Searches for rare and forbidden kaon decays at the NA62 experiment at CERN

By

Joel C. Swallow

A thesis submitted to the University of Birmingham
for the degree of *Doctor Of Philosophy*

UNIVERSITY OF
BIRMINGHAM

Particle Physics Research Group
School of Physics and Astronomy
College of Engineering and Physical Sciences
University of Birmingham
February 2021

Abstract

Studies of rare and forbidden K^+ decays are presented based on analysis of data taken by the NA62 experiment between 2016 and 2018 (Run 1).

The primary physics objective of the NA62 experiment is the study of the ultra-rare decay $K^+ \to \pi^+\nu\bar{\nu}$. Studies of backgrounds are presented and the results of the Run 1 $K^+ \to \pi^+\nu\bar{\nu}$ analysis are summarised.

A search for the $K^+ \to \pi^+ X$ decay, where X is a long-lived feebly interacting particle, is performed through an interpretation of the $K^+ \to \pi^+\nu\bar{\nu}$ analysis of 2017 data. Upper limits are established on the branching ratio, $\mathcal{B}(K^+ \to \pi^+ X)$ for ranges of X masses, 0–110 MeV/c^2 and 154–260 MeV/c^2, and lifetimes above 100 ps.

Searches for the lepton number violating $K^+ \to \pi^-\mu^+e^+$ decay and lepton flavour violating $K^+ \to \pi^+\mu^-e^+$ and $\pi^0 \to \mu^+e^+$ decays are described. No signals are observed and upper limits on the branching ratios of these decays at 90% confidence level are obtained: $\mathcal{B}(K^+ \to \pi^-\mu^+e^+) < 4.2 \times 10^{-11}$, $\mathcal{B}(K^+ \to \pi^+\mu^-e^+) < 6.6 \times 10^{-11}$, and $\mathcal{B}(\pi^0 \to \mu^-e^+) < 3.2 \times 10^{-10}$, improving by an order of magnitude on previous results.

Acknowledgements

This thesis would not have been possible without the tuition of, and fruitful collaboration with, my colleagues and it is a pleasure to identify just a few of them in particular.

First I would like to thank my supervisors Evgueni Goudzovski and Cristina Lazzeroni from who I have learned a great deal.

I extend warm thanks to a few of my NA62 collaborators in particular with whom I have worked closely: Giuseppe Ruggiero and Rado Marchevski for the $K^+ \to \pi^+\nu\bar{\nu}$ analysis and Elisa Minucci for the studies of lepton number and lepton flavour violation in K^+ decays. In addition I acknowledge my appreciation for the time and effort taken by colleagues and reviewers of the analyses presented in this thesis, for asking the difficult questions which prompted improvement and consolidation. For discussions of countless drafts and re-drafts of various documents I thank Evgueni, Cristina, Brigitte Bloch-Devaux and John Fry.

I have been lucky enough to work in great environments at the University of Birmingham and CERN, so to everybody who made these last few years entertaining in and out of the office, thanks.

Finally, to my friends, family and Henriette, those who have been with me for this PhD journey, what can I say? In the absence of words sufficiently fitting my gratitude for your support, I will simply say thank you.

Contents

List of Figures

List of Tables

Declaration of Author's Contribution

The theoretical foundations of my work are presented in chapter 2 to provide context. The NA62 detector, used to collect the data analysed for results reported, is described in chapter 3. My contribution to detector studies is presented in chapter 4, showing simulation studies of the KTAG detector.

The analysis of the $K^+ \to \pi^+\nu\bar{\nu}$ decay, presented in chapter 5, represents the collaborative efforts of a group of people. My personal contributions to this, directly related to analysis of 2016 and 2017 NA62 data, are presented in sections 5.2, 5.3[1] and 5.5. The (preliminary) final result for the full 'Run 1' NA62 data-set (taken from 2016-18) is reported in section 5.7.

The search for $K \to \pi^+X$ decays presented in chapter 6 represents my work. An independent analysis (not my work) with the same goal was performed in parallel and in section 6.4 the results are compared and found to be in good agreement. The preliminary results of this analysis were first presented by the collaboration at the Phenomenology 2020 Symposium (PHENO20) in May 2020 [1] and reported by me at an invited talk at the FIPs workshop in September 2020 [2]. Final results are described in a paper accepted for publication in JHEP [3].

The search for $K^+ \to \pi\mu e$ decays presented in chapter 7 represents my work. Again an independent analysis was performed with the same objective and found to be in good agreement. The preliminary result was first presented at the ICHEP20 conference on July the 29^{th} 2020 by me [4] and the final results are reported in a paper to be submitted in February 2021 for publication in JHEP [5].

[1] I was responsible for simulation studies reported in this section (and appendix C) which are compared to data studies by collaborators and integrated into the results.

Chapter 1

Introduction

The Standard Model of particle physics (SM) represents the best description currently available of the fundamental constituents of the universe and their interactions. It stands as an outstanding achievement of science, developed over decades and its predictions tested over a century. The SM includes the most precise testable, and experimentally confirmed, predictions of the workings of nature. However, it is not enough to explain all of our experimental results and therefore is not a complete model of physics.

The goal of modern particle physics is to push the SM to breaking point, to test its predictions as precisely as possible and, hopefully, in the way its limitations are exposed a discovery of the underlying theoretical framework will be revealed. The search for this 'New Physics' (NP), beyond the standard model (BSM), involves theoretical and phenomenological predictions of potential differences between the SM and BSM scenarios which may be tested experimentally, and the corresponding experimental efforts. At this moment in time no clear theoretical consensus exists and now, perhaps more than ever, experimental searches are vital to unlocking the next advance in our knowledge of fundamental physics. These experimental searches may take the form of direct searches for new particles, either at the highest energies or for weakly interacting particles in very large datasets, or through precision tests of SM predictions which may be modified by the presence of additional interactions or particles.

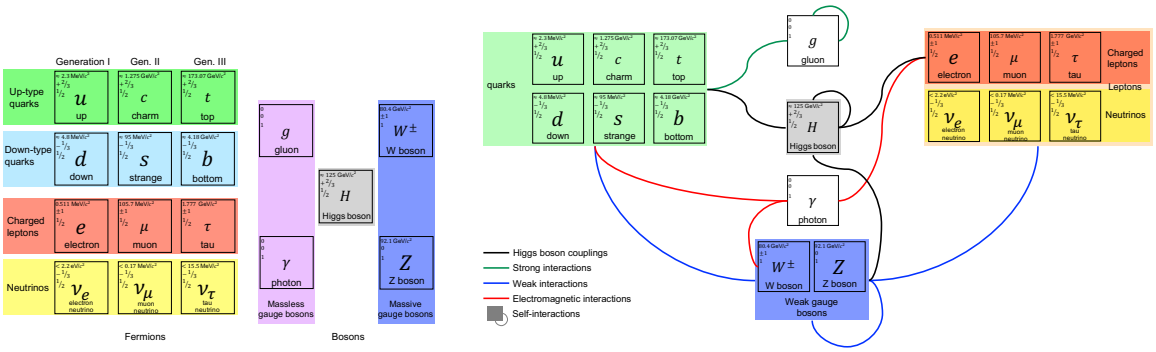
In chapter 2 a summary of the fundamentals of the SM will be presented in the context of the search for NP in rare and forbidden K^+ decays. Potential BSM scenarios which can be tested by these searches are introduced. The NA62 experiment at the European Organization for Nuclear Research (Conseil Européen pour la Recherche Nucléaire, CERN) is described in chapter 3. Specific searches for NP in precision tests of the standard model in rare kaon decays, decays to an exotic particle not in the SM, and forbidden lepton number violation and/or lepton flavour number violating decays, are presented in chapters 5, 6 and 7 respectively.

Chapter 2

Theoretical Foundations

2.1 The Standard Model of Particle Physics

The Standard Model of particle physics [6] [7] [8] is a mathematically self-consistent renormalisable gauge quantum field theory description of fundamental fields (usually depicted as corresponding to 18 particles: 12 spin-1/2 fermions, 5 spin-1 bosons and the spin-0 Higgs boson) and three interactions (Electromagnetic, weak and strong), illustrated in figure 2.1. The SM is constructed from symmetry principles and has symmetry group $SU(3)_c \times SU(2)_L \times U(1)_Y$, with unbroken non-Abelian gauge group $SU(3)_c$ related to Quantum ChromoDynamics (QCD) and strong interactions and spontaneously broken $SU_2 \times U(1)_Y$ local gauge symmetry associated with electroweak interactions. The Higgs field is responsible for this spontaneous symmetry breaking and allows a mechanism through which the associated electroweak gauge bosons ($W^{\pm}$ and Z^0) acquire masses. There are two chiralities of fermions, left-handed (with associated $SU(2)$ doublets) and right-handed ($SU(2)$ singlets) and to preserve the $SU_2 \times U(1)_Y$ symmetry a Yukawa coupling to the Higgs field is introduced, generating fermion masses as free parameters of the SM.



(a) SM particles. (b) Interactions between SM particles.

Figure 2.1: Fundamental particles of the SM and their interactions.

2.1.1 Flavour Structure of the SM

In the SM only left-handed chiral states are subject to vector − axial vector ($V-A$) charged current interactions (mediated by $W^{\pm}$ bosons). This leads to violation of Parity (P) conservation, as first measured experimentally in beta decays [9]. In the quark sector there exists a left-handed doublet under $SU(2)_L$

$$Q_L = \begin{pmatrix} U_L \\ D_L \end{pmatrix} , \tag{2.1.1}$$

(where U_L and D_L are spinors representing up and down type quarks respectively) with hypercharge $Y_Q = 1/3$, and right-handed singlets U_R and D_R with $Y_U = 4/3$ and $Y_D = -2/3$ respectively. After symmetry breaking the quark mass (strong) and interaction (weak) eigenstates do not coincide. The original two-generation formulation of this, the Cabibbo hypothesis [10], is parameterised by a unitarity matrix with controlling parameter $\theta_c \approx 13°$, the Cabibbo angle,

$$\underbrace{\begin{pmatrix} d' \\ s' \end{pmatrix}}_{\text{weak eigenstates}} = \underbrace{\begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix}}_{\text{Cabibbo matrix}} \underbrace{\begin{pmatrix} d \\ s \end{pmatrix}}_{\text{mass eigenstates}} . \tag{2.1.2}$$

By convention diagonalization of the Cabibbo mixing matrix is performed such that the down-type quark eigenstates are rotated. This two-generation formalism led to the prediction of the charm quark through the GIM mechanism [11] which also results in suppression of the rate of rare decays. Including third generation quarks into this mixing paradigm [12] results in the CKM (Cabibbo-Kobayashi-Maskawa) complex unitary matrix V_{CKM}, parameterised by four physical parameters. Equation 2.1.2 therefore becomes

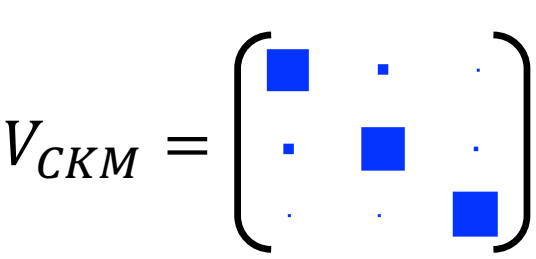
$$\underbrace{\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}}_{\text{weak eigenstates}} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{\text{CKM matrix}} \underbrace{\begin{pmatrix} d \\ s \\ b \end{pmatrix}}_{\text{mass eigenstates}} . \tag{2.1.3}$$

The relative size of the CKM matrix elements determines the coupling strength entering the matrix element at a quark vertex. The off-diagonal elements representing couplings between generations are small and lead to suppression of decay rates (CKM suppression). The current experimentally determined magnitudes for the matrix elements are [13]

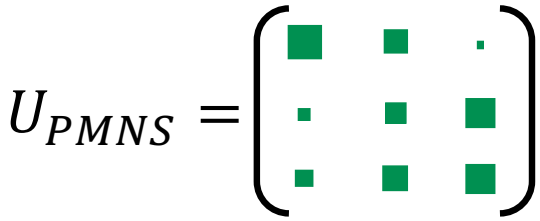
$$\begin{pmatrix} |V_{ud}| = 0.97420 \pm 0.00021 & |V_{us}| = 0.2243 \pm 0.0005. & |V_{ub}| = (3.94 \pm 0.36) \times 10^{-3} \\ |V_{cd}| = 0.218 \pm 0.004 & |V_{cs}| = 0.997 \pm 0.017 & |V_{cb}| = 0.0422 \pm 0.0008 \\ |V_{td}| = (8.1 \pm 0.5) \times 10^{-3} & |V_{ts}| = 0.0394 \pm 0.0023 & |V_{tb}| = 1.019 \pm 0.025 \end{pmatrix} . \tag{2.1.4}$$

The hierarchical structure of the matrix, displayed schematically in figure 2.2a, is clear when writing V_{CKM} in the Wolfenstein parameterisation (up to 4^{th} order in λ) as

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ \lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) , \tag{2.1.5}$$



(a) CKM Matrix, box sizes proportional to measured magnitude of elements [13].



(b) PMNS matrix, box sizes approximately the size of the central value of current 95% confidence interval bounds [15].

Figure 2.2: Schematic illustration of the relative magnitude of mixing matrix elements.

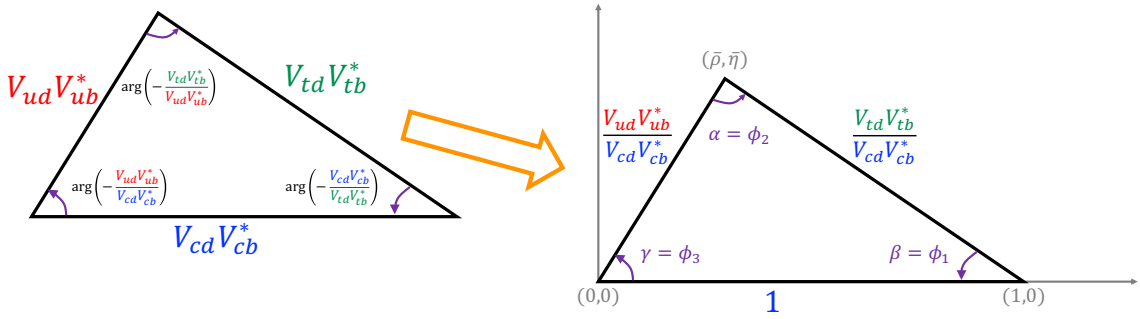


Figure 2.3: The unitarity triangle (UT).

where $\lambda = \sin\theta_c = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2+|V_{us}|^2}}$, $A = \frac{1}{\lambda}\frac{|V_{cb}|}{|V_{us}|}$ and $\bar{\rho}+i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}$ and where $\bar{\rho}$ and $\bar{\eta}$ are defined as $\bar{\rho} = \rho(1-\frac{\lambda^2}{2}+\mathcal{O}(\lambda^4)+...)$ and $\bar{\eta} = \eta(1-\frac{\lambda^2}{2}+\mathcal{O}(\lambda^4)+...)$. From the unitarity relations, $\sum_i V_{ij}V_{ik}^* = \delta_{jk}$ and $\sum_j V_{ij}V_{jk}^* = \delta_{ik}$, the six cases equal to zero allow a closed triangle to be defined in the complex plane where side lengths are given by $V_{ij}V_{ik}^*$ and with area equal to the Jarlskog invariant [14], $J = \frac{\mathrm{Im}[V_{ij}V_{kl}V_{il}^*V_{kj}^*]}{\sum_{m,n}\varepsilon_{ikm}\varepsilon_{jln}}$. Of particular interest are the cases where all three sides are of order λ^3 and often the case of

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{2.1.6}$$

is used to define 'the' unitarity triangle as shown in figure 2.3.

In the lepton sector a similar picture emerges with the chiral left-handed electron and electron neutrino forming a doublet under $SU(2)_L$

$$\chi_L = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, \tag{2.1.7}$$

with hypercharge $Y_L = -1$. There exists an analogous right handed singlet for charged leptons, for example e_R for the electron, with hypercharge $Y_R = -2$ (and therefore electric charge $Q = -1$ for both electron components). As in the quark sector, the weak interaction eigenstates and mass eigenstates do not coincide and here this is parameterised by

the PMNS mixing matrix (U_{PMNS}) [16] [17] [18] with

$$\underbrace{\begin{pmatrix}\nu_e\\ \nu_\mu\\ \nu_\tau\end{pmatrix}}_{\text{weak eigenstates}} = \underbrace{\begin{pmatrix}U_{e1} & U_{e2} & U_{e3}\\ U_{\mu1} & U_{\mu2} & U_{\mu3}\\ U_{\tau1} & U_{\tau2} & U_{\tau3}\end{pmatrix}}_{\text{PMNS matrix}} \underbrace{\begin{pmatrix}\nu_1\\ \nu_2\\ \nu_3\end{pmatrix}}_{\text{mass eigenstates}} . \tag{2.1.8}$$

The weak eigenstates take part in interactions, and by definition the electron neutrino, ν_e, is the neutrino that emerges in conjunction with a positron in a charged current interaction. However, during propagation neutrinos exist as a coherent linear superposition of the mass eigenstates and it is only when a subsequent interaction occurs that the wavefunction collapses and a specific weak state is expressed. This property, where a neutrino of one flavour can be produced but can later interact with another flavour, is designated as 'neutrino oscillations' and is only possible if there is a mass difference between each pair of neutrinos (and therefore only if at least two neutrino flavours have mass). The mixing matrix U_{PMNS} can be written as

$$U_{PMNS} = \begin{pmatrix}1 & 0 & 0\\ 0 & c_{23} & s_{23}\\ 0 & -s_{23} & c_{23}\end{pmatrix}\begin{pmatrix}c_{13} & 0 & s_{13}e^{i\delta}\\ 0 & 1 & 0\\ -s_{13}e^{i\delta} & 0 & c_{13}\end{pmatrix}\begin{pmatrix}c_{12} & s_{12} & 0\\ -s_{12} & c_{12} & 0\\ 0 & 0 & 1\end{pmatrix}, \tag{2.1.9}$$

where $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$ and the parameterisation highlights the three mixing angles θ_{12}, θ_{23} and θ_{13} and the complex phase δ. The hierarchy in U_{PMNS} is very different to V_{CKM}, as shown in figure 2.2b. Strictly within the standard model there is no mechanism for neutrinos to obtain mass and therefore a minimal extension of the SM is required to describe this behaviour, see section 2.2.2.

The SM has 19 parameters, four from the electroweak scalar sector which may be mapped to physical observables:

Z boson mass:	$m_Z = 92.1\,\text{GeV}/c^2$,	(2.1.10)
Higgs boson mass:	$m_H = 125\,\text{GeV}/c^2$,	(2.1.11)
Fine structure (EM coupling) constant	$\alpha \equiv \frac{e^2}{4\pi} \simeq \frac{1}{137}$,	(2.1.12)
Weinberg (weak mixing) angle, with	$\sin^2\theta_W = 0.231$,	(2.1.13)

one from strong interactions

$$\text{strong coupling constant} \quad g_s\,, \tag{2.1.14}$$

nine masses (6 quarks plus 3 charged leptons) with 4 parameters from the CKM matrix (3 mixing angles plus one complex phase) and finally a QCD phase θ_{QCD}.

2.1.2 Emergent Symmetries

In addition to the gauge symmetry principles used to construct the SM it is found to have other (global) symmetries. These are not required for its construction but apparently

emerge naturally from it. The Noether theorem [19] then implies conserved quantities. First, the SM Lagrangian is invariant under simultaneous phase rotation of all quarks (with corresponding opposite rotation for antiquarks), with corresponding conservation of baryon number:

$$B = \frac{1}{3}(N_q - N_{\bar{q}}) = N_{\text{baryons}} - N_{\text{anti-baryons}} \ . \tag{2.1.15}$$

As a consequence the lightest baryon, the proton, is predicted to be stable. Similarly, conservation of lepton number is enforced with

$$L = N_\ell - N_{\bar{\ell}} \, . \tag{2.1.16}$$

Moreover, in the lepton sector without neutrino masses no mixing between lepton generations is possible and so individual lepton flavour/family number conservation is enforced as

$$L_\ell = N_\ell + N_{\nu_\ell} - (N_{\bar{\ell}} + N_{\bar{\nu_\ell}}) \, , \tag{2.1.17}$$

with $\ell = e, \, \mu, \, \tau$.

2.2 Beyond the Standard Model

2.2.1 The Limits of the SM and Requirement for BSM Physics

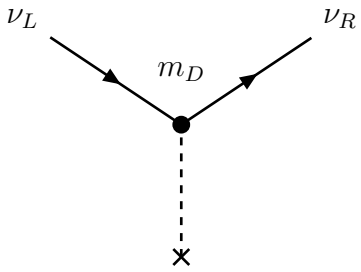
The SM is not able to explain all experimental observations and is therefore not a complete theory of nature. The SM does not describe all four fundamental forces, with the best description of gravity provided by general relativity. Apparently it does not describe all fundamental particles either, with the known presence of Dark Matter (DM) in the universe. Moreover, complete descriptions of interactions and decays may not be fully contained in the SM, as evidenced by the dramatic baryon asymmetry in the universe, explanation of which requires more CP violation than is present in the SM [1] [20] [21].

The phenomenon of neutrino oscillations, as introduced and defined in section 2.1.1, was demonstrated in neutrinos produced in the upper atmosphere [22] and the sun [23] [24] by 2002. This is the only completely blatant experimental proof of BSM physics to date and the implications of including neutrino masses is discussed in section 2.2.2. Sections 2.2.3 and 2.2.4 review relevant important known open issues in modern particle physics. A brief discussion of BSM model frameworks is given in sections 2.2.5 and 2.2.6. The current prominent hints for NP, the so-called 'anomalies', are then reviewed in section 2.2.7.

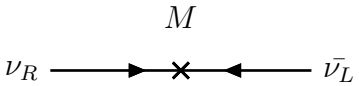
2.2.2 Standard Model Extensions for Neutrino Masses

No right handed neutrinos exist in the SM. The weak interaction $V-A$ only interacts with left-handed fermions and therefore in the minimal case no right handed neutrinos are predicted and left-handed neutrinos are massless. However, since neutrino oscillations prove

[1] The three 'Sakharov conditions' [20] which must be fulfilled to allow baryon asymmetry in the universe: baryon number violation, C and CP violation, and interactions taking place out of thermal equilibrium.



(a) Dirac neutrinos.



(b) Majorana neutrinos.

Figure 2.4: Interaction vertices for Dirac and Majorana neutrinos.

the existence of non-zero neutrino mass, terms must be added to the SM Lagrangian representing their mass. These mass terms can take two forms, Dirac or Majorana expressed respectively as

$$\mathcal{L}_D = -m_D(\bar{\nu_R}\nu_L + \bar{\nu_L}\nu_R)\,, \tag{2.2.1}$$

with Dirac mass m_D and

$$\mathcal{L}_M = -\frac{1}{2}M(\bar{\nu_R^c}\nu_R + \bar{\nu_R}\nu_R^c)\,, \tag{2.2.2}$$

where M is the Majorana mass and ν_R^c is the CP conjugate field for a right handed neutrino, corresponding to a left handed antineutrino. Corresponding interactions are illustrated in figure 2.4. Both respect the local gauge invariance of the SM, but the Majorana term implies a direct coupling of a particle and its antiparticle, in this way such 'Majorana neutrinos' may be considered to be their own antiparticle. In the most general scenario both Dirac and Majorana masses may be included

$$\mathcal{L}_{m_\nu} = -\frac{1}{2}\left[\underbrace{m_D\bar{\nu_L}\nu_R + m_D\bar{\nu_R^c}\nu_L^c}_{\text{Dirac}} + \underbrace{M\bar{\nu_R^c}\nu_R}_{\text{Majorana}}\right] + h.c.\ . \tag{2.2.3}$$

Physical neutrino states are then obtained from a diagonalized mass matrix $\mathbf{M} = \begin{pmatrix} 0 & m_D \\ m_D & M \end{pmatrix}$ (for each neutrino flavour) leading to physical mass

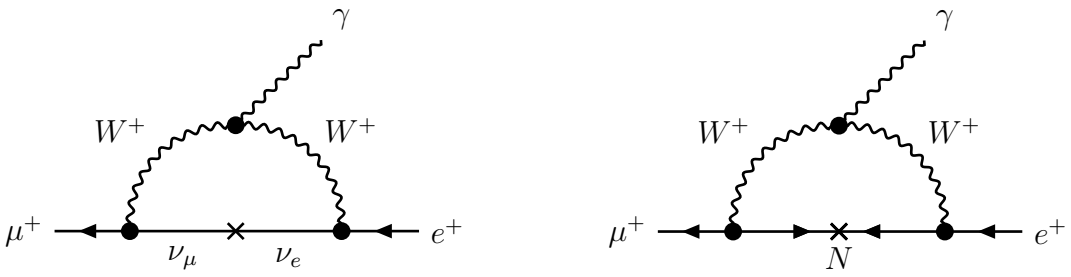
$$m_\nu = \frac{M \pm \sqrt{M^2 + 4m_D^2}}{2}\,. \tag{2.2.4}$$

The observation that the observed left handed neutrinos must have a mass several orders of magnitude smaller than the other fermions [13] may then be explained if the Dirac mass m_D is in fact of $\mathcal{O}(\mathrm{GeV}/c^2)$ but the Majorana mass M is very large, giving

$$m_{\nu_L} \approx \frac{m_D^2}{M} \sim \mathcal{O}(0.01\,\mathrm{eV}/c^2) \text{ and } m_N \approx M \sim \mathcal{O}(10^{11}\,\mathrm{GeV}/c^2)\,, \tag{2.2.5}$$

with heavy neutrino state(s) N. This is the (type I) 'see-saw' mechanism [25] [26] [27].

The introduction of new right-handed neutrinos, with Majorana mass terms and corresponding interactions, has potentially observable implications for particle interactions including a mechanism for lepton number violation (LNV), (charged) lepton flavour/family number violation ((C)LFV) and neutrinoless double beta decay ($0\nu\beta\beta$).



(a) Mediated via neutrino oscillations. (b) Mediated by Majorana neutrinos.

Figure 2.5: Feynman diagrams for CLFV decay $\mu^+ \to e^+\gamma$.

2.2.2.1 LFV and LNV

Mixing of the 3 known active massive left-handed neutrinos constitutes LFV and allows mediation of decay processes forbidden in the minimal SM, for example the CLFV $\mu^\pm \to e^\pm\gamma$ decays, see figure 2.5a. However, if this mixing is the only source of LFV such processes are highly suppressed, via the GIM mechanism, leading to unobservably low rates [13] [28]

$$\mathcal{B}(\mu^\pm \to e^\pm\gamma) = \frac{3\alpha}{32\pi}\left|\sum_i U_{\mu i}U_{ei}\frac{m_{\nu_i}^2 - m_{\nu_1}^2}{m_W^2}\right|^2 < 10^{-54}\,. \tag{2.2.6}$$

However, LNV and LFV processes are predicted in many BSM scenarios. For example, in this case Majorana neutrinos can act as a mediator, see figure 2.5b, and enhance the decay rate to $\mathcal{B}(\mu^\pm \to e^\pm\gamma) \sim 10^{-13}$ [28] or similar, depending on the construction of the model [29], which is within range of experimental sensitivity with current upper limit $\mathcal{B}(\mu^+ \to e^+\gamma) < 4.2 \times 10^{-13}$ (from the MEG experiment at PSI) [30] .

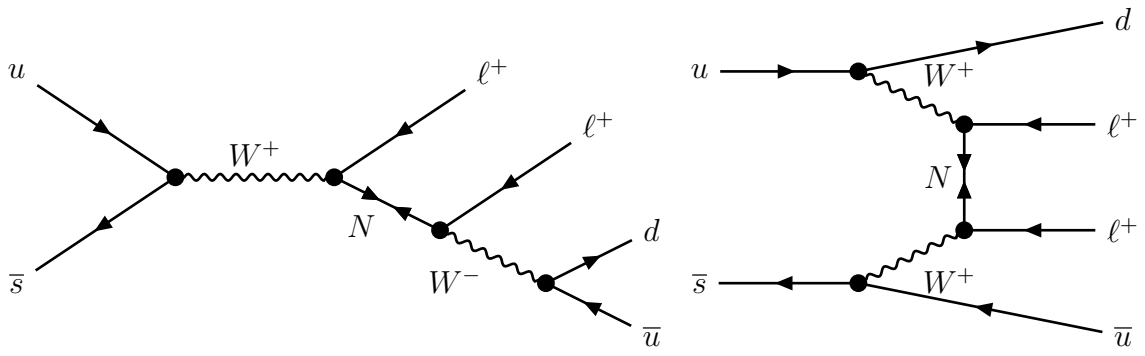
Majorana neutrino mediated interactions can also lead to a violation of total lepton number, for example in the decays $K^+ \to \pi^-\ell^+\ell^+$ with $\Delta L = 2$ (and $\Delta L_\ell = 2$), see figure 2.6. In the case where an on-shell Majorana neutrino (resonantly produced with $m_\pi + m_\ell < m_N < m_K - m_\ell$) the branching ratio may be enhanced to observable levels [31]. This scenario was specifically studied by the NA48/2 experiment for $\ell = \mu$ [32] [33] and improved upper limits on the branching ratio were recently reported by the NA62 collaboration (for $\ell = \mu$ and e) [34].

2.2.2.2 Neutrinoless Double Beta Decay

Certain even-even (A and Z even) radioactive nuclei have an intermediate state before the lowest energy level for which transition to is energetically forbidden via standard $\beta^\pm$ decay ($e^\pm$ emission), being at higher energy than the state before. However, a double beta decay can allow transition to the lowest stable energy state. In the SM this process always produces two companion neutrinos:

$$(Z, A) \to (Z + 2, A) + 2e^- + 2\bar{\nu}_e\,, \tag{2.2.7}$$

a rare process with half-life $\tau_{1/2} \approx 10^{19}$–$10^{25}$ years which has been observed. However, if neutrinos are Majorana particles then their exchange, similar to figure 2.6b, can lead to



(a) s-channel diagram.

(b) t-channel diagram.

Figure 2.6: Feynman diagrams for LNV decay $K^+ \to \pi^-\ell^+\ell^+$. The s-channel process alone contributes to the decay via mediation of an on-shell Majorana neutrino [31].

an equivalent neutrinoless double beta decay ($0\nu\beta\beta$) process

$$(Z, A) \to (Z + 2, A) + 2e^- \,, \tag{2.2.8}$$

which is experimentally distinct because the electrons must be monoenergetic. The theoretical $0\nu\beta\beta$ decay rates depend on the size of U_{PMNS} elements. Currently $0\nu\beta\beta$ decays have not been observed and most stringent upper limits are set by the GERDA experiment for the ^{76}Ge isotope at $\tau_{1/2} > 8.0 \times 10^{25}$ years [13] [35].

Existence of $0\nu\beta\beta$ decay requires a Majorana neutrino mass regardless of the specific model. The decay rate, $(\tau_{1/2})^{-1}$ is proportional to the square of the effective Majorana mass, $< m_{\beta\beta} >^2 = |\sum_i U_{ei}^2 m_{\nu_i}|^2$, [13]

$$(\tau_{1/2})^{-1} = G^{0\nu\beta\beta}|\mathcal{M}^{0\nu\beta\beta}|^2 < m_{\beta\beta} >^2 \tag{2.2.9}$$

where $G^{0\nu\beta\beta}$ is a phase-space factor [36] and $\mathcal{M}^{0\nu\beta\beta}$ is the nuclear matrix element. Current sensitivity leads to upper limits on the effective Majorana mass of $\mathcal{O}(100\,\text{meV})$, with recent CUORE results giving upper limits between 75 and 350 meV [37]. This relatively large spread of upper limits is primarily due to the range of results of theoretical nuclear matrix element calculations. Next generation experiments, using a range of isotopes and experimental techniques, aim at sensitivities to $\tau_{1/2} = \mathcal{O}(10^{27}\,\text{years})$ and $< m_{\beta\beta} >= \mathcal{O}(10\,\text{meV})$ [38] [39].

2.2.3 The Strong CP problem

The only experimentally observed origin of CP violation in the SM is in weak interactions in the quark sector and is attributed to the complex phase in the CKM matrix. In principle however CP violation is also possible in the strong interaction, with associated parameter θ_{QCD}. This should generate electric dipole moments (edm) d (conventionally measured in units of electron charge $\times$ cm: $e\,$cm) for the neutron [40] [13]

$$|d_n|_{\theta_{CQD}} \approx (0.9 - 1.2) \times 10^{-16}\theta_{QCD}\, e\,\text{cm}\,. \tag{2.2.10}$$

In principle the CP violating phase from the CKM matrix gives rise to an edm at three and four loop orders for quarks and leptons respectively [41] [42] and contributes an additional negligible amount to the neutron edm [13] [40]

$$|d_n|_{CKM} \approx (1-6) \times 10^{-32}\, e\,\mathrm{cm}\,. \tag{2.2.11}$$

The uncertainty in the theoretical predictions is indicated by the ranges quoted which include the impacts of hadronic, nuclear, and atomic theory uncertainties. It is clear that the CKM contribution may be neglected since it is not observable with current or foreseen technologies, however the component from θ_{QCD} could be observed depending on the size of the parameter. In a model without fine-tuning one would expect $\theta_{QCD} \sim \mathcal{O}(1)$, however the current experimental bounds on the neutron edm are [13] [43]

$$|d_n| < 3.6 \times 10^{-26}\, e\,\mathrm{cm}\,(@95\%\, C.L.) \quad \text{implying} \quad \theta_{QCD} < 3 \times 10^{-10}\,. \tag{2.2.12}$$

This improbably low value of θ_{QCD} suggests fine-tuning and is described as the 'strong CP problem', implying that there should be some deeper and more natural explanation for the size of θ_{QCD}.

A natural solution to the strong CP problem is given by the Peccei-Quinn (PQ) mechanism [44] [45] [46] where a spontaneously broken global (PQ) symmetry, $U(1)_{PQ}$, is introduced. Explicit breaking of this global symmetry at energy scale f_a , through anomalous triangle coupling to gluons, leads to a massive pseudo-Nambu-Goldstone boson, which is called an axion, a [13] [46] [47]. At next-to-next-to leading order in chiral perturbation theory the axion mass, m_a, is related to the PQ scale, f_a, by [48].

$$m_a = (5.691 \pm 0.051) \times \left(\frac{10^9\,\mathrm{GeV}}{f_a}\right)\,\mathrm{meV}/c^2\,. \tag{2.2.13}$$

A generic class of models exists with no tree level coupling to axions for quarks and leptons but invoking new heavy quarks, Q_H, carrying $U(1)_{PQ}$ charges, the archetype being the KSVZ model [49] [50] with Q_H being electrically neutral. An alternative class of models, requiring at least two Higgs doublets and where quarks and leptons carry $U(1)_{PQ}$ charge, has archetypal DFSZ model [51] [52]. In any case, axion models with large f_a (needed to evade experimental bounds) include only very feeble interactions between the axion and SM particles.

A broader class of BSM models predict pseudo-Nambu-Goldstone bosons arising from breaking of $U(1)$ symmetries, these generally have similar properties to axions and are labelled axion-like particles (ALPs). An important difference is that while for axions there is a strict relationship between coupling to photons and axion mass, no such condition is imposed for ALPs [13] [53]. Experimental searches for axions and ALPs cover a large phase-space and include cosmological, astrophysical and laboratory experimental constraints, with reviews given in references [13] [54].

2.2.4 Dark Matter

Astronomical observations demonstrate that more gravitational matter exists than is visible or can be inferred to be constituted of SM particles. This Dark Matter (DM) does not

interact electromagnetically and so is not visible but its existence can be inferred from the gravitational effect it has.

For a disc galaxy composed of only conventional matter, and a gravitational force inversely proportional to the square of the distance between objects, the velocities of stars are expected to decrease towards larger radii. However, observations show much higher velocities than expected at larger radii, indicating the presence of a significant amount more mass, see figure 2.7a.

On larger scales of galaxy clusters the majority of the 'visible mass' is made up of hot X-ray emitting gas, with a small fraction from stars and smaller planetary bodies. Observations of the bullet cluster (1E 0657-56), see figure 2.7b, reveal that two colliding galaxy clusters have passed through each other, with the well-separated stars passing past each other and hot gas colliding and becoming somewhat separated in the wake of the stellar structure. The majority of the mass is then expected to be concentrated towards the centre, between the galaxy clusters and around the hot X-ray emitting gas. However, gravitational lensing proves that the majority of the mass is still clustered around the stars, demonstrating the presence of an additional collisionless 'DM halo' which forms the dominant component of a galaxy cluster's mass.

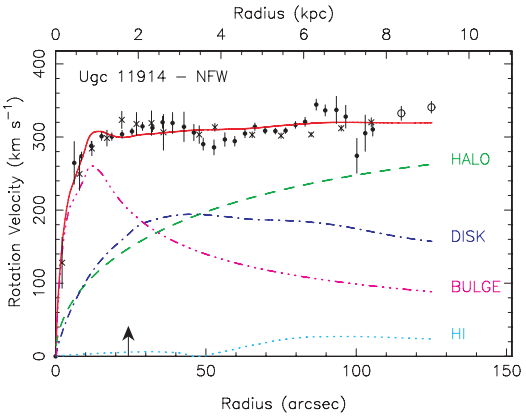
From observations of the anisotropies in the Cosmic Microwave Background (CMB) and the observed universal elemental composition it is established that only $(4.86 \pm 0.10)\%$ of the energy density of the universe is in the form of baryonic matter (composed of any SM particles) with an additional $(25.89 \pm 0.57)\%$ attributed to DM [55]. Additionally, for observed structure formation in the universe, the majority of dark matter particles must have been non-relativistic at the time they de-coupled from the baryonic matter. Such DM candidates are called 'cold dark matter' (CDM). The requirement for majority CDM rules out (light) neutrinos as a viable candidate, and indeed the derived cosmological constraints on the neutrino masses (under the assumption of the standard ΛCDM cosmological model) is $\sum m_\nu < 0.25\,\mathrm{eV}$ @ $90\%\,CL$ [56].

One of the most experimentally interesting DM scenarios is a weakly interacting massive particle (WIMP) produced thermally in the early universe. The DM relic abundance (observed today) depends on the mass and couplings of the WIMP and can be matched to observations with a WIMP with thermally averaged annihilation cross-section of $< \sigma v >= 10^{-9}\,\mathrm{GeV}^{-2} = 10^{-26}\,\mathrm{cm}^{-3}/\mathrm{s}$, of an order expected from electroweak interactions. The 'WIMP miracle' is then that this implies a WIMP mass $\mathcal{O}(100\,\mathrm{GeV}/c^2)$ which corresponds naturally with experimental sensitivity [61]. However, with many stringent experimental bounds being placed on WIMP DM the experimental and theoretical explorations are now broadening, especially towards lower mass candidates [54].

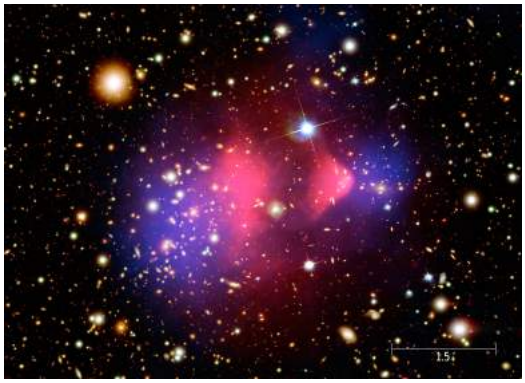
2.2.5 Extra Symmetries and Dimensions

To construct BSM models a common starting point is to consider a new symmetry principle and/or embedding models in higher dimensions.

For example, in Grand Unified Theories (GUTs) all three forces are combined at



(a) Example of Galaxy rotation curve for the UGC 12060 galaxy from [57]. Dots, open circles and crosses correspond to different observational techniques and best-fit models are superimposed. The bulge component represents visible matter with characteristic profile $v \propto \frac{1}{\sqrt{r}}$, with the disk and halo components of the DM halo indicated (using a NFW profile [58], with halo scale radius indicated by arrow), along with neutral (HI) gas.



(b) Composite Image of the Bullet Cluster showing visible light observation (Magellan and HST observatories [59]) with superimposed X-ray (pink, from NASA's Chandra X-ray Observatory [60]) and gravitational lensing map (blue [59]). The separation of the hot X-ray gas (making up the majority of visible matter) and the gravitational barycentres (seen from lensing) provides clear evidence for DM.

Figure 2.7: Observational evidence for Dark Matter.

some high energy GUT scale $\Lambda_{GUT} \sim \mathcal{O}(10^{16}\,\text{GeV})$ where the three interaction couplings of the SM, that run with energy according to renormalisation group equations (RGE), are found to approximately coincide. This scenario may be described with a single symmetry group ($SU(5)$ or $SO(10)$) where the symmetry is spontaneously broken at lower energy scales, including the $SU(3) \times SU(2) \times U(1)$ symmetry of the SM. However, in such GUT scenarios proton decay and magnetic monopoles are generally predicted and the absence of observation of these constrains the possible model construction. Furthermore, such models can leave some problems with the SM unresolved.

Alternatively, a highly developed class of BSM models predict a new space-time symmetry, 'supersymmetry' (SUSY), which maps integer spin particles (bosons) to half integer spin particles (fermions) and vice versa. The particles of the SM therefore become just one component in a supermultiplet with their superpartners which, due to breaking of the supersymmetry, can have much larger masses. The Minimal Supersymmetric standard model (MSSM) realises this new symmetry with a minimal number of new fields and interactions [13]. The 'R-Parity' is often introduced ($R_P = +1$ for SM particles and -1 for their superpartners) and if it is conserved then the lightest SUSY particle and proton become stable. This lightest SUSY particle is then a potential dark matter candidate and the proton stability from the SM is preserved. SUSY models that are covariant under general coordinate transforms encapsulate supergravity (SUGRA), including a spin-3/2 gravitino to mediate gravitational interactions in a quantum theory of gravity. In general, for SUSY theories to maintain their simplicity and theoretical motivations, the scale of SUSY breaking is required to be at relatively low energy scales such that the lightest

SUSY particles exist at scale $\Lambda_{SUSY} \sim \mathcal{O}(\text{TeV})$. Since this is within, or just beyond, current experimental energy reach there has been much experimental effort to search for candidate SUSY particles. However, with no observation yet reported and limits being pushed to higher and higher masses (generally $> 1\,\text{TeV}$) the remaining allowed parameter space of SUSY is becoming limited.

Further theories exist in which the number of space-time dimensions is increased and the phenomenology made possible in this extended space is used to attempt to unify quantum field theories describing particles and quantum systems with general relativity and spacetime curvature. String theory formalisms, where particles are replaced by strings whose vibrations determine their quantum properties and form part of a larger general family of 'branes', have been used to build the phenomenology of the SM plus additional BSM effects.

2.2.6 The Effective Field Theory Approach

New BSM theories must replicate the well-tested results of the SM which may be considered as a low-energy limit of a more general theory, with analogy to the correspondence between classical and relativistic kinematics or between the effective Fermi interaction and the weak interaction. In this scenario the SM Lagrangian is the first term of a more general Lagrangian involving higher order operators (the SM includes operators up to dimension 4). This may be expressed as [62]

$$\mathcal{L}_{EFT} = \mathcal{L}_{SM} + \frac{a_5}{\Lambda^{(5)}} O_5 + \sum_i \frac{a_6^i}{(\Lambda_i^{(6)})^2} O_6^i + \sum_j \frac{a_7^j}{(\Lambda_j^{(7)})^3} O_7^j + \dots . \tag{2.2.14}$$

There is only one possible dimension 5 operator, O_5, which violates lepton flavour numbers and gives masses to neutrinos, however there are a plethora of possible higher dimension operators. Each operator is accompanied by an arbitrary constant a_D^k which is expected to be of order one in units of the energy scale $(\Lambda^{(D)})^{D-4}$. Leading order new physics effects are therefore expected to arise from new dimension 6 operators in this standard model effective field theory (SMEFT) formalism.

2.2.7 Hints of New Physics: Anomalies

2.2.7.1 B Physics Anomalies

In decays of B mesons several measurements have been reported which are in tension with SM predictions regarding lepton flavour universality (LFU), either in τ/μ or τ/e differences in $b \to c\ell\nu$ charged current transitions or in μ/e differences in $b \to s\ell\ell$ neutral current processes [63]. Experimental measurements are performed for ratios

$$R_{K^{(*)}} = \frac{\mathcal{B}(B \to K^{(*)}\mu\bar{\mu})}{\mathcal{B}(B \to K^{(*)}e\bar{e})} \tag{2.2.15}$$

and

$$R_{D^{(*)}} = \frac{\mathcal{B}(B \to D^{(*)}\tau\bar{\nu_\tau})}{\mathcal{B}(B \to D^{(*)}\ell\bar{\nu_\ell})}\,, \quad \text{with}\, \ell = \mu, e \tag{2.2.16}$$

which can be precisely predicted theoretically, with uncertainties cancelling in the ratio both theoretically and experimentally [64]. Theoretically the $R_{K^{(*)}}$ ratios are close to unity, with small corrections arising from lepton masses. Experimentally a lower value than expected for R_K and R_{K^*} is observed [65] [66], suggesting a suppression in $b \to s\ell\ell$ neutral current transitions at one loop [67]. Conversely, larger values are observed of R_D and R_{D^*} than predicted in the SM [65] [66], suggesting an enhancement of (tree-level in SM) $b \to c\ell\nu$ charged current processes. Each individual measurement has a tension of order 2–3σ with respect to the SM prediction, combined the two cases reach a 3–4σ tension. Other measurements of observables which are more challenging to predict theoretically with high precision, such as P_5', also show some tension with the SM and in several BSM scenarios a coherent model can be found which describes all of these tensions. If LFU violation (LFUV) can be demonstrated beyond a reasonable doubt (the standard in particle physics being a significance corresponding to $> 5\sigma$) a clear BSM signature will have been found. Interestingly recent updates [68] [69] have maintained the tension at a similar $\mathcal{O}(3\sigma)$ level and the particle physics community awaits further updates.

2.2.7.2 Muon Anomalous Magnetic Dipole Moment $(g-2)$ Anomaly

Muons are fermion particles, satisfying the Dirac equation, and therefore have intrinsic magnetic dipole moment

$$\mu = g_\mu \frac{e}{2m_\mu}\mathbf{S}\,, \tag{2.2.17}$$

with gyromagnetic ratio g_μ, which is 2 except from a small deviation caused by quantum loop effects parameterised by anomalous magnetic moment [13]

$$a_\mu \equiv \frac{g_\mu - 2}{2}\,. \tag{2.2.18}$$

Experimentally this is measured through the known relationship between the rate of change of angle between the muon spin vector and its momentum, studied by measuring the energy spectrum of $e^\pm$ produced in polarised decays of $\mu^\pm$ (which is modulated by precession of the muon spin) in a storage ring with precisely measured magnetic field [70]. The E821 (often called $g-2$) experiment at BNL recorded results for μ^+ and μ^- with average result [13]

$$a_\mu^{exp} = (11659209.1 \pm 5.4_{stat} \pm 3.3_{syst}) \times 10^{-10}\,. \tag{2.2.19}$$

The theoretical prediction for a_μ is constructed from QED, electroweak, and hadronic components including: loops with leptons; $W^\pm$, Z^0 or Higgs bosons; and quark and gluon contributions respectively. High precision calculations lead to a $\sim 3\sigma$ tension with the experimental result, as shown by figure 2.8. A recent theoretical consensus result was published [71] with

$$a_\mu^{SM} = (11659181.0 \pm 4.3) \times 10^{-10}\,, \tag{2.2.20}$$

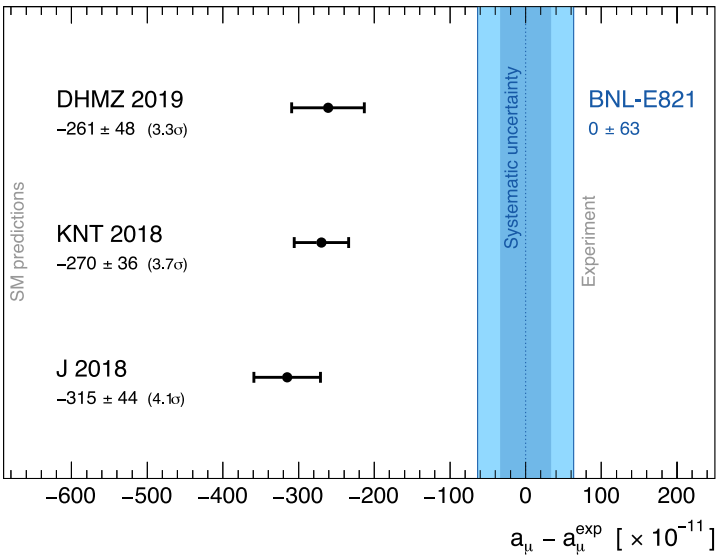


Figure 2.8: [13] Difference between experimental measurement (see equation 2.2.19) and theoretical predictions (DHMZ [74], KNT [75] and J [76]) for the muon anomalous magnetic moment a_μ.

representing a 3.7σ tension with experiment. After the transportation of the E821 storage ring to fermilab the E989 experiment is in progress with goal of further reducing experimental uncertainty by a factor of four [70] [72] [73].

Multiple explanations for this anomaly have been posited, including SUSY or dark photon contributions to loops. However, SUSY particles would require mass $\mathcal{O}(100\,\mathrm{GeV}/c)$ making them directly detectable at the LHC and no evidence of these has yet been found. Similarly, dark photon searches, assuming dominant decay rates to charged leptons [77] disfavour this scenario. Dark photon searches continue with recently announced production search results from NA62 [78] and other proposed searches [54] [79] [80]. Models under study may also concurrently explain the B physics R_{K^*} anomalies [54].

2.3 Theory of $K^+ \to \pi^+\nu\bar{\nu}$

2.3.1 $K^+ \to \pi^+\nu\bar{\nu}$ in the Standard Model

The $K^+ \to \pi^+\nu\bar{\nu}$ decay is a flavour changing neutral current (FCNC) process which is highly suppressed in the standard model through CKM and GIM suppression, with decay width accounted for dominantly by a Z^0 penguin diagram with significant contributions from W box diagrams, shown in figure 2.9.

The branching ratio, after summation over lepton flavours, is given by [81]

$$\mathcal{B} = \kappa_+(1+\Delta_{EM})\left[\left(\frac{\mathrm{Im}[\lambda_t]}{\lambda^5}\right)^2 + \left(\frac{\mathrm{Im}[\lambda_c]}{\lambda}P_c(X) + \frac{\mathrm{Re}[\lambda_t]}{\lambda^5}X(x_t)\right)^2\right], \qquad (2.3.1)$$

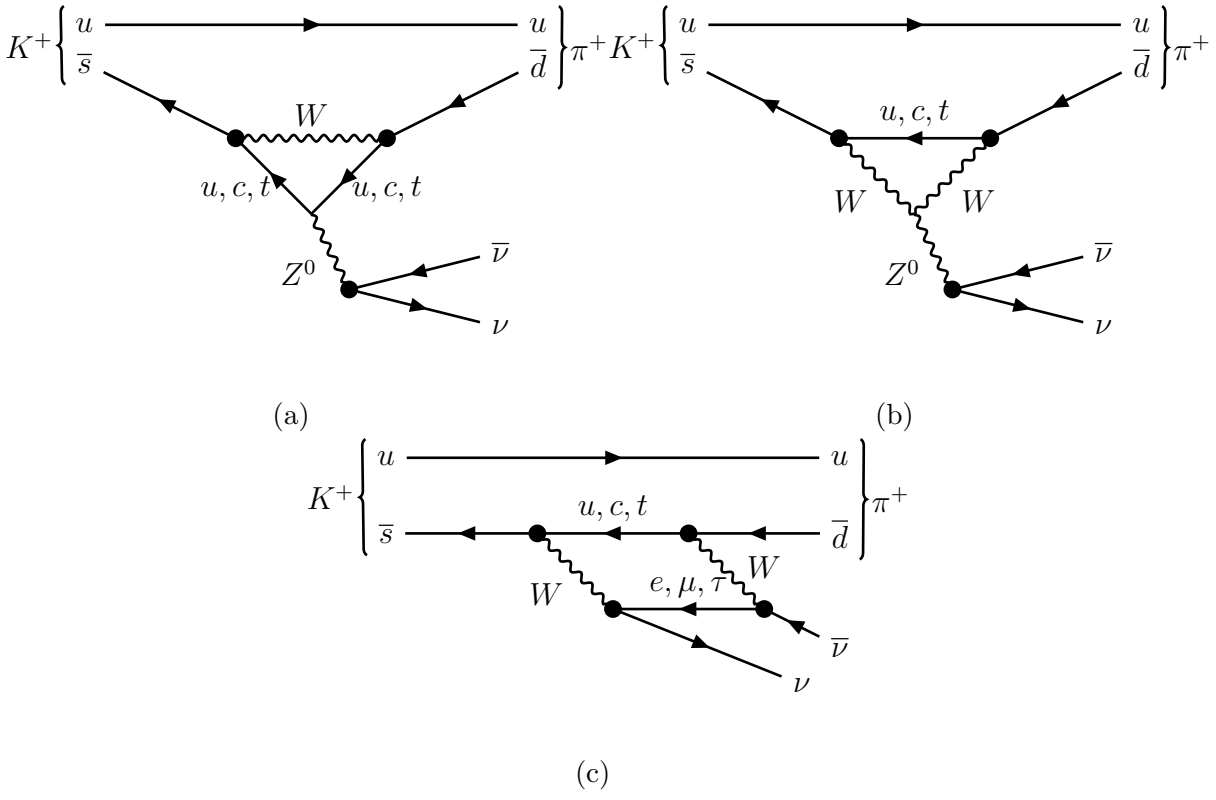


Figure 2.9: Feynman diagrams for the $K^+ \to \pi^+ \nu\bar{\nu}$ decay.

where κ_+ summarises relevant hadronic matrix elements, Δ_{EM} represents the radiative correction from photon exchange,

$$\Delta_{EM} = -0.003\,, \tag{2.3.2}$$

x_t is the ratio of the top quark and W boson masses,

$$x_t = \frac{m_t^2}{m_W^2}\,, \tag{2.3.3}$$

and

$$\lambda = |V_{us}|\,, \tag{2.3.4}$$

$$\lambda_q = V_{qs}^* V_{qd}\,, \tag{2.3.5}$$

where V_{ij} are CKM matrix elements with $X(x_t)$ and $P_c(X)$ being the loop functions for the top and charm contributions respectively.

The κ_+ term may be expressed as [82] [83]

$$\kappa_+ = (5.173 \pm 0.025) \times 10^{-11} \left[\frac{\lambda}{0.225}\right]^8 = \frac{G_F^2 m_K^5 \alpha(m_Z)^2}{256\pi^5 \sin^4\theta_W} |\lambda|^8 \tau_{K^+} (r_K |\lambda f_+^{K^0\pi^+}(0)|_{exp})^2 \mathcal{I}_\nu^+\,, \tag{2.3.6}$$

where G_F is the Fermi constant, m_K and m_Z are the K^+ and Z masses respectively, with $\alpha(m_Z)$ being the value of the running EM coupling constant at m_Z, θ_W is the Weinberg

angle, τ_{K^+} is the K^+ lifetime, the term $(r_K|\lambda f_+^{K^0\pi^+}(0)|_{exp})^2$ includes the relevant K^+ form factor and $\mathcal{I}_\nu^+$ is the phase space integral. The last two terms can be derived to high precision using measurements of $K \to \pi\ell\nu_\ell$ decays (specifically $K_L \to \pi^0 e^+\nu_e$, $K_L \to \pi^0\mu^+\nu_\mu$ and $K_S \to \pi^0 e^+\nu_e$). Indeed, κ_+ may be written explicitly in terms of the branching ratio for $K \to \pi\ell\nu_\ell$ decays as [84]

$$\kappa_+ = r_K \frac{3\alpha^2 \mathcal{B}(K^+ \to \pi^0 e^+\nu_e)}{2\pi^2 \sin^4\theta_W}\lambda^8\,, \tag{2.3.7}$$

where factor $r_K = 0.901$ summarises isospin breaking corrections. The use of these measurements from $K \to \pi\ell\nu_\ell$ decays means uncertainties arising from hadronic matrix elements are negligible.

The t quark loop function is given by

$$X(x_t) = \underbrace{X_0(x_t)}_{\text{leading order}} + \underbrace{\frac{\alpha_s}{4\pi}X_1(x_t)}_{\text{NLO QCD correction}} + \underbrace{\frac{\alpha}{4\pi}X_{EW}(x_t)}_{\text{2-loop EW correction}} = 1.481 \pm 0.005_{th} \pm 0.008_{exp} \tag{2.3.8}$$

where the theory uncertainty quoted arises from scale and scheme uncertainties from renormalisation and associated parameters (evaluated at NNLO precision) and experimental uncertainty arising from experimental inputs to the calculation from the top and W mass entering from x_t.

The corresponding c quark loop function can be written as [81]

$$P_c(X) = P_c^{SD}(X) + \delta P_{c,u} = 0.404 \pm 0.024\,, \tag{2.3.9}$$

with the two terms representing the short and long-distance components respectively with [85] $\delta P_{c,u} = 0.04 \pm 0.02$ and [81]

$$P_c^{SD} = \frac{1}{\lambda^4}\left[\frac{2}{3}X_{NNL} + \frac{1}{3}X_{NNL}^\tau\right] = 0.365 \pm 0.012\,. \tag{2.3.10}$$

Here X_{NNL}^ℓ terms arise from NLO QCD and NNLO calculations including 2-loop EW contributions with distinction between lepton flavours important for the charm case where $m_c < m_\tau$, but not for top quarks with $m_t \gg m_\tau$.

Using the above, the predicted SM branching ratio is [81]

$$\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})_{SM} = (8.4 \pm 1.0) \times 10^{-11}\,. \tag{2.3.11}$$

This precision is possible because the process is dominated by short-distance effects (calculable in perturbation theory) with only small contributions from (non-perturbative) long-distance effects [13] [84] [86]. The components of this uncertainty are shown in figure 2.10a with dominant contributions from the CKM matrix element $|V_{cb}|$ and unitarity triangle angle $\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right)$, in terms of which the branching ratio may be written as

$$\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu}) = (8.39 \pm 0.30) \times 10^{-11}\left[\frac{|V_{cb}|}{40.7 \times 10^{-3}}\right]^{2.8}\left[\frac{\gamma}{73.2^\circ}\right]^{0.74}. \tag{2.3.12}$$

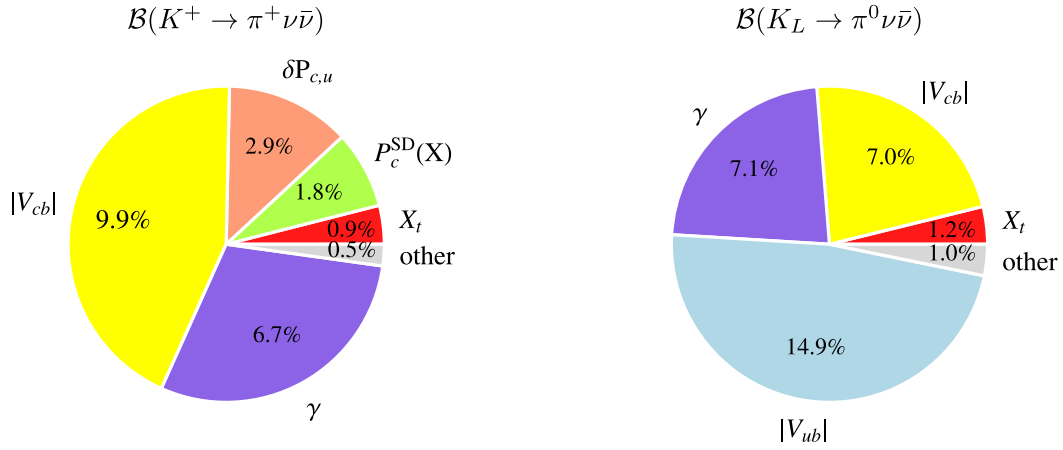


Figure 2.10: [81]Uncertainty budget for $\mathcal{B}(K \to \pi\nu\bar{\nu})$ theoretical predictions. Percentages display individual fractional uncertainties which are added in quadrature to reach total fractional uncertainty of 12% and 18% for the charged and neutral modes respectively.

Further links to the CKM parameters and UT become apparent if the approximate relations [81] [87]

$$\mathrm{Re}[\lambda_t] \simeq |V_{ub}||V_{cb}|\cos\gamma(1-2\lambda^2) + (|V_{ub}|^2 - |V_{cb}|^2)\lambda\left(1-\frac{\lambda^2}{2}\right), \tag{2.3.13}$$

$$\mathrm{Im}[\lambda_t] \simeq |V_{ub}||V_{cb}|\sin\gamma, \tag{2.3.14}$$

$$\mathrm{Re}[\lambda_c] \simeq -\lambda\left(1-\frac{\lambda^2}{2}\right), \tag{2.3.15}$$

are used to re-write the branching ratio in terms of CKM Wolfenstein parameters as [13] [87]

$$\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu}) \approx 1.6\times10^{-5}|V_{cb}|^4(\sigma\bar{\eta} + (\rho_c+\bar{\rho})^2), \tag{2.3.16}$$

where $\sigma = \frac{1}{\left(1-\frac{\lambda^2}{2}\right)^2}$ and with $\rho_c \approx 1.45$. Therefore measurement of $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ describes an ellipse in the UT $(\bar{\rho},\bar{\eta})$ plane with centre $(\rho_c, 0)$ and horizontal semimajor axis of $a \approx \frac{1}{|V_{cb}|^2}\sqrt{\frac{\mathcal{B}}{1.6\times10^{-5}}}$ and vertical semiminor axis of $b \approx \frac{1}{\sigma|V_{cb}|^2}\sqrt{\frac{\mathcal{B}}{1.6\times10^{-5}}}$. The theoretical and experimental ellipses are displayed in figure 2.11.

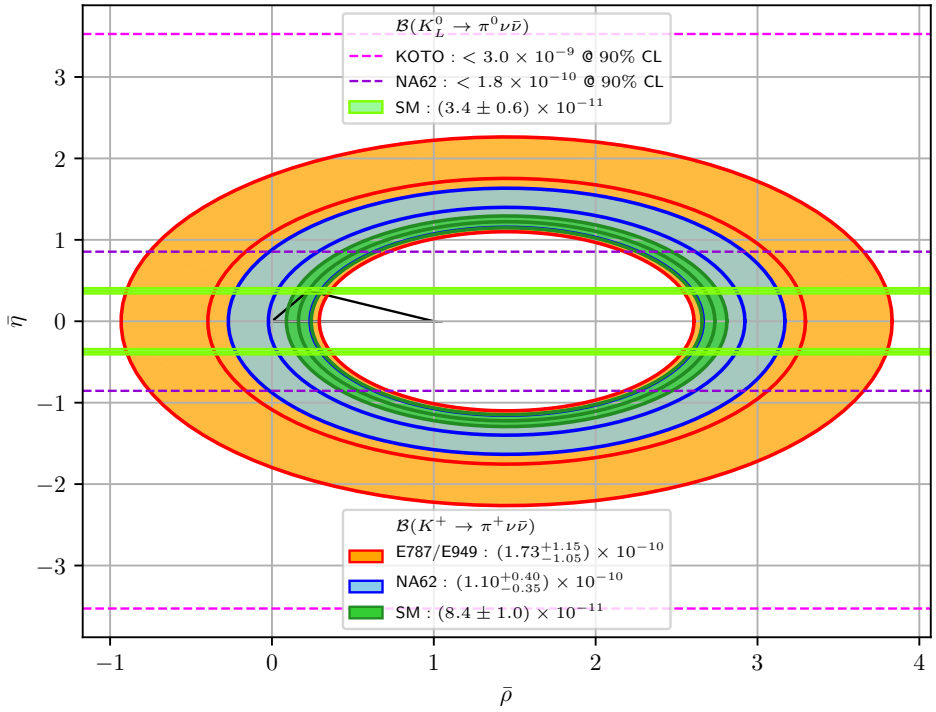


Figure 2.11: UT constraint ellipses and allowed regions derived from SM predictions [81] and experimental results from $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ [88] [89] and $\mathcal{B}(K_L^0 \to \pi^0\nu\bar{\nu})$ [89] [90] . The UT for the SM prediction is shown in black.

2.3.2 Relationship with $K_L^0 \to \pi^0\nu\bar{\nu}$

Similar theoretical analysis has been performed for the equivalent neutral mode $K_L^0 \to \pi^0\nu\bar{\nu}$ with branching ratio given by [81]

$$\mathcal{B}(K_L^0 \to \pi^0\nu\bar{\nu}) = \kappa_L \left(\frac{\mathrm{Im}[\lambda_t]}{\lambda^5} X(x_t) \right) \tag{2.3.17}$$

$$= (3.36 \pm 0.05) \times 10^{-11} \left[\frac{|V_{ub}|}{3.86 \times 10^{-3}} \right]^2 \left[\frac{|V_{cb}|}{40.7 \times 10^{-3}} \right]^2 \left[\frac{\sin\gamma}{\sin(73.2^\circ)} \right]^2 , \tag{2.3.18}$$

$$\simeq 7.6 \times 10^{-5} |V_{cb}|^4 \bar{\eta}^2 , \tag{2.3.19}$$

$$\mathcal{B}(K_L^0 \to \pi^0\nu\bar{\nu})_{SM} = (3.4 \pm 0.6) \times 10^{-11} , \tag{2.3.20}$$

with $\kappa_L = (2.231 \pm 0.013) \times 10^{-10} \left[\frac{\lambda}{0.225} \right]^8$. The top contributions are dominant with charm corrections well below 1% [2] [81] and the main contribution to the uncertainty is from $|V_{ub}|$, see figure 2.10b. Through the second relation a further constraint is added in the UT

[2] The K_L^0 is an admixture of the K^0 and $\bar{K}^0$ flavour eigenstates, $K_L^0 = \frac{1}{\sqrt{2(1+|\varepsilon|^2)}} \left[(1+\varepsilon)K^0 + (1-\varepsilon)\bar{K}^0 \right]$ (where ε is a small complex parameter), and this picks out the imaginary part of the amplitude and hence effectively removes the charm contribution [84]. For the same reason the $K_L^0 \to \pi^0\nu\bar{\nu}$ decay is CP violating, with amplitude proportional to Wolfenstein parameter η (equation 2.3.19).

$(\bar{\rho}, \bar{\eta})$ parameter space and combined with the constraints from the charged mode the UT can be constructed, see figure 2.11.

In the SM, or any BSM model including only isospin changes of $\Delta I = 1/2$, the decay widths of the charged and neutral $K \to \pi\nu\bar{\nu}$ decays are related by [91] [13]

$$\frac{\mathcal{B}(K_L \to \pi^0\nu\bar{\nu})}{\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})} = \frac{\Gamma(K^+)_{tot}}{\Gamma(K_L)_{tot}} \frac{1}{r_{is}} \frac{1}{2} \left(1 + |\lambda^2| - 2\mathrm{Re}[\lambda]\right) , \tag{2.3.21}$$

$$= \frac{\tau_{K_L}}{\tau_{K^+}} \frac{1}{r_{is}} \sin^2\theta \, , \tag{2.3.22}$$

$$\lesssim 4.3 \, , \tag{2.3.23}$$

where $\Gamma(K)_{tot} = \frac{1}{\tau_K}$ is the total decay width of the K^+ or K_L in terms of the lifetime (in units of GeV^{-1}), $\lambda = e^{2i\theta}$ with θ being the CP violating relative phase between the K^0-$\bar{K}^0$ mixing and $s \to d\nu\bar{\nu}$ amplitude and $r_{is} = 0.954$ is the isospin breaking factor [91]. This relationship (equation 2.3.23) defines the Grossman-Nir bound [91] The current theoretical predictions give $\frac{\mathcal{B}(K_L \to \pi^0\nu\bar{\nu})}{\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})} = 0.41 \pm 0.18$, well within this bound. Experimental measurements are currently consistent with the theoretical predictions, however a future observation of $K_L \to \pi^0\nu\bar{\nu}$ with corresponding branching ratio measurement is required to confirm that only $\Delta I = 1/2$ transitions are allowed.

2.3.3 $K^+ \to \pi^+\nu\bar{\nu}$ Beyond the Standard Model

In a variety of BSM scenarios $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ may be modified, either enhanced or suppressed, and is generally a sensitive probe for potential NP. Of particular interest at this time are models where $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ may be enhanced and therefore may be more accessible to experimental measurements which may provide a precision confirmation of a deviation from the SM prediction. If new BSM interactions or particles arise in loop diagrams such enhancements may result.

In general the charged and neutral modes are tied closely together with strong correlation between them, even in BSM scenarios, due to the dominant short-distance dynamics arising from t loops. Dominance of left or right handed currents, or balanced cases with similar magnitude, in the NP will also affect results. Three relatively general classes of BSM scenarios are shown in figure 2.12 [92]. First, in a case where the flavour structure is CKM-like, in minimal flavour violation (MFV) or models with a $U(2)^3$ flavour symmetry (potentially in a SUSY framework) [93], values in the green bands of figure 2.12 are allowed. Second, in models with new CP violating interactions fully dominated by left or right handed interactions the blue regions of figure 2.12 are allowed. The horizontal branch indicates a case where the NP (contribution to λ_t, see equations 2.3.1 and 2.3.17) is entirely real (according to the phase definition of Ref. [92]) and so vanishes for $K_L \to \pi^0\nu\bar{\nu}$, while the vertical branch is for purely imaginary NP contribution and is parallel to the Grossman-Nir bound. This structure is present in Littlest Higgs models with T-parity (LHT) [94] where electroweak symmetry breaking is naturally achieved with a composite Higgs, with the Higgs boson being a pseudo Goldstone boson, such models naturally predict the (relatively) light Higgs boson mass. Alternatively, models with pure

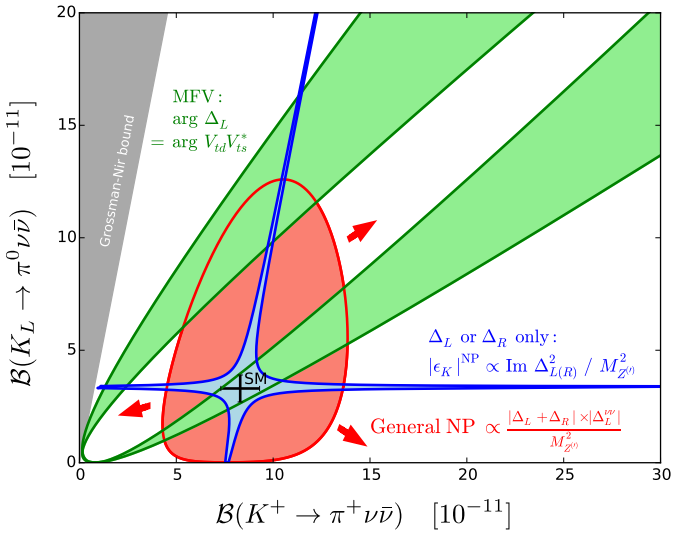
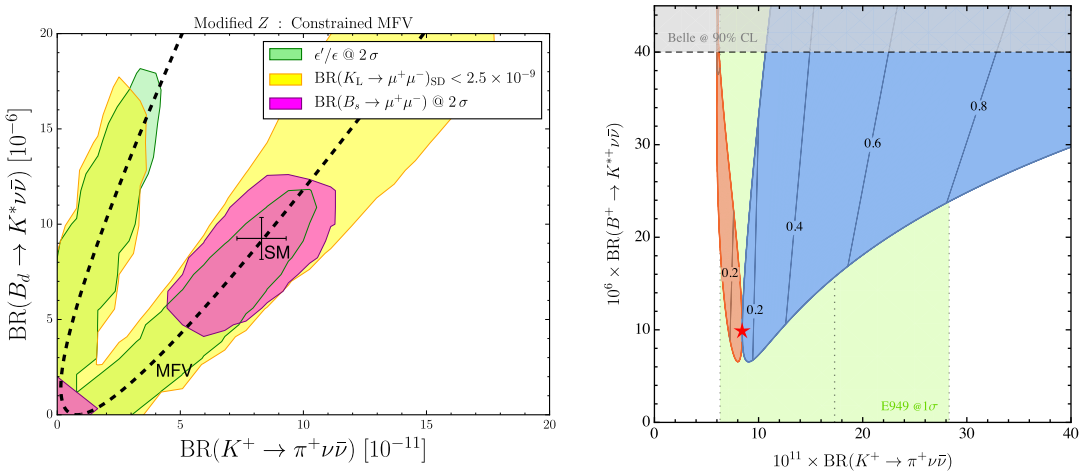


Figure 2.12: [92] Allowed values for the pair of $K \to \pi\nu\bar{\nu}$ branching ratios in three general NP scenarios. Here the green region is allowed in models with a CKM-like flavour structure (MFV scenario), the blue region by models with new flavour violating interactions with fully dominant left- or right-handed couplings (e.g. modified Z/Z' or LHT models) and the red region by NP without flavour constraints (e.g. Randall-Sundrum models).

left or right-handed couplings of modified Z or Z' (a new vector boson with similar or the same couplings as the Z boson but, in general, a significantly higher pole mass) for FCNC [95] [96]. Finally in a third class of models, where flavour constraints are absent or LH-RH operators can contribute, any general region may be allowed (illustrated by the red area in figure 2.12). An example is a Randall-Sundrum model with custodial protection (RSc) [97] with a warped extra dimension with SM fields able to propagate in the bulk except the Higgs field, providing a geometrical explanation of the observed hierarchy of scales and a coherent picture of flavour [98]. Custodial protection refers to a custodial symmetry group, present after electroweak symmetry breaking which maintains a symmetry, in this case of the modified Z/Z' interaction potential, and therefore prevents tree level FCNC processes and ensures FCNC remains a highly suppressed process.

Other correlations exist in NP scenarios between $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ and $\frac{\varepsilon'}{\varepsilon}$ [92] as well as B-physics [99], and models have been constructed which contain explanations for some or all coherently. In figure 2.13 correlations between $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ and $\mathcal{B}(B_d^+ \to K^{*(+)}\nu\bar{\nu})$ are shown, while correlations with the $R_{D^{(*)}}$ parameter(s) linked to the observed B anomalies (see section 2.2.7.1) are shown in figure 2.14.



(a) [92] (b) [99]

Figure 2.13: Correlations between $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ and $\mathcal{B}(B_d^{(+)} \to K^{*(+)}\nu\bar{\nu})$ in NP models.
(a) [92] Allowed ranges for the $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ and $\mathcal{B}(B_d \to K^*\nu\bar{\nu})$ (see also [100]) in a simplified Z model with constraints (coloured 2σ confidence contours) from CP violation parameters ε'/ε and rare $K_L \to \mu^+\mu^-$ and $B_s \to \mu^+\mu^-$ decays.
(b) [99] Allowed regions and correlations between $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ and $\mathcal{B}(B^+ \to K^{*+}\nu\bar{\nu})$ with $R_{D^{(*)}} = 1.25$ and LFU violating NP operators giving rise to $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu}) = \mathcal{B}(K^+ \to \pi^+\nu_e\bar{\nu}_e)_{SM} + \mathcal{B}(K^+ \to \pi^+\nu_\tau\bar{\nu}_\tau)_{SM}\left|1 + \frac{R_0\theta_q^2(1-c_{13}^2)}{\frac{\alpha}{\pi}C}\right|^2$ with $R_0 = \frac{1}{\sqrt{2}G_F\Lambda^2} = \frac{1}{2(1+\theta_q\cos\phi_q)}(R^{\tau/\nu}_{D^{(*)}} - 1)$, C is a complex variable and θ_q, c_{13} and ϕ_q are parameters of the model. The red (blue) regions are where $c_{13} = 0(2)$, lines with associated numbers indicate constant values of θ_q and the star indicates the SM prediction.

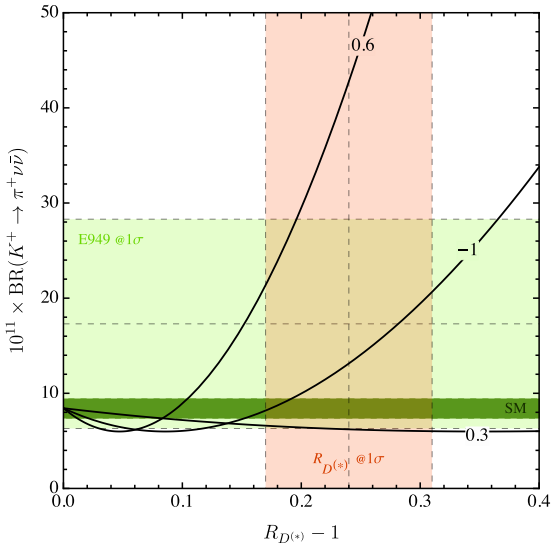


Figure 2.14: [99] Correlation between $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ and $R_{D^{(*)}}$ for different values of parameter θ_q (with $\phi_q = c_{13} = 0$, see caption of figure 2.13b), with experimental results shown in light green and red (indicating 1σ band) and SM prediction of $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ shown in dark green.

2.3.4 Historical and Experimental Context

The historical evolution of the theoretical prediction for $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ and the experimental efforts to measure this quantity are shown in figure 2.15. All experimental measurements preceding NA62 were performed with a stopped-kaon methodology with a K^+ brought effectively to rest at the centre of a detector before decaying, with the π^+ identified through decay chain $\pi^+ \to \mu^+ \to e^+$ and attempting to veto any other activity in hermetic detector systems. Early experiments were carried out at bubble chamber [101] [102] and spark chamber [103] [104] experiments before moving to proton synchrotrons [105] with the E787 [106] [107] [108] [109] [110] [111] and E949 [90] [112] experiments at Brookhaven National Laboratory (BNL) providing a set of measurements from data taken between 1989 and 2002 providing the experimental result before NA62 of [90] [113]

$$\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})_{E787+E949} = (1.73^{+1.15}_{-1.05}) \times 10^{-10}\,, \tag{2.3.24}$$

with observation of seven candidate events with four potentially being due to background (with calculated likelihood of signal/background less than unity). The final E949+E787 results are shown in figure 2.16. The latest experimental progress at the NA62 experiment is reported in detail in chapter 5, with the most recent results from the full Run 1 (2016-18) data, with 20 observed candidate events, with 10.01 ± 0.90 and $6.9^{+1.0}_{-0.8}$ expected signal and background events respectively, giving [89]

$$\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})_{NA62} = (1.10^{+0.40}_{-0.35}) \times 10^{-10}\,. \tag{2.3.25}$$

The ongoing KOTO experiment at J-PARC is designed to study the $K^0_L \to \pi^0\nu\bar{\nu}$ decay. The most recent results reported, from analysis of data taken in 2015, give an upper limit on the branching ratio of [88]

$$\mathcal{B}(K^0_L \to \pi^0\nu\bar{\nu})_{exp,KOTO} < 3.0 \times 10^{-9}\ @90\%\,CL \tag{2.3.26}$$

alongside upper limits on the process $K^0_L \to \pi^0 X^0$, where X^0 is some new BSM particle undetected in the experiment, as a function of its mass. The results are shown in figure 2.17. Preliminary results from analysis of data collected in 2016–18 have been reported [127] [128] with three candidate signal events and a background expectation of 1.05 ± 0.28 [128]. A lower initial background estimate (of 0.05 ± 0.02) [127] led to a flurry of papers seeking a theoretical explanation to a potential 'KOTO anomaly' [129] [130] [131] [130] [132] [133] [134] [135] [136] [137] [138] [139] [140] [141] [142]. Final results have not yet been reported but with current background estimates [128] there is no significant anomaly. In the future combining branching ratio measurements from KOTO and NA62 will allow constraints to be placed on the full UT triangle [143], see figure 2.11.

Using the Grossman-Nir bound [91], see equation 2.3.23, the most recent NA62 results improve upon the KOTO limits obtaining

$$\mathcal{B}(K^0_L \to \pi^0\nu\bar{\nu})_{exp,NA62GNl} < 1.8 \times 10^{-10}\ @90\%\,CL\,, \tag{2.3.27}$$

see also figure 2.18.

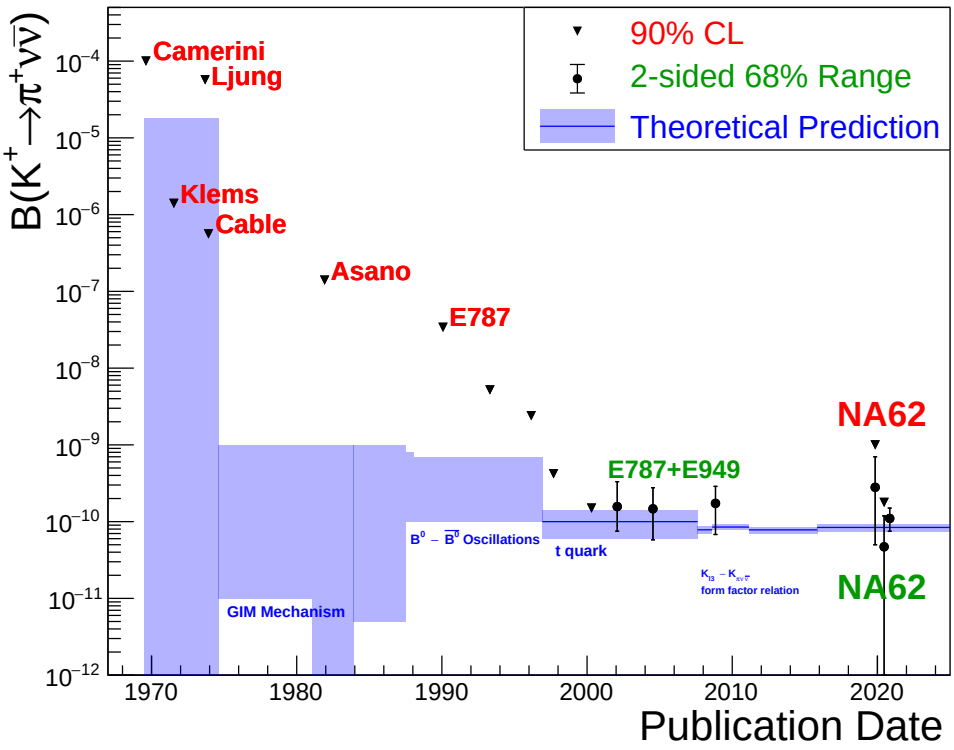


Figure 2.15: Historical evolution of the SM theoretical prediction of $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ from references (from left to right) [114] [115] [116] [117] [118] [119] [120] [83] [121] [122] [81], with milestone theoretical developments highlighted, and experimental measurements from references [101] [103] [102] [104] [123] [106] [107] [108] [109] [110] [111] [111] [112] [90] [124] [125] [89].

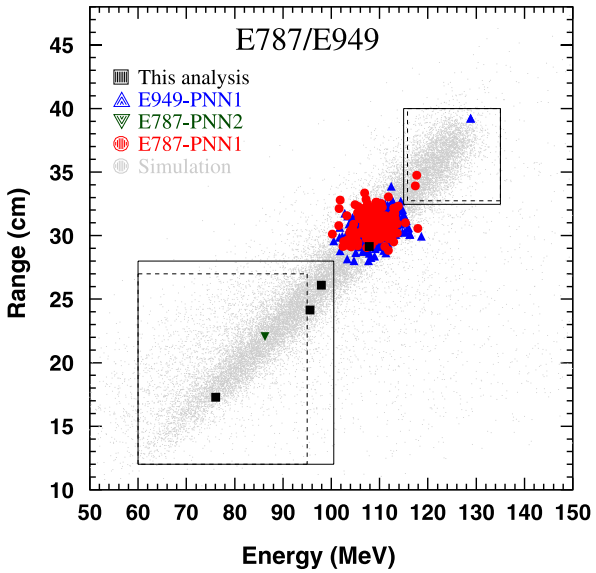
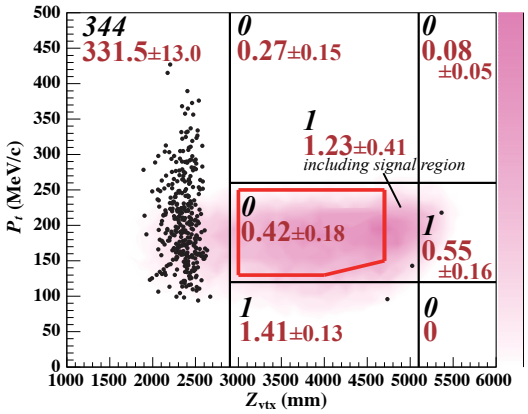
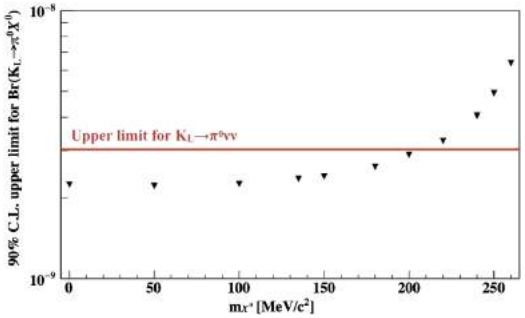


Figure 2.16: [90] [113]Final E949+E787 $K^+ \to \pi^+\nu\bar{\nu}$ results. Points labelled 'this analysis' are from [90], E949-PNN1 [112], E787-PNN2 [126] and E787-PNN1 [111] where PNN1 and PNN2 are the signal regions with p_π below the $K^+ \to \pi^+\pi^0$ peak and above the $K^+ \to \pi^+\pi^0$ peak but below the $K^+ \to \mu^+\nu_\mu$ peak respectively.



(a) Results of search for $K_L^0 \to \pi^0\nu\bar{\nu}$ with selected events (black dots) and expected signal from simulations (pink shaded) displayed in the plane of reconstructed transverse momentum P_t versus K_L^0/π^0 vertex position (Z_{vtx}). In the central signal region 0 events are observed with a background expectation of 0.42 ± 0.18.



(b) Upper limits on the exotic $K_L^0 \to \pi^0 X^0$ decay as a function of m_{X^0}, with upper limit on $K_L^0 \to \pi^0\nu\bar{\nu}$ superimposed for reference.

Figure 2.17: [88]Latest results from the KOTO experiment studying rare and exotic K_L^0 decays.

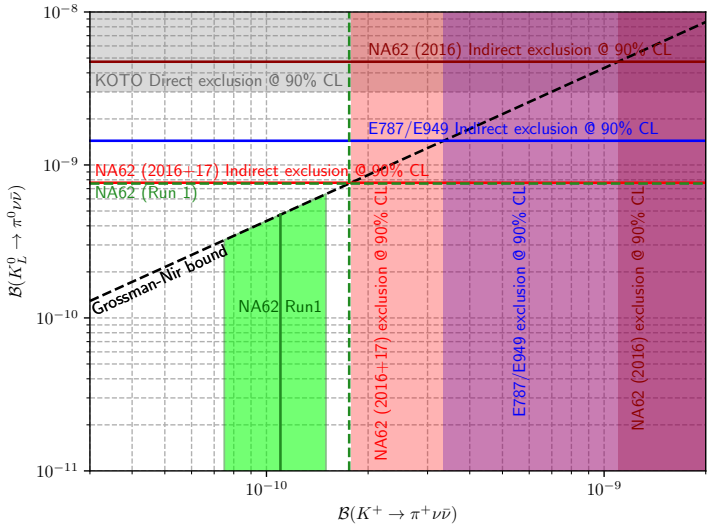


Figure 2.18: Updated constraints on $K \to \pi\nu\bar{\nu}$ [124] [88] [125] [89] with correlations allowing improvement on neutral mode limit from Grossman-Nir bound [91].

2.4 $K^+ \to \pi^+ X$

In BSM scenarios the existence is postulated of X, an electromagnetically neutral long-lived new light scalar (fitting into a scalar portal model [54]) or pseudoscalar particle with only very feeble interactions with SM particles. A massless X particle could naturally be explained as a Nambu-Goldstone boson arising from the spontaneous breaking of a global $U(1)$ symmetry [13] which may then acquire mass through explicit symmetry breaking (in this case being a pseudo-Nambu-Goldstone boson). An important example is that X could be an axion, arising from breaking of a PQ symmetry $U(1)_{PQ}$, which would be a signature of the PQ mechanism which credibly solves the strong CP problem [44] [45] [46], see section 2.2.3. Alternatives, from breaking of the lepton number and flavour/family symmetries respectively, are majorons [144] [145] or familons [146] (more recently, axiflavons [147]). Alternatively X could be an sgoldstino, the superpartner of the longitudinal component of the gravitino arising in models with low energy spontaneously broken supersymmetry [148]. Instead, X may arise as a gauge boson of a new $U(1)'$ symmetry [149] [150]. Moreover, as a feebly interacting long lived particle (FILLP) X can be considered as a light dark matter candidate in a variety of models [151] [152] [153] or an inflaton [154].

The detectable manifestation of a $K^+ \to \pi^+ X$ decay is an incoming K^+, outgoing π^+ and missing energy-momentum, as is the rare $K^+ \to \pi^+ \nu\bar{\nu}$ decay. Alongside results on the $K^+ \to \pi^+ \nu\bar{\nu}$ decay the BNL E787/E949 collaboration reported upper limits on the branching ratio $\mathcal{B}(K^+ \to \pi^+ X)$ [113]. New upper limits were calculated through interpretation of the NA62 analysis of 2017 data as an extension to the $K^+ \to \pi^+ \nu\bar{\nu}$ study. This study and the results are provided in chapter 5. As part of these studies reinterpretation of limits on the branching ratio $\mathcal{B}(K^+ \to \pi^+ X)$ allows constraints on BSM models to be established. In particular three scenarios were investigated with each discussed in turn in the following sections.

2.4.1 X as a Dark Scalar Mixing with the Higgs Boson

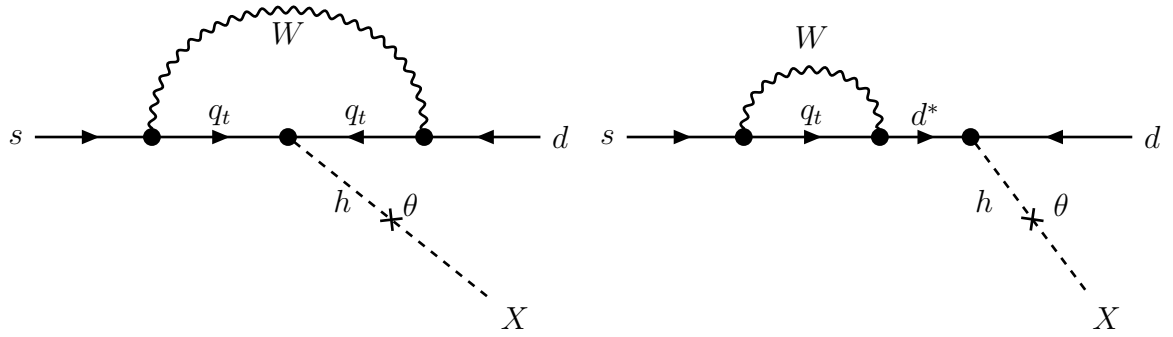
In the general BSM portal model framework operators composed of SM fields, O_{SM}, and new BSM 'dark sector' fields, O_{DS} are combined in new Lagrangian terms

$$\mathcal{L}_{portal} = \sum O_{SM} \times O_{DS} \,, \tag{2.4.1}$$

where the sum is over the possible operators. The most accessible NP contributions will arise in the lowest dimension renormalisable operators. For the minimal scalar portal the dark sector, with one additional singlet field Φ_S, is coupled to the (SM) Higgs boson field H via two types of coupling μ and λ giving overall Lagrangian [54]

$$\mathcal{L} = \mathcal{L}_{SM} + \mathcal{L}_{DS} - (\mu\Phi_S + \lambda\Phi_s^2)H^\dagger H \,, \tag{2.4.2}$$

with terms representing, respectively, the SM and dark sector Lagrangians and the coupling between the visible (SM) and dark sector. After electroweak symmetry breaking and diagonalizing the scalar mass terms two mass eigenstates remain, the (SM) Higgs



(a) (b)

Figure 2.19: Diagram for FCNC processes with production of scalar X in transitions $s \to dX$.

boson h and a new scalar X [155]. At low energies the Higgs field

$$H = \frac{\nu + h}{\sqrt{2}}\,, \tag{2.4.3}$$

where $\nu \approx 246\,\text{GeV}$ is the electroweak vacuum expectation value and h corresponds to the physical Higgs boson which has been observed at the LHC [156] [157] (with $m_h = 125\,\text{GeV}/c^2$ and properties that are so far compatible with the SM Higgs boson [158] [159]). If $\mu \neq 0$ then mixing of h (scalar component of Higgs field) and X states proceeds, and if $\lambda \neq 0$, hXX vertices can be formed allowing pair production (but not decay) of X [54]. A benchmark model (labelled BC4) from [54] assumes that $\lambda = 0$ and therefore all production and decay of X is governed by mixing as expressed through parameter θ where, in the limit of small mixing,

$$\theta = \frac{\mu\nu}{m_h^2 - m_X^2}\,. \tag{2.4.4}$$

Such a framework leads to additional phenomenology, importantly allowing a mechanism for FCNC processes as shown in figure 2.19. The current landscape of constraints on this class of models in the $\sin^2\theta$ versus m_X phase-space is reviewed in [54].

A concrete realisation of this scalar portal scenario is presented in [155] [160] where the new scalar X is a mediator for light thermal dark matter particles (part of the dark/hidden sector) and their interactions with SM particles. The decay width for $K^+ \to \pi^+ X$ is given by [161] [162]

$$\Gamma(K^+ \to \pi^+ X) = \frac{p_X}{8\pi m_{K^+}^2}|\mathcal{M}|^2\,, \tag{2.4.5}$$

where p_X is the X momentum in the K^+ decay rest frame

$$p_X = \frac{\sqrt{(m_X^2 + m_{K^+}^2 - m_{\pi^+}^2)^2 - 4m_{K^+}m_X}}{2m_{K^+}}\,, \tag{2.4.6}$$

and the amplitude squared is given by

$$|\mathcal{M}|^2 = \sin^2\theta \frac{\left(m_{K^+}^2 - m_{\pi^+}^2\right)^2}{v^2} \left(\frac{m_s + m_d}{m_s - m_d}\right)^2 \left(\gamma_1 \frac{7}{18} \frac{m_{K^+}^2 - m_X^2 + m_{\pi^+}^2}{m_{K^2}} - \gamma_2 \frac{7}{9} + \frac{3}{32\pi^2 v^2} \sum_{i=u,c,t} V_{id}^* m_i^2 V_{is}\right)^2 , \tag{2.4.7}$$

where v is the Higgs vacuum expectation value, $\gamma_1 \approx 3.1 \times 10^{-7}$ and $\gamma_2 \approx 0$. The final term dominates, this arises from W loop diagrams, see figure 2.19a, which are dominated by the t quark (according to the GIM mechanism [11]). Additionally, a simplified mass dependence can be used [160]. Therefore the amplitude can be expressed approximately as [163] [160]

$$|\mathcal{M}|^2 = \sin^2\theta \frac{3m_{K^+}^2}{32\pi^2 v^2} \frac{m_s^2}{(m_s - m_d)^2} |V_{td}^* m_t^2 V_{ts}|^2 . \tag{2.4.8}$$

The branching ratio is then given by [162] [164]

$$\mathcal{B}(K^+ \to \pi^+ X) = \frac{\Gamma(K^+ \to \pi^+ X)}{\Gamma(K^+)_{tot}}, \tag{2.4.9}$$

$$\mathcal{B}(K^+ \to \pi^+ X) = \frac{1}{\Gamma(K^+)_{tot}} \frac{p_X}{m_{K^+}} \frac{|\mathcal{M}|^2}{8\pi m_{K^+}} \approx 3 \times 10^{-3} \frac{p_X}{m_{K^+}} \sin^2\theta , \tag{2.4.10}$$

and therefore $\sin^2\theta$ may be expressed in terms of the branching ratio as

$$\sin^2\theta = 330 \frac{m_{K^+}}{p_X} \mathcal{B}(K^+ \to \pi^+ X) \tag{2.4.11}$$

or, for small angles where equation 2.4.4 may be applied, the coupling strength, μ, is given by

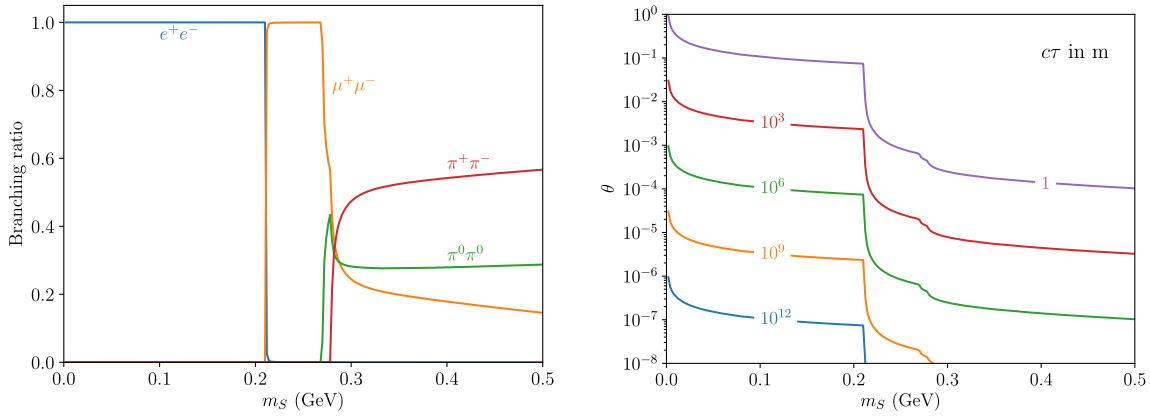
$$\mu = \frac{m_h^2 - m_X^2}{v} \sqrt{330 \frac{m_{K^+}}{p_X} \mathcal{B}(K^+ \to \pi^+ X)} . \tag{2.4.12}$$

Therefore by experimentally setting limits on the branching ratio the mixing parameter and coupling strength may be constrained, see section 6.3.8.

In scenarios where X decays predominantly to invisible particles, for example dark matter [154], the final state is not observed and experimentally this is equivalent to a long lived (or stable) X which escapes the detector apparatus. In this case limits set on the coupling strength and mixing angle through use of equations 2.4.12 and 2.4.11 represent upper limits and all values above are excluded.

If instead X decays predominately to visible (SM) particles, the partial width of X decays is dominated by $X \to e^+e^-$ for $2m_e < m_X < 2m_\mu$ and $X \to \mu^+\mu^-$ for $2m_\mu < m_X \lesssim 2m_\pi$ (the coupling of X to the more massive muons is greater, resulting in a larger decay width) [163] [154], see figure 2.20a. The X lifetime in such scenarios where $X \to \ell^+\ell^-$ ($\ell = \mu, e$) decays dominate is related to the coupling strength via the mixing parameter and the X mass according to [164]

$$\tau_X = \frac{1}{\sin^2\theta} \frac{8\pi v^2 \hbar}{m_\ell^2 m_X} \left(1 - \frac{4m_\ell^2}{m_X^2}\right)^{-3/2} , \tag{2.4.13}$$



(a) (b)

Figure 2.20: [163] For a scenario where $X = S$ is a scalar mixing with the Higgs boson Yukawa couplings arising from this mixing can lead to decays to fermions, dominated below the di-pion threshold by decays to lepton anti-lepton pairs as shown by (a) giving branching ratios for $S \to f\bar{f}$ decays as a function of the scalar mass. The model-dependent lifetime τ (and therefore the mean free path of the scalar in the lab frame, $\beta\gamma c\tau$) is related to the mass and mixing parameter, θ as shown by equation 2.4.13 and illustrated in (b) for several choices of τ ($c = 3.00 \times 10^8\,\mathrm{m/s}$ is the speed of light).

see figure 2.20b. Experimentally the $\ell^+\ell^-$ final state may be observed with sensitivity to $X \to \ell^+\ell^-$ decays depending on the lifetime of X and therefore where it decays in the experimental apparatus. In this scenario the model-dependent lifetime must be considered alongside experimental sensitivity to determine an upper bound for the exclusion in the coupling strength or mixing parameter.

2.4.2 X as a Pseudo-Scalar Mediator for Dark Matter Interactions

An alternative case is considered, following [165], where X is a light (sub-GeV/c^2 mass) pseudoscalar mediator for DM. This is well-motivated theoretically since natural extensions of the Higgs sector give rise to pseudoscalar states, for example two-Higgs doublet models [166] in the context of SUSY. Phenomenologically it is favourable since it suppresses interactions between DM and SM particles and therefore can evade strong bounds from direct detection experiments (as well as collider monojet searches) while potentially explaining some anomalies from astronomical observations [165]. Interactions between the pseudoscalar X and SM fermions are described by Lagrangian term

$$\mathcal{L}_{X,SM} = \sum_{f=q,\ell,\nu} i g_f X \bar{f} \gamma^5 f \,, \tag{2.4.14}$$

where the coupling g_f is real and couplings to neutrinos are considered negligible, $g_\nu \simeq 0$. Three possible coupling structures are then considered [165]:

- **Yukawa-like couplings.**
  Couplings to charged fermions are proportional to the SM Yukawa couplings y_f so

$$g_f = \frac{\sqrt{2} m_f y_f}{\nu} \,. \tag{2.4.15}$$

  This coupling structure is expected for pseudoscalars arising from an extended Higgs sector, with mixing to the SM Higgs naturally introducing a link to SM Yukawa couplings. Using the conversion $g'_f = g_f 2\sqrt{2}$ this model is identical to the BC10 model of [54] which explicitly considers X as an Axion-Like Particle (ALP).

- **Universal coupling to Quarks.**
  If X couples universally to quarks but not to leptons then the sum in equation 2.4.14 runs only over quarks and $g_f = g_q$.

- **Coupling to third generation quarks only.**
  In a more limited case X only couples to t and b quarks (equally) with $g_f = g_Q$.

In each scenario FCNC processes, including $K^+ \to \pi^+ X$, can proceed according to diagrams in figure 2.21. The decay width for $K^+ \to \pi^+ X$ is given by [165]

$$\Gamma(K^+ \to \pi^+ X) = \frac{1}{16\pi m_{K^+}^3} \xi(m_{K^+}, m_{\pi^+}, m_X) \left(\frac{m_{K^+}^2 - m_{\pi^+}^2}{m_s - m_d} \right)^2 |h_{sd}^S|^2 \,, \tag{2.4.16}$$

where $\xi(a, b, c)$ is given by

$$\xi(a, b, c) = \sqrt{(a^2 - b^2 - c^2)^2 - 4b^2 c^2} \,, \tag{2.4.17}$$

and $h_{sd}^S = c g_f$ is a complex number c multiplied by the coupling strength which, in the three cases of interest, is given by [165]

$$\text{Yukawa-like couplings: } h_{ds}^S = (3.5 \times 10^{-9} + 1.5 \times 10^{-9} i) g_Y \,, \tag{2.4.18}$$

$$\text{Quark universal couplings: } h_{ds}^S \approx (4.6 \times 10^{-6} + 2.0 \times 10^{-6} i) g_q \,, \tag{2.4.19}$$

$$\text{3rd-Gen. quark couplings: } h_{ds}^S \approx (1.7 \times 10^{-6} + 7.6 \times 10^{-10} i) g_Q \,. \tag{2.4.20}$$

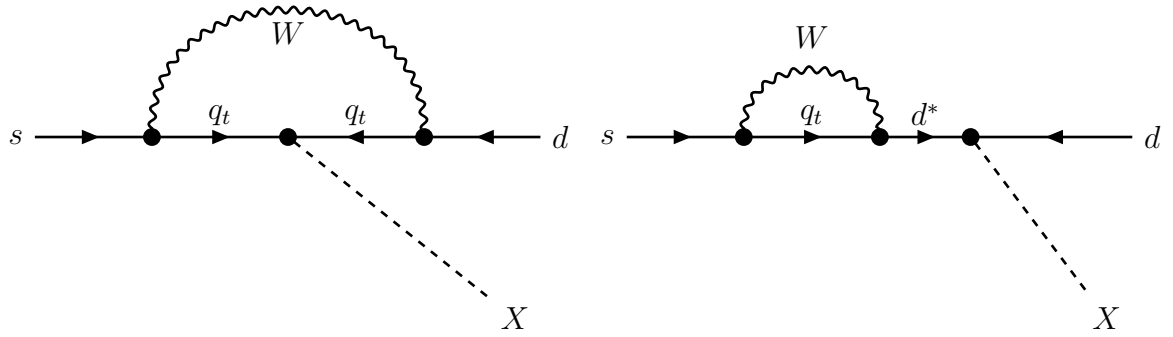
Using equation 2.4.9 the coupling strength is then given by

$$g = \sqrt{\frac{1}{|c|^2} \mathcal{B}(K^+ \to \pi^+ X) \Gamma(K^+)_{tot} \frac{16\pi m_{K^+}^3}{\xi(m_{K^+}, m_{\pi^+}, m_X)} \left(\frac{m_s - m_d}{m_{K^+}^2 - m_{\pi^+}^2} \right)^2} \,, \tag{2.4.21}$$

with $g = g_Y$, g_q or g_Q with corresponding complex coefficients c from equations 2.4.18, 2.4.19 and 2.4.20. In section 6.3.8 constraints are placed on the three cases in this framework.

As for the scenario of scalar X, described in section 2.4.1, there are two possibilities:

1. The pseudosclar X decays predominantly to invisible particles (e.g. DM) or is long-lived relative to the size of a detector. Experimentally no final state is recorded and equation 2.4.21 allows limits to be set on the coupling strength, excluding all larger couplings.



(a) (b)

Figure 2.21: Diagram for FCNC processes with production of pseudoscalar X in transitions $s \to dX$.

2. The decay width of pseudoscalar X is dominated by decays to visible (SM) particles which may be detected. At low X masses the branching ratio is dominated by $X \to \ell^+\ell^-$ decays, with $\ell = e,\ \mu$, as shown by figure 2.22a. For the case of Yukawa-like couplings the model-dependent X lifetime is given by [165]

$$\tau_X = \frac{4\pi\hbar v^2}{m_\ell^2 m_X}\left(1 - \frac{4m_\ell^2}{m_X}\right)^{-\frac{1}{2}} \frac{1}{g_Y^2}\,, \tag{2.4.22}$$

see figure 2.22b. The $\ell^+\ell^-$ final state is detectable experimentally with sensitivity depending on the model-dependent X lifetime. These considerations lead to the derivation of an upper bound for the experimental exclusion in coupling strength.

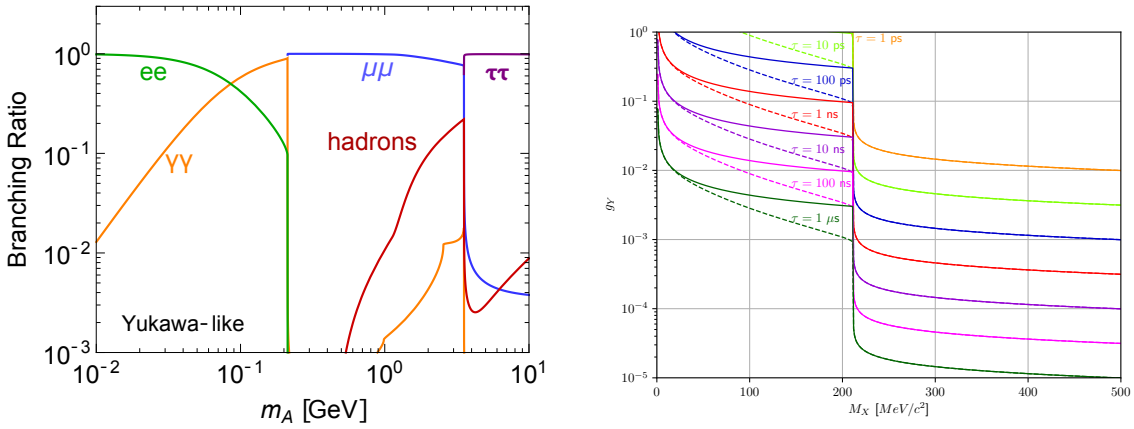
2.4.3 X as an Axiflavon

A QCD axion arising from breaking of a new global $U(1)$ flavour symmetry is predicted by [147] [167] [168]. In [147] X is called an axiflavon and is a Nambu-Goldstone boson with effectively zero mass except that obtained from the QCD anomaly which is given by [48] [13]

$$m_X = (5.691 \pm 0.051) \times 10^{-13}\,\text{eV}\frac{1}{f_a}\,, \tag{2.4.23}$$

(see also equation 2.2.13) where f_a is the axion decay constant, the energy scale that characterises the symmetry breaking, which is usually much greater than the EW symmetry breaking scale $\nu_{EW} \simeq 247\,\text{GeV}$ [13]. In the axifavon model FCNC $K^+ \to \pi^+ X$ processes arise with decay width [147]

$$\Gamma(K^+ \to \pi^+ X) \simeq \frac{1}{f_a^2}\frac{m_{K^+}}{64\pi} m_d m_s \left(1 - \frac{m_{\pi^+}^2}{m_{K^+}^2}\right) B_s^2 \left(\frac{\kappa_{SD}}{N}\right)^2\,, \tag{2.4.24}$$



(a) [165] (b)

Figure 2.22: For a scenario where $X = A$ is a pseudoscalar (ALP) with Yukawa-like couplings the branching ratio for X decays is shown by (a). The relationship between the coupling strength, g_Y, the model-dependent lifetimes and the X mass is described by equation 2.4.22 and illustrated in (b) for several choices of X lifetime, τ, with solid lines accounting for all X decays and dashed lines following the assumption of X decays only to lepton pairs.

where $B_s = 4.6 \pm 0.8$ and $\frac{\kappa_{SD}}{N}$ is a model-dependent $\mathcal{O}(1)$ constant. Using equation 2.4.9 the axion decay constant is then given by

$$f_a = \sqrt{\frac{1}{\mathcal{B}(K^+ \to \pi^+ X)\Gamma(K^+)_{tot}} m_d m_s \frac{m_{K^+}}{64\pi}\left(1 - \frac{m_{\pi^+}^2}{m_{K^+}^2}\right) B_s \frac{\kappa_{SD}}{N}}\,, \tag{2.4.25}$$

and therefore the axiflavon mass by

$$m_X = \frac{5.7 \times 10^{-13}\,\text{eV}}{\left[\sqrt{\frac{1}{\mathcal{B}(K^+ \to \pi^+ X)\Gamma(K^+)_{tot}} m_d m_s \frac{m_{K^+}}{64\pi}\left(1 - \frac{m_{\pi^+}^2}{m_{K^+}^2}\right) B_s \frac{\kappa_{SD}}{N}}\right]} \tag{2.4.26}$$

To set a limit on the axiflavon mass from $K^+ \to \pi^+ X$ decays the limit on the experimental branching ratio input to equation 2.4.26 should correspond to the value measured for a mass hypothesis of $m_X = 0$ since experimentally an axi(flav)on mass will be too small to resolve (and $\frac{N}{\kappa_{SD}} = 1$ is assumed). Results for this scenario are shown in section 6.3.8.

2.5 $K^+ \to \pi\mu e$: Searches for LNV/LFV Decays

The observation of neutrino oscillations demonstrates that the conservation of lepton flavour numbers is only approximate, however no evidence has yet been demonstrated for CLFV. Observation of either CLFV or LNV is clear evidence of NP. A selection of LNV and (C)LFV decays (of K^+, K_L^0 and π^0), in the light meson sector, are summarised in table 2.1 along with the current experimental upper limits.

Table 2.1: Current experimental branching ratio limits on LNV/LFV light meson (K^+, K_L^0 and π^0) decays.

Decay	$\mathcal{B}$ UL (@ 90% CL)	Experiment	Publication & Year	Notes
$K^+ \to \mu^- \nu_\mu e^+ e^+$	2.1×10^{-8}	CERN Geneva-Saclay	[169]([13]) , 1976	LFV
$K^+ \to e^- \nu_e \mu^+ \mu^+$	no limit	–	–	LFV
$K^+ \to \pi^- \mu^+ \mu^+$	8.6×10^{-11}	CERN NA48/2	[170] , 2017	pre-NA62 limit
	4.2×10^{-11}	CERN NA62	[34] , 2019	LNV, LFV
$K^+ \to \pi^- e^+ e^+$	6.4×10^{-10}	BNL E865	[171] , 2000	pre-NA62 limit
	2.2×10^{-10}	CERN NA62	[34] , 2019	LNV, LFV
$K^+ \to \pi^- \mu^+ e^+$	5.0×10^{-10}	BNL E865	[171] , 2000	LNV, LFV
	4.2×10^{-11}	CERN NA62	[4] [5] , 2020	this work
$K^+ \to \pi^+ \mu^- e^+$	5.2×10^{-10}	BNL E865	[171] , 2000	LFV
	6.6×10^{-11}	CERN NA62	[4] [5] , 2020	this work
$K^+ \to \pi^+ \mu^+ e^-$	1.3×10^{-11}	BNL (E777+)E865	[172] , 2005	LFV
$\pi^0 \to \mu^\pm e^\mp$	3.6×10^{-10}	FNAL KTeV	[173] , 2008	LFV
$\pi^0 \to \mu^- e^+$	3.2×10^{-10}	CERN NA62	[5] , 2020	this work
$K_L^0 \to \mu^\pm e^\mp$	4.7×10^{-12}	BNL E871	[174] , 1998	LFV
$K_L^0 \to \pi^0 \mu^\pm e^\mp$	7.6×10^{-11}	FNAL KTeV	[173] , 2008	LFV

The LFV decays $K^+ \to \pi^+ \mu^\mp e^\pm$ may be mediated by exchange of a new neutral heavy 'horizontal' gauge boson X_B which facilitates changes in lepton and quark generation [175] [176]. In such models a new 'Generation' quantum number, G, is established where $G = 2, 1, 0$ for 1st, 2nd and 3rd generation fermions respectively, with inverse signs for antifermions [175] [177]. This quantum number is approximately, but not completely, conserved with transitions involving $\Delta G \neq 0$ being suppressed by powers of an 'inter-generational mixing angle' ϵ. This leads to the interesting prediction that the branching ratios for the two $K^+ \to \pi^+ \mu^\mp e^\pm$ decays could be very different,

$$\underbrace{K^+}_{\substack{\underbrace{\bar{s}u}_{G=-1+2}}} \to \underbrace{\pi^+}_{\substack{\underbrace{\bar{d}u}_{G=2-2}}} \underbrace{\mu^+}_{G=-1} \underbrace{e^-}_{G=+2} , \Delta G = 0 , \tag{2.5.1}$$

$$\underbrace{K^+}_{\substack{\underbrace{\bar{s}u}_{G=-1+2}}} \to \underbrace{\pi^+}_{\substack{\underbrace{\bar{d}u}_{G=2-2}}} \underbrace{\mu^-}_{G=1} \underbrace{e^+}_{G=-2} , \Delta G = -2 . \tag{2.5.2}$$

Therefore the latter $|\Delta G| = 2$ process is expected to be suppressed [175] [176] [177], and the fact that K^0–$\bar{K^0}$ oscillations are a $|\Delta G| = 2$ process means the precision measurements that exist of, for example the mass difference $m_{K_L} - m_{K_S}$, do not exclude the possibility of such a model [177].

The $K^+ \to \pi^+ \mu^\mp e^\pm$ decays may be facilitated by vector or scalar interactions, the corresponding operators to be added to the Lagrangian (see section 2.2.6) can be written

generally as [178]

$$\mathcal{O}_{V,(A)} = \frac{g_{X_B}^2}{2M_{X_B}^2}\bar{d}\gamma_\alpha \left(C_{Lq}P_L + C_{Rq}P_R\right) s \cdot \bar{\mu}\gamma^\alpha \left(C_{L\ell}P_L + C_{R\ell}P_R\right) e + h.c.\,, \tag{2.5.3}$$

$$\mathcal{O}_{S,(P)} = \frac{g_{X_B}^2}{2M_{X_B}^2}\bar{d}\left(C'_{Lq}P_L + C'_{Rq}P_R\right) s \cdot \bar{\mu}\left(C'_{L\ell}P_L + C'_{R\ell}P_R\right) e + h.c.\,, \tag{2.5.4}$$

where the mass and coupling strength of the new X_B boson is given by M_{X_B} and g_{X_B}, $P_L = \frac{1+\gamma_5}{2}$ and $P_R = \frac{1-\gamma_5}{2}$ are the left and right-handed chiral operators and C and C' are constants. These are generation number conserving operators but exchanging e and μ allows the suppressed generation changing operators to be defined [178]. The vector and scalar operators relevant for $K^+ \to \pi^+\mu^\mp e^\pm$ decays are also of the correct structure for axial vector or pseudoscalar operators required for the $K_L \to \mu^\pm e^\mp$ decays (as indicated by subscripts).

Assuming instead that the transition is of a vector − axial vector $(V-A)$ structure analogous to the weak interaction and comparing with similar diagram for the common $K_L^0 \to \pi^+\mu^-\nu_\mu$ decay a branching ratio can be calculated [179]. The Feynman diagrams for the processes with coupling strengths associated with vertices for $K^+ \to \pi^+\mu^\mp e^\pm$ and $K_L^0 \to \pi^+\mu^-\nu_\mu$ decays are shown in figure 2.23. The ratio of decay widths is then given by [179]

$$\frac{\Gamma(K^+ \to \pi^+\mu^\mp e^\pm)}{\Gamma(K_L^0 \to \pi^+\mu^-\nu_\mu)} \simeq \left(\frac{\frac{g_{X_B}g'_{X_B}}{M_{X_B}^2}}{\frac{g^2\sin\theta_C}{m_{W^2}^2}}\right)^2, \tag{2.5.5}$$

where $|V_{us}| \simeq \sin\theta_C$, g is the weak coupling constant and g_{X_B} and g'_{X_B} are coupling constants of the new X_B boson with quarks and leptons respectively, with $g'_{X_B} \propto |\epsilon^{|\Delta G|}|^2$, suppressing $|\Delta G| \neq 0$ processes. One can then express the branching ratio as

$$\mathcal{B}(K^+ \to \pi^+\mu^\mp e^\pm) = \frac{1}{M_{X_B}^4}\left(\frac{g_{X_B}g'_{X_B}}{g^2 sin\theta_C}\right)^2 m_W^4 \frac{\tau_{K^+}}{\tau_{K_L^0}}\mathcal{B}(K_L^0 \to \pi^+\mu^+\nu_\mu)\,, \tag{2.5.6}$$

and setting $g_{X_B} = g'_{X_B} = g$. Figure 2.24 shows the branching ratio as a function of M_{X_B} calculated for this $V-A$ case and the cases of vector (V) and scalar (S) from [178]. With the difference in predicted branching ratios expected with approximate G quantum number conserving models, the strong limit on $\mathcal{B}(K^+ \to \pi^+\mu^+e^-)$ (see table 2.1) is able to exclude these scenarios for a $V-A$ interaction for particular reasonable choices of coupling strengths.

The possibility of a model with extra dimensions is explored by [180], where LFV processes are mediated by Kaluza-Klein modes of new gauge bosons in a 6 dimensional space in which all three SM generations are unified into one. Here there is a 4D spacetime plus a manifold compactified within a sphere of radius R and the three SM generations are localised in different regions of multidimensional space [177]. The generation quantum number is then associated with some angular momentum in compactified space. The structure can be tuned to match SM predictions and the generation mixing parameters (ϵ_q for quarks and ϵ_L for leptons) are free parameters. In this model $\mathcal{B}(K^+ \to \pi^+\mu^\mp e^\pm)$

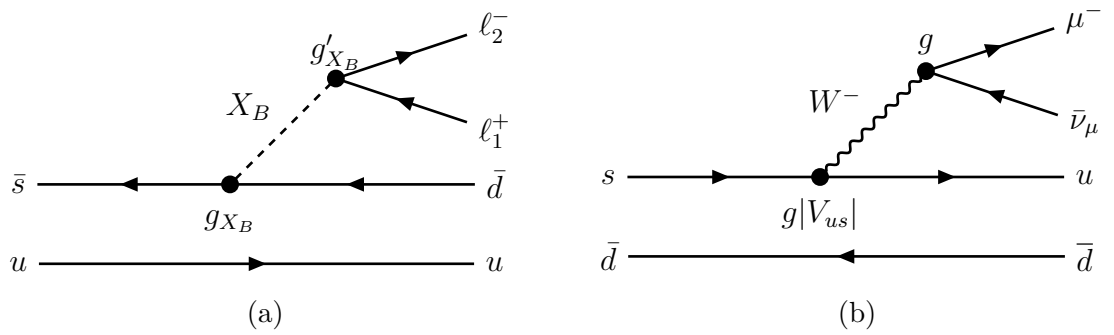


Figure 2.23: Feynman diagrams for the LFV $K^+ \to \pi^+\mu^\pm e^\mp$ decays via horizontal boson X compared to the SM $K_L^0 \to \pi^+\mu^-\bar{\nu}_\mu$ decay via the weak interaction.

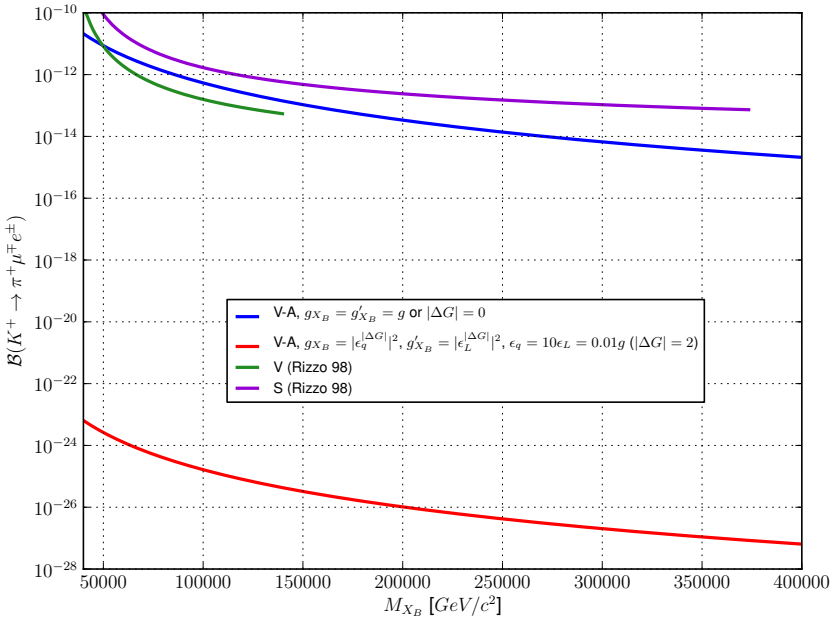


Figure 2.24: Relationship between predicted $\mathcal{B}(K^+ \to \pi^+\mu^\mp e^\pm)$ and mass of mediator boson X_H for different interaction types, vector-axial vector (V-A), calculated according to equation 2.5.6, and vector (V) and scalar(S) from [178]. Note that the $|\Delta G| = 0$ and $|\Delta G| = 2$ curves are relevant for the modes with e^- and μ^- respectively.

is related to the compactification mass scale $\frac{1}{R}$ which may be chosen based on the largest allowed value given other present limits. The most stringent alternative condition comes from the limit on $\mathcal{B}(K_L^0 \to \mu^\pm e^\mp)$, see table 2.1, which implies $\frac{1}{R} = 64\,\text{TeV}$. The value of $\mathcal{B}(K^+ \to \pi^+\mu^\mp e^\pm)$ is given by [180]

$$\mathcal{B}(K^+ \to \pi^+\mu^+e^-) = \left(\frac{1}{1/R}\frac{m_W}{\cos\theta_W}\right)^2 \left(\frac{\zeta}{2\sin\theta_c}\right)^2 \xi_w \mathcal{B}(K^+ \to \pi^0\mu^+\nu_\mu)\,, \tag{2.5.7}$$

where $\xi_w = 1.38$, $\zeta \approx 0.4$ and with $\frac{1}{R} = 64\,\text{TeV}$ the predicted upper limit is

$$\mathcal{B}(K^+ \to \pi^+\mu^+e^-)_{BSM,6D} < 1.75 \times 10^{-13}\,, \tag{2.5.8}$$

with a further suppression for $|\Delta G| = 2$ process $K^+ \to \pi^+\mu^-e^+$. Experimentally these are not ruled out and are currently at least two orders of magnitude beyond current and projected sensitivity.

In certain BSM models new leptoquark, Y_{LQ}, particles arise which can couple to both quarks and leptons. These particles can facilitate both LFV $K^+ \to \pi^+\mu^\mp e^\pm$ and LNV $K^+ \to \pi^-\mu^+e^+$ decays as shown in figure 2.25. Measurements of the branching ratio can be used to constrain parameters of the models, including leptoquark masses and couplings [181]. If one assumes the leptoquark is a scalar then the ratio of the decay width for $K^+ \to \pi^+\mu^\mp e^\pm$ decays width respect to $K^+ \to \pi^0\mu^+\nu_\mu$ is given by [175]

$$\frac{\Gamma(K^+ \to \pi^+\mu^\mp e^\pm)}{\Gamma(K^+ \to \pi^0\mu^+\nu_\mu)} = 16\left(\frac{g_{LQ}g'_{LQ}}{m_{LQ}^2}\right)^2 \left(\frac{m_W^2}{g^2\sin\theta_c}\right)^2\,, \tag{2.5.9}$$

where the mass and coupling constant for leptoquarks are m_{LQ} and $g_{LQ}^{(\prime)}$ respectively and is suppressed by a factor 16 relative to equation 2.5.5 due to being only a vector current as opposed to $V - A$. The branching ratio is then given by

$$\mathcal{B}(K^+ \to \pi^+\mu^\mp e^\pm) = \frac{16m_W^4}{\sin^2\theta_c}\mathcal{B}(K^+ \to \pi^0\mu^+\nu_\mu)\frac{1}{m_{LQ}^4}\left(\frac{g_{LQ}g'_{LQ}}{g^2 sin\theta_C}\right)^2\,, \tag{2.5.10}$$

and re-arranging in terms of the leptoquark mass allows a lower limit on m_{LQ} to be established from an upper limit on the branching ratio according to

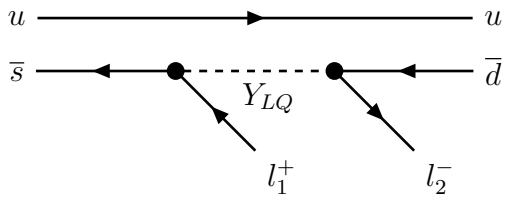
$$m_{LQ} = \left(\frac{\mathcal{B}(K^+ \to \pi^0\mu^+\nu_\mu)}{\mathcal{B}(K^+ \to \pi^+\mu^\mp e^\pm)}\frac{16m_W^4}{\sin^2\theta_c}\right)^{\frac{1}{4}} \frac{\sqrt{g_{LQ}g'_{LQ}}}{g}\,. \tag{2.5.11}$$

Taking present bound, see table 2.1, gives constraint that a vector leptoquark must have mass

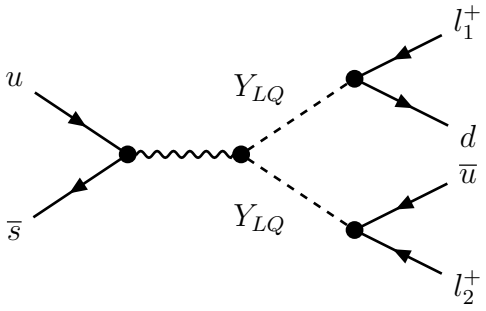
$$m_{LQ} > 76\,\text{TeV}/c^2\,\frac{\sqrt{g_{LQ}g'_{LQ}}}{g}\,. \tag{2.5.12}$$

From this result it is clear that searches for $K^+ \to \pi\mu e$ decays are able to probe indirectly up to very high energy scales.

Right handed neutrinos with majorana mass are able to mediate LNV decay $K^+ \to \pi^-\mu^+e^+$, analogous to $K^+ \to \pi^-\ell^+\ell^+$ with same-flavour $\ell = \mu$ or e shown in section 2.2.2.1



(a) Transition via scalar Leptoquark.



(b) Transition via scalar and vector Leptoquarks.

Figure 2.25: Feynman diagrams for the LFV $K^+ \to \pi^+\mu^\pm e^\mp$ and LNV $K^+ \to \pi^-\mu^+e^+$ decays via Leptoquarks.

with Feynman diagrams shown in figure 2.6. The calculation of the t channel diagram is non-trivial and since each virtual right-handed neutrino is accompanied by a complex mixing matrix parameter which may add constructively or destructively, there is a degeneracy of solutions and a limit on the branching ratio cannot directly constrain the mass or mixing parameter [182]. However, since the leptonic contribution to the amplitude is related by crossing symmetry to that of conversion of a μ in the electromagnetic field of a nucleus, $\mu^- + (Z, A) \to e^+ + (Z-2, A)$, an indirect limit can be established as [182]

$$\mathcal{B}(K^+ \to \pi^-\mu^+e^+) \lesssim 10^{-11}\,. \tag{2.5.13}$$

This indirect limit requires a theoretical estimate of the hadronic matrix element and therefore a direct experimental limit, now able to reach a similar sensitivity, as shown in section 7.7, is of interest [183].

If instead, following [184] [185] [186], the effective Majorana masses $< m_{\ell\ell} >$ are defined by

$$< m^{-1}_{\ell\ell'} > = \sum_N U_{\ell N} U_{\ell' N} \eta_N m_N^{-1}\ \text{, for heavy sterile Majorana neutrinos}\, m_N \gg m_{K^+}\,, \tag{2.5.14}$$

$$< m_{\ell\ell'} > = \sum_N U_{\ell N} U_{\ell' N} \eta_N m_N\ \text{, for light sterile Majorana neutrinos}\, m_N \ll m_{K^+}\,, \tag{2.5.15}$$

with sum over Majorana neutrino species and with mixing parameters $U_{\ell N}$ between standard leptons and Majorana neutrinos with true mass m_N. The predicted values of $\mathcal{B}(K^+ \to \pi^-\mu^+e^+)$ in the two scenarios is then [185]

$$\mathcal{B}(K^+ \to \pi^-\mu^+e^+) = \tau_{K^+} \frac{G_F^2 m_{K^+}^3}{128\pi^3} f_K^2 f_\pi^2 |V_{us}V_{ud}|^2 \Phi^H_{\ell_1\ell_2} | < m^{-1}_{\ell\ell'} > |^2\ ,\ m_N \gg m_{K^+}\,, \tag{2.5.16}$$

$$\sim 2.3 \times 10^{-18}\,\text{MeV}^{-2} \Phi^H_{\ell_1\ell_2} | < m^{-1}_{\ell\ell'} > |^2 \tag{2.5.17}$$

$$\mathcal{B}(K^+ \to \pi^-\mu^+e^+) = \tau_{K^+} \frac{G_F^2 m_{K^+}^3}{16\pi^3} f_K^2 f_\pi^2 |V_{us}V_{ud}|^2 \Phi^L_{\ell_1\ell_2} | < m_{\ell\ell'} > |^2\ ,\ m_N \ll m_{K^+}\,, \tag{2.5.18}$$

$$\sim 1.7 \times 10^{-18}\,\text{MeV}^{-2} \Phi^L_{\ell_1\ell_2} | < m_{\ell\ell'} > |^2 \tag{2.5.19}$$

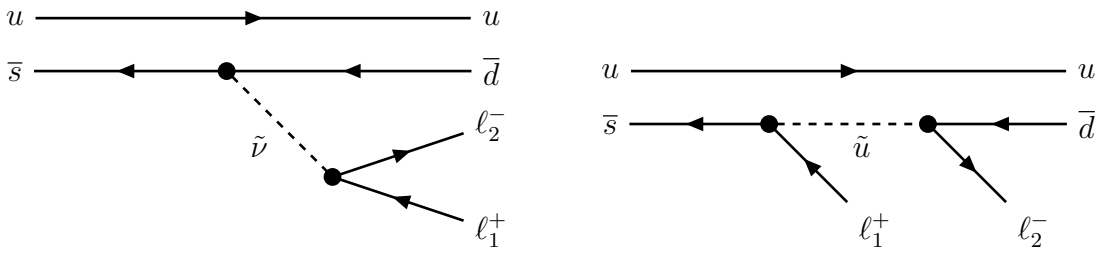


Figure 2.26: Feynman diagrams for the LFV $K^+ \to \pi^+\mu^\pm e^\mp$ decays via SUSY particles, neutralinos $\tilde{\nu}$ and up-type squarks $\tilde{u}$.

where $\Phi^{H,L}_{\ell_1\ell_2}$ are reduced phase space integrals. Through indirect limits on the neutrino masses from precision electroweak tests, neutrino oscillation results, cosmological parameters, and limits from neutrinoless double beta decay searches [184] predictions can be made for the branching ratio in the two scenarios of heavy and light majorana neutrinos as [184]

$$\mathcal{B}(K^+ \to \pi^-\mu^+e^+)_{N_H,indir.} < 5.1 \times 10^{-24}\ , m_N \gg m_{K^+}\ , \tag{2.5.20}$$
$$\mathcal{B}(K^+ \to \pi^-\mu^+e^+)_{N_L,indir.} < 2.0 \times 10^{-33}\ , m_N \ll m_{K^+}\ . \tag{2.5.21}$$

These far exceed sensitivity for direct experimental measurement for the foreseeable future.

In minimal supersymmetry models with R-parity violation ($\not{R}$MSSM), the R-parity violation leads to LNV and LFV mediated by SUSY particles [182]. For the LFV decays the relevant Feynman diagrams are shown in figure 2.26 , mediated by up-type squarks and neutralinos. For LNV decay $K^+ \to \pi^-\mu^+e^+$ the relevant Feynman diagrams are shown in [187] and are mediated by bottom squarks giving predicted branching ratio [187]([182])

$$\mathcal{B}(K^+ \to \pi^-\mu^+e^+) = 4.70 \times 10^{-12} V^2_{LR}(\lambda'_{2K2}\lambda'_{11K}) \left(\frac{100\,\mathrm{GeV}/c^2}{m_{\tilde{b}_K}}\right)^2 , \tag{2.5.22}$$

with V_{LR} being the left-right mixing matrix element, λ terms being coupling strengths and $m_{\tilde{b}_K}$ is the mass of the bottom squark. In the limit of $\lambda = V_{LR} = 1$ and $m_{\tilde{b}_K} = 100\,\mathrm{GeV}/c^2$ this corresponds to a limit just below current experimental sensitivity, however in general even more suppression is expected.

Models with axion-like particles (see sections 2.2.3 and 2.4.2) may also lead to LFV decays [188] [189] [190]. In such models $K^+ \to \pi^+\mu^\pm e^\mp$ decays can occur via an ALP, a, through the process $K^+ \to \pi^+ a$ followed by $a \to \mu^\pm e^\mp$. The branching ratio is then given by [190]

$$\begin{aligned}\mathcal{B}(K^+ \to \pi^+\mu^\pm e^\mp) = &\left[\frac{\tau_{K^+} m_{K^+} f_0(m_a^2)^2}{16\pi}\left(1 - \frac{m^2_{\pi^+}}{m^2_{K^+}}\right)^2 \frac{|v^d_{21}|^2}{\Lambda^2}\xi(m_{K^+}, m_{\pi^+}, m_a)\right] \\ &\times \left[\frac{\tau_a m_a m_\mu^2}{2\pi}\frac{|a^d_{21}|^2}{\Lambda^2}\sqrt{1 - \frac{4m_\mu^2}{m_a^2}}\right] ,\end{aligned} \tag{2.5.23}$$

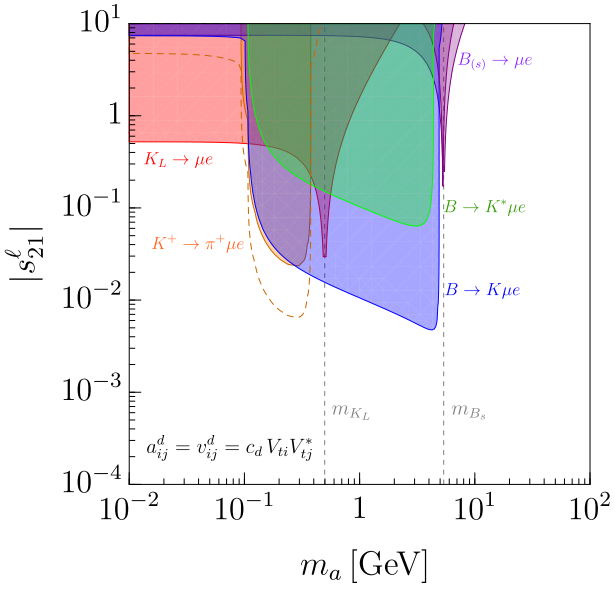


Figure 2.27: [190] Constraints on the effective coupling strengths, $|s^\ell_{ij}|^2$, as a function of ALP mass, m_a, from experimental searches for LFV processes in a scenario where $|v^d_{21}|^2 = |a^d_{21}|^2$ and decay width $\Gamma_a = \frac{\hbar}{\tau_a} = 10^{-6}$. The constraint labelled $K^+ \to \pi^+ \mu e$ is derived from the experimental limit on $\mathcal{B}(K^+ \to \pi^+ \mu^+ e^-)$ [172] and the corresponding dashed line corresponds to $\mathcal{B}(K^+ \to \pi^+ \mu^+ e^-) < 10^{-12}$.

where m_{K^+}, m_{π^+}, m_μ, m_a are the masses of the K^+, π^+, $\mu^\pm$ and a respectively, τ_{K^+} and τ_a are the lifetimes of the K^+ and a respectively, $f_0(m_a^2)$ is the relevant form factor, $\xi(a, b, c)$ is defined by equation 2.4.17 and effective coupling strengths $|s^\ell_{21}|^2 = |v^d_{21}|^2 = |a^d_{21}|^2 = \frac{8\pi}{3}\frac{\Lambda^2}{m_\mu^2}$, when the ALP is produced on-shell ($m_\mu + m_e < m_a < m_{K^+} - m_{\pi^+}$) and using approximation $m_\mu \gg m_e$. Constraints on $|s^\ell_{21}|^2$ as a function of m_a are given for an example scenario in figure 2.27.

Finally, an effective field theory constructed to contain an explanation for the B physics anomalies, potentially showing LFUV, is found to also make predictions pertinent to $K^+ \to \pi^+ \mu^\pm e^\mp$ LFV decays [191]. Considering a CKM-like quark flavour structure with $\lambda^q_{ij} = b_q V^*_{ti} V_{tj}$ where b_q is a flavour-blind coupling and with a lepton-sector coupling λ^ℓ_{ij} the branching ratio is predicted to be

$$\mathcal{B}(K^+ \to \pi^+ \mu^\pm e^\mp) \simeq 0.136 \frac{G_F}{|V_{us}|^2} G_F^2 \left|\frac{C}{\Lambda}\right|^2 |\lambda^\ell_{12}|^2 \left(\mathrm{Re}[\lambda^q_{21}]^2 + \mathrm{Im}[\lambda^q_{21}]^2\right) , \tag{2.5.24}$$

where C is a constant proportional to b_q and Λ is the energy scale of the EFT. For different possible values of λ^ℓ_{12} figure 2.28 shows the expected $\mathcal{B}(K^+ \to \pi^+ \mu^\pm e^\mp)$ for a given Λ. Solid lines in the figure correspond to parameter space compatible with explanations of the B anomalies (specifically R_K) in this EFT, strong limits on the e^- mode already reject some parameter choices.

To summarise, even in BSM models where LFV and LNV are predicted enhancements to reach potentially experimentally detectable levels are rare. However, this does not mean these decay modes should be overlooked since their observation would not only go beyond the SM but require some further consideration in BSM models.

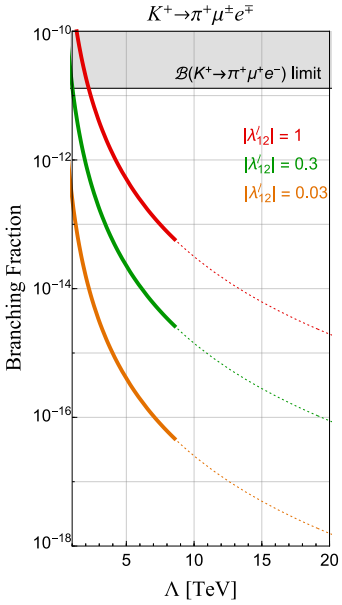


Figure 2.28: [191]Predicted values of $\mathcal{B}(K^+ \to \pi^+ \mu^\pm e^\mp)$ as a function of the NP scale Λ of an effective field theory model able to also explain the current B physics anomalies. The solid lines represent values which are compatible with the explanation of current tension in R_K (to within 2σ) and 3 choices of constants controlling coupling to lepton λ'^{ℓ}_{12} are shown along with the experimental limit [172] (see table 2.1).

2.6 Review of Relevant Decays

A collection of the decays, and their branching ratios, relevant to the work presented in this thesis are given in table 2.2.

Table 2.2: Review of particle decays.

Decay	Abbr.	$\mathcal{B}$	Notes
$K^+ \to \pi^+\pi^+\pi^-$	$K_{3\pi}$	0.05583 ± 0.00024	
$K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$	$K_{2\pi D}$	$(2.427 \pm 0.073) \times 10^{-3}$	with $\mathcal{B}(\pi^0 \to e^+e^-\gamma) = 0.01174 \pm 0.00035$
$K^+ \to \mu^+\nu_\mu[e^+e^-\gamma]_{\pi^0}$	$K_{\mu 3D}$	$(3.94 \pm 0.12) \times 10^{-4}$	with $\mathcal{B}(\pi^0 \to e^+e^-\gamma) = 0.01174 \pm 0.00035$
$K^+ \to e^+\nu_e[e^+e^-\gamma]_{\pi^0}$	K_{e3D}	$(5.95 \pm 0.18) \times 10^{-4}$	with $\mathcal{B}(\pi^0 \to e^+e^-\gamma) = 0.01174 \pm 0.00035$
$K^+ \to \pi^+\pi^-\mu^+\nu_\mu$	$K_{\mu 4}$	$(4.5 \pm 0.2) \times 10^{-6}$ *	
$K^+ \to \pi^+\pi^-e^+\nu_e$	K_{e4}	$(4.247 \pm 0.024) \times 10^{-5}$	
$K^+ \to \pi^+\mu^+\mu^-$	$K_{\pi\mu\mu}$	$(9.4 \pm 0.6) \times 10^{-8}$	
$K^+ \to \pi^+e^+e^-$	$K_{\pi ee}$	$(3.00 \pm 0.09) \times 10^{-7}$	
$K^+ \to \mu^+\nu_\mu\mu^+\mu^-$	$K_{\mu\nu\mu\mu}$	$< 4.1 \times 10^{-7}$	χPT prediction: $(1.350 \pm 0.005) \times 10^{-8}$♣
$K^+ \to e^+\nu_e\mu^+\mu^-$	$K_{e\nu\mu\mu}$	$(1.7 \pm 0.5) \times 10^{-8}$	χPT prediction: $(1.120 \pm 0.005) \times 10^{-8}$♣
$K^+ \to \mu^+\nu_\mu e^+e^-$	$K_{\mu\nu ee}$	$(7.06 \pm 0.31) \times 10^{-8}$	χPT prediction: $(4.200 \pm 0.005) \times 10^{-6}$♣
$K^+ \to e^+\nu_e e^+e^-$	$K_{e\nu ee}$	$(2.48 \pm 0.20) \times 10^{-8}$	χPT prediction: $(1.220 \pm 0.005) \times 10^{-7}$♣
$K^+ \to \pi^+\pi^0\pi^0$	$K_{3\pi^0}$	0.01760 ± 0.00023	
$K^+ \to \pi^+\pi^+\pi^-\gamma$	$K_{3\pi\gamma}$	$(1.04 \pm 0.31) \times 10^{-4}$	
$K^+ \to \pi^+\gamma\gamma$	$K_{\pi\gamma\gamma}$	$(1.01 \pm 0.06) \times 10^{-6}$	
$K^+ \to \pi^+e^+e^-\gamma$	$K_{\pi ee\gamma}$	$(1.19 \pm 0.13) \times 10^{-8}$	
$K^+ \to \pi^-\mu^+\mu^+$	$K^{LNV}_{\pi\mu\mu}$	$< 4.2 \times 10^{-11}$	recent NA62 limit [34]
$K^+ \to \pi^-e^+e^+$	$K^{LNV}_{\pi ee}$	$< 2.2 \times 10^{-10}$	recent NA62 limit [34]
$K^+ \to \pi^-\mu^+e^+$	$K_{\pi^-\mu e}$	$< 5.0 \times 10^{-10}$	this work: $< 4.2 \times 10^{-11}$
$K^+ \to \pi^+\mu^-e^+$	$K_{\pi\mu^- e}$	$< 5.2 \times 10^{-10}$	this work: $< 6.6 \times 10^{-11}$
$K^+ \to \pi^+\mu^+e^-$	$K_{\pi\mu e^-}$	$< 1.3 \times 10^{-11}$	
$K^+ \to \mu^-\nu_\mu e^+e^+$	$K^{LFV}_{\mu\nu ee}$	$< 2.1 \times 10^{-8}$	
$K^+ \to e^-\nu_e\mu^+\mu^+$	$K^{LFV}_{e\nu\mu\mu}$	no limit	
$K^+ \to \pi^+\pi^0$	$K_{2\pi}$	0.2067 ± 0.0008	
$K^+ \to \mu^+\nu_\mu$	$K_{\mu 2}$	0.6356 ± 0.0011	
$K^+ \to e^+\nu_e$	K_{e2}	$(1.582 \pm 0.007) \times 10^{-5}$	
$K^+ \to \pi^0\mu^+\nu_\mu$	$K_{\mu 3}$	0.03352 ± 0.00033	
$K^+ \to \pi^0e^+\nu_e$	K_{e3}	0.0597 ± 0.0004	
$K^+ \to \pi^+\nu\bar{\nu}$	$K_{\pi\nu\nu}$	$(11.0^{+4.0}_{-3.5}) \times 10^{-11}$†	SM prediction: $(8.4 \pm 1.0) \times 10^{-11}$
$K^+ \to \pi^+\nu\bar{\nu}$	$K^0_{\pi\nu\nu}$	$< 3.0 \times 10^{-9}$	SM prediction: $(3.4 \pm 0.6) \times 10^{-11}$
$\pi^0 \to \gamma\gamma$	$\pi^0_{\gamma\gamma}$	0.98823 ± 0.00034	
$\pi^0 \to e^+e^-\gamma$	π^0_D	0.01174 ± 0.00035	
$\pi^0 \to e^+e^-$	π^0_{ee}	$(6.46 \pm 0.33) \times 10^{-8}$	
$\pi^0 \to \mu^\pm e^\mp$	$\pi^0_{\mu e}$	$< 3.6 \times 10^{-10}$	this work: $\mathcal{B}(\pi^0 \to \mu^-e^+) < 3.2 \times 10^{-10}$
$\pi^+ \to \mu^+\nu_\mu$	π^+_ν	0.9998770 ± 0.0000004	
$\pi^+ \to e^+\nu_e$	π^+_e	$(1.230 \pm 0.004) \times 10^{-4}$	

All information from PDG review [13] unless otherwise stated.

♣ Estimate from Chiral Perturbation theory [192]. The assigned uncertainty is 0.5 in the next decimal place after the reported value.

∗ Using the predicted branching ratio $\mathcal{B}(K^+ \to \pi^+\pi^-\mu^+\nu_\mu) = (4.5 \pm 0.2) \times 10^{-6}$ [193] [194] [195], as in the NA48/2 $K^+ \to \pi\mu\mu$ analysis [196]. This is used instead of the measured result, given in [13], of $(1.4 \pm 0.9) \times 10^{-5}$ based on a 1967 result with claimed observation of only 7 events [197].

†Preliminary Run 1 NA62 result [89]. The PDG reports the previous experimental result from the E787/E949 experiment of $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu}) = (17.3^{+11.5}_{-10.5}) \times 10^{-11}$ [90] [113].

Chapter 3

The NA62 Experiment at CERN

3.1 Overview

The NA62 experiment is located at the CERN north area with a high energy particle beam provided by the Super Proton Synchrotron (SPS), see figure 3.1. The experiment is constructed on a site steeped in the history of kaon physics with the previous users of the experimental hall being NA31 and NA48 [198] [199]. The NA62 experiment's primary objective is the study of the ultra-rare $K^+ \to \pi^+\nu\bar{\nu}$ decay, aiming to first provide unequivocal evidence of its existence and then measure its branching ratio at a precision of order 10%, testing the SM prediction (discussed in section 2.3.1) of

$$\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})_{SM} = (8.4 \pm 1.0) \times 10^{-11}\,. \tag{3.1.1}$$

The NA62 detector layout and strategy were purpose-built to meet this challenging physics goal. The signature for the $K^+ \to \pi^+\nu\bar{\nu}$ decay is a K^+ beam particle upstream, a lone π^+ downstream and no additional activity associated spatially or temporally, with 'missing' energy-momentum (carried away by the undetected neutrino anti-neutrino pair). With a lack of total hermeticity and imperfect particle identification (PID) systems this signature is shared by a number of K^+ decay modes with branching ratios up to 10 orders of magnitude greater than the signal, see figure 3.2.

It is vital to match the parent K^+ and π^+ daughter, requiring precise direction and timing measurements for both particles from tracking detectors (see sections 3.2.3 and 3.2.5). Using the measured 3-momenta of K^+ and π^+ candidates, and using the corresponding mass hypotheses to reconstruct the 4-momenta, P_{K^+} and P_π, the squared missing mass variable

$$m^2_{miss} = (P_{K^+} - P_{\pi^+})^2\,, \tag{3.1.2}$$

is reconstructed. Selecting restricted signal regions in this variable (see section 5.1 and figure 3.2) leads to a kinematic suppression of backgrounds to a level of $\mathcal{O}(10^{-3}$–$10^{-4})$, limited by non-Gaussian misreconstruction tails. The momentum and angular resolution for measurements of the K^+ and π^+ are $\frac{\sigma_p}{p}|_{K^+} = 0.2\%$, $\sigma_{\theta_{K^+}} = 16\,\mu$rad and $\frac{\sigma_p}{p}|_{\pi^+} \lesssim 0.5\%$, $\sigma_{\theta_{\pi^+}} \approx 0.060\,$mrad respectively.

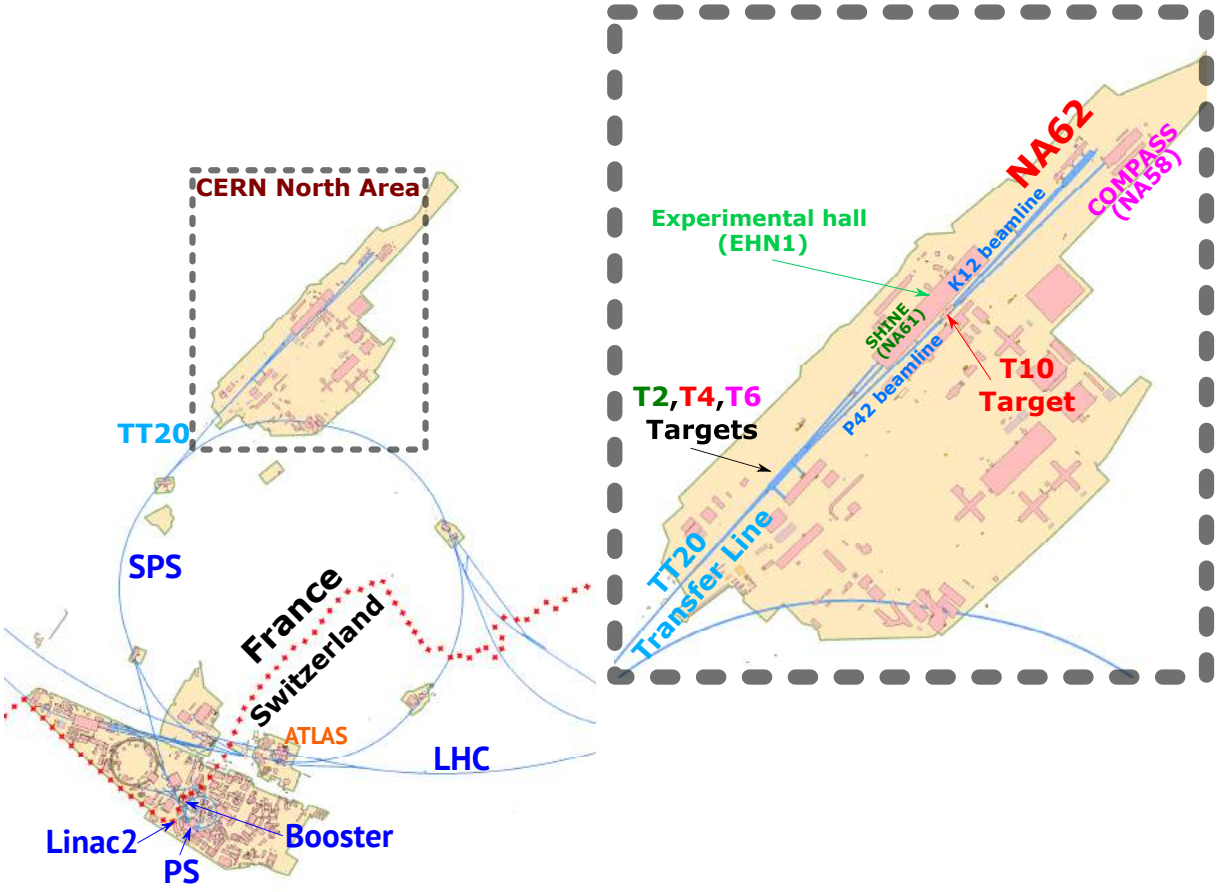


Figure 3.1: The CERN accelerator complex and North Area. Protons are initially accelerated by linear accelerator Linac2 then fed into a series of synchrotron rings, the Booster, the PS (proton synchrotron) then the SPS (super proton synchrotron) where they reach 400 GeV. The TT20 transfer line carries the protons to the CERN north area and to a set of targets. To reach NA62, protons pass through target T4 then travel down the P42 beamline to the T10 target where a secondary hadron beam is created. This secondary beam is carried along the K12 beamline and through the NA62 experimental set-up. (This figure was produced using CERN Geographic Information System plans.)

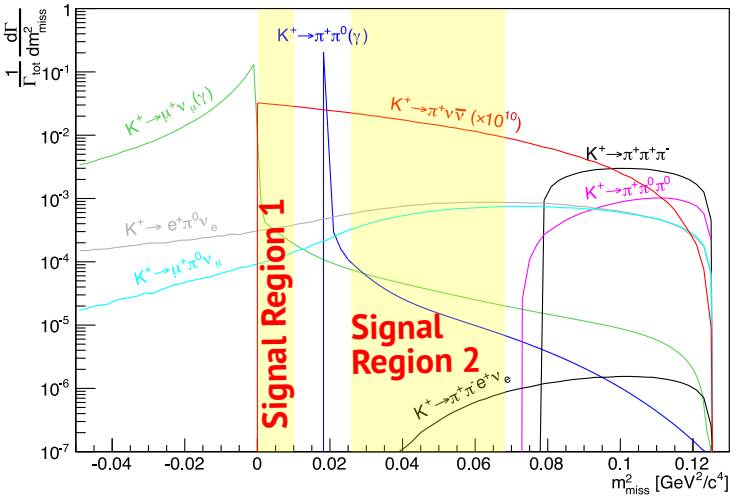


Figure 3.2: (Edited from [200].) Expected m^2_{miss} distributions of K^+ decays. The signal $K^+ \to \pi^+\nu\bar{\nu}$ decay is enhanced by 10 orders of magnitude for visibility above the significant contributions of other background processes. The two signal regions are indicated with borders set to kinematically reject $K^+ \to \mu^+\nu_\mu$, $K^+ \to \pi^+\pi^0$ and $K^+ \to \pi^+\pi^+\pi^-$ decays.

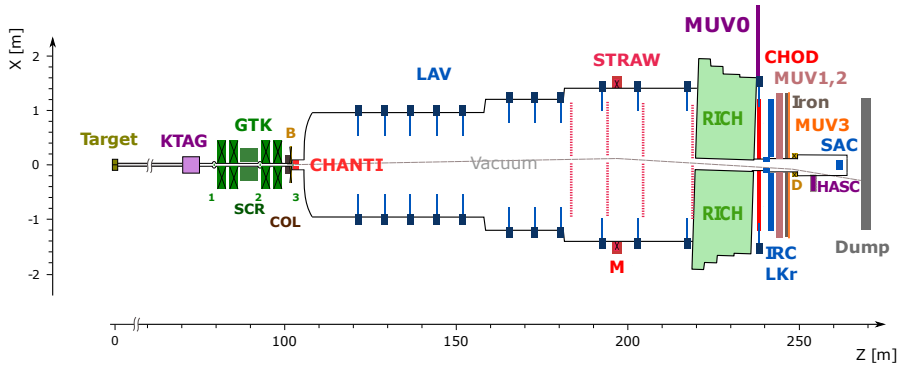
Veto systems (see primarily sections 3.2.4, 3.2.7 and 3.2.10) must reject any activity associated, in time or space, with a candidate K^+–π^+ pair. Especially important is the rejection of photons produced in the common decay chain of $K^+ \to \pi^+\pi^0$ followed by $\pi^0 \to \gamma\gamma$ (which may be written as $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$). A $\pi^0(\to \gamma\gamma)$ detection inefficiency of $\leq \mathcal{O}(10^{-8})$ is required.

Particle identification systems (sections 3.2.6, 3.2.7.1, 3.2.9 and 3.2.10) are needed to ensure powerful separation between π^+ and μ^+ to reject the most common charged kaon decay process $K^+ \to \mu^+\nu_\mu$. The probability of misidentification of a μ^+ as a π^+ must be $\leq \mathcal{O}(10^{-7})$.

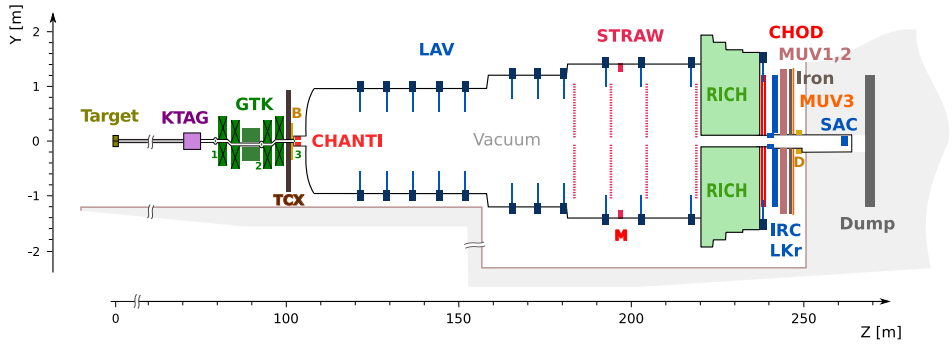
The acute rarity of the $K^+ \to \pi^+\nu\bar{\nu}$ decay means a large number of K^+ decays, $\mathcal{O}(10^{13})$, must be collected; to do this in a reasonable few year time-scale requires a high intensity beam. This further increases the importance of precise timing (with K^+–π^+ time matching with 100 ps precision required) and redundancy in veto and detector systems. Furthermore an efficient triggering system is essential, selecting only a small fraction ($\mathcal{O}(10^{-4})$, see section 3.2.12) of events which are of potential interest to be stored for analysis.

3.2 The NA62 Detector and Data Acquisition Systems

The NA62 experimental apparatus, depicted in figure 3.3, is formed from a number of sub-systems described below. The coordinate system for the experiment is arranged with the origin at the T10 target, with z increasing along the beamline and a right-handed



(a) x vs z view.



(b) y vs z view.

Figure 3.3: [200]Schematic diagram of the NA62 experiment with horizontal (3.3a, 2016 and 2017 configuration) and vertical (3.3b, 2018) perspectives. In June 2018 the final collimator (COL) was replaced by a larger fixed collimator (TCX) [201].

arrangement of x and y axes.

The experiment has a fixed-target design with a 400 GeV beam of protons from the CERN SPS impinging on a cylindrical beryllium target, of diameter 2 mm and 400 mm in length, producing an unseparated secondary hadron beam. This secondary beam is composed of 70% pions, 23% protons and 6% kaons with mean momentum of +75 GeV/c [200].

The experimental interest for NA62 is K^+ decays, and the length of the experiment is fixed by the mean lifetime of the K^+, $\tau = (1.2380 \pm 0.0020) \times 10^{-8}$ s [13], and the special relativistic effects (time dilation/length contraction) from its high momentum. A 75 GeV/c K^+ has Lorentz gamma factor $\gamma = \frac{1}{\sqrt{1-\beta^2}} = 151.9$ (where velocity in the lab frame, in units of the speed of light c, is $\beta = \frac{p}{E} = \frac{p}{\sqrt{p^2+m^2}} = 0.999978$). This results in a mean path length in the lab reference frame of $l = \beta\gamma c\tau = 564$ m (compared to that for a 75 GeV/c π^+ of $l = 4194$ m). With the beam settings and detector strategy designed in parallel a detector system of about 250 m in length is used with primary sensitivity for decays of K^+ in a decay volume of about 75 m in length housed in a vacuum tank starting 105 m downstream of the target.

3.2.1 NA62 Beam-line & Overview

In a slow extraction cycle [202] [203], with effective time 3 s, protons from the SPS are directed along the P42 extraction line to the T10 beryllium target. The secondary hadron beam produced is then directed along the K12 beam line to the NA62 detector systems as displayed diagrammatically in figure 3.4. Following the target a collimator selects the beam particles to be transmitted. These next reach a triplet of quadrupole magnets which are used to focus the beam. Each quadrupole focuses in one plane while defocusing in the perpendicular plane, but by combining a pair (and additionally a triplet as for this beam-line) the overall effect is focusing in both planes. Subsequently an upstream achromat is encountered, this consists of four dipole magnets which each bend (deflect) the beam. The first pair creates an overall parallel displacement of the beam and the second pair returns it to the original trajectory. However, this is only true for 75 GeV/c momentum hadrons while muons are swept away, being directed out of the main particle beam. During the parallel achromat deflection the beam passes through a series of graduated holes in two large movable metal blocks (TAX1 and 2) which may be closed to act as a secondary target in beam dump operation or for safety when work is carried out in the experimental hall. Remnants of the original proton beam and other superfluous particles are absorbed by the TAXes. Between TAX1 and TAX2 a set of tungsten plates with optional thickness up to 5 mm (corresponding to 1.3 X_0) is positioned at a common horizontal and vertical beam focus to create energy loss for positrons from bremsstrahlung (such that they can be subsequently rejected) while minimising losses through scattering of beam hadrons [200].

Subsequently another set of quadrupoles and collimators further remove unwanted particles from the beam and focus the 75 GeV/c hadronic component of interest. This includes a set of 3 dipole magnets specifically designed to sweep out remaining $\mu^{\pm}$.

The first detector system, the KTAG (see section 3.2.2), is encountered about 70 m downstream of the target. Upstream components critically ensure the beam is parallel as it passes through the KTAG, with minimal angular divergence below 100 μrad (see section 4). Two final focusing quadrupoles follow the KTAG (Q9 and Q10) before the upstream spectrometer achromat system is reached. Four identical dipole magnets (BEND4, 4a, 5 and 6, with 6 inverted vertically with respect to the alignment of the other three) displace the 75 GeV/c beam vertically down by 60 mm before returning it to the original path. The three stations of the GTK (see section 3.2.3) are arranged to perform momentum measurements forming a beam (upstream) spectrometer, with stations 1 and 3 just before and after the achromat with station 2 in the parallel vertically displaced section between the second and third dipole magnets. Also present between BEND4a and 5 is a toroidally magnetized iron collimator referred to as a scraper magnet (SCR1) which defocuses muons and removes them from the central beam. A final set of cleaning collimators (C6, 7, labelled COL) preceding the third GTK station intercept particles just outside the standard beam, including for example $\pi^{\pm}$ arising from decays of K^{+} upstream of this point. During 2018 an additional fixed collimator was added to the beamline (TCX) which further assists in blocking particles arising from upstream decays (specifically the 'mambas' mechanism, see section 5.3.1).

The beam is then deflected by a horizontal steering magnet (B) which introduces

a 90 MeV/c kick to 75 GeV/c beam particles deflecting the beam by 1.2 mrad in the positive x direction. The decay volume then begins at $z = 104.2$ m as particles enter a large vacuum tank.

Downstream a dipole magnet (M), positioned between two pairs of tracking detectors (see section 3.2.5) to form a downstream spectrometer, introduces a kick of 270 MeV/c to 75 GeV/c beam particles deflecting the beam towards negative x by -3.6 mrad. The magnitude of the momentum kick is important because it directly correlates with the momentum resolution achievable by the spectrometer. The resolution depends on the displacement of a track at the final spectrometer tracking station relative to its position without the magnetic field, as determined by the magnitude of the momentum kick, and the point resolution for the tracking detector. After the end of the vacuum tank at $z = 219.6$ m remaining 75 GeV/c beam particles are contained in an evacuated beam pipe following the beam trajectory through the centre of other downstream detector systems. These remaining beam particles are not of interest and are transmitted through the experimental apparatus without further interactions with active detector material. Finally a bending magnet (D) deflects the remaining charged beam particles through angle -13.2 mrad into a beam dump.

3.2.2 Cedar/KTAG

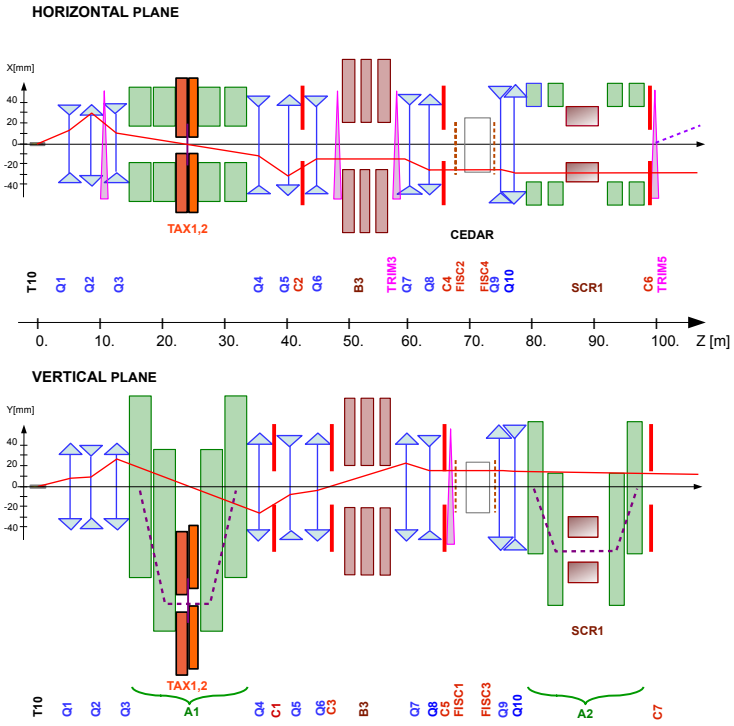
The Cedar/KTAG system provides a fast kaon-tagging system to identify K^+ particles in the unseparated secondary hadron beam entering the NA62 detector. The Cedar/KTAG is a composite system formed from a Cedar module, based on a CERN West Area CEDAR [1] [204], coupled to a specially designed photon detection and readout KTAG detector module. This forms a differential Cherenkov counter with achromatic ring focus with a bespoke photon detection and readout system designed for the high intensity NA62 environment [205]. The optics of the Cedar and KTAG construction are shown in figures 3.5 and 3.6 respectively. Further details of the Cedar/KTAG system, its optics and simulation, are presented in chapter 4.

The Cedar has a cylindrical gas volume of length 7 m and diameter 60 cm operated at room temperature (relatively stable ambient temperature in the experimental cavern). The design allows use of hydrogen (H_2) or nitrogen (N_2) gas and the pressure may be varied (with different working points for the two gases). However, for the 2016–18 data-taking nitrogen gas at 1.73 bar was used. Charged particles passing through the gas at faster than the speed of light in this medium emit Cherenkov photons at an angle θ_C governed by

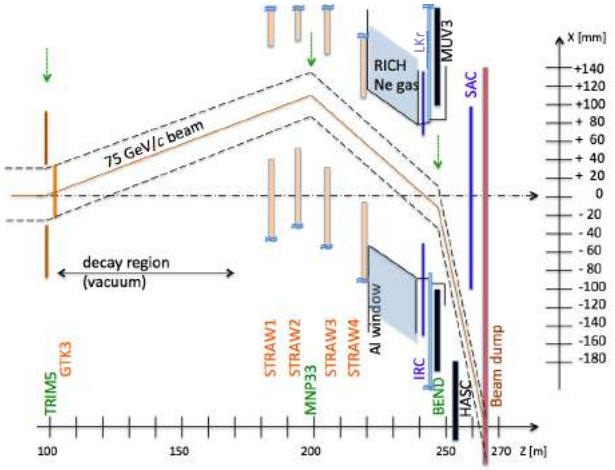
$$\cos(\theta_C) = \frac{1}{\beta n} \ , \tag{3.2.1}$$

where $\beta = \frac{v}{c} = \frac{p}{\gamma mc}$ is the velocity of the particle and n is the refractive index of the gas. The NA62 beam particles have momentum $p = 75$ GeV/c (with only 1% rms spread) so the velocity is different for particle species of different mass, leading to a different emission angle of photons. Cherenkov photons are reflected from the downstream end of

[1] There have been two Cedar designs used at the CERN west and north areas, named Cedar-West and Cedar-North, and optimised for different purposes [204].



(a) Upstream region from T10 target to decay volume with both horizontal and vertical views.



(b) Downstream region (x vs z view) from decay volume entry to beam dump.

Figure 3.4: [200]Schematic diagram of the K12 beam line in the upstream (3.4a), with both horizontal and vertical perspectives displayed and downstream (3.4b) regions.

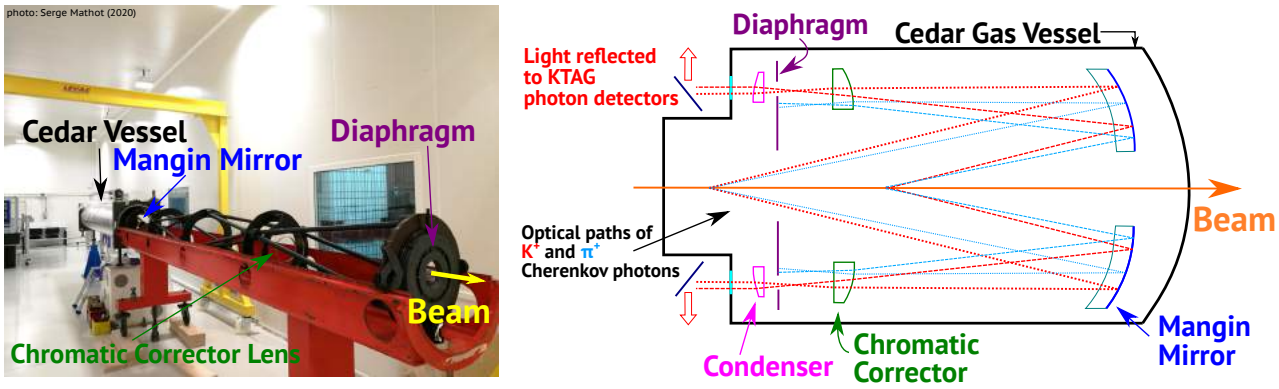


Figure 3.5: Annotated photograph and diagram of Cedar optics. The photograph shows a Cedar-N [204], after it was opened at CERN during 2020 in preparation for design of a new Cedar-H for future NA62 data-taking.

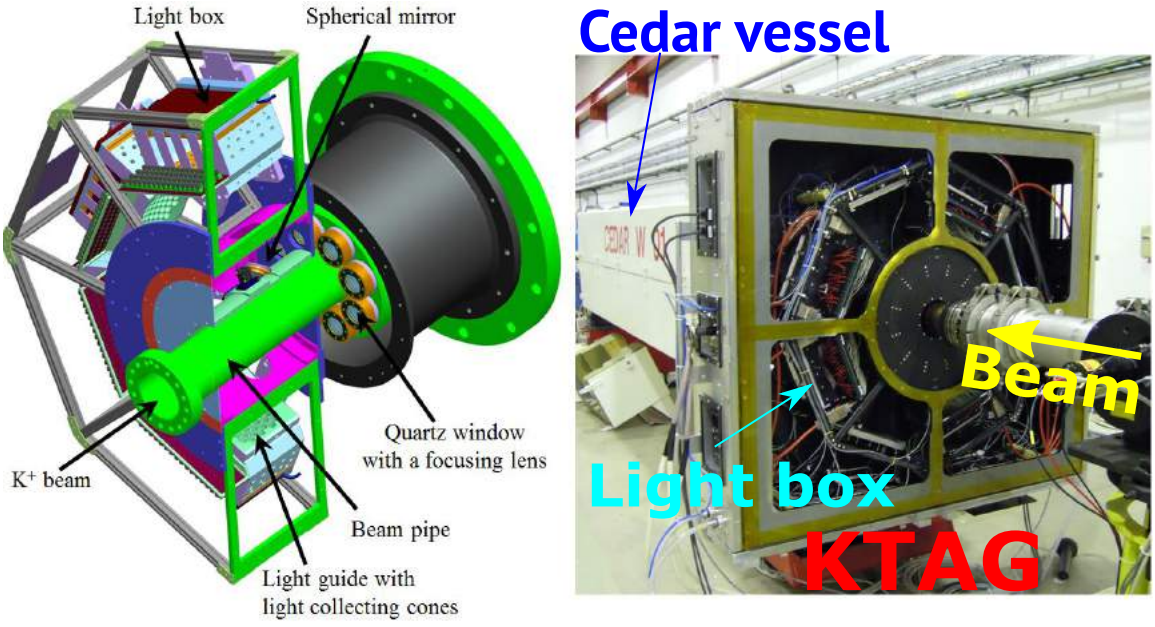


Figure 3.6: (Adapted from [200].) KTAG schematic cut-away diagram (left) and annotated photograph (right) showing installation of the detector with 4 sectors completed.

the Cedar by a Mangin mirror and passed through a chromatic corrector, a system of lenses, and an annular diaphragm. The central radius of the diaphragm is fixed, while the outer edge is movable, meaning the width of the annular opening can be modified. By controlling the refractive index (by choosing the gas type and pressure contained in the Cedar vessel) and the width of the diaphragm opening, light due to a specific particle species can be selected, K^+ for standard NA62 operation. Light passing through the diaphragm is then refracted at the upstream end of the Cedar and steered through eight quartz exit windows into the corresponding eight sectors of the KTAG photon detection system. Light is focused onto a spherical mirror then reflected onto eight arrays of 48 photomultiplier tubes (PMTs) [205]. This distribution of light reduces the single-channel occupancy allowing operation at high intensities while maintaining high precision time measurements for beam particles.

Alignment of the Cedar optical axis and the beam, to 100 μrad precision, is performed, minimising left-right and top-bottom asymmetries in the detected light. The pressure is optimised to maximise light yield (through a pressure scan) and at the 2016–18 data-taking working point on average 20 photon detections were recorded per kaon. The RMS time resolution on a single channel is 300 ps, translating to a K^+ time resolution of 70 ps. The K^+ identification probability is 98% when requiring coincident signals in at least five KTAG sectors with a pion misidentification estimated (at low intensity) to be $\mathcal{O}(10^{-4})$.

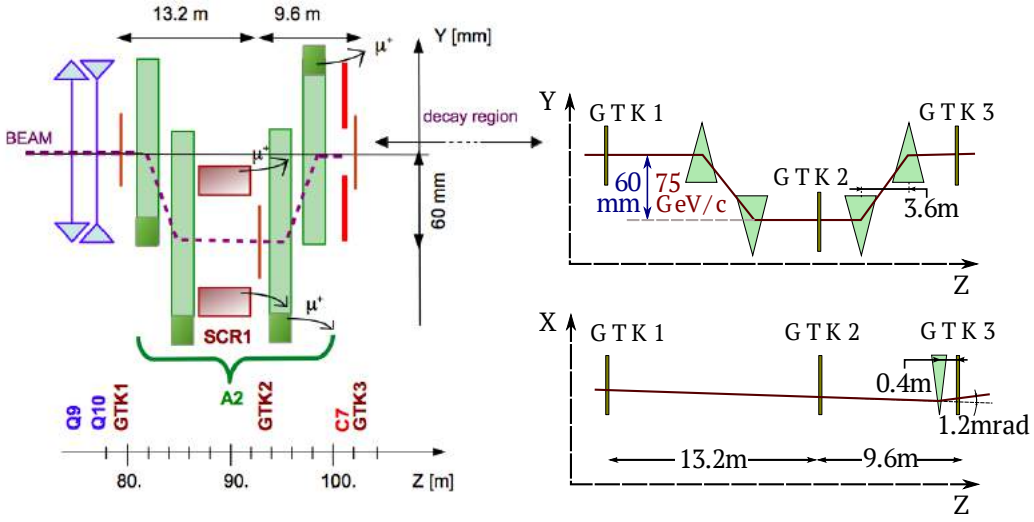
3.2.3 GigaTracker (GTK)

The GigaTracker or GTK is formed of three hybrid silicon pixel detectors positioned around an achromat to form an upstream spectrometer (see figure 3.7), providing precision momentum, direction and timing measurements of beam particles. Each station is constructed from a 200×90 array of 18,000 pixels with dimensions $300\,\mu\text{m} \times 300\,\mu\text{m}$ read out by application-specific integrated circuits (ASICs). Particles of different momentum are deflected by different amounts in a magnetic field. By making precision measurements of the position of the track in GTK2 compared to GTK1 and 3 the deflection due to the magnetic fields is derived and therefore the momentum calculated with hits in three stations matched to reconstruct an upstream beam track. The precision achieved on the momentum measurement is 0.2% with angular measurement to 16 μrad precision, a hit time resolution of 200 ps and therefore a track time resolution of order 130 ps.

The full beam rate of, nominally, 0.75 GHz is experienced by the GTK stations with a corresponding maximum rate in the central region of 1.5 MHz/mm^2. The detector therefore must be able to sustain high levels of radiation, with damage limited by cooling to $-15°C$ with single phase C_6F_{14} [2] pumped through 150 channels (heat exchangers) on the back of a silicon wafer (cooling plate) bonded to the silicon sensor. To minimise multiple Coulomb scattering [3], especially in the final station (GTK3), the radiation length

[2] This coolant was chosen for inertness, non-flammability and radiation resistance properties, however a drawback is its high greenhouse potential.

[3] The deflection of charged particles through multiple electrostatic interactions with atoms of material traversed.



(a) (b)

Figure 3.7: [200]Schematic diagrams showing the positions of the GTK stations in relation to the final upstream beam-line components, importantly the achromat (BEND4,4a,5 & 6) magnets.

(see section 3.2.7.1) of the material in the beam is minimised, with each station being $0.005\,X_0$ ($\sim 500\,\mu$m of silicon). This helps maximise position resolution and minimise potential backgrounds for the $K^+ \to \pi^+\nu\bar{\nu}$ study. Alignment and momentum calibration are performed using fully reconstructed $K^+ \to \pi^+\pi^+\pi^-$ decays.

3.2.4 CHarged ANTI-coincidence detector (CHANTI)

The primary usage of the CHANTI [206] is to veto particles which have undergone inelastic scattering interactions with GTK3 which may otherwise produce a fake $K^+ \to \pi^+\nu\bar{\nu}$ signal.

The CHANTI is constructed from six hodoscope stations (arrays of scintillator bars, each with two perpendicular 'x' and 'y' layers) located in a vacuum vessel shared with GTK3. When a charged particle passes through a scintillator material its interactions excite electrons into higher energy levels from which they later de-excite, emitting scintillation light. Each scintillator bar has a triangular cross section and is read out, via wavelength shifting (WLS) fibres, by silicon photomultipliers (SiPMs). The stations have a square cross section with side length 300 mm with a central hole through which the beam passes. The CHANTI provides hermetic coverage for particles originating from GTK3 from 49 mrad to to 1.34 rad [200]. The detection efficiency is measured to be at least 99% [200] [206] and has time resolution (with respect to the KTAG reference time) of 830 ps and position resolution of 2.5 mm. Good resolution allows veto conditions to be localised in space and time and therefore minimise rejection of events through random activity (random veto).

3.2.5 Straw Spectrometer (STRAW)

The Straw Spectrometer (STRAW) provides the momentum and direction measurements for particles produced in K^+ decays in the decay volume. The STRAW momentum resolution is $\frac{\sigma_p}{p} = 0.3\% \oplus 0.005\%p$, with p measured in units of GeV/c [200].

Four chambers, each constructed from four layers (views) of straw drift tubes , are positioned around the dipole bending magnet (M), two before and two after, to form a spectrometer. For the first time these straw chambers are operated in a vacuum tank, the same one that contains the decay volume. This, along with the light-weight construction, minimises the material traversed by particles (limited to $0.018X_0$ in total) thus minimising multiple scattering and allowing the best precision measurements with lowest probability of misreconstruction. This is important since scattering in STRAW chambers, along with potential pileup of beam tracks upstream, can lead to background mechanisms for the $K^+ \to \pi^+\nu\bar{\nu}$ studies.

Each STRAW chamber has four views, X,Y,V and U aligned at $0, 90°, -45°$ and $+45°$ respectively and collectively covering a region with outer circular profile of diameter 2.1 m. A hole is present in each chamber centred on the mean beam position and in the central region only one or two views may be traversed by a track. The straw geometry, see figure 3.8, ensures at least two straw crossings per view per track. With two or more views (each view providing one coordinate) in chambers before and after the magnet a track can be reconstructed.

The straw drift tubes are 36 μm thick polyethylene terephthalate (PET), with 50 nm of copper and 20nm of gold coated on the inside which form the cathode. The tube is filled with a gas mixture of 70% argon and 30% carbon dioxide (CO_2) with a gold-plated tungsten anode wire, 30μm in diameter, at the centre. Charged particles passing through the gas create electron-ion pairs which drift to opposite electrodes, inducing charge on the central wire. As electrons approach the central anode wire where the electric field strength grows rapidly a multiplication (avalanche) process occurs leading to the creation of a macroscopic charge pulse (current) on the anode wire which is read out. The time taken for the drift to the central anode wire (in the known electric field) allows the radial position from the wire to be derived. Combined with information from other straw tubes this allows the position of a particle to be derived.

3.2.6 RICH

The Ring Imaging CHerenkov (RICH) detector is designed for particle identification (PID), specifically to distinguish between π^+ and μ^+ of 15–35 GeV/c. Efficiencies for π^+ identification and μ^+ misidentification achieved are momentum dependent and of order $\mathcal{O}$(60–80%) and $\mathcal{O}(10^{-4}$–$10^{-3})$ respectively [125]. The physical principles are the same as for the KTAG, with charged particles passing through a gas medium and emitting Cherenkov photons at an angle determined by their mass and momentum (see equation 3.2.1). By measuring this angle and using the momentum measurement provided by the Straw spectrometer the mass of a particle can be derived, which uniquely identifies

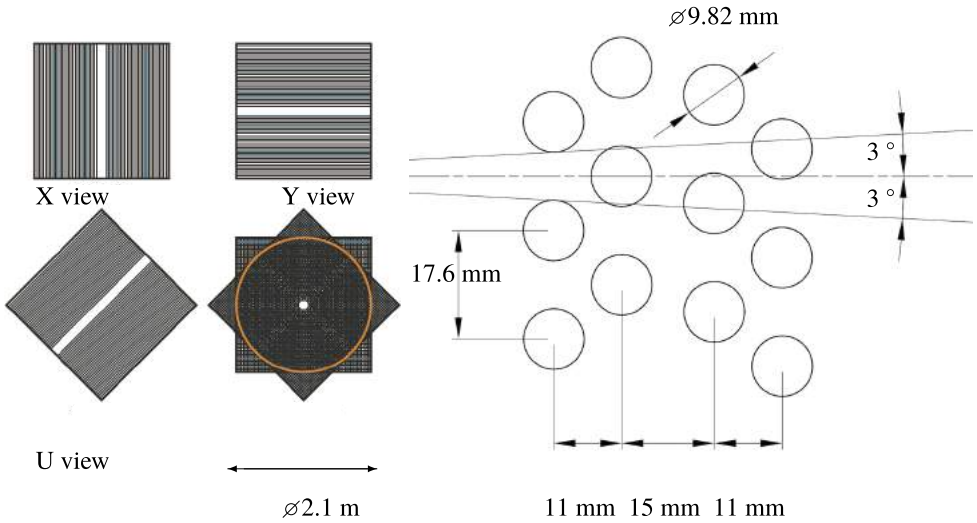


Figure 3.8: [200] Sketches of the STRAW station geometry showing the arrangement of the four views (left) and the arrangement of STRAWs in each view (right) to ensure at least two STRAW tubes are traversed per view.

it. It is also possible to use the RICH as a velocity spectrometer, measuring momentum of a track given an assumed particle identity. The RICH moreover provides a time measurement with 100 ps precision.

A sketch of the RICH is given in figure 3.9, it consists of a 17.5 m long cylindrical vessel, with four sections of decreasing diameter towards larger z and filled with neon gas at atmospheric pressure (990 mbar) and room temperature. This Cherenkov radiator medium has refractive index n such that $(n-1) = 62 \times 10^{-6}$ meaning a Cherenkov threshold momentum of 12.5 GeV/c for π^+. Cherenkov photons produced in a cone as the particle passes through the neon gas are reflected at a mosaic of 20 hexagonal mirrors and focused onto two arrays of 976 PMTs forming rings.

The vacuum tank containing the decay volume ends with the entrance window of the RICH with a conical shape and constructed from 2 mm thick aluminium. A central beam pipe is included inside the RICH meaning beam particles do not create Cherenkov rings at the PMT arrays. The mirror array is split, left-right astride the beam pipe, focusing light on the two PMT arrays and meaning reflected Cherenkov photons avoid absorption by the beam pipe.

3.2.7 Photon Detection and Veto Systems

The decay chain $K^+ \to \pi^+\pi^0$ followed by $\pi^0 \to \gamma\gamma$ produces a signature like $K^+ \to \pi^+\nu\bar{\nu}$ if the photons are not detected. To detect and veto such events a photon veto system is employed with hermetic coverage for these photons from $0-50$ mrad. This angular coverage is shared between the large angle veto (LAV, section 3.2.7.2) $8.5-50$ mrad, liquid krypton calorimeter (section 3.2.7.1) $1-8.5$ mrad and small angle vetos (IRC & SAC, section 3.2.7.3) $0-1$ mrad, see figure 3.10. By requiring that a π^+ must have 15–35 GeV/c the π^0 is constrained to have at least 40 GeV/c meaning high energy photons are produced

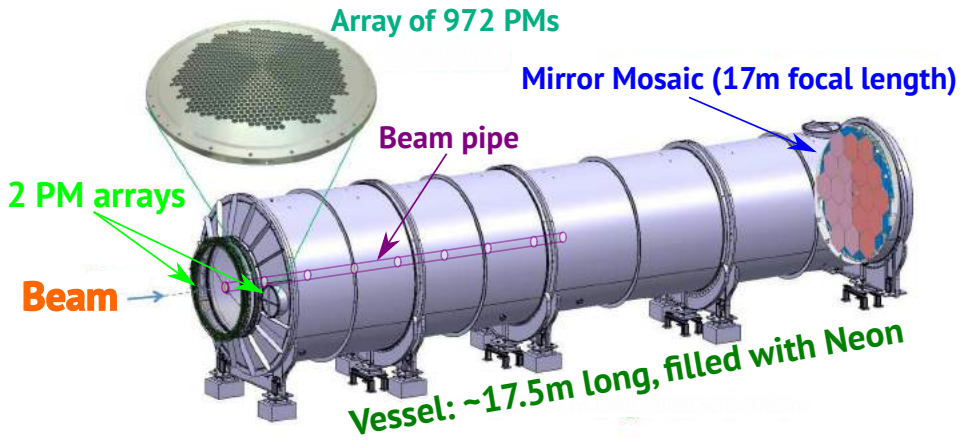


Figure 3.9: (Adapted from [200].) Annotated schematic view of the RICH.

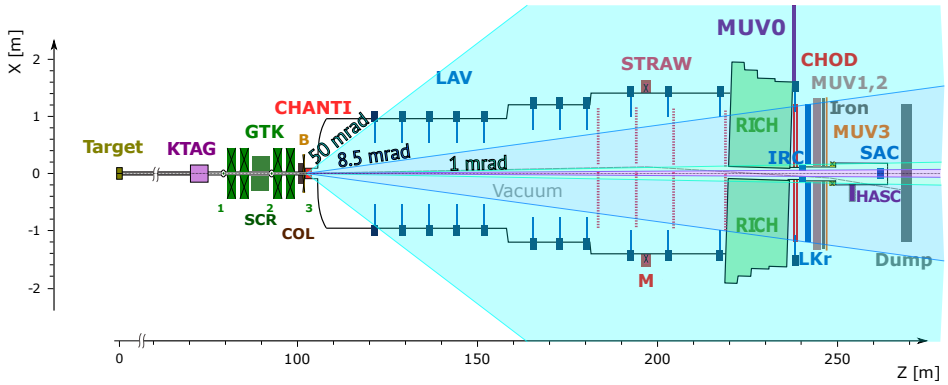


Figure 3.10: Detector schematic([200]) with shaded overlay indicating angular coverage of photon veto systems: LAV $8.5 - 50$ mrad (cyan), LKr $1 - 8.5$ mrad (blue), and < 1 mrad IRC (light green) and SAC (purple).

which can be detected efficiently. A critical photon veto rejection power with $\pi^0(\to \gamma\gamma)$ rejection of order $\mathcal{O}(10^8)$ is achieved through combined use of these detector systems. The measured photon veto inefficiency for individual systems is shown in figure 3.11.

3.2.7.1 Liquid Krypton Calorimeter (LKr)

The Liquid Krypton calorimeter (LKr) is a quasi-homogeneous electromagnetic calorimeter formed of a cryostat filled with 9000 litres of liquid krypton at 120 K. The LKr is re-used from the NA48 experiment [209] but with updated readout systems to match the high rate NA62 environment. The LKr energy resolution is

$$\frac{\sigma_E}{E} = \frac{4.8\%}{\sqrt{E}} \oplus \frac{11\%}{E} \oplus 0.9\%\,, \tag{3.2.2}$$

(where E is measured in GeV), as evaluated with simulations and degraded with respect to the NA48 operating conditions due to additional upstream material and non-linear energy response characteristics.

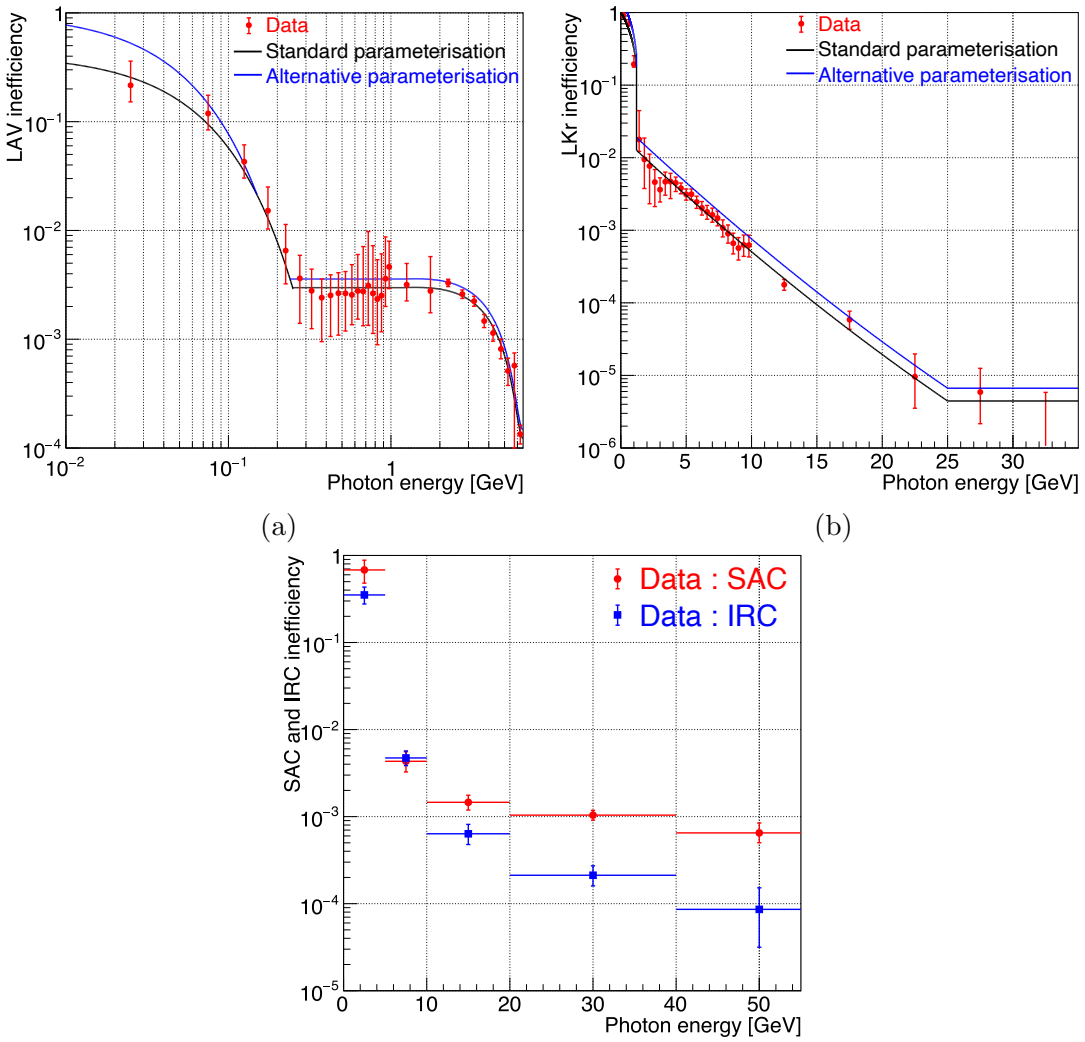


Figure 3.11: Measured photon veto inefficiency as a function of energy for the LAV (a), LKr (b) and SAC (c) subsystems, provided by T. Spadaro and L. Peruzzo as part of the search for $\pi^0 \to$ invisible decays [207]. The LAV and LKr inefficiencies are shown with a 'standard' parameterisation best fitting the data points and an 'alternative' parameterisation which is constrained to reach unity at a photon energy of 0 GeV and the difference with respect to the standard parameterisation contains the known systematic uncertainties. For the SAV detectors the standard and alternative models are taken to be the central values and central values multiplied by 1.2 respectively in each photon energy bin. These models were constructed by E. Goudzovski as part of the search for production of heavy neutral leptons [208].

Electrically charged particles and photons passing through the liquid krypton instigate electromagnetic showers starting with bremsstrahlung or pair production interactions respectively. Through repeated bremsstrahlung and pair production interactions a shower of many electrons and photons are produced. In the process energy is lost by the original particle and transferred to the calorimeter medium. The energy deposition in the calorimeter is measured through the charge produced in the shower which is picked up by preamplifiers in 13248 longitudinal cells of Cu-Be electrodes with size $2 \times 2\,\mathrm{cm}^2$ (c.f. the liquid krypton Moliere radius of 4.7 cm).

The LKr has an active fiducial volume with a regular octagon cross section with apothem 1.130 m and a depth (along the z direction) of 1.27 m which corresponds to 27 radiation lengths ($27\,X_0$). Electromagnetic showers develop logarithmically, each photon or $e^{\pm}$ creating a pair of new particles. This continues until photons fall below the electron pair production threshold and energy losses of $e^{\pm}$ starts to dominate. The longitudinal development of the electromagnetic shower is characterised by the radiation length, X_0, which is equal to the distance over which a high energy electron loses a fraction $1/e$ of its energy through bremsstrahlung and 7/9 of the mean free path of pair production for a photon in matter. With a depth of $27\,X_0$ the LKr therefore contains effectively the full shower produced by photons and $e^{\pm}$. However, pions instigate both electromagnetic and hadronic showers and this shower is unlikely to be fully contained in the LKr. Muons act as minimum ionising particles (MIPs), depositing only a small amount of energy as they traverse the detector material. Because of these characteristics, the ratio, E/p, of the energy deposited in the LKr to the momentum measured by the STRAW for a charged particle can be used for particle identification, alongside its function as part of the photon veto system.

3.2.7.2 Large Angle Photon Vetos (LAV)

Twelve Large Angle photon Veto stations are positioned along the length of the experiment, with LAV1-11 interspersed along the vacuum tank surrounding the decay volume and LAV12 located after the RICH and 3 m upstream of the LKr. The positioning ensures full angular coverage for photons emitted in π^0 decays from $8.5 - 50$ mrad.

Each LAV station is constructed from four or five staggered layers of annuli made up of lead glass blocks, recovered from the OPAL experiment's electromagnetic calorimeter barrel [210] wrapped in a new light-tight Tyveck covering, see figure 3.12. Cherenkov photons produced in electromagnetic showers in the lead glass blocks are detected by PMTs optically coupled to the blocks via light guides. The design ensures incident particles traverse active material in at least three rings corresponding to $21\,X_0$.

During standard operation the beam muon halo is responsible for a dominant fraction of the detected LAV signals. To keep the probability of accidental coincidence to the percent level the LAV time and energy resolutions for $\geq 1\,\mathrm{GeV}/c$ photons are 1 ns and 10% respectively.

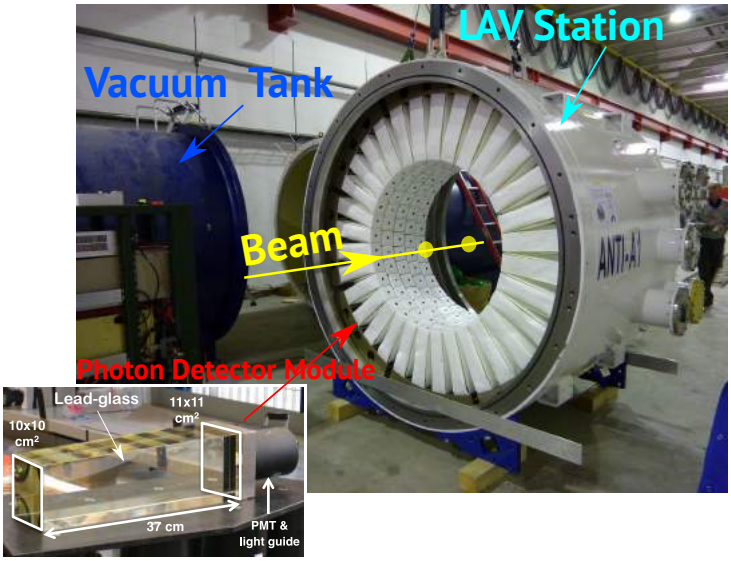


Figure 3.12: Annotated photograph of a LAV station during its installation and photon detector module during construction (adapted from [211] and [200] respectively).

3.2.7.3 Small Angle Photon Vetos (SAV)

The intermediate ring calorimeter (IRC) and small angle calorimeter (SAC) are sampling calorimeters with a shashlik design of alternating 1.5 mm layers of lead absorbers and active plastic-scintillator plates traversed by WLS fibres and read out by PMs, see figure 3.13. Photons from $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$ decays in the decay volume produced at low angles have energies $E_\gamma > 5\,\text{GeV}$ and are detected with inefficiency of $\leq 10^{-3}$.

The IRC has an eccentric cylindrical shape, with outer diameter 290 mm and central bore of diameter 120 mm. It surrounds the beam pipe upstream of the LKr, filling the photon veto coverage down to the beam pipe. The total depth corresponds to $19\,X_0$. There are two modules with upstream and downstream sections of length 98 mm and 154 mm with 25 and 45 layers respectively. Each has four optically isolated quadrants reducing cross-talk between read-out channels.

The SAC is constructed from 70 pairs of scintillator and lead square-shaped layers (interspersed with tyveck creating diffuse reflection and thus increasing light collection) with side length 205 mm and total depth corresponding to $19\,X_0$. The measured energy resolution is [212]

$$\frac{\sigma_E}{E} = \frac{8.8\%}{\sqrt{E}} \oplus \frac{7.1\%}{E}\,. \tag{3.2.3}$$

WLS fibres are grouped into four bundles, each read out by a PM. The SAC is installed centrally at the end of the beam line inside the vacuum such that photons which travel along the beam direction are intercepted while charged particles are deflected into the beam dump. A rotation of the SAC by 23 mrad ensures photons cannot remain undetected by traversing only WLS fibres.

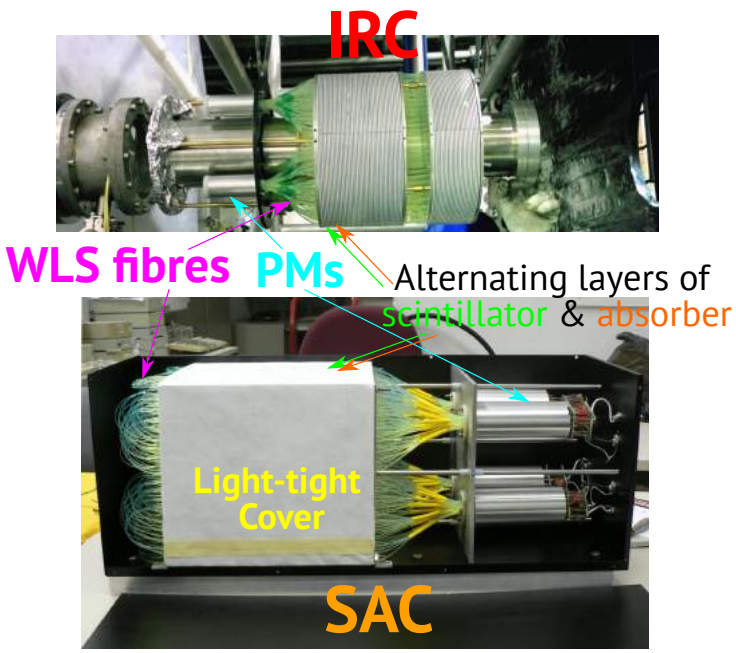


Figure 3.13: [200]Annotated pictures of the IRC and SAC during assembly.

3.2.8 Charged Particle Hodoscopes (CHODs)

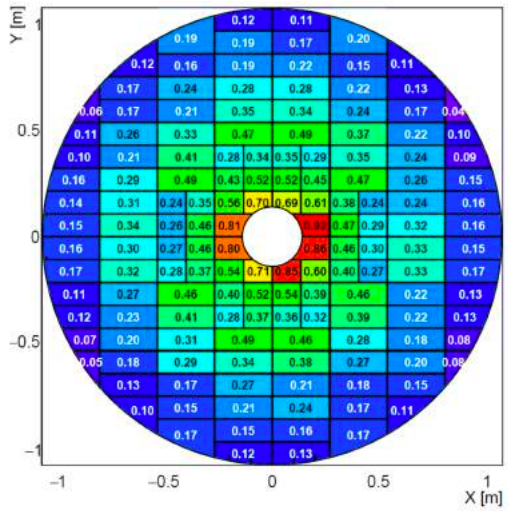
The charged particle hodoscopes, the CHOD and NA48-CHOD, are scintillator counters positioned between the RICH and LKr with active area covering radii between the IRC outer radius (145 mm) and LAV12 inner radius (1070 mm).

The NA48-CHOD is re-used from the NA48 experiment [209], see figure 3.14b. Two planes of 64 slabs of 20 mm ($0.10\,X_0$) thick scintillators, arranged vertically in the first plane and horizontally in the second and split at the centre, are read out by PMs at either side of the detector. For trigger logic the detector is divided into four quadrants defined by perpendicular pairs of half-planes. Two time measurements (reducing tails in time distributions) and a position measurement are provided, associating signals produced in the two planes with a charged particle track. A time resolution of 270 ps is measured in 2017 and 2018 data.

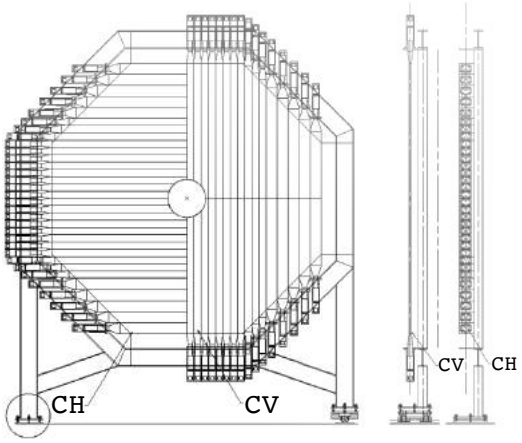
The CHOD is constructed from a mosaic of 152 plastic scintillator tiles covering an annular region between radii of 0.140 m and 1.070 m. Tiles are arranged with a 1 mm overlap with their neighbours and vary in sizes, increasing towards the edges to optimise the distribution of signal rates between readout channels, see figure 3.14a. The tiles are 30 mm thick, with the active region having a total longitudinal thickness of $0.13\,X_0$. Scintillation light produced in each tile is read out via WLS fibres by SiPM pairs. The time resolution measured in 2017 and 2018 data is 0.8 ns.

3.2.9 Hadronic Calorimeters

A hadronic sampling calorimeter is formed from the combination of the MUV2, the refurbished front module of the NA48 hadron calorimeter [209] (but reversed in direction along



(a) [200] Rate (in MHz) of signals from each CHOD tile during data-taking at the design instantaneous beam intensity of 750 MHz. Central high-occupation tiles are a result of beam π^+ decays.



(b) [200] Drawing of the NA48-CHOD displaying half-sections of the horizontal and vertical planes (labelled CH and CV respectively) in the front view, with beam at the centre (left), and the arrangement of the two planes in the side view (right).

Figure 3.14: Diagrams showing the CHOD tiles and NA48-CHOD slabs.

the beam line), and MUV1, with finer transversal segmentation for greater discrimination between hadronic and electromagnetic shower components. The NA62 calorimetric system is illustrated in figure 3.15.

Showers instigated by hadronically interacting particles, for example $\pi^{\pm}$, begin when an interaction with a nucleus produces secondary hadrons, including p, n, $\pi^{\pm}$ and π^0. Further interactions of hadrons follow, creating more secondaries logarithmically until their energy drops sufficiently to make further interactions unlikely. The development of this hadronic shower is characterised by the nuclear interaction length, λ_I, which is generally much larger than the radiation length X_0. The MUV1 and MUV2 combined have a depth corresponding to 8 interaction lengths, with 4.1 and 3.7 interaction lengths contributed by the MUV1 and MUV2 respectively. Furthermore, since the π^0 produced in the hadronic shower decays as $\pi^0 \to \gamma\gamma$, electromagnetic showers can be instigated by the high energy photons. In this way hadronic showers are accompanied by an electromagnetic shower component. This characteristic is used to distinguish between $e^{\pm}$ and $\pi^{\pm}$ at NA62: $e^{\pm}$ will only instigate electromagnetic showers in the LKr and will not reach the MUV1 and MUV2, while the π^+ may begin to shower in the LKr or in MUV1 and MUV2 with a combination of hadronic and electromagnetic showers. The shower characteristics further assist in distinguishing $\pi^{\pm}$ from $\mu^{\pm}$ because $\mu^{\pm}$ act as MIPs. Full exploitation of the calorimetric systems (LKr, MUV1 and MUV2) for PID results in efficiencies for π^+ identification and μ^+ misidentification of order $\mathcal{O}$(75–80%) and $\mathcal{O}(10^{-7}$–$10^{-6})$ respectively [125].

The MUV1 consists of interleaved layers of 26.8 mm thick steel plates and 9 mm thick scintillator strips, 60 mm in width and 2.620 m long (spanning the full detector width or height). Consecutive scintillator planes are alternately aligned horizontally and vertically to provide both x and y coordinate measurements for shower development.

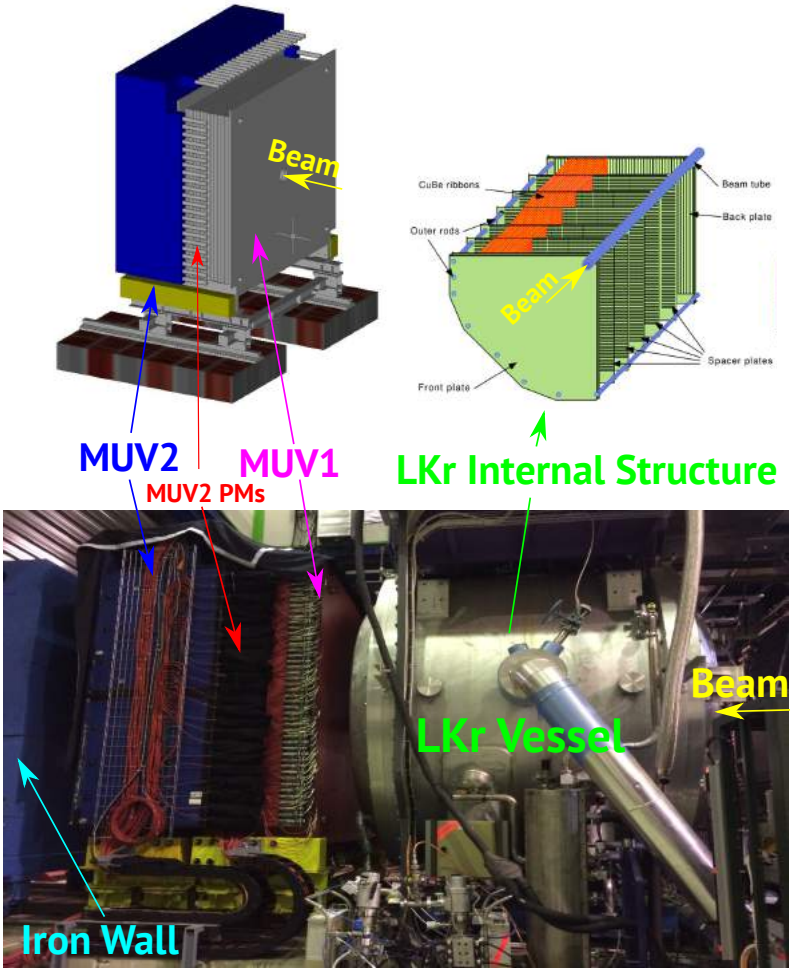


Figure 3.15: [200] [213] Annotated photograph and illustrations of the NA62 calorimetric system.

Each scintillator strip is read out at both ends by WLS fibres coupled to PMs, with all parallel strips with analogous positions in all layers read out by the same PM. For a MIP, depending on the impact point, 35 to 50 photoelectrons are detected in one PM.

The MUV2 is positioned downstream of the MUV1 with PMs arranged externally and protruding forwards around the MUV1. MUV2 has a similar construction with 24 iron plates 25 mm thick followed by plastic scintillators read out via lightguides by PMs.

The performance, evaluated using 2015 data, is characterised by the time resolution of 0.9 ns and energy resolution

$$\frac{\sigma_E}{E} = \frac{0.38}{\sqrt{E}} \oplus \frac{1.37}{E} \oplus 0.115\,, \tag{3.2.4}$$

where E is measured in GeV.

3.2.10 Fast Muon Veto (MUV3)

The MUV3 detector is positioned downstream of the hadron calorimeter and an 0.8 m iron wall (constituting an additional 4.8 λ_I), see figure 3.15. Only charged particles which traverse the full calorimetric system, representing over 14 λ_I, are detected. While $e^{\pm}$

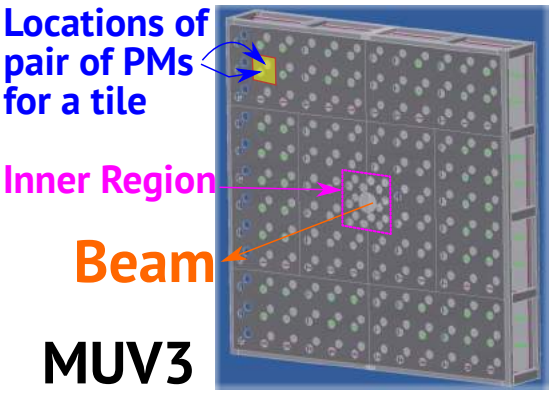


Figure 3.16: [213] Annotated drawing of the MUV3 detector structure.

are absorbed in the LKr and $\pi^{\pm}$ showers are contained in the calorimetric systems, $\mu^{\pm}$ act as MIPs and reach the MUV3. The MUV3 thus tags $\mu^{\pm}$: useful either for particle identification or vetoing.

The MUV3 is a square plane, of side length 2.640 m, constructed from an array of 148 scintillator tiles 50 mm thick, 140 'outer tiles' with size 200 mm × 200 mm and eight smaller 'inner tiles' surrounding the central beam pipe, see figure 3.16. Each tile is read out by a pair of PMs. If a $\mu^{\pm}$ passes through a PM window (probability of 8% for a standard tile) Cherenkov photons are generated which arrive before scintillation photons and affect time reconstruction. Despite this, a time resolution of 0.5 ns is obtained, assisted by the presence of PM pairs for each tile helping distinguish early signals from Cherenkov photons. For $\mu^{\pm}$ of momenta over 15 GeV/c the measured identification efficiency is in excess of 99.5%.

3.2.11 Additional Veto Detectors

Additional veto detectors, the peripheral muon veto (MUV0) and hadronic sampling calorimeter (HASC), are used to detect $\pi^{\pm}$ from $K^{+} \to \pi^{+}\pi^{+}\pi^{-}$ decays which do not traverse the STRAW chambers. The MUV0 is a scintillator hodoscope detector positioned to the left of the RICH exit flange and is used to detect π^{-} with momenta less than 10 GeV/c deflected into this region by the magnet. The HASC is a hadronic sampling calorimeter formed from 9 modules each with 60 pairs of lead plates and scintillators 100 mm × 100 mm in size. The HASC is located downstream of the MUV3 and final bend magnet and is designed to detect π^{+} of momentum greater than 50 GeV/c which pass through the holes at the centres of the STRAW chambers.

3.2.12 Trigger and Data Acquisition System (TDAQ)

In order to efficiently collect a large volume of high quality data in the high intensity NA62 environment a trigger system selectively reads out detector information for events with properties of potential interest for physics analysis. The trigger operates with a two-level structure with a hardware-based L0 selection followed by a L1 decision based on the

outcomes of software algorithms run on partially reconstructed data. During data-taking the beam is provided to the experiment in 'bursts' lasting about 3 s. The TDAQ system subdivides this into 'clock cycles' lasting 24.951 ns (40.079 MHz) defined by a high quality signal generator with quartz crystal phase-locked loop jitter-cleaning system, guaranteeing high accuracy and stability. This is distributed using unidirectional optical fibres with two multiplexed channels, one carrying synchronous L0 trigger signals and the other carrying asynchronous trigger type (8-bit trigger word) and reset information.

Packets of information, called trigger primitives, from the NA48-CHOD, CHOD, RICH, LKr and MUV3 detectors are sent to the L0 trigger processor (L0TP) for each event. These primitives contain information regarding which pre-defined conditions (trigger bits) are satisfied independently in each system. The L0TP then will accept the 'event' (defined by coincident trigger primitives) if the collection of primitives satisfy a set of conditions (trigger masks). In the simplest example of the minimum bias control (CTRL) trigger an event is accepted at L0 if coincident signals exist in matching x and y tiles of the NA48-CHOD (from the passage of a charged particle). Given a positive decision by the L0TP a request is sent for readout of the rest of the detectors for a level 1 decision. Software algorithms, run on dedicated machines in a PC farm, use information read out from the KTAG, STRAW and LAV detectors to make a L1 decision which, if passed, means the full detector information will be stored to tape for later analysis.

The L0 trigger decision reduces the rate of data output from the expected rate of K^+ decays in the decay volume of $\mathcal{O}(5\,\text{MHz})$ (plus additional rate from beam $\pi^\pm$ and their decays and the muon halo leading to a total rate of $\mathcal{O}(10\,\text{MHz})$) to a design maximum output rate of 1 MHz. This is further reduced by the L1 trigger to write data to permanent storage at $\mathcal{O}(10\,\text{kHz})$. The trigger system has in-built flexibility, in principle the High Level Trigger System (HLT) is composed of both a level 1 and 2 software trigger, however in practise only L1 algorithms have been used during 2016–18 data-taking.

A set of triggers with different requirements are run simultaneously. However, due to limited bandwidth a downscaling factor, D, is applied to almost all triggers such that only every D^{th} event fulfilling the requirements of the trigger is accepted.

The data acquisition relies on readout boards for each detector subsystem, with a TDC(time to digital converters)-based TDAQ system for all detectors except the GTK, STRAW and calorimeters which each have bespoke systems. In the general case a FPGA(field-programmable gate array)-based TEL62 integrated circuit board is used with TDC board (mezzanine cards) which can digitise leading and trailing edge time signals of 128 channels with 100 ps resolution. For the GTK a modular system of custom boards is used to interface between the trigger system and the on-detector readout boards. The STRAW uses CARIOCA readout chips [214] with FPGAs and custom back-end Straw Readout Boards (SRBs). Calorimeters use a bespoke CREAM (Calorimeter REadout Module) readout system [215].

3.2.13 Detector Development

Driven by continuous analysis of data the NA62 detector apparatus is undergoing continuous development. The additional veto detectors, discussed in section 3.2.11, were added during 2016 to cover gaps in detection acceptance for $K^+ \to \pi^+\pi^+\pi^-$ decays identified from analysis of 2015 data.

By design possible backgrounds for the study of the $K^+ \to \pi^+\nu\bar{\nu}$ decay arising from decays and interactions in the upstream region (before the decay volume) were minimised. However, upon detailed analysis of the the 2016 and 2017 data such backgrounds were revealed to be the largest single component, as discussed in section 5.3. To further block a specific background mechanism where upstream K^+ decays produce π which pass through a gap in the BEND6 magnet (called 'mambas', see section 5.3.1) a 'plug' was installed during 2017 [216], followed by replacement of the final collimator (Col 6/7) with one of larger dimensions (TCX) in 2018 to fully block such π^+ [201].

In addition, new upstream veto (scintillator) detectors are planned for installation in 2021 to detect particles in the upstream region related to possible background mechanisms. The ANTI0 detector is a plane of scintillator tile detectors, using technology similar to the CHOD and MUV3, to be positioned just upstream of the start of the vacuum tank. Its inclusion into the L0 trigger in anti-coincidence is expected to reduce the L0 trigger rate by 30%. A veto-counter detector will be positioned above and below the beam pipe after the final GTK magnet and before the final collimator with two scintillator stations separated by an absorber, with a third station after the final collimator. This will detect decay products from upstream decays allowing suppression of background.

The GTK stations, despite radiation-hard design, must be replaced after $\mathcal{O}(1\,\text{year})$ of beam exposure. During continued production of GTK stations the thickness has been decreased further reducing the amount of material traversed by beam particles. Addition of a fourth GTK station is planned for 2021 which will provide improved time and spatial resolution and assist with rejection of background arising from multiple particles passing through the detector concurrently.

3.3 Physics Programme

Beyond the primary physics goal of studying the $K^+ \to \pi^+\nu\bar{\nu}$ decay the NA62 physics programme can be divided into three thematic areas:

1. Studies and precision measurements of rare K^+ decays.
2. Searches for forbidden K^+ decays.
3. Searches for production and/or decays of 'exotic' (BSM) particles.

Moreover, since the NA62 beam contains a large number of π^+ and the common $K^+ \to \pi^+\pi^0$ leads to large numbers of π^0, decays of π^+ and π^0 can also be studied. A review

of relevant decays is given in section 2.6. Within this work contributions to each of these areas is presented, with contributions to the study of $K^+ \to \pi^+\nu\bar{\nu}$ in chapter 5 (sections 5.2, 5.3 and 5.5) fitting the first thematic area, searches for LNV/LFV $K^+ \to \pi\mu e$ decays, reported in chapter 7, fitting the second, and searches for production of a BSM particle X in chapter 6 relating to the third.

3.3.1 Precision Measurements and Rare Decays

The rare FCNC $K^+ \to \pi^+\mu^+\mu^-$ decay is under investigation. Comparison with theoretical calculations [217] [218] [219] [220] [221] [222], generally in the context of chiral perturbation theory, provides an opportunity to test the SM and the complex computational techniques needed to realise the predictions relevant to this decay. Moreover, in analogy to the R_K LFU ratio defined in the B physics sector an analogous ratio for K decays can be defined,

$$R_K^{K\ell\ell} = \frac{\mathcal{B}(K^+ \to \pi^+\mu^+\mu^-)}{\mathcal{B}(K^+ \to \pi^+e^+e^-)}\,. \tag{3.3.1}$$

A precision measurement of $\mathcal{B}(K^+ \to \pi^+\mu^+\mu^-)$ from NA62 [223], combined with previous precision measurement of $\mathcal{B}(K^+ \to \pi^+e^+e^-)$ from NA48/2 [224] provides an independent test of LFU in the K sector. The NA62 experiment has collected the world's largest ever sample of $K^+ \to \pi^+\mu^+\mu^-$ and the precision of the measurement is currently limited by systematics [223] but still able to reach comparable precision to the NA48/2 measurement [225]. A further LFU test can be provided by studying the ratio

$$R_K^{K\ell\nu} = \frac{\mathcal{B}(K^\pm \to e^\pm\nu_e)}{\mathcal{B}(K^\pm \to \mu^\pm\nu_\mu)}\,, \tag{3.3.2}$$

as already performed using 2007 NA62-R_K data [226] [227] with further possibilities to improve with the new NA62 dataset [228].

Precision measurements of the $\pi^0 \to e^+e^-$ decay are planned [228], which may help confirm or refute a slight tension that has recently arisen between the SM prediction [229] and experimental result [230].

3.3.2 Searches for Forbidden Decays

Searches for forbidden K^+ multi-body decays are actively pursued at NA62 with recent published results on LNV and LFV processes $K^+ \to \pi^-\mu^+\mu^+$ and $K^+ \to \pi^-e^+e^+$ [34]. Searches for LNV/LFV $K^+ \to \pi\mu e$ and $\pi^0 \to \mu^\pm e^\mp$ decays are addressed in this thesis. Beyond this, other searches are in progress such as for the 4-body LNV and LFV decays $K^+ \to \ell_1^-\nu_{\ell_1}\ell_2^+\ell_2^+$ with $\ell_{1,2} = \mu$ or e [228].

3.3.3 Searches for Production and Decay of BSM Particles

The NA62 experiment primarily takes data with the TAXes open allowing the secondary hadron beam to pass into the experiment and allowing studies of K^+ decays to be per-

formed. Recent searches for exotic particles using normal data taking conditions include searches for dark photons [78] and heavy neutral leptons [231] [232]. However, the experiment is also able to operate in 'beam dump mode' where the beryllium target is removed and the TAXes are closed and act as a secondary target in which exotic BSM particles could be produced. Operationally the transition between standard and beam dump operation is simple and is achieved in a matter of minutes. In 2016–18 data-taking a beam dump sample was collected corresponding to 3×10^{16} protons on target (POT). This already allows NA62 to begin to challenge current bounds on NP models, constrain the BSM phase space or potentially discover a BSM signal [228].

Chapter 4

Cedar/KTAG Simulation Studies

4.1 Cedar Optics

4.1.1 Simplified 'Single Mirror' Cedar

Cherenkov photons are emitted at constant angle θ, given by equation 3.2.1, along the path of a charged particle as it traverses the radiator gas inside the Cedar vessel. In a simplified scenario, if a spherical mirror is positioned at the end of the Cedar, a ring image with fixed radius R is produced in the focal plane of the mirror. The diaphragm is then located in this focal plane and

$$R \approx f\theta\,, \tag{4.1.1}$$

where f is the focal length and a small angle approximation has been made. The velocity of a particle of mass m and momentum p is

$$\beta = \left(1 + \frac{m^2}{p^2}\right)^{-\frac{1}{2}}\,, \tag{4.1.2}$$

and since there is a 1% RMS spread of momentum of beam particles, while the difference between beam particle masses is large ($m_p \gg m_{K^+} \gg m_{\pi^+}$), the difference in velocities between two particle species is [204]

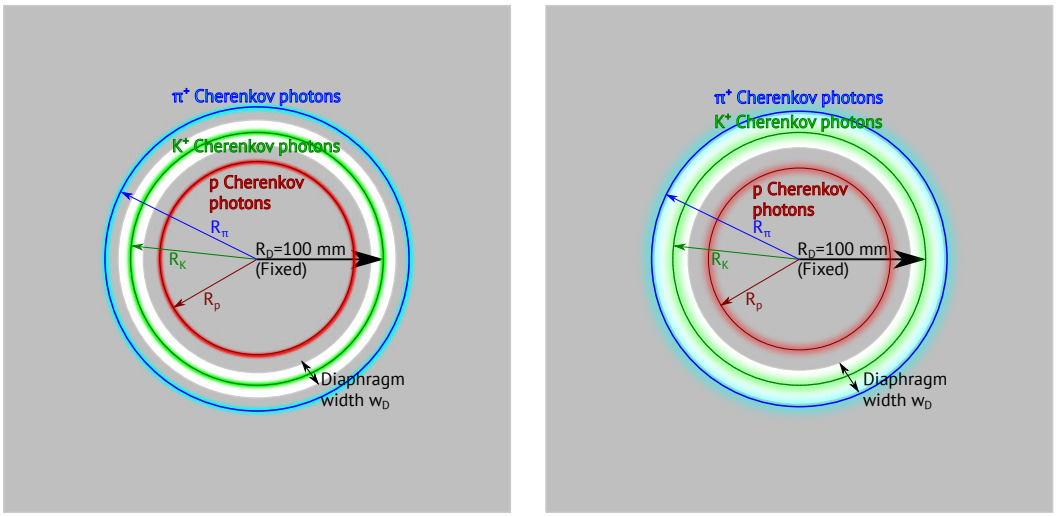
$$\frac{\Delta\beta}{\beta} \approx \beta_2 - \beta_1 \approx \frac{m_2^2 - m_1^2}{2p^2}\,, \tag{4.1.3}$$

valid for high momentum beams $\frac{m}{p} \ll 1$. Differentiating equation 3.2.1 and applying a small angle approximation the difference in angle $\Delta\theta$ is given by

$$\Delta\theta = \frac{1}{\tan\theta}\frac{\Delta\beta}{\beta} \approx \frac{1}{\theta}\frac{\Delta\beta}{\beta}\,, \tag{4.1.4}$$

and therefore the difference in ring radius for two particle species is

$$\Delta R \approx f\Delta\theta \approx \frac{f}{\theta}\frac{m_2^2 - m_1^2}{2p^2}\,. \tag{4.1.5}$$



(a) (b)

Figure 4.1: Sketches of Cherenkov rings focused on the diaphragm plane arising from passage of π^+, K^+ and p beam particles through the Cedar. In case (a) rings are well-focused with relatively low width and well-separated, as expected in standard N_2 data-taking conditions. In case (b) rings are broadened by, for example chromatic dispersion or incompatible chromatic corrector and radiator gas, and with the diaphragm width increased with respect to (a).

The optics of the Cedar-West used for the NA62 Cedar/KTAG system are optimised for use with nitrogen gas at 1.7 bar [204], operated at the ambient temperature of the experimental hall, approximately $20°C$. The NA62 Cedar effective optical focal length is $f = 3.24\,\text{m}$ and the gas pressure (which determines the refractive index) is tuned to give Cherenkov angle $\theta = 30.9\,\text{mrad}$, which leads to a K^+ producing a ring of radius 100 mm to match the fixed central diaphragm radius, see figure 4.1. With this operating point the difference between ring radii of K^+ and π^+ or protons are

$$\Delta R_{K\pi} = 2.1\,\text{mm} \tag{4.1.6}$$
$$\Delta R_{Kp} = 5.9\,\text{mm}\,. \tag{4.1.7}$$

In principle therefore, by controlling the width of the diaphragm, w_D see figure 4.1a, K^+ rings alone can be selected. However, in this simplified 'single-mirror Cedar' scenario both spherical aberration of the mirror and chromatic aberration arising from photons travelling through the radiator gas, broaden the ring image at the diaphragm which can lead to overlap of the photon distributions produced by two particle species, see figure 4.1b.

4.1.2 Full Cedar Optics

Elimination of Spherical Aberration

By using a Mangin mirror instead of a spherical mirror spherical aberration is eliminated through specific choice [233] of radii of the front (refracting) surface, r_1, and rear (reflecting)

surface, r_2. The focal length of the Mangin mirror, f_M, is given by [1] [233]

$$\frac{1}{f_M} = \frac{2n}{r_2} - \frac{2(n-1)}{r_1} \ , \tag{4.1.8}$$

where n is the refractive index of the lens which is constructed from fused silica and can be parameterised by the empirical Sellmeier formula [234] as [235]

$$n = \sqrt{1 + \frac{B_1\lambda^2}{\lambda^2 - C_1} + \frac{B_2\lambda^2}{\lambda^2 - C_2} + \frac{B_3\lambda^2}{\lambda^2 - C_3}} \ , \tag{4.1.9}$$

where λ is the wavelength expressed in μm and the appropriate numerical factors are $B_1 = 0.6961663$, $C_1 = 0.0684043$, $B_2 = 0.4079426$, $C_2 = 0.1162414$, $B_3 = 0.8974794$, $C_3 = 9.8961610$. For the NA62 Cedar $r_1 = 8.610\,\text{m}$, $r_2 = 6.615\,\text{m}$ and the Mangin mirror focal length is $f_M = 5.034\,\text{m}$ for photons of wavelength $\lambda = 300\,\text{nm}$.

Chromatic Dispersion/Aberration and Chromatic Correction

When using a Mangin mirror the primary cause of ring broadening is chromatic aberration, with different wavelengths of light being refracted by slightly different angles and therefore causing a spread in the ring image at the diaphragm. The resulting divergence is given by [204]

$$\Delta R_{ch} = f\frac{1}{2\nu}\left(1 + \frac{1}{\theta^2\gamma^2}\right) \ . \tag{4.1.10}$$

The Abbe number is $\nu = \frac{n_D - 1}{n_F - n_C}$, with n_i being refractive indices of a gas at wavelengths $\lambda_F = 486.1\,\text{nm}$, $\lambda_D = 589.3\,\text{nm}$, and $\lambda_C = 656.3\,\text{nm}$ and for N_2 gas in the NA62 Cedar [204] $\nu = 35$. Under standard NA62 operation $\Delta R_{ch} = 1.9\,\text{mm}$ which is comparable to $\Delta R_{K\pi}$. However, for high momenta beams and for a fixed choice of gas in a detector f, θ and ν are fixed so ΔR_{ch} is approximately constant and depends only on γ. For the case of K^+–π^+ separation at high momentum a specific fixed chromatic corrector can therefore be used. Correction of transverse and longitudinal chromatic aberrations are performed by different optical components, the first being the Mangin mirror itself and the second being a separate chromatic corrector lens, see figure 3.5. Using this system a correction factor k is added such that $\Delta R_{ch}^{corr} = \frac{\Delta R_{ch}}{k}$. For the NA62 Cedar $k = 11$ and so $\Delta R_{ch}^{corr} = 0.17\,\text{mm}$.

The chromatic corrector is a planoconvex (converging) lens with radius of the curved surface, $r_C = 1.385\,\text{m}$ for the NA62 Cedar. The focal length is given by

$$\frac{1}{f_C} = \frac{n-1}{r_C} = \frac{1}{t} - \frac{1}{f_M - d} \tag{4.1.11}$$

where n is given by equation 4.1.9, $d = 3.508\text{m}$ is the distance between the Mangin mirror and Chromatic corrector and $t = 0.998\,\text{m}$ is the distance between the chromatic corrector and the final focal plane of the optical system, as shown in figure 4.2. The combination of the Mangin mirror and chromatic corrector produces a ring image at the diaphragm plane (the focal plane of the optical system) with spherical and chromatic aberration effects eliminated to a good precision for all optical wavelengths. However, additional effects, described below, can still broaden the ring images at the diaphragm plane.

[1] With sign convention chosen to give $f_M > 0$.

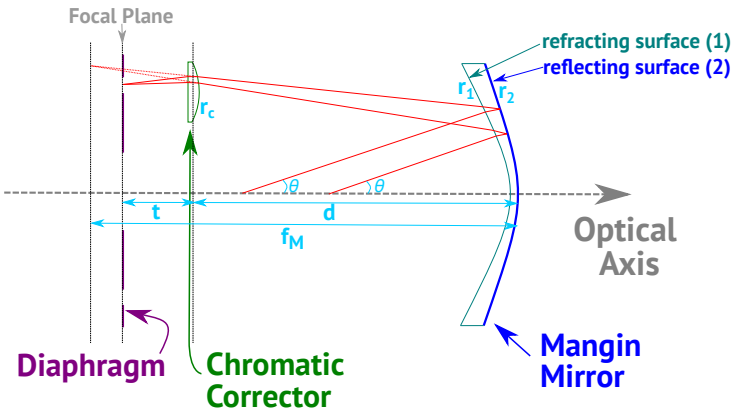


Figure 4.2: Schematic diagram of Cedar optics.

Multiple scattering of beam particles

After the chromatic correction the largest effect is due to multiple scattering of beam particles traversing the gas which leads to several non-concentric ring images with RMS scatter given by [204] [2]

$$\sigma_{R_{scat}} = f\frac{15}{\beta p}\sqrt{\frac{LP}{3X_0}}\,, \tag{4.1.12}$$

where momentum p is expressed in MeV, L is the length of gas traversed expressed in radiation lengths (5.6 m of N_2 for the NA62 Cedar corresponds to $3\%X_0$) and P is the gas pressure in bar. For effective focal length $f = 3.24\,\text{m}$ this gives $\sigma_{R_{scat}} = 0.09\,\text{mm}$.

Beam divergence

A similar effect occurs if beam particles do not always follow parallel trajectories. For an angular beam divergence characterised by σ_x and σ_y the corresponding change in R is

$$\sigma^{x,y}_{R_{div}} = f\sigma_{x,y}\,. \tag{4.1.13}$$

The NA62 beam divergence is required to be less than $100\,\mu\text{rad}$ [200] such that $\sigma^{x,y}_{R_{div}} < 0.32\,\text{mm}$ and is held to $\sigma^{x,y}_{R_{div}} < 0.3\,\text{mm}$ in practise.

Inhomogineity of refractive index

Inhomogeneities in the refractive index of the radiator gas can create refraction and degrade the focus at the diaphragm plane. The refractive index gradient is proportional to the temperature and pressure gradients [204]: $\frac{\Delta n}{n} = \frac{n-1}{n}(\frac{\Delta P}{P} - \frac{\Delta T}{T})$. Pressure is kept fixed but small temperature gradients are possible, but unlikely due to the stable cavern temperature in which the NA62 Cedar is operated. The standard deviation of the spread of photons at the diaphragm due to a linear temperature gradient is given by [204]

$$\sigma_{R_{\Delta T}} = \frac{1}{2\sqrt{3}}\frac{1-n}{n}\frac{f}{\tan\theta}\frac{\Delta T}{T}\,, \tag{4.1.14}$$

For a large temperature differential of $1K$ a spread $\sigma_{R_{\Delta T}} \approx 0.05\,\text{mm}$ results, which is smaller than the above effects and this effect can generally be neglected to first order.

[2] This is a simplified empirical relation consistent with calculating the RMS angular scatter using the standard formula [13] $\theta_{RMS} = \frac{13.6\,\text{MeV}}{\beta cp} z\sqrt{\frac{x}{X_0}}\left[1 + 0.038\ln\left(\frac{xz^2}{X_0\beta^2}\right)\right]$ and equation 4.1.1.

4.2 Choice of Radiator Gas: N_2 versus H_2

A critical aspect of the NA62 strategy is to precisely tag a K^+ with the Cedar/KTAG system and measure the associated track momentum in the GTK. Beam particle interactions with the Cedar gas can mean a K^+ is not properly tagged and can instigate a far-upstream background contribution. Use of H_2 gas instead of N_2 reduces the material budget for the gas traversed by the beam to $0.3\%X_0$ from $3\%X_0$ [213] [3] . The multiple scattering effect broadening rings at the diaphragm is also reduced by approximately a factor of 4.

However, the chromatic corrector lens designed for N_2 gas no longer corrects the chromatic dispersion and in fact can worsen the effect. This is a wavelength-dependent effect, as illustrated by figure 4.3 which shows the ring radii at the diaphragm for π^+ and K^+ as a function of wavelength. The increased chromatic dispersion means the diaphragm opening width w_D must be larger when using H_2 gas. However, to maintain reasonable K^+/π^+ discrimination the w_D must be limited and some component of the K^+ ring is lost, reducing the overall light yield, with respect to the standard operational use of N_2, by 30%. To reach the same operational angle $\theta = 30.9\,\text{mrad}$ a higher pressure is required for H_2 of approximately 3.7 bar compared to 1.7 bar for N_2. Operation with H_2 at higher pressure carries safety risks but these are accounted for in the NA62 design.

4.3 Pressure Scan

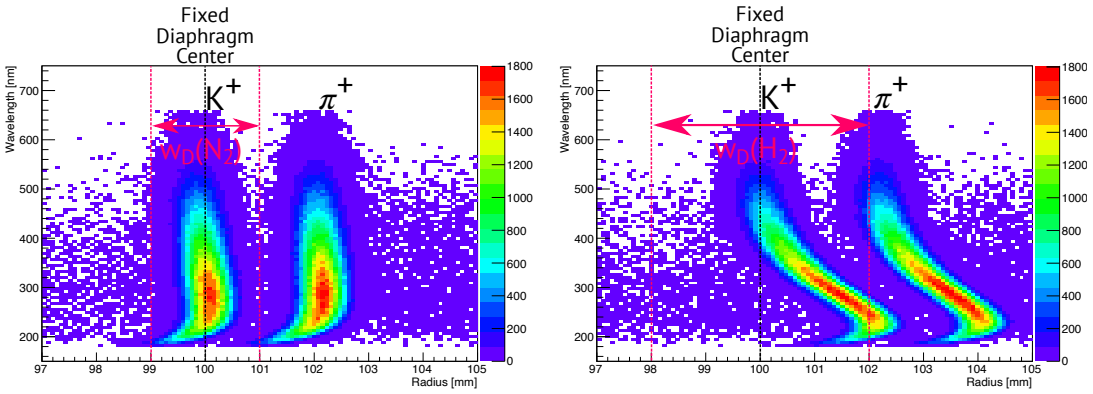
There are therefore positive and negative consequences of switching from the standard operation with N_2, used for 2016–18 data-taking, to use of H_2. However, in both cases a working point must be found which maximises light yield while maintaining sufficiently high K^+/π^+ discrimination efficiency. Assuming the detector is well-aligned (see section 3.2.2), all that can be changed is the gas choice (N_2 or H_2), pressure of the gas, P, and w_D. To determine a working point for a given gas choice a pressure scan is performed, varying the pressure of the gas over a fixed range and measuring the number of photons, and multi-sector coincidences, detected. Results of an experimental pressure scan performed in 2015 with N_2 radiator gas are shown in figure 4.4. The value of w_D is tuned to obtain good separation of the K^+ and π^+ peaks whose (central) positions are known from the pressure scan. However the shape (width) of peaks observed does depend on the value of w_D used for the pressure scan.

To predict a reasonable working point for future operation with H_2 gas simulated pressure scans were performed. Four specific cases will be shown below for concreteness:

Case 1: 2016–18 standard data-taking conditions: N_2, $w_D = 2.0\,\text{mm}$.

Case 2: Potential H_2 data-taking parameters: H_2, $w_D = 3.0\,\text{mm}$.

[3] The total material budget contributed by the Cedar includes the Aluminium entrance and exit windows ($150\,\mu\text{m}$ and $200\,\mu\text{m}$ thick respectively) which dominates the material budget when using H_2 radiator gas. The total material budget is $0.666\%\,X_0$ and $3.33\%\,X_0$ when using H_2 and N_2 radiator gas respectively.



(a) (b)

Figure 4.3: (Edited from [205].) Simulated position of Cherenkov photons produced by K^+ beam particles at the diaphragm plane, normalised for the quantum efficiency of the photon detection system, in the cases where the Cedar gas radiator is N_2 (a) and $H_2(b)$. The chromatic correction optics are designed for the N_2 radiator, high photon yield for K^+ with low contamination is achieved with diaphragm annulus width $w_D(N_2) = 2\,\text{mm}$ (for gas pressure 1.7 bar). In contrast for H_2 the chromatic dispersion is worse and a larger width, for example $w_D(N_2) = 4\,\text{mm}$, must be used (for H_2 at 3 bar). This includes larger contamination from π^+ Cherenkov photon at high wavelength and loss of light originating from K^+ at lower wavelength.

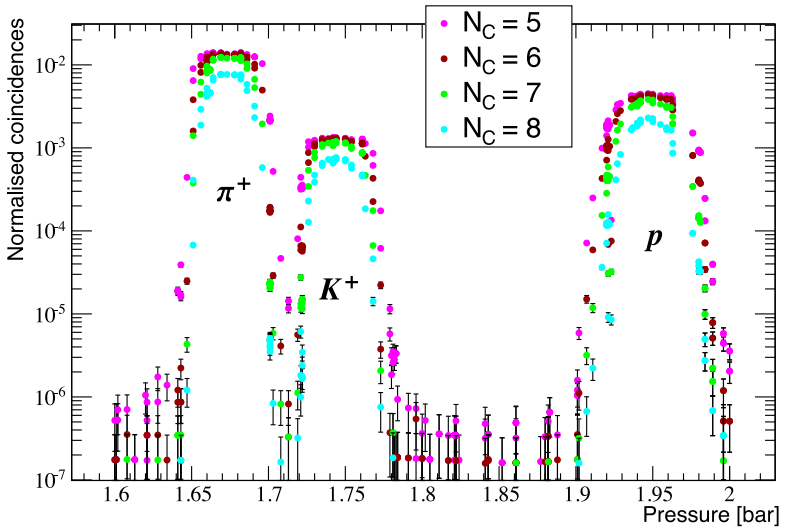


Figure 4.4: [200] Results from an experimental pressure scan performed in 2015 showing the three peaks due to Cherenkov photons associated with π^+, K^+ and p for cases of coincident detection in $5, 6, 7$ or 8 sectors.

Case 3: Standard N_2 pressure scan parameters: N_2, $w_D = 1.3\,\text{mm}$.

Case 4: Potential H_2 pressure scan parameters: H_2, $w_D = 1.3\,\text{mm}$.

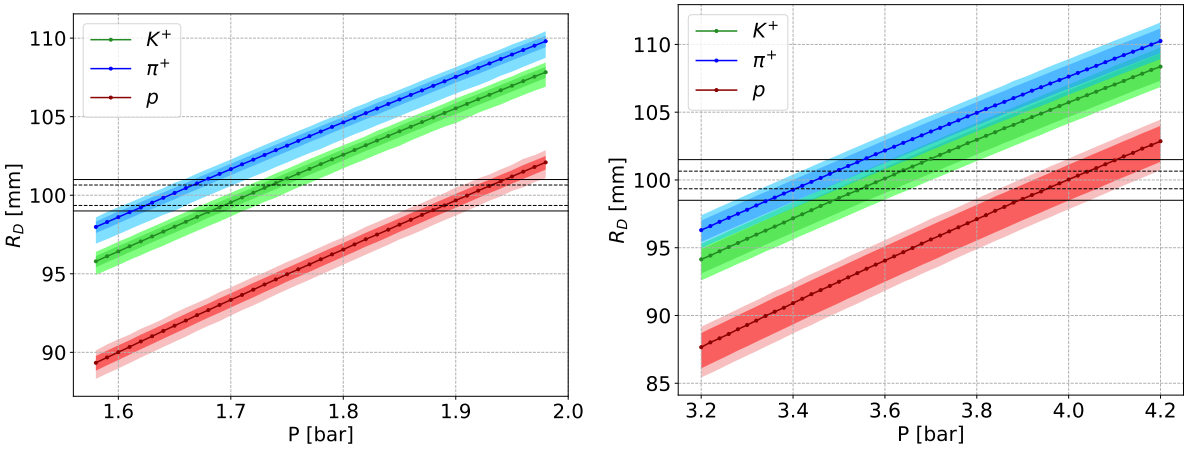
When performing a pressure scan the light yield need not be as high since efficiency of detection is not of primary concern, instead the objective is finding the position of the three peaks corresponding to π^+, K^+ and p. This explains the lower value of w_D used for H_2 pressure scan and provides the caveat that the fourth case is not necessarily an optimal data-taking set-up for H_2.

A sample of 5000 events is simulated for each of the three species of beam particles, π^+, K^+ and p, for a relevant range of gas pressure in steps of 0.01 (0.02) bar for $N_2(H_2)$ gas. The radii of photons reaching the diaphragm plane when using N_2 and H_2 radiator gas is shown in figures 4.5a and 4.5b respectively, with the mean shown by a solid line and two-tone bands illustrating the one and two standard deviation spread. Full simulation of the Cedar/KTAG system, including consideration of the optical transmissivity of components and quantum efficiency of the PM photon detectors, allows the expected values of measurable quantities to be determined. For each event the number of photons detected (hits) and coincident n-sector signals (candidates) are counted. The number of hits and candidates is first normalised to the number of beam particles per event, with average values of $< N_{hits} >= 23$ and $< N_{cand}^{5-sector} >= 0.997$ at the K^+ peak for standard data-taking case 1. The beam composition is then considered by scaling by the fraction of each species in the real beam - giving an overall normalisation per beam particle. The total normalised number of hits/candidates per beam particle is then provided by summing the three components, this is equivalent to what may be observed experimentally.

Results for the three stages described above for case 1 are shown in figures 4.6a and 4.6b for hits and $\geq$ 5-sector candidates respectively. For cases 2, 3 and 4 the results are each summarised with three plots, equivalent to those highlighted in figure 4.6, in figures 4.7a, 4.7b and 4.7c respectively.

Comparison of simulated pressure scans using N_2 (cases 1 and 3) and H_2 (cases 2 and 4) as the radiator gas, shows that in the latter case peaks are generally broader and show asymmetry, especially in the distribution of number of hits, due to chromatic dispersion. In cases 1, 2 and 4 the K^+ and π^+ peaks are not well separated, leading to contamination of π^+ at the standard K^+-peak-centre working point (at 1.71[3.61] bar for case 1[2]). In case 2 in particular at this standard working point the more numerous beam π^+ dominates and to reach a regime dominated by K^+ a higher pressure must be used, towards the upper trailing edge of the K^+ distribution. This leads to additional loss of light yield; shown by the lower mean number of photons detected (hits) shown in the left panels of figure 4.7; beyond the loss from chromatic dispersion.

In summary, the simulations show that a change to use H_2 gas will lead, in general, to greater overlap between π^+ and K^+ peaks, increasing π^+ mistag probability. To reach a lower mistag probability a higher pressure can be used but this leads to lower light yield which degrades K^+ tagging efficiency as well as time resolution. However, the benefits of switching to the lighter H_2 gas with reduced scattering were not a part of this study and are under further investigation. Only qualitative comparison with data is possible due to



(a) (b)

Figure 4.5: Photon radii at the diaphragm plane as a function of gas pressure for a Cedar-W using N_2 (a) or H_2 (b) as a Cherenkov radiator. Radii of Cherenkov photons produced by beam K^+, π^+ and protons, are shown and in each case the mean radii is shown by solid lines and points and the one and two standard deviation spread is illustrated by the two-tone bands. The diaphragm aperture is illustrated by the horizontal lines with: solid lines for $w_D = 2[3]$ mm, for standard data-taking conditions case 1[2]; and dashed lines for $w_D = 1.3[1.3]$ mm, for pressure scan conditions case 3[4]; when using $N_2[H_2]$, in (a)[(b)].

the special data-taking conditions of a pressure scan and potential shifts in peak positions due to temperature effects.

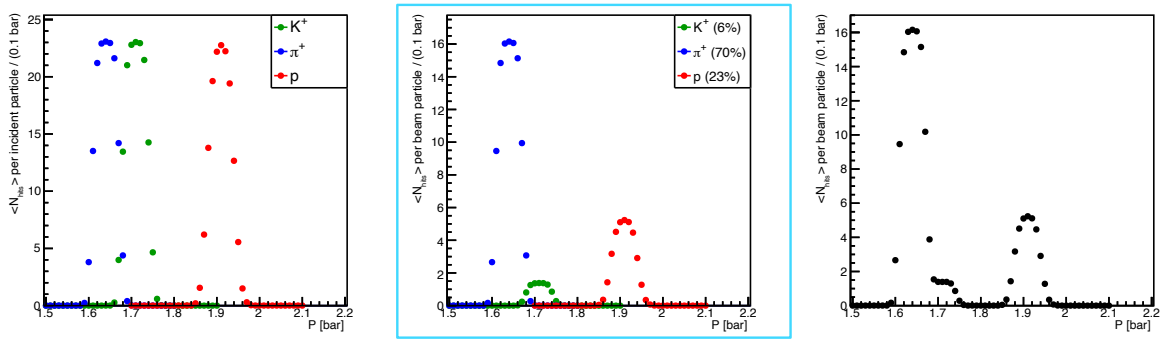
Following this study a project has been pursued to improve and enhance the Cedar/KTAG simulation, study the optical properties in detail and consider designs for a potential replacement Cedar with hydrogen, or other light radiator gases, as well as possible use of a Cedar-North [204] instead of a Cedar-West[4]. An available Cedar-North at CERN is under consideration for this purpose and a design study is underway for operation with hydrogen gas with modifications to the optics of this Cedar-North to mitigate issues highlighted by the above and subsequent studies. This has been designated the 'Cedar-H'.

4.4 Cedar Internal Optical Alignment

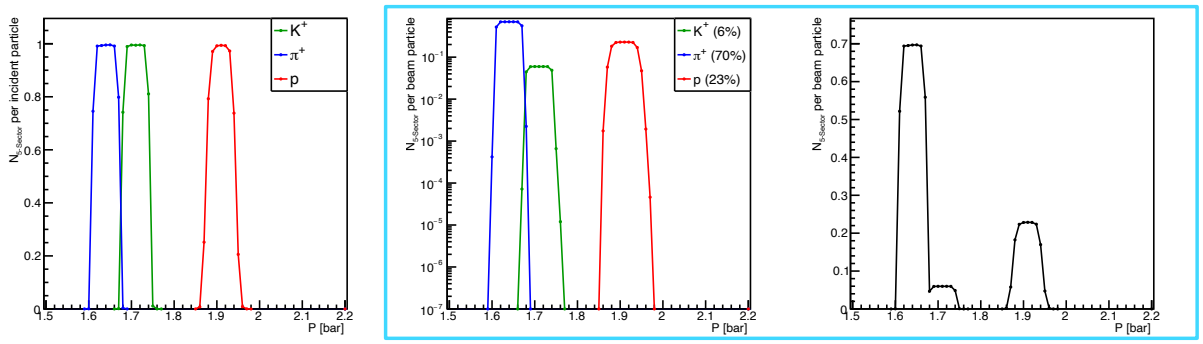
As described in section 3.2.2, the Cedar optical axis is aligned to the beam axis using asymmetry in detected photon numbers between opposing KTAG sectors. However, this assumes that the internal optical components of the Cedar are themselves aligned to the Cedar optical axis. Such an internal optical alignment was last performed for the CERN Cedars in the 1970s, however the original apparatus has been retained and will be re-used to re-align the optics after modifications are made for the Cedar-H.

The internal alignment strategy is depicted schematically in figure 4.8. A photon

[4]These studies were performed primarily by collaborators at the University of Birmingham and will be reported in future NA62 theses.

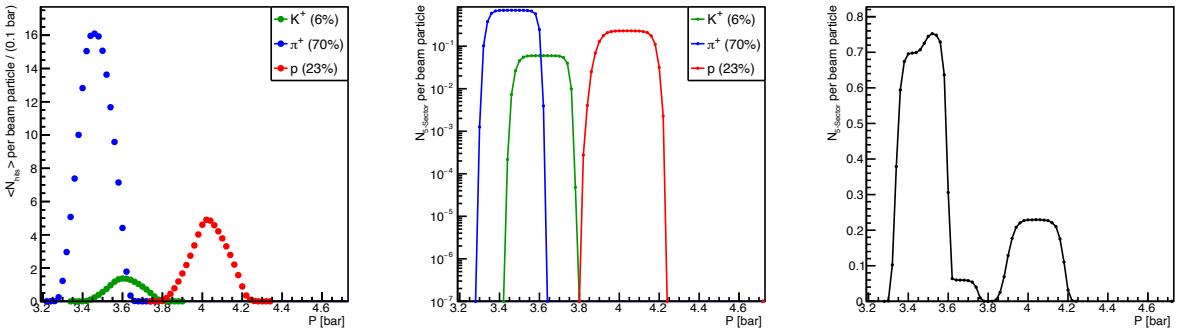


(a) Case 1 : N_2, $w_D = 2.0\,\text{mm}$: number of hits.

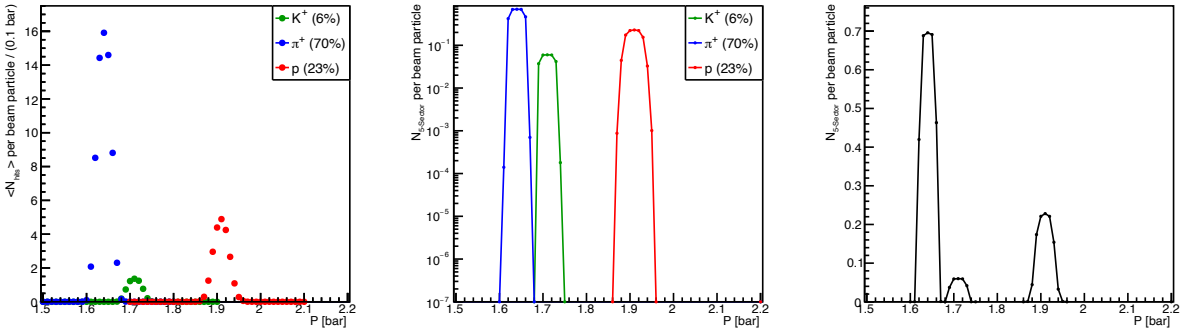


(b) Case 1 : N_2, $w_D = 2.0\,\text{mm}$: number of 5-sector candidates.

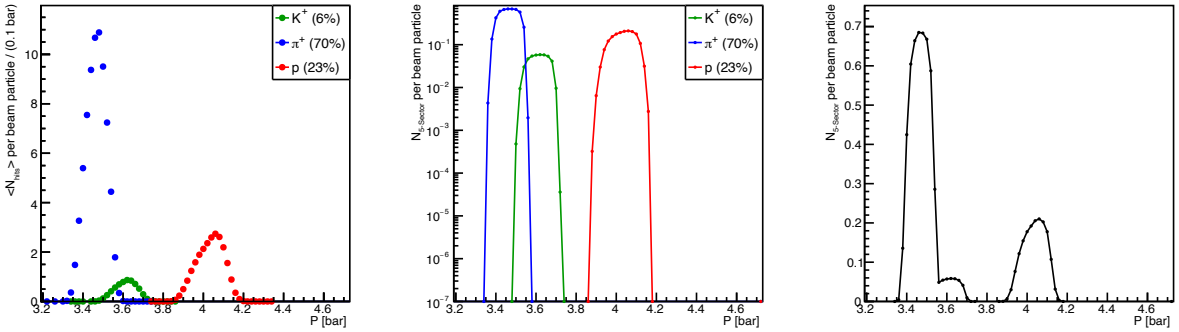
Figure 4.6: Simulated pressure scan (Case 1 : N_2, $w_D = 2.0\,\text{mm}$) showing the mean number of: (a) hits and (b) ≥ 5 sector coincidence candidates; normalised to number of incident particles simulated (left) and number of beam particles, plus scaling to account for beam composition (centre and right). Cases are shown with Individual π^+, K^+ and p peaks (left and centre) and with the three components summed (right), as would be observed in data.



(a) Case 2 : H_2, $w_D = 3.0\,\text{mm}$

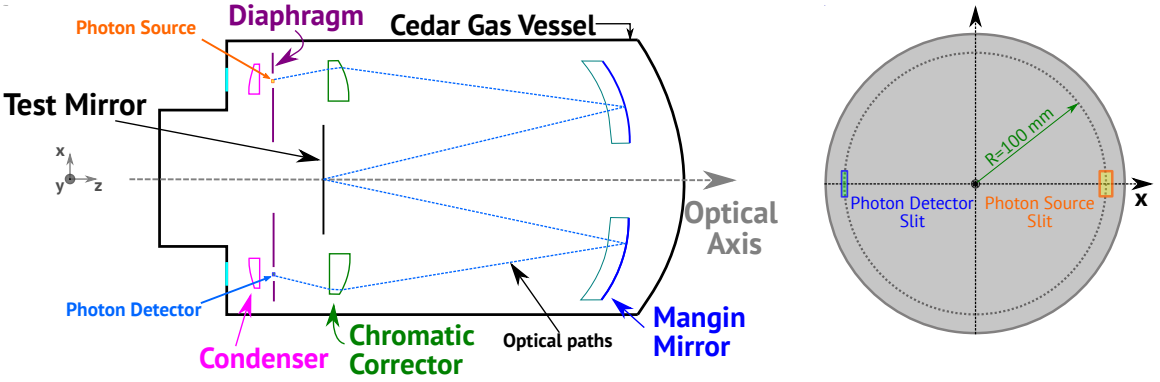


(b) Case 3 : N_2, $w_D = 1.3\,\text{mm}$



(c) Case 4 : H_2, $w_D = 1.3\,\text{mm}$

Figure 4.7: Simulated pressure scan results for: (a) Case 2 : H_2, $w_D = 3.0\,\text{mm}$, (b) Case 3 : N_2, $w_D = 1.3\,\text{mm}$ and (c) Case 4 : H_2, $w_D = 1.3\,\text{mm}$. The mean number of hits and 5-sector coincidence candidates, normalised to the number of incident beam particles are displayed in the left and centre plots respectively, with scaling to account for beam composition. Cases are shown with individual π^+, K^+ and p peaks (left and centre) and with the three components summed (right), in terms of mean number of 5 sector coincidence candidates, as would be observed in data.



(a) (b)

Figure 4.8: (a) Schematic of internal Cedar optical alignment procedure. A Cedar with all components perfectly aligned is depicted, with the Mangin and test mirror both perpendicular to the optical axis. Photons emitted from the photon source traverse the optical components and reach the detector on the opposite side of the optical axis in the diaphragm plane. (b) Schematic view of the diaphragm plane for alignment with photon source and detector slits illustrated.

source is positioned at the diaphragm plane at a radius of 100 mm (position $x = 100$ mm, $y = 0$ mm), the diaphragm itself may be removed. Photons traverse the chromatic corrector and are reflected at the Mangin mirror (in a direction opposite to Cherenkov photons in standard operation). A test mirror is introduced which reflects photons which then traverse the remainder of the Cedar optics, imitating Cherenkov photons originating from beam particles travelling along the optical axis in standard operation. A detector, positioned opposite the source (at $x = -100$ mm, $y = 0$ mm), records photons if the optical components are aligned. The photon source has a slit with dimensions 1 mm $\times$ 0.02 mm and the detector entrance slit is 1 mm $\times$ 0.01 mm. Positions of the photon source and detector may be changed by rotating the alignment set-up about the optical (z) axis.

The simulation of this internal Cedar alignment set-up has been added to the NA62 software, for both Cedar-West and Cedar-North optics and with programmable tilts about the x and y axes of the Mangin mirror (MM) and test mirror (TM). The location of the test mirror is different for the two types of Cedar, due to the differences in the optical components. Simulation studies using existing Cedar designs, presented below, are assisting in the understanding of the alignment system in preparation for alignment of the Cedar-H.

The variation in x position at the detector plane over a range of visible wavelengths 290–635 nm is 0.072 mm and 2.3 μm for the Cedar-West and Cedar-North respectively. For the Cedar-West in particular (c.f. x-width of the detector slit, 0.01 mm), this means the choice of light source wavelength is important. The use of a green light source is proposed with wavelength $\lambda = 532$ nm and this is used for simulation studies reported below.

Photons emitted at the centre of the source slit at angles, from the z axis in the x-z plane, of $\theta_x^{emit} = 37.18(18.12)$ mrad reach the centre of the detector slit for the Cedar-

West(North) if optical components are perfectly aligned. The full area of the source slit is illuminated using an optical fibre. Diffraction effects from the thin slit cause a spread in the angle of emitted light, the first diffraction minima are located at an angle

$$\theta_{dm} = \sin^{-1}\left(\frac{\lambda}{w}\right) , \tag{4.4.1}$$

with respect to the central emission angle, where w is the slit width. For wavelength $\lambda = 532\,\text{nm}$ and the emission slit widths in the x and y direction of $w_x = 0.02\,\text{mm}$ and $w_y = 1\,\text{mm}$, respectively, the corresponding first diffraction minima occur at angles $\theta_{dm,x} = 26.6\,\text{mrad}$ and $\theta_{dm,y} = 0.532\,\text{mrad}$. The range of photon emission angles which traverse the fully aligned optical system and reach the detector plane are

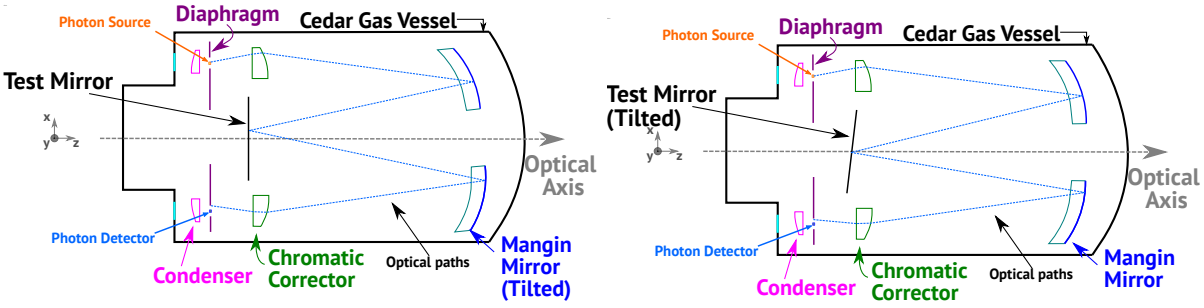
$$\theta_x^{emit} = 24.1\text{–}50.0\,\text{mrad} \text{ for Cedar-West} , \tag{4.4.2}$$

$$\theta_x^{emit} = 8.3\text{–}26.8\,\text{mrad} \text{ for Cedar-North} . \tag{4.4.3}$$

With respect to the central emission angles given above this angular range is within the central diffraction maximum, however the intensity of light detected will vary as a function of emission angle. The 'light spot' produced at the detector plane is therefore elliptical with height $s_y = 1.01(1.001)\,\text{mm}$ and width $s_x = 0.505(0.0485)\,\text{mm}$ for the Cedar-West(North). This exceeds the size of the detector slit dimensions, in particular the x-width of 0.01 mm. The detector slit may be moved in the x direction which allows a scan of the diffraction maximum intensity distribution to be performed, allowing finer determination of the centre of the 'light spot'.

If the Mangin mirror is tilted, for example a positive (anticlockwise) rotation about the y axis, the path of the photon test beam is deflected in the positive x direction and misses the photon detector, as shown by figure 4.9a. This however assumes that the test mirror is perfectly aligned (perpendicular to the optical axis), if this is not the case a tilt of the test mirror, rotating in the clockwise direction about the y axis, causes the photon test beam to be deflected in the same direction, as shown in figure 4.9b, assuming that the Mangin mirror is aligned perpendicular to the optical axis.

Simulations have allowed the impact of tilts of the MM or TM to be understood qualitatively and quantitatively. Denoting the tilt of the MM about the x and y axes as $(\theta_{MM,x}, \theta_{MM,y})$, where $\theta_{MM,x/y}$ are expressed in milliradians, and analogously $(\theta_{TM,x}, \theta_{TM,y})$ for TM tilts, the position of optical photons (with wavelength 532 nm) at the diaphragm plane for the cases of: (a) MM tilt only, (b) TM tilt only and (c) MM tilt $(0, 1)$ plus TM tilt; are shown in figure 4.10 for the Cedar-West (left) and Cedar-North (right). Qualitatively it is observed that tilts of the TM produce a deflection a factor of two less and oppositely directed to a tilt in the MM. This is because two reflections occur at the MM compared to one at the TM and the reflecting surfaces are on opposite sides of the optical components. Moreover, the effects of tilting one component is independent of the tilt of the other, with the same pattern observed in figures 4.10b and 4.10c except the same constant displacement towards negative y due to the $(1, 0)$ MM rotation. Quantitatively the displacement of the photons at the diaphragm plane as a function of MM and TM tilts are shown in figure 4.11 for the Cedar-West (left) and Cedar-North (right). An empirical model for the displacement in the x and y planes is derived (for $\lambda = 532\,\text{nm}$ photons) to



(a) MM Tilt (anticlockwise about y axis) (b) TM Tilt (clockwise about y axis)

Figure 4.9: Schematic of internal Cedar optical alignment set-up with optical components aligned perpendicular to the optical axis except: (a) Mangin mirror tilted anticlockwise about the y axis; or (b) test mirror tilted clockwise about the y axis. Both simulate misalignment which causes deflection of the optical photons at the diaphragm plane towards positive x (smaller radius) causing them to miss the photon detector.

be

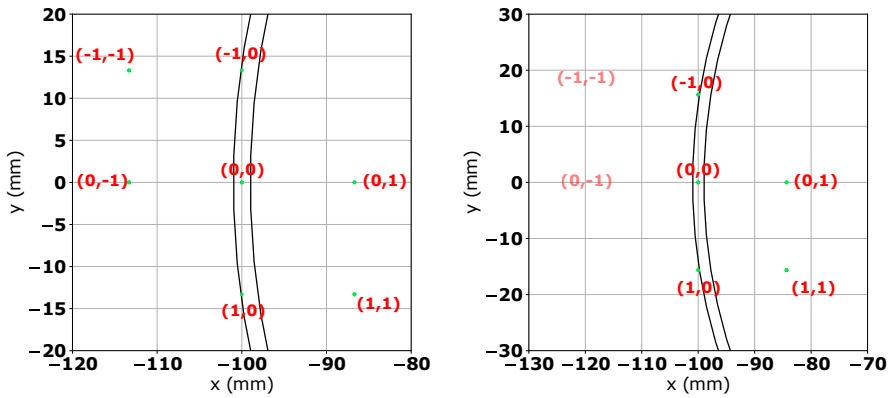
$$\Delta x = a\,\theta_{MM,y} + b\,\theta_{TM,y}\,, \tag{4.4.4}$$
$$\Delta y = -a\,\theta_{MM,x} - b\,\theta_{TM,x}\,, \tag{4.4.5}$$

where

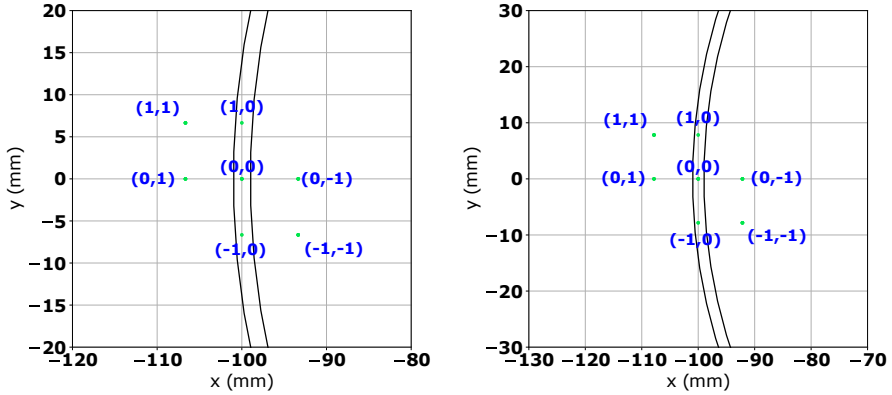
$$a = 13.3066 \text{ and } b = -6.6574 \text{ for Cedar-West}\ , \tag{4.4.6}$$
$$a = 15.6523 \text{ and } b = -7.8320 \text{ for Cedar-North}\,, \tag{4.4.7}$$

where angles are expressed in milliradians and displacements in millimetres. From these models it can be concluded that the Cedar-North is more sensitive to misalignment since the same tilt leads to approximately 20% larger displacement of photons at the diaphragm plane.

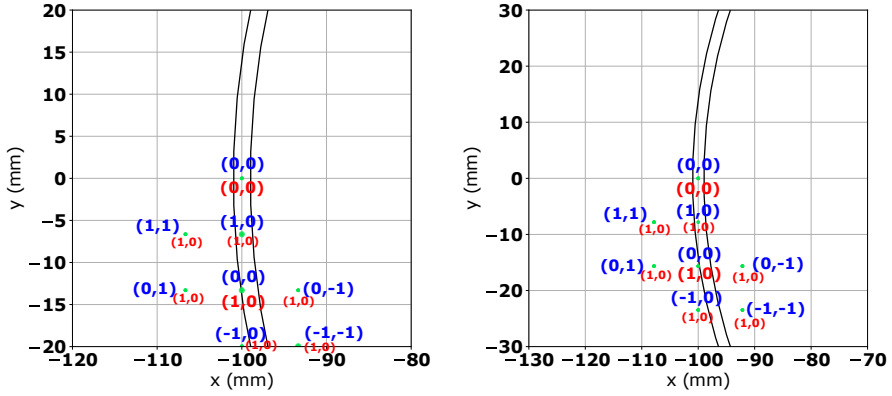
An independent study has shown that the MM must be aligned with a tolerance of 0.1 mrad. The alignment system is, in principle, sensitive to angular shifts of $|\Delta\theta_{MM}| = 0.039(0.0037)$ mrad for Cedar-West(North) meaning it is capable of measuring significantly smaller tilts than the required tolerance. This assumes however that the TM can be aligned perfectly, and because the same effect can be achieved tilting the MM and TM differently (creating degenerate solutions for minimising equations 4.4.4 and 4.4.5), the mechanical precision of TM alignment is important.



(a) Tilting MM with TM perfectly aligned.

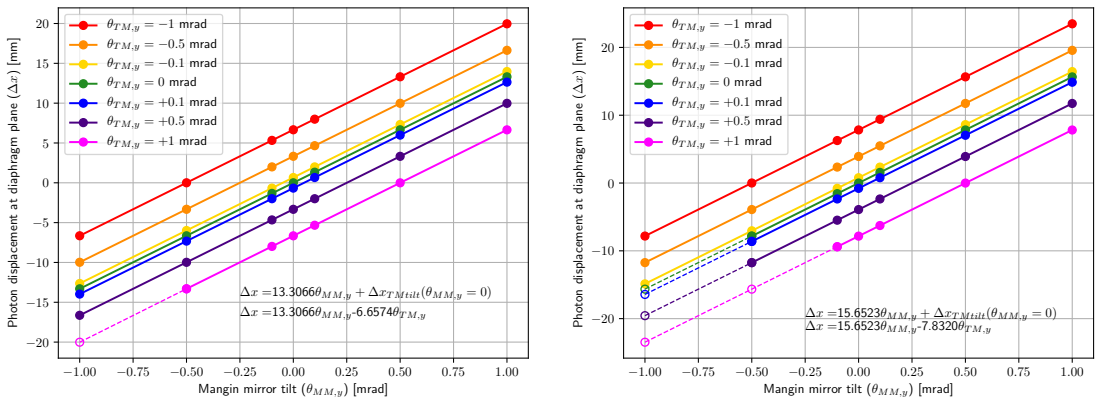


(b) Tilting TM with MM perfectly aligned.

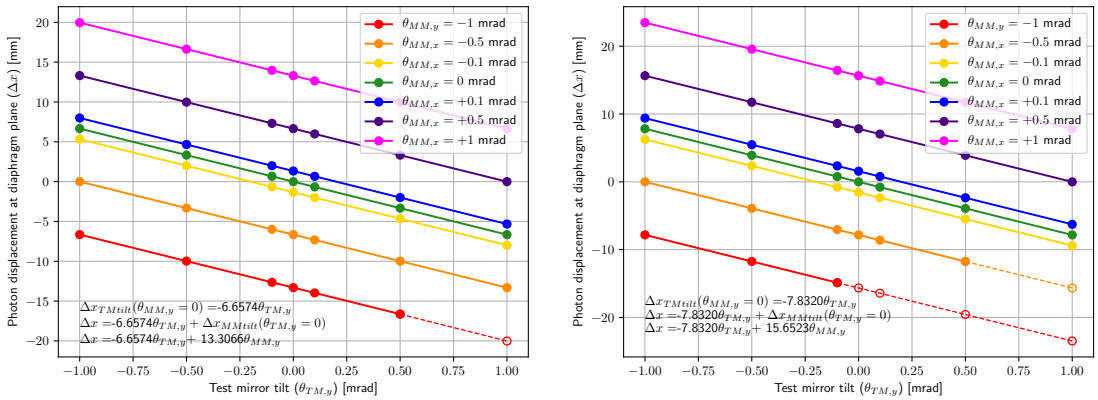


(c) Tilting TM with MM tilted by $(1, 0)$.

Figure 4.10: Positions of photons at the diaphragm plane in the vicinity of the photon detector for alignment ($x = -100\,\text{mm}$, $y = 0\,\text{mm}$) after traversing the Cedar-West (left) or Cedar-North (right) optics with Mangin mirror tilts ($\theta_{MM,x}, \theta_{MM,y}$) shown in red, and test mirror tilts ($\theta_{TM,x}, \theta_{TM,y}$) shown in blue. Cases are shown with (a) the test mirror aligned and a set of Mangin mirror tilts of 1 miliradian, (b) vice versa, and (c) constant Mangin mirror tilt of $(1\,[\text{mrad}], 0)$ and a set of test mirror tilts. In case (a) for Cedar-North with $\theta_{MM,y} = 1\,\text{mrad}$ photons miss the chromatic corrector.



(a)



(b)

Figure 4.11: Measured displacement of the alignment test beam of photons at the diaphragm plane in the x direction with respect to the centre of the photon detector (at $x = -100\,\text{mm}$, $y = 0\,\text{mm}$) as a function of (a) Mangin mirror and (b) test mirror tilts in the y direction. Results are displayed for the Cedar-West (left) and Cedar-North (right) optics, and a range of tilts for the other mirror are shown in different colour for each case. Open markers indicate the expected locations in scenarios where photons do not reach the detector plane due to missing the chromatic corrector after reflection from the Mangin mirror. The linear relationships observed are encapsulated in the empirical models displayed on the figures and equation 4.4.5.

Chapter 5

Studies of the Rare Decay $K^+ \to \pi^+\nu\bar{\nu}$

In the following the study of the $K^+ \to \pi^+\nu\bar{\nu}$ decay is presented. The analysis of the full Run 1 (2016-18) NA62 dataset has been completed and the preliminary result has been reported [89]; this is presented in section 5.7 together with future prospects. However, the focus of this chapter is the analysis of data collected in 2016 and 2017 (with associated publications [124] and [125]) and contributions of the author, presented in sections 5.2, 5.3 and 5.5. These contributions are to background studies, specifically for: $K^+ \to \pi^+\pi^+\pi^-$ decays (section 5.2); upstream backgrounds through use of simulations (section 5.3); and cross-checks of background expectations for $K^+ \to \pi^+\pi^0$ and $K^+ \to \mu^+\nu_\mu$ with an alternative method (section 5.5).

5.1 Fundamentals of the NA62 $K^+ \to \pi^+\nu\bar{\nu}$ Studies

5.1.1 Analysis Procedure

The signature of the $K^+ \to \pi^+\nu\bar{\nu}$ decay is an incident K^+ and an outgoing π^+ with missing energy/momentum. This may be kinematically isolated from potential background K^+ decay processes using the squared missing mass variable defined as

$$m^2_{miss} = (P_K - P_\pi)^2 \ , \tag{5.1.1}$$

where P_K and P_π are the 4-momenta of the K^+ and π^+ respectively, determined from the measured 3-momenta of the associated GTK and STRAW tracks and using the K^+ and π^+ mass hypotheses. Background suppression of order $\mathcal{O}(10^{-4})$ is achieved by definition of two signal regions such as to avoid the peaking two-body $K^+ \to \pi^+\pi^0$ and $K^+ \to \mu^+\nu_\mu$ decays and the kinematic endpoint of the $K^+ \to \pi^+\pi^+\pi^-$ decay. Signal regions 1 and 2 are shown in figure 5.1 and cover m^2_{miss} ranges 0–0.01 GeV^2/c^4 and 0.026–0.068 GeV^2/c^4 respectively. The momentum range for the π^+ candidate track is limited to 15–35 GeV/c, matching the design of the RICH and photon veto systems. The two-dimensional view of

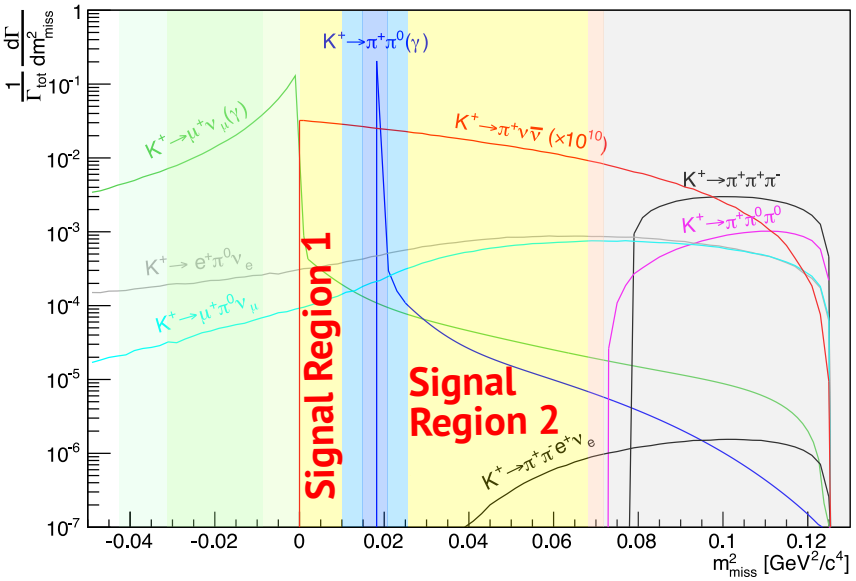


Figure 5.1: Squared missing mass regions definition, edited from [200]. The yellow regions are the signal regions, the pale blue regions are control regions either side of the $\pi^+\pi^0$ region (dark blue). The grey region is the 3π region and a small apricot control region separates this from signal region 2. Below $m^2_{miss} = 0$ the two pale overlapping regions indicate the $\mu\nu$ region and the associated control region with exact range being momentum dependent, see figure 5.2.

the signal regions in m^2_{miss} versus momentum is shown in figure 5.2. For the signal region definitions additional restrictions are applied based on alternative squared missing mass variables, as discussed in appendix A.

To suppress backgrounds with alternative final state particle species, particle identification (PID) requirements are applied, importantly distinguishing π^+ from μ^+ with inefficiency of $\mathcal{O}(10^{-7})$, and powerfully rejecting $K^+ \to \mu^+\nu_\mu$ decays. Photon veto conditions suppress $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$ decays, rejecting $\pi^0 \to \gamma\gamma$ with inefficiency $\mathcal{O}(10^{-8})$. Multi-body final states are distributed in m^2_{miss} but are rejected by a combination of PID requirements and multiplicity cuts, which reject additional activity recorded in the detector caused by additional visible particles from, for example $K^+ \to \pi^+\pi^+\pi^-$ or $K^+ \to \pi^+\pi^-e^+\nu_e$ decays. The selection procedure is reported in [124] and [125] with a short summary presented in section 5.1.3.

5.1.2 Data Samples, Trigger and Simulation Technique

Three results have been presented, analysing the data taken at the NA62 experiment in 2016 [124], 2017 [125] and 2018 [89]. Discussion of analysis of 2018 data is deferred exclusively to section 5.7. The 2016 dataset corresponds to 30 days of data-taking, and represents 2% of the full Run 1 (2016–18) sample, in terms of effective number of K^+ decays collected [228]. The 2017 data, 30% of the Run 1 sample, was taken at higher average intensity over a four month period [125] [228].

The first stage towards isolation of candidate $K^+ \to \pi^+\nu\bar{\nu}$ events is the trigger

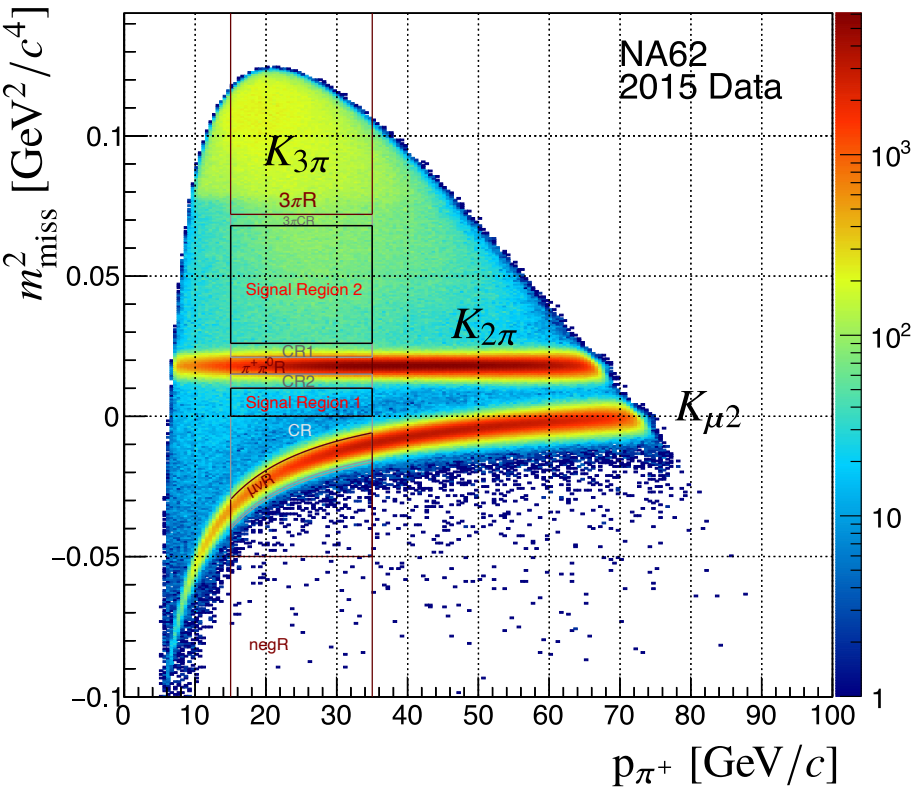
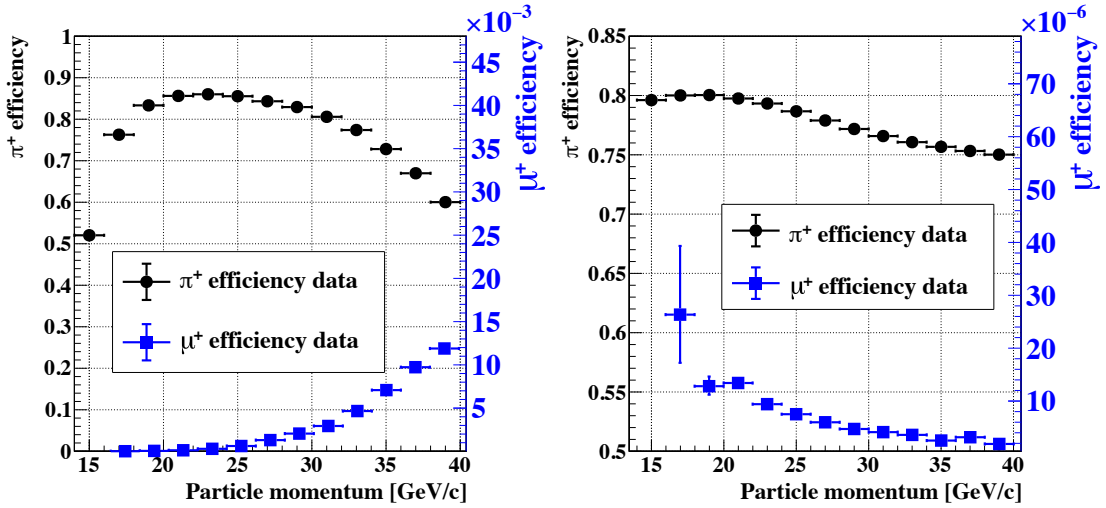


Figure 5.2: (Edited from [200], produced from single-track analysis of 2015 data without PID, photon veto or full multiplicity rejection requirements.) Two dimensional definitions of signal regions (bounded by black lines) control regions (grey lines) and background regions (dark red lines) in the m^2_{miss} versus momentum plane. Here the 2-body $K^+ \to \pi^+\pi^0$ decay is seen as a peak at $m^2_{\pi^0}$, while the $K^+ \to \mu^+\nu_\mu$ peak is seen as a curved distribution because of the misassignment of mass when calculating m^2_{miss}. The 3-body $K^+ \to \pi^+\pi^+\pi^-$ background is distributed diffusely at high m^2_{miss}.

system, as outlined in section 3.2.12, which instigates readout and storage of information from detectors if and only if certain criteria indicative of a signal-like event signature are met. The 'PNN' trigger was used for collection of the signal and is defined as follows. At L0 a time window of 6.3 ns is opened around a L0 RICH reference signal from a charged track, which defines the trigger time. Within this time window: a signal must exist in at least one CHOD tile but not in opposite quadrants (suppressing 3-track decays such as $K^+ \to \pi^+\pi^+\pi^-$); there are no signals in the MUV3 (rejecting $K^+ \to \mu^+\nu_\mu$); and there is a lone cluster topology in the LKr with energy deposit of less than 30 GeV (rejecting $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$ decays). At L1 coincident hits in ≥ 5 KTAG sectors are required with at least one STRAW track of momenta less than 50 GeV/c forming a vertex with the beam upstream of STRAW1 and signals in no more than two LAV blocks within 10 ns of the trigger time are allowed. The minimum bias control trigger, applying only a L0 condition of a signal in the NA48-CHOD and run concurrently with downscaling factor $D = 400$, is used to collect normalisation $K^+ \to \pi^+\pi^0$ events.

A GEANT4-based [236] [237] [238] Monte-Carlo (MC) simulation package is used to study the expected properties of the signal and background events and their detection in the NA62 detector. All relevant detector active and passive material is included in the simulation and readout properties and inefficiencies are accounted for. Decays of K^+ are simulated in the K^+ rest frame using appropriate form factors and matrix elements [125].



(a) (b)

Figure 5.3: [125] Particle identification probabilities (efficiencies) for π^+ identification and μ^+ misidentification as π^+ using RICH (a) and calorimeters (b) measured with 2017 data.

5.1.3 $K^+ \to \pi^+\nu\bar{\nu}$ Selection

Less than three STRAW tracks are allowed with none being negatively charged. Only one track can fulfil additional criteria to become a π^+ candidate but, for example, an additional out-of-time halo muon track may exist in the readout window of the STRAW reconstruction (the probability of which is intensity dependent). A π^+ candidate track must have momentum in the range 15–35 GeV/c and traverse the sensitive regions of downstream detectors (RICH, CHODs, LKr, and MUV1,2,3) with spatially and temporally matched signals in RICH, CHODs and LKr. The track time is measured to 100 ps precision and is calculated using the mean times measured in the STRAW, NA48-CHOD and RICH weighted by their respective measured resolutions. The particle identity is determined by the RICH, using the ratio of likelihoods under the π^+ and μ^+ hypotheses and particle mass calculated from the ring radius, and using a multivariate classifier using calorimetric information. Integrated over momentum a π^+ identification efficiency, of 82% and 78% respectively, is obtained, with probability of μ^+ misidentification as π^+ of 2.3×10^{-3} and 6.3×10^{-6}, see figure 5.3. No geometrically associated MUV3 signals within 7 ns of the track time are allowed. Events are rejected if any LKr clusters are present outside a 100 mm radius (c.f. LKr Moliere radius of 47 mm) from the π^+ impact point and in time windows, depending on cluster energies, of half width between 5 and 50 ns. Events are rejected if there is additional activity in detectors upstream of the LKr associated with either the decay vertex or LKr clusters in excess of 40 MeV. Background $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$ events are rejected by photon veto requirements with average inefficiency of 1.3×10^{-8}, see figure 5.4.

A K^+ is tagged upstream if coincident signals are detected in at least five KTAG sectors within 2 ns of the π^+ track time. This is matched to a GTK track within 0.6 ns of the KTAG time and closest distance of approach (CDA) with the π^+ candidate of less than 4 mm. A discriminant based on temporal and spatial information is used for matching of

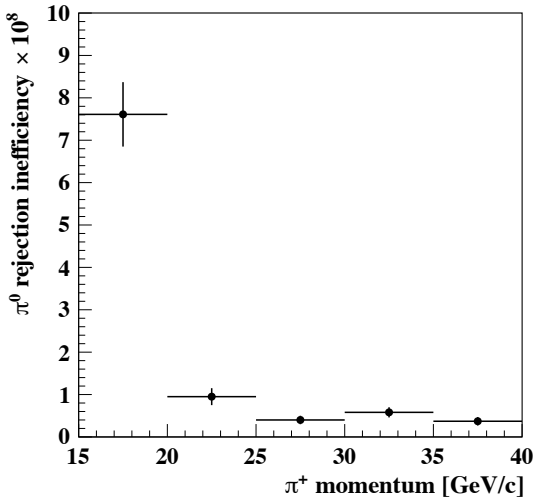


Figure 5.4: [125] π^0 rejection inefficiency as a function of π^+ momentum for $K^+ \to \pi^+\pi^0$ decays measured with 2017 data.

the K^+ and π^+. This has an efficiency of 75% and a 1.3% (3.5%) probability of incorrect associations when a K^+ is (is not) correctly reconstructed. Incorrect associations and misreconstruction can contribute to background mechanisms, specifically as discussed in section 5.3.3.

Upstream backgrounds, see section 5.3, arise from a combination of early decays (upstream of the start of the decay volume vacuum tank), beam particle interactions with GTK stations, pileup GTK tracks and high-angle π^+ scattering in STRAW1. To minimise these backgrounds the K^+–π^+ vertex must be formed inside the fiducial volume (FV, with $105 < z_{vtx} < 165\,\text{m}$) with no additional CHANTI activity within 3 ns of the π^+ track time. Additionally a 'box cut' is applied requiring that the projected position of the π^+ at the final collimator (COL) plane is outside a central $200 \times 1000\,\text{mm}^2$ area. This is effective to reject upstream backgrounds which primarily arise from π^+ produced upstream passing through the excluded 'box cut' spatial region at the final collimator. In contrast, the projected positions in this plane of π^+ from signal events are expected to be broadly distributed due to the decay kinematics.

5.1.4 Single Event Sensitivity

The single event sensitivity, $\mathcal{B}_{SES}$, is defined as the branching ratio corresponding to observation of one signal event. For the $K^+ \to \pi^+\nu\bar{\nu}$ analysis the $K^+ \to \pi^+\pi^0$ decay is used as a normalisation channel, with a sample selected with a procedure similar to the signal selection but without upstream background, photon veto and some multiplicity rejection requirements and with $0.01 < m^2_{miss} < 0.026\,\text{GeV}^2/c^4$. The $\mathcal{B}_{SES}$ is then defined by

$$\mathcal{B}_{SES} = \frac{1}{A_{\pi\nu\nu}} \frac{\mathcal{B}_{\pi\pi} A_{\pi\pi}}{N_{\pi\pi}} \frac{1}{D} \frac{\varepsilon^{CTRL}_{\pi\pi}}{\varepsilon^{PNN}_{\pi\nu\nu}} \varepsilon_{RV} \ , \tag{5.1.2}$$

where $A_{\pi\nu\nu}$ and $A_{\pi\pi}$ are the acceptances for selection of signal and normalisation events (including kinematic and selection acceptance, assuming the SM phase-space distribution

for the $K^+ \to \pi^+\nu\bar{\nu}$ decay, and corrected to account for detector and reconstruction efficiencies), $\mathcal{B}_{\pi\pi} = \mathcal{B}(K^+ \to \pi^+\pi^0)$, $N_{\pi\pi}$ is the number of normalisation events recorded using control (CTRL) trigger with efficiency $\varepsilon_{\pi\pi}^{CTRL}$ and downscaling factor $D = 400$ and signal events are collected with the PNN trigger with efficiency $\varepsilon_{\pi\nu\nu}^{PNN}$. The random veto efficiency term is $\varepsilon_{RV} = 1 - f_{RV}$, where f_{RV} is the fraction of events rejected by the selection veto conditions due to accidental activity. This term is added explicitly because pileup, which can lead to additional activity, is not simulated. Random veto effects are different for the signal and normalisation selections, the largest effect being because of the photon veto conditions applied in the signal selection only, however effects that cancel in the ratio $\frac{\varepsilon_{RV}^{\pi\pi}}{\varepsilon_{RV}^{\pi\nu\nu}}$ can be neglected. In general the terms of 5.1.2 depend on π^+ momentum, p_π, and instantaneous beam intensity, I. The analysis is performed in four momentum bins, each $5\,\mathrm{GeV}/c^2$ wide between 15 and 35 GeV and five instantaneous intensity bins and the total single event sensitivity is given by

$$\mathcal{B}_{SES} = \frac{D}{\mathcal{B}_{\pi\pi}} \sum_{p_\pi \text{ bins}} \sum_{I \text{ bins}} \left[\frac{1}{A_{\pi\nu\nu}(p_\pi, I)} \frac{A_{\pi\pi}(p_\pi, I)}{N_{\pi\pi}(p_\pi, I)} \frac{\varepsilon_{\pi\pi}^{CTRL}(p_\pi, I)}{\varepsilon_{\pi\nu\nu}^{PNN}(p_\pi, I)} \varepsilon_{RV}(p_\pi, I) \right] . \tag{5.1.3}$$

The intensity dependence of the signal and normalisation acceptance are expected to cancel and their dependence on momenta is shown in figure 5.5. The intensity and momentum dependence of the PNN trigger efficiency is shown in figure 5.6a while the minimum bias CTRL trigger has efficiency of over 99% and is approximately flat in p_π and I. The variation of random veto efficiency with intensity is shown in figure 5.6b.

The single event sensitivities obtained for the analysis of 2016 and 2017 data are

$$\mathcal{B}_{SES}^{16} = (3.15 \pm 0.01_{stat} \pm 0.24_{syst}) \times 10^{-10} , \tag{5.1.4}$$

$$\mathcal{B}_{SES}^{17} = (3.89 \pm 0.024_{syst}) \times 10^{-11} . \tag{5.1.5}$$

Simulated $K^+ \to \pi^+\nu\bar{\nu}$ events are generated assuming a SM description and therefore the number of expected $K^+ \to \pi^+\nu\bar{\nu}$ events is

$$N_{\pi\nu\nu,SM}^{exp} = \frac{\mathcal{B}_{SM}(K^+ \to \pi^+\nu\bar{\nu})}{\mathcal{B}_{SES}} . \tag{5.1.6}$$

Therefore, the expected numbers of events in the 2016 and 2017 data samples are [124] [125]

$$N_{\pi\nu\nu,SM}^{exp,16} = 0.267 \pm 0.001_{stat} \pm 0.020_{syst} \pm 0.032_{ext} , \tag{5.1.7}$$

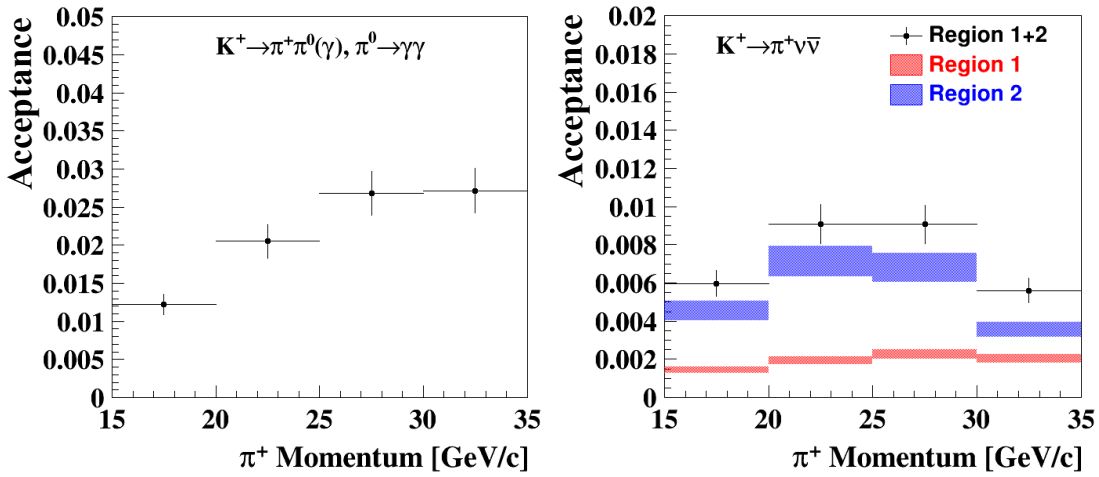
$$N_{\pi\nu\nu,SM}^{exp,17} = 2.16 \pm 0.13_{syst} \pm 0.26_{ext} , \tag{5.1.8}$$

with negligible statistical uncertainty (in 2017) and external uncertainty arising from the uncertainty on the SM branching ratio calculation (see equation 2.3.11).

5.2 Studies of the $K^+ \to \pi^+\pi^+\pi^-$ Background

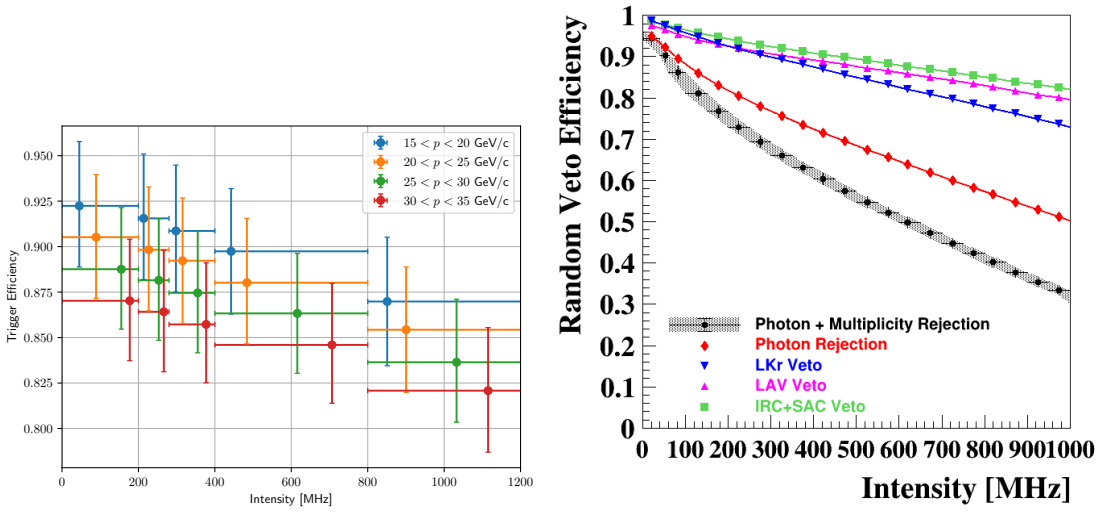
5.2.1 Preliminaries: The $K^+ \to \pi^+\pi^+\pi^-$ Background

To illustrate the required rejection of $K^+ \to \pi^+\pi^+\pi^-$ background events with a signal-like signature an order-of-magnitude consideration can be followed. If acceptance for



(a) (b)

Figure 5.5: ([125], edited.) Acceptance of normalisation (a) and signal (b) selections derived from simulations. The signal acceptance, and its associated uncertainty, for each signal region is shown individually by the coloured boxes.



(a) (b) [125]

Figure 5.6: (a) PNN trigger efficiency as a function of instantaneous beam intensity for each of the four momentum bins. Points are displaced horizontally in each intensity region for visibility and vertical bars represent the total uncertainty, with results from [125]. Trigger efficiency decreases with increasing intensity due to increased probability of additional in-time activity. Efficiency also decreases for higher momentum tracks because straw momentum reconstruction efficiency degrades for higher momenta for which resolution worsens (see section 3.2.5).
(b) [125] Random veto efficiency as a function of instantaneous intensity showing results after individual components of the veto selection and the combined result (with total uncertainty indicated).

$K^+ \to \pi^+\nu\bar{\nu}$ is of order 10% and $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu}) \sim \mathcal{O}(10^{-10})$, to get a 1% background contribution $K^+ \to \pi^+\pi^+\pi^-$ events must be rejected at the level of at least 2×10^{-12}, given that $\mathcal{B}(K^+ \to \pi^+\pi^+\pi^-) \approx 5\%$ (see table 2.2). This is achieved by kinematic and multiplicity rejection.

The upper limit of signal region 2 is defined to be below the kinematic endpoint of the $K^+ \to \pi^+\pi^+\pi^-$ decay, specifically to exclude it kinematically. However, in principle misreconstruction of event kinematics or misassociation of reconstructed tracks could produce a signal-like signature with a non-Gaussian 'kinematic tail' in the m^2_{miss} distribution which extends into signal region 2.

Remaining rejection is achieved through the multiplicity rejection cuts in the selection, vetoing events with excess associated activity in time or space in the detector with a candidate signal-like event. The required sensitivity for this analysis is such that rare exclusive pathological situations and phenomena must be considered including decays, hadronic interactions, inefficiencies and some compound effect of a combination of these processes. Only scenarios with probability lower than 10^{-13} may be considered negligible.

In the following a study is presented which determines the kinematic rejection of the $K^+ \to \pi^+\pi^+\pi^-$ background. This is achieved by determining the 'kinematic tail fraction', the fraction of $K^+ \to \pi^+\pi^+\pi^-$ decays which pass the $K^+ \to \pi^+\nu\bar{\nu}$ selection and enter signal region 2. The methodology of calculating background estimates using this information is described in section 5.2.2. This study was performed first in the context of the NA62 analysis of 2016 data and was then used in the estimation of the $K^+ \to \pi^+\pi^+\pi^-$ background for both this and the analysis of 2017 data.

5.2.2 Background Estimation from Kinematic Tail Fraction

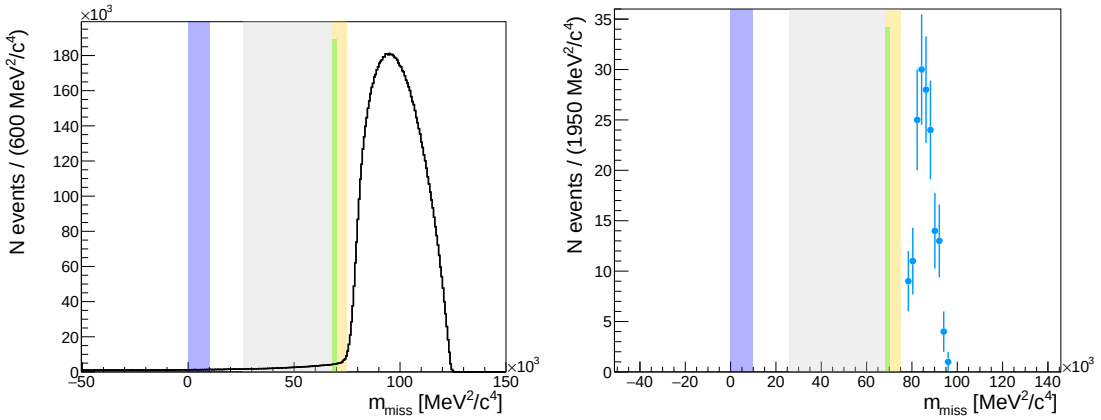
Non-Gaussian tails in the m^2_{miss} spectrum arise due to kinematic misreconstruction. To calculate the expected number of background events in the signal region the kinematic tail fraction, f_{kin}, can be established from a control sample. In figure 5.7a an example exaggerated case is shown of a control sample of $K^+ \to \pi^+\pi^+\pi^-$ decays with a non-Gaussian kinematic tail extending into the signal regions, here signal region 2 will be considered. From this control sample

$$f_{kin} = \frac{N(R2)}{N(3\pi R)}\,, \tag{5.2.1}$$

where $N(R2)$ and $N(3\pi R)$ are the number of events in signal region 2 and the 3π region (see figure 5.1), respectively. Then, when the full signal selection is applied to data the number of events in the 3π region can be determined, figure 5.7b shows the selected $n^{obs}_{3\pi R} = 159$ data events in this region in 2017 data. The predicted number of events in signal region 2 is then

$$n^{exp}_{R2} = f_{kin} n^{obs}_{3\pi R}\,. \tag{5.2.2}$$

For this example, for signal region 2 if $f_{kin} < 1.4 \times 10^{-4}$ then the background from $K^+ \to \pi^+\pi^+\pi^-$ decays is equivalent to less than 1% of the expected signal in 2017 data. This conclusion also holds for 2016 data.



(a) (b)

Figure 5.7: Illustration of background estimation from kinematic tail fraction. (a) An exaggerated view of a kinematic non-Gaussian tail from a selected control sample of $K^+ \to \pi^+\pi^+\pi^-$ decays (from an early stage of the selection). The kinematic tail fraction is determined from the final control sample selected to be representative of, but orthogonal to, the selected signal sample. (b) Events passing the $K^+ \to \pi^+\nu\bar{\nu}$ signal selection in the 3π region (see figure 5.1). Multiplying the number of events shown in (b) by the kinematic tail fraction then gives the background expectation in, for example signal region 2. Following the colour code of results in this subsection signal regions 1 and 2 are shown by the blue and grey shaded regions respectively with control regions, defined in section 5.2.6, shown in orange and cyan.

5.2.3 The $K^+ \to \pi^+\pi^+\pi^-$ Background Mechanism

To pass the signal $K^+ \to \pi^+\nu\bar{\nu}$ selection a $\pi^+\pi^-$ pair from a $K^+ \to \pi^+\pi^+\pi^-$ decay must go undetected while a lone π^+ (henceforth referred to as π_3) is detected and matched to a track upstream. It is expected that the probability for this scenario is low, however this must be quantified. The expected momentum spectrum of π_3 candidates from simulations is shown in figure 5.8.

An example background mechanism scenario is a $K^+ \to \pi^+\pi^+\pi^-$ decay event with a $15\,\mathrm{GeV}/c$ π^+ inside the geometric acceptance of all relevant downstream detectors, an undetected $50\,\mathrm{GeV}/c$ π^+ which travels along the beamline and misses downstream detectors, with a $10\,\mathrm{GeV}/c$ π^- which decays in flight to a $\sim 6\,\mathrm{GeV}/c$ μ^- which is swept out of acceptance of downstream detectors by the STRAW dipole magnet (M) and is not detected. Only the first π^+ is detected and, if matched to an upstream track, the event could have signal-like properties and be selected, meaning a background event enters the signal region. A second example scenario has a $15\,\mathrm{GeV}/c$ π^+, which again mimics a signal π^+, the other π^+ has $35\,\mathrm{GeV}/c$ and interacts hadronically with the RICH entrance flange, while the $25\,\mathrm{GeV}/c$ π^- is not reconstructed due to detector inefficiency (known to be $\mathcal{O}(1\%)$ in STRAW).

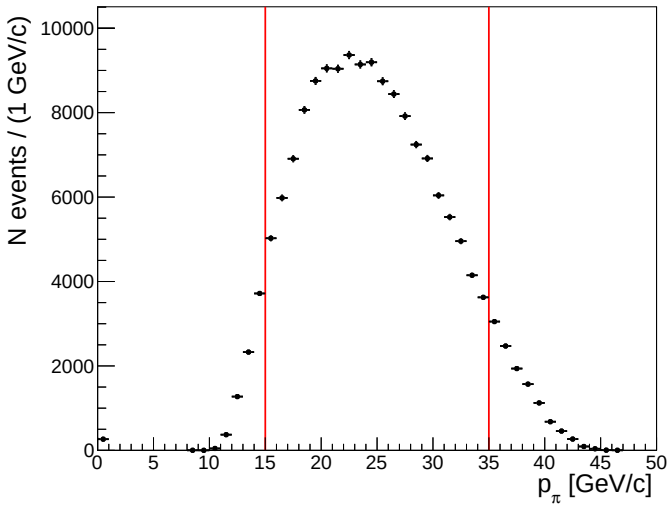


Figure 5.8: Momentum spectrum of π^+ from a control sample of simulated $K^+ \to \pi^+\pi^+\pi^-$ decays. The 15–35 GeV/c momentum range allowed in the $K^+ \to \pi^+\nu\bar{\nu}$ signal selection is between the vertical red lines.

5.2.4 Background Study Procedure

The selection of potential background events proceeds in two stages: first, a representative sample of $K^+ \to \pi^+\pi^+\pi^-$ events must be isolated, then a signal-like ($\pi\nu\bar{\nu}$-like) selection must be applied to a candidate K^+–π_3 pair. However, in selecting $K^+ \to \pi^+\pi^+\pi^-$ events kinematic information about the candidate π_3 cannot be used since this may bias the subsequent $\pi\nu\bar{\nu}$-like selection run for the candidate K^+–π_3 pair. Therefore a selection was developed to select $K^+ \to \pi^+\pi^+\pi^-$ events using only kinematic information from a $\pi^+\pi^-$ pair which will be referred to as '2π tagging'. The selection applied to the candidate K^+–π_3 pair is described as $\pi\nu\bar{\nu}$-like because some parts of the full selection procedure are not possible in this case. However, carefully applying the relevant selection cuts from the full selection allows the $\pi\nu\bar{\nu}$-like selection to preserve all relevant properties of the full signal selection. The primary analysis procedure, as outlined here, is detailed in section 5.2.5. Results were validated using similar but modified analysis procedures as described in appendix B.1.

5.2.5 Details of The Primary Analysis Procedure

Events passing the control trigger (for data) with at least one two-track vertex reconstructed are considered. For correctly reconstructed three-track $K^+ \to \pi^+\pi^+\pi^-$ events four two-track vertices should be formed and each is considered in turn.

5.2.5.1 2π Tagging: Selection of $K^+ \to \pi^+\pi^+\pi^-$ Control Sample

A control sample of $K^+ \to \pi^+\pi^+\pi^-$ events is identified by the following, using only the kinematic information from a $\pi^+\pi^-$ forming a two-track vertex. A cut-based procedure is applied with selection requirements described below.

1. Tracks forming the two-track vertex have charges +1 and −1 and total vertex charge is 0.
2. The two-track vertex has good quality with $\chi^2_{vtx} < 40$.
3. Two-track vertex z position lies in range $115 < z_{vtx} < 165\,\text{m}$.
4. Both vertex tracks must lie inside the geometric acceptance of all four STRAW chambers, the CHODs, LKr and MUV3.
5. Both vertex tracks are reconstructed from signals in all four STRAW chambers.
6. $E/p < 0.9$, where E/p is the ratio of energy deposited in the LKr to the STRAW track momentum. This rejects $e^\pm$ which have $E/p \approx 1$.
7. Events are rejected if either vertex track has an associated signal in the MUV3. This rejects muons from other decays or where $\pi^\pm \to \mu^\pm\nu_\mu$ decays in flight occur (with probability of 7.5%, see appendix H).
8. At least one K^+ is identified with coincident signals in at least 5 sectors, with associated timing information (t_{KTAG}) used later.
9. The two vertex tracks each have associated signals in the CHODs.
10. A track time, $t_{trk,i}$, is defined for each track i using the mean of the CHOD and NA48-CHOD time measurements weighted by the measured time resolutions (see appendix E). The smallest difference between this and the time associated to a tagged K^+ candidate in the KTAG must be less than 10 ns.
11. Photon veto conditions are applied rejecting events with LAV, SAC or IRC hits within 3 ns, 3 ns and 4 ns respectively of the KTAG K^+ reference time best matching (in time) the pair of downstream tracks.

The squared missing mass is then calculated as

$$m^2_{miss,2\pi} = ((P_{\pi^+} + P_{\pi^-}) - P_{K^+})^2\,, \tag{5.2.3}$$

where P_{π^+},P_{π^-} and P_{K^+} are the 4-momenta of the π^+ and π^- tracks identified and forming the two-track vertex and the K^+. If $m^2_{miss,2\pi} \approx m^2_{\pi^+}$ this suggests a $K^+ \to \pi^+\pi^+\pi^-$ event has been selected only using the $\pi^+\pi^-$ pair. The third track is then the π_3 candidate. A Gaussian fit to the peak in $m^2_{miss,2\pi}$ was performed with mean consistent with $m^2_{\pi^+}$ and standard deviation $\sigma = 790\,\text{MeV}^2/c^4$. Events with $m_{\pi^+} - 3\sigma < m^2_{miss,2\pi} < m_{\pi^+} + 3\sigma$ are selected, as shown by figure 5.9.

5.2.5.2 $K^+ \to \pi^+\nu\bar{\nu}$-like Selection For K^+–π_3 Pairs

The π_3 candidate track must have:

1. Charge +1.

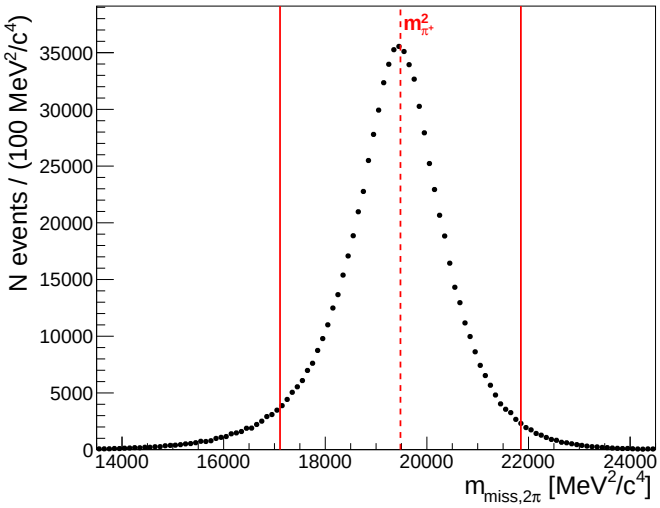


Figure 5.9: Distribution of $m^2_{miss,2\pi}$, defined by equation 5.2.3, after 2π tagging selection requirements 1–11 for simulated $K^+ \to \pi^+\pi^+\pi^-$ decays. Events between the solid red lines are selected and pass the full 2π tagging procedure.

2. Good quality: $\chi^2_{trk} < 20$.
3. Signals associated to it in all four STRAW chambers.
4. A position inside the geometric acceptance of all four STRAW chambers, CHODs, LKr and MUV3 and outside LAV(12) and IRC active area.
5. An associated cluster in the LKr.
6. Associated signals in the CHODs.
7. No associated signal in the MUV3.
8. Momentum in range 15–35 GeV/c.
9. Impact point at LKr more than 20 mm from any dead cells.
10. Associated signals in CHOD with a time within 5 ns of reference KTAG K^+ time.

K–π matching is then performed (using a standard matching tool in the NA62 software framework designed by the author) with requirements that: at least one successful match is made with best matched pair having CDA less than 7 mm and with K–π vertex in region $115 < z < 165$ m, see figure 5.10.

The π_3 candidate track is then compared to the projected properties of a π^+ using information from the $\pi^+\pi^-$ pair. The predicted 3-momentum of the π_3 is calculated according to

$$p^{BM}_{\pi_3} = p_{K^+} - (p^{BM}_{\pi^+} + p^{BM}_{\pi^-}), \tag{5.2.4}$$

where p_{K^+} is the K^+ beam particle 3-momentum and $p^{BM}_{\pi^+}$ and $p^{BM}_{\pi^-}$ are the 3-momenta of the $\pi^+\pi^-$ track pair before the straw spectrometer magnet. Using this the expected position of π_3 is then predicted at the NA48-CHOD and CHOD detectors and compared

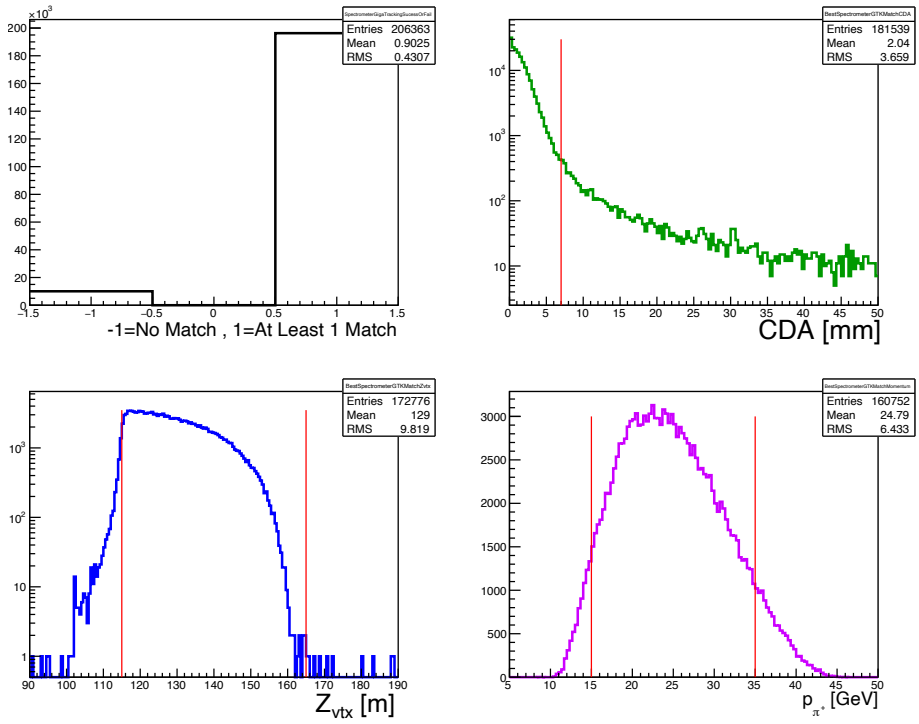


Figure 5.10: Illustrations of selection criteria, shown by red vertical lines, in the relevant variable: CDA (top right), z_{vtx} (bottom left) and p_{π_3} (bottom right). The number of candidate π_3 tracks matched to a K^+ candidate upstream is shown in the top left histogram.

with the position information recorded by the CHODs for the selected candidate π_3 track. The difference between predicted and observed position, ΔR, must be less than 100 mm, see figure 5.11. The π_3 position is first extrapolated to $z_{BTT} = 183.331$, accounting for effects of charge and magnetic field in the vacuum tank using a standard NA62 tool. Beyond this point, the maximum allowed by the tool, a linear extrapolation is performed using the final position and momenta ($p_{\pi_3}^{BMaBTT}$) given by the full extrapolation at $z_{BTT} = 183.331$. Following a linear extrapolation from $z_L = z_{BTT}$ to the spectrometer magnet (M) at $z_m = 196.995$ m, the momentum kick is accounted for, followed by a further linear extrapolation to the position of the CHOD or NA48-CHOD $z_C = 238.130$ m and $z_N = 239.029$ m respectively. The final positions at the CHOD and NA48-CHOD are therefore given by vectors

$$C = \begin{pmatrix} x_L - \left[(z_C - z_L)\frac{p_x}{p_y} - \frac{270}{p_z}(z_C - z_M)\right] \\ y_L + (z_C - z_L)\frac{p_y}{p_z} \\ z_C \end{pmatrix}, \quad N = \begin{pmatrix} x_L - \left[(z_N - z_L)\frac{p_x}{p_y} - \frac{270}{p_z}(z_N - z_M)\right] \\ y_L + (z_N - z_L)\frac{p_y}{p_z} \\ z_N \end{pmatrix}, \tag{5.2.5}$$

where the momenta have units of MeV/c, $p = p_{\pi_3}^{BMaBTT}$ and spatial coordinates have units of metres, see also figure 5.12.

If all above conditions are satisfied then the squared missing mass is calculated for the K^+–π_3 pair as

$$m_{miss}^2 = (P_{K^+} - P_{\pi_3})^2 . \tag{5.2.6}$$

This is the proxy for the $K^+ \to \pi^+\nu\bar{\nu}$ squared missing mass and any events found to be

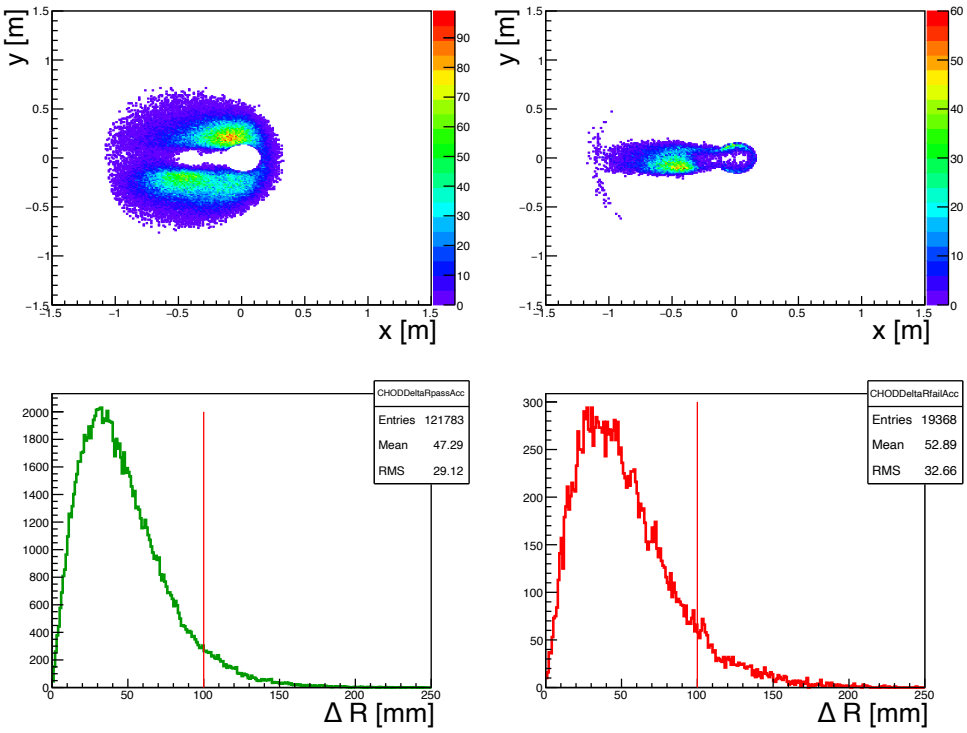


Figure 5.11: Predicted π_3 position at the CHOD (top) and difference between predicted and actual position, ΔR, (bottom) for tracks passing all geometric acceptance criteria (left) or not (right).

in signal region 2 represent the background signature of interest for this study.

5.2.6 Results

For this analysis the 2016A data-set is used. This data-set was collected in a 4-week period during September and October 2016 [216] and is the data analysed for the first NA62 $K^+ \to \pi^+\nu\bar{\nu}$ results [124].

The primary results of the study are the 'tail fraction' of events which lie in regions of interest, specifically:

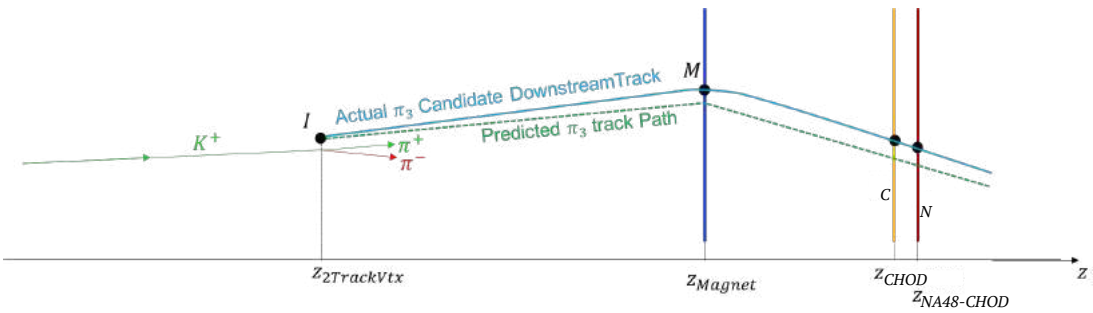


Figure 5.12: Sketch of the procedure to predict the position of a candidate π_3 track at the CHOD (C) and NA48-CHOD (N) planes.

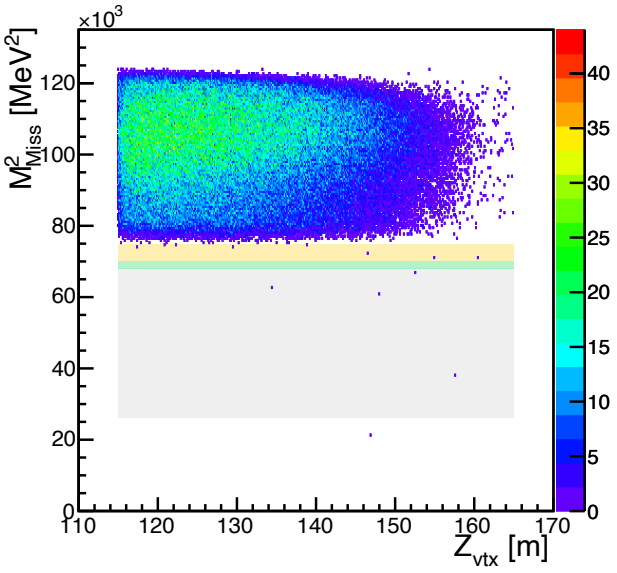


Figure 5.13: K^+–π_3 squared missing mass versus vertex z position two dimensional histogram for control sample of $K^+ \to \pi^+\pi^+\pi^-$ isolated in 2016 data. The grey, orange and cyan regions illustrate respectively signal region 2, the full control region and restricted control region.

- Signal Region 2 (RII): $26 \times 10^3 < m_{miss} < 68 \times 10^3\,\mathrm{MeV}^2/c^4$.
- $K^+ \to \pi^+\pi^+\pi^-$ control region (CR): $68 \times 10^3 < m_{miss} < 75 \times 10^3\,\mathrm{MeV}^2/c^4$.
- Restricted control region (RCR): $68 \times 10^3 < m_{miss} < 70 \times 10^3\,\mathrm{MeV}^2/c^4$.

This is compared to the 3π region above $75 \times 10^3\,\mathrm{MeV}^2/c^4$. The restricted control region was studied specifically for the 2016 analysis in which the region $70 \times 10^3 < m^2_{miss} < 75 \times 10^3\,\mathrm{MeV}^2/c^4$ was unblinded separately during background studies.

The analysis was performed separately in four momentum regions of width $5\,\mathrm{GeV}/c$ covering the standard $K^+ \to \pi^+\nu\bar{\nu}$ pion momentum range 15–35 GeV/c. No significant dependence on momentum for the tail fraction is observed, however the shape of the m^2_{miss} distribution in the 3π background region is found to be different. Because of a known difference in the distribution of $K^+ \to \pi^+\pi^+\pi^-$ as a function of the z position of the decay, the analysis was also performed independently in five regions of vertex position for the K^+–π_3 pair. These regions are 10 m long and cover the full z_{vtx} range 115–265 m. The two-dimensional histogram of m^2_{miss} versus z_{vtx} for the control sample in data is shown in figure 5.13. Dependency of the tail fraction on z_{vtx} could indicate some bias since a similar relationship is not expected in principle for the signal $K^+ \to \pi^+\nu\bar{\nu}$. Figures 5.14 and 5.15 show the squared missing mass distributions for three key control samples (data, MC simulations and $\pi\nu\bar{\nu}$-like MC, see appendix B.1) in the four momentum and five z_{vtx} regions defined above.

The final tail fraction results are shown in table 5.1, broken down separately into z_{vtx} and momentum bins and displaying the overall average results.

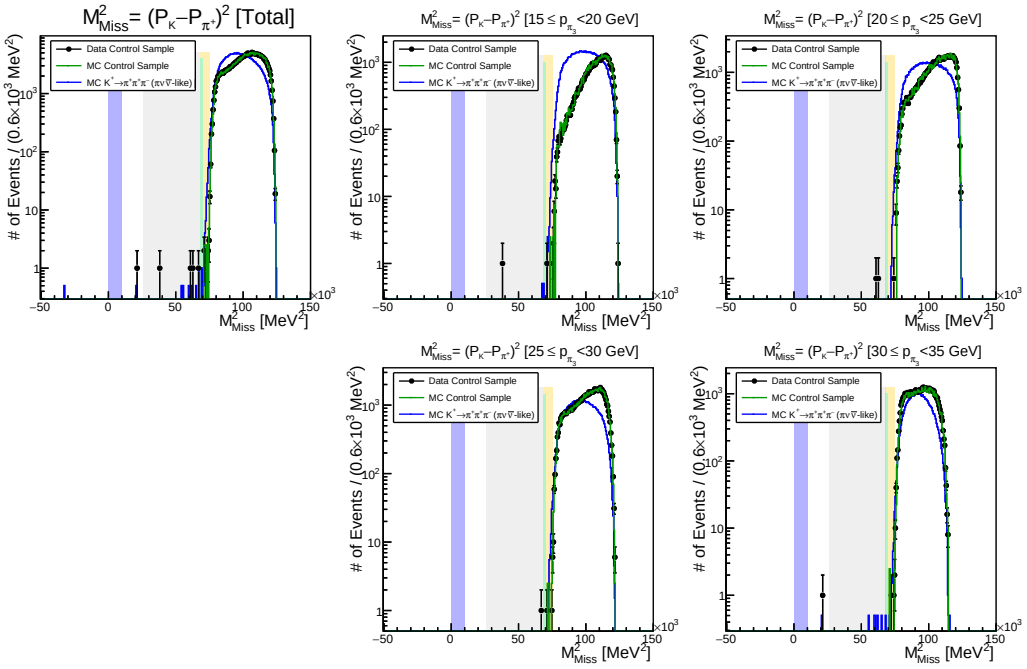


Figure 5.14: The K^+–π_3 pair m^2_{miss} distributions obtained from the three key analysis procedures (1,2 and 5) described in section B.1 for the four momentum regions and the total (full 15–35 GeV/c range). The grey, orange and cyan regions illustrate respectively signal region 2, the full control region and restricted control region.

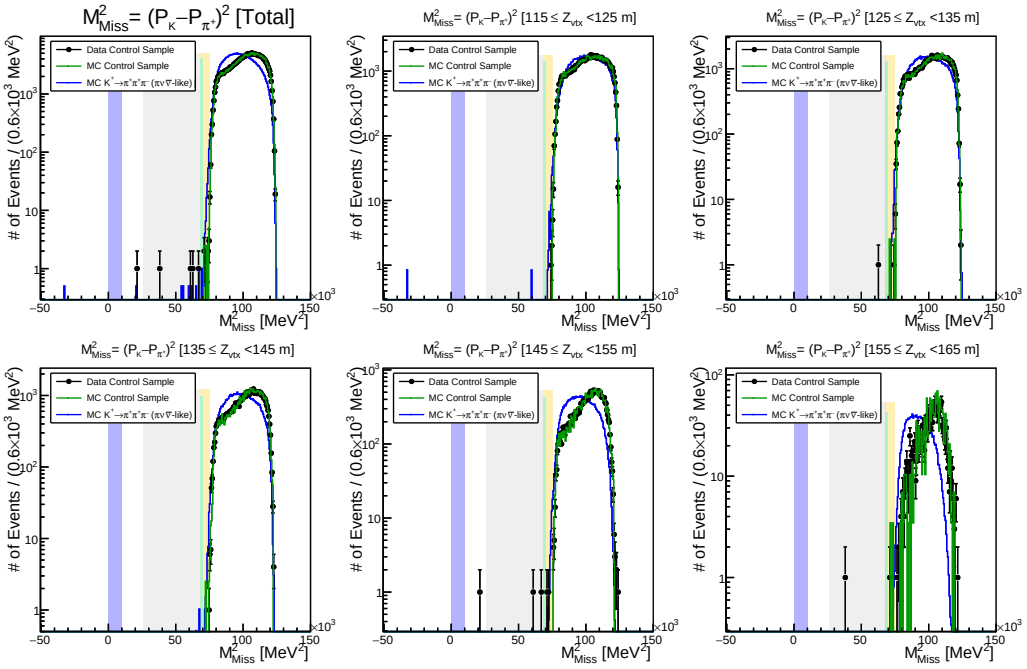


Figure 5.15: The K^+–π_3 pair m^2_{miss} distributions obtained from the three key analysis procedures (1,2 and 5) described in section B.1 for the five z_{vtx} regions and the total (full 115–165 m range). The grey, orange and cyan regions illustrate respectively signal region 2, the full control region and restricted control region.

Table 5.1: Final results for fraction of events in regions of interest for analysis of 2016 data. The uncertainties displayed represent 90% confidence intervals considering potential statistical fluctuations.

	Signal Region 2 (RII) [$\times 10^{-5}$]	3π Control Region (CR) [$\times 10^{-5}$]	Restricted Control Region (RCR) [$\times 10^{-5}$]
$115 < Z_{vtx} < 125$ m	$0^{+3.2}_{-0}$	$8.6^{+6.9}_{-4.3}$	$0^{+3.2}_{-0}$
$125 < Z_{vtx} < 135$ m	$1.2^{+4.6}_{-1.2}$	$8.5^{+7.5}_{-4.5}$	$0^{+3.6}_{-0}$
$135 < Z_{vtx} < 145$ m	$0^{+5.5}_{-0}$	13^{+11}_{-7}	$0^{+5.4}_{-0}$
$145 < Z_{vtx} < 155$ m	10^{+20}_{-8}	10^{+20}_{-8}	0^{+14}_{-0}
$155 < Z_{vtx} < 165$ m	60^{+230}_{-60}	40^{+230}_{-60}	0^{+180}_{-0}
$15 < p_{\pi_3} < 20$ GeV/c	$2.3^{+8.5}_{-2.1}$	$4.5^{+9.7}_{-3.7}$	$0^{+6.8}_{-0}$
$20 < p_{\pi_3} < 25$ GeV/c	$2.6^{+5.6}_{-2.1}$	$1.3^{+4.9}_{-1.2}$	$0^{+3.9}_{-0}$
$25 < p_{\pi_3} < 30$ GeV/c	$1.3^{+4.8}_{-1.2}$	$1.03^{+0.83}_{-0.52}$	$0^{+3.9}_{-0}$
$30 < p_{\pi_3} < 35$ GeV/c	$0^{+5.5}_{-0}$	26^{+15}_{-10}	$0^{+5.5}_{-0}$
Total	$1.6^{+2.0}_{-1.0}$	$9.9^{+4.0}_{-3.0}$	$0^{+1.2}_{-0}$

The analysis concept is non-standard and therefore it is important to perform checks to ascertain its validity. Clear potential for bias arises from the selection of the control sample of $K^+ \to \pi^+\pi^+\pi^-$ events including, despite remaining independent of the candidate π_3 kinematics, the 2π tagging procedure. Therefore four alternative analysis strategies were followed, including cases without 2π tagging with details given in appendix B.1, and results compared to identify any bias. The summary of the results, averaged across all z_{vtx} and momentum bins, is given in table 5.2. Differences in results are observed between procedures with and without 2π tagging, especially in the control region (CR). This difference is due to changes in the m^2_{miss} spectrum when applying a requirement that at least one two-track vertex is formed, the first selection condition in the 2π tagging procedure, see also appendix B.2. However, the difference between results in signal region 2 (RII) is less significant and the tail fraction values determined with and without 2π tagging are consistent considering their uncertainty.

Table 5.2: Tail fraction results summary for RII, CR and RCR using each of the five analysis strategies described in appendix B.1. The uncertainties displayed represent 90% confidence intervals considering potential statistical fluctuations.

Analysis	RII [$\times 10^{-5}$]	CR [$\times 10^{-5}$]	RCR [$\times 10^{-5}$]
1. Data	$1.6^{+2.0}_{-1.0}$	$9.9^{+4.0}_{-3.0}$	$0^{+1.2}_{-0}$
2. $K_{3\pi}$ MC	$0^{+3.0}_{-0}$	$13.8^{+7.8+}_{-5.5}$	$0^{+3.0}_{-0}$
3. $K_{3\pi}$ MC+!2πTag	$2.47^{+0.89}_{-0.75}$	$151.1^{+6.5}_{-6.3}$	$1.09^{+0.71}_{-0.48}$
4. MC+!2πTag+AL1π^+	$2.33^{+0.97}_{-0.74}$	$151.7^{+6.6}_{-6.4}$	$1.01^{+0.71}_{-0.46}$
5. MC+!2πTag+E1π^+	$2.0^{+1.4}_{-0.9}$	$153.4^{+9.4}_{-9.0}$	$1.2^{+1.6}_{-0.7}$

5.2.7 Implication of Results for $K^+ \to \pi^+\nu\bar{\nu}$

Good agreement between analysis of data and simulations is achieved, demonstrating validity of the description of the tail over at least four orders of magnitude, see figures 5.14

and 5.15. Differences between the simulation studies with and without applying 2π tagging is understood to be due to the effect of the requirement for at least one two-track vertex in the event. The consistency of the tail fraction, without dramatic variation, in bins of z_{vtx}, shown in figure B.5, for each analysis procedure, suggests a lack of bias and reinforces the validity of the results presented.

For the main $K^+ \to \pi^+\nu\bar{\nu}$ analysis the final background estimates for the 2016 and 2017 data analyses were based on the results of this study. From full consideration of the results, and accounting for uncertainties arising from the study of a limited sample (in both data and simulation) of $K^+ \to \pi^+\pi^+\pi^-$ kinematic tail events, a conservative estimation for the kinematic tail fraction of $f_{3\pi}^{kin} = (5 \pm 5) \times 10^{-5}$ was adopted, with a 100% uncertainty [124] [125]. The conclusion was that the kinematic and multiplicity rejection combined achieve suppression of the $K^+ \to \pi^+\pi^+\pi^-$ background to an effectively negligible level (compared to the signal and largest backgrounds, see sections 5.4 and 5.6). Based on observation of 20 and 159 events passing the signal selection in the 3π region (see figures 5.1 and 5.7b), the background predictions for the $K^+ \to \pi^+\pi^+\pi^-$ background in the 2016[1] and 2017 data analyses are

$$N_{3\pi}^{exp.\,16} = 0.002 \pm 0.001_{stat} \pm 0.002_{syst}\,, \tag{5.2.7}$$

$$N_{3\pi}^{exp.\,17} = 0.008 \pm 0.008\,. \tag{5.2.8}$$

The procedure is validated in the analysis of 2017 data by comparing estimation of background in a 3π control m^2_{miss} region, (68–72) $\times 10^3\,\mathrm{MeV}^2/c^4$, of < 0.24 events with the observation of 0 events.

5.3 Studies of the Upstream Background

5.3.1 Upstream Background Mechanisms

Upstream backgrounds arise from a combination of early decays (upstream of the decay volume), beam particle interactions with the GTK stations, pileup GTK tracks and high-angle π^+ scattering in the STRAW1. The resulting background mechanisms have low probabilities, however when compared to the requirement for a rejection factor of $\mathcal{O}(10^{11})$ they become significant. Five specific mechanisms have been identified:

Type 1: (Known as 'snakes'.) A beam K^+ decays (e.g. as $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$) in the upstream region between GTK2 and GTK3. The daughter π^+ produced has momentum less than the beam average and is therefore deflected differently to beam particles in the final two GTK spectrometer (BEND 5a and 6) magnets. The π^+ then passes through

[1] For the 2016 analysis with 20 observed events in the 3π region and $f_{3\pi}^{kin} = (5 \pm 5) \times 10^{-5}$ a background expectation of $N_{3\pi}^{exp.\,16} = 0.001 \pm 0.001_{stat}$ is derived. However, since this is negligible with respect to the signal expectation (equation 5.1.7) and the result based on limited statistics with limited cross-checks available an even more conservative approach was adopted, doubling the predicted number of events and adding a 100% systematic uncertainty on the result. This gives the final result shown by equation 5.2.7. With additional cross-checks possible in 2017 with the larger data-set this further step was not required.

the central part of the final collimator and yoke of the BEND 6 magnet and enters the vacuum tank. This π^+ may undergo (relatively) large-angle scattering in the STRAW1. Simultaneously, a pileup track must be present in the GTK (most likely a π^+ beam particle) to which the downstream π^+ discussed above can be matched, forming a fake vertex. With scattering in the STRAW1 the position of the fake vertex and the projected position of the π^+ at the upstream reference plane (see discussion below) can pass signal selection criteria. This mechanism is illustrated in figure 5.16.

Type 2: (Known as 'mambas'.) This case is similar to the type 1/snakes scenario except that the deviated path of the daughter π^+ due to the GTK spectrometer magnets is more dramatic and so it passes through through holes in the yoke of the final magnet (BEND 6) and enters the vacuum tank. Again, a pileup GTK track is required upstream and a large angle scattering for the π^+ downstream in STRAW1. This mechanism is illustrated in figure 5.17.

Type 3: An inelastic interaction of a beam particle with a GTK station (GTK2 or GTK3) produces a π^+ which enters the vacuum tank and is detected downstream. The π^+ undergoes scattering in STRAW1 and is matched to a pileup GTK track upstream.

Type 4: An interaction of a K^+ beam particle with a GTK station (GTK2, or most importantly, GTK3) produces a prompt π^+ which is detected downstream and matched to the GTK track associated with the parent K^+. Coupled with π^+ scattering in STRAW1 this pushes the reconstructed vertex downstream and into the range selected and as such leads to a signature indistinguishable from the signal.

Type 5: In interactions of beam particles with upstream material (especially GTK3) a long-lived unstable particle (K^+, K_L^0, K_S^0) is produced which subsequently decays producing a π^+. Depending on the decay kinematics and the detection of the beam particle upstream, this can immediately produce a signal-like signature.

A few clarifications can be made in addition to the above:

- $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$ decays were explicitly considered, in principle other decays such as $K^+ \to \mu^+\nu_\mu$ or $K^+ \to \pi^+\pi^+\pi^-$ in the upstream region can also lead to background mechanisms as above with the addition of misidentification or inefficient multiplicity rejection. Nevertheless in practice the efficiency of PID and multiplicity rejection is such that these are held to a negligible level. Therefore, $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$ decays will be discussed exclusively here but studies of other possibilities were considered implicitly and explicitly for the analysis of 2016 and 2017 data respectively [124] [125].

- For $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$ decays to enter as a background the two photons must be lost. Depending on the position of the decay (for mechanism type 1) the photons may or may not hit upstream material. In the case of decays after the end of the final collimator photons will almost certainly enter the decay volume, and for such events the photon rejection cuts will be effective and reduce the background to negligible levels.

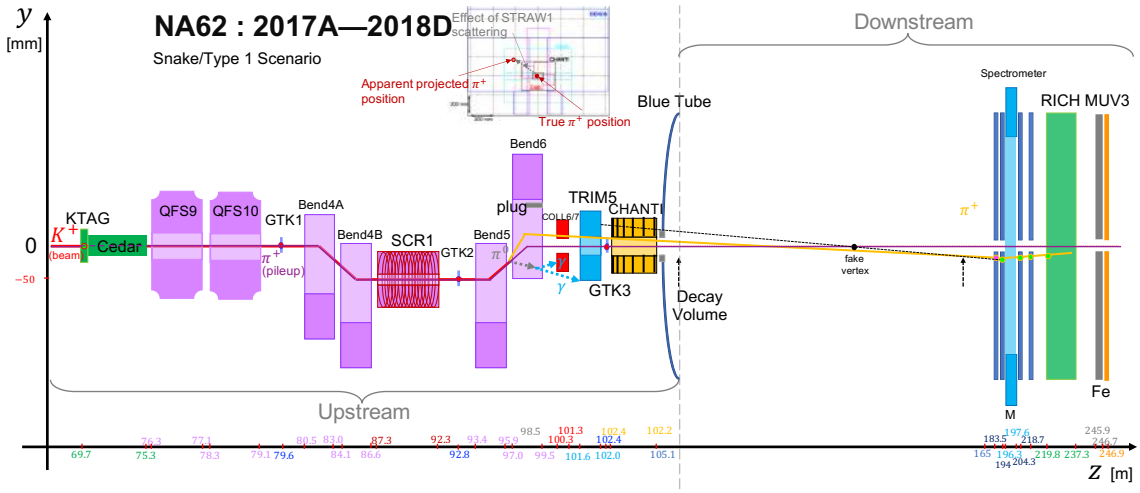


Figure 5.16: A schematic diagram of NA62 is presented with particular focus on the upstream region in which all components are drawn approximately in proportion to their size. The dipole bending magnet, labelled as B in figure 3.3, is given its CERN component name 'TRIM5' to avoid confusion with GTK spectrometer BEND magnets. An event is sketched with properties corresponding to the snake/type 1 upstream background mechanism. A K^+ decays between GTK2 and GTK3 and is bent upwards more than the standard beam path (purple line) but still passes through the central region of the collimator (see inset of (x, y) plane at final collimator) and enters the decay volume. A pileup track is also present upstream (following purple line) producing signals in the GTK out of time from those from the K^+ detected in the KTAG and its daughter π^+ detected by the RICH (which are in time with each other). Scattering in STRAW1 occurs and a fake vertex is created in the decay volume with projection outside the box cut region.

- In terms of the way the mechanism presents itself, mechanism type 3 with GTK2 interactions can be considered equivalent to type 1 but with the production of the π^+ specifically at GTK2 instead of at any point between GTK2 and GTK3.
- Mechanisms 1–4 require relatively large angle scattering in the STRAW1 to pass the full signal selection (due to the box cut discussed below). Scattering in STRAW2, 3 or 4 are not possible mechanisms because in these cases the change in direction can be identified, and such events would be rejected. Whereas, for scattering in STRAW1 the original direction of the π^+ track is always lost.
- Events without a matched pileup track or scattering in STRAW1 are still relevant to the discussion below because such events enter the control samples used to study the upstream background.

To reject the upstream background two selection components are crucial: the K^+–π^+ matching and the box cut. The box cut rejects events where the position of the π^+ projected back onto an upstream reference plane at the exit of the final collimator (Col6/7, labelled COL in figure 3.3) lies within a central region $|x| < 100\,\text{mm}$ and $|y| < 500\,\text{mm}$. This plane, along with the outlines of the relevant upstream material projected into this plane are shown in figure 5.19. The effect of decays or, as illustrated in figure 5.19, scattering in STRAW1 is to shift the projected position of the π^+ from the 'true' location through which the beam particle passed. The shift associated to the kinematics of $K^+ \to \pi^+\nu\bar{\nu}$ decays is expected, in general, to lead to a projected π^+ position which is distributed broadly across the upstream projection plane and so often (with approximately 60%

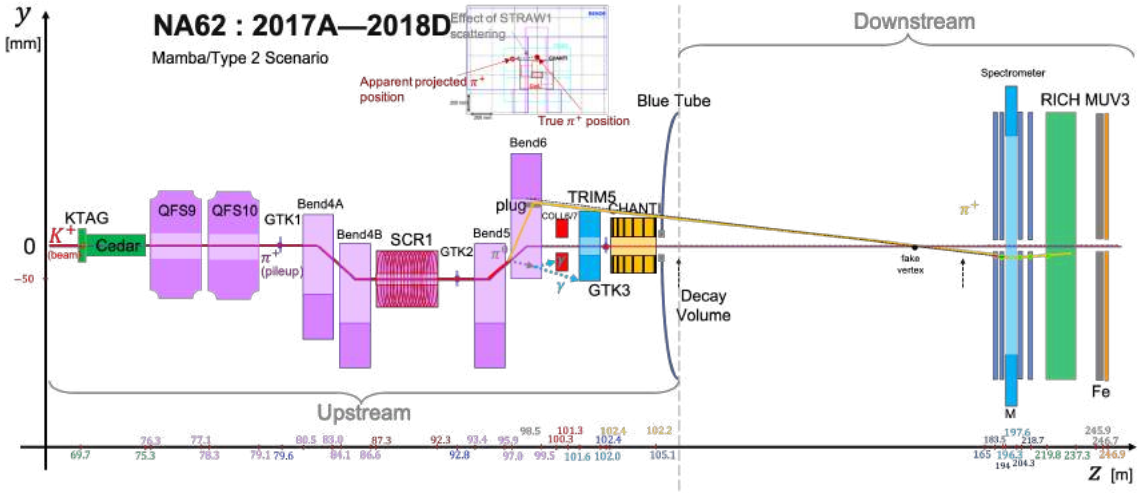


Figure 5.17: A schematic diagram of NA62 is presented with particular focus on the upstream region in which all components are drawn approximately in proportion to their size. The dipole bending magnet, labelled as B in figure 3.3, is given its CERN component name 'TRIM5'. An event is sketched with properties corresponding to the mamba/type 2 upstream background mechanism. A K^+ decays between GTK2 and GTK3 and is deflected upwards and passes through the hole in the final GTK spectrometer BEND 6 magnet yoke (see inset of (x, y) plane at final collimator) and enters the decay volume. During 2017 (before the 2017A sample) a plug was added to block exactly these π^+. A pileup track is also present upstream (following purple line) producing signals in the GTK out of time from those from the K^+ detected in the KTAG and its daughter π^+ detected by the RICH (which are in time with each other). Scattering in STRAW1 occurs and a fake vertex is created in the decay volume with projection outside the box cut region.

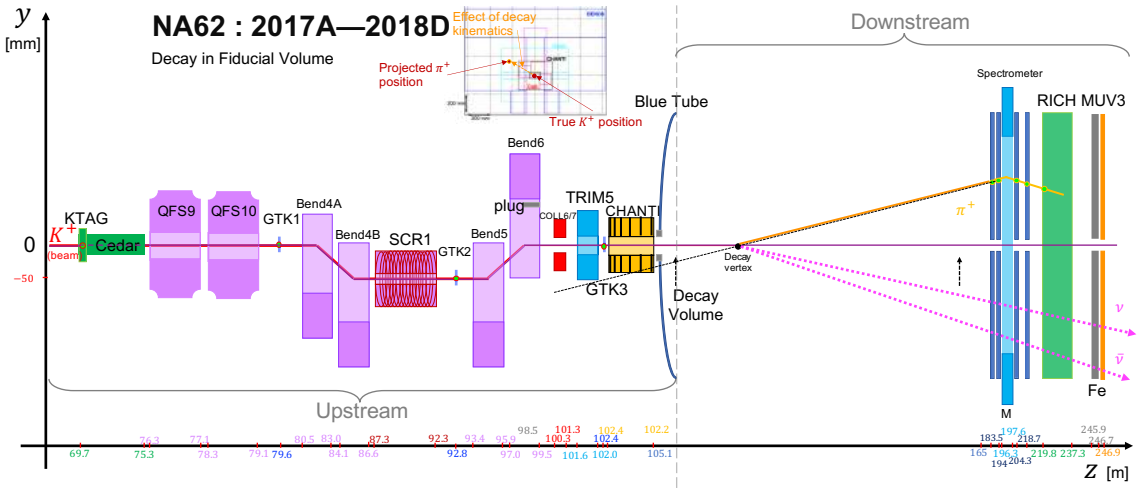


Figure 5.18: A schematic diagram of the NA62 apparatus is presented with particular focus on the upstream region in which all components are drawn approximately in proportion to their size. The dipole bending magnet, labelled as B in figure 3.3, is given its CERN component name 'TRIM5'. A $K^+ \to \pi^+\nu\bar{\nu}$ signal event is sketched. The K^+ passes through the upstream region and decays in the decay volume. The π^+ is detected downstream and its projected position in the final collimator plane is outside the central region due to the decay kinematics. Signals in the KTAG, GTK and RICH are all in time.

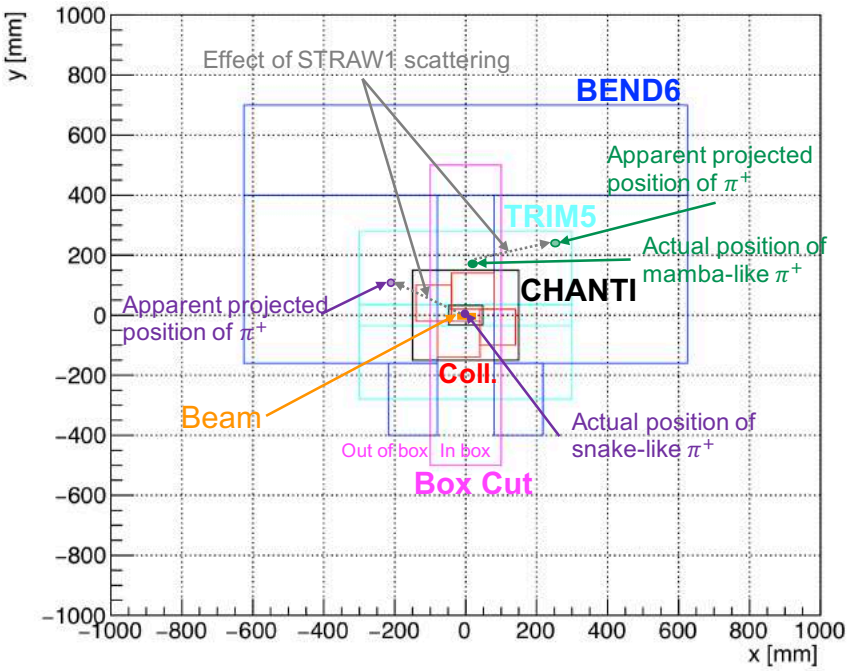


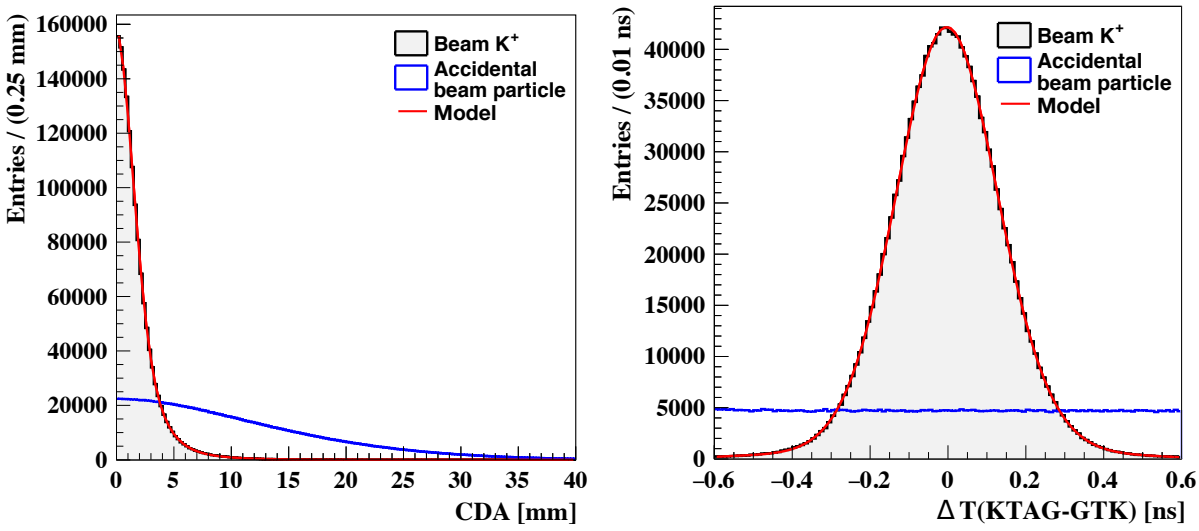
Figure 5.19: Projection of π^+ position on the (x, y) plane at the final collimator. Relevant material is shown; the yoke of the BEND6 and TRIM5 (labelled as B in figure 3.3) magnets, the material of the final collimator (Coll6,7 as used data taken before June 2018 and displayed in figures 3.3, 3.4a and 3.4) and the active area of the CHANTI. In addition the outline of the box cut accepted region is shown. The beam is rectangular in shape and passes through the centre. Also illustrated are the actual and apparent projected positions of π^+ for snake and mamba background mechanisms (see text).

probability) this lies outside the box cut region, see figure 5.18. To mimic this case for upstream backgrounds the STRAW1 scattering must similarly shift the projected π^+ position to the 'out of box' region, see figure 5.19.

The K^+–π^+ matching includes closest distance of approach (CDA) and time matching for which two time differences are defined

$$\Delta t_{KK} = t_{KTAG} - t_{GTK}\,, \tag{5.3.1}$$
$$\Delta t_{K\pi} = t_{KTAG} - t_{RICH}\,, \tag{5.3.2}$$

based on measured times associated to the K^+ upstream using the KTAG and GTK and to the π^+ downstream from the RICH. Presence of pileup tracks upstream tends to give larger Δt_{KK} since any coincidence between signals associated to the beam particle and pileup track is purely accidental. However, for an event with a K^+ beam particle which decays and its daughter π^+ is detected downstream (as in the type 1/snakes mechanism for example), the $\Delta t_{K\pi}$ will be small because the signals in the KTAG and RICH are associated to the parent and daughter particles respectively of the same $K^+ \to \pi^+\pi^0$ decay. In contrast, if scattering or interactions occur in GTK3 then signals in the KTAG and GTK will correspond to the same K^+ and so Δt_{KK} will be small. The different distributions of CDA and Δt_{KK} are shown in figure 5.20 for cases with and without a pileup GTK track.



(a) (b)

Figure 5.20: [125] Distributions of (a) the closest distance of approach (CDA) between a candidate K^+–π^+ pair and (b) time difference between signals in GTK and KTAG, Δt_{KK} associated to the K^+. Samples for correctly associated K^+ (black) and pileup GTK tracks (blue) are displayed. The model describing correctly associated K^+ is shown in red.

Figure 5.21 and its caption explain how the composition of the upstream backgrounds is seen in the two-dimensional plane formed by the two time differences $\Delta t_{K\pi}$ and Δt_{KK} defined above.

To study the upstream background an enriched control sample is defined where selection criteria employed to suppress upstream backgrounds are relaxed. This means the box and K^+–π matching criteria and additionally, the limit on the number of allowed GTK tracks and CHANTI veto conditions. Four independent bifurcation samples are then defined, as illustrated by figure 5.22a based on whether these events pass or fail the box and K^+–π^+ matching conditions. A different combination of upstream background mechanisms enter the four bifurcation samples and events fulfilling partial criteria of the full background mechanisms also enter. For example, a snake-like event without π^+ scattering in STRAW1 has a π^+ projected position at the collimator plane inside the box and so will enter bifurcation sample B. If additionally no pileup track is well-matched then it will enter sample D. These possibilities are illustrated in figure 5.22b.

5.3.2 Simulation Studies of Snakes and Mambas Upstream Backgrounds

Studies of the 2017 data showed that the dominant component of the upstream background lies in the horizontal arm of the $\Delta t_{K\pi}$ versus Δt_{KK} plane. Events inside the box cut region are rejected which will remove all mamba events without high angle scattering in STRAW1. Therefore the primary focus of the study was on snakes. There were two goals of the study: first, to qualitatively understand the origin and features of the upstream background by reproducing the observations from analysis of data using simula-

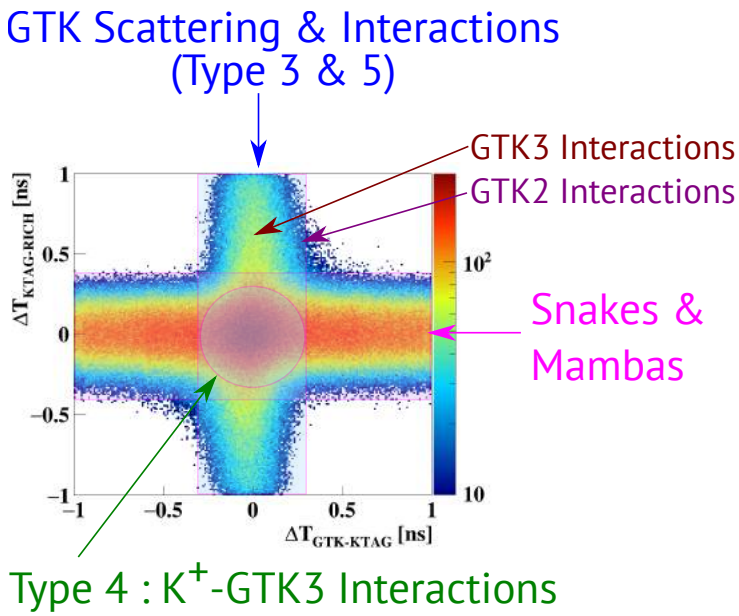


Figure 5.21: (Modified from [125].) Time differences $\Delta t_{K\pi}$ versus Δt_{KK} (defined by equations 5.3.2 and 5.3.1 respectively) are plotted for a control sample of data enriched with upstream background. The horizontal arm, with in-time signals in KTAG and RICH but not necessarily matching the GTK time, is dominated by snakes and mambas. The vertical arm is dominated by GTK scattering and interactions from type 3 or 5 upstream backgrounds, with matching time in KTAG and GTK but not necessarily with the RICH. At the centre there is a contribution is from interactions with GTK3 with significantly smaller admixture, more likely in the tails, from GTK2 interactions. The central spot also contains type 4 upstream background with KTAG, GTK and RICH times all matching.

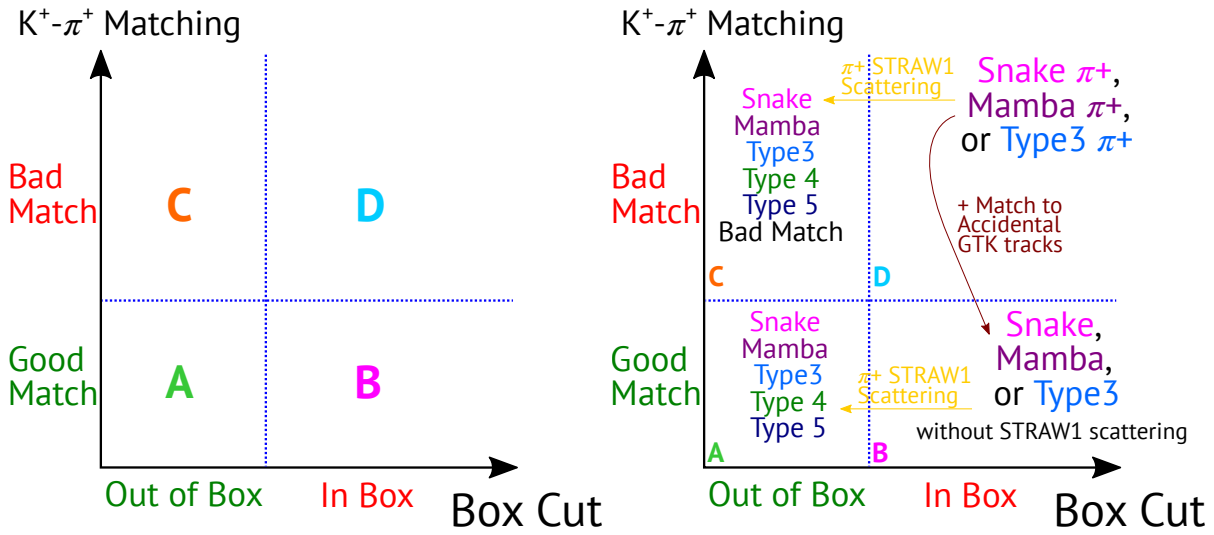


Figure 5.22: Definition of the bifurcation samples (a) and illustration of the composition of the samples (b) from each upstream background mechanism and how scattering in STRAW1 and accidental matching to pileup tracks can change the bifurcation sample an event will enter.

tions; and second, to build a quantitative model for the CDA distribution of the upstream background.

5.3.2.1 Bifurcation Method for Background Estimation

The bifurcation method of background estimation requires the definition of four samples through inverting two independent (uncorrelated) selection criteria, see figure 5.22a. If the hypothesis of uncorrelated cuts is fulfilled then the ratio of number of events in opposing pairs of bifurcation samples is equal with $\frac{n_A}{n_C} = \frac{n_B}{n_D}$ and $\frac{n_A}{n_B} = \frac{n_C}{n_D}$. The number of events in the signal bifurcation sample A can then be estimated as

$$n_A = \frac{n_B n_C}{n_D} \,. \tag{5.3.3}$$

In a blind analysis of data the signal sample A remains hidden until unblinding but the orthogonal samples B, C and D can be studied. In contrast, if the study is performed using simulations then no regions are blinded and the observed value of n_A and the value predicted from equation 5.3.3 can be compared.

5.3.2.2 Bifurcation Method and Beyond for Upstream Background Studies

For the analysis of 2016 data the bifurcation method was used to estimate the upstream background [124]. These estimations were validated using variations in the bifurcation cuts, with precision limited by the size of the bifurcated samples. For analysis of the larger 2017 data sample a more complete study was both possible and required, given the increased sensitivity.

Of particular interest was bifurcation sample C which in principle can contain events from all five upstream background mechanism types, however with an unknown mixture. It is known that mambas should contribute negligibly to this sample, this is because mamba-like events without STRAW1 scattering can be studied in a reasonably sized sample in data where the π^+ projects back to the hole in the BEND6 dipole. In bifurcation sample B events from interactions with GTK3 (mechanism type 3 and 4) are at least 10 times less frequent than matching to accidental GTK tracks (snakes and mambas)[2]. The contribution to sample C for such events is expected to be lower. Moreover, for an orthogonal selection with inverted CHANTI selection criteria, actively selecting events with associated CHANTI signals, no events were found in the signal region. This gives further evidence that the contribution from interaction mechanisms is low. In summary therefore, the dominant upstream background appears to arise from the type 1/snakes mechanism, with smaller admixture within the bifurcated samples from other mechanisms. A deeper study of the snakes background in particular, and background from K^+ decays upstream in general, was therefore performed using simulations.

[2]The factor of 10 here is an upper limit estimated by studies of control samples in data. When selection criteria are relaxed or inverted to attempt to select such events 0 are found, in contrast to the $\mathcal{O}(10)$ found to be from upstream decay mechanisms.

5.3.2.3 Upstream Decays Simulations

The primary challenge of the simulation study is to obtain a sufficiently large representative sample, challenging because of the rarity of the mechanisms and stringency of the selection. This necessitates simulation of a large number of events, naively equivalent to $\mathcal{O}(10^{13})$ K^+ decays in the standard decay region, to perform not only a counting study but also one differential in one and two-dimensional parameter spaces. In particular the CDA distribution was required for use in the final background estimation, as will be discussed in section 5.3.3. In order to generate a sufficiently large simulation sample a fast MC procedure was adopted where simulation of Cherenkov light propagation in the Cedar and all detectors downstream of the STRAWs is not performed. This increases the rate of simulation production by a factor of 60 but removes a significant amount of information usually vital to running the $K^+ \to \pi^+\nu\bar{\nu}$ selection which uses all detector systems. In order to use these fast MC simulations at each stage of the $K^+ \to \pi^+\nu\bar{\nu}$ selection where information is not available an emulated or approximated detector response must be injected to mimic the behaviour of components not simulated. Included in the fast MC simulations are emulated responses of the CHOD, NA48-CHOD and Cedar.

The decay process simulated was $K^+ \to \pi^+\pi^0$ with π^0 decaying invisibly, the latter feature further speeding up simulations since photon tracking is not required. This approach effectively assumes that the photons produced in upstream $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$ decays are always lost. The extent to which this is true depends on the position of the K^+ decay: for decays upstream of BEND6 it is expected that the assumption will be valid, while for decays after the final collimator photons will reach downstream detectors and so may be detected and the event vetoed. The assumption of lost photons holds approximately for K^+ decays at $z < 100\,\text{m}$.

To become a background event upstream decays require pileup GTK tracks to be matched to the candidate π^+. A library of $60,000$ upstream tracks and upstream detector responses from data are stored and can then be injected into a simulated event. More details are given in appendix C.1. In order to replicate the temporal properties of the upstream background from K^+ decays time measurements of the K^+ by the KTAG upstream and π^+ by the RICH downstream are emulated based on the resolutions of the detectors measured in the 2017 data sample (with $\sigma_t \approx 100\,\text{ps}$ for both detectors). By default (at the time of the analysis) simulated events were generated with times effectively set to 0 by default and therefore by emulating the KTAG and RICH time resolution and injecting a GTK pileup track at $t \approx 0$, this leads to a correlation in the $\Delta t_{K\pi}$ versus Δt_{KK} plane since the Δt_{KK} value reduces to t_{KTAG}. By instead injecting pileup tracks at a random time, according to a uniform distribution with half-width $t_{\frac{1}{2}w}$, the correlation is lost and the horizontal line in the $\Delta t_{K\pi}$ versus Δt_{KK} plane is restored, see figure 5.23. Therefore, emulated KTAG and RICH resolutions, combined with the simulated GTK response to injected pileup tracks, replicates the horizontal band in the $\Delta t_{K\pi}$ versus Δt_{KK} plane expected for upstream backgrounds from early K^+ decays, see figure 5.24.

Scattering in STRAW1 is required for the snakes and mambas background mechanisms. The process of scattering through large enough angles is relatively rare and therefore to enhance this in the simulation studies an emulated additional scattering can

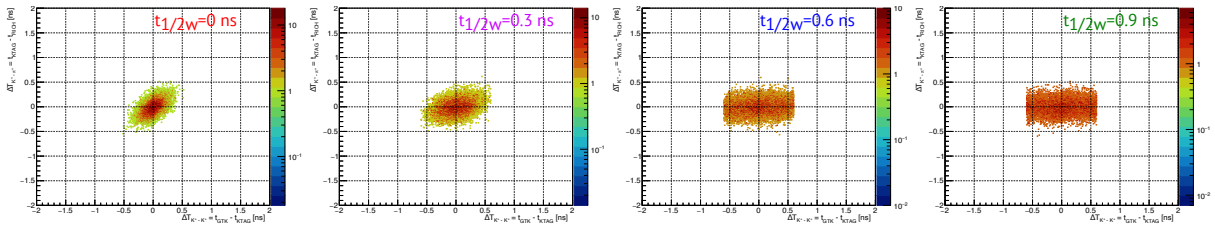


Figure 5.23: Four $\Delta t_{K\pi}$ versus Δt_{KK} distributions when injecting upstream pileup tracks (which determine the GTK time only) in different time windows of half-width $t_{\frac{1}{2}w} = 0,\ 0.3,\ 0.6,\ 0.9$ ns.

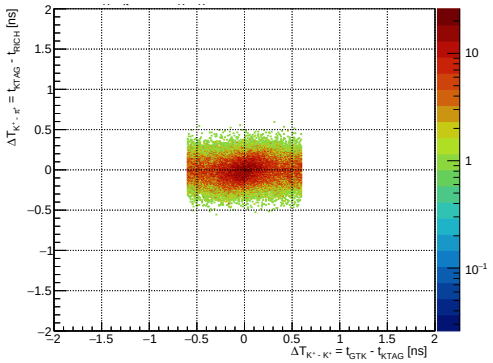


Figure 5.24: Results from simulations of upstream K^+ decays for $\Delta t_{K\pi}$ versus Δt_{KK} with emulated Cedar and RICH responses injected into the fast MC.

be added. The scattering process is symmetric in the ϕ_s angle (in the (x, y) plane) and includes rotation of the momentum vector $\mathbf{p}_\pi$ for the π^+ through an angle θ_s with respect to its direction of flight, see figure 5.25. For each event a random angle ϕ_s is drawn from a uniform distribution between $-\pi$ and π. A scattering axis is defined in the (x, y) plane according to unit vector $r_s = (1\,, 0\,, 0)$ which is then rotated by random angle ϕ_s in the (x, y) plane, perpendicular to z. The momentum vector $\mathbf{p}_\pi$ is then rotated by angle θ_s, a random value drawn from a Gaussian distribution with mean $\mu_s = 0$ and standard deviation σ_s. The impact of the additional scattering is shown in figure 5.26, with a broader distribution of projected π^+ positions at the final collimator and larger differences between the true π^+ position at the collimator and the projected back position. This leads to a migration of a subset of events from being inside the box cut to being outside, most clearly seen in events migrating from bifurcation sample D to C.

Through comparison of results of analysis of the simulations and 2017 data, quali-

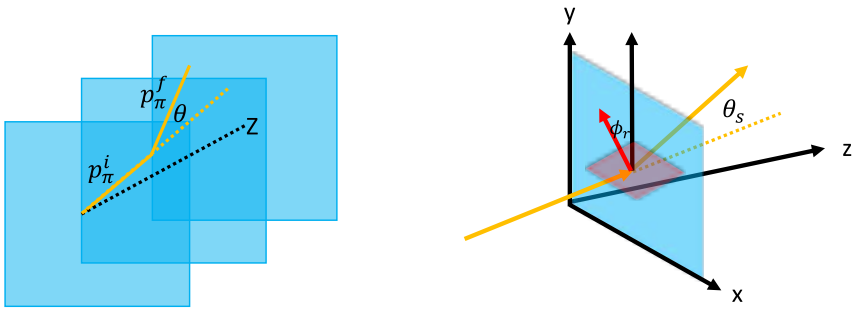
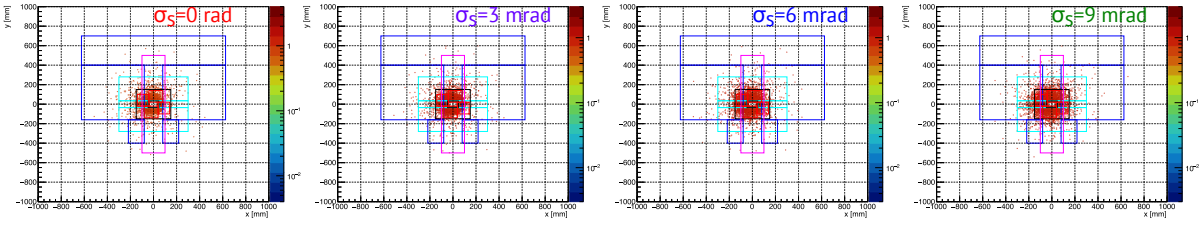
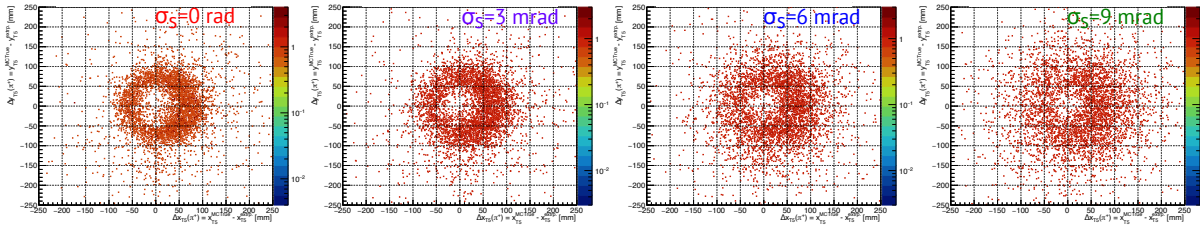


Figure 5.25: Schematic illustration of the procedure of adding emulated additional scattering of π^+ at the STRAW1, see text.



(a)



(b)

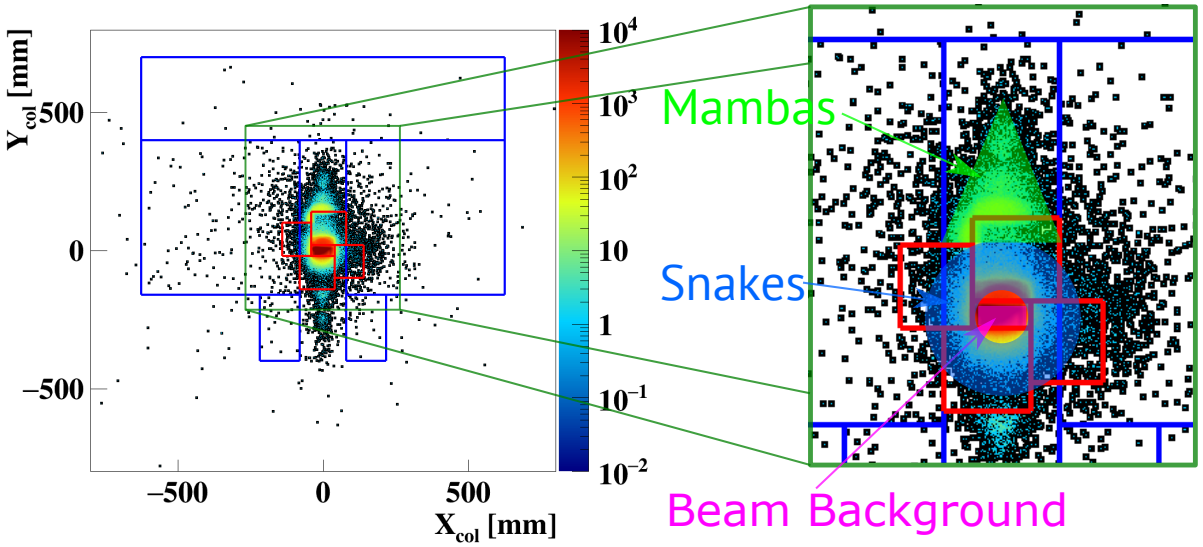
Figure 5.26: Projected position of π^+ track at reference upstream plane after the final collimator (a) and difference between this position and the true position of the π^+ when it passed through the same plane (b). Four instances are shown with increasing additional scattering from left to right as parameterised by σ_s, see text.

tatively as shown by figure 5.27 and importantly quantitatively for the CDA distribution, see section 5.3.2.4, it was established that at least the primary components (snakes and mambas) of the upstream background are replicated successfully[3]. This provided evidence that the mechanisms were well-understood. The properties of this background from upstream K^+ decays can therefore be studied with the simulations, along with the impact of selection requirements. Additional evidence and further detailed qualitative studies of the upstream background simulations are presented in appendix C. Quantitative results, used in the calculation of the expected upstream background, are presented in section 5.3.2.4.

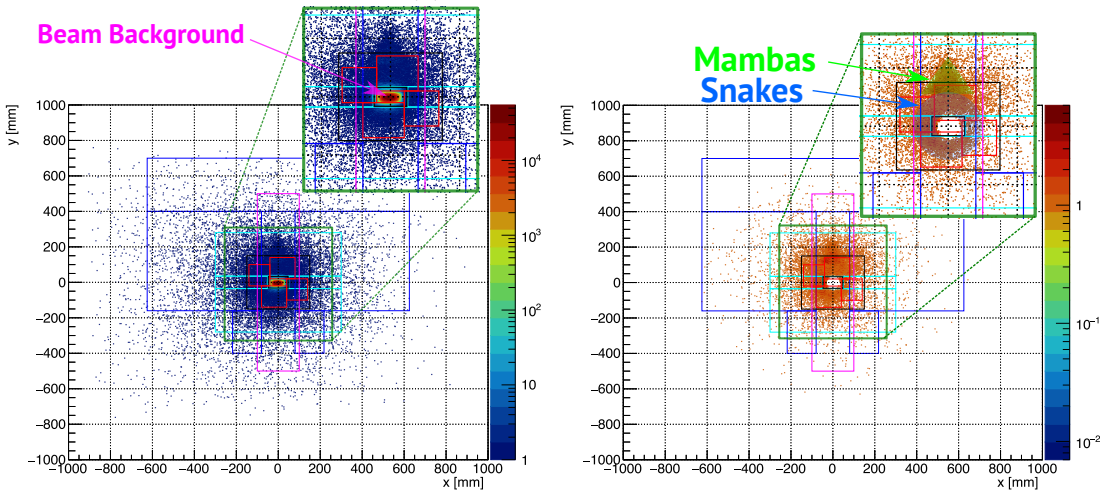
5.3.2.4 CDA Distribution Studies

To enter into specific bifurcation samples requires events to have specific geometric properties and it is because of this that simulations can provide a key input to the background predictions. The shape of the CDA distribution has an almost purely geometric origin, determined by the position of the K^+ decay, magnetic fields upstream and STRAW1 scattering downstream. The K^+–π^+ matching ensures that events with good matching (bifurcation samples A and B) have CDA$<$ 4 mm while events with bad matching generally have much larger CDA, see figure 5.28. Combining samples A+C, where the π^+ projected position is outside the box cut region (labelled OutOfBox), and samples B+D (InBox) the CDA distributions fit the same model with different values of the free parameters, as shown by figure 5.29. The normalised OutOfBox distribution is shown to be broader and flatter than the InBox case, as expected, due to scattering in STRAW1, which is required to enter the OutOfBox sample. Importantly these models derived from simulation studies can be compared to specifically isolated samples in data and are found to be in good agreement, as shown by figure 5.30b and discussed in section 5.3.3.

[3]Studies of mechanisms involving interactions with GTK stations were performed with independent dedicated simulations.

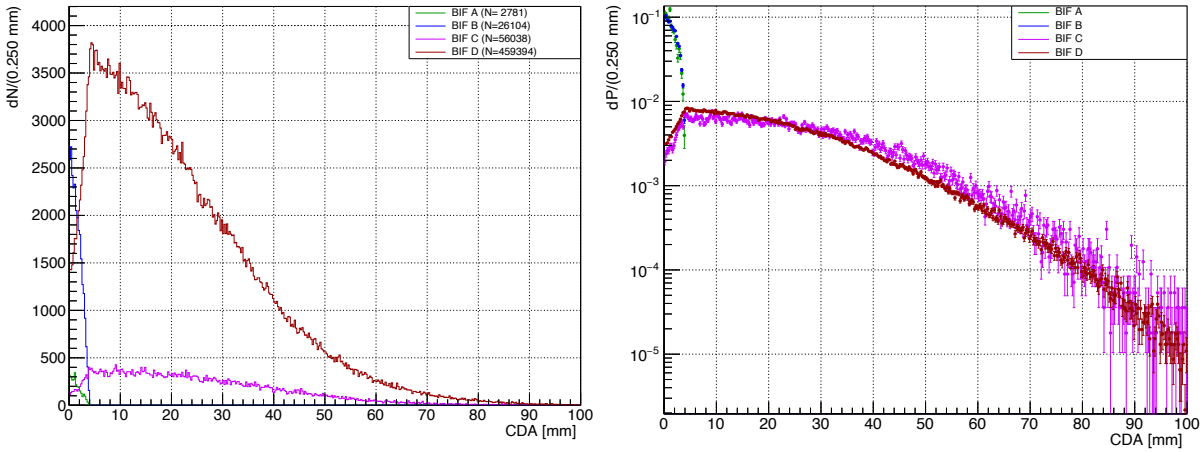


(a) Modified from [125].



(b) (c)

Figure 5.27: (a) Sample of upstream events from 2017 data with features highlighted in expanded view on right. Results of simulations are shown in (b) and (c) before and after selection cuts to reject beam background. Qualitative agreement between data and simulations demonstrates the primary mechanisms of upstream background are understood.



(a) (b)

Figure 5.28: CDA distributions for all events in bifurcation samples (a) and normalised CDA distributions (b). Samples A and B, with good K^+–π^+ matching, must have CDA< 4 mm while B and D samples, with bad matching, generally have much larger CDA. The normalised CDA show that sample C has a broader shape than D. This is expected because of the geometric nature of the mechanism and the scattering in STRAW1 which is required to move the projected π^+ position outside the box cut region.

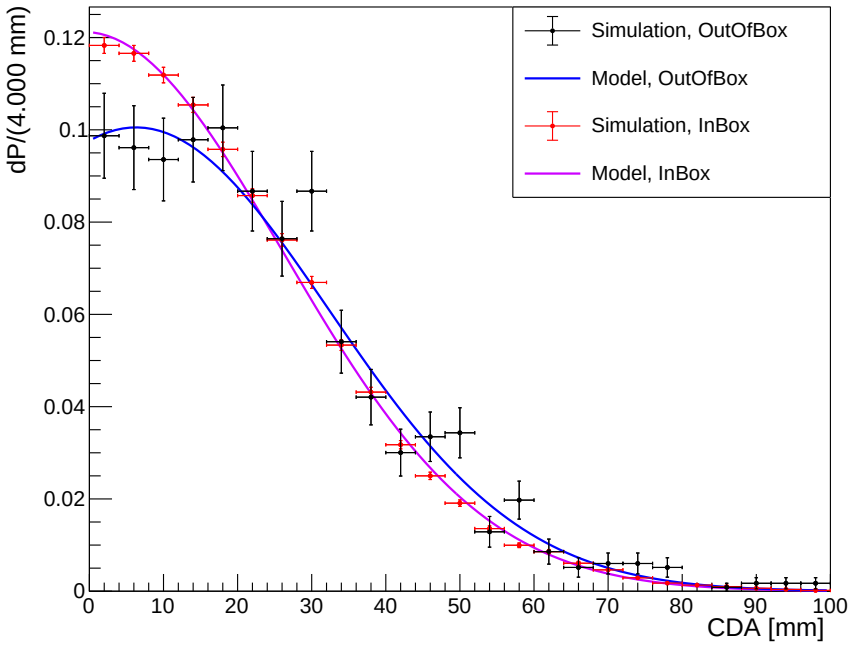


Figure 5.29: CDA distributions obtained from analysis of simulations in bifurcation regions A+C (OutOfBox) and B+D (InBox) with associated derived models. Distributions are for events in signal regions 1 and 2.

5.3.3 Implications of Upstream K^+ Decay Simulation Studies for the $K^+ \to \pi^+\nu\bar{\nu}$ Analysis

5.3.3.1 Procedure for Upstream Background Expectation for 2017 Data

The numerical prediction for the upstream background is given by

$$N^{upstr} = N^{upstr}_{\pi^+}(\lambda) P_{mistag}(\lambda)\,, \tag{5.3.4}$$

where $N^{upstr}_{\pi^+}$ is the number of events with a π^+ reconstructed downstream with momentum in range 15–35 GeV/c originating at, or upstream of, GTK3 and P_{mistag} is the probability that this π^+ is matched to a GTK track. Both are dependent on beam characteristics, instantaneous intensity, λ, momentum and beamline geometry. Additionally P_{mistag} depends on time resolution and efficiency of detectors associated with K^+–π^+ matching. Mis-tagging occurs if a) a pileup GTK track is reconstructed and selected and b) is matched to a π^+ track downstream. The probability of the two components may be expressed separately such that

$$P_{mistag}(\lambda) = P^{reco}_{pileup}(\lambda) P^{match}_{K^+-\pi^+}(\lambda)\,. \tag{5.3.5}$$

Pileup probability $P^{reco}_{pileup}(\lambda)$ depends on: if a beam track is in time, $|\Delta t| < 0.6$ ns (where Δt is Δt_{KK} or $\Delta t_{K\pi}$ for events in the horizontal and vertical band respectively of figure 5.21); well-reconstructed in GTK; and survives GTK track quality requirements. Meanwhile matching probability $P^{match}_{K^+-\pi^+}$ depends on detector time resolution (measurement of $\Delta t_{K\pi}$) and, crucially, on geometric matching effects influenced by the beamline design and encoded in the CDA distribution, see figure 5.22. The number of events in bifurcation sample C after the $K^+ \to \pi^+\nu\bar{\nu}$ selection is applied is given by

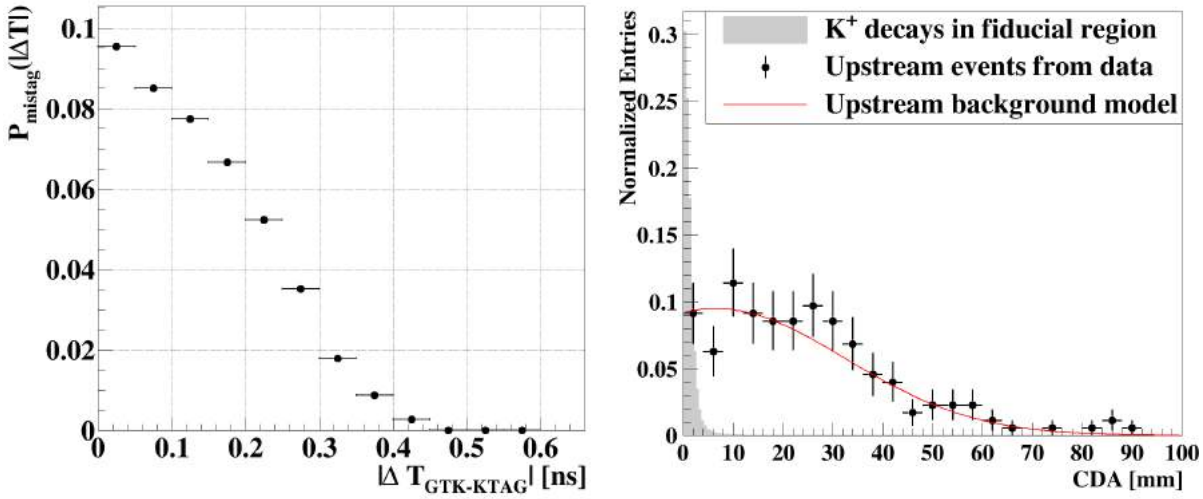
$$N_C(\lambda) = N^{upstr}_{\pi^+} P^{reco}_{pileup}(\lambda)\,, \tag{5.3.6}$$

and both this and $P^{match}_{K^+-\pi^+}(\lambda)$ can be measured in data allowing a data-driven background estimate of $N^{upstr} = N_C P^{match}_{K^+-\pi^+}$, where intensity dependence is included in the measurements of data.

The mistag probability is evaluated in twelve 100 ps wide Δt regions covering the $-0.6 < \Delta t < +0.6$ ns range allowed by the signal selection and is shown in figure 5.30a. The CDA distribution for events in bifurcation sample C from data is shown in figure 5.30b alongside the model derived from simulations, as shown in figure 5.29. The model is validated using a data sample with GTK and CHANTI veto conditions removed, considering timing side-bands 0.6 ns $< |t_{KTAG} - t_{GTK}| <$ 3 ns and without CDA$>$ 4 mm conditions (equivalent to an OutOfBox sample). Agreement of the model and the data is achieved within statistical uncertainties.

The final background expectation is given by

$$N^{upstr}_{bg} = f_{cor} \sum_{i=1}^{12} N^i_{data} P^i_{mistag}\,, \tag{5.3.7}$$



(a) (b)

Figure 5.30: [125] Mistag probability (a) of reconstructing a pileup GTK track and matching to a downstream π^+ and CDA distributions (b) for upstream decays (in bifurcation sample C) compared to K^+ decays in the standard fiducial volume.

where N^i_{data} is the number of data events selected in bifurcation sample C in bin i of ΔT_{KK}, with corresponding mistag probabilities P^i_{mistag} for each bin shown in figure 5.30a. A correction factor $f_{cor} = 1.06$ is applied to account for the fact that in bifurcation sample C events with $CDA \leq 4$ are not included. In total 16 data events are found in bifurcation sample C for 2017 data [125].

5.3.3.2 Validation of Background Expectations

The procedure described in section 5.3.3.1 is carried out under seven modified conditions, described below, for which a background prediction is provided and compared to the number of observed events in a control sample of data. The modified conditions are changes in the signal selection:

1. The box cut is replaced by requirement that the projected position of the π^+ in the collimator plane is $|x| < 100\,\text{mm}$ and $|y| < 140\,\text{mm}$. This primarily isolates snakes/type 1 upstream background events.

2. The box cut is replaced similarly with $|x| < 100\,\text{mm}$ and $|y| > 140\,\text{mm}$. This primarily isolates mambas/type 2 upstream background events.

3. Instead of the signal regions 1 and 2 the region $m^2_{miss} < -0.05\,\text{GeV}^2/c^4$ is used, where no physical events are expected except from upstream (of any mechanism).

4. The box cut is replaced as in case 1 and GTK and CHANTI veto conditions are removed. This sample includes beam-GTK interactions (primarily of upstream background mechanisms of type 3 and 4).

5. The box cut is replaced as in case 2 and GTK and CHANTI veto conditions are removed. This isolates a sample involving beam-GTK interactions (including of upstream background mechanisms of type 3, 4 and 5).

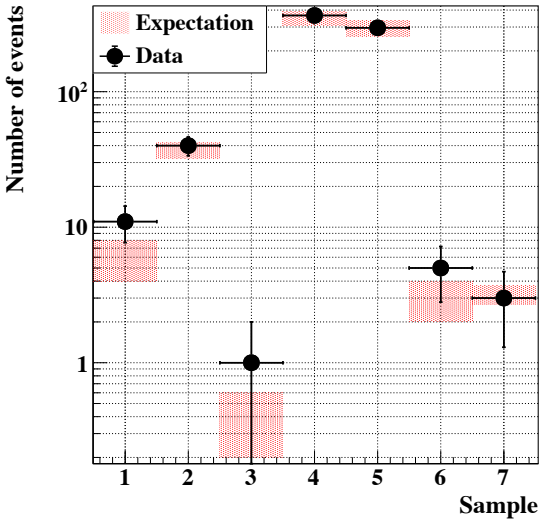


Figure 5.31: [125] Comparison between data and expectation for upstream background in the seven validation samples described in section 5.3.3.2.

6. The region $m^2_{miss} < -0.05\,\mathrm{GeV}^2/c^4$ is used (as in case 3) with GTK and CHANTI veto conditions removed.

7. GTK and CHANTI veto conditions are removed. This is a super-set of the upstream background including events passing the full selection and those usually rejected by GTK and CHANTI upstream background veto conditions.

Between these seven cases all upstream background mechanisms described in section 5.3.1 are investigated. The final case includes events in the standard signal region and therefore was initially blinded. The expectation and observation in data for the six other cases were found to agree within 1 standard deviation, case 7 was therefore unblinded and found to similarly match, see figure 5.31.

5.3.3.3 Expected Upstream Background for 2017 Data

Based on the above procedure the expected number of upstream background events in 2017 data is

$$N^{upstr} = 0.89 \pm 0.24_{stat} \pm 0.18_{syst}\,, \tag{5.3.8}$$

with statistical uncertainty from the number of data events in bifurcation sample C and systematic uncertainty of 12% from modelling of the CDA distribution and derived from the comparison between data and simulations. A further 20% systematic uncertainty is assigned based on half the difference between expected and observed number of events in validation sample 6 with similar statistical sensitivity to the upstream sample in 2017 data. This accounts for the assumption that all types of upstream background have similar CDA distributions, implicitly made by using a CDA model derived from upstream decays simulations only and somewhat validated by quality of the match of this model to events in data bifurcation sample C which can include all five types of upstream background events.

5.4 Summary of Other Backgrounds

The dominant background for the study of $K^+ \to \pi^+\nu\bar{\nu}$ is the upstream background discussed in section 5.3 while the $K^+ \to \pi^+\pi^+\pi^-$ background is shown in section 5.2 to be small and effectively negligible. In the following backgrounds from other K^+ decays in the standard decay volume are discussed.

5.4.1 $K^+ \to \pi^+\pi^0$ and $K^+ \to \mu^+\nu_\mu$ Backgrounds

Both the $K^+ \to \pi^+\pi^0$ and $K^+ \to \mu^+\nu_\mu$ decays can enter the signal region through non-Gaussian kinematic misreconstruction tails (as with $K^+ \to \pi^+\pi^+\pi^-$, see section 5.2). However, additionally for $K^+ \to \pi^+\pi^0$ inefficiency in π^0 rejection is required and for $K^+ \to \mu^+\nu_\mu$ the μ^+ must be misidentified as a π^+. The method for estimation of the background is therefore similar and is data-driven with expected number of background events in the signal regions N^{exp}_{decay} given by

$$N^{exp}_{decay} = N_{decay} f_{kin} \,, \tag{5.4.1}$$

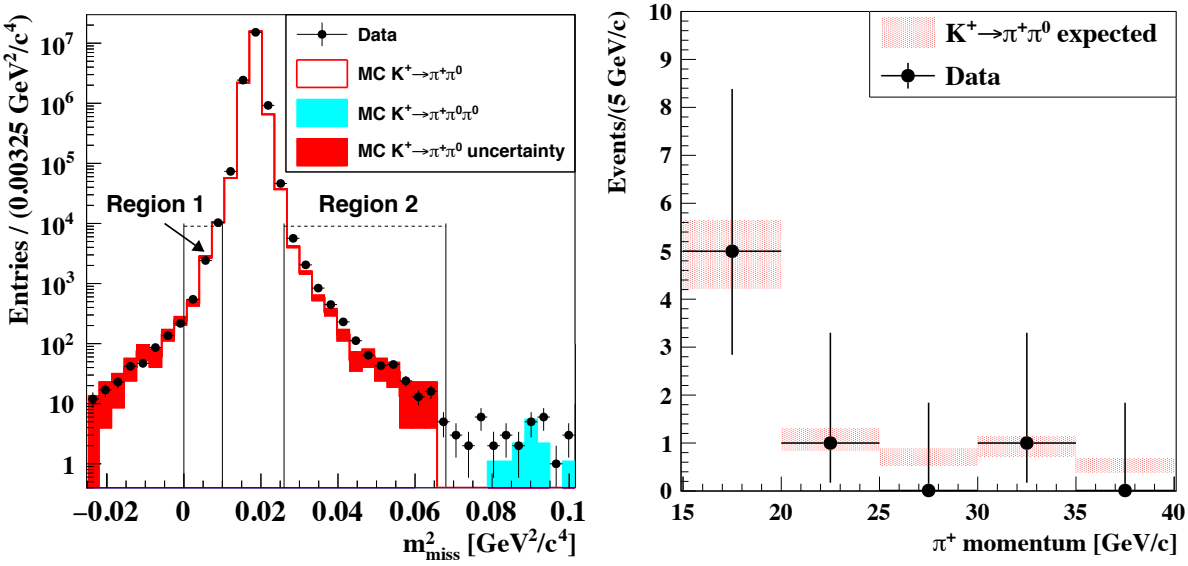
with N_{decay} being the number of a particular decay selected by the signal selection in the corresponding background m^2_{miss} versus momentum region (see figure 5.2) and f_{kin} is the kinematic tail fraction defined as in section 5.2.1. This procedure assumes f_{kin} represents the probability of N_{decay} events entering the signal region while the two terms are uncorrelated and with the sample of N_{decay} events only being due to the decay of interest. The extended momentum range of 15–40 GeV/c is used for background studies while only range 15–35 GeV/c is used for the signal, giving additional control regions. The measurement of f_{kin} is based on control samples of control (minimum bias) trigger events with the K^+–π^+ vertex at $115 < z_{vtx} < 165$ m.

The full $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$ decay chain is reconstructed using detection of both photons in the LKr, allowing isolation of an almost background-free control sample, shown by figure 5.32a, from which f_{kin} is determined and background predictions are derived as shown in figure 5.32b. Additional details are given in appendix D. Radiative $K^+ \to \pi^+\pi^0\gamma$ decays are found to enter signal region 2 but are rejected by a factor 30 greater than $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$ decays due to photon vetos. The combined total background estimate for both signal regions is

$$N^{exp,16}_{\pi\pi} = 0.064 \pm 0.007_{stat} \pm 0.006_{syst} \,, \tag{5.4.2}$$
$$N^{exp,17}_{\pi\pi} = 0.29 \pm 0.03_{stat} \pm 0.03_{syst} \,, \tag{5.4.3}$$

in the 2016 and 2017 data samples respectively.

For $K^+ \to \mu^+\nu_\mu$ decays a control sample is similarly isolated but with positive μ^+ identification from calorimetric PID and without box cut and kinematic requirements on m^2_{miss}. The control sample is shown in figure 5.33a, f_{kin} is measured and the background predictions shown in figure 5.33b. Additional details are given in appendix D. Radiative decays $K^+ \to \mu^+\nu_\mu\gamma$ are again considered and included in the prediction. Additionally, the decay chain $K^+ \to [e^+\nu_e\bar{\nu_\mu}]_{\mu^+}\nu_\mu$ with $\mu^+ \to e^+\nu_e\bar{\nu_\mu}$ decays in flight is considered and



(a) (b)

Figure 5.32: [125] m^2_{miss} distribution for $K^+ \to \pi^+\pi^0$ control sample used to study the non-Gaussuan tails entering the signal regions (a) and resulting background predictions (b) for 2017 data.

simulations show it contributes to signal region 1 only, with a predicted 0.04 ± 0.02 events in 2017 data and a negligible amount in 2016 data. The total number of $K^+ \to \mu^+\nu_\mu(\gamma)$ decays is therefore predicted, for the 2016 and 2017 data samples respectively, to be

$$N_{\mu\nu}^{exp,16} = 0.020 \pm 0.003_{syst} \pm 0.006_{syst}\,, \tag{5.4.4}$$

$$N_{\mu\nu}^{exp,17} = 0.15 \pm 0.02_{exp} \pm 0.04_{syst}\,. \tag{5.4.5}$$

The background expectations are validated by prediction of number of events in the control regions and comparison to the number of observed events, which are found to be compatible in all cases [124] [125].

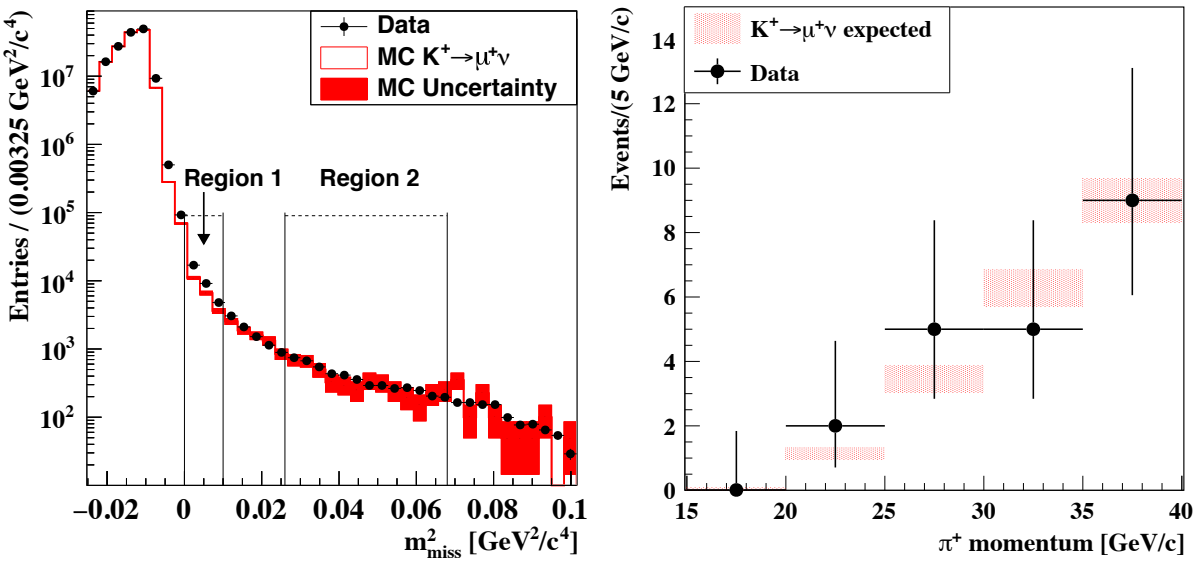
5.4.2 $K^+ \to \pi^+\pi^-e^+\nu_e$ Background

The background from $K^+ \to \pi^+\pi^-e^+\nu_e$ decays enters signal region 2 only (see figure 5.1) when two charged particles are not detected due to multiplicity rejection inefficiencies which depend on the kinematics of the charged particles. Estimation of the background is based on studies of simulations and uses the same normalisation procedure as for the calculation of the single event sensitivity, giving

$$N_{Ke4}^{exp,16} = 0.013^{+0.017}_{-0.012}|_{stat} \pm 0.009_{syst}\,, \tag{5.4.6}$$

$$N_{Ke4}^{exp,17} = 0.12 \pm 0.05_{stat} \pm 0.06_{syst}\,, \tag{5.4.7}$$

for the 2016 and 2017 data-sets respectively. The systematic uncertainty is derived from the level of agreement between data and simulations in four control samples with modified selections which are more likely to contain $K^+ \to \pi^+\pi^-e^+\nu_e$ events [125].



(a) (b)

Figure 5.33: [125] m^2_{miss} distribution for $K^+ \to \mu^+\nu_\mu$ control sample used to study the non-Gaussuan tails entering the signal regions (a) and resulting background predictions (b) for 2017 data.

5.4.3 Other K^+ Decays

The contributions from other K^+ decays are small in comparison to the others discussed in this section. The semileptonic decays $K^+ \to \pi^0\mu^+\nu_\mu$ and $K^+ \to \pi^0 e^+\nu_e$ are found to be effectively negligible, contributing less than 0.001 events in 2017 data [125]. The $K^+ \to \pi^+\gamma\gamma$ decay has the same final state as the $K^+ \to \pi^+[\gamma\gamma]_{\pi^0}$ decay chain, but has a branching ratio 5 orders of magnitude smaller (see table 2.2) and has decay kinematics which preferentially produce higher energy photons and larger m^2_{miss}. The background from the $K^+ \to \pi^+\gamma\gamma$ decay is suppressed by factors of 10^{-1} and 10^{-7} by kinematic and photon veto conditions respectively [125] and is estimated with simulations to be

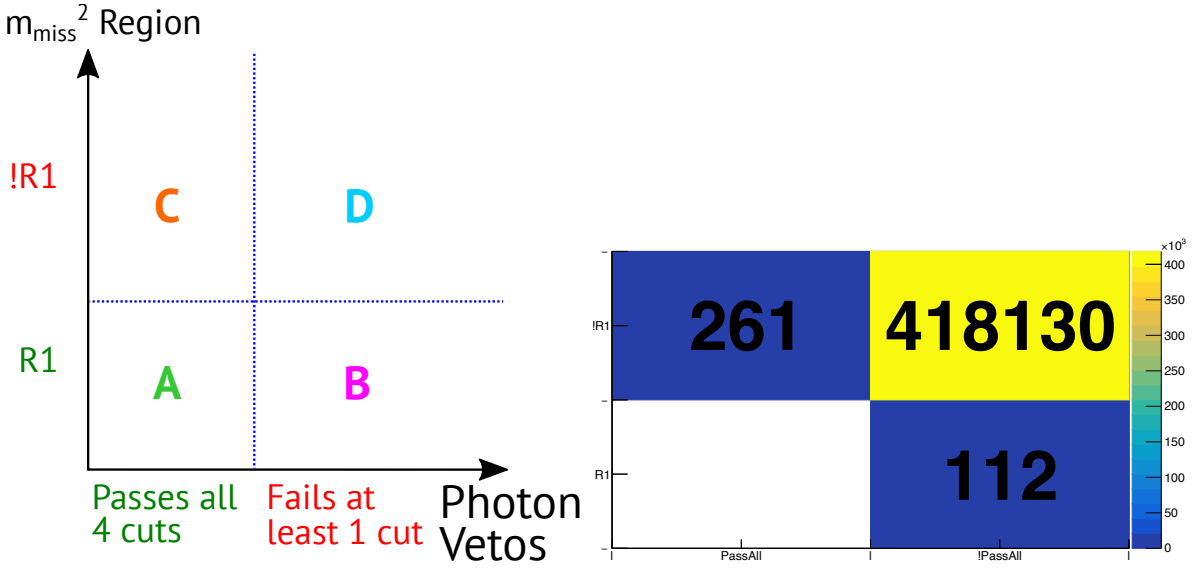
$$N^{exp,16}_{\pi\gamma\gamma} < 0.002\,, \tag{5.4.8}$$

$$N^{exp,17}_{\pi\gamma\gamma} = 0.005 \pm 0.005\,, \tag{5.4.9}$$

for the 2016 and 2017 data samples respectively.

5.5 Bifurcation Method Cross-Check of $K^+ \to \pi^+\pi^0$ and $K^+ \to \mu^+\nu_\mu$ Background Predictions For 2016 Data Analysis

For the 2016 analysis an *a posteriori* cross-check of the $K^+ \to \pi^+\pi^0$ and $K^+ \to \mu^+\nu_\mu$ background estimates was performed using a bifurcation strategy. Where the procedure could be performed, in signal region 1, the results found were consistent with those derived following the standard procedure described in section 5.4.1.



(a) (b)

Figure 5.34: Definition of bifurcation regions for $K^+ \to \pi^+\pi^0$ background estimation (a) and observed number of events in each sample in 2016 data (b).

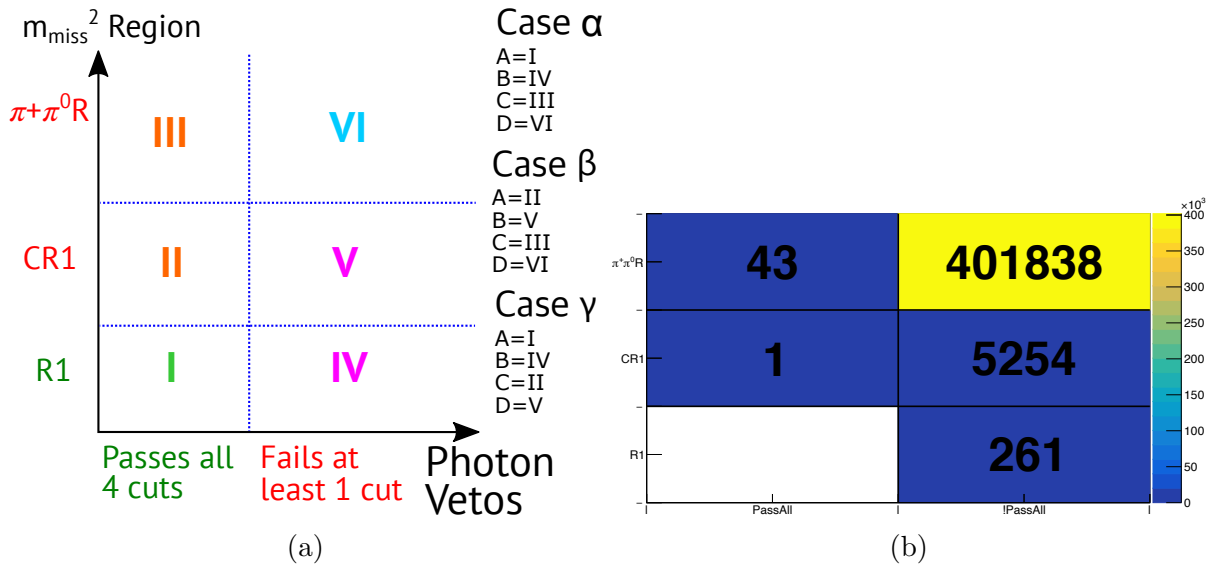
5.5.1 Bifurcation Background Prediction for $K^+ \to \pi^+\pi^0$ Decays

The bifurcation samples were defined in this case by the two orthogonal (independent) requirements of

1. Signal region definition (in terms of m^2_{miss}, see figure 5.1).
   - The standard signal region 1 (R1): $0 < m^2_{miss} < 0.01\,\mathrm{GeV}^2/c^4$.
   - Control region 1 (CR1): $0.01 < m^2_{miss} < 0.015\,\mathrm{GeV}^2/c^4$.
   - $\pi^+\pi^0$ region ($\pi^+\pi^0$R) $0.015 < m^2_{miss} < 0.021\,\mathrm{GeV}^2/c^4$.
   - Control region 2 (CR2): $0.021 < m^2_{miss} < 0.026\,\mathrm{GeV}^2/c^4$.
   - Signal region 2 (R2): $0.026 < m^2_{miss} < 0.068\,\mathrm{GeV}^2/c^4$.
2. Photon veto rejection conditions
   - Detection of ≤ 2 photons in LAVs.
   - Detection of photon(s) in LKr.
   - Photon detection in LKr via auxiliary reconstruction.
   - multiplicity rejection.

5.5.1.1 Signal Region 1

The standard bifurcation grid, along with the number of events selected in each sample is shown in figure 5.34.



(a) (b)

Figure 5.35: Definition of extended bifurcation grid for $K^+ \to \pi^+\pi^0$ background estimation (a) and observed number of events in each sample in 2016 data (b).

An alternative study can be performed using the further independent m^2_{miss} region CR1, and defining a more abstract 'extended bifurcation grid' as displayed in figure 5.35. By performing the standard bifurcation estimate of $n_A = \frac{n_B n_C}{n_D}$ with different choices of A,B,C,D three results are obtained:

- **Case** α **:** estimation of $N^{exp}_{\pi\pi}(R1)$ using $\pi^+\pi^0$R as bifurcated region. This is the standard case with: $N^{exp}_{\pi\pi,BIF}(R1|\pi^+\pi^0 R) = 0.0279 \pm 0.046_{stat}$.
- **Case** β **:** estimation of the number of events in CR1 using $\pi^+\pi^0$R as bifurcated region: $N^{exp}_{\pi\pi,BIF}(CR1|\pi^+\pi^0 R) = 0.562 \pm 0.086_{stat}$.
- **Case** γ **:** estimation of $N_{\pi\pi}(R1)$ using CR1 as bifurcated region: $N^{exp}_{\pi\pi,BIF}(R1|CR1)' = 0.050 \pm 0.050_{stat}$.

The statistical uncertainty arises from the Poisson uncertainties on the number of events in bifurcation samples used in each case. The results of case α and γ are consistent, albeit with much larger uncertainty for the latter case due to the small number of events in CR1. Case β allows a check of the background prediction in the control region which is unblinded separately and validated before the signal region is unblinded. This is therefore essential in principle if this procedure was applied before unblinding and acts as a further cross-check in this *a posteriori* analysis.

A set of additional cross-checks were performed by further fragmenting into an extended bifurcation grid, breaking down the photon rejection cuts into five categories, passing all four, any three, two, one or none of the conditions, or into 16 categories based on exact combinations of which veto requirements are passed or failed. This allowed a total of nine studies to be performed, with other cases not possible because zero events were recorded in one of the four relevant bifurcation samples. Results for case α prediction of number of events in R1 are shown in figure 5.36 and summarised in table 5.3. Study 8 involves very low statistics and is therefore excluded from detailed considerations,

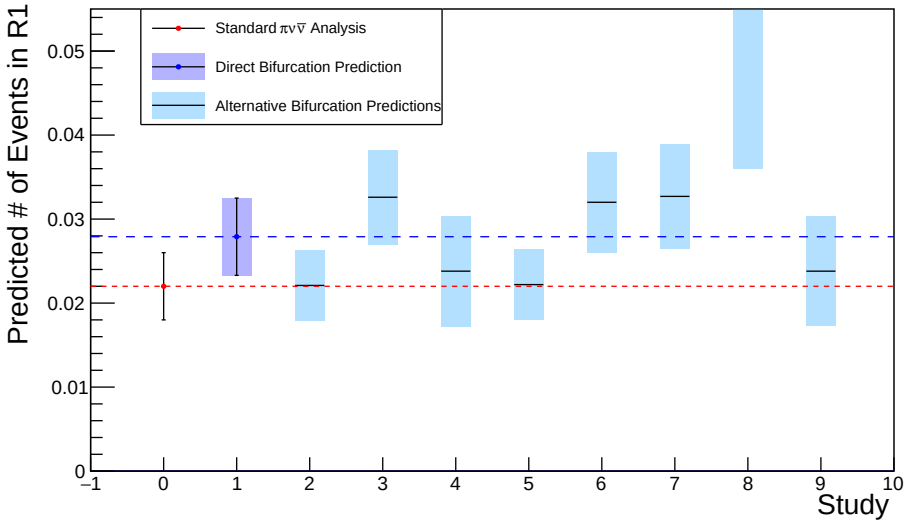


Figure 5.36: Predicted number of $K^+ \to \pi^+\pi^0$ events in signal region 1 determined by eight bifurcation studies (1-9) of case α (see text) and compared to the result using the standard procedure from section 5.4.1.

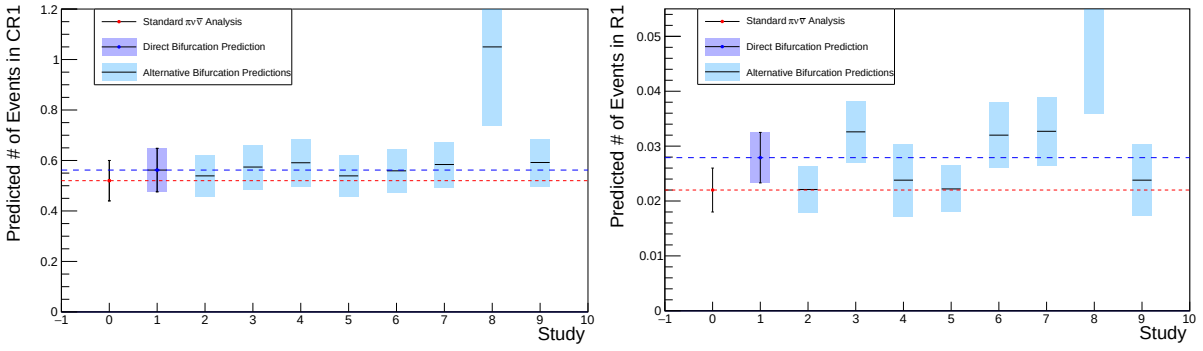
although is still consistent with all other predictions to within 1.1 standard deviations. Corresponding results for case β and γ are shown in figures 5.37a and 5.37b respectively and are included in table 5.3. Based on the variation of the studies a systematic uncertainty is assigned as the mean deviation from the initial 'standard case' prediction giving result

$$N^{exp,16}_{\pi\pi,BIF}(R1) = 0.0279 \pm 0.046_{stat} \pm 0.0056_{syst}\,, \tag{5.5.1}$$

for R1 with results for all cases (α, β and γ) given in final row of table 5.3 and compared to results from the standard procedure (section 5.4.1) and observed number of events.

Table 5.3: Number of predicted events of $K^+ \to \pi^+\pi^0$ in signal region 1 and control region 1 in 2016 data [124] using the bifurcation method, with comparison to result from standard procedure ($\pi\nu\bar{\nu}$ exp.) and observed number of events. Uncertainties are statistical unless explicitly indicated.

Study #	α : $N^{exp}_{\pi\pi}(R1\|\pi^+\pi^0 R)$	β : $N^{exp}_{\pi\pi}(CR1\|\pi^+\pi^0 R)$	γ : $N^{exp}_{\pi\pi}(R1\|CR1)$
Observed	1	0	1
$\pi\nu\bar{\nu}$ exp.	$0.022 \pm 0.004_{stat} \pm 0.002_{syst}$	$0.52 \pm 0.08_{stat} \pm 0.03_{syst}$	-
1	0.0279 ± 0.046	0.562 ± 0.086	0.050 ± 0.050
2	0.0221 ± 0.0042	0.539 ± 0.082	0.057 ± 0.041
3	0.0236 ± 0.0056	0.574 ± 0.088	0.057 ± 0.057
4	0.0238 ± 0.0066	0.591 ± 0.094	0.40 ± 0.041
5	0.0222 ± 0.042	0.539 ± 0.087	0.057 ± 0.058
6	0.0320 ± 0.060	0.559 ± 0.087	0.057 ± 0.058
7	0.0327 ± 0.0062	0.584 ± 0.090	0.056 ± 0.056
8	0.131 ± 0.095	1.05 ± 0.031	0.13 ± 0.16
9	0.0238 ± 0.0065	0.592 ± 0.094	0.040 ± 0.041
BIF	$0.0279 \pm 0.046_{stat} \pm 0.056_{syst}$	$0.562 \pm 0.086_{stat} \pm 0.020_{syst}$	$0.050 \pm 0.050_{stat} \pm 0.009_{syst}$



(a) (b)

Figure 5.37: Predicted number of $K^+ \to \pi^+\pi^0$ events in control region 1 (a) and signal region 1 (b) determined by eight bifurcation studies (1-9) of case β and γ respectively (see text) and compared to the result using the standard procedure from section 5.4.1.

5.5.1.2 Validation Tests

To validate the results consistency checks were performed, based on a procedure of relaxing the bifurcation cuts to define a new set of samples, A', B', C' and D' and an associated background prediction region given by $A' - A$, see figure 5.38. The predicted number of events in this region can be compared to the observed number. Since this does not require consideration of signal region A itself this procedure could in principle be done before final unblinding and therefore represents a reasonable validation test.

For the $K^+ \to \pi^+\pi^0$ background in R1 four validation tests were possible with the statistics available and results are presented in figure 5.39. The consistency between predicted and observed results for this $A' - A$ region provides additional evidence to validate the bifurcation background estimate procedure.

5.5.1.3 Signal Region 2

The radiative tail (from $K^+ \to \pi^+\pi^0\gamma$ decays), as discussed in section 5.4.1, extends into signal region 2 and will disproportionately enter the bifurcated samples with loosened or inverted photon veto requirements. This means that following the same approach as for signal region 1 fails due to the bias of the bifurcated sample with contamination from radiative decays. If the procedure is attempted for signal region 2 it becomes clear that inconsistent results are obtained, shown by figure 5.40a, and the key highest-statistic validation test is also failed, see figure 5.40b.

5.5.2 Bifurcation Background Prediction for $K^+ \to \mu^+\nu_\mu$ Decays

For the study of $K^+ \to \mu^+\nu_\mu$ decays the first selection requirement to define the bifurcation samples is the same (although there are different regions of interest) while the second becomes PID conditions.

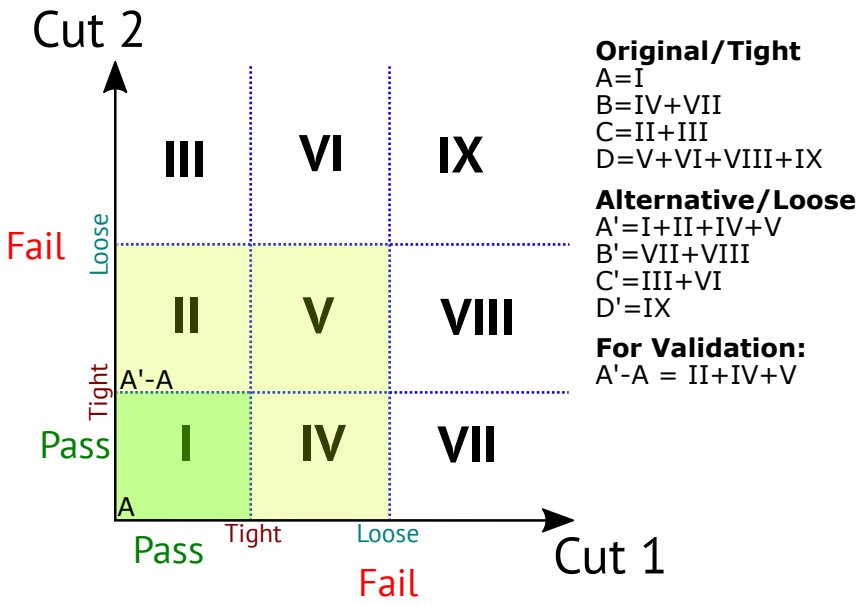


Figure 5.38: Illustration of bifurcation sample definitions for validation using standard (tight) and relaxed (loose) selection cuts to define control region $A' - A$.

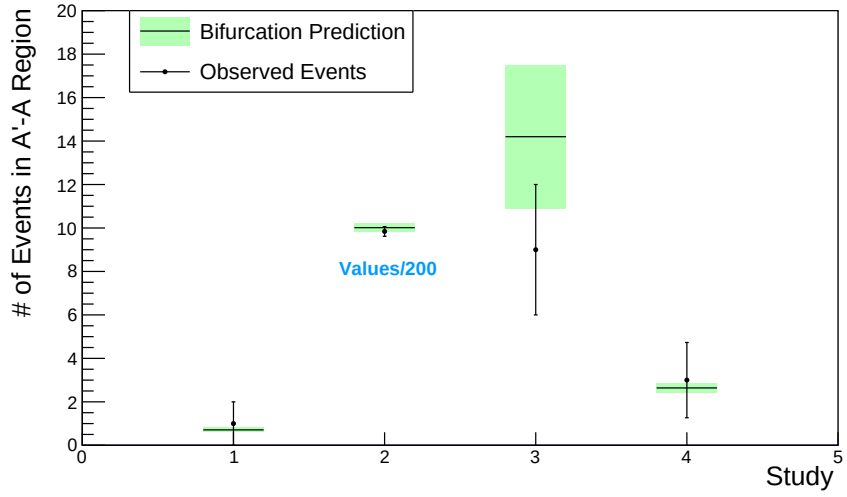
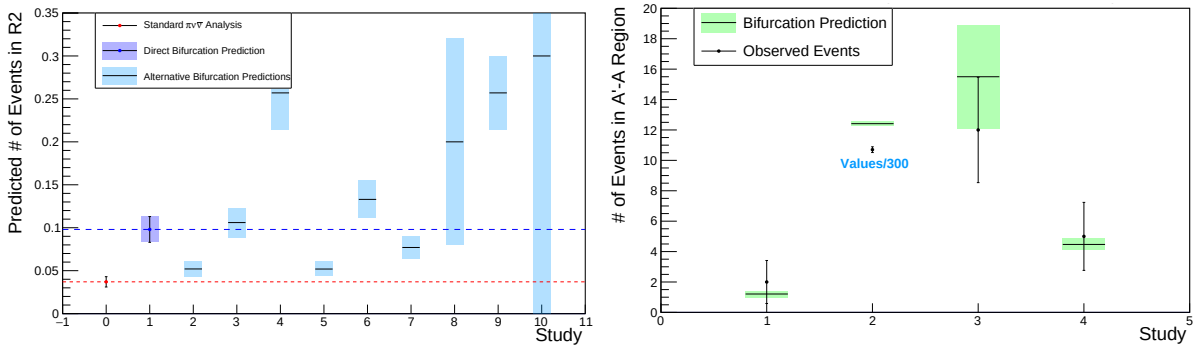


Figure 5.39: Results of bifurction background prediction validation study for the number of $K^+ \to \pi^+\pi^0$ in signal region 1. The number of expected and observed events in control region $A' - A$ is compared for four validation studies.



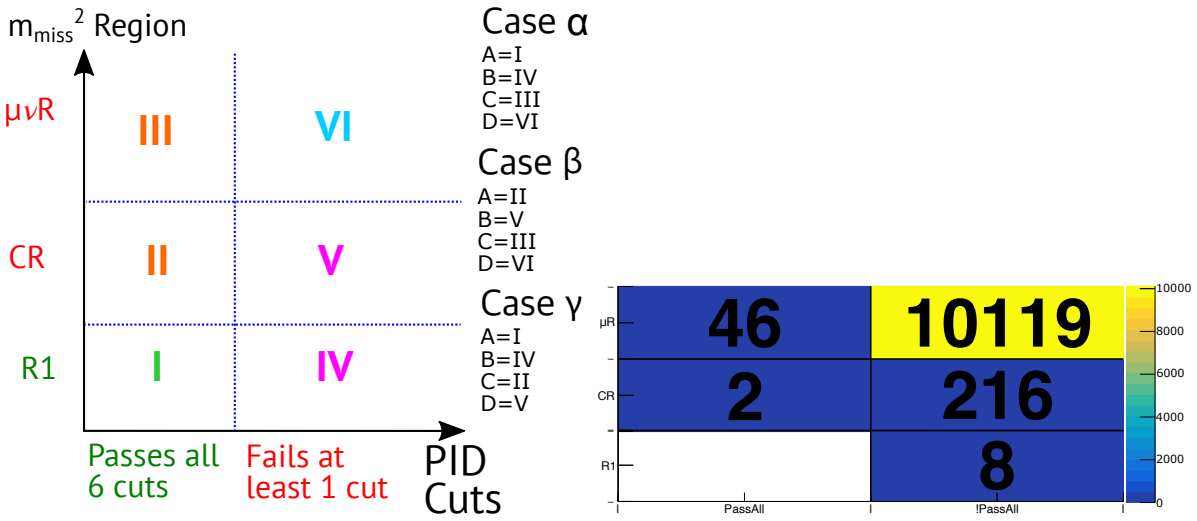
(a) (b)

Figure 5.40: Bifurcation background predictions for the number of $K^+ \to \pi^+\pi^0$ events in signal region 2 from 10 studies (of case α) compared to the standard procedure from section 5.4.1 (a) and corresponding validation studies (b). Large variations seen in (a) and inconsistency in (b) of critical test 2 demonstrate the procedure does not produce valid predictions in this case.

1. Signal region definition (in terms of m^2_{miss} versus momentum, see figure 5.1).
   - Signal region 1 (R1): $0 < m^2_{miss} < 0.01\,\mathrm{GeV}^2/c^4$ for 15–35 GeV/c^2.
   - Control region (CR): upper boundary at the border of CR1 with lower boundary curved as a function of momentum to match and exclude almost all of the expected distribution of $K^+ \to \mu^+\nu_\mu$, see figure 5.2.
   - $\mu\nu$ Region ($\mu\nu$R): Curved region as a function of momentum following expected distribution of $K^+ \to \mu^+\nu_\mu$ with m^2_{miss} calculated under the pion mass hypothesis, see figure 5.2.
   - Signal region 2 (R2): $0.026 < m^2_{miss} < 0.068\,\mathrm{GeV}^2/c^4$ for 15–35 GeV/c^2.
2. Particle identification conditions.
   - Rejection of MIPs.
   - Multivariate calorimetric discriminant probability of muon $P_{ML}(\mu) < 0.01$.
   - Equivalent probability of pion $P_{ML}(\pi) > 0.96$.
   - Set of requirements using LKr shower shape.
   - E/p ratio requirement based on total energy deposition in calorimeters.
   - E/p ratio requirement based on total energy deposited in MUV1 and MUV2.

The study is more challenging in this case because of a lower number of events in the bifurcated samples in general. However, as before, a bifurcation grid can be defined and similar cases α, β and γ defined for calculation of the number of events in R1 from $\mu\nu$R, CR from $\mu\nu$R and R1 from CR respectively, see figure 5.41.

For signal region 1 six studies are possible, however again one (study 5) is excluded because of bias from low statistics, with results shown in table 5.4 and figures 5.42 and 5.43. Validation studies are difficult due to limited statistics. In the two tests that



(a) (b)

Figure 5.41: Definition of extended bifurcation grid for $K^+ \to \mu^+\nu_\mu$ background estimation (a) and observed number of events in each sample in 2016 data (b).

could be performed by loosening calorimetric probability requirements the expected and observed number of events in the $A' - A$ regions are:

$$\text{Validation test 1: } n_{exp} = 2.03 \pm 0.32 \quad n_{obs} = 5\,, \tag{5.5.2}$$

$$\text{Validation test 2: } n_{exp} = 4.38 \pm 0.67 \quad n_{obs} = 5\,. \tag{5.5.3}$$

In general the results for bifurcation background prediction for $K^+ \to \mu^+\nu_\mu$ decays appears consistent with the standard procedure (see section 5.4.1) but, with limited statistics and therefore limited validation, is less favourable as an alternative strategy than the case of $K^+ \to \pi^+\pi^0$.

Table 5.4: Number of predicted events of $K^+ \to \mu^+\nu_\mu$ in signal region 1 and control region for 2016 data [124] using the bifurcation method, with comparison to result from standard procedure ($\pi\nu\bar{\nu}$ exp.) and observed number of events. Uncertainties are statistical unless explicitly indicated.

Study #	α : $N_{\pi\pi}^{exp}(R1\|\pi^+\pi^0 R)$	β : $N_{\pi\pi}^{exp}(CR1\|\pi^+\pi^0 R)$	γ : $N_{\pi\pi}^{exp}(R1\|CR1)$
Observed	0	2	0
$\pi\nu\bar{\nu}$ exp.	$0.036 \pm 0.003_{stat} \pm 0.006_{syst}$	$1.02 \pm 0.16_{stat}$	-
1	0.036 ± 0.014	0.98 ± 0.16	0.074 ± 0.059
2	0.028 ± 0.017	0.74 ± 0.14	0.076 ± 0.070
3	0.046 ± 0.021	1.18 ± 0.20	0.077 ± 0.065
4	0.019 ± 0.014	0.73 ± 0.14	0.053 ± 0.053
5	0.49 ± 0.50	1.47 ± 0.89	0.67 ± 0.90
6	0.088 ± 0.041	1.49 ± 0.27	0.12 ± 0.10
BIF	$0.036 \pm 0.014_{stat} \pm 0.022_{syst}$	$0.98 \pm 0.16_{stat} \pm 0.30_{syst}$	$0.074 \pm 0.059_{stat} \pm 0.018_{syst}$

For signal region 2 it is again found that the bifurcated samples for $K^+ \to \mu^+\nu_\mu$ decays are contaminated, potentially also by radiative or upstream decays. The procedure therefore similarly fails to produce consistent results, which are therefore not reported.

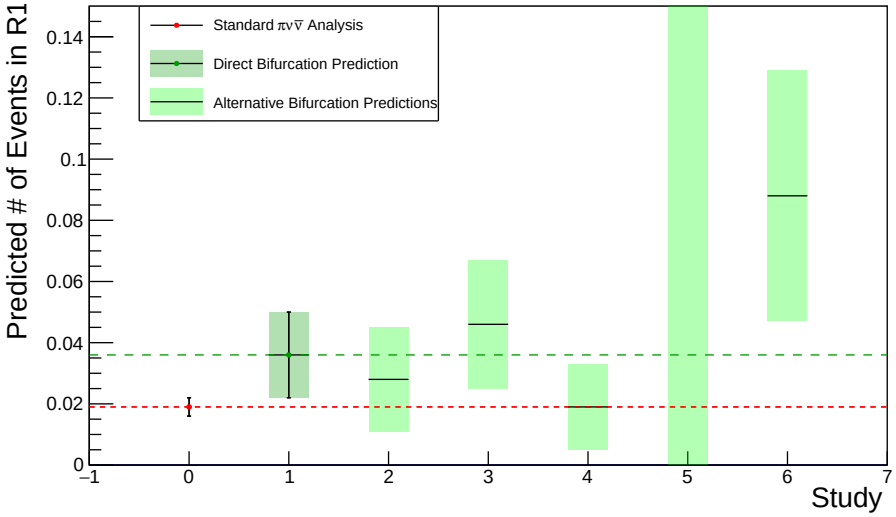
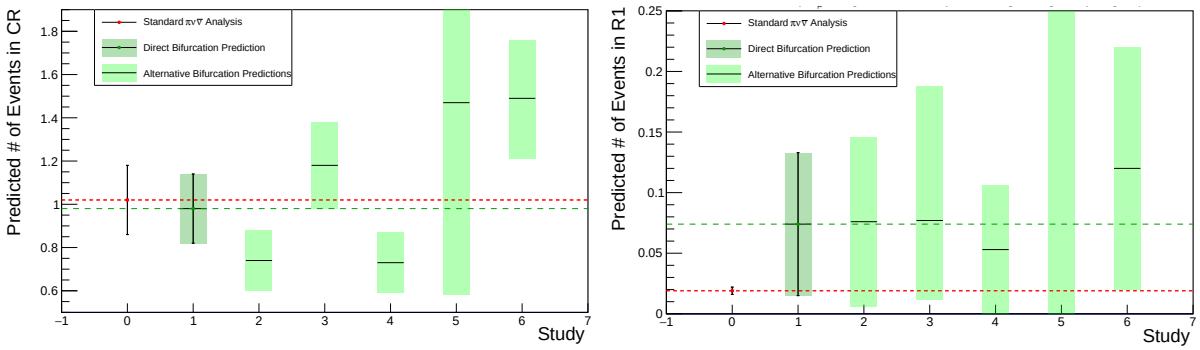


Figure 5.42: Predicted number of $K^+ \to \mu^+\nu_\mu$ events in signal region 1 determined by six bifurcation studies (1-6) of case α (see text) and compared to the result using the standard procedure from section 5.4.1.



(a) (b)

Figure 5.43: Predicted number of $K^+ \to \mu^+\nu_\mu$ events in $\mu\nu$ control region (a) and signal region 1 (b) determined by six bifurcation studies (1-6) of case β and γ respectively (see text) and compared to the result using the standard procedure from section 5.4.1.

5.6 Results for Analysis of 2016 & 2017 Data

The results of background estimations, discussed in the preceding sections, are summarised in table 5.5 along with the number of expected standard model $K^+ \to \pi^+\nu\bar{\nu}$ events based on the single event sensitivity from equations 5.1.4 and 5.1.5. Three candidate $K^+ \to \pi^+\nu\bar{\nu}$ events were found in signal region 2 with one in 2016 and two in 2017 data. For analysis of 2016 data a statistical treatment based on a hybrid Bayesian-frequentist approach [239] and the CL_S method [240], interpreting the analysis as a counting experiment with 0.267 and 0.152 ± 0.090 expected signal and background events respectively, gives an observed upper limit on the branching ratio of [124]

$$\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})_{16,NA62} < 11 \times 10^{-10} \,@\, 90\%\, CL\,. \tag{5.6.1}$$

For the combined 2016 and 2017 data-set a fully frequentist hypothesis test with profile likelihood ratio test statistic [241] was performed with upper limit obtained using the CL_S method of [125]

$$\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})_{16+17,NA62} < 1.78 \times 10^{-10} \,@\, 90\%\, CL\,. \tag{5.6.2}$$

Using the Grossman-Nir relationship of equation 2.3.23 an indirect upper limit on the neutral decay counterpart is established of [125]

$$\mathcal{B}(K_L^0 \to \pi^0\nu\bar{\nu})_{16+17,NA62} < 7.8 \times 10^{-10} \,@\, 90\%\, CL\,, \tag{5.6.3}$$

improving the current direct limit obtained by the KOTO experiment [88], as shown by figure 2.18.

Table 5.5: Number of predicted events of background and SM $K^+ \to \pi^+\nu\bar{\nu}$ signal events in 2016 [124] and 2017 [125] data.

Decay	Expected Events 2016	Expected Events 2017
SM $K^+ \to \pi^+\nu\bar{\nu}$	$0.267 \pm 0.001_{stat} \pm 0.020_{syst} \pm 0.032_{ext}$	$2.16 \pm 0.13_{syst} \pm 0.26_{ext}$
$K^+ \to \pi^+\pi^0(\gamma)$	$0.064 \pm 0.007_{stat} \pm 0.006_{syst}$	$0.29 \pm 0.03_{stat} \pm 0.26_{syst}$
$K^+ \to \mu^+\nu_\mu$	$0.020 \pm 0.003_{stat} \pm 0.006_{syst}$	$0.15 \pm 0.02_{stat} \pm 0.04_{syst}$
$K^+ \to \pi^+\pi^+\pi^-$	$0.002 \pm 0.001_{stat} \pm 0.002_{syst}$	$0.008 \pm 0.008_{syst}$
$K^+ \to \pi^+\pi^- e^+\nu_e$	$0.013^{+0.017}_{-0.012}\vert_{stat} \pm 0.009_{syst}$	$0.12 \pm 0.05_{stat} \pm 0.06_{syst}$
$K^+ \to \pi^+\gamma\gamma$	< 0.002	$0.005 \pm 0.005_{syst}$
$K^+ \to \pi^0\ell^+\nu_\ell$ ($\ell = \mu$ or e)	< 0.001	< 0.001
Upstream Background	$0.050^{+0.090}_{-0.030}\vert_{stat}$	$0.89 \pm 0.24_{stat} \pm 0.18_{syst}$
Total Background	$0.152^{+0.092}_{-0.033}\vert_{stat} \pm 0.013_{syst}$	$1.46 \pm 0.25_{stat} \pm 0.20_{syst}$

5.7 Full NA62 Run 1 Results and Future Prospects at CERN

5.7.1 Results of Analysis of 2018 Data and Full Run 1 Results

During the 216 days of data-taking in 2018 NA62 operated at higher instantaneous intensity, approximately 10% higher than in 2017 on average. In addition in June 2018

the final collimator was replaced with a larger fixed collimator (TCX) which significantly suppresses upstream background [201], removing the mamba mechanism discussed in section 5.3.1. As a result for data after the installation of the TCX the 'box cut' applied to reject upstream background could be significantly relaxed leading to an increase in acceptance. In addition updated PID and upstream background rejection, based on machine learning boosted decision tree classifiers, were applied in the selection. It was found that the momentum range for signal region 2 could be expanded, to the range 15–45 GeV/c, with sensitivity to the $K^+ \to \pi^+\nu\bar{\nu}$ signal with reasonable signal/background ratio. The analysis was then formally split into seven categories, the 2018 sample before the TCX installation (both signal regions in full momentum ranges), and in six momentum bins of 5 GeV/c width after the TCX installation (combining both signal regions for momenta 15–35 GeV/c and just signal region 2 for 35–45 GeV/c). Some selection criteria, including the R_{STRAW1} versus z_{vtx} cut discussed in appendix C.2, were optimised in specific categories to maximise signal acceptance and background rejection.

These analysis developments lead to a significant improvement in acceptance, especially for the analysis after the TCX installation, and with the increased data sample size the single event sensitivity achieved is

$$\mathcal{B}^{18}_{SES} = (1.11 \pm 0.07_{syst}) \times 10^{-11}\,, \tag{5.7.1}$$

an improvement by a factor 3.5 over the 2017 analysis. The signal/background ratio is maintained at the same relative level as for the 2017 analysis leading to a total expected number of SM signal and background events of

$$N^{exp,18}_{\pi\nu\nu,SM} = 7.58 \pm 0.40_{syst} \pm 0.75_{ext}\ , \tag{5.7.2}$$

$$N^{18}_{bg} = 5.28^{+0.99}_{-0.74}\ . \tag{5.7.3}$$

Procedures for background expectations were updated and further control regions added. In all control regions expectations were found to satisfactorily match the observations, see figure 5.44, and therefore the signal regions were unblinded. In total 17 events were observed, four and thirteen events in signal regions 1 and 2 respectively, as shown by figure 5.45 [89].

Statistical combination, of the seven categories in 2018 data results with those from 2017 and 2016 as two more categories, leads to a preliminary NA62 Run 1 measurement of [89]

$$\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu}) = (11.0^{+4.0}_{-3.5}|_{stat} \pm 0.3_{syst}) \times 10^{-11}\,. \tag{5.7.4}$$

This corresponds to evidence for observation of the $K^+ \to \pi^+\nu\bar{\nu}$ decay, excluding the background-only hypothesis with a 3.5σ significance.

Based on this measurement the Grossman-Nir bound, see equation 2.3.23, predicts a constraint on the $K_L \to \pi^0\nu\bar{\nu}$ decay of

$$\mathcal{B}(K^0_L \to \pi^0\nu\bar{\nu})_{exp,NA62GNl} < 7.6 \times 10^{-10}\ @90\%\,CL\,, \tag{5.7.5}$$

which is consistent with the SM prediction given in equation 2.3.17. This constraint is two orders of magnitude less than the present direct limit on the branching ratio [88]. These updated results are shown in figure 5.46.

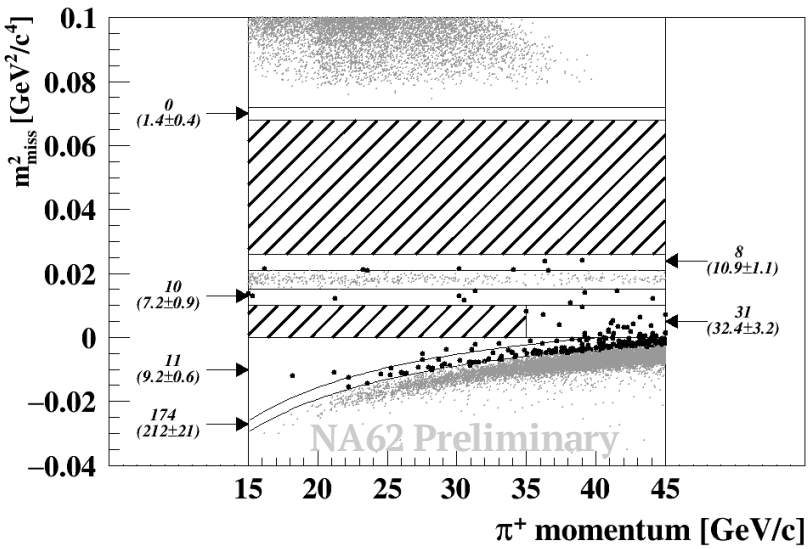


Figure 5.44: [89]Preliminary squared missing mass versus momentum distribution showing expected and observed number of events in control regions which are in good agreement.

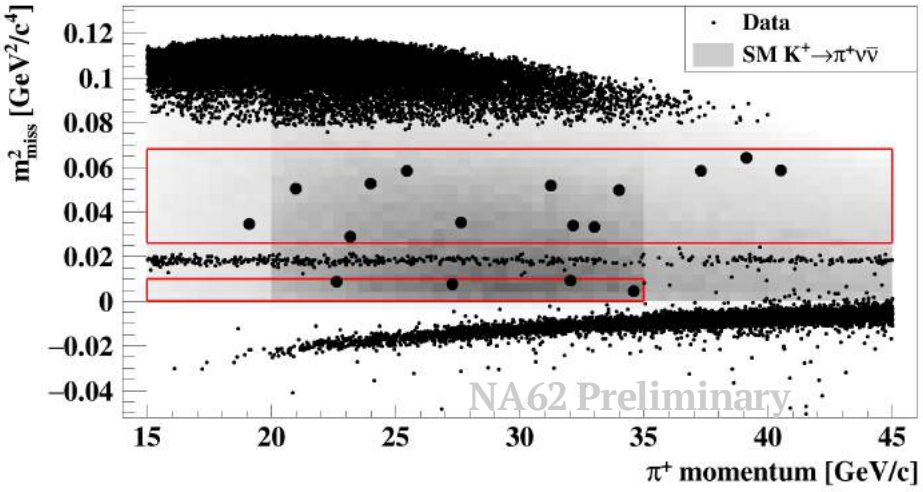


Figure 5.45: [89]Preliminary squared missing mass versus momentum results plot for the analysis of the 2018 data-set showing the 17 observed events.

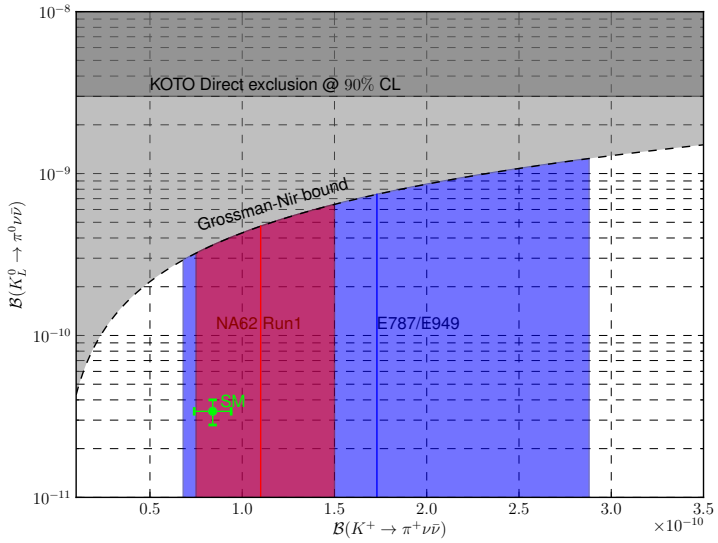


Figure 5.46: Preliminary updated landscape of constraints on the branching ratios of $K^+ \to \pi^+\nu\bar{\nu}$ and $K_L \to \pi^0\nu\bar{\nu}$ with comparison to the SM predictions.

5.7.2 Future Prospects

The NA62 experiment will resume data-taking in 2021 at higher intensity and with new veto detectors (see section 3.2.13) to further reduce the upstream background, the dominant background component in the Run 1 analyses. With the significant new data sample to come in Run 2, and additional analysis developments under investigation, the experiment is on track for discovery of the $K^+ \to \pi^+\nu\bar{\nu}$ decay at 5σ significance and proceeding to a branching ratio measurement with approximately 10% precision.

Looking further ahead, beyond 2024, investigations are underway considering the possibility for higher precision branching ratio measurements using an upgraded experimental apparatus and intensity approximately four times higher than expected in 2021. Such precision would go beyond that currently achieved with theoretical calculations and could lead to clear evidence for new physics effects.

A complementary advanced design study is ongoing for the proposed KLEVER experiment [242] which would run in the same experimental hall at CERN as NA62, with a new detector (based on a similar layout) for the study of $K_L \to \pi^0\nu\bar{\nu}$ decays. This challenging measurement would require significant beam intensity increase and detector development but could lead to a measurement of $\mathcal{B}(K_L \to \pi^0\nu\bar{\nu})$ at $\sim 20\%$ precision [242].

Chapter 6

Search for a Feebly Interacting Particle X in the Decay $K^+ \to \pi^+ X$

6.1 Analysis Outline

In section 2.4 the search for the production of a feebly interacting long lived particle X in the decay $K^+ \to \pi^+ X$ was motivated. The experimental signature for $K^+ \to \pi^+ X$ is the same as for $K^+ \to \pi^+ \nu\bar{\nu}$ except the two-body decay is expected to peak in the m^2_{miss} variable. An experimental search has been performed through interpretation of the NA62 $K^+ \to \pi^+ \nu\bar{\nu}$ analysis of 2017 data (referred to hereafter in this chapter as the '$\pi\nu\bar{\nu}_{2017}$ analysis'). No modifications were made to the selection or signal regions. While in principle a dedicated analysis would be able to achieve improved results, this direct interpretation analysis was possible exploiting the studies for the $\pi\nu\bar{\nu}_{2017}$ analysis with small additions, following the analysis procedure described below.

1. The signal ($K^+ \to \pi^+ X$) selection is identical to the $K^+ \to \pi^+ \nu\bar{\nu}$ signal selection.

2. A scan is performed, evaluating a range of X mass hypotheses by testing for the presence of a peaking signal in a 'search window' around each mass hypothesis with a width determined by the resolution of the observable.

3. Simulations of $K^+ \to \pi^+ X$ decays are analysed, providing measurements of the resolution of the observable, signal selection acceptance and single event sensitivity for each mass hypothesis.

4. The background expectation is established for each search window. This includes all sources of background from the $\pi\nu\bar{\nu}_{2017}$ analysis, see chapter 5, with the addition of the $K^+ \to \pi^+ \nu\bar{\nu}$ decay itself which becomes the dominant background by construction.

5. The number of events observed in each search window is determined after applying the full $\pi\nu\bar{\nu}_{2017}$ selection.

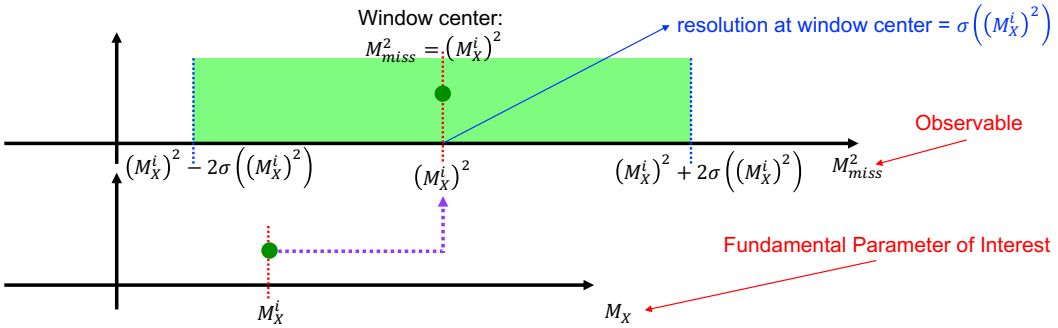


Figure 6.1: Schematic illustration of the definition of mass scan windows for the $K^+ \to \pi^+ X$ search showing the relationship between the observable M^2_{miss} and the parameter of interest M_X.

6. Expected and observed upper limits are then set on $\mathcal{B}(K^+ \to \pi^+ X)$ at 90% confidence level for each mass hypothesis using the CL_S procedure with a log likelihood ratio test statistic [13] [240] [243] [244]. This requires the information gathered above: the total expected background and associated uncertainty $N_b \pm \sigma_b$, the observed number of events n and the single event sensitivity.

6.2 Analysis Strategy

6.2.1 Mass Scan Procedure

The observable for the analysis is the squared missing mass, as defined by equation 5.1.1. This observable is defined exactly at $m^2_{miss} = 0\,\mathrm{GeV}^2/c^4$, relevant to the theoretically interesting case of $M_X = 0\,\mathrm{MeV}/c^2$, and for negative values[1] and can be reconstructed experimentally.

While the observable is m^2_{miss} the goal of the study is to search for production of X with a set of mass hypotheses. The parameter of interest is therefore the X mass, M_X. For a given mass hypothesis M^i_X the corresponding value of the observable is $m^2_{miss}|_i = (M_X)^2$, this defines the centre of the search window. The width of the search window is chosen to be four times the resolution on the m^2_{miss} observable, $\sigma(m^2_{miss})$, giving range $|m^2_{miss} - (M^i_X)^2| < 2\sigma\left((M^i_X)^2\right)$, as shown by figure 6.1. Therefore, to define the width of the windows the observable resolution $\sigma(m^2_{miss})$ must be determined, see section 6.3.1.

6.2.2 Simulation Strategy

Simulations were performed of $K^+ \to \pi^+ X$ decays with stable X for a range of masses 0–277.2 MeV/c with a mass step of 1.4 MeV/c^2. This covers the full range of sensitivity with a step size which, in the corresponding observable, is always smaller than the resolution

[1] Note the convention that squared missing mass is measured in units of GeV^2/c^4 while the X mass is measured in MeV/c^2.

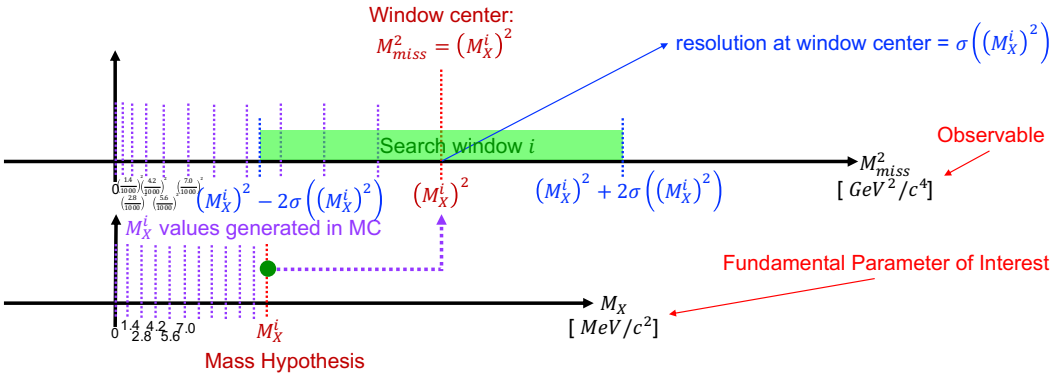


Figure 6.2: Schematic sketch of simulation strategy and matching to window definitions. The generated X mass values are taken to be equal to the X mass hypotheses to be tested M^i_X and this sets the the window centre in the observable $m^2_{miss}|_i = (M^i_X)^2$ and the window width $\left|m^2_{miss} - (M^i_X)^2\right| < 2\sigma\left((M^i_X)^2\right)$ using the measured resolution at the window centre $\sigma\left((M^i_X)^2\right)$.

of the observable (see section 6.3.1). The generated M^i_X values can be taken as the mass hypotheses for the mass scan and therefore the window centres and widths in the squared missing mass observable are set according to the prescription in section 6.2.1 and illustrated in figure 6.2.

6.2.3 Determination of Signal Acceptance and Single Event Sensitivity

Using the simulation strategy described in section 6.2.2 there is a sample of simulated events for each mass hypothesis M^i_X for which the signal acceptance can be determined according to

$$A(M^i_X) = \frac{N_{MC}(M^i_X,\ \text{PnnFV}, \text{PnnSel}, \text{In Window}\, i)}{N_{MC}(M^i_X,\ \text{PnnFV})} = \frac{N_{MC}(M^i_X,\ \text{PnnFV}, \text{PnnSel})}{N_{MC}(M^i_X,\ \text{PnnFV})} f_i\ , \tag{6.2.1}$$

where the denominator is the number of simulated events for mass hypothesis M^i_X with the $K^+ \to \pi^+ X$ decay happening in the $\pi\nu\bar{\nu}_{2017}$ analysis decay volume $105 < z < 165\,\text{m}$. The numerator is the number of these events which pass the $\pi\nu\bar{\nu}_{2017}$ selection and have reconstructed squared missing mass inside the i^{th} search window:

$$\left|m^2_{miss}|_{reco} - \left(M^i_X\right)^2\right| < 2\sigma\left((M^i_X)^2\right)\ . \tag{6.2.2}$$

The fraction of selected events contained within this search window can also be expressed by factor

$$f_i = \begin{cases} \frac{N_{MC}(M^i_X, \text{PnnFV}, \text{PnnSel}, \text{In window}\, i)}{N_{MC}(M^i_X, \text{PnnFV}, \text{PnnSel})} & \text{PnnSel} \\ 0 & \text{!PnnSel (event does not pass } \pi\nu\bar{\nu}_{2017} \text{ selection)}\ . \end{cases} \tag{6.2.3}$$

The choice of a window four times the resolution wide ($\pm 2\sigma$) means 95% of events with true squared missing mass equal to the window centre, $m^2_{miss}|_{true} = (M^i_X)^2$, will have reconstructed squared missing mass inside the window. Therefore in general $f_i \approx 0.95$. Events must have $m^2_{miss}|_{reco}$ within the signal regions to pass the full $\pi\nu\bar{\nu}_{2017}$ selection, therefore search windows overlapping with the boundaries of the signal regions will have lower values of f. The value of f falls away, passing through 0.5×0.95 at the signal region boundary and reaching zero when a search window no longer overlaps with either signal region. For completeness it is noted that both terms in equation 6.2.2 tend to zero since to enter the numerator in the first term an event must pass the $\pi\nu\bar{\nu}_{2017}$ selection.

The single event sensitivity, $\mathcal{B}_{SES}(K^+ \to \pi^+ X)$, can then be defined with respect to the determination used for the $\pi\nu\bar{\nu}_{2017}$ analysis (see section 5.1.4) through a ratio of acceptances for each search window

$$\mathcal{B}_{SES}(\pi X, M^i_X) = \mathcal{B}_{SES}(\pi\nu\bar{\nu}, R1+R2)\frac{A(\pi\nu\bar{\nu}, R1+R2)}{A(\pi X, M^i_X)} \ , \tag{6.2.4}$$

From the $\pi\nu\bar{\nu}_{17}$ analysis results: $\mathcal{B}_{SES}(\pi\nu\bar{\nu}, R1+R2) = (3.89 \pm 0.21) \times 10^{-11}$ and $A(\pi\nu\bar{\nu}, R1+R2) = 0.030 \pm 0.004$ [125]. The $\mathcal{B}_{SES}$ for the $\pi\nu\bar{\nu}$ analysis is defined based on ratios relative to the normalisation $K^+ \to \pi^+\pi^0$ decay, as

$$\mathcal{B}_{SES}(\pi\nu\bar{\nu}, R1+R2) = \frac{A(\pi^+\pi^0)}{A(\pi\nu\bar{\nu}, R1+R2)}\frac{\mathcal{B}(K^+ \to \pi^+\pi^0)}{N_{\pi^+\pi^0}}\frac{\varepsilon_{RV}}{D}\frac{\varepsilon^{CTRL}_{\pi\pi}}{\varepsilon^{PNN}_{\pi\nu\nu}} \ , \tag{6.2.5}$$

(see section 5.1.4) and therefore the SES for the $K^+ \to \pi^+ X$ search is linked directly to this normalisation according to

$$\mathcal{B}_{SES}(\pi X, M^i_X) = \frac{\mathcal{B}(K^+ \to \pi^+\pi^0)}{N_{\pi^+\pi^0}}\frac{A(\pi^+\pi^0)}{A(\pi X, M^i_X)}\frac{\varepsilon_{RV}}{D}\frac{\varepsilon^{CTRL}_{\pi\pi}}{\varepsilon^{PNN}_{\pi\nu\nu}} \ . \tag{6.2.6}$$

The $\pi\nu\bar{\nu}_{2017}$ analysis is performed in 4 bins of momentum and 5 bins of intensity. Since this result is an interpretation of the same dataset the intensity dependence is the same and will cancel in the ratio (equation 6.2.4) but the momentum spectra of the $K^+ \to \pi^+\nu\bar{\nu}$ and $K^+ \to \pi^+ X$ decays are different (and the $K^+ \to \pi^+ X$ spectrum will change with X mass) and therefore this must be accounted for when calculating $\mathcal{B}_{SES}(\pi X)$, see section 6.3.5.

6.2.4 Treatment of Unstable X

The treatment above assumes X is stable, meaning its lifetime $\tau_x = \infty$, (or X decays to 'invisible' undetectable particles such as Dark Matter) however in principle X may have a finite lifetime and decay to visible SM particles before the end of the NA62 detector apparatus. K^+ decays are required to occur between 105 and 165 m, if X subsequently decays before MUV3 ($z_{MUV3} = 246.8$ m) then it is assumed the decay products would be detected and the event vetoed. The inefficiency for MUV3 detection of single muons is measured to be 0.15% and for an event with a $X \to \mu^+\mu^-$ decay the average inefficiency would be smaller than per-mille. Therefore in the case of muon decays this assumption is well-motivated. For decays to other light SM particles ($e^\pm, \pi^\pm, \pi^0, \gamma$) in general the

sensitive photon rejection or calorimetric systems will still detect decay products, however the efficiency depends somewhat on location of the X decay.

Using the above assumption, acceptance and $\mathcal{B}_{SES}(\pi X)$ results can be derived for X with finite lifetime τ_X by weighting each simulated event (entering the numerator of the acceptance, equation 6.2.1) by the probability X does not decay before the MUV3,

$$P(X\,\text{survives}) = e^{-\left(\frac{z_{MUV3}-z_{Xprod}^{MCTrue}}{\beta\gamma c\tau}\right)} = e^{-\left(\frac{z_{MUV3}-z_{Xprod}^{MCTrue}}{\frac{p_X}{M_X}c\tau}\right)}. \tag{6.2.7}$$

This approach follows that adopted for the E787/E949 experimental results [113] and, despite very different experimental set-ups (stopped K^+ decay at E787/E949 [245] versus highly boosted K^+ decay in flight at NA62 [200]), it is found that the lifetimes both experiments are sensitive to is similar. This is due to a similar ratio of Lorentz factor to decay length for a potential X particle. Seven X lifetimes are investigated explicitly: $\tau = \infty, 5, 2, 1, 0.5, 0.2, 0.1$ ns, with greatest sensitivity for larger τ.

6.3 Studies and Results

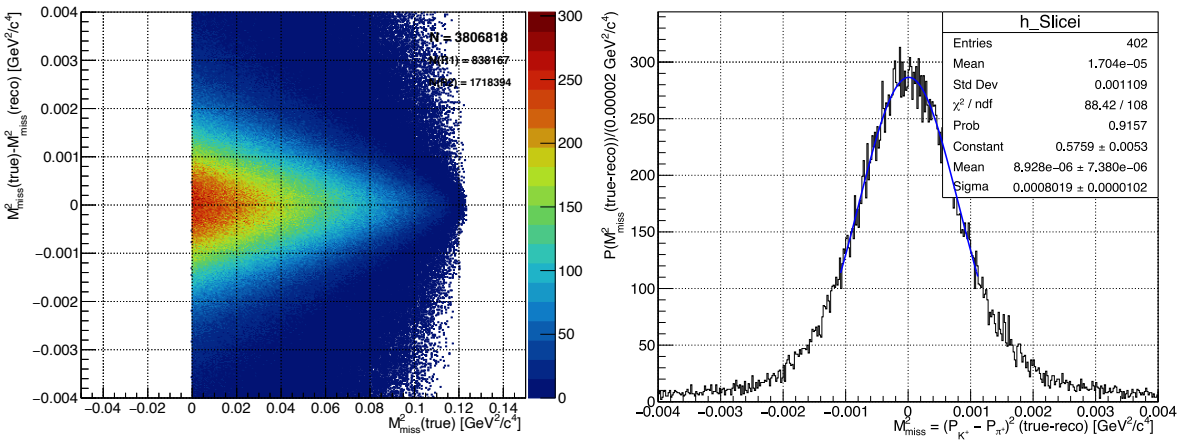
6.3.1 Squared Missing Mass Resolution

6.3.1.1 m^2_{miss} Resolution Measurements

The m^2_{miss} resolution can be measured directly using simulations by comparing the 'truth' generated value $m^2_{miss}|_{true}$ to the reconstructed value $m^2_{miss}|_{reco}$. A two dimensional histogram of the difference $\Delta m^2 = m^2_{miss}|_{true} - m^2_{miss}|_{reco}$ versus $m^2_{miss}|_{true}$ is filled for simulated events, see figure 6.3a. For each $m^2_{miss}|_{true}$ bin a projection onto the y (Δm) axis is performed, see figure 6.3b. A Gaussian fit to this projection provides a model for the distribution and its standard deviation is used as a measurement of the resolution $\sigma(m^2_{miss}|_{true})$, with uncertainty taken to be equal to the fitting uncertainty. Plotting the mean of the Gaussian fits, see figure 6.4, provides a test of calibration - showing how closely the reconstructed value matches the 'true' generated value. The deviations observed at the level of $\leq 0.02 \times 10^{-3}$ demonstrates a good calibration. The resolution as a function of m^2_{miss} is shown in figure 6.5a and is modelled by a fourth order polynomial function. The resolution on the m^2_{miss} observable corresponding to mass hypothesis M_X is shown in figure 6.5b. This study was performed using simulations of $K^+ \to \pi^+\nu\bar{\nu}$ decays for which a large continuous range of m^2_{miss} can be studied with high statistics. Equivalent studies for simulated $K^+ \to \pi^+ X$ decays give consistent results at $m^2_{miss}|_{true} = (M^i_X)^2$ values.

6.3.2 m^2_{miss} Resolution Differences Between Simulations and Data

Modelling inaccuracies in the simulations lead to differences between the m^2_{miss} resolution measured in data and simulations. Results of studies of the squared missing mass reso-



(a) (b)

Figure 6.3: (a) Two dimensional histogram plotting the difference between the true and reconstructed m^2_{miss} on the vertical axis against the true m^2_{miss} on the horizontal axis. To determine the resolution at a given true m^2_{miss} a slice (e.g. one m^2_{miss}(true) bin) is taken and projected onto the y-axis. An example is shown by (b) for a slice with $0.05400 < m^2_{miss}(\text{true}) < 0.05450\,\text{GeV}/c^2$ with associated Gaussian fit.

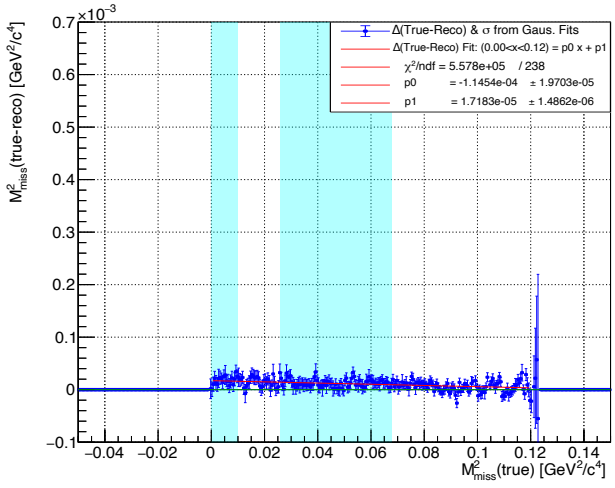
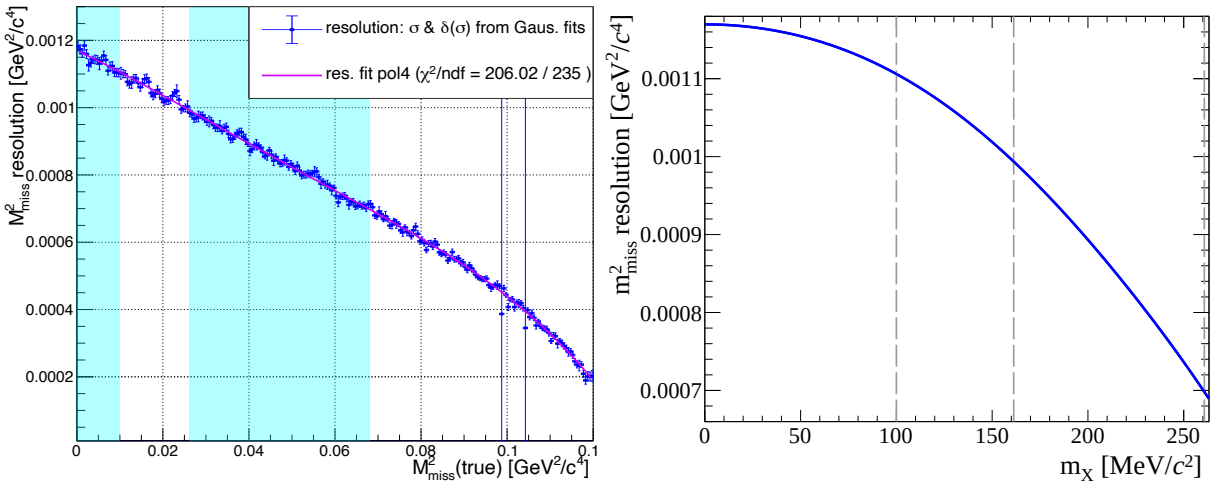


Figure 6.4: Measured mean of Gaussian fits to $m^2_{miss}(\text{true}) - m^2_{miss}(\text{reco})$ projections at different values of $m^2_{miss}(\text{true})$. Differences from zero suggest residual miscalculation.



(a) (b)

Figure 6.5: Resolution on the (reconstructed) m^2_{miss} variable as a function of: (a) true m^2_{miss} from σ of Gaussian fits with uncertainties combining statistical and fit uncertainties with model shown in magenta and (b) the assumed m_X. The signal regions are shaded in cyan in (a) with boundaries shown in (b) by vertical dashed lines.

lution in control samples selected for data and simulations for different decays, relevant to different squared missing mass regions, are reported in table 6.1. A consistent picture emerges in which the resolution in simulations is better than for the data with

$$R_{res} = \frac{\sigma(m^2_{miss})_{Data}}{\sigma(m^2_{miss})_{Sim}} = 1.04\,. \tag{6.3.1}$$

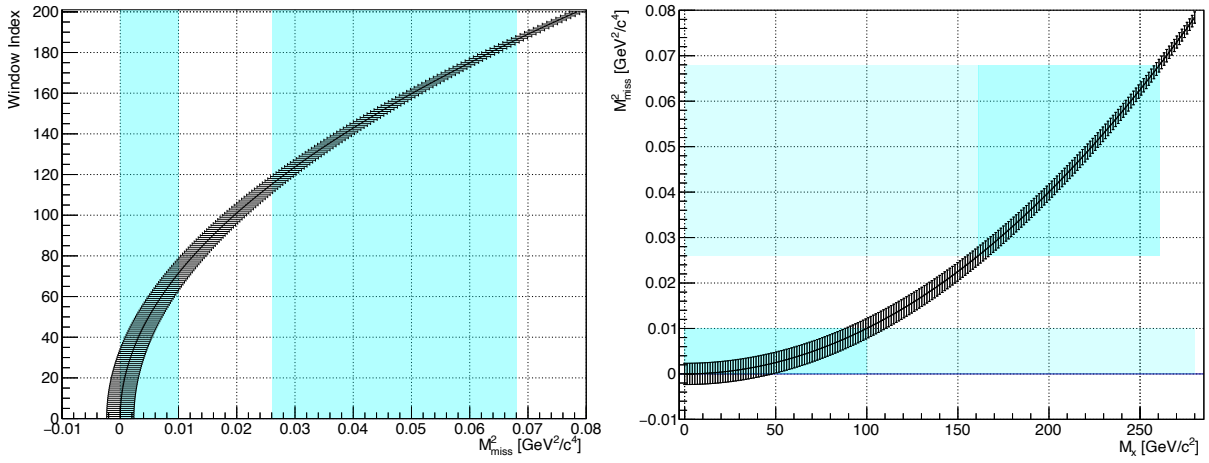
Uncertainties arising from fitting procedures mostly cancel in the ratio, however R_{res} is sensitive to changes in calibration. Conservatively this factor is used to apply a correction to the acceptance accounting for the fact that the window widths are defined to be $\pm 2\sigma(m^2_{miss})_{sim}$ which is equivalent to $\frac{\pm 2\sigma(m^2_{miss})_{Data}}{R_{res}} \approx \pm 1.923\sigma(m^2_{miss})_{Data}$.

Table 6.1: Ratio of measured squared missing mass resolution for 2017 minimum bias data with respect to simulations, $R_{res} = \frac{\sigma(m^2_{miss})_{Data}}{\sigma(m^2_{miss})_{sim}}$

Case	R_{res}
$K^+ \to e^+\nu_e$ at $m^2_{\nu_e} = 0\,\mathrm{GeV}^2/c^4$	1.049
$K^+ \to \mu^+\nu_\mu$ at $m^2_{\nu_\mu} = 0\,\mathrm{GeV}^2/c^4$	1.030
$K^+ \to \pi^+\pi^0$ at $m^2_{\pi^0} \approx 0.018$	1.046
$K^+ \to \pi^+\pi^+\pi^-$ (proxy for $0.085 < m^2_{miss} < 0.1225\,\mathrm{GeV}^2/c^4$)	1.032

6.3.3 Search Window Definitions

Using the procedure described in section 6.2.1, search windows are defined centred on $m^2_{miss}|^i_{reco} = M^i_X$ with width equal to four times the resolution at that point, accounting for the correction from section 6.3.2. The search window definitions are illustrated in figure 6.6.



(a) (b)

Figure 6.6: Graphic depiction of the window definitions in the m^2_{miss} observable (a) and versus the mass hypothesis m_X (b). Each point (connected by a line) shows the centre of the associated window and the lower and upper bars correspond to the lower and upper bounds of the window. Blue shaded areas correspond to the signal regions.

6.3.4 Signal Acceptance and Single Event Sensitivity

The acceptance and single event sensitivity is determined for each mass hypothesis as described in section 6.2.3 according to equation 6.2.1. The factor f_i, the fraction of events in window i, is shown in figure 6.7. The difference in measured squared missing mass resolution between data and MC is accounted for in the analysis by applying a correction, C_{res} to this factor

$$f'_i = f_i C^i_{res} = f_i \frac{\int_{\mu-\frac{2\sigma}{R_{res}}}^{\mu+\frac{2\sigma}{R_{res}}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)}{2\sigma^2}} dy}{\int_{\mu-2\sigma}^{\mu+2\sigma} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)}{2\sigma^2}} dy} , \tag{6.3.2}$$

where $\mu = \left(M^i_X\right)^2$ and $\sigma = \sigma\left(\left(M^i_X\right)^2\right)_{Data}$. It can be shown that all dependence on μ and σ cancel and

$$C^i_{res} = C_{res} = \frac{\mathrm{erf}\left(\frac{\sqrt{2}}{R_{res}}\right)}{\mathrm{erf}\left(\sqrt{2}\right)} = 0.9906 \ . \tag{6.3.3}$$

The measured acceptance and corresponding single event sensitivity are shown in figure 6.8 (and with statistical fluctuations smoothed out in figure 6.9) for the case of stable or invisibly decaying X, and for a range of lifetimes in figure 6.10. The $\mathcal{B}_{SES}$ results displayed include the correction described in section 6.3.5.

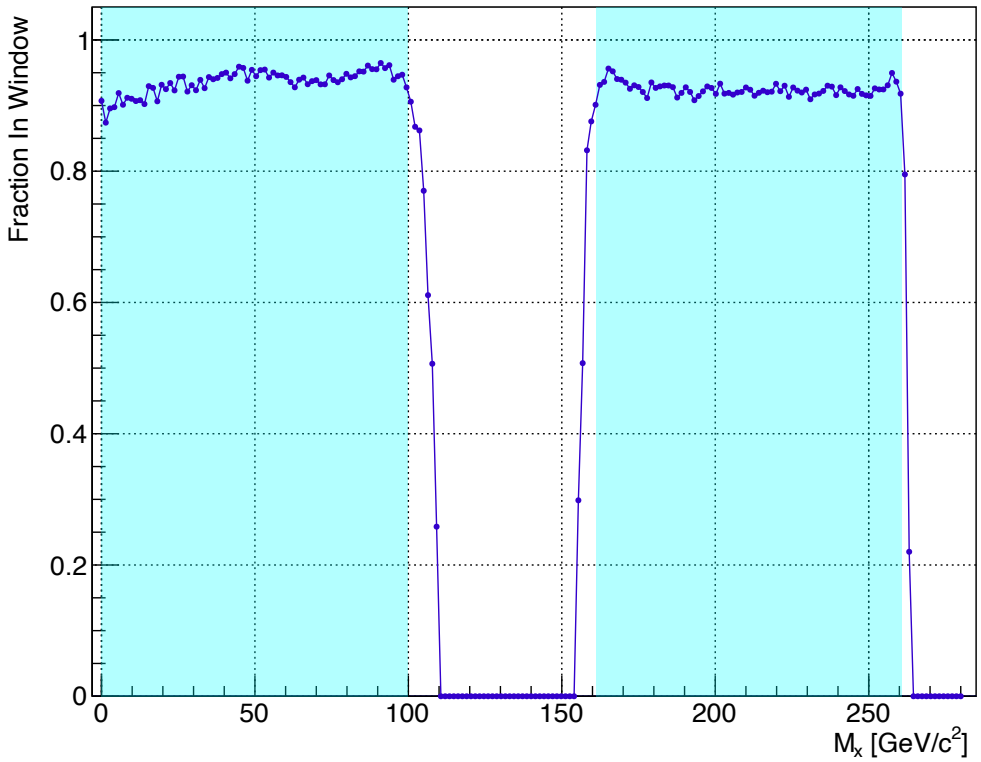
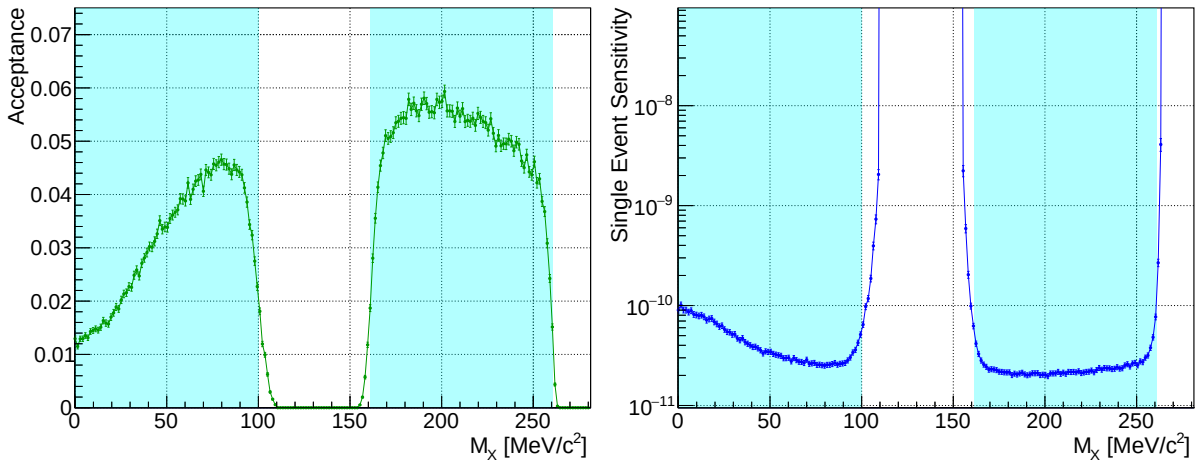
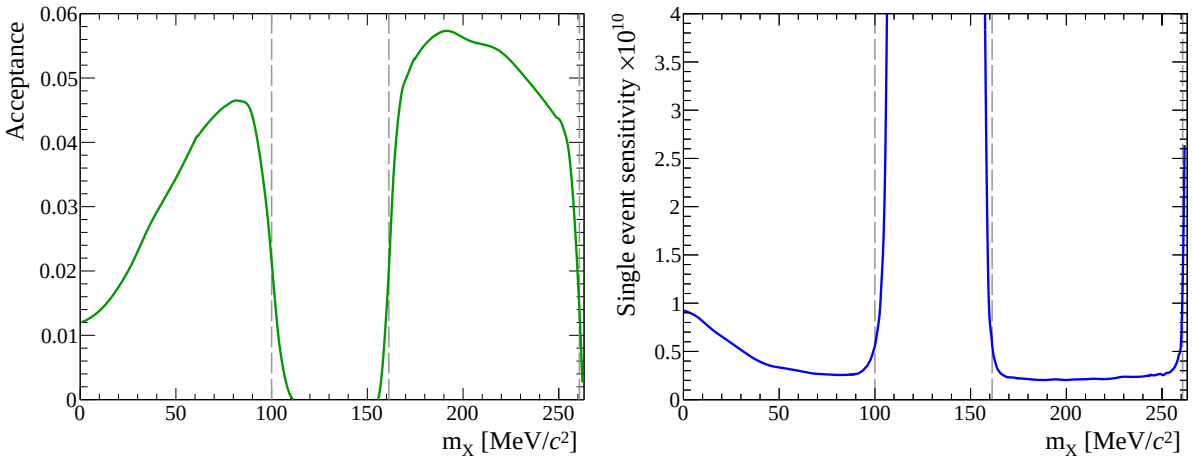


Figure 6.7: Corrected fraction of events contained in the i^{th} window, f_i' (equation 6.3.2). The fraction of events contained inside a 4σ wide ($\pm 2\sigma$) window is expected to be 95% on average but close to the hard cuts at the edges of signal regions 1 and 2 the fraction drops reaching exactly 0 once no part of the window overlaps either signal region (equation 6.2.3). A further correction (amounting to a reduction by $\sim 1\%$, see equations 6.3.2 and 6.3.3) is applied to account for the difference in M^2_{miss} resolution between data and MC.



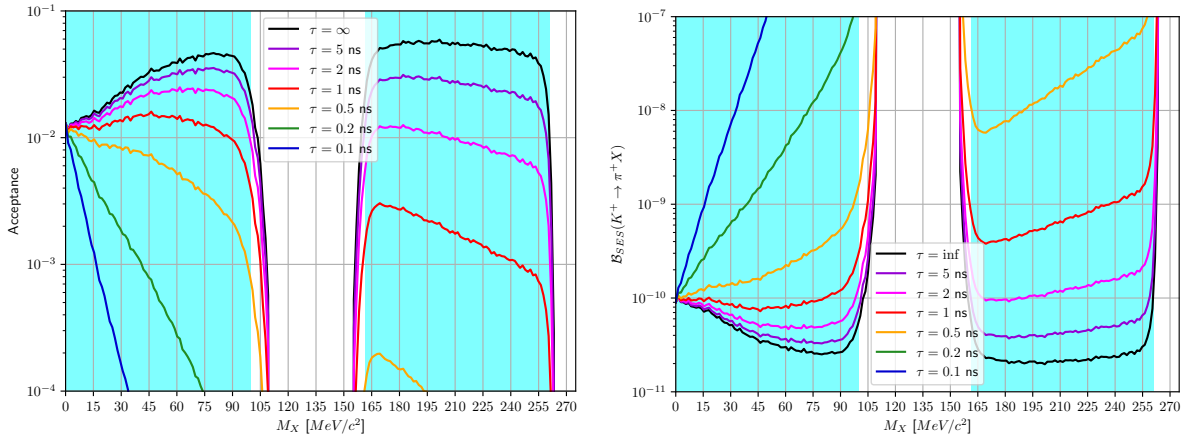
(a) (b)

Figure 6.8: Signal acceptance (a) and single event sensitivity (b) for $K^+ \to \pi^+ X$ decays with stable X.



(a) (b)

Figure 6.9: Signal acceptance (a) and single event sensitivity (b) for $K^+ \to \pi^+ X$ decays with stable X, displayed as a smooth function.



(a) (b)

Figure 6.10: Signal acceptance (a) and single event sensitivity (b) for $K^+ \to \pi^+ X$ decays with X of lifetime τ_X, under the assumption described in section 6.2.4.

6.3.5 Momentum Dependence $\mathcal{B}_{SES}$ Correction

Differences in momentum spectra of $K^+ \to \pi^+\nu\bar{\nu}$ and $K^+ \to \pi^+ X$, which also change with X mass, are accounted for by modifying equation 6.2.4 to become

$$\begin{aligned}&\mathcal{B}_{SES}(\pi X, M_X^i, p_1 < p_{\pi^+} < p_2) \\ &\quad = \mathcal{B}_{SES}(\pi\nu\bar{\nu}, R1+R2, p_1 < p_{\pi^+} < p_2)\frac{A(\pi\nu\bar{\nu}, R1+R2, p_1 < p_{\pi^+} < p_2)}{A(\pi X, M_X^i, p_1 < p_{\pi^+} < p_2)}\,,\end{aligned} \tag{6.3.4}$$

where four momentum bins are considered of width 5 GeV/c starting at $p_1 = 15, 20, 25, 30$ GeV/c. The results from the $\pi\nu\bar{\nu}_{2017}$ analysis needed for calculations are given in table 6.2. Figure 6.11 displays the acceptance and $\mathcal{B}_{SES}$ for each mass hypothesis for each momentum bin individually, a result using simple integration over the full range 15–35 GeV/c and equation 6.2.4 (labelled int.) and the result combining the results and accounting for momentum dependence according to equation 6.3.4 (labelled combi.). The acceptance adds directly

$$A(\pi\nu\bar{\nu}, R1+R2, 15 < p_{\pi^+} < 35) = A_1 + A_2 + A_3 + A_4\,, \tag{6.3.5}$$

where A_i is the acceptance for momentum bin i, so there is no difference between the combined and integrated cases. However the combined $\mathcal{B}_{SES}$ is calculated according to

$$\mathcal{B}_{SES}(\pi\nu\bar{\nu}, R1+R2, 15 < p_{\pi^+} < 35) = \frac{1}{\frac{1}{\mathcal{B}^1_{SES}} + \frac{1}{\mathcal{B}^2_{SES}} + \frac{1}{\mathcal{B}^3_{SES}} + \frac{1}{\mathcal{B}^4_{SES}}}\,, \tag{6.3.6}$$

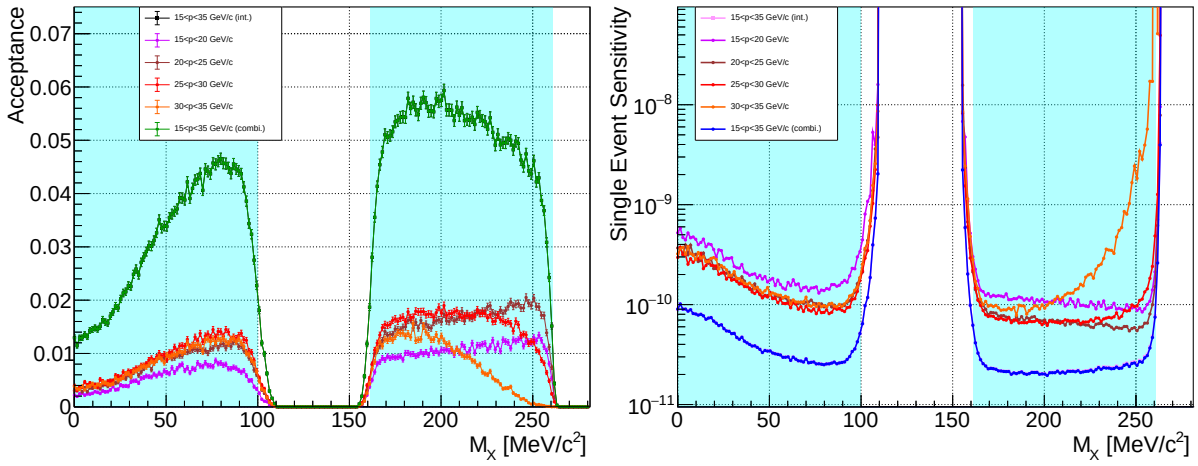
with uncertainty

$$\sigma_s = \mathcal{B}^2_{SES}\sqrt{\frac{\sigma_1^2}{(\mathcal{B}^1_{SES})^4} + \frac{\sigma_2^2}{(\mathcal{B}^2_{SES})^4} + \frac{\sigma_3^2}{(\mathcal{B}^3_{SES})^4} + \frac{\sigma_4^2}{(\mathcal{B}^4_{SES})^4}}\,, \tag{6.3.7}$$

where $\mathcal{B}^i_{SES}$ and σ_i are the $\mathcal{B}_{SES}$ and its associated uncertainty for momentum bin i. In this case a small effect is visible with the combined case, correctly accounting for momentum dependence, giving marginally lower (improved) $\mathcal{B}_{SES}$ (visible in the upper third of signal region 2). The patterns observed in figure 6.11 can be understood by considering the kinematics of the situation. The pion momentum $p_{\pi^+} = p_{K^+} - p_X = p_{K^+} - \sqrt{E_X^2 - M_X^2}$ is limited to lower values for higher M_X. This leads to the reduced acceptance at higher M_X for higher momentum bins (significantly for final bin $30 < p_{\pi^+} < 35$ GeV/c) and correspondingly degraded sensitivity.

Table 6.2: Single event sensitivity and acceptance for $K^+ \to \pi^+\nu^+\overline{\nu}$ from $\pi\nu\bar{\nu}_{2017}$ analysis.

Momentum Range [GeV/c]	$A(\pi\nu\bar{\nu}, R1+R2, p_1 < p_{\pi^+} < p_2)$	$\mathcal{B}_{SES}(\pi\nu\bar{\nu}, R1+R2, p_1 < p_{\pi^+} < p_2)$
$15 < p_{\pi^+} < 20$	$(5.975 \pm 0.080) \times 10^{-3}$	$(1.90 \pm 0.20) \times 10^{-10}$
$20 < p_{\pi^+} < 25$	$(9.081 \pm 0.099) \times 10^{-3}$	$(1.23 \pm 0.13) \times 10^{-10}$
$25 < p_{\pi^+} < 30$	$(9.070 \pm 0.099) \times 10^{-3}$	$(1.30 \pm 0.14) \times 10^{-10}$
$30 < p_{\pi^+} < 35$	$(5.600 \pm 0.078) \times 10^{-3}$	$(2.18 \pm 0.23) \times 10^{-10}$
$15 < p_{\pi^+} < 35$	0.02973 ± 0.00018	$(2.89 \pm 0.21) \times 10^{-11}$



(a) (b)

Figure 6.11: Signal acceptance (a) and single event sensitivity (b) for $K^+ \to \pi^+ X$ decays with stable X for four momentum bins, an integrated result over the full momentum range (int.) and results combined accounting for momentum dependence (combi.). For the acceptance the combined and integrated cases (black squared and green dots) coincide since acceptances simply add. For the $\mathcal{B}_{SES}$ the integrated case (pale magenta) is slightly higher (corresponding to worse sensitivity) than the combined case (blue), this is visible in the upper third of signal region 2.

6.3.6 Background Expectation

The complete set of background studies for the $\pi\nu\bar{\nu}_{2017}$ analysis, see chapter 5, are directly applicable to this analysis which has the same backgrounds plus the $K^+ \to \pi^+\nu\bar{\nu}$ decay itself. For the $\pi\nu\bar{\nu}_{2017}$ analysis a counting experiment methodology was adopted for the statistical treatment used to set (one or two-sided) limits on the branching ratio, see section 5.6. This means only the integral number of events expected in the signal regions is required. However, of critical importance for this analysis is the distribution of these backgrounds as a function of the m^2_{miss} observable since an individual background estimate is required for each search window. Using a combination of simulations and data-driven procedures these distributions were derived and are presented in figure 6.12 along with the observed distribution of 2017 data.

Using the components of this stacked histogram models are constructed as a function of m^2_{miss} to describe each background component individually. Based on these models the number of expected background events in any given range of m^2_{miss} can be calculated. The total background for a given search window is then established by summing these components for the relevant m^2_{miss} range. This procedure is used for all K^+ decay backgrounds with an additional 20% systematic uncertainty applied, accounting for the range of results obtained for individual search windows when varying model parameters. Specific additional studies were performed for the $K^+ \to \pi^+\nu\bar{\nu}$ and upstream backgrounds.

For this analysis the SM description of the $K^+ \to \pi^+\nu\bar{\nu}$ decay is assumed. Using the same simulations as used in figure 6.12 explicit predictions for the number of $K^+ \to \pi^+\nu\bar{\nu}$ events expected in each search window are derived as shown in figure 6.13.

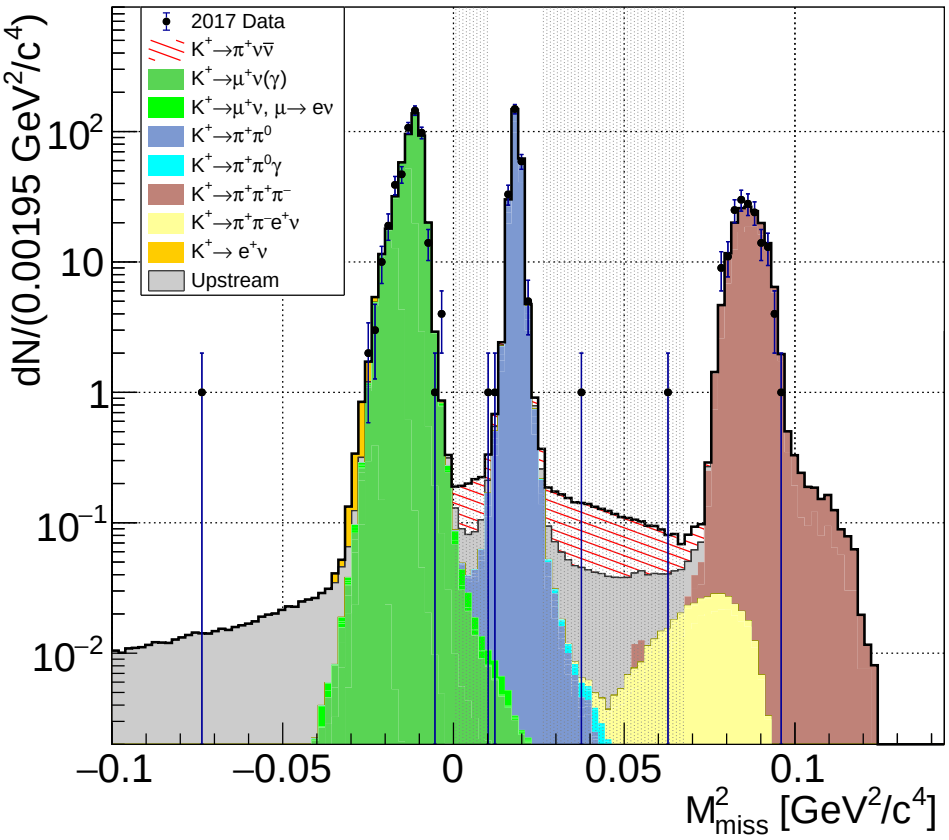


Figure 6.12: Observed data and predicted background distributions as a function of m^2_{miss} for the $\pi\nu\bar{\nu}_{2017}$ analysis.

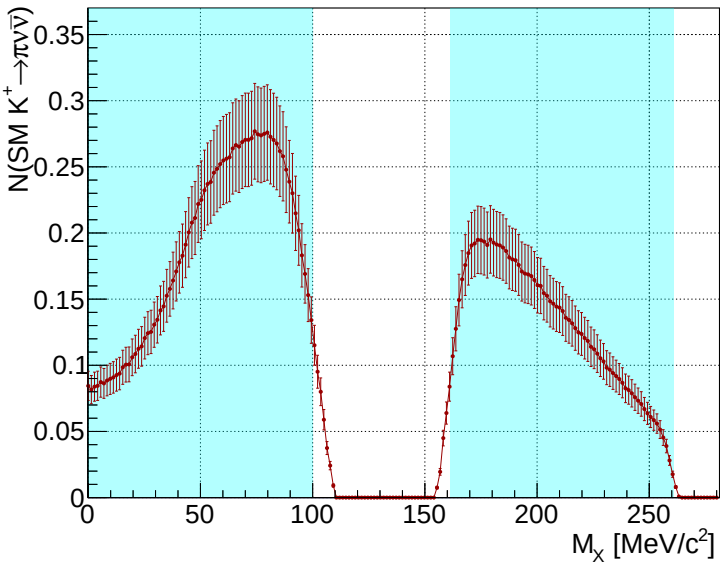
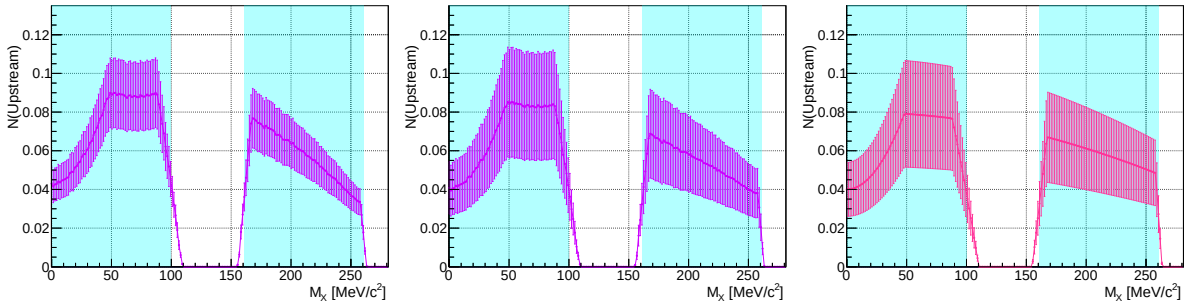


Figure 6.13: The number of expected SM $K^+ \to \pi^+\nu\bar{\nu}$ events in each search window, with results shown at the corresponding mass hypothesis M_X^i.



(a) (b) (c)

Figure 6.14: Upstream background predictions for each search window displayed as a function of the corresponding mass hypothesis m_X. The models used are: (a) data-driven from bifurcation sample C data, (b) simulation-driven model from OutOfBox bifurcation samples C+A and (c) flat-in-m^2_{miss} model.

For the upstream background three different models have been studied:

1. Data-driven model constructed from information from main $\pi\nu\bar{\nu}_{2017}$ analysis (i.e. shape as in figure 6.12, uses shape of bifurcation sample C from data), see figure 6.14a.

2. A fully simulation-driven model constructed from a fit to the m^2_{miss} distribution for selected MC events from upstream simulations (using 'OutOfBox' bifurcation samples A+C), see figure 6.14b.

3. Approximation assuming the upstream background is flat in the m^2_{miss} variable with integral in R1+R2 equal to that from the $\pi\nu\bar{\nu}_{2017}$ analysis (table 5.5), see figure 6.14c.

The three different models are seen to produce compatible results and share a similar shape. The data-driven model is used for the final results.

The total background predictions and a breakdown of all components is shown in figure 6.15 for each search window.

6.3.7 Setting upper limits on $\mathcal{B}(K^+ \to \pi^+ X)$

Upper limits are established on the branching ratio $\mathcal{B}(K^+ \to \pi^+ X)$ at 90% confidence level using the CL_S method with a log likelihood test statistic and performing counting experiments in each search window. A standard implementation of the CL_S method from the NA62 software framework is utilised assuming the total background estimate uncertainty can be treated as Gaussian, with an additional 20% systematic uncertainty assigned to background estimates obtained from models derived from the m^2_{miss} distributions, see figure 6.12. When using the CL_S method the background prediction can only change the limit obtained if the number of observed events is non-zero. The expected upper limit and one and two standard deviation bands were calculated before unblinding by carrying out an ensemble of pseudoexperiments. Results of these pseudoexperiments are arranged

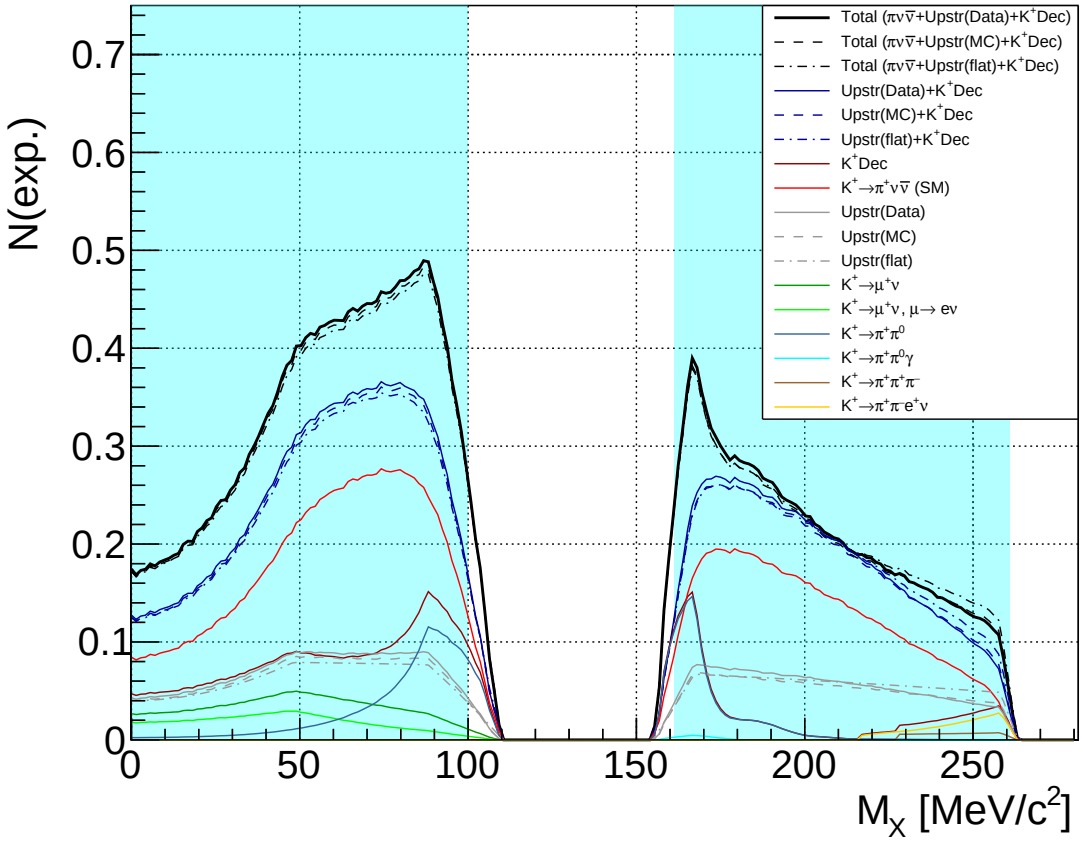


Figure 6.15: Background prediction in each search window as a function of the corresponding mass hypothesis, including the total contribution from all K^+ decays (except $K^+ \to \pi^+\nu\bar{\nu}$), the three upstream background models and SM $K^+ \to \pi^+\nu\bar{\nu}$ separately (the data-driven model is used for results).

in size order and elements of the list corresponding to the median and upper and lower 68 and 95 percentiles set the expected limit and upper and lower one and two standard deviation results respectively. In cases with low statistics (small number of expected and observed events) many pseudoexperiments have identical results and therefore the one and two standard deviation results may coincide.

When the signal regions were unblinded two candidate $K^+ \to \pi^+ X$ events were found with squared missing mass (and associated square root for comparison to M_X) of

$$\text{Event 1: } m^2_{miss} = 0.0383\,\text{GeV}^2/c^4 \;\;,\;\; \sqrt{m^2_{miss}} = 195\,\text{MeV}/c^2\,, \tag{6.3.8}$$

$$\text{Event 2: } m^2_{miss} = 0.0635\,\text{GeV}^2/c^4 \;\;,\;\; \sqrt{m^2_{miss}} = 253\,\text{MeV}/c^2\,. \tag{6.3.9}$$

The observed and expected upper limits for each mass hypothesis are shown in figure 6.16 for stable X and observed upper limits in figure 6.17 with comparison to previous experimental results from E787/E949 [113].

6.3.8 Interpretation in Terms of BSM Scenarios

The upper limits established on the branching ratio can be translated into exclusion limits in parameter spaces of BSM scenarios. Three specific cases were considered and are described in section 2.4.

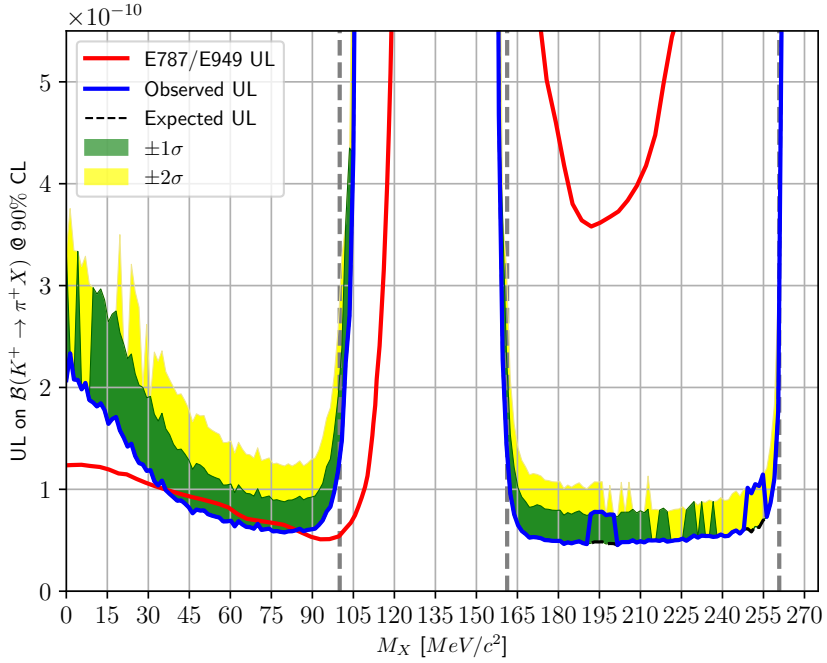


Figure 6.16: Observed (blue) and expected (dashed black) upper limits on the branching ratio $\mathcal{B}(K^+ \to \pi^+ X)$ at 90% confidence level, with one and two standard deviations indicated by green and yellow regions respectively, as a function of X mass hypothesis. For comparison, observed upper limits from the BNL E787/E949 experiment [113] are shown in red.

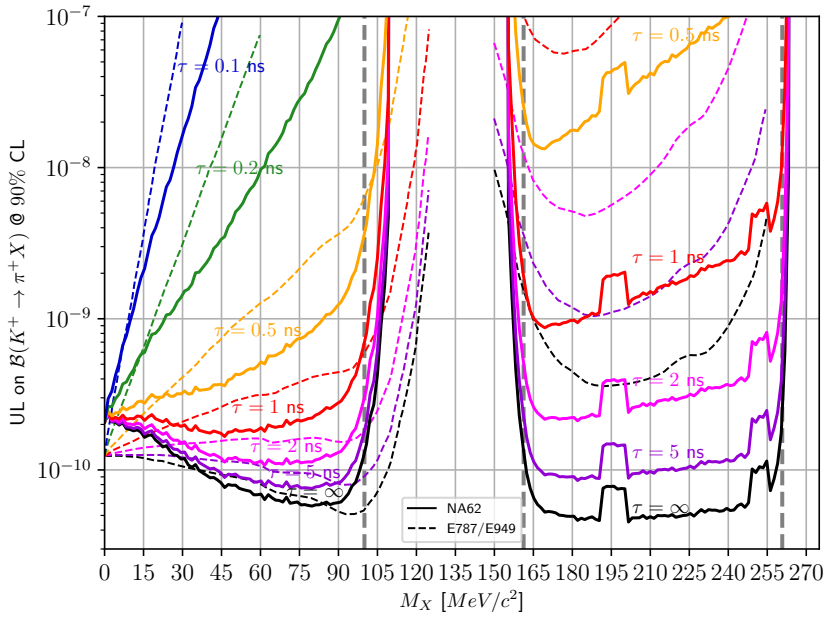


Figure 6.17: Observed upper limits on $\mathcal{B}(K^+ \to \pi^+ X)$ at 90% confidence level as a function of X mass hypothesis and for a set of lifetimes, τ_X for NA62 (this analysis, solid lines) and the BNL E787/E949 experiment [113] (dashed lines). These are model-independent limits set accounting for experimental sensitivity to decays of X with mass M_X lifetime τ_X.

Interpreting X as a dark scalar mixing with the Higgs boson upper limits on the mixing strength, expressed as $\sin^2\theta$, are calculated according to equation 2.4.11 with results shown in figure 6.18. These results may be compared to the other constraints on this model reported in [54] (and therein called BC4) as shown in figure 6.19. If X is considered to be a pseudoscalar dark matter mediator upper limits on its coupling strength in the three scenarios described in section 2.4.2 are displayed in figure 6.20a for stable X. In the BC10 model of reference [54] (equivalent to the Yukawa-like coupling case of section 2.4.2 up to a numerical constant) X is an ALP and corresponding limits are shown in figure 6.20b and compared to other constraints on this model reported in [54] in figure 6.21.

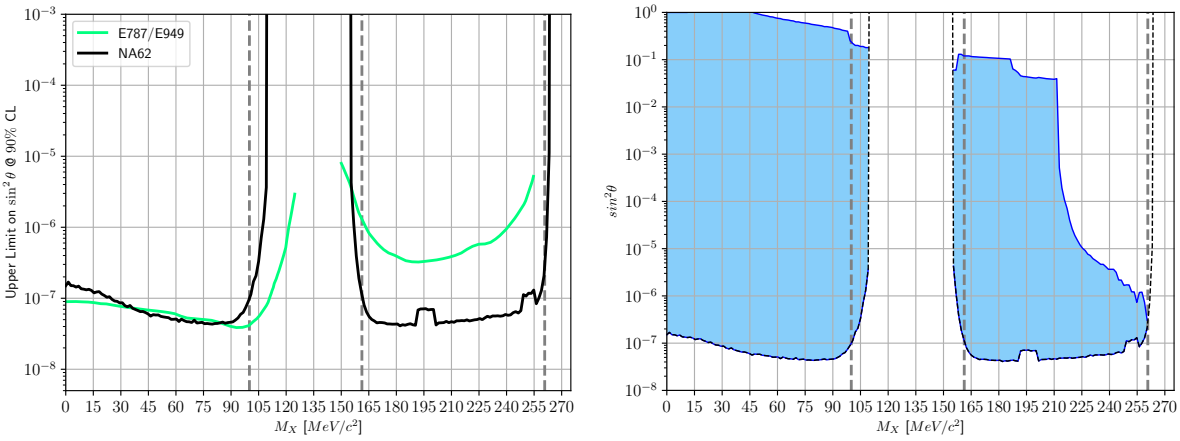
As described in section 2.4, if X decays predominantly to invisible particles, for example dark matter, bounds on the coupling parameter ($\sin^2\theta$ or g_Y for the scalar and pseudoscalar/ALP models respectively) are directly derived from its relationship with the branching ratio, with results shown by the dashed back lines in figures 6.18b and 6.20b. If X decays predominantly to visible SM particles τ_X is inversely proportional to the coupling parameters [164] [165] (see equations 2.4.13 and 2.4.22), limiting the reach of this analysis for large coupling because of lower acceptance for shorter lifetimes. The $X \to e^+e^-$ decays dominate the visible decay width up to the di-muon threshold beyond which the additional $x \to \mu^+\mu^-$ channel opens (see figure 2.20a) and τ_X decreases, limiting the sensitivity of this search. The model-dependent relationship between the lifetime and coupling therefore determines the shape of the exclusion regions shown in blue in figures 6.18b and 6.20b. The global picture of constraints on the coupling parameter versus m_X phase-space are shown in figures 6.19 and 6.21 showing the cases of X decaying predominantly to (a) invisible particles and (b) visible particles.

Alternatively, identifying X with mass experimentally consistent with zero with an Axiflavon, an upper limit on the Axiflavon mass can be derived as described in section 2.4.3 with results shown in figure 6.22.

6.3.9 Discussion

Two candidate $K^+ \to \pi^+X$ events are observed [125] in agreement with the background expectations, and therefore upper limits are established on $\mathcal{B}(K^+ \to \pi^+X)$ at 90% CL for each X mass hypothesis. Two perturbations as a function of M_X in the observed upper limits are formed around the two $K^+ \to \pi^+X$ candidates at reconstructed m_{miss} values of 196 and 252 MeV/c^2. Sharp boundaries of these perturbations are a result of the search windows procedure with the jump occurring between consecutive windows with 0 and 1 events. Consideration of the signal shape tends to smooth these out, see section 6.4.

Results for the lowest mass hypotheses from this measurement do not improve on E787/E949 results, including the interesting case where M_X is experimentally consistent with zero which is relevant to the interpretation of X as a QCD axion, for example an Axiflavon, see figure 6.22. Combined effects of kinematic rejection requirements lead to a decrease in acceptance for $K^+ \to \pi^+X$ events at $M_X = 0$ as seen in figure 6.9. A dedicated analysis may be able to improve upon this by extending the signal region below



(a) (b)

Figure 6.18: Exclusion limits for the scenario of X being a scalar mixing with the Higgs boson with production and decay determined by mixing parameter θ (BC4 model reported in [54]). The black line (solid and dashed in (a) and (b) respectively) represents the upper limit at 90% confidence level on $\sin^2\theta$ for each X mass hypothesis, and is relevant to the case where X is stable or decays to invisible particles (e.g. dark matter). This is calculated using equation 2.4.10 following reference [164] and results from similar treatment for the previous result from the E747/E949 experiment [113] is shown for comparison in (a). If X decays predominantly to visible particles the model-dependent [164] lifetime, given by equation 2.4.13, is considered along with the experimental sensitivity as a function of lifetime (see figure 6.10a) and the excluded region (b, blue) includes an upper bound.

$m^2_{miss} = 0$, however here the $K^+ \to \mu^+\nu_\mu$ background increases and the extension is limited by the need for control and blinded regions in the analysis.

Under the assumption of stable X these new upper limits improve over those reported by the E787/E949 experiments by a factor of $\mathcal{O}(10)$ in signal region 2, and are competitive in region 1. For unstable X, assuming that it predominantly decays to visible SM particles, the same pattern holds in general. However, in region 1 the new limits improve over those from E787/E949 across an increasingly large range of mass hypotheses as the assumed lifetime becomes shorter. Despite differences in experimental set-up between E787/E949 (stopped K^+ decay-at-rest) and NA62 (highly boosted K^+ decay-in-flight), the two results exhibit similar dependence on τ_X. This is due to a similar ratio of Lorentz factor to decay length in experimental apparatus for an X particle.

6.4 Setting upper limits on $\mathcal{B}(K^+ \to \pi^+ X)$ Using a Shape Analysis

6.4.1 The Shape Analysis Procedure and Results

Full consideration of the shape (distribution in observable m^2_{miss}) of the signal and background allows additional discrimination power and can improve sensitivity. The search

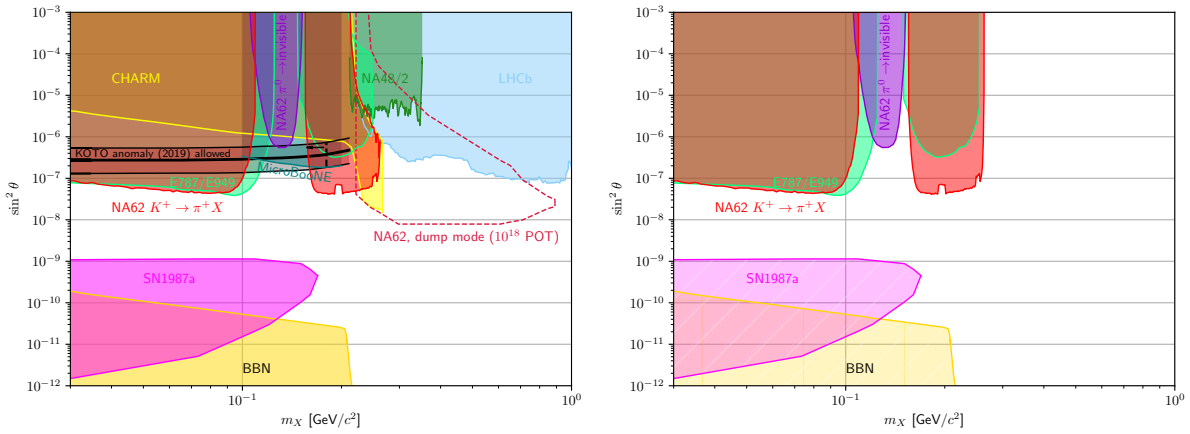


Figure 6.19: Global picture for excluded regions in the $\sin^2\theta$ versus m_X parameter-space of the BC4 model of reference [54] for the cases of X decays to visible (a) and invisible (b) final states. Results from this work are shown in red with bounds from a recent NA62 search for $\pi^0 \to invisible$ [207] shown in purple while projected NA62 exclusion from 'dump mode' operation with 10^{18} POT ($\sim$ 90 days of data-taking, expected in the next data-taking period from 2021 [228]) [54] is shown by dashed red lines. The KOTO anomaly (section 2.3.4) 2019 allowed values are shown by a thick black line with $\pm 1\sigma$ band contained between thinner black lines, with full explanation for KOTO's observed events below the vertical dashed line, as indicated by the arrow [246]. This is excluded at 95% confidence level by recent MicroBooNE results [247] shown in teal. Reinterpretation of results from the CHARM experiment [248] [160] give excluded region shown in yellow, while studies of $K \to \pi\mu\mu$ from NA48/2 [170] and $B \to K\mu\mu$ from LHCb [249] [250] are shown in green and pale blue respectively. Constraints from supernova 1987a [251] [252] [253] [254] and Big Bang Nucleosynthesis [54] are also shown, hatching in (b) indicates that the exact shape of the bounds is unclear (no detailed treatment available in the literature) for invisible decays, however if X is sufficiently short-lived these bounds can be evaded.

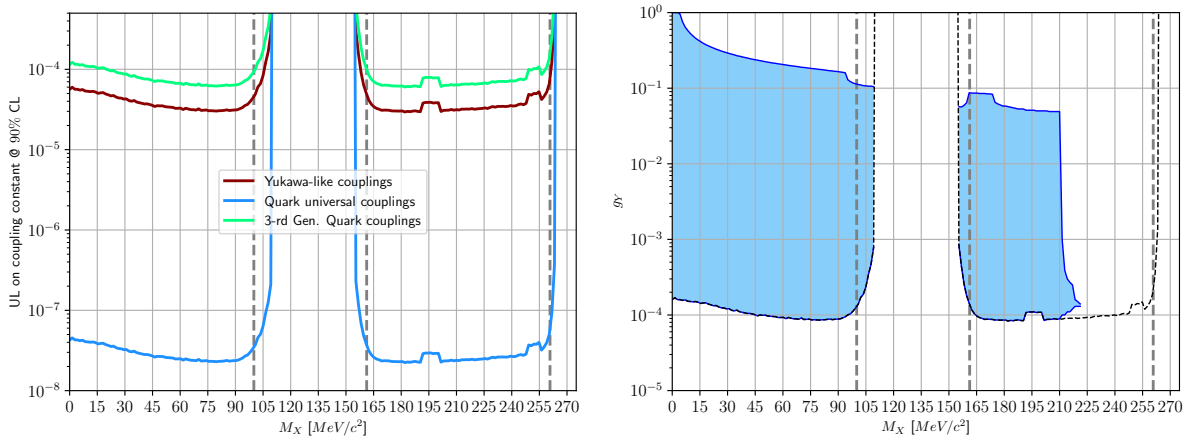


Figure 6.20: (a) Upper limits on coupling strength for the scenario of X being a pseudoscalar mediator for dark matter under three cases considered in section 2.4.2. The case of Yukawa-like couplings is equivalent (up to a factor $2\sqrt{2}$) to the BC10 ALP model of reference [54] and the corresponding upper limit on coupling strength g_Y is shown by the black dashed line in (b). Considering the model-dependent lifetime of X [165], expressed by equation 2.4.22, and experimental sensitivity as a function of lifetime (figure 6.10a), upper boundaries on the exclusion are derived and the excluded region is shown in blue.

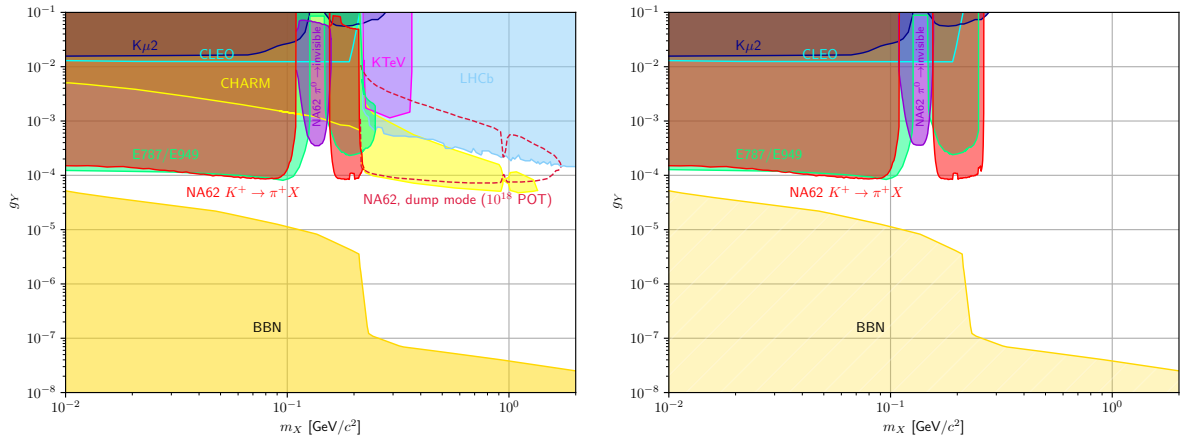


Figure 6.21: Global picture for excluded regions in the g_Y versus m_X parameter-space of the BC10 model of reference [54] for the cases of X decays to visible (a) and invisible (b) final states. Results from this work are shown in red with bounds from a recent NA62 search for $\pi^0 \to invisible$ [207] shown in purple while projected NA62 exclusion from 'dump mode' operation with 10^{18} POT ($\sim$ 90 days of data-taking, expected in the next data-taking period from 2021 [228]) [255] is shown by dashed red lines. Older upper limits on the $K^+ \to \pi^+ X$ branching ratio from the $K\mu 2$ experiment [256] are interpreted in the same way as for this result [165] and lead to excluded region shown in dark blue. Reinterpretation of results from the CHARM experiment [248] [255] give excluded region shown in yellow, while exclusion from CLEO searches for $B \to K + invisible$ [257] [165] are shown in cyan. Interpretation [165] of KTeV searches for $K_L \to \pi^0 \mu\mu$ [258] [259]and of LHCb searches for $B \to K\mu\mu$ [249] [250] result in exclusion shown in magenta and pale blue respectively. Constraints from Big Bang Nucleosynthesis [54] are also shown, hatching in (b) indicates that the exact shape of the bound is unclear (no detailed treatment available in the literature) for invisible decays and can be evaded for sufficiently short-lived X.

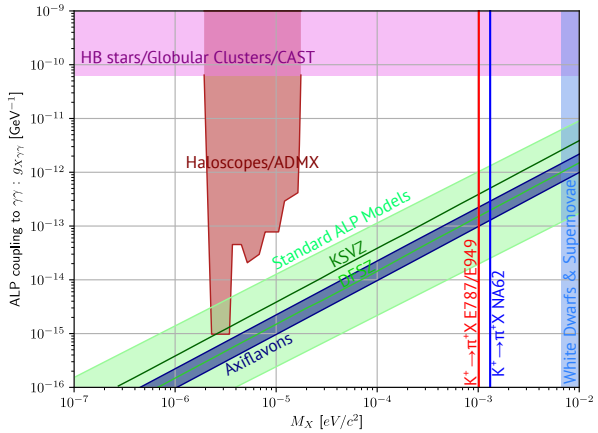


Figure 6.22: Constraints on ALP X in the phasespace $g_{X\gamma\gamma}, M_X$. Constraints derived from astronomical observations are shown along with Haloscopes/ADMX results as reported in [147]. The light green band contains standard ALP models with benchmark KSVZ and DFSZ models marked (as shown in [13]) and the Axiflavon model occupies the dark blue band. Upper limits on axiflavon mass, M_X, are derived using equation 2.4.26 (see section 2.4.3) using E787/E949 [113] and NA62 upper limits on $\mathcal{B}(K^+ \to \pi^+ X)$ for mass hypothesis $M_X = 0$ (since such a small mass would not be resolved experimentally).

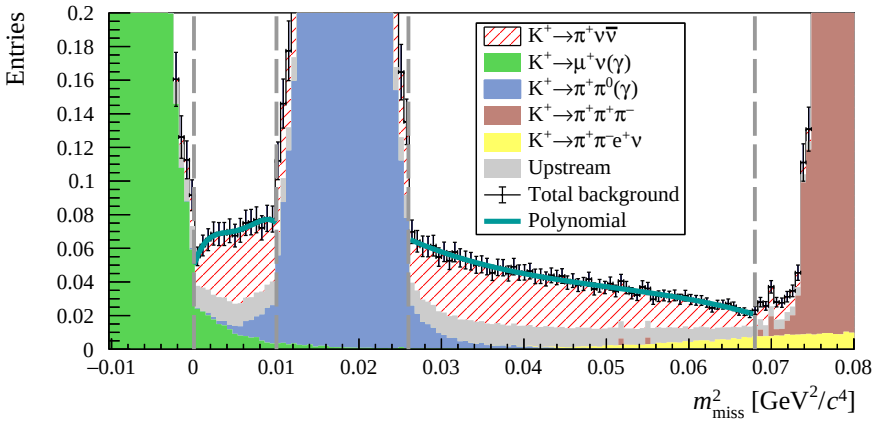
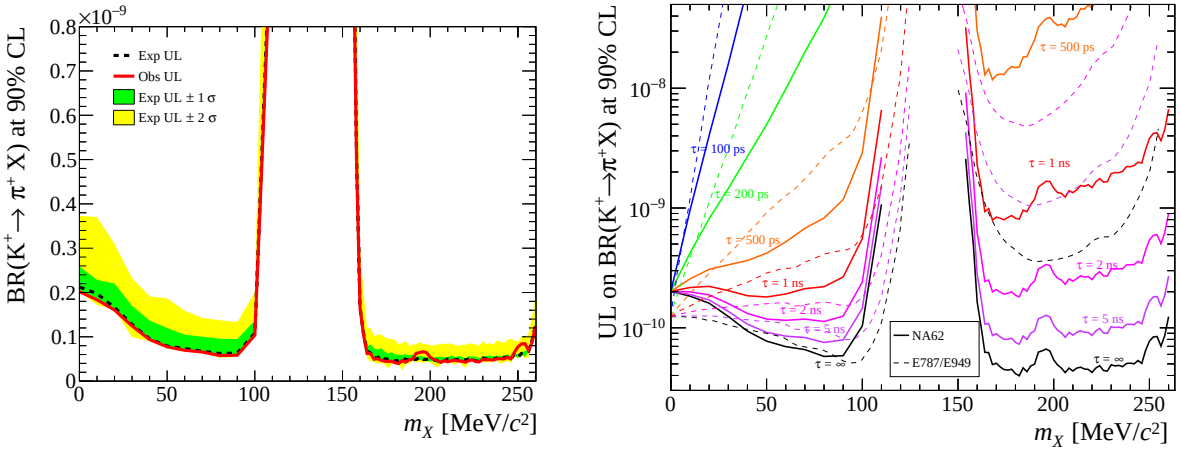


Figure 6.23: [3] Distributions of the expected reconstructed squared missing mass, m^2_{miss}, for background processes, obtained from simulations and data-driven procedures [125]. In each signal region, with boundaries indicated by vertical dashed lines, the polynomial model used to describe the total background is shown.



(a) (b)

Figure 6.24: [3] Upper limits on the branching ratio $\mathcal{B}(K^+ \to \pi^+ X)$ at 90% confidence level derived from the fully frequentist shape analysis. Expected upper limits (black) with 1 and two standard deviation bands (green and yellow) and observed upper limits (red) are shown in (a) for stable or invisibly decaying X, while model-independent branching ratio limits are shown for a range of lifetimes in (b).

for $K^+ \to \pi^+ X$ decays includes a peaking signal which can be described by a Gaussian as shown in figure 6.3b. This is distinct from the relatively smooth (total) background distribution which may be modelled with polynomial functions as shown in figure 6.23. Exploiting this information an alternative search procedure was performed as a fully frequentist shape analysis using an unbinned profiled likelihood ratio test statistic. The profiling procedure accounts for uncertainties associated with the observable resolution, background expectation and single event sensitivity. The parameter of interest is $\mathcal{B}(K^+ \to \pi^+ X)$, and each mass hypothesis m_X is treated independently. The expected and observed upper limits using this procedure are shown for stable or invisibly decaying X in figure 6.24a and converted to limits for a range of lifetimes using the procedure described in section 6.2.4 with corresponding observed upper limits shown in figure 6.24b. Corresponding limits in the case of interpretation of X as a dark scalar or ALP are shown in figure 6.25.

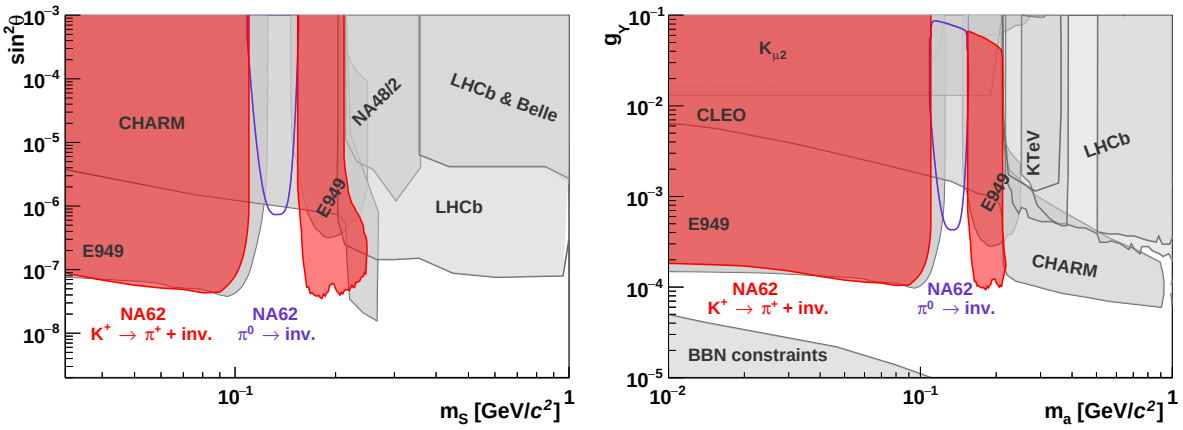
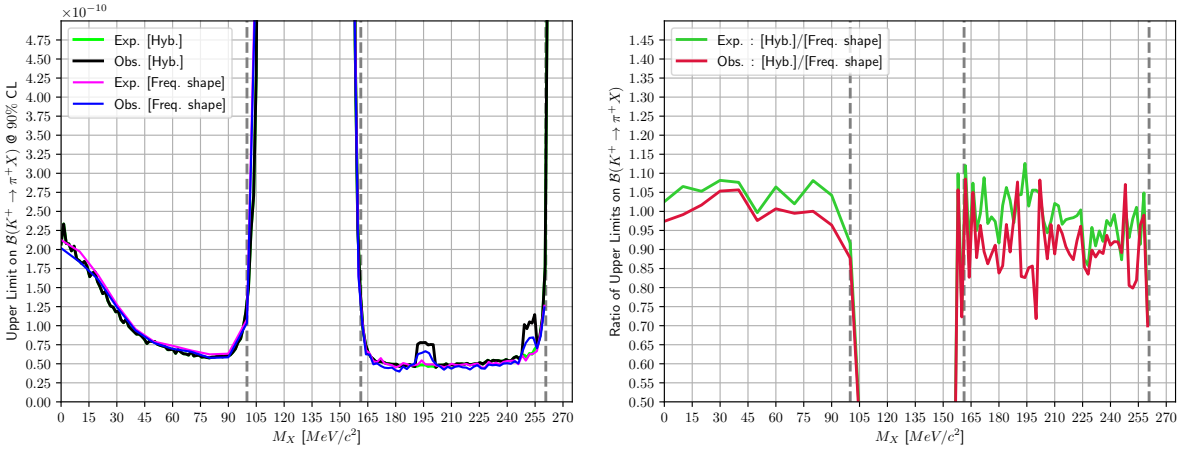


Figure 6.25: [3]Excluded regions of phase space for BC4 (a and c) and BC10 (b and d) models of [54] where X is a scalar mixing with the Higgs boson or a pseudoscalar ALP respectively. Excluded regions are displayed in Scenarios where X decays to visible SM particles in upper panels (a and b) while X decays to invisible (e.g. dark matter) particles is shown in lower panels (c and d). Published NA62 limits from the $K^+ \to \pi^+ X$ search [3] (reported here) are shown in red, alongside exclusion from NA62 $\pi^0 \to invisible$ searches [207] (purple) in the $K^+ \to \pi^+\pi^0$ region. References for other excluded regions displayed in grey are given in [3] and [207], see also figure 6.19 and 6.21.



(a) (b)

Figure 6.26: Comparison between expected and observed upper limits on $\mathcal{B}(K^+ \to \pi^+ X)$ using the method described in sections 6.1– 6.3 with hybrid bayesian-frequentist statistical treatment using a log likelihood ratio test statistic and the fully frequentist shape analysis with profile log likelihood ratio test statistic described in section 6.4.1 [3]. Absolute values are compared in (a) while the ratio between results is shown in (b).

6.4.2 Comparison of Results

In general a fully frequentist procedure sets a more conservative limit than a hybrid (Bayesian-Frequentist) method as used in section 6.3.7, however the shape analysis is expected to improve sensitivity as explained in section 6.4.1. Comparison of results obtained with the procedures described in sections 6.1– 6.3 and 6.4.1, shown in figure 6.26, reveals that comparable or slightly improved sensitivity is obtained with the frequentist shape analysis. Moreover, the consideration of signal shape smooths the boundaries of the perturbations observed as a result of the two observed events. The shape analysis was chosen for the final result and used for publication [3].

Chapter 7

Search for Lepton Number and Lepton Flavour Violating $K^+ \to \pi\mu e$ decays

7.1 Analysis Overview

7.1.1 $K^+ \to \pi\mu e$ decays and Experimental Landscape

As discussed in section 2.5, observation of LNV/(C)LFV decays of the form $K^+ \to \pi\mu e$, forbidden in the SM, would provide clear evidence for BSM physics. The five modes studied are [1]:

1. $K^+ \to \pi^-\mu^+e^+$, the π^- channel
2. $K^+ \to \pi^+\mu^-e^+$, the μ^- channel
3. $K^+ \to \pi^+\mu^+e^-$, the e^- channel
4. $\pi^0 \to \mu^-e^+$ by searching for $K^+ \to \pi^+\pi^0$ followed by $\pi^0 \to \mu^-e^+$ (denoted $K^+ \to \pi^+[\mu^-e^+]_{\pi^0}$)
5. $\pi^0 \to \mu^+e^-$ by searching for $K^+ \to \pi^+\pi^0$ followed by $\pi^0 \to \mu^+e^-$ (denoted $K^+ \to \pi^+[\mu^+e^-]_{\pi^0}$)

The experimental landscape prior to this analysis is summarised in table 7.1. The E865 experiment was specifically designed to search for the $K^+ \to \pi^+\mu^+e^-$ decay and succeeded in setting a stronger limit for this mode. The limits reported in table 7.1 for the two $\pi^0 \to \mu^\pm e^\mp$ modes are taken to be equal to the combined limit[2] set on $\mathcal{B}(\pi^0 \to \mu^-e^+) + \mathcal{B}(\pi^0 \to \mu^+e^-)$ from searches for $K_L^0 \to \pi^0\pi^0\mu^\pm e^\mp$ at the KTeV experiment [173].

[1] The NA62 beam includes K^+ only so the charge conjugate $K^- \to \pi\mu e$ processes are not studied.
[2] The branching ratio of each process cannot be greater than the sum.

Table 7.1: Previous branching ratio upper limits for $K^+ \to \pi\mu e$ Decays.

Decay	$\mathcal{B}$ Limit (@ 90% CL)	Experiment	Publication & Year
$K^+ \to \pi^-\mu^+e^+$	5.0×10^{-10}	BNL E865	[171] , 2000
$K^+ \to \pi^+\mu^-e^+$	5.2×10^{-10}	BNL E865	[171] , 2000
$K^+ \to \pi^+\mu^+e^-$	1.3×10^{-11}	BNL (E777+)E865	[172] , 2005
$\pi^0 \to \mu^-e^+$	3.6×10^{-10}	FNAL KTeV	[173] , 2008
$\pi^0 \to \mu^+e^-$	3.6×10^{-10}	FNAL KTeV	[173] , 2008

As a result of the work reported here upper limits on the branching ratios for the π^- and μ^- channels are improved with respect to previous results by a factor of $\mathcal{O}(10)$, these results are the main focus of this chapter. An improvement on the branching ratio upper limit for the $\pi^0 \to \mu^-e^+$ decay is also reported. The current NA62 data-set does not allow improved sensitivity to the e^- or $\pi^0 \to \mu^+e^-$ modes, although current and projected sensitivities are established. The recent NA62 thesis [260] includes a preliminary independent study of these decay modes (using a smaller data set and preliminary background studies) and clearly demonstrates the NA62 sensitivity to these modes. The study reported in this chapter brings the search for $K^+ \to \pi\mu e$ in NA62 Run 1 data to its conclusion.

7.1.2 Analysis Procedure Overview

The search for $K^+ \to \pi\mu e$ decays is performed using NA62 data collected in 2017 and 2018 from three multi-track trigger streams (see sections 7.1.3 and 7.2.5) and uses the $K^+ \to \pi^+\pi^+\pi^-$ decay collected with the minimum bias multi-track trigger for normalisation. A blind analysis strategy is adopted with a range of reconstructed $\pi\mu e$ candidate final state invariant mass, surrounding the K^+ mass, being masked. Within the blinded region a signal region is defined based on the invariant mass resolution (see section 7.2.3). The background expectations (see section 7.5) are almost exclusively calculated using simulation-driven procedures with significant data-driven corrections which account for limitations in reproducing with simulations the effect of complex processes observed in data. These background expectations are validated using two control regions below and above the signal region. The sensitivity to the signal modes is evaluated (section 7.6) in terms of discovery potential and expected upper limits on the branching ratio, relevant for cases of the observation of existence or lack of an excess of events above the background expectation respectively. Once the analysis procedure and background estimates are finalised and validated the blinded region is opened the result is obtained (section 7.7).

7.1.3 Data Sample And Triggers

Candidate signal $K^+ \to \pi\mu e$ events are collected with the minimum bias multi-track (MT), multi-track μ ($MT\mu$) and multi-track e (MTe) triggers, while only the MT trigger is used for collection of normalisation $K^+ \to \pi^+\pi^+\pi^-$ events. The $MT\mu$ trigger is specifically designed for this analysis requiring both a significant energy deposit in the

LKr, characteristic of a $e^{\pm}$, and a signal in the MUV3, identifying a $\mu^{\pm}$. The MTe trigger requires a greater energy deposit in the LKr and is for collection of multi-track final states including an $e^{\pm}$. Data samples are broken down into periods indicated by the year of collection and letters (e.g. sample 17A) and further into enumerated experimental runs. The precise trigger definitions and the downscaling factors have changed over the course of the data-taking as detailed in table 7.2 with each trigger component defined below and illustrated schematically in figure 7.1. The trigger downscaling factors as a function of run number are shown in figure 7.2.

L0 Trigger Components

RICH: At least two RICH PM signals exist within the trigger time window.

Qx: Signals are recorded from charged particle impacts in at least two diagonally adjacent CHOD quadrants within the trigger time window.

MO1: At least one signal from an outer MUV3 tile within the trigger time window.

LKr10: An energy deposit of at least 10 GeV is recorded in the LKr within the trigger time window.

LKr20: As LKr10 but with a 20 GeV LKr energy threshold.

L1 Trigger Algorithms

KTAG: At least one K^+ candidate, with coincident signals in at least five KTAG sectors, is recorded within the trigger time window.

StrEx: A negative track is reconstructed within the STRAW readout time window.

StrMT: Three charged tracks are reconstructed within the STRAW readout time window.

!LAV: Fewer than two signals are recorded in LAV modules within the trigger time window.

The search for $K^+ \to \pi\mu e$ decays uses data collected during 2017 and 2018. Data collected in 2016 could also be used in principle, however in practice there were significant changes to the trigger and there are known features of the 2016 data which mean special treatment is required [260]. Simulations used for the analysis are tuned to represent data-taking conditions in 2017 and while these can also be used to successfully model 2018 data this is not always true for 2016 data since the difference in data-taking conditions was more significant between 2016 and 2017 than between 2017 and 2018. Because of these reasons, and the fact that the usable 2016 sample is small comprising only 5% of the 2016–18 dataset, it has been decided to neglect the 2016 data and consider 2017+2018 data as the full Run 1 sample for this search.

To streamline the analysis a central data filtering (pre-selection) is performed which selects three track events with at least one $\mu^{\pm}$ or $e^{\pm}$ in the final state. The filter conditions applied are given below, more details on the variables used and the physical motivations are provided in section 7.2.

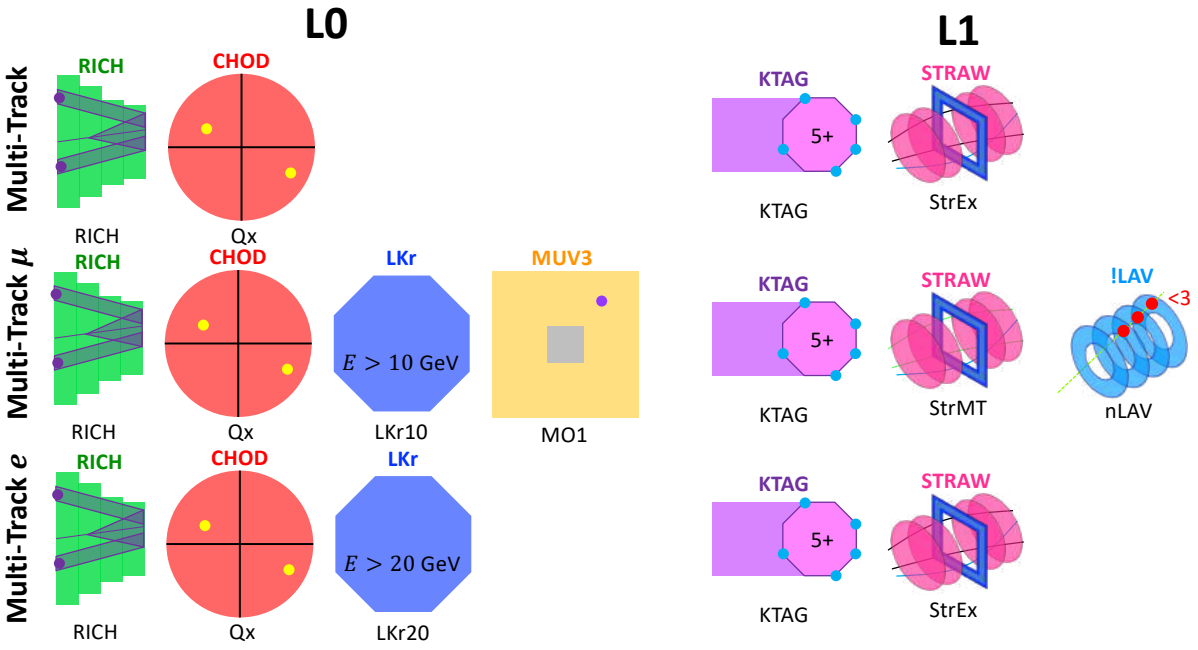


Figure 7.1: Schematic illustrations summarising the trigger conditions imposed for the three triggers used for the $K^+ \to \pi\mu e$ search.

Table 7.2: Trigger definitions and downscaling factors, D.

Sample	Run Range	Trigger		
		Multi-Track	Multi-Track μ	Multi-Track e
2017D	7615–7721	L0: RICH*Qx $D = 10$, 50 or 75 L1: KTAG*StrawExotics	L0: RICH*Qx*M01*LKr10 $D = (1, 15$ or $20)$ L1: KTAG*StrawExotics	L0: RICH*Qx*LKr20 $D = (2$ or $12)$ L1: KTAG*StrawExotics
2017C	7735–7873	L0: RICH*Qx $D = 10$, 50 or 100 L1: KTAG*StrawExotics	L0: RICH*Qx*M01*LKr10 $D = (1$ or $15)$ L1: KTAG*StrawExotics	L0: RICH*Qx*LKr20 $D = (2$ or $10)$ L1: KTAG*StrawExotics
2017B	7876–8107	L0: RICH*Qx $D = 100$ L1: KTAG*StrawExotics	L0: RICH*Qx*M01*LKr10 $D = (10$ or $15)$ L1: KTAG*StrawExotics	L0: RICH*Qx*LKr20 $D = (8$ or $10)$ L1: KTAG*StrawExotics
2017A	8134–8282	L0: RICH*Qx $D = 100$ L1: KTAG*StrawExotics	L0: RICH*Qx*M01*LKr10 $D = 10$ L1: KTAG*StrawExotics	L0: RICH*Qx*LKr20 $D = 8$ L1: KTAG*StrawExotics
2018A	8518–8535	L0: RICH*Qx $D = 100$ L1: KTAG*(!LAV$_{flag}$*)StrawExotics	L0: RICH*Qx*M01*LKr10 $D = 10$ L1: KTAG*StrawExotics	L0: RICH*Qx*LKr20 $D = 8$ L1: KTAG*StrawExotics
2018B	8548–8740	L0: RICH*Qx $D = 100$ L1: KTAG*(!LAV$_{flag}$*)StrawExotics	L0: RICH*Qx*M01*LKr10 $D = (2, 5,$ or $8)$ L1: KTAG*!LAV*StrawMT	L0: RICH*Qx*LKr20 $D = (4$ or $8)$ L1: KTAG*StrawExotics
2018C	8777–8799	L0: RICH*Qx $D = 100$ L1: KTAG*(!LAV$_{flag}$*)StrawExotics	L0: RICH*Qx*M01*LKr10 $D = 5$ L1: KTAG*!LAV*StrawMT	L0: RICH*Qx*LKr20 $D = 8$ L1: KTAG*StrawExotics
2018D	8802–8890	L0: RICH*Qx $D = 100$ L1: KTAG*(!LAV$_{flag}$*)StrawExotics	L0: RICH*Qx*M01*LKr10 $D = 5$ L1: KTAG*!LAV*StrawMT	L0: RICH*Qx*LKr20 $D = 8$ L1: KTAG*StrawExotics
2018E	8891–9040	L0: RICH*Qx $D = 100$ L1: KTAG*(!LAV$_{flag}$*)StrawExotics	L0: RICH*Qx*M01*LKr10 $D = 5$ L1: KTAG*!LAV*StrawMT	L0: RICH*Qx*LKr20 $D = 8$ L1: KTAG*StrawExotics
2018F	9047–9134	L0: RICH*Qx $D = 100$ L1: KTAG*(!LAV$_{flag}$*)StrawExotics	L0: RICH*Qx*M01*LKr10 $D = 5$ L1: KTAG*!LAV*StrawMT	L0: RICH*Qx*LKr20 $D = 8$ L1: KTAG*StrawExotics
2018G	9147–9271	L0: RICH*Qx $D = 100$ L1: KTAG*(!LAV$_{flag}$*)StrawExotics	L0: RICH*Qx*M01*LKr10 $D = 5$ L1: KTAG*!LAV*StrawMT	L0: RICH*Qx*LKr20 $D = 8$ L1: KTAG*StrawExotics
2018H	9304–9462	L0: RICH*Qx $D = (50$ or $100)$ L1: KTAG*(!LAV$_{flag}$*)StrawExotics	L0: RICH*Qx*M01*LKr10 $D = (1, 5,$ or $10)$ L1: KTAG*!LAV*StrawMT	L0: RICH*Qx*LKr20 $D = (8, 12,$ or $16)$ L1: KTAG*StrawExotics

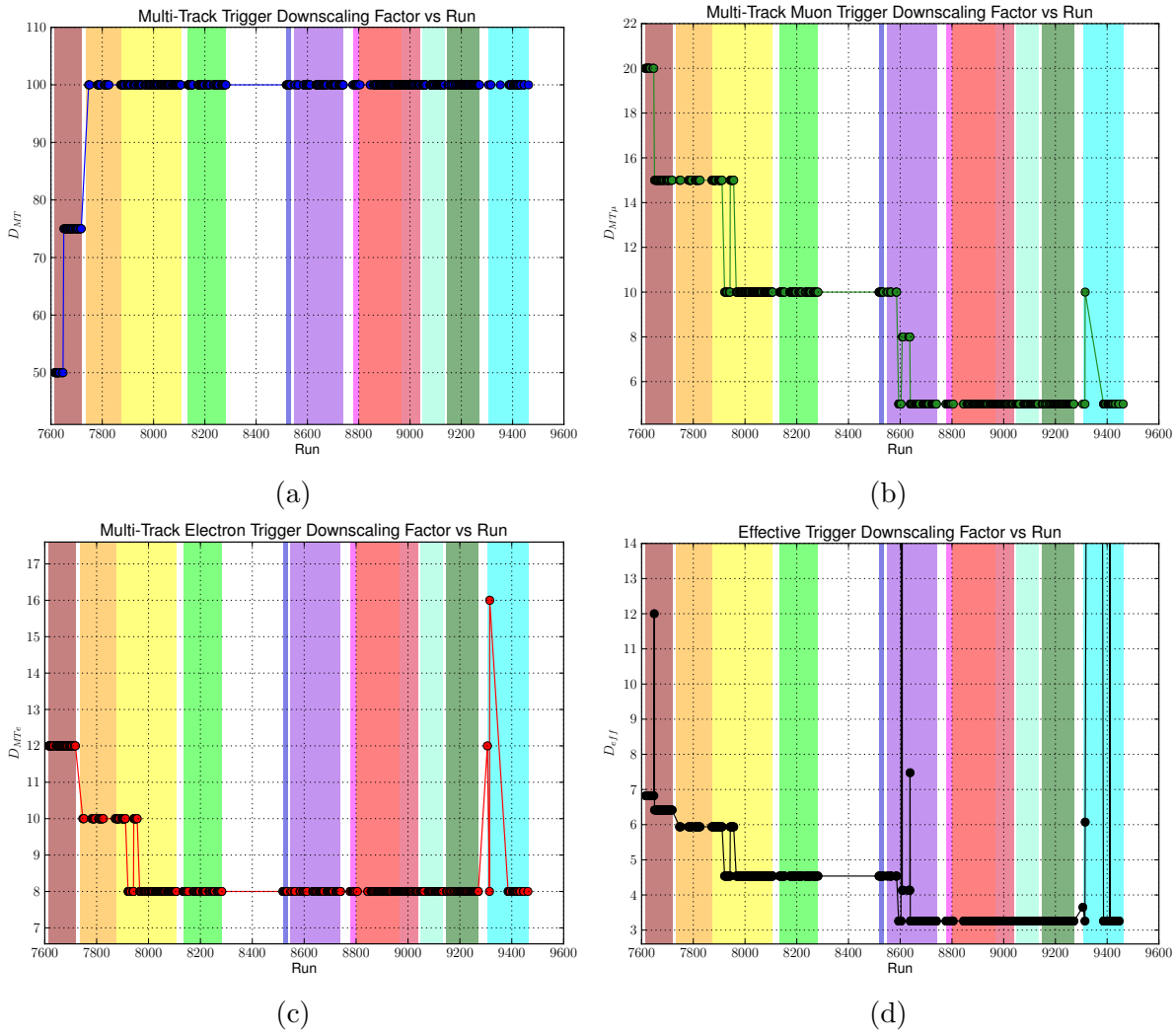


Figure 7.2: Trigger downscaling factors for: a) MT; b) $MT\mu$; c) MTe triggers; and d) effective downscaling factor, as a function of run number with data taking periods contained in shaded bands for full 2017+2018 sample. The four cases in d) where the point is above the y-range of the plot $D_{eff} = D_{MT} = 100$. Coloured bands indicate the run ranges for the named data-taking periods defined in table 7.2.

Data Filtering Criteria

- At least one of the $MT\mu$ and MTe triggers must be in operation.
- An event must be selected by the CTRL, MT, $MT\mu$ or MTe trigger.
- At least one three track vertex must be formed in the event with good quality ($\chi^2_{vtx} < 40$) and downstream of the start of the vacuum tank, ($z_{vtx} > 102\,\text{m}$).
- Each of the three tracks forming a candidate vertex must be reconstructed from signals in at least three of four STRAW chambers.
- At least one track in a candidate vertex is a candidate $e^{\pm}$, with energy deposited in the LKr equal to at least 80% of its momentum ($E/p > 0.8$), or a candidate $\mu^{\pm}$, with a geometrically associated signal in the MUV3 detector.

7.2 Selection

The selection is performed in two stages: a core selection common to the normalisation and signal, followed by separate additional selections for the normalisation and individual signal channels. Filtered data is analysed, with filter pre-selection conditions described in section 7.1.3. The selection criteria applied in the filter are matched by more stringent conditions in the selection meaning there is no bias arising from this when comparing selected data and simulated events.

7.2.1 Common Selection

At least three tracks must be reconstructed by the STRAW in the event, with exactly three of these fulfilling the following conditions:

Track Conditions

1. The track remains inside the geometric active region of all downstream detectors used for the analysis: all four STRAW chambers; the CHODs, LKr, MUV3; and avoiding LAV12 material (see figure 7.3) which otherwise leads to mismeasurement of the particle energy in the LKr.
2. The track is reconstructed from signals in at least three STRAW chambers.
3. The STRAW track reconstruction assigns a good track quality, with $\chi^2_{trk} < 30$, where χ^2_{trk} expresses the quality of the fit achieved through a recursive Kalman filter track reconstruction technique [261].
4. No other track passes within a 25 mm radius around the impact point of the track at the STRAW1.

5. No other track passes within a 200 mm radius around the impact point of the track at the LKr front plane.

6. The momentum of the track calculated before and after the fit performed during track reconstruction must be consistent within 20 GeV/c.

7. the track momentum after the fit p_{trk} must satisfy $10 < p_{trk} < 65\,\text{GeV}/c$.

8. A signal is associated to the track in the CHODs. The track time, t_{trk} is defined based on the NA48-CHOD time if it exists or, otherwise, the CHOD time.

9. The track time must be consistent with the trigger time: $|t_{trk} - t_{trig}| < 5\,\text{ns}$.

A vertex is built from the three selected tracks which must satisfy conditions:

Vertex Conditions

1. The vertex charge $Q_{vtx} = +1$.

2. The vertex is located at a longitudinal position $105 < z_{vtx} < 180\,\text{m}$.

3. The vertex fit quality is good with $\chi^2_{vtx} < 30$, where χ^2_{vtx} expresses the goodness-of-fit for the Billoir-Fruhwirth-Regler implementation [262] of vertex least-squares fitting algorithm used to reconstruct the vertex [260].

Timing Conditions

1. A vertex time, t_{vtx}, is defined from the weighted mean of the track times where weights are assigned based on the time resolution of the detector(s) used. The resolutions measured for the 2017+2018 data sample for the NA48-CHOD and CHOD are $\sigma_N = 0.271\,\text{ns}$ and $\sigma_C = 0.809\,\text{ns}$ respectively, see appendix E. The vertex time must be consistent with the trigger time: $|t_{vtx} - t_{trig}| < 5\,\text{ns}$.

2. An associated KTAG candidate, with coincident signals in at least five sectors, must exist with time consistent with the vertex time: $|t_{KTAG} - t_{vtx}| < 3\,\text{ns}$.

Additional Requierments

- A LAV photon veto condition is applied, rejecting events with signals in any LAV station within 3 ns of the trigger time.

- A 'Qx-like' cut is applied requiring that signals are found in at least two diagonally adjacent CHOD quadrants matched to selected tracks. With this requirement the efficiency of the Qx L0 trigger component for selected events should be close to 100%.

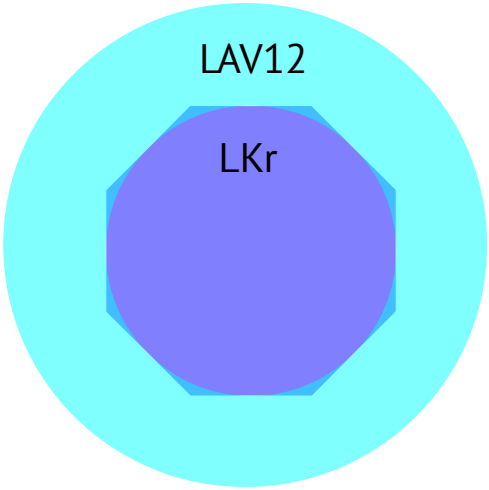


Figure 7.3: Schematic illustration of the slightly overlapping active regions of the LAV12 and LKr detectors in the (x, y) plane.

- The magnitude of the total resultant momentum of the three selected tracks after the vertex fit correction, $|\mathbf{p_{\pi\mu e}}|$ must be consistent with the beam K^+ momentum magnitude (measured run-by-run for data and stored in the BeamParameters database) $|\mathbf{p^{BP}_{K^+}}|$ such that: $||\mathbf{p_{\pi\mu e}}| - |\mathbf{p^{BP}_{K^+}}|| < 2.5\,\mathrm{GeV}/c$. This will reject decays with additional particles in the final state which carry away at least $2.5\,\mathrm{GeV}/c$ of the K^+ momentum.

- The transverse momentum of the total resultant three track vertex momentum with respect to the beam momentum must be small: $p_T < 0.030\,\mathrm{GeV}/c$. This is because the signal decay should have a transverse momentum consistent with zero in an ideal detector.

7.2.2 Normalisation $K^+ \to \pi^+\pi^+\pi^-$ Selection

The invariant mass of the three selected final state tracks is calculated under the hypothesis that all three are pions,

$$M_{3\pi} = \sqrt{\left(\sqrt{p_1^2 + m_\pi^2} + \sqrt{p_2^2 + m_\pi^2} + \sqrt{p_3^2 + m_\pi^2}\right)^2 - (\mathbf{p_1} + \mathbf{p_2}^2 + \mathbf{p_3})^2}\,, \qquad (7.2.1)$$

where $\mathbf{p_i}$ are the track three-momenta measured by the STRAW, with magnitude p_i, and m_π is the charged pion mass. A $K^+ \to \pi^+\pi^+\pi^-$ event is then selected by requiring $|M_{3\pi} - m_{K^+}| < 2.55\,\mathrm{MeV}/c^2$, where $2.55\,\mathrm{MeV}/c^2 = 3\sigma_{M_K^{3\pi}}$ and $\sigma_{M_K^{3\pi}}$ is the corresponding invariant mass resolution, measured to be $\sigma_{M_K^{3\pi}} = (0.85 \pm 0.05_{syst})\,\mathrm{MeV}/c^2$ in the 2017+2018 data sample. The systematic uncertainty is determined by the difference between data and MC simulations as well as the range of values obtained by varying possible fit ranges.

7.2.3 Signal $K^+ \to \pi\mu e$ Selection

A particle identification (PID) selection is applied requiring positive identification of the three tracks as $\pi^\pm$, $\mu^\pm$ and $e^\pm$ with charge combination defining which signal sample the event enters. The PID requirements are:

$\pi^\pm$: The ratio between the energy of the geometrically closest matching LKr cluster, E, and the track momentum, p, must satisfy $E/p < 0.9$.[3] This specifically rejects $e^\pm$, see below and figures 7.4a and 7.5. The geometrically closest matching LKr cluster must be within 50 mm of the track impact point at the LKr front plane and have a time consistent with the track time, $|t_{LKrClust} - t_{track}| < 5$ ns. It is also possible no LKr cluster is associated to the track, this gives $E/p = 0$, the time check is not applicable and this is allowed since a $\pi^\pm$ may act as a MIP. No geometrically associated signals in any MUV3 tiles within 5 ns of the track time are allowed.

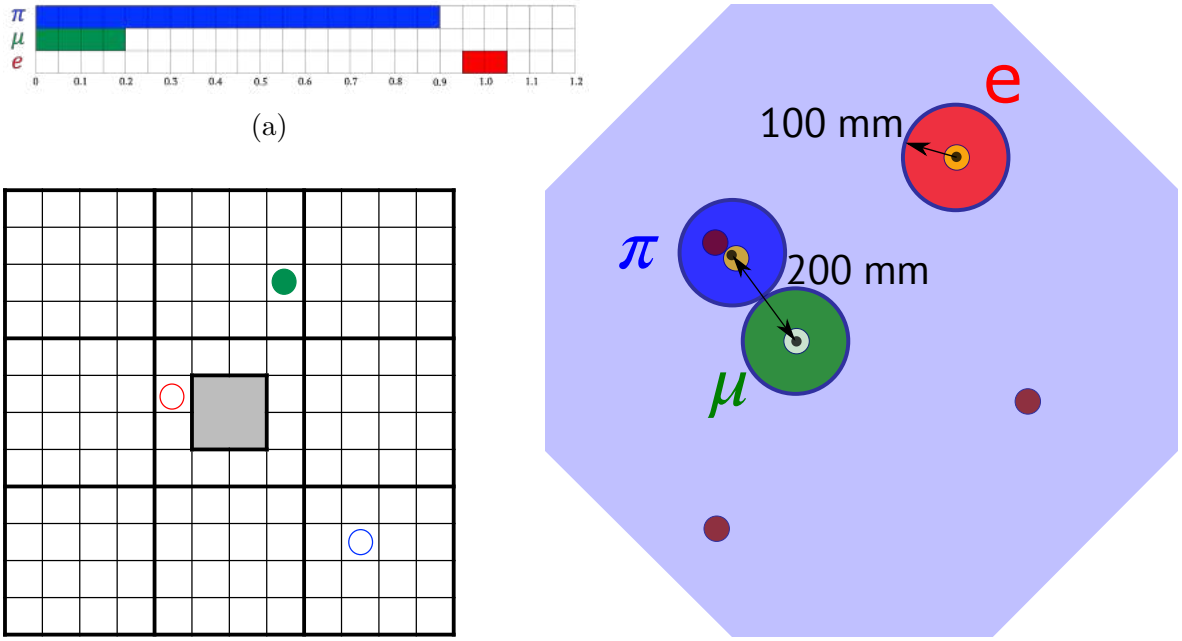
$\mu^\pm$: The E/p ratio must be small and consistent with a MIP: $E/p < 0.2$. If a cluster is associated to the track its time must be consistent with the track time within 5 ns. Exactly one geometrically associated signal must exist in an outer MUV3 tile within 5 ns of the track time. Only outer tiles are used to remove potential accidental background including a hit from a μ^+ from beam pion decay (see section 3.2.10) and for consistency with the MO1 trigger requirement, which means for signal events the MO1 condition should be effectively 100% efficient.

$e^\pm$: The E/p ratio must be consistent with unity: $0.95 < E/p < 1.05$. This is because $e^\pm$ are expected to deposit all of their energy in the LKr and $p \gg m_e$, see figure 7.5. The data sample is dominated by $\pi^\pm$ from the abundant $K^+ \to \pi^+\pi^+\pi^-$ decay and therefore the E/p range is restricted and separated from the allowed range for $\pi^\pm$ to minimise the probability of misidentification. For $e^\pm$ exactly one LKr cluster is expected, in contrast to $\pi^\pm$ where the possibility of hadronic showers can lead to more LKr clusters, each containing a fraction of the energy. To discriminate between these cases and select $e^\pm$ an exclusion zone with 100 mm radius is established around a candidate $e^\pm$ track impact point at the LKr front plane and only one LKr cluster is allowed in this region within 5 ns of the track time.

The PID requirements use the LKr and MUV3, see schematic illustrations in figures 7.4a and 7.4b, the performance for positive identification and misidentification cases relevant for the current study are evaluated in section 7.5. The LKr cluster conditions and track separation requirements at the LKr are illustrated schematically in figure 7.4c. The use of the RICH for additional discrimination power was considered but found to give worse sensitivity with similar discovery potential, due to reduced acceptance, and additionally makes background estimates more imprecise.

The invariant mass, $M_{\pi\mu e}$, of a selected $K^+ \to \pi\mu e$ candidate event is calculated,

[3] Considering the dimensions of the variables the E/p ratio should be expressed in units of the speed of light, c. However, in this thesis the short-hand 'E/p' should be taken to mean $E/(pc)$ which is expressed as a dimensionless number.



(a) (b) (c)

Figure 7.4: a) Schematic illustration of the allowed LKr E/p ranges for π, μ and e candidates.
b) Illustration of the MUV3 requirement, with exactly one outer MUV3 tile signal associated to a μ candidate and within 5 ns of the track time and no associated in-time signals in any MUV3 tile for π or e candidates.
c) Schematic illustrations of the track separation and LKr cluster conditions from the PID selection. The impact points of the three candidate tracks are shown by black points, it is required that no other track is within 200 mm of these points, with limiting case shown for the π and μ tracks. LKr clusters best matched to tracks geometrically are shown in orange, while other clusters are shown in brown, in the case illustrated there is no LKr cluster reconstructed for the μ^+ track, shown by the light blue circle. Two clusters are matched to the π track and the best matched geometrically has energy $E < 0.9p_\pi c$ and is in time with the track time while the other cluster is generally a low energy counterpart. Only one LKr cluster is reconstructed within 100 mm of the e impact point with its time consistent with the track time. Additional pileup or photon clusters can exist in the LKr.

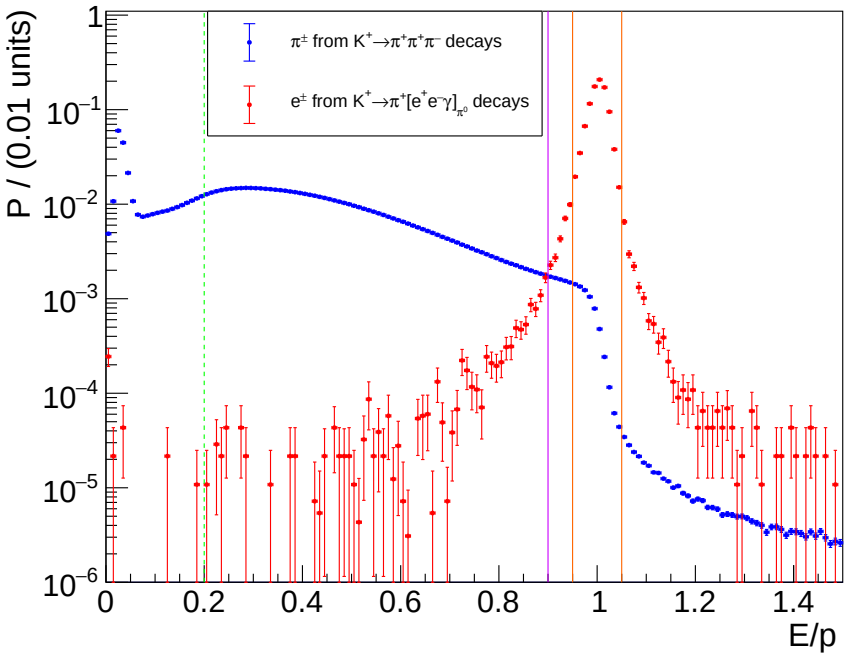


Figure 7.5: Probability density functions for LKr E/p for $\pi^{\pm}$ and $e^{\pm}$ from selected control samples of $K^+ \to \pi^+\pi^+\pi^-$ and $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ collected with the CTRL trigger for the full 2017+2018 data sample. The allowed ranges for $e^{\pm}$ (0.95–1.05), $\pi^{\pm}$ (< 0.9) and μ (< 0.2) are indicated by orange, purple and green vertical lines respectively.

where

$$M_{\pi\mu e} = \sqrt{\left(\sqrt{p_\pi^2 + m_\pi^2} + \sqrt{p_\mu^2 + m_\mu^2} + \sqrt{p_e^2 + m_e^2}\right)^2 - (\mathbf{p_\pi} + \mathbf{p_\mu} + \mathbf{p_e})^2}\,. \tag{7.2.2}$$

The region $487 < M_{\pi\mu e} < 510\,\mathrm{MeV}/c^2$ is blinded for data and the signal region is established as $490 < M_{\pi\mu e} < 498\,\mathrm{MeV}/c^2$ with choice of width informed by the invariant mass resolution, measured for the signal modes to be $1.34 \pm 0.05_{syst}\,\mathrm{MeV}/c^2$. The selected samples of $K^+ \to \pi^+[\mu^\mp e^\pm]_{\pi^0}$ candidates are a subset of the corresponding $K^+ \to \pi^+\mu^\mp e^\pm$ candidates which additionally satisfy the requirement $|M_{\mu e} - m_{\pi^0}| < 3\,\mathrm{MeV}/c^2$, where $M_{\mu e}$ is the invariant mass of the $\mu^\mp e^\pm$ pair,

$$M_{\mu e} = \sqrt{\left(\sqrt{p_\mu^2 + m_\mu^2} + \sqrt{p_e^2 + m_e^2}\right)^2 - (\mathbf{p_\mu} + \mathbf{p_e})^2}\,. \tag{7.2.3}$$

The choice of width is again informed by the invariant mass resolution, measured to be $0.41 \pm 0.02_{syst}\,\mathrm{MeV}/c^2$.

The background arising from $K^+ \to \pi^+\pi^+\pi^-$ decays upstream of the fiducial volume is found to be asymmetric between the decay modes and therefore channel-dependent requirements are made on the minimum value of z_{vtx}. Specifically, the three-track vertex must be formed in the range defined by $z_{min,i} < z_{vtx} < 180\,\mathrm{m}$ where $z_{min,\pi^-} = 107\,\mathrm{m}$, $z_{min,\mu^-} = 111\,\mathrm{m}$ and $z_{min,e^-} = 110\,\mathrm{m}$, the choice of these values and the upstream background are discussed in section 7.5.7.1.

For the π^- and e^- channels important backgrounds exist from decays including a π^0 in the final state which undergoes a Dalitz decay, $\pi^0 \to e^+e^-\gamma$, with the photon undetected

and an $e^{\pm}$ misidentified as a $\pi^{\pm}$. These backgrounds can be rejected by requiring that the invariant mass of the identified $\pi^{\pm}e^{\mp}$ pair (which is actually a e^+e^- pair) has an invariant mass, under the hypothesis of the electron mass for both particles, $M_{\pi e}(m_e, m_e)$, larger than the π^0 mass. This 'Dalitz rejection cut' requires $M_{\pi e}(m_e, m_e) > 140\,\mathrm{MeV}/c^2$. While this reduces acceptance by 20% in relative terms it completely removes the backgrounds including a Dalitz decay. The Dalitz rejection cut is not applied to the μ^- channel where this background mechanism does not exist.

7.2.4 Selection Acceptances

The acceptance for the signal and normalisation selections are evaluated using simulations according to

$$A = \frac{N_{sel}}{N_{MC}}\,, \tag{7.2.4}$$

where N_{MC} is the number of simulated events with a K^+ decay in the fiducial volume, defined as $105 < z < 180\,\mathrm{m}$, and N_{sel} is the number of events passing the signal selection. The acceptances are presented in table 7.3. Uncertainties are statistical and determined by the size of the sample of simulated events with approximately 50×10^6 events for normalisation, 1×10^6 events for direct $K^+ \to \pi\mu e$ signal decays and 0.5×10^6 events for the signal $K^+ \to \pi^+[\mu^{\mp}e^{\pm}]_{\pi^0}$ channels. The Dalitz rejection cut is responsible for the lower acceptance of the π^- and e^- channels [4] when compared to the μ^- channel while the decay kinematics lead to significantly lower geometric acceptance for the $K^+ \to \pi^+[\mu^{\mp}e^{\pm}]_{\pi^0}$ channels.

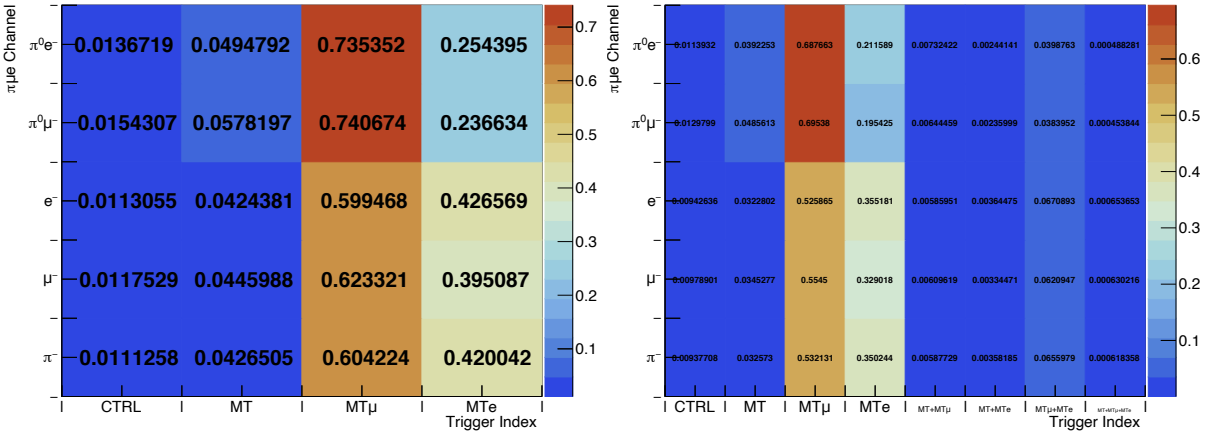
Table 7.3: Signal and normalisation selection acceptances.

Case	Acceptance
$K^+ \to \pi^+\pi^+\pi^-$	$(10.177 \pm 0.002)\%$
$K^+ \to \pi^-\mu^+e^+$	$(4.90 \pm 0.02)\%$
$K^+ \to \pi^+\mu^-e^+$	$(6.21 \pm 0.02)\%$
$K^+ \to \pi^+\mu^+e^-$	$(4.64 \pm 0.02)\%$ [4]
$K^+ \to \pi^+[\mu^-e^+]_{\pi^0}$	$(3.11 \pm 0.02)\%$
$K^+ \to \pi^+[\mu^+e^-]_{\pi^0}$	$(2.73 \pm 0.02)\%$

7.2.5 Trigger Mixture

The combination of trigger downscaling factors (see section 7.1.3 and figure 7.2) and efficiencies (see section 7.3) determine the mixture of triggers observed for selected candidate $K^+ \to \pi\mu e$ events. The observed fraction of selected events arising from each trigger or combination of triggers is shown in figure 7.6. The largest sample is selected from $MT\mu$ triggers, exclusively (54%) or which also pass the MTe trigger (6%), the second largest contributor (34%), while the events passing the MT trigger represent a small ($\sim$ 4%

[4] If the Dalitz rejection cut is not applied for the e^- channel the acceptance obtained is $(6.16 \pm 0.02)\%$.



(a) (b)

Figure 7.6: The fraction of selected candidate $K^+ \to \pi\mu e$ events (in control regions) which pass each trigger with an inclusive measurement (a) and exclusive measurement (b). In the inclusive case (a) an event passing more than one trigger count towards the result for each trigger and the fractions are taken with respect to the total number of events selected in the control region from all triggers. Because of this the total for a given channel does not add up to 1, however the sum does not exceed 1 by more than $\sim 7\%$ indicating that the overlap between triggers is not large meaning a significant gain in sample size is achieved by combining them. The exclusive case (b) shows the fractions of events collected from each each possible trigger mixture. This illustrates that around 6% of selected events pass both the $MT\mu$ and MTe, while approximately half of events pass only the $MT\mu$ trigger which was designed for this analysis. The CTRL trigger is also included in the MUE3TV filter and so the number of events from this trigger is also studied, although it is not used for the analysis.

inclusive) addition. The relative contributions from each trigger and the overlap for the three signal triggers is illustrated in figure 7.7.

7.3 Trigger Efficiency Studies

The efficiency for collection of normalisation $K^+ \to \pi^+\pi^+\pi^-$ and $K^+ \to \pi\mu e$ events with the corresponding triggers must be measured to evaluate the sensitivity of the search for $K^+ \to \pi\mu e$ decays and calculate the expected background. The normalisation trigger efficiency, ε_n, is the probability that the MT trigger selects the event given that it passes the normalisation $K^+ \to \pi^+\pi^+\pi^-$ selection. The signal trigger efficiency, ε_s is the probability that at least one of the three signal triggers selects a signal event given that it passes the signal $K^+ \to \pi\mu e$ selection. Because of the blind analysis strategy, and because $K^+ \to \pi\mu e$ decays are forbidden in the SM so are not expected to exist by default, the signal trigger efficiency cannot be measured directly using data. However, a representative measurement can be made for signal-like events, selected in data with the signal selection minus the Dalitz rejection cut, and combining all three $\pi\mu e$ charge combinations outside the blinded regions. The trigger efficiency for collection of signal is measured for each

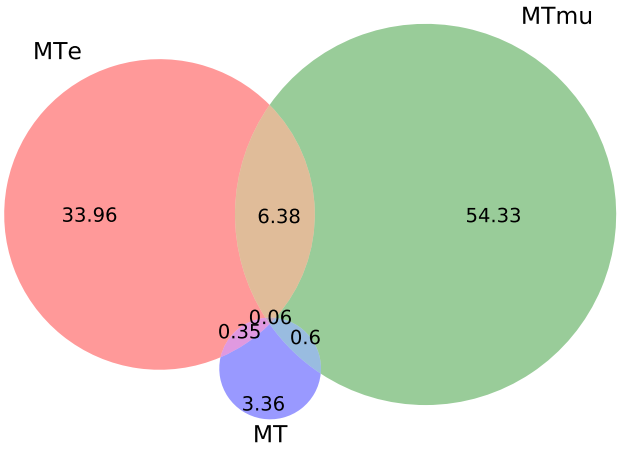


Figure 7.7: Venn diagram showing the percentage of selected candidate $K^+ \to \pi\mu e$ events in control regions which pass each trigger or trigger mix.

of the three signal triggers independently and the total effective signal trigger efficiency calculated considering the downscaling factors, is given by

$$\varepsilon_s = \frac{1 - \left(1 - \frac{\varepsilon^s_{MT}}{D_{MT}}\right)\left(1 - \frac{\varepsilon^s_{MT\mu}}{D_{MT\mu}}\right)\left(1 - \frac{\varepsilon^s_{MTe}}{D_{MTe}}\right)}{1 - \left(1 - \frac{1}{D_{MT}}\right)\left(1 - \frac{1}{D_{MT\mu}}\right)\left(1 - \frac{1}{D_{MTe}}\right)}\,, \tag{7.3.1}$$

where D_x and ε^s_x are the downscaling factor and trigger efficiency for collection of signal-like events for trigger x. The dataset is divided into periods (groups of runs, which approximately coincide with the named data-taking periods, e.g. sample 17A) with the same downscaling factors and equation 7.3.1 is valid explicitly for each period with its associated downscaling factors. As shown in section 7.2.5 there is a correlation between trigger streams. Using the measured correlation (factors $\rho_{x,y}$), and using the fact that the downscaling factors are, in principle, exact numbers with no associated uncertainty, the uncertainty associated with the effective signal trigger efficiency, ε_s, is

$$\begin{aligned}\sigma_{\varepsilon_s} = \varepsilon_s \Bigg(& \underbrace{\left(\frac{\sigma_{\varepsilon^s_{MT}}}{\varepsilon^s_{MT}}\right)^2 + \left(\frac{\sigma_{\varepsilon^s_{MT\mu}}}{\varepsilon^s_{MT\mu}}\right)^2 + \left(\frac{\sigma_{\varepsilon^s_{MTe}}}{\varepsilon^s_{MTe}}\right)^2}_{\text{uncorrelated term}} \\ & + \underbrace{2\rho_{MT,MT\mu}\frac{\sigma_{\varepsilon^s_{MT}}}{\varepsilon^s_{MT}}\frac{\sigma_{\varepsilon^s_{MT\mu}}}{\varepsilon^s_{MT\mu}} + 2\rho_{MT,MTe}\frac{\sigma_{\varepsilon^s_{MT}}}{\varepsilon^s_{MT}}\frac{\sigma_{\varepsilon^s_{MTe}}}{\varepsilon^s_{MTe}} + 2\rho_{MT\mu,MTe}\frac{\sigma_{\varepsilon^s_{MT\mu}}}{\varepsilon^s_{MT\mu}}\frac{\sigma_{\varepsilon^s_{MTe}}}{\varepsilon^s_{MTe}}}_{\text{correlations between trigger pairs}} \\ & + \underbrace{3\rho_{MT,MT\mu,MTe}\frac{\sigma_{\varepsilon^s_{MT}}}{\varepsilon^s_{MT}}\frac{\sigma_{\varepsilon^s_{MT\mu}}}{\varepsilon^s_{MT\mu}}\frac{\sigma_{\varepsilon^s_{MTe}}}{\varepsilon^s_{MTe}}}_{\text{correlation between all three triggers}} \Bigg)^{0.5}\,.\end{aligned} \tag{7.3.2}$$

To construct $\varepsilon^{n/s}_x$, the normalisation/signal trigger efficiency for a trigger x, the

	Trigger Component	Normalisation ($K_{3\pi}$)	Signal ($K_{\pi\mu e}$)
Individual L0 Trigger Bits	L0 RICH		
	L0 Qx		
	L0 M01		
	L0 LKr10		
	L0 LKr20		
Full L0 Trigger	**L0 MT** : RICH*Qx		
	L0 MTμ : RICH*Qx*M01*LKr10		
	L0 MT*e* : RICH*Qx*LKr20		
Individual L1 Trigger Bits	L1 KTAG		
	L1 StrawExotics		
	L1 StrawMT		
	L1 !LAV		
Full L1 Trigger	**L1 MT** : KTAG*StrawExotics		
	L0 MTμ : KTAG*[StrawExotics/StrawMT*!LAV]		
	L0 MT*e* : KTAG*StrawExotics		
Full Trigger	**MT**		
	MTμ		
	MT*e*		

Figure 7.8: Schematic table indicating the trigger efficiency measurements required for the analysis with shaded boxes.

efficiency of individual L0 trigger components and L1 algorithms are evaluated and combined. The measurements that must be made are summarised in figure 7.8 and details of these measurements are given in the following.

7.3.1 Trigger Efficiency Measurement Methodology

Trigger efficiency measurements are performed using events which pass the minimum bias CTRL trigger and pass either the normalisation selection or enter the signal-like sample defined in section 7.3. To decouple detector effects which are accounted for in the selection acceptances additional checks are performed, with an offline selection matching the trigger requirements, to determine if an event is expected to produce a L0 trigger primitive or pass a L1 algorithm. Efficiencies of L0 trigger components are measured relative to the RICH. This means that the RICH trigger primitive is used as the reference and existence of other L0 trigger primitives are searched for in a time window centred on the RICH primitive. To measure the efficiency of the L0 RICH a CHOD primitive is used as a reference.

The trigger efficiency for a specific L0 component x for the collection of signal /normalisation events is then given by

$$\varepsilon_x^{s/n} = \frac{N(s/n\,\text{event}\,,\,x\,\text{primitive}\,,\,r\,\text{primitive})}{N(s/n\,\text{event}\,,\,r\,\text{primitive})}\,. \tag{7.3.3}$$

The denominator is the number of selected signal/normalisation events which have the properties expected to produce the trigger primitive of interest, x, and a reference trigger primitive, r, exists. The numerator is the number of these events for which a trigger primitive of type x is found in the time window opened around the reference trigger primitive. For composite L0 trigger efficiencies the individual L0 trigger bit efficiencies are multiplied together.

To measure the efficiency of the L1 trigger algorithms it is first required that an event passes the L0 conditions of a given trigger and is autopassed at L1. Such events, with properties expected to also pass the L1 algorithm(s) of interest enter the denominator while those which are also found to pass the L1 algorithm enter the numerator. The total L1 efficiency is calculated by looking at events that pass the full chain of L1 algorithms rather than multiplying those of independent components. This is because the L1 trigger is defined by a sequential application of algorithms, in contrast to the L0 trigger primitives which are created independently.

7.3.2 Trigger Efficiency Results

Trigger efficiencies are calculated for each experimental run and an average efficiency is calculated for each data-taking period, results are shown in section 7.3.2.1. For the LKr10 and LKr20 L0 trigger components the efficiency is dependent on the energy deposited in the LKr and the study of this dependence is presented in section 7.3.2.2. The study of the impact of the change in the L1 $MT\mu$ trigger, see section 7.1.3, is described in section 7.3.2.3.

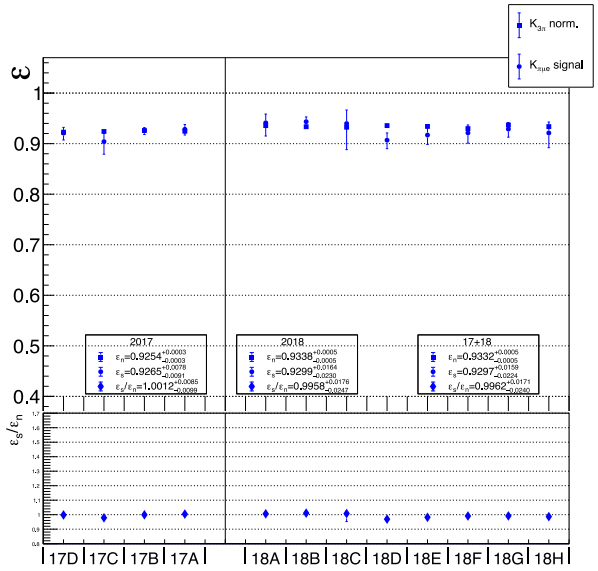
7.3.2.1 Summary of Results

The measured trigger efficiency for signal and normalisation events for the MT, $MT\mu$ and MTe triggers is shown in figures 7.9a, 7.9b and 7.9c respectively for the full 2017+2018 data set. For the MT trigger the measured efficiencies are consistent, to 1% precision, with the hypothesis that the trigger is equally efficient for signal and normalisation events. In contrast, for the $MT\mu$ and MTe triggers there are noticeable differences in efficiency, accounted for primarily by the L0 LKr10 and LKr20 conditions in the $MT\mu$ and MTe triggers respectively, see section 7.3.2.2. The efficiency of the $MT\mu$ trigger was also lower in 2018 than 2017, this is because of the change in the L1 algorithms used.

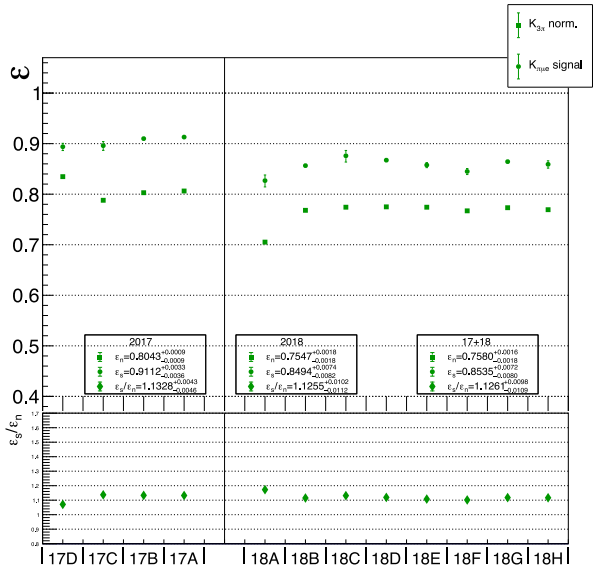
The overall signal and normalisation trigger efficiencies, as defined in section 7.3, are summarised in table 7.4. The efficiencies change from period to period due to changes in trigger configuration, either differences in downscaling factors or L0/L1 conditions/algorithms. For the normalisation MT trigger the trigger configuration was, to good approximation, not changed in this dataset and the normalisation trigger efficiency can therefore be specified as

$$\varepsilon_n = (93.18 \pm 0.04_{stat} \pm 0.49_{syst})\% \ , \tag{7.3.4}$$

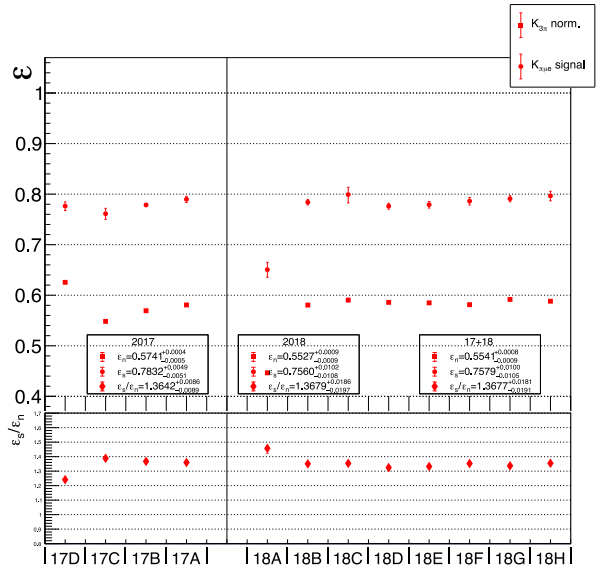
where the systematic uncertainty quoted accounts for sample-to-sample variations.



(a) MT trigger efficiency results.



(b) $MT\mu$ trigger efficiency results.



(c) MTe trigger efficiency results.

Figure 7.9: Overall trigger efficiency for a) MT, b) $MT\mu$ and c) MTe triggers for normalisation (ε_n) and signal-like (ε_s) events in each data sample and $\frac{\varepsilon_s}{\varepsilon_n}$ ratio.

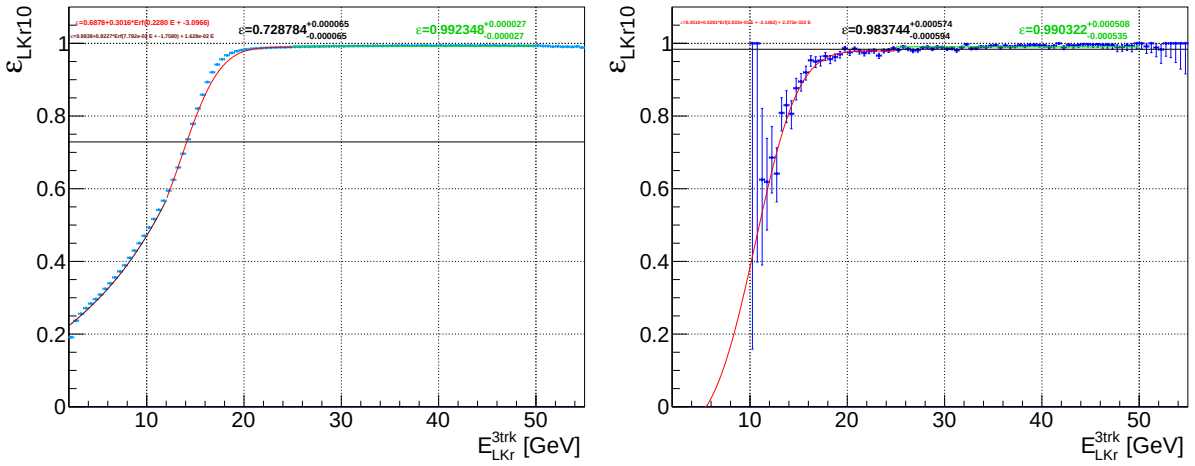
Table 7.4: Trigger efficiency results summary. The normalisation trigger efficiency, ε_n, is the efficiency of the MT trigger for collection of $K^+ \to \pi^+\pi^+\pi^-$ events passing the normalisation selection. The signal trigger efficiency, ε_s, is the effective efficiency (defined by equation 7.3.1) of the MT, $MT\mu$ and MTe triggers for collection of signal-like events.

Sample	ε_n [%]	ε_s [%]	$\frac{\varepsilon_s}{\varepsilon_n}$
17D	92.260 ± 0.035	84.0 ± 2.0	0.911 ± 0.022
17C	92.400 ± 0.055	82.6 ± 3.1	0.894 ± 0.033
17B	92.605 ± 0.022	84.7 ± 1.0	0.914 ± 0.011
17A	92.445 ± 0.033	85.6 ± 1.5	0.926 ± 0.016
18A	93.567 ± 0.065	75.0 ± 3.6	0.801 ± 0.038
18B	93.348 ± 0.036	84.2 ± 1.5	0.902 ± 0.016
18C	93.25 ± 0.10	86.2 ± 5.6	0.924 ± 0.060
18D	93.568 ± 0.036	84.8 ± 1.9	0.907 ± 0.020
18E	93.413 ± 0.042	84.4 ± 2.1	0.904 ± 0.022
18F	92.985 ± 0.049	83.8 ± 2.3	0.902 ± 0.024
18G	93.686 ± 0.040	85.2 ± 1.8	0.910 ± 0.020
18H	93.367 ± 0.062	85.0 ± 3.2	0.911 ± 0.035
17	92.511 ± 0.027	84.6 ± 1.4	0.915 ± 0.015
18	93.425 ± 0.046	84.4 ± 2.3	0.904 ± 0.024
17+18	93.178 ± 0.041	84.2 ± 2.1	0.903 ± 0.022

7.3.2.2 Study of energy dependence of LKr10(20) L0 Trigger Components

The LKr10(20) L0 trigger components have an efficiency which depends on the energy deposited in the LKr. A 'turn on' curve is expected around the threshold set at trigger level of 10(20) GeV. To measure this energy dependence no check is made on the event passing the energy threshold and numerator and denominator counts, of events selected and all events respectively, are made in 0.5 GeV wide bins of energy. The trigger is sensitive to all energy deposition in the LKr inside the trigger window, including energy associated to the three selected tracks as well as any other pileup tracks or photons in the event, as well as being sensitive to detector timing glitches. The latter effects are difficult to simulate precisely, while the energy deposited in the LKr associated with the three selected charged tracks, E^{3trk}_{LKr}, can be simulated and therefore this variable is used for the study of trigger efficiency energy dependence. This allows the resulting measurement to be applied to simulations, which do not contain trigger decisions, to emulate the response of the trigger. More details on this, and its use in calculating background expectations, will be given in section 7.5.8.1.

The efficiency of the L0 LKr10 and LKr20 trigger components as a function of E^{3trk}_{LKr} are shown by figures 7.10a(7.10b) and 7.11a(7.11b) respectively for normalisation(signal-like) events. Results are consistent between normalisation and signal-like events, as expected since the dependence is on the energy deposited in the LKr and not the species of particle. Because of the higher statistics available for normalisation events the associated models are used to describe the energy dependence of the LKr10 and LKr20 L0 trigger components. The overall efficiency is higher for the signal-like events, due to the



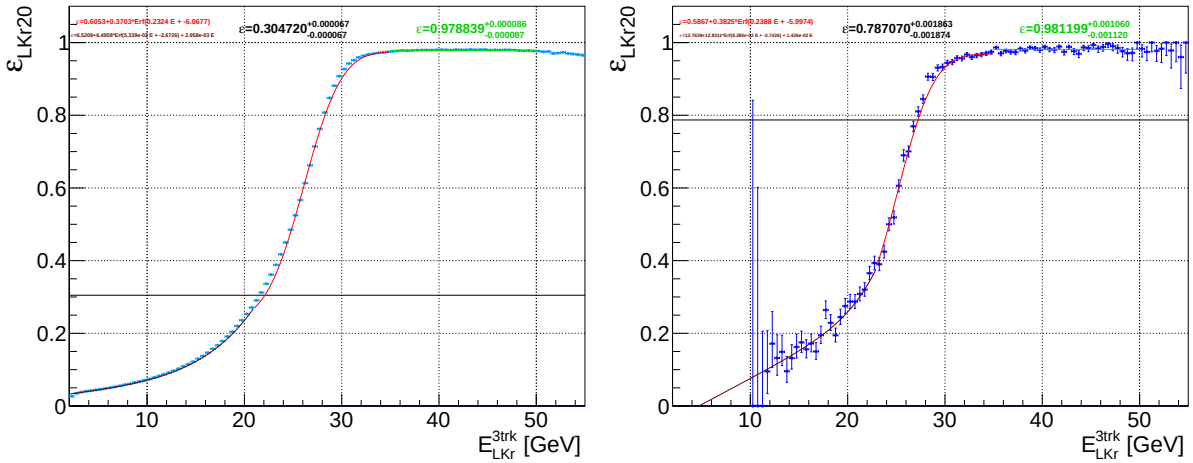
(a) (b)

Figure 7.10: L0 LKr10 trigger component efficiency as a function of energy deposited in the LKr associated to the selected 3 charged tracks, E^{3trk}_{LKr}, for normalisation $K^+ \to \pi^+\pi^+\pi^-$ (a) and signal-like (b) events in the full 2017+2018 data-set. The black line and label gives the overall average efficiency and models describing the shape as a function of E^{3trk}_{LKr} are shown.

Table 7.5: LKr10(20) trigger efficiency for each signal mode. Uncertainties include statistical and systematic components, the latter dominates and is due to data sample-to-sample variations and the precision of the data-driven correction to the energy deposition spectra (see section 7.5.8.1).

Channel	ε_{LKr10} [%]	ε_{LKr20} [%]
$K^+ \to \pi^-\mu^+e^+$	97.5 ± 1.3	74.1 ± 1.6
$K^+ \to \pi^+\mu^-e^+$	97.5 ± 1.3	73.3 ± 1.6
$K^+ \to \pi^+\mu^+e^-$	97.6 ± 1.3	75.4 ± 1.6
$K^+ \to \pi^+[\mu^-e^+]_{\pi^0}$	93.3 ± 1.2	45.3 ± 1.0
$K^+ \to \pi^+[\mu^+e^-]_{\pi^0}$	92.7 ± 1.2	44.8 ± 1.0

presence of a tagged electron track with $0.95 < E/p < 1.05$, when compared to normalisation $K^+ \to \pi^+\pi^+\pi^-$ events which deposit less energy on average. This is illustrated by figure 7.12 which shows the measured LKr10 and LKr20 trigger efficiency curves superimposed on the E^{3trk}_{LKr} distributions expected (from simulations) for normalisation and signal events. Distributions for the three $K^+ \to \pi\mu e$ and two $K^+ \to \pi[\mu^\mp e^\pm]_{\pi^0}$ decays are similar so averages are displayed. The distribution for $K^+ \to \pi[\mu^\mp e^\pm]_{\pi^0}$ decays is shifted to lower energies because the $e^\pm$, which is responsible for the largest energy deposition, is produced in the intermediate π^0 decay and therefore has a lower average momentum. Measurements of the effective LKr10(20) trigger efficiency for each signal mode are summarised in table 7.5.



(a) (b)

Figure 7.11: L0 LKr20 trigger component efficiency as a function of energy deposited in the LKr associated to the selected 3 charged tracks, E^{3trk}_{LKr}, for normalisation $K^+ \to \pi^+\pi^+\pi^-$ (a) and signal-like (b) events in the full 2017+2018 data-set. The black line and label gives the overall average efficiency and models describing the shape as a function of E^{3trk}_{LKr} are shown.

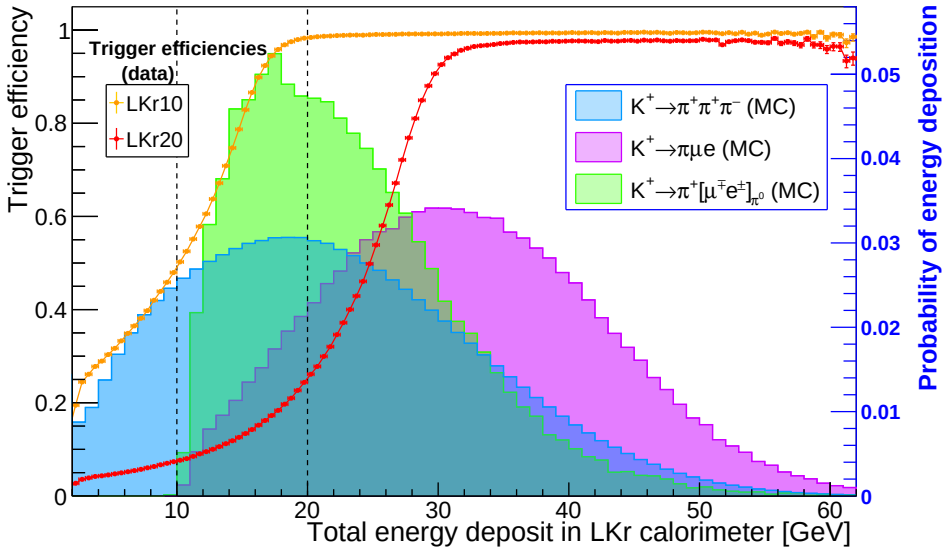
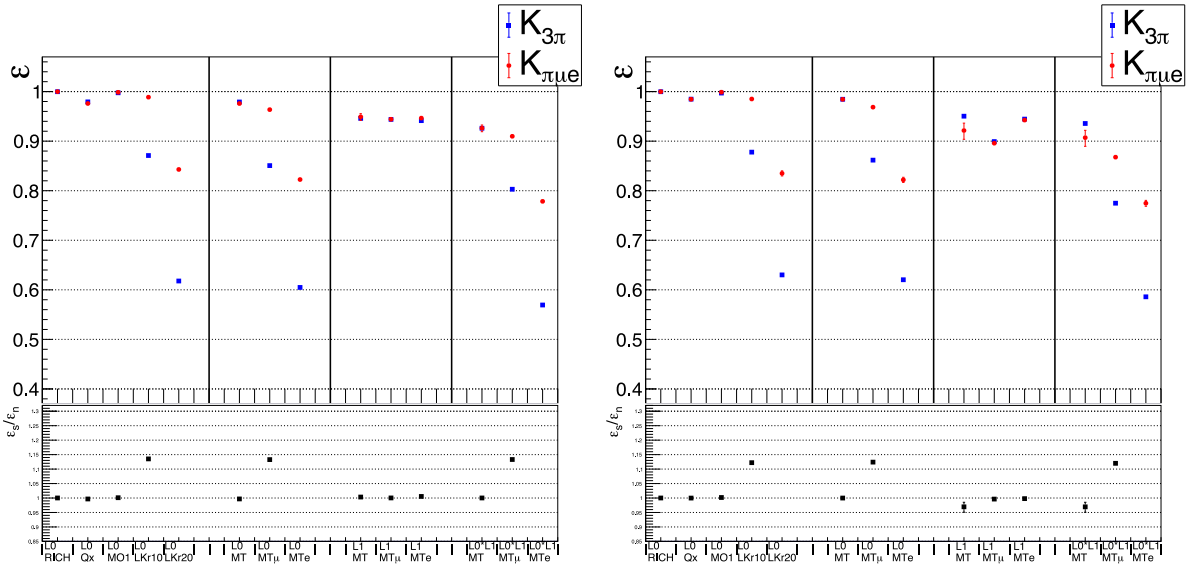


Figure 7.12: Summary plot showing the trigger efficiency of the LKr10 and LKr20 L0 trigger components as a function of the energy deposited by the selected three charged tracks, E^{3trk}_{LKr}, and the expected distribution of E^{3trk}_{LKr} for normalisation $K^+ \to \pi^+\pi^+\pi^-$ and signal $K^+ \to \pi\mu e$ or $K^+ \to \pi[\mu^\mp e^\pm]_{\pi^0}$ events. The E^{3trk}_{LKr} distributions are derived from simulations with corrections applied to E_π according to procedure described in section 7.5.8.1.



(a) (b)

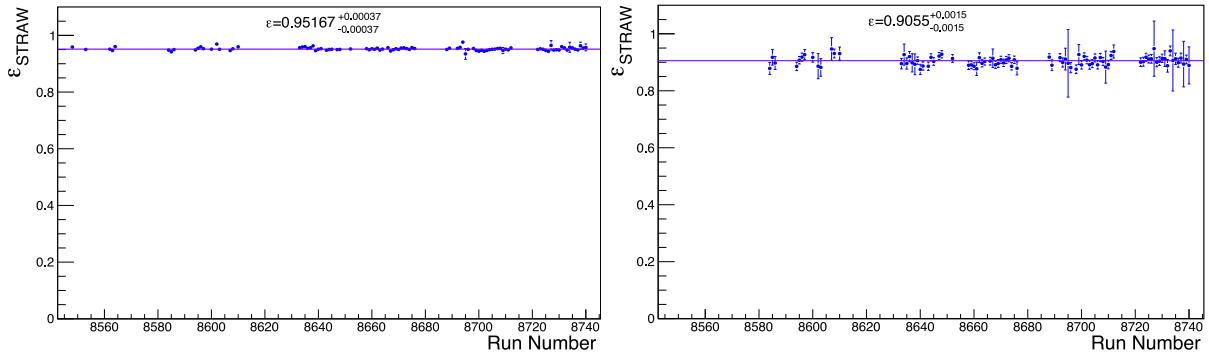
Figure 7.13: Trigger efficiency measurements for each trigger component and their combination for the 2017B (a) and 2018D (b) data samples for normalisation ($\varepsilon_n = \varepsilon_{MT}$) and signal-like ($\varepsilon_s$) events with $\frac{\varepsilon_s}{\varepsilon_n}$ ratio.

7.3.2.3 Studying Changes to the L1 Component of the $MT\mu$ Trigger

The $MT\mu$ L1 trigger definition was changed, from run 9462 (in the 2018B data sample), from KTAG*StrEx to KTAG*StrMT*!LAV while the other two triggers maintain the former L1 definition. This change has implications for the trigger efficiencies and the calculation of sample size and single event sensitivities, see section 7.4.2, of importance because of the significant fraction of events arising from this trigger, see section 7.2.5. The change decreased the output rate from the trigger and, because of independent changes to other triggers, bandwidth was available for the L0 downscaling factor to be reduced from 10 to 5, see figure 7.2b. This significantly increased the fractional contribution of the $MT\mu$ trigger to the search for $K^+ \to \pi\mu e$ for the 2018 data sample relative to that from 2017 where contributions from the MTe and $MT\mu$ triggers were approximately equal.

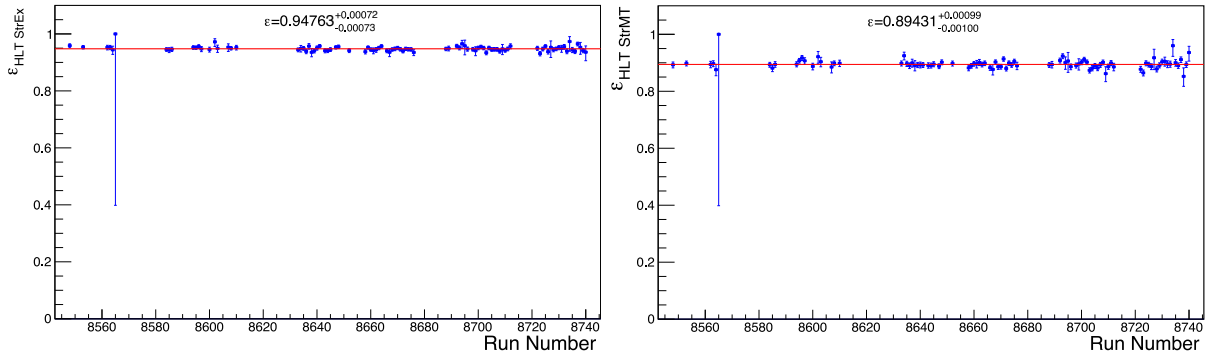
To study the impact of this change the large data samples 17B and 18D (one before and one after the change) are considered. The efficiency breakdown for the trigger components and their combination for these samples are shown in figures 7.13a and 7.13b. In general good agreement can be seen, for both signal-like and normalisation events, except cases including the L1 $MT\mu$ trigger where a clear decrease in efficiency is observed.

To further understand the L1 trigger the HLT (high level trigger) software can be re-run for both data and simulations at analysis level. This allows identical algorithms to those run for the L1 trigger in the PCfarm during data-collection to be applied and a detailed analysis of the behaviour of the L1 trigger to be performed. Importantly in this case it allows the two STRAW L1 algorithms, StrEx and StrMT defined in section 7.1.3, to be run on the same data-sample and the efficiency compared directly. The L1 MT and



(a) (b)

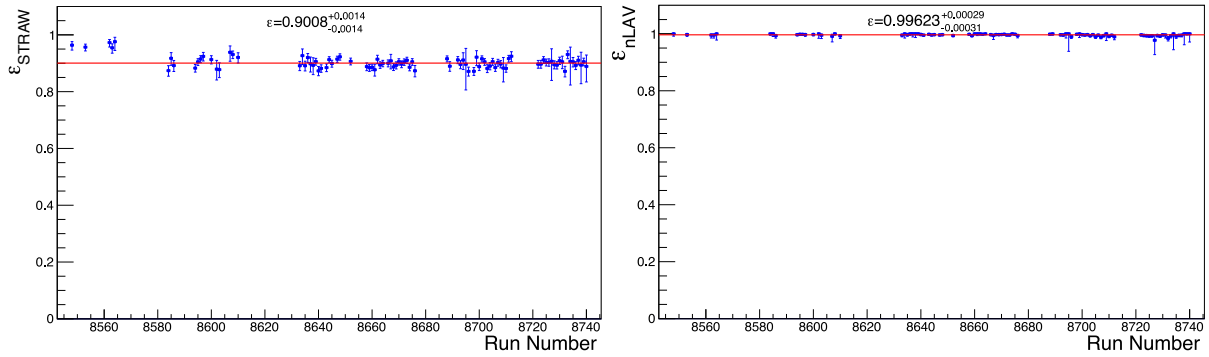
Figure 7.14: The efficiency of L1 STRAW trigger algorithms StrEx (a) and StrMT (b) measured for normalisation events passing the L0 MT and $MT\mu$ triggers respectively and shown as a function of run number for the 2018B data sample.



(a) (b)

Figure 7.15: The efficiency of L1 STRAW trigger algorithms StrEx (a) and StrMT (b) evaluated using the HLT for normalisation events passing the L0 MT and $MT\mu$ triggers respectively and shown as a function of run number for the 2018B data sample.

$MT\mu$ STRAW trigger efficiencies (which use the StrEx and StrMT algorithms respectively) are compared in figures 7.14a and 7.14b using direct measurements from the data and in figures 7.15a and 7.15b using HLT measurements. The efficiency for the StrMT algorithm is consistently seen to be approximately 90% as opposed to the approximately 95% efficiency found for the StrEx algorithm. The measured $MT\mu$ L1 STRAW trigger efficiency for normalisation events for the 2018B sample is shown in figure 7.16a and the effect of the change in algorithm (from run 9462) can be clearly observed. This change in algorithm definition accounts for effectively all of the decrease in $MT\mu$ trigger efficiency. In contrast, the decrease in efficiency by addition of the L1 !LAV algorithm is small because of the LAV veto applied in the common selection (and which is therefore part of both the signal and normalisation selections). The L1 !LAV trigger algorithm efficiency for the $MT\mu$ trigger is shown in figure 7.16b for the 2018B sample as an example, with value of above 99%.



(a) (b)

Figure 7.16: (a) The measured L1 STRAW trigger algorithm efficiency for normalisation events passing the L0 $MT\mu$ trigger as a function of run number for the 2018B data sample. The change in efficiency between the first runs and latter runs is due to the change in algorithm and the efficiencies of approximately 95% and 90% before and after the change respectively are consistent with measurements shown in figures 7.14 and 7.15.
(b) The measured L1 !LAV trigger algorithm efficiency for normalisation events passing the L0 $MT\mu$ trigger as a function of run number for the 2018B data sample.

7.4 Sample Size and Sensitivity

7.4.1 Bad Bursts

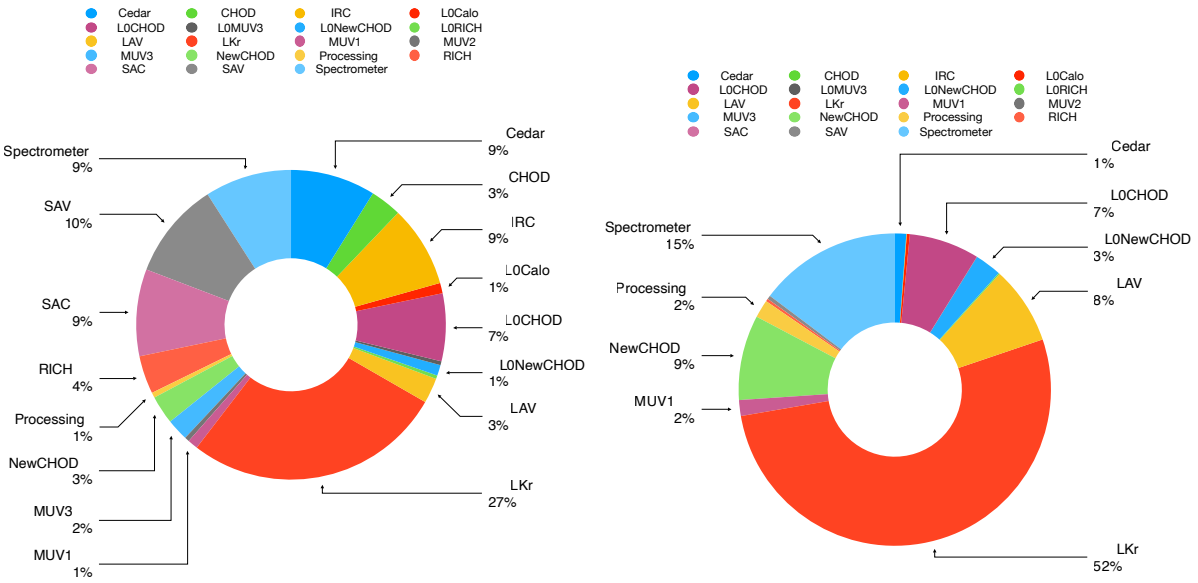
During central processing of NA62 data automatic quality checks are performed for each system and if problems are identified then the relevant bursts or events are marked as 'bad bursts' or 'bad events'. During analysis of the data sample bursts or events marked as bad for systems used in the analysis are rejected. There are 82.5×10^3 bad bursts rejected in the $K^+ \to \pi\mu e$ search and the breakdown of the source of the bad bursts is given in figure 7.17.

7.4.2 Sample Size

The size of the data sample is expressed in terms of N_K, the effective number of K^+ decays in the fiducial decay volume. In general the value of N_K for a given data sample collected with signal trigger(s), T_s, is calculated with respect to the number of selected events N_n of a normalisation decay, K_n, collected with a normalisation trigger T_n, correcting for the normalisation selection acceptance, A_n, the downscaling factors of the signal and normalisation triggers, D_n and D_s, and the normalisation trigger efficiency, ε_n, according to

$$N_K(T_s|T_n, K_n) = \frac{N_n}{\mathcal{B}_n A_n} \frac{D_n}{D_s} \frac{1}{\varepsilon_n} . \tag{7.4.1}$$

For the $K^+ \to \pi\mu e$ search $K_n = K_{3\pi}$, with normalisation selection acceptance $A_{3\pi} = (10.177 \pm 0.002)\%$ and branching ratio $\mathcal{B}_{3\pi}$ given in table 2.2, $T_s = MT + MT\mu + MTe$, $T_n = MT$ and correspondingly $D_n = D_{MT}$ and $D_s = D_{eff}$, the effective trigger downscal-



(a) (b)

Figure 7.17: Detailed breakdown of the systems responsible for bad bursts, for only systems checked for the $K^+ \to \pi\mu e$ search (2017+2018 data), considering (a) exclusive bad bursts, due exclusively to one system alone and (b) all bad bursts inclusively.

ing factor given by

$$D_{eff} = \frac{1}{1 - \left(1 - \frac{1}{D_{MT}}\right)\left(1 - \frac{1}{D_{MT\mu}}\right)\left(1 - \frac{1}{D_{MTe}}\right)}, \tag{7.4.2}$$

as derived in appendix F. Therefore, for this case, the general equation 7.4.1 becomes

$$N_K(MT + MT\mu + MTe|MT, K_{3\pi}) = \frac{N_{3\pi}}{\mathcal{B}_{3\pi}A_{3\pi}} \frac{D_{MT}}{D_{eff}} \frac{1}{\varepsilon_n}. \tag{7.4.3}$$

Since downscaling factors can and do vary during data-taking, see figure 7.2d, the calculation of N_K is strictly only valid for a group of runs with constant trigger configurations and to find the total N_K for a data sample the individual values for each period are summed. Therefore

$$N_K = \sum_i N_K^i = \frac{1}{\mathcal{B}_{3\pi}A_{3\pi}} \cdot \sum_i \left(N_{3\pi}^i \frac{D_{\mathrm{MT}}^i}{D_{\mathrm{eff}}^i \cdot \varepsilon_{\mathrm{n}}^i} \right), \tag{7.4.4}$$

where the index i runs over data-taking periods (consecutive groups of runs) defined by constant trigger downscaling factors. The total number of selected normalisation events collected with the MT trigger in the full data-set is $\sum_i N_{3\pi}^i = 2.73 \times 10^8$.

The uncertainty of N_K is composed of three terms: a statistical term, $\sigma_{N_K}^{stat}$, from the normalisation selection acceptance and measurement of MT trigger efficiency for normalisation $K_{3\pi}$ events (combined this is approximately 0.05%), an external term, $\sigma_{N_K}^{ext}$ from the uncertainty on the branching ratio of the normalisation channel ($\mathcal{B}_{3\pi} =$

$(5.583 \pm 0.024)\%$ [13], a 0.4% relative uncertainty) and a systematic uncertainty $\sigma_{N_K}^{syst}$ arising from differences between MC simulations and data (primarily due to description of the beam momentum and STRAW efficiency) and estimated to be 1%.

The N_K results for each data sample are summarised in table 7.6. For the full 2017+2018 data-set $N_K = (1.321 \pm 0.014) \times 10^{12}$.

Table 7.6: Summary of sample size results showing N_K for data collected with the MT, $MT\mu$ and MTe triggers and normalising to $K^+ \to \pi^+\pi^+\pi^-$ events collected with the MT trigger. The total uncertainties quoted with breakdown into statistical, external and systematic components given in brackets.

Sample	$N_K \times 10^{11}$
17D	$0.3703 \pm 0.0040_{tot}$ $(\pm 0.0002_{stat} \pm 0.0016_{ext} \pm 0.0037_{syst})$
17C	$0.2527 \pm 0.0028_{tot}$ $(\pm 0.0002_{stat} \pm 0.0011_{ext} \pm 0.0025_{syst})$
17B	$1.982 \pm 0.022_{tot}$ $(\pm 0.001_{stat} \pm 0.009_{ext} \pm 0.020_{syst})$
17A	$1.000 \pm 0.011_{tot}$ $(\pm 0.001_{stat} \pm 0.004_{ext} \pm 0.011_{syst})$
18A	$0.2336 \pm 0.0025_{tot}$ $(\pm 0.0002_{stat} \pm 0.0010_{ext} \pm 0.0023_{syst})$
18B	$1.940 \pm 0.021_{tot}$ $(\pm 0.001_{stat} \pm 0.008_{ext} \pm 0.019_{syst})$
18C	$0.2789 \pm 0.0031_{tot}$ $(\pm 0.0003_{stat} \pm 0.0012_{ext} \pm 0.0028_{syst})$
18D	$1.989 \pm 0.022_{tot}$ $(\pm 0.001_{stat} \pm 0.009_{ext} \pm 0.020_{syst})$
18E	$1.523 \pm 0.017_{tot}$ $(\pm 0.001_{stat} \pm 0.007_{ext} \pm 0.015_{syst})$
18F	$1.165 \pm 0.013_{tot}$ $(\pm 0.001_{stat} \pm 0.005_{ext} \pm 0.012_{syst})$
18G	$1.739 \pm 0.019_{tot}$ $(\pm 0.001_{stat} \pm 0.008_{ext} \pm 0.017_{syst})$
18H	$0.7292 \pm 0.0080_{tot}$ $(\pm 0.0005_{stat} \pm 0.0031_{ext} \pm 0.0073_{syst})$
17	$3.605 \pm 0.039_{tot}$ $(\pm 0.001_{stat} \pm 0.016_{ext} \pm 0.036_{syst})$
18	$9.60 \pm 0.10_{tot}$ $(\pm 0.01_{stat} \pm 0.04_{ext} \pm 0.10_{syst})$
17+18	$13.21 \pm 0.14_{tot}$ $(\pm 0.01_{stat} \pm 0.06_{ext} \pm 0.06_{ext} \pm 0.13_{syst})$

7.4.3 Single Event Sensitivity

The analysis ultimately relies on the measurement of, or setting an upper limit on, the branching ratio for a signal decay, $\mathcal{B}_s$, relative to the normalisation channel $K^+ \to \pi^+\pi^+\pi^-$. The number of signal events observed (or upper limit on the number of signal events), n_s, is given by

$$n_s = \frac{\mathcal{B}_s}{\mathcal{B}_{SES}} \ , \tag{7.4.5}$$

where the single event sensitivity, the branching ratio corresponding to observation of one signal event, is defined as

$$\mathcal{B}_{SES} = \frac{1}{N_K \varepsilon_s A_s} = \frac{1}{A_s} \frac{\mathcal{B}_{3\pi} A_{3\pi}}{N_{3\pi}} \frac{D_{eff}}{D_n} \frac{\varepsilon_n}{\varepsilon_s} \ , \tag{7.4.6}$$

where A_s is the signal acceptance, and ε_s is the signal trigger efficiency, see also appendix F.2. This definition has the same form as equation 5.1.2 from the $K^+ \to \pi^+\nu\bar{\nu}$ analysis, although here no random veto term is added explicitly since there is no significant

additional random veto introduced by the signal selection with respect to the normalisation selection. The single event sensitivities for each data-taking period are determined individually according to

$$\mathcal{B}^i_{SES} = \frac{1}{N^i_K A_s \varepsilon^i_s} = \mathcal{B}_{3\pi} \frac{A_{3\pi}}{A_s} \frac{1}{N^i_{3\pi}} \frac{D^i_{eff}}{D^i_{MT}} \frac{\varepsilon^i_n}{\varepsilon^i_s} \ , \tag{7.4.7}$$

and combined for the full data-set according to

$$\mathcal{B}_{SES} = \left[\sum_i (\mathcal{B}^i_{SES})^{-1} \right]^{-1} \ . \tag{7.4.8}$$

The uncertainty on $\mathcal{B}_{SES}$ primarily arises from the uncertainty on N_K which is of order 1% (total) and the signal trigger efficiency which is of order 3% (added to statistical term). To this is added (to the statistical term) the uncertainty on the signal acceptance which has relative uncertainty of order 0.5% and (to the systematic term) a factor accounting for fractional differences in results calculated using the measured effective trigger efficiency (using signal-like events passing $K^+ \to \pi\mu e$ selection in the control region) and the emulated trigger efficiency (as applied as event weights to simulated events, see section 7.5.8.1), which is channel-dependent and of order 1%.

Using the acceptances displayed in table 7.3 and the N_K and trigger efficiency results for each data-taking period, reported in tables 7.6 and 7.4 respectively (and accounting for channel-specific effects encapsulated in event weights), the $\mathcal{B}_{SES}$ is calculated for each period and combined to give an overall figure for the search for each $K^+ \to \pi\mu e$ decay mode in the 2017+2018 data-set, shown in table 7.7.

In the case of decay $K^+ \to \pi^+\mu^+e^-$ the sensitivity has clearly not reached the level of the previous world best result (see table 7.1), and this is without applying the Dalitz rejection cut that would be required to reject background in the signal region but decreases acceptance significantly. Since the previous limit for this mode is from a dedicated experiment it is understandable that for this mode NA62 is not yet able to compete, however the fact that the same order of magnitude of single event sensitivity is achieved is promising for potential future NA62 work including data taken after the current CERN long shutdown (in 2021–2024), see section 7.7.4.

For the $\pi^0 \to \mu^\mp e^\pm$ decay modes N_K must be replaced in equation 7.4.6 with $N_{\pi^0} = N_K \mathcal{B}(K^+ \to \pi^+\pi^0)$ which decreases the sensitivity to these modes by approximately a factor of five, in addition with the acceptance loss by approximately a factor of two. Combined these effects mean the sensitivity for these modes reaches a similar level to the previous charged-combined upper limits ($\mathcal{B}(\pi^0 \to \mu^\mp e^\pm) < 36 \times 10^{-11}$ @90% CL [13] [173]) and potential for improvement on charge-specific limits [13] [171] [263] exists.

For the LNV and LFV decay $K^+ \to \pi^-\mu^+e^+$ and LFV decay $K^+ \to \pi^+\mu^-e^+$ the sensitivity achieved is sufficient for a significant improvement over previous world-best limits which are $\mathcal{O}(50 \times 10^{-11})$. Because of the above facts these two modes are the primary focus of the remainder of the chapter, with secondary focus on the $\pi^0 \to \mu^-e^+$ channel. To fully evaluate the sensitivity of the search the background must be understood

and only if this is held to an acceptably small level can improved results be reached. The evaluation of background expectations is the subject of section 7.5, using these results the ultimate sensitivity for the searches is derived in section 7.6.

Table 7.7: Summary of the single event sensitivities for the searches for $K^+ \to \pi\mu e$ decays in the 2017+2018 data-set.

Case	$\mathcal{B}_{SES} \times 10^{-11}$
$K^+ \to \pi^-\mu^+e^+$	$1.836 \pm 0.084_{tot}$ $(\pm 0.046_{stat} \pm 0.008_{ext} \pm 0.070_{syst})$
$K^+ \to \pi^+\mu^-e^+$	$1.449 \pm 0.049_{tot}$ $(\pm 0.036_{stat} \pm 0.006_{ext} \pm 0.032_{syst})$
$K^+ \to \pi^+\mu^+e^-$	$1.461 \pm 0.061_{tot}$ $(\pm 0.036_{stat} \pm 0.006_{ext} \pm 0.049_{syst})$
$K^+ \to \pi^+[\mu^-e^+]_{\pi^0}$	$14.00 \pm 0.94_{tot}$ $(\pm 0.35_{stat} \pm 0.08_{ext} \pm 0.86_{syst})$
$K^+ \to \pi^+[\mu^+e^-]_{\pi^0}$	$15.9 \pm 1.1_{tot}$ $(\pm 0.4_{stat} \pm 0.1_{ext} \pm 1.0_{syst})$

7.5 Background Studies

7.5.1 Calculation of Expected Number of Background Events

Background studies are performed using simulations (with one exception described in section 7.5.6). If n^{obs}_{MC} events are observed to pass the signal selection the number of events expected in the data sample is predicted to be

$$N_{bg} = n^{obs}_{MC} f_{scale} w_{MC} w_t w_p w_S w_{PID} \,, \tag{7.5.1}$$

where f_{scale} is the scale factor used to normalise the simulation sample to the data, defined as

$$f_{scale} = \frac{N_K \mathcal{B}}{N_{MC}} \,, \tag{7.5.2}$$

where N_{MC} is the number of simulated events with a K^+ decay, with branching ratio $\mathcal{B}$, in the FV ($105 < z < 180\,\text{m}$). The w_x weight terms in equation 7.5.1 encode data-driven corrections for special MC simulation settings, trigger efficiency, momentum, STRAW inefficiency and PID respectively and are discussed in sections 7.5.7.1, 7.5.8.1, 7.5.8.2, 7.5.8.3 and 7.5.4. These weight factors are applied event-by-event so the total number of background events predicted may be more accurately described as

$$N_{bg} = f_{scale} \sum_i w^i_{MC} w^i_t w^i_p w^i_S w^i_{PID} \,. \tag{7.5.3}$$

7.5.2 Background Mechanisms

Potential backgrounds for the search for $K^+ \to \pi\mu e$ decays arise from two types of mechanisms: decays in flight (DIF) and misidentification (misID). The potential background mechanisms studied are summarised in table 7.8 where decays in flight, for example, $\pi^\pm \to \mu^\pm\nu_\mu$ are displayed with shorthand $\pi^\pm \to \mu^\pm$ and misidentification of, for example, $\pi^\pm$ as an $e^\pm$ is shown with shorthand $\pi^\pm \Rightarrow e^\pm$.

Table 7.8: Summary of background mechanisms for $K^+ \to \pi^\mp \mu^\pm e$ searches.

Background Process	Abbr.	Case	π^- Channel (FS = $\pi^-\mu^+e^+$)	μ^- Channel (FS = $\pi^+\mu^-e^+$)
$K^+ \to \pi^+\pi^+\pi^-$	$K_{3\pi}$	A	$\pi^+ \to \mu^+$ DIF + $\pi^+ \Rightarrow e^+$ misID	$\pi^- \to \mu^-$ DIF + $\pi^+ \Rightarrow e^+$ misID
		B	$\pi^+ \Rightarrow \mu^+$ misID + $\pi^+ \Rightarrow e^+$ misID	$\pi^- \Rightarrow \mu^-$ misID + $\pi^+ \Rightarrow e^+$ misID
		C	$\pi^+ \to \mu^+$ DIF + $\pi^+ \to e^+$ DIF	$\pi^- \to \mu^-$ DIF + $\pi^+ \to e^+$ DIF
		D	$\pi^+ \Rightarrow \mu^+$ misID + $\pi^+ \to e^+$ DIF	$\pi^- \Rightarrow \mu^-$ misID + $\pi^+ \to e^+$ DIF
$K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$	$K_{2\pi D}$	A	$\pi^+ \to \mu^+$ DIF + $e^- \Rightarrow \pi^-$ misID	
		B	$\pi^+ \Rightarrow \mu^+$ misID + $e^- \Rightarrow \pi^-$ misID	
		C		$e^- \Rightarrow \mu^-$ misID
		D	$\pi^+ \Rightarrow e^+$ misID + $e^+ \Rightarrow \mu^+$ misID + $e^- \Rightarrow \pi^-$ misID	$\pi^+ \Rightarrow e^+$ misID + $e^+ \Rightarrow \pi^+$ misID + $e^- \Rightarrow \mu^-$ misID
		E	$\pi^+ \to e^+$ DIF + $e^+ \Rightarrow \mu^+$ misID + $e^- \Rightarrow \pi^-$ misID	$\pi^+ \to e^+$ DIF + $e^+ \Rightarrow \pi^+$ misID + $e^- \Rightarrow \mu^-$ misID
$K^+ \to \mu^+\nu_\mu[e^+e^-\gamma]_{\pi^0}$	$K_{\mu 3D}$	A	$e^- \Rightarrow \pi^-$ misID	
		B		$\pi^+ \Rightarrow e^+$ misID + $e^- \Rightarrow \mu^-$ misID
		C	$\mu^+ \Rightarrow e^+$ misID + $e^+ \Rightarrow \mu^+$ misID + $e^- \Rightarrow \pi^-$ misID	$\mu^+ \Rightarrow e^+$ misID + $e^+ \Rightarrow \pi^+$ misID + $e^- \Rightarrow \mu^-$ misID
$K^+ \to e^+\nu_e[e^+e^-\gamma]_{\pi^0}$	K_{e3D}	A	$e^+ \Rightarrow \mu^+$ misID + $e^- \Rightarrow \pi^-$ misID	$e^+ \Rightarrow \pi^+$ misID + $e^- \Rightarrow \mu^-$ misID
$K^+ \to \pi^+\pi^-\mu^+\nu_\mu$	$K_{\mu 4}$	A	$\pi^+ \Rightarrow e^+$ misID	
		B		$\pi^- \to \mu^-$ DIF + $\pi^+ \Rightarrow e^+$ misID
		C	$\pi^+ \to \mu^+$ DIF + $\mu^+ \Rightarrow e^+$ misID	$\pi^- \to \mu^-$ DIF + $\mu^+ \Rightarrow e^+$ misID
		D	$\pi^+ \Rightarrow \mu^+$ misID + $\mu^+ \Rightarrow e^+$ misID	$\pi^- \Rightarrow \mu^-$ misID + $\mu^+ \Rightarrow e^+$ misID
$K^+ \to \pi^+\pi^-e^+\nu_e$	K_{e4}	A	$\pi^+ \to \mu^+$ DIF	$\pi^- \to \mu^-$ DIF
		B	$\pi^+ \Rightarrow \mu^+$ misID	$\pi^- \Rightarrow \mu^-$ misID
		C	$\pi^+ \Rightarrow e^+$ misID + $e^+ \Rightarrow \mu^+$ misID	
		D	$\pi^+ \to e^+$ DIF + $e^+ \Rightarrow \mu^+$ misID	
$K^+ \to \pi^+\mu^+\mu^-$	$K_{\pi\mu\mu}$	A	$\pi^+ \Rightarrow e^+$ misID + $\mu^- \Rightarrow \pi^-$ misID	$\pi^+ \Rightarrow e^+$ misID + $\mu^+ \Rightarrow \pi^+$ misID
		B	$\pi^+ \Rightarrow e^+$ misID + $\mu^- \Rightarrow \pi^-$ misID	$\pi^+ \Rightarrow e^+$ misID + $\mu^+ \Rightarrow \pi^+$ misID
		C		$\mu^+ \Rightarrow e^+$ misID
$K^+ \to \pi^+e^+e^-$	$K_{\pi ee}$	A	$\pi^+ \to \mu^+$ DIF + $e^- \Rightarrow \pi^-$ misID	
		B	$\pi^+ \Rightarrow \mu^+$ misID + $e^- \Rightarrow \pi^-$ misID	
		C	$e^+ \Rightarrow \mu^+$ misID	$e^- \Rightarrow \mu^-$ misID
		D	$\pi^+ \Rightarrow e^+$ misID + $e^- \Rightarrow \pi^-$ misID + $e^+ \Rightarrow \mu^+$ misID	$\pi^+ \Rightarrow e^+$ misID + $e^+ \Rightarrow \pi^+$ misID + $e^- \Rightarrow \mu^-$ misID
		E	$\pi^+ \to e^+$ misID + $e^- \Rightarrow \pi^-$ misID + $e^+ \Rightarrow \mu^+$ misID	$\pi^+ \to e^+$ misID + $e^+ \Rightarrow \pi^+$ misID + $e^- \Rightarrow \mu^-$ misID
$K^+ \to \mu^+\nu_\mu\mu^+\mu^-$	$K_{\mu\nu\mu\mu}$	A	$\mu^+ \Rightarrow e^+$ misID + $\mu^- \Rightarrow \pi^-$ misID	$\mu^+ \Rightarrow e^+$ misID + $\mu^+ \Rightarrow \pi^+$ misID
$K^+ \to e^+\nu_e\mu^+\mu^-$	$K_{e\nu\mu\mu}$	A	$\mu^- \Rightarrow \pi^-$ misID	$\mu^+ \Rightarrow \pi^+$ misID
$K^+ \to \mu^+\nu_\mu e^+e^-$	$K_{\mu\nu ee}$	A	$e^- \Rightarrow \pi^-$ misID	
		B		$\mu^+ \Rightarrow \pi^+$ misID
$K^+ \to e^+\nu_e e^+e^-$	$K_{e\nu ee}$	A	$e^+ \Rightarrow \mu^+$ misID + $e^- \Rightarrow \pi^-$ misID	$e^+ \Rightarrow \pi^+$ misID + $e^- \Rightarrow \mu^-$ misID

7.5.3 Decay In Flight Studies

For the $K^+ \to \pi\mu e$ analysis the only important decay in flight processes are pion decays $\pi^\pm \to \ell^\pm\nu_\ell$ (where $\ell = e, \mu$). The probability of $\pi^+ \to \mu^+\nu_\mu$ and $\pi^+ \to e^+\nu_e$ decays as a function of the momentum of the parent pion is displayed in figures 7.18a and 7.18b. As expected the relationship seen in the two cases is similar. Of interest for the analysis is the amount of energy/momentum that a neutrino can carry away (since in the analysis the total momentum must be within 2.5 GeV/c of the beam momentum and the transverse momentum less than 0.030 GeV/c), this is shown by figures 7.18c and 7.18d. The DIF probabilities for $\pi^\pm$ are momentum-dependent, due to the relationship between momentum and mean free path due to the fixed mean $\pi^\pm$ lifetime, with a probability of 7.5% on average, with suppression by approximately four orders of magnitude for $\pi^\pm \to e^\pm\nu_e$ decays with respect to $\pi^+ \to \mu^+\nu_\mu$ decays due to the relative size of the branching ratios, [13]

$$\mathcal{B}(\pi^+ \to \mu^+\nu_\mu) = 0.9998770 \pm 0.0000004\,, \tag{7.5.4}$$

$$\mathcal{B}(\pi^+ \to e^+\nu_e) = (1.230 \pm 0.004) \times 10^{-4}\,. \tag{7.5.5}$$

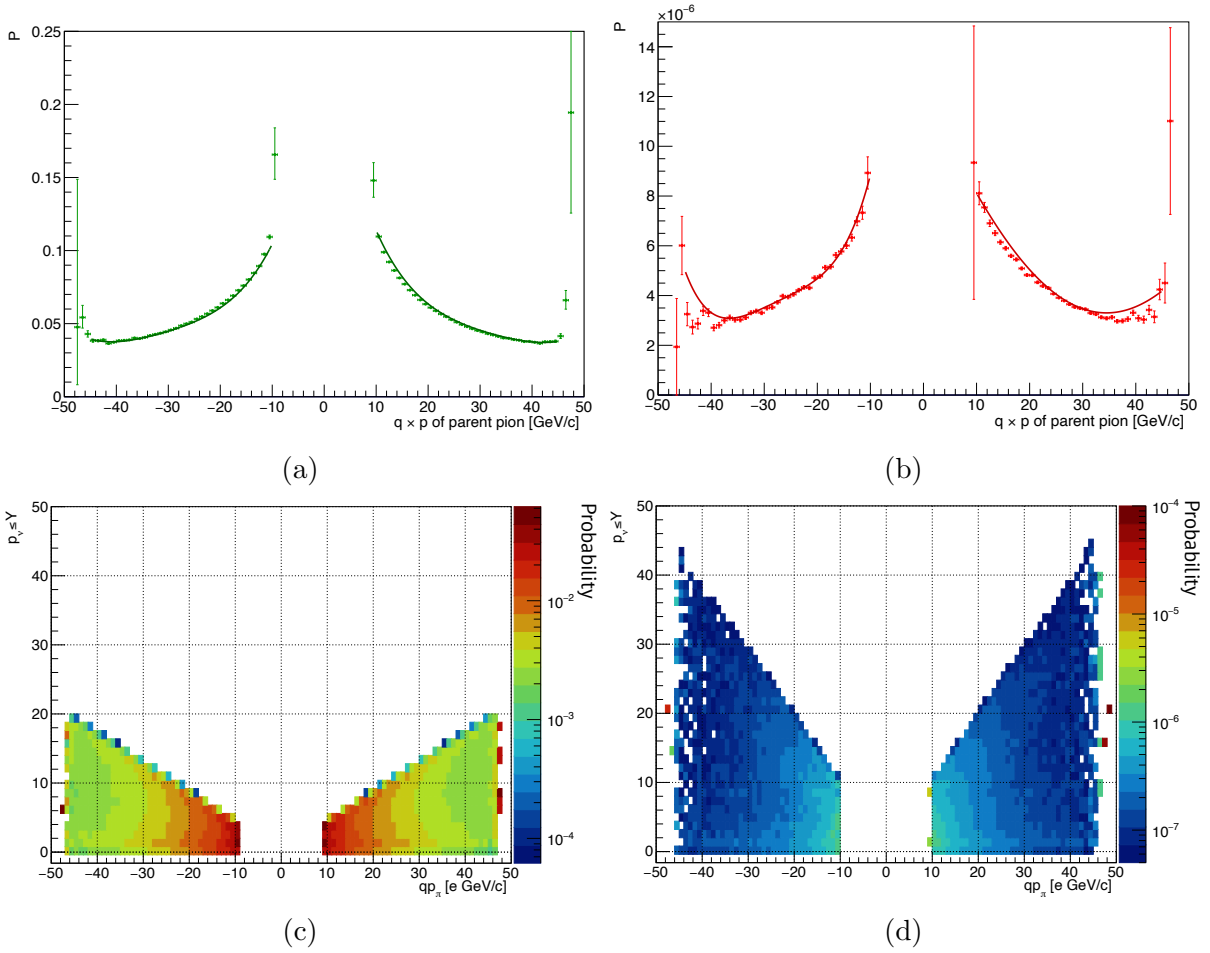


Figure 7.18: Probability of (a) $\pi^+ \to \mu^+\nu_\mu$ or (b) $\pi^+ \to e^+\nu_e$ decays in flight (in the FV) as a function of the charge multiplied by the momentum of the parent $\pi^\pm$, qp. The DIF probability is shown in the two dimensional plane of neutrino momentum, p_ν, and qp for (c) $\pi^+ \to \mu^+\nu_\mu$ or (d) $\pi^+ \to e^+\nu_e$ decays in flight. These results are calculated by studying 'true' information from simulations before reconstruction from (for a and c) standard $K^+ \to \pi^+\pi^+\pi^-$ simulations and (for b and d) a special simulation sample with modified $\pi^\pm$ branching ratios so $Br(\pi^\pm \to \mu^\pm\nu_\mu) = Br(\pi^\pm \to e^\pm\nu_e) = 0.5$ (the corresponding bias is accounted for when calculating the probability with the 'true' branching ratios).

7.5.4 Particle Identification and Misidentification Studies

7.5.4.1 Introduction and Overview

The objective of the study is to find the probability that, given a particle has a true identity $X^{\pm}$, an identity $Y^{\pm}$ is assigned using the PID selection applied in the $K^{+} \to \pi\mu e$ selection, $P(X^{\pm}_{ID}|Y^{\pm}_{true})$ for $X^{\pm}, Y^{\pm} \in [\pi^{\pm}, \mu^{\pm}, e^{\pm}]$. In general these probabilities depend most strongly on the momentum of the charged track and may differ between charges, for this reason studies are generally performed as a function of the track charge times momentum, qp.

For the analysis four key misidentification mechanisms enter, to construct background models the corresponding misidentification probabilities must be evaluated:

1. $\pi^{\pm} \Rightarrow e^{\pm}$ misidentification (LKr only): the probability of a $\pi^{\pm}$ track having LKr $0.95 < E/p < 1.05$. This is momentum dependent with value $P \sim 4 \times 10^{-3}$.
2. $e^{\pm} \Rightarrow \pi^{\pm}$ misidentification (LKr only): the probability of a $e^{\pm}$ track having LKr $E/p < 0.9$. This is momentum dependent with value $P \sim 0.02$.
3. Accidental association of tracks with MUV3 candidates. This is one component of $e^{\pm} \Rightarrow \mu^{\pm}$ misidentification (the other being $E/p < 0.2$ with probability $\sim 10^{-4}$). This has significant momentum and (x, y) dependence with value $P \leq 10^{-3}$.
4. $\pi^{\pm} \Rightarrow \mu^{\pm}$ misidentification (LKr + MUV3): the probability of a $\pi^{\pm}$ track having $E/p < 0.2$ and with exactly 1 matching in-time MUV3 candidate. This is momentum dependent and charge asymmetric with value $P \sim 3 \times 10^{-3}$.

Each of these are addressed in detail in sections 7.5.4.3, 7.5.4.4, 7.5.4.5 and 7.5.4.6 respectively. To measure these quantities, and the PID efficiency (the probability of correctly identifying a given particle species), dedicated studies have been performed using control samples of data for all but the last case where simulations must be used. The selections used to isolate these control samples are given in appendix G. These selections are designed to isolate samples of a particular particle species of interest, with as little contamination as possible from other particle species (see section 7.5.4.2) while using selection criteria which do not include the detectors used for the PID selection.

It is important to use control selections which match the signal selection as closely as possible to avoid introducing bias into background estimates which use PID models constructed for a different 'PID selection environment'. For example measurements using control samples with and without use of photon veto requirements are found to change results by $\sim 5\%$ and introducing exclusion requirements on the distance between an LKr cluster and a dead cell may change results by up to 20% (found to be most significant for $e^{\pm} \Rightarrow \pi^{\pm}$ misID measurements). Pileup can also create noticeable differences, especially in $\mu^{\pm}$ ID or misID scenarios using the MUV3 and in simulation studies the effect of pileup is emulated based on a data-driven model.

Data collected with the minimum bias multi-track (MT) and general purpose minimum bias (CTRL) triggers have been used for these studies to remove bias from PID

requirements applied at trigger level. Results from the two trigger streams are found to be compatible but more statistics are available when using the MT trigger, this also has the advantage that the same data-set as used for the main analysis can be used.

7.5.4.2 Control Sample Contamination studies

Where possible data-driven measurements are used and it is therefore important to understand the purity of the selected control sample. Contamination of the control samples limits the accuracy and precision of the PID studies. Consider the study of misidentification process $X_{true} \Rightarrow Y_{ID}$. If the selected control sample of X is contaminated by species Y then assigning a track the identity Y could mean a true Y particle has been correctly identified, instead of observation of the $X_{true} \Rightarrow Y_{ID}$ misidentification mechanism.

The contamination is studied by applying the control selections to a variety of simulation samples of possible 'background' decay processes and the number of events passing each selection is determined. Then, taking into account the relative branching ratio of the 'signal' control sample mode and the potential 'background' simulated mode the level at which this 'background' can contribute to the selected control sample can be established. Specifically the relative contamination or contamination factor $C(K_X, K_c)$ is calculated for each pair of control decay $K_c = K_{3\pi}, K_{\mu 2}, K_{2\pi D}$ and potential background K_X. This contamination factor is defined as

$$C(K_X, K_c) = \frac{\mathcal{B}(K_X)A_c(K_X)}{\mathcal{B}(K_c)A_c(K_c)}\,, \tag{7.5.6}$$

where $\mathcal{B}(K_X)$ and $\mathcal{B}(K_c)$ are the branching ratios of the control and background decays respectively and A_c is the acceptance for the control sample selection for the control sample decay K_c.

The control sample selection acceptances are shown in table 7.9 and the contamination factors are given in table 7.10. When studying $\pi^{\pm} \Rightarrow e^{\pm}$ or $e^{\pm} \Rightarrow \pi^{\pm}$ misID the MUV3 conditions are common between the two particle identities and therefore only the LKr PID component must be studied. This means a MUV3 veto condition can be applied in the control sample selections to isolate $\pi^{\pm}$ (from $K^+ \to \pi^+\pi^+\pi^-$ decays) and $e^{\pm}$ (from $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ decays). The stringent $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ control sample selection results in limited statistics to test the contamination. In table 7.10, cases where no events pass the control selection for a 'background' simulation sample are included as if 1 event was observed with 100% uncertainty. This approximation is used in a number of cases for the contamination study of the $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ selection, importantly in the cases of $K^+ \to \mu^+\nu_\mu$, $K^+ \to \pi^+\pi^0$ and $K^+ \to \pi^+\pi^0\pi^0$ which are likely smaller than the approximate results. However, in all cases the level of contamination identified is sufficiently small to allow the required misidentification studies with negligible bias from contamination.

The contamination has also been studied as a function of the momentum of the track. This is important because of the different momentum spectra of charged particle tracks from different decays and potential differences in rejection of 'background' processes as a function of momentum using (primarily) kinematic control selections. The results

are shown in figures 7.19a, 7.19c and 7.19b for the $K^+ \to \pi^+\pi^+\pi^-$, $K^+ \to \mu^+\nu_\mu$ and $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ control samples (for studies of $\pi^\pm$, μ^+ and $e^\pm$ respectively). A clean sample of $e^\pm$ is difficult to obtain, to select the $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ control sample with sufficiently low contamination requires a stringent selection with a small acceptance. An alternative case, selecting a control sample of single-track $K^+ \to \pi^0 e^+\nu_e$ decays, was considered where a larger acceptance of $(5.07 \pm 0.04)\%$ was achieved but which suffers from more significant and momentum dependent contamination (from $K^+ \to \pi^+\pi^0$ which cannot be reduced further kinematically), as shown by figure 7.19d. This alternative case results in an e^+ sample with contamination dominated by π^+ at higher momenta, in contrast using the $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ control sample gives both e^+ and e^- samples with contamination primarily from other decays with $e^\pm$ in the final state.

Table 7.9: Control sample selection acceptances.

Case	Acceptance
$K^+ \to \pi^+\pi^+\pi^-$ (with MUV3 veto)	$(9.80567 \pm 0.0025)\%$
$K^+ \to \mu^+\nu_\mu$ (purely kinematic selection)	$(16.43970 \pm 0.00040)\%$
$K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ (with MUV3 veto)	$(0.4828 \pm 0.0010)\%$

Table 7.10: Control sample contamination factor results. The control sample selections for $K_{3\pi} = K^+ \to \pi^+\pi^+\pi^-$ and $K_{2\pi D} = K^+ \to \pi^+[e^+e^-]_{\pi^0}$ include a MUV3 veto while a purely kinematic selection is used for $K_{\mu 2} = K^+ \to \mu^+\nu_\mu$.

Decay	$C(K_X, K_{3\pi})$	$C(K_X, K_{\mu 2})$	$C(K_X, K_{2\pi D})$
$K^+ \to \pi^+\pi^+\pi^-$	1	$(4.76 \pm 0.13) \times 10^{-6}$	$(3.3 \pm 3.3) \times 10^{-5}$
$K^+ \to \mu^+\nu_\mu$	$(5.03 \pm 0.80) \times 10^{-5}$	1	$(5.9 \pm 5.9) \times 10^{-4}$
$K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$	$(2.16 \pm 0.46) \times 10^{-7}$	$(6.45 \pm 0.26) \times 10^{-7}$	1
$K^+ \to \pi^+\pi^0$	$(2.03 \pm 0.42) \times 10^{-6}$	$(6.50 \pm 0.18) \times 10^{-5}$	$(4.1 \pm 4.1) \times 10^{-4}$
$K^+ \to \pi^0\mu^+\nu_\mu$	$(2.72 \pm 0.30) \times 10^{-6}$	$(2.737 \pm 0.074) \times 10^{-6}$	$(1.5 \pm 1.5) \times 10^{-5}$
$K^+ \to \pi^0 e^+\nu_e$	$(3.39 \pm 0.82) \times 10^{-6}$	$(4.45 \pm 0.22) \times 10^{-6}$	$(6.5 \pm 2.5) \times 10^{-4}$
$K^+ \to \mu^+\nu_\mu[e^+e^-\gamma]_{\pi^0}$	$(4.8 \pm 1.8) \times 10^{-8}$	$(3.04 \pm 0.34) \times 10^{-8}$	$(3.2 \pm 3.2) \times 10^{-4}$
$K^+ \to e^+\nu_e[e^+e^-\gamma]_{\pi^0}$	$(2.3 \pm 1.6) \times 10^{-8}$	$(4.15 \pm 0.52) \times 10^{-8}$	$(8.95 \pm 0.79) \times 10^{-4}$
$K^+ \to \pi^+\pi^0\pi^0$	$(2.5 \pm 1.1) \times 10^{-8}$	$(1.39 \pm 0.19) \times 10^{-6}$	$(2.3 \pm 2.3) \times 10^{-4}$
$K^+ \to \pi^+\pi^-\mu^+\nu_\mu$	$(2.2 \pm 1.4) \times 10^{-8}$	$(1.12 \pm 0.72) \times 10^{-9}$	$(2.5 \pm 2.5) \times 10^{-8}$
$K^+ \to \pi^+\pi^-e^+\nu_e$	$(1.310 \pm 0.011) \times 10^{-6}$	$(3.458 \pm 0.078) \times 10^{-9}$	$(5.92 \pm 0.20) \times 10^{-5}$
$K^+ \to \pi^+\mu^+\mu^-$	$(3.6 \pm 3.6) \times 10^{-12}$	$(8.0 \pm 1.3) \times 10^{-12}$	$(1.7 \pm 1.7) \times 10^{-6}$
$K^+ \to \pi^+e^+e^-$	$(5.2 \pm 5.2) \times 10^{-12}$	$(1.534 \pm 0.050) \times 10^{-9}$	$(1.056 \pm 0.048) \times 10^{-5}$
Total	$(8.08 \pm 0.91) \times 10^{-5}$	$(7.91 \pm 0.18) \times 10^{-5}$	$(2.90 \pm 0.48) \times 10^{-3}$
Main Contaminants	$K_{\mu 2}$	$K_{2\pi}$	K_{e3} , K_{e3D}

7.5.4.3 $\pi^\pm \Rightarrow e^\pm$ Misidentification Studies

A control sample of $K^+ \to \pi^+\pi^+\pi^-$ events is selected without using the LKr (but applying a MUV3 veto condition to reject events with $\pi^\pm \to \mu^\pm\nu_\mu$ decays in flight), as discussed above and with full selection details given in appendix G. The contamination of the $\pi^\pm$ sample from other particles is less than 10^{-4} (and less than 10^{-5} from $e^\pm$, see table 7.10 and figure 7.19a), and because the probability of $\pi^\pm \to e^+\nu_e$ decays in flight is $< 10^{-5}$

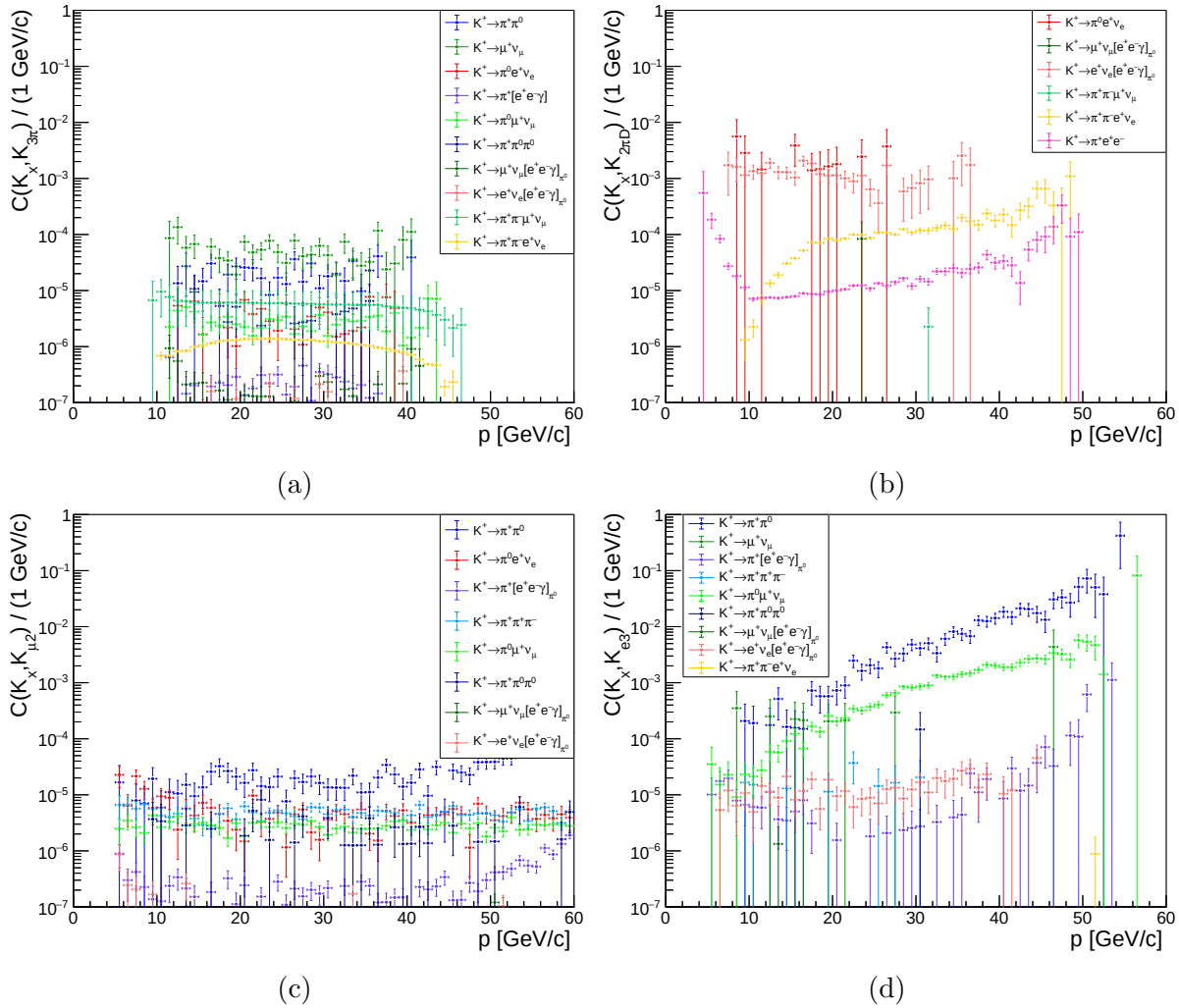
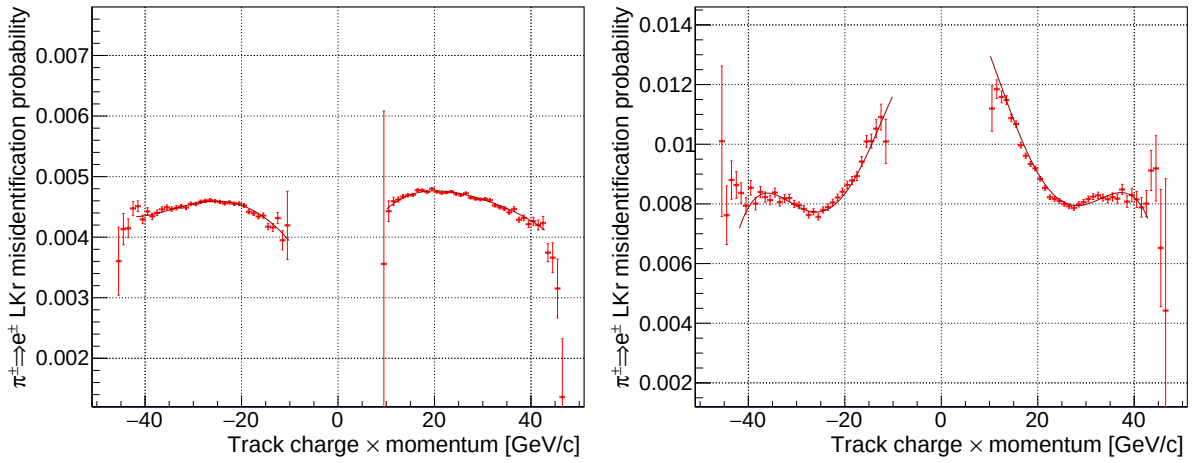


Figure 7.19: Particle identification study control sample contamination factors (defined by equation 7.5.6) as a function of track momentum, p, for: a) $K^+ \to \pi^+\pi^+\pi^-$ providing $\pi^\pm$ samples, c) $K^+ \to \mu^+\nu_\mu$ providing a μ^+ sample b) $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ providing $e^\pm$ samples and d) $K^+ \to \pi^0 e^+\nu_e$, a selection with higher acceptance providing an alternative e^+ sample but suffering from larger contamination.



(a) (b)

Figure 7.20: Probability of $\pi^{\pm} \Rightarrow e^{\pm}$ misidentification as a function of track charge times momentum for the control sample of $\pi^{\pm}$ tracks from a $K^{+} \to \pi^{+}\pi^{+}\pi^{-}$ sample selected in: (a) data collected with the minimum-bias three track (MT) trigger and (b) $K^{+} \to \pi^{+}\pi^{+}\pi^{-}$ simulations. Note that the maximum of the y axis scale has been doubled in (b). The models describing the distributions, from fits to the measurements, are shown by the dark red lines.

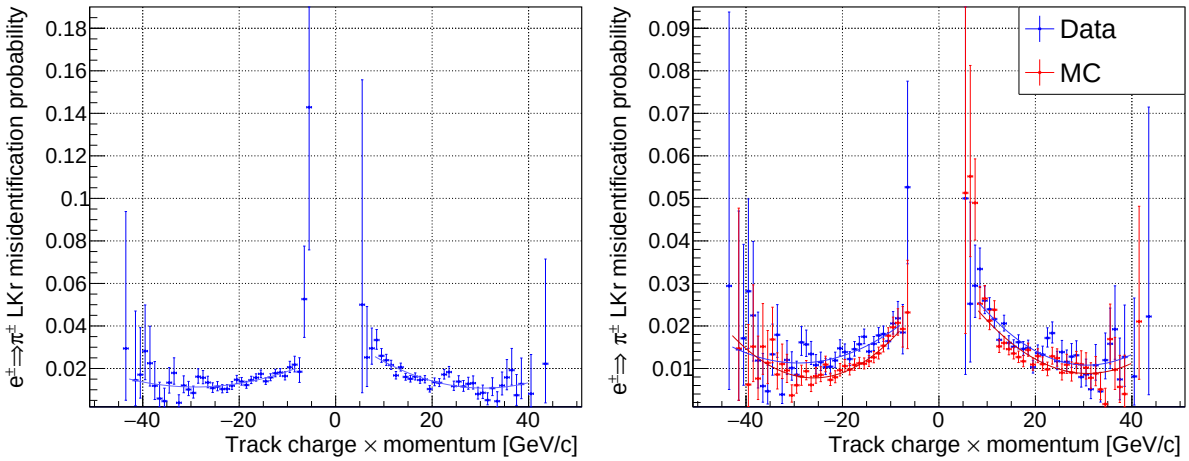
in each momentum bin (see figure 7.18b) this study is dominated by genuine cases of $\pi^{\pm} \Rightarrow e^{\pm}$ misidentification.

The precise measurement made is the $\pi^{\pm} \Rightarrow e^{\pm}$ misidentification probability (of $0.95 < E/p < 1.05$ and exactly one LKr cluster within 100 mm of the impact point at the LKr front plane and 5 ns of the track time) given that there were exactly 0 in-time MUV3 associations: $P(\pi^{\pm} \Rightarrow e^{\pm}|0\,\text{MUV3}) = P(0.95 < E/p < 1.05|0\,\text{MUV3})$ as a function of the charge times momentum of the reconstructed track, for results see figure 7.20a. The results using simulations are shown for comparison in figure 7.20b and a clear difference is observed in both magnitude and distribution with momentum. This makes using a data-driven model for $\pi^{\pm} \Rightarrow e^{\pm}$ essential, by running the selection on the simulations without taking this into account the results differ by approximately a factor of two.

This method was developed in parallel with, and is the same as, that used for the search for LNV decay $K^{+} \to \pi^{+}e^{+}e^{-}$ [34] and cross checks were performed (using the same PID requirements) with results in good agreement.

7.5.4.4 $e^{\pm} \Rightarrow \pi^{\pm}$ Misidentification Studies

To study $e^{\pm} \Rightarrow \pi^{\pm}$ misidentification $e^{\pm}$ tracks from a control sample of $K^{+} \to \pi^{+}[e^{+}e^{-}\gamma]_{\pi^0}$ decays are used, as described above. The selection is detailed in appendix G.2, and, as for the $\pi^{\pm} \Rightarrow e^{\pm}$ misidentification study, it does not use the LKr but includes a MUV3 veto. This study explicitly measures the $e^{\pm} \Rightarrow \pi^{\pm}$ misidentification probability (of $E/p < 0.9$ with the closest geometrically associated cluster time being within 5 ns of the track time) given that that there were exactly 0 in-time MUV3 associations: $P(e^{\pm} \Rightarrow \pi^{\pm}|0\,\text{MUV3}) =$



(a) (b)

Figure 7.21: Probability of $e^{\pm} \Rightarrow \pi^{\pm}$ misidentification as a function of track charge times momentum for the control sample of $e^{\pm}$ tracks from a $K^{+} \to \pi^{+}[e^{+}e^{-}\gamma]_{\pi^0}$ sample selected in: (a) data collected with the minimum-bias three track (MT) trigger and (b) data and $K^{+} \to \pi^{+}[e^{+}e^{-}\gamma]_{\pi^0}$ MC simulations. Note that the maximum of the y axis scale has been reduced in (b) for finer comparison and the blue points and model in (b) are the same as those in (a). The models describing the distributions, from fits to the measurements, are shown by the blue and red lines overlaid on top of the measurement points.

$P(E/p < 0.95|0\,\mathrm{MUV3})$ as a function of the charge times momentum of the reconstructed track, for results see figure 7.21a. A comparison between results of studying data and simulations is shown in figure 7.21b

The control sample of $K^{+} \to \pi^{+}[e^{+}e^{-}\gamma]_{\pi^0}$ decays is used, as opposed to the alternative of a control sample of $K^{+} \to e^{+}\nu_e\pi^0$ events. Particle identification study results can differ between single-track and three-track cases and therefore the usage of the three-track $K^{+} \to \pi^{+}[e^{+}e^{-}\gamma]_{\pi^0}$ decay better represents the environment under which possible misidentifications can occur in the signal selection, reducing bias. Furthermore, the $K^{+} \to e^{+}\nu_e\pi^0$ sample suffers from contamination (with significant momentum dependence) at the level of $\sim 10^{-3}$, while the $K^{+} \to \pi^{+}[e^{+}e^{-}\gamma]_{\pi^0}$ sample has a contamination from non-electrons of $< 10^{-4}$. This ensures that the dominant effect measured is due to genuine $e^{\pm} \Rightarrow \pi^{\pm}$ misidentification.

The methods used here are identical to those used above for $\pi^{\pm} \Rightarrow e^{\pm}$ misidentification which have been verified by comparison to the LNV $K^{+} \to \pi^{-}e^{+}e^{+}$ analysis [34]. By extension this procedure, and the resulting measurement, are equally valid.

7.5.4.5 $e^{\pm} \Rightarrow \mu^{\pm}$ Misidentification

To misidentify an $e^{\pm}$ as a $\mu^{\pm}$ the measured LKr E/p must be < 0.2. Generally this means the $e^{\pm}$ is 'lost', either due to multiple scattering corrupting geometric matching of clusters to tracks or showers starting upstream of the LKr or other problems with the LKr or tracking. Additionally, there must be an accidental hit in the outer tiles of the

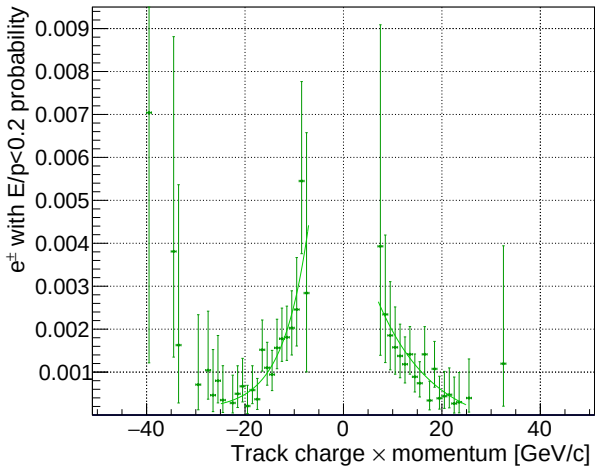


Figure 7.22: Probability of candidate $e^{\pm}$ track, from a the control sample of $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ decays in 2017+2018 data collected with the minimum-bias three track (MT) trigger, being associated with $E/p < 0.2$ as a function of track charge times momentum.

MUV3 which is within 5 ns of the track time.

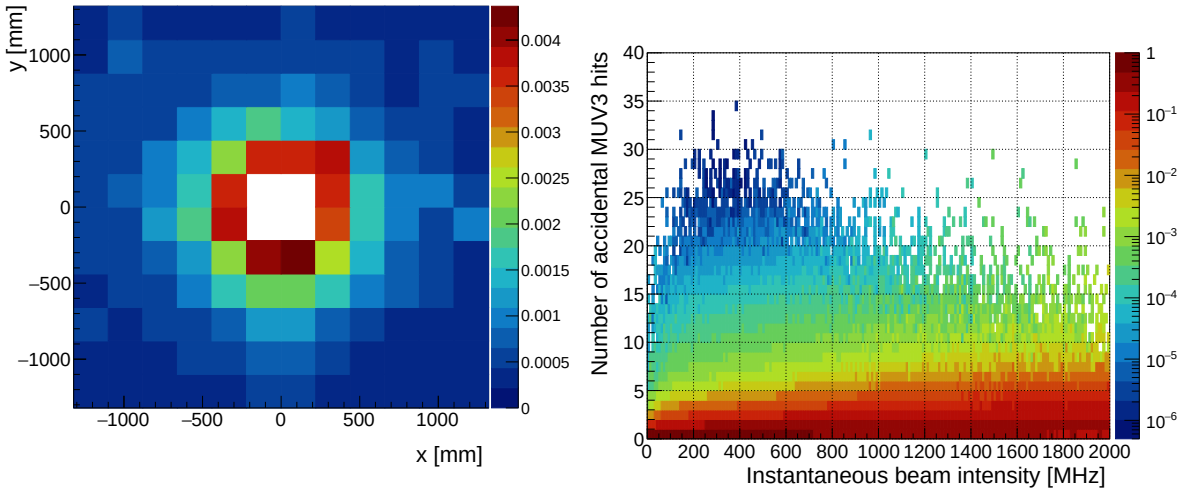
7.5.4.5.1 Low E/p

The probability of the first component, $E/p < 0.2$, is shown as a function of momentum in figure 7.22, with the probability being $\mathcal{O}(10^{-3})$ overall. This measurement is severely limited by the number of selected control sample events with $e^{\pm}$ candidates with $E/p < 0.2$. There is also a possibility of bias from contamination which could exist at the $\mathcal{O}(10^{-3})$ level as shown in section 7.5.4.2, however simulation studies show similar results and momentum dependence.

7.5.4.5.2 Accidental MUV3 Association

While the first component arises from potentially complex processes the latter, accidental MUV3 matching, effect is possible to study in detail since it is a function of well-defined quantities: a) the extrapolated position (x, y) of the track at the MUV3 plane, b) the track momentum (which determines the search radius around the projected MUV3 position (x, y) in which MUV3 hits may be associated to the track) and c) the time window for matching between a track and MUV3 candidate. From the selection criteria a potential accidental match of a track to a MUV3 hit occurs when: a) the extrapolated track (x, y) at the MUV3 plane is within the MUV3 geometric acceptance, b) there is a MUV3 hit within the search radius of the extrapolated MUV3 impact point which is c) in an outer MUV3 tile and d) within 5 ns of the track time.

To study the probability of accidental associations information is used from standard templates of accidental MUV3 signals recorded in data in a 50 ns time window, in sidebands with $45 < |t_{MUV3} - t_{trig}| < 75$ ns, as a function of instantaneous beam intensity and corresponding channel occupancy maps. The probability that a given MUV3 accidental signal is in a certain outer MUV3 tile, is shown in figure 7.23a. This probability is divided by 5 with respect to the template because the time window in the signal PID



(a) (b)

Figure 7.23: (a) The probability of an accidental signal in each outer MUV3 tile in a 10 ns time window. (b) The probability of having n accidental MUV3 signals for a given instantaneous beam intensity, I. Each column of bins with the same I value is normalised to 1, showing the most likely scenario is less than 5 accidental hits, and in general the probability of having a larger number of hits increases with intensity.

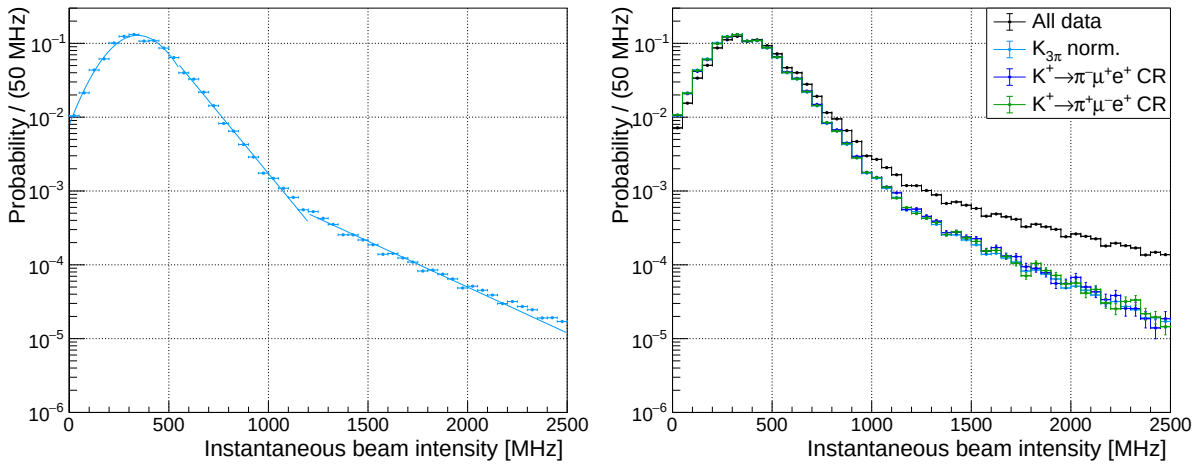
selection is 10 ns wide, 5 times smaller than the one in which accidental hits are considered in the template. The probability for n accidental hits for a given instantaneous intensity I is displayed in figure 7.23b.

The instantaneous beam intensity for selected $K^+ \to \pi^+\pi^+\pi^-$ normalisation events has been measured and a model constructed to describe it, as shown in figure 7.24a. The distribution measured for $K^+ \to \pi^+\pi^+\pi^-$ normalisation events is consistent with that for selected $K^+ \to \pi\mu e$ events outside the blinded regions, as shown by figure 7.24b, and therefore the same model can be used.

The probability of n accidental MUV3 signals, given the measured beam intensity profile for the selected sample of data events, is shown by figure 7.25.

Using these measurements the probability of having n accidental hits is found by multiplying $P(n \text{ accidental hits} \,|\, I)P(I)$ (effectively integrating over each horizontal n bin in figure 7.25) and this is displayed in figure 7.26a. The probability of exactly one accidental signal in the MUV3 in a given outer tile, accounting for the intensity dependence using the above and in the 10 ns time window around the track time, is shown in figure 7.26b.

The probability of having exactly one accidental MUV3 signal in an outer tile inside the search radius (momentum dependent) for a given extrapolated (x, y) position at the MUV3 is shown for 4 example cases in figure 7.27 for tracks of momenta 10, 20, 30 and 50 GeV/c. To provide a more granular description each bin in these plots corresponds to a virtual MUV3 'sub-tile', where a 10×10 array of 'sub-tiles' is contained in each



(a) (b)

Figure 7.24: (a) The probability instantaneous beam intensity distribution for normalisation $K^+ \to \pi^+\pi^+\pi^-$ events, selected in 2017+2018 data collected with the MT trigger, with model from fit to the measured distribution. (b) The probability instantaneous beam intensity distribution for events passing the normalisation and signal selections (in the control regions, CR) as well as for all events in the 2017+2018 data-set.

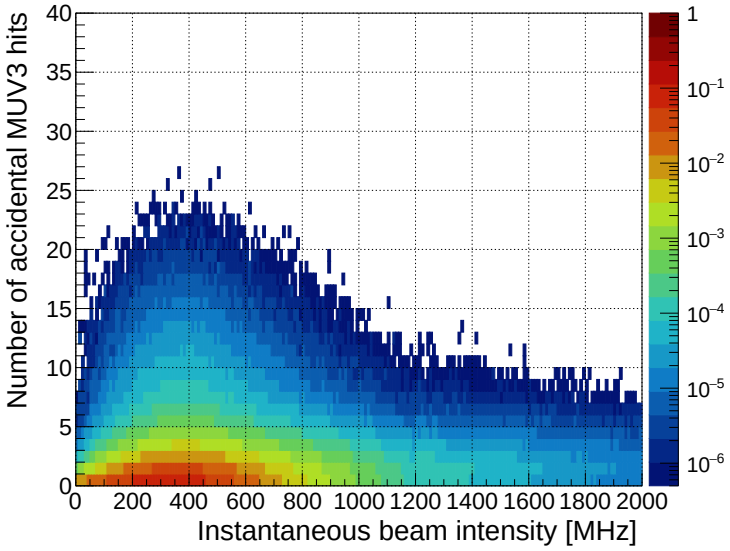
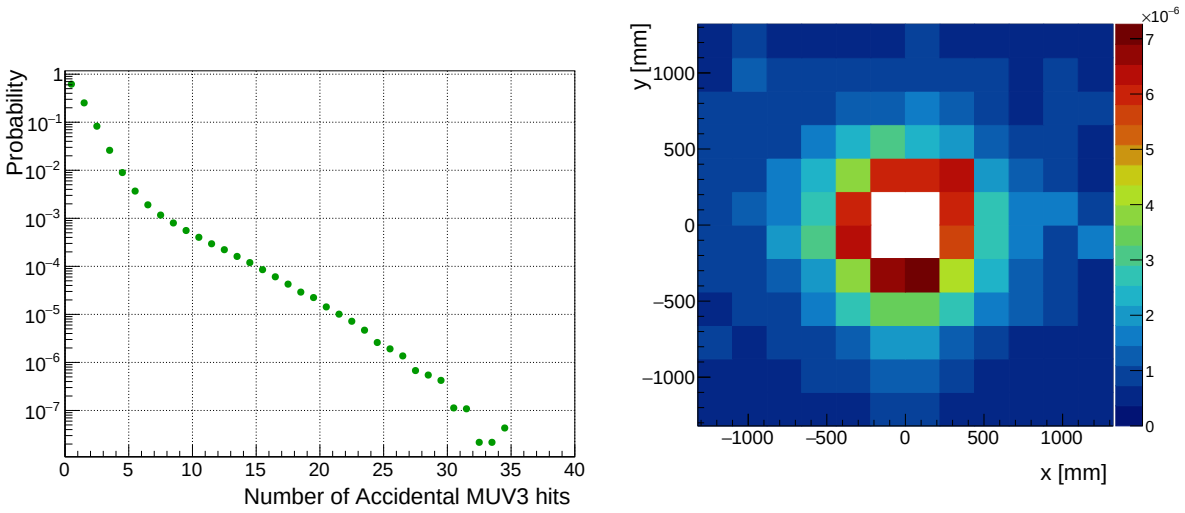


Figure 7.25: The probability of having n accidental MUV3 signals and a given instantaneous beam intensity, I. Each column of bins with the same I value is normalised to the probability of observing that intensity in a given data event in the selected sample, determined by model shown in figure 7.24a.



(a) (b)

Figure 7.26: (a) The probability of n accidental MUV3 signals, accounting for the intensity profile observed in data. (b) Probability of having exactly one in-time accidental signal in each outer MUV3 tile.

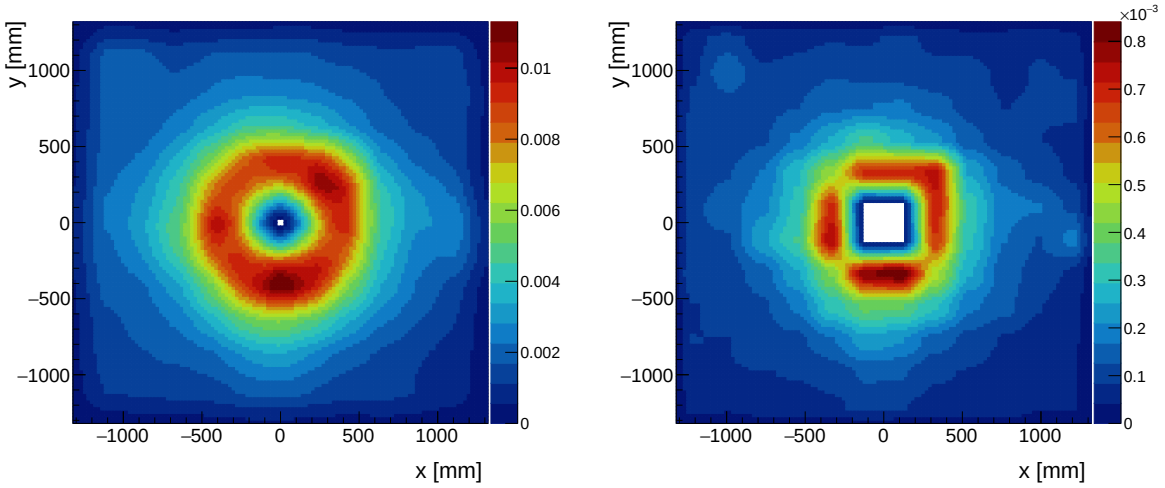
physical MUV3 tile. A full model, such as the four examples shown, is constructed for track momenta in the range 10–55 GeV/c with 1 GeV/c precision.

7.5.4.6 $\pi^{\pm} \Rightarrow \mu^{\pm}$ Misidentification Probability

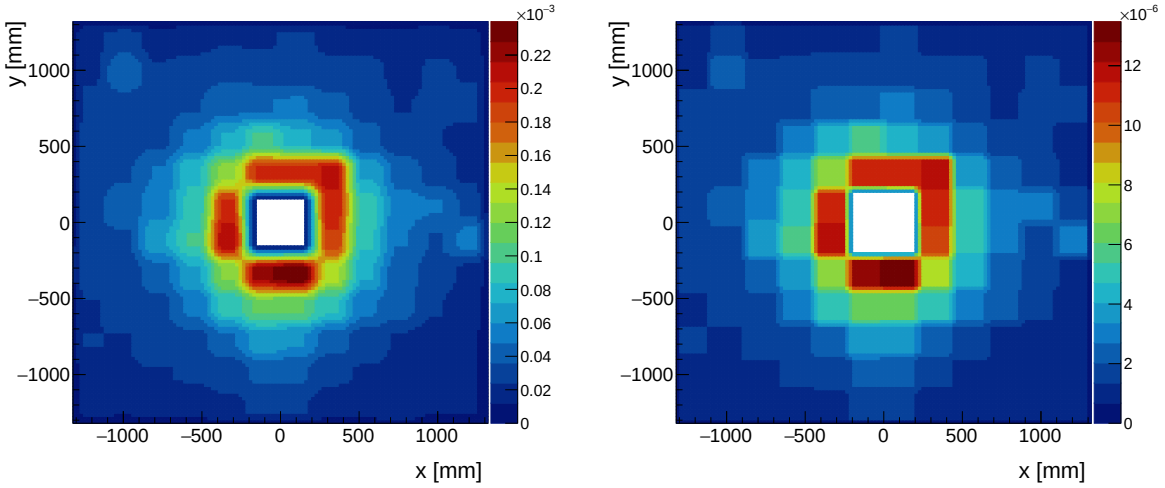
The experimental signature of $\pi^{\pm} \to \mu^{\pm}\nu_{\mu}$ decays in flight and $\pi^{\pm} \Rightarrow \mu^{\pm}$ misidentification is effectively identical when the neutrino carries away a small fraction of the parent $\pi^{\pm}$ momentum. This is the case preferentially selected by the analysis requirement that the total three-charged-track momentum is within 2.5 GeV of the beam K^{+} momentum. Simulations show that the $\pi^{\pm} \to \mu^{\pm}\nu_{\mu}$ decay in flight probability is $\sim 7.5\%$ (see appendix H) and is larger than the $\pi^{\pm} \Rightarrow \mu^{\pm}$ misidentification probability, meaning any measurement of the probability of a $\pi^{\pm}$ becoming or 'looking like' a $\mu^{\pm}$ in data is dominated by decays in flight. Because of these facts the $\pi^{\pm} \Rightarrow \mu^{\pm}$ misidentification process is studied using simulations.

To perform the study the sample used is $\pi^{\pm}$ from simulated $K^{+} \to \pi^{+}\pi^{+}\pi^{-}$ decay events passing the same PID control sample selection as for the $\pi^{\pm} \Rightarrow e^{\pm}$ misidentification studies (but without applying the MUV3 veto), and with no π^{+} decays in flight. The selection is followed by a check to see if the $\pi^{\pm}$ tracks pass the $\mu^{\pm}$ PID criteria. The misidentification probability is calculated as a ratio where all selected events form the denominator and those selected as $\mu^{\pm}$ make up the numerator (binned in 1 GeV/c intervals in the track charge times momentum, qp). The derived results for $\pi^{\pm} \Rightarrow \mu^{\pm}$ misidentification probability are shown in figure 7.28.

Two mechanisms are responsible for $\pi^{\pm} \Rightarrow \mu^{\pm}$ misidentification: pileup and hadronic



(a) (b)



(c) (d)

Figure 7.27: Probability of having exactly one in-time accidental signal associated with a track with momenta 10, 20, 30 and 50 GeV/c, (a), (b), (c), (d) respectively, extrapolated to a given (x, y) location in MUV3.

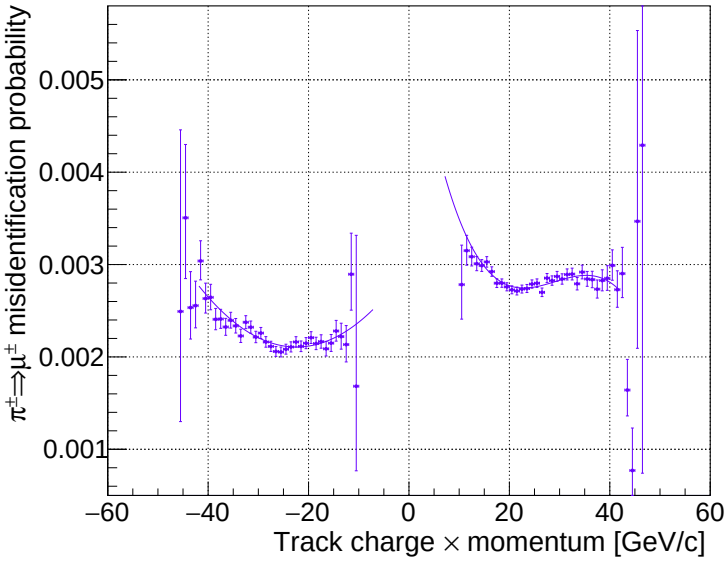


Figure 7.28: The $\pi^{\pm} \Rightarrow \mu^{\pm}$ misidentification probability as a function of the track charge times momentum, qp derived from a study of simulated $K^{+} \rightarrow \pi^{+}\pi^{+}\pi^{-}$ decays. The charge asymmetry is a result of the left/right asymmetry in MUV3 hit occupation.

showers, both inducing in-time activity in the MUV3. Each mechanism is responsible for roughly half of the overall misID probability. The probability of accidental association with pileup signals in the MUV3 is momentum dependent due to the search radius which is inversely proportional to the track momentum (see also section 7.5.4.5), and this is responsible for a decrease in the $\pi^{\pm} \Rightarrow \mu^{\pm}$ misID probability as momentum magnitude increases. In contrast, the probability of in-time MUV3 activity from hadronic showers, included in the simulations, increases somewhat with increasing momentum. The competing dependencies create different results for positive and negative tracks, with left-right asymmetry in the MUV3 channel occupation from pileup being a key reason for charge-asymmetric results observed in figure 7.28.

7.5.5 Forced Misidentification Procedure for Background Predictions

Background studies for a number of possible mechanisms are limited or impossible due to the simulated sample size of this type of event. One critical case is the primary background mechanism for $K^{+} \rightarrow \pi^{+}\pi^{+}\pi^{-}$ decays which includes a $\pi^{+} \Rightarrow e^{+}$ misID and a $\pi^{\pm} \rightarrow \mu^{\pm}$ DIF, see section 7.5.7.1 for details. In this case approximately 5×10^{7} simulated $K^{+} \rightarrow \pi^{+}\pi^{+}\pi^{-}$ events are available, this means the scale factor (see equation 7.5.2) $f_{scale} \approx \frac{10^{12} \times 0.05}{5 \times 10^{7}} \approx 10^{3}$. Therefore each simulated event corresponds to an expected $\mathcal{O}(10^{3})$ events in the data sample and this immediately precludes any detailed background analysis. Moreover as shown by figures 7.20 and 7.21 the measured misidentification probability for data and simulation studies disagree, by approximately a factor of 2 for the $\pi^{+} \Rightarrow e^{+}$ misID relevant to this case. Both of these issues with using simulations, limited statistical power and poor reproduction of the misidentification mechanism, can

be solved using the forced misidentification procedure.

To perform forced misidentification, for the $X^{\pm} \Rightarrow Y^{\pm}$ misID process, the PID selection decision is not considered for the candidate $X^{\pm}$ track and the identity $Y^{\pm}$ is assigned and an event weight, w_{PID}, is applied based on the models reported in section 7.5.4. The specific application of the forced misidentification procedure for each case is discussed in the following sections (7.5.5.1–7.5.5.3). Considering the misidentification probabilities, assigning the identity allows an increase in statistical power by a factor of $\mathcal{O}(10^3)$, depending on the specific case (the order of magnitude of misID probabilities are given in section 7.5.4.1). Data-driven models are used (where possible) and this removes the issue of differences between simulations and data.

7.5.5.1 Forced $\pi^{\pm} \Rightarrow e^{\pm}$ or $e^{+} \Rightarrow \pi^{\pm}$ (LKr only)

As described, in section 7.5.4.3 and 7.5.4.4 respectively, the misidentification probabilities for $\pi^{\pm} \Rightarrow e^{\pm}$ or $e^{+} \Rightarrow \pi^{\pm}$ are measured from control sample selections with the MUV3 veto already applied meaning the probability is the probability of obtaining LKr E/p in a certain range given no in-time MUV3 association exists, $P(\text{LKr}\, E/p\, \text{misID}\,|\,!\text{MUV3})$. To apply forced misidentification in these two cases it is first required that there is no in-time MUV3 association (in inner or outer tiles) for the track. Then, the misidentified identity is assigned (a $e^{\pm}$ ID for $\pi^{\pm} \Rightarrow e^{\pm}$ forced misID and $\pi^{\pm}$ ID for $e^{+} \Rightarrow \pi^{\pm}$ forced misID), and an event weight set for the given charge times momentum, qp, of the track according to models built in the studies above.

This procedure is applied in the following cases for forced $\pi^{\pm} \Rightarrow e^{\pm}$ misidentification:

- $K^{+} \to \pi^{+}\pi^{+}\pi^{-}$ in both π^{-} and μ^{-} channels.
- $K^{+} \to \pi^{+}[e^{+}e^{-}\gamma]_{\pi^0}$ in π^{-} channel (as part of triple forced misID).
- $K^{+} \to \pi^{+}\pi^{-}\mu^{+}\nu_{\mu}$ in both π^{-} and μ^{-} channels.
- $K^{+} \to \pi^{+}\mu^{+}\mu^{-}$ in both π^{-} and μ^{-} channels.

Forced $e^{+} \Rightarrow \pi^{\pm}$ misidentification is applied for the following cases:

- $K^{+} \to \pi^{+}[e^{+}e^{-}\gamma]_{\pi^0}$ in π^{-} channel (as part of triple forced misID).
- $K^{+} \to \mu^{+}\nu_{\mu}[e^{+}e^{-}\gamma]_{\pi^0}$ in both π^{-} and μ^{-} channels (as part of double forced misID).
- $K^{+} \to e^{+}\nu_{e}[e^{+}e^{-}\gamma]_{\pi^0}$ in both π^{-} and μ^{-} channels (as part of double forced misID).
- $K^{+} \to \pi^{+}e^{+}e^{-}$ in π^{-} channel.
- $K^{+} \to \mu^{+}\nu_{\mu}e^{+}e^{-}$ in π^{-} channel.

7.5.5.2 Forced $\mu^{\pm} \Rightarrow e^{\pm}$ Misidentification (Accidental MUV3 Association)

Misidentification mechanism $\mu^{\pm} \Rightarrow e^{\pm}$ requires $E/p < 0.2$ and an accidental in-time MUV3 association in an outer tile. From section 7.5.4.5.2 the accidental MUV3 association probability is known for a given projected (x_{MUV3}, y_{MUV3}) position at MUV3, matching time window, Δt_w (5 ns) and track momentum p_{trk}. To apply forced misidentification it is first required that the candidate track has $E/p < 0.2$,[5] then the $\mu^{\pm}$ identity is assigned and an event weight applied (given the (x_{MUV3}, y_{MUV3}), Δt_w and p_{trk}).

Forced $\mu^{\pm} \Rightarrow e^{\pm}$ misidentification is applied for:

- $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ in π^- channel (as part of triple forced misID) and, importantly, in the μ^- channel.
- $K^+ \to \mu^+\nu_\mu[e^+e^-\gamma]_{\pi^0}$ in both π^- and μ^- channels (as part of double forced misID).
- $K^+ \to e^+\nu_e[e^+e^-\gamma]_{\pi^0}$ in both π^- and μ^- channels (as part of double forced misID).
- $K^+ \to \pi^+e^+e^-$ in μ^- channel.
- $K^+ \to \mu^+\nu_\mu e^+e^-$ in μ^- channel (as part of double forced misiD).

7.5.5.3 Forced $\pi^{\pm} \Rightarrow \mu^{\pm}$ Misidentification (LKr and MUV3)

For the case of $\pi^{\pm} \Rightarrow \mu^{\pm}$ misidentification the only fully simulation-driven measurement of the misID probability is performed (the rest being data-driven). When applying forced $\pi^{\pm} \Rightarrow \mu^{\pm}$ misidentification the 'truth' information available in the simulation is first checked to ensure the candidate track ($\pi^{\pm}$) did not undergo a decay in flight process. A $\mu^{\pm}$ identity is assigned and the misidentification probability (of having LKr $E/p < 0.2$ and exactly 1 in-time MUV3 association, as a function of the charge times momentum of the track from models established in section 7.5.4.6) set as an event weight.

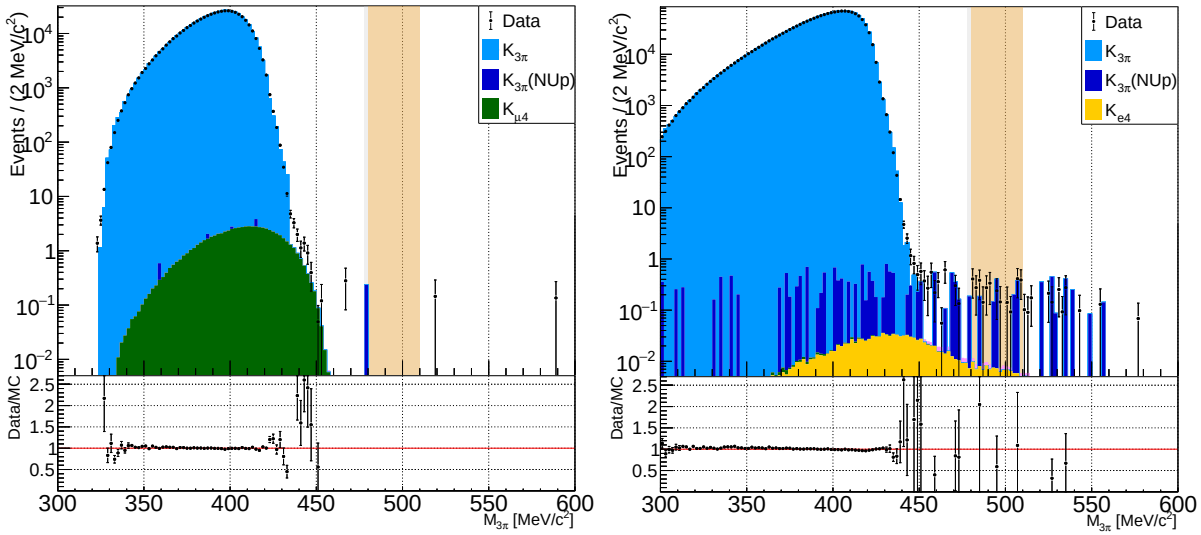
Forced $\pi^{\pm} \Rightarrow \mu^{\pm}$ misidentification is used to study backgrounds:

- $K^+ \to \pi^+\pi^+\pi^-$ in both π^- and μ^- channels.
- $K^+ \to \pi^+\pi^-e^+\nu_e$ in both π^- and μ^- channels.

7.5.6 Explicitly Data-Driven $K^+ \to \pi^+\pi^+\pi^-$ Background Estimate

For the $K^+ \to \pi^+\pi^+\pi^-$ background the large branching ratio means simulation of a sample representative of the expected number of such decays in the data sample is unfeasible. As shown in section 7.5.7.1 in most cases the use of forced misidentification

[5] In this case the results from simulations and data are in agreement within uncertainties so the simulation response is used. However, the uncertainties are large for both data and simulation studies due to very limited statistics available.



(a) (b)

Figure 7.29: Selected alternative PID control samples (a) $\pi^-\mu^+\pi^+$ and (b) $\pi^+\pi^+\mu^-$ collected with the MT trigger with comparison to simulations (uncertainties are statistical). With additional scaling, w, see text, (a) and (b) are used to provide a data-driven background estimate for the $K^+ \to \pi^+\pi^+\pi^-$ background in the π^- and μ^- signal channels respectively. The high invariant mass tail is dominated by $K^+ \to \pi^+\pi^+\pi^-$ decays upstream of the standard FV, shown by the comparison of data to the upstream simulations sample shown in dark blue.

and special simulation studies allow sufficient sensitivity to be reached. However, as both a cross-check and an important measurement where statistics are lacking (most notably in the π^- channel above 478 MeV), an explicitly data-driven background estimate for the (combination of the) two main $K^+ \to \pi^+\pi^+\pi^-$ background mechanisms is performed. To do this an 'alternative PID' selection is run which is identical to the signal selection except requiring positive identification of $\pi^+\mu^\pm\pi^\mp$, followed by forced $\pi^+ \Rightarrow e^+$ misidentification. This must be done only using events passing the MT trigger to remove bias introduced by additional requirements in the other trigger streams, and a weight is assigned to each event including this effect and the $\pi^+ \Rightarrow e^+$ misidentification probability: $w = P(\pi^+ \Rightarrow e^+ \,\text{misID}\,|\,p)\frac{D_{MT}}{D_{eff}}$. The comparison between data and simulations for this 'alternative PID' $\pi^+\mu^\pm\pi^\mp$ selection is shown in figures 7.29a and 7.29b, for the π^- and μ^- channels respectively, and it is clear that the dominant component in the high $M_{\pi\mu e}(> 478\,\text{MeV}/c^2)$ region arises from $K^+ \to \pi^+\pi^+\pi^-$ upstream decays.

It is noted that a similar approach was attempted to constrain the combination of sub-dominant background mechanisms including $\pi^\pm \Rightarrow \mu^\pm$ misidentification, however here the alternative PID selection (selecting $\pi^+e^+\pi^-$) is contaminated by $K^+ \to \pi^+e^+e^-$ events in the region of interest at high $M_{\pi\mu e}$.

7.5.7 Procedures Used To Study Each Background

The treatment of each background is summarised in table 7.11. Key cases are described in the following sections, 7.5.7.1–7.5.7.4.

In simulations pileup, critically in the MUV3, is emulated. The MUV3 inefficiency is measured for data to be 0.15%, with weak geometric and momentum dependence, while inefficiency is negligible in simulations. To account for this difference 0.15% of the reconstructed candidate signals in the MUV3 are rejected randomly when analysing simulated events.

Table 7.11: Summary of techniques for evaluating background predictions for each background mechanism (see table 7.8) in the search for $K^+ \to \pi\mu e$ decays.

Background Abbr.	Case	π^- Channel (FS = $\pi^-\mu^+e^+$)	μ^- Channel (FS = $\pi^+\mu^-e^+$)
$K_{3\pi}$	A	$K_{3\pi}$ (standard ($K_{3\pi}$ event class 1) + near upstream) MC with forced $\pi^+ \Rightarrow e^+$ misID & Data-driven for $M_{\pi\mu e} > 478\text{MeV}/c^2$.	$K_{3\pi}$ (standard ($K_{3\pi}$ event class 3) + near upstream) MC with forced $\pi^+ \Rightarrow e^+$ misID & Data-driven for $M_{\pi\mu e} > 478\text{MeV}/c^2$.
	B	$K_{3\pi}$ standard MC ($K_{3\pi}$ event class 0) [near upstream and Data-driven for combination of A+B]	$K_{3\pi}$ standard MC ($K_{3\pi}$ event class 0) [near upstream for combination of A+B]
	C	$K_{3\pi}$ 5050 MC ($K_{3\pi}$ event class 9)	$K_{3\pi}$ 5050 MC ($K_{3\pi}$ event class 7)
	D	$K_{3\pi}$ pie1 MC ($K_{3\pi}$ event class 2) with forced $\pi^+ \Rightarrow \mu^+$ misID	$K_{3\pi}$ pie1 MC ($K_{3\pi}$ event class 2) with forced $\pi^- \Rightarrow \mu^-$ misID
$K_{2\pi D}$	A	negligible (removed by DalitzRejection cut)	- N/A -
	B	negligible (removed by DalitzRejection cut)	- N/A -
	C	- N/A -	forced $e^+ \Rightarrow \mu^+$ misID (accidental MUV3 asso.)
	D	forced triple misID ($\pi^+e^+e^- \Rightarrow e^+\mu^+\pi^-$)	forced triple misID ($\pi^+e^+e^- \Rightarrow e^+\pi^+\mu^-$)
	E	double forced misID ($e^+e^- \Rightarrow \mu^+\pi^-$)	double forced misID ($e^+e^- \Rightarrow \pi^+\mu^-$)
$K_{\mu 3D}$	A	negligible (removed by DalitzRejection cut)	- N/A -
	B	- N/A -	forced $e^- \Rightarrow \mu^-$ misID [B+C]
	C	forced $e^- \Rightarrow \pi^-$ misID	[B+C combined]
K_{e3D}	A	double forced misID ($e^+e^- \Rightarrow \mu^+\pi^-$)	double forced misID ($e^+e^- \Rightarrow \pi^+\mu^-$)
$K_{\mu 4}$	A	forced $\pi^+ \Rightarrow e^+$ misID	- N/A -
	B	- N/A -	standard analysis [for A+C+D]
	C	standard analysis [also covers case A&D], found to be negligible	standard analysis [also covers case A,B&D], found to be negligible
	D	standard analysis [also covers case A&C], found to be negligible	standard analysis [also covers case A,B&C], found to be negligible
K_{e4}	A	standard analysis [for A+C+D]	standard analysis [for A+C+D]
	B	forced $\pi^+ \Rightarrow \mu^+$ misID	forced $\pi^- \Rightarrow \mu^-$ misID
	C	standard analysis [combined with A+D]	- N/A -
	D	standard analysis [combined with A+C]	- N/A -
$K_{\pi\mu\mu}$	A	forced $\pi^+ \to e^+$ misID	forced $\pi^+ \Rightarrow e^+$ misID
	B	(standard analysis, considered to be negligible)	(standard analysis, considered to be negligible)
	C	- N/A -	(forced $\mu^+ \Rightarrow e^+$ misID, found to be negligible)
$K_{\pi ee}$	A	forced $e^- \Rightarrow \pi^-$ misID	- N/A -
	B	forced $e^- \Rightarrow \pi^-$ misID (+ forced $\pi^+ \Rightarrow \mu^+$ misID)	- N/A -
	C	(forced $e^+ \Rightarrow \mu^+$ misID) considered negligible	($e^- \Rightarrow \mu^-$ misID) considered negligible
	D	triple forced misID ($\pi^+e^+e^- \Rightarrow e^+\mu^+\pi^-$) - found to be negligible	triple forced misID ($\pi^+e^+e^- \Rightarrow e^+\pi^+\mu^-$) - found to be negligible
	E	considered to be negligible	considered to be negligible
$K_{\mu\nu\mu\mu}$	A	standard analysis	standard analysis
$K_{e\nu\mu\mu}$	A	standard analysis	standard analysis
$K_{\mu\nu ee}$	A	forced $e^- \Rightarrow \pi^-$ misID	- N/A -
	B	- N/A -	forced $e^- \Rightarrow \mu^-$ misID
$K_{e\nu ee}$	A	double forced misID ($e^+e^- \Rightarrow \mu^+\pi^-$)	double forced misID ($e^+e^- \Rightarrow \pi^+\mu^-$)

7.5.7.1 Treatment of $K^+ \to \pi^+\pi^+\pi^-$ Background

There are four background mechanisms from $K^+ \to \pi^+\pi^+\pi^-$ decays as shown schematically in figure 7.30. Decays upstream of the FV, of any of the four types, can contribute to the background. Each mechanism is studied separately and the four components are combined to give a total background estimate. Since $N_K \sim 10^{12}$ and $\mathcal{B}(K^+ \to \pi^+\pi^+\pi^-) = (5.583 \pm 0.024)\%$ [13] the data sample is expected to include approximately 5×10^{10} $K^+ \to \pi^+\pi^+\pi^-$ decays. Simulation of $\mathcal{O}(50)$ billion SM $K^+ \to \pi^+\pi^+\pi^-$ decay events is unfeasible and so additional techniques must be applied to understand these background

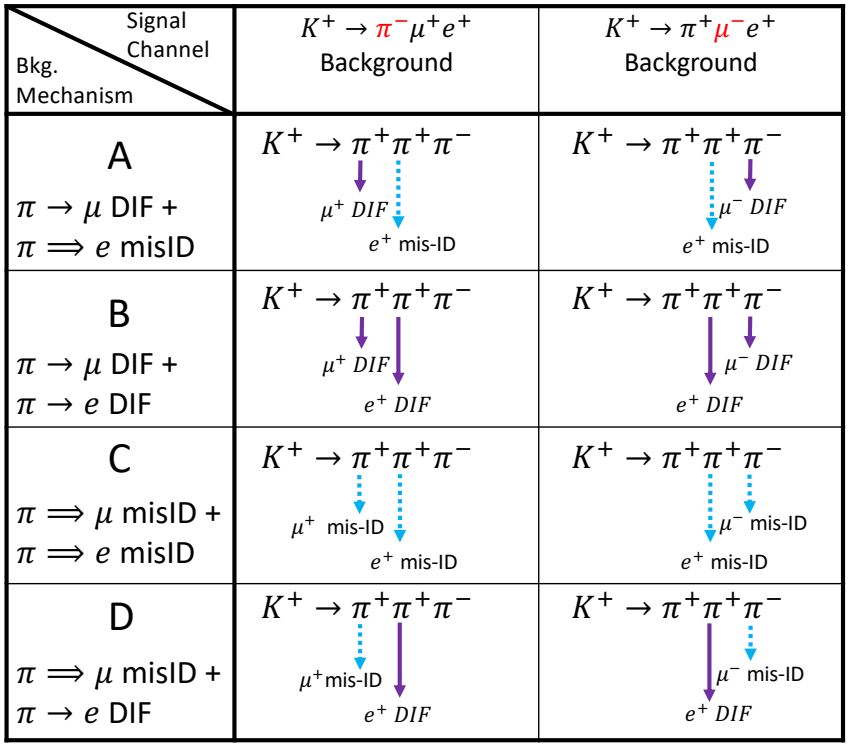


Figure 7.30: Schematic illustration of the background mechanisms through which $K^+ \to \pi^+\pi^+\pi^-$ decays can obtain a particle identity signature like a signal $K^+ \to \pi\mu e$ decay.

mechanisms. For the primary background mechanisms (A and B) which include $\pi^\pm \Rightarrow e^\pm$ misidentification, use of the forced misID procedure, described in section 7.5.5, increases the equivalent sample size by a factor $1/P(\pi^+ \to e^\pm \text{misID}) \sim 1/(4 \times 10^{-4}) \sim 250$. Using this technique the available simulation sample, of approximately 50 million SM $K^+ \to \pi^+\pi^+\pi^-$ decay events, becomes sufficient (scale factor $f_{scale} \sim \mathcal{O}(1)$). This forced misID procedure is also mandatory because of the differences between data and simulation misidentification performance (see section 7.5.4.3), and therefore to correctly reproduce the data the correct misidentification probabilities must be injected into the simulations.

There are 18 possible event classes for $K^+ \to \pi^+\pi^+\pi^-$ decays with up to three decays in flight possible, see table 7.12, figure 7.31 and appendix H. Each of the four background mechanisms can be matched to these event types, as shown in table 7.12.

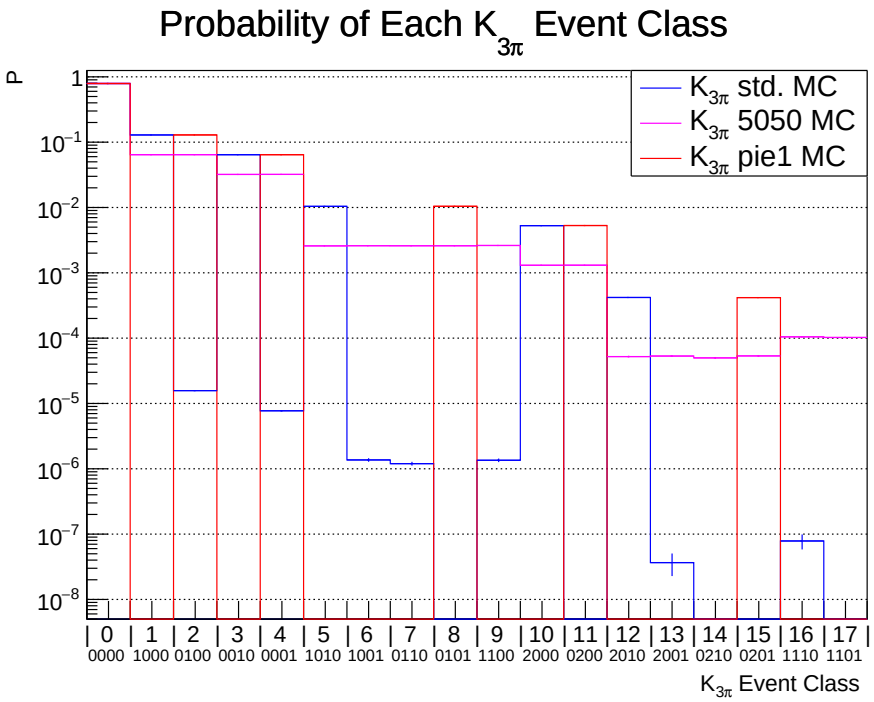


Figure 7.31: Probability of each $K^+ \to \pi^+\pi^+\pi^-$ event class (see table 7.12) for standard simulations (blue, with SM branching ratios), with $\mathcal{B}(\pi^\pm \to e^\pm\nu_e) = \mathcal{B}(\pi^\pm \to \mu^\pm\nu_\mu) = 0.5$ (magenta) and $\mathcal{B}(\pi^\pm \to e^\pm\nu_e) = 1$ (red).

Table 7.12: $K^+ \to \pi^+\pi^+\pi^-$ Event Class.

Event Class	$n_\mu^+ n_e^+ n_\mu^- n_e^-$	Final Particle Content	$K^+ \to \pi^\pm\mu^\mp e$ Background Mechanism		Probability of Event Class $P(n_\mu^+ n_e^+ n_\mu^- n_e^-)$		
			π^- Channel	μ^- Channel	K3pi Std. MC	K3pi5050 MC	K3pi_pie1 MC
0	0000	$\pi^+\pi^+\pi^-$	B) $\pi^+ \Rightarrow \mu^+ + \pi^+ \Rightarrow e^+$	B) $\pi^- \Rightarrow \mu^- + \pi^+ \Rightarrow e^+$	0.791	0.791	0.791
1	1000	$\mu^+\pi^+\pi^-$	A) $\pi^+ \to \mu^+ + \pi^+ \Rightarrow e^+$	-	0.129	0.643	0
2	0100	$\pi^+e^+\pi^-$	D) $\pi^+ \Rightarrow \mu^+ + \pi^+ \to e^+$	D) $\pi^- \Rightarrow \mu^- + \pi^+ \to e^+$	1.57×10^{-5}	0.643	0.129
3	0010	$\pi^+\pi^+\mu^-$	-	A) $\pi^- \to \mu^- + \pi^+ \Rightarrow e^+$	0.0643	0.0321	0
4	0001	$\pi^+\pi^+e^-$	-	-	7.73×10^-6	0.0321	0.0643
5	1010	$\mu^+\pi^+\mu^-$	-	-	0.0104	2.58×10^{-3}	0
6	1001	$\mu^+\pi^+e^-$	-	-	1.37×10^{-6}	2.60×10^{-3}	0
7	0110	$\pi^+e^+\mu^-$	-	C) $\pi^- \to \mu^- + \pi^+ \to e^+$	1.20×10^{-6}	2.60×10^{-3}	0
8	0101	$\pi^+e^+e^+$	-	-	$< 1.2 \times 10^{-8}$	2.60×10^{-3}	0.0104
9	1100	$\mu^+e^+\pi^-$	C) $\pi^+ \to \mu^+ + \pi^+ \to e^+$	-	1.35×10^{-6}	2.62×10^{-3}	0
10	2000	$\mu^+\mu^+\pi^-$	-	-	5.27×10^{-3}	1.31×10^{-3}	0
11	0200	$e^+e^+\pi^-$	-	-	$< 1.2 \times 10^{-8}$	$1.31 \times 016{-}3$	5.30×10^{-3}
12	2010	$\mu^+\mu^+\mu^-$	-	-	4.19×10^{-4}	5.22×10^{-5}	0
13	2001	$\mu^+\mu^+e^-$	-	-	3.66×10^{-8}	5.32×10^{-5}	0
14	0210	$e^+e^+\mu^-$	-	-	$< 1.2 \times 10^{-8}$	4.98×10^{-5}	0
15	0201	$e^+e^+e^-$	-	-	$< 1.2 \times 10^{-8}$	5.35×10^{-5}	4.18×10^{-4}
16	1110	$\mu^+e^+\mu^-$	-	-	7.85×10^{-8}	1.05×10^{-4}	0
17	1101	$\mu^+e^+e^-$	-	-	$< 1.2 \times 10^{-8}$	1.02×10^{-4}	0

A set of NA62 simulation features, as described below, are exploited for the $K^+ \to \pi^+\pi^+\pi^-$ background studies.

- **Modified $\pi^\pm$ decay branching ratio**
  For sub-dominant background mechanisms including at least one $\pi^\pm \to e^\pm\nu_e$ decay in flight, the probability of a specific event type can be enhanced in simulations by modifying the $\pi^\pm$ branching ratio or the pion lifetime. In NA62 simulations the

approximation

$$\mathcal{B}(\pi^{\pm} \to \mu^{\pm}\nu_{\mu}) + \mathcal{B}(\pi^{\pm} \to e^{\pm}\nu_{e}) = 1 \tag{7.5.7}$$

is made, which holds with high precision (see equations 7.5.4 and 7.5.5). The value of $\mathcal{B}(\pi^{\pm} \to e^{\pm}\nu_{e})$ can be modified and $\mathcal{B}(\pi^{\pm} \to \mu^{\pm}\nu_{\mu})$ is then adjusted according to equation 7.5.7.

- **Capped $\pi^{\pm}$ lifetime**
  NA62 simulations allow the proper lifetime of a $\pi^{\pm}$ to be capped meaning it is forced to decay within the NA62 simulations fiducial volume (defined as $102.425 < z < 180\,\text{m}$).

- **Upstream Decays**
  Upstream $K^{+} \to \pi^{+}\pi^{+}\pi^{-}$ decays are simulated by modifying the z range in which K^{+} decays are allowed. A 'near upstream' region is defined with K^{+} decays occurring at $96.950 < z < 102.425\,\text{m}$ as opposed to the standard NA62 simulations fiducial volume.

Mechanisms A and B can be studied using standard NA62 simulations of SM $K^{+} \to \pi^{+}\pi^{+}\pi^{-}$ decays, using forced $\pi^{+} \Rightarrow e^{+}$ misidentification. Mechanism A, with one $\pi^{+} \to \mu^{\pm}\nu_{\mu}$ decay in flight ($K_{3\pi}$ event class 1 or 3 for the π^{-} and μ^{-} channels respectively) can be studied with the standard simulations since the probability of a π^{+} decaying before the end of the standard NA62 simulations fiducial volume is about 7.5%.

However, for mechanisms C and D, which include a $\pi^{\pm} \to e^{\pm}\nu_{e}$ decay in flight, the smaller branching ratio of 1.23×10^{-4} (see equations 7.5.5 and 7.5.4) means an insufficient number of such events arise in the standard simulations sample to provide adequate background estimates. By modifying the branching ratios of $\pi^{\pm}$ decays, setting $\mathcal{B}(\pi^{\pm} \to e^{\pm}\nu_{e}) = \mathcal{B}(\pi^{\pm} \to \mu^{\pm}\nu_{\mu}) = 0.5$, the number of events with C or D-like decays in flight is enhanced ($K_{3\pi}$ event class 9 and 7 for the π^{-} and μ^{-} channels respectively, this simulation is referred to as 'K3pi5050'). Two decays in flight happening in the same event is relatively rare, with probability of approximately 0.5%, and so for mechanism C additional enhancement from imposing a capped pion lifetime on two of the three pions is exploited (this simulation is referred to as 'K3piCapped5050'). For mechanism C, with both types of decay in flight, it is optimal to use $\mathcal{B}(\pi^{\pm} \to e^{\pm}\nu_{e}) = \mathcal{B}(\pi^{\pm} \to \mu^{\pm}\nu_{\mu}) = 0.5$, as shown in appendix H, while for mechanism D with only $\pi^{\pm} \to e^{\pm}\nu_{e}$ decays ($K_{3\pi}$ event class 2 for both π^{-} and μ^{-} channels) further enhancement is achieved by setting $\mathcal{B}(\pi^{\pm} \to e^{\pm}\nu_{e}) = 1$ and therefore $\mathcal{B}(\pi^{\pm} \to \mu^{\pm}\nu_{\mu}) = 0$ (this simulation is referred to as 'K3pi_$\pi e1$'). The scaling factor obtained from modifying the $\pi^{\pm}$ branching ratio is given by

$$f_{mod}(n_{\mu}, n_{e}) = \left(\frac{Br(\pi^{\pm} \to \mu^{\pm}\nu_{\mu})_{SM}}{Br(\pi^{\pm} \to \mu^{\pm}\nu_{\mu})_{mod}}\right)^{n_{\mu}} \cdot \left(\frac{Br(\pi^{\pm} \to e^{\pm}\nu_{e})_{SM}}{Br(\pi^{\pm} \to e^{\pm}\nu_{e})_{mod}}\right)^{n_{e}} . \tag{7.5.8}$$

where n_{μ} and n_{e} are the numbers of $\pi^{\pm} \to \mu^{\pm}\nu_{\mu}$ and $\pi^{-} \to e^{\pm}\nu_{e}$ decays in the event respectively and the modified and standard model (measured) branching ratios are labelled with subscripts 'mod' and 'SM' respectively. This factor is used, as weight w_{MC} introduced in section 7.5.1, when scaling the number of MC simulation events for background

estimates. The enhancement factor ($1/f_{mod}$) for K3pi event class 7 or 9 in K3pi5050 simulations is $1/f_{mod} = 2032.520$ (for background mechanism C) and for K3pi event class 2 in K3pi_$\pi e1$ MC is $1/f_{mod} = 4065.041$ (for background mechanism D).

The procedures used and their associated enhancement factors using specific simulation samples are given in table 7.13.

Table 7.13: $K^+ \to \pi^+\pi^+\pi^-$ background mechanisms, probability of occourance, and MC enhancement factor

Bkg. Mech.	Mechanism Description	Prob. of Mechanism	K3pi Event Class (π^-)[μ^-] and Probability	Enhancement Method	Enhancement Factor
A	$\pi^\pm \to \mu^\pm + \pi \Rightarrow e^+$	$\sim 4 \times 10^{-3}$	(1, $P = 0.129$), [3, $P = 0.0643$]	Forced $\pi \Rightarrow e^+$ misID	~ 250
B	$\pi^\pm \Rightarrow \mu^\pm + \pi \Rightarrow e^+$	$\sim 1 \times 10^{-5}$	(0) [0] $P = 0.791$	Forced $\pi \Rightarrow e^+$ ($+\pi^\pm \Rightarrow \mu^\pm$) misID	~ 250 ($\sim 1 \times 10^5$)
C	$\pi^\pm \to \mu^\pm + \pi \to e^+$	$\sim 1 \times 10^{-6}$	(9, $P = 1.35 \times 10^{-6}$) [7, $P = 1.20 \times 10^{-6}$]	K3pi5050 MC or K3piCapped5050 MC	~ 2000 or $\sim 3 \times 10^4$
D	$\pi^\pm \Rightarrow \mu^\pm + \pi \to e^+$	$\sim 3 \times 10^{-9}$	(2) [2], $P = 1.57 \times 10^{-5}$	K3pi_pie1 MC + Forced $\pi^\pm \Rightarrow \mu^\pm$ misID	$\sim 4 \times 10^6$

For the upstream component it is sufficient to constrain only mechanisms A and B since the standard z_{vtx} cuts reduce the background to a low level so mechanisms C and D become negligible. The sum of mechanisms A and B can be constrained using only forced $\pi^\pm \Rightarrow e^\pm$ misID and studied using upstream decay simulations. The K^+ is produced upstream at the target, $z = 0$, and the probability of it remaining undecayed at a given z is given by the exponential decay law

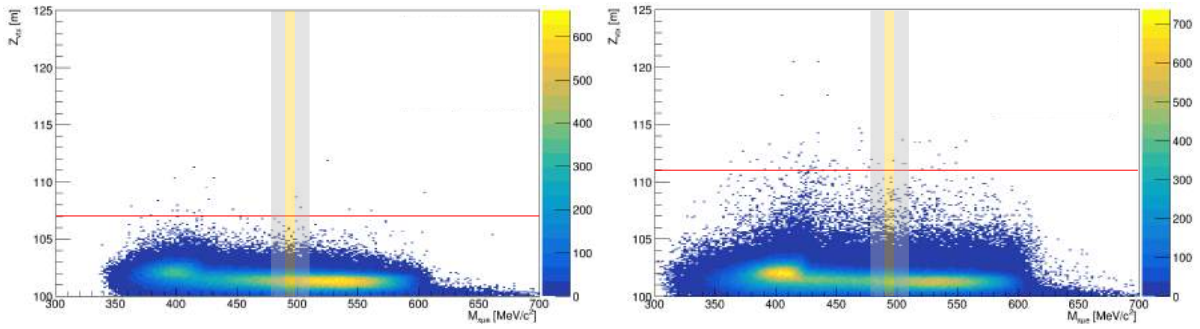
$$P(z) = e^{-\frac{z}{c\tau\sqrt{\gamma^2-1}}}, \tag{7.5.9}$$

where $z = \beta\gamma c\tau$ is the distance travelled by the K^+ in the laboratory reference frame and Lorentz factor $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ where $\beta = \frac{v}{c}$ is the K^+ velocity and [13] $\tau = (1.2380 \pm 0.0020) \times 10^{-8}\,\text{s}$ is the mean K^+ lifetime. Therefore the ratio of K^+ decays occurring in the near upstream region relative to the standard NA62 simulations fiducial volume, R_{NUp}, assuming the beam K^+ has momentum $75\,\text{GeV}/c$, is given by integrals of $P(z)$ over the respective ranges of z (expressed in m) by

$$R_{NUp} = \frac{P(96.950 < z < 102.425)_{NUp}}{P(102.425 < z < 180.000)} = \frac{\int_{96.950}^{102.425} e^{-\frac{z}{c\tau\sqrt{\gamma^2-1}}} dz}{\int_{102.425}^{180.000} e^{-\frac{z}{c\tau\sqrt{\gamma^2-1}}} dz} = \frac{4.58776}{60.4359} = 0.076\,. \tag{7.5.10}$$

Therefore each near upstream decay event is equivalent to $1/R_{NUp} = 13.2$ decays in the standard decay volume and the weight $w_{MC} = R_{NUp}$ is applied to upstream simulation events (see section 7.5.1). This effective enhancement factor, coupled with the factor ~ 250 enhancement from forced $\pi^\pm \Rightarrow e^\pm$ misID, means the available upstream simulation sample of 1×10^8 events is sufficient to study the background.

Very few simulated $K^+ \to \pi^+\pi^+\pi^-$ events pass the signal selection with candidate $M_{\pi\mu e} > 478\,\text{MeV}/c^2$. For this invariant mass regime the data-driven background estimates, described in section 7.5.6, are used for both the π^- and μ^- signal channels. The data-driven procedure estimate contains mechanisms A and B for decays in both the standard FV and upstream regions. By comparing the estimates from the data-driven approach and upstream simulations in the μ^- channel (where there are sufficient statistics to do so meaningfully) good agreement is observed, see figures 7.34b and 7.35b. This comparison confirms that the primary mechanism for the 'high invariant mass tail' of



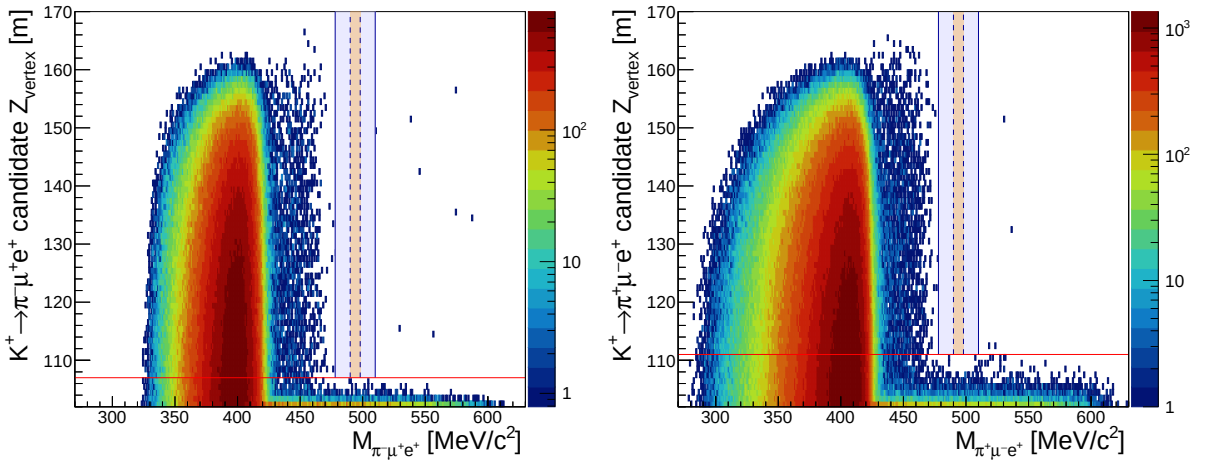
(a) (b)

Figure 7.32: Reconstructed vertex position, z_{vtx}, versus invariant mass, $M_{\pi\mu e}$, for simulated upstream $K^+ \to \pi^+\pi^+\pi^-$ decays selected with no z_{vtx} cut and using forced $\pi^\pm \Rightarrow e^\pm$ misID for events passing signal selections for (a) $K^+ \to \pi^-\mu^+e^+$ and (b) $K^+ \to \pi^+\mu^-e^+$. The signal and blinded regions are indicated by the yellow and grey regions respectively and the minimum z_{vtx} cut chosen for each case is indicated by the red line.

the $K^+ \to \pi^+\pi^+\pi^-$ background, which can enter the blinded region or above, is from upstream decays with a $\pi^\pm \Rightarrow e^\pm$ misID.

The way the $K^+ \to \pi^+\pi^+\pi^-$ background predictions, determined using the various methods described above, are combined is summarised in figures 7.34 and 7.35. By comparing figure 7.34, which contains only simulation studies, to its counterpart including data-driven background estimates in the high invariant mass tail, figure 7.35, the similar background description of the tail can be observed. The relative contributions of the four mechanisms is apparent from these figures, and matches the expectations derived from misidentification and decay in flight probabilities shown in table 7.13. Furthermore the fact that systematic uncertainties are added in the 'simulation studies only' figure (7.34) while only statistical errors are shown in the other (figure 7.35) illustrates that for the $K^+ \to \pi^+\pi^+\pi^-$ background systematic uncertainties, primarily from the forced misidentification procedure, dominate. These uncertainties are highly correlated between bins. It is also clear from these figures, which include the observed number of data events outside blinded regions, that the $K^+ \to \pi^+\pi^+\pi^-$ background is overwhelmingly dominant below $\sim 430\,\mathrm{MeV}/c^2$ and represents the dominant background towards the signal region and above in the μ^- channel. The final background estimates for $K^+ \to \pi^+\pi^+\pi^-$ are summarised in tables 7.14 and 7.15 for the π^- and μ^- channels respectively. For these results the corrections discussed in section 7.5.8 are included where applicable.

In the selection a cut is applied on $z_{min,i}$, the minimum z at which a vertex for signal channel i can be selected. The values of $z_{min,i}$ were chosen based on studies of simulations to minimise the expected background in the signal region. Figure 7.32 shows the distribution of z_{vtx} versus $M_{\pi\mu e}$ for the selected sample of simulated upstream $K^+ \to \pi^+\pi^+\pi^-$ decays with no requirement for z_{vtx} applied and using forced $\pi^\pm \Rightarrow e^\pm$ misID. The chosen values of $z_{min,i}$ are set at the smallest integer number of meters for which less than two events remain in the signal region. This leads to the choice of $z_{min,\pi} < 107\,\mathrm{m}$ and $z_{min,\mu} < 111\,\mathrm{m}$ as shown in figure 7.33 with events passing the $K \to \pi\mu e$ selection (and the blinded region masked).



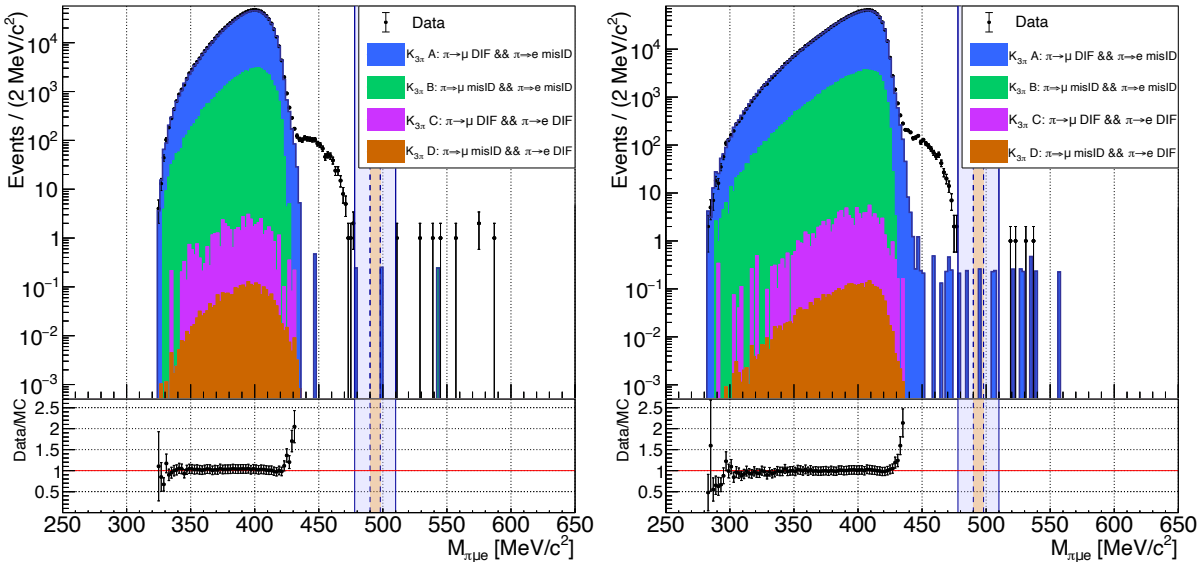
(a) (b)

Figure 7.33: Reconstructed vertex position versus $M_{\pi\mu e}$, for events passing the $K^+ \to \pi^-\mu^+e^+$ (a) and $K^+ \to \pi^+\mu^-e^+$ (b) selections with blinded regions (shown in grey with signal region in yellow) masked. The minimum z_{vtx} cut for each case is indicated by the red line.

Note on Presentation of Background Predictions

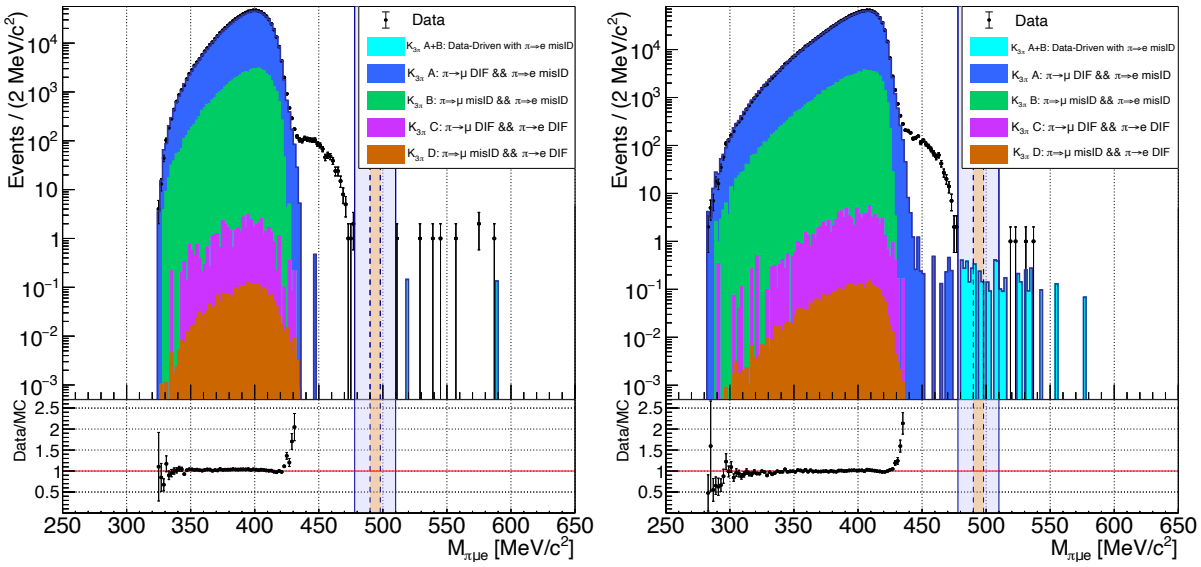
In cases with limited statistics, as with the studies of $K^+ \to \pi^+\pi^+\pi^-$ background mechanisms C and D in regions above $478\,\mathrm{MeV}/c^2$ for example, situations arise where no events pass the signal selection. For these cases the result is displayed as if one event was observed with an uncertainty of 100% (this is found to be statistically equivalent when using Bayesian methods when summing background contributions), appropriate scaling factors are then applied. The result, in the form $x \pm x$, is then combined with other results as normal, summing the central value and propagating the uncertainties.

This rule is adopted consistently for this analysis. This does lead to some figures in some tables combining in a non-trivial fashion. As a concrete example consider the $K^+ \to \pi^+\pi^+\pi^-$ background mechanism C for the μ^- channel in regions above $478\,\mathrm{MeV}/c^2$, no simulated events of this type pass the selection in the full blinded region (which includes the signal region). Following the above rule it is assumed 1 ± 1 event is selected and scaling factors are applied as outlined in section 7.5.1, where the $\pi^\pm \Rightarrow e^\pm$ misID probability weight, w_{PID}, is taken to be a maximal value giving an answer of 0.11 ± 0.11. The argument is the same for the signal region, so here the prediction 0.11 ± 0.11 is reported. This does not imply that the whole background is contained in the signal region alone and does not imply one should sum the two, each is evaluated independently and is relevant specifically to the region it is calculated for.



(a) (b)

Figure 7.34: Detailed $K^+ \to \pi^+\pi^+\pi^-$ background estimates before unblinding, broken down into the four mechanisms (see table 7.8 and figure 7.30) obtained using studies of simulations with various special settings (see text). The π^- and μ^- channels are shown by (a) and (b) respectively. The uncertainties displayed are a combination of statistical and systematic and the observed data is shown in the control regions for comparison.



(a) (b)

Figure 7.35: Detailed $K^+ \to \pi^+\pi^+\pi^-$ background estimates before unblinding, broken down into the four mechanisms (see table 7.8 and figure 7.30) obtained using studies of simulations and the data-driven approach for $M_{\pi\mu e} > 487\,\mathrm{MeV}/c^2$. The π^- and μ^- channels are shown by (a) and (b) respectively. The uncertainties displayed are statistical and the observed data is shown in the control regions.

Table 7.14: Background predictions: detailed breakdown of $K^+ \to \pi^+\pi^+\pi^-$ background mechanisms for $K^+ \to \pi^-\mu^+e^+$ search.

Sample	Control Regions		Blinded Region	Signal Region
	N_{CRB}	N_{CRA}	N_{BR}	N_{SR}
Data	857074	8	BLIND	BLIND
A (+B Data-Driven + Upstream)	$(7.77 \pm 0.78) \times 10^5$	0.28 ± 0.20	0.098 ± 0.098	0.098 ± 0.098
B (FV MC)	$(5.48 \pm 0.55) \times 10^4$	–	–	–
C (MC)	46.2 ± 5.4	0.11 ± 0.11	0.34 ± 0.20	0.11 ± 0.11
D (MC)	2.48 ± 0.25	$(3.2 \pm 3.2) \times 10^{-3}$	$(3.2 \pm 3.2) \times 10^{-3}$	$(3.2 \pm 3.2) \times 10^{-3}$
Total	$(8.34 \pm 0.78) \times 10^5$	0.36 ± 0.21	0.40 ± 0.23	0.22 ± 0.15
Data/Prediction	1.031 ± 0.097	–	–	–

Table 7.15: Background predictions: detailed breakdown of $K^+ \to \pi^+\pi^+\pi^-$ background mechanisms for $K^+ \to \pi^+\mu^-e^+$ search.

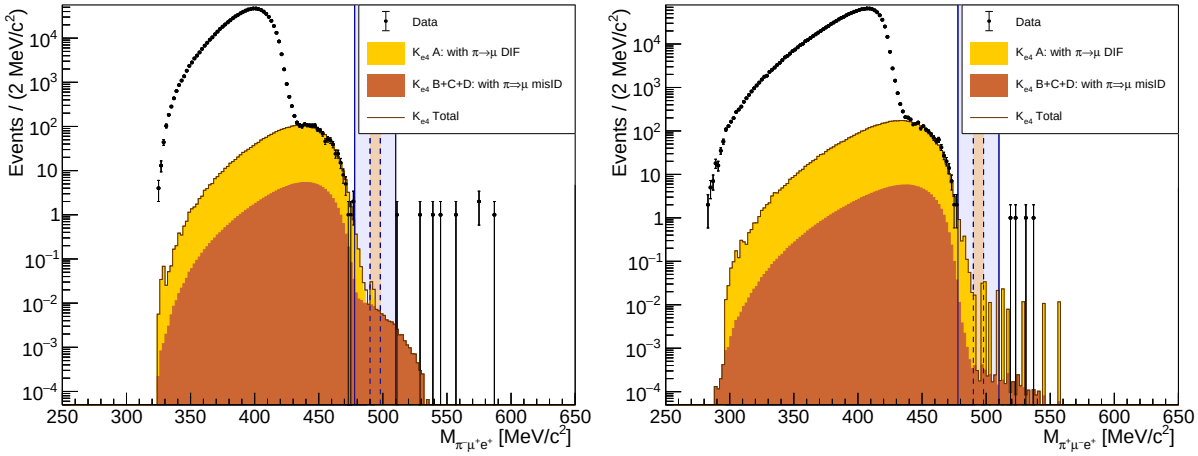
Sample	Control Regions		Blinded Region	Signal Region
	N_{CRB}	N_{CRA}	N_{BR}	N_{SR}
Data	1376260	4	BLIND	BLIND
A (+B Data-Driven + Upstream)	$(1.29 \pm 0.13) \times 10^5$	1.64 ± 0.46	3.24 ± 0.73	0.72 ± 0.32
B (FV MC)	$(7.80 \pm 0.78) \times 10^4$	–	–	–
C (MC)	97 ± 10	0.11 ± 0.11	0.11 ± 0.11	0.11 ± 0.11
D (MC)	3.32 ± 0.34	$(3.2 \pm 3.2) \times 10^{-3}$	$(3.2 \pm 3.2) \times 10^{-3}$	$(3.2 \pm 3.2) \times 10^{-3}$
Total	$(1.37 \pm 0.13) \times 10^6$	1.76 ± 0.48	3.35 ± 0.74	0.84 ± 0.34
Data/Prediction	1.007 ± 0.095	–	–	–

7.5.7.2 Treatment of $K^+ \to \pi^+\pi^-e^+\nu_e$ Background

There are two primary background mechanisms for $K^+ \to \pi^+\pi^-e^+\nu_e$ decays, with two much more suppressed mechanisms for the π^- channel, as shown schematically in figure 7.36. To first order the background from $K^+ \to \pi^+\pi^-e^+\nu_e$ decays can be well-described by analysing the large sample of 2.8×10^9 simulated events available. This is because the background is dominated by the primary mechanism (A) with a single $\pi^\pm \to \mu^\pm$ decay in flight producing a signal-like signature and this kinematic mechanism is well simulated. However, better precision and understanding of the sub-leading background mechanism (B) with $\pi^\pm \Rightarrow \mu^\pm$ misID can be achieved by applying forced $\pi^\pm \Rightarrow \mu^\pm$ misID to events with no decays in flight. To sum the two mechanisms and avoid double counting $K^+ \to \pi^+\pi^-e^+\nu_e$ events must be split into two categories : 1) with no decays in flight, where $\pi^\pm \Rightarrow \mu^\pm$ misID gives background mechanism B (and C) contribution, and 2) other events (at least 1 decay in flight), where standard analysis of simulations constrains other background mechanisms (A and D). Mechanisms C and D are very unlikely and so to good approximation the two components can be combined with mechanisms B and A respectively and considered negligible in comparison. The two components of the background are displayed in figures 7.37a and 7.37b and the corresponding numerical background estimates are provided in tables 7.16 and 7.17 for the π^- and μ^- channels respectively. The $K^+ \to \pi^+\pi^-e^+\nu_e$ background makes up the majority of the shoulder observed, between the main $K^+ \to \pi^+\pi^+\pi^-$ component and the blinded region, in data for both signal channels of interest.

Bkg. Mechanism \ Signal Channel	$K^+ \to \pi^-\mu^+e^+$ Background	$K^+ \to \pi^+\mu^-e^+$ Background
A $\pi \to \mu$ DIF	$K^+ \to \pi^+\pi^-e^+\nu_e$ ↓ μ^+ DIF	$K^+ \to \pi^+\pi^-e^+\nu_e$ ↓ μ^- DIF
B $\pi \Longrightarrow \mu$ misID	$K^+ \to \pi^+\pi^-e^+\nu_e$ ⇣ μ^+ mis-ID	$K^+ \to \pi^+\pi^-e^+\nu_e$ ⇣ μ^- mis-ID
C $\pi \Longrightarrow \mu$ misID + $\pi \Longrightarrow e$ misID	$K^+ \to \pi^+\pi^-e^+\nu_e$ ⇣ ⇣ e^+ mis-ID μ^+ mis-ID	
D $\pi \Longrightarrow \mu$ misID + $\pi \to e$ DIF	$K^+ \to \pi^+\pi^-e^+\nu_e$ ↓ ⇣ e^+ DIF μ^+mis-ID	

Figure 7.36: Schematic illustration of the background mechanisms through which $K^+ \to \pi^+\pi^-e^+\nu_e$ decays can obtain a particle identity signature like a signal $K^+ \to \pi\mu e$ decay.



(a) (b)

Figure 7.37: Detailed $K^+ \to \pi^+\pi^-e^+\nu_e$ background estimates, before unblinding, showing the two primary mechanisms (see table 7.8 and figure 7.36) obtained using studies of simulations. The π^- and μ^- channels are shown by (a) and (b) respectively. The uncertainties displayed are statistical and the observed data is shown in the control regions for comparison.

Table 7.16: Background predictions: detailed breakdown of $K^+ \to \pi^+\pi^-e^+\nu_e$ background mechanisms for $K^+ \to \pi^-\mu^+e^+$ search.

Sample	Control Region		Blinded Regions	Signal Region
	N_{CRB}	N_{CRA}	N_{BR}	N_{SR}
With decays in flight	$(2.71 \pm 0.27) \times 10^3$	0.019 ± 0.019	0.622 ± 0.094	0.034 ± 0.020
With misidentification	142 ± 14	0.0135 ± 0.0014	0.127 ± 0.013	0.0314 ± 0.0032
Combined	$(2.85 \pm 0.28) \times 10^3$	0.033 ± 0.019	0.75 ± 0.12	0.065 ± 0.021

Table 7.17: Background predictions: detailed breakdown of $K^+ \to \pi^+\pi^-e^+\nu_e$ background mechanisms for $K^+ \to \pi^+\mu^-e^+$ search.

Sample	Control Region		Blinded Regions	Signal Region
	N_{CRB}	N_{CRA}	N_{BR}	N_{SR}
With decays in flight	$(5.13 \pm 0.51) \times 10^3$	0.064 ± 0.031	1.86 ± 0.17	0.047 ± 0.028
With misidentification	174 ± 17	$(2.48 \pm 0.33) \times 10^{-3}$	0.0219 ± 0.0023	$(1.29 \pm 0.21) \times 10^{-3}$
$K^+ \to \pi^+\pi^-e^+\nu_e$	$(5.30 \pm 0.53) \times 10^3$	0.067 ± 0.032	1.88 ± 0.25	0.048 ± 0.028

7.5.7.3 Treatment of $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ Background (μ^- Channel)

The Dalitz decays are potentially dangerous backgrounds because of their large branching ratios and 3-charged-track signature. However, in the π^- channel case the DalitzRejection cut reduces these to a negligible level. A numerical estimate, or powerful upper limit, is obtained using multiple forced misID.

Meanwhile in the μ^- channel these backgrounds are generally very small because the negative track is an e^- and the probability for $e^- \Rightarrow \mu^-$ misID is low. Similar multiple forced misID procedures can constrain $K_{\mu 3D}$ and K_{e3D} background and show they are negligible. However, as shown above the $e^- \Rightarrow \mu^-$ misID probability has been measured and can be $\mathcal{O}(10^{-8})$ and provides a mechanism for non-negligible $K_{2\pi D}$ background in the μ^- channel where only this single misidentification is required. The background in this case is estimated using forced $e^- \Rightarrow \mu^-$ misID as discussed in section 7.5.5.2. This background contribution is reduced significantly by the use of the LAV photon veto requirement, as shown by figure 7.38.

7.5.7.4 Treatment of $K^+ \to \pi^+e^+e^-$ and $K^+ \to \mu^+\nu_\mu e^+e^-$ Backgrounds (π^- Channel)

The dominant background mechanisms for $K^+ \to \pi^+e^+e^-$ (A & B, the others are negligible) in the π^- channel include a $\pi^+ \to \mu^+$ DIF (A) or $\pi^+ \Rightarrow \mu^+$ misID (B) combined with an $e^- \Rightarrow \pi^-$ misID. The sum can be constrained analysing standard simulations with forced $e^- \Rightarrow \pi^-$ misID. A number of the events selected in analysis of data in in the upper control region (CRA) in the π^- channel have properties consistent with $K^+ \to \pi^+e^+e^-$ events, see figure 7.39. Some clearly enter through $\pi^+ \Rightarrow \mu^+$ misID (mechanism B) since the invariant mass calculated in the $\pi^+e^+e^-$ hypothesis is consistent with the kaon mass,

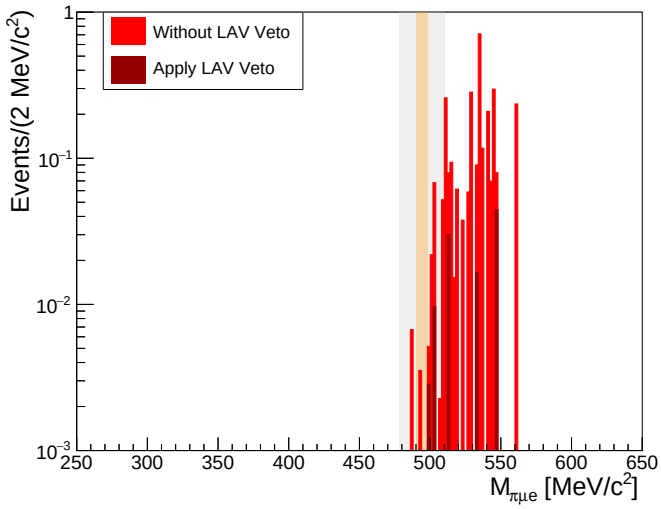


Figure 7.38: Comparison of expected number of $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ background events as a function of invariant mass for the μ^- channel with and without applying LAV photon veto.

while others with this invariant mass further from the kaon mass may be of mechanism A but with a soft neutrino in the $\pi^+ \to \mu^+\nu_\mu$ DIF.

The situation for $K^+ \to \mu^+\nu_\mu e^+e^-$ decays is effectively the same but only $e^- \Rightarrow \pi^-$ misID is required, with no additional decay in flight or misidentification needed since a μ^+e^+ pair is already present. Forced $e^- \Rightarrow \pi^-$ misID therefore is used for background estimation in the π^- channel as for $K^+ \to \pi^+e^+e^-$.

An independent analysis has been performed which cross-checks and validates the results reported for the analysis described here. For the study of these backgrounds which involve $e^\pm \Rightarrow \pi^\pm$ misidentification some tension was observed between the two analyses. This was found to be due to differences in the misID model (shown for this analysis in figure 7.21) of up to 30%, but with momentum dependence consistent within uncertainties. The statistics available for the measurement are relatively low and this misID model is measured in a sample most prone to potential contamination. To account for these differences, and reflect the impact on the background estimate, the average of the results of the two analyses is used for the final background estimates, with an additional systematic uncertainty equal to half the difference. The results of the two analyses and the averaged results are shown in tables 7.18 and 7.19 for the $K^+ \to \pi^+e^+e^-$ and $K^+ \to \mu^+\nu_\mu e^+e^-$ decays respectively.

Table 7.18: Background predictions: detailed breakdown of $K^+ \to \pi^+e^+e^-$ background mechanisms for $K^+ \to \pi^-\mu^+e^+$ search.

Sample	Control Regions		Blinded Region	Signal Region
	N_{CRB}	N_{CRA}	N_{BR}	N_{SR}
This Analysis	0.317 ± 0.034	4.26 ± 0.43	2.04 ± 0.21	0.517 ± 0.054
Alternative Analysis	0.376 ± 0.046	4.57 ± 0.46	2.81 ± 0.29	0.738 ± 0.080
Averaged Results	0.346 ± 0.047	4.42 ± 0.47	2.42 ± 0.46	0.63 ± 0.13

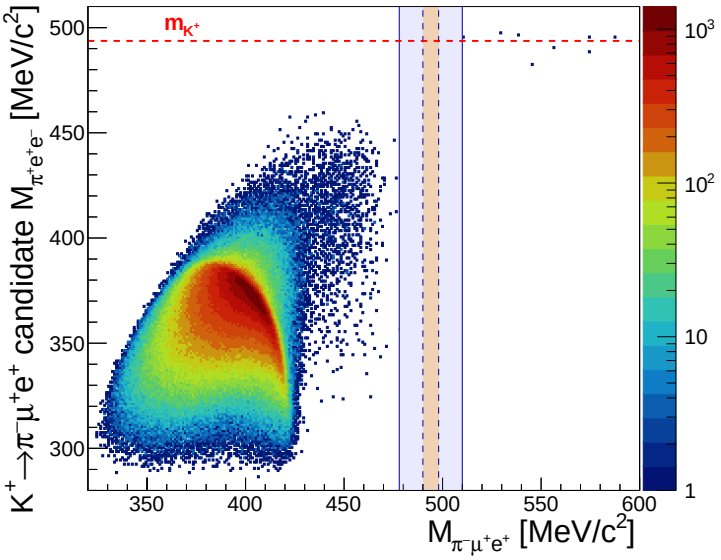


Figure 7.39: Two dimensional histogram showing invariant mass of events passing the $K^+ \to \pi^-\mu^+e^+$ selection under the $\pi^-\mu^+e^+$ and $\pi^+e^+e^-$ particle hypotheses. The blinded region is shown in grey and is masked including the signal region shown in yellow, the K^+ mass is indicated in the $\pi^+e^+e^-$ hypothesis by the red dashed line.

Table 7.19: Background predictions: detailed breakdown of $K^+ \to \mu^+\nu_\mu e^+e^-$ background mechanisms for $K^+ \to \pi^-\mu^+e^+$ search.

Sample	Control Regions		Blinded Region	Signal Region
	N_{CRB}	N_{CRA}	N_{BR}	N_{SR}
This Analysis	0.850 ± 0.092	0.622 ± 0.069	0.454 ± 0.052	0.120 ± 0.018
Alternative Analysis	1.12 ± 0.13	0.653 ± 0.082	0.547 ± 0.070	0.141 ± 0.027
Averaged Results	0.98 ± 0.17	0.639 ± 0.072	0.501 ± 0.073	0.131 ± 0.021

7.5.8 Corrections Applied to Simulations

Equation 7.5.1 shows that the prediction of the number of background events derived from simulation studies contains a number of weight terms, these are corrections applied to simulations because of measured differences between simulations and data. The w_{MC} term is a weight applied to account for special simulation settings, such as modified $\pi^\pm$ decay branching ratio or modified K^+ decay volume, and has already been described in preceding sections (principally section 7.5.7.1). The w_{PID} term is the PID probability, as measured in section 7.5.4, importantly applied in the forced misidentification procedure described in section 7.5.5. The remaining weight terms are described below.

7.5.8.1 Accounting for Trigger Efficiencies

The trigger efficiencies for the data samples used for the $K^+ \to \pi\mu e$ searches were measured in section 7.3. The inefficiencies of the two primary triggers for the analysis, $MT\mu$

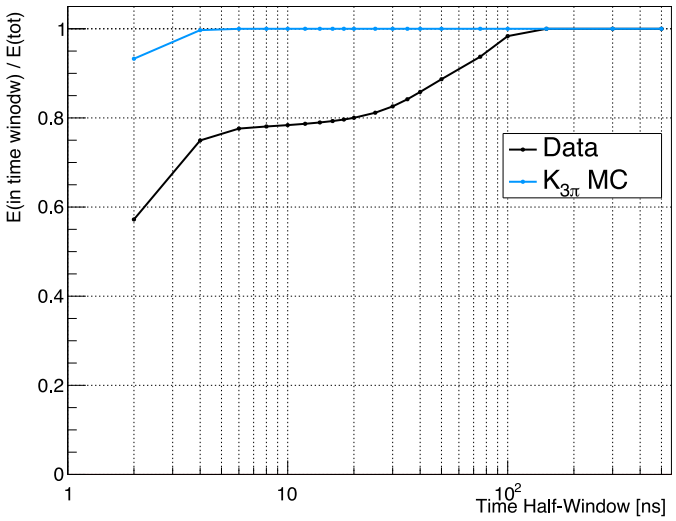


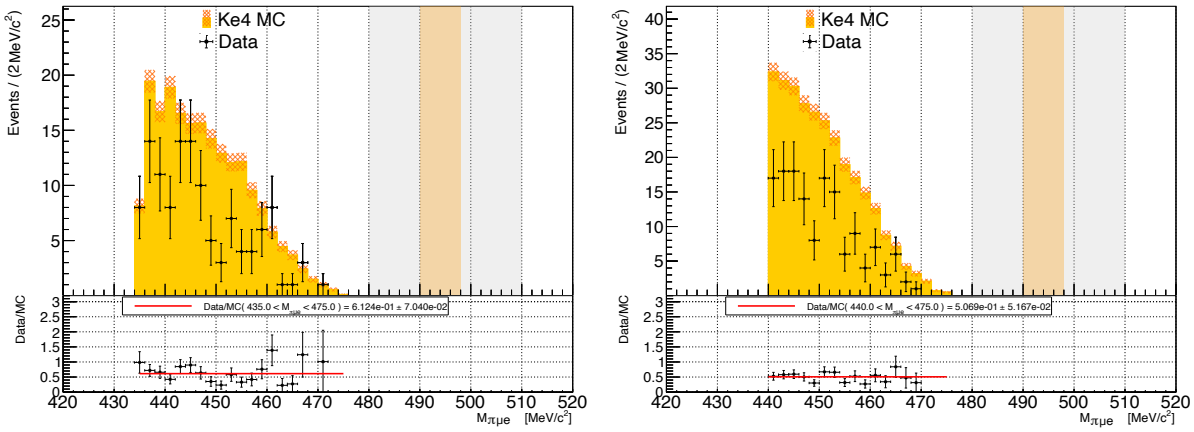
Figure 7.40: Fraction of total energy deposited in LKr inside a time window, defined by half window width t_{hw} ($|t_{trig} - t_{LKr}| < t_{hw}$) for all data from the MUE3TV filter (see section 7.1.3) and $K^+ \to \pi^+\pi^+\pi^-$ MC simulations. The $K^+ \to \pi^+\pi^+\pi^-$ decay is the normalisation channel and the main control region background component.

and MTe, are non-negligible and must be accounted for. However, the trigger system is not present in simulations, emulation of L0 components is in development and L1 algorithms can be run using the HLT software (discussed in section 7.3.2.3) but only in the latest simulations. For this analysis it was therefore mandatory to produce a bespoke emulation of the trigger system able to reproduce its performance for the specific case of the $K^+ \to \pi\mu e$ searches.

As shown in section 7.3 most trigger components, L0 RICH, Qx, MO1 and L1 KTAG and !LAV, have high efficiency over 97%, with only weak dependence over time (as a function of run number) and on event properties (as a function of properties such as track momentum, charge, or position). The L1 STRAW algorithm efficiencies have some momentum dependence but to first order these effects cancel because they appear in both the normalisation and signal triggers. However, the efficiency of the L0 LKr10 and LKr20 trigger components have a significant energy dependence and since these appear only in the signal $MT\mu$ and MTe triggers respectively these must be accounted for.

The L0 LKr10 and LKr20 trigger components are satisfied if a certain total energy deposition in the LKr, E^{tot}_{LKr} is recorded within a certain time window. However, this E^{tot}_{LKr} variable is poorly reproduced in simulations because of additional pileup activity, noise and choice of timing windows. This is illustrated by figure 7.40 which shows the fraction of E^{tot}_{LKr} deposited as a function of time window widths for data and $K^+ \to \pi^+\pi^+\pi^-$ simulations for selected normalisation events. The significant difference observed here results in a dramatic mis-modelling of the background when using E^{tot}_{LKr} to define a trigger efficiency event weight. This is seen most clearly for the $K^+ \to \pi^+\pi^-e^+\nu_e$ background in figure 7.41 since for this background the trigger efficiency weight is the most significant correction. The total energy deposit is therefore not an appropriate variable to use to perform the correction for trigger inefficiencies.

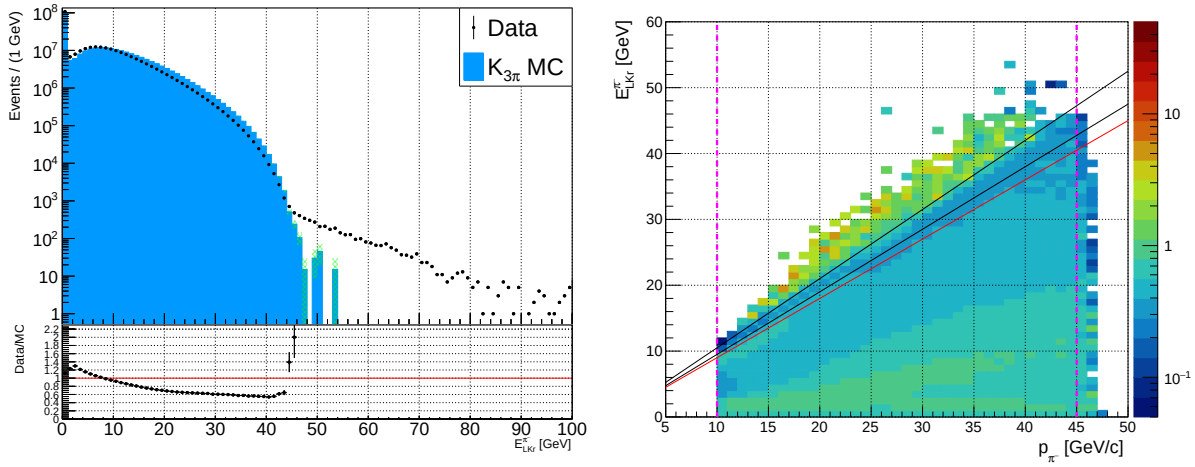
The total energy deposited in the LKr which is associated to the three selected



(a) (b)

Figure 7.41: Comparison, in the selected range of the control region dominated by the $K^+ \to \pi^+\pi^-e^+\nu_e$ (K_{e4}) background, between a sub-sample of 2017 data (samples 17A+17B) and MC simulations as a function of invariant mass for the π^- (a) and μ^- (b) channels. Uncertainties for simulations are indicated by hatching. The significant mis-modelling of the background when applying trigger efficiency event weights based on total LKr energy deposition is displayed. The integrals ratio of observed/predicted number of events is 0.61 ± 0.07 and 0.51 ± 0.05 for the π^- and μ^- channels respectively.

tracks from the common selection, E^{3trk}_{LKr}, is also not well-reproduced by simulations. For a candidate signal event $E^{3trk}_{LKr} = E_\pi + E_\mu + E_e$. Simulations reproduce well E_μ, which is low since μ acts as a MIP, and E_e where the e deposits effectively all of its energy in the LKr through fully electromagnetic showers. However, studies comparing the pion energy deposition in the LKr, E_π, for well-separated π^- from $K^+ \to \pi^+\pi^+\pi^-$ decays, shown in figure 7.42a, demonstrate that E_π is not well-reproduced in the simulations. For $\pi^\pm$ a more complex mixture of hadronic and electromagnetic showers are possible and clusters in the LKr can be split and therefore some fraction of the energy deposition may not be associated to the candidate $\pi^\pm$ track. Taking the aforementioned sample of π^- candidates the energy deposition in the LKr is measured as a function of the momentum of the track. Figure 7.42b shows the ratio between data and simulations for the two dimensional plane of E_π versus p_π (for π^- tracks). A measurement is made of the probability of an energy deposition E_π for a given momentum p_π, $P(E_\pi, p_\pi)$, in two-dimensional bins of size $1\,\text{GeV} \times 1\,\text{GeV}/c$. The results are shown in figure 7.43 where each momentum bin/column in the 2D histogram is normalised independently to 1 such that the entry in a given (x, y) bin corresponds to the probability $P(E_\pi, p_\pi)$, effectively displaying a family of probability density functions of E_π for a range of p_π. The track momentum is well reproduced by simulations, and therefore based on the reconstructed momentum of a π track and a random number generator an 'emulated' pion energy, E^{emul}_π, can be assigned based on drawing a value from the E_π PDF for the corresponding p_π (see figure 7.43). Since this procedure is based on the reconstructed track momentum it may be performed for both data and simulations. Applying the procedure to data acts as a check: the measured E_{π^-} is replaced with an emulated value for normalisation $K^+ \to \pi^+\pi^+\pi^-$ events, and the resulting E_π distributions are compared. The procedure is verified if the reconstructed and emulated distributions are compatible. When applied to $K^+ \to \pi^+\pi^+\pi^-$ simulations



(a) (b)

Figure 7.42: (a) Comparison of the pion energy deposition spectrum for π^- tracks in selected normalisation $K^+ \to \pi^+\pi^+\pi^-$ events between data collected with the minimum bias three-track (MT) trigger in data and $K^+ \to \pi^+\pi^+\pi^-$ MC simulations. The simulations distribution is normalised to the integral of the data but clear differences in shape are seen in the ratio plot (lower panel). Uncertainties are statistical and are indicated by hatched green regions for the $K^+ \to \pi^+\pi^+\pi^-$ simulations. (b) Ratio between number of data and simulation events (as in (a)) in the two dimensional plane of π^- energy deposition versus reconstructed momentum.

the emulated E_{π^-} distribution can be compared to the distribution for reconstructed E_{π^-} from data. These tests are summarised in figure 7.44 which shows the reconstructed and emulated E_{π^-} distributions in both data and simulations and comparisons. From this figure evidence is provided which: verifies the procedure since the emulated and reconstructed distributions for data are well-matched (middle panel), and demonstrates that emulation with MC simulations reproduces the distribution from data since the ratio (emulated MC)/(reconstructed data) is mostly flat and compatible with 1 (lower panel).

Since the LKr energy deposition is well-reproduced for $\mu^\pm$ and $e^\pm$ and it has now been demonstrated that E_π can be emulated, it is possible to construct a corrected three-track LKr energy: $E^{3trk\prime}_{LKr} = E^{reco}_{\mu} + E^{reco}_{e} + E^{emul}_{\pi}$. Models for the LKr10 and LKr20 trigger component efficiencies have been established as a function of reconstructed three-track LKr energy in data (meaning effects of accidental in-time clusters are accounted for), as shown in section 7.3.2.2. When using simulations for calculating background expectations the inefficiencies of the LKr10 and LKr20 trigger components can be accounted for by applying a weight w_t which is dependent on $E^{3trk\prime}_{LKr}$. The difference of this weight from unity (where $w_t = 1$ corresponds to an assumption of 100% efficient triggers) is mostly determined by the LKr10 and LKr20 components but all other trigger components are included. Moreover, the relative mixture of triggers in the data sample must be considered; the MTe trigger has, in general, a lower downscaling factor than the $MT\mu$ trigger but makes up a smaller fraction of the selected signal-like sample due to the higher inefficiency arising from the LKr20 trigger component. Moreover it is not enough to simply assign a simulated event to one trigger at random based on the relative downscaling

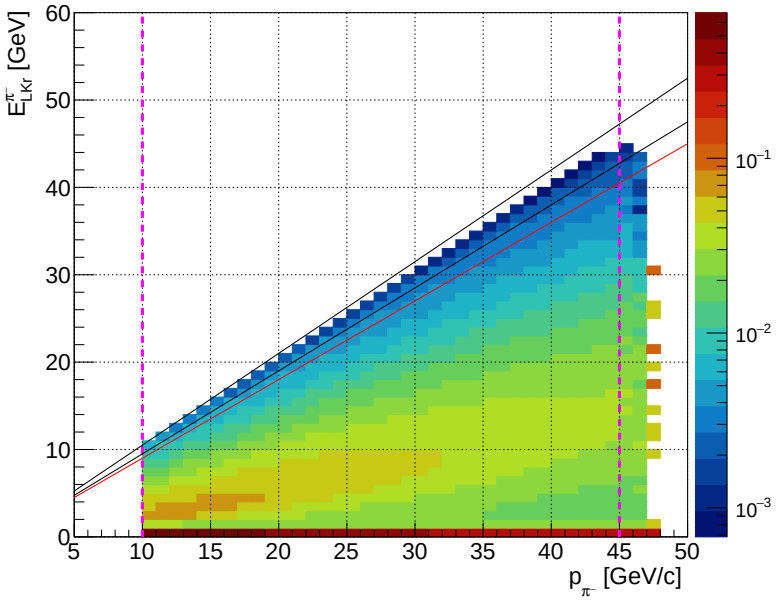


Figure 7.43: LKr pion energy deposition model showing probability for a deposition of energy $E^{\pi^-}_{LKr}$ for a given track momentum p_{π^-}.

factors because events may or may not have properties which 'should' be selected by a particular trigger and the selection may in principle introduce a further bias between triggers (however the analysis cuts are chosen to minimise the latter effect). To account for these features an algorithm has been developed to mimic the trigger selection and inefficiency, the workflow is illustrated in figure 7.45 and this produces an 'emulated trigger efficiency', $\varepsilon^{s}_{eff,emul}(E^{3trk}_{LKr})$, for a specific event which is taken to be the trigger efficiency weight $w_t = \varepsilon^{s}_{eff,emul}(E^{3trk}_{LKr})$. The procedure is found to reproduce the effect of the trigger to percent precision and residual differences are accounted for when calculating the single event sensitivity, as referenced in section 7.4.3.

7.5.8.2 Momentum Corrections

Comparison studies of selected normalisation $K^+ \to \pi^+\pi^+\pi^-$ events, in data and simulations, show differences between the reconstructed Δp and p_T distributions, see figure 7.46. Here Δp is the difference in magnitude between the resultant momentum of the three selected tracks and the average beam K^+ momentum and p_T is the transverse momentum of the three track final state with respect to the average beam momentum (see also section 7.2.1).

To account for the differences the ratio of data/simulations was found in $1\,\mathrm{GeV}/c \times 1\,\mathrm{GeV}/c$ bins of a two-dimensional plane of p_T versus Δp, see figure 7.47a. The ratio in each bin defines a weight $w_{p1}(\Delta p, p_T)$ which is then applied to events in simulation, correcting for the difference. While measurements are made comparing reconstructed variables the weight is applied based on the 'true' (generated) simulation event properties. With a single correction the agreement between simulations and data in the region of interest $|\Delta p| < 2.5\,\mathrm{GeV}/c$ and $p_T < 0.030\,\mathrm{GeV}/c$, improves as shown by figure 7.47b. A

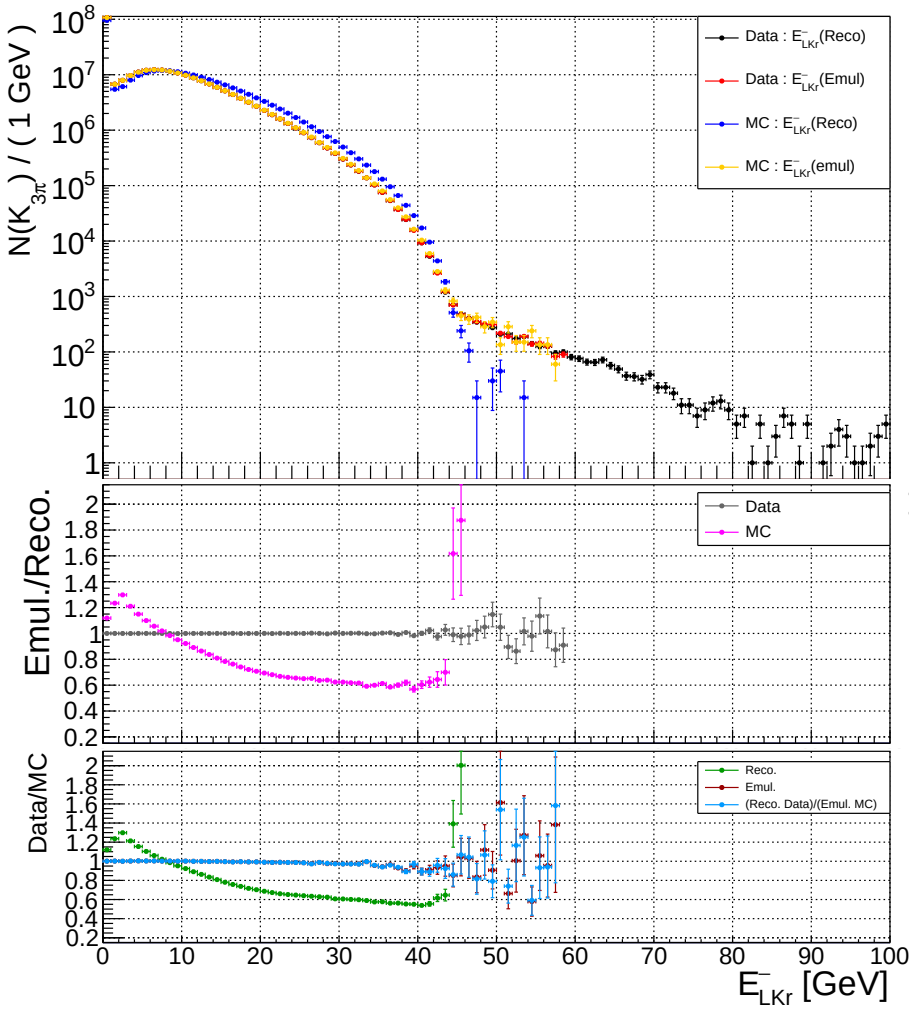


Figure 7.44: Comparison between probability of an LKr energy deposition associated to a π^- track, E^-_{LKr}, for selected normalisation $K^+ \to \pi^+\pi^+\pi^-$ events for cases of reconstructed and emulated values in data and simulations. The central plot shows a comparison between emulated and reconstructed values for data (grey) and MC (magenta). The fact that the ratio for data is close to one verifies that the emulated values closely match the reconstructed values, while the poor agreement of the MC ratio with one illustrates the scale of correction needed for MC. The lower plot shows a comparison between data and MC, first for reconstructed (green) and emulated (dark red) quantities - the former showing the scale of the Data/MC disagreement without emulation and the latter that the emulation behaves similarly between data and MC. It also importantly includes the comparison between reconstructed data and emulated MC which is seen to have a ratio close to one, demonstrating the good performance of the emulation of pion energies.

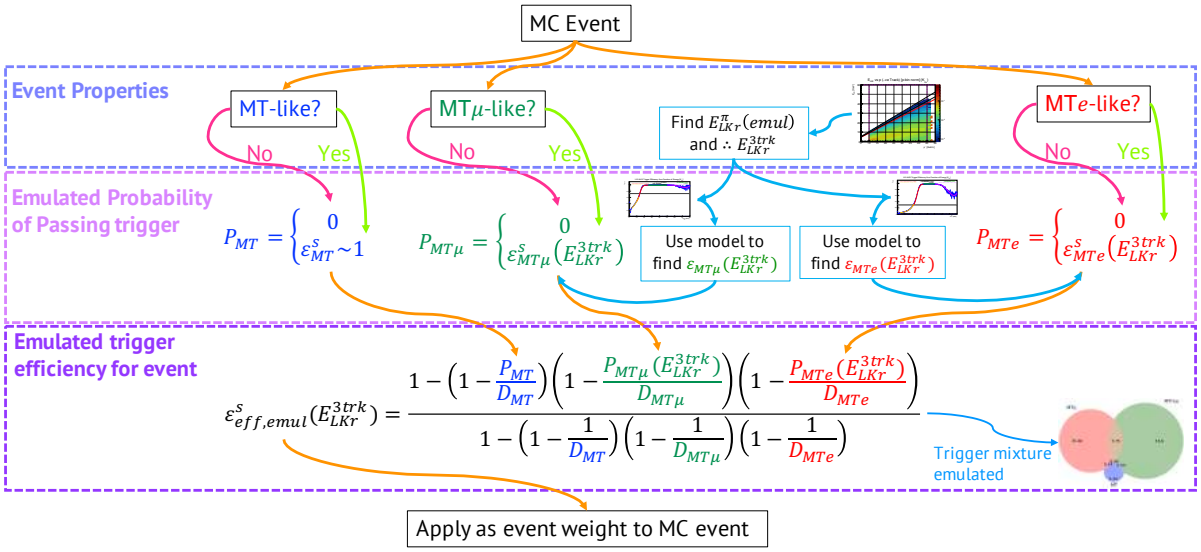


Figure 7.45: Schematic algorithm for applying event weights to MC events to account for trigger inefficiencies.

second round of corrections is then derived from the ratio in the 2D plane after the first correction (figure 7.47b), which defines weight $w_{p2}(\Delta p, p_T)$. Following application of the combined momentum correction weight, $w_p = w_{p1}(\Delta p, p_T)w_{p2}(\Delta p, p_T)$, good agreement is obtained for both variables simultaneously, shown by the ratio in the 2D plane in figure 7.48, and projections onto Δp and p_T in the region of interest, see figure 7.49.

The simultaneous correction of both variables is found to be mandatory since correction of either variable alone tends improve agreement between data and simulations for this variable while degrading agreement for the other variable. This is due to the correlations between them. The momentum weights are applied only in the region of interest for the $K^+ \to \pi\mu e$ search and no events outside this region are used in the analysis.

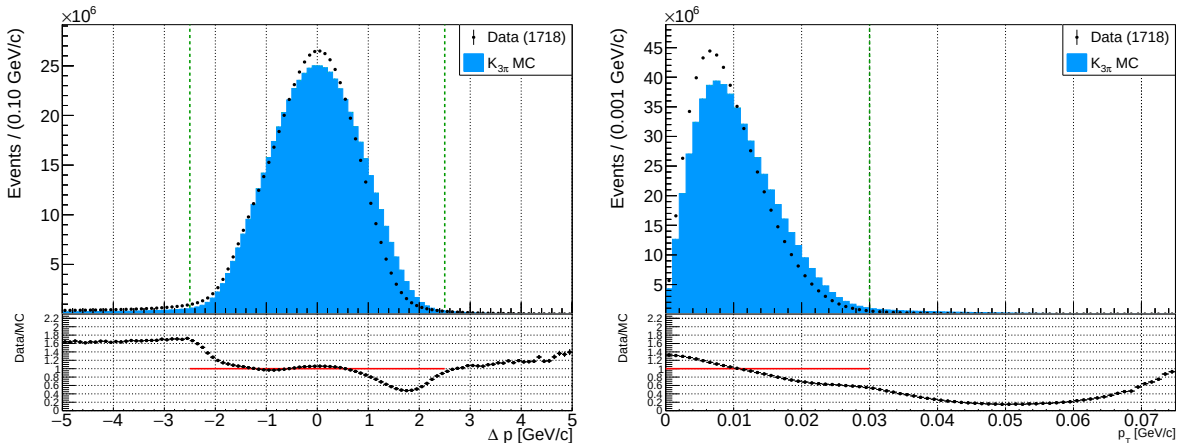
7.5.8.3 Straw Inefficiency Corrections

As part of the common selection, described in section 7.2.1, geometric acceptance conditions are applied to each track, requiring that it passes through a region covered by at least two views at each STRAW station, see figure 7.50. When comparing data and simulations a discrepancy was observed for tracks which pass through the central region of STRAW chambers covered by only two views. The difference has two potential origins:

1. The inefficiency of L1 trigger STRAW algorithm(s).

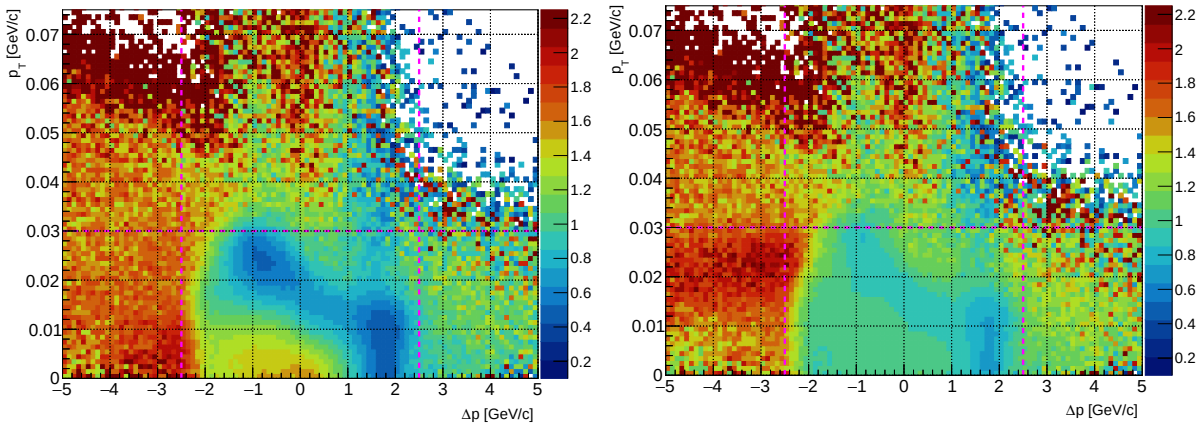
2. Differences between spectrometer reconstruction efficiency in data and simulations, arising from the features of the digitisation of real and simulated detector signals.

By imposing a stricter requirement that tracks must traverse regions covered by at least three views agreement between simulations and data improves but acceptance drops by



(a) (b)

Figure 7.46: Comparison, before applying corrections, between reconstructed Δp (a) and p_T (b) distributions between selected normalisation $K^+ \to \pi^+\pi^+\pi^-$ decay events in data and MC simulations.



(a) (b)

Figure 7.47: Two dimensional map showing ratio, in bins of p_T versus Δp, between selected normalisation $K^+ \to \pi^+\pi^+\pi^-$ decay events in data and MC simulations before any corrections (a) and after first 2D momentum weight, $w_{p1}(\Delta p, p_T)$, is applied (b).

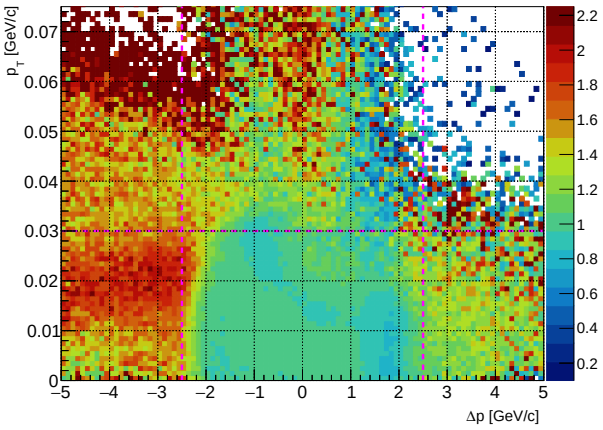
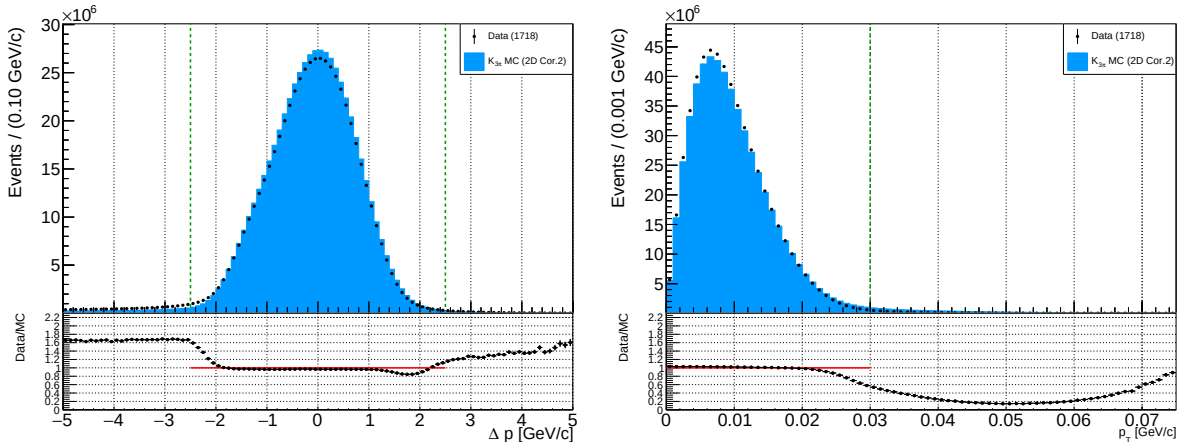
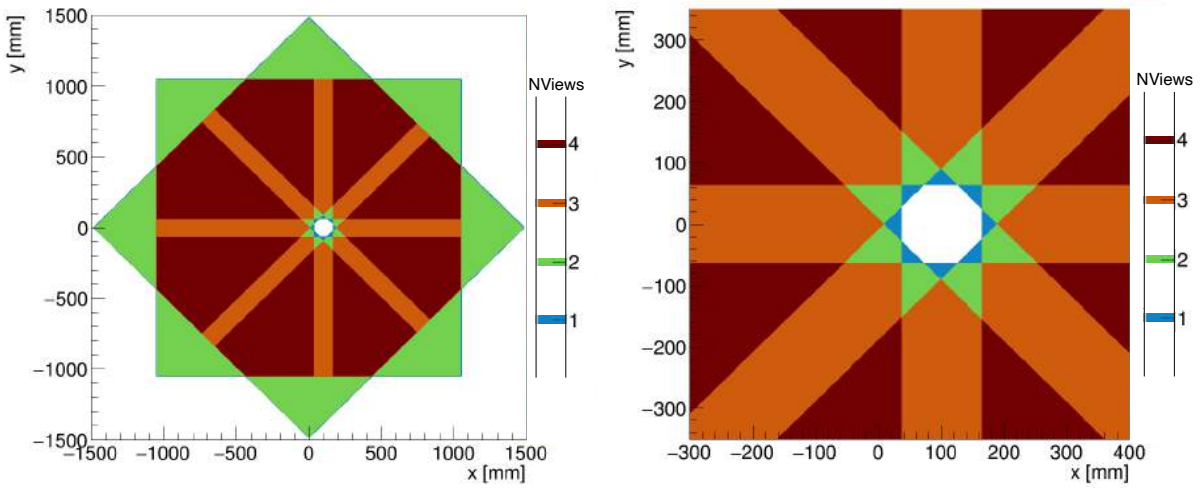


Figure 7.48: Final two dimensional map showing ratio, in bins of p_T versus Δp, between selected normalisation $K^+ \to \pi^+\pi^+\pi^-$ decay events in data and MC simulations after full 2D momentum corrections are applied.



(a) (b)

Figure 7.49: Comparison, after before applying 2D momentum corrections, between reconstructed Δp (a) and p_T (b) distributions between selected normalisation $K^+ \to \pi^+\pi^+\pi^-$ decay events in data and MC simulations.

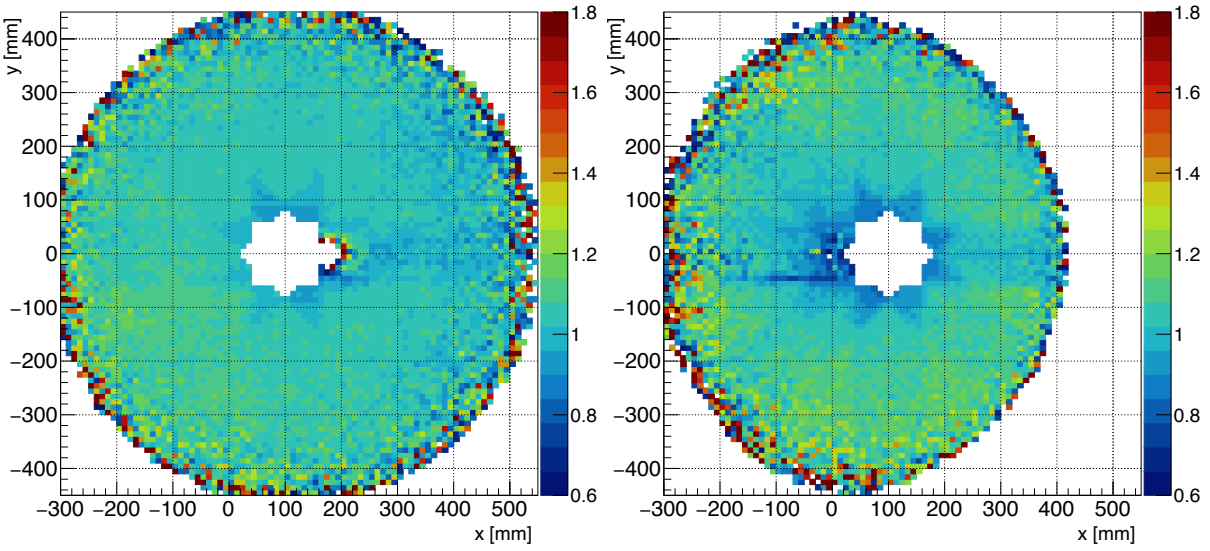


(a) (b)

Figure 7.50: Geometric region covered by 1,2,3 or 4 STRAW views in chamber 1, showing the full plane (a) and a zoomed view of the central region (b). These figures were edited from original versions produced by M. Koval.

45% in relative terms because of the nature of three three-track decays studied. To better account for the above effects, without significant acceptance loss, the original selection requirement is used and an event weight w_S is applied.

Two dimensional maps of (x, y) track position at STRAW1 were produced for selected normalisation $K^+ \to \pi^+\pi^+\pi^-$ events in data and simulations and the ratio data/simulations evaluated for each bin of the map. For this study two of the three pion tracks are required to pass through at least three views, and as such are considered to be a 'good reference track pair' and then the third track is studied as a function of its (x, y) position at STRAW1 and the number of views it traverses in each STRAW chamber. The data/simulations ratio observed as a function of (x, y) at STRAW1 is shown in figures 7.51a and 7.51b for positive and negative pion tracks respectively. The star-like structure seen close to the centre coincides with the region covered by only 2 views, see figure 7.50b. To simplify the event weights, and because the efficiency is not expected to be dependent on (x, y) but rather the number of chambers where only 2 views are traversed, weights are assigned for each track with charge q passing through only 2 views in n STRAW chambers. Weights, w_S, are assigned based on the measured ratio of integrals for data/simulations in each of the eight cases, as shown in table 7.20. By performing the measurements discussed above using MT trigger data (which includes the L1 STRAW algorithm) and comparing to simulations both effects enumerated above are covered simultaneously.



(a) (b)

Figure 7.51: Ratio of the number of normalisation $K^+ \to \pi^+\pi^+\pi^-$ events collected with the MT trigger divided by (scaled) simulations, in 2D bins of (x, y) position at STRAW1. (a) and (b) show the results for for π^+ and π^- tracks respectively, where the other $(\pi^+\pi^\pm)$ tracks both traverse at least 3 views in all four STRAW chambers.

Table 7.20: Straw weights, w_S, measured using integral data/simulations ratio for selected normalisation $K^+ \to \pi^+\pi^+\pi^-$ events

n 2-view chambers	Positive tracks	Negative tracks
1	1.	0.966
2	0.986	0.902
3	0.943	0.835
4	0.915	0.639

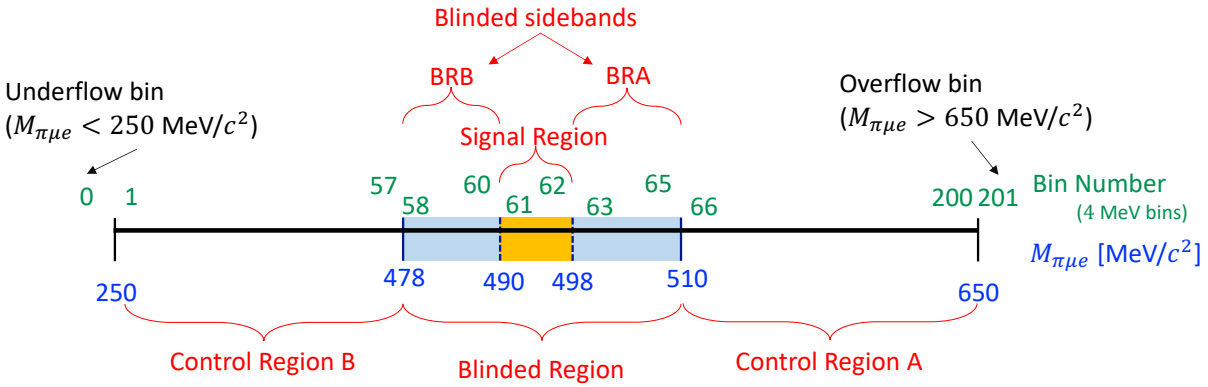


Figure 7.52: Definition of regions of interest in histograms of $M_{\pi\mu e}$.

7.5.9 Background Prediction and Comparison to Data in Control Regions

Using the methods described in preceding sections the number of background events is predicted for six regions of interest (ROI) defined in figure 7.52. The predicted number of events in ROI relevant before unblinding for the π^- and μ^- channels are shown in tables 7.21 and 7.22 respectively (see also tables 7.23 and 7.24 for the remaining ROI) and displayed as a function of $M_{\pi\mu e}$ in corresponding figures 7.53a and 7.53b with comparison to data in control regions (BR masked). The note on presentation of backgrounds from section 7.5.7.1 applies to all presented background tables in this chapter, including tables 7.21 and 7.22.

Agreement in CRB is achieved to better than 3% and 0.3% precision in the π^- and μ^- channels, this is as good as can be expected given the uncertainty on N_K of order 1%. The uncertainty associated with the predicted number of $K^+ \to \pi^+\pi^+\pi^-$ background events is dominated by systematics, specifically from forced misID model uncertainties of 10%, which are highly correlated between adjacent bins. For illustration in figures the $K^+ \to \pi^+e^+e^-$ result using only the analysis presented here is used while for tables the averaged result (see section 7.5.7.4) is displayed.

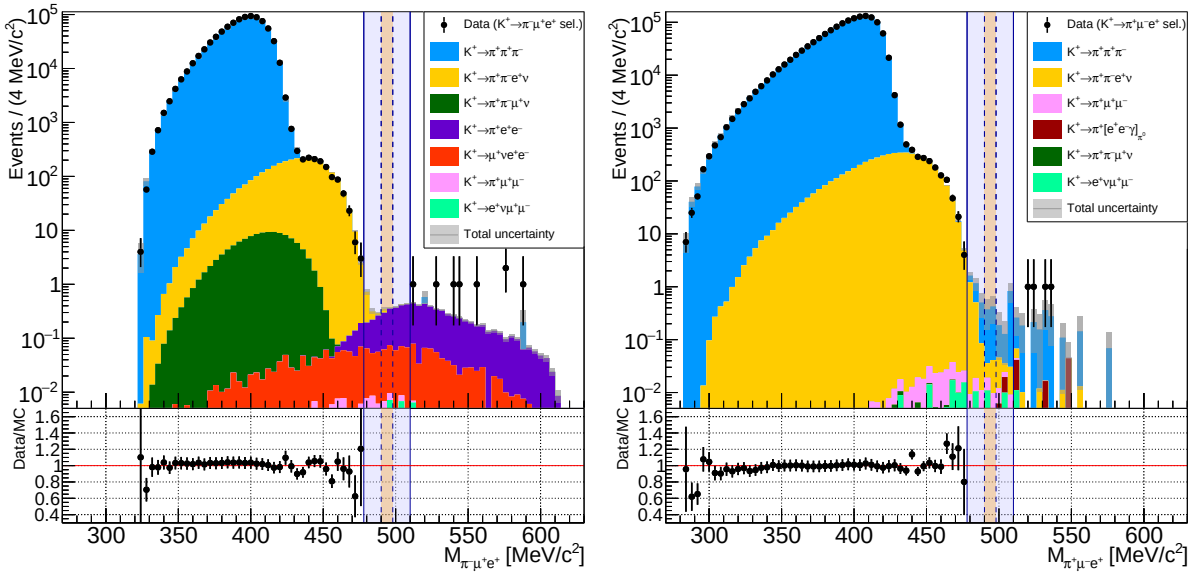
Table 7.21: Background predictions for $K^+ \to \pi^-\mu^+e^+$ search.

Sample	Control Regions		Blinded Region	Signal Region
	N_{CRB}	N_{CRA}	N_{BR}	N_{SR}
Data	857074	8	BLIND	BLIND
$K^+ \to \pi^+\pi^+\pi^-$	$(8.31 \pm 0.78) \times 10^5$	0.40 ± 0.23	0.22 ± 0.15	0.22 ± 0.15
$K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$	$(1.4 \pm 1.4) \times 10^{-9}$	$(4.0 \pm 4.0) \times 10^{-6}$	$(1.4 \pm 1.4) \times 10^{-9}$	$(1.4 \pm 1.4) \times 10^{-9}$
$K^+ \to \mu^+\nu_\mu[e^+e^-\gamma]_{\pi^0}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$
$K^+ \to e^+\nu_e[e^+e^-\gamma]_{\pi^0}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$
$K^+ \to \pi^+\pi^-\mu^+\nu_\mu$	121 ± 12	$(2.7 \pm 2.7) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$
$K^+ \to \pi^+\pi^-e^+\nu_e$	$(2.85 \pm 0.28) \times 10^3$	0.033 ± 0.019	0.75 ± 0.12	0.065 ± 0.021
$K^+ \to \pi^+\mu^+\mu^-$	0.0502 ± 0.0052	$(1.66 \pm 0.30) \times 10^{-3}$	0.0241 ± 0.0026	$(5.33 \pm 0.71) \times 10^{-3}$
$K^+ \to \pi^+e^+e^-$	0.346 ± 0.047	4.42 ± 0.47	2.42 ± 0.46	0.63 ± 0.13
$K^+ \to \mu^+\nu_\mu\mu^+\mu^-$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$
$K^+ \to e^+\nu_e\mu^+\mu^-$	0.0309 ± 0.0095	0.0105 ± 0.0063	0.0310 ± 0.0098	0.0111 ± 0.0067
$K^+ \to \mu^+\nu_\mu e^+e^-$	0.98 ± 0.17	0.638 ± 0.072	0.501 ± 0.073	0.131 ± 0.021
$K^+ \to e^+\nu_e e^+e^-$	$(5.5 \pm 5.5) \times 10^{-7}$	$(1.9 \pm 1.2) \times 10^{-6}$	$(1.6 \pm 1.6) \times 10^{-6}$	$(1.6 \pm 1.6) \times 10^{-6}$
Total background expected	$(8.34 \pm 0.78) \times 10^5$	5.50 ± 0.53	3.95 ± 0.50	1.06 ± 0.20
Data/MC	1.027 ± 0.096	1.46 ± 0.53	–	–

Table 7.22: Background predictions for $K^+ \to \pi^+\mu^-e^+$ search.

Sample	Control Region		Blinded Region	Signal Region
	N_{CRB}	N_{CRA}	N_{BR}	N_{SR}
Data	1376260	4	BLIND	BLIND
$K^+ \to \pi^+\pi^+\pi^-$	$(1.37 \pm 0.13) \times 10^6$	1.76 ± 0.48	3.35 ± 0.74	0.84 ± 0.34
$K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$	$(6.2 \pm 6.2) \times 10^{-6}$	0.091 ± 0.057	0.013 ± 0.010	$(6.2 \pm 6.2) \times 10^{-3}$
$K^+ \to \mu^+\nu_\mu[e^+e^-\gamma]_{\pi^0}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$
$K^+ \to e^+\nu_e[e^+e^-\gamma]_{\pi^0}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$
$K^+ \to \pi^+\pi^-\mu^+\nu_\mu$	0.0512 ± 0.0062	$(2.7 \pm 2.7) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$
$K^+ \to \pi^+\pi^-e^+\nu_e$	$(5.30 \pm 0.53) \times 10^3$	0.067 ± 0.032	1.88 ± 0.25	0.048 ± 0.028
$K^+ \to \pi^+\mu^+\mu^-$	0.203 ± 0.021	$(7.20 \pm 0.89) \times 10^{-3}$	0.0572 ± 0.0059	0.0151 ± 0.0017
$K^+ \to \pi^+e^+e^-$	$(3.4 \pm 3.4) \times 10^{-9}$	0.0140 ± 0.0017	$(2.5 \pm 2.0) \times 10^{-7}$	$(3.4 \pm 3.4) \times 10^{-9}$
$K^+ \to \mu^+\nu_\mu\mu^+\mu^-$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$
$K^+ \to e^+\nu_e\mu^+\mu^-$	0.095 ± 0.019	0.0111 ± 0.0071	0.043 ± 0.011	0.0136 ± 0.0061
$K^+ \to \mu^+\nu_\mu e^+e^-$	$(1.1 \pm 1.1) \times 10^{-8}$	$(1.1 \pm 1.1) \times 10^{-8}$	$(1.1 \pm 1.1) \times 10^{-8}$	$(1.1 \pm 1.1) \times 10^{-8}$
$K^+ \to e^+\nu_e e^+e^-$	$(1.6 \pm 1.6) \times 10^{-10}$	$(6.2 \pm 3.2) \times 10^{-5}$	$(2.9 \pm 2.7) \times 10^{-5}$	$(1.6 \pm 1.6) \times 10^{-10}$
Total background expected	$(1.37 \pm 0.13) \times 10^6$	1.95 ± 0.48	5.35 ± 0.78	0.92 ± 0.34
Data/MC	1.003 ± 0.094	2.1 ± 1.1	–	–

The p value is calculated to quantify the agreement between the predicted and observed number of events. The p value, viewed from a frequentist prospective, quantifies the fraction of times one would observe data as inconsistent or more so with the null hypothesis if the experiment was repeated many times. It is therefore the probability of obtaining data at least as incompatible with the null hypothesis given that the null hypothesis is true [264]. The p value is calculated by simulating a set of pseudo-experiments where the number of events 'observed' is drawn from a Poisson distribution of mean determined (in each pseudo-experiment) by a randomly drawn value from a Gaussian distribution of mean and standard deviation equal to the central value and uncertainty respectively of the total background prediction. The fraction of these pseudo-experiments where the number of 'observed' events is greater than the actual number of events observed in this analysis is the p value (for null hypothesis of observed events being due only to



(a) (b)

Figure 7.53: Expected number of events from background sources and comparison to the selected data events in control regions as a function of invariant mass for the π^- (a) and μ^- (b) channels. The full blinded region, indicated by the grey shaded region is masked and the signal region is indicated by the orange shaded band, see also figure 7.52. The uncertainty of the background prediction is the combined statistical plus systematic uncertainty and is shown as grey boxes overlaid on top of the total background prediction histogram, summing all components.

background sources considered in the total background expectation). As an approximate figure of merit if $p < 0.05$ then the number of observed events appears to be somewhat in tension with the hypothesis that these are due only to the mechanisms considered in the total background expectation. In CRA the p values are 0.14 and 0.19 in the π^- and μ^- channel respectively so there is no clear evidence of inconsistency with the background expectations. Based on these results the decision was made to proceed with a staged opening of the blinded regions as discussed in section 7.6.2.

7.6 Statistical Treatment and Unblinding Procedure

7.6.1 Statistical Treatment

Once the signal region is opened two possibilities exist:

1. A significant excess of events is observed above the background expectations. The significance of the potential observation of signal decays is evaluated and if this is in excess of 5σ discovery can be claimed [13] and the branching ratio measured. The number of events required for observation is evaluated in section 7.6.1.1.

2. The number of events observed is consistent with the background expectations. In this case an upper limit will then be established at 90% confidence level as described in section 7.6.1.2.

7.6.1.1 Discovery Potential

If n events are observed with a total background expectation of $n_b \pm \sigma_b$ the p value and significance for rejection of the background-only hypothesis are calculated performing an ensemble of 10^8 pseudo-experiments with method described in section 7.5.9. The number of events required to claim discovery, meaning to reach a 5σ significance, corresponding to a p-value for the background-only hypothesis of $p < 2.9 \times 10^{-7}$, is found by determining the lowest integer number of events required to obtain a p value at least this small.

For both the π^- and μ^- channels the number of events required to claim discovery is 11. For $n < 5$ no evaluation for potential signal significance will be performed since in this case the p values are greater than 4.6×10^{-3} and 2.7×10^{-3} for the π^- and μ^- channels respectively, corresponding to an excess above the background-only hypothesis of less than 3σ.

7.6.1.2 Setting Upper Limits

After unblinding if $n < 5$ upper limits are established on the branching ratios of $K^+ \to \pi^\pm \mu^\mp e^+$ decays, with a size or significance level of $\alpha = 10\%$ chosen for the statistical test [13]. Upper limits are established using a modified frequentist (hybrid frequentist-Bayesian) 'CLs procedure' [240] treating the two channels independently as single-bin counting experiments with a log likelihood ratio function as the test statistic. The full description of the procedure used is given in appendix I which accounts for uncertainty on the expected background and single event sensitivity. The expected and observed upper limits were calculated for $n_{obs} = 0, 1, 2, 3, 4$ before unblinding.

7.6.2 Unblinding Procedure and Opening Sidebands

Before unblinding the single event sensitivity was determined, full background expectations were developed for all regions of interest, the discovery potential was evaluated and expected and observed upper limits for $n_{obs} < 5$ calculated. Full cross-checks were performed with the independent analysis.

A two-stage unblinding procedure was then used, first opening the sideband blinded regions (BRA and BRB) only and keeping the signal region masked, then opening the signal region itself. Background expectations are further validated using the sidebands. The number of expected and observed events and the p value for the observation under the (null) background-only hypothesis are shown in tables 7.23 and 7.24 for the π^- and μ^- channels respectively.

To further evaluate the compatibility of the observed number of events with the background expectations in the three control regions CRA, BRB and BRA for both channels the p values for combinations of cases were considered. Combination of individual results, which may be treated as independent since regions are mutually exclusive, was achieved using Fisher's method [265] [266] as described in appendix J. The results of this

study are given in table 7.25.

Table 7.23: Background predictions for $K^+ \to \pi^-\mu^+e^+$ with sidebands unblinded.

Sample	Control Region	Blinded Sidebands		Signal Region
	N_{CRA}	N_{BRB}	N_{BRA}	N_{SR}
Data	8	2	4	BLIND
$K^+ \to \pi^+\pi^+\pi^-$	0.40 ± 0.23	0.22 ± 0.15	0.22 ± 0.15	0.22 ± 0.15
$K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$	$(4.0 \pm 4.0) \times 10^{-6}$	$(1.4 \pm 1.4) \times 10^{-9}$	$(1.4 \pm 1.4) \times 10^{-9}$	$(1.4 \pm 1.4) \times 10^{-9}$
$K^+ \to \mu^+\nu_\mu[e^+e^-\gamma]_{\pi^0}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$
$K^+ \to e^+\nu_e[e^+e^-\gamma]_{\pi^0}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$
$K^+ \to \pi^+\pi^-\mu^+\nu_\mu$	$(2.7 \pm 2.7) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$
$K^+ \to \pi^+\pi^-e^+\nu_e$	0.033 ± 0.019	0.658 ± 0.011	0.045 ± 0.019	0.065 ± 0.021
$K^+ \to \pi^+\mu^+\mu^-$	$(1.66 \pm 0.30) \times 10^{-3}$	0.0137 ± 0.0016	$(5.08 \pm 0.68) \times 10^{-3}$	$(5.33 \pm 0.71) \times 10^{-3}$
$K^+ \to \pi^+e^+e^-$	4.417 ± 0.47	0.61 ± 0.12	1.18 ± 0.21	0.63 ± 0.13
$K^+ \to \mu^+\nu_\mu\mu^+\mu^-$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$
$K^+ \to e^+\nu_e\mu^+\mu^-$	0.0105 ± 0.0063	$(8.8 \pm 4.2) \times 10^{-3})$	0.0111 ± 0.0068	0.0111 ± 0.0067
$K^+ \to \mu^+\nu_\mu e^+e^-$	0.638 ± 0.072	0.168 ± 0.026	0.202 ± 0.030	0.131 ± 0.021
$K^+ \to e^+\nu_e e^+e^-$	$(5.5 \pm 5.5) \times 10^{-7}$	$(1.9 \pm 1.2) \times 10^{-6}$	$(1.6 \pm 1.6) \times 10^{-6}$	$(1.6 \pm 1.6) \times 10^{-6}$
Total background expected	5.50 ± 0.53	1.68 ± 0.20	1.66 ± 0.26	1.06 ± 0.20
p value	0.19	0.50	0.09	–

Table 7.24: Background predictions for $K^+ \to \pi^+\mu^-e^+$ with sidebands unblinded.

Sample	Control Region	Blinded Sidebands		Signal Region
	N_{CRA}	N_{BRB}	N_{BRA}	N_{SR}
Data	4	2	0	BLIND
$K^+ \to \pi^+\pi^+\pi^-$	1.76 ± 0.48	1.60 ± 0.49	1.15 ± 0.40	0.84 ± 0.34
$K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$	0.091 ± 0.057	$(6.2 \pm 6.2) \times 10^{-3}$	0.013 ± 0.010	$(6.2 \pm 6.2) \times 10^{-3}$
$K^+ \to \mu^+\nu_\mu[e^+e^-\gamma]_{\pi^0}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$
$K^+ \to e^+\nu_e[e^+e^-\gamma]_{\pi^0}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$
$K^+ \to \pi^+\pi^-\mu^+\nu_\mu$	$(2.7 \pm 2.7) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$
$K^+ \to \pi^+\pi^-e^+\nu_e$	0.067 ± 0.032	1.77 ± 0.24	0.067 ± 0.035	0.048 ± 0.028
$K^+ \to \pi^+\mu^+\mu^-$	$(7.20 \pm .0.89) \times 10^{-3}$	0.0274 ± 0.0029	0.0147 ± 0.0017	0.0151 ± 0.0017
$K^+ \to \pi^+e^+e^-$	0.0140 ± 0.0017	$(3.4 \pm 3.4) \times 10^{-9}$	$(2.5 \pm 2.0) \times 10^{-7}$	$(3.4 \pm 3.4) \times 10^{-9}$
$K^+ \to \mu^+\nu_\mu\mu^+\mu^-$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$
$K^+ \to e^+\nu_e\mu^+\mu^-$	0.0111 ± 0.0071	0.0117 ± 0.0063	0.0182 ± 0.0070	0.0136 ± 0.0061
$K^+ \to \mu^+\nu_\mu e^+e^-$	$(1.1 \pm 1.1) \times 10^{-8}$	$(1.1 \pm 1.1) \times 10^{-8}$	$(1.1 \pm 1.1) \times 10^{-8}$	$(1.1 \pm 1.1) \times 10^{-8}$
$K^+ \to e^+\nu_e e^+e^-$	$(6.2 \pm 3.2) \times 10^{-5}$	$(1.8 \pm 1.8) \times 10^{-6}$	$(2.8 \pm 2.8) \times 10^{-5}$	$(1.6 \pm 1.6) \times 10^{-10}$
Total background expected	1.95 ± 0.48	3.41 ± 0.54	1.27 ± 0.40	0.92 ± 0.34
p value	0.14	0.85	1.0	–

Table 7.25: Study of p values for quantitative assessment of the agreement between the number of observed events in control regions (CRA, BRB and BRA) and the background expectations under the null (background-only) hypothesis. Combined p values are calculated with Fisher's method.

ROI(s)	$K^+ \to \pi^-\mu^+e^+$ search	$K^+ \to \pi^+\mu^-e^+$ search	Combined
CRA	$n_b = 5.50 \pm 0.53$, $n_{obs} = 8$ $p = 0.19$	$n_b = 1.95 \pm 0.48$, $n_{obs} = 4$ $p = 0.14$	$p = 0.12$
BRB	$n_b = 1.68 \pm 0.21$, $n_{obs} = 2$ $p = 0.50$	$n_b = 3.41 \pm 0.54$, $n_{obs} = 2$ $p = 0.85$	$p = 0.79$
BRA	$n_b = 1.66 \pm 0.26$, $n_{obs} = 4$ $p = 0.09$	$n_b = 1.27 \pm 0.40$, $n_{obs} = 0$ $p = 1.0$	$p = 0.30$
Combined BRB+BRA	$p = 0.18$	$p = 0.99$	$p = 0.58$
Combined CRA+BRB+BRA	$p = 0.13$	$p = 0.63$	$p = 0.31$

7.7 Results and Discussion

7.7.1 Results: Opening Signal Regions For $K^+ \to \pi^\mp\mu^\pm e^+$ Searches

The final invariant mass spectra, with all regions open, are displayed in figure 7.54 for the searches for $K^+ \to \pi^\mp\mu^\pm e^+$ decays. The expected and observed number of events in the signal region are:

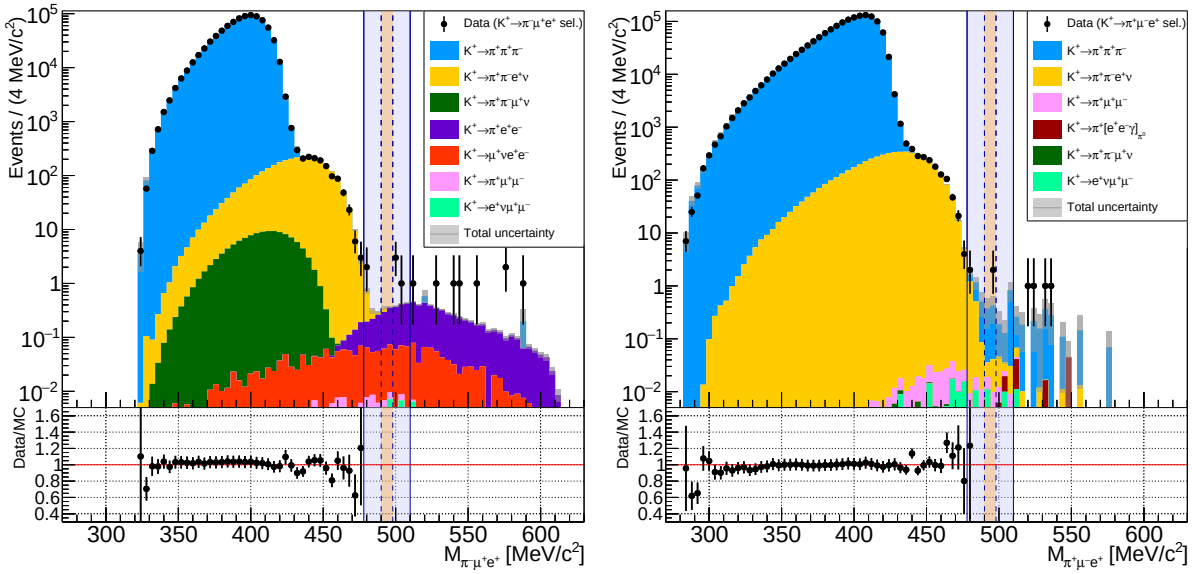
$$K^+ \to \pi^-\mu^+e^+ \text{ search}: \quad n_b = 1.06 \pm 0.20\,, \quad n_{obs} = 0\,, \tag{7.7.1}$$
$$K^+ \to \pi^+\mu^-e^+ \text{ search}: \quad n_b = 0.92 \pm 0.34\,, \quad n_{obs} = 2\,, \tag{7.7.2}$$

which are compatible with the background-only hypothesis (with p values of 1.0 and 0.24 respectively), with no significant excess observed. Therefore, following the procedure described in section 7.6.1.2 and appendix I, upper limits are set on the branching ratios, as shown by figure 7.55, at

$$\mathcal{B}(K^+ \to \pi^-\mu^+e^+) < 4.2 \times 10^{-11} \ @\,90\%\,CL \tag{7.7.3}$$
$$\mathcal{B}(K^+ \to \pi^+\mu^-e^+) < 6.6 \times 10^{-11} \ @\,90\%\,CL. \tag{7.7.4}$$

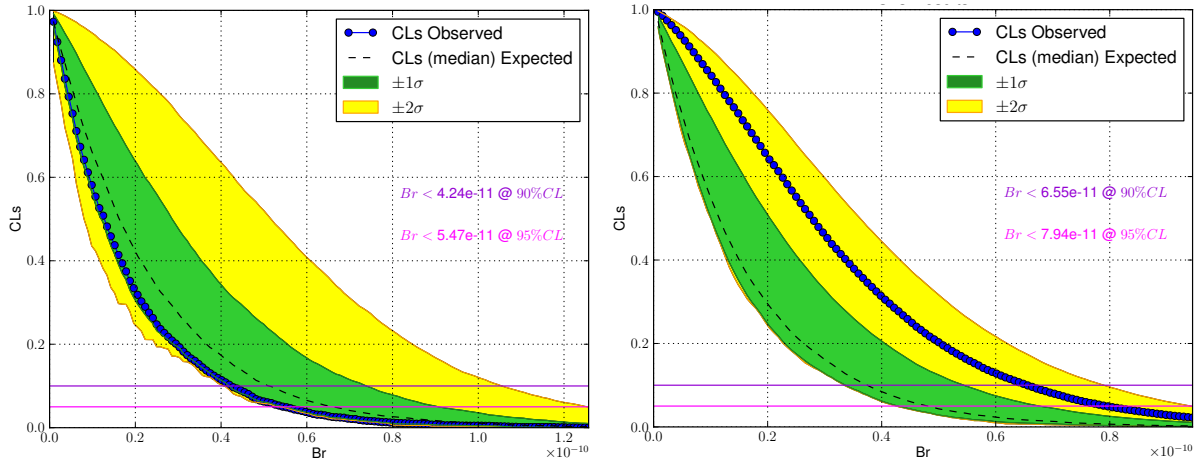
7.7.2 Discussion of $K^+ \to \pi^\mp\mu^\pm e^+$ Results

A summary of the key numerical results from the search for $K^+ \to \pi^\mp\mu^\pm e^+$ decays is given in table 7.26. These results improve over previous upper limits by factors of 12 and 8 for the π^- and μ^- channels respectively.



(a) (b)

Figure 7.54: The full (unblinded) invariant mass spectra of selected events in the 2017+2018 data-set for searches for $K^+ \to \pi^-\mu^+e^+$ (a) and $K^+ \to \pi^+\mu^-e^+$ (b) with comparison to the expected number of events from background sources. The uncertainty of the background prediction is the combined statistical plus systematic uncertainty and is shown as grey boxes overlaid on top of the total background prediction histogram, summing all components. For these final results uncertainty bars displayed for data points in each bin are Poisson 68% confidence intervals, as opposed to the standard case adopted elsewhere where the normal (Wald) approximation is made treating Poissonian counting uncertainties as Gaussian ($\sigma^+ = \sigma_- = \sqrt{N}$).



(a) (b)

Figure 7.55: Expected (black dashed line, with one and two standard deviation bands shown in green and yellow respectively) and observed (blue points and line) CLs as a function of $\mathcal{B}(K^+ \to \pi^-\mu^+e^+)$ (a) and $\mathcal{B}(K^+ \to \pi^+\mu^-e^+)$ (b). Observed upper limits on the branching ratio at 90% and 95% confidence levels are displayed, being the branching ratio values where the observed CLs is equal to 0.1 and 0.05 respectively.

Table 7.26: Numerical summary of $K^+ \to \pi^\mp \mu^\pm e^+$ searches.

	$K^+ \to \pi^- \mu^+ e^+$	$K^+ \to \pi^+ \mu^- e^+$
Normalisation acceptance, A_n	$(10.1772 \pm 0.0023)\%$	
N_K	$(1.321 \pm 0.014) \times 10^{12}$	
Signal acceptance, A_s	$(4.90 \pm 0.02)\%$	$(6.21 \pm 0.02)\%$
Single event sensitivity, $\mathcal{B}_{SES}$	$(1.836 \pm 0.084) \times 10^{-11}$	$(1.449 \pm 0.049) \times 10^{-11}$
Total background expected (SR)	1.06 ± 0.20	0.92 ± 0.34
Number of events for 5σ discovery	11	11
Observed number of events (SR)	0	2
Observed $\mathcal{B}$ upper limit @90% CL	4.2×10^{-11}	6.6×10^{-11}
Previous $\mathcal{B}$ upper limit @90% CL [171]	5.0×10^{-10}	5.2×10^{-10}
(old limit)/(new limit)	12	8

The two events observed in the μ^- channel signal region are identical in this and the independent analysis and their details are provided in table 7.27. Event B in particular is on the upper border of the signal region which, given the resolution on the invariant mass and the background expectation, means the likelihood of it being due to background is favoured over a signal event. No evidence for lepton number or lepton flavour violations is observed, compatible with the SM expectation.

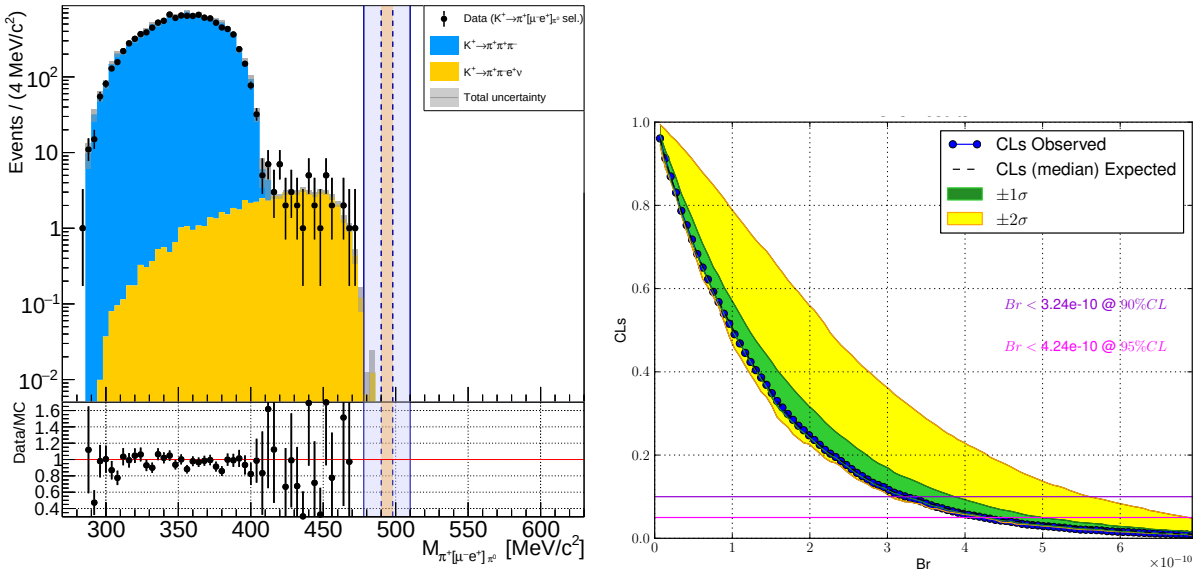
Table 7.27: Information about the two observed events in the signal region in the search for $K^+ \to \pi^+ \mu^- e^+$.

Property	Event A	Event B
Data sample	17B	18H
Run number	8051	9314
Burst number	227	175
Event number	997000	1200503
Trigger(s)	MTe	$MT\mu$
Invariant mass $M_{\pi\mu e}$ [MeV/c^2]	496.67	497.94
Z_{vertex} [m]	141.8	162.7
Track Momenta [GeV/c]	$p_\pi = 14.9$, $p_\mu = 34.3$, $p_e = 25.2$	$p_\pi = 33.7$, $p_\mu = 26.4$, $p_e = 12.2$
Total Momentum [GeV/c]	74.5	74.4
Transverse Momentum [GeV/c]	0.0107	0.0072
E/p	π: 0.85, μ: 0, e: 0.98	π: 0, μ: 0, e: 1.01

7.7.3 Results of $\pi^0 \to \mu^- e^+$ Search

A search for the $\pi^0 \to \mu^- e^+$ decay can be performed as a direct spin-off from the $K^+ \to \pi^+ \mu^- e^+$ search with the additional requirement that $|M_{\mu e} - m_{\pi^0}| < 3\,\mathrm{MeV}/c^2$. The acceptance is lower by approximately a factor of two (see table 7.3) but the background is also lower. The single event sensitivity is competitive with the previous result [173], see tables 7.1 and 7.7. Details of the search are given in appendix K and results are summarised here. The total background expectations and observed number of events in the signal region are:

$$K^+ \to \pi^+ [\mu^- e^+]_{\pi^0} \text{ search}: \quad n_b = 0.24 \pm 0.15\,, \quad n_{obs} = 0\,, \tag{7.7.5}$$



(a) (b)

Figure 7.56: (a) The full (unblinded) invariant mass spectra of selected events in the 2017+2018 data-set for searches for $K^+ \to \pi^+[\mu^- e^+]_{\pi^0}$ with comparison to the expected number of events from background sources. The uncertainty of the background prediction is the combined statistical plus systematic uncertainty and is shown as grey boxes overlaid on top of the total background prediction histogram, summing all components. The uncertainty bars displayed for data points in each bin are Poisson 68% confidence intervals. (b) Expected and observed CLs as a function of $\mathcal{B}(K^+ \to \pi^+[\mu^- e^+]_{\pi^0})$.

with final invariant mass spectra shown in figure 7.56a for the full $\pi^+\mu^- e^+$ final state (and figure K.1b for candidate π^0 decay daughters $\mu^- e^+$ pair). An upper limit is therefore established (see also figure 7.56b) at

$$\mathcal{B}(\pi^0 \to \mu^- e^+) < 3.2 \times 10^{-10} \ @\, 90\%\, CL \tag{7.7.6}$$

This new limit represents a marginal improvement over the strongest previous limit, 3.6×10^{-10} [173], or an improvement by a factor of 11 on the previous charge-specific limit, see appendix K.1.

7.7.4 Future Prospects For $K^+ \to \pi\mu e$ Searches at NA62

These are the final NA62 Run 1 results for $K^+ \to \pi^\mp \mu^\pm e^+$ and $\pi^0 \to \mu^- e^+$ searches. Data-taking resumes in 2021, potentially for three or four years and at higher intensity than before. The focus will remain on detailed measurements of the $K^+ \to \pi^+ \nu\bar{\nu}$ decay so the PNN trigger will occupy the majority of the trigger bandwidth and this may limit the availability of the triggers for future $K^+ \to \pi\mu e$ searches. To significantly improve the single event sensitivity a factor of $\mathcal{O}(5)$ more K^+ decays would need to be collected. However, assuming the analysis procedure is kept effectively fixed the $\mathcal{B}_{SES}/n_b$ ratio will remain the same and so while the searches are currently not limited by background such

an increase in statistics would mean a background expectation $\mathcal{O}(5)$. Further suppression of backgrounds is possible, for example with careful use of the RICH for PID, however this lowers acceptance. In summary, with the current analysis strategy, with potential minor modifications, analysis of future Run 2 data may lead to further improvement in sensitivity by a up to a factor of $\mathcal{O}(5)$.

With a factor of $\mathcal{O}(5)$ improvement in sensitivity the $\mathcal{B}_{SES}$ reaches $\sim 4 \times 10^{-12}$ and NA62 may become competitive with the previous limit of $\mathcal{B}(K^+ \to \pi^+\mu^+e^-) < 1.3 \times 10^{-11}$ [172]. However, to reach a significant improvement would require moderate development of the current analysis strategy.

Much better prospects exist for the searches for $\pi^0 \to \mu^\pm e^\mp$ decays. In these channels the expected background is small ($\lesssim 0.3$ events) and current predictions are limited by simulation and control sample statistics. The single event sensitivity is already comparable to the previous searches [173], see tables 7.1 and 7.7. Therefore, an increase even by a factor of 2 in statistics, coupled with some optimisation of the analysis for these channels to gain acceptance, may provide the possibility for good improvement over the previous limits. Moreover the analysis could be expanded to profit not only from π^0 produced in $K^+ \to \pi^+\pi^0$ decays but also in the $K^+ \to \pi^0 e^+ \nu_e$ and $K^+ \to \pi^0 \mu^+ \nu_\mu$ decays with branching ratios of $(5.07 \pm 0.04)\%$ and $(3.352 \pm 0.033)\%$ respectively [13]. Together these potentially increase the number of π^0 that can be studied by about 40% in relative terms.

Chapter 8

Conclusions

The goal of modern particle physics is to test the SM to its limits, identify new phenomena which the SM cannot explain and investigate new models to find an underlying theory. The work presented in this thesis contributes to this effort through tests of a precision SM prediction in studies of the ultra-rare $K^+ \to \pi^+\nu\bar{\nu}$ decay, searches for new exotic particles, X, not contained in the SM framework in $K^+ \to \pi^+ X$ decays, and searches for decays $K^+ \to \pi\mu e$ which violate emergent conservation laws of the SM. No evidence of new BSM physics has been identified but parameter spaces of potential models have been ruled out and these results feed into the growing global picture.

The NA62 experiment was specifically designed to measure the branching ratio of the ultra-rare $K^+ \to \pi^+\nu\bar{\nu}$ decay, which is precisely predicted in the SM and may be enhanced or suppressed significantly in a variety of BSM scenarios. The analysis techniques have developed with each year of data being investigated in independent analyses. Particular techniques for background evaluation for analysis of 2016 and 2017 data have been presented, specifically for the $K^+ \to \pi^+\pi^+\pi^-$, upstream and *a posteriori* cross-checks for $K^+ \to \pi^+\pi^0$ and $K^+ \to \mu^+\nu_\mu$ backgrounds. Further developments have led to the Run 1 result with measurement of the branching ratio

$$\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu}) = (11.0^{+4.0}_{-3.5}|_{stat} \pm 0.3_{syst}) \times 10^{-11}\,, \tag{8.0.1}$$

compatible with the SM prediction (see equation 2.3.11) within 1σ and representing evidence for the $K^+ \to \pi^+\nu\bar{\nu}$ decay, rejecting the background-only hypothesis, with significance of 3.5σ.

Alongside the headline measurement of $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$, and within the same analysis framework, a search was performed for $K^+ \to \pi^+ X$ where X is a new feebly interacting, potentially long-lived, particle. New model-independent upper limits on the branching ratio are placed which are competitive at low masses (0–100 MeV/c^2) and improve by a factor of $\mathcal{O}(10)$ at higher masses (155–260 MeV/c^2). Specific BSM scenarios are constrained with new exclusion bounds added to the relevant parameter spaces.

Searches for lepton number and flavour violation in $K^+ \to \pi\mu e$ decays have been

presented with new world-leading upper limits placed on three modes:

$$\mathcal{B}(K^+ \to \pi^- \mu^+ e^+) < 4.2 \times 10^{-11} \ @\, 90\%\, CL, \tag{8.0.2}$$

$$\mathcal{B}(K^+ \to \pi^+ \mu^- e^+) < 6.6 \times 10^{-11} \ @\, 90\%\, CL, \tag{8.0.3}$$

$$\mathcal{B}(\pi^0 \to \mu^- e^+) < 3.2 \times 10^{-10} \ @\, 90\%\, CL\,. \tag{8.0.4}$$

These fit into a broader set of searches for lepton number and flavour violating processes in K^+ decays at NA62 including recent results on $K^+ \to \pi^- \ell^+ \ell^+$ ($\ell = \mu$ or e)) decays [34]. These analyses are not limited by background and further improvements are expected with additional data-taking beginning in 2021.

Bibliography

[1] R. Vople for the NA62 Collaboration. *Latest measurement of $K^+ \to \pi^+\nu\bar{\nu}$ with the NA62 experiment at CERN*. The Phenomenology 2020 Symposium (PHENO20), University of Pittsburgh. May 2020. URL: https://indico.cern.ch/event/858682/contributions/3837170/.

[2] Joel Swallow. *Search for light feebly-interacting scalar particles at extracted beam lines (SeaQuest FNAL, NA62, SHiP,...)* The Workshop on Feebly Interacting Particles (FIPs20), 31$^{\text{th}}$ August – 4$^{\text{th}}$ September. Sept. 2020. URL: https://indico.cern.ch/event/864648/contributions/3896867/.

[3] Eduardo Cortina Gil et al. "Search for a feebly interacting particle X in the decay $K^+ \to \pi^+ X$". In: (Nov. 2020). Accepted for publication in JHEP. arXiv: 2011.11329 [hep-ex].

[4] J. Swallow for the NA62 Collaboration. "Searches for lepton flavour and lepton number violation in K^+ decays". In: *ICHEP20*. Prague, July 2020. URL: https://indico.cern.ch/event/868940/contributions/3815640/attachments/2081977/3497128/ICHEP20_JSwallow_NA62_LNVLFV_vF.pdf.

[5] NA62. *Search for lepton number and flavour violation in K^+ and π^0 decays*. To be submitted for publication in JHEP. Feb. 2021.

[6] Sheldon L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (1961), pp. 579–588. ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(61)90469-2. URL: http://www.sciencedirect.com/science/article/pii/0029558261904692.

[7] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[8] Abdus Salam. "Weak and Electromagnetic Interactions". In: *Conf. Proc. C* 680519 (1968), pp. 367–377. DOI: 10.1142/9789812795915_0034.

[9] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413. URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413.

[10] Nicola Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531. URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[11] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: *Phys. Rev. D* 2 (7 Oct. 1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285. URL: https://link.aps.org/doi/10.1103/PhysRevD.2.1285.

[12] Makoto Kobayashi and Toshihide Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652. eprint: https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf. URL: https://doi.org/10.1143/PTP.49.652.

[13] M. Tanabashi et al. "Review of Particle Physics (Particle Data Group)". In: *Phys. Rev. D* 98.3 (2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.

[14] C. Jarlskog. "Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Nonconservation". In: *Phys. Rev. Lett.* 55 (10 Sept. 1985), pp. 1039–1042. DOI: 10.1103/PhysRevLett.55.1039. URL: https://link.aps.org/doi/10.1103/PhysRevLett.55.1039.

[15] Ivan Esteban et al. "Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of θ_{23}, δ_{CP}, and the mass ordering". In: *JHEP* 01 (2019), p. 106. DOI: 10.1007/JHEP01(2019)106. arXiv: 1811.05487 [hep-ph].

[16] B. Pontecorvo. "Mesonium and anti-mesonium". In: *Sov. Phys. JETP* 6 (1957), p. 429. URL: http://www.jetp.ac.ru/cgi-bin/e/index/e/6/2/p429?a=list.

[17] B. Pontecorvo. "Inverse beta processes and nonconservation of lepton charge". In: *Sov. Phys. JETP* 7 (1958), pp. 172–173. URL: http://www.jetp.ac.ru/cgi-bin/dn/e_007_01_0172.pdf.

[18] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. "Remarks on the Unified Model of Elementary Particles". In: *Progress of Theoretical Physics* 28.5 (Nov. 1962), pp. 870–880. ISSN: 0033-068X. DOI: 10.1143/PTP.28.870. eprint: https://academic.oup.com/ptp/article-pdf/28/5/870/5258750/28-5-870.pdf. URL: https://doi.org/10.1143/PTP.28.870.

[19] E. Noether. "Invariante Variationsprobleme". ger. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1918 (1918), pp. 235–257. URL: http://eudml.org/doc/59024.

[20] Andrei D Sakharov. "Violation ofCPin variance,Casymmetry, and baryon asymmetry of the universe". In: *Soviet Physics Uspekhi* 34.5 (May 1991). re-published, origional JETP Letters, 5, 1, p. 24 (1967), Erratum: JETP Letters 5(4), p. 109, pp. 392–393. DOI: 10.1070/pu1991v034n05abeh002497.

[21] Werner Bernreuther. "CP violation and baryogenesis". In: *Lect. Notes Phys.* 591 (2002), pp. 237–293. arXiv: hep-ph/0205279.

[22] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[23] Q. R. Ahmad et al. "Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (7 July 2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.071301.

[24] Q. R. Ahmad et al. "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 89 (1 June 2002), p. 011301. DOI: 10.1103/PhysRevLett.89.011301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.89.011301.

[25] Peter Minkowski. "$\mu \to e\gamma$ at a rate of one out of 109 muon decays?" In: *Physics Letters B* 67.4 (1977), pp. 421–428. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(77)90435-X. URL: http://www.sciencedirect.com/science/article/pii/037026937790435X.

[26] Murray Gell-Mann, Pierre Ramond, and Richard Slansky. "Complex Spinors and Unified Theories". In: *Conf. Proc. C* 790927 (1979), pp. 315–321. arXiv: 1306.4669 [hep-th].

[27] Rabindra N. Mohapatra and Goran Senjanović. "Neutrino Mass and Spontaneous Parity Nonconservation". In: *Phys. Rev. Lett.* 44 (14 Apr. 1980), pp. 912–915. DOI: 10.1103/PhysRevLett.44.912. URL: https://link.aps.org/doi/10.1103/PhysRevLett.44.912.

[28] T. P. Cheng and Ling-Fong Li. "$\mu \to e\gamma$ in Theories with Dirac and Majorana Neutrino-Mass Terms". In: *Phys. Rev. Lett.* 45 (24 Dec. 1980), pp. 1908–1911. DOI: 10.1103/PhysRevLett.45.1908. URL: https://link.aps.org/doi/10.1103/PhysRevLett.45.1908.

[29] J. Hisano and Daisuke Nomura. "Solar and atmospheric neutrino oscillations and lepton flavor violation in supersymmetric models with right-handed neutrinos". In: *Phys. Rev. D* 59 (11 Apr. 1999), p. 116005. DOI: 10.1103/PhysRevD.59.116005. URL: https://link.aps.org/doi/10.1103/PhysRevD.59.116005.

[30] Toshinori Mori. "Final Results of the MEG Experiment". In: *Nuovo Cim. C* 39.4 (2017), p. 325. DOI: 10.1393/ncc/i2016-16325-7. arXiv: 1606.08168 [hep-ex].

[31] Anupama Atre et al. "The Search for Heavy Majorana Neutrinos". In: *JHEP* 05 (2009), p. 030. DOI: 10.1088/1126-6708/2009/05/030. arXiv: 0901.3589 [hep-ph].

[32] Karim Massri. "Kaon identification and search for lepton number violation in $K^{\pm}$ decay-in-flight experiments at CERN". Presented on 2014-08-11. 2014. URL: https://cds.cern.ch/record/2136901.

[33] J.R. Batley et al. "Searches for lepton number violation and resonances in $K^{\pm} \to \pi\mu\mu$ decays". In: *Phys. Lett. B* 769 (2017), pp. 67–76. DOI: 10.1016/j.physletb.2017.03.029. arXiv: 1612.04723 [hep-ex].

[34] Eduardo Cortina Gil et al. "Searches for lepton number violating K^+ decays". In: *Phys. Lett. B* 797 (2019), p. 134794. DOI: 10.1016/j.physletb.2019.07.041. arXiv: 1905.07770 [hep-ex].

[35] M. Agostini et al. "Improved Limit on Neutrinoless Double-β Decay of ^{76}Ge from GERDA Phase II". In: *Phys. Rev. Lett.* 120 (13 Mar. 2018), p. 132503. DOI: 10.1103/PhysRevLett.120.132503. URL: https://link.aps.org/doi/10.1103/PhysRevLett.120.132503.

[36] J. Kotila and F. Iachello. "Phase-space factors for double-β decay". In: *Phys. Rev. C* 85 (3 Mar. 2012), p. 034316. DOI: 10.1103/PhysRevC.85.034316. URL: https://link.aps.org/doi/10.1103/PhysRevC.85.034316.

[37] D. Q. Adams et al. "Improved Limit on Neutrinoless Double-Beta Decay in ^{130}Te with CUORE". In: *Phys. Rev. Lett.* 124 (12 Mar. 2020), p. 122501. DOI: 10.1103/PhysRevLett.124.122501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.124.122501.

[38] Matteo Agostini, Giovanni Benato, and Jason A. Detwiler. "Discovery probability of next-generation neutrinoless double-β decay experiments". In: *Phys. Rev. D* 96 (5 Sept. 2017), p. 053001. DOI: 10.1103/PhysRevD.96.053001. URL: https://link.aps.org/doi/10.1103/PhysRevD.96.053001.

[39] A. Giuliani et al. "Double Beta Decay APPEC Committee Report". In: (Oct. 2019). arXiv: 1910.04688 [hep-ex].

[40] Chien-Yeah Seng. "Reexamination of the standard model nucleon electric dipole moment". In: *Phys. Rev. C* 91 (2 Feb. 2015), p. 025502. DOI: 10.1103/PhysRevC.91.025502. URL: https://link.aps.org/doi/10.1103/PhysRevC.91.025502.

[41] E.P. Shabalin. "THE ELECTRIC DIPOLE MOMENT OF THE NEUTRON IN A GAUGE THEORY". In: *Soviet Physics Uspekhi* 26 (1983), p. 297. DOI: 10.1070/PU1983v026n04ABEH004331.

[42] Maxim Pospelov and Adam Ritz. "CKM benchmarks for electron electric dipole moment experiments". In: *Phys. Rev. D* 89.5 (2014), p. 056006. DOI: 10.1103/PhysRevD.89.056006. arXiv: 1311.5537 [hep-ph].

[43] J. M. Pendlebury et al. "Revised experimental upper limit on the electric dipole moment of the neutron". In: *Phys. Rev. D* 92.9 (2015), p. 092003. DOI: 10.1103/PhysRevD.92.092003. arXiv: 1509.04411 [hep-ex].

[44] R. D. Peccei and Helen R. Quinn. "CP Conservation in the Presence of Pseudoparticles". In: *Phys. Rev. Lett.* 38 (25 June 1977), pp. 1440–1443. DOI: 10.1103/PhysRevLett.38.1440. URL: https://link.aps.org/doi/10.1103/PhysRevLett.38.1440.

[45] R. D. Peccei and Helen R. Quinn. "Constraints imposed by CP conservation in the presence of pseudoparticles". In: *Phys. Rev. D* 16 (6 Sept. 1977), pp. 1791–1797. DOI: 10.1103/PhysRevD.16.1791. URL: https://link.aps.org/doi/10.1103/PhysRevD.16.1791.

[46] Steven Weinberg. "A New Light Boson?" In: *Phys. Rev. Lett.* 40 (4 Jan. 1978), pp. 223–226. DOI: 10.1103/PhysRevLett.40.223. URL: https://link.aps.org/doi/10.1103/PhysRevLett.40.223.

[47] F. Wilczek. "Problem of Strong P and T Invariance in the Presence of Instantons". In: *Phys. Rev. Lett.* 40 (5 Jan. 1978), pp. 279–282. DOI: 10.1103/PhysRevLett.40.279. URL: https://link.aps.org/doi/10.1103/PhysRevLett.40.279.

[48] Marco Gorghetto and Giovanni Villadoro. "Topological Susceptibility and QCD Axion Mass: QED and NNLO corrections". In: *JHEP* 03 (2019), p. 033. DOI: 10.1007/JHEP03(2019)033. arXiv: 1812.01008 [hep-ph].

[49] Jihn E. Kim. "Weak-Interaction Singlet and Strong CP Invariance". In: *Phys. Rev. Lett.* 43 (2 July 1979), pp. 103–107. DOI: 10.1103/PhysRevLett.43.103. URL: https://link.aps.org/doi/10.1103/PhysRevLett.43.103.

[50] M.A. Shifman, A.I. Vainshtein, and V.I. Zakharov. "Can confinement ensure natural CP invariance of strong interactions?" In: *Nuclear Physics B* 166.3 (1980), pp. 493–506. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(80)90209-6. URL: http://www.sciencedirect.com/science/article/pii/0550321380902096.

[51] Michael Dine, Willy Fischler, and Mark Srednicki. "A simple solution to the strong CP problem with a harmless axion". In: *Physics Letters B* 104.3 (1981), pp. 199–202. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(81)90590-6. URL: http://www.sciencedirect.com/science/article/pii/0370269381905906.

[52] A.R. Zhitnitsky. "On Possible Suppression of the Axion Hadron Interactions. (In Russian)". In: *Sov. J. Nucl. Phys.* 31 (1980), p. 260.

[53] Joerg Jaeckel and Andreas Ringwald. "The Low-Energy Frontier of Particle Physics". In: *Ann. Rev. Nucl. Part. Sci.* 60 (2010), pp. 405–437. DOI: 10.1146/annurev.nucl.012809.104433. arXiv: 1002.0329 [hep-ph].

[54] J. Beacham et al. "Physics Beyond Colliders at CERN: Beyond the Standard Model Working Group Report". In: *J. Phys.* G47.1 (2020), p. 010501. DOI: 10.1088/1361-6471/ab4cd2. arXiv: 1901.09966 [hep-ex].

[55] P.A.R. Ade et al. "Planck 2015 results. XIII. Cosmological parameters". In: *Astron. Astrophys.* 594 (2016), A13. DOI: 10.1051/0004-6361/201525830. arXiv: 1502.01589 [astro-ph.CO].

[56] Zackaria Chacko et al. "Cosmological Limits on the Neutrino Mass and Lifetime". In: *JHEP* 04 (2020), p. 020. DOI: 10.1007/JHEP04(2020)020. arXiv: 1909.05275 [hep-ph].

[57] M. Spano et al. "GHASP: an Hα kinematic survey of spiral and irregular galaxies - V. Dark matter distribution in 36 nearby spiral galaxies". In: *Monthly Notices of the Royal Astronomical Society* 383.1 (Jan. 2008), pp. 297–316. DOI: 10.1111/j.1365-2966.2007.12545.x. arXiv: 0710.1345 [astro-ph].

[58] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. "A Universal Density Profile from Hierarchical Clustering". In: *The Astrophysical Journal* 490.2 (Dec. 1997), pp. 493–508. DOI: 10.1086/304888. URL: https://doi.org/10.1086%2F304888.

[59] Douglas Clowe et al. "A direct empirical proof of the existence of dark matter". In: *Astrophys. J. Lett.* 648 (2006), pp. L109–L113. DOI: 10.1086/508162. arXiv: astro-ph/0608407.

[60] M. Markevitch et al. "A Textbook example of a bow shock in the merging galaxy cluster 1E0657-56". In: *Astrophys. J.* 567 (2002), p. L27. DOI: 10.1086/339619. arXiv: astro-ph/0110468.

[61] Maira Dutra. "Origins for dark matter particles : from the "WIMP miracle" to the "FIMP wonder"". Theses. Université Paris-Saclay, Feb. 2019. URL: https://tel.archives-ouvertes.fr/tel-02100637.

[62] Zhengkang Zhang. "Effective Field Theory Approaches to Particle Physics Beyond the Standard Model". PhD thesis. Michigan U., 2018.

[63] Dario Buttazzo et al. "B-physics anomalies: a guide to combined explanations". In: *JHEP* 11 (2017), p. 044. DOI: 10.1007/JHEP11(2017)044. arXiv: 1706.07808 [hep-ph].

[64] Sokratis Trifinopoulos. "B -physics anomalies: The bridge between R -parity violating supersymmetry and flavored dark matter". In: *Phys. Rev. D* 100.11 (2019), p. 115022. DOI: 10.1103/PhysRevD.100.115022. arXiv: 1904.12940 [hep-ph].

[65] Elena Graverini. "Flavour anomalies: a review". In: *J. Phys. Conf. Ser.* 1137.1 (2019). Ed. by Fernando Barao et al., p. 012025. DOI: 10.1088/1742-6596/1137/1/012025. arXiv: 1807.11373 [hep-ex].

[66] Simone Bifani et al. "Review of Lepton Universality tests in B decays". In: *J. Phys. G* 46.2 (2019), p. 023001. DOI: 10.1088/1361-6471/aaf5de. arXiv: 1809.06229 [hep-ex].

[67] Marzia Bordone, Gino Isidori, and Andrea Pattori. "On the Standard Model predictions for R_K and R_{K^*}". In: *Eur. Phys. J. C* 76.8 (2016), p. 440. DOI: 10.1140/epjc/s10052-016-4274-7. arXiv: 1605.07633 [hep-ph].

[68] R. Aaij et al. "Search for Lepton-Universality Violation in $B^+ \to K^+\ell^+\ell^-$ Decays". In: *Phys. Rev. Lett.* 122 (19 May 2019), p. 191801. DOI: 10.1103/PhysRevLett.122.191801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.191801.

[69] A. Abdesselam et al. "Test of lepton flavor universality in $B \to K^*\ell^+\ell^-$ decays at Belle". In: (Apr. 2019). arXiv: 1904.02440 [hep-ex].

[70] A.T. Fienberg. "The Status and Prospects of the Muon $g-2$ Experiment at Fermilab". In: *54th Rencontres de Moriond on QCD and High Energy Interactions*. May 2019, pp. 163–166. arXiv: 1905.05318 [hep-ex].

[71] T. Aoyama et al. "The anomalous magnetic moment of the muon in the Standard Model". In: (June 2020). arXiv: 2006.04822 [hep-ph].

[72] Diktys Stratakis et al. "Accelerator performance analysis of the Fermilab Muon Campus". In: *Phys. Rev. Accel. Beams* 20 (11 Nov. 2017), p. 111003. DOI: 10.1103/PhysRevAccelBeams.20.111003. URL: https://link.aps.org/doi/10.1103/PhysRevAccelBeams.20.111003.

[73] J. Grange et al. "Muon (g-2) Technical Design Report". In: (Jan. 2015). arXiv: 1501.06858 [physics.ins-det].

[74] M. Davier et al. "A new evaluation of the hadronic vacuum polarisation contributions to the muon anomalous magnetic moment and to $\alpha(m_Z^2)$". In: *Eur. Phys. J. C* 80.3 (2020), p. 241. DOI: 10.1140/epjc/s10052-020-7792-2. arXiv: 1908.00921 [hep-ph].

[75] Alexander Keshavarzi, Daisuke Nomura, and Thomas Teubner. "Muon $g-2$ and $\alpha(M_Z^2)$: A new data-based analysis". In: *Phys. Rev. D* 97 (11 June 2018), p. 114025. DOI: 10.1103/PhysRevD.97.114025. URL: https://link.aps.org/doi/10.1103/PhysRevD.97.114025.

[76] Fred Jegerlehner. "Muon g - 2 theory: The hadronic part". In: *European Physical Journal Web of Conferences*. Vol. 166. European Physical Journal Web of Conferences. Jan. 2018, p. 00022. DOI: 10.1051/epjconf/201816600022. arXiv: 1705.00263 [hep-ph].

[77] J.R. Batley et al. "Search for the dark photon in π^0 decays". In: *Phys. Lett. B* 746 (2015), pp. 178–185. DOI: 10.1016/j.physletb.2015.04.068. arXiv: 1504.00607 [hep-ex].

[78] Eduardo Cortina Gil et al. "Search for production of an invisible dark photon in π^0 decays". In: *JHEP* 05.arXiv:1903.08767 (Mar. 2019), p. 182. 17. DOI: 10.1007/JHEP05(2019)182. URL: https://cds.cern.ch/record/2655539.

[79] S.N. Gninenko, N.V. Krasnikov, and V.A. Matveev. "Muon g-2 and searches for a new leptophobic sub-GeV dark boson in a missing-energy experiment at CERN". In: *Phys. Rev. D* 91 (2015), p. 095015. DOI: 10.1103/PhysRevD.91.095015. arXiv: 1412.1400 [hep-ph].

[80] S.N. Gninenko and N.V. Krasnikov. "Probing the muon $g_\mu - 2$ anomaly, $L_\mu - L_\tau$ gauge boson and Dark Matter in dark photon experiments". In: *Physics Letters B* 783 (2018), pp. 24–28. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2018.06.043. URL: http://www.sciencedirect.com/science/article/pii/S0370269318304994.

[81] Andrzej J. Buras et al. "$K^+ \to \pi^+\nu\overline{\nu}$ and $K_L \to \pi^0\nu\overline{\nu}$ in the Standard Model: status and perspectives". In: *JHEP* 11 (2015), p. 033. DOI: 10.1007/JHEP11(2015)033. arXiv: 1503.02693 [hep-ph].

[82] Andrzej J. Buras et al. "Improved anatomy of ϵ'/ϵ in the Standard Model". In: *JHEP* 11 (2015), p. 202. DOI: 10.1007/JHEP11(2015)202. arXiv: 1507.06345 [hep-ph].

[83] Federico Mescia and Christopher Smith. "Improved estimates of rare K decay matrix-elements from Kl3 decays". In: *Phys. Rev. D* 76 (2007), p. 034017. DOI: 10.1103/PhysRevD.76.034017. arXiv: 0705.2025 [hep-ph].

[84] Andrzej J. Buras. "Weak Hamiltonian, CP violation and rare decays". In: *Les Houches Summer School in Theoretical Physics, Session 68: Probing the Standard Model of Particle Interactions*. June 1998, pp. 281–539. arXiv: hep-ph/9806471.

[85] Gino Isidori, Federico Mescia, and Christopher Smith. "Light-quark loops in $K \to \pi\nu\bar{\nu}$". In: *Nucl. Phys. B* 718 (2005), pp. 319–338. DOI: 10.1016/j.nuclphysb.2005.04.008. arXiv: hep-ph/0503107.

[86] Adam F. Falk, Adam Lewandowski, and Alexey A. Petrov. "Effects from the charm scale in $K^+ \to \pi^+\nu\bar{\nu}$". In: *Physics Letters B* 505.1 (2001), pp. 107–112. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(01)00343-4. URL: http://www.sciencedirect.com/science/article/pii/S0370269301003434.

[87] Andrzej J. Buras and Robert Fleischer. "Quark mixing, CP violation and rare decays after the top quark discovery". In: *Heavy flavours II*. Ed. by A.J. Buras and M. Lindner. Vol. 15. 1997, pp. 65–238. DOI: 10.1142/9789812812667_0002. arXiv: hep-ph/9704376.

[88] 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". In: *Phys. Rev. Lett.* 122 (2 Jan. 2019), p. 021802. DOI: 10.1103/PhysRevLett.122.021802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.021802.

[89] R. Marchevski for the NA62 Collaboration. "Evidence for the decay $K^+ \to \pi^+\nu\bar{\nu}$ from the NA62 experiment at CERN". In: *ICHEP20*. Prague, July 2020. URL: https://indico.cern.ch/event/868940/contributions/3815641/attachments/2080353/3496097/RadoslavMarchevski_ICHEP_2020.pdf.

[90] A. V. Artamonov et al. "New Measurement of the $K^+ \to \pi^+\nu\bar{\nu}$ Branching Ratio". In: *Phys. Rev. Lett.* 101 (19 Nov. 2008), p. 191802. DOI: 10.1103/PhysRevLett.101.191802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.101.191802.

[91] Yuval Grossman and Yosef Nir. "$K_{(L)} \to \pi^0\nu\bar{\nu}$ beyond the standard model". In: *Phys. Lett. B* 398 (1997), pp. 163–168. DOI: 10.1016/S0370-2693(97)00210-4. arXiv: hep-ph/9701313.

[92] Andrzej J. Buras, Dario Buttazzo, and Robert Knegjens. "$K \to \pi\nu\bar{\nu}$ and ϵ'/ϵ in simplified new physics models". In: *JHEP* 11 (2015), p. 166. DOI: 10.1007/JHEP11(2015)166. arXiv: 1507.08672 [hep-ph].

[93] Riccardo Barbieri et al. "Flavour physics and flavour symmetries after the first LHC phase". In: *JHEP* 05 (2014), p. 105. DOI: 10.1007/JHEP05(2014)105. arXiv: 1402.6677 [hep-ph].

[94] Monika Blanke et al. "FCNC Processes in the Littlest Higgs Model with T-Parity: a 2009 Look". In: *Acta Phys. Polon. B* 41 (2010), pp. 657–683. arXiv: 0906.5454 [hep-ph].

[95] Andrzej J. Buras, Fulvia De Fazio, and Jennifer Girrbach. "The Anatomy of Z' and Z with Flavour Changing Neutral Currents in the Flavour Precision Era". In: *JHEP* 02 (2013), p. 116. DOI: 10.1007/JHEP02(2013)116. arXiv: 1211.1896 [hep-ph].

[96] Andrzej J. Buras et al. "Can we reach the Zeptouniverse with rare K and Bs, d decays?" In: *JHEP* 11 (2014), p. 121. DOI: 10.1007/JHEP11(2014)121. arXiv: 1408.0728 [hep-ph].

[97] Monika Blanke et al. "Rare K and B Decays in a Warped Extra Dimension with Custodial Protection". In: *JHEP* 03 (2009), p. 108. DOI: 10.1088/1126-6708/2009/03/108. arXiv: 0812.3803 [hep-ph].

[98] Monika Blanke et al. "Δ F=2 Observables and Fine-Tuning in a Warped Extra Dimension with Custodial Protection". In: *JHEP* 03 (2009), p. 001. DOI: 10.1088/1126-6708/2009/03/001. arXiv: 0809.1073 [hep-ph].

[99] Marzia Bordone et al. "Probing Lepton Flavour Universality with $K \to \pi\nu\bar{\nu}$ decays". In: *Eur. Phys. J. C* 77.9 (2017), p. 618. DOI: 10.1140/epjc/s10052-017-5202-1. arXiv: 1705.10729 [hep-ph].

[100] Andrzej J. Buras et al. "$B \to K^{(*)}\nu\bar{\nu}$ decays in the Standard Model and beyond". In: *JHEP* 02 (2015), p. 184. DOI: 10.1007/JHEP02(2015)184. arXiv: 1409.4557 [hep-ph].

[101] U. Camerini et al. "Experimental Search for Semileptonic Neutrino Neutral Currents". In: *Phys. Rev. Lett.* 23 (6 Aug. 1969), pp. 326–329. DOI: 10.1103/PhysRevLett.23.326. URL: https://link.aps.org/doi/10.1103/PhysRevLett.23.326.

[102] D. Ljung and D. Cline. "Experimental Study of the Rare K^+ Decay Modes: $K^+ \to \pi^+\pi^0\gamma$, $K^+ \to \mu^+\pi^0\nu\gamma$, $K^+ \to \pi^+\gamma\gamma$, $K^+ \to \pi^+\nu\overline{\nu}$, $K^+ \to \pi^0\pi^0e^+\nu$, and $K^+ \to e^+\pi^0\nu\gamma$". In: *Phys. Rev. D* 8 (5 Sept. 1973), pp. 1307–1330. DOI: 10.1103/PhysRevD.8.1307. URL: https://link.aps.org/doi/10.1103/PhysRevD.8.1307.

[103] J. H. Klems, R. H. Hildebrand, and R. Stiening. "Limits on the $K^+ \to \pi^+ + \nu + \overline{\nu}$ and $K^+ \to \pi^+ + n\gamma$ Decay Rates". In: *Phys. Rev. D* 4 (1 July 1971), pp. 66–80. DOI: 10.1103/PhysRevD.4.66. URL: https://link.aps.org/doi/10.1103/PhysRevD.4.66.

[104] G. D. Cable et al. "Search for Rare K^+ Decays. II. $K^+ \to \pi^+\nu\overline{\nu}$". In: *Phys. Rev. D* 8 (11 Dec. 1973), pp. 3807–3812. DOI: 10.1103/PhysRevD.8.3807. URL: https://link.aps.org/doi/10.1103/PhysRevD.8.3807.

[105] Takehiko Asaka and Mikhail Shaposhnikov. "The νMSM, dark matter and baryon asymmetry of the universe". In: *Physics Letters B* 620.1 (2005), pp. 17–26. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2005.06.020. URL: http://www.sciencedirect.com/science/article/pii/S0370269305008087.

[106] M. S. Atiya et al. "Search for the decay $K^+ \to \pi^+\nu\overline{\nu}$". In: *Phys. Rev. Lett.* 64 (1 Jan. 1990), pp. 21–24. DOI: 10.1103/PhysRevLett.64.21. URL: https://link.aps.org/doi/10.1103/PhysRevLett.64.21.

[107] M. S. Atiya et al. "Search for the Decay $K^+ \to \pi^+\nu\overline{\nu}$". In: *Phys. Rev. Lett.* 70 (17 Apr. 1993), pp. 2521–2524. DOI: 10.1103/PhysRevLett.70.2521. URL: https://link.aps.org/doi/10.1103/PhysRevLett.70.2521.

[108] S. Adler et al. "Search for the Decay $K^+ \to \pi^+\nu\overline{\nu}$". In: *Phys. Rev. Lett.* 76 (9 Feb. 1996), pp. 1421–1424. DOI: 10.1103/PhysRevLett.76.1421. URL: https://link.aps.org/doi/10.1103/PhysRevLett.76.1421.

[109] S. Adler et al. "Evidence for the Decay $K^+ \to \pi^+\nu\overline{\nu}$". In: *Phys. Rev. Lett.* 79 (12 Sept. 1997), pp. 2204–2207. DOI: 10.1103/PhysRevLett.79.2204. URL: https://link.aps.org/doi/10.1103/PhysRevLett.79.2204.

[110] S. Adler et al. "Further Search for the Decay $K^+ \to \pi^+\nu\overline{\nu}$". In: *Phys. Rev. Lett.* 84 (17 Apr. 2000), pp. 3768–3770. DOI: 10.1103/PhysRevLett.84.3768. URL: https://link.aps.org/doi/10.1103/PhysRevLett.84.3768.

[111] S. Adler et al. "Further Evidence for the Decay $K^+ \to \pi^+\nu\overline{\nu}$". In: *Phys. Rev. Lett.* 88 (4 Jan. 2002), p. 041803. DOI: 10.1103/PhysRevLett.88.041803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.88.041803.

[112] V. V. Anisimovsky et al. "Improved Measurement of the $K^+ \to \pi^+\nu\overline{\nu}$ Branching Ratio". In: *Phys. Rev. Lett.* 93 (3 July 2004), p. 031801. DOI: 10.1103/PhysRevLett.93.031801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.93.031801.

[113] A. V. Artamonov et al. "Study of the decay $K^+ \to \pi^+\nu\overline{\nu}$ in the momentum region $140 < P_\pi < 199\,\mathrm{MeV}/c$". In: *Phys. Rev. D* 79 (9 May 2009), p. 092004. DOI: 10.1103/PhysRevD.79.092004. URL: https://link.aps.org/doi/10.1103/PhysRevD.79.092004.

[114] O. I. Dahl et al. "Strange-Particle Production in $\pi^- p$ Interactions from 1.5 to 4.2 BeV/c. II. Two-Body Final States". In: *Phys. Rev.* 183 (5 July 1969), pp. 1520–1520. DOI: 10.1103/PhysRev.183.1520.4. URL: https://link.aps.org/doi/10.1103/PhysRev.183.1520.4.

[115] M. K. Gaillard and Benjamin W. Lee. "Rare decay modes of the K mesons in gauge theories". In: *Phys. Rev. D* 10 (3 Aug. 1974), pp. 897–916. DOI: 10.1103/PhysRevD.10.897. URL: https://link.aps.org/doi/10.1103/PhysRevD.10.897.

[116] Takeo Inami and C. S. Lim. "Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K_L \to \mu\mu$, $K^+ \to \pi\nu\nu$ and $K^0 \leftrightarrow \bar{K}^0$ ". In: *Progress of Theoretical Physics* 65.1 (Jan. 1981), pp. 297–314.

[117] F.J. Gilman and J.S. Hagelin. "Improved bounds on some weak amplitudes from the b lifetime". In: *Physics Letters B* 133.6 (1983), pp. 443–448. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)90825-0. URL: http://www.sciencedirect.com/science/article/pii/0370269383908250.

[118] John Ellis, J.S. Hagelin, and S. Rudaz. "Re-examination of the standard model in the light of B meson mixing". In: *Physics Letters B* 192.1 (1987), pp. 201–206. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(87)91168-3. URL: http://www.sciencedirect.com/science/article/pii/0370269387911683.

[119] John Ellis et al. "Implications of recent measurements of B meson mixing and ϵ'/ϵ_K". In: *Nuclear Physics B* 304 (1988), pp. 205–235. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(88)90625-6. URL: http://www.sciencedirect.com/science/article/pii/0550321388906256.

[120] Gerhard Buchalla and Andrzej J. Buras. "$K \to \pi\nu\overline{\nu}$ and high precision determinations of the CKM matrix". In: *Phys. Rev. D* 54 (11 Dec. 1996), pp. 6782–6789. DOI: 10.1103/PhysRevD.54.6782. URL: https://link.aps.org/doi/10.1103/PhysRevD.54.6782.

[121] Joachim Brod and Martin Gorbahn. "Electroweak corrections to the charm quark contribution to $K^+ \to \pi^+\nu\overline{\nu}$". In: *Phys. Rev. D* 78 (3 Aug. 2008), p. 034006. DOI: 10.1103/PhysRevD.78.034006. URL: https://link.aps.org/doi/10.1103/PhysRevD.78.034006.

[122] Joachim Brod, Martin Gorbahn, and Emmanuel Stamou. "Two-loop electroweak corrections for the $K \to \pi\nu\overline{\nu}$ decays". In: *Phys. Rev. D* 83 (3 Feb. 2011), p. 034030. DOI: 10.1103/PhysRevD.83.034030. URL: https://link.aps.org/doi/10.1103/PhysRevD.83.034030.

[123] Y. Asano et al. "Search for a rare decay mode $K^+ \to \pi^+\nu\bar{\nu}$ and axion". In: *Physics Letters B* 107.1 (1981), pp. 159–162. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(81)91172-2. URL: http://www.sciencedirect.com/science/article/pii/0370269381911722.

[124] Eduardo Cortina Gil et al. "First search for $K^+ \to \pi^+\nu\bar{\nu}$ using the decay-in-flight technique". In: *Phys. Lett. B* 791.arXiv:1811.08508 (Nov. 2018). 18 pages, 13 figures, 156–166. 11 p. DOI: 10.1016/j.physletb.2019.01.067. URL: https://cds.cern.ch/record/2638467.

[125] Eduardo Cortina Gil et al. "An investigation of the very rare $K^+ \to \pi^+\nu\bar{\nu}$ decay". In: *J. High Energ. Phys.* 11.42 (Nov. 2020). DOI: https://doi.org/10.1007/JHEP11(2020)042. arXiv: 2007.08218 [hep-ex].

[126] S. Adler et al. "Further search for the decay $K^+ \to \pi^+\nu\bar{\nu}$ in the momentum region $P < 195\text{MeV}/c$". In: *Phys. Rev. D* 70 (3 Aug. 2004), p. 037102. DOI: 10.1103/PhysRevD.70.037102. URL: https://link.aps.org/doi/10.1103/PhysRevD.70.037102.

[127] KOTO Collaboration S. Shinohara. "Search for the rare decay $K_L^0 \to \pi^0\nu\bar{\nu}$ at J-PARC KOTO experiment". In: *KAON19*. Perugia, Italy. Sept. 2019. URL: https://indico.cern.ch/event/769729/contributions/3510939/attachments/1904988/3145907/KAON2019_shinohara_upload.pdf.

[128] Nobuhiro Shimizu for the KOTO Collaboration. "Search for New Physics via the $K_L \to \pi^0\nu\bar{\nu}$ decay at the J-PARC KOTO experiment". In: *ICHEP20*. Prague, July 2020. URL: https://indico.cern.ch/event/868940/contributions/3815582/.

[129] Teppei Kitahara et al. "New physics implications of recent search for $K_L \to \pi^0\nu\bar{\nu}$ at KOTO". In: *Phys. Rev. Lett.* 124.7 (2020), p. 071801. DOI: 10.1103/PhysRevLett.124.071801. arXiv: 1909.11111 [hep-ph].

[130] Robert Ziegler, Jure Zupan, and Roman Zwicky. "Three Exceptions to the Grossman-Nir Bound". In: (May 2020). arXiv: 2005.00451 [hep-ph].

[131] Yi Liao et al. "An imprint of a new light particle at KOTO?" In: (May 2020). arXiv: 2005.00753 [hep-ph].

[132] Matheus Hostert, Kunio Kaneta, and Maxim Pospelov. "Pair production of dark particles in meson decays". In: (May 2020). arXiv: 2005.07102 [hep-ph].

[133] Daniel Egana-Ugrinovic, Samuel Homiller, and Patrick Meade. "Light Scalars and the KOTO Anomaly". In: *Phys. Rev. Lett.* 124.19 (2020), p. 191801. DOI: 10.1103/PhysRevLett.124.191801. arXiv: 1911.10203 [hep-ph].

[134] P.S. Bhupal Dev, Rabindra N. Mohapatra, and Yongchao Zhang. "Constraints on long-lived light scalars with flavor-changing couplings and the KOTO anomaly". In: *Phys. Rev. D* 101.7 (2020), p. 075014. DOI: 10.1103/PhysRevD.101.075014. arXiv: 1911.12334 [hep-ph].

[135] James M. Cline, Matteo Puel, and Takashi Toma. "A little theory of everything, with heavy neutral leptons". In: *JHEP* 05 (2020), p. 039. DOI: 10.1007/JHEP05(2020)039. arXiv: 2001.11505 [hep-ph].

[136] Xiao-Gang He et al. "Evading the Grossman-Nir bound with $\Delta I = 3/2$ new physics". In: (May 2020). arXiv: 2005.02942 [hep-ph].

[137] Stefania Gori, Gilad Perez, and Kohsaku Tobioka. "KOTO vs. NA62 Dark Scalar Searches". In: (May 2020). arXiv: 2005.05170 [hep-ph].

[138] Tong Li, Xiao-Dong Ma, and Michael A. Schmidt. "Implication of $K \to \pi\nu\bar{\nu}$ for generic neutrino interactions in effective field theories". In: *Phys. Rev. D* 101.5 (2020), p. 055019. DOI: 10.1103/PhysRevD.101.055019. arXiv: 1912.10433 [hep-ph].

[139] Xiao-Gang He et al. "Breaking the Grossman-Nir Bound in Kaon Decays". In: *JHEP* 04 (2020), p. 057. DOI: 10.1007/JHEP04(2020)057. arXiv: 2002.05467 [hep-ph].

[140] Yongsoo Jho et al. "Light gauge boson interpretation for $(g-2)_\mu$ and the $K_L \to \pi^0$ + (invisible) anomaly at the J-PARC KOTO experiment". In: *JHEP* 04 (2020), p. 086. DOI: 10.1007/JHEP04(2020)086. arXiv: 2001.06572 [hep-ph].

[141] Jia Liu et al. "A light scalar explanation of $(g-2)_\mu$ and the KOTO anomaly". In: *JHEP* 04 (2020), p. 197. DOI: 10.1007/JHEP04(2020)197. arXiv: 2001.06522 [hep-ph].

[142] Alakabha Datta, Saeed Kamali, and Danny Marfatia. "Dark sector origin of the KOTO and MiniBooNE anomalies". In: (May 2020). arXiv: 2005.08920 [hep-ph].

[143] CKMfitter Collaboration. *Prospective studies on $K^+ \to \pi^+ \nu\bar{\nu}$ and $K_L \to \pi^0 \nu\bar{\nu}$ rare kaon decays*. 2015. URL: http://ckmfitter.in2p3.fr/www/studies/plots_kpinunu15/ckm_plots_kpinunu15.html.

[144] Y. Chikashige, Rabindra N. Mohapatra, and R. D. Peccei. "Are There Real Goldstone Bosons Associated with Broken Lepton Number?" In: *Phys. Lett.* 98B (1981), pp. 265–268. DOI: 10.1016/0370-2693(81)90011-3.

[145] G.B. Gelmini and M. Roncadelli. "Left-handed neutrino mass scale and spontaneously broken lepton number". In: *Phys. Lett. B* 99.5 (1981), pp. 411–415. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(81)90559-1. URL: http://www.sciencedirect.com/science/article/pii/0370269381905591.

[146] Frank Wilczek. "Axions and Family Symmetry Breaking". In: *Phys. Rev. Lett.* 49 (Nov. 1982), pp. 1549–1552. URL: https://link.aps.org/doi/10.1103/PhysRevLett.49.1549.

[147] Lorenzo Calibbi et al. "Minimal axion model from flavor". In: *Phys. Rev. D* 95.9 (2017), p. 095009. DOI: 10.1103/PhysRevD.95.095009. arXiv: 1612.08040 [hep-ph].

[148] D.S. Gorbunov. "Light sgoldstino: precision measurements versus collider searches". In: *Nuclear Physics B* 602.1 (2001), pp. 213–237. URL: http://www.sciencedirect.com/science/article/pii/S0550321301001225.

[149] T.M. Aliev, M.L. Dobroliubov, and A.Yu. Ignatiev. "The decay $K^+ \to \pi^+ X$ in $SU(2) \times U(1) \times U'(1)$ gauge theories". In: *Nucl. Phys.* 335.2 (1990), pp. 311–333. URL: http://www.sciencedirect.com/science/article/pii/055032139090496Z.

[150] Maxim Pospelov. "Secluded U(1) below the weak scale". In: *Phys. Rev. D* 80 (9 Nov. 2009), p. 095002. DOI: 10.1103/PhysRevD.80.095002. URL: https://link.aps.org/doi/10.1103/PhysRevD.80.095002.

[151] Pierre Fayet. "U-boson production in e^+e^- annihilations, ψ and Υ decays, and light dark matter". In: *Phys. Rev. D* 75 (11 June 2007), p. 115017. DOI: 10.1103/PhysRevD.75.115017. URL: https://link.aps.org/doi/10.1103/PhysRevD.75.115017.

[152] John F. Gunion, Dan Hooper, and Bob McElrath. "Light neutralino dark matter in the next-to-minimal supersymmetric standard model". In: *Phys. Rev. D* 73 (1 Jan. 2006), p. 015011. DOI: 10.1103/PhysRevD.73.015011. URL: https://link.aps.org/doi/10.1103/PhysRevD.73.015011.

[153] Maxim Pospelov, Adam Ritz, and Mikhail Voloshin. "Secluded WIMP dark matter". In: *Phys. Lett. B* 662.1 (2008), pp. 53–61. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2008.02.052. URL: http://www.sciencedirect.com/science/article/pii/S0370269308002402.

[154] F. Bezrukov and D. Gorbunov. "Light inflaton hunter's guide". In: *JHEP* 2010.5 (May 2010), p. 10.

[155] Gordan Krnjaic. "Probing Light Thermal Dark-Matter With a Higgs Portal Mediator". In: *Phys. Rev. D* 94.7 (2016), p. 073009. DOI: 10.1103/PhysRevD.94.073009. arXiv: 1512.04119 [hep-ph].

[156] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 1–29. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2012.08.020. URL: http://www.sciencedirect.com/science/article/pii/S037026931200857X.

[157] S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 30–61. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2012.08.021. URL: http://www.sciencedirect.com/science/article/pii/S0370269312008581.

[158] G. Aad et al. "Combined Measurement of the Higgs Boson Mass in *pp* Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments". In: *Phys. Rev. Lett.* 114 (19 May 2015), p. 191803. DOI: 10.1103/PhysRevLett.114.191803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.114.191803.

[159] Georges Aad et al. "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV". In: *JHEP* 08 (2016), p. 045. DOI: 10.1007/JHEP08(2016)045. arXiv: 1606.02266 [hep-ex].

[160] Martin Wolfgang Winkler. "Decay and detection of a light scalar boson mixing with the Higgs boson". In: *Phys. Rev. D* 99.1 (2019), p. 015018. DOI: 10.1103/PhysRevD.99.015018. arXiv: 1809.01876 [hep-ph].

[161] H. Leutwyler and M.A. Shifman. "Light Higgs particle in decays of K and η mesons". In: *Nuclear Physics B* 343.2 (1990), pp. 369–397. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(90)90475-S. URL: http://www.sciencedirect.com/science/article/pii/055032139090475S.

[162] Jonathan L. Feng et al. "Dark Higgs bosons at the ForwArd Search ExpeRiment". In: *Phys. Rev. D* 97 (5 Mar. 2018), p. 055034. DOI: 10.1103/PhysRevD.97.055034. URL: https://link.aps.org/doi/10.1103/PhysRevD.97.055034.

[163] Brian Batell, Joshua Berger, and Ahmed Ismail. "Probing the Higgs portal at the Fermilab short-baseline neutrino experiments". In: *Phys. Rev. D* 100 (11 Dec. 2019), p. 115039. DOI: 10.1103/PhysRevD.100.115039. URL: https://link.aps.org/doi/10.1103/PhysRevD.100.115039.

[164] Jackson D. Clarke, Robert Foot, and Raymond R. Volkas. "Phenomenology of a very light scalar ($100\,\text{MeV} < m_h < 10\,\text{GeV}$) mixing with the SM Higgs". In: *JHEP* 02 (2014), p. 123. DOI: 10.1007/JHEP02(2014)123. arXiv: 1310.8042 [hep-ph].

[165] Matthew J. Dolan et al. "A taste of dark matter: Flavour constraints on pseudoscalar mediators". In: *JHEP* 03 (2015). [Erratum: JHEP 07, 103 (2015)], p. 171. DOI: 10.1007/JHEP03(2015)171. arXiv: 1412.5174 [hep-ph].

[166] John F. Gunion et al. *The Higgs Hunter's Guide*. Vol. 80. 2000.

[167] Yohei Ema et al. "Flaxion: a minimal extension to solve puzzles in the standard model". In: *JHEP* 01 (2017), p. 096. DOI: 10.1007/JHEP01(2017)096. arXiv: 1612.05492 [hep-ph].

[168] Tommi Alanne, Simone Blasi, and Florian Goertz. "Common source for scalars: Flavored axion-Higgs unification". In: *Phys. Rev. D* 99.1 (2019), p. 015028. DOI: 10.1103/PhysRevD.99.015028. arXiv: 1807.10156 [hep-ph].

[169] A.M. Diamant-Berger et al. "Study of some rare decays of the K+ meson". In: *Physics Letters B* 62.4 (1976), pp. 485–490. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(76)90690-0. URL: http://www.sciencedirect.com/science/article/pii/0370269376906900.

[170] J.R. Batley et al. "Searches for lepton number violation and resonances in $K^{\pm} \to \pi\mu\mu$ decays". In: *Phys. Lett. B* 769.CERN-EP-2016-310 (Dec. 2016). 15 pages, accepted for publication in Physics Letters B, 67–76. 10 p. DOI: 10.1016/j.physletb.2017.03.029. URL: https://cds.cern.ch/record/2239327.

[171] R. Appel et al. "Search for Lepton Flavor Violation in K^+ Decays into a Charged Pion and Two Leptons". In: *Phys. Rev. Lett.* 85 (14 Oct. 2000), pp. 2877–2880. DOI: 10.1103/PhysRevLett.85.2877. URL: https://link.aps.org/doi/10.1103/PhysRevLett.85.2877.

[172] Aleksey Sher et al. "Improved upper limit on the decay $K^+ \to \pi^+\mu^+e^-$". In: *Phys. Rev. D* 72 (1 July 2005), p. 012005. DOI: 10.1103/PhysRevD.72.012005. URL: https://link.aps.org/doi/10.1103/PhysRevD.72.012005.

[173] E. Abouzaid et al. "Search for Lepton-Flavor-Violating Decays of the Neutral Kaon". In: *Phys. Rev. Lett.* 100 (13 Apr. 2008), p. 131803. DOI: 10.1103/PhysRevLett.100.131803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.131803.

[174] D. Ambrose et al. "New Limit on Muon and Electron Lepton Number Violation from $K_L^0 \to \mu^{\pm}e^{\mp}$ Decay". In: *Phys. Rev. Lett.* 81 (26 Dec. 1998), pp. 5734–5737. DOI: 10.1103/PhysRevLett.81.5734. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.5734.

[175] Robert N. Cahn and Haim Harari. "Bounds on the masses of neutral generation-changing gauge bosons". In: *Nuclear Physics B* 176.1 (1980), pp. 135–152. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(80)90067-X. URL: http://www.sciencedirect.com/science/article/pii/055032138090067X.

[176] L.G. Landsberg. "Overview of the CKM physics opportunities beyond K+ –¿ pi+ nu anti-nu". In: *Phys. Atom. Nucl.* 65 (2002), pp. 1749–1770. DOI: 10.1134/1.1515838.

[177] L.G. Landsberg. "Is it still worth searching for lepton flavor violation in rare kaon decays?" In: *Phys. Atom. Nucl.* 68 (2005), pp. 1190–1210. DOI: 10.1134/1.1992575. arXiv: hep-ph/0410261.

[178] Thomas G. Rizzo. "Rare K decays and new physics beyond the standard model". In: *Workshop on CP Violation*. Sept. 1998, pp. 138–147. arXiv: hep-ph/9809526.

[179] Jack L. Ritchie and Stanley G. Wojcicki. "Rare K decays". In: *Rev. Mod. Phys.* 65 (4 Oct. 1993), pp. 1149–1197. DOI: 10.1103/RevModPhys.65.1149. URL: https://link.aps.org/doi/10.1103/RevModPhys.65.1149.

[180] J.M. Frere et al. "Flavor violation with a single generation". In: *JHEP* 03 (2004), p. 001. DOI: 10.1088/1126-6708/2004/03/001. arXiv: hep-ph/0309014.

[181] Rusa Mandal and Antonio Pich. "Constraints on scalar leptoquarks from lepton and kaon physics". In: *JHEP* 12 (2019), p. 089. DOI: 10.1007/JHEP12(2019)089. arXiv: 1908.11155 [hep-ph].

[182] Laurence S. Littenberg and Robert Shrock. "Implications of improved upper bounds on —Delta L— = 2 processes". In: *Phys. Lett. B* 491 (2000), pp. 285–290. DOI: 10.1016/S0370-2693(00)01041-8. arXiv: hep-ph/0005285.

[183] Laurence S. Littenberg and Robert E. Shrock. "Upper bounds on lepton number violating meson decays". In: *Phys. Rev. Lett.* 68 (1992), pp. 443–446. DOI: 10.1103/PhysRevLett.68.443.

[184] A.V. Borisov. "Lepton-Number Violating Meson Decays in Theories Beyond the Standard Model". In: *Helmholz International Summer School on Heavy Quark Physics*. Dec. 2011, pp. 165–174. arXiv: 1112.3269 [hep-ph].

[185] A. Ali, A.V. Borisov, and M.V. Sidorova. "Majorana neutrinos in rare meson decays". In: *Phys. Atom. Nucl.* 69 (2006), pp. 475–484. DOI: 10.1134/S1063778806030100.

[186] Asmaa Abada et al. "Effective Majorana mass matrix from tau and pseudoscalar meson lepton number violating decays". In: *JHEP* 02 (2018), p. 169. DOI: 10.1007/JHEP02(2018)169. arXiv: 1712.03984 [hep-ph].

[187] A. Belyaev et al. "Charged lepton flavor violation in kaon decays in supersymmetric theories". In: *Eur. Phys. J. C* 22 (2002), pp. 715–726. DOI: 10.1007/s100520100824. arXiv: hep-ph/0008276.

[188] Fredrik Björkeroth, Eung Jin Chun, and Stephen F. King. "Flavourful Axion Phenomenology". In: *JHEP* 08 (2018), p. 117. DOI: 10.1007/JHEP08(2018)117. arXiv: 1806.00660 [hep-ph].

[189] Martin Bauer et al. "Axionlike Particles, Lepton-Flavor Violation, and a New Explanation of a_μ and a_e". In: *Phys. Rev. Lett.* 124.21 (2020), p. 211803. DOI: 10.1103/PhysRevLett.124.211803. arXiv: 1908.00008 [hep-ph].

[190] Claudia Cornella, Paride Paradisi, and Olcyr Sumensari. "Hunting for ALPs with Lepton Flavor Violation". In: *JHEP* 01 (2020), p. 158. DOI: 10.1007/JHEP01(2020)158. arXiv: 1911.06279 [hep-ph].

[191] Martino Borsato et al. "Effective-field-theory arguments for pursuing lepton-flavor-violating K decays at LHCb". In: *Phys. Rev. D* 99.5 (2019), p. 055017. DOI: 10.1103/PhysRevD.99.055017. arXiv: 1808.02006 [hep-ph].

[192] J. Bijnens, G. Ecker, and J. Gasser. "Radiative semileptonic kaon decays". In: *Nucl. Phys.* B396 (1993), pp. 81–118. DOI: 10.1016/0550-3213(93)90259-R. arXiv: hep-ph/9209261 [hep-ph].

[193] J. Bijnens, G. Colangelo, and J. Gasser. "K(l4) decays beyond one loop". In: *Nucl. Phys.* B427 (1994), pp. 427–454. DOI: 10.1016/0550-3213(94)90634-3. arXiv: hep-ph/9403390 [hep-ph].

[194] The NA48/2 Collaboration et al. "Precise tests of low energy QCD from K_{e4} decay properties". In: *The European Physical Journal C* 70.3 (Dec. 2010), pp. 635–657. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1480-6. URL: https://doi.org/10.1140/epjc/s10052-010-1480-6.

[195] J. R. Batley et al. "New measurement of the charged kaon semileptonic $K^{\pm} \to \pi^{+}\pi^{-}e^{\pm}\nu$ (K_e4) decay Branching Ratio and Hadronic Form Factors". In: *Phys. Lett.* B715 (2012). [Addendum: Phys. Lett.B740,364(2015)], pp. 105–115. DOI: 10.1016/j.physletb.2012.07.048,10.1016/j.physletb.2014.12.020. arXiv: 1206.7065 [hep-ex].

[196] J. R. Batley et al. "Searches for lepton number violation and resonances in $K^{\pm} \to \pi\mu\mu$ decays". In: *Phys. Lett.* B769 (2017), pp. 67–76. DOI: 10.1016/j.physletb.2017.03.029. arXiv: 1612.04723 [hep-ex].

[197] V. Bisi et al. "K+ Rare Decay Modes". In: *Phys. Lett.* 25B (1967), pp. 572–573. DOI: 10.1016/0370-2693(67)90150-5.

[198] Heinrich Wahl. "First observation and precision measurement of direct CP violation: the experiments NA31 and NA48". In: *Physics Reports* 403-404 (2004), pp. 19–25. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2004.08.007. URL: http://www.sciencedirect.com/science/article/pii/S0370157304003199.

[199] L. Iconomidou-Fayard and Daniel Fournier. "The Discovery of Direct CP Violation". In: *Adv. Ser. Direct. High Energy Phys.* 23 (2015), pp. 205–235. DOI: 10.1142/9789814644150_0009.

[200] Eduardo Cortina Gil et al. "The Beam and detector of the NA62 experiment at CERN". In: *JINST* 12.05 (2017), P05025. DOI: 10.1088/1748-0221/12/05/P05025. arXiv: 1703.08501 [physics.ins-det].

[201] NA62 Collaboration. *2020 NA62 Status Report to the CERN SPSC*. Tech. rep. CERN-SPSC-2020-007. SPSC-SR-266. Geneva: CERN, Mar. 2020. URL: https://cds.cern.ch/record/2713499.

[202] *CAS - CERN Accelerator School: Beam Injection, Extraction and Transfer: Erice, Italy 10 - 19 Mar 2017*. CERN. Geneva: CERN, Dec. 2018. DOI: 10.23730/CYRSP-2018-005. URL: https://cds.cern.ch/record/2646489.

[203] M.A. Fraser et al. "Experimental results of crystal-assisted slow extraction at the SPS". In: *8th International Particle Accelerator Conference*. 2017, MOPIK048. DOI: 10.18429/JACoW-IPAC2017-MOPIK048. arXiv: 1707.05151 [physics.acc-ph].

[204] C. Bovet et al. "The Cedar Counters for Particle Identification in the SPS Secondary Beams: A Description and an Operation Manual". In: (1982). CERN-82-13, CERN-YELLOW-82-13.

[205] Evgueni Goudzovski et al. "Development of the kaon tagging system for the NA62 experiment at CERN". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 801 (2015), pp. 86–94. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2015.08.015. URL: http://www.sciencedirect.com/science/article/pii/S0168900215009493.

[206] F. Ambrosino et al. "CHANTI: a Fast and Efficient Charged Particle Veto Detector for the NA62 Experiment at CERN". In: *JINST* 11.03 (2016), P03029. DOI: 10.1088/1748-0221/11/03/P03029. arXiv: 1512.00244 [physics.ins-det].

[207] Eduardo Cortina Gil et al. "Search for π^0 decays to invisible particles". In: (Oct. 2020). Accepted for publication in JHEP. arXiv: 2010.07644 [hep-ex].

[208] Eduardo Cortina Gil et al. *Search for heavy neutral lepton production in K^+ decays to muons*. Publication in preparation. Jan. 2021.

[209] V. Fanti et al. "The Beam and detector for the NA48 neutral kaon CP violations experiment at CERN". In: *Nucl. Instrum. Meth. A* 574 (2007), pp. 433–471. DOI: 10.1016/j.nima.2007.01.178.

[210] K. Ahmet et al. "The OPAL detector at LEP". In: *Nucl. Instrum. Meth. A* 305 (1991), pp. 275–319. DOI: 10.1016/0168-9002(91)90547-4.

[211] A. Antonelli CERN. *LAV 1 Installation*. CERN Document Server. Available online. Nov. 2008. URL: https://cds.cern.ch/record/1357868?ln=en.

[212] A. Antonelli et al. "Study of the performance of the NA62 small-angle calorimeter at the DAΦNE Linac". In: *Nucl. Instrum. Meth. A* 877 (2018), pp. 178–182. DOI: 10.1016/j.nima.2017.09.065. arXiv: 1610.03827 [physics.ins-det].

[213] F Hahn et al. *NA62: Technical Design Document*. Tech. rep. NA62-10-07. Geneva: CERN, Dec. 2010. URL: https://cds.cern.ch/record/1404985.

[214] D. Moraes et al. "The CARIOCA Front End Chip for the LHCb muon chambers". In: (Jan. 2003).

[215] A Ceccucci et al. "The NA62 Liquid Krypton calorimeter readout module". In: *Journal of Instrumentation* 6.12 (Dec. 2011), pp. C12017–C12017. DOI: 10.1088/1748-0221/6/12/c12017. URL: https://doi.org/10.1088%2F1748-0221%2F6%2F12%2Fc12017.

[216] NA62 Collaboration. *2018 NA62 Status Report to the CERN SPSC*. Tech. rep. CERN-SPSC-2018-010. SPSC-SR-229. Geneva: CERN, Apr. 2018. URL: http://cds.cern.ch/record/2312430.

[217] Gerhard Ecker, Antonio Pich, and Eduardo de Rafael. "$K \to \pi\ell^+\ell^-$ Decays in the Effective Chiral Lagrangian of the Standard Model". In: *Nucl. Phys. B* 291 (1987), pp. 692–719. DOI: 10.1016/0550-3213(87)90491-3.

[218] G. D'Ambrosio et al. "Radiative nonleptonic kaon decays". In: Nov. 1994, pp. 265–313. arXiv: hep-ph/9411439.

[219] G. D'Ambrosio et al. "The Decays $K \to \pi\ell^+\ell^-$ beyond leading order in the chiral expansion". In: *JHEP* 08 (1998), p. 004. DOI: 10.1088/1126-6708/1998/08/004. arXiv: hep-ph/9808289.

[220] Samuel Friot, David Greynat, and Eduardo De Rafael. "Rare kaon decays revisited". In: *Phys. Lett. B* 595 (2004), pp. 301–308. DOI: 10.1016/j.physletb.2004.05.069. arXiv: hep-ph/0404136.

[221] A.Z. Dubnickova et al. "Kaon decay probe of the weak static interaction". In: *Phys. Part. Nucl. Lett.* 5 (2008), pp. 76–84. DOI: 10.1007/s11497-008-2003-9. arXiv: hep-ph/0611175.

[222] Norman H. Christ et al. "First exploratory calculation of the long-distance contributions to the rare kaon decays $K \to \pi\ell^+\ell^-$". In: *Phys. Rev. D* 94.11 (2016), p. 114516. DOI: 10.1103/PhysRevD.94.114516. arXiv: 1608.07585 [hep-lat].

[223] L. Bican for the NA62 Collaboration. "New measurement of the $K^+ \to \pi^+\mu^+\mu^+$ decay at NA62". In: *ICHEP20*. Prague, July 2020. URL: https://indico.cern.ch/event/868940/contributions/3815686/.

[224] J.R. Batley et al. "Precise measurement of the $K^\pm \to \pi^\pm e^+e^-$ decay". In: *Phys. Lett. B* 677.SLAC-PUB-14772. CERN-PH-EP-2009-005. 5 (Mar. 2009), 246–254. 16 p. DOI: 10.1016/j.physletb.2009.05.040. URL: http://cds.cern.ch/record/1167596.

[225] J.R. Batley et al. "New measurement of the $K^+ \to \pi^+\mu^+\mu^-$ decay". In: *Physics Letters B* 697.2 (2011), pp. 107–115. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2011.01.042. URL: http://www.sciencedirect.com/science/article/pii/S0370269311000797.

[226] C. Lazzeroni et al. "Precision Measurement of the Ratio of the Charged Kaon Leptonic Decay Rates". In: *Phys. Lett. B* 719 (2013), pp. 326–336. DOI: 10.1016/j.physletb.2013.01.037. arXiv: 1212.4012 [hep-ex].

[227] C. Lazzeroni et al. "Test of lepton flavour universality in $K^+ \to \ell^+\nu$ decays". In: *Phys. Lett. B* 698 (2011), pp. 105–114. DOI: 10.1016/j.physletb.2011.02.064. arXiv: 1101.4805 [hep-ex].

[228] NA62 Collaboration. *ADDENDUM I TO P326 Continuation of the physics programme of the NA62 experiment*. Tech. rep. CERN-SPSC-2019-039. SPSC-P-326-ADD-1. Geneva: CERN, Oct. 2019. URL: https://cds.cern.ch/record/2691873.

[229] Tomáš Husek, Karol Kampf, and Jiří Novotný. "Rare decay $\pi^0 \to e^+e^-$: on corrections beyond the leading order". In: *Eur. Phys. J. C* 74.8 (2014), p. 3010. DOI: 10.1140/epjc/s10052-014-3010-4. arXiv: 1405.6927 [hep-ph].

[230] E. Abouzaid et al. "Measurement of the Rare Decay $\pi^0 \to e^+e^-$". In: *Phys. Rev. D* 75 (2007), p. 012004. DOI: 10.1103/PhysRevD.75.012004. arXiv: hep-ex/0610072.

[231] Eduardo Cortina Gil et al. "Search for heavy neutral lepton production in K^+ decays". In: *Phys. Lett. B* 778.CERN-EP-2017-311 (Jan. 2018). Version accepted for publication in Phys.Lett.B (12/01/2018), 137–145. 9 p. DOI: 10.1016/j.physletb.2018.01.031. URL: https://cds.cern.ch/record/2284919.

[232] Eduardo Cortina Gil et al. "Search for heavy neutral lepton production in K^+ decays to positrons". In: *Phys. Lett. B* 807 (May 2020), p. 135599. arXiv: 2005.09575 [hep-ex].

[233] Max J. Riedl. "The Mangin Mirror and Its Primary Aberrations". In: *Appl. Opt.* 13.7 (July 1974), pp. 1690–1694. DOI: 10.1364/AO.13.001690. URL: http://ao.osa.org/abstract.cfm?URI=ao-13-7-1690.

[234] W. Sellmeier. "Ueber die durch die Aetherschwingungen erregten Mitschwingungen der Körpertheilchen und deren Rückwirkung auf die ersteren, besonders zur Erklärung der Dispersion und ihrer Anomalien". In: *Annalen der Physik* 223.11 (1872), pp. 386–403. DOI: 10.1002/andp.18722231105. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.18722231105. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.18722231105.

[235] I. H. Malitson. "Interspecimen Comparison of the Refractive Index of Fused Silica". In: *J. Opt. Soc. Am.* 55.10 (Oct. 1965), pp. 1205–1209. DOI: 10.1364/JOSA.55.001205. URL: http://www.osapublishing.org/abstract.cfm?URI=josa-55-10-1205.

[236] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: http://www.sciencedirect.com/science/article/pii/S0168900203013688.

[237] J. Allison et al. "Geant4 developments and applications". In: *IEEE Transactions on Nuclear Science* 53.1 (2006), pp. 270–278.

[238] J. Allison et al. "Recent developments in Geant4". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 835 (2016), pp. 186–225. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2016.06.125. URL: http://www.sciencedirect.com/science/article/pii/S0168900216306957.

[239] Robert D. Cousins, James T. Linnemann, and Jordan Tucker. "Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process". In: *Nucl. Instrum. Meth. A* 595.2 (2008), pp. 480–501. DOI: 10.1016/j.nima.2008.07.086. arXiv: physics/0702156.

[240] A L Read. "Presentation of search results: the CLs technique". In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (Sept. 2002), pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313. URL: https://doi.org/10.1088%2F0954-3899%2F28%2F10%2F313.

[241] Kyle Cranmer. "Practical Statistics for the LHC". In: arXiv:1503.07622 (Mar. 2015). Comments: presented at the 2011 European School of High-Energy Physics, Cheile Gradistei, Romania, 7-20 September 2011 I expect to release updated versions of this document in the future, 267–308. 41 p. DOI: 10.5170/CERN-2014-003.267. URL: https://cds.cern.ch/record/2004587.

[242] F. Ambrosino et al. "KLEVER: An experiment to measure BR($K_L \to \pi^0\nu\bar{\nu}$) at the CERN SPS". In: (Jan. 2019). arXiv: 1901.03099 [hep-ex].

[243] Thomas Junk. "Confidence level computation for combining searches with small statistics". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 434.2 (1999), pp. 435–443. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(99)00498-2. URL: http://www.sciencedirect.com/science/article/pii/S0168900299004982.

[244] A L Read. "Modified frequentist analysis of search results (the CL_s method)". In: CERN-OPEN-2000-205 (2000). DOI: 10.5170/CERN-2000-005.81. URL: https://cds.cern.ch/record/451614.

[245] S. Adler et al. "Measurement of the $K^+ \to \pi^+ \nu\overline{\nu}$ branching ratio". In: *Phys. Rev. D* 77 (5 Mar. 2008), p. 052003. DOI: 10.1103/PhysRevD.77.052003. URL: https://link.aps.org/doi/10.1103/PhysRevD.77.052003.

[246] Daniel Egana-Ugrinovic, Samuel Homiller, and Patrick Meade. "Light Scalars and the KOTO Anomaly". In: *Phys. Rev. Lett.* 124.19 (2020), p. 191801. DOI: 10.1103/PhysRevLett.124.191801. arXiv: 1911.10203 [hep-ph].

[247] MicroBooNE Collaboration. *Search for a Higgs Portal scalar decaying to electron-positron pairs in MicroBooNE*. MICROBOONE-NOTE-1092-PUB. July 2020. URL: https://microboone.fnal.gov/wp-content/uploads/MICROBOONE-NOTE-1092-PUB.pdf.

[248] F. Bergsma et al. "Search for Axion Like Particle Production in 400-GeV Proton - Copper Interactions". In: *Phys. Lett. B* 157 (1985), pp. 458–462. DOI: 10.1016/0370-2693(85)90400-9.

[249] R. Aaij et al. "Search for long-lived scalar particles in $B^+ \to K^+ \chi(\mu^+\mu^-)$ decays". In: *Phys. Rev. D* 95.7 (2017), p. 071101. DOI: 10.1103/PhysRevD.95.071101. arXiv: 1612.07818 [hep-ex].

[250] Roel Aaij et al. "Search for hidden-sector bosons in $B^0 \to K^{*0}\mu^+\mu^-$ decays". In: *Phys. Rev. Lett.* 115.16 (2015), p. 161802. DOI: 10.1103/PhysRevLett.115.161802. arXiv: 1508.04094 [hep-ex].

[251] Michael S. Turner. "Axions from SN 1987a". In: *Phys. Rev. Lett.* 60 (1988), p. 1797. DOI: 10.1103/PhysRevLett.60.1797.

[252] Joshua A. Frieman, Savas Dimopoulos, and Michael S. Turner. "Axions and Stars". In: *Phys. Rev. D* 36 (1987), p. 2201. DOI: 10.1103/PhysRevD.36.2201.

[253] Adam Burrows, Michael S. Turner, and R.P. Brinkmann. "Axions and SN 1987a". In: *Phys. Rev. D* 39 (1989), p. 1020. DOI: 10.1103/PhysRevD.39.1020.

[254] Rouven Essig et al. "Discovering New Light States at Neutrino Experiments". In: *Phys. Rev. D* 82 (2010), p. 113008. DOI: 10.1103/PhysRevD.82.113008. arXiv: 1008.0636 [hep-ph].

[255] Babette Döbrich et al. "Model-independent bounds on light pseudoscalars from rare B-meson decays". In: *Phys. Lett. B* 790 (2019), pp. 537–544. DOI: 10.1016/j.physletb.2019.01.064. arXiv: 1810.11336 [hep-ph].

[256] T. Yamazaki et al. "Search for a neutral boson in a two-body decay of $K^+ \to \pi^+ X^0$". In: *Phys. Rev. Lett.* 52 (1984), pp. 1089–1091. DOI: 10.1103/PhysRevLett.52.1089.

[257] R. Ammar et al. "Search for the familon via $B^{\pm} \to \pi^{\pm} X^0$, $B^{\pm} \to K^{\pm} X^0$, and $B^0 \to K^0_{(S)} X^0$ decays". In: *Phys. Rev. Lett.* 87 (2001), p. 271801. DOI: 10.1103/PhysRevLett.87.271801. arXiv: hep-ex/0106038.

[258] E. Abouzaid et al. "Final Results from the KTeV Experiment on the Decay $K_L \to \pi^0\gamma\gamma$". In: *Phys. Rev. D* 77 (2008), p. 112004. DOI: 10.1103/PhysRevD.77.112004. arXiv: 0805.0031 [hep-ex].

[259] T. Alexopoulos et al. "Measurements of K(L) branching fractions and the CP violation parameter $|\eta_{\pm}|$". In: *Phys. Rev. D* 70 (2004), p. 092006. DOI: 10.1103/PhysRevD.70.092006. arXiv: hep-ex/0406002.

[260] Elisa Minucci. "Search for Lepton Number and Flavor violations in K decays at the NA62 experiment". PhD thesis. Université catholique de Louvain, Center for Cosmology, Particle Physics and Phenomenology, Louvain-La-Neuve, Nov. 2018. URL: http://cp3.irmp.ucl.ac.be/upload/theses/phd/Thesis_EMinucci.pdf.

[261] R. Frühwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2 (1987), pp. 444–450. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(87)90887-4. URL: http://www.sciencedirect.com/science/article/pii/0168900287908874.

[262] Pierre Billoir, R. Fruhwirth, and M. Regler. "TRACK ELEMENT MERGING STRATEGY AND VERTEX FITTING IN COMPLEX MODULAR DETECTORS". In: *Nucl. Instrum. Meth. A* 241 (1985), pp. 115–131. DOI: 10.1016/0168-9002(85)90523-6.

[263] R. Appel et al. "Improved Limit on the Rate of the Decay $K^+ \to \pi^+\mu^+e^-$". In: *Phys. Rev. Lett.* 85 (12 Sept. 2000), pp. 2450–2453. DOI: 10.1103/PhysRevLett.85.2450. URL: https://link.aps.org/doi/10.1103/PhysRevLett.85.2450.

[264] Olaf Behnke et al. *Data Analysis in High Energy Physics: A Practical Guide to Statistical Methods*. Wiley-VCH, 2013.

[265] R. A. Fisher. *Statistical methods for Research Workers*. 4th ed. Biological Monographs and Manuals ; no 3. Edinburgh: Oliver and Boyd, 1934.

[266] Robert D. Cousins. "Annotated Bibliography of Some Papers on Combining Significances or p-values". In: *arXiv e-prints*, arXiv:0705.2209 (May 2007), arXiv:0705.2209. arXiv: 0705.2209 [physics.data-an].

[267] NA62 Collaboration. *2017 NA62 status report to the CERN SPSC*. Tech. rep. CERN-SPSC-2017-013. SPSC-SR-208. Geneva: CERN, Mar. 2017. URL: https://cds.cern.ch/record/2256924.

[268] Thomas J. Roberts and Daniel M. Kaplan. "G4Beamline Simulation Program for Matter dominated Beamlines". In: (2007). Ed. by C. Petit-Jean-Genaz, pp. 3468–3470. DOI: 10.1109/PAC.2007.4440461.

[269] Luca Lista. "Statistical Methods for Data Analysis in Particle Physics". In: *Lect. Notes Phys.* 941 (2017), pp. 1–257. DOI: 10.1007/978-3-319-62840-0.

[270] A. Sher. "An Improved Limit on the Decay $K^+ \to \pi^+\mu^+e^-$". Available online. PhD thesis. The University of Zurich, 2004. URL: https://www.phy.bnl.gov/~hma/e865/thesis.html%22.

Appendix A

Signal Region Definition for the Study of $K^+ \to \pi^+\nu\bar{\nu}$

The signal regions for the study of the $K^+ \to \pi^+\nu\bar{\nu}$ decay are shown in figure 5.2 in terms of the squared missing mass, m^2_{miss} defined in equation 5.1.1, and the momentum of the candidate π^+ track. The squared missing mass may be calculated in three ways:

1. m^2_{miss}: The K^+ and π^+ four momenta are reconstructed using the 3-momentum measurement of the GTK and STRAW respectively and the corresponding mass hypotheses.

2. m^2_{miss}(No GTK): The average beam K^+ three-momentum and K^+ mass hypothesis is used to build the K^+ 4-momentum while the π^+ is treated as in case 1.

3. m^2_{miss}(RICH): The π^+ 4-momentum is reconstructed using the RICH as a velocity spectrometer to measure the momentum under the assumption of the π^+ mass, with the K^+ treated as in case 1.

The resolution of alternative squared missing masses m^2_{miss}(No GTK) and m^2_{miss}(RICH) is degraded with respect to the default case, by a factor of approximately three [267]. However, this gives an additional tool to reject backgrounds originating from non-Gaussian tails which arise from misreconstruction of GTK or STRAW tracks, for which m^2_{miss}(No GTK) and m^2_{miss}(RICH) respectively are not sensitive.

The full definition of kinematic regions, the signal regions and background regions, is given in table A.1, which includes restrictions in terms of these alternative squared missing mass variables. All regions must have π^+ momentum (as reconstructed by the STRAW) in the range 15–35 GeV/c. The definition of signal region 1 is differentiated between momentum bins. For regions around the $K^+ \to \mu^+\nu_\mu$ background definitions vary as a function of momentum according to function

$$f_{\mu\nu}(m^2_{miss}) = (m^2_{\pi^0} - m^2_\mu)\left(1 - \frac{p_{K^+}}{p_\pi}\right), \tag{A.0.1}$$

which depends on the π^0 and μ^+ masses, $m^2_{\pi^0}$ and m^2_μ, the nominal beam kaon momentum magnitude $p_{K^+} = 75\,\text{GeV}/c$ and the π^+ momentum, p_π. All regions are shown in the 2D plane of m^2_{miss} versus p_π in figure 5.2 with boundaries motivated by physical kinematic boundaries and the resolution of the squared missing mass variables.

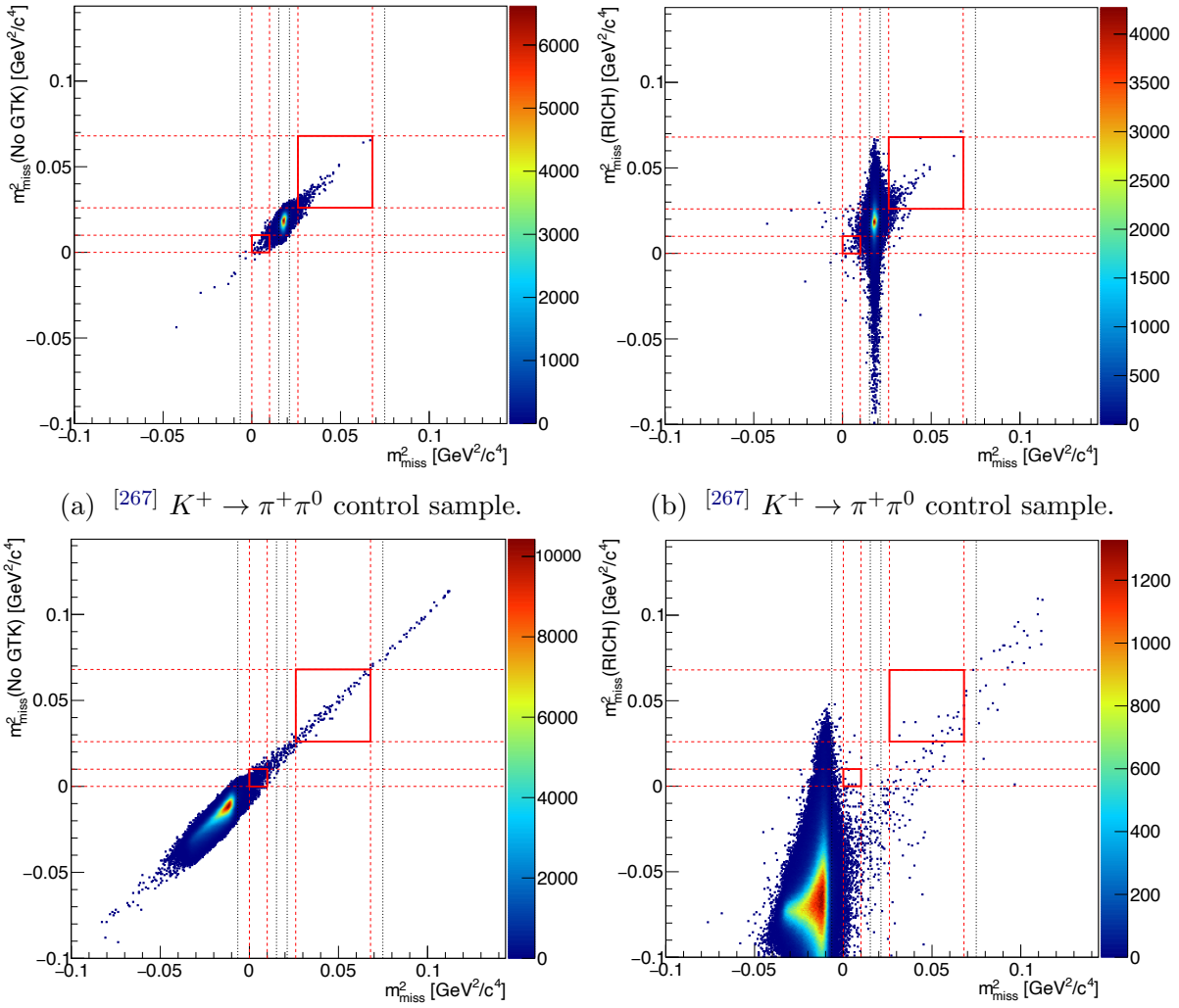
Table A.1: Definition of kinematic regions for the studies of the $K^+ \to \pi^+\nu\bar{\nu}$ decay. In all cases the π^+ candidate momentum reconstructed by the STRAW, p_π must be in the range 15–35 GeV/c, slight modifications are made to signal region 1 for different momentum regions. See figure 5.2 for two dimensional region definitions in the m^2_{miss} versus p_π plane.

Regions		m^2_{miss}	m^2_{miss}(No GTK)	m^2_{miss}(RICH)
Signal Regions	R1 : $p_\pi < 20\,\text{GeV}/c$	0–0.01	−0.005–0.0135	0–0.01
	R1 : $20 < p_\pi < 25\,\text{GeV}/c$	0–0.01	−0.005–0.0135	0–0.02
	R1 : $p_\pi > 25\,\text{GeV}/c$	0–0.01	0–0.0135	0–0.02
	R2	0.026–0.068	0.024–0.068	0.02–0.07
Control Regions	CR1 ($\pi^+\pi^0$)	0.015–0.021	–	–
	CR2 ($\pi^+\pi^0$)	0.021–0.026	–	–
	CR ($\mu\nu$)	$0 - [f_{\mu\nu}(m^2_{miss}) - 0.006]$	–	–
	3πCR	0.068–0.072	–	–
	negR	$< [f_{\mu\nu}(m^2_{miss}) - 0.007]$	–	–
Bkg. Regions	$\pi^+\pi^0$R	0.015–0.021	–	0–0.07
	$\mu\nu$R	$-0.05 - [f_{\mu\nu}(m^2_{miss}) + 0.036]$	–	–
	3πR	> 0.72	–	–

Figure A.1 shows the distribution of $K^+ \to \pi^+\pi^0$ and $K^+ \to \mu^+\nu_\mu$ control samples in the plane of alternative definitions of the squared missing mass versus the standard definition. The signal regions and $\pi^+\pi^0$ background region are illustrated in the m^2_{miss}(RICH) versus m^2_{miss} plane in figure A.2. Figure A.2 shows a sub-sample of 2016 data passing the signal $K^+ \to \pi^+\nu\bar{\nu}$ selection before π^0 rejection with π^+ candidates in momentum range 15–35 GeV/c. By comparison to figure A.1 the background from $K^+ \to \pi^+\pi^0$ decays can be seen including the diagonal line entering into signal region 2 as well as $K^+ \to \pi^+\pi^+\pi^-$ beyond the signal regions in the upper right corner and $K^+ \to \mu^+\nu_\mu$ in the lower left corner.

The three-dimensional squared missing mass definitions of the signal regions leads to slight decreases in the signal acceptance and improved rejection of background especially at the boundaries of the signal regions which can be visible when results are illustrated in terms of m^2_{miss} alone.

Table A.1 and details presented in this section apply directly to analysis of 2016 and 2017 data. For the 2016 $K^+ \to \pi^+\pi^+\pi^-$ background studies, see section 5.2, the 3πR border was slightly shifted to $0.075\,\text{GeV}^2/c^4$ (set prior to final 3πR definition), the impact of this is negligible and was accounted for in results. In the analysis of 2017 data additional control regions were added at higher π^+ candidate momenta. For analysis of 2018 data signal region 2 was expanded to cover the momentum range 15–45 GeV/c, further control regions were added and some were redefined considering the signal region change.



(a) [267] $K^+ \to \pi^+\pi^0$ control sample.

(b) [267] $K^+ \to \pi^+\pi^0$ control sample.

(c) [267] $K^+ \to \mu^+\nu_\mu$ control sample.

(d) [267] $K^+ \to \mu^+\nu_\mu$ control sample.

Figure A.1: Control samples of $K^+ \to \pi^+\pi^0$ [upper, (a) and (b)] and and $K^+ \to \mu^+\nu_\mu$ [lower, (c) and (d)] selected in a subset of 2016 data, showing their distributions in two-dimensional planes of an alternative squared missing mass variable, m^2_{miss}(No GTK) [(a) and (c)] or m^2_{miss}(RICH) [(b) and (d)], versus the default m^2_{miss}. The signal regions are indicated by thick red boxes and kinematic boundaries are indicated by grey dotted lines, from right to left: $f_{\mu\nu}(m^2_{miss})_{max}(p_{pi} = 35\,\mathrm{GeV}/c) - 3\sigma_{m^2_{miss}}$, two lines around the π^0 mass squared at $|m^2_{miss} - m^2_{\pi^0}| = 3\sigma_{m^2_{miss}}$ and below the 3π kinematic endpoint at $4m^2_{\pi^+} - 3\sigma_{m^2_{miss}}$. These define the boundaries of control and background regions.

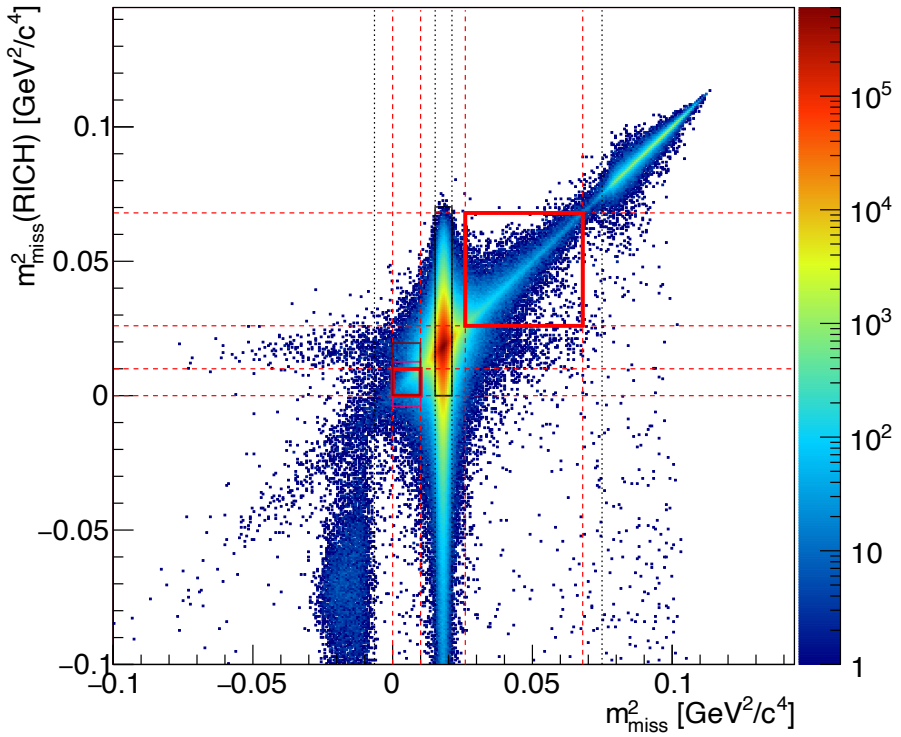


Figure A.2: ([267], edited) Distribution of selected events in the m^2_{miss}(RICH) versus m^2_{miss} plane for a subsample of 2016 data at an intermediate stage of the signal $K^+ \to \pi^+\nu\bar{\nu}$ selection before π^0 rejection and with π^+ candidate track momentum between 15 and 35 GeV/c. The signal regions are shown by thick red boxes (with the slight variations for different momentum bins, see table A.1, indicated by overlaid dark red and dark magenta boxes) and the $\pi^+\pi^0$ background region is indicated by the black box. Vertical dotted grey lines indicate kinematic boundaries, see caption of figure A.1.

Appendix B

Additional Information Regarding The $K^+ \to \pi^+\pi^+\pi^-$ Background In $K^+ \to \pi^+\nu\bar{\nu}$ Studies

B.1 Avoiding Bias and Establishing the Validity of the Analysis Procedure

The analysis concept is non-standard and therefore it is important to perform checks to ascertain its validity. In particular there is a clear risk of introducing bias in selection of the control sample of $K^+ \to \pi^+\pi^+\pi^-$ events with a possible π_3 candidate. The 2π tagging procedure aims to mitigate this by remaining independent of the π_3 kinematics, although this in itself could introduce bias. Therefore alongside the standard analysis procedure (selecting $K^+ \to \pi^+\pi^+\pi^-$ events using 2π tagging followed by a $\pi\nu\bar{\nu}$-like selection) four other similar procedures were performed. Comparison of the results of the different procedures should reveal any bias introduced. The five analysis procedures carried out were:

1. **$K^+ \to \pi^+\pi^+\pi^-$ selection via 2π tagging followed by a $\pi\nu\bar{\nu}$-like selection.**
   This is the standard analysis procedure, it is applied to the data, and details are provided in section 5.2.5.

2. **Simulation study A.**
   The same standard analysis procedure is applied to simulated $K^+ \to \pi^+\pi^+\pi^-$ events. Comparison with case 1 provides a test of the extent to which simulations can describe the data for this specific scenario.

3. **Simulation study B: no 2π tagging, $\pi\nu\bar{\nu}$-like selection applied to all downstream tracks.**
   The 2π tagging selection is bypassed and all downstream tracks are considered as potential π_3 candidates. However, the photon veto conditions, usually applied in the 2π tagging selection, are maintained since these must be applied for a $\pi\nu\bar{\nu}$-like

selection. This applies to the subsequent two cases which also skip the 2π tagging selection. This removal of the 2π tagging selection is possible for simulation studies since these explicitly include only $K^+ \to \pi^+\pi^+\pi^-$ events, while with data the pre-selection of the $K^+ \to \pi^+\pi^+\pi^-$ control sample must be performed to isolate such events.

4. **Simulation study C: no 2π tagging, $\pi\nu\bar{\nu}$-like selection applied to events with at least one π^+.**
Here the 2π tagging selection is bypassed and, for any event with at least one π^+ identified according to the criteria below, each downstream track is considered as a potential π_3 candidate. This accounts for the fact that misreconstruction may misassociate detector signals to incorrect tracks. The criteria to identify a π^+ are:

   - Track has charge +1.
   - LKr $E/p < 0.9$.
   - No associated MUV3 signals.
   - Track momentum in range 15–35 GeV/c.
   - RICH most likely hypothesis is π^+.

5. **Simulation study D: no 2π tagging, $\pi\nu\bar{\nu}$-like selection applied to events with exactly one π^+.**
The 2π tagging selection is bypassed and, for events with exactly one π^+ (passing conditions above), the selected track is considered as a potential π_3 candidate. This is the closest $\pi\nu\bar{\nu}$-like analogue considered without 2π tagging.

B.2 $K^+ \to \pi^+\pi^+\pi^-$ Background Kinematic Tail Fraction Results From All Analysis Procedures

The total number of events observed in each of the five analysis procedures for the regions of interest is given in table B.1 and the corresponding m^2_{miss} distributions displayed in figure B.1. The two-dimensional histograms of m^2_{miss} versus z_{vtx} for all five analysis procedures are shown in figure B.2.

Table B.1: Number of Cases Passing Selection in RII, CR and RCR.

Analysis	RII	CR	RCR	Total
1. Data	4	25	0	252355
2. $K_{3\pi}$ MC	0	14	0	101535
3. $K_{3\pi}$ MC+!2πTag	25	1529	11	1011823
4. MC+!2πTag+AL1π^+	23	1496	10	986447
5. MC+!2πTag+E1π^+	10	770	6	501920

The tail fraction results for signal region 2, the full $K^+ \to \pi^+\pi^+\pi^-$ control region and restricted control region are given for each of the five analysis procedures in tables B.2, B.3

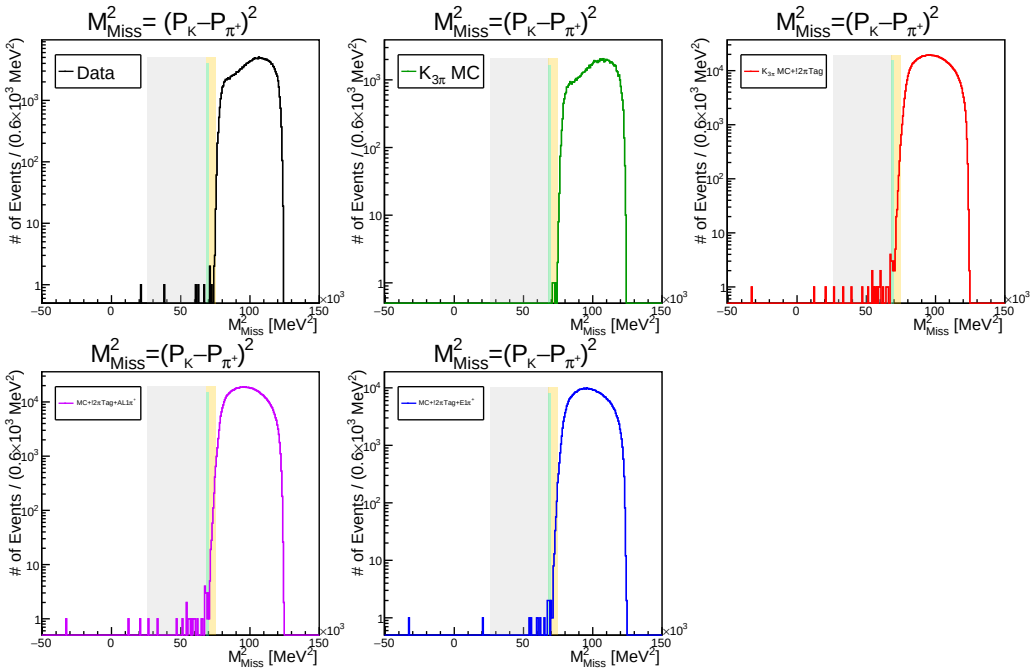


Figure B.1: The K^+–π_3 pair m^2_{miss} distributions obtained from the five analysis procedures described in section B.1. The grey, orange and cyan regions illustrate respectively signal region 2, the full control region and restricted control region.

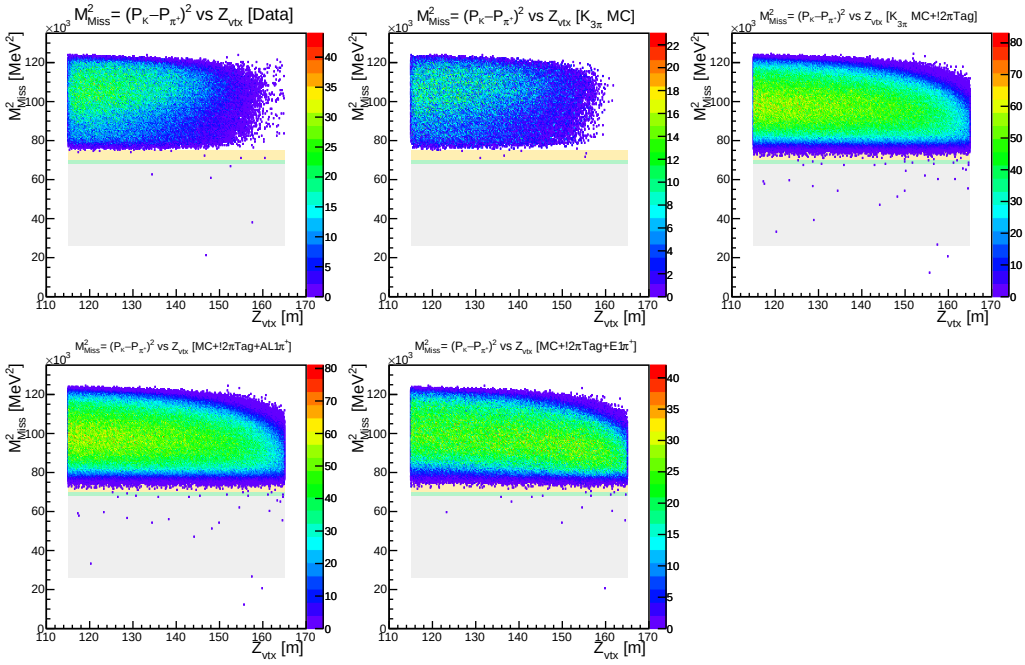


Figure B.2: K^+–π_3 squared missing mass versus vertex z position two dimensional histograms for the five analysis procedures described in section B.1. The grey, orange and cyan regions illustrate respectively signal region 2, the full control region and restricted control region.

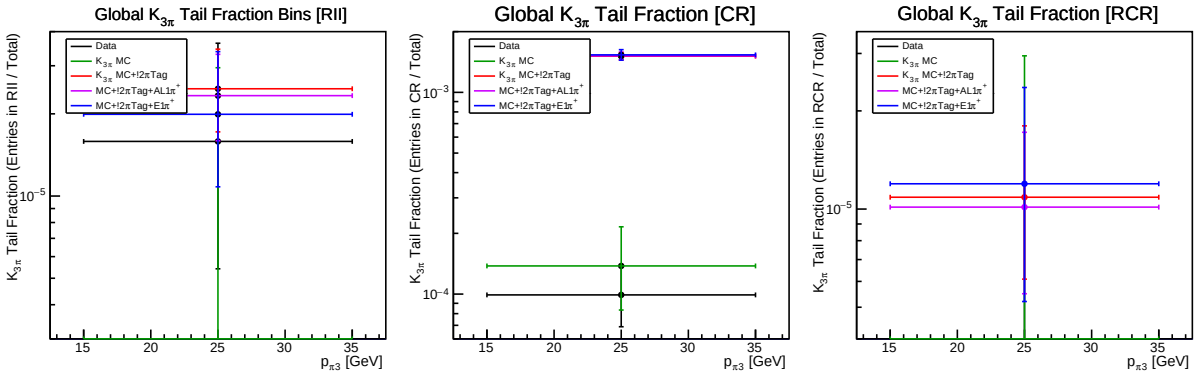


Figure B.3: Total tail fraction results for the three regions of interest and five analysis procedures.

and B.4 respectively. The results are depicted graphically for the three regions of interest in figure B.3 and broken down as a function of momentum in figure B.4 and of z_{vtx} in figure B.5.

Table B.2: Final Results for RII ($26 \times 10^3 < M^2_{Miss} < 68 \times 10^3$ MeV$^2/c^4$).

$z_{vtx}[m]$ \Analysis	Data	$K_{3\pi}$ MC	$K_{3\pi}$ MC +!2πTag	MC +!2πTag+AL $1\pi^+$	MC +!2πTag+E $1\pi^+$
$115-125$ m	$\left(0^{+3.2}_{-0}\right)\times10^{-5}$	$\left(0^{+7.9}_{-0}\right)\times10^{-5}$	$\left(1.6^{+2.0}_{-1.0}\right)\times10^{-5}$	$\left(1.6^{+2.1}_{-1.1}\right)\times10^{-5}$	$\left(9^{+34}_{-9}\right)\times10^{-6}$
$125-135$ m	$\left(1.2^{+4.6}_{-1.2}\right)\times10^{-5}$	$\left(0^{+9.0}_{-0}\right)\times10^{-5}$	$\left(2.1^{+2.3}_{-1.3}\right)\times10^{-5}$	$\left(1.7^{+2.2}_{-1.1}\right)\times10^{-5}$	$\left(0^{+2.8}_{-0}\right)\times10^{-5}$
$135-145$ m	$\left(0^{+5.5}_{-0}\right)\times10^{-5}$	$\left(0^{+1.3}_{-0}\right)\times10^{-4}$	$\left(1.4^{+2.2}_{-1.0}\right)\times10^{-5}$	$\left(1.9^{+2.4}_{-1.2}\right)\times10^{-5}$	$\left(2.8^{+4.5}_{-2.0}\right)\times10^{-5}$
$145-155$ m	$\left(1.0^{+2.0}_{-0.8}\right)\times10^{-4}$	$\left(0^{+3.6}_{-0}\right)\times10^{-4}$	$\left(2.7^{+3.0}_{-1.7}\right)\times10^{-5}$	$\left(2.3^{+2.9}_{-1.5}\right)\times10^{-5}$	$\left(3.0^{+4.7}_{-2.2}\right)\times10^{-5}$
$155-165$ m	$\left(6^{+23}_{-6}\right)\times10^{-4}$	$\left(0^{+6.4}\right)\times10^{-3}$	$\left(7.0^{+5.6}_{-3.5}\right)\times10^{-5}$	$\left(6.3^{+5.6}_{-3.4}\right)\times10^{-5}$	$\left(3.9^{+6.2}_{-2.9}\right)\times10^{-5}$
Total	$\left(1.6^{+2.0}_{-1.0}\right)\times10^{-5}$	$\left(0^{+3.0}_{-0}\right)\times10^{-5}$	$\left(2.47^{+0.98}_{-0.75}\right)\times10^{-5}$	$\left(2.33^{+0.97}_{-0.74}\right)\times10^{-5}$	$\left(2.0^{+1.4}_{-0.9}\right)\times10^{-5}$

Table B.3: Final Results for CR ($68 \times 10^3 < M^2_{Miss} < 75 \times 10^3$ MeV$^2/c^4$).

$z_{vtx}[m]$ \Analysis	Data	$K_{3\pi}$ MC	$K_{3\pi}$ MC +!2πTag	MC +!2πTag+AL $1\pi^+$	MC +!2πTag+E $1\pi^+$
$115-125$ m	$\left(8.6^{+6.9}_{-4.3}\right)\times10^{-5}$	$\left(1.6^{+1.5}_{-0.9}\right)\times10^{-4}$	$\left(1.15^{+0.12}_{-0.11}\right)\times10^{-3}$	$\left(1.15^{+0.12}_{-0.11}\right)\times10^{-3}$	$\left(8.4^{+1.6}_{-1.4}\right)\times10^{-4}$
$125-135$ m	$\left(8.5^{+7.5}_{-4.5}\right)\times10^{-5}$	$\left(6^{+12}_{-5}\right)\times10^{-5}$	$\left(1.27^{+0.13}_{-0.12}\right)\times10^{-3}$	$\left(1.28^{+0.13}_{-0.12}\right)\times10^{-3}$	$\left(8.9^{+1.6}_{-1.4}\right)\times10^{-4}$
$135-145$ m	$\left(1.3^{+1.1}_{-0.7}\right)\times10^{-4}$	$\left(1.8^{+2.4}_{-1.2}\right)\times10^{-4}$	$\left(1.43^{+0.14}_{-0.13}\right)\times10^{-3}$	$\left(1.42^{+0.14}_{-0.13}\right)\times10^{-3}$	$\left(1.31^{+0.20}_{-0.18}\right)\times10^{-3}$
$145-155$ m	$\left(1.0^{+2.0}_{-0.8}\right)\times10^{-4}$	$\left(0^{+3.64}_{-0}\right)\times10^{-4}$	$\left(1.84^{+0.17}_{-0.16}\right)\times10^{-3}$	$\left(1.85^{+0.18}_{-0.16}\right)\times10^{-3}$	$\left(2.02^{+0.35}_{-0.23}\right)\times10^{-3}$
$155-165$ m	$\left(6^{+23}_{-6}\right)\times10^{-4}$	$\left(4.3^{+9.1}_{-3.5}\right)\times10^{-3}$	$\left(2.46^{+0.25}_{-0.24}\right)\times10^{-3}$	$\left(2.49^{+0.26}_{-0.24}\right)\times10^{-3}$	$\left(3.11^{+0.35}_{-0.32}\right)\times10^{-3}$
Total	$\left(9.9^{+4.0}_{-3.0}\right)\times10^{-5}$	$\left(1.38^{+0.78}_{-0.55}\right)\times10^{-4}$	$\left(1.511^{+0.065}_{-0.063}\right)\times10^{-3}$	$\left(1.517^{+0.066}_{-0.064}\right)\times10^{-3}$	$\left(1.534^{+0.094}_{-0.090}\right)\times10^{-3}$

Table B.4: Final Results for RCR ($68 \times 10^3 < M^2_{Miss} < 70 \times 10^3$ MeV$^2/c^4$).

$z_{vtx}[m]$ \Analysis	Data	$K_{3\pi}$ MC	$K_{3\pi}$ MC +!2πTag	MC +!2πTag+AL $1\pi^+$	MC +!2πTag+E $1\pi^+$
$115-125$ m	$\left(0^{+3.2}_{-0}\right)\times10^{-5}$	$\left(0^{+8.0}_{-0}\right)\times10^{-5}$	$\left(0^{+1.2}_{-0}\right)\times10^{-5}$	$\left(0^{+1.2}_{-0}\right)\times10^{-5}$	$\left(0^{+2.7}_{-0}\right)\times10^{-5}$
$125-135$ m	$\left(0^{+3.6}_{-0}\right)\times10^{-5}$	$\left(0^{+9.0}_{-0}\right)\times10^{-5}$	$\left(1.2^{+2.0}_{-9.3}\right)\times10^{-6}$	$\left(1.3^{+2.0}_{-0.9}\right)\times10^{-5}$	$\left(0^{+2.8}_{-0}\right)\times10^{-5}$
$135-145$ m	$\left(0^{+5.4}_{-0}\right)\times10^{-5}$	$\left(0^{+1.4}_{-0}\right)\times10^{-4}$	$\left(0^{+1.4}_{-0}\right)\times10^{-5}$	$\left(0^{+1.4}_{-0}\right)\times10^{-5}$	$\left(0^{+2.8}_{-0}\right)\times10^{-5}$
$145-155$ m	$\left(0^{+1.4}_{-0}\right)\times10^{-4}$	$\left(0^{+3.6}_{-0}\right)\times10^{-4}$	$\left(1.6^{+2.6}_{-2.0}\right)\times10^{-5}$	$\left(1.1^{+2.4}_{-0.9}\right)\times10^{-5}$	$\left(1.0^{+3.7}_{-0.9}\right)\times10^{-5}$
$155-165$ m	$\left(0^{+1.8}_{-0}\right)\times10^{-3}$	$\left(0^{+6.4}_{-0}\right)\times10^{-3}$	$\left(4.4^{+4.8}_{-2.6}\right)\times10^{-5}$	$\left(4.5^{+5.0}_{-2.7}\right)\times10^{-5}$	$\left(6.6^{+7.2}_{-4.0}\right)\times10^{-5}$
Total	$\left(0^{+1.2}_{-0}\right)\times10^{-5}$	$\left(0^{+3.0}_{-0}\right)\times10^{-5}$	$\left(1.09^{+0.71}_{-0.48}\right)\times10^{-5}$	$\left(1.01^{+0.71}_{-0.46}\right)\times10^{-5}$	$\left(1.2^{+1.6}_{-0.7}\right)\times10^{-5}$

Informed by the results obtained for signal region 2 (table B.2) a conservative upper limit on the tail fraction is proposed as $F^{RII}_{K_{3\pi}\ texttail} < 5 \times 10^{-5}$.

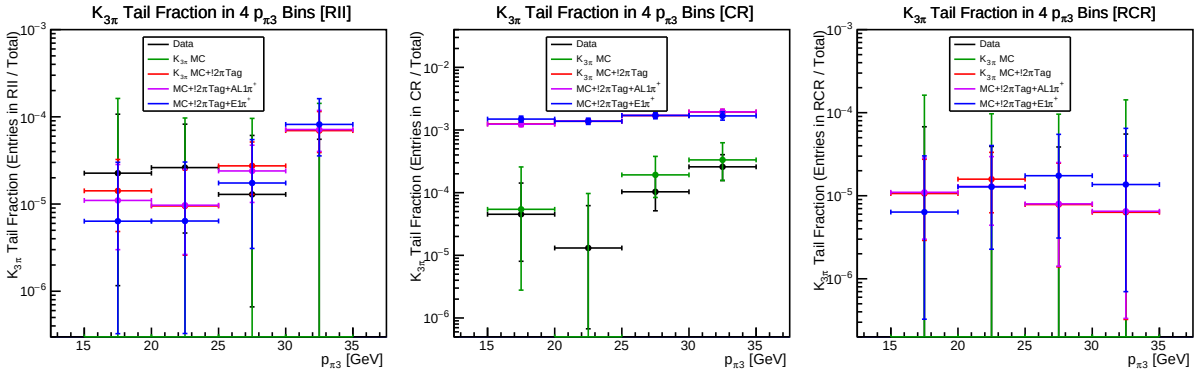


Figure B.4: Tail fraction results in the four momentum regions for the three regions of interest and five analysis procedures.

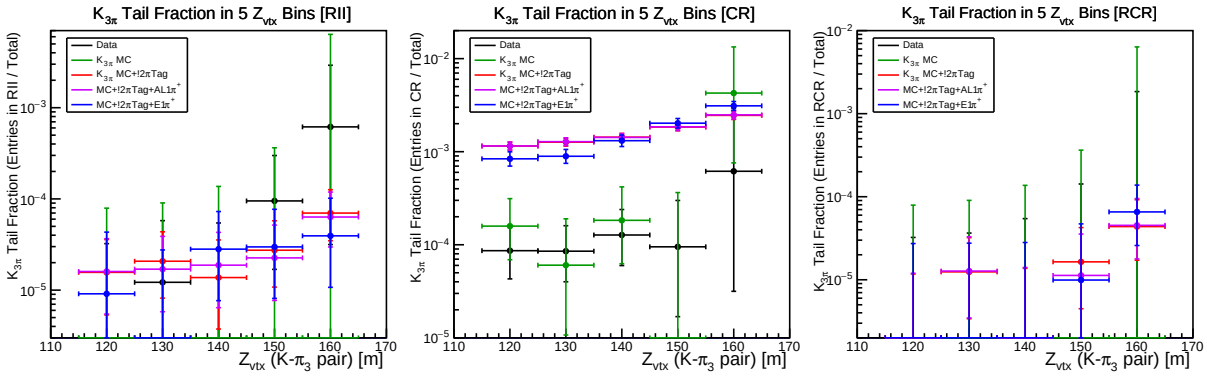


Figure B.5: Tail fraction results in the z_{vtx} regions for the three regions of interest and five analysis procedures.

For the $K^+ \to \pi^+\pi^+\pi^-$ control region two separate upper limits are established because of the clear difference between the results with and without the 2π tagging procedure. This difference is a side-effect of the significant difference in the m^2_{miss} distribution observed with and without the requirement that at least one two-track vertex is formed, the first selection condition for 2π tagging. Figure B.6 illustrates the difference in the m^2_{miss} spectrum before and after this single requirement. This is understood to create the discrepancy between the analysis with and without 2π tagging clearly seen in the global tail fraction in the control region (figure B.3, centre). The conservative upper limits are: $F^{CR+2\pi T}_{K_{3\pi}Tail} < 3\times10^{-4}$ and $F^{CR!2\pi T}_{K_{3\pi}Tail} < 3\times10^{-3}$ for the cases with and without the application of the 2π tagging procedure respectively.

For the restricted control region similar consideration of results (table B.4) lead to a limit on the tail fraction of $F^{RCR}_{K_{3\pi}Tail} < 4\times10^{-5}$. Here the differences in results between use or not of the 2π tagging procedure are less pronounced due to the narrow width of the region and its position at the lower end of the control region. However, while 0 events are observed in RCR for the two analyses using 2π tagging, events are observed for cases without it. This suggests some impact of the systematic change in kinematic properties of events considered if the requirement is made for at least one two-track vertex to be formed.

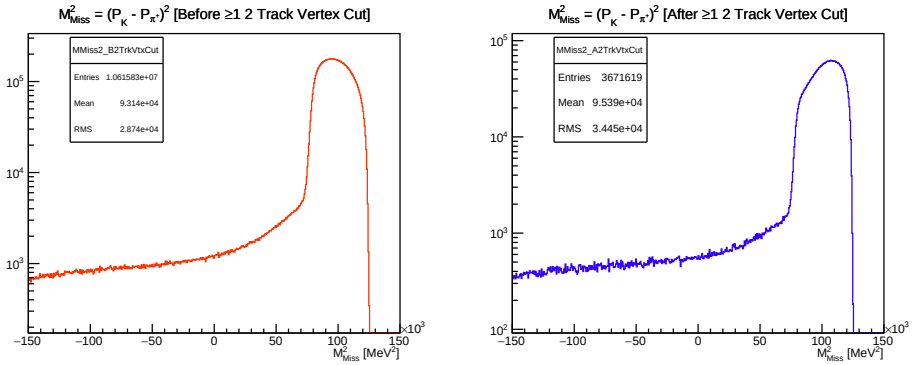


Figure B.6: Reconstructed m^2_{miss} distribution before (left) and after (right) the requirement that at least one two-track vertex is formed.

The tail fraction results can be summarised as:

$$\text{RII} : F^{RII}_{K_{3\pi}Tail} < 5 \times 10^{-5}\,, \tag{B.2.1}$$

$$\text{CR } (+2\pi \text{ Tagging}) : F^{CR+2\pi T}_{K_{3\pi}Tail} < 3 \times 10^{-4}\,, \tag{B.2.2}$$

$$\text{CR (No } 2\pi \text{ Tagging)} : F^{CR!2\pi T}_{K_{3\pi}Tail} < 3 \times 10^{-3}\,, \tag{B.2.3}$$

$$\text{RCR} : F^{RCR}_{K_{3\pi}Tail} < 4 \times 10^{-5}\,. \tag{B.2.4}$$

Appendix C

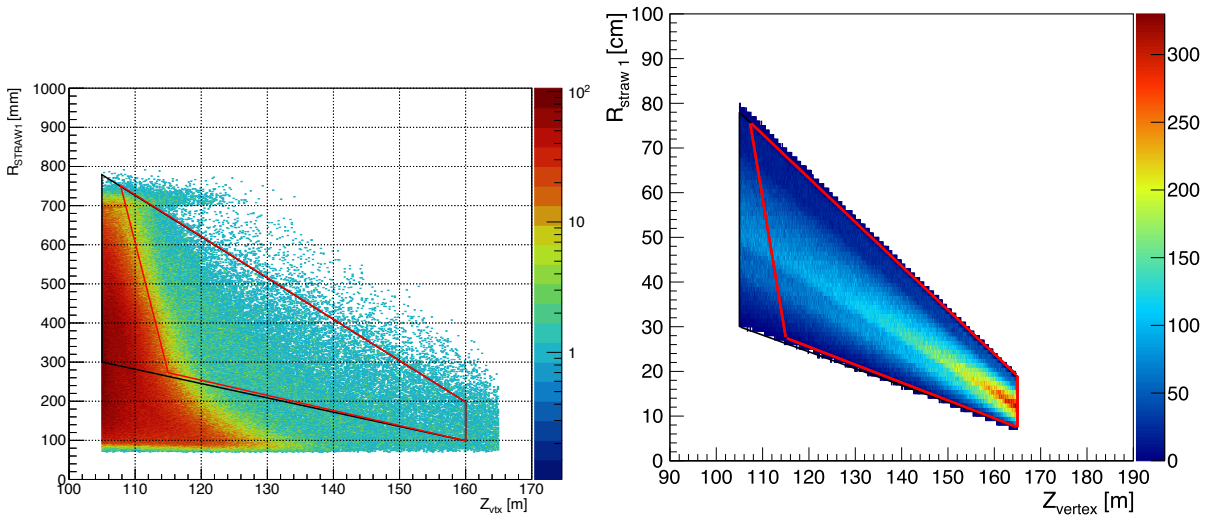
Detailed Qualitative Study of Upstream Background Simulations

C.1 Boosting Statistical Power Using a Library of Upstream Pileup Tracks

To become a background event upstream decays require pileup GTK tracks to be matched to the candidate π^+. A library of 60,000 upstream tracks and upstream detector responses, out-of-time with physics events in 2017 data, are stored and can then be injected into a simulated event by the `UpstreamPileupGenerator` tool. By default this injection of pileup is random and based on a template intensity profile derived from 2017 data. However, a system was developed allowing injection of a specific pileup event from the library and in this way, at least in principle, for every $K^+ \to \pi^+[\text{invisible}]_{\pi^0}$ decay generated 60,000 possible upstream pileup scenarios can be injected. This dramatically boosts the statistical power of each simulated decay event. Unfortunately this leads to bias; because of the very specific geometric conditions needed to form a snake or mamba-like signature if a good match is made between the π^+ and one pileup GTK track it is likely that many other pileup GTK tracks in the pileup library will also match. This results in nonphysical results with localised spikes in differential distributions in one and two dimensions and therefore cannot be used for detailed studies. However, it does allow qualitative studies to be performed, with bias being negligible when using each simulated event less than 10 times and injecting different upstream pileup tracks. Moreover this functionality allowed the study of time profiles, as shown in section 5.3.2.3, by varying the size of the half-window in which pileup tracks are injected, $t_{\frac{1}{2}w}$, as shown in figure 5.25.

C.2 Impact of R_{STRAW1} versus z_{vtx} Cut

In the selection a two dimensional region is isolated in the radius of a track at STRAW1, R_{STRAW1}, and the z position of the K^+–π^+ vertex, z_{vtx}, as illustrated in figure C.1, to



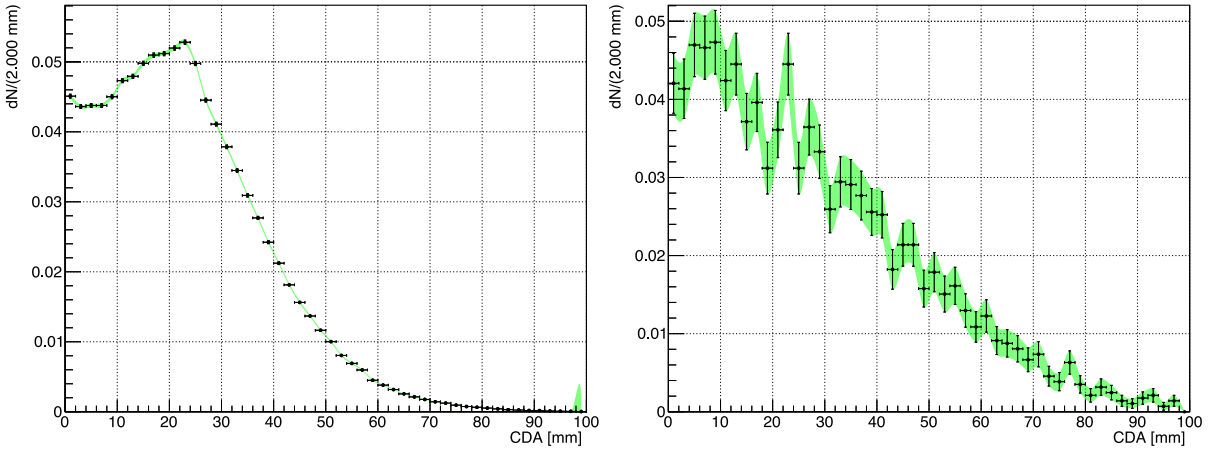
(a) (b) [267]

Figure C.1: Distribution of reconstructed π^+ track radius at STRAW1 versus the reconstructed K^+–π^+ vertex z position from simulations of upstream K^+ decays (left) and signal $K^+ \to \pi^+\nu\bar{\nu}$ (right) with π^+ track momentum in range 15–35 GeV/c and requiring $105 < z_{vtx} < 165$ m. The black line indicates the expected distribution for $K^+ \to \pi^+\nu\bar{\nu}$ events and the red line indicates the region within which events are selected, rejecting upstream background.

reject upstream and multi-body backgrounds. If this requirement is not enforced a larger sample size of of this class of events is available for studies, however the properties of this super-set no longer match the upstream background events. As an example, the change in the CDA distribution is shown in figure C.2. This requirement also rejects events with no scattering in STRAW1, as illustrated by figure C.3, for simulation of snakes only. For simulated events the actual position of the π^+ at the collimator exit plane is known and can be compared to the position projected back using the reconstructed π^+ track. If there is no scattering in STRAW1 then the actual and projected positions match (to within the precision of the reconstruction), seen in figure C.3c. If the R_{STRAW1} versus z_{vtx} cut is applied then such events are rejected as shown by figure C.3d. Corresponding distributions of π^+ projected positions (figures C.3a and C.3d respectively) show a greater fractional population of events closer to the central region when the cut is not applied.

C.3 Study of Events With no STRAW1 Scattering

Using the plane of difference in true π^+ position and projected position at the collimator exit, simulated events where no STRAW1 scattering has taken place can be isolated by selecting the central spot with $\Delta R = \sqrt{\Delta x^2 - \Delta y^2} < 5$ mm (see figure C.3c) only. If this is done a strange distribution emerges in the m^2_{miss} versus momentum plane as shown by figure C.4a. The two curving structures observed may be understood once it is found that they arise from decays in the final two BEND magnets (5 and 6), as shown by figure C.4b. The decay inside the magnetic field causes the path of the π^+ to deviate from



(a) (b)

Figure C.2: CDA distributions for $K^+ \to \pi^+\nu\bar{\nu}$ signal regions with (right) and without (left) R_{STRAW1} versus z_{vtx} cut.

the beam axis and, with a small perturbation, can hit the GTK3 in a slightly different position. Depending on the precise nature of the event, matching between pileup hits in GTK1 and this deviated hit in GTK3 lead to mismeasurement of the beam particle momentum (by up to $\mathcal{O}(10\,\text{GeV}/c)$) and therefore a squared missing mass inconsistent with the expectation of $m^2_{\pi^0}$ for standard $K^+ \to \pi^+\pi^0$ decays. This effect is momentum dependent and depends on the direction of perturbation of the upstream track, with the BEND5 and 6 being responsible for the positive and negative tails respectively. The higher the momentum of the π^+ the closer to the expected position of the beam particle the deviated track will be and so the closer m^2_{miss} will be to the expected value of $m^2_{\pi^0}$.

C.4 Impact of Beam Background Cut

Rejection of beam background is included in the selection, effectively removing decays with π^+ projecting back to the very central region at the final collimator corresponding to the expected beam position (see figure 5.19). If this selection criteria is relaxed when analysing upstream K^+ decay simulations the selected events are dominated by those arising from the central beam region, see figure C.5a. These events have a similar two-arm structure in the m^2_{miss} versus momentum plane, figure C.5b, as observed for the no-scattering events discussed above (figure C.4) and is similarly explained by the effects of decays in the magnetic fields, which is clear from figure C.5c.

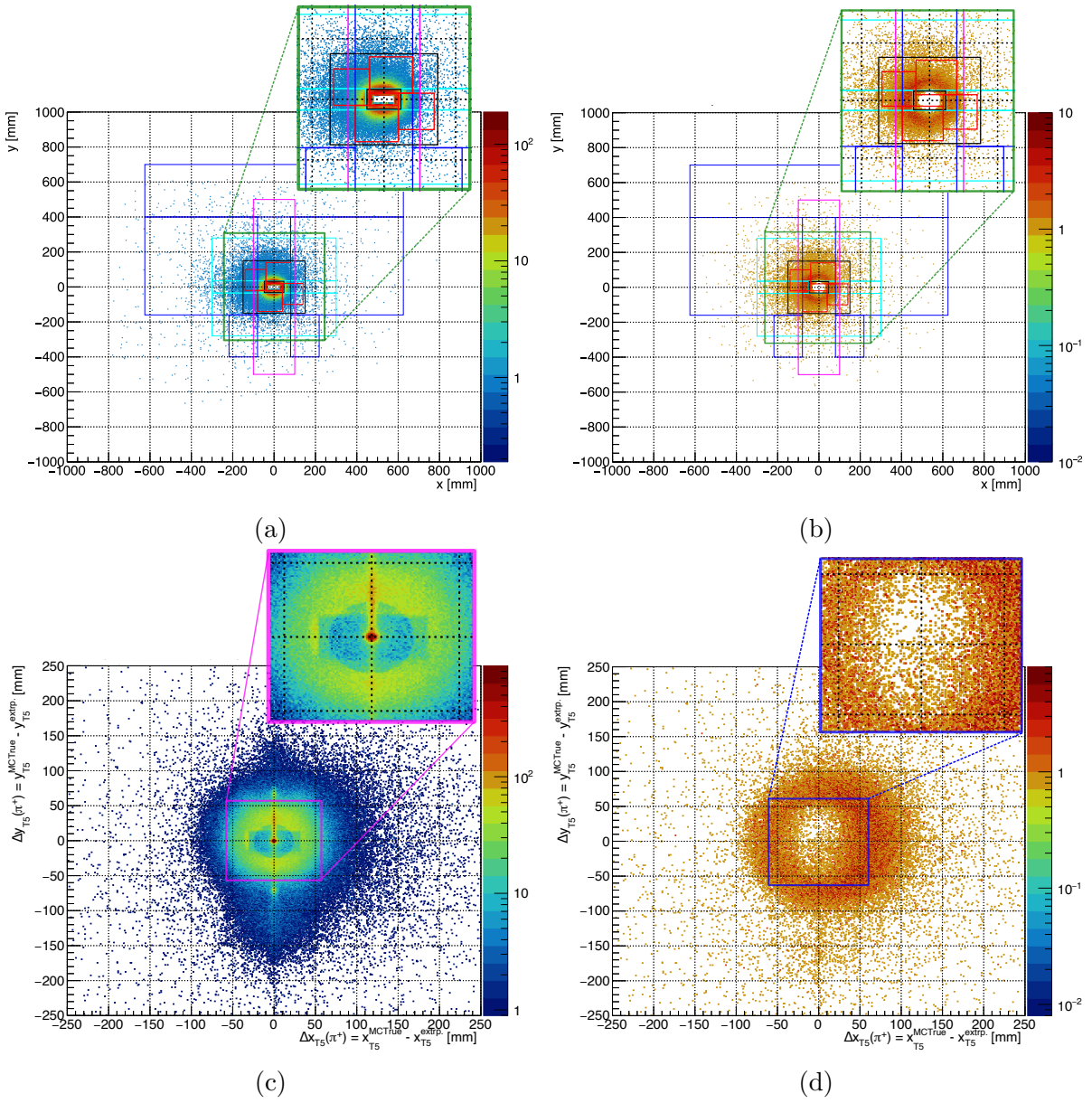
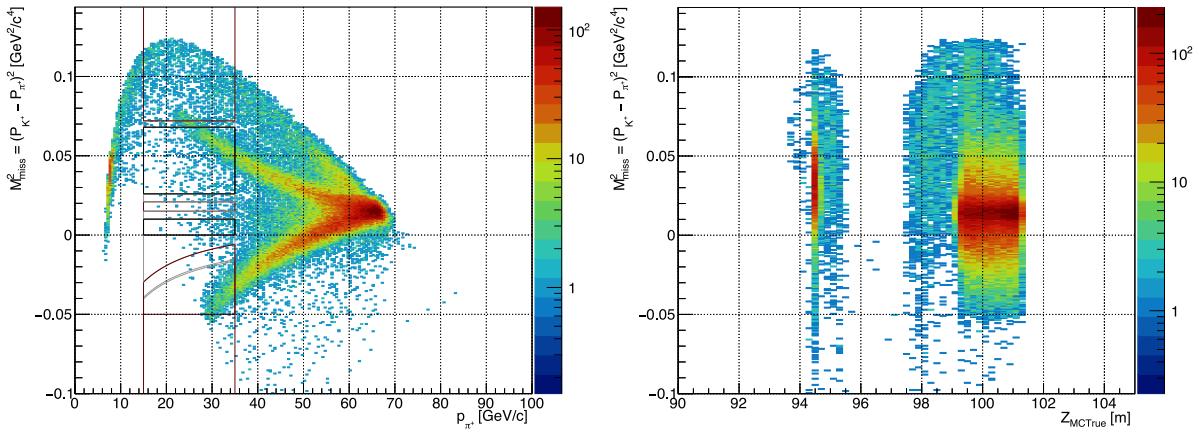
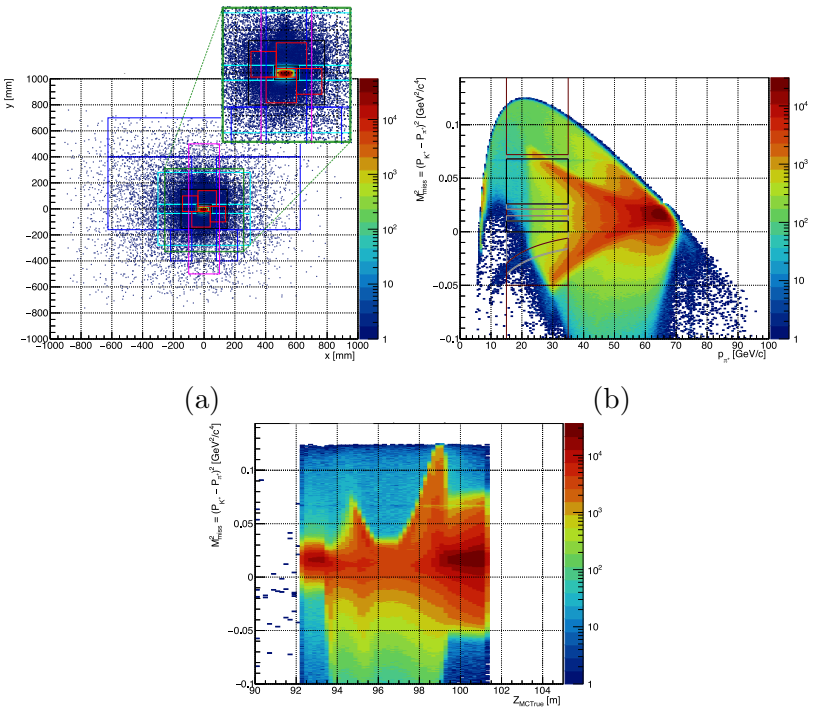


Figure C.3: Projected position of π^+ track at reference upstream plane after the final collimator (top, a and b) and difference between this position and the true position of the π^+ when it passed through the same plane (bottom, c and d). Distributions are shown for cases without (left, a and c) and with (right, b and d) the z_{vtx} versus STRAW1 radius selection condition (see figure C.1).



(a) (b)

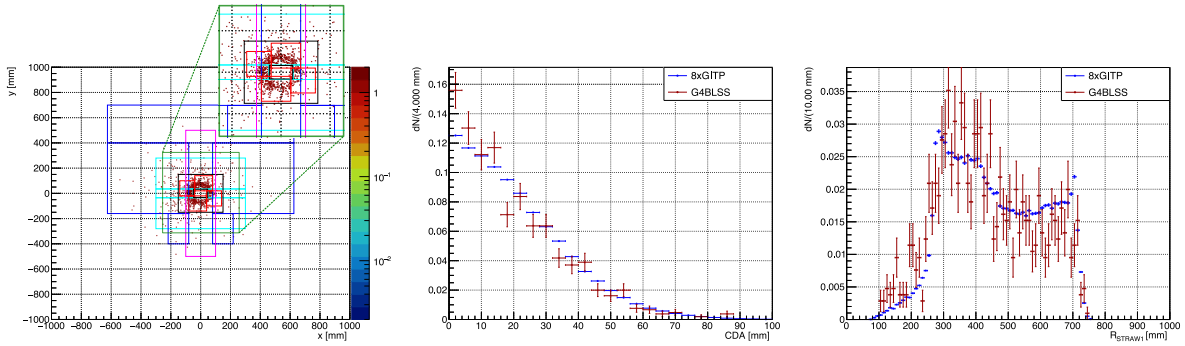
Figure C.4: Squared missing mass versus π^+ momentum (left, a) and true K^+ decay z position (right, b) for simulated upstream $K^+ \to \pi^+[\text{invisible}]_{\pi^0}$ decays without STRAW1 scattering. The two structures observed in the left plot correspond to decays in the magnetic field of BEND5 and BEND6 for the upper and lower structures respectively, with the former seen clearly on the right hand plot.



(a) (b)

(c)

Figure C.5: Selected distributions from upstream decays (snakes) simulations without beam background rejection criteria enforced. The (x, y) projected position at the final collimator exit of the π^+ is shown in (a), with the corresponding m^2_{miss} versus momentum (b) and m^2_{miss} versus true decay z position (c). The two large structures, distributed in m^2_{miss}, observed in (b) and (c) are seen, in (c), to originate around the final dipole BEND 5 and 6 magnets (located at $93.36 < z_{B5a} < 95.86$ m and $96.96 < Z_{B6} < 99.46$ m respectively) from decays in the magnetic fields.



(a) (b) (c)

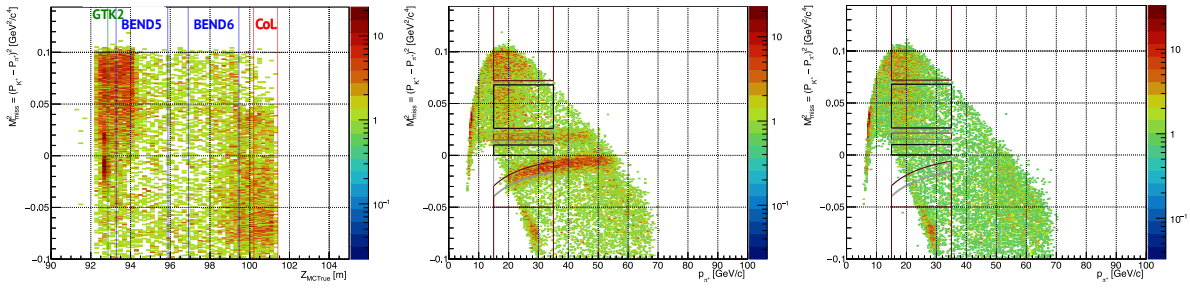
Figure C.6: (a) Projected positions of π^+ at the final collimator exit for selected events from simulations using Geant4beamline for the region upstream of the Cedar/KTAG and for injecting pileup tracks, to be compared to figure C.3c using the standard simulations strategy. Direct comparisons between standard simulations, injecting 8 pileup tracks (GITPx8) and the Geant4beamline simulations (G4BLSS), are shown for the CDA (b) and radius of π^+ tracks at STRAW1 (c).

C.5 Pileup Treatment and Comparison with Geant4 BeamLine treatment

An alternative simulation strategy was trialled for this analysis using Geant4beamline (G4BL) [268] simulations which include a description of the beamline preceding the Cedar/KTAG starting from simulation of interactions of protons in the target and tracking of all particles to the Cedar entrance plane. To replicate the snakes scenario of upstream background beam K^+ from G4BL are forced to decay as $K^+ \to \pi^+[\text{invisible}]_{\pi^0}$ in the upstream region (between GTK2 and GTK3) with fast MC simulation strategy for downstream detectors. Any particles produced concurrently during beam interactions upstream are tracked by G4BL and injected into the simulation providing associated upstream pileup tracks. Comparison with the standard simulation strategy, injecting in-time tracks from the `UpstreamPileupGenerator` library (with 8 iterations, labelled 8xGITP) and validated against data, demonstrated that this G4BL 'snakes scenario' (labelled G4BLSS) simulation was able to reproduce the basic event properties. See figure C.6. A limited G4BL simulation sample size was available for these tests but this was a powerful cross-check of the alternative simulation strategy for the upstream region since events of these kind are highly sensitive to upstream material and magnetic field simulation.

C.6 Decays Upstream of GTK2

Simulated K^+ decays upstream of the GTK2 are observed to pass the enriched upstream selection, as shown in figure C.7a. These decays are responsible for two features observed in the m^2_{miss} versus momentum distribution shown in figure C.7b: a horizontal band at $m^2_{miss} \approx m^2_{\pi^0}$ and a curved band which follows the curve expected for μ^+ from $K^+ \to \mu^+\nu_\mu$



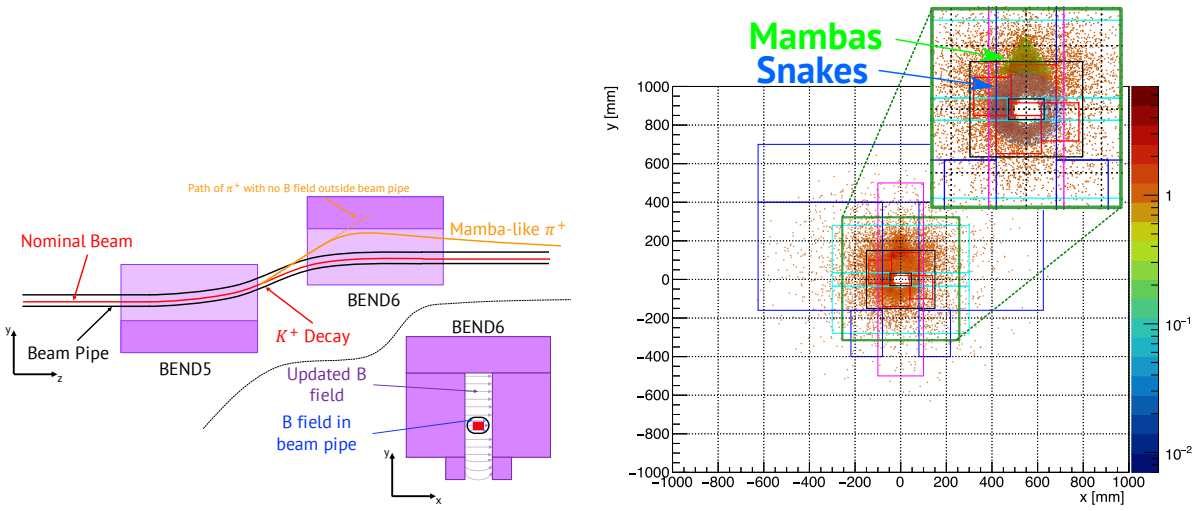
(a) (b) (c)

Figure C.7: m^2_{miss} versus simulated K^+ decay z positions (a) and momenta (b) including decays before GTK2 ($z_{GTK2} = 92.8\,\text{m}$). Decays before GTK2 lead to a horizontal band in (b) at $m^2_{miss} \approx m^2_{\pi^0}$ and curved $K^+ \to \mu^+\nu_\mu$-like band. The analogous distribution to (b) for analysis of simulations of K^+ decays after GTK2 only is shown in (c).

decays. The former represents situations where, since the π^+ from the decay upstream of GTK2 traverses the full length of both BEND5 and BEND6, the overall effect of the pair of magnets is a parallel displacement upwards with respect to a $75\,\text{GeV}/c$ particle by an amount dependant on its momentum. This π^+ is then detected downstream, matched to a pileup GTK track from the nominal beam giving precisely the signature of a $K^+ \to \pi^+\pi^0$ decay, observed as the horizontal band. The latter curved $K^+ \to \mu^+\nu_\mu$-like band is caused by the same type of event but with a $\pi^+ \to \mu^+\nu_\mu$ decay in flight, the μ^+ is mis-tagged as a π^+ because of the fast MC emulation of PID which was not calibrated for such early decays. Because of this, and since PID rejection of μ^+ has a low inefficiency of $\mathcal{O}(10^{-7})$ (on top of the rarity of upstream backgrounds) the reasonable approximation is made that if a $\pi^+ \to \mu^+\nu_\mu$ decay in flight occurs and the μ^+ is reconstructed by the STRAWs then the event is rejected. Both features seen in figure C.7b are removed if only events with K^+ decays after GTK2 are analysed, as shown by figure C.7c.

C.7 Magnetic Field Simulations and Mambas

The two classes of upstream background arising from K^+ decays upstream are snakes and mambas. The simulation of snakes has already been demonstrated in the above sections using several simulation strategies. The mambas mechanism is slightly different since the daughter π^+ track must diverge more from the standard beam path and its projection points back to the gap in the final GTK dipole magnet BEND6. For this situation to arise the π^+ track must be bent in the BEND6 magnetic field present outside the beam pipe and in the hole in the magnet yoke. Originally this field was not simulated and only snake-like events were seen upon analysis, however once the magnetic field was added a population of mamba-like events was discovered, see figure C.8. This result demonstrates that the mambas mechanism requires π^+ to deviate significantly from the standard beam path and experience different magnetic fields.



(a) (b)

Figure C.8: (a) Illustration of the path of mamba-like π^+ upstream with and without simulated magnetic field outside the beam pipe. (b) (x, y) projection back to the collimator exit for π^+ tracks reconstructed downstream with both snakes-like and mamba-like populations highlighted. Without the inclusion of the magnetic (B) field outside the beam pipe the mambas population is not observed, see for example figure C.3b.

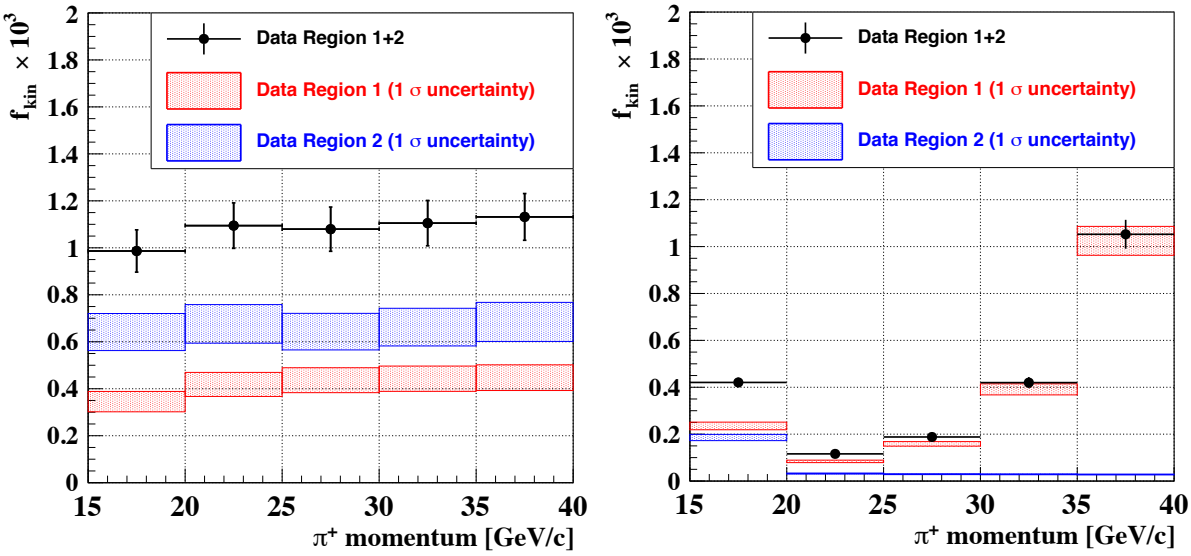
Appendix D

Details of Background Predictions from Tail Fractions

From the control samples for $K^+ \to \pi^+\pi^0$ and $K^+ \to \mu^+\nu_\mu$ decays in 2017 data the kinematic tail fractions derived, for each momentum bin, are shown in figure D.1.

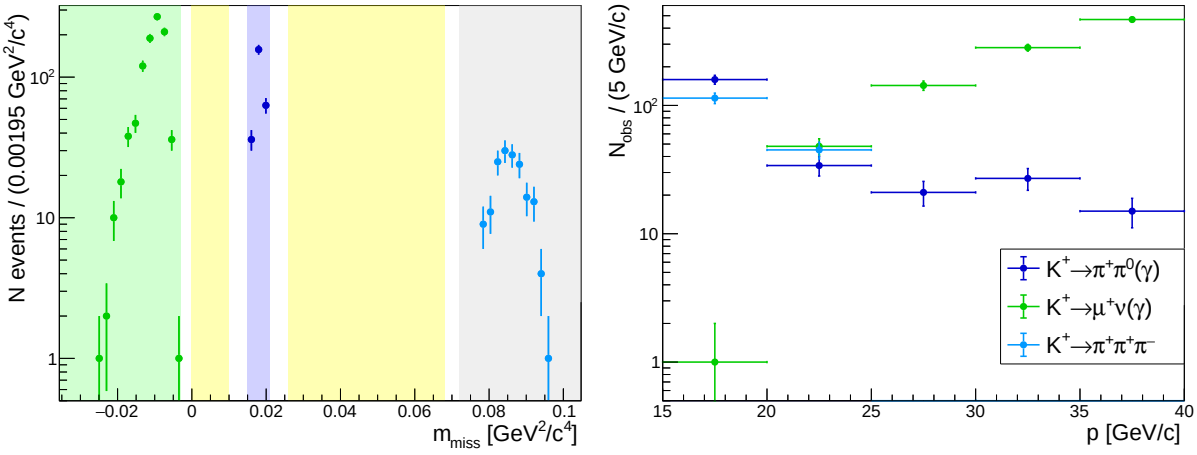
After the signal selection is applied to 2017 data the m^2_{miss} spectrum observed in the background regions ($\mu\nu R$, $\pi^+\pi^0 R$ and $3\pi R$, see figures 5.1 and 5.2), due to $K^+ \to \mu^+\nu$, $K^+ \to \pi^+\pi^0$ and $K^+ \to \pi^+\pi^+\pi^-$ decays, are shown in figure D.2a. The number of observed events in each momentum bin are shown in figure D.2b. In total there are 474, 241 and 159 events due to $K^+ \to \mu^+\nu$, $K^+ \to \pi^+\pi^0$ and $K^+ \to \pi^+\pi^+\pi^-$ decays respectively in the 15–35 GeV/c momentum range used for signal regions in 2017.

Using the tail fraction results from figures D.1b and D.1a and section 5.2, the number of observed events shown in figure D.2b, and equation 5.4.1 the expected background is calculated with results shown in figures 5.33b and 5.32b for $K^+ \to \mu^+\nu$ and $K^+ \to \pi^+\pi^0$ decays for each momentum bin. Total expected background results, integrated over all momentum bins, are given for the analysis of 2017 data by equations 5.4.5, 5.4.3 and 5.2.8 for $K^+ \to \mu^+\nu$, $K^+ \to \pi^+\pi^0$ and $K^+ \to \pi^+\pi^+\pi^-$ decays respectively. The same method was followed for the 2016 analysis.



(a) (b)

Figure D.1: [125] Measured kinematic tail fractions for signal regions 1 (blue), 2 (red) and total (points) for (a) $K^+ \to \pi^+\pi^0$ and (b) $K^+ \to \mu^+\nu_\mu$ decays in 2017 data.



(a) (b)

Figure D.2: (a) Squared missing mass spectrum for selected events in 2017 data in the regions $\mu\nu R$ (green), $2\pi R$ (blue) and $3\pi R$ (light blue) attributed to be due approximately exclusively to $K^+ \to \mu^+\nu$, $K^+ \to \pi^+\pi^0$ and $K^+ \to \pi^+\pi^+\pi^-$ decays (indicated by corresponding colours of data points). Signal regions are indicated by the yellow shaded regions. (b) Observed number of events from (a) in each momentum bin.

Appendix E

Time Resolution Measurements

The time resolution for each data-taking period for the CHOD, NA48-CHOD, RICH and STRAW Spectrometer (full track time and leading edge time) have been measured with respect to the trigger with results in table E.1. Improvements in CHOD reconstruction lead to an improvement in resolution after 2016. Representative plots are shown in figures E.1–E.5 for the weighted mean of the full 17+18 sample (as used for the search for $K^+ \to \pi\mu e$ decays reported in chapter 7) showing the $t_{detector} - t_{trig}$ distribution and the Gaussian fit used to determine the resolution. The precision of these measurements is limited by reasonable choice of fit ranges and choice of a simple Gaussian function for fitting and is estimated to be of order 5%, with this assumption figure E.6 summarises the measurements of time resolution versus data-taking period.

Table E.1: Track time resolutions for detectors for each data-taking period.

Period	σ_t(NA48-CHOD) [ns]	σ_t(CHOD) [ns]	σ_t(RICH) [ns]	σ_t(Spectrometer) [ns]	σ_t(Spectrometer Leading) [ns]
16A	0.261	1.01	0.161	5.32	0.901
17D	0.279	0.741	0.162	5.39	0.906
17C	0.275	0.764	0.163	5.39	0.936
17B	0.275	0.771	0.162	5.40	0.943
17A	0.270	0.774	0.157	5.43	0.960
18A	0.271	0.800	0.188	5.43	0.960
18B	0.270	0.819	0.181	5.43	0.952
18C	0.263	0.834	0.176	5.45	0.968
18D	0.265	0.840	0.177	5.43	0.966
18E	0.266	0.850	0.175	5.44	0.964
18F	0.267	0.847	0.171	5.48	0.975
18G	0.272	0.850	0.168	5.48	0.975
18H	0.280	0.853	0.168	5.43	0.994
17+18 weighted mean	0.271	0.809	0.170	5.43	0.958

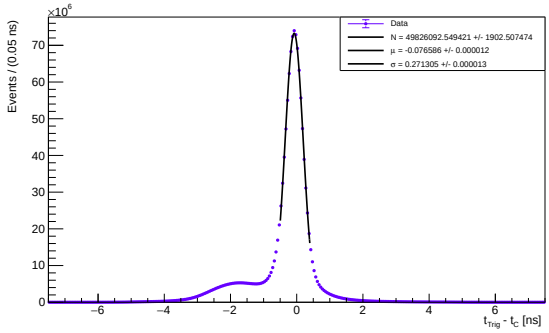


Figure E.1: Distribution of NA48-CHOD time (matched to tracks) minus trigger time with Gaussian fit used to measure time resolution.

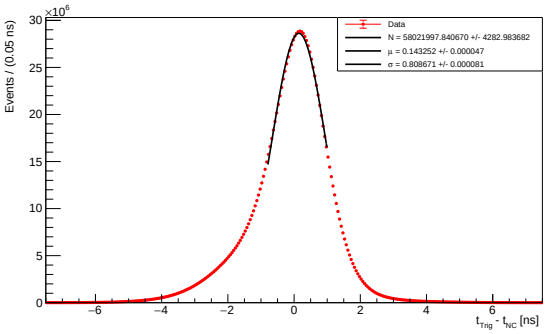


Figure E.2: Distribution of CHOD time (matched to tracks) minus trigger time with Gaussian fit used to measure time resolution.

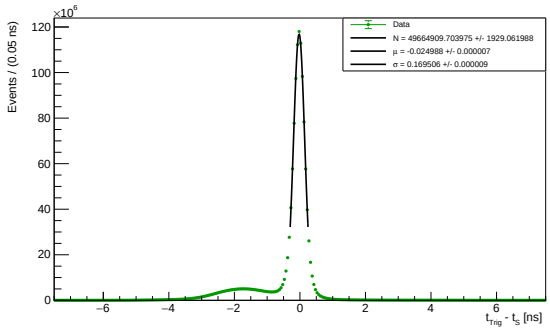


Figure E.3: Distribution of RICH time (matched to tracks) minus trigger time with Gaussian fit used to measure time resolution.

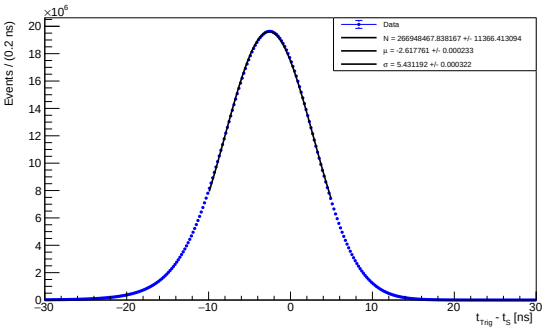


Figure E.4: Distribution of spectrometer track time minus trigger time with Gaussian fit used to measure time resolution.

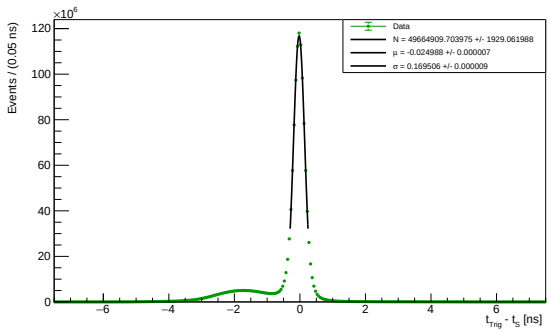


Figure E.5: Distribution of spectrometer leading time (for tracks) minus trigger time with Gaussian fit used to measure time resolution.

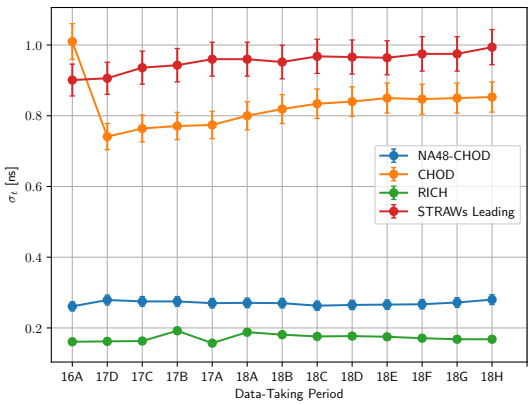


Figure E.6: Time resolution for detectors with respect to trigger time as a function of data-taking period.

Appendix F

Effective Downscaling Factor For $K^+ \to \pi\mu e$ Search

F.1 Ideal Case: 100% Efficient Triggers

Each trigger (labelled X) has a downscaling factor D_X at level 0. This means that only every D_X events that pass the L0 trigger requirements are recorded. The probability of an event being stored, given that it has properties that fulfil the trigger requirements, is $P_{trig,X} = \frac{1}{D_X}$, the inverse of the downscaling factor. Conversely, the probability that such an event is not recorded is $1 - P_{trig,X} = 1 - \frac{1}{D_X}$. If one combines 3 trigger streams (now using labels MT, MTμ and MTe to label the three triggers used in the $K^+ \to \pi\mu e$ search) then the probability of an event not being recorded by any trigger is $1 - P_{trig,MT}P_{trig,MT\mu}P_{trig,MTe}$. One can then define the effective downscaling factor, D_{eff}, as the inverse of this quantity giving

$$D_{eff} = \frac{1}{1 - \left(1 - \frac{1}{D_{MT}}\right)\left(1 - \frac{1}{D_{MT\mu}}\right)\left(1 - \frac{1}{D_{MTe}}\right)}, \tag{F.1.1}$$

where the downscaling factors for the MT, MTμ and MTe triggers are D_{MT}, $D_{MT\mu}$ and D_{MTe} respectively.

F.2 General Case: Imperfect Triggers

In the general case the probability of an event with properties that fulfil the trigger requirements (for trigger X) being stored is $P_{trig,X} = \frac{\varepsilon_X}{D_X}$. Here ε_X is the efficiency of the trigger - the probability that a trigger signal is recorded given that the event has properties which fulfil the trigger requirements.

In this case the probability of an event not being recorded is the same, $1 - P_{trig,MT}P_{trig,MT\mu}P_{trig,MTe}$, while the individual probabilities are modified, as shown above, by the trigger efficiencies. This means that the probability of an event being

recorded is given by $P = \frac{\varepsilon_{eff}}{D_{eff}}$ where the combination of effective trigger downsacling and efficiency becomes

$$\frac{D_{eff}}{\varepsilon_{eff}} = \frac{1}{1-\left(1-\frac{\varepsilon_{MT}}{D_{MT}}\right)\left(1-\frac{\varepsilon_{MT\mu}}{D_{MT\mu}}\right)\left(1-\frac{\varepsilon_{MTe}}{D_{MTe}}\right)} = \frac{(\varepsilon_{eff})^{-1}}{1-\left(1-\frac{1}{D_{MT}}\right)\left(1-\frac{1}{D_{MT\mu}}\right)\left(1-\frac{1}{D_{MTe}}\right)}. \tag{F.2.1}$$

where ε_{eff} is the effective signal trigger efficiency defined by equation 7.3.1. This ratio is included in the definition of the single event sensitivity, equation 7.4.6, accounting for the effective downscaling and trigger efficiency effects arising from combining data from three trigger streams for the search for $K^+ \to \pi\mu e$ decays. Equation 7.4.6 can be expanded to give

$$\mathcal{B}_{SES} = \frac{A_{K_{3\pi}}}{A_s}\frac{\mathcal{B}_{K_{3\pi}}}{N_{K_{3\pi}}}\frac{D_{MT}}{\varepsilon_{MT}^{K_{3\pi}}}\left[1-\left(1-\frac{\varepsilon_{MT}^s}{D_{MT}}\right)\left(1-\frac{\varepsilon_{MT\mu}^s}{D_{MT\mu}}\right)\left(1-\frac{\varepsilon_{MTe}^s}{D_{MTe}}\right)\right]^{-1}, \tag{F.2.2}$$

$$\mathcal{B}_{SES} = \frac{A_{K_{3\pi}}}{A_s}\frac{\mathcal{B}_{K_{3\pi}}}{N_{K_{3\pi}}}\frac{D_{MT}}{\varepsilon_{MT}^{K_{3\pi}}}\left[\underbrace{\frac{\varepsilon_{MT}^s}{D_{MT}}+\frac{\varepsilon_{MT\mu}^s}{D_{MT\mu}}+\frac{\varepsilon_{MTe}^s}{D_{MTe}}}_{\text{event passes 1 trigger}}\underbrace{-\frac{\varepsilon_{MT}^s\varepsilon_{MT\mu}^s}{D_{MT}D_{MT\mu}}-\frac{\varepsilon_{MT}^s\varepsilon_{MTe}^s}{D_{MT}D_{MTe}}-\frac{\varepsilon_{MT\mu}^s\varepsilon_{MTe}^s}{D_{MT\mu}D_{MTe}}}_{\text{passes 2 triggers}}+\underbrace{\frac{\varepsilon_{MT}^s\varepsilon_{MT\mu}^s\varepsilon_{MTe}^s}{D_{MT}D_{MT\mu}D_{MTe}}}_{\text{passes all 3 triggers}}\right]^{-1}, \tag{F.2.3}$$

where terms arising from signal events passing one, two or three triggers simultaneously are highlighted.

Appendix G

PID study control sample selections

G.1 $K^+ \to \pi^+\pi^+\pi^-$ Control Sample Selection

1. Exactly three tracks in the event.
2. Exactly one three-track vertex reconstructed in the event.
3. Vertex charge = +1.
4. Vertex has good quality with $\chi^2_{vtx} < 30$.
5. All 3 tracks are inside geometric acceptance of: all 4 STRAWs, NA48-CHOD, CHOD, LKr, miss the LAV and inside MUV3 acceptance.
6. An NA48-CHOD and/or CHOD association exists for each track.
7. The vertex time (constructed from the weighted mean of the three track times which are individually calculated from a weighted mean of the NA48-CHOD and CHOD times with weights corresponding to the measured time resolution for that detector for the relevant data-taking period) must be consistent with the trigger time: $|t_{vtx} - t_{trig}| < 5\,\text{ns}$.
8. Each track time must be consistent with the vertex time: $|t_{trk,i} - t_{vtx}| < 5\,\text{ns}$.
9. There is at least one KTAG candidate with at least five-sector coincidence.
10. The KTAG time best matching the vertex time satisfies: $|t_{KTAG} - t_{vtx}| < 3\,\text{ns}$.
11. NA48-CHOD shower-like events are rejected.
12. The total resultant three-track momentum must be consistent with the beam average (taken from the BeamParameters database): $|p_{3trk} - p_K^{BP}| < 2.5\,\text{GeV}/c$.
13. The transverse momentum of the three-track final state with respect to the beam average momentum must be small: $p_T < 0.030\,\text{GeV}/c$.

14. The three-track vertex must be formed inside the standard decay volume: $105 < Z_{vtx} < 180\,\text{m}$.

15. The invariant mass, in the 3π hypothesis, must be consistent with the K^+ mass: $|M_{3\pi} - m_{K^+}| < 3.5\,\text{MeV}/c^2$.

16. Additionally if any track has a MUV3 association (both inner and outer tiles) within 5 ns of the track time, then the track is tagged as being a decay in flight ($\pi^\pm \to \mu^\pm\nu_\mu$). For $\pi^\pm \Rightarrow e^\pm$ misidentification studies events must have zero MUV3 associations (no decays in flight).

The selection has an acceptance of $(11.0610 \pm 0.0027)\%$ and the selected $\pi^\pm$ sample has a contamination of less than 10^{-4} for all momenta, as shown by figure 7.19a.

G.2 $K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$ Control Sample Selection

1. Exactly three tracks in the event.

2. Exactly one three-track vertex reconstructed in the event.

3. Vertex charge = +1.

4. Vertex has good quality with $\chi^2_{vtx} < 30$.

5. All 3 tracks are inside geometric acceptance of: all 4 STRAWs, NA48-CHOD, CHOD, LKr, miss the LAV and inside MUV3 acceptance.

6. Exactly four in-time ($|t_{cluster} - t_{track}| < 5\,\text{ns}$) LKr clusters ($\pi^+e^+e^-\gamma$), three ($\pi^+e^+e^-$) within 100 mm from the projected impact point of the three tracks on the LKr front plane and exactly 1 outside these three circles.

7. An NA48-CHOD and/or CHOD association exists for each track.

8. There is at least one KTAG candidate with at least five-sector coincidence.

9. The KTAG time best matching the vertex time (constructed from the weighted mean of the three track times which are individually calculated from a weighted mean of the NA48-CHOD and CHOD times with weights corresponding to the measured time resolution for that detector for the relevant data-taking period) satisfies: $|t_{KTAG} - t_{vtx}| < 3\,\text{ns}$.

10. Each track time must be consistent with the vertex time: $|t_{trk,i} - t_{vtx}| < 5\,\text{ns}$.

11. The KTAG time best matching the vertex time satisfies: $|t_{KTAG} - t_{vtx}| < 3\,\text{ns}$.

12. The vertex time must be consistent with the trigger time: $|t_{vtx} - t_{trig}| < 5\,\text{ns}$.

13. The invariant mass of the $e^+e^-\gamma$ candidates (using energy from LKr cluster to determine the photon momentum $p_\gamma = E_\gamma$) must be consistent with the π^0 mass, $|M_{e^+e^-\gamma} - m_{\pi^0}| < 13.1\,\mathrm{MeV}/c^2$ (where the window is set as three times the resolution on the measured peak from MC) and the $\pi^+e^+e^-\gamma$ candidates invariant mass must also be consistent with the K^+ mass $|M_{\pi^+e^+e^-\gamma} - m_{K^+}| < 6.7\,\mathrm{MeV}/c^2$ (with the window similarly chosen). This must be true in one of the two possible assignments of positive tracks as π^+ and e^+.

14. The squared missing mass of the $e^+e^-\gamma$ candidates must be consistent with the squared π^0 mass $M^2_{miss}(e^+e^-\gamma) - m^2_{\pi^0}| < 8670\,\mathrm{MeV}^2/c^4$ (again with the window chosen from the measured resolution in MC).

15. The three-track vertex must be formed inside the standard decay volume: $105 < Z_{vtx} < 180\,\mathrm{m}$.

16. Using the above information the π^+ and e^+ tracks are identified and then the momentum and direction of the γ is calculated and the position of impact at the LKr predicted. The actual cluster tagged as being due to a γ (because it is over 100 mm away from the impact point of charged tracks at the LKr) must be within 150 mm of the predicted position.

17. Additionally if any track has a MUV3 association (both inner and outer tiles) within 5 ns of the track time, then the track is tagged as being a decay in flight ($\pi^\pm \to \mu^\pm\nu_\mu$). For $\pi^\pm \Rightarrow e^\pm$ misidentification studies events must have zero MUV3 associations (no decays in flight).

The selection has an acceptance of $(0.5882 \pm 0.0011)\%$ and the selected $e^\pm$ sample has a contamination of less than 10^{-4} for all momenta, as shown by figure 7.19b.

G.2.1 $K^+ \to \mu^+\nu_\mu$ Control Sample Selection

- Require exactly one reconstructed STRAW spectrometer track.
- The track has charge +1.
- The track must have good quality: $\chi^2_{track} < 40$.
- The track is inside the geometric acceptance of all four STRAW chambers, the LKr, MUV3 and does not traverse LAV12 or IRC material.
- A K^+ candidate must be tagged upstream by the KTAG, with coincident signals in at least five sectors. A KTAG time, t_{KTAG} is defined from the measured candidate K^+ best matching the downstream track time.
- The spectrometer track downstream (μ^+ candidate) is matched to a K^+ candidate track upstream using a standard tool developed by the author. It is required that a match is made with the closest distance of approach (CDA) between the upstream K^+ and downstream μ^+ candidate tracks satisfying $CDA < 7\,\mathrm{mm}$. The K–μ vertex

is further required to be be formed at $105 < z_{vtx} < 180\,\text{m}$. The best matching GTK candidate is used to define a GTK time t_{GTK}.

- The muon track can be matched to 0 or 1 in-time (within $\pm 5\,textns$ of trigger time) LKr clusters within 100 mm of the muon track position at LKr but events with any other in-time cluster (more than 100 mm from the muon track position at LKr) are rejected.
- An NA48-CHOD and/or CHOD association is required for the candidate μ track, the associated NA48-CHOD and CHOD time are used to calculate a track time from the average time weighted by resolutions of the detectors (see appendix E).
- Timing consistency checks are applied requiring:
  - $|t_{NA48-CHOD} - t_{CHOD}| < 3.0\,\text{ns}$.
  - $|t_{trk} - t_{trigger}| < 3.0\,\text{ns}$.
  - $|t_{KTAG} - t_{trk}| < 3.0\,\text{ns}$.
  - $|t_{GTK} - t_{CHOD}| < 3.0\,\text{ns}$.
- The squared missing mass, defined as

$$M^2_{miss}(K-\mu) = (P_{K^+} - P_{\mu^+})^2 \ , \tag{G.2.1}$$

  where P_{K^+} and P_{μ^+} are the 4-momenta of the K^+ and μ^+ respectively, is calculated. For a $K^+ \to \mu^+\nu_\mu$ event it is expected that $M^2_{miss}(K-\mu) \approx 0$ with some spread due to the resolution. The resolution is measured to be $\sigma_{M^2_{miss}(K-\mu)} = 3830\,\text{MeV}^2/c^4$, and it is required that the squared missing mass is within $3\sigma_{M^2_{miss}(K-\mu)}$ of $0\,\text{MeV}^2/c^4$.
- Finally, the μ^+ track momentum must be $5 < p < 70\,\text{GeV}/c$.

The selection is fully kinematic and has an acceptance of $(16.4397 \pm 0.0039)\%$. The selection isolates a sample with contamination below 5×10^{-5} for $< 50\,\text{GeV}/c$ and remaining less than 10^{-4} despite increases at higher momenta, as shown by figure 7.19c.

Appendix H

Additional $K^+ \to \pi^+\pi^+\pi^-$ Studies

H.1 $K^+ \to \pi^+\pi^+\pi^-$ Event Classes And Probabilities

There are 18 potential (physical) scenarios to consider for $K^+ \to \pi^+\pi^+\pi^-$ decays since each pion has the potential to decay in flight within the NA62 detector. As discussed in section 7.5.7.1 to good approximation there are two possible decays: $\pi^\pm \to \ell^\pm\nu_\ell$ with $\ell = \mu, e$. One can express the number of these decays in each event with variables $n_\ell^\pm$ where charge conservation requires $n_\mu^+ + n_e^+ \leq 2$ and $n_\mu^- + n_e^- \leq 1$ and overall $\sum_{\ell,\pm} n_\ell^\pm \leq 3$. The 18 possible event classes are then uniquely defined by four numbers and expressed as $(n_\mu^+ n_e^+ n_\mu^- n_e^-)$ and can be classified according to the number of decays by charge or lepton flavour as shown in tables H.1 and H.2.

Table H.1: Possible decay in flight configurations for $K^+ \to \pi^+\pi^+\pi^-$ in terms of $n^\pm$.

	$n^+ = 0$	$n^+ = 1$	$n^+ = 2$
$n^- = 0$	(0000)	(1000), (0100)	(2000), (0200), (1100)
$n^- = 1$	(0010), (0001)	(1010), (1001), (0110), (0101)	(2010), (2001), (0210), (0201), (1110), (1101)

Table H.2: Possible decay in flight configurations for $K^+ \to \pi^+\pi^+\pi^-$ in terms of n_ℓ.

	$n_\mu = 0$	$n_\mu = 1$	$n_\mu = 2$	$n_\mu = 3$
$n_e = 0$	(0000)	(1000) (0010)	(2000) (1010)	(2010)
$n_e = 1$	(0100) (0001)	(1100), (1001), (0110), (0101)	(1110) (2001)	X
$n_e = 2$	(0200) (0101)	(1101)	X	X
$n_e = 3$	(0201)	X	X	X

The probability for each of the 18 possible scenarios can be calculated given that the probability of a $\pi^{\pm}$ kaon daughter particle decaying in flight (within the volume to which the NA62 experimental apparatus is sensitive), P_{DIF}, and the branching ratios $\mathcal{B}(\pi^{\pm} \to \ell\nu_{\ell})$ are known. Under the (very good) approximation that

$$\mathcal{B}(\pi^{\pm} \to \mu^{\pm}\nu_{\mu}) + \mathcal{B}(\pi^{\pm} \to e^{\pm}\nu_{e}) = \mathcal{B}_{\mu} + \mathcal{B}_{e} = 1 \ , \tag{H.1.1}$$

one can express the probability of a scenario specified by $(n_{\mu}^{+} n_{e}^{+} n_{\mu}^{-} n_{e}^{-})$ as

$$\begin{aligned} &P(n_{\mu}^{+} n_{e}^{+} n_{\mu}^{-} n_{e}^{-} | \mathcal{B}_{e}) \\ &= \begin{cases} (P_{DIF})^{n_{DIF}} \left(\left[(1-\mathcal{B}_{e})^{n_{\mu}^{+}} (\mathcal{B}_{e})^{n_{e}^{+}} \right] \right) \left((1-\mathcal{B}_{e})^{n_{\mu}^{-}} (\mathcal{B}_{e})^{n_{e}^{-}} \right) \ , & n_{DIF} > 0, \ n^{+} = 0 \\ (P_{DIF})^{n_{DIF}} \left(2 \left[(1-\mathcal{B}_{e})^{n_{\mu}^{+}} (\mathcal{B}_{e})^{n_{e}^{+}} \right] \right) \left((1-\mathcal{B}_{e})^{n_{\mu}^{-}} (\mathcal{B}_{e})^{n_{e}^{-}} \right) \ , & n_{DIF} > 0, \ n^{+} > 0 \\ 1 - \sum_{i=1}^{17} P(i) \ , & n_{DIF} = 0 \ , \end{cases} \end{aligned} \tag{H.1.2}$$

where i represents the $18-1$ cases $(n_{\mu}^{+} n_{e}^{+} n_{\mu}^{-} n_{e}^{-})$ in which $N_{DIF} > 0$. The number of events of each type expected in a MC sample of N_{MC} events is therefore

$$\begin{aligned} &N(n_{\mu}^{+} n_{e}^{+} n_{\mu}^{-} n_{e}^{-} | \mathcal{B}_{e}) = N_{MC} P(n_{\mu}^{+} n_{e}^{+} n_{\mu}^{-} n_{e}^{-} | \mathcal{B}_{e}) \\ &= \begin{cases} N_{MC} (P_{DIF})^{n_{DIF}} \left(\left[(1-\mathcal{B}_{e})^{n_{\mu}^{+}} (\mathcal{B}_{e})^{n_{e}^{+}} \right] \right) \left((1-\mathcal{B}_{e})^{n_{\mu}^{-}} (\mathcal{B}_{e})^{n_{e}^{-}} \right) \ , & n_{DIF} > 0, \ n^{+} = 0 \\ N_{MC} (P_{DIF})^{n_{DIF}} \left(2 \left[(1-\mathcal{B}_{e})^{n_{\mu}^{+}} (\mathcal{B}_{e})^{n_{e}^{+}} \right] \right) \left((1-\mathcal{B}_{e})^{n_{\mu}^{-}} (\mathcal{B}_{e})^{n_{e}^{-}} \right) \ , & n_{DIF} > 0, \ n^{+} > 0 \\ N_{MC} \left(1 - \sum_{i=1}^{17} P(i) \right) \ , & n_{DIF} = 0 \ . \end{cases} \end{aligned} \tag{H.1.3}$$

If the branching ratio for $\pi^{\pm} \to e^{\pm}\nu_{e}$ decays is modified in MC simulations (see section 7.5.7.1) then the number of events of each type will change. The number of events of each type observed in a MC sample, $N^{MC,obs}_{n_{\mu}^{+} n_{e}^{+} n_{\mu}^{-} n_{e}^{-}}$, allows a measurement of the probability for each case as

$$P(n_{\mu}^{+} n_{e}^{+} n_{\mu}^{-} n_{e}^{-} | \mathcal{B}_{e})_{MC,obs} = \frac{N^{MC,obs}_{n_{\mu}^{+} n_{e}^{+} n_{\mu}^{-} n_{e}^{-}}}{N_{MC}} \ . \tag{H.1.4}$$

The probability of a pion decaying in flight for $K^{+} \to \pi^{+}\pi^{+}\pi^{-}$ events can be measured by calculating

$$P_{DIF} = \frac{1}{3} \frac{\sum_{i=0}^{17} n_{DIF}(i) N(i)}{N_{MC}} \ , \tag{H.1.5}$$

where $n_{DIF}(i)$ is the number of decays in flight for event class i (see table 7.12) and $N(i)$ is the number of simulated events observed of that event class. The multiplicative factor $\frac{1}{3}$ accounts for the 3 pions in each event. For the MC simulation sample studied the result is $P_{DIF} = 7.53\%$.

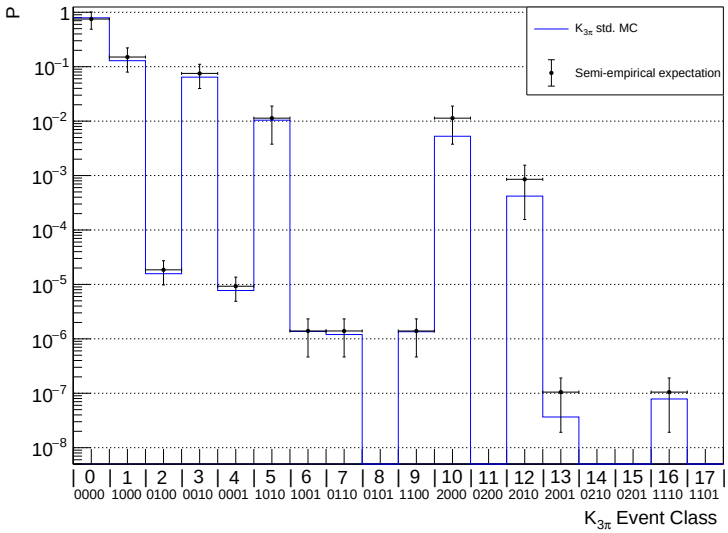


Figure H.1: Probability of $K^+ \to \pi^+\pi^+\pi^-$ of each event class (see table 7.12) for standard MC compared to the semi-empirical prediction.

Given equation H.1.5 and H.1.3 a 'semi-empirical' prediction[1] can be calculated for the expected number of each type of event class in a MC simulation sample with N_{MC} events and a specific configuration of π decay branching ratios satisfying H.1.1. Applying this procedure to standard $K^+ \to \pi^+\pi^+\pi^-$ simulations, and simulations with $\mathcal{B}_e = 0.5$ or $\mathcal{B}_e = 1$, good agreement is obtained between observed and 'semi-empirical' predicted probabilities of each event class, see figures H.1, H.2a and H.2b respectively. This demonstrates a good understanding of the simulations, the probability of each event class and therefore the event weights that should be applied to simulated events in the special $K^+ \to \pi^+\pi^+\pi^-$ samples.

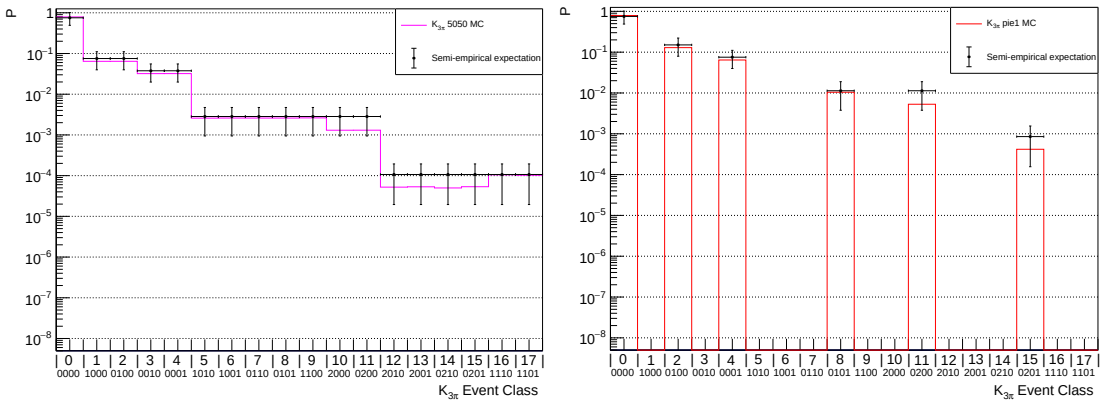
The definition of each of each event class, how they can lead to a background within the search for $K^+ \to \pi\mu e$ decays and the measured probabilities of each from simulations is given in table 7.12.

H.2 Enhancement Factors For Dedicated $K^+ \to \pi^+\pi^+\pi^-$ Background Studies

By modifying the branching ratios $\mathcal{B}_e$ and $\mathcal{B}_\mu$ the relative probabilities of obtaining different final state particle content, through the 18 possible mechanisms discussed above, can be modified. When studying specific background mechanisms for the $K^+ \to \pi\mu e$ searches arising from $K^+ \to \pi^+\pi^+\pi^-$ decays it is crucial to be able to enhance certain rare mechanisms for background estimates with sufficient precision to be determined.

If $\mathcal{B}_e$ is increased then the number of $\pi^\pm \to e^\pm\nu_e$ decays increases and this simulation enhances this process. However, since $\mathcal{B}_\mu = 1 - \mathcal{B}_e$, the number of $\pi^\pm \to \mu^\pm\nu_\mu$ is decreased

[1] This is considered 'semi-empirical' because the probability of decay is evaluated with a simulated sample of $K^+ \to \pi^+\pi^+\pi^-$ decays.



(a) (b)

Figure H.2: Probability of $K^+ \to \pi^+\pi^+\pi^-$ of each event class (see table 7.12) for MC with (a) $\mathcal{B}_e = 0.5$ and (b) $\mathcal{B}_e = 1$ compared to the semi-empirical prediction.

meaning such scenarios are suppressed in these simulations. An overall 'enhancement factor', $f_{n_\mu^+ n_e^+ n_\mu^- n_e^-}$, for each scenario can be calculated as a function of the chosen $\mathcal{B}_e$ as

$$f_{n_\mu^+ n_e^+ n_\mu^- n_e^-} = \frac{P(n_\mu^+ n_e^+ n_\mu^- n_e^- | \mathcal{B}_e)}{P(n_\mu^+ n_e^+ n_\mu^- n_e^-)_{SM}} \,, \tag{H.2.1}$$

where in the denominator the standard model branching ratio [13] $\mathcal{B}_e = 1.230 \times 10^{-4}$ is used ($P(n_\mu^+ n_e^+ n_\mu^- n_e^-)_{SM} = P(n_\mu^+ n_e^+ n_\mu^- n_e^- | 1.230 \times 10^{-4})$) and the numerator value may be chosen. Figure H.3 shows the enhancement factors for each event class as a function of $\mathcal{B}_e$. For events with 0 decays in flight there is never any enhancement, thus $f_{0000} = 1$ for any $\mathcal{B}_e$. Cases with only $\pi^\pm \to \mu^\pm \nu_\mu$ decays are always suppressed $f_{n_\mu^+ n_e^+ n_\mu^- n_e^-} < 1$, becoming more strongly suppressed as $\mathcal{B}_e \to 1$. Conversely all cases with only $\pi^\pm \to e^\pm \nu_e$ decays are enhanced $f_{n_\mu^+ n_e^+ n_\mu^- n_e^-} > 1$, rising steeply at low $\mathcal{B}_e$ but tending to plateau as $\mathcal{B}_e \to 1$. However, most interesting are cases with a mixture of decay types, here the enhancement factor varies with $\mathcal{B}_e$ in a non-monotonic way, first rising at low $\mathcal{B}_e$ before falling again as $\mathcal{B}_e \to 1$. In these cases the behaviour is being driven by the convolution of the two cases discussed above. To see more clearly the trends considered here figure H.4 shows a set of 'averaged' curves which take up the described situations.

It is further instructive to view the 'normalised' enhancement factor curves displayed in figure H.5. Here each curve is normalised such that its maximum point is equal to unity so the trend with $\mathcal{B}_e$ becomes apparent for each case using a linear scale. Combining information of the trends and magnitude of the enhancement effect for each case (using figures H.5 and H.3 respectively) it becomes apparent that using $\mathcal{B}_e = 0.5$ is an effective and efficient choice. By moving to larger values the suppression of $\pi^\pm \to \mu^\pm \nu_\mu$ decays becomes increasingly significant and the mixed cases (which are of significant importance for this study) tend to plateau before starting to decrease. By generating at, for example $\mathcal{B}_e = 0.66$ some small gains may be made is some cases however the most general and effective strategy is to use a single value for a relatively large simulation sample. For these reasons the value $\mathcal{B}_e = 0.5$ is adopted in general for the $K^+ \to \pi\mu e$ analysis and can be used to study background mechanisms C and D (defined in figure 7.30). Further

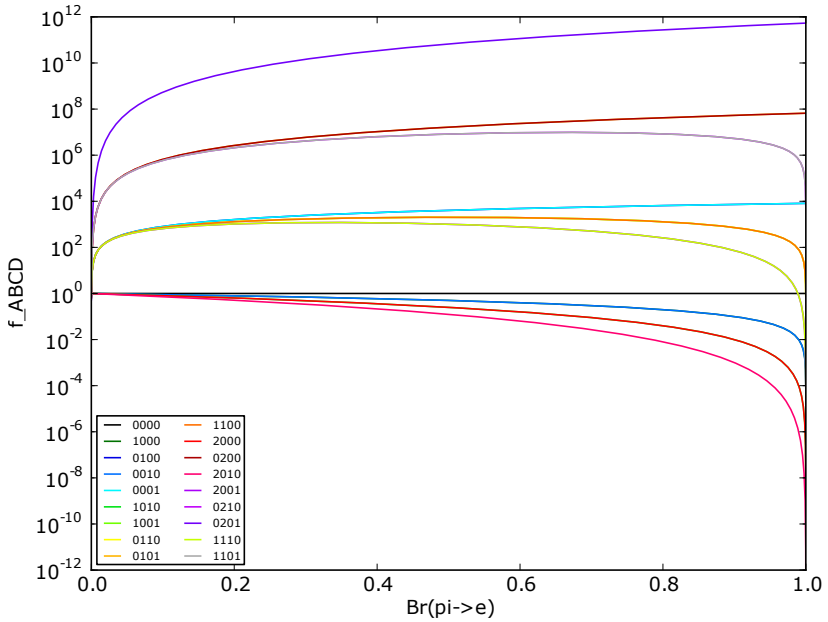


Figure H.3: Enhancement factors for each of the 18 event classes (see table 7.12) as a function of the modified $\pi^{\pm} \to e^{\pm}\nu_e$ branching ratio.

enhancement is achieved for study of the specific case with exactly one $\pi^+ \to e^+\nu_e$ is achieved using $\mathcal{B}_e = 1$, and this is used for study of background mechanism D (defined in figure 7.30).

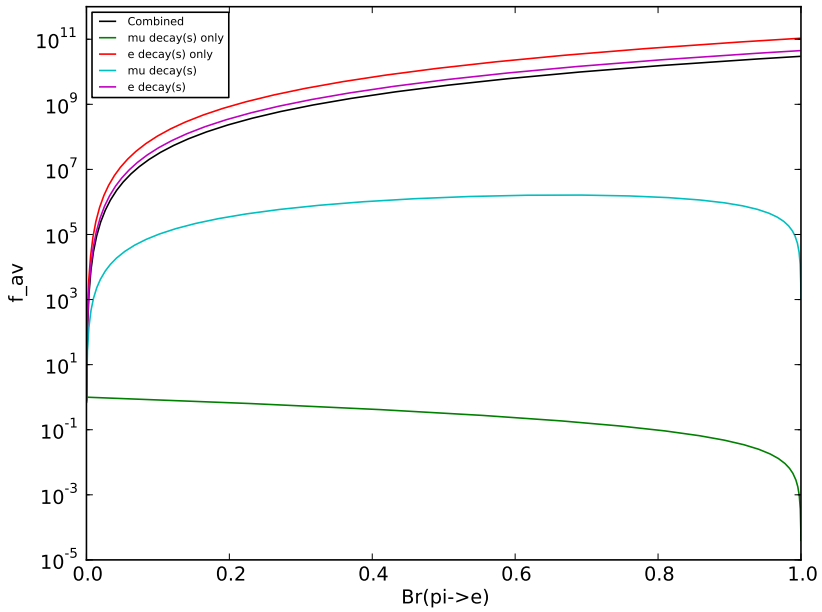


Figure H.4: Averaged enhancement factors for sets of event classes (see text) as a function of the modified $\pi^{\pm} \to e^{\pm}\nu_e$ branching ratio.

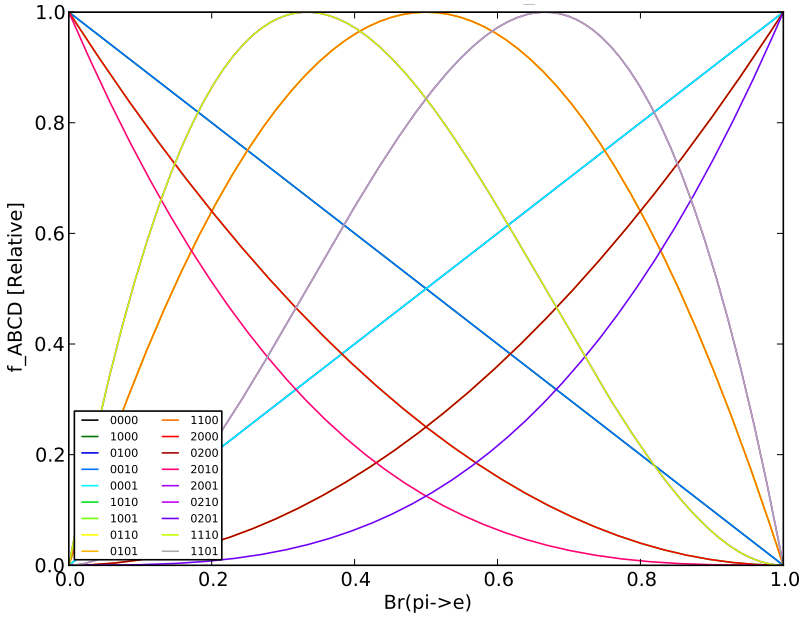


Figure H.5: Normalised enhancement factors for each of the 18 event classes (see table 7.12) as a function of the modified $\pi^{\pm} \to e^{\pm}\nu_e$ branching ratio. Each case is normalised such that the maximum is equal to unity. It should be noted that several curves are overlaid due to following the same trend when normalised.

Appendix I

Setting Upper Limits on $\mathcal{B}_S$ With the CLs Method

The CLs method [240] [243] [244] is widely used in the statistical interpretation of particle physics analyses for setting upper limits in searches, especially in low-statistics scenarios where it will not exclude regions of parameter space for which the search has little or no sensitivity [269].

For the $K^+ \to \pi\mu e$ searches the parameter of interest is the branching ratio of the decay, $\mathcal{B}_s$, (which in reality is a relative measurement with respect to the normalisation mode branching ratio). This is related to the single event sensitivity through relation

$$\mathcal{B}_s = s \times \mathcal{B}_{SES} \ , \tag{I.0.1}$$

where s is the number of signal events (see also equation 7.4.5). To simplify the statistical treatment s is considered as the parameter of interest and an upper limit set on this as s^{UL} and converted to the upper limit on the branching ratio using equation I.0.1. It is noted that the uncertainty on the single event sensitivity is then taken into account within the procedure itself (however with a fractional uncertainty of order 5% the effect is small).

The number of events observed in data, n, is a discrete number and can be considered as a Poisson distributed variable. In a search experiment one result is observed which is considered as a single trial of drawing a number from a Poisson distribution with mean equal to $s+b$, the mean number of expected events made up of the mean number of signal, s, and total background, b, events. Therefore s and b are continuous variables where b is determined for $K^+ \to \pi\mu e$ searches along with an associated uncertainty σ_b (which is considered to be a Gaussian error and therefore the background expectation is Gaussian distributed with mean b and sigma σ_b), as described in section 7.5, while in the statistical analysis a scan is performed over possible values of s. The likelihood function describing this is scenario is

$$L(s, b, n) = \frac{(s+b)^n}{n!} e^{-(s+b)} \ . \tag{I.0.2}$$

To account for the uncertainty on the single event sensitivity (which only relates to the signal and not the background, since the background uncertainty is contained in σ_b), a

factor f is included which multiplies s. This is a Gaussian distributed continuous variable with mean of one and standard deviation equal to the fractional uncertainty on the single event sensitivity ($\frac{\sigma_{\mathcal{B}_{SES}}}{\mathcal{B}_{SES}}$). This means the likelihood becomes

$$L(s, b, n, \mathcal{B}_{SES}, \sigma_{\mathcal{B}_{SES}}) = \frac{(fs+b)^n}{n!} e^{-(fs+b)} \ . \tag{I.0.3}$$

The likelihood ratio is then constructed as the ratio of the signal + background hypothesis and the background-only hypothesis, $\frac{L_{s+b}}{L_b}$, given by

$$Q = \frac{\frac{(fs+b)^n}{n!} e^{-(fs+b)}}{\frac{b^n}{n!} e^{-b}} = \left(\frac{(fs+b)}{b}\right)^n e^{-fs} \ . \tag{I.0.4}$$

Taking the natural logarithm of this gives the log likelihood ratio test statistic[1],

$$t = \ln(Q) = n \ln\left(\frac{fs+b}{b}\right) - fs \ . \tag{I.0.5}$$

So far the construction of the equations assumes that factors s, b and f are known precisely. Since s is a parameter that is scanned over in this study this can be considered to be known precisely, however the factors b and f are, as discussed above, actually subject to uncertainties. This is accounted for through a Monte Carlo approach where multiple pseudo-experiments are performed and in each one a value is chosen for b and f drawing numbers form Gaussian distributions $\mathcal{N}(b, \sigma_b)$ and $\mathcal{N}(1, \sigma_f)$ respectively. Therefore, in each pseudo-experiment there is an exact value for b and f.

For each test value of s the following procedure is performed. A set of n_T pseudo-experiments is performed, in each a random number is drawn from Gaussian distributions to set values of b and f,[2] and the value of the test statistic is calculated under three hypotheses:

1. The observed case, t_{obs}. Here $n = n_{obs}$, the observed number of events from the $K^+ \to \pi\mu e$ analysis. Since this is a blind analysis the standard procedure is to perform the study using $n_{obs} = 0, 1, 2, 3, 4$ and then report upper limits obtained in each case. Once the signal region is unblinded and the true number of observed events is known the result is already fixed.

[1] By convention the log likelihood ratio is often multiplied by a factor of -2 [13] [264] [269]. This is useful if one is looking for interval estimation or ascertaining the uncertainty on a maximum likelihood estimator, by looking at the range of the likelihood function which is below unity. If this is done then the minus sign flips the order of the signal+background and background-only hypotheses in the following discussion, so the background-only will be at higher values of the test statistic instead of lower as in the construction considered here. It is also noted that the absolute value of the likelihood and therefore the test statistic is not important in of itself, rather what matters is the distribution shape and comparison between signal+background and background-only cases. Because of this, in principle, the constant term could be removed from equation I.0.5.

[2] Note that it is required that these values are positive and if a negative value is drawn the process is repeated until a positive value is chosen. This has the effect of truncating the Gaussian distribution at 0 which can be a significant difference if, for example σ_b is a reasonable fraction of b.

2. The background-only hypothesis, t_b. Here $n = n_b$ where n_b is a discrete integer number of events drawn from a Poisson distribution of mean equal to b.

3. The signal+background hypothesis, t_{s+b}. Here $n = n_{s+b}$ where n_{s+b} is a discrete integer number of events drawn from a Poisson distribution of mean equal to $s + b$.

The CL parameters are then calculated from the fraction of pseudo-experiments fulfilling different criteria:

CL_b : $CL_b = 1 - p_b$ where p_b is the background-only p-value - the probability for the test statistic t to be as signal-like or less so [264]. In the set-up here that means the fraction of times $t_b \leq t_{obs}$.

CL_{s+b} : which is the probability that the test statistic is as background-like or more so [264]. Here this means the fraction of times $t_{s+b} \leq t_{obs}$.

From these the CL_s is calculated as

$$CL_s = \frac{CL_{s+b}}{CL_b} \ . \tag{I.0.6}$$

Therefore, for each trial value of s one value of observed CL_s is obtained for a given observed number of events n_{obs}. For a given n_{obs} a scan is performed over s and an upper limit is set, s^{UP}, at confidence level $1 - \alpha$ as the value of s for which CL_s is α (where the default is $\alpha = 0.1$ with optional use of $\alpha = 0.05$). The observed limit is presented in figure 7.55 as a collection of blue points connected by a blue line. Relevant α values are highlighted by horizontal lines, the upper limit is then read from the x-axis (s^{UL}) at the point when the observed CL_s blue points/line crosses these horizontal lines.

To set an upper limit one is trying to exclude the signal+background hypothesis at a certain level of probability. This means calculating the probability of the observed value of the test statistic being at least as incompatible or more or with the signal+background hypothesis. This would suggest using CL_{s+b} and excluding signals where $CL_{s+b} < 1 - \alpha$. However close to the limit of sensitivity, where the distribution of test statistic for the signal+background and background-only hypotheses are not well-separated, a downward fluctuation in the data for the calculation of the background expectation can result in exclusion of a signal where the analysis is not actually sensitive. The use of CL_s is a common solution to this problem. It is not technically a confidence level and results using it must be interpreted with this in mind [264].

To determine the expected upper limit and the one and two standard deviation bands, often displayed when reporting such search upper limits, one needs to consider the ensemble of pseudo-experiments, producing a set of n_T values of CL_s. The mean, upper and lower one and two standard deviation values (for a given trial s and observed n_{obs}) can be defined by the median CL_s value and the percentiles which correspond to the one and two sigma upper and lower points in a Gaussian distribution . Technically this is realised by creating a sorted list of CL_s values from the n_T pseudo-experiments and choosing the element of the list, $1 \leq n_l \leq n_T$, most closely matching the percentiles:

$$
\begin{aligned}
-2\sigma: &\quad n_T\left(\tfrac{1}{2} - \left(\tfrac{0.95448}{2}\right)\right) \\
-\sigma: &\quad n_T\left(\tfrac{1}{2} - \left(\tfrac{0.68269}{2}\right)\right) \\
\text{mean}: &\quad \tfrac{n_T}{2} \\
+\sigma: &\quad n_T\left(\tfrac{1}{2} + \left(\tfrac{0.68269}{2}\right)\right) \\
+2\sigma: &\quad n_T\left(\tfrac{1}{2} + \left(\tfrac{0.95448}{2}\right)\right)
\end{aligned}
$$

The (median) expected limit is presented in figure 7.55 as a dashed line and the one and two standard deviation envelope results are displayed as green and yellow bands respectively.

Appendix J

Combination of p Values Using Fisher's Method

For n somewhat independent tests of null hypotheses with individual p values p_i a combined p value can be calculated using Fisher's method [265] [266]. A test statistic variable, t, is constructed according to

$$t = -2 \sum_{i=1}^{n} ln(p_i) \,, \tag{J.0.1}$$

which is distributed as a χ^2 with $2n$ degrees of freedom. The combined p value is then calculated, using a Monte Carlo method, as the fraction of 10^6 simulation trials in which the value t is greater than or equal to random values drawn from a χ^2 distribution with $2n$ degrees of freedom.

Appendix K

Searches for $\pi^0 \to \mu^\mp e^\pm$

K.1 Particle Data Group Limits for $\pi^0 \to \mu^\mp e^\pm$ Searches

The Particle Data Group (PDG) reports three results for $\pi^0 \to \mu^\mp e^\pm$ searches [13]:

$$\mathcal{B}(\pi^0 \to \mu^- e^+) < 34 \times 10^{-10}\,@90\%\,CL \quad [171]\,, \tag{K.1.1}$$

$$\mathcal{B}(\pi^0 \to \mu^+ e^-) < 3.8 \times 10^{-10}\,@90\%\,CL \quad [263]\,, \tag{K.1.2}$$

$$\mathcal{B}(\pi^0 \to \mu^\mp e^\pm) < 3.6 \times 10^{-10}\,@90\%\,CL \quad [173]\,. \tag{K.1.3}$$

The first two searches for a specific charge configuration come from searches for $K^+ \to \pi^+[\mu^\pm e^\pm]_{\pi^0}$ decay chains at the BNL E865 experiment [171] [263] [172]. Similar searches are possible at NA62 and are discussed in section 7.7.4 and subsequent appendix sections K.2 and K.3. The E865 experiment was designed to search for $K^+ \to \pi^+\mu^+e^-$ decays and therefore is able to set a stronger limit on the $\pi^0 \to \mu^+e^-$ channel.

The third result was reported by the FNAL KTeV (E799-II) experiment [173] and arises from searches from the $K_L \to \pi^0\pi^0[\mu^\mp e^\pm]_{\pi^0}$ decay chain, normalised to $K_L \to \pi^0\pi^0[e^+e^-\gamma]_{\pi^0}$ decays. It is reported [173] that the search is equally sensitive to both charge combinations ($\mu^\mp e^\pm$) and the result is set on the sum of the two charge combinations giving result

$$\mathcal{B}(\pi^0 \to \mu^\mp e^\pm)_{KTeV} = \mathcal{B}(\pi^0 \to \mu^- e^+) + \mathcal{B}(\pi^0 \to \mu^+ e^-) < 3.8 \times 10^{-10}\,. \tag{K.1.4}$$

The limit on either individual charge configuration cannot be weaker than the limit on the sum. If the two charge combinations have equal contributions the limit on either is $\mathcal{B} < 1.9 \times 10^{-9}$, however there is no clear motivation for this, especially in BSM scenarios with such LFV decays.

The three results reported can therefore all be considered. The interpretation of the limits on individual charge combinations is clear and well-defined and can be directly compared to searches at NA62. The third result appears to demonstrate that neither charge combination can have a branching ratio in excess of 3.6×10^{-10} and so this can be considered the benchmark on which to improve for either specific charge combination.

K.2 Search for $\pi^0 \to \mu^- e^+$ in the Decay Chain $K^+ \to \pi^+[\mu^- e^+]_{\pi^0}$

The signature of a $\pi^0 \to \mu^- e^+$ decay could be seen in the search for K^+ decays to $\pi^+\mu^- e^+$ final states where a $K^+ \to \pi^+\pi^0$ decay is followed by a $\pi^0 \to \mu^- e^+$ decay. Candidate events would have a $M_{\mu e}$ invariant mass consistent with the π^0 mass. From simulations of the $K^+ \to \pi^+[\mu^- e^+]_{\pi^0}$ decay chain the π^0 mass resolution is determined to be $0.41 \pm 0.02_{syst}$, with systematic uncertainty arising from the variation in results possible with different Gaussian fit ranges. Informed by this, the subset of candidate K^+ decays to $\pi^+\mu^- e^+$ final states with $|M_{\mu e} - m_{\pi^0}| < 3\,\text{MeV}/c^2$ are selected as $K^+ \to \pi^+[\mu^- e^+]_{\pi^0}$ candidates for the $\pi^0 \to \mu^- e^+$ search.

The additional invariant mass requirement reduces acceptance from $(6.21 \pm 0.02)\%$ for the $K^+ \to \pi^+\mu^- e^+$ search to $(3.11 \pm 0.02)\%$ for the $K^+ \to \pi^+[\mu^- e^+]_{\pi^0}$ search, see table 7.3. The sensitivity is decreased by a further factor of approximately five due to the branching ratio $\mathcal{B}(K^+ \to \pi^+\pi^0) = (20.67 \pm 0.08)\%$ giving (see also table 7.7)

$$\mathcal{B}_{SES}(K^+ \to \pi^+[\mu^- e^+]_{\pi^0}) = (1.400 \pm 0.094) \times 10^{-10}\,. \tag{K.2.1}$$

The observed data for candidate K^+ decays to $\pi^+\mu^- e^+$ final states is shown in figure K.1a with the magenta band showing the region containing $\pi^0 \to \mu^- e^+$ candidates. A projection onto the $M_{\mu e}$ invariant mass is shown in figure K.1b with good agreement demonstrated between data and simulations.

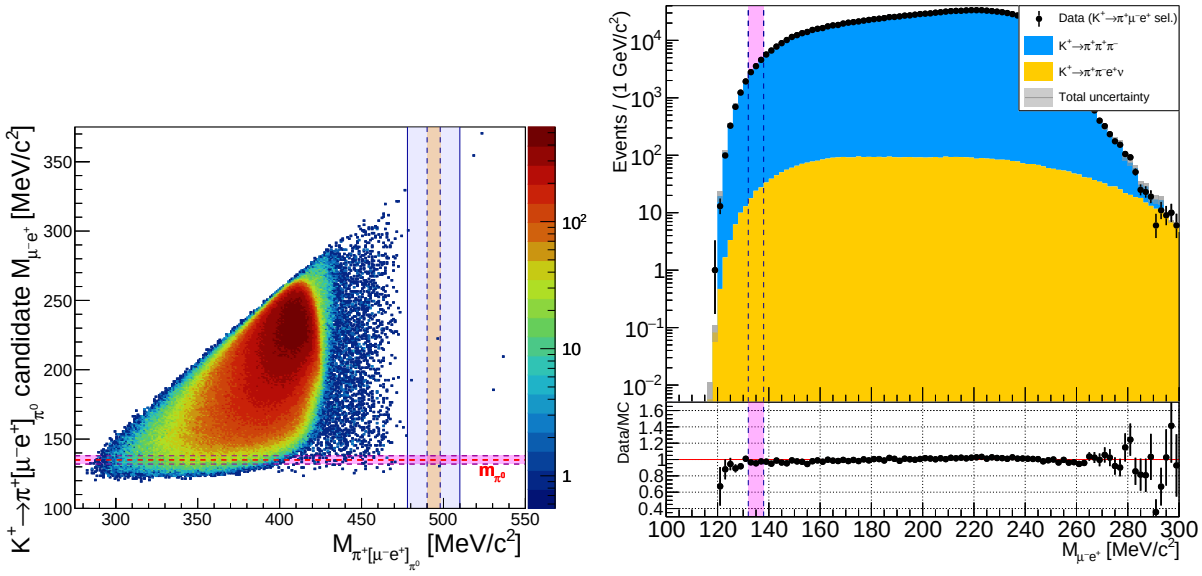
The full background expectation is presented in table K.1, however only backgrounds from $K^+ \to \pi^+\pi^+\pi^-$ and $K^+ \to \pi^+\pi^- e^+\nu_e$ decays are significant. As shown in table K.1 and figure 7.56a the background expectation and observation in the lower control region are in good agreement. The total background expectation in the signal region, as reported in section 7.7.3, is

$$n_{bg} = 0.24 \pm 0.15\,. \tag{K.2.2}$$

No candidate events are observed in the combined $K^+ \to \pi^+[\mu^- e^+]_{\pi^0}$ signal region ($|M_{\mu e} - m_{\pi^0}| < 3\,\text{MeV}/c^2$ and $490 < M_{\pi\mu e} < 498\,\text{MeV}/c^2$) in the 2017+2018 dataset. An upper limit is therefore established at 90% confidence level using the same statistical treatment as for the $K^+ \to \pi^\mp\mu^\pm e^+$ searches, as discussed in section 7.6 and appendix I, of

$$\mathcal{B}(\pi^0 \to \mu^- e^+) < 3.2 \times 10^{-10}\,. \tag{K.2.3}$$

This represents a marginal improvement by a factor of 1.13 over the KTeV combined limit of [173] $\mathcal{B}(\pi^0 \to \mu^\pm e^\pm) < 3.6 \times 10^{-10}$ but an improvement by a factor of 15 over the BNL E865 experiment result from a similar $K^+ \to \pi^+[\mu^- e^+]_{\pi^0}$ search of [171] $\mathcal{B}(\pi^0 \to \mu^- e^+) < 34 \times 10^{-10}$. See section K.1 for discussion of these two previous limits, both reported in the PDG [13].



(a) (b)

Figure K.1: Selected candidate $K^+ \to \pi^+\mu^-e^+$ events, the sub-set with $|M_{\mu e} - m_{\pi^0}| < 3\,\text{MeV}/c^2$ additionally being candidate $K^+ \to \pi^+[\mu^-e^+]_{\pi^0}$ events. The two dimensional distribution of $M_{\mu e}$ versus $M_{\pi\mu e}$ for all events is shown in (a) for data, with projection onto the $M_{\mu e}$ axis shown in (b) alongside the simulations of background contributions. Candidate $K^+ \to \pi^+[\mu^-e^+]_{\pi^0}$ events enter the magenta band in (a) and (b) and for these events the one dimensional projection onto $M_{\pi\mu e}$ is shown for data and background simulations in figure 7.56a. The final signal region for the search for $K^+ \to \pi^+[\mu^-e^+]_{\pi^0}$ events is the intersection of the magenta and orange shaded regions shown in (a).

Table K.1: Background predictions for $K^+ \to \pi^+[\mu^-e^+]_{\pi^0}$.

Sample	Control Region		Blinded Region	Signal Region
	N_{CRB}	N_{CRA}	N_{BR}	N_{SR}
Data	10882	0	0	0
$K^+ \to \pi^+\pi^+\pi^-$	$(1.12 \pm 0.11) \times 10^4$	0.22 ± 0.15	0.22 ± 0.15	0.22 ± 0.15
$K^+ \to \pi^+[e^+e^-\gamma]_{\pi^0}$	$(6.2 \pm 6.2) \times 10^{-6}$	$(6.2 \pm 6.2) \times 10^{-6}$	$(6.2 \pm 6.2) \times 10^{-6}$	$(6.2 \pm 6.2) \times 10^{-6}$
$K^+ \to \mu^+\nu_\mu[e^+e^-\gamma]_{\pi^0}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$	$(2.2 \pm 2.2) \times 10^{-7}$
$K^+ \to e^+\nu_e[e^+e^-\gamma]_{\pi^0}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$	$(3.7 \pm 3.7) \times 10^{-7}$
$K^+ \to \pi^+\pi^-\mu^+\nu_\mu$	$(7.8 \pm 3.8) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$	$(2.7 \pm 2.7) \times 10^{-4}$
$K^+ \to \pi^+\pi^-e^+\nu_e$	67.2 ± 6.8	0.014 ± 0.014	0.019 ± 0.014	0.014 ± 0.014
$K^+ \to \pi^+\mu^+\mu^-$	$(1.44 \pm 0.27) \times 10^{-3}$	$(4.3 \pm 4.3) \times 10^{-5}$	$(4.3 \pm 4.3) \times 10^{-5}$	$(4.3 \pm 4.3) \times 10^{-5}$
$K^+ \to \pi^+e^+e^-$	$(3.4 \pm 3.4) \times 10^{-9}$	$(3.4 \pm 3.4) \times 10^{-9}$	$(3.4 \pm 3.4) \times 10^{-9}$	$(3.4 \pm 3.4) \times 10^{-9}$
$K^+ \to \mu^+\nu_\mu\mu^+\mu^-$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$	$(1.7 \pm 1.7) \times 10^{-3}$
$K^+ \to e^+\nu_e\mu^+\mu^-$	$(4.4 \pm 4.4) \times 10^{-3}$	$(4.4 \pm 4.4) \times 10^{-3}$	$(4.4 \pm 4.4) \times 10^{-3}$	$(4.4 \pm 4.4) \times 10^{-3}$
$K^+ \to \mu^+\nu_\mu e^+e^-$	$(1.1 \pm 1.1) \times 10^{-8}$	$(1.1 \pm 1.1) \times 10^{-8}$	$(1.1 \pm 1.1) \times 10^{-8}$	$(1.1 \pm 1.1) \times 10^{-8}$
$K^+ \to e^+\nu_e e^+e^-$	$(1.6 \pm 1.6) \times 10^{-10}$	$(1.6 \pm 1.6) \times 10^{-10}$	$(1.6 \pm 1.6) \times 10^{-10}$	$(1.6 \pm 1.6) \times 10^{-10}$
Total background expected	$(1.12 \pm 0.11) \times 10^4$	0.24 ± 0.15	0.24 ± 0.15	0.24 ± 0.15
Data/MC	0.967 ± 0.097	–	–	–

K.3 Evaluation of Sensitivity for $\pi^0 \to \mu^+e^-$ Searches in the Decay Chain $K^+ \to \pi^+[\mu^+e^-]_{\pi^0}$

Using the same approach as discussed above in appendix K.2 a search for $\pi^0 \to \mu^+e^-$ can be performed considering K^+ decays to $\pi^+\mu^+e^-$ final states where a $K^+ \to \pi^+\pi^0$ decay is followed by a $\pi^0 \to \mu^+e^-$ decay. The $K^+ \to \pi^+\mu^+e^-$ search has not been reported in detail because, with the current dataset, the previous limit set by a dedicated experiment cannot be improved upon. However for the $K^+ \to \pi^+[\mu^+e^-]_{\pi^0}$ search, due to the low background and $M_{\mu e}$ invariant mass requirement, the Dalitz rejection cut (see section 7.2.3) is not required which leads to a relative increase of acceptance. The acceptance of $(2.73 \pm 0.02)\%$ (table 7.3) is still lower than for the $\pi^0 \to \mu^-e^+$ channel and leads to a single event sensitivity of (table 7.7)

$$\mathcal{B}_{SES}(K^+ \to \pi^+[\mu^+e^-]_{\pi^0}) = (1.59 \pm 0.11) \times 10^{-10}\,. \tag{K.3.1}$$

The expected background from $K^+ \to \pi^+\pi^+\pi^-$ is similar to the $K^+ \to \pi^+[\mu^+e^-]_{\pi^0}$ search but all other backgrounds are negligible (in the signal region, and only $K^+ \to \pi^+\pi^+\mu^+\nu_\mu$ decays additionally contribute in CRB). This leads to a total background expectation of 0.22 ± 0.15 in the signal region. In the likely scenario of 0 events observed an upper limit could be set at $\mathcal{B}(\pi^0 \to \mu^+e^-) < 3.8 \times 10^{-10}$. This limit would be weaker than the KTeV combined limit of [173] $\mathcal{B}(\pi^0 \to \mu^\pm e^\pm) < 3.6 \times 10^{-10}$ and identical to the BNL E865 experiment result from a similar $K^+ \to \pi^+[\mu^+e^-]_{\pi^0}$ search of [263] [270]. See section K.1 for discussion of these two previous limits. In either case no improvement can be made at this time but clearly with future data-taking at NA62 improvements can be expected, especially as the background is low, see section 7.7.4.