



UNIVERSITÀ DEGLI STUDI "ROMA TRE"

DIPARTIMENTO DI MATEMATICA E FISICA

Measurement of the K_S meson semileptonic branching ratios with the KLOE experiment

February 25, 2019

Candidato:
Andrea Selce

Relatore:
Prof. Filippo Ceradini

Corso di Dottorato in Fisica
XXXI Ciclo
Anno Accademico 2018-2019



Ai miei genitori.

Contents

1	Introduction	7
2	Phenomenological frame	9
2.1	The neutral kaon system	9
2.1.1	Mass and decay matrices	9
2.1.2	Semileptonic decay amplitudes	11
2.2	The V_{CKM} mixing matrix	13
2.2.1	Determining V_{us}	14
3	The KLOE experiment at DAΦNE	17
3.1	The DAΦNE collider	17
3.2	The KLOE detector	18
3.2.1	The drift chamber	18
3.2.2	The calorimeter	21
3.2.3	The trigger	23
3.2.4	FILFO	25
3.2.5	The data acquisition	25
3.3	Data reconstruction	25
3.3.1	Cluster reconstruction	25
3.3.2	Track reconstruction	26
3.3.3	Track-to-cluster association	27
3.3.4	Event classification	27
3.4	Beam parameters and luminosity	28
3.5	MC event simulation	29
3.5.1	MC decay channels	30
4	Analysis scheme	31
4.1	Data sample	32
4.2	K_L -crash and K_S identification	32
4.2.1	K_L -crash	32
4.2.2	K_S selection	34
5	Normalisation sample	35
5.1	Efficiency	35

6	$K_s \rightarrow \pi e \nu$	37
6.1	Selection	37
6.1.1	Preselection	37
6.1.2	Multivariate analysis	38
6.1.3	Selection based on time of flight	43
6.1.4	Extraction of the signal	46
6.2	Determination of the efficiencies	48
6.2.1	$K_L \rightarrow \pi e \nu$ control sample	49
6.2.2	MC efficiency	54
6.2.3	Determination of the ratio R_ϵ	54
6.2.4	Summary of efficiencies	56
6.3	Systematic uncertainties	57
6.3.1	BDT scan and systematics	58
6.3.2	Scan and systematic on $\delta DTOF$	59
6.3.3	Others systematic uncertainties	61
6.4	Result	62
7	$K_S \rightarrow \pi \mu \nu$	63
7.1	Selection	63
7.1.1	Preselection and multivariate analysis	63
7.1.2	Selection based on time of flight	64
7.1.3	Extraction of the signal	66
7.2	Determination of efficiencies	68
7.2.1	$K_S \rightarrow \pi \mu \nu$ control sample	68
7.2.2	Monte Carlo efficiency	71
7.2.3	Determination of R_ϵ	72
7.2.4	Summary of efficiencies	73
7.3	Systematic uncertainties	73
7.3.1	BDT scan and systematic uncertainty	73
7.3.2	TOF scan and systematic uncertainty	73
7.4	Result	75
8	Conclusions	77
8.1	V_{us} determination	78
8.1.1	$K_S \rightarrow \pi e \nu$	78
8.1.2	$K_S \rightarrow \pi \mu \nu$	79
8.2	Lepton universality test	79
8.3	Perspectives	80
	Appendices	81
	A Boosted Decision Tree	83

Chapter 1

Introduction

The Cabibbo-Kobayashi-Maskawa quark mixing matrix is a main experimental input to the Standard Model of fundamental interactions. Measuring with high precision the matrix elements is important both to increase our knowledge, and to look for signals of new physics, testing the unitarity of the matrix. The Cabibbo angle $V_{us} = \lambda = \sin \theta_C$ is a critical ingredient for these tests, and it is best extracted from the kaon semileptonic decay amplitudes. Among the six semileptonic decays of K^\pm , K_L and K_S mesons, the worse determination of $|V_{us}|$ is from to the $K_S \rightarrow \pi e \nu$ decay, due to the higher experimental uncertainty of this decays channel respect to the other kaon decay channels. There is no measurement thus far of the $K_S \rightarrow \pi \mu \nu$ branching ratio.

The first observation of the $K_S \rightarrow \pi e \nu$ decay was done at CERN in a heavy-liquid bubble chamber in 1963 [1]. Neutral kaons were produced by charge exchange of an 800 MeV K^+ beam from the proton-synchrotron. The measurement of the time distribution of K^0 and \bar{K}^0 decays into $\pi e \nu$ gave $BR(K_S \rightarrow \pi e \nu) = (8.3 \pm 2.2) 10^{-4}$.

A comparable value was obtained by the CMD-2 experiment [2] at the e^+e^- collider VEPP-2M in Novosibirsk in 1998. They exploited for the first time the K_S -tagging method applied to $\phi \rightarrow K_L K_S$ decays. The result with 75 ± 13 events is $BR(K_S \rightarrow \pi e \nu) = (7.2 \pm 1.4) 10^{-4}$.

The first KLOE measurement [3] was obtained with an integrated luminosity of 17 pb^{-1} recorded during the DAΦNE pilot run in 2001. The K_S mesons from ϕ decays were tagged by identifying the K_L interactions in the calorimeter. The result with 624 ± 30 events is $BR(K_S \rightarrow \pi e \nu) = (6.91 \pm 0.34 \pm 0.15) 10^{-4}$.

The two most recent and precise results were obtained by the NA48 and KLOE experiments. NA48 exploited the high-energy neutral kaon beam at CERN produced with a 400 GeV proton beam on a Beryllium target [4]. The ratio $\frac{K_S \rightarrow \pi e \nu}{K_L \rightarrow \pi e \nu}$ was obtained from the measurement of the decay time distribution, providing $BR(K_S \rightarrow \pi e \nu) = (7.046 \pm 0.18 \pm 0.16) 10^{-4}$.

The updated KLOE result is based on a new event sample corresponding to an integrated luminosity of 410 pb^{-1} [5]. The K_S decays into $\pi^+ e^- \bar{\nu}$ and $\pi^- e^+ \nu$ were measured separately to extract the charge asymmetry. The event yields were normalised to the main decays $K_S \rightarrow \pi^+ \pi^-$ and $K_S \rightarrow \pi^0 \pi^0$ recorded in the same data sample. The result is $BR(K_S \rightarrow \pi e \nu) = (7.046 \pm 0.077 \pm 0.050) 10^{-4}$.

The first goal of this work is to do the measurement on 1.63 fb^{-1} of data collected by KLOE in 2004-2005. The number of counts of the $K_S \rightarrow \pi e \nu$ will be normalized to the number of $K_S \rightarrow \pi^+ \pi^-$, and the $\text{BR}(K_S \rightarrow \pi e \nu)$ will be obtained multiplying the ratio to $\text{BR}(K_S \rightarrow \pi^+ \pi^-) = (69.0196 \pm 0.051)\%$ measured by KLOE [5].

The K_S semileptonic decay into the muon final state has not been measured yet. This is due to the fact that a pure K_S beam is difficult to select at high energy, while at low energy is more difficult to separate the signal from the abundant $K_S \rightarrow \pi^+ \pi^-$ background for muon than for electron, and because of the decay $K_S \rightarrow \pi^+ \pi^- \rightarrow \pi \mu \nu$ is a background very similar to the signal. The second goal of this thesis is to measure the $\text{BR}(K_S \rightarrow \pi \mu \nu)$ for the first time.

In Chapter 2 the theoretical and phenomenological frame of semileptonic decays and V_{us} matrix element is described. Chapter 3 contains a brief description of the KLOE experiment, focusing on the KLOE detector, trigger and the data reconstruction. The analysis scheme is presented in Chapter 4, with particular attention on the tag of the K_S for the K_L identification. The way the normalisation sample is selected and counted is described in Chapter 5. Chapter 6 is dedicated to the $K_S \rightarrow \pi e \nu$ analysis, the data selection (Sec. 6.1) and the evaluation of efficiencies (Sec. 6.2) and systematics uncertainties (Sec. 6.3). Chapter 7 is dedicated to the measurement $K_s \rightarrow \pi \mu \nu$ decay, organized in the same way as the chapters before. Finally Chapter 8 summarises the results of this thesis work.

Chapter 2

Phenomenological frame

2.1 The neutral kaon system

Kaons were discovered by G. D. Rochester and C. C. Butler in 1947 studying cosmic ray interactions in a cloud chamber [6]. The problem of different production rate and decay time of these new particles ended with the introduction of a new quantum number S , called strangeness, that is conserved in both electromagnetic and strong interactions, while first order weak decay are generated by transitions with $|\Delta S| = 1$. Kaons are part of the isospin $I = \frac{1}{2}$ doublets: $(K^+ K^0)$ with $S = +1$ and (\bar{K}^0, K^-) with $S = -1$, and are produced in strong interaction processes like:

$$\begin{aligned}\pi^+ p &\rightarrow K^+ \bar{K}^0 p \\ \pi^- p &\rightarrow K^0 \Lambda \\ p \bar{p} &\rightarrow K^- \pi^+ K^0 \\ p \bar{p} &\rightarrow K^+ \pi^- \bar{K}^0\end{aligned}$$

Considering the strong interaction, the K^0 meson is a particle and \bar{K}^0 is the corresponding antiparticle. The weak interaction allows non conservation of strangeness and transitions like $K^0 \rightarrow 2\pi \rightarrow \bar{K}^0$ or $K^0 \rightarrow 3\pi \rightarrow \bar{K}^0$. The two strangeness eigenstates can oscillate one into another via the $|\Delta S| = 2$ second order weak process, through virtual 2π or 3π states. Therefore, neutral kaon states have defined mass and width when they evolve in free space, as a mixture of K^0 and \bar{K}^0 strong eigenstates.

2.1.1 Mass and decay matrices

The time evolution of the neutral kaon system in the particle rest frame can be described using an effective 2×2 Hamiltonian [7]:

$$|\psi(t)\rangle = A(t) |K^0\rangle + B(t) |\bar{K}^0\rangle \quad (2.1)$$

$$i \frac{\partial}{\partial t} \begin{bmatrix} A(t) \\ B(t) \end{bmatrix} = \begin{bmatrix} h & l \\ m & n \end{bmatrix} \begin{bmatrix} A(t) \\ B(t) \end{bmatrix} = [M - \frac{i}{2}\Gamma] \begin{bmatrix} A(t) \\ B(t) \end{bmatrix}$$

where h, l, m and n are complex numbers and the matrices M and Γ are hermitian. The number of parameters is eight. If CPT holds, the number of parameters reduces to six, since CPT requires that both the masses as well as the decay rates of K^0 and \bar{K}^0 be the same:

$$M_{11} \stackrel{CPT}{=} M_{22}$$

$$\Gamma_{11} \stackrel{CPT}{=} \Gamma_{22}$$

If CP holds, in the phase convention in which $CP|K^0\rangle = |\bar{K}^0\rangle$, all matrix elements are invariant under the exchange of the indices $1 \leftrightarrow 2$:

$$M_{11} \stackrel{CP}{=} M_{22}$$

$$\Gamma_{11} \stackrel{CP}{=} \Gamma_{22}$$

$$M_{12} \stackrel{CP}{=} M_{21} = M_{12}^*$$

$$\Gamma_{11} \stackrel{CP}{=} \Gamma_{21} = \Gamma_{12}^*$$

Hence CP implies CPT. If CP is assumed, the number of parameters in the effective hamiltonian is reduced to four.

The eigenstates of the Hamiltonian are the physical K_S and K_L states, whose mass and width are, respectively, m_S and m_L , Γ_S and Γ_L . These states are the following superpositions a K^0 and \bar{K}^0 :

$$|K_S\rangle = \frac{1}{[2(1 + |\epsilon_K + \delta_K|^2)]^{1/2}} \left[(1 + \epsilon_K + \delta_K) |K^0\rangle + (1 - \epsilon_K - \delta_K) |\bar{K}^0\rangle \right]$$

$$|K_L\rangle = \frac{1}{[2(1 + |\epsilon_K - \delta_K|^2)]^{1/2}} \left[(1 + \epsilon_K - \delta_K) |K^0\rangle - (1 - \epsilon_K + \delta_K) |\bar{K}^0\rangle \right] \quad (2.2)$$

with eigenvalues:

$$\lambda_S = m_s - \frac{i}{2}\Gamma_s = \frac{1}{2} \frac{n - h + \sqrt{(h-2)^2 + 4lm}}{2l}$$

$$\lambda_L = m_L - \frac{i}{2}\Gamma_L = \frac{1}{2} \frac{n - h - \sqrt{(h-2)^2 + 4lm}}{2l}$$

The parameters ϵ_k and δ_k are small complex numbers which characterise CP and CPT violation in the Hamiltonian. Ignoring negligible quadratic terms:

$$\epsilon_K = \frac{\Im M_{12} - i\Im \Gamma_{12}/2}{[(m_S - m_L) - i(\Gamma_S - \Gamma_L)/2]}$$

$$\delta_K = \frac{(M_{11} - M_{22}) - i(\Gamma_{11} - \Gamma_{22})/2}{2[(m_s - m_L) - i(\Gamma_s - \Gamma_L)/2]}$$

2.1.2 Semileptonic decay amplitudes

The decay amplitudes for K^0 and \bar{K}^0 into the final state $\pi\ell\nu$, called the semileptonic decay, can be expressed as a function of four complex amplitudes. Each of these amplitudes parameterises the conservation or violation in the decay of a symmetry such as CP , T , CPT , or the $\Delta S = \Delta Q$ rule. These amplitudes are defined in the following way:

$$\begin{aligned}
\langle \ell^+ \pi^- \nu | \mathcal{H}_{WK} | K^0 \rangle &= a + b \equiv \mathcal{A}_+ \\
\langle \ell^- \pi^+ \bar{\nu} | \mathcal{H}_{WK} | \bar{K}^0 \rangle &= a^* - b^* \equiv \bar{\mathcal{A}}_- \\
\langle \ell^- \pi^+ \bar{\nu} | \mathcal{H}_{WK} | K^0 \rangle &= c + d \equiv \mathcal{A}_- \\
\langle \ell^+ \pi^- \nu | \mathcal{H}_{WK} | \bar{K}^0 \rangle &= c^* - d^* \equiv \bar{\mathcal{A}}_+
\end{aligned} \tag{2.3}$$

The two amplitudes c and d both violate the $\Delta S = \Delta Q$ rule. The symmetry properties of these amplitudes can be understood by applying the rules for symmetry operators C , P and T . If A and B are two spin zero systems, and \bar{A} and \bar{B} the corresponding antiparticle systems, the following relations hold for two amplitudes $M(A \rightarrow B)$ and $M(\bar{A} \rightarrow \bar{B})$:

$$\langle TB | T \mathcal{H}_{WK} T^{-1} | TA \rangle = \langle B | T \mathcal{H}_{WK} T^{-1} | A \rangle^* \tag{2.4}$$

$$\langle CPB | CP \mathcal{H}_{WK} CP^{-1} | CPA \rangle = \langle \bar{B} | CP \mathcal{H}_{WK} CP^{-1} | \bar{A} \rangle \tag{2.5}$$

$$\langle CPTB | CPT \mathcal{H}_{WK} CPT^{-1} | CPTA \rangle = \langle \bar{B} | CPT \mathcal{H}_{WK} CPT^{-1} | \bar{A} \rangle^* \tag{2.6}$$

Therefore if T is conserved ($[H_{WK}, T] = 0$) all amplitudes are real; if CP is conserved ($[H_{WK}, CP] = 0$) the decay amplitudes $M(A \rightarrow B)$ and $M(\bar{A} \rightarrow \bar{B})$ are identical; if CPT is conserved ($[H_{WK}, CPT] = 0$) the relation $M(A \rightarrow B)$ and $M(\bar{A} \rightarrow \bar{B})^*$ holds. The symmetry properties of the semileptonic amplitudes, derived using these results, are summarised in Table 2.1.

Table 2.1: Symmetry properties of the semileptonic amplitudes

T	$\Im a = \Im b = \Im c = \Im d = 0$
CP	$\Im a = \Re b = \Im c = \Re d = 0$
CPT	$b = d = 0$
$\Delta S = \Delta Q$	$c = d = 0$

It is useful to define the quantities:

$$x = \frac{\bar{\mathcal{A}}_-}{\mathcal{A}_+} ; \quad \bar{x} = \left(\frac{\mathcal{A}_+}{\bar{\mathcal{A}}_-} \right)^* ; \quad y = \frac{\bar{\mathcal{A}}_-^* - \mathcal{A}_+}{\bar{\mathcal{A}}_-^* + \mathcal{A}_+} \tag{2.7}$$

where x parameterises the violation of the $\Delta S = \Delta Q$ rule in the decay to positive-charged leptons, while \bar{x} parameterises the violation in the decay to negative-charged

leptons; $y = -\frac{b}{a}$ parameterises the violation of CPT when $\Delta S = \Delta Q$ holds. It is seen from Table ?? that $\Re a$ is the only term invariant under T , CP , and CPT transformations in decays where $\Delta S = \Delta Q$. Therefore all the other amplitudes can be considered much smaller than $\Re a$. Thus, two parameters which violate the $\Delta S = \Delta Q$ rule can be defined: $x_+ = (x + \bar{x})/2 \approx c^*/a$ when CPT is conserved, and $x_- = x_+ = (x - \bar{x})/2 \approx -d^*/a$ when CPT is violated. The decay amplitudes for K_L and K_S decays into final states of opposite lepton charge are obtained by using the definitions in Eq. 2.3 and the amplitudes in Eq. 2.2. The square of the amplitudes is then obtained keeping only the symmetry-conserving terms and, to first order, the symmetry-violating terms. It is useful to introduce also the charge asymmetry

$$\delta_{S,L} = \frac{N_{S,L}^+ - N_{S,L}^-}{N_{S,L}^+ + N_{S,L}^-} = \frac{\Gamma_{S,L}^+ - \Gamma_{S,L}^-}{\Gamma_{S,L}^+ + \Gamma_{S,L}^-} \quad (2.8)$$

Again, using the above definitions of x_+ , x_- and y :

$$\delta_S = 2\Re\epsilon_K + 2Re\delta_K - 2\Re y + 2\Re x_- \quad (2.9)$$

$$\delta_L = 2\Re\epsilon_K - 2\Re\delta_K - 2\Re y - 2\Re x_- \quad (2.10)$$

$K_S \rightarrow \pi e\nu$ and $K_S \rightarrow \pi\mu\nu$ decays

Starting from the semileptonic decay width

$$\Gamma_{K\ell 3} = \frac{G_F^2 M_K^5}{192\pi^3} S_{EW} (1 + \delta_K^\ell + \delta_{SU2}) C^2 |V_{us}| f_+^2(0) I_K^\ell \quad (2.11)$$

and assuming lepton universality, the ratio of the two semileptonic decay amplitudes can be expressed as:

$$r_{\mu e} = \frac{\Gamma(K_s \rightarrow \pi\mu\nu)}{\Gamma(K_s \rightarrow \pi e\nu)} = \frac{1 + \delta_K^\mu I_K^\mu}{1 + \delta_K^e I_K^e} \quad (2.12)$$

where δ_K^ℓ are mode-dependent long-distance radiative correction, and I_K^ℓ are decay phase space integrals. Using $\frac{I_K^\mu}{I_K^e} = 0.6622 \pm 0.0018$ from the measurement of the semileptonic decays of the KTeV experiment [8] and $\frac{1 + \delta_K^\mu}{1 + \delta_K^e} = 1.0058 \pm 0.0010$ [9], the result is

$$R_{\mu e}^{SM} = 0.6660 \pm 0.0019 \quad (2.13)$$

Thus the expected value is [10]:

$$BR(K_S \rightarrow \pi\mu\nu) = (4.69 \pm 0.05) \times 10^{-4} \quad (2.14)$$

A different result could be associated to Lepton Universality Violation (LUV).

2.2 The V_{CKM} mixing matrix

The masses and mixings of quarks have a common origin in the Standard Model (SM)[10]. They arise from the Yukawa interactions with the Higgs field

$$L_Y = -Y_{ij}^d \overline{Q_{Li}^I} \phi d_{Rj}^I - Y_{ij}^u \overline{Q_{Li}^I} \varepsilon \phi^+ u_{Rj}^I \quad (2.15)$$

where $Y^{u,d}$ are 3×3 complex matrices, ϕ is the Higgs field, i, j are generation labels, and ε is the 2×2 antisymmetric tensor. Q_L^I are left-handed quark doublets, and d_R^I and u_R^I are right-handed down-type and up-type quark singlets, respectively, in the weak-eigenstate basis. When ϕ acquires a vacuum expectation value, $\langle \phi \rangle = (0, v/\sqrt{2})$, Eq. 2.15 yields mass terms for the quarks. The physical states are obtained by diagonalising $Y^{u,d}$ by four unitary matrices, $V_{L,R}^{u,d}$ as $M_{diag}^f = V_L^f Y^f V_R^{f\dagger} (v/\sqrt{2})$, $f = u, d$. As a result, the charged current W^\pm interactions couple to the physical u_{Lj} and d_{Lk} quarks with couplings given by

$$\frac{g}{\sqrt{2}} (\bar{u}_L \quad \bar{c}_L \quad \bar{t}_L) \gamma^\mu W_\mu V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + h.c., \quad V_{CKM} \equiv V_L^u V_L^{d\dagger} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad (2.16)$$

The Cabibbo-Kobayashi-Maskawa mixing matrix V_{CKM} [ref] is a 3×3 unitary matrix. It can be parameterised by three mixing angles and the CP-violating KM phase [ref]. Of the many possible conventions, a standard choice has become [ref]

$$V_{CKM} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.17)$$

where $s_{ij} = \sin \theta_{ij}$, $c_{ij} = \cos \theta_{ij}$, and δ is the phase responsible for all CP-violating phenomena in flavour-changing processes. The angles θ_{ij} can be chosen to lie in the first quadrant, $s_{ij} \geq 0, c_{ij} \geq 0$.

It is known experimentally that $s_{13} \ll s_{23} \ll s_{12}$; it is convenient to use the Wolfenstein parameterisation:

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, \quad s_{12} = A\lambda^2 = \lambda \frac{|V_{cb}|}{|V_{us}|}, \quad (2.18)$$

$$s_{13}e^{i\delta} = V_{ub}^* = A\lambda^3(\rho + i\eta) = \frac{A\lambda^3(\bar{\rho} + i\bar{\eta})\sqrt{1 - A^2\lambda^4}}{\sqrt{1 - \lambda^2[1 - A^2\lambda^4]}(\bar{\rho} + i\bar{\eta})} \quad (2.19)$$

These relations ensure that $\bar{\rho} + i\bar{\eta} = -(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)$ is phase convention independent, and the CKM matrix written in terms of $\lambda, A, \bar{\rho}$ and $\bar{\eta}$ is unitary to all orders in λ .

The Wolfenstein parameterisation of the CKM matrix is

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4) \quad (2.20)$$

The unitarity of the CKM matrix imposes

$$\sum_i V_{ij} V_{ik}^* = \delta_{jk} \quad \text{and} \quad \sum_j V_{ij} V_{kj}^* = \delta_{ik} \quad (2.21)$$

The six vanishing combinations can be represented as triangles in a complex plane, and those obtained by taking scalar products of neighbouring rows or columns are nearly degenerate. The areas of all triangles are the same, and is a phase-convention-independent measure of CP violation.

2.2.1 Determining V_{us}

The parameter V_{ud} is measured with good precision from the $0^+ \rightarrow 0^+$ transitions of nuclear β decay. The measurement of the Cabibbo angle, $V_{us} = \lambda = \sin \theta_C$ is a critical ingredient for precise tests of CKM unitarity [10, 11]. For many years, the precise value of V_{us} was somewhat controversial, with kaon decays suggesting $\lambda \approx 0.220$, while indirect determinations via nuclear β -decays implied a somewhat larger value, $\lambda \approx 0.225 - 0.230$. This difference resulted in a 2–2.5 sigma deviation from the unitarity requirement

$$|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 1 \quad (2.22)$$

with $|V_{ub}|^2 \approx 1.7 \times 10^{-5}$ giving a negligible contribution to the sum. Eq. (2.22) is currently the most stringent test of unitarity in the CKM matrix.

The parameter $|V_{us}|$ can be determined from kaon decays, hyperon decays, and tau decays. For $K\ell 3$ decays:

$$\Gamma_{K\ell 3} = \frac{G_F^2 M_K^5}{192\pi^3} S_{EW} (1 + \delta_K^\ell + \delta_{SU2}) C^2 |V_{us}|^2 f_+^2(0) I_K^\ell \quad (2.23)$$

where, $\ell = e, \mu$, G_F is the Fermi constant, M_K is the kaon mass, S_{EW} is the short-distance radiative correction, δ_K^ℓ is the mode-dependent long-distance radiative correction, $f_+(0)$ is the form factor at zero momentum transfer to the $\ell\nu$ system, $C=1$ is the isospin factor for neutral Kaon decay and I_K^ℓ is the phase-space integral. For the neutral kaons, $\delta_{SU2} = 0$, while for charged kaons it is the deviation from one of the ratio of $f_+(0)$ for the charged and neutral kaon. The values of S_{EW} , δ_K^ℓ , δ_{SU2} and $f_+(0)$ are calculated from theory.

The current status of $|V_{us}|f_+(0)$ measurement are reported in Figure 2.1. Among the six kaon semileptonic decays $K_S \rightarrow \pi\ell\nu$ gives the less precise value of $|V_{us}|f_+(0)$ and the decay $K_S \rightarrow \pi\mu\nu$ has not been measured yet. Thus the picture presented in Figure 2.1 motivates a precise measurement of both semileptonic decays of the K_S meson.

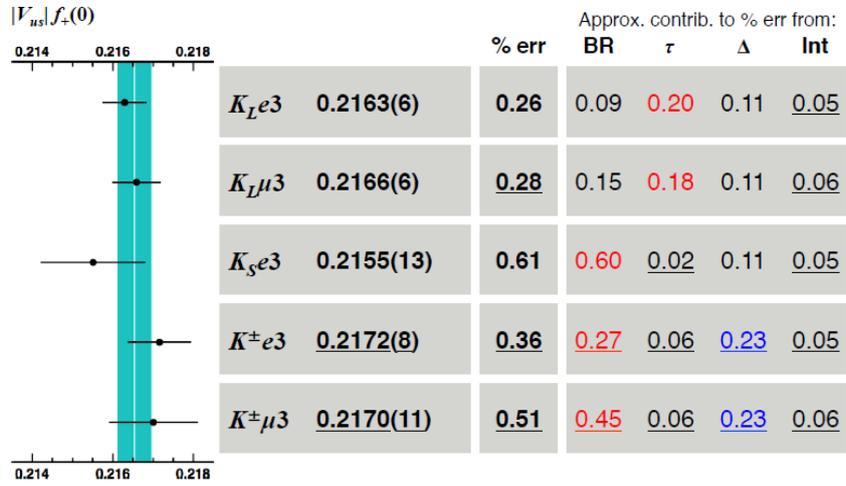


Figure 2.1: Contribution of the kaon semileptonic decay channels to the V_{us} matrix element determination, with relative errors and main contribution to the errors [12].

Chapter 3

The KLOE experiment at DAΦNE

The KLOE¹ experiment was installed at the DAΦNE ϕ -factory, in the Laboratori Nazionali di Frascati of the Italian Institute for Nuclear Physics (INFN). DAΦNE is an electron–positron collider operating at around 1020 MeV, the ϕ meson mass. The main decays of the ϕ meson are: neutral kaon pairs (34%) , charged kaon pairs (49%), $\rho\pi$ (15%) and $\eta\gamma$ (1.3%) [10]. Detecting ϕ decay products with the KLOE apparatus it is possible to perform studies on discrete symmetries, charged and neutral kaon decay parameters and investigating light-hadrons properties.

In this chapter a short description of the DAΦNE accelerator complex is presented, together with the KLOE detector with a brief overview of the trigger, data acquisition system and offline software.

3.1 The DAΦNE collider

DAΦNE² is an e^+e^- collider designed to operate at the centre-of-mass energy $\sqrt{s} = m_\phi = 1019.45$ MeV. The layout of the accelerating complex [13] is shown in Figure 3.1.

The accelerator complex consists of a LINAC, an accumulator ring and a two-ring collider. Electrons are accelerated to the energy of about 510 MeV in the LINAC, accumulated and cooled in the accumulator and transferred to the electron ring. Positrons are generated in an intermediate station in the LINAC, where 250 MeV electrons hit a tungsten target to produce positrons, these then follow the same processing as electrons in order to be injected in the positron ring. Up to 120 electron and positron bunches circulate in the two main rings and collide in the interaction region once per turn, thus minimising the mutual perturbations. In the interaction region electrons and positrons collide at an angle of $\pi - 0.025$ rad in the horizontal plane, with a frequency of 356 MHz, corresponding to a bunch crossing period of $T_{\text{rf}} = 2.7$ ns. The ϕ mesons thus move in the laboratory frame with a momentum of about 13 MeV corresponding to $\beta_\phi \sim 0.015$, $\gamma_\phi \sim 1.0001$. Therefore, neutral kaons from ϕ decays are not monochromatic in the laboratory, their momentum range is between 104 and 116 MeV and it is a single-valued

¹ K–LONg Experiment

² Double Annular ring For Nice Experiments

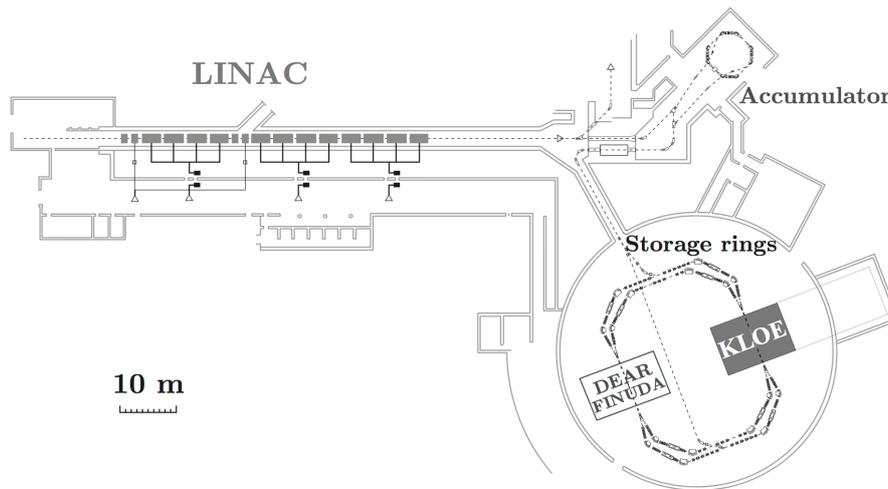


Figure 3.1: The DAΦNE complex schematic view.

function of the angle between the kaon momentum in the laboratory frame and the ϕ momentum along the x -axis³. Knowing the kaon direction with a few degrees accuracy allows to transform to the ϕ center-of-mass.

3.2 The KLOE detector

The KLOE detector consists of a large cylindrical drift chamber, to reconstruct the trajectories of charged particles, and a hermetic calorimeter, with a barrel and two end-cap modules, to measure the energy and the entry point of particles. Both are immersed in an axial magnetic field of 0.52 T provided by a superconducting coil and an iron yoke surrounding the calorimeter. The beam pipe at the interaction region (IR) is spherically shaped, with a 10 cm radius, to preserve the K_L - K_S interference, and it is made of a Beryllium-Aluminum alloy of 0.5 mm thickness.

The detector, whose transverse view is shown in Figure 3.2, is designed to perform high-precision measurements, while its size is demanded by the long decay path of the K_L meson, 3.4 m. Thus, in order to capture approximately 40% of K_L decays, the detector was built with a core volume of 2 meter radius.

3.2.1 The drift chamber

The drift chamber (DC) [14] is designed to detect all charged particles produced in the interactions, and the secondary products from meson decays, and measure their properties with high precision. It provides:

³ The coordinate system is defined such that the x -axis is in the horizontal plane, towards the center of DAΦNE, the y -axis is vertical, pointing upward, and the z -axis bisects the angle between the two beam lines.

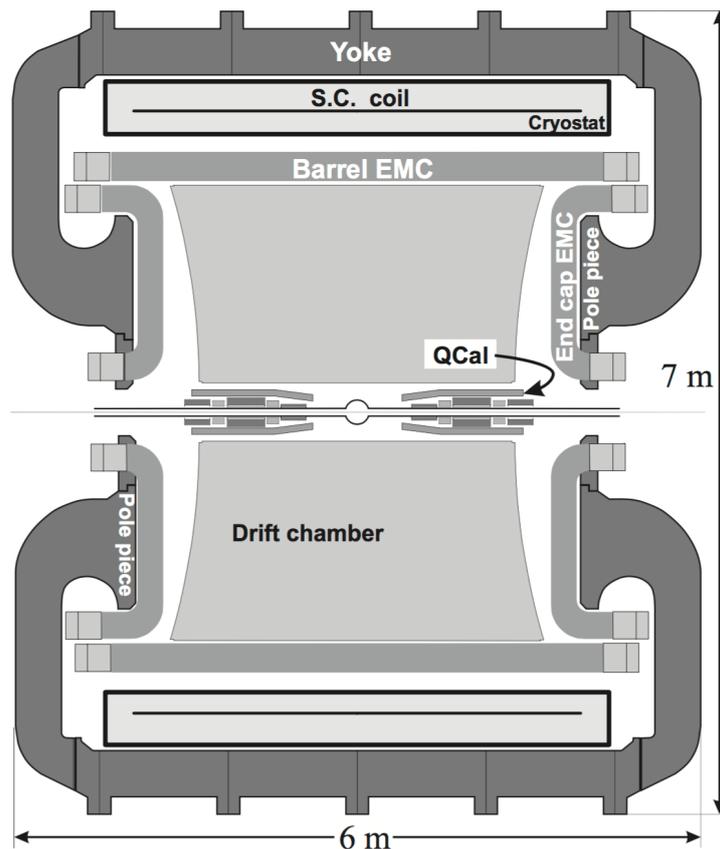


Figure 3.2: Vertical cross section of the KLOE detector, showing the interaction region, the drift chamber (DC), the electromagnetic calorimeter (EMC), the superconducting coil, the return yoke of the magnet and the structure of the beam pipe.

- high and uniform reconstruction efficiency over a large volume;
- good momentum resolution ($\delta p_T/p_T \sim 0.4\%$), the dominant contribution to the momentum resolution being due to multiple Coulomb scattering;
- a track resolution in the transverse plane $\delta\rho \sim 200 \mu\text{m}$, with $\rho = \sqrt{x^2 + y^2}$, a vertex resolution $\delta\rho_{\text{vtx}} \sim 1 \text{ mm}$, and $\delta z_{\text{vtx}} \sim 2 \text{ mm}$ over the whole sensitive volume;
- transparency to low energy photons (down to 20 MeV), and minimisation of K_L regeneration;
- fast trigger for charged particles.

To minimise the K_L regeneration, the Coulomb multiple scattering and the photon absorption, the drift chamber is constructed out of carbon fiber composite with low- Z and low density and is filled with a gas mixture of Helium (90%) and Isobutane (10%); this permits to obtain a radiation length (gas + 52140 wires) of about 900 m. Gold-plated tungsten wires are used as anodes and silver-plated aluminum as field-shaping wires. The

signals from the sense wires are amplified, discriminated and transmitted to the read-out system, consisting in ADCs for dE/dx measurement and TDCs for time measurement. Wires are strung in an all-stereo geometry in order to obtain a high and uniform track and vertex reconstruction efficiency. The stereo angle varies with the radius from 50 mrad to 120 mrad going outward. This design results in a uniform filling of the sensitive volume with almost square drift cells arranged in 58 concentric rings, with shape slowly changing along the z -axis. The track density is higher at small radii, thus the dimensions of the cells were designed to be of about $2 \times 2 \text{ cm}^2$ for the 12 innermost wire layers and of about $3 \times 3 \text{ cm}^2$ for the remaining 46 layers.

To extract the spatial position from the measured drift times of the incident particles, 232 space-to-time relations are used. They are parameterised in terms of two angles β and $\tilde{\phi}$ defined in Figure 3.3. The β angle characterises the geometry of the cell, directly related to the electric field responsible for the electron drift and avalanche multiplication; $\tilde{\phi}$ gives the orientation of the particle trajectory in the cell reference frame, defined in the transverse plane, with origin in the sense wire of the cell. Using the wire geometry, the space-to-time relations and the magnetic field, one can reconstruct the tracks and the vertices of charged particles.

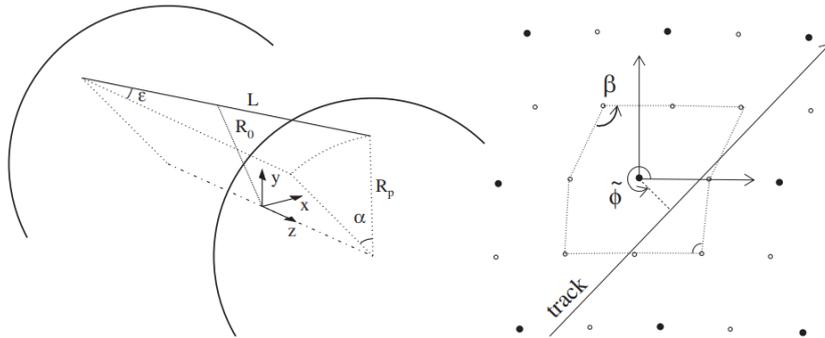


Figure 3.3: *Left*: wire geometry with the definition of the stereo angle ϵ between the wire of length L and the z -axis. *Right*: definition of β and $\tilde{\phi}$ angles characterising the shape of the cell and the angle of the incident track.

Periodical calibrations with cosmic ray muon samples (large enough for measuring more than 200 different space-to-time relations) are performed to monitor the stability in time of the drift chamber performance. The calibration is performed at the beginning of each run and selects about 8×10^4 cosmic ray muon events. These events are tracked using the existing space-to-time relations and the average value of the residuals for hits in the central part of the cells is monitored. If these residuals exceed $40 \mu\text{m}$, then additional 3×10^5 cosmic-ray muon events are collected, and a new set of calibration constants is provided.

The DC provides three dimensional tracking with a resolution in the transverse plane of about $200 \mu\text{m}$, a resolution along z of about 2 mm. The resolution of the decay vertex position is 1 mm for vertices inside the chamber volume and 3 mm for vertices reconstructed outside. The particle momentum is determined from the curvature of its

trajectory in the magnetic field with a fractional accuracy $\sigma_p/p = 0.4\%$ for polar angles $45^\circ < \theta < 135^\circ$.

3.2.2 The calorimeter

The electromagnetic calorimeter (EMC) [15] is designed to fulfil several requirements:

- good time resolution (~ 100 ps) and good spatial determination of the particle hit point (~ 1 cm);
- hermeticity (98% of the solid angle), good energy resolution and high efficiency for energy above 20 MeV;
- hermetic detection of low energy photons with high efficiency, adequate energy resolution and excellent time resolution to reconstruct the vertex of K_L neutral decays;
- particle identification for electrons, muons and charged pions;
- fast first-level trigger.

The calorimeter is composed of three modules: one barrel and two end-cap modules. The barrel calorimeter is a cylinder with inner diameter of 4 m, made of 24 trapezoidal modules, 4.3 m long and 23 cm thick. Each end-cap calorimeter consists in 32 vertical C-shaped modules. Each module consists of a mixture of lead (48% of the volume), scintillating fibers (42%) and epoxy glue (10%). The fibers, with a diameter of 1 mm, are embedded in 0.5 mm lead foils to allow the showering processes. The special care in designing and assembling of the Pb-scintillating fiber composite ensures that the light propagates along the fiber in a single mode with velocity ~ 17 cm/ns, which greatly reduces the spread of the arrival time of the light signals at the fiber ends. The calorimeter modules are read out at both sides by 4.4×4.4 cm² light guides coupled to photomultipliers, each defining a calorimeter cell. The calorimeter cells are grouped to form five planes and twelve columns. Signals from the photomultipliers are sent to ADCs for energy measurements and trigger, and to TDCs for time measurements.

When a particle hits the calorimeter, for each cell the charge and time of arrival of the photomultiplier signals are recorded. The amplitude of the signals, A_i , is proportional to the amount of the deposited energy, while the recorded times, t_i , are related to the time of flight of the particle. For each cell, the position of the readout elements and the difference of the arrival times at the two fiber ends determine the shower position with about 1 cm accuracy.

As a first step, the reconstruction program makes the average of time and energy of the recorded t_i and A_i for the two sides of each cell and computes the hit position. At this stage corrections for attenuation length, energy scale, time offsets and light propagation speed are taken into account. Then a clustering procedure is used to find and group nearby clumps of energy deposition; average quantities over all the participating cells is also calculated. The calibration constants to transform t_i and A_i from raw quantities to

time in nanosecond and energy in MeV are evaluated with dedicated online and offline algorithms.

The energy calibration starts by a first equalisation in cell response to minimum ionising particles (MIP) at the calorimeter center and by determining the attenuation length of each single cell with a dedicated cosmic-ray trigger. This is done before the start of each long data-taking period. The determination of the absolute energy scale in MeV relies instead on a monochromatic source of 510 MeV photons from $e^+e^- \rightarrow \gamma\gamma$ events. The determination of the linearity of the response is done with radiative Bhabha scattering ($e^+e^- \rightarrow e^+e^-\gamma$) and $\phi \rightarrow \pi^+\pi^0\pi^-$ events. This calibration is routinely carried out every 200–400 nb^{-1} of integrated luminosity.

DAΦNE operates with a bunch-crossing period equal to the machine radio frequency (RF) period, $T_{\text{rf}} = 2.715$ ns. Due to the spread of the particles arrival times, the trigger is not able to identify the bunch crossing related to each event, which has to be determined offline. The common start signal to the calorimeter TDC boards is provided by the first-level trigger, which will be described in the next section. The stop is given, instead, by the photomultiplier signals delayed because of the electronics and light propagation in the fibers. The time associated to an energy cluster, T_{cl} , is related to the time of flight of particles from the interaction point (IP) to the calorimeter, T_{tof} , by the relation:

$$T_{\text{cl}} = T_{\text{tof}} + \delta_{\text{c}} - N_{\text{bc}}T_{\text{rf}},$$

where δ_{c} is a single number accounting for the overall electronics offsets and cable delays and N_{bc} is the number of bunch-crossing periods needed to generate the TDC start. The values of δ_{c} and N_{bc} are determined for each data-taking run with $e^+e^- \rightarrow \gamma\gamma$ events by looking at the $T_{\text{cl}} - R_{\text{cl}}/c$ distribution, which exhibits well separated peaks corresponding to different values of N_{bc} . The constant δ_{c} is arbitrarily defined as the position of the most populated peak, and T_{rf} is obtained from the distance between the peaks. Both quantities are evaluated with a precision better than 4 ps for a typical run of 200 nb^{-1} of integrated luminosity. This measurement of T_{rf} allows to set the absolute calorimeter time scale to few ps. During offline processing, to allow the cluster times to be related to the particle time of flight, δ_{c} is determined and, on an event-by-event basis, the global event start time $T_0 = N_{\text{bc}}T_{\text{rf}}$ is set, so that the corrected cluster time is obtained as follows:

$$t_{\text{cl}} = T_{\text{cl}} - (\delta_{\text{c}} - T_0).$$

A starting value for all analyses is evaluated by assuming that the earliest cluster in the event is due to a photon originating at the interaction point. Further corrections are analysis dependent.

The high photon yield and the fine sampling of the calorimeter enable cluster energies to be measured with a resolution of $\sigma(E)/E = 0.057/\sqrt{E}$ (GeV), as determined with the DC using Bhabha scattering events. The absolute time resolution $\sigma(t) = 57$ ps/ \sqrt{E} (GeV) is dominated by photoelectron statistics, which is well parameterised by the energy scaling law. A constant term of 140 ps has to be added in quadrature, as determined from $e^+e^- \rightarrow \gamma\gamma$, radiative ϕ decays and $\phi \rightarrow \pi^+\pi^-\pi^0$ data control samples. This constant term is shared between a channel-by-channel uncorrelated term and

a common term to all channels. The uncorrelated term is mostly due to the calorimeter calibration while the common term is related to the uncertainty of the event T_0 , arising from the DAΦNE bunch length and from the jitter in the trigger phase-locking to the machine RF. By measuring the average and the difference of $T_{\text{cl}} - R_{\text{cl}}/c$ for the two photons in $\phi \rightarrow \pi^+\pi^-\pi^0$ events, a similar contribution of about 100 ps for the two terms has been estimated. The cluster position is measured with a resolution of 1.3 cm in the coordinate transverse to the fibers, and, by timing, of $1.2 \text{ cm}/\sqrt{E} \text{ (GeV)}$ in the longitudinal coordinate.

3.2.3 The trigger

Event rates at DAΦNE, with a luminosity of $10^{32} \text{ cm}^{-2}\text{s}^{-1}$, amount to about 300 ϕ mesons per second and 3×10^4 Bhabha scattering events per second within the KLOE acceptance.

The trigger system [16] is based on local energy deposits in the calorimeter and hit multiplicity information from the drift chamber. It has been optimised to retain almost all ϕ decays and to provide efficient rejection of the two main sources of background: small angle Bhabha scattering events and particles lost from the DAΦNE beams, resulting in very high photon and electron fluxes in the interaction region. Moreover, all Bhabha scattering and $\gamma\gamma$ events produced at large polar angles are gathered for detector monitoring and calibration purposes, as well as cosmic-ray muon events, which cross the detector at a rate of ~ 3 kHz.

Since the DAΦNE bunch crossing period is $T_{\text{rf}} = 2.7$ ns, the KLOE trigger must operate in continuous mode. A two-level scheme has been adopted. A first-level trigger, T1, is produced with a delay of ~ 200 ns and is synchronised with the DAΦNE master clock. The T1 signal initiates conversion in the front-end electronics modules, which are subsequently read out following a fixed time interval of about $2.6 \mu\text{s}$, driven by the typical drift time of electrons in the drift chamber cells. After the arrival of a first-level trigger, additional information is collected from the drift chamber, which is used together with the calorimeter information as a second-level trigger, T2. It confirms the first-level trigger, initialises digitisation of the drift chamber electronics and starts the data acquisition readout. If no T2 signal arrives before the end of a $2.6 \mu\text{s}$ dead time, all readout is reset.

The T1 and T2 triggers are based on the topology of energy deposits in the calorimeter and on the number and spatial distribution of the drift chamber hits. Since ϕ decay events have a relatively high multiplicity, they can be efficiently selected by the calorimeter trigger by requiring two isolated energy deposits above a threshold of 50 MeV in the barrel and 150 MeV in the end-caps. Events with only two fired sectors in the same end-cap are rejected, because this topology is dominated by beam background. Moreover, 15 hits in the drift chamber within a time window of 250 ns from beam crossing are required. The trigger identifies Bhabha scattering events requiring two clusters with energy above 350 MeV (BBT). An event which satisfies at least one of the two above conditions, and is not recognised as a Bhabha event, generates the T1 signal⁴. The second-level trigger, T2, requires further multiplicity or geometrical conditions for the energy deposits in the

⁴ Part of the $e^+e^- \rightarrow e^+e^-$ events are gathered for detector monitoring and calibration.

calorimeter, or at least 120 drift chamber wire signals within a 1.2 μ s time window. At this level, the trigger recognises also cosmic-ray muon events by the presence of two energy deposits above 30 MeV in the outermost calorimeter layers (CRT). A fraction of about 80% of the cosmic ray events are identified and rejected at the trigger level with this technique.

To be more more precise, first level trigger is defined as

$$T1 = (TC1 \oplus TD1) \odot NOT(BBT) \quad (3.1)$$

and second level one is:

$$T2 = (TC2 \oplus TD2) \odot T1 \odot NOT(CRT) \quad (3.2)$$

where BBT and CRT are the BhaBha Trigger and the Cosmic Trigger, TC1 (TD1) and TC2 (TD2) are the first and second level EMC (DC) trigger signals, respectively.

EMC trigger

A *trigger sector* consist of six calorimeter columns (half calorimeter sector) , and their signal are added togherer. Sectors are organized in two series, *normal* and *overlap*, staggered by half a sector width each other. The outer layer (of the five) is used as CR detector. The analog signals from both sides A and B of each calorimeter sectors are compared with high (hi) and low (lo) programmable threshold, and added to generate the T signals:

$$T = (T_A^{lo} \odot T_B^{lo}) \odot (T_A^{hi} \oplus T_B^{hi}) \quad (3.3)$$

This logical combination minimizes the non uniformity in response due to light attenuation in the fibers. As already said, barrel has typical threshold value of 50 MeV, while endcaps have different threshold values to suppress machine background events, from 50 up to 150 MeV: low values are used far from beam axis (socalled *cold* region), and higher values close to the beam axis (*warm* and *hot* regions).

The EMC trigger operates in the first level mode, so TC2=TC1. Standard trigger condition requires two sectors above threshold according to:

$$TC1 = TC2 = B2 \oplus B1 \odot (W1 \oplus E1) \oplus (E1 \odot W1) \quad (3.4)$$

where *B* stands for hit barrel sector, *E* for east endcap and *W* for west endcap, while 1 and 2 stand for one or two hit sectors respectively.

DC trigger

The signals from the DC wires are summed plane-wise and are then organized in nine superlayer signals. Each superlayer signal is obtained from contiguous plane signals. The superlayer signal is clipped to a relatively small value of fired wires (~ 5) to suppress trigger generated by spiraling particles with a high multiplicity of hits. The final trigger decision is made by summing togheter the superlayer signals. The first-level trigger (TD1) requires 13 hits in a 250 ns time window; the second-level trigger (TD2) requires 120 hits in the following 850 ns. This is done to take into account the typical formation time of the DC signals, which is about 1 μ s.

3.2.4 FILFO

Further suppression of the DAΦNE background events and cosmic ray particles is performed by the FILFO⁵ filter [17]. FILFO is an offline filter used to recognise and reject cosmic rays, beam background events and Bhabha scattering events with electrons (positrons) emitted at small polar angles that interact with the low-beta focusing quadrupoles. To reject background events, cuts are applied on the number of EMC clusters, the number of DC hits, the total energy deposited in the calorimeter, the position of the most energetic clusters, and the ratio of the number of hits in the internal DC layers to the total number of hits. A prescaled 1/20 sample of unfiltered data is retained to control the filter efficiency.

3.2.5 The data acquisition

The data acquisition (DAQ) [18] has been designed to cope with a rate of 10^4 events per second, due to ϕ decays, downscaled Bhabha events, non vetoed cosmic rays and beam backgrounds. An average event size of 5 kbyte is estimated, corresponding to a total bandwidth requirement of 50 Mbyte/s. The DAQ readout system involves some 23000 channels of front-end electronics (FEE) from EMC, DC and trigger system. For each event, relevant data from the whole FEE system are concentrated in a single CPU where a dedicated process builds the complete event. A three-level scheme has been implemented. The first level reads data from single FEE crates. The second level combines information from different crates. The third level, responsible for final event building, relies on standard network media and protocols (TCP/IP).

3.3 Data reconstruction

The data reconstruction starts immediately after the completion of the calibration jobs. The reconstruction program, DATAREC [17], provides additional data-quality and monitoring information, and consists of several modules, among which EMC reconstruction, DC reconstruction, track-to-cluster association and event-classification streaming procedure.

3.3.1 Cluster reconstruction

The calorimeter is segmented into 2440 cells, which are read out by photomultipliers (PMTs) at both ends (A, B). This segmentation provides the determination of the position of energy deposits in r - ϕ for the barrel and in x - y for the end-caps. Both charges $Q_{\text{ADC}}^{\text{A,B}}$, from ADCs, and times $t_{\text{TDC}}^{\text{A,B}}$, from TDCs, are recorded. For each cell, the particle arrival time t and its coordinate s along the fiber direction (the zero being taken at the fiber

⁵ FILtro di FOndo

center) are obtained using the times at the two ends as:

$$\begin{aligned} t(\text{ns}) &= \frac{1}{2}(t^{\text{A}} + t^{\text{B}} - t_0^{\text{A}} - t_0^{\text{B}}) - \frac{L}{2v} , \\ s(\text{cm}) &= \frac{v}{2}(t^{\text{A}} - t^{\text{B}} - t_0^{\text{A}} + t_0^{\text{B}}) , \end{aligned}$$

with $t^{\text{A,B}} = c^{\text{A,B}} \times t_{\text{TDC}}^{\text{A,B}}$, where $c^{\text{A,B}}$ are the TDC calibration constants, $t_0^{\text{A,B}}$ denotes overall time offsets, L stands for the length of the cell (cm) and v for the light velocity in the fibers (cm/ns). The energy on each side of a cell i is obtained as:

$$E_i^{\text{A,B}}(\text{MeV}) = k_E \times g_i(s) \times \frac{S_i^{\text{A,B}}}{S_{i,\text{MIP}}^{\text{A,B}}} ,$$

here $S = Q_{\text{ADC}} - Q_{0,\text{ADC}}$ is the charge collected after subtraction of the zero-offsets, the so-called ADC pedestals, and S_{MIP} is the response to a minimum ionising particle crossing the calorimeter center. The correction factor $g(s)$ accounts for light attenuation as a function of the impact position s along the fiber, while k_E is the energy scale factor, obtained from showers of particles of known energy (see sec. 3.2.2). The cell energy E_i is taken as the mean of the energies at each end:

$$E_i(\text{MeV}) = \frac{E_i^{\text{A}} + E_i^{\text{B}}}{2} .$$

The calorimeter reconstruction starts by applying the calibration constants to transform the measured quantities Q_{ADC} and t_{TDC} into the physical quantities s and t . Position reconstruction and energy/time corrections are applied to each fired cell. Then a clustering algorithm looks for groups of cells contiguous in r - ϕ or x - y and groups them into pre-clusters. In a second step, the longitudinal coordinates and arrival times of the pre-clusters are used for further merging and/or splitting. The cluster energy, E_{cl} , is the sum of the energies of all cells assigned to a cluster. The cluster position, $(x, y, z)_{\text{cl}}$, and time, t_{cl} , are evaluated as energy-weighted averages over the contributing cells. Cells are included in the cluster search only if times and amplitudes are available on both sides; otherwise, they are recorded as *incomplete* cells. The available information from most of the incomplete cells is added to the existing clusters at a later stage, by comparing the positions of such cells with the cluster centroid.

3.3.2 Track reconstruction

Track reconstruction is performed in three steps: pattern recognition, track fit, and vertex fit. Each step is managed separately and produces the inputs for the subsequent step.

Pattern recognition The pattern recognition algorithm searches for track candidates.

It begins by associating hits, working inward from the outermost layer of the DC, and then obtains track segments and approximate trajectories parameters. The DC wires form alternating positive and negative stereo angles with respect to the z

direction. When the hits are projected on the x - y plane, they are seen in the stereo views as two distinct images. The pattern recognition procedure first associates separately the hits of each projection, using only two dimensional information, and in a second step combines the track candidates of the two views, according to their curvature values and geometrical compatibility.

Track fit The track-fit procedure minimises the function $\chi_{\text{trk}}^2 = \sum_{i=1}^n (d_i - d_i^{\text{fit}})^2 / \sigma_i^2$ defining the comparison between the measured and the expected drift distance for each hit. In this formula, n is the number of hits, $d_i(t_{\text{drift}})$ is the drift distance, obtained via the space-to-time relation from the measured drift time (see sec. 3.2.1), d_i^{fit} is the result of the fit and σ_i is the estimate of the hit resolution. The procedure is iterative because the space-to-time relation depends on the track parameters. At each tracking step, the effects due to energy loss and multiple scattering are estimated.

Vertex fit The track parameters are used to look for primary and secondary vertices. For each track pair, a χ_{vtx}^2 function is computed from the distances of closest approach between tracks; the covariance matrices from the track-fit stage are used to evaluate the errors. The vertex position is obtained minimising the χ_{vtx}^2 .

3.3.3 Track-to-cluster association

The track-to-cluster association module makes correspondences between tracks in the DC and clusters in the EMC. The procedure starts by assembling the reconstructed tracks and vertices into decay chains and by isolating the tracks at the end of these chains. For each of these tracks, the measured momentum and the position of the last hit ⁶ in the DC are used to extrapolate the track to the EMC. The extrapolation gives the track length L_{ex} from the last hit in the chamber to the calorimeter surface, and the momentum \vec{p}_{ex} and the position \vec{x}_{ex} of the particle at the surface. The resulting impact point is then compared with the positions \vec{x}_{cl} of the reconstructed cluster centroids. A track is associated to a cluster if the distance to the centroid in the plane orthogonal to the direction of incidence of the particle on the calorimeter, $D = |(\vec{x}_{\text{cl}} - \vec{x}_{\text{ex}}) \cdot \vec{p}_{\text{ex}} / |\vec{p}_{\text{ex}}||$, is less than 60 cm.

3.3.4 Event classification

At reconstruction level, data are divided into streams [17]. A stream is a collection of events which are all identified by a streaming algorithm. The streaming algorithms are implemented in such a way that they are not mutually exclusive, *i.e.* one event can be found in none, one or more than one stream.

The streaming procedure has been introduced in the KLOE event reconstruction in order to save CPU time and disk space. In fact, among all the acquired events during the data taking, only part of them belongs to, e.g., the $\phi \rightarrow K_L K_S$ or $\phi \rightarrow \eta\gamma$ categories.

⁶ Conventionally, the last hit is defined as the track hit closest to the DC outer wall.

Furthermore, by exploiting the streaming algorithms, users can select a definite stream for analysis purposes. For this thesis the interesting stream is $\phi \rightarrow K_L K_S$.

There are seven streams defined:

- stream 1 (KPM): $\phi \rightarrow K^+ K^-$;
- stream 2 (KSL): $\phi \rightarrow K_L K_S$;
- stream 3 (RPI): $\phi \rightarrow \rho\pi, \pi^+ \pi^0 \pi^-$;
- stream 4 (RAD): ϕ radiative decays, such as $\phi \rightarrow \eta\gamma, \phi \rightarrow \eta'\gamma, \phi \rightarrow \pi^0\gamma$;
- stream 5 (CLB): Bhabha scattering and cosmic-ray muon events, to be used for calibration. This streams also contains $e^+e^- \rightarrow \mu^+\mu^-$ and $e^+e^- \rightarrow \pi^+\pi^-$ events;
- stream 6 (UFO): unidentified events;
- stream 7 (BHA): Bhabha scattering events.

The KSL stream has eight sub-algorithms, each devoted to identify a special kaon decay and/or special event topologies. Among these algorithms, the KLCRASH has been chosen to identify K_L interactions in the calorimeter, through which K_S mesons are tagged, as described in next chapter.

3.4 Beam parameters and luminosity

KLOE monitors continuously the beams working point, providing continuous feedback to DAΦNE. The most important parameters are the beam energies, beam energy spread, beam crossing angle and the e^+e^- transverse momentum, which are obtained from the analysis of Bhabha scattering events with electron (positron) polar angles $45^\circ < \theta < 135^\circ$.

The average value of the center-of-mass energy is evaluated online during data taking with a precision of ~ 50 keV for each 200 nb^{-1} of integrated luminosity. This determination is further refined with offline analysis to achieve a precision of ~ 20 keV, as discussed later. The average position of the e^+e^- interaction primary vertex, with coordinates x_{PV} , y_{PV} , and z_{PV} , is reconstructed run-by-run from the same sample of Bhabha events. x_{PV} and y_{PV} are determined with an accuracy of about $10 \mu\text{m}$, and have widths Δx_{PV} and Δy_{PV} which are about 1 mm and few tens of microns, respectively. z_{PV} is also reconstructed online with $100\text{--}200 \mu\text{m}$ accuracy, and has a width of $12\text{--}14$ mm, determined by the bunch length.

An improved determination of the center-of-mass energy is obtained for each run by fitting the e^+e^- invariant mass distribution for Bhabha scattering events to a Monte Carlo generated function, which includes radiative effects from initial- and final-state radiation corrections (ISR, FSR). The absolute energy scale is calibrated by measuring the visible cross section for $\phi \rightarrow K_L K_S$ events. The cross section peak is fit to a theoretical function, which depends on the ϕ parameters, takes into account the effect of ISR, and includes the interference with the $\rho(770)$ and the $\omega(782)$ mesons. The ϕ mass, total width, and peak

cross section are the only free parameters of the fit, the $\rho(770)$ and $\omega(782)$ parameters being fixed. The ϕ mass value obtained from the fit is compared with the ϕ -meson mass measured with high precision by the CMD-2 experiment at VEPP-2M [19]. The ratio $m_\phi^{\text{CMD}}/m_\phi^{\text{KLOE}} = 1.00015$ is then used to correct the KLOE determination of the center-of-mass energy.

An online luminosity measurement is performed by selecting a sample of Bhabha scattering events within the acceptance of the barrel calorimeter, asking for two trigger sectors fired and using a high energy threshold. These selections allow to strongly reduce machine background and to provide DAΦNE with a value of the instantaneous luminosity with a 3% statistical uncertainty, when operating at $L = 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$.

A more accurate measurement of the integrated luminosity is performed offline [20], by selecting Bhabha scattering events with polar angle $55^\circ < \theta < 125^\circ$, the so-called very large-angle Bhabha (VLAB) events. The effective cross section for these events, about 430 nb, is large enough to reduce the statistical error at a negligible level. The integrated luminosity is obtained by counting the number of VLAB candidates, N_{VLAB} , and normalising it to the effective Bhabha scattering cross section, $\sigma_{\text{VLAB}}^{\text{MC}}$, obtained from MC simulation, after subtraction of the background

$$L_{\text{int}} = \frac{N_{\text{VLAB}}}{\sigma_{\text{VLAB}}^{\text{MC}}} (1 - \delta_{\text{bkg}})$$

This method allows to estimate the integrated luminosity with a 0.3% accuracy.

3.5 MC event simulation

The KLOE Monte Carlo program, *GEANTFI* [17], is based on the *GEANT3.21* library, and incorporates a detailed description of the interacting region, the drift chamber, the calorimeter (endcaps and barrel), and the superconducting magnet with the return yoke structure. Specific routines, starting from the *GEANT* routines for particle-tracking and energy-deposition, are developed to simulate the response of the KLOE detector. For the physics process generation, *BHAGEN* and *BABAYAGA* are used to generate Bhabha events; *PHOKHARA* and *EVA* are also used. Beside the most common ϕ , K^\pm and K^0 decays, more rare decays are also simulated. For kaon hadronic, semileptonic and leptonic decays, radiative corrections are also taken into account. Whenever a decay is generated in which the radiated energy is more than 0.1 MeV, a final state photon is explicitly simulated.

A detailed simulation of the detector activity due to accidental coincidences produced by beam backgrounds is required in order to obtain an accurate control of systematic effects, as needed for most KLOE analyses. This activity consists mainly of noise hits in the DC and low-energy clusters in the EMC, mostly produced at small polar angles. Background hits in the drift chamber and calorimeter are added to the simulated events at the event-reconstruction stage. For the analysed data set, this background was obtained from $e^+e^- \rightarrow \gamma\gamma$ events satisfying specific topological cuts. These events are selected from KLOE data with a cross-section of ~ 40 nb. Since $e^+e^- \rightarrow \gamma\gamma$ events are fully neutral, all

DC hits in these events are considered background, in addition to all EMC clusters not identified as belonging to the $\gamma\gamma$ topology (care is taken to correctly distinguish clusters from initial state radiation or from cluster splitting, which actually belong to the $\gamma\gamma$ topology, from those due to beam-generated background). A file containing background hits is created for each raw-data file in the data set.

3.5.1 MC decay channels

The following decays channel are the most important in this work (related colour in the legend in parenthesis):

- $\phi \rightarrow K^+K^-$ (green)
- $\phi \rightarrow \pi^+\pi^-\pi^0$ (indigo)
- $K_S \rightarrow \pi^+\pi^-$ (blue)
- $K_S \rightarrow \pi^0\pi^0$ (light blue)
- $K_S \rightarrow \pi^\pm e^\pm \nu$ (red)
- $K_S \rightarrow \pi^\pm \mu^\pm \nu$ (light brown/dark green)
- $K_S \rightarrow \pi^+\pi^-e^+e^-$ (dark gray)
- other K_S decay channel (light gray)

All decay channels are inclusive in emitted photons from charged particle, and the same criteria is used for the data sample.

Chapter 4

Analysis scheme

The plan is to measure the branching ratio (BR) of the K_S semileptonic decay $K_S \rightarrow \pi\ell\nu$ normalising to $K_S \rightarrow \pi^+\pi^-$ decay through:

$$BR(K_S \rightarrow \pi\ell\nu) = \frac{N_{\pi\ell\nu}}{\epsilon_{\pi\ell\nu}} \times \frac{\epsilon_{\pi^+\pi^-}}{N_{\pi^+\pi^-}} \times R_\epsilon \times BR(K_S \rightarrow \pi^+\pi^-) \quad (4.1)$$

where $\ell = e, \mu$; $N_{\pi\ell\nu}$ and $N_{\pi^+\pi^-}$ are the number of $K_S \rightarrow \pi\ell\nu$ (analysis sample) and $K_S \rightarrow \pi^+\pi^-$ (normalisation sample), counted after the selection, which is the same for the two decays only in the first part of the analysis; R_ϵ is the ratio of the efficiencies for the common selections of the two decays, while $\epsilon_{\pi\ell\nu}$ and $\epsilon_{\pi^+\pi^-}$ are the efficiencies of the selections not in common; $BR(K_S \rightarrow \pi^+\pi^-)$ is the branching ratio measured by KLOE [5].

$$BR(K_S \rightarrow \pi^+\pi^-) = (69.020 \pm 0.051)\% ,$$

which is dominating the value quoted by the PDG, $BR(K_S \rightarrow \pi^+\pi^-) = (69.02 \pm 0.05)\%$. All analysis is inclusive in emitted photons from charged particle.

The analysis scheme is divided in two parts. The first part is the same for the analysis ($K_S \rightarrow \pi\ell\nu$) and the normalisation ($K_S \rightarrow \pi^+\pi^-$) samples and consists of:

- Trigger and FILFO selection (Sec. 3.2.3) and the event classification (Sec. 3.3.4);
- K_L crash selection and K_S identification (Sec. 4.2).

The second part is different for the two samples:

- Simple cuts to select the normalisation sample and count of the number of $K_S \rightarrow \pi^+\pi^-$ decays (Sec. 5);
- Selection of the analysis sample from the background (Sec. 6.1) based on:
 - Track-to-cluster association (Sec. 3.3.3) ;
 - Preselection (Sec. 6.1.1);
 - Selection of variables for the multivariate analysis (Sec. 6.1.2);
 - Multivariate analysis (Sec. 6.1.2);
 - Time of flight analysis (Sec. 6.1.3).

4.1 Data sample

The measurement of the branching ratios for both $K_S \rightarrow \pi e \nu$ and $K_S \rightarrow \pi \mu \nu$ decays was performed on the data collected during 2004–2005. Data were reconstructed using the official reconstruction algorithms and classified into different streams, and events from the KSL stream were analysed (see Sec. 3.3). The runs used in the analysis are from run 30300 to run 41902, reconstructed with DATAREC version 26, with total luminosity 1.63 pb^{-1} . For the analysis all the Monte Carlo production was used corresponding to a total luminosity scale factor $\times 2$.

4.2 K_L -crash and K_S identification

The first step of the analysis consists in identifying a pair of neutral kaons. This is performed with the tagging method: the identification of one neutral kaon in one hemisphere guarantees the presence of the other kaon in the opposite hemisphere. Since the interest is in the K_S decay, the tag is performed by detecting the K_L meson, via its interaction in the calorimeter that provides the most efficient and unbiased way. Neutral kaons are produced with a velocity in the ϕ rest frame $\beta^* \approx 0.22$, corresponding to a time of flight of about 31 ns from the interaction point to the calorimeter. The different lifetime of the two neutral mesons, ($\tau_S = 0.089 \text{ ns}$, $\tau_L = 51.0 \text{ ns}$) causes that the K_S decay close to the interaction point, with an mean flight path of 6 mm, while the K_L reach the calorimeter before decaying in about 60% of the cases. The interaction of the K_L meson (called K_L -crash in the following) releases energy up to $\sim 500 \text{ MeV}$; this signature, together with the delayed signal, provides a clear tag of the presence of the K_S meson. The position of the K_L -crash cluster permits to reconstruct the direction of the K_L with $\sim 1^\circ$ accuracy and, knowing the ϕ four-momentum, to estimate event-by-event the four-momentum of the tagged K_S meson.

4.2.1 K_L -crash

The recognition of the K_L -crash cluster is accomplished requiring the presence of a neutral cluster, not associated to any track, with an energy: $E_{\text{clu}}(K_L\text{-crash}) > 100 \text{ MeV}$. After detecting such a cluster, the velocity is computed in the laboratory frame: $\beta_{\text{cl}} = R_{\text{cl}}/ct_{\text{cl}}$ where R_{cl} is the distance of the cluster from the interaction point, t_{cl} the time of flight and c is the velocity of light. In the case of more than one cluster fulfilling the above criteria, the cluster with the smaller time of arrival is selected as the K_L -crash cluster. Since the velocity of the kaon in the ϕ rest frame is fixed ($\beta^* \approx 0.22$), β_{cl} is transformed accordingly

$$\beta^* = \frac{\sqrt{\beta_\phi^2 + \beta_{\text{cl}}^2 + 2\beta_\phi\beta_{\text{cl}}\cos\alpha}}{1 - \beta_\phi\beta_{\text{cl}}\cos\alpha}, \quad (4.2)$$

where α is the angle between the ϕ momentum and the K_L -crash direction.

The β^* distribution peaks at two different values. This is due to the procedure to determine the event T_0 , that assumes the first cluster in time to be produced by a prompt photon. For the decays we are looking for, this hypothesis is wrong in most of the cases, because of first cluster is generated either by an electron or by a charged pion. Electrons reach the calorimeter about 3 ns before charged pions, so on average they are associated to the preceding bunch crossing, causing the double-peak structure of the distribution. The MC simulation accounts for this effect, but the MC resolution of β^* does not reproduce well the data. Thus a smearing procedure on the MC is applied to improve the MC–data agreement:

$$\beta^{*'} = \beta^* \times \text{ran}G(\mu, \sigma) \quad (4.3)$$

where β^* is before the smearing procedure, $\beta^{*'}$ is after the smearing procedure (called just β^* from now on) and $\text{ran}G$ is a Gaussian distribution with mean value $\mu = 1$ and $\sigma = 0.0012$ determined by minimising the difference between the data and MC distribution. The result of the smearing procedure is shown in Figure 4.1. A cut on the β^* value was applied to select both peaks

$$0.17 < \beta^* < 0.28 \quad (4.4)$$

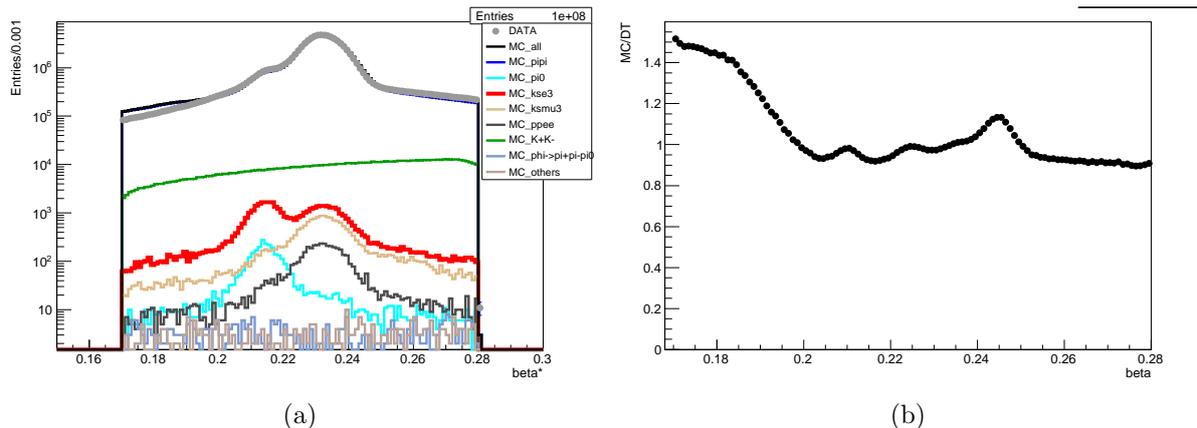


Figure 4.1: (a): Distribution of β^* for the data and MC signal events, background events and their sum; (b) ratio of MC to data

Determination of the K_L momentum

After the K_L identification, its four-momentum is determined based on the two-body decay in the ϕ rest frame, knowing the four-momentum P_ϕ of the ϕ meson and the angle between the ϕ and K_L directions. The average value of P_ϕ is determined run-by-run by large angle Bhabha scattering events and P_{K_L} by the centroid of the K_L -crash cluster. The four-momentum of the tagged K_S meson is $P_{K_S} = P_\phi - P_{K_L}$.

4.2.2 K_S selection

After K_S tagging, the goal is to identify the K_S decay of the signal candidates and the normalisation channel, $K_S \rightarrow \pi^+\pi^-$. The requirement is to have two tracks of opposite charge forming a vertex in a cylinder

$$\rho_{\text{VTX}} = \sqrt{x_{\text{VTX}}^2 + y_{\text{VTX}}^2} < 5 \text{ cm} \quad |z_{\text{VTX}}| < 10 \text{ cm}.$$

This requirement allows reducing the background from $K_S \rightarrow \pi^0\pi^0$ (absence of vertex) and $\phi \rightarrow K^+K^-$ (long decay path) with high efficiency for both the signal and normalisation samples.

Chapter 5

Normalisation sample

After the K_L -crash recognition and the vertex requirement, the fraction of $K_S \rightarrow \pi^+\pi^-$ events in the selected sample is 93%, as predicted by simulation. It is then easy to count the number of normalisation events with high efficiency and small background. This is done applying a cut on the momentum distribution (Figure 6.3) of the two tracks connected to the vertex

$$140 < p < 280 \text{ MeV}$$

with a MC efficiency of 97.4%, reaching a purity of 99.9%. Result is:

$$N_{\pi^+\pi^-} = 282\,314\,000 \pm 17\,000$$

5.1 Efficiency

The normalisation sample $K_S \rightarrow \pi^+\pi^-$ is selected with a cut on the momentum distribution of the two tracks that form the K_S vertex: $140 < p_i < 280 \text{ MeV}$, where $i = 1, 2$ are the two tracks. In this case, the control sample is a subsample of the analysis sample, selected to maximise the purity. The selection of the control sample is made with a cut on the position of the reconstructed vertex, ρ_{VTX} , that is the most independent variable from the tracks momenta used for the selection of the normalisation sample, as shown in Figure (5.1).

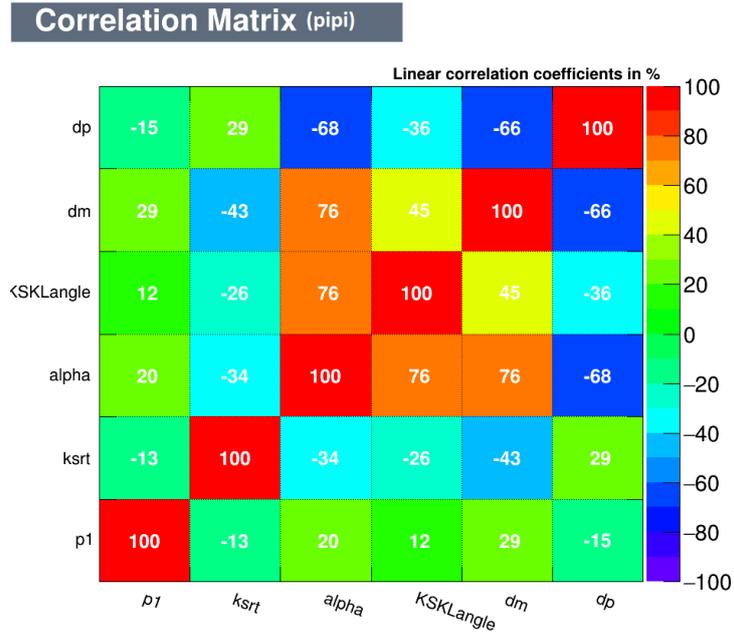
The correlation between ρ_{VTX} and the particles momentum is 13%. Cutting with different values of ρ_{VTX} it is possible to select different control samples, and compute the efficiency for each sample with Eq. (6.7). The results are shown in Table 5.1.

A small dependence of the efficiency computed from selecting different ρ_{VTX} ranges is observed. For this reason, the efficiency used in the analysis is the extrapolation to the value of the ρ_{VTX} cut used in the analysis of the signal. The efficiency is:

$$\epsilon_{\pi\pi}^{DT} = (96.569 \pm 0.004)\%$$

A second way to compute this efficiency is directly from the analysis data sample, correcting for the purity. Using this method (*i.e.* using Eq. (6.7) with $\epsilon_{AS}^{MC} = \epsilon_{CS}^{MC}$) the result is:

$$\epsilon_{\pi\pi}^{DT} = (96.657 \pm 0.002)\%$$

Figure 5.1: Correlation matrix for some $\pi^+\pi^-$ variables

The agreement between the two results is at the level of permil. For the simplicity of the method, that minimises the possibility of biasing the result, the second value is preferred, and the difference between the two values is taken as systematic error. Finally:

$$N_{\pi^+\pi^-}/\epsilon_{\pi^+\pi^-} = (292.10 \pm 0.26) \times 10^6$$

ρ_{VTX} [cm]	purity %	$140 < p < 280$ MeV %
$\rho_{\text{VTX}} < 5$	96.79	96.657 ± 0.002
$\rho_{\text{VTX}} < 4$	97.12	96.772 ± 0.002
$\rho_{\text{VTX}} < 3$	97.63	96.933 ± 0.002
$\rho_{\text{VTX}} < 2$	98.23	97.125 ± 0.002
$\rho_{\text{VTX}} < 1.5$	98.58	97.234 ± 0.002
$\rho_{\text{VTX}} < 1$	98.96	97.339 ± 0.002

Table 5.1: Efficiency for selecting $K_S \rightarrow \pi^+\pi^-$ events as function of the cut on the two-tracks vertex.

Chapter 6

$K_S \rightarrow \pi e \nu$

6.1 Selection

After the K_L -crash recognition and the vertex requirement, the data sample contains 301,645,500 events and its composition, as predicted by simulation, is shown in Table 6.1

	MC events	fraction [%]
DATA	301 645 500	
MC	312 018 500	
$K_S \rightarrow \pi e \nu$	259 264	0.08
$K_S \rightarrow \pi^+ \pi^-$	301 976 400	96.78
$\phi \rightarrow K^+ K^-$	9 565 465	3.07
$K_S \rightarrow \pi^0 \pi^0$	30 353	0.01
$K_S \rightarrow \pi \mu \nu$	139 585	0.04
$K_S \rightarrow \pi^+ \pi^- e^+ e^-$	18 397	$6 \cdot 10^{-3}$
$\phi \rightarrow \pi^+ \pi^- \pi^0$	24 153	$8 \cdot 10^{-3}$
others	4 852	$2 \cdot 10^{-3}$

Table 6.1: Number of events by channel after K_L -crash and K_S selection

The main background is $K_S \rightarrow \pi^+ \pi^-$, because of the large branching ratio, about 1000 times the semileptonic branching ratio. The strategy is to perform first a multivariate analysis, using as input only variables built from DC information; then to use calorimeter information to exploit as discriminant the time of flight difference between electrons and charged pions. Before the multivariate analysis, some cuts are applied to the variables used as input to improve the agreement between MC and data. For exploiting the time of flight for e - π separation, track-to-cluster association is required for both tracks.

6.1.1 Preselection

The first selection requirement is the track-to-cluster association (TCA): both tracks are required to be associated to a cluster.

The distribution of the polar angle is shown in Fig. 6.1 for the clusters associated to the two tracks and for the neutral K_L -crash cluster. The two clusters are selected with polar angle

$$\theta_{\text{clu}}^i > 15^\circ$$

to avoid the end-cap calorimeter region close the beam pipe, where beam backgrounds are higher. The same requirement is applied to the K_L -crash cluster (Fig. 6.1):

$$\theta_{\text{crash}} > 15^\circ$$

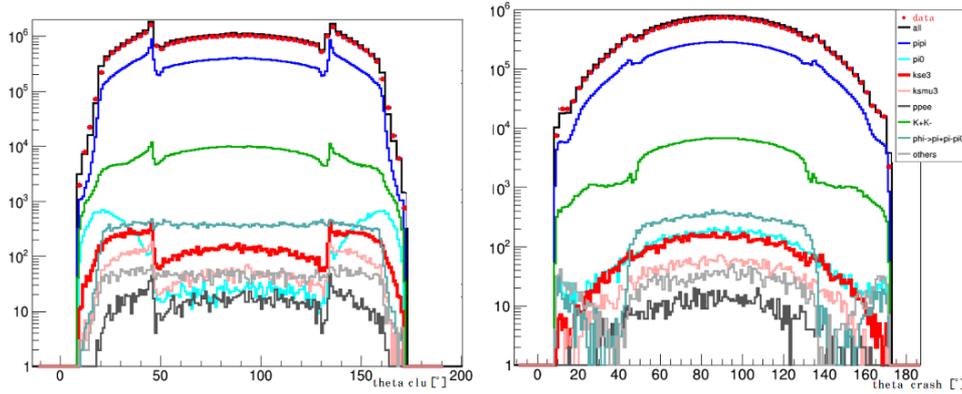


Figure 6.1: Theta angle distribution for the cluster associated to K_S daughter track (a) and for the K_L -crash cluster (b)

Moreover, to ensure a similar selection for the $K_S \rightarrow \pi e \nu$ sample and a data control sample from the abundant $K_L \rightarrow \pi e \nu$ decay a momentum cut is applied on δp , the difference between the absolute value of the momentum of the K_S , reconstructed from the two tracks, and value reconstructed from the tag.

$$\delta p > -95 \text{ MeV}$$

The introduction and use $K_L \rightarrow \pi e \nu$ control sample is discussed in Sec. 6.2.

At this stage, after the requirement on the cluster polar angles, the track to cluster association and the track momentum selection, the number of events in the data and MC samples is 212,720,400 and 234,516,600, respectively.

6.1.2 Multivariate analysis

The multivariate analysis consists of an algorithm applied in a multidimensional space instead of multiple selections on single variables. The algorithm (classifier in the following) exploits some variables in input, and gives one single value in output, the classifier value. This value is function of all the variables in input, chosen to give good separation between signal and background. For this, the classifier needs to be trained with simulation, separately for signal and background, to *learn* how to separate the two categories.

The algorithm is trained with simulated event samples, using a high purity sample for the signal and another sample with all the background channels. Three main steps are followed to perform a multivariate analysis:

- training of the classifier with two separate samples (signal and background);
- test the classification algorithm on a MC sample containing both signal and background, to check if it is working properly;
- run it on the data sample.

Number of events used for training and testing are listed in App. A (Tab. A.1). In this analysis, the algorithm used is a boosted decision tree (BDT), for its good performance and short time of training. More information on BDT are in App. A.

Because the training is done with MC samples, but the interest is in classifying the data sample, the power of a multivariate classifier depends strongly on the agreement between data and MC. If the distributions of the input variables are not in agreement, the classification could not work properly. The first step is to select the variables to be used. The choosing criteria are a high separation power between signal and background, and a low correlation among the variables. The following five variables are chosen:

- two tracks momenta (p_i);
- angle between the two tracks in the center of mass of the K_S (α);
- angle between the K_S and the K_L -crash (α_{SL});
- difference between the momentum of the K_S , reconstructed from the two tracks, and the momentum determined from the tag (δp);
- difference δm between the invariant mass reconstructed from the two tracks ($M_{inv}(pipi)$), in the hypothesis of pion mass, and the K_S mass (497.65 MeV).

The distributions of these variables for data and MC are shown in Figures 6.3–6.7, labeled with different colours for the different channels. The correlation between the variables for the signal and background samples is shown in Fig. 6.2.

The distributions in Figures 6.3–6.7 show that there are regions where data and MC are not in good agreement. This has two main motivations:

- processes not simulated (or not simulated properly) in the MC;
- accidental background events, that are inserted in the MC *ad hoc*.

However, the disagreement appears in regions that do not affect much the signal and where the background contribution is small, this allows to apply cuts to avoid those regions. Two cuts are applied:

- $p < 320$ MeV for both tracks;

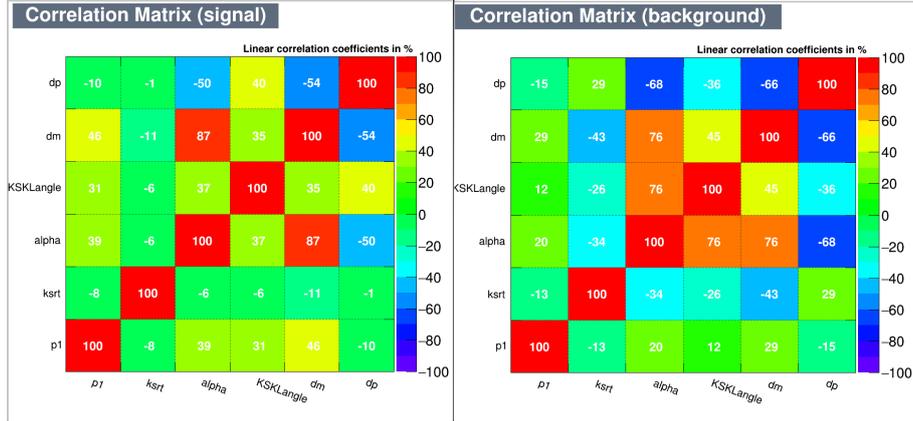


Figure 6.2: Correlation matrix for signal (left) and background (right).

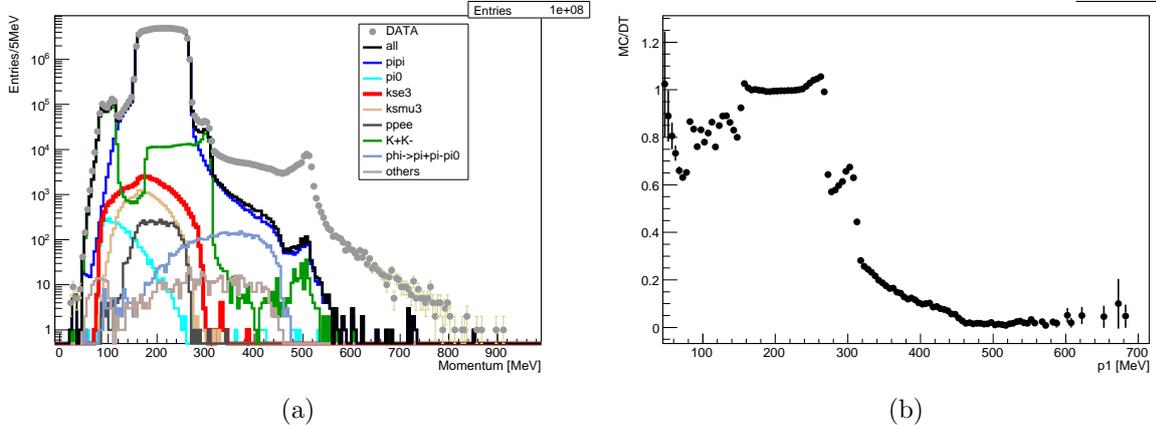


Figure 6.3: (a) Distribution of the momentum of the tracks; (b) ratio of MC to data.

	MC events	fraction [%]	relative efficiency [%]
DATA	211 368 600		70.07
MC	234 250 000		75.08
$K_S \rightarrow \pi e \nu$	120 171	0.05	46.35
$K_S \rightarrow \pi^+ \pi^-$	231 861 100	98.98	76.78
$\phi \rightarrow K^+ K^-$	2 211 535	0.94	23.12
$K_S \rightarrow \pi^0 \pi^0$	7 965	$3 \cdot 10^{-3}$	26.24
$K_S \rightarrow \pi \mu \nu$	46 706	0.02	33.46
$K_S \rightarrow \pi^+ \pi^- e^+ e^-$	10 993	$5 \cdot 10^{-3}$	59.75
$\phi \rightarrow \pi^+ \pi^- \pi^0$	786	$8 \cdot 10^{-4}$	3.25
others	752	$3 \cdot 10^{-4}$	15.49

Table 6.2: Number of events by channel after preselection

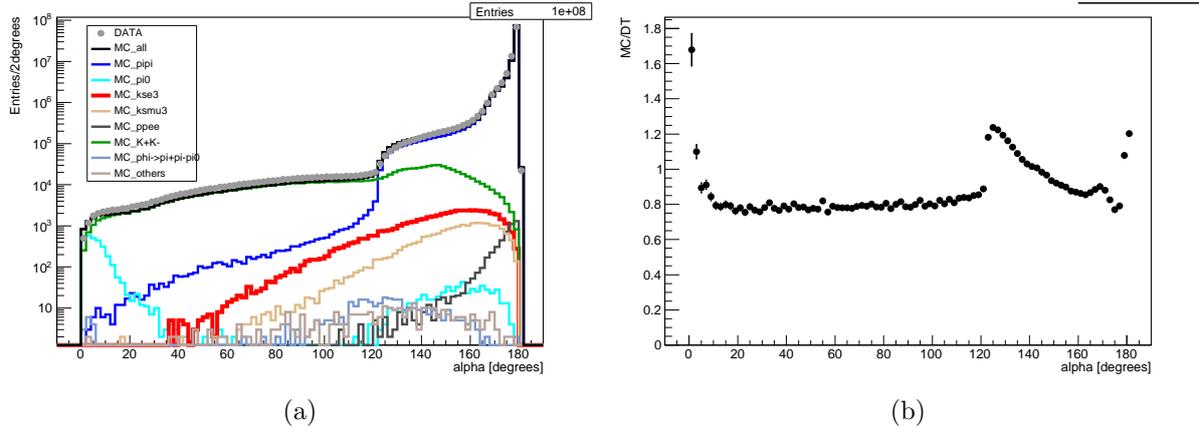


Figure 6.4: (a) Distribution of the angle between the tracks momenta in the K_S rest frame; (b) ratio of MC to data.

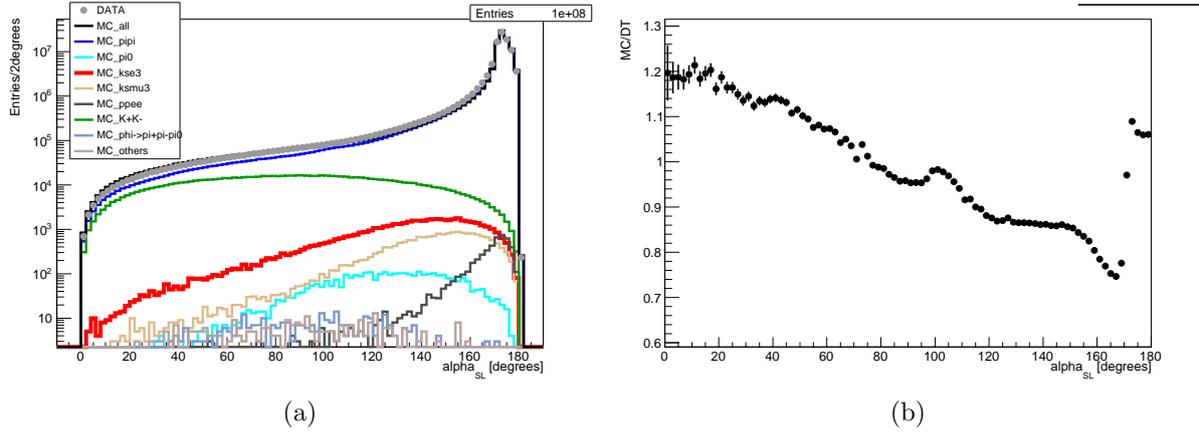


Figure 6.5: (a) Distribution of the angle between the K_L and K_S -crash; (b) ratio of MC to data.

- $\delta p < 190$ MeV.

The number of data and MC events are reduced to 211,368,600 and 234,250,000, respectively, and its composition as predicted by simulation is listed in Table 6.2.

After this, the classifier is ready to be trained. Details on the values of the BDT parameters are presented in Appendix. The signal training was done on 15,000 events, while the background test on 50,000 events. These numbers were chosen as a compromise to have a statistically significant sample, short time for the training and to avoid overtraining. Similar samples were chosen for the test. After training and testing, the classification was run on all events of the MC and data sample. The results are shown in Figure 6.8. The BDT classifier gives high BDT values for the signal and lower BDT values for background events. However, due to the high contamination of background in the signal sample, a

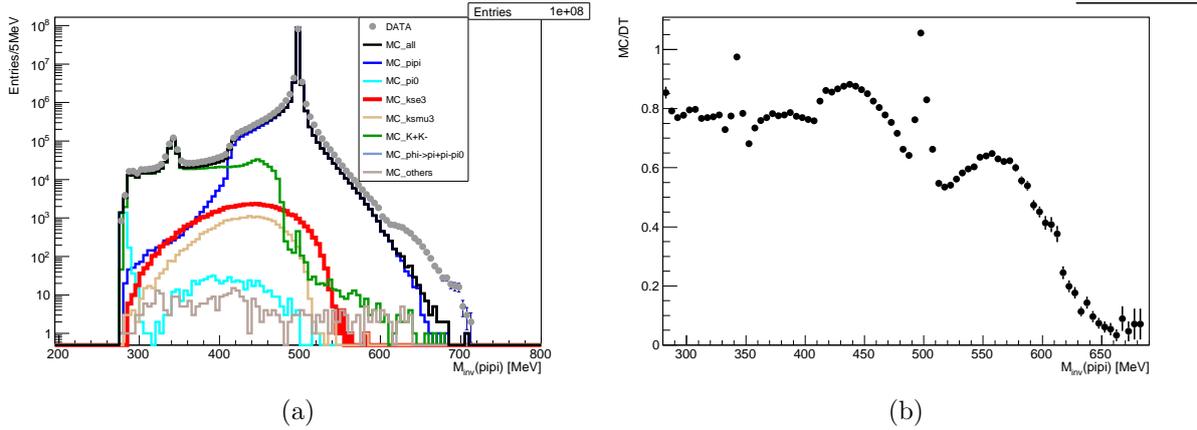


Figure 6.6: (a) Distribution of the invariant mass reconstructed with the two tracks momenta in the hypothesis of charged pion mass $M_{inv}(pipi) = \delta m + 497.65 \text{ MeV}$; (b) ratio of MC to data.

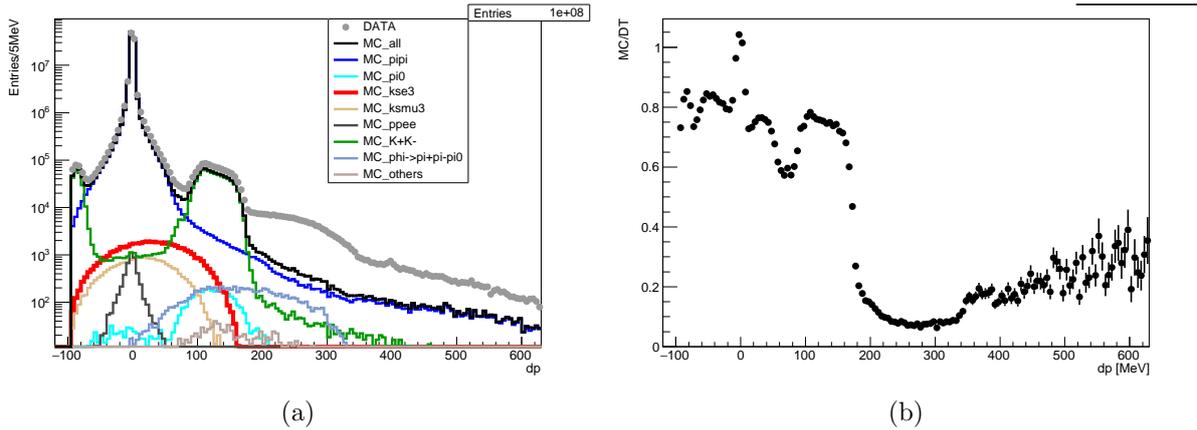


Figure 6.7: (a) Distribution of the difference between the K_S momentum reconstructed from the tag and the "nominal" momentum; (b) ratio of MC to data.

separation based on the BDT output is not sufficient.

The agreement between the data and MC distributions is not good enough. The agreement gets better for high BDT values, when the background contamination (mainly from $K_S \rightarrow \pi^+ \pi^-$) is lower. This has two main important consequences:

- the efficiency of a BDT-based selection, as determined from MC, cannot be trusted, thus a control sample is needed to evaluate the efficiency;
- a high value for the BDT cut is needed, to reject a large fraction of background and to improve the agreement between MC and data, without losing too much signal.

To define the cut of the BDT output variable, that ensures good stability and minimises

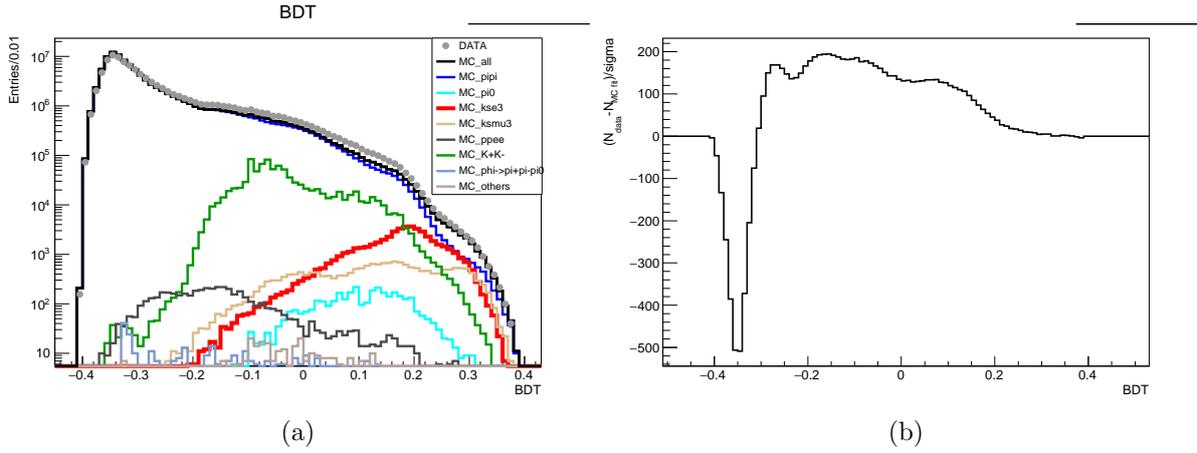


Figure 6.8: (a): Distribution of the BDT classifier output for the data and MC signal events, background events and their sum; (b) significance defined as the difference between data and MC divided by the statistical error.

the systematic error of the selection, a scan over a wide interval was performed, $0.135 < BDT < 0.170$, corresponding to seven BDT standard deviation (Sec. 6.3.1). The value chosen is

$$BDT > 0.15 . \tag{6.1}$$

After this selection, the number of events in the data sample is 788,238 and its composition as predicted by simulation is listed in Table 6.3

	MC events	fraction [%]	relative efficiency [%]
DATA	788 328		0.37
MC	645 096		0.28
$K_S \rightarrow \pi e \nu$	79 684	12.35	66.18
$K_S \rightarrow \pi^+ \pi^-$	467 037	72.40	0.20
$\phi \rightarrow K^+ K^-$	73 912	11.46	3.20
$K_S \rightarrow \pi^0 \pi^0$	2 359	0.37	28.24
$K_S \rightarrow \pi \mu \nu$	21 708	3.37	46.40
$K_S \rightarrow \pi^+ \pi^- e^+ e^-$	283	0.04	2.57
$\phi \rightarrow \pi^+ \pi^- \pi^0$	43	< 0.01	0.24
others	70	0.01	3.05

Table 6.3: Number of events by channel after BDT selection

6.1.3 Selection based on time of flight

Because the BDT classification is not sufficient to separate the signal from background, a time of flight analysis is performed. For this, both tracks are required to be associated

to a cluster (track-to-cluster association). For each track associated with a cluster, the variable time-of-flight difference, DTOF, is computed as

$$DTOF_i = T_{cl_i} - L_i/c\beta_i$$

where $i = 1, 2$ are the two selected tracks, T_{cl_i} is the time of the cluster associated to track i , L_i is the length of the track, and $\beta_i = p_i/\sqrt{p_i^2 + m_i^2}$ with m_i being the mass hypothesis for track i . As shown in Sec. 3.2.2, T_{cl_i} depends on the T_0 calculated for each event and thus the same T_0 value is assigned to both tracks. To reduce the uncertainty from the determination of T_0 , the quantity $\delta DTOF$ is used:

$$\delta DTOF = DTOF_1 - DTOF_2$$

The DTOF value, and consequently $\delta DTOF$, depends on the mass hypothesis, that can be different for the two tracks. Since the main goal is to discriminate $K_S \rightarrow \pi e \nu$ against $K_S \rightarrow \pi^+ \pi^-$, $\delta DTOF$ is first computed in the $\pi\pi$ hypothesis:

$$\delta DTOF(\pi\pi) = DTOF_1(\pi) - DTOF_2(\pi)$$

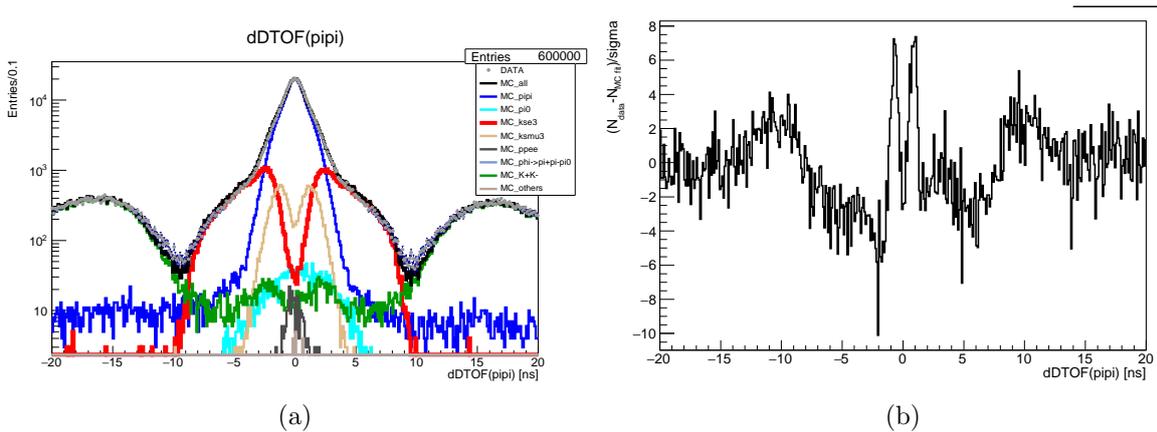


Figure 6.9: (a) Distribution of $\delta DTOF(\pi\pi)$ for the data and MC signal events, background events and their sum; (b) significance defined as the difference between data and MC divided by the statistical error.

The distribution of the $\delta DTOF(\pi\pi)$ variable is shown in Figure 6.9. The central part of the distribution, for $|\delta DTOF(\pi\pi)| < 5$ ns, is dominated by the $K_S \rightarrow \pi^+ \pi^-$ background while the wings, for $|\delta DTOF(\pi\pi)| > 10$ ns, by $K^+ K^-$ events. The signal is concentrated in the region $|\delta DTOF(\pi\pi)| < 10$ ns. To select the signal a cut is chosen:

$$2.5 \text{ ns} < |\delta DTOF(\pi\pi)| < 10 \text{ ns}$$

Figure 6.9 shows a non satisfactory agreement between data and MC, thus a control sample is needed to determine the efficiency of the $\delta DTOF$ selection.

After the applied selection, both the πe and $e\pi$ hypotheses are tested:

$$\delta DTOF(\pi e) = DTOF_1(\pi) - DTOF_2(e)$$

$$\delta DTOF(e\pi) = DTOF_1(e) - DTOF_2(\pi)$$

where the assignment as track-1 and track-2 is chosen at random. The two-dimensional distribution of $\delta DTOF(\pi e) \times \delta DTOF(e\pi)$ is shown in Figure 6.10, where the signal populates either band around $\delta DTOF \sim 0$ ns. A combined cut is chosen as

$$|\delta DTOF(\pi e)| < 1 \text{ ns} \quad |\delta DTOF(e\pi)| < 1 \text{ ns}$$

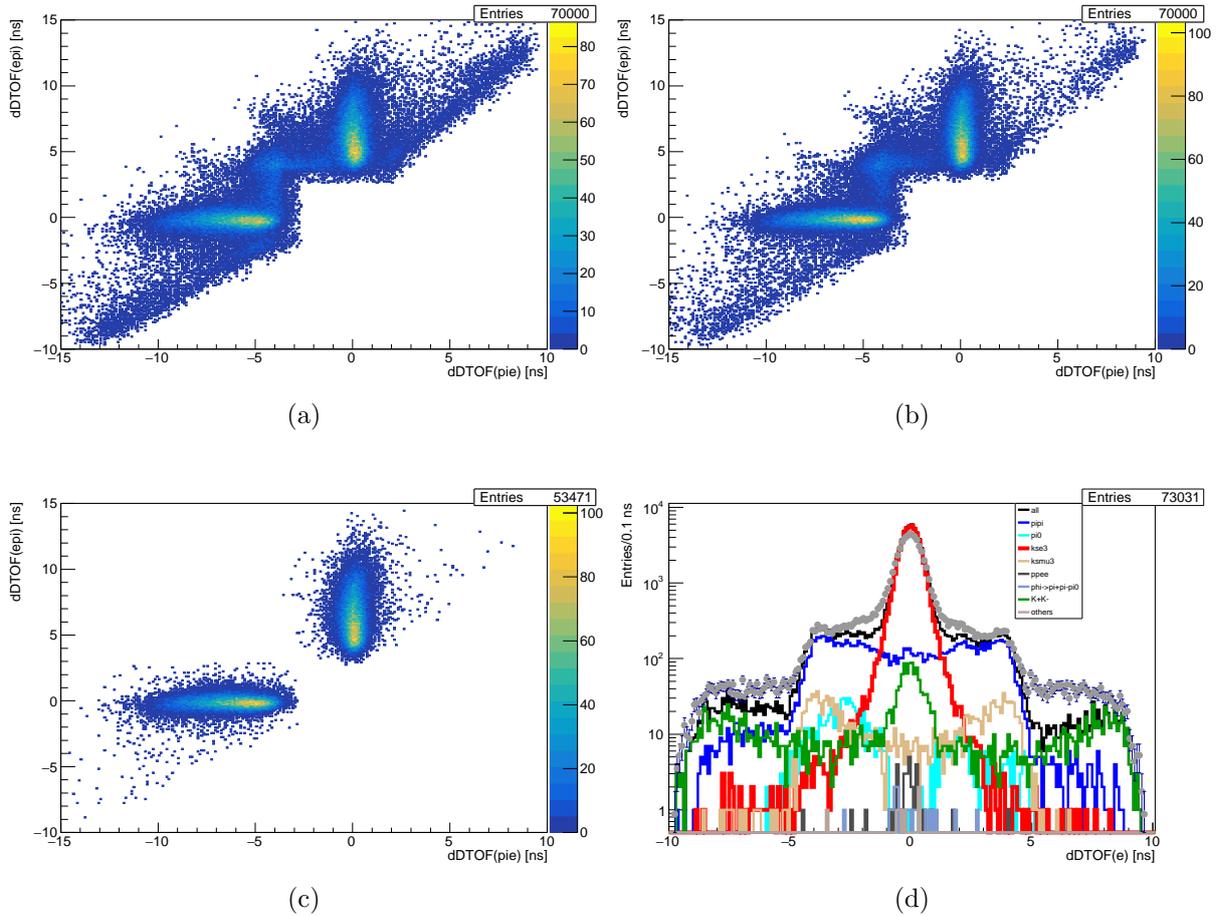


Figure 6.10: Two-dimensional distribution $\delta DTOF(\pi e) \times \delta DTOF(e\pi)$ for data (a), MC signal and background (b), and for signal MC only (c); $\text{Min}|\delta DTOF(\pi e), \delta DTOF(e\pi)|$ (d), which is the correct mass hypothesis.

The two-dimensional distribution allows for $e\pi$ separation: the lower of the two $\delta DTOF$ values corresponds to the correct mass hypothesis. The particle identification is made in this way: if $|\delta DTOF(\pi e)| < |\delta DTOF(e\pi)|$, the first particle is the pion and the second is the electron, otherwise if $|\delta DTOF(e\pi)| < |\delta DTOF(\pi e)|$ the opposite mass assignment is assumed.

6.1.4 Extraction of the signal

After this selection, the number of events in the data sample is 57,577 and its composition as predicted by simulation is listed in Table 6.4

	MC events	fraction [%]	relative efficiency [%]	tot relative efficiency [%]
DATA	57 577		7.30	0.02
MC	56 843		8.81	0.02
$K_S \rightarrow \pi e \nu$	53 559	94.22	67.21	20.66
$K_S \rightarrow \pi^+ \pi^-$	2 157	3.83	0.47	< 0.001
$\phi \rightarrow K^+ K^-$	903	1.58	1.22	< 0.01
$K_S \rightarrow \pi^0 \pi^0$	23	0.04	0.97	0.08
$K_S \rightarrow \pi \mu \nu$	136	0.24	0.62	0.10
$K_S \rightarrow \pi^+ \pi^- e^+ e^-$	28	0.05	9.89	0.15
$\phi \rightarrow \pi^+ \pi^- \pi^0$	5	0.02	32.56	0.06
others	5	< 0.01	7.14	0.10

Table 6.4: Number of events by channel after $\delta DTOF$ selection

After particle identification (Sec. 6.1.3) the invariant mass of the particle identified as the electron is calculated as:

$$m_e^2 = (E_{K_{S\text{tag}}} - E_\pi - p_{\text{mis}})^2 - p_e^2 \quad (6.2)$$

where $p_{\text{mis}}^2 = (\vec{p}_{K_{S\text{tag}}} - \vec{p}_e - \vec{p}_\pi)^2$, $p_{K_{S\text{tag}}}$ and $E_{K_{S\text{tag}}}$ are the K_S momentum and energy estimated from the tag, and p_e and p_π are the momenta of the electron and the pion track. Figure 6.11 shows the distribution of m_e^2 for data and MC. The distribution identifies a large electron peak at $m_e^2 \sim 0$ and two small peaks at the m^2 value of the muon and the charged pion.

Since there are small differences between the momentum distributions of data and MC, a smearing procedure is applied on the MC momentum [21] as:

$$p_i(j)' = p_i(j) \times (1 + p_{\text{shift}}) \times (1 + \text{ranG}(0, 0.004)) \quad i = e, \pi; \quad j = x, y, z \quad (6.3)$$

where $p_i(j)$ and $p_i(j)'$ are the components of the momentum of the selected tracks before and after the smearing procedure, $p_{\text{shift}} = 0.000137$ and $\text{ranG}(\mu, \sigma)$ is a Gaussian distributed number. Figure 6.11 shows the distribution of m_e^2 after the smearing procedure.

After this, to extract the number of the signal events, a fit is performed to the m_e^2 data distribution using the MC shapes of three components: $\pi e \nu$, $\pi^+ \pi^-$ and all others channels. The `TFractionFitter` class of Root [22] is used to perform the fit. It minimises a binned likelihood function and gives in output the χ^2 of the fit and, for each of the components, the fraction of the total number of events and the error of the fraction. The fit is performed in the interval $[-30000, +30000]$ MeV², the distributions are divided into 100 bins, the number of degrees of freedom is 97. Table 6.5 lists the result of the fit, and Figure 6.12

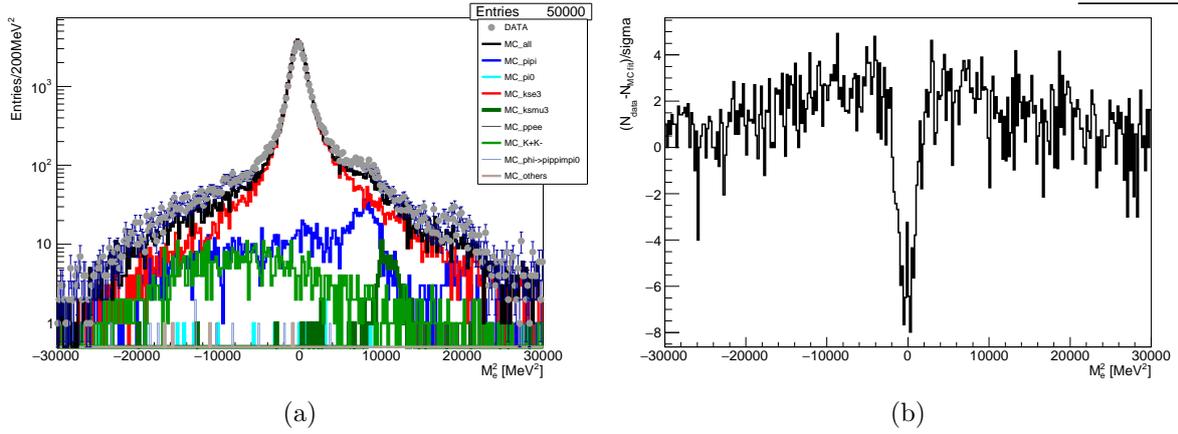


Figure 6.11: (a) Distribution of m_e^2 after applying the correction of Eq. (6.3) for the data and MC signal events, background events and their sum; (b) significance defined as the difference between data and MC divided by the statistical error.

shows the fitted m_e^2 distribution and the distribution of the residuals. The number of signal events resulting from the fit is:

$$N_{\pi e \nu} = 49647 \pm 316 \text{ events} \quad (6.4)$$

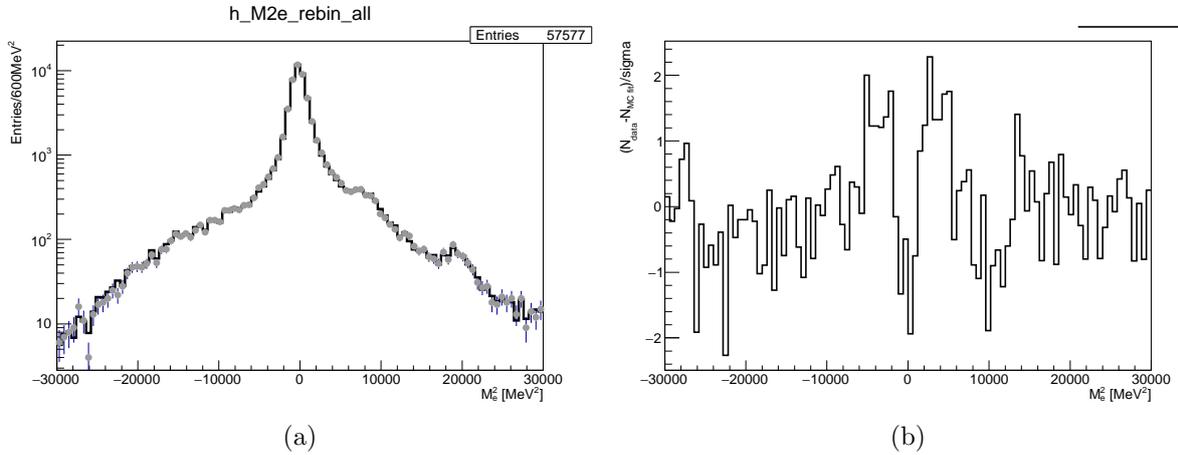


Figure 6.12: (a) Distribution of m_e^2 with the fit superimposed, (b) distribution of the difference between data and fit divided by the statistical errors.

Fit error correction

As explained in Ref. [23], the error calculation in the software package `TFractionFitter` could be affected by a bug. This is related to the fact that in this software the identity

	fraction	events	relative error [%]
$\pi e \nu$	0.8652	49 652 \pm 351	0.71
$\pi^+ \pi^-$	0.0758	4 350 \pm 392	9.00
all others	0.0590	3 388 \pm 384	11.33
Total		57 389	
χ^2/ndf	1.96		

Table 6.5: Fit output

$f_i = f_i^*$ (where f_i^* are the fractions calculated after the minimisation procedure and f_i are the true fractions) it is supposed to hold everywhere. In reality, that identity is true only at the likelihood minimum, thus only after the minimisation procedure, while it is only an approximation elsewhere. Because the errors of the fitted fractions are obtained moving around the minimum in a small range, using that identity could enlarge the errors of the fractions obtained as output of the fit. The correct errors can be computed through the error propagation starting from:

$$f_i = \frac{f_i^*}{\sum_i f_i^*} \quad (6.5)$$

using the output covariance matrix of the fit.

Despite of possible improvements of the software package, this seems to be the correct method to compute the error of the calculated fractions, then it is considered valid here. Applying this method, the error of the fit for the signal fraction is:

$$fit\ error = 0.384\% \quad (6.6)$$

6.2 Determination of the efficiencies

The agreement between data and MC is not good enough to use the MC for evaluating the efficiencies for most of the selections. Then only few efficiencies are derived from simulation. Most of the efficiencies are obtained from different control samples (CS). In each case, the efficiency used for the analysis is calculated as:

$$\epsilon_{AS}^{DT} = \epsilon_{CS}^{DT} \times \frac{\epsilon_{AS}^{MC}}{\epsilon_{CS}^{MC}} \times \frac{p'}{p} \quad (6.7)$$

where ϵ_{AS}^{DT} is the efficiency used in the analysis sample (AS) for the specific channel, ϵ_{CS}^{DT} is the efficiency of the data control sample (CS), ϵ_{AS}^{MC} is the MC efficiency of the analysis sample, ϵ_{CS}^{MC} is the efficiency computed from the MC for the control sample, p and p' are the MC purities before and after a given selection.

6.2.1 $K_L \rightarrow \pi e \nu$ control sample

Most of the efficiencies of the signal selection are computed from the abundant $K_L \rightarrow \pi e \nu$ decay. This decay is kinematically identical to the $K_S \rightarrow \pi e \nu$ decay, the only difference being the much larger lifetime. In this case the tag is done with $K_S \rightarrow \pi^+ \pi^-$ decay, selected in the same way as for the analysis sample (Section 4.2.2) and adding a cut on the δm to have high purity:

$$|\delta m| < 15 \text{ MeV}.$$

To minimise the misidentification of the K_L with the K_S vertex, the radial distance of the K_L vertex is required to be greater than 1 cm, and smaller than 5 cm to best match the K_S selection

$$1 \text{ cm} < \rho_{\text{VTX}}^{K_L} < 5 \text{ cm}$$

At this level the sample is composed mainly of $K_L \rightarrow \pi e \nu$ (BR = 0.405), $K_L \rightarrow \pi \mu \nu$ (BR = 0.270) and $K_L \rightarrow \pi^+ \pi^- \pi^0$ (BR = 0.125), while most of the $K_L \rightarrow \pi^0 \pi^0 \pi^0$ events are rejected requiring the vertex.

The goal is to separate $K_L \rightarrow \pi e \nu$ from the other decays, minimising the bias of the selection. For this reason two different sets of selection are used to evaluate the two efficiencies. The analysis selection is divided in two steps: selection based on tracking variables (preBDT and BDT), and selection based on time of flight (TOF). The variables used in the two steps are quite independent of each other, as shown in Fig. 6.13

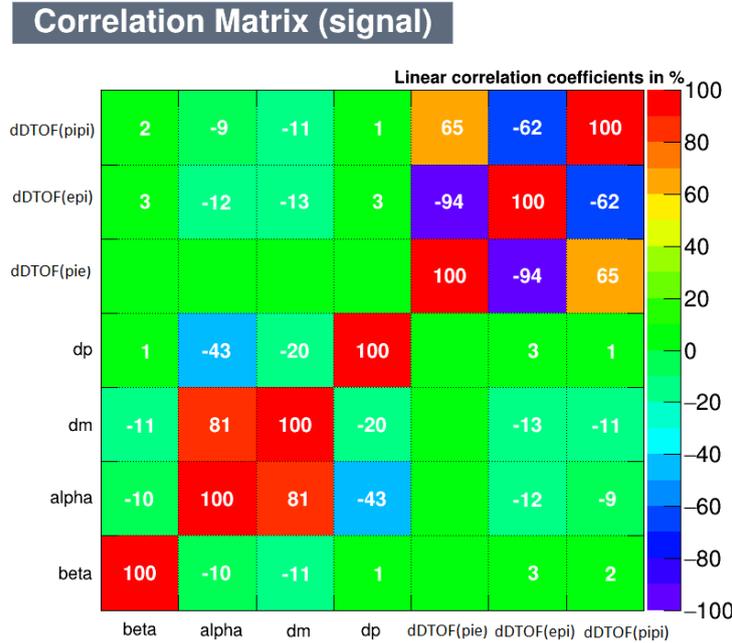


Figure 6.13: Correlation between the variables used in the BDT and TOF selections.

Then the selection of the control sample follows a similar scheme:

- a common cut on the missing mass distribution, shown in Figure 6.14, $m_{\text{mis}}^2 < 15,000 \text{ MeV}^2$ to reject $K_L \rightarrow \pi^+ \pi^- \pi^0$ events;
- a cut on the TOF variables is applied to evaluate the efficiency of the BDT pre-selection and BDT selection;
- a cut on the $\delta m \times m_{\text{mis}}^2$ distribution is applied to evaluate the efficiency of the TCA and TOF selections.

Here the missing mass is defined as $m_{\text{mis}}^2 = E_{\text{mis}}^2 - p_{\text{mis}}^2$, where $p_{\text{mis}}^2 = (\vec{p}_{K_L \text{tag}} - \vec{p}_1 - \vec{p}_2)^2$; $p_{K_L \text{tag}}$ is the K_L momentum estimated from the tag, p_1 and p_2 are the momenta of the tracks forming the K_L vertex, and $E_{\text{mis}} = m_\phi/2 - E_1 - E_2$ computed with the charged pion mass hypothesis.

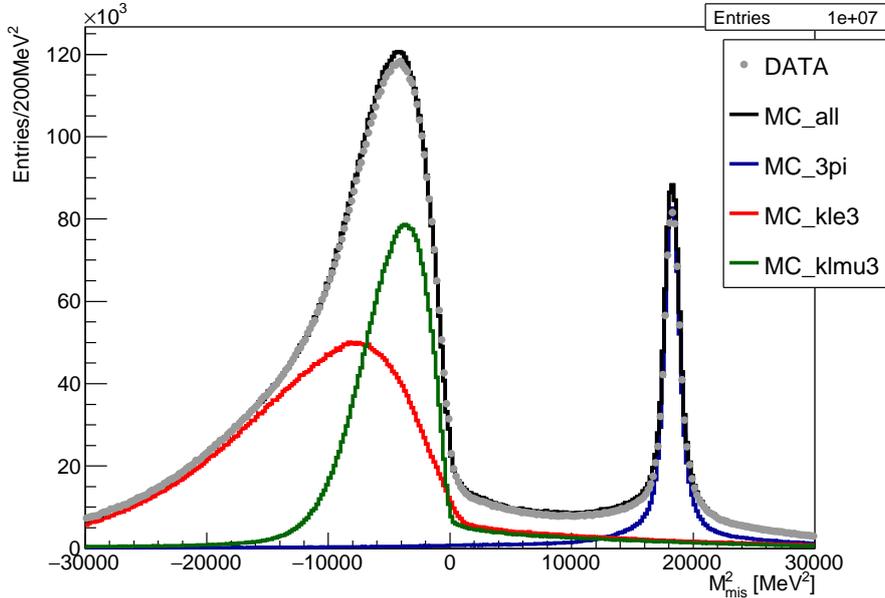


Figure 6.14: Distribution of the missing mass squared for two-track events of the $K_L \rightarrow \pi e \nu$ selection for data and MC. A clear peak at the π^0 mass is visible.

Control sample tag

Two different high statistics tags are possible for the $K_L \rightarrow \pi e \nu$ sample: $K_S \rightarrow \pi^+ \pi^-$ and $K_S \rightarrow \pi^0 \pi^0$. In principle, the neutral decay better emulates the neutral tag (K_L -crash) of the analysis sample. In reality, there are two reasons to prefer $K_S \rightarrow \pi^+ \pi^-$. First, the resolution of the direction of the tag should be similar for the control and analysis samples. This resolution is important for some variables used as input to the BDT, in particular those determining the Lorentz transformation to the K_S centre of mass. The resolution of the neutral vertex of the $K_S \rightarrow \pi^0 \pi^0$ is much worse than for $K_S \rightarrow \pi^+ \pi^-$. Second, to compute correctly the TOF efficiencies, the cluster giving the T_0 should not be associated

with the tagging kaon, as for the K_L -crash in the analysis sample. In case of $K_S \rightarrow \pi^0 \pi^0$, the first cluster arriving at the calorimeter is often a photon from π^0 decays, such that it is almost impossible to fulfil this requirement, while for $K_S \rightarrow \pi^+ \pi^-$, because of the small ρ_{VTX} for K_L , only about half of the time is a pion from the tagging kaon determining the event T_0 .

PreMVA and MVA control sample and efficiency

The control sample for evaluating the efficiency of the preselection and BDT cut is selected with a cut on the δDTOF variable. As for the signal, TCA is required for both tracks. The variables used are:

$$\delta\text{DTOF}(\pi e) = \text{DTOF}_1(\pi) - \text{DTOF}_2(e) \quad ; \quad \delta\text{DTOF}(e\pi) = \text{DTOF}_1(e) - \text{DTOF}_2(\pi)$$

$$\delta\text{DTOF}(\pi\mu) = \text{DTOF}_1(\pi) - \text{DTOF}_2(\mu) \quad ; \quad \delta\text{DTOF}(\mu\pi) = \text{DTOF}_1(\mu) - \text{DTOF}_2(\pi)$$

where 1 and 2 are the two tracks from the reconstructed K_L vertex. The distributions of these variable are shown, separately for data, total MC simulation and for the MC signal only, in Figure 6.15 and Figure 7.6.

The following selection, in OR of each other, is applied, to select $K_L \rightarrow \pi e \nu$:

- $\delta\text{DTOF}(\pi e) > -0.4 \otimes \delta\text{DTOF}(\pi e) < 0.7 \otimes \delta\text{DTOF}(e\pi) > 2 \otimes \delta\text{DTOF}(e\pi) < 11$ ns
- $\delta\text{DTOF}(e\pi) > -0.7 \otimes \delta\text{DTOF}(e\pi) < 0.2 \otimes \delta\text{DTOF}(\pi e) > -11 \otimes \delta\text{DTOF}(\pi e) < -2$ ns
- $\delta\text{DTOF}(\pi\mu) < -2 \otimes \delta\text{DTOF}(\mu\pi) < 1$ ns
- $\delta\text{DTOF}(\pi\mu) > 1 \otimes \delta\text{DTOF}(\mu\pi) < 0.8$ ns

where all values are in ns.

After this selection, the purity of the sample determined from MC is about 97%. The distributions show good agreement, except for low values of δp . The difference in the δp distribution of the K_L control sample is probably due to a mis-assignment of the tracks to the $K_S \rightarrow \pi\pi$ and K_L vertices. To avoid consequences of this disagreement, a cut $\delta p > -95$ MeV is applied both to the analysis sample (Sec. 6.1.1) and the control sample. The comparison between the MC BDT output distributions of the analysis and control samples is shown in Figure 6.17.

After this last selection, the efficiencies of the preselection and of the BDT cut are evaluated with the same method described in Eq. (6.7). The results are shown in the second and fourth row of Table 6.7.

TCA and DTOF control sample and efficiency

The control sample for evaluating the TCA and TOF efficiencies is selected using only tracking variables. To increase the purity without biasing the sample, a triangular cut in the $\delta m \times m_{\text{mis}}^2$ variables is applied, as shown in Figure 6.18:

- $\delta m > 0.015 \times m_{\text{mis}}^2 + 15.5$ MeV
- $\delta m < 0.00555 \times m_{\text{mis}}^2 + 32.4$ MeV
- $\delta m > -170$ MeV

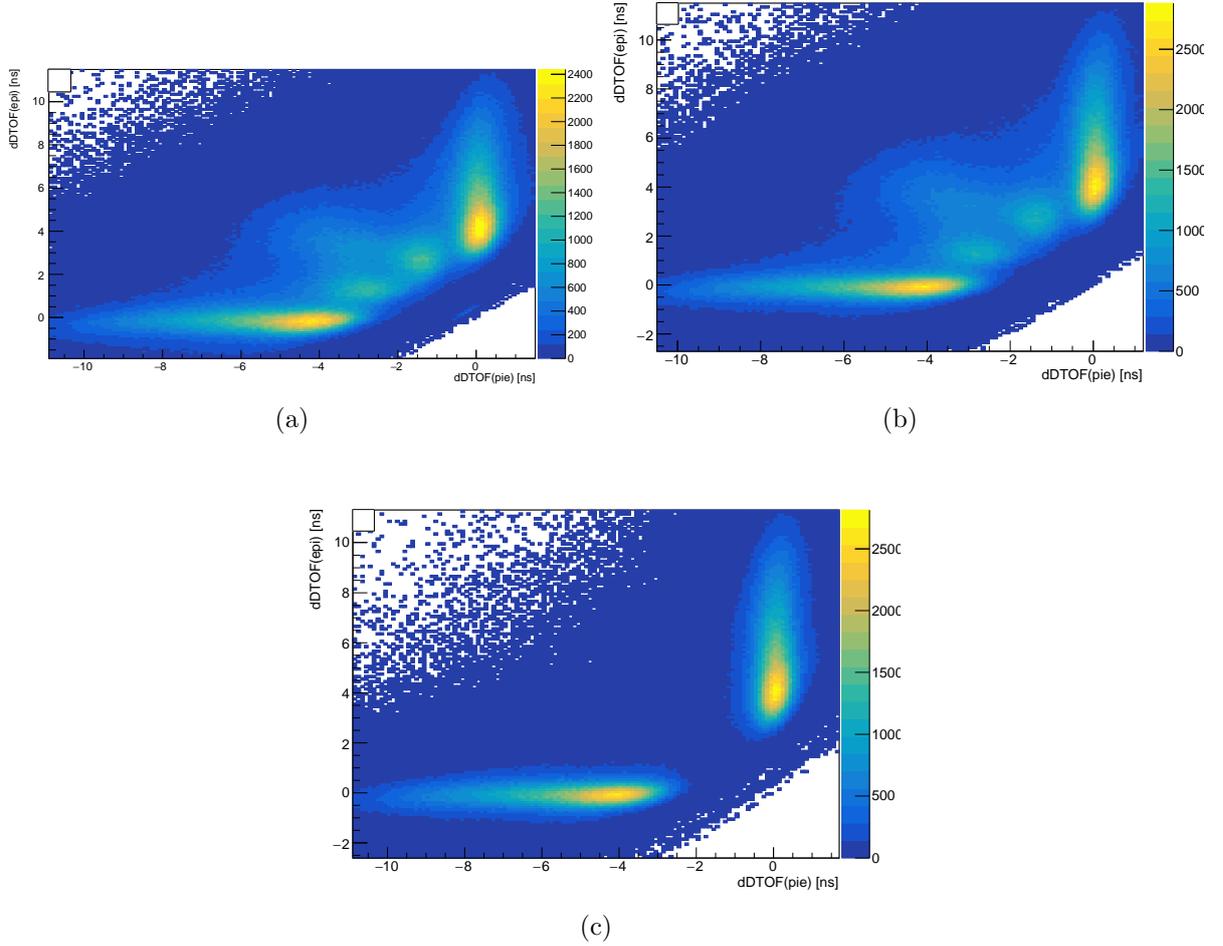


Figure 6.15: Two-dimensional distribution $\delta DTOF(\pi e) \times \delta DTOF(e\pi)$ for data (a), MC signal and background (b), and for signal MC only (c).

To reduce the dependence on the number of clusters in the event, different in the analysis sample (at least three clusters, one from the K_L -crash and two from the K_S) respect to the control sample (at least two from K_S and two from K_L), the track-to-cluster association is required for both tracks associated to the K_S decay. Then the efficiency for the TCA selection is computed in the usual way from Eq. 6.7.

The result is $\epsilon_{TCA} = 0.46391 \pm 0.00197$ and is reported Table 6.7.

To check that the higher number of clusters in the control sample is not affecting the calculation, the efficiency is also calculated after weighting every event of the control sample to match the number of clusters of the analysis sample.

The new value is $\epsilon_{TCA} = 0.46387 \pm 0.00250$, equal to the previous result in the errors. The first method is chosen to avoid possible biases from the weighting procedure. The difference between the two values is taken as systematic error of the TCA efficiency.

After evaluating the TCA efficiency, the control sample is selected in the same way as

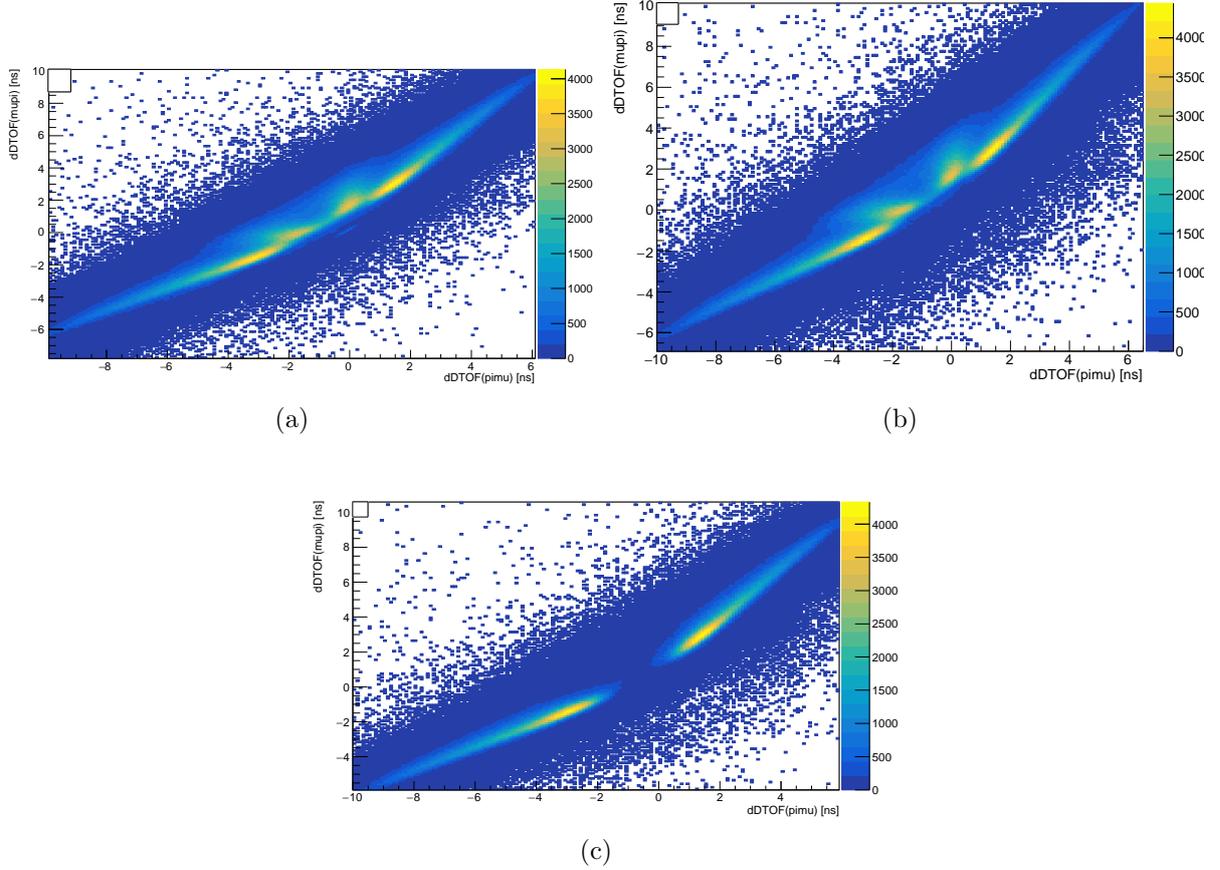


Figure 6.16: Two-dimensional distribution $\delta DTOF(\pi\mu) \times \delta DTOF(\mu\pi)$ for data (a), MC signal and background (b), and for signal MC only (c).

the analysis sample before the TOF selections (Section 6.1.3):

- preselection, as described in Section 6.1.1;
- same BDT cut in the BDT classifier output, $BDT > 0.15$.

There is still a difference between the analysis and the control samples, because in the first case the T_0 of the event is always determined by a K_S daughter particle, while for the control sample the first cluster in time originates from the $K_S \rightarrow \pi\pi$ tag in about half of the events. The T_0 is not taking part in the TOF variables, but the event selection could bias the result. To better match the two samples, it is required that the first cluster in time arriving at the calorimeter has to be associated to the K_L decay. The comparison between the MC $\delta DTOF(\pi\pi)$, $\delta DTOF(\pi e)$ and $\delta DTOF(e\pi)$ distributions of the analysis and control samples is shown in Figure 6.19.

The results of the TOF efficiencies for the $\delta DTOF(\pi\pi)$ and the $\delta DTOF(\pi e) \times \delta DTOF(e\pi)$ selections are shown in Table 6.7.

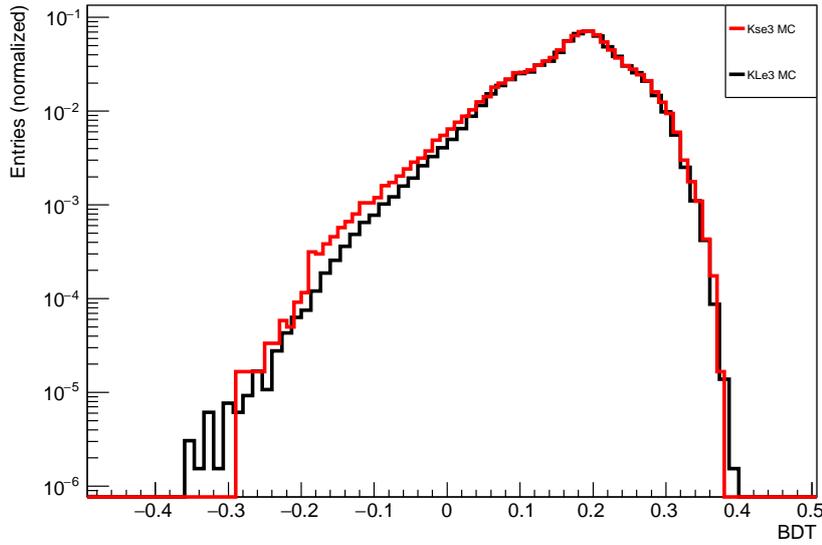


Figure 6.17: Distribution of the BDT classifier output for the $K_S \rightarrow \pi e \nu$ and the control sample of $K_L \rightarrow \pi e \nu$ events.

6.2.2 MC efficiency

Only a couple of efficiencies are calculated directly from MC, these efficiencies are high and are related to:

- the selection on the angle of the crash cluster, $\theta_{\text{crash}} > 15^\circ$ (see Sec.4.2.1);
- the cut $\delta p < -95$ introduced in Section 6.1.1 to equalize control the δp distributions of the data and control sample.

For both selections, the K_L control sample is not useful, because in the first case there is no K_L -crash, and in the second the cut δp is used both for the signal and the control samples. The combination of these two efficiency is reported in the third row of Table 6.7.

6.2.3 Determination of the ratio R_ϵ

For all common selections (Sec. 4) there is no need to compute separately the efficiencies for the control and the normalisation sample. In fact in Eq. 4.1 only the ratio of the two is needed. The ratio of efficiencies R_ϵ , as explained in Sec. 4, is composed of several factors: Trigger, FILFO, Event Classification, K_L -crash and K_S identification. The single efficiency ratio of these selections, and the total R_ϵ ratio, are reported in Table 6.6. All these efficiencies are taken from MC simulation.

Trigger

To check the agreement between data and MC on the trigger efficiency, the two-trigger method is used [16]. This method is only usable to check the agreement between the

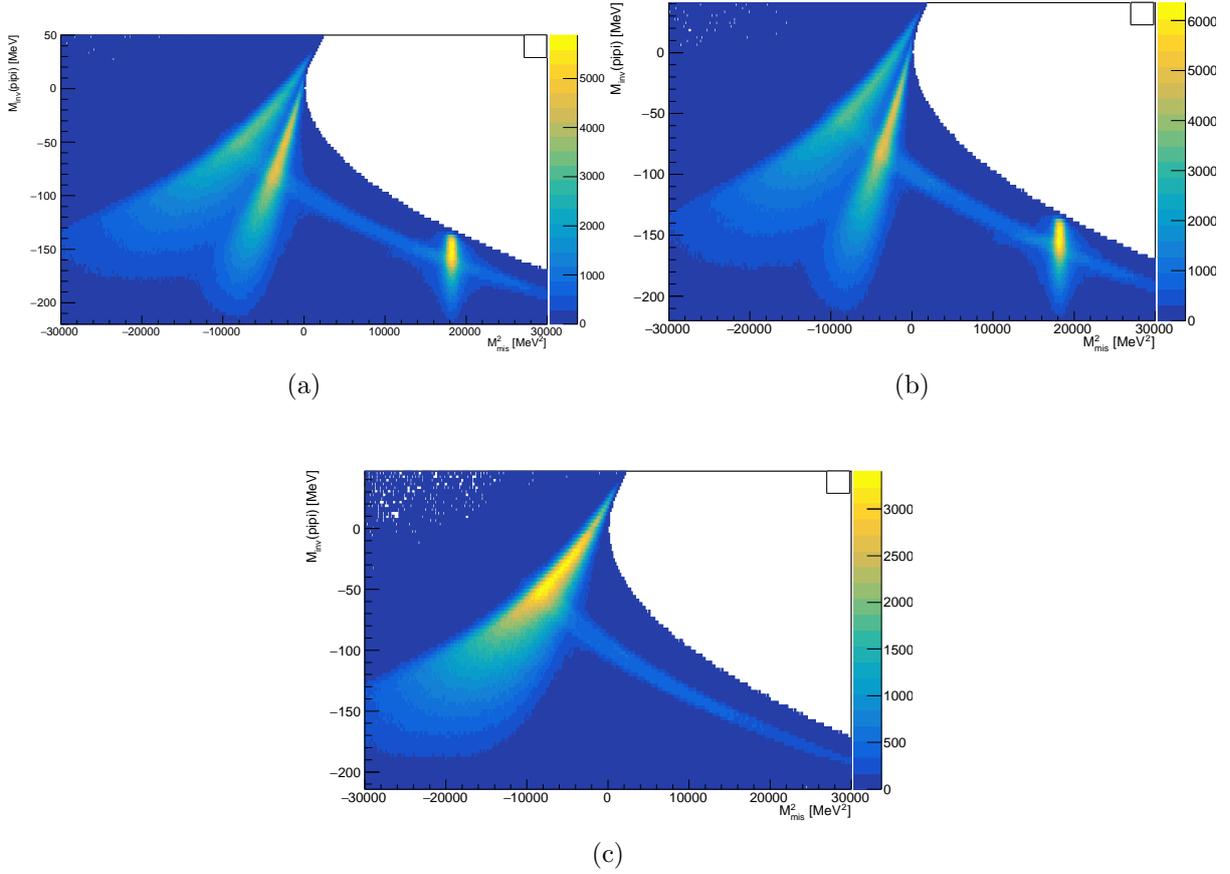


Figure 6.18: Two-dimensional distribution $\delta m \times m_{\text{mis}}^2$ for data (a), MC signal and background (b), and for signal MC only (c).

two, but cannot be used to compute trigger efficiency, because it is not taking in account events not triggered with any trigger, due for example to geometrical acceptance.

As explained in Section 3.2.3, two separate triggers are used in the experiment: the DC trigger and the EMC trigger. This gives the possibilities to compute relative trigger efficiencies:

$$\epsilon_{EMC} = \frac{N_{BOTH}}{C_T * N_{DC}} \quad \epsilon_{DC} = \frac{N_{BOTH}}{C_T * N_{EMC}} \quad (6.8)$$

where N_{EMC} and N_{DC} and N_{BOTH} are the observed number of event triggered by EMC, DC and both, respectively; $C_T = 1.01$ is the correlation factor between the two triggers, determined from MC simulation [16]. Then the total trigger efficiency can be computed as:

$$\epsilon_{TRG} = 1 - (1 - \epsilon_{EMC})(1 - \epsilon_{DC}) \quad (6.9)$$

Using MC information, the results for signal and normalisation channels are:

$$\epsilon_{TRG}(\pi^+ \pi^-) = 99.893 \pm 0.002\% \quad \epsilon_{TRG}(\pi e \nu) = 99.589 \pm 0.008\% \quad (6.10)$$

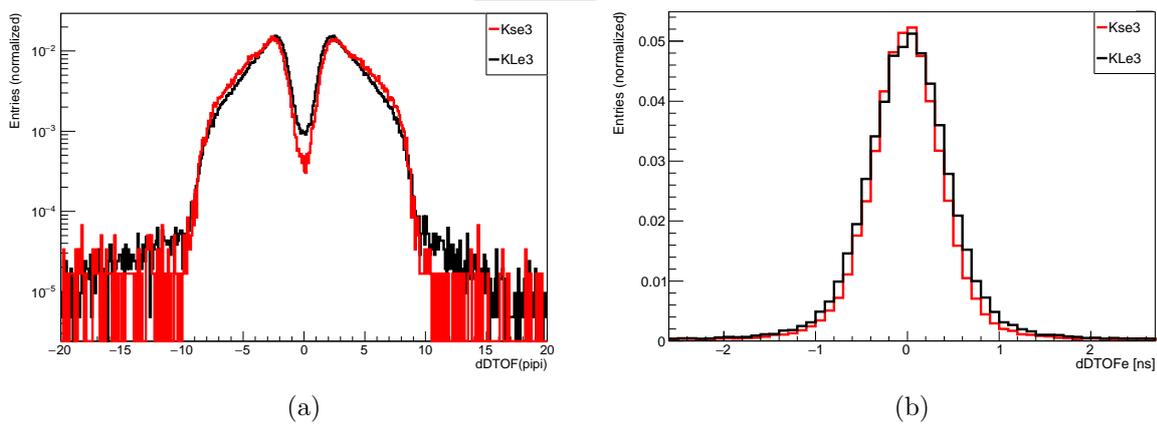


Figure 6.19: MC distributions of $\delta DTOF(\pi\pi)$, $\delta DTOF(\pi e)$ and $\delta DTOF(e\pi)$ for the analysis ($K_S \rightarrow \pi e \nu$) and control ($K_L \rightarrow \pi e \nu$) sample.

Selection	Ratio of efficiency
Trigger	1.0297 ± 0.0003
FILFO	1.0054 ± 0.0001
Event classification	1.0635 ± 0.0004
T0 time fix	1.0063 ± 0.0001
K_L -crash	1.0295 ± 0.0010
K_S ID	1.0418 ± 0.0009
R_ϵ	1.1882 ± 0.0012

Table 6.6: R_ϵ ; single errors are strongly correlated, so the error on R_ϵ is smaller than the quadratic sum of the errors of the single efficiencies.

Using the data sample, that is mainly composed of $\pi^+ \pi^-$ with small contamination of other channels ($\sim 0.6\%$), the value $\epsilon_{TRG}(\pi^+ \pi^-) = 99.819 \pm 0.003\%$ is obtained. The agreement between the two estimates is better than 0.1%, the difference is considered negligible for the final result. So the efficiency of the MC is considered trustable.

6.2.4 Summary of efficiencies

Selection	Efficiency
TCA	0.4639 ± 0.0009
CS preselection	0.9720 ± 0.0007
MC preselection	0.9961 ± 0.0002
BDT selection	0.6534 ± 0.0013
TOF selection	0.6605 ± 0.0012
FIT interval	0.9985 ± 0.0001
Total	0.1935 ± 0.0004

Table 6.7: Efficiencies of the analysis cuts; single efficiency errors are strongly correlated, so the error of the total efficiency is smaller than the quadratic sum of the errors of the single efficiencies. For ϵ_{DTOF} a correction will be applied (See Section 6.3.2 for details).

6.3 Systematic uncertainties

The main systematic errors are related to:

- the statistical uncertainties in evaluating the efficiencies, both from MC and data control samples;
- the cuts applied in the analysis, related to the resolution of the variables to which a cut is applied. The relative systematic error is evaluated by varying the selection cut by several standard deviation, recomputing the BR for the different cuts, and taking the standard deviation of the different results as systematic error;
- the difference of the result when an efficiency is calculated in different ways. In this case, if the two results are in agreement, the simplest procedure is chosen to minimise a possible bias, and the difference between the two results is taken as systematic error.

The systematic error of the efficiency is evaluated in part with the control sample, in part with the MC simulation sample. In both cases, the statistical error due to the number of events, if not negligible, is taken into account in the result. Three main samples are used, beside the signal sample, in the analysis (Sec. 6.2):

- two different data control samples selected from the same K_L sample;
- five subsamples of the data used as control samples for evaluating the selection efficiencies of the normalisation sample;
- the MC sample, used for few efficiencies and for the determination of R_c .

Several efficiencies are computed for each control sample, and the related statistical error is evaluated. The largest statistical error of each sample is taken as the systematic error. The results are shown in Table 6.8.

Selection	Relative systematic error [%]
TOF	0.672
BDT	0.276
K_L CS statistics	0.108
$\pi^+\pi^-$ CS statistics	0.002
MC sample statistics	0.143
TCA efficiency	0.009
$\pi^+\pi^-$ efficiency	0.092
Total	0.754

Table 6.8: Systematic errors

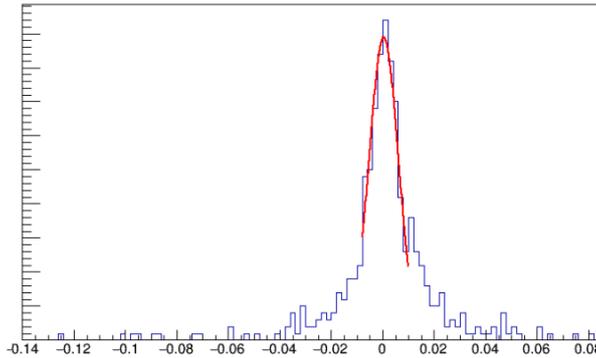
6.3.1 BDT scan and systematics

BDT standard deviation

The BDT output value is determined by a combination of several input variables, that have different resolutions, and it is not an analytic function of the five inputs, such that the method of error propagation is not applicable. Beside the measurement of the K_L -crash position determined by the calorimeter, all variables used as BDT input are derived by tracking measurements and the simulated distribution of each variable fairly well reproduces that of the data, at least for the bulk of the distributions. To evaluate the resolution of the BDT classifier, the MC sample is used by computing event-by-event the distribution of the difference between the BDT output before (at truth level) and after the event reconstruction for those MC events that have a BDT output value in the interval 0.130–0.180. Fig. 6.20 shows the distribution of the difference, the r.m.s. width of the distribution is determined by a fit with a Gaussian function around the peak:

$$\sigma_{BDT} = 0.005$$

This value is used as the step for the BDT scan to evaluate the systematic error related to the choice of the BDT cut.

Figure 6.20: Value of $BR(K_S \rightarrow \pi e \nu)$ as function of the cut on the BDT classifier output.

BDT scan

To select the best value for the selection based on the BDT classifier output, and to check whether there is a dependence of the result, a scan of the BDT cut is performed. The step of the BDT cut scan is taken equal to one standard deviation, and for each step the BDT and DTOF efficiencies are recomputed using the control sample. The interval considered for the scan, $0.135 < \text{BDT cut} < 0.170$, corresponds to about 7 standard deviations around the chosen value of the cut ($\text{BDT} > 0.150$).

The results are presented in Figure 6.21, that shows good stability by varying the BDT cut. An error band is defined as the difference between the maximum and the minimum value in Figure 6.21, and half of the band width is taken as systematic error due to BDT cut choice.

$$\text{BDT syst} = 0.276\% \tag{6.11}$$

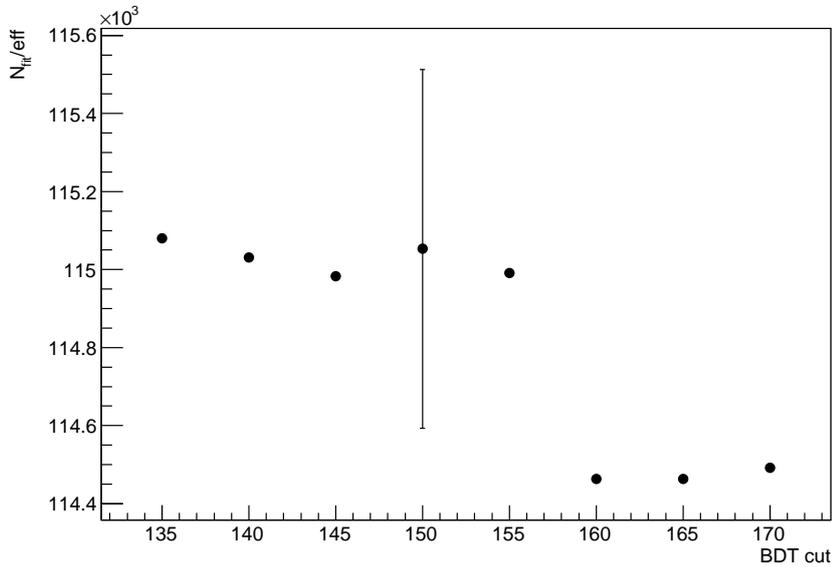


Figure 6.21: Number of fit counts divided by total efficiency as function of the cut on the BDT classifier output.

6.3.2 Scan and systematic on δDTOF

Resolution on δDTOF

The variables used in the TOF selection are differences of time of flight measurements, defined as

$$\text{DTOF}_i = T_{\text{cli}} - L_i/c\beta_i$$

The resolution is the combination of the time resolution of the calorimeter and the tracking resolution of the drift chamber. The impact of the resolution in measuring the track length has a much smaller effect than the impact of the calorimeter time resolution; thus

the resolution in measuring the TOF variables is dominated by the time resolution.

The resolution for time difference [24, 25] can be parameterised as:

$$\sigma_{\Delta T} = \frac{a}{\sqrt{E^*}} \oplus \sqrt{2}b$$

where $E^* = \frac{E_1 E_2}{E_1 + E_2}$ is in GeV, a is found to be 57.1 ps in the barrel and 62.5 ps in the endcap calorimeter, and b is 105 ps in the barrel and 120 ps in the endcap.

The average $\delta DTOF$ resolution is estimated to be:

$$\sigma_{\delta DTOF} = 270 \text{ ps} \quad (6.12)$$

Scan on $\delta DTOF$

Looking at the $\delta DTOF(\pi\pi)$ distribution of (Figure 6.9, it is clear that choice of the lower cut is delicate in a region where the signal and background distributions are steep and have opposite slopes. To choose the cut on $\delta DTOF(\pi\pi)$ that ensures good stability of the result, a scan of the lower cut of $\delta DTOF(\pi\pi)$ is done. The $\delta DTOF(\pi\pi)$ cut is varied steps is 0.25 ns, approximately equal to one standard deviation, and the efficiency is recomputed at each step. The result is shown in Figure 6.22.

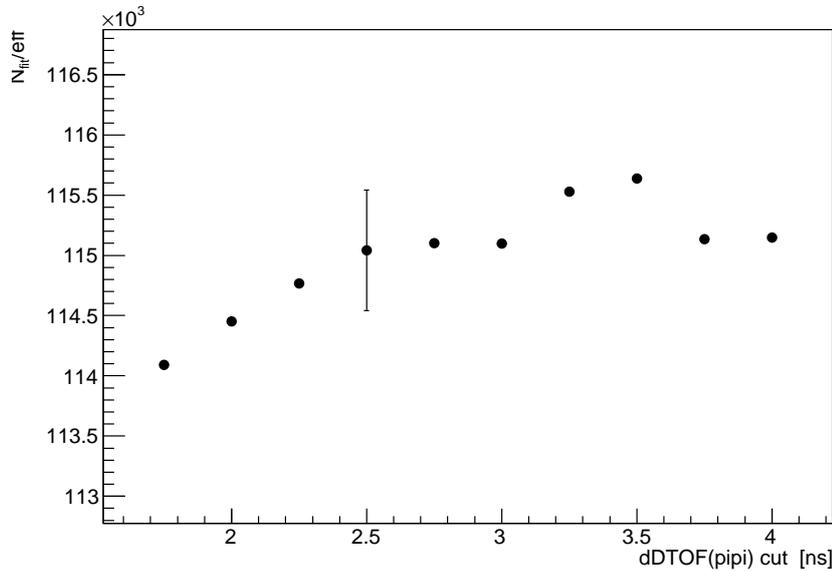


Figure 6.22: Value of $BR(K_S \rightarrow \pi e \nu)$ as function of the $\delta DTOF$ cut.

A dependence of the result from the $\delta DTOF(\pi\pi)$ cut is observed for low values of the cut, while increasing the cut the result seems to reach a plateau. An error band is defined as the difference between the maximum and the minimum value in Figure 6.22; the central value is taken as final number of counts, while half of the band width is taken

as systematic error due to $\delta DTOF(\pi\pi)$ cut choice.

The result is:

$$\delta DTOF(\pi\pi)_{syst} = 0.672\%$$

In comparison, varying by one standard deviation the higher cut of $\delta DTOF(\pi\pi)$, and the $\delta DTOF(\pi e) \times \delta DTOF(e\pi)$ selection has a negligible effect on the result. Thus the value of $\pm 0.672\%$ is taken as the relative systematic uncertainty of the DTOF analysis.

TOF efficiency correction

A correction factor $C_\epsilon = 1.0029 \pm 0.0067$ to the efficiency of $\delta DTOF$ selection is applied, to take in account for the shift of the central value. Corrected value is:

$$\epsilon_{DTOF} = (66.24 \pm 0.44)\% \quad (6.13)$$

6.3.3 Others systematic uncertainties

Those efficiencies that are not determined using the $K_L \rightarrow \pi e \nu$ control sample, *i.e.* for the track-to-cluster association (Section 6.2.1) and for the $\pi^+\pi^-$ selection (Section 5.1), are evaluated with two alternative methods. In both cases the two results are in good agreement and the value obtained with the method that is less biased is taken as related efficiency, and the difference of the two values as the systematic error. The results are listed in Table 6.8.

6.4 Result

Using the formula introduced before, the number of events result of the fit to the m_e^2 distribution, and the efficiencies evaluated for the analysis selections (Table 6.7 and Section 6.3.2 for $\epsilon_{D\text{TOF}}$ correction), the branching ratio for the $K_S \rightarrow \pi e \nu$ decay is derived as

$$BR(K_S \rightarrow \pi e \nu) = \frac{N_{\pi e \nu}}{\epsilon_{\pi e \nu}} \times \frac{\epsilon_{\pi^+ \pi^-}}{N_{\pi^+ \pi^-}} \times R_\epsilon \times BR(K_S \rightarrow \pi^+ \pi^-) \quad (6.14)$$

with:

$$BR(K_S \rightarrow \pi^+ \pi^-) = (69.020 \pm 0.051)\%$$

$$N_{\pi^+ \pi^-} / \epsilon_{\pi^+ \pi^-} = (292.10 \pm 0.26) \times 10^6 \text{ events};$$

$$N_{\pi e \nu} = 49\,647 \pm 191 \text{ events};$$

$$\epsilon_{\pi e \nu} = (19.41 \pm 0.14)\%$$

$$R_\epsilon = 1.1882 \pm 0.0012$$

$$BR(K_S \rightarrow \pi^+ \pi^-) = (69.0196 \pm 0.051)\%.$$

The result of the branching ratio is:

$$BR(K_S \rightarrow \pi e \nu) = (7.181 \pm 0.028_{\text{stat}} \pm 0.054_{\text{syst}}) \times 10^{-4} = (7.181 \pm 0.061) \times 10^{-4} \quad (6.15)$$

where the statistical uncertainty is the statistical error of $N_{\pi e \nu}$ and all other uncertainties are combined in the systematic uncertainty. The relative error is:

$$\pm 0.39\%_{\text{stat}} \pm 0.75\%_{\text{syst}} = \pm 0.85\% \quad (6.16)$$

Chapter 7

$$K_S \rightarrow \pi\mu\nu$$

7.1 Selection

After the K_L -crash recognition and the vertex requirement, the data sample contains 301,645,500 events and its composition, as predicted by simulation, is shown in Table 6.1. The analysis scheme is the same as for the $K_S \rightarrow \pi e\nu$.

7.1.1 Preselection and multivariate analysis

The preselection criteria for the $K_S \rightarrow \pi\mu\nu$ analysis are the same of the $K_S \rightarrow \pi e\nu$ analysis:

- $\theta_{\text{crash}} > 15^\circ$;
- track-to-cluster association (TCA): for both tracks are required to be associated to a cluster, with $\theta_{\text{clu}}^i > 15^\circ$;
- $\delta p > -95$ MeV.

At this stage, after the requirement on the cluster polar angles, the track to cluster association and the track momentum selection, the number of events in the data and MC samples is 212,720,400 and 234,516,600, respectively.

The same variables used in the $K_S \rightarrow \pi e\nu$ analysis are used for building the BDT classifier and the same cuts are applied to the tracks momenta and the K_S momentum determined from the tag

- $p < 320$ MeV for both tracks;
- $\delta p < 190$ MeV.

The last two cuts do not affect the signal, the number of data reduces to 211,368,600. The event composition is the same as in Table 6.2. The training of the BDT classifier was done on a sample of 5,000 $K_S \rightarrow \pi\mu\nu$ events and a sample of 50,000 background events and samples of same size were used for the test. After training and test, the classification

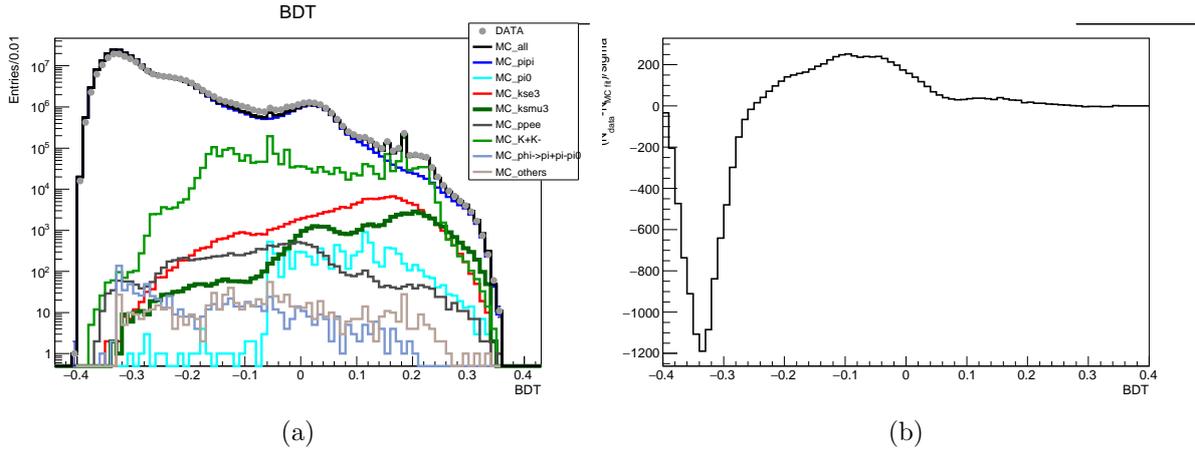


Figure 7.1: Distribution of the BDT classifier output for data, simulated signal and backgrounds

was run on all events of the MC and data sample. The distribution of the BDT classifier output is shown in Figure 7.1 for data, simulated signal and backgrounds.

Given the lower discrimination power relative to the $K_S \rightarrow \pi e\nu$ analysis a harder cut is needed to reject the $K_S \rightarrow \pi^+\pi^-$ background. The chosen value is:

$$BDT > 0.18 . \quad (7.1)$$

After this selection, the number of events in the data sample is 585,661 and its composition as predicted by simulation is listed in Table 7.1

	MC events	fraction [%]	relative efficiency [%]
DATA	585 661		0.28
MC	553 295		0.24
$K_S \rightarrow \pi\mu\nu$	20 505	3.71	43.90
$K_S \rightarrow \pi^+\pi^-$	174 937	31.62	0.08
$\phi \rightarrow K^+K^-$	331 120	59.85	14.97
$K_S \rightarrow \pi^0\pi^0$	857	0.15	10.76
$K_S \rightarrow \pi e\nu$	25436	4.60	46.40
$K_S \rightarrow \pi^+\pi^-e^+e^-$	374	0.07	3.40
$\phi \rightarrow \pi^+\pi^-\pi^0$	4	0.001	0.51
others	62	0.01	8.24

Table 7.1: Number of events divided by channel after the BDT selection.

7.1.2 Selection based on time of flight

Since the BDT classification is not sufficient to separate the signal from background, a time of flight analysis was performed. The time of flight selection follows the same paths

as for the $K_S \rightarrow \pi e\nu$ analysis (Sec. 6.1.3), the variable used is:

$$\delta DTOF = DTOF_1 - DTOF_2$$

that depends on the mass hypothesis for the two tracks.

Since the main goal is to discriminate $K_S \rightarrow \pi\mu\nu$ against $K_S \rightarrow \pi^+\pi^-$, $\delta DTOF$ is first computed in the $\pi\pi$ hypothesis:

$$\delta DTOF(\pi\pi) = DTOF_1(\pi) - DTOF_2(\pi)$$

The distribution of the $\delta DTOF(\pi\pi)$ variable is shown in Figure ???. It is evident, given the small difference in mass between muon and charged pion, that the distributions of the signal and the main background are in the same range though they show different shapes.

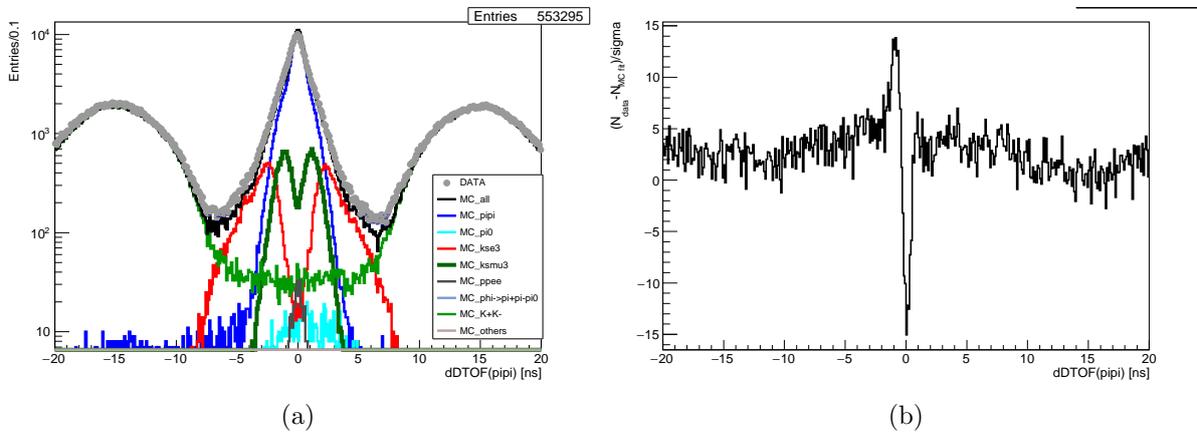


Figure 7.2: (a) Distribution of $\delta DTOF(\pi\pi)$ for the data and MC signal events, background events and their sum; (b) significance defined as the difference between data and MC divided by the statistical error.

The cut chosen is:

$$1 \text{ ns} < |\delta DTOF(\pi\pi)| < 3 \text{ ns}$$

to avoid the large $\pi\pi$ peak close to $\delta DTOF = 0$ and the contribution of K^+K^- and $K_S \rightarrow \pi e\nu$ for $|\delta DTOF| > 3$ ns, and to have good efficiency for the signal. Figure 7.2 shows a non satisfactory agreement between data and MC, thus a control sample of $K_L \rightarrow \pi\mu\nu$ events is used to determine the efficiency of the $\delta DTOF$ selection.

After the applied selection, both $\pi-\mu$ and $\mu-\pi$ hypotheses are tested:

$$\delta DTOF(\pi\mu) = DTOF_1(\pi) - DTOF_2(\mu)$$

$$\delta DTOF(\mu\pi) = DTOF_1(\mu) - DTOF_2(\pi)$$

The two-dimensional distribution in the two hypotheses, $\delta DTOF(\pi\mu) \times \delta DTOF(\mu\pi)$, and the variable $\delta DTOF(\mu)$ defined as:

$$\delta DTOF(\mu) = \min(|\delta DTOF(\pi\mu)|, |\delta DTOF(\mu\pi)|) \quad (7.2)$$

are represented in Figure 7.3. The cut chosen is:

$$|\delta DTOF(\mu)| < 0.5 \text{ ns}$$

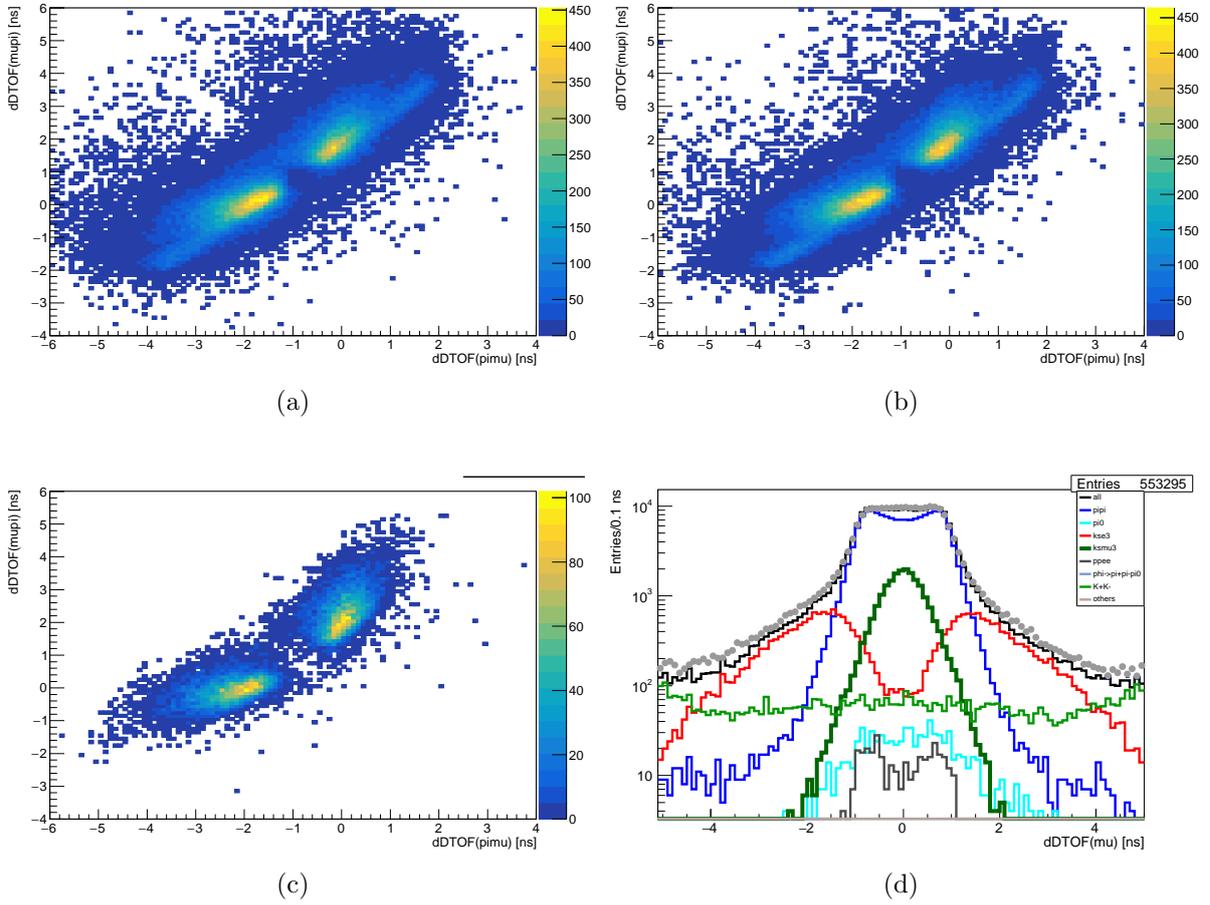


Figure 7.3: Two-dimensional distribution $\delta DTOF(\pi\mu) \times \delta DTOF(\mu\pi)$ for data (a), MC signal and background (b), and for signal MC only (c), and $\delta DTOF(\mu)$ define by Eq. 7.2 (d).

7.1.3 Extraction of the signal

After this selection, the number of events in the data sample is 38,686 and its composition as predicted by simulation is listed in Table 7.2.

After particle identification (Sec. 6.1.3) the invariant mass of the particle identified as the muon is calculated as:

$$m_\mu^2 = (E_{K_S\text{tag}} - E_\pi - p_{\text{mis}})^2 - p_\mu^2 \quad (7.3)$$

where $p_{\text{mis}}^2 = (\vec{p}_{K_S\text{tag}} - \vec{p}_\mu - \vec{p}_\pi)^2$, $p_{K_S\text{tag}}$ and $E_{K_S\text{tag}}$ are the K_S momentum and energy estimated from the tag, and p_μ and p_π are the momentum of the muon and the pion track.

	MC events	fraction [%]	relative efficiency [%]	tot relative efficiency [%]
DATA	38 686		6.61	0.01
MC	36 444		6.59	0.01
$\bar{K}_S \rightarrow \pi\mu\nu$	9 424	25.86	45.96	6.75
$K_S \rightarrow \pi^+\pi^-$	25 853	70.94	14.78	< 0.001
$\phi \rightarrow K^+K^-$	475	1.30	1.14	< 0.001
$K_S \rightarrow \pi^0\pi^0$	215	0.59	25.09	0.71
$K_S \rightarrow \pi e\nu$	448	1.23	1.76	0.17
$K_S \rightarrow \pi^+\pi^-e^+e^-$	29	0.08	7.75	0.16
$\phi \rightarrow \pi^+\pi^-\pi^0$	0	0	0	0
others	0	0	0	0

Table 7.2: Number of events by channel after $\delta DTOF$ selection

The same smearing procedure of the MC momenta as in Sec. 6.1.4 is applied. Figure 7.4 shows the resulting distribution of m_μ^2 .

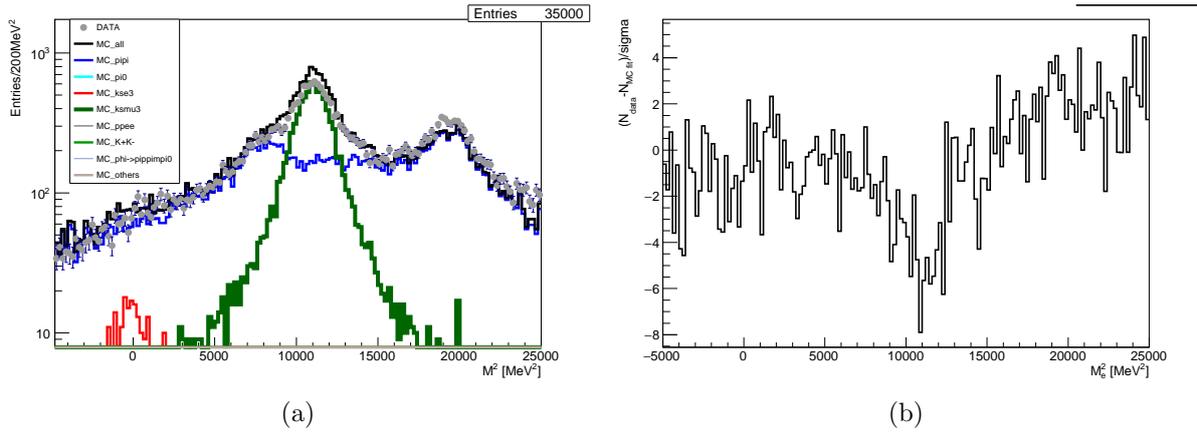


Figure 7.4

The number of the signal events is extracted by a fit to the m_μ^2 distribution with the MC shapes of three components: $\pi\mu\nu$, $\pi^+\pi^-$ and all others channels. Table 7.3 lists the results of the fit, and Figure 7.5 shows the fitted distributions for data and MC. The result is:

$$N_{\pi\mu\nu} = 7223 \pm 180 \text{ events} . \tag{7.4}$$

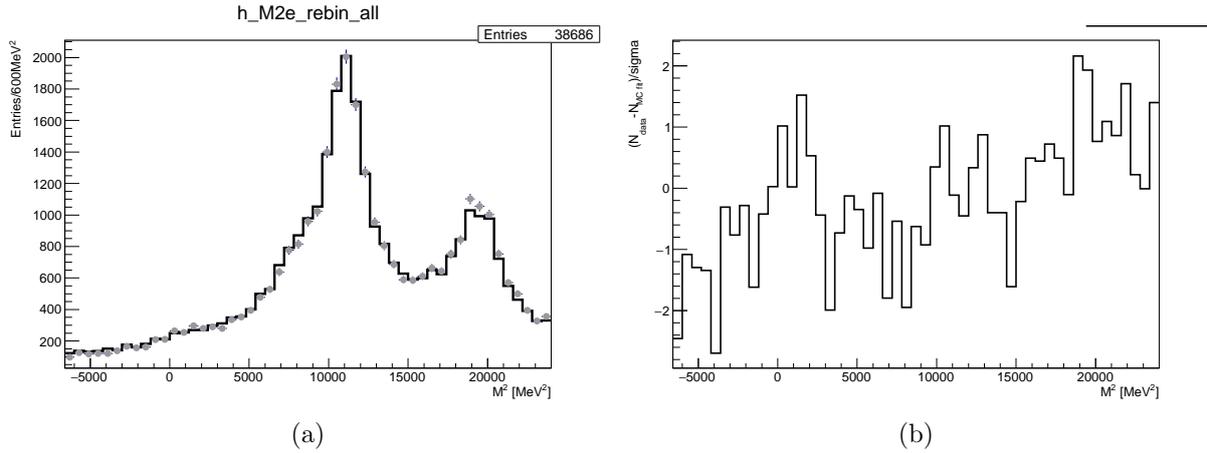


Figure 7.5

	fraction	events	relative error [%]
$\pi\mu\nu$	0.23	$7\,223 \pm 180$	2.49
$\pi^+\pi^-$	0.77	$23\,764 \pm 270$	1.13
Total		30 987	
χ^2/ndf	2.59		

Table 7.3: Fit output

7.2 Determination of efficiencies

7.2.1 $K_S \rightarrow \pi\mu\nu$ control sample

The control sample uses $K_L \rightarrow \pi\mu\nu$ events tagged by $K_S \rightarrow \pi^+\pi^-$ decays, in the same way as discussed in Sec. 6.2.1. The same selection on the radial distance of the K_L vertex is required:

$$1 \text{ cm} < \rho_{\text{VTX}}^{K_L} < 5 \text{ cm}$$

Beside $K_L \rightarrow \pi\mu\nu$ events (with an expected BR of 0.270) the sample contains $K_L \rightarrow \pi e\nu$ (BR = 0.405) and $K_L \rightarrow \pi^+\pi^-\pi^0$ (BR = 0.125) events. To suppress these backgrounds without introducing biases the event selection is made with the same method as in Sec. 6.2.1. The selection of the control sample is made as follows:

- a common cut on the missing mass distribution shown in Figure 6.14, $m_{\text{mis}}^2 < 15,000 \text{ MeV}^2$ to reject $K_L \rightarrow \pi^+\pi^-\pi^0$ events;
- a cut on the TOF variables is applied to evaluate the efficiency of the preBDT and BDT selection;
- a cut on $\delta m \times m_{\text{mis}}^2$ is applied to evaluate the efficiency of the TCA and TOF selections.

The missing mass and the others variable are defined in Sec. 6.2.1.

PreMVA and MVA control sample and efficiency

The control sample for evaluating the efficiency of the preselection and BDT cut is selected with a cut on the $\delta DTOF$ variable. More information are in Sec. 6.2.1 Distribution of these variable are shown, separately for data, total MC simulation and for the MC signal only, in Figure 7.6.

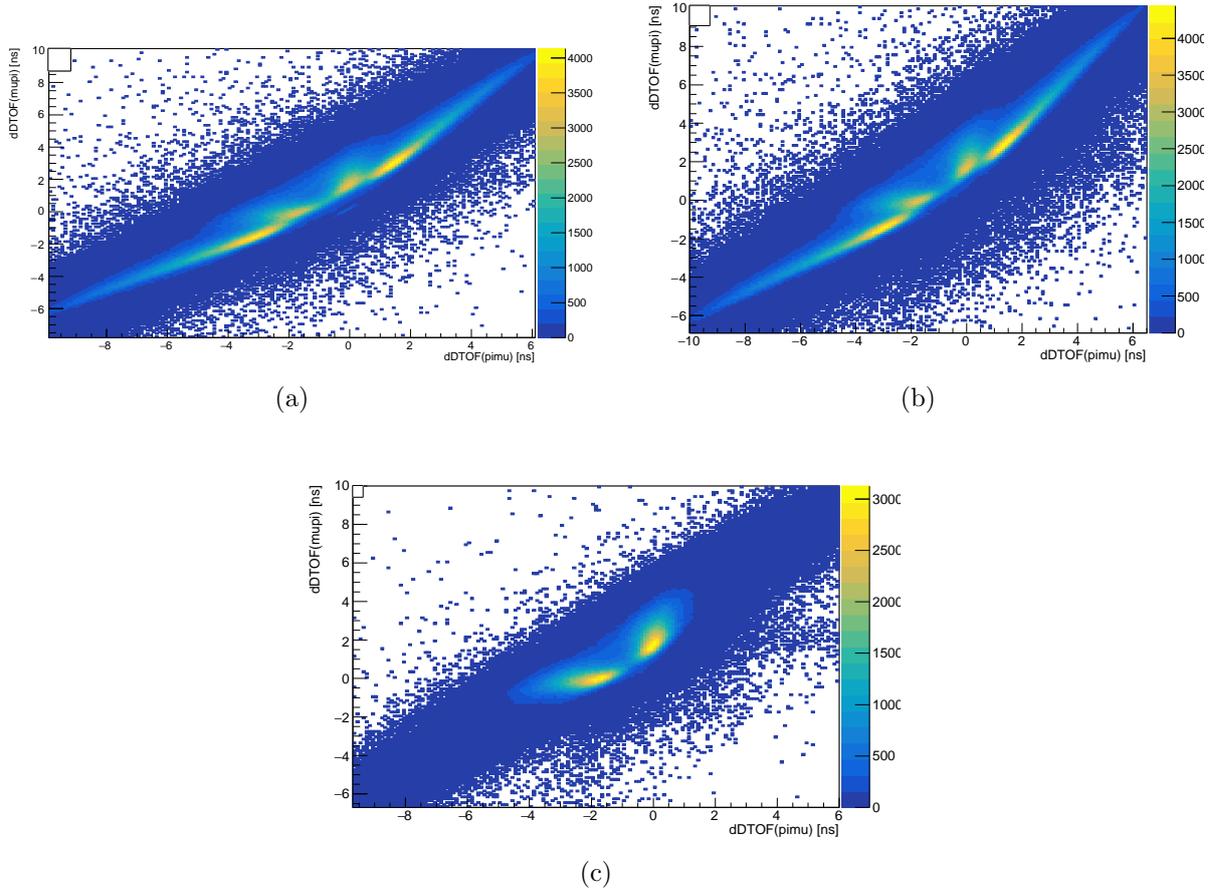


Figure 7.6: C provvisoria

The following selection, in OR of each other, is applied, to select the $K_L \rightarrow \pi e \nu$ events:

$$[\delta DTOF(\pi\mu) + 0.075]^2 + [\delta DTOF(\mu\pi) - 1.8]^2 < 0.4^2$$

$$[\delta DTOF(\pi\mu) + 1.7]^2 + [\delta DTOF(\mu\pi) - 0.12]^2 < 0.4^2$$

where all values are in ns.

After this selection, the purity of the sample determined from MC is about 87%. The comparison between the MC BDT output distributions of the analysis and control samples is shown in Figure 7.7. Then the efficiencies of the preselection and of the BDT cut are

evaluated with the method described in Eq. 6.7. The results are shown in Table 7.5, second and fourth row.

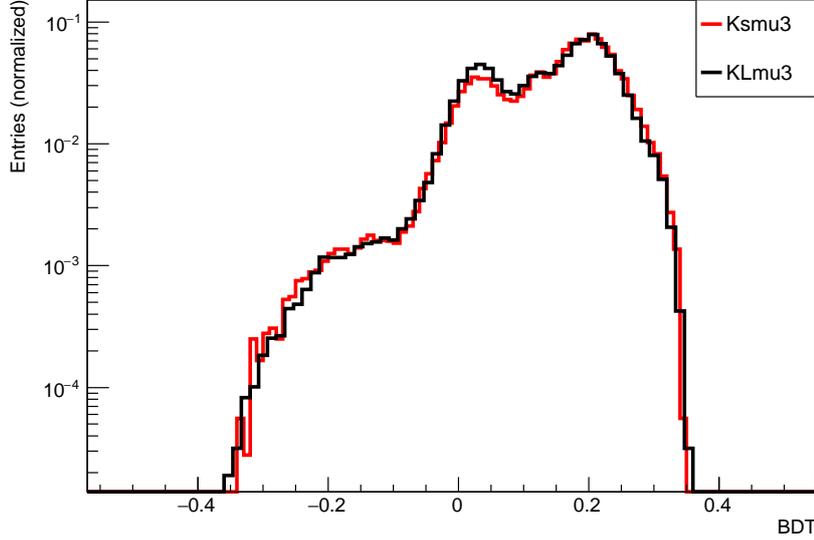


Figure 7.7: Monte Carlo BDT distributions for the $K_L \rightarrow \pi\mu\nu$ and $K_S \rightarrow \pi\mu\nu$ selected events.

TCA and DTOF control sample and efficiency

The control sample for evaluating the TCA and TOF efficiencies is selected using only tracking variables, as for the $K_S \rightarrow \pi e\nu$ analysis in Section 6.2.1. To suppress background a triangular cut in the $\delta m \times m_{\text{mis}}^2$ variables is applied, as shown in Figure 7.8:

- $\delta m < 0.015 \times (m_{\text{mis}}^2 - 1200) + 15.5 \text{ MeV}$
- $\delta m > 0.055 \times m_{\text{mis}}^2 + 32.4 \text{ MeV}$
- $-180 \text{ MeV} < \delta m < -5 \text{ MeV}$

Track-to-cluster association is required for both tracks of the tag, $K_S \rightarrow \pi^+\pi^-$ decay. Then the efficiency for the TCA selection is computed in the usual way from Eq. 6.7. The result is shown in Table 7.5.

After evaluating the TCA efficiency, the control sample is selected in the same way as the analysis sample before the TOF selections (Section 6.1.3):

- preselection (Sec. 6.1.1)
- same BDT cut ($BDT > 0.18$)

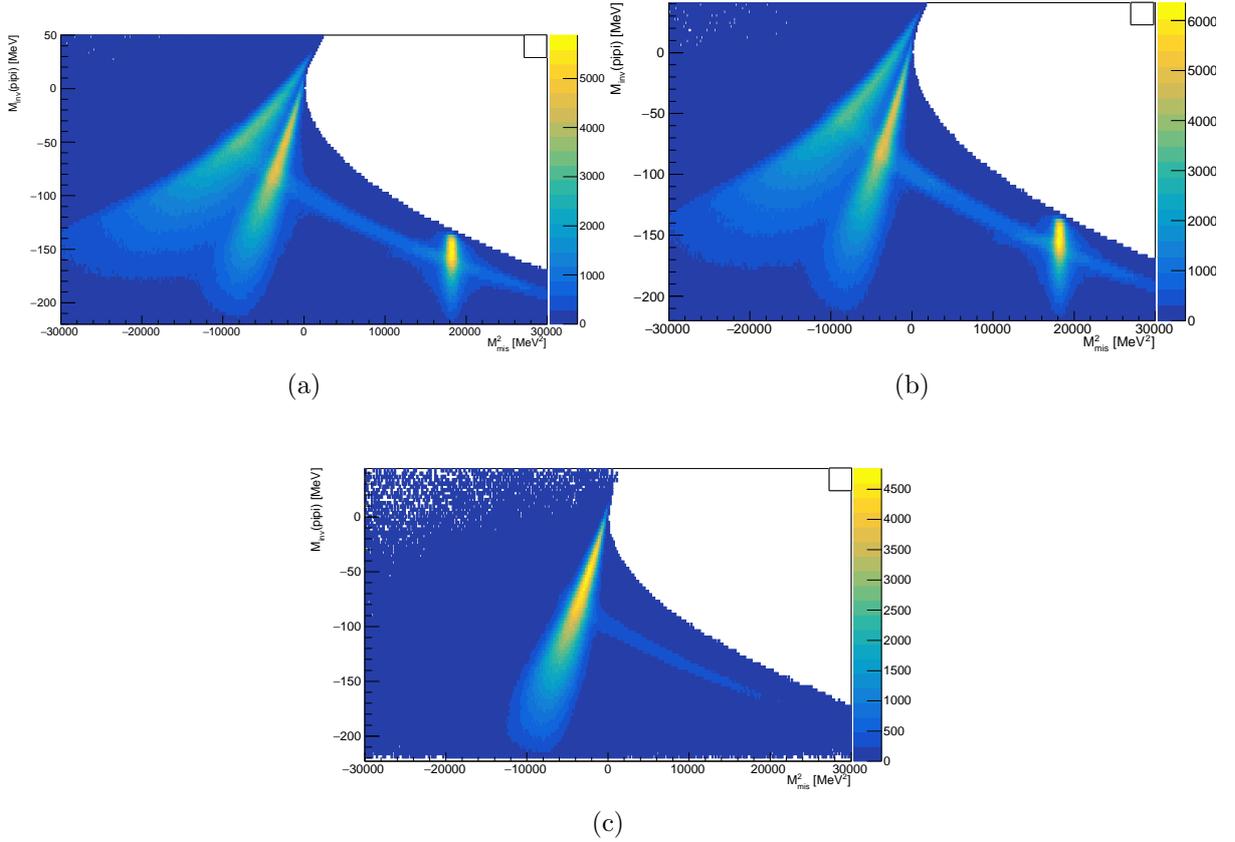


Figure 7.8

As in Section 6.2.1, the request that the first cluster in time arriving at the calorimeter has to be associated to the K_L decay is added. The comparison between the MC $\delta DTOF(\pi\pi)$ and $\delta DTOF(\mu)$ distributions of the analysis and control samples is shown in Figure 7.9.

The results of the TOF efficiencies for the $\delta DTOF(\pi\pi)$ and the $\delta DTOF(\mu)$ selections are shown in Table 7.5.

7.2.2 Monte Carlo efficiency

Only a couple of efficiencies are calculated directly from MC. These efficiencies are high and are related to:

- the selection on the angle of the crash cluster, $\theta_{\text{crash}} > 15^\circ$ (see Sec. 4.2.1);
- the lower cut on $\delta p < -95$ (see Sec. 6.1.1), to equalise the control sample and data sample distributions of δp .

The combination of these two efficiencies is reported in Table 7.4, third row.

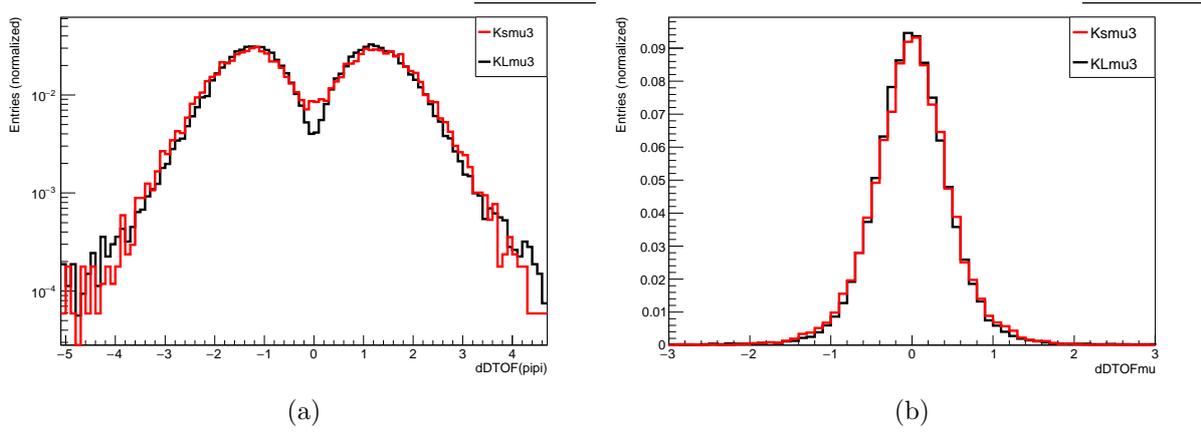


Figure 7.9: MC distributions of $\delta DTOF(\pi\pi)$ and $\delta DTOF(\mu)$ for the analysis ($K_S \rightarrow \pi\mu\nu$) and control ($K_L \rightarrow \pi\mu\nu$) sample.

7.2.3 Determination of R_ϵ

For all common selections there is no need to compute separately the efficiencies for the control and the normalisation sample. In fact, in Eq. 4.1 only the ratio of the two is needed. The ratio R_ϵ , as explained in Sec. 4, is composed of several factors: Trigger, FILFO, Event Classification, K_L -crash and K_S identification efficiency. All these efficiencies are taken from MC simulation and listed in Table 7.4. The result is:

$$R_\epsilon = 1.472 \pm 0.004 \quad (7.5)$$

Selection	Ratio of efficiency
Trigger	1.0649 ± 0.0005
FILFO	1.0113 ± 0.0002
Event classification	1.1406 ± 0.0007
T0 time fix	1.0135 ± 0.0002
K_L -crash	1.1283 ± 0.0022
K_S ID	1.0481 ± 0.0012
R_ϵ	1.472 ± 0.003

Table 7.4: R_ϵ ; single errors are strongly correlated, so the error on R_ϵ is smaller than the quadratic sum of the errors of the single efficiencies.

7.2.4 Summary of efficiencies

Selection	Efficiency
TCA	0.347 ± 0.002
CS preselection	0.986 ± 0.004
MC preselection	0.996 ± 0.002
BDT selection	0.417 ± 0.003
TOF selection	0.401 ± 0.003
FIT interval	0.989 ± 0.001
Total	0.0563 ± 0.0005

Table 7.5: Efficiencies of the analysis cuts; single efficiency errors are strongly correlated, so the error of the total efficiency is smaller than the quadratic sum of the errors of the single efficiencies. For $\epsilon_{D\text{TOF}}$ a correction will be applied (See Section 7.3.2 for details)

7.3 Systematic uncertainties

7.3.1 BDT scan and systematic uncertainty

Due to the good agreement between data and MC on the BDT distribution in the region of the applied cut (Figure 7.1), only a variation of one standard deviation on the BDT cut is done ($\sigma_{BDT} = 0.005$, see Sec. 6.3.1). The largest variation is taken as the systematic uncertainty related to the BDT cut. The result is:

$$BDT \text{ syst} = 0.30\%$$

and is reported in Table 7.6.

7.3.2 TOF scan and systematic uncertainty

As explained in Sec. 6.3.2, the average $\delta D\text{TOF}$ resolution is estimated to be $\sigma_{\delta D\text{TOF}} = 270$ ps. To choose the cut on $\delta D\text{TOF}(\pi\pi)$ that ensures good stability of the result, a scan of the lower cut of $\delta D\text{TOF}(\pi\pi)$ is performed. The $\delta D\text{TOF}(\pi\pi)$ cut is varied in steps of 0.25 ns, approximately equal to one standard deviation, and the efficiency is recomputed at each step. The interval chosen is $0.5 < \delta D\text{TOF}(\pi\pi) < 1.5$ ns, corresponding to about four times the resolution. A lower value of the cut brings too large background, while a higher value brings too small signal. The result is shown in Figure 7.10.

A dependence of $N_{\pi\mu\nu}/\epsilon_{\pi\mu\nu}$ from the $\delta D\text{TOF}(\pi\pi)$ cut is observed. An error band is defined as the difference between the maximum and the minimum value in Figure 7.10; the central value is taken as the number of counts, while half of the band width is taken as systematic error due to the choice of the $\delta D\text{TOF}(\pi\pi)$ cut. The result is:

$$\delta D\text{TOF}(\pi\pi) \text{ syst} = 2.97\%$$

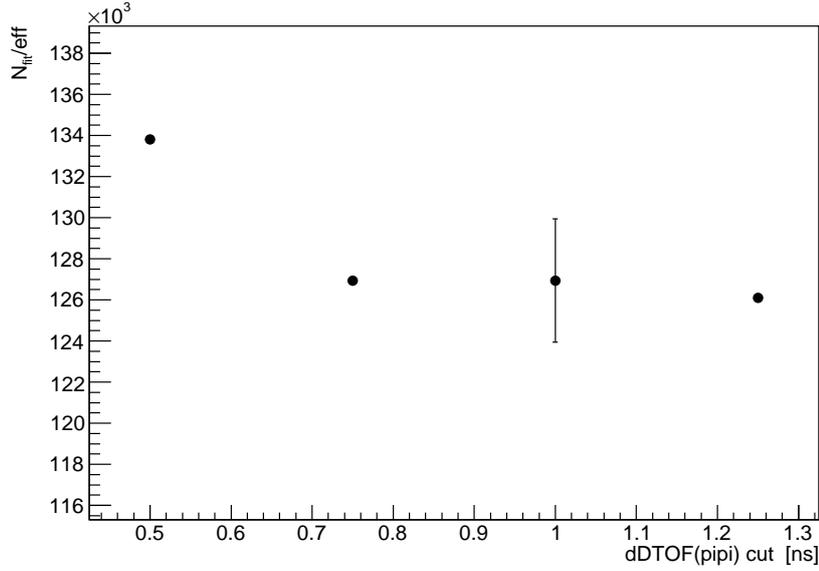


Figure 7.10: Number of fit counts divided by total efficiency as function of the cut on the BDT classifier output.

In comparison, the variation by one standard deviation of the higher cut of $\delta DTOF(\pi\pi)$, and of the $\delta DTOF(\pi\mu) \times \delta DTOF(\mu\pi)$ selection are found to have a negligible effect. Thus the value of $\pm 2.97\%$ is taken as the relative systematic uncertainty of the TOF analysis, listed in Table 7.6.

TOF efficiency correction

A correction factor $C_\epsilon = 0.9770 \pm 0.0290$ to the efficiency of $\delta DTOF$ selection is applied, to take in account for the shift of the central value. Corrected value is:

$$\epsilon_{DTOF} = (39.2 \pm 1.2)\% \quad (7.6)$$

Table 7.6: Systematic errors

Selection	Relative systematic error [%]
TOF	2.97
BDT	0.30
K_L CS statistics	0.45
$\pi^+\pi^-$ CS statistics	< 0.01
MC sample statistics	0.68
$\pi^+\pi^-$ efficiency	0.09
Total	3.09

7.4 Result

Using the formula introduced before, the number of events result of the fit to the m_e^2 distribution, and the efficiencies evaluated for the analysis selections (Table 7.5 and Section 7.3.2 for $\epsilon_{D_{TOF}}$ correction), the branching ratio for the $K_S \rightarrow \pi e \nu$ decay is derived as

$$BR(K_S \rightarrow \pi\mu\nu) = \frac{N_{\pi\mu\nu}}{\epsilon_{\pi\mu\nu}} \times \frac{\epsilon_{\pi^+\pi^-}}{N_{\pi^+\pi^-}} \times R_\epsilon \times BR(K_S \rightarrow \pi^+\pi^-) \quad (7.7)$$

with:

$$N_{\pi^+\pi^-}/\epsilon_{\pi^+\pi^-} = (292.10 \pm 0.26) \times 10^6 \text{ events}$$

$$N_{\pi\mu\nu} = 7\,223 \pm 180 \text{ events;}$$

$$\epsilon_{\pi\mu\nu} = (5.50 \pm 0.16)\%$$

$$R_\epsilon = 1.472 \pm 0.003$$

$$BR(K_S \rightarrow \pi^+\pi^-) = (69.020 \pm 0.051)\%$$

The final results is:

$$BR(K_S \rightarrow \pi\mu\nu) = (4.57 \pm 0.11_{\text{stat}} \pm 0.14_{\text{syst}}) \times 10^{-4} = (4.57 \pm 0.18) \times 10^{-4} \quad (7.8)$$

where the statistical uncertainty is the statistical error of $N_{\pi\mu\nu}$ and all other uncertainties are combined in the systematic uncertainty. The relative error is:

$$\pm 2.5\%_{\text{stat}} \pm 3.1\%_{\text{syst}} = \pm 4.0\% \quad (7.9)$$

Chapter 8

Conclusions

Using the data sample collected in 2004–2005 by the KLOE experiment at DAΦNE, corresponding to 1.63 pb^{-1} of e^+e^- integrated luminosity the Branching Ratios of the two semileptonic decay are computed through the formula:

$$BR(K_S \rightarrow \pi\mu\nu) = \frac{N_{\pi\mu\nu}}{\epsilon_{\pi\mu\nu}} \times \frac{\epsilon_{\pi^+\pi^-}}{N_{\pi^+\pi^-}} \times R_\epsilon \times BR(K_S \rightarrow \pi^+\pi^-) \quad (8.1)$$

where $BR(K_S \rightarrow \pi^+\pi^-) = (69.0196 \pm 0.051)\%$.

The counts for the $K_S \rightarrow \pi e\nu$, $K_S \rightarrow \pi\mu\nu$ and for the $K_S \rightarrow \pi^+\pi^-$ normalisation sample are:

$$N_{\pi^+\pi^-}/\epsilon_{\pi^+\pi^-} = (292.10 \pm 0.26) \times 10^6 \text{ events};$$

$$N_{\pi e\nu} = 49\,647 \pm 191 \text{ events};$$

$$N_{\pi\mu\nu} = 7\,223 \pm 180 \text{ events};$$

The total efficiencies for the semileptonic decays are:

$$\epsilon_{\pi e\nu} = (19.41 \pm 0.07)\%$$

$$\epsilon_{\pi\mu\nu} = (5.50 \pm 0.16)\%$$

The ratio of the common efficiencies are:

$$R_\epsilon^e = 1.1882 \pm 0.0012$$

$$R_\epsilon^\mu = 1.472 \pm 0.003$$

The result for the branching ratio of $K_S \rightarrow \pi e \nu$ is:

$$BR(K_S \rightarrow \pi e \nu) = (7.181 \pm 0.028_{\text{stat}} \pm 0.054_{\text{syst}}) \times 10^{-4} = (7.181 \pm 0.061) \times 10^{-4}$$

with a relative error of $\pm 0.39\%_{\text{stat}} \pm 0.75\%_{\text{syst}} = \pm 0.85\%$.

The previous result from KLOE, based on an independent data sample corresponding to 0.40 pb^{-1} of e^+e^- integrated luminosity, is

$$BR(K_S \rightarrow \pi e \nu) = (7.046 \pm 0.077_{\text{stat}} \pm 0.050_{\text{syst}}) \times 10^{-4} = (7.046 \pm 0.091) \times 10^{-4}$$

For the $K_S \rightarrow \pi \mu \nu$ branching ratio, the result is:

$$BR(K_S \rightarrow \pi \mu \nu) = (4.57 \pm 0.11_{\text{stat}} \pm 0.14_{\text{syst}}) \times 10^{-4} = (4.57 \pm 0.18) \times 10^{-4}$$

with a relative error of $\pm 2.5\%_{\text{stat}} \pm 3.1\%_{\text{syst}} = \pm 4.0\%$.

This is the first measurement of this decay and the result is in good agreement with the expected value of

$$BR(K_S \rightarrow \pi \mu \nu) = (4.69 \pm 0.05) \times 10^{-4}.$$

8.1 V_{us} determination

8.1.1 $K_S \rightarrow \pi e \nu$

Starting from Eq. 2.23 and from the branching ratio, it is possible to determine $|V_{us}|f_+^2(0)$:

$$|V_{us}|f_+(0) = \left[\frac{192\pi^3 BR(K_S \rightarrow \pi e \nu)}{\tau_s G_F^2 M_K^5 S_{EW} I_K^e (1 + \delta_K^e)} \right]^{1/2} \quad (8.2)$$

where G_F is the Fermi constant, M_K is the kaon mass, S_{EW} is the short-distance radiative correction, δ_K^e is the mode-dependent long-distance radiative correction, $f_+(0)$ is the form factor at zero momentum transfer to the $e\nu$ system, $I_K^e = 0.15480 \pm 0.00030$ is the phase-space integral, determined starting from the value in [5] multiplied for 1.5 due to different normalization convention. The values of $S_{EW} = 1.0232$ [26], $\delta_K^\ell = (0.50 \pm 0.10) \times 10^{-2}$ [27] are calculated from theory. The result is :

$$|V_{us}|f_+(0) = 0.2181 \pm 0.0010 \quad (8.3)$$

Using $f_+(0) = 0.9704 \pm 0.0032$ from [29]:

$$|V_{us}| = 0.2247 \pm 0.0012 \quad (8.4)$$

Using the updated value of $V_{ud} = 0.97417(21)$ from the superallowed nuclear beta decay [30] and $V_{ub} = 0.00413(49)$ from B-meson decays [31], it is possible to perform a unitarity test on CKM matrix first row, defining:

$$\Delta = 1 - |V_{ud}|^2 - |V_{us}|^2 - |V_{ub}|^2 = (4.7 \pm 6.9) \times 10^{-4} \quad (8.5)$$

equal to zero within 0.68σ , demonstrating the unitarity of the CKM matrix.

8.1.2 $K_S \rightarrow \pi\mu\nu$

Similarly:

$$|V_{us}|f_+(0) = \left[\frac{192\pi^3 BR(K_S \rightarrow \pi\mu\nu)}{\tau_s G_F^2 M_K^5 S_{EW} I_K^\mu (1 + \delta_K^\mu)} \right]^{1/2} \quad (8.6)$$

where G_F is the Fermi constant, M_K is the kaon mass, S_{EW} is the short-distance radiative correction, δ_K^μ is the mode-dependent long-distance radiative correction, $f_+(0)$ is the form factor at zero momentum transfer to the $\mu\nu$ system, $I_K^\mu = 0.10165 \pm 0.00080$ [8] is the phase-space integral. The values of $S_{EW} = 1.0232$ [26], $\delta_K^\ell = (1.38 \pm 0.10) \times 10^{-2}$ [27] are calculated from theory. The result is :

$$|V_{us}|f_+(0) = 0.2138 \pm 0.0043 \quad (8.7)$$

Using $f_+(0) = 0.9704 \pm 0.0032$ from [29], the V_{us} matrix element is:

$$|V_{us}| = 0.220 \pm 0.0045 \quad (8.8)$$

8.2 Lepton universality test

With the results of this work, it is possible to perform a lepton universality test. If lepton universality holds:

$$r_{\mu e} = \frac{R_{\mu e}^{exp}}{R_{\mu e}^{SM}} = 1 \quad (8.9)$$

where $R_{\mu e}^{exp}$ and $R_{\mu e}^{SM}$ are the experimental and theoretical Standard Model value, respectively, defined as:

$$r_{\mu e}^{exp} = \frac{\Gamma(K_s \rightarrow \pi\mu\nu)}{\Gamma(K_s \rightarrow \pi e\nu)} \quad \text{and} \quad r_{\mu e}^{th} = \frac{1 + \delta_K^\mu}{1 + \delta_K^e} \frac{I_K^\mu}{I_K^e} \quad (8.10)$$

where δ_K^ℓ are mode-dependent long-distance radiative correction, and I_K^ℓ are decay phase space integrals (see Section 2.1.2 for details). Using $\frac{I_K^\mu}{I_K^e} = 0.6622 \pm 0.0018$ from the measurement of the semileptonic decays of the KTeV experiment [8] and $\frac{1 + \delta_K^\mu}{1 + \delta_K^e} = 1.0058 \pm 0.0010$ [9]:

$$R_{\mu e}^{SM} = 0.6660 \pm 0.0019 \quad (8.11)$$

Starting from the result of this work, the experimental value of the ratio is:

$$R_{\mu e}^{exp} = 0.636 \pm 0.026 \quad (8.12)$$

and then:

$$r_{\mu e} = 1.047 \pm 0.043 \quad (8.13)$$

consistent with 1 in 1.1σ .

8.3 Perspectives

After completion of the KLOE data-taking, an upgrade program of the DAFNE accelerator complex and of the detector was undertaken. The KLOE2 Collaboration has collected new data at the ϕ peak corresponding to an integrated luminosity of about $5 fb^{-1}$, thus increasing the data sample by a factor of three respect to that analysed here. The increased data sample will contribute in improving the measurements of the branching fraction of kaon semileptonic decays and the other inputs to Eq.8.2 whose present value is mostly determined by KLOE analyses, e.g. the K_S lifetime and semileptonic decay form factors. The increased data sample will also allow to develop methods to reduce the systematic errors that at present limit the value of the $K_S \rightarrow \pi e \nu$ branching fraction, reaching a precision similar to that of the K^\pm decay and a measurement of $|V_{us}|f_+(0)$ competitive with the present K_L results from fixed target experiments. The measurement of $K_S \rightarrow \pi \mu \nu$ is a new entry in the table of kaon semileptonic kaon decays and will surely benefit of the increase in statistics, a reduction of the error by a factor two is reachable with the two independent data samples.

Appendices

Appendix A

Boosted Decision Tree

Boosted Decision Tree is one of the most used algorithm for MultiVariate Analysis (MVA). This is due manly to:

- Shorter training and testing time respect to other algorithm (for example Neural Network)
- Good separation power between signal and background
- Easiness to understand the basical mechanism of classificatio (respect to other algorithm)

Like every MVA, the BDT algorithm have to be trained to learn how to discriminate between signal and background and tested, both stage on MC simulation, as explained in Sec. 6.1.2; number of events used in the different stage are listed in Tab. A.1.

BDT is the evolution of the Decision Tree, and solve the problem of the instability of a single tree, where the output could be really different in case of fluctation of the sample, while the output of weighted average of $O(1000)$ tree will be stable. After every tree, the misclassified event are reweighted and the procedure repeated; each tree is scored. Every event is classified with the avarage output (signal o background) of all the tree, weighted with the tree-scores:

$$T(x_i) = \sum_{m=1}^{N_{tree}} \alpha_m T_m(x_i) \quad (\text{A.1})$$

where $T_m(x_i)$ is +1 (-1) if the x_i event is classified as signal (background), and α_m is the score of each tree.

Different method can be used to boost the tree, in this analysis Adaptive Boost (Adaboost) is used. Score of each tree is defined as:

$$\alpha_m = \beta \times \ln \frac{1 - err_m}{err_m} \quad (\text{A.2})$$

where $err_m = \sum_j W_j$, W_j is the weight of the event and j runs over bad classified events, β is a constant (Tab. A.1). Misclassified event are reweighted: $W_j \rightarrow W_j \exp_m^\alpha$

In Tab. A.1 main parameter used are listed [32]; *Min node size* is the minimum percentage of training events required in a leaf node, while *nCuts* is the number of grid points in variable range used in finding optimal cut in node splitting.

Table A.1: BDT parameter

Parameter	value
N. of signal training events	10 000
N. of background training events	30 000
N. of signal test events	5 000
N. of background test events	30 000
Boost type	AdaBoost
AdaBoost β	0.5
N. of trees	850
Max tree depth	3
Min node size	2.5%
nCuts	20

Bibliography

- [1] B. Aubert et al., Leptonic decays of K^0 mesons and PC violation, *Physics Letters* 17, 59 (1965)
- [2] R. R. Akhmetshin et al., Observation of K^0 semileptonic decays with the CMD-2 detector, *Physics Letters B* 456, 145 (1999)
- [3] A. Aloisio et al., Measurement of the branching fraction for the decay $K_S \rightarrow \pi e \nu$, *Physics Letters B* 535 (2002) 37
- [4] J. R. Batley, Determination of the relative decay rate $K_S \rightarrow \pi e \nu / K_L \rightarrow \pi e \nu$, *Physics Letters B* 653, 145 (2007)
- [5] F. Ambrosino et al., Study of the branching ratio and charge asymmetry for the decay $K_S \rightarrow \pi e \nu$ with the KLOE detector, *Physics Letters B* 636 (2006) 173.
- [6] G. D. Rochester and C. C. Butler, Evidence for the Existence of New Unstable Elementary Particles, *Nature* 160 (1947) 855.
- [7] P. K. Kabir, *The CP Puzzle*, Academic Press, 1968.
- [8] KTeV Collaboration, T. Alexopoulos, et al., *Phys. Rev. D* 70 (2004) 092007.
- [9] T.C. Andre, hep-ph/0406006.
- [10] M. Tanabashi et al. (Particle Data Group), Review of Particle Physics, *Phys. Rev. D* 98 (2018) 030001.
- [11] M. Antonelli, et al., An evaluation of $|V_{us}|$ and precise tests of the Standard Model from world data on leptonic and semileptonic kaon decays, *Eur. Phys. J. C* 69 (2010) 399.
- [12] E. Passemar, Semileptonic Kaon Decays, Kaon 2016, University of Birmingham Birmingham, UK, September 14-17 2016
- [13] A. Gallo et al., DAFNE Status Report, *Conf. Proc. C060626* (2006) 604.

- [14] M. Adinolfi et al., The tracking detector of the KLOE experiment, Nucl. Instrum. Meth. A 488 (2002) 51.
- [15] M. Adinolfi et al., The KLOE electromagnetic calorimeter, Nucl. Instrum. Meth. A 482 (2002) 364.
- [16] M. Adinolfi et al., The trigger system of the KLOE experiment, Nucl. Instrum. Meth. A 492 (2002) 134.
- [17] F. Ambrosino et al., Data handling, reconstruction and simulation for the KLOE experiment, Nucl. Instrum. Meth. A 534 (2004) 403.
- [18] A. Aloisio et al., Data acquisition and monitoring for the KLOE detector, Nucl. Instrum. Meth. A 516 (2004) 288.
- [19] R. R. Akhmetshin et al., Reanalysis of hadronic cross-section measurements at CMD-2, Phys. Lett. B 578 (2004) 285.
- [20] F. Ambrosino et al., Measurement of the DAFNE luminosity with the KLOE detector using large angle Bhabha scattering, Eur. Phys. J. C 47 (2006) 589.
- [21] A. Anastasi et al., Measurement of the charge asymmetry for the $K_S \rightarrow \pi e \nu$ decay and test of CPT symmetry with the KLOE detector, JHEP 09 (2018) 021.
- [22] R. Barlow and C. Beeston, Comp. Phys. Comm. 77 (1993) 219.
- [23] A. Nappi, A Pitfall in the use of extended likelihood for fitting fractions of pure samples in mixed samples, Comput. Phys. Commun. 180 (2009) 269.
- [24] C. Gatti, T. Spadaro, KLOE Note 176, 2002, unpublished, <http://www.lnf.infn.it/kloe/pub/knote/kn176.ps.gz>.
- [25] C. Gatti, T. Spadaro, KLOE Note 208, 2006, unpublished, <http://www.lnf.infn.it/kloe/pub/knote/kn208.ps>.
- [26] W.J. Marciano, A. Sirlin, Phys. Rev. Lett. 71 (1993) 3629.
- [27] V. Cirigliano et al. JHEP 11 (2008) 006, arXiv:0807.4507.
- [28] KTeV Collaboration, T. Alexopoulos, et al., Phys. Rev. D 70 (2004)092007.
- [29] [FNAL/MILC 13E] A. Bazavov et al., Determination of $|V_{us}|$ from a lattice-QCD calculation of the $K \rightarrow \pi e \nu$ semileptonic form factor
- [30] J. C. Hardy and I. S. Towner, Phys. Rev. C 91, 025501 (2015).
- [31] K. A. Olive et al. (Particle Data Group Collaboration), Chin. Phys. C 38, 090001 (2014)
- [32] A. Hoecker et al., TMVA - Toolkit for Multivariate Data Analysis, arXiv:physics/0703039.