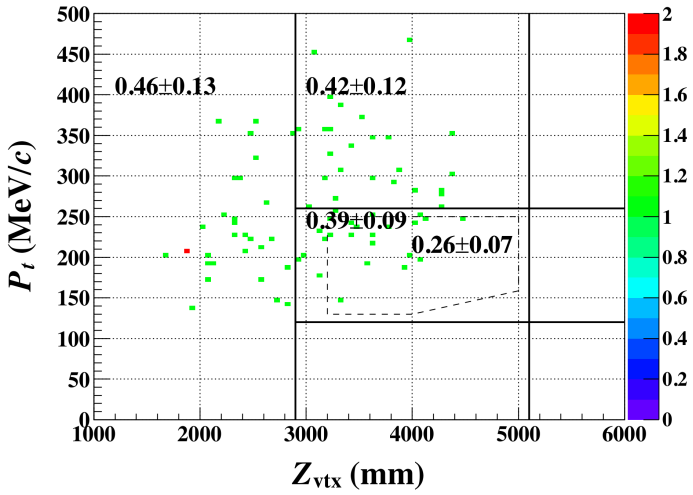


Figure 9.17: The estimated background contribution in each region of the P_T–Z plane for the halo K_L background. Regions without values have an estimated contribution of zero.

9.4.3 Other Backgrounds and Updated Background Estimations

In order to ensure a thorough and accurate estimation of backgrounds, several additions and careful checks were carried out. In double checking the background study for the upstream π^0 background, an error was found in the calibration factors for the MBCV veto detector. After updating the cut threshold from 0.5 to 1.0 MeV to keep consistency between data and MC, a background event appeared in the signal region. For this reason, the upstream π^0 background estimation was updated from 0.001 $\pm$ 0.001 to 0.03 $\pm$ 0.03. The estimation for the $K_L^0 \to \pi^{\pm}e^{\mp}\nu_e$ masking background was previously estimated at < 0.09 and after finding an incorrect parameter for the timing and correcting it, this estimation was updated to < 0.08.

We reevaluated the background estimations for the $K_L^0 \to 2\pi^0$ background and the $K_L^0 \to 3\pi^0$ masking background with higher statistics. The $K_L^0 \to 2\pi^0$ MC sample was increased by a factor of two and the $K_L^0 \to 3\pi^0$ MC seed sample was increased while the accidental overlay recycling was decreased for a more accurate sample. This caused the background estimations for $K_L^0 \to 2\pi^0$ to be updated from < 0.18 to < 0.08 and for $K_L^0 \to 3\pi^0$ to be updated from < 0.04 to a central value of 0.01 $\pm$ 0.01.

BG Source	Estimated # of BG	Note
K_L Decay Background		
$K_L^0 \to \pi^+\pi^-\pi^0$	< 0.02	
$K_L^0 \to 2\pi^0$	< 0.08	updated, increased MC stat.
$K_L^0 \to 2\gamma$ (vacuum window)	0.005 ± 0.005	
$K_L^0 \to 3\pi^0$ (masking)	0.01 ± 0.01	updated, increased MC stat.
$K_L^0 \to \pi^\pm e^\mp \nu_e$ (masking)	< 0.08	updated, 5% diff. in timing
$K_L^0 \to \pi^\pm e^\mp \nu_e$ ($\pi^\pm \to \pi^0$ conversion)	< 0.04	estimated after unblinding
$K_L^0 \to \pi^\pm e^\mp \nu_e$ ($\pi^\pm$ beta decay)	< 0.01	estimated after unblinding
$K_L^0 \to \pi^\pm e^\mp \nu_e \gamma$	< 0.05	estimated after unblinding
$K_L^0 \to \pi^0\pi^\pm e^\mp \nu_e$	< 0.04	estimated after unblinding
$K_L^0 \to \pi^+\pi^-$	< 0.03	estimated after unblinding
$K_L^0 \to ee\gamma$	< 0.09	estimated after unblinding
$K_L^0 \to K^\pm e^\mp \nu_e$	< 0.04	estimated after unblinding
$K_L^0 \to 2\gamma$ (core-like)	< 0.11	estimated after unblinding
$K_L^0 \to 2\gamma$ (halo-K_L)	0.26 ± 0.07	estimated after unblinding
$K^\pm$ Decay Background		
$K^\pm \to \pi^0 e^\pm \nu_e$	0.84 ± 0.25	estimated after unblinding
$K^\pm \to \pi^0 \mu^\pm \nu_\mu$	0.02 ± 0.02	estimated after unblinding
$K^\pm \to \pi^0\pi^\pm$	0.004 ± 0.004	estimated after unblinding
Total $K^\pm$ BG	**0.87 ± 0.25**	
Neutron-Induced Background		
Hadron-Cluster	0.017 ± 0.002	
Upstream π^0	0.03 ± 0.03	updated, wrong veto threshold
CV-π^0	< 0.10	
CV-η	0.03 ± 0.01	
Total	**1.22 ± 0.26**	

Table 9.5: The updated summary table of the estimated background contribution in the signal region for each background source. The largest sources that contribute to the total are the $K^\pm$ background and the halo K_L background. The different $K_L^0 \to 2\gamma$ background sources result from 1) K_Ls scattering at the vacuum window (as described in Section 8.3.3), 2) K_Ls that are at the edge of the beam core (core-like) and so have a small amount of P_T but usually not enough to exceed the COE radius threshold, and 3) halo K_Ls that scatter off the edge of the second collimator as described in Section 9.4.2. The backgrounds estimated with an upper limit were estimated at the 90% CL.

In addition to these updates on previously estimated backgrounds, studies were done on other K_L decay modes that may contribute background in the signal region. These other decay modes that were newly considered are listed in Table 9.5 and they all had upper limits set with a 90% CL. In particular, these backgrounds

were estimated with MC simulation and studied by loosening various veto detector thresholds. Since they all have charged particles in the final state, as long as the CV detector and other veto detectors such as CC04 and new BHCV continue to have a high efficiency in rejecting events with charged particles, these backgrounds should remain reasonably small and well-suppressed.

Before opening the box the background estimation in the signal region was 0.05 $\pm$ 0.02. After three events were observed in the signal region, the background estimations were reevaluated and two new backgrounds were considered– the halo K_L background which was estimated to be 0.26 $\pm$ 0.07 and the $K^{\pm}$ background which was estimated to be 0.87 $\pm$ 0.25. Various other backgrounds were updated and newly studied and the total background estimation in the signal region was updated to be 1.22 $\pm$ 0.26 by adding the central values of each background source in Table 9.5.

9.5 Final Results

The results of the 2016–2018 analysis are summarized here. With the 2016–2018 data set, we achieved a SES of $(7.20 \pm 0.05_{\text{stat}} \pm 0.66_{\text{syst}}) \times 10^{-10}$ which is the best sensitivity for the $K_L^0 \to \pi^0\nu\bar{\nu}$ search to date. The box was opened and three candidate events were observed in the signal region. One of the events (event 0) was found to have overlapped pulses in the NCC detector, and so is likely an accidental masking background. Based on the estimation for the masking background in the signal region (0.01 $\pm$ 0.01), the probability of observing this kind of event is 2.2%, which was determined from the p-value calculated using a Poisson distribution. The total number of background events expected in the signal region was estimated to be 1.22 $\pm$ 0.26 with the largest background contribution coming from $K^{\pm}$ decays. The probability of observing three events in the signal region is 13%. Thus, the number of observed events is statistically consistent with the background expectation.

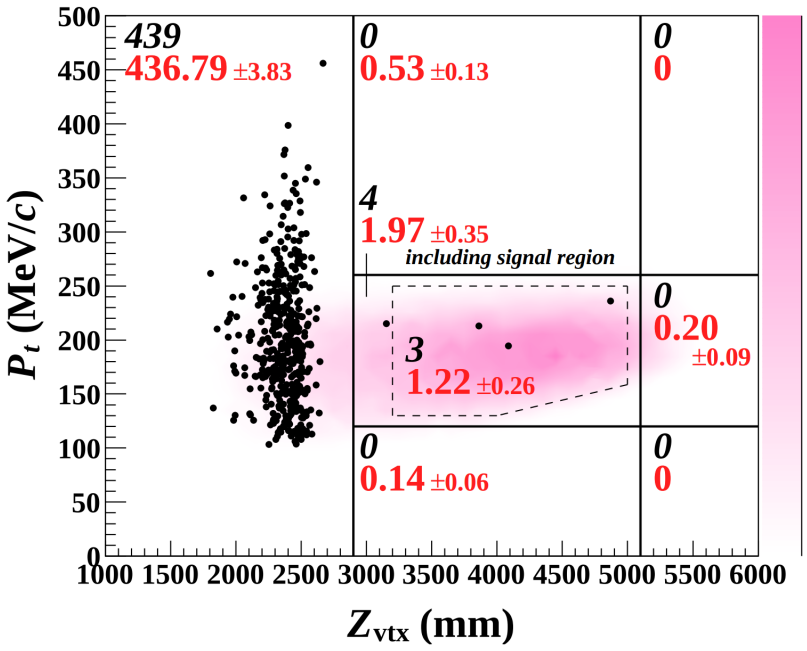


Figure 9.18: The open box results of the 2016–2018 $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis. The expected number of background events (red) and the number of observed events (black) are shown in each region of the P_T–Z plane, and the MC signal distribution is shown in pink. Three events were observed in the signal region with a background expectation of 1.22 ± 0.26.

Figure 9.18 shows the final, unblinded PtZ plot for the 2016–2018 data set with the expected number of background events (red), the number of observed events (black), and the MC signal distribution (pink) shown in each region. The background distributions vs the observed candidate events are also shown in Figure 9.19 for the P_T and z distributions within the signal region. In particular, the red histogram represents the expected number of signal events at this sensitivity level based on the SM BR of $K_L^0 \to \pi^0\nu\bar{\nu}$ (that is, 3×10^{-11} / $7.2 \times 10^{-10} \sim 0.04$ SM events) and the black data points are the candidate events.

Thus, an upper limit (UL) on the branching ratio of $K_L^0 \to \pi^0\nu\bar{\nu}$ was calculated assuming Poisson statistics and considering uncertainties, using Equation 9.15 at the 90% CL [69].

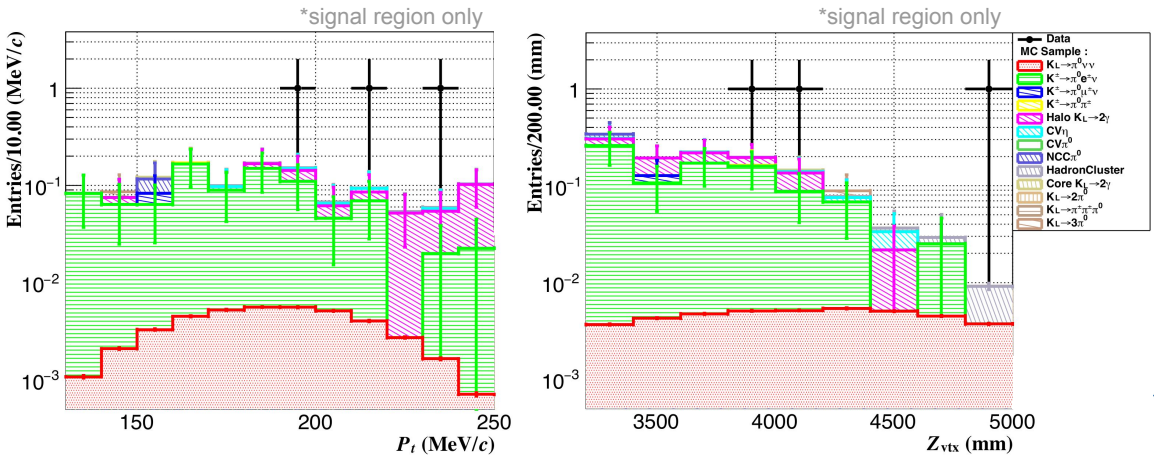


Figure 9.19: The P_T (left) and z (right) distributions in the signal region for the expected background and the observed candidate events. The red histogram represents the expected signal distribution at this sensitivity level for the SM predicted BR of $K_L^0 \to \pi^0\nu\bar{\nu}$. The black data points represent the observed candidate events. Background distributions are shown in the other colored histograms (green = $K^\pm$ background, pink = halo K_L background, cyan = CV-η background, etc.).

$$\mathrm{BR}(K_L^0 \to \pi^0\nu\bar{\nu})\ \mathrm{UL} = 6.68 \times \left(1 + \frac{1 - \sqrt{1 - 6.68^2 \times \sigma_r^2}}{6.68}\right) \times \mathrm{SES} \tag{9.15}$$

where σ_r is the total uncertainty, which was 9.2% as described in Section 9.1.9. Thus, the upper limit on the BR for the 2016–2018 data set is

$$\mathrm{BR}(K_L^0 \to \pi^0\nu\bar{\nu}) < 4.9 \times 10^{-9} \tag{9.16}$$

This upper limit is on the same order of magnitude as the previous BR upper limit set by KOTO, which was $< 3.0 \times 10^{-9}$ [36]. The discussion and conclusions on these results as well as the future prospects for the KOTO experiment and the $K_L^0 \to \pi^0\nu\bar{\nu}$ search are covered in Chapter 10.

Chapter 10

Discussion and Conclusions

After opening the box and observing four events in the signal region when the estimated number of background was 0.05 ± 0.02, these results were surprising. Even after updating the number of candidate events to three (one event remained due to an error) and the estimated background to 1.22 ± 0.26, the number of candidate events observed at this sensitivity level is unexpected. This is perhaps best illustrated in Figure 9.19 where the number of expected SM $K_L^0 \to \pi^0\nu\bar{\nu}$ events at a SES of 7.2×10^{-10} is around 0.04.

It should also be noted that in the 2015 data set with a SES of 1.3×10^{-9}, the estimated background was 0.42 ± 0.18 and no candidate events were observed in the signal region [36]. While the probability of observing three events in the signal region with a background expectation of 1.22 ± 0.26 is 13% and thus statistically consistent, a conclusion with absolute certainty cannot be drawn about the 2016–2018 results. However, possible explanations for these results are explored in this chapter and the impact of these results are discussed in Section 10.2. In addition, the upgrades and future prospects for the KOTO experiment and the $K_L^0 \to \pi^0\nu\bar{\nu}$ search are covered in the final sections.

10.1 Discussion of Results

The details of the properties of the candidate events are covered in Section 9.3.1. In summary, one of the three candidate events (event 0) is likely due to the masking background as it had overlapped pulses in the NCC detector, though the other two candidate events cannot be confirmed or denied to be background or signal. The possibilities are either that the background in the signal region was underestimated, that the observed candidate events are actual signal events and the branching ratio of $K_L^0 \to \pi^0\nu\bar{\nu}$ is higher than the SM prediction ($\sim 3 \times 10^{-11}$), or that some new physics process involving a new particle or particles was observed. The latter two cases involving physics beyond the SM are explored in Section 10.1.1. Based on Equation 9.1, if all three candidate events are taken to be signal events, the corresponding branching ratio can be calculated to be $\mathrm{BR}(K_L^0 \to \pi^0\nu\bar{\nu}) = 2.16 \times 10^{-9}$, if two events are taken to be signal events then $\mathrm{BR}(K_L^0 \to \pi^0\nu\bar{\nu}) = 1.44 \times 10^{-9}$, and if one event is taken to be signal then $\mathrm{BR}(K_L^0 \to \pi^0\nu\bar{\nu}) = 7.2 \times 10^{-10}$.

It is also important to note here that the NA62 experiment reports recent results on the BR limit of the charged decay, $K^+ \to \pi^+\nu\bar{\nu}$ to be $< 1.78 \times 10^{-10}$ at the 90% CL [70]. This translates to an updated Grossman-Nir bound[1] of $\mathrm{BR}(K_L^0 \to \pi^0\nu\bar{\nu}) < 7.8 \times 10^{-10}$ and a measurement of $\mathrm{BR}(K^+ \to \pi^+\nu\bar{\nu}) = (0.48\ ^{+0.72}_{-0.48}) \times 10^{-10}$ at the 68% CL which is consistent with the E949 previous experimental measurement and the SM prediction. Thus, the conflicting results between the NA62 experiment which is consistent with the SM, and KOTO's results that may indicate a deviation from the SM, drive the exploration of new physics theories.

[1] Discussed in Section 1.6, the previous Grossman-Nir bound was reported to be 1.5×10^{-9} based on the E949 experiment's measurement of $K^+ \to \pi^+\nu\bar{\nu}$ [18].

10.1.1 Possible Interpretations of Candidate Events

Several possible explanations for the candidate events observed in the 2016–2018 $K_L^0 \to \pi^0\nu\bar{\nu}$ analysis are explored in the following subsections. While many beyond the SM theories are presented in this section, a strong emphasis is placed on the need for more experimental data and subsequent studies in order to understand the true nature of the three observed candidate events.

10.1.1.1 Violating the Grossman-Nir Bound

Considering all three events to be signal events is unlikely, as one candidate event is likely due to the masking background, but also because a BR of 2.16×10^{-9} conflicts with the Grossman-Nir bound set by the E949 experiment and the NA62 experiment. Still, there are a couple of BSM theories that violate the Grossman-Nir bound and allow for the BR of $K_L^0 \to \pi^0\nu\bar{\nu}$ to be higher than this limit. In one case, the Grossman-Nir bound can be violated with the inclusion of isospin $\Delta I = 3/2$ operators as explored in [71] and [72] and as a result, this could point to the existence of an invisible, light scalar particle. Another BSM theory shows that the Grossman-Nir bound can be violated in the presence of light new physics with flavor violating couplings, and consequently these various new physics models require at least one light mediator around the GeV mass scale [73].

10.1.1.2 New Light Particles

Several theories to interpret "the KOTO anomoly" as it has been referred to, call for the existence of new, lighter particles that would explain the three candidate events that KOTO observed without violating the Grossman-Nir bound. One paper [74] assuming a BR of 2.1×10^{-9}, explores the possibility of a new, light, long-lived particle, X, in which the decay of the kaon proceeds as $K_L^0 \to \pi^0 X$ and then the invisible particle X escapes the KOTO detectors before decaying into two photons,

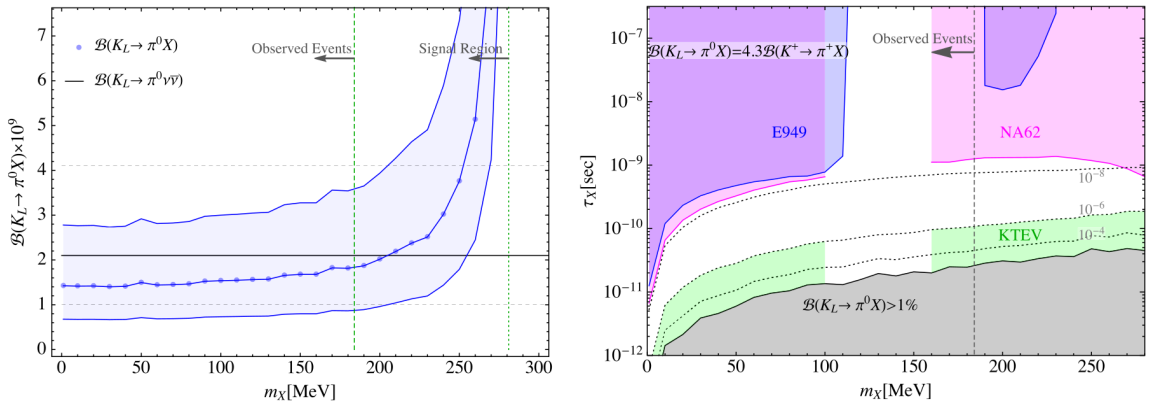


Figure 10.1: Constraints on the branching ratio of $K_L^0 \to \pi^0 X$ (left) and the lifetime of X (right) that can accommodate the KOTO events, the Grossman-Nir bound, and the $K^+ \to \pi^+ \nu\bar{\nu}$ experimental results. In the left figure, the dotted blue (solid grey) line represents the central value of the $K_L^0 \to \pi^0 X(\pi^0\nu\bar{\nu})$ interpretation with the blue shaded band (dashed horizontal lines) for the two-sided 68% confidence interval. In the right figure, The $K^+ \to \pi^+ X$ bound is translated to the K_L bound assuming a saturation of the Grossman-Nir bound. The pink (blue) shaded region is constrained by NA62 at 95% CL (E949 at 90% CL). Too short of a lifetime leads to $\mathrm{BR}(K_L^0 \to \pi^0 X) > 1\%$, which is inconsistent with the untagged K_L branching ratio. The $\mathrm{BR}(K_L^0 \to \pi^0 X) = 10^{-4}$, 10^{-6}, and 10^{-8} are indicated on the plot. The green shaded region is constrained from the KTeV search for $K_L^0 \to \pi^0\gamma\gamma$ assuming $\mathrm{BR}(X \to \gamma\gamma) = 1$. Figure courtesy of [74].

for example. The constraints on the BR of $K_L^0 \to \pi^0 X$ and the lifetime of X based on theoretical models and experimental constraints are shown in Figure 10.1. Based on these constraints, the possible underlying models for the existence of X point to a Higgs portal that induces $K_L^0 \to \pi^0 X$ in which X is a possible leptophobic and/or photophilic scalar [74].

Another paper [75] theorizes weakly-coupled, light scalars (ϕ) produced in kaon decays (specifically, $K_L^0 \to \pi^0\phi$) that escape KOTO undetected due to their long-lived and weakly-interacting nature. Two possible models that agree with current experimental bounds for this new particle are the minimal Higgs portal (the real-scalar singlet extension of the SM) [76] and a hadrophilic scalar model with flavor-aligned, generation-specific couplings to up-type quarks. Further possibility for the observed events can be explained by an invisible gauge-boson, X^0, with a mass around the

mass of the π^0 in which $K_L^0 \to \pi^0 X^0$ [77; 78]. This process could occur at tree-level (as opposed to FCNCs which are forbidden at tree-level) and account for events seen at a higher sensitivity than the SM prediction for $K_L^0 \to \pi^0 \nu\bar{\nu}$. Many other papers [79; 80; 81; 82; 83] also theorize the existence of a light particle that could explain the KOTO results.

10.1.1.3 Dark Sectors

There is also potential for new physics from dark sectors to explain the observed events at this sensitivity level above the Standard Model. The interactions between possible dark sector particles and the SM particles are thought to be weak, indirect and typically mediated through gravity and so would occur infrequently. Because KOTO probes the SM at such a high sensitivity, it is possible that particles from a dark sector could be observed in KOTO if they exist. One paper [84] proposes a dark scalar to explain the observed KOTO events through the process $K_L^0 \to \pi^0 \phi'$. In this model, ϕ' is a dark scalar that arises in a dark neutrino sector with a hidden $U(1)'$ gauge symmetry that is broken by the vacuum expectation value of ϕ'.

Two papers propose the existence of a light scalar that arises from different dark sector models. In the first [85], a light scalar singlet, S, couples to SM fermions via Yukawa couplings and to photons via higher-dimensional coupling. The KOTO events can be explained by a flavor-changing penguin diagram process, $K_L^0 \to \pi^0 S$ followed by the decay of S to neutrinos. In the second [86], a dark sector model has appropriate strange flavor symmetry and an additional light, flavor-like, complex scalar field, ϕ which is only accessible to K_L. Thus, this model leads to a strong violation of the Grossman-Nir bound.

Another possibility is a light, dark fermion, Q, in which the process $K_L^0 \to \pi^0 Q\bar{Q}$ could explain the events observed in the signal region [87]. Pair production of dark particles (X_1, X_2) in K_L decays has also been proposed in which $K_L^0 \to X_1 X_2$ or $X_1 X_2$

[88]. If either of these two dark states is unstable, they could mimic the $K_L^0 \to \pi^0\nu\bar{\nu}$ signature. Two papers also propose the existence of an axion-like particle (ALP) in which either the new ALP particle is long-lived and produced at the T1 target ($\phi \to \gamma\gamma$), travels to the KOTO detector, and decays off-axis [74] or the ALP (a) is produced by the decay $K_L^0 \to \pi^0 a$ and decays to two photons in which case KOTO could search for this process (four photons) and constrain this dark sector model [86]. The authors of [74] also explore the possibility of new heavy physics boosting the SM signal to explain the KOTO results.

10.1.1.4 Other BSM Explanations

It should also be noted that other BSM theories have been explored, such as extensions and revisions [89; 90] to the Z' model [26] mentioned in Section 1.7, an extended Higgs model with sterile neutrinos [91], pionium ($\pi^+\pi^-$ atom) as a source of false events in KOTO [92], generic neutrino interactions in effective field theories [93], constraints on lepton number violating interactions with $K_L^0 \to \pi^0\nu\bar{\nu}$ [94], and new particles from supersymmetry [95] to explain the observed events. The large number of theory papers (over 30) that have been published to offer explanations of the KOTO 2016–2018 results indicate the importance of these results as well as the importance of the $K_L^0 \to \pi^0\nu\bar{\nu}$ search.

10.2 Conclusions

The KOTO experiment collected data from 2016 to 2018 to search for the $K_L^0 \to \pi^0\nu\bar{\nu}$ decay. The total accumulated POT for this data set was 3.05×10^{19} POT which corresponded to $(6.83 \pm 0.05) \times 10^{12}$ K_L at the beam exit. The SES achieved with this data set was the best sensitivity reached for the $K_L^0 \to \pi^0\nu\bar{\nu}$ search to date at $(7.20 \pm 0.05_{\text{stat}} \pm 0.66_{\text{syst}}) \times 10^{-10}$. Before unblinding the data in the signal region, the total expected background was estimated to be 0.05 ± 0.02. After opening

the box, four candidate events were observed in the signal region and after further investigation, one event was removed due to an incorrect timing parameter setting. Out of the three remaining events in the signal box, one event has overlapped pulses in the NCC detector and the probability of observing this kind of event is 2.2%.

Two new backgrounds were considered, the $K^{\pm}$ background (Section 9.4.1) and the halo K_L background (Section 9.4.2), and the background estimation in the signal region was updated to be 1.22 ± 0.26 with the largest background contribution coming from $K^{\pm}$ decays. Based on the total estimated background and the SES, the probability of observing three events in the signal region is 13%. We conclude that the number of candidate events observed is statistically consistent with the background expectation. Finally, with the 2016–2018 data set we set an upper limit on the branching ratio of $K_L^0 \to \pi^0\nu\bar{\nu}$ to be 4.9×10^{-9} which is on the same order of magnitude as the limit set with data collected in 2015 [36].

The impact of these results are shown in Figure 10.2 and are also illustrated by the large number of theoretical papers generated to offer possible explanations of the observed candidate events. Though one candidate event seems to be due to the masking background, a conclusion about the candidate events observed well above the SM expectation cannot be drawn at this time. As Carl Sagan noted, "Science demands a tolerance for ambiguity. Where we are ignorant, we withhold belief. Whatever annoyance the uncertainty engenders serves a higher purpose: It drives us to accumulate better data." [96]. This is the most important conclusion to be drawn– that it is essential to collect more data and higher statistics which will improve the systematic uncertainties and allow us to understand the true nature of the observed candidate events in the 2016–2018 analysis. Furthermore, the identification of two new background sources is an important step for KOTO to reach SM sensitivity and improvements to background reduction methods and detector upgrades are necessary for the continued $K_L^0 \to \pi^0\nu\bar{\nu}$ search. The upgrades and improvements made to the

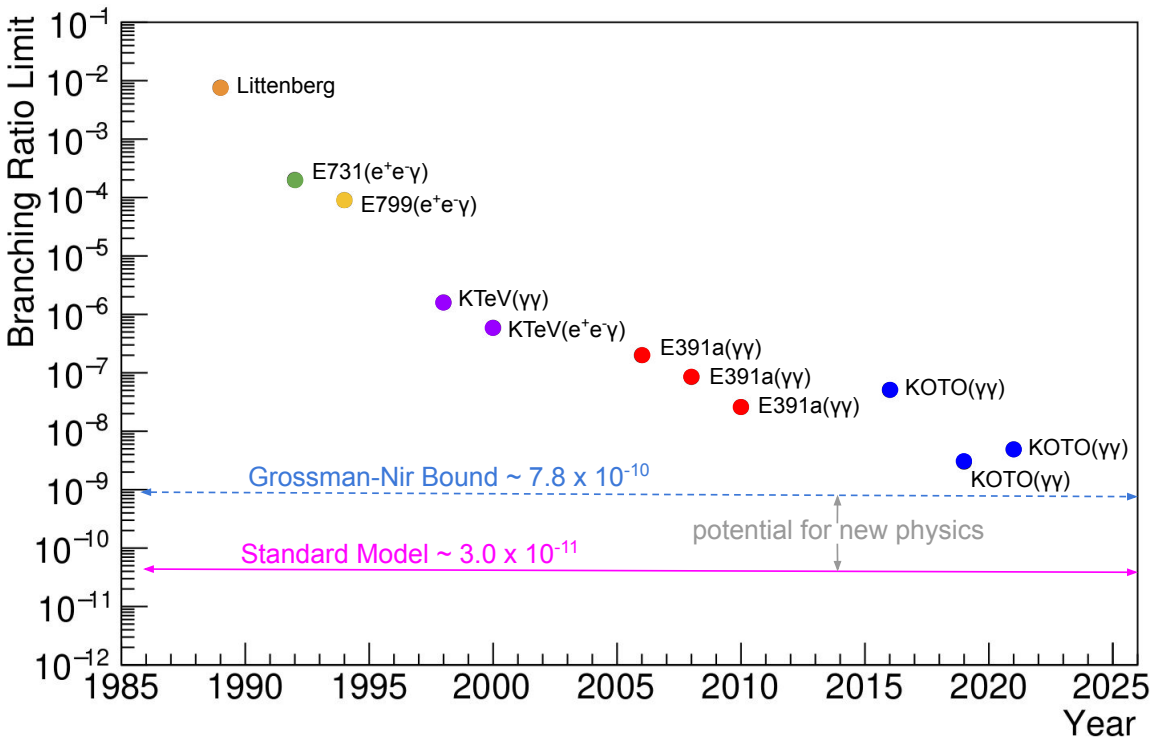


Figure 10.2: The updated search history of $K_L^0 \to \pi^0\nu\bar{\nu}$. The branching ratio upper limit is shown for various experiments and the KOTO experiment set an upper limit of 4.9×10^{-9} with the 2016–2018 data which is on the same order of magnitude as the best experimental upper limit set by KOTO previously. The Grossman-Nir bound has been updated to reflect the new results from the NA62 experiment [70].

KOTO experiment after the 2016–2018 data was collected are covered in Section 10.3.

In summary, the purpose of this thesis was to report on the search for $K_L^0 \to \pi^0\nu\bar{\nu}$ with the data collected from 2016–2018 and the objectives laid out in Chapter 1 have been completed. The importance of the $K_L^0 \to \pi^0\nu\bar{\nu}$ search was covered in Chapter 1 and the experimental methods and data collection of the 2016–2018 data were detailed in Chapters 2–4. The analysis strategy was explained in Chapter 5 and carried out in Chapters 6–9 where the total number of K_L was calculated (Chapter 7), the background was estimated (Chapters 8 and 9), and the Single Event Sensitivity was determined (Chapter 9). Finally, an upper limit on the branching ratio of $K_L^0 \to \pi^0\nu\bar{\nu}$ using the 2016–2018 data set was reported.

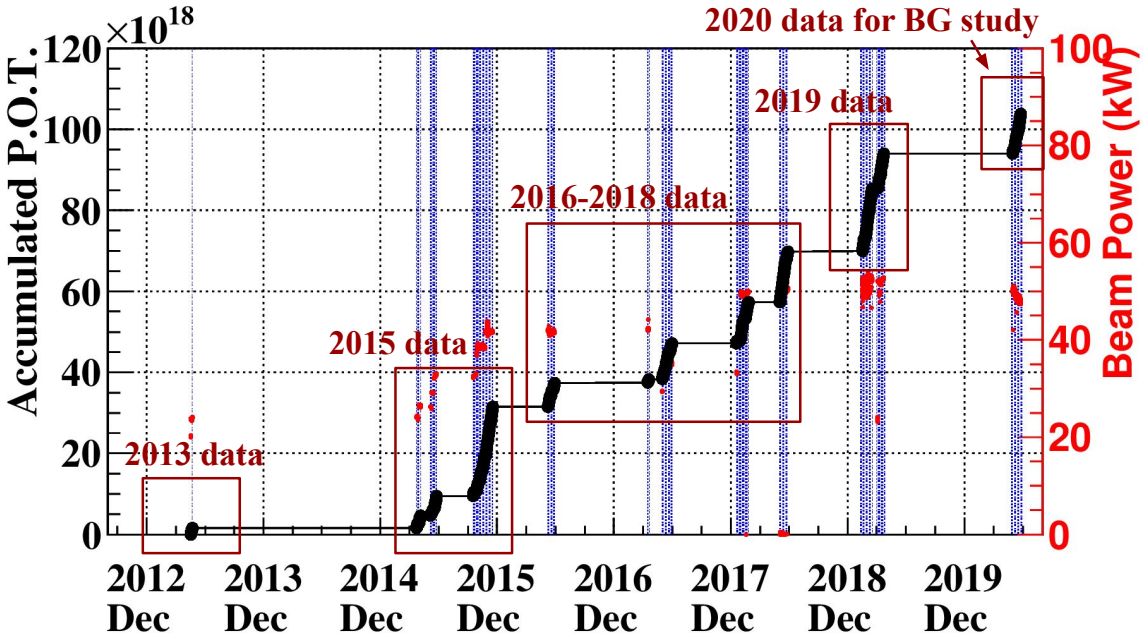


Figure 10.3: The accumulated POT (black) and beam power (red) over time shows KOTO's data collection history. The 2019 data set is currently being analyzed and the 2020 data was used for the $K^{\pm}$ background estimation.

10.3 Outlook and Future Work

Since 2018 the KOTO experiment has continued to collect data periodically from year to year. In February and March of 2019 KOTO had two runs (Run 81 and 82) and collected around 2.2×10^{19} POT. KOTO also collected data in June of 2020 (Run 85) to gather a control sample of $K^{\pm}$ events for the $K^{\pm}$ background estimation mentioned in Section 9.4.1 and collected a small sample of physics trigger data corresponding to 0.7×10^{19} POT. These additional data sets can be seen in Figure 10.3 along with the rest of KOTO's data collection history. The 2019 data set is currently being analyzed and has a preliminary SES estimation of 1.02×10^{-9}. KOTO is also collecting data currently (February/March 2021) and is estimated to collect around one month of physics data before the J-PARC accelerator goes into a shutdown period for the remainder of the year for an upgrade of the main ring. The

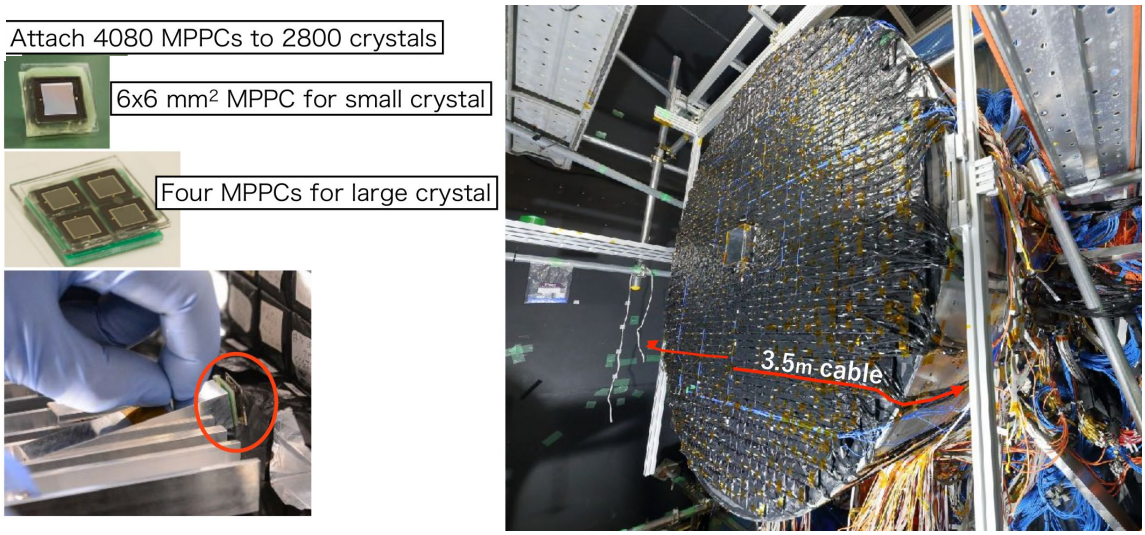


Figure 10.4: Installation photos (Dec. 2018) of MPPCs on the upstream end of the CsI calorimeter crystals. The MPPCs were attached to the upstream end of the CsI crystals for dual-ended readout.

upgrades and improvements since 2018 for the detectors and the DAQ system as well as new background suppression methods are discussed in the following subsections.

10.3.1 Detector and Accelerator Upgrades

After the 2018 runs, the calorimeter was upgraded to have dual-ended readout of the CsI crystals to further suppress the hadron-cluster background and improve discrimination between neutrons and photons. This was accomplished by attaching MPPCs (Multi-Pixel Photon Counters) to the upstream end of each of the CsI crystals. Figure 10.4 shows photos of the MPPC installation on the calorimeter. For the small crystals (2.5×2.5 cm^2) one MPPC (6×6 mm^2) was attached to each crystal and for the large crystals (5×5 cm^2) four MPPCs were attached to each crystal. In total 4080 MPPCs were attached to 2800 crystals and were used to read out signals from the CsI calorimeter.

Because neutrons tend to have deeper interactions than photons, there is depth information available by measuring the timing difference between the MPPCs and the

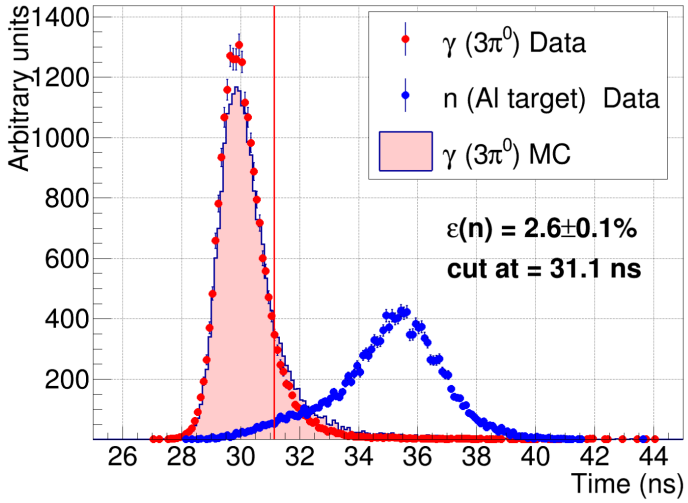


Figure 10.5: Distribution of the timing difference seen for $K_L^0 \to 3\pi^0$ events (photons, red) and Aluminum target data (neutrons, blue) with the MPPCs on the calorimeter. The line at 31.1 ns represents the cut made to accept 90% of photon events while rejecting over 98% of neutron events.

PMTs attached at the back of the CsI crystals. By using this dual-ended readout, a clear separation can be made between photon and neutron events, as shown in Figure 10.5. The performance of the MPPCs were tested in the 2019 runs and we were able to achieve a 90% photon acceptance with only 2.6% of neutrons remaining by making a cut on the timing difference. This corresponds to 0.5 hadron-cluster background events at the SM sensitivity for $K_L^0 \to \pi^0\nu\bar{\nu}$. More details on the MPPC upgrade of the calorimeter can be found in [97].

In addition to the calorimeter upgrade, a new detector was also installed in the downstream region to reduce the $K_L^0 \to \pi^+\pi^-\pi^0$ background. The Downstream Charged Veto (DCV) was installed before the 2019 runs and consists of two plastic scintillator pipes with WLS fibers embedded that are read out with MPPCs. Figure 10.6 shows a picture of the DCV detector and its location with respect to the other KOTO detectors. The DCV reduces the $K_L^0 \to \pi^+\pi^-\pi^0$ background especially in the lower P_T region as shown in Figure 10.7. More details on the DCV detector can be found in [98].

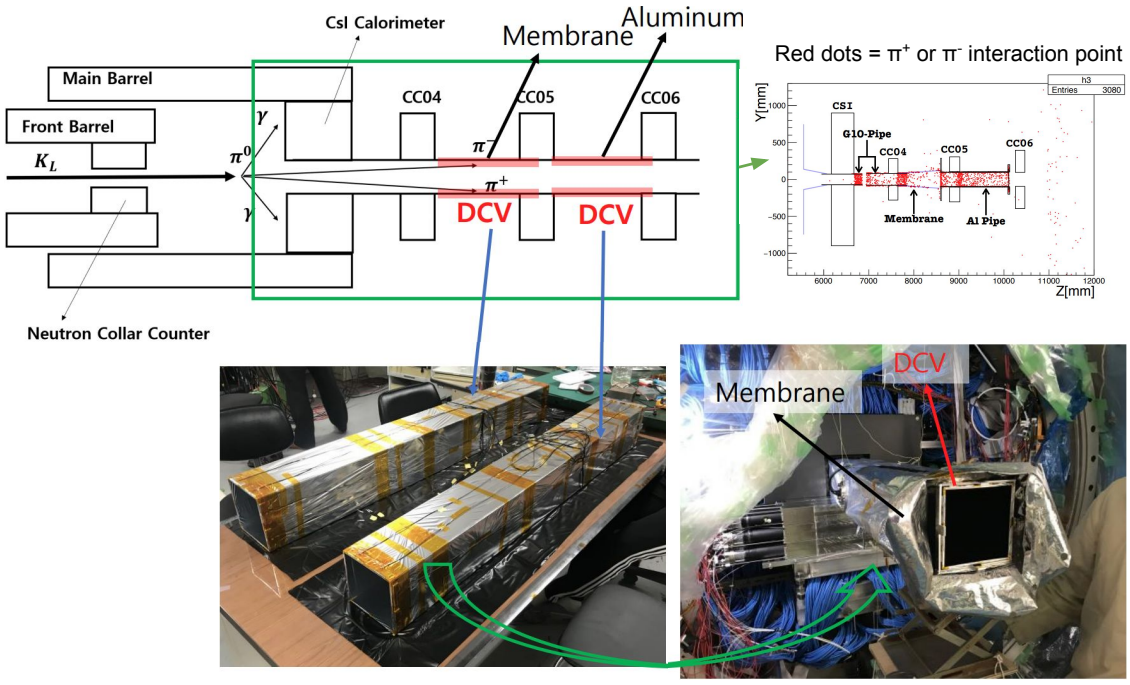


Figure 10.6: Installation photos of the DCV detector and its location within the KOTO detector system. The top right shows the interaction points of charged pions within the downstream beam pipe where the DCV was installed.

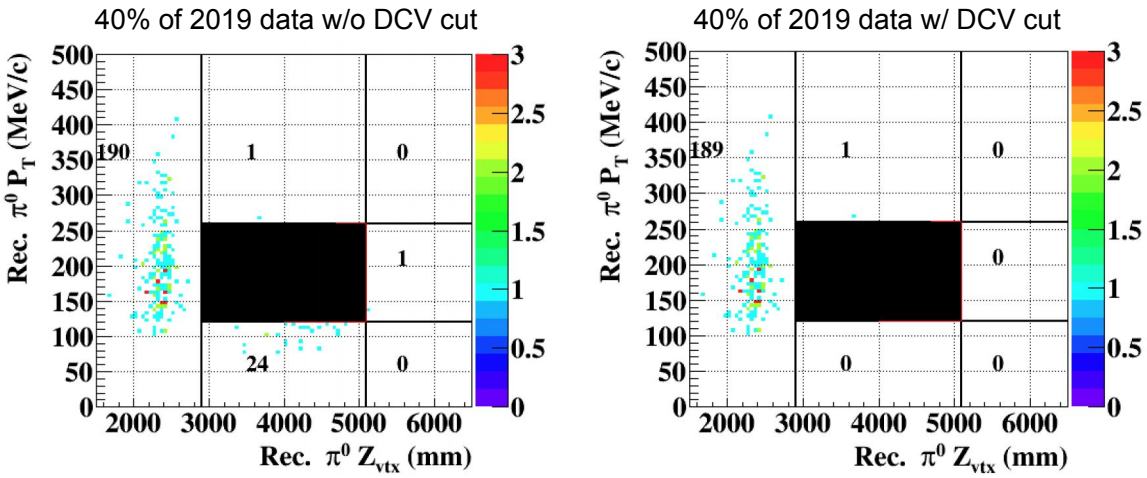


Figure 10.7: PtZ plots of 40% of the 2019 data with (right) and without (left) the DCV cut applied. The DCV cut rejects any event with more than 5 MeV of energy deposited in the DCV detector.

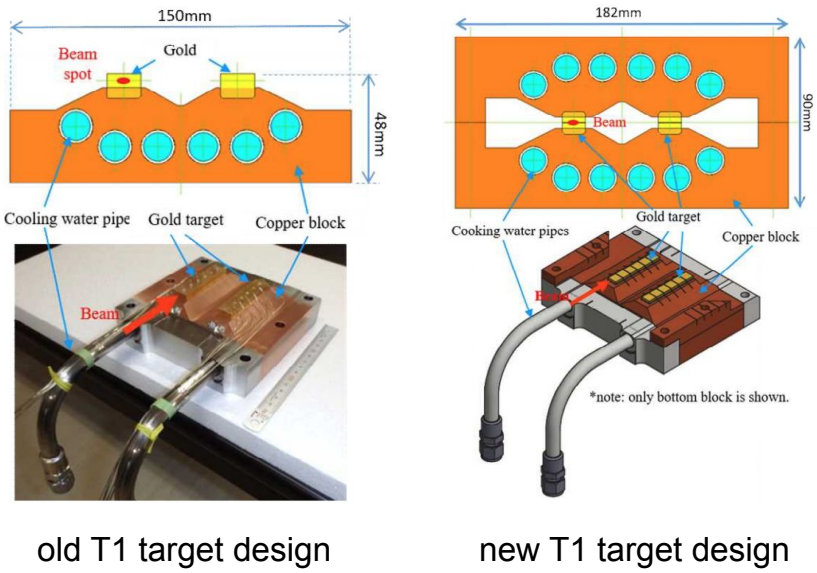


Figure 10.8: The old (left) and new (right) T1 production targets in Hadron Hall used to produce the neutral K_L beam. The new T1 target replaced the old T1 target in the fall of 2019. Figure modified from [99].

After the 2019 runs, the T1 target was upgraded in the fall of 2019 to have a higher cooling efficiency by doubling the copper cooling block to which the gold target is attached. Figure 10.8 shows the old and new T1 target designs. While the old T1 target design had a maximum beam power of 50 kW, the new T1 target is capable of a primary proton beam power up to 90 kW for a 5.52 second repetition cycle which is important for KOTO's continued data collection. A higher beam power means more K_Ls produced, providing higher statistics for the $K_L^0 \to \pi^0\nu\bar{\nu}$ search. More information about the new T1 target is covered in [99]. Additionally, in December of 2019 a new iron wall was installed on the north side of the KOTO vacuum vessel (Figure 10.9) to reduce accidental activity from the primary beam line. The iron wall reduced the accidental rate from the primary beam line by around 20% and consequently, the signal acceptance lost due to the accidental activity decreased by 7% going from 15% to 8%. Another iron wall was added in 2020 in the downstream barrel region for the same purpose.

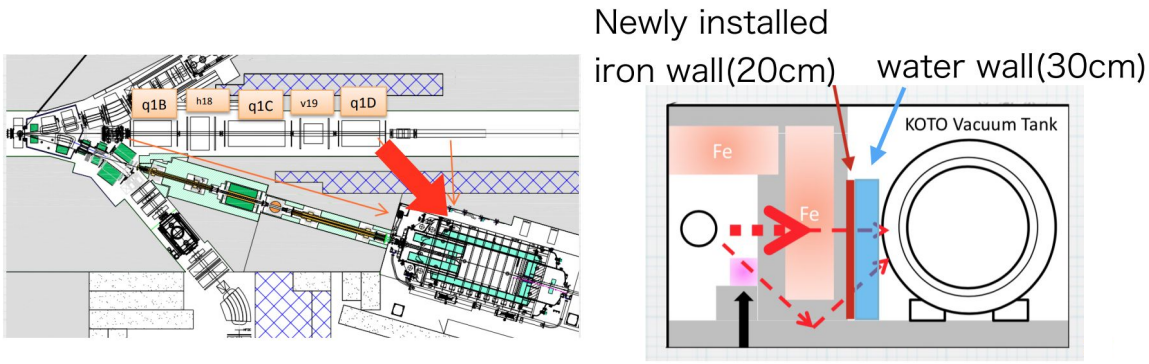


Figure 10.9: The location of the new iron wall installed in December of 2019 to reduce accidental activity from the primary beam line.

Before the June 2020 run, a prototype of a new detector called the UCV (Upstream Charged Veto) was designed, built, and installed to verify the existence of the $K^{\pm}$ background, as well as to test the design of this new detector to reduce the $K^{\pm}$ background. The prototype UCV was installed on the upstream side of the FB and consisted of 12 plastic scintillator fibers read out with MPPCs as shown in the left of Figure 10.10. The inefficiency of the prototype UCV against charged particles was checked by using $K^{\pm} \to \pi^0\pi^{\pm}$ events as shown in the right of Figure 10.10 and was found to be about 30%. This inefficiency was attributed to the limited coverage of the prototype, gaps in the scintillator fibers, and noise fluctuations in the detector readout system.

In designing the new UCV detector, these inefficiencies were reduced by increasing the detector size, tilting the new UCV detector with respect to the beam axis to avoid gaps in the fibers, and locating the MPPCs used for readout further away from the beam core to reduce the effects of irradiation on the MPPCs. The new UCV detector shown in Figure 10.11 was built with these specifications in mind and was installed to be utilized during data collection in 2021 to reduce the $K^{\pm}$ background. The new UCV is expected to perform with an efficiency of a few percent. Since the $K^{\pm}$ background was estimated to be the largest background, we are also considering

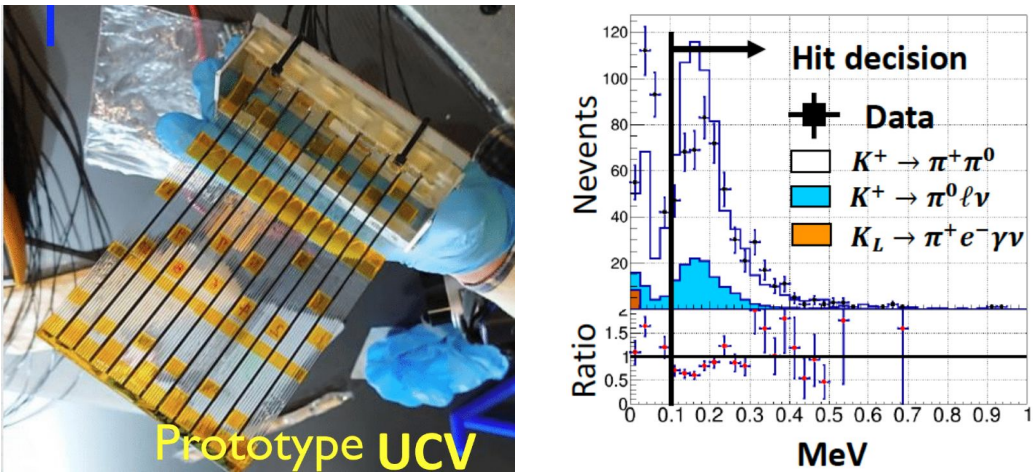


Figure 10.10: Picture of the prototype UCV (left) and the energy distribution of the prototype UCV for data and MC. The inefficiency of the prototype was around 30% and can be seen in the disagreement between the data and MC. The black line shows the 0.1 MeV energy threshold for the UCV veto decision.

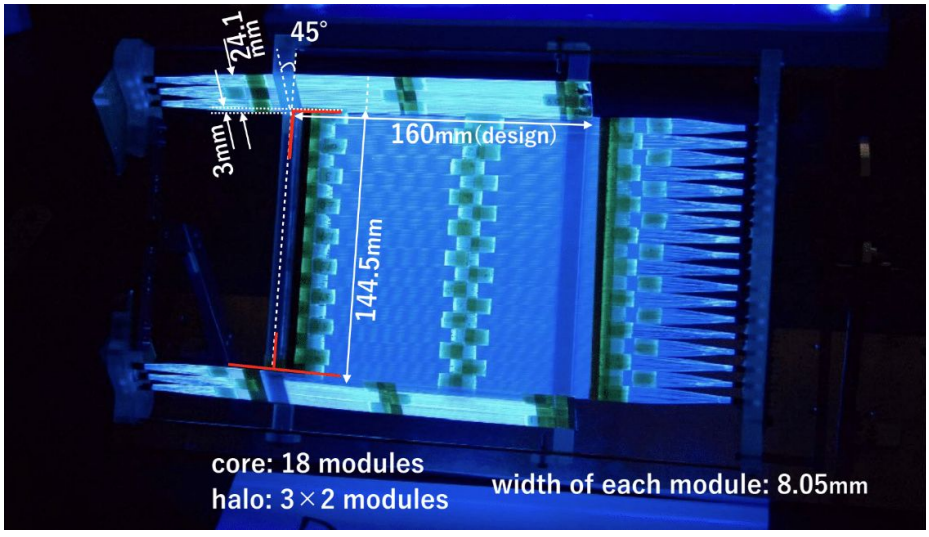


Figure 10.11: The new UCV detector was designed to reduce the $K^{\pm}$ background. The UCV was installed before data collection in 2021.

installing another sweeping magnet at the downstream end of the second collimator to reduce the $K^{\pm}$ background even further. Finally, a long shutdown of the J-PARC accelerator is planned for 2021 to upgrade the power supplies in the main ring. This upgrade will increase the beam power up to 100 kW, allowing for an even higher intensity K_L beam. This will enable KOTO to collect even higher statistics in the fall of 2022 when data collection resumes.

10.3.2 DAQ Upgrades

Since 2018, the DAQ system has performed reliably during data collection periods with a high livetime ratio of around 99%. Minor upgrades have been made to replace faulty ADC boards or other parts, but the DAQ system has remained largely the same as described in Section 3.7. However, after the planned accelerator upgrade to the MR power supplies, the beam power will double and the trigger rates will increase. Therefore, the DAQ must be upgraded to further reduce the trigger rates by moving the event building from the current L3 trigger (described in Section 3.8) earlier in the DAQ chain. Then the current L3 computer cluster will be able to perform additional online analysis on complete events to further reduce unwanted events in the data.

The proposed upgrade for the DAQ after the 2021 shutdown includes the addition of event-building OFC modules with multiple optical connectors as shown in Figure 10.12. The event building OFC modules are arranged in a pyramid structure where 16 ADC packages are transmitted to an OFC1 module at 2 Gbps and then assembled. There are 18 groups of ADCs and as a result, 18 OFC1 modules perform the local assembly in parallel and then send the resulting packages to the OFC2 module at 4 Gbps which builds the final event. The complete event is then sent to the L3 computer cluster via 10 Gbps Ethernet and the L3 trigger performs more sophisticated analysis on full events. Several tests have been carried out in the 2020 June run to test the performance and data transferring of the OFC modules and more tests are being

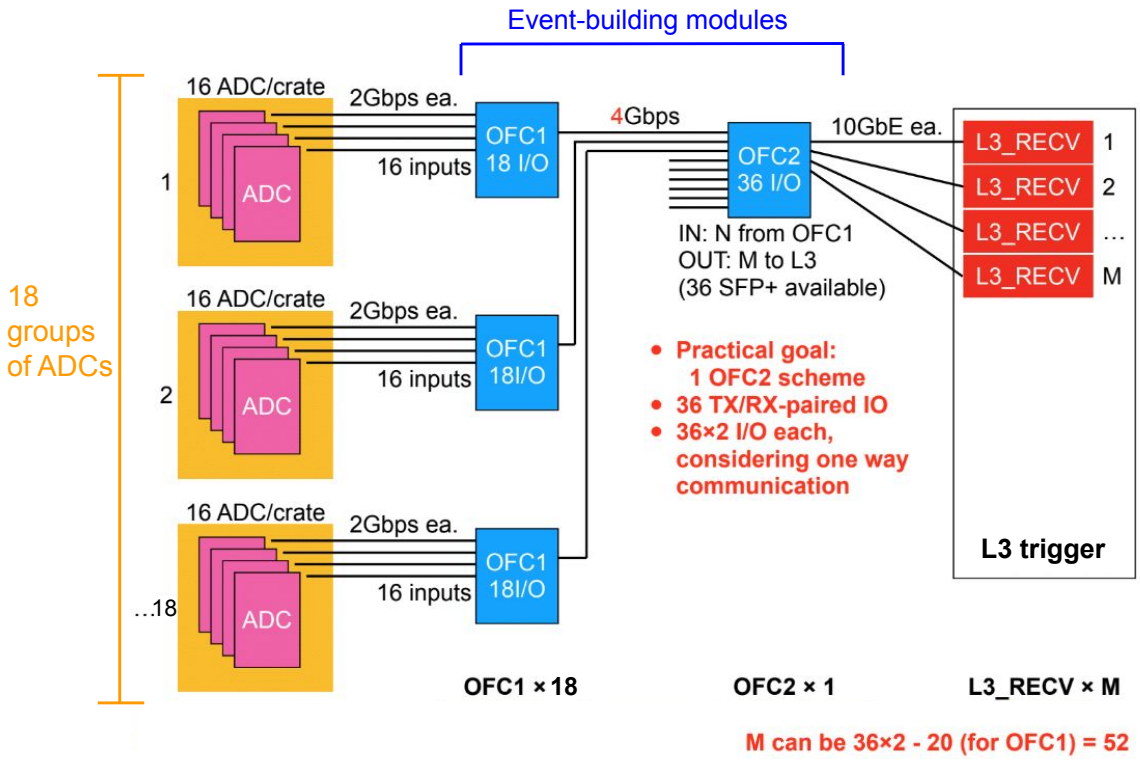


Figure 10.12: Schematic of the proposed upgrade to the DAQ system with events being fully built in the OFC module pyramid structure (blue) and then being passed to the L3 trigger for further online analysis.

performed in the 2021 run. Preliminary results indicate by simulating and testing these modules with triggers of the real beam structure at a doubled rate, that the data loss due to any bandwidth limitations is estimated to be negligible. After the 2021 shutdown the DAQ system will be upgraded with this new proposed system. Also, since the current L3 computer cluster is over 10 years old, we are also considering an upgrade to the L3 PCs to ensure stable data collection in the future.

10.3.3 Further Background Suppression

Aside from upgrades to the detector system to reduce background, several cuts are being developed to discriminate between signal and background. In particular, to reduce the halo K_L background, we are developing cuts based on cluster shape

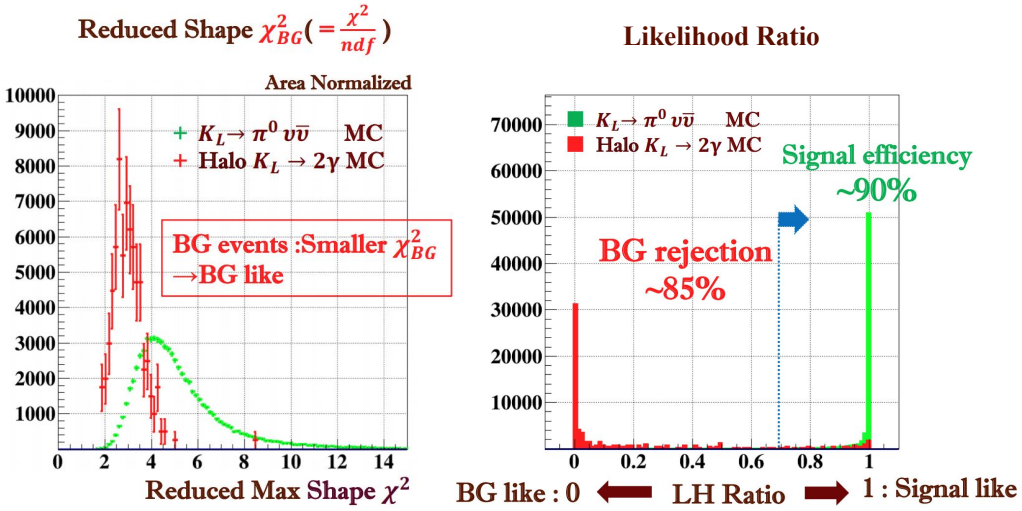


Figure 10.13: Preliminary results of a background reduction method to reduce the halo K_L background using cluster shape. Signal MC is shown in green and halo $K_L \to 2\gamma$ MC is shown in red. This method has a background rejection of 85% with a signal retention of 90%.

and kinematic parameters as well as cuts that use machine learning to discriminate between events that come from halo K_Ls in the beam and signal events. One discrimination method calculates the reduced shape χ^2 value for each cluster where clusters coming from halo K_L events are likely to have a smaller shape χ^2 value, as they decay off-axis. Preliminary results are shown in Figure 10.13 which give a background rejection power of around 85% with a signal efficiency of 90%. This cut can drastically reduce the halo K_L background which is the second largest background. Other cuts are being developed to reduce the $K^{\pm}$ background as well.

Many of the advanced cuts and discrimination methods that KOTO uses are based on classifying an event as either signal or background. For example, the CSDDL cut developed to reduce the hadron-cluster background uses machine learning to classify events as either neutron-like or photon-like based on their cluster shape. However, preliminary work has been completed on a new background reduction method that uses machine learning to return the incident angles of a photon cluster hit on the calorimeter based on the transverse shower profile. In this way, some directional

information can be obtained about the cluster, despite the inability to track the path of neutral photons with the KOTO detector system. Having directional information about where particles originated would be largely valuable to identifying and rejecting the many backgrounds that originate off-axis, such as the CV-η background, the CV-π^0 background, and the halo K_L background, as well as other possible backgrounds that may occur in KOTO. Details on the work that has been completed for this new background reduction method are described in Appendix A.

10.4 Continuing the Search

The KOTO experiment has continued to take data after 2018 and will continue collecting data in the future for the $K_L^0 \to \pi^0\nu\bar{\nu}$ search. Several upgrades have been completed to increase statistics and new methods are being developed to reduce background while retaining as much signal as possible. With the 2021 data we expect to reach a sensitivity of $\sim 3 \times 10^{-10}$ by suppressing the $K^{\pm}$ background and after 2021, the beam power will increase up to 100 kW. Therefore, with these projections the KOTO experiment expects to reach the Standard Model sensitivity by 2026.

After KOTO reaches the SM sensitivity, a next generation experiment at J-PARC to measure the branching ratio to order 10^{-12} called KOTO Step-2 is being considered [100]. This new experiment would include a new beam line with a smaller production angle of 5% to increase the K_L flux and a larger detector system that would require an extension to Hadron Hall. With these improvements, KOTO Step-2 would be able to measure the BR of $K_L^0 \to \pi^0\nu\bar{\nu}$ even more precisely by collecting 100 $K_L^0 \to \pi^0\nu\bar{\nu}$ events and set more stringent constraints on the Standard Model, or discover new physics. In addition, another experiment to measure $K_L^0 \to \pi^0\nu\bar{\nu}$ called KLEVER is proposed at the CERN SPS for 2025 [101].

Appendices

Appendix A

Photon Angle Reconstruction with Machine Learning

The reduction of background is the primary challenge of the KOTO analysis and in order to reach the SM sensitivity, it is essential to continue developing ways to reduce and reject background events while retaining as much signal as possible. This appendix details preliminary work that has been completed on a new background reduction method using machine learning (ML) to reconstruct the incident angles of the photons that hit the calorimeter using information about the transverse shower profile. In order to do this, two neural networks were designed, built, and trained and their performance was evaluated and compared. As this is a preliminary study, further studies and improvements to this work are discussed in Section A.6.

A.1 Purpose and Objectives

The motivation for this work was to develop a new method to reduce some of KOTO's largest backgrounds– events that come from off-axis particle interactions and decays. The signal signature is two photons from the π^0 that hit the calorimeter, and it is expected that the K_L decays on the beam axis. Because the photons are neutral, only information about the photon energies and positions on the calorimeter

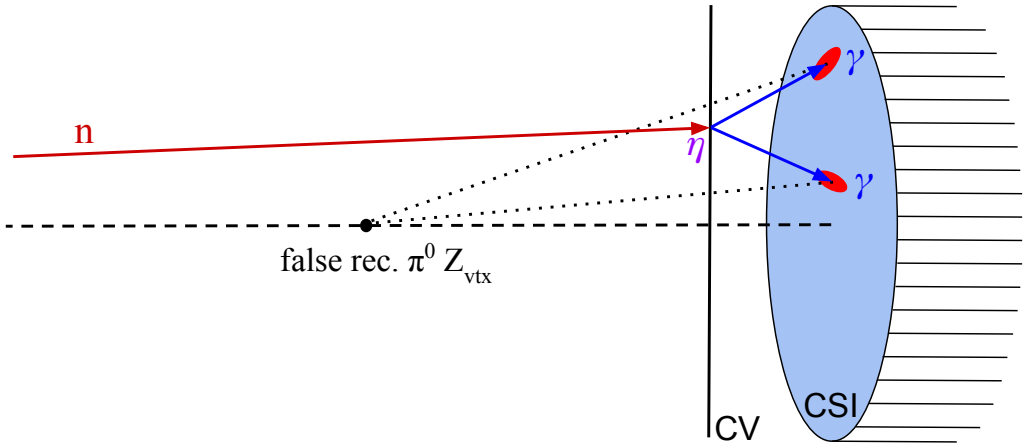


Figure A.1: Example of the CV-η background mechanism which occurs when a halo neutron in the beam hits the CV detector off-axis and creates and η particle which decays into two photons and mimics a signal event. The two photon hits are then used to erroneously reconstruct the π^0 decay position (Z_{vtx}) on the beam axis.

is known, and not the incident angles at which they hit the calorimeter. For this reason, it is not possible to identify whether an event comes from an off-axis decay or interaction, or is truly a signal event that originates on the beam axis. Many backgrounds such as the CV-η background, the CV-π^0 background, and the halo K_L background would be eliminated if the incident angles of the photons could be identified. For example, if we had directional information about the photons from the CV-η background in Figure A.1, it would be straightforward to identify as a background as one of the photons comes from a direction that is off-axis. This would also allow for the identification of any unknown backgrounds that come from off-axis interactions.

The objective of this study was to use a multidimensional, regression neural network (NN) to obtain the incident angles of the photons hitting the calorimeter based on the transverse shower profile. "Multidimensional" indicates that the network returns more than one output and "regression" indicates that the network solves a regression problem (returns a value) as opposed to a classification problem (returns a label). For example, there are two angles (θ and ϕ) that describe the incident direction

of the photon, and so the goal is to have a network return these two values. In contrast, the CSDDL cut (described in Section 5.3.3.3) which also uses a neural network, simply classifies whether a cluster is photon-like or neutron-like (binary classification problem). In fact, all of the advanced cuts that use ML in the KOTO analysis solve a classification problem, and so this study attempts to take a new approach and actually return the values of the incident angles to identify background events. With this objective in mind, I designed two different types of neural networks and compared their performance. An overview of these two networks, their inputs, outputs, and training samples used are detailed in Section A.2. The details and performance of each network are covered in Sections A.3 and A.4.

A.2 Methodology and Setup

Two types of neural networks were designed to reconstruct the incident angles of photons, a Fully Connected Network (FCN) or Multi-Layer Perceptron (MLP)[1] and a Convolutional Neural Network (CNN). The network inputs for the FCN were features of the cluster while the inputs to the CNN were cluster images. It is also important to note that in both of the networks, the inputs were individual clusters from photons from $K_L^0 \to \pi^0\nu\bar{\nu}$ MC and so the networks do not have information about the full event, they simply have information about single cluster hits on the CsI. Both networks were designed to output the incident angles of the photon, θ and ϕ, where θ is defined as the polar angle and ϕ is defined as the azimuthal angle in spherical coordinates as shown in Figure A.2. These angles were defined specifically from the point of view of the cluster, as opposed to from the beam axis, to prevent either of the networks from having global position information about the cluster hits on the CsI. In this way, the networks remain unbiased about the cluster angles that they are given for training, as all of the photons from the $K_L^0 \to \pi^0\nu\bar{\nu}$ MC come from

[1] A specific class of feed-forward NNs that are fully connected

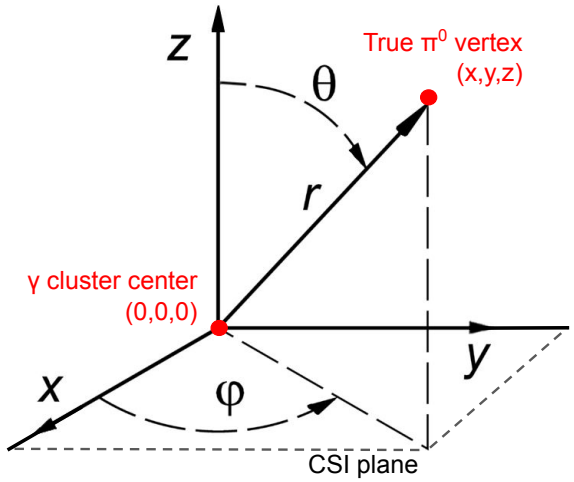


Figure A.2: The angles θ and ϕ are defined from the point of view of the cluster to prevent network bias. The origin (0, 0, 0) is defined as the center of the photon cluster on the CsI plane and the point (x, y, z) is the true π^0 decay vertex.

the beam line. The angle θ has a range from 0° to 45° as anything past 45° would have an angle behind the calorimeter and the angle ϕ has a range from −180° to 180°. In addition to the incident angles, the energy of the photon was also included as an output as a sanity check to ensure that the networks could perform as well as KOTO's own photon energy reconstruction (described in Section 5.2.3).

A.2.1 Training Sample

A sample of photon clusters from $K_L^0 \to \pi^0\nu\bar{\nu}$ MC was used for training, validation, and testing of the networks. In particular the MC sample used was from Run 79 (51 kW beam power) and a total of 327,738 photon clusters were used. The MC cluster sample was cleaned before being utilized by the networks such that only $K_L^0 \to \pi^0\nu\bar{\nu}$ MC events were used and a loose event selection criteria was applied[2]. The photon clusters from the π^0s were separated as individual inputs for training, however, the reconstructed π^0 Z_{vtx} and reconstructed π^0 P_T for the $K_L^0 \to \pi^0\nu\bar{\nu}$ training sample are shown in Figure A.3. For each cluster in the training sample, the true energy of

[2]All trigger bias removal cuts, γ selection cuts, background source cuts, and all veto cuts (older veto condition used; "MyVetoCondition==0").

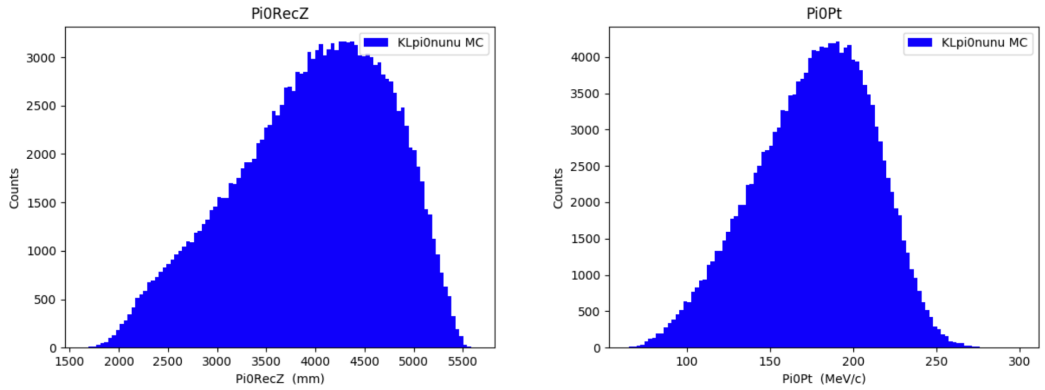


Figure A.3: The reconstructed π^0 Z_{vtx} (left) and the reconstructed π^0 P_T (right) for the $K_L^0 \to \pi^0\nu\bar{\nu}$ MC training sample with a loose cut selection applied.

the photon as well as the true incident angles, θ and ϕ, were calculated using the true decay vertex information from the MC. These distributions are shown in Figure A.4.

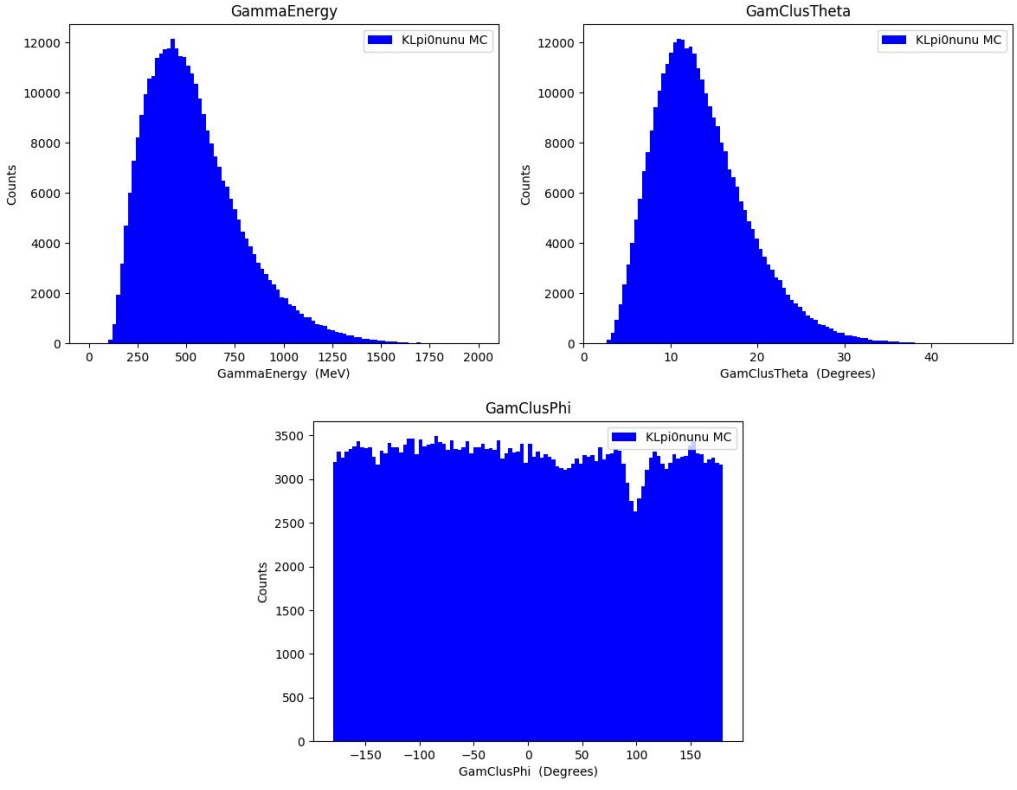


Figure A.4: Distributions of the true photon energy, true θ, and true ϕ for the $K_L^0 \to \pi^0\nu\bar{\nu}$ photon cluster training sample. These are the outputs of the NNs. It is unknown why there is a dip at 90° for the ϕ distribution.

A.3 Fully Connected Network (FCN)

This section details the input features, network architecture, and results and performance of the FCN.

A.3.1 Input Features

The FCN inputs were features or variables for each individual cluster that describe the shower shape in position and energy. There were eight input features that were calculated for each cluster and input into the network. These are listed in Table A.1 and the distributions are shown in Figure A.5. These features were only calculated with local position information such that the center of the cluster (x=0, y=0) was defined using the Center of Energy (COE) position. In this way, the network did not learn any global position information about where the cluster was on the CsI.

Input Features	**Definition**
Energy (E) deposited	Raw energy deposited in the cluster (not corrected)
Cluster size	Number of crystals in a cluster
E weighted RMS in x	$RMSx_\gamma = \sqrt{\sum_i \frac{e_i X_i^2}{\sum_j e_j}}$ where $X_i^2 = (x_i - x_{COE,\gamma})^2$
E weighted RMS in y	$RMSy_\gamma = \sqrt{\sum_i \frac{e_i Y_i^2}{\sum_j e_j}}$ where $Y_i^2 = (y_i - y_{COE,\gamma})^2$
E weighted skewness in x	$SkewX = \frac{\sum_i^n (x_{weighted,i} - \bar{x}_{weighted})^3}{(n-1)\sigma^3}$ where $x_{weighted,i} = \frac{e_i x_i}{E_{tot}}$
E weighted skewness in y	$SkewY = \frac{\sum_i^n (y_{weighted,i} - \bar{y}_{weighted})^3}{(n-1)\sigma^3}$ where $y_{weighted,i} = \frac{e_i y_i}{E_{tot}}$
E weighted kurtosis in x	$KurtosisX = \frac{\sum_i^n (x_{weighted,i} - \bar{x}_{weighted})^4}{(n-1)\sigma^4}$
E weighted kurtosis in y	$KurtosisY = \frac{\sum_i^n (y_{weighted,i} - \bar{y}_{weighted})^4}{(n-1)\sigma^4}$

Table A.1: Input features for the FCN for each photon cluster. E indicates energy, $x_i(y_i)$ is the $x(y)$ position of the ith crystal, e_i is the energy of the ith crystal, $\bar{x}_{weighted}(\bar{y}_{weighted})$ is the weighted mean in $x(y)$, n is the number of crystals in a cluster, and σ is the standard deviation.

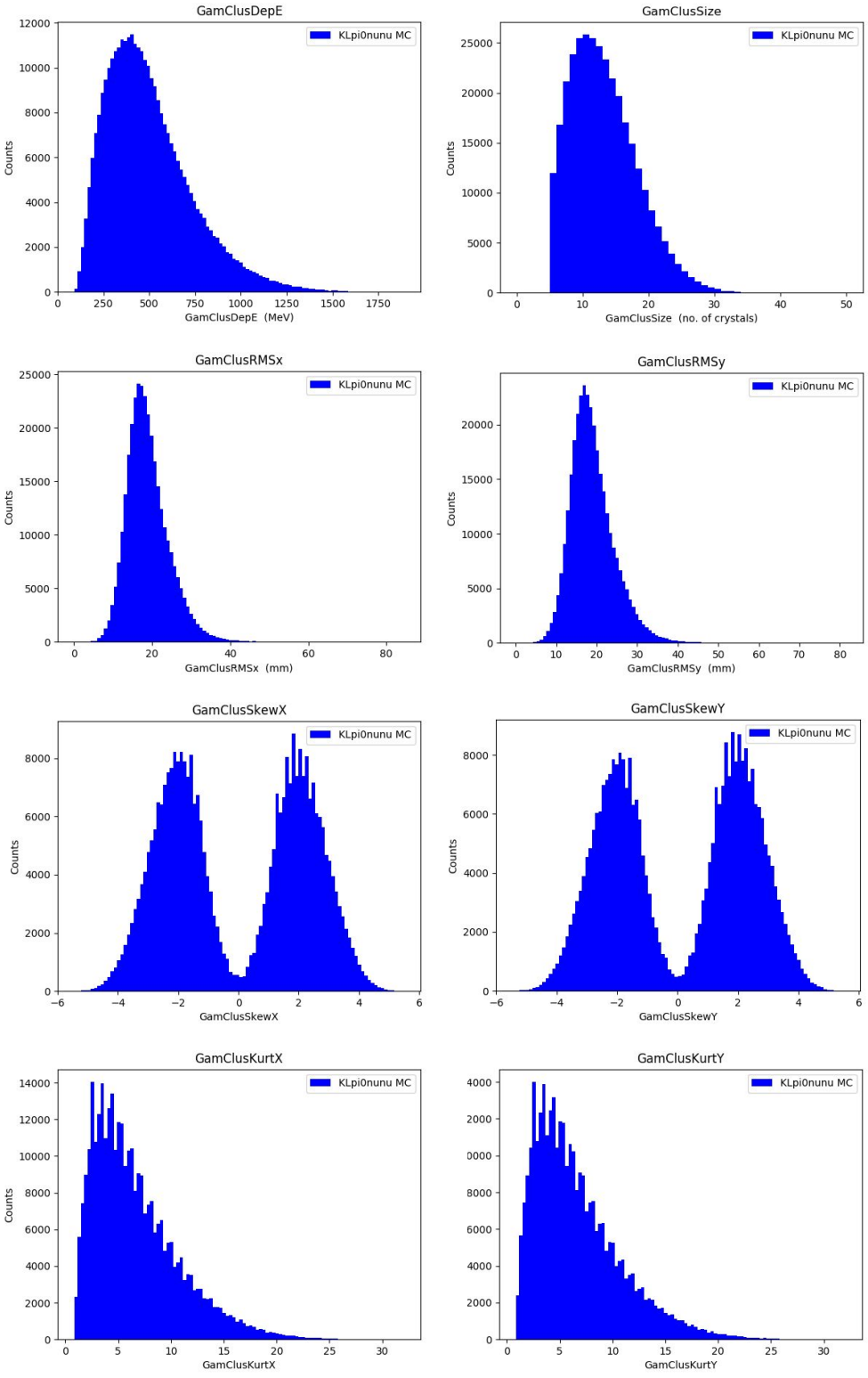


Figure A.5: Distributions of the input features for the FCN.

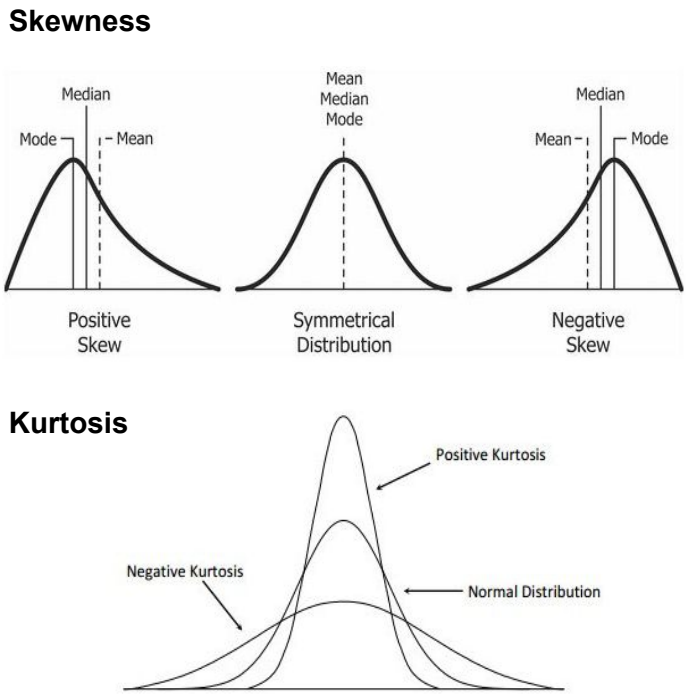
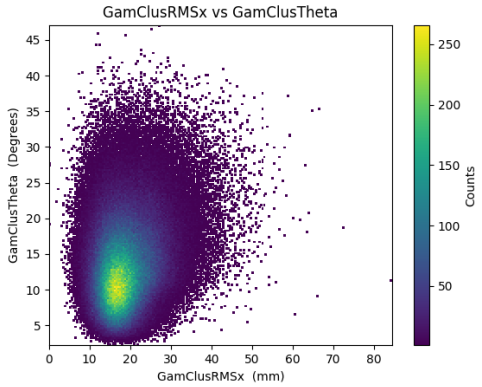
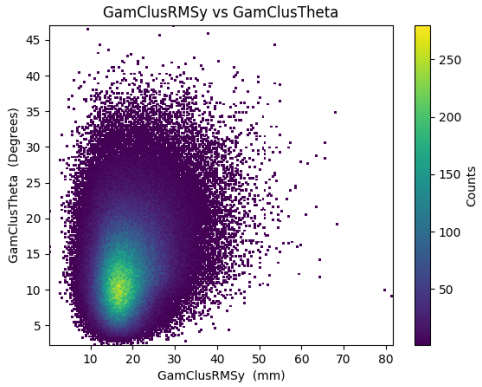


Figure A.6: Examples of sknewness and kurtosis relative to a Gaussian distribution.

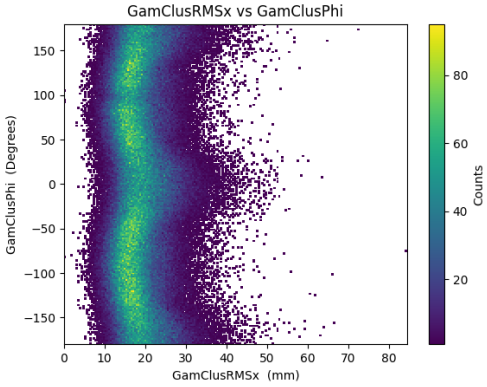
The input features of the cluster included the raw energy deposited in the cluster (not corrected for calibration), the size of the cluster, and variables that described the energy distribution in x and y. These were the RMS, skewness, and kurtosis of the energy distribution. The skewness and kurtosis are the third and fourth moments of a probability distribution, respectively. Skewness is a measure of asymmetry in a distribution relative to a Gaussian, and kurtosis is a measure of whether the data is heavy-tailed or light-tailed relative to a Gaussian. Figure A.6 shows examples of skewness and kurtosis with respect to a Gaussian distribution. The distributions of the moments RMS, skewness, and kurtosis vs the angles θ and ϕ are shown in Figures A.7 and A.8. These distributions demonstrate that there are correlations between the cluster shape and the incident angles. For example, in Figures A.8a and A.8b, there is a clear correlation between positive and negative skewness based on the angle ϕ.



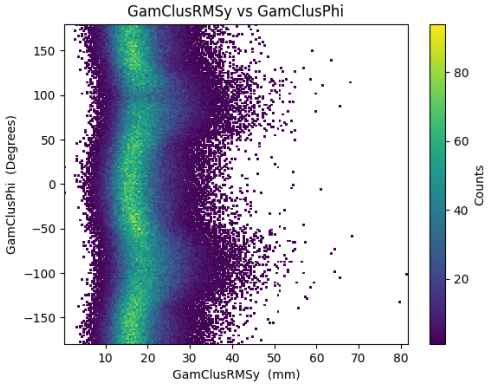
(a) RMS_x vs θ



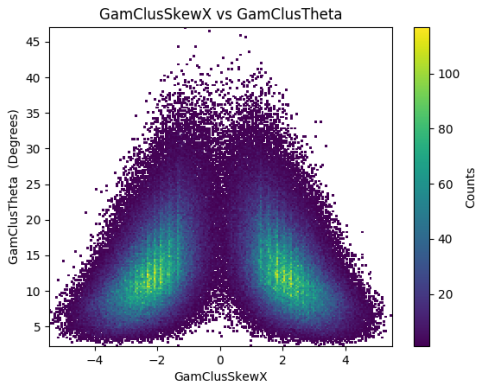
(b) RMS_y vs θ



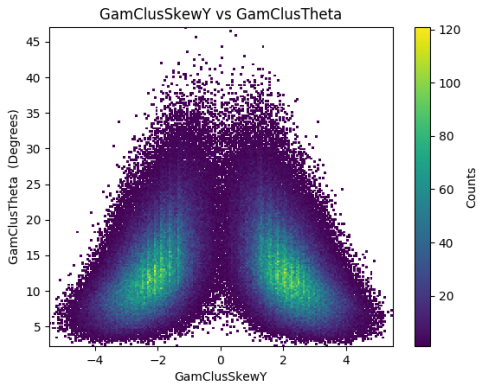
(c) RMS_x vs ϕ



(d) RMS_y vs ϕ

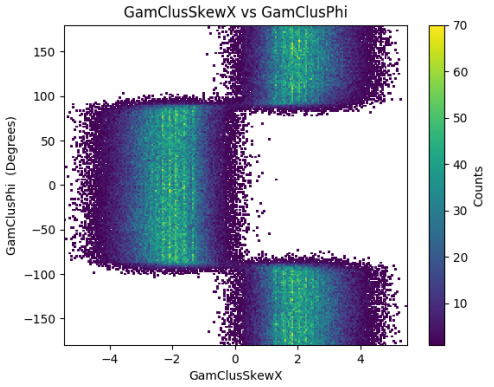


(e) SkewX vs θ

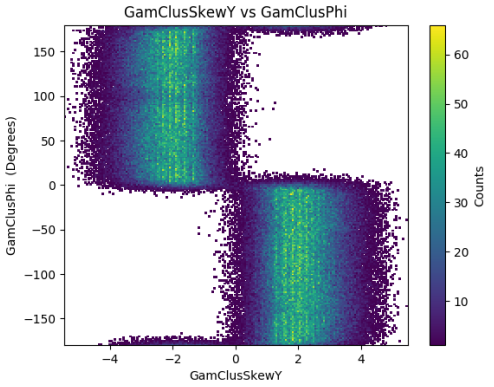


(f) SkewY vs θ

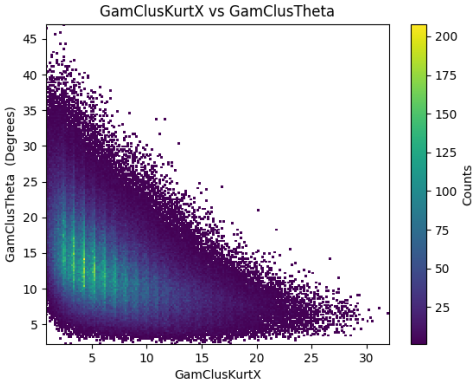
Figure A.7: Distributions of the moments (RMS, skewness, and kurtosis) vs the angles θ and ϕ.



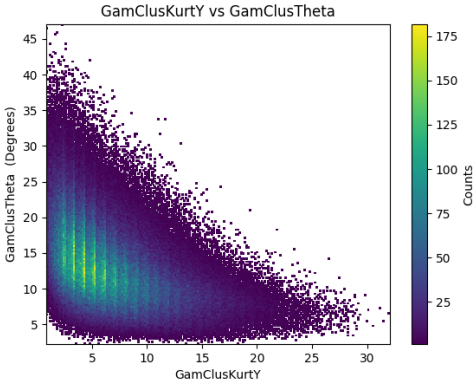
(a) SkewX vs ϕ



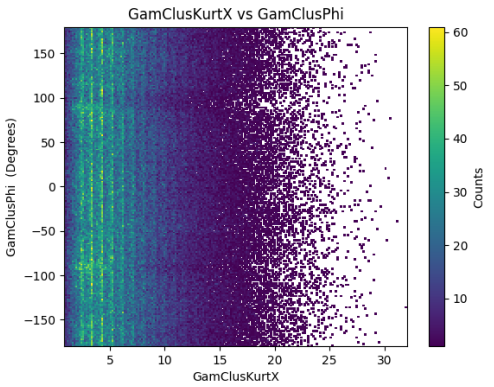
(b) SkewY vs ϕ



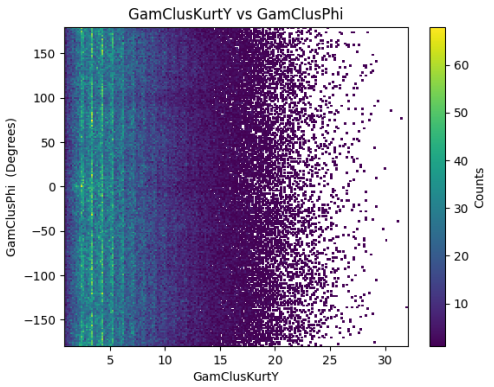
(c) KurtX vs θ



(d) KurtY vs θ



(e) KurtX vs ϕ



(f) KurtY vs ϕ

Figure A.8: Distributions of the moments (RMS, skewness, and kurtosis) vs the angles θ and ϕ.

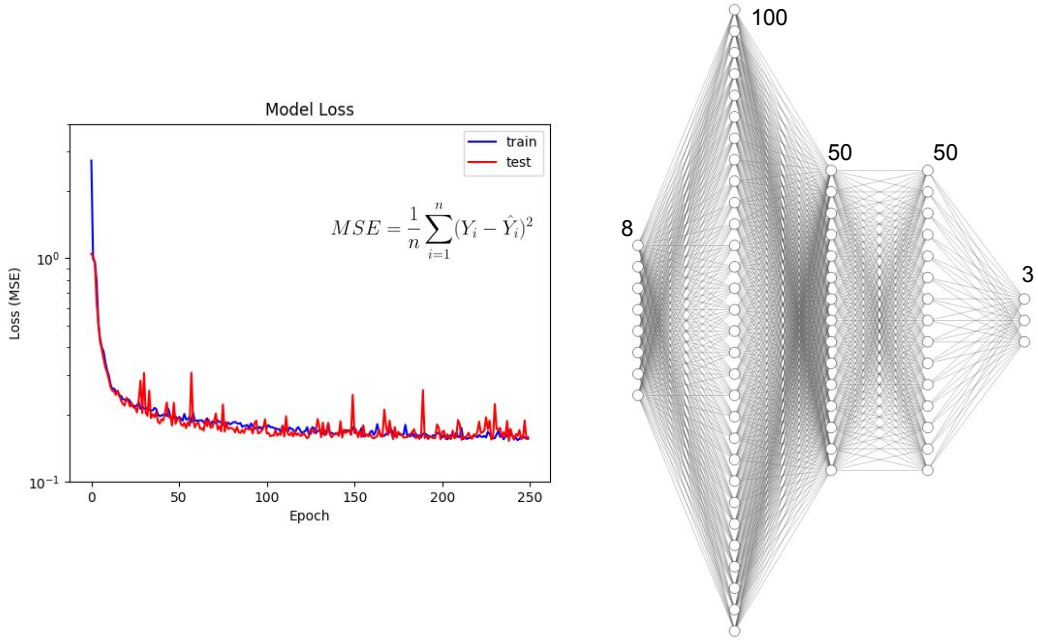


Figure A.9: The architecture of the FCN (right) and the model loss (MSE) (left) after training for 250 epochs.

A.3.2 Network Architecture and Training

In order to input the cluster features into the network, a data pipeline was designed to 1) convert the raw data contained in ROOT ntuples from the $K_L^0 \to \pi^0\nu\bar{\nu}$ MC files to numpy arrays using a package called "uproot" and to 2) calculate the input features and true output targets for each photon cluster. Once the input features and true outputs were calculated, these variables were saved in numpy arrays to be easily fed into the network. The FCN was implemented with Keras[3] with a Tensorflow backend. The right of Figure A.9 shows the architecture of the FCN which had eight input nodes in the first layer, three hidden layers with 100, 50, and 50 nodes, and three output nodes in the final layer. 64% of the data sample was used to train the network, 16% was used for validation, and 20% of the data was used to test the network's performance. The network was trained for 250 epochs and the loss was evaluated using the Mean Squared Error (MSE) as shown in the left of Figure

[3]A library that provides a Python interface for NNs

A.9. The output target θ was also reweighted to eliminate bias in the network. In Figure A.4, most of the photon clusters had an incident angle θ around 10° or 15° and so to prevent the network from learning that these angles were more probable, a reweighting was applied to θ such that the network equally considered all values of θ.

A.3.3 Results and Performance

To evaluate the performance of the FCN network, the resolution was calculated as

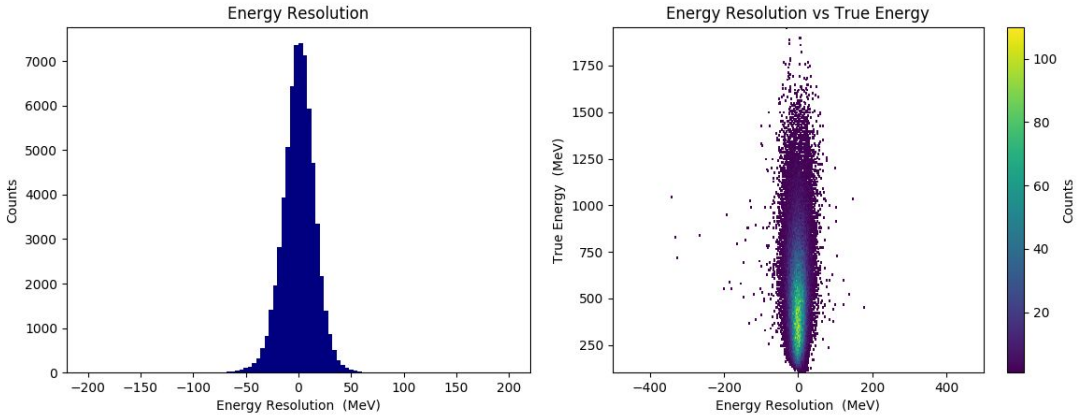
$$\text{resolution} = \text{network value} - \text{true value} \tag{A.1}$$

and the resolution was evaluated for the photon energy, θ, and ϕ. The mean and the RMS error of the resolution was also calculated where the RMS error is defined in Equation A.2.

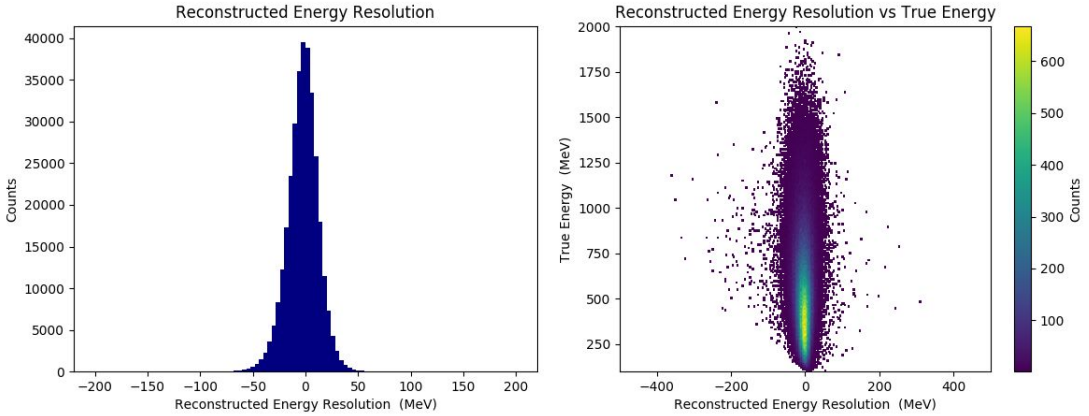
$$\text{RMS}_{\text{error}} = \sqrt{\frac{\sum_{i=1}^{n}(\text{predicted}_i - \text{true}_i)^2}{n}} \tag{A.2}$$

where n is the number of data events (photon clusters), and "predicted" and "true" represent the network value and the true value, respectively. The results for the energy resolution are shown in Figure A.10. Figure A.10a shows a histogram of the energy resolution (left) as well as the energy resolution vs the true energy (right) for the network's performance. In Figure A.10b, the same plots are shown instead for KOTO's own reconstruction of the energy. These plots show that the energy resolution of the network was comparable to our own reconstruction method with a difference of about 2%. This can be seen in Figure A.10c, where the energy resolution is compared for both the network energy and KOTO's usual energy reconstruction. A profile plot[4] to

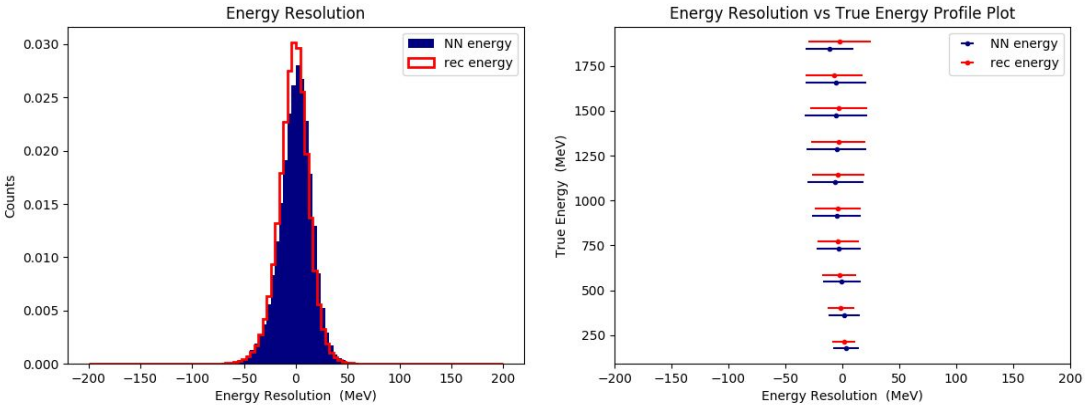
[4]In which the data point is the mean and the width of the error bar is the RMS



(a) Energy resolution (left) and energy resolution vs true energy (right) for the FCN. Mean = −0.75 MeV and RMS = 16.03 MeV.



(b) Energy resolution (left) and energy resolution vs true energy (right) for the KOTO energy reconstruction. Mean = −2.53 MeV and RMS = 15.70 MeV.



(c) Energy resolution (left) and energy resolution vs true energy (right) compared for the FCN (blue) and the KOTO reconstruction (red). The profile plot on the right shows the mean (data point) and the RMS (error bar) for each horizontal slice. The difference between the two methods is ∼2%.

Figure A.10: Energy resolution of the FCN compared to the KOTO reconstruction.

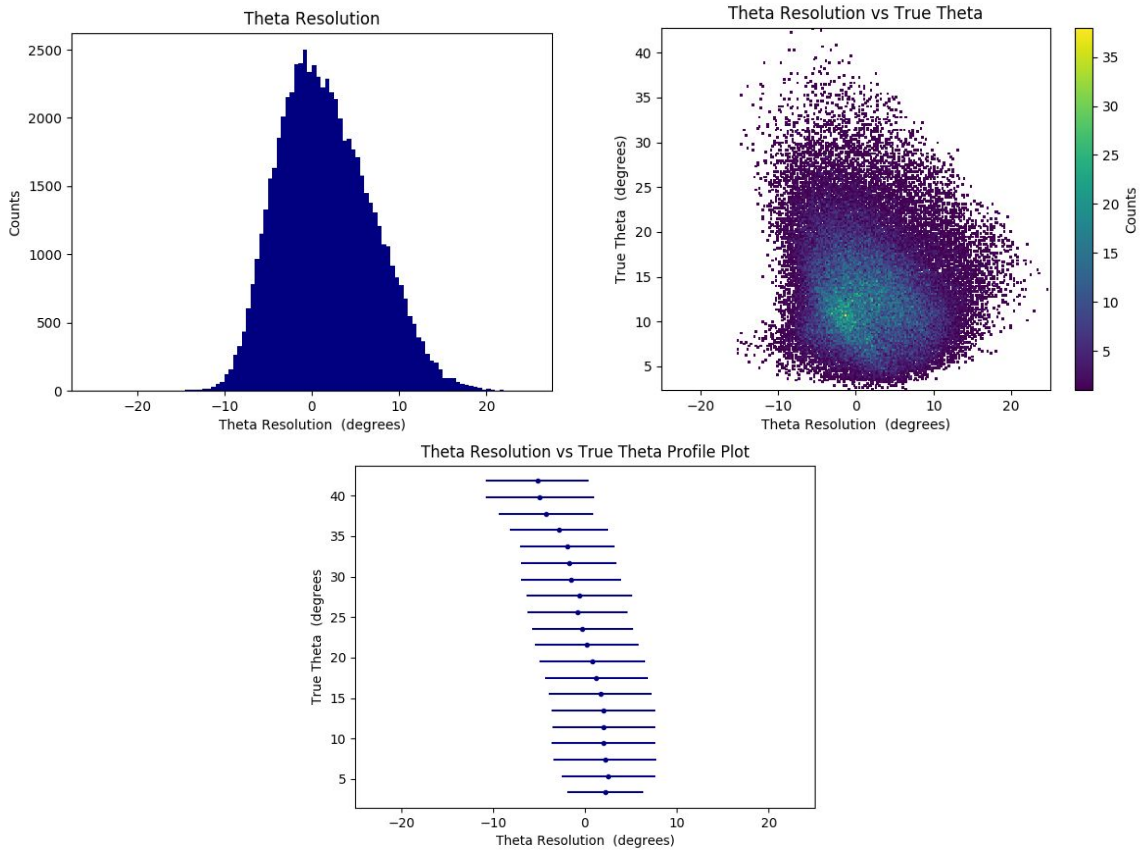


Figure A.11: The θ resolution for the FCN. Top left shows the θ resolution distribution, top right shows the θ resolution vs the true θ, and the bottom figure shows the profile plot for θ resolution vs true θ. Mean $= -0.72°$ and RMS $= 6.68°$.

compare the FCN energy resolution with the reconstructed energy resolution shows good agreement between the two methods which indicates the network behaved as expected.

The results for the θ resolution for the FCN are shown in Figure A.11. As seen in the profile plot of Figure A.11, there was still a small amount of bias in the θ resolution vs the true θ. However, despite this small bias, the RMS for the θ resolution was around 7° (recall the range of θ is 0° to 45°) and the mean was $-0.72°$.

The results for the ϕ resolution for the FCN are shown in Figure A.12. The plots are the same types as what is shown for the θ resolution. As illustrated in the ϕ resolution vs the true ϕ, there was a bias in ϕ and a 90° structure that can be seen in

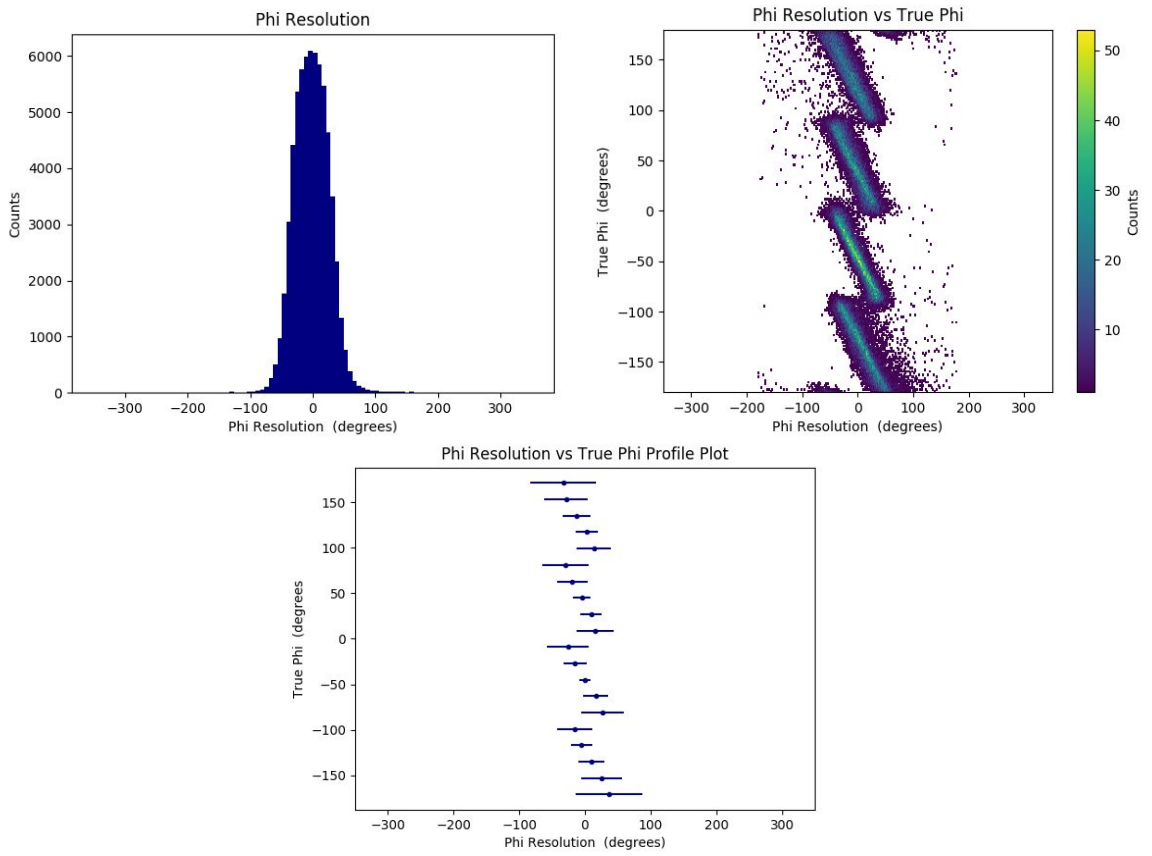


Figure A.12: The ϕ resolution for the FCN. Top left shows the ϕ resolution distribution, top right shows the ϕ resolution vs the true ϕ, and the bottom figure shows the profile plot for ϕ resolution vs true ϕ. Mean = $-7.43°$ and RMS = $41.78°$. The $90°$ structure was attributed to the network only receiving separate information about x and y.

the different quadrants of ϕ. This structure was attributed to the fact that the FCN only received information about x and y separately and wasn't given any information about the correlation between x and y. For example, no information was given about the diagonal energy distribution. Despite this structure, the ϕ resolution RMS was around $42°$ in total, but within the $90°$ structure, the resolution was closer to around $20°$. The mean of the ϕ resolution was $-7.43°$ for the FCN.

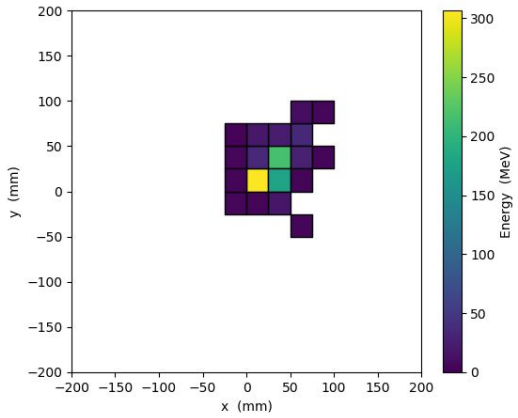


Figure A.13: An example of a cluster image containing only small crystals. The highest energy crystal is centered at the origin.

A.4 Convolution Neural Network (CNN)

This section details the cluster images, network architecture, and results and performance of the CNN.

A.4.1 Cluster Images

The inputs to the CNN were images of energy deposited in a cluster and an example is shown in Figure A.13. To create these images, the raw energy and position information for each crystal in the cluster was used to create a 27×27 pixel image. To prevent the network from learning any global position information, the crystal with the highest energy was placed at the origin. Because the CsI calorimeter is made up of both small (2.5×2.5 cm^2) and large (5×5 cm^2) crystals, the image was split into two different channels in order to account for the different sized crystals. This treatment of small and large crystals in an image is shown in Figure A.14. If a cluster had any large crystals in it, they were split up into four smaller crystals and the energy was divided between them. Then, two separate channels of the image were input into the network. The purpose of this step was to have all of the pixels that were input into the network the same size, as this CNN required inputs with the same dimension.

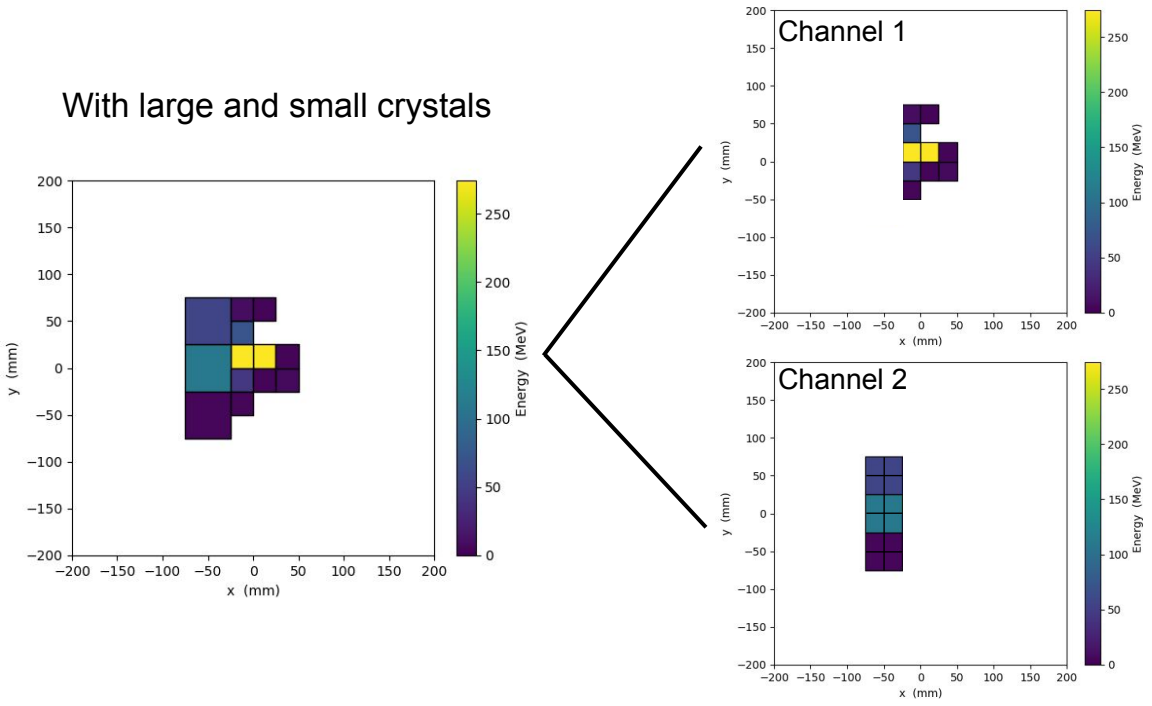


Figure A.14: An example of the treatment of clusters with large crystals. If the cluster contained any large crystals, they were split up into four smaller crystals and the energy was shared between them. Two channels of small crystals (channel 1) and larger crystals split up into smaller ones (channel 2) were then input into the CNN.

A.4.2 Network Architecture and Training

A similar data pipeline was designed to input data into the CNN as described in Section A.3.2. The cluster images were input into the CNN as numpy arrays in which each element in the array represented a pixel or crystal in the cluster image. The CNN, like the FCN, was also implemented in Keras with a Tensorflow backend. The architecture of the CNN is shown in the right of Figure A.15 and consisted of a 5×5 2D convolutional layer with 32 channels, a 2×2 max pooling layer with 32 channels, and a dropout layer. The filter maps were then flattened and input into a 128 node layer and then the output layer had three nodes. Similar to the FCN, 64% of the data sample was used to train the network, 16% was used for validation, and 20% of the data was used to test the network's performance. The CNN was trained for 50 epochs (anything past that lead to over-training) and the loss was evaluated using

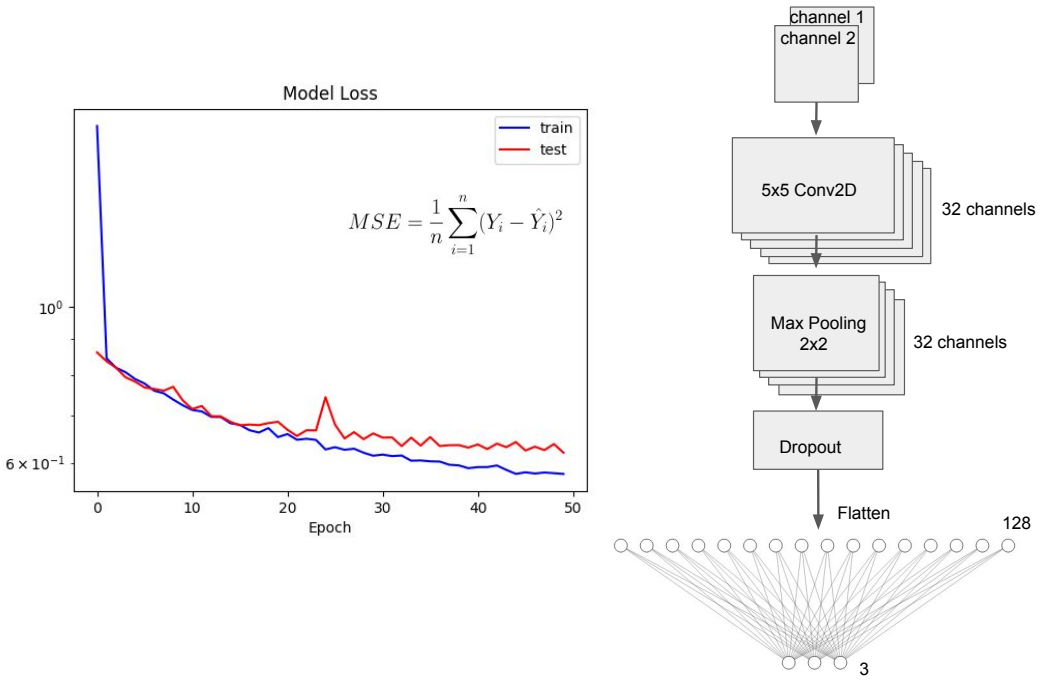


Figure A.15: The architecture of the CNN (right) and the model loss (MSE) (left) after training for 50 epochs.

the Mean Squared Error (MSE) as shown in the left of Figure A.15. The output θ was also reweighted to eliminate bias in the network, as it was in the FCN.

A.4.3 Results and Performance

The performance of the CNN was evaluated in the same way that the FCN was– by calculating the resolution (defined in Equation A.1) for the photon energy and incident angles θ and ϕ. The mean and RMS error (defined in Equation A.2) of the resolution were also calculated to evaluate the network performance. The results for the energy, θ, and ϕ resolution are shown in Figures A.16, A.17, and A.18, respectively. As with the FCN, the energy resolution and RMS for the CNN were comparable to KOTO's own energy reconstruction method. The θ resolution, like the FCN, had a small amount of bias as shown in Figure A.17, however the RMS of the θ resolution was slightly better than the FCN at around 5°. The ϕ resolution had a larger bias

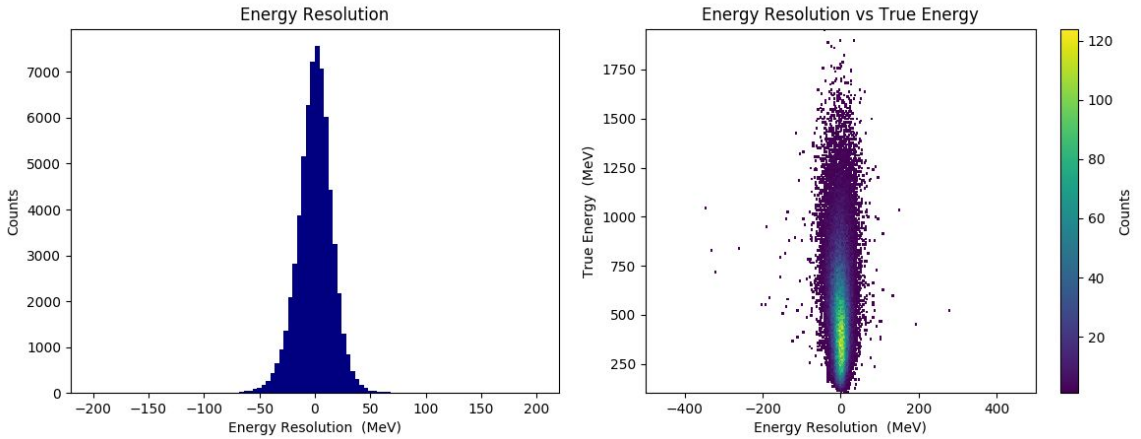


Figure A.16: The energy resolution for the CNN. The left figure shows the energy resolution distribution and the right shows the energy resolution vs the true energy. Mean = -0.03 MeV and RMS = 16.27 MeV.

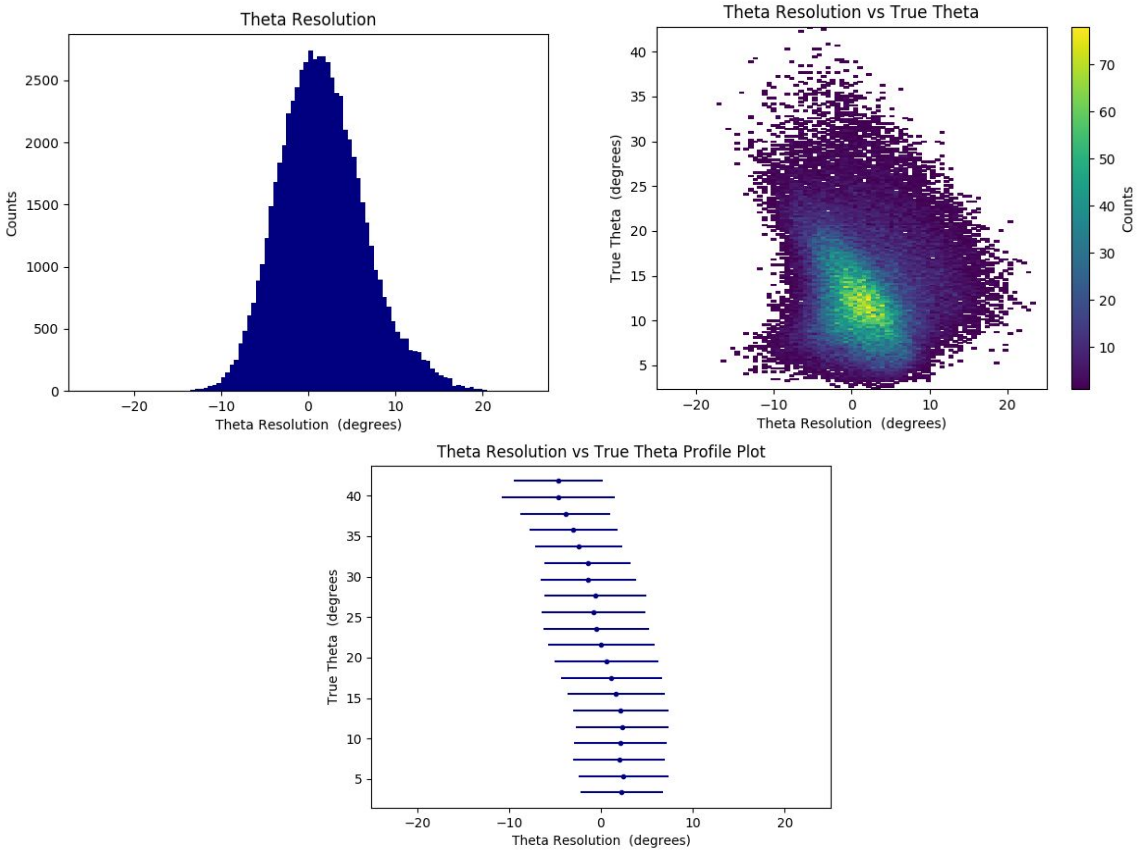


Figure A.17: The θ resolution for the CNN. Top left shows the θ resolution distribution, top right shows the θ resolution vs the true θ, and the bottom figure shows the profile plot for θ resolution vs true θ. Mean = 1.64° and RMS = 5.22°.

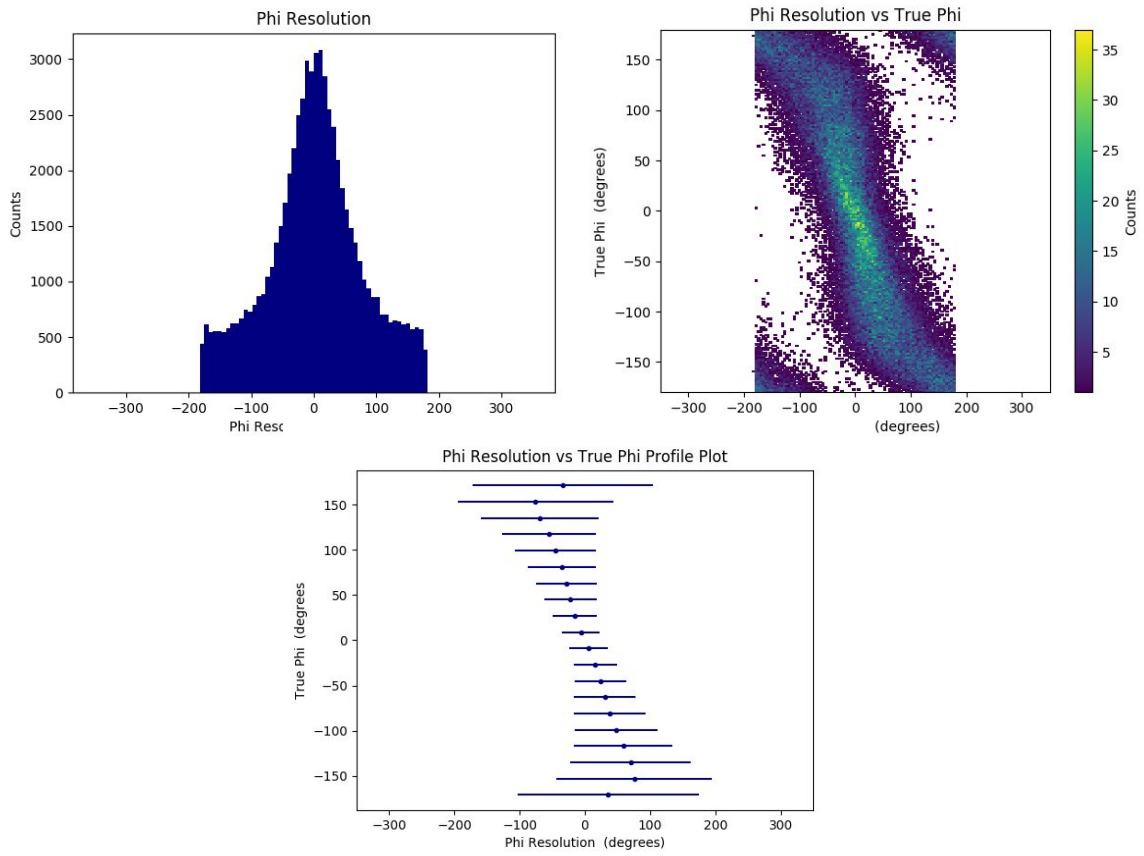


Figure A.18: The ϕ resolution for the CNN. Top left shows the ϕ resolution distribution, top right shows the ϕ resolution vs the true ϕ, and the bottom figure shows the profile plot for ϕ resolution vs true ϕ. Mean = 1.88° and RMS = 84.34°.

than the FCN which can be seen in the bottom of Figure A.18 and the RMS of this resolution was worse than that of the FCN at around 80°. Around a true ϕ of 0°, the RMS of the ϕ resolution was at its smallest point, which is about 40°. The mean of the ϕ resolution was 1.88° and the 90° structure that was seen in the FCN ϕ resolution was not present for the CNN, due to the fact that the CNN received images with information about both x and y as opposed to the FCN.

A.5 Comparison of FCN and CNN

A comparison of the performance of the two networks is discussed here. Figure A.19 shows profile plots for the FCN (blue) and the CNN (red) that show the resolu-

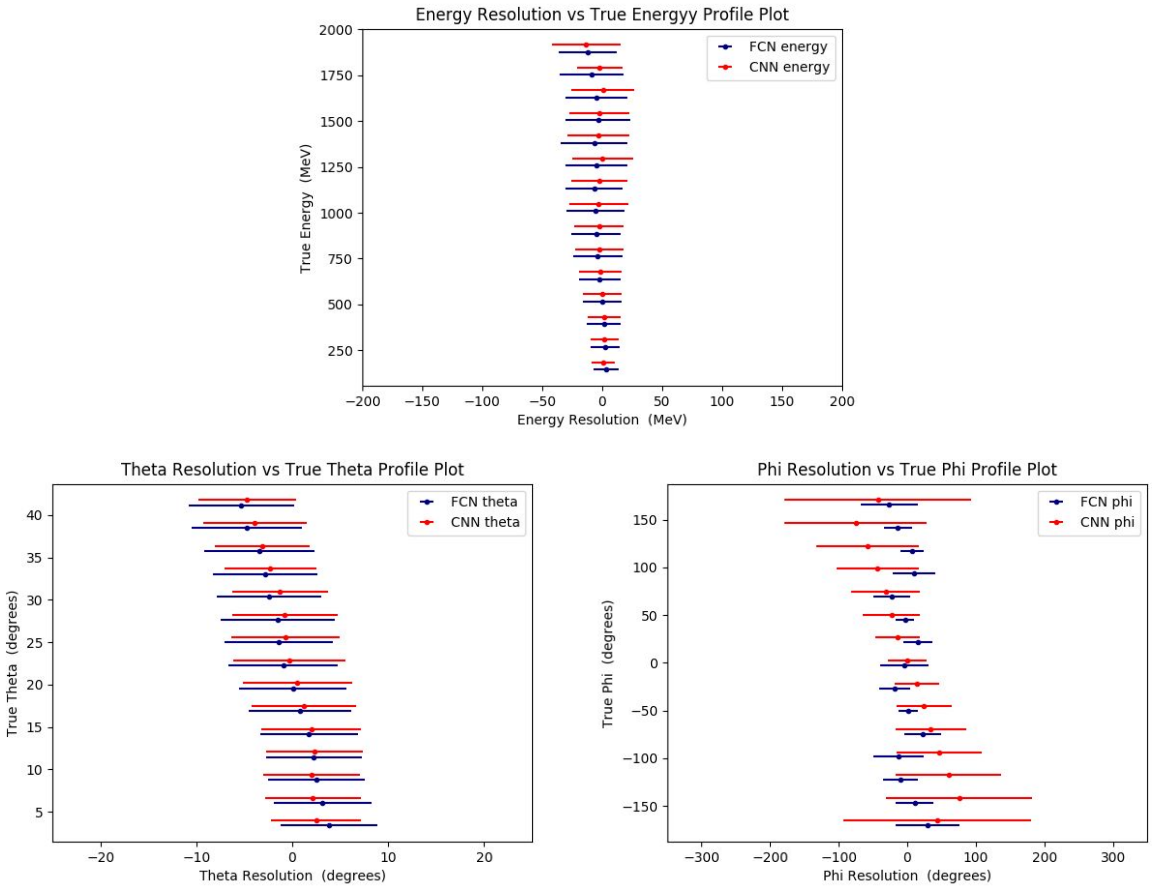


Figure A.19: Profile plots comparing the energy (top), θ (bottom left), and ϕ (bottom right) resolutions for the FCN (blue) and the CNN (red).

tion vs the true value for each output target (energy, θ, and ϕ). Table A.2 also lists the mean and RMS of the resolutions for each target for the FCN and the CNN. In comparing the two networks, the FCN gave a better RMS for the energy resolution than the CNN by 1.5% and also gave a better RMS for the ϕ resolution by 50%. However, the CNN outperformed the FCN and gave a better RMS for the θ resolution by 22%. Overall, the FCN performed better, despite having the 90° structure in the ϕ resolution due to the x and y information being provided to the network separately. An RMS of 7° in θ and around 20° in ϕ is quite good, considering there was no previous algorithm to obtain the incident angles of photon clusters. Furthermore, if the bias in ϕ can be eliminated, the RMS of the resolutions could be even better.

	Energy (MeV)		θ (degrees)		ϕ (degrees)	
	mean	RMS	mean	RMS	mean	RMS
FCN	-0.75	16.03	-0.72	6.68	-7.43	41.78
CNN	-0.03	16.27	1.64	5.22	1.88	84.34

Table A.2: Comparing the mean and RMS of the energy, θ, and ϕ resolutions for the FCN and the CNN. It should be noted that the usual energy reconstruction that KOTO uses yielded a mean of -2.53 MeV and an RMS of 15.70 MeV for the energy resolution. Both networks agree with this RMS within 4% and the networks both have a better mean resolution than the KOTO energy reconstruction.

In conclusion, this study shows that it is indeed possible to obtain directional information about the photons from their shower profile with machine learning. Because we must use the beam axis as a constraint in reconstructing the π^0 decay vertex in the KOTO analysis, knowing the incident angles of the photons would be very valuable and could help reduce backgrounds that come from off-axis decays and interactions. Even with this preliminary work, an RMS of 7° in θ and 20° in ϕ would be a narrow enough range to reject some backgrounds that come from an off-axis direction. However, further improvements can be made to this work, and these are discussed in the following section.

A.6 Further Studies and Improvements

The objective of this work was to use NNs to obtain the incident angles of the photons that hit the calorimeter so that this information could be used to reduce backgrounds that originate off-axis. There are several improvements that are necessary in order to fully implement this work as a background reduction method. For one, the networks need further optimization to eliminate the biases in both θ and ϕ, and while reweighting in θ eliminated some of the bias (see Figure A.20), some still remains. Several attempts were made to eliminate the bias in ϕ, however more work on this needs to be done. Other improvements could be made by optimizing the learn-

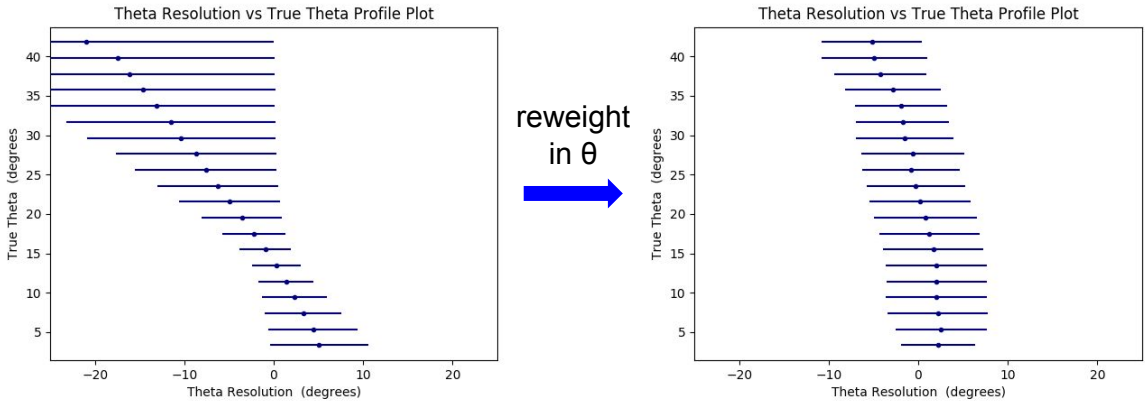


Figure A.20: Eliminating the bias in θ by reweighting (FCN).

ing rate, changing the network architecture, and increasing the training sample size. Since the FCN was limited by the input features only containing x and y information, combining the input features for the FCN and the output of the CNN into another FCN would likely improve the performance as well and this step is recommended for the continuation of this work.

An additional step would be to add timing information to the cluster images. This study only included energy information, however, each crystal has an associated hit time with the energy deposition. By adding two additional channels for the hit time of each crystal in the 27×27 pixel image, this would provide the network with information about which crystals were hit first which is likely correlated with the incident angle. It should also be noted that splitting up the small and large crystals into two separate channels introduces bias into the network since the large crystals are further from the beam axis and thus have a larger θ, and so this would need to be considered. One way to eliminate the need for two channels would be to use a different network type, such as a graph network that isn't limited by the data needing to have the same dimensions.

Another important improvement to be made is to use data to train the network instead of MC. Since the MC and the data don't always agree, it is important that

the network be trained on data if it will be used to evaluate data. This work was completed with $K_L^0 \to \pi^0\nu\bar{\nu}$ MC, and so the true vertex of the π^0 was known and was used to calculate the true incident angles of the photon. It is possible to use $K_L^0 \to 3\pi^0$ data to calculate the incident angles of the photons that hit the calorimeter, as there are enough constraints with six photon hits to calculate the decay vertex accurately. Additionally, it's possible to use reconstructed π^0s generated at the Al target from the Al target data (covered in Section 4.4.1). By using the π^0 mass as a constraint and reconstructing the decay position to be at the Al target, a data sample of π^0s could be used to train the network.

In order to fully implement this study as a background reduction method, the above steps would be necessary to complete. After a network was sufficiently trained with data, then a cut could be developed to reduce backgrounds originating off the beam axis. The incident angles from the network could be used to reconstruct the π^0 decay vertex position and this could then be compared to the KOTO π^0 reconstructed decay vertex. If any events reconstructed with the network incident angles were not close to the KOTO reconstructed decay vertex, a cut (perhaps using a χ^2 value) could be applied to remove these events. Finally, since this study only considered photon clusters, it could potentially be extended to include neutron clusters as well to further reduce the hadron-cluster background.

Appendix B

Overview of B Physics and Lepton Universality

Measuring the kaon decays $K_L^0 \to \pi^0\nu\bar{\nu}$ and $K^+ \to \pi^+\nu\bar{\nu}$ are considered excellent ways to test CP violation in the quark sector. In addition to these decays, B mesons also provide a very clean way to investigate CP violation in the quark sector, as does lepton universality in the lepton sector. Sections 1.5 and 1.6 cover details of the rare kaon decays, and this appendix gives a brief overview of the B meson processes and lepton universality.

B.1 B Mesons

CP violation has been heavily studied through B physics from experiments including LHCb at CERN, Belle/BelleII at KEK, and BABAR at SLAC. A major focus of these experiments is working to measure CP violation in the golden decay, $B^0 \to J/\psi K_s^0$. This process includes contributions from both tree level diagrams and penguin diagrams. The penguin diagrams are similar to the loop diagram for $K_L^0 \to \pi^0\nu\bar{\nu}$ but changes a $\bar{b}$ quark to an $\bar{s}$ quark. The penguin level contributions affect the sin2β quantity where β is one of the angles in the unitary triangle. Determining sin2β requires measuring the time-dependent difference in decay rates for

$B^0 \to J/\psi K_s^0$ and $\bar{B}^0 \to J/\psi K_s^0$. The SM predicts sin2$\beta$ to be $= 0.74 \pm 0.02$ [102] and LHCb reports a value of sin2$\beta = 0.73 \pm 0.04$ [103]. Because of the efforts of LHCb, Belle, and BABAR, β is currently one of the most accurately determined CP parameters.

Additionally, B_d and B_s meson systems allow for fundamental testing of the SM through particle-antiparticle mixing. In both of these systems, a neutral B meson, B_q^0 (where $q = d, s$) can decay into a $\bar{B}_q^0$ meson or vice versa, which leads to oscillations between the mass eigenstates. These transitions are due to weak interactions and are described by box diagrams involving two W bosons and two up-type quarks. By measuring the masses and decay widths of these systems, CP violation and CPT symmetry can be tested. Specifically, if CPT is conserved, the masses and the decay rates of B_q^0 and $\bar{B}_q^0$ should be identical. CP violation can also be investigated via mixing through the mass differences, Δm_d and Δm_s, which constrain the apex of the unitary triangle.

B.2 Lepton Universality

The SM predicts that the different flavors of leptons should behave in the same way and be produced equally as often in weak decays. Along with other semileptonic decay modes, one way of testing this idea of lepton universality is by looking at the ratio of BR($B^0 \to K^* e^+ e^-$) to BR($B^0 \to K^* \mu^+ \mu^-$). These rare B meson decays provide a particularly good laboratory for testing universality because the decays are FCNC processes (similar to the loop diagrams of $K_L^0 \to \pi^0 \nu \bar{\nu}$) in which unseen particles could affect the decay rates, such as a yet undiscovered charged Higgs boson or another undiscovered particle all together. Searches for deviations from lepton universality have been performed by the BABAR and Belle experiments, as well as LHCb. In particular, LHCb reports [104] a measurement of the ratio,

$$\frac{BR(B^0 \to K^* \mu^+ \mu^-)}{BR(B^0 \to K^* e^+ e^-)} = 0.745 \pm 0.036 \tag{B.1}$$

which deviates from the SM by 2.6 standard deviations. Earlier measurements by Belle [105] and BABAR [106] had much larger uncertainties but found the ratio to be consistent with the SM. While these experimental measurements are still within the SM, these results could be an early indication of physics BSM. If this is the case, such results challenging lepton universality would point to new particles or interactions and would require a major revision to our understanding of physics.

In summary, FCNC processes while rare, are an important tool for probing physics BSM. Rare kaon and B meson decays allow us to directly measure CP violating parameters and constrain the CKM model of weak interactions. Thus, experiments studying these processes are essential for investigating physics that we do not fully understand. The current limits for the CKM parameters are shown in Figure B.1.

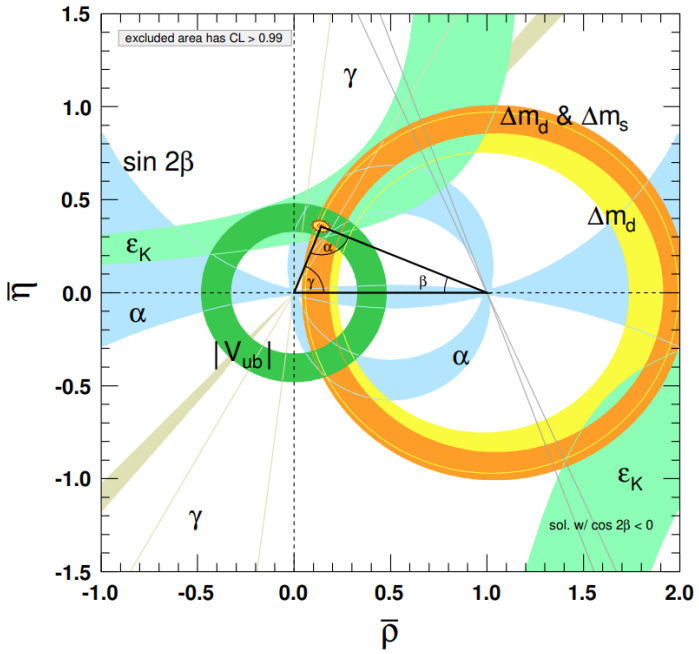


Figure B.1: Constraints on the unitarity triangle in the $\bar{\rho}$–$\bar{\eta}$ plane. The shaded areas have 99% CL. Figure courtesy of [107].

Appendix C

Acronyms and Glossary

ADC– Analog-to-Digital Converter

ALP– Axion-Like Particle

BNL– Brookhaven National Laboratory

BR– Branching Ratio

BSM– Beyond the Standard Model

Banjo– The name of the L3 computer cluster of 47 nodes

Beam Exit– The end of the KL beam line defined at the end of the second collimator, at $z = -1.507$ m

CDT– Clock Distribution and Trigger

CERN– European Organization for Nuclear Research

CKM matrix– Cabibbo-Kobayashi-Maskawa matrix

CL– Confidence Level

CNN– Convolutional Neural Network

COE– Center of Energy

CP violation– Charge-parity violation

CSD– Cluster Shape Discrimination

Cluster– Energy deposited in a group of CsI crystals from a particle hitting the calorimeter

Clustering– The process of identifying clusters on the CsI calorimeter

CsI– Cesium Iodide

Cuts– Selection criteria that are used to retain signal and reduce background

DAQ– Data Acquisition

DL– Deep Learning

DVU– Decay Volume Upstream

$\mathbf{E}_t$– Total energy

ESS1(2)– Electro-Static Septum 1(2)

Even pairing– In pairing photons to reconstruct π^0s, the even pairing occurs when two photons are paired from the same π^0

FCN– Fully Connected Network

FCNC– Flavor Changing Neutral Current

FIFO– First-In-First-Out

FPGA– Field Programmable Gate Array

FTT– Fast Fourier Transform

FX– Fast Extraction

Fermilab– Fermi National Accelerator Laboratory

Fourier PSD– Fourier Pulse Shape Discrimination

Gbps– Gigabits per second

Geant3/Geant4– A package written in C++ that is widely used in high energy physics to simulate particle interactions in matter

HD– Hadron Experimental Facility (AKA Hadron Hall)

Halo neutrons– Neutrons outside the beam width in the beam halo

J-PARC– Japan Proton Accelerator Research Complex

Kμ3– The decay mode $K_L^0 \to \pi^{\pm}\mu^{\mp}\nu_{\mu}$

KEK– High Energy Accelerator Research Organization

KEKCC– KEK Computer Cluster

Ke3– The decay mode $K_L^0 \to \pi^{\pm}e^{\mp}\nu_{e}$

Kurtosis– The fourth moment of a probability distribution which measures whether the data is heavy-tailed or light-tailed compared to a Gaussian

L1 Accept (L1A)– Event that passed the L1 trigger decision and L2 is not busy

L1 Gate (L1Gate)– The number of clocks that the L1 trigger requires to process events

L1 Raw (L1Raw)– Raw trigger/event coming into the DAQ system

L1 Reject (L1Rej)– Event that was rejected by the L1 trigger decision

L1 Request (L1Req)– Event that is far enough apart from other triggers and passed the L1 trigger decision

L1– Level 1

L2 Accept (L2A)– Event that passed the L2 trigger decision

L2 Reject (L2Rej)– Event that was rejected by the L2 trigger decision

L2– Level 2

L3– Level 3

LHT– Littlest Higgs model with T-parity

Livetime– The fraction of time the DAQ system (or subsystem) is processing data (live) without any delays or loss of data

MACTRIS+– MAster Control and TRIgger Supervisor (the plus indicates an upgraded version of MACTRIS)

MC– Monte Carlo

MFV– Minimum Flavor Violation

MIP– Minimum-Ionizing Particle

ML– Machine Learning

MLF– Material and Life Science Experimental Facility

MLP– Multi-Layer Perceptron

MPI– Message Passing Interface

MPPC– Multi-Pixel Photon Counter

MR– Main Ring

MSE– Mean Squared Error

MSSM– Minimal Supersymmetric Model

Masking– Masking or the masking background occurs when accidental activity in the beam cause overlapped pulses where the true hit pulse is "masked" by an accidental pulse and the veto timing isn't calculated correctly

NN– Neural Network

NU– Neutrino Experimental Facility

Normalization modes– The K_L decay modes used in the normalization analysis; $K_L^0 \to 3\pi^0$, $K_L^0 \to 2\pi^0$, and $K_L^0 \to 2\gamma$

OFC– Optical Fiber Center

Odd pairing– In pairing photons to reconstruct π^0s, the odd pairing occurs when two photons are paired from different π^0s

Off-spill– The time during a spill cycle where there is no beam extracted

On-spill– The time during a spill cycle where there is beam extracted; typically two seconds

Opening the box– Unblinding the the signal region to reveal the data inside

P$_T$– Transverse momentum

PMT– Photomultiplier Tube

POT– Protons on Target

Pedestal– The baseline of a waveform

Profile plot– A plot in which the data point represents the mean and the width of the error bar represents the RMS error for a particular slice in the data

PtZ plot– A plot in the reconstructed π^0 P_T vs reconstructed π^0 Z_{vtx} plane

QCD– Quantum Chromodynamics

RCS– Rapid Cycling Synchrotron

RMS– Root Mean Square

RSc– Randall-Sundrum model with custodial symmetry

Reduction power– The fraction of events that survive a particular cut

SES– Single Event Sensitivity; defined in Section 5.1.2

SINET– Science Information NETwork; a Japanese academic network for universities and research institutions

SM– Standard Model

SM4– SM with a 4th generation of quarks and leptons

SX– Slow Extraction

SY– Switch Yard

Skewness– The third moment of a probability distribution which measures the asymmetry of the distribution compared to a Gaussian

Spill cycle– A full acceleration cycle of the proton extraction, typically lasting 5.2 to 6.0 seconds

Spill– A shot of extracted protons from the MR onto the T1 target in a single acceleration cycle, typically lasting two seconds

TMon– Target Monitor

UL– Upper Limit

VME– Versa Module Europa; a computer bus standard

WLS– Wavelength Shifting

Bibliography

Bibliography

[1] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.13.321

[2] P. Higgs, "Broken symmetries, massless particles and gauge fields," *Phys. Lett.*, vol. 12, no. 2, pp. 132–133, 1964. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0031916364911369

[3] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett. B*, vol. 716, no. 1, pp. 1–29, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S037026931200857X

[4] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys. Lett. B*, vol. 716, no. 1, pp. 30–61, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0370269312008581

[5] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2π Decay of the K_2^0 Meson," *Phys. Rev. Lett.*, vol. 13, pp. 138–140, Jul 1964. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.13.138

[6] J. P. Lees *et al.*, "Observation of Time-Reversal Violation in the B^0 Meson System," *Phys. Rev. Lett.*, vol. 109, p. 211801, Nov 2012. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.109.211801

[7] G. D. Rochester and C. C. Butler, "Evidence for the Existence of New Unstable Elementary Particles," *Nature*, vol. 160, pp. 855–857, 1947.

[8] M. Gell-Mann and A. Pais, "Behavior of Neutral Particles under Charge Conjugation," *Phys. Rev.*, vol. 97, pp. 1387–1389, Mar 1955. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.97.1387

[9] K. Lande, E. T. Booth, J. Impeduglia, L. M. Lederman, and W. Chinowsky, "Observation of Long-Lived Neutral V Particles," *Phys. Rev.*, vol. 103, pp. 1901–1904, Sep 1956. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.103.1901

[10] M. Tanabashi *et al.*, "Review of Particle Physics," *Phys. Rev.*, vol. D98, no. 3, p. 030001, 2018.

[11] N. Cabibbo, "Unitary Symmetry and Leptonic Decays," *Phys. Rev. Lett.*, vol. 10, pp. 531–533, Jun 1963. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.10.531

[12] M. Kobayashi and T. Maskawa, "CP-Violation in the Renormalizable Theory of Weak Interaction," *Prog. Theor. Phys.*, vol. 49, p. 652, 1973. [Online]. Available: http://ptp.oxfordjournals.org/content/49/2/652.full.pdf+html

[13] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa Matrix," *Phys. Rev. Lett.*, vol. 51, pp. 1945–1947, Nov 1983. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.51.1945

[14] A. Buras, "Flavour Changing Neutral Current Processes," *Proceedings, 28th International Conference on High Energy Physics, Warsaw, Poland, 25-31 July 1996*, 1996. [Online]. Available: https://cds.cern.ch/record/313473/files/9610461.pdf

[15] A. Buras, D. Buttazzo, J. Girrbach-Noe, and R. Knegjens, "$K^+_L \to \pi^+\nu\bar{\nu}$ and $K^0_L \to \pi^0\nu\bar{\nu}$ in the Standard Model: Status and Perspectives," *J. High Energy Phys.*, vol. 1511, no. 033, 2015. [Online]. Available: https://arxiv.org/pdf/1503.02693v2.pdf

[16] T. Inami and C. S. Lim, "Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K_L \to \mu\bar{\mu}$, $K^+_L \to \pi^+\nu\bar{\nu}$ and $K^0 \leftrightarrow \bar{K^0}$," *Prog. Theor. Phys.*, vol. 65, no. 5, pp. 1772–1772, 05 1981. [Online]. Available: https://doi.org/10.1143/PTP.65.1772

[17] Y. Grossman and Y. Nir, "$K^0_L \to \pi^0\nu\bar{\nu}$ Beyond the Standard Model," *Phys. Lett. B*, vol. 398, p. 163, 1997. [Online]. Available: https://cds.cern.ch/record/318376/files/9701313.pdf

[18] A. Artamonov *et al.*, "New measurement of the $K^+_L \to \pi^+\nu\bar{\nu}$ branching ratio," *Phys. Rev. Lett.*, vol. 101, p. 191802, 2008. [Online]. Available: https://arxiv.org/pdf/0808.2459.pdf

[19] E. C. Gil *et al.*, "First search for $K^+_L \to \pi^+\nu\bar{\nu}$ using the decay-in-flight technique," *Phys. Lett. B*, vol. 791, pp. 156 – 166, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0370269319301121

[20] D. M. Straub, "New physics correlations in rare decays," in *CKM unitarity triangle. Proceedings, 6th International Workshop, CKM 2010, Warwick, UK, September 6-10, 2010*, 2010. [Online]. Available: http://inspirehep.net/record/881964/files/arXiv:1012.3893.pdf

[21] M. Blanke, A. J. Buras, B. Duling, K. Gemmler, and S. Gori, "Rare K and B Decays in a warped extra dimension with custodial protection," *J. High Energy Phys.*, vol. 2009, no. 03, p. 108–108, Mar 2009. [Online]. Available: http://dx.doi.org/10.1088/1126-6708/2009/03/108

[22] M. Blanke, A. J. Buras, B. Duling, S. Recksiegel, and C. Tarantino, "FCNC Processes in the Littlest Higgs Model with T-Parity: an Update," 2009.

[23] T. Hurth, G. Isidori, J. F. Kamenik, and F. Mescia, "Constraints on new physics in MFV models: A model-independent analysis of $\Delta F = 1$ processes," *Nucl. Phys. B*, vol. 808, no. 1-2, p. 326–346, Feb 2009. [Online]. Available: http://dx.doi.org/10.1016/j.nuclphysb.2008.09.040

[24] A. J. Buras, B. Duling, T. Feldmann, T. Heidsieck, C. Promberger, and S. Recksiegel, "Patterns of flavour violation in the presence of a fourth generation of quarks and leptons," *J. High Energy Phys.*, vol. 2010, no. 9, Sep 2010. [Online]. Available: http://dx.doi.org/10.1007/JHEP09(2010)106

[25] G. Isidori, F. Mescia, P. Paradisi, C. Smith, and S. Trine, "Exploring the flavour structure of the MSSM with rare K decays," *J. High Energy Phys.*, vol. 2006, no. 08, pp. 064–064, Aug 2006. [Online]. Available: https://doi.org/10.1088%2F1126-6708%2F2006%2F08%2F064

[26] A. J. Buras, D. Buttazzo, J. Girrbach-Noe, and R. Knegjens, "Can we reach the Zeptouniverse with rare K and $B_{s,d}$ decays?" *J. High Energy Phys.*, vol. 2014, no. 11, Nov 2014. [Online]. Available: http://dx.doi.org/10.1007/JHEP11(2014)121

[27] L. S. Littenberg, "CP-violating decay $K_L^0 \to \pi^0 \nu\nu$," *Phys. Rev. D*, vol. 39, pp. 3322–3324, Jun 1989. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.39.3322

[28] G. E. Graham *et al.*, "Search for the decay $K_L^0 \to \pi^0 \nu\bar{\nu}$," *Phys. Lett. B*, vol. 295, no. 1, pp. 169 – 173, 1992. [Online]. Available: http://www.sciencedirect.com/science/article/pii/037026939290107F

[29] M. Weaver *et al.*, "Limit on the Branching Ratio of $K_L \to \pi^0 \nu\overline{\nu}$," *Phys. Rev. Lett.*, vol. 72, pp. 3758–3761, Jun 1994. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.72.3758

[30] J. Adams *et al.*, "Search for the decay $K_L^0 \to \pi^0 \nu\bar{\nu}$," *Phys. Lett. B*, vol. 447, no. 3, pp. 240 – 245, 1999. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0370269398015937

[31] A. Alavi-Harati *et al.*, "Search for the decay $K_L \to \pi^0 \nu\overline{\nu}$ using $\pi^0 \to e^+e^-\gamma$," *Phys. Rev. D*, vol. 61, p. 072006, Mar 2000. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.61.072006

[32] J. K. Ahn *et al.*, "New limit on the $K_L^0 \to \pi^0 \nu\overline{\nu}$ decay rate," *Phys. Rev. D*, vol. 74, p. 051105, Sep 2006. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.74.051105

[33] J. K. Ahn *et al.*, "Search for the Decay $K_L^0 \to \pi^0\nu\overline{\nu}$," *Phys. Rev. Lett.*, vol. 100, p. 201802, May 2008. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.100.201802

[34] J. K. Ahn *et al.*, "Experimental study of the decay $K_L^0 \to \pi^0\nu\overline{\nu}$," *Phys. Rev. D*, vol. 81, p. 072004, Apr 2010. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.81.072004

[35] J. Ahn *et al.*, "A New Search for the $K_L^0 \to \pi^0\nu\bar{\nu}$ and $K_L^0 \to \pi^0 X^0$ decays," *Prog. Theor. Exp. Phys.*, 2016. [Online]. Available: https://arxiv.org/pdf/1609.03637v1.pdf

[36] J. K. Ahn *et al.*, "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.*, vol. 122, p. 021802, Jan 2019. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.122.021802

[37] J. Comfort *et al.*, "Proposal for $K_L^0 \to \pi^0\nu\bar{\nu}$ Experiment at J-PARC," 2006. [Online]. Available: http://koto.kek.jp/pub/p14.pdf

[38] S. Nagamiya, "Introduction to J-PARC," *Prog. Theor. Exp. Phys.*, vol. 2012, no. 1, 10 2012, 02B001. [Online]. Available: https://doi.org/10.1093/ptep/pts025

[39] M. Ikegami, "Beam commissioning and operation of the J-PARC linac," *Prog. Theor. Exp. Phys.*, vol. 2012, no. 1, 09 2012, 02B002. [Online]. Available: https://doi.org/10.1093/ptep/pts019

[40] H. Hotchi *et al.*, "Beam commissioning and operation of the Japan Proton Accelerator Research Complex 3-GeV rapid cycling synchrotron," *Prog. Theor. Exp. Phys.*, vol. 2012, no. 1, 09 2012, 02B003. [Online]. Available: https://doi.org/10.1093/ptep/pts021

[41] T. Koseki *et al.*, "Beam commissioning and operation of the J-PARC main ring synchrotron," *Prog. Theor. Exp. Phys.*, vol. 2012, no. 1, 12 2012, 02B004. [Online]. Available: https://doi.org/10.1093/ptep/pts071

[42] H. Takahashi *et al.*, "Construction and beam commissioning of Hadron Experimental Hall at J-PARC," *J. Phys. Conf. Ser.*, vol. 312, no. 5, p. 052027, Sep 2011. [Online]. Available: https://doi.org/10.1088%2F1742-6596%2F312%2F5%2F052027

[43] H. Takahashi, K. Agari, K. Aoki *et al.*, "Indirectly water-cooled production target at J-PARC Hadron facility," *J. Radioanal. Nucl. Chem.*, vol. 305, pp. 803–809, 2015. [Online]. Available: https://doi.org/10.1007/s10967-015-3940-9

[44] T. Masuda *et al.*, "Long-lived neutral-kaon flux measurement for the KOTO experiment," *Prog. Theor. Exp. Phys.*, vol. 2016, no. 1, 01 2016, 013C03. [Online]. Available: https://doi.org/10.1093/ptep/ptv171

[45] T. Shimogawa, "Design of the neutral K_L^0 beamline for the KOTO experiment," *Nucl. Instrum. Methods Phys. Res.*, vol. 623, pp. 585–587, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016890021000642X?via%3Dihub

[46] K. Sato, "Measurement of the CsI calorimeter performance and K_L momentum spectrum for the J-PARC KOTO experiment," Ph.D. dissertation, Osaka University, 2015.

[47] Y. Maeda, "Search for the Decay $K_L^0 \to \pi^0\nu\bar{\nu}$ with a Neutron-Intensive GeV-Energy Photon Detector," Ph.D. dissertation, Kyoto University, 2016.

[48] Y. Yanagida, "Reusing KTeV CsI crystals for J-PARC KOTO experiment," *Proceedings, International Conference on Kaon Physics (KAON09): Tsukuba, Japan, June 09-12, 2009*, vol. 083, 2010. [Online]. Available: https://pos.sissa.it/083/021

[49] K. Sato, E. Iwai, K. Shiomi, Y. Sugiyama, M. Togawa, and T. Yamanaka, "Performance of KOTO CsI Calorimeter," *Proceedings of the 2nd International Symposium on Science at J-PARC — Unlocking the Mysteries of Life, Matter and the Universe —*. [Online]. Available: https://journals.jps.jp/doi/abs/10.7566/JPSCP.8.024007

[50] Y. Tajima *et al.*, "Barrel photon detector of the KEK $K_L^0 \to \pi^0\nu\bar{\nu}$ experiment," *Nucl. Instrum. Methods Phys. Res.*, vol. 592, pp. 261–272, 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900208006505

[51] M. Togawa, "A new cylindrical photon-veto detector for the KOTO experiment," *J. Phys. Conf. Ser.*, vol. 800, p. 012043, Jan 2017. [Online]. Available: https://doi.org/10.1088%2F1742-6596%2F800%2F1%2F012043

[52] N. Kawasaki, "Gamma Neutron Counter with WLSFiber Readout of undoped CsI crystals for KOTO Experiment," *Proceedings, International Conference on Kaon Physics (KAON13): Ann Arbor, MI, USA, 29 April - 1 May, 2013*, vol. 181, 2013. [Online]. Available: https://pos.sissa.it/181/040

[53] D. Naito *et al.*, "Development of a low-mass and high-efficiency charged-particle detector," *Prog. Theor. Exp. Phys.*, vol. 2016, no. 2, 02 2016, 023C01. [Online]. Available: https://doi.org/10.1093/ptep/ptv191

[54] S. Su, "Search for the Neutral Long-Lived Kaon Rare Decay $K_L^0 \to \pi^0\nu\bar{\nu}$," Ph.D. dissertation, University of Michigan, 2019.

[55] T. Matsumura *et al.*, "Photon-veto counters at the outer edge of the endcap calorimeter for the KOTO experiment," *Nucl. Instrum. Methods Phys. Res.*, vol. 795, pp. 19–31, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900215006762?via%3Dihub

[56] K. Nakagiri, "Search for the Decay $K_L^0 \to \pi^0\nu\bar{\nu}$ at the J-PARC KOTO Experiment," Ph.D. dissertation, Kyoto University, 2019.

[57] Y. Maeda *et al.*, "An aerogel Cherenkov detector for multi-GeV photon detection with low sensitivity to neutrons," *Prog. Theor. Exp. Phys.*, vol. 2015, no. 6, 06 2015, 063H01. [Online]. Available: https://doi.org/10.1093/ptep/ptv074

[58] S. Shinohara, "Beam-edge Photon Detector with Low Sensitivity to Neutrons for the KOTO Experiment," *J. Phys. Conf. Ser.*, vol. 800, p. 012044, Jan 2017. [Online]. Available: https://doi.org/10.1088%2F1742-6596%2F800%2F1%2F012044

[59] C. Carruth, "Monte Carlo Study of a Level II Trigger Cut for the K0TO Experiment," undergraduate honors thesis, University of Michigan, 2013.

[60] C. Lin *et al.*, "The Data-Acquisition System with Cluster-Finding Trigger at the J-PARC KOTO Experiment," 2019, *Poster presented at the International Conference on Kaon Physics 2019, Perugia, Italy, Sept. 10-13.*

[61] Japan Atomic Energy Agency, "Accelerator Facility Accident Report," May 2013. [Online]. Available: https://j-parc.jp/en/topics/HDAccident20130531.pdf

[62] K. Gorai *et al.*, "J-PARC Annual Report 2017," 2018. [Online]. Available: http://j-parc.jp/documents/annual_report/a_report_2017.pdf

[63] T. Masuda, "Development and Experimental Study of the KOTO Detector System using Three K_L Neutral Decay Modes," Ph.D. dissertation, Kyoto University, 2014.

[64] C. Lin *et al.*, "The Data-Acquisition System of the KOTO Experiment," *Proceedings of Science, ICHEP2020: Virtual, July 28 - Aug 6, 2020*, vol. 390, 2020. [Online]. Available: https://pos.sissa.it/390/914/

[65] B. Beckford, "Progress on the search for $K_L^0 \to \pi^0\nu\bar{\nu}$ with the KOTO detector," 2019, HEP Seminar, University of Michigan, Nov 13-14, 2019.

[66] S. Agostinelli *et al.*, "Geant4—a simulation toolkit," *Nucl. Instrum. Methods Phys. Res.*, vol. 506, no. 3, pp. 250 – 303, 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900203013688

[67] J. Allison *et al.*, "Geant4 developments and applications," *IEEE Transactions on Nuclear Science*, vol. 53, no. 1, pp. 270–278, 2006.

[68] J. Allison *et al.*, "Recent developments in Geant4," *Nucl. Instrum. Methods Phys. Res.*, vol. 835, pp. 186 – 225, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900216306957

[69] R. D. Cousins and V. L. Highland, "Incorporating systematic uncertainties into an upper limit," *Nucl. Instrum. Methods Phys. Res.*, vol. A320, pp. 331–335, 1992. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/0168900292907945

[70] The NA62 Collaboration, E. Cortina Gil, A. Kleimenova *et al.*, "An investigation of the very rare $K^+ \to \pi^+ \nu\bar{\nu}$ decay," *J. High Energy Phys.*, vol. 2020, no. 42, 2020. [Online]. Available: https://doi.org/10.1007/JHEP11(2020)042

[71] X. G. He, X. D. Ma, J. Tandean, and G. Valencia, "Breaking the Grossman-Nir bound in kaon decays," *J. High Energy Phys.*, vol. 2020, no. 57, 2020. [Online]. Available: https://doi.org/10.1007/JHEP04(2020)057

[72] X. G. He, X. D. Ma, J. Tandean, and G. Valencia, "Evading the Grossman-Nir bound with $\Delta I = 3/2$ new physics," *J. High Energy Phys.*, vol. 2020, no. 34, 2020. [Online]. Available: https://doi.org/10.1007/JHEP08(2020)034

[73] R. Ziegler, J. Zupan, and R. Zwicky, "Three exceptions to the Grossman-Nir bound," *J. High Energy Phys.*, vol. 2020, no. 229, 2020. [Online]. Available: https://doi.org/10.1007/JHEP07(2020)229

[74] T. Kitahara, T. Okui, G. Perez, Y. Soreq, and K. Tobioka, "New Physics Implications of Recent Search for $K_L \to \pi^0 \nu\bar{\nu}$ at KOTO," *Phys. Rev. Lett.*, vol. 124, p. 071801, Feb 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.124.071801

[75] D. Egana-Ugrinovic, S. Homiller, and P. Meade, "Light Scalars and the Koto Anomaly," *Phys. Rev. Lett.*, vol. 124, p. 191801, May 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.124.191801

[76] D. O'Connell, M. J. Ramsey-Musolf, and M. B. Wise, "Minimal extension of the standard model scalar sector," *Phys. Rev. D*, vol. 75, p. 037701, Feb 2007. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.75.037701

[77] K. Fuyuto, W.-S. Hou, and M. Kohda, "Loophole in $K \to \pi\nu\overline{\nu}$ Search and New Weak Leptonic Forces," *Phys. Rev. Lett.*, vol. 114, p. 171802, Apr 2015. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.114.171802

[78] Y. Jho *et al.*, "Light gauge boson interpretation for $(g\text{-}2)_\mu$ and the $K_L \to \pi^0$ + (invisible) anomaly at the J-PARC KOTO experiment," *J. High Energy Phys.*, vol. 2020, no. 86, 2020. [Online]. Available: https://doi.org/10.1007/JHEP04(2020)086

[79] W. Altmannshofer, B. V. Lehmann, and S. Profumo, "Cosmological implications of the KOTO excess," *Phys. Rev. D*, vol. 102, p. 083527, Oct 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.102.083527

[80] P. S. B. Dev, R. N. Mohapatra, and Y. Zhang, "Constraints on long-lived light scalars with flavor-changing couplings and the KOTO anomaly," *Phys. Rev. D*, vol. 101, p. 075014, Apr 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.101.075014

[81] J. Liu, N. McGinnis, C. Wagner, and X. Wang, "A light scalar explanation of $(g\text{-}2)_\mu$ and the KOTO anomaly," *J. High Energy Phys.*, vol. 2020, no. 197, 2020. [Online]. Available: https://doi.org/10.1007/JHEP04(2020)197

[82] Y. Liao, H.-L. Wang, C.-Y. Yao, and J. Zhang, "Imprint of a new light particle at KOTO?" *Phys. Rev. D*, vol. 102, p. 055005, Sep 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.102.055005

[83] J. Cline, M. Puel, and T. Toma, "A little theory of everything, with heavy neutral leptons," *J. High Energy Phys.*, vol. 2020, no. 39, 2020. [Online]. Available: https://doi.org/10.1007/JHEP05(2020)039

[84] P. Ballett, M. Hostert, and S. Pascoli, "Dark neutrinos and a three-portal connection to the standard model," *Phys. Rev. D*, vol. 101, p. 115025, Jun 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.101.115025

[85] A. Datta, S. Kamali, and D. Marfatia, "Dark sector origin of the KOTO and MiniBooNE anomalies," *Phys. Lett. B*, vol. 807, p. 135579, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S037026932030383X

[86] S. Gori, G. Perez, and K. Tobioka, "KOTO vs. NA62 dark scalar searches," *J. High Energy Phys.*, vol. 2020, no. 110, 2020. [Online]. Available: https://doi.org/10.1007/JHEP08(2020)110

[87] M. Fabbrichesi and E. Gabrielli, "Dark-sector physics in the search for the rare decays $K^+ \to \pi^+ \nu\bar{\nu}$ and $K_L \to \pi^0 \nu\bar{\nu}$," *Eur. Phys. J. C*, vol. 2020, no. 532, 2020. [Online]. Available: https://doi.org/10.1140/epjc/s10052-020-8103-7

[88] M. Hostert, K. Kaneta, and M. Pospelov, "Pair production of dark particles in meson decays," *Phys. Rev. D*, vol. 102, p. 055016, Sep 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.102.055016

[89] L. Calibbi, A. Crivellin, F. Kirk, C. A. Manzari, and L. Vernazza, "Z' models with less-minimal flavor violation," *Phys. Rev. D*, vol. 101, p. 095003, May 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.101.095003

[90] J. Aebischer, A. Buras, and J. Kumar, "Another SMEFT story: Z' facing new results on ϵ'/ϵ, $_K$ and $K \to \pi\nu\bar{\nu}$," *J. High Energy Phys.*, vol. 2020, no. 97, 2020. [Online]. Available: https://doi.org/10.1007/JHEP12(2020)097

[91] B. Dutta, S. Ghosh, and T. Li, "Explaining $(g-2)_{\mu,e}$, the KOTO anomaly, and the MiniBooNE excess in an extended Higgs model with sterile

neutrinos," *Phys. Rev. D*, vol. 102, p. 055017, Sep 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.102.055017

[92] P. Lichard, "Pionium as a source of false events in the $K \to \pi\nu\overline{\nu}$ decays," *Phys. Rev. D*, vol. 102, p. 113005, Dec 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.102.113005

[93] T. Li, X.-D. Ma, and M. A. Schmidt, "Implication of $K \to \pi\nu\overline{\nu}$ for generic neutrino interactions in effective field theories," *Phys. Rev. D*, vol. 101, p. 055019, Mar 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.101.055019

[94] F. F. Deppisch, K. Fridell, and J. Harz, "Constraining lepton number violating interactions in rare kaon decays," *J. High Energy Phys.*, vol. 2020, no. 186, 2020. [Online]. Available: https://doi.org/10.1007/JHEP12(2020)186

[95] X. Liu, Y. Li, T. Li, and B. Zhu, "The light sgoldstino phenomenology: explanations for the muon (g - 2) deviation and KOTO anomaly," *J. High Energy Phys.*, vol. 2020, no. 197, 2020. [Online]. Available: https://doi.org/10.1007/JHEP12(2020)197

[96] C. Sagan, *Pale Blue Dot: A Vision of the Human Future in Space*. New York: Random House, 1994.

[97] K. Kotera, "Construction and performance of 4-D CsI calorimeter for the $K_L \to \pi^0\nu\bar{\nu}$ search of KOTO experiment," *Proceedings of Science, ICHEP2020: Virtual, July 28 - Aug 6, 2020*, vol. 390, 2020. [Online]. Available: https://pos.sissa.it/390/749/

[98] H. Kim, "A new charged particle detector for the KOTO experiment at J-PARC," *J. Phys. Conf. Ser.*, vol. 1526, p. 012032, Apr 2020. [Online]. Available: https://doi.org/10.1088/1742-6596/1526/1/012032

[99] H. Watanabe *et al.*, "Development of a New Production Target at the J-PARC Hadron Experimental Facility," *Proceedings of the 14th International Workshop on Spallation Materials Technology*. [Online]. Available: https://journals.jps.jp/doi/abs/10.7566/JPSCP.28.041004

[100] T. Nomura, "A future $K_L \to \pi^0\nu\bar{\nu}$ experiment at J-PARC," *J. Phys. Conf. Ser.*, vol. 1526, p. 012027, Apr 2020. [Online]. Available: https://doi.org/10.1088/1742-6596/1526/1/012027

[101] M. Moulson, "Prospects for an experiment to measure BR($K_L \to \pi^0\nu\bar{\nu}$) at the CERN SPS," *J. Phys. Conf. Ser.*, vol. 800, no. 1, p. 012037, 2017. [Online]. Available: https://arxiv.org/pdf/1611.04864.pdf

[102] J. Charles *et al.*, "Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics," *Phys. Rev.*, vol. D91, no. 7, p. 073007, 2015.

[103] R. Aaij *et al.*, "Measurement of CP violation in $B^0 \to J/\psi K_s^0$ and $B^0 \to \psi(2S)K_s^0$ decays," *J. High Energy Phys.*, vol. 11, no. 170, 2017. [Online]. Available: https://cds.cern.ch/record/2284162/files/LHCb-PAPER-2017-029.pdf

[104] R. Aaij *et al.*, "Test of Lepton Universality Using $B^+ \to K^+\ell^+\ell^-$ Decays," *Phys. Rev. Lett.*, vol. 113, p. 151601, Oct 2014. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.113.151601

[105] J.-T. Wei *et al.*, "Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for $B \to K^{(*)}l^+l^-$," *Phys. Rev. Lett.*, vol. 103, p. 171801, Oct 2009. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.103.171801

[106] J. P. Lees *et al.*, "Measurement of branching fractions and rate asymmetries in the rare decays $B \to K^{(*)}\ell^+\ell^-$," *Phys. Rev. D*, vol. 86, p. 032012, Aug 2012. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.86.032012

[107] P.A. Zyla *et al.*, "Review of Particle Physics," *Prog. Theor. Exp. Phys.*, vol. 2020, no. 8, 2020, 083C01. [Online]. Available: https://doi.org/10.1093/ptep/ptaa104