

# DISSERTATION

submitted to the  
Combined Faculties for the Natural Sciences and Mathematics  
of the Ruperto-Carola-University of Heidelberg, Germany  
for the degree of

**Doctor of Natural Sciences**

put forward by

**Lucia Grillo**

born in Udine, Italy

Oral examination: 22.07.2015



Measurement of  $CP$  violation in  
 $B^0 - \overline{B}^0$  mixing using  
semileptonic decays at LHCb

Referees:

Prof. Dr. Ulrich Uwer

Prof. Dr. Norbert Herrmann



## Abstract

This thesis presents the measurement of the  $CP$ -violating asymmetry in  $B^0 - \bar{B}^0$  mixing using data corresponding to an integrated luminosity of  $3.0 \text{ fb}^{-1}$  collected at the LHCb experiment in proton-proton collisions at the center-of-mass energies of 7 TeV and 8 TeV. This analysis uses untagged, semileptonic  $B^0 \rightarrow D^- \mu^+ \nu_\mu$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$  decays, where the  $D^-$  decays into  $K^+ \pi^- \pi^-$ , and the  $D^{*-}$  decays into  $\bar{D}^0 (\rightarrow K^+ \pi^-) \pi^-$ . The neutrino in the semileptonic  $B$  decays is not reconstructed. A decay time dependent fit allows to disentangle the  $CP$  asymmetry from the possible  $B^0 - \bar{B}^0$  production asymmetry. Detection and reconstruction asymmetries are calibrated using promptly produced Cabibbo-favored  $D^+$  decays, and inclusive secondary  $J/\psi$  decays. The  $CP$ -violating asymmetry is measured to be

$$a_{\text{sl}}^d = (-0.02 \pm 0.19 \text{ (stat)} \pm 0.30 \text{ (syst)})\% .$$

This result is consistent with the Standard Model prediction, and it is the most precise measurement from a single experiment to date. This measurement is published in *Physical Review Letters* [1].

## Kurzfassung

In der vorliegenden Arbeit wird die Messung der  $CP$ -Asymmetrie in der  $B^0 - \bar{B}^0$  Mischung vorgestellt. Die benutzten Daten wurden bei Proton-Proton Kollisionen mit Schwerpunktsenergien von 7 TeV und 8 TeV mit dem LHCb-Experiment aufgenommen und entsprechen einer integrierten Luminosität von  $3.0 \text{ fb}^{-1}$ . Diese Analyse verwendet untagged, semileptonische  $B^0 \rightarrow D^- \mu^+ \nu_\mu$  und  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$  Zerfälle, in denen das  $D^-$  in  $K^+ \pi^- \pi^-$  zerfällt bzw. das  $D^{*-}$  in  $\bar{D}^0 (\rightarrow K^+ \pi^-) \pi^-$  zerfällt. Das Neutrino in den semileptonischen  $B$  Zerfällen wird nicht rekonstruiert. Mit einem Zerfallszeit abhängigen Fit kann die  $CP$ -Asymmetrie von einer möglichen  $B^0 - \bar{B}^0$  Produktionsasymmetrie unterschieden werden. Detektions- und Rekonstruktionsasymmetrien werden mit  $D$ -Meson Zerfällen und inklusiven  $J/\psi$  Zerfällen kalibriert. Die gemessene  $CP$ -Asymmetrie ist

$$a_{\text{sl}}^d = (-0.02 \pm 0.19 \text{ (stat)} \pm 0.30 \text{ (syst)})\% .$$

Dieses Ergebnis ist konsistent mit der Standard-Modell-Vorhersage. Es stellt die präziseste Messung dieser Observable dar. Diese Messung wurde in *Physical Review Letters* [1] veröffentlicht.



# Contents

<b>1</b>	<b>What's the matter?</b>	<b>1</b>
<b>2</b>	<b>Neutral meson mixing and <math>CP</math> violation</b>	<b>5</b>
2.1	Standard Model and searches for New Physics . . . . .	5
2.1.1	Elementary particles . . . . .	6
2.1.2	Fundamental interactions . . . . .	8
2.2	Flavor in the Standard Model . . . . .	11
2.3	Neutral meson mixing . . . . .	13
2.4	$CP$ violation . . . . .	18
2.5	$CP$ violation in mixing . . . . .	19
2.6	$B^0 - \overline{B}^0$ system . . . . .	23
2.6.1	Theoretical predictions . . . . .	23
2.6.2	Experimental status . . . . .	24
2.6.3	Bounds on new physics . . . . .	26
2.7	Summary . . . . .	28
<b>3</b>	<b>The LHCb experiment</b>	<b>31</b>
3.1	The Large Hadron Collider . . . . .	31
3.1.1	LHC design . . . . .	31
3.1.2	LHC performance . . . . .	33
3.1.3	Interaction Point 8 . . . . .	34
3.1.4	$b$ quark production in $pp$ collisions . . . . .	35
3.1.5	$B$ mesons production in $pp$ collisions . . . . .	36
3.2	The LHCb experiment . . . . .	39
3.2.1	Track reconstruction system . . . . .	39
3.2.2	Particle identification system . . . . .	46
3.2.3	Trigger system . . . . .	51
3.3	Event Reconstruction . . . . .	53
3.3.1	Track reconstruction . . . . .	53
3.3.2	Particle Identification . . . . .	54

<b>4</b>	<b>Measurement strategy</b>	<b>57</b>
4.1	Methods to measure $a_{\text{sl}}^d$ and LHCb strategy . . . . .	57
4.1.1	Untagged asymmetry . . . . .	58
4.1.2	$B^0$ - $\bar{B}^0$ production asymmetry . . . . .	61
4.1.3	Detection asymmetries . . . . .	63
4.1.4	Time-dependent fit strategy . . . . .	65
4.1.5	Final result determination . . . . .	67
4.2	Summary . . . . .	68
<b>5</b>	<b>Samples and Selection</b>	<b>71</b>
5.1	Signal data samples . . . . .	71
5.1.1	Signal decay topology . . . . .	72
5.1.2	Trigger selection . . . . .	72
5.1.3	Decay chain reconstruction: Decay Tree Fitter . . . . .	75
5.1.4	Central offline selection (“Stripping”) . . . . .	76
5.1.5	Offline selection . . . . .	76
5.1.6	Event yields . . . . .	79
5.2	Monte Carlo simulation . . . . .	83
5.2.1	Signal Samples . . . . .	84
5.2.2	$B^+$ decays samples . . . . .	85
5.2.3	Inclusive $D$ samples . . . . .	85
5.3	$B^+$ fraction estimation . . . . .	88
5.3.1	$B^+$ fraction estimated from MC . . . . .	89
5.3.2	$B^+$ fraction estimated from data . . . . .	90
5.4	Control Samples . . . . .	93
5.5	Summary . . . . .	96
<b>6</b>	<b>Detection asymmetries</b>	<b>97</b>
6.1	Interaction of particles with matter . . . . .	97
6.2	Hadronic interaction asymmetries . . . . .	99
6.3	Detector, trigger, reconstruction effects . . . . .	99
6.4	Detection asymmetries for $a_{\text{sl}}^d$ . . . . .	102
6.5	$\mu - \pi$ detection asymmetry . . . . .	102
6.5.1	$\mu - \pi$ tracking asymmetry . . . . .	103
6.5.2	Muon mis-identification and trigger asymmetry . . . . .	111
6.5.3	Pion mis-identification asymmetry . . . . .	112
6.6	$K - \pi$ asymmetry . . . . .	112
6.7	Input values for $a_{\text{sl}}^d$ measurement . . . . .	114
<b>7</b>	<b>Decay time description</b>	<b>117</b>
7.1	Partially reconstructed decays . . . . .	117
7.1.1	Reconstruction of the missing momentum . . . . .	118
7.1.2	Decay time resolution . . . . .	123

7.2	Signal decay time model . . . . .	129
7.2.1	Decay time resolution . . . . .	129
7.2.2	Decay time acceptance . . . . .	131
7.3	Summary . . . . .	134
<b>8</b>	<b>Time-dependent fit</b>	<b>135</b>
8.1	Maximum likelihood fits . . . . .	135
8.1.1	Weighted datasets . . . . .	137
8.1.2	Quality of a binned maximum likelihood fit . . . . .	138
8.2	Determination of $a_{\text{sl}}^d$ and $A_{\text{P}}$ . . . . .	140
8.3	Fit to the decay time . . . . .	142
8.3.1	Signal decays . . . . .	142
8.3.2	$B^+$ decays . . . . .	144
8.3.3	$D^\pm/D^0$ mass sidebands decays . . . . .	145
8.4	Fit to the $K^+\pi^-\pi^-/K^+\pi^-$ invariant mass . . . . .	147
8.5	Fit validation . . . . .	148
8.6	The $a_{\text{sl}}^d$ time-dependent fit summary . . . . .	150
8.7	Fit results . . . . .	152
<b>9</b>	<b>Systematics and Crosschecks</b>	<b>157</b>
9.1	Systematic uncertainties . . . . .	157
9.1.1	Uncertainty on $A_{\text{D}}$ . . . . .	158
9.1.2	$B^+$ background description . . . . .	159
9.1.3	Correlations crosscheck . . . . .	160
9.1.4	Other backgrounds . . . . .	161
9.1.5	Signal description . . . . .	163
9.2	Summary of the systematic uncertainties . . . . .	167
9.3	Crosschecks . . . . .	167
<b>10</b>	<b>Results and Conclusions</b>	<b>173</b>
<b>A</b>	<b>Analytical form of PDFs and relative normalizations</b>	<b>177</b>
A.1	Decay time PDFs . . . . .	177
A.2	$K^+\pi^-\pi^-/K^+\pi^-$ invariant mass PDFs . . . . .	179
<b>B</b>	<b>Fit projections</b>	<b>181</b>
B.1	$B^0 \rightarrow D^-\mu^+\nu_\mu X$ data sample . . . . .	182
B.2	$B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$ data sample . . . . .	186
<b>C</b>	<b>Crosschecks</b>	<b>191</b>
C.1	Alternative fit parametrization . . . . .	193
C.2	Fiducial regions . . . . .	196
C.3	Data taking periods . . . . .	198
C.4	Number of Primary Vertices . . . . .	200

C.5 Random pions background in $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$ sample . . . . .	201
<b>References</b>	<b>205</b>

# Chapter 1

## What's the matter?

*The most exciting phrase to hear in science, the one that heralds new discoveries, is not "Eureka!" but "that's funny!"* (Isaac Asimov)

Observation and investigation of nature have always been among the primary interests of human beings. The study of fundamental particles and interactions has its origin in ancient Greek and Indian philosophies, and developed over the centuries. During the last decades, the Standard Model (SM) of particle physics, which describes our current knowledge of nature, has been probed by a broad variety of experiments, and it successfully explains most of the physical phenomena so far observed. The SM provides a mathematical description of the three fundamental interactions, strong, weak and electromagnetic, that act among the elementary particles, quarks and leptons, which are classified in three families. A well known example of "Eureka!" in the case of the SM is the recent discovery of the Higgs boson [2], [3]. However, just as in the ancient times, there are still "funny" phenomena, which are not completely understood in our theoretical description, and are the subjects of today's searches for SM-extensions in particle physics. For example: How was the preference of matter over antimatter produced in the universe?

One of the conditions proposed to explain a different production rates of matter compared to antimatter is the violation of the  $CP$  symmetry [4]. This asymmetry consists of a different behavior of particles and antiparticles. The phenomenon is unique to the weak interaction, and it was observed for the first time in 1964 in the neutral kaon system [5]. The SM has sufficient complexity to accommodate  $CP$ -violating effects, in the quark-mixing parameters of the Cabibbo-Kobayashi-Maskawa (CKM) matrix, but does not predict  $CP$  violation sufficient to explain the baryon asymmetry of the universe. In the past two decades, several precise measurements of strange and beauty hadron decays, along with accurate theoretical predictions, allowed a redundant determination of the CKM parameters. So far all the measurements performed are consistently described by few CKM parameters [6], and are consistent with SM predictions. Nevertheless small effects of New Physics and new additional sources of  $CP$  violation are not yet ruled out by the current experimental precision.

Searches for new particles can either exploit their direct production in high-energy collisions, as in the case of the Higgs boson discovery by the ATLAS and CMS experiments at

the Large Hadron Collider (LHC), or processes affected by quantum corrections. In this complementary approach precision measurements of processes sensitive to potential contributions of unknown particles are necessary. The LHCb experiment at the LHC is designed to measure such processes in charm and beauty hadron decays.

In this thesis  $B^0$  decays with a lepton and a neutrino in the final state, called semileptonic decays, collected by the LHCb experiment during Run-I are used to measure the  $CP$  asymmetry in the flavor oscillations between  $B^0$  and  $\bar{B}^0$  mesons. Neutral  $B^0$  mesons are bound states of a  $d$  quark and a  $\bar{b}$  antiquark, and the corresponding antiparticles,  $\bar{B}^0$ , are bound states of a  $\bar{d}$  antiquark and a  $b$  quark. Meson-antimeson oscillations are a good example of processes sensitive to the contributions of new particles. These oscillations in the  $B^0$  system, observed for the first time in 1987 [7], are a well established phenomenon. The goal of the analysis presented in this thesis is to determine the  $a_{\text{sl}}^d$  parameter, that measures the asymmetry between the probability of a  $\bar{B}^0$  meson to oscillate into a  $B^0$  meson and the probability of a  $B^0$  meson to oscillate into a  $\bar{B}^0$  meson. The SM prediction of  $CP$  violation in this process is extremely small and precisely determined, therefore this measurement represents an excellent test of the SM. Moreover, among the previous determinations of this  $CP$  asymmetry, the measurement by the DØ experiment [8], which has a  $3\sigma$  tension with the SM prediction, is one of the “that’s funny!” experimental results. It demands further investigation to clarify the  $B^0$  picture.

The first measurement of  $a_{\text{sl}}^d$  performed by the LHCb experiment is described in this thesis. This measurement is challenging, given the proton-proton initial state, as it might produce an asymmetry in the number of  $B^0$  and  $\bar{B}^0$  produced that could be wrongly interpreted as a sign of  $CP$  violation. To disentangle this production asymmetry, the  $a_{\text{sl}}^d$  parameter is extracted from the  $B$  decay-time dependent charge-asymmetry of the products of the  $B^0$  and  $\bar{B}^0$  decays. The development of time-dependent analysis for partially reconstructed decays allowed for this measurement. A crucial input for this measurement is the precise evaluation of the asymmetry in the reconstruction and detection of particles and antiparticles performed at LHCb for  $CP$  asymmetry measurements [9], [10]. The result of this analysis is the most precise measurement of  $a_{\text{sl}}^d$  to date, and it is published in a letter to *Physical Review* [1].

In the following, the organization of the thesis is outlined. The motivation for the measurement is presented in Chapter 2. Chapter 3 briefly describes the LHCb experiment at the LHC, focusing especially on the aspects relevant for this measurement. Once clarified the goal and the given experimental conditions, Chapter 4 provides an overview of the analysis strategy. The reconstruction and selection of the data samples are described in Chapter 5; the samples of simulated events used in the analysis are introduced in the same chapter. Chapter 6 is dedicated to the important inputs for the  $a_{\text{sl}}^d$  analysis, which are the measured values of the detection asymmetries of the particles in the final state. A method to describe the decay time dependent distributions suitable for partially reconstructed decays is described in Chapter 7. The maximum likelihood estimator for measuring the observable of interest is explained in detail in Chapter 8. The validation of the method and relative results are reported in the same chapter. Chapter 9 is devoted to the evaluation of

the systematic uncertainties and summarizes the stability checks performed. In Chapter 10 the final result is presented and the conclusions from this work are drawn.



# Chapter 2

## Neutral meson mixing and $CP$ violation

In this chapter, the physics motivation for the measurement presented in this thesis is explained. The Standard Model of particle physics is introduced, along with basics of the current strategies to search for physics beyond the current affirmed theoretical description. This introduction focuses on the quark sector of the Standard Model, and on the meson mixing dynamics.  $CP$  violation phenomena are introduced and most of the attention is devoted to  $CP$  violation in the  $B^0 - \bar{B}^0$  system. The experimental landscape of the measurements of  $CP$  violation in  $B^0 - \bar{B}^0$  mixing is presented and compared to the theoretical predictions. The possibility of New Physics observations is discussed.

### 2.1 Standard Model and searches for New Physics

The Standard Model (SM) of particle physics provides a mathematical description of three fundamental interactions, namely the strong, weak, and electromagnetic interactions, that act among elementary particles. It is the unification of two theories: the *Glashow-Weinberg-Salam* (GWS) model (also known as *Electroweak theory*) describing the electroweak interactions, and *Quantum Chromo Dynamics* (QCD) describing the strong interactions.

During the last decades, the SM has been probed by a broad variety of experiments. The recent discovery of the Higgs boson announced in July 2012 [2], [3] completes the validation of the SM, which successfully explains most of the physical phenomena so far observed in nature. On the other hand, it is well known that the SM needs to be extended for the following reasons. Classical gravity, well described by general relativity, should break down at the energy scales close to the Planck scale ( $10^{19}$  GeV), where effects of quantum gravity may appear [11]. At that energy scale the SM should then be replaced by a theory including quantum gravity. Another conceptual problem is the value of the Higgs mass. In the current model, the Higgs mass receives large quantum corrections, that

would result in a larger value, compared to the measured Higgs mass, unless there is a so-called *fine-tuning* cancellation of the radiative corrections.

The energy scale above which the model must be replaced by a more fundamental theory is so called *cut-off* of the effective theory. In the current theory description, either a fine-tuning of the model parameters could bring the cut-off to the Planck scale, or alternatively new particles generated from new symmetry principles could be present between the Planck scale and the electroweak scale [12]. The question whether new particles are present in the energy range from  $\mathcal{O}(1)$  TeV to the Planck scale is therefore still open. Moreover experimental results, such as the observation that 84% of the mass of the Universe is made of non-baryonic and non-luminous matter, are not explained by the current theoretical description.

A first strategy to search for new particles exploits their *direct* production in high-energy collisions. In these direct searches the crucial quantity is the center-of-mass energy of the collisions: the higher the energy, the heavier the particles that can be produced. The production rate of the particles and the background given by the SM processes with signatures similar to the new physics process of interest also have an impact on these studies. The relevant quantities depend on the New Physics scale to be probed.

A complementary approach is provided by *indirect* searches. In this case the exchange of virtual new particles within SM processes is probed. The presence of new particles can then be inferred from deviations of the measured values of observed quantities from the SM predictions. As an example, a decay is predicted within the SM to occur with a certain branching ratio. Then a measurement of a branching ratio significantly different can indicate the presence of a new process. The quantum corrections to SM processes are smaller the heavier the particles involved. For this reason the key ingredient for this second type of searches is the precision. Higher New Physics scales are explored by increasing the precision of the measurements while controlling at sufficiently accurate level the SM contributions and the experimental effects.

Flavor physics is an excellent candidate for indirect searches. A wide variety of measurable processes precisely predicted by the theory allow to determine with complementary approaches the parameters left free by the SM.

The introduction to the Standard Model of particle physics reported in these pages is by no means complete. More detailed descriptions are given in Ref. [13], [14], Ref. [15] focuses on QCD, while Ref. [16] devotes more attention to the electroweak sector.

### 2.1.1 Elementary particles

The elementary constituents of the SM are either *fermions*, particles with spin half-integer, or *bosons*, particles with integer spin. Fermions are the constituents of all visible matter of the Universe within the SM framework. They are described in terms of fields, and are divided in two categories, *quarks* and *leptons*. Bosons are the force-carriers, the particles that are exchanged between the fermions and amongst each other.

Fermions						
Quarks			Leptons			
Gen.	Type	Mass	Electric Charge	Type	Mass	Electric Charge
1 <sup>st</sup>	$u$	$2.3_{-0.5}^{+0.7}$ MeV/ $c$	$+2/3e$	$\nu_e$	$<2\text{eV}/c^2$	0
	$d$	$4.8_{-0.3}^{+0.5}$ MeV/ $c$	$-1/3e$	$e$	0.511 MeV/ $c$	$-e$
2 <sup>nd</sup>	$c$	$1.275 \pm 0.025$ GeV/ $c$	$+2/3e$	$\nu_\mu$	$<190\text{keV}/c^2$	0
	$s$	$95 \pm 5$ MeV/ $c$	$-1/3e$	$\mu$	105.7 MeV/ $c$	$-e$
3 <sup>rd</sup>	$t$	$173.21 \pm 0.87$ GeV/ $c$	$+2/3e$	$\nu_\tau$	$<18.2$ MeV/ $c$	0
	$b$	$4.18 \pm 0.03$ GeV/ $c$	$-1/3e$	$\tau$	$1776.82 \pm 0.16$ MeV/ $c$	$-e$

Table 2.1: Measured masses [17] and electrical charges of the SM fermions.

The SM includes six different quarks: *up* ( $u$ ), *down* ( $d$ ), *charm* ( $c$ ), *strange* ( $s$ ), *top* ( $t$ ), and *bottom/beauty* ( $b$ ) and the corresponding anti-quarks. They can be classified according to their fractional electric charge as up-type quarks with electric charge  $+2/3$  ( $u,c,t$ ), and down-type quarks with electric charge  $-1/3$  ( $d,s,b$ ). The lightest two quarks are thought to be the elementary constituents<sup>1</sup> of the protons and neutrons in the atomic nuclei of the ordinary matter. Quarks can be sorted in three families, and the order of the families conventionally follows the order of the masses of the quarks, as shown in Tab. 2.1. The masses of the quarks are free parameters of the Standard Model and range from a few MeV/ $c^2$ , up and down quark, up to about 173 GeV/ $c^2$  for the top quark. The *bottom* ( $b$ ) quark is especially relevant for this thesis. With a mass of about 4.2 GeV/ $c^2$ , it is the second heaviest quark in the SM, the heaviest known to form bound states, and  $b$ -flavored hadrons have a very rich decay structure leading to many different final states. Since its discovery and the discovery of the  $B$  mesons (bound states of a  $b$  antiquark and a lighter quark), studies with particles containing a  $b$  quark allowed the determination of several SM observables. An example it is given in this thesis.

Leptons can also be classified in three families. Each family includes a charged lepton ( $e, \mu, \tau$ ) and its neutral partner ( $\nu_e, \nu_\mu, \nu_\tau$ ), called neutrino. Neutrinos can interact only weakly, while charged leptons can interact also electromagnetically. A difference between the quark sector and the lepton sector of the SM regards the family number conservation, i.e. the conservation of the number of leptons belonging to each family in the processes. Since neutrino oscillations have been observed, neutrinos do have a tiny nonzero mass and the lepton family conservation law is therefore only approximate. This means the conservation laws is violated, although because of the smallness of the neutrino mass, it still holds to a very large degree for interactions containing charged leptons. The situation is very different for the transitions between quarks, as will be shown in Sec.2.2.

<sup>1</sup>Quarks are considered *pointlike* particles, meaning that no sub-structure has been observed so far.

### 2.1.2 Fundamental interactions

The SM is formulated as a quantum field theory. The structure of the model stems from symmetries under transformations of a gauge group, which uniquely determines the interactions and the number of gauge bosons which correspond to the generators of the group. The quarks and the leptons must reside in representations of the gauge group, i.e. quantum fields must be invariant under local transformations of the a gauge group. In the case of the SM, the symmetry group is

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \quad (2.1)$$

where the subscripts  $C$ ,  $L$  and  $Y$  indicate the color, left handed chirality and weak hypercharge respectively. The force carriers implied by the structure of the model mediate the various interactions: eight massless gluons, each having a different combination of color and anti-color, for the strong interaction; two charged massive bosons,  $W^\pm$ , and a single neutral massive boson,  $Z^0$ , for the weak interaction; and a massless photon,  $\gamma$ , for the electromagnetic interaction. The different behavior of the strong and electroweak interactions as function of the distance derives from the properties of the mediators of the interactions. The presence of the self-interactions of the gluons, while they are not possible in the case of the photons, leads to the different behavior of the strong and electromagnetic interactions. While the strong coupling becomes stronger for larger distances, the electromagnetic force becomes weaker. The short range character of the weak interactions is due to the massive  $W^\pm$  and  $Z^0$  mediators. Finally the SM includes a spin-zero particle, the Higgs boson, that allows for the generation of the particles masses from the spontaneous symmetry breaking of the gauge group of the electroweak interaction.

#### Quantum Chromo Dynamics (QCD)

$SU(3)_C$  is the symmetry group of QCD, the theory describing the strong interactions. The generators of the QCD symmetry group correspond to the eight massless gluons, and the charge to which they couple is the color charge. The color quantum number can assume the values red, green and blue and respective anti-colors. Gluons couple to quarks and to themselves according to their color charge. The self-coupling of the gluons is the source of the short range of the strong interactions.

Notable properties of QDC are *asymptotic freedom* and *confinement*. Both originate from a dependence of the measured strength of the force as function of the transferred four-momentum squared  $Q^2$  among the participants of the interaction. Asymptotic freedom means that the coupling constant  $\alpha_s(Q^2)$  decreases for increasing values of  $Q^2$ , corresponding to short interaction scales, until it vanishes asymptotically. On the other hand, at large distances or small transferred momenta, the coupling becomes strong, leading to non-perturbative phenomena. A consequence of confinement is the impossibility of separating color charges, like individual quarks and gluons. This is a consequence of the QCD interaction potential increasing linearly with the distance, when considering long

distances. As a result, we only can observe *hadrons*, tightly bound composite states of quarks, overall neutral in color charge. The typical length scale of the quark confinement is given by the size of the hadrons  $R_{\text{had}} \sim \hbar c / \Lambda_{QCD} \sim 1$  fm. Hadrons can be distinguished in *mesons*, bound states of a quark and an anti-quark, and *baryons*, bound states of three quarks.

The two regimes of large and small values of  $\alpha_s(Q^2)$  are separated by a reference scale,  $\Lambda_{QCD}$ . Quarks are considered *heavy* when their mass is large in comparison to  $\Lambda_{QCD}$ . In this case  $\alpha_s(m_q)$  is small and perturbative method can be used to make theoretical predictions of related QCD processes. Heavy quarks in the SM are  $c$ ,  $b$  and  $t$ .

### Electroweak interactions

$SU(2)_L \otimes U(1)_Y$  is the gauge group of the electroweak interactions in the SM. The Electroweak theory unifies the electromagnetic and weak interactions. The gauge fields related to the generators of the weak isospin  $SU(2)_L$  group and the generator of the weak hypercharge  $U(1)_Y$  symmetry group, are indicated with  $W_i^\mu$  ( $i = 1, 2, 3$ ) and  $B^\mu$  respectively.

The charges to which these gauge fields couple are the weak isospin  $T$ , with third component  $T_3$  and the weak hypercharge, related to the electric charge  $Q$  via the Gell-Mann-Nishijima formula  $Y = 2(Q - T)$ . The left-handed components of the quark fields carry weak isospin charge ( $T = 1/2$ ) and are represented by doublets of one up-type and the corresponding down-type quark:

$$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \quad (2.2)$$

Analogously the left-handed components of the leptonic fields are represented with the following doublets

$$L_L = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix} \quad (2.3)$$

The right-handed components, singlets with  $T = 0$ , of the fermionic fields exist, an exception is made for the right-handed neutrino, which does not exist in the original formalism of the SM<sup>2</sup>. Table 2.2 reports the SM fermions and their quantum numbers. For antiparticles the right-handed fields are weak isospin doublets and the left-handed fields are the singlets. The  $W_i^\mu$  fields couple only to left-handed particles and right-handed antiparticles, while the  $B^\mu$  field couple to particles carrying weak hypercharge, regardless the weak isospin. In the SM, the electroweak  $SU(2)_L \otimes U(1)_Y$  symmetry is spontaneously broken to  $U(1)_Q$ , as consequence of the so-called *Higgs mechanism* [20], [21]. It introduces the scalar Higgs field  $\phi$ , with a non-zero vacuum expectation value  $v$ . If the potential  $V(\phi)$  is chosen with minimum not symmetric respect to  $\phi$ , mass terms can be generated without breaking gauge invariance.

A first implication of this mechanism is that the fields mediating the electroweak interactions

---

<sup>2</sup>The existence of right-handed neutrinos is required by several SM extensions [18], [19].

Table 2.2: Overview of the fermions predicted by the Standard Model (SM) and their quantum numbers.  $Q$  refers to the electric charge,  $T$  indicates the weak isospin and  $T_3$  shows its third component.  $Q$ ,  $T_3$  and the weak hypercharge  $Y$  are related via the Gell-Mann-Nishijima formula  $Q = T_3 + \frac{Y}{2}$

symbol	Family			$Q$	$T$	$T_3$
<b>quarks</b>						
$Q_L$	$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} c_L \\ s_L \end{pmatrix}$	$\begin{pmatrix} t_L \\ b_L \end{pmatrix}$	$\begin{pmatrix} +2/3 \\ -1/3 \end{pmatrix}$	1/2	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$
$U_R$	$u_R$	$c_R$	$t_R$	+2/3	0	0
$D_R$	$d_R$	$s_R$	$b_R$	-1/3	0	0
<b>leptons</b>						
$L_L$	$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1/3 \end{pmatrix}$	1/2	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$
$E_R$	$e_R$	$\mu_R$	$\tau_R$	-1	0	0

Table 2.3: Fundamental interactions in the SM with relative mediators.

Interaction	Mediator Type	Mediator Mass	Strength
Strong	$g (\times 8)$	0	$\alpha_S \sim \mathcal{O}(1)$ <sup>3</sup>
Electromagnetic	$\gamma$	0	$\alpha_{em} = 1/137$
Weak	$Z^0$	90.2 GeV/ $c$	$G_F \sim 10^{-5} 1/\text{GeV}$
	$W^\pm$	80.4 GeV/ $c$	

are no longer the massless  $W_i^\mu$  and  $B^\mu$ . The mediator of the electromagnetic interaction is the photon  $\gamma$ , a linear combination of  $B^\mu$  and  $W_3^\mu$ , that couples to the particles with a strength according to their electrical charge, and not distinguishing between left-handed or right-handed particles. For the weak interactions the mediators are the massive charged  $W^\pm$  bosons and the neutral  $Z^0$  boson, explaining the short range of the weak interactions. The  $Z^0$  boson is also a linear combination of  $B^\mu$  and  $W_3^\mu$ , it mediates weak interactions so-called *neutral weak currents*.  $W^\pm$  bosons are instead linear combinations of  $W_1^\mu$  and  $W_2^\mu$  bosons, and they mediate the so-called *charged weak currents*. These currents allow transitions between up- and down-type quarks and between charged and neutral leptons. Tab. 2.3 reports the fundamental interactions predicted by the SM, together with their mediators.

A second important implication of the Higgs mechanism is the generation of fermion masses via the Yukawa terms in the Lagrangian.

As third implication, the Higgs mechanism predicts the existence of a massive, spin 0 particle, so-called Higgs boson  $H$ , discovered in 2012 by CMS [3] and ATLAS experiments [2].

<sup>3</sup>The coupling strength depends on the energy. See Sec. 2.1.2.

## 2.2 Flavor in the Standard Model

The spontaneous symmetry breaking of the SM allows the quarks to acquire mass via *Yukawa* interactions with the Higgs field, without breaking gauge invariance:

$$\mathcal{L}_{Yukawa} = Y_d^{ij} \bar{Q}_L^i \phi D_R^j + Y_u^{ij} \bar{Q}_L^i \phi U_R^j + (h.c.) \quad (2.4)$$

With the Higgs field denoted with  $\phi$ , and  $Y_{d,u}^{i,j}$  representing the coupling constants. The mass of the quarks  $m_q$  are related to their coupling to the Higgs field:  $m_q = Y_q \frac{v}{\sqrt{2}}$ . To write proper mass terms for quarks, the  $Y_{d,u}^{i,j}$  matrices need to be diagonalized, that is possible using four independent matrices. Only three of them can be freely chosen (redefining the quark fields with a different phase), therefore if the up-type quarks are diagonalized, the down-type quarks are left non-diagonal. By convention, the interaction eigenstates and the mass eigenstates are chosen to be equal for the up-type quarks, whereas the down-type quarks are chosen to be rotated, going from the flavor (o interaction) basis to the mass basis:

$$Y_u = \mathbb{I} \cdot \begin{pmatrix} u \\ c \\ t \end{pmatrix} ; \quad Y_d = \mathbb{I} \cdot \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

with  $\mathbb{I}$  the identity matrix and  $V_{CKM}$  the Cabibbo-Kobayashi-Maskawa (CKM) matrix, that relates the flavor eigenstates  $(d', s', b')$  to the mass eigenstates  $(d, s, b)$ .

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

The CKM matrix is unitary. The off-diagonal element show a strong hierarchical order:  $|V_{us}|$  and  $|V_{cd}|$  are about 0.22,  $|V_{cb}|$  and  $|V_{ts}|$  of order  $4 \cdot 10^{-2}$  and  $|V_{ub}|$  and  $|V_{td}|$  of order  $5 \cdot 10^{-3}$ . As the matrix is unitary and global phases are not observable, four free parameters remain. Three are the quark mixing angles and one is a complex phase. This complex phase gives rise to *CP violation* in the Standard Model, i.e. the different behavior of particles and anti-particles in the weak interaction. The CKM-matrix can be expressed in terms of  $\lambda = |V_{us}|$ , up to  $\mathcal{O}(\lambda^4)$  terms <sup>4</sup>

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

---

<sup>4</sup> For the *CP* violating measurement in the  $B^0$  sector presented in this thesis it is sufficient to write the CKM-matrix including terms up to  $\mathcal{O}(\lambda^3)$ . For measurements in the  $B_s^0$  sector, it is helpful to include terms up to  $\mathcal{O}(\lambda^5)$ , given that the phase of the matrix element  $V_{ts}$  is playing a role in that case:

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

where  $A$ ,  $\rho$  and  $\eta$  are the remaining real parameters, all of order unity ( $\lambda \approx 0.23, A \approx 0.80, \rho \approx 0.14, \eta \approx 0.34$ ). This form highlights the hierarchy of the CKM matrix elements, which is important to understand the probability of processes involving quarks to occur. The transition probability between an up-type quark  $x$  and a down-type  $y$  quark is indeed proportional to the matrix element  $|V_{xy}|^2$ . This explains, for example, why hadrons with a  $b$  quark component have a long lifetime. As the decay to a top quark is forbidden by energy conservation, since the  $t$  quark mass is larger, the  $b$  quark can only decay to lighter flavors, with transitions that are suppressed by the CKM-elements. This is commonly referred to as *Cabibbo suppression*. The effect is even more pronounced in processes which, for example, involve the transition of a  $b$  to a  $u$  quark.

Expressing the charged weak current part of the Lagrangian (i.e. mediated by  $W^\pm$  gauge bosons only) using the mass eigenstates instead of the flavor eigenstates, the quark mixing between families (i.e. the off-diagonal elements) appears:

$$\mathcal{L}_{cc}(Q_L) = \frac{g}{\sqrt{2}} \bar{u}_L^i \gamma_\mu W^{-\mu} d_L^j + \frac{g}{\sqrt{2}} \bar{d}_L^i \gamma_\mu W^{+\mu} u_L^j + \dots \quad (2.5)$$

$$= \frac{g}{\sqrt{2}} \bar{u}_L^i (V_{CKM}) \gamma_\mu W^{-\mu} d_L^j + \frac{g}{\sqrt{2}} \bar{d}_L^i (V_{CKM}) \gamma_\mu W^{+\mu} u_L^j + \dots \quad (2.6)$$

As a result of the fact that  $V_{CKM}$  is not diagonal, the  $W^\pm$  gauge bosons couple to quark (mass eigenstates) of different generations. For this reason the Flavor Changing Charged Currents (FCCCs) occur at tree-level<sup>5</sup>, while  $b \rightarrow s$  or  $s \rightarrow d$  transitions, namely Flavor Changing Neutral Currents (FCNCs), can occur only at higher than tree-level. The only possibility to study FCNCs is to analyze loop-processes. These are rarer than tree-level transitions, since they involve at least two FCCCs, and one of them is CKM suppressed by at least one power of  $\lambda$ , given that it has to be between two different families.

The unitarity of the CKM-matrix,  $V_{CKM} V_{CKM}^\dagger = \mathbb{I}$ , leads to a set of 9 equations:

$$\sum_{k=u,c,t} V_{ki} V_{kj}^\dagger = \delta_{ij} \quad (i, j = d, s, b) \quad (2.7)$$

When  $i \neq j$ , these relations are known as *orthogonality conditions* or *unitarity triangles* because they define triangles in the complex plane. The area of all these triangles is constant and equals to  $J_{CP}/2$ . The symbol  $J_{CP}$  stands for the *Jarlskog invariant*, a combination of the CKM elements. This quantity also quantifies the possible violation of the charge-parity ( $CP$ ) symmetry in the SM, i.e. the non-invariance of physics processes when all spatial coordinates are inverted and particles are replaced by their antiparticles. The  $CP$  symmetry is violated only if  $J_{CP} \neq 0$ . Any  $CP$ -violating quantity in the SM must be related to  $J_{CP}$ , reflecting the fact that the complex phase, responsible to  $CP$  violation appears in the  $3 \times 3$  CKM-matrix.

One triangular equation of special interest is known as the *unitarity triangle* (UT):

$$V_{ud} V_{ub}^* + V_{cd} V_{cb}^* + V_{td} V_{tb}^* = 0 \quad (2.8)$$

---

<sup>5</sup>Tree-level means first order in perturbation theory.

That can be rewritten in the form:

$$R_t e^{-i\beta} + R_u e^{i\gamma} = 1 \quad (2.9)$$

where

$$R_t = \left| \frac{V_{tb} V_{tb}^*}{V_{cd} V_{cb}^*} \right|, R_u = \left| \frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*} \right|, \beta = \text{arg} \left( - \frac{V_{cd} V_{cb}^*}{V_{td} V_{tb}^*} \right), \gamma = \text{arg} \left( - \frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*} \right) \quad (2.10)$$

$R_t$  and  $R_u$  give two sides of the unitarity triangle, and the third side is given by the unitary vector, while  $\beta$  and  $\gamma$  are two angles, and  $\alpha = \pi - \beta - \gamma = \text{arg}(V_{td} V_{tb}^* / V_{ud} V_{ub}^*)$  represents the third angle of the triangle. Given the unitarity constraint (definition in Eq. 2.9), the unitarity triangle is unambiguously defined by the complex number:

$$\bar{\rho} + i\bar{\eta} = R_u e^{i\gamma} \quad (2.11)$$

which corresponds to the coordinates in the complex plane  $(\bar{\rho}, \bar{\eta})$  of the only non trivial apex of the UT. The parameters  $\bar{\rho}$ , and  $\bar{\eta}$  can be expressed in terms of the Wolfenstein parameters  $\rho$  and  $\eta$ :  $\bar{\rho} = \rho(1 - \lambda^2/2) + \mathcal{O}(\lambda^4)$ ;  $\bar{\eta} = \eta(1 - \lambda^2/2) + \mathcal{O}(\lambda^4)$ . One purpose of flavor physics is to overconstrain the CKM-matrix elements. For this reason it is interesting to represent all the possible observables in function of the two parameters  $\bar{\rho}$  and  $\bar{\eta}$ , as shown in Fig 2.1.

## 2.3 Neutral meson mixing

This section aims to describe the neutral meson mixing, which is unique to the neutral  $K$ ,  $D$  and  $B$  mesons. The focus will be on neutral  $B$  mesons, which are especially relevant for this thesis. A more complete description, can be found in Ref. [23], [24], [25].

In general neutral mesons  $K^0$ ,  $D^0$ ,  $B^0$  (or  $B_d^0$ ),  $B_s^0$  are flavor eigenstates and the corresponding  $\bar{K}^0$ ,  $\bar{D}^0$ ,  $\bar{B}^0$  (or  $\bar{B}_d^0$ ),  $\bar{B}_s^0$  are characterized by opposite quantum numbers

$$\begin{aligned} K^0 &= |\bar{s}d\rangle & D^0 &= |\bar{c}u\rangle & B^0 &= |\bar{b}d\rangle & B_s^0 &= |\bar{b}s\rangle \\ \bar{K}^0 &= |s\bar{d}\rangle & \bar{D}^0 &= |c\bar{u}\rangle & \bar{B}^0 &= |b\bar{d}\rangle & \bar{B}_s^0 &= |b\bar{s}\rangle \end{aligned} \quad (2.12)$$

Neutral  $B$  mesons are bound states of a  $b$  antiquark and a  $d$  or  $s$  quark, they are denoted as  $B^0 = |\bar{b}d\rangle$  and  $B_s^0 = |\bar{b}s\rangle$  (and corresponding antiparticles  $\bar{B}^0 = |b\bar{d}\rangle$  and  $\bar{B}_s^0 = |b\bar{s}\rangle$ ). Given the similar phenomenology described in this paragraph,  $B^0$  and  $B_s^0$  are here indicated with the common notation:  $B_q = |\bar{b}q\rangle$  with  $q = s, d$  and the corresponding antiparticle  $\bar{B}_q = |b\bar{q}\rangle$  with  $q = s, d$ .

It has been known since a long time (e.g. [7]) that due to weak interactions, transitions like  $B_q \rightarrow \bar{B}_q$  are possible in the neutral  $B$ -system. This requires transitions, at quark level, of the form  $b \rightarrow d, s$ . As mentioned in Sec. 2.2, FCNCs are forbidden at tree-level in the SM, but allowed in higher order processes, like loop diagrams. The diagrams responsible for

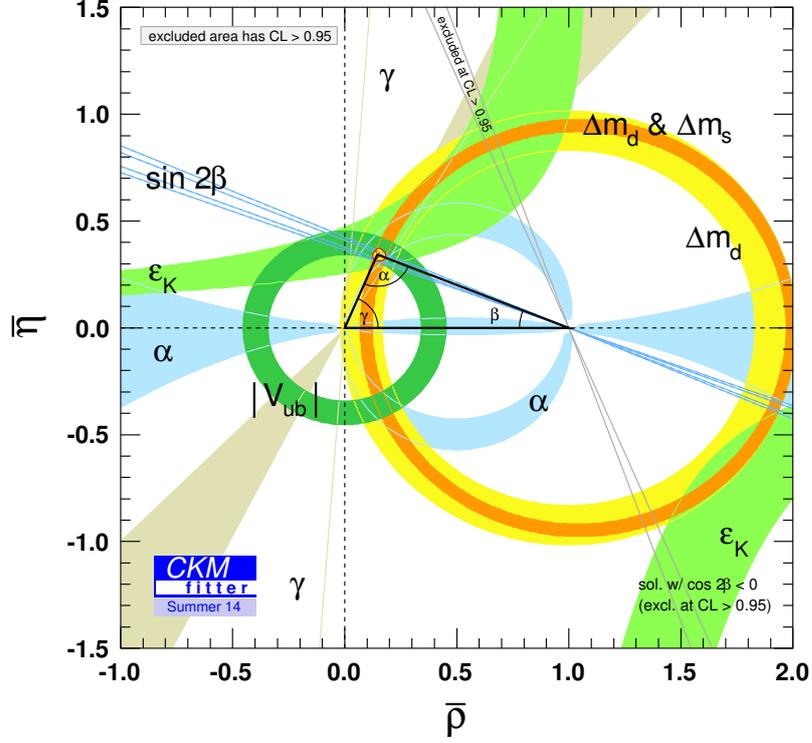


Figure 2.1: Unitarity Triangle: Constraints in the  $(\bar{\rho}, \bar{\eta})$  plane [22].

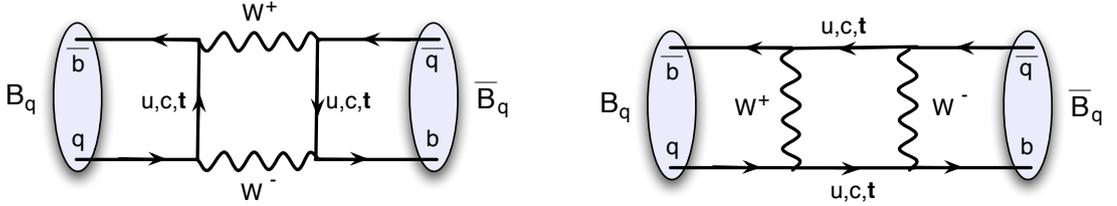


Figure 2.2: Dominant Feynman diagrams responsible for neutral  $B$  meson mixing in the SM.

$B_q \rightarrow \bar{B}_q$  transitions involve the exchange of two  $W$  bosons. They are the so called *box diagrams*, shown in Fig. 2.2.

Due to GIM suppression [26], in these diagrams the leading contribution is given by the top quark. The amplitude of the sum of the diagrams including all the up-type quark contributions to the  $\bar{b} \rightarrow q$  transition, is proportional to:

$$m_u^2 V_{uq} V_{ub}^* + m_c^2 V_{cq} V_{cb}^* + m_t^2 V_{tq} V_{tb}^* \quad (2.13)$$

where with  $m_u, m_c, m_t$ , the mass of the corresponding quarks are represented and  $V_{xy}$  corresponds to the  $x, y$  element of the CKM-matrix. In the case of degenerate quark masses, the amplitude vanishes, since it reduces to one of the orthogonality relations in Eq. 2.7, in the case  $i \neq j$ :

$$V_{uq}V_{ub}^* + V_{cq}V_{cb}^* + V_{tq}V_{tb}^* \quad (2.14)$$

Neglecting the mass difference between the  $u$ -quark mass and the  $c$ -quark mass, and assuming  $m_u^2 \approx m_c^2 \ll m_t^2$ , the amplitude is dominated by the top quark contribution <sup>6</sup>:

$$\mathcal{A} \propto m_u^2 \underbrace{(V_{uq}V_{ub}^* + V_{cq}V_{cb}^* + V_{tq}V_{tb}^*)}_{\approx 0} + \underbrace{(m_t^2 - m_u^2)}_{\approx m_t^2} V_{tq}V_{tb}^* \quad (2.15)$$

$B_q - \overline{B}_q$  oscillations in the SM are allowed though loop processes, therefore they are suppressed at each vertex (See Fig. 2.2) by the coupling constants. In addition the mixing processes are suppressed by a factor  $|V_{tq}|^2$ , that results in a mixing frequency of the  $B_s^0$  mesons much higher than the mixing frequency of the  $B^0$  mesons (See Tab. 2.4).

$B_q$  and  $\overline{B}_q$  are flavor eigenstates of the 2-state system, whose time evolution can be described by means of an effective formalism:

$$i \frac{d}{dt} \begin{pmatrix} |B_q(t)\rangle \\ |\overline{B}_q(t)\rangle \end{pmatrix} = \underbrace{\left( \mathbf{M}_q + i \frac{\mathbf{\Gamma}_q}{2} \right)}_{\mathcal{H}_q^{eff}} \begin{pmatrix} |B_q(t)\rangle \\ |\overline{B}_q(t)\rangle \end{pmatrix} \quad (2.16)$$

$$= \begin{pmatrix} M_{11} - i \frac{\Gamma_{11}}{2} & M_{12} - i \frac{\Gamma_{12}}{2} \\ M_{12}^* - i \frac{\Gamma_{12}^*}{2} & M_{22} - i \frac{\Gamma_{22}}{2} \end{pmatrix} \begin{pmatrix} |B_q(t)\rangle \\ |\overline{B}_q(t)\rangle \end{pmatrix} \quad (2.17)$$

The effective Hamiltonian  $\mathcal{H}_q^{eff}$  describing the process includes two components,  $\mathbf{M}_q = \mathbf{M}_q^\dagger$ , called mass matrix and  $\mathbf{\Gamma}_q = \mathbf{\Gamma}_q^\dagger$ , the decay matrix. The diagonal elements  $M_{11}$  and  $M_{22}$  of the mass matrix are the meson and anti-meson masses  $m_q$  and  $m_{\bar{q}}$ , that are generated from the quark mass terms in the Lagrangian, as explained in Sec. 2.2 and from the strong interaction binding energy within the mesons. The diagonal elements of the decay matrix,  $\Gamma_{11}$  and  $\Gamma_{22}$  are the decay widths, inverse of the meson and anti-meson lifetimes  $\Gamma_q$  and  $\Gamma_{\bar{q}}$  described by the weak interaction.  $CPT$  invariance requires  $M_{11} = M_{22}$  and  $\Gamma_{11} = \Gamma_{22}$ . Meson mixing phenomena imply that the off-diagonal elements of  $\mathbf{M}_q$  and  $\mathbf{\Gamma}_q$  are different from zero. The off-diagonal elements of the mass matrix account for the dominant, dispersive contributions of the box diagrams in Fig 2.2, i.e. diagrams with the internal top-quark. The off-diagonal elements of the decay matrix account for absorptive contributions, i. e. real intermediate decays to a state  $f$ , (e.g.  $B_q \rightarrow f \rightarrow \overline{B}_q$ ) common to the states  $B_q$  and  $\overline{B}_q$  and box diagrams with up-quark and charm-quark. As consequence of these off-diagonal elements, the  $B_q$  and  $\overline{B}_q$  are not eigenstates for the weak interaction.

Both the mass matrix  $\mathbf{M}_q$  and the decay matrix  $\mathbf{\Gamma}_q$  can be complex due to their dependence on CKM-elements. Diagonalizing the effective Hamiltonian ( $\mathbf{M}_q + i \frac{\mathbf{\Gamma}_q}{2}$ ), the

<sup>6</sup>A more complete and rigorous argument can be found for example in Ref. [23].

mass eigenstates of the system,  $|B_L\rangle$  and  $|B_H\rangle$ , with masses  $M_L$  and  $M_H$  and decay widths  $\Gamma_L$  and  $\Gamma_H$  respectively, can be expressed as superpositions of the flavor eigenstates:

$$|B_L\rangle = p|B_q\rangle + q|\overline{B_q}\rangle \quad (2.18)$$

$$|B_H\rangle = p|B_q\rangle - q|\overline{B_q}\rangle \quad (2.19)$$

with  $p$  and  $q$  complex numbers satisfying the normalization condition  $|p|^2 + |q|^2 = 1$ . The time evolution of the mass eigenstates of the effective Hamiltonian ( $\mathbf{H}_q = \mathbf{M}_q + i\frac{\Gamma_q}{2}$ ) is given by:

$$|B_L(t)\rangle = e^{-iM_L t - \frac{\Gamma_L}{2}t} |B_L(0)\rangle \quad (2.20)$$

$$|B_H(t)\rangle = e^{-iM_H t - \frac{\Gamma_H}{2}t} |B_H(0)\rangle \quad (2.21)$$

The average mass and average lifetime, the mass difference  $\Delta M_q$  and the decay width difference  $\Delta\Gamma_q$ <sup>7</sup> between the two mass eigenstates can be defined:

$$m_q = m_{\bar{q}} = \frac{M_H + M_L}{2} \quad \Gamma_q = \frac{\Gamma_H + \Gamma_L}{2} \equiv \frac{1}{\tau} \quad (2.22)$$

$$\Delta M_q = M_H - M_L \quad \Delta\Gamma_q = \Gamma_L - \Gamma_H \quad (2.23)$$

From Eq. 2.19 the flavor eigenstates can be expressed as:

$$|B_q\rangle = \frac{1}{2p}(|B_H\rangle + |\overline{B_L}\rangle) \quad (2.24)$$

$$|\overline{B_q}\rangle = \frac{1}{2q}(|B_H\rangle - |\overline{B_L}\rangle) \quad (2.25)$$

and the time evolution of the flavor eigenstates results:

$$|B_q(t)\rangle = g_+(t)|B_q\rangle + \frac{q}{p}g_-(t)|\overline{B_q}\rangle \quad (2.26)$$

$$|\overline{B_q}(t)\rangle = \frac{p}{q}g_+(t)|B_q\rangle + g_-(t)|\overline{B_q}\rangle \quad (2.27)$$

with

$$g_+(t) = e^{-im_q t} e^{-i\Gamma_q t/2} \left[ \cosh \frac{\Delta\Gamma_q t}{4} \cos \frac{\Delta M_q t}{2} - i \sinh \frac{\Delta\Gamma_q t}{4} \sin \frac{\Delta M_q t}{2} \right] \quad (2.28)$$

$$g_-(t) = e^{-im_q t} e^{-i\Gamma_q t/2} \left[ -\sinh \frac{\Delta\Gamma_q t}{4} \cos \frac{\Delta M_q t}{2} + i \cosh \frac{\Delta\Gamma_q t}{4} \sin \frac{\Delta M_q t}{2} \right] \quad (2.29)$$

---

<sup>7</sup>  $\Delta M_q$  is chosen to be positive, while  $\Delta\Gamma_q$  can be either positive or negative.

Table 2.4: Mass and lifetime parameters of  $B^0$  and  $B_s^0$  systems [17]

Meson	$B^0$	$B_s^0$
$M_q$	$(5279.58 \pm 0.17)$ MeV	$(5366.77 \pm 0.24)$ MeV
$\Delta M_q$	$(0.510 \pm 0.003)$ ps $^{-1}$	$(17.761 \pm 0.022)$ ps $^{-1}$
$\tau_q$	$(1.519 \pm 0.005)$ ps	$(1.512 \pm 0.007)$ ps
$\Delta\Gamma_q$	-	$(0.091 \pm 0.008)$ ps $^{-1}$
$x_d = \Delta M_q/\Gamma_q$	$(0.774 \pm 0.006)$	$(26.85 \pm 0.13)$

It is interesting to calculate the probability of finding a  $\bar{B}$  meson after a time  $t$ , in a beam of  $B$  mesons  $B(t=0)$  initially produced, or vice versa:

$$\begin{aligned}
 |\langle \bar{B}_q | B_q(t) \rangle|^2 &= \frac{e^{-\Gamma_q t}}{2} \left( \cosh\left(\frac{\Delta\Gamma_q t}{2}\right) - \cos(\Delta M_q t) \right) \left| \frac{q}{p} \right|^2 \\
 |\langle B_q | \bar{B}_q(t) \rangle|^2 &= \frac{e^{-\Gamma_q t}}{2} \left( \cosh\left(\frac{\Delta\Gamma_q t}{2}\right) + \cos(\Delta M_q t) \right) \left| \frac{p}{q} \right|^2
 \end{aligned} \tag{2.30}$$

These relations give an idea about the expected time distributions of a sample of  $B_q$  and  $\bar{B}_q$  decays. If the flavor of the  $B_q$  meson at the production and at the decay is known, the expected asymmetry between mesons that did change flavor before decaying and those that did not, is expected to be an oscillating distribution, with characteristic parameters  $\Delta M_q$ , the oscillation frequency, and  $\Delta\Gamma_q$  (with the dependence in Eq. 2.30). Assuming here  $|q/p| = 1$  for simplicity, the next paragraphs focus on the phenomenology, meaning and theoretical implication of having  $|q/p| \neq 1$ .

In Table 2.4 the values of  $\Delta M_q$  and  $\Delta\Gamma_q$  for the  $B^0$  and  $B_s^0$  mesons are reported. In the case of the  $B^0$  mesons, the mixing frequency is comparable to their lifetime, meaning that the mesons oscillate at most once before decaying ( $\Delta M_d/\Gamma_d \sim 0.77$ ).  $\Delta\Gamma_d$  is compatible with zero <sup>8</sup>, meaning that the heavy and light mass eigenstates have the same lifetime ( $y_d \equiv \Delta\Gamma_d/2\Gamma_d \approx 0$ ). In the case of the  $B_s^0$  meson instead, the mixing frequency is much higher ( $\Delta M_s/\Gamma_s \approx 27$ ) and  $\Delta\Gamma_s$  has a large value (about 15% of the lifetime itself).

The formalism and the observables until now introduced are sufficient to describe the  $B_q - \bar{B}_q$  oscillation phenomenon, but in order to determine the SM predictions of the characteristic observables of the process, it is necessary to relate  $\Delta M$ ,  $\Delta\Gamma$ ,  $\Gamma$  and  $M$  to  $M_{12}$  and  $\Gamma_{12}$ . These relations can be determined for example calculating the diagrams involved (the box diagrams in Fig. 2.2 give the dominant contributions). More precisely, mixing observables as  $\Delta M_q$ , and  $\Delta\Gamma_q$ , can be expressed in terms of  $|M_{12}|$ ,  $|\Gamma_{12}|$  and the relative phase  $\phi_q = \arg(-M_{12}/\Gamma_{12})$ . Section 2.5 discusses in detail the observable that measures the  $CP$  violation in  $B_q - \bar{B}_q$  mixing.

<sup>8</sup>A value significantly different from zero could be compatible with some New Physics scenarios, see Sec. 2.6.3.

## 2.4 $CP$ violation

$CP$  violation has been introduced in Sec. 2.2 as different behavior of particles with respect to antiparticles in weak interactions. Namely the Charge-conjugation and Parity ( $CP$ ) transformation, transforms a particle in its own antiparticle (the electrical charge change signs and the spatial coordinates are inverted). As mentioned in Sec. 2.2, the SM accommodates a source of  $CP$  violation in weak interactions with the CKM-matrix and the three families of quarks, the question is whether the amount of  $CP$  violation observed in the variety of processes we can probe, is in agreement or not with the SM predictions. It is necessary to select the processes that are affected by  $CP$  violation, and in order to do this, it is convenient to start from the common classification of possible types of  $CP$  violating effects. Moreover, it is useful to understand, the implications of measurements of  $CP$  violating observables on the constraints of the CKM picture. A brief summary of the classification of  $CP$  violating effects in this Section is followed by a more complete description of  $CP$  violation in the process of interest for this thesis, the  $B^0 - \bar{B}^0$  mixing. For this process, also a hint about the relation between the measurable  $CP$  observable and the CKM constraints is given.

The  $CP$  violating effects mentioned here are described for  $B$  mesons, but they are generally valid.

*CP violation in decay* or direct  $CP$  violation. This type of  $CP$  violation occurs when the time integrated decay rate of a  $B$  meson into a final state  $f$ , is different with respect to the  $CP$  conjugate time integrated decay rate, i.e. the decay rate of  $\bar{B}$  to  $\bar{f}$ :

$$\Gamma(B \rightarrow f) \neq \Gamma(\bar{B} \rightarrow \bar{f}) \quad (2.31)$$

An example of this process could be the  $B \rightarrow K^+\pi^-$  decay (See for example Ref. [27]). Introducing the commonly used notation for the decay amplitudes

$$A_f = \mathcal{A}(B \rightarrow f); \quad \bar{A}_{\bar{f}} = \mathcal{A}(\bar{B} \rightarrow \bar{f}); \quad \bar{A}_f = \mathcal{A}(\bar{B} \rightarrow f); \quad A_{\bar{f}} = \mathcal{A}(B \rightarrow \bar{f}) \quad (2.32)$$

the direct  $CP$  violation condition can be expressed as:

$$\left| \frac{A_f}{\bar{A}_{\bar{f}}} \right| \neq 1. \quad (2.33)$$

*CP violation in mixing.* Measurements of this type of  $CP$  violation, probe if the oscillation of a neutral meson into its antimeson is different from its revers process. In Sec. 2.3 the meson-antimeson transitions are described assuming no  $CP$  violation. It can be shown that  $CP$  violation in mixing is observed when

$$|q/p| \neq 1 \quad (2.34)$$

$CP$  violation in  $B^0$ - $\bar{B}^0$  and  $B_s^0$ - $\bar{B}_s^0$  mixing is predicted to be very small (see Sec. 2.5 and 2.6), but, for example, in the neutral Kaons system, it is the dominant  $CP$  violating effect.

$CP$  violation in the interference between a decay with and without mixing. This form of  $CP$  violation is measured in decays to a final,  $f$ , state accessible from both mesons and relative antimesons. The necessary condition is:

$$\Gamma(B_q^0 \rightarrow \bar{B}_q^0 \rightarrow f)(t) \neq \Gamma(\bar{B}_q^0 \rightarrow B_q^0 \rightarrow f)(t) \quad (2.35)$$

That can be equivalently expressed by means of the  $CP$  violating parameter  $\lambda_f$ , defined in function of  $q$ ,  $p$ , and amplitudes of the processes considered ( $B \rightarrow f$  and  $\bar{B} \rightarrow f$ ):

$$\mathcal{I}m(\lambda_f) \equiv \mathcal{I}m\left(\frac{q \bar{A}_f}{p A_f}\right) \neq 0 \quad (2.36)$$

Well known examples are given by the  $B^0 \rightarrow J/\psi K_S^0$  decays and  $B_s^0 \rightarrow J/\psi \phi$  decays. The second and third type of  $CP$  violation presented here are also known as *indirect CP violation*.

## 2.5 $CP$ violation in mixing

For a better understanding of indirect  $CP$  violation, it is necessary to focus on the eigenvalue problem of neutral meson mixing, and on expressing the characteristic observables ( $\Delta M_q$ ,  $\Delta \Gamma_q$ ) in terms of the more theoretical quantities  $M_{12}$ ,  $\Gamma_{12}$ . When calculating the eigenvalues  $M_{L/H} - i\Gamma_{L/H}$  of the effective Hamiltonian  $\mathbf{M}_q + i\frac{\Gamma_q}{2}$

$$M_{L/H} - i\Gamma_{L/H} = M_q - i\frac{\Gamma_q}{2} \mp \sqrt{\left(M_{12} - i\frac{\Gamma_{12}}{2}\right)\left(M_{12}^* - i\frac{\Gamma_{12}^*}{2}\right)} \quad (2.37)$$

it is convenient to define also the relative phase between  $M_{12}$  and  $\Gamma_{12}$ <sup>9</sup>:  $\phi_{12} = \arg\left(-\frac{M_{12}}{\Gamma_{12}}\right)$ . The following relations can be derived

$$(\Delta M_q)^2 - \frac{1}{4}(\Delta \Gamma_q)^2 = 4|M_{12}^q|^2 - |\Gamma_{12}^q|^2 \quad (2.38)$$

$$\Delta M_q \Delta \Gamma_q = -4|M_{12}^q||\Gamma_{12}^q|\cos\phi_{12}^q, \quad (2.39)$$

where the index  $q = s, d$ , to indicate the  $B^0$  or  $B_s^0$  system. From experimental observations in  $B_q^0$  systems,  $\Delta M_q \gg \Delta \Gamma_q$ , and theoretically the inequality  $|\Gamma_{12}^q| \ll \Delta M_q$  is established, since possible New Physics effect must be small or would be in conflict with the precisely measured semileptonic branching ratio [24]. Eq. 2.38 and Eq. 2.39 imply  $\Delta M \simeq 2|M_{12}|$ .

---

<sup>9</sup> $|B\rangle$  and  $|\bar{B}\rangle$  can always be rotated by an arbitrary phase, without changing the physics, the individual phases of  $M_{12}$  and  $\Gamma_{12}$  therefore don't have a physical meaning.

As a consequence, for the  $B_q^0$  systems, it is possible to assume  $|M_{12}^q| \gg |\Gamma_{12}^q|$  [23], and relate  $\Delta\Gamma_q$  and  $\Delta M_q$  to  $M_{12}^q$ ,  $\Gamma_{12}^q$ ,  $\phi_{12}^q$

$$\Delta M_q \simeq 2|M_{12}^q| \quad (2.40)$$

$$\Delta\Gamma_q \simeq -2|\Gamma_{12}^q|\cos\phi_{12}^q \quad (2.41)$$

There is a third characteristic observable of neutral meson mixing, which is related to  $CP$  violation in the process. First of all,  $q$  and  $p$ , coefficients of the linear combinations in 2.19 can be expressed as:

$$\frac{q}{p} = -\frac{\Delta M_q + i\Delta\Gamma_q/2}{2M_{12}^q - i\Gamma_{12}^q} = -\frac{2M_{12}^{q*} - i\Gamma_{12}^{q*}}{\Delta M_q + i\Delta\Gamma_q/2} \quad (2.42)$$

and using the relations 2.39:

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^{q*}}{M_{12}^q} \frac{1 + i\left|\frac{\Gamma_{12}^q}{2M_{12}^q}\right|e^{i\phi_{12}^q}}{1 + i\left|\frac{\Gamma_{12}^q}{2M_{12}^q}\right|e^{-i\phi_{12}^q}} \quad (2.43)$$

From this relation it is clear that values of  $\phi_{12}^q \neq 0, \pi$ , imply  $|q/p| \neq 1$ , and vice versa. From Eq.2.30, the value of  $|q/p|$  determines whether the probability of a  $B_q^0$  meson to oscillate into a  $\bar{B}_q^0$  meson is equal to the reverse transition. A difference in these two probabilities, given that  $B_q^0$  and  $\bar{B}_q^0$  are  $CP$  eigenstates, translates in  $CP$  violation in the  $B_q^0$  mesons oscillation process.  $CP$  violation in mixing can be quantified by the parameter  $a$ , defined as

$$\left|\frac{q}{p}\right|^2 = 1 - a \quad (2.44)$$

For the  $B_q^0$  meson systems the quantity  $a$  is known to be small, for this reason, to get the relation between  $a$  and the theoretical quantities  $M_{12}^q$  and  $\Gamma_{12}^q$ , it is possible to expand  $(q/p)^2$  in Eq. 2.43 in terms of  $\phi_{12}^q$  and  $\Gamma_{12}^q/M_{12}^q$ .

$$a = \mathcal{I}m\frac{\Gamma_{12}^q}{M_{12}^q} + \mathcal{O}\left(\left(\mathcal{I}m\frac{\Gamma_{12}^q}{M_{12}^q}\right)^2\right) = \left|\frac{\Gamma_{12}^q}{M_{12}^q}\right|\sin\phi_{12}^q \quad (2.45)$$

The formalism so far introduced is used to express the decay rates of  $B_q^0$  and  $\bar{B}_q^0$  mesons to a final state  $f$  or to the  $CP$  conjugate final state  $\bar{f}$ . The time dependent decay rate  $d\Gamma/dt(B(t) \rightarrow f)$  of a  $B$  meson into a final state  $f$ , is defined as:

$$\frac{d\Gamma}{dt}(B \rightarrow f) = \frac{1}{N_B} \frac{dN(B(t) \rightarrow f)}{dt} \quad (2.46)$$

with  $N_B$  the total number of  $B$  mesons produced at the time  $t = 0$ . A commonly used notation for the decay amplitudes is,

$$A_f = \mathcal{A}(B \rightarrow f) = \langle f|\mathcal{H}_W|B\rangle, \quad \bar{A}_f = \mathcal{A}(\bar{B} \rightarrow f) = \langle f|\mathcal{H}_W|\bar{B}\rangle \quad (2.47)$$

and

$$\Gamma(B(t) \rightarrow f) = \mathcal{N}_f |\langle f | \mathcal{H}_W | B(t) \rangle|^2, \quad \Gamma(\bar{B}(t) \rightarrow f) = \mathcal{N}_f |\langle f | \mathcal{H}_W | \bar{B}(t) \rangle|^2 \quad (2.48)$$

where  $H_W$  indicates the Hamiltonian operator of the weak interaction and  $\mathcal{N}_f$  represents a normalization factor accounting for the integration over the phase-space. Exploiting these notations, it is possible to write the time decay rates in terms of  $A_f$  and  $\bar{A}_f$ . The time evolution of the meson flavor eigenstates  $|B(t)\rangle$  and  $|\bar{B}(t)\rangle$  is reported in Eq. 2.27. Eq. 2.29 gives the definition of  $g_+(t)$  and  $g_-(t)$ . Finally it is possible to write  $\bar{A}_f$  in terms of  $\lambda_f$  (using the definition in Eq. 2.36) and  $|q/p|^2$  can instead be expressed in terms of  $a$  (Eq. 2.44).

$$\Gamma(B(t) \rightarrow f) = \mathcal{N}_f |A_f|^2 \frac{e^{-\Gamma t}}{2} \left\{ (1 + \lambda_f^2) \cosh \frac{\Delta\Gamma t}{2} + (1 - \lambda_f^2) \cos(\Delta Mt) - 2\mathcal{R}e(\lambda_f) \sinh \frac{\Delta\Gamma t}{2} - 2\mathcal{I}m(\lambda_f) \sin(\Delta Mt) \right\} \quad (2.49)$$

$$\Gamma(\bar{B}(t) \rightarrow f) = \mathcal{N}_f |A_f|^2 \frac{e^{-\Gamma t}}{2} \left\{ (1 + \lambda_f^2) \cosh \frac{\Delta\Gamma t}{2} - (1 - \lambda_f^2) \cos(\Delta Mt) - 2\mathcal{R}e(\lambda_f) \sinh \frac{\Delta\Gamma t}{2} + 2\mathcal{I}m(\lambda_f) \sin(\Delta Mt) \right\} \quad (2.50)$$

Considering the  $CP$ -conjugated finale state of  $f$ , indicated with  $\bar{f}$ , the corresponding decay rates are obtained (writing  $A_{\bar{f}}$  in terms of  $\bar{\lambda}_{\bar{f}}$ ):

$$\Gamma(B(t) \rightarrow \bar{f}) = \mathcal{N}_f |\bar{A}_{\bar{f}}|^2 \frac{e^{-\Gamma t}}{2} \left\{ (1 + \lambda_{\bar{f}}^2) \cosh \frac{\Delta\Gamma t}{2} - (1 - \lambda_{\bar{f}}^2) \cos(\Delta Mt) - 2\mathcal{R}e(\bar{\lambda}_{\bar{f}}) \sinh \frac{\Delta\Gamma t}{2} + 2\mathcal{I}m(\bar{\lambda}_{\bar{f}}) \sin(\Delta Mt) \right\} \quad (2.51)$$

$$\Gamma(\bar{B}(t) \rightarrow \bar{f}) = \mathcal{N}_f |\bar{A}_{\bar{f}}|^2 \frac{e^{-\Gamma t}}{2} \left\{ (1 + \lambda_{\bar{f}}^2) \cosh \frac{\Delta\Gamma t}{2} + (1 - \lambda_{\bar{f}}^2) \cos(\Delta Mt) - 2\mathcal{R}e(\bar{\lambda}_{\bar{f}}) \sinh \frac{\Delta\Gamma t}{2} - 2\mathcal{I}m(\bar{\lambda}_{\bar{f}}) \sin(\Delta Mt) \right\} \quad (2.52)$$

These are the general time-dependent decay rates for neutral  $B$  mesons, without making any particular assumption of the decay mode. In the interest of defining an experimentally accessible quantity carrying information on the  $CP$  nature of  $B \rightarrow \bar{B}$  and the reverse transition, it is necessary to consider so called *flavor specific* decay modes. A decay is flavor specific if  $B \rightarrow f$  transitions are allowed, but  $B \rightarrow \bar{f}$  are forbidden. In other words, the final state identifies unambiguously the flavor of the  $B$  meson at the

moment of decaying. Given a final state  $f$ , the meson at decay is necessarily a  $B$  meson, it could have been a  $\bar{B}$  meson at the production, only in the case that it oscillated to  $B$  before decaying. When considering flavor specific decays, the decay rates in Eq. 2.49, 2.52 result substantially simplified. Flavor specific decays satisfy the conditions:  $\bar{A}_f = A_{\bar{f}} = 0$  leading to  $\lambda_f = \bar{\lambda}_{\bar{f}} = 0$ <sup>10</sup>. In addition, given that the main interest of this thesis is to quantify the  $CP$  violation in the  $B^0$  mixing process only, a further assumption is made: only tree-level decays without direct  $CP$  violation are considered. This last assumption implies  $|\bar{A}_{\bar{f}}| = |A_f|$ . Under these hypotheses, the decay rates read:

$$\Gamma(B(t) \rightarrow f) = \mathcal{N}_f |A_f|^2 \frac{e^{-\Gamma t}}{2} \left\{ \cosh \frac{\Delta\Gamma t}{2} + \cos(\Delta Mt) \right\} \quad (2.53)$$

$$\Gamma(\bar{B}(t) \rightarrow f) = \mathcal{N}_f |A_f|^2 \frac{e^{-\Gamma t}}{2(1-a)} \left\{ \cosh \frac{\Delta\Gamma t}{2} - \cos(\Delta Mt) \right\} \quad (2.54)$$

$$\Gamma(B(t) \rightarrow \bar{f}) = \mathcal{N}_f |A_f|^2 \frac{e^{-\Gamma t}}{2} (1-a) \left\{ \cosh \frac{\Delta\Gamma t}{2} - \cos(\Delta Mt) \right\} \quad (2.55)$$

$$\Gamma(\bar{B}(t) \rightarrow \bar{f}) = \mathcal{N}_f |A_f|^2 \frac{e^{-\Gamma t}}{2} \left\{ \cosh \frac{\Delta\Gamma t}{2} + \cos(\Delta Mt) \right\} \quad (2.56)$$

The  $CP$  asymmetry in flavor specific decays is defined as:

$$a_{\text{fs}} = \frac{\Gamma(\bar{B}(t) \rightarrow f) - \Gamma(B(t) \rightarrow \bar{f})}{\Gamma(\bar{B}(t) \rightarrow f) + \Gamma(B(t) \rightarrow \bar{f})} = \frac{1 - (1-a)^2}{1 + (1-a)^2} = a + \mathcal{O}(a^2) \quad (2.57)$$

The asymmetry  $a_{\text{fs}}$  is between the  $CP$  conjugated final states, when considering decays where the  $B$  meson has changed flavor with respect to the production flavor. It is an experimentally measurable quantity, but requires the knowledge of the  $B$  meson flavor at the production.

Decays of  $B$  mesons, and  $b$  hadrons in general, can be classified according to the structure of the charged current interaction involved, and therefore according to the final state: leptonic, semileptonic and hadronic decays. The asymmetry  $a_{\text{fs}}$  is often called also *semileptonic  $CP$  asymmetry*,  $a_{\text{sl}}$ , because semileptonic decays are ideal decays to measure these asymmetries. First of all they are flavor specific and have no direct  $CP$  violation, necessary conditions in order to derive Eq. 2.57. Moreover, the  $B$  meson has a semileptonic branching fraction of about 10%, resulting in high statistics data samples, which provide a unique opportunity to measure very small asymmetries (See Sec. 2.6.1). Fig. 2.3 illustrates the Feynman diagram of a semileptonic decay. As last remark,  $a_{\text{sl}}$  is defined through the time-dependent decay rates, but as shown in Eq. 2.57 has the property of being time-invariant.

---

<sup>10</sup> $\lambda_f = \frac{q \bar{A}_f}{p A_f}$ ,  $\bar{\lambda}_f = \frac{1}{\lambda_f}$ ,  $\lambda_{\bar{f}} = \frac{q \bar{A}_{\bar{f}}}{p A_{\bar{f}}}$ ,  $\bar{\lambda}_{\bar{f}} = \frac{1}{\lambda_{\bar{f}}}$

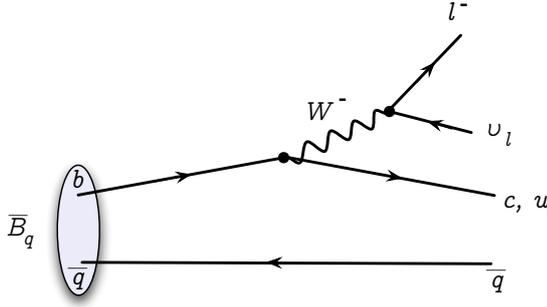


Figure 2.3: Feynman diagram of the semileptonic  $B_q^0$  decay.

## 2.6 $B^0 - \bar{B}^0$ system

The purpose of this section is to present in quantitative terms the current theoretical and experimental picture of the  $B^0 - \bar{B}^0$  system.

### 2.6.1 Theoretical predictions

In order to make theoretical predictions for the characteristic observables of neutral meson mixing, it is necessary to take into account that quarks are bound in mesons, which means that the external quarks in the diagrams in Fig. 2.2 are forced to form mesons. Weak processes in the SM are described with the help of an *effective Hamiltonian*  $\mathcal{H}^{eff}$ . In particular for  $B_q - \bar{B}_q$  mixing amplitudes, the effective Hamiltonian  $\mathcal{H}^{|\Delta B|=2}$  describing only the processes in which the bottom quantum number  $B$  changes by two units<sup>11</sup>. In this formalism the  $\mathcal{H}^{|\Delta B|=2}$  does not account for W-boson or t-quark fields explicitly, as represented in the diagrams in Fig. 2.2, but contains an effective point-like coupling of four quark fields, commonly denoted with  $Q$ . The  $|\Delta B| = 2$  interaction can be expressed in terms of an effective coupling constant multiplying the four-quark operator  $Q$ :

$$\mathcal{H}^{|\Delta B|=2} = \frac{G_F^2}{4\pi^4} (V_{tq}^* V_{tb})^2 C^{|\Delta B|=2}(m_t, M_W, \mu) Q(\mu) \quad (2.58)$$

The coupling constant can be split in few factors. The first contains the Fermi constant  $G_F$ . The second includes the CKM matrix elements entering the box diagram. Lastly  $C^{|\Delta B|=2}(m_t, M_W, \mu)$  is the so-called *Wilson coefficient*, which carries the information of the  $M_W$  and  $m_t$  energy scales involved in the process (recalling that the top quark contribution dominates the  $B - \bar{B}$  mixing phenomenon). The factorization between short-distance coefficients and hadronic long-distance matrix element is also known as *Operator Product*

<sup>11</sup>  $B^0 = |\bar{b}d\rangle$  mesons have  $B$  quantum number = -1 while  $\bar{B}^0 = |b\bar{d}\rangle$  mesons have  $B$  quantum number = +1.

*Expansion* [28]. For example, the amplitude of the  $|\Delta B| = 2$  transition reads

$$\mathcal{A}(\bar{B} \rightarrow B) = \langle B | \mathcal{H}^{|\Delta B|=2} | \bar{B} \rangle = \frac{G_F^2}{4\pi^4} (V_{tb} V_{tq}^*)^2 C^{|\Delta B|=2}(m_t, M_W, \mu) \langle B | Q(\mu) | \bar{B} \rangle \quad (2.59)$$

The problem of calculating this amplitude can be separated in two parts: the short distance perturbative calculation of the Wilson coefficients and the long distance (generally non-perturbative) calculation of the matrix elements  $\langle B_q | Q(\mu) | \bar{B}_q \rangle$ . The energy scale  $\mu$  separates the physics of the two contributions. The value of  $\mu$  is an arbitrary choice, but the amplitude must  $\mu$ -independent, therefore the  $\mu$ -dependence of the Wilson coefficients should cancel the  $\mu$ -dependence of  $\langle B_q | Q(\mu) | \bar{B}_q \rangle$ . The hadronic matrix element is conventionally parametrized in terms of the decay constant  $f_{B_q}$  and the so-called *bag factor*  $\hat{B}_{B_q}$ , quantities calculated with the help of lattice gauge theory computations [29], [30], [31].

In order to make a theoretical prediction for the  $B - \bar{B}$  mixing observables,  $M_{12}^q$  and  $\Gamma_{12}^q$  need to be evaluated.

$$M_{12}^q = \frac{\langle B_q | \mathcal{H}^{|\Delta B|=2} | B_q \rangle}{2M_{B_q}} \propto (V_{tq}^* V_{tb})^2 \hat{B}_{B_q} f_{B_q}^2 \quad (2.60)$$

From  $M_{12}^q$  it is straight forward to calculate the prediction of  $\Delta M_q$ , using the relation in Eq. 2.40. In order to find the predicted values for  $\Delta\Gamma_q$  and  $a_{\text{fs}}^q$ , it is necessary to calculate  $\Gamma_{12}^q$ . To include the diagrams with light internal quarks that are responsible for  $\Gamma_{12}$ , and therefore  $\Delta\Gamma_q$ , the effective Hamiltonian formalism has to be extended to include also  $\Delta B = 1$  physics. Calculations of  $\Gamma_{12}^q$  for the  $B$  system can be found in Refs. [32] [33]. From Eq. 2.41 and Eq. 2.45, the predictions for  $\Delta\Gamma_q$  and  $a_{\text{fs}}^q$  are derived. For a recent review of the SM prediction in flavor physics, including  $B_q$  mixing observables, see Ref. [34]. The SM predictions for  $CP$  violation in  $B_q$  meson mixing systems are [35] :

$$a_{\text{sl}}^d = (-4.1 \pm 0.6) \times 10^{-4} \quad (2.61)$$

$$a_{\text{sl}}^s = (1.9 \pm 0.3) \times 10^{-5}. \quad (2.62)$$

They are tiny and precise compared to the current experimental sensitivity.

## 2.6.2 Experimental status

An overview of existing measurements and the world average is provided by the Heavy Flavor Averaging Group (HFAG Summer 2014) [36] and they are reported in Table 2.5 . In Fig. 2.4 the experimental results are visualized in the  $\{a_{\text{sl}}^d - a_{\text{sl}}^s\}$  ( $a_{\text{sl}}$  for the  $B^0$  and  $B_s^0$  systems) plane. The average value of  $a_{\text{sl}}^d$  obtained from CLEO [37], BaBar [38–40], Belle [41] measurements results in  $a_{\text{sl}}^d = (0.02 \pm 0.32)\%$ . At the  $\Upsilon(4S)$  center-of-mass energy of the  $e^+e^-$  machines it is impossible to produce  $B_s^0$  mesons, for this reason the  $B_s^0$  sector is not accessible for  $B$ -factories <sup>12</sup> The experimental average for  $a_{\text{sl}}^s$  is therefore

<sup>12</sup>Part of the dataset of the Belle experiment was collected from  $\Upsilon(5S)$  center-of-mass energy collisions. Nevertheless there is no published measurement of  $a_{\text{sl}}^s$  by the Belle experiment.

obtained from the  $D\bar{O}$  [42] and LHCb [43] measurements to be  $a_{\text{sl}}^s = (-0.48 \pm 0.48)\%$ . Finally, a new measurement of  $a_{\text{sl}}^d$  by  $D\bar{O}$  [44], and an inclusive asymmetry measurement of:

$$A_{\text{sl}} = C_d a_{\text{sl}}^d + C_s a_{\text{sl}}^s \quad \text{with } C_d = 0.58, C_s = 0.42 \quad (2.63)$$

considering a sample including  $B_s^0$  and  $B^0$  semileptonic decays performed by the  $D\bar{O}$  experiment [8], should also be included in the overall picture. It is common to refer to the latter, common  $a_{\text{sl}}^d$  and  $a_{\text{sl}}^s$  measurement as *dimuon* asymmetry measurement, due to the method used (Sec. 4.1 explains the measurement strategy for this measurement). The average of all public experimental results before 31 March 2014 [36] is  $a_{\text{sl}}^d = (-0.09 \pm 0.21)\%$  and  $a_{\text{sl}}^s = (-0.77 \pm 0.42)\%$ , where the correlation coefficient between these numbers is found to be  $-0.195$ .

The most precise measurements of  $a_{\text{sl}}^d$  and  $a_{\text{sl}}^s$  are shown in Table 2.5. Nowadays two  $a_{\text{sl}}^d$  measurements need to be added to this picture. These are the measurement presented in this thesis, and a new measurement by the BaBar collaboration [45] finalized at about the same time.

The  $D\bar{O}$  dimuon measurement, shows a discrepancy of about 3 standard deviations from the SM prediction of  $A_{\text{sl}}$  (see Eq. 2.63). If confirmed, this would be a hint of New Physics in the  $B^0 - \bar{B}^0$  mixing. This experimental result therefore provides an additional motivation to further investigate  $CP$  violation in both  $B_s^0$  and  $B^0$  systems.

Table 2.5: Overview of measurements of  $a_{\text{sl}}^d$  and  $a_{\text{sl}}^s$  (first uncertainty is statistical, second systematic), and averages from the Heavy Flavor Averaging Group. Earlier measurements from LEP, which do not distinguish between the  $B^0$  and  $B_s^0$  contribution are not included.

Exp. & Ref.	Method	Measurement
CLEO [37]	Dileptons + partial hadronic	$a_{\text{sl}}^d = (1.4 \pm 4.1 \pm 0.6)\%$
Belle [41]	Dileptons	$a_{\text{sl}}^d = (-0.11 \pm 0.79 \pm 0.85)\%$
BaBar [38]	Full hadronic rec.	$a_{\text{sl}}^d = (-5.8 \pm 2.6 \pm 2.2)\%$
BaBar [39]	Dileptons	$a_{\text{sl}}^d = (0.16 \pm 0.54 \pm 0.38)\%$
BaBar [40]	Partial semilept.	$a_{\text{sl}}^d = (0.06 \pm 0.17_{-0.32}^{+0.38})\%$
Average of $B$ factories above [36]		$a_{\text{sl}}^d = (0.02 \pm 0.32)\%$
BaBar [45]	Partial semilept.	$a_{\text{sl}}^d = (-0.39 \pm 0.35(\text{stat}) \pm 0.19(\text{syst}))\%$
$D\bar{O}$ [44]	Partial semilept.	$a_{\text{sl}}^d = (0.68 \pm 0.45 \pm 0.14)\%$
$D\bar{O}$ [8]	Dimuon	$a_{\text{sl}}^d = (-0.62 \pm 0.42)\%$ $a_{\text{sl}}^s = (-0.86 \pm 0.74)\%$ $\rho = -0.79$
$D\bar{O}$ [42]	Partial semilept.	$a_{\text{sl}}^s = (-1.12 \pm 0.74 \pm 0.17)\%$
LHCb [10]	Partial semilept.	$a_{\text{sl}}^s = (-0.06 \pm 0.50 \pm 0.36)\%$
Average of all measurements above [36]		$a_{\text{sl}}^d = (-0.09 \pm 0.21)\%$ $a_{\text{sl}}^s = (-0.77 \pm 0.42)\%$

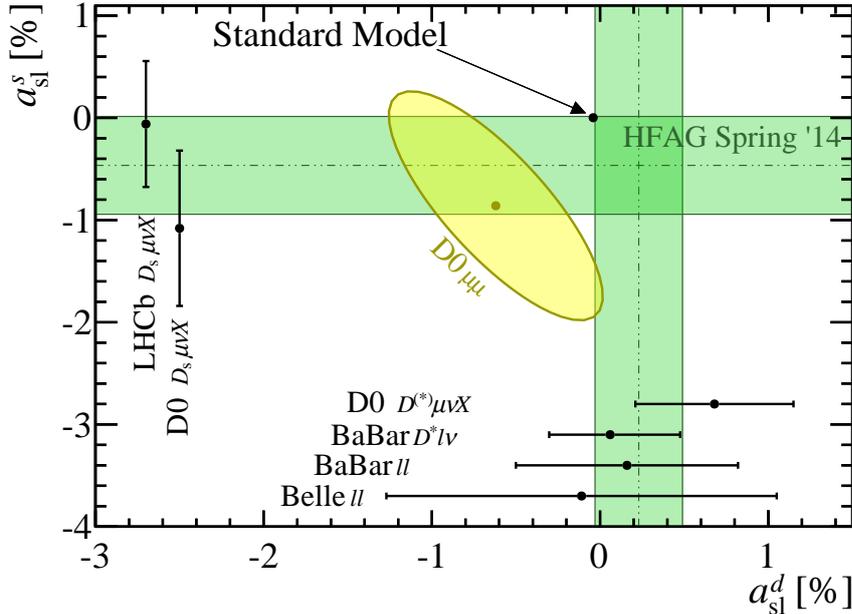


Figure 2.4: Overview of measurements in the  $\{a_{\text{sl}}^d - a_{\text{sl}}^s\}$  plane. The 68% confidence regions are displayed. The HFAG averages do not include the  $a_{\text{sl}}^d$  result of this thesis, the BaBar  $a_{\text{sl}}^d$  measurement [45] and the DØ dimuon result [8]. The numerical values of the measurements are given in Table 2.5.

### 2.6.3 Bounds on new physics

A broad range of New Physics models accommodates additional sources of  $CP$  violation in  $B_q - \bar{B}_q$  mixing. For example supersymmetry scenarios [46] can predict different particles in the box diagrams of Fig 2.2. Another possibility is offered by extra-dimensions scenarios [47] [48] or quark flavor violation [49], where also tree-level transitions could contribute to the mixing amplitudes.

A model independent parametrization of the new effects entering the mixing process is helpful in order to understand the size of  $CP$  violating effects can possibly be observed in the semileptonic  $CP$  asymmetries and be consistent at the same time with all the other measurements probing the  $B_q$  sector. Deviations from the SM mixing amplitude induced by New Physics (NP) effects might be conveniently parametrized in terms of the shift induced in the modulus,  $|\Delta|_q$ , and in the  $CP$ -violating phase  $\phi_q$  [50]

$$\Delta_q \equiv \frac{\langle B_q | \mathcal{L}_{SM} + \mathcal{L}_{NP} | B_q \rangle}{\langle B_q | \mathcal{L}_{SM} | B_q \rangle} = |\Delta_q| e^{i\phi_q^\Delta} \quad (2.64)$$

The New Physics parametrized as  $\Delta_q = |\Delta_q| e^{i\phi_q^\Delta}$ , can be assumed to act only in  $M_{12}^q$ , while  $\Gamma_{12}^q$  is given by the SM prediction, i.e.

$$\Gamma_{12}^q = \Gamma_{12}^{q,SM}, \quad M_{12}^q = M_{12}^{q,SM} \cdot \Delta_q. \quad (2.65)$$

This assumption is expected to describe correctly any type of New Physics within the hadronic uncertainties [51]. The mixing observables can be expressed in terms of the SM predictions and the complex parameter  $\Delta_q$

$$\Delta M_q = 2|M_{12}^{q,SM}| \cdot |\Delta_q| \quad (2.66)$$

$$\Delta\Gamma_q = 2|\Gamma_{12}^q| \cdot \cos(\phi_q^{SM} + \phi_q^\Delta) \quad (2.67)$$

$$a_{\text{fs}}^q = \frac{|\Gamma_{12}^q|}{|M_{12}^{q,SM}|} \cdot \frac{\sin(\phi_q^{SM} + \phi_q^\Delta)}{|\Delta_q|} \quad (2.68)$$

Experimental determinations of  $\Delta M_q$ ,  $\Delta\Gamma_q$  and  $a_{\text{fs}}^q$  represent therefore bounds in the complex  $\Delta_q$ -plane. Eq. 2.66-2.68 are helpful to understand these constraints. Fig 2.5 represents the current experimental constraints on New Physics in the  $B^0$  system. Measurements of  $\Delta M_q$  are sensitive to the modulus  $|\Delta_q|$  (orange cycle in Fig. 2.5), while measurements of  $\Delta\Gamma_q$  give direct bounds on the phase of  $\Delta_q$ , which correspond to the light-blue radial regions in Fig. 2.5. Finally,  $a_{\text{fs}}^q$  constrains simultaneously modulus and phase of  $\Delta_q$  (grey area in Fig. 2.5).

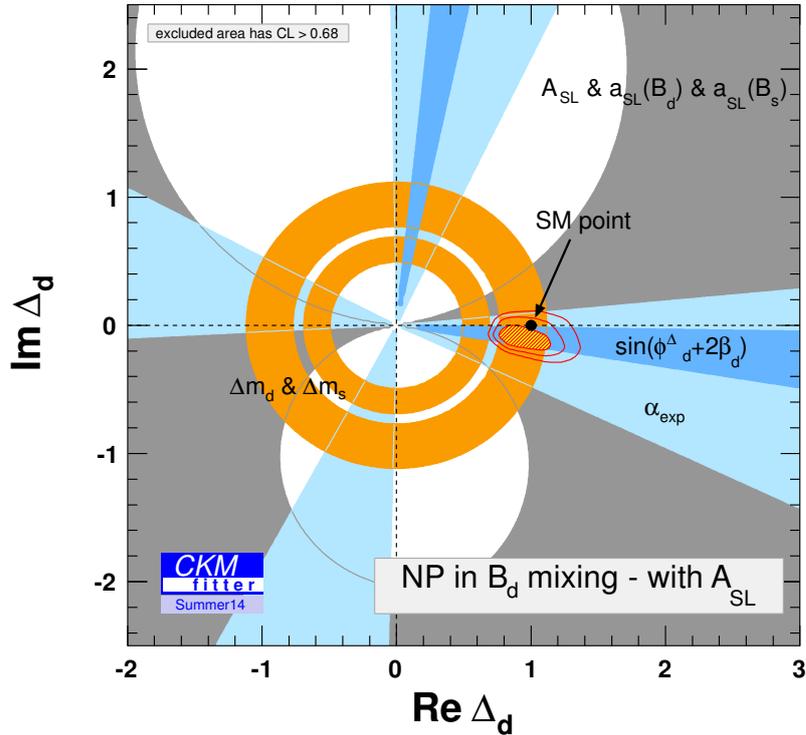


Figure 2.5: Experimental bounds on new physics in  $B^0$  system. [22].

The possibility of having new physics in  $\Gamma_{12}^q$  has also been considered [51]. This hypothesis is especially interesting, since it could point to a possible explanation for the  $D\bar{O}$  dimuon asymmetry. As found in Ref. [52], the dimuon asymmetry measured by the  $D\bar{O}$  experiment, accounts not only for contributions proportional to  $a_{sl}^d$  and  $a_{sl}^s$ , as assumed in Eq. 2.63, but it is also modified by other  $CP$ -violating processes. For example the process

$$\begin{aligned} p\bar{p} &\rightarrow B^+\bar{B}^0 X, \\ B^+ &\rightarrow \mu^+ X, \\ \bar{B}^0 &\rightarrow D^+D^-, D^- \rightarrow \mu^+ X, \end{aligned}$$

produces a  $\mu^+\mu^+$  pair that contributes to the dimuon asymmetry. The  $D^+D^-$  final state is a  $CP$ -even eigenstates and is accessible from both  $B^0$  and  $\bar{B}^0$  mesons. The additional contribution to the dimuon asymmetry, according to Ref. [52] originates from the interference between mixing and decay and that might be approximated by being proportional to  $\Delta\Gamma_d$ <sup>13</sup>. A more detailed discussion about the size of this contribution can be found in Ref. [53]. While an large enhancement of  $\Delta\Gamma_s$  without violating other experimental constraints is not possible [54],  $\Delta\Gamma_d$  is only weakly constrained by other measurements [35]. Different scenarios as the violation of CKM unitarity, the existence of new  $b\bar{d}\tau\tau$  operators or NP affecting differently  $b \rightarrow c\bar{c}d$ ,  $b \rightarrow c\bar{u}d$  and  $b \rightarrow u\bar{c}d$  transitions, are analyzed [55], [56].

## 2.7 Summary

$CP$  violation in  $B_q$  mixing means that the probability that a  $B_q$  meson mixes into a  $\bar{B}_q$  meson is different from the probability that a  $\bar{B}_q$  mixes into a  $B_q$ . The flavour specific or “semileptonic” asymmetry is defined as

$$a_{sl}^q = \frac{\Gamma(\bar{B}_q \rightarrow B_q \rightarrow f) - \Gamma(B_q \rightarrow \bar{B}_q \rightarrow \bar{f})}{\Gamma(\bar{B}_q \rightarrow B_q \rightarrow f) + \Gamma(B_q \rightarrow \bar{B}_q \rightarrow \bar{f})}, \quad (2.69)$$

with the subscript  $q$  distinguishing the two species of neutral  $B$  mesons, namely the  $B_s^0$  and  $B_d^0$ .

The Standard Model predictions [35]:  $a_{sl}^d = (-4.1 \pm 0.6) \times 10^{-4}$  and  $a_{sl}^s = (1.9 \pm 0.3) \times 10^{-5}$  are tiny compared to the current experimental sensitivity of a few per mille. This makes the precise measurement of these asymmetries an excellent test of the Standard Model. In the case that a significant deviation from zero is measured, this would be a signal for New Physics. Among the past experimental results, the  $D\bar{O}$  dimuon measurement [8] shows a discrepancy of about 3 standard deviations from the SM prediction. This result provides

<sup>13</sup> $\Gamma_{12}^q$  stems from the decays into final states common to  $B_q^0$  and  $\bar{B}_q^0$  mesons, as in the case of the  $D^+D^-$  final state, and from Eq. 2.39  $|\Gamma_{12}^q|$  relates to  $\Delta\Gamma_q$ .

an additional motivation to further investigate  $CP$  violation in both  $B_s^0$  and  $B^0$  systems. The analysis presented in this thesis aims to clarify the  $B^0$  picture by providing a precise measurement of  $a_{\text{sl}}^d$ .



# Chapter 3

## The LHCb experiment

The LHCb (Large Hadron Collider beauty) experiment is one of the four large experiments at the Large Hadron Collider (LHC) at CERN. It is designed to perform precision measurements of  $b$ - and  $c$ - hadron decays.

This chapter is meant to briefly describe the experiment, focusing on the information necessary to understand the main experimental challenges of the  $a_{\text{sl}}^d$  measurement at LHCb. After having introduced the LHC, the  $b$  quark and  $B$  meson production mechanisms are discussed. LHCb detector and performance are discussed, including an overview of the subsystems. The trigger system and the event reconstruction is subsequently summarized, defining some quantities useful for the understanding of the following chapters.

### 3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [57] is a proton-proton collider designed to reach a center-of-mass energy up to 14 TeV and housed in the 27 km long Large Electron-Positron collider (LEP) tunnel. This accelerator and collider is the largest scientific instrument in the world, mostly aimed at the discovery of the Higgs boson and the study of rare decays, which provide crucial tests of the Standard Model consistency and excellent searches for New Physics. A more detailed overview of the physics program is found in the physics programs of the single experiments.

#### 3.1.1 LHC design

The LHC consists of eight arcs with a length of about 2.8 km and eight 500 m long straight sections. It lies at a depth between 45 m and 170 m from the surface, and has an inclination of 1.42% with respect to the horizontal plane, to allow an easier civil engineering between the two different geological zones present in the area. In order to achieve the design energy of 7 TeV per beam, the bending magnets which keep the protons on the designed path need to produce a field strength of 8.3 Tesla. This requires the usage of superconducting magnets cooled with superfluid helium at 1.9 K. The diameter of 3.8 m of the LEP tunnel requires two separate beam pipes for the beams in a compact

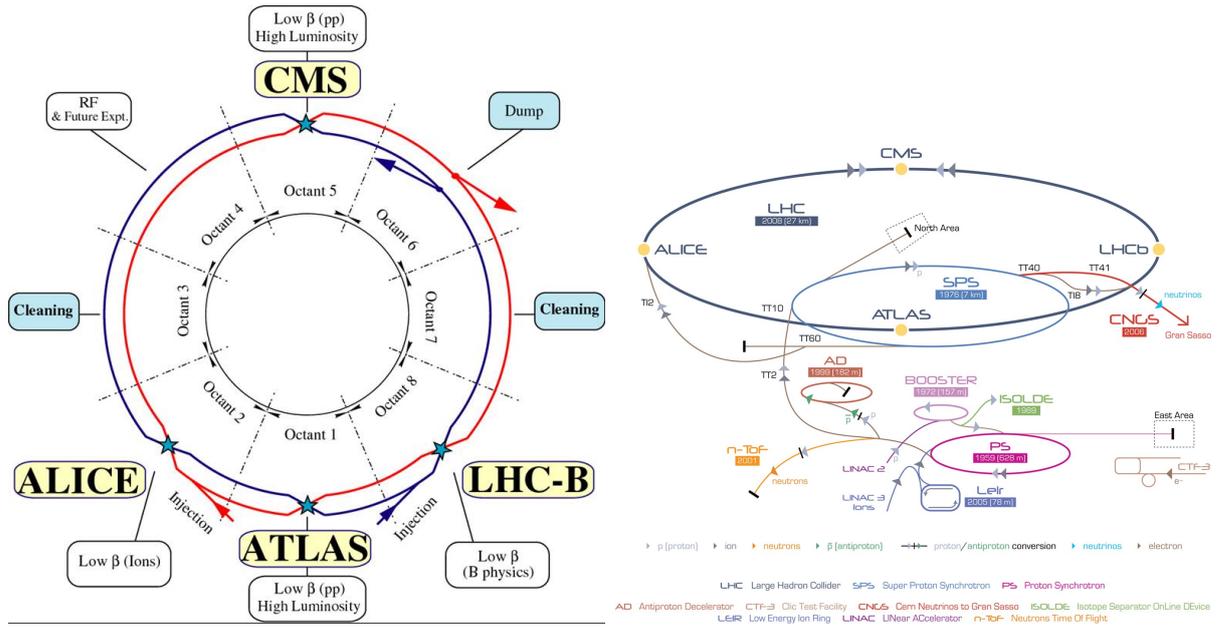


Figure 3.1: *Left* Schematic layout of the LHC. Beam 1 (red) circulates clockwise while beam 2 (blue) counter-clockwise. *Right* Accelerator chain at CERN. To obtain collisions at the LHC, the steps are: LINAC2, BOOSTER, PS, SPS, LHC.

structure. The two pipes are separated by only 19 cm and are placed in a single cryostat. In order to house the four main experiments, four collision points were chosen. Fig. 3.1 shows the layout of the LHC, highlighting the four collision points as well as the remaining access points. The four main experiments at the LHC are designed for different scopes. The two *General Purpose Detectors*, ATLAS and CMS are hosted in caverns at Point 1 and Point 5, respectively. These detectors are designed for high transverse momentum ( $p_T$ ) physics. The ALICE experiment, located at Point 2, investigates instead the primordial states of matter such as the *quark gluon plasma*, by exploiting heavy ion collisions also provided by the LHC. LHCb is instead a dedicated flavor physics experiment, which focuses on  $CP$  violation measurements and rare decays of beauty and charm hadrons, complemented by a broad physics program including electroweak measurements and other topics. It is located at the Point 8 interaction point.

While proton beams circulate in the LHC, a halo of particles is formed around the core of the beam. These particles move towards the external region of the beam due to the non-linearities of the magnetic field, or due to the interactions between the beams. Beam losses, i.e. particles that hit the wall of the vacuum chambers are dangerous for the machine, because they can lead to a quench<sup>1</sup> due to the heat locally produced. It is therefore important to remove the halo particles. This is done by the *collimation system* located at Point 3 and Point 7, where graphite blocks can be moved into the beam pipe

<sup>1</sup>Sudden loss of the superconductive regime in a magnet

to absorb the halo particles. This is not the only technology used, a more complete description can be found in Ref. [58]. Radio Frequency (RF) cavities are placed at Point 4, with the purpose of accelerating the protons and correct the longitudinal spread of the proton distribution. The RF use longitudinally oscillating electric fields operating at 400 MHz to accelerate the protons. This frequency needs to be tuned to the revolution frequency of the LHC in order to provide a constant acceleration. The *beam-dump* facility is situated at Point 6, and consists of a quick firing kicker magnet which steers the beam away from its normal path and deflects it in a 700 m long evacuated pipe. The beam is then dumped onto a block of graphite.

The LHC is only the last of the chain of devices to produce, bunch and accelerate protons. Figure 3.1 illustrates the full accelerator chain, which starts with extraction of protons as nuclei of hydrogen atoms with an energy of 50 keV, which are guided to a linear accelerator (LINAC2), where they reach the energy of 50 MeV. They are subsequently injected into a booster synchrotron, where they are accelerated up to 1.4 GeV, and injected in the Proton Synchrotron (PS). At 2 GeV protons are transferred to the Super Proton Synchrotron (SPS), where they reach the energy of 450 GeV, before they are injected into the LHC by means of one of the two transfer lines. After filling both the LHC beams, which requires about 4 cycles of the PS and about 12 cycles of the SPS [57], the magnetic field in the bending magnets is raised, bringing the beam energy from 450 GeV to the collision energy. The beams are then "squeezed" to maximize the number of collisions at the collision points. From the technical design, the full process is expected to require a minimum time of about 70 min. A beam lifetime of several hours can be achieved in the LHC<sup>2</sup>.

### 3.1.2 LHC performance

Energy per beam and luminosity are the most important figures of merit for the performance of the LHC. The luminosity indicates how many collisions occur per time unit and given the cross-section. It is related to the number of collisions per second and to the cross section

$$\frac{dN}{dt} = \mathcal{L}\sigma \quad (3.1)$$

Where  $dN/dt$  is the number of collisions per second,  $\mathcal{L}$  the instantaneous (machine) luminosity and  $\sigma$  the cross-section of the process considered. The total amount of data is given by the integrated luminosity  $L = \int \mathcal{L}dt$ . The machine luminosity depends on the beam parameters, and for a Gaussian beam distribution, it can be written as:

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta^*} F \quad (3.2)$$

where  $N_b$  indicates the number of particles per bunch,  $n_b$  is the number of bunches per beam,  $f_{rev}$  is the revolution frequency,  $\gamma_r$  the relativistic gamma factor,  $\epsilon_n$  the normalized

---

<sup>2</sup>The fill duration is designed to be about 15 h, from experimental data collected by LHCb, the fill duration distribution follows an exponential decay distribution with average lifetime of few hours.

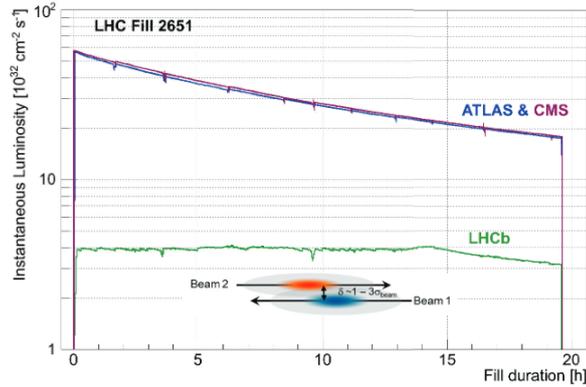


Figure 3.2: Evolution of the instantaneous luminosity for LHCb, compared to ATLAS and CMS during the LHCb fill 2651. After reaching the desired value, the instantaneous luminosity is kept constant at LHCb in a range of 5% for about 15 hours by adjusting the transversal beam overlap. The different behavior among the experiments at the end of the fill is due to the different focusing at the collision points, commonly referred to as  $\beta^*$ . Figure taken from [59]

transverse beam emittance,  $\beta^*$  the beta function at the collision point <sup>3</sup> and  $F$  is the geometrical luminosity reduction factor due to the crossing angle at the interaction point. The emittance of a beam indicated the average spread of the particles in position and momentum phase-space. In a low emittance beam the particles are confined to a small area and have nearly the same momentum.

### 3.1.3 Interaction Point 8

At the interaction points the two proton beams are bent to collide with each other with a crossing-angle. A low value of the  $\beta$  function is required, in order to maximize the number of the collisions, and it is achieved with a dedicated string of three quadrupole magnets. From the machine point of view, the main difference between LHCb and the other experiments is the shifted interaction point. In order to accommodate the single arm spectrometer (described in detail in Sec.3.2) in the existing hall, the interaction point and the focussing quadrupoles are displaced by  $3\lambda_{RF}/2$  ( $\sim 11.22$  m) towards Point 7. This shift has some implications on beam-beam effects [60]. Moreover, the LHCb experiment features a dipole magnet (see Sec. 3.2.1) whose polarity is reversed regularly. This additional magnetic field complicates the accelerator optics at the interaction point Point 8, and results in different crossing angles of the beams for the two magnet polarities. In addition, at collision it is desirable to adjust the optics to fulfill the luminosity requirements dynamically, i.e. to keep the luminosity approximately constant during one run despite decaying beam intensities. This goal is reached with a *luminosity leveling* technique,

<sup>3</sup>the  $\beta$  function in accelerator physics is a function related to the transverse size of the beam along the trajectory. The value of the  $\beta$  function at the collision point is indicated as  $\beta^*$ , and it is used to indicate how squeezed the beam is at the collision point

that consists of shifting the beams with respect to each other according to the desired instantaneous luminosity (See Fig. 3.2). The LHCb detector was designed for an average instantaneous luminosity of  $2 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ <sup>4</sup>, and a peak luminosity of  $5 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$  with the center-of-mass energy of  $\sqrt{s} = 14 \text{ TeV}$ . LHCb has decided to keep a constant luminosity, lower than the full luminosity provided by the LHC, as exploited by CMS and ATLAS experiments, mainly for three reasons. First of all, the forward region is characterized by a high flux of particles, which leads to high occupancies in the detectors. Secondly, LHCb is devoted to the study of beauty and charm decays, therefore the ability to distinguish the primary vertex from the secondary vertex is crucial for many analysis performed. This task becomes more difficult with a larger pile-up. Finally, high occupancies in the tracking detectors result in a degradation of the track reconstruction ability. The luminosity requirements for LHCb were therefore chosen as a compromise between all these effects and the need for high statistics samples to perform precision measurements.

### 3.1.4 $b$ quark production in $pp$ collisions

In  $pp$  collisions, quarks are produced as  $q\bar{q}$  pairs, since strong interactions are flavor conserving. The leading order (LO) processes for the production of  $b\bar{b}$  pairs are quark-antiquark annihilation ( $q\bar{q} \rightarrow b\bar{b}$ ) and gluon-gluon fusion ( $gg \rightarrow b\bar{b}$ ), known as *pair creation* processes. At the next-to-leading order (NLO) level, the contributions from gluon-splitting and flavor-excitation become relevant. The contribution of leading-order processes to the total  $b$ -cross section decreases for increasing center-of-mass energy values. The flavor-excitation contribution becomes dominant at the LHC for center-of-mass energies of 7 TeV or 14 TeV, which exceeds the pair-creation LO processes as shown in Fig. 3.3 (*Left*). Considering that the  $b\bar{b}$  production threshold, given by the invariant mass of the pair ( $\sim 10 \text{ GeV}/c^2$ ), is small compared to the center-of-mass energy of the LHC, the favorite pair-production mechanism is the gluon-gluon fusion (see Fig. 3.3 (*Right*), [61]). Different contributions are simulated in event-generators as PYTHIA [62]. The  $b\bar{b}$  production threshold is small when compared to the center-of-mass energy. This implies that when two protons collide, the partons originating the  $b\bar{b}$  pair have likely very different momenta. As result  $b\bar{b}$  quark pairs are often created with a large boost and therefore tend to fly along the axis of one incoming proton. Furthermore there is a strong correlation between the  $b$  and the  $\bar{b}$  quark which causes them both to end up in the forward or backward direction. This is shown in Fig. 3.4 and makes clear the reason for the forward design of the LHCb experiment. The total  $b\bar{b}$  cross section in  $pp$  collisions at center-of-mass energy  $\sqrt{s} = 7 \text{ TeV}$  has been measured at LHCb using semileptonic  $b$  hadron decays [63]:

$$\sigma(p\bar{p} \rightarrow (b\bar{b})) = (284 \pm 20 \pm 49) \mu\text{b} \quad (3.3)$$

At the center-of-mass energy  $\sqrt{s} = 8 \text{ TeV}$ , the total  $b\bar{b}$  cross section is measured to be

$$\sigma(p\bar{p} \rightarrow (b\bar{b})) = (298 \pm 2 \pm 36) \mu\text{b} \quad (3.4)$$

---

<sup>4</sup>This  $2 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$  value is the LHCb design value for average instantaneous luminosity, LHCb took data at a luminosity of  $3.5 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$  in 2010 and 2011 with 7 TeV center-of-mass energy and  $4 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$  in 2012 with 8 TeV center-of-mass energy.

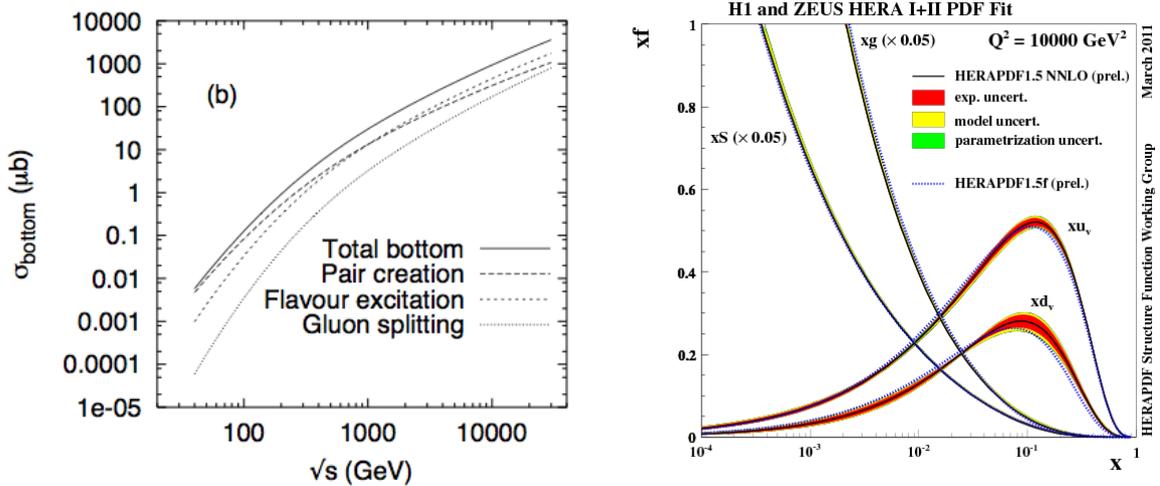


Figure 3.3: *Left* Total  $b$ -cross section as function of  $E_{CM} = \sqrt{s}$ . The contributions from pair creation, flavor excitation and gluon splitting are shown separately. Figure taken from [65]. *Right* The parton distribution functions from HERAPDF1.5 NNLO and HERAPDF1.0 NNLO at a region relevant for the hadron colliders Tevatron and LHC ( $Q^2 = 10000 \text{ GeV}^2$ ). The gluon and sea distributions are scaled down by a factor 20. The experimental, model and parametrization uncertainties are shown separately. Figure taken from [61].

from the measurement of the  $J/\psi$  from  $b$  decays production cross section [64]

### 3.1.5 $B$ mesons production in $pp$ collisions

The  $b$  quarks produced according to the processes mentioned in Sec. 3.1.4 subsequently combine with a lighter antiquark or two quarks to form hadrons, bound states of a quark and an antiquark (mesons) or of three quarks (baryons). This process is known as *hadronization*. The lighter quarks come from the proton remnants or from the fragmentation process in the initial interaction. Table 3.1 shows the measured fractions of hadronizations of a  $\bar{b}$  quark in  $B^+$  ( $\bar{b}u$ ),  $B^0$  and  $B_s^0$  mesons. Given that the LHC is colliding two proton beams, and the hadrons formed by the  $b\bar{b}$  pairs depend on the quarks of the proton remnants of fragmentation processes, an asymmetry in the production of the hadrons with respect to the anti-hadrons is expected.

Different mechanisms lead to an asymmetry between the number of mesons produced compared to the number of antimesons for a given species.

First of all an excess of baryons with respect to anti-baryons is expected, a produced  $b$  quark likely combines with proton remnants or quarks from the fragmentation process to form heavy baryons, or with antiquarks from the fragmentation to form antimesons. A produced  $\bar{b}$  antiquark essentially hadronizes instead to a meson. This leads to an excess of the  $B^+$  and  $B^0$  with respect to the corresponding anti-mesons. In the case of  $B_s^0$  mesons, most of  $\bar{b}$  quarks combine with  $u$  and  $d$  quarks, so the formation of  $\bar{B}_s^0$  mesons, which

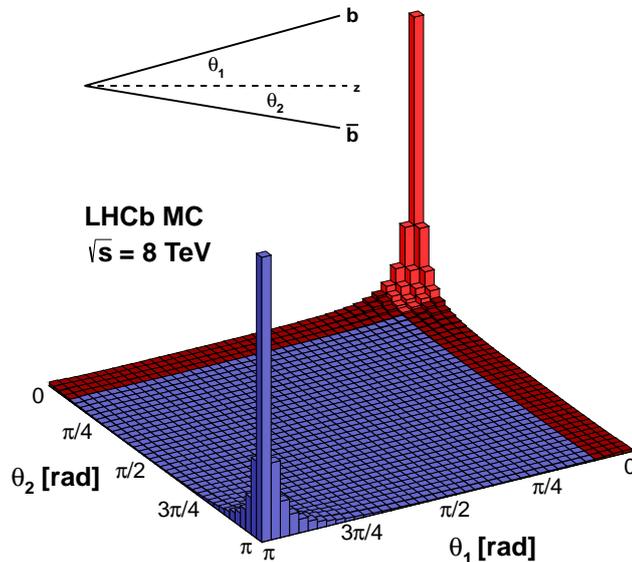


Figure 3.4: Polar angle distribution of  $b$  and  $\bar{b}$  collisions at  $\sqrt{s} = 8$  TeV. The beam line lies on the  $z$ -axis and the red region corresponds to the LHCb acceptance. The figure is taken from [66]

Table 3.1: Fragmentation fractions for different hadrons containing a  $b$  quark. For the first row the fraction of  $B^+$  and  $B^0$  are set to be equal for isospin symmetry among  $u$  and  $d$  quarks and the number applies for one type of  $B$  meson and not for both together. The numbers are taken from [36]

Fraction	Value
$B^+ \text{ or } B^0, f_u = f_d$	$0.406 \pm 0.005$
$B_s^0, f_s$	$0.105 \pm 0.005$
$b$ -baryons, $f_{baryon}$	$0.083 \pm 0.010$

contain a  $b$  quark and an  $\bar{s}$  quark from vacuum is favored with respect to the  $B_s^0$  mesons. While this argument provides a reason of the larger number of produced  $B$  mesons, it does not describe the production asymmetries in multiple parton shower. A phenomenological model that describes the hadronization process is the *Lund string model* [68], which describes the color flow in the process by means of strings formed by self-interacting gluons. The currently used event generators as the PYTHIA package [69] are based on this model. Two different sources of meson-antimeson production asymmetry can be distinguished [65], [67].

The first effect is more important at high transverse momenta and is represented in Fig. 3.5

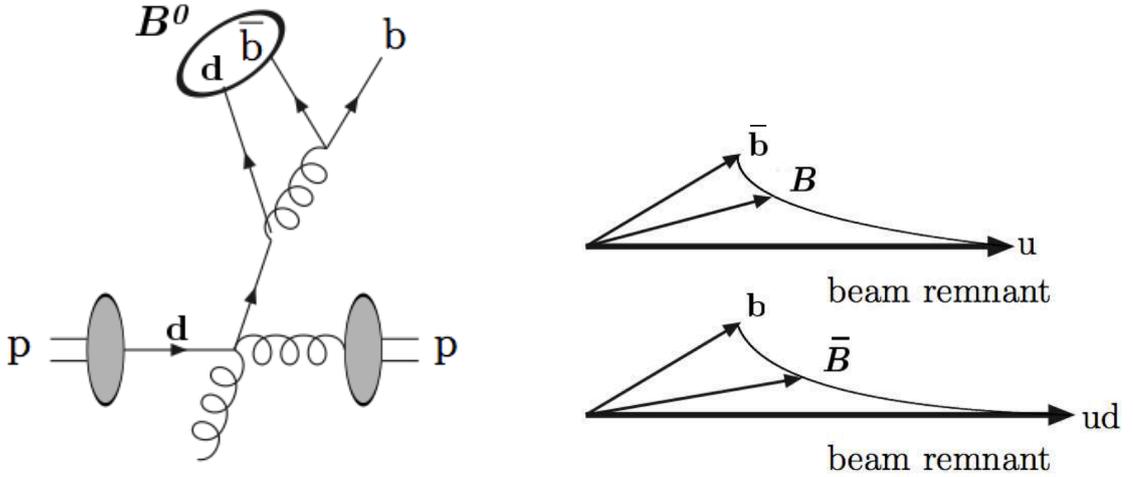


Figure 3.5: Collapse to a  $B^0$  meson at high  $p_T$  (*right*) [65] and beam drag effect (*left*) [67].

(*left*). A  $\bar{b}$  antiquark produced in a  $b\bar{b}$  pair as explained in Sec. 3.1.4 and a scattered valence quark from a proton, can collapse to a  $B$  meson. This effect favors the production of heavy mesons with a light quark content in common with the proton at high momenta. The second effect, depicted in Fig. 3.5 (*right*), is the so called *beam drag* effect. The  $b$  quark and two beam remnants can be in a color-neutral state, while the  $\bar{b}$  can be in color-neutral state only with one beam remnant, since the beam remnants have different color charge. As consequence the  $b$  quark is “dragged” by two partons to higher rapidities, while the  $\bar{b}$  can be “dragged” to higher rapidities by only one parton. This effect favors the production of heavy antimesons at higher rapidity<sup>5</sup> regions.

In summary, the asymmetries between the production of hadrons with respect to the corresponding anti-hadrons depends on the kinematics of the  $b$  quark. For this reason having a detector acceptance covering kinematics regions symmetric with respect to the collision point (as occurs in the case of  $4\pi$  detectors or would occur in an hypothetical detector having the same coverage in the forward and backward regions), would help to obtain a partial cancelation of the effect. In the case of LHCb the hadron production asymmetry effect is expected to be at the percent level, depending on the meson/baryon of interest and on the kinematic region considered, and it is an important effect, clearly competing with  $CP$  violating effects, the measurement presented in this thesis represents an important example. Furthermore, production asymmetries arise from soft QCD processes involved in the hadronization, and theory currently needs some input from the experimental side to achieve a better understanding of the processes involved to provide predictions. Therefore, it is important to measure the hadron - anti-hadron production asymmetries at LHCb. Measurements of  $\bar{B}^0$ - $B^0$  and  $\bar{B}_s^0$ - $B_s^0$  production asymmetries performed with

<sup>5</sup>The rapidity is defined as  $y = \frac{1}{2} \ln \frac{E+p_z c}{E-p_z c}$ , where the  $z$ -axis is along the beam line. More widely-used is the pseudo-rapidity  $\eta$ , defined as  $\eta = -\ln \left( \tan \frac{\theta}{2} \right)$ .

fully reconstructed decays at LHCb are described in [70]. The use of the high statistics semileptonic samples collected by LHCb will allowed a more precise determination of these important parameters.

## 3.2 The LHCb experiment

The LHCb experiment [71], [59] is located at Point 8 of the LHC in the same cavern where the DELPHI experiment [72] at LEP was located. It is a single-arm spectrometer, with a forward geometry that allows to reconstruct a large fraction of produced particles containing a  $b$  or  $\bar{b}$  quark while covering only a small solid angle (see Sec. 3.1.4 and Fig. 3.4). Its geometrical coverage ranges from about 10 mrad to approximately 300 (250) mrad in the horizontal (vertical) direction. LHCb uses a coordinates system in which the  $z$ -axis is along the beam pipe, the positive direction pointing from the interaction point to the muon system (see Fig. 3.6), the  $y$ -coordinate is vertical, oriented from the interaction point to the surface and perpendicular to the LHC tunnel <sup>6</sup>, the  $x$ -coordinate is defined such that the three axes form a right-handed coordinate system. Three parts can be identified within the LHCb detector. The *track reconstruction system*, which determines particle trajectories and momenta, the *particle identification system*, that determines the type of the particles, and the *trigger system*, which selects the interesting events.

### 3.2.1 Track reconstruction system

Precise vertex reconstruction and high momentum resolution are key ingredients for the LHCb performance. The tracking system is designed to reconstruct the trajectories of charged particles, so-called *tracks*, achieving these two purposes. All the detectors devoted to track reconstruction need to have a high spatial resolution and a low material budget. All the tracking detectors are built in two halves, placed left and right of the beam pipe, that during data taking are closed to ensure a complete coverage, but they can be opened if needed for intervention.

The LHCb track reconstruction system includes the dipole magnet and the tracking detectors: the Vertex Locator (VELO) and the Tracker Turicensis (TT) <sup>7</sup> before the magnet and the T stations behind. The VELO is a detector surrounding to the interaction point, that measures with very high precision the position of the Primary Vertices (PV) and Impact Parameter (IP) of the tracks. The TT, a second tracking detector placed before the dipole magnet, helps in improving the momentum resolution and rejection of the fake tracks. The dipole magnet plays an important role in the determination of the momentum of the particles: the magnetic fields bends the trajectories of the particles, the comparison between the slopes before and after the magnet provides a momentum determination. After the magnet the T stations provide the measurements used for the

---

<sup>6</sup>the LHC tunnel is inclined by 3.6 mrad with respect to the horizontal plane due to geological constraints.

<sup>7</sup>Also called “Trigger Tracker”. It was used in the trigger starting from 2012.

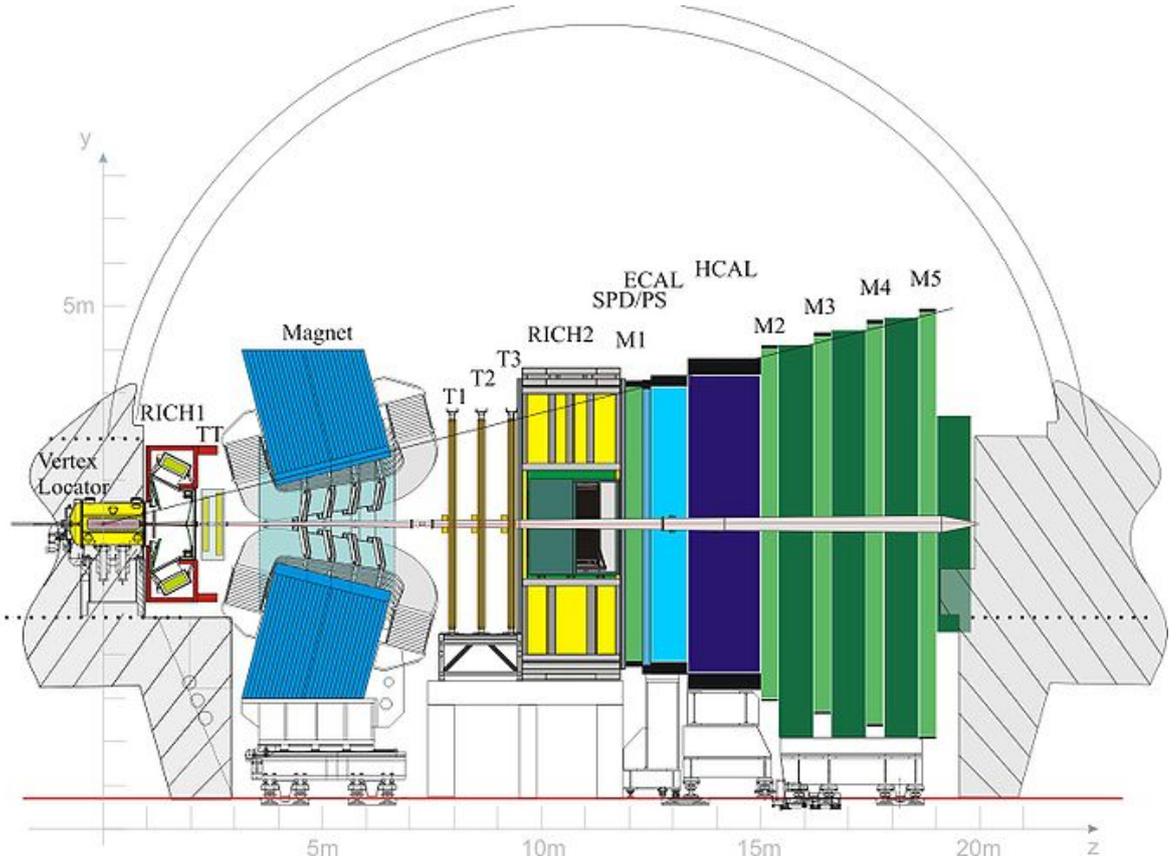


Figure 3.6: The LHCb detector

track reconstruction.

VELO and TT are silicon microstrip detectors, while the T stations use different technologies to trace the particles: silicon microstrips in the inner region (IT), where the detector occupancy is higher and straw-based gas detector in the outer region (OT).

### Dipole magnet

The LHCb dipole magnet is located about 5 m from the interaction region. The main component of the magnetic field,  $B_y$  shown in Fig. 3.17, is pointing in the  $y$  direction, therefore particles are predominantly bent in the horizontal plane (containing the  $x$  and  $z$  axes) The integrated magnetic field is  $\int \vec{B} d\vec{l} = 4 \text{ Tm}$ . All the tracking detectors are located outside of the magnetic field or where only a weak field is present. The momentum resolution for particles transversing the whole tracking system  $\Delta p/p = 0.4\%$  at 2 GeV and  $0.6\%$  at 100 GeV. Due to the detector geometry and the dipole field, charged particles, depending on the charge, are bent preferentially to one side of the detector. In order to properly reconstruct the trajectories of particles inside LHCb, a precise knowledge of the magnetic field map is needed. The magnetic field is measured with Hall probes before the

data taking periods and the results are compared to the magnetic field simulations.

An important characteristic of the LHCb magnet is that the polarity can be periodically reversed. This allows to control the detector systematics related to the bending direction of the tracks. Charges particles are deflected in  $x - z$  plane in the LHCb coordinates system such that for examples, as shown in Fig. 3.7, muon of one charge preferentially hit the muon stations from the left-hand side, while muons of the opposite charge prefer the right-hand side of the detector. A difference in performances of the left and right sides of the detectors would lead to different efficiencies to detect particles of one charge with respect to the relative opposite charge particles. Similarly a left-right misalignment of the detector with respect to the nominal position would lead to a different acceptance on the right side and left side of the detector. Different acceptances also lead to an asymmetry between the detection of particles of a certain charge with respect to the particles of the opposite charge. If data samples of equal size are collected with the two magnet polarities and the operating conditions for the two polarities are the same, the effects of detection asymmetries between particles with opposite charges due to the bending of the tracks are expected to cancel. The hypotheses of having datasets with exactly the same size and exactly the same operation conditions are hard to achieve, but the LHCb strategy is to change the polarity every two weeks of data acquisition, in order to have the same amount of data (about  $120 \text{ pb}^{-1}$ ) recorded for each polarity. Frequent polarity reversal also helps to keep the same configurations of detector and trigger. In addition between two technical stops <sup>8</sup>, the number of polarity flips is chosen to ensure the same size data sets acquired with each polarity. This frequency of polarity changes is shown to allow a control of the detection asymmetries at the level of  $\mathcal{O}(10^{-3})$  [73]. The  $a_{\text{sl}}^d$  measurement does not rely on the LHCb magnet polarity reversal to cancel the charge detection asymmetries. On the other hand the full analysis is performed independently on the dataset acquired with one polarity and with the other. This ensures a better understanding of the determination of the detection asymmetries and a powerful crosscheck to verify the reliability of the measurement. The final result for  $a_{\text{sl}}^d$  is obtained as average of the results obtained for single magnet polarities, which cancels the eventual higher order effects. With this procedure the detection asymmetries are correctly accounted for, despite the single polarities samples acquired in 2011 are not of the same size.

## Vertex Locator

The VELO subdetector [74] is located as close as possible to the proton-proton interactions, and aims to measure the position of the primary vertex and the trajectories of the particles in that region. The VELO consists of in total 21 “stations” of silicon strip sensors. Each station has a detector module on the left and on the right of the beam axis and each module includes an  $r$ -sensor and a  $\phi$ -sensors (see Fig. 3.8).  $r$ -sensors measure the radial coordinate with circular-shaped strips, subdivided into four 45 degrees sectors.  $\phi$ -sensors have strips in approximately radial direction subdivided into inner and outer regions, to

---

<sup>8</sup>Periods without data taking in order to make the major maintenance interventions on the detector

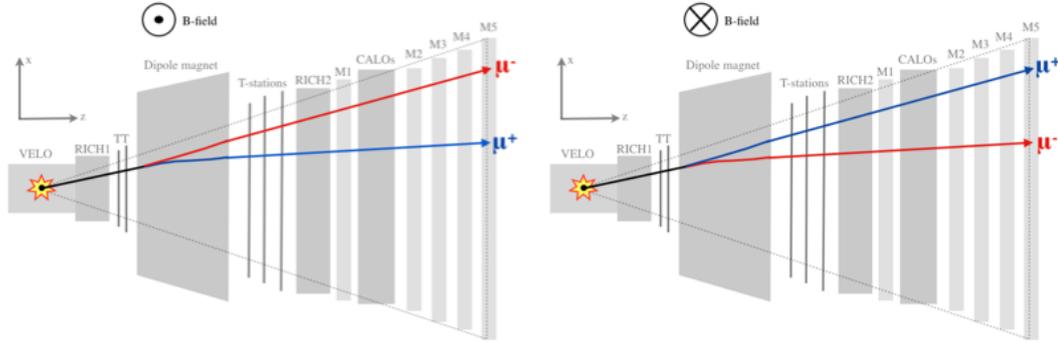


Figure 3.7: Illustrative example of trajectories of two muons with the same initial kinematics but opposite charges. The left and right figure correspond to the polarities up and down respectively of the LHCb dipole magnet. Figure taken from [73].

determine the azimuthal angle  $\phi$  defined as the angle between the x-axis and a direction vector in the  $x - y$  plane. The strip pitch varies from  $40 \mu\text{m}$  to  $100 \mu\text{m}$ , with the finer granularity close to the beam. A VELO module has a diameter of about  $90 \text{ mm}$  and covers a bit more than  $180$  degrees in azimuthal angle, both  $r-$  and  $\phi-$  sensors are  $300 \mu$  thick. Two additional stations, located upstream of the nominal interaction point, were designed to veto pile-up events. They consist of  $r$ -sensors only. The detector modules on each VELO half are contained in an alluminum-walled box, which is under vacuum. An RF foil, made of an aluminium alloy, separates the beam vacuum region from the vacuum inside the VELO-box. This RF-foil represents about  $40\%$  of the material budget of the VELO. Given that the beams are not always stable, especially in the transition between injection-state and stable-beams-state, the VELO halves can be moved away from the beams, to avoid radiation damage. In this position the two halves are about  $6 \text{ cm}$  from each other, while the VELO sensors stand at a radial distance of about  $7 \text{ mm}$  from the beam when taking data. The VELO achieves a best spatial resolution of about  $4 \mu\text{m}$ , which is the best vertex detector resolution reached at the LHC.

### Tracker Turicensis

The Tracker Turicensis (TT) is a  $150 \text{ cm}$  wide and  $130 \text{ cm}$  high silicon microstrip detector located right before the magnet. It consists of two stations with two layers each. Out of these four layers, two are tilted by an angle of  $-5^\circ$  and  $+5^\circ$  with respect to the vertical axis. to be able to perform three dimensional measurements. This configuration allows for a stereo view, which results in a reconstruction of the tracks in three dimensions. The vertical orientation of the strips is chosen to obtain the best spatial resolution in the horizontal plane (bending plane of the magnet), allowing for a more accurate momentum estimate. Figure 3.9 shows the geometry of the TT modules. The layers of the TT are built up from half modules, each of those consists of seven silicon sensors, which are read out in sectors. Each sector contains three or four sensors located far away from the beam

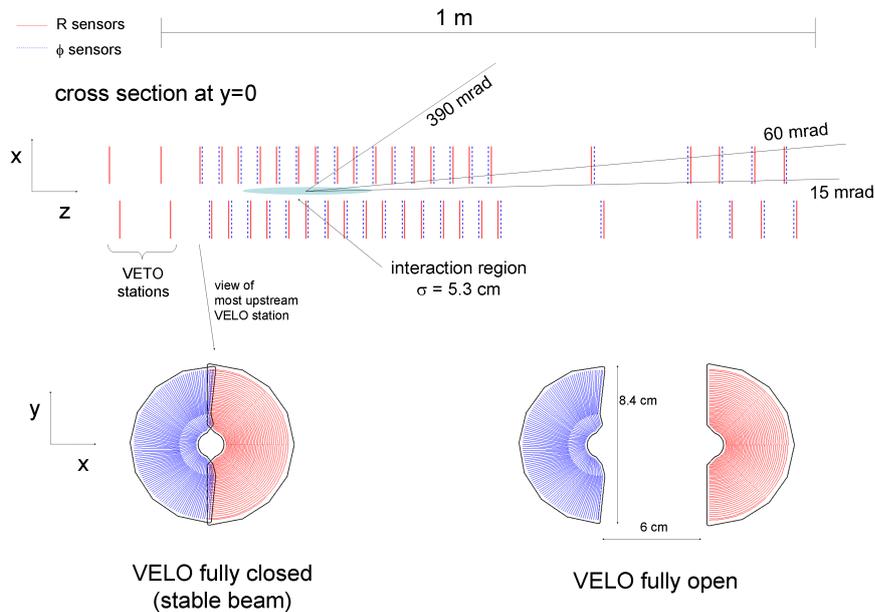


Figure 3.8: VELO Overview diagram. Diagram showing spacing of modules along Z, and open and closed configurations of the sensors.  $\phi$ -sensors (blue) and  $r$ -sensors (red) are displayed.

pipe or one or two sensors close to the beam pipe, given that the occupancy close to the beam is much higher than in the outer regions. Each sensor is 9.44 cm long, 9.46 cm wide and has a pitch of 183  $\mu\text{m}$ . The single hit resolution of the TT is about 50  $\mu\text{m}$ .

### Inner Tracker

Downstream of the magnet, the T stations use two different technologies according to the distance from the beam pipe. Closer to the beam pipe, where the detector occupancy is higher, the Inner Tracker (IT) [76], consisting of silicon strip sensors, is placed. In the outer regions the modules of the Outer Tracker (OT) are situated. Each of the three stations of the IT includes four detector boxes, arranged around the beam pipe, see Fig. 3.11. Each box houses four detector layers including seven modules each. As for the TT the two inner layers are tilted by  $\mp 5^\circ$  with respect to the vertical direction<sup>9</sup>. The modules on the bending plane consist of two sensors while those on the vertical plane include only a single sensor. The IT silicon sensors are single-sided  $p^+$ -on- $n$  sensors with a size of 7.6 cm  $\times$  11 cm and thickness of 320  $\mu\text{m}$  in the case of one-sensor modules and 410  $\mu\text{m}$  in the case of two-sensors modules. With a strip pitch of 198  $\mu\text{m}$  and 384 readout strips per

<sup>9</sup>For all the detectors downstream of the magnet, the detector layers are placed vertically with respect to the center of gravity, not the LHCb coordinate system. This simplifies the construction of the support structures of the detectors, that are larger and heavier with respect to the detectors upstream of the magnet.

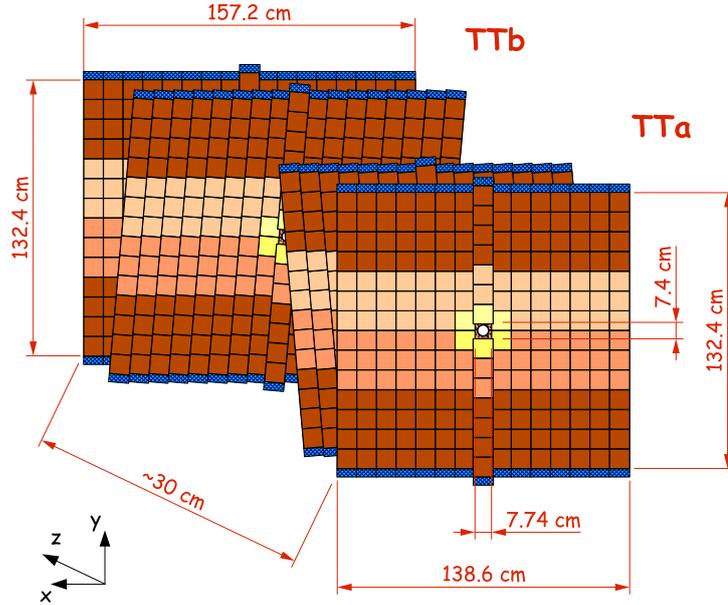


Figure 3.9: Layout of the four TT layers. Second and Third layers are tilted by  $\mp 5^{\text{deg}}$ . Different colors correspond to different readout sectors, the blue rim indicates the position of the readout electronics. Figure taken from [75].

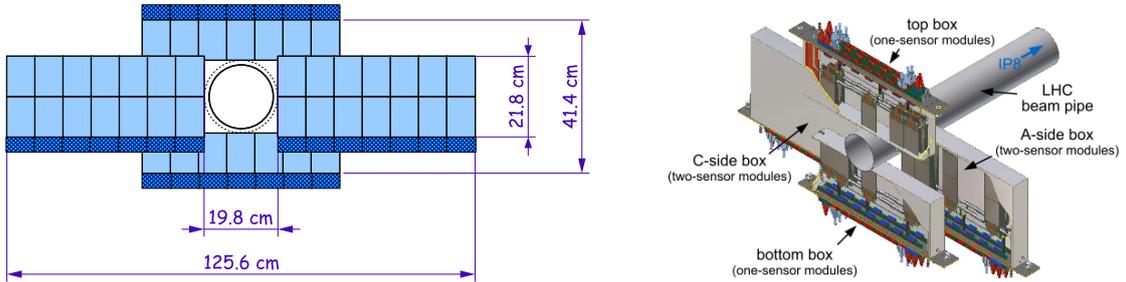


Figure 3.10: Layout of the IT. (Left) The silicon sensors are highlighted with light blue color, while the readout electronics is identified with dark blue color. (Right) The four IT boxes and the layers structure are shown. Figures taken from [75], [71]

sensors, the achieved resolution is similar to the TT resolution. In the case of the IT the readout electronics and the cooling system, cannot be placed outside the acceptance, for this reason the material budget of the IT is not uniform (up to  $0.035 X_0$ <sup>10</sup> per station close to the beam pipe while for narrow regions close to the cooling pipes it can add up to  $0.3 X_0$  per station).

<sup>10</sup>The *radiation length*  $X_0$  is defined as the average distance that an electron has to fly in a given material to see its energy reduced by a factor  $e$ .

## Outer Tracker

The outer part of the T stations is named Outer Tracker (OT) [77], [78] and is a gaseous straw tube detector. It covers an area of approximately  $5 \times 6 \text{ m}^2$  with 12 double layers of straw tubes. Each module of the OT is composed of two staggered layers (monolayers) of 64 drift tubes. In the longest modules, the monolayers are split in the middle into two independent readout sections composed of individual straw tubes. Both sections are read out from the outer ends. The splitting in two sections is done at a different position for the two monolayers to avoid insensitive regions in the middle of the module. Long modules have an active length of 4850 mm and contain twice 128 straws in the upper and lower part. Short modules are used to cover the region above and below the beam pipe. They contain 128 straws and are read from one side only. One detector layer is built from 14 long and 8 short modules. The region closer to the beam pipe is covered by the IT. The separation between the region covered by OT and IT respectively was chosen such that the occupancy in the OT does not exceed the 10% for an instantaneous luminosity of  $2 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ . Modules are arranged in three stations, each of those includes four layers, with the already familiar so-called  $x - u - v - x$  geometry, for the first and last layer the modules are oriented vertically, while the inner layers are tilted by  $\mp 5^\circ$ . The complete OT detector consists of 168 long and 96 short modules and comprises 53,760 single straw-tube channels, arranged in 24 straw layers.

Straw tube are 2.4 m long, with an inner diameter of 4.9 mm, and they are filled with a gas mixture of  $Ar/CO_2/O_2$  (70/28.5/1.5) which guarantees a drift-time below 50 ns. The anode wire is set to +1550 V and it is made of gold plated tungsten of 25  $\mu\text{m}$  diameter, while the cathode consists of an inner foil of electrically conducting carbon doped polyimide film and an outer foil of the same material laminated together with a layer of aluminium. Despite the fact that the OT readouts and support structures can be placed outside the LHCb acceptance, the OT contributes significantly to the material budget (up to 3.2% of  $X_0$  per station).

The position of the so-called *hits*, points belonging to the trajectory of a particle, in the OT is determined by measuring the drift-time to the wire of the ionization clusters created in the gas inside the straws. The distance from the tracks to the anode wire can be determined by using the so-called *tr-relation*, which depends on the gas mixture used. The measured drift-time can be affected by any change in the readout electronics or by a change in the synchronization between the LHCb clock and the collision time. These time offsets are regularly monitored. The spatial position of the OT detector also has an impact in the determination of the hit position, for this reason the OT spatial position is periodically ascertained with the alignment procedures. The effective time resolution of the OT is about 3 ns, which corresponds to spatial hit position resolution of 200  $\mu\text{m}$ . The hit efficiency is above 99%.

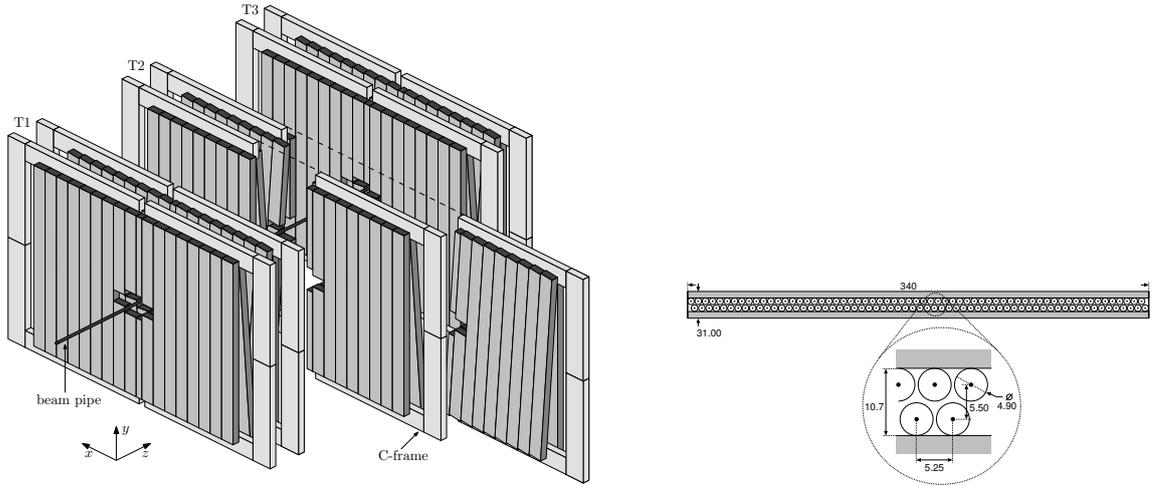


Figure 3.11: OT. (*Left*) view of the T stations and layers configuration. (*Right*) Cross section of an OT module, showing the two straw-tubes monolayers structure.

### 3.2.2 Particle identification system

The particle identification system takes advantage of different physical principles to identify the type of the particles transversing LHCb. Two Ring Imaging Cherenkov detectors (RICH1 and RICH2) exploit the so-called Cherenkov radiation emitted by a particle transversing a medium with a speed higher than the speed of light in the medium, to discriminate between the different hadron species. The calorimeter system aims to measure the energy of the particles by fully absorbing them. The LHCb calorimeter system consists of an Electromagnetic calorimeter (ECAL) designed to detect electrons and photons and a Hadronic calorimeter (HCAL) to measure the energy of the hadrons passing through. They are assisted by two smaller detectors, the Scintillating Pad Detector (SPD) and Pre-Shower detector (PS). The muon system is devoted to the detection of the muons. All the muon stations but one are placed in the most remote position from the interaction point, and separated by iron shields, allowing for the detection of only the particles unaffected by the passage through this amount of material. This property differentiates the muons from the other charged particles.

#### Ring Imaging Cherenkov detectors

The Ring Imaging Cherenkov (RICH) detectors exploit the Cherenkov effect to identify particles, mainly to provide a discrimination between pions and kaons, which is challenging considering the intense hadron production at the LHC. A charged particle transversing with velocity  $v$  a medium with refraction index  $n$ , when  $v$  is larger than the speed of light in a medium ( $c' = c/n$ ), photons are emitted in the shape of a cone, with opening angle the Cherenkov angle  $\theta_{\text{Ch}}$ , that depends on the velocity of the particle and on  $n$ :  $\cos\theta_{\text{Ch}} = 1/\beta n$  (with  $\beta = v/c$ ). Combining the measured opening angle with the momentum of the

particle  $p$  measured by the tracking system, allows to calculate the mass of the particle:

$$\cos\theta_{\text{Ch}} = \frac{1}{n} \sqrt{\left(\frac{m}{p}\right)^2 + 1} \quad (3.5)$$

LHCb uses two RICH detectors [79], to cover different momentum ranges. RICH1, placed before the magnet, exploits two different radiators during Run-I: aerogel ( $n = 1.03$ ) and  $C_4F_{10}$  ( $n = 1.0014$ ), to ensure  $K - \pi$  separation in a momentum range from approximately  $1 \text{ GeV}/c$  to  $60 \text{ GeV}/c$ . Fig 3.13 shows the different Cherenkov bands for muons, pions, kaons and protons when a  $C_4F_{10}$  radiator is used. RICH1 covers the full LHCb angular acceptance of 25-300 mrad in x-direction and 25-250 mrad in y-direction. RICH2, located instead behind the tracking stations, uses  $CF_4$ , with  $n = 1.0005$  as radiator, and covers a momentum range between  $15 \text{ GeV}/c$  and  $100 \text{ GeV}/c$ . Both RICH detectors use two sets of mirrors. The *primary* mirrors are spherical and reflect the Cherenkov photons on one of the *secondary* mirrors, that are much planer and deflect the photons out of the LHCb acceptance, where the plane of the photon detectors are placed. That plane coincides with the focal plane of the given part of the optical system (see Fig. 3.12). The photon sensors used are Hybrid Photo Detectors (HPD), that detect photons in the 200-600 nm spectrum. As example, the identification (ID) efficiency for kaons is almost 95%, with about 5% probability of misidentification of the pions, for a particle momentum range roughly between 10 and 40  $\text{GeV}/c$ , the precise momentum dependence is reported in Fig. 3.13.

### Calorimeter system

The LHCb calorimeter system [81] measures energy depositions and their positions and performs electron, photon and hadron identification. In addition the calorimeter system is used in the first level (L0) trigger. This puts strict requirements on the readout speed of the calorimeters, since the L0 decision needs to occur after only  $4 \mu\text{s}$ . Charged particles and photons when interacting with the detector material produce electromagnetic showers via bremsstrahlung and pair production. Hadrons instead produce hadronic showers. The LHCb calorimeter system consists of alternating layers of absorbing material placed between layers of active scintillating material. The showers develop in the absorbing layers, while the particles produce photons in the scintillating layers, that are then read out with photomultiplier tubes. The calorimeter system of LHCb consists of four sub-detectors: A scintillating pad detector (SPD), which is used to separate photons from electrons, followed by 15 mm of lead absorber and subsequent a pre-shower detector (PS). The PS is used to separate charged pions from electrons. An electromagnetic calorimeter (ECAL), provides informations about the energy and the position of the electromagnetic showers produced by electrons and photons, and finally a hadronic calorimeter (HCAL) measures the energy and the shape of showers produced by the hadrons. For an optimal energy resolution, electromagnetic showers from high energy photons have to be fully contained by the ECAL, which has been designed to have a thickness of 25 radiation lengths. Such a good energy resolution is not needed for triggering on hadrons. Moreover, an ECAL with

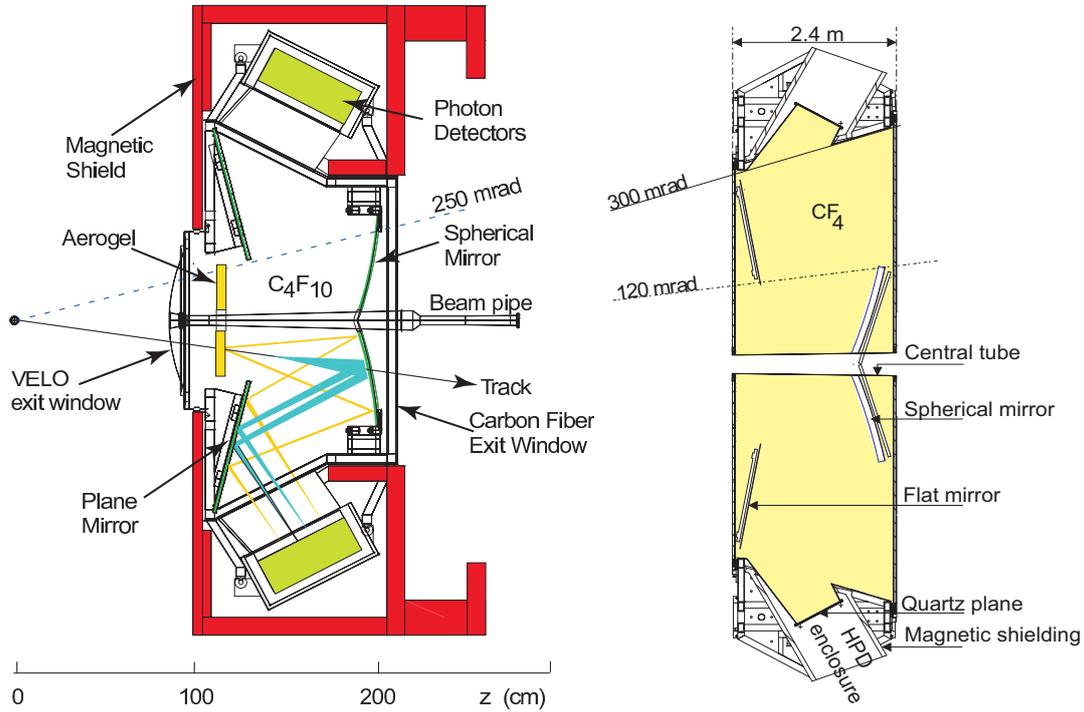


Figure 3.12: Schematic view of RICH1 (*left*) and RICH2 (*right*) detectors and their optical systems.

this thickness ensures that almost all the electrons are stopped at that stage, without transversing the downstream detectors. So, the HCAL thickness is only 5.6 interaction lengths, minimizing the occupied space. The whole calorimeter system is segmented in the x-y plane and, due to the larger particle density close to the beam pipe, the segmentation increases in dimension moving away from the beam-pipe with a projective design.

*Scintillating Pad Detector and Pre-Shower.* SPD and PS are located behind the first muon station M1 on the two opposite sides of a 15 mm thick lead absorber. Both detectors are constructed out of scintillating pads. The detectors are subdivided in three different regions with different segmentation. The segmentation decreases from 4 cm  $\times$  4 cm large pads in the inner region to 12 cm  $\times$  12 cm large pads in the outer region. The purpose of the SPD detector is to separate electrons from photons. The principle behind the electron-photon separation is that while electrons will give a signal in the SPD the electrically neutral photons don't. The photon as electron misidentification rate is found to be below 3%. The PS detector is built for electron-pion separation.

*Electromagnetic Calorimeter.* The sampling structure of the ECAL is built from alternating layers of 2 mm thick lead absorber and 4 mm thick scintillator material read out via wavelength shifting fibers. The segmentation of the ECAL is the same as SPD/PS,

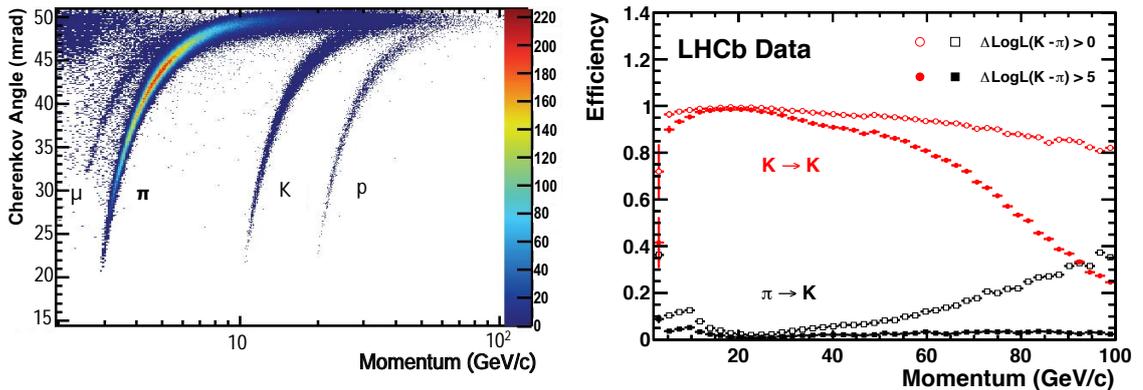


Figure 3.13: *(left)* Reconstructed Cherenkov angle  $\theta_{\text{Ch}}$  in function of momentum of the tracks transversing  $C_4F_{10}$  radiator. The Cherenkov bands for the different particles are visible. *(right)* Kaon identification efficiency and pion misidentification rate as function of the particle momentum, measured on data. The different marker indicate the different  $\Delta\log\mathcal{L}(K-\pi)$  used. More details about this quantities can be found in the reconstruction Sec. 3.3.2. Figs taken from reference [80], [59].

and features a finer granularity in the regions close to the beam pipe where the particle density is high. Three different types of modules are produced for this purpose. The ECAL has a depth corresponding to 25 radiation lengths  $X_0$  and 1.1 hadronic interaction lengths  $\lambda_I$ .

*Hadronic Calorimeter.* The HCAL is located after the ECAL and uses an iron/scintillator sampling structure. The segmentation of the HCAL is coarser than for the ECAL with the cells having a size of about  $13\text{ cm}\times 13\text{ cm}$  in the inner and  $26\text{ cm}\times 26\text{ cm}$  in the outer region. In total the HCAL has a depth corresponding to 5.6 hadronic interaction lengths  $\lambda_I$ .

## Muon system

The muon chambers are vital for muon identification and triggering of B meson decays containing muons in the final state. The LHCb muon system [83], [84] consists of five muon stations (M1-M5), see Fig. 3.15. M1 is located upstream of the calorimeters to improve the  $p_T$  resolution in the muon trigger by minimizing uncertainties caused by multiple scattering in the calorimeter material. Iron absorbers of 80 cm thick material are placed between the muon stations M2-M5 located after the calorimeter to filter out all particles except muons. To traverse all muon chambers and absorbers muons need to have a minimum momentum of 6 GeV. Each of the muon stations is divided in four regions R1-R4 with finer segmentation in the regions with higher particle multiplicity close to the beam pipe. All regions use multiwire proportional chambers (MWPC) except for the region of M1 closest to the beam pipe, which uses triple-GEM detectors. This choice is

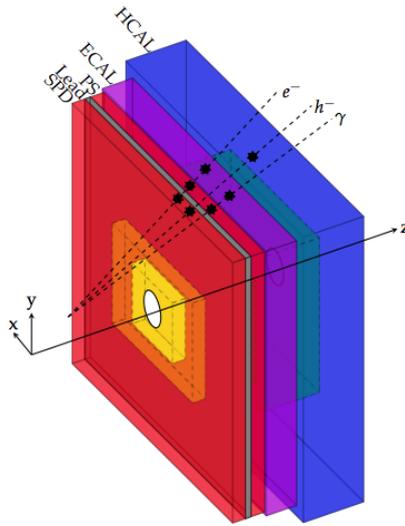


Figure 3.14: Layout of the LHCb calorimeter system. The segmentation and the interactions of different particle species are also shown. The relative dimensions of the ECAL and HCAL are correct, but the z-dimension of the SPD/PS is exaggerated. Fig. taken from Ref. [82].

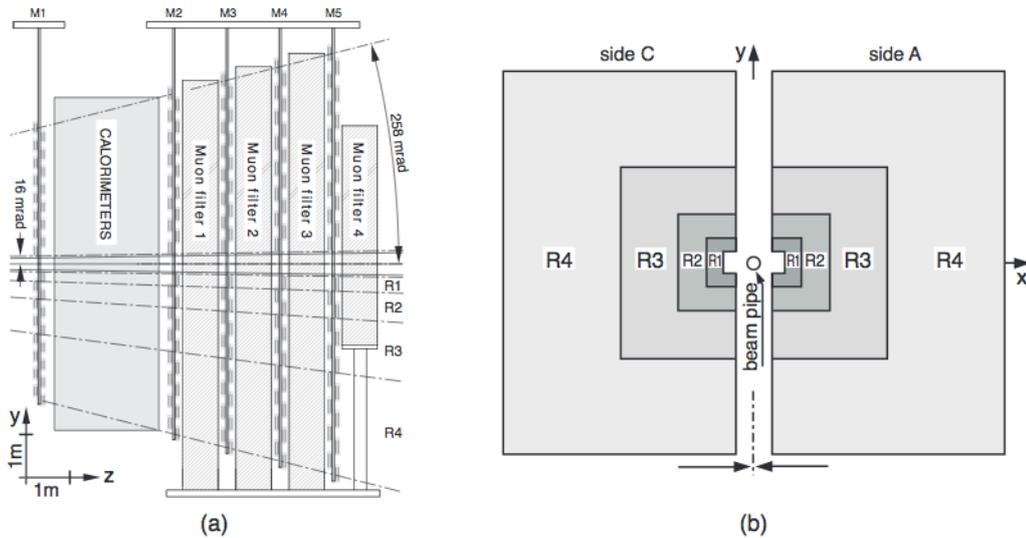


Figure 3.15: *Left* Side view of the LHCb muon system. *Right* Four regions R1-R4 of the stations layout.

due to the particle flux in that region that exceeds the radiation tolerance of the MWPC. Both detector types are able to collect the signal in less than 20 ns with an efficiency larger than 95%, which is important for the hardware trigger.

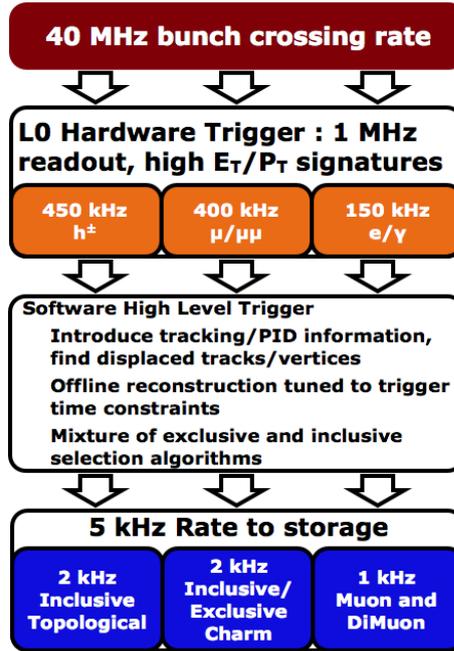


Figure 3.16: The three levels of the LHCb trigger system in 2012.

### 3.2.3 Trigger system

The trigger system is the heart of the experiment, since it determines which physics processes can be studied. The nominal bunch-crossing rate of the LHC is 40 MHz, definitely too high to allow data to be written to storage. The LHCb trigger system [85] has the task to reduce this rate to about 5 kHz while recording the  $pp$  interactions that are interesting for physics analysis. In particular the LHCb physics program focuses on heavy flavor physics, therefore the events are selected according to the characteristic features of  $c$ - and  $b$ -decays: displacement of the vertices and tracks of the decay products, high momenta and high transverse energy<sup>11</sup>. The LHCb trigger system, sketched in Fig.3.16 is developed in three stages, the first consists of a hardware trigger, while the second and third act at software level.

#### Level 0

The hardware trigger reduces the event rate to about 1 MHz, which is the maximum rate for the detector to be read out. Two different systems are used to select the interesting events for physics: L0-HADRON and L0-CALORIMETER.

<sup>11</sup>Most of the particles selected for physics analysis at LHCb have a momentum of about 20 GeV. The particles are required to have a momentum larger than a couple of GeV, according to the stage of the selection, trigger or offline selection for example, and decay mode. Sec. 5.1 reports the selection criteria applied to the samples used in this thesis.

L0-CALORIMETER is based on the informations collected from the calorimeter system (PS, SPD, ECAL and HCAL). The events with a transverse energy of a cluster of  $2 \times 2$  cells larger than a certain threshold are selected. The transverse energy is calculated as

$$E_T = \sum_{i=1}^4 E_i \sin \theta_i$$

where  $E_i$  is the energy deposited in the  $i$  cell and  $\theta_i$  is the angle between the  $z$ -axis of the LHCb experiment and the direction joining the primary interaction point to the cell  $i$ . Hadrons, photons and electrons are distinguished according to the calorimeter system where the particle deposits the energy <sup>12</sup>.

L0-MUON takes advantage of the detections in the muon stations. The two muons with highest transverse momentum in each quadrant of the muon stations are selected. The positions where the muons interacted with the material of the muon stations are used to reconstruct the trajectory of the particles. Assuming the muons originate from the primary  $pp$  interaction and get a single kink from the dipole magnet, the track and the transverse momentum can be inferred. Either the single muon or a muon pair with transverse momentum above a threshold determine a positive trigger decision.

If one or both systems of the hardware trigger return a positive trigger decision, then the full detector is read out by the data acquisition system (DAQ).

## High Level Trigger

Events selected by the L0 are then processed by the software trigger (HLT). The first stage, HLT1 takes advantage of the prominent features of  $c$ - and  $b$ - decays as the displacement of vertices and tracks and high momenta of the tracks. The rate is reduced to 40-80 kHz. The second stage, HLT2 exploits an event reconstruction of about the same quality as the off-line reconstruction. The difference between the reconstruction at HLT level and off-line has to be ascribed only to the timing requirements that are strict in the case of the HLT, and relaxed when the data has already been written to storage. At HLT2 level the event rate is lowered to 3 kHz in 2011 and 5 Hz in 2012. A number of different event selection strategies can be used at the HLT level, according to the topology of the decays of interest. A “trigger line” is a sequence of reconstruction and selection algorithms to trigger an event. Different trigger lines select decays with different signatures. Being fully developed at software level, the HLT ensures the freedom to broaden the physics program of the experiment in different directions and according to the different experimental conditions.

Specific choices of the trigger selection used for the  $a_{\text{sl}}^d$  measurement are discussed in Sec. 5.1.2.

---

<sup>12</sup>The different trigger decisions are called L0HADRON, L0PHOTON and L0ELECTRON

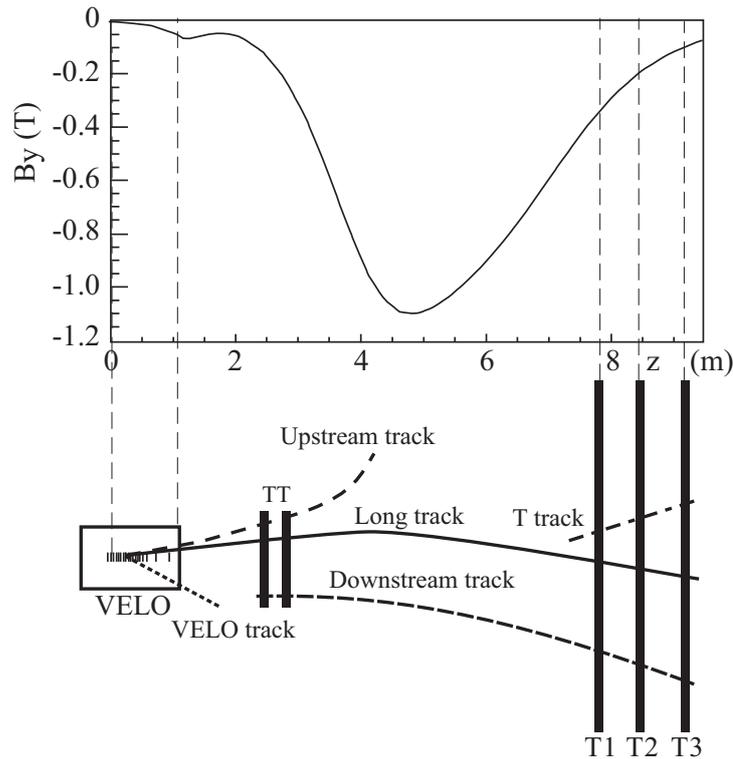


Figure 3.17: *Top* Strength of the main component of the magnetic field ( $B_y$ ) as function of the  $z$ -position. The position of the tracking detectors is also indicated. *Bottom* Illustration of the different type of tracks reconstructed at LHCb. Figure taken from [59].

### 3.3 Event Reconstruction

This section is a short description of the path from the electrical signals of the detector to the *reconstructed event* used for the physics analyses in LHCb. A good event for physics analysis contains information about one or more interesting decays. For example a  $b$ -hadron produced in the  $pp$  interaction and particles produced in its subsequent decay. Fundamental information to reconstruct and measure the decay are the reconstruction of the trajectories of the particles, the identification of the vertices, and the identification of the type of the particles involved.

#### 3.3.1 Track reconstruction

Charged particle trajectories, here referred to as tracks, are reconstructed in LHCb combining the electronic signals, hits, coming from the different tracking detectors (VELO, TT, IT and OT). The first step of the track reconstruction is the *pattern recognition*, that identifies a sequence of hits produced by the charged particle in the different detectors. According to the detectors transversed by the track, different type of tracks are distinguished

in LHCb, as shown in Fig. 3.17:

- *Velo tracks* are tracks with only VELO hits. They are used as input to reconstruct the long and upstream tracks. When they cannot be extended after the VELO, they are exploited in the primary vertex reconstruction.
- *T tracks* use hits located in the T stations. They are used as input for long and downstream tracks.
- *Long tracks* belong to particles that transversed the full tracking system. They are defined combining hits in the VELO and in the T stations: this leads to the most precise momentum resolution. Hits in the TT are added when possible, improving the track resolution and the rejection of tracks not associated to a real particle. They are the tracks mostly used for physics analyses, including the  $a_{\text{sl}}^d$  measurement.
- *Upstream tracks* are low momentum tracks that transversed only VELO and TT, and bent by the magnetic field out of the LHCb acceptance.
- *Downstream tracks* are reconstructed only in the TT and T stations. They are useful to reconstruct the decays of long lived resonances that decay after the VELO. An example are the neutral kaons. <sup>13</sup>

Two algorithms are used to reconstruct long tracks. Starting from the results of the VELO pattern recognition, a first option is to prolong the track into the T stations by using a “thin lens” approximation of the dipole field (*forward tracking*). A second option is to reconstruct the tracks in the T stations independently (*seeding*) and then match the VELO and the T stations segments to produce long tracks (*matching*) [86]. Then a Kalman Filter [87] is used to reconstruct the trajectory drawn by the hits. It takes into account effects from multiple scattering and energy loss due to ionization. The quality of the track is judged by means the  $\chi^2/dof$ . A selection requirement on this variable helps to remove the tracks not associated to real particles, so-called *ghosts*. The last step consists of the removal of tracks which share most of the hits, so called *clones*. Due to some redundancies in the reconstruction algorithm, it is possible that track segments of a same long track are also reconstructed as tracks. This step eliminates the duplicated segments and tracks.

### 3.3.2 Particle Identification

Particle Identification at LHCb takes advantage of different sub-systems, the RICH detectors, muon system and calorimeter system. The measured Cherenkov angle in the RICH detectors is combined with the measured momentum of the track, to provide the information about the mass of the particle (See Eq. 3.5). In the muon system the distance between the hits in the muon detectors and the extrapolation of the track in the same region provides an additional particle identification variable for the muons. Lastly the

<sup>13</sup>In this  $a_{\text{sl}}^d$  measurement  $D^+ \rightarrow \overline{K^0}\pi^+$  decays are used, but the data sample is restricted to long tracks (See Sec 5.4)

energy deposits in the ECAL are compared with the extrapolation of the tracks in the same region, providing an identification variable for the electrons. The calorimeter system helps to separate electrons from photons. The information is combined to provide the best achievable separation between the charged particles types ( $e, \mu, K, \pi, p$ ). The probability of a specific particle hypothesis  $x$  is calculated, and a likelihood value  $\mathcal{L}(x)$  is associate to the respective track. The likelihood for the hypothesis  $x$  is usually calculated relatively to the pion hypothesis, since pions are the most abundant species produced and detected at LHCb. The difference between the logarithms of the two likelihood is calculated  $\Delta \log \mathcal{L}_{x\pi} = \log \mathcal{L}_x - \log \mathcal{L}_\pi$ . This is the quantity referred to  $DLL_{x\pi}$  when discussing the selection of the data samples for the  $a_{\text{sl}}^d$  measurement (see Sec.5.1.4). The larger the  $DLL_{x\pi}$  value is, the more likely is for the particle to belong to the  $x$  species, rather than to be a pion. To select pions small or negative values can be required.



# Chapter 4

## Measurement strategy

This chapter describes the strategy and main experimental challenges of the  $a_{\text{sl}}^d$  measurement at LHCb. The idea is to use a sample of flavor specific (semileptonic) decays and measure the asymmetry in the number of decays in the final state  $f$  with respect to the number of decays in the  $CP$  conjugate final state  $\bar{f}$ . This asymmetry is induced by  $CP$  violation in mixing, but also from other sources. The first additional source is the production asymmetry between  $B^0$  and  $\bar{B}^0$  mesons, and the second is the asymmetry in the detection and reconstruction of particles in the final state with respect to the relative antiparticles. An additional experimental challenge is that semileptonic decays, since they involve a neutrino in the final state, can only be partially reconstructed. This implies the impossibility to reconstruct the  $B$  momentum and leads to a difficult signal versus background separation.

### 4.1 Methods to measure $a_{\text{sl}}^d$ and LHCb strategy

$CP$  violation in  $B$  mixing means that the probability that a  $B$  meson oscillates into a  $\bar{B}$  meson is different from the probability that a  $\bar{B}$  oscillates into a  $B$ , and the flavor specific or semileptonic asymmetry  $a_{\text{sl}}^q$  measures the asymmetry of these probabilities. It is defined in Eq. 2.57, here reported for simplicity:

$$a_{\text{sl}} = \frac{\Gamma(\bar{B} \rightarrow B \rightarrow f) - \Gamma(B \rightarrow \bar{B} \rightarrow \bar{f})}{\Gamma(\bar{B} \rightarrow B \rightarrow f) + \Gamma(B \rightarrow \bar{B} \rightarrow \bar{f})}$$

The use of semileptonic decays identifies the flavor of the  $B$  meson at the decay. The determination of the flavor of the  $B$  meson at the production is also needed in this expression. A first possible approach is used at the  $B$ -factories and by the  $D\bar{O}$  experiment, it exploits the symmetric production of  $B$  hadrons and the ability of identify and reconstruct the decay products of both  $B$  mesons. Assuming the  $B$ -factories case, the  $B^0\bar{B}^0$  pair created in the  $\Upsilon(4S)$  decay evolves coherently until one of the  $B$  mesons decays. When selecting semileptonic decays, the electric charge of the lepton in the final state identifies the flavor of the parent  $B$  meson. If one  $B$  meson has oscillated to its antiparticle, while the other  $B$  meson has not, the leptons produced in the decay have the same charge. The

same-sign dilepton events, allow to measure the  $CP$  asymmetry:

$$A_{\text{meas}} = \frac{N(l^+l^+) - N(l^-l^-)}{N(l^+l^+) + N(l^-l^-)} = a_{\text{sl}} \quad (4.1)$$

In the case of the  $e^+e^-$  collisions at the center-of-mass energy at the  $\Upsilon(4S)$  resonance peak ([41], [38], [39], [40], [45]), only  $B^0\bar{B}^0$  pairs are produced, therefore the asymmetry measured in Eq. 4.1 is  $a_{\text{sl}}^d$ . In the case of the  $D\bar{D}$  experiment, a general purpose detector exploiting the  $p\bar{p}$  collisions produced at Tevatron, Fermilab, both  $B^0$  and  $B_s^0$  mesons are produced. For this reason the inclusive asymmetry  $A_{\text{sl}}$  of Eq. 2.63 is measured at the  $D\bar{D}$  experiment [8]. This strategy relies on the assumption that the probability to produce a  $B$  meson of a given species is the same as to produce the corresponding antimeson. In the case of the  $B$ -factories, when the center-of-mass energy is at the  $\Upsilon(4S)$  resonance peak,  $B^0$  and  $\bar{B}^0$  are always produced in pairs. In the case of the  $D\bar{D}$  experiment an asymmetry between the number of mesons produced for a given species compared to the number of the relative anti-mesons is expected (see Sec. 3.1.5 for meson-antimesons production asymmetry at hadron colliders). Two factors help to cancel the meson-antimeson production asymmetry. First of all  $D\bar{D}$  exploits  $p\bar{p}$  collisions. This  $CP$  symmetric initial state favors the symmetric production of hadrons and anti-hadrons, but does not eliminate the possible production asymmetry, since hadronization processes depend on the kinematic region. For this reason the second important feature of the  $D\bar{D}$  experiment is the  $4\pi$  acceptance. In the individual kinematic regions (only the forward region for example) the meson-antimeson production asymmetry can take not negligible values, but when considering the whole acceptance, the production asymmetry cancels (for example the excess of  $B^0$  mesons in the forward region is expected to be compensated by the excess of  $\bar{B}^0$  mesons in the backward region).

As explained in Chapter 3, LHCb is a forward spectrometer, exploiting the  $pp$  collisions at the LHC. The  $a_{\text{sl}}^q$  measurement strategy needs to be different from the approach used at the  $B$ -factories and  $D\bar{D}$  experiment, because of the geometry of LHCb and of the production asymmetry between mesons and antimesons of the same species. The only method to determine the flavor of the  $B$  mesons at the production at LHCb is provided by the so-called *flavor tagging* algorithms. In Sec. 4.1.1 is explained why this is a challenging task at hadron colliders, and also how it can be avoided for the semileptonic asymmetries measurements. As first step, it is defined an asymmetry related to  $a_{\text{sl}}^q$  that doesn't need flavor-tagged decays, and this is crucial in order to aim for a competitive experimental sensitivity. Moreover, the  $B - \bar{B}$  production asymmetry and the detection asymmetries of all the final state particles need to be determined. Finally, a partial reconstruction technique needs to be developed to cope with the mass and momentum measurement of semileptonic decays.

### 4.1.1 Untagged asymmetry

The measurement of  $a_{\text{sl}}^d$  as defined in Eq. 2.57 requires a method to determine the flavor of the  $B$  mesons at the production. This is possible at hadron colliders exploiting either the

information from the hadronization and decay of the  $b$ -quark produced in the  $pp$  collision along with the  $b$ -quark belonging to the  $B$  meson of interest (so-called *Opposite Side (OS)* flavor tagging algorithms), or the fragmentation tracks of the  $B$  meson candidate of interest (*Same Side (SS)* flavor tagging algorithms). Based on this information, the flavor tagging algorithms, with a *tagging efficiency*,  $\epsilon$ , associate to each  $B$  candidate a flavor decision, along with an estimate of the probability of assigning an incorrect decision, so-called *mistag probability*, and denoted with  $\omega$ . The effective tagging power of the algorithms is defined as  $\epsilon(1 - 2\omega)^2$ . At hadron colliders, flavor tagging is a very challenging task, because the high number of fragmentation tracks and tracks from previous collisions. The tagging power at LHCb is of the order of  $\mathcal{O}(3\%)$ <sup>1</sup>. The tagging power affects the effective statistic of the data sample. If  $N$  decays are analyzed, and a tagger with effective power  $\epsilon(1 - 2\omega)^2$  is used, the effective number of events of the sample is  $\epsilon(1 - 2\omega)^2 N$ . In order to make a precision measurement of  $a_{\text{sl}}^d$ , high statistics data samples are needed. For this reason it is convenient to use a measurement strategy that does not need flavor tagging. For a better understanding of the reasons why a measurement strategy that does not exploit flavor tagging is preferred, few more considerations are necessary.

In the case of a *tagged* analysis, the following asymmetry between the final state  $f$  and  $\bar{f}$  is derived, using the definition of  $a_{\text{sl}}^d$  in Eq. 2.57, and the decay rates Eq. 2.53

$$A_{\text{meas}} = \frac{\Gamma(\bar{B} \rightarrow B \rightarrow f) - \Gamma(B \rightarrow \bar{B} \rightarrow \bar{f})}{\Gamma(\bar{B} \rightarrow B \rightarrow f) + \Gamma(B \rightarrow \bar{B} \rightarrow \bar{f})} = a_{\text{sl}}, \quad (4.2)$$

where the effects of the production asymmetry of  $\bar{B}$  antimesons with respect to  $B$  mesons and of the different efficiency of detecting and reconstructing the final state  $f$ , compared to the final state  $\bar{f}$  are not included<sup>3</sup>. The asymmetries  $A_{\text{meas}}$  can be measured by determining the effective number of signal decays with oscillation of the  $B$  meson. Each event needs to be weight with the probability of the tagging decision to be correct.

<sup>1</sup>This tagging power can be compared to the tagging power at the  $B$ -factories,  $\mathcal{O}(30\%)$

<sup>2</sup>In addition, in the case of  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  semileptonic decays, the background of  $B^+ \rightarrow D^- \mu^+ \nu_\mu X^+$  decays affects also the tagging performances. The additional pion track of the  $B^+$  decays can be identified as belonging to the signal from the SAME SIDE PION tagging algorithm, leading to a wrong tagging decision, given that indicates a opposite quark content for the  $B$  meson candidate. A similar problem is experienced with the OPPOSITE SIDE VERTEX CHARGE algorithm, that uses an inclusive reconstruction of the vertex of the  $B$  meson produced in association of the  $B$  candidate of interest. The charge of the particles included in the vertex is calculated. The decay products of the  $B$  candidate of interest are excluded when reconstructing the vertex. This problem is overcome in the measurement of the  $B^0 - \bar{B}^0$  mixing frequency using semileptonic decays at LHCb by using a multivariate analysis isolation tool to search for additional charged tracks in the vicinity of the signal. The same approach is complicated to use for the  $a_{\text{sl}}^d$  measurement. Since the signal and control samples may respond differently when applying the isolation tool, the detection asymmetries measured using the control samples can result different from the detection asymmetry of the signal. This could lead to a biased  $a_{\text{sl}}^d$  measure. A different solution would need to be found in order to measure the semileptonic asymmetry. A possibility is to exclude the usage of tagging algorithms with problematic behavior, obtaining a further reduction of effective statistic of the signal sample.

<sup>3</sup>In order to describe the possible strategy of *tagged* analysis, the following asymmetries between the

Decays of neutral  $B$  mesons (namely  $B^0$  and  $B_s^0$ ) to a final state  $f$  are considered. The *untagged decay rates* are defined as

$$\Gamma[f, t] = \Gamma(B(t) \rightarrow f) + \Gamma(\bar{B}(t) \rightarrow f) \quad (4.3)$$

$$\Gamma[\bar{f}, t] = \Gamma(B(t) \rightarrow \bar{f}) + \Gamma(\bar{B}(t) \rightarrow \bar{f}) \quad (4.4)$$

Using the expressions in Eq. 2.53, it is possible to calculate the untagged decay rates:

$$\Gamma[f, t] = \mathcal{N}_f |A_f|^2 \frac{e^{-\Gamma t}}{2} \left\{ \left(1 + \frac{1}{1-a}\right) \cosh\left(\frac{\Delta\Gamma t}{2}\right) + \left(1 - \frac{1}{1-a}\right) \cos(\Delta Mt) \right\} \quad (4.5)$$

$$\Gamma[\bar{f}, t] = \mathcal{N}_f |A_f|^2 \frac{e^{-\Gamma t}}{2} \left\{ (1 + (1-a)) \cosh\left(\frac{\Delta\Gamma t}{2}\right) + (1 - (1-a)) \cos(\Delta Mt) \right\} \quad (4.6)$$

and define the asymmetry:

$$\begin{aligned} A_{\text{meas}}(t) &= \frac{\Gamma[f, t] - \Gamma[\bar{f}, t]}{\Gamma[f, t] + \Gamma[\bar{f}, t]} \quad (4.7) \\ &= \frac{\left\{ \frac{2-a}{1-a} - (2-a) \right\} \cosh\left(\frac{\Delta\Gamma t}{2}\right) + \left\{ -\frac{a}{1-a} - a \right\} \cos(\Delta Mt)}{\left\{ \frac{2-a}{1-a} + (2-a) \right\} \cosh\left(\frac{\Delta\Gamma t}{2}\right) + \left\{ -\frac{a}{1-a} + a \right\} \cos(\Delta Mt)} \\ &= \frac{a}{2} - \frac{a}{2} \frac{\cos(\Delta Mt)}{\cosh(\Delta\Gamma t/2)} + \mathcal{O}(a^2) \\ &= \frac{a_{\text{sl}}}{2} - \frac{a_{\text{sl}}}{2} \frac{\cos(\Delta Mt)}{\cosh(\Delta\Gamma t/2)} \quad (4.8) \end{aligned}$$

where the definition of  $a_{\text{sl}}$  in Eq. 2.57 is used. Measurements of this asymmetry do not require to identify the flavor of the  $B$  meson at production, since  $A_{\text{meas}}(t)$  is simply the time-dependent asymmetry between the charge-conjugated final states. However, Eq. final state  $f$  and  $\bar{f}$  are derived, using the definition of  $a_{\text{sl}}^d$  in Eq. 2.57, and the decay rates Eq. 2.53

$$\begin{aligned} A_{\text{osc}} &= \frac{\Gamma(\bar{B} \rightarrow B \rightarrow f) - \Gamma(B \rightarrow \bar{B} \rightarrow \bar{f})}{\Gamma(\bar{B} \rightarrow B \rightarrow f) + \Gamma(B \rightarrow \bar{B} \rightarrow \bar{f})} = a_{\text{sl}} + A_{\text{P}} + A_{\text{D}}, \\ A_{\text{NOTosc}} &= \frac{\Gamma(B \rightarrow f) - \Gamma(\bar{B} \rightarrow \bar{f})}{\Gamma(B \rightarrow f) + \Gamma(\bar{B} \rightarrow \bar{f})} = A_{\text{D}} - A_{\text{P}}. \end{aligned}$$

The two asymmetries refer to events where the  $B$  meson oscillated into its antimeson ( $A_{\text{osc}}$ ) and where the  $B$  meson did not oscillate ( $A_{\text{NOTosc}}$ ). With  $A_{\text{P}}$  the production asymmetry of  $\bar{B}$  antimesons with respect to  $B$  mesons is denoted,  $A_{\text{P}} \equiv \frac{N(\bar{B}) - N(B)}{N(\bar{B}) + N(B)}$ , and  $A_{\text{D}}$  indicates the asymmetry in the efficiency of detecting and reconstructing the final state  $f$ , compared to the final state  $\bar{f}$ ,  $A_{\text{D}} \equiv \frac{\epsilon(f) - \epsilon(\bar{f})}{\epsilon(f) + \epsilon(\bar{f})}$ . The asymmetries  $A_{\text{osc}}$  and  $A_{\text{NOTosc}}$  can be measured by determining the effective number of signal decays with and without oscillation of the  $B$  meson. Each event needs to be weight with the probability of the tagging decision to be correct. An external input as the value of  $A_{\text{D}}$  (or equivalently  $A_{\text{P}}$ ) is needed. The same limit affects the  $a_{\text{sl}}^d$  measurement that does not relies on flavor tagging techniques.

4.7, assumes to have an equal number of  $B$  and  $\bar{B}$  mesons at production, but this is not true in the case of LHCb. No difference between the reconstruction efficiencies of  $f$  and  $\bar{f}$  is here assumed. In order to compare the expected sensitivities of the tagged and of the *untagged* analysis <sup>4</sup>, a sample of  $N = 2.5 \times 10^6$  events is considered. The effective statistics of the tagged sample, considering a tagging power of 2.5% [88] is about  $N_{\text{eff}} = \epsilon(1 - 2\omega)^2 N = 6.25 \times 10^4$ , but since only events where the  $B$  meson changed flavor between production and decay provide information about  $a_{\text{sl}}^d$ , the effective statistics decreases by a factor 1/3 in the case of the  $B^0$  mesons <sup>5</sup>, giving  $N_{\text{eff}} = 2.1 \times 10^4$ . On the other hand, for the untagged sample  $N_{\text{eff}} = N$ . The uncertainty on an asymmetry obtained by simply measuring the event yields, as in Eq.4.2, is  $1/\sqrt{N_{\text{eff}}}$ , which gives an estimate of  $7 \times 10^{-3}$  for the uncertainty on  $a_{\text{sl}}^d$  with the tagged analysis. In the untagged analysis instead, the uncertainty on  $a_{\text{sl}}^d$  needs to be scaled at least with a factor of two, from Eq. 4.7, that leads to an uncertainty of  $1.3 \times 10^{-3}$ . This factor is expected to be more than two, given that the value of  $a_{\text{sl}}^d$  is estimated with a more complicated fit procedure, including 14 floating parameters (a complete description is given in Chapter 8). Considering these additional effects that contribute to the loss of precision, the statistical uncertainty expected with a tagged analysis is a factor 4.2 larger than the statistical uncertainty expected with an untagged analysis. For this reason the untagged analysis strategy is used. The following sections describe the challenges of this measurement at LHCb.

### 4.1.2 $B^0$ - $\bar{B}^0$ production asymmetry

As explained in Sec 3.1.5, in  $pp$  collisions, the production rate of specific  $B$  mesons is different from the production rate of the corresponding antimesons<sup>6</sup>. The decay rates can be written for different production rates for  $B$  and  $\bar{B}$  mesons  $N(B) \neq N(\bar{B})$ :

---

<sup>4</sup>With *untagged* analysis is meant the analysis strategy that exploits the untagged decay rates (Eq. 4.3) and does not use flavor tagging techniques.

<sup>5</sup>Given  $\Delta M_d = 0.51$  [6], about 1/3 of the events are identified as events where the  $B$  meson oscillate to its antiparticle (or vice versa) before decaying [88]. The decay time interval considered ranges from 0.4 ps to 15 ps. Different is the picture for  $B_s^0$  decays, where the high mixing frequency  $\Delta M_s = 17.8$  [6] ensures that half of the  $B$  candidates changed flavor between production and decay [89].

<sup>6</sup>While  $b$  and  $\bar{b}$  are produced in pairs, the hadronization in the neighborhood of the proton remnants leads to a different production rate for a given  $B$  species and its charge-conjugated counterpart

$$\Gamma(B(t) \rightarrow f) = N(B)\mathcal{N}|A_f|^2 \frac{e^{-\Gamma t}}{2} \left\{ \cosh \frac{\Delta\Gamma t}{2} + \cos(\Delta Mt) \right\} \quad (4.9)$$

$$\Gamma(\bar{B}(t) \rightarrow f) = N(\bar{B})\mathcal{N}|A_f|^2 \frac{e^{-\Gamma t}}{2(1-a)} \left\{ \cosh \frac{\Delta\Gamma t}{2} - \cos(\Delta Mt) \right\} \quad (4.10)$$

$$\Gamma(B(t) \rightarrow \bar{f}) = N(B)\mathcal{N}|A_f|^2 \frac{e^{-\Gamma t}}{2} (1-a) \left\{ \cosh \frac{\Delta\Gamma t}{2} - \cos(\Delta Mt) \right\} \quad (4.11)$$

$$\Gamma(\bar{B}(t) \rightarrow \bar{f}) = N(\bar{B})\mathcal{N}|A_f|^2 \frac{e^{-\Gamma t}}{2} \left\{ \cosh \frac{\Delta\Gamma t}{2} + \cos(\Delta Mt) \right\} \quad (4.12)$$

With these decay rates it is possible to calculate the untagged decay rates:

$$\Gamma[f, t] = \mathcal{N} \frac{|A_f|^2 e^{-\Gamma t}}{2} \left\{ \left( N(B) + \frac{N(\bar{B})}{1-a} \right) \cosh \frac{\Delta\Gamma t}{2} + \left( N(B) - \frac{N(\bar{B})}{1-a} \right) \cos(\Delta Mt) \right\} \quad (4.13)$$

$$\Gamma[\bar{f}, t] = \mathcal{N} \frac{|A_f|^2 e^{-\Gamma t}}{2} \left\{ (N(B)(1-a) + N(\bar{B})) \cosh \frac{\Delta\Gamma t}{2} + (N(\bar{B}) - N(B)(1-a)) \cos(\Delta Mt) \right\} \quad (4.14)$$

The meson-antimeson production asymmetry  $A_P$  is defined as:

$$A_P \equiv \frac{N(\bar{B}) - N(B)}{N(\bar{B}) + N(B)} \quad (4.15)$$

Using the untagged decay rates in Eq. 4.13 and 4.14 and the definition of  $A_P$  above (Eq. 4.15) instead of  $N(\bar{B})$  and  $N(B)$ , the time dependent measured asymmetry relates to  $a_{\text{sl}}$  as follows <sup>7</sup>:

$$A_{\text{meas}}(t) = \frac{\Gamma[f, t] - \Gamma[\bar{f}, t]}{\Gamma[f, t] + \Gamma[\bar{f}, t]} = \frac{a_{\text{sl}}}{2} - \left( A_P + \frac{a_{\text{sl}}}{2} \right) \frac{\cos(\Delta Mt)}{\cosh(\Delta\Gamma t/2)} \quad (4.16)$$

As already noticed, the production asymmetry  $A_P$  and  $a_{\text{sl}}$  need to be determined simultaneously, since its effect modifies the observed charge asymmetry, and it depends on the kinematics of the selected sample. At this point it is necessary to distinguish between the  $a_{\text{sl}}^d$  and  $a_{\text{sl}}^s$  measurements at LHCb. The difference is mainly given by the mixing frequency of the mesons, that allows different strategies in the two cases. For the  $a_{\text{sl}}^s$  measurement a time-integrated analysis is performed. After integrating  $A_{\text{meas}}$  over decay time, the time dependent term in Eq. 4.16 is suppressed to a negligible level by the large value of  $\Delta M_s$  and a production asymmetry expected to be at the percent level. In the case of the slowly oscillating  $B^0$  meson, a time-dependent analysis needs to be performed.

<sup>7</sup> $\mathcal{O}(a^2)$  are neglect as before, as in  $a_{\text{sl}}$  definition in Eq. 2.57

### 4.1.3 Detection asymmetries

The detector can have a different efficiency for reconstructing the final state  $f$  compared to  $\bar{f}$ . Different detection and reconstruction efficiencies could mimic a  $CP$  asymmetry. The untagged decay rates can be written including efficiency factors  $\epsilon(f)$  and  $\epsilon(\bar{f})$  for the final states  $f$  and  $\bar{f}$

$$\Gamma[f, t] = \epsilon(\bar{f})\mathcal{N}\frac{|A_f|^2 e^{-\Gamma t}}{2} \left\{ \left( N(B) + \frac{N(\bar{B})}{1-a} \right) \cosh\frac{\Delta\Gamma t}{2} + \left( N(B) - \frac{N(\bar{B})}{1-a} \right) \cos(\Delta Mt) \right\} \quad (4.17)$$

$$\Gamma[\bar{f}, t] = \epsilon(f)\mathcal{N}\frac{|A_f|^2 e^{-\Gamma t}}{2} \left\{ (N(B)(1-a) + N(\bar{B}))\cosh\frac{\Delta\Gamma t}{2} + (N(\bar{B}) - N(B)(1-a))\cos(\Delta Mt) \right\} \quad (4.18)$$

The detection asymmetry between the final states  $f$  and  $\bar{f}$

$$A_D \equiv \frac{\epsilon(f) - \epsilon(\bar{f})}{\epsilon(f) + \epsilon(\bar{f})}. \quad (4.19)$$

From Eq. 4.17 - 4.18 and the  $A_{\text{meas}}$  definition in Eq. 4.7, the detection asymmetry  $A_D$  is an additive term to the raw asymmetry  $A_{\text{meas}}$ :

$$A_{\text{meas}}(t) = \frac{\Gamma[f, t] - \Gamma[\bar{f}, t]}{\Gamma[f, t] + \Gamma[\bar{f}, t]} = A_D + \frac{a_{\text{sl}}}{2} - \left( A_P + \frac{a_{\text{sl}}}{2} \right) \frac{\cos(\Delta Mt)}{\cosh(\Delta\Gamma t/2)} \quad (4.20)$$

Eq. 4.20 is underconstrained, for this reason  $A_D$  has to be measured independently. The determination of  $A_D$  is crucial for the  $a_{\text{sl}}^d$  determination.

To perform the  $a_{\text{sl}}^d$  measurement described in this thesis, data samples of  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays collected with LHCb during Run-I are used, where  $X$  denotes possible additional particles. These might be due to  $\tau^+$  decays into  $\mu^+ X$  or higher  $D$  resonance decays to  $D^{(*)-} X$ . The inclusion of charge-conjugate processes is implied. Considering the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  sample,  $D^-$  meson decays in the Cabibbo-favored mode  $D^- \rightarrow K^+ \pi^- \pi^-$  are selected. For the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  sample, decays of the  $D^{*-}$  into  $\bar{D}^0 \pi^-$  and subsequently  $\bar{D}^0$  decays into a  $K\pi$  pair are selected. These choices ensure no additional sources of  $CP$  violation in the decays. The reconstructed final state particles in both the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decay modes are  $K^+ \pi^- \pi^- \mu^+$  (see Fig. 4.2).

A precise evaluation of the detection asymmetries is vital: the decay modes used in this analysis contain a single charged kaon in the final state. The different interaction of the charged kaons with the detector material alone is expected to be at the percent

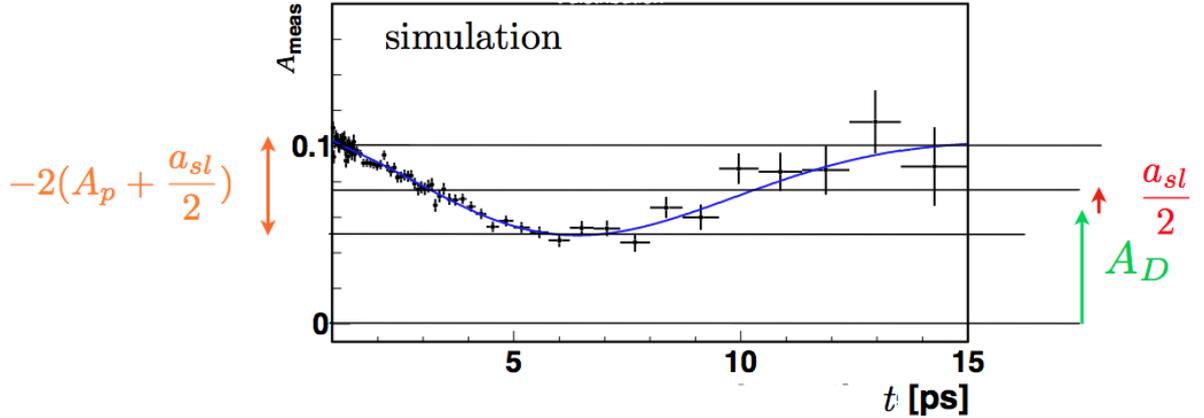


Figure 4.1: Expected behavior of  $A_{\text{meas}}(t)$ , assuming  $a_{\text{sl}}^d = 0.1\%$ ,  $A_P = -2.5\%$ ,  $A_D = 7.5\%$ . The amplitude of the oscillation is determined by the production asymmetry and  $a_{\text{sl}}^d/2$ , while  $A_D$  and  $a_{\text{sl}}^d/2$  are responsible for the offset of the oscillating curve.

level, to be compared with the target experimental precision of few per-mille. Focusing on the detection asymmetry measurement, in Eq. 4.19  $A_D$  is defined as the asymmetry in the detection efficiency of all the reconstructed particles in the final state  $f$  (with respect to  $\bar{f}$ ). In this analysis the detection asymmetry evaluation is performed separately for the oppositely charged  $K^+\pi^-$  and  $\pi^-\mu^+$  pairs. For  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  decays the pion with the lowest transverse momentum is paired with the kaon, and the pion with the highest momentum is paired with the muon. For  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$  decays the pion from the  $D^{*-}$  decay, also called *slow pion* is paired with the kaon, and the other pion is paired with the muon (See Fig. 4.2).

The  $K^+\pi^-$  asymmetry is measured by determining the charge asymmetries of Cabibbo-favored  $D^- \rightarrow K^+\pi^-\pi^-$  and  $D^- \rightarrow K^0\pi^-\pi^-$  decays originating from the primary interaction. The idea is that a measure of the charge asymmetry of the  $D^- \rightarrow K^+\pi^-\pi^-$  decays gives an estimate of the detection asymmetry of the final state  $K^+\pi^-\pi^-$  polluted by the contribution of the  $D$  meson production asymmetry.

The effects of the asymmetry in the number of  $D^-$  mesons produced compared to the  $D^+$  mesons, and the detection asymmetry of the additional pion present in the  $K^+\pi^-\pi^-$  final state, is measured by using a second control sample of  $D^- \rightarrow K^0\pi^-\pi^-$  decay, with the  $K^0$  decaying in two pions.

A further correction for the neutral kaon detection asymmetry is used, and is taken from [90]. This method accounts for all the effects that could possibly cause detection asymmetries: the different interaction cross section of positive and negative kaons with the detector material, inefficiencies and misalignment of the detector.

For the  $\pi^-\mu^+$  asymmetry, different contributions are evaluated with different methods: the muon identification and trigger component are measured using  $J/\psi$  decays. The pion identification is measured using  $D^0 \rightarrow K\pi$  decays from  $D^{*\pm} \rightarrow D^0\pi^\pm$  decay samples. The effect from different tracking efficiencies, due to the momentum imbalance between the

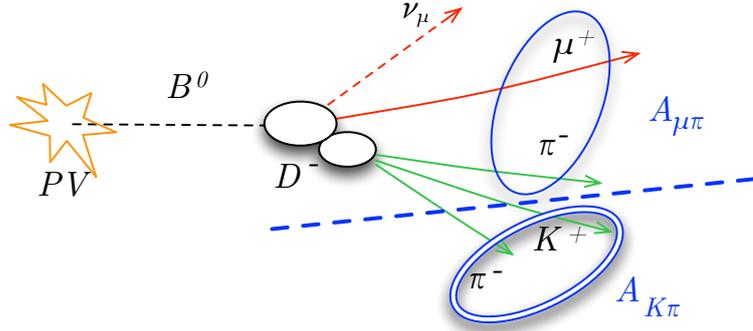


Figure 4.2: Sketch of a  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  decay. The illustrated separation in the final state particles is used for the evaluation of the detection asymmetries: the lower  $p_T$  pion is paired to the kaon while the higher  $p_T$  pion is paired to the muon.  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays have a similar topology and same particles in the final state.

muon and pion is reduced to a negligible level by re-weighting the signal. A detailed description of all the methods involved is given in Chapter 6.

The LHCb magnet polarity is reversed roughly every two weeks of data taking. When reversing the polarity of the magnet, the detection asymmetries related to the bending of particles due to the magnetic field change sign. The method used to evaluate the detection asymmetries does not rely on the cancelation obtained with the reversal of the magnet polarity. The full analysis is performed separately on data acquired with one polarity and with the other to control the detection asymmetries, and the result is obtained as average of the obtained results.

#### 4.1.4 Time-dependent fit strategy

For a time-dependent analysis, it is crucial to correctly determine the  $B$  decay time of the events.  $B$  hadrons are discriminated against other particles by means of their long lifetime. The  $B^0$  meson for example has a lifetime  $\tau(B^0) = (1.519 \pm 0.007)$  ps [17]. The method to calculate the lifetime of the  $B$  mesons detected by the LHCb detector uses the information of the measured momentum and mass of the  $B$  meson, and the *flight distance*,  $L$ . The latter is defined as the distance between the Primary Vertex (PV) and Decay Vertex (DV) of the  $B$  meson (see Fig. 4.3). The  $B$  meson momentum is instead calculated as the sum of the momenta of the final state particles. The decay time of the  $B$  meson is:

$$t = \frac{L_B \cdot M_B}{|\vec{p}_B|} = \frac{L_B}{\beta_B \gamma_B} \quad (4.21)$$

It can also be expressed as function of the relativistic quantities describing the Lorentz boost,  $\beta_B = v_B/c$  and  $\gamma_B = 1/\sqrt{1-\beta_B^2}$ . It is important to note that the decay

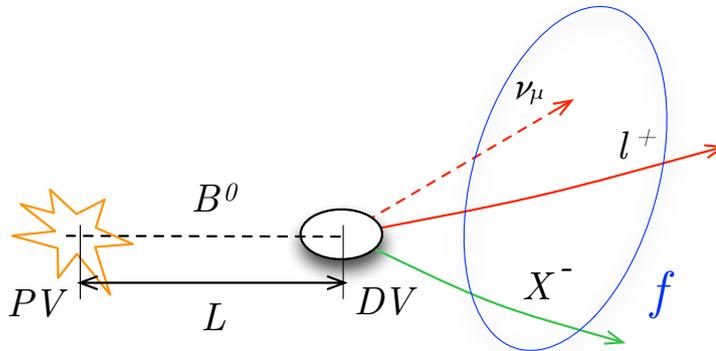


Figure 4.3: Sketch of a generic semileptonic  $B^0 \rightarrow f$  decay. The final state  $f$  is here written including a lepton, a neutrino and a hadronic part indicated with  $X$ . The flight distance  $L$  definition is emphasized.

time reconstruction is correct and precise when all the particles in the final state are reconstructed, since the momentum of the  $B$  meson is obtained as sum of the momenta of the particles reconstructed in the final state. In the case of semileptonic decays, the neutrino cannot be reconstructed at LHCb. Experiments with hermetic geometry, like ATLAS and CMS, can infer the energy of missing particles using their calorimeters, obtaining the so-called *missing*  $E_T$ . The forward coverage of LHCb excludes any possible estimation of missing  $E_T$ . The sum of the momenta of the daughter particles, in the case of semileptonic decays, does not determine the momentum of the  $B$  meson. It is possible to estimate the missing momentum on average, as shown in detail in Sec. 7.1.2, but not for the individual event. This leads to a poor momentum resolution, which dominates the time resolution. In order to cope with this problem several strategies can be used. Here the choice is to rely on simulated events to understand the resolution and the bias of the reconstructed  $B$  momentum,  $\vec{p}_B$ , in data. It should be emphasized that the raw charge asymmetry, related to  $a_{\text{sl}}^d$  by means of the formula 4.20, is not expected to be extremely sensitive to the time description. The partial reconstruction issues are expected to affect the decays to the final state  $f$  and to the final state  $\bar{f}$  similarly. The effect of an inaccurate method to reconstruct the  $B$  decay time is expected to have an impact on the value of  $\Delta M_d$  that is assumed for the measurement.

The poor time reconstruction is not the only consequence of partial reconstruction, more serious is the poor signal versus physics background separation. For fully reconstructed decays, the reconstructed invariant  $B$  mass provides a good handle to separate  $B$  decays from other decays. In the case of semileptonic decays, the poor momentum resolution leads to a poor  $B$  mass resolution (see for example Fig 5.6). Decays of the  $B^+$  meson

are a physical background that is difficult to distinguish from the signal in this analysis. Considering the largest statistics sample used for this analysis, the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$ , reconstructed with the final state  $K^+ \pi^- \pi^- \mu^+$ , the data sample includes about 10% of  $B^+$  decays in the final state  $K^+ \pi^- \pi^- \pi^+ \mu^+$ , where the  $\pi^+$  is not reconstructed. This component of the data sample is not oscillating and is affected by the production asymmetry of  $B^+$ ,  $B^-$  mesons, therefore it directly contributes to the raw charge asymmetry  $A_{\text{meas}}$ . The strategy pursued here is the inclusion of a realistic model of the  $B^+$  component.

Background originating from random combinations of charged tracks that accidentally satisfy the selection requirements, is the so-called *combinatorial background*. In the  $a_{\text{sl}}^d$  analysis this type of background and  $B$  decays not containing a  $D^\pm$  or  $D^0$  in the final state are not separated, as is explained in Sec. 8.2.

To determine the values of  $a_{\text{sl}}^d$  and  $A_P$  ( $B^0$ ), a binned maximum likelihood fit in the  $B$  decay time, charge of the final state, and  $D^-(\bar{D}^0)$  mass for the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  decay mode (for  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decay mode) is performed. Combinatorial and  $B^+$  background components are included in the fit. The effect of other possible backgrounds is evaluated as systematic uncertainty.

#### 4.1.5 Final result determination

The full analysis is performed independently on the different magnet polarity, different center-of-mass energies and different decay modes data samples. This strategy allows for an optimal control of the detection asymmetries. Moreover the comparison of the fit results among the different sub-samples represents a powerful crosscheck of the consistency and reliability of the analysis strategy used. In addition, the  $B - \bar{B}$  meson production asymmetry is expected to show a dependence in function of the center-of-mass energy. For this reason the parameter measuring the production asymmetry is allowed to take different values in 2011 and 2012 datasets. In order to obtain the final measured value of  $a_{\text{sl}}^d$ , the result on a single year is obtained by performing an arithmetic average of the results obtained on the single magnet polarity datasets:

$$a_{\text{sl}}^d(\text{year}) = \frac{a_{\text{sl}}^d(\text{year}) \uparrow + a_{\text{sl}}^d(\text{year}) \downarrow}{2}, \quad (4.22)$$

$$\sigma_{a_{\text{sl}}^d}(\text{year}) = \frac{\sqrt{\sigma_{a_{\text{sl}}^d(\text{year})\uparrow}^2 + \sigma_{a_{\text{sl}}^d(\text{year})\downarrow}^2}}{2}, \quad (4.23)$$

where the arrows indicate the magnet polarity. The arithmetic average is used to obtain the cancellation of possible higher order detection asymmetries. The result for a single decay mode is provided by the weighted average of the results obtained on the dataset

split by year.

$$\begin{aligned}
a_{\text{sl}}^d(\text{decay mode}) &= \frac{1}{\frac{1}{\sigma_{2011}^2} + \frac{1}{\sigma_{2012}^2}} \left( \frac{a_{\text{sl},2011}^d}{\sigma_{2011}^2} + \frac{a_{\text{sl},2012}^d}{\sigma_{2012}^2} \right), \\
\sigma_{a_{\text{sl}}^d(\text{decay mode})} &= \sqrt{\frac{1}{\frac{1}{\sigma_{2011}^2} + \frac{1}{\sigma_{2012}^2}}}.
\end{aligned}
\tag{4.24}$$

The final result for  $a_{\text{sl}}^d$  is obtained with a weighted average of the results obtained for each decay mode. The systematic uncertainty on  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  are estimated for each decay mode (see Chapter 9). The weighted average is taken as systematic uncertainty on the final results. The final result for  $A_{\text{P}}$  is obtained with an analogous procedure, exception made for the weighted average of the results obtained for the different center-of-mass energies (that coincide with the different years). Two different results for  $A_{\text{P}}$  at 7 TeV and 8 TeV are reported. These values measure the the  $B^0 - \bar{B}^0$  production asymmetry in the selected kinematic region, and are not corrected for the  $p_{\text{T}}$ -dependent and  $\eta$ -dependent reconstruction efficiencies.

## 4.2 Summary

This analysis uses  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  semileptonic decays samples collected during Run-I by the LHCb experiment and corresponding to an integrated luminosity of  $3 \text{ fb}^{-1}$ . Cabibbo-favored  $D$  decays are selected to eliminate sources of  $CP$  violation different from  $CP$  violation in  $B^0 - \bar{B}^0$  mixing. The final state reconstructed for both decay channel is  $K^+ \pi^- \pi^- \mu^+$ .

The raw charge asymmetry  $A_{\text{meas}}$  is measured. Several asymmetries contribute to the raw charge asymmetry of the decays in Eq. 4.20: the  $B^0 - \bar{B}^0$  meson production asymmetry  $A_{\text{P}}$ , the charge asymmetry in the detection of the particles in the final state  $A_{\text{D}}$  and the  $CP$  asymmetry  $a_{\text{sl}}^d$ .  $A_{\text{P}}$  and  $a_{\text{sl}}^d$  are extracted from a time-dependent multidimensional maximum likelihood fit, while  $A_{\text{D}}$  is measured using dedicated data samples.

The maximum likelihood fit used to determine  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  includes a decay time description able to correct the decay time of the partially reconstructed decays on average and model the resulting poor momentum resolution. A component describing  $B^+$  decays, the main physics background present in the data samples is also included.

The full analysis is performed independently for the different magnet polarities and different center-of-mass energies. This is done primarily in order to properly control the detection asymmetries. In addition the  $B - \bar{B}$  meson production asymmetry is expected to behave differently for different center-of-mass energies, therefore it is necessary to allow

the parameter to take different values for the two datasets. The independent analysis of the different subsamples constitutes a powerful consistency check for the measurement.



# Chapter 5

## Samples and Selection

This chapter introduces the data and simulation samples used for the  $a_{\text{sl}}^d$  measurement. Sec. 5.1 describes the signal data samples, focusing on the selection procedure. The Monte Carlo simulation samples used are subsequently introduced in Sec. 5.2. Simulated events are very important in this analysis, as input and validation of the technique used to model the partially reconstructed decays. Finally, in Sec. 5.4, the control samples used to evaluate the detection asymmetries and the relative selection criteria are introduced.

### 5.1 Signal data samples

The  $a_{\text{sl}}^d$  measurement presented here uses  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays collected in 2011 and 2012 by the LHCb experiment, at centre-of-mass energies of 7 TeV and 8 TeV, respectively. These datasets correspond to an integrated luminosity of  $3 \text{ fb}^{-1}$ , the full luminosity acquired in Run-I. The selection of the signal candidates proceeds in several steps.

The crucial strategy behind the selection of the signal events for this measurement is to restrict the sample to a kinematic region where the detection asymmetries can be measured in a reliable and precise manner. The selection strategy at trigger level and the selection requirements here defined as *calibration cuts* aim to select the kinematical region of the signal decays to be as similar as possible to the control samples: Cabibbo-favored  $D^- \rightarrow K^+ \pi^- \pi^-$  and  $D^+ \rightarrow \bar{K}^0 \pi^+$  decays and  $J\psi \rightarrow \mu^+ \mu^-$  decays. The largest statistical reduction of the signal data samples is precisely due to these calibration cuts. A second important feature of partially reconstructed samples is that a considerable amount of physics background can contaminate the signal samples. In Sec. 5.1.5, a method to reduce to a negligible level the  $D$  decays originating directly from the primary interaction is discussed. In Sec. 5.3.2, a data-driven method to estimate the pollution of  $B^+$  decays the data sample is presented. Furthermore, the fraction of several types of physics background is estimated using simulated events, and discussed in the next section.

### 5.1.1 Signal decay topology

Fig. 5.1 shows the the topology of the signal decays used for this measurement. Two different samples are considered:  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  decays with the  $D$  meson decaying in the Cabibbo-favored mode  $D^- \rightarrow K^+ \pi^- \pi^-$ , and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  with the  $D^{*-}$  meson decaying to  $(\bar{D}^0 \pi^-)$  and  $\bar{D}^0 \rightarrow K^+ \pi^-$ .

As already mentioned, the lifetime of  $B$  mesons is large enough to allow them to fly on average about 1 cm before decaying. Also  $D^\pm$  and  $D^0$  mesons have a significant lifetime that helps to identify their decays. The second important feature of the semileptonic decays considered for this measurement is that the reconstructed  $B$  momentum doesn't point back to the Primary Vertex (PV).

The main handles to select these decays are large impact parameters (IP) of the reconstructed tracks (See. Fig. 5.1), displaced vertices with good quality and a long track associated with a muon candidate in the muon stations. The hadrons' species is identified by the LHCb PID system. Selection criteria based on the PID variables help to distinguish kaons from pions. Lastly the final state particles exhibit generally a harder transverse momentum spectrum with respect to the light quark background from the primary interaction. This justifies the minimum  $p_T$  requirements applied to the final state particles.

### 5.1.2 Trigger selection

The trigger system decides whether an event is interesting for physics analyses. For an event classified as interesting, all the detector hits and informations are saved when writing the event to storage. According the specific signature, an event can be acquired because of the positive trigger decision of a trigger line or another. "Trigger line" refers to a sequence of reconstruction and selection algorithms to trigger an event. Selected signal candidates can be classified as:

- *Triggered On Signal (TOS)*, when the signal candidate or its daughters cause the event to be triggered.
- *Triggered Independently of Signal (TIS)*, when a positive trigger decision is reached independently of the signal candidate or its daughters, i.e. some other track in the event caused the positive trigger decision.

An event can be triggered by a TOS and a TIS line at the same time, which allows a measurement of the trigger efficiencies. The TIS/TOS distinction can be made for trigger lines at the L0 level, or at HLT level.

The signal candidates analyzed are required to fulfill the following trigger conditions.

#### Level-0

A semileptonic signal event for this analysis must be TOS for the L0MUON trigger line. This choice is justified by the fact that L0MUON is the only line at Level-0 for which a

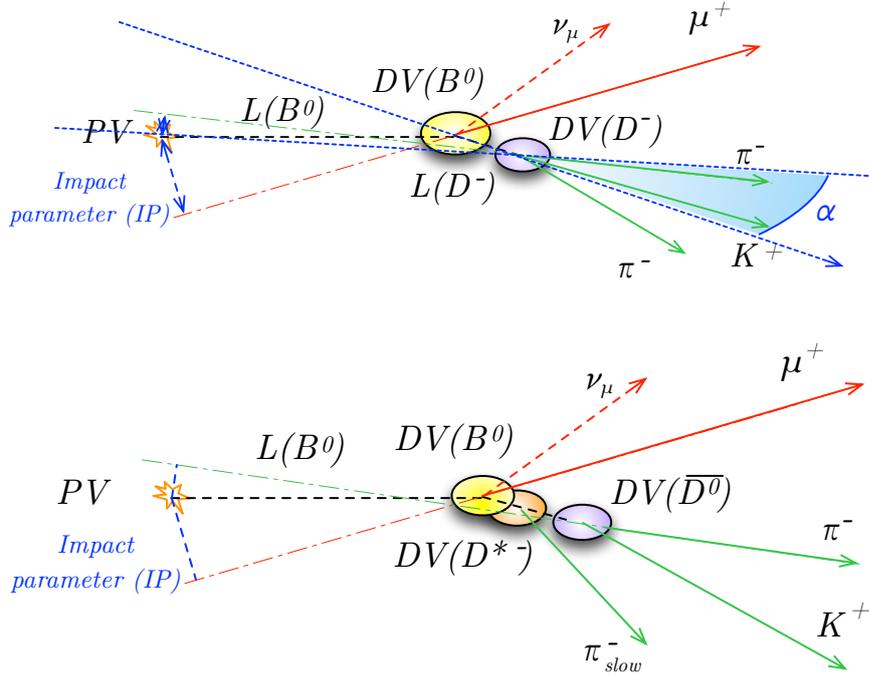


Figure 5.1: Signal decays topology. The flight distance of a mesons  $M$  is indicated with  $L(M)$  and the decay vertex with  $DV(M)$ . The decay vertices of different mesons are indicated with different colors. The impact parameter (IP) for the muon track and the pointing angle  $\alpha$  for the  $D^-$  meson in sketched.

precise calibration of the charge asymmetry is available.

## HLT1

The first software level of the LHCb trigger takes advantage of the higher transverse momentum of the final state particles with respect to the light quark background and their displacement for the primary interaction. The first step of HLT1 algorithms is to reconstruct a VELO track, measure the IP and subsequently extend the VELO track to the detector hits behind the magnet. The latter step can be done with two different approaches: HLT1TRACKALL0 and HLT1TRACKMUON. In the second case the L0MUON positive decision is required and the VELO track is matched with hits in the muon system. The track is then upgraded to a long track. The track is required to have an  $IP > 0.1$  mm and  $p_T > 1$  GeV. In  $a_{\text{sl}}^d$  signal decays at HLT1 level the muon track must be TOS for HLT1TRACKMUON.

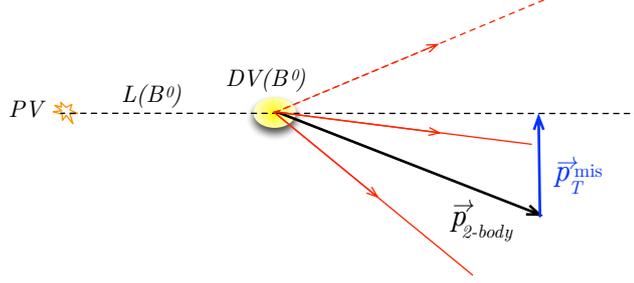


Figure 5.2: The definition of  $p_T^{\text{mis}}$ . The back dashed line indicates the flight direction of the  $B^0$  meson in this case. The solid red lines represent the track reconstructed in the final state. The solid black line is the momentum of the 2-body combination. The dashed red line belongs to a not reconstructed particle.

## HLT2

HLT2 trigger lines, according to the strategy that they adopt, can be distinguished in *inclusive* and *exclusive* lines. The former lines are suitable for a variety of partially reconstructed  $B$  decays, while the second type of lines is optimal for fully reconstructed decays. The so-called *topological* trigger lines [91] take advantage of combinations of 2, 3, or 4 tracks to partially reconstruct  $b$ -hadrons. A set of  $B$  daughter tracks is identified thanks to selection cuts on the track quality and IP. Subsequently two tracks are combined in a 2-body object by requiring that their distance of closest approach (DOCA) is less than 0.2 mm. Up to two additional tracks can be added using the same criteria to form 3-body and 4-body objects. Since not all the products of a  $B$  decays originate from the same vertex, even if the DOCA requirement is not strict, the topological trigger is designed to make a decision on partially reconstructed decays. For partially reconstructed decays, it is convenient to define a missing transverse momentum  $p_T^{\text{mis}}$  relative to the flight direction of the  $n$ -body reconstructed object, as depicted in Fig. 5.2 in the case of two reconstructed particles in the final state.

Fig. 5.5 visualizes the missing transverse momentum for example in the case of the signal decays considered. The  $p_T^{\text{mis}}$  can be used to correct the invariant mass of the reconstructed  $n$ -body system as follows

$$m_{\text{corr}} = \sqrt{m^2 + |p_T^{\text{mis}}|^2 + |p_T^{\text{mis}}|} \quad (5.1)$$

and it can be used to suppress background of  $D$  decays produced in the  $pp$  interaction by requiring a minimum value for  $m_{\text{corr}}$ . The  $m_{\text{corr}}$  variable indeed takes values closer to the real  $B$  mass when more tracks from the  $B$  decay are added. In Sec. 5.3.2 another example of usage of this corrected mass variable is reported. The final decision in the topological lines is reached by using a multivariate selection. For the  $a_{\text{sl}}^d$  selection at HLT2 level, the  $B$  candidate must be TOS for one of the three topological lines HLT2MUNBODYBBDT, ( $N=2, 3$  or  $4$ ).

In addition to the trigger line selection, some data acquired with special configurations, is excluded from the selected sample:

- *Exclusion of the data acquired between the 20th and 24th of April 2012.*<sup>1</sup> Data collected during 2011 is affected by a detection asymmetry which originates from the Level-0 muon trigger. This is the effect of having assumed a position of the muon pads that was shifted with respect to the real position of the detector. The effect is consistently reduced in 2012 data, since a calibration was run on the first data acquired. The runs used for this calibration are excluded for the  $a_{\text{sl}}^d$  measurement, they correspond to approximately  $50 \text{ pb}^{-1}$ . For both the data acquired in 2011 and MC this effect is corrected for by emulating the L0 with the correct muon pad positions. This results in a look-up-table (LUT) which converts the L0 muon coordinates into a new  $p_{\text{T}}$  value (see Sec. 6.3 for more details).
- *Exclusion of few TCKs.*<sup>2</sup> A particular choice of trigger settings and thresholds is given a unique identifier, known as Trigger Configuration Key (TCK). Some TCKs are excluded for the  $a_{\text{sl}}^d$  measurement because of having thresholds in the L0 trigger different with respect to the other TCKs. The excluded events correspond to less than 0.1% of the signal data sample.

### 5.1.3 Decay chain reconstruction: Decay Tree Fitter

In high energy experiments there are two strategies to proceed when reconstructing a decay chain.

The first is known as *bottom-up approach*, and consists of consecutive steps. The particles reconstructed in the final state are combined to form the intermediate particles by constraining them to originate from a common vertex. The same procedure is performed on the intermediate particles to reconstruct the upstream decays. This method doesn't exploit the whole information available, since the constraints upstream of a decay vertex don't affect the determination of the parameters of the decay vertex itself. An approach that doesn't have this disadvantage is the so-called *Decay Tree Fitter (DTF)* [92], that extracts all parameters in a decay chain simultaneously. In this case, the positions of the vertices and the momenta of all particles, are degrees of freedom of the decay tree. Momentum conservation at each vertex and the relation between the decay vertex of a particle and the production vertex of its daughters establish the internal constraints that eliminate the redundant degrees of freedom. The external constraints are given by the reconstructed final state particles, namely their momentum vector. Each intermediate particle is modeled by a four momentum vector and a decay vertex position vector. In the case where the decay length of the intermediate particle is comparable to (or larger than) the vertex detector resolution, also a parameter indicating the decay time of the particle  $\theta = L/|\vec{p}|$  is assigned. Otherwise the particle is identified as "resonance" and the decay vertex position

---

<sup>1</sup>Run numbers between 110000 and 114000.

<sup>2</sup>0x4a0033, 0x5d0033, 0x740036 in 2011 and 0x95003d, 0x9a0042 in 2012

coincides with the decay vertex of the mother particles. Additional constraints can be used, for example constraints can be placed on the mass of the intermediate particles. The decay parameters and the corresponding covariance matrix is extracted from the constraints using a Kalman filter [87]. The DTF is used to reconstruct the decay chain for the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  signal decays. In this case the DTF constrains simultaneously the  $K^\mp - \pi^\pm$  to form a  $D^0$  candidate and the  $D^0 - \pi - \mu$  combination to form a  $B$  candidate. The  $\chi^2/\text{ndof}$  from this fit is required to be below 10, otherwise the candidate is rejected (see Table.5.1).

### 5.1.4 Central offline selection (“Stripping”)

Before the final offline selection performed by each analysis group, a central offline selection is performed, in order to reduce the datasets to a manageable size. This central offline selection is named “stripping” within the LHCb collaboration. According to the signal of interest, different central selection are possible. Table. 5.1 summarizes the selection criteria at stripping level used for the selection of the signal  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays.

### 5.1.5 Offline selection

Further offline selection criteria are applied. A distinction can be made between the selection cuts designed to reduce the background level and the selection cuts that aim to equalize the kinematics of the signal samples with the kinematics of the control samples.

#### $J/\psi$ and $\Lambda_c^+$ vetos

Two types of physics background are easily identified and rejected by applying vetoes:

- $B \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)X$  decays, where one of the decay muons from the  $J/\psi$  decay mimics the muon of the semileptonic decay, and the other muon, combined with the other tracks, mimics a  $D^{(*)\pm}$  meson. Events having a reconstructed  $\mu\pi$  invariant mass between 3070 MeV/ $c^2$  and 3150 MeV/ $c^2$  and the pion fulfilling the muon particle identification criteria are rejected.
- $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu^- X$  decays, with  $\Lambda_c^+ \rightarrow pK^- \pi^+$ , in which the proton is misidentified as a  $\pi^+$ . Events having a  $\mu\pi K$  invariant mass between 2260 MeV/ $c^2$  and 2310 MeV/ $c^2$  and the pion fulfilling proton particle identification criteria <sup>3</sup> are rejected. This veto is applied only to the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  decay mode. It is not necessary for the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  data sample.

---

<sup>3</sup>DLL<sub>pπ</sub> > 10

## Promptly produced $D$ mesons

$D$  meson decays where the  $D$  meson is not a daughter of a  $B$  hadron, but is produced directly from the combination of a  $c$  or  $\bar{c}$  quark with a  $\bar{u}$  or  $u$  quark following the  $pp$  collision, are named *prompt* charm decays. These decays, with a muon associated, are a source of background. A variable that helps discriminating between the prompt  $D$  decays

<b>Kaon</b>		<b>Pions from <math>D^0</math> or <math>D^+</math></b>		<b>Muon</b>	
$p$	$> 2.0 \text{ GeV}/c$	$p$	$> 2.0 \text{ GeV}/c$	$p$	$> 3.0 \text{ GeV}/c$
$p_T$	$> 0.3 \text{ GeV}/c$	$p_T$	$> 0.3 \text{ GeV}/c$	$p_T$	$> 0.8 \text{ (1.2) GeV}/c$
$\mathcal{P}(\text{ghost})$	$< 0.5$	$\mathcal{P}(\text{ghost})$	$< 0.5$	$\mathcal{P}(\text{ghost})$	$< 0.5$
$\chi_{\text{track}}^2/\text{ndf}$	$< 4$	$\chi_{\text{track}}^2/\text{ndf}$	$< 4$	$\chi_{\text{track}}^2/\text{ndf}$	$< 4$
$\chi_{\text{IP}}^2$	$> 9$	$\chi_{\text{IP}}^2$	$> 9$	$\chi_{\text{IP}}^2$	$> 9$
$\text{DLL}_{K\pi}$	$> 4$	$\text{DLL}_{K\pi}$	$< 10 \text{ (4)}$	$\text{DLL}_{\mu\pi}$	$> 0$
<b><math>D^0</math> or <math>D^\pm</math></b>					
$ M(D) - M_{PDG}(D) $		$< 80 \text{ MeV}/c^2$			
$\sum p_T \text{ } D \text{ daughters}$		$> 1800 \text{ (1400) MeV}/c$			
$\chi_{\text{DOCA}}^2 \text{ } D \text{ daughters}$		$< 20$			
$\chi_{\text{vertex}}^2/\text{ndf}$		$< 6$			
$\chi_{\text{distance}}^2 \text{ } D \text{ vertex} - \text{PV}$		$> 100$			
$\cos \alpha$		$> 0.99$			
<b><math>D^{*\pm}</math> (only for <math>B^0 \rightarrow D^{*-}\mu^+\nu_\mu X</math>)</b>					
$\chi_{\text{track}}^2/\text{ndf}$ slow pion		$< 5$			
$p_T$ slow pion		$> 180 \text{ MeV}/c^2$			
$M_{\text{reco}}(D^{*\pm}) - M_{\text{reco}}(D^0) - M_{PDG}(\pi^\pm)$		$\in [0.0, 170] \text{ MeV}/c^2$			
$\chi_{\text{vertex}}^2/\text{ndf}$		$< 8$			
<b><math>B</math></b>					
$M(B)$		$\in [2.5, 6.0] \text{ GeV}/c^2$			
$\chi_{\text{vertex}}^2/\text{ndf}$		$< 6$			
$\cos \alpha$		$> 0.999$			
$z(D) - z(B)$		$> 0 \text{ mm (no cut)}$			
$\chi_{\text{vertex}}^2/\text{ndf}$ (DTF)		no cut ( $< 10$ )			

Table 5.1: Stripping selection for both decay channels. When different, the requirement for the  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$  mode is given between brackets. For the definition of Impact Parameter (IP) and pointing angle  $\alpha$  refer to Fig. 5.1 (where the pointing angle for the  $D^\pm$  meson and the IP of the muon are illustrated.). The ghost probability  $\mathcal{P}(\text{ghost})$  is a multivariate classifier, using kinematic variables and track reconstruction parameters as inputs, to identify reconstructed tracks which do not correspond to a real particle [93].  $\text{DLL}_{x\pi}$  is a particle identification variable (see Sec. 3.3.2). With “soft pion” for the  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$  decays, the bachelor of the  $D^*$  is meant, while the pion with lower transverse momentum is meant for  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  decays.

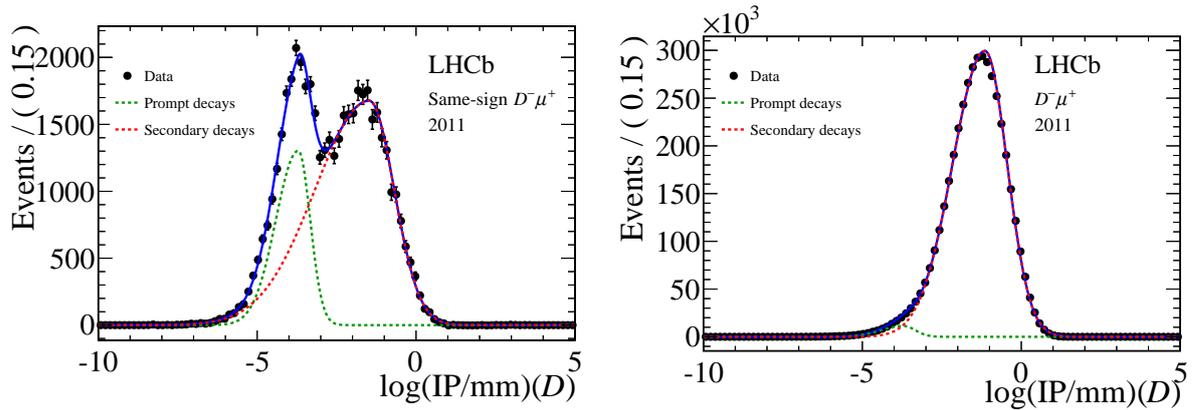


Figure 5.3: Logarithm of the Impact Parameter(IP) of the  $D^\pm$  meson distributions for  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  same-sign decays (Left) and right-sign decays (Right).

and  $D$  decays from  $B$  decays is the Impact Parameter (IP) of the  $D^\pm$ , or even better, the logarithm of the impact parameter of the  $D^\pm$  meson. Fig. 5.3 (Left) shows the  $\log(\text{IP})(D)$  distribution for the so-called *same-sign* events. These are events with exactly the same topology as the signal, but with the “wrong” charge combinations, i.e. events with a  $D^\pm$  and a  $\mu^\pm$  with the same charge. This sample is enriched with prompt  $D$  decays combined with a random muon. In these events, the  $D$  meson points back to the primary vertex, giving on average lower IP values, compared to the non-prompt charm decays. The data points of both prompt and secondary decays can be described with a bifurcated Gaussian. The number of prompt same-sign decays from this fit can be compared with the total number of right-sign signal decays, to obtain an estimate of the fraction of prompt decays contaminating the data sample. The fraction of prompt  $D$  decays present in the signal data sample is about 0.7% from this estimate.

The same-sign sample accounts for only random muons associated to prompt  $D$  decays of the same charge, the random muon associated could also have the opposite charge with respect to the  $D$  decay. In this case the prompt  $D$  decay would be selected in the signal sample. Fig. 5.3 (Right) shows the same  $\log(\text{IP})(D)$  for the signal sample. Fitting the data points with the same models used for the same-sign samples, the fraction of prompt  $D$  decays results to be about 2%. The fraction of prompt  $D$  decays is expected to be smaller for the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  sample. Fig.5.3 shows also that by requiring among the selection criterium  $\log(\text{IP})(D) > -3.0$  would reduce the prompt  $D$  decays contribution essentially at per-mille level. With this requirement, the amount of prompt  $D$  decays in the data sample becomes negligible. In this analysis the signal decay candidates are required to fulfill the selection requirement  $\log(\text{IP})(D) > -3.0$ . The background from fake  $D$  mesons is kept under control by including the  $K^\pm \pi^\mp \pi^\mp$  invariant mass distribution among the global fit dimensions. The amount of this background is further reduced by requiring a  $D$  decay time larger than 0.1 ps.

## Calibration cuts

The calibration requirements are listed in Table 5.2. These include restrictions to the kinematics of the final state particles and PID selection criteria. The muon detection asymmetry is only measured down to  $p_T = 1.2 \text{ GeV}/c$  and  $p = 6 \text{ GeV}/c$ . Therefore these limits are applied as selection cuts to the signal candidates. Having similar kinematics between the muon and the higher-momentum pion (the pion from the  $D^0$  in the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decay mode and the higher- $p_T$  pion in the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  decay mode) is an advantage when measuring the detection asymmetry of the two particles, as is explained in Chapter 6. For this reason the same kinematic cuts applied to the muon are also applied to the higher-momentum pion. Similarly, a  $300 \text{ MeV}/c$  cut on  $p_T$  has to be applied to both the kaon and lower-momentum pion, to match the minimum requirements in the prompt  $D^\pm$  samples used to calibrate their detection asymmetries. In addition, some PID selection criteria are used, to obtain the same PID selection as in the control samples, and the correction look-up-table (see Sec. 5.1.2) is applied to reduce the muon detection asymmetry due to the L0 trigger. The selection efficiency of the calibration requirements is given in Table 5.3. Table 5.3 shows the reduction of the data sample due to the application of each group of selection criteria. Calibration cuts and the kinematical weighting procedure used and explained in Sec. 6.5.1 determine most of the statistical degradation of the samples. The novel strategies that are currently under development to evaluate the detection asymmetries, will allow to avoid the kinematical weighting procedure and likely to relax the calibration cuts.

### 5.1.6 Event yields

The signal yields are listed in Table 5.4, split by year and magnet polarity. Absolute signal yields and effective yields are reported. The effective yields account for the statistical degradation due to the  $\mu - \pi$  kinematical weighting procedure (Sec. 6.5.1). In the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  channel, this weighting reduces the effective statistics by about 30%. These yields still include backgrounds from semileptonic  $B^+$  decays. For the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  channel, an additional offline cut is used: the difference between the invariant mass of the  $D^{*\pm}$  and the  $D^0$ , defined as  $\delta M = M(D^{*\pm}) - M(D^0)$ , is limited to be between  $144 \text{ MeV}/c^2$  and  $147 \text{ MeV}/c^2$ . This cut is included in the full analysis (reported as “Offline cut” in Tab.5.2), and it is included when calculating the yields in Tab.5.4.

Fig. 5.4 shows an example of  $D^0$  and  $D^\pm$  invariant mass distributions (magnet-up, 2012) having all selection requirements described in this Section applied. Fit results and projections for all subsets of the data are presented in Chapter 10 and in Appendix B, respectively. For both modes, the signal yields are obtained from a fit to these mass distributions. The signal is modeled with a sum of a Crystal Ball [94] and a Gaussian functions having the same mean, and the background with an exponential function.

Table 5.2: Offline selection on top of the stripping and the trigger requirements.

	<b>Variable</b>	$B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$	$B^0 \rightarrow D^- \mu^+ \nu_\mu X$
Offline cuts	$M(D^{*\pm} \mu^\mp)$	[3.0, 5.2] GeV/ $c^2$	[3.0, 5.2] GeV/ $c^2$
	$M(D^{*\pm}) - M(D^0)$	[144, 147] MeV/ $c^2$	-
	$\tau(D^{0,\pm})$	> 0.1 ps	> 0.1 ps
	$\log(\text{IP})D^{0,\pm}$	> -3.0	> -3.0
	$J/\psi$ veto	See Sec. 5.1.5	See Sec. 5.1.5
	$A_c^+$ veto	-	See Sec. 5.1.5
Calibration cuts	$p$ all tracks	> 3 GeV/ $c$	> 3 GeV/ $c$
	$p$ ( $\mu^\pm$ and higher $p_T$ $\pi^\pm$ )	> 6 GeV/ $c$	> 6 GeV/ $c$
	$p_T(\pi_s$ or lower $p_T$ $\pi^\pm$ )	> 300 MeV/ $c$	> 300 MeV/ $c$
	$p_T(K^\pm)$	[300, 7000] MeV/ $c$	[300, 7000] MeV/ $c$
	$p_T(\mu^\pm)$	1.2 GeV/ $c$	> 1.2 GeV/ $c$
	$p_T(\pi^\pm$ or higher $p_T$ $\pi^\pm$ )	1.2 GeV/ $c$	> 1.2 GeV/ $c$
	$DLL_{K\pi}(K^\pm)$	> 7	> 7
	$DLL_{K\pi}(\pi_s$ or lower $p_T$ $\pi^\pm$ )	< 3	< 3
	$DLL_{K\pi}$ (higher $p_T$ $\pi^\pm$ )	< 10	< 4
	$DLL_{\mu\pi}(\mu^\pm)$	> 0	> 0
ISMUON ( $\mu^\pm$ )	true	true	
	TCK selection	See Sec. 5.1.2	See Sec. 5.1.2
	Excluded runs	[110000,114000]	[110000,114000]

Table 5.3: The fraction of candidates passing each group of selection requirements. Each row corresponds to a group of selection requirements, for which it is shown show the ratio of the number of events passing all selection requirements to those passing all except this group. I.e. it is the efficiency of each group of selection requirements w.r.t. all others. The stripping selection is always included in the denominator. The row labelled "μπ reweighting" indicates the statistical degradation of the sample due to the kinematic weights applied on the data sample. The last row shows the total efficiency with respect to the stripping, with kinematics weights applied. The efficiencies are calculated using  $D^+/D^0$  sideband subtracted data (see Note 6).

Selection	$B^0 \rightarrow D^- \mu^+ \nu_\mu X$		$B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$	
	2011	2012	2011	2012
Trigger selection	72.4%	71.8%	74.0%	73.5%
TCK and run selection	99.3%	96.3%	99.3%	96.4%
Calibration cuts	75.9%	75.8%	40.8%	40.1%
Offline cuts	81.6%	81.5%	69.6%	69.1%
PID cuts + LUT for 2011	82.3%	88.7%	73.9%	82.7%
B decay time (fit range)	76.9%	77.3%	75.0%	75.6%
<b>Total efficiency (before μπ reweighting)</b>	18.7%	20.9%	7.1%	8.5%
μπ reweighting	64.5%	65.0%	78.8%	79.9%
<b>Total efficiency</b>	12.1%	13.5%	5.6%	6.7%

Table 5.4: Signal yields for the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  obtained from fits to the  $D^0$  and  $D^\pm$  mass distributions after all selection cuts. The yields are obtained given the number of candidates after the full selection and the fraction of combinatorial background determined by means of  $a_{\text{sl}}^d$  nominal fit. The first two rows show the absolute signal yields. The second two rows show the effective signal yields after applying the  $\mu - \pi$  re-weighting that is introduced in Sect. 6.5.1

Channel	2011 Up	2011 Down	2012 Up	2012 Down	Total
$B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$ (no μπ weights)	53,089	71,462	179,411	168,005	471,967
$B^0 \rightarrow D^- \mu^+ \nu_\mu X$ (no μπ weights)	352,705	477,661	1,187,579	1,103,893	3,121,838
$B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$	40,020	54,208	140,515	131,647	366,389
$B^0 \rightarrow D^- \mu^+ \nu_\mu X$	226,153	306,205	768,088	714,399	2,014,845

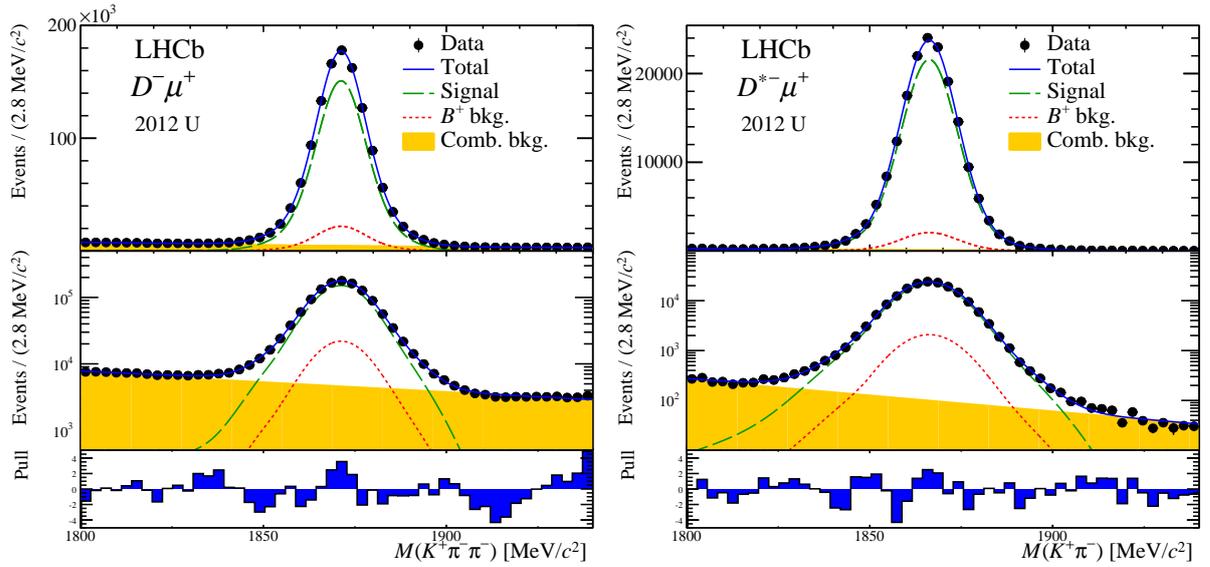


Figure 5.4: The invariant mass distributions of the  $K^\mp \pi^\pm \pi^\pm$  combinations in the (*left*)  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  magnet-up 2012 data sample and invariant  $K^\mp \pi^\pm$  mass distribution for the (*right*)  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  magnet-up 2012 data sample. The first two top rows show the invariant mass distributions using a logarithmic and a not-logarithmic y-axis, respectively. The last row shows the so-called *Pull* values, i.e. the difference between the number of events observed in data for each bin of the histogram and the number of events predicted by the model in the same bin, divided by the uncertainty on the number of events of the bin. These three plots are used for most of the fit projections in this thesis, in order to show the agreement between the data distribution and the fit model.

## 5.2 Monte Carlo simulation

Estimation of the fraction of events of a certain type that escape the detector acceptance, or detailed studies of the expected response of the detector to the passage of particles are common needs in several analyses. Usually, complex detector geometries and the numerous effects that need to be accounted for in predicting their response make the analytical derivation of the relevant distributions impractical or impossible. An alternative method is to use numerical simulation to perform these studies. The algorithms used involve random sampling to simulate processes, and are collectively called Monte Carlo (MC) simulation. The simulation involves several steps in order to describe the  $pp$  collisions, decay processes, detector response and data processing and selection [95].

The first step is the modeling of the  $pp$  collision environment, the outgoing quark and gluon collision products, followed by the simulation of the fragmentation and hadronization processes which produces hadrons and associated jets. This is done by the PYTHIA package [69]. A second step is the description of the time evolution and decay of the generated hadrons. This is delegated to the EVTGEN package <sup>4</sup> [96], specialized in heavy flavor processes. The subsequent step incorporates the interaction of the decay products with the detector material. This is achieved using the GEANT4 toolkit [97], [98]. The geometry and detector conditions are used by GEANT4 to correctly describe LHCb geometry and running conditions (e. g. the magnet polarity). The detector response is then determined from the interaction of the particles with the material and used as input for trigger decision, track reconstruction, central offline analysis and selection, which are performed as on data.

Simulated event samples are extremely useful for the  $a_{\text{sl}}^d$  measurement. The decay time model used (see Chapter. 7.1.2) relies on Monte Carlo input. In addition, several types of physics background can pollute the data samples. The abundance of these components can be estimated by using MC samples selected with the same criteria imposed for the data. MC samples generated with 2011 and 2012 data acquisition conditions are used. An equal number of events generated with two PYTHIA versions (6 and 8), and both magnet-polarities are simulated. Table 5.5 shows the generation parameters for the  $B^0$  mixing frequency and  $B^0$  and  $B^+$  decay widths.

Different simulated samples, listed in Tab.5.6, are used:

---

<sup>4</sup>EVTGEN was originally designed by the BaBar collaboration, it is therefore specialized in describing  $B$  mesons time evolution and decays. LHCb needs to model the incoherent production of  $B$  mesons properly, for this reason a customized version of EVTGEN is used.

Parameter	Generated value
$\Delta m_d$	$0.507 \text{ ps}^{-1}$
$\Gamma(B^0)$	$0.656 \text{ ps}^{-1}$
$\Gamma(B^+)$	$0.609 \text{ ps}^{-1}$

Table 5.5: Parameters used in the generation of the used simulation samples.

Table 5.6: MC samples with 2011 and 2012 conditions. The data size is given as the total number of generated events (including both magnet polarities and Pythia versions).

Sample	Data size	
	2011	2012
Signal $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$ cocktail	20M	40M
Signal $B^0 \rightarrow D^-\mu^+\nu_\mu X$ cocktail	10M	20M
Bkg $B^+ \rightarrow D^0\mu^+\nu_\mu X$ cocktail	5M	17.5M
Bkg $B^+ \rightarrow D^-\mu^+\nu_\mu X$ cocktail	2.5M	7.5M
Bkg inclusive $D^- \rightarrow K^+\pi^-\pi^-$	5M	10M
Bkg inclusive $D^{*-} \rightarrow \bar{D}^0\pi^-$ with $\bar{D}^0 \rightarrow K^+\pi^-$	5M	10M

- *Signal samples.*  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  and  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$  samples. These simulated events are used to model the signal decay time in the time dependent fit. These samples are not only including the most common signal decay chain, but also decays through higher charm resonances. The knowledge of the branching fractions of the different modes allows to build these so-called *cocktail* samples. By imposing the same selection procedure of the data samples, simulated events are expected to correctly reproduce the data samples.
- *$B^+$  decays samples.*  $B^+ \rightarrow D^-\mu^+\nu_\mu X^+$  and  $B^+ \rightarrow D^{*-}(\rightarrow D^0\pi^-)\mu^+\nu_\mu X^+$  are the main physics background component in the  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  and  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$  data samples, respectively. The same procedure as for the signal decay samples allows to estimate the level of contamination of the data samples with  $B^+$  decays.
- *inclusive  $D$  samples.* These inclusive samples are used to infer the fraction of background decays present in the selected signal data samples.

In the next sections these MC samples are presented, the attention is mostly devoted to their composition and selection efficiencies.

### 5.2.1 Signal Samples

In the two signal samples  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  and  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$  (charge-conjugation implied), the  $B^0$  mesons are forced to decay to a  $D^-$  or  $D^{*-}$  respectively. The so-called cocktail includes feed-down through higher  $D$  resonances, based on the known branching fractions [99]. When measured values are not available, the branching fraction values are inferred by using isospin arguments [100], [101]. Subsequently, the  $D^-$  mesons are required to decay to  $K^+\pi^-\pi^-$  and the  $D^{*-}$  mesons to  $\bar{D}^0(\rightarrow K^+\pi^-)\pi^-$ . Tables 5.8 and 5.7 list the branching fractions [99] that are included in the generation of these MC cocktails. For both signal samples, the cocktail includes the main decay chain, several decay chains where the  $B^0$  decays to some higher  $D$  resonance, a contribution from decays through

Table 5.7: The main contributions to the total branching fractions used in the generation of the signal simulation sample in the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  mode. In red the main  $B^0 \rightarrow D^- \mu^+ \nu_\mu$  chain, blue feed-down through higher  $D$  resonances contribution, green decays through non resonant  $D^\pm \pi^0$ ,  $D^\pm \pi^0 \pi^0 \dots$  states, yellow contribution of tau decays. The fractions,  $\Gamma_i$ , are with respect to the total decay width of the  $B^0$ ,  $\Gamma$ . The values are taken from the PDG [99], and the branching fraction of the subsequent decay of the  $D$  mesons produced in the  $B$  decay, indicated between brackets, are taken into account.

Process	Branching fraction ( $\Gamma_i/\Gamma$ )
$B^0 \rightarrow D^- \mu^+ \nu_\mu$	$(2.17 \pm 0.12)\%$
$B^0 \rightarrow D^{*-} (\rightarrow D^- \pi^0 / \gamma) \mu^+ \nu_\mu$	$(1.62 \pm 0.04)\%$
$B^0 \rightarrow D_0^{*-} \mu^+ \nu_\mu \rightarrow D^- \mu^+ \nu_\mu X$	$(0.14 \pm 0.05)\%$
$B^0 \rightarrow D_1^- \mu^+ \nu_\mu \rightarrow D^- \mu^+ \nu_\mu X$	$(0.06 \pm 0.02)\%$
$B^0 \rightarrow D_1^- \mu^+ \nu_\mu \rightarrow D^- \mu^+ \nu_\mu X$	$(0.18 \pm 0.03)\%$
$B^0 \rightarrow D_2^{*-} \mu^+ \nu_\mu \rightarrow D^- \mu^+ \nu_\mu X$	$(0.17 \pm 0.07)\%$
$B^0 \rightarrow D^{*-} (\rightarrow D^- \pi^0 / \gamma) \pi^+ \pi^- \mu^+ \nu_\mu$	$(0.09 \pm 0.09)\%$
$B^0 \rightarrow D^- \pi^+ \pi^- \mu^+ \nu_\mu$	$(0.12 \pm 0.12)\%$
$B^0 \rightarrow D^- \tau^+ \nu_\tau \rightarrow D^- \mu^+ \nu_\mu X$	$(0.19 \pm 0.04)\%$
$B^0 \rightarrow D^{*-} \tau^+ \nu_\tau \rightarrow D^- \mu^+ \nu_\mu X$	$(0.08 \pm 0.03)\%$
<b>overall</b> $B^0 \rightarrow D^- \mu^+ X$	$(4.95 \pm 0.23)\%$

non-resonant states  $D^\pm \pi^0$ ,  $D^\pm \pi^0 \pi^0$ ,  $D^\pm \pi^+ \pi^-$  or  $D^{*\pm} \pi^0$ ,  $D^{*\pm} \pi^0 \pi^0$ ,  $D^{*\pm} \pi^+ \pi^-$ , and decays where the muon originates from a  $\tau$  decay.

The selection efficiency for these signal MC samples is given in Table 5.9. The total efficiency is compatible with the observed yields and the measured  $B$  cross sections [102].

## 5.2.2 $B^+$ decays samples

Two  $B^+$  background samples are considered. In one sample the  $B^+$  is required to decay into  $D^- (\rightarrow K^+ \pi^- \pi^-) \mu^+ X$ , while the other sample is forced to decay into  $\bar{D}^0 (\rightarrow K \pi)$ . The MC cocktail samples used are built with the same method as the simulated signal decays. The branching fractions are given in Tables 5.11 and 5.10. In Table 5.11, only the branching fractions for decays with a  $D^{*\pm}$  resonance are given, since these are the only decays contributing to the  $B^+$  background in the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decay mode. The MC cocktail used to investigate  $B^+ \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays is a more inclusive  $B^+ \rightarrow D^0 \mu^+ \nu_\mu X$  sample.

## 5.2.3 Inclusive $D$ samples

Two inclusive  $D$  background samples, containing  $D^{*+}$  decays into  $D^0 \pi^+$ , and  $D^0 \rightarrow K^- \pi^+$ , are used. These samples allow to study backgrounds containing real  $D^\pm$  and  $D^{*-}$  decays, for example promptly produced  $D^\pm$  and  $D^{*-}$  with a fake or random muon. These inclusive  $D$

Table 5.8: The main contributions to the total branching fractions used in the generation of the signal simulation sample in the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  mode. In red the main  $B^0 \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu$  chain, blue feed-down through higher  $D$  resonances contribution, green decays through non resonant  $D^\pm \pi^0$ ,  $D^\pm \pi^0 \pi^0 \dots$  states, yellow contribution of tau decays. The fractions,  $\Gamma_i$  are with respect to the total decay width of the  $B^0$ ,  $\Gamma$ . The values are taken from the PDG [99], and the branching fraction of the subsequent decay of the  $D$  mesons produced in the  $B$  decay, indicated between brackets, are taken into account.

Process	Branching fraction ( $\Gamma_i/\Gamma$ )
$B^0 \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu$	$(3.39 \pm 0.07)\%$
$B^0 \rightarrow D_0^{*-} \mu^+ \nu_\mu \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu X$	$(0.04 \pm 0.01)\%$
$B^0 \rightarrow D_1^{*-} \mu^+ \nu_\mu \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu X$	$(0.04 \pm 0.02)\%$
$B^0 \rightarrow D_1^- \mu^+ \nu_\mu \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu X$	$(0.12 \pm 0.02)\%$
$B^0 \rightarrow D_2^{*-} \mu^+ \nu_\mu \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu X$	$(0.04 \pm 0.02)\%$
$B^0 \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \pi^0 \mu^+ \nu_\mu$	$(0.03 \pm 0.03)\%$
$B^0 \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \pi^0 \pi^0 \mu^+ \nu_\mu$	$(0.04 \pm 0.04)\%$
$B^0 \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \pi^+ \pi^- \mu^+ \nu_\mu$	$(0.17 \pm 0.17)\%$
$B^0 \rightarrow D^{*-} \tau^+ \nu_\tau \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu X$	$(0.18 \pm 0.06)\%$
<b>overall <math>B^0 \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu X</math></b>	<b><math>(4.07 \pm 0.20)\%</math></b>

events include about 15 million events. They include charm from  $B$  decays.<sup>5</sup>. Reconstructed

<sup>5</sup>The simulated background from secondary  $B$  decays does not contain semileptonic  $\Lambda_b^0$  decays to final states with a  $D$  meson (only to charmed baryons). The effect of such backgrounds is estimated in

Table 5.9: Efficiency for the simulated signal samples after applying each selection. Both magnet polarities and Pythia versions are added. Each row corresponds to a group of selection requirements, for which it is shown the ratio of the number of events passing all selection requirements to those passing all except this group.

Selection	$B^0 \rightarrow D^- \mu^+ \nu_\mu X$		$B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$	
	2011	2012	2011	2012
Stripping cuts	5.61%	4.97%	3.70%	2.52%
Trigger selection	49.1%	52.0%	60.9%	62.5%
Offline cuts	82.0%	82.3%	74.4%	74.6%
Calibration cuts	75.7%	76.1%	41.2%	41.6%
PID cuts	93.0%	92.1%	86.7%	84.8%
B decay time cut	77.4%	77.7%	76.0%	76.5%
Efficiency wrt stripping	13.8%	13.7%	8.7%	8.8%
$\mu\pi$ weights	62.6%	62.4%	92.6%	96.1%
Effective Efficiency wrt stripping	8.0%	8.5%	7.9%	8.5%
<b>Total efficiency</b>	<b>0.258%</b>	<b>0.239%</b>	<b>0.214%</b>	<b>0.164%</b>
<b>Selected events</b>	<b>69,041</b>	<b>130,333</b>	<b>78,593</b>	<b>149,421</b>
<b>Effective selected events</b>	<b>43,220</b>	<b>81,328</b>	<b>72,777</b>	<b>143,595</b>

Table 5.10: The main contributions to the total branching fractions used in the generation of the  $B^+$  background sample in the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  mode. The values are taken from the PDG [99], and the branching fraction of the subsequent decay of the  $D$  mesons produced in the  $B$  decay, indicated between brackets, are taken into account.

Process	Branching fraction
$B^+ \rightarrow D_0^{*0} \mu^+ \nu_\mu \rightarrow D^- \mu^+ \nu_\mu X$	$(0.26 \pm 0.09)\%$
$B^+ \rightarrow D_1^0 \mu^+ \nu_\mu \rightarrow D^- \mu^+ \nu_\mu X$	$(0.07 \pm 0.02)\%$
$B^+ \rightarrow D_1^0 \mu^+ \nu_\mu \rightarrow D^- \mu^+ \nu_\mu X$	$(0.21 \pm 0.03)\%$
$B^+ \rightarrow D_2^{*0} \mu^+ \nu_\mu \rightarrow D^- \mu^+ \nu_\mu X$	$(0.23 \pm 0.10)\%$
$B^+ \rightarrow D^- \pi^+ \mu^+ \nu_\mu$	$(0.04 \pm 0.04)\%$
$B^+ \rightarrow D^- \pi^0 \pi^+ \mu^+ \nu_\mu$	$(0.05 \pm 0.05)\%$
$B^+ \rightarrow D^{*-} (\rightarrow D^- \pi^0 / \gamma) \pi^+ \mu^+ \nu_\mu$	$(0.03 \pm 0.03)\%$
$B^+ \rightarrow D^{*-} (\rightarrow D^- \pi^0 / \gamma) \pi^0 \pi^+ \mu^+ \nu_\mu$	$(0.04 \pm 0.04)\%$
$B^+ \rightarrow D_2^{*0} \tau^+ \nu_\tau \rightarrow D^- \mu^+ \nu_\mu X$	$(0.02 \pm 0.01)\%$
<b>overall</b> $B^+ \rightarrow D^- \mu^+ X$	$(0.97 \pm 0.16)\%$

Table 5.11: The main contributions to the total branching fractions used in the generation of the  $B^+$  background sample in the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  mode. The values are taken from the PDG [99], and the branching fraction of the subsequent decay of the  $D$  mesons produced in the  $B$  decay, indicated between brackets, are taken into account.

Process	Branching fraction
$B^+ \rightarrow D_0^{*0} \mu^+ \nu_\mu \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu X$	$(0.03 \pm 0.1)\%$
$B^+ \rightarrow D_1^0 \mu^+ \nu_\mu \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu X$	$(0.09 \pm 0.03)\%$
$B^+ \rightarrow D_1^0 \mu^+ \nu_\mu \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu X$	$(0.22 \pm 0.03)\%$
$B^+ \rightarrow D_2^{*0} \mu^+ \nu_\mu \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu X$	$(0.03 \pm 0.03)\%$
$B^+ \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \pi^+ \mu^+ \nu_\mu$	$(0.06 \pm 0.06)\%$
$B^+ \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \pi^0 \pi^+ \mu^+ \nu_\mu$	$(0.08 \pm 0.08)\%$
<b>overall</b> $B^+ \rightarrow D^{*-} (\rightarrow \bar{D}^0 \pi^-) \mu^+ \nu_\mu X$	$(0.58 \pm 0.12)\%$

candidates, after passing the complete signal selection, are classified according to the generated decay. Background events can be categorized as follows:

- *Prompt  $D$  decays associated with a random or fake muon*, which cannot be associated with a true muon from a  $B$  decay. The fraction of these events is estimated to be  $(0.6 \pm 0.3)\%$  in the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  sample, and  $(1.2 \pm 0.4)\%$  in the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  sample, when applying the full selection except the  $\log(\text{IP})(D)$  cut. When including the  $\log(\text{IP})(D)$  cut, no background decays of this type are left in the simulated samples.

- *Misidentified decays*, when the reconstructed  $D$  decay products are the products of the generated decay, but are not correctly identified as kaon and pion of the reconstructed decay. This background is negligible in both samples.
- *Combinatorial candidates*, when the reconstructed  $D$  daughters do not originate from a simulated  $D$  meson. The fraction of this background is about 3% for the  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$  sample and about 15% for the  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  sample. These fractions are measured more precisely from data exploiting the signal versus background separation in the  $D^-$  ( $\bar{D}^0$ ) mass distribution.
- *$D$  decays generated from other  $B$  meson decays*, when the  $D$  meson, at generation level, is not a decay product of a  $B^0$ . In most of the cases these background decays are  $B^+$  decays. Sec. 5.3 explains in detail how the fraction of  $B^+$  decays is estimated using simulated samples and data-driven techniques. These decays are separately modeled in the analysis. Inclusive MC samples are useful to estimate the fraction of  $B_s^0$  decays possibly present in the signal samples. These decays look similar to the  $B^+$  decays when considering the charge asymmetry of the final state, with the difference that the production asymmetry is expected to be zero (see Sec. 9.1.4). The expected fraction of  $B_s^0$  decays is  $(1.6 \pm 0.5)\%$  for the  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$  decay mode, and  $(1.5 \pm 0.5)\%$  for the  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  decay mode. This component is not included when determining the central value of  $a_{\text{sl}}^d$ , but a systematic uncertainty is evaluated by including a contribution of 2%  $B_s^0$  decays.
- *Other  $B^0 \rightarrow D$  decays reconstructed as  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  or  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$* . This is the case of true  $B^0 \rightarrow D^{(*)+}D_{(s)}^{(*)+}X$  decays where the  $D_{(s)}^{(*)+}$  decays semi-muonically. These are real  $B^0$  decays, and can therefore be treated as signal. The only inconvenience can be that these are not included in the signal cocktail MC samples used to model the decay time resolution. A systematic uncertainty is calculated to account for this effect. These events are expected to contribute  $(0.9 \pm 0.4)\%$  to the selected sample.

### 5.3 $B^+$ fraction estimation

One of the major challenges of this analysis is the treatment of the  $B^+$  background. Semileptonic  $B^+$  decays have a topology very similar to the signal decays, therefore they are not easily distinguished. On the other hand,  $B^+$  decays are expected to have a different charge asymmetry with respect to the signal: they do not oscillate, therefore no  $CP$  asymmetry from mixing is expected, and the production asymmetry of  $B^+$  mesons is different from the production asymmetry of  $B^0$  mesons. It is crucial to understand which fraction of the data sample selected as signal is contaminated by  $B^+$  decays. It is possible to obtain this number using two different methods. The first exploits known branching ratio values and selection efficiencies evaluated on simulated events. The second approach is data driven. The two methods are explained in detail in the following two sections.

Table 5.12: The branching ratios corresponding to the generated MC cocktails [100], [101]. The first uncertainty comes from the measured uncertainties on the branching fractions [99]. The second uncertainty on the semileptonic  $B$  cocktails is obtained from varying the branching fractions for the different  $D^{**}$  decay models.

Sample	Branching fraction
Signal $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$ cocktail	$(4.07 \pm 0.20 \pm 0.08)\%$ $\times \mathcal{B}(D^0 \rightarrow K^- \pi^+)$
Signal $B^0 \rightarrow D^- \mu^+ \nu_\mu X$ cocktail	$(4.95 \pm 0.23 \pm 0.08)\%$ $\times \mathcal{B}(D^- \rightarrow K^+ \pi^- \pi^-)$
Bkg $B^+ \rightarrow D^0 \mu^+ \nu_\mu X$ cocktail	$(10.33 \pm 0.41 \pm 0.10)\%$ $\times \mathcal{B}(D^0 \rightarrow K^- \pi^+)$
Bkg $B^+ \rightarrow D^{*-} \mu^+ \nu_\mu X$ cocktail	$(0.58 \pm 0.12 \pm 0.06)\%$ $\times \mathcal{B}(D^0 \rightarrow K^- \pi^+)$
Bkg $B^+ \rightarrow D^- \mu^+ \nu_\mu X$ cocktail	$(0.97 \pm 0.16 \pm 0.08)\%$ $\times \mathcal{B}(D^- \rightarrow K^+ \pi^- \pi^-)$

Given the insufficient handle in the global fit to determine simultaneously the yield of the  $B^+$  decays, the fraction of  $B^+$  decays in the sample is fixed to the value estimated. A systematic uncertainty accounting for this assumption is evaluated.

### 5.3.1 $B^+$ fraction estimated from MC

The inclusive semileptonic branching fractions of  $B^0$  and  $B^+$  into the reconstructed final state are given in Table 5.12. They are obtained by summing the branching fractions, based on the PDG [99], used to produce the cocktails [100], [101]. The first uncertainty is due to the uncertainties on the branching fraction measurements. The second uncertainty is obtained by varying the branching fractions for the different  $D^{**}$  decay models.

The total efficiencies calculated on Monte Carlo samples for 2012 data acquisition conditions and accounting for  $\mu\pi$  kinematic weights, for each mode are (including the statistical uncertainty only)

$$\begin{aligned}
 B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X &: (0.1633 \pm 0.0006)\% , \\
 B^+ \rightarrow D^0 \mu^+ \nu_\mu X &: (0.0085 \pm 0.0002)\% , \\
 B^0 \rightarrow D^- \mu^+ \nu_\mu X &: (0.2401 \pm 0.0011)\% , \\
 B^+ \rightarrow D^- \mu^+ \nu_\mu X &: (0.1777 \pm 0.0015)\% .
 \end{aligned} \tag{5.2}$$

$$\tag{5.3}$$

Because of isospin symmetry,  $B^+$  and  $B^0$  should have the same production cross section. The fraction of  $B^+$  events in the final selection can then be estimated as f

$$f_{B^+} = \frac{N_{B^+}}{N_{B^0} + N_{B^+}} , \tag{5.4}$$

where  $N_{B^0}$  is the number of  $B^0$  events and  $N_{B^+}$  is the number of  $B^+$  events. The number of events are obtained, taking as example the  $D^- \mu^+$  final state, as

$$N_{B^+} = \epsilon(B^+ \rightarrow D^- \mu^+ \nu_\mu X) \times \mathcal{B}(B^+ \rightarrow D^- \mu^+ \nu_\mu X); \tag{5.5}$$

$$N_{B^0} = \epsilon(B^0 \rightarrow D^- \mu^+ \nu_\mu X) \times \mathcal{B}(B^0 \rightarrow D^- \mu^+ \nu_\mu X), \tag{5.6}$$

where  $\epsilon$  indicates the total selection efficiency, reported for the signal in Tab. 5.9, and  $\mathcal{B}$  indicate the branching fractions used in the MC cocktails reported in Tab. 5.12.

For the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  mode we determine the number of  $B^+ \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays  $N_{B^+}$ , using a  $B^+ \rightarrow D^0 \mu^+ \nu_\mu X$  MC sample. Therefore,

$$N_{B^+} = \epsilon(B^+ \rightarrow D^{*-} \mu^+ \nu_\mu X) \times \mathcal{B}(B^+ \rightarrow D^{*-} \mu^+ \nu_\mu X), \quad (5.7)$$

becomes

$$N_{B^+} = \epsilon(B^+ \rightarrow D^{*-} \mu^+ \nu_\mu X | B^+ \rightarrow D^0 \mu^+ \nu_\mu X) \quad (5.8)$$

$$\times \frac{\mathcal{B}(B^+ \rightarrow D^0 \mu^+ \nu_\mu X)}{\mathcal{B}(B^+ \rightarrow D^{*-} \mu^+ \nu_\mu X)} \times \mathcal{B}(B^+ \rightarrow D^{*-} \mu^+ \nu_\mu X), \quad (5.9)$$

where  $\epsilon(B^+ \rightarrow D^{*-} \mu^+ \nu_\mu X | B^+ \rightarrow D^0 \mu^+ \nu_\mu X)$  represents the total selection efficiency of  $B^+ \rightarrow D^{*-} \mu^+ \nu_\mu X$  in the more inclusive sample of  $B^+ \rightarrow D^0 \mu^+ \nu_\mu X$  simulated events. Although,  $\mathcal{B}(B^+ \rightarrow D^{*-} \mu^+ \nu_\mu X)$  cancels in Eq.5.8, the uncertainty on the ratio of the branching ratios,  $\sigma\left(\frac{\mathcal{B}(B^+ \rightarrow D^0 \mu^+ \nu_\mu X)}{\mathcal{B}(B^+ \rightarrow D^{*-} \mu^+ \nu_\mu X)}\right) \sim 0.2107$ , is taken into account. A small difference is observed in the selection efficiencies between simulated events with PYTHIA 6 and PYTHIA 8. The difference in the  $B^+$  fraction is  $(1.0 \pm 0.3)\%$  for the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  mode, and  $(0.61 \pm 0.13)\%$  for the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  mode. The estimated  $B^+$  fractions are

$$f_{B^+}(B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X) = (8.8 \pm 2.0 \pm 1.0)\%, \quad (5.10)$$

$$f_{B^+}(B^0 \rightarrow D^- \mu^+ \nu_\mu X) = (12.7 \pm 2.1 \pm 0.6)\%, \quad (5.11)$$

where the first uncertainty is due to the branching fractions, and the second is the systematic uncertainty ascribed to the difference between PYTHIA versions.

### 5.3.2 $B^+$ fraction estimated from data

The  $B^+$  fraction estimated from the MC samples, is cross-checked with a data-driven method. The difference between the corrected mass spectra for  $B^+$  and  $B^0$  mesons is used. The corrected mass is defined as

$$m_{\text{corr}} = \sqrt{M_{D^{\pm(*)}\mu}^2 + |p_T^{\text{mis}}|^2 + |p_T^{\text{mis}}|}, \quad (5.12)$$

where  $p_T^{\text{mis}}$  is the missing transverse momentum relative to the flight direction of the  $B$  meson. The corrected mass variable is more suitable than the reconstructed mass to account for a possibly missing decay product, and in this case provides a good separation between the  $B^+$  and  $B^0$ . The invariant mass of the decay products, in this case  $K^+ \pi^- \pi^- \mu^+$ , is also referred to as *visible*  $B^0$  mass or reconstructed  $B^0$  mass, and it is shown, for the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  sample in Fig.5.6

The  $B^+$  fraction is extracted from a fit to the corrected mass distribution in data, using the shapes of the corrected mass distributions in simulated  $B^+$  and  $B^0$  events as

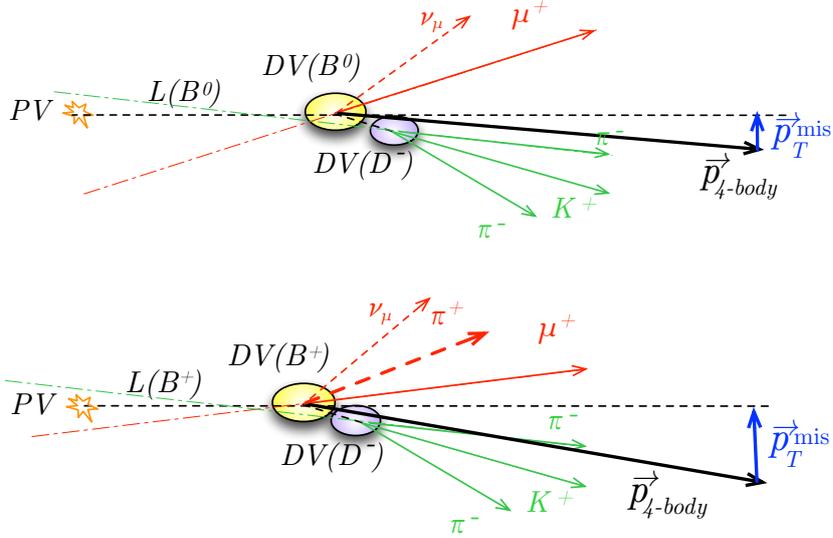


Figure 5.5: Sketch of signal and  $B^+$  background decays reconstructed in  $D^- \mu^+ X$  sample. Solid lines are track reconstructed, dashed lines are not (except for the  $B$  flight direction indicated with the dashed black line). The definition of  $p_T^{\text{mis}}$  is highlighted. More tracks are missed in the reconstruction of a decay, larger  $|p_T^{\text{mis}}|$  values are expected and lower values are expected for the corrected mass (see Eq. 5.12). For these two examples the corrected mass distributions can be visualized from simulated events in Fig. 5.7.

models. Moreover, a combinatorial background component needs to be accounted for: The corrected mass distribution of the same-sign events (events with  $D^\pm$  and  $\mu^\pm$  with the

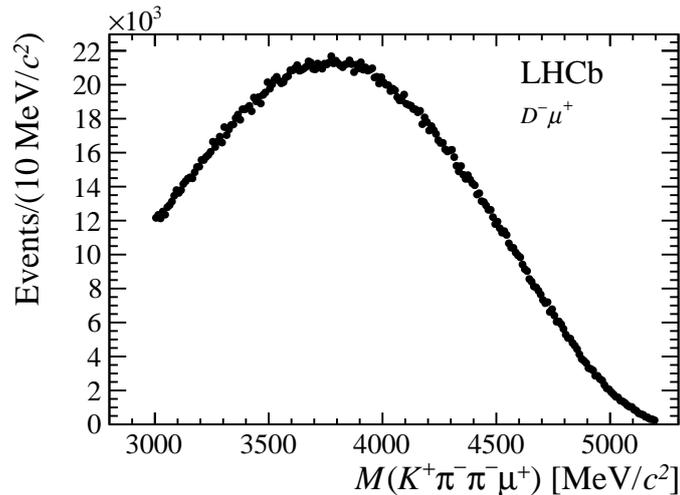


Figure 5.6: Reconstructed  $B^0$  mass distribution for the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  data sample.

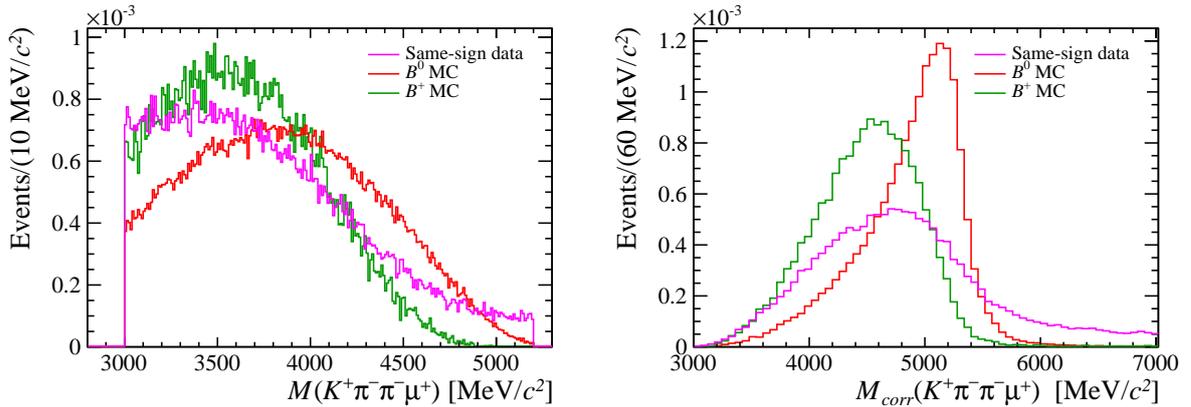


Figure 5.7: Templates for the various components of fits to the visible  $B$  mass (*Left*) and corrected  $B$  mass (*Right*). The distributions are normalized to 1. Despite a separation between  $B^0$  and  $B^+$  component is already present in the visible  $B$  mass, the corrected  $B$  mass variable provides definitely a better separation. A possible template for the combinatorial background is also reported. It is obtained from the same-sign data sample.  $B^+$  and  $B^0$  templates are taken from MC. Decays to  $D^-\mu^+X$  are reconstructed.

same charge, already used in Sec. 5.1.5) are used.

Fig. 5.7 shows the visible and corrected  $B$  mass shapes that are used in a fit to extract the  $B^+$  fraction. Decays to  $D^-\mu^+X$  are reconstructed, and the full statistics of MC samples and same-sign data events is used. The corrected  $B$  mass exhibits a more clear separation between the simulated  $B^0$  events and  $B^+$  events, with respect to the  $B$  visible mass. As expected, the corrected mass of  $B^+$  decays reconstructed as signal presents a peaking structure at lower values than the corrected mass of signal decays. This is due to the charged track that is missing in the reconstruction. The tail at higher corrected mass values is due to combinatorial background, as the model obtained from the same-sign decays shows.

Fig 5.8 shows the corrected  $B$  mass data distributions with the fit models (the histograms reported in Fig. 5.7(*right*)) superimposed, for both  $D^-\mu^+X$  and  $D^{*-}\mu^+X$  data samples. For  $D^-\mu^+X$  decays the simple models presented here are good enough to obtain an estimate of the  $B^+$  fraction and the fraction of combinatorial background, even though a large systematic uncertainty on the fraction is expected. The estimate for the  $B^+$  fraction is  $f_{B^+}(B^0 \rightarrow D^-\mu^+\nu_\mu X) = (16.1 \pm 0.1(\text{stat}))\%$ , while the combinatorial background is estimated to be  $(1.8 \pm 0.6(\text{stat}))\%$ , reporting only the statistical uncertainty. The sources of systematic uncertainty have to be identified essentially in the uncertainty on the templates, that is not included in these fits, and in the lack of components which are present in the samples, but not modeled in the fit. The correct procedure to account for the first effect would be to allow for variations of the height of each bin of the histograms used as models, within the uncertainty. This is normally obtained with the Beeston-Barlow approach [103]. Concerning the components not included in the

fits, the expected contributions are at the percent level.  $B_s^0$  decays are provided as an example. These decays are expected to be about  $(1.5 \pm 0.5)\%$  of the data sample, from Monte Carlo studies. Another possible background processes are  $\Lambda_b \rightarrow D^- \mu^+ X_n$ , with  $X_n$  any neutral baryonic state, that could contribute up to  $2_{-1}^{+2}\%$  of the selected sample (see Sec. 9.1.4). In addition also  $B \rightarrow D^{(*)-} D_s^+$  background events could also be about 2% of the selected data sample (see Sec. 9.1.4). Consequently, the systematic uncertainty of the  $B^+$  fraction determined with this method is expected to be at the few per-cent level. This estimate of the  $B^+$  fraction should be compared to the result of Sec. 5.3.1,  $f_{B^+}(B^0 \rightarrow D^- \mu^+ \nu_\mu X) = (12.7 \pm 2.2)\%$ . When considering the possible systematics, the two results are compatible.

For  $D^{*-} \mu^+ X$  the conclusions are similar. This sample includes only about 20% of the statistics of the  $D^- \mu^+ X$  sample, while featuring an extremely low level of combinatorial background. The first characteristics results in larger uncertainties on the MC templates, while the second leads to very low sensitivity to the fraction of combinatorial background. For this reason, an optimal fit result is obtained when excluding the combinatorial background component. To reduce the combinatorial background present in the data sample, the background events with a  $D^0$  mass higher and lower than the  $D^0$  mass peak, have been statistically subtracted (the  $B$  corrected mass distribution considered is  $D^0$  *sideband-subtracted*<sup>6</sup>). The result for the  $B^+$  fraction is  $f_{B^+}(B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X) = (8.0 \pm 0.2(\text{stat}))\%$  using the data-driven method, to be compared to  $f_{B^+}(B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X) = (8.8 \pm 2.2)\%$  using the estimate from MC. Also in this case a systematic uncertainty of the few per-cent level is expected. Therefore, improvements are needed before this data-driven method could be used to reliably extract the  $B^+$  fraction, but these studies show that the corrected  $B$  mass can be used in the future to discriminate the  $B^0$  signal against  $B^+$  background, once all the systematic uncertainties are under control.

## 5.4 Control Samples

Control samples are used for the measurement of the detection asymmetries. For the  $K\pi$  charge detection asymmetry, prompt Cabibbo-favored  $D^- \rightarrow K^+ \pi^- \pi^-$  and  $D^+ \rightarrow \bar{K}^0 \pi^+$  decays are used, while  $J/\psi$  decays are used to evaluate the muon trigger and

---

<sup>6</sup>The so-called sideband-subtraction is a common procedure used to statistically subtract background events in a given distribution. The necessary condition is that the type of background to be rejected exhibits a distribution very different from the signal in at least one variable. The simplest case, from which the sideband-subtraction took its nomenclature, is when one variable is a clean gaussian mass peak, as the  $D^\pm$  or  $D^0$  in this analysis, and the variable of interest has any shape, including the  $B$  visible or corrected mass as in the present case. The assumption is that the background events in the mass range of the  $D^\pm$  or  $D^0$  mass peak exhibit the same behavior in the variable of interest as the events in regions at the left and right of the mass peak. The events in the  $D^0$  sidebands include both a pure combinatorial background, and real  $B$  decays that don't include a  $D^0$  meson in the final state. For this reason it is not optimal to use a  $D^0$  sidebands subtraction to estimate the  $B^+$  fraction in this case, nevertheless the combinatorial component is reduced by this sideband subtraction.

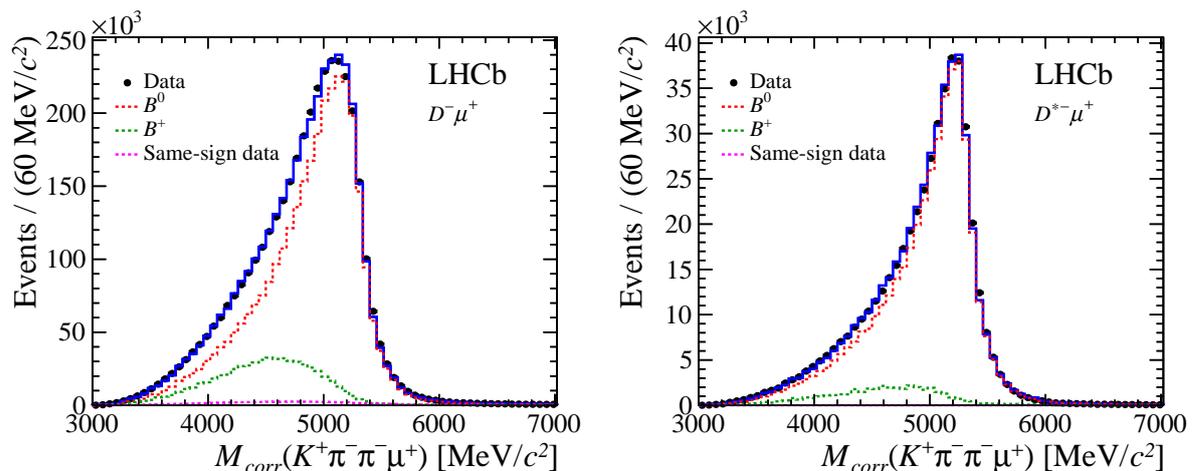


Figure 5.8: *Top* Fits to the  $B$  corrected mass distribution used to extract the  $B^+$  fraction for the  $D^-\mu^+X$  (Left) and  $D^{*-}\mu^+X$  (Right) samples with a data driven method. In red the  $B^0$  component, green the  $B^+$  component, both taken from MC. In purple the combinatorial background shape obtained from "same-sign" events.

misidentification asymmetries. In this Section these decay samples are introduced, the studies performed and the results used in the  $a_{\text{sl}}^d$  measurement are presented in Chapter 6.

In  $D^+ \rightarrow \overline{K^0}\pi^+$  decays the flavor of the neutral kaon is determined at the production, and subsequently evolves in a superposition of  $K_S^0$  and  $K_L^0$  states. In this analysis only the decay to the  $CP$  eigenstate  $\pi^+\pi^-$  is reconstructed, for this reason the neutral kaon is referred to as  $K_S^0$ . Consequently, this sample is here indicated also as  $D^+ \rightarrow K_S^0\pi^+$ . The lifetime of the  $K_S^0$  is quite long, resulting in only about 25% of the  $K_S^0$ s decaying inside the VELO. In this case the pions are reconstructed as long tracks, while if the decay occurs between VELO and TT, the pion tracks are reconstructed as downstream tracks. Decays occurring after the TT are not reconstructed.  $K_L^0$  mesons have a lifetime about a factor 50 larger than the  $K_S^0$  lifetime, therefore they hardly decay within the LHCb acceptance. Neutral  $K^0$  and  $\overline{K^0}$  transverse a large amount of the material of the LHCb detector, and similarly to charged kaons, also  $K^0$  and  $\overline{K^0}$  have different material interaction rates. In addition  $CP$  violation in the  $K_S^0 \rightarrow \pi^+\pi^-$  decay needs to be accounted for. All these effects need to be kept under control, as explained in Sec. 6.6. For  $CP$  violation measurements as  $a_{\text{sl}}^d$  the downstream reconstructed  $K_S^0$  decays exhibit large detection asymmetries, for this reason only decays reconstructed using long tracks are considered for this measurement. The measurement of the  $K\pi$  pair detection asymmetry using prompt  $D^- \rightarrow K^+\pi^-\pi^-$  and  $D^+ \rightarrow \overline{K^0}\pi^+$  decays, is explained in Sec. 6.6. The method exploited relies on the assumption that some charge asymmetries are exactly the same in the two samples, and therefore cancel when considering the difference of the two. One of these is the charge asymmetry induced by the selection. For this reason the selection criteria for these two

samples need to ensure the same charge asymmetry. Concerning the trigger requirements, at Level-0 it is required that events are triggered independently of the signal candidate (LOGLOBAL TIS), given that the resolution of the hadronic calorimeter is not sufficient to trigger on a single particle, and in addition any charge asymmetry of TIS triggered events affects  $D^- \rightarrow K^+\pi^-\pi^-$  and  $D^+ \rightarrow K_s^0\pi^+$  decays in the same manner. Subsequently in HLT1 only candidates where a pion of the  $D^-$  decay triggered the HLT1TRACKALL0 line are selected. In HLT2 exclusive lines are used. Further selection requirements can be summarized as follows:

- $D^- \rightarrow K^+\pi^-\pi^-$  decays.<sup>7</sup> In addition the  $D$  meson is required to be TOS, and the pion with highest momentum is required to have  $1.6 \text{ GeV}/c < p_T < 7 \text{ GeV}/c$  and  $1.9 < \eta < 4.9$ , TOS on the HLT1TRACKALL0 line, and to satisfy the PID requirement  $DLL_{K\pi} < 0$ .
- $D^+ \rightarrow K_s^0\pi^+$  decays.<sup>8</sup> In addition the  $D$  meson is required to be TOS on the HLT2 trigger line used<sup>9</sup>, and the pion is required to have  $1.6 \text{ GeV}/c < p_T < 7 \text{ GeV}/c$  and  $1.9 < \eta < 4.9$ , TOS on the HLT1TRACKALL0 line, and to satisfy the PID requirement  $DLL_{K\pi} < 0$ .

For the muon trigger and misidentification asymmetries, the trigger selection of the signal decays needs to be probed, therefore it needs to be compared to a global TIS selection, as is discussed in Sec. 6.5.2. The central offline selection is here summarized. Two different data samples have been used:

- *sample 1*) inclusive detached  $J/\psi$ .<sup>10</sup> Both the muons are required to be long tracks and have  $p_T > 0.8 \text{ GeV}/c$  and  $p > 3.0 \text{ GeV}/c$ , and satisfy some quality requirements as  $\text{Track } \chi^2/\text{ndf} < 3$ ,  $\text{min. IP}\chi^2(\text{PV}) > 10$ . One of the two muons is required to be identified as muon, have  $p_T > 1.5 \text{ GeV}/c$ ,  $p > 6.0 \text{ GeV}/c$ ,  $\text{IP}\chi^2 > 25$ . Concerning the selection requirements of the  $J/\psi$ , the reconstructed mass should be in a  $400 \text{ MeV}/c^2$  window centered on the known mass of the  $J/\psi$ , the  $\text{Vertex } \chi^2/\text{ndf} < 8.0$ , and the decay vertex should be displaced by more than  $5\sigma$  from the PV.
- *sample 2*)  $J/\psi$  making a vertex with an extra displaced high  $p_T$  track.<sup>11</sup> This line foresees several additional requirements on the  $J/\psi$ , additional track and  $B$  mother, that result in a better background rejection. These criteria include higher  $p$  and  $p_T$  requirements for the muons, tighter mass window and minimum  $p_T$  for the  $J/\psi$ , and the parent particle, likely a  $B$  meson. Displacement of the  $J/\psi$  and  $B$  and quality requirements on the additional track originating from the  $B$  also belong to the selection used.

The two samples have a large overlap, therefore the results obtained are not combined, but are each crosscheck for the other.

<sup>7</sup>TheD2HHH KPPLINE stripping line is used.

<sup>8</sup>The D2KS0K\_PIONLINE stripping line is used.

<sup>9</sup>HLT2CHARMHADD2KS0H\_D2KS0PI

<sup>10</sup>The JPSIFROMBNOPIDNOMIP stripping line is used.

<sup>11</sup>The SEMIINCJPSI2MUMUJPSIFORSL stripping line is used.

## 5.5 Summary

The first  $a_{\text{sl}}^d$  measurement at LHCb uses semileptonic samples of  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays collected at LHCb in 2011 and 2012 at a center-of-mass energy of 7 TeV and 8 TeV, respectively. These datasets correspond to the full  $3 \text{ fb}^{-1}$  acquired in Run-I. In the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  sample,  $D^-$  meson decays are selected in the Cabibbo-favored mode  $D^- \rightarrow K^+ \pi^- \pi^-$ , and in the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  sample,  $D^{*-}$  meson decays to  $(\bar{D}^0 \pi^-)$  and  $\bar{D}^0 \rightarrow K^+ \pi^-$  are selected. These choices ensure no additional sources of  $CP$  violation in the decays.

The selected signal sample needs to satisfy the same kinematical and particle identification requirements as the control samples, which are used to measure the detection asymmetries. Precise and reliable measurements of the detection asymmetries are indeed a vital ingredient of the  $a_{\text{sl}}^d$  measurement. A second goal of the signal sample selection is to reduce to negligible level the physics background that could contaminate the sample. This second task becomes more challenging when the signal sample considered is partially reconstructed. An example of physics background events eliminated by imposing selection requirements are  $D$  decays produced directly in the primary interaction. After the whole selection, the effective  $B^0$  signal yields is 1.8 million events in the  $D^- \mu^+$  sample and 0.33 million events in the  $D^{*-} \mu^+$  sample.

Physics backgrounds not excluded from the signal selection and expected to be a non negligible component of the sample, need to be properly modeled in the analysis. This is the case of the  $B^+$  background events in this analysis. Two independent strategies are used to determine the fraction of these events present in the selected data samples. The first exploits simulated events, while the second is a data-driven technique. The fraction used in analysis determining the  $a_{\text{sl}}^d$  parameter are  $f_{B^+}(D^- \mu^+) = (12.7 \pm 2.1 \pm 0.6) \%$  for the  $D^- \mu^+$  sample and  $f_{B^+}(D^{*-} \mu^+) = (8.8 \pm 2.0 \pm 1.0) \%$  for the  $D^{*-} \mu^+$  sample.

Fractions of other possible physics backgrounds polluting the signal samples are estimated by means of Monte Carlo simulation. The values obtained are used to evaluate the systematic uncertainties on the  $a_{\text{sl}}^d$  measurement. An example is the  $B_s^0$  fraction estimated to be up to 2% of the selected data samples, a variation of the nominal fit used to extract  $a_{\text{sl}}^d$  and  $A_P$  contains also this component.

Monte Carlo studies are not only useful to estimate background fractions, but also to properly model the decay time and the decay time resolution of the signal decays. For this reason it is important that all the decay chains contributing to the signal samples are properly modeled in the MC samples. Variations of the simulated event samples compositions are also used to evaluate the systematic uncertainties.

# Chapter 6

## Detection asymmetries

In an experiment, an asymmetry in the number of reconstructed particles with respect to the number of antiparticles can arise from different interactions with the detector material. This chapter provides an overview of the detector and reconstruction effects that have an impact on the reconstructed particle versus antiparticle asymmetry relevant for the  $a_{\text{sl}}^d$  measurement (Sec. 6.1-6.3). Moreover, the description of the methods used to account for these asymmetries is given (Sec. 6.4-6.7). The development of these techniques was demanded by several analysis, including the  $a_{\text{sl}}^d$  measurements. Different collaborators provided extremely important inputs for the  $a_{\text{sl}}^d$  analysis, which are briefly presented in this chapter, together with the relevant references.

The efficiency of reconstructing a particle  $x$  with a momentum  $p$  is defined as

$$\epsilon(x, \vec{p}) = \frac{N_{\text{reconstructed}}(x, \vec{p})}{N_{\text{created}}(x, \vec{p})}. \quad (6.1)$$

In general, the reconstruction efficiency depends on the momentum of the particle. The following sections provide more details about this dependence. An asymmetry in the reconstruction efficiency of positively and negatively charged particles  $x$  with a momentum  $\vec{p}$  is defined as

$$a_{\text{P}}(x^+, \vec{p}) = \frac{\epsilon(x^+, \vec{p}) - \epsilon(x^-, \vec{p})}{\epsilon(x^+, \vec{p}) + \epsilon(x^-, \vec{p})}. \quad (6.2)$$

$A_{\text{D}}$  indicates the detection asymmetry integrated over the momentum region of interest

$$A_{\text{D}}(x^+) = \frac{\epsilon(x^+) - \epsilon(x^-)}{\epsilon(x^+) + \epsilon(x^-)}. \quad (6.3)$$

### 6.1 Interaction of particles with matter

The different interactions particles can have with matter are [17]:

*Hadronic interactions* Hadrons, charged and neutral, interact with the nuclei of the detector material via the strong interaction. The distance the hadrons have to travel in a

given material before interacting, elastically or inelastically, with the nuclei is the so-called *nuclear collision length*,  $\lambda_T$ . The rate of collisions per travelled distance  $d$  is

$$\frac{dN}{dd} = -\frac{d}{\lambda_T} = \sigma_T n d. \quad (6.4)$$

It can also be expressed as function of the cross section  $\sigma_T$  and the nuclear density  $n$  of the transversed material. The *nuclear interaction length*  $\lambda_I$  is defined excluding elastic and quasi-elastic processes. Therefore  $\lambda_I$  is normally larger than  $\lambda_T$ , but the latter quantity is more relevant, given the large angles of elastic collisions. At LHCb, about 20% of the hadrons undergo hadronic interactions before they reach the last tracking station.

*Bremsstrahlung* Charged particles lose energy in the form of photons when they are deflected by other charged particles. This effect is called Bremsstrahlung and depends on the rest mass and velocity of the deflected particle. The electrons detected in LHCb are ultra-relativistic, since particles with momentum smaller than 2 GeV are bent outside the LHCb acceptance by the dipole magnet. For ultra-relativistic electrons the energy loss is mainly due to Bremsstrahlung effects. The amount of energy lost by an electron in a material depends on the *radiation length*  $X_0$  of the material. For an electron that has transversed a distance,  $d$ , in a material with radiation length  $X_0$ , the energy loss is

$$E_e(d) = E(0)e^{-\frac{d}{X_0}}, \quad (6.5)$$

where  $E(0)$  is the initial energy of the electron. For heavier particles with mass  $M$  the energy loss due to Bremsstrahlung effect is suppressed by a factor  $\left(\frac{m_e}{M}\right)^2$ .

*Ionization* Heavy charged particles also lose energy due to the interaction with the electrons of the detector material. Atoms are either excited or ionized by the interactions. The mean rate of energy loss,  $\langle dE/dd \rangle$ , due to this type of interaction is described by the Bethe Bloch formula. For particles with a momentum magnitude in the range between 2 GeV and 100 GeV the effect is minimal, and the energy loss  $\Delta E$  in a specific material can be approximated with a linear dependence on the distance traveled by the particle in the material

$$\Delta E = \left. \frac{dE}{dd} \right|_{\min} d. \quad (6.6)$$

The typical energy loss in the detector material is few MeV per cm.

*Multiple scattering* Charged particles are deflected by the electromagnetic field of the nuclei in the detector material. A sequence of Coulomb scatterings is known as multiple scattering. This effect results in an angular dispersion that can be modeled with a Gaussian. The modified flight direction of the incident particle affects the momentum resolution in the tracking system. The relative momentum resolution at LHCb is limited by multiple scattering and is 0.4% .

The first of these effects, the hadronic interaction, introduces an asymmetry in the number of detected particles compare to antiparticles, as is explained in the following paragraph. All other effects are essentially equal for particles and antiparticles.

## 6.2 Hadronic interaction asymmetries

The final states  $K^+\pi^-\pi^-\mu^+$  and  $K^-\pi^+\pi^+\mu^-$  reconstructed for the  $a_{\text{sl}}^d$  measurement contain a single charged kaon, which is responsible for the leading contribution to the detection asymmetry of the two  $CP$  conjugated final states. The negatively charged kaon has quark content ( $\bar{u}s$ ), while the positively charged kaon has ( $u\bar{s}$ ). To a good approximation the detector contains the same number of protons and neutrons (isoscalar target). The valence quarks of the proton are ( $uud$ ) and the valence quarks of the neutron are ( $udd$ ). The sea quarks always appear in  $q\bar{q}$  pairs. Different cross sections for  $K^-$  and  $K^+$  are expected. First of all negative kaons can react with the neutrons of the material via strangeness-exchange producing a  $\Lambda^0$  baryon and a negative pion, while the analog process involving positive kaons does not exist. Moreover, while the  $s$  quark in the case of  $K^-$  and  $\bar{s}$  anti-quark of the  $K^+$  can annihilate with the respective anti-particle of the sea, the interaction is asymmetric when considering  $u$  and  $\bar{u}$ . In the case of  $K^-$ , the  $\bar{u}$  can annihilate with the valence  $u$  quark of protons and neutrons as well as the sea  $u$  quarks. On the other hand the  $u$  quark of the  $K^+$  can only annihilate with the sea  $\bar{u}$  quarks. Moreover, the parton density function of sea quarks increases at low Bjorken  $x$  (see Fig. 3.4). Therefore, the relative fraction of valence quarks which take part in the scattering process decreases for higher values of the kaon momentum, leading to a decrease of the interaction asymmetry. In the case of negative and positive pions ( $\pi^-$  contains  $\bar{u}d$  quarks and  $\pi^+$ ,  $u\bar{d}$ ) the interaction is expected to be symmetric on isoscalar targets, given that the antiquark for both  $\pi^-$  and  $\pi^+$  can annihilate with a valence quark. Fig. 6.1 shows the cross sections for negative and positive kaons and pions on deuterium, that is an isoscalar target, therefore can be used as good approximation of the detector material. Considering kaons with a momentum of about 20 GeV, the expected difference in the cross section is about 15%, which leads to an expected asymmetry of the order of (-1%), considering that 20% of the hadrons with 20 GeV momentum undergo hadronic interactions. This represents the larger detection asymmetry expected to be measured in the final states used for the  $a_{\text{sl}}^d$  measurement.

## 6.3 Detector, trigger, reconstruction effects

Besides the different interactions of particles compared to antiparticles with the detector material, also detector inefficiencies and misalignments give rise to asymmetries in detecting particles and antiparticles.

Several effects in LHCb are the consequence of having a *left-right* asymmetric detector. For example few dead sensors on one side of the detector lead to a lower detection efficiency of that side with respect to the other. Detector halves not installed fully symmetrically with

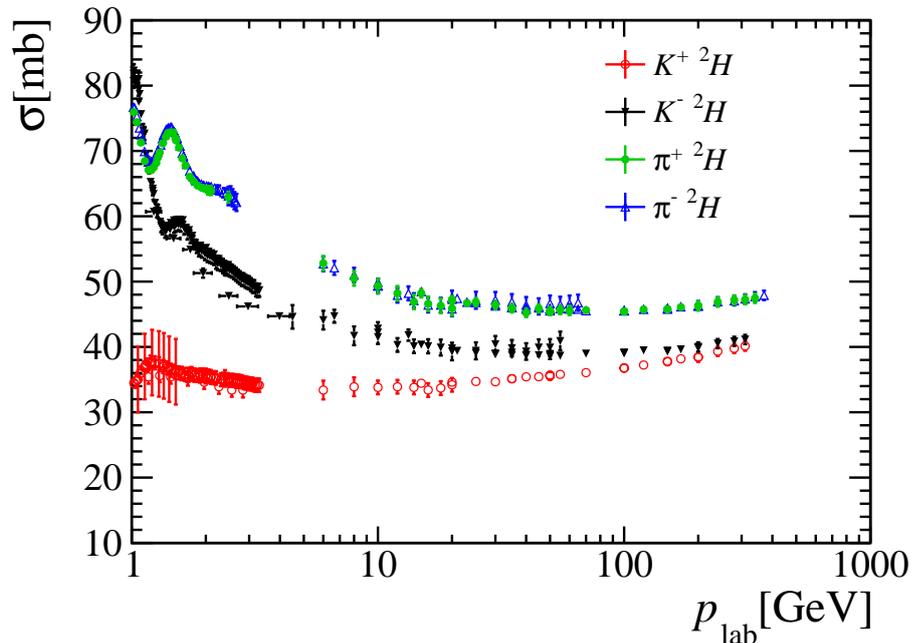


Figure 6.1: Measured kaon and pion cross sections on deuterium target, as a function of momentum. The data are taken from the COMPAS group [17]

respect to the beam pipe also lead to different detection efficiencies. This is the case of the muon stations, for example. Moreover a left-right asymmetry is introduced by the slightly asymmetric field of the dipole magnet [104], that leads to slightly different bending on the tracks to the left and right half of the detector.

The examples mentioned until now are related to the hardware, but also the event reconstruction algorithms can be source of detection asymmetries. For example slightly different requirements in the tracking algorithms [105] (Sec. 3.3.1) for tracks bent to the right side of the detector with respect to the left side, result in a charge-asymmetric reconstruction efficiency. Another example is the muon asymmetry in the Level 0 trigger (see Sec. 3.2.3), that affected the data acquired in 2011 [106].

The L0MUON hardware trigger selects an event if the largest transverse momentum measured is higher than a certain threshold. In order to understand the L0MUON charge asymmetry in 2011 data, it is necessary to understand how the transverse momentum is estimated on which the L0MUON decision is based. This is done in the L0MUON dedicated processors, by means of look-up tables which use the coordinates of the first muon stations, M1 and M2 <sup>1</sup>(see Sec. 3.2.2) . If the coordinates in the look-up tables do not correspond to the real coordinates, a wrong momentum is estimated. A relative misalignment between M1 and M2 leads to an over-estimation or under-estimation of the transverse momentum, depending on the charge of the muons. The effect is reversed

<sup>1</sup>From the position of the hits in M1 and M2 the track trajectory is determined and therefore the  $p_T$  is estimated, assuming a single kick from the magnetic field.

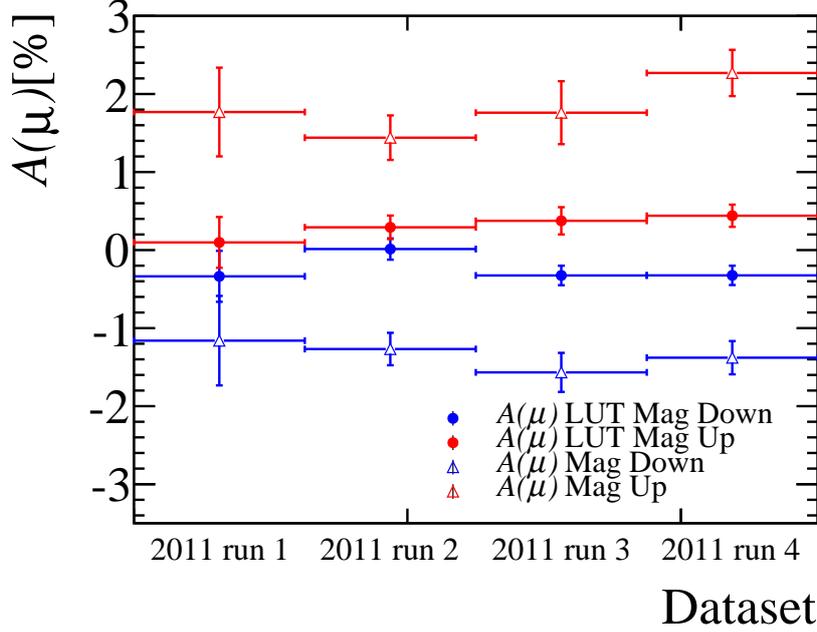


Figure 6.2: Measured muon trigger and mis-identification asymmetry,  $A(\mu) = \frac{\epsilon(\mu^+) - \epsilon(\mu^-)}{\epsilon(\mu^+) + \epsilon(\mu^-)}$ , in 2011 run periods, before and after L0MUON calibration (LUT). The method used is described in Sec. 6.5.2. The selection of the  $a_{\text{sl}}^d$  measurement is used. Before applying the LUT correction,  $A(\mu)$  is of the order of 1% and has opposite signs for different magnet polarities. After the LUT correction, the asymmetries obtained are at the per-mille level. The measured values are taken from Ref. [107].

with opposite magnet polarities. Wrong look-up tables lead to the same effect. Both problems have been present in the 2011 data taking. The cut on the largest  $p_T$  introduces different detection efficiencies, depending on the charge of the muon, and on the transverse momentum of the muon.

A correction for this dominant detection asymmetry has been made available. The L0MUON  $p_T$  is emulated with the correct look-up table. Since events that were not triggered during the data taking, cannot be recovered, a tighter cut on the calibrated L0MUON  $p_T$  is applied<sup>2</sup>, in order to compensate for the inefficiencies implied by the wrong momentum estimation (in Sec. 5.1.2 this selection requirement has been denoted with “LUT”). Fig. 6.2 shows the measured values of the muon trigger and mis-identification asymmetry,  $A(\mu) = \frac{\epsilon(\mu^+) - \epsilon(\mu^-)}{\epsilon(\mu^+) + \epsilon(\mu^-)}$ , in 2011 data, evaluated with the method explained in Sec. 6.5.2, including or not the LUT correction. For the nominal  $a_{\text{sl}}^d$  measurement the LUT selection requirement is applied. Before applying the LUT correction,  $A(\mu)$  is of the order of 1% and has opposite signs for different magnet polarities. After the LUT correction, the asymmetries obtained are at the per-mille level. This issue is not present in the 2012 data, since the first data acquired in 2012, have been used to align the muon stations.

The polarity reversal of LHCb dipole magnet plays an important role in the suppression

<sup>2</sup>The L0MUON  $p_T$  cut is tightened from 1.48 GeV to 1.64 GeV.

of the left-right asymmetries. With one polarity the positively charged particles are bent to the right side of the detectors after the magnet, and the negatively charged to the other side. With the opposite polarity the directions are swapped. The negatively charged particles are bent to the right and the positively charged to the left (see Fig. 3.7). The periodical reversal of the polarity of the magnet allows for a cancellation of the left-right effects when considering the datasets acquired with both polarities and as similar as possible detector conditions. However, not all the sources of charge detection asymmetry are perfectly canceled. Efficiencies related to the detector performance which change over time (e.g. calorimeter response that degrades with the time due to radiation effects, efficiency loss of the HPDs of the RICH detectors) are expected to give small residual asymmetries. Measurements where the detection asymmetries need to be known at the sub-percent level, cannot rely on detection asymmetry cancellation through the magnet polarity reversal.

## 6.4 Detection asymmetries for $a_{\text{sl}}^d$

The effects described in Sec. 6.2 and 6.3 play a role in the reconstruction of the final states used for the  $a_{\text{sl}}^d$  measurement. The final states  $K^+\pi^-\pi^-\mu^+$  and  $K^-\pi^+\pi^+\mu^-$  contain a single charged kaon, that leads to a detection asymmetry independent for the bending direction of the tracks and dependent on the momentum of the kaon. The detection asymmetry of  $K^-$  with respect to  $K^+$  can only be evaluated with data driven techniques. The other particles present in the reconstructed final states are also effected by the detector or reconstruction effects. The detection asymmetries of the final states need to be known at the sub-percent level, therefore it is not possible to rely to the magnet polarity reversal cancellation. The detection asymmetries are evaluated with data-driven techniques, separately for data taken with one magnet polarity and the other and in the two years of data acquisition. The final states are split in two pairs of particles, in order to measure the detection asymmetries (see Fig. 4.2). The lower momentum pion in  $B^0 \rightarrow D^-[\rightarrow K^+\pi^-\pi^-]\mu^+\nu_\mu X$  decays, or the pion from the  $D^*$  decay in the case of the  $B^0 \rightarrow D^{*-}[\rightarrow \bar{D}^0[\rightarrow K^+\pi^-]\pi^-]\mu^+\nu_\mu X$ , is paired to the kaon, and the remaining pion is paired to the muon

$$A_D = A(K\pi) + A(\mu\pi). \quad (6.7)$$

The detection asymmetry of the  $K\pi$  pair is described in Sec. 6.6, and accounts for all the possible sources of charge-asymmetric detection efficiencies. Concerning the  $\mu\pi$  pair (see Sec. 6.5 and Subsections), the detection asymmetry is evaluated separately for each contributing effect, with different procedures.

## 6.5 $\mu - \pi$ detection asymmetry

The detection asymmetry of the  $\mu\pi$  pair is defined as

$$A(\mu\pi) = \frac{\epsilon(\mu^+\pi^-) - \epsilon(\mu^-\pi^+)}{\epsilon(\mu^+\pi^-) + \epsilon(\mu^-\pi^+)}. \quad (6.8)$$

Several effects contribute to this asymmetry. The asymmetry in the nuclear interaction cross sections of positively and negatively charged pions can be neglected (see Sec. 6.2 and Fig. 6.1). The uncertainty of the PDG combination [17] for the pion nuclear interaction asymmetry is assumed. With a simulation which uses the kinematical distributions from data and tracks the particles transversing the LHCb detector using the LHCb material map, the uncertainty on the  $a_{\text{sl}}^d$  parameter has been estimated. The uncertainty obtained is 0.07%, and it is added in quadrature to the systematic uncertainty. The systematic uncertainty assumed on the  $B^0 - \bar{B}^0$  production asymmetry parameter is 0.035%. Differences in the trigger efficiency and particle identification of the muon and the particle identification of the pion contribute to the detection asymmetry of the  $\mu - \pi$  pair. Concerning the pion trigger asymmetry, no significant asymmetry is expected to be generated by the used HLT2 topological lines, as shown in [10]. Moreover the tracking efficiency of the  $\mu - \pi$  pair gives an additional contribution.  $A_{\text{D}}(\mu\pi)$  can be decomposed in

$$A(\mu\pi) = A(\mu) + A_{\text{PID}}(\pi) + A_{\text{tracking}}(\mu\pi), \quad (6.9)$$

where  $A(\mu)$  indicates the muon trigger and mis-identification asymmetry. These contributions are evaluated separately and described below.

### 6.5.1 $\mu - \pi$ tracking asymmetry

Efficiencies of reconstruct tracks depend on the momentum of the particles. At low momenta, lower tracking efficiencies are expected, due to multiple scattering effects, while at high momenta, the tracking efficiency is limited by the angular resolution, which depends on the momentum of the tracks. Hence  $A_{\text{tracking}}(\mu\pi)$  can be negligible or assume very large values, depending on the agreement between kinematic distributions of the muon and pion involved. Assuming similar momentum behaviors of the tracking efficiencies for muon and pion in function of the transverse momentum, the overlap between the  $p_{\text{T}}$  spectra of muon and pion determines the tracking asymmetry of the pair integrated over the kinematical region considered.

$$A_{\text{tracking}}(\mu^+\pi^-) = \int_{\mu \text{ } p_{\text{T}} \text{ range}} P_{\mu}(p_{\text{T}}) dp_{\text{T}} A_{\text{tracking}}(\mu^+, p_{\text{T}}) \quad (6.10)$$

$$+ \int_{\pi \text{ } p_{\text{T}} \text{ range}} P_{\pi}(p_{\text{T}}) dp_{\text{T}} A_{\text{tracking}}(\pi^-, p_{\text{T}}) \quad (6.11)$$

$$(6.12)$$

where  $P_{\mu}(p_{\text{T}})$  and  $P_{\pi}(p_{\text{T}})$  denote the muon and pion  $p_{\text{T}}$  distributions. The asymmetry due to tracking performance of a particle  $x^+$ ,  $A_{\text{tracking}}(x^+, p_{\text{T}})$  is expected to be of the opposite sign, compared to the relative antiparticle with the same transverse momentum  $p_{\text{T}}$ ,  $A_{\text{tracking}}(x^+, p_{\text{T}}) = -A_{\text{tracking}}(x^-, p_{\text{T}})$ . In addition, when neglecting the different interactions of the different particle species with the detector material, the

same tracking efficiencies can be assumed. In this case the pion interaction asymmetry is estimated to be negligible [6], and a systematic uncertainty is associated to this assumption (see Sec. 6.5). Therefore, it is assumed  $A_{\text{tracking}}(\mu^+, p_T) = A_{\text{tracking}}(\pi^+, p_T)$ , and when considering the  $\mu^+\pi^-$  pair, where muon and pion are oppositely charged,  $A_{\text{tracking}}(\mu^+, p_T) = -A_{\text{tracking}}(\pi^-, p_T)$ . Hence, Eq. 6.10 can be written as function of the muon tracking asymmetry

$$A_{\text{tracking}}(\mu^+\pi^-) = \int_0^\infty (P_\mu(p_T) - P_\pi(p_T))A_{\text{tracking}}(\mu^+, p_T)dp_T, \quad (6.13)$$

The choice of considering the pion tracking asymmetry would have been equivalently valid for the method. Since the tracking asymmetry of one of the two particles has to be measured from data, and the measurement of the muon tracking asymmetry gives a smaller uncertainty,  $A_{\text{tracking}}(\mu, p_T)$  is preferred.

Fig. 6.3 shows the kinematical distributions of the muon and highest momentum kaon for the  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  selected sample. Given the small overlap between the distributions, particularly relevant when considering the transverse momentum, and realistic tracking efficiency of pions  $\epsilon(\pi^+)$  and efficiency ratio  $\epsilon(\pi^+)/\epsilon(\pi^-)$ , as shown in Fig.6.9, the  $\mu\pi$  tracking asymmetry would be at the percent level, becoming possibly the leading contribution to the detection asymmetries<sup>3</sup>. Here is reported the approach that has been used for the first  $a_{\text{sl}}^d$  measurement at LHCb. A weighting procedure is applied to the events in order to obtain similar kinematics for the muon and the pion and therefore reduce the tracking asymmetry of the  $\mu\pi$  pair to a negligible level. It is necessary to show that this procedure does not introduce a bias in the  $a_{\text{sl}}^d$  measurement, and fulfill the task of reduce the  $\mu\pi$  tracking asymmetry to a negligible level. First of all the procedure used to assign the weights is explained, followed by the tests that show how using these weights these purposes are achieved. The procedure is different in  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  and  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$  samples, because of different kinematics.

### Determination of $\mu - \pi$ weights

Fig. 6.4 compares the muon and pion kinematic distributions in the  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  channel, after a simple one-dimensional reweighting in the  $p_T$  of the pion. The events are assigned the weight

$$w(p_T(\pi)) = \frac{P_\mu(p_T(\pi))}{P_\pi(p_T(\pi))}, \quad (6.14)$$

where  $P_\mu(p_T)$  and  $P_\pi(p_T)$  are the normalized muon and pion  $p_T$  distributions. The idea is that the pion  $p_T$  distribution moves closer to that of the muon. This re-weighting

---

<sup>3</sup>The tracking efficiency for pions  $\epsilon(\pi^+)$  and  $\epsilon(\pi^+)/\epsilon(\pi^-)$  in Fig. 6.9 are taken from the studies needed for the  $a_{\text{sl}}^d$  measurement [10] at LHCb, reported in Ref. [108]. The tracking asymmetry is the leading contribution to the systematic uncertainty on the  $a_{\text{sl}}^s$  measurement performed by LHCb [10]

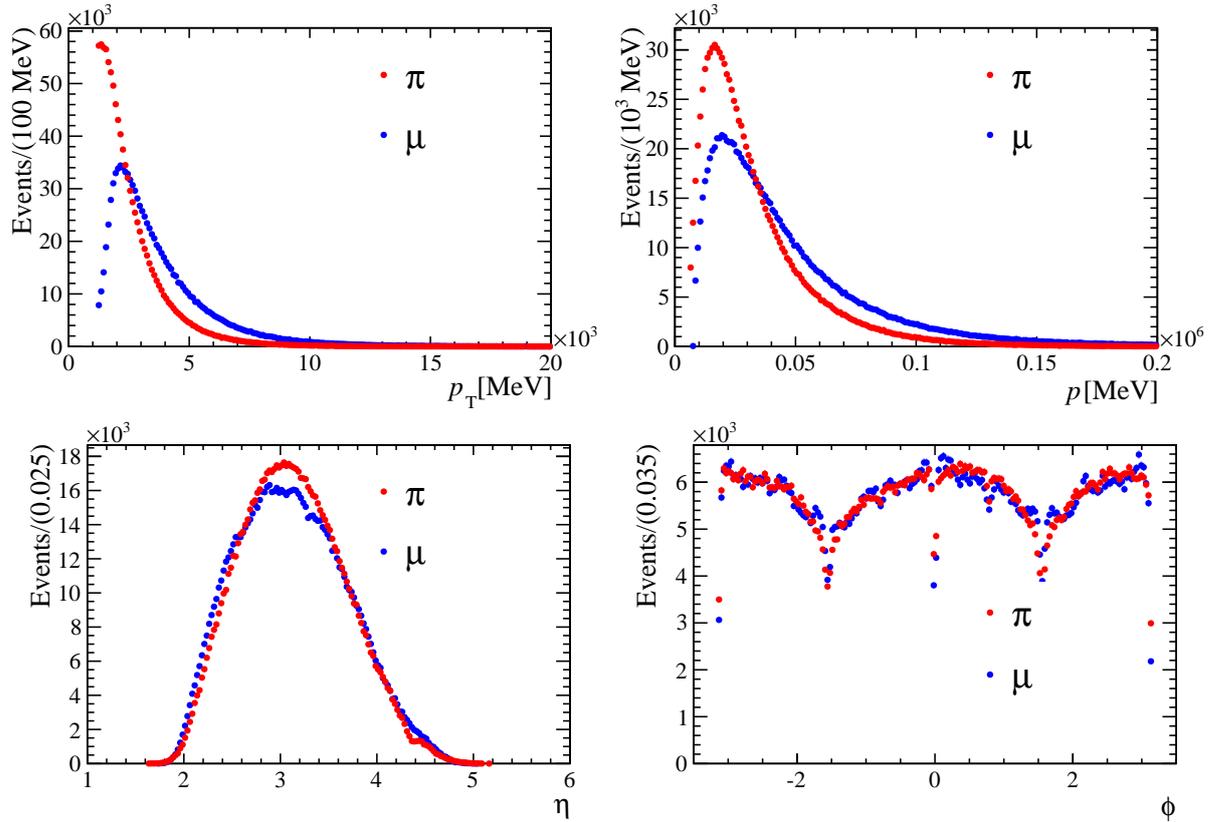


Figure 6.3: Before reweighting: kinematic distributions of the muon compared to the leading pion for  $D^+$  data sample collected in 2011

ignores that fact that the muon and pion kinematics are correlated, such that the muon distributions are also modified. This is the reason why, in Fig. 6.4 the transverse momenta spectra of muon and pion do not agree perfectly. While this could be accounted for with additional iterations of the weighting, the level of agreement is considered to be sufficient.

The weighting approach used here reduces the size of  $A_{\text{tracking}}(\mu\pi)$ , and therefore of the dependence of the measured  $a_{\text{sl}}^d$  value, on the value of  $A_{\text{tracking}}(\mu\pi)$  (see ). On the other hand, the weighting reduces the effective power of the dataset by about 40% (See Tab. 5.3) for the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  channel. Fig. 6.5 shows the distribution of weights for the 2011  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  data sample. The weights are truncated at the value of 5, in order to avoid divergent weights in the tails of the distribution that would unfairly degrade the effective statistical power of the sample. The choice to move the pion closer to the muon kinematics (rather than muon to the pion kinematics) is preferable since it assigns a lower weight to candidates in which the leading pion is soft, thus improving the purity of the sample.

In the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  channel, the muon and pion kinematics are naturally much closer, as can be seen in Fig. 6.6. For this sample a simple one-dimensional single-iteration reweighting in  $p_T$  does not improve the agreement in all distributions. Therefore, a 2-

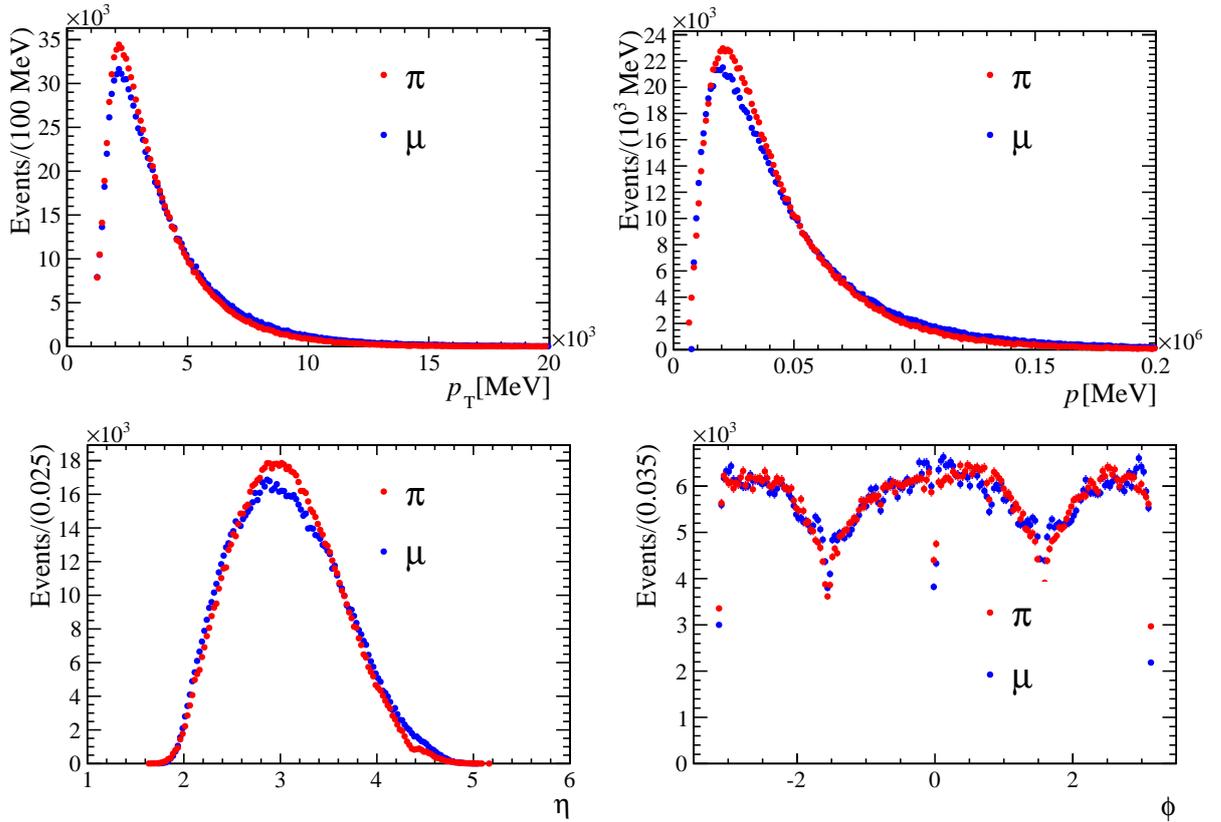


Figure 6.4: After reweighting: kinematic distributions of the muon compared to the leading pion for  $D^+$  data sample collected in 2011. The transverse momenta distribution do not show perfect agreement since muon and pion distribution are correlated, and after applying the weights not only the pion distribution is modified, but also the muon distribution changes. More iterations cure this effect, but one iteration is sufficient to reduce  $A_{\text{tracking}}(\mu\pi)$  to a negligible level.

dimensional reweighting in  $p_T$  and  $\eta$  is used (see 6.7). Again the choice is made to move the pion to be closer to the muon. Fig. 6.8 shows the muon and leading pion kinematic distributions after applying this weighting. The agreement of the pion and muon weighted distributions is better for the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decay mode, because of the better initial agreement between the distribution and because of the two-dimensional weighting scheme, which is more suitable to the decays.

### Validation of the weighting technique

A possible method to validate this weighting procedure of the signal sample makes use of so-called *pseudo-experiments* or *toys*. A pseudo-experiment is a sample of simulated events, obtained with a simplified generation of the characteristic distributions of the decays of interest. In this case the weighting procedure is verified for  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays, because on that sample the effect of the reweighting is larger. The  $B$  decay time distribution and the muon and pion momentum and transverse momentum distributions

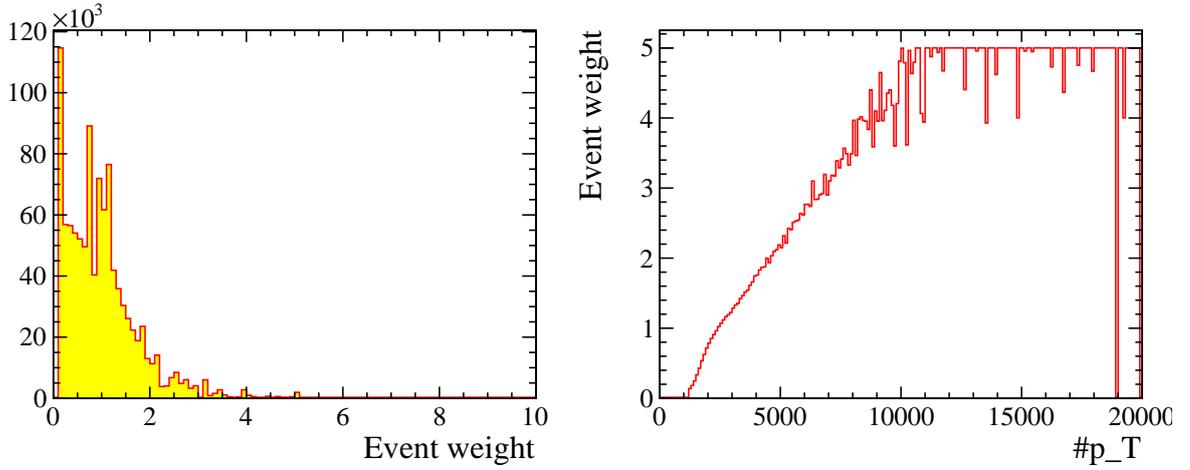


Figure 6.5:  $p_T$  weights. Event weights (*left*) and Event weights in function of  $p_T$  (*right*),  $D^+$  data sample collected in 2011

are generated for  $B^0 \rightarrow D^- \mu^+ \nu_\mu$  decays and for the  $CP$  conjugated,  $\bar{B}^0 \rightarrow D^+ \mu^- \nu_\mu$  decays. The decay time distribution of the two final states is generated according to the decay

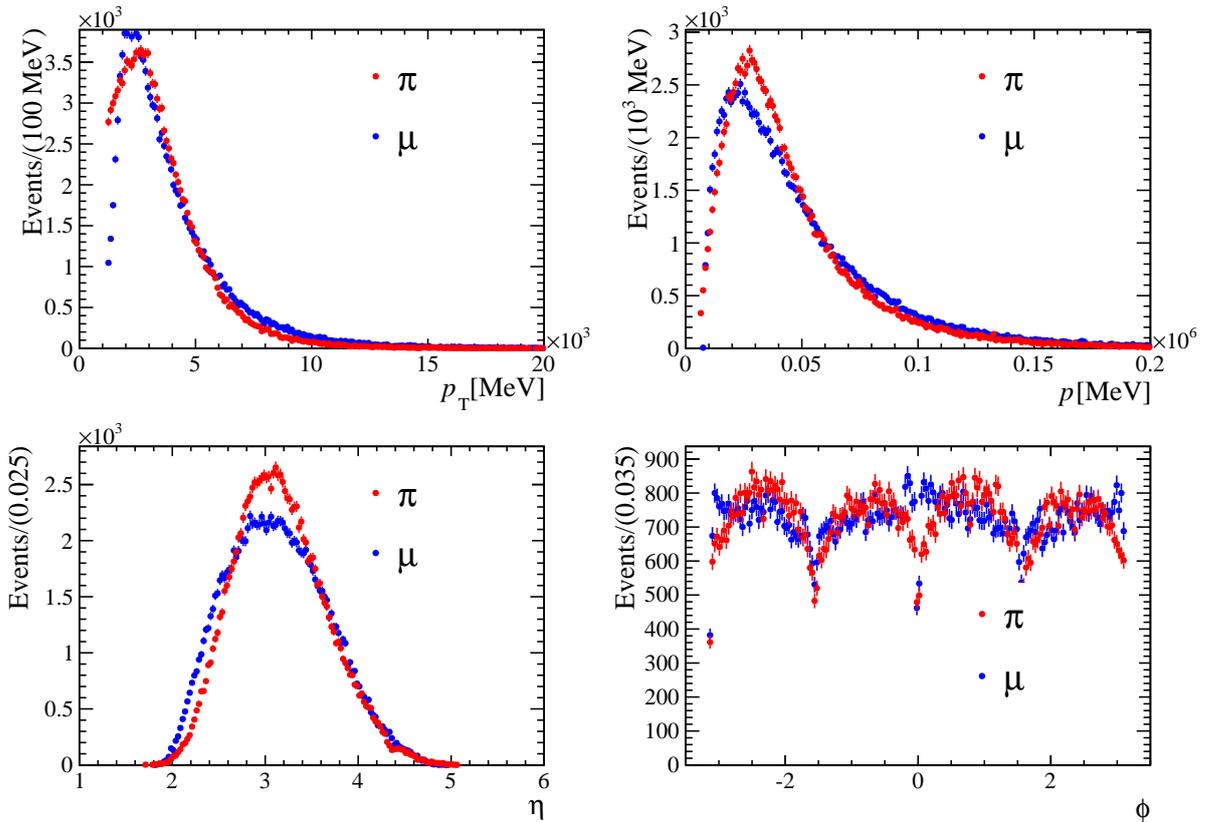


Figure 6.6: Kinematic distributions of the muon compared to the leading pion for  $D^*$  data sample collected in 2011

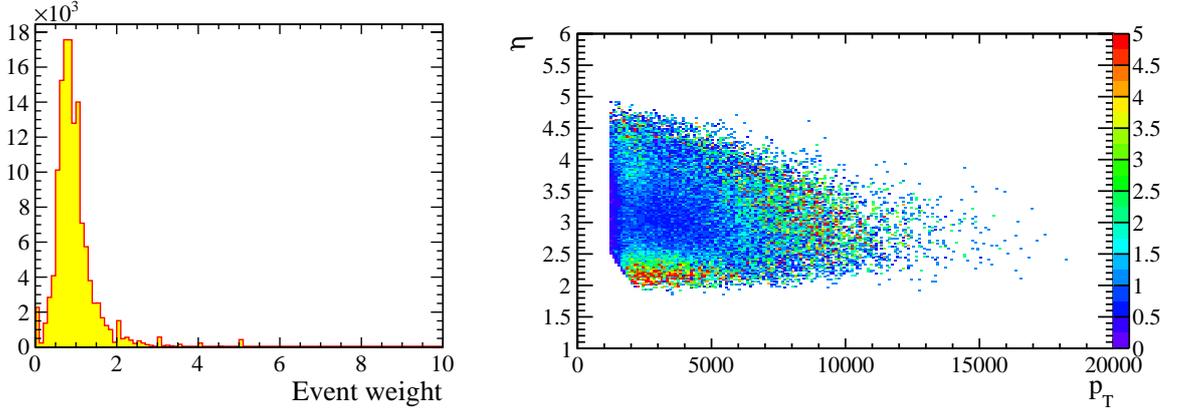


Figure 6.7:  $p_T - \eta$  weights. Event weights (left) and Event weights in function of  $p_T$  and  $\eta$  (right),  $D^*$  data sample collected in 2011

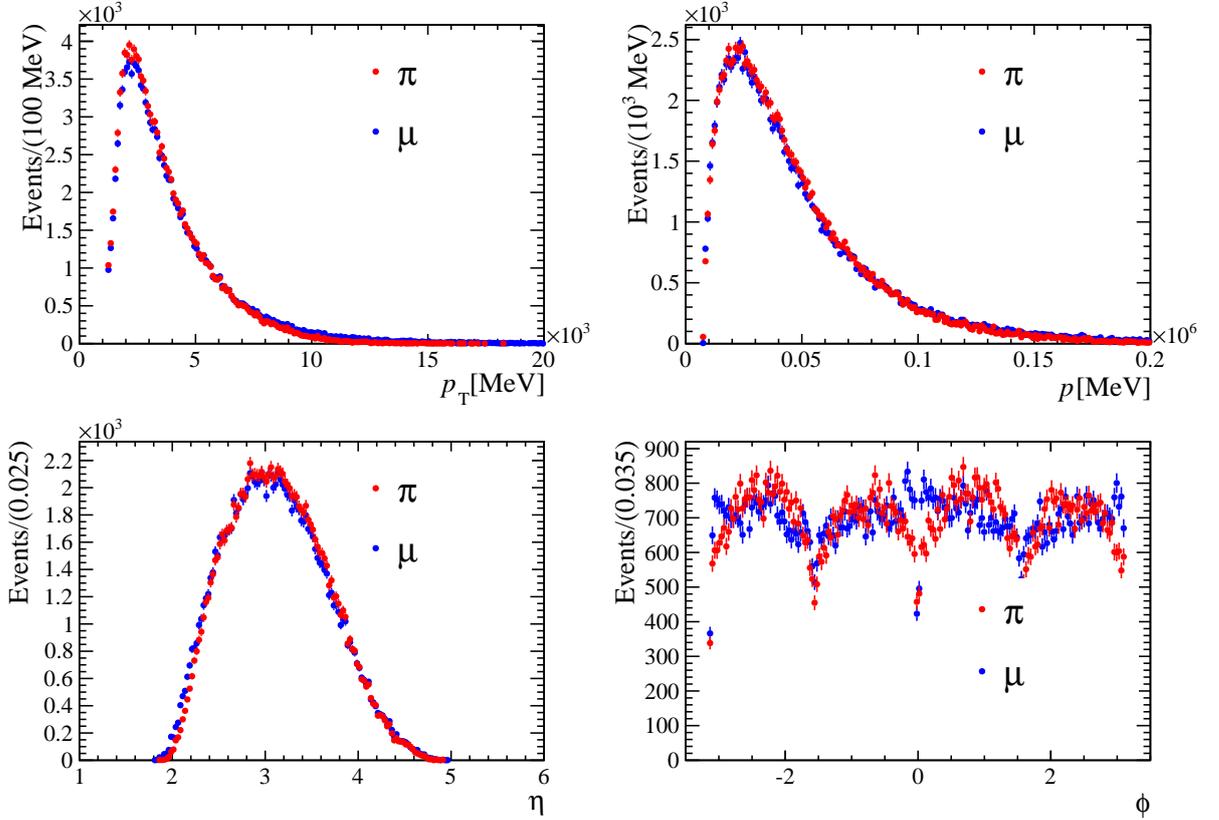


Figure 6.8: After reweighting: kinematic distributions of the muon compared to the leading pion for  $D^*$  data sample collected in 2011

rates

$$\mathcal{N}(D^- \mu^+, t) = \mathcal{N}_{\text{sig}} e^{-\Gamma_{dt}} \left( 1 + A_D + \frac{a_{\text{sl}}^d}{2} + \left( A_P - \frac{a_{\text{sl}}^d}{2} \right) \cos \Delta m_{dt} \right), \quad (6.15)$$

$$\mathcal{N}(D^+ \mu^-, t) = \mathcal{N}_{\text{sig}} e^{-\Gamma_{dt}} \left( 1 - A_D - \frac{a_{\text{sl}}^d}{2} - \left( A_P - \frac{a_{\text{sl}}^d}{2} \right) \cos \Delta m_{dt} \right), \quad (6.16)$$

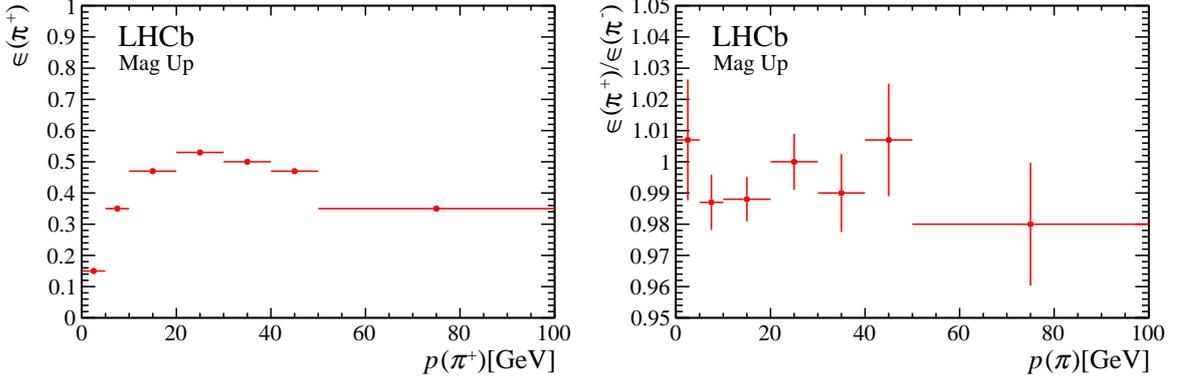


Figure 6.9: Positive pion tracking efficiency (*left*) and  $\pi^+/\pi^-$  tracking efficiency ratio (*right*). Data points taken from Ref. [108], studies needed for the  $a_{\text{sl}}^d$  measurement [10] at LHCb,

respectively. The detection asymmetry is indicated with the parameter  $A_D = \frac{\epsilon(K^+\pi^-\pi^-\mu^+) - \epsilon(K^-\pi^+\pi^+\mu^-)}{\epsilon(K^+\pi^-\pi^-\mu^+) + \epsilon(K^-\pi^+\pi^+\mu^-)}$ , and the production asymmetry. The parameter  $A_P$  denotes instead the  $B^0 - \bar{B}^0$  production asymmetry,  $A_P = \frac{N(\bar{B}^0) - N(B^0)}{N(\bar{B}^0) + N(B^0)}$ . The transverse momenta of muon and pion is generated according to the distributions in Fig.6.3. The idea is to apply a momentum dependent selection of the tracks, as expected in data, and determine the parameter  $A_D$  in Eq. 6.15 from a fit to the pseudo-experiment using the model in Eq. 6.15. The subsequent step is to apply the kinematical weights to be applied to the data and measure  $A_D$ . A weight is associated to an event according to the  $p_T$  of the pion generated (see Eq. 6.14 and Fig. 6.5 (*right*)) The value of  $A_D$  measured including the weights should show the reduction of the  $\mu - \pi$  tracking asymmetry achieved with the weights. The fit software used for these checks is fully described in Chapter 8.

For this validation it is sufficient to generate signal decays, and the number of events generated in the pseudo-experiment has to be chosen according to the desired uncertainty on the charge asymmetry determination. The tracking asymmetry  $A_{\text{tracking}}(\mu\pi)$  is expected to be at the percent level, therefore a precision on the charge asymmetry at the per-mille level is acceptable. Given a sample of  $N$  events, the uncertainty on an asymmetry is given by  $1/\sqrt{N}$ , therefore a million of events is the statistics chosen for the pseudo experiment. Previous studies [108] performed for the first  $a_{\text{sl}}^s$  measurement, show the tracking efficiency, Fig 6.9(*left*) and the order of magnitude of the asymmetry, Fig. 6.9(*right*) that could be expected. In this case the investigated tracks are pion tracks, but since no contribution to the charge asymmetry is expected from the pion interaction, the result can be used also for muons. The difference, as previously discussed, is expected in the uncertainties.

These momentum dependent tracking efficiencies are used to model a realistic tracking efficiency in the pseudo-experiments. In Eq.6.15 the charge asymmetry is described by different parameters:  $A_D$ ,  $a_{\text{sl}}^d$  and  $A_P$ . The production asymmetry is not relevant for this purpose, therefore a realistic value of 1% is assumed. The  $a_{\text{sl}}^d$  parameter is assumed to be zero, in order to evaluate whether the weights reduce the tracking asymmetry. Different values are assumed for  $a_{\text{sl}}^d$  in subsequent tests. The detection asymmetry introduced by the

momentum dependent efficiency in Fig.6.9 is determined by the fit to be  $A_D = (1.5 \pm 0.2)\%$ . After applying the kinematical weights to the events of the pseudo-experiment, the fitted value for the detection asymmetry is compatible with zero. When generating and fitting a set of 100 pseudo-experiments, the average measured value of  $A_D$  is  $-0.03\%$ , with an uncertainty of  $0.24\%$ . This confirms that the kinematical weighting of muon and pion reduces the asymmetry.

The goal of the second test is to show that the value of  $a_{\text{sl}}^d$  is not modified by the weighting procedure. The time model assumed is Eq. 6.15 as previously used, but a non-zero value is assumed for the  $a_{\text{sl}}^d$  parameter:  $a_{\text{sl}}^d = 0.5\%$ . Both  $A_D$  and  $a_{\text{sl}}^d$  parameters are determined by means of fits, to the not-weighted and weighted pseudo-experiment. For the not-weighted fit the resulting parameters are  $A_D = (1.9 \pm 0.2)\%$  and  $a_{\text{sl}}^d = (0.9 \pm 1.0)\%$ . The results obtained with the weighted fit,  $A_D = (-0.2 \pm 0.5)\%$  and  $a_{\text{sl}}^d = (0.6 \pm 1.4)\%$ . The two parameters are very correlated, and higher statistics checks would be needed, nevertheless this result does not point to significant biases.

Other checks are performed to confirm the validity of the method. The stability of the  $a_{\text{sl}}^d$  result is verified using different weightings. In this case the complete  $a_{\text{sl}}^d$  fit, described in Chapter 8 is used. Different weighting schemes to match the kinematic of the pion to the kinematic of the muon have been applied on the data samples and the nominal analysis is repeated. The stability of the  $a_{\text{sl}}^d$  result ensures that the simple uni-dimensional  $p_T$  weighting ensures already that the tracking asymmetry can be neglected. As last validity test, the residual  $\mu - \pi$  tracking asymmetry is measured with a different method, as explained in the next section. A residual tracking asymmetry could be due to the not perfect agreement of muon and pion kinematics after the weighting procedure, but it is expected to be negligible.

## Residual asymmetry

The residual asymmetry is estimated using  $J/\psi \rightarrow \mu^+\mu^-$  decays, with a procedure based on the same method that was used to measure the tracking efficiencies [109].  $J/\psi \rightarrow \mu^+\mu^-$  decays are used, since they are clean, abundant and cover the momentum spectrum typical for most of the analysis at LHCb. A so called *tag-and-probe* method is used to determine the track reconstruction efficiency. One of the tracks originating from the  $J/\psi$ , the *tag* track is fully reconstructed and is identified as a muon. The probe track is only partially reconstructed, not using information from at least one subdetector which is probed. The track reconstruction efficiency is determined by checking for the existence of a fully reconstructed track corresponding to the probe track. This allows to determine the efficiency of the subdetector that is not used in the reconstruction of the probe track. For this analysis, a weighting procedure is applied to a  $J/\psi$  sample in order to modify the  $\mu^+\mu^-$  to reproduce the kinematic distribution of the  $\mu^+\pi^-$  pair of the signal decays. With this approach, it is possible to directly extract  $A_{\text{tracking}}(\mu\pi)$  following Eq. 6.13. Using this method, the residual asymmetry after applying the  $\mu - \pi$  weighting to the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$

channel is estimated to be

$$A_{\text{tracking}}(\mu\pi) = (0.00 \pm 0.02)\%.$$

The resulting uncertainty is accounted for when evaluating the systematic uncertainty on  $a_{\text{sl}}^d$ . The same uncertainty is assigned to the  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$ . This method is currently being finalized and will be used for the future measurements of  $a_{\text{sl}}^s$  and  $a_{\text{sl}}^d$ , since it allows to obtain a precise and reliable estimation of the tracking asymmetries, avoiding the event weighting approach used here.

## 6.5.2 Muon mis-identification and trigger asymmetry

This section summarizes the method used to extract the muon mis-identification and trigger asymmetry. This approach was used for the  $a_{\text{sl}}^s$  measurement at LHCb [10]. Further improvements developed for the  $a_{\text{sl}}^d$  measurements are described in Ref. [110], [107]. The muon mis-identification and trigger asymmetry, defined as

$$A(\mu) = \frac{\epsilon(\mu^+) - \epsilon(\mu^-)}{\epsilon(\mu^+) + \epsilon(\mu^-)}, \quad (6.17)$$

and is measured with  $J/\psi \rightarrow \mu^+\mu^-$  decays (see Sec. 5.4) using a so called *tag-and-probe* method. This method relies on the selection of an unbiased sample of physics objects, the *probe* objects, which are used to calculate the efficiencies and resolutions. The data sample is selected by means of an independent *tag* selection. For the muon mis-identification and trigger asymmetry in order to select the sample, the probe tracks in the events are required to fulfill the global TIS (“triggered independently on signal” see Sec. 5.1.2) trigger requirement on the three trigger levels (see Sec. 3.2.3). A probe track is defined to be efficient (“passed”) when it satisfies the signal selection criteria (trigger: TOS on L0MUON and HLT1TRACKMUON, muon identification requirements as  $DLL_{\mu\pi}$ ). The method is not sensitive to asymmetries possibly generated by the HLT2 topological lines. No significant asymmetry is expected to be generated by these lines, as shown in [10]. Four subsamples can be identified:

$$\begin{array}{ll} (A)\mu^- \text{ tagged, } \mu^+ \text{ probe passed} & (C)\mu^+ \text{ tagged, } \mu^- \text{ probe passed} \\ (B)\mu^- \text{ tagged, } \mu^+ \text{ probe NOT passed} & (D)\mu^+ \text{ tagged, } \mu^- \text{ probe NOT passed} \end{array} \quad (6.18)$$

The  $J/\psi \rightarrow \mu^+\mu^-$  decay yields are determined by means of fits to the  $\mu^+\mu^-$  invariant mass. From the fits to the yields of these four subsamples, the muon and anti-muon efficiencies can be calculated. Where  $A$  is the  $J/\psi$  yield of the subsample (A) in Eq 6.18,  $B$  the  $J/\psi$  yield of the subsample (B) and so on, the efficiencies are calculated as  $\epsilon(\mu^+) = \frac{A}{A+B}$  and  $\epsilon(\mu^-) = \frac{C}{C+D}$ . The datasets are weighted with the set of weights determined on the signal samples in order to suppress the  $\mu\pi$  tracking asymmetry (see Sec. 6.10). The statistical

correlation between the  $A$  and  $C$  samples is accounted for. A number of variations of the method are used in order to assign the systematic uncertainty to the  $A_D(\mu)$  measured value. These include variation in the  $\mu^+\mu^-$  invariant mass fit model and in binning scheme used. The sum in quadrature of the deviations of the the results of the alternative fits with respect to the nominal fits are taken as systematic uncertainty. It results to be 0.06%. A number of cross-checks have been performed:

- *Different  $J/\psi \rightarrow \mu^+\mu^-$  decay samples*  
Both  $J/\psi \rightarrow \mu^+\mu^-$  decays with the  $J/\psi$  produced directly in the  $pp$  collision and produced from a  $B$  decay are used (see Sec. 5.4). They give compatible  $A(\mu)$  results.
- *Data taking periods and magnet polarity*  
Values of  $A(\mu)$  are calculated separately in subsets of the data samples. The datasets are split according to the data taking period and magnet polarity. The values of the muon asymmetry have been compared and no particular deviation of one or more sub-dataset with respect to the overall trend is found. On each subsample the full  $a_{\text{sl}}^d$  analysis is performed as well (see Appendix. C).
- *$B$  decay time dependence*  
 $A(\mu)$  is calculated in  $B$  decay time bins, in order to understand if it features a dependence in function of the  $B$  decay time. No  $B$  decay time dependence is observed.

The resulting muon asymmetry used for the nominal  $a_{\text{sl}}^d$  fits are shown in Tab. 6.1, together with the other detection asymmetries accounted for.

### 6.5.3 Pion mis-identification asymmetry

The PID selection requirement  $DLL_K < 10.0(4.0)$  for the  $D^\pm(D^{*\pm})$  on the fastest pion (pion from the  $\bar{D}^0$ ) could introduce an additional detection asymmetry, that should be taken into account. The efficiencies of this PID cut are measured using  $D^0 \rightarrow K\pi$  decays from  $D^{*\pm}$  mesons, selected independently of any PID selection criteria. Particle Identification performances depend on the kinematics of the decays, for this reason the  $D^{*\pm}$  tagged  $D^0 \rightarrow K\pi$  decays are weighted to match the kinematic distributions of the signal decays. The  $\mu - \pi$  weights (Sec. 6.5.1) are already applied to the signal samples. The values obtained for the  $A_{\text{PID}}(\pi) = \frac{\epsilon(\pi^-) - \epsilon(\pi^+)}{\epsilon(\pi^-) + \epsilon(\pi^+)}$  are reported Tab. 6.1, with the other detection asymmetries.

## 6.6 $K - \pi$ asymmetry

This section describes the method used to measure the detection asymmetry of the kaon and lower  $p_T$  pion (in the case of  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$  decays, the pion originating from the  $D^*$ ). A more detailed description of the method can be found in Ref. [9], [106], [110].

The effects of all possible sources of asymmetry are accounted for with this method. These

sources include the nuclear interactions, tracking and PID performances. The detection asymmetry of the  $K\pi$  pair is defined as

$$A(K\pi) = \frac{\epsilon(K^+\pi^-) - \epsilon(K^-\pi^+)}{\epsilon(K^+\pi^-) + \epsilon(K^-\pi^+)}. \quad (6.19)$$

The main idea is to measure the charge asymmetry between  $K^+\pi^-$  and  $K^-\pi^+$  reconstructed pairs from charm decays where the only source of charge asymmetry can be found in the detection and reconstruction effects. Among the sources of possible charge asymmetry, it is easy to exclude  $CP$  violation. This is done by choosing a control sample of Cabibbo favored decays, where no  $CP$  violation is expected. As explained in Chapter. 4, the production of hadrons at LHCb is affected by a production asymmetry of particles with respect to antiparticles. This affects  $B$  and  $D$  mesons. For this reason it is necessary to use a second control sample, in order to cancel the  $D$  production asymmetry effect. The first control sample includes  $D^- \rightarrow K^+\pi^-\pi^-$  decays, while the second is constituted by  $D^- \rightarrow K^0\pi^-$  decays (see Sec. 5.4 for further selection details). The  $K - \pi$  asymmetry is obtained as

$$A(K\pi) = A(D^- \rightarrow K^+\pi^-\pi^-) - A(D^- \rightarrow K^0\pi^-) - A(K^0). \quad (6.20)$$

The first two asymmetries on the right hand side are measured in the  $a_{\text{sl}}^d$  analysis and the correction for the neutral kaon interaction and  $CP$  asymmetry,  $A(K^0)$ , is calculated in [9] and is taken into account. The two asymmetries are measured from the event yields, which are determined from the invariant mass distribution of the  $K^+\pi^-$  (and  $K^-\pi^+$ ) pairs and  $K^0\pi^-$  (and  $\bar{K}^0\pi^-$ ) pairs. Before performing these fits, a reweighting procedure is necessary.

First of all, it is necessary that the kinematics and the PID requirements of the kaon and the slowest pion in the  $D^- \rightarrow K^+\pi^-\pi^-$  decays are the same of the kaon and slowest pion in the nominal samples of signal decays (after that the  $\mu - \pi$  weights (Sec. 6.5.1) are applied). Moreover, it is necessary that the kinematics of the  $D$  mesons in the two control samples are equalized, in order to be able to cancel the  $D$  meson production asymmetry. Finally, the fastest pion (here also defined as "trigger pion", since it is used in the trigger decision) in the two control samples needs to have the same PID, trigger requirements (TOS on HLTTRACKALL0) and kinematics. Figure 6.10 helps to visualize the method.

The weighting procedure involves two steps:

- *$K\pi$  weighting*

This uses a 3D distribution in  $p_T$  and  $\eta$  of the slow pion and  $p_T$  of the kaon and a 1D distribution of the  $\phi$  of the kaon. These two distributions are determined for the signal sample and  $D^- \rightarrow K^+\pi^-\pi^-$  sample. Based on the comparison between the two, weights are assigned. The 3D distribution and the  $\phi$  distribution of the kaon are not strongly correlated, so the two weightings can be applied sequentially.

- *$D\pi$  weighting*

This uses a 3D distribution in  $p_T$  and  $\eta$  of the  $D$  meson and  $p_T$  of the trigger pion. Weights are assigned based on the comparison of this distribution between the  $D^- \rightarrow K^+\pi^-\pi^-$  sample and  $D^- \rightarrow K^0\pi^-$  decays.

In order to avoid large weights that lead to a severe degradation of the effective statistics of the samples, a weight truncation is applied at both stages. This compromises the degree of possible agreement between the kinematic distribution. The effects related to the weighting scheme chosen are accounted for in the evaluation of the systematic uncertainties.

A number of variations of the procedure are used to access the systematic uncertainty on the  $A(K\pi)$  measurement. These include variations in both the  $K\pi$  weighting and  $D\pi$  weighting and in the fit model used in the fit to extract the yields used to calculate the asymmetry. The overall systematical uncertainty assigned to  $A(K\pi)$  is 0.07% in the case of  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  decays and 0.1% for the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays.

Several cross-checks are also performed on the  $A(K\pi)$  measurement.

- *Stability of  $A(K\pi)$  values measured using orthogonal subsamples*  
The dataset is split according to number of primary vertices, range of impact parameter values, trigger decision. No unexpected variation is observed.
- *Data taking periods and magnet polarity*  
Values of  $A(K\pi)$  are calculated separately in subsets of the data samples. The datasets are split according to the data taking period and magnet polarity. The values of the muon asymmetry have been compared and no particular deviation of one or more sub-dataset with respect to the overall trend is found. On each subsample the full  $a_{\text{sl}}^d$  analysis is performed as well (see Appendix. C).
- *$B$  decay time dependence*  $A(K\pi)$  is calculated in  $B$  decay time bins, in order to understand if it features a dependence in function of the  $B$  decay time. No  $B$  decay time dependence is observed.

The results of the  $A(K\pi)$  measurement used as input for the  $a_{\text{sl}}^d$  measurement are reported in Tab 6.1.

## 6.7 Input values for $a_{\text{sl}}^d$ measurement

Tab. 6.1 summarizes the detection asymmetry values used as input for the  $a_{\text{sl}}^d$  measurement. Uncertainties statistical and systematics are added in quadrature and the result is reported. The  $K - \pi$  pair detection asymmetry,  $A(K\pi)$  is measured to be at the order of 1%, as expected given the different cross sections between the positively and negatively charged kaons. The uncertainty is at the per-mille level, which is the uncertainty aimed for the  $a_{\text{sl}}^d$  measurement. This is the leading contribution to the detection asymmetry correction  $A_{\text{D}}$  used in the  $a_{\text{sl}}^d$  measurement. The muon mis-identification and trigger asymmetries,  $A(\mu)$  are measured to be at the few per-mille level, with uncertainty below the per-mille. The measured pion PID asymmetry,  $A_{\text{PID}}(\pi)$  is in almost in all the cases below the per-mille. The  $\mu\pi$  charge asymmetry due to tracking performances is not reported, since the weighting procedure used for this analysis reduces this asymmetry to negligible level, therefore a correction of the measured charge-asymmetry between the final state particle charges is not needed. Different approaches are currently under investigation to provide an accurate

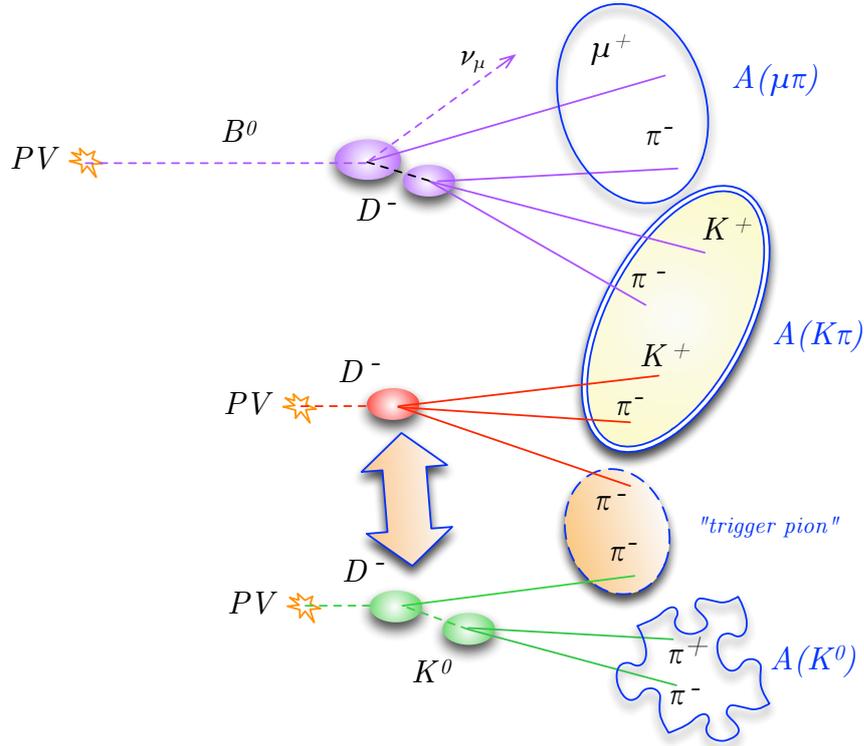


Figure 6.10: Visualization of the strategy used to estimate  $A(K\pi)$ . The detection asymmetry of the  $K\pi$  pair in the final state of the signal decay (highlighted in yellow) is obtained by measuring the charge asymmetry if the  $K\pi$  pair in the Cabibbo favored  $D^- \rightarrow K^+\pi^-\pi^-$  decays. It is then necessary to subtract from this measurement the effects of the  $D$  production asymmetry, the additional “trigger pion” (which has triggered the LHT1TRACKALL0 line) and the neutral kaon detection and  $CP$  asymmetry. The first two corrections are performed using the second control sample  $D^- \rightarrow K^0\pi^-\pi^-$ , while the latter correction is taken from [9]. A complicated weighting procedure is necessary, and it is performed in two steps:  $K\pi$  weighting (the particles used to determine the weights are highlighted in yellow) and  $D\pi$  weighting (the particles whose used to determine the weights are highlighted in orange) described in Sec. 6.6.

measurement of the  $\mu\pi$  tracking asymmetry to be used in precision measurements. . The uncertainty, statistical and systematic, on the detection asymmetry is the source of the leading systematic uncertainty on the  $a_{\text{sl}}^d$  measurement. For future  $a_{\text{sl}}^d$  measurements, the uncertainty on the detection asymmetry determination is going to decrease, because of the larger statistics of the control sample and because of the improvements in the techniques used. The development of a reliable method to evaluate the  $\mu\pi$  tracking asymmetry will allow to avoid the weighting procedure that is currently reducing the effective size of the data sample by about 40%, leading to a further reduction of the statistical uncertainty on  $a_{\text{sl}}^d$ .

Table 6.1: Detection asymmetries[%]. Values used as input in the time dependent fits for the four sub samples of  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays. The uncertainties include both the statistical and systematic components added in quadrature.

Detection Asymmetry	$B^0 \rightarrow D^- \mu^+ \nu_\mu X$	$B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$
<b>2011 Magnet Down:</b>		
$A(\mu)$	$-0.2210 \pm 0.0770$	$-0.2120 \pm 0.0750$
$A(K\pi)$	$1.0558 \pm 0.2403$	$-0.2206 \pm 0.1891$
$A_{\text{PID}}(\pi)$	$0.0440 \pm 0.0020$	$0.0880 \pm 0.0040$
<b>total</b>	$0.8788 \pm 0.2529$	$-0.3446 \pm 0.2041$
<b>2011 Magnet Up:</b>		
$A(\mu)$	$0.3750 \pm 0.0900$	$0.3140 \pm 0.0890$
$A(K\pi)$	$1.5072 \pm 0.2766$	$1.9062 \pm 0.2242$
$A_{\text{PID}}(\pi)$	$-0.0490 \pm 0.0030$	$-0.3230 \pm 0.0050$
<b>total</b>	$1.8332 \pm 0.2913$	$1.8972 \pm 0.2418$
<b>2012 Magnet Down:</b>		
$A(\mu)$	$-0.0300 \pm 0.0460$	$-0.0110 \pm 0.0470$
$A(K\pi)$	$1.0027 \pm 0.1527$	$0.5167 \pm 0.2044$
$A_{\text{PID}}(\pi)$	$0.0020 \pm 0.0020$	$-0.0280 \pm 0.0030$
<b>total</b>	$0.9747 \pm 0.1603$	$0.4777 \pm 0.2104$
<b>2012 Magnet Up:</b>		
$A(\mu)$	$0.0370 \pm 0.0450$	$-0.0630 \pm 0.0460$
$A(K\pi)$	$1.1335 \pm 0.1503$	$1.3250 \pm 0.2721$
$A_{\text{PID}}(\pi)$	$-0.0290 \pm 0.0020$	$-0.0480 \pm 0.0030$
<b>total</b>	$1.1415 \pm 0.1577$	$1.2140 \pm 0.2764$

# Chapter 7

## Decay time description

This chapter describes the method used to reconstruct the  $B$  meson decay time and properly account for the time resolution. Although the strategy reported and the results are valuable for any time-dependent analysis exploiting data samples of decays with a partially reconstructed final state, the attention is devoted to the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decay samples used for the  $a_{\text{sl}}^d$  measurement.

### 7.1 Partially reconstructed decays

A reliable reconstruction of the  $B$  meson decay time is the first ingredient for the measurement of any time-dependent observable. The decay time of a  $B$  meson is obtained as

$$t = \frac{L(B)M(B)}{|\vec{p}(B)|}, \quad (7.1)$$

where  $L$  indicates the flight distance of the  $B$  meson,  $M$  denotes the reconstructed mass of the  $B$  meson, and the momentum of the  $B$  meson,  $\vec{p}$ , is reconstructed summing the momenta of the particles in the final state. When considering partially reconstructed decays, one or more particles of the final state are not reconstructed, therefore the sum of the momenta of the  $B$  daughters does not result in the momentum of the  $B$  meson. In this analysis semileptonic samples are considered. In the case of signal decays, the particle of the final state which is not reconstructed is the neutrino.

Fig. 7.1 helps to visualize the problem, taking as example a  $B^0 \rightarrow D^- \mu^+ \nu_\mu$  decay. The first evident effect of the partial reconstruction is that the sum of the momenta of the final state particles result in a momentum which is systematically lower than the momentum of the original  $B$  meson. A second remarkable effect is the uncertainty that is affecting the determination of the  $B$  decay time. These two topics are examined in the following two sections.

For the development and validation of the methods summarized here, simulated events are used. The advantage of MC samples is that for every reconstructed particle, also the “truth” information is available. For example the reconstructed momentum of a particle is in general different from the value in generation of the momentum of the same particle,



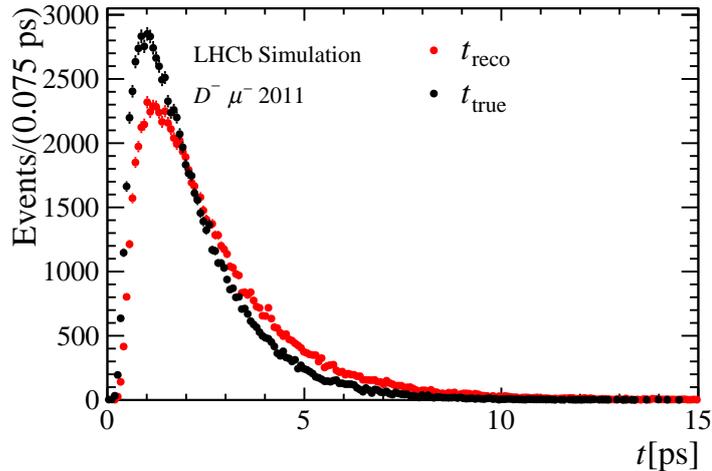


Figure 7.2: Reconstructed and true  $B$  decay time.  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  simulated events with 2011 conditions. The rising edge at low decay times is due to acceptance effects, which are described in Sec. 7.2.2.

### Data driven techniques

It is convenient to write all the momenta playing a role in Fig. 7.1 in terms of components parallel and perpendicular to the flight direction  $\hat{F}$  of the  $B$  meson.

$$p_{\parallel} = \vec{p} \cdot \hat{F}, \quad p_{\perp} = |\vec{p} - p_{\parallel} \hat{F}| \quad (7.2)$$

The unknown quantity of the system is the momentum of the neutrino (or equivalently the momentum of the  $B$ ). As shown in Fig. 7.1 For momentum conservation the component of the neutrino momentum which is perpendicular to the flight direction,  $p_{\perp}(\nu)$  has the same magnitude of the component of the resulting momentum of the reconstructed final state perpendicular to the flight direction of the  $B$ , here indicated as  $p_{\perp}(D^- \mu^+)$ .

The knowledge of  $p_{\perp}(\nu)$  helps already for an evaluation of the  $B$  momentum. This is exploited by the empirical relation used by the E653 experiment [111] to calculate the  $B$  momentum:

$$p_{corr}(B^0) = \frac{m(B^0)E(D^- \mu^+)}{\sqrt{p_{\perp}(D^- \mu^+)^2 + m(D^- \mu^+)^2}} \quad (7.3)$$

The component of the neutrino momentum parallel to the flight direction is still not accounted for. The estimate of the  $B$  momentum in Eq. 7.3 is surely closer to the true  $B$  momentum, if compared to the reconstructed  $B$  momentum, but a further step towards the reconstruction of the full final state is possible.

The method known as *neutrino reconstruction* or *neutrino closure* calculates  $p_{\parallel}(\nu)$  given the mass and energy conservation of the system:

$$\begin{cases} p_{\perp}(D^{-}\mu^{+}) = -p_{\perp}(\nu) \\ p(B^0) = p_{\parallel}(D^{-}\mu^{+}) + p_{\parallel}(\nu) \\ E(B^0) = E(D^{-}\mu^{+}) + E(\nu) \end{cases}$$

Imposing the  $B^0$  mass constraint  $m(B^0)^2 = E^2(B^0) - p^2(B^0)$ , it is possible to obtain the following second order equation for  $p_{\parallel}(\nu)$ :

$$ap_{\parallel}(\nu)^2 + bp_{\parallel}(\nu) + c = 0 \quad (7.4)$$

with

$$\begin{aligned} a &= 4[m(D^{-}\mu^{+})^2 + p_{\perp}(D^{-}\mu^{+})^2] \\ b &= 4p_{\parallel}(D^{-}\mu^{+})[m(D^{-}\mu^{+})^2 - m(B^0)^2 + 2p_{\perp}(D^{-}\mu^{+})^2] \\ c &= 4p_{\perp}(D^{-}\mu^{+})^2[p_{\parallel}(D^{-}\mu^{+})^2 + m(B^0)^2] - [m(B^0)^2 - m(D^{-}\mu^{+})^2]^2 \end{aligned}$$

$p_{\parallel}(\nu)$  can be determined up to a two-fold ambiguity

$$p_{\parallel}(\nu) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (7.5)$$

This method presents mainly two disadvantages. The first is that when using semileptonic decays as  $B^0 \rightarrow D^{-}\mu^{+}\nu_{\mu}$ , there is no handle to solve the two-fold ambiguity. This is not the case for all the possible decays. For example if considering  $B^+$  semileptonic decays, with the  $B^+$  originating from  $B_{s2}^* \rightarrow B^+K^-$  decays, the mass of the  $B_{s2}^*$  parent of the  $B^+$  would provide an additional constraint useful to choose the correct solution between the two in Eq. 7.5. But for the data samples used for the  $a_{\text{sl}}^d$  measurement there are no such additional constraints. A second problem of the neutrino reconstruction method is that, due to resolution effects, for a large fraction of the sample the discriminant  $b^2 - 4ac$  in Eq. 7.5 is negative, leading to the impossibility of finding a solution for  $p_{\parallel}(\nu)$ .

Effects of the approach used by the E653 collaboration and neutrino reconstruction method on the  $B$  decay time distribution and time resolution are reported in Fig. 7.3, together with the performances of the  $k$ -factor method, explained in the next section. For the neutrino reconstruction method, in case negative discriminant  $b^2 - 4ac$ , only the component of the neutrino momentum perpendicular to the flight direction is accounted for in the  $B$  momentum correction. In case of positive discriminant, a random solution between the two is chosen.

### **$k$ -factor method**

The so-called *k-factor method* takes advantage of the ratio between the magnitude of reconstructed and true  $B$  decay momentum in simulated events. This ratio is called  $k$ -factor

$$k = \frac{p_{\text{reco}}}{p_{\text{true}}} \quad (7.6)$$

This  $k$ -factor can be calculated event by event in simulation. The distribution of the  $k$ -factors allows for an average correction of the  $B$  reconstructed decay time and a proper

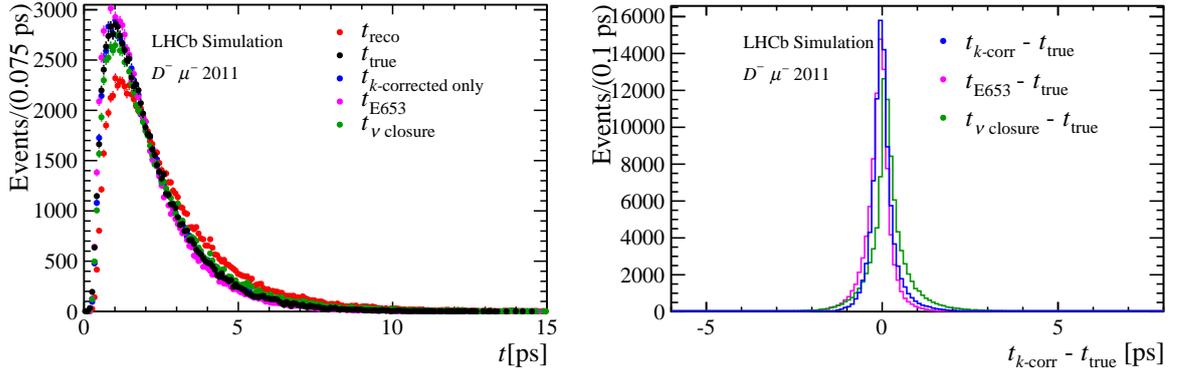


Figure 7.3: Different  $B$  momentum reconstruction techniques.  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  simulated events with 2011 conditions.

modeling of the resolution. A distinction should be made within the  $k$ -factor method. It is possible to use the  $k$ -factors distribution to model the decay time resolution of semileptonic decays, and it is a well known strategy, see Ref. [112] for example. In addition, some LHCb analysis as Ref. [113], use an average  $k$ -factor,  $\langle k \rangle$  to correct the  $B$  momentum and therefore the  $B$  decay time. This average  $k$ -factor correction is done by assigning to each event a  $k$  factor according to the reconstructed  $B$  mass ( $m(D^- \mu^+)$ ) of the event. This correction reproduces the distribution of the  $B$  decay time, but introduces also a decay-time dependent time resolution, that is problematic to model. In this analysis the average  $k$ -factor correction is used and applied at early stage on the data. In order to account properly for the decay time resolution, the  $k$ -factor distribution is used. This second procedure allows to properly account for the decay time resolution and correctly extract the physics parameters of interest. The average  $k$ -factor,  $\langle k \rangle$ , correction basically allows only to adjust the time scale of the time-dependent data distribution to the expected  $B^0$  lifetime scale.

In order to calculate the event by event  $B$  decay time, the  $k$ -factors distribution as function of the reconstructed  $B$  mass ( $m(D^- \mu^+)$ ) is used.

The  $k$ -factors distribution in different reconstructed  $B$  mass ( $m(D^- \mu^+)$ ) ranges is shown in Fig. 7.4.

Fig. 7.5 shows  $\langle k \rangle$ , and the RMS of the  $k$ -factors distribution, as function of the reconstructed  $B$  mass. As expected, the higher the visible  $B$  mass, the closer to 1 are the resulting values of the  $k$ -factors. The higher the visible  $B$  mass, the smaller the magnitude of the missing momentum is allowed to be, leading to narrower  $k$ -factors distributions.

The  $B$  meson decay time is calculated as

$$t = \frac{L(B^0)M(B^0)_{PDG}}{|\vec{p}(B^0)|} \langle k \rangle (M(B^0)) = t_{k\text{-corr}}. \quad (7.7)$$

The known value of the  $B^0$  mass is used,  $M(B^0)_{PDG}$ , in order to reduce the uncertainty

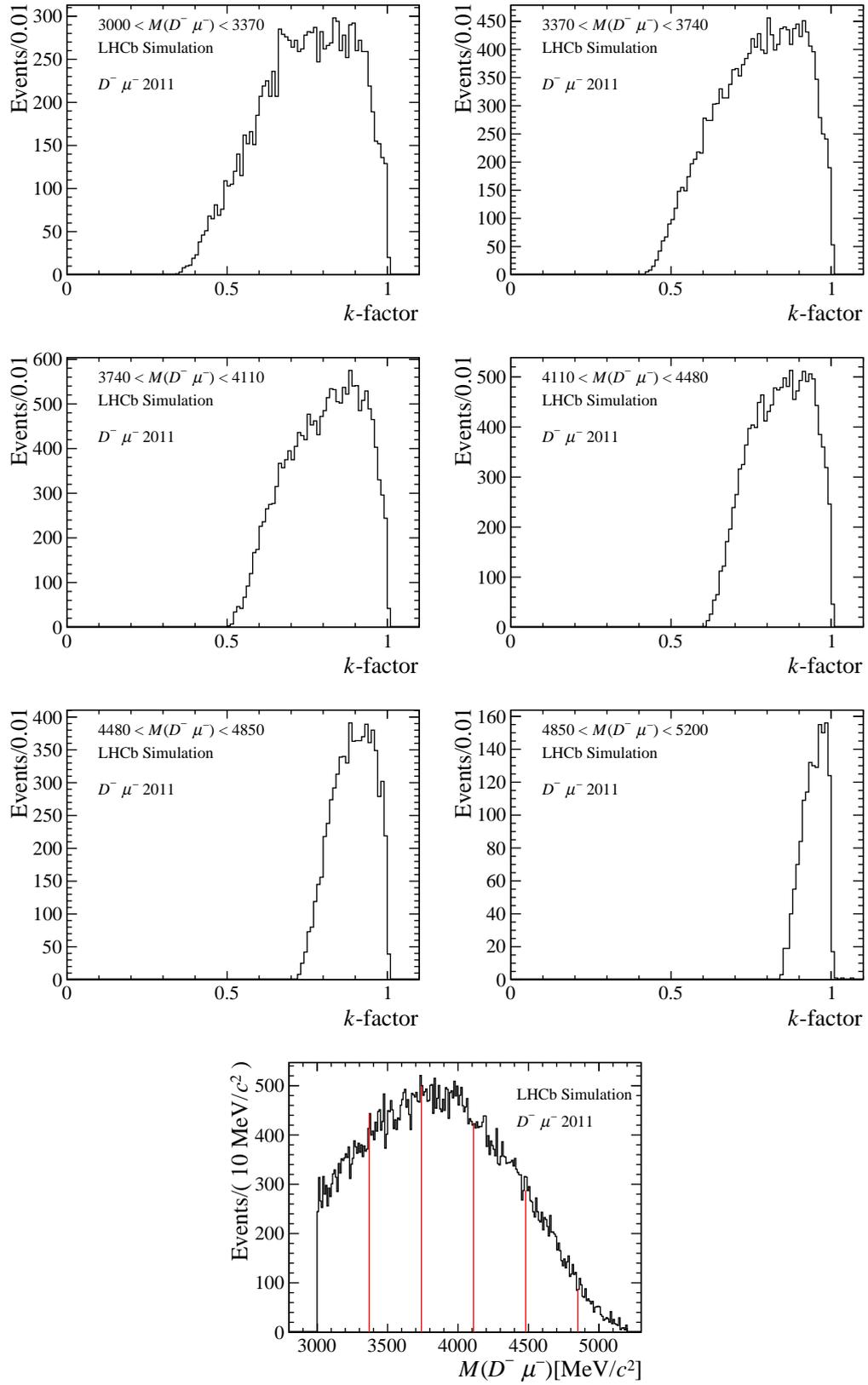


Figure 7.4:  $k$ -factor distribution in different mass ranges, correspondently indicated on the reconstructed  $B^0$  mass distribution.

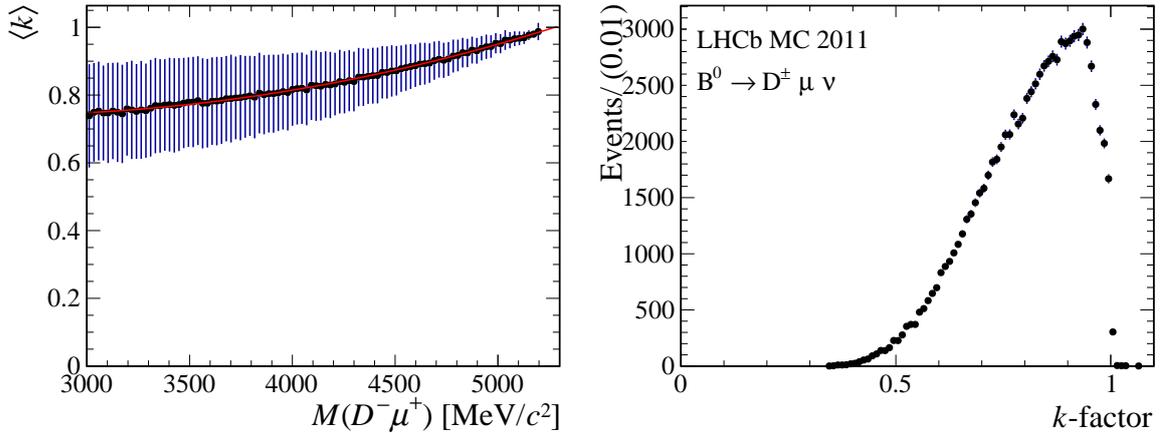


Figure 7.5: *Left*  $\langle k \rangle$  as function of the reconstructed  $B$  mass ( $m(D^- \mu^+)$ ). *Right*  $k$ -factor distribution of all the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  simulated events with 2011 conditions.

due to the mass In this Chapter, the  $B$  decay time carries the indication of the method of correction applied e.g.  $t_{k\text{-corr}}$ , in the following chapters the  $B$  decay time obtained after the missing neutrino correction is indicated as  $t$ . As indicated in Eq. 7.7 the mass obtained to calculate the  $B^0$  decay time is the known value of the  $B^0$  mass [17]. The value of  $\langle k \rangle(M(B^0))$  is calculated for each event given the  $m(D^- \mu^+)$  mass, and the parametrization of  $\langle k \rangle$  in function of  $m(D^- \mu^+)$  shown in Fig. 7.5. The dependence of  $\langle k \rangle$  in function of  $m(D^- \mu^+)$  is described by an empirical second order polynomial.

$$\langle k \rangle = p_0 + p_1 m(D^- \mu^+) + p_2 m^2(D^- \mu^+) \quad (7.8)$$

Applying the  $k$ -factor correction to the reconstructed  $B$  decay time distribution, reproduces the true  $B$  decay time distribution, as shown in Fig. 7.6 . A proper understanding of the effects of this correction on the  $B$  decay time distribution is essential to be able to use this method.

## 7.1.2 Decay time resolution

The  $k$ -factor corrected decay time  $t_{k\text{-corr}}$  can be written as

$$t_{k\text{-corr}} = \frac{L(B^0)M(B^0)_{PDG}}{|p_{k\text{-corr}}^{\vec{}}(B^0)|}, \quad \text{with } |p_{k\text{-corr}}^{\vec{}}(B^0)| = |\vec{p}(B^0)|/\langle k \rangle(M(B^0)) \quad (7.9)$$

For a time dependent analysis, it is crucial to understand and properly model the time resolution. Typically there are several effects playing a role in the decay time resolution for decays with fully reconstructed final states. These are the tracking detectors spatial resolution, the amount of material that the particle traverses, given that a large amount of material increases the effects of multiple scattering, and the boost of the particles

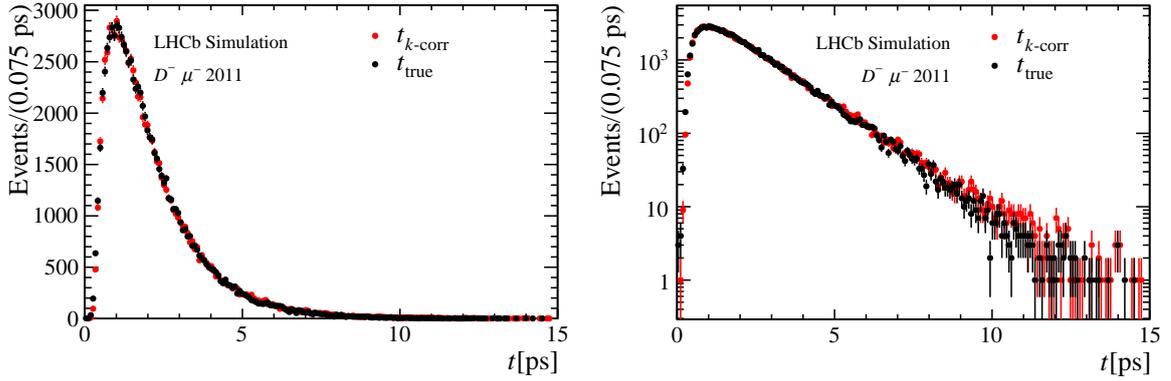


Figure 7.6:  $k$ -corrected and true  $B$  decay time. The disagreement between the distributions at high  $B$  decay times is a resolution effect, that is properly modeled in the  $a_{\text{sl}}^d$  analysis, as explained in Sec. 7.2.1. The  $B$  decay time distribution after accounting also for resolution effects is shown in Fig. 7.15.  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  simulated events with 2011 conditions.

in the forward region. The decay time resolution is typically estimated from simulated event samples. The distribution of the difference between the reconstructed decay time and the true decay time provides a model for the decay time resolution. In the case of partially reconstructed decays, the resolution of the decay time needs in addition to account for the effect of the missing momenta. In the present case the  $k$ -factor correction is applied to the partially recorded  $B$  decay time obtained from the sum of the momenta of the reconstructed decay products. Therefore it is necessary to understand the  $k$ -factor corrected  $B$  decay time resolution. Fig. 7.7 shows the distribution of the difference between the  $k$ -factor corrected  $B$  decay time and true  $B$  decay time obtained from one of the MC signal samples used in the  $a_{\text{sl}}^d$  analysis. The width of this distribution is about an order of magnitude larger than the resolution observed for fully reconstructed decay modes, and the shape is rather asymmetric with respect to zero. This second feature is due to the time dependence of the  $k$ -factor corrected  $B$  decay time resolution. Fig. 7.8 shows the  $t_{k\text{-corr}} - t_{\text{true}}$  distribution in different bins of  $k$ -factor corrected  $B$  decay time. The distributions can be described with a gaussian shape with larger mean and larger  $\sigma$  for increasing values of  $t_{k\text{-corr}}$ . For a more complete overview of the behavior of the time resolution, Fig. 7.9 reports the mean and width of the  $t_{k\text{-corr}} - t_{\text{true}}$  distribution as function of  $t_{k\text{-corr}}$  and  $t_{\text{true}}$ .

The feature of this decay time resolution can be explained as follows. Using the standard error propagation the uncertainty on  $t_{k\text{-corr}}$  is given by

$$\sigma_{t_{k\text{-corr}}} = \sqrt{\left(\frac{M(B^0)_{PDG}}{p_{k\text{-corr}}}\right)^2 \sigma_L^2 + t_{k\text{-corr}}^2 \left(\frac{1}{p_{k\text{-corr}}}\right)^2 \sigma_{p_{k\text{-corr}}}^2} \quad (7.10)$$

The first and second terms in this expression indicate the uncertainty on the flight distance measurement and the uncertainty on the momentum measurement. In fully

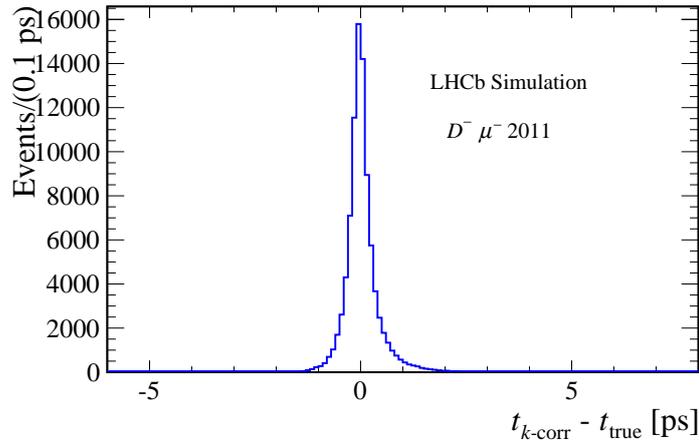


Figure 7.7:  $k$ -corrected  $B$  decay time resolution.  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  simulated events with 2011 conditions.

reconstructed decay modes the first term dominates, leading to a decay time resolution independent on the decay time itself, and mainly being the result of the effects enumerated at the beginning of this section. In the case of partially reconstructed decays instead, the second terms is clearly the dominant. This leads to a time-dependent decay time resolution as shown in Fig. 7.9. The factor  $t_{k\text{-corr}}^2$  in Eq. 7.10 explains the observed quadratic dependence of the width of the resolution in function of the decay time in Fig.7.9. The increasing value of the mean of the time resolution is a result of “bin-to-bin migration”. It is an effect arising from the increasing width of the resolution, that allows events from the higher decay time bins to migrate in the time bin of interest, and an effect of the average  $k$ -factor correction that can be visualized in the tail distribution of Fig. 7.6. This dependence is introduced by the  $k$ -factor correction.

At this stage, the issue is to correctly model the features of the decay time resolution. The approach used in this analysis is to separate the contributions of the uncertainty on the flight distance measurement from the uncertainty on the momentum measurement. These effects in principle are independent, any dependence that might be observed in the data sample is introduced by the selection criteria applied to the data sample. Fig. 7.10 shows the average of the  $k$ -factor distribution in bins of true  $B$  decay time. The behavior can be described by a first order polynomial  $p_0 + p_1 t_{true}$ , with the slope  $p_1$  compatible with zero. In this  $a_{sl}^d$  measurement the mean value of the  $k$ -factor distribution is assumed to be independent of the true  $B$  decay time. The resolution on the flight distance is assumed to be independent of the  $B$  momentum.

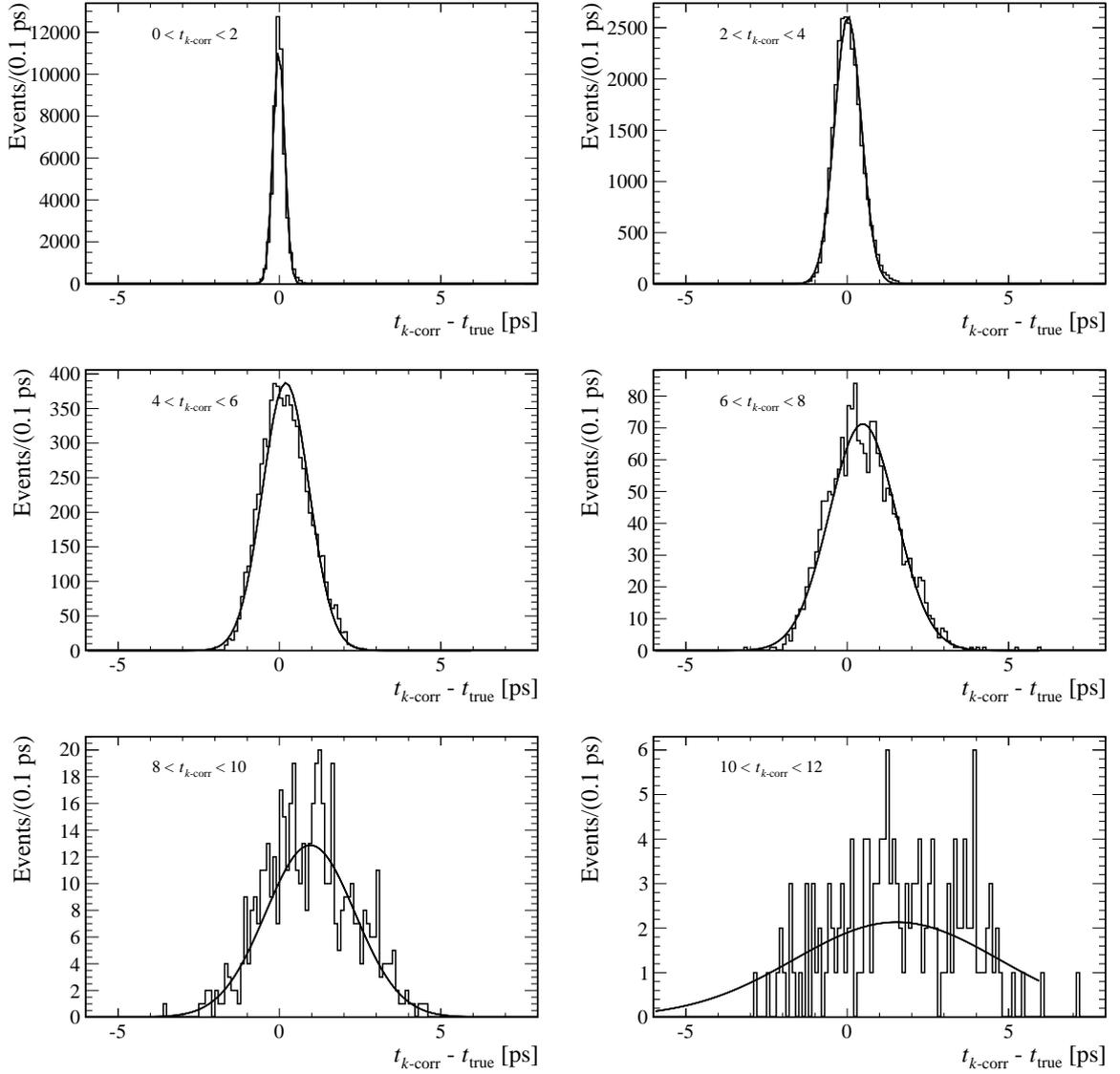


Figure 7.8:  $k$ -corrected  $B$  decay time resolution in  $t_{k\text{-corr}}$  bins.  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  simulated events with 2011 conditions.

### Flight distance resolution

The effect on the  $B$  decay time due to the uncertainty on the measurement of the flight distance can be isolated from the uncertainty on the momentum by evaluating for each event the  $B$  decay time  $t' = \frac{L_{\text{reco}} M_{PDG}}{p_{\text{true}}}$  and comparing  $t'$  to the true  $B$  decay time. The  $t' - t_{\text{true}}$  distributions for the two decay modes are reported in Fig. 7.11. The width of the distributions is of about 70 fs and the shape is well described by a triple Gaussian shape.

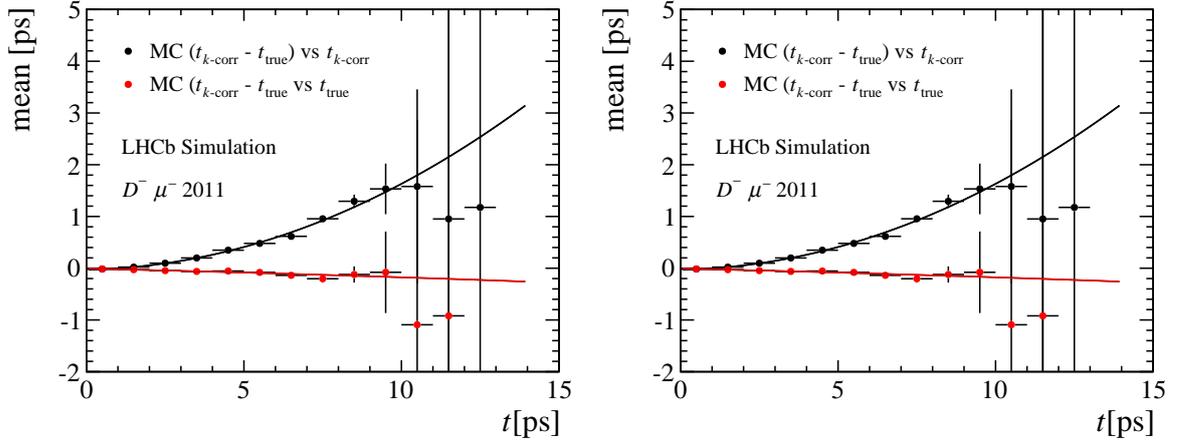


Figure 7.9:  $k$ -corrected  $B$  decay time resolution mean and width in function of  $t_{k\text{-corr}}$  and  $t_{\text{true}}$ .  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  simulated events with 2011 conditions.

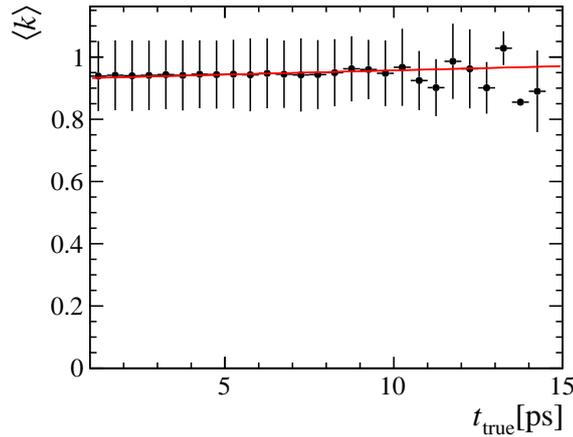


Figure 7.10:  $\langle k \rangle$  in function of the true  $B$  decay time. The fit function is a first order polynomial  $p_0 + p_1 t_{\text{true}}$ . Fit results :  $p_0 = 0.93 \pm 0.05$  and  $p_1 = 0.003 \pm 0.005$ .  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  simulated events with 2011 conditions.

## Momentum resolution

When correcting the sum of the momenta of the final state particles with the  $\langle k \rangle$  value, only the mean value of the  $k$ -factor distribution is exploited. It is possible to take advantage of the full distribution in order to account for the uncertainty on the  $B$  momentum determination. The idea is to learn from the  $k$ -factor distribution the fraction of events that should be corrected with a certain value of  $k$ , and make the corresponding  $k$ -factor correction for only that fraction of events. For example if the choice of using a  $k$ -factor distribution with only two bins is made, and 60% of the MC events have a  $k$ -factor

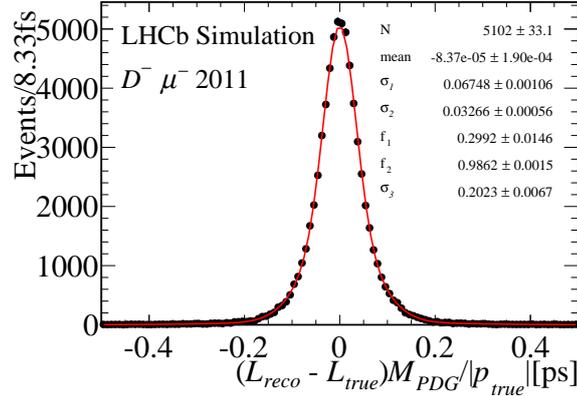


Figure 7.11: Decay time resolution from the MC simulation when taking the true  $B^0$  momentum for both decay modes. A triple Gaussian is fitted to the data. The average resolution is indicated in the plots.

belonging to the first bin centered at the value  $k_1$ , while the other 40% of the MC events can be associated to  $k_2$ , 60% of the events will be corrected with  $k_1$  and 40% with  $k_2$ . This represents the idea of the multiplicative convolution, that will be further explained in the next section.

Since the  $B$  decay time used is corrected with  $\langle k \rangle$  independently from the resolution, the distribution relevant to account properly for the momentum resolution is not  $k$ , but  $k/\langle k \rangle$  displayed in Fig. 7.12, to make the proper correction only once. Using a simplified notation, the reconstructed decay time  $t_{\text{reco}}$  is properly modeled using the multiplicative convolution with the  $k$ -factor distribution in Fig.7.5,

$$t_{\text{reco}} = t_{\text{true}} \otimes F(k), \quad (7.11)$$

where the distributions  $t_{\text{true}}$  and  $t_{\text{reco}}$  are shown in Fig. 7.13. Since the average correction  $\langle k \rangle$  is applied to the data at an early stage, the purpose is now to properly describe the  $\langle k \rangle$ -corrected decay time distribution,  $t_{k\text{-corr}}$ , reproducing the resolution features introduced with this correction. This is done by means of a multiplicative convolution as in the case without  $\langle k \rangle$  correction, but using the  $k/\langle k \rangle$  distribution in Fig. 7.12.

$$t_{k\text{-corr}} = t_{\text{reco}} \langle k \rangle = t_{\text{true}} \otimes F\left(\frac{k}{\langle k \rangle}\right), \quad (7.12)$$

with the  $t_{k\text{-corr}}$  and  $t_{\text{true}}$  are shown in Fig. 7.14. Studies showed that there is no difference in using Eq.7.11 and Eq.7.12 when determining the physics parameters of interest. The convolution with the  $k$ -factor distribution is explained in Sec. 7.2. Details on the mathematical expressions are reported in the Appendix A.

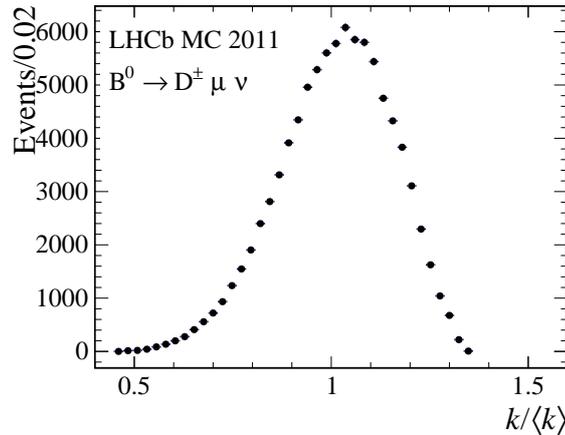


Figure 7.12:  $k/\langle k \rangle$  distribution.  $B^0 \rightarrow D^+ \mu^- \nu_\mu X$  simulated events with 2011 conditions.

## 7.2 Signal decay time model

This section aims to provide a detailed description of the experimental effects that affect the  $B$  decay time distribution. The theoretically predicted distribution, i.e. the decay rate, is here indicated with  $T(t'')$  to be as generic as possible. Where  $t''$  indicate the true decay time:  $t'' = \frac{L_{true} M}{p_{true}}$ . In the case of a not oscillating  $B$  mesons,  $T(t'')$  has a simple exponential shape. For the  $a_{sl}^d$  measurement the decay time distributions theoretically expected are the untagged decay rates reported in Eq. 4.17, that can be written in terms of the detection and production asymmetries:

$$T(f, t'') = \mathcal{N} e^{-\Gamma_d t''} \left( 1 + A_D + \frac{a_{sl}^d}{2} + \left( A_P - \frac{a_{sl}^d}{2} \right) \cos \Delta m_d t'' \right), \quad (7.13)$$

$$T(\bar{f}, t'') = \mathcal{N} e^{-\Gamma_d t''} \left( 1 - A_D - \frac{a_{sl}^d}{2} - \left( A_P - \frac{a_{sl}^d}{2} \right) \cos \Delta m_d t'' \right), \quad (7.14)$$

where  $\mathcal{N}$  is a relative normalization factor and  $\Gamma_d = 1/\tau_d$  is the decay width of the  $B^0$  meson.

### 7.2.1 Decay time resolution

To properly describe the two resolution effects described above, the true distribution of the  $B$  decay time is smeared using a double convolution.

In the first convolution the decay rate is convoluted with a triple Gaussian resolution function with a width of about 70 fs. This convolution is done analytically for the  $a_{sl}^d$  measurement. In Fig. 7.11 the decay time resolution is shown when using the true  $B^0$  momentum for the calculation of the reconstructed time. The second convolution is a multiplicative convolution to describe the effect of the missing momentum. This means

that the true decay time is multiplied by the  $k$ -factor and integrated over the all possible values of  $k$ . Since the  $k$ -factor distribution is obtained from the simulation as a histogram, this convolution is performed numerically. The double convolution is expressed as

$$N(t) = [T(t'') \otimes_t R(t' - t'')] \otimes_k F(k), \quad (7.15)$$

where  $T(t'')$  is the expected decay time distribution without accounting any experimental effect, and  $N(t)$  is the reconstructed decay time distribution.

The first convolution in Eq. 7.15,  $\otimes_t R(t' - t'')$  is analytical and accounts only for the flight distance resolution, and  $\otimes_k F(k)$  is the second convolution of the  $k$ -factor. To explicitly derive Eq. 7.15, the starting point are the convolution integrals. The reconstructed decay time distribution  $N(t)$  of the data is described in terms of the true decay time distribution ( $f(t'')$ ) and resolutions as:

$$N(t) = \int_{-\infty}^{+\infty} \left( \int_{-\infty}^{+\infty} T(t'') R(t' - t'') dt'' \right) F(t'/t) dt', \quad (7.16)$$

where  $t''$  is the true decay time  $t'' = \frac{L_{true}M}{p_{true}}$ .

$t'$  indicates the decay time accounting only for the flight distance resolution effect  $t' = \frac{L_{reco}M}{p_{true}}$ , and  $t$  is the measured time  $t = \frac{L_{reco}M}{p_{reco}}$ . The integral between brackets is nothing but the usual convolution of the true decay time distribution with a resolution function (in general of Gaussian shape) used to obtain the reconstructed time distribution. In this case this integration is done analytically, and gives a distribution in  $t'$ :  $H(f')$ <sup>1</sup>. The second convolution is left:

$$N(t) = \int_{-\infty}^{+\infty} H(t') F(t'/t) dt' = \int_{-\infty}^{+\infty} H\left(\frac{L_{reco}M}{p_{true}}\right) F\left(\frac{L_{reco}M}{p_{true}} / \frac{L_{reco}M}{p_{reco}}\right) d\left(\frac{L_{reco}M}{p_{true}}\right) \quad (7.17)$$

Using the  $k$ -factor definition  $p_{true} = p_{reco}/k$ , and the definition of  $t = \frac{L_{reco}M}{p_{reco}}$ , the integral becomes:

$$N(t) = \int_{-\infty}^{+\infty} H\left(\frac{L_{reco}M}{p_{reco}}k\right) F(k) d\left(\frac{L_{reco}M}{p_{reco}}k\right) = \int_{-\infty}^{+\infty} H(tk) F(k) d(tk) \quad (7.18)$$

Now we want to integrate in the  $k$ -factor distribution, obtaining:

$$N(t) = \int_{k\text{-factor range}} H(tk) F(k) k d(k) \quad (7.19)$$

In the case of having a  $k$ -factor shape easy to model and integrable with the  $H(tk)$  functional shape, this analytical integral should be used. In our case the  $k$ -factor histogram is used, which makes the integral discrete:

$$N(t) = \sum_{k_i} H(tk_i) F_i k_i \Delta k_i \quad (7.20)$$

---

<sup>1</sup>The analytical expression of  $H(f')$  is given in Appendix A.

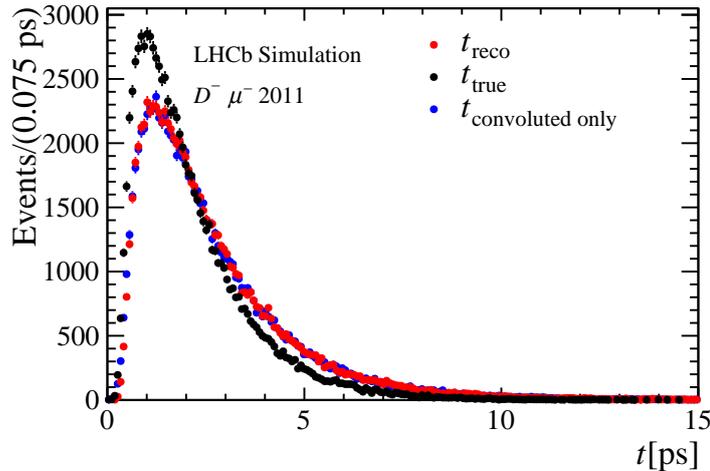


Figure 7.13: Reconstructed and true  $B$  decay time.  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  simulated events with 2011 conditions.

where  $k_i$  is the position of each bin of the  $k$ -factor histogram,  $F_i$  is the height of the bin, and  $\Delta k_i$  the bin width. With a  $k$ -factor histogram normalized to the unity,  $F_i \Delta k_i$  correspond to the fraction of events in MC having a  $k$ -factor contained in the bin  $i$ . In this explicit derivation the normalization of the decay time probability density function has not been considered <sup>2</sup>.

In order to show that this double convolution and the  $k$ -factor correction properly models the decay time distributions, accounting for the resolution effects, a first check has been made by using the  $t_{true}$  distribution from MC samples in place of  $T(t'')$ . Fig. 7.13 shows that the  $k$ -factor corrected and convoluted  $B$  decay time reproduces the decay time reconstructed on the partially reconstructed data sample well. Fig. 7.14 shows that the  $k$ -factor convolution reproduces the behavior of the mean and width of the decay time resolution introduced by the  $k$ -factor correction. This represents a first validation of this method. Subsequent steps are the substitution of the  $t_{true}$  distribution with an analytical  $T(t'')$  distribution (Eq. 8.23 in the case of the  $a_{sl}^d$  measurement). Fig. 7.15 shows how applying the  $k$ -factor correction and convolution is possible to obtain the true  $B$  decay time distribution.

## 7.2.2 Decay time acceptance

Some of the selection criteria applied introduce a distortion of the decay time distributions of the signal  $B^0$  decays. These criteria are for example the IP requirements and the effect of the reconstruction itself affecting the decay time distribution for large decay times. Due

<sup>2</sup> The normalization used for the fit is  $\mathcal{N}(t_{min}, t_{max}) = \int_{t_{min}}^{t_{max}} dt \sum_{k_i} A(t) H(tk) F_i k_i \Delta k_i$ , where  $a(t)$  is the acceptance shape described later in the text and  $H(tk)$  is the convolution of the true time decay rate description and the flight distance triple Gaussian distribution.

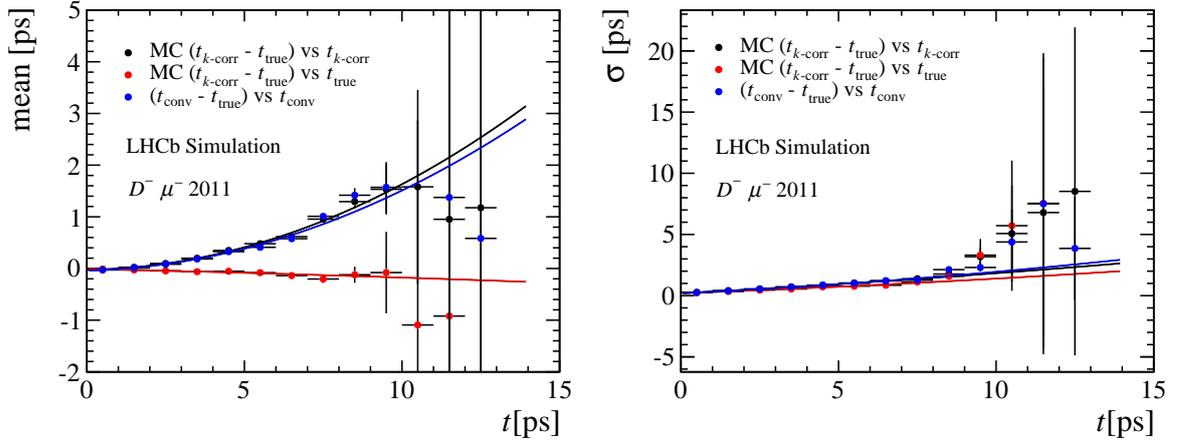


Figure 7.14:  $k$ -corrected and convoluted  $B$  decay time resolution mean and width in function of  $t_{k\text{-corr}}$  and  $t_{\text{true}}$ .  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  simulated events with 2011 conditions.

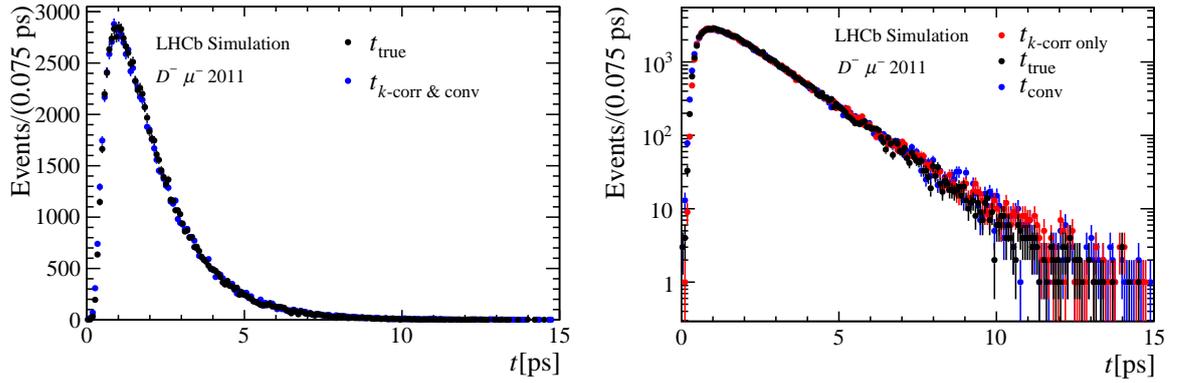


Figure 7.15:  $k$ -corrected and convoluted  $B$  decay time compared to the  $k$ -corrected only  $B$  decay time and the true  $B$  decay time. These plots should be compared to Fig. 7.6, where the momentum resolution effect (modeled by means of the  $k$ -factor convolution), is not taken into account.  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  simulated events with 2011 conditions.

to these effects, the time distribution cannot be described by a simple exponential. An acceptance function has to be included in the model. This acceptance describes the decay time dependence of the signal selection efficiency. It can be obtained from the ratio of the signal MC reconstructed  $B$  decay time distribution over a theoretical distribution of the  $B$  decay time, not sculpted by any selection criteria, but accounting for the resolution effects

$$a(t) = \frac{N(t)}{([T(t'') \otimes_t R(t' - t'')] \otimes_k F(k))}, \quad (7.21)$$

where the  $T(t'')$  is simply the theoretical expression of the decay rates, and  $N(t)$  the reconstructed decay time distribution. A simple acceptance function is chosen to parametrize

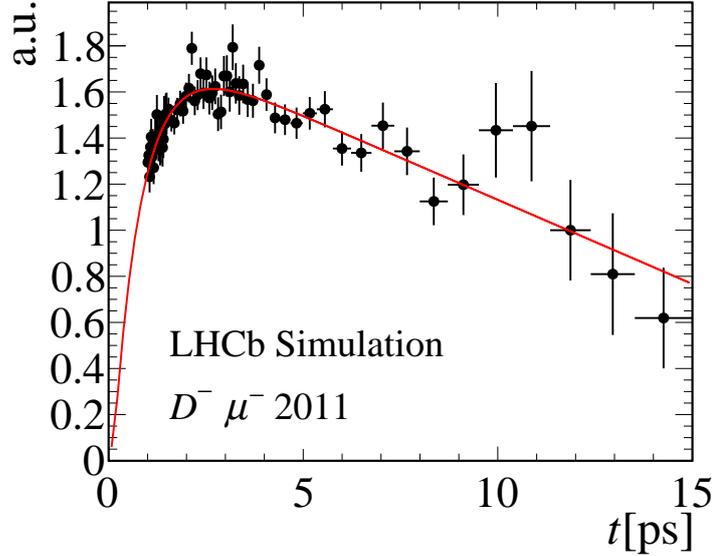


Figure 7.16: Reconstructed and true  $B$  decay time.  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  simulated events with 2011 conditions.

this ratio between the two distributions, shown in Fig. 7.16,

$$a(t) = (1 - e^{-(t-t_{\text{shift}})/\alpha})(1 + \beta t)\Theta(t - \zeta) , \quad (7.22)$$

where  $t_{\text{shift}}$  and  $\alpha$  describe the curvature of the rising edge at low decay times, and  $\beta$  models the linear decrease in efficiency at higher decay times. The Heaviside step function  $\Theta$  indicates the cut-off at low decay times. For decay times below  $\zeta$  the acceptance  $a(t)$  is zero. Since this simple acceptance function is unable to accurately describe the high statistics data samples at low lifetimes, the range of the measurement has been restricted to  $[1, 15]$  ps. This choice has a large impact on the sensitivity of  $A_P$ , given that the time zero intercept gives a better handle in order to determine the amplitude of an oscillation. On the other hand, the effect on the determination of  $a_{\text{sl}}^d$  is given only by the statistical reduction of the data sample, which is negligible for this measurement (see Sec. 9.1.5 and Tab. 9.2). Variations of the low limit of the decay time range have been considered in the evaluation of the systematic uncertainties. The acceptance parameters are not fixed in the fit procedure that determines the  $a_{\text{sl}}^d$  and  $A_P$  values, but they are allowed to vary, in order to compensate any possible difference between data and simulated events and for the low statistics of the MC samples. The upper lifetime acceptance effect originates from the trigger and reconstruction due to (loose) assumptions on the pointing direction of the tracks. This effect is described by the parameter  $\beta$  in this model, and it is fully correlated with the lifetime of the  $B^0$ . For this reason the parameter  $\beta$  is left free to float in the fit, while the  $B^0$  lifetime is fixed [17].

In summary, the total signal decay time model can be expressed as the product of the

double convolution with the acceptance function,

$$N(t) = \mathcal{N} \times ([T(t'') \otimes_t R(t' - t'')] \otimes_k F(k)) \times a(t) , \quad (7.23)$$

where  $\mathcal{N}$  is a normalization factor.

The assumption behind this decay time description is that the  $k$ -factor distribution does not depend on the true time of the decays. Experimentally the time acceptance, however, can introduce a dependence. A non-negligible dependence of the  $k$ -factor distribution can cause a bias in the measured value of physical parameters sensitive to the decay time scale, e.g. in the  $B$  lifetime or mixing frequency. Quantities like  $a_{\text{sl}}^d$  are less dependent on the absolute scale of the  $B$  decay time. More relevant is the relative scale of  $f$  and  $\bar{f}$  final states in this case. The effect of having a small dependence of the  $k$ -factor distribution on the true time is equivalent to assume a slightly different value for  $\Delta m_d$ . This effect is estimated to be  $0.003 \text{ ps}^{-1}$ . This variation is considered in the systematics of the  $a_{\text{sl}}^d$  analysis (the maximum variation considered for  $\Delta m_d$  is the uncertainty from the PDG average:  $0.004 \text{ ps}^{-1}$ ).

### 7.3 Summary

The  $k$ -factor method presented here accounts for the missing neutrino momentum and properly describes the  $B$  decay time.

It relies on MC simulation and consists of two steps. The first is the correction of the sum of the momenta of the  $B$  reconstructed daughters with the average  $\langle k \rangle (M(B^0)) = \langle p_{\text{reco}}/p_{\text{true}} \rangle (M(B^0))$ .

The second step is a multiplicative convolution of the decay time distribution with the  $k/\langle k \rangle$  overall distribution to properly account for the time resolution.

In spite of being less important than the momentum resolution, also the flight distance resolution is accounted for by convolving the time distribution with a triple gaussian shape of about 70 fs width as determined from simulated events.

The effect of the overall selection efficiency of the data samples is accounted for by an acceptance function, determined from MC samples. The acceptance parameters are allowed to vary in the fit to the data.

# Chapter 8

## Time-dependent fit

This chapter describes the time-dependent fit used to extract  $a_{\text{sl}}^d$ , the physics parameter of interest. A standard maximum likelihood fit procedure commonly applied in experimental analysis on large data samples is used, with the additional challenge of the description of the  $B$  decay time resolution in partial reconstructed decays.

Firstly the maximum likelihood fit method in general is discussed, focusing on the aspects relevant for the fit used in the  $a_{\text{sl}}^d$  measurement: Proper treatment of weighted datasets and the goodness-of-fit evaluation for maximum likelihood fits of binned datasets. A more complete and detailed explanation of the maximum likelihood method can be found for example in Ref. [17].

Subsequently the fit used to extract  $a_{\text{sl}}^d$  is described in detail, focusing on the description used for each component of the data sample. Moreover, the validation of the fit is presented. Finally, the results obtained in each sub-set of the data samples are reported.

### 8.1 Maximum likelihood fits

Maximum likelihood fits are largely used to estimate parameters in analyses of experimental data. Suppose a set of  $N$  independent measurements,  $x_i$  (with  $i = 1, 2, \dots, N$ ), is given and the purpose is to determine a set of unknown parameters  $\boldsymbol{\theta}$  describing the distribution of the  $x_i$  measured values. The Probability Density Function (PDF), denoted with  $p(x_i|\boldsymbol{\theta})$ , describes the expected distribution of the measurements in the measured variable  $x$ , given a vector of unknown parameters  $\boldsymbol{\theta}$ . Estimators for  $\boldsymbol{\theta}$  are those values that maximize the likelihood function

$$L(\boldsymbol{\theta}) = \prod_{i=1}^N p(x_i|\boldsymbol{\theta}). \quad (8.1)$$

The PDF,  $p(x_i|\boldsymbol{\theta})$  has to be normalized such that  $\sum_i p(x_i|\boldsymbol{\theta}) = 1$ . A common procedure is to perform a minimization of the negative logarithm of the likelihood instead of the maximization of the likelihood itself, in order to avoid numerical precision problems that can arise when multiplying very small numbers. Both  $\ln L(\boldsymbol{\theta})$  and  $L(\boldsymbol{\theta})$  are maximized by the same parameter values  $\hat{\boldsymbol{\theta}}$ . Finding the maximum likelihood estimators means to solve

the so-called *maximum likelihood equations*:

$$\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \theta_i} = 0, \text{ with } \theta_i \in \boldsymbol{\theta}. \quad (8.2)$$

In the large sample limit,  $L$  has a Gaussian form and  $\ln L$  is parabolic. In this case, a possible numerical way to determine the  $s$ -standard deviation errors on the estimated parameters is from the hypersurface defined by  $\boldsymbol{\theta}'$  such that  $\ln L(\boldsymbol{\theta}') = \ln L_{\max} - s^2/2$ , where  $\ln L_{\max}$  is the value of  $\ln L$  at the solution point. The minimum and maximum values of  $\theta_i$  on the hypersurface then give an approximate  $s$ -standard deviation confidence interval for  $\theta_i$ <sup>1</sup>. [17]

The same strategy can be exploited to determine the best values of the parameters  $\boldsymbol{\theta}$ , when considering a sample of  $N$  events, and for each event the vector  $\boldsymbol{x}_i$  collects the  $k$  variables characterizing the event (like kinematic and identification, or more specifically to this case decay time  $t$ , the charge of the final state  $\xi$  and mass  $m$ ). In this case, the PDF  $p(\boldsymbol{x}_i|\boldsymbol{\theta})$  can also be understood as the probability to observe an event characterized by the observed quantities  $\boldsymbol{x}_i$ , in the hypothesis that such event is belonging to the class of events described by the vector of parameters with values  $\boldsymbol{\theta}$ . In this case the method is called *unbinned* since the likelihood function is based on event-by-event values of the measured quantities. On the other hand it is possible to construct also *binned* likelihoods. In this second case a histogram (one- or multi-dimensional) is considered, and the observed population of each bin of the histogram is compared to expected value for the population of the bin, given a set of unknown parameters  $\boldsymbol{\theta}$ . The best estimate for the parameters  $\boldsymbol{\theta}$  is the set of values that maximizes the likelihood function.

Suppose a simple case of a one-dimensional histogram with  $M$  bins describing the decay time  $t$  distribution of a sample of  $B^0$  decays. The number of events in each bin  $i$  is indicated with  $O_i$  (with  $i = 1, 2, \dots, M$ ), while the expected number of events  $E_i$  for each  $i$  bin, can be calculated given the PDF  $p(t|\boldsymbol{\theta})$ :

$$E_i(\boldsymbol{\theta}) = \int_{\text{bin } i} p(t|\boldsymbol{\theta}) dt. \quad (8.4)$$

Depending on the nature of the experimental data, two different statistics can be assumed for the number of observed events in each bin,  $O_i$ . Here the *Poisson* statistics is assumed.<sup>2</sup> The likelihood function writes as follows

---

<sup>1</sup>In order to estimate the uncertainty on the parameters of an  $n$  parameter estimation, the full covariance matrix  $V_{ij} = \text{cov}[\hat{\theta}_i, \hat{\theta}_j]$  is required. The inverse of the covariance matrix is estimated as

$$(\hat{V}_{ij}) = \left. \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \right|_{\hat{\boldsymbol{\theta}}} \quad (8.3)$$

<sup>2</sup> For counting experiments, where  $N$  is one of the parameters to be measured, the *Poisson* statistics applies. Alternatively, if the total number of measured events  $N$  is fixed by the design of the experiment, the distribution of the events among the bins will be *multinomial*. The Poisson and multinomial likelihood

$$L(\boldsymbol{\theta}) = L(\mathbf{E}(\boldsymbol{\theta}); \mathbf{O}) = \prod_{i=1}^M \frac{E_i(\boldsymbol{\theta})^{O_i} e^{-E_i(\boldsymbol{\theta})}}{O_i!}, \quad (8.6)$$

the negative logarithm of the likelihood becomes

$$-\ln L(\boldsymbol{\theta}) = -\sum_{i=1}^M (O_i \ln E_i(\boldsymbol{\theta}) - E_i(\boldsymbol{\theta}) - \ln(O_i!)). \quad (8.7)$$

The last term doesn't depend on the parameters  $\boldsymbol{\theta}$ , so it doesn't affect the minimization procedure, being just a constant off-set.

For the  $a_{\text{sl}}^d$  measurement, the high statistics of the data samples analyzed requires the usage of a binned fit, in order to obtain a reasonable computation time. A binned multidimensional maximum likelihood fit is used to extract the parameters of interest. The function in Eq. 8.7 is the function minimized, and for the numerical minimization procedure the Minuit package [115] and the ROOT framework [116] are used.

The following two paragraphs aim to describe few delicate aspects of maximum likelihood fits implementation and usage. When fitting an experimental distribution with a function, typically there are three goals. The first is to determine the “best fit” parameters for the function to describe the data distribution. The second is the determination of the error on the parameter. The third is to judge the quality of the fit. The first goal has been addressed in this section, some selected indications of the second are given in Sec. 8.1.1 and the fit quality for binned maximum likelihood fits is discussed in Sec. 8.1.2.

### 8.1.1 Weighted datasets

The statistical significance of a data sample is reduced when applying a weighting procedure. Given a sample of  $N$  events, and having associated a weight  $w_i$  to each event by means of some weighting procedure, the yield obtained is

$$N' = \sum_{i=1}^N w_i. \quad (8.8)$$

According to the weighting procedure, and limits applied on the weights,  $N'$  could result larger than  $N$ . This would mean that applying different weights to the events it is possible

---

functions relate to each other as

$$L(\boldsymbol{\theta})_p = \frac{N!}{N^N} \prod_{i=1}^M \frac{E_i(\boldsymbol{\theta})^{O_i}}{O_i!} \frac{e^{-N_0} N_0^N}{N!} = L(\boldsymbol{\theta})_m \frac{e^{-N_0} N_0^N}{N!}, \quad (8.5)$$

where  $N_0 = \sum_{i=1}^M E_i$  is the total number of events predicted by the model. Further discussion on the most appropriate statistics to be assumed in a physics analysis can be found in Ref. [114] and [17]

to increase not only the statistical power of the single events, but also of the whole sample. To prevent this, every event is normalized with the factor

$$\alpha = \frac{\sum_i w_i}{\sum_i w_i^2}, \quad (8.9)$$

that ensures that the effective sample size is always smaller than, or equal to  $N$ . For a sample of  $N$  events, the effective size is

$$N_{eff} = \alpha \sum_i w_i. \quad (8.10)$$

One can compare  $N_{eff}$  to  $N$ . Given the inequality

$$\sum_i w_i \leq \sqrt{N \sum_i w_i^2}, \quad (8.11)$$

it follows  $N_{eff} \leq N$ . The equal sign is obtained when all the weights have the same value. In the case of the kinematical weights used on the data samples for the  $a_{s1}^d$  measurements, as explained in Sec. 6.5.1, the reduction of the statistical power of the data sample reflects the initial agreement between the kinematical distributions of the muon and pion.

In the case of maximum likelihood fits the weights need to be properly accounted for, in order to obtain the correct error coverage. When considering a weighted dataset, a weighted likelihood [117] fit needs to be performed

$$L_W(\boldsymbol{\theta}) = \prod_{i=1}^N p(x_i|\boldsymbol{\theta})^{\alpha w_i}. \quad (8.12)$$

Denoting with  $O_i^W$  the observed number of events in the  $i$ -th bin with weights  $w_i$  applied, not yet normalized with  $\alpha$ ,  $O_i^W = \sum_i w_i$ , the function minimized by the fit is

$$-\ln L_W(\boldsymbol{\theta}) = -\sum_{i=1}^M (\alpha O_i^W \ln(\alpha E_i(\boldsymbol{\theta})) - \alpha E_i(\boldsymbol{\theta})). \quad (8.13)$$

This approach is verified to give the correct error coverage (see Sec. 8.5). Here the focus is on the kinematical weights applied in the  $a_{s1}^d$  analysis, but the same treatment is possible for any kind of weight. A well-known example is given by the sWeights, that associate to each event a probability of being a signal event, calculated on the distribution of a variable independent on the fit variables. More details can be found in Ref. [117].

### 8.1.2 Quality of a binned maximum likelihood fit

In the case of maximum likelihood fits on binned datasets, it is possible to obtain a global goodness-of-fit indicator, similar to the  $\chi^2/d.o.f.$  in the case of  $\chi^2$  fits. Moreover, in case

of multidimensional fits, there is a number of tests that can be performed on the single fit projections.

The *fit projection* of a  $k$ -dimensional fit onto the variable  $x^j$  of the variables vector  $\mathbf{x}$  (where  $\mathbf{x}$  is the continuous distribution of the measurements  $\mathbf{x}_i$  previously introduced) is defined as the following one-dimensional function:

$$P(x^j|\hat{\boldsymbol{\theta}}) = \int p(\vec{x}|\hat{\boldsymbol{\theta}})dx^1 \dots dx^{j-1}, dx^{j+1} \dots dx^k, \quad (8.14)$$

which is the predicted distribution for  $x^j$  under the assumed values for the fit parameters, and it can be overlaid to the experimental data. The case of binned fits is analogous. Considering for example for simplicity a two-dimensional fit in  $x^1$  and  $x^2$  dimensions. The expected number of events predicted by the model in the  $ij$ -bin is

$$E_{ij} = \int_{\text{bin } ij} p(\mathbf{x}|\boldsymbol{\theta})dx^1 dx^2 \quad (8.15)$$

and the fit projection onto the variable  $x^1$  is

$$P(x^j|\hat{\boldsymbol{\theta}}) = \sum_{n=1}^{\text{bins in } x^2} E_{i,n}(\boldsymbol{\theta}). \quad (8.16)$$

Quality tests on the projections are normally used as qualitative goodness-of-fit indicators in the case of unbinned maximum likelihood fits, where it is not possible to take advantage of a global goodness-of-fit indicator. The following sections develop these goodness-of-fit tests in the case of maximum likelihood fits on binned datasets.

### Global goodness-of-fit: likelihood ratio

When considering maximum likelihood fits of binned datasets, the likelihood function can be converted into a general  $\chi^2$  statistics, by making use of the likelihood ratio test [114]. The likelihood ratio  $\lambda$  is defined as

$$\lambda = L(\mathbf{E}(\boldsymbol{\theta}); \mathbf{O})/L(\mathbf{T}; \mathbf{O}), \quad (8.17)$$

where  $\mathbf{O} = (O_1, O_2, \dots, O_i, O_N)$  represents in this case the observed bin content in the data histogram,  $\mathbf{E} = (E_1, E_2, \dots, E_i, E_N)$  stands for the number of expected (predicted by the model) events in each bin, and  $\mathbf{T}$  refers to the true, unknown values of  $\mathbf{O}$ . Using the likelihood ratio theorem, the ‘‘likelihood  $\chi^2$ ’’ can be defined as  $\chi_\lambda^2 = -2\ln\lambda$  and asymptotically obeys a chi-square distribution. For Poisson-distributed histograms,  $\mathbf{T}$  can be replaced by  $\mathbf{O}$ , and then the likelihood ratio results from the comparison of the value of the likelihood function at the minimum, and the likelihood function calculated assumed the expected bin content to be equal to the observed bin content. This leads to the Poissonian likelihood  $\chi^2$ :

$$\chi_\lambda^2 = 2 \sum_i E_i(\boldsymbol{\theta}) - O_i + O_i \cdot \ln(O_i/E_i(\boldsymbol{\theta})). \quad (8.18)$$

This test represents a global goodness-of-fit estimation. In multidimensional problems with high statistics data samples as the determination of  $a_{\text{sl}}^d$ , it might happen that the fit model in one fit dimension is not optimal to describe the data. In such cases it is crucial to establish whether the distribution that is not perfectly described affects or not the estimation of the parameter of interest.

### Quality of the fit projections

For each projection of the binned multidimensional fit, the simple  $\chi^2/d.o.f.$  gives an indication of the goodness of the fit model in the dimension examined, where the number of degrees of freedom in this case is the number of bins in the dimension under study  $D$ , and  $\chi^2 = \sum_{i=1}^D (O_i - E_i)^2 / \sigma_i^2$ , with  $(i = 1, \dots, D)$ ,  $O_i$  and  $E_i$  the observed and expected numbers of events in the  $i$ -th bin, and  $\sigma_i$  indicates the uncertainty of the  $i$ -th bin.

Given a fit projection, also the distribution of the so-called pulls can be used for quality checks. The pull for the  $i$ -th bin is defined as  $(O_i - E_i) / \sigma_i$ . In addition, it is possible to perform a *runs test* or *Wald-Wolfowitz test* on the pull distribution.

The runs test aims to test a randomness hypothesis for a two-valued data sequence. In particular it is interesting to test the hypothesis that the elements of the sequence are mutually independent. In the present case, it can be used to verify whether the pulls are randomly distributed or there is some pattern due to some problems of the fit model used. In this case the two-valued data in our sequence are positive and negative pulls. A “run” or “sequence” is defined by a number of either positive or negative pulls in bins next to each other. Given the number of negative ( $N_-$ ) and positive ( $N_+$ ) pulls, the mean number of sequences expected when assuming the random hypothesis, is estimated to be equal to  $\mu$  and with a variance  $\sigma^2$ :

$$\mu = 1 + 2 \frac{N_- N_+}{N_- + N_+}, \quad (8.19)$$

$$\sigma^2 = \frac{(\mu - 1)(\mu - 2)}{N_- + N_+ - 1}. \quad (8.20)$$

The number of sequences observed on the pulls of the charge asymmetry projection on the data fits, should be compared with the expected value  $\mu$  within the uncertainty  $\sigma$ .

## 8.2 Determination of $a_{\text{sl}}^d$ and $A_P$

In order to determine the value of  $a_{\text{sl}}^d$  a binned maximum likelihood fit is used. As explained in Sec. 6.5.1, it is necessary to associate to each event a kinematical weight to equalize the muon and the pion kinematics in the decay sample. These weights are calculated and associated to the events before the fitting procedure, that is therefore performed on the weighted data sample. The  $B^0$  decay time distribution of the events is described separately for the  $f = (K^+ \pi^- \pi^- \mu^+)$  and  $\bar{f} = (K^- \pi^+ \pi^+ \mu^-)$  final states, in order to extract the time dependent asymmetry. The mass of the  $D^\pm$  for the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  data sample and of

the  $D^0$  in the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  data sample, is used in order to separate signal decays from events of combinatorial background and decays not involving a  $D^\pm/D^0$  meson, as it is explained later in this section. Therefore, a multidimensional fit in  $B$  decay time, charge of the final state particles and  $D^\pm/D^0$  mass is used. When the dimensions of the fit are independent, the PDFs factorize. This is the case for the  $D^\pm/D^0$  mass which can be treated fully independently from the  $B^0$  decay time and charge of the final state. For each component of the fit it is therefore possible to write the PDF as:

$$\mathcal{P}(m, t, f) = \mathcal{P}(t, f) \times \mathcal{P}(m), \quad (8.21)$$

where  $m$  indicates the  $K^+ \pi^- \pi^- / K^+ \pi^-$  invariant mass and  $t$  refers to the  $B$  decay time. There are three categories of events in the sample that can be distinguished in the fit procedure:

- *Signal decays.* These are the events of interest, their decay time distribution is the  $B^0$  decay time distribution containing the information on  $CP$  violation in  $B^0$  mixing. The  $D^\pm/D^0$  meson in the final state leads to an enhancement of the number of events in the  $K^+ \pi^- \pi^- / K^+ \pi^-$  invariant mass spectrum, at values corresponding to the  $D^\pm/D^0$  mass.
- *$B^+$  decays.* As previously discussed their topology is very similar to the signal decays. The structure of their  $D^\pm/D^0$  mass distribution is not distinguishable from the signal shape, and the decay time distribution does not provide a strong separation. The fraction of this component is fixed in order to gain in stability of the global fits and to provide a more reliable result. This type of background is particularly dangerous because of the  $B^+$  production asymmetry directly affecting the time dependent charge asymmetry of the final state particles. This has a sizable impact on  $a_{\text{sl}}^d$ , and it is accounted for by a systematic uncertainty, see Sec. 9.1.2.
- *$D^\pm/D^0$  mass sidebands decays.* These are combinatorial background events and decays not involving a  $D^\pm/D^0$  meson from the signal decays. The shape in the  $D^\pm/D^0$  mass allows to separate this background from the rest of the data sample. An effective model is used to describe their decay time behavior.

In Sec. 8.3 the detailed description of all the components of the sample and models in the fit is reported. Not all the possible backgrounds are accounted for in the nominal fit, but the effect of neglecting these background contributions is estimated with the systematic uncertainties described in Sec. 9.1.4.

The full PDF used to describe the data sample reads

$$\mathcal{P}(m, t, f) = (1 - f_{\text{sb}}) \{ f_{\text{sig}} \cdot \mathcal{P}_{\text{sig}}(m, t, f) + (1 - f_{\text{sig}}) \cdot \mathcal{P}_{B^+}(m, t, f) \} + f_{\text{sb}} \cdot \mathcal{P}_{\text{sb}}(m, t, f). \quad (8.22)$$

The fraction  $f_{\text{sb}}$ , of the sideband background is extracted from the fit to the data sample, while the  $B^+$  background fraction,  $(1 - f_{\text{sig}})$  is fixed as explained in Sec. 5.3.1.

## 8.3 Fit to the decay time

In this section the model used to describe the decay time distribution of each component separately for  $f$  and  $\bar{f}$  final states, is described. All the PDFs are reported without including the normalization factor and do not include the convolutions with resolution effects, when they are necessary. These technical informations can be found in Appendix A.

### 8.3.1 Signal decays

The untagged decay rates, as derived in Chapter 4 are expressed in terms of the production and detection asymmetries  $A_P$  and  $A_D$  as follows

$$\mathcal{P}_{\text{sig}}(f, t) = \mathcal{N}_{\text{sig}} e^{-\Gamma_d t} \left( 1 + A_D + \frac{a_{\text{sl}}^d}{2} - \left( A_P + \frac{a_{\text{sl}}^d}{2} \right) \cos \Delta m_d t \right), \quad (8.23)$$

$$\mathcal{P}_{\text{sig}}(\bar{f}, t) = \mathcal{N}_{\text{sig}} e^{-\Gamma_d t} \left( 1 - A_D - \frac{a_{\text{sl}}^d}{2} + \left( A_P + \frac{a_{\text{sl}}^d}{2} \right) \cos \Delta m_d t \right). \quad (8.24)$$

These terms are convoluted with the flight distance resolution and  $k$ -factor distributions as explained in Sec. 7.2.1 and multiplied by the acceptance function detailed in Sec. 7.2.1. Here  $f$  denotes the charge of the reconstructed final state particles. It is possible to write the PDFs in a more compact form by defining  $\xi$  such that to the final state  $f = (K^+\pi^-\pi^-\mu^+)$  a value  $\xi = 1$  is associated and to the  $CP$  conjugate final state  $\bar{f} = (K^-\pi^+\pi^+\mu^-)$ ,  $\xi = -1$  is associated. The untagged decay rates become

$$\mathcal{P}_{\text{sig}}(t, \xi) = \mathcal{N}_{\text{sig}} e^{-\Gamma_d t} \left( 1 + \xi A_D + \xi \frac{a_{\text{sl}}^d}{2} - \xi \left( A_P + \frac{a_{\text{sl}}^d}{2} \right) \cos \Delta m_d t \right). \quad (8.25)$$

An important validation test for this fit model is the fit on MC signal samples, which are generated assuming no  $CP$  violation in  $B^0 - \bar{B}^0$  mixing ( $a_{\text{sl}}^d = 0.0$ ). Fig. 8.1 shows the time distribution for the two final states with the fit model overlaid and Tab. 8.1 reports the fit results. Only the  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  MC sample corresponding to the 2011 conditions is included here, but similar results are obtained with the other signal MC samples. The parameters  $a_{\text{sl}}^d$  and  $A_P$  are the physics parameters of interest, while the remaining parameters describe the acceptance shape  $(1 - e^{-t-t_{\text{shift}}}/\alpha)(1 + \beta t)$ , see Eq. 7.22. The quality of the charge asymmetry fit projection, in Fig. 8.1 (*right*), and the value of  $a_{\text{sl}}^d$  obtained from the fit which are compatible with zero, support the suitability of this model to describe the signal decays. Tab. 8.2 reports the correlations between the parameters involved in the fit model. The physics parameters of interest are not correlated with other parameters, which is a desirable feature of stable fits. The two parameters of the acceptance describing the rising edge  $t_{\text{shift}}$  and  $\alpha$  have a strong correlation. This doesn't compromise the convergence of the fit on the MC signal sample, but is a source of instability for the full fit on data. For this reason the choice of fixing this parameter to the value determined by the fit on the signal MC sample has been made. Different choices

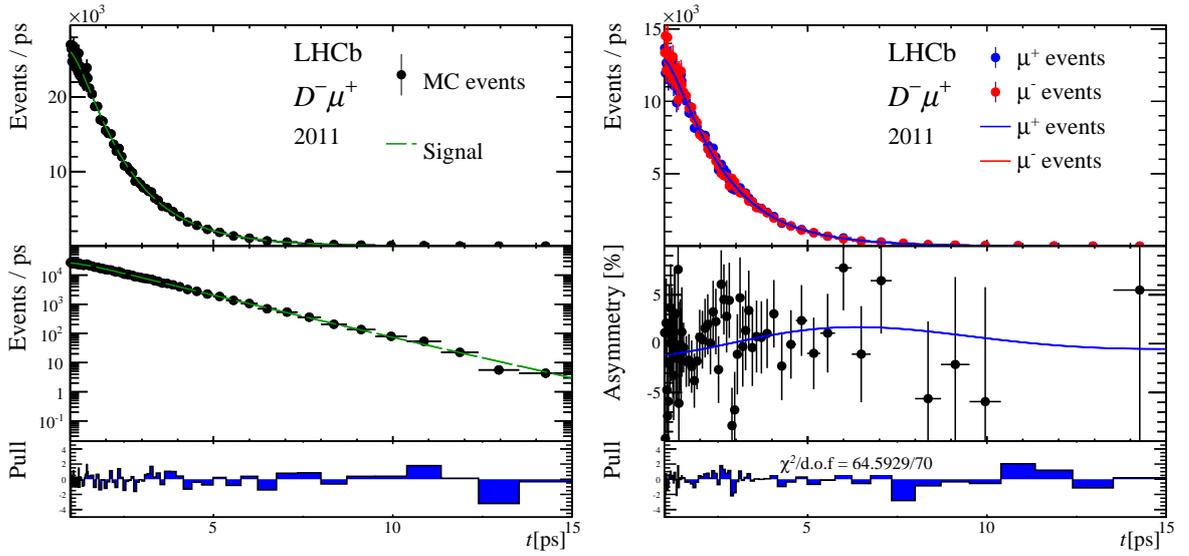


Figure 8.1: Fit to the decay time of  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  simulated events with 2011 conditions. *Left*: Decay time distribution for all the events, in linear and logarithmic scale. The last row shows the so-called *Pull* values, i.e. the difference between the number of events observed in data for each bin of the histogram and the number of events predicted by the model in the same bin, divided by the uncertainty on the number of events. *Right*: decay time distribution for  $f$  and  $\bar{f}$  final states shown separately and charge asymmetry of the final state particles. The *Pull* values in the last row refer to the charge asymmetry projection in the second row.

Table 8.1: Time-dependent fit results for the simulated  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  signal decays.

Parameter	MC 2011
$\alpha$	$0.554 \pm 0.094$
$\beta$	$-0.036 \pm 0.003$
$t_{shift}$	$0.197 \pm 0.139$
$a_{sl}^d$	$0.012 \pm 0.011$
$A_P$	$-0.006 \pm 0.008$

of the fit interval and acceptance description have been probed to evaluate the systematic uncertainty related to the acceptance. As shown in Fig. 8.1 a dynamical binning scheme in the decay time fit dimension is used. This allows to have a sufficiently large number of events in every bin to describe both the rapidly-varying time distribution at short decay times and the behavior at large decay times.

Table 8.2: Time-dependent fit correlation matrix for the simulated  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  signal decays.

	$\alpha$	$\beta$	$t_{shift}$	$a_{sl}^d$	$A_P$
$\alpha$	1.00	-0.59	0.96	0.00	0.00
$\beta$		1.00	-0.45	-0.01	0.00
$t_{shift}$			1.00	0.00	0.00
$a_{sl}^d$				1.00	0.05
$A_P$					1.00

### 8.3.2 $B^+$ decays

The PDF of the  $B^+$  background has a functional shape very similar to that of the signal description, except that it is not oscillating and no  $CP$  violation is expected,

$$\mathcal{P}_{B^+}(f, t) = \mathcal{N}_{B^+} e^{-t/\tau_{B^+}} (1 + A_D - A_{P,B^+}) (1 - e^{-(t-t_{shift,B^+})/\alpha_{B^+}}), \quad (8.26)$$

$$\mathcal{P}_{B^+}(\bar{f}, t) = \mathcal{N}_{B^+} e^{-t/\tau_{B^+}} (1 - A_D + A_{P,B^+}) (1 - e^{-(t-t_{shift,B^+})/\alpha_{B^+}}). \quad (8.27)$$

The subscript “ $B^+$ ” indicates that these parameters are specific to the  $B^+$  background model. This is an effective model used for simplicity, where there is no double convolution and no upper decay time acceptance, which means that their effects are now absorbed in  $\tau_{B^+}$ ,  $t_{shift,B^+}$ , and  $\alpha_{B^+}$  parameters. This choice is reasonable, since there is no attempt to measure the lifetime of this background, and the final state charge asymmetry in the case of the  $B^+$  is not dependent on the decay time. The acceptance parameters are taken from the simulation.

The detection and production asymmetry for the  $B^+$  events cannot be disentangled. The detection asymmetry,  $A_D$ , can be assumed to be the same as that of the  $B^0$  signal decays. Due to the selection the momentum spectra of the decay products for the  $B^0$  signal and  $B^+$  background are nearly identical, as can be seen for example from the distributions reported in Fig.8.2. This guarantees that both detection asymmetries are the same.

On the other hand, the  $B^+$  production asymmetry,  $A_{P,B^+}$ , has to be taken from external inputs. In Ref. [118] the  $B^+$  production asymmetry is measured with  $B^+ \rightarrow J/\psi K^+$  decays using 2011 and 2012 data. A raw charge asymmetry of  $A_{raw}(J/\psi K^+) = (-1.3 \pm 0.1)\%$  is measured, which, after correcting for the measured  $CP$  asymmetry  $A_{CP}(B^+ \rightarrow J/\psi K^+) = (0.3 \pm 0.6)\%$  from the PDG [6] and the detection (kaon) asymmetry  $(-1.0 \pm 0.2)\%$  [119], gives

$$A_{P,B^+} = (-0.6 \pm 0.6)\% . \quad (8.28)$$

The new PDG average of the  $CP$  asymmetry in  $B^+ \rightarrow J/\psi K^+$  decays includes a recent measurement of the  $CP$  asymmetry by DØ [120],  $A_{CP}(B^+ \rightarrow J/\psi K^+) = (0.59 \pm 0.37)\%$ . Note that the PDG scales the uncertainties by 1.8 to account for the difference in the results from DØ and Belle.

Fig. 8.3 shows the fit projections obtained when using the model reported, to describe

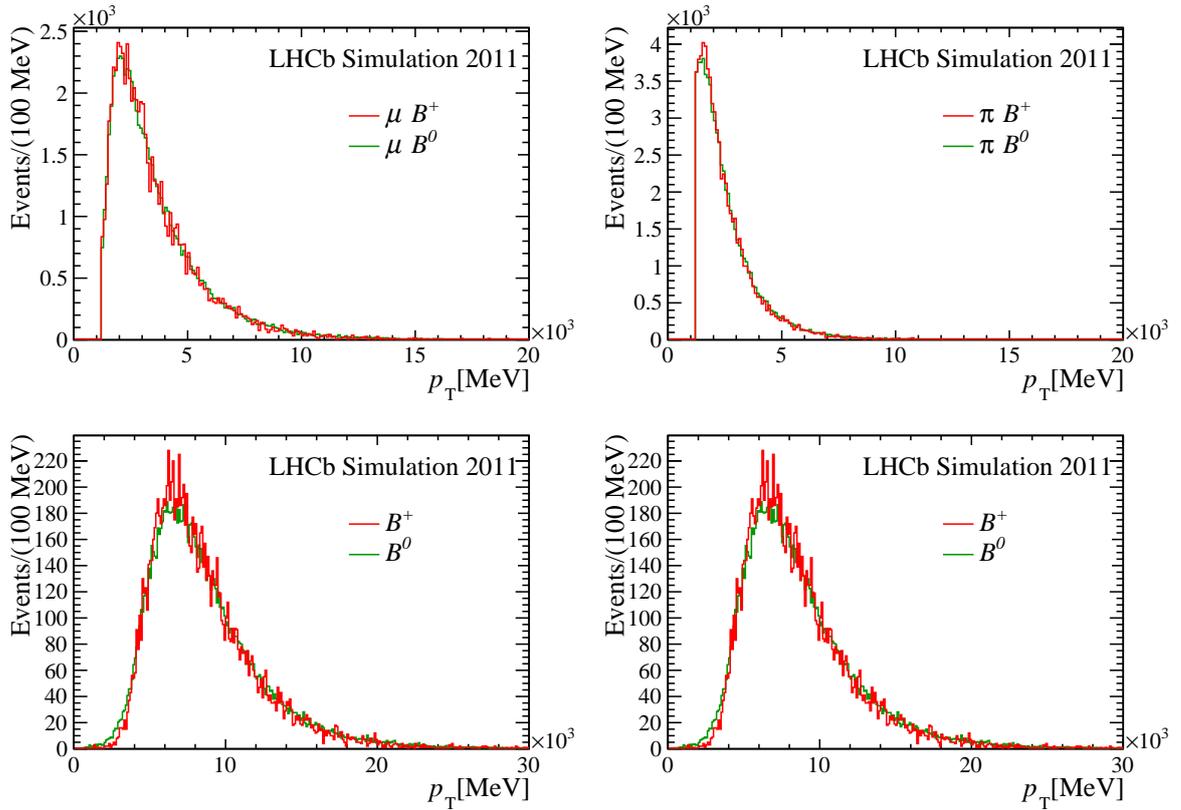


Figure 8.2: Muon and highest  $p_T$  pion transverse momenta spectra, reconstructed  $B$  meson momentum and  $p_T$  for  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^+ \rightarrow D^+ \mu^- X^+$  simulated events with 2011 conditions. The  $\mu\pi$  kinematical weights discussed in Sec 6.5.1 are not applied for these plots.

simulated  $B^+$  events with 2011 conditions and reconstructed in the  $D^+ \mu^-$  mode. The values obtained for the acceptance parameters for all the datasets can be found in Tab. 8.4, together with the other parameters that are not determined with the nominal  $a_{\text{sl}}^d$  fit.

### 8.3.3 $D^\pm/D^0$ mass sidebands decays

In order to understand the proper description of the combinatorial background events and decays not containing a  $D^\pm/D^0$  meson, the regions of the reconstructed  $D^\pm/D^0$  candidate mass far from the peaking structure are probed. These regions at left and right for the peak are so-called “sidebands”. The sidebands regions are defined as  $M(D^0) < 1820$  and  $M(D^0) > 1905$  for the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  mode, and  $M(D^\pm) < 1840$  and  $M(D^\pm) > 1900$  for the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  mode. The mixing asymmetry is commonly defined as the difference between the number decays in which the  $B$  meson changed flavor between the production and the decay and the number of decays where the  $B$  flavor did not change, divided by the sum. The mixing asymmetry plot of the events in the sideband region of the  $D^\pm/D^0$  candidate mass shows an oscillating behavior, as shown in Fig. 8.4. This means

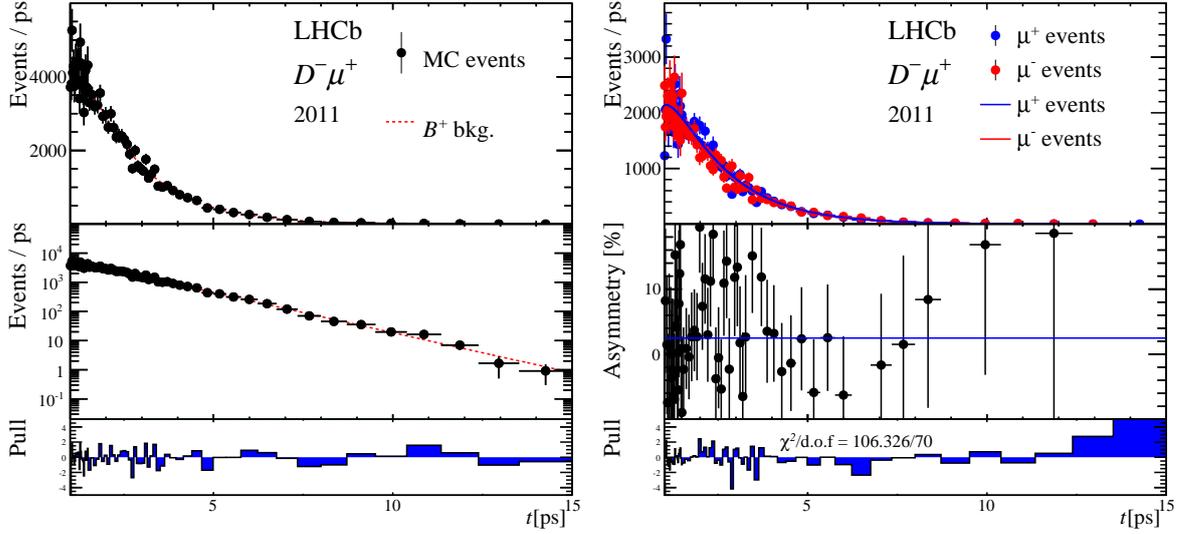


Figure 8.3: Fit to the decay time of  $B^+ \rightarrow D^- \mu^+ X^+$  simulated events with 2011 conditions. *Left*: Decay time distribution for all the events, in linear and logarithmic scale. The last row shows the so-called *Pull* values, i.e. the difference between the number of events observed in data for each bin of the histogram and the number of events predicted by the model in the same bin, divided by the uncertainty on the number of events. *Right*: decay time distribution for  $f$  and  $\bar{f}$  final states shown separately and charge asymmetry of the final state particles. The *Pull* values in the last row refer to the charge asymmetry projection in the second row.

that this background consists of true  $B^0$  meson decays. For this reason, the same decay time model as for the signal is used. No  $CP$  violating effect is parametrized. Any value for  $a_{sl}^d$  would be absorbed into the  $A_{P, sb}$  and  $A_{D, sb}$  parameters, that are free parameters determined by the fit. The PDFs used are

$$\mathcal{P}_{sb}(f, t) = \mathcal{N}_{sb} e^{-t/\tau_{sb}} (1 + A_{D, sb} - A_{P, sb} \cos \Delta m_{d, sb} t) , \quad (8.29)$$

$$\mathcal{P}_{sb}(\bar{f}, t) = \mathcal{N}_{sb} e^{-t/\tau_{sb}} (1 - A_{D, sb} + A_{P, sb} \cos \Delta m_{d, sb} t) , \quad (8.30)$$

where the subscript “sb” denotes the parameter for the sideband component. The effect from both the  $B^+$  production asymmetry and any detection asymmetry in the sidebands is absorbed in the non-oscillating  $A_{D, sb}$  term, while any  $B^0$  production asymmetry is absorbed by the oscillating term. In summary, this is an effective model that does not allow to distinguish the fraction of oscillating background in the sidebands from the  $B^0$  production asymmetry and the effect of detection asymmetries and  $B^+$  production asymmetry. This is perfectly acceptable given that there is no interest to extract any of the parameters describing these effects from the sidebands distribution. A charge asymmetry fit is performed on the sideband data, to verify the accuracy of the model and to determine the values of the acceptance parameters  $\alpha_{sb}$  and  $t_{shift, sb}$ , that can be then fixed in the nominal fit to improve its stability. It is also useful to find proper starting values for the other parameters of the sideband component. The value of the mixing frequency  $\Delta m_{sb}$  is

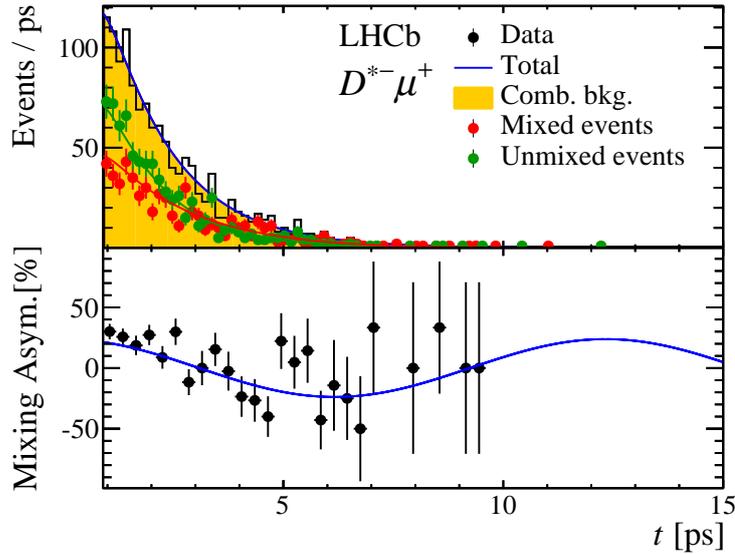


Figure 8.4: Fit to the decay time and mixing asymmetry of the data selected in the sideband region of the reconstructed  $D^0$  candidate mass.

instead fixed from the fit of the mixing asymmetry on data. The fit projections in Fig. 8.5 show the good agreement between the data and the fit model used.

## 8.4 Fit to the $K^+\pi^-\pi^-/K^+\pi^-$ invariant mass

The  $K^+\pi^-\pi^-/K^+\pi^-$  invariant mass distributions for the signal and  $B^+$  decays are described with the sum of a Gaussian and a Crystal Ball function [94] distributions. The latter function is needed in order to model the left tail of the  $K^+\pi^-\pi^-/K^+\pi^-$  invariant mass spectra. A Crystal Ball function consists of a power-law tail added to a Gaussian shape. The parameters  $n_{CB}$  and  $\alpha_{CB}$  describe the steepness of the tail and the transition point between the tail and the Gaussian shape (see Appendix A for the exact parametrization). The Gaussian and the Crystal Ball functions share the same mean and the widths of the two distributions are related to each other by the factor  $s_{12}$ , which is determined by the fit. This parametrization choice reduces the correlation between the parameters describing the width of the peak. The parameter  $n_{CB}$  is fixed in the full fit to a value determined with a preparatory fit to the  $D^\pm/D^0$  mass only, in order to give more stability to the fit.

The sideband background is instead modeled with an exponential function. Fig. 8.6 shows the projection of this  $D^\pm$  mass fits for one of the data samples analyzed.

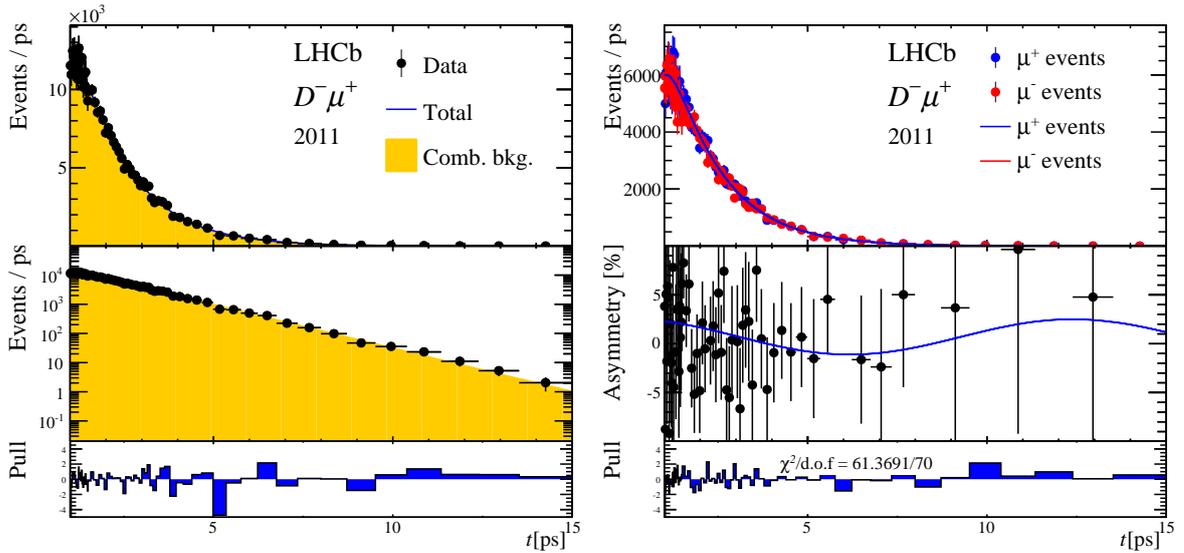


Figure 8.5: Fit to the decay time of  $D^\pm$  mass sidebands events. Data collected in 2011 with down magnet polarity are used. *Left*: Decay time distribution for all the events, in linear and logarithmic scale. The last row shows the so-called *Pull* values, i.e. the difference between the number of events observed in data for each bin of the histogram and the number of events predicted by the model in the same bin, divided by the uncertainty on the number of events. *Right*: decay time distribution for  $f$  and  $\bar{f}$  final states shown separately and charge asymmetry of the final state particles. The *Pull* values in the last row refer to the charge asymmetry projection in the second row.

## 8.5 Fit validation

A well known method to validate a fit procedure is the so-called *pull study* (or *toy study*). The goal is to show that the fit procedure developed provides the correct unbiased results for  $a_{\text{sl}}^d$  and with the correct uncertainty estimate. This means that one has to verify the correctness of the likelihood function maximized in the fit, which includes also the correct treatment of the weighted dataset and the consistency of all the PDFs presented in the previous sections. A set of pseudo-experiment (also called *toys*) is generated according to the PDFs used in the fit and with in general the same statistics of the data sample used in the measurement. Each pseudo-experiment is then fitted with the same nominal fit used for the data. For each parameter involved in the PDFs,  $\theta$ , the difference between the generated value and the estimated value from the fit is calculated. This difference is compared to the uncertainty on the parameter estimated by the fit, and this quantity is denominated *pull*:

$$\text{Pull}(\theta^i) = \frac{\theta_{\text{measured}}^i - \theta_{\text{generated}}^i}{\sigma_{\text{measured}}^i}. \quad (8.31)$$

For each pseudo-experiment a value for the pull for each parameter is obtained. The distribution of the pulls for a given parameter provides the information concerning the

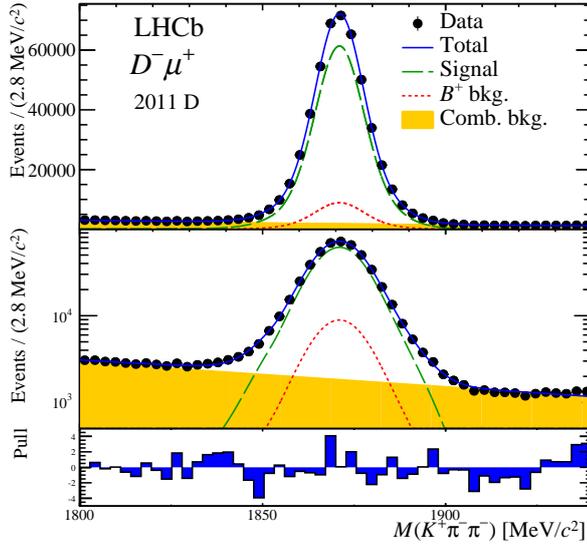


Figure 8.6: The invariant  $K^\mp \pi^\pm \pi^\pm$  mass distributions of the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  magnet-down 2011 data sample, with fit models overlaid. The first two top rows show the invariant mass distributions using a linear and a logarithmic y-axis, respectively. The last row shows the so-called *Pull* values, i.e. the difference between the number of events observed in data for each bin of the histogram and the number of events predicted by the model in the same bin, divided by the uncertainty on the number of events of the bin.

reliability of the estimate of that parameter. The relevant features for a pull distribution are:

- *The shape.* For a well behaved parameter estimation, the pull distribution is expected to exhibit a Gaussian distribution. This might not be the case for example if the parameter of interest is expected to take a value close to the limits of the allowed range for that parameter. In this cases the likelihood used is not considered a good estimator for the parameter considered.
- *The mean value.* The mean value is expected to be zero for an unbiased fit. A mean pull value different from zero indicates that the parameter considered is systematically overestimated or underestimated.
- *The pull width.* For a correct uncertainty estimation in the fit, the width of the pull distribution is expected to be compatible with 1. A smaller (larger) value for the width indicates that the error on the parameter is systematically overestimated (underestimated).

Several toy studies have been performed to check the implementation of the binned maximum likelihood fit and the consistency of each PDF used in the fit. In Fig. 8.7 the pull distributions relevant for the  $a_{\text{sl}}^d$  measurement are shown. A set of 380 pseudo-experiments of 1.1 million events each is generated including all the fit components of

the nominal fit, in both the decay time and  $D^\pm$  mass dimensions. The same weights distribution as used for the data samples in the measurement is used for the toy generation. The statistics of the toys is the same as of the highest statistic data sample considered for the measurement. The pseudo-experiments are subsequently fitted with the nominal fit algorithm, with exactly the same configurations as used for the  $a_{\text{sl}}^d$  measurement. The pull

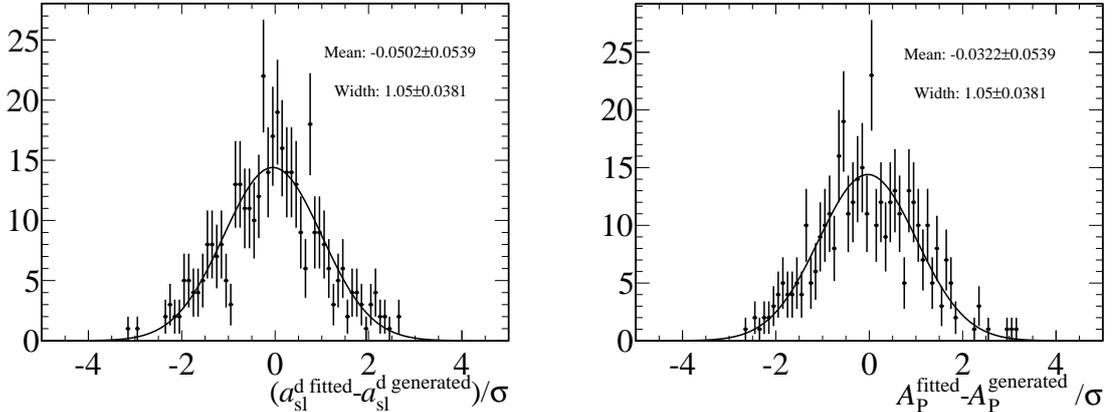


Figure 8.7: Fig:  $a_{\text{sl}}^d$  and  $A_P$  pull distributions for 380 toys accounting for all the components, dimensions, weights of the nominal fits.

distribution obtained for the  $a_{\text{sl}}^d$  parameter, in Fig. 8.7, shows a Normal distribution. This confirms that both the value of the parameter and of the uncertainty on the parameter are reliable. The likelihood used is therefore a good estimator for  $a_{\text{sl}}^d$ . The same conclusions can be drawn for the  $A_P$  parameter, also shown in Fig. 8.7.

## 8.6 The $a_{\text{sl}}^d$ time-dependent fit summary

A multidimensional maximum likelihood fit on a weighted and binned dataset is performed in order to determine the value of  $a_{\text{sl}}^d$ . The dimensions of the fit are the decay time, the charge of the final state and the  $K^+\pi^-\pi^-$  and  $K^+\pi^-$  invariant mass for  $B^0 \rightarrow D^-\mu^+\nu_\mu X$  and  $B^0 \rightarrow D^{*-}\mu^+\nu_\mu X$  respectively.

The reconstructed  $D^\pm/D^0$  candidate mass gives the handle to separate combinatorial background and decays not containing a  $D^\pm/D^0$  meson from the signal. The charge of the final state and the time provide the handle to determine  $a_{\text{sl}}^d$  and  $A_P$ , provided that the detection asymmetry is independently determined and the  $B^+$  component of the data sample is properly constrained from external inputs. The PDFs used to describe the data distribution for each component of the data sample and in each dimension are described in Sec 8.3-8.4.

The decay time distribution for the signal decays is described with the help of a double convolution in order to account for the experimental resolution (see Chapter. 7).

Table 8.3 summarizes the parameters which are determined in the fit (*floating parameters*) and the parameters for which the value is assumed and kept constant (*fixed parameters*).

Table 8.3: Fit parameters summary. The flight distance resolution and the  $k$ -factor distribution, the first described with a triple Gaussian fully fixed from MC and the second with a histogram extracted from MC events, are not included in the table. The upper part of the table contains the parameters which are left free to float (floating), while the second part contains the parameters which are fixed (fixed) in the fit.

Parameter	Description
<b>Floating</b>	
Physics	$a_{sl}^d$ $A_P$
	$CP$ asymmetry in $B^0$ mixing $B^0$ production asymmetry
Time acceptance	$\alpha$ $\beta$
	Lower decay time acceptance Upper decay time acceptance ( $\beta$ -factor)
Sideband background	$f_{sb}$ $\Gamma_{sb}$ $A_{P, sb}$ $A_{D, sb}$
	Fraction Effective lifetime Effective production asymmetry Effective detection asymmetry
Mass model $D^\pm(D^0)$	$m_D$ $\sigma_D$ $s_{12}$ $f_{12}$ $\alpha_{CB}$ $\alpha_{mass}$
	Mass peak position Width $\sigma$ of the Gaussian component Ratio between the $\sigma$ of Gaussian and Crystal Ball Fraction of the Gaussian model component $\alpha$ parameter of the Crystal Ball Slope of the sideband background in $m$ spectrum
<b>Fixed</b>	
Physics	$\Delta m$ $A_D$
	$B^0$ mixing frequency Detection asymmetry
Time acceptance	$t_{shift}$
	Zero crossing at low decay times
Sideband background	$t_{shift, sb}$ $\alpha_{sb}$ $\Delta m_{sb}$
	Zero crossing at low decay times Lower decay time acceptance Mixing frequency
$B^+$ background	$f_{sig}$ $\alpha_{B^+}$ $t_{shift} B^+$ $\Gamma_{B^+}$ $A_{P, B^+}$ $A_{D, B^+}$
	Fraction of signal wrt the fraction of peaking events Lower decay time acceptance Zero crossing at low decay times $B^+$ effective lifetime Production asymmetry Detection asymmetry (same value as for the signal)
Mass model	$n_{CB}$
	$n$ parameter of the Crystal Ball of $D^\pm(D^0)$ mass peak

The values of the fixed parameters are reported in Tab. 8.4. They are determined either by means of fits on simulated data samples or with preparatory simpler fits on data. The systematic uncertainty on  $a_{sl}^d$  due to this parameter choice is calculated and reported in Sec. 9.1.5.

Table 8.4: Values of the parameters fixed in the nominal fit. Physics parameters like  $\Delta m$  are fixed at the same value for the two decay modes and subsamples. Parameters evaluated for the  $B^+$  component on MC simulation are different for the two decay modes, and parameters evaluated with preliminary fits are different for each data sample. The parameters of the sidebands background acceptance are determined with preliminary fits on the data sidebands region only, while the signal  $t_{shift}$  and  $n_{CB}$  parameters are evaluated from preliminary one-dimensional fit on the full data sample. Systematic uncertainties due to this assumptions are considered.

$B^0 \rightarrow D^- \mu^+ \nu_\mu X$				
Parameter	2011 Mag Down	2011 Mag Up	2012 Mag Down	2012 Mag Up
$f_{sig}$	0.873	0.873	0.873	0.873
$t_{shift}$	0.180	0.180	0.000	0.000
$\Gamma_{B^+}$	0.634	0.634	0.612	0.612
$\alpha_{B^+}$	0.522	0.522	0.639	0.639
$t_{shift} B^+$	0.274	0.274	0.000	0.000
$\Delta m_{sb}$	0.510	0.510	0.510	0.510
$\alpha_{sb}$	0.386	0.764	0.480	0.526
$t_{shift, sb}$	0.452	-0.072	0.317	0.230
$\Delta m$	0.510	0.510	0.510	0.510
$A_{P, B^+}$	-0.006	-0.006	-0.006	-0.006
$n_{CB}$	0.486	0.486	0.486	0.486
$B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$				
Parameter	2011 Mag Down	2011 Mag Up	2012 Mag Down	2012 Mag Up
$f_{sig}$	0.912	0.912	0.912	0.912
$t_{shift}$	0.012	0.000	0.000	0.000
$\Gamma_{B^+}$	0.602	0.602	0.602	0.602
$\alpha_{B^+}$	0.200	0.200	0.200	0.200
$t_{shift} B^+$	0.650	0.650	0.650	0.650
$\Delta m_{sb}$	0.510	0.510	0.510	0.510
$\alpha_{sb}$	0.353	0.469	0.782	0.122
$t_{shift, sb}$	0.511	0.503	0.127	0.770
$\Delta m$	0.510	0.510	0.510	0.510
$A_{P, B^+}$	-0.006	-0.006	-0.006	-0.006
$n_{CB}$	3.266	3.266	3.266	3.266

## 8.7 Fit results

In this section, the fit results obtained for each sub-set of the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  data samples are reported. Fig. 8.8 shows the decay time and final state charge fit projections for the sub-sample of  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  decays, acquired in 2011 with magnet polarity down. The mass projection of the same fit is shown in Fig. 8.6. The same projections for the other sub-samples are reported in Appendix B, while the combination

of the data from 2011 and 2012 run periods and magnet polarities with the fit results overlaid are reported in Sec. 10, in Fig. 10.2.

The charge asymmetry projection is particularly interesting in order to judge the quality of the fits. Tab. 8.5 reports the  $\chi^2$  values obtained for this dimension of the fit. They indicate in general a good agreement between the experimental data and the model used. In addition, the runs test (see Sec. 8.1.2) is performed on the pull sequence obtained on these projections. The results are reported in Tab. 8.6 and Tab. 8.7. As a reminder, a “run” or a “sequence” is defined by a number of either positive or negative pulls in bins next to each other. The number of sequences observed on the pulls of the charge asymmetry projection on the data fits, indicated with  $N_{\text{runs}}$  in Tab. 8.6 and Tab. 8.7, should be compared with the expected value  $\mu$  within the uncertainty  $\sigma$  (see Eq. 8.19 and Eq. 8.20). Exception made for a subsample, the number of sequences expected and observed on the pull distributions are in a good agreement. Concerning the overall quality of the binned maximum likelihood fits, Tab. 8.8 reports the  $\chi^2_\lambda$  values (see Sec. 8.1.2) and probabilities. The  $\chi^2_\lambda$  values are not optimal for all the subsets of the data samples. In particular for the highest statistics data samples (as the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  samples), it is common to obtain a not optimal  $\chi^2_\lambda$  value when using a simple model as the one used for the  $a_{\text{sl}}^d$  measurement. In this case it is possible for example that the model used for  $D^\pm/D^0$  mass dimension of the fits is too simple, or the binning is sub-optimal to perfectly describe the data distribution, but as long as the same inaccuracy affects both the final states  $K^+ \pi^- \pi^- \mu^+$  and  $K^- \pi^+ \pi^+ \mu^-$ , the fit is giving a reliable estimate of the  $a_{\text{sl}}^d$  and  $A_P$  parameters. The quality of the charge asymmetry projections (that can be judged from the  $\chi^2$  in Fig. 8.8) guarantees that the effects of sub-optimal description of the single final states are the same. Tab. 8.9 reports the correlations between the parameters obtained for one of the nominal fits. The  $a_{\text{sl}}^d$  parameter does not show correlations of concern with other parameters.

Tab. 8.10 reports the fit results for all the parameters and all the data samples. Only the statistical uncertainties from the fits are reported. Concerning the  $a_{\text{sl}}^d$  and  $A_P$  measured values, the final results for the values and uncertainties are obtained by means of an arithmetic average of the results for each magnet polarity and a weighted average of the results obtained on different center-of-mass energies and decay modes, as explained in detail in Sec. 4.1.5.

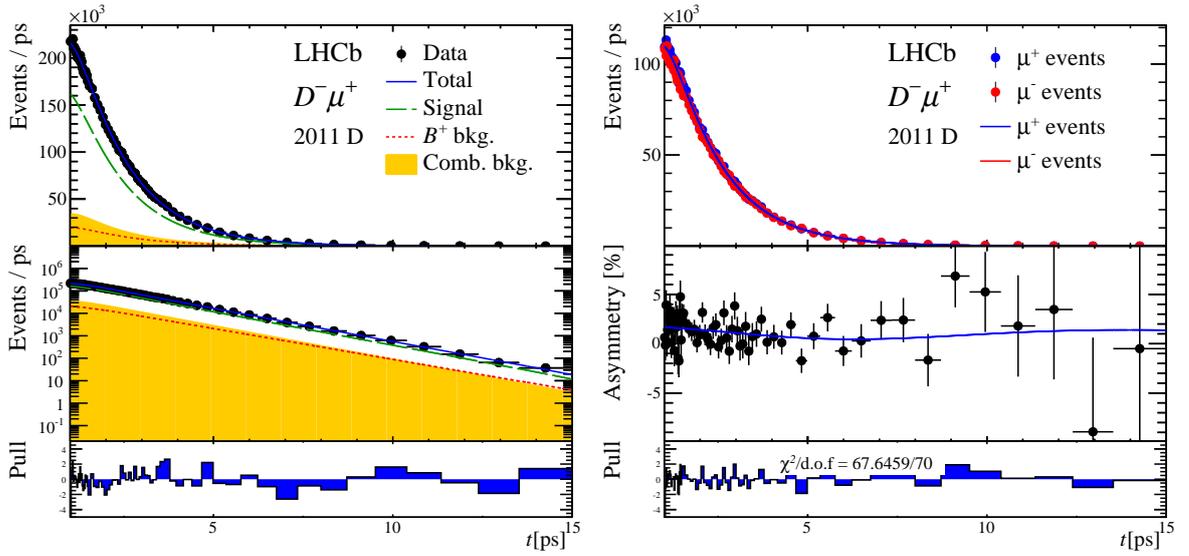


Figure 8.8: Final state charge and decay time projections for the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  decay mode, using 2011 data collected with magnet Down polarity.

Table 8.5: Chi-square test on the charge asymmetry projections of the nominal fits.

Data sample	$\chi^2$	
	$B^0 \rightarrow D^- \mu^+ \nu_\mu X$	$B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$
2012 MagDown	48/70	63/70
2012 MagUp	98/70	62/70
2011 MagDown	68/70	70/70
2011 MagUp	49/70	62/70

Table 8.6: Run test on the charge asymmetry pulls for the nominal fits of the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  sample.

Data Samples	$N_+$	$N_-$	$\mu$	$\sigma$	$N_{\text{runs}}$
2011 U	33	37	35.886	4.139	37
2011 D	37	33	35.886	4.139	37
2012 U	35	35	36.000	4.153	39
2012 D	35	35	36.000	4.153	39

Table 8.7: Run test on the charge asymmetry pulls for the nominal fits of the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  sample.

Data Samples	$N_+$	$N_-$	$\mu$	$\sigma$	$N_{\text{runs}}$
2011 U	30	40	35.286	4.067	36
2011 D	36	34	35.971	4.149	36
2012 U	37	33	35.886	4.139	40
2012 D	35	35	36.000	4.153	38

Table 8.8: Poissonian likelihood chi-square values for the nominal fits. The number of degrees of freedom to be considered for all the fits is 6986, obtained subtracting the 14 floating parameters from the 7000 bins used (50  $D^\pm(D^0)$  mass bins, 70  $B$  decay time bins, and 2 bins for the final state charge).

Data sample	$B^0 \rightarrow D^- \mu^+ \nu_\mu X$		$B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$	
	$\chi_\lambda^2$	Probability	$\chi_\lambda^2$	Probability
2012 MagDown	7999.520	$1.2 \times 10^{-16}$	6589	1
2012 MagUp	7821.535	$4.9 \times 10^{-12}$	6448	1
2011 MagDown	7540.578	$2.3 \times 10^{-6}$	5832	1
2011 MagUp	7711.491	$1.4 \times 10^{-9}$	5503	1

Table 8.9: Correlation matrix for the nominal  $a_{\text{sl}}^d$  fit to the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  2012 Magnet Down data sample.

	$f_{sb}$	$\alpha$	$\beta$	$\Gamma_{sb}$	$\alpha_{mass}$	$a_{\text{sl}}^d$	$A_P$	$A_{P,\text{sb}}$	$A_{D,\text{sb}}$	$\sigma_{D^+}$	$\alpha_{CB}$	$f_{12}$	$s_{12}$	$m_{D^+}$
$f_{sb}$	1.00	-0.01	-	-0.01	0.52	-	-	0.01	0.01	-0.07	0.73	0.13	-0.04	-0.09
$\alpha$		1.00	-0.74	0.03	-0.01	-	-	-	-	-	-0.01	-	-	-
$\beta$			1.00	0.11	-	-	-	-	-	-	-	-	-	-
$\Gamma_{sb}$				1.00	-0.02	-	-	-	-	-0.01	-0.02	0.01	0.02	-
$\alpha_{mass}$					1.00	-	-	0.01	-	0.30	0.76	-0.15	-0.41	-0.07
$a_{\text{sl}}^d$						1.00	0.05	0.11	-0.20	-	-	-	-	-
$A_P$							1.00	-0.16	-0.02	-	-	-	-	-
$A_{P,\text{sb}}$								1.00	-0.49	0.01	0.01	-0.01	-0.01	-
$A_{D,\text{sb}}$									1.00	-0.01	0.01	0.01	0.01	-
$\sigma_{D^+}$										1.00	0.42	-0.94	-0.84	-0.09
$\alpha_{CB}$											1.00	-0.27	-0.51	-0.15
$f_{12}$												1.00	0.63	0.07
$s_{12}$													1.00	0.10
$m_{D^+}$														1.00

Table 8.10: Time dependent fit results for the eight subsamples considered. Only the parameters determined by means of the nominal fit are reported. The uncertainties include only the statistical uncertainty determined by the fit.

$B^0 \rightarrow D^- \mu^+ \nu_\mu X$				
Parameter	2011 Mag Up	2011 Mag Down	2012 Mag Down	2012 Mag Up
$f_{sb}$	$0.167 \pm 0.002$	$0.165 \pm 0.002$	$0.166 \pm 0.001$	$0.169 \pm 0.001$
$\alpha$	$0.495 \pm 0.012$	$0.500 \pm 0.014$	$0.629 \pm 0.010$	$0.625 \pm 0.010$
$\beta$	$-0.051 \pm 0.001$	$-0.047 \pm 0.001$	$-0.041 \pm 0.001$	$-0.041 \pm 0.001$
$\Gamma_{sb}$	$0.691 \pm 0.004$	$0.697 \pm 0.004$	$0.672 \pm 0.002$	$0.673 \pm 0.002$
$\alpha_{mass}$	$0.007 \pm 0.000$	$0.006 \pm 0.000$	$0.007 \pm 0.000$	$0.007 \pm 0.000$
$a_{sl}^d$	$0.005 \pm 0.005$	$-0.019 \pm 0.006$	$0.004 \pm 0.003$	$-0.004 \pm 0.003$
$A_P$	$-0.011 \pm 0.004$	$-0.002 \pm 0.004$	$-0.008 \pm 0.002$	$-0.000 \pm 0.002$
$A_{P, sb}$	$-0.004 \pm 0.010$	$0.010 \pm 0.011$	$0.006 \pm 0.006$	$0.000 \pm 0.006$
$A_{D, sb}$	$0.005 \pm 0.006$	$0.001 \pm 0.007$	$0.006 \pm 0.004$	$0.004 \pm 0.004$
$\sigma_{D+}$	$10.835 \pm 0.156$	$10.967 \pm 0.178$	$10.859 \pm 0.106$	$10.959 \pm 0.098$
$\alpha_{CB}$	$2.749 \pm 0.053$	$2.728 \pm 0.059$	$2.750 \pm 0.034$	$2.830 \pm 0.038$
$f_{12}$	$0.386 \pm 0.017$	$0.381 \pm 0.018$	$0.390 \pm 0.012$	$0.387 \pm 0.011$
$s_{12}$	$0.569 \pm 0.005$	$0.565 \pm 0.006$	$0.577 \pm 0.003$	$0.571 \pm 0.003$
$m_{D+}$	$1871.013 \pm 0.016$	$1871.143 \pm 0.019$	$1871.034 \pm 0.011$	$1871.222 \pm 0.010$
$B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$				
Parameter	2011 Mag Up	2011 Mag Down	2012 Mag Down	2012 Mag Up
$f_{sb}$	$0.031 \pm 0.001$	$0.032 \pm 0.002$	$0.032 \pm 0.001$	$0.031 \pm 0.001$
$\alpha$	$0.559 \pm 0.030$	$0.565 \pm 0.038$	$0.601 \pm 0.021$	$0.590 \pm 0.020$
$\beta$	$-0.056 \pm 0.002$	$-0.050 \pm 0.003$	$-0.044 \pm 0.002$	$-0.045 \pm 0.002$
$\Gamma_{sb}$	$0.764 \pm 0.027$	$0.722 \pm 0.028$	$0.705 \pm 0.015$	$0.667 \pm 0.014$
$\alpha_{mass}$	$0.014 \pm 0.001$	$0.015 \pm 0.001$	$0.014 \pm 0.001$	$0.015 \pm 0.001$
$a_{sl}^d$	$0.026 \pm 0.012$	$-0.013 \pm 0.014$	$0.018 \pm 0.007$	$-0.002 \pm 0.007$
$A_P$	$-0.009 \pm 0.008$	$-0.003 \pm 0.009$	$-0.015 \pm 0.005$	$-0.002 \pm 0.005$
$A_{P, sb}$	$0.008 \pm 0.068$	$-0.130 \pm 0.074$	$-0.018 \pm 0.041$	$-0.024 \pm 0.040$
$A_{D, sb}$	$0.003 \pm 0.042$	$0.029 \pm 0.045$	$0.020 \pm 0.025$	$-0.022 \pm 0.025$
$\sigma_{D+}$	$12.587 \pm 0.489$	$13.633 \pm 0.684$	$13.295 \pm 0.304$	$13.470 \pm 0.389$
$\alpha_{CB}$	$2.027 \pm 0.055$	$2.219 \pm 0.093$	$2.144 \pm 0.045$	$2.125 \pm 0.040$
$f_{12}$	$0.207 \pm 0.030$	$0.174 \pm 0.028$	$0.208 \pm 0.016$	$0.174 \pm 0.017$
$s_{12}$	$0.582 \pm 0.017$	$0.546 \pm 0.022$	$0.556 \pm 0.010$	$0.561 \pm 0.013$
$m_{D+}$	$1866.066 \pm 0.041$	$1866.213 \pm 0.047$	$1866.012 \pm 0.026$	$1866.250 \pm 0.025$

# Chapter 9

## Systematics and Crosschecks

This chapter describes the single contributions to the total systematic uncertainty on the  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  parameters. Moreover, the crosschecks performed in order to ensure the stability and the reliability of the measurement are summarized.

### 9.1 Systematic uncertainties

The  $a_{\text{sl}}^d$  fit model uses several assumptions and is affected by some limitations. Here the impact of these effects is evaluated, and a corresponding systematic uncertainty is assigned to the  $a_{\text{sl}}^d$  measurement.

The different contributions to the overall systematic uncertainty on  $a_{\text{sl}}^d$  are determined by means of *alternative fits*. As general strategy, a pseudo-experiment of 6 million events is generated according to the nominal fit model. The pseudo-experiment is subsequently fitted with a fit model modified according to the systematic effect to be estimated, and with the nominal fit. The difference between the value of  $a_{\text{sl}}^d$  obtained with the modified fit model and the nominal fit result is taken as systematic uncertainty on  $a_{\text{sl}}^d$ . The high-statistic pseudo-experiment is chosen to provide the nominal systematic uncertainties and to eliminate statistical pollution in the evaluation of genuine systematic biases. As crosschecks, the same set of alternative fits is performed on a lower statistics pseudo-experiment and on the data. The systematic uncertainties are evaluated separately for the  $B^0 \rightarrow D^{*-} \mu^+ \nu_{\mu} X$  and  $B^0 \rightarrow D^{-} \mu^+ \nu_{\mu} X$  decay modes, since some effects are expected to have a different impact on the different samples.

The effect of some sources of systematic uncertainties is not properly estimated when using pseudo-experiments, and in this case different approaches are exploited, as is specified in the relevant following sections.

The sources of systematic uncertainties are expected to be uncorrelated. For this reason the single contributions are evaluated separately and the overall systematic error on the physics parameters of interest is obtained by means of sum in quadrature of the single effects. However, it is verified that the leading systematic uncertainties are uncorrelated, as shown in Sec. 9.1.3. A number of variation of the nominal fit have been considered,

Table 9.1: Systematic uncertainty contributions in [%], on  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  at 7 TeV and 8 TeV center-of-mass energy. Contributions below  $10^{-4}$  level are considered negligible (indicated with “-”).

Source of uncertainty	$a_{\text{sl}}^d$	$A_{\text{P}}(7 \text{ TeV})$	$A_{\text{P}}(8 \text{ TeV})$
Detection asymmetry	0.26	0.20	0.14
$B^+$ background	0.13	0.06	0.06
$\Lambda_b$ background	0.07	0.03	0.03
$B_s^0$ background	0.03	0.01	0.01
Combinatorial $D$ background	0.03	–	–
$k$ -factor distribution	0.03	0.01	0.01
Decay-time acceptance	0.03	0.07	0.07
Knowledge of $\Delta m_d$	0.02	0.01	0.01
Quadratic sum	0.30	0.22	0.17

and are described in the following sections.

The largest source of systematic uncertainty in this measurement is the overall uncertainty of the measurement of the detection asymmetry. The second largest systematic is ascribed to the description of the  $B^+$  component of the data sample. The physics backgrounds not included in the fit model lead also to sizable systematic uncertainties. Among these backgrounds the largest source of systematic uncertainty is given by  $\Lambda_b$  decays. Even though the yield of this background is expected to be small, the poor knowledge of the production asymmetry of  $\Lambda_b$  baryons leads to the third largest contribution to the total systematic uncertainty on  $a_{\text{sl}}^d$ . Systematics related to the signal decay time description give a smaller contribution. Tab. 9.1 shows an overview of the systematic uncertainties of this analysis. A more detailed breakdown of the different contributions is provided in Tab. 9.2.

### 9.1.1 Uncertainty on $A_{\text{D}}$

The largest source of systematic uncertainty in this measurement is the uncertainty, statistical and systematic, of the measurement of the detection asymmetry,  $A_{\text{D}}$ . The systematic uncertainty due to the detection asymmetry precision in total is evaluated by means of alternative fits to a high statistics toy MC sample. The data samples used for the  $a_{\text{sl}}^d$  measurement are fitted assuming values of  $A_{\text{D}}$  different from the nominal value by the uncertainty on  $A_{\text{D}}$ . The systematic uncertainty assigned to  $a_{\text{sl}}^d$  for this effect is  $\pm 0.26\%$  ( $0.30\%$ ) in the case of the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  ( $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$ ) data sample. The systematic uncertainty assigned for the  $A_{\text{P}}$  determination at 7 TeV is  $\pm 0.20\%$  ( $0.21\%$ ) when using  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  ( $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$ ) decays, and a systematic uncertainty of  $\pm 0.14\%$  ( $0.17\%$ ) is assigned to the  $A_{\text{P}}$  measurement using  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  ( $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$ ) decays acquired at a center-of-mass energy of 8 TeV.

### 9.1.2 $B^+$ background description

A number of sources of systematic uncertainty can be ascribed to the treatment of the  $B^+$  decays in the fit.

#### Production asymmetry

The value assumed  $A_{P,B^+} = (-0.6 \pm 0.6)\%$  is discussed in Sec. 8.3.2. The systematic uncertainty due to the uncertainty on the assumed value is evaluated. A high statistics pseudo-experiment is generated assuming the nominal value of  $A_{P,B^+}$  and fitted assuming different values for  $A_{P,B^+}$ . The difference between the  $a_{\text{sl}}^d$  determination assuming  $A_{P,B^+} = (-0.6 + 0.6)\%$  and the nominal value is taken as systematic uncertainty. Also the opposite variation ( $A_{P,B^+} = (-0.6 - 0.6)\%$ ) is considered as crosscheck. The same procedure is repeated on the highest statistics data sample. The systematic uncertainty obtained when comparing the nominal and the alternative fits is  $\pm 0.12\%$  on the  $a_{\text{sl}}^d$  parameter and  $\pm 0.06\%$  on  $A_P$ . This is the largest contribution to the systematic uncertainty on  $a_{\text{sl}}^d$  that can be ascribed to the  $B^+$  background description, and the second largest systematic when considering all the effects.

#### Fraction

Since the  $B^+$  background is almost indistinguishable from the signal in the time-dependent fit, it is necessary to fix its fraction and decay time shape from simulation. The estimate of the  $B^+$  fraction is described in Sect. 5.3.1, including its uncertainty, which is propagated to  $a_{\text{sl}}^d$  and  $A_P$ . A high statistics toy is generated assuming the nominal value for the fraction of  $B^+$  decays and fitted with values for the fraction corresponding to the nominal value  $\pm$  its uncertainty. The comparison of the  $a_{\text{sl}}^d$  result of the varied fit with the nominal fit provides a systematic uncertainty of  $\pm 0.03\%$ . From the same comparison of the  $A_P$  result, a systematic uncertainty of  $\pm 0.01\%$  is assigned to the  $A_P$  parameter.

#### Fit model

The  $B^+$  decay time acceptance parameters are fixed to the values obtained from the MC. The statistical uncertainty on these parameters is considered. These parameters are correlated, since they are used in the same model, to describe the same distribution. The most conservative choice in order to evaluate the systematic uncertainty due to this model, is to vary all the parameters by their uncertainty in the direction given by the correlations among them, which is obtained from the fit on the simulated events. Alternative fits assuming varied values for the parameters are used to estimate the systematic uncertainty on  $a_{\text{sl}}^d$  and  $A_P$ , which results to be negligible.

There are two possibilities to describe the decay time distribution of  $B^+$  decays. One option is to use a model similar to the signal model, involving the convolution with the flight distance resolution and the k-factor distribution, and an acceptance model of the form  $(1 - e^{-t-t_{\text{shift}}}/\alpha})(1 + \beta t)$ , as for signal decays (see Eq. 7.22). The other option is

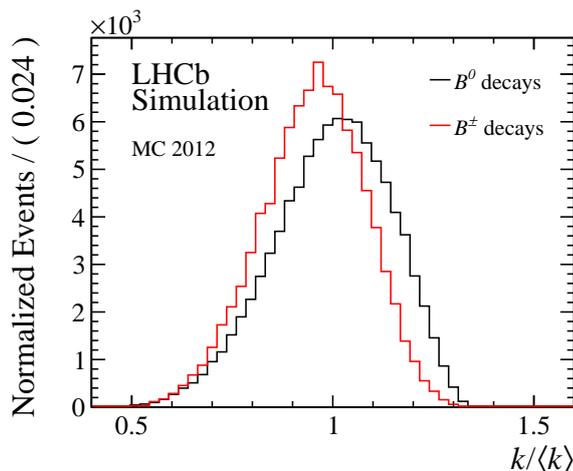


Figure 9.1:  $k/\langle k \rangle$  distributions for simulated  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^+ \rightarrow D^- \mu^+ \nu_\mu X^+$  decays with 2012 conditions. Both distribution are normalized to 1. The  $k/\langle k \rangle$  for  $B^+$  decays is used in the alternative model of  $B^+$  component described in Sec. 9.1.2.

to use a simpler effective model that just excludes the k-factor distribution convolution. Fig 9.1 shows the  $k/\langle k \rangle$  distribution for  $B^+$  decays compared to the  $k/\langle k \rangle$  distribution for signal decays.

The advantage of the first approach is that the lifetime will be the  $B^+$  lifetime with its physical meaning and, in principle, the shape of the distribution is closer to the real  $B^+$  distribution in the data sample. Unfortunately the latter advantage is lost due to the low statistics of the  $B^+$  MC samples and to the possible data/MC differences. In the case of the second approach, the advantage is mainly in terms of computational costs. Given the statistics of the  $B^+$  MC samples available, the two models should be equally good to describe the  $B^+$  background component. The difference on the physical parameters of interest between two fits using the two different models is taken as systematic uncertainty. The systematic uncertainty estimated for  $a_{\text{sl}}^d$  is  $\pm 0.02\%$ , while the uncertainty estimated for  $A_{\text{P}}$  is  $\pm 0.01\%$ .

### 9.1.3 Correlations crosscheck

This section includes a procedure to check the leading systematics uncertainties and their possible correlations. The external measurements of the  $B^+$  production asymmetry and of the detection asymmetries are uncorrelated, but how the two parameters interplay in the fit with all the floating parameters needs to be determined from the data. A so-called *Gaussian-constrained* parameter is used for each of these external values ( $A_{\text{P},\text{B}^+}$  and  $A_{\text{D}}$ ), where the constraints are kept independent since the external measurements are uncorrelated.

A parameter is Gaussian-constrained when it is not fixed at a given value, but it is assumed to take a value, according to a Gaussian distribution with a certain mean and sigma,

within a certain interval <sup>1</sup>.

The systematic uncertainty in  $a_{\text{sl}}^d$  obtained from this fit (the difference in quadrature between the statistical uncertainty obtained in this case and in the nominal fit) is compared to the systematic uncertainties obtained when using the gaussian constraint for a single parameter each time, while fixing the other one. For this check the highest statistics data sample, consisting of  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  decays acquired in 2012, is used. The systematic uncertainty obtained when using Gaussian-constrained parameters for both  $A_{\text{P},\text{B}^+}$  and  $A_{\text{D}}$  is 0.325%, which should be compared to the sum in quadrature of the systematic uncertainty obtained when using the gaussian constraint only for  $A_{\text{P},\text{B}^+}$ : 0.170% , and when using the Gaussian constraint for  $A_{\text{D}}$ : 0.277%.

The uncertainties obtained with the single Gaussian constraint are the same obtained when fixing the parameter at a value different from the nominal one by the uncertainty on the parameter.

The approach of using two Gaussian constraints for the parameters  $A_{\text{D}}$  and  $A_{\text{P},\text{B}^+}$ , and summing in quadrature of the uncertainties obtained by varying the parameters in two separate fits, leads to the same systematic uncertainty. From this it is possible to conclude that the two systematic uncertainties are uncorrelated and the strategy of adding in quadrature the single contributions is justified. No correlation is expected among the sources of systematic uncertainty, exception made for the parameters involved in the  $B^+$  acceptance model. For this latter case the most conservative choice is assumed, as explained in Sec.9.1.2.

### 9.1.4 Other backgrounds

Not all possible types of physics background are included in the nominal fit model. In this section the systematic uncertainty implied by this choice is evaluated.

#### $\Lambda_b$ background component

A potentially dangerous background process is  $\Lambda_b \rightarrow D^- \mu^+ X_n$ , with  $X_n$  any neutral baryonic state. The potential bias on  $a_{\text{sl}}^d$  is twice (See Eq. 4.20) the product of the fraction of  $\Lambda_b$  decays present in the sample,  $f_{\Lambda_b}$ , and the production asymmetry of the  $\Lambda_b$ ,  $2 \cdot f_{\Lambda_b} \cdot A_{\text{P}}(\Lambda_b)$ .

No precision measurement of  $\Lambda_b$  production asymmetry has been finalized yet at LHCb, but an estimate from the measured raw charge asymmetry in  $\Lambda_b \rightarrow J/\psi p^+ K^-$  decays of  $(1.1 \pm 1.0)\%$  in [121] is possible. Assuming the kaon and proton detection asymmetries

$$A_{\text{D}}(K) \equiv \frac{\epsilon(K^-) - \epsilon(K^+)}{\epsilon(K^-) + \epsilon(K^+)} = (-1.0 \pm 0.5)\%, \quad A_{\text{D}}(p) \equiv \frac{\epsilon(p) - \epsilon(\bar{p})}{\epsilon(p) + \epsilon(\bar{p})} = (3.0 \pm 1.0)\%, \quad (9.1)$$

---

<sup>1</sup>A PDF (see Sec. 8.1) including a Gaussian-constrained parameter  $\zeta_G$  is written as  $p(\mathbf{x}|\boldsymbol{\theta}, \zeta_G) = p(\mathbf{x}|\boldsymbol{\theta}, \zeta_G) \cdot \mathcal{G}(\zeta_G)$ . Where  $\mathcal{G}(\zeta_G)$  indicates the gaussian probability density function of the parameter  $\zeta_G$ .

and considering that the measured raw charge asymmetry results from the detection asymmetry of the particle in the final state and the production asymmetry,  $A(\Lambda_b \rightarrow J/\psi p^+ K^-) = A_{rmD}(K) + A_D(p) + A_P(\Lambda_b)$ , the estimate for  $A_P(\Lambda_b)$  is

$$A_P(\Lambda_b) \equiv \frac{N(\Lambda_b) - N(\bar{\Lambda}_b)}{N(\Lambda_b) + N(\bar{\Lambda}_b)} = (-0.9 \pm 1.5)\% \quad (9.2)$$

The background fraction is given by the product of the production fraction, the branching fraction and the selection efficiency.

The  $\Lambda_b$  production fraction,  $f_{\Lambda_b}^{\text{prod}}/(f_u^{\text{prod}} + f_d^{\text{prod}})$ , has been measured in LHCb as function of the transverse momentum [122]. Considering the range of  $p_T$  of the  $B^0$  mesons used in this analysis, the production fraction is estimated to be about 60%.

Relative to the signal, this decay is color suppressed since a  $d\bar{d}$  pair must be formed between the  $c$  and  $u$  quarks. A factor  $1/9$  from the color factor, with a factor of two uncertainty due to the unknown relative phase space, is assumed. LHCb has performed a measurement of a related mode  $\Lambda_b \rightarrow D^0 p \pi^-$  and its branching fraction ratio with respect to  $\Lambda_b \rightarrow \Lambda_c^+ \pi^-$  [123]. Correcting this measurement with the latest value of  $\mathcal{B}(\Lambda_c^+ \rightarrow p^+ K^- \pi^+)$  from Belle [124], and the PDG average for  $\mathcal{B}(D^0 \rightarrow K^- \pi^+)$  [125], the obtained branching ratio is

$$\frac{\mathcal{B}(\Lambda_b \rightarrow D^0 p \pi^-)}{\mathcal{B}(\Lambda_b \rightarrow \Lambda_c^+ \pi^-)} = (14 \pm 1)\%.$$

The selection efficiency, compared to the selection efficiency of the signal samples, is estimated from MC samples of  $\Lambda_b$  and  $B^0$  particles decaying into  $D^+ n \mu \nu$  and  $D^+ \mu \nu$  final states, respectively, with  $D^+ \rightarrow K^- \pi^+ \pi^+$ . Using the same kinematic selection requirements as for the signal sample, the efficiency ratio results to be about 35%. The  $\Lambda_b$  fraction in the selected candidates results to be  $f_{\Lambda_b} = 2_{-1}^{+2}\%$ . Combining the estimate of the fraction with the estimate of the production asymmetry, the correction on  $a_{\text{sl}}^d$  would be  $(-0.04 \pm 0.06_{A_P} \pm 0.03_{f_{\Lambda_b}})\%$ . For this reason a systematic uncertainty of  $\pm 0.07\%$  is assigned to  $a_{\text{sl}}^d$ , and  $\pm 0.035\%$  to  $A_P$ .

### $B_s$ background component

To evaluate this systematic uncertainty a different high statistics pseudo-experiment is generated: in this case an additional  $B_s$  background component is included. As explained in Sec. 5.2.3, about 2% of the data samples is expected to consist of  $B_s^0$  decays, this is the fraction of  $B_s^0$  decays generated in this pseudo-experiment. This component is analogous to the  $B^+$  background component, but with the corresponding production asymmetry value set to zero. The pseudo experiment is fitted with the nominal fit model and with an alternative fit model including the  $B_s$  component. The difference of the  $a_{\text{sl}}^d$  values is 0.03% and the difference of the  $A_P$  values is 0.01%. These differences are taken as systematic uncertainties.

## $B \rightarrow D^{(*)-} D_s^+$ background decays

A small contribution from  $B \rightarrow D^{(*)-} D_s^+$  decays, with the  $D_s$  decaying semileptonically needs also to be taken into account (see Sec. 5.2.3). Decays as  $B \rightarrow D^{(*)} \tau X$  have a  $k$ -factor distribution very similar to  $B \rightarrow D^{(*)-} D_s^+$  decays. Therefore, the effect of this contribution is estimated by enlarging the contribution of  $B \rightarrow D^{(*)} \tau X$  decays by 2% in the  $k$ -factor distribution used for the alternative fit. The systematic uncertainty results to be negligible for both  $a_{\text{sl}}$  and  $A_{\text{P}}$  parameters.

### 9.1.5 Signal description

These uncertainties are evaluated by performing alternative fits, each accounting for a specific systematic effect, describing the same pseudo-experiment sample. The effects here investigated involve the decay time or the  $D^-/\bar{D}^0$  variables, that are expected to be mostly uncorrelated to the charge of the final state particles. The effect of not using an optimal fit model, is expected to have the same impact on one final state and on the  $CP$  conjugated.

#### The $k$ -factor model

The  $a_{\text{sl}}^d$  fit is relying on MC simulation, for the  $k$ -factor correction and resolution. Possible mis-modeling of both the production and decay models are considered in the following.

The production model can be validated by studying fully reconstructed  $B$  decays as  $B^+ \rightarrow J/\psi K^+$ . A data versus MC comparison of this decay mode, allows for extracting a set of kinematic weights, which are shown in Fig. 9.2. These weights can be applied to the semileptonic signal MC sample, and the effect on the  $k/\langle k \rangle$  distribution can be seen in Fig. 9.2. This varied  $k/\langle k \rangle$  distribution is used in an alternative fit, and the difference of the  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  estimation with respect to the nominal value is negligible. Therefore no systematic uncertainty is assigned for this effect.

The decay model clearly affects the  $k/\langle k \rangle$  distribution, for example the relative branching fractions of the decays with additional particles (both resonant and non resonant). In the  $D^\pm$  channel, after all selection criteria applied, about 44% of the decays are  $B^0 \rightarrow D^\pm \mu^\mp \nu_\mu$ , while the other decays included are mainly  $B^0 \rightarrow D^{*\pm} \mu^\mp \nu_\mu$ , or happening through a higher resonance  $D^{0*}, D_1', D^{**}$ , instead of the  $D^{*\pm}$ . These decays also contribute for about 44% to the MC sample. A histogram of these different contributions is shown in Fig. 9.3. A small contribution come from decays through non-resonant states  $D^\pm \pi^0$ ,  $D^\pm \pi^0 \pi^0$ ,  $D^\pm \pi^+ \pi^-$  or  $D^{*\pm} \pi^0$ ,  $D^{*\pm} \pi^0 \pi^0$ ,  $D^{*\pm} \pi^+ \pi^-$  (6% of the full sample) and decays of the  $D^\pm$  or of the other resonances quoted before in  $\tau \nu$  (the last 6% of the sample). The uncertainties on the fractions of the different decays included can be calculated given the PDG uncertainty on the branching fraction and the fraction of events actually present in the sample. Concerning the main decay chain ( $B^0 \rightarrow D^\pm \mu^\mp \nu_\mu$ ), the uncertainty on the 44% fraction quoted above is about 2.5% (absolute). The uncertainty on the

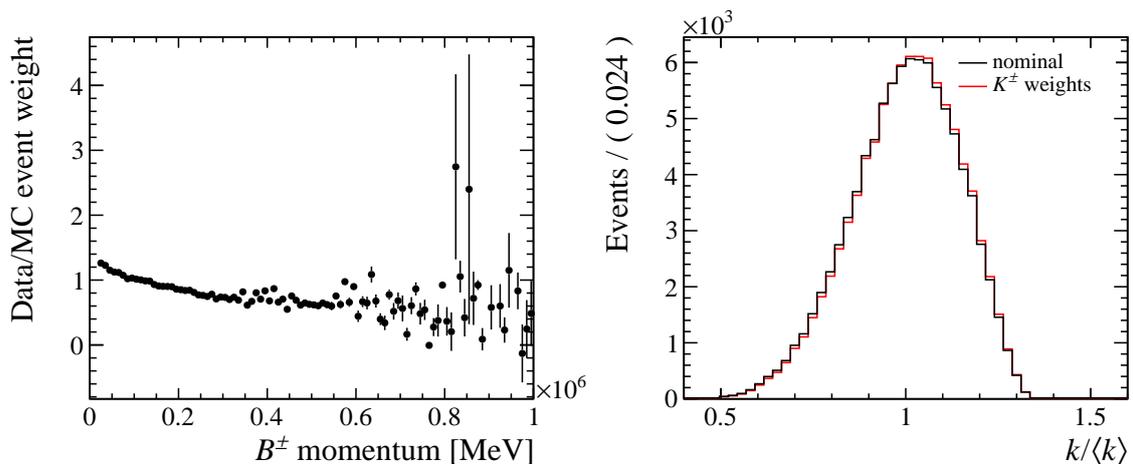


Figure 9.2: *Left*: Weights for data/MC versus  $p(B^+)$  obtained from the  $B^+ \rightarrow J/\psi K^+$  decay mode. *Right*: different  $k/\langle k \rangle$  distribution obtained when weighting the  $p(B^0)$  according to the data/MC differences observed in the reconstructed  $p(B^+)$  distribution with  $B^+ \rightarrow J/\psi K^+$  decay mode.

$B^0 \rightarrow D^{*\pm} \mu^\mp \nu_\mu$  fraction is instead about 2%. The systematic uncertainty is evaluated by splitting the MC sample into these two components, and varying their relative fraction by  $\pm 10\%$ . While this variation is larger than the uncertainty from the branching fraction ratios, the effect on  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  is small. The different  $k/\langle k \rangle$  distributions are shown in Fig. 9.4. This more conservative variation ( $\pm 10\%$ ) is considered, for instance to include the uncertainties on the decay modes not considered in this approximation and decay modes not included in the MC cocktail sample. The effect of this variation on the  $a_{\text{sl}}^d$  fitted value is  $\pm 0.02\%$ , and the effect on the  $A_{\text{P}}$  fitted value is  $\pm 0.01\%$  as reported in Table 9.2. Fig. 9.4 shows the  $k/\langle k \rangle$  distributions for the main decay chain contained in the MC sample ( $B^0 \rightarrow D^\pm \mu^\mp \nu_\mu$ ), compared to the  $k/\langle k \rangle$  distribution for the events of all the other decays contained in the MC sample, including the contributions mentioned before. For completeness, the uncertainty on the non-resonant fraction of the sample is about 3.2% and the uncertainty on the  $B \rightarrow D^{(*)} \tau X$  decays fraction is about 1%.

For the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  sample, the MC cocktail composition is dominated by the main decay chain. The  $\pm 10\%$ <sup>2</sup> variation of the abundance of this component of the MC cocktail relatively to the other components causes negligible variations of the  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  results. Finally, the effect of the time-dependence of the  $k$ -factor is considered by varying  $\Delta m_d$  with  $0.003 \text{ ps}^{-1}$  as explained in Sect. 7.2.2. The effect on the determination of  $A_{\text{P}}$  is negligible, while an uncertainty of  $\pm 0.01\%$  is assigned to the  $a_{\text{sl}}^d$  parameter.

<sup>2</sup>The  $\pm 10\%$  variation is even more conservative in this case, the systematic uncertainty estimated is nevertheless negligible.

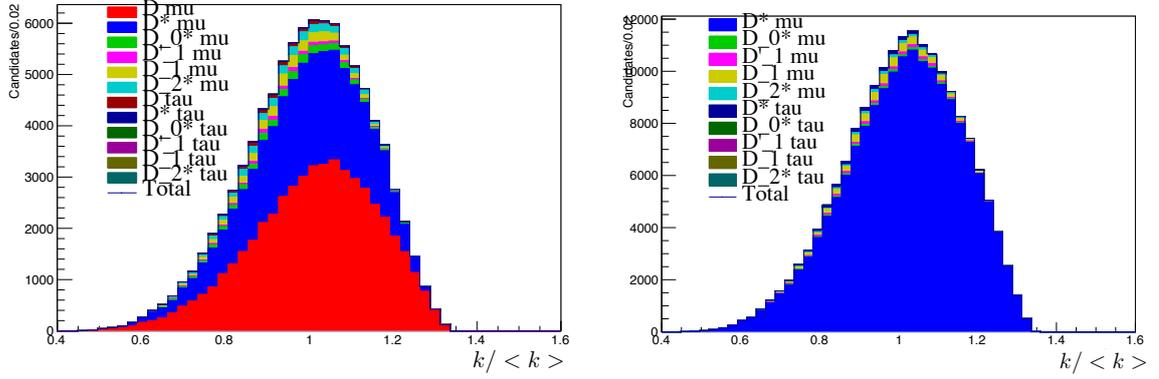


Figure 9.3: *Left*: MC cocktail composition for  $D^\pm$  MC sample. Decays happening through non-resonant states (for example  $D^\pm\pi^0$ ,  $D^\pm\pi^0\pi^0$ ) are not separated. *Right*: same composition plot for the  $D^*$  sample.

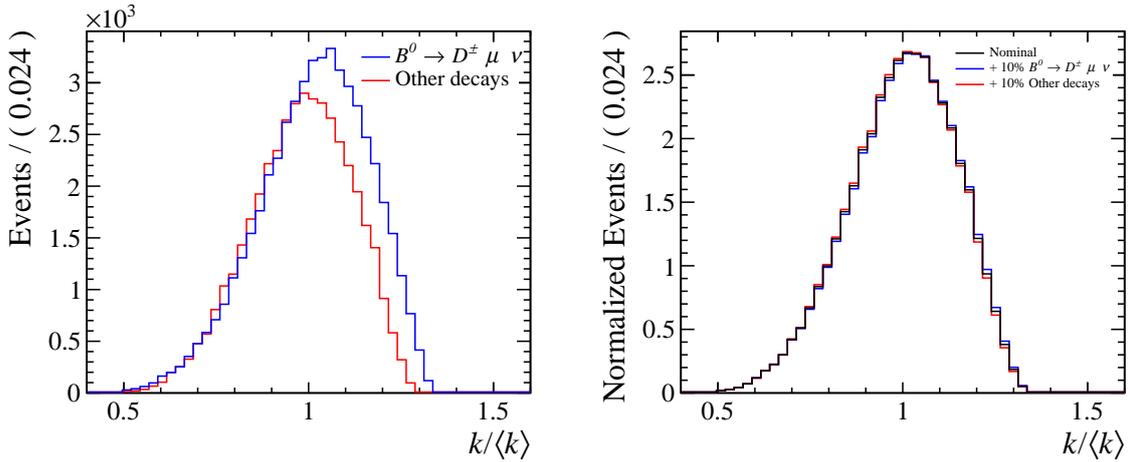


Figure 9.4: *Left*: Sample of simulated  $k/\langle k \rangle$  distributions for  $B^0$  mesons decaying according to the main decay chain ( $B^0 \rightarrow D^\pm\mu^\mp\nu_\mu$ ) in blue and all the other decay chains in red. *Right*: in blue the main decay chain ( $B^0 \rightarrow D^\pm\mu^\mp\nu_\mu$ ) component has been weighted to be 10% more than in the standard cocktail, and the other component 10% less, while in red the same weights are applied but on the opposite direction. The distribution obtained exhibits only small differences when compared to the nominal.

### Knowledge of $\Delta m_d$

In order to determine  $a_{\text{sl}}^d$  and  $A_P$  from the fit, it is necessary to assume a value of  $\Delta m_d$ . This assumption is made for both the signal and the sidebands background models, as explained in Chapter 8. In the nominal fits the value of  $\Delta m_d$  is fixed to the PDG average, for the alternative fits the value of  $\Delta m_d$  is constrained at the PDG average  $\pm$  its uncertainty.

Comparing the nominal fit results, to the alternative fit results, a systematic uncertainty of  $\pm 0.02\%$  is assigned to  $a_{sl}^d$  and a systematic uncertainty of  $\pm 0.01\%$  is assigned to  $A_P$ .

### Signal acceptance model

As explained in Sect. 7.2.2, the decay time fit interval is limited to times above 1.0 ps, to avoid the need for a more complicated acceptance function, with a negligible loss in statistical sensitivity to  $a_{sl}^d$ . The shifts in  $a_{sl}^d$  and  $A_P$ , with a fit range down to 0.4 ps are considered as contributions to the systematic uncertainty. The value of 0.4 ps is chosen, since at lower decay times an even more complicated acceptance model is needed to ensure a good description of the distribution. The effect of extending the  $B$  decay time fit interval to 0.4 ps has a negligible impact on  $a_{sl}^d$ , therefore no systematic uncertainty is assigned. On the other hand it has a large effect on the determination of  $A_P$ . A systematic uncertainty of  $\pm 0.07\%$  is assigned to  $A_P$ . For a precise determination of the production asymmetries as function of transverse momenta and pseudo-rapidity, improvements of the  $B$  decay time acceptance description at low times are needed.

### Flight distance resolution

An alternative fit assuming a flight distance resolution wide twice as much as the nominal one is used to access this effect. The results of the alternative fit and nominal fit are compared and the difference is taken as systematic. Since the effect is found to be negligible, no systematic uncertainty is assigned.

### Mass model

The shape of the  $D^+$  ( $D^0$ ) mass is described by default by a model including a Gaussian and a Crystal Ball. The model used in the alternative fit is a double Gaussian. The alternative fit is performed on the high-statistic pseudo-experiment and on the data. Since the effect is found to be negligible, no systematic uncertainty is assigned.

### Parameters that are fixed in the nominal fit

Few parameters are fixed in the nominal fit to values obtained from preliminary unidimensional fits. For example the  $n$  parameter of the crystal ball shape used for the mass distribution and the acceptance parameter  $t_{shift}$  for the signal decay time description are fixed from preliminary one-dimensional fits to data. Alternative fits are performed fixing those parameters at values  $\pm 1\sigma$  obtained with those fits. The only non-negligible effect found has to be ascribed to the  $n_{CB}$  parameter of the signal  $D^+$  ( $D^0$ ) candidate mass description. A systematic uncertainty of  $\pm 0.03\%$  is assigned to the  $a_{sl}^d$  parameter for this reason.

## Quadratic acceptance at large decay times

It has been observed a quadratic dependence of the upper decay time acceptance, of the type  $(1 + \beta t + \gamma t^2)$ , in decay modes as  $B_s \rightarrow J/\psi\phi$  at LHCb [126]. A pseudo-experiment is generated with this acceptance shape and fit once with the same model and once with the nominal acceptance model. The difference in  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  values contributes to the systematics. Fig. 9.5 compares the nominal and alternative fit projections, and shows how this effect in spite of being clearly visible in the decay time description, has a small impact on the charge asymmetry. The values used for the generation of the pseudo experiment are  $\beta = 0.003$  and  $\gamma = -0.002$ . The measured values on LHCb  $B_s \rightarrow J/\psi\phi$  MC are  $\beta = 0.0032 \pm 0.0018$  and  $\gamma = -0.0023 \pm 0.0002$ , obtained with the techniques explained in Ref. [127]. A systematic uncertainty of  $\pm 0.03\%$  is assigned to  $a_{\text{sl}}^d$ , and  $\pm 0.01\%$  to  $A_{\text{P}}$ .

## Binning choice

An alternative fit is performed in order to evaluate the effect of the binning choice in the decay time dimension of the fit. The same high statistic pseudo-experiment, generated on event-by-event basis, i.e. without assuming any binning, is fitted using the nominal dynamical binning scheme (70 bins in the decay time dimension) and a uniform binning including 200 bins in the decay time dimension. The difference between the measured values of  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  is assigned as systematic uncertainty. Fig. 9.6 shows that the binning choice has a negligible effect on the charge asymmetry projections.

## 9.2 Summary of the systematic uncertainties

Tab. 9.2 summarizes the individual contributions to the systematic uncertainties on  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  at 7 TeV and 8 TeV. The sources of systematic uncertainties are assumed to be uncorrelated, therefore the total systematic uncertainty on the parameters is calculated as the sum in quadrature of the single contributions. In the case of the leading systematic uncertainties, due to the uncertainty on  $A_{\text{D}}$  and on the  $B^+$  production asymmetry, it is verified that the hypothesis of having uncorrelated systematic effects holds (see Sec. 9.1.3). The systematic uncertainties that result to have different values for  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decay modes, are quoted separately in Tab. 9.2. The leading systematic is ascribed to the uncertainty, statistic and systematic, on the measured detection asymmetry, while the second large systematic in this analysis is due to the  $B^+$  background. Systematics related to the decay time description give a smaller contribution.

## 9.3 Crosschecks

Numerous crosschecks are performed in order to ensure the stability and consistency of the result. Relevant plots and details are reported in Appendix. C.

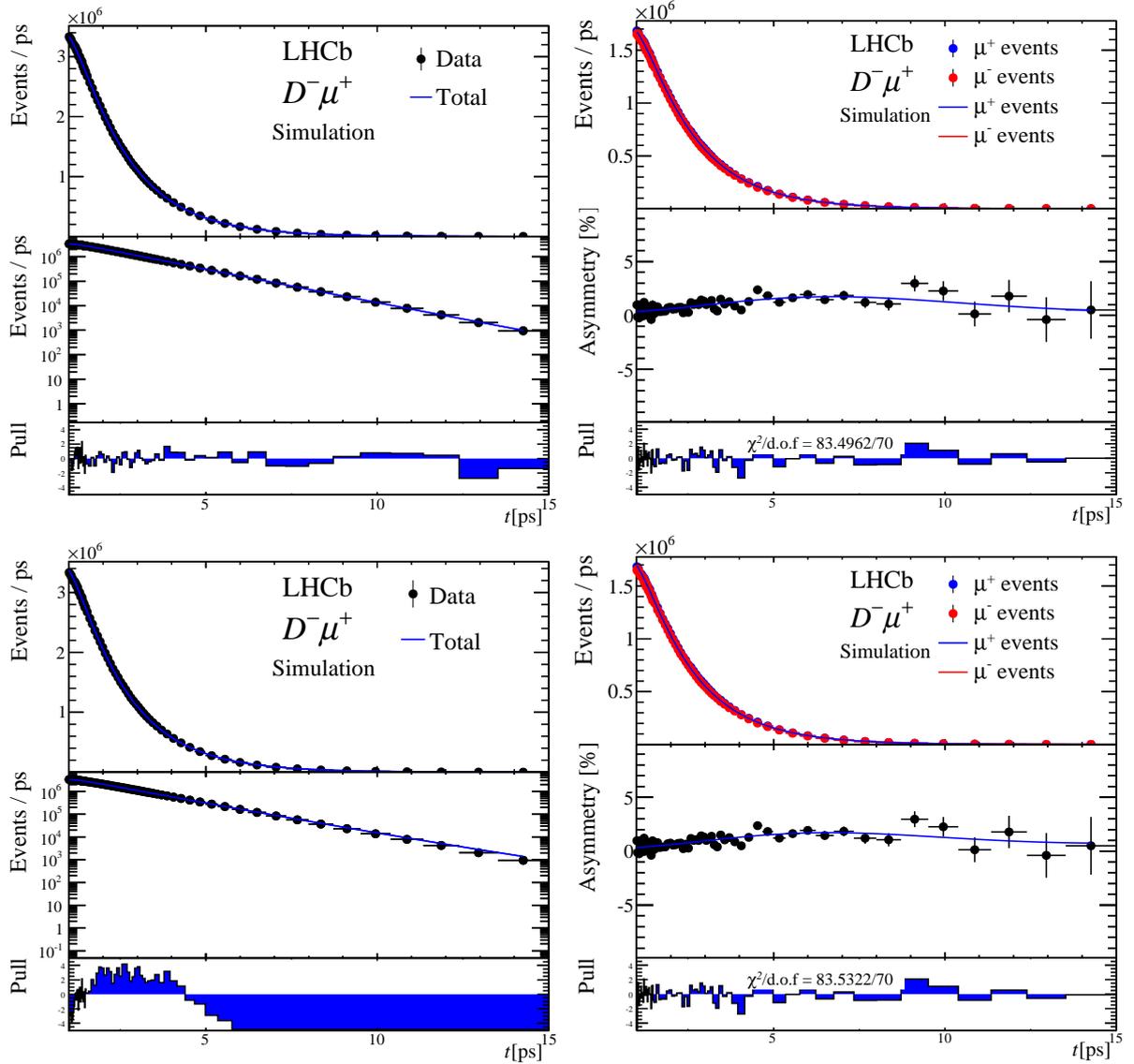


Figure 9.5: High statistics pseudo-experiment generated with a quadratic acceptance model and fitted with quadratic acceptance shape (*top*) and the nominal PDF used for the  $a_{\text{sl}}^d$  fits (*bottom*). Decay time projections (*left*) and charge asymmetry projections (*right*). The quadratic shape is modeled with realistic parameters (see Sec. 9.1.5), hence the effect is evident only in the pull distributions. In spite of the evident effect in the decay time projection, the effect has negligible impact on the charge asymmetry.

- The dataset for each decay mode is split by year of acquisition and magnet polarity. By doing this we observe a  $2.2 \sigma$  discrepancy in the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  decay mode ( $2.0 \sigma$  in  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays) in the measurements of  $a_{\text{sl}}^d$  in 2011 between the two magnet polarities, as it can be seen from Fig.9.7 Due to the detection asymmetry, which is evaluated on the same control samples, these numbers are

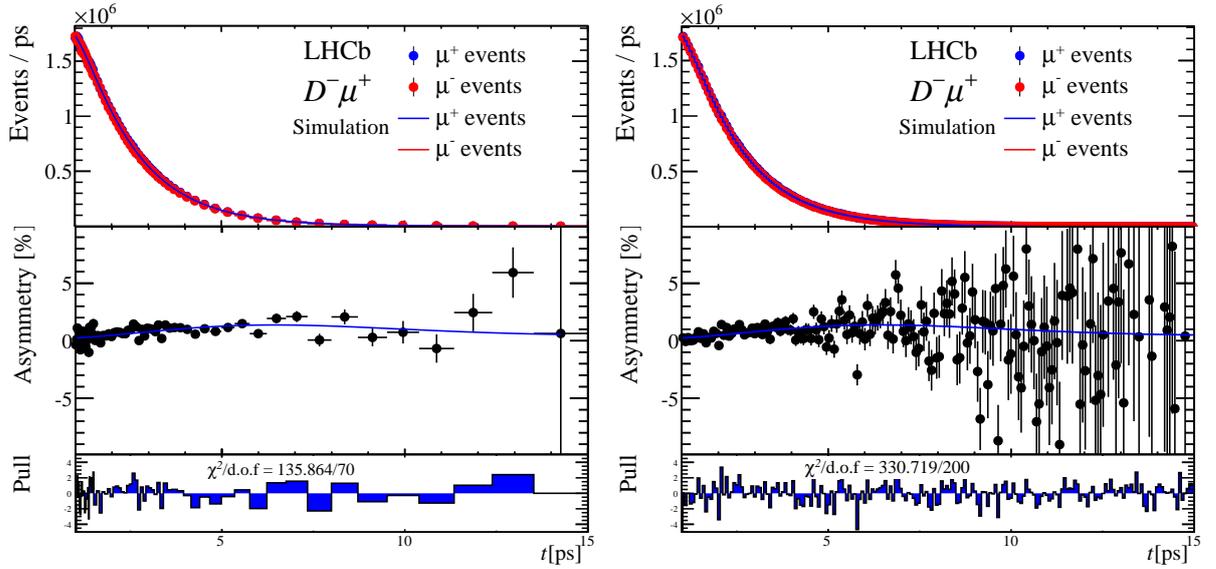


Figure 9.6: Fits with different binning schemes of the same 6 million events pseudo experiment generated according to the unbinned PDFs.

very correlated. For 2012 data instead the results are perfectly compatible. No discrepancy is observed in the measured values of  $A_P$ . Further details are given in Appendix. C.

- *The dataset for  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  is split by periods of data taking.* 2011 data are split in 4 run blocks while 2012 data are split in 3 different run blocks according to the data taking periods. Also the different magnet polarities are separated. No particular structure is observed, the  $a_{sl}^d$  and  $A_P$  results are compatible.
- *The  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  dataset is split according to the number of Primary Vertices.* Events with 1, 2 and  $>2$  PVs are separately analyzed. The measured values of  $a_{sl}^d$  and  $A_P$  are stable.
- *A different parametrization for the fit is used.* Further details are reported in Appendix C. This check shows that the raw asymmetry values are quite stable for the different sub samples of the dataset.
- *Fiducial cuts.* Two fiducial cuts are applied to the data sample in order to exclude the regions where (1) either positive or negative muons are bent out of the acceptance of the detector (2) a large raw asymmetry between positive and negative muons is observed. In spite of variations on the results per year and polarity, the final value of  $a_{sl}^d$  is compatible with the nominal within less than  $0.5\sigma$  (statistical uncertainty only from the fits).
- *Random pion background.* In the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decay mode, there is a small amount of background (percent level) of random slow pions associated with a real

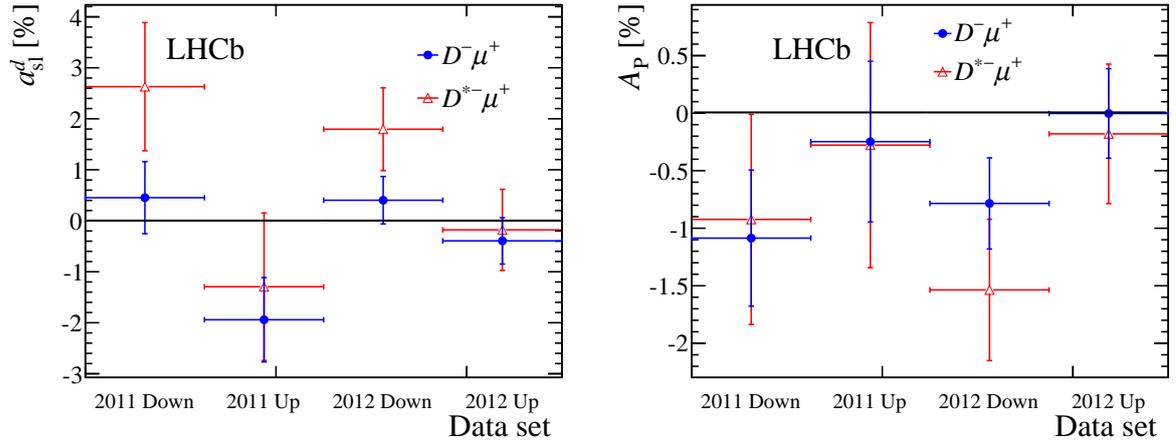


Figure 9.7: *Left*  $a_{sl}^d$ , *Right*  $A_P$ . Results on samples divided by year of acquisition and magnetic field polarity are compared. In blue the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  results, in red the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  results. The statistical uncertainty on the single fit is added in quadrature to the uncertainty due to the statistical component of the uncertainty on the detection asymmetry measurements.

$D^0$  candidate. This type of background is not described in the nominal fits. To understand the size of this effect on  $a_{sl}^d$  and  $A_P$ , an alternative fit on data is performed. From the data, the fraction of this component is extracted, and the time behavior is determined with a simple fit, using the same model used for the  $D^0$  sidebands. The effect on the final results of  $a_{sl}^d$  and  $A_P$  is negligible.

Table 9.2: Overview of all contributions to the systematic uncertainty on  $a_{\text{sl}}^d$  and  $A_{\text{P}}$ . Systematics evaluated with a high statistics toy. Contributions below  $10^{-4}$  level are considered negligible (indicated with "-"). Differences to the  $B^0 \rightarrow D^{*-}\mu^+\nu_{\mu}X$  sample are given in parentheses. individual detection asymmetries are multiplied by two to estimate the effect on  $a_{\text{sl}}^d$ . The actual effect on  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  is estimated with toys for the total detection asymmetry.

Source of uncertainty	$a_{\text{sl}}^d$	$A_{\text{P}}$ (7 TeV)	$A_{\text{P}}$ (8 TeV)
<b><math>B^+</math> background:</b>			
$B^+$ production asymmetry	0.12	0.06	0.06
$B^+$ fraction	0.03	0.01	0.01
$B^+$ acceptance	-	-	-
$B^+$ decay time model	0.02	0.01	0.01
<b>Other backgrounds:</b>			
$B_s^0$ component	0.03	0.01	0.01
$A_b^0$ component	0.07	0.03	0.03
$B^0 \rightarrow D^{(*)-}D_s^+$ decays	-	-	-
<b><math>k</math>-factor distribution:</b>			
Momentum distributions	-	-	-
Branching ratio from $D^{(*)}$ and $\tau$	0.02 (-)	0.01 (-)	0.01 (-)
Time dependence	0.01	-	-
<b>Knowledge of <math>\Delta m_d</math>:</b>			
for signal (true $B^0$ ) ( $\pm 1\sigma$ )	0.02	0.01	0.01
for sideband background ( $\pm 1\sigma$ )	-	-	-
<b>Other fit related systematics:</b>			
binning	-	-	-
mass model	-	-	-
flight distance resolution	-	-	-
B decay time fit interval	-	0.07	0.07
quadratic acceptance	0.03	0.01	0.01
$t_{\text{shift}}$ variation	-	-	-
$n_{CB}$ variation	0.03 (-)	-	-
<b>Detection asymmetry:</b>	0.26 (0.30)	0.20 (0.21)	0.14 (0.17)
<b><math>K - \pi</math> detection asymmetry</b>			
Statistical uncertainty	$\sim 0.16$ ( $\sim 0.18$ )	$\sim 0.15$ ( $\sim 0.14$ )	$0.09$ ( $\sim 0.10$ )
Systematic uncertainty	$\sim 0.13$ ( $\sim 0.20$ )	$\sim 0.07$ ( $\sim 0.10$ )	$\sim 0.07$ ( $\sim 0.10$ )
<b>Muon detection asymmetry</b>			
Statistical uncertainty	$\sim 0.06$	$\sim 0.06$	$\sim 0.03$
Systematic uncertainty	$\sim 0.06$	$\sim 0.03$	$\sim 0.03$
<b>Pion detection asymmetry</b>			
$\mu - \pi$ tracking asymmetry	$\sim 0.07$	$\sim 0.04$	$\sim 0.04$
$\mu - \pi$ tracking asymmetry	$\sim 0.04$	$\sim 0.02$	$\sim 0.02$
Quadratic sum	0.30 (0.34)	0.22 (0.23)	0.17 (0.20)



# Chapter 10

## Results and Conclusions

The  $CP$ -violating asymmetry in  $B^0-\bar{B}^0$  mixing is measured in untagged semileptonic  $B$  decays using data collected by the LHCb experiment in Run-I. The charge asymmetry is determined as a function of the decay time, which allows to separate the contribution from the production asymmetry. The detection asymmetries are subtracted to obtain the  $CP$ -violating mixing asymmetry, which is found to be

$$a_{\text{sl}}^d = (-0.02 \pm 0.19 (\text{stat}) \pm 0.30 (\text{syst}))\% ,$$

where the first uncertainty is statistical and second systematic. This is the most precise  $a_{\text{sl}}^d$  measurement from a single experiment to date, and the result is compatible with the Standard Model prediction. An overview of the current measurements of the semileptonic asymmetries  $a_{\text{sl}}^d$  and  $a_{\text{sl}}^s$  is shown in Fig. 10.1. The most recent measurements of the single asymmetries do not have the sensitivity necessary to exclude the measurement of the di-muon semileptonic asymmetry reported by the DØ experiment [8], which was found to be in about  $3\sigma$  tension with the Standard Model. Larger data samples and novel strategies to control the detection asymmetries to sub-per-mille level are needed to challenge the Standard Model, and confirm or disprove the DØ measurement.

The effect of  $B^0-\bar{B}^0$  production asymmetry is determined by means of a time-dependent fit, simultaneously to  $a_{\text{sl}}^d$ . The fit projections for the combined datasets of  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays are shown in Fig.10.2. The production asymmetry of  $B^0$  and  $\bar{B}^0$  mesons with transverse momentum and pseudo-rapidity in the range  $2 < p_T < 30$  GeV and  $2.0 < \eta < 4.8$  at center-of-mass energy of 7 TeV and 8 TeV are determined

$$A_P(7 \text{ TeV}) = (-0.66 \pm 0.26 (\text{stat}) \pm 0.22 (\text{syst}))\% .$$

$$A_P(8 \text{ TeV}) = (-0.48 \pm 0.15 (\text{stat}) \pm 0.17 (\text{syst}))\% .$$

These values are not corrected for reconstruction efficiencies depending on  $p_T$  and  $\eta$ . The measured value at 7 TeV is compatible with the previous result at LHCb exploiting fully reconstructed decays [128] and it is compatible with the measured value at 8 TeV.

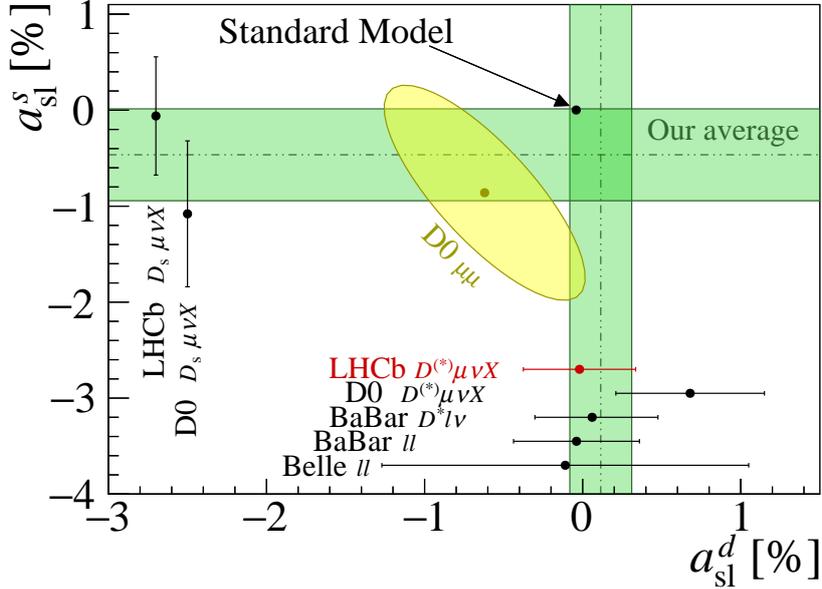


Figure 10.1: Overview of the semileptonic asymmetries measurements. The measurement presented in this thesis is indicated in red. The bands correspond to the average of the pure  $a_{\text{sl}}^d$  and  $a_{\text{sl}}^s$  measurements, which excludes the  $D\bar{0}$  di-muon result. The  $a_{\text{sl}}^d$  measurement presented in this thesis and the recent result from the BaBar collaboration [45] are included in the average.

The analysis presented in this thesis is the first measurement of the  $a_{\text{sl}}^d$  parameter performed by the LHCb experiment. Novel techniques were pioneered in order to deal with the experimental challenges and achieve this result. These techniques include methods to evaluate detection and reconstruction asymmetries at the per-mille level, and the development of a time-dependent analysis strategy for partially reconstructed decays. Reliable methods to determine precise detection and reconstruction asymmetries allow for precision  $CP$  violation measurements, even exploiting the future higher statistics datasets. The  $a_{\text{sl}}^d$  measurement is limited by the systematic uncertainties, that include, as leading contribution, the statistical and systematic uncertainties of the measured values of the detection asymmetries. The high statistics data samples that the LHCb upgraded experiment is going to collect, will allow to reduce the uncertainties on the detection asymmetries measurements and will reduce drastically the statistical uncertainty of the  $a_{\text{sl}}^d$  parameter. In addition, new methods to determine the detection asymmetries are currently under development and are expected to reduce also the systematic uncertainties on the detection asymmetry measurements.

Moreover, a new method to account for the tracking asymmetries will allow to avoid the weighting strategy of the signal samples that has been used in this analysis (see Sec. 6.5.1), leading to a gain in the effective statistics of the signal samples of about 35% for the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$ , and 22% for the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$ . The possibility of using a more inclusive sample selecting the  $\bar{D}^0 (\rightarrow K^+ \pi^-) \mu^+ X^-$  final state, with the idea to separate the  $B^+$  decays from the  $B^0$  decays using the corrected  $B$  mass is as well under investigation.

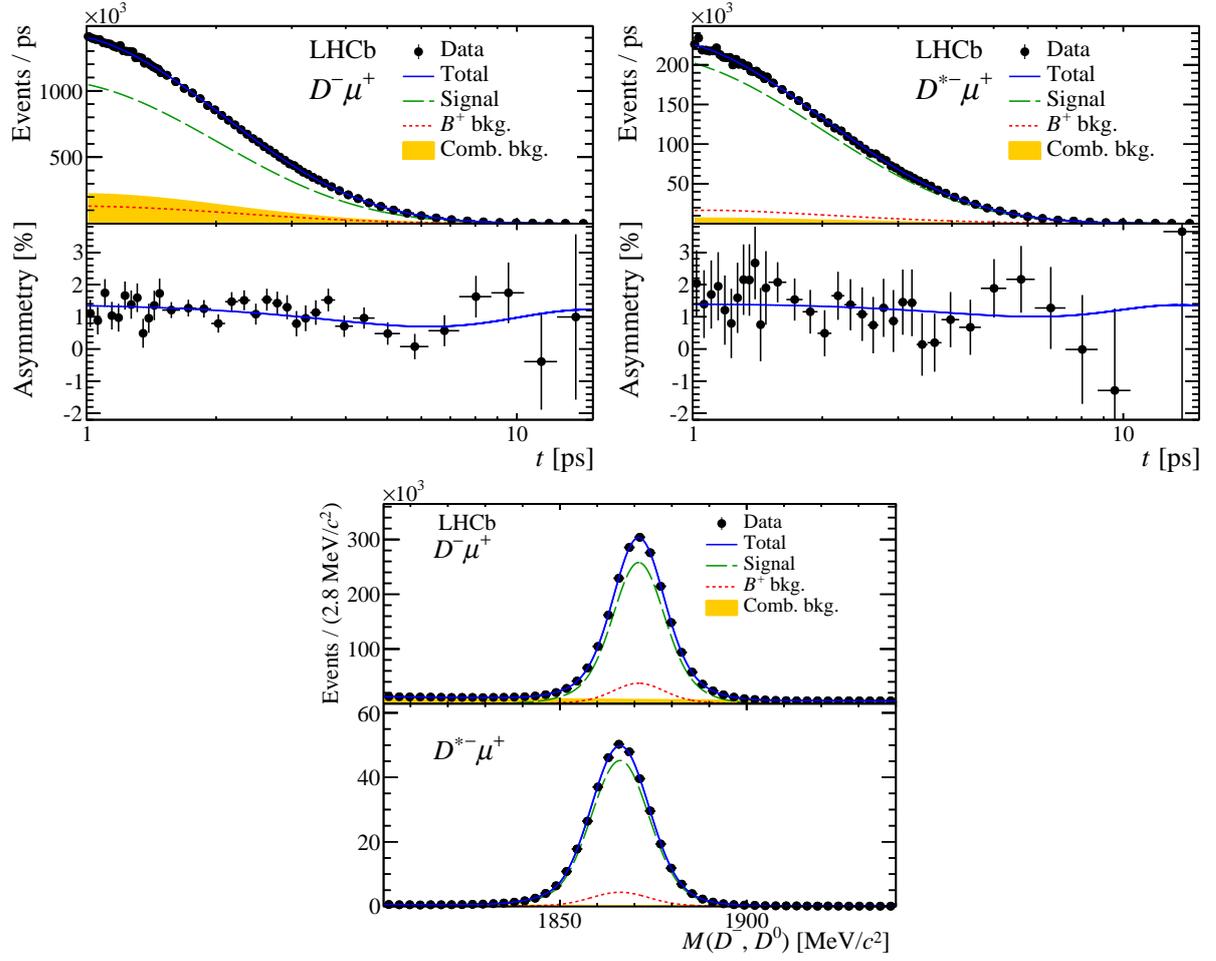


Figure 10.2: *Top:*  $B$  decay time distribution and charge asymmetry for  $D^- \mu^+$  (left) and  $D^{*-} \mu^+$  (right) samples. The datasets of each year and magnet polarity are combined and the fit results are overlaid. The number of bins for the charge asymmetry is reduced for illustration purposes. The non-zero charge asymmetry is due to the detection asymmetries (offset of the oscillating behavior) and  $B^0$  production asymmetry (amplitude of the oscillating behavior). *Bottom:* Mass distributions of  $D^-$  candidates in the  $D^- \mu^+$  sample and  $\overline{D}^0$  candidates in the  $D^{*-} \mu^+$  sample. The datasets of each year and magnet polarity are combined and the fit results are overlaid. All the data distributions shown are obtained after applying the kinematical weights (see Sec. 6.5.1).

Besides a reduction of the statistical uncertainty of  $a_{\text{sl}}^d$  provided by the more inclusive sample, the technique is expected to allow the extraction of the  $B^+$  decays fraction from the data, leading to a reduction of the systematic uncertainty associated to  $a_{\text{sl}}^d$ . Finally, more precise determinations of the  $B^+$  production asymmetry are currently finalized, and will lead to the reduction of the second largest contribution to the systematic uncertainty assigned to  $a_{\text{sl}}^d$ . In addition, studies on the  $\Lambda_b$  background could provide more precise constraints for the description of this physics background, determining an additional reduction of the systematic uncertainty of  $a_{\text{sl}}^d$ . Accounting for additional datasets corresponding to the integrated luminosity of  $6 \text{ fb}^{-1}$  ex-

pected to be acquired by LHCb in Run-II and  $50 \text{ fb}^{-1}$  planned for the LHCb upgrade [129], the statistical uncertainty on  $a_{\text{sl}}^d$  is expected to reduce to 0.024%. Taking into account the reduction of the statistical uncertainty of the measured detection asymmetries, the total uncertainty on  $a_{\text{sl}}^d$  is expected to be about 0.17%. Given that LHCb will be able to provide an  $a_{\text{sl}}^s$  measurement expected to have about the same precision, and the two measurements are expected to be completely uncorrelated, the combination of the two measurements could represent a challenge for the SM. In the case of “DØ -dimuon-like” measured values for  $a_{\text{sl}}^d$  and  $a_{\text{sl}}^s$ , the discrepancy with the SM predictions is approaching the  $5\sigma$  significance. In the case of “SM-like” measured values for  $a_{\text{sl}}^d$  and  $a_{\text{sl}}^s$ , the tension with the DØ dimuon measurement is expected to be about  $2.8\sigma$ . These numbers do not account yet for the improvements concerning the  $B^+$  and  $\Lambda_b$  background decays description, that aim to further increase the precision of the  $a_{\text{sl}}^d$  measurement.

The decay time description studied, developed and validated in this work not only allows for performing the precision measurement of  $a_{\text{sl}}^d$ , but provides a method for a variety of precision time-dependent measurements using partially reconstructed decays. A very interesting example is the determination of the  $B^0 - \bar{B}^0$  production asymmetry as a function of the kinematics of the  $B^0$  meson, using semileptonic samples. This will help towards an improved understanding of hadronization processes in proton-proton collisions, where theory currently needs more input from the experimental side. Moreover a better knowledge of the  $B^0 - \bar{B}^0$  production asymmetry is an essential input for many  $CP$  analyses at LHCb, and helps to reduce the systematic uncertainties in other measurements. As an example, the knowledge of this production asymmetry would decrease significantly the systematic uncertainty in the measurement of the decay width difference  $\Delta\Gamma_d$ , which has recently been proposed as potential explanation for the anomalous dimuon asymmetry measured from the DØ experiment. Moreover, new ideas for searches for  $CPT$  invariance with LHCb semileptonic data samples [130] are also based on a reliable decay time description as developed in this work.

# Appendix A

## Analytical form of PDFs and relative normalizations

### A.1 Decay time PDFs

The analytical form of the decay time PDF is here reported. The PDFs used to describe the other components of the sample can be obtained from the PDFs of the signal, by making the necessary simplifications and changes, as explained in Chapter .8.

First, the interest is to express the analytical shape of the  $H(tk_i)$  function used in Eq. 7.20

$$N(t) = \sum_{k_i} H(tk_i) F_i k_i \Delta k_i, \quad (\text{A.1})$$

where  $k_i$  is the position of each bin of the  $k$ -factor histogram,  $F_i$  is the height of the bin, and  $\Delta k_i$  the bin width. With a  $k$ -factor histogram normalized to the unity,  $F_i \Delta k_i$  correspond to the fraction of events in MC having a  $k$ -factor contained in the bin  $i$ . As reminder,  $H(tk_i)$  is defined in Eq. 7.17 as

$$H(tk) = \left( \int_{-\infty}^{+\infty} T(t'') R(tk - t'') dt'' \right), \quad (\text{A.2})$$

using  $tk$  instead of  $t' = \frac{L_{reco} M}{p_{true}} = tk$ , and indicating with  $T(t'')$  the decay rates in Eq. 8.23, and with  $R(tk - t'')$  a flight distance resolution function. In this case a simple single gaussian resolution function with mean  $m$  and width  $\sigma$  is considered. Given the functional shape of  $T(t'')$ , the first step is to write the convolution of an exponential function with the gaussian resolution (the cosine term can be written in complex exponential form, the exponential form here considered is the simplest possible case)

$$H(tk) = \int e^{-\frac{t''}{\tau}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(tk-t''+m)^2}{2\sigma^2}} dt'' \quad (\text{A.3})$$

$$= \frac{1}{2} e^{-\frac{tk}{\tau} + \frac{\sigma^2 k^2}{2\tau^2} - \frac{mk}{\tau}} \text{Erfc} \left( \frac{\sigma k}{\sqrt{2}\tau} - \frac{t}{\sqrt{2}\sigma} - \frac{m}{\sqrt{2}\sigma} \right). \quad (\text{A.4})$$

The subsequent step is again analytical and consists of multiplying the obtained  $H(tk)$  function by the acceptance function  $a(t) = (1 - e^{-(t-t_{\text{shift}})/\alpha})(1 + \beta t)$ . The PDF for an unbinned fit could be written as

$$\mathcal{P}(tk) = \mathcal{N} \frac{1}{2} e^{-\frac{tk}{\tau} + \frac{\sigma^2 k^2}{2\tau^2} - \frac{mk}{\tau}} \text{Erfc} \left( \frac{\sigma k}{\sqrt{2}\tau} - \frac{t}{\sqrt{2}\sigma} - \frac{m}{\sqrt{2}\sigma} \right) (1 - e^{-(t-t_{\text{shift}})/\alpha}) (1 + \beta t), \quad (\text{A.5})$$

with  $\mathcal{N}$  normalization factor, defined such that  $\int \mathcal{P}(t) = 1$ . The dependence on  $k$  is kept so far, the last step will be the numerical convolution as written in Eq. 7.20. The function of Eq. A.5 needs to be integrated on the fit interval in order to determine the normalization  $\mathcal{N}$ . In the case of the binned fit used to measure  $a_{\text{sl}}^d$  in this thesis, the integral of the function in Eq. A.5 is not only performed on the overall fit interval in order to determine the normalization, but also in each bin. Binned fit procedures might use approximations to estimate the number of expected events in each bin, in this case the integral in each bin is calculated. The expression for the undefined integral  $\mathcal{I}(tk)$  if the function in Eq. A.5 is

$$\mathcal{I}(tk) = \frac{1}{2} e^{\frac{\sigma^2 k^2}{2\tau^2}} e^{-\frac{mk}{\tau}} ((I_{0A} + I_{0B}) + \beta(I_{1A} + I_{1B})) \quad (\text{A.6})$$

A list of symbols is introduced for illustration purposes

$$\begin{aligned} a &= \frac{1}{\sigma\sqrt{2}}, & b &= \frac{k}{\tau}, \\ b' &= \frac{k}{\tau} + \frac{1}{\alpha}, & c &= \frac{\sigma k}{\tau\sqrt{2}} - \frac{m}{\sigma\sqrt{2}}, \end{aligned}$$

where  $k$  is the value of the  $k$  factor corresponding to the bin of the  $k$ -factor distribution considered. When considering the complete PDF to extract  $a_{\text{sl}}^d$  (Eq. 8.23), the  $k$  to be considered is the complex number  $k(1 + i\Delta M\tau)$ . The contributions to  $\mathcal{I}(tk)$  in Eq. A.6, are

$$\begin{aligned} I_{0A} &= \frac{1}{b} \left( e^{\frac{b^2}{4a^2} - \frac{bc}{a}} \text{Erf} \left[ \frac{b}{2a} - c + at \right] - e^{-bt} \text{Erfc} \left[ -at + c \right] \right) \\ I_{0B} &= -e^{\frac{t_{\text{shift}}}{\alpha}} \frac{1}{b'} \left( e^{\frac{b'^2}{4a^2} - \frac{b'c}{a}} \text{Erf} \left[ \frac{b'}{2a} - c + at \right] - e^{-b't} \text{Erfc} \left[ -at + c \right] \right) \\ I_{1A} &= \left( \frac{1}{b^2} - \frac{1}{2a^2} + \frac{c}{ab} \right) e^{\frac{b^2}{4a^2} - \frac{bc}{a}} \text{Erf} \left[ at - c + \frac{b}{2a} \right] \\ &\quad - \frac{1}{ab\sqrt{\pi}} e^{-((-at+c)^2 - bt)} - \frac{(bt+1)}{b^2} e^{-bt} \text{Erfc} \left[ -at + c \right] \\ I_{1B} &= -e^{\frac{t_{\text{shift}}}{\alpha}} \left[ \left( \frac{1}{b'^2} - \frac{1}{2a^2} + \frac{c}{ab'} \right) e^{\frac{b'^2}{4a^2} - \frac{b'c}{a}} \text{Erf} \left[ at - c + \frac{b'}{2a} \right] \right. \\ &\quad \left. - \frac{1}{ab'\sqrt{\pi}} e^{-((-at+c)^2 - b't)} - \frac{(b't+1)}{b'^2} e^{-b't} \text{Erfc} \left[ -at + c \right] \right] \end{aligned} \quad (\text{A.7})$$

For each time bin, defined with lower and upper limits  $a$  and  $b$  respectively, the multiplicative convolution over the  $k$ -factor convolution is numerically performed at the end

$$N(t) = \sum_{k_i} (\mathcal{I}(bk_i) - \mathcal{I}(sk_i)) F_i k_i \Delta k_i. \quad (\text{A.8})$$

## A.2 $K^+\pi^-\pi^-/K^+\pi^-$ invariant mass PDFs

The PDF used to describe the  $K^+\pi^-\pi^-/K^+\pi^-$  invariant mass in the  $\alpha_{\text{sl}}^d$  measurement are widely-used, therefore there is no need to report them here. The Crystal Ball function is reported, to help the reader in the identification of the parameters reported in the tables of Chapter .8

$$CB(m, m_{D^+}, \sigma_{CB}, \alpha_{CB}, n_{CB}) = \mathcal{N} \begin{cases} e^{-\frac{(m-m_{D^+})^2}{2\sigma_{CB}^2}} & \text{for } \frac{m-m_{D^+}}{\sigma_{CB}} > -\alpha_{CB} \\ A \cdot \left(B - \frac{m-m_{D^+}}{\sigma_{CB}}\right)^{-n_{CB}} & \text{for } \frac{m-m_{D^+}}{\sigma_{CB}} \leq -\alpha_{CB} \end{cases}$$

where  $\mathcal{N}$  is the normalization and  $A$  and  $B$  are defined as follows

$$A = \left(\frac{n_{CB}}{|\alpha_{CB}|}\right)^{n_{CB}} e^{-\frac{|\alpha_{CB}|^2}{2}}$$

$$B = \frac{n_{CB}}{|\alpha_{CB}|} - |\alpha_{CB}|$$



# Appendix B

## Fit projections

In this section the nominal fit projections, in the  $D/D^0$  mass, decay time, charge of the muon in the final state, are reported. The two data samples,  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$ , are split in subsets according to the center-of-mass energy and magnet polarity. Fit results and fit quality indications are reported in Sec. 8.7.

## B.1 $B^0 \rightarrow D^- \mu^+ \nu_\mu X$ data sample

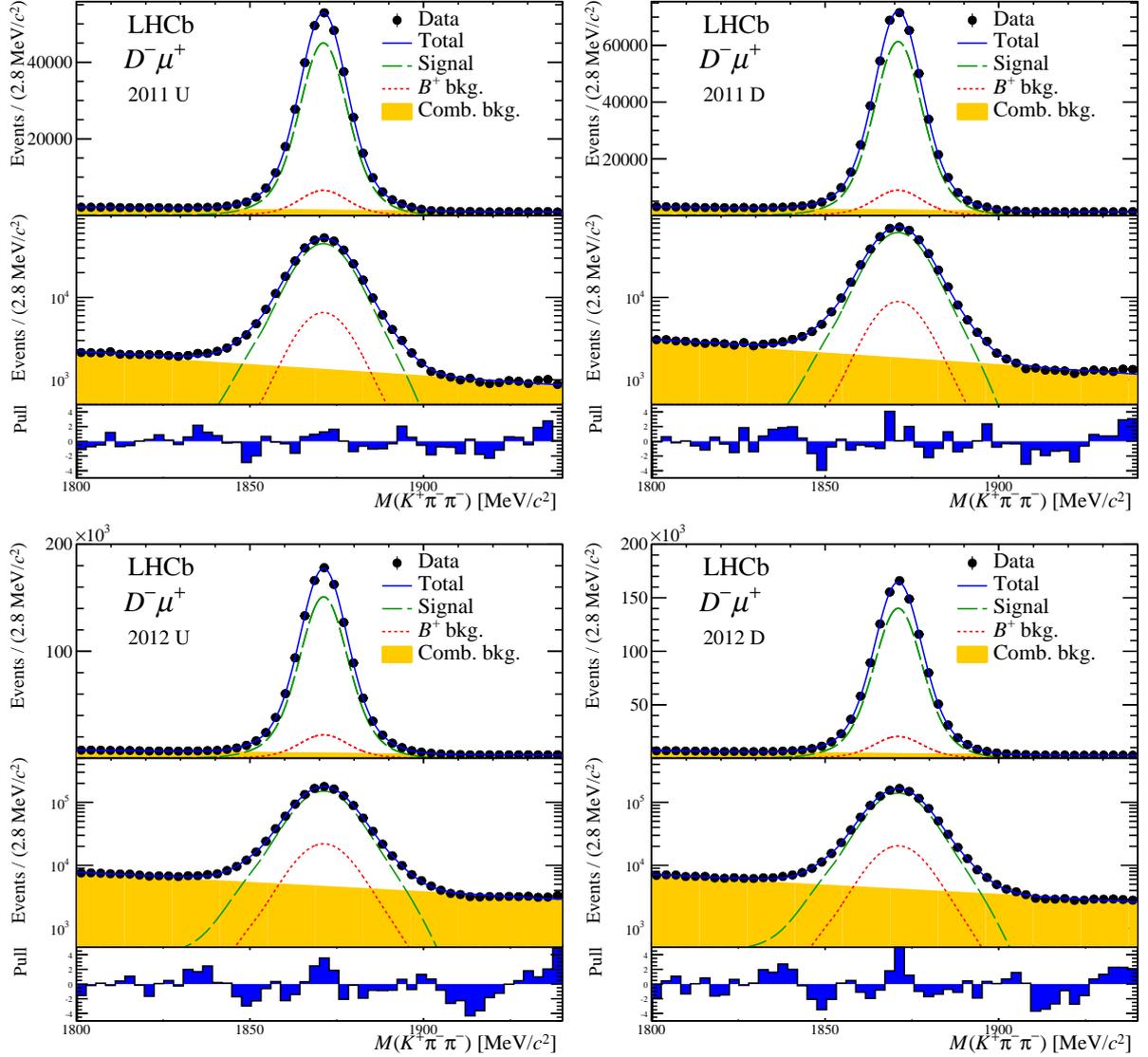


Figure B.1: Invariant mass distributions of the  $K^\mp \pi^\pm \pi^\pm$  combinations. The fit results are overlaid to the distribution of the dataset of each year and magnet polarity. The data distributions shown are obtained after applying the kinematical weights (see Sec. 6.5.1). For each subsample, the first two top rows show the invariant mass distributions using a linear and a logarithmic y-axis, respectively. The last row shows the so-called *Pull* values, i.e. the difference between the number of events observed in data for each bin of the histogram and the number of events predicted by the model in the same bin, divided by the uncertainty on the number of events.

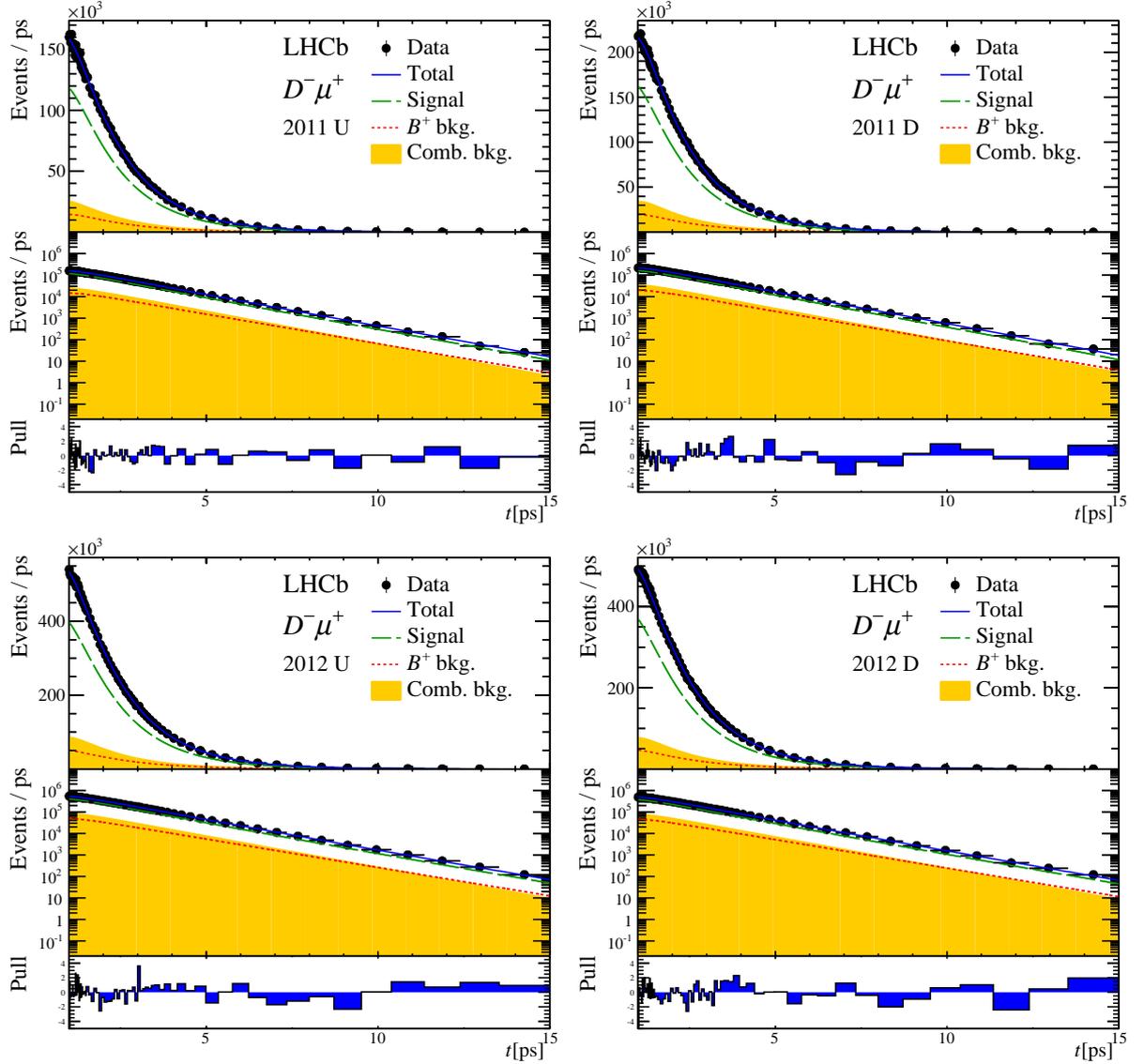


Figure B.2:  $B$  decay time distributions. The results are overlaid to the distribution of the dataset of each year and magnet polarity. The data distributions shown are obtained after applying the kinematical weights (see Sec. 6.5.1). For each subsample, the first two top rows show the  $B$  decay time using a linear and a logarithmic scale. The different components of the sample are highlighted with different colors and styles. The last row shows the so-called *Pull* values, i.e. the difference between the number of events observed in data for each bin of the histogram and the number of events predicted by the full model in the same bin, divided by the uncertainty on the number of events.

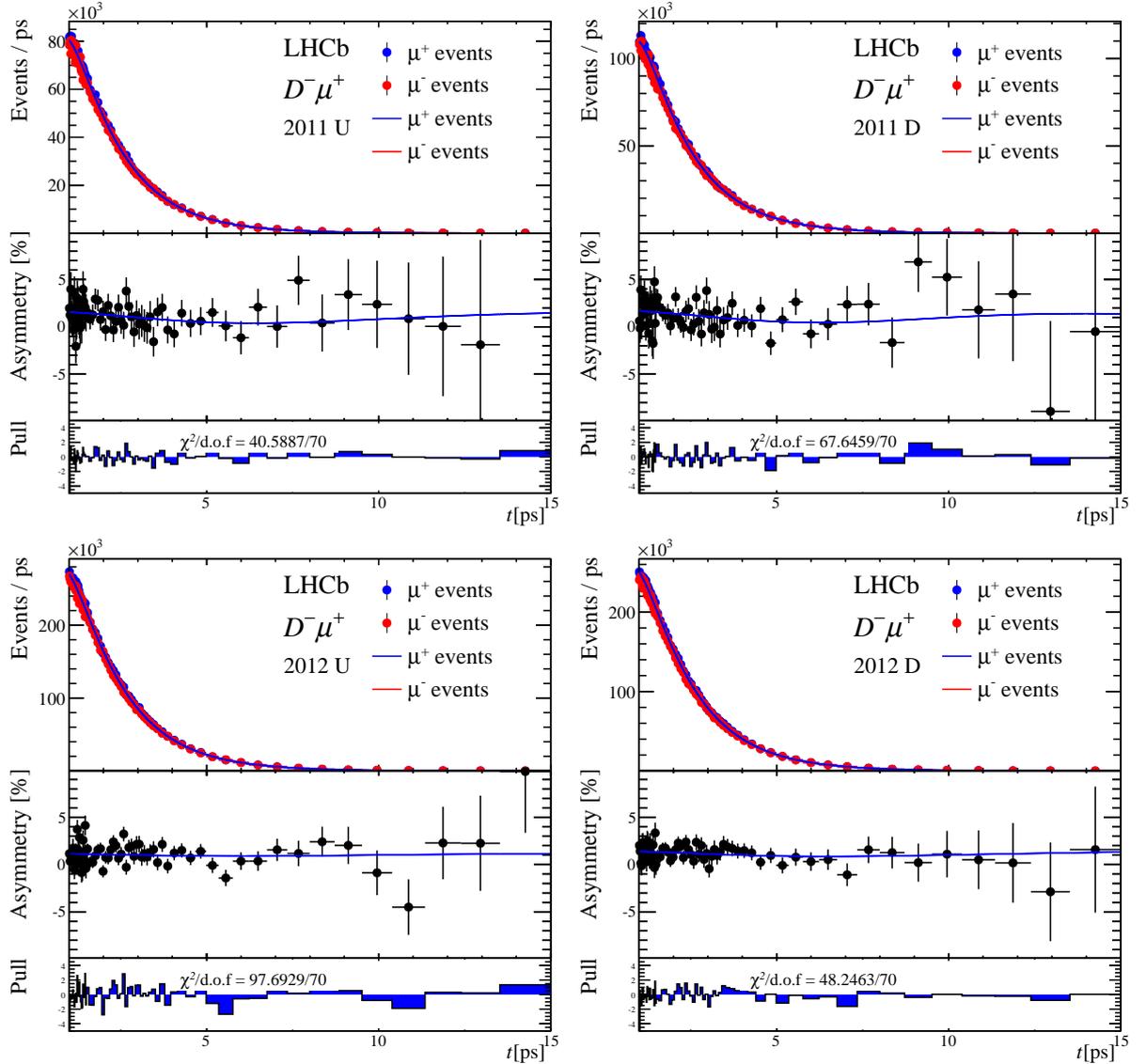


Figure B.3:  $B$  decay time distributions for  $K^+\pi^-\pi^-\mu^+$  and  $K^-\pi^+\pi^+\mu^-$  reconstructed final states and charge asymmetry. The fit results are overlaid to the distribution of the dataset of each year and magnet polarity. The data distributions shown are obtained after applying the kinematical weights (see Sec. 6.5.1). For each subsample, the first row shows the  $B$  decay time distributions for  $K^+\pi^-\pi^-\mu^+$  and  $K^-\pi^+\pi^+\mu^-$  final states separately. The second row shows the charge asymmetry of the particles in the final states, with the full fit model overlaid. The *Pull* values in the last row refer to the charge asymmetry projection in the second row.

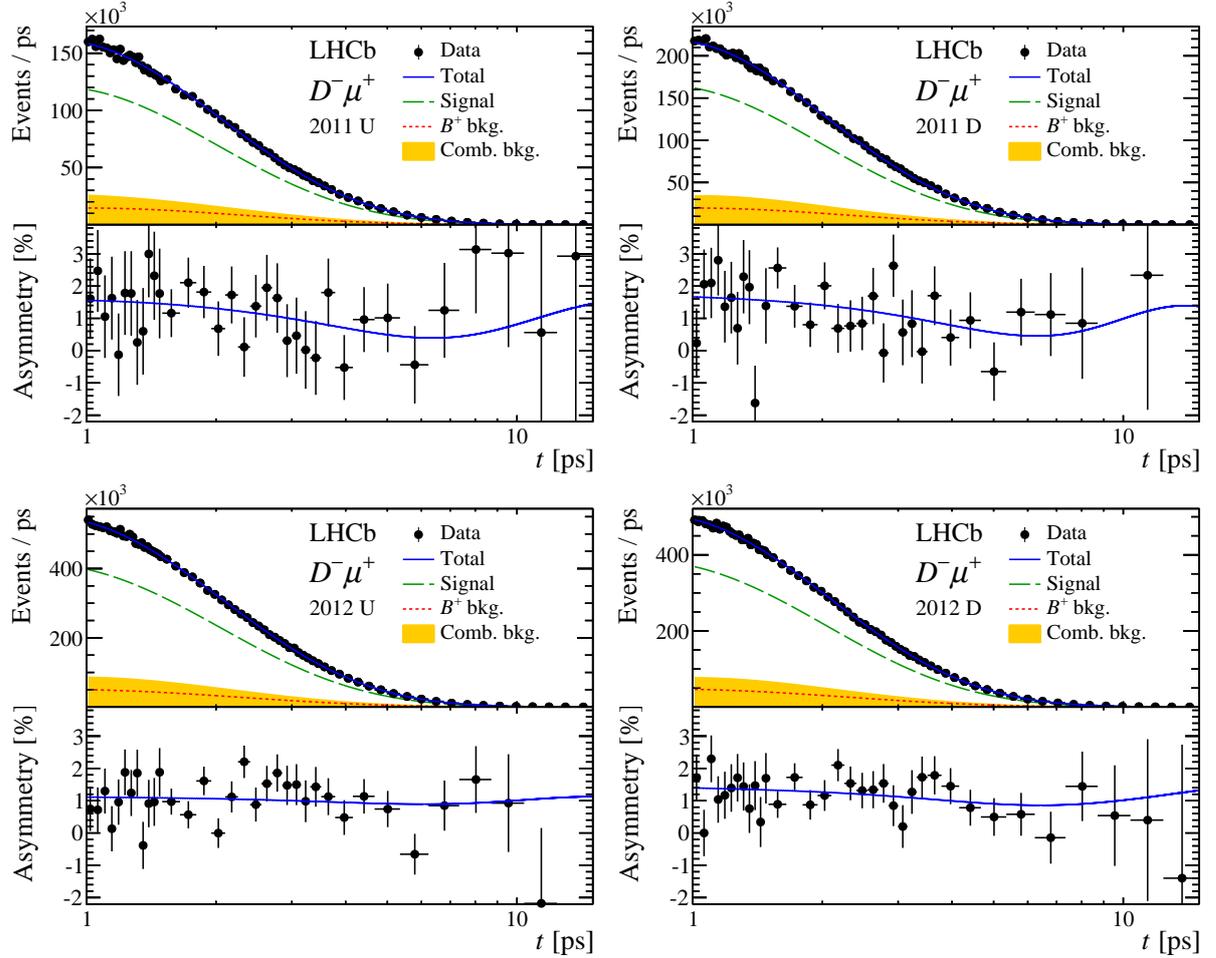


Figure B.4:  $B$  decay time distribution for and charge asymmetry. The fit results are overlaid to the distribution of the dataset of each year and magnet polarity. The number of bins for the charge asymmetry is reduced for illustration purposes, as well as the logarithmic scale of the decay time axis. The non-zero charge asymmetry is due to the detection asymmetries (offset of the oscillating behavior) and  $B^0$  production asymmetry (amplitude of the oscillating behavior). The data distributions shown are obtained after applying the kinematical weights (see Sec. 6.5.1).

## B.2 $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$ data sample

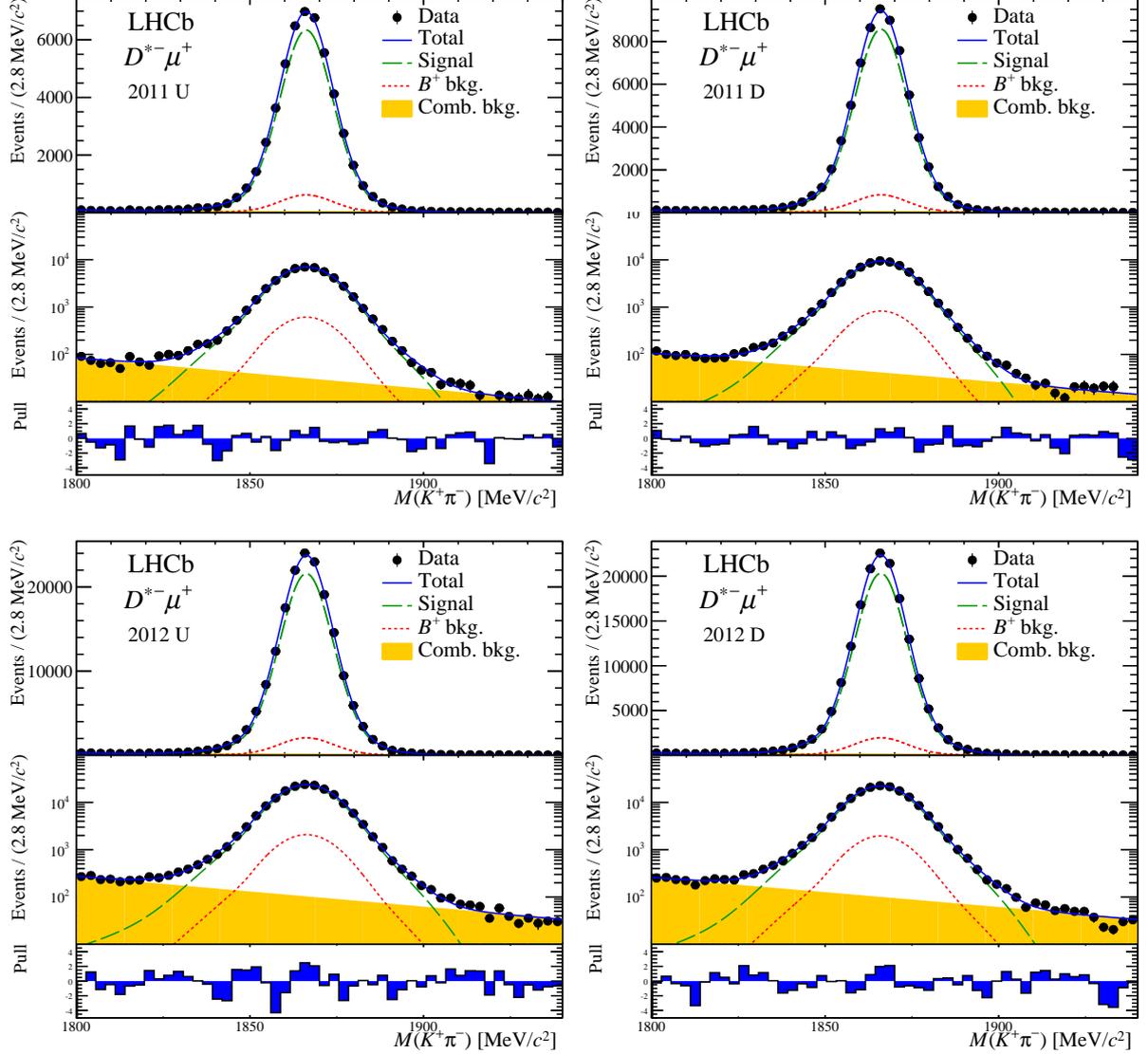


Figure B.5: Invariant mass distributions of the  $K^\mp \pi^\pm$  combinations. The fit results are overlaid to the distribution of the dataset of each year and magnet polarity. All the data distributions shown are obtained after applying the kinematical weights (see Sec. 6.5.1). For each subsample, the first two top rows show the invariant mass distributions using a logarithmic and a not-logarithmic y-axis, respectively. The last row shows the so-called *Pull* values, i.e. the difference between the number of events observed in data for each bin of the histogram and the number of events predicted by the model in the same bin, divided by the uncertainty on the number of events.

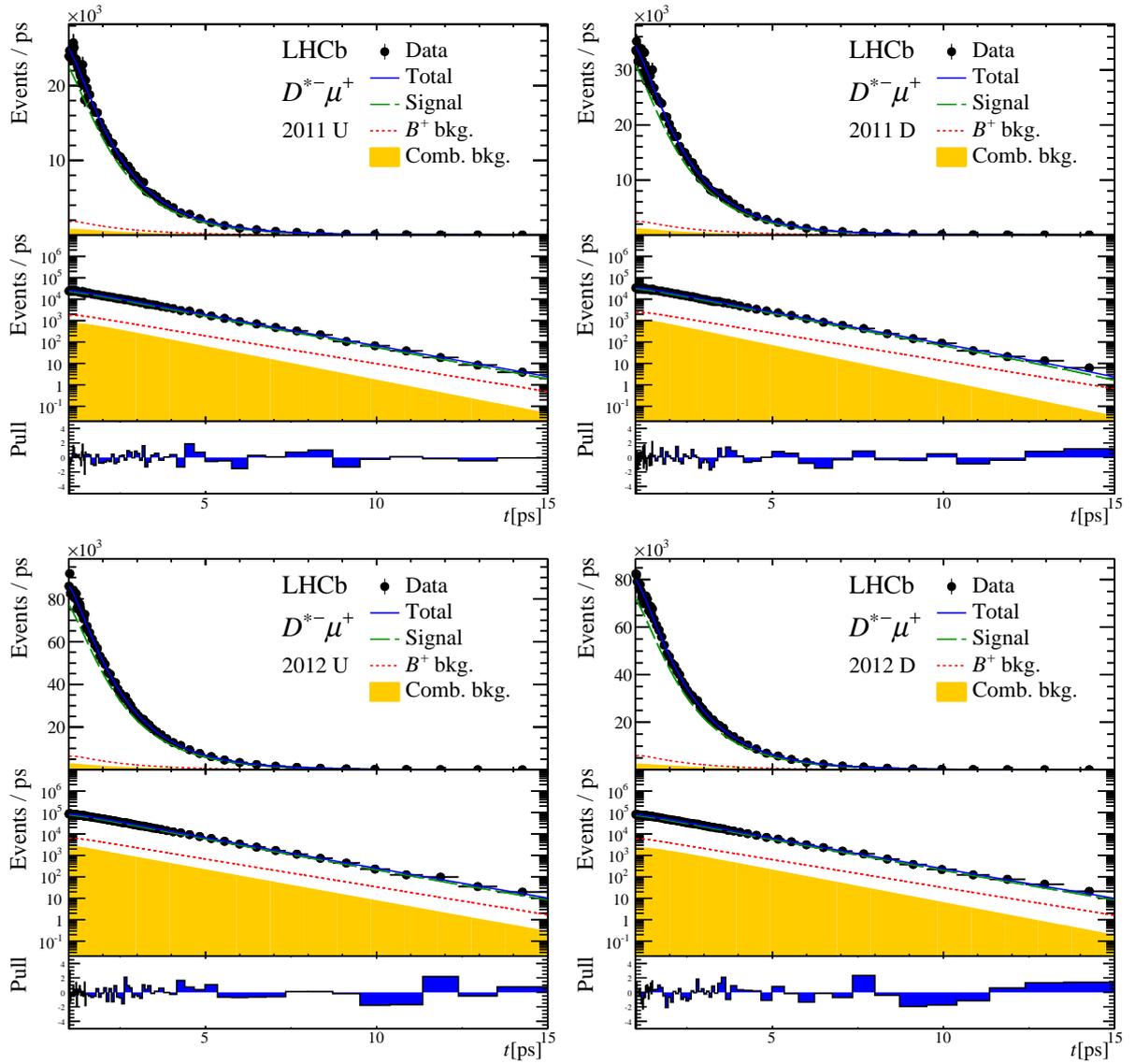


Figure B.6:  $B$  decay time distributions. The fit results are overlaid to the distribution of the dataset of each year and magnet polarity. The data distributions shown are obtained after applying the kinematical weights (see Sec. 6.5.1). For each subsample, the first two top rows show the  $B$  decay time using a linear and a logarithmic scale. The different components of the sample are highlighted with different colors and styles. The last row shows the so-called *Pull* values, i.e. the difference between the number of events observed in data for each bin of the histogram and the number of events predicted by the full model in the same bin, divided by the uncertainty on the number of events.

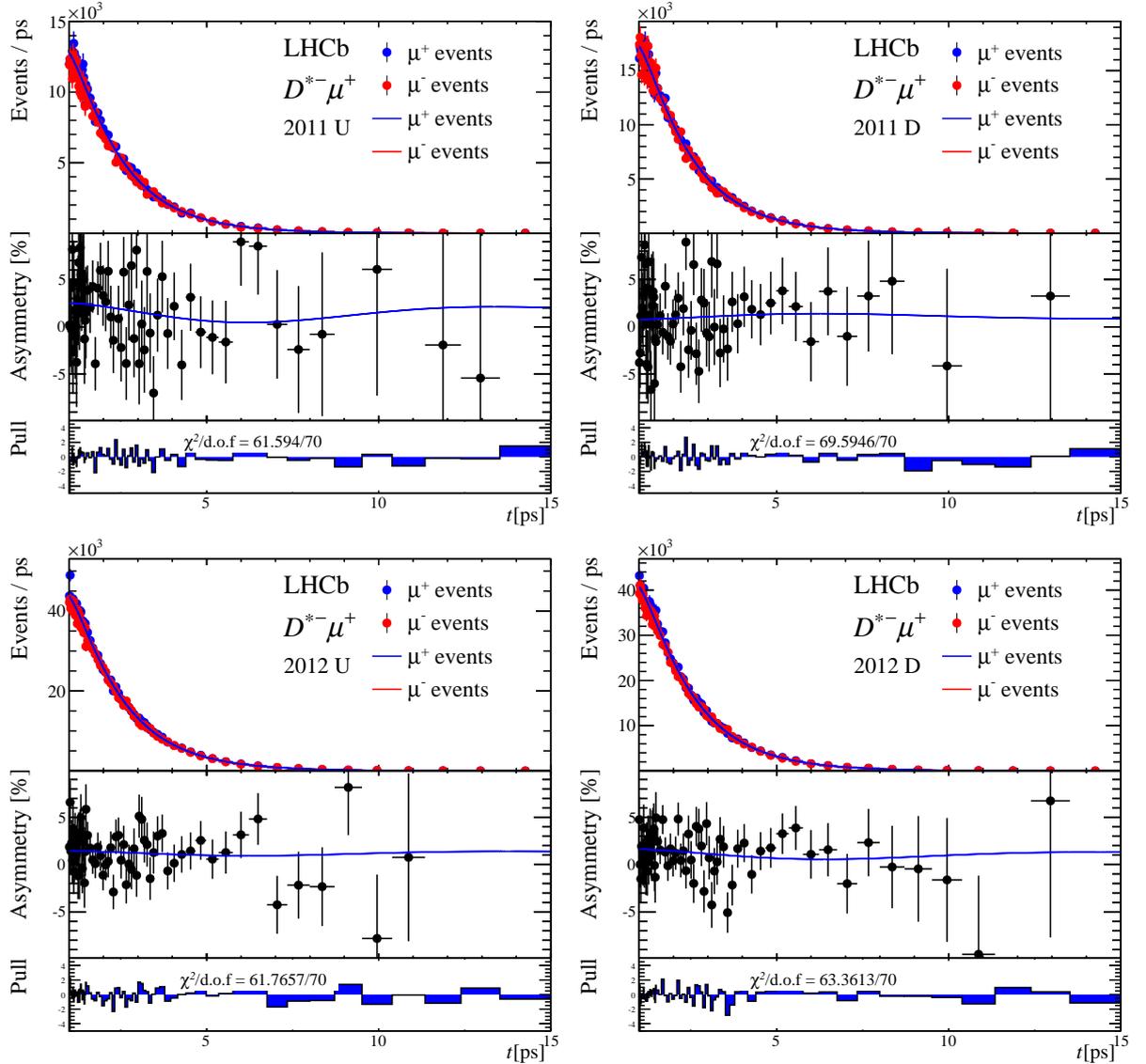


Figure B.7:  $B$  decay time distributions for  $K^+\pi^-\pi^-\mu^+$  and  $K^-\pi^+\pi^+\mu^-$  reconstructed final states and charge asymmetry. The fit results are overlaid to the distribution of the dataset of each year and magnet polarity. The data distributions shown are obtained after applying the kinematical weights (see Sec. 6.5.1). For each subsample, the first row shows the  $B$  decay time distributions for  $K^+\pi^-\pi^-\mu^+$  and  $K^-\pi^+\pi^+\mu^-$  final states separately. The second row shows the charge asymmetry of the particles in the final states, with the full fit model overlaid. The *Pull* values in the last row refer to the charge asymmetry projection in the second row.

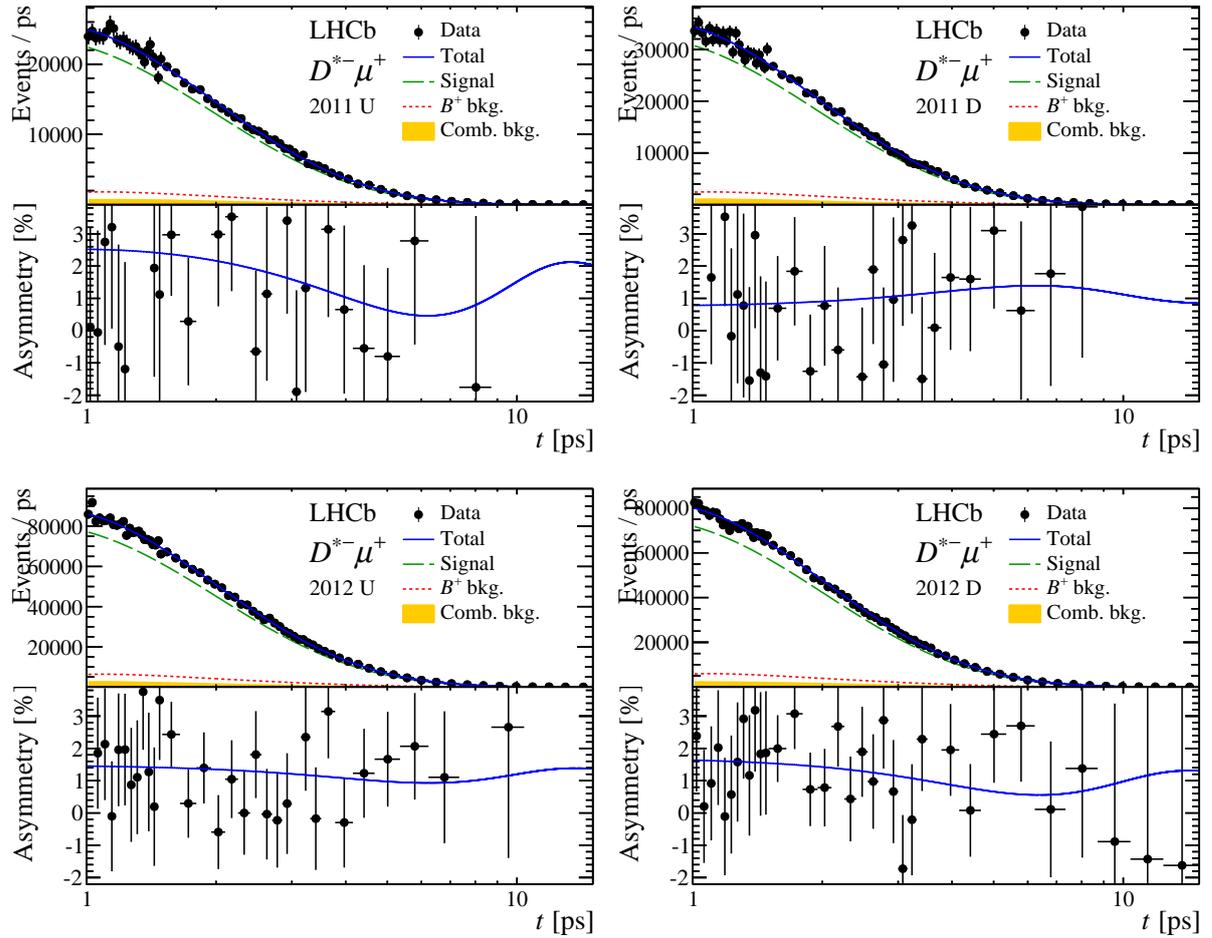


Figure B.8:  $B$  decay time distribution for and charge asymmetry. The fit results are overlaid to the distribution of the dataset of each year and magnet polarity. The number of bins for the charge asymmetry is reduced for illustration purposes as well as the logarithmic scale of the decay time axis. The non-zero charge asymmetry is due to the detection asymmetries (offset of the oscillating behavior) and  $B^0$  production asymmetry (amplitude of the oscillating behavior). The data distributions shown are obtained after applying the kinematical weights (see Sec. 6.5.1).



# Appendix C

## Crosschecks

As reported in Sec. 9.3 and shown in Fig. 9.7, a difference between the  $a_{\text{sl}}^d$  values measured with data acquired in 2011 with the two magnet polarities. This difference, calculated separately for the two decay modes is

$$\begin{aligned}
 2011 \quad \text{data} \quad B^0 \rightarrow D^- \mu^+ \nu_\mu X \quad \delta(a_{\text{sl}}^d) &= (-2.41 \pm 1.09)\%, \\
 B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X \quad \delta(a_{\text{sl}}^d) &= (-3.81 \pm 1.89)\%, \quad (\text{C.1})
 \end{aligned}$$

where the uncertainty on the difference is calculated as the sum in quadrature of statistical uncertainty estimated by the fit on the data and the statistical component of the uncertainty on the detection asymmetry measurement (the single contributions of the detection asymmetry are accounted for). The statistical uncertainties obtained with this definition are used in Fig. 9.7, and the fit results are reported in Tab. C.1 and Tab. C.2. The discrepancy between the  $a_{\text{sl}}^d$  measurements using the two different magnet polarities in 2011 is  $2.2\sigma$  when considering the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  data samples and  $2.0\sigma$  when considering the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  data samples.

Table C.1: Time-dependent fit  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  results for the four sub samples of the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  decays. The statistical uncertainty here considered accounts for the uncertainty of the fit on data and the statistical component of the uncertainty on the detection asymmetries measurements.

Data Sample	$a_{\text{sl}}^d$	$A_{\text{P}}$
2011 Magnet Down	$0.0045 \pm 0.0071$	$-0.0109 \pm 0.0044$
2011 Magnet Up	$-0.0194 \pm 0.0083$	$-0.0025 \pm 0.0051$
2012 Magnet Down	$0.0040 \pm 0.0047$	$-0.0078 \pm 0.0029$
2012 Magnet Up	$-0.0039 \pm 0.0045$	$-0.0000 \pm 0.0028$
Overall $B^0 \rightarrow D^- \mu^+ \nu_\mu X$	$-0.0019 \pm 0.0028$	$-0.0047 \pm 0.0017$

A number of crosschecks is performed, in order to understand the origin of this difference, and to ensure that the final  $a_{\text{sl}}^d$  result is not affected by experimental effects not taken into account.

The nominal  $a_{\text{sl}}^d$  fit determines the value of the parameters  $a_{\text{sl}}^d$  and  $A_{\text{P}}$ , but it is possible

Table C.2: Time-dependent fit  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  results for the four sub samples of the  $B^0 \rightarrow D^{*-} \mu^+ \nu_{\mu} X$  decays. The statistical uncertainty here considered accounts for the uncertainty of the fit on data and the statistical component of the uncertainty on the detection asymmetries measurements.

Data Sample	$a_{\text{sl}}^d$	$A_{\text{P}}$
2011 Magnet Down	$0.0263 \pm 0.0126$	$-0.0092 \pm 0.0083$
2011 Magnet Up	$-0.0129 \pm 0.0145$	$-0.0028 \pm 0.0097$
2012 Magnet Down	$0.0180 \pm 0.0081$	$-0.0154 \pm 0.0054$
2012 Magnet Up	$-0.0018 \pm 0.0080$	$-0.0018 \pm 0.0053$
Overall $B^0 \rightarrow D^{*-} \mu^+ \nu_{\mu} X$	$0.0077 \pm 0.0049$	$-0.0079 \pm 0.0033$

to use a different parametrization of the decay rates, in order to extract the offset and the amplitude of the charge asymmetry of the particles in the final state,  $A_{\text{mea}}(t)$ , as described in Sec. C.1. The stability of the measured values for the offset and amplitude of  $A_{\text{mea}}(t)$  on data acquired with different magnet polarities, suggests that the cause of the difference in the  $a_{\text{sl}}^d$  measured values, could be ascribed to the determination of the detection asymmetries.

Since the detection asymmetries values used as inputs for the  $a_{\text{sl}}^d$  measurement using  $B^0 \rightarrow D^- \mu^+ \nu_{\mu} X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_{\mu} X$  data samples are evaluated using the same control samples, a correlation is expected, and it is also observed when comparing the  $a_{\text{sl}}^d$  measured values using the two decay modes.

As crosscheck, the kinematical regions where the detection asymmetries are expected to be large or maximal, have been excluded from the data samples, as explained in Sec. C.2. In spite of variations of the  $a_{\text{sl}}^d$  results on the single samples of about  $1\sigma$ , applying or not these fiducial cuts has no significant impact on the final  $a_{\text{sl}}^d$  result.

In order to probe the stability of the  $a_{\text{sl}}^d$  measurement as function of the data taking conditions, the  $B^0 \rightarrow D^- \mu^+ \nu_{\mu} X$  data sample is split in sub-samples according to the data taking period, as shown in Sec. C.3. The results obtained on the sub-samples exhibit a stable behavior.

Also when splitting the  $B^0 \rightarrow D^- \mu^+ \nu_{\mu} X$  data sample according to the number of the Primary Vertices of the events, no particular structure is observed, as shown in Sec. C.4. In addition, for the  $B^0 \rightarrow D^{*-} \mu^+ \nu_{\mu} X$  data sample, the effect of neglecting possible background of random slow pions associated with a real  $D^0$  is estimated in Sec. C.5. The effect on the  $a_{\text{sl}}^d$  measurement is negligible.

On the other hand, the better agreement between the higher-statistics data samples acquired in 2012,

$$\begin{aligned}
 2012 \quad \text{data} \quad B^0 \rightarrow D^- \mu^+ \nu_{\mu} X \quad \delta(a_{\text{sl}}^d) &= (-0.79 \pm 0.64)\%, \\
 B^0 \rightarrow D^{*-} \mu^+ \nu_{\mu} X \quad \delta(a_{\text{sl}}^d) &= (-1.89 \pm 1.13)\%, \quad (\text{C.2})
 \end{aligned}$$

corresponding to  $1.2 \sigma$  in the  $B^0 \rightarrow D^- \mu^+ \nu_{\mu} X$  decay mode and  $1.6 \sigma$  in the  $B^0 \rightarrow D^{*-} \mu^+ \nu_{\mu} X$  decay mode, support the hypothesis of a statistical effect.

## C.1 Alternative fit parametrization

In the nominal fit parametrization, the untagged decay rates for the two final states,  $f$  and  $\bar{f}$  are expressed in terms of  $A_D$ ,  $A_P$ ,  $a_{\text{sl}}^d$ , as in Eq. 8.23, and here reported for simplicity

$$\begin{aligned}\mathcal{P}_{\text{sig}}(f, t) &= \mathcal{N}e^{-\Gamma t} \left( 1 + A_D + \frac{a_{\text{sl}}^d}{2} + \left( A_P - \frac{a_{\text{sl}}^d}{2} \right) \cos \Delta m_d t \right) , \\ \mathcal{P}_{\text{sig}}(\bar{f}, t) &= \mathcal{N}e^{-\Gamma t} \left( 1 - A_D - \frac{a_{\text{sl}}^d}{2} - \left( A_P - \frac{a_{\text{sl}}^d}{2} \right) \cos \Delta m_d t \right) .\end{aligned}$$

An alternative is to describe the same decay rates not any more as function of the parameters  $A_P$ ,  $A_D$  and  $a_{\text{sl}}^d$ , but in terms of the offset and amplitude of the charge asymmetry as function of the  $B$  decay time (Eq. 4.20). Eq. (4.20) can be written as:

$$A_{\text{meas}}(t) = \frac{N(f, t) - N(\bar{f}, t)}{N(f, t) + N(\bar{f}, t)} \approx \text{OFFSET} + \text{AMPLITUDE} \cos(\Delta m_d t) , \quad (\text{C.3})$$

And for the decay rates the following parametrization can be used

$$\mathcal{P}_{\text{sig}}(f, t) = \mathcal{N}e^{-\Gamma t} (1 + \text{OFFSET} + \text{AMPLITUDE} \cos \Delta m_d t) , \quad (\text{C.4})$$

$$\mathcal{P}_{\text{sig}}(\bar{f}, t) = \mathcal{N}e^{-\Gamma t} (1 - \text{OFFSET} - \text{AMPLITUDE} \cos \Delta m_d t) . \quad (\text{C.5})$$

This parametrization is interesting, because allows for describing the data sample, without making any assumption on the detection asymmetries during the fit procedure. The measured values for the detection asymmetries can be used lately, when calculating the values for the physical parameters  $a_{\text{sl}}^d$  and  $A_P$ :

$$\begin{aligned}a_{\text{sl}}^d &= 2(\text{OFFSET} - A_D) \\ A_P &= \text{AMPLITUDE} + \text{OFFSET} - A_D\end{aligned}$$

Naturally the correlation between the parameters of interest in this case is different from the nominal case. Tab. C.4 and Tab. 8.9 report two correlation matrices. The correlation between the AMPLITUDE and OFFSET parameters is about -55%, while the correlation between  $A_P$  and  $a_{\text{sl}}^d$  is 5%. This shows that the nominal parametrization of the fit model used to measure  $a_{\text{sl}}^d$  is preferable. The OFFSET-AMPLITUDE parametrization is used as crosscheck of the fit procedure. Tab. C.3 and Tab. C.5 show the comparison between the four nominal fits on the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  samples, respectively, and the values of  $a_{\text{sl}}^d$  and  $A_P$  calculated from the OFFSET-AMPLITUDE fit results. The selection of the samples used for this crosscheck is slightly different from the nominal selection, for this reason the nominal  $a_{\text{sl}}^d$  values are different from the values reported in Tab. 8.10. Small differences between the  $a_{\text{sl}}^d$  values determined with the nominal fit procedure and the  $a_{\text{sl}}^d$  values calculated from the OFFSET determined with the alternative fits can be ascribed to small differences in the fit, for instance the different correlations between the OFFSET and AMPLITUDE parameters and  $a_{\text{sl}}^d$  and  $A_P$  can have an influence.

Table C.3:  $a_{sl}^d$  results for the four fits of  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  decays. The result obtained with the nominal parametrization is compared to value of  $a_{sl}^d$  calculated from the OFFSET fitted value.

dataset	nominal $a_{sl}^d$	$a_{sl}^d$ from OFFSET AMPLITUDE fit	OFFSET	$A_D$
2011 D	$0.0083 \pm 0.0069$	0.0109	$0.0144 \pm 0.0033$	0.0089
2011 U	$-0.0237 \pm 0.0080$	-0.0183	$0.0091 \pm 0.0038$	0.0183
2012 D	$0.0047 \pm 0.0044$	0.0071	$0.0133 \pm 0.0022$	0.0097
2012 U	$-0.0009 \pm 0.0043$	0.0009	$0.0116 \pm 0.0021$	0.0112

Table C.4: Correlation matrix  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  2012 Magnet Down, OFFSET-AMPLITUDE parametrization.

	$f_{sb}$	$\alpha$	$\beta$	$\Gamma_{sb}$	$\alpha_{mass}$	OFF.	AMPL.	$A_{P,sb}$	$A_{D,sb}$	$\sigma_{D+}$	$\alpha_{CB}$	$f_{12}$	$s_{12}$	$m_{D+}$
$f_{sb}$	1.00	-0.01	-	-0.01	0.52	-	-	0.01	0.01	-0.07	0.73	0.13	-0.04	-0.09
$\alpha$		1.00	-0.74	0.03	-0.01	-	-	-	-	-	-0.01	-	-	-
$\beta$			1.00	0.11	-	-	-	-	-	-	-	-	0.01	-
$\Gamma_{sb}$				1.00	-0.02	-	-	-	-	-0.02	-0.02	0.01	0.02	-
$\alpha_{mass}$					1.00	-	-	0.01	-	0.30	0.76	-0.15	-0.41	-0.07
OFF.						1.00	-0.55	0.11	-0.20	-	-	-	-	-
AMPL.							1.00	-0.20	0.10	-	-	-	-	-
$A_{P,sb}$								1.00	-0.49	0.01	0.01	-0.01	-0.01	-
$A_{D,sb}$									1.00	-0.01	0.01	0.01	0.01	-
$\sigma_{D+}$										1.00	0.42	-0.94	-0.84	-0.09
$\alpha_{CB}$											1.00	-0.27	-0.51	-0.15
$f_{12}$												1.00	0.63	0.07
$s_{12}$													1.00	0.10
$m_{D+}$														1.00

Fig. C.1 and Fig. C.2 show the measured values for the OFFSET and AMPLITUDE parameters on the different data samples according to the year of acquisition and magnet polarity. The stability of the OFFSET value in particular, supports the hypothesis that the discrepancy between the  $a_{sl}^d$  results obtained in the 2011 samples with the two magnet polarities is due to a fluctuation in the detection asymmetry measured values.

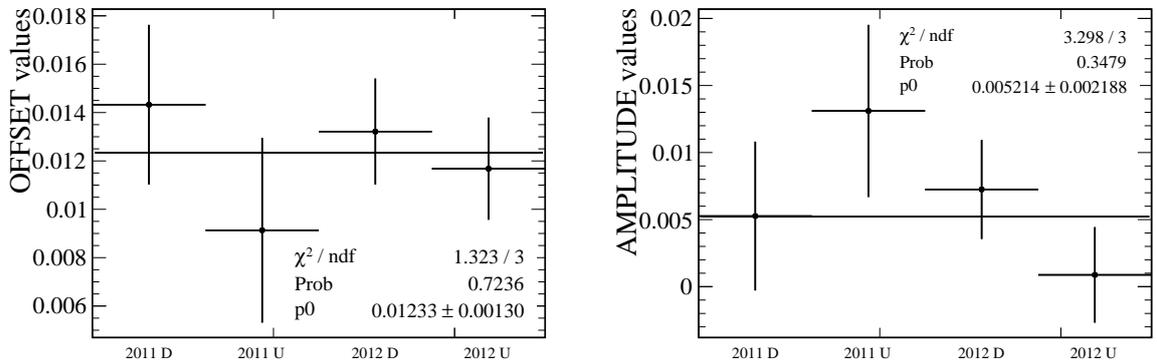


Figure C.1: OFFSET and AMPLITUDE measured values for the different  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  samples.

Table C.5:  $a_{\text{sl}}^d$  results for the four fits of  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays. The result obtained with the nominal parametrization is compared to value of  $a_{\text{sl}}^d$  calculated from the OFFSET fitted value.

dataset	nominal $a_{\text{sl}}^d$	$a_{\text{sl}}^d$ from OFFSET e AMPLITUDE fit	OFFSET	$A_D$
2011 D	$0.0307 \pm 0.0119$	0.0335	$0.0133 \pm 0.0056$	-0.0035
2011 U	$-0.0101 \pm 0.0137$	-0.0123	$0.0128 \pm 0.0065$	0.0189
2012 D	$0.0160 \pm 0.0074$	0.0167	$0.0131 \pm 0.0036$	0.0047
2012 U	$0.0028 \pm 0.0072$	0.0039	$0.0142 \pm 0.0035$	0.0122

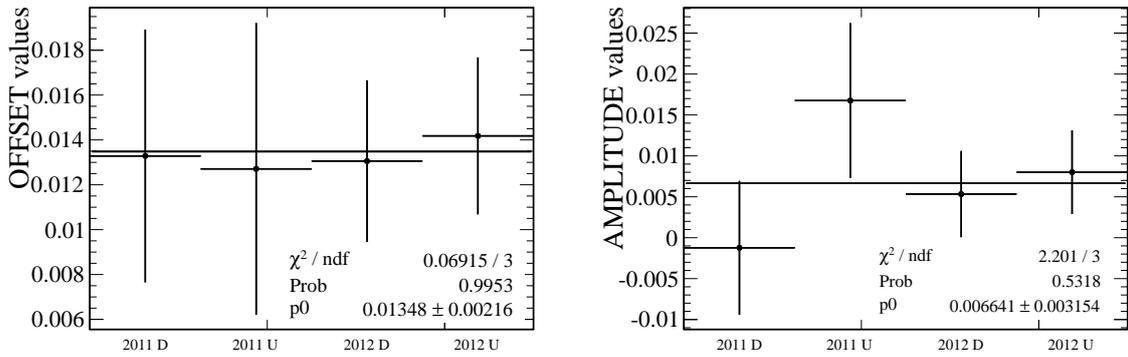


Figure C.2: OFFSET and AMPLITUDE measured values for the different  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  samples.

## C.2 Fiducial regions

As shown from previous studies [131], in two regions in the  $p$  versus  $p_x$  plane either the negative or the positive muons are bent out of the acceptance of the detector, resulting on a maximal asymmetry for a given magnet polarity. Moreover two regions, for values of  $|p_x|$  between 600 MeV/ $c$  and 1100 MeV/ $c$ , feature a large asymmetry depending on the polarity of the magnetic field. The latter effect is understood to be caused by the L0 trigger hardware: the four quadrants in the muon stations operate independently of each other, resulting in a trigger inefficiency for muons that cross from the left to right side in the muon detector. These muons have mostly  $|p_x|$  around 800 MeV/ $c$ . These effects depend on the charge of the muon, therefore they are expected to affect the asymmetry measurements on data acquired with a single magnet polarity, but they are expected to vanish when considering the full dataset (i.e. the average of the two magnet polarities in this analysis.). As crosschecks two fiducial cuts have been applied to the data samples, in order to exclude those regions expected to have large or maximal detection asymmetries

$$\begin{aligned} \text{fiducial cut 1} & : |p_x| < 0.317(p - 3400 \text{ MeV}/c), \\ \text{fiducial cut 2} & : |p_x| < 600 \text{ MeV}/c \text{ or } |p_x| > 1100 \text{ MeV}/c. \end{aligned}$$

Both these fiducial cuts are applied, and they reduce the size of the signal sample by about 10% (comparing the data sample with the full selection and the fiducial cuts applied to the data sample with the full nominal selection applied). The detection asymmetries as input in the time dependent fit have not been changed, given that the effect on the final result is expected to be small. The crosscheck is performed on the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  data sample, which gives the largest contribution to the final result. The four fits on the data sample split by magnet polarity and year are performed and the average is determined as for the nominal results. In Table. C.6 are reported the  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  values obtained on the single fits and on the total average, together with the nominal reference. Nominal reference and fiducial cuts applied differ only for the sample selection as explained above.

As additional crosscheck, the same set of fiducial cuts is required for the  $\pi_{\text{slow}}$  of the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays. In Tab. C.7 are reported the  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  values obtained on the single fits and on the total average, together with the nominal reference. Nominal reference and fiducial cuts applied differ only for the sample selection as explained above. In spite of the variations observed for single year and magnet polarity, the difference between the final averages of  $a_{\text{sl}}^d$  is below  $0.5\sigma$  (where only the statistical uncertainty of the fits is considered). In the case of  $A_{\text{P}}$  the variation of the single fits is smaller, and the variation on the average is below the statistical uncertainty. The values for the detection asymmetries used for this crosscheck are the nominal values, i.e. they are calculated on samples without the fiducial cuts applied. The agreement is expected to be better when using the detection asymmetries with the fiducial cuts included.

Table C.6:  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  [blinded] resulting values when applying or not the fiducial cuts on the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  data samples.

Fiducial Cuts applied		
Data Sample	$a_{\text{sl}}^d$	$A_{\text{P}}$
2011 Magnet Down	$0.0204 \pm 0.0056$	$-0.0247 \pm 0.0039$
2011 Magnet Up	$-0.0120 \pm 0.0065$	$-0.0123 \pm 0.0045$
2012 Magnet Down	$0.0178 \pm 0.0036$	$-0.0185 \pm 0.0026$
2012 Magnet Up	$0.0053 \pm 0.0035$	$-0.0085 \pm 0.0025$
All $B^0 \rightarrow D^- \mu^+ \nu_\mu X$	$0.0096 \pm 0.0022$	$-0.0148 \pm 0.0015$
Nominal		
Data Sample	$a_{\text{sl}}^d$	$A_{\text{P}}$
2011 Magnet Down	$0.0165 \pm 0.0054$	$-0.0210 \pm 0.0037$
2011 Magnet Up	$-0.0076 \pm 0.0062$	$-0.0126 \pm 0.0043$
2012 Magnet Down	$0.0163 \pm 0.0034$	$-0.0181 \pm 0.0024$
2012 Magnet Up	$0.0084 \pm 0.0033$	$-0.0103 \pm 0.0023$
All $B^0 \rightarrow D^- \mu^+ \nu_\mu X$	$0.0103 \pm 0.0021$	$-0.0149 \pm 0.0014$

Table C.7:  $a_{\text{sl}}^d$  and  $A_{\text{P}}$  resulting values when applying or not the fiducial cuts on the  $\pi_{\text{slow}}$  on the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  data samples.

Fiducial Cuts applied		
Data Sample	$a_{\text{sl}}^d$	$A_{\text{P}}$
2011 Magnet Down	$0.0353 \pm 0.0127$	$-0.0169 \pm 0.0086$
2011 Magnet Up	$-0.0102 \pm 0.0146$	$-0.0003 \pm 0.0100$
2012 Magnet Down	$0.0296 \pm 0.0080$	$-0.0204 \pm 0.0056$
2012 Magnet Up	$-0.0118 \pm 0.0078$	$-0.0026 \pm 0.0054$
All $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$	$0.0098 \pm 0.0048$	$-0.0107 \pm 0.0034$
Nominal		
Data Sample	$a_{\text{sl}}^d$	$A_{\text{P}}$
2011 Magnet Down	$0.0262 \pm 0.0118$	$-0.0092 \pm 0.0080$
2011 Magnet Up	$-0.0131 \pm 0.0135$	$-0.0027 \pm 0.0093$
2012 Magnet Down	$0.0178 \pm 0.0074$	$-0.0153 \pm 0.0052$
2012 Magnet Up	$-0.0019 \pm 0.0072$	$-0.0017 \pm 0.0050$
All $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$	$0.0076 \pm 0.0045$	$-0.0079 \pm 0.0031$

Table C.8: Run number ranges for the data sub-samples split according to the data taking conditions.

sub-sample name	run number range	
	2011	2012
run 1	Run number < 91000	Run number < 119500
run 2	91000 < Run number < 95000	119500 < Run number < 129000
run 3	95000 < Run number < 101000	Run number > 129000
run 4	Run number > 101000	

### C.3 Data taking periods

In order to check the stability of the  $a_{\text{sl}}^d$  result as function of the data taking conditions, the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  data sample is split in sub-samples according to the data taking period (identified with the *Run number*) and magnet polarity. The  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  sample is used, given that is providing most of the statistics for the measurement. In case that a measured value of  $a_{\text{sl}}^d$  significantly different from the average is found using the data of a specific run period, probably the data acquisition conditions of that run period are not completely understood. The data taking periods considered are reported in Tab.C.8 and the distribution of the run numbers of the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  data is displayed in Fig. C.3.

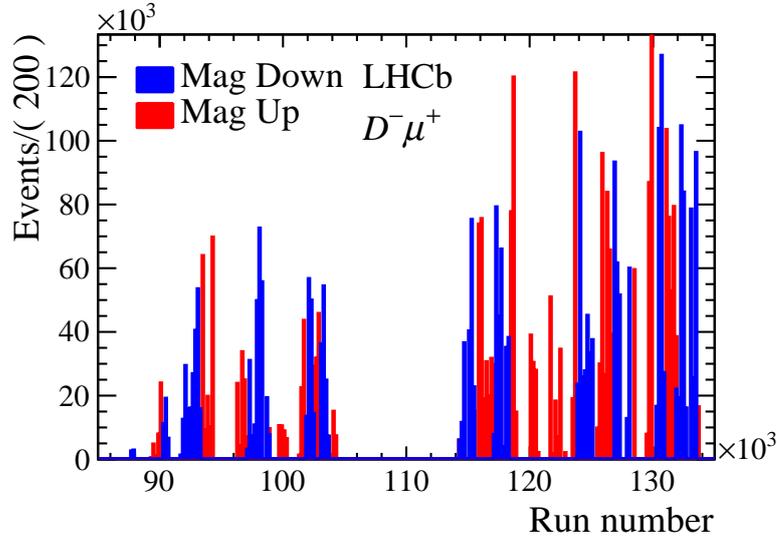


Figure C.3: Run numbers of the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  events recorded by LHCb in 2011 and 2012.

Fig.?? reports the measured values of  $a_{\text{sl}}^d$  and  $A_P$  obtained using the data acquired in the different data taking periods. The statistical uncertainty obtained from the fit is added in quadrature to the statistical uncertainty of the detection asymmetry measurements. No particular structure is observed, therefore the difference between the  $a_{\text{sl}}^d$  measured values with the different magnet polarities in 2011 cannot be ascribed to the data taking condition in a specific time range.

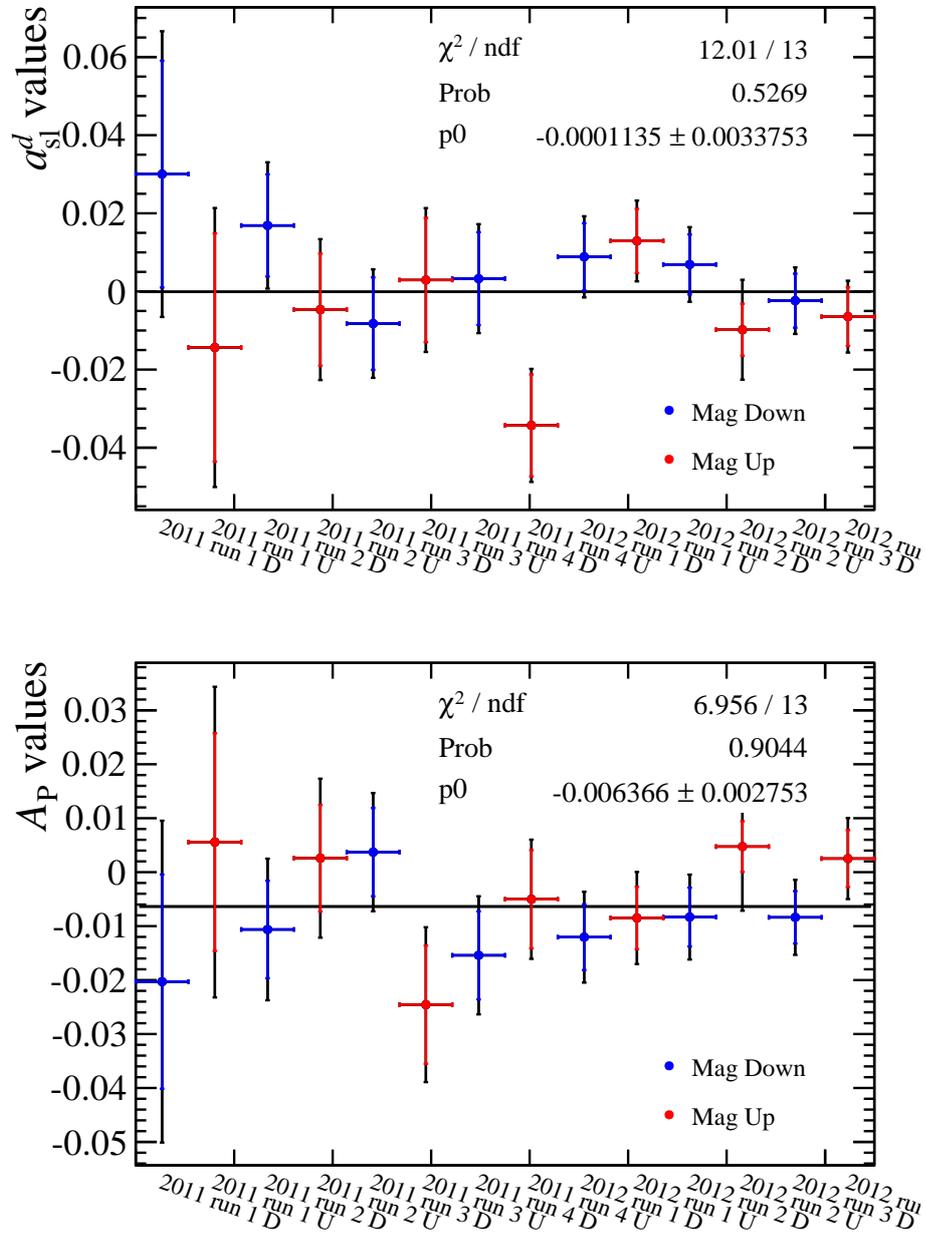


Figure C.4: Measured values for  $a_{sl}^d$  and  $A_p$  using data subsamples, defined according to the data taking periods and magnet polarity. The black error-bar show the sum in quadrature of the statistical uncertainty of the  $a_{sl}^d$  fit and the statistical uncertainty on the measured detection asymmetries. The colored error-bars indicate the statistical uncertainty on the  $a_{sl}^d$  fit only. No particular structure is observed.

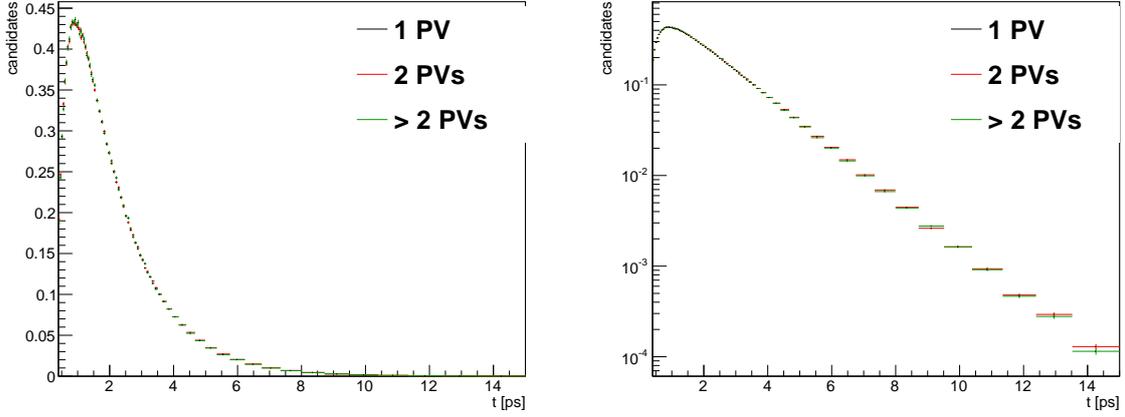


Figure C.5: Decay time distribution for different numbers of PVs, in nonlogarithmic and logarithmic scale.

## C.4 Number of Primary Vertices

In Fig. C.5 the  $B$  decay time distribution is reported for three sub-samples of the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  data sample, defined according to the number of primary vertices reconstructed in the events. Events with only one reconstructed PV are distinguished from events with two reconstructed PVs, and from events with three or more reconstructed PVs. Any difference of the  $B$  decay time distribution according to the number of PV is observed. On each of these subsamples the nominal fit is performed. In this case the polarity of the magnet are not separate and the value used for the detection asymmetry is measured on sub-samples of the control samples, defined according to the number of PVs as for the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  sample. The results found are found statistically compatible, and are shown in Fig. C.6.

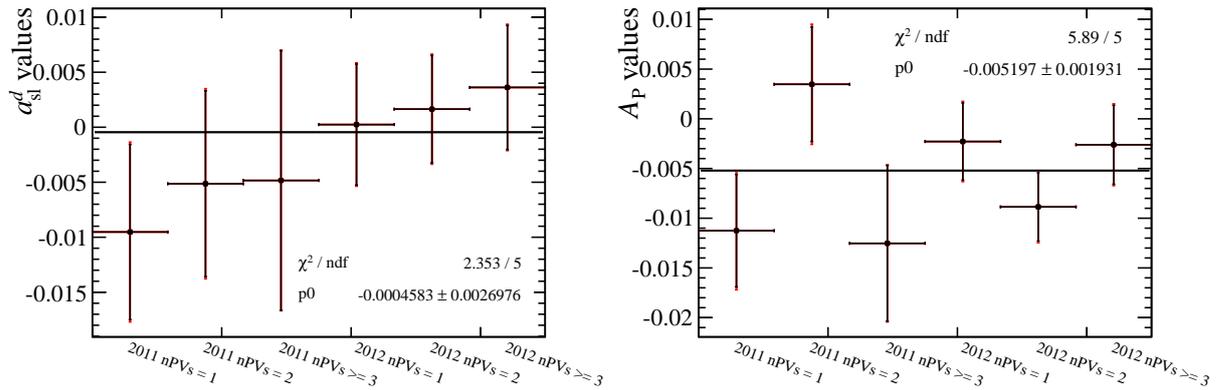


Figure C.6:  $a_{sl}^d$  and  $A_P$  measured values using the  $B^0 \rightarrow D^- \mu^+ \nu_\mu X$  dataset split according to the number of PVs.

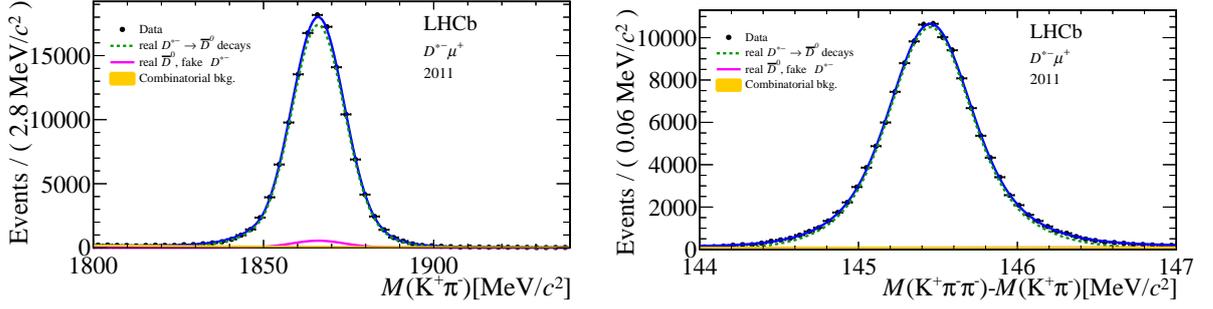


Figure C.7:  $D^0$  and  $(D^* - D^0)$  mass distributions in  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  decays collected in 2012.

## C.5 Random pions background in $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$ sample

Events with a slow pion not produced by a  $D^{*-}$  decay and associated to a real  $\overline{D^0}$  meson are a source of background included in the  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  data sample and not modeled in the nominal fit procedure. This background is estimated to be about 0.5% of the data sample. This crosscheck aims to verify the assumption made that this background component is negligible and does not affect the final measured value of  $a_{\text{sl}}^d$ .

By performing a 2-dimensional fit on the  $D^0$  mass and  $(D^* - D^0)$  mass difference distributions, the fraction of events that contribute to the peak in the  $D^0$  mass distribution, but are not produced by a  $D^*$  meson decaying to a  $D^0 \pi_{\text{slow}}$ , is extracted. Fig. C.7 show the projections of this fit on 2012 data, after applying the  $\mu - \pi$  kinematical weights (see Sec. 6.5.1). The mass fit range in  $(D^* - D^0)$  used is same as in the nominal analysis. The fraction of events that looks like signal in the  $D^0$  mass, but not in the  $(D^* - D^0)$  mass distribution, is estimated to be  $(3.0 \pm 0.4)\%$ . This fraction includes  $B^+$  and  $B^0$  decays, since it is determined by looking only at the daughters.

The selection cut on the  $(D^* - D^0)$  mass difference is relaxed, to allow to fit the time and charge asymmetry behavior of the decays populating the sidebands of the  $(D^* - D^0)$  mass difference, and the  $D^0$  mass signal region. The model used is the same as for the  $D^0$  mass sidebands background in the nominal analysis.

This new component of random-pion background is added to the nominal fit. The fraction is taken from the 2-dimensional fit to  $D^0$  and  $(D^* - D^0)$  mass distributions, and the time-charge asymmetry shape is taken from the simple fit on data. The alternative fit on data is performed for the 2012  $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu X$  dataset, keeping the magnet polarities separated as in the nominal strategy. The results are reported in Table C.10, together with the results of the nominal fits with the same detection asymmetries to be compared. The already negligible difference in the  $a_{\text{sl}}^d$  results from the single fits, is further reduced when calculating the average between the two magnet polarities.

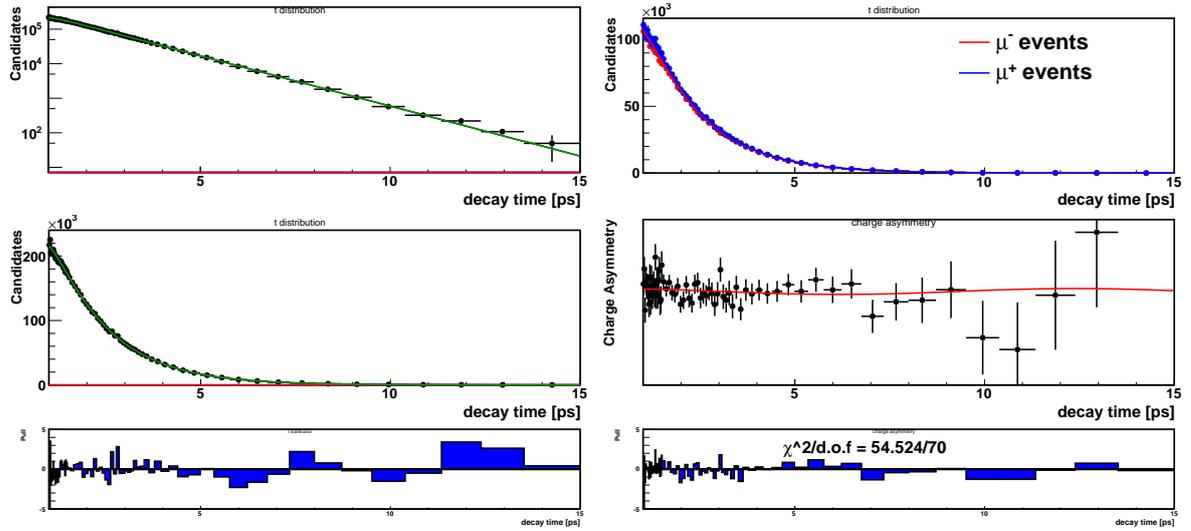


Figure C.8: Decay time and charge asymmetry of the random pion background. The model for the random-pions used is the same as the  $D^0$  mass sidebands background.

Table C.9: Time-dependent fit results when including the random pions background. Results obtained for the 2012 data, the results divided by magnet polarity are highly compatible; for this reason only this set of values has been used.

parameter	Value
$\alpha_{\text{random}\pi}$	$0.402 \pm 0.056$
$\Gamma_{\text{random}\pi}$	$0.672 \pm 0.002$
$t_{\text{shift,random}\pi}$	$0.171 \pm 0.119$
$A_{P,\text{random}\pi}$	$0.0036 \pm 0.0034$
$A_{D,\text{random}\pi}$	$0.013 \pm 0.002$

Table C.10: [blinded] Results of the fits, comparing the model without random pion background (nominal) and with random pion background.

Data Sample	$a_{\text{sl}}^d$	$A_P$
random pions included		
2012 Magnet Down	$0.0275 \pm 0.0077$	$-0.0246 \pm 0.0054$
2012 Magnet Up	$0.0111 \pm 0.0075$	$-0.0125 \pm 0.0052$
Average	$0.0193 \pm 0.0054$	$-0.0185 \pm 0.0037$
Nominal		
2012 Magnet Down	$0.0276 \pm 0.0075$	$-0.0245 \pm 0.0052$
2012 Magnet Up	$0.0110 \pm 0.0072$	$-0.0125 \pm 0.0050$
Average	$0.0193 \pm 0.0052$	$-0.0185 \pm 0.0036$





# Bibliography

- [1] LHCb, R. Aaij *et al.*, *Measurement of the semileptonic CP asymmetry in  $B^0 - \bar{B}^0$  mixing*, Phys. Rev. Lett. **114** (2015) 041601, arXiv:1409.8586.
- [2] ATLAS Collaboration, G. A. et al. *First observation of CP violation in the decays of Bs mesons*, Phys. Lett. . **B716** (2012) 1, arXiv:1207.7214.
- [3] CMS Collaboration, S. C. et al. *Observation of a new boson at a mass of 125GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235.
- [4] A. D. Sakharov, *Violation of cp invariance, c asymmetry, and baryon asymmetry of the universe*, Journal of Experimental and Theoretical Physics **5** 24.
- [5] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the  $2\pi$  decay of the  $k_2^0$  meson*, Phys. Rev. Lett. **13** 138.
- [6] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001.
- [7] ARGUS collaboration, H. Albrecht *et al.*, *Observation of  $B^0-\bar{B}^0$  mixing*, Phys. Lett. **B192** (1987) 245, arXiv:1408.0275.
- [8] D0 Collaboration, V. M. Abazov *et al.*, *Study of CP-violating charge asymmetries of single muons and like-sign dimuons in p pbar collisions*, Phys. Rev. **D89** (2014) 012002, arXiv:1310.0447.
- [9] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in  $D^0 \rightarrow K^- K^+$  and  $D^0 \rightarrow \pi^- \pi^+$  decays*, J. High Energy Phys. **07** (2014) 01419 p, Comments: 19 pages, 5 figures.
- [10] LHCb, R. Aaij *et al.*, *Measurement of the flavour-specific CP-violating asymmetry  $a_{sl}^s$  in  $B_s^0$  decays*, Phys. Lett. **B728** (2014) 607, arXiv:1308.1048.
- [11] C. Quigg, *Unanswered Questions in the Electroweak Theory*, Ann. Rev. Nucl. Part. Sci. **59** (2009) 505, arXiv:0905.3187.
- [12] M. S. Chanowitz, *Electroweak Symmetry Breaking: Unitarity, Dynamics, Experimental Prospects*, Ann. Rev. Nucl. Part. Sci. **38** (1988) 323.

- [13] D. H. Perkins, *Introduction to high energy physics*, (Cambridge University Press, Cambridge, 2000), 4th edition. .
- [14] G. Altarelli, *The standard model of particle physics*, [arXiv:0510281v1](#).
- [15] P. Skands, *Introduction to QCD*, [arXiv:1207.2389](#).
- [16] G. Altarelli, *The standard electroweak theory and beyond*, [arXiv:0011078v1](#).
- [17] K. A. O. et al. (Particle Data Group) *Chin. Phys. C* **38** (2014) 090001, [arXiv:http://pdg.lbl.gov/](#).
- [18] R. N. Mohapatra and G. Senjanovic, *Neutrino Mass and Spontaneous Parity Violation*, *Phys. Rev. Lett.* **44** (1980) 912.
- [19] M. Gell-Mann, P. Ramond, and R. Slansky, *Complex Spinors and Unified Theories*, *Conf. Proc.* **C790927** (1979) 315, [arXiv:1306.4669](#).
- [20] P. Higgs, *Broken symmetries, massless particles and gauge fields*, .
- [21] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, .
- [22] J. Charles et al. , *CKMfitter Group*, *Eur. Phys. J.* **C41** (2005) 1, [arXiv:0406184](#).
- [23] U. Nierste, *Three Lectures on Meson Mixing and CKM phenomenology*, 0904.1869.
- [24] K. A. et al. *B physics at the tevatron: Run ii and beyond*, [arXiv:0201071v2](#).
- [25] BaBar collaboration, *The babar physics book:physics at an asymmetric b factory*, .
- [26] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, *Phys. Rev. D* **2** (1970) 1285.
- [27] LHCb collaboration, R. Aaij *et al.*, *First observation of CP violation in the decays of Bs mesons*, *Phys. Rev. Lett.* **110** (2013) 221601, [arXiv:1304.6173](#).
- [28] A. J. Buras, *Flavour dynamics: Cp violation and rare decays*, [arXiv:0101336v1](#).
- [29] G. Colangelo, S. Durr, A. Juttner, L. Lellouch, H. Leutwyler, V. Lubicz, S. Necco and C. T. Sachrajda et al. , *webupdate: itpwiki.unibe.ch/flag*, *Eur. Phys. J. C* **71** (2011) 1695, [arXiv:1011.4408](#).
- [30] S. Aoki, Y. Aoki, C. Bernard, T. Blum, G. Colangelo, M. Della Morte, S. Drr and A. X. E. Khadra et al. [arXiv:1310.8555](#).
- [31] A. A. P. P. Gelhausen, A. Khodjamirian and D. Rosenth, *Decay constants of heavy-light vector mesons from qcd sum rules*, *Phys. Rev. D* **88** (2013) 1, [arXiv:1305.5432](#).

- [32] A. S. Dighe (MPI Munich), T. Hurth (CERN), C. S. Kim (Yonsei), T. Yoshikawa (Chapel Hill), *Measurement of the Lifetime Difference of  $B_d$  Mesons: Possible and Worthwhile?*, Nucl. Phys. B **624** (2002) 377, [arXiv:0109088v2](#).
- [33] M. Ciuchini, E. Franco, V. Lubicz, F. Mescia, C. Tarantino, *Lifetime Differences and CP Violation Parameters of Neutral B Mesons at the Next-to-Leading Order in QCD*, JHEP **0308** (2003) 031, [arXiv:0308029](#).
- [34] A. Lenz, *Selected topics in heavy flavour physic*, [arXiv:1404.6197](#).
- [35] A. Lenz and U. Nierste, *Numerical Updates of Lifetimes and Mixing Parameters of B Mesons*, [arXiv:1102.4274](#).
- [36] Heavy Flavor Averaging Group (HFAG), Y. Amhis *et al.*, *Averages of  $b$ -hadron,  $c$ -hadron, and  $\tau$ -lepton properties as of summer 2014*, [arXiv:1412.7515](#), updated results and plots available at: <http://www.slac.stanford.edu/xorg/hfag/>.
- [37] CLEO Collaboration, D. E. Jaffe *et al.*, *Bounds on the CP asymmetry in like sign dileptons from  $B^0\bar{B}^0$  meson decays*, Phys. Rev. Lett. **86** (2001) 5000, [arXiv:hep-ex/0101006](#).
- [38] BaBar collaboration, B. Aubert *et al.*, *Limits on the decay rate difference of neutral-B mesons and on CP, T, and CPT violation in  $B^0$ - $\bar{B}^0$  oscillations*, Phys. Rev. **D70** (2004) 012007, [arXiv:hep-ex/0403002](#).
- [39] BaBar collaboration, B. Aubert *et al.*, *Search for T, CP and CPT violation in  $B^0\bar{B}^0$  mixing with inclusive dilepton events*, Phys. Rev. Lett. **96** (2006) 251802, [arXiv:hep-ex/0603053](#).
- [40] BaBar Collaboration, J. P. Lees *et al.*, *Search for CP Violation in  $B^0$ - $\bar{B}^0$  Mixing using Partial Reconstruction of  $B^0 \rightarrow D^{*-} X \ell^+ \nu_\ell$  and a Kaon Tag*, Phys. Rev. Lett. **111** (2013), no. 10 101802, [arXiv:1305.1575](#).
- [41] Belle Collaboration, E. Nakano *et al.*, *Charge asymmetry of same-sign dileptons in  $B^0 - \bar{B}^0$  mixing*, Phys. Rev. **D73** (2006) 112002, [arXiv:hep-ex/0505017](#).
- [42] D0 Collaboration, V. M. Abazov *et al.*, *Measurement of the semileptonic charge asymmetry using  $B_s^0 \rightarrow D_s \mu X$  decays*, Phys. Rev. Lett. **110** (2013) 011801, [arXiv:1207.1769](#).
- [43] LHCb collaboration, R. Aaij *et al.*, *Search for direct CP violation in  $D^0 \rightarrow h^- h^+$  modes using semileptonic B decays*, Phys. Lett. **B723** (2013) 33, [arXiv:1303.2614](#).
- [44] D0 Collaboration, V. M. Abazov *et al.*, *Measurement of the semileptonic charge asymmetry in  $B^0$  meson mixing with the D0 detector*, Phys. Rev. **D86** (2012) 072009, [arXiv:1208.5813](#).

- [45] BaBar collaboration, J. P. Lees *et al.*, *Study of CP asymmetry in  $B^0 - \bar{B}^0$  mixing with inclusive dilepton events*, Phys. Rev. Lett. **114** (2015) 081801, [arXiv:1411.1842](#).
- [46] Riccardo Barbieri, Dario Buttazzo, Filippo Sala, David M. Straub, *Flavour physics and flavour symmetries after the first LHC phase*, [arXiv:1402.6677](#).
- [47] Monika Blanke, Andrzej J. Buras, Bjoern Duling, Stefania Gori, Andreas Weiler, *Delta F=2 Observables and Fine-Tuning in a Warped Extra Dimension with Custodial Protection*, JHEP **0903** (2009) 001, [arXiv:0809.1073](#).
- [48] M. Bauer (Mainz U. ), S. Casagrande (TU Munich), U. Haisch (Mainz U. ), M. Neubert (Mainz U. ), *Flavor Physics in the Randall-Sundrum Model: II. Tree-Level Weak-Interaction Processes*, JHEP **1009** (2010) 017, [arXiv:0912.1625](#).
- [49] Andrzej J. Buras, Jennifer Girrbach, *Towards the Identification of New Physics through Quark Flavour Violating Processes*, Rep. Prog. Phys. **77** (2014) 086201, [arXiv:1306.3775](#).
- [50] G. P. Gino Isidori, Yosef Nir, *Flavor physics constraints for physics beyond the standard model*, Ann. Rev. Nucl. Part. Sci. **60** (2010) 355, [arXiv:1002.0900](#).
- [51] Alexander Lenz, *Theoretical update of B-Mixing and Lifetimes*, [arXiv:1205.1444](#).
- [52] G. Borissov and B. Hoeneisen, *Understanding the like-sign dimuon charge asymmetry in  $p\bar{p}$  collisions*, Phys. Rev. **D87** (2013), no. 7 074020, [arXiv:1303.0175](#).
- [53] U. Nierste, *Effect of  $\Delta\Gamma$  on the dimuon asymmetry in B decays*, .
- [54] C. Bobeth and U. Haisch, *New Physics in  $\Gamma_{12^s}$ :  $(\bar{s}b)(\bar{\tau}\tau)$  Operators*, Acta Phys. Polon. **B44** (2013) 127, [arXiv:1109.1826](#).
- [55] C. Bobeth *et al.*, *On new physics in  $\Delta\Gamma_d$* , JHEP **1406** (2014) 040, [arXiv:1404.2531](#).
- [56] A. Lenz, *B-mixing in and beyond the Standard model*, [arXiv:1409.6963](#).
- [57] *LHC design report*, <http://ab-div.web.cern.ch/ab-div/Publications/LHC-DesignReport.html> .
- [58] *LHC design report, Chapter 18*, [lhc-collimation-project.web.cern.ch/lhc-collimation-project/files/Chapter18b.pdf](http://lhc-collimation-project.web.cern.ch/lhc-collimation-project/files/Chapter18b.pdf) .
- [59] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, International Journal of Modern Physics A **30** (2015) 1530022, [arXiv:1412.6352](#).
- [60] W. Herr, *Effect of PACMAN bunches in the LHC*, LHC Project Report **39** (1996) .
- [61] R. *Parton Distribution Functions*, proceedings of XXXI Physics in Collision, Vancouver, BC Canada, 28 Aug-1 Sep, 2011 [arXiv:0603175](#).

- [62] P. S. T. Sjöstrand, S. Mrenna, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, [arXiv:0603175](#).
- [63] LHCb Collaboration, R. Aaij *et al.*, *Measurement of  $\sigma(p\bar{p} \rightarrow b\bar{b}x)$  at  $\sqrt{s}=7$  TeV in the forward region*, Phys. Lett. B **694** (2010) 209, [arXiv:1009.2731](#).
- [64] LHCb Collaboration, R. Aaij *et al.*, *Production of  $j/\psi$  and  $v$  mesons in  $pp$  collisions at  $\sqrt{s} = 8$  tev*, J. High Energy Phys. **06** (2013) 064, [1304.6977](#).
- [65] E. Norrbin and T. Sjostrand, *Production and hadronization of heavy quarks*, Eur. Phys. J. **C17** (2000) 137, [arXiv:hep-ph/0005110](#).
- [66] *LHCb webpage*, <http://lhcb.web.cern.ch/lhcb/> .
- [67] E. Norrbin, *Heavy quark production asymmetries*, [arXiv:hep-ph/9909437](#).
- [68] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31.
- [69] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, [arXiv:hep-ph/0603175](#).
- [70] LHCb collaboration, R. Aaij *et al.*, *Measurement of the  $\bar{B}^0 - B^0$  and  $\bar{B}_s^0 - B_s^0$  production asymmetries in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*, Phys. Lett. **B739** (2014) 218, [arXiv:1408.0275](#).
- [71] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.
- [72] DELPHI collaboration, PA Aarnio, H Abie, P Abreu, W Adam, P Adrianos *et al.* , *The DELPHI detector at LEP*, Nucl Instrum Meth **303** (1991) 233.
- [73] M. Vesterinen, *Considerations on the LHCb dipole magnet polarity reversal*, Tech. Rep. LHCb-PUB-2014-006. CERN-LHCb-PUB-2014-006, CERN, Geneva, Apr, 2014. On behalf of the LHCb collaboration.
- [74] LHCb Collaboration, Barbosa-Marinho *et al.*, *LHCb VELO (Vertex Locator): Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001.
- [75] *LHCb Silicon Tracker webpage*, <http://lhcb.physik.uzh.ch/SiliconTracker/> .
- [76] LHCb Collaboration, Barbosa-Marinho *et al.*, *LHCb inner tracker: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2002. revised version number 1 submitted on 2002-11-13 14:14:34.
- [77] LHCb Collaboration, Barbosa-Marinho *et al.*, *LHCb outer tracker: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001.

- [78] LHCb Outer Tracker group, Arink *et al.*, *Performance of the LHCb Outer Tracker*, J. Instrum. **9** (2013) P01002. 30 p, Comments: 30 pages, 20 figures.
- [79] LHCb Collaboration, Amato *et al.*, *LHCb RICH: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2000.
- [80] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C 73** (2013) 2431, [arXiv:1211.6759](#).
- [81] LHCb Collaboration, Amato *et al.*, *LHCb calorimeters: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2000.
- [82] J. R. Harrison and G. Lafferty, *Radiation damage studies in the LHCb VELO detector and searches for lepton flavour and baryon number violating tau decays*, PhD thesis, Manchester U., 2014, Presented 16 05 2014.
- [83] LHCb Collaboration, Barbosa-Marinho *et al.*, *LHCb muon system: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001.
- [84] A. A. A. Jr *et al.*, *Performance of the lhcb muon system*, JINST **015** (2013) 1112, [arXiv:1211.1346](#).
- [85] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, [arXiv:1211.3055](#).
- [86] M. T. Schiller, S. Hansmann-Menzemer, and N. Herrmann, *Track reconstruction and prompt  $K_S^0$  production at the LHCb experiment*, PhD thesis, Heidelberg U., Heidelberg, 2011, Presented 12 Jul 2011.
- [87] W. Hulsbergen, *The global covariance matrix of tracks fitted with a kalman filter and an application in detector alignment*, [arXiv:0810.2241](#).
- [88] C. Bozzi *et al.*,  *$\Delta m_d$  measurement in the  $B_d \rightarrow D^{*-} \mu^+ \nu$  and  $B_d \rightarrow D^- \mu^+ \nu$  channels*, LHCb-ANA-2013-079.
- [89] LHCb Collaboration, R. Aaij *et al.* New J. Phys. **15** (2013) 053021. 18 p.
- [90] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in  $D^0 \rightarrow K^- K^+$  and  $D^0 \rightarrow \pi^- \pi^+$  decays*, [arXiv:1405.2797](#), submitted to PLB.
- [91] M. Williams *et al.*, *The HLT2 Topological Lines*, Tech. Rep. LHCb-PUB-2011-002, CERN, Geneva, Jan, 2011.
- [92] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, [arXiv:physics/0503191](#).
- [93] J. Brehmer, J. Albrecht, and P. Seyfert, *Ghost probability: an efficient tool to remove background tracks*, Tech. Rep. LHCb-INT-2012-025. CERN-LHCb-INT-2012-025, eCERN, Geneva, Sep, 2012.

- [94] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.
- [95] M. Clemencic *et al.*, *The LHCb simulation application, GAUSS: design, evolution and experience*, J. Phys.: Conf. Ser. **331** (2011) 032023.
- [96] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [97] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [98] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.
- [99] Particle Data Group, K. Nakamura *et al.*, *Review of particle physics*, J. Phys. **G37** (2010) 075021.
- [100] C. Bozzi, *LHCb MC cocktails*, Twiki page.
- [101] P. Urquijo, *Semileptonic B decay branching fractions*, Unpublished internal note.
- [102] LHCb collaboration, R. Aaij *et al.*, *Measurement of  $\sigma(pp \rightarrow b\bar{b}X)$  at  $\sqrt{s} = 7$  TeV in the forward region*, Phys. Lett. **B694** (2010) 209, [arXiv:1009.2731](#).
- [103] R. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Computer Physics Communications **77** (1993) 219.
- [104] A. Keune, *Field map parametrisation: Current status*, 02.06.2010. .
- [105] S. Stahl, *Charge asymmetry workshop, hints from detector: tracking system*, 02.06.2010. .
- [106] S. Stahl, S. Hansmann-Menzemer, and J. van Tilburg, *Measurement of CP asymmetry in muon-tagged  $D^0 \rightarrow K^-K^+$  and  $D^0 \rightarrow \pi^-\pi^+$  decays at LHCb*, PhD thesis, Heidelberg U., Jun, 2014, Presented 23 Jul 2014.
- [107] J. de Vries, *Dissertation (expected in 2017)*, .
- [108] Xing, Zhou and Artuso, Marina, *Measurement of the semileptonic CP violating asymmetry  $a_{sl}^s$  in  $B_s^0$  decays and the  $D_s^+ - D_s^-$  production asymmetry in 7 TeV pp collisions*, PhD thesis, Syracuse U., Jun, 2013, Presented 22 May 2013.
- [109] LHCb, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015), no. 02 P02007, [arXiv:1408.1251](#).
- [110] C. Bozzi *et al.*, *Measurement of the semileptonic CP violating asymmetry  $a_{sl}^d$* , LHCb-ANA-2013-050.

- [111] Fermilab E653 Collaboration, K. Kodama and others, *Measurement of  $\Gamma(D^0 \rightarrow K^- \mu^+ \nu) / \Gamma(D^0 \rightarrow \mu + X)$  using a  $D^*$  tag*, Physics Letters B **336** (1994) 605.
- [112] Gavril A. Giurgiu, *B Flavor Tagging Calibration and Search for  $B_s$  Oscillations in Semileptonic Decays*, 2005.
- [113] LHCb collaboration, Aaij, R and others, *Observation of  $B_s^0$ - $\bar{B}_s^0$  mixing and measurement of mixing frequencies using semileptonic B decays*, Eur. Phys. J. C **73** (2013) 2655. 20 p.
- [114] Steve Baker and Robert D. Cousins, *Clarification of the use of chi-square and likelihood functions in fits to histograms*, Nuclear Instruments and Methods in Physics Research **221** (1984) 437.
- [115] F James, M Roos, *MINUIT - A system for function minimization and analysis of the parameter errors and correlations*, Computer Physics Communications **10** (1975) 343.
- [116] R. Brun (CERN) , F. Rademakers (NIKHEF, Amsterdam & Hewlett-Packard, Geneva), *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81.
- [117] Y. Xie, *sFit: a method for background subtraction in maximum likelihood fit*, arXiv:0905.0724.
- [118] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetries in the decays  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  and  $B^+ \rightarrow K^+ \mu^+ \mu^-$* , arXiv:1408.0978, submitted to Phys. Rev. Lett.
- [119] LHCb collaboration, R. Aaij *et al.*, *First evidence of direct CP violation in charmless two-body decays of  $B_s^0$  mesons*, Phys. Rev. Lett. **108** (2012) 201601, arXiv:1202.6251.
- [120] D0 Collaboration, V. M. Abazov *et al.*, *Measurement of direct CP violation parameters in  $B^\pm \rightarrow J/\psi K^\pm$  and  $B^\pm \rightarrow J/\psi \pi^\pm$  decays with  $10.4 \text{ fb}^{-1}$  of Tevatron data*, Phys. Rev. Lett. **110** (2013) 241801, arXiv:1304.1655.
- [121] LHCb collaboration, R. Aaij *et al.*, *Observation of the  $\Lambda_b^0 \rightarrow J/\psi p \pi^-$  decay*, JHEP **1407** (2014) 103, arXiv:1406.0755.
- [122] LHCb Collaboration, R. Aaij *et al.*, *Measurement of b-hadron production fractions in 7 TeVpp collisions*, Phys. Rev. **D85** (2012) 032008, arXiv:1111.2357.
- [123] LHCb collaboration, R. Aaij *et al.*, *Studies of beauty baryon decays to  $D^0 p h^-$  and  $\Lambda_c^+ h^-$  final states*, Phys. Rev. D **89** (2013) 032001. 19 p, Comments: 19 pages, 6 figures, 5 tables.

- [124] Belle Collaboration, A. Zupanc *et al.*, *Measurement of the Branching Fraction  $Br(\Lambda_c + - \rightarrow pK^- \pi^+)$* , Phys. Rev. Lett. **113** (2014) 042002.
- [125] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001.
- [126] LHCb Collaboration, R. Aaij *et al.*, *Precision measurement of CP violation in  $B_s^0 \rightarrow J/\psi K^+ K^-$  decays*, Phys. Rev. Lett. **114** (2014) 041801. 19 p, Comments: 6 figures.
- [127] LHCb, R. Aaij *et al.*, *Measurements of the  $B^+$ ,  $B^0$ ,  $B_s^0$  meson and  $\Lambda_b^0$  baryon lifetimes*, JHEP **1404** (2014) 114, arXiv:1402.2554.
- [128] LHCb collaboration, R. Aaij *et al.*, *Measurement of the  $\bar{B}^0 - B^0$  and  $\bar{B}_s^0 - B_s^0$  production asymmetries in pp collisions at  $\sqrt{s} = 7$  TeV*, Phys. Lett. B **739** (2014) 218. 24 p, Comments: Updated to v2 after journal review.
- [129] LHCb collaboration, Bediaga *et al.*, *Framework TDR for the LHCb Upgrade: Technical Design Report*, Tech. Rep. CERN-LHCC-2012-007. LHCb-TDR-12, CERN, Geneva, Apr, 2012.
- [130] J. van Tilburg and M. van Veghel, *cpt violation searches and prospects for lhcb*, Phys. Lett. **B742** (2015) 236, arXiv:1407.1269.
- [131] L. Darmé *et al.*, *Search for CP violation in  $D^0 \rightarrow K^- K^+, \pi^- \pi^+$  using semileptonic B decays on  $3 fb^{-1}$* , LHCb-ANA-2012-012.



# Acknowledgments

*The noblest pleasure is the joy of understanding.* (Leonardo da Vinci)

Sometimes all the words of a language are not enough to express the gratitude, especially if the gift received is an opportunity to learn and understand more. Nevertheless, I would like to thank

My supervisors Prof. Ulrich Uwer and Prof. Stephanie Hansmann-Menzemer, for all the corrections and advices that made me learn more and improve the quality of my work.

Mika Vesterinen, Jeroen van Tilburg, Jacco de Vries, for the work done together for the  $a_{s1}^d$  measurement, and for all I could learn from them.

My second referee, Prof. Norbert Herrmann, for the time and attention dedicated to this work.

The Outer Tracker team, with whom I had the chance to collaborate taking care of the drift time calibration procedure.

The alignment and calibration team, with whom I could develop the OT drift time calibration algorithm for Run-II.

Prof. Ulrich Uwer and the IMPRS-PTFS graduate school, for giving me the opportunity to join this group for my PhD in first place, and to spend a long period at CERN in second place. The latter was by far the most interesting and productive period of my PhD.

I am enormously thankful to my colleagues, from whom and with whom I learned so much in the last three years.

I will never be able to express my gratitude to my parents, family and friends, that have always been able to support me in the life of the last three years.

I cannot avoid to remember (and miss) Moritz Karbach. “Don’t give up” was simply the advice of a friend.

Since I am very thankful thinking about my last three years, and the gifts received are simply too big, such that the “debt” cannot be paid (not for sure in a page of a thesis), with this cross crossword puzzle (Fig. C.9) I try at least to make you smile for a moment. What I would like to tell you is “hidden” in the yellow row. You have just to fill the [definitions](#) below. No spaces between the words are included (but they should be easy to be understood).

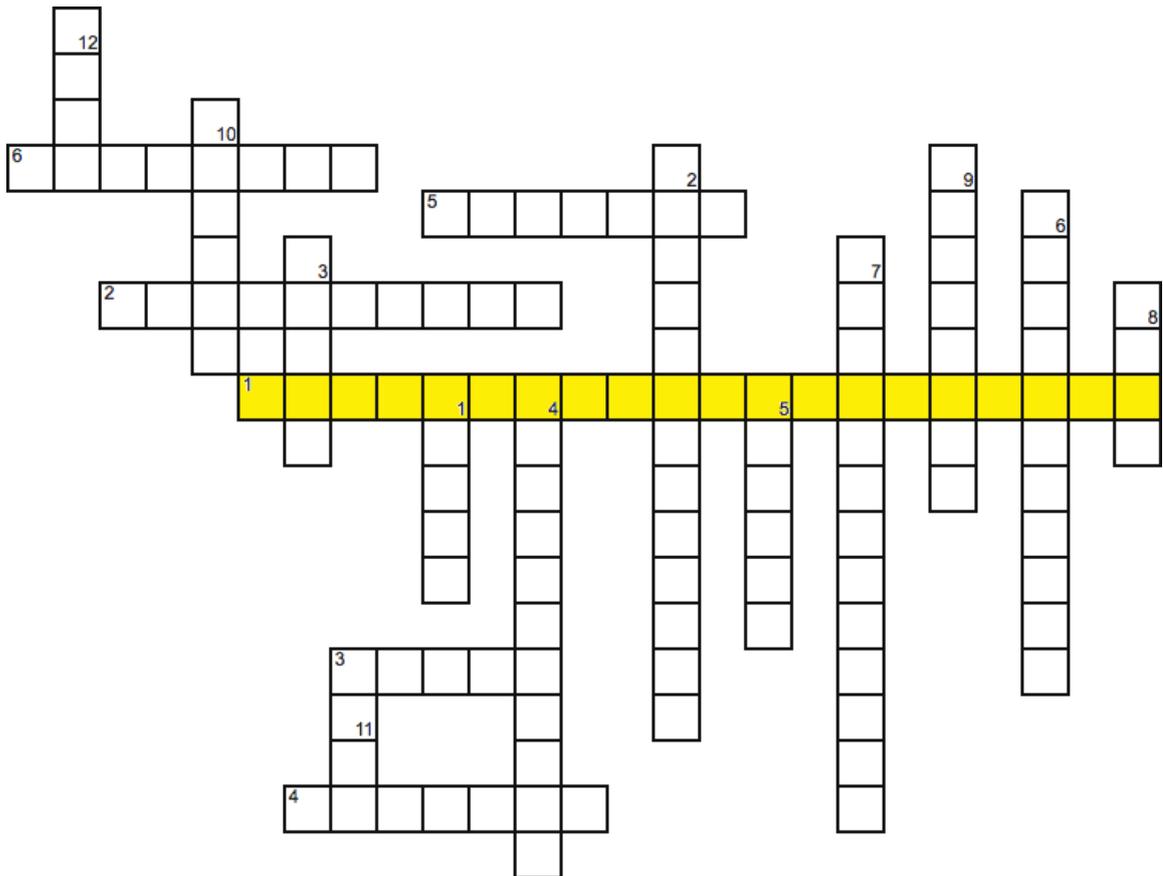


Figure C.9: Complete the definitions below to read the yellow row. The spaces between the words are not included. Each definition should be easy for the group of people I take the opportunity to thank with that definition. Not all the definitions are easy for everyone!

**Across**

1. [Main message](#)
2. [Where I go to ask for the Polizeiliches Führungszeugnis](#). Thanks to all the office mates I had, especially to Thomas Nikodem, for the constant help, advice and support in physics and daily office life. (I make fun of my insufficient German skills, of course not of the super efficient German office!)
3. [176 ... was the length of the reading effort](#). Thanks again to my referees for this thesis, for the time and attention they dedicated to this work.

4. [Often brings me back home.](#) Thanks a lot to my fantastic family. It is not only about the support I have always received, it is about my education first of all as human being, and later as physicist.
5. [LHCb, see Figure C.10 below.](#) No need to add anything else.. thanks to all the people I met and I work with in these three years!
6. [I wish I learned a bit how to do this in a good way.](#) I would like to thank my supervisor Prof. Ulrich Uwer for giving me the chance to make my PhD in this group, and for all the chances I got in these three years. For the corrections and advices that made me learn more, and become a better physicist. I would like to thank also Prof. Stephanie Hansmann-Menzemer and all my colleagues: I learnt a lot from you and from your questions and corrections!

### Down

1. [My favorite calibration \(informal name\).](#) Thanks again to Niels Tuning and all the Outer Tracker team. I am thankful for the incredibly collaborative and constructive work environment I have experienced with you. Thanks also for the shifts! they are always a learning adventure...
2. [OT and RICH calibration tasks \(they use SaveSets produced by Brunel\).](#) A special thank you to Alexandr Kozlinsky, the OT drift time calibration expert from whom I got the code and the knowledge. Thanks to Roel Aaij and Francesco Dettori, with whom the story continued for Run-II
3. ["Sorry for my ..." is the standard incipit of many of my emails.](#) Thanks to all the secretaries, especially to Ms. Claudia Wallenwein. For all the help and patience when receiving wrong forms, incomplete forms and many many questions.
4. [Neologism, funniest mistake of my thesis, found and corrected by Svende.](#) Thanks to Mika Vesterinen, Manuel Schiller, Paul Seyfert, Blake Leverington, Christian Färber, Katharina Kreplin, Francesca Dordei, Sebastian Wandernoth, Sevda Esen, Svende Braun, Michel De Cian, Jacco de Vries that helped me a lot improving this document.
5. [Nickname for the first person of the Heidelberg group \(and LHCb experiment\) I worked with.](#) I am truly thankful to whom teaches me something. I learned a lot from many people, probably from everyone I met, but I would like to say a special thank you to Sebastian Wandernoth, colleague and friend.
6. [Mika, Jeroen, Jacco, what did we measure in  \$B^0 - \bar{B}^0\$  mixing?](#) Thanks for the analysis we made, and for all I could learn from you. We should find something else to work together on! Thanks also to Sascha Stahl and Marian Stahl who contributed to the analysis with inputs and studies, to the Semileptonic working group and the reviewers of our analysis Angelo Di Canto and Prof. Ulrik Egede, for all the useful comments, requests and suggestions.
7. [Hosts the daily "game of doors", hosted dinners and birthday celebrations... it is where you can find me if I am not in the office or around.](#) Thanks to the friends in Heidelberg... no one as a friend makes me look at the reality in its completeness, and shows me what really matters in the life of a human being.

8. [Imperative addressed to the Heidelberg batch system \(when you want a job to be done\)](#). Thanks for all the help I got with coding issues (Paul, Thomas ...), and support of the computing (Alexey Zhelezov).
9. [Fun for Katharina, myself and others](#). Thanks for working together. An additional special thank you to Christoph Langenbruch, for the fitting algorithm he implemented in Heidelberg, that is the origin of the algorithm used in this analysis, and for the fruitful discussions I could always have with.
10. [Very good dish of our favorite restaurant in Crozet](#). Thanks to the friends in Geneva. It is simply amazing how the life path becomes approachable and beautiful with a friend.
11. [... does this distribution look like this? ... can you write this?](#) Thanks for all the questions, comments, constructive criticism. In physics and in life, coming from supervisors, colleagues, students and friends. Because a question that you can answer make you more confident in what you are doing, a question that you cannot answer forces you to make a further step.
12. [I borrowed it from Khoi, and used it for more than a year](#). Thanks to my flatmates and friends, Simona and Maria Concetta in Heidelberg and Khoi and Elisabetta in Geneva. Thanks for sharing with me the difficult and joyful moments of the daily life. The daily challenges show who we are and what our heart seeks.

02-Jun-2014 18:44:54		Fill #: 3582	Energy: 450 GeV	-	-
Experiment Status	ATLAS	ALICE	CMS	LHCb	
Instantaneous Lumi [(ub.s) <sup>-1</sup> ]	CALIBRATION	STANDBY	STANDBY	Amusing!	
BRAN Luminosity [(ub.s) <sup>-1</sup> ]	-	0.000	-	0.000	
Fill Luminosity (nb) <sup>-1</sup>	-	0.000	-	0.000	
BKGD 1	0.002	0.000	-	-	
BKGD 2	0.000	0.020	-	-	
BKGD 3	0.614	0.000	-	-	
LHCb VELO Position	OUT	Gap: 58.0 mm	NO BEAM		TOTEM: No info

Figure C.10: Useful for definition 5. across.