UNIVERSITÀ
DEGLI STUDI
FIRENZE

Scuola di
Scienze Matematiche,
Fisiche e Naturali

Corso di laurea in
Scienze Fisiche e
Astrofisiche

Ricerca del decadimento raro $K^0_S \to \mu^+\mu^-$ a LHCb

Search for the $K^0_S \to \mu^+\mu^-$ rare decay at LHCb

Candidato: Alessio Borgheresi

Relatore: Giacomo Graziani

Correlatore: Giovanni Passaleva

Anno accademico 2014-2015

Contents

List of Figures

List of Tables

Introduction

The Standard Model (SM) is the theory that describes our current understanding of fundamental particles and their interactions and it gives an explanation to most of the phenomena involving three of the four fundamental forces: electromagnetic, weak and strong.

However there are same experimental evidences that cannot be explained by the SM. The most typical examples of these are the cosmological evidence for Dark Matter and the baryonic asymmetry of the universe. Indeed, although the main ideas of baryogenesis are accommodated within the SM of particle physics, additional mechanisms are required to explain the imbalance between matter and antimatter currently observed. In particular, among the several features that motivate an extension of the SM, the search for new sources of CP violation is of great importance. New Physics (NP) models have been developed to cope with these problems and these models can be tested in different High Energy Physics (HEP) experiments at accelerator. The largest and most important accelerator build to test the SM and NP models is the Large Hadron Collider (LHC), located at CERN, in Geneva (Switzerland). The LHC collides protons head-on at a nominal center-of-mass energy of $\sqrt{s} = 14$ TeV ($\sqrt{s} = 7$ TeV during 2011 and $\sqrt{s} = 8$ TeV during 2012, $\sqrt{s} = 13$ TeV presently), which are used by four main experiments (ALICE, ATLAS, CMS and LHCb) to perform HEP searches to a precision never reached before. The LHCb [1] experiment is specialized in CP-violation measurements and in the search for rare decays with charm and beauty hadrons. Both aims are motivated by searches for NP beyond the SM, since departures from the SM parameters constraints and enhancements on the branching fraction of rare decays dominated by loop diagrams can reveal NP contributions at a higher energy scale than what can be reached in particle production directly.

So far, some of the most remarkable LHCb results come in the area of rare decays. Amongst those, a peculiar role is played by the $B^0_{(s)} \to \mu^+\mu^-$ and $B^0_s \to \mu^+\mu^-$ decays since both the branching ratios are very interesting from a theoretical point of view. The latter was found to be compatible with the SM predictions [2] resulting in useful constraint to NP scenarios.

The flavour-changing neutral-current (FCNC) $K^0_S \to \mu^+\mu^-$ decay is also a rare decay whose search has been proposed as a novel way to profit from the excellent performance of the LHCb detector. While very suppressed in SM (second order weak interaction), such a decay can have a different probability according to several NP models, so that it turns out to be extremely powerful to constrain existing NP scenarios. This channel shares some of the experimental characteristics of $B_s(B^0) \to \mu^+\mu^-$, but with some remarkable differences mainly due to fact that the $B_s(B^0)$ masses are a factor of more than 10 larger than that of K^0_S and they both fly a factor of more than 50 less, which implies very different reconstruction and trigger performances.

An analysis of the $K^0_S \to \mu^+\mu^-$ decay had been previously performed with the 2011 LHCb data set (1 fb^{-1} of integrated luminosity), yielding an upper limit [3] $\mathcal{B}(K^0_S \to \mu^+\mu^-)$

$< 11.2(9.0) \cdot 10^{-9}$ at 95 (90)% confidence level. Before the measurement of LHCb the most sensitive search, realized at the CERN PS in the early 1970's [4], set an upper limit $\mathcal{B}(K_S^0 \to \mu^+\mu^-) < 3.1 \cdot 10^{-7}$ at 90% confidence level. The LHCb results improves the previous value by a factor exceeding 30. The branching fraction of $K_S^0 \to \mu^+\mu^-$ in the SM is expected [5] to be $(5.0 \pm 1.5) \cdot 10^{-12}$, three orders of magnitudes below the present experimental limit. The measurement is limited by the statistics and it the possibility to approach the SM level could be achieved during the upgrade phase, currently planned from 2019 onward, aiming to collect up to 50 fb^{-1} of integrated luminosity. However, there are theoretical models that indicate possible contributions of new physics that could rise the branching fraction to a value of the order of $10^{-10} - 10^{-11}$, motivating the effort of an update of the analysis.

This thesis is devoted to the search for the very rare decays $K_S^0 \to \mu^+\mu^-$, performed using the 2 fb^{-1} of data taken by LHCb during 2012 from the LHC proton-proton collisions at an energy of $\sqrt{s} = 8$ TeV. The sensitivity is expected to be improved, other than from the larger integrated luminosity collected in 2012, by the increase of the trigger efficiency, which is a weak point for this channel. Indeed, in the 2012 data taking it was possible to increase the trigger bandwidth from 2.5 to 5 kHz. This allowed to adapt the software trigger level to select the di-muon candidates of any invariant mass, removing the 1 GeV/c^2 lower threshold that was removing the K_S^0 region during the 2011 run. Some of the important points of the analysis are the understanding of the trigger efficiency and the background contributions and kinematic dependencies associated to the trigger selection. In this work of thesis, I had responsibility in the study of the trigger for the $K_S^0 \to \mu^+\mu^-$ channel, as also in the improving of the background rejection achieved with the development of a Multi-Variate-Analysis (MVA), for which particular attentions are dedicated to find the best classifier algorithm and to chose the variables that allow to achieve the best possible the signal to background discrimination. Eventually, I have studied the optimization of the muon identification requirement needed to minimize the background from $K_S^0 \to \pi^+\pi^-$ decays with double particle misidentification. My work in the $K_S^0 \to \mu^+\mu^-$ analysis dealt also with several technical aspects, such as the ntuple creation and the study of the TMVA toolkit integrated in the ROOT framework. The contributions of my work were presented at two internal meeting of the very rare decay group of LHCb.

This thesis is composed of three main parts. In Chapter 1 it is given a theoretical introduction to the interactions and particles that compound the SM and to the CP violation sector, focusing on the contribution given by the kaon system. It is presented also an explanation of the SM prediction on the $K_S^0 \to \mu^+\mu^-$ decay process.

The second part (Chapter 2) is dedicated to a detailed description of the LHCb experiment at the LHC with a summary on the most important results achieved so far. The performance of LHCb during 2012 is also briefly reported, underlying the differences between the 2011 and 2012 conditions.

The following chapters deal with my personal work on the search for the $K_S^0 \to \mu^+\mu^-$ decay. An overview of the analysis strategy is given in Chapter 3, while the Chapter 4, 5 are dedicated to my main contributions to the analysis work, including the selection of events from the LHCb data, the trigger studies and the multivariate analysis developed to discriminate the combinatorial background. Finally, in Chapter 6 conclusions about the expected sensibility with respect to the 2011 measurement are achieved from a preliminary background expectation following the optimization of the muon identification requirement on the daughters particles. In this last chapter a brief illustration of the following analysis

steps, such as the measure of the efficiencies entering in the normalization factor, is given.

Chapter 1

Theoretical Background

In this chapter a brief description about the elementary particle scenario, studied in High Energy Physics (HEP), is given.

The first section (1.1) is composed by an introduction of the Standard Model (SM) theory, focusing on the electroweak interaction and the quarks generations. The most important composite particles are briefly described in 1.1.2. The CP violation in the SM is presented in section 1.2, with the historical role played by the kaons and the heavy neutral mesons emphasized in 1.2.2 and 1.2.3. Finally, the theoretical prediction of the branching ratio of $K_S^0 \to \mu^+\mu^-$, highlighting the difference with the $K_L^0 \to \mu^+\mu^-$ case, is described in 1.3.

1.1 Introduction to the Standard Model

The Standard Model (SM) is a quantum field theory (QFT) which describes our understanding of the fundamental particles and their interactions.

In the SM theory, fields interactions are based on local gauge invariance[1], i.e. it is possible to apply a set of transformations which keep the SM Lagrangian density unchanged. A gauge transformation with constant parameters at every point in space and time is analogous to a rigid rotation of the geometric coordinate system; it represents a global symmetry of the gauge representation. As in the case of a rigid rotation, this gauge transformation affects expressions that represent the rate of change along a path of some gauge-dependent quantity in the same way as those that represent a truly local quantity. A gauge transformation whose parameters depend on the space-time coordinates is referred to as a local symmetry.

SM has proven to be an excellent model to describe and predict phenomena involving three of the four fundamental forces: electromagnetism, strong and weak interaction. Each of these interactions has its own basic charges: electric charge for the electrodynamics, two

[1] A local gauge symmetry is a group of local transformations of the fields. The local adjective means that the parameters of the transformations depends on the space-time point $x = (t, \overrightarrow{x})$. If the symmetry is conserved, by definition, the Lagrangian is left unchanged by the local transformations. In the simplest form (U(1) gauge symmetry as can be the case for the electromagnetism), the gauge transformation on a scalar field $\phi(x)$ and on a vectorial field $A^\mu(x)$ can be writte as

$$\phi(x) \to e^{i\theta(x)}\phi(x), \quad A^\mu(x) \to A^\mu(x) + ie\partial^\mu(x) \tag{1.1}$$

in such a way to maintain unchanged the Lagrangian $\mathcal{L} = -F^{\mu\nu}F_{\mu\nu} + D^\mu\phi D_\mu\phi$, where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ and $D^\mu = \partial^\mu + ieA^\mu$.

weak charges (weak isospin) for the weak interaction and three colour charges (red, green and blue) for the strong force. Gravity is not included in the SM since its contribution is negligible compared to the other forces at the mass value studied in HEP (up to $\sim$ 1 TeV/c^2).

Both the Quantum Electrodynamics (QED), which describes the charged particle interactions, and the Quantum Chromodynamic (QCD), related to the strong force, are part of the SM. QED and weak interactions have been unified in the Electroweak theory (EW) by Glashow [6], Weinberg [7] and Salam [8].

In the SM the Lagrangian $\mathcal{L}$ has to be local gauge invariant under the group $SU(3)_C \otimes SU(2)_{weak} \otimes U(1)_Y$, with the QCD corresponds to the symmetry group $SU(3)_C$ of color (C), while the EW interaction is described by the group $SU(2)_T \otimes U(1)_Y$ of weak-isospin (T) and hypercharge (Y).

The aforementioned symmetry is broken spontaneously[2] according to the Higgs mechanism to give particles their masses.

Indeed, if the symmetry had been exact the gauge bosons would not have masses, going against the experimental fact that both W and Z are massive particles. The Higgs mechanism predicts a scalar boson field corresponding to a massive scalar boson particle, the Higgs particle, that has been discovered, with a mass of 125 GeV, by the ATLAS and CMS experiments at LHC. On July the 4th, 2012, they announced the discovery of the Higgs boson particle [9] [10], whose existence was predicted by François Englert and Peter W. Higgs, awarded by the 2014 Nobel Prize.

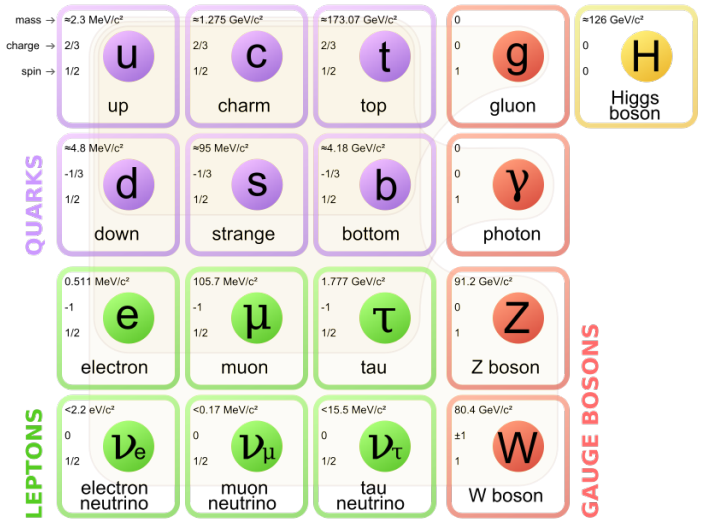


Figure 1.1: Summary of some proprieties (charge, mass, spin) of the elementary particles known so far. The fermions are classified in three generation of increasing mass, the bosons are the mediator of the three fundamental interactions described in the SM. The Higgs boson is responsible for the mechanism to give mass to all the particles.

The Standard Model (SM) involves two classes of fundamental particles: fermions and bosons, whose key properties and interactions are given in Figure 1.1.

[2] The symmetry breaking, in QFT, can appear in two main ways, one called explicitly symmetry breaking and the other spontaneous symmetry breaking. The former is the case of a Lagrangian which is not invariant with respect to the symmetry, while the latter happens when the Lagrangian is invariant with respect to the symmetry but not the vacuum value that is the minimum energy state.

Fermions can be divided in two groups, depending if they are affected by strong interaction (quarks) or not (leptons). In both groups, fermions are grouped into three generations, according to their masses, with four particles in each generation. Quarks have fractional electric charge, while leptons have integer (zero for neutrinos). Each elementary particle is characterized by a mass, m, and a spin, s, which is related to the helicity if the particle mass is null.

The three lepton families are e, μ, τ with the corresponding associated neutrino. While charged leptons interact both electromagnetically and weakly, neutrinos, which are neutral, do not undergo the electromagnetic force and are so characterized only by the weak interaction. Each lepton has its own anti-lepton, which differs for the electric charge, as it has all quantum numbers inverted. It is still debated whether neutrino and anti-neutrimo are the same particle, differentiated only by the helicity, or not.

In its original formulation, the SM included massless neutrinos, so their helicity becomes equivalent to chirality. This means that the right-handed neutrinos in the SM do not interact with anything and, equivalently, for the left-handed anti-neutrinos. It is today clear that at least two of the neutrino eigenstates are massive. Indeed, the process of neutrino oscillations, that consists of the variation of the neutrino flavour during its time evolution, has been observed and confirmed, and it is only possible if the free Hamiltonian eigenstates are a linear combination of the states with defined flavour ν_e, ν_μ, and ν_τ and if their masses are different. The absolute mass value of the neutrinos is still unknown as well as the hierarchy of the mass states because the measure of the oscillations gives only informations about the difference between the square of the masses. The upper limit on the ν_e neutrino mass set with tritium decay studies is 2eV [11].

Quarks are divided in six different flavours: down (d), up (u), strange (s), charm (c), bottom (b), and top (t). The two latter flavours are alternatively named beauty and truth. Quarks are the only particles which interact via strong force, that is responsible for their confinement in bound states called hadrons. Hadrons are divided into mesons and baryons according to the number of quarks and anti-quarks: one quark and one anti-quark for mesons and three quarks for baryons.

In addition to the fermions, the SM also describes a set of bosons which act as force mediators in interaction processes:

- Photon (zero mass and two states of helicity $\lambda = \pm 1$) is the mediator of the electromagnetic interaction.
- $W^\pm$ and Z bosons with mass of 80 GeV/c^2 and 91 GeV/c^2, respectively, and spin 1 are the mediators of the weak force.
- The strong force is mediated by eight vectorial particles with null mass called gluons, which are differentiated by their colour combination.

The mass values for fermion particles are free parameters in the SM.

1.1.1 Electroweak Force

The electroweak interaction is based on the $SU(2)_T \otimes U(1)_Y$ symmetry, thus the Lagrangian has to be invariant under the gauge transformations of the fields (the corresponding transformations are reported some lines below). The first (non abelian $SU(2)_T$ group) is a gauge interaction generated by the weak isospin $\overrightarrow{T}$, while the second (abelian $U(1)_Y$) is a gauge interaction generate by the hypercharge Y.

After the spontaneous broken symmetry due to the Higgs mechanism, only one gauge symmetry remains: $U(1)_{em}$ with the electromagnetic field A_μ as generator.

There is a relation between the third component of the weak isospin, the hypercharge Y and the electric charge of the particle Q, expressed in unit of e (Gell-Mann Nishijima formula): $Q = T_3 + \frac{Y}{2}$.

The quantum numbers for leptons and quarks are summarized in the Table 1.1, where L and R are the left and right chiral projection of the fermion field.

	Family					
	1	**2**	**3**	$\mathbf{T_3}$	**Y/2**	**Q**
Leptons	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	$1/2$	$-1/2$	0
				$-1/2$	$-1/2$	-1
	e_R	μ_R	τ_R	0	-1	-1
Quarks	$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} \mathrm{t} \\ b' \end{pmatrix}_L$	$1/2$	$1/6$	$2/3$
				$-1/2$	$1/6$	$-1/3$
	u_R	c_R	t_R	0	$2/3$	$2/3$
	d_R	s_R	b_R	0	-1/3	-1/3

Table 1.1: The values of the third component of the weak isospin T_3, the hypercharge Y and the electric charge for the lepton and quark families.

The local gauge transformation associated with the $SU(2)_T$ and $U(1)_Y$ are:

$$\begin{pmatrix} \nu \\ l \end{pmatrix}_L^{'} = exp(ig_2 \overrightarrow{T} \cdot \overrightarrow{\beta}(x)) \begin{pmatrix} \nu \\ l \end{pmatrix} \tag{1.2}$$

$$\begin{pmatrix} \nu \\ l \end{pmatrix}_L^{'} = exp(ig_1 \frac{Y}{2} \theta(x)) \begin{pmatrix} \nu \\ l \end{pmatrix} \tag{1.3}$$

$$\nu_R^{'} = exp(ig_1 \frac{Y}{2} \theta(x)) \nu_R \tag{1.4}$$

$$l_R^{'} = exp(ig_1 \frac{Y}{2} \theta(x)) l_R \tag{1.5}$$

where the constants g_1 and g_2 that are characteristics of the electroweak interaction and similar relations hold for quarks (universality of the EW interaction), with the difference that now the autostates of the weak interaction are not the same of the flavor autostates (the doublets are shown in the Table 1.1). This is taken into account with the mixing matrix (described in 1.1.1.1) between the flavor d, s, b.

The $U(1)_Y$ symmetry is generated by one field $B_\mu(x)$ that transforms as $B_\mu^{'} = B_\mu - \partial_\mu \theta(x)$, while for the $SU(2)_T$ there are three fields $\overrightarrow{W}_\mu = (W_\mu^1, W_\mu^2, W_\mu^3)$ that transform as $\overrightarrow{W}_\mu = \overrightarrow{W}_\mu - \partial_\mu \overrightarrow{\beta}(x) + g_2 \overrightarrow{\beta} \wedge \overrightarrow{W_\mu}$.

The invariant Lagrangian for global transformation of $SU(2)_T \otimes U(1)_Y$ can be written as:

$$\mathcal{L} = i \sum_f \bar{\psi}_f \gamma^\mu D_\mu \psi_f - \frac{1}{4} \overrightarrow{W}_{\mu\nu} \cdot \overrightarrow{W}^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \tag{1.6}$$

where $\overrightarrow{W}_{\mu\nu} = \partial_\mu \overrightarrow{W}_\nu - \partial_\nu \overrightarrow{W}_\mu + g_2(\overrightarrow{W}_\mu \wedge \overrightarrow{W}_\nu)$, $B_{\nu\mu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ and D^μ is the covariant derivative of fermion fields: $D_\mu = \partial_\mu + ig_2 \overrightarrow{W}_\mu \cdot \overrightarrow{T} - ig_1 \frac{Y}{2} B_\mu$.

The last two terms are for the free gauge fields while the fermion part can be divided into:

$$\sum_f \bar{\psi}_f \gamma^\mu \partial_\mu \psi_f - g_1 \sum_f Y_f \bar{\psi}_f \gamma^\mu \psi_f B_\mu - g_2 \sum_f i \bar{\psi}_{fL} \gamma^\mu \overrightarrow{T} \cdot \overrightarrow{W}_\mu \psi_{fL} \tag{1.7}$$

where the first term is simply the free kinetic fermionic part.

The physical fields, associated to the gauge bosons, are obtained by the following linear combinations of the generators:

$$W^\pm = \frac{1}{\sqrt{2}}(W_1 \mp iW_2) \tag{1.8}$$

$$A^\mu = \cos\theta_W B^\mu + \sin\theta_W W_3^\mu \tag{1.9}$$

$$Z^\mu = -\sin\theta_W B^\mu + \cos\theta_W W_3^\mu \tag{1.10}$$

with θ_W called the Weinberg angle defined by the following relation:

$$\sin\theta_W = \frac{g_1}{\sqrt{g_1^2 + g_2^2}} \tag{1.11}$$

and measured experimentally $\sin^2\theta_W \simeq 0.231$.

The interactions, implied in the covariant derivative, can (after some calculations not reported here) be expressed in a more clear relation with the particle fields A^μ, Z^μ and $W^\pm$:

EM $-eQ_f \bar{\psi}_f \gamma^\mu \psi_f A_\mu = -eQ_f J^\mu_{em} A_\mu$
where Q_f is the electric charge of the fermion in units of e, meaning that the electromagnetic (EM) interaction is only felt by the charged leptons and not by neutrinos. The EM interaction introduces, in the EM current J_{em}, the left and right fields in the same way.

Weak NC $-\frac{g_2}{2\cos\theta_W} \bar{\psi}_f \gamma^\mu (g_f^V - g_f^A \gamma_5) \psi_f Z_\mu = -\frac{e}{\sin 2\theta_W} J^0_\mu Z^\mu$
with $g_f^V = T_3 + 2q\sin^2\theta_W$, $g_f^A = T_3$.

In this term can be recognized the weak neutral current (J^0_μ) that couple a lepton-antilepton pair without allowing change in the lepton flavor.

Weak CC $-\frac{g_2}{2\sqrt{2}}(J^\mu_{CC} W\,\dagger_\mu + h.c.)$
with $J^\mu_{CC} = \bar{l}\gamma^\mu(\frac{1-\gamma_5}{2})\nu$.

This Lagrangian describes a vertex where the electric charge of the lepton changes, meaning that the mediator bosons ($W^\pm$) have charge themselves. This part is called weak charged current (CC).

There is a link between the θ_W parameter and the masses of the bosons that is $\cos(\theta_W) = M_W/M_Z$.

Concerning the experimental confirmations, the success of the EW theory started with the discovery of the NC. This was made in mid 1973 in the Gargamelle bubble chamber at CERN [12]. The experiment used separate neutrino and anti-neutrino beams, with the

beams overwhelmed by muon-neutrinos, so the task was to demonstrate the occurrence of events without a final muon state.

The W and Z bosons were discovered by the two collaborations UA1 and UA2 working at the $Sp\bar{p}S$ collider [13]. The W^+ decays into $e^+\nu_e$ about 8% of the time, while a more spectacular signal (due to less background contributions) could be obtained from Z's that decays into e^+e^- or $\mu^+\mu^-$. The mass of the W could be estimated from the observed transverse momenta, and was found to be in good agreement with SM predictions. Later the two collaborations detected the Z through its decay into e^+e^- and $\mu^+\mu^-$ [14]. With both the Z and W mass measured, the formula $\cos(\theta_W) = M_W/M_Z$ was tested and found to perfectly agree with the previous measured value for the θ_W angle in neutral-current experiments.

1.1.1.1 V_{CKM} Matrix

The weak interaction allows hadrons to decay to lighter states with different flavours. This peculiar behaviour is associated only to the weak charge current, that is when a W boson is exchanged between quarks.

In 1963, when only the u, d, and s quarks were known, Nicola Cabibbo [15] suggested that in the weak interaction the current associated to the u quark was not due to the d quark only but to a linear combination of the quarks d and s: $d\cos(\theta_C) + s\sin(\theta_C)$. The experimental parameter θ_C is a mixing angle, named Cabibbo angle, and its experimental value is $\sin(\theta_C) \sim 0.22$.

Thanks to this mixing it was possible to explain both the $\Delta S = \pm 1$ and $\Delta S = 0$ proceses with the caveat that if the same quark association held also for the neutral current, it was possible to have proceses of flavour changing neutral currents (FCNC, $\Delta S \neq 0$) with comparable probability of the charged one. This is not confirmed by the experimental facts.

In 1970, Glashow, Iliopoulos, and Maiani predicted the existence of a fourth quark, named charm, to explain the suppression of FCNC through the contribution of this quark to the mixing, which can then be expressed by the matrix:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}$$

In this way the FCNC was forbidden in the neutral weak current interactions and was due only to the quark loop diagrams involving the W, that would also give a vanishing contribution if the masses of the up-type and down-type quarks were degenerate (see Figure 1.2). This effect is called GIM mechanism.

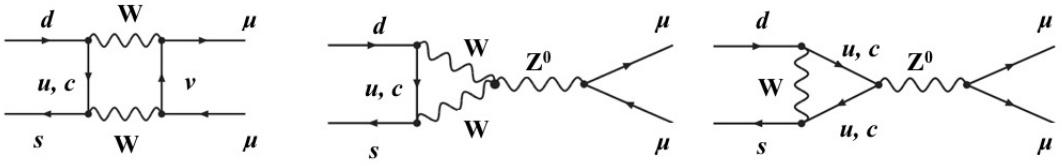


Figure 1.2: An example of box diagrams contributing to FCNC $K^0 \to \mu^+\mu^-$ decay, that would give zero contribution if the masses of the up-type and down-type quarks were degenerate.

After three years, in 1973, Kobayashi and Maskawa, in order to incorporate a proposed third generation of quarks to accommodate the new experimental results of CP violation

[16], expanded the mixing matrix to be a 3x3 unitary matrix.

This matrix (known as the Cabibbo-Kobayashi-Maskawa matrix, V_{CKM}) can be expressed as:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.12}$$

Where the elements V_{ij} represent the amplitude of probability of a transition of a quark i to a quark j.

It can be show that with the unitarity condition of the V_{CKM} and taking into account the arbitrarity in the choice of the phases for the quark fields the only physical parameters left in the matrix are 3 real mixing angles and a complex phase. The complex phase is the one responsible for the CP-violating phenomena in flavour-changing processes in the SM. It has to be pointed out that if any two up-type or down-type quarks were degenerate in mass, they could not be distinguished, and an arbitrary rotation among them two would still be allowed and the parameters of this rotation could be used to remove the single phase in V_{CKM}. Also if one family would be completely decoupled from the other two (zero mixing angle), the V_{CKM} would become block-diagonal, with a 2x2 unitary sub-matrix that does not support CP violation.

About the representation of the V_{CKM}, many parametrizations exist. Of course the physics results do not depend on the choice of parametrization.

An useful parametrization (named Wolfenstein parametrization) is based on the very hierarchical characteristics of the V_{CKM}, which are:

- It is almost diagonal, so the decay channels inside the same quark generations are preferred
- The strength of transition between generations follows this order: first to second, second to third, first to third
- The CP violation effect is small and is more effective in the first to third transition.

Starting from the definitions of $s_{12} \equiv \lambda \equiv \frac{|V_{us}|}{\sqrt{|V_{ud}|^2+|V_{us}|^2}}$, $s_{23} \equiv A\lambda^2 \equiv \lambda \left|\frac{V_{cb}}{V_{us}}\right|$ and $s_{13} \equiv A\lambda^3(\rho + i\eta) \equiv \frac{A\lambda^3(\bar{\rho}+i\bar{\eta})\sqrt{1-A^2\lambda^4}}{\sqrt{1+\lambda^2[1-A^2\lambda^4(\bar{\rho}+i\bar{\eta})]}}$, that ensure the relation $\bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}$, the V_{CKM} can be expressed as an expansion in term of λ (the unitarity is obtained at all orders in λ):

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} - \frac{\lambda^4}{8} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda - A^2\lambda^5(\rho + i\eta - \frac{1}{2}) & 1 - \frac{\lambda^2}{2} - (\frac{1}{8} + \frac{A}{2})\lambda^4 & A\lambda^2 \\ A\lambda^3[1 - (\rho + i\eta)(1 - \frac{\lambda^2}{2})] & -A\lambda^2 - A\lambda^4(\rho + i\eta - \frac{1}{2}) & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix} + O(\lambda^6)$$

The condition $\eta \neq 0$ is required to generate CP violation. The experimental values of the 4 parameters used in the Wolfstein parameterizations are $\lambda = 0.22537 \pm 0.00061$, $A = 0.814^{+0.023}_{-0.024}$, $\bar{\rho} = 0.117 \pm 0.021$, $\bar{\eta} = 0.353 \pm 0.013$.

Due to unitarity constrain ($VV^\dagger = I$) the matrix elements of 1.12 are subject to 12, not independent, relations:

$$\sum_i V_{ij}V_{ik}^* = \delta_{jk}, \quad \sum_j V_{ij}V_{kj}^* = \delta_{ik}, \tag{1.13}$$

where i, j assume the values 1, 2, 3 and correspond to quarks u, c, t and d, s, b respectively.

The six vanishing combinations, when represented in plex plane, are triangles whose sides are the three contributions concatenated. Because none of the angles is so small to become difficult to measure, the db triangle is particularly relevant. This is defined by the equation:

$$\underbrace{V_{ud}V_{ub}^*}_{A\lambda^3(\rho+i\eta)} + \underbrace{V_{cd}V_{cb}^*}_{-A\lambda^3} + \underbrace{V_{td}V_{tb}^*}_{A\lambda^3(1-\rho-i\eta)} = 0, \tag{1.14}$$

where the value of the members are written using the third order Wolfenstein parametrization to obtain an easy comparison with what reported in Figure 1.3. By dividing each side by the best-known one ($V_{cd}V_{cb}^*$), its vertices are exactly (0,0), (1,0) and ($\bar{\rho}$,$\bar{\eta}$). The internal angles of the triangle would parametrize the mixing and the area would be a measure of CP violation in the quark sector[3]. This triangle is called "the"unitary triangle (as it represents the unitary of the CKM matrix, see Figure 1.3).

In principle, the squares of the various matrix elements can be measured by observing a variety of weak decays and η can be determined also by the modulus of the elements of the V_{CKM} that completely defines it. E.g. if A is known, the knowledge of $|V_{ub}|$ and $|V_{td}|$ gives ρ and η. The experimental values of the modulus are:

$$V_{CKM} = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} =$$

$$\begin{pmatrix} 0.97427 \pm 0.00014 & 0.22536 \pm 0.00061 & 0.00355 \pm 0.00015 \\ 0.22522 \pm 0.00061 & 0.97343 \pm 0.00015 & 0.0414 \pm 0.0012 \\ 0.00886^{+0.00033}_{-0.00032} & 0.0405^{+0.0011}_{-0.0012} & 0.99914 \pm 0.00005 \end{pmatrix}$$

Standard experimental measurements used to extract the modulus of CKM matrix elements are:

- $|V_{ud}|$ can be obtained from nuclear β decays with $J^P = O^+ \to O^+$ (Fermi type)
- $|V_{us}|$ (the Cabibbo angle) is determined from semi-leptonic K decays into electrons $K^+ \to \pi^0 e^+ \nu$ and $K_L^0 \to \pi^\pm e^\mp \nu$
- $|V_{ub}|$ can be obtained from the semi-leptonic decays of B mesons
- $|V_{cd}|$ is determined from the inclusive production of charm in neutrino deep inelastic scattering on nucleons, $\nu d \to c\mu^-$ followed by $c \to d\mu^+\bar{\nu}$
- $|V_{cs}|$ semi-leptonic decays of D mesons into strange particles and from charm decays of W bosons

[3] It is because of the non-null value of the complex phase that the area of the unitarity triangles is not zero. As a consequence of the fact that there is a single source of CP violation the area is the same for all the triangles and equal to $J_{CP}/2$ where

$$J_{CP} \equiv |Im(V_{ij}V_{kl}V_{il}^*V_{kj}^*)| = s_{12}s_{13}s_{23}c_{12}c_{13}^2c_{23}\sin(\Delta_{CP}) \simeq \lambda^6 A^2 \eta \tag{1.15}$$

is defined as the only CP violation term in the CKM matrix. Its measured value is $J_{CP} = (3.08 \pm 0.17) \cdot 10^{-5}$. Notice that c_{ij} stands for $\cos(\theta_{ij})$ and s_{ij} for $\sin(\theta_{ij})$.

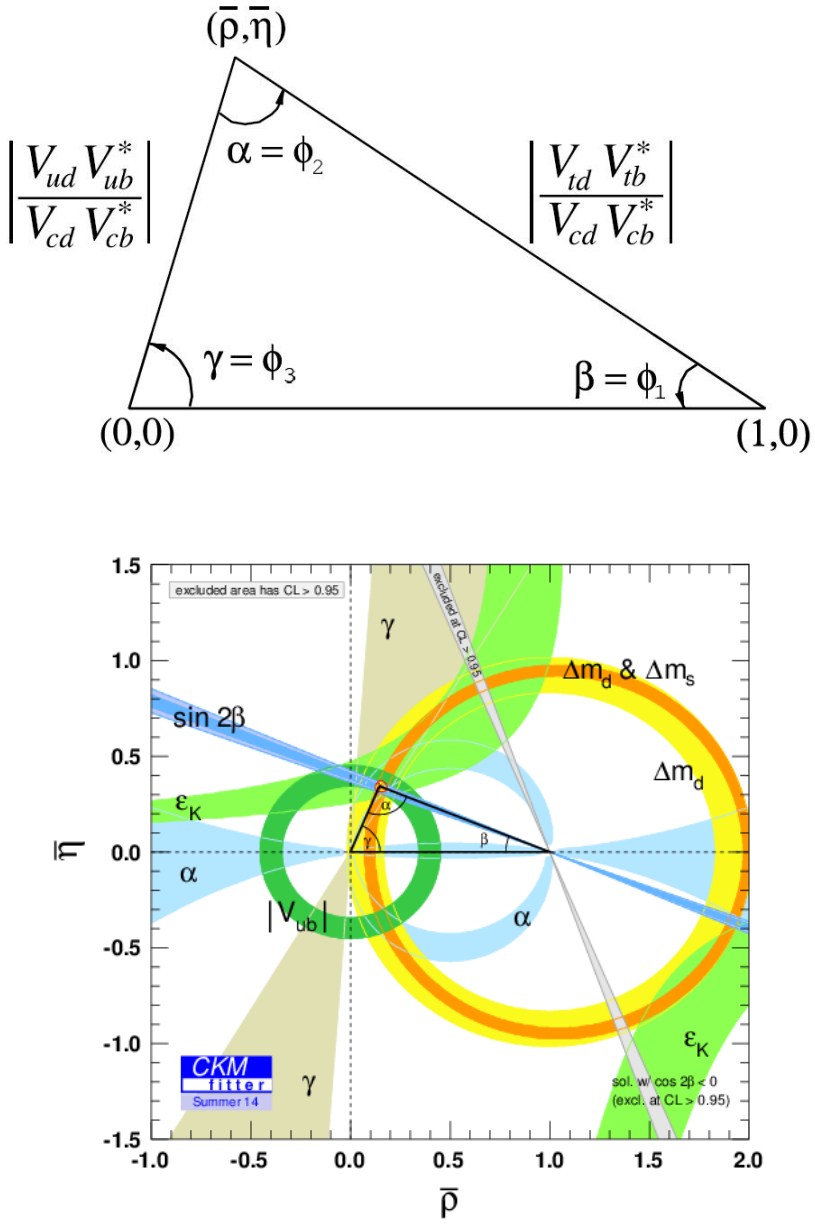


Figure 1.3: At the top: the unitary triangular as defined in the text. In the bottom: unitarity triangular constrain from various experimental results. To date no discrepancies with the SM constrains have been found.

- $|V_{cb}|$ is extracted from exclusive and inclusive semi-leptonic decays of B mesons
- $|V_{td}|$ and $|V_{ts}|$ can be obtained from processes in which the t quark is virtual, such as the longitudinal muon polarization in $K^+ \to \pi^+\mu^+\mu^-$ decays, or other rare K decays
- $|Vtb|$ is roughly constrained by limits on top decays into particles not containing a b quark.

The three angles, showed in Figure 1.3 at the vertices 0, 1, and $\rho + i\eta$, are traditionally called α (also known as φ_2), β (φ_1) and γ (φ_3). These angles are defined as

$$\alpha = \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right], \quad \beta = \arg\left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right], \quad \gamma = \arg\left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right] \tag{1.16}$$

The measure of these parameters, that are forced to stay in the triangular relation within the SM, is carried on to look for NP contribution[4]. One of the key goals of LHCb is to improve the measurement of the angle γ.

For a qualitative discussion of CP violation in B-meson systems, V_{CKM} at λ^3 order is sufficient and higher orders are usually ignored, while for CP violation in $K^0 - \bar{K}^0$ oscillations, the higher orders corrections to V_{cd} are important. For B-meson system, the correction to V_{td} and V_{ts} becomes relevant once the sensitivity of experiments to measure CP-violation parameters becomes 10^{-2} or less.

The extraordinary consistency of several independent measurements of the unitary triangular angles (see Figure 1.3) is one of the great triumph of the Kobayashi-Maskawa theory, which is able to explain all the CP violation observed so far.

1.1.2 Quark Model

The variety of the hadrons is huge due to the number of flavors and the complexity of QCD.

A quantum number, named baryon number, is defined to be $\frac{1}{3}$ for quarks and $-\frac{1}{3}$ for anti-quarks, and it is found to be conserved in all experimental results observed so far. Using its definition, the baryon number for mesons is zero, 1 for baryons and -1 for anti-baryons. The additive quantum number of the quarks, called flavors, are the strong isospin(I_3), the strangeness (S), the charm (C), the topness (T) and the bottomness (B) and are related to the charge Q (in units of e) through the generalized Gell-Mann-Nishijima formula

$$Q = I_3 + \frac{\mathcal{B} + S + C + B + T}{2}, \tag{1.17}$$

where $\mathcal{B}$ is the baryon number. The convention is that the flavor of a quark (I_3, S, C, B, T) has the same sign as its charge Q with the isospin that is the only semiinteger quantiy and it deals with the up and down quark only. The hypercharge is defined as Y = $\mathcal{B}$ + S - $\frac{C-B+T}{3}$ thus Y is equal to 1/3 for the u and d quarks, -2/3 for the s quark and 0 for all the other quarks. It has to be pointed out that there are two different isospins in the SM: the weak isospin T that is responsible for the $SU_T(2)_{weak}$ group symmetry and the strong isospin I that is a consequence of the almost mass degeneracy for the u-d quarks, resulting in an approximate $SU(2)_I$ symmetry.

When including the s quark, the nine possible $q\bar{q}$ combinations containing the light u, d and s quarks are grouped into an octet and a singlet of light quark mesons following

[4] E.g. if the angles extracted by the two different methods disagree, this would indicate NP.

an approximate $SU(3)$ symmetry ($3 \otimes \bar{3} = 8 \oplus 1$). A highlight on the meson physics is reported in section 1.1.2.1.

The $SU(3)$ symmetry is not exact (the quarks mass are not equal) so the mass of the particles in each invariant space are not exactly the same.

Since the baryons are fermions, the state function must be antisymmetric under the interchange of any two equal-mass quarks (up and down quarks in the limit of isospin symmetry and also the s quark if $SU(3)$ flavour symmetry is considered). Thus it can be written as $|qqq\rangle_{\mathrm{A}} = |color\rangle_{\mathrm{A}} \otimes |space, spin, flavor\rangle_{\mathrm{S}}$, where the subscripts S and A indicate symmetric and antisymmetric under interchange of any two equal-mass quarks.

Using only the quarks in the first generation, it is possible to describe the proton (uud) and the neutron (udd). The approximate flavor $SU(3)$ symmetry requires that baryons made of the u, d, s quarks belong to the symmetric multiplets of $3 \otimes 3 \otimes 3 = 10_{\mathrm{S}} \oplus 8_{\mathrm{M}} \oplus 8_{\mathrm{M}} \oplus 1_{\mathrm{A}}$. Here the subscripts indicate symmetric, mixed-symmetric or antisymmetric states under interchange of any two quarks. This is confirmed experimentally. Indeed, the light baryons can be organized in an octet and a decuplet as predicted from the decomposition above. The intrinsic parity of quarks, by convention, is positive while is negative for anti-quark. About the baryons, since they are constituted by three particles with spin $\frac{1}{2}$ they can only have semi-integer spin value. The fundamental states (L = 0) have S = $\frac{1}{2}$ (organized in the octet) or $\frac{3}{2}$ (organized in the decuplet).

When the baryon decuplet was proposed only 9 of the 10 particles that fill the spot were found with the predicted proprieties. The remaining particle had to be characterized by S =-3, $I = 0$ and $J^P = \frac{3}{2}^+$ and should have had a mass of about 1680 MeV/c^2. The discovery of this last particles, named Ω^- [17], was a strong triumph of the quark model and established the eightfold way as the way to describe all observed hadronic particles at that time.

With the introduction of the c and b quarks, new composite particles were found and catalogized [18]. The name of these particles is the same of that of the strange hadrons with the addition of a pedix indicating which quark is substituted to an s-quark, e.g. Λ_b is composed by udb quarks and Ω_b by ssb quarks. This notation emphasizes the fact that when one b or c quark is present, the composite particle is very unstable and hard to be produced and detected. For the mesons a similar approach is used as they are named after the heaviest quark inside them, so mesons containing a strange, charm or bottom quark are named K, D and B respectively. If the lighter quark are not the u or d, the meson's name is augmented so that a meson formed of a bottom and a strange quark would be named a strange B meson (B_s).

Another supporting evidence of the quark model has been found at the end of the '60s by R.E. Taylor at SLAC in deep-inelastic scattering experiments. The cross section of high energy electrons (25 GeV) on protons as a function of the transferred momentum was measured. These experiments showed that proton could not be described as a continuous charge distribution. On the contrary he had to assume that inside proton there were point constituents, called partons and only later identified as quarks and gluons, the real scattering centers. In general, a parton distribution function (PDF), $f_i(x)$, states the probability density that a parton of the i type has an hadron momentum fraction x. The PDF is experimentally determined at certain energies, while its trend as a function of the energy can be calculated through appropriate evolution equations. To predict a cross section value at a proton-proton collider as LHC it is important to know the PDFs. In Figure 1.4 some PDF trends are shown, for the valence quarks, u and d, peaks appear at $x \sim 0.2$, while

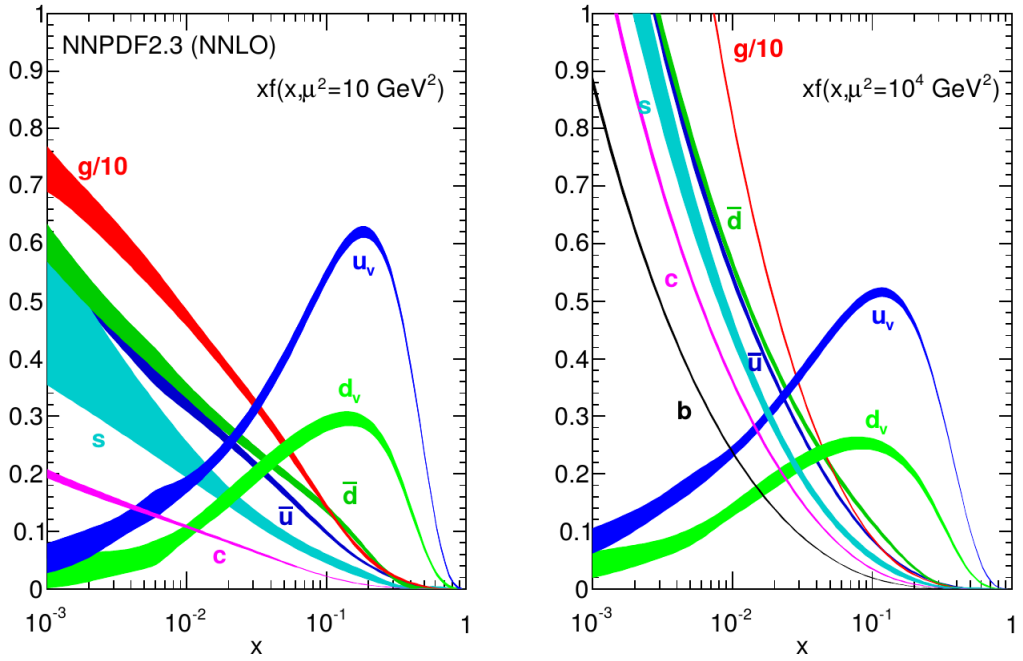


Figure 1.4: The bands show the unpolarized parton distribution functions $f(x)$ multiplied by the x variable (where $f = u_v$, d_v, $\bar{u}$, $\bar{d}$, $s \sim \bar{s}$, $c = \bar{c}$, $b = \bar{b}$, g for gluons) obtained from PDG [18].

for the other particles the probability decreases for large x. Integrals over the structure functions, called sum rules, gave additional confidence in the quark model. As because the structure functions represent the sums over the various probabilities of an electron encountering each kind of parton (multiplied by the square of its charge), specific parton models give definite and different predictions for these sum rules. Fractional charges are favored by the data, but certain sum rules still came in about a factor of 2 smaller than was expected based on a simple three-quark model of the proton (uud). In addition, sum rules evaluated using neutrino structure functions confirmed that half the nucleon's momentum is carried by charged partons leaving the remain half to neutral partons. These neutral partons gave a first suggestion about the existence of the gluons, force mediators of the strong interaction.

1.1.2.1 More about Mesons

A large variety of mesons is known but no stable meson has been observed. The mesons are composed by two spin 1/2 fermions, so they have an integer spin and if orbital angular momentum of the $q\bar{q}$ state is l, then the parity P is $(-1)^{l+1}$. The lighter mesons have zero angular momentum and so $J^P = 0^+$ (pseudoscalar) or 1^- (pseudovector).

The main difference between the two possibilities is due to higher mass of pseudovector particles. Indeed, this allows the decay to the lighter mesons via strong interaction giving them a short lifetime compared to the pseudoscalar mesons that can only decay via weak interaction or electromagnetic in the case of $q\bar{q}$ flavour-less combination like π^0, η.

A short description of the discoveries of some mesons is given here, focusing on the charmed and bottomed systems and following [19].

The strangeness was introduced to explain the fact that certain particles, such as the

kaons or certain hyperons, were created easily (and in pair) in particle collisions, but decayed much more slowly than expected for their large masses and large production cross sections. It was postulated that a new conserved quantity, the "*strangeness*", which was preserved during their creation (strong interactions), but not conserved in their decay (weak interaction). Consequently, the lightest particles containing a strange quark cannot decay by the strong interaction and must instead decay via the much slower weak interaction. The same story was repeated for the other flavours, like the charm and the bottomness.

In November 1974, Burton Richeter at SLAC and Samuel Ting at Brookhaven were leading two different experiments, one studying e^+e^- annihilation and the other the e^+e^- pairs produced in proton-beryllium collisions. The experiments simultaneously discovered a new resonance with a mass of 3.1 GeV/c^2 [20] [21]. The resonance was called J by Ting and ψ by Richter, today it is mentioned as J/ψ. Shortly after, the confirmation of the discovery was given at Frascati [22].

The resonance was found to be extremely narrow, with width completely dominated by the energy resolution of the experiments (today it is known that the value is 87 KeV to be compared for example with the mass and width value of the ϕ: 1020 MeV and 4 MeV).

Prior to the announcement of the J/ψ, there were theoretically studies on the possibility of binding a charmed and an anticharmed quark (the charm quark was supposed to be the quark predicted by the GIM mechanism). The QCD predicted that there would be a series of bound states with very small widths analogous to e^+e^- bound states. Ten day after the discovery of J/ψ, a second narrow resonance (named ψ') was found [23]. The $c\bar{c}$ bound states immediately became the leading explanation of J/ψ and this interpretation was strenghtened by the discovery of the ψ', giving also a confirmation of the GIM mechanism.

The $c\bar{c}$ interpretation also explains the narrowness of the widths because both the J/ψ and ψ' are too light to decay into 2 charm mesons and the lower allowed QCD level is at 3 gluons exchange, while for the ϕ one gluon exchange is possible and the strange mesons are kinetically accessible too.

If the J/ψ was a bound state of a charmed quark and charmed anti-quark, there would have to be mesons with the composition $c\bar{u}$ or $c\bar{d}$ etc... The same situation previously found for the kaon family should reappear here: the new mesons are stable against strong decays but can decay via weak interaction into particles containing a s or d quark in place of a c quark (and the same for anti-quark), with the first favorite by the quark hierarchy. The first mesons seen were the D^0 and D^+ by the SLAC-LBL Mark I group [24]. Their masses were large enough to forbid the decay of the ψ' into a $D\bar{D}$ pair and these particles come in doublets with a decay width less than 2 MeV, indicating the existence of a quantum number conserved only in strong and electromagnetic interaction. In addition the decays were shown to violate parity.

The first $c\bar{c}$ resonance above the D meson threshold is the $\psi(3772)$, and in fact it decays entirely to $D\bar{D}$. The $\psi(3772)$ is thus a D-meson factory and has been the base to study the charmed mesons. To complete the lower D meson possible states the last possible combination of $c\bar{q}$ is the $c\bar{s}$, which is associated by the meson called D_s^+ [25].

The same story happened for the J/ψ and ψ' was replicated in mid 1977, when a significant $\mu^+\mu^-$ peak was observed at the 9.5 GeV/c^2 by Leon Lederman and his coworkers [26]. A more detailed analysis showed better agreement with two peaks at 9.44 GeV/c^2 and 10.17 GeV/c^2 which were given the names Υ and Υ' and were interpreted as a bond system of $b\bar{b}$, thus giving evidence of the b quark. The discovery was confirmed in may 1978 by two groups at the DORIS e^+e^- storage ring at DESY in the PLUTO and

DASP II detectors [27].

These states are very narrow and the first $B\bar{B}$ broader state is the Υ''' at 10.577 GeV/c^2 that is interpreted as the 4^3S_1 state, lying above the threshold for $B\bar{B}$ production, where B is a meson containing the quark $\bar{b}$ and u or d. By focusing on the Υ''' (or $\Upsilon(4S)$) it was possible to obtain a good sample of $B\bar{B}$ events [28]. The CKM picture predicts that the dominant B meson decays are due to $b \to c\bar{u}d$ and $b \to c\bar{c}s$ so that the B meson will most probably produce a D meson. Thus, these decays had proven to be an excellent opportunity to test the models proposed to explain the flavour physics.

Because of the coupling of b to lighter quarks is very small, the lifetime of B hadrons is longer ($\tau \sim 1.5 \cdot 10^{-12}$s that means a flight distance, if the meson can be considered relativistic, of $c\tau \sim 450$ μm) than what could be expected from the decay phase space, if the coupling was comparable to that of lighter quarks. This has important experimental consequence: the B meson can be identified with a precise measure of the flight distance, achieved in the modern system of detection by dedicated vertex detectors.

1.1.3 The Strong Force

In this subsection the peculiar characteristics of the strong interaction are reported without going into the full mathematics formalism, following [29].

The strong force allows the quarks to form composite particles and thus all the known matter. A problem of the quark model was that it seemed to violate the principle of Pauli about identical fermions when the model was used to interpret the Δ^{++} baryon as a combination of 3 u quarks with a symmetric spin-spatial state (spin $= \frac{3}{2}$ and L $= 0$). As quark theory evolved, it became clear that quarks had an extra quantum number, named colour, which affected how quarks interact with each other via the strong force. In fact, in 1972, the problem was fully solved by Bardeen, Fritzsch and Gell-Mann that concluded that the quarks have tree different states (color states called red, blue and green). The new SU(3) symmetry introduced was considered exact and the compound systems of quarks could only appear in an colorless combination, that means in a SU(3) singlet.

The theory describing the strong force is called Quantum Chromodynamic (QCD) and was formuled by Fritzsch [30], Gross and Wilczec [31], Weinberg [32] in 1973 in an attempt to describe the phenomenology of hadronic interactions. The gauge bosons of QCD are massless particles of spin 1, called gluons, and have eight possible color states. Differently from the weak interaction the strong force conserves the quark flavor but changes their color and only quarks are subjected to it.

The non-abelian nature of the group of symmetry of the strong and weak interaction ($SU(3)$ and $SU(2)$) allows the particles that mediate the interactions to bring charge and interact between themselves. This is the main reason that makes the QCD very different from the QED. Due to these auto interactions of the gluons, the strength of the force exerted in the interaction becomes small at high energies (or small distances). This effect is known as asymptotic freedom. Indeed, going at short distance (less than 0.1 fm), the strong force becomes very weak, the quarks can be considered as almost free with the dominant interaction due to a single gluon exchange. This is the situation when the transfered four-momentum (an useful quantity is Q^2, defined as $Q^2 = q^2 - E_q^2$ where $Q = (E_q, \overrightarrow{q})$ is the transfered four-momentum) between two quarks becomes very high allowing to use a perturbative approach for the theory which is not possible when the strength of the interaction is comparable with the kinetic term (long distance between quarks).

When the distance between two coloured particles increases, it is energetically favorable

to produce a new quark-antiquark pair out of the vacuum, thus reducing the separation between the quark and anti-quark. In this way new pairs of quark-antiquark or gluons are generated and can aggregate to create hadrons and mesons: this process is called hadronization. This fact implies that quarks are confined inside physical hadrons, which are always color singlets and means that only colourless composite particles can exist freely. This property, known as confinement, is a peculiarity of the strong interaction.

Thus, a single quark or single gluon, that cannot be a color singlet, should not occur as a physical particle, but can be observed looking at the jets of hadrons moving nearly along the directions of the quarks, because in a very high energy collision the hadrons would lie nearly along this single axis, with momenta transverse to it of a few hundred MeV/c.

Evidence for jets arising from quarks coming from e^+e^- collider was first obtained using the SLAC-LBL Mark I detector at the SPEAR storage ring located at SLAC [33].

The detection of gluons in e^+e^- collisions can be made when the final state is $q\bar{q}g$ and when the gluon and the quark are not moving in nearly the same direction, because it becomes difficult to discern the presence of the gluon. To differentiate a quark jet between a gluon jet the difference in the angular dependence due to the different spin (1 for gluon and $\frac{1}{2}$ for quark) can be used. The identification of $q\bar{q}g$ state, and so the detection of a gluon jet, was achieved first at PETRA e^+e^- collider located at DESY, which was able to reach more than 30 GeV total center-of-mass energy [34].

1.2 CP Violation

In Quantum Field Theory (QFT) it is usual to distinguish between continuous and discrete symmetries. On the one hand continuous symmetries imply operators which act on continuous variables, e.g. time translation, position translations and rotations. On the other hand discrete symmetries concern a finite and numerable set of states. Relatively to the latter type of symmetries, the Noether's theorem states that any continuous symmetry corresponds to a physical quantity conserved in the system evolution. For example, the translational space-time invariance gives rise to the quadri-momentum conservation, whereas the rotational invariance leads to the angular momentum conservation.

The main examples of discrete symmetries are C, P and T, named the charge conjugation, parity and time-reversal symmetries, respectively. The P symmetry changes the sign of the three space coordinates $x_i \to -x_i$ where $i = 1, 2, 3$. The charge conjugation turns a particle into its own antiparticle (or vice versa) as it inverts the sign of the electric charge and all the internal quantum numbers of the particle. Both C and P symmetries are conserved in the strong and electromagnetic interactions but not in the weak interaction [35]. Finally, T is an anti-unitary operator which makes the system evolve using the parameter -t instead of t.

The combination of the charge conjugation and parity transformations results in the CP operator. There would be no difference in behavior between a particle and its antiparticle, if a process was invariant under the C and P transformations individually. A symmetry violation means that the physical process is not invariant under the process obtained from the application of the symmetry transformation.

Even if CP is violated in some physical processes, the CPT symmetry is still considered to be an exact symmetry as no CPT violation has been observed experimentally yet[5]. Indeed,

[5]For example measurements of the masses of neutral K mesons have shown that $|m_{K^0} - m_{\bar{K}^0}| < 5.1 \cdot 10^{-19}$ GeV/c^2 at a 90% confidence level [36].

the CPT symmetry constitutes a fundamental property of any Quantum Field Theory[6] requiring, among other things, that the particle mass and lifetime must be the same as its antiparticle.

At first, the P symmetry was considered conserved, but in 1956 T.D. Lee and C.N. Yang asserted that none of experiments realized until then had proven the P and C invariance in weak interactions [37]. They also suggested some possible experiments to test P and C symmetry conservation in weak interactions and in the following year two groups found experimental evidences for parity violation in the weak interaction (and that intensity of the violation was maximum). The violation of P symmetry was observed in 1957 both by M.me Wu et al. [38] in the β decays of ^{60}Co nuclei and by Garwin and Lederman in sequential π^+ and μ^+ decays [39].

In 1964 Cronin and Fitch [16] observed in the neutral kaons system that not only the weak interaction violates the two symmetries individually but also their combination, CP symmetry, was slightly violated. More details about the CP violation are given in section 1.2.2, following substantially [40].

In the SM, the only source for CP violation is the phase of the CKM matrix. The experimental studies on CP violation are still an important area of research, since testing the CKM paradigm for quark mixing, allows to set stringent constraints on NP scenarios.

In the early universe it is assumed that matter and anti-matter were created in equal amounts by processes like boson $\rightarrow$ matter + anti-matter and destroyed in equal amounts as matter + anti-matter $\rightarrow$ boson, which are perfectly symmetric processes. The imbalance of matter and anti-matter observed in the present universe, with baryons predominant over anti-baryons, raises the question about asymmetric behavior in the matter and anti-matter evolution or about the early universe nature. The CP violation in the SM is not enough to account for the matter asymmetry[7]. In fact, the level of CP violation required in the early universe to observe the asymmetry in the present universe is several orders of magnitude bigger than that observed in the quark flavour sector according to the CKM picture.

Theories beyond the SM such as the Minimal Supersymmetric Standard Model [42] provide a promising framework for weak baryogenesis. Still remaining within the SM, contribution to CP violation can be found in neutrino mixing [43]. Indeed, to explain the neutrino flavour oscillations, a lepton mixing matrix called Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, which describes the relation of the physical (mass) eigenstates to the flavour eigenstates, can be defined. As for quarks, complex matrix elements in the matrix could induce CP violation effects. These effects can be experimentally measured by comparing the probability of a defined flavour neutrino state to be detected, after propagation in vacuum, as a different flavour, with the same probability for an antineutrino. The difference

$$P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta) \neq P(\nu_\alpha \rightarrow \nu_\beta) \tag{1.18}$$

indicates CP violation. This could be performed using future intense muon storage facilities (neutrino factories), which would provide very large fluxes of neutrinos of the type ν_μ, $\bar{\nu}_e$ or $\bar{\nu}_\mu$, ν_e.

[6] The CPT theorem says that CPT symmetry holds for all physical phenomena, or more precisely, that any Lorentz invariant local quantum field theory with a Hermitian Hamiltonian must have CPT symmetry.

[7] In 1967, Andrei Sakharov came up with three necessary conditions for the universe to have been able to create matter and anti-matter at different rates [41]. One of these so-called Sakharov conditions is that charge conjugation symmetry and CP symmetry must be violated.

1.2.1 General Aspects

In general [44], CP symmetry holds if the Hamiltonian is invariant under the CP transformation that is when

$$H_{CP} \equiv (CP)^{\dagger} H (CP) = H \text{ or } [H, CP] = 0. \tag{1.19}$$

Assuming CPT, the CP violation can be described by complex terms appearing in the Lagrangian density, which correspond to complex coefficients a_i for the operators O_i that compound the Lagrangian

$$\mathcal{L} = \sum_i a_i O_i + h.c. \tag{1.20}$$

where the operators are such that $(CP)^{\dagger} O_i (CP) = O_i^{\dagger}$. The only important complex coefficients for CP violation are those that cannot be make real by just redefining the unphysical phase of the fields that enter inside the operators. Complex parameters in the Lagrangian density correspond to phase factors in transition amplitudes that have opposite values for two CP-conjugated processes. In the SM there is only one source of CP violation due to the weak interaction (as explained in section 1.1.1.1), for this reason the appearing phase in the amplitudes that change sign under CP is often called weak phase. Even in absence of CP violation, due to absorptive parts in amplitudes, complex phase can appear in the transition amplitude and usually it is associated to the strong interactions and is called strong phase or scattering phase. This kind of phase has the same sign for CP-conjugated processes since the strong force conserves CP.

Because only phase differences between amplitudes can be observed, at least two interfering amplitudes, A_1-A_2 with associated weak phases ϕ_1-ϕ_2 and strong phases δ_1-δ_2, are required to observe a phase-related effect and so to study CP symmetry in a given process. To exalt the CP violation term in a process with the initial state i to the final state f ($\bar{i} = CPi$ and $\bar{f} = CPf$), the two CP-conjugate amplitudes

$$A(i \to f) = |A_1| e^{i\phi_1} e^{i\delta_1} + |A_2| e^{i\phi_2} e^{i\delta_2} \tag{1.21}$$

and

$$A(\bar{i} \to \bar{f}) = |A_1| e^{-i\phi_1} e^{i\delta_1} + |A_2| e^{-i\phi_2} e^{i\delta_2} \tag{1.22}$$

are separately observed (their squared modules) and then the transition rates difference is calculated, that is

$$|A(\bar{i} \to \bar{f})|^2 - |A(i \to f)|^2 = 2|A_1||A_2| \sin(\phi_1 - \phi_2) \sin(\delta_1 - \delta_2). \tag{1.23}$$

The difference in transition rates vanishes if the two interfering amplitudes have the same phase (strong or weak). The presence of the scattering phases is a necessary condition to observe CP violation but it can introduce some difficulties, since being induced by strong interaction effects that cannot be reliably computed, it is usually difficult to extract information on the weak phases from the measurement of CP asymmetries. Because the asymmetry in transition rates is proportional to the product of two interfering amplitudes, it is only sizeable when such amplitudes have comparable magnitude.

CP violation can be studied looking, for example, to the following signatures:

- Decay selection rules: that is when a CP eigenstate decays into a final state with a different CP eigenvalue. This is how CP violation was discovered, with the K_L^0 state decaying both into 2π and 3π states, as described in section 1.2.2.

- Partial decay widths: CP symmetry requires the equality of any partial decay width for particles and antiparticles. Differences in CP conjugate partial decay widths are a clear sign of CP violation and since CPT symmetry enforces the equality of the lifetimes for CP-conjugate states, a comparison of the branching ratios is sufficient for the test.

- Interference effects: a non exponential decay for a final CP eigenstate observed for an initial meson state given by the coherent superposition of physical eigenstates with different CP is a signal of CP violation.

Regarding the flavored neutral mesons K^0, D^0, B^0, B_s^0, because the full Hamiltonian does not commute with flavour due to weak interaction, the flavour eigenstates $|M^0\rangle$ and $|\bar{M}^0\rangle$ are not expected to be the physical states (eigenstates of the full Hamiltonian) with defined lifetimes. Indeed, the physical states $|M_{a,b}\rangle$ time evolution is

$$|M_{a,b}(t)\rangle = e^{im_{a,b}t}e^{\Gamma_{a,b}t/2}|M_{a,b}(0)\rangle \tag{1.24}$$

with masses $m_{a,b}$ and total decay widths $\Gamma_{a,b}$ and can be related in full generality to the flavour eigenstates

$$|M_a\rangle = p|M^0\rangle + q|\bar{M}^0\rangle \quad |M_b\rangle = p|M^0\rangle + q|\bar{M}^0\rangle \tag{1.25}$$

where p and q are parameters fixed by experiment, satisfying the condition $|p|^2 + |q|^2 = 1$. If CP symmetry is valid

$$\left|\frac{q}{p}\right| = 1. \tag{1.26}$$

The kind of CP violation discovered in 1964 with the $K_L^0 \to \pi^+\pi^-$ decay was linked to the fact that neutral kaons can transform into their antiparticles and vice versa, but such transformation does not occur with exactly the same probability in both directions. This is called indirect CP violation. CP violation can also occur directly in the decay. Three types of CP violation can thus be defined:

- CP violation in the mixing: this occurs when physical states do not coincide with CP eigenstates, that is when $|q| \neq |p|$. As a consequence of this, the transition rates of $M^0 \to \bar{M}^0$ and $\bar{M}^0 \to M^0$ are different. This type of CP violation is of the indirect one.

- CP violation in the decay: it is when the decay amplitudes for CP conjugate processes into final states f and $\bar{f}$ are different in modulus, that is when $|A(i \to f)| = |A(\bar{i} \to \bar{f})|$ is not satisfied leading to different BR for CP-conjugated decays. This type of CP violation is of the direct type and is the only one which is possible for charged particles.

- CP violation in the interference of mixing and decay: this affects only the flavour neutral mesons and it requires the presence of a final state f that can be reached by both M^0 and $\bar{M}^0$. The processes in which the meson does or does not oscillate before decaying cannot be distinguished and can therefore interfere in the overall amplitude

  $$A(M^0 \to f) + A(M^0 \to \bar{M}^0)A(\bar{M}^0 \to f). \tag{1.27}$$

  The interference term gives the CP asymmetry. This type of CP violation cannot be unambiguously classified as direct or indirect, but any difference of this kind of CP violation between different final states is an evidence of direct CP violation, since the single phase in the mixing amplitude is the same for all decays.

When dealing with decays into a state f which can be reached by both flavour eigenstates, it is convenient to introduce the complex quantity

$$\lambda_f = \frac{q}{p}\frac{A(\bar{i} \to f)}{A(i \to f)} \tag{1.28}$$

that is used as a key parameter to study the CP violation since it contains the important features of the interference of the two meson states to final state f.

If $\lambda_f = 0$ then the decay is flavour-specific. The phase of λ_f is a physical quantity because λ_f is invariant with respect to the choice of the arbitrary relative phase between $|M^0\rangle$ and $|\bar{M}^0\rangle$. All three kinds of CP violation discussed above can be conveniently expressed in terms of λ_f. The condition imposed by CP symmetry is

$$\lambda_f = \frac{1}{\lambda_{\bar{f}}} \tag{1.29}$$

that become $\lambda_f = \pm 1$ in the case of final states which are CP eigenstates ($f = \bar{f}$). For a CP eigenstate, CP violation in either mixing or decay is indicated by

$$|\lambda_f| \neq 1 \tag{1.30}$$

while CP violation in the interference of mixing and decay corresponds to

$$Im(\lambda_f) \neq 0. \tag{1.31}$$

1.2.2 The Role of Kaons

The neutral kaon system, the lightest particles containing a quark from the second family, can be considered as the minimal flavor laboratory. Two neutral K particles exist: one with strangeness +1 (K^0) and its own antiparticle with strangeness -1 ($\bar{K}^0$). In this discussion, if not specifically mentioned the CP symmetry is assumed to be conserved. With an appropriate choice of phases CP $|K^0\rangle = |\bar{K}^0\rangle$ and CP$|\bar{K}^0\rangle = |K^0\rangle$.

Both kaons can decay into $\pi^+\pi^-$ via weak interaction, so it is allowed that they can couple through processes such as $K^0 \to \pi^+\pi^- \to \bar{K}^0$ (second order in weak interaction box diagrams with two W boson exchanges, $\Delta S = 2$, see Figure 1.5). In general, transitions between the two states would be allowed through the common virtual decay states. This shows that the K^0 and $\bar{K}^0$ are not the eigenstates of the Hamiltonian that governs the decays.

Considering CP as a good symmetry means that, in the neutral kaons system, the CP eigenstates are the physical states $|K_1\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\bar{K}^0\rangle)$ and $|K_2\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\bar{K}^0\rangle)$ for which CP$|K_1\rangle = +|K_1\rangle$ and CP$|K_2\rangle = -|K_2\rangle$. Experimentally, there are two observed kaon states called short-lived, K_S^0, and long-lived, K_L^0, with lifetimes $\tau_S = 0.89 \cdot 10^{-10}$s and $\tau_L = 5.17 \cdot 10^{-8}$s, respectively.

If the CP symmetry is respected in decay process, K_S^0 would only decay to states with CP = +1, while K_L^0 to states with CP = -1. The K_S^0 decay is dominated by $K_S^0 \to \pi^+\pi^-$, $\pi^0\pi^0$, whereas for the other one $K_L^0 \to \pi^+\pi^-\pi^0$, $\pi^0\pi^0\pi^0$ dominate because, due to angular momentum conservation, the CP-parity of the 2π final state is +1 while for the 3π final state is -1. Decaying in two or three pions leads to different phase space volumes, a smaller one in the second case resulting in a lifetime difference factor of about 600.

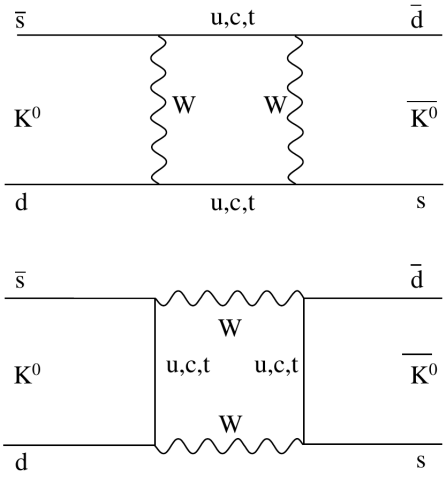


Figure 1.5: The box diagrams contributing to $K^0 - \bar{K}^0$ mixing.

Starting from an arbitrary mixture[8] of K_S^0 and K_L^0, the long lived component can be separated by just waiting long enough: far from the production point a beam will contain exclusively the long-lived K_L^0, while for a short time interval most decay come from the K_S^0 component. The existence of a long lived neutral kaon decay was observed in 1956 by Lande et al. [45] using a 3 GeV beam from the Brookhave Cosmostron.

K_1 and K_2 are not a particle-antiparticle pair and so their masses could also be different without violating the CPT symmetry. It has to be reminded that K^0 and $\bar{K}^0$ are eigenstates of the strong interaction while K_S^0 and K_L^0 are eigenstates of the weak interaction which governs the decay, meaning that only the weak eigenstates have a proper lifetime.

K^0 and $\bar{K}^0$ are not physical states with defined lifetime, so their time evolution is not a simple exponential. Starting with a state of K^0 at $t = 0$, the probability of measure positive or negative strangeness at the generic time t is:

$$P[K^0(t=0) \to K^0](t) = \frac{1}{4}[e^{-\Gamma_S t} + e^{-\Gamma_L t}] + 2e^{-(\Gamma_S+\Gamma_L)t/2} cos(\Delta m t) \tag{1.32}$$

$$P[K^0(t=0) \to \bar{K}^0](t) = \frac{1}{4}[e^{-\Gamma_S t} + e^{-\Gamma_L t}] - 2e^{-(\Gamma_S+\Gamma_L)t/2} cos(\Delta m t) \tag{1.33}$$

where $\Delta m = m_{K_L^0}$ - $m_{K_S^0}$. If the mass difference Δm is not zero, the time evolution of the strangeness eigenstates exhibits an interference oscillatory term, which allows a measurement of Δm ($3.5 \cdot 10^{-6}$ eV) but gives no information about its sign. These oscillations show that strangeness is not conserved in time and it has to be emphasized the empirical fact that, for neutral kaons, the oscillation frequency Δm is comparable in magnitude to the inverse of the decay lifetime allowing the effect of the oscillation to be sizable.

At first order in weak decay the $\Delta S = \Delta Q$ rule holds, which states that the change in strangeness and in electric charge of the hadrons in the initial and final states are linked. For example, in the neutral kaons, only the $K^0 \to \pi^- e^+ \nu_e$ and $\bar{K}^0 \to \pi^+ e^- \bar{\nu}_e$ are possible, while the decays with opposite charge of the electron and pion do not occur. The measurement of the charge of the electron allows to determine the strangeness of a neutral kaon at the time of its decay. This fact was used to experimentally study the strangeness oscillation effects

[8] This is a real situation because in the kaons production process neither K_S^0 nor K_L^0 are produced but a K^0 or a $\bar{K}^0$ instead.

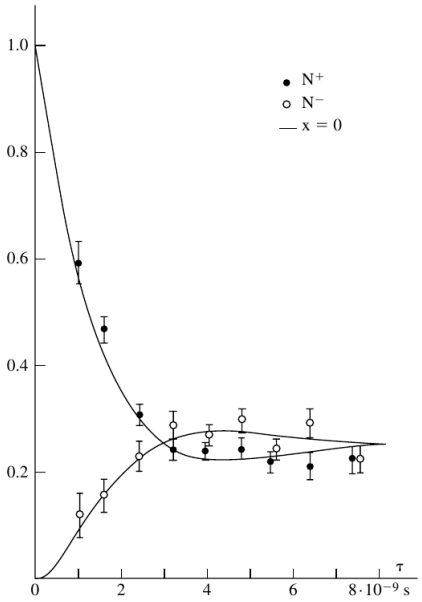


Figure 1.6: Strangeness oscillations in neutral K decays, starting with a pure K^0 beam (S = +1) at production (proper time $\tau = 0$) obtained from $K^+p \to K^0p\pi^+$. The points show the relative rates of $\pi^-e^+\nu$ decays (N^+, filled circles, S = +1) and $\pi^+e^-\nu$ decays (N^-, empty circles, S = -1) [46]. The lines indicate the expectation if the $\Delta S = \Delta Q$ rule holds.

(which is shown in Figure 1.6) counting the number of positive and negative electrons as a function of time when starting ($t = 0$) with a pure K^0 beam obtained from the $K^+p \to K^0p\pi^+$ reaction.

The K^0 and $\bar{K}^0$ interact, via strong force, differently with the matter, which has not a charge-conjugation symmetric composition. Indeed, a $\bar{K}^0$ has more final states available than a K^0 when interacting with a baryon, e.g. besides the charge-exchange process $\bar{K}^0n \to K^-p$ (which has the counterpart $K^0p \to K^+n$) also $\bar{K}^0p \to \Lambda\pi^+$ is possible, while an analogous reaction does not exist for K^0. Thus, the $\bar{K}^0$ cross section interaction is higher than that of the K^0.

If a neutral kaon beam travels a long distance, only K_L^0 are left. When a pure K_L beam traverses matter the differential interaction probability of K^0 and $\bar{K}^0$ leads to the regeneration of a K_S^0 component. This phenomenon, called regeneration of K_S^0 in a medium, was verified experimentally by O.Piccioni and collaborators at Berkeley in 1960 [47].

If CP is a good symmetry, the longer lived neutral kaon is strictly forbidden to decay into two pions. However, Cronin and Fitch measured decays of the K_L^0 into a $\pi\pi$ system reporting the decay ratio

$$\frac{\Gamma(K_L \to \pi^+\pi^-)}{\Gamma(K_L \to \text{charged particles})} = (2.0 \pm 0.4) \cdot 10^{-3}. \tag{1.34}$$

They supposed that it was the K_1^0 component which was decaying in the case they observed and thus they concluded that it was the existence of the CP forbidden K_1^0 component in the K_L^0 wave function which provided indirect CP violation.

The experiment was soon repeated and confirmed at several laboratories, but CP violation was not immediately accepted by the whole physics community. A conclusive piece of evidence came just one year later, when the interference between the 2π state obtained

from K_S^0 and K_L^0 decays was detected. For a state which is produced at $t = 0$ as a $K^0(\bar{K}^0)$ and propagates freely in vacuum, the decay rate to $\pi^+\pi^-$ is

$$\Gamma[K^0(\bar{K}^0)(t=0) \to \pi\pi](t) \propto e^{-\Gamma_S t} + |\eta_{\pi\pi}|^2 e^{-\Gamma_L t} \pm 2|\eta_{\pi\pi}| e^{-(\Gamma_L+\Gamma_S)t/2} \cos(\Delta m t - \phi_{\pi\pi}) \tag{1.35}$$

where

$$\eta_{\pi\pi} = \mid \eta_{\pi\pi} \mid e^{i\phi_{\pi\pi}} = \frac{A(K_L \to \pi\pi)}{A(K_S \to \pi\pi)}$$

is the CP-violating ratio of $K \to \pi\pi$ decay amplitudes and the +(-) sign applies to an initial $K^0(\bar{K}^0)$ state. The interference term, of opposite sign for initial K^0 and $\bar{K}^0$, is an indication of CP violation.

Another manifestation of CP violation was later detected as charge asymmetry in semi-leptonic decays $K_L^0 \to \pi^+ l^- \bar{\nu}_l$ vs. $K_L^0 \to \pi^- l^+ \nu_l$ (where $l = e, \mu$), first at Brookhaven [48] and then at [49]. These semi-leptonic decays of neutral kaons are flavour specific because of the $\Delta S = \Delta Q$ rule and are called K_{e3} and $K_{\mu 3}$, respectively for the electron and muon case. For K^0 corresponds a positive lepton charge, while negative lepton arises from $\bar{K}^0$ only, thus no direct CP violation is possible in semi-leptonic decays and CP violation arises only from the mixing. The semi-leptonic charge asymmetries for physical states are defined as

$$\delta_{L,S}^{(l)} \equiv \frac{\Gamma(K_{L,S} \to \pi^- l^+ \nu) - \Gamma(K_{L,S} \to \pi^+ l^- \bar{\nu})}{\Gamma(K_{L,S} \to \pi^- l^+ \nu) + \Gamma(K_{L,S} \to \pi^+ l^- \bar{\nu})} \tag{1.36}$$

and they are expected to be equal for both K_{e3} and $K_{\mu 3}$ and for K_S^0 and K_L^0. The asymmetry is zero if CP holds in the mixing, that is if the weights of K^0 and $\bar{K}^0$ are the same in the physical state expression.

The average measured asymmetries are

$$\delta_L^{(e)} = (3.34 \pm 0.07) \cdot 10^{-3} \quad \delta_L^{(\mu)} = (3.04 \pm 0.25) \cdot 10^{-3}. \tag{1.37}$$

Semi-leptonic decays of K_S^0 were also observed for the first time by KLOE at Frascati [50].

The small CP asymmetries means that the physical states are not too far from being the CP eigenstates and the indirect CP violation could be explained by introducing a complex parameters ϵ (intended to be small, $|\epsilon|^2 \ll 1$) in the wave function of K_S^0 and K_L^0:

$$|K_S^0\rangle = \frac{1}{\sqrt{2(1+|\epsilon|^2)}}[(1+\epsilon)|K^0\rangle + (1-\epsilon)|\bar{K}^0\rangle] \tag{1.38}$$

$$|K_L^0\rangle = \frac{1}{\sqrt{2(1+|\epsilon|^2)}}[(1+\epsilon)|K^0\rangle - (1-\epsilon)|\bar{K}^0\rangle]. \tag{1.39}$$

The mixing parameter ϵ is related to the q and p factors by:

$$\left(\frac{q}{p}\right) = \frac{1-\epsilon}{1+\epsilon}, \quad \epsilon = \frac{p-q}{p+q}. \tag{1.40}$$

Out of the four real parameters which are in principle required to describe K_S^0 in terms of K^0 and $\bar{K}^0$, one is not physically significant, as the relative phase of K^0 and $\bar{K}^0$ can be changed at will, and another is fixed by the normalization condition, so that a single complex parameter ϵ appears above. Moreover, the CP phase convention for K^0 and $\bar{K}^0$ is arbitrary and so another degree of freedom is removed leaving only the real part of ϵ as

physical quantity. This is consistent with the fact that CP violation in H is described (in case of CPT symmetry) by a single real parameter that is defined as

$$\delta = \frac{2Re(\epsilon)}{1+|\epsilon|^2}. \tag{1.41}$$

If $\epsilon \neq 0$ (and so $\delta \neq 0$), the physical states do not coincide with the CP eigenstates, indicating CP violation.

It has to be pointed out that flavour oscillations occur independently from the validity of CP symmetry and due to CPT symmetry the probability that at time t a state has the same flavour eigenvalue which it had at $t = 0$ is the same for K^0 or $\bar{K}^0$. Conversely, because of CP violation in the mixing, the probability that an initial K^0 becomes a $\bar{K}^0$ at time t is not the same as for a $\bar{K}^0$ becoming a K^0.

In decays to final states f with $\eta_{CP}(f) = +1$, forbidden for K_L^0 if CP symmetry holds, it is useful to define the following quantity:

$$\eta_f = \frac{A(K_L \to f)}{A(K_S \to f)} \tag{1.42}$$

which at first order in ϵ, for the $\pi^+\pi^-$ and $\pi^0\pi^0$ final states, can be expressed like

$$\eta_{+-} \simeq \epsilon + \epsilon' \quad \eta_{00} \simeq \epsilon - 2\epsilon' \tag{1.43}$$

with ϵ' a parameter describing the direct CP violation.

Thus, the two types of CP violation in the neutral kaon system are parametrized into ϵ and ϵ' with the measured quantity $Re(\epsilon'/\epsilon)$ is seen depending on both.

$$Re(\epsilon'/\epsilon) \simeq \frac{1}{3}(1 - |\eta_{00}/\eta_{+-}|) \tag{1.44}$$

If CP violation were only due to an asymmetry in the $K^0 - \bar{K}^0$ mixing all $\pi\pi$ decays of neutral kaons would be CP-conserving decays of the K_1 component, both for K_S^0 and K_L^0. In this case the ratio of $\pi^+\pi^-$ (charged) to $\pi^0\pi^0$ (neutral) decays should be the same for both physical states. The above expression clearly shows that ϵ' cannot be associated to $K^0 - \bar{K}^0$ mixing, as it affects differently the $\pi^+\pi^-$ and $\pi^0\pi^0$ final states. The modulus of the CP-violating parameters $\eta_{\pi\pi}$ are extracted from the measurement of $\Gamma(K_L^0 \to \pi\pi)/\Gamma(K_S^0 \to \pi\pi) =\mid \eta_{\pi\pi} \mid^2$ and the quantity

$$R \equiv \frac{\Gamma(K_L \to \pi^0\pi^0)\Gamma(K_S \to (\pi^+\pi^-)}{\Gamma(K_S \to \pi^0\pi^0)\Gamma(K_L \to (\pi^+\pi^-)} = \left|\frac{\eta_{00}}{\eta_{+-}}\right|^2 \tag{1.45}$$

is used to look for direct CP violation (R $\neq$ 1). The measurement of R is performed counting the number of K_S^0, K_L^0 decays into $\pi^+\pi^-$ and $\pi^0\pi^0$ and analyzing their ratio. The difficulty of the measure lies in avoiding any normalization bias in the translation from event numbers to partial decay widths. A round of dedicated experiments to measure $Re(\epsilon'/\epsilon)$ was completed at the end of the 90s, resulting in an unclear situation, but with more data collected, the experiments KTeV at FNAL [51] and NA48 at CERN [52] showed a clear $\epsilon^{'} \neq 0$ value obtaining

$$Re(\epsilon'/\epsilon) = (20.71 \pm 1.48 \pm 2.39) \cdot 10^{-4} = (20.7 \pm 2.8) \cdot 10^{-4} \tag{1.46}$$

$$Re(\epsilon'/\epsilon) = (14.7 \pm 1.4 \pm 0.9 \pm 1.5) \cdot 10^{-4} = (14.7 \pm 2.2) \cdot 10^{-4} \tag{1.47}$$

where the first quoted error is purely statistical, the central (if present) is the one induced by the finite statistic of the control samples used to study systematic effects and the last is purely systematic.

As a summer of CP violation in the kaon system in terms of ϵ and ϵ':

- CP violation in the mixing is expressed by $Re(\epsilon) \neq 0$ (or more precisely by $\delta \neq 0$). The CP violation appearing in the semi-leptonic decays are totally explained by this term.
- CP violation in the decays is expressed by $Re(\epsilon') \neq 0$.
- CP violation in the interference between decays with and without mixing is expressed by $Im(\epsilon) \neq 0$ and $Im(\epsilon') \neq 0$.

Indirect CP violation dominates in such decays because $|\epsilon'|$ is smaller than $|\epsilon|$. Indeed, from the measurements of the ratio of two neutral pions and two charged pions it is found that the direct CP violation contribution is the 0.14% with respect to the indirect one.

The decay $K_S^0 \to 3\pi^0$ violates CP just as $K_L^0 \to 2\pi$, but in this case the expected branching ratio is

$$BR(K_S \to 3\pi^0) \simeq |\epsilon|^2(\tau_S/\tau_L)BR(K_L \to 3\pi^0) \sim 10^{-9} \tag{1.48}$$

becoming challenging to be measured. This decay mode was searched at kaon-factories such as the $DA\Phi NE$ machine in Frascati where the properties of the coherent initial state ($K^0\bar{K}^0$ pairs produced from a ϕ resonance) allowed the possibility of tagging the presence of a K_S^0 by the identification of a K_L^0 decay obtaining a unique K_S^0 beam. The best current limit is $BR(K_S^0 \to 3\pi^0) < 2.6 \cdot 10^{-8}(90\%CL)$, that is translated into the CP-violating amplitude ratio $\eta_{000} = \frac{A(K_S \to 3\pi^0)}{A(K_L \to 3\pi^0)}$ and $|\eta_{000}| < 0.0088(90\%CL)$.

In the case of charged kaons where no mixing is possible, the only way to look for CP violation is to compare the decay properties of K^+ and K^-. The channel $K^\pm \to \pi^\pm\pi^+\pi^-$ and $K^\pm \to \pi^\pm\pi^0\pi^0$ have been studied and the results are consistent with no CP violation.

1.2.3 CP Violation in Heavy Neutral Mesons System

The neutral B and D mesons are the basis for some of the measurements performed at LHCb. Four families of flavoured neutral mesons, those with a net non-zero strangeness, charm or beauty content, are known: neutral kaons, neutral D mesons and neutral B_d and the strange B_s^0 mesons. All of them have the lowest mass in each family and are pseudo-scalar particles, while the higher mass states rapidly decay into these via strong interactions. Neutral mesons built from heavier quarks are not simple higher-mass replicas of the kaon system but, in contrast to kaons, they have no single dominant decay mode. The diagrams contributing to the oscillation of the heavy neutral mesons have the same structure of the one shown in Figure 1.5. While no qualitatively new features of CP violation were discovered in the B system, large effects were observed in many decay modes allowing significant quantitative test of the SM predictions. Moreover, in the B system it is possible to find not suppressed decays that allow a clear measurement of the weak phases appearing in the unitary triangle (e.g. $B^0 \to J/\psi K_S^0$ discussed in section 1.2.3.1). This situation does not arise in the kaon system, where the clearer channel to measure the weak phases is the $K^0 \to \pi\nu\bar{\nu}$ but this decay is very suppressed.

The technique to get the knowledge of the flavour state of a meson is called tagging. A way to obtain the flavour of meson at the production time is to identify the flavour quantum number of the particles produced together with it as the production process (strong or electromagnetic) is usually flavour-conserving and the total initial state has F = 0. For example this procedure is possible when the neutral meson is produced in association with a flavoured charged meson since the latter does not exhibit flavour oscillations and the observation of semi-leptonic decays of the charged meson identifies the flavour of the partner meson at the production time. Flavour mixing is often studied using semi-leptonic decays, which in the SM are flavour-specific decays as, e.g., only $B^0 \to X^- l^+ \nu_l$ or $\bar{B}^0 \to X^+ l^- \bar{\nu}_l$ are allowed and not $B^0 \to X^+ l^- \bar{\nu}_l$ or $\bar{B}^0 \to X^- l^+ \nu_l$. Any error in the flavour tagging dilutes the observed mixing ratio.

The techniques used for flavour tagging depend on the nature of the meson pair. For example, if a flavoured neutral meson is observed, it usually gives no information on the flavour of its partner because of the flavour oscillation with the exception of correlated meson pairs for which the oscillations are completely coherent. The observation of a flavour-specific decay for a meson of the pair at a given time t allows to tag the flavour of the other one at the same instant of time. From that moment on, the undecayed meson evolves as a meson of a given flavour and the situation for the decays of single tagged mesons is recovered. If the decay of the tagging meson occurs before the other meson, t is positive, otherwise is negative. Experimentally, the difference $\Delta t = t_1$ - t_2 of the two decay times is used to check for asymmetries to get information about CP violation in the time evolution of the pair. Comparing the time dependence of a flavoured neutral meson with its own antiparticle it is possible to extract the interference term between the two decays, which only depends on the CP-violating phase introduced in the mixing matrix V_{CKM}. This method allows a precise measurement of the α and β angles of the unitarity triangle, as discussed in section 1.1.1.1.

Correlated meson pairs can be produced at an e^+e^- collider running at the centre of mass energy at the production threshold for a resonance that can decay into the meson pair (for example the $\Upsilon(4S)$ for the B mesons).

The CP violation can contribute to decays in which both tree and penguin diagrams are present, because the strong and weak phases appearing in the penguin diagram amplitudes are in general different for different quarks appearing in the loop. The best situation occurs when the tree and penguin amplitudes share the same phase as is the case for $B_d \to J/\psi K^0_S$.

The remarkable consistency of data from B, B_s, and D mesons in mixing and CP violation provides enormous circumstantial evidences in favor of the CKM model for the weak interactions and CP violation in particular.

1.2.3.1 B Mesons

The two physical eigenstates of the neutral B mesons, due to the presence of many common decay channels, have very similar lifetimes, unlike the situation for neutral K mesons. Indeed, for the B meson the two states are named by their mass and called heavy (B_H) and light (B_L) instead of using the lifetimes as for the kaon system. However, the mass differences are of the order of $10^{-4}eV$ and $10^{-2}eV$ for B^0_d and B^0_s so, experimentally, the two mesons are distinguished by the different flavour, this means that a high efficiency flavour tagging is a necessary feature at the B physics experiments. The decays of B mesons can involve all three quark generations at tree levels and are thus a privileged system for the study of CP violation.

The study of B mesons has been carried out both at hadronic colliders and, since 2000, at dedicated B-factories. In the former, large B samples can be obtained with high backgrounds, while the latter, despite lower production cross sections, takes advantages from the coherent nature of meson pairs produced from the first $b\bar{b}$ resonance above the $b\bar{b}$ production threshold, that is the $\Upsilon(4S)$ which decays almost exclusively into B^+B^- or $B^0\bar{B}^0$. Another advantage of the B-factories is the low hadronic background below the $\Upsilon(4S)$ resonance compared to the B meson pairs produced and to what is achieved at the hadronic colliders.

If the lab frame coincides with the center-of-mass frame, the $\Upsilon(4S)$ is produced almost at rest and the B mesons travel only about 30 μm before decaying. Such decay lengths are too short to be measured with sufficient accuracy to see the oscillations. To overcome this, asymmetric-energy e^+e^- colliders were built[9] obtaining a typical decay length of 250 μm, which is enough to be measured with a silicon vertex detector.

If a $B^0\bar{B}^0$ pair is created and both mesons decay semileptonically, the B^0 would be expected to give a positive lepton ($\bar{b} \to \bar{c}l^+\nu$) and the $\bar{B}^0$ a negative one. With the $B - \bar{B}^0$ mixing, it is possible that both leptons have the same sign. The first hints of B flavour mixing were obtained by the UA1 experiment at the CERN $p\bar{p}$ collider [53] and by the ARGUS experiment at the DESY DORIS II e^+e^- storage ring looking for like-sign di-lepton pairs.

The asymmetry of the $B^0 - \bar{B}^0$ oscillation (indirect CP violation) was measured using inclusive events with two leptons in the final state. The lepton charge of the semi-leptonic decays was used to tag the neutral meson flavour and the asymmetry between like-sign lepton pairs,

$$\frac{N(l^+l^+) - N(l^-l^-)}{N(l^+l^+) + N(l^-l^-)} \tag{1.49}$$

is the quantity of interest, from which measurement the BaBar and Belle experiment obtained the results of $|q/p| = 0.9992 \pm 0.0027 \pm 0.0019$ and $|q/p| = 1.0005 \pm 0.0040 \pm 0.0043$ respectively. These results show no appreciable CP violation in the mixing at the level of few ‰.

The first manifestation of CP violation outside the kaon system was seen in 2001 in the B mesons by BaBar and Belle analyzing the decay asymmetry of the channel $B^0 \to J/\psi K^0_S$, arising in the interference of mixing and decay amplitudes. The decay $B_d \to J/\psi K^0_S$ [10] is the golden mode for B-factories, as it allows a clean extraction of $\sin(2\beta)$ of the unitarity triangle. Moreover the part which is difficult to compute, namely the hadronic matrix element linking the quark process to the observed hadronic process, drops out in the amplitude ratio $\bar{A}_f/A_f$ which modulus is 1 because there is only one dominant weak phase in the process forbidding the CP violation in the decay. Thanks to the fact that the ratio $|q/p|$ can be approximate as 1, it is possible to show that

$$\lambda_{J/\psi K_S} \simeq -\frac{V^*_{tb}V_{td}}{V_{tb}V^*_{td}}\frac{V^*_{cb}V_{cd}}{V_{cb}V^*_{cd}} = e^{2i\beta} \tag{1.50}$$

$$Im(\lambda_{J/\psi K_S}) \simeq \sin(2\beta) \tag{1.51}$$

[9] In the '90s two asymmetric B-factories were built: the PEP-II at SLAC and the KEK-B at KEK that hosted, respectively, the BaBar and Belle experiments, both designed to look for CP violation in B mesons.

[10] Strictly speaking the CP-eigenstates $K_{1,2}$ should be used instead of $K^0_{S,L}$, but in this case the difference is quantitatively irrelevant.

allowing a clear measure of the weak phase 2β from the measurement of the interference term obtained by the time asymmetry

$$a_f(t) = \frac{\Gamma(\bar{B}^0(t) \to f) - \Gamma(B^0(t) \to f)}{\Gamma(\bar{B}^0(t) \to f) + \Gamma(B^0(t) \to f)} = \frac{2Im\lambda_f}{|\lambda_f|^2 + 1} \sin(\Delta mt) + \frac{|\lambda_f|^2 - 1}{|\lambda_f|^2 + 1} cos(\Delta mt)$$
$$\simeq \sin(2\beta) \sin(\Delta mt).$$

Besides the tree process, penguin diagrams (Figure 1.7) containing quark loops also contribute to the decay, but the dominant one (unsuppressed by CKM matrix elements) shares the same phase as the tree amplitude, while the size of the other one is only 1% of the tree amplitude, so that the above conclusion is valid to a high degree of accuracy.

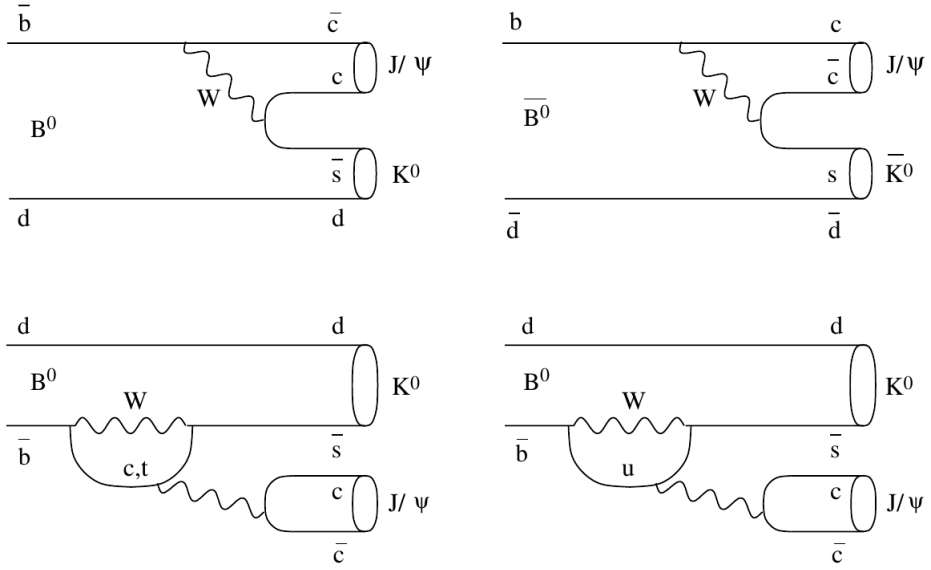


Figure 1.7: Diagrams contributing to the $B_d \to J/\psi K^0$ decay. Top row: the tree decays $B_d \to J/\psi K^0$ and $B_d \to J/\psi K^0$ mixing into B_d, $B_d \to J/\psi K^0_S$. Bottom row: the penguin decays, the left-hand one having the same phase as the tree decay.

By March 2001 both the Belle and BaBar collaborations reported values for $\sin 2\beta$. The Belle measured $0.58^{+0.32}_{-0.34}(stat)\ ^{+0.09}_{-0.10}(sys)$ [54], while BaBar $0.34 \pm 0.20 \pm 0.05$ [55]. A few months later, the BaBar collaboration announced an updated result $\sin(2\beta) = 0.59 \pm 0.14 \pm 0.05$ [56], which taken alone was enough to establish CP violation in the B system. The current result of BaBar is $\sin(2\beta) = 0.741 \pm 0.067 \pm 0.034$ [57], confirmed at the hadron colliders by LHCb that obtained a results of $\sin(2\beta) = 0.731 \pm 0.035 \pm 0.020$ [58].

In 2004, evidences for direct CP violation were observed in the decay of $B^0 \to K^+\pi^-$ [59, 60] by the B-factory BaBar and Belle. If CP invariance holds, the number of decays of $B^0 \to K^+\pi^-$ ($N_{K^+\pi^-}$) and $\bar{B}^0 \to K^-\pi^+$ ($N_{K^-\pi+}$) out of $B^0\bar{B}^0$ pairs should be identical except for contribution from CP violation in the mixing that is known to be small for B mesons. The partial decay rate asymmetry

$$A_{CP}(K\pi) = \frac{N(K^-\pi^+) - N(K^+\pi^-)}{N(K^-\pi^+) + N(K^+\pi^-)} \tag{1.52}$$

was measured to be

$$A_{CP}(K^\pm\pi^\mp) = -0.133 \pm 0.030 \pm 0.009 \quad \text{(BaBar)} \tag{1.53}$$

$$A_{CP}(K^{\pm}\pi^{\mp}) = -0.113 \pm 0.022 \pm 0.008 \quad \text{(Belle)} \tag{1.54}$$

thus establishing the existence of direct CP violation in the decay of B mesons with a significance exceeding 5 standard deviations. Up to now, direct CP violation in the interference of mixing and decay has been measured in several decay modes.

The measurement of the much faster B_s^0 mixing required an higher sensitivity in a time-dependent analysis which was achieved only in 2006 by the CDF experiment at the TeVatron collider [61]. Oscillations of B_s^0 are similar in principle to those of the non-strange B^0, but replacing the d quark with an s quark causes that V_{ts}^2 instead of V_{td}^2 governs the oscillation resulting in an increase of Δm by a large factor. The B factories were unable to study B_s^0 oscillation because they run at the $\Upsilon(4S)$ resonance and so at a center-of-mass energy that kinematically forbids the creation of $B_s^0\bar{B}_s^0$ pair but only the $B_d\bar{B}_d$ pair. The $\Upsilon(5S)$ is the first resonance that can be used to obtain a coherent B_s^0-$\bar{B}_s^0$ pair and indeed it will be used by the super B-factories (e.g. the Super KEKB collider) [62].

Today, the measurement of Δm_s is lead by LHCb which obtain a value of $17.768 \pm 0.023 \pm 0.006\text{ps}^{-1}$ [63]. For the B_s^0 mesons case the phase of the CKM factors appearing in the box diagrams is smaller than for B^0 by two orders of magnitude and the phase difference, which could be measured from the decay $B_s^0 \to J/\psi\phi$, is

$$\beta_s \equiv \arg\left[-\frac{V_{tb}^* V_{ts}}{V_{cb}^* V_{cs}}\right] \tag{1.55}$$

and it is a small angle of the squashed triangle corresponding to the b-s unitarity relation.

1.2.3.2 D Mesons

D mesons contain charm as the heavier quark and behave somewhat different from K and B mesons. Neutral D mesons contain u quarks and charged D mesons either d or s quarks. The lifetimes are ~ 1 ps for $D^{\pm}$ and ~ 0.5 ps for D^0 and $D_s^{\pm}$, much shorter than those of B mesons when the quark mass scaling ($\Gamma \propto m_q^5$) is taken into account because D mesons decays are not suppressed by small CKM angles. Such lifetimes are however long enough to allow a direct experimental observation of the meson decay lengths and so time dependent measurements.

Only a small contribution from the FCNC is expected in $D^0 - \bar{D}^0$ oscillations (no contribution from the top quark) and the indirect CP violation in the SM are supposed to be very small. The fact that mixing and CP violation are highly suppressed in the SM (the last has not yet seen experimentally) makes the D system a good laboratory to search for NP effects and large CP violation would indicate contributions from new particles.

CP-violating effects in the interference of mixing and decay can be searched in decays into CP-eigenstates such as $\pi^+\pi^-$ and K^+K^-, and also in other flavour non specific states such as $K^{\pm}\pi^{\mp}$. For example, both $\bar{D}^0$ and D^0 can decay into $K^+\pi^-$, the latter being a doubly Cabibbo-suppressed decay and thus with a better chance to compete with the oscillation amplitude $D^0 \to \bar{D}^0 \to K^+\pi^-$. Measurements of A_{CP} in $D \to KK$ and $D \to \pi\pi$ decays are under study (also at LHCb) and up to date no clear evidence of CP violation has been found.

D mesons are studied both at hadronic and at e^+e^- colliders, the latter can be used as a D-factory when the center of mass energy is at the $\psi(3770)$ resonance, which is just above the $D\bar{D}$ threshold but below the $D^*\bar{D}$. However, B-factories also produce very large samples of D mesons, both from continuum and from B decays and, currently, hadronic colliders

provide the most precise results and LHCb has become one of the leading experiment in this sector.

1.3 The $K_S^0 \to \mu^+\mu^-$ Decay

In this section a brief description of the $K_S^0 \to \mu^+\mu^-$ and $K_L^0 \to \mu^+\mu^-$ SM prediction is provided following [5] [64] [65].

The $K_L^0 \to \mu^+\mu^-$ BR has been observed to be $(6.84 \pm 0.11)10^{-9}$ [66] and perfectly agree with the SM prediction, while the $K_S^0 \to \mu^+\mu^-$ BR is expected to be $(5.1 \pm 1.5)10^{-12}$ in the SM and the present upper limit on it is $9.0 \cdot 10^{-9}$ at 90%CL obtained by LHCb [67], still very far from the SM expectations.

The non observation of the $K_S^0 \to \mu^+\mu^-$ decay played an historical role in particle physics, inspiring the GIM mechanism which predicts the suppression of FCNC decay with the introduction of a new quark, later identified as the c quark. Indeed, the $K_S^0 \to \mu^+\mu^-$ decay is a FCNC which does not arise at tree level in SM because the two final state muons can only be directly produced from a photon, Higgs or Z, but none of these bosons can be originated by a *s-d* quark interaction directly, since they can be produced only by particles of the same flavour. The $K_S^0 \to \mu^+\mu^-$ and $K_L^0 \to \mu^+\mu^-$ decays were also studied to test the understanding of CP violation in the kaon system in the '60-'70 years during which the last measurement (before that obtained by LHCb) was performed.

The most general amplitude for the $K^0 \to l^+l^-$ (l being a lepton) processes includes a s-wave (A) and p-wave (B) component:

$$\mathcal{A}(K^0 \to l^+l^-) = \bar{u}_l(i\mathrm{B} + \mathrm{A}\gamma_5)v_l \tag{1.56}$$

which yields as corresponding decay rate:

$$\Gamma(K^0 \to l^+l^-) = \frac{m_K\beta_l}{8\pi}(|\mathrm{A}|^2 + \beta_l^2|\mathrm{B}|^2) \tag{1.57}$$

where $\beta_l = \sqrt{1 - \frac{4m_l^2}{m_K^2}}$, m_K is the K^0 mass and m_l is the mass of the lepton in the final state.

These two components have opposite CP such that CP-conserving contributions to K_L^0 and K_S^0 decays are generated by A and B, respectively.

Long-distance[11](LD) contributions are driven mainly by virtual intermediate two-photon states and leads to neglect the B amplitude in the case of $K_L^0 \to \mu^+\mu^-$ because this contribution has a negligible CP-violating component. On the other hand, in the $K_S^0 \to \mu^+\mu^-$ case we need to keep both types of amplitudes. Short-distance(SD) SM like contributions will only enter in the A term and in this case it is expected to be non-negligible, so that the SD part of A has to be taken into account both in K_S^0 and K_L^0 decays.

The proportion between the LD and the SD contribution is model dependent. In the SM the LD contributions are dominant. The Figure 1.8 shows examples of diagrams contributing to the SD and LD components.

The amplitude for K_S^0 can be written as

$$\Gamma(K_S^0 \to l^+l^-) = \frac{m_K\beta_l}{8\pi}[(ImA_{SD})^2 + \beta_l^2(ReB_{\gamma\gamma})^2 + \beta_l^2(ImB_{\gamma\gamma})^2] \tag{1.58}$$

[11]The long-distance scales correspond to masses below that of the c quark, while short-distance scales correspond to masses above or equal to that of the c quark.

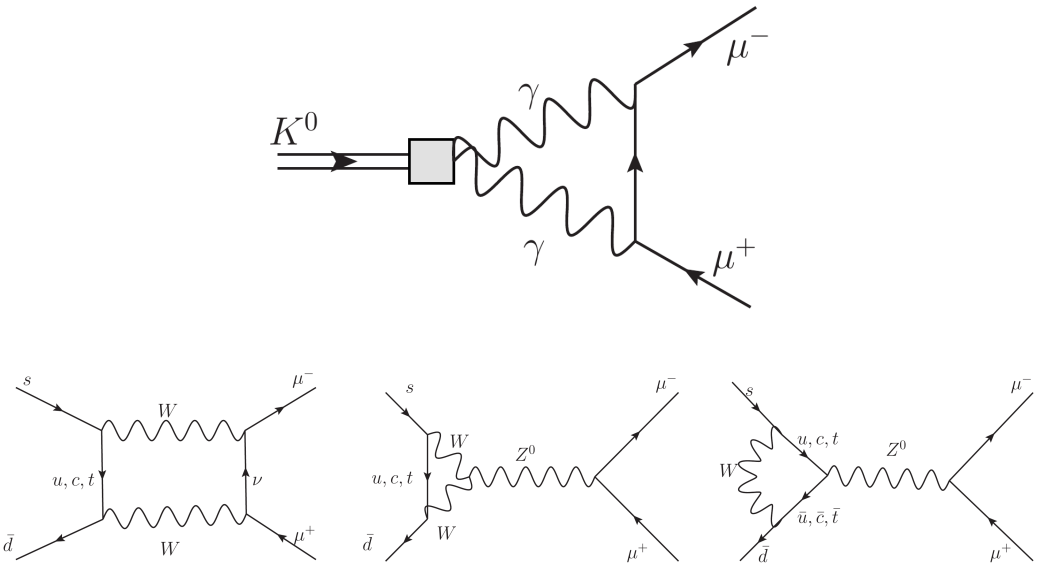


Figure 1.8: Examples of diagrams contributing to $K_S^0 \to \mu^+\mu^-$. Top: long distance contribution, generated by two intermediate photons. In principle, the meson loop involves both charged pions and kaons but it is found that only pions contribute to the amplitude. Bottom: short distance, with penguins and boxes.

where A_{SD} is the SD component of the A amplitude and $B_{\gamma\gamma}$ the B amplitude (LD), dominated by diagrams with two intermediate virtual photons.

Concerning the LD contributions, the subprocess $K_1^0 \to \gamma^*\gamma^*$ is determined by the one-loop diagrams in Figure 1.9 and, normalizing to the rate for $K_S^0 \to \gamma\gamma$ [12], one obtains the relative branching ratios

$$\frac{\Gamma(K_S \to \mu^+\mu^-)}{\Gamma(K_S \to \gamma\gamma)} = 1.9 \cdot 10^{-6} \tag{1.59}$$

that turns into

$$\mathcal{B}(K_S^0 \to \mu^+\mu^-)_{LD} = 5.1 \cdot 10^{-12} \tag{1.60}$$

with an error at the level of 30%. For the consideration of CP violating effects, the SM SD gives only the contributions arising by the second-order electroweak interactions, in particular Z^0 penguins and $W^\pm$ box (see Figure 1.9) which are largely dominated by the top-quark contribution receiving small QCD corrections. Within SM the SD contribution is

$$\mathcal{B}(K_S^0 \to \mu^+\mu^-)_{SD} = 10^{-5} \cdot |Im(V_{ts}^* V_{td})|^2 = 1.4 \cdot 10^{-12} \left|\frac{V_{cb}}{0.041}\right|^4 \cdot \left|\frac{\lambda}{0.223}\right|^2 \bar{\eta}^2 \tag{1.61}$$

and, using the current value of $\bar{\eta}$ from the global CKM fits [18], a value of $B(K_S^0 \to \mu^+\mu^-)_{SD}$ $\sim 10^{-13}$ is obtained, which is smaller than the estimated error in the LD component.

Merging both SD and LD results yields a SM completely dominated by LD contributions, so that the final BR estimation is $B(K_S^0 \to \mu\mu) = (5.1 \pm 1.5) \cdot 10^{-12}$.

[12] The theoretical branching ratio is $BR(K_S \to \gamma\gamma) = 2.1 \cdot 10^{-6}$, while the experimental branching ratio [68] $BR(K_S \to \gamma\gamma) = (2.4 \pm 1.2) \cdot 10^{-6}$, in good agreement with the prediction.

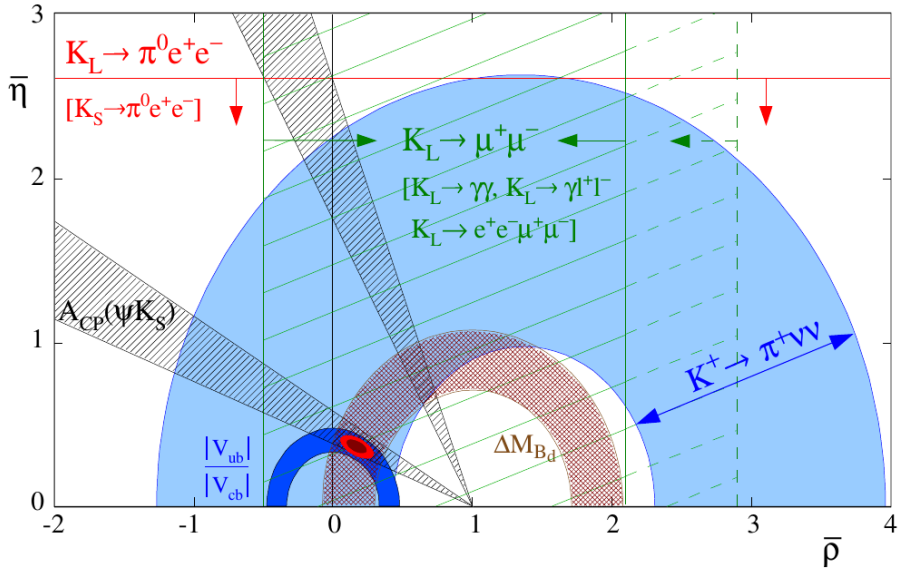


Figure 1.9: Summary of the present constraints in the $\bar{\eta}$ vs. $\bar{\rho}$ plane from rare K decays. The bounds in $\mathrm{B}(K_S^0 \to \mu^+\mu^-)$, which is proportional to $\bar{\eta}$, and in $\mathrm{B}(K_L^0 \to \mu^+\mu^-)$, which is proportional to $\bar{\rho}$, are orthogonal. The constraints on $|V_{ub}|$ and $B - \bar{B}$ mixing are also shown for comparison. Figure taken from [5].

On the theory side, the $K_S^0 \to \mu^+\mu^-$ amplitude is particularly interesting since its SD component is dominated by the CP-violating part of the $s \to dl^+l^-$ amplitude, which is very sensitive to NP and which is poorly constrained so far. An evidence of $K_S^0 \to \mu^+\mu^-$ well above 10^{-11} would be a clear signal of NP, while the non observation of such enhancement will provide significant constraints on several consistent NP scenarios. Indeed, in $K_S^0 \to \mu^+\mu^-$ has also been suggested as a possible way to look for new light scalars particles. The SD contribution to the $\mathrm{B}(K_S^0 \to \mu^+\mu^-)$ is proportional to $\bar{\eta}$ and so a limit in its contribution can be expressed as effective bound on the CKM parameter $\bar{\rho}$. In Figure 1.9 the current bounds in the $\bar{\eta}$ vs. $\bar{\rho}$ plane are reported. This limit in the $\mathrm{B}(K_S^0 \to \mu^+\mu^-)$ could also constrain other important channels of the kaon sector, such as $K^+ \to \pi^+\nu\bar{\nu}$. As can be seen in Figure 1.9, this bound is not comparable in precision with present CKM constraints from B physics, still competes with the other constraints derived only from $\Delta S = 1$ FCNC processes which are rather weak.

Thus, we can conclude that the rare decay $K_S^0 \to \mu^+\mu^-$ is a very useful source of information for the SD structure of new $\Delta S = 1$ FCNC transitions as possible new high mass particles can contribute in the loop as virtual process.

The neutral kaons are still today, an important laboratory to look for insight of NP. A full dedicated experiment [69] to study the very rare CP-violating $K_L^0 \to \pi^0\nu\nu$ decay is in progress. The peculiar characteristic of this decay is that it is one of the cleanest weak decay from the theoretical point of view, with a predicted BR in the SM of about $3 \cdot 10^{-11}$ and theoretical uncertainties at 1% level. In particular, the theoretical difficulties related to strong interactions are limited by the presence of only a single hadron in the final state and they can be constrained from the well-measured $K \to \pi e\nu$ decay.

Chapter 2

The LHCb Experiment at the LHC

In this chapter the LHCb experiment is described. In the first section (2.1) the accelerator system is briefly introduced, showing the luminosity achieved during the years 2010, 2011, 2012. In 2.2 an overview of the physics studied at LHCb is given. The description of the detector is reported in 2.3, where the characteristics of each sub-detector are shown. The particle identification in LHCb is discussed in 2.4 and the trigger aspects in 2.5. A short introduction about the software used by the LHCb Collaboration is presented at 2.6, while the LHC Computing Grid is described at 2.6.2. Finally, the differences between the 2011 and 2012 experimental conditions are clarified in 2.7.

2.1 LHC

The Large Hadron Collider (LHC) [70] is currently the man-made world most powerful particle accelerator. It is located at CERN [71], near Geneva (CH). The LHC is installed in the tunnel that previously contained the Large Electron Positron collider (LEP). The ring is 27 km long and it is situated approximately a hundred meters underground.

The LHC has two beam pipes to guide the accelerated protons in the ring in opposite directions. At these energies, most of the interesting hard processes between the partons that constitute the proton (the same its true for an anti-proton) are dominated by interactions bewtween gluons (or sea quarks), so that there is only a small difference between $p-p$ and p-$\bar{p}$ collisions but proton-proton collisions allow to reach high luminosity that would be more difficult to achieved by a proton-anti-proton accelerator.

In Figure 2.1 the CERN acceleration complex, including the LHC and all the preaccelerators, is sketched. The protons are produced from a hydrogen duoplasmatron source and are accelerated to 50 MeV by a linear accelerator (LINAC 2). They are fed into the Proton Synchrotron Booster where they are further accelerated up to 1.4 GeV. From here they are passed into the Proton Synchrotron (PS) where they are separated into bunches and accelerated to 25 GeV. After the PS, the protons are accelerated by the Super Proton Synchrotron (SPS) to an energy of 450 GeV before being injected into the main ring of the LHC, where they are boosted to the final energy.

Under nominal conditions, the LHC would contain 2340 bunches, each composed by about $1.15 \cdot 10^{11}$ particles. The bunches are spaced by 25 ns or multiples. In 2011 and 2012 it has been preferred to run with a bunch spacing of 50 ns, to optimize the achievable luminosity, and the number of bunches per beam was 1380. The bunches move in opposite directions around the ring and collide at 4 distinct points, where the detectors are

located. The LHC is designed to run at a centre-of-mass energy of 14 TeV with a maximum instantaneous luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$ [72].

Presently, protons collide at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. During 2011, both the proton beams were accelerated up to 3.5 TeV, while in 2012 run, when the data used for this thesis were taken, the collisions took place at $\sqrt{s} = 8$ TeV.

Inside the beam pipe an ultra high vacuum is maintained at a pressure of about 10^{-13} atm in order to avoid collisions between the beam and residual gas molecules.

At these high energies, to maintain proton orbits a very intense magnetic field is needed (8.33 T for a 7 TeV beam [72]). This field is provided by superconducting dipole magnets (for a total of 1232, 15 m long), kept at a temperature of 1.9 K using a complex cryogenic system based on superfluid helium. Thanks to the superconducting dipole magnets a current of about 11700 A is provided. Other kinds of magnets in the LHC (for a total of about 8000 magnets) are quadrupoles, exapole, octupoles and decapoles used to focus the beams and to control their orbits. The main contribution to the focusing is given by the quadrupoles.

The acceleration (from 450 GeV to 7 TeV) is achieved by superconducting radiofrequency (RF) cavities which also give contribution to maintain each bunch grouped.

The LHC is divided into eight linear sections and eight circular sections. In the former the beam collimators, the RF cavities, the experimental areas and also the location where the beams are deflected at the end of a fill (beam dump) are placed, while in the latter the dipole and other magnets are located.

Because of the two proton beams travel in opposite directions through the ring, opposite magnetic field directions are needed. To achieved this condition the magnet coils surrounding the two beam channels are embodied inside the same iron yoke.

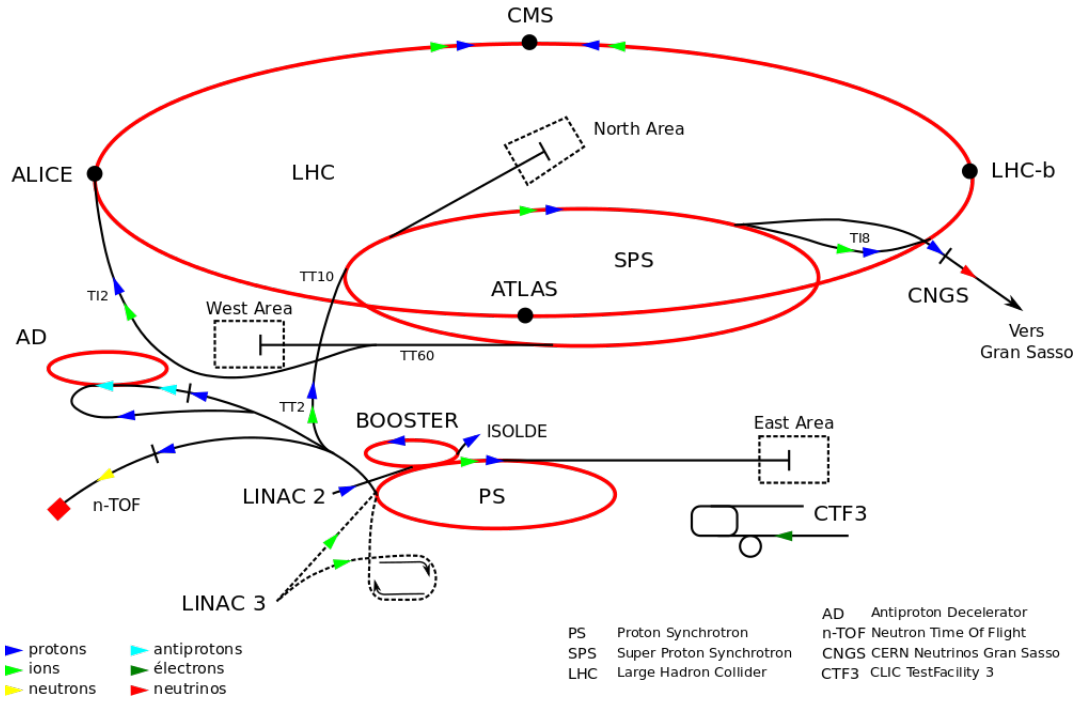


Figure 2.1: The preacceleration scheme of the LHC system.

The aim of the LHC project is to test the SM and to look for signals of physics beyond it.

The four main experiments at the LHC are located at each of the four interaction points, see Figure 2.1, where the proton beams collide. These four detectors sit underground in huge caverns on the LHC ring. The biggest experiments are ATLAS [73] and CMS [74] which are general-purpose experiments. This allows cross-confirmation of any new discoveries between the two experiments. Both are central detectors constructed by large and international Collaborations.

ALICE and LHCb are not general-purpose experiment, which means they do not investigate the largest range of physics possible, in fact they are designed for particular physical processes. In ALICE the study of strongly interacting matter at the extreme energy densities in heavy-ion collisions are performed, realizing measurements of the phase transition between hadronic matter and the quark-gluon plasma. For this latter aspect, the LHC is filled in dedicated runs with heavy ions (e.g. Pb) instead of protons.

Finally, LHCb [75] focuses on B physics, rare decays and CP violation measures. The physics performed by LHCb will be better introduced in section 2.2.

2.1.1 Luminosity

The instantaneous luminosity is an important parameter for an accelerator, as it allows to estimate the expected event rate for a process with a known cross section and to measure the cross section from the rate of the process. Indeed, it is defined as the ratio between the event rate and the cross section.

In order to estimate the number of events expected in a given data sample, the integrated luminosity $L_{int} = \int L(t)dt$ is usually preferred with respect to the instantaneous luminosity is L. The integrated luminosity achieved by LHCb during the Run-1 of the LHC is showed in Figure 2.2. The data-taking is divided in three sets according to the year of data-taking: 2010, 2011 and 2012.

To ease the multiplication by a known cross sections to estimate the number of events in a given period, the integrated luminosity is usually expressed in inverse picobarn (pb^{-1}) or its multiples with one barn that corresponds to 10^{-24} cm^2.

The point where a proton-proton collision takes place is called a primary vertex (PV), while the decaying point of a flying particle produced in the PV is called secondary vertex (SV). For the physics performed at LHCb the association of the PVs and the SVs need to be as good as possible. In LHCb, the luminosity is maintained smaller than the LHC nominal one ($10^{34}\text{cm}^{-2}\text{s}^{-1}$) because of the high rate of particle produced at high rapidity which can saturate and damage the sub-detectors and the associated electronics. Indeed, LHCb works at an instantaneous luminosity of $2 - 4 \cdot 10^{32}\text{cm}^{-2}\text{s}^{-1}$ allowing to improve the association of the PVs and SVs.

To maintain the luminosity level constant at the desired value, LHCb uses a technique named luminosity leveling. This technique is implemented through a misalignment of the beams near to the interaction point followed by a progressive realignment, dynamically driven by the luminosity online monitoring. An example of monitoring for the instantaneous luminosity during a bunch fill run is shown in Figure 2.3, while the average number of proton-proton collisions and the peak luminosity achieved during the whole first run of LHC are shown in Figure 2.4. An approximately constant instantaneous luminosity allows to increase the integrated luminosity and preserving stable trigger conditions during the run.

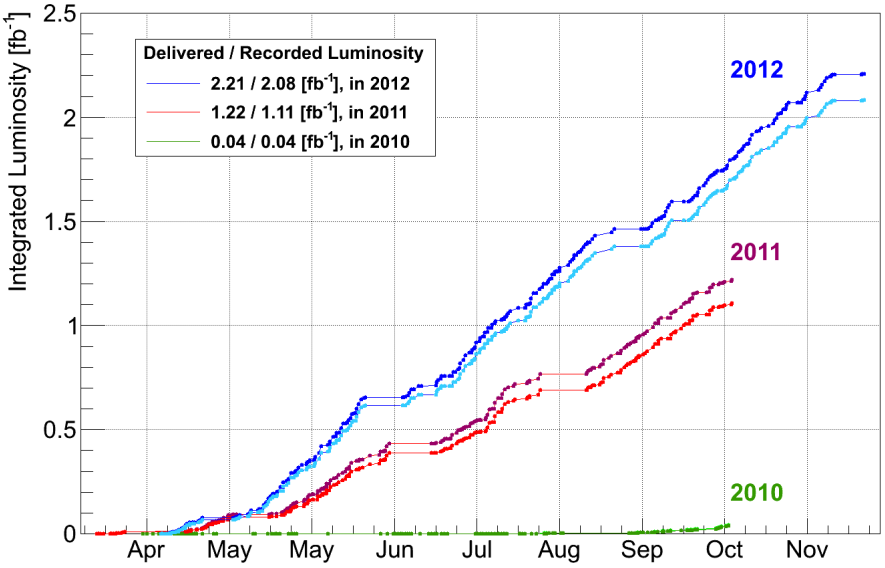


Figure 2.2: Integrated luminosity in LHCb during the three years of LHC Run-1. The figure shows the curves for the delivered (dark coloured lines) and recorded (light coloured lines) integrated luminosities.

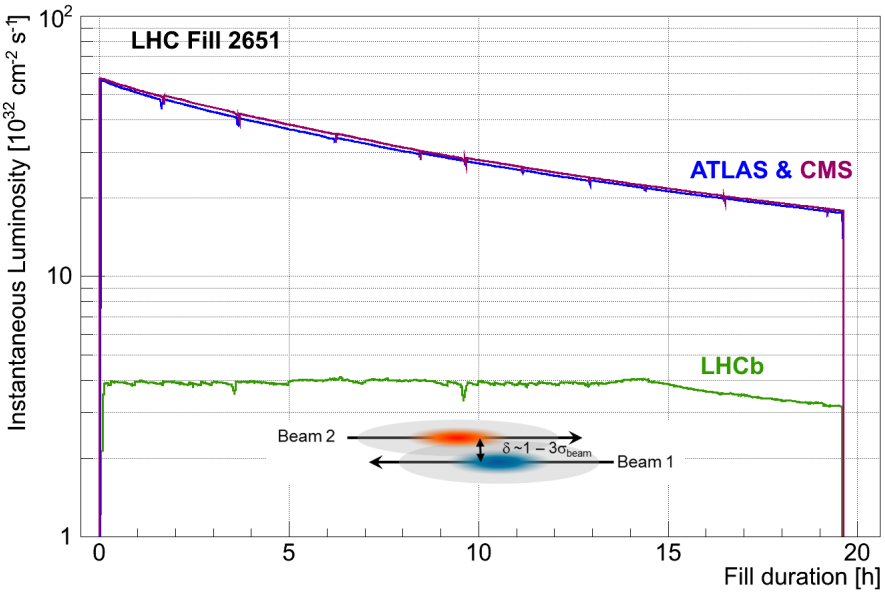


Figure 2.3: Development of the instantaneous luminosity for ATLAS, CMS and LHCb during LHC fill 2651. After ramping to the desired value of $4 \cdot 10^{32}\text{cm}^{-2}\text{s}^{-1}$ for LHCb, the luminosity is kept stable in a range of 5% for about 15 hours by adjusting the transversal beam overlap. The difference in luminosity towards the end of the fill between ATLAS, CMS and LHCb is due to the difference in the final focusing at the collision points.

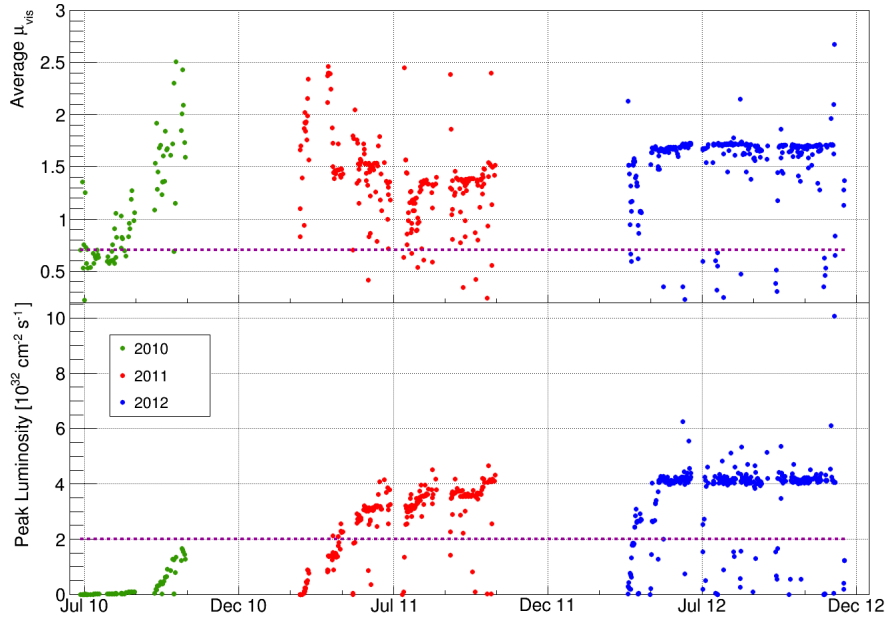


Figure 2.4: Average number of visible interactions per bunch crossing (top) and instantaneous luminosity (bottom) at the LHCb interaction point in the period 2010-2012. The dotted lines show the design values. In the last year of the Run-1 both the quantities was well under control to the fixed value of 1.7 average primary vertices and $4 \cdot 10^{32}\text{cm}^{-2}\text{s}^{-1}$.

2.2 The Physics Studied at LHCb

The SM gives remarkable results describing the elementary particles world, but there are experimental evidences that cannot be explained by the SM. Two examples of these open questions are the cosmological evidence for Dark Matter and the baryonic asymmetry of the universe. None particles in the SM are satisfactory candidates for the Dark Matter and the CP violation in the SM is not enough to explain the asymmetry of the matter seen in the universe. To search for NP in High Energy Physics (HEP), two complementary approaches are carried out at experiments using accelerators: a direct search (e.g. ATLAS and CMS), looking for the production of new particles, and an indirect search (e.g. LHCb) that looks for effects due to the presence of new particles in the loop (because the contribution is virtual, the mass of the new particles can be higher than what is directly achievable). Some of the physics studied at the LHCb experiment to discriminate between NP scenarios lies in the indirect searches for NP via precise measurements of quantum loop induced processes in the b and c-quark systems. A way to observe these contributions is to search for enhancement on the BR of very rare decays that are theoretically under control. NP can also manifest itself as violations of the tight SM constraints in the flavour sector, notably in CP-violating processes. For this reason precise measurements (e.g. the four parameters of the CKM mass mixing matrix, in particular ρ and η) are necessary.

Compared to other accelerators that are in operation or under construction, the LHC will be the most copious source of B mesons, due to the high $b\bar{b}$ production cross section and the high luminosity. About 10^{12} $b\bar{b}$ pairs are expected to be produced in one year of data taking. The LHCb detector is designed to exploit this large number of b-hadrons in order to make precision studies of CP asymmetries and of rare decays in the B-meson systems. LHCb has

the possibility, due to the high statistics available, to study B_s^0 decays with high accuracy allowing, notably, to measure the mixing-induced CP violation phase in $B_s^0 \to J/\psi\varphi$ decays and the oscillation frequency between B_s^0 and $\bar{B}_s^0$ ($\Delta m_s = (17.768 \pm 0.023 \pm 0.006)$ ps^{-1} [61]). In addition to investigating CP violation in B meson decays, the physics programme of the LHCb experiment includes studies of rare B, D, K, τ decays, $D - \bar{D}$ oscillations, hadron spectroscopy, notably on b hadrons and more (EW physics, heavey ions, ...).

The first Technical Proposal of the LHCb experiment was submitted in February 1998 [76] and the definitive version in September 2003 [77]. In December 2009, a roadmap document [78] was published to describe six of the key measurements concerning CP asymmetries and rare B decays, that the LHCb Collaboration was expecting to achieve during the first run of the LHC. The six measurements included in the document are:

- The tree level determination of the unitarity triangle angle γ with several techniques, both time dependent and time independent. The angle γ is the worst measured angle of the CKM unitarity triangle. The world average of this measurement is now lead by the LHCb results [79]. The observation of a discrepancy among the values of γ determined using processes involving loops or tree diagram processes would imply that NP has appeared in the loops. The results obtained by LHCb show no discrepancy between the γ angle obtained from loop and tree diagram.

- Charmless charged two-body B decays. The family of B hadrons, that decays into pairs of charmless charged mesons or baryons, has a rich set of channels, each one characterized by charge or time dependent CP asymmetries providing many different ways to test the SM picture of CP violation. The decays under study are of the type $H_b \to h^+h^-$, where H_b can be either a B^0 meson, a B_s^0 meson or a Λ_b baryon, while h stands for π, K or p. These decays provide an interesting way to determine the angle γ with loops, for which LHCb measurements are more precise than with tree-diagram decay only [80].

  The results obtained by LHCb in this field are summarized in [81], with particular emphasis on the first observation of direct CP violation in the B_s system.

- Measurement of mixing-induced CP violation in $B_s^0 \to J/\psi\phi$ caused by interference between the decays with and without B_s oscillations. The interference between the direct amplitude and the amplitude via B_s^0-$\bar{B}_s^0$ oscillation gives rise to the CP violating phase $\phi_s = arg[-V_{ts}V_{tb}^*/V_{cs}V_{cb}^*]$, that is the smallest angle of the b-s unitarity triangle. This phase is one of the CP observables with the smallest theoretical uncertainty in the SM and can be significantly modified by new particles contribute to the B_s^0 -$\bar{B}_s^0$ box diagram. The measurement $\phi_s = -0.010 \pm 0.039$ rad, published in [82], represents the world best measurement of the phase ϕ_s, which is consistent with the SM expectations (see Figure 2.5).

- Measuring the BR of the rare $B_s(B^0) \to \mu^+\mu^-$ decay. This decay belongs to the family of the FCNC, and its BR is precisely predicted by the SM to be $BR(B_s^0 \to \mu^+\mu^-) = (3.66 \pm 0.23) \cdot 10^{-9}$ and $BR(B^0 \to \mu^+\mu^-) = (1.06 \pm 0.09) \cdot 10^{-10}$. The main conclusion of the first generation of B-decay experiments can be expressed by saying that the CKM description of flavour-changing processes has been confirmed in $b \to d$ transitions, but on the other hand, NP effects can still be large in flavour-changing $b \to s$ transitions. So the decay $B_s^0 \to \mu^+\mu^-$ has been identified as a very interesting

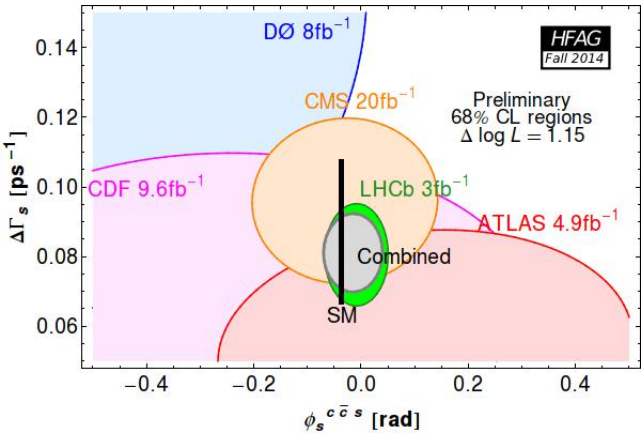


Figure 2.5: Likelihood confidence regions in the $\Delta\Gamma_s - \phi_s$ plane. The results obtained from many experiments are shown. The combined results are dominated by the LHCb one, which is the most precise. The black square and error bar correspond to the SM prediction.

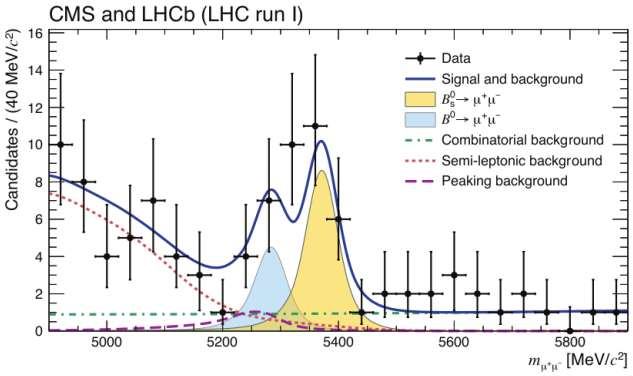


Figure 2.6: Weighted distribution of the dimuon invariant mass for all categories. The combined fit (solid blue) and its components are superimposed on the data points in black: the B_s^0 (yellow shaded) and B^0 (light-blue shaded) signal components; the combinatorial background (dash-dotted green); the sum of the semileptonic backgrounds (dotted salmon); the peaking backgrounds (dashed violet).

measurement that could show a clear indication of NP and/or constrain the parameter space of models describing physics beyond the SM.

The first evidence for the $B_s^0 \to \mu^+\mu^-$ decay was reported by LHCb in [83] with the 2011 data, now the BR has been measured by combining the results of the CMS and LHCb Collaborations [2](see Figure 2.6), resulting in perfect agreement with the SM expectation. The analysis of this channel has an approach similar to that of the $K_S^0 \to \mu^+\mu^-$ as they share the same request for an high muon identification performance and the reduction of the combinatorial background with a MVA, while the specific background is estimated with a fit in the data sidebands as for the $K_S^0 \to \mu^+\mu^-$.

- Measuring the forward-backward asymmetry of the muon pair in the FCNC $B^0 \to K^{*0}\mu^+\mu^-$ decay process. Such a FCNC in the SM only occurs through box and loop diagrams so the properties of the decay can be strongly modified by NP. The decay

$B^0 \to K^*\mu^-\mu^+$ is interesting for the same reasons of the $B_s \to \mu^+\mu^-$ but, moreover, it is possible to extract a number of angular observables which are well calculable theoretically and offers discrimination between different NP models. The LHCb experiment measured a significant deviation from the SM predictions in one of these angular observables [84] which has risen a considerable interest in the theory community. A class of existing models could explain these deviations and can accommodate also for the other LHCb results of the anomaly found in the ratio of the decay rates of the two modes $B \to K\mu^+\mu^-$ and $B \to Ke^+e^-$ [85]. However, it is too early to draw a firm conclusion, with the $B \to K^*\mu^+\mu^-$ observables being potentially susceptible to poorly known QCD effects.

- Analysis of $B_s^0 \to \phi\gamma$ and other radiative B decays such as, e.g., $B_d^0 \to K^*\gamma$, $B_s^0 \to \phi\gamma$ and $B^+ \to \phi K^+\gamma$. These analysis consist in measuring properties of radiative B decays (B meson decays with photons in the final states) such as the photon polarisation. Radiative decays are associated to FCNC so that they are another interesting probe for physics beyond the SM. The analysis published in [86] has observed the first direct evidence of photon polarization in a $b \to s\gamma$ transition wich is found to be consistent with the SM.

The roadmap of the physics analyses accessible at LHCb has been expanded thanks to new ideas and to the excellent performance of the LHC and of the LHCb detector, such as the vertex, the trigger and the particle identification capabilities (e.g. hadronic physics in the forward direction, exotic states production, kaon physics, soft-QCD, etc...).

Several important physics achievements have been reached. For example, LHCb has performed world leading measurements of the lifetime of many b-hadrons and has given important contributions in the field of charm physics. The production of quarkonia states is also an attractive field where LHCb is playing a prominent role.

Examples of the key physics measurements performed by LHCb during the first Run can be found in http://lhcb-public.web.cern.ch/lhcb-public/ and in the list below are reported some of the most important one:

- Recently, the LHCb Collaboration has published the first evidence for a new state of matter called pentaquark [87]. The pentaquark-charmonium (P_c^+) states has been observed to decay into a J/ψ meson and a proton p ($\Lambda_b \to (P_c^+ \to J/\psi p)K^-$) analyzing the Run-1 data. In his fundamental 1964 paper [88], in which he proposed the quark model, Gell-Mann mentioned the possibility of adding a quark-antiquark pair to a minimal meson or baryon quark configuration. It has taken 50 years, however, for measurements to unambiguously demonstrate the existence of these exotic states. The internal mechanism of quarks interactions inside pentaquarks is still to be understood with both possibilities, that the quarks are tight bound or loosely bound in a meson-baryion molecule, still open.

- Study of the branching fraction ratio $\mathrm{R}(D^*) = \bar{B}_0 \to D^{*+}\tau^-\bar{\nu}_\tau / \bar{B}_0 \to D^{*+}\mu^-\bar{\nu}_\mu$ which, due to the lepton universality and mass difference of τ and μ, is expected to be 0.252 ± 0.003 in the SM. The preliminary results have shown a value of $0.336 \pm 0.027 \pm 0.030$, resulting in a 2.1 standard deviation discrepancy from the SM prediction [89]. This ratio is considered to be precisely calculable thanks to the cancellation of uncertainties associated with the B to D^* meson transition.

- Precise measurement of the unitarity triangle angle β [58]. The measurement is made studying the decays of about 41500 B^0 and $\bar{B}^0$ mesons to J/ψ and K^0_S mesons obtaining a value of $sin(2\beta) = 0.731 \pm 0.035 \pm 0.020$. The results found are consistent with the expectation provided by the measurements of the other parameters of the unitarity triangle and hence with the predictions of the SM itself.

- Measure of the ratio R_K of the probability that a B^+ meson decays to a $K^+\mu^+\mu^-$ or a $K^+e^+e^-$. In the SM of particle physics this ratio expected to be very close to one thanks to the so called lepton universality. The ratio R_K is measured in function of the $\mu^+\mu^-$ and e^+e^- invariant mass squared q^2. The LHCb result, which is the most precise, is $R_K = 0.745^{+0.090}_{-0.074} \pm 0.036$ measured in the q^2 range between 1 and 6 GeV^2, different from 1 with a 2.6 σ significance giving hints of possible NP[1] [90]. It is interesting to note that the result of BaBar Collaboration at low q^2 favors also a value below one, while the measurement by the Belle Collaboration is consistent with one in the whole q^2 range up to $22GeV^2$. The Belle result does not, however, contradict a possible q^2 dependence of R_K which may be indicated by the BaBar results.

- Observation of the $D^0 - \bar{D}^0$ oscillations with more precision with respect to the previous experiments (e.g. CDF) obtaining the first observation above the 5 σ level [91]. However, no evidence for CP violation is found, providing the most stringent bounds on the parameters and thus constraining the room for physics beyond the SM.

- First observation of CP violation in the decays of B^0_s mesons looking at the asymmetry in the $B^0_s \to K^-\pi^+$ decay. The result obtained is $A_{CP}(B^0_s \to K^-\pi^+) = +0.27 \pm 0.04 \pm 0.01$ giving a significance exceeding five Gaussian standard deviations [92].

- Improved search for CP violation in charm decays. The difference of CP asymmetry ($\Delta A_{CP} = A_{CP}(\pi^+\pi^-) - A_{CP}(K^+K^-)$) between the decay rates of D^0 and $\bar{D}^0$ mesons into K^+K^- pairs and into $\pi^+\pi^-$ pairs has been measured to be compatible with no CP violation in the charm system [93].

- The B^0 decays into $D^+\pi^-$ and $J/\psi K^{*0}$ are used to study $B^0 - \bar{B}^0$ oscillations. The measured value $\Delta m_d = (0.5156 \pm 0.0051 \pm 0.0033)$ ps^{-1} [94] is about 37 times slower than $B^0_s - \bar{B}^0_s$ oscillations while the $B^0 - \bar{B}^0$ oscillations have been previously measured at LEP, Tevatron and B factories, the LHCb result is currently the most precise.

- CP violation asymmetry in the charmless three body B meson decays $B^\pm \to K^\pm\pi^+\pi^-$ and $B^\pm \to K^\pm K^+ K^-$ obtaining a significance of 2.8 σ for the first and 3.7 σ for the second channel [95]. The latter is the first evidence of inclusive CP asymmetry in charmless three-body $B^\pm$ decays. It is interesting to note that much larger asymmetries are observed in some small special regions, like invariant mass squared of the $\pi^+\pi^-$ pair in the $K^\pm\pi^+\pi^-$ decay lower than 1 GeV/c^2, or of the K^+K^- pair in the $K^\pm K^+K^-$ final state between 1.2 and 2 GeV/c^2. Further study of this intriguing feature are ongoing to make light on these results.

[1] For example the existence of additional Higgs bosons, new scalar or new pseudo-scalar particles that violate lepton universality, or "heavy Z", Z', which were suggested as a possible explanation of deviations observed the analysis of $B^0 \to K^*\mu^+\mu^-$ decays.

To date no results from the experimental particle physics performed have shown strong inconsistencies with the SM prediction[2].

LHCb has yet a data sample of b decays that is larger than the sum of all previous experiments. Moreover, the LHCb Collaboration is working on the upgrade of the detector, which should be installed in 2018 and 2019 [96]. The aim of the upgrade is to increase the instantaneous luminosity and possibly acquire about 50 fb^{-1} of data and this need an upgrade of most detectors and of the trigger system. LHCb promises a rich spectrum of rare and precise measurements that have the potential to fundamentally affect our understanding of the SM and CP-violationg phenomena.

The subject of this thesis is an example of search for NP contributions in the very rare decay process $K_S^0 \to \mu^+\mu^-$ which is a FCNC expected only in second-order in the electroweak interaction within the SM and so sensitive to heavy virtual particle. Indeed, $K_S^0 \to \mu^+\mu^-$ decay has been proposed as a novel way to profit from the excellent performance of the LHCb detector. While very suppressed in SM, such a decay can have NP contributions (e.g. coupled to the strange quark).

2.3 The LHCb Detector

The LHCb detector [1] (Figure 2.7) is located in the cavern previously occupied by the DELPHI LEP experiment. It is a single arm spectrometer with angular coverage from 10 mrad to 250 mrad in the vertical plane (non-bending plane, corresponding to a pseudorapidity range $2 < \eta < 5$) and ± 300 mrad in the horizontal (bending plane).

LHCb is designed for the detection of hadrons containing the b or $\bar{b}$ quark. The choice of the detector geometry is justified by the fact that at the high energies of the LHC both the b and $\bar{b}$ hadrons are predominantly produced in the same forward or backward cone (see Figure 2.8).

In order to be able to discriminate the $b\bar{b}$ events against the hadronic background, which have respectively cross sections of about 500 μb^{-1} and 100 mb, optimal detector performances are required. The focus on B physics study imposes many requirements on the design of the detector such as a need for excellent primary and secondary vertex resolution to allow for precise measurements of the proper decay time as well as good momentum resolution and excellent particle identification.

Indeed, the key features of LHCb include: an excellent vertex and proper time resolution; a precise particle identification; a precise invariant mass reconstruction which is required to efficiently reject background due to random combinations of tracks (combinatorial background) and implies a good momentum resolution; a versatile trigger scheme with an High efficiency in both leptonic and hadronic B decay channels, in order to collect high statistics samples and study the variety of modes with small BR.

In the following subsections the main elements of the LHCb detector, listed below, are described.

- Vertex locator (VELO) system, subsection 2.3.1;
- Spectrometer magnet, subsection 2.3.2;
- Tracking System, subsection 2.3.4;

[2]Some hints, that are still under study, of possible SM deviations have been obtained in the $B \to K^{\pm}\mu^+\mu^-$ amplitude, $B^+ \to K^+\mu^+\mu^- \neq B^+ \to K^+e^+e^-$, $B \to D^+\tau/D^+\mu$, V_{ub} inclusive $\neq$ exclusive.

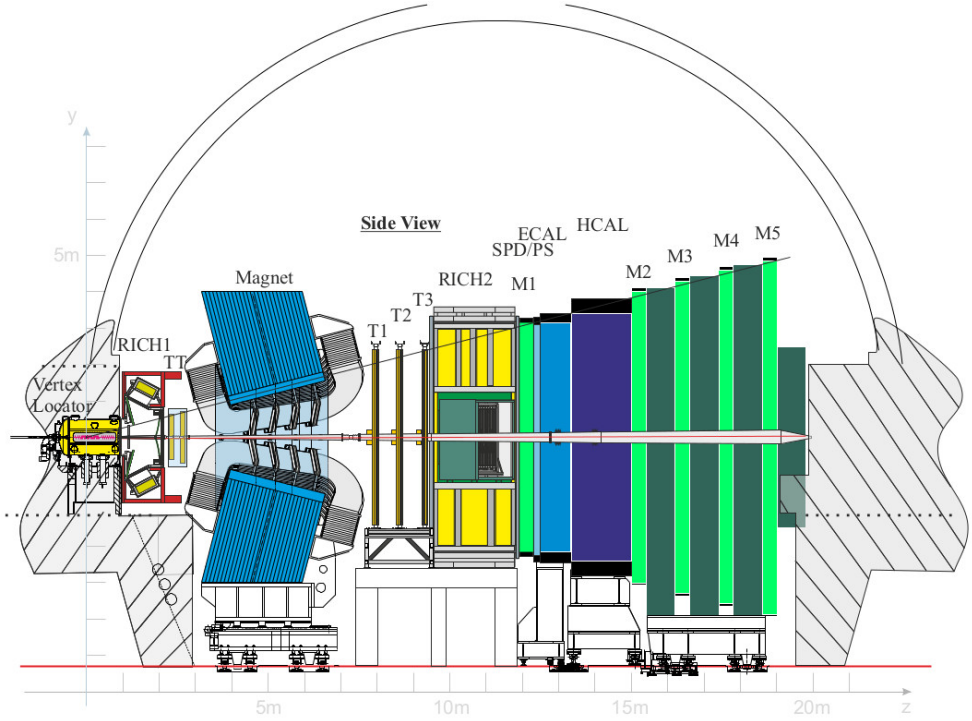


Figure 2.7: View of the LHCb detector [97]. From left to right: Vertex Locator (VELO), RICH1, tracker station (TT), Magnet, tracker stations (T1-T3), RICH2, preshower (SPD/PS), Electromagnetic and Hadronic calorimeters, Muon stations (M1-M5). The right-handed coordinate system adopted has the z axis along the beam, and the y axis almost along the vertical since there is a $\sim$3.7 mrad tilt compared to the actual geometrical vertical. Thus, the x axes is almost horizontal.

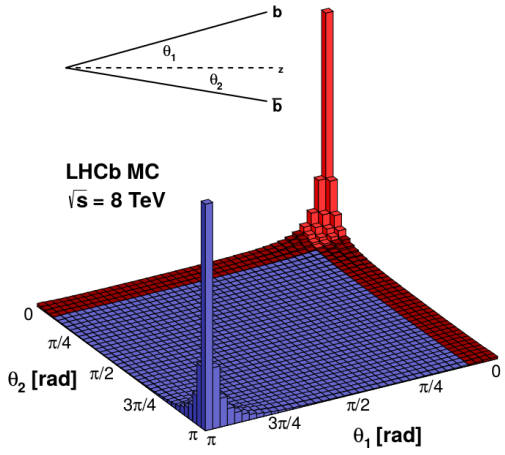
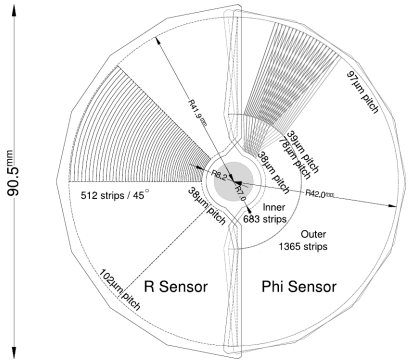
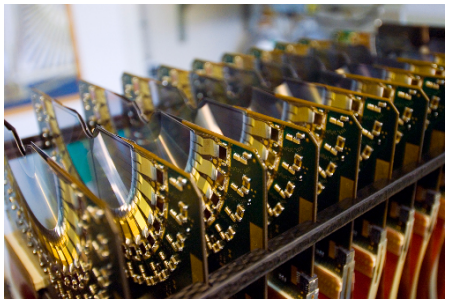


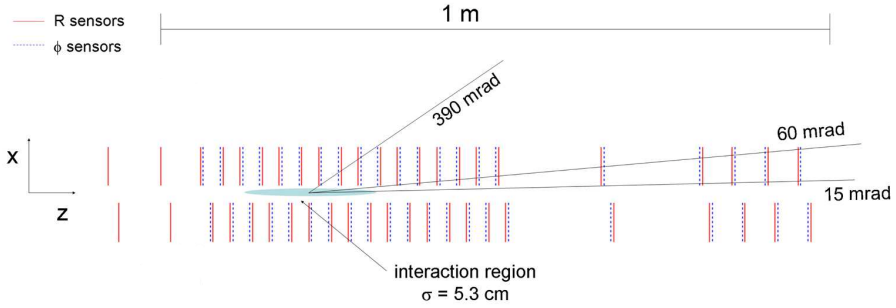
Figure 2.8: The two-dimensional histogram representing the correlation in the directions of angles the B mesons θ_1 (θ_2) is events the angle generated between by the a PYTHIA $B(\bar{B})$ meson simulation flight direction of $pp \to B\bar{B}X$ and the events beam at axis.



(a) The scheme of a station of the VELO detector. The rho and phi pitch are reported as also the dimension of the entire sensor.



(b) A picture of half of the sensor that make one of the 23 stations of the VELO. When the bunch of protons in the LHC reach a stable condition, the two half move closer to increase the precision in the PV and SV measure.



(c) Representation of the rho and phi sensors together to form the VELO. The angular acceptance are shown respect to the bunch crossing point.

Figure 2.9

- Two Ring Imaging Cherenkov detectors (RICH1 and RICH2), subsection 2.3.3;
- Calorimeter system, subsection 2.3.5;
- Muon detection System, subsection2.3.6.

The LHCb detectors can be conceptually subdivided in two subsystems: particle identification system (RICH, calorimeters and muon system) and tracking system (vertex locator, spectrometer magnet and tracking stations).

2.3.1 Vertex Locator

The Vertex Locator (VELO) [98] [77] is a silicon detector composed by 23 stations perpendicular to the beam axis. In Figure 2.9 are reported the schemes of a single station and of all the stations that compose the detector.

The function of the VELO detector is to measure the position of the vertices, which is very important to study the b and c-hadron decays because of their distinctive feature of a well detached decay vertex.

The more stations that a particle travels though, the more accurate the measurement of the vertex will be. Particles created from the PV itself or from a decaying particle close to the PV will leave more hits in the VELO. For example, muons produced in $J/\psi \to \mu^+\mu^-$, due to the short lifetime of the J/ψ, have a different reconstruction and trigger efficiency with respect to muons from $K^0_S \to \mu^+\mu^-$. The VELO resolution on PV varies between 9 and 35 μm for the x and y coordinates, while for the z coordinate a range of 50 and 280 μm is achieved, depending on the number of tracks used to reconstruct the PV.

A station consists of a series of silicon strips modules, each providing a measure of the radial distance from the origin, r, and of the azimuthal angle, φ. The half-disc sensors are arranged in pairs of r and φ sensors. The radial dimension of each module is around 42 mm and 8 mm for the outer and inner radius, respectively. To measure the r coordinate circular strips centered around the beam axis are arranged. The pitch of the strips increase linearly from a value of 38 μm at the innermost radius to 101.6 μm at the outermost one. With this, a spatial cluster resolution of about 4 μm is achieved for 100 mrad tracks.

The φ measure is obtained from radial strips and the sensor is divided into two radial sections to keep the occupancy low. With this division the strip pitch, that increases linearly towards the outer edge of the sensor, can be lowered in the outer part. In the inner region the pitch varies from 35.5 μm to 78 μm, while in the outer starts from 39.3 μm increasing up to 97 μm. The strips are not aligned perfectly in the radial direction but are skewed at an angle of 20° in the inner section and -10° in the outer section with respect to the radial direction. In alternating stations the angles are reversed to create a stereo effect. This geometry is summarized in the top left (a) picture of Figure 2.9.

The acceptance of the VELO covers the pseudorapidity range $1.6 < \eta < 4.9$ for particles coming from PVs in the range -10.6 < z < 10.6 cm. Only particles with very small or very large (15 mrad $> \theta >$ 300 mrad) polar angle will escape without being tracked from the VELO detector.

Another important information obtained from the VELO detector is the impact parameter (IP) of a charged track. The resolution of the VELO on this quantity ranges between 10 and 80 μm depending on the transverse momentum of the considered track. To quickly determine the IP for the trigger, only the r measurements can be used to identify tracks with an high impact parameter that will be better reconstructed in the offline procedure.

The closer to the PV the sensors are, the smaller is the error on the reconstructed position of the vertex. For this reason the VELO works at the shorter possible distance (about 8 mm) respect to the beam axis. The transversal dimension of the beam during the start of a fill variates until the stable data-acquisition situation is reached. Before this condition is obtained the VELO halves are moved away from the beam in the horizontal direction to avoid radiation damage or, in a worse case, that beam goes straight through the sensors. The safe minimal distance in this phase is 3 cm. Once the beams reach stable conditions, the detectors are placed into the nominal position transversally centered around the interaction region, with the two halves a little overlapped to cover the whole space.

2.3.2 Magnet

To achieve a high momentum resolution for the charged particles, the LHCb dipole magnet [99] provides a vertical (y axes) magnetic field with a peak of about 1.1 T (see Figure 2.10).

No restriction for the acceptance is caused by the opening in the center of the magnet, which is large enough to see the acceptance of the rest of the detector. The dipole is composed by a Fe yoke surrounded by two identical coils of conical saddle shape produced

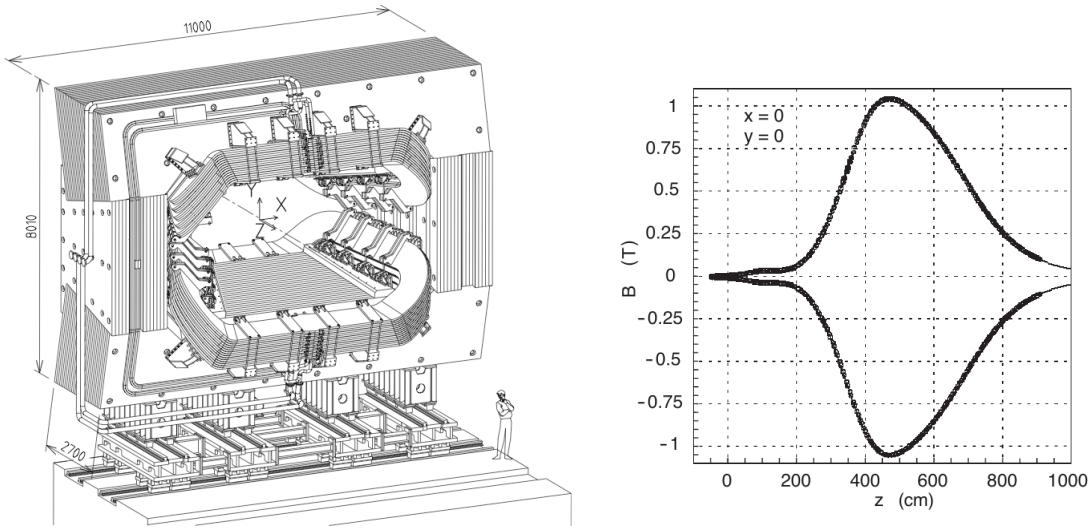


Figure 2.10: On the left a sketch of the LHCb dipole magnet is reported with the Cartesian system coordinates used. On the right is shown the map of the most relevant magnetic field component, B_y, that bends the particle in the horizontal plane, as a function of z coordinate, for both polarity.

of pure Al-99.7.

The strength of the field has to be known with excellent precision, in order to yield a momentum resolution as good as possible that is necessary to perform high quality physics analysis. The precision of the measurement obtained for the field mapping in the tracking volume is about $4 \cdot 10^{-4}$.

The measurements of CP asymmetries can be affected by systematic effects due to asymmetries in the detector. To prevent these systematic effects the magnetic field is inverted periodically during the data-taking, collecting data for approximately equal amounts of time for each polarity.

2.3.3 RICH

The LHCb detector includes two RICH detectors (scheme in Figure 2.11), both used for particle identification, in particular to differentiate between pions and kaons in the decay of B mesons.

RICH detectors are based on the Cherenkov light emission of a particle traversing a material with refractive index n. A charged particle passing into a material with a higher speed respect to the velocity of the light in the material has a probability to irradiate photons which are called Cherenkov light. The Cherenkov angle, θ_c, is fixed by the relation $\cos\theta_c = \frac{1}{n\beta}$. For a fixed radiator material, θ_c depends only on the particle velocity (see Figure 2.12).

The two RICH cover different momentum ranges, the first one (RICH1, closer to the VELO) works in ∼1-70 GeV/c range, while the other (RICH2) from ∼12 GeV/c up to and beyond 150 GeV/c. Also the acceptance is different: RICH1 covers the full LHCb acceptance (from ±25 mrad to ±300 mrad in the horizontal direction and ±250 mrad in the vertical one), while RICH2 has a more limited angular acceptance (from ±15 mrad to ±120 mrad in the horizontal direction and ±100 mrad in the vertical one). RICH2 is optimized for high momentum particles, which are supposed to remain in the geometrical acceptance after the magnetic deflection, justifying its placement after the magnet.

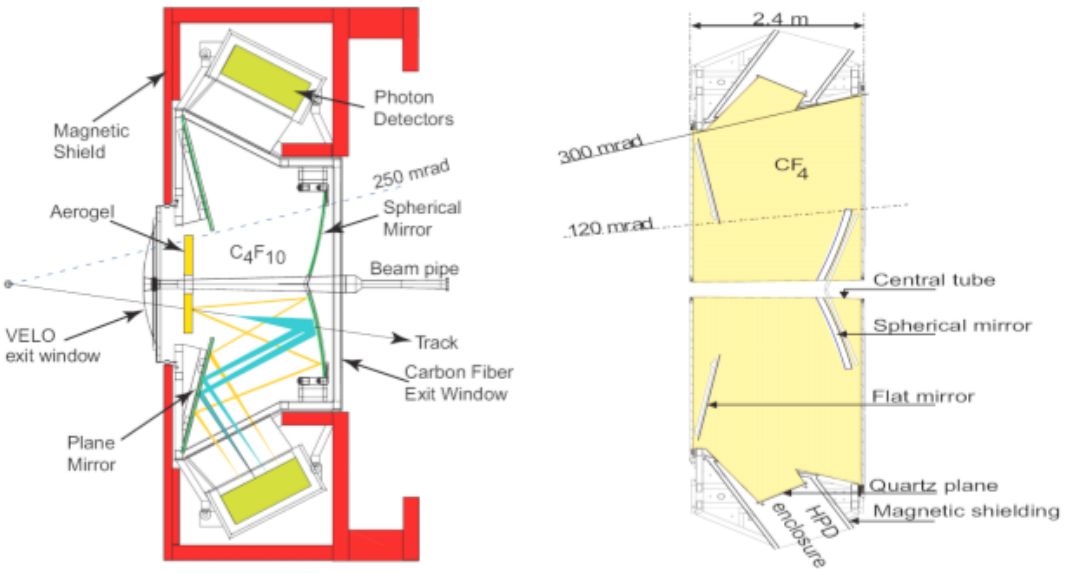


Figure 2.11: Scheme of the RICH1 (left) and RICH2 (right) detector showing the travel of the cherenkov light up to the HPDs that convert the photons in electronic signals.

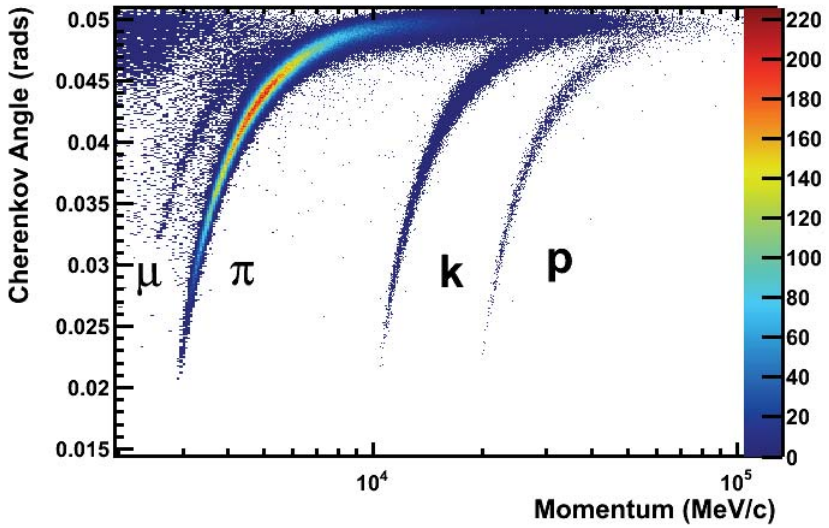


Figure 2.12: The relation between the Cherenkov angle and the momentum for different long live particles. When the momentum of the particle is well above its mass value, the Cherenkov angle saturates to $\frac{1}{n}$.

The optical path of the Cherenkov light for both RICHs is focused via spherical and flat mirrors in a region out of the spectrometer acceptance. The photons, in the wavelength range of 200-600 nm, are detected by 196 and 288 Hybrid Photon Detectors (HPDs) respectively in RICH1 and RICH2. These are surrounded by external μmetal shields, as shown in Figure 2.11. The spatial resolution obtained is 2.5 mm with a time resolution of 25 ns.

To identify a charged particle, the informations on its direction and the momentum are taken from the tracking system. One ring for each possible mass hypothesis is constructed and the likelihood of each test ring against the positions of the activated HPDs channels is evaluated. This RICH likelihood value, combined with the information from the calorimeters and the muon system, is associated to each track, allowing to define a particle identification criteria in offline analyses.

2.3.4 The Tracking System

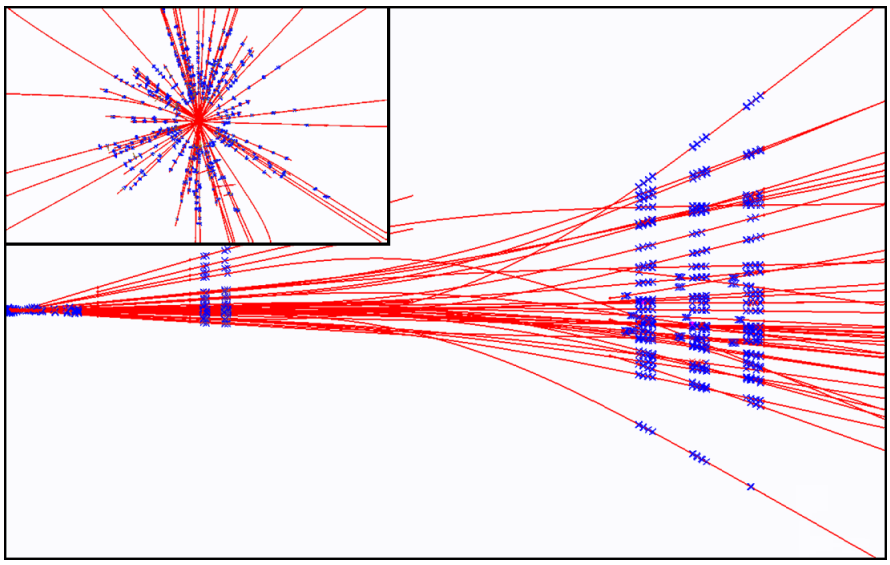


Figure 2.13: Display of the reconstructed tracks and assigned hits in an event in the x-z plane [100]. The insert shows a zoom into the VELO region in the x-y plane.

The tracking system is based on a set of silicon strip tracking stations (Figure 2.14). A station is placed upstream (TT) the magnet and three stations downstream (T1-T2-T3) the magnet and before RICH2. In this way informations before and after the magnetic deflection are used to evaluate the momentum and to select events with heavy-particle daughter tracks asking for a large transversal momentum component. Each TT stations have two different substructures: the IT [101] in the most inner part and the OT [102] surrounding it. The TT covers the full acceptance of LHCb and it is 150 cm wide and 130 cm high, while IT covers the area near the beam pipe and it is 120 cm wide and 40 cm high. TT and IT are together called the Silicon Tracker (ST) because they are both based on the same technology of silicon microstrips.

The four ST stations have a strip pitch of about 200 μm (giving a single hit resolution of 50 μm). In order to get a 3D reconstruction, each of the station includes four detection layers in an (x-u-v-x) arrangement with vertical strips in the first and the last layer and strips rotated by a stereo angle of -5° and +5° in the second and the third layer. The xuvx

geometry has been chosen to enhance resolution in the bending plane and it allows to solve ambiguities in case of two or more tracks.

The OT is a drift-time detector that uses as counting gas a mixture of Argon (70%) and CO_2 (30%). The gas mixture has been chosen to guarantee a fast drift-time, below 50 ns, and good aging properties. The detector modules are arranged in three stations and each station consists of four layers arranged in the same fashion of ST (see Figure 2.15). Each module is composed of two layers of 64 drift tubes each which are 2.4 m long with 4.9 mm inner diameter. In the longest modules, the straw tubes are split in the middle into two independent readout. The resolution obtained with the OT system is 200 μm on the x-axis.

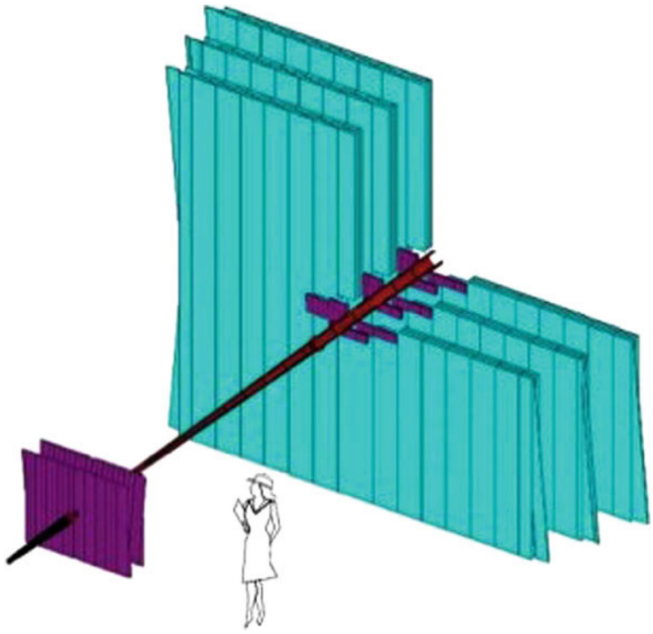


Figure 2.14: Layout of the LHCb tracking system. The silicon microstrips of the TT and the IT are colored in purple, the OT planes are shown in blue.

The classification of the reconstructed tracks in LHCb depends on which sub-detectors of the tracking system are crossed (see Figure 2.16):

- VELO tracks: only hits in the VELO. Particle of this kind are the backward and the large polar angle particles, which are rarely used for the physics studies but turn out to give an important contribution to the resolution of the PV.

- Upstream tracks: hits in the VELO and the TT. Particles with low momentum are subject to a stronger bending by the magnetic field and can escape from the detector before passing through the T stations.

- Long tracks: hits in the VELO, the TT and the T-stations. These are the most important tracks used for the LHCb analyses because their momentum can be measured with the best precision. In the $K^0_S \to \mu^+\mu^-$ analysis, which is the subject of this thesis, the resolution on the K^0_S mass reconstructed from the two muons is a very important parameter of the analysis. For this reason only the muons which correspond to long tracks are considered.

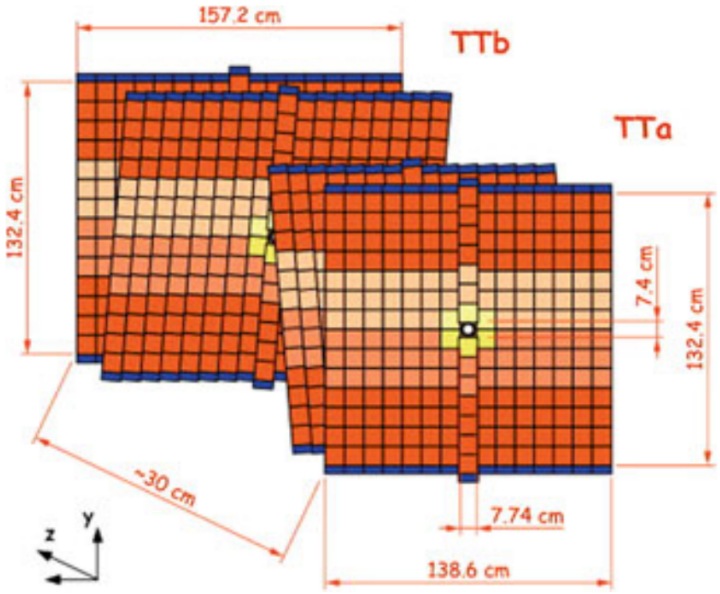


Figure 2.15: A view of the OT.

- Downstream tracks: the TT and the T-stations. Products from long-lived particles that decay between the VELO and the TT are a typical source of this category (e.g. daughters of K_s^0 and Λ). Compared to the long tracks, a lower resolution on momentum and on the impact parameter is achieved. Also, there is an higher misreconstructed track rate because of fewer hits in the tracking system.
- T tracks: only hits in the T-stations. Their use is often restricted to the calibration and detector studies and they are typically produced by secondary interactions.

2.3.4.1 Measure of the Momentum

The reconstruction process looks all along the tracking sub-detectors for hits which can be assigned to the same particle with high confidence. A fit of the hits is realized and a track is obtained with an associated χ^2 that states the quality of the track. High χ^2 can correspond to no real particle at all (mixture of hits from different particles), in this case the track is classified as a ghost.

The momentum of the charged particles is obtained from the detected tracks traversing the magnetic field. The mass of the decaying particles, which is an important parameter of the analysis, is obtained from the daughters tracks ($M = \frac{1}{c^2}\sqrt{(\sum_i E_i)^2 - |c\sum_i \overrightarrow{P_i}|^2}$ for i daughters tracks). That means that the momentum resolution needs to be as good as possible (the ultimate limits is due to multiple scattering). Decays of well known $c\bar{c}$, $b\bar{b}$ and Z_0 resonances are used to calibrate the momentum measurements.

The decay length of a particle is obtained from the distance between the PV and the SV associated at the particle itself. Using the decay length d, the momentum of the particle P, assuming the mass of the particle M, the time of flight Δt of the particle in its rest frame is $\Delta t = cMd/P$. An excellent proper decay time resolution allows to perform important physics studies, e.g. the time dependent CP asymmetries of B meson system. Typical values

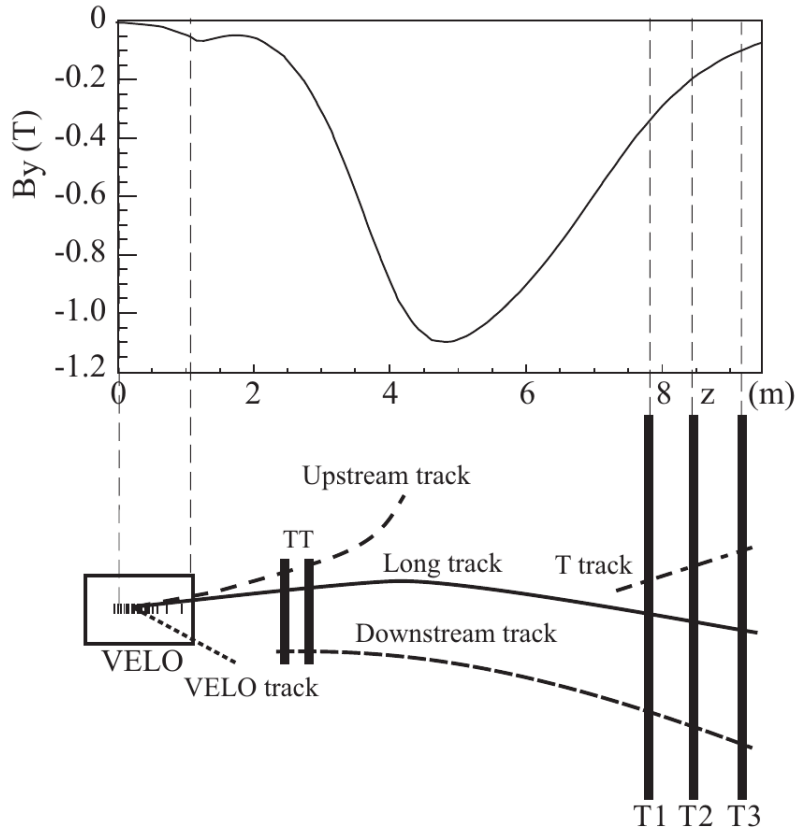


Figure 2.16: A schematic illustration of the various track types [100]: long, upstream, downstream, VELO and T tracks. For reference the main B-field component (B_y) is plotted above as a function of the z coordinate.

for the proper decay time resolution for B meson range between 40 and 50 fs, depending on the quality of the reconstruction process, to be compared with the B meson lifetime of around 1 ps.

With the interaction of the material, particles can deviate randomly from their trajectories (multiple scattering process). As the main difference between Monte Carlo (MC) and data comes from the material description and interaction, it is always preferred to obtain the efficiencies of the tracking system using real data.

2.3.5 Calorimeter System

The role of the LHCb calorimeter system [103] is to measure the transverse energy, the energy and the position of hadrons, electrons and photons candidates. It provides informations about the identification of electrons, photons and hadrons with their measured energy and position to the first trigger level (L0), allowing to select events with a candidate particle of high transverse energy.

The energy deposition in the material depends on the nature of the particle considered (see Figure 2.17). For this reason more substructures of the calorimeter detector are needed: the particles at first encounter the Scintillator Pad Detector (SPD), then a plane of lead and the Preshower detector (PS) from which the Electromagnetic Calorimeter (ECAL) and the Hadron Calorimeter (HCAL) follow.

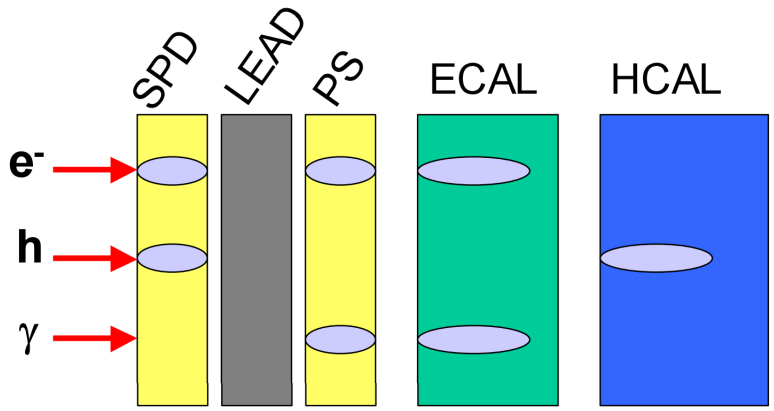


Figure 2.17: Signal deposited on the different parts of the calorimeter by an electron, a hadron and a photon allowing the calorimeter system to discriminate them.

The Readout of the entire calorimeter system is achieved with wavelength-shifting fibres, embedded in the scintillator tiles which are read out by Photo-Multipliers-Tubes (PMTs).

The SPD and PS detectors together consist of two layers of scintillator pads sandwiching a 15 mm Pb converter plate, with a sensitive area of 7.6m·6.2m. Both the SPD and PS are used to validate and cross-check any signals of the ECAL and HCAL, associating an input position to charged and neutral particles. The SPD is not sensible to photons while, thanks to the conversion of a photon in an electron pair in the Pb plate, the PS is able to detect them allowing to discriminate electromagnetic and hadronic showers and disentangle overlapping electromagnetic showers casted by different particles. The SPD is also used to count the multiplicity in the event and so rejects the events with high occupancy in the sub-detectors.

The ECAL is a sampling calorimeter of shashlik type composed of multiple and alternated layers of lead (2 mm thick) as absorber and scintillator tiles as active material (4 mm thick). Its thickness is 25 radiation lengths and the overall dimension are 7.8 m × 6.3 m. Finally, the HCAL is a sampling device with dimensions 8.4 m × 6.8 m. It is made of iron and scintillating tiles, as absorber and active material respectively. Its thickness is equal to 5.6 interaction lengths, which are not sufficient to ensure the whole containment of the hadronic shower, but are enough to achieve a reasonable measurement of the energy.

The energy resolution of ECAL and HCAL are described respectively by the following formulas:

$$\frac{\sigma_E}{E} = \frac{8.5\% - 9.5\%}{\sqrt{E(GeV)}} \oplus 0.8\%, \quad \frac{\sigma_E}{E} = \frac{(69 \pm 5)\%}{\sqrt{E(GeV)}} \oplus (9 \pm 2)\% \tag{2.1}$$

where E is expressed in GeV, the first term is the statistical error on the energy deposition and the second a constant contribution.

2.3.6 Muon Detector

The muon system [104] is the last stage in the LHCb detector. It consists of five tracking stations (see Figure 2.18) placed along the beam axis with the first station (M1) situated in front of the calorimeter preshower and the remaining four stations (M2, M3, M4 and M5) located downstream the calorimeter. The stations from M2 to M5 are interspersed with three iron filters 80 cm thick which are distanced of about 1.2 m. The total absorber-

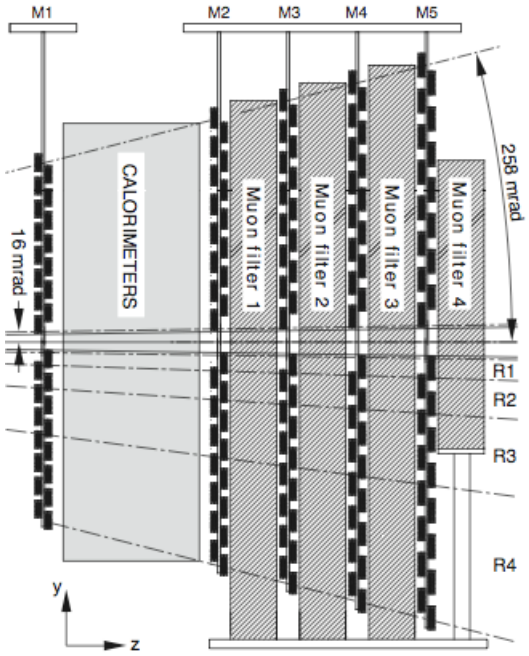


Figure 2.18: Scheme of the five muons stations labeled as M-station-number. The geometrical division in 4 region, reported in details in Figure 2.19, is shown as well as the acceptance in the polar angles.

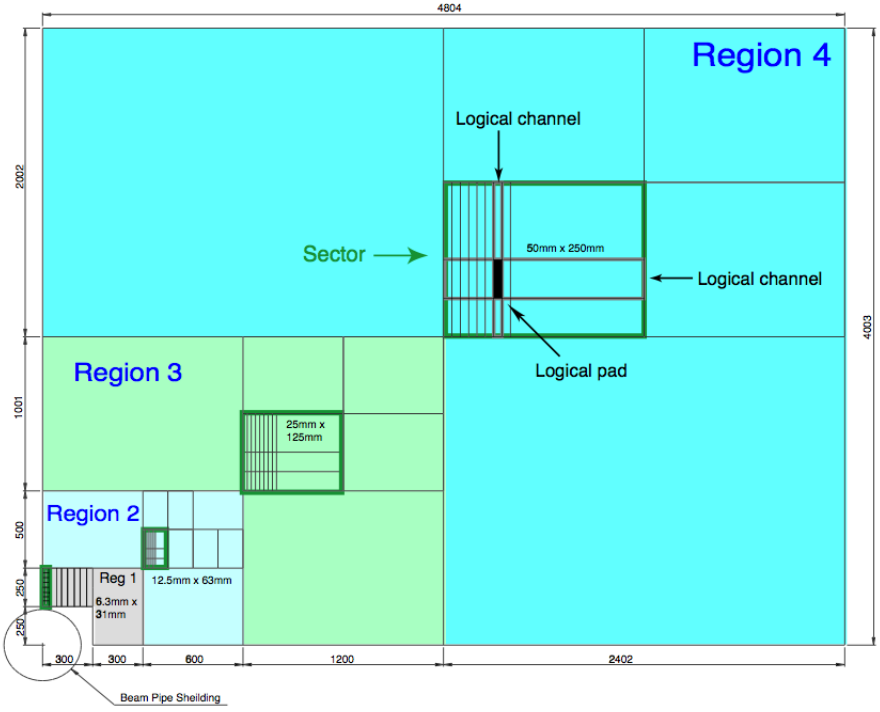


Figure 2.19: Schematic division of the four regions for one of the four quadrants of a muon station.

thickness is about 20 nuclear interaction lengths. The angular acceptance for the muon system in the bending (non-bending) plane is similar to the tracking system, acquiring tracks from 20 (16) mrad to 306 (258) mrad.

As shown in Figure 2.19, each station is structured in four quadrants subdivided in four regions R1, R2, R3 and R4. The logical pad size scales with the region in the ratio 1:2:4:8 as well as the region width; in this way the contribution to the PT resolution is approximately equal to the multiple scattering contribution. In fact, the multiple scattering in the absorber increases with the distance from the beam axis. Thanks to the scaling of the regions, a roughly constant occupancy through the detector is achieved, despite the large variation of the particle flux passing from the central part respect to the detector border.

Multi Wire Proportional Chambers (MWPCs), with a gas mixture of Ar:CO_2:CF_4 (40%:55%:5%), make up the muon system. The M1 station is under a high rate of particles (up to 500 kHz/cm^2 of charged particles), for this reason its R1 region uses different detector composed of three gas electron multipliers (GEM) with a gas mixture of Ar:CO_2:CF_4 (45%:15%:40%).

The M1 station has an important role in the transverse momentum measurement of the muon track in the hardware trigger stage (discussed in section 2.5.1) in order to require events with an high-PT muon. Indeed, using the informations from M1-M3 station, a good spatial resolution in the x direction allows to calculate the PT and the track direction of the muon with a resolution of about 20% for PT.

2.4 Particle Identification in LHCb

After the reconstruction of all the tracks an hypothesis on the nature of each particle associated to a track is done. This procedure is called particle identification (PID). The informations from the PID detectors (RICH1, RICH2, calorimeters and the muon system) are all used to state if the charged particles are pions, kaons, protons, muons or electrons.

Summary of the main feature for the PID in LHCb:

- **Electron and Neutral Particle Identification.** The main contribution is given by the ECAL detector, since neutral particles do not interact with the tracking system (e.g. the reconstruction of neutral pions and prompt photons in the calorimeter is essential for studies of many decays). The energy deposited in the ECAL is studied to build a likelihood $L^{CALO}(h)$. More details on electron and neutral particle identification (such as electron, γ and π^0) can be found in [105].

- **Hadron Identification.** The main contribution is given by the RICH detectors and the HCAL. The comparison between the circular rings observed in the phototube planes and the pattern expected under a given particle hypothesis, united with the informations from the energy deposition in the HCAL, allows to calculate the likelihood $L^{RICH}(h)$ [106] [107]. To calculate the performance of the PID, real data are used from independent and high statistics calibration samples. For example pions and kaons samples can be obtained from the decay $D^{+*} \to D^0(K^-\pi^+)\pi^+$ with an high purity.

- **Muon Identification.** MuonID uses the fact that the muons are the only particles able to go through the calorimeter and hit the muon system stations. The main source of misidentified muons are pions and hadrons decaying in flight to a muon and

a neutrino. In fact, when the kink of the muon with respect to the initial hadron is low, it becames very hard to distinguish between the hadron decaying in flight and a real muon. For this identification process an important role is carried out by both the second stage of the trigger (discussed in 2.5.2 and 2.5.3) and the offline analysis.

The minimum momentum of a muon required to pass the calorimeters and reach the M2 station is 3 GeV/c. Tracks are flagged as muons if they are associated to a certain number of hits in the five muon stations which depends on their momentum (see Table 2.1). The complete procedure is described in [108, 109, 110] and the main aspects are reported in the following.

Momentum range	**Muon stations**
3 GeV/c < p < 6 GeV/c	M2 + M3
6 GeV/c < p < 10 GeV/c	M2 + M3 + (M4 or M5)
p > 10 GeV/c	M2 + M3 + M4 + M5

Table 2.1: Muon stations (from 1 to 5) required to trigger the muon decision as a function of momentum range. The higher the momentum the stronger the requirement.

The muon identification is an important part of the analysis studied in this thesis and is here emphasized with more details respects to the hadron or electron identification.

The muon identification strategy can be divided in three steps (that create the IsMuon and a likelihood criteria which can be applied separately):

- Based on the penetration of the muons through the calorimeters and iron filters, a loose binary selection of muon candidates is created (called IsMuon, some details are given below).
- Computation of a likelihood for the muon and non-muon hypotheses, based on the pattern of hits around the extrapolation to the different muon stations of the charged particles trajectories reconstructed in the tracking system.
- Computation of a combined likelihood for the different particle hypotheses, including informations from the calorimeter and RICH systems.

The IsMuon binary selection provides high efficiency while reducing the misidentification probability of hadrons to the percent level.

The number of stations where a hit is found within a field of interest (FOI) defined around the track extrapolation is used to define IsMuon. The spatial size of the FOI is parametrized as a function of momentum and the region R of the 4 different stations downstream the calorimeter crossed by the track. The FOI sizes is optimized using a data-driven method on the two muons coming from $J/\psi \to \mu^+\mu^-$ decays. The data-driven method is called tag and probe and consists in selecting track pairs with invariant mass compatible with the J/ψ using and requiring tight muon identification on one track, called tag muon, and taking the other track as probe. The event is required to have been triggered indipendently on the probe track. A final sample of high purity probe muons without muon identification requirements is so obtained.

After the IsMuon requirement, the muon identification can be further improved with a selection that uses the logarithm of the ratio between the likelihoods for the muon and

non-muon (an hadron h) hypotheses, called $\Delta LL_{\mu/h}$. The likelihoods are computed as the cumulative probability distributions of the average squared distance significance D^2 of the hits in the muon chambers with respect to the linear extrapolation of the tracks from the tracking system. The average squared distance significance is defined as:

$$D^2 = \frac{1}{N}\sum_i \left\{ \left(\frac{x^i_{closest} - x^i_{track}}{pad^i_x} \right)^2 + \left(\frac{y^i_{closest} - y^i_{track}}{pad^i_y} \right)^2 \right\} \tag{2.2}$$

where the index i runs over the stations containing hits within the FOI, ($x^i_{closest}$, $y^i_{closest}$) are the coordinates of the closest hit to the track extrapolation point for each station, (x^i_{track}, y^i_{track}) and $pad^i_{x,y}$ correspond to one half of the pad sizes in the x, y directions. The total number of stations containing hits within their FOI is denoted by N.

For true muons is expected a D^2 distribution closer to zero than the other particles incorrectly selected by the IsMuon requirement.

The D^2 distribution for protons is used to calibrate the likelihood for the non-muon hypothesis because the other charged hadrons (pions or kaons), incorrectly selected by IsMuon, have a D^2 distribution with a component identical to the protons and a component very similar to the true muons, due to decays in flight before the calorimeter.

For each track (candidate), the likelihood number associated for the muon (or non-muon) hypothesis is then defined as the integral of the calibrated muon (or proton) D^2 probability density function from 0 to the measured D^2 value. The procedure end up with the creation of two likelihoods: $L^{MUON}(\mu)$ and $L^{MUON}(\text{non } \mu)$.

To increase the particle identification capabilities, a combined log-likelihood is then obtained by summing the logarithms of the likelihoods obtained using the muon system, the RICH and the calorimeter. For example for the muon identification at low momentum (below 5 GeV/c) the contribution from the RICH likelihood becomes important. With all the likelihood from the PID detectors a log-likelihood difference ($\Delta LL_{h/\pi}$) between a given particle hypothesis h (with $h = K$, p, μ or e) and the pion hypothesis is created. This global likelihood function is defined as:

$$\Delta LL_{h/K} = lnL(h) - lnL(\pi) = ln\left[\frac{L(h)}{L(\pi)}\right] \tag{2.3}$$

where L is the global likelihood that is formed multiplying the different sub-detector contributions.

2.5 LHCb Trigger

When the bunches of protons collide, particles are detected in LHCb on almost every collision. The main limitation to the number of acquired events is due to storage space constraints, so only the most interesting events have to be stored. A high efficiency trigger to minimize the loss of important physics events (in particular events where B and D mesons are present) is needed. Two main signatures allowing the identification of these kinds of events are the high transverse momenta and the non-zero impact parameters (IP).

The trigger is divided in three stages in order to reduce the rate of visible interactions to a value of about 4.5 kHz. The total output of the trigger in 2011 has been about 3 kHz, exceeding the designed value by 50%. In 2012, the trigger output rate has been increased to 4.5 kHz. This represents the experimental bandwidth limit for writing events to offline

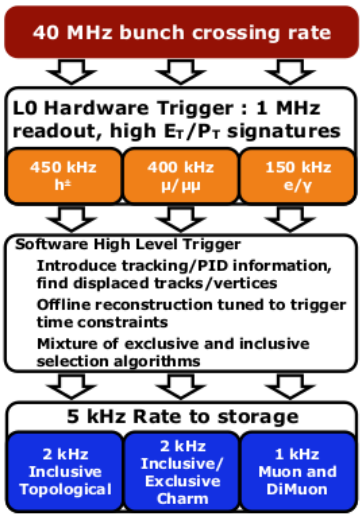


Figure 2.20: The LHCb trigger is divided into two main levels: the hardware (L0) and the software (HLT), the latter is further splitted into two steps (HLT1-HLT2). In the picture a schematic vision of the LHCb trigger is provided.

storage. A scheme of the LHCb trigger is shown in Figure 2.20. The three levels of the LHCb trigger are: Level-zero (L0, hardware trigger), High Level Trigger One (HLT1, first software level), High Level Trigger Two (HLT2, second software level). The reconstruction of the events needs to be as fast as possible in the trigger phase, so different algorithms are implemented in the online and offline event reconstruction.

For only about 30% of the time, the LHC beams are in a stable condition while the rest of the time is used for injection or tuning. During the idle time, the hard disk of the Event Filter Farm nodes are used to store about 20% of the accepted events by L0 trigger. The events are then analysed during the inter-fill gaps [111]. With this approach, called deferred trigger, it was possible to increase the bandwidth of the 2012 with respect to 2011. Indeed, the PT threshold on the reconstructed tracks was lowered from 500 MeV/c in 2011 to 300 MeV/c in 2012 and the dimuon mass restriction in HLT2 that was cutting off the K^0_S was removed in 2012.

In the following subsection (2.5.1, 2.5.2, 2.5.3) a short description of the three trigger stages is given.

2.5.1 LO

The L0 is an hardware trigger synchronous with the LHC clock and it is integrated in the front-end electronics of the detector. The buffer time of the read-out chips is 4 μs, so decisions must be provided with this latency. The L0 deals with 40 MHz input rate and reduces the output to 1 MHz which is the maximum rate possible for the readout electronics. At this rate, the whole detector can be read out.

A good experimental signature for heavy particles consists of daughter particles having large PT compared to the background of low energy QCD proceses. Indeed, the L0 trigger asks for large PT objects determining in a fast event reconstruction the highest ET candidate for hadrons, electrons and photons, and the highest PT for the muons. Only a part of the whole detector is used to take decisions in L0, in particular the calorimeter and muon systems informations are used.

Using ECAL and HCAL it is possible to select tracks with high transverse energy (ET) which can correspond to photons, electrons, neutral pions and hadrons candidates. For particle identification also the information from the PS and SPD detector are used. For an estimation of the charged track multiplicity in the event the number of hits in the SPD is considered.

Muon candidates are selected with the request that the track must be present in all the five muon stations, roughly pointing to the interaction vertex. Thanks to the position of M1 before the calorimeters, the effect from the multiple scattering is reduced improving the quality of the momentum measurement. An estimation of the momentum is provided assuming the muon track originate at the interaction point. The momentum resolution obtained is $\delta p/p \sim 20\%$, while for PT is about 25%. The evaluation of the muon PT is obtained using only the first two stations M1 and M2. High PT are required for the muon or di-muon candidates (L0Muon and L0DiMuon lines): 1760MeV/c for the former and 1600MeV/c for the latter. The two highest PT muon tracks are used for the decision.

The time required by the reconstruction algorithms increases with the number of tracks (and so hits) that are present in the event. To avoid these situations, events that contain too many hits or tracks in some specific sub-detectors are processed. These cuts applied on the events are called Global Event Cuts (GEC). For example in the L0Muon case it is required that the event has less than 600 SPD hits while for the L0DiMuon case the threshold is 900.

2.5.2 HLT1

The LHCb High level trigger (HLT1 and HLT2) is a software that processes, in parallel, the output of the L0 on more than 20,000 CPU cores on a PC farm close to the detector. HLT1 reduces its input rate (output of the L0) to about 40/80 kHz (in 2011/2012) to allow the HLT2 to make a complete reconstruction of events selected by the HLT1.

A partial reconstruction of the data is performed at this level to be able to have access to higher level informations such as the reconstructed tracks. An estimation of the position of PVs is obtained thanks to HLT1 particles reconstruction in the VELO. Several trigger decision lines are run in parallel, with a total decision time per event of the order of 10-15 ms, each one makes a decision based on the presence of one or two tracks with a good track fit quality (such as two muons or a single hadron) that match certain criteria, for example having an high PT or an high impact parameter.

Only if the event has been triggered by L0Muon or L0DiMuon lines, the muon HLT1 triggers are executed. A distinction based on low and high particles momentum and the invariant mass of the composed objects of di-muon candidates is provided. Using the informations coming from the ECAL, electron and photon trigger lines are created.

The GEC cut applied at HLT1 are upper cuts in the tracking system hits. For example, it is required that the number of the VELO hits and the IT hits are less then 10000 and 3000, respectively.

2.5.3 HLT2

The HLT2 stage runs over events which have passed the HLT1 and stores its output as Raw data. The trigger rate output of HLT2 is about 3/4.5 kHz (2011/2012). In HLT2, due to the lower input rate with respect to L0 and HLT1, the reconstruction and the particle identification algorithms performed are more similar to the one performed offline.

For example, the HLT2 muon lines are similar to those run in the HLT1, but with a more refined and complete track reconstruction for all tracks with a PT of at least 300 MeV/c. Thanks to the more complete reconstruction process, the HLT2 is able to access the full event information, allowing the selection of SVs and composed particles with the presence of charm and beauty hadron candidates.

In HLT2 the main signatures looked for in the event are:

- Displaced vertices formed by two, three or four tracks
- Single muons with high PT and IP
- Pairs of leptons (muons or electrons) with moderate PT and IP and forming a good vertex
- A photon cluster in the calorimeter and a pair of tracks forming a secondary vertex, aiming for radiative B decays ($B \to X\gamma$)
- Several exclusive selections of B and D channels

The only GEC imposed at HLT2 is to have 350 or less VELO tracks.

To associate an event to a trigger configuration, an hexadecimal word called Trigger Configuration Key (TCK) that identifies the set of trigger decisions (e.g. trigger algorithms run in L0, in HLT1 and in HLT2, the selection criteria, etc..) is used. TCKs is 32-bit long, with the first 16 bits reserved for the L0 configuration and the second 16 bits for the HLT. A single change in a cut decision will create a new TCK, but for most of the relevant physics trigger lines the changes are minimal or not present.

2.5.4 Trigger Efficiency

Candidates for a given physical signal, as the $K^0_S \to \mu^+\mu^-$ decay, can be classified according to how they were triggered as

TOS Triggered On Signal: when the signal process under study is able, alone, to trigger the event

TIS Triggered Independent of Signal: when the event would also have been triggered without the studied signal process.

An event can be classified as both TIS and TOS simultaneously (TIS&TOS) or neither TIS nor TOS. All the informations needed to classify the event are recorded by the trigger system.

The trigger efficiency for the events that are selected by the selection procedure (see section 2.6.1) or by the user custom selection is defined as $\epsilon_{trg} = N_{TRG\&SEL}/N_{SEL}$, where $N_{TRG\&SEL}$ is the number of triggered and selected events and N_{SEL} is the number of events that would be selected offline if the trigger was able to acquire all the events. A general data-driven method can be used to extrapolate the trigger performance with respect to a defined process (e.g. $K^0_S \to \mu^+\mu^-$) from the data (avoiding the possible discrepancies between the reality and the simulation).

If an analysis relies only on TOS events (as it is often the case but not for the analysis performed in this thesis, $K^0_S \to \mu^+\mu^-$) then the TOS requirement can be considered a part of the offline selection.

If it is possible to consider the TIS efficiency as independent from the selection criteria applied on the signal studied, then the total trigger efficiency can be obtained with $N_{TIS\&TOS}/N_{TIS}$. Using this definition it is possible to measure the trigger line efficiency using real data, thus without relying on the accuracy of the simulation.

2.6 LHCb Software Framework

All the LHCb software environment is based upon the full Object Oriented Gaudi framework [112]. The analysis, performed into the Gaudi architecture, can be splitted in problems involving only a single protons collision that is considered independent from the previous and the next one except for effects due to dead-times, treated statistically. Because of the non mixing events procedure, the workload can be easily split on many computing nodes in exactly the same way for all the purposes.

The main LHCb software tools are:

Gauss Both the event generation and detector simulation are handled by Gauss [113]. A tuned version of Pythia [114] is used to generate the *pp* interactions, while the decays of the particles is described by the EvtGen package [115]. The Geant4 toolkit [116] simulates the interaction of particles in the detector. To save time and increase the production efficiency, cuts in the acceptance can be applied at the generator level (e.g. for $K^0_S \to \mu^+\mu^-$ MC a cut in the decay position of K^0_S inside the VELO is applied because these events are the only one used in the analysis).

Boole The Boole software package [117] is in charge to simulate each sub-detector response. It recreates the digitization of the energy deposited in the LHCb detector active material. At this stage, the hardware trigger (L0) is also emulated. The output format of Boole is exactly that used in real raw data files, allowing the following software chain to be the same for MC and real data.

Moore The HLT1 and HLT2 are software triggers which are implemented by the Moore package [118] for both the online system and offline simulated events. The trigger settings are defined via the unique TCK as explained in subsection 2.5.3. To ensure a correct reproduction of the trigger performance in the simulated events, each TCK must be processed with a specific version of the Moore software (exactly the same that has been run when the real data was taken).

Brunel Brunel [119] is the reconstruction program that creates physics object starting from raw detector hits. The physics objects consist of charged tracks, photons, calorimeter clusters and Cherenkov rings. Also informations about the tracks are obtained, such as PID likelihoods and track quality.

DaVinci The last reconstruction step is DaVinci [120]. With this program it is possible to create composed objects (e.g. decay chain such $B^+ \to J/\psi K^+$ with $J/\psi \to \mu^+\mu^-$) and to perform candidate selection. Selected candidates can be stored in different file formats useful for statistical analyses (e.g. histograms or ntuples) or visualization (using the Panoramix software package [121]). DaVinci is also used for the selection procedure (described in 2.6.1), in which case an output format readable by DaVinci itself is used.

In this thesis, I have used the DaVinci software to obtain the ntuples with the informations of the particles involved in the $K_S^0 \to \mu^+\mu^-$ decay. This informations are: kinematic variables, particle identification variables, the fired trigger lines, tracks fit results such as the quality of the fit, the time of flight distance and parameters to state the isolation of the vertex. For the MC sample, the MC truth information is added. The ntuples are created requiring that the events have passed the stripping line to select $K_S^0 \to \mu^+\mu^-$ events, whose selection cuts are described in section 3.3 both for $K_S^0 \to \mu^+\mu^-$ and $K_S^0 \to \pi^+\pi^-$ candidates.

Each contribution to the code by the members of the analysis group is merged to have a common code for the creation of the ntuples with all the informations needed for the analysis studies.

2.6.1 Monte Carlo Simulation and Stripping Process

MC simulated events are a very important piece for the studies of the detector and physics analyses. The advantage of the simulated events is that the true full knowledge is known in advance, so it is possible to train the tools that are later used on data or to obtain parameters very difficult to be extracted directly from real data (e.g. the calculation of the reconstruction efficiency for a certain process).

To simulate the physics events it is necessary that the MC and real data are treated in the same way. The simulation of all the imperfections of the detector (which could change during the experiment) is the step that creates the larger discrepancy between MC and real data. For this reason it is always better to use a data-driven method (that means to evaluate the values needed from real data) to obtain the numbers used in the physics analysis such as trigger efficiencies or PID efficiencies, while using the MC to validate the procedure.

The simulation procedure is divided into several steps: the generation of the particles produced in proton-proton interaction, the propagation of the particles in the detector and the whole detector response to convert the event into simulates readout channel responses that mimic the real recorded events. The last mentioned procedure is called digitization. At this point the data format is the same of that of real data, so the same algorithms can be run in both the simulated and real data. The data retained by the trigger process are saved on tape as Raw data, which are spread world wide to store copies in large computing centers connected through the LHC Computing Grid (LCG). In this format only hits and raw information from the trigger and sub-detectors are stored. The reconstruction runs in parallel on the LCG with different machines running on different events, creating output files that are stored on the LCG computing centers.

Concerning the reconstruction it has to be pointed out that only the particles, called basic particles, having a flight distance larger than or comparable to the detector size are directly reconstructed from the detector hits. The π^0 is an exception because even if it is short lived it produces two collinear photons that are hard to be distinguished from a single photon, therefore the two photon system, is considered a basic particle. From the basic particles other common candidates are created using the specified decay chain of the particle (e.g. K_S^0 using $K_S^0 \to \pi^+\pi^-$ or J/ψ resonances from $J/\psi \to \mu^+\mu^-$).

To filter data relevant for the analyses and to save space on disk, which is a critical issues in the experiments at LHC, a last selection, named Stripping, is performed. The Stripping project is composed by a collection of many algorithms, called Stripping lines, that search and create candidates combining tracks and/or neutral objects. Selection cuts on

the candidate decay chain are applied to increase the purity of the Stripping lines output. In general each physics analysis has its own dedicated Stripping lines. The stripping procedure can be applied also to MC and real datasets, to evaluate its efficiency.

The complete flow software is summarized in Figure 2.21. The left arm represents the simulated data flow, while the right arm represents the real data processing.

2.6.2 The LHC Computing Grid

The grid computing is an hardware and software infrastructure that consists in a collection of computer resources from multiple locations, which can be seen as a special type of parallel computing connected to a network by a conventional network interface. The grid can be thought as a distributed system with non-interactive workloads that involve a large number of files. This arrangement is thus well-suited for applications in which multiple parallel computations can take place independently, without the need to communicate intermediate results between processors.

A grid computing is established, for example, in situations where the computing power required is too large to be affordable by a single institution or country. The multi-location characteristic provides some advantages in the availability and reliability of the data. Indeed, with the multi-location structure the data will be still accessed even if a whole computing center gets off-line, because of the replicas of the data in different locations.

The LCG [122, 123], developed to support experiments using the CERN Large Hadron Collider, is an international collaborative project that consists of a grid-based computer network infrastructure. As of 2012, the LHC Computing Grid had become the world's largest computing grid comprising over 170 computing facilities in a world wide network across 36 countries. The aim of the LCG is to handle the prodigious volume of data produced, which are more then an order of magnitude higher than what was available to previous particle physics experiments. The order of magnitude of the data produced by the LHC is approximately of 10 petabytes per years. The Tier 0 is the CERN computer center that sends the data, using a dedicated 10 Gbit/s connection, to eleven Tier 1 academic institutions in Europe, Asia and North America. Each Tier 1 receives only a specific subsets of raw data to provide a backup repository for CERN Tier 0. Connected to the Tier 1, there are more than 150 Tier 2 institutions. The primary configuration for the computers used in the grid is based on Scientific Linux.

A dedicated middleware, called DIRAC, has been developed by LHCb for the production and user jobs running on heterogeneous architecture including the LCG computing center. To simplify the user job management, a further layer, called Ganga and implemented in Python, has been developed to configure the jobs to be run on the LCG. Ganga can easily switch between test on a local batch system and large-scale processing on Grid resources.

For the creation of the ntuple used in this thesis, both for MC and real data, I have used the Ganga interface to configure and run jobs in the LCG system. In this way the large amount of data could be analyzed with a reasonable effective time.

2.7 2011 and 2012 Experimental Differences Conditions

In this section I want to briefly outline the main different experimental conditions between the 2011 and 2012 LHCb runs that are important for the $K^0_S \to \mu^+\mu^-$ analysis. These differences are:

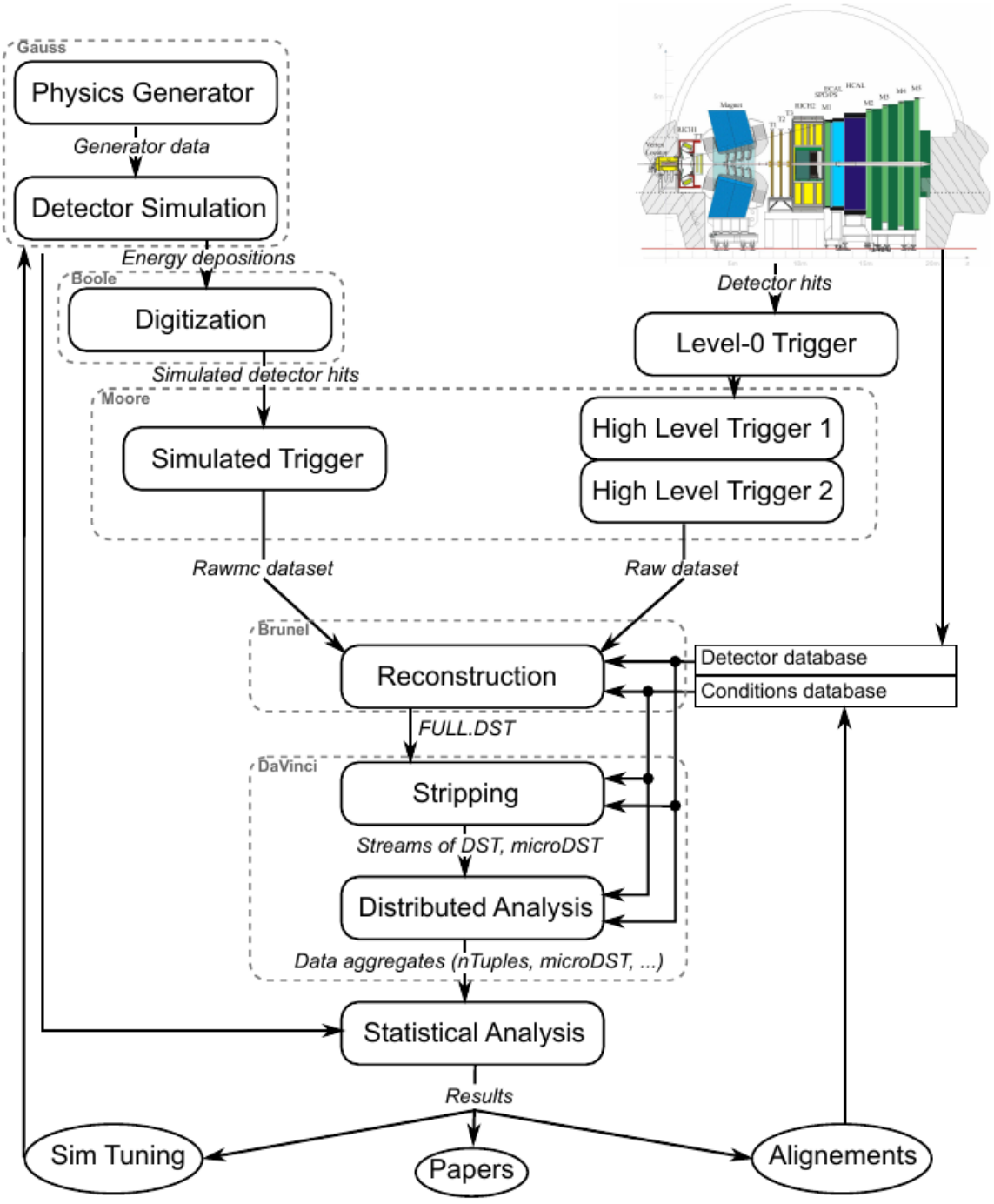


Figure 2.21: Data flow in LHCb for MC and real data. After the HLT stage the flow for simulated and real data is the same. The blocks enclosed in gray dashed boxes are implemented using the Gaudi architecture.

- The integrated luminosity for the 2012 run (2.082 fb^{-1}) is about double than the one collected in the 2011 run (1.107 fb^{-1}), so a factor of 2 in the $K_S^0 \to \mu^+\mu^-$ statistics is expected.

- The trigger output bandwidth was increased in the 2012 as described in section 2.5, allowing to loosen the trigger selection cuts and enlarge the LHCb physics program. In particular, for the $K_S^0 \to \mu^+\mu^-$ decay an increase in the trigger efficiency can be expected by the removal of the cut on the di-muon mother mass in the HLT2DiMuonLowMass line.

- The capability to maintaining the instantaneous luminosity at a fixed value was introduced in 2012, allowing to have a better knowledge on the unbiased trigger rate whose acquisition frequency is regulated depending on the value of the instantaneous luminosity.

- The aging and radiation damage of the detector after the 2011 run is not expected to have a significance effect on the detector performances, assuring the same reconstruction capabilities for the 2011 and 2012 data.

- The LHCb software algorithms for reconstruction and particle identification are under constant improvement (but of course the main structure is maintained).

Chapter 3

The Search for $K^0_S \to \mu^+\mu^-$

In this chapter an overview of the $K^0_S \to \mu^+\mu^-$ experimental search at LHCb is given. An illustration of the analysis scheme, as also the data and MC samples used are introduced in section 3.2, while the sections 3.3 is devoted to the description of the initial selections applied to both $K^0_S \to \mu^+\mu^-$ and its normalization channel $K^0_S \to \pi^+\pi^-$.

3.1 Kaon Physics at LHCb

The yet unobserved decay $K^0_S \to \mu^+\mu^-$ is described in the Standard Model as a Flavor Changing Neutral Current (FCNC) with a BR expected to be $(5.0 \pm 1.5) \cdot 10^{-12}$. Before LHCb, the most sensitive search for the $K^0_S \to \mu^+\mu^-$ was performed at the CERN PS in 1973 obtaining an upper limit of $3.1 \cdot 10^{-7}$ at 90% confidence level (CL).

Today the limit is dominated by the result obtained by LHCb with the 2011 data that correspond to an integrated luminosity of 1.0 fb^{-1} of p-p collisions at$\sqrt{s} = 7$ TeV. From them an upper limit of $11(9) \cdot 10^{-9}$ at 95(90)% CL was achieved, still above the SM expectation by more than a factor 10^3. This limit is below the previous measurement by more than a factor of thirty, proving the capabilities of LHCb to contribute to the physics of rare kaon decay.

Indeed, although the LHCb experiment is optimized for the decays in the heavy quark sector (b and c hadrons), the LHC is also a kaon factory, with about 10^{13} K^0 produced within the LHCb acceptance per fb^{-1} of integrated luminosity and about 40% of those kaons decaying within the length of the VELO.

A summary of the situation for the upper limit of $K^0_S \to \mu^+\mu^-$ rare decay is reported in Figure 3.1, where the BR range expected if new physics contributions (such as new light scalars) are present is highlighted. Indeed, enhancements up to one order of magnitude above the SM level would be well compatible with the present bounds from other $\Delta S = 1$ FCNC processes, while enhancements above 10^{-10} are less likely.

The data used in the LHCb analysis were only those acquired during the 2011 campaign, while there are still 2 fb^{-1} of integrated luminosity of proton-proton collision at a center-of-mass energy of $\sqrt{s} = 8$ TeV, acquired by LHCb in the 2012, to be analyzed. The new analysis on the data acquired in the 2012 is the subject of this thesis, for which it is expected to achieve a better sensibility as explained in section 2.7.

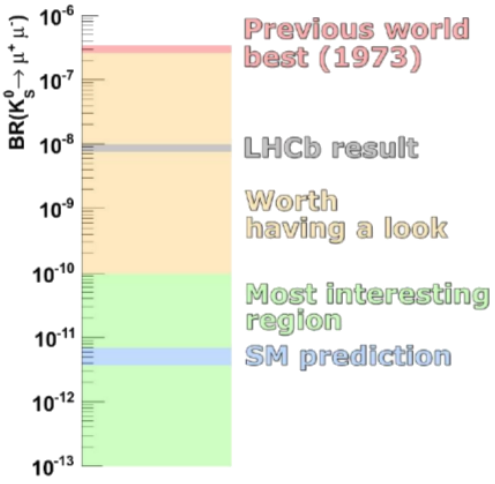


Figure 3.1: $\mathcal{B}(K^0_S \to \mu^+\mu^-)$ in SM is expected to be $(5.0 \pm 1.5) \cdot 10^{-12}$. The pre-LHCb world best upper limit $\mathcal{B}(K^0_S \to \mu^+\mu^-) < 3.1 \cdot 10^{-7}$ (PS at CERN, 1973) [124] and LHCb sets an upper limit $\mathcal{B}(K^0_S \to \mu^+\mu^-) < 9 \cdot 10^{-9}$ at 90%CL with the 2011 data. Still there are many orders of magnitude to search for NP, although the most interesting region is below 10^{-10}.

3.2 Analysis Scheme

The analysis is performed on 2 fb^{-1} of 2012 data not merged with the already analyzed 1 fb^{-1} 2011 data because of the different trigger condition. The result with the 2011 data will be combined in the final computation of the upper limit. The main structure of the analysis is the same adopted for the 2011 data and the scheme to select the $K^0_S \to \mu^+\mu^-$ candidates is represented in Figure 3.2.

In the LHCb L0 trigger, the nominal threshold on the transverse momentum for the muons is of the order of 1 GeV/c, allowing to cutoff the hadron background while maintaining high trigger efficiency for the muons coming from the b-quark decays (e.g. in the $B^0_s \to \mu^+\mu^-$ situation the trigger efficiency is at the level of 90%). This relative low threshold for the b-quark physics results into a very high cut level for the muons coming from the K^0_S decay, which spectrum for the offline selected events before the trigger request, taken from MC events, is shown in Figure 3.3. The same situation happens for the threshold in the di-muon reconstructed transverse momentum.

The critical point of the analysis is the low trigger efficiency due to the low momentum and transverse momentum of the particles involved in the decay. For this reason, there is not a single trigger selection with high efficiency and so, to achieve a statistics as big as possible, several trigger selections are investigated to search for $K^0_S \to \mu^+\mu^-$ candidates. The analysis is performed separately for each different trigger paths included, due to major differences in the signal and background properties, which will be explained in the next sections.

Since handling all data acquired by LHCb would be too demanding in terms of informatic resources, the data sample available to offline analysis is reduced by about a factor 20 thorough a filtering step called stripping (the stripping lines concept was introduced in section 2.6.1). In this way only the events that have passed this preliminary selection are used in the various analysis, and this is the case for the $K^0_S \to \mu^+\mu^-$ too. Indeed, the $K^0_S \to \mu^+\mu^-$ candidates pass a first dedicated selection implemented as a stripping line, requiring a detached SV with a reconstructed momentum pointing to one PV in the event, asking for an invariant mass in the region of the K^0_S mass ([492, 504.] MeV/c^2). The stripping line is based on simple cuts on some key variables: the mentioned ones, and variables related to reconstruction quality, both for the muons as for the total decay process.

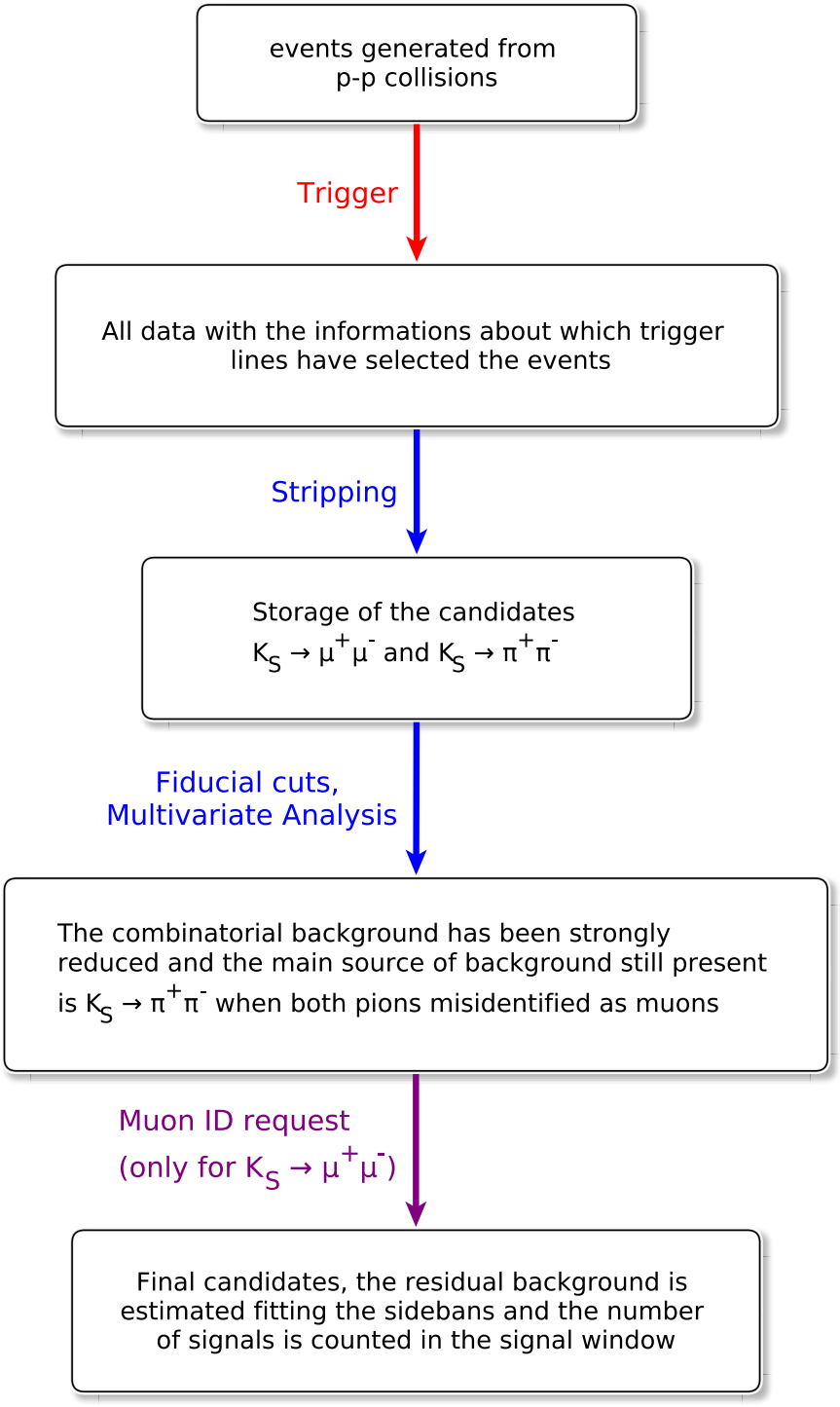


Figure 3.2: Data flow to search for $K_S^0 \to \mu^+\mu^-$ (signal) and $K_S^0 \to \pi^+\pi^-$ (normalization channel) candidates. The most interesting events are acquired via the trigger and then the $K_S^0 \to \mu^+\mu^-/K_S^0 \to \pi^+\pi^-$ candidates are created performing the selection cuts called stripping. After this selection, some fiducial cuts and a multivariate analysis are applied to reduce the background contribution. Finally, a specific request of muon identification allows to reduce the contamination coming from pions misidentified as muons. The calculation of the efficiencies for the selection of the candidates is divided into three parts which correspond to the process described by different colors in the picture.

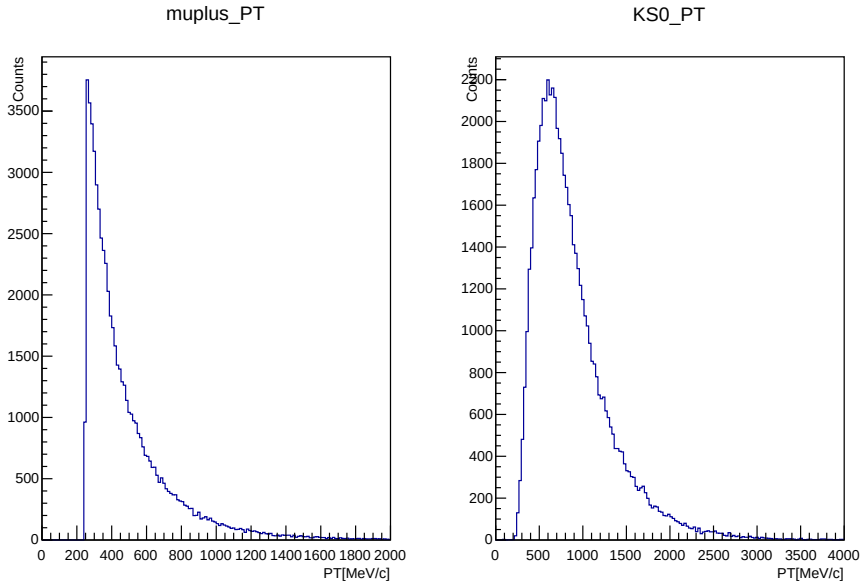


Figure 3.3: The transverse momentum spectra for the muons (left) and the K^0_S in the MC sample of the $K^0_S \to \mu^+\mu^-$ process at the generation level that pass the selection phase, but before the trigger request and so without any trigger bias. The nominal threshold at LHCb for the muons are of the order of 1 GeV/c on PT which results in a very low efficiency for the $K^0_S \to \mu^+\mu^-$ decay.

This stripping line was realized for the 2011 analysis trying to optimize the selection of candidates to achieve an high efficiency on the events that are more likely to pass the LHCb trigger. This efficiency is measured to be of about 19%/20% for the 2012 events respectively for $K^0_S \to \mu^+\mu^-$ and $K^0_S \to \pi^+\pi^-$ (numbers obtained from the MC) and being the PT cut of 250 MeV/c the dominant inefficiency source. Due to the low momentum of the muons the reconstruction and selection efficiencies is low (few percentage) respect, for example, to muons coming from J/ψ or B semileptonic decays. This selection is also able to reduce all the possible background contribution. Here an introduction to the the main backgrounds is given to better understand the decisions taken in the analysis that are explained in the rest of this chapter. A component of the background is the combinatorial background, which, after the trigger and selection procedure, is further rejected with a multivariate analysis (MVA). The mass distribution for these candidates is expected to be non-peaking, so that the residual background in the signal region can be estimated from sidebands.

The other relevant source of background is the $K^0_S \to \pi^+\pi^-$ decay ($BR_{K^0_S\to\pi+\pi-} = 69.2\% \gg BR^{SM}_{K^0_S\to\mu+\mu-} = 5{\cdot}10^{-12}$), when both pions are misidentified as muons. It has to be pointed out that the misidentification is substantially due to the case when the pions decay in flight into muons and the corresponding neutrinos. For this background, the invariant mass of the kaon candidate is underestimated on average by ~ 40 MeV/c^2 because of the wrong mass assignment to the daughter particles produced in the decay (the difference in mass between muons and pions is not negligible when compared to the K^0_S one). To reduce this contribution, only the muon tracks that give the best mass resolution are used, that is only long track muons (the track categories have been defined in section 2.3.4) for which the K^0_S mass resolution is about 4 MeV/c^2, 10 times smaller than the shift length. Thus, only the far tail of the $K^0_S \to \pi^+\pi^-$ distribution is expected to enter into the signal region.

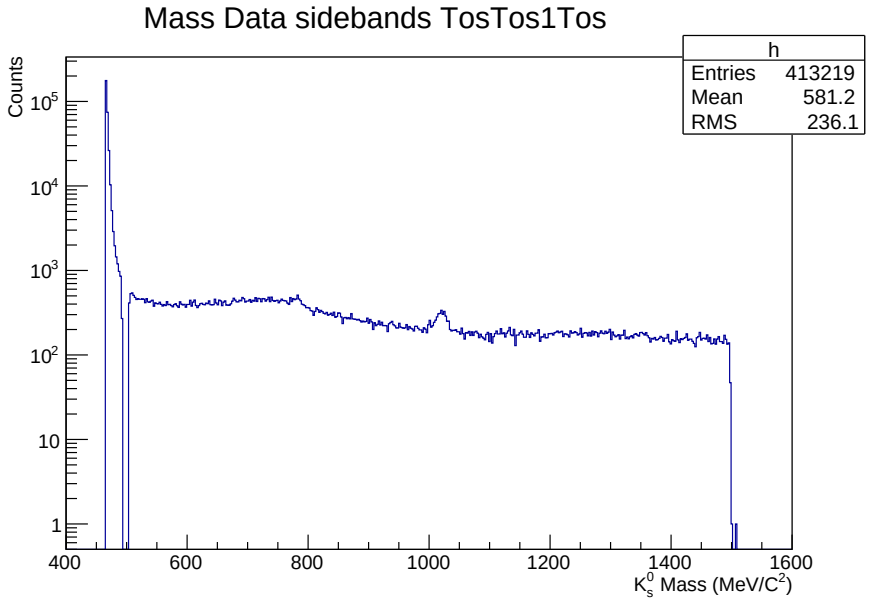


Figure 3.4: Mass distribution of reconstructed and selected $K_S^0 \to \mu^+\mu^-$ candidates in real data. The signal region [492, 504] MeV/c^2 is blinded at this point of the analysis. Region above the mass window is composed mainly by combinatorial background, while $K_S^0 \to \pi^+\pi^-$ with pions identified as muons end up in the leftside mass region. Also the peak of the $\phi \to \mu^+\mu^-$ is clear at $\sim$ 1050 MeV/c^2. Near 800 MeV/c^2 there is a small peak that can be associated to the decay $\rho \to \pi^+\pi^-$ with pions misidentified with muons, no clear peak is seen because of the large decay width ($\sim$ 150 MeV/c^2) of the ρ.

The contributions of both the main backgrounds to the $K_S^0 \to \mu^+\mu^-$ decay process can be seen in Figure 3.4, where the signal region [492, 504] MeV/c^2 is blinded. The right sideband is substantially composed of combinatorial background and in the left sideband it is possible to see the tail of the $K_S^0 \to \pi^+\pi^-$ decay when the mass of the muons is assigned to the pions. In the picture the $K_S^0 \to \mu^+\mu^-$ candidates after the stripping and fiducial cuts are shown. The muon identification request has not been applied yet to allow the study of the backgrounds and to do not discriminate the $K_S^0 \to \mu^+\mu^-$ and $K_S^0 \to \pi^+\pi^-$ sample at this stage of the analysis (before the MVA, as can be see from Figure 3.2). The other distinctive feature of the LHCb detector, the particle identification capabilities, can be exploited to identify the muons from the pions.

After the application of a loose cut on the MVA output, the signal candidates, and so the analysis, are divided in ten MVA bins which have different sensibility.

To estimate the background contribution, the mass sidebands are fitted and the number of background events that enter into the signal region is predicted, for each MVA bin. After the removal of the blinding of the signal region, the candidate events will be counted without using the shape of the distribution within the window, that can only be predicted using simulation for the $\mu^+\mu^-$ mode.

To convert the yield of selected K_S^0 decays into a branching fraction one needs to normalize the number of counted events to a known process. Due to its abundance, similar topology and well known branching fraction, the $K_S^0 \to \pi^+\pi^-$ decay is the ideal normalization mode. The same selection cuts, including requirements on the reconstruction quality and on kinematic variables, are applied to the signal and normalization channel mode,

except for the μ identification criteria. The formula to obtain the BR of $K_S^0 \to \mu^+\mu^-$ is

$$\frac{\mathcal{B}(K_S^0 \to \mu^+\mu^-)}{\mathcal{B}(K_S^0 \to \pi^+\pi^-)} = \frac{\epsilon_{\pi\pi}}{\epsilon_{\mu\mu}} \frac{N_{K_S^0 \to \mu^+\mu^-}}{N_{K_S^0 \to \pi^+\pi^-}} \tag{3.1}$$

where $\epsilon_{\pi\pi}$ and $\epsilon_{\mu\mu}$ are the efficiencies for the $K_S^0 \to \pi^+\pi^-$ and $K_S^0 \to \mu^+\mu^-$ channel respectively. They include the reconstruction, the selection and the trigger process, with the particle identification efficiency only taken into account for the $\mu\mu$ case. Data-driven methods are used as much as possible to determine the efficiencies, while simulated events are mainly used for the validation of the analysis procedure. For example, real muons from $J/\psi \to \mu^+\mu^-$ are used to get the efficiency of the muon identification and of the trigger.

If no events are found in the signal region, a limit is computed using the CLs method [125], based on the background and signal expectations and the observed events.

Hereafter the contributions I have provided so far on the $K_S^0 \to \mu^+\mu^-$ analysis are illustrated, but also the other part of the analysis, that I have not directly elaborated, will be briefly described in Chapter 6 to do not lose the general structure of the analysis. The first task of my thesis was the study of the best trigger lines categories to be included in the analysis. After this, I was in charge to implemented the MVA to discriminate the combinatorial background contribution from the $K_S^0 \to \mu^+\mu^-/K_S^0 \to \pi^+\pi^-$ decay. These tasks were implemented using the LHCb software and the programming languages as C++ and python, as also the ROOT software [126].

3.2.1 2012 Monte Carlo and Data Samples

The MC generation process has been explained in section 2.6.1 and here are reported the available simulation samples used for the $K_S^0 \to \mu^+\mu^-$ analysis so far. The samples used are listed in Table 3.1 for MC. The generation conditions are also reported, namely the generation-level cuts in the SV length and z coordinate. For the analysis only the MC where these cuts were applied are used for the $K_S^0 \to \mu^+\mu^-$ and $K_S^0 \to \pi^+\pi^-$ samples. Also samples of $K_S^0 \to \pi^+\pi^-$ and $K_S^0 \to \pi^+\mu^-\overline{\nu}$ where the pions are forced to decay into muons are available to study the background. In the exclusive samples, only K_S^0 in the forward direction are passed through the full detector simulation, while the K_S^0 in the "MinimumBias" sample, simulating a generic collision event only requiring that at least one p-p collision takes place, are generated in the full solid angle. The trigger emulated in pass-through mode (only a flag on the events stating with trigger line have been fired without remove the events if not triggered) corresponds to Trigger Configuration Key (TCK) 0x409f0045. This TCK includes the cuts used in most of 2012 data taking, and removes the prescales existing in several trigger lines. The stripping $K_S^0 \to \mu^+\mu^-$ and $K_S^0 \to \pi^+\pi^-$ selections, which are the same used in the data, were applied.

After the stripping and the reconstruction process from the 5M events for $K_S^0 \to \mu^+\mu^-$ only 60k events are left, and from these only a few percent pass the trigger request.

Particular attentions, to ensure that the computational time for the creation and the analysis of ntuple do not become prohibitive are taken, such as store into the ntuple only the variable needed for the analysis.

3.3 Stripping Line for $K_S^0 \to (\mu^+\mu^-/\pi^+\pi^-)$

All the triggered events are stripped with specific stripping lines to select candidates for the analysis purpose of the LHCb collaboration. The stripping selections for the $K_S^0 \to$

Name	Nu. of events	Generation cuts
Ksmumu	1.3M	No cuts
Ksmumu=tightcut	2.5M	$z_{K^0_S SV} \in$ [-1, 1]m
		$\rho_{K^0_S SV} <$ 30 mm
Kspipi=tightcut, rho	2.5M	$z_{K^0_S SV} \in$ [-1, 1]m
		$\rho_{K^0_S SV} <$ 30 mm
Kspimunu	28k	No cuts
Kspipi,mm=tightcut	250k	both π forced to $\mu\nu$,
		with $z_{K^0_S SV} \in$ [-1, 1]m
Kspimunu,m=tightcut	250k	π forced to $\mu\nu$,
		with $z_{K^0_S SV} \in$ [-1, 1]m
MinimumBias	20M	-

Table 3.1: MC 2012 generated and used for the study reported in this analysis. In the Name column the nomenclature tightcut and rho reflect the cuts applied on the samples at the generation level respectively in the z coordinate of the SV ($z_{K^0_S SV}$) and in the transverse decay length of the K^0_S ($\rho_{K^0_S SV}$). The m label is introduced for the samples where the pion is forced to decay into a muon and corresponding neutrino. The value of [-1, 1] m and 30 mm for these cuts select only K^0_S that decay into the VELO detector. The statistics reported is for only for one configuration of the magnet and so it is half of the entire available data.

$\mu^+\mu^-$ and $K^0_S \to \pi^+\pi^-$ candidates are reported in Table 3.2. These cuts are designed to be the same between the signal and the control channel, with the exception of the muon identification requirements and the mass windows. Indeed, the mass windows for the $K^0_S \to \mu^+\mu^-$ case is larger ([465, 1500] MeV/c^2) with respect to the $K^0_S \to \pi^+\pi^-$ ([400, 600] MeV/c^2) to allow the usage of the mass regions (called sidebands) outside the range defined as signal windows ([492, 504] MeV/c^2) where the final candidates of $K^0_S \to \mu^+\mu^-$ decay will be counted. The sidebands are very useful to study and extrapolate the background contributions (both the combinatorial background and the $K^0_S \to \pi^+\pi^-$ misID).

Another difference between the $K^0_S \to \pi^+\pi^-$ and $K^0_S \to \mu^+\mu^-$ is in the rate of analyzed events (prescale factors). For the $K^0_S \to \mu^+\mu^-$ all the events are in input to the stripping procedure, while for the $K^0_S \to \pi^+\pi^-$ situation the selection algorithm is run only in one every 1000 events to limit the retention of the stripping selection, due to the abundance of these decays.

The stripping requests select $K^0_S \to \mu^+\mu^-$ candidates require:

- Detached secondary vertex position with respect to the collision point asking for mother particles time of flight in its rest frame higher that 0.1 of the K^0_S lifetime. This variable is computed using the reconstructed momentum of the K^0_S candidate and the distance between the reconstructed secondary (SV) and primary (PV) vertices. When more PV are present it is used the one that gives the minimum K^0_S impact parameter significance.

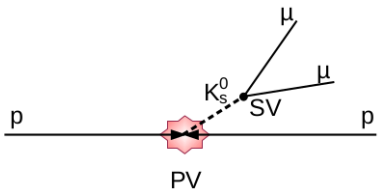


Figure 3.5: Schematic illustration of the $K_S^0 \to \mu^+\mu^-$ decay. The K_S^0 comes directly from the proton-proton interaction or from a decay of another strange particle produced (such as the K^{0*} that is the excitation of the K_S^0 itself); both the situations happens near the PV and so the K_S^0 momentum points in that direction. After a measurable flight distance the K_S^0 decays in two muons that can be tracked by the detector. A candidate $K_S^0 \to \mu^+\mu^-$ is searched requiring the selection cuts in Table 3.2.

Variables	$\boldsymbol{K_S^0 \to \mu^+\mu^-}$	$\boldsymbol{K_S^0 \to \pi^+\pi^-}$
Input Protoparticle	StdLooseMuons	StdNoPidsPions
Prescale	1.0	0.001
K_S^0 M	[465, 1500] MeV/c^2	[400, 600] MeV/c^2
K_S^0 tof	$> 0.1\ \tau$	$> 0.1\ \tau$
K_S^0 IP	< 0.4 mm	< 0.4 mm
K_S^0 DIRA	> 0	> 0
K_S^0 DOCA	< 0.3 mm	< 0.3 mm
Daughters Track χ^2/ndof	< 5	< 5
Daughters IP_{χ^2}	> 100	> 100

Table 3.2: The selection cuts performed in the stripping phase. $K_S^0 \to \mu^+\mu^-$ and $K_S^0 \to \pi^+\pi^-$ share the same cuts except for the mass window. In both Std Muons/Pions input protoparticles samples (that are the candidates created in the first stage of the reconstruction procedure of DaVinci software) the most important cuts, not reported in the table, are $P_T > 250$ MeV/c and $IP\chi^2 > 4$. Moreover, for the $K_S^0 \to \mu^+\mu^-$ case, there is, for both the muons, a muonID loose cut that yields an high efficiency of more than 99% (the IsMuon algorithm introduced in section 2.4). Definition of the variables entering in the stripping process is given in the text.

- Tracks with opposite charge identified as muons, forming a vertex with invariant mass compatible with the known K_S^0 mass (~ 497 MeV/c^2).
- Good quality reconstructed track of the daughters is required using the χ^2/ndof of the track fit.
- Good quality SV reconstruction realized asking for a low distance of closest approach (DOCA) of the two muons.
- Forward K_S^0 obtained requiring a positive cosine of the polar direction angle (DIRA).
- A K_S^0 coming from the PV asking for a low impact parameter with respect to PV.
- Daughters not related to PV (high impact parameter χ^2, which is defined as the difference of the χ^2 of the PV fit obtained with and without the considered track).

The search for $K_S^0 \to \mu^+\mu^-$ is performed by using only the long-long category, i.e. both daughters of the K_S^0 being long tracks. The mass resolution is crucial for this analysis, as it allows the separation between the $K_S^0 \to \pi^+\pi^-$ with both pions misidentified as muons from the real signal. Another reason to discard the downstream-downstream category is the larger influence (with respect to long-long) of the $K_L^0 \to \mu^+\mu^-$, which is an irreducible background for $K_S^0 \to \mu^+\mu^-$.

Studies (not included in this thesis) to include also the down-down category have shown that it is expected a $\sim 76\%$ increase of statistics at reconstruction and muon identification level but the signal over misid background is worse by a factor ~ 10 (accurate number would depend on selection, trigger, etc...) and for this reason the effective luminosity gain introducing downstream K_S^0 would be only at $\sim 15\%$ level.

The efficiency to reconstruct and select a $K_S^0 \to \mu^+\mu^-$ candidate (obtained from MC data) is $\sim 1.2\%$, mainly due to poor reconstruction for events where the momentum of K_S^0 is low. The total number of events, from the 2012 data acquisition, which have passed the stripping lines are reported in Table 3.3 for the $K_S^0 \to \mu^+\mu^-$, $K_S^0 \to \pi^+\pi^-$] case and for the NoBias case which is not affected by the trigger. Indeed, sometimes, the LHCb trigger acquires also events randomly to obtain a sample that is not biased by the trigger (NoBias sample). Thanks to this characteristic, this kind of events can be used to obtain important informations for the analysis in LHCb. Indeed, although this sample is negligible in terms of potential $K_S^0 \to \mu^+\mu^-$ signal, it provides a sample of $K_S^0 \to \pi^+\pi^-$ which can be used to understand the distributions that the signal would have in the case of no trigger bias.

Due to the fact that the reconstruction and the stripping procedure do not require the muon identification, it is expected that the efficiency for $K_S^0 \to \pi^+\pi^-$ is close to that of $K_S^0 \to \mu^+\mu^-$ because they have the same kinematic and topology. This has been confirmed in the simulation data.

Name	Ksmm	Kspp	NoBias
Nu. of events	4M	1.4M	90k

Table 3.3: The numbers here reported are relative to the events that have been selected from the stripping procedure without any further background suppression.

Chapter 4

Offline Selection and Trigger

In this chapter my contribution to the analysis related to the fiducial cuts applied offline and the aspects of the trigger relevant for the $K_S^0 \to \mu^+\mu^-$ channel are described in details.

4.1 Fiducial Cuts

After the stripping selection, and before the training and application of the multivariate classifier, some extra fiducial cuts, described in the list below, are applied offline to both $K_S^0 \to \mu^+\mu^-$ and $K_S^0 \to \pi^+\pi^-$ candidates. These cuts help to discriminate the $K_S^0 \to \mu^+\mu^-$ and $K_S^0 \to \pi^+\pi^-$ decay from the combinatorial background as it can be seen from Figure 4.1 where the distribution of the variables is shown, while keeping high efficiency for the signal. In the figure the trigger unbiased $K_S^0 \to \pi^+\pi^-$ events are compared to the combinatorial background extracted from the data right sideband. The criteria used for this further offline selection are:

- Muon track momentum smaller than 100 GeV/c.
- SV χ^2 smaller than 25.
- Time of flight in the rest frame of the mother smaller than 130 ps.
- Impact parameter of daughters larger than 0.5 mm. This cut has almost no cost on efficiency and simplifies the trigger studies as the IP cuts applied by the trigger become transparent.
- A cut in ProbNNK (confidence that the track can be associated to a kaon) to remove the small contribution from $K^{*0} \to K^{\pm}\pi^{\mp}$ (see Figure 4.2).
- Λ veto via a cut in the Armenteros-Podolanski plot (see Figure 4.5). The derivation of the Armenteros-Podolanski plot is explained in detail in this section.
- Cut in the ProbNNghost (confidence that the track do not correspond to a real particle) variable to remove the background contribution.

The last three requests listed above are new with respect to the 2011 analysis.

The total efficiency for these cuts is 97% measured in the NoBias distributions $K_S^0 \to \pi^+\pi^-$.

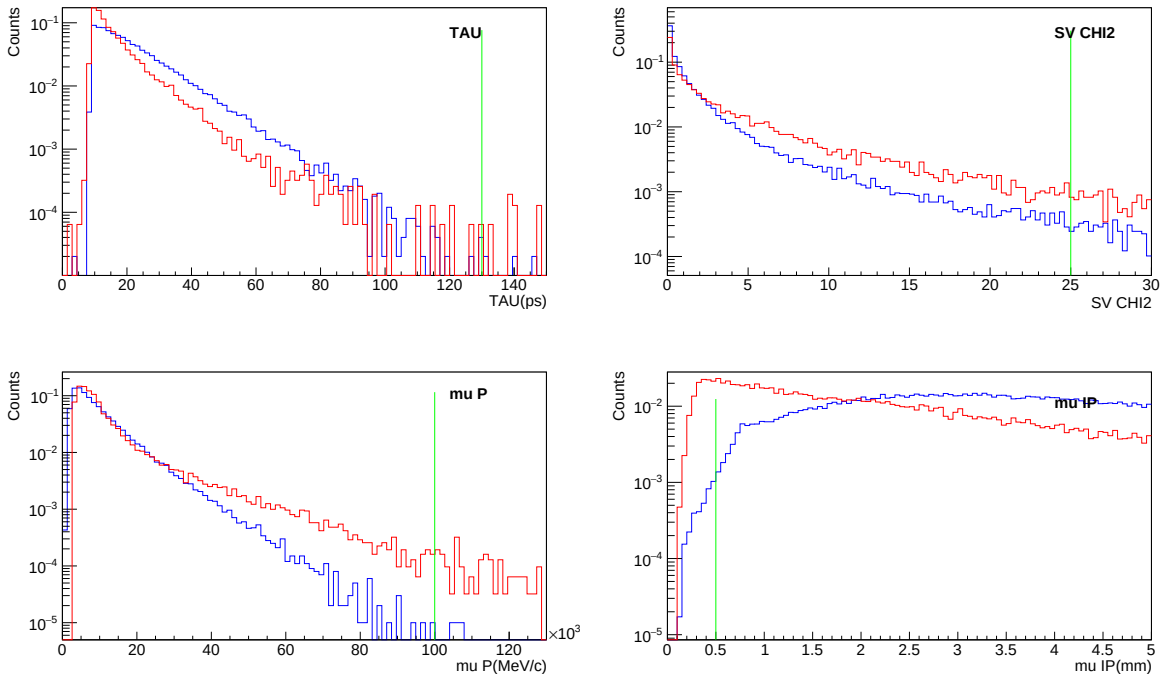


Figure 4.1: Fiducial cuts performed on NoBias $K^0_S \to \pi^+\pi^-$ data sample (blue line) and real data $K^0_S \to \mu^+\mu^-$ right sideband which model the combinatorial background (red line). The green lines represents the cut performed. In order to increase the statistical sample, μ^+ and μ^- distribution have been merged, as the possible discrimination effects between the two from the detector response are eliminated from the fact that there are measures with both the configurations of the magnetic field (B-up and B-down).

Here, the Armenteros-Podolanski plot (AP plot) theory is illustrated in order to show how this method can distinguish between two different decays simply using kinematic variables. In general, the AP plot is used to classify the V^0-particles, for example it makes possible to differentiate processes like $K^0_S \to \pi^+\pi^-$ and $\Lambda \to p\pi^-$ without a direct identification of the daughters.

In 1954, Podolanski and Armenteros [127] showed a method to analyse the dynamics of two-body V^0 decays to search for new unstable neutral particles in terms of measured quantities of the charged daughters particles for which the index 1 and 2 are used. These quantities are the three components of the daughters momentum: $\overrightarrow{p}_1$, $\overrightarrow{p}_2$.

The method is based on the fact that a two body decay has a exactly determined kinematic, for a chosen mass of the daughter particles, by the energy-momentum conservation. New variables are defined to study the process: $1/P$, α, P_T where P is the module of the total momentum $\overrightarrow{P} = \overrightarrow{P}_1 + \overrightarrow{P}_2$, α is called the Podolanski-Arementeros parameter and P_T is the transverse momentum of the daughters particles respect to the flight direction of the mother in the lab frame. Sometimes instead of P_T the variable $\epsilon = 2P_T/P$ is used.

The kinematic nomenclature is illustrated in Figure 4.3 where the * affix is used for the center-of-mass-system values. In particular,

$$P^* = \frac{\sqrt{(s - M_1^2 - M_2^2)^2 - 4M_1^2M_2^2}}{2\sqrt{s}} \tag{4.1}$$

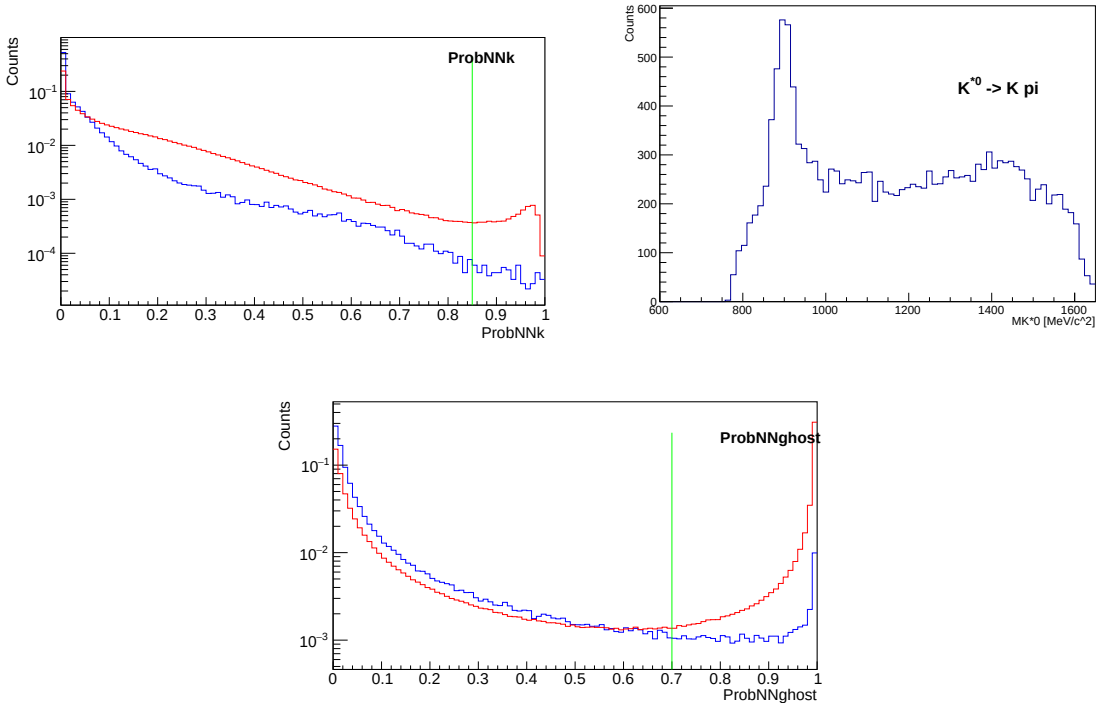


Figure 4.2: Top left: fiducial cut in the ProbNNK variable performed on NoBias $K^0_S \to \pi^+\pi^-$ data sample (blue line) and real data $K^0_S \to \mu^+\mu^-$ right sideband which correspond to combinatorial backgrounds (red line). The green lines represents the cut performed. Top right: reconstructed with the daughter with higher ProbNNK associate to K, the other to π for candidates NoBias $K^0_S \to \pi^+\pi^-$ showing a peak around the K^* mass (892 MeV/c^2). Only events with ProbNNK value of a daughter higher that the fiducial cut applied are considered for the plot. Bottom: fiducial cut in the ProbNNghost variable of both the daughter particles using the same sample as for the ProbNNK. The peak in the NoBias distribution is related to the small background contribution entering in this sample.

where $s = M_V^2$ is the square of the decaying neutral particle.

The Lorentz transformations link the energy and the particle momentum in the center-of-mass-system and the lab one.

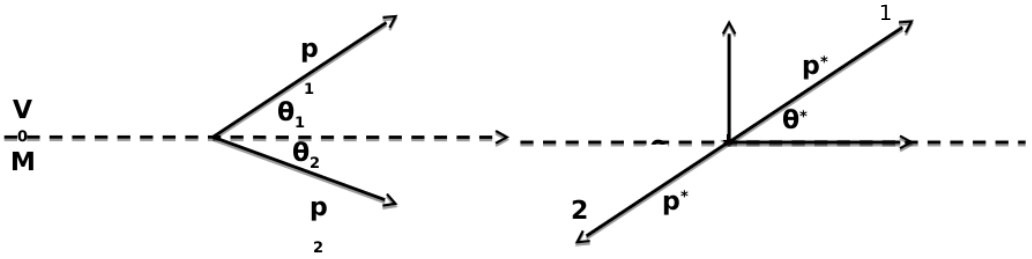
$$\begin{cases} E = \gamma E^* + \beta\gamma P_L^* \\ P_L = \beta\gamma E^* + \gamma P_L^* \\ P_T = P_T^* \end{cases} \tag{4.2}$$

and $P_T^* = P^* \sin\theta^*$, $P_L^* = P^* \cos\theta^*$.

In the lab system it is possible to define[1] $\alpha = \frac{P_L^1 - P_L^2}{P_L^1 + P_L^2}$, which expressed using the center-of-mass quantities becomes

$$\alpha = \frac{\gamma\beta(E_1^* - E_2^*) + 2\gamma P^* \cos\theta^*}{\gamma\beta(E_1^* + E_2^*)} \tag{4.3}$$

[1]It is convention to identify with the index 1 the positive charged particle while the index 2 is associated with the negative charged one. In this way the coniugate final states have the opposite value of α by definition.



(a) Two particle decay in lab system **(b)** Two particle decay in center-of-mass system

Figure 4.3: Two particle decay in both reference systems.

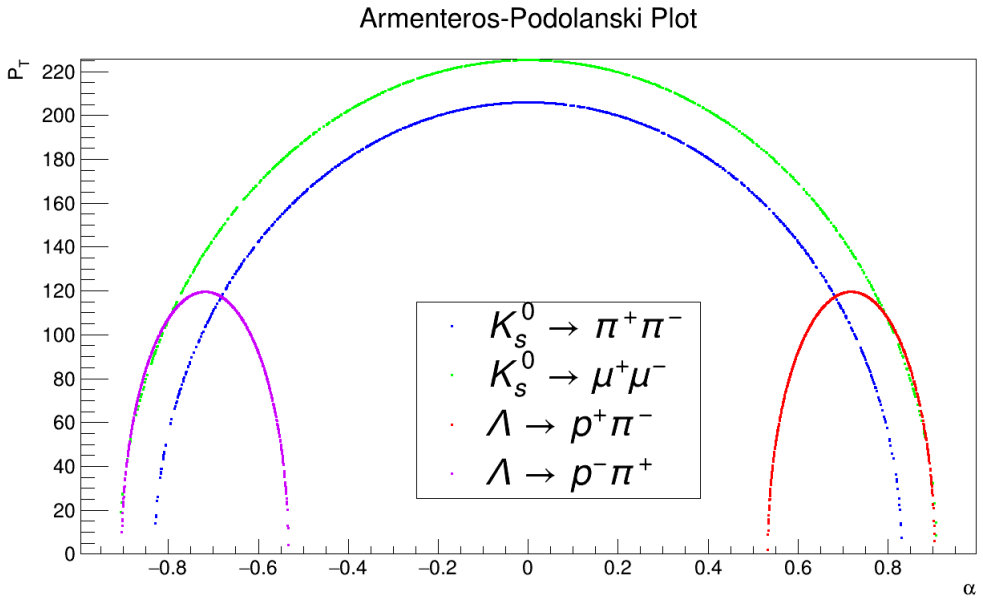


Figure 4.4: Simulated examples of AP plots for different K_S^0 and Λ two particle decays.

where $E^* = E_1^* + E_2^* = M_V$.
Defining $\bar{\alpha}$ as $\frac{E_1^* - E_2^*}{M_V} = \frac{M_1^{*2} - M_2^{*2}}{M_V^2}$, it is possible to rewrite α such as $\alpha = \bar{\alpha} + a\cos\theta^*$, where $a = \frac{2P^*}{M_V\beta}$. In the $\alpha - P_T$ plane (see Figure 4.4 where the quantities for simulated decays are plotted), the points that correspond to a given decay scheme lie on a family of ellipses determined by the particle mass and the velocity of the secondaries. This can be demonstrated expressing the relation using P_T^* and P^* instead of the angle θ^*:

$$\left(\frac{\alpha - \bar{\alpha}}{a}\right)^2 + \left(\frac{P_T}{P^*}\right)^2 = 1. \tag{4.4}$$

It is important to note that on the one hand P_T^* does not depend on β_1 and β_2 so it does not distinguish between the daughter particle masses. On the other hand α strongly depends on β_1 and β_2 as it is written with P_L^1 and P_L^2. So α differentiates between different daughter particles. Indeed, the average of α is $\bar{\alpha}$ which is 0 if $M_1 = M_2$, making the ellipse symmetrical respect to 0, as can be seen in Figure 4.4 for the $K_S^0 \to \pi^+\pi^-$ and $K_S^0 \to \mu^+\mu^-$

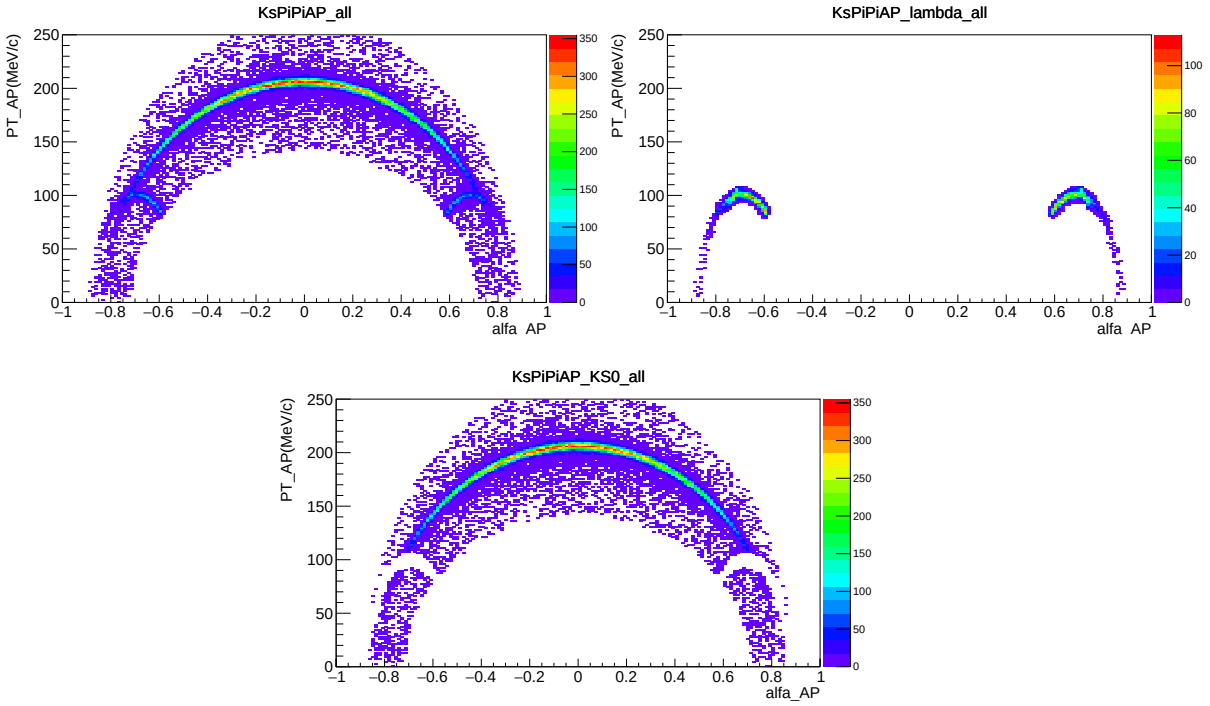


Figure 4.5: The AP plot obtained from $K_S^0 \to \pi^+\pi^-$ real data requiring the K_S^0 mass in the mass signal window. Top left: no cuts performed. Top right: removed events corresponding to Λ and $\overline{\Lambda}$ decay. Bottom: Accepted events.

decays. Instead, if the mother particle can produce its charge conjugated final state too, the corresponding plots are two similar ellipses with opposite centers respect to 0. This case is shown in Figure 4.4 for the $\Lambda^0 \to p\pi^-$ and $\bar{\Lambda}^0 \to \bar{p}\pi^+$ decays.

The AP plot is used here to reject the Λ decay contamination (when both p and π are identified as muons or pions) in the $K_S^0 \to \mu^+\mu^-$ and $K_S^0 \to \pi^+\pi^-$ samples through the cut illustrated in Figure 4.5.

4.2 Trigger Studies

This section summarized the main trigger aspects of the analysis. This includes investigating the lines mainly responsible for the acquisition $K_S^0 \to \mu^+\mu^-$ decay and classify the different samples of events which are achieved by different trigger lines.

A very important difference of this analysis with respect to others analysis performed in LHCb (e.g. $B_s^0 \to \mu^+\mu^-$) is the very low trigger efficiency for signal. For example the L0 muon level selects muons with transverse momentum (PT) larger than 1.5 GeV/c, a relatively low threshold for b decays that results, on the contrary, in a low efficiency for $K_S^0 \to \mu^+\mu^-$ decays. Indeed, the need for tight trigger requirements in the LHC environment turns out to be the most limiting experimental factor for this search.

The LHCb trigger has been explained in section 2.5 and here only the important aspects for the analysis will be retained. The LHCb trigger is divided into three levels: L0 (hardware trigger), HLT1 (first software level), HLT2 (second software level). The muons from K_S^0 can cause the trigger to fire at any level (e.g. the L0 level if they have a PT above $\sim$ 1.5

GeV/c), and in this case the corresponding events are referred as Triggered On Signal (Tos) in the corresponding trigger level.

Due to the low efficiency of the LHCb trigger for the low PT muons coming from $K_S^0 \to \mu^+\mu^-$ candidates, there are strong biases in the distributions of these events. For example, in order to fire the trigger, the K_S^0 must have very high PT[2], but high PT K_S^0 will decay outside the VELO unless they are very short-lived and so the Tos requirement biases the time of flight distribution towards smaller values.

Alternatively, the events can also have been triggered by an object different to the signal K_S^0, in this case it is referred to as Triggered Independently of Signal (TIS) at that trigger level. The trigger biases in this case are small and this kind of samples can also be used to obtain a sample of events with minimal trigger bias[3].

It is possible, but the probability is low, that an event can be classified as both Tis and Tos at the same time; these shared events between the two categories will be taken into consideration only in the Tis sample to minimize the trigger bias. In this way the overlap with Tis and Tos is excluded and the different results for the different trigger paths can be considered as statistically independent when combined into a final result. In fact, a way of studying if a line is well suited to trigger a given process is to check the number of events Tos and Tis which are selected on that line: if the fraction of Tis events is small with respect to Tos events, it means that the line has not a high trigger efficiency for the searched processes. This is a quick method to compare and understand the importance of a line and it has been used to eliminate other lines, not mentioned here, for which it was obvious that their contribution is negligible.

The trigger lines that can contribute to the $K_S^0 \to \mu^+\mu^-$ search in LHCb are those which require the presence of 1 or 2 muon tracks and from those the lines with the lowest PT muon and di-muon mass cuts will be predominant. Indeed, it is found that the best trigger lines are for the hardware level the L0Muon and the L0DiMuon (see Table 4.1), for the first software trigger level the Hlt1TrackMuon and Hlt1DiMuonLowMass (see Table 4.2) and for the second software trigger level the Hlt2DiMuonDetached and the Hlt2SingleMuon (see Table 4.3). In the mentioned Tables the differences between the 2011 case and the 2012 case are bold to point out the increase of trigger efficiency for the $K_S^0 \to \mu^+\mu^-$ decay, mainly at the Hlt2 level where the cuts on the di-muons invariance mass was removed and a reduction in PT request for the muons was applied.

At the L0 trigger level it is found that the only contributing line is the L0Muon while the L0DiMuon suffers of a higher PT cut that involves both the muons. This situation is different for both the Hlt1 and Hlt2 stage. In the former both the lines are found to have a comparable efficiency, while in the latter the DiMuon line prevails respect to the SingleMuon (called TrackMuon) one due to the lower PT cuts.

In general, the Tis sample does not add a substantial contribution to the overall statistics if the trigger efficiencies when the Tos efficiency is high enough as in the case for the $B_s^0 \to \mu^+\mu^-$ decay. Indeed, in the 2011 analysis Tis-selected events at the three levels were found to add a significant contribution to the result. Because the Hlt2 has been changed in the 2012 with respect to the 2011 and now the efficiency on that level is higher, it is expected that the Tis request at the Hlt2 level will now give a negligible contribution with respect to the Tos case.

[2]If the K_S^0 goes straight in the z direction the PT of the daughter muons will have, at maximum, the value of 200 MeV/c, so to fire the trigger the K_S^0 must have a PT different from zero.

[3]Kinematical correlations with the particles firing the trigger can in principle produce a bias. This will be investigated in Section 5.

As reported in the Table 4.2, the Hlt1DiMuonLowMass line has no explicit cuts on the P and PT of the muons and therefore looks very promising for our analysis, but it has to be reminded that the MuonID algorithm implemented in the trigger level (see section 2.4 in particular the Table 2.1) requires hits on all 5 muon stations, resulting in a $\gtrsim 6$ GeV/c cut on the muon momentum. Thus, also the Hlt1DiMuonLowMass is affected by non explicit P-PT cuts. A MC sample of 60k candidates that has passed the stripping line and the fiducial cuts described in 3.3 has been used.

To understand the best trigger lines that can contribute to $K^0_S \to \mu^+\mu^-$, an accurate study has been performed and here described.

Variables	**L0Muon**		**L0DiMuon**	
	2011	**2012**	**2011**	**2012**
Prescale	1.0	1.0	1.0	1.0
SPD Mult	< 600	< 600	< 900	< 900
P_T^{fist}	> 1480 MeV/c	$>$ **1760 MeV/c**	-	-
$\sqrt{P_T^{first} \cdot P_T^{second}}$	-	-	> 1296 MeV/c	$>$ **1600 MeV/c**

Table 4.1: L0 trigger cuts for the most relevant lines to select $K^0_S \to \mu^+\mu^-$ candidates. Changes between 2011 and 2012 are bold.

Variables	**Hlt1TrackMuon**		**Hlt1DiMuonLowMass**	
	2011	**2012**	**2011**	**2012**
Prescale	1.0	1.0	1.0	1.0
VeLo Hits	> 0	> 0	-	-
VeLo Missing Hits	< 999	< 999	-	-
Hits	> 0	> 0	-	-
Track χ^2/ndof	< 2	$<$ **2.5**	< 4	$<$ **3**
P	> 8 GeV/c	$>$ **3 GeV/c**	> 6 GeV/c	$>$ **0 GeV/c**
P_T	> 1.0 GeV/c	> 1.0 GeV/c	> 500 MeV/c	$>$ **0 MeV/c**
IP	-	$>$ **0.1 mm**	-	-
IPχ^2	> 16	> 16	> 3	$>$ **6**
DiMuons DOCA	-	-	< 0.2 mm	< 0.2 mm
DiMuons M	-	-	> 1000 MeV/c^2	$>$ **0 MeV/c^2**
DiMuons χ^2_{vtx}	-	-	< 25	<25

Table 4.2: Hlt1 trigger cuts for the most relevant lines to select $K^0_S \to \mu^+\mu^-$ candidates. Changes between 2011 and 2012 are bold.

Variables	Hlt2DiMuonDetached		Hlt2SingleMuon	
	2011	**2012**	**2011**	**2012**
Prescale	1.0	1.0	0.5	0.5
Muon P_T	-	-	$>$ 1300 MeV/c	$>$ 1300 MeV/c
Muon Track χ^2/ndof	-	-	< 2	< 2
Muon IPχ^2	> 9	> 9	> 200	> 200
Muon IP	-	-	> 0.5 mm	> 0.5 mm
DiMuon P_T	$>$ 1500 MeV/c^2	$>$ **600 MeV/c**	-	-
DiMuon M	$>$ 1000 MeV/c^2	$>$ **0 MeV/c^2**	-	-
DiMuon V_s/σ_{Vs}	> 7	> 7	-	-

Table 4.3: Hlt2 trigger cuts for the most relevant lines to select $K^0_S \to \mu^+\mu^-$ candidates. Changes between 2011 and 2012 are bold. The predominant Hlt2 line is the DiMuonDetached line, its bandwidth has been increased in 2012 to include the K^0_S mass (∼497 MeV/c^2).

In general, for a trigger line the cuts performed can be subject to change during the data taking period. When a line is changed a new trigger configuration key (TCK, see subsection 2.5.3) is created. When a MC sample is generated the trigger requirements referred to a specific TCK and so our trigger analysis is associated with the TCK used for the MC $K^0_S \to \mu^+\mu^-$ sample. Our MC TCK is 0x409f0045 which has the same cuts at all the trigger levels (L0, Hlt1, Hlt2) of the most used TCK of the 2012 data taking and the cut for the lines used in this analysis had not been changed during all the 2012 data taking.

Hereafter the trigger path of a candidate will be expressed in the form L0-Hlt1-Hlt2 with the Tis or Tos nomenclature to state if the candidates has been triggered independently from the presence of the candidate $K^0_S \to \mu^+\mu^-$ or not. In the ultimate analysis the lines considered for the L0 is only the L0Muon, for the Hlt1 both the lines named Tos1 (Hlt1TrackMuon) and Tos2 (Hlt1DiMuonLowMass), while the Hlt2DiMuonDetached line for the Hlt2.

Out of the sample of ∼ 60k MC $K^0_S \to \mu^+\mu^-$ candidates only 2.7% pass the trigger (logic OR of all the mentioned trigger lines) with the efficiencies for different trigger paths reported in Table 4.4. The relative errors on all these measurements are at the few % level and, because of its low value it is not considered for the decision of the best trigger lines. It is omitted in the following Tables to make the results more clear and visible. The fraction of shared events, triggered simultaneously by two different paths, are also shown. In the last row of the Table the percentages of events that are only selected by the specific trigger path are shown, allowing to understand which line can introduce a bigger unique sample of $K^0_S \to \mu^+\mu^-$ candidates.

The efficiencies reported in Table 4.4 are affected by the shared events so to understand which is the best trigger lines, the following procedure has been applied: in the first step the number of triggered events for the most prevalent lines are obtained, then the line with the bigger number is chosen and all the events that this line shares with the others are removed from them and the whole procedure is reiterated for all the lines except the chosen one. The numbers obtained at the end of the process are reported in Table 4.5 and normalized

Trigger category	L0&HLT1&HLT2	shared with	only in this line
1)TosTos1Tos	1.3%	2) 34% 3) 21%	0.7%
2)TosTos2Tos	0.7%	1) 62% 4) 26%	0.2%
3)TisTos1Tos	0.5%	1) 61% 4) 36%	0.1%
4)TisTos2Tos	0.4%	2) 51% 3) 43%	0.1%
5)TisTisTos	0.8%	3) 11% 4) 7%	0.7%

Table 4.4: Tigger category studies: first Tos stands for L0Muon, Tos1 = HLT1TrackMuon, Tos2 = HLT1DiMuonLowMass, Tos = HLT2DiMuonDetached. In 2nd and 4th column are reported the relative fractions to total selected and reconstructed events, the 3th column show the two higher shared events percentage with other lines.

to the total number of all the lines.

Trigger category	Number of events	Percentage respect to all
TosTos1Tos	828	(50%)
TosTos2Tos	165	(10%)
TisTos1Tos	58	(4%)
TisTos2Tos	96	(6%)
TisTisTos	491	(30%)

Table 4.5: The total number, obtained with the OR of the all lines, is 1638 (2.7% of the total reconstructed and stripped candidates). The study is performed in the MC $K^0_S \to \mu^+\mu^-$ data which contain a statistics of $\sim$ 60k reconstructed and stripped candidates. It's clear which are the best trigger lines to be included in the analysis.

Only the most prevalent lines (TosTos1Tos, TosTos2Tos, TisTisTos) for a total of 90% of all the triggered events are considered for the analysis. These lines are also the simplest (all Tos, nearly all Tis) to be analyzed. The introduction of TisTos1Tos and TisTos2Tos (10%) would add little statistics and complicate the analysis (different and more complex normalization, MVA studies, etc..). As the Tos request introduces a bias in the selected candidates, the final samples in the analysis are the entire TisTisTos (so no trigger bias introduced but Hlt2), TosTos1Tos without the shared events ($\sim$ 3% with TisTisTos) and the not shared TosTos2Tos (shared Tos1-Tos2 = 34%, and no substantial shared events with TisTisTos).

Hereafter only the bigger statistic trigger category (TosTos1Tos) is analyzed, due to lack of statistics in the MC sample to study the other two trigger categories that will be considered when larger MC samples will become available. However, consideration about the TisTisTos and TosTos2Tos trigger categories will be discussed.

The difference bias introduced by the trigger for the events in the three trigger paths is shown in Figure 5.2, 5.3 and 5.4. It is important to notice that when a K^0_S goes in the forward direction (z axis) the PT of the muon is at maximum $\sim$ 200 MeV/c and, therefore, the event is discarded by the trigger. These not selected events are the one that have their

SV in the z axis (x = 0, y = 0) and so the fact that they cannot be selected by the trigger creates a bias in the SV x and y distribution. This situation is not true for the Tis case where the K_S^0 does not contribute to the trigger and so its direction is allowed in the forward one too.

Chapter 5

Suppression of the Combinatorial Background

This chapter focus on the Multivariate Analysis that is a very important step for the analysis, allowing to suppress the combinatorial background.

5.1 General aspects

A MVA aims at extracting the maximum available information from the data. For each event we measure a set of numbers $\overrightarrow{x} = (x_1, \cdots, x_n)$, where x_i correspond to variables such as kinematic parameters or geometrical properties. $\overrightarrow{x}$ follows some n-dimensional joint probability density, which depends on the type of event produced, that is signal or background. The goal of MVA is to find an optimal way to separate (classify) the event types in a way that exploits the information carried in $\overrightarrow{x}$.

From the n-dimensional space of $\overrightarrow{x}$ it is possible to create a one dimensional variable $y(\overrightarrow{x})$ that can be treated as a discriminating function and can be optimized so that its distribution has the maximum possible separation between the event types. In all multivariate analyses it is important to consider the best choice of variables to use and the functional form of decision boundary (type of classifier). The discrimination between the signal and the background is now effectively a single cut on $y(\overrightarrow{x})$. To obtain the parameters which optimize the discrimination power, $y(\overrightarrow{x})$ is trained and tested with the use of data sets for which the desired output is known for each event.

For the analysis the TMVA toolkit, which is integrated into the analysis framework ROOT and hosts a large variety of multivariate classification algorithms, has been used. Concerning the implementation of the various binders used, reference is made to the TMVA manual [128], which has been consulted to understand how to optimize and modify the various parameters of the provided classifiers. Multivariate classification methods based on machine learning techniques have become a fundamental ingredient to most analyses and are here developed to reduce the combinatorial background that is still present after the selection cuts. To solve a concrete problem, all methods require at least some specific tuning to deploy their maximum classification capabilities.

A complete TMVA analysis consists of two steps: the training, where the ensemble of available and optimally customized MVA methods are trained and tested on independent signal and background data samples, and an application phase, that, using the best trained MVA methods, performs the classification of data samples with unknown signal and back-

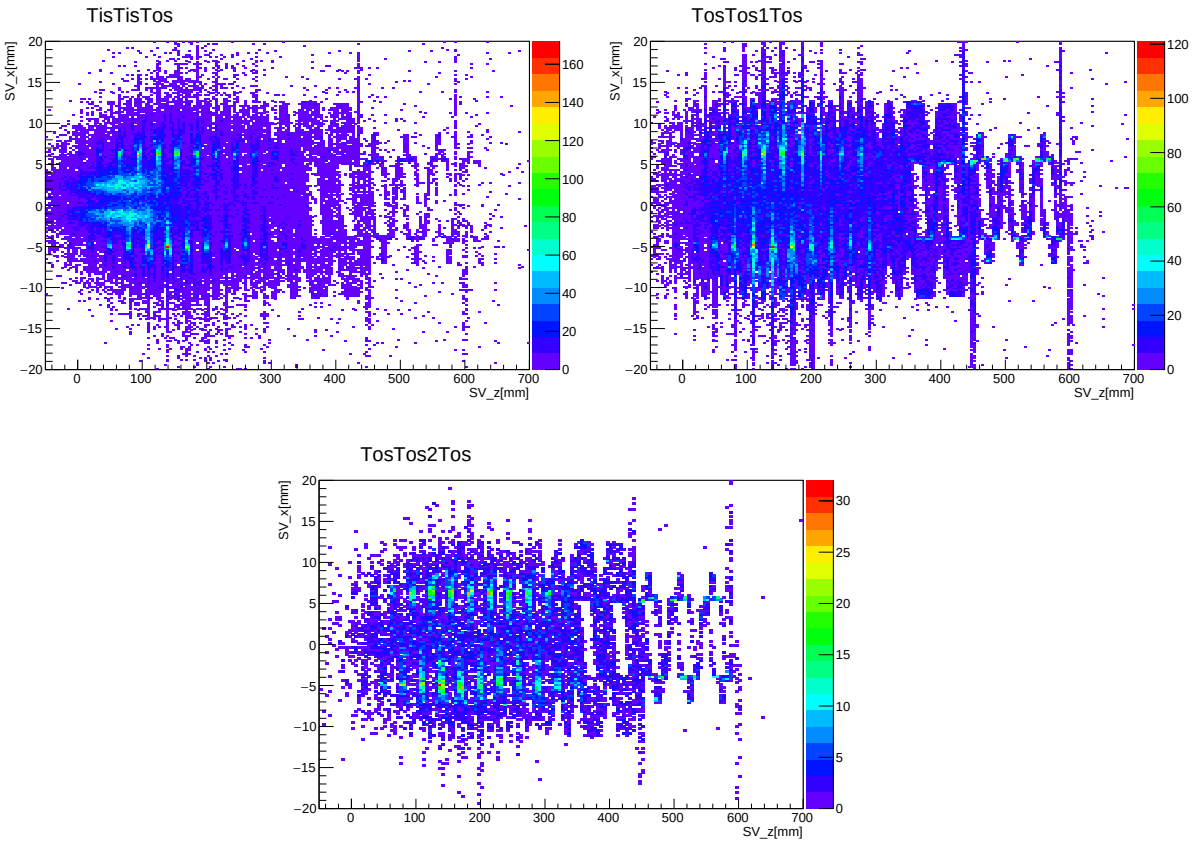


Figure 5.1: Secondary vertex x coordinate vs z coordinates for the thee trigger categories obtained in the rightside data, thus associated with the combinatorial background. The contribution from the material is evident but, while for the all Tos cases the distributions are comparable, for the TisTisTos case it is clear that the vertex distribution is different due to the contributions from other part of materials. This is because the fraction of background from material interactions is bigger in the Tos category due to the harder transverse momentum spectrum of signal particles.

ground composition. The output of a classifier may then be used to set a cut that increases the signal purity of the sample.

The variables that allow the separation of signal from background have very different distributions for events coming from different trigger categories. Furthermore, also the nature of the background changes, in particular an important amount of background comes from interactions of primary particles in the detector material, mimicking a detached decay vertex. This contribution is found to depend on the trigger bias condition. These scattering events can be misidentified as $K^0_S \to \mu^+\mu^-$ with the SV of the K^0_S situated inside the material. A well known characteristic of material interaction events is that their measured transverse momentum is higher (due to the scattering process) with respect to the real particles an so more suitable to be selected by the trigger. This means that, if we put the SV x and y coordinates in a scatter plot, if the material contribution is present the pattern of the VELO stations has to be somehow reproduced. Indeed, an important discriminating variable turns out to be the position of the decay vertex, since sideband data show a relevant fraction of events located inside the VELO sensors, proving a substantial contribution from interactions with the detector material (see Figure 5.1).

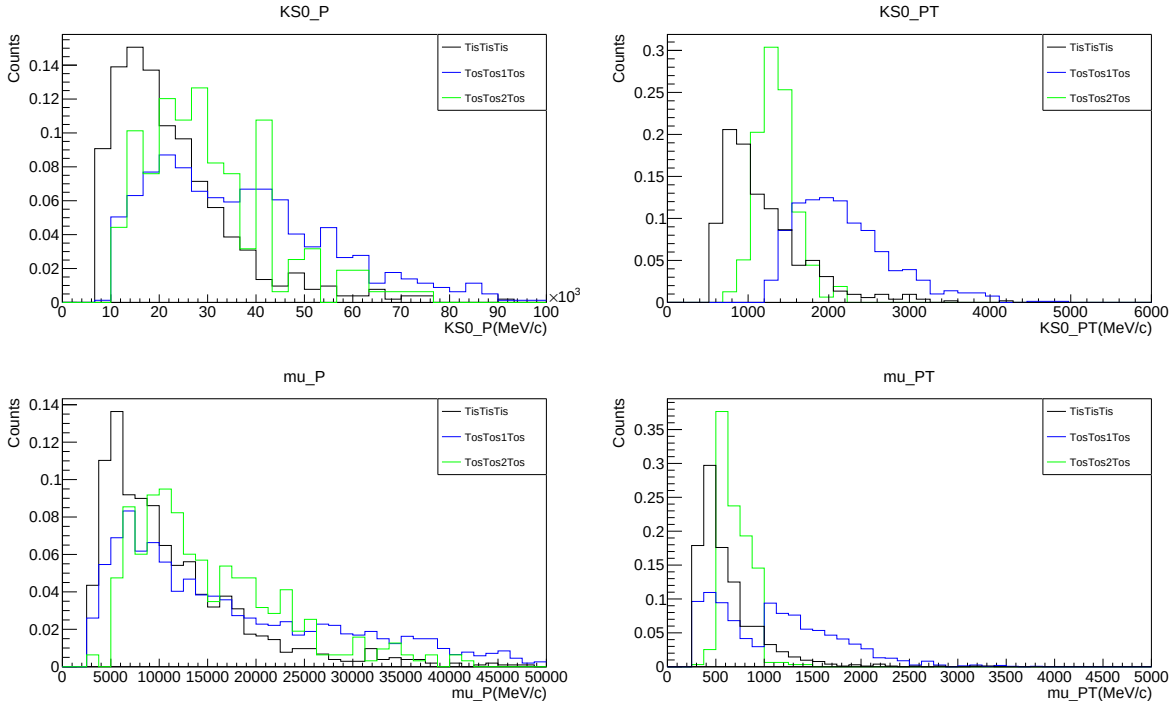


Figure 5.2: Momentum and transverse momentum distributions for K_S^0 and the daughter muons in different trigger categories. The distributions differ from each other forcing the MVA study to be split. In order to increase the statistical sample, μ^+ and μ^- distribution have been merged, as the possible discrimination effects between the two from the detector response are eliminated from the fact that there are measures with both the configuration of the magnetic field (B-up and B-down).

The K_S^0 kinematic distributions and so the MVA input variables depend on the trigger categories (see Figure **??**, 5.3, 5.4). For example, strong dependencies from the trigger category are present both in the time of flight variable and in the polar angular direction of the muons in the reference system where the K_S^0 is at rest. In the first case, it is lower for higher transverse momentum and momentum due to the effect of time dilation between inertial references and to the fact that the K_S^0 can go outside the VELO detector transversely. As only the K_S^0 decaying in the VELO are considered, these are necessarily characterized by a smaller average time of flight. Regarding the angular distribution of the muons, the trigger require the muons to have an high transverse momentum, favoring the muons produced at large angles with respect to the K_S^0 direction.

This forces us to implement and optimize different specific MVA for each trigger category, checking which is the best variables combinations that can give the best discrimination performance separately in each case.

Due to the lack of MC statistics for the TisTisTos and TosTos2Tos samples, the only studied path is the TosTos1Tos, but the programs created are able to deal with the other two trigger categories too.

To perform the discrimination in a blind analysis, proxy samples of data with MVA input variables distributions as similar as possible to the real signal and background samples are needed. Because of the not complete reliability of the MC in the reproduction of the real data, the usage of simulated samples should be avoided in favor of data-driven methods but still the MC can be used to validate the procedure (comparison between two MC samples

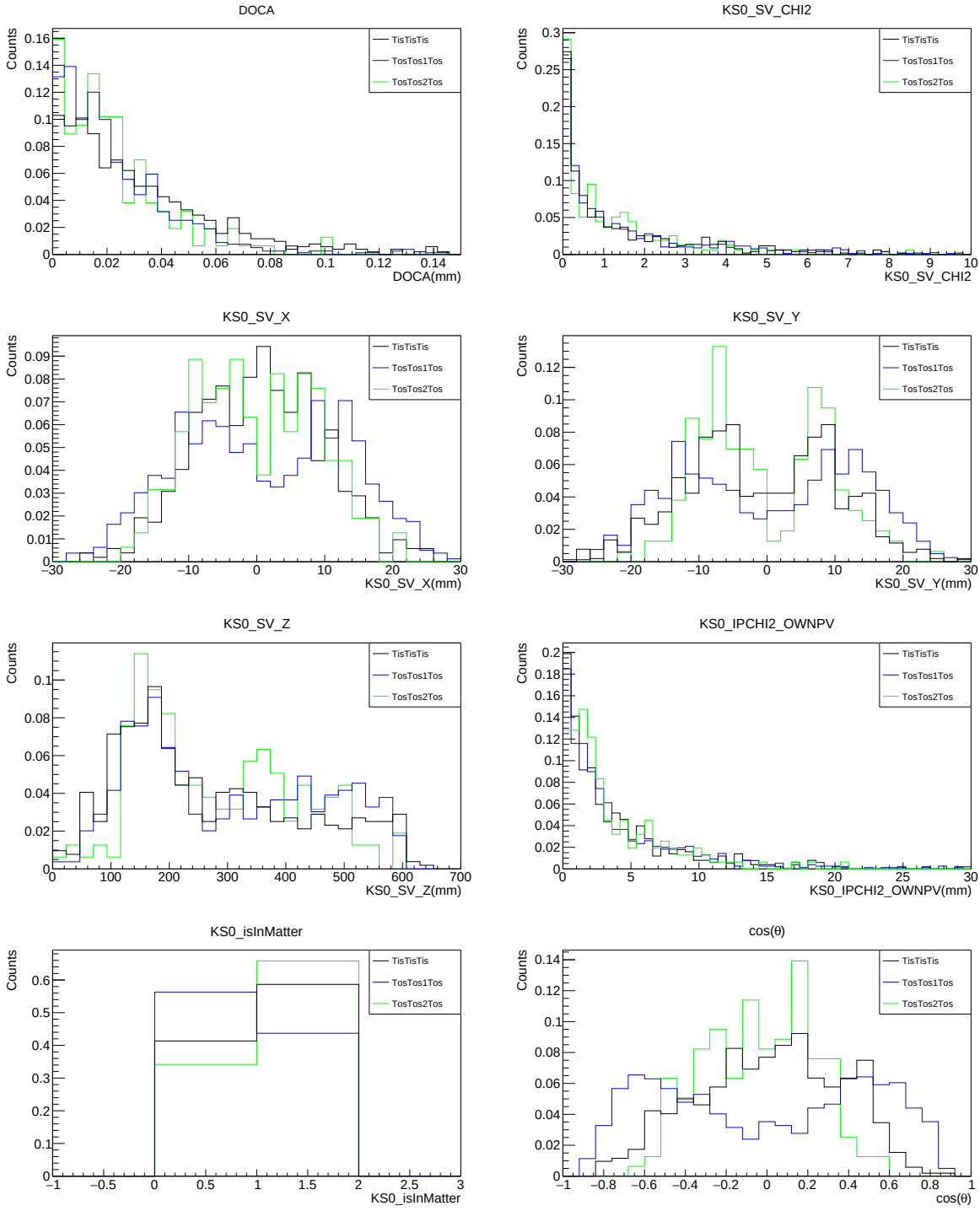


Figure 5.3: The distribution of the MVA input variables compared for different trigger categories. The distributions differ from each other forcing the MVA study to be split.

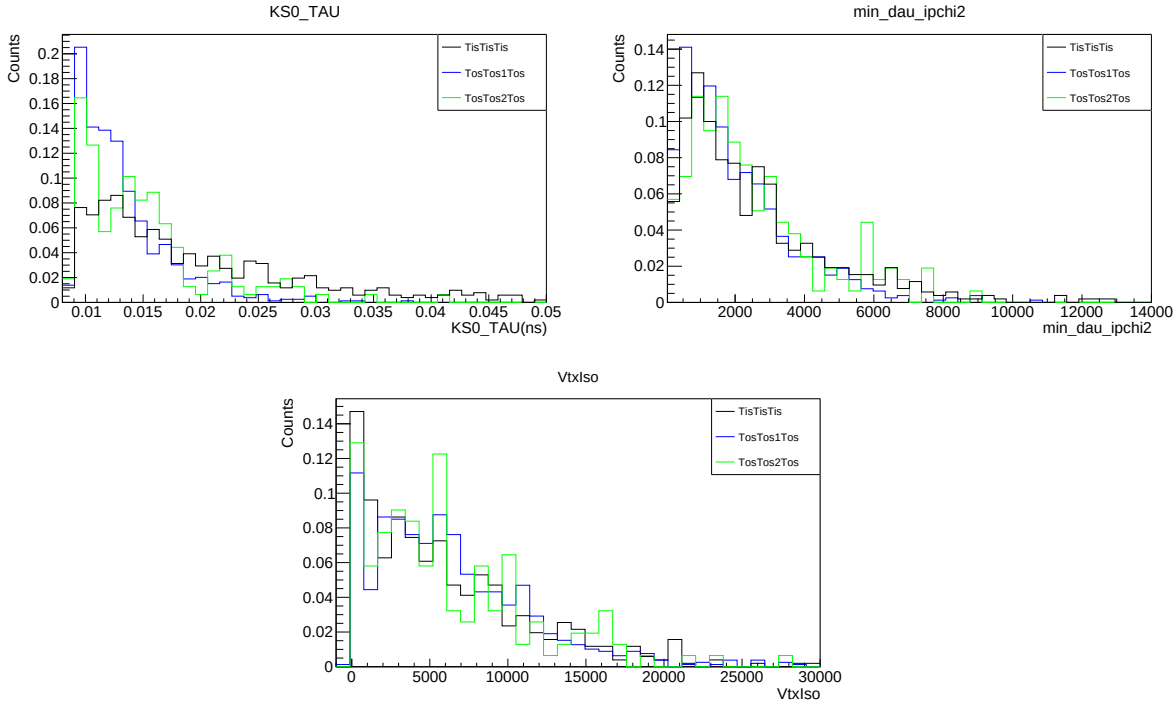


Figure 5.4: The distribution of the MVA input variables compared for different trigger categories. The distributions differ from each other forcing the MVA study to be split.

are still reliable because both samples are affected by the same possible simulation defects). To simulate the signal we make use of $K^0_S \to \pi^+\pi^-$ events (the procedure is detailed in 5.2).

The combinatorial background is composed by two reconstructed tracks not corresponding to the K^0_S daughter particles that accidentally simulate the signal expected from $K^0_S \to \mu^+\mu^-$ or $K^0_S \to \pi^+\pi^-$, passing the stripping line and the fiducial cuts described previously. The signal $K^0_S \to \mu^+\mu^-$ is expected to give a clear vertex detached from the PV not ending into the material of the detector, while for the combinatorial background the quality of the fit of the reconstruction procedure is expected to be worse compared to the signal. So the input variables are chosen from kinematic and reconstruction quality variables only, without involving the invariant mass and the μID-related variables in order to avoid that the MVA will be able to distinguish the $K^0_S \to \mu^+\mu^-$ case from the $K^0_S \to \pi^+\pi^-$ (normalization channel and sample used to simulate the signal distributions).

The input MVA variables used as minimal set to describe the topology and kinematics of the signal decay for all the trigger paths are:

- The lowest impact parameter significance of the two muon candidates with respect to any of the PVs reconstructed in the event. The impact parameter significance ($IP\chi^2$) of a track is defined as the difference in the χ^2 of the fit of the vertex obtained with and without the introduction of the track in the fit.

- K^0_S impact parameter χ^2.

- Distance of closest approach between muons, which is expected to be larger for the combinatorial background.

- Secondary vertex χ^2, that adds complementary informations with respect to DOCA, as it uses information on the error of the vertex fit.
- Cos(θ) with θ the decay angle of the daughters particles in the K^0_S frame; the angular distribution has to be isotropic in the solid angle for signal, but not necessarily for background.
- Isolation criteria of the SV defined as the difference in the χ^2 of the vertex fit with only the two muons tracks and when the next nearest track is added. This is a new variable with respect to the 2011 analysis allowing a better discrimination performance in the new data because the signal should present a very clear isolated SV while the background could not. For example, in the VELO stations we expect for two isolated hits for the signal, while in a nominal situations the hits multiplicity in the VELO stations are higher than two.

Some of these variables are the same used to select the candidates in the stripping phase where only rigid cuts were applied.

Other variables, depending on trigger category are:

- A boolean variable IsMaterial that states if the SV of the K^0_S coincides with a point in the material of the detector or not
- Coordinates (x, y, z) of the SV in the laboratory rest frame. As shown in the Figure 5.1, the SV position is an important parameter to discriminate the fraction of the background coming from material interactions.
- K^0_S PT.
- The time of flight of the K^0_S candidate in its rest frame.

From the combination of more variables the K^0_S mass can be obtained (for example if you know the momentum and the flight distance of a particle) and so not all the variables combinations can be used to not allow the MVA to learn about the K^0_S mass. When the SV coordinates are introduced the PT of the K^0_S is not included because the simultaneous use of the time of flight, the PT and the SV position will allow the MVA to learn which is the mass of the candidate, and thus to use it to discriminate from the background. In such a case, the background rejection will be very different in the sidebands than in the signal region, and a fake signal peak could be artificially created. Alternatively, the PT of the K^0_S can be included if the time of flight is dropped. The formula that makes possible to obtain the mass from the knowledge of momentum, transverse momentum, time of flight and SV coordinates position is (here we suppose that the PVs are compatible with the $(0,0,0)$ point that is not so different from the reality): $M = \frac{\tau PT}{\sqrt{x^2+y^2}} = \frac{\tau P_z}{z}$ that means that, when τ is used, if PT is introduced then only the z coordinates can be included and be sure that the MVA is not able to recreate the information of the mass of the mother, while if P is introduced then x and y can be included. Four different variables combinations are studied:

- $PT_{K^0_S}, TAU_{K^0_S}$, IsMaterial.
- $PT_{K^0_S}, TAU_{K^0_S}$, IsMaterial, SV z coordinate.
- $PT_{K^0_S}$, SV coordinates.

- $TAU_{K^0_S}$, SV coordinates.

The first two cases differ only for the introduction of the SV z coordinate and their comparison is used to be sure that the MVA is not able to learn about the mass of the mother particle. Indeed, if the mass can be guessed the discrimination power between signal and background has a big artificial improvement, as is observed when are used variables that are known to allow the reconstruction of the mass.

The use of other variables has been investigated, as for example the number of hits in the TT stations from which the muons are reconstructed, that has proved to be useful for the characterizing of low momentum muons. No significant improvement has been seen after its introduction in the MVA and so these variables are not included in the previous list[1]. One of this tested variables is the probNNghost one that states the confidence of a track to be a causal recombination of hits, so we decided, although its contribution in the MVA was not effective, to apply a fiducial cuts on it to do not include ghost tracks in the analysis.

5.2 Signal and Background Proxy Creation

The goal of the proxy is to simulate the distribution of the signal $K^0_S \to \mu^+\mu^-$ for the input variables of the MVA and so both the kinematic aspects and the trigger bias need to be taken into account. The procedure to create the signal proxy sample is first validated in the MC and then applied in the data sample.

The sample of $K^0_S \to \pi^+\pi^-$ with pions misidentified as muons can recreate (with some cuts in the muonID variables) the muon trigger bias but it suffers of different kinematic distribution with respect to the $K^0_S \to \mu^+\mu^-$ signal. This can be understood remembering that the muon trigger requires that all the five muon stations are crossed, creating a cut in the momentum of the muons. Moreover, the pions misidentified as muons are mostly the one that decay in flight before the muon stations that means that their time of flight, in the lab system, is lower than signal muons. Difference in the kinematic means that the MVA input variables of the proxy have a different distribution with respect to the real signal sample. The approach used in this analysis consists in using $K^0_S \to \pi^+\pi^-$ without trigger bias, and then applying offline cuts that mimic the effect of the trigger selections used for the signal.

The proxy creation procedure is not the same of that adopted in the 2011 data where, to simulate the trigger effects, samples of $K^0_S \to \pi^+\pi^-$ with pions misidentified as muons was used requiring some cuts in the muonID variables.

The trigger requirements create a bias in the signal sample, in order to get a realistic proxy the triggers effects are emulated applying cuts on the Muon P and PT. The creation of the proxy sample is validated on the MC checking if the kinematic variables of the proxy become closer to the $K^0_S \to \mu^+\mu^-$ one. It is expected that the similarity of the kinematic distributions will bring also the input MVA variables to look similar, as checked after the tuning of the proxy cuts.

The muon P and PT value are not the same for the online and offine reconstruction due to different algorithm with the exception of the Hlt2 (but the harder cuts are in the L0Muon and Hlt1 lines). This aspect is taken into account introducing a smearing on these

[1] Of course this situation is still consistent with the results obtained in other LHCb analysis, because there are different conditions with respect to the $K^0_S \to \mu^+\mu^-$ case.

Variable	**TosTos1Tos**
Triggered P_μ	> 7500 MeV/c
Both muons P	> 4500 MeV/c
Triggered PT_μ	> 1130 MeV/c
Di Muon $PT_{\mu\mu}$	> 600 MeV/c

Table 5.1: The proxy cuts applied on $K^0_S \to \pi^+\pi^-$ data to create the training sample for the MVA. The cuts on P_μ is used to simulate the MuonID at L0-Hlt1 trigger level, the cut on a smeared PT_μ try to emulate the bias introduced by L0 level where the PT resolution is $\sim 20\%$ and the algorithm to calculate PT is different from the offline one. Finally the $PT_{\mu\mu}$ request is a Hlt2 replica.

variables. When the trigger acts only on a muon, the cuts are applied on the daughter with the higher PT.

The proxy cuts performed on MC MinimumBias with 70k $K^0_S \to \pi^+\pi^-$ candidates that have passed the same selection of the signal, are reported in Table 5.1 and the results obtained in the validation phase are shown in Figure 5.6. To better simulate also the reconstruction quality of the real signal events that will be counted in the signal region, only the $K^0_S \to \pi^+\pi^-$ with K^0_S mass into this window are accepted for the proxy signal sample.

It is important to stress that for the validation procedure only the kinematic distributions are looked at and that statistical fluctuations are present, so that it is not mandatory to achieve exactly the same distributions, that would be an adapting to a particular statistical situation and this would decrease the proxy quality. So, the distributions are not exactly reproduced, in particular some peaks are considered to be solely a fluctuation and not a trigger or physical effects. I want to point out that the reproduction of the trigger bias can not be realized with a simple re-weight procedure of the proxy sample using the $K^0_S \to \mu^+\mu^-$ MC because this method will suffer of the possible errors in the simulation and there is no sample from which it is possible to obtain the weight from data (the $\phi \to \mu^-\mu^+$ would not work because of the higher mass with respect to the K^0_S).

From the Figure 5.7, 5.8 and 5.9 it can be seen that after the proxy cuts the $K^0_S \to \pi^+\pi^-$ sample variables distributions become more similar giving confidence on the performance of the method.

For the other trigger categories a similar analysis will be performed with the same approach to create the signal proxy sample. In the TisTisTos it is expected that only the Hlt2 cut replica is needed, while in the TosTos2Tos case more complications can arise for possible bias due to the removal of the shared events already selected by the other two lines. In particular, for triggered muon of high PT the probability to be triggered by both TosTos1Tos and TosTos2Tos lines is higher (see Figure 5.5), thus this effect can be emulated with an upper smeared PT cut.

These considerations are true also for the momentum P. The removal of high PT in the shared events of TosTos2Tos leads to net changes in other variable distributions as shown, for some of them, in Figure 5.5. The four diagrams show the events shared separately between the TosTos2Tos and the other two specific categories (red line for TisTisTos, blue line for TosTos1Tos), the total events of TosTos2Tos (black line) and the same events that are considered in this line when the shared elements are removed (green line). In

particular, it is clear that this effect changes the distribution of the distance of flight, favoring the suppression of events with short flight distance. Morover, as the shared events between TisTisTis and TosTos2Tos are minimal and negligible compared to those between TosTos1Tos and TosTos2Tos allow to restrict the studies to only the shared contributions from the TosTos2Tos events. More studies need to be done and presently the MC statistics for the TosTos2Tos line does not allow any reliable conclusion.

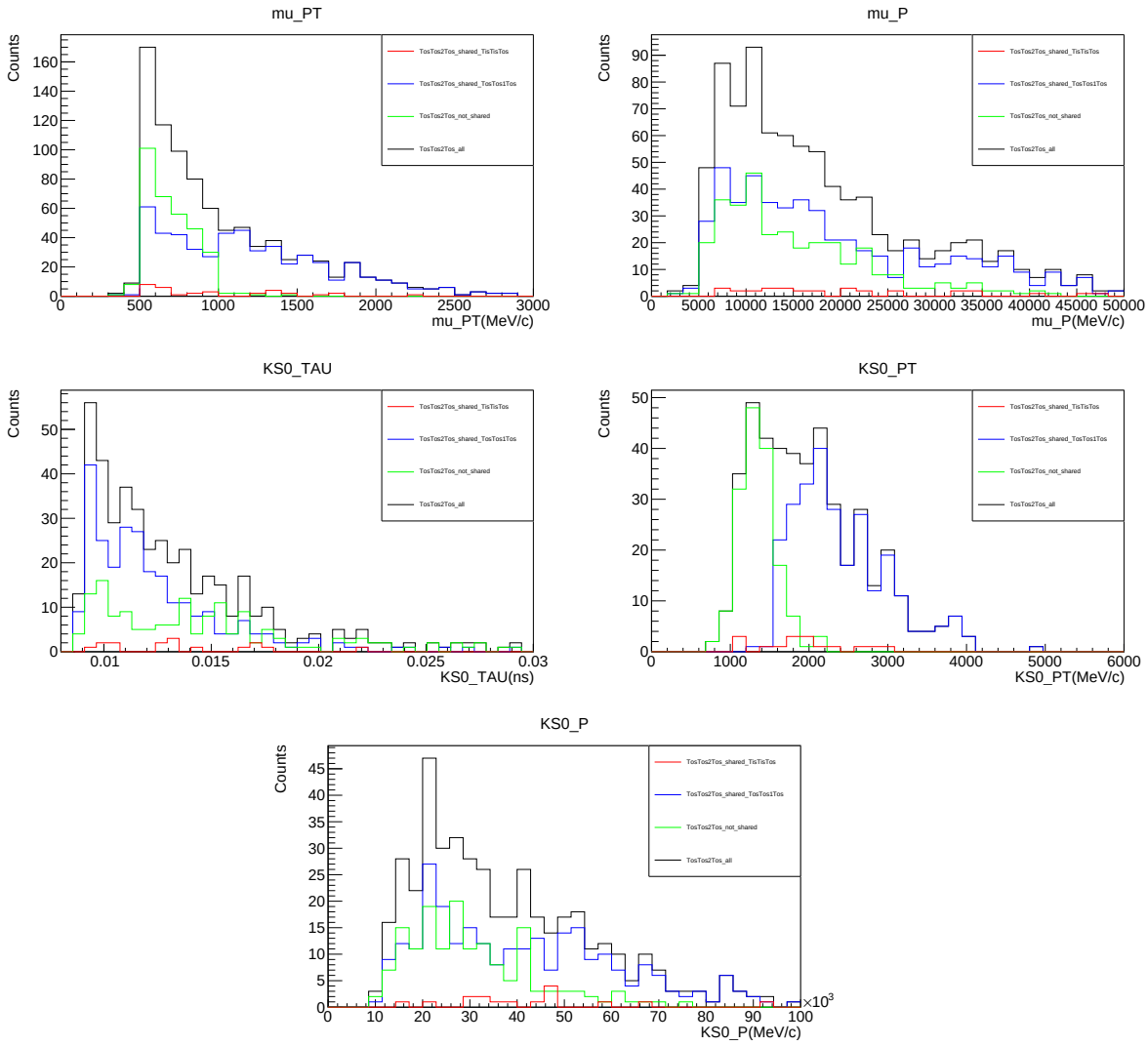


Figure 5.5: The four diagrams show the events shared separately between the TosTos2Tos and the other two specific categories (red line for TisTisTos, blue line for TosTos1Tos), the total events of TosTos2Tos (black line) and the same events that are considered on this line when the shared elements are removed (green line). In order to increase the statistical sample, μ^+ and μ^- distribution have been merged, as the possible discrimination effects between the two from the detector response are eliminated from the fact that there are measures with both the configuration of the magnetic field (B-up and B-down).

To create the proxy, a sample of triggered unbiased data is needed, but, in the real dataset, the statistics achievable in the NoBias lines for $K^0_S \to \pi^+\pi^-$ is too low after the stripping line and proxy cuts. For this reason $K^0_S \to \pi^+\pi^-$ TisTisTis sample are used,

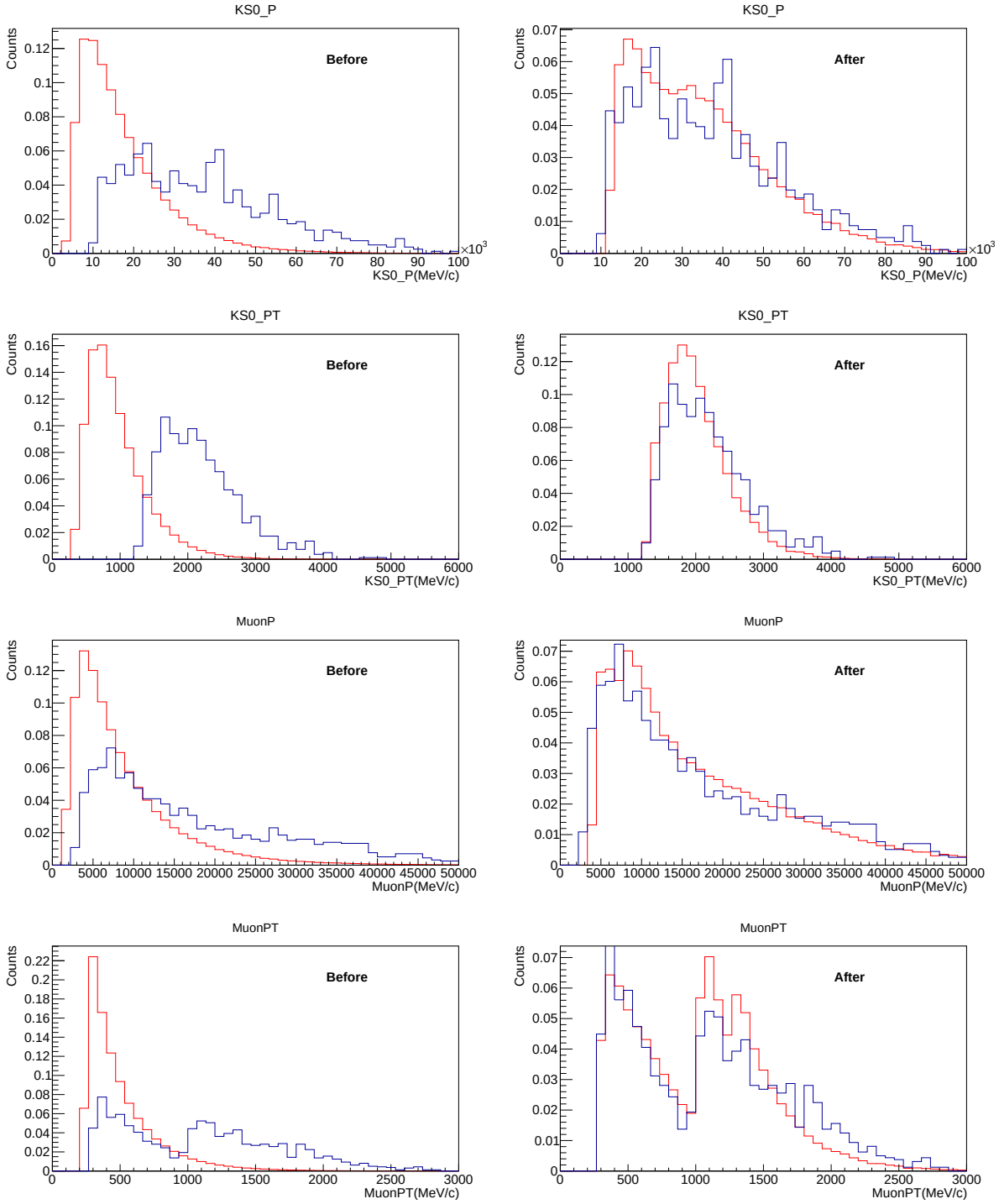


Figure 5.6: Kinematic variables after and before the proxy cuts described in Table 5.1. These variables are used to validate the value of the cuts. **TosTos1Tos** trigger category. Red: Proxy, Blue: Signal.

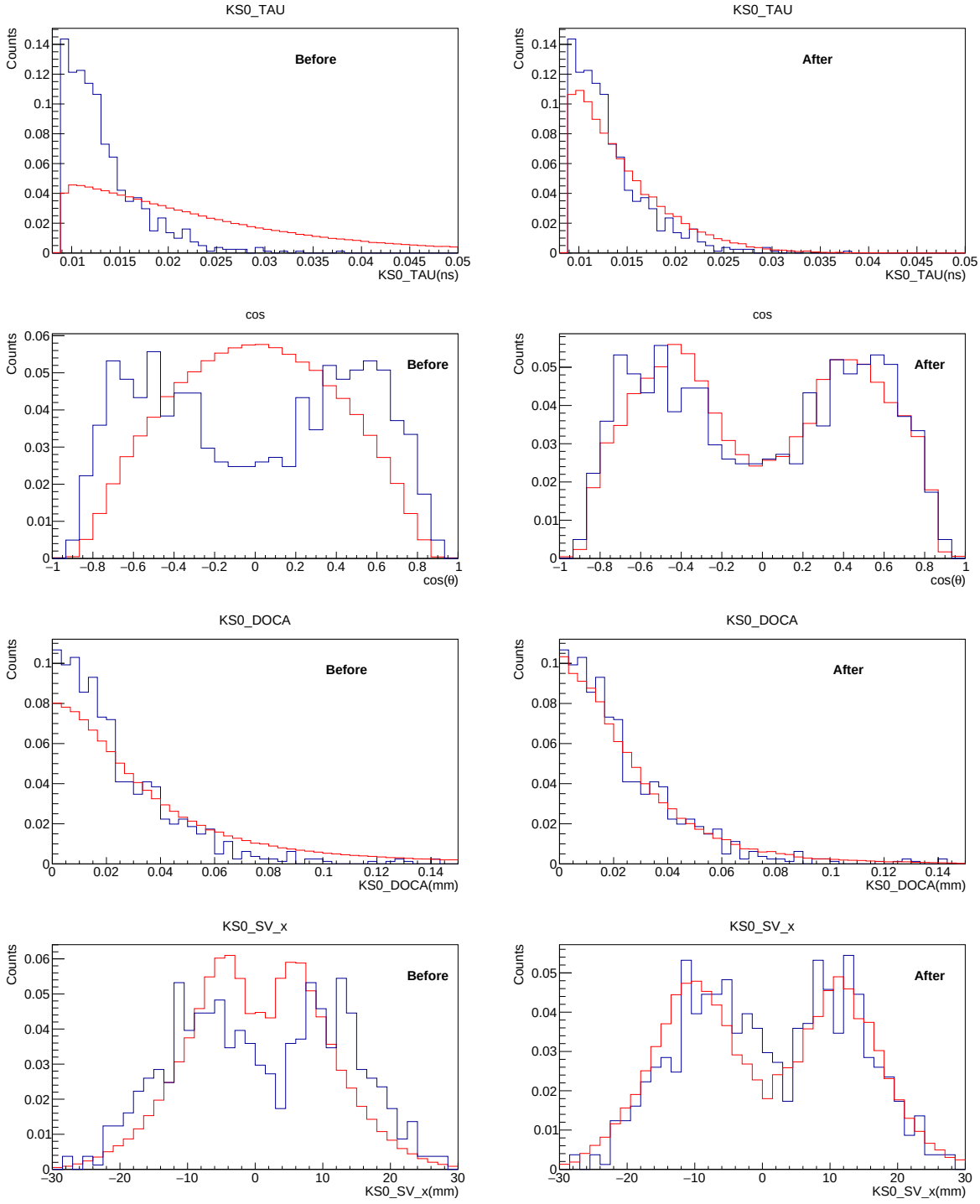


Figure 5.7: Cross check of the MVA variables after and before the proxy cuts described in Table 5.1. **TosTos1Tos** trigger category. Red: Proxy, Blue: Signal.

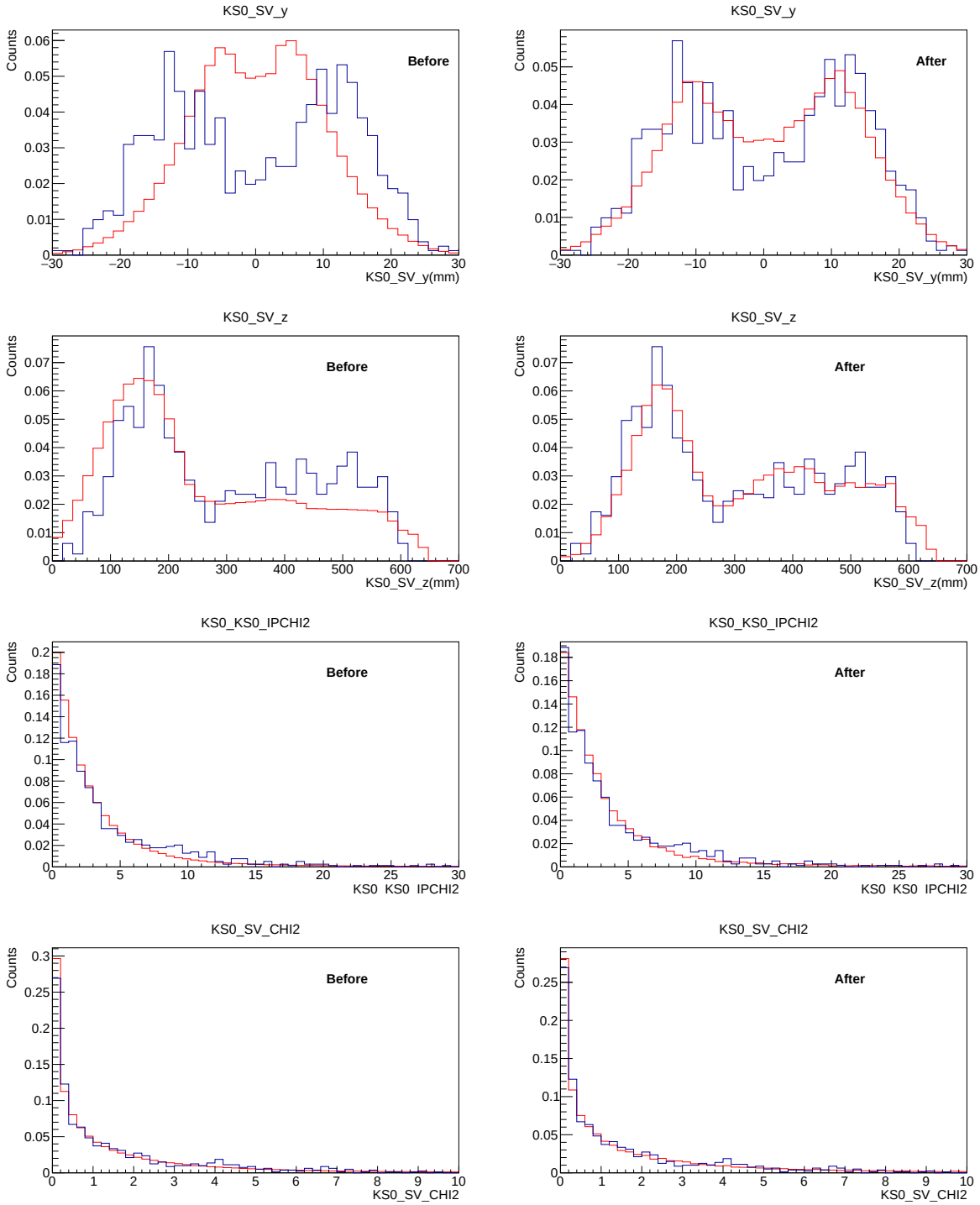


Figure 5.8: Cross check of the MVA variables after and before the proxy cuts described in Table 5.1. **TosTos1Tos** trigger category. Red: Proxy, Blue: Signal.

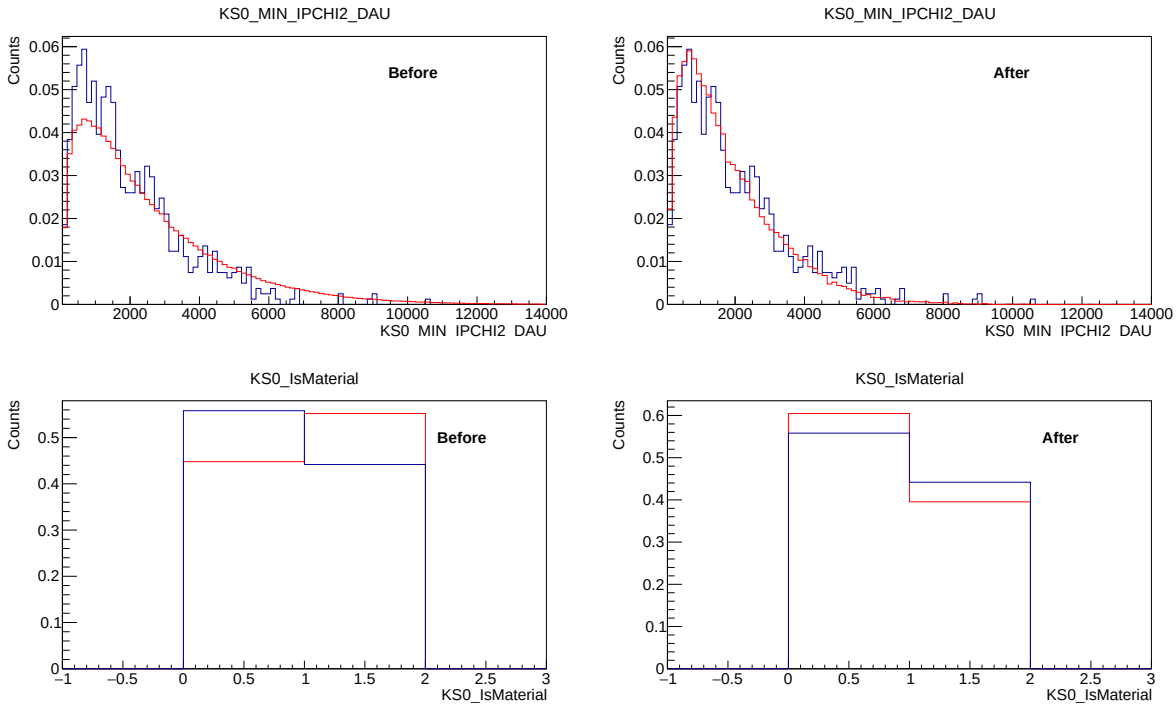


Figure 5.9: Cross check of the MVA variables after and before the proxy cuts described in Table 5.1. **TosTos1Tos** trigger category. Red: Proxy, Blue: Signal.

removing their very small trigger bias with a re-weight on the PT of the K^0_S with the one obtained from the NoBias line (see Figure 5.10, 5.11).

In the figures are reported the NoBias (black) and the TisTisTis distributions before (red) and after (blue) the re-weighing and so the plots are read looking at the difference in the comparison of the blue line with the NoBias and the red line with the NoBias. Thanks to the re-weighing the distributions of the variables, which already were not very different from the ones unbiased by the trigger, become even more similar. This aspect is particularly evident for the time of flight variable. The re-weighing procedure has been proven in both P bins and PT ones separately, so that the best one could be used (the PT one). After their calculations, the weights are passed to the TMVA package that counts the events using their weigh. Then, the proxy cuts are applied and the sample is used as training in the MVA.

Thanks to this procedure real data sample can be used as proxy to train the MVA, avoiding the use of MC. The signal sample is obtained from the procedure explained so far, while for the combinatorial background the right sideband $K^0_S \to \mu^+\mu^-$ data is used with the exclusion of the ϕ mass range ($\sim$ 1020 MeV/c^2).

Two regions can be defined, one with K^0_S mass in [600, 950] MeV/c^2 and another with K^0_S mass above 1100 MeV/c^2. The MVA input variables exhibit a dependence on the mass (see Figure 5.12, 5.13) of the candidate so, to have a more realistic proxy of the background expected in the signal region, only the region closer to the signal window is used. In these figures profile histograms are shown displaying the mean value of a variable and its error for each bin in the K^0_S mass. The leftside part (below $\sim$ 500 MeV/c^2) should not be considered as it is contaminated by $K^0_S \to \pi^+\pi^-$ tail as it is clear, e.g., looking at the IsMaterial

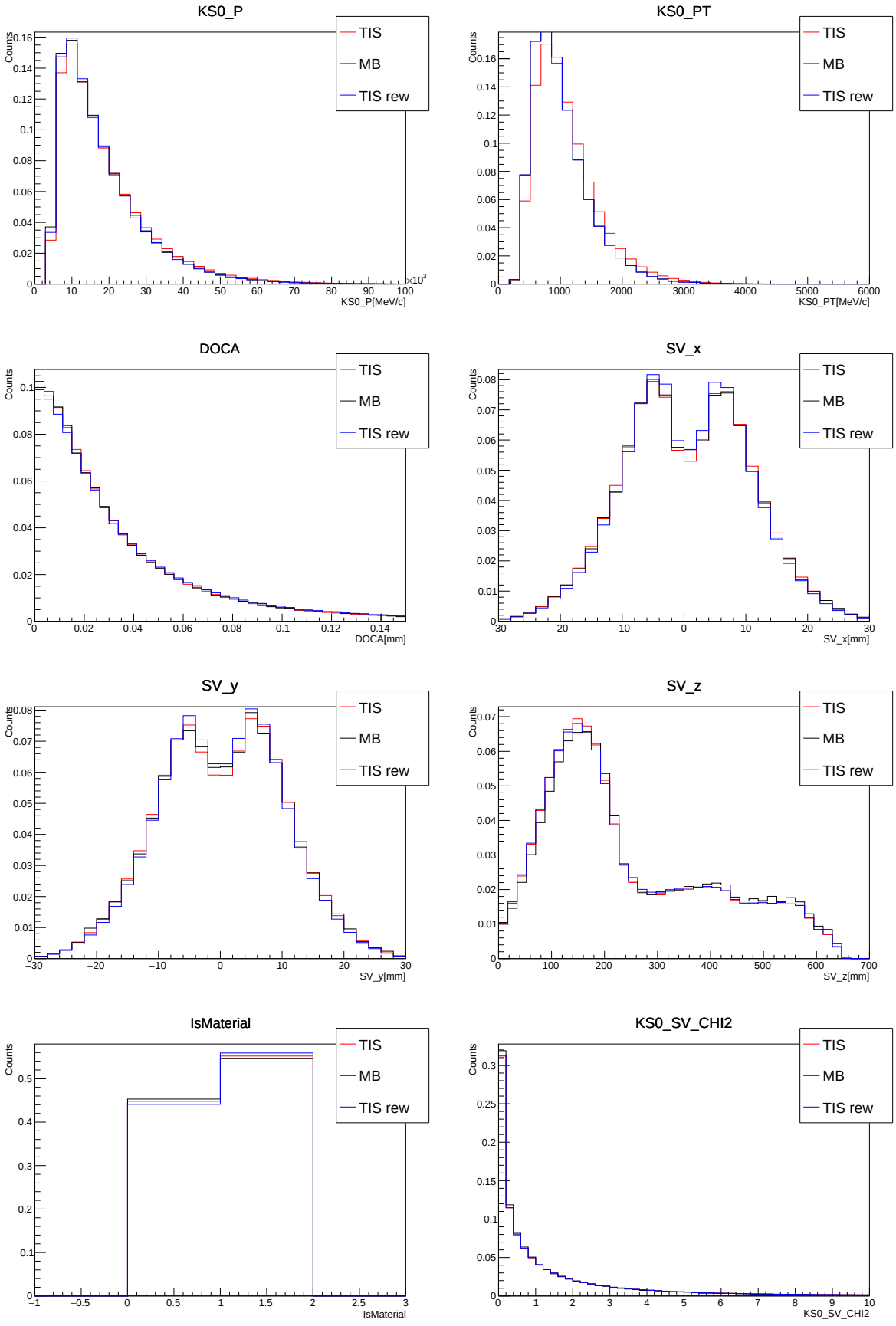


Figure 5.10: TosTos1Tos trigger category. The TisTisTis samples is re-weighted in the transverse momentum of K_S^0 coming from NoBias sample to correct the small trigger bias present and achieve a bigger statistics for the proxy sample in data with respect to the NoBias itself. It can be seen from the picture that the re-weight has the effect of make the distribution of the two sample closer

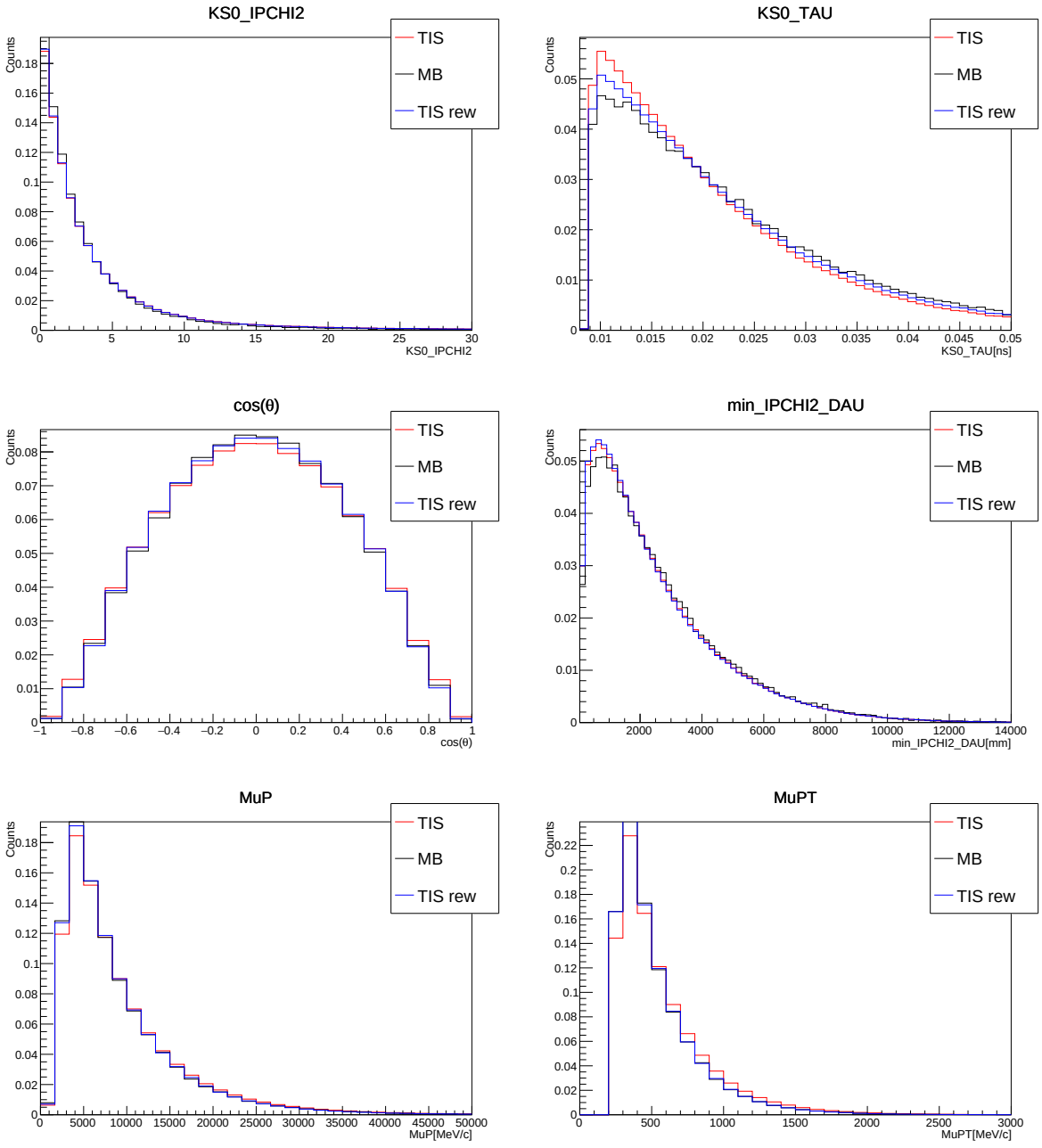


Figure 5.11: TosTos1Tos trigger category. The TisTisTis samples is re-weighted in the transverse momentum of K^0_S coming from NoBias sample to correct the small trigger bias present and achieve a bigger statistics for the proxy sample in data with respect to the NoBias itself. It can be seen from the picture that the re-weight has the effect of make the distribution of the two sample closer proving the procedure.

variable which goes to zero in the leftside because of the presence of real events that do not arise from the interaction with the material. The Ismaterial variable has also a contribution that goes to zero (vertex not in the material) in the region further to the right, indicating a contribution different from that of the combinatorial background present in the signal region. It can be seen from the graphics that certain variables vary with the mass more than others but the only relevant aspect is that for high mass the reconstruction of the SV loses in quality because events are strongly different from the signal for which the selection is optimized. This thing can be seen in the trend of DOCA and SV_{χ^2}.

The lower limit of the region is 600 MeV/c^2 to not use the same data from which the expected background in the signal region will be estimated (see Section **??**) and to have a low contamination from the $K^0_S \to \pi^+\pi^-$ tail.

It is important to notice that a small dependence of the MVA output from the mass can still be present without invalidating the procedure, but simply resulting in a slightly suboptimal combinatorial background rejection (due to the fact that the combinatorial proxy sample used do not exactly recreate the real one present in the signal region).

5.3 Tuning, Training and Check of the Overtraining

In summary, the training samples of the BDT are the right sidebands [600, 950] MeV/c^2 for the background and the $K^0_S \to \pi^+\pi^-$ TisTisTis, re-weighted to reproduce the NoBias events, with some cuts applied to simulate the muon trigger effects.

All the performance results here reported are obtained during the tuning phase, where there is enough statistics for the test sample, which is independent from the training procedure. At this stage the MVA studies are split into three parts: a first one where the tuning of the classifier is developed followed by the overtraining and mass dependency cross check, then the final training and application to the data is realized. In the tuning phase, to achieve reliable performance results, the proxy samples for both signal and background are randomly split into two parts, one used for train and the other to test the MVA.

The distributions of the input MVA variables for both background (red histograms) and signal (blue histograms) proxies are shown in Figure 5.14 and, for few of them, in Figure 5.15 the scatter plots illustrate the non linear dependency between these variables. From those plots it is clear that it is not possible to discriminate the two types of events using simple cuts and, due to the non linear dependence between the input variables, the best discrimination cannot be achieved by a linear classifier (such as the Fisher discriminator) and the necessity of a multivariate non-linear classifier is evident.

If two variables are totally correlated the presence of both does not add informations (and then discriminating power) to the classifier with respect to the use of only one variable. It can happen for some classifiers (e.g. the neural networks) that the presence of strongly correlated variables could end up with a reduction of the performances due to the way the classification algorithm is developed. In Figure 5.16 the linear correlations of the input variables of the MVA are shown to check if we are introducing useless variables. It is possible to see a strong correlation between DOCA and SV_{χ^2} as both refer to the SV. The effect of the introduction of only one of these variables on the MVA output has been checked and it was noted a slightly improvement when both the variables are used instead of the case where only one of them is considered. So we decided to keep both of them in. With regard to the other variables no particular correlations that worths a further investigation was considered, concerning both the signal and the background, because all the other correlations are well

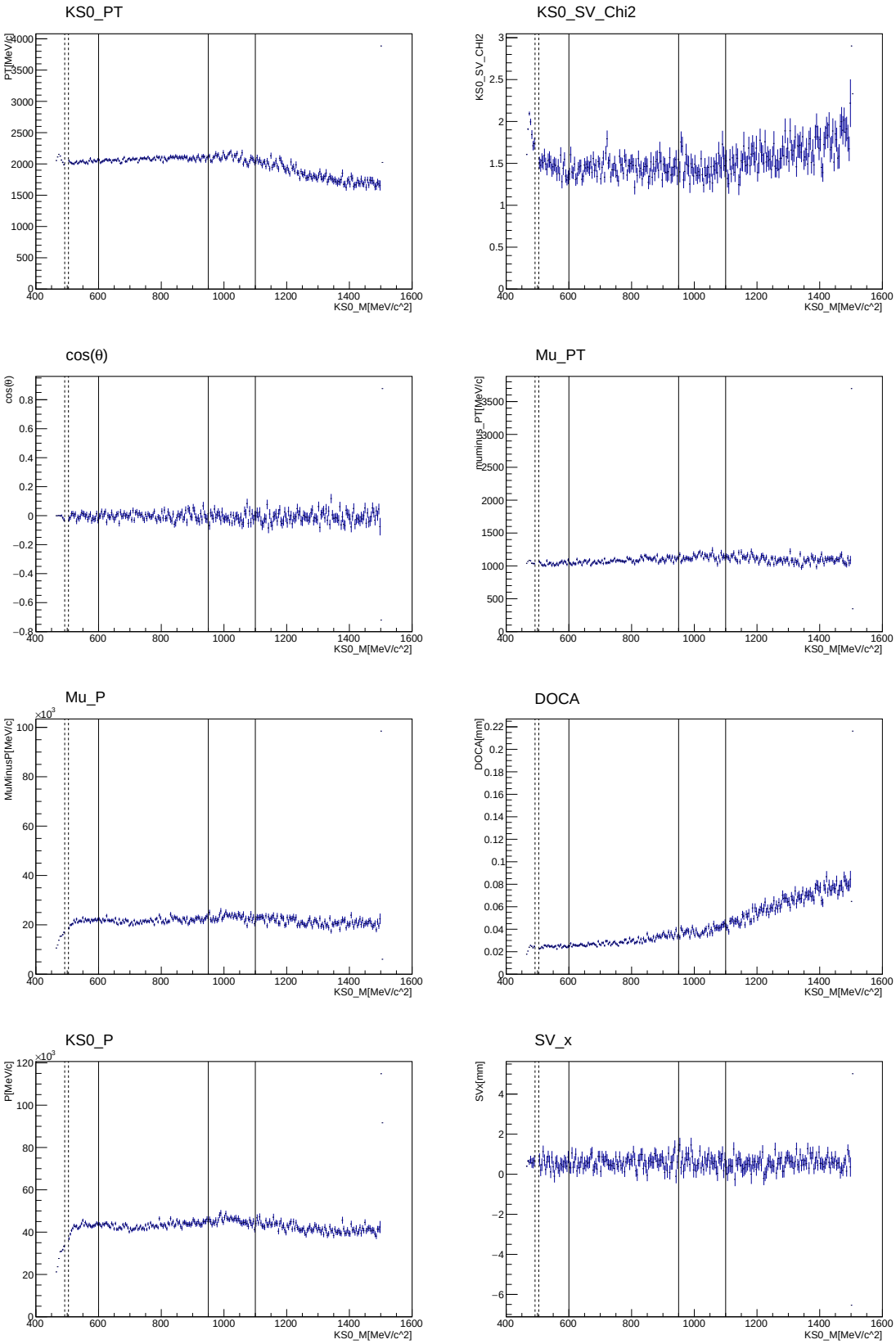


Figure 5.12: TosTos1Tos trigger category. In the pictures is shown the dependence of the input MVA variables for the sidebands $K^0_S \to \mu^+\mu^-$ from the mass of the K^0_S. The blinded signal window is identified by the dashed lines, while the two defined regions between [600, 950] MeV/c^2 and above 1100 MeV/c^2 are delimited by straight lines.

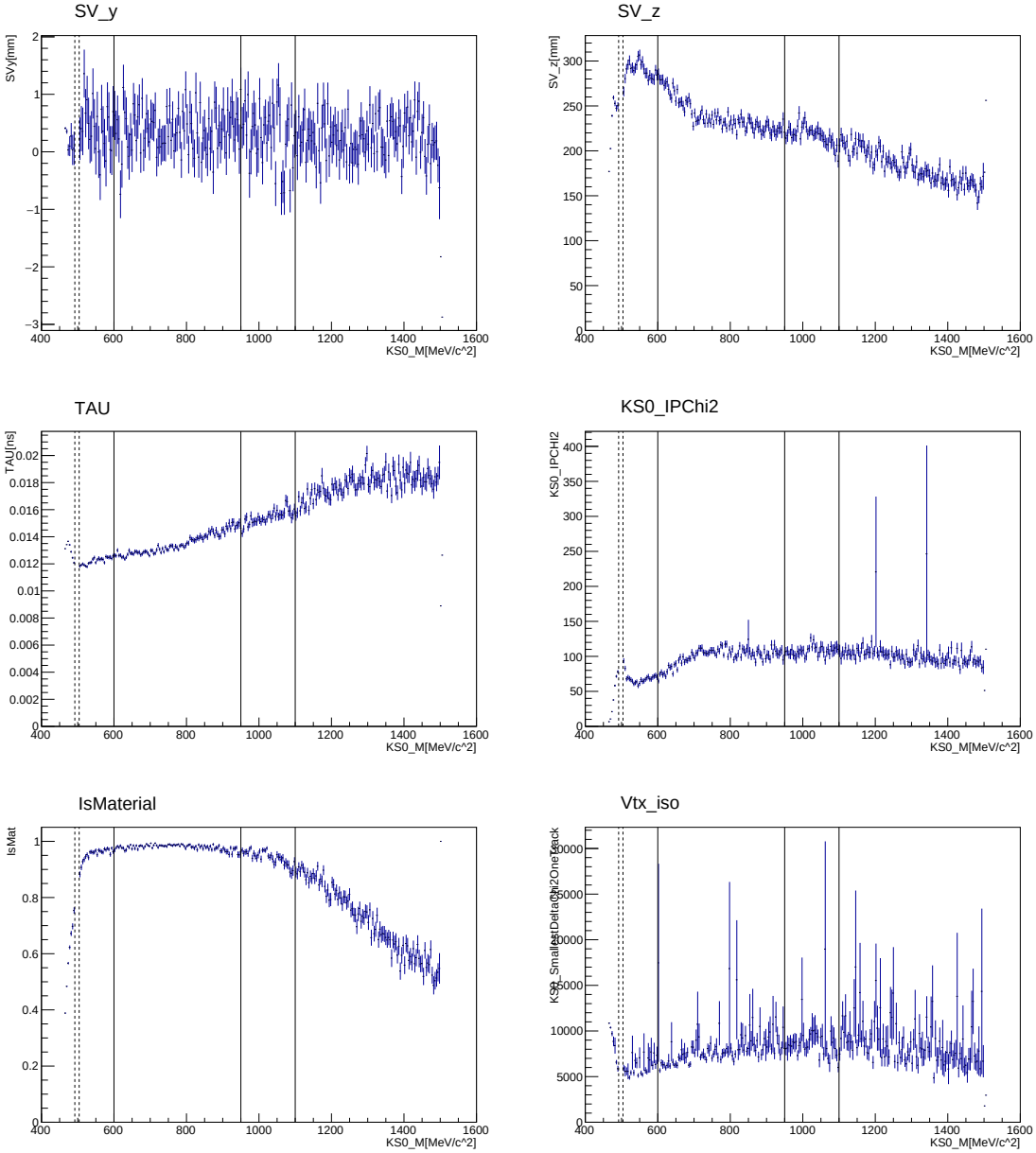


Figure 5.13: TosTos1Tos trigger category. In the pictures is shown the dependence of the input MVA variables for the sidebands $K^0_S \to \mu^+\mu^-$ from the mass of the K^0_S. The blinded signal window is identified by the dashed lines, while the two defined regions between [600, 950] MeV/c^2 and above 1100 MeV/c^2 are delimited by straight lines.

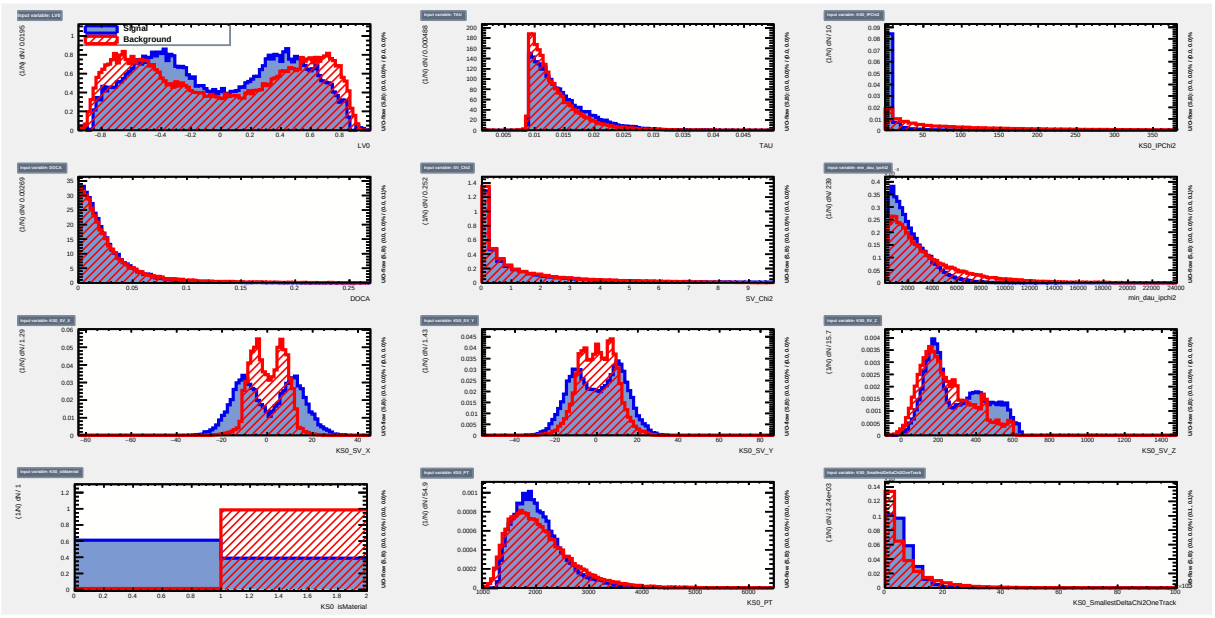


Figure 5.14: Inputs MVA variables in the tuning phase. **TosTos1Tos** trigger category. Red: Background, Blue: Signal. Not all the variables can be used at the same time to do not let the MVA to learn the K^0_S mass.

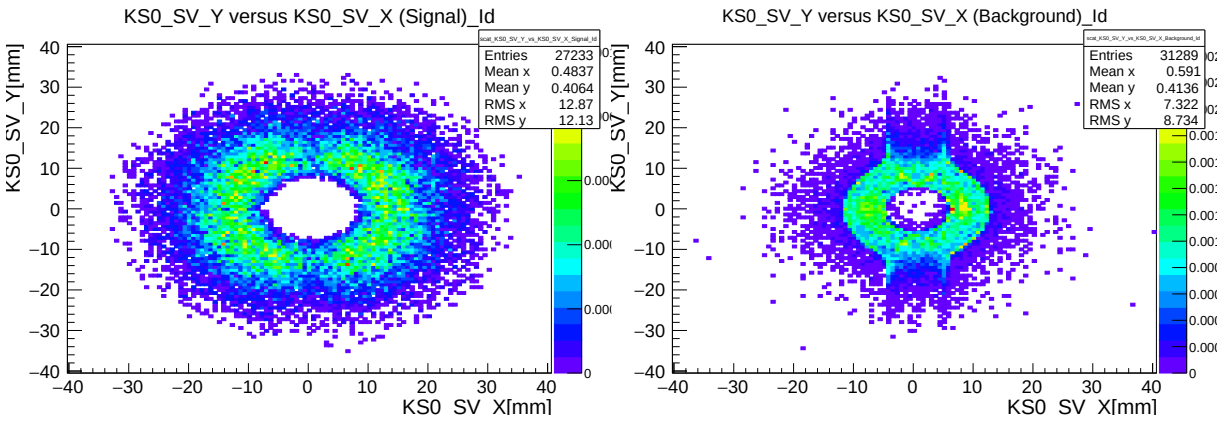


Figure 5.15: Scatter plots for signal (left) and background (right) proxy samples for few of the inputs MVA variables in the tuning phase to demonstrate the need for a non linear classifier to reject the background contributions. **TosTos1Tos** trigger category.

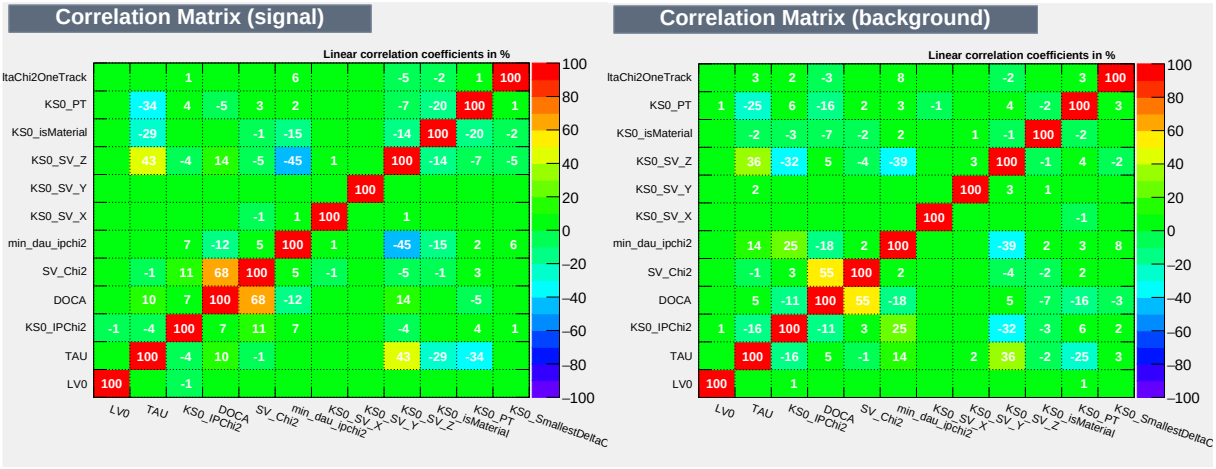


Figure 5.16: Correlation matrices for the inputs MVA variables in the tuning phase for the signal (left) and for the background (right). **TosTos1Tos** trigger category.

explained and not as extreme as the one between the two variables aforementioned. For example the z coordinate of the SV and the τ are positively correlated as particles with a greater time of flight correspond to a SV more detached from the PV. It is important to note that all the correlations concerning the background are inhibited as it can be deduced from the fact that these tracks are not real $K^0_S \to \mu^+\mu^-$ or physics processes and so their correlations are not dictated by the decay constrains.

From the tuning phase the best classifier is chosen out of the many MVA algorithms of classification that have been studied in this thesis.

To achieve the best background rejection for a fixed signal efficiency the parameters of these algorithms need to be tuned. In the tuning phase more MVA algorithm of classification for the four different input variables combinations are studied and tested:

- many Likelihood methods (Projective Likelihood Estimation, Multi-dimensioal Likelihood Estimation, K Nearest Neighbors)
- Boosted Decision Tree
- Neural Networks with multi layer perception (MLP)
- Fisher discriminant.

There are many likelihood methods developed in the TMVA package, each one differs for the way to extrapolate the probability densities estimation of the variables from the dataset. The k-NN algorithm was found to get a big improvement with the tuning of its parameters, while the others were found to do not perform well in every parameters configurations, the default too.

A single decision tree is a binary tree structured classifier made of sub-sequential cuts on one single variable at a time until a stop criterion is fulfilled. They are in this respect similar to rectangular cuts. However, whereas a cut-based analysis is able to select only one hypercube as region of phase space, the decision tree is able to split the phase space into a large number of hypercubes. Indeed, the variables space is split this way into many rectangular regions that are eventually classified as signal or background, depending on

the majority of training events that end up in the region. A single decision tree suffers of instability with respect to statistical fluctuations of the training sample from which the tree structure is derived. For example, if two input variables exhibit similar separation power, a fluctuation in the training sample may cause the tree growing algorithm to decide to split on one variable, while the other variable could have been selected without that fluctuation. Boosting stabilizes the response of the decision trees with respect to fluctuations in the training sample and is able to considerably enhance the performance. Boosting is a way of enhancing the classification (and increasing the stability with respect to statistical fluctuations in the training sample) by sequentially applying an MVA algorithm to re-weighted (boosted) versions of the training data using the information from the previous step (the previous MVA) about the misclassifed events which are given a higher weight in the training of the following tree. The final MVA is a weighted sum of all the MVA created in the boosting sequence with the weight that enhances the contribution of MVA with a better discrimination power.

Different kind of boost algorithms (such as Gradient boost or AdaBoost) are tried with the AdaBoost one that gives the best response. In general, the boosting process performs better on weak classifiers such as a single decision tree with a tree depth of 2 or 3, that have very little discrimination power by themselves with the performance further enhanced by forcing a slow learning rate and allowing a larger number of boost steps. The parameters of the BDT classifier are tuned considering all these informations.

The parameters that have been tuned are: the number of boosting steps, the max numbers of repetitive cuts for a signal decision tree, the minimum percentage of training events required in a final region and the learning rate for AdaBoost algorithm.

BDTs are sometimes referred to as the best out of the box classifiers because little tuning is required in order to obtain reasonably good results. Decision trees are also insensitive to the inclusion of poorly discriminating input variables.

Their implementation on the TMVA package are explained in [128]. When the tuning phase is completed the best classifier and variables combination are chosen and used to perform the MVA. To choose a particular MVA algorithm, the proxy samples was divided into two part, the first one used for training and the second for test.

Many of the classifiers present in the TMVA tool have been tried to take more confidence with these analysis methods which are having an increasing popularity in HEP. Most of these, in comparison to the best obtained, are very weak in discriminating between signal and background when a high rejection of the background is required and so these classifiers have not been reported in the plots of this section. For comparison also the default TMVA classifiers were considered to understand if the introduced parameters optimizations was able to give us a better performance (in the figure they are labeled with BDT and BDTD for the Boosted Decision Tree algorithm case). From these comparisons it can be noted, for the BDT case, that, although many attempts to vary the parameters of the classifier, the result gives no notable differences, giving a substantial proof of the robustness of the method. The same result is not true for the the other tested methods, for which only the best achieved parameters option setting is here reported, as they give a lower performance with respect o the BDT one.

The main tool to check the performance of a classifier and of the input variables used, that are included in the TMVA package, are: the area of the Received Operating Characteristic (ROC) curve (in an range chosen by the user), the ranking of the variables (e.g. in the BDT the ranking is based on which variable was most used in the classification process).

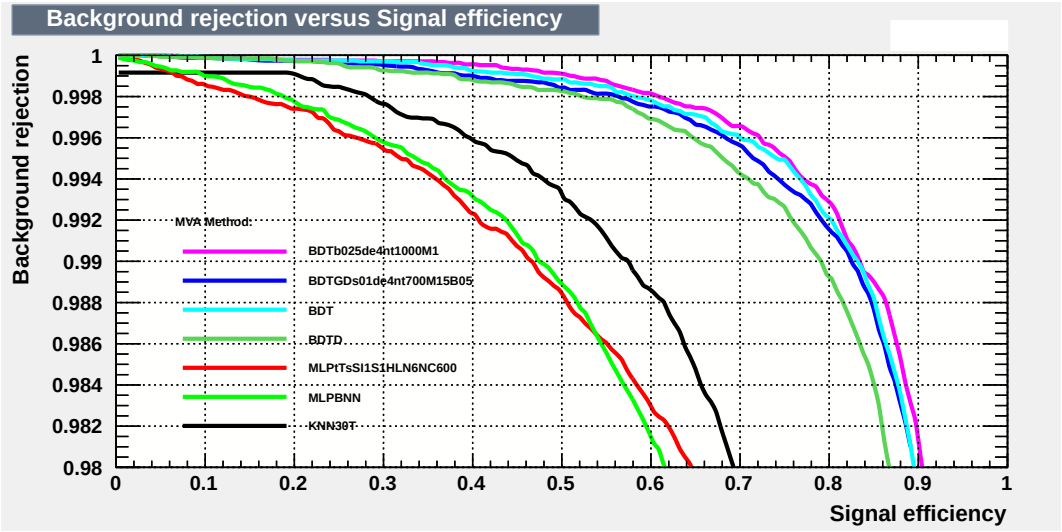


Figure 5.17: ROC curves for all the tuned methods for a fixed input variables choice (the one that includes the $PT_{K^0_S}$ and the SV coordinates). **TosTos1Tos** trigger category.

In general, in HEP the important quantities when a selection is applied on a sample to separate the signal from the background are the signal efficiencies, the background rejection and the purity. The first one is defined as the fraction of signal events that are selected by the cut, the second is the fraction of background which is removed with the cut and the last is the fraction of signal events present in the sample obtained after the cuts. The ROC curve is a plot that reports the background rejection (percentage of the background events removed with this cut) as a function of the signal efficiency achieved for different cuts on the MVA output and it is calculated using the testing sample. Other ways to check the performance, such as the rarity curve, are not so clear and powerful as the ROC curve and so they have not been taken into consideration. For all the four variables combinations the best (larger ROC area) classifier is the BDT. In Figure 5.17 the ROC curves for some of the MVAs tested for the combination that includes $PT_{K^0_S}$ and the SV coordinates are shown.

Choosing different combinations of the input variables, without letting the MVA learning about the mass candidate, can give a different response. In Figure 5.18 the comparison of the ROC curves of the four variables combinations is illustrated. The slightly better situation is obtained when $TAU_{K^0_S}$, $PT_{K^0_S}$, IsMaterial, SV z coordinate are used if a high background rejection is required. It can be seen that the option taken (let IsMaterial in the BDT) is the one leading to the best rejection power, although all the approaches would lead to very similar results, as the performaces in the high rejection region are almost identical.

The dependence on the K^0_S mass is checked comparing the output of the MVA for different masses (see Figure 5.19) using the right sideband data and taking in account intervals of 12 MeV/c^2 for different values of the reconstructed mass (600, 750, 900, 1300) MeV/c^2. Despite the effort to minimize it, it can be seen that there is still a dependence, but we can state that better results have been achieved in the final real signal efficiency with respect to the old analysis, where also the region exceeding 1100 MeV/c^2 was used for the MVA studies. As the aforementioned dependence is present in all the four different combinations of variables studied and it does not give us any clear clue about which variables combination should be used to give the lower mass dependence and so the choice is left only to the ROC curve.

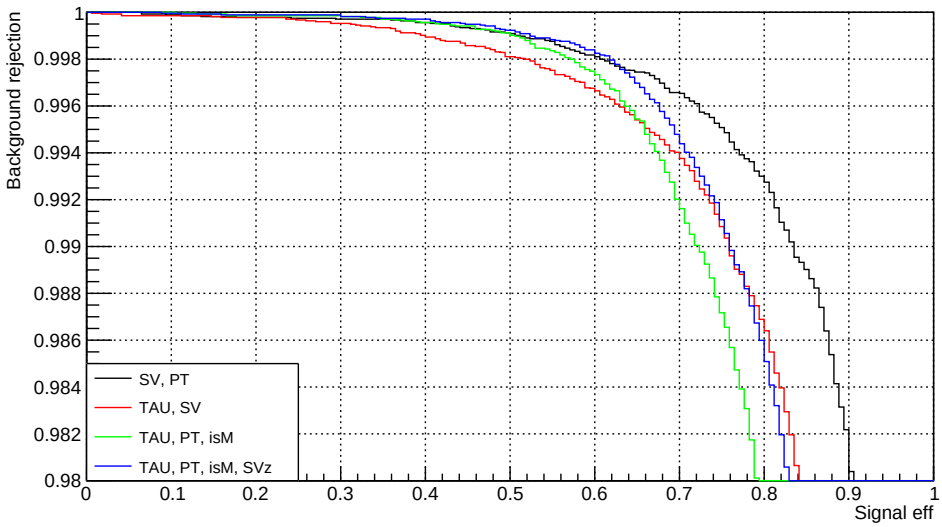
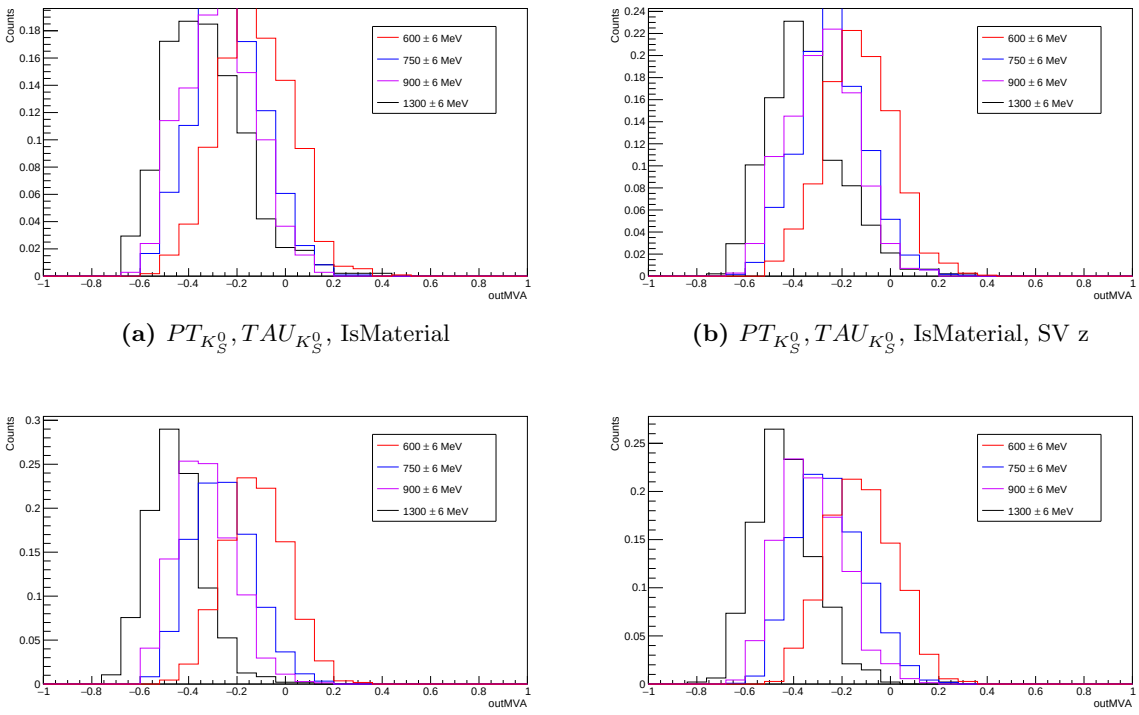


Figure 5.18: Comparison between the four variables combinations for the best tuned classifier. The response for the four cases is very similar in the high background rejection region which is the region we decided to use for the analysis. **TosTos1Tos** trigger category.



(a) $PT_{K^0_S}, TAU_{K^0_S}$, IsMaterial

(b) $PT_{K^0_S}, TAU_{K^0_S}$, IsMaterial, SV z

(c) $PT_{K^0_S}$, SV coord.

(d) $TAU_{K^0_S}$, SV coord.

Figure 5.19: TosTos1Tos trigger category. BDT distribution for different variables combinations and for different mass ranges. Intervals of 12 MeV/c^2 around the mass values of 600 (red), 750 (blue), 900 (violet), 1300 (black) are reported to check the dependence of the BDT output from the K^0_S mass.

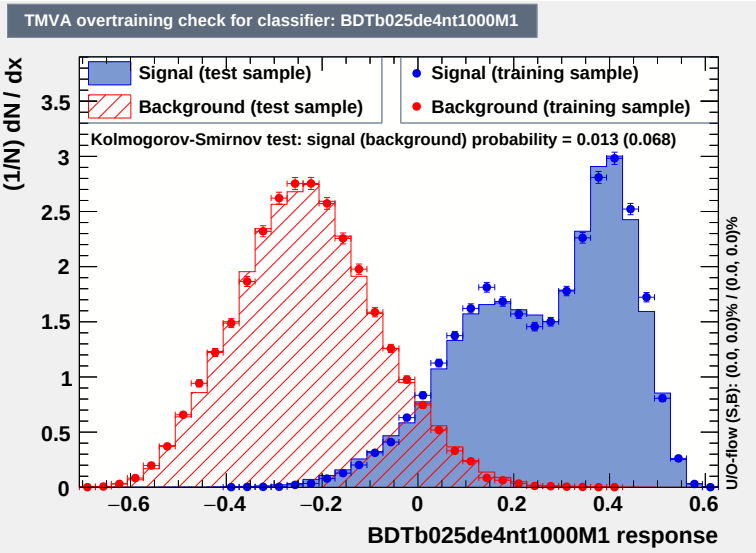


Figure 5.20: Distribution of the signal and background MVA output for both the training and test phase during the tuning phase to cross check the possibility of overtraining effects. **TosTos1Tos** trigger category. Red: Background, Blue: Signal.

Both the proxy samples have a limited statistics and so they suffer from fluctuations. Due to the statistical fluctuations, it is possible that the train procedure of the MVA could adapt the learning process to these fluctuations, giving a fake high discrimination power that would not be present in another independent sample (as is the real signal in the blind region). In this situation the real efficiency is lower than the one achieved looking at the training sample. This effect is called overtraining and it is checked by looking at the MVA output distribution of the training sample and of another independent sample called test sample. Indeed, the overtrainig leads to a seeming increase in the classification performance over the objectively achievable one when measured on the training sample and to an effective performance decrease when measured with an independent test sample. This check is done for all the MVA outputs during the tuning phase and in the Figure 5.20 it is reported only the final case. As the cross check between the training and the test sample seems to give no hints of possible overtraining a 90% fraction of the full proxy samples is used as training and the remaining as test. Moreover, a possible overtraining would not bias the measurement because the background sample used for the training (all Tis) is independent from that used in the normalization (NoBias) and in the background extrapolation (different sideband range). The output of the BDT (see Figure 5.20) demonstrates the validity of the MVA to discriminate signal from background.

A cut on the BDT output is chosen to have a high rejection efficiency ($\sim 99.8\%$) against combinatorial background. For this cut the signal efficiency is $\sim 66.2\%$.

The decision of the cut on BDT has been taken by reference to the previous analysis work of 2011, which aimed at reducing drastically the contribution of the background. The expected amount of combinatorial background has not an easy estimation as it does not depend on a particular physical process and so no BR has been associated with it, but from the data it is clear that its contribution is more than 3 orders of magnitude higher than the signal and the cut decision needs to take this into account. The background contamination and signal efficiency is not the same for the type of events accepted but it depends on the

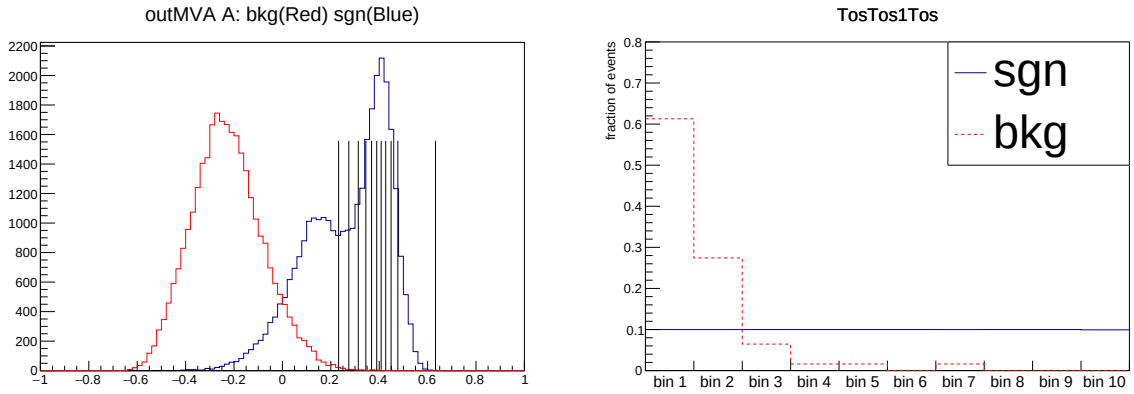


Figure 5.21: On the left the boundaries of the bins in the BDT output are shown with each bin containing 10% of the total events that passed the cut in the BDT (the left boundary of the first bin). On the right the signal and background percentage with respect to the all passed events entering in each bin are shown. **TosTos1Tos** trigger category.

MVA output value. To further exploit the information provided by the MVA classifier for the selected events, the analysis is performed in 10 bins of the BDT output, with increasing signal/background ratio. The ten bins are chosen in such a way to contain almost the same number of events (so about 10% of the events that have passed the BDT cut). Each bin is treated as independent in the rest of the analysis. The bin division is illustrated in Figure 5.21, where it is clear that a different amount of background contamination is present for each bin. This leads the bins to have different sensibility. So considering the other two trigger paths samples, after the whole MVA phase there are 30 independent samples (ten for each category) for which the normalization factors and the background contribution are calculated.

The mass distribution of the events that have passed the BDT cut is shown in Figure 5.22 for each bin. From the illustration it is clear that the background contribution is reduced not only the right sideband where combinatorial background prevails in the training phase but also in the left sideband where the $K^0_S \to \pi^+\pi^-$ tail is still present. Even if the BDT is not effective for rejecting the double misID, it has an effect in the size of the tail entering in the signal region because there is a correlation between the mass resolution and the quality of the reconstruction. Indeed, high BDT output double misID is composed mainly by candidates of very good quality, i.e., in which the tracks are hardly affected by the decay in flight. This implies that the momentum resolution is also better, so there is a lower probability to enter in the signal region.

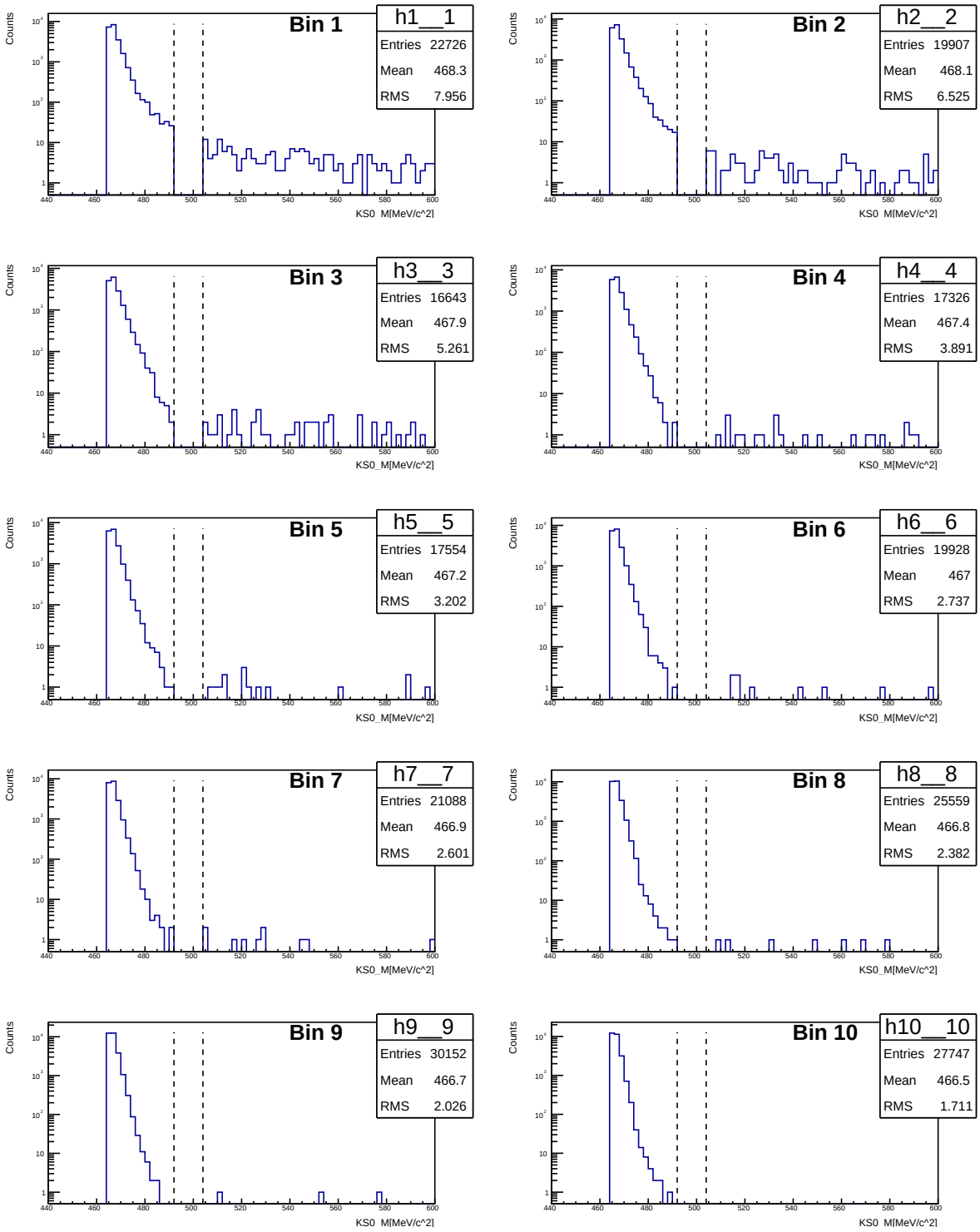


Figure 5.22: TosTos1Tos trigger category. The K_S^0 mass distribution for all the 10 bins in which the data are split after the BDT cut. The signal region is blind and it can be seen that the combinatorial background (right side) is reduced after the BDT cut if we compare these distributions to those shown in Figure 3.4.

Chapter 6

Final selection and outlook

In this chapter the conclusion of this work of thesis are given using the results from the preliminary estimations of the background inside the signal region for each of the ten bin after the application of muon identification requirement, that is optimized to maximize the signal sensitivity. An outlook for the subsequent steps of the analysis is given in Section 6.3, focusing on the normalization procedure which allows converting the observed number of $K_S^0 \to \mu^+\mu^-$ candidates into a branching fraction.

6.1 Muon Identification

After the strong reduction of the combinatorial background through the dedicated MVA analysis, the background is composed mainly by $K_S^0 \to \pi^+\pi^-$ decays with pions misidentified as muons. The request that the daughter particles are identified as muons in the stripping line consists only on a loose match between the track extrapolation and hits in the muon stations (IsMuon condition, detailed in Section 2.4). In other words, to the daughters tracks is associated the mass of the muon but there is no application of a strong muonID request in the characteristic of the tracks. The IsMuon requirement has an high efficiency (more than 97% [129]) and also implies a momentum above 3 GeV/c. The standard MuonID algorithm was optimized for high efficiency without focusing on the combinatorial background rejection which is still important at low momentum (< 8 GeV), where not all muon stations are reached. This kind of misidentified pions affects the $K_S^0 \to \mu^+\mu^-$ candidates, mostly in the left sideband.

To reduce the $K_S^0 \to \pi^+\pi^-$ contribution, a cut in an improved muonID algorithm is applied, increasing the discrimination between muons and pions that decay in flight before the muon stations. In this analysis the muonID algorithm has been optimized using low-momentum pions from well reconstructed $K_S^0 \to \pi^+\pi^-$ decays as a model of background. A new dedicated MVA classifier, trained on real data, has been developed (I gave no contribution on this development) using the present muon classifier plus twenty new variables that can have a different distribution between muons and pions. The additional variables, not used by the standard LHCb classifier, include the timing of muon station system, the quality of the match in the two detector views (x and y), the cluster size of muon hits, some isolation variables, and the result of extending the Kalman-filter based tracking algorithm to the muon stations.

The background contribution enter with a different intensity into each BDT bin due to the fact that higher BDT values have a better track quality and in fact in the last bins

the contribution of the $K^0_S \to \pi^+\pi^-$ tail is lower than the first ones. For each bin an optimized cut has been studied following the Punzi figure of merit [130]. It is particularly suitable for optimization in new physics searches, being independent of a-priori expectations about the presence of a signal, thus allowing the determination of a single set of cuts that is optimal both for setting limits and for making a discovery. In our case we are dealing with a counting experiment in presence of background. If n is the number of events observed, which is Poisson-distributed with a mean determined by B (the expected number of background event) and the possible contribution of signal events S. Both the background B and the signal events S will depend on the cut applied. The optimal cuts to to achieve the best sensitivity to the $K^0_S \to \mu^+\mu^-$ decay is obtained maximizing the formula $\frac{S}{x/2+\sqrt{B}}$ where x is the number of sigmas corresponding to the target significance of a signal evidence (it is custumary to use x=3). We can compare this formula with two other common used formulas such as $\frac{S}{\sqrt{B}}$ and $\frac{S}{\sqrt{B+S}}$. The latter expression cannot be maximized without knowing explicitly the cross section for the searched signal and so it need a guess from some theoretical prediction (but if we use the SM, the expected one is simply zero and so the formula does not work); the same is not true for the first two expressions where the result only depend on by the signal efficiency which is somehow measurable. The problem with the $\frac{S}{\sqrt{B}}$ one is that it breaks down at small values of B an so the maximization of the ratio may push the experiment efficiency down to very small values (e.g. it prefer an expectation of 0.1 signal events with a background of 10^{-5} over a situation with 10 signal events expected and a background of 1 event). All the three mentioned formulas behave as $1/\sqrt{B}$ at large B.

So I have used this approach to decide the cut on the muonID variable trying to maximized the quantity $\frac{\epsilon}{3/2+B}$. In a fixed bin, for each cut in the muonID the signal efficiency and the background contribution are calculated, the latter obtained with a fit of the sidebands in the range [470, 492]-[504, 600] MeV/c^2 and then calculating the number B of background events expected from the fit in the signal region. This calculation is is a preliminary estimation of the expected background level to deduce the improvement in sensibility of the 2012 LHCb data with respect to the 2011 situation.

Rather than relying on MC simulations, the signal efficiency ϵ is calculated using a sample of pure muons from the $B^+ \to J/\psi(\mu\mu)K^+$ decay chain with the tag and probe method that are re-weighted in P-PT bins to match the muons spectra for muons from $K^0_S \to \mu^+\mu^-$. This last statement means that the $\epsilon = \epsilon(P, PT)$ is a function of P and PT of the muons. In each BDT bin (j), the expected muonID efficiency for a given cut is computed as

$$\frac{\sum_i^{N_j} \epsilon^j(P_i^+, PT_i^+)\epsilon^j(P_i^-, PT_i^-)}{N_j} \tag{6.1}$$

where the index i run over the N_j $K^0_S \to \mu^+\mu^-$ candidates and the $P_i^\pm$-$PT_i^\pm$ are the momentum and the transverse momentum of the positive and negative muons for the i-candidate. The signal efficiency for each bin as a function of the cut in the muonID is shown in Figure 6.1. The background mass distribution was fitted with a parametric model developed for the 2011 analysis and validated on MC data. It consists in a power low $(\frac{p[0]}{(x-p[1])^{p[2]}})$ for the $K^0_S \to \pi^+\pi^-$ tail and an exponential for the combinatorial background and it is applied directly on the data sidebands.

From these efficiency and from the B values obtained from the fit the significance figure of merit is constructed and showed in Figure 6.2. Optimal values are typically in the range

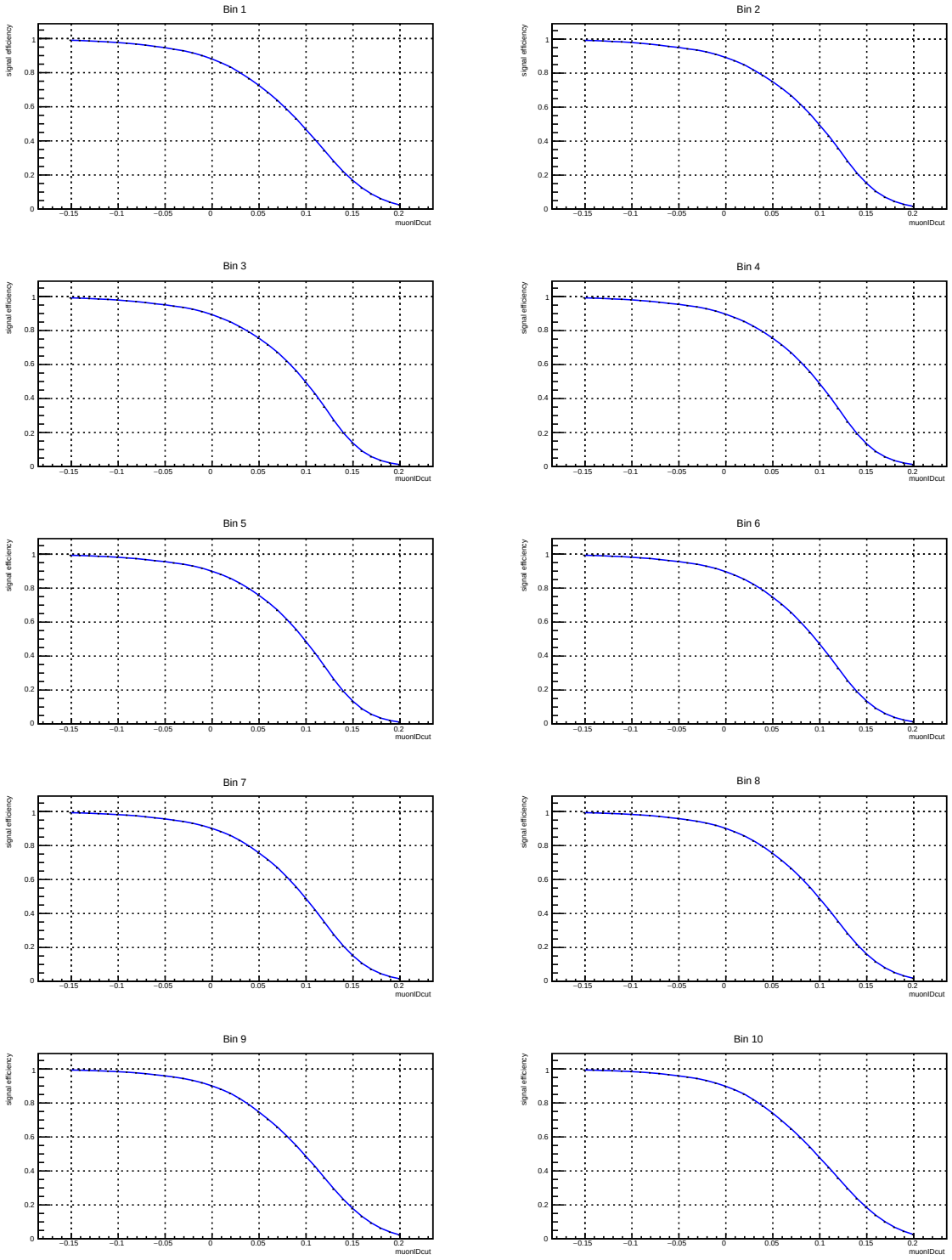


Figure 6.1: In the plot the trend of the signal efficiency with respect to the muonID cut applied is shown for each bin. the efficiency drop to zero for high values of the muonID cut. The way this quantity has been calculated is explained in the text.

0-0.05 of the muonID variable, with a sizable statistical uncertainty due to the very low background statistics, and without any clear dependence on the BDT value. We conclude that a uniform cut at the value 0.05 is close to an optimal requirement. The signal efficiency for this cut is about 75% and slightly depend on the bin, while the background B numbers calculated after this cut are shown in Figure 6.3.

The results of the muonID application and of the fit are shown in Figure 6.4 for each bin, where the number of events before end after the muonID request are illustrated. It is clear that the muonID request reduce the fraction of $K_S^0 \to \pi^+\pi^-$ tail that enter into the signal region.

6.2 Other Background sources

In the 2011 analysis some decay channels that could be potential background for the $K_S^0 \to \mu^+\mu^-$ search were investigated obtaining the result that only the $K_S^0 \to \pi^+\pi^-$ double-misID case was relevant for the significance reachable with the old statistics (which is half of the one analyzed in this thesis). The conclusion reached for the previous analysis are thought to be still valid for this analysis, because the rejected decay channel was found to give contribution only when the significance of the $K_S^0 \to \mu^+\mu^-$ upper limit will reach the SM one (or at least the 10^{-11} order of magnitude) which is clearly far from the present capabilities. Here a brief summary of these negligible sources is given, referencing to [3] for the complete study.

The two main other specific backgrounds are $K_L^0 \to \mu^+\mu^-$ and $K_S^0 \to \pi^+\mu^-\overline{\nu}$ with the pion decaying to muon and neutrino. The first, which would peak in the $K_S^0 \to \mu^+\mu^-$ signal region, is suppressed by the lifetime difference between K_S^0 and K_L^0 but the $\mathcal{B}(K_L^0 \to \mu^+\mu^-) = (6.84 \pm 0.11) \cdot 10^{-9}$ is much larger than the SM expectation of $\mathcal{B}(K_S^0 \to \mu^+\mu^-)$. $K_L^0 \to \mu^+\mu^-$ events are topologically equivalent to $K_S^0 \to \mu^+\mu^-$ and can be disentangled from $K_S^0 \to \mu^+\mu^-$ only by the lifetime (or variables which depend on it).

To estimate the order of magnitude of this background an estimation of the efficiencies ratio for for $K_S^0 \to \mu^+\mu^-$ and $K_L^0 \to \mu^+\mu^-$ was obtained assuming that it depend only on the lifetimes of K_S^0 and K_L^0, through a non-flat lifetime acceptance that was estimated using real data $K_S^0 \to \pi^+\pi^-$ Tis events and was found to behaves approximately roughly like an exponential. Using the lifetime range accepted by the selection and fiducial cuts (8.95 to 130 ps) it was found an an efficiency ratio of $\sim 2-3\cdot10^{-3}$ resulting in an effective branching fraction for $K_L^0 \to \mu^+\mu^-$, including the K_L^0-K_S^0 interference contribution, of about $1.6\cdot10^{-11}$ and so this background can therefore be neglected also for the current analysis for which it is not expected to go below 10^{-10}.

A variant of the case is the $K_L^0 \to \mu^+\mu^-\gamma$ channel since the $\mathrm{B}(K_L^0 \to \mu^+\mu^-\ \gamma) = 3.6 * 10^{-7}$ [23] is a factor ∼50 larger than that of $K_L^0 \to \mu^+\mu^-$, but again it was found to be negligible due to the lifetime acceptance as for the $K_L^0 \to \mu^+\mu^-$ case and also because this background does not peak in the signal region. Indeed, only the fraction of events where the γ energy is low enough to allow that the mother mass reconstructed using only the two muons enter in the signal region. It was concluded that the $K_L^0 \to \mu^+\mu^-\gamma$ should give a contribution below the $K_L^0 \to \mu^+\mu^-$ one.

Concerning the $K_S^0 \to \pi^+\mu^-\overline{\nu}$ (BR of $(4.69 \pm 0.05) * 10^{-4}$), it is expected to lie in the low-mass sideband due to the missing neutrino and so only the the region of the phase space in which the momentum of the neutrino is very small can affect the signal region. Moreover, the pion has to be misidentified as muon and so the mass left shift is increased

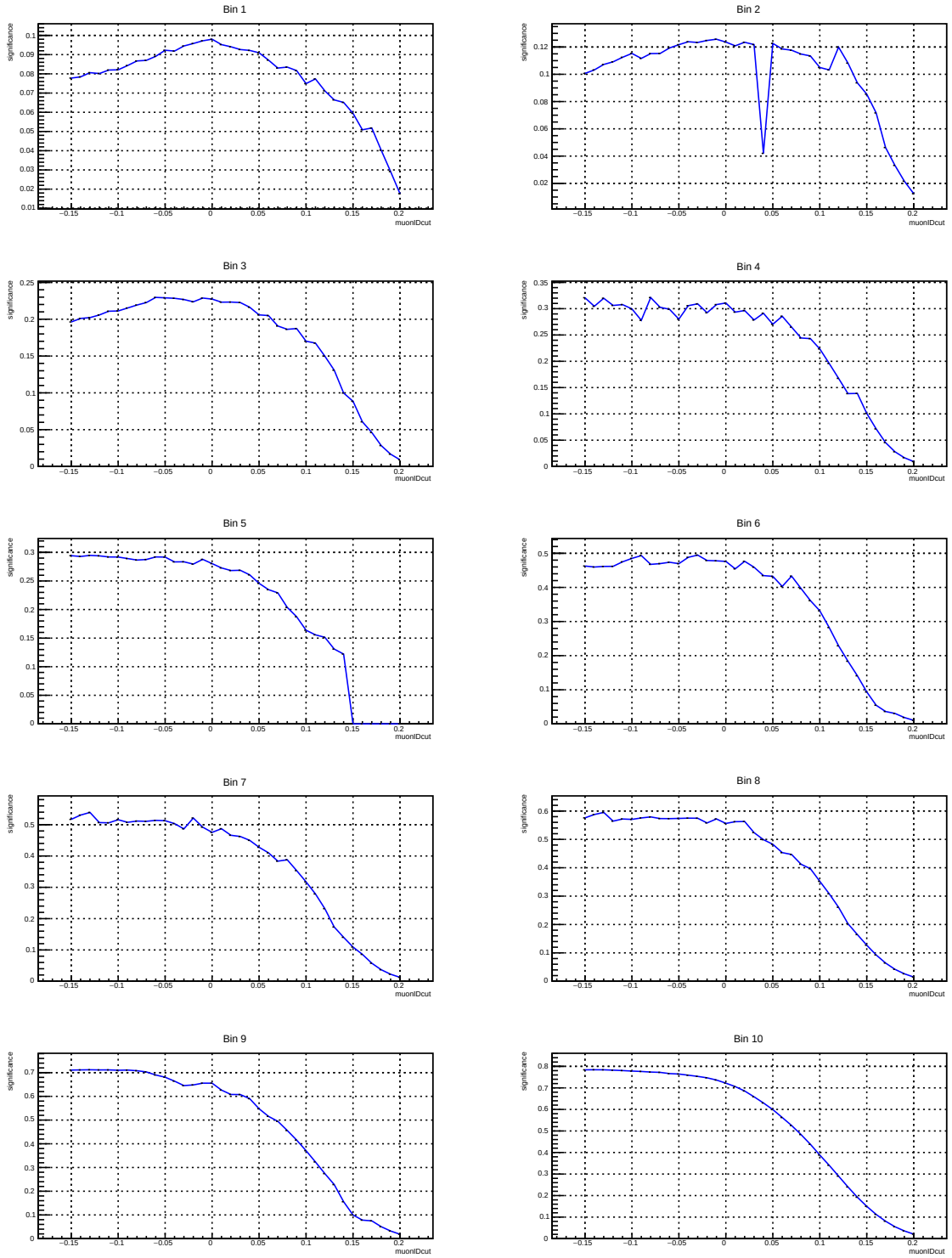


Figure 6.2: In the plot the trend of the significance defined by the formula $\epsilon/(\frac{3}{2}+B)$ with respect to the muonID cut applied is shown for each bin. The significance show that the region at high (> 0.1) muonID values can be excluded due to the drop of the signal efficiency, while, for each bin, there is a no ideal choice of the cut because the calculation is dominated by the fluctuation in the B quantity.

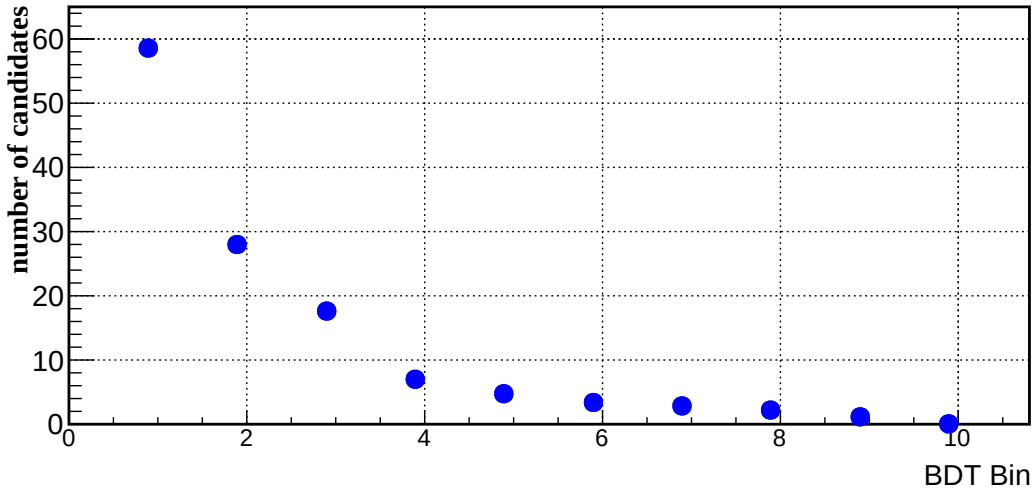


Figure 6.3: Background obtained from the integral in the signal region of the fit function after the cut in the muonID at the value of 0.05. **TosTos1Tos** trigger category.

by this wrong mass assignment.

Given the fact that the neutrino is never reconstructed, the initial K_S^0 is built only with the pion and the muon. In the previous analysis the contribution of this background was studied in an dedicated MC sample where the the expected yield of $K_S^0 \to \pi^+\mu^-\overline{\nu}$ in 1fb^{-1} integrated luminosity was estimated normalizing to the $K_S^0 \to \pi^+\pi^-$ Tis. After the application of whole selection, it was found that this background is negligible with respect to the the double misID $K_S^0 \to \pi^+\pi^-$.

Another similar channel, $K_L^0 \to \pi^+\mu^-\overline{\nu}$, was investigated because of the much larger BR($K_L^0 \to \pi^+\mu^-\overline{\nu}$) = 0.27 with respect that of $K_S^0 \to \pi^+\mu^-\overline{\nu}$. The decay is expected to occur mainly outside VELO (because of the large K_L^0 lifetime) and it was found that, because of the same effect that inibits the $K_L^0 \to \mu^+\mu^-$ respect to the $K_S^0 \to \mu^+\mu^-$ mode, the effective BR($K_L^0 \to \pi^+\mu^-\overline{\nu}$) has the same order of BR($K_S^0 \to \pi^+\mu^-\overline{\nu}$) and so negligible for the current dataset.

All these background would need to be revisited for a search for $K_S^0 \to \mu^+\mu^-$ when the upper limit on the BR could reach 10^{-11} which is cleary not expected with the 2 fb^{-1} analyzed in this thesis. Indeed, this sensitivity is supposed to be achievable in the future upgrade phase of LHC, currently planned for 2019 onward.

6.3 Normalization and Outlook

The most important next analysis step which is currently under study is the normalization, which allows to convert an observed number of $K_S^0 \to \mu^+\mu^-$ candidates in a branching fraction. As already explained, the normalization sample will consist in $K_S^0 \to \pi^+\pi^-$ decays candidates, obtained after applying all selection criteria except the muonID requirement. A possible strategy consists in counting normalization events in the randomly acquired events (NoBias sample). The normalization will be independent for each trigger paths considered

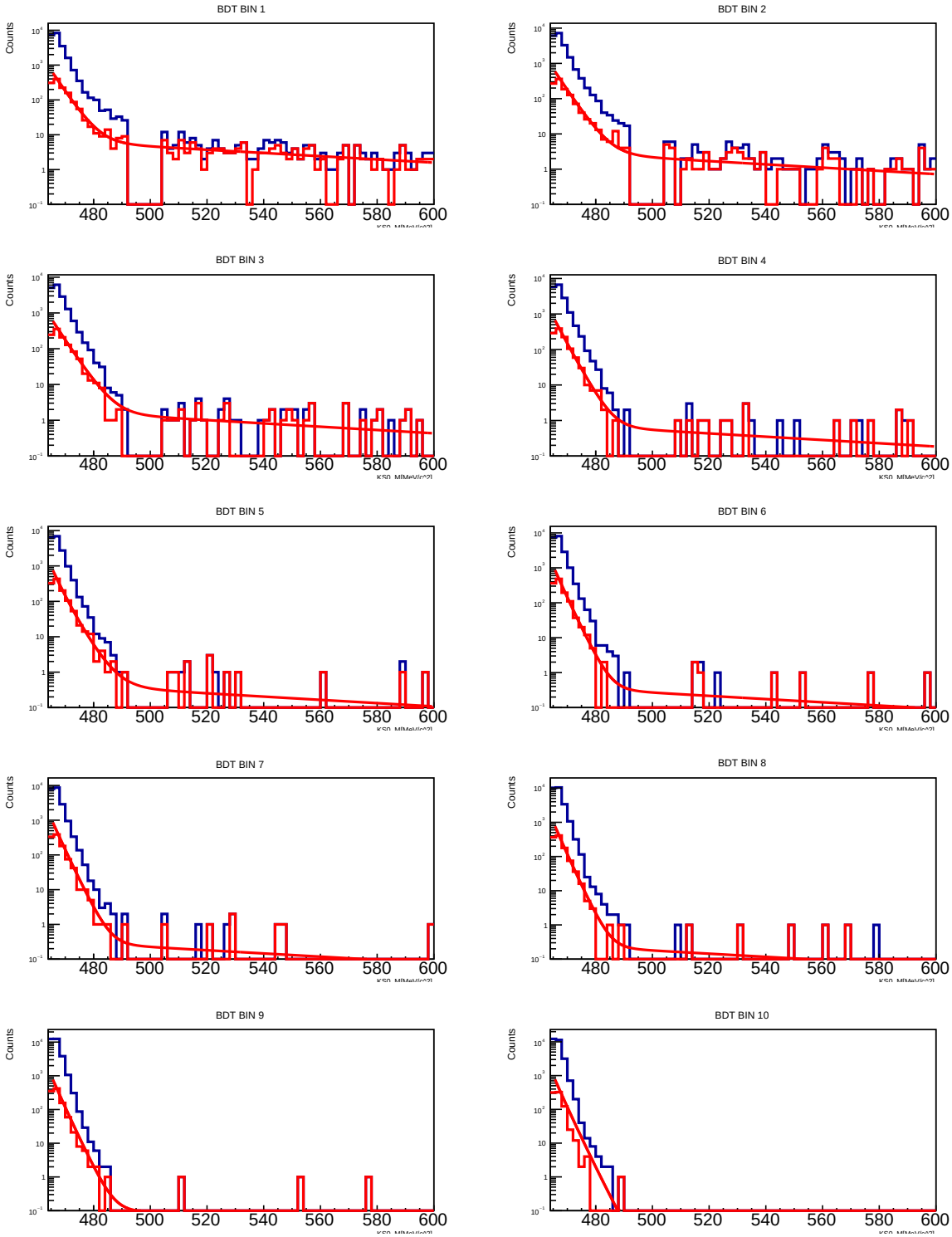


Figure 6.4: Mass distribution in each bin before (blue) and after (red) the muonID request. Both the combinatorial (exponential) and the $K^0_S \to \pi^+\pi^-$ tail (power law) are fitted to the sidebands is also reported with the signal region, delimited by the vertical dashed lines, still blind at this stage of the analysis. The cut in the muonID algorithm is able to reduce the $K^0_S \to \pi^+\pi^-$ tail component in the left side. **TosTos1Tos** trigger category.

and the formula to convert the yield of selected K^0_S decays into a branching fraction is

$$\frac{\mathcal{B}(K^0_S \to \mu^+\mu^-)}{\mathcal{B}(K^0_S \to \pi^+\pi^-)} = \frac{\epsilon^{SEL}_{\pi\pi}}{\epsilon^{SEL}_{\mu\mu}} \times \frac{1}{\epsilon^{PID}_{\mu\mu}} \times \frac{s^{\text{NoBias}}}{\epsilon^{TRG/SEL}_{\mu\mu}} \times \frac{N^{TRG}_{K^0_S \to \mu^+\mu^-}}{N^{NoBias,GECS}_{K^0_S \to \pi^+\pi^-}}. \tag{6.2}$$

where the efficiencies are factorized for convenience in three terms:

- ϵ_{SEL} is the offline efficiency including reconstruction and selection but excluding the particle identification i.e. the probability for a $K^0_S \to \pi^+\pi^-$ ($K^0_S \to \mu^+\mu^-$) decay generated in a $p-p$ collision to have been reconstructed and selected in the case that the trigger is able to select all the events ($\epsilon^{SEL} = \frac{N_{SEL}}{N_{GEN}}$);
- ϵ^{PID} is the efficiency of muon identification with respect to the selected candidates ($\epsilon^{PID} = \frac{N_{PID\&SEL}}{N_{SEL}}$) and it affects only the $K^0_S \to \mu^+\mu^-$ channel.
- $\epsilon^{TRG/SEL}$ is the trigger efficiency with respect to offline selected events ($\epsilon^{TRG/SEL} = \frac{N_{SEL\&TRG\&PID}}{N_{SEL}}$), while the s^{NoBias} is the trigger efficiency for the NoBias sample, which is simply the inverse of the scaling factor used in the acquisition of these events.

The calculation of these parameters is under study and here I will only give the main idea to calculate them.

The ratio of selection and reconstruction efficiencies is the only factor that needs to be obtained from simulation and, due to the same kinematic characteristics of the $K^0_S \to \mu^+\mu^-$ and $K^0_S \to \pi^+\pi^-$ decays, it is expected to be depart from unity only due to the different material interaction cross-section for pions and muons.

The efficiency of the muonID is measured from real data, with a tag and probe method applied on a clean sample of $J/\psi \to \mu^+\mu-$ with J/ψ from $B^+ \to J/\psi K^+$. This calibration sample is selected requiring the identification of only one of the two muons and with the trigger that only relies on this tag muon. The other muon, not biasing the trigger and the selection at any level, is used to probe the muonID procedure.

Thanks to the new generated MC statistics it will be possible to carry out the analysis also for the TisTisTos and TosTos2Tos trigger lines as also conclude the improvement in the muonID algorithm. In the description of this chapters the role I played in the analysis was underlined and the general scheme of the analysis and the remain works which are here summarized

- creation of all the proxies for each trigger category
- check of the best MVA variables and classifiers for each trigger category
- studies on the possible background contributions for $K^0_S \to \pi^+\pi^-$ (both π misID), $K^0_L \to \mu^+\mu^-$, combinatorial, $K^0_S \to \pi\mu\nu_\mu$ (π misID)
- conclusion of the muonID improvement, increasing background rejection which is the most important new factor with respect to the previous analysis to discriminate the $K^0_S \to \pi^+\pi^-$ background
- research of a muonID cuts optimization for a better BR sensibility in each bins for each trigger categories
- calculation of the normalization factors

where the blue points indicate the analysis phases I carried out personally.

We have shown in this thesis that the single event sensitivity, namely the branching ratio corresponding to the observation of one signal event, is expected to be increased with respect to the 2011 result by a factor 2 from the larger data sample acquired in 2012, and by an additional factor 3 from the improved trigger efficiency. The reconstruction and selection efficiency are expected to be comparable to the 2011 analysis. However, thanks to the improvements in the procedure to build an MVA classifier to reject the combinatorial background, the improved muon identification algorithm, and the optimization of the muonID requirement, the expected background in the most sensitive BDT bins is of order unity, as was the case in the 2011 result. This achievement only costed a limited reduction of the muonID requirement efficiency (from about 98% to about 75%, note that this requirement was not optimized in the previous analysis). The evaluation of systematic uncertainties for the current analysis has still to be finalized. With respect to the 2011 analysis, where the result was mostly limited by the available statistics, systematic uncertainties can only improve thanks to the larger available control samples and the more stable data acquisition conditions. We can conclude that, if no signal events above background expectations will be observed, we can expect that the limit on the branching ratio will be improved by more than a factor 4 by the inclusion of 2012 data.

Conclusions

The SM is the theory that describes our current understanding of fundamental particles and their interactions. It has proven to be excellent in describing and predicting the phenomena studied at HEP. However, there are same experimental evidences that cannot be explained by the SM, such as the cosmological evidence for Dark Matter and the baryonic asymmetry of the universe. To explain these results, NP models have been developed and they are tested in HEP experiments at accelerator such as the LHC. LHC collides protons almost head-on at a nominal center-of-mass energy of $\sqrt{s} = 14$ TeV ($\sqrt{s} = 7$ TeV during 2011 and $\sqrt{s} = 8$ TeV during 2012). One of the four main detector at LHC is LHCb, with a physics programme originally focused on b and c quark physics, CP violating processes and rare decays, intended for indirect searches for NP as unexpected contribution arising in loop diagrams. The excellent performances of the LHC and of the LHCb experiment allowed to significantly extend its physics program to include, e.g., studies on top, heavy ion, and kaon physics.

The flavour-changing neutral-current (FCNC) $K_S^0 \to \mu^+\mu^-$ decay is a rare decay whose search has been proposed as a novel way to profit from the excellent performance of the LHCb detector. The non observation of the $K_S^0 \to \mu^+\mu^-$ decay played an historical role in particle physics, inspiring the GIM mechanism which predicts the suppression of FCNC decay with the introduction of a new quark, later identified as the c quark and observed with the discovery of the J/ψ meson. Indeed, the $K_S^0 \to \mu^+\mu^-$ decay is a FCNC which does not arise at tree level in SM because the two final state muons can only be directly produced from a photon, Higgs or Z, but none of these bosons can be originated by a s-d quark interaction directly, since they can be produced only by particles of the same flavour.

The $K_S^0 \to \mu^+\mu^-$ amplitude is particularly interesting since its short distance (SD) component is dominated by the CP-violating part of the $s \to dl^+l^-$ amplitude, which is very sensitive to NP and which is poorly constrained so far, while the long distance (LD) contributions are well predicted and driven mainly by virtual intermediate two-photon states. The proportion between the LD and the SD contribution is model dependent and in the SM the LD contributions are dominant, so evidence of $K_S^0 \to \mu^+\mu^-$ well above the SM expected branching ratio ($(5.1 \pm 1.5) \cdot 10^{-12}$) would be a clear signal of NP, while the non observation of such enhancement will provide significant constraints on several consistent NP scenarios. Indeed, $K_S^0 \to \mu^+\mu^-$ has also been suggested as a possible way to look for new light scalars particles. Thus, we can conclude that the rare decay $K_S^0 \to \mu^+\mu^-$ is a very useful source of information for the SD structure of new $\Delta S = 1$ FCNC transitions as possible new high mass particles can contribute in the loop as virtual process.

The search for $K_S^0 \to \mu^+\mu^-$ is still challenging at LHCb because the LHCb detector is optimised for B mesons which show different features to K_S^0, as, for example, the K_S^0 mass is around 10 times smaller than that of B^0, while it flies a factor of more than 50 more, impling lower reconstruction and trigger performances. However, the LHC is a great kaon

factory, with about 10^{13} neutral kaons being produced within the LHCb acceptance per fb^{-1} of integrated luminosity and about 40% of these kaons decay inside its vertex detector.

An analysis of the K_S^0 decay was performed with the 2011 LHCb data set, yielding an upper limit $\mathcal{B}(K_S^0 \to \mu^+\mu^-) < 11.2(9.0) \cdot 10^{-9}$ at 95 (90) % confidence level. Before LHCb the most sensitive search had been performed at the CERN PS in the 1973 yielding an upper limit $\mathcal{B}(K_S^0 \to \mu^+\mu^-) < 3.1 \cdot 10^{-7}$ at 90 % confidence level.

In this thesis, a contribution to the upgrade of the $K_S^0 \to \mu^+\mu^-$ analysis at LHCb, using the not yet analyzed 2 fb^{-1} data acquired in 2012, is performed. The main different experimental condition between the 2011 and 2012 LHCb runs is the increased trigger efficiency achieved thanks to some changes in the trigger line that search for di-muon candidates.

The analysis strategy for the $K_S^0 \to \mu^+\mu^-$ at LHCb starts with an efficient selection to get rid of most of the background, requiring two muons track with a common and detached vertex, with a reconstructed total momentum pointing to one of primary vertex in the event. However, the strong biases caused by the low trigger efficiency force to split the analysis depending on which trigger lines have selected the $K_S^0 \to \mu^+\mu^-$ candidate. The two main lines are the one where an high transverse momentum single muon is observed and the one that contain two muons with a common vertex.

The different trigger selections were studied with the help of MonteCarlo simulations. In order to maximize the sensitivity of the analysis, three different selections have been identified. The kinematics of the three paths is different, forcing us to implement a separate analysis for each one of them.

The background is mainly due to combinatorial candidates and $K_S^0 \to \pi^+\pi^-$ decays with both pions misidentified as muons. For the former an optimized multivariate analysis (MVA) is developed, checking which are the best variables combinations that can give the best discrimination performance separately for each of the three trigger categories. A tuned non linear classifier developed with the boosted decision tree (BDT) technique is employed for this purpose. The BDT is trained, with a data driven method, developed within this work, where training input samples are used to simulate the signal and the combinatorial background. The signal candidates passing the MVA selection are classified in bins of the BDT. Concerning the other background, thanks to the excellent K_S^0 mass resolution achieved ($\sim$ 4 MeV/c^2), it does not peak into the signal mass region ([492, 504] MeV/c^2) because of the mass shift created by the wrong mass association of the daughter pions and so it contributes only with a tail. To further reduce this tail, a tuned muonID algorithm for this analysis is developed and the cut request on the muonID parameter for the two daughters is optimized. Finally, the residual background expectation is obtained from a fit into the data mass sidebands ([440, 492] MeV/c^2 and [504, 600] MeV/c^2).

To convert the number of $K_S^0 \to \mu^+\mu^-$ decay observed into a branching fraction, $K_S^0 \to \pi^+\pi^-$ events from a trigger-unbiased sample (randomly acquired) will be used as normalization channel. This channel has the advantage to share the same topology and most of the kinematics aspects with the signal one, making it the natural candidate. The normalization is done in in bins of BDT and if no significant signal excess will be found in the $K_S^0 \to \mu^+\mu^-$ signal region, an upper limit, using the CLs method, will be set.

In this thesis work I took confidence with the LHCb software, studied the best triggers categories, developed the MVA classifier to suppress combinatorial background, and optimized the muon identification requirement, making use of programming languages as C++ and python, as also the ROOT software and the TMVA toolkit.

In this elaborate, my original contributions to the analysis have been underlined: the

procedure to build a data sample as close as possible in principle to the expected signal using $K^0_S \to \pi^+\pi^-$ decays; the improved combination of input variables feeding the MVA classifier, and the optimization of the muon identification requirement.

The result from this new analysis on the integrated luminosity of 2 fb^{-1} acquired by LHCb in the 2012, will be published by the LHCb Collaboration when the analysis will be completed and approved after a LHCb internal review. Only at that stage, events in the signal region will be counted and a result for the branching ratio will be given. From the results of this work we can predict that, in case of no signal observation, a limit will be set that will likely overcome the current result by more than a factor 4. The large datasets that will be collected in the ongoing run of LHCb will further improve the precision of the measurement, but the much larger statistics needed to push the sensitivity closer to the SM level will be only reachable during the future upgrade phase (currently planned for 2019 onward) aiming at collecting up to 50 fb^{-1} of integrated luminosity.

Bibliography

[1] Jr. Alves, A. Augusto et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[2] Vardan Khachatryan et al. Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data. *Nature*, 522:68–72, 2015.

[3] Xabier Cid Vidal. *Search for the rare decays* $B^0_{(s)} \to \mu^+\mu^-$ *and* $K^0_S \to \mu^+\mu^-$ *with 1* fb^{-1} *at LHCb*. PhD thesis, Santiago de Compostela U., 2012-07-20.

[4] S. Gjesdal, G. Presser, P. Steffen, J. Steinberger, F. Vannucci, et al. Search for the decay $K(s) \to 2muon$. *Phys.Lett.*, B44:217–220, 1973.

[5] Gino Isidori and Rene Unterdorfer. On the short distance constraints from K(L,S) $\to$ mu+ mu-. *JHEP*, 0401:009, 2004.

[6] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

[7] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[8] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc.*, C680519:367–377, 1968.

[9] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.

[10] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[11] E. W. Otten and C. Weinheimer. Neutrino mass limit from tritium beta decay. *Rept. Prog. Phys.*, 71:086201, 2008.

[12] F. J. Hasert et al. Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment. *Phys. Lett.*, B46:138–140, 1973.

[13] G. Arnison et al. Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV. *Phys. Lett.*, B122:103–116, 1983.

[14] G. Arnison et al. Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider. *Phys. Lett.*, B126:398–410, 1983.

[15] Nicola Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys.Rev.Lett.*, 10:531–533, 1963.

[16] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2 pi Decay of the k(2)0 Meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[17] VE Barnes, PL Connolly, DJ Crennell, BB Culwick, WC Delaney, WB Fowler, PE Hagerty, EL Hart, N Horwitz, PVC Hough, et al. Observation of a hyperon with strangeness minus three. *Physical Review Letters*, 12(8):204, 1964.

[18] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014.

[19] Robert N Cahn and Gerson Goldhaber. *The experimental foundations of particle physics*. Cambridge University Press, 2009.

[20] J. J. Aubert et al. Experimental Observation of a Heavy Particle J. *Phys. Rev. Lett.*, 33:1404–1406, 1974.

[21] J.E. Augustin et al. Discovery of a Narrow Resonance in e+ e- Annihilation. *Phys.Rev.Lett.*, 33:1406–1408, 1974.

[22] C. Bacci et al. Preliminary Result of Frascati (ADONE) on the Nature of a New 3.1-GeV Particle Produced in e+ e- Annihilation. *Phys. Rev. Lett.*, 33:1408, 1974. [Erratum: Phys. Rev. Lett.33,1649(1974)].

[23] G. S. Abrams et al. The Discovery of a Second Narrow Resonance in e+ e- Annihilation. *Phys. Rev. Lett.*, 33:1453–1455, 1974. [Adv. Exp. Phys.5,150(1976)].

[24] G. Goldhaber et al. Observation in e+ e- Annihilation of a Narrow State at 1865-MeV/c**2 Decaying to K pi and K pi pi pi. *Phys. Rev. Lett.*, 37:255–259, 1976.

[25] A. Chen et al. Evidence for the F Meson at 1970-MeV. *Phys. Rev. Lett.*, 51:634, 1983.

[26] S. W. Herb et al. Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions. *Phys. Rev. Lett.*, 39:252–255, 1977.

[27] Christoph Berger et al. Observation of a Narrow Resonance Formed in e+ e- Annihilation at 9.46-GeV. *Phys. Lett.*, B76:243–245, 1978.

[28] G Finocchiaro, G Giannini, J Lee-Franzini, RD Schamberger Jr, M Sivertz, LJ Spencer, PM Tuts, T Böhringer, F Costantini, J Dobbins, et al. Observation of the υ''' at the cornell electron storage ring. *Physical Review Letters*, 45(4):222, 1980.

[29] Brian R Martin and Graham P Shaw. *Particle physics*. John Wiley & Sons, 2013.

[30] H. Fritzsch, Murray Gell-Mann, and H. Leutwyler. Advantages of the Color Octet Gluon Picture. *Phys. Lett.*, B47:365–368, 1973.

[31] David J Gross and Frank Wilczek. Asymptotically free gauge theories. i. *Physical Review D*, 8(10):3633, 1973.

[32] Steven Weinberg. Non-abelian gauge theories of the strong interactions. *Physical Review Letters*, 31(7):494, 1973.

[33] G. Hanson et al. Evidence for Jet Structure in Hadron Production by e+ e- Annihilation. *Phys. Rev. Lett.*, 35:1609–1612, 1975.

[34] R. Brandelik et al. Evidence for Planar Events in e+ e- Annihilation at High-Energies. *Phys. Lett.*, B86:243, 1979.

[35] IJR Aitchison and AJG Hey. Gauge theories in particle physics. *Adam Hilger*, 1984.

[36] Juerg Beringer, JF Arguin, RM Barnett, K Copic, O Dahl, DE Groom, CJ Lin, J Lys, H Murayama, CG Wohl, et al. Review of particle physics. *Physical Review D*, 86(1), 2012.

[37] T. D. Lee and Chen-Ning Yang. Question of Parity Conservation in Weak Interactions. *Phys. Rev.*, 104:254–258, 1956.

[38] C.S. Wu, E. Ambler, R.W. Hayward, D.D. Hoppes, and R.P. Hudson. Experimental Test of Parity Conservation in Beta Decay. *Phys.Rev.*, 105:1413–1414, 1957.

[39] R. L. Garwin, L. M. Lederman, and Marcel Weinrich. Observations of the Failure of Conservation of Parity and Charge Conjugation in Meson Decays: The Magnetic Moment of the Free Muon. *Phys. Rev.*, 105:1415–1417, 1957.

[40] Marco Sozzi. *Discrete Symmetries and CP Violation: From Experiment to Theory: From Experiment to Theory.* Oxford University Press, 2008.

[41] A.D. Sakharov. Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe. *Pisma Zh.Eksp.Teor.Fiz.*, 5:32–35, 1967.

[42] Howard E Haber and Gordon L Kane. The search for supersymmetry: probing physics beyond the standard model. *Physics Reports*, 117(2):75–263, 1985.

[43] Evgeny K. Akhmedov, V. A. Rubakov, and A. Yu. Smirnov. Baryogenesis via neutrino oscillations. *Phys. Rev. Lett.*, 81:1359–1362, 1998.

[44] Gustavo C Branco, Luis Lavoura, and Joao Paulo Silva. *CP violation.* Number 103. Oxford University Press, 1999.

[45] K. Lande, E. T. Booth, J. Impeduglia, L. M. Lederman, and W. Chinowsky. Observation of Long-Lived Neutral V Particles. *Phys. Rev.*, 103:1901–1904, 1956.

[46] F. Niebergall et al. Experimental study of the delta-s/delta-q rule in the time dependent rate of k0 —¿ pi e nu. *Phys. Lett.*, B49:103–108, 1974.

[47] F. Muller, R. W. Birge, W. B. Fowler, R. H. Good, W. Hirsch, R. P. Matsen, L. Oswald, W. M. Powell, H. S. White, and O. Piccioni. Regeneration and Mass Difference of Neutral K Mesons. *Phys. Rev. Lett.*, 4:418–421, 1960.

[48] Sheldon Bennett, David Nygren, Harry Saal, Jack Steinberger, and John Sunderland. Nonorthogonality of the long- and short-lived neutral kaon states and phenomenological analysis of experiments on CP nonconservation in K^0 decay. *Phys. Rev. Lett.*, 19:997–999, Oct 1967.

[49] David E. Dorfan, J. Enstrom, D. Raymond, Melvin Schwartz, Stanley G. Wojcicki, Donald H. Miller, and M. Paciotti. CHARGE ASYMMETRY IN THE MUONIC DECAY OF THE K0(2). *Phys. Rev. Lett.*, 19:987–993, 1967.

[50] A. Aloisio et al. Measurement of the branching fraction for the decay K(S) —> π e ν. *Phys. Lett.*, B535:37–42, 2002.

[51] A. Alavi-Harati et al. Measurements of direct CP violation, CPT symmetry, and other parameters in the neutral kaon system. *Phys. Rev.*, D67:012005, 2003. [Erratum: Phys. Rev.D70,079904(2004)].

[52] J Richard Batley, RS Dosanjh, TJ Gershon, GE Kalmus, C Lazzeroni, DJ Munday, E Olaiya, M Patel, MA Parker, TO White, et al. A precision measurement of direct cp violation in the decay of neutral kaons into two pions. *Physics Letters B*, 544(1):97–112, 2002.

[53] C Albajar, MG Albrow, OC Allkofer, G Arnison, A Astbury, B Aubert, T Axon, C Bacci, T Bacon, N Bains, et al. Search for B^0-$\bar{B}^0$ oscillations at the cern proton-antiproton collider. *Physics Letters B*, 186(2):247–254, 1987.

[54] A. Abashian et al. Measurement of the CP violation parameter $\sin 2\phi_1$ in B^0_d meson decays. *Phys. Rev. Lett.*, 86:2509–2514, 2001.

[55] Bernard Aubert et al. Measurement of CP violating asymmetries in B^0 decays to CP eigenstates. *Phys. Rev. Lett.*, 86:2515–2522, 2001.

[56] Bernard Aubert et al. Observation of CP violation in the B^0 meson system. *Phys. Rev. Lett.*, 87:091801, 2001.

[57] Bernard Aubert et al. Measurement of the CP-violating asymmetry amplitude $\sin 2\beta$. *Phys. Rev. Lett.*, 89:201802, 2002.

[58] Julian Wishahi. Measurements of Δm_d, Δm_s, and $\sin 2\beta$ with LHCb. In *7th Workshop on the CKM Unitarity Triangle (CKM 2012) Cincinnati, Ohio, USA, September 28-October 2, 2012*, 2012.

[59] Bernard Aubert et al. Observation of direct CP violation in $B^0 \to K^+\pi^-$ decays. *Phys. Rev. Lett.*, 93:131801, 2004.

[60] Y. Chao et al. Evidence for direct CP violation in B0 —> K+ pi- decays. *Phys. Rev. Lett.*, 93:191802, 2004.

[61] A. Abulencia et al. Measurement of the $B^0_s - \bar{B}^0_s$ Oscillation Frequency. *Phys. Rev. Lett.*, 97:062003, 2006.

[62] T. Aushev et al. Physics at Super B Factory. 2010.

[63] R. Aaij et al. Precision measurement of the B^0_s–$\overline{B}^0_s$ oscillation frequency in the decay $B^0_s \to D^-_s\pi^+$. *New J. Phys.*, 15:053021, 2013.

[64] G. Ecker and A. Pich. The Longitudinal muon polarization in $K_L \to \mu^+\mu^-$. *Nucl.Phys.*, B366:189–208, 1991.

[65] G. D'Ambrosio, G. Ecker, G. Isidori, and H. Neufeld. Radiative nonleptonic kaon decays. pages 265–313, 1994.

[66] Alan J. Schwartz. Measurement of BR (K(L) $\to \mu^+\mu^-$) : Overall results from BNL E791. *Conf. Proc.*, C921110:816–820, 1992.

[67] R. Aaij et al. Search for the rare decay $K_S \to \mu^+\mu^-$. *JHEP*, 1301:090, 2013.

[68] H. Burkhardt et al. Observation of the Decay $K_S \to 2\gamma$ and Measurement of the Decay Rates $K_L \to 2\gamma$ and $K_S \to 2\gamma$. *Phys. Lett.*, B199:139, 1987.

[69] Na62 web site. https://na62.web.cern.ch/na62/Home/Home.html.

[70] Lyndon Evans and Philip Bryant. LHC Machine. *JINST*, 3:S08001, 2008.

[71] www.cern.ch. http://home.web.cern.ch/.

[72] R. Aaij et al. LHCb Detector Performance. *Int. J. Mod. Phys. A*, 30(arXiv:1412.6352. CERN-PH-EP-2014-290. LHCB-DP-2014-002. CERN-LHCB-DP-2014-002):1530022. 82 p, Dec 2014.

[73] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[74] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[75] Lhcb web site. http://lhcb.web.cern.ch/lhcb/.

[76] S. Amato et al. LHCb technical proposal. 1998.

[77] LHCb technical design report: Reoptimized detector design and performance. 2003.

[78] B. Adeva et al. Roadmap for selected key measurements of LHCb. 2009.

[79] R Aaij et al. Measurement of the CKM angle γ from a combination of $B^\pm \to Dh^\pm$ analyses. *Phys. Lett.*, B726:151–163, 2013.

[80] Roel Aaij et al. Determination of γ and $-2\beta_s$ from charmless two-body decays of beauty mesons. *Phys. Lett.*, B741:1–11, 2015.

[81] Aurelien Martens. Time integrated and time dependent asymmetries in B $\to$ hh' (h = K , π) decays at LHCb. *PoS*, Beauty2013:013, 2013.

[82] R Aaij et al. Measurement of CP violation and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi\pi^+\pi^-$ decays. *Phys. Rev.*, D87(11):112010, 2013.

[83] R Aaij et al. First Evidence for the Decay $B_s^0 \to \mu^+\mu^-$. *Phys.Rev.Lett.*, 110(2):021801, 2013.

[84] Angular analysis of the $B_d^0 \to K^{*0}\mu^+\mu^-$ decay. (LHCb-CONF-2015-002), Mar 2015.

[85] Bhubanjyoti Bhattacharya, Alakabha Datta, David London, and Shanmuka Shivashankara. Simultaneous Explanation of the R_K and $R(D^{(*)})$ Puzzles. *Phys. Lett.*, B742:370–374, 2015.

[86] Roel Aaij et al. Observation of Photon Polarization in the b$\to$sγ Transition. *Phys. Rev. Lett.*, 112(16):161801, 2014.

[87] Roel Aaij et al. Observation of J/ψ p Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi$ K$^-$p Decays. *Phys. Rev. Lett.*, 115:072001, 2015.

[88] Murray Gell-Mann. A Schematic Model of Baryons and Mesons. *Phys. Lett.*, 8:214–215, 1964.

[89] Roel Aaij et al. Measurement of the ratio of branching fractions $\mathcal{B}(\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau)/\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)$. *Phys. Rev. Lett.*, 115(11):111803, 2015.

[90] R. Aaij et al. Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays. *Phys. Rev. Lett.*, 113:151601, 2014.

[91] R Aaij et al. Observation of $D^0 - \overline{D}^0$ oscillations. *Phys.Rev.Lett.*, 110(10):101802, 2013.

[92] R Aaij et al. First observation of CP violation in the decays of B_s^0 mesons. *Phys. Rev. Lett.*, 110(22):221601, 2013.

[93] Roel Aaij et al. Measurement of indirect CP asymmetries in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays using semileptonic B decays. *JHEP*, 04:043, 2015.

[94] R. Aaij et al. Measurement of the B^0–$\overline{B}^0$ oscillation frequency Δm_d with the decays $B^0 \to D^-\pi^+$ and $B^0 \to J/\psi K^{*0}$. *Phys. Lett.*, B719:318, 2013.

[95] R Aaij et al. Measurement of CP violation in the phase space of $B^\pm \to K^\pm\pi^+\pi^-$ and $B^\pm \to K^\pm K^+K^-$ decays. *Phys. Rev. Lett.*, 111:101801, 2013.

[96] Paula Collins. The LHCb Upgrade. In *Flavor physics and CP violation. Proceedings, 9th International Conference, FPCP 2011, Maale HaChamisha, Israel, May 23-27, 2011*, 2011.

[97] LHCb reoptimized detector design and performance: Technical Design Report. (CERN-LHCC-2003-030), 2003. LHCb-TDR-009.

[98] LHCb VELO TDR: Vertex locator. Technical design report. 2001.

[99] LHCb magnet: Technical design report. 2000.

[100] Jr. Alves, A. Augusto et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[101] LHCb inner tracker: Technical Design Report. (CERN-LHCC-2002-029), 2002. LHCb-TDR-008.

[102] LHCb outer tracker: Technical Design Report. (CERN-LHCC-2001-024), 2001. LHCb-TDR-006.

[103] LHCb calorimeters: Technical Design Report. (CERN-LHCC-2000-036), 2000. LHCb-TDR-002.

[104] LHCb muon system: Technical Design Report. (CERN-LHCC-2001-010), 2001. LHCb-TDR-004.

[105] Roel Aaij et al. LHCb Detector Performance. *Int. J. Mod. Phys.*, A30(07):1530022, 2015.

[106] R. Forty. RICH pattern recognition for LHCb. *Nucl. Instrum. Meth.*, A433:257–261, 1999.

[107] M. Benayoun and C. Jones. RICH reconstruction and particle identification using ring fit methods: Application to the RICH2 detector. 2005.

[108] F. Archilli et al. Muon Identification performance at LHCb with the 2010 data. *LHCb-INT-2011-048*, 2011.

[109] R. Aaij et al. LHCb detector performance.

[110] A. A. Alves, Jr. et al. Performance of the LHCb muon system. *JINST*, 8:P02022, 2013.

[111] Markus Frank, C. Gaspar, E. v. Herwijnen, B. Jost, and N. Neufeld. Deferred High Level Trigger in LHCb: A Boost to CPU Resource Utilization. *J.Phys.Conf.Ser.*, 513:012006, 2014.

[112] G. Barrand, I. Belyaev, P. Binko, M. Cattaneo, R. Chytracek, et al. GAUDI - A software architecture and framework for building HEP data processing applications. *Comput.Phys.Commun.*, 140:45–55, 2001.

[113] I. Belyaev, P. Charpentier, S. Easo, P. Mato, J. Palacios, W. Pokorski, F. Ranjard, and J. van Tilburg. Simulation application for the LHCb experiment. *eConf*, C0303241:TUMT003, 2003.

[114] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. Pythia 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026, 2006.

[115] David J Lange. The evtgen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 462(1):152–155, 2001.

[116] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl.Instrum.Meth.*, A506:250–303, 2003.

[117] The lhcb collaboration, the boole project. http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/.

[118] The lhcb collaboration, the moore project. http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/.

[119] The lhcb collaboration, the brunel project. http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/.

[120] The lhcb collaboration, the davinci project. http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/.

[121] The lhcb collaboration, the panoramix project. http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/panoramix/.

[122] Lhc computing grid web site. http://wlcg.web.cern.ch.

[123] Jürgen Knobloch, L Robertson, et al. Lhc computing grid technical design report. *CERN/LHCC*, 24:2005, 2005.

[124] S. Gjesdal, G. Presser, P. Steffen, J. Steinberger, F. Vannucci, H. Wahl, H. Filthuth, K. Kleinknecht, V. Lüth, and G. Zech. Search for the decay ks $\rightarrow$ 2µ. *Physics Letters B*, 44(2):217 – 220, 1973.

[125] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nucl.Instrum.Meth.*, A434:435–443, 1999.

[126] Brun and Fons Rademakers. Root - an object oriented data analysis framework. *Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86. See also http://root.cern.ch/.*

[127] J. Podolanski and R. Armenteros. Phil. Mag. 45 (1954) 13.

[128] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[129] F. Archilli et al. Performance of the Muon Identification at LHCb. *JINST*, 8:P10020, 2013.

[130] Giovanni Punzi. Sensitivity of searches for new signals and its optimization. *eConf*, C030908:MODT002, 2003. [,79(2003)].