Department of Physics and Astronomy

University of Heidelberg

Master thesis

in Physics

submitted by

Dominik Mitzel

born in Baden-Baden

December 2014

Study of semileptonic D^0 decays for a measurement of charm mixing at LHCb

This master thesis has been carried out by Dominik Mitzel
at the
Physikalisches Institut
under the supervision of
Prof. Dr. Ulrich Uwer
and
Dr. Angelo Di Canto
CERN

ABSTRACT

Precise measurements of neutral D meson mixing is a powerful probe for physics beyond the Standard Model, as quantum loops governing charm mixing transitions are sensitive to potential new heavy degrees of freedom and additional sources of CP violation. In this thesis, semileptonic decays of the form $D^{*+} \to D^0(\to K^+\mu^-\nu)\pi^+$ are for the first time explored at LHCb to measure charm mixing, using a sample of proton-proton collisions recorded in 2012 corresponding to an integrated luminisity of $2\,fb^{-1}$. The flavour of the D^0 meson at the moment of prodution and decay is determined by the charges of the pion and the muon, respectively. In the scope of this work a full event selection and fitting procedure is developed and validated, as well as the investigation of different correction algorithms for the undetected neutrino in the final states. A sensitivity study is performed with pseudo-experiments mimicking the observed data distributions, showing that with the available sample a statistical precision of 0.01% on the determination of the mixing rate $R_M = (x^2 + y^2)/2$ can be achieved.

ZUSAMMENFASSUNG

Präzise Messungen der Mischung von neutralen D Mesonen ist ein wichtiger Test auf Physik jenseits des Standard Modells, da die involvierten Quantenschleifen sensitiv für neue schwere Freiheitsgrade und zusätzlichen Quellen von CP Verletzung sind. In dieser Thesis werden zum ersten mal semileptonische Zerfälle der Form $D^{*+} \to D^0(\to K^+\mu^-\nu)\pi^+$ im Rahmen einer Charm Mischungsanalyse am LHCb Experiment untersucht. Dabei wird ein Datensatz von Proton-Proton Kollisionen aus dem Jahr 2012 verwendet, welcher einer integrierten Luminosität von $2\,fb^{-1}$ entspricht. Die Quarkzusammensetzung der D^0 Mesonen zur Zeit der Produktion und des Zerfalls wird durch die Ladungen des Pions und des Myons bestimmt. Im Rahmen dieser Arbeit wird eine komplette Ereignisselektion und eine Methode zur statistischen Trennung von Singal- und Untergrundereignissen entwickelt und validiert. Außerdem werden verschiedene Algorithmen zur Rekonstruktion des nicht detektierbaren Neutrinos im Endzustands untersucht. Mit Hilfe von Pseudoexperimenten, welche die beobachteten Datenverteilungen imitieren, wird in einer Sensitivitätsstudie gezeigt, dass mit dem verfügbaren Datansatz eine statisitische Präzision von 0.01% auf die Bestimmung der Mischungsrate $R_M = (x^2 + y^2)/2$ erreicht werden kann.

CONTENTS

INTRODUCTION

It has been 40 years already since the current theory describing the properties of the known elementary particles and their mutual interactions has been developed. Called the Standard Model (SM) of particle physics, this theory has withstood all experimental probes with increasing accuracy over the past years. The latest experimental confirmation of the theory happened recently with the discovery of the Higgs boson, being the scalar field responsible of the electroweak-symmetry breaking. Nevertheless, it is clear that the theory is incomplete as there are phenomena, such as dark matter, dark energy or the observed matter-antimatter asymmetry in the Universe, which can not be explained and therefore drive the development of new experiments to search for physics beyond the SM.

At the *European Organization for Nuclear Research* (CERN), the *Large Hadron Collider* (LHC), being the highest energy proton-proton collider ever built, provides the opportunity to search for physics beyond the SM by direct production of new heavy particles. In addition, a much higher energy scale can be probed by the investigation of processes involving quantum loops where the existence of new particles may contribute to the observable amplitudes even if the available energy is not sufficient for direct production. One of the four main experiments at the LHC, the LHCb experiment, is dedicated to these indirect searches through the study of c- and b-hadrons decays.

A powerful indirect probe for non-SM particles contributing to loops involving mesons with charm or beauty is the phenomenon of periodical matter-antimatter transitions, called *oscillation* or *mixing*. The process tests the structure of neutral flavour changing interactions, which happen only at higher order in the SM, but also allows to search for new sources of CP violation. CP violation describes the different behavior of matter and antimatter, which is believed to be a key ingredient to explain the observed matter-antimatter asymmetry in the Universe. The amount of CP violation predicted by the SM is not however sufficient to explain the observations, therefore alternative physic models providing additional sources of CP violation need to be probed.

Over the past years most of the experimental activity focused on indirect searches with kaon and beauty decays. Now LHCb has the potential to perform also indept investigations of the charm sector. Only two years ago, LHCb reported the first observation of charm mixing from a single measurement, using the hadronic two-body decay $D^0 \to K^+\pi^-$. This thesis represents the first attempt to study charm mixing using semileptonic charm decays. Contrary to hadronic two-body final states, where the same final state is accessible through mixing and highly suppressed direct decays, semileptonic final states, such as $D^0 \to K^+\mu^-\nu$, can unambiguously indicate charm-anticharm transitions as the direct decay is in this

case not possible. This allows to measure directly the mixing rate using a simpler time-integrated analysis. However, semileptonic final states at LHCb are only partially reconstructible due to the undetected neutrino and therefore suffer from large background contamination. This introduces non-trivial experimental challenges, which are dealt with for the first time in this thesis.

The thesis is organized as follows. Chapter 1 presents the general formalism and theoretical motivation for the study charm mixing, with special focus on the semileptonic decays, as well as a summary of the current experimental status. Chapter 2 gives a short description of the experimental apparatus. The topology and experimental features of semileptonic charm decays at LHCb are introduced in Chapter 3. Several algorithms which can be used to account/correct for the missing neutrino are discussed in Chapter 4. Chapter 5 describes the data sample and the specific selection requirement used to isolate the signal and suppress the relevant backgrounds. Chapter 6 explains the analysis procedure developed to statistically separate the semileptonic signal from remaining backgrounds. A first attempt to fit the data is also there presented, before the final conclusions and prospects are drawn.

1

CHARM MIXING WITH SEMILEPTONIC DECAYS

This chapter introduces the theoretical background and the motivation for the study of charm mixing with semileptonic decays. It first discusses the general formalism describing neutral meson mixing and then focuses the attention on the charm system, explaining why it is a sensible probe for physics beyond the Standard Model. An overview of how charm mixing is experimentally measured and a summary of the available results is also given at the end, with particular interest in the semileptonic decays here studied.

1.1 NEUTRAL MESON-ANTIMESON MIXING

Mesons are subatomic particles consisting of a quark-antiquark state, bound together by strong interactions. Mesons made of quark and antiquarks of different flavour are usually referred to as *flavoured* mesons and can be identified by an associated quantum number. Flavoured mesons composed by pairs of the first-generation quarks, *i.e.* the up (u), the down (d) and the corresponding antiquarks ($\overline{u}$, $\overline{d}$), are called *light* mesons; flavoured mesons containing a strange (s), a charm (c), a bottom (b) quark or the corresponding antiquarks ($\overline{s}$, $\overline{c}$ and $\overline{b}$) are called respectively kaons (or K), D and B mesons.

Contrarily to the neutral pion, which is composed by a linear combination of $u\,\overline{u}$ and $d\,\overline{d}$ pairs, flavoured mesons with neutral electric charge appear in two different charge-conjugated[1] states with distinct quantum numbers, hereafter denoted as P^0 and $\overline{P}^0$ respectively. As originally postulated by Gell-Mann and Pais in 1955 [1], this allows for flavour-changing $P^0 \to \overline{P}^0$ and $\overline{P}^0 \to P^0$ transitions occurring before the decay. In particular, for times t much larger than the typical strong interaction scale, the time evolution of the neutral meson-antimeson system is described by the following effective Schrödinger equation:

$$i\frac{\partial}{\partial t}\begin{pmatrix}|P^0(t)\rangle\\ |\overline{P}^0(t)\rangle\end{pmatrix} = \left[\mathbf{M} - \frac{i}{2}\mathbf{\Gamma}\right]\begin{pmatrix}|P^0(t)\rangle\\ |\overline{P}^0(t)\rangle\end{pmatrix}, \tag{1}$$

where $\mathbf{M}$ and $\mathbf{\Gamma}$ are 2×2 Hermitian matrices,

$$\mathbf{M} = \begin{pmatrix}M_{11} & M_{12}\\ M_{12}^* & M_{22}\end{pmatrix} \quad \text{and} \quad \mathbf{\Gamma} = \begin{pmatrix}\Gamma_{11} & \Gamma_{12}\\ \Gamma_{12}^* & \Gamma_{22}\end{pmatrix},$$

1 Charge conjugation (C) inverts the sign of all charges, thus transforming a particle into its antiparticle.

associated with transitions via *off-shell* (dispersive) and *on-shell* (absorptive) intermediate states, respectively. The diagonal elements of the two matrices are associated with the flavour-conserving time-development $P^0 \to P^0$ and $\overline{P}^0 \to \overline{P}^0$ and must satisfy $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$ to be consistent with CPT invariance[2]. The off-diagonal elements are instead associated with flavour-changing transitions $P^0 \to \overline{P}^0$ and $\overline{P}^0 \to P^0$. The eigenstates resulting from Equation 1, $|P_1\rangle$ and $|P_2\rangle$, can be expressed as linear combinations of the flavour eigenstates

$$|P_{1,2}\rangle = p\left|P^0\right\rangle \pm q\left|\overline{P}^0\right\rangle,$$

where p and q are complex parameters satisfying the relations $|p^2| + |q^2| = 1$ and $q/p = \pm\sqrt{2M_{12}^* - i\Gamma_{12}^*)}/\sqrt{(2M_{12} - i\Gamma_{12})}$. The real and imaginary part of the corresponding eigenvalues,

$$\omega_{1,2} = (M_{11} - \frac{i}{2}\Gamma_{11}) \pm \frac{q}{p}(M_{12} - \frac{i}{2}\Gamma_{12}) \equiv m_{1,2} - \frac{i}{2}\Gamma_{1,2},$$

identify the masses, $m_{1,2}$, and the decay widths, $\Gamma_{1,2}$, of the two eigenstates, respectively. Since the flavour-changing transitions occur only if the two eigenstates have different masses or widths, flavour mixing is usually characterized by the following two dimensionless parameters:

$$x = \frac{m_1 - m_2}{\Gamma}, \qquad y = \frac{\Gamma_1 - \Gamma_2}{2\Gamma}, \tag{2}$$

with the average decay width defined as $\Gamma = (\Gamma_1 + \Gamma_2)/2$. The probability for an initially (at $t = 0$) produced P^0 meson to oscillate or not into $\overline{P}^0$ are then given respectively by

$$\left|\left\langle P^0(t)\,\middle|\,P^0(0)\right\rangle\right|^2 \propto e^{-\Gamma t}[\cosh(y\Gamma t) + \cos(x\Gamma t)],$$
$$\left|\left\langle \overline{P}^0(t)\,\middle|\,P^0(0)\right\rangle\right|^2 \propto \left|\frac{q}{p}\right|^2 e^{-\Gamma t}[\cosh(y\Gamma t) - \cos(x\Gamma t)].$$

In the presence of flavour mixing, the typical exponential decay is then modified with an oscillating term governed by the mass difference x and modulated by the hyperbolic cosine term with argument proportional to the decay-width difference y.

The CP symmetry is violated in the mixing process if the probability for the transition $P^0 \to \overline{P}^0$ is different than that of $\overline{P}^0 \to P^0$. This is only possible if $|q/p| \neq 1$. CP violation can also happen in the interference between a decay following

2 The CPT operator is the product of the three symmetry transformations C (charge conjugation), P (parity) and T (time reversal) interchanging particles with antiparticles, changing sign to the spatial coordinates ($x \to -x$) and reversing the time direction ($t \to -t$), respectively. While any of the three individual symmetries may be violated, the CPT symmetry must be conserved in any local field theory with Lorentz invariance [2].

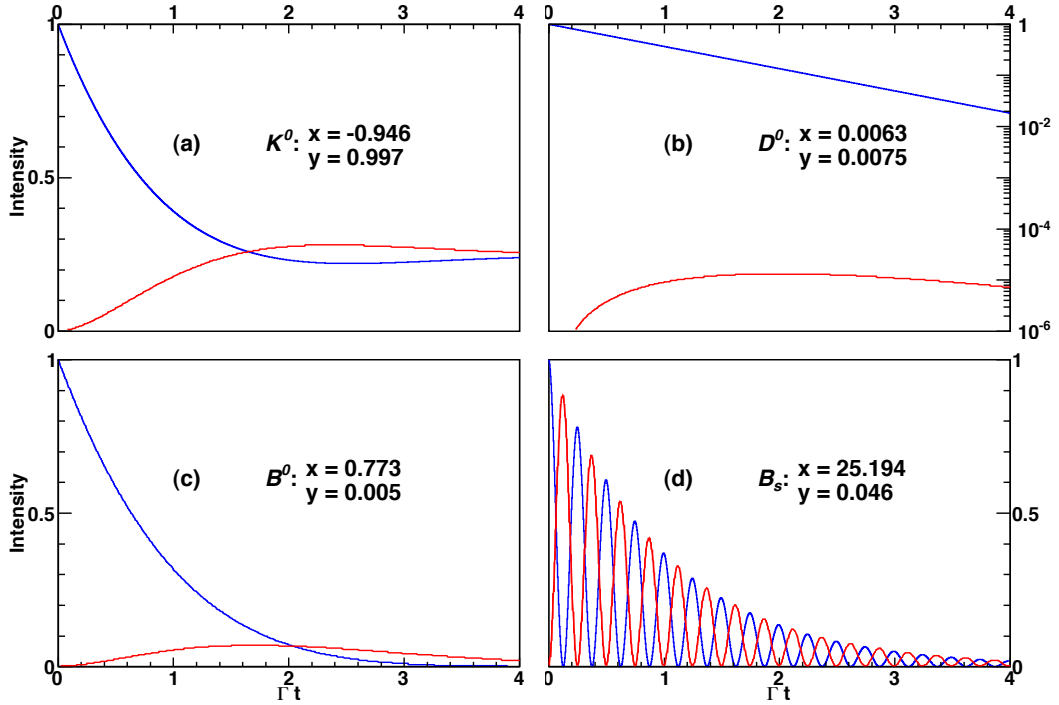


Figure 1: The (red) mixed and (blue) unmixed intensities as a function of normalized decay time Γt, for initially pure (a) K^0, (b) D^0, (c) B^0 and (d) B_s^0 mesons. The plot is taken from Ref. [3] and assumes $|q/p| = 1$.

mixing and a decay without mixing to the same final state. In that case, the phase $\arg(q/p) \neq 0$.[3]

Each neutral meson-antimeson system has its own set of mixing parameters (Equation 2), thus flavour oscillations result in different behaviours depending on the system considered, as shown in Figure 1. Neutral kaons have both very different masses and widths, with $|x|, |y| \sim \mathcal{O}(1)$, and were the system where flavour mixing was firstly observed in 1956 [4]. Oscillations in the B^0-$\bar{B}^0$ are instead dominated by the large mass difference, with $|x| \sim \mathcal{O}(1)$ and $|y| \sim 0.5\%$, and were observed in 1987 [5]. The B_s^0-$\bar{B}_s^0$ mixing, where $|x| \sim 20$ and $|y| \sim 5\%$, exhibits the fastest oscillation frequency and required experiments with very accurate decay-time resolution and was observed in 2006 [6]. The year after, several measurements showed first evidence for D^0-$\bar{D}^0$ oscillations [7, 8, 9]. These, with $|x|, |y| \lesssim 0.5\%$, occur with the smallest rate and require huge data samples to be observed. Indeed, only in 2012 the LHCb collaboration published the first observation of neutral charm mixing from a single measurement [10].

1.2 CHARM MIXING IN THE STANDARD MODEL AND BEYOND

Neutral meson-antimeson oscillations are interesting processes as they are sensitive to potential not-yet-observed heavy particles which propagate in the under-

3 Setting the weak phase of the ratio of decay amplitudes $\mathcal{A}_f/\bar{\mathcal{A}}_f$ to final state f to zero.

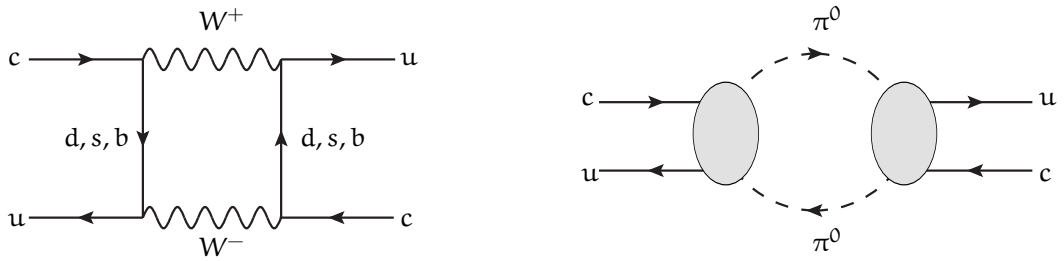


Figure 2: Two Feynman diagrams describing processes contributing to neutral charm mixing. Left: example for a short-range box diagram, where all possible down-type quarks can run in the loop. Right: transition between the D^0 and the $\overline{D}^0$ states via long-range intermediate-meson states (here neutral pions).

lying flavour changing amplitudes. The observation of mixing in the K^0 and B^0 systems implied the existence respectively of the charm and top quarks before these particles were actually discovered. Similarly, by comparing observed meson mixing with predictions of the Standard Model (SM), modern experimental studies have been able to constrain new physics models. Thanks to many years of experimental investigations dedicated to the K^0, B^0 and B_s^0 oscillations, it is today clear that the initial hopes for large new physics contributions affecting these processes have come to naught, as the SM successfully describes all the available experimental results. While these investigations have and will continue to play a central role in our quest to understand flavour physics, in-depth examinations of the D mesons sector have only recently started reaching the needed level of precision to be sensitive to possible NP contributions.

To better quantify this statement and motivate why studying of charm mixing provides a unique, not-yet-exploited probe for physics beyond the SM, it is useful to explain briefly how charm flavour-changing neutral currents are generated in the SM framework. Flavour-changing processes in the SM are only possible through weak interactions mediated by an electrically charged $W^\pm$ boson (*charged currents*). These transitions are of the form $q \to W^+ q'$ ($q' \to W^- q$), where q and q' are respectively an up- (u, c, t) and a down-type (d, s, b) quark, and happen with probabilities proportional to the Cabibbo-Kobayashi-Maskawa (CKM) matrix element $V_{qq'}$ ($V^*_{qq'}$) [11, 12]. As these processes change the flavour by only one unit, flavour-changing neutral currents such as $D^0 \to \overline{D}^0$, where the charm flavour is changed by two units, have to involve processes of higher order. Figure 2 shows two examples of SM processes contributing to the charm mixing amplitude. The first is a short-distance contributions made by a box diagrams, where a simultaneous exchange of two weak bosons allows for the transition between a charm and and an anticharm quark. In the box diagrams, all possible down-type quarks can participate. This is contrary to neutral kaon or $B^0_{(s)}$ meson systems, where only the up-type quarks enter, and therefore provides a unique opportunity to study effects of physics beyond the SM in the down-type quark sector. The second possible form of flavour-changing neutral currents in the SM, also shown in Figure 2, are transitions via intermediate states accessible to both the D^0 and the $\overline{D}^0$ mesons, referred to as long-distance contributions.

The SM predicts the contributions of the box diagrams to be very small. This is due to the small CKM coupling to the b quark, which is proportional to the factor $|V_{ub}V_{cb}^*|^2/|V_{us}V_{cs}^*|^2 \sim \mathcal{O}(10^{-6})$, and the destructively interfering contributions of the light d and s quarks, historically known as the Glashow-Illiopoulos-Maiani (GIM) mechanism [13]. In total, the prediction for the short-distance SM contributions are $\mathcal{O}(10^{-5})$ for x and $\mathcal{O}(10^{-7})$ for y [14]. For this reason, the dominant contributions to charm mixing are expected to come from long-distance intermediate-meson states, which cannot be calculated perturbatively and thus are hard to estimate. Different approaches for the SM contributions yield estimations differing very much (*e.g.* $|x|$, $|y| \leqslant 10^{-3}$ in Ref. [15, 16, 17] or $|y| \approx 10^{-2}$, $|x| \in (-1.0, -0.1) \times y$ in Ref. [18, 19]). Estimations for contributions of models extending the SM also differ by orders of magnitude, but can enhance the mixing rate significantly [20]. However, the large theoretical uncertainties for the SM contributions make it hard to judge, if measured values of x and y are hints of physics beyond the SM or not.

In the charm system, to a very good approximation, only two families of quarks contribute to mixing and decay [21]. The CKM mechanism, however, needs at least three families to generate CP violation [22]. For that reason, SM estimations of CP violation in the charm system are tiny and the measurement CP-violating asymmetries can be a powerful probe to test models beyond the SM, providing additional sources of CP violation.

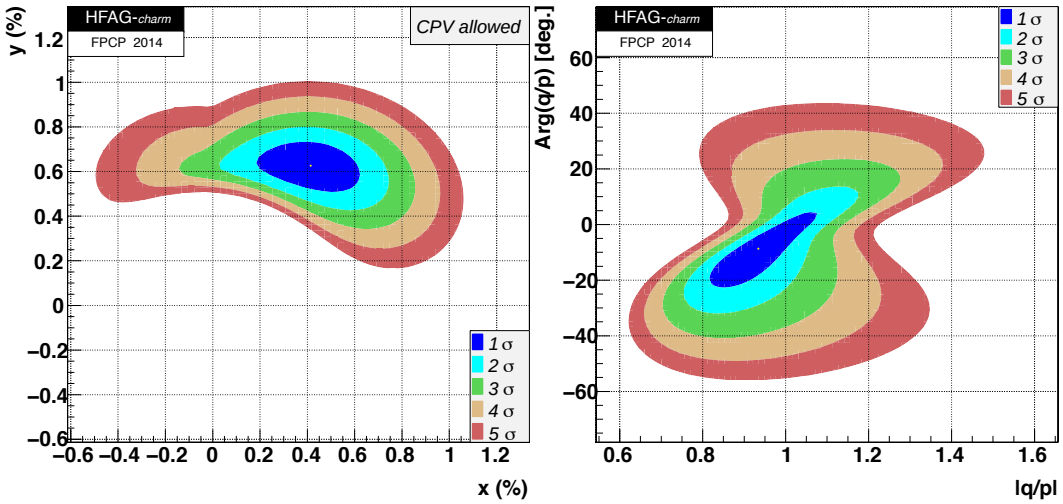


Figure 3: Two dimensional confidence-level regions for the charm mixing parameters (left) $x - y$ and (right) $|p/q| - \arg(q/p)$, corresponding to $1 - 5$ Gaussian standard deviations (σ) [23]. The no-mixing [no-CP-violation] hypothesis correspond to the $(0, 0)$ [$(0, 1)$] point.

To date, the combination of all available experimental measurements, as performed by the Heavy Flavour Averaging Group [23], yield mixing parameters,

$$x = \left(0.41^{+0.14}_{-0.15}\right)\,\% \quad \text{and} \quad y = \left(0.63^{+0.07}_{-0.08}\right)\,\%,$$

and CP-violating parameters,

$$|q/p| = 0.93^{+0.09}_{-0.08} \quad \text{and} \quad \arg(q/p) = \left(-8.7^{+8.7}_{-9.1}\right)^{\circ},$$

compatible with the SM expectations and with conservation of CP symmetry (Figure 3). These results do not yet close the possibility to observe new-physics effects in the neutral D meson sector. Hence, additional and more precise measurements are needed.

1.3 EXPERIMENTAL ASPECTS OF CHARM MIXING MEASUREMENTS

The cross section for $c\,\bar{c}$ production in proton-proton interactions is huge, $\sigma(c\bar{c}) \approx 1.4\,\text{mb}$ at a center-of-mass energy of $\sqrt{s} = 7$ TeV [24][4]. Therefore the LHCb experiment at the CERN's Large Hadron Collider, that thanks to its dedicated design to study heavy-flavoured mesons collected so far the world's largest sample of charm decays, is the ideal place to search for new physics through measurements of mixing and CP violation in charm decays.

To determine whether a decay of a neutral charm meson followed a flavour oscillation or not, it is necessary to determine (*tag*) the flavour of the neutral D at production time and compare it to that at the moment of its decay. Contrarily to the case of the other neutral meson systems, where flavour tagging requires dedicated experimental setups or complicated analyses procedures, it is possible to determine the production flavour of a neutral D meson by simply restricting the analysis to those produced in the decay of a $D^*(2010)^{\pm}$ meson, hereafter indicated simply as $D^{*\pm}$. This meson decays, with approximately 68% branching ratio [25], through strong (flavour-conserving) interactions as $D^{*+} \to D^0\pi^+$ or $D^{*-} \to \bar{D}^0\pi^-$. Hence, the charge of the charged pion unambiguously indicates the flavour of the neutral D meson at production time.[5] The small mass difference between the D^{*+} and its decay products, $Q = m_{D^{*+}} - m_{D^0} - m_{\pi^+} \approx 5.8$ MeV [25], limits the available kinetic energy of the daughter particles. For this reason the charged pion from the $D^{*\pm}$ decay is usually referred as *soft* pion. The low Q-value is also experimentally useful as it allows smaller mass resolutions, helping to drastically reduce the amount of combinatorial background. As better described in Chapter 3, requiring the presence of the parent $D^{*\pm}$ is essential for reconstructing decays, such as those studied in this thesis, where the D^0 final state is partially reconstructed.

The flavour at decay time can only be inferred by analysing the decay products. For final states that are not CP self-conjugated, it is conventional to define *right-sign* (RS) and *wrong-sign* (WS) decays by comparing the charge of a special probe

4 Measured in the phase space available at LHCb

5 Another possibility would be to look at D mesons coming from the semileptonic decays of b hadrons, where the charge of the lepton tags the D production flavour. This method was firstly used in LHCb, where is particularly useful as it leads to very small mistag rates [26].

particle in a multi-particle final state with the charge of the soft pion.[6] As charm mesons decay mostly through Cabibbo-favoured $c \to W^+ s$ transition, the probe particle is usually identified with a charged kaon. Decay chains where the kaon and the soft pion have opposite charge,[7]

$$D^{*+} \to D^0(\to K^- X^+)\pi^+, \tag{3}$$

are called RS, while those where the kaon and the soft pion have same charge,

$$D^{*+} \to D^0(\to K^+ X^-)\pi^+, \tag{4}$$

are called WS. The latter can also proceed through a $D^0 \to \overline{D}^0$ transition followed by the Cabibbo-favoured decay $\overline{D}^0 \to K^+ X^-$ and is sensitive to mixing. If the nature of the other final-state particles, indicated generically in Equation 3 and 4 as $X^\pm$, is purely hadronic, the WS process can also occur through a direct doubly-Cabbibo-supressed $c \to W^+(\to u\bar{s})d$ transition.[8] In this case, the mixing can be distinguished from the doubly-Cabibbo-suppressed contribution by studying the time-dependent WS rate. For an initially tagged D^0 mesons, this reads as

$$N^+_{WS}(t) \propto \left|\frac{q}{p}\right|^2 \left| g_+(t)\, \frac{p}{q}\mathcal{A}(D^0 \to K^+X^-) + \mathcal{A}(\overline{D}^0 \to K^+X^-) g_-(t)\right|^2, \tag{5}$$

where $\mathcal{A}$ indicates the direct decay amplitudes and the time-dependent coefficients are

$$g_\pm(t) = \frac{1}{2}\left(e^{-i\omega_1 t} \pm e^{-i\omega_2 t}\right).$$

It is experimentally convenient to normalize the time-dependent WS rate to that of the RS decays,

$$N^+_{RS}(t) \propto |\mathcal{A}(\overline{D}^0 \to K^+X^-)|^2 e^{-\Gamma t},$$

so that many systematic uncertainties cancel:

$$R^+(t) = \frac{N_{WS}(t)}{N_{RS}(t)} = \frac{1}{e^{-\Gamma t}}\left|\frac{q}{p}\right|^2 \left| g_+(t)\, \frac{p\mathcal{A}(D^0 \to K^+X^-)}{q\mathcal{A}(\overline{D}^0 \to K^+X^-)} + g_-(t)\right|^2. \tag{6}$$

The analogous expression for initially tagged $\overline{D}^0$ mesons, $R^-(t)$, can be obtained by replacing q/p with p/q and $\mathcal{A}(D^0 \to K^+X^-)/\mathcal{A}(\overline{D}^0 \to K^+X^-)$ with $\mathcal{A}(\overline{D}^0 \to K^-X^+)/\mathcal{A}(D^0 \to K^-X^+)$.

6 For CP-self-conjugated final states different analysis strategies are used. For an overview see for example Ref. [25].

7 Charge-conjugated decays are from now implied unless stated otherwise.

8 Similarly, for these decays the RS final state can be reached by $D^0 \to \overline{D}^0$ followed by the doubly-Cabibbo-suppressed decay $\overline{D}^0 \to K^-X^+$. Because of the small charm mixing rate, this process is however completely negligible when compared to the direct Cabibbo-favored decays and is therefore ignored.

1.3.1 *Charm mixing with semileptonic decays*

Contrarily to hadronic decays, where there is also a doubly-Cabibbo-suppressed amplitude contributing to the WS process, semileptonic final states, such as $K^+\ell^-\bar{\nu}$, can only be reached through mixing. In this case, the number of WS events per decay time of Equation 6 takes the following simplified form

$$R^+(t) = \frac{1}{e^{-\Gamma t}} \left|\frac{q}{p}\right|^2 |g_-(t)|^2 \approx \left|\frac{q}{p}\right|^2 \frac{x^2+y^2}{4} (\Gamma\, t)^2, \tag{7}$$

where the right-hand side of the equation has been expanded to second order for $|x|\,|y| \ll 1/(\Gamma t)$. As a consequence, for semileptonic decays no time-dependent analysis is needed and the mixing rate $R_M \equiv (x^2+y^2)/2$ can be accessed directly with a time-integrated measurement. The time-integrated WS-to-RS ratio R^+_{int} is the time integral of Equation 7, including the time-dependent selection and reconstruction efficiency $\epsilon(t)$:

$$\begin{aligned} R^+_{\text{int}} = \frac{\int N^+_{WS}(t)dt}{\int N^+_{RS}(t')dt'} &= \frac{\int e^{-\Gamma t} \left|\frac{q}{p}\right|^2 \frac{x^2+y^2}{4} (\Gamma t)^2 \epsilon(t) dt}{\int e^{-\Gamma t'} \epsilon(t') dt'} \\ &= \int R^+(t) D(t) dt \\ &= \left|\frac{q}{p}\right|^2 \frac{R_M}{2} \Gamma^2 \int t^2 D(t) dt \\ &= \left|\frac{q}{p}\right|^2 \frac{R_M}{2} \Gamma^2 \langle t^2 \rangle \end{aligned} \tag{8}$$

where $\langle t^2 \rangle$ by definition represents the average value of t^2 computed over the observed normalized distribution of decay times,

$$D(t) = \frac{e^{-\Gamma t}\epsilon(t)}{\int e^{-\Gamma t'}\epsilon(t')dt'}.$$

The time-dependent efficiency $\epsilon(t)$ and therfore also $D(t)$ are subject to the experimental environment. As the measured value of $\langle t^2 \rangle$ is depending on $D(t)$, the observed time-integrated WS-to-RS ratios may be different from experiment to experiment. In experiments where the reconstruction efficiency does not depend on decay time, as it is the case at the B-factories, the observed distribution coincides with the simple exponential decay, $D(t) = \Gamma e^{-\Gamma\, t}$. It follows that $\Gamma^2 \langle t^2 \rangle = 2$ and thus $R_{\text{int}} = R_M$ (assuming $|q/p| = 1$). On the contrary, as better detailed in Chapter 5, the data used in this analysis were collected with an online event selection (*trigger*) that imposes requirements on the displacement of the D^0 meson decay point from the production point, thus rejecting candidates with short decay times. This results in larger values of $\langle t^2 \rangle$, thus providing better experimental sensitivities to the mixing rate R_M.

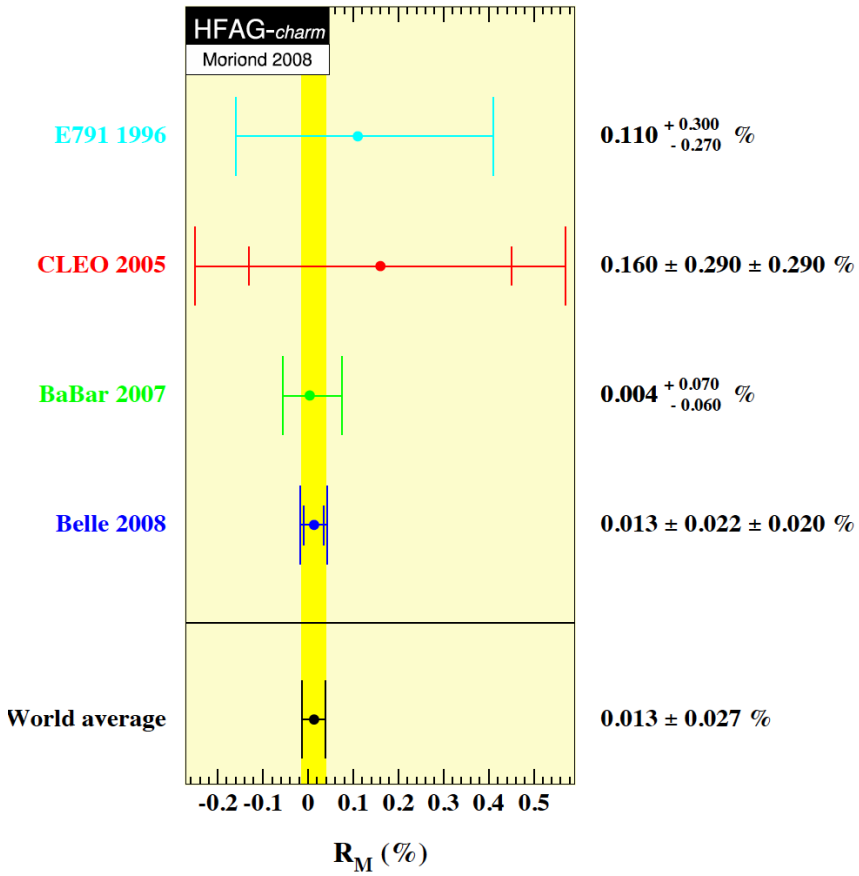


Figure 4: Experimental status on the measurement of R_M from semileptonic D^0 decays [23].

Measuring the WS-to-RS time-integrated ratio separately for initially tagged D^0 and $\overline{D}^0$ events gives sensitivity to CP violation in mixing. The experimentally convenient observable is in this case the semileptonic asymmetry:

$$A_{SL} = \frac{R^+_{int} - R^-_{int}}{R^+_{int} + R^-_{int}} = \frac{|q/p|^2 - |p/q|^2}{|q/p|^2 + |p/q|^2}. \tag{9}$$

In the presence of CP violation in the mixing $|q/p| \neq 1$ and $A_{SL} \neq 0$.

Although the semileptonic decay channel looks very attractive because of its theoretical simplicity, it also implies difficult experimental challenges. Notably, the presence of the undetectable neutrino in the final state does not allow for a complete reconstruction of the decay chain. This, particularly at hadron-collider experiments, greatly reduces the ability to isolate the rare WS decays from other D^0 decays, where final state particles are misreconstructed or not detected, and from irreducible combinatorial backgrounds. These challenges require sophisticated analysis strategies, as it will be better described in the rest of this thesis, which represents the first approach to measure R_M using semileptonic decays at the LHCb experiment.

Currently available experimental results are summarised in Figure 4. Measurements have been performed by different experiment: E791 [27], CLEO [28], BaBar [29,

30] and Belle [31]. The world average is dominated by the Belle result with a precision of $\approx 0.03\%$ and is compatible with $R_M = 0$, further confirming that measuring charm mixing with semileptonic decays requires huge data samples and challenging analysis strategies.

2

EXPERIMENTAL APPARATUS

This chapter gives a brief description of the Large Hadron Collider and the LHCb detector, focusing on the subsystems most important for the study of semileptonic charm decays such as the tracking, the particle identification and the trigger systems.

2.1 THE LARGE HADRON COLLIDER

The *Large Hadron Collider* (LHC) [32] is a circular hadron collider located at the European Organization for Nuclear Research (CERN) close to Geneva, Switzerland. It represents the final stage of a more complex accelerators system, entirely represented in Figure 5, which provides beams to different typologies of experiments. In its main operational mode the LHC receives bunches of 450 GeV protons from the *Super Proton Synchrotron* (SPS) and further accelerates them in two storage rings, working in opposite directions. At four positions along the LHC 27 km length the two beams cross each other, producing collisions at a frequency of 40 MHz. These interaction points are surrounded by particle detectors, which record the products of the collision selecting interesting events in a fraction of a second.

The LHC operated with proton-proton collisions at an energy of 3.5 TeV per beam in 2010 and 2011 and at 4 TeV in 2012. Beyond this main operation mode, the LHC also collided protons with lead ions for two months in 2013 and used lead-lead collisions for about one month each in 2010 and 2011. The LHC is currently in a shutdown period for major upgrades to increase proton beam energy to 6.5 TeV, and further approaching its goal center-of-mass energy of $\sqrt{s} = 14$ TeV, with restarting planned for early 2015.

2.2 THE LHCb EXPERIMENT

The *Large Hadron Collider beauty* (LHCb) detector [34] is the only experiment at the LHC having as primary goal the search for physics beyond the SM through the study of heavy flavoured mesons containing a b or a c quark. It is mainly build to operate during pp collisions, but it has also been proficiently used to study proton-lead collisions [35].

LHCb (see Figure 6) is a single-arm spectrometer with forward angular coverage from approximately 10 to 300 (250) mrad in the bending (non-bending) plane. The bending plane is defined by the magnetic field provided by a large dipole magnet that deflects trajetories of charged particles. The special geometry is justified by the fact that at the LHC energies both $b\bar{b}$ and $c\bar{c}$ pairs are predominately produced

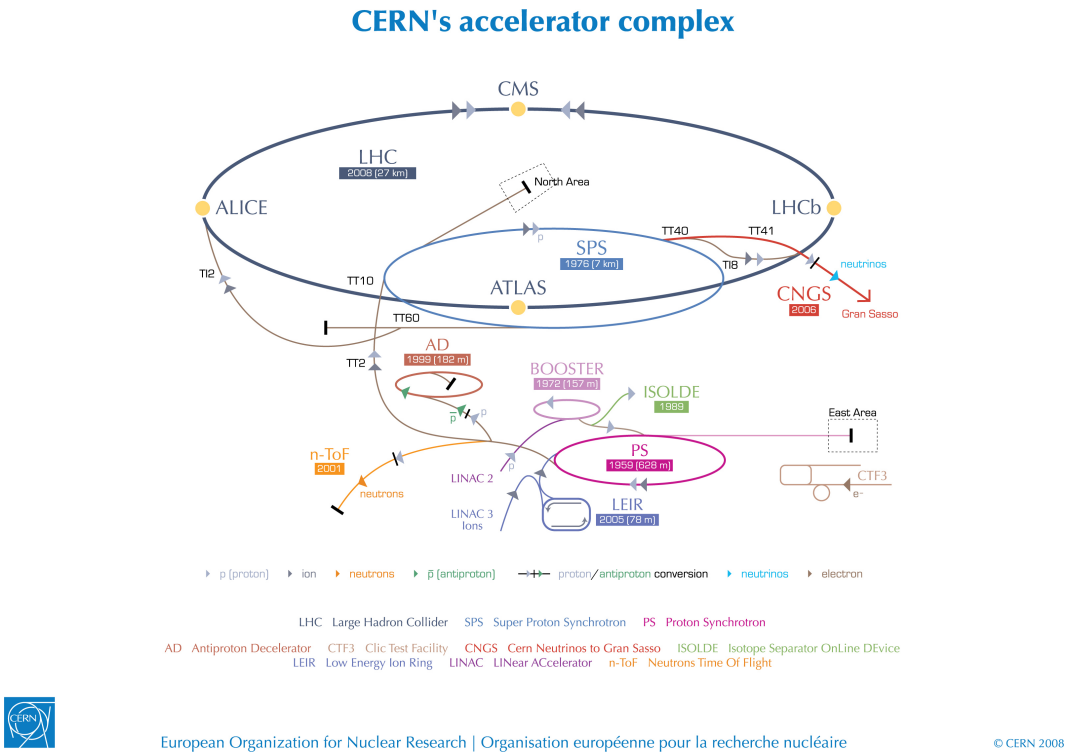


Figure 5: A schematic view of the CERN's accelerator complex and the location of the main experiments [33].

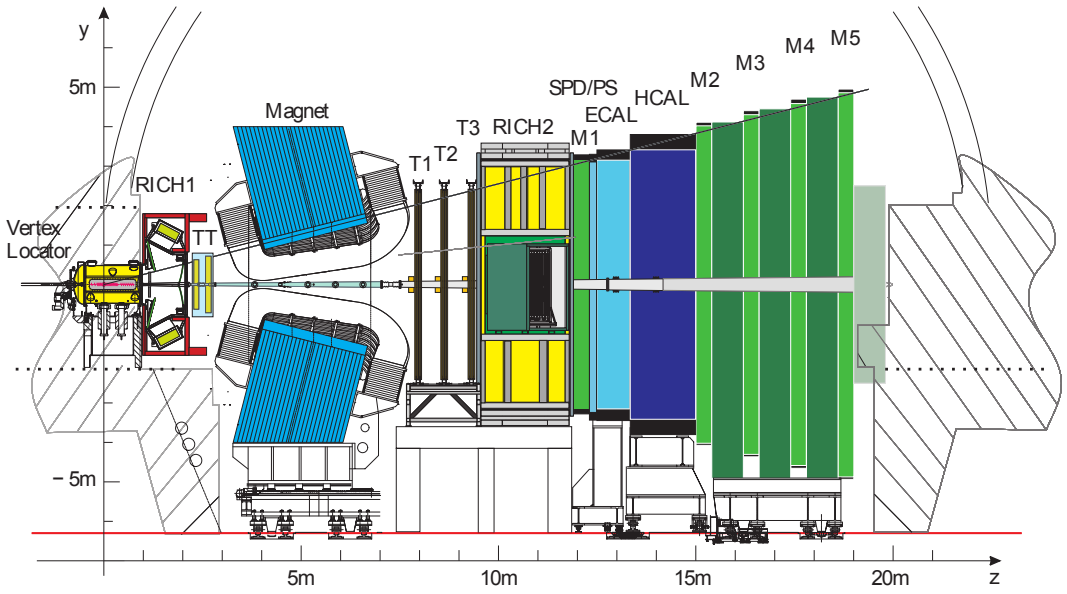


Figure 6: Schematic vertical cross section through the LHCb detector, showing the various subsystems [34].

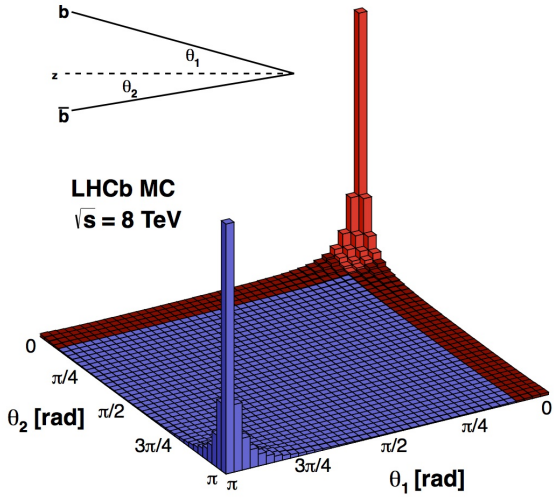


Figure 7: Simulated angular distribution of $b\bar{b}$ pair production in pp collisions at center-of-mass energy of 8 TeV, corresponding to the conditions of the 2012 data-taking period. The LHCb angular acceptance is coloured in red [36].

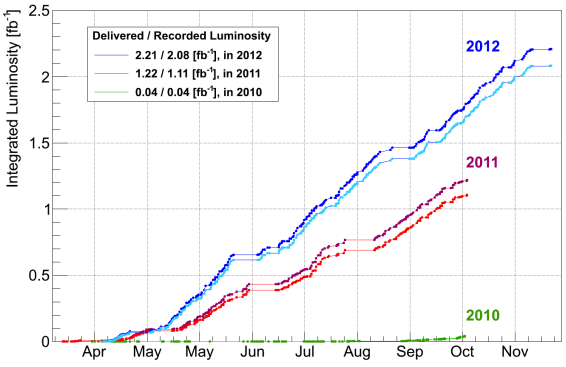


Figure 8: Delivered and recorded luminosity by the LHCb detector as a function of time, separately for 2010, 2011 and 2012 data-taking periods [36].

in the forward direction, as shown in Figure 7. Heavy flavoured mesons flying in this direction are highly boosted and travel measurable distances in detector before they decay. Several subsystems allow for high-resolution track reconstruction, measurement of engery and momentum as well as particle identification.

To efficiently operate at the LHC energies and rates, the instantaneous luminosity provided by the accelerator is reduced at the LHCb intersection point by slightly separating the two colliding proton beams. This procedure allows to keep a constant instantaneous luminosity profile during data taking and maintain stable detector performances. Figure 8 shows the delivered luminosity, as well as that recorded by LHCb, as function of time. The detector operated with high efficiency during the whole data-taking period and recorded in total approximately

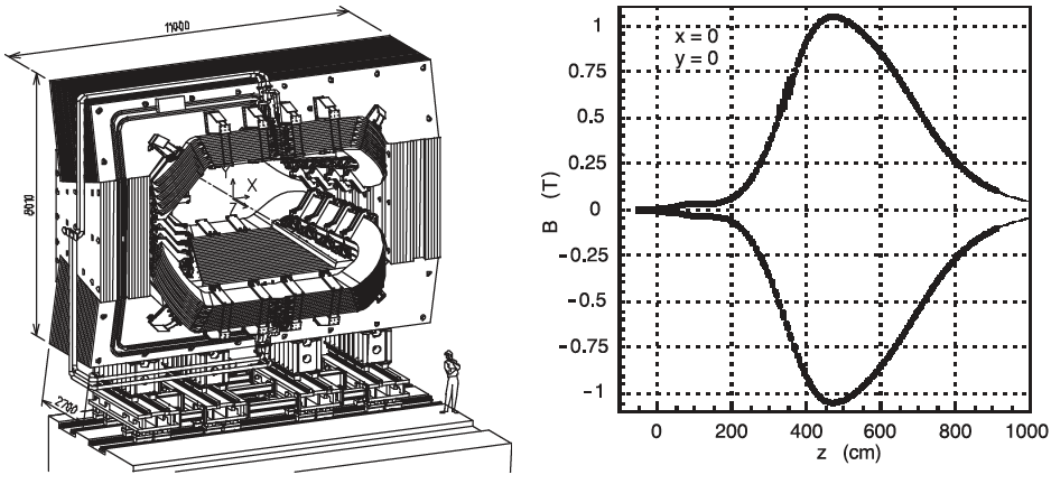


Figure 9: The LHCb dipole magnet (left) and the magnetic field along the z-axis (right) [34].

$40\,\mathrm{pb}^{-1}$, $1\,\mathrm{fb}^{-1}$ and $2\,\mathrm{fb}^{-1}$ of integrated luminosity during 2010, 2011 and 2012, respectively.

Before giving a more detailed description of the subdetectors that are relevant for the analysis presented here, it is useful to introduce the coordinate system adopted at LHCb (see also Figure 6). This is a right-handed Cartesian system with z-axis along the beam and y-axis pointing vertically upwards.[1] From this also a spherical coordinate system is deduced with azimuthal and polar angle indicated by ϕ and θ, respectively, where the polar angle also defines the pseudorapidity as $\eta = -\ln \tan(\theta/2)$.

2.2.1 *The tracking system*

Tracking means measuring the trajectories of charged particles (*tracks*) and, in the presence of a magnetic field, their charge and momentum. In LHCb this is achieved with a system of subdetectors comprising a large non-superconducting dipole magnet [37]. The magnet, as shown in Figure 9, provides a magnetic field oriented mainly in y-direction corresponding to bending power of $4\,\mathrm{Tm}$ for a track of $10\,\mathrm{m}$ length in z-direction. To mitigate detector-induced left-right asymmetries the magnet polarity is regularly switched and a roughly equal amount of data is recorded with each polarity.

Figure 9 shows a schematic view of the LHCb dipole and magnetic field along the z-axis.

Two subdetectors are placed upstream of the dipole magnet: the *VErtex LOcator* (VELO) and the *Tracker Turicensis* (TT). The VELO [38] is the detector situated closest to the interaction point, as it is dedicated to the precise reconstruction of the position where heavy-flavoured hadrons decay. Because of their typical lifetimes of $\sim 1\,\mathrm{ps}$ and their heavy boost, b- and c-hadrons travel distances of the order of

1 As a result, the x-axis points away from the centre of the beam ring

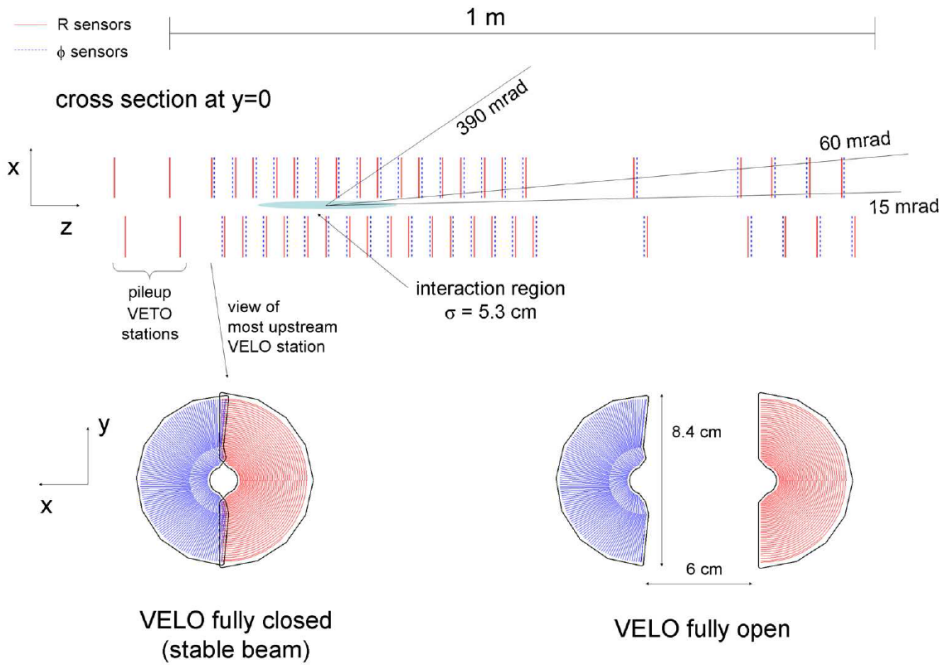


Figure 10: The upper part shows a schematic view of the arrangement of the 21 velo stations as well as the 2 pileup veto stations in the x-z plane at y=0. The lower part shows one of the stations in the x-y plane. The two semicircle shaped parts can can be moved apart when the beam is unstable [34].

1 cm before decaying. The resulting decay products form then *secondary* vertices that are well displaced from the vertex of the primary pp interaction. The VELO allows to precisely measure the position of these vertices and determines the decay time with a resolution for a fully reconstructed decay of approximately 45 fs. The detector consists of 21 stations arranged in a row, as shown in Figure 10. Each stations is divided into two semicircle shaped modules. On each module two different types of silicon-strip sensors are mounted. One of them is measuring radial distances (R-sensors), the other is used to determine azimuthal angles (ϕ-sensors). Information of the third dimension is inferred from the position of the sensors along the z-axis. Four additional R-sensors, placed behind the interaction point in negative z-direction, form the pileup veto station which is used for triggering purposes. The VELO is designed to measure particles in the range $1.6 < |\eta| < 4.9$ and $|z| < 10.6\,\text{cm}$, as for particles in this region at least three modules are hit. In order to separate the vacuum inside the VELO from the ultra-high vacuum of the LHC, the modules are covered by a thin-walled aluminum box. The two-halves design allows to move the detector away from the beam when this is unstable and threatens the sensitive electronics. It is possible to measure the minimum distance of a track to a primary vertex with a spatial resolution of $(15 + 29/p_T)\,\mu\text{m}$. Here, p_T denotes the component of the momentum in GeV/c, which is transverse to the beam axis.

The TT [34], situated approximately at $z = 2.5\,\text{m}$, is a silicon microstrip detector consisting of two pairs of single-sided detector layers for a total size of 150 cm in

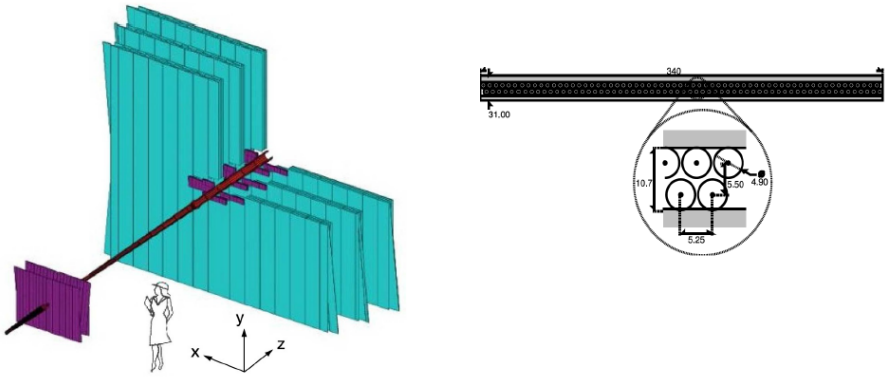


Figure 11: Left: Schematic view of the TT and IT silicon strip detectors in purple and the OT straw-tube detector in cyan. Right: View inside an OT module, showing the arangement of the straw-tubes inside [34].

width and 130 cm in height. In each pair one layer is tilted by a relative rotation of $\pm 5^\circ$ to allow for a three dimensiomal track measurement with a spatial resolution of 50 µm. The main purpose of the detector is to allow the reconstruction of both low-momentum particles, that are bent by the magnetic field out of the acceptance of the downstream tracking stations, and long-lived particles, such as K^0_S mesons and Λ^0 baryons, that mostly decay after the VELO.

The tracking system is completed by three main tracking stations (T1-T3) located downstream of the magnet. Each tracking station comprises two subsystems and consists of four layers arranged in the tilted configuration like the TT. The first subsystem is called *Inner Tracker* (IT) [39] as it covers the 120 cm wide and 40 cm high region closest to the beam pipe, where the largest particle-flux is expected. It consists of high-resolution silicon-microstrip sensors which provide a spatial resolution of 50 µm. The second subsystem is called *Outer Tracker* (OT) [40]. It is made of drift tubes with an inner diameter of 4.9 mm, filled with a mixture of Ar, CO_2 and O_2. The spatial resolution of 200 µm is not as good as for the IT, but on the other hand a large area of approximately $6 \times 5\,m^2$ is covered. Each of the four layers of one station consists of two layers of drift tubes, as shown in the right part of Figure 11. The momentum resolution provided by the tracking system is 0.4% at low momentum and 0.6% at 100 GeV/c.

2.2.2 *The particle identification system*

Particle identification (PID) is an essential tool for many flavour-physics analyses. For example, as explained in Section 1.3, the flavour of the neutral D mesons at decay time is tagged by the charge of a kaon in the final state. It is therefore necessary to identify the type of the particles in the final state with high reliability. In LHCb several subdetectors are devoted to PID: two ring-imaging Cherenkov

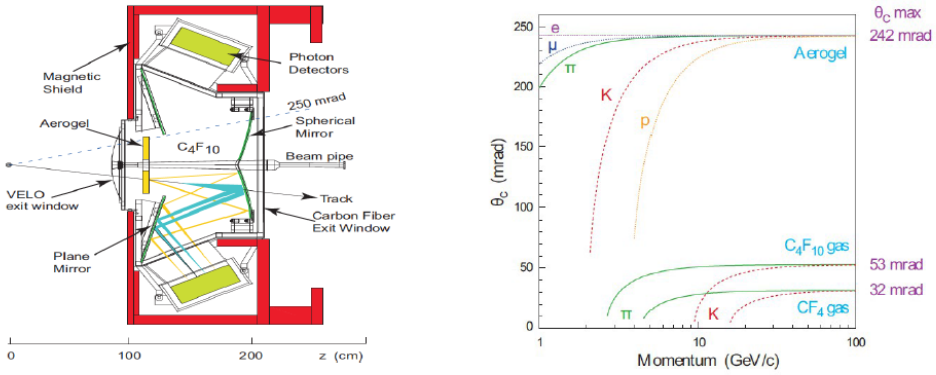


Figure 12: Left: schematic view of the RICH1 detector, showing how the combination of a spherical and a plane mirror reflects the emitted Cherenkov light to the photo-detectors. Right: the Cherenkov angle as function of the particle momentum for different radiators and particle species [34].

detectors [41] (RICH1 and RICH2) are used for charged hadrons, a calorimeter system identifies electrons and photons, the muon system takes care of muons.

The technique to identify hadrons is based on the Cherenkov effect. If a charged particle travels through a medium faster than the speed of light in that medium, electromagnetic radiation is emitted in the form of a cone around the particle trajectory. The cone opening angle, θ_C, depends on the speed of the particle v and on the refractive index of the medium n as $\cos\theta_C = c/(nv)$, so that its knowledge, together with the particle momentum measured by the tracking system, can be used to infer the mass of the particle. The RICH detectors consist then in suitable radiator material and in a system of mirrors and photo-detectors which guide and collect the emitted Cherenkov radiation, as shown in Figure 12. Two detectors are employed to cover a larger momentum range, by using different radiators. RICH1 uses aerogel and is designed to identify particles in the momentum range $1-60$ GeV/c, while RICH2 with gaseous CF_4 covers the range $50-150$ GeV/c. Figure 12 also shows the obtained separation power between different particle species over the full momentum range.

Calorimetry refers to the measurement of a particle's energy. The LHCb calorimeter system [42] consists of an electromagnetic (ECAL) followed by an hadronic calorimeter (HCAL). In the ECAL, layers of 2 mm thick lead and 4 mm thick scintillating material are installed alternately to drive the shower production and detect the rising secondary particles. The energy resolution of the ECAL is $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$, where E is in GeV and $\oplus$ indicates the quadratic sum. In front of the ECAL a scintillating pad detector (SPD) provides separation between photons and electrons, as only the latter release a signal in the SPD. Separated by a layer of lead, but still in front of the ECAL, the preshower detector measures the amount of energy deposited in the lead, which is different for electrons and charged pions. The ECAL is followed by the HCAL, used to measure the energy of hadrons.

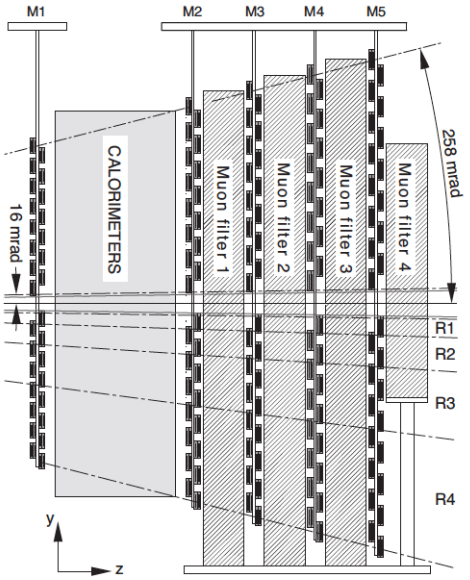


Figure 13: A schematic view of the muon system in the $y - z$ plane [43].

It consist of alternating layers of iron absorbers and scintillating tiles and has a resolution of $\sigma_E/E = 80\%/\sqrt{E} \oplus 10\%$.

Muons produced inside LHCb act as minimum ionizing particles, that means they penetrate through all the inner detector subsystems, including the calorimeters. That is why the outermost part of the detector is specialized to detect muons. As visible in Figure 13, the LHCb muon system [43] comprises one station before (M1) and four stations after (M2-M5) the calorimeter system. The M2-M5 stations are separated by 80 cm thick iron plates to further ensure that only muons can reach the last station of the system. The stations are mainly made of multi-wire proportional chambers that allow for a fast readout. This allows the muon system to play an important role in the trigger.

2.2.3 *The trigger system*

The LHC is designed to bring two proton bunches to collision every 25 ns, translating to an operating frequency of 40 MHz. This frequency has to be reduced drastically to shrink the amount of data to a recordable level. LHCb uses a two-stage trigger system [44] to quickly process the event and decide if it is worth storing it. The trigger flow is graphically summarized in Figure 14.

The first stage (called level zero, L0) is hardware based and reduces the 40 MHz bunch-crossing rate down to approximately 1 MHz. The limited time available does not allow for a sophisticated event reconstruction based on the full information provided by the detector. The decision is then taken based on the following simple signatures:

- activity in the pileup veto stations of the VELO;

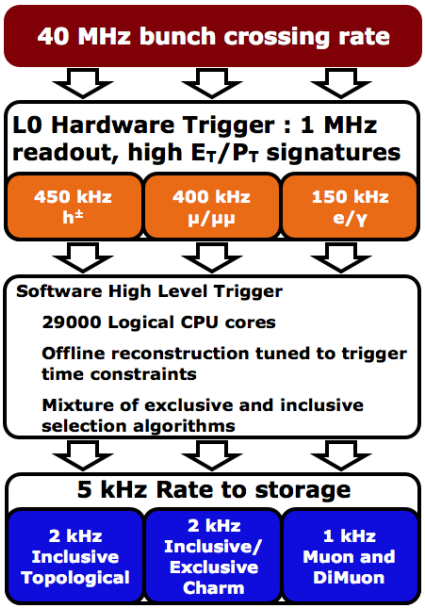


Figure 14: Summary of the LHCb trigger system, reducing the operating frequency of 40 MHz down to 5 kHz in 3 steps [36].

- large enough transverse energy depositions in the calorimeter system;
- track segments in the muon system with sufficiently large transverse momentum.

Momentum and transverse energy thresholds vary across the different data-taking periods and are better documented in Ref. [44].

The L0 trigger is followed by a software based high-level trigger, further divided in two consecutive steps (HLT1 and HLT2). In HLT1, the events passing L0-selection are enhanced by information of the tracking system. A first and very fast fitting procedure of the tracks is performed and allows to impose certain track quality requirements and select for events with displaced secondary vertices. The output rate of HTL1 is approximately 50 kHz. In HLT2 the rate is sufficiently low to perform a full event reconstruction making use of all the LHCb subsystems. Still the trigger has to meet strict timing requirements, so the reconstruction is very similar to but also somehow simplified with respect to the offline algorithm to be less CPU consuming. The total HLT2 output rate, corresponding to the rate of events stored, is 5 kHz. These 5 kHz have to be divided between different lines dedicated for various physical analysis: 2 kHz are taken by inclusive lines that select the typical topology of b-hadrons final states; another 2 kHz are reserved for inclusive and exclusive charm decays; and the rest is taken by dimuon or single-muon final states. Inclusive lines are triggering on more generic event classes, as for a example the presence of a $D^{*\pm}$. Contrarily, exclusive trigger lines allow to select for very particular final states. Exclusive lines are needed in the charm sector because of the very high $c\bar{c}$ production rate at LHCb, far exceeding the $b\bar{b}$ rate.

More details about the trigger used to collect the sample analised in this thesis are given in Section 5.1.

3

SEMILEPTONIC CHARM DECAYS AT LHCb

This chapter defines the variables used in the analysis to reconstruct and select semileptonic charm decays at LHCb. It also introduces the various sources of background contamination which can mimic the topology of the signal decay, preparing the discussion on the sample selection carried out in the next chapter.

3.1 DECAY TOPOLOGY AND EXPERIMENTAL OBSERVABLES

Figure 15 shows the topology of the $D^{*+} \to D^0 (\to K^\mp \mu^\pm \nu) \pi^+$ decay chain. The D^{*+} meson is produced in the primary pp interaction vertex, where it instantaneously decays through strong interactions into a D^0 meson and a charged pion. The charged pion interacts with the LHCb tracking detectors producing a track with typical momentum of $5\,\mathrm{GeV/c}$. The D^0 meson has a lifetime τ_{D^0} of approximately $0.41\,\mathrm{ps}$ and a typical momentum of $50\,\mathrm{GeV/c}$, hence it flies before decaying. The tracks resulting from its charged decay products intersect each other in a decay vertex, that is a few cm displaced from the primary vertex. This feature is common to all weakly-decaying particles and is widely used to identify decays of heavy-flavoured mesons. In addition to the flight distance, the impact parameter (IP) of the track, *i.e.* the minimum distance between the track and the primary vertex, can also help to judge whether a particle was produced in the primary interaction or not. As an example, the pion from the D^{*+} decay will have small IP, whereas the charged kaon and the muon are more likely to have large IP values.

The neutrino does not interact within LHCb and remains undetected. As a consequence, the kinematical properties of the D^0 decay, such as momentum, decay time and invariant mass, are only partially accessible. Chapter 4 describes different techniques that can be used to overcome this limitation and estimate the momentum carried by the neutrino. Meanwhile, it is worth to define those quantities that can be directly reconstructed:

- the visible D^0 mass, *i.e.* the invariant mass of the $K\mu$ pair

$$m(K\mu) = \sqrt{\left[E(K) + E(\mu)\right]^2 - \left[\vec{p}(K) + \vec{p}(\mu)\right]^2},$$

  where $E^2(K,\mu) = m^2_{K,\mu} + p^2(K,\mu)$ with m_K and m_μ being the known kaon and muon masses [25], respectively;

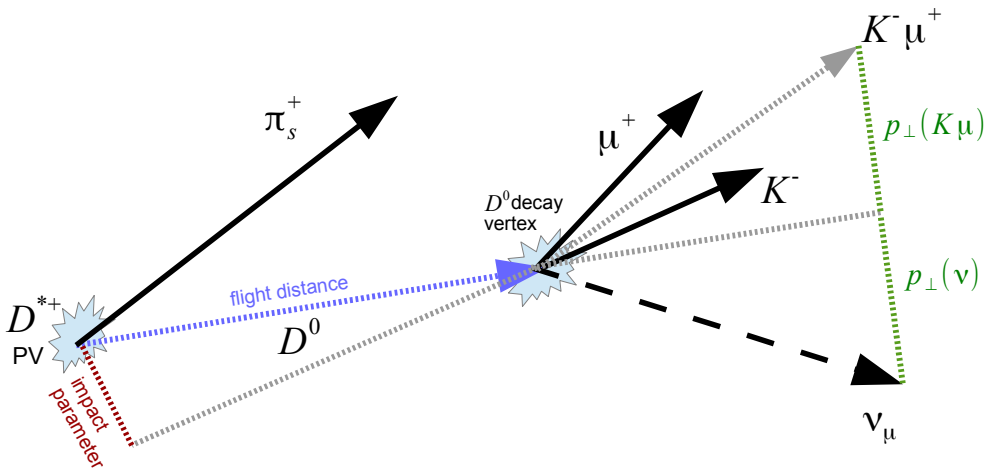


Figure 15: Schematic view in the y-z plane of the $D^{*+} \to D^0(\to K^-\mu^+\nu)\pi^+$ decay chain topology. The relative scale is arbitrary and not realistic

- the corrected D^0 mass

$$m_{corr}(D^0) = \sqrt{m(K\mu)^2 + p_\perp^2(K\mu)} + p_\perp(K\mu),$$

where the component of the $K\mu$ pair momentum perpendicular to the flight direction, $\vec{p}_\perp(K\mu)$, is used to partially account for the missing neutrino (see also Chapter 4);

- the mass of the $D^0\pi^+$ pair

$$m(D^0\pi) = \sqrt{E^2(K\mu\pi) - p^2(K\mu\pi)},$$

where

$$\begin{aligned} E(K\mu\pi) &= \sqrt{m_{D^0}^2 + p^2(K\mu)} + \sqrt{m_{\pi^+}^2 + p^2(\pi)}, \\ \vec{p}(K\mu\pi) &= \vec{p}(K\mu) + \vec{p}(\pi), \\ \vec{p}(K\mu) &= \vec{p}(K\mu) + \vec{p}(\pi), \end{aligned}$$

and m_{D^0} being the known D^0 mass [25]. This variable presents the same feature as the more customary difference between the visible D^{*+} and D^0 masses, $\Delta m = m(K\mu\pi) - m(K\mu)$, that the contribution of the momentum of the undetected neutrino cancels to a large extent. This means that the $m(D^0\pi)$ distribution of signal decays will have a peaking structure approximately at the known value of the D^{*+} mass, $m_{D^{*+}}$ [25], providing a powerful discrimination with respect to the background. In addition, contrary to Δm,

the $m(D^0\pi)$ variable has the advantage of being independent of the mass hypotheses of the D^0 daughters.

Although the D^0 is not fully reconstructed, due to its large boost, the direction of the D^0 momentum is supposed to be fairly aligned with the reconstructed flight distance. That is the reason, why the angle between the reconstructed D^0 momentum and flight distance has to be small, if all the particles in the decay chain are combined correctly. The cosine of this angle is defined as direction angle (DIRA) and expected to be close to one for signal events. Another variable used in this analysis to decide whether two particles arise from the same mother particle or not is the minimal three dimensional distance measured between two tracks, called DOCA (distance of closest apprach).

Final state particles are identified at LHCb [45] using two different sets of observables: log-likelihood differences (DLL) and neural network probabilities (ProbNN). For each charged particle, likelihood functions for several mass hypotheses are constructed using information mainly from the RICH detectors and the tracking stations. To test a certain mass hypothesis, say muon, for a particle candidate, the difference between the logarithms of the likelihoods under such hypothesis and under the pion hypothesis, $DLL_{\mu-\pi}$, is computed. This difference allows to separate muons from pions, as its value will be on average larger (smaller) than zero if the candidate is really a muon (pion). To better profit from all particle identification capabilities of the LHCb detector, a multivariate analysis based on an artificial neural network is used to combine in a single probability, for each mass hypothesis, the information coming from the tracking, RICH, muon and calorimeter systems. While DLLs can separate one particle specie from another, these ProbNN variables can be used to isolate the specified specie independently from other hypothesis. In addition, to standard particle hypothesis, the neural network based algorithms also allows to calculate the probability that a certain track is a *ghost*, *i.e.* a random association of hits accidentally combined to form a track and not caused by the passage of a real particle in the detector.

3.2 EXPECTED BACKGROUND SOURCES

To isolate the very rare mixing signal, it is important to understand which backgrounds can possibly mimic the typical signature of a WS semileptonic charm decay. Here is a list of different categories for the relevant background contributions.

OTHER SEMILEPTONIC CHARM DECAYS – Besides the $D^0 \to K^-\mu^+\nu$ decay investigated in this thesis, other semileptonic decay modes of the D^0 are possible. These can be decays with an associated electron being misidentified as a muon, a pion misidentified as a muon or other decays with same final-state particles and additional missing hadrons. An example would be

$D^0 \to K^{*-}(892)\mu^+\nu$, with branching ratio $\mathcal{B} \sim 1.9\%$, where the intermediate $K^*(892)$ state further decays to $K^-\pi^0$ and the neutral pion is not reconstructed. These decays would not bias the WS-to-RS ratio, but will have a different decay-time distribution.

HADRONIC TWO-BODY DECAYS – Hadronic two-body decays of D^0 mesons can mimic a semileptonic decay when one or both final state particles are misidentified. In particular, decays to CP-conjugated final states, such as $D^0 \to K^+K^-$ or $D^0 \to \pi^+\pi^-$, with $\mathcal{B}$ of $\sim 4 \times 10^{-3}$ and $\sim 1 \times 10^{-3}$ respectively, can contribute to both RS and WS decays through a double misidentification of the final state hadrons. Cabibbo-favoured $D^0 \to K^-\pi^+$ ($\mathcal{B} \sim 4 \times 10^{-2}$) decays can enter the RS sample if the pion is misidentified as a muon, and the WS sample if both the kaon is misidentified as a muon and the pion is misidentified as a kaon. Contamination from mixed $D^0 \to \overline{D}^0 \to K^+\pi^-$ or doubly-Cabibbo-suppressed $D^0 \to K^+\pi^-$ decays ($\mathcal{B} \sim 10^{-5}$) can be neglected for RS decays, but may be present in WS decays if the pion is misidentified as a muon. In addition to particle identification selection, the background from these decays can be easily isolated using kinematics, as explained in Section 5.2.2, since the final state is, in this case, fully reconstructed.

HADRONIC MULTI-BODY DECAYS – Hadronic D^0 decays to multi-body final states are, as the signal, partially reconstructed and therefore provide a very dangerous source of *peaking* background. They have in general large branching fractions and may be kinematically very similar to the semileptonic signal. However, hadronic decays can mimic the signal candidates only if at least one of the decay products is misidentified. Table 1 lists possible contributions to the RS and WS samples, explaining which misidentification should happen. As an example, a large contribution may be expected from $D^0 \to K^+\pi^-\pi^0$ decays, where the pion is misidentified as a muon and the neutral pion remains undetected.

RANDOM PIONS – In proton-proton collisions, lots of charged pions are produced in every interaction as they are the lightest existing mesons. If a charged pion not coming from the signal process is accidentally associated to a D^0 candidate, the decay of a charged D^* can be mimicked. This irreducible background is expected to provide the largest contribution to the WS sample, but it can be easily separated from the signal as it features a different $m(D^0\pi)$ distribution as explained in Chapter 6.

PURE COMBINATORICS – Pure combinatorial background arises when the D^{*+} candidate is made of three randomly associated tracks. Again, this background will feature different kinematical properties that allow to statistically separate it from the signal.

Channel	Branching ratio [%]	Contribution to RS	Contribution to WS
$D \to K^-\pi^+\pi^0$	13.9 ± 0.5	$\pi^+ \to \mu^+$	$\pi^+ \to K^+$,$K^- \to \mu^-$
$D \to K^-2\pi^+\pi^-$	$8.1^{+0.21}_{-0.19}$	$\pi^+ \to \mu^+$	$\pi^+ \to K^+$,$K^- \to \mu^-$
$D \to K^-2\pi^+\pi^-\pi^0$	4.2 ± 0.4	$\pi^+ \to \mu^+$	$\pi^+ \to K^+$,$K^- \to \mu^-$
$D \to \pi^-\pi^+\pi^0$	1.43 ± 0.06	$\pi^- \to K^-$,$\pi^+ \to \mu^+$	$\pi^+ \to K^+$,$\pi^- \to \mu^-$
$D \to \pi^-\pi^+2\pi^0$	1.0 ± 0.09	$\pi^- \to K^-$,$\pi^+ \to \mu^+$	$\pi^+ \to K^+$,$\pi^- \to \mu^-$
$D \to K^+K^-\pi^0$	0.329 ± 0.014	$K^+ \to \mu^+$	$K^- \to \mu^-$
$D \to K^+\pi^-\pi^0$ (DCS)	0.0304 ± 0.0017	$K^+ \to \mu^+$,$\pi^- \to K^-$	$\pi^- \to \mu^-$
$D \to K^+\pi^-\pi^0$ (through mixing)	0.073 ± 0.005	$K^+ \to \mu^+$,$\pi^- \to K^-$	$\pi^- \to \mu^-$
$D \to K^+\pi^+2\pi^-$ (DCS)	0.0262 ± 0.0021	$K^+ \to \mu^+$,$\pi^- \to K^-$	$\pi^- \to \mu^-$
$D \to K^+\pi^+2\pi^-$ (through mixing)	$<0.04(\mathrm{CL} = 90\%)$	$K^+ \to \mu^+$,$\pi^- \to K^-$	$\pi^- \to \mu^-$
$D \to K^+2K^-\pi^+$	0.0221 ± 0.0031	various	$K^- \to \mu^-$

Table 1: D^0 decays to hadronic multi-body final states and how they can possibly contribute to RS and WS samples through misidentification of final state particles. Reported branching ratios are from Ref. [25].

SECONDARY D DECAYS – If the D^{*+} is not produced in the primary pp interaction, but as decay product of a b-hadron decay, the decay length is not correctly computed. These *secondary* D^{*+} decays are therefore a background to the promptly-produced semileptonic signal in that they result in a biased decay-time distribution. They, however, do not change the observed WS-to-RS ratio.

4

RECONSTRUCTION OF THE MISSING NEUTRINO

This chapter presents prospects and problems of different reconstruction algorithms for a missing particle in decays with partially reconstructed final states. Although these techniques may be simply adapted to any kind of missing particle in different experimental environments, the main focus is on $D^0 \to K^-\mu^+\nu$ decays at LHCb, where the neutrino remains undetected.

4.1 PARTIALLY RECONSTRUCTED DECAYS AT LHCb

Neutrinos hardly interact within LHCb and are not directly detectable. This considerably complicates the study of semileptonic decays, as it is not possible to determine the total momentum, and therefore the mass, of the D^0 candidate, thus limiting the capabilities to separate the signal from backgrounds. Although the mass variable $m(D^0\pi)$ is defined such that the neutrino's missing momentum contribution cancels to a large extent, accounting for the neutrino would provide an improved mass resolution and a better signal-to-background discrimination. In addition, an unbiased and precise knowledge of the D^0 decay time, usually derived dividing the D^0 flight distance F by its velocity,

$$t = \frac{F \cdot m_{D^0}}{p(D^0)}, \tag{10}$$

is desirable to enhance the sensitivity to mixing, as the WS-to-RS ratio grows quadratically with time (see Equation 7).

Experiments at e^+e^- machines benefit from the precise knowledge of the collision energy to conveniently derive the momentum of a missing particle in the final state using momentum conservation. At hadron colliders, however, the momentum of the interacting partons remains unknown and other techniques have to be developed. In this chapter, four different methods to correct for the missing neutrino are presented:

- an average correction based on simulation, referred to as k-factor method;
- an empirical estimation derived by the E653 collaboration, refered to as E653 estimation;
- an analytical derivation of the neutrino momentum called neutrino closure method;
- an alternative analytical derivation called cone closure method.

The different methods are then compared quantitatively using a simulated sample of $D^{*+} \to D^0(\to K^-\mu^+\nu)\pi^+$ decays in terms of the difference between the momentum estimated by the correction method, p_{cor}, and true momentum of the D^0 candidate, p_{true}. The bias and resolution obtained with each correction method are estimated by the mean and root mean square (r.m.s.) values of the resulting $p_{cor} - p_{true}$ distributions.

4.1.1 k-*factor method*

The k-factor method is a well established technique to correct for the momentum of a missing particle on a statistical basis using Monte Carlo simulated data. The k factor is defined as the ratio between the momentum of the reconstructed daughter particles of a decay and the true momentum of the mother particle. In the present analysis this reads:

$$k = \frac{p(K\mu)}{p_{true}}, \tag{11}$$

where $p(K\mu)$ is the magnitude of the momentum of the combined kaon-muon system, referred to as visible D^0 momentum.

The visible momentum is correlated with the reconstructed invariant mass of the system $m(K\mu)$. The larger the fraction of total momentum carried by the neutrino, the smaller the values of $m(K\mu)$ and $p(K\mu)$ and hence the smaller the k factor. Figure 16 shows the two dimensional distribution of the k factor and the visible mass. Overlaid, a profile plot, showing the average value $\langle k \rangle$ in bins of the $K\mu$ mass, is fit with a simple second-order polynomial. The resulting fit function, $\langle k \rangle [m(K\mu)]$, can be used in data as an average event-by-event correction to determine the magnitude of the D^0 momentum:

$$p_{cor}(D^0) = \frac{p(K\mu)}{\langle k \rangle [m(K\mu)]}. \tag{12}$$

The spatial orientation of the momentum vector is then assumed to be aligned with the reconstructed D^0 flight direction $\hat{F}$. Since the D^{*+} decays instantaneously, this is inferred from the measured positions of the primary pp interaction vertex, $\vec{v}_{PV}$, and the D^0 decay vertex, $\vec{v}_{D^0}$, as $\hat{F} = (\vec{v}_{D^0} - \vec{v}_{PV})/|\vec{v}_{D^0} - \vec{v}_{PV}|$.

Figure 16 shows that the k-factor method reproduces the true D^0 momentum with a small bias of $-0.3\,\text{GeV/c}$ and a resolution of $14\,\text{GeV/c}$. The main disadvantage of this method is that, being based on simulation, it relies on an accurate description of the decay under study. For semileptonic decays this may lead to unwanted systematic uncertainties, as the form factors used to describe the decay amplitude are not always well known. Additionally, the accuracy of the k-factor correction may be limited by the size of the available simulated sample, which

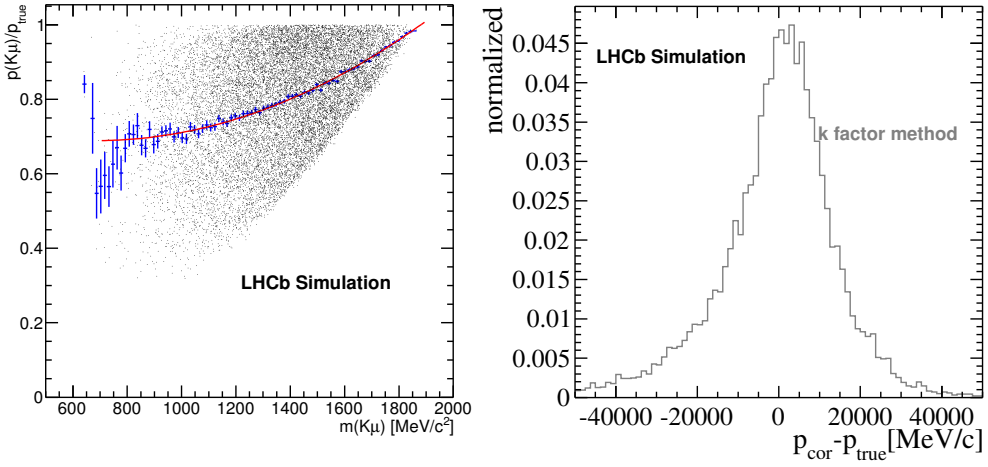


Figure 16: Left: (black points) two-dimensional distribution of the k factor and the visible mass $m(K\mu)$; (blue points) profile of the distribution, showing the average k-factor value in bins of $m(K\mu)$; (red curve) fit result to second-order polynomial. Right: distribution of the relative difference between k-factor-corrected and true D^0 momenta.

when studying charm decays it is generally much smaller then the corresponding data sample.

4.1.2 *E653 estimation*

The E653 collaboration derived the following empirical formula to estimate the corrected momentum, neglecting the component of the neutrino momentum parallel to the reconstructed D^0 flight distance [46]:

$$p_{cor}(D^0) = \frac{m_{D^0}\, E(K\mu)}{\sqrt{p_\perp^2(K\mu) + m^2(K\mu)}}, \tag{13}$$

where $E(K\mu)$ is the energy of the kaon-muon system and $p_\perp(K\mu)$ is the component of $\vec{p}(K\mu)$ perpendicular to the reconstructed D^0 flight direction $\hat{F}$. Due to momentum conservation, the magnitude of this component is equal to the perpendicular component of the neutrino momentum and therefore helps to partially account for the missing momentum. However, the component of the neutrino momentum parallel to the flight distance remains unknown.

The momentum resolution obtained by this empirical formula, as shown in Figure 17, is comparable in both size and shape to that of the k-factor method. A bias of 2 GeV/c and a resolution of 15 GeV/c is observed. The advantage of this method with respect to the k-factor is that the correction does not rely on simulation.

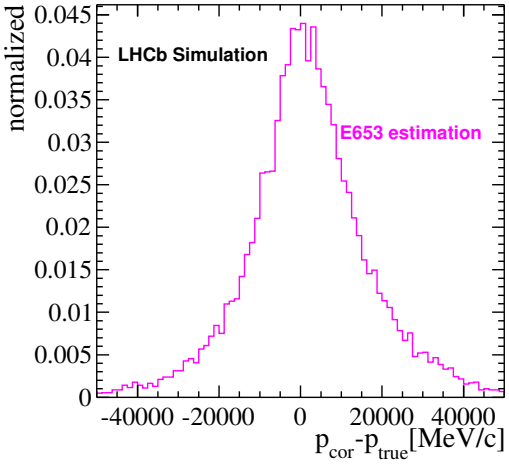


Figure 17: Distribution of torethe difference between the E653-corrected and the true D^0 momenta.

4.1.3 *Neutrino closure method*

In addition to require the D^0 momentum to be aligned with the reconstructed flight direction $\hat{F}$, the *neutrino closure* method requires that the three-body mass $m(K\mu\nu)$ has to coincide with the known D^0 mass. To explain how this constraint is used, it is convenient to express the momentum vectors in terms of their components parallel and perpendicular to $\hat{F}$:

$$p_{\parallel} = \vec{p} \cdot \hat{F}, \quad p_{\perp} = \left|\vec{p} - \vec{p}_{\parallel}\right| = \left|\vec{p} - (p_{\parallel} \cdot \hat{F})\right|.$$

From momentum conservation it follows that

$$p_{\perp}(K\mu) = -p_{\perp}(\nu) \quad \text{and} \quad p(D^0) = p_{\parallel}(K\mu) + p_{\parallel}(\nu).$$

Now using also energy conservation, $E(D^0) = E(K\mu) + E(\nu)$, and imposing the D^0 mass constraint such that $m^2_{D^0} = E^2(D^0) - p^2(D^0)$, it is possible to derive $p_{\parallel}(\nu)$ from a quadratic equation of the form

$$\alpha\, p^2_{\parallel}(\nu) + \beta\, p_{\parallel}(\nu) + \gamma = 0, \tag{14}$$

where the coefficients are defined by

$$\begin{aligned}
\alpha &= 4\, \left[p^2_{\perp}(K\mu) + m^2(K\mu)\right], \\
\beta &= 4\, p_{\parallel}(K\mu)\, \left[2\, p^2_{\perp}(K\mu) - m^2_{D^0} + m^2(K\mu)\right], \\
\gamma &= 4\, p^2_{\perp}(K\mu)\, \left[p^2_{\parallel}(K\mu) + m^2_{D^0}\right] - \left[m^2_{D^0} - m^2(K\mu)\right]^2.
\end{aligned}$$

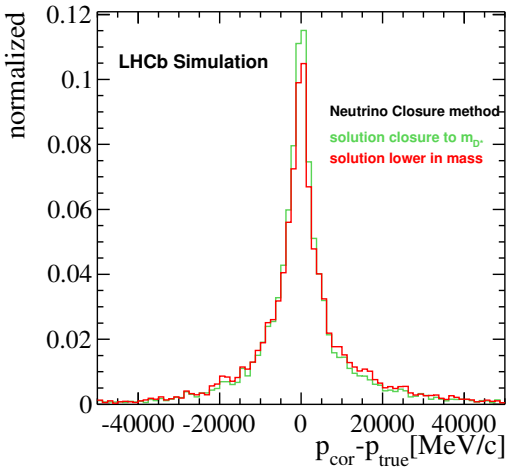


Figure 18: Distribution of the relative difference between the corrected and the true D^0 momenta resulting from the neutrino closure method with $p_{\parallel}(\nu)$ solution corresponding to $m(D^0\pi)$ closer to (green) m_{D^*} and (red) to threshold.

The neutrino's momentum parallel to the D^0 flight direction can then be calculated up to a two-fold ambiguity as

$$p_{\parallel}(\nu) = \frac{-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}. \tag{15}$$

Resolution effects however lead, in approximately 40% of the cases for the decay under study, to negative values of $\beta^2 - 4\alpha\gamma$, making Equation 14 unsolvable.

For $D^{*+} \to D^0(\to K^-\mu^+\nu)\pi^+$ decays, the two-fold ambiguity of Equation 15 can be resolved thanks to the presence of the parent D^{*+} meson. For these decays, the right solution for $p_{\parallel}(\nu)$ is indeed more likely to be that resulting in a neutrino-corrected $m(D^0\pi)$ mass closer to the known D^{*+} mass. From simulation, we see that only in 16.6% of the cases this does not happen because of resolution effects. As discussed in Section 4.2, this approach however induces a bias in the neutrino-corrected $m(D^0\pi)$ mass distribution of background events and is therefore discarded. To avoid this problem, one can select the $p_{\parallel}(\nu)$ solution resulting in neutrino-corrected $m(D^0\pi)$ closer to the kinematical threshold of $m_{D^0} + m_{\pi^+}$. In simulation we see indeed that, because of the low Q value of the $D^{*+} \to D^0\pi^+$ decay, this procedure allows to identify the right solution in approximately 75% of the cases with just a limited loss in momentum resolution with respect to choosing the solution closer to $m_{D^{*+}}$ (see Figure 18).

The *neutrino closure method* allows to reconstruct the D^0 momentum with much better resolution (12.6 GeV/c) then the k-factor or E653 estimations, but does not provide any correction in the $\sim 40\%$ of the cases where $\beta^2 - 4\alpha\gamma < 0$.

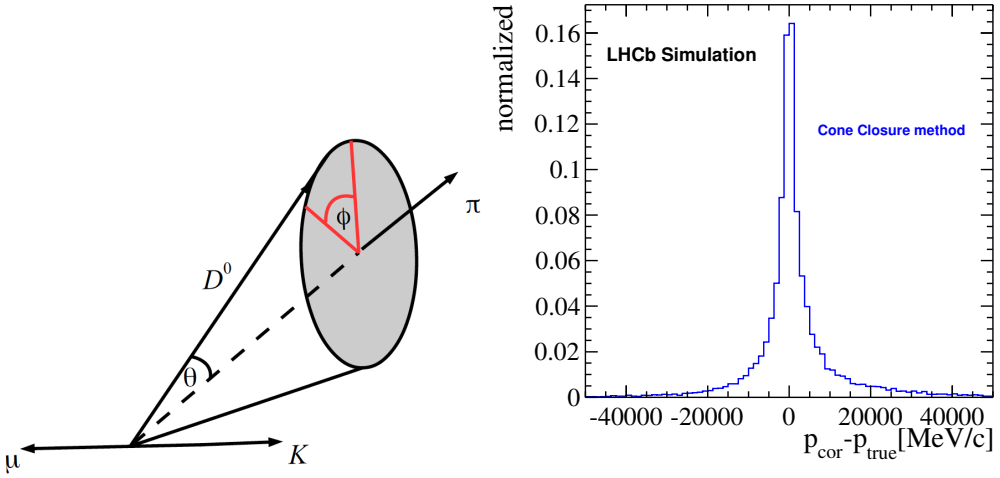


Figure 19: Left: schematic view of the $D^{*+} \to D^0(\to K^-\mu^+\nu)\pi^+$ decay in the $K\mu$ rest frame; the magnitude of the D^0 momentum is fixed, while its spatial orientation is restricted to be on a cone around the pion momentum. Right: distribution of the relative difference between the corrected and true D^0 momenta resulting from the cone closure method.

4.1.4 *Cone closure method*

The *cone closure* method [47] makes use of the same constraints of the neutrino closure method and, in addition, also requires the invariant mass of the D^0 π^+ system to be equivalent to the known mass of the D^*. To derive the corrected D^0 momentum with this method, it is convenient to boost the decay in the rest frame of the $K\mu$ system, so that

$$\vec{p}(K\mu) = 0, \quad \vec{p}(D^0) = \vec{p}(\nu) \quad \text{and} \quad E(D^0) = m(K\mu) + p(D^0).$$

Using the D^0 mass constraint allows, in this reference frame, to determine the magnitude of the D^0 momentum:

$$\begin{aligned} m^2_{D^0} &= E^2(D^0) - p^2(D^0) = m^2(K\mu) + 2\, m(K\mu)\, p(D^0) \\ \implies \quad p(D^0) &= \frac{m^2_{D^0} - m^2(K\mu)}{2\, m(Km\mu)}. \end{aligned} \tag{16}$$

Additionally, using the D^{*+} mass constraint restricts the D^0 momentum vector to lie on a cone with respect to the momentum of the pion (see Figure 19 for a visual representation), whose opening angle θ is determined as

$$\begin{aligned} m^2_{D^{*+}} &= \left[E(D^0) + E(\pi)\right]^2 - \left[\vec{p}(D^0) + \vec{p}(\pi)\right]^2 \\ &= m^2_{D^0} + m^2_{\pi^+} + 2\, E(D^0)\, E(\pi) - 2\, p(D^0)\, p(\pi) \cos\theta \\ \implies \quad \cos\theta &= \frac{2E(D^0)E(\pi) - (m^2_{D^{*+}} - m^2_{D^0} - m^2_{\pi^+})}{2p(D^0)p(\pi)}. \end{aligned} \tag{17}$$

The full D^0 momentum vector is then determined when choosing the value of the azimuthal angle ϕ (see Figure 19) that results in the best alignment between $\vec{p}(D^0)$ and the flight direction $\hat{F}$ in the laboratory frame. This is done numerically by scanning the angle ϕ between 0 and 360° in steps of one degree.

Also in this case, because of finite experimental resolution, in ~ 5% of the cases the method cannot be applied as Equation 17 results in $|\cos(\theta)| > 1$. When applicable, the cone closure shows the best momentum resolution among tested methods (10 GeV/c). In addition, it is worth mentioning that this approach, because of the D^{*+} mass constraint, cannot be used to determine the neutrino-corrected $m(D^0\pi)$ distribution.

4.2 EFFECT ON THE $m(D^0\pi)$ DISTRIBUTION

The expected signal of the WS sample is tiny. Hence, any improvement in $m(D^0\pi)$ resolution resulting from the neutrino reconstruction technique is precious. However, it is absolutely important to understand what is the effect of the employed correction on the background to avoid the creation of artificial structures in the $m(D^0\pi)$ distributions of the background that resembles the signal peak. The effect of the various neutrino reconstruction algorithms described in this chapter is then tested also on the sample of artificially created random-pion combinations, described in more details in Section 6.2, that reliably reproduces the kinematical features of the random-pion background.

Figure 20 shows the neutrino-corrected $D^0\pi^+$ mass distribution separately for simulated signal decays and random-pion background for the algorithms discussed. The effect on the signal is what expected as better mass resolutions are achieved by the algorithms providing better momentum resolutions. For the background, instead, most methods do not result in dangerous changes on the mass distribution. The only exception being the neutrino closure method with solution of $p_{\parallel}(\nu)$ yielding the mass closest to $m_{D^{*+}}$, where an artificial peak is created. On the other hand, taking the solution resulting in a lower neutrino-corrected mass also changes the mass shape (moving more events towards the threshold) but does not create artificial peaks which prevent any discrimination between signal and background candidates.

4.3 DECAY TIME RESOLUTION

The corrected D^0 momentum is used to calculate the corrected D^0 decay time with Equation 10. Figure 21 shows the decay timer resolution defined as $t_{reco} - t_{true}$ in units of the known D^0 lifetime [25] integrated over the full decay-time range. Resolution and bias of the reconstructed D^0 decay time of the presented correction algorithms itself are depending on the corrected decay. Figure 22 shows the two dimensional scatter plot of the difference $t_{reco} - t_{true}$ versus the corrrected decay

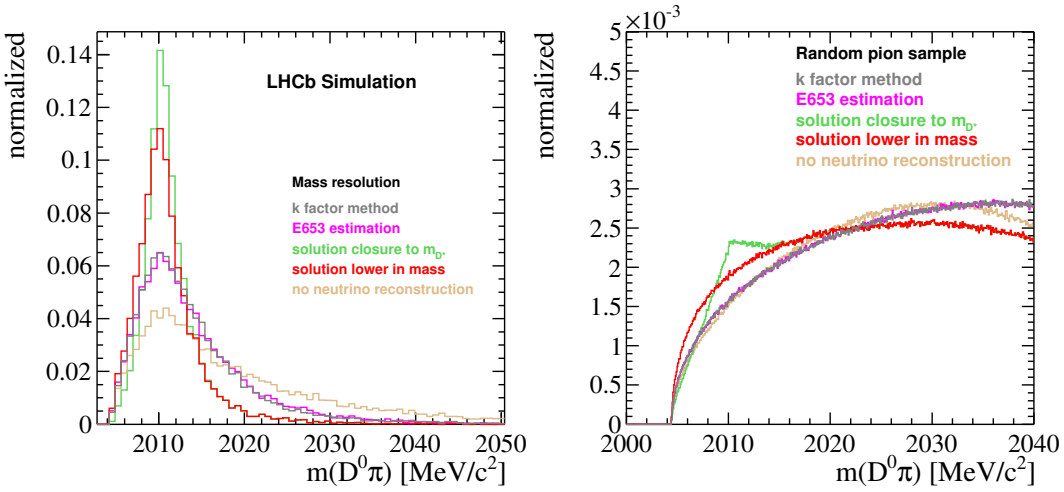


Figure 20: Neutrino-corrected $m(D^0\pi)$ distributions for (left) simulated signal decays and (right) random-pion combinations as resulting from the different algorithms discussed in this chapter.

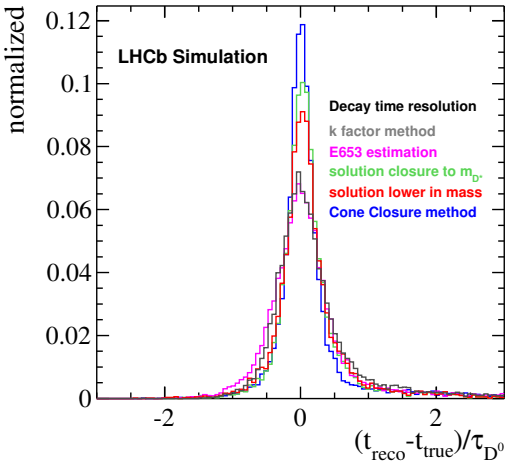


Figure 21: Decay time resolution $t_{cor} - t_{true}$ in units ot the D^0 lifetime as resulting from the different neutrino reconstruction algorithms discussed in this chapter.

time with the profile plot overlaid, showing the mean and r.m.s. for $t_{reco} - t_{true}$ in slices of the corrected decay time. The red lines indicate mean and r.m.s. integrated over the whole decay-time range. All reconstruction algorithms show inreasing bias and resolution at higher reconstructed lifetimes.

4.4 SUMMARY AND FINAL RECONSTRUCTION STRATEGY

Being able to reconstruct the missing neutrino momentum allows to determine the D^0 decay time and helps to reduce the background thanks to an improved $m(D^0\pi)$ resolution. The performances of the neutrino-reconstruction methods just described are compared in Table 2. Each method has advantages and disadvantages. Clearly, the cone closure technique, using all possible constraints from the

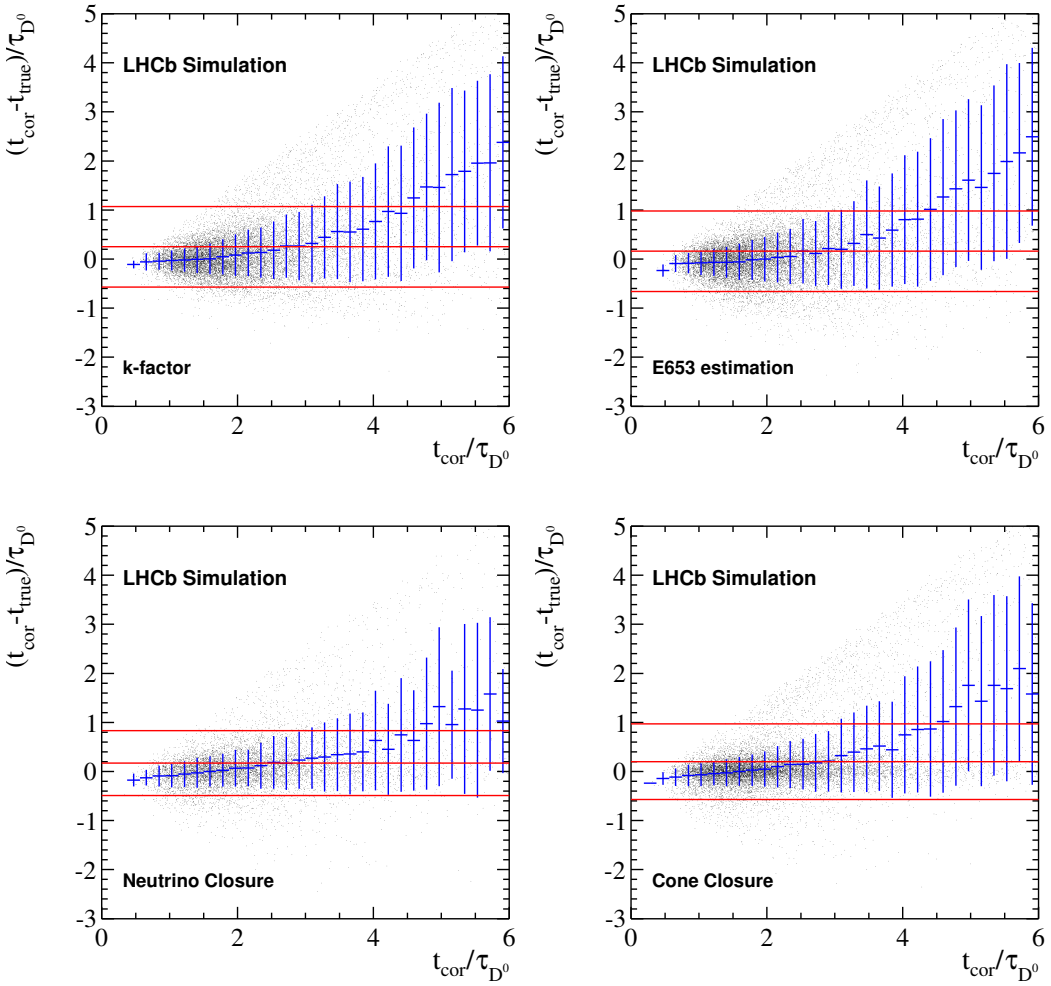


Figure 22: Time dependence of the reconstruction methods explained in the text. The red lines indicate the decay time bias and the resolution achievied by the algorithms integrated over the full decay time range.

decay topology/kinematics, allows for the best resolution and will therefore be used in the analysis to reconstruct the decay time. However, this method cannot be employed to correct the $D^0\pi^+$ mass. The mass will then be corrected using the neutrino closure method, taking the $p_{\parallel}(\nu)$ solution resulting in $m(D^0\pi)$ closer to threshold. Unless otherwise specified, all distributions showing $m(D^0\pi)$ will be corrected from now on. In all cases where the above two methods cannot be applied because of resolution effect, *i.e.* when $|\cos\theta| > 1$ for the cone closure and when $\beta^2 - 4\alpha\gamma < 0$ for the neutrino closure, the E653 estimation is instead used.

Method	$m(D^0\pi)$ width [MeV/c^2]	Decay-time resolution [τ_{D^0}]	Decay-time bias [τ_{D^0}]	Failure rate	Other comments
k factor	7.49	0.82	0.25	0%	Based on simulation
E653 estimation	7.49	0.82	0.16	0%	–
neutrino closure					
solution closest to $m_{D^{*+}}$	4.42	0.65	0.18	40%	Bias on background mass shape
solution closest to threshold	4.64	0.66	0.17	40%	–
cone closure	-	0.77	0.20	5%	Unusable to correct the mass

Table 2: Comparison of different neutrino-reconstruciton methods. Resolutions are determined as r.m.s. values of the corresponding corrected minus true quantity distribution. The failure rate represents the rate of cases where, because of resolution effects, the method cannot be applied. For the mass the r.m.s value of neutrino corrected $m(D^0\pi)$ distributions is quoted.

5

DATA SAMPLE SELECTION

This chapter describes the reconstruction and selection of $D^{*+} \to D^0(\to K^\pm\mu^\mp\nu)\pi^+$ *decays at both online and offline level. The selection aims at isolating signal candidates from expected backgrounds, mainly random pions, for which a multivariate analysis has been developed.*

5.1 ONLINE RECONSTRUCTION AND SELECTION

Events used in this analysis are triggered by signatures consistent with the $D^{*+} \to D^0(\to K^\pm\mu^\mp\nu)\pi^+$ decay.

At L0 the signature could be either an hadronic energy deposition with a transverse component of at least 3.5 GeV left by the charged kaon, or track segments in the muon stations compatible with a muon with $p_T \gtrsim 1.5\,\text{GeV/c}$.[1] In addition, at L0, also events triggered independently of the candidate $D^{*+} \to D^0(\to K^\pm\mu^\mp\nu)\pi^+$ decay are considered. It is possible that for example decay products of the charmed hardon formed by the other charm of the $c\bar{c}$ pair produced by the pp interaction triggered the event.

In HLT1, where tracking information becomes available, the kaon or the muon is required to have large momentum, $p \gtrsim 3\,\text{GeV/c}$ and $p_T \gtrsim 1.6\,\text{GeV/c}$, and to be incompatible with being produced in the primary vertex, IP> 0.1 mm and $\chi^2(\text{IP}) > 16$. The $\chi^2(\text{IP})$ is defined as the difference between the χ^2 obtained when fitting the primary vertex with and without the considered particle, and is a measure of consistency with the hypothesis that the particle originates from the primary vertex. Events triggered at L0 by the muon can also be selected at HLT1 if a track with $p_T \gtrsim 1\,\text{GeV/c}$ is matched to hits in the muon stations, where the reduced p_T threshold is possible due to a lower rate.

In HLT2, a full event reconstruction is performed and the trigger decision aims at selecting the specific final state of interest. For semileptonic neutral charm decays, two trigger lines are available. The first employs an exclusive selection of the $D^0 \to K^\pm\mu^\mp\nu$ decay, solely based on the presence of two displaced oppositely charged tracks fulfilling basic momentum and quality requirements. The two tracks have to form a good-quality decay vertex with flight distance $F > 4$ mm and $\vec{F}\cdot\hat{z} > 0$ mm to ensure that the D^0 candidate is flying in forward direction. As the D^0 candidate is only partially reconstructed, only very loose requirements are imposed on the visible and corrected D^0 masses: $m(K\mu) < 2500\,\text{MeV/c}^2$ and

1 The thresholds changed during data taking period. Here, only the values from the second half of 2012 are mentioned.

$1400\,\mathrm{MeV/c^2} < m_{cor} < 2700\,\mathrm{MeV/c^2}$. A variant of this trigger allows to reconstruct also candidates where the D^0 daughters have the same charge. The second HLT2 trigger line is designed to select $D^{*+} \to D^0\pi^+$ decays where the soft pion has momentum larger 3 GeV/c and $p_T > 300\,\mathrm{MeV/c}$ and the D^0 decays inclusively in any final state with at least two charged tracks having $\chi^2(\mathrm{IP}) > 36$. The soft pion is also required to come from the primary vertex with $\chi^2(\mathrm{IP}) < 9$. This selection criterion suppresses randomly associated pions and secondary D^{*+} decays. Compared to the exclusive line, here no cut on the corrected mass is applied to account for the missing neutrino, but the candidates are selected only by $\Delta m < 285\,\mathrm{MeV/c^2}$ to inclusively detect D^0 decays to multi-body final states. This trigger was implemented in 2012 and therefore does not contribute to the sample collected during 2011.

5.2 OFFLINE RECONSTRUCTION AND SELECTION

At LHCb, the event reconstruction from the collected data is done centrally and only events passing a pre-selection (called in jargon *stripping*) are processed in this analysis. The stripping reconstructs the candidates of interest by applying or tightening, with offline-reconstructed quantities, the same selection requirements used in the trigger. In addition, as particle-identification information is not available in the trigger, requirements on the kaon, muon and pion identification are now implemented. The stripping-selection criteria are based on previous studies of semileptonic charm decays at LHCb [48] and are summarized, together with the other requirements used in this analysis, in Table 3.

The resulting sample can be seen in Figure 23. Figure 23a shows the $m(D^0\pi)$ distribution for RS candidates. The peaking structure close to the kinematic threshold, with maximum at the nominal D^* mass value, is mostly due to RS $D^0 \to K^-\mu^+\nu$ signal candidates, but also includes $D^{*+} \to D^0\pi^+$ candidates where the D^0 decays to a different final state. In the corresponding distribution for WS candidates, shown in Figure 23b, no clear peaking structure can be distinguished from the much broader distribution due to random-pions background. The spectra of the visible D^0 mass for RS and WS candidates can be seen in Figure 23c and Figure 23d, respectively. The narrow peak close to the known D^0 mass value is due fully reconstructed $D^0 \to K^-\pi^+$ decays, where the charged pion is misidentified as a muon. For partially reconstructed candidates from semileptonic D^0 decays, the visible mass is depending on the fraction of momentum carried by the neutrino, thus resulting in a very broad distribution peaking at lower mass values. The fact that RS and WS samples have very similar $m(K\mu)$ distributions indicates that the composition of latter is dominated by background consisting of real semileptonic D^0 decays accidentally associated to uncorrelated soft pions.

Figure 24 compares the $m(D^0\pi)$ distribution of the RS sample, normalized to the same area, for different trigger categories for the data set collected during

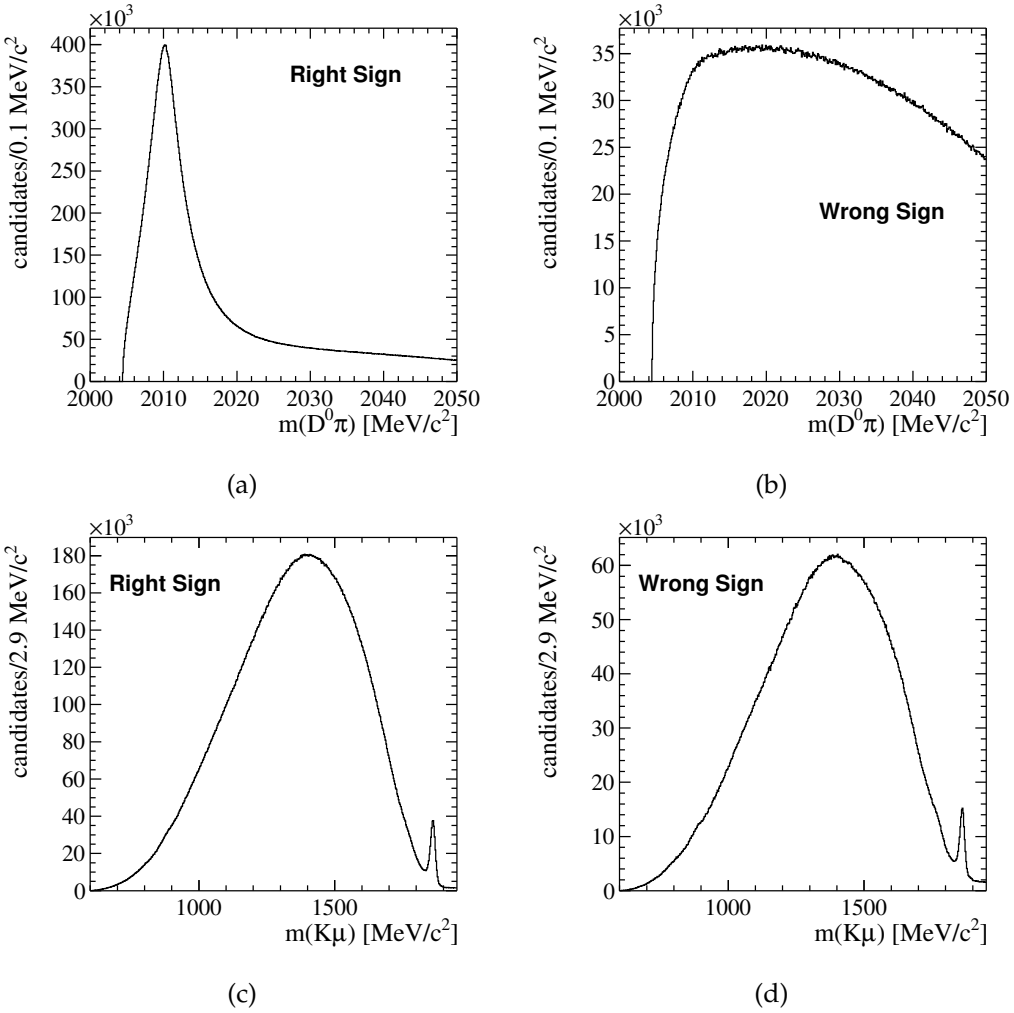
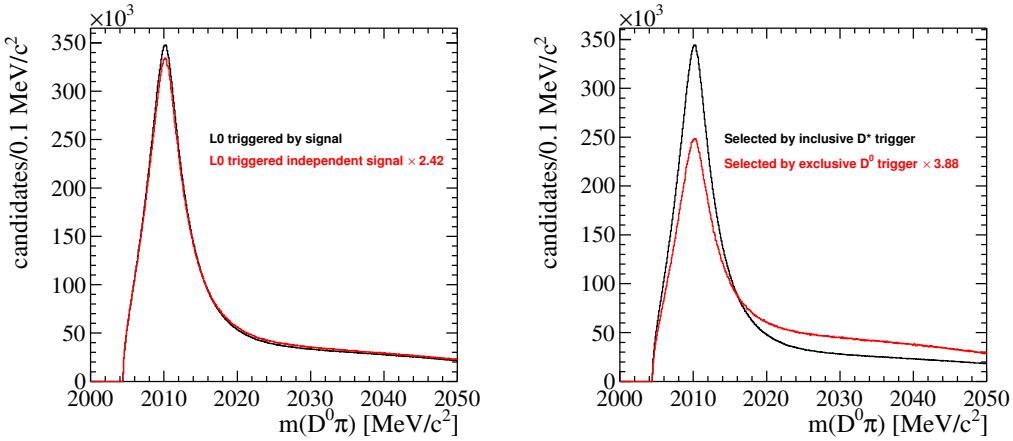


Figure 23: Spectra in $m(D^0\pi)$ and $m(K\mu)$ after the stripping selection for RS and WS candidates. The narrow peak in $m(K\mu)$ close to the D^0 mass is due to misidentified two body decays (see Section 5.2.2).

2012. Figure 24a compares events triggered by signal candidates at L0 and those triggered independently of the signal. The two samples consists of approximately 34×10^6 and 14×10^6 events, respectively, and have similar signal purities. The overlap between the two samples is 10×10^6. Figure 24b compares instead candidates selected by the inclusive and the exclusive HLT2 triggers. The inclusive sample consists of approximately 32×10^6 events, all collected during 2012. The exclusive trigger collected instead approximately 13×10^6 events, 36% of which during 2011. The exclusively triggered sample, as clearly visible in the plot, provides a larger background contamination and adds only a limited amount of signal (when the overlap between the samples is considered). For these reasons, and to simplify the treatment of the data, the analysis is limited only to candidates selected by the inclusive HLT2 trigger during 2012.

Specific selections to suppress the relevant background described in Section 3.2 have been developed for this analysis, as described in the following.



(a) Comparison of candidates triggered by and indepedent of signal candidates at L0.

(b) Comparison of candidates selected by the inclusive D^0 and the exclusive D^* HLT2 trigger.

Figure 24: Comparison of the different trigger categories. The distributions are normalized to the same area. Scaling factors are given in the legend.

5.2.1 *Suppression of random-pion background*

To reduce the amount of purely combinatorial candidates and of those where a soft pion has been associated accidentally to a correctly reconstructed D^0, a multivariate analysis is performed to combine information of several discriminating variables into a single observable. The available information provided by each variable, and by the correlations between the different variables, is this way used most efficiently allowing for a better signal-to-background discriminating power.

Several methods have been tested from those available in the *Toolkit for Multivariate Analysis* (TMVA) [49]. Among those, the *boosted decision tree* (BDT) [50] provides the best performances and is therefore chosen for the analysis. More details on how a BDT works are provided in Appendix A.

The BDT is trained directly on approximately 10% of the RS data sample, where signal and random-pion background are disentangled using the ${}_s\mathcal{P}lot$ technique [51]. The ${}_s\mathcal{P}lot$ is a statistical tool that allows to separate signal and background contributions in various *control variables* from the knowledge of the distributions of a *discriminating variable*, in this case the $m(D^0\pi)$. A weight is calculated on an event-by-event basis, so that the weighted distributions in the control variables reproduce on average the corresponding true signal and background components. The signal and background ($i = \text{sig, bkg}$) weights are calculated as [51]

$$ {}_s\mathcal{P}_i(m) = \frac{V_{i,sig}\ \mathcal{P}_{sig}(m) + V_{i,bkg}\ \mathcal{P}_{bkg}(m)}{N_{sig}\ \mathcal{P}_{sig}(m) + N_{bkg}\ \mathcal{P}_{bkg}(m)}, \tag{18} $$

where N_{sig} (N_{bkg}) is the total number of signal (background) candidates, $\mathcal{P}_{sig}$ ($\mathcal{P}_{bkg}$) is the $m(D^0\pi)$ probability distribution function (PDF) for signal (background) and V_{ij} are the elements of the covariance matrix relating N_{sig} and N_{bkg}. These param-

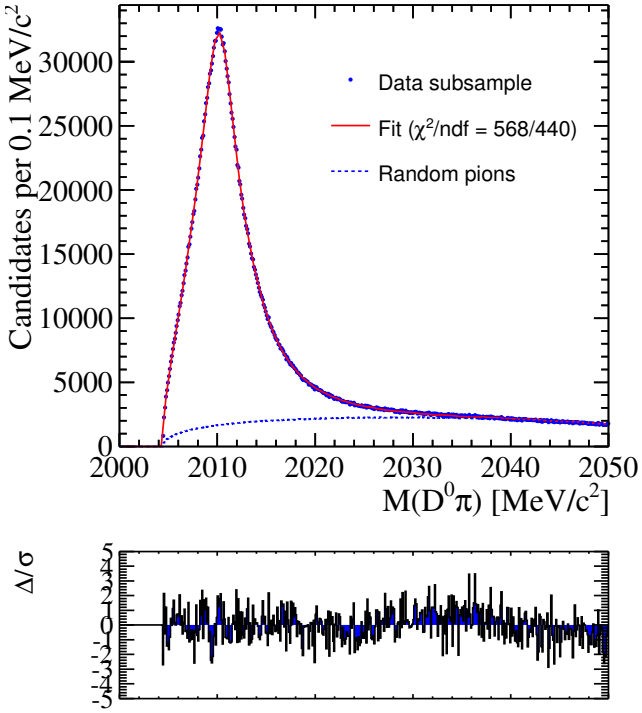


Figure 25: $m(D^0\pi)$ distribution for the RS subsample used in the BDT training, with fit projections overlaid. The fit results are used to calculate the weights of Equation 18.

eters are all estimated from a minimum χ^2 fit to the $m(D^0\pi)$ distribution of the RS sample used in the training. Figure 25 shows the result of the fit, whose details can be found in Chapter 6.

Input variables for the BDT are chosen to maximize the statistical separation between signal and background, and to avoid correlations between the BDT output and the mass variables $m(D^0\pi)$ and $m(K\mu)$, later used to extract the signal yield (see Chapter 6). The chosen set of input variables is shown in Figure 26 and listed below:

- Log $\text{ProbNN}_{\text{ghost}}(\pi)$, logarithm of the probability for the soft pion to be a ghost track ($\delta = 1.2\%$);
- $\sqrt{\text{DOCA}(K\mu, \pi)}$, square root of the distance of closest approach between the $K\mu$ pair and soft pion flight directions ($\delta = 0.5\%$);
- $\sqrt{\text{DOCA}(K, \mu)}$, square root of the distance of closest approach between the kaon and muon flight directions ($\delta = 1.2\%$);
- D^0 flight distance ($\delta = 0.6\%$);
- D^0 transverse momentum ($\delta = 5.2\%$);
- D^0 pseudorapidity ($\delta = 0.4\%$);

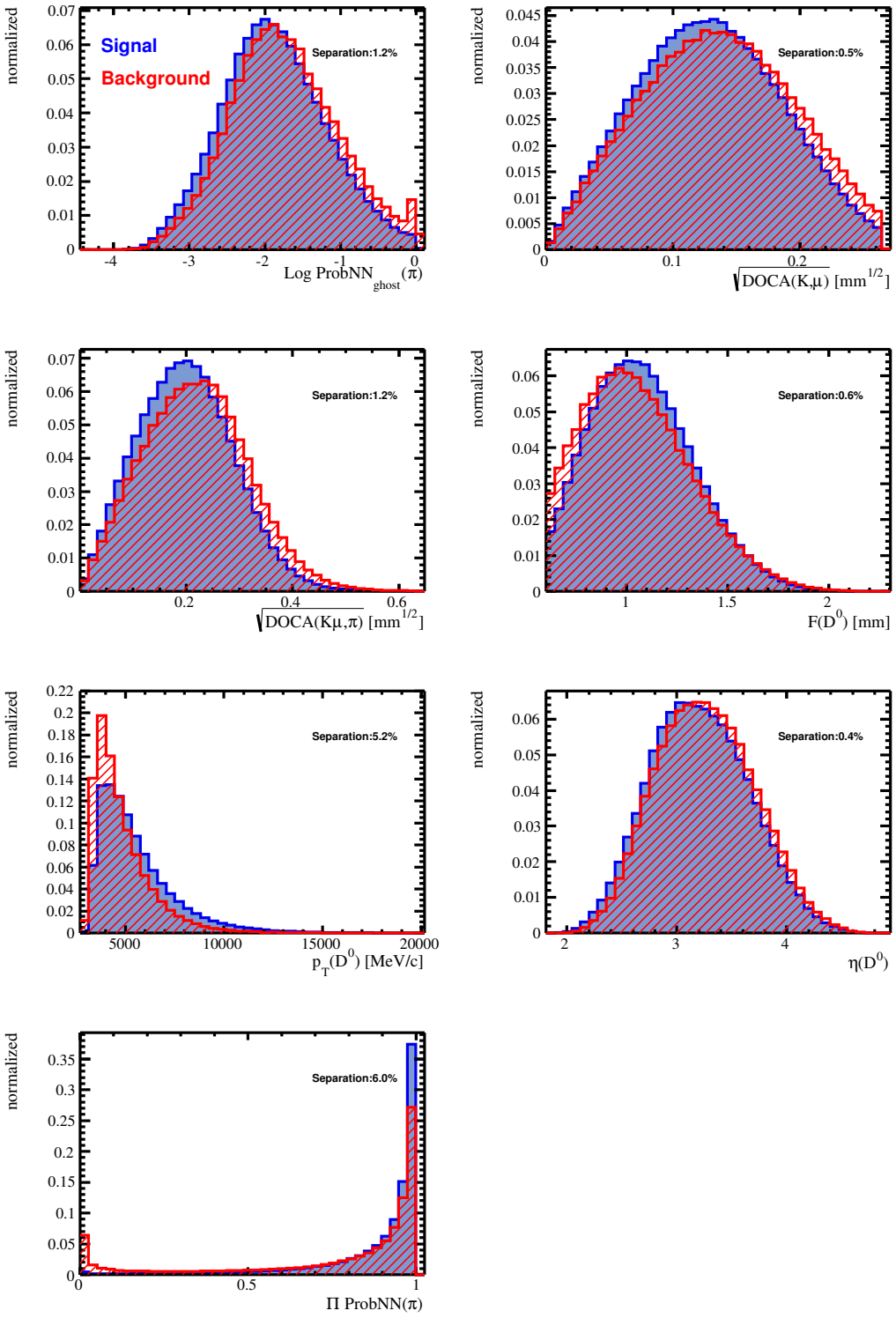


Figure 26: Signal (blue) and background (red) distributions of the variables used in the BDT training.

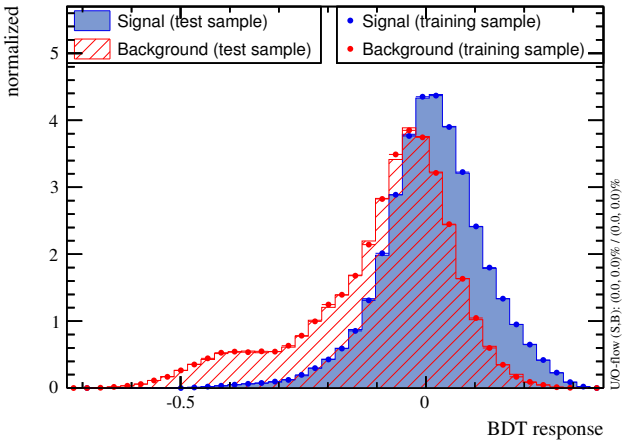


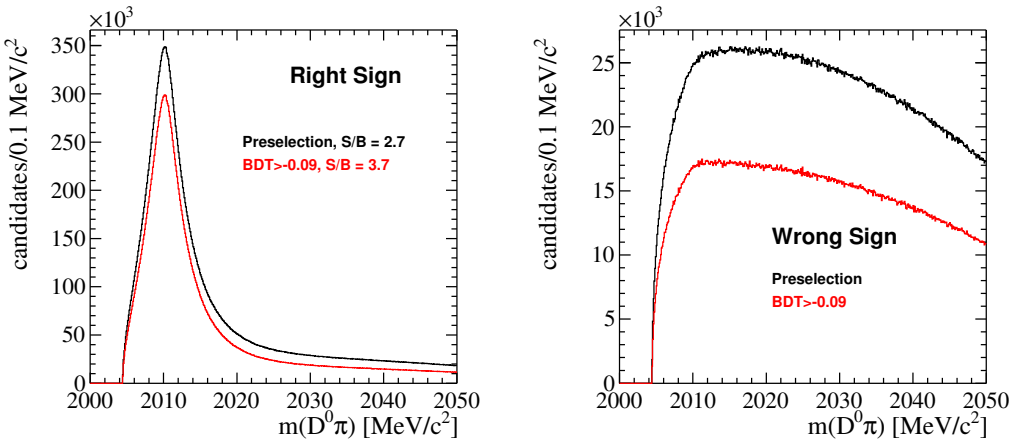
Figure 27: Signal (blue) and backgorund (red) distributions of the BDT response.

- $\prod \text{ProbNN}(\pi) = \text{ProbNN}_\pi(\pi)\ \left[1 - \text{ProbNN}_\mu(\pi)\right]\ \left[1 - \text{ProbNN}_K(\pi)\right]$ ($\delta = 6.0\%$).

Here, the quantity δ indicates the statistical separation provided by each variable, defined as

$$\delta = \frac{1}{2}\int \frac{(\mathcal{P}_{sig}(y) - \mathcal{P}_{bkg}(y))^2}{\mathcal{P}_{sig}(y) + \mathcal{P}_{bkg}(y)}dy$$

with the signal ($\mathcal{P}_{sig}(y)$) and background ($\mathcal{P}_{bkg}(y)$) PDF for variable y [49]. For some variables, simple transformations such as the square root or the logarithm of the original variable provide better signal-to-background separation and are therefore used.



(a) $m(D^0\pi)$ distribution for the RS sample, before and after the BDT selection

(b) $m(D^0\pi)$ distribution for the WS sample, before and after the BDT selection.

Figure 28: Application of the BDT. Preselection refers to stripping and trigger selection.

Figure 27 shows the obtained BDT response separately for the training sample and for an independent testing sample. Large differences in the BDT response for

training the testing sample arise if the BDT is trained on statistical fluctuations; here, the distributions look very similar and no indication for training on fluctuations is visible. The achieved signal-to-background separation is $\delta = 13.6\%$. An optimal BDT selection requirement of BDT > -0.09 is chosen, corresponding to the maximum for the figure of merit[2] [52]

$$S = \frac{N_{sig}}{1.5 + \sqrt{N_{bkg}}}.$$

Figure 28 shows the $m(D^0\pi)$ distribution of the RS sample before and after the BDT selection. The number of RS signal candidates decreases by approximately 12% from $\approx 18.6 \times 10^6$ to $\approx 16.4 \times 10^6$, while the background is reduced by 35% from $\approx 6.8 \cdot 10^6$ to 4.4×10^6, improving the signal-to-background ratio from 2.7 to 3.6.

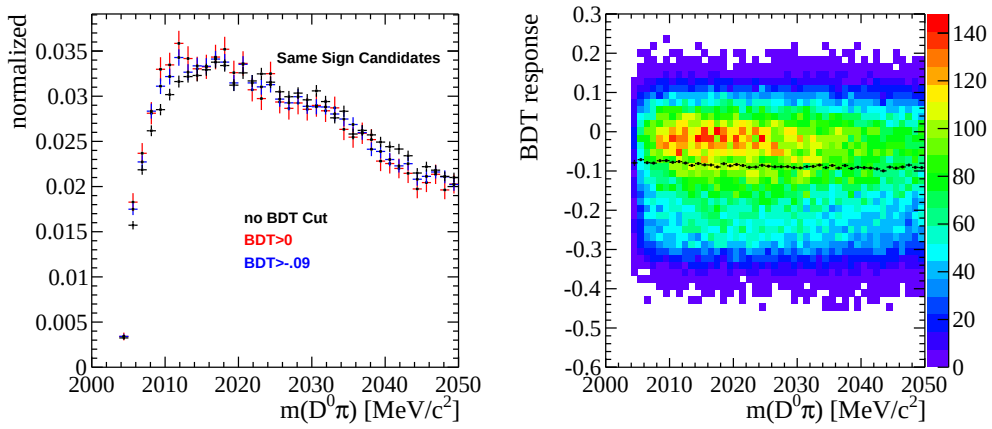


Figure 29: Testing the effect of BDT cut on $m(D^0\pi)$ with candidates, where D^0 daughters and the soft pion have all the same charge. Left: $m(D^0\pi)$ distribution applying different BDT cuts. Richt: two dimensional plane of mass and BDT response.

The effect of the BDT selection on the $m(D^0\pi)$ distribution is checked using candidates where the D^0 daughters and the soft pion have all the same charge. As this selection is only available for events collected by the exclusive trigger, the BDT was retrained on such sample before performing the check. Figure 29 shows the comparison between the D^0 π^+ mass distibutions obtained with and without the BDT selection requirement, as well as the two-dimensional distribution of $m(D^0\pi)$ versus BDT. The plots confirm that the BDT response is independent of $m(D^0\pi)$ and therefore no artificially peaking structures, which may fake the signal, are created by the BDT selection.

2 The form of this f.o.m is: $S = \frac{N_{sig}}{a/2+\sqrt{N_{bkg}}}$, with a being the significance aimed to optimize for, chosen to be $a = 3$ in this case.

5.2.2 *Suppression of misidentified decays*

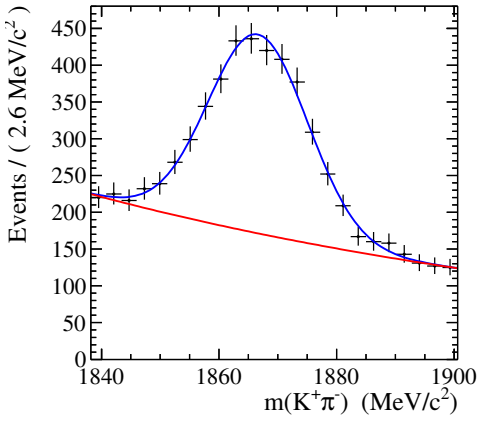
The muon track is required to match a segment of hits in the muon stations by the *IsMuon* algorithm [53], while both the kaon and the soft pion are not. Tight requirements on the muon identification variables, $\mathrm{ProbNN}_\mu(\mu) > 0.6$, are then added to further suppress hadronic D^0 decays where one of the final state hadron is misidentified as a muon. Misidentified kaons are reduced by requiring $\mathrm{ProbNN}_K(K) > 0.6$. These requirements, however, do not completely suppress the background from misidentified hadronic decays.

As an example, Figure 30a and Figure 30b show the distribution of $m(K^+\pi^-)$ and $m(\pi^+K^-)$, *i.e.* the invariant mass computed when assigning the pion hypothesis to the muon and both assigning pion mass to the kaon and kaon mass hypothesis to the muon, respectively, for the WS sample with fit projection overlaid. In Figure 30a, a peak from $D^0 \to K^+\pi^-$ decays, where the pion was misidentified as a muon, is clearly visible. Figure 30b shows the peak from doubly misidentified $D^0 \to K^-\pi^+$ decays. Figure 30c and Figure 30d show the number of candidates in the peaks as a function of the $\mathrm{ProbNN}_\mu(\mu)$ and $\mathrm{ProbNN}_K(K)$ cut value, in comparison with the corresponding number of expected WS signal candidates. The latter is estimated by scaling the RS signal yiel by the mixing rate $x^2 + y^2 \approx 10^{-4}$. The plots make evident that misidentified backgrounds from hadronic D^0 decays to multi-body final states, such as $D^0 \to K^+\pi^-\pi^0$, can be certainly reduced imposing tighter particle identification requirements but never to a negligible level. They are therefore accounted for when measuring the signal yield, as described in Chapter 6. On the contrary, two-body hadronic decays are vetoed by excluding candidates with the relevant two-body mass within $24\,\mathrm{MeV}/c^2$ (corresponding to approximately 3σ in mass resolution) of m_{D^0}.

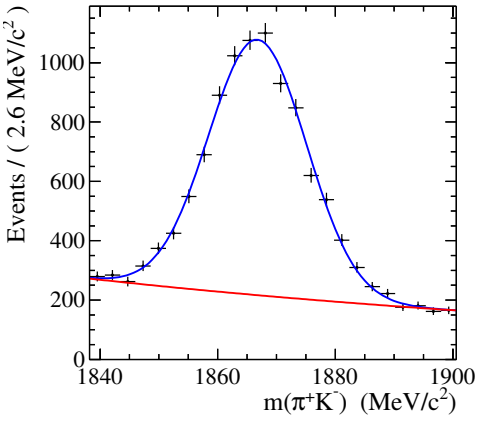
5.2.3 *Multiple candidates removal*

More than one RS or WS D^{*+} candidate can be reconstructed in the same event. Usually, this happens because the same D^0 candidate can be easily associated to different soft pions produced in the primary vertex. Multiple candidates are removed from the selected sample, by randomly choosing only one candidate per event, to avoid any potential bias they may create on the result. This removes approximately 3.0% of both RS and WS candidates from the sample.

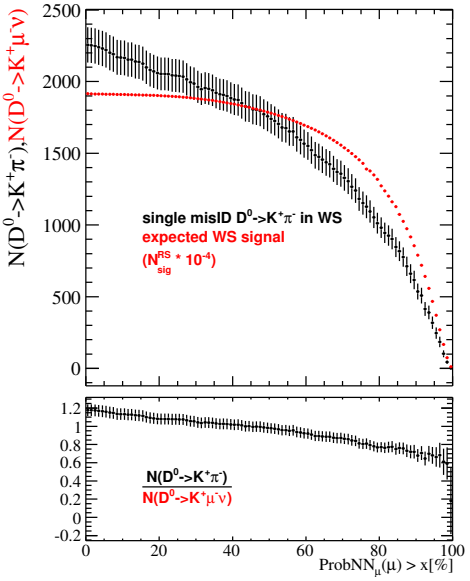
In addition, it is possible to form both a RS and a WS D^{*+} candidate by associating soft pions of opposite charge to the same D^0 candidate. Since it is much more probable for such D^0 candidates to be real RS than real WS D^{*+} decays, they are identified and removed from the WS sample. This procedure removes $\sim 9.5\%$ of the WS random-pion background, with negligible loss in the expected signal.



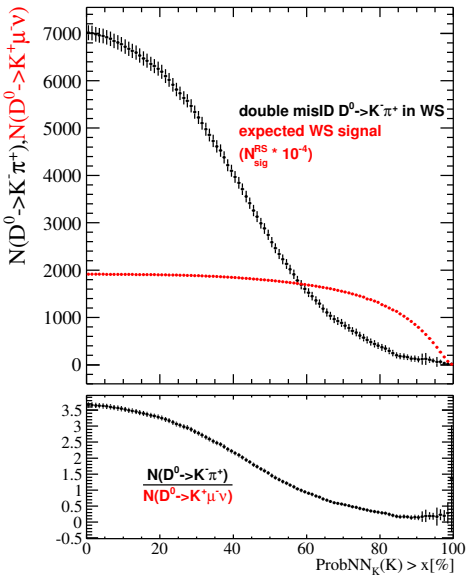
(a) Peak in the visible mass when pion mass hypothesis is assigned to the muon candidate, resulting from singly misidentified $D^0 \to K^+\pi^-$ decays .



(b) Visible mass when pion mass hypothesis is assigned to the kaon, and kaon mass hypothesis to the muon candidate, respectively. The peak is resulting from doubly misidentified $D^0 \to \pi^+K^-$ decays



(c) Singly misidentified candidates as function of $\text{ProbNN}_\mu(\mu)$ cut value, compared to expexted WS signal.

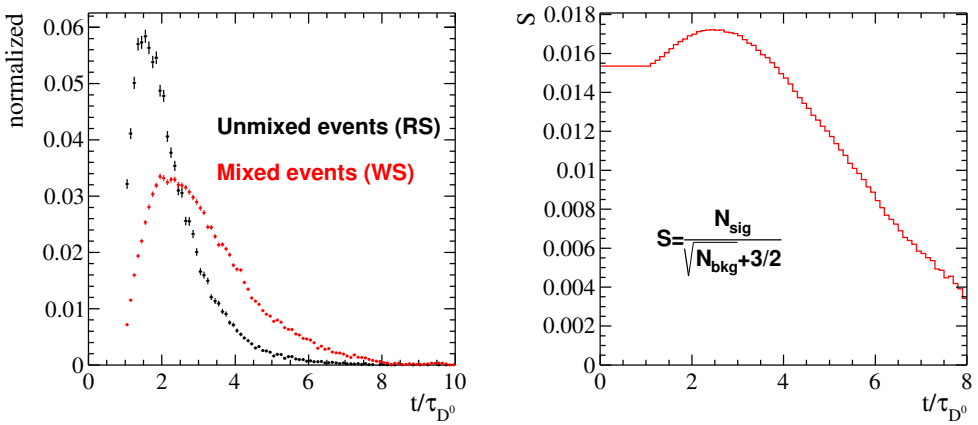


(d) Doubly misidentified candidates as function of $\text{ProbNN}_K(K)$ cut value, again compared to expexted WS signal.

Figure 30: Study of singly and doubly misidentified D^0 two body decays.

5.2.4 *Optimal decay-time selection*

As already explained in Chapter 1, a biased decay-time selection helps to increase the sensitivity to mixing because the WS-to-RS ratio increase quadratically with decay time (see Equation 7). Moreover, a selection cut on the D^0 decay time can also help to reduce the random-pion background because, consisting mainly of unmixed D^0 candidates associated to uncorrelated pions, this background features a different decay-time distribution with respect to the WS signal.



(a) Mixed (WS) and unmixed (RS) decay-time distribution as estimated from simulation.

(b) Significance of the WS signal as a function of the minimum decay-time threshold.

Figure 31: Optimizing the decay time cut.

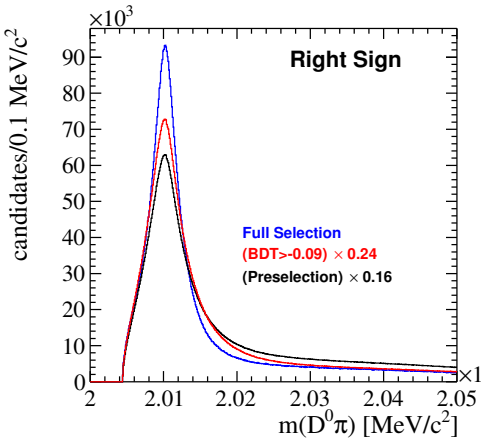
Figure 31a shows the expected decay-time distribution of the mixed (WS) and unmixed (RS) signals, normalized to the same area, as derived from simulation studies. Scaling the distributions to the expected number of WS signal and random-pion background events allows to compute the significance as a function of the minimum decay-time cut, as shown in Figure 31b. The optimal value, corresponding to the maximum of the curve, is $t > 2.3\tau_{D^0}$. At this point, the average of the squared decay time $\langle t^2 \rangle$ as introduced in Section 1.3.1 is estimated by using the sample of simulated signal candidates to be $\langle t^2 \rangle \sim 11.11\,\tau^2_{D^0}$ after the decay time cut.

In addition, candidates with decay times in excess of $20\,\tau_{D^0}$ are removed from the sample, as they are expected to be dominated by badly reconstructed events rather than by signal.

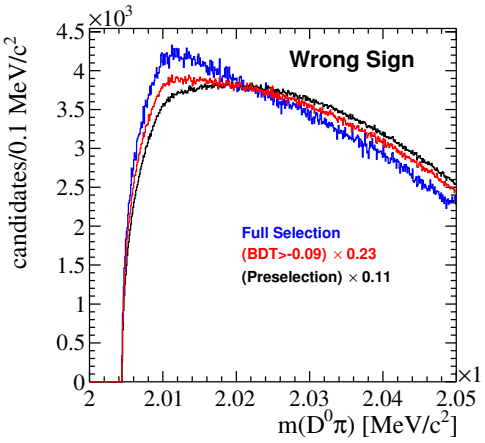
The final set of selection criteria are listed in Table 3. Figure 32 shows the $m(D^0\pi)$ and $m(K\mu)$ distributions for RS and WS candidates after the different steps of the selection.

Candidate	Variable	HLT2 selection	Stripping	Offline selection	Unit
Muon	p_T	> 800			MeV/c
	χ^2_{track}	> 2.25			
	$\chi^2(IP)$	> 36			
	$DLL_{\mu-\pi}$		> 0		
	IsMuon		true		
	$ProbNN_{\mu}$			> 0.6	
	$ProbNN_{ghost}$			> 0	
Kaon	p_T	> 600			MeV/c
	χ^2_{track}	> 2.25			
	$\chi^2(IP)$	> 36			
	$DLL_{K-\pi}$		> 0		
	IsMuon		false		
	$ProbNN_{k}$			> 0.6	
	$ProbNN_{ghost}$			> 0	
D^0	DIRA	> 0.99			
	χ^2_{FD}	> 100			
	$m(K\mu)$	< 2500		$< m_{D0}$	MeV/c^2
	χ^2_{vertex}/ndf	< 10	< 9		
	p		> 20000		MeV/c
	p_T		> 2800		MeV/c
	$DOCA(K\mu)$		< 0.07		mm
	m_{cor}		$[1400, 2700]$		MeV/c^2
	t_{cor}/τ_{D^0}			$[2.3, 20]$	
	$\lvert m(KK) - m_{D0}\rvert$			> 24	MeV/c^2
	$\lvert m(K\pi) - m_{D0}\rvert$			> 24	MeV/c^2
	$\lvert m(\pi K) - m_{D0}\rvert$			> 24	MeV/c^2
	F		> 4		mm
	F_Z		> 0		mm
	F_{XY}			< 4	mm
Pion	p	> 3000			MeV/c
	p_T	> 300			MeV/c
	$\chi^2(IP)$	< 9			
	$DLL_{\mu-\pi}$		< 1		
	IsMuon		false		
	$ProbNN_{ghost}$			> 0	
D^*	p_T	> 3750			MeV/c
	$DOCA(D^0\pi_s)$	< 100	< 0.4		mm
	Δm	$[0, 285]$	< 200		MeV/c^2
	$m(D^0\pi)$			< 2050	MeV/c^2
	$m(K\mu\pi)$		> 300		MeV/c^2
	χ^2_{vertex}/ndf		< 9		

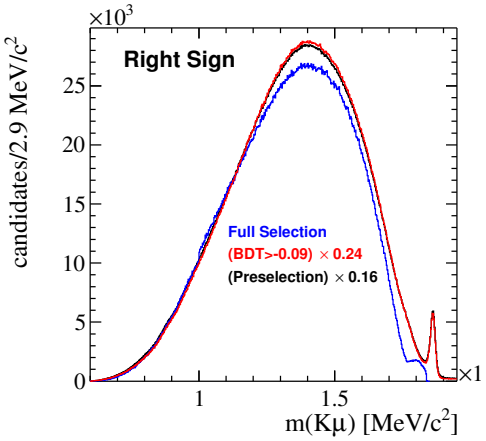
Table 3: List of all selection criteria used in the analysis, separately for HLT2 trigger, preselection (*stripping*) and final offline selection. If not explicitly mentioned, the trigger cuts have been adopted also in the stripping.



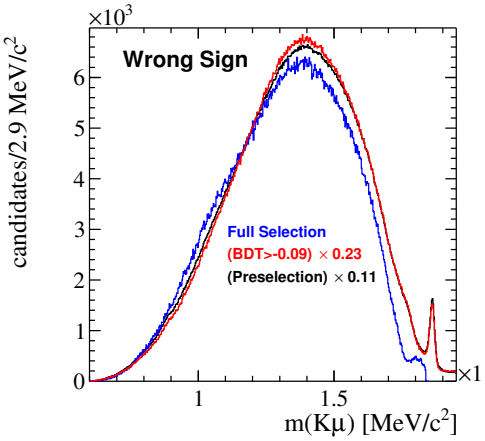
(a) Right sign $m(D^0\pi)$ distribution after different steps of the selection.



(b) Wrong sign $m(D^0\pi)$ distribution after different steps of the selection.



(c) Right sign $m(K\mu)$ distribution after different steps of the selection.



(d) Wrong sign $m(K\mu)$ distribution after different steps of the selection.

Figure 32: The spectra are normalized to the area of the distributions after the full selection. Scaling factors are given in the legends.

6

FIT OF SAMPLE COMPOSITION

This chapters describes the strategy developped to extract the signal yields on a statistical basis. First, the general strategy is explained and then applied to the RS sample. Second, a preliminary fit to the WS sample is shown.

6.1 GENERAL STRATEGY

Although the selection of Chapter 5 is designed to minimize the background contamination in the sample, the remaining events are a mixture of unknown composition and the amount of signal events has to be identified statistically. No single variable is found to be sufficient to separate the signal candidates from all different kinds of backgrounds discussed in Section 3.2 at once. However, the two dimensional space spanned by the invariant masses $m(D^0\pi)$ and $m(K\mu)$ shows promising separation to isolate the signal as explained in the following.

Figure 33 schematically shows the expected shapes for signal, misidentified and multi-body D^0 decays, random pions and purely combinatorial background. Random pions and combinatorial background can easily be identified in $m(D^0\pi)$, because they feature a different non-peaking shape compared to signal and misidentified multi-body decays. Having both similar peaking distributions in $m(D^0\pi)$, misidentified multi-body decays and signal candidates can instead be separated in $m(K\mu)$, where they are expected to have different shapes as the amount of missing momentum does depend on the number and species of particles not reconstructed. Combining the information of $m(D^0\pi)$ and $m(K\mu)$ in a bidimensional fit would allow to disentangle the individual components. While this is the preferable approach in terms of statistical precision, it would require to know the two-dimensional distribution of each component, as correlations between $m(D^0\pi)$ and

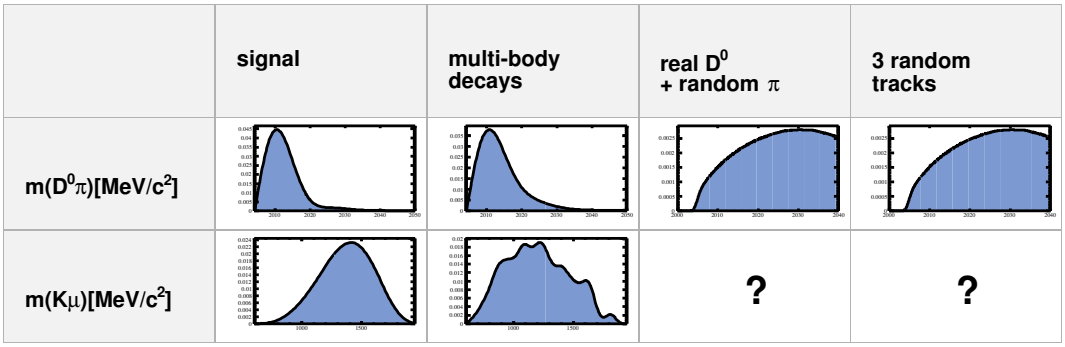


Figure 33: Schematic mass shapes for signal and backgound components in $m(D^0\pi)$ and $m(K\mu)$.

$m(K\mu)$ do not allow to factorize the joint PDFs. This makes the bidimensional fit difficult and subject to unwanted systematic uncertainties, as for some of the background component, such as the purely combinatorial background, the two-dimensional distributions are not known and cannot be simply extracted from simulation or control samples. Also, to reliably extract the two-dimensional distribution of the physics backgrounds from the simulation does requires huge samples of simulated events, which, as discussed later, are not always available.

To overcome these limitations, in this analysis a different fitting strategy is employed. This consists of fitting the sample in two successive steps. First, the data sample is divided in bins of $m(K\mu)$. In each bin, a one-dimensional fit to $m(D^0\pi)$ is performed to get the number of peaking events. This number includes signal and misidentified multi-body D^0 decays, while non-peaking random pions and combinatorial background are statistically subtracted. Plotting these yields in the bins of the visible mass reveals the $m(K\mu)$ projection for signal and multi-body decays only. Next, signal decays are separated from background due to misidentified multi-body decays by a one-dimensional fit to the obtained $m(K\mu)$ distribution. With this procedure, the number of fit components which needs to be described is drastically reduced and correlations are implicitly accounted for.

Details of the fit in $m(D^0\pi)$ are presented in Section 6.2, where the peaking signal-like part is described by a phenomenological analytical function and the non-peaking component is modeled by an artificially created random-pion sample derived from data. Section 6.3 describes instead the subsequent fit in $m(K\mu)$, where the mass shapes are extracted from simulated events. Therefore, a huge sample of inclusive D^* decays has been generated which is briefly described in Section 6.3 and in more detail in Appendix B.

6.2 A MODEL TO FIT $m(D^0\pi)$

As just explained, the purpose of the $m(D^0\pi)$ fit is to isolate the peaking component, including signal and misidentified multibody decays, from the non-peaking component, which is mainly composed by random associations of real D^0 candidates and uncorrelated soft pions from the primary vertex. For this a reliable model for the random pions is required. The idea is to mimic fake D^* candidates by associating D^0 candidates from one event to soft pions from another event and extract the $m(D^0\pi)$ shape from them.

A naive association is not sufficient, as it is only possible to sample from already selected candidates and the kinematics of the final state particles are this way not correctly reproduced. An example, is shown in Figure 34, where the transverse momentum distribution of the soft pion from a sideband of the naively mixed random-pion sample is compared to the same sideband of the WS data[1]. Since the mass shape may depend on the kinematical features of the sample, the association

1 It has been checked that RS and WS sideband do not show significant differences.

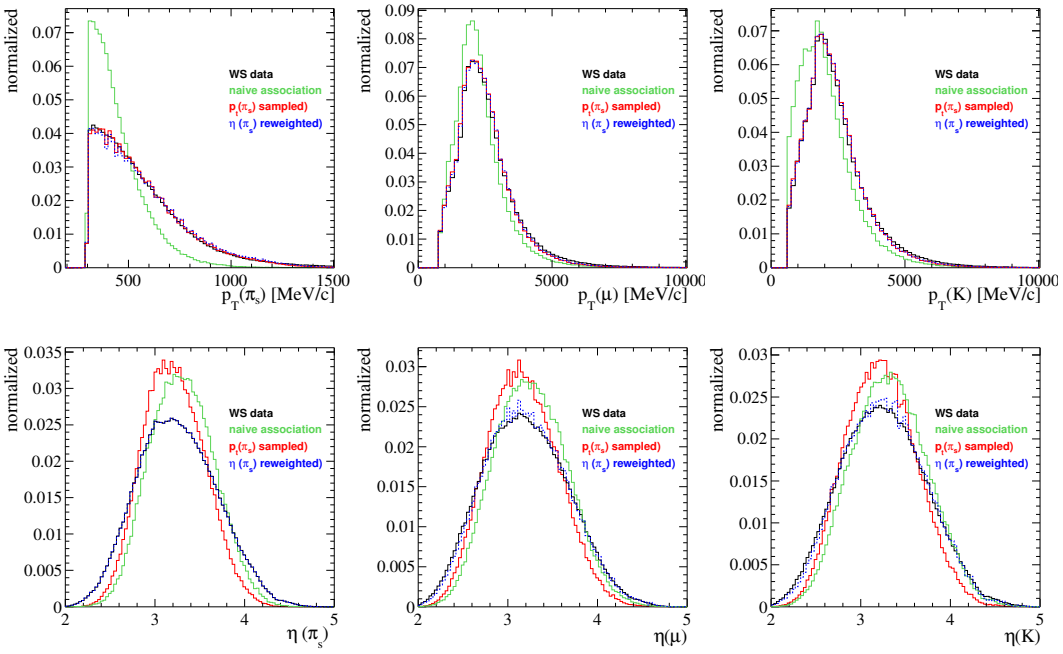


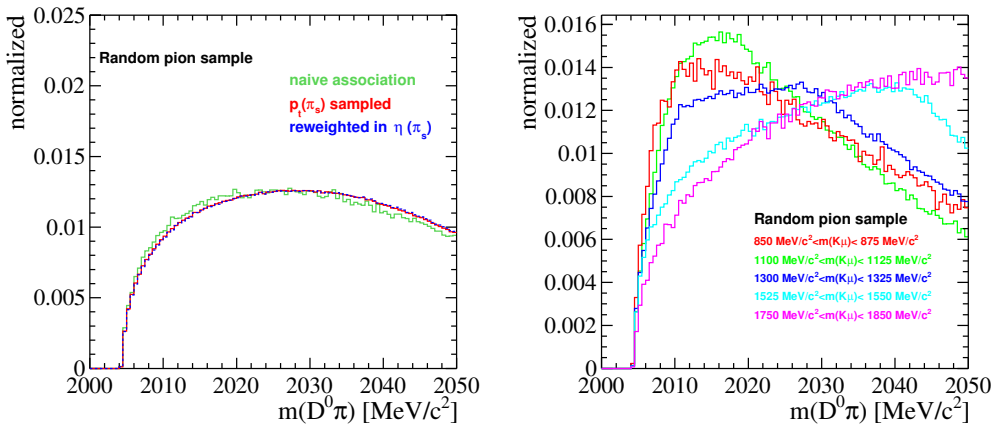
Figure 34: Transverse momentum and η distibution of the final state particles at the different steps of the procedure.

procedure is tuned to create a background sample consistent with the kinematics of the data sideband defined by $m(D^0\pi) > 2035\,\text{MeV}/c^2$. In particular, the artificial association of D^0 and soft pions is done in the following steps:

1. a soft pions from one event is combined to a D^0 candidates from another event;
2. all selection cuts on the kinematics of the formed D^* candidate, *e.g.* $\Delta m < 200\,\text{MeV}/c^2$, are applied;
3. the combinations surviving the cuts are sampled according to the $p_T(\pi_s)$ distribution observed in the sideband of the WS data (with a keep-or-reject method).

The $D^0\pi$ combinations obtained with this procedure shows increased agreement with the WS data sideband not only in $p_T(\pi)$ but also in the distributions of kaon and muon momenta, as shown in Figure 34. However, residual discrepancies remain in the pseudorapidity distributions. These are adjusted by assigning a weight to each event. The weight is calculated from the ratio of the normalized $\eta(\pi)$ distributions observed in the WS and random-pion samples. Again, the procedure also corrects for the deviations observed in $\eta(K)$ and $\eta(\mu)$ as shown in Figure 34.

Figure 35a shows the $m(D^0\pi)$ distibution of the mixed-events associations at the different steps of the procedure. While correctly reproducing all kinematical variables is important to ensure a reliable model for the $m(D^0\pi)$ distribution of the random pions, it is worth noticing that the effect of the p_T sampling and the η-reweighting on the mass shape is marginal (especially the latter). Figure 35b shows



(a) $m(D^0\pi)$ distibution of the mixed-events associations at the different steps of the procedure.

(b) $m(D^0\pi)$ distribution of the mixed-events associations in different $m(K\mu)$ bins.

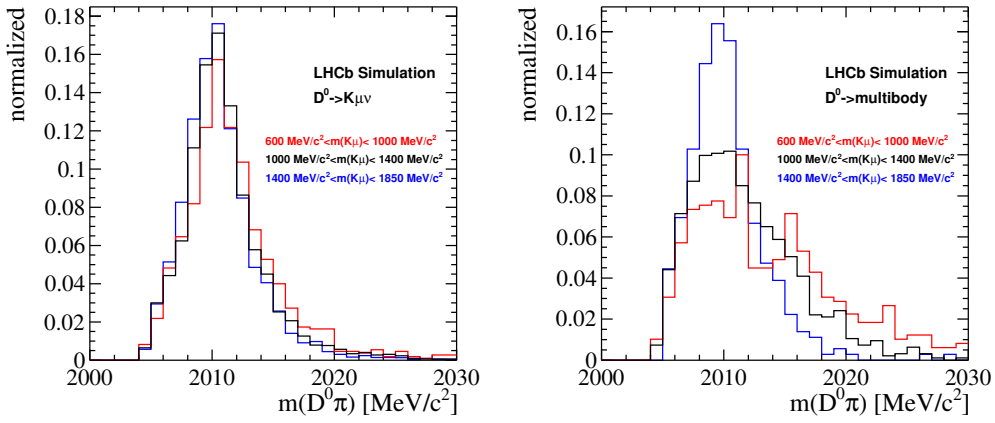
Figure 35

the $m(D^0\pi)$ distribution of the mixed-events associations for different $m(K\mu)$ bins. The changes in shapes are mainly caused by the stripping selection requirement on Δm, which introduces correlations between $m(D^0\pi)$ and $m(K\mu)$.

In total, the random-pion sample consists of approximately 150×10^6 candidates and therefore the available statistic is much higher than the data sample. These events are used to build histograms that are used in the $m(D^0\pi)$ fit as templates for the random-pion background PDF:

$$\mathcal{P}_{\mathrm{non-peaking}}(m) = \frac{1}{\mathcal{N}_{\mathcal{R}}}\mathcal{R}(m),$$

where $\mathcal{R}(m)$ returns the value of the histogram filled with random pions and $\mathcal{N}_{\mathcal{R}}$ is an appropriate normalization factor.



(a) Simulated $m(D^0\pi)$ distribution for signal candidates.

(b) Simulated $m(D^0\pi)$ distribution for misidentified multi-body decays.

Figure 36

The events peaking in $m(D^0\pi)$ are instead described by an empirical analytic function, with sufficient degrees of freedom to account for the broad peaking structure, the long tail and the sharp turn-on observed in simulated decays. The mass shapes for a simulated sample of misidentified multi-body D^0 decays and for the $D^0 \to K\mu\nu$ signal, in different $m(K\mu)$ bins, are shown in Figure 36. For signal candidates, the $m(D^0\pi)$ mass shape is only subject to small changes within different $m(K\mu)$ bins. Comparing misidentified multi-body decays, however, the shape in $m(D^0\pi)$ is varying a lot, because the composition of the different multi-body components is depending on the considered $m(K\mu)$ bin. Generally, the shape for the misidentified multi-body decays is broader compared to signal. The PDF used to account for all these effects is the following:

$$\begin{aligned}\mathcal{P}_{peaking}(m|\vec{\mu},\vec{\sigma},\vec{a},\vec{b},\vec{f}) = {} & f_1\mathcal{G}_1(m|\mu_1,\sigma_1)\\ & + (1-f_1)f_2\mathcal{G}_2(m|\mu_2,\sigma_2)\\ & + (1-f_1)(1-f_2)f_3\mathcal{G}_3(m|\mu_3,\sigma_3)\\ & + (1-f_1)(1-f_2)(1-f_3)f_4\mathcal{T}_1(m|a_1,b_1)\\ & + (1-f_1)(1-f_2)(1-f_3)(1-f_4)\mathcal{T}_2(m|a_2,b_2),\end{aligned}$$

where $\mathcal{G}_i$ are three Gaussian distributions describing the bulk and the high-mass tail of the peaking structure,

$$\mathcal{G}_i(m|\mu_i,\sigma_i) = \frac{1}{\mathcal{N}_{\mathcal{G}_i}} e^{-\frac{(m-\mu_i)^2}{2\sigma_i^2}},$$

and $\mathcal{T}_i$ are used to model the sharp turn-on close to threshold,

$$\mathcal{T}_i(m|a_i,b_i) = \frac{1}{\mathcal{N}_{\mathcal{T}_i}}(m-m_{min})^{a_i}\cdot e^{-b_i(m-m_{min})}.$$

The factors $\mathcal{N}_{\mathcal{G}_i}$ and $\mathcal{N}_{\mathcal{T}_i}$ normalize the corresponding functions to unitary area and the fractions $f_i \in [0,1]$ determine the relative contribution of the individual functions. The kinematic threshold at $m_{min} = m_{D^0} + m_{\pi^+} \approx 2004.41\,\mathrm{MeV/c^2}$ is fixed in the fit.

The final fitting function is then the sum of the peaking part, including signal and misidentified multi-body decays, and the non-peaking part including random pions and pure combinatorial background:

$$\mathcal{F}_{total}(m|\vec{\theta}) = \mathcal{N}_{peaking}\mathcal{P}_{peaking}(m|\vec{\theta}) + \mathcal{N}_{non-peaking}\mathcal{P}_{non-peaking}(m)$$

Here, $\vec{\theta}$ includes all the free parameters used to describe the signal shape. Since all the components are properly normalized, the coefficients $\mathcal{N}_{peaking}$ and $\mathcal{N}_{non-peaking}$ correspond to the yields of candidates peaking and not peaking in $m(D^0\pi)$, respectively. A plot of $\mathcal{P}_{peaking}$ and $\mathcal{P}_{non-peaking}$ is shown in Figure 37.

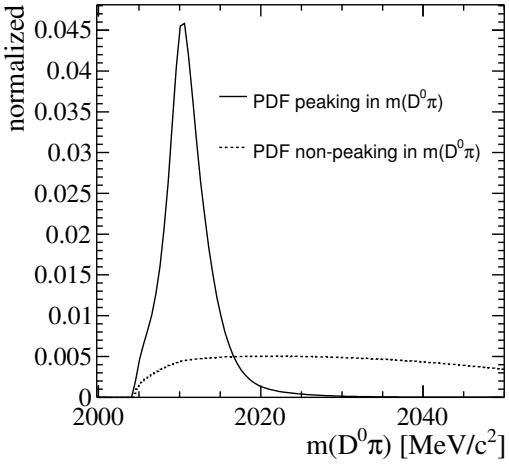


Figure 37: PFDs used in the fit to $m(D^0\pi)$.

6.3 A MODEL TO FIT $m(K\mu)$

As mentioned in Section 3.2, a large contamination from multi-body decays can enter both the RS and the WS samples by misidentification of one or more final state particles. The shape in $m(K\mu)$ depends on the actual composition and therefore can only be studied with simulation. A sample of simulated inclusive $D^{*+} \to D^0\pi^+$ decays has been produced, where all possible decays of the D^0 are generated according to the known branching ratios [25].

A reliable model is only possible if not only the generation but also the full detector response is correctly simulated to account for resolution effects and misidentification probabilities. Generally, this is very CPU intensive and requires time (typically one minute per event). To accelerate the generation step, a private Monte Carlo production chain is set up, where the PARTICLE GUN generator is used. Contrary to the standard production chain, where PYTHIA [54] simulates the full proton-proton interaction, PARTICLE GUN generates just the decay of interest allowing to simulate one event in approximately three seconds. Details on this alternative Monte Carlo production are described in Appendix B. A total of approximately 7×10^8 decays have been generated, of which only approximately 73×10^3 (18×10^3) RS candidates survive the stripping (offline) selection. Of these approximately 63×10^3 (17×10^3) are $D^0 \to K^-\mu^+\nu$ signal, while the remaining are misidentified background from multi-body D^0 decays. As mixed and unmixed D^0 mesons feature the same kinematical properties, the signal mass shape for candidates reconstructed as RS and WS is the same. The situation is different for misidentified multi-body decays. Here, the relative composition in the RS and WS samples can differ as the relevant contributions enter the sample with branching ratios different misidentification probabilities. For example, consider the suppressed $D^0 \to K^+\pi^-\pi^0$ decay ($\mathcal{B} \approx 3 \times 10^{-4}$), which contributes to the WS sample when the pion is misidentified as a muon. It will instead contribute to

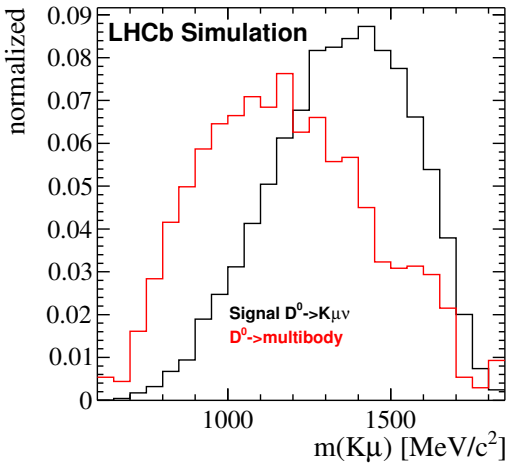


Figure 38: Visible mass for signal candidates and misidentified multi-body decays taken from a sample of inclusive D^0 decays.

the RS sample only if both the kaon and the pion are misidentified. The opposite happens for the favoured $D^0 \to K^-\pi^+\pi^0$ decay ($\mathcal{B} \approx 14 \times 10^{-2}$). Therefore, while the suppressed decay will be negligible for the RS but not for the WS signal, the favoured one will contribute to both with different relative contamination. As a result, the $m(K\mu)$ shape for the misidentified multi-body background is expected to be different for between RS and WS candidates. The number simulated misidentified WS candidates surviving the full selection is only 13 and thus unusable for the determination of the corresponding $m(K\mu)$ shape. For this reason, while the analysis procedure is fully completed on the RS sample, no attempt is performed to fit the $K\mu$-mass distribution of the WS data. Nevertheless, a preliminary estimation of the expected sensitivity on the mixing ratio is provided in Section 6.5.1, using pseudo-experiments that mimic the observed data distributions.

Figure 38 shows the $m(K\mu)$ distributions for the simulated events reconstructed as RS candidates, which are used as PDF templates in the fit. The distributions are well separated and expected to be identifiable by a fit. The corresponding distribution for multi-body decays reconstructed as WS candidates remains unknown due to low Monte Carlo statistics.

The total fitting function is therefore the sum of two components,

$$\mathcal{P}_{total}(m) = \mathcal{N}_{sig}\mathcal{P}_{sig}(m) + \mathcal{N}_{bkg}\mathcal{P}_{bkg}(m),$$

where the signal and background components,

$$\mathcal{P}_{sig}(m) = \frac{1}{\mathcal{N}_{\mathcal{S}}}\mathcal{S}(m)$$

and

$$\mathcal{P}_{bkg}(m) = \frac{1}{\mathcal{N}_{\mathcal{M}}}\mathcal{M}(m),$$

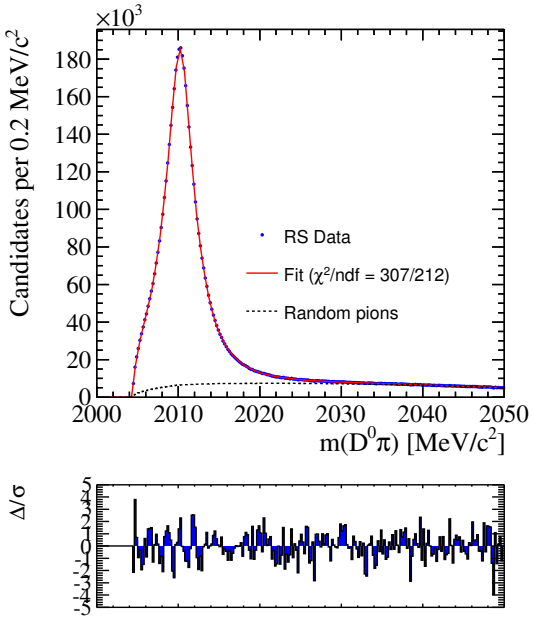


Figure 39: Fit to $m(D^0\pi)$ to the integrated RS sample.

correspond to the appropriately normalized histograms for signal and multi-body decays background, respectively. In this definition, the parameters N_{sig} and N_{bkg} directly correspond to signal yield and background from misidentified multi-body decays , respectively.

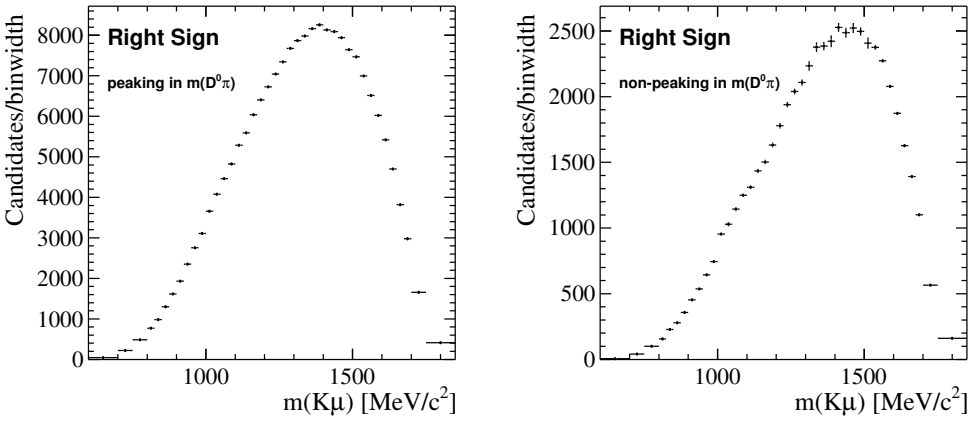
6.4 RESULTS OF THE FIT TO THE RIGHT-SIGN SAMPLE

The fit model described in Section 6.2 is now used to determine the signal yield of the RS sample. Figure 39 shows the fit in $m(D^0\pi)$ to the RS sample, integrated over the full $m(K\mu)$ range. The yields of peaking events obtained by this fit is:

$$N^{RS}_{peaking} = (4.944 \pm 0.006) \times 10^6.$$

This number includes both signal and misidentified multi-body decays. The analytical function defined in Section 6.3 and the model for the random combinations are sufficiently describing the shape of the observed peaking and non-peaking components in $m(D^0\pi)$.

Following the procedure described in Section 6.3, the data are divided in 41 bins of $m(K\mu)$, with variable bin widths to ensure reasonable amount of candidates in each bin. The same binning is used to divide the random-pion sample which is used as background template for the $m(D^0\pi)$ fits. In the individual fits, all parameters of the signal PDF explained in Section 6.2 are left free to float. The $m(D^0\pi)$ signal shape is expected to vary as a function of $m(K\mu)$, as it depends on the relative composition of semileptonic and multi-body decays. Low-$m(K\mu)$ bins



(a) RS yields of the peaking component obtained by the individual fits in the corresponding $m(K\mu)$ bins.

(b) RS yields of the non-peaking component obtained by the individual fits in the corresponding $m(K\mu)$ bins.

Figure 40

are expected to be dominated by peaking background, whereas at higher masses the semileptonic signal contribution is enhanced (see Figure 38).

The projections of the $m(D^0\pi)$ fits for each $m(K\mu)$ bin can be found in Appendix C. The corresponding yields for peaking and non-peaking components are shown in Figure 40, together with their statistical uncertainties. The extracted $m(K\mu)$ shapes for the peaking and non-peaking components in $m(D^0\pi)$ are very similar and both resemble the Monte Carlo signal shape of Figure 38. This again confirms that most of the subtracted non-peaking component is due to real semileptonic D^0 decays with randomly attached pions and also that most of the peaking component is made of signal.

To determine the RS signal yield, the histogram shown in Figure 40a is fitted as explained in Section 6.3. Figure 41 shows the resulting projection, from which a fraction of $(10.55 \pm 0.09)\%$ of background from misidentified multi-body decays is extracted. The fit quality is satisfactory and the number of $D^{*+} \to D^0(\to K^-\mu^+\nu)\pi^+$ decays estimated is

$$N_{sig}^{RS} = (4.44 \pm 0.05) \times 10^6.$$

A larger sample of simulated inclusive D^{*+} decays can certainly allow a more refined fitting model and improved precision.

6.5 PRELIMINARY FIT TO THE WRONG-SIGN SAMPLE

To measure the ratio of WS and RS events, the same procedure has to be applied to the WS sample. Figure 42 shows the fit projection to $m(D^0\pi)$ for the WS data,

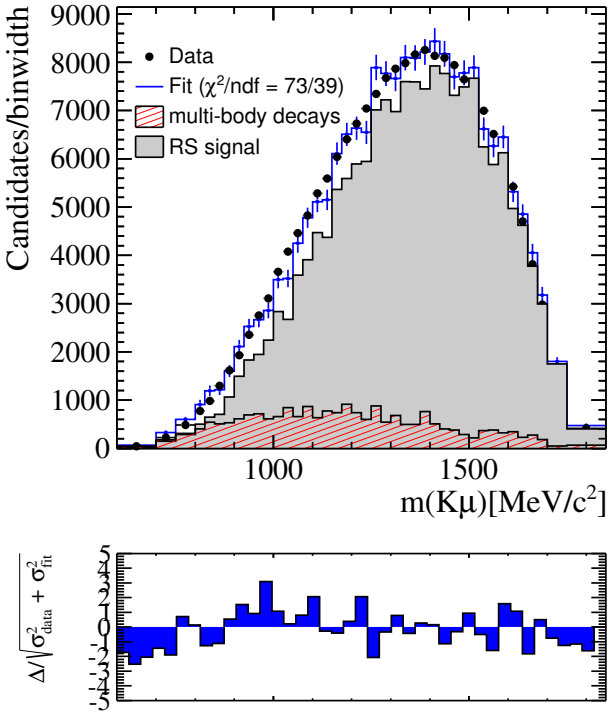


Figure 41: Fit to $m(K\mu)$ to RS yields peaking in $m(D^0\pi)$.

integrated over the full $m(K\mu)$ range. The total amount of peaking events, again including signal and misidentified multi-body decays, is

$$N^{WS}_{peaking} = (2.03 \pm 0.04) \times 10^5.$$

The same binning as defined for the RS fit is used to divide the WS sample in subsamples according to the visible D^0 mass. As the statistic is significantly lower compared to the RS sample, the number of free parameters for the signal component in the individual $m(D^0\pi)$ fits is reduced. A double-Gaussian function is sufficient in the first 23 bins, while a single-Gaussian distribution is used for the rest.

Figure 43 shows the yields of the peaking and the non-peaking components obtained by the individual fits in the corresponding $m(K\mu)$ bins. The projections of the fits for each $m(K\mu)$ bin are in Appendix C. Also in this case (see Figure 44), the distribution of the non-peaking component looks consistent to what observed in the RS and signal Monte Carlo samples, confirming again that the dominant non-peaking background is due to random pions. The $m(K\mu)$ distribution for the candidates peaking in $m(D^0\pi)$, however, differs completely from both the signal and the multi-body decays shapes shown in Figure 38. As explained in Section 6.3, the composition, and therefore the mass distribution, of multi-body decays reconstructed as RS candidates is expected to differ from that of the decays reconstructed as WS candidates. A measurement of the WS signal yield requires the

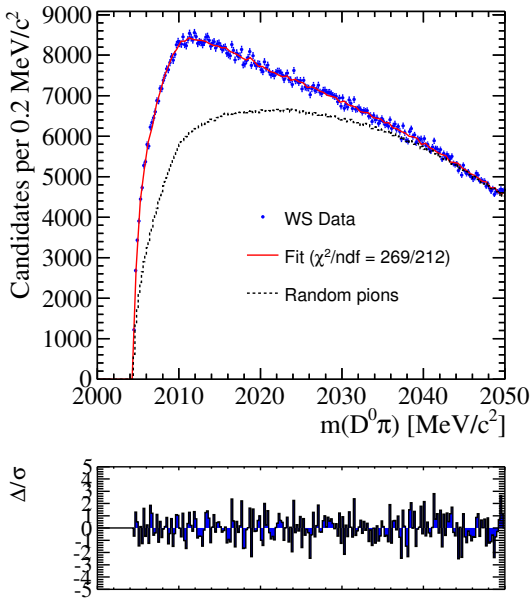


Figure 42: Fit to $m(D^0\pi)$ to the integrated WS sample.

knowledge of this distribution and is at this stage not performed due to lack of Monte Carlo statistics.

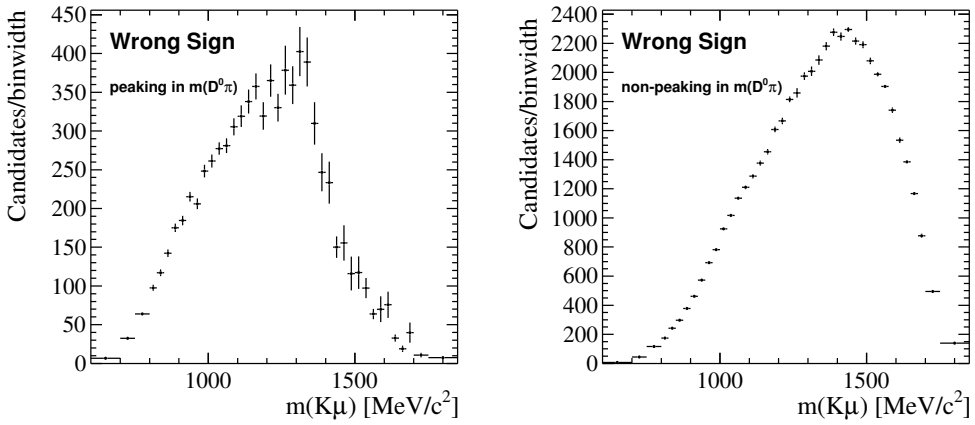
6.5.1 *Expected sensitivity to the mixing ratio*

To estimate the expected sensitivity to the mixing ratio $R_M = (x^2 + y^2)/2$, pseudo-experiments that mimic the observed data distributions are performed.

Using the RS signal yield measured in Section 6.4, the world average values of the mixing parameters [23] and the value of $\Gamma^2 \langle t^2 \rangle = 11.11$ determined from simulation, it is possible to calculate the expected number of mixed WS signal events with Equation 8 to be

$$N^{WS}_{\text{exp sig}} = N^{RS}_{\text{sig}} \frac{R_M}{2} \Gamma^2 \langle t^2 \rangle \approx 697. \tag{19}$$

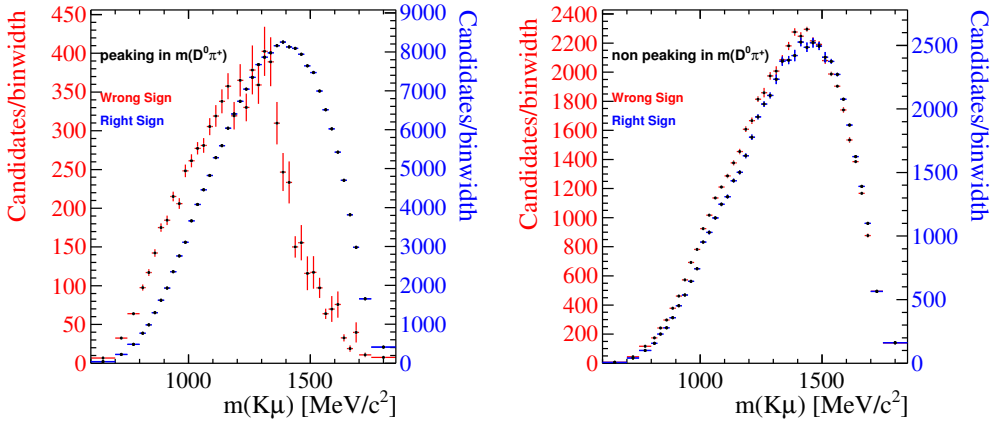
These correspond to $\sim 0.3\%$ of the measured yield of WS candidates peaking in $m(D^0\pi)$. Assuming therefore that the measured $m(K\mu)$ distribution of Figure 40a is esentially composed of only background, it is possible to generate pseudo-experiments acoording to the observed features of the WS data where a known and fluctuated mixing signal is injected. These pseudo-experiments can then be used to estimate the expected precision on R_M provided by the analysis procedure developed here. While performing the $m(K\mu)$ fit, the mass template for the multibody background of the WS sample is assumed to be known with a precision corresponding to the statistic of the available data. Therfore the statistic uncertainty due to limited Monte Carlo is taken into account.



(a) WS yields of the peaking component obtained by the individual fits in the corresponding $m(K\mu)$ bins.

(b) WS yields of the non-peaking component obtained by the individual fits in the corresponding $m(K\mu)$ bins.

Figure 43

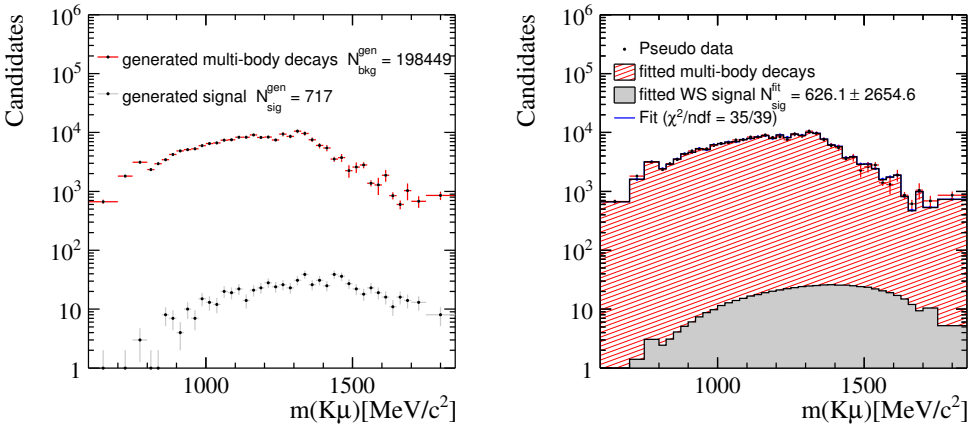


(a) Signal yields extracted in the individual $m(K\mu)$ bins.

(b) Corresponding background yields.

Figure 44

Figure 45a shows an example for generated $m(K\mu)$ distributions for signal and misidentified mulit-body decays in a pseudo-experiment. In Figure 45b, the corresponding fit to the combined distribution is shown. The extracted mixing ratio is $R_M = (0.0025 \pm 0.0107)\%$ and therfore still compatible with zero. To better estimate the statisitical precision, 10^4 independent pseudo-experiments are performed, where the distribution of the resulting uncertainties on R_M are shown in Figure 46. The estimated statistical precision on the mixing ratio is 0.01%, which corresponds to a factor of two improvement with respect to the current world's best determination of $R_M = (0.013 \pm 0.022)\%$ using semileptonic final states [31]. The estimation does not include systematic uncertainties due to the mass shapes extracted from simulations, such as form factors used to describe the decay amplitutes which are not always well known. These systematic uncertainties have to be evaluated carfully when sufficient Monte Carlo statistic is available.



(a) Example for generated signal and background distribution in a pseudo-experiment.

(b) Fit to the generated pseudo data.

Figure 45

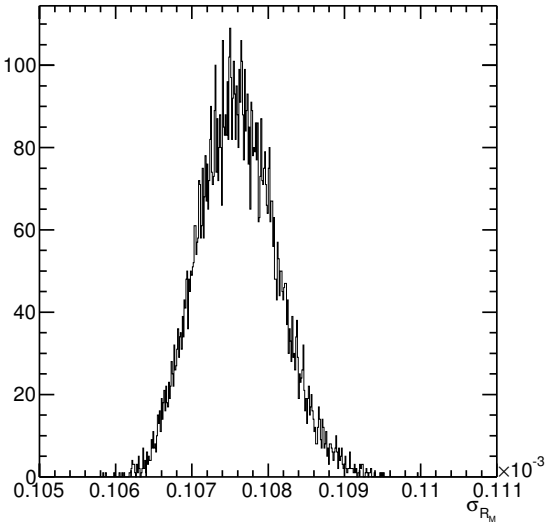


Figure 46: Distribution of the statistical uncertainties obtained in 10^4 independent pseudo-experiments.

SUMMARY AND OUTLOOK

This thesis presents a study of $D^0 \to K^{\pm}\mu^{\mp}\nu$ decays for a possible measurement of the charm mixing rate at LHCb. Charm mixing describes periodical meson-antimeson transitions in the time evolution of a neutral D decays. It is a powerful indirect probe for non-SM particles contributing to quantum loops and allows to search for new sources of CP violation. The semileptonic decay channel opens the chance to measure the mixing rate directly through a time-integrated analysis. By selecting D^0 mesons arising from strong-interaction decays of $D^{*+} \to D^0\pi^+$, the flavour of the D^0 meson at production time can be determined by the charge of the accompanying pion. With semileptonic final states, the flavour of the D^0 at the time of decay is unambiguously determined by the charge of the muon and allows to decide if flavour changing transitions occurred or not. The work of this thesis represents the first attempt to measure charm mixing in this channel at LHCb. The entire analysis strategy is developed from the start.

In LHCb the presence of the neutrino does not allow to fully reconstruct the semileptonic final state, introducing much larger background contamination than with fully reconstructed charm decays. The thesis presents therefore a detailed study of several algorithms that may be used to correct for the undetected neutrino and determine the D^0 momentum. This study shows that substantial improvements in momentum resolution can be achieved, which also result in improved mass as well as decay-time resolutions, hence providing an important handle to suppress backgrounds. Expected sources of backgrounds are studied and an specific selection requirements are developed to isolate the semileptonic signal. These include a multivariate analysis technique based on a boosted decision tree, an optimized lower decay-time threshold, particle-identification requirements and multiple candidate removal.

An original method to statistically separate the signal from all background components surviving the final selection is developed. The two-dimensional space $m(D^0\pi) \times m(K\mu)$ is found to show sufficient separation between the distinct components. First, fits to the $m(D^0\pi)$ distribution in bins of $m(K\mu)$ are performed to remove background from fake D^{*+} candidates. The $m(D^0\pi)$ shape of this non-peaking background is modeled by an artificially created sample, where soft pion candidates from one event are randomly associated to D^0 candidates from another. In this work, an new association procedure is developed to ensure a reliable background sample consistent with the observed kinematics of the data sideband. The $m(K\mu)$ distribution of the resulting yields of candidates peaking in $m(D^0\pi)$ are fit, using templates derived from simulation, to disentangle the semileptonic signal from misidentified D^0 decays to multi-body final states. For these purpose, large simulated samples of inclusive D^{*+} decays are needed to account for all possible

sources of contamination. The thesis includes therefore the development of a new Monte Carlo production chain, which allows to reduce by a factor 20 the CPU-time required to simulate an event. This part of the work, partially based on already existing tools, is in the process to be adopted as official Monte Carlo production within the LHCb collaboration. The analysis procedure is fully tested on the data corresponding to $2\,\mathrm{fb}^{-1}$ of integrated luminosity, with unmixed $D^0 \to K^-\mu^+\nu$ decays, and yields satisfactory results. No attempt is performed to fit the mixing signal, as further studies are needed to better understand the background faking the $D^0 \to K^+\mu^-\nu$ decay. These require simulated samples of larger statistics which will become available when the new simulation chain will be adopted as official production chain.

A sensitivity study has been performed with pseudo-experiments that mimic the features of the real data. The study shows that with the available data a statistical precision of 0.01% on the determination of the mixing rate $R_M = (x^2 + y^2)/2$ can be achieved. Comparing to the current world's most precise measurement of charm mixing in semileptonic decays from Belle [31], $(0.013 \pm 0.022)\%$, this would be an improvement by a factor of two. In this sensitivity study, systematic uncertainties related to mass shapes taken from simulations are not considered and need to be evaluated.

While this first and very preliminary study already shows that a competitive measurement of charm mixing with semileptonic D^0 decays is possible at LHCb, it is clear that many aspects of the analysis can be further improved in the near future. Moreover, larger data samples will become available during the upcoming Run II, where more efficient trigger and offline selections are foreseen to be used. During this period, LHCb is expected to collect more than $5\,\mathrm{fb}^{-1}$ of integrated luminosity and will therefore be in an excellent position to contribute to our understanding of mixing and CP violation with semileptonic charm decays.

A

BOOSTED DECISION TREES

This chapter explains the general principle of boosted decision trees, focusing on the configuration chosen in this analysis.

A.1 THE IDEA OF MULTIVARIATE ANALYSES

A boosted decision tree is an example for a multivariate analysis. Multivariate techniques combine information of several discriminating variables into one single response variable, called a *multivariate classifier*. The signal-to-background discriminating information is this way fully exploited as all correlations between the discriminating variables are used to separate signal and background events. In the analysis described in this thesis, the *boosted dicision tree* (BDT) method [50] available in the *Toolkit for Multivariate Analysis* (TMVA) [49] is used. The following explains how the BDT work and how it has been configured.

A.2 BOOSTED DECISION TREES

A BDT is based on binary decision trees. Starting from the original sample, called the *root node*, the *best cut* on one of the discriminating variables is defined that divides the sample in two branches, as shown in Figure 47. The criterion used in this analysis to determine the best cut value is that minimizing the *Gini Index* [49], which is defined by $p(1-p)$, where p is the signal purity. This procedure is repeated subsequently for all other discriminating variables. At each step a new *node* of the tree is formed. The same variable can be used at several different nodes. The number of cuts per variable tested at a node is set to 20, which is the standard setting that represents a compromise between computing time and step size. The last sample at the end of a branch is called *leaf*. Depending of the composition of a leaf, it is labeled as signal-like or background-like. This classification is done according to the purity of the individual samples. The procedure of splitting is repeated until a convergence criterion is fulfilled. For this analysis, the convergence is reached when a node containing 2.5% of the total training sample is created.

Compared to a selection which is based on rectangular cuts, not only one hypercube in the multidimensional space spanned by the selection variables is accessible, but several signal- and background-like hypercubes at the end of the tree. This makes the decision tree a powerful tool, which is unfortunately very sensitive to statistical fluctuations of the training sample. Slightly different compositions in the root node can, indeed, cause large differences in the structure of the tree. For

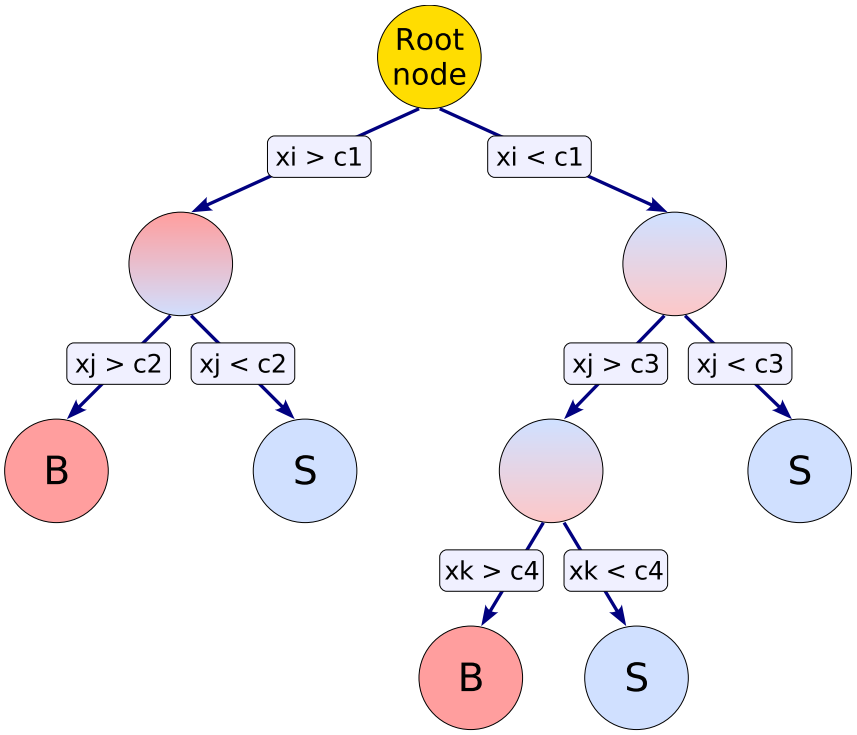


Figure 47: The principle of a binary decision tree [49].

this reason, not only one tree is derived from the training sample, but a so-called *forest* consisting of various trees. In the forest, every tree is subsequently subject to a *boosting* algorithm. After having derived a tree, the events that has been misclassified within a signal- or background-like leaf are assigned a weight. Several boosting algorithms are available, but in this analysis the *AdaBoost* algorithm [55] is used.

A.2.1 *The AdaBoost algorithm*

In the *AdaBoost* algorithm, the boosting is done in the following way. Let N be the total number of events in a training sample, x_i the set of discriminating variables for the ith event, m the index of the tree within a forest and w_i^m the weight of an event in tree m. The function $T_m(x_i)$ returns $(+1)$ (-1) if event i ended up in a signal-like (background-like) node for the tree m. If event i was really a signal (background) event the variable y_i is set to $+1$ (-1). To boost only misclassified events, the weight for the event i in the tree $m+1$ is defined as

$$w_i^{m+1} = w_i^m e^{\alpha_m \, I(y_i \neq T_m(x_i))},$$

where

$$\alpha_m = \beta \, \ln \frac{1-\mathrm{err}_m}{\mathrm{err}_m} \quad \text{with} \quad \mathrm{err}_m = \frac{\sum_{i=1}^{N} w_i I(y_i \neq T_m(x_i))}{\sum_{i=1}^{N} w_i}$$

and

$$I(y_i \neq T_m(x_i)) = \begin{cases} 1 & \text{if} \quad y_i \neq T_m(x_i) \\ 0 & \text{if} \quad y_i = T_m(x_i) \end{cases}.$$

The final weight for an event i is then calculated as the weighted sum over all the individual trees:

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

In this analysis, the boosting factor β is set to 0.5 and the number of trees is decided to be 850 to be robust against statistical fluctuations in the training sample.

B

MONTE CARLO SIMULATION USING PARTICLE GUN

This is chapter presents in more details the Particle Gun *generator and the Monte Carlo simulation chain used in this analysis. A comparison with the standard Monte Carlo production, based on the* Pythia *generator is also provided.*

B.1 MONTE CARLO SIMULATION AT LHCB

At LHCb, the production of Monte Carlo simulations is processed in five subsequent steps using different software packages based on the Gaudi framwork [56]:

gauss – The Gauss package [57] is used for generation of the signal process and the simulation part, where the the passage of the particles through the detector is simulated with Geant4 [58].

boole– Boole [59] is simulating the detector response due to the interaction of the generated and secondary particles with the detector components.

moore– The simulation of the LHCb trigger is done by Moore [60], which emulates the hardware trigger and executes the same algorithm used for real data in the software stage.

brunel– The offline reconstruction (tracking, particle identifications, *etc.*) is done by Brunel [61] in exactly same way as for real data.

davinci– DaVinci [62] is used to reconstruct and select signal topologies of a desired process by combining the particle candidates created by Brunel. The stripping selection is part of the DaVinci step.

In the generation phase of Gauss, the standard production tool used at LHCb is Pythia [54], where the proton-proton collision is simulated including the hard process and hadronization. The decay of the produced particles is governed by the particle decay simulation package EvtGen [63]. This produces large multiplicities of particles, whose passage through the detector has to be simulated by Geant4. The simulation process becomes therefore very CPU and time consuming. To speed up the production of the large Monte Carlo samples of simulated inclusive $D^{*+} \to D^0\pi^+$ decays required for this analysis, a simpler generation of only the particle of interest has been used in place of Pythia. This alternative generator is called Particle Gun and is described in the following.

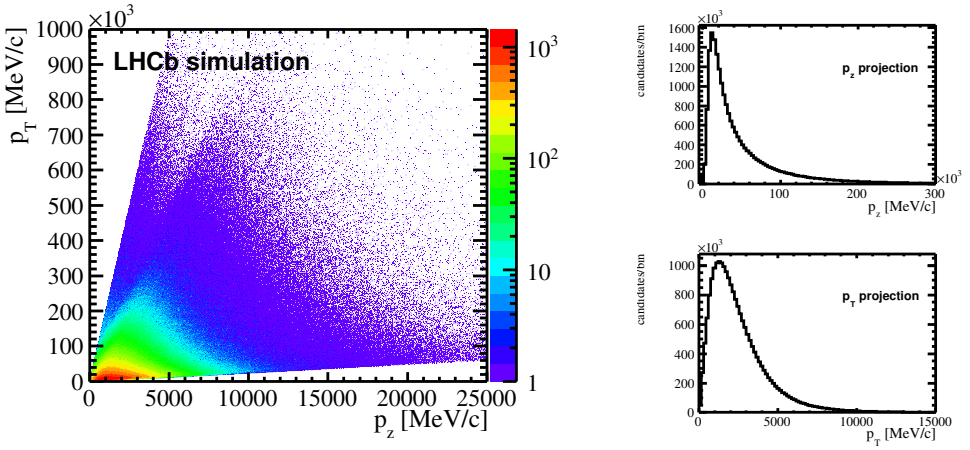


Figure 48: Two dimensional p_t,p_z distribution taken from alrady generated PYTHIA D^* decays and projections.

B.2 THE PRODUCTION CHAIN USING PARTICLE GUN

The idea of PARTICLE GUN is not to simulate the full proton-proton interaction, but to only send a single signal particle into the detector. The advantage is to benefit from time saving during the generation itself and mainly in the simulation part due to a very low track multiplicity.

As no proton-proton collision is simulated, the starting momentum of the signal particle has to be given as input. One possibility is to sample from an multidimensional histogram provided externally by the user. In this analysis, solely D^* decays are generated. A sample of already existing PYTHIA generated D^* decays is used to fill a two-dimensional histogram with the transverse momentum and the z-component of the momentum of the D^* candidates. The histogram shown in Figure 48 is used during the generation process to randomly sample p_T and p_z and assigning the components p_x, p_y, p_z to the D^* candidate, assuming a flat distribution in ϕ. To be more efficient, only (p_T,p_z) configurations leading to a D^* within the LHCb detector acceptance are allowed as shown by the visible kinematical borders of Figure 48. The D^* is forced to decay into a $D^0\pi^+$ pair by EVTGEN, followed by all possible D^0 decays according to their known branching ratios [25].

For simplicity only one D^{*+} candidate is generated per event. Its generated position in the detector will then coincide with the primary vertex. To simulate a realistic distribution, the primary vertex position is extracted event-by-event following the same parameterization of the luminous region used by PYTHIA. Figure 49 shows the three-dimensional distribution of the primary vertex position, as well as the corresponding projections along the three axes.

Following the generation, the digitalisation by BOOLE can be processed with standard configurations. Since, in the work here presented, we are not interested in determining precise trigger efficiencies, which may depend on the full event and not only on the signal-candidate particles, the MOORE step is skipped and the

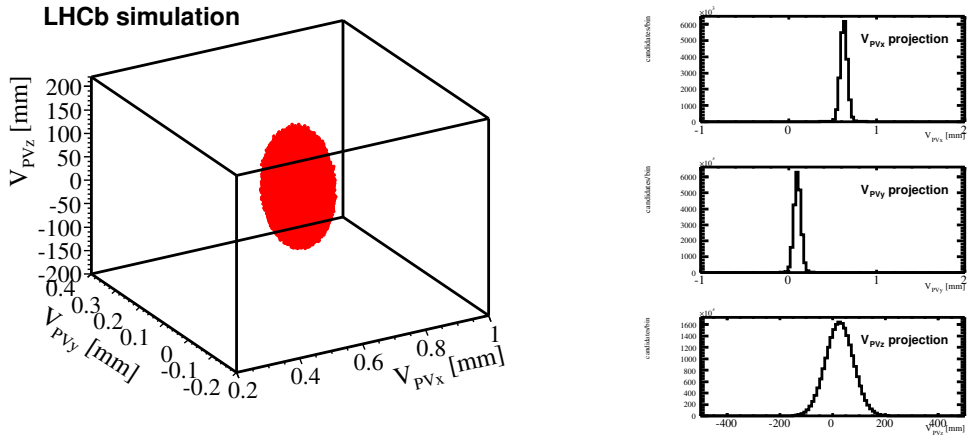


Figure 49: Three dimensional PV position (V_{PVx},V_{PVy},V_{PVz}) taken from alrady generated PYTHIA D^* decays and projections in each dimension.

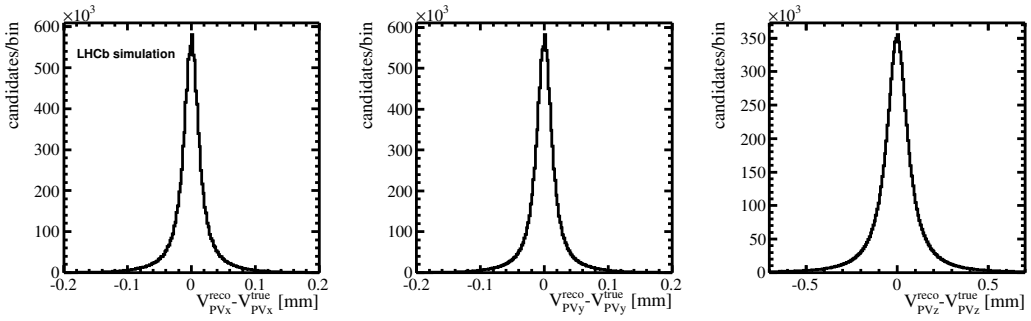


Figure 50: Projections of the LHCb PV resolution taken from minimum bias MC samples.

trigger selection is emulated later on offline-reconstructed quantities. During the BRUNEL reconstruction, the low track multiplicity does not allow to fit the primary vertex position. At this stage, a new tool has been implemented to smear the position of the generated primary vertex with realistic resolutions. The smearing is done according to an externally provided three-dimensional histogram, filled with the (x, y, z) resolutions extracted from a high-statistics simulated sample of minimum bias events. The projections of the three dimensional histogram are shown in Figure 50, where the resolution in each dimension $i = x, y, z$ is defined as $\sigma_{PVi} = V^{reco}_{PVi} - V^{true}_{PVi}$, where V^{reco}_{PVi} and V^{true}_{PVi} are the reconstructed and generated position of the primary vertex, respectively. Finally, in the DAVINCI step, the stripping selection explained in Chapter 5 is performed together with the emulation of the trigger.

B.3 COMPARISON PYTHIA AND PARTICLE GUN

The PARTICLE GUN production chain was tested by comparing a sample of 5×10^6 simulated $D^{*+} \to D^0(K^-\mu^+\nu)\pi^+$ candidates to an existing sample of equal amount produced with the standard Monte Carlo production based on PYTHIA.

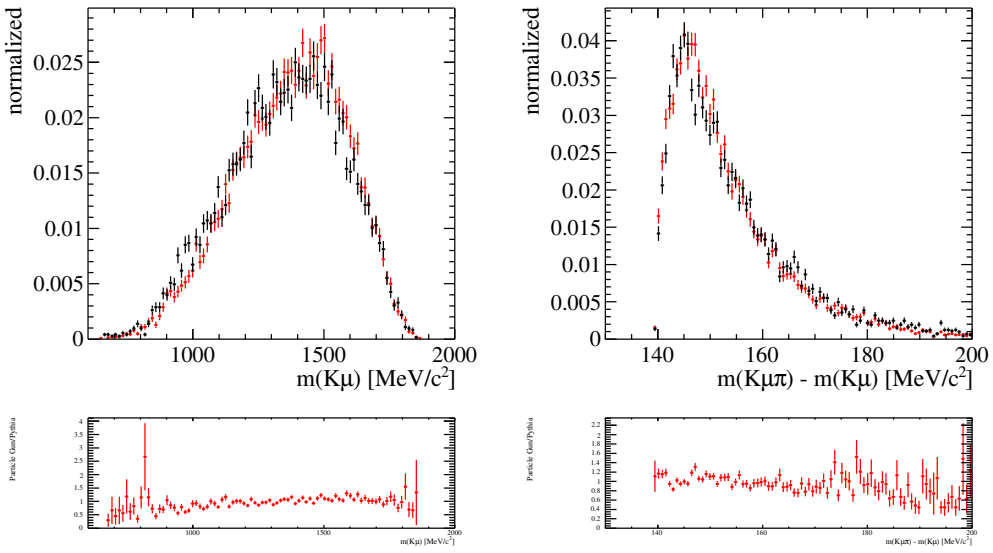


Figure 51: Comparison of visible D^0 mass and Δm for Monte Carlo samples generated using PYTHIA (black) and PARTICLE GUN (red).

Figure 51 and Figure 52 show the comparison of $m(K\mu)$, Δm, reconstructed primary-vertex position, transverse momenta and impact parameter of the reconstructed candidates. No significant differences are observed in the distribution, confirming that the PARTICLE GUN simulation chain is effective in reproducing the same kinematical features as that resulting from the standard simulation chain. Differences between PARTICLE GUN and PYTHIA are instead visible when comparing the $\chi^2(IP)$ and the particle-identification variables (see Figure 53). This is somewhat expected as these variables are correlated with the track multiplicity of the event. For the scope of this thesis, where a mass model for background studies is needed, the agreement obtained is however sufficient and these differences can be ignored.

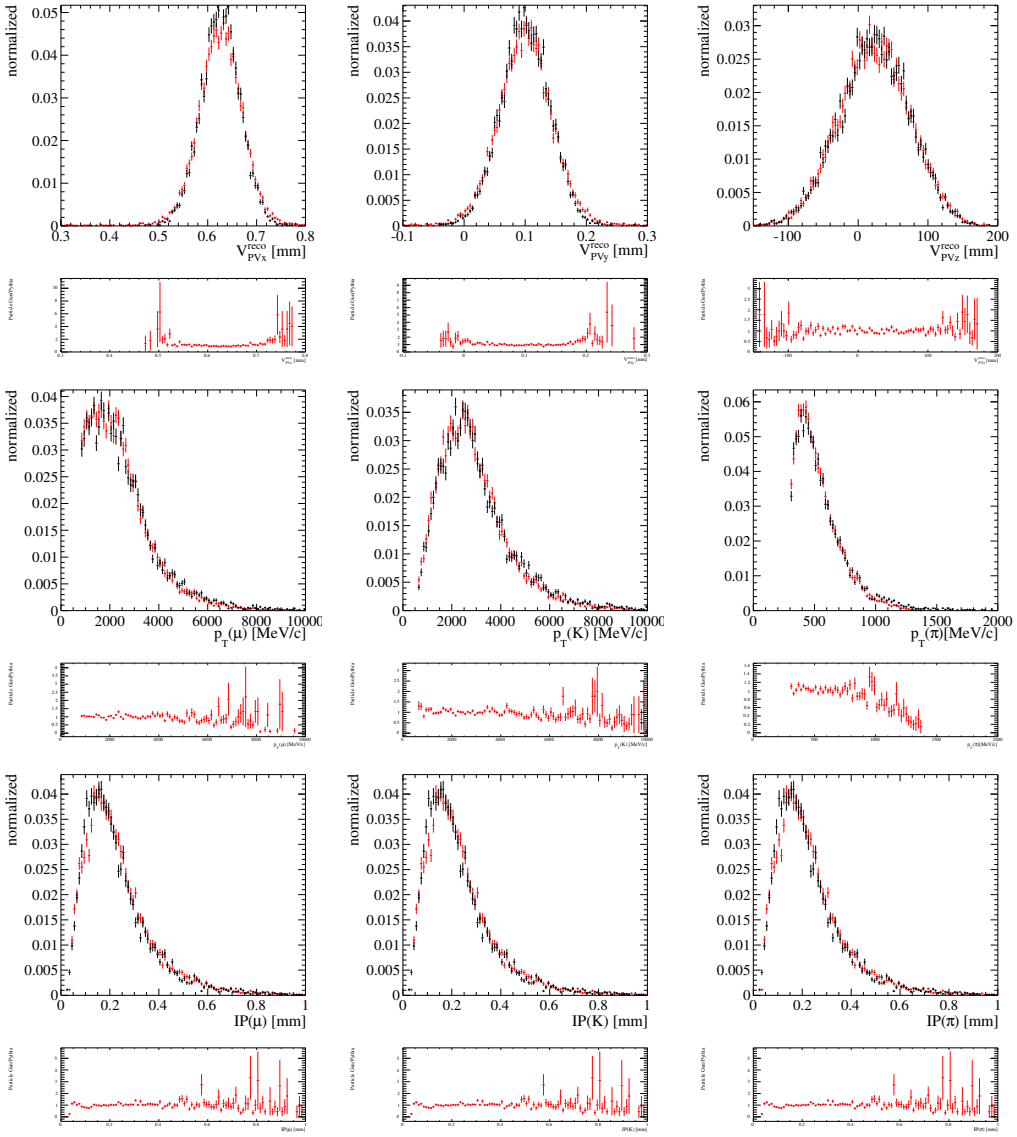


Figure 52: Comparison of reconstructed PV position, transverse momentum of the final state particles and impact parameter for reconstructed candidates generated by PYTHIA (black) and PARTICLE GUN (red). All distributions are in good agreement.

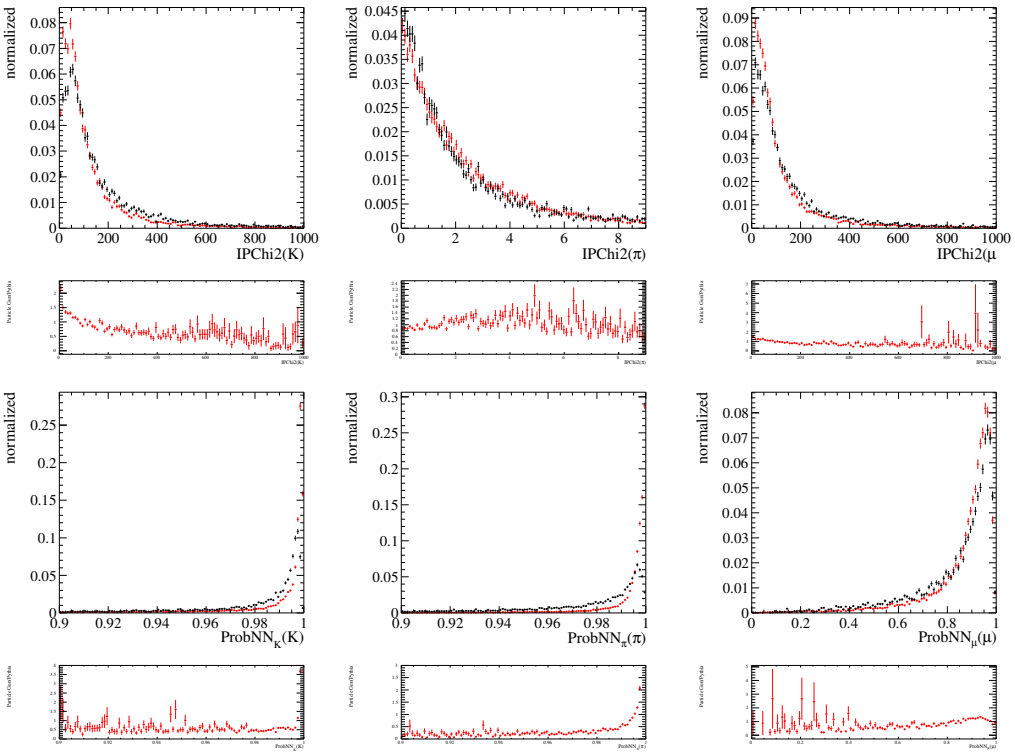


Figure 53: Impact parameter significance and PID variable for reconstructed candidates generated by PYTHIA (black) and PARTICLE GUN (red), where expected deviations are visible.

C

FITS IN BINS OF $m(K\mu)$

In this chapter, the fits in $m(D^0\pi)$ for all bin in $m(K\mu)$ are shown. See Chapter 6 for details.

C.1 RIGHT SIGN SAMPLE

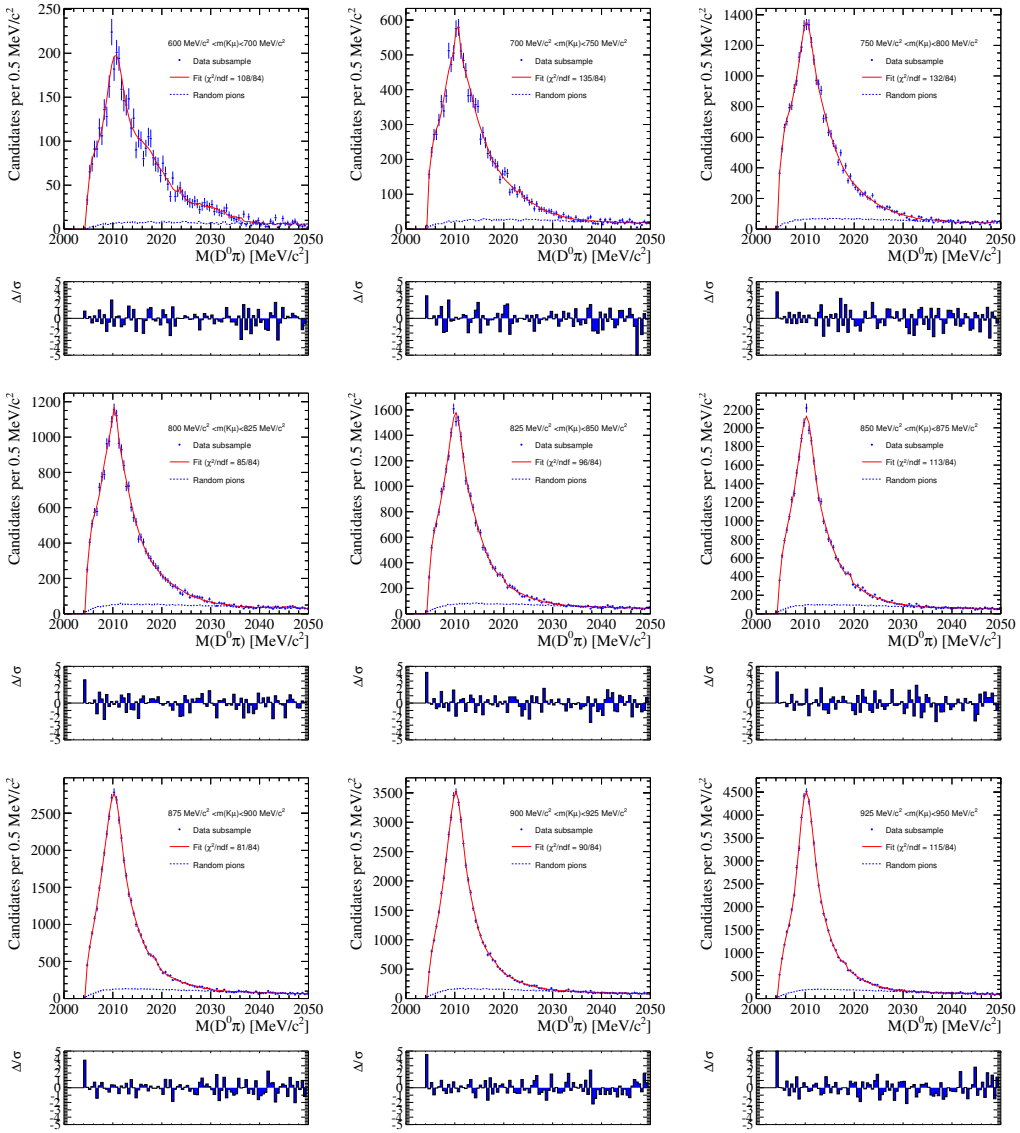


Figure 54: Fits in bins of $m(K\mu)$ RS sample (part 1 of 4)

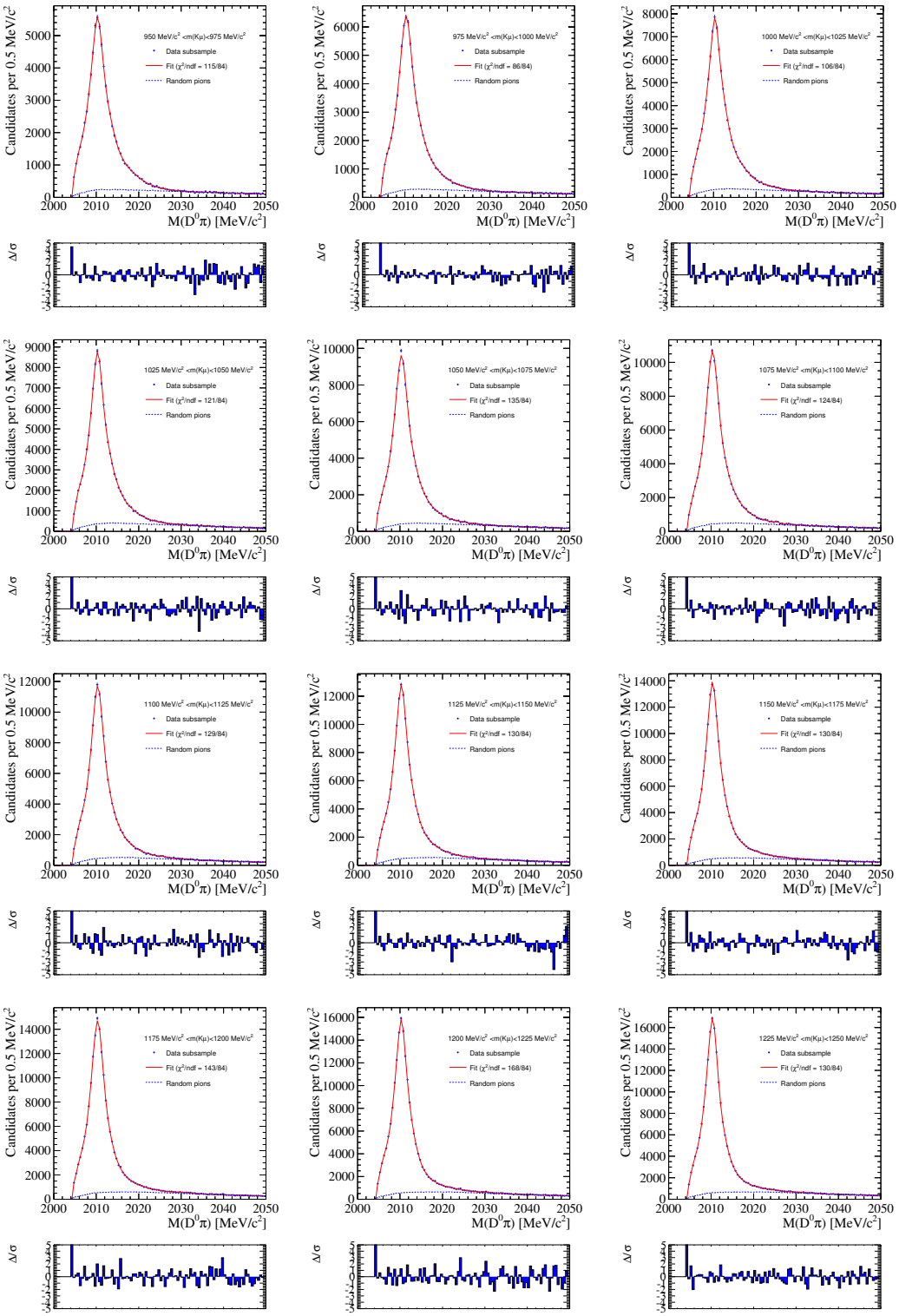


Figure 55: Fits in bins of m(Kμ) RS sample (part 2 of 4)

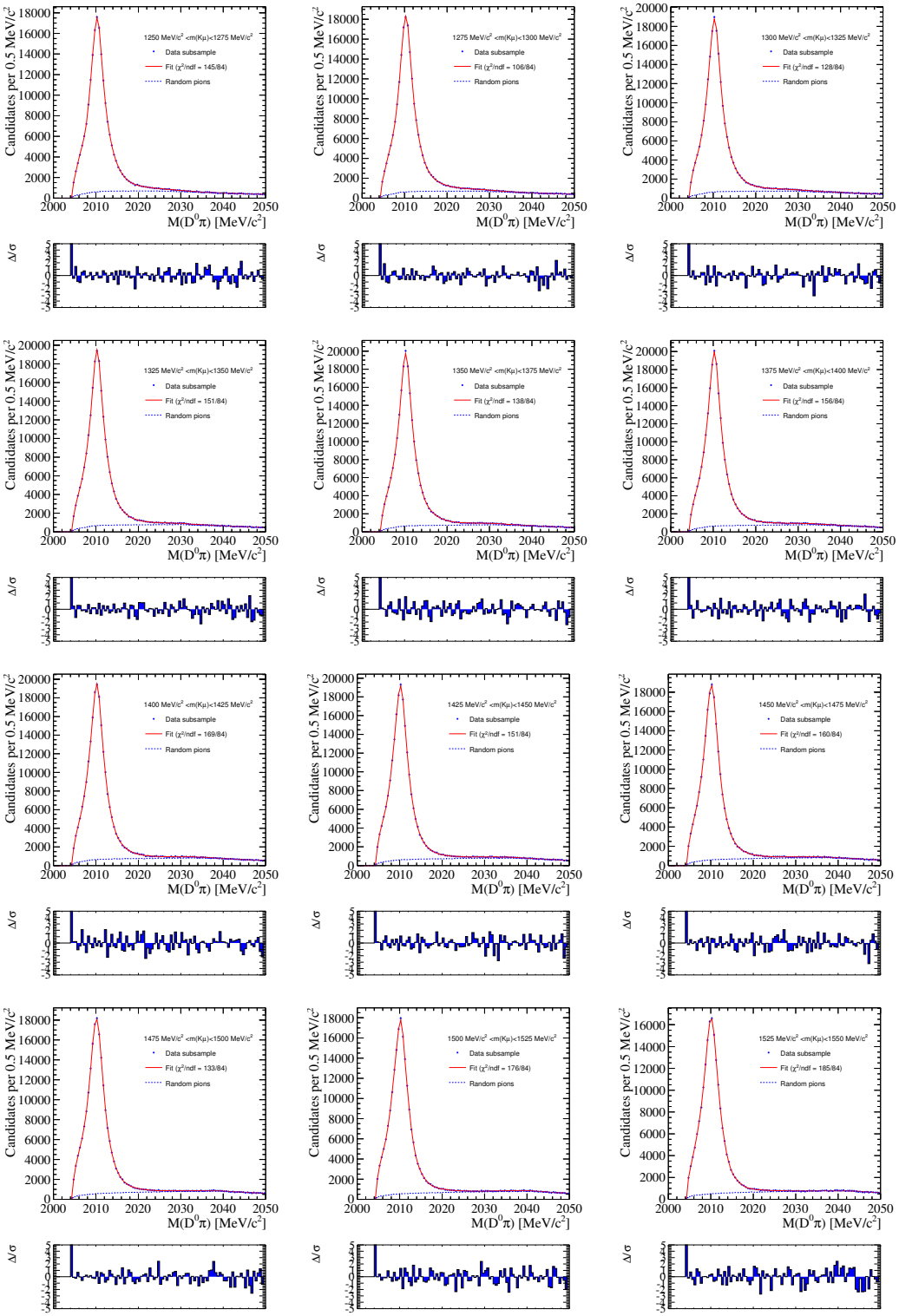


Figure 56: Fits in bins of $m(K\mu)$ RS sample (part 3 of 4)

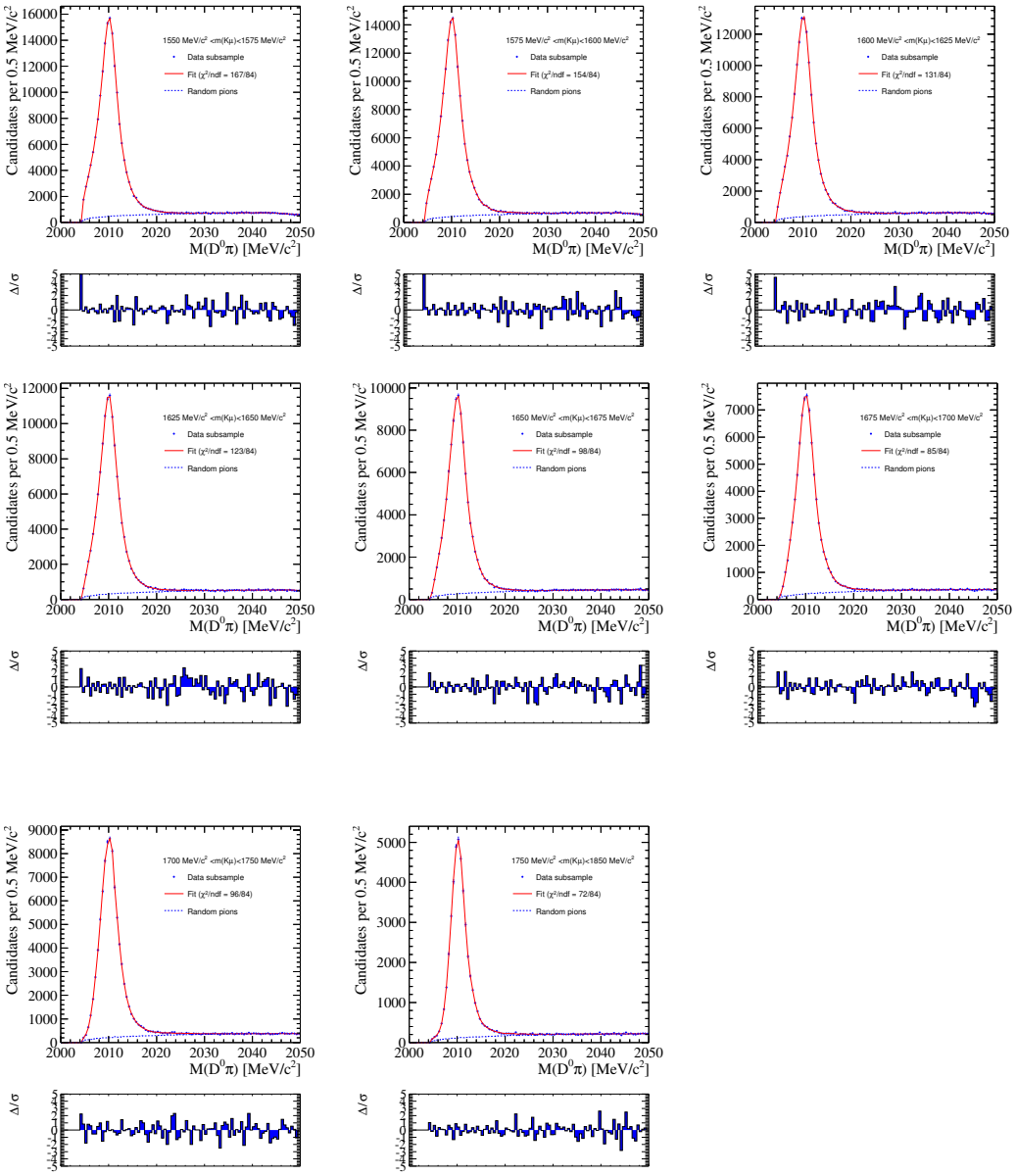


Figure 57: Fits in bins of m(Kμ) RS sample (part 4 of 4)

C.2 WRONG SIGN SAMPLE

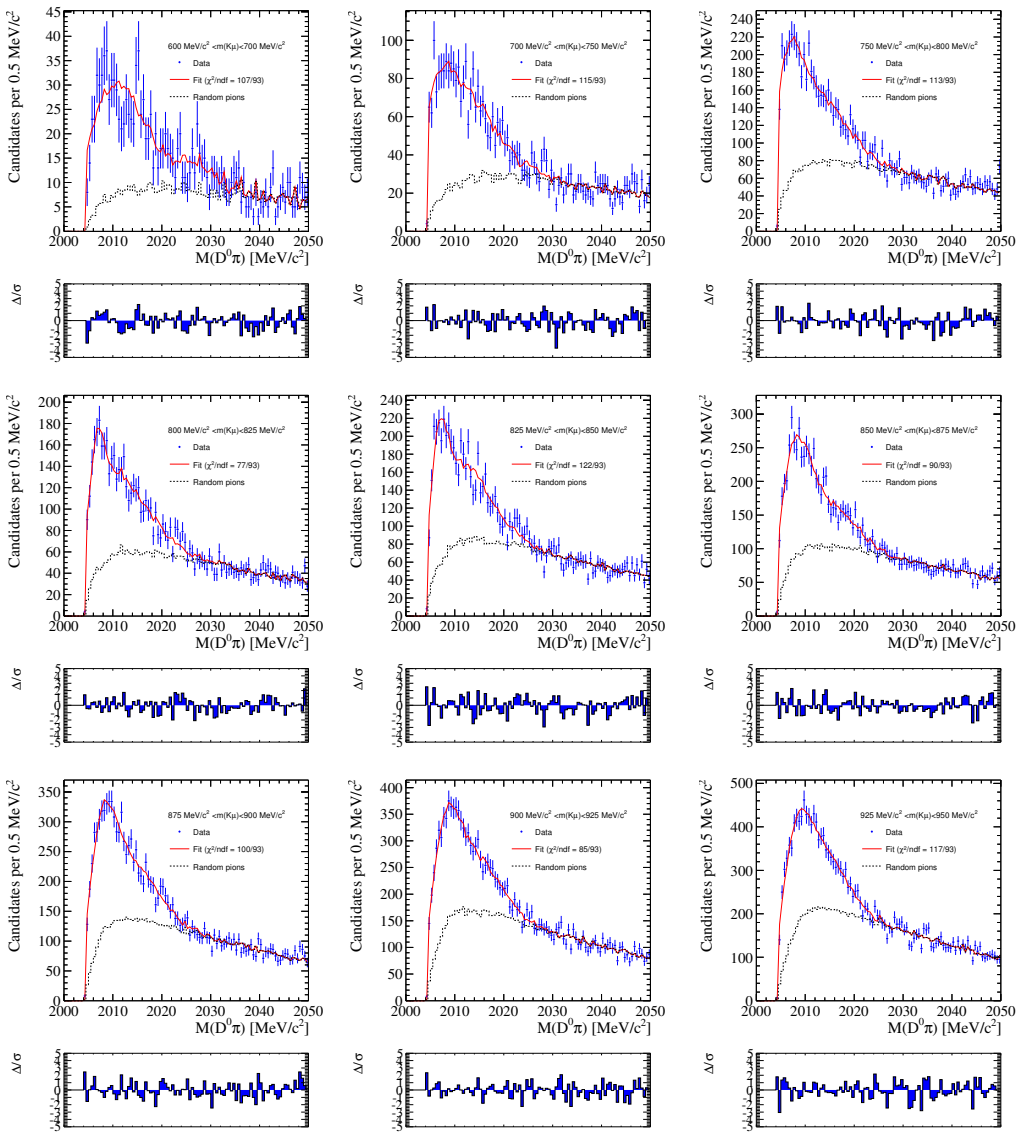


Figure 58: Fits in bins of $m(K\mu)$ WS sample (part 1 of 4)

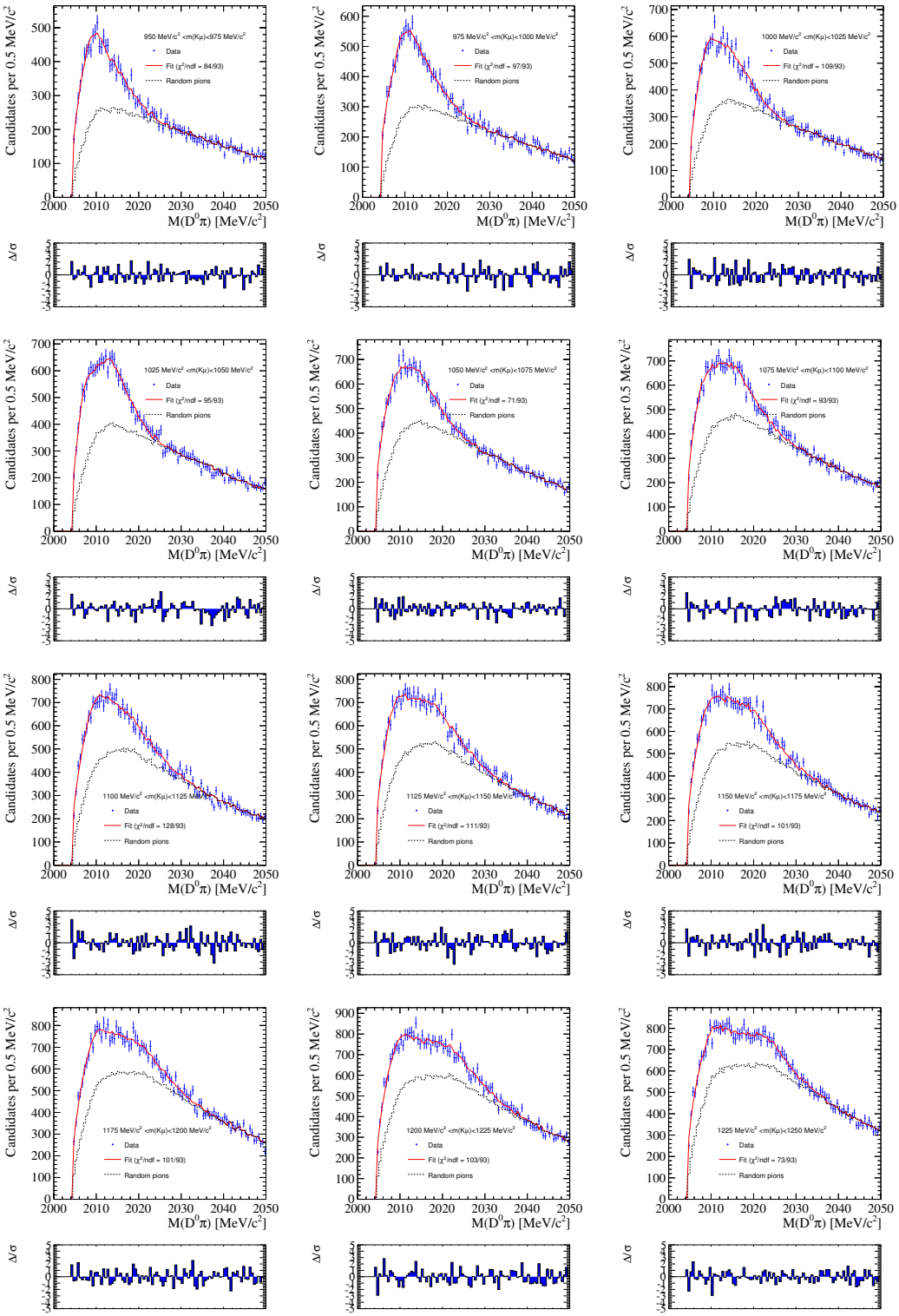


Figure 59: Fits in bins of m(Kμ) WS sample (part 2 of 4)

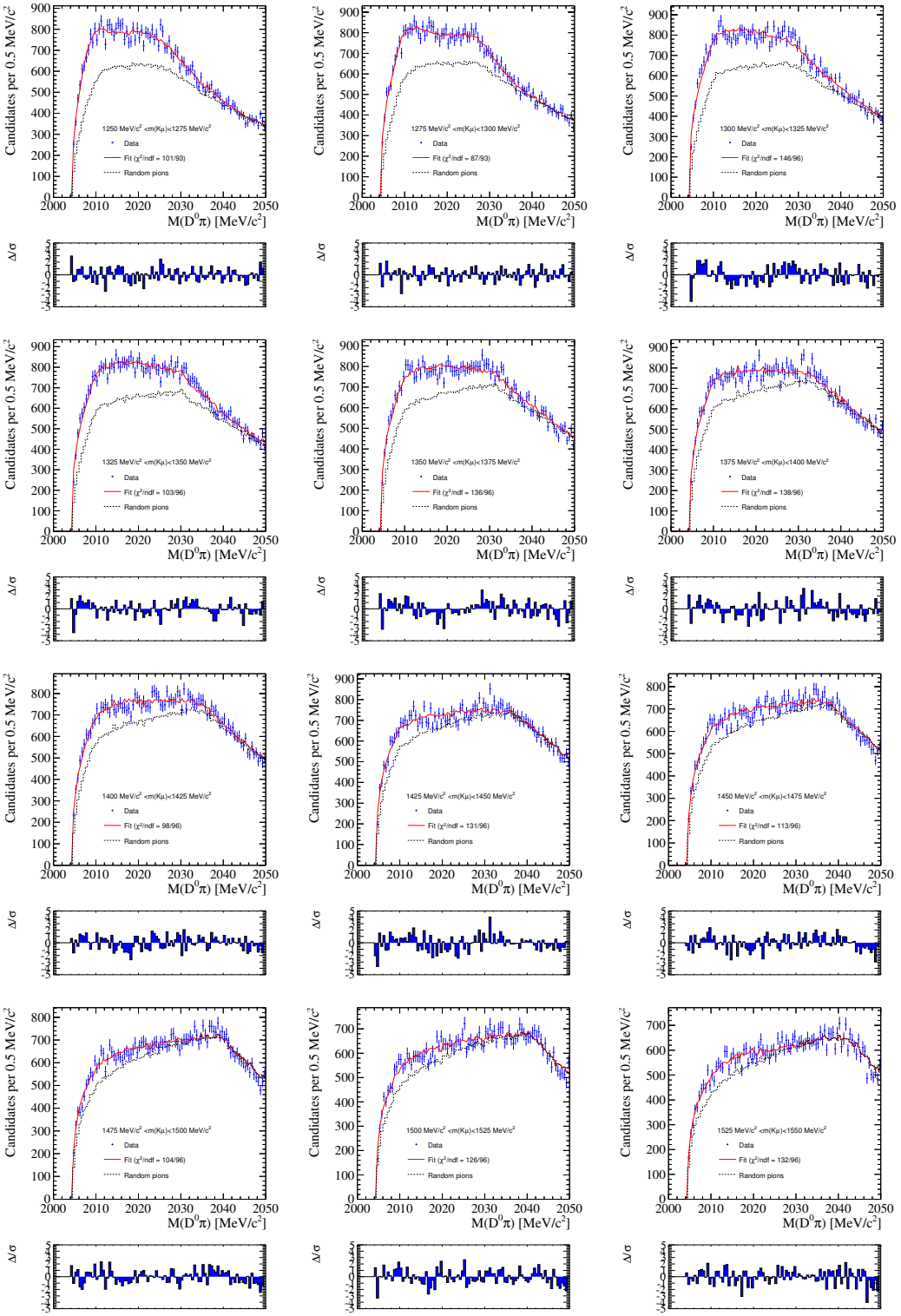


Figure 60: Fits in bins of $m(K\mu)$ WS sample (part 3 of 4)

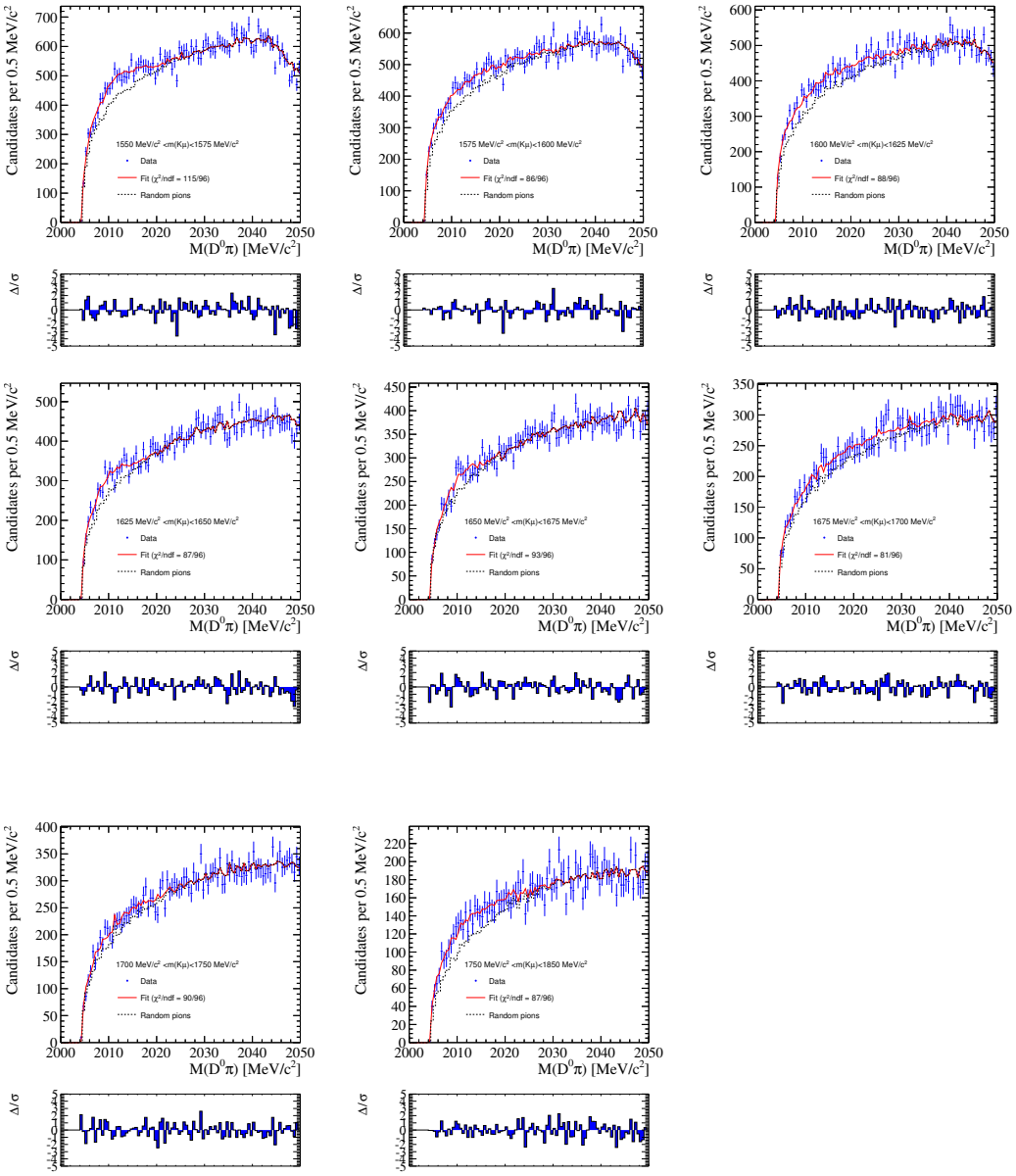


Figure 61: Fits in bins of m(Kμ) WS sample (part 4 of 4)

BIBLIOGRAPHY

[1] M. Gell-Mann and A. Pais. Behavior of neutral particles under charge conjugation. *Phys. Rev.*, 97:1387–1389, 1955.

[2] G. Luders. On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories. *Kong.Dan.Vid.Sel.Mat.Fys.Med.*, 28N5:1–17, 1954.

[3] C.A. Chavez, R.F. Cowan, and W.S. Lockman. Charm meson mixing: An experimental review. *Int.J.Mod.Phys.*, A27:1230019, 2012.

[4] K. Lande, E. T. Booth, J. Impeduglia, L. M. Lederman, and W. Chinowsky. Observation of long-lived neutral ν particles. *Phys. Rev.*, 103:1901–1904, 1956.

[5] H. Albrecht et al. (ARGUS collaboration). Observation of B^0-$\bar{B}^0$ mixing. *Physics Letters B*, 192(1-2):245 – 252, 1987.

[6] A. Abulencia et al. (CDF Collaboration). Observation of B_s^0 - $\bar{B}_s^0$ Oscillations. *Phys. Rev. Lett.*, 97:242003, 2006.

[7] B. Aubert et al. (BaBar Collaboration). Evidence for D^0-$\overline{D}^0$ Mixing. *Phys.Rev.Lett.*, 98:211802, 2007.

[8] M. Staric et al. (Belle Collaboration). Evidence for D^0-$\overline{D}^0$ Mixing. *Phys.Rev.Lett.*, 98:211803, 2007.

[9] T. Aaltonen et al. (CDF Collaboration). Evidence for D^0-$\overline{D}^0$ mixing using the CDF II Detector. *Phys.Rev.Lett.*, 100:121802, 2008.

[10] R. et al. (LHCb collaboration) Aaij. Observation of D^0–$\overline{D}^0$ oscillations. *Phys. Rev. Lett.*, 110:101802, 2013.

[11] M. Kobayashi and T. Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog.Theor.Phys.*, 49:652–657, 1973.

[12] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, 1963.

[13] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak interactions with lepton-hadron symmetry. *Phys. Rev. D*, 2:1285–1292, 1970.

[14] E. Golowich and A.A. Petrov. Short distance analysis of D^0 - $\overline{D}^0$ mixing. *Phys.Lett.*, B625:53–62, 2005.

[15] H. Georgi. D-$\bar{D}$ mixing in heavy quark effective field theory. *Physics Letters B*, 297(3-4):353 – 357, 1992.

[16] T. Ohl, G. Ricciardi, and E. H. Simmons. D^0-$\overline{D}^0$ mixing in heavy quark effective field theory: The Sequel. *Nucl.Phys.*, B403:605–632, 1993.

[17] I.I.Y. Bigi and N.G. Uraltsev. D^0-$\overline{D}^0$ oscillations as a probe of quark hadron duality. *Nucl.Phys.*, B592:92–106, 2001.

[18] A. F. Falk, Y. Grossman, Z. Ligeti, Y. Nir, and A. A. Petrov. The D^0-$\overline{D}^0$ mass difference from a dispersion relation. *Phys.Rev.*, D69:114021, 2004.

[19] A. F. Falk, Y. Grossman, Z. Ligeti, and A. A. Petrov. SU(3) breaking and D^0-$\overline{D}^0$ mixing. *Phys.Rev.*, D65:054034, 2002.

[20] A. A Petrov. Charm mixing in the Standard Model and beyond. *Int.J.Mod.Phys.*, A21:5686–5693, 2006.

[21] S. Bianco, F.L. Fabbri, D. Benson, and I. Bigi. A Cicerone for the physics of charm. *Riv.Nuovo Cim.*, 26N7:1–200, 2003.

[22] P.F. Harrison and H.R. Quinn. The BABAR physics book: Physics at an asymmetric B factory. 1998.

[23] Charm Physics sub group. Hfag, 2008. URL http://www.slac.stanford.edu/xorg/hfag/charm/FPCP14/results_mixing.html.

[24] R. Aaij et al. (LHCb collaboration). Prompt charm production in pp collisions at $\sqrt{s} = 7$ TeV. *Nucl.Phys.*, B871:1–20, 2013.

[25] K.A. Olive and Particle Data Group. Review of particle physics. *Chinese Physics C*, 38(9):090001, 2014.

[26] R. Aaij et al. (LHCb collaboration). Evidence for CP violation in time-integrated $D^0 \to h^- h^+$ decay rates. *Phys.Rev.Lett.*, 108:111602, 2012.

[27] E.M. Aitala et al. (E791 Collaboration). Search for D^0-$\overline{D}^0$ mixing in semileptonic decay modes. *Phys.Rev.Lett.*, 77:2384–2387, 1996.

[28] C. Cawlfield et al. (CLEO Collaboration). Limits on neutral D mixing in semileptonic decays. *Phys.Rev.*, D71:077101, 2005.

[29] B. Aubert et al. (BaBar Collaboration). Search for D^0-$\overline{D}^0$ mixing using semileptonic decay modes. *Phys.Rev.*, D70:091102, 2004.

[30] B. Aubert et al. (BaBar Collaboration). Search for D^0-$\overline{D}^0$ mixing using doubly flavor tagged semileptonic decay modes. *Phys.Rev.*, D76:014018, 2007.

[31] U. Bitenc et al. (BELLE Collaboration). Improved search for D^0-$\overline{D}^0$ mixing using semileptonic decays at Belle. *Phys.Rev.*, D77:112003, 2008.

[32] O.S. Brüning et al. (LHCb Collaboration). *LHC Design Report*. CERN, 2004.

[33] C. Lefèvre. The CERN accelerator complex. complexe des accélérateurs du CERN. Dec 2008.

[34] D.R. Alves et al. (LHCb Collaboration). The LHCb Detector at the LHC. *J. Instrum.*, 3:S08005, 2008.

[35] R. Aaij et al. (LHCb Collaboration). Observation of Z production in proton-lead collisions at LHCb. *JHEP*, 1409:030, 2014.

[36] LHCb Speakers' Bureau. URL http://lhcb.web.cern.ch/lhcb/speakersbureau/excel/default.html.

[37] S. Amato et al. (LHCb Collaboration). *LHCb magnet: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2000.

[38] P. R. Barbosa-Marinho et al. (LHCb Collaboration). *LHCb VELO (VErtex LOcator): Technical Design Report*. Technical Design Report LHCb. CERN, 2001.

[39] Barbosa-Marinho et al. (LHCb Collaboration). *LHCb inner tracker: Technical Design Report*. Technical Design Report LHCb. CERN, 2002.

[40] P. R. Barbosa-Marinho et al. (LHCb Collaboration). *LHCb outer tracker: Technical Design Report*. Technical Design Report LHCb. CERN, 2001.

[41] S. Amato et al. (LHCb Collaboration). *LHCb RICH: Technical Design Report*. Technical Design Report LHCb. CERN, 2000.

[42] S. Amato et al. *LHCb calorimeters: Technical Design Report*. Technical Design Report LHCb. CERN, 2000.

[43] P. R. Barbosa-Marinho et al. (LHCb Collaboration). *LHCb muon system: Technical Design Report*. Technical Design Report LHCb. CERN, 2001.

[44] LHCb Collaboration. *LHCb trigger system: Technical Design Report*. Technical Design Report LHCb. CERN, 2003.

[45] LHCb Collaboration. LHCb PID Upgrade Technical Design Report. Technical report, CERN, 2013.

[46] K. Kodama et al. (E653 Collaboration). Limits for four and five prong semimuonic charm meson decays. *Phys.Lett.*, B313:260–266, 1993.

[47] W.E. Johns. *Measurements of the semileptonic decay of the neutral charmed meson* $D^0 \to K^- \mu^+$ *muon-neutrino*. PhD thesis, University of Colorado, 1995.

[48] G.D. McGregor and G. Lafferty. *Calibration of the LHCb VELO Detector and Study of the Decay Mode* $D^0 \to K^- \mu^+ \nu_\mu$. PhD thesis, Manchester U., 2011.

[49] J. Therhaag. TMVA Toolkit for multivariate data analysis in ROOT. *PoS*, ICHEP2010:510, 2010.

[50] L. Breiman, J.H. Friedman, R.A. Olshen, and C.J. Stone. *Classification and Regression Trees*. Belmont, CA, 1984.

[51] M. Pivk and F.R. Le Diberder. sPlot: a statistical tool to unfold data distributions. *Nucl. Instrum. Meth.*, A555:356, 2005.

[52] G. Punzi. Sensitivity of searches for new signals and its optimization. *eConf*, C030908:MODT002, 2003.

[53] F. Archilli et al. (LHCb MuonID group). Performance of the Muon Identification at LHCb. *J. Instrum.*, 8:P10020. 17 p, 2013.

[54] T. Sjöstrand, S. Mrenna, and P. Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications* , 178(11):852 – 867, 2008.

[55] Y. Freund and R.E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119 – 139, 1997.

[56] LHCb Collaboration. The Gaudi project, . http://proj-gaudi.web.cern.ch/proj-gaudi/.

[57] LHCb Collaboration. The Gauss project, . http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/.

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

[59] LHCb Collaboration. The Boole project, . http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/.

[60] LHCb Collaboration. The Moore project, . http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/.

[61] LHCb Collaboration. The Brunel project, . http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/.

[62] LHCb Collaboration. The daVinci project, . http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/.

[63] D.J. Lange. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 462(1-2):152 – 155, 2001.

ERKLÄRUNG

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den .....