ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE
Institut de Physique de l'Energie et des Particules
Laboratoire de Physique des Hautes Energies

Exclusive Trigger Selections and Sensitivity to the B_s–$\overline{B}_s$ Mixing Phase at LHCb

Thèse de Doctorat

présentée à la Section de Physique de la Faculté des Sciences de Base
de l'Ecole Polytechnique Fédérale de Lausanne
pour l'obtention du grade de Docteur ès Sciences

par

Luis Fernández

Physicien diplômé de l'Université de Lausanne, Suisse

Jury

Président:	Prof. **Harald Brune**
Directeur de thèse:	Prof. **Olivier Schneider**
Expert interne:	Prof. **Mikhail Shaposhnikov**
Expert externe:	Dr **Hans Dijkstra**
Expert externe:	Dr **Gerhard Raven**

CH - Lausanne
2006

Abstract

THE LHCb experiment is one of the main experiments that will be hosted at the Large Hadron Collider (LHC) at CERN, in the area of Geneva, and is scheduled to start after the Summer 2007. The LHCb detector is a single-arm forward spectrometer dedicated to precise measurements of CP violation and rare decays in the b sector. The primary goal is to test the Standard Model description of flavor physics, and possibly to look for New Physics beyond it.

The event rate resulting from the LHC proton collisions will be tremendous, and the production of b hadrons will be copious, thus providing the required statistics for the study of flavor physics. However, not all the events are relevant to LHCb physics program, and a dedicated trigger is required. In particular, the High-Level Trigger will take the final decision whether to accept or discard an event before sending it to permanent storage, in a limited period of time. We present in this dissertation the implementation and design of the High-Level Trigger exclusive selections, and assess their performance in selecting the b and c decays of interest, based on a full Monte Carlo simulation.

The physics program of LHCb is vast, with the study of a large collection of b decays offering the possibility to investigate CP violation, and any deviation from the Standard Model. In this quest, the neutral B_s–$\overline{B}_s$ system plays a predominant role. The B_s–$\overline{B}_s$ mixing phase, denoted by ϕ_s, has not yet been measured, and represents a crucial probe of New Physics. This electroweak phase may be tested through theoretically clean $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions to CP eigenstates, by performing a time-dependent measurement of mixing-induced CP violation.

Among the B_s decays to pure CP eigenstates mediated by the $\bar{b} \to \bar{c}c\bar{s}$ transitions, the channel $B_s \to \eta_c\phi$ yields one of the best sensitivities to ϕ_s. We present the reconstruction and selection of $B_s \to \eta_c\phi$ events using a full Monte Carlo simulation, and we determine the characteristics relevant to a CP violation measurement with this channel.

The sensitivity of LHCb to the B_s–$\overline{B}_s$ mixing parameters is studied, using a fast parameterized Monte Carlo simulation. This simulation uses the outputs of the realistic full simulation to determine the statistical precision to the mixing observables. We consider the $B_s \to J/\psi\phi$ decay to an admixture of CP eigenstates requiring an angular analysis to disentangle the different CP components, as well as the $B_s \to \eta_c\phi$, $B_s \to D_sD_s$, and $B_s \to J/\psi\eta$ decays to pure CP eigenstates. Each of these signal samples are simultaneously fitted with a flavor-specific control sample, $B_s \to D_s\pi$, allowing the extraction of the B_s–$\overline{B}_s$ mixing frequency ΔM_s. We demonstrate that the sensitivity to ϕ_s is dominated by $B_s \to J/\psi\phi$ events, with the statistical precision required to uncover New Physics

effects. In comparison, the decays to pure CP eigenstates have an order of magnitude smaller event yields, but their contribution to the determination of ϕ_s, although small compared to $B_s \to J/\psi\phi$, is non negligible.

Keywords: CERN, LHC, LHCb, Standard Model, CP violation, B_s–$\overline{B}_s$ mixing phase.

Résumé

L'EXPERIENCE LHCb, en cours d'installation auprès du futur grand collisionneur de hadrons LHC (Large Hadron Collider) du CERN, dans la région de Genève, commencera après l'été 2007. Le détecteur LHCb est un spectromètre à un seul bras consacré aux mesures de précision de la violation de CP et des désintégrations dans le secteur des quarks b. Le but principal est de tester la description de la physique de la saveur décrite par le Modèle Standard, ainsi que de rechercher de la nouvelle physique au-delà de ces prédictions.

Le taux d'événements résultant des collisions de protons dans le collisionneur LHC sera faramineux, et la production des hadrons de type b sera copieuse, fournissant ainsi la statistique nécessaire à l'étude de la physique de la saveur. Tous les événements ne sont cependant pas pertinents pour le programme de physique de LHCb, et un système de déclenchement s'impose. En particulier, le système de déclenchement de haut niveau (HLT) prendra la décision finale quant à envoyer ou non un événement sur les disques d'enregistrement permanent, dans un temps imparti limité. Nous présentons dans ce mémoire l'implémentation et la conceptation des sélections exclusives du système de déclenchement de haut niveau, et nous évaluons leurs performances pour la sélection des désintégrations de hadrons de type b et c, sur la base d'une simulation de Monte-Carlo complète.

Le programme de physique de LHCb est vaste, avec l'étude d'un large éventail de désintégrations de hadrons de type b offrant la possibilité d'investiguer la violation de CP, et toute déviation du Modèle Standard. Dans cette quête, le système neutre B_s–$\overline{B}_s$ occupe une place prédominante. La phase du mélange B_s–$\overline{B}_s$, dénotée par ϕ_s, n'a pas encore été mesurée, et représente une sonde cruciale pour la recherche de nouvelle physique. Cette phase de l'interaction électrofaible peut être testée d'une manière théoriquement propre à l'aide des transitions de quarks du type $\bar{b} \to \bar{c}c\bar{s}$ en des états propres de CP, par une mesure en fonction du temps de la violation de CP induite par le phénomène de mélange.

Parmi les désintégrations contrôlées par les transitions $\bar{b} \to \bar{c}c\bar{s}$, la désintégration $B_s \to \eta_c\phi$ offre l'une des meilleures sensibilités à la phase de mélange ϕ_s de toutes les désintégrations en des états propres purs de CP. Nous présentons la recontruction et la sélection d'évènements $B_s \to \eta_c\phi$ sur la base d'une simulation de Monte-Carlo complète, et nous déterminons les caractéristiques pertinentes à la mesure d'une violation de CP à l'aide de ce canal.

La sensibilité de LHCb aux paramètres du mélange B_s–$\overline{B}_s$ est étudiée par le biais d'une

simulation rapide de type Monte-Carlo. Cette simulation s'appuie sur les résultats de la simulation complète afin de déterminer la précision statistique sur les paramètres du mélange. Nous considérons la désintégration $B_s \to J/\psi\phi$ en une superposition d'états propres de CP nécessitant une analyse angulaire pour permettre la distinction entre les différentes composantes de CP, ainsi que les désintégrations $B_s \to \eta_c\phi$, $B_s \to D_sD_s$, et $B_s \to J/\psi\eta$ en des états propres purs de CP. Chacun de ces canaux est analysé simultanément à un canal de contrôle, $B_s \to D_s\pi$, permettant l'extraction de la fréquence d'oscillation du système B_s–$\overline{B}_s$. Nous démontrons que la sensibilité à ϕ_s est dominée par le canal $B_s \to J/\psi\phi$, qui jouit de la statistique nécessaire pour mettre en évidence les effets d'une éventuelle nouvelle physique. En comparaison, les désintégrations en des états propres purs de CP ont des taux évènementiels d'un ordre de grandeur plus faible, mais leur contribution à la détermination de ϕ_s, bien que petite, est cependant non negligeable.

Mots-clés: CERN, LHC, LHCb, Modèle Standard, violation de CP, phase du mélange B_s–$\overline{B}_s$.

Acknowledgments

DURING the five years of research to prepare this Ph. D. thesis I benefited from advice and support of numerous people, both in and outside the physical society, and I wish to thank them.

I am indebted to Olivier Schneider, my supervisor, for his constant help and encouragements. His outstanding knowledge of physics considerably improved the quality of my work, as well as his scientific rigor and critiques.

My gratitude goes to our "big boss" at the LPHE, Aurelio Bay, for offering me the opportunity to prepare this thesis. I would also like to thank Thomas Schietinger for his guidance during my first steps in the LHCb trigger.

Thanks to all the LPHE members for the pleasant work atmosphere, and to our secretaries who are always keen to help. I reserve special thanks to our "superusers" (Christian, Hubert, Jean, Odie) for their technical support, and in particular to my first officemate, Hubert Degaudenzi, for initiating me to programing.

As for the members of the LHCb collaboration, I am particularly thankful to:

Benjamin Carron and Gerhard Raven, for their precious help and comments on the sensitivity studies.

Patrick Koppenburg, for our fruitful collaboration in the development of the High-Level Trigger.

Olivier Callot, for his availability and for advising me during my tracking work.

A few members of the trigger and reconstruction groups, in random order: Hans Dijkstra, Patrick Koppenburg, José-Angel Hernando, Frederic Teubert, Hugo Ruiz, Olivier Callot, Thomas Ruf, ...

People I have not yet mentioned, but who were of valuable help as well: Gloria Corti, Vanya Belyaev, Joel Closier, Patrick Robbe, Florence Ranjard, ..., and our spoke-person, Tatsuya Nakada.

I wish to thank the jury of the thesis: Harald Brune, Olivier Schneider, Mikhail Shaposhnikov, Hans Dijkstra and Gerhard Raven.

I want to thank my family for their support during my many years as a student. I do not forgot all my friends for the time we spent together, and for their friendship. Special thanks to Deep, Fifi, Juju, Mié, ...

Finally, I am deeply and truly thankful to the person who most significantly contributed to this thesis work by her presence and love, Timea. Thank you for always being there for me, from the very beginning.

August, 2006

Contents

Abstract **I**

Résumé **III**

Acknowledgments **V**

Introduction and Foreword **1**

1 CP Violation and the Beauty System **5**
1.1 Standard Model Basics . . . 6
1.2 Flavor Physics in the Standard Model . . . 9
1.2.1 Parameterization of the CKM Matrix . . . 10
1.2.2 The Unitarity Triangles . . . 11
1.2.3 Constraining the CKM Picture . . . 13
1.3 The Neutral B–$\overline{\text{B}}$ System . . . 15
1.3.1 Particle–Antiparticle Formalism . . . 15
1.3.2 B_q–$\overline{\text{B}}_\text{q}$ Mixing . . . 19
1.3.3 Mixing Parameters . . . 21
1.3.4 Time-Dependent Decay Rates . . . 23
1.3.5 CP Violation in the B_q–$\overline{\text{B}}_\text{q}$ System . . . 24
1.3.6 Effective Hamiltonians . . . 27
1.4 B_s Decays through $\bar{\text{b}} \to \bar{\text{c}}\text{c}\bar{\text{s}}$ Quark-Level Transitions . . . 30
1.4.1 Penguin Pollution . . . 31
1.4.2 B_s Decay Rates in $\bar{\text{b}} \to \bar{\text{c}}\text{c}\bar{\text{s}}$ Transitions . . . 32
1.4.3 Angular Analysis of the CP Components . . . 34
1.4.4 Control Channel . . . 39
1.5 Experimental Tests of the CKM Picture . . . 40
1.5.1 En Route Towards New Physics with B_s Decays? . . . 40

2 The LHCb Detector **45**
2.1 The Large Hadron Collider . . . 45
2.1.1 LHC Parameters . . . 46
2.1.2 LHCb Luminosity . . . 47
2.1.3 Production Cross-Sections . . . 48
2.2 The Large Hadron Collider Beauty Experiment . . . 49
2.2.1 Magnet . . . 51
2.2.2 Vertex Locator . . . 51
2.2.3 Trigger Tracker . . . 53

2.2.4 Tracking Stations . . . 54
2.2.5 RICH Counters . . . 55
2.2.6 Calorimeters . . . 56
2.2.7 Muon System . . . 57
2.2.8 Trigger System . . . 57

3 LHCb Monte Carlo Simulation 63
3.1 Simulation Framework . . . 64
3.1.1 MC Truth Data and Reconstructed Data . . . 67
3.2 Event Generation . . . 68
3.2.1 Proton-Proton Collisions . . . 68
3.2.2 Bottom Production . . . 69
3.2.3 Samples Generation . . . 70
3.3 Off-line Track Reconstruction . . . 71
3.3.1 Track Types . . . 72
3.3.2 Pattern Recognition and Track Fitting . . . 73
3.3.3 Long Tracks Performance . . . 75
3.4 Flavor Tagging . . . 75
3.4.1 Tagging Categories . . . 78
3.5 Data Samples and MC Technicalities . . . 79
3.5.1 Software Versions . . . 79
3.5.2 Data Samples . . . 80
3.5.3 HLT Technicalities . . . 81

4 Exclusive Trigger Selections 85
4.1 On-line and HLT Environment . . . 86
4.1.1 From Tracks to Particles . . . 86
4.1.2 Dedicated On-line Software . . . 88
4.1.3 Selection Criteria Definitions . . . 89
4.2 On-line Reconstruction Performance . . . 90
4.2.1 On-line VELO-TT Tracking . . . 90
4.2.2 On-line Track-Finding . . . 97
4.3 HLT Selections . . . 98
4.3.1 Final States . . . 101
4.3.2 Shared Composite Particles . . . 101
4.3.3 Exclusive HLT Selection Channels . . . 103
4.3.4 Exclusive HLT Selection Criteria . . . 104
4.3.5 Final State Filtering . . . 105
4.4 HLT Selection Performance . . . 111
4.4.1 HLT Selection Performance on Signal Events . . . 111
4.4.2 HLT Selections Performance on Minimum-Bias Events . . . 120
4.5 Outlook of the Exclusive Trigger Selections . . . 125

5 The $B_s \to \eta_c \phi$ Event Selection 127
5.1 Annual Production . . . 127
5.1.1 Signal Sample Generation . . . 130
5.2 Particle Identification . . . 131
5.3 Algorithmic Sequence . . . 132
5.4 Selection Variables . . . 133

5.4.1 Plotting Selection Variable Distributions . . . 134
5.5 Event Selection Tunes . . . 135
5.6 Signal Event Yield . . . 145
5.6.1 Reconstruction and Selection Efficiencies . . . 145
5.6.2 Trigger Performance on Signal . . . 147
5.6.3 Annual Yield . . . 152
5.6.4 Flavor Tagging . . . 153
5.7 Background Studies . . . 154
5.7.1 Preselection on $b\overline{b}$ . . . 155
5.7.2 $b\overline{b}$ Background Level . . . 156
5.7.3 Specific Backgrounds . . . 158
5.7.4 Inclusive D_s background . . . 161
5.8 Reconstruction Quality and Selection Bias . . . 163
5.8.1 Mass and Momentum Resolutions . . . 163
5.8.2 Vertex Resolutions . . . 166
5.8.3 B_s Proper Time . . . 167
5.9 Full Monte Carlo Simulation Results for $B_s \to \eta_c\phi$ . . . 174

6 Sensitivity to B_s–$\overline{B}_s$ Mixing Parameters **175**
6.1 Experimental Aspects of the B_s Decays of Interest . . . 176
6.2 Toy Monte Carlo Simulation . . . 178
6.3 Likelihood Modelling . . . 180
6.3.1 Description of the Mass Model . . . 181
6.3.2 Description of the Angular Model . . . 183
6.3.3 Modelling of the Time-Dependent Decay Rates . . . 185
6.3.4 Final Likelihood Functions . . . 190
6.4 Extracted Parameters . . . 191
6.4.1 Fit Strategy . . . 191
6.4.2 Fit Results . . . 192
6.4.3 Scanned Parameters . . . 199
6.5 Outlook and Future Improvements . . . 202

Conclusion **205**

A Full MC Results **207**
A.1 $B_s \to J/\psi\phi$ Event Selection . . . 207
A.2 $B_s \to D_s\pi$ Event Selection . . . 209
A.3 $B_s \to D_sD_s$ Event Selection . . . 210
A.4 $B_s \to J/\psi\eta$ Event Selection . . . 211

B Full Sensitivity Results **213**
B.1 Results with the Nominal Parameters . . . 214
B.2 Results with other Parameters . . . 216

Bibliography **221**

Introduction and Foreword

SYMMETRY is a characteristic feature of Nature. For instance, it reveals itself when we look at our image in a mirror, and symmetry is intimately related to the aesthetic appeal of objects and persons. However, we obviously live in an asymmetric world: Nature has clearly made a preference between matter and antimatter. Everything surrounding us, such as our body or the planets, is made of matter. Even though antimatter can be observed (e.g. radioactive decays) or created (e.g. in accelerators), it cannot remain for a long period of time. This imbalance between matter and antimatter has intrigued physicists for quite some time, without any explanation accounting for the observed difference.

In physics, symmetry is a generalized concept that refers to the invariance of the physical properties and laws under given transformations. This is reflected by Noether's paper [1], which states that there is a one-to-one correspondence between a symmetry and a conservation law, provided the (differentiable) symmetry was generated by a local action (in the Lagrangian field theory meaning). This means that to every symmetry satisfying Noether's theorem, there is a conserved quantity or current. For instance, time invariance leads to energy conservation of the physical laws, the momentum and the angular momentum are conserved under spatial translations and rotations, respectively.

In particle physics, the concept of symmetry plays a central role. For instance, the symmetries of the Lagrangian describing the dynamics of the particles determine the properties of the particles found in nature, or in other words, each particle is an irreducible representation of the (non-broken) symmetry groups describing the Standard Model (SM). This principle can be further exploited: a supersymmetry (SUSY) can be hypothetically introduced to relate the fermion and boson sectors of the SM.

Physicists are thus keen to test if the underlying theory of a physical process conserves or violates a given symmetry. Besides the usual continuous transformations of space-time preserving the Minkowski interval of special relativity, namely translations, rotations, and Lorentz boosts (Poincaré group), there is a set of discrete quantum transformations also conserving the space-time interval. These local operators act on a quantum state and they are the charge conjugation C (the transformation of a particle into its antiparticle), the parity transformation P (the inversion of the space component of the state), and the time reversal T (the inversion of the time component of the state). According to the CPT theorem [2, 3, 4], the CPT transformation is preserved in any Poincaré invariant local quantum field theory described by a Hermitian Hamiltonian (or Lagrangian). In other words, given $\mathcal{L} = \mathcal{L}^\dagger$, one has the invariance of the action:

$$S = \int \mathrm{d}^4x\, \mathcal{L}(x) = \int \mathrm{d}^4x'\, \mathrm{CPT}\, \mathcal{L}(x')\, (\mathrm{CPT})^{-1} = \mathrm{CPT}\, S\, (\mathrm{CPT})^{-1}\,.$$

In particular, this fundamental symmetry implies that masses and lifetimes of particles and antiparticles are equal, and their charges are equal and opposite. Any deviation

from an exact CPT symmetry would mean a violation of relativity. However, the discrete symmetries can be violated separately even if CPT is conserved.

The violation of the P symmetry was first observed in 1957 in nuclear β decays [5], and the weak interactions were found to maximally violate the P and the C symmetries. For instance, there are no left-handed antineutrinos. Despite these violations, it was believed that the product CP was preserved. However, in 1964, CP violation in $\mathrm{K}^0_\mathrm{L} \to \pi^+\pi^-$ was established [6]. CP violation was then observed outside the neutral kaon system, in the decays of $\mathrm{B_d}$ mesons. The first measurements were made in 2001 at the B factories, BABAR (at SLAC, Stanford, California, USA) and Belle (at KEK, Tsukuba, Japan), using $\mathrm{B_d} \to \mathrm{J}/\psi\mathrm{K}^0_\mathrm{S}$ decays [7, 8]. The description of CP violation, within the Standard Model of electroweak interactions with three fermion generations, naturally arises from the complex couplings between the quarks [9], the so-called Cabbibo–Kobayashi–Maskawa (CKM) matrix [9, 10]. All the CP violation measurements are so far compatible with the SM expectations.

The limited number of measurements in flavor physics and of CP violation, together with the uncertainties of the SM expectations, do not exclude the possibility of New Physics (NP), i.e. physics lying beyond the SM. These limitations make the study of b hadrons an exciting place where to look for NP with new sources of CP violation and flavor couplings. Besides particle physics, the study of CP violation is also of interest to cosmology. CP violation is one of the necessary conditions to generate an excess of matter over antimatter – in addition to baryon number violation and departure from thermal equilibrium – as pointed out by Sakharov [11]. Nevertheless, results from baryogenesis show that the observed cosmological baryon asymmetry $\mathcal{O}(10^{-10})$ cannot solely be generated by the CP violation present in the SM [12], and would require new sources of CP violation.

We are now at the beginning of a new era of the study of b physics. Whereas B factories (asymmetric $\mathrm{e^+e^-}$ colliders) operating at the $\Upsilon(4\mathrm{S})$ resonance only have access to $\mathrm{B_d}$ mesons, the experiments at hadron colliders will enable the study of strange $\mathrm{B_s}$ mesons with sufficiently sensitivity to further constrain the CKM picture. The long awaited measurement of the fast $\mathrm{B_s}$ mixing frequency (ΔM_s) will enable to determine the strength of this mixing and thus constrain the NP contributions, if any. The first significant results on ΔM_s started to appear this year (2006) by the experiments at the Tevatron collider (at Fermilab, Illinois, USA). After a mild upper bound from the DØ experiment [13], the CDF collaboration has announced the first measurement of the $\mathrm{B_s}$–$\overline{\mathrm{B}}_\mathrm{s}$ oscillation frequency [14]. This result is in agreement with the SM. However, given the large uncertainties affecting the SM predictions, there might still be limited deviations from the SM. Moreover, the NP contributions to the phase of the $\mathrm{B_s}$–$\overline{\mathrm{B}}_\mathrm{s}$ mixing (ϕ_s) may still be large, even with a SM value of ΔM_s. Finally, the presence of new CP-violating phases is as well essentially not constrained [15].

The Large Hadron Collider (LHC), scheduled to start during 2007, will collide protons at higher energies compared to the Tevatron. The production rates of b hadrons will open the study of rare b decays, and the search for CP violation in $\mathrm{B_s}$ mesons. The LHC (at CERN, Geneva) will host several experiments, among which the Large Hadron Collider beauty (LHCb) experiment dedicated to the study of rare b decays and CP violation. LHCb will significantly contribute to the precise measurement of ΔM_s, but will also provide the first measurements of CP violation in the $\mathrm{B_s}$ system. The experimental situation of $\mathrm{B_s}$ mesons will drastically change, as LHCb will be able to measure the yet unconstrained CP-violating phase associated with the $\mathrm{B_s}$–$\overline{\mathrm{B}}_\mathrm{s}$ mixing. This phase, ϕ_s, is

expected to be negligibly small in the SM since it is doubly Cabibbo-suppressed. The B_s system thus represents an exciting place to look for New Physics.

The LHCb experiment is currently in its final development phase, before the first collisions will become available. The design and the performance of the detector have been optimized using realistic Monte Carlo simulations.

This thesis is divided into following chapters:

Chapter 1 sets the theoretical framework of CP violation and B_q mixing, with a special care for the B_s system, and in particular $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions. The latter enable the determination of the B_s–$\overline{B}_s$ mixing phase through the measurement of mixing-induced CP violation. Moreover, $\bar{b} \to \bar{c}c\bar{s}$ amplitudes are to a very good approximation dominated by single weak phase, such that the corresponding CP asymmetry directly measures the B_s–$\overline{B}_s$ mixing phase ϕ_s.

Chapter 2 describes the detector setup and its characteristics in the LHC environment.

Chapter 3 introduces the simulation framework used to assess the physics performance of LHCb.

Chapter 4 shows the design and performance of the exclusive High-Level Trigger selections. The trigger is one of the most challenging parts of the experiment given the high collision rates, the timing constraints, the storage limitations, and the importance to keep as many interesting events as possible.

Chapter 5 illustrates the reconstruction and the event selection of the $B_s \to \eta_c(\to h^+h^+h^-h^-)\phi\ (\to K^+K^-)$ decay channel using the full Monte Carlo simulation. The annual event yield, the background levels, the tagging performance, and the resolutions involved will be presented. Moreover, a dedicated study of the proper time is provided. The $B_s \to \eta_c\phi$ channel is caused by $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions to pure CP-even eigenstates, and they may therefore be used to probe the B_s–$\overline{B}_s$ mixing phase. The results obtained from the full simulation will then be used in the determination of the sensitivity to ϕ_s.

Chapter 6 gives the performance of LHCb to the B_s–$\overline{B}_s$ mixing parameters, and in particular to ϕ_s, considering several channels. The sensitivities are assessed by means of a fast (toy) simulation, using as inputs results of the full Monte Carlo: yields, background levels, proper time acceptance, per-event proper time errors, resolutions and flavor tagging. The channels considered are the $B_s \to \eta_c\phi$, $B_s \to D_sD_s$, $B_s \to J/\psi\eta$ decays to CP-even eigenstates, as well as the $B_s \to J/\psi\phi$ decay, an admixture of CP eigenstates. All these channels are simultaneously fitted with the flavor specific $B_s \to D_s\pi$ decay channel used as control channel for the B_s–$\overline{B}_s$ oscillations. Several scans of the physical parameters will be performed in order to explore a few interesting points of the parameter space.

The ordering of the chapters is not chronological and was chosen to propose an original journey through the LHCb experiment. It starts from the theoretical motivations followed by the description of the experimental apparatus required to enable the measurements of the physics purposes. Then the feasibility of the experiment is tested using the Monte Carlo simulation. Before being able to perform any physics analysis, the interesting events have to be reconstructed on-line and selected by the trigger system, to

finally become available for off-line physics analysis. Then the $B_s \to \eta_c\phi$ selection illustrates how to extract signal events from the rough background, and finally the off-line selected events can be used to extract the physical information we are interested in, namely the B_s–$\overline{B}_s$ mixing parameters.

The contents of this thesis constitute selected parts of the five years I worked on the LHCb experiment. After my graduation in theoretical particle physics in 2001, I decided to leave the world of branes and extra dimensions to join the experimental physics at the IPHE (Institut de Physique des Hautes Energies, Université de Lausanne), and then the LPHE (Laboratoire de Physique des Hautes Energies, EPFL). During my first year, I worked on the data acquisition system for some readout chips and electronics. At the same time, I started working at the Monte Carlo generator level. I became the QQ event generator librarian until the replacement by its current successor, EVTGEN.

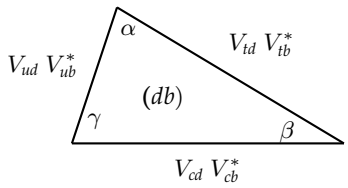
During the Fall 2002, I was freed from my hardware tasks and I could begin looking at physics analyses. I performed the $B_s \to \eta_c\phi$ event selection [16], which was finalized for the Reoptimized Technical Design Report [17] in the Summer 2003. After the analysis of $B_s \to \eta_c\phi$, the natural continuation was the study of its performance for the determination of CP violation. I therefore started at the end of 2003 to look at the sensitivity to the B_s–$\overline{B}_s$ parameters with $\bar{b} \to \bar{c}c\bar{s}$ transitions to pure CP eigenstates, and then extended the study to $B_s \to J/\psi\phi$ decays.

At the same time, I decided to work for the trigger group, with as primary goal the High-Level Trigger. I started by studying the retention rate of the generic Level-1 trigger [18], work which was finalized in the beginning of 2004. From the Spring 2004, I joined the effort in the design and implementation of the High-Level Trigger. As the border line between the on-line tracking and the trigger is thin, I naturally contributed to the tracking with the improvement of the VELO-TT tracking.

This thesis summarizes my work on the High-Level Trigger, the $B_s \to \eta_c\phi$ off-line analysis, and the determination of the B_s–$\overline{B}_s$ mixing parameters at LHCb using $B_s \to \eta_c\phi$, $B_s \to D_sD_s$, $B_s \to J/\psi\eta$, $B_s \to J/\psi\phi$, and $B_s \to D_s\pi$ decays.

Chapter 1

CP Violation and the Beauty System



We introduce in this chapter the Standard Model of elementary particles, and the corresponding description of CP violation within the CKM picture. The formalism of neutral B–$\overline{B}$ mixing is described, and all the formulae needed for the study of mixing-induced CP violation with B_s decays proceeding through $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions are derived. Finally, the experimental status of CP violation in the SM is given, with a small excursion towards New Physics in the B_s mixing.

THE Standard Model (SM) of particle physics is a quantum field theory describing the fundamental particles that make up all matter, as well as their dynamics through the electromagnetic, weak, and strong forces. However, the Standard Model is not a complete theory of fundamental interactions, primarily because it does not accommodate the fourth interaction, namely gravity. Moreover, the SM has many parameters whose values are not explained, such as the (running) coupling constants or the particle masses and their hierarchy. The number of particle generations is also not explained. Finally, the predicted Higgs boson, introduced to generate the mass of the fundamental particles, has not been detected yet. The Higgs field is furthermore the only non-gauge boson of the SM. All these arguments naturally raise the question whether the SM describes the reality, or if it is just some low energy effective theory of some more fundamental theory, or New Physics (NP).

The study of CP violation provides excellent tests of the SM. Furthermore, the decays of heavy bottom particles offer some theoretically clean measurements of CP violation, with a small influence from the strong interaction. CP violation is expected in many b decays, such that a systematic comparison with the SM predictions is eased. Consequently, the b sector provides fine insights into the flavor structure of the electroweak interactions in order to (over-) constrain the SM description of CP violation, and thus to look for NP.

In this chapter we present the basics of the Standard Model and its flavor structure in Sections 1.1 and 1.2. Then in Section 1.3, the formalism of the B_q–$\overline{B}_q$ system is introduced and the description of CP violation in the b sector is discussed, with a particular devotion to $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions in B_s decays in Section 1.4. In Section 1.5 a survey of the

latest experimental results is presented. Finally, Section 1.5.1 is a brief excursion towards model-independent New Physics in B_s mixing.

There are many textbooks and reviews on CP violation. Among those, we chose to base this chapter on References [19, 20, 21]. Moreover, as pointed out in [19, 21], we should be careful with the phase conventions and the signs of the different quantities used to describe CP violation. The CP transformation involves an arbitrary phase as a result of the invariance of the Lagrangian densities under rephasing of the quark fields. In particular, the phase of the B_q–$\overline{B}_q$ mixing amplitude does depend on the phase convention and is therefore *not* a physical observable. However, after a convention choice, the B_s–$\overline{B}_s$ phase can be assimilated to CKM elements and thus represents a physical observable that we will denote by ϕ_s. In this dissertation, we assume the phase conventions and use the notations from [21, 22].

1.1 Standard Model Basics

The Standard Model (SM) of particle physics is a theory describing the elementary particles and their interactions through the strong, weak, and electromagnetic fundamental forces. This theory is consistent with both quantum mechanics and special relativity, and its expectations are so far in good agreement with the experimental tests. The only missing particle to be discovered in the spectrum is the Higgs boson, and gravity is not accommodated in the SM.

The SM combines the theory of strong interactions, Quantum Chromo Dynamics (QCD), and the Glashow–Salam–Weinberg electroweak theory [23, 24, 25]. The SM is characterized by the gauge group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, spontaneously broken to $SU(3)_C \otimes U(1)_Q$. This Spontaneous Symmetry Breaking (SSB) of the $SU(2)_L$ group is obtained by the introduction of a non-gauge boson, the Higgs, and its corresponding asymmetric ground state. The propagators of the strong interaction are the eight gluons G, for the weak interaction the mediating bosons are the massive $W^{\pm}$ and Z^0 particles, and the photon γ is the carrier of the electromagnetic interaction. As the SM is fairly discussed in any textbook on particle physics, we will hereafter just summarize the spectrum of the theory, the electroweak interactions and the SSB.

The gauge fields and elementary particles in the SM are listed hereafter:[2]

Gauge fields

- $SU(3)_C$: 8 gluons $G^k_{\mu\nu}$, $k = 1 \dots 8$. The subscript C stands for the color quantum numbers.
- $SU(2)_L$: 3 W^a_ν bosons, $a = 1, 2, 3$. The subscript L stands for left-handed.
- $U(1)_Y$: B_μ boson. The subscript Y stands for the hypercharge.

Non-gauge fields

- There are three generations of quarks, each generation with an up-type quark ($U' \in \{u', c', t'\}$) and a down-type quark ($D' \in \{d', s', b'\}$).[3] The up-type quarks have electric charge $Q = +2/3$, whereas the down-type quark have electric charge $Q = -1/3$. All quarks are spin-$1/2$ fermions. Furthermore, each quark

[2] Greek letters denote space-time (Lorentz) indices.

[3] The primes denote the weak eigenstates; as we shall see later, we will denote the mass eigenstates without the primes.

comes with an antiquark with opposite quantum numbers. The mass hierarchy is such that the quarks are classified with increasing mass within each up- or down-types. Note that the t quark is the most massive of all, while the b quark is the second heavier. The different flavors are called up (u), charm (c), top (t, or truth), down (d), strange (s), and bottom (b, or beauty). The quark fields (for each generation) are grouped in two components, characterized by their helicity:

  - left-handed (L) doublets of $\mathrm{SU(2)_L}$: $\begin{pmatrix} \mathrm{U}' \\ \mathrm{D}' \end{pmatrix}_{\mathrm{L}}$;
  - right-handed (R) singlets of $\mathrm{SU(2)_L}$: $(\mathrm{U}')_{\mathrm{R}}$ and $(\mathrm{D}')_{\mathrm{R}}$.

- As for the quarks, leptons are grouped in three generations. Each generation has a neutrino ν_ℓ ($\nu_\ell \in \{\nu_{\mathrm{e}}, \nu_\mu, \nu_\tau\}$) with no electric charge, and a charged lepton ℓ^- ($\ell^- \in \{\mathrm{e}^-, \mu^-, \tau^-\}$). In addition, each particle has its antiparticle. The lepton fields are grouped in left-handed and right-handed representations:

  - doublets $\begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix}_{\mathrm{L}}$;
  - singlets $(\ell^-)_{\mathrm{R}}$.

  The absence of any right-handed neutrino field leads to the non-existence of neutrino masses and to the conservation of the individual lepton flavors. However, observations suggest the existence of neutrino masses, thus requiring a modified neutrino sector.

- Scalar Higgs field doublet: $\Phi(x^\mu) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$.

There exist several hundreds of other observed particles, but they all come as bound states of quarks, called hadrons, which are subdivided into:

- Mesons are built from a quark–antiquark pair with opposite color. For instance, the $\mathrm{B_s}$ meson has a ($\mathrm{\bar{b}}$ s) quark content, where $\mathrm{\bar{b}}$ denotes the anti b quark.

- Baryons are made of three quarks of different color. For instance, protons and neutrons are made of different combinations of up and down quarks.

In the SM, the electromagnetic and weak interactions are unified into a single electroweak theory, which is based on the SSB scheme:

$$\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y} \xrightarrow{\text{SBB}} \mathrm{U(1)_Q}\,,$$

where $\mathrm{U(1)_Q}$ is identified with the gauge group of the electromagnetic interaction. The SSB is induced by the Higgs potential:

$$V(\Phi^\dagger\Phi) = -\mu^2(\Phi^\dagger\Phi) + \lambda(\Phi^\dagger\Phi)^2 = -\mathcal{L}_{\mathrm{Higgs}},$$

where $\mathcal{L}_{\mathrm{Higgs}}$ denotes the self-interaction Higgs Lagrangian density, without the usual kinetic term. The asymmetric ground state Φ_0 is chosen such that the lower component of Φ acquires a vacuum expectation value v (complex number), constant over the whole Minkowski space. The minimum of the potential and the ground state are then given by:

$$\langle\Phi^\dagger\Phi\rangle = \frac{v^2}{2} = \frac{\mu^2}{2\lambda}\,, \quad \Phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}.$$

The price to pay for the disappearance of the $\mathrm{SU(2)_L}$ symmetry is the introduction of an hypothetical new field, which is the only boson in the SM which is not a gauge field. However, the remarkable feature of the SBB is that the couplings of the different fields to the Higgs field generate the mass of the particles given in terms of v. This is the so-called Higgs mechanism.

After spontaneous symmetry breaking, we may parameterize the Higgs field as:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \longrightarrow \begin{pmatrix} G^+ \\ \frac{1}{\sqrt{2}}(v + H^0 + iG^0) \end{pmatrix},$$

where H^0 is the physical Higgs field. In the unitary gauge, the G^0 is the Goldstone boson to be absorbed in the longitudinal component of the Z^0 boson, and G^+ (and the conjugate G^-) is the Goldstone boson to become the longitudinal component of the W^+ boson (and of W^-). The charged $\mathrm{W}^\pm$ are defined as $\mathrm{W}^\pm_\mu = (W^1_\mu \mp i W^2_{\mu)}/\sqrt{2}$. The neutral $\mathrm{SU(2)_L}$ gauge boson W^3_μ and the $\mathrm{U(1)_Y}$ gauge boson B_μ are mixed into the massless $\mathrm{U(1)_Q}$ gauge boson A_μ and another neutral massive gauge boson Z_μ. This rotation is characterized by the weak mixing Weinberg angle θ_W given in terms of the $\mathrm{SU(2)_L}$ coupling constant g and of the $\mathrm{U(1)_Y}$ coupling constant g', such that $g \sin\theta_\mathrm{W} = g' \cos\theta_\mathrm{W} = e$, where e is the $\mathrm{U(1)_Q}$ coupling constant (i.e. the electric charge). The orthogonal rotation is given by:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_\mathrm{W} & -\sin\theta_\mathrm{W} \\ \sin\theta_\mathrm{W} & \cos\theta_\mathrm{W} \end{pmatrix} \begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix}.$$

Finally, the masses of the gauge bosons after SBB are $m^2_A = 0$, $m^2_W = g^2 v^2/4$ and $m^2_Z = v^2(g^2 + g'^2)/4$. The Higgs mechanism thus gives mass to the gauge bosons in terms of fundamental parameters of the theory. Note that the coupling constant g and g' are not independent, and thus we have $m_W = m_Z \cos\theta_\mathrm{W}$.

The original massless fermions obtain their mass in Yukawa couplings (i.e. between Dirac and scalar fields) to the Higgs field. These couplings give mass to the charged leptons, and yield massless neutrinos since they have no interacting term with the Higgs field. The latter results from the non-existence of a right-handed neutrino singlet.

The couplings of the quarks to the Higgs field are not diagonal, for any weak basis. In order to diagonalize the Yukawa couplings, the so-called Cabbibo–Kobayashi–Maskawa (CKM) *unitary* matrix V_CKM is introduced [9, 10]. The CKM matrix connects the electroweak eigenstates $(\mathrm{d}', \mathrm{s}', \mathrm{b}')$ of the down-type quarks with their mass eigenstates $(\mathrm{d}, \mathrm{s}, \mathrm{b})$ through the following unitary transformation:

$$\begin{pmatrix} \mathrm{d}' \\ \mathrm{s}' \\ \mathrm{b}' \end{pmatrix} = \begin{pmatrix} V_\mathrm{ud} & V_\mathrm{us} & V_\mathrm{ub} \\ V_\mathrm{cd} & V_\mathrm{cs} & V_\mathrm{cb} \\ V_\mathrm{td} & V_\mathrm{ts} & V_\mathrm{tb} \end{pmatrix} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \\ \mathrm{b} \end{pmatrix} = V_\mathrm{CKM} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \\ \mathrm{b} \end{pmatrix}. \tag{1.1}$$

In this new basis, the charged current (CC) interactions mediated by the $\mathrm{W}^\pm$ bosons are purely left-handed, which is responsible for parity violation. Moreover, the lack of flavor neutral changing current (FCNC) at the tree level is due to the unitarity of the CKM matrix, since the Z_μ interaction terms are now flavor diagonal. The fact that the CKM matrix elements can a priori be complex numbers could – in principle – allow CP violation in the quark sector of the SM, provided we have complex physical phases.

1.2 Flavor Physics in the Standard Model

After spontaneous symmetry breaking, the charged-current interactions of quarks have the structure $\mathrm{D} \to \mathrm{U}\ \mathrm{W}^-$, where $\mathrm{U} \in \{\mathrm{u}, \mathrm{c}, \mathrm{t}\}$, $\mathrm{D} \in \{\mathrm{d}, \mathrm{s}, \mathrm{b}\}$. The coupling strength corresponds to the element V_{UD} of the CKM matrix. The $\mathrm{D} \to \mathrm{U}\ \mathrm{W}^-$ vertex and its CP conjugate are shown in Figure 1.1, where a complex phase could enable CP violation in the SM.

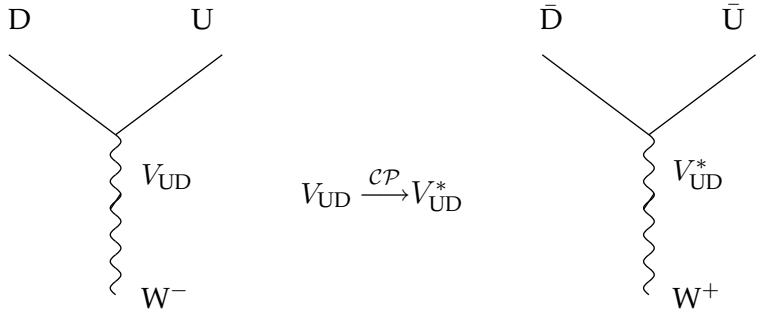


Figure 1.1: Charged-current quark-level interaction process $\mathrm{D} \to \mathrm{U}\ \mathrm{W}^-$ and its CP conjugate in the Standard Model.

Following the notations of Section 1.1, the charged-current interaction Lagrangian density in terms of the mass eigenstates reads:

$$\mathcal{L}^{\mathrm{CC}}_{\mathrm{int}} = -\frac{g}{\sqrt{2}} \left(\begin{array}{ccc} \bar{\mathrm{u}}_{\mathrm{L}}, & \bar{c}_{\mathrm{L}}, & \bar{\mathrm{t}}_{\mathrm{L}} \end{array} \right) \gamma^\mu V_{\mathrm{CKM}} \begin{pmatrix} \mathrm{d}_{\mathrm{L}} \\ \mathrm{s}_{\mathrm{L}} \\ \mathrm{b}_{\mathrm{L}} \end{pmatrix} \mathrm{W}^+_\mu \ + \ \mathrm{h.c.}, \tag{1.2}$$

where the γ^μ are the Dirac matrices, and the weak bosons only couple to purely left-handed spinors. The different interaction terms in $\mathcal{L}^{\mathrm{CC}}_{\mathrm{int}}$ show that each vertex has a coupling given by a CKM element, as illustrated above. Applying $CP\, \mathcal{L}^{\mathrm{CC}}_{\mathrm{int}}\, (CP)^{-1}$, the density remains the same if we have $V_{\mathrm{UD}} = V^*_{\mathrm{UD}}$. Thus CP violation is related to complex phases, but this is not sufficient. The quark fields could be rotated such that $\mathcal{L}^{\mathrm{CC}}_{\mathrm{int}}$ is CP invariant. As we shall see hereafter, this is not possible once there are more than two quark generations.

A general unitary matrix has N^2 independent parameters. In our case, we have the freedom of rephasing $2N$ quarks fields such that the number of independent parameters for a $N \times N$ unitary quark-mixing matrix is given by:

$$N^2 - (2N-1) = \underbrace{\frac{1}{2}N(N-1)}_{\text{Euler angles}} + \underbrace{\frac{1}{2}(N-1)(N-2)}_{\text{Complex phases}} = (N-1)^2.$$

We see that for $N = 2$ fermion generations we only have one rotation angle (Cabibbo angle). The matrix is thus real, preventing CP violation to occur. For $N = 3$ generations, we can parameterize the matrix with 3 Euler angles and 1 complex phase, which is the source of CP violation in the SM.

The study of CP violation is directly related to fundamental questions. The phase structure of the CKM matrix is connected to the Yukawa couplings with the Higgs field, and thus intimately related to the SSB. The number of generations is directly concerned,

since we need at least three generations to accommodate CP violation. From an experimental point of view, the observation of three quark families and thus of only one CP-violating phase makes it a predictive theory, that can be constrained by independent measurements of the CKM matrix elements.

1.2.1 Parameterization of the CKM Matrix

An extremely useful parameterization of the CKM matrix for phenomenological applications is that of Wolfenstein [26]. This parameterization corresponds to an expansion in terms of the small quantity $\sin\theta_{\rm C} \equiv \lambda$, where $\theta_{\rm C}$ is the Cabibbo angle, with additional three real parameters (ρ, η, A), which are of the order of unity ($\rho \approx 0.1$, $\eta \approx 0.4$, $A \approx 0.8$). Using the experimental result:

$$|V_{\rm us}|^3 \approx |V_{\rm cb}|^{3/2} \approx |V_{\rm ub}|\,, \quad \lambda \equiv |V_{\rm us}| \approx 0.22\,,$$

and the unitarity of the CKM matrix, Wolfenstein's parameterization reads:

$$V_{\rm CKM} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)\,, \tag{1.3}$$

where the series expansions are truncated at order λ^3. Clearly, η is CP-violating while the three (ρ, η, A) conserve CP. Furthermore, we need to consider next-to-leading order corrections in λ, typically for the study of the $\rm B_s$–$\overline{\rm B}_{\rm s}$ mixing phase. Starting from the parameterization of the CKM matrix advocated by the Particle Data Group (PDG) [27], namely the Chau–Kueng parameterization [28], an *exact* parameterization of the CKM matrix can be found in terms of Wolfenstein's expansion [29, 20]. The higher order corrections to (1.3) $\delta V_{\rm CKM}$ are:

$$\begin{pmatrix} -\frac{1}{8}\lambda^4 + \mathcal{O}(\lambda^6) & \mathcal{O}(\lambda^7) & 0 \\ \frac{1}{2}A^2\lambda^5\left[1-2(\rho+i\eta)\right] + \mathcal{O}(\lambda^7) & -\frac{1}{8}\lambda^4(1+4A^2) + \mathcal{O}(\lambda^6) & \mathcal{O}(\lambda^8) \\ \frac{1}{2}A\lambda^5(\rho+i\eta) + \mathcal{O}(\lambda^7) & \frac{1}{2}A\lambda^4\left(1-2(\rho+i\eta)\right) + \mathcal{O}(\lambda^6) & -\frac{1}{2}A^2\lambda^4 + \mathcal{O}(\lambda^6) \end{pmatrix}. \tag{1.4}$$

Note that there are *no* corrections to $V_{\rm ub} \equiv A\lambda^3(\rho - i\eta)$. If we introduce the generalized Wolfenstein's parameters [29]:

$$\bar\rho \equiv \rho\left(1-\frac{1}{2}\lambda^2\right)\,, \quad \bar\eta \equiv \eta\left(1-\frac{1}{2}\lambda^2\right)\,, \tag{1.5}$$

we simply have $V_{td} = A\lambda^3(1-\bar\rho - i\,\bar\eta) + \mathcal{O}(\lambda^7)$. These parameters will be used in the next subsection when rescaling the unitarity relations.

The existence of a CP-violating phase in order to accommodate CP violation in the SM can be translated to the Jarlskog parameter J_{CP} [30]:

$$J_{\rm CP} = |\Im(V_{i\alpha}V_{j\beta}V_{i\beta}^*V_{j\alpha}^*)|\,, \quad (i \neq j,\, \alpha \neq \beta)\,. \tag{1.6}$$

If the imaginary part of the products of the CKM matrix elements is different from zero, then CP is violated. The parameter J_{CP} thus measures the strength of CP violation in the SM, and CP violation requires all CKM matrix elements to be non-zero, when taking into account the unitarity of the CKM matrix. Using (1.3) and (1.4), we have that:

$$J_{\rm CP} \approx \lambda^6 A^2\eta(1-\lambda^2/2) + \mathcal{O}(\lambda^{10}) \sim 10^{-5}\,.$$

CP violation is therefore a small effect in the SM. However, new complex couplings arising from NP would constitute additional sources of CP violation.

1.2.2 The Unitarity Triangles

The SM quark mixing matrix is unitary, $V_{\rm CKM}^\dagger V_{\rm CKM} = V_{\rm CKM} V_{\rm CKM}^\dagger = \mathbb{1}$, which leads to a set of twelve equations given by $\sum_k V_{ki} V_{kj}^* = \delta_{ij}$. The six equations among the magnitudes express the normalization of the columns and rows of the CKM matrix, while the six relations involving both the magnitudes and phases express the orthogonality of different columns and rows. The six orthogonality relations can be represented in the complex plane [31], all having the same area, $2A_\Delta = J_{\rm CP}$ [32]. The angles and sides of these triangles are what we need to measure in order to test the CKM picture. The six unitarity triangles are artistically (i.e. not to scale) drawn in Figure 1.2, where we have indicated the size of each side using Wolfenstein's parameterization up to the leading non-vanishing order in λ. The orthogonality relations are:

$$\begin{array}{lccccccc}
(\mathrm{ds}) & V_{\rm ud}V_{\rm us}^* & + & V_{\rm cd}V_{\rm cs}^* & + & V_{\rm td}V_{\rm ts}^* & = & 0\,, \\
(\mathrm{sb}) & V_{\rm us}V_{\rm ub}^* & + & V_{\rm cs}V_{\rm cb}^* & + & V_{\rm ts}V_{\rm tb}^* & = & 0\,, \\
(\mathrm{db}) & V_{\rm ud}V_{\rm ub}^* & + & V_{\rm cd}V_{\rm cb}^* & + & V_{\rm td}V_{\rm tb}^* & = & 0\,, \\
(\mathrm{cu}) & V_{\rm ud}^*V_{\rm cd} & + & V_{\rm us}^*V_{\rm cs} & + & V_{\rm ub}^*V_{\rm cb} & = & 0\,, \\
(\mathrm{tc}) & V_{\rm cd}^*V_{\rm td} & + & V_{\rm cs}^*V_{\rm ts} & + & V_{\rm cb}^*V_{\rm tb} & = & 0\,, \\
(\mathrm{tu}) & V_{\rm ud}^*V_{\rm td} & + & V_{\rm us}^*V_{\rm ts} & + & V_{\rm ub}^*V_{\rm tb} & = & 0\,,
\end{array} \tag{1.7}$$

where the first three relations describe the orthogonality of different columns of the CKM matrix, whereas the last relations are associated with the orthogonality of different rows. As seen on Figure 1.2, only two triangles have sides of comparable size, of $\mathcal{O}(\lambda^3)$, namely the (db) and the (tu) triangles. All the other triangles are squashed. The (db) is often referred to as *The* unitarity triangle, as it represents the central target of the experimental tests of CP violation in the SM, at the B factories. However, we also need to test other triangles. For instance, LHCb will be able to probe the other triangles, in particular the (sb) triangle.

The CKM matrix may also be parameterized using four independent phases. A useful parameterization is given in terms of the following (rephasing invariant) angles [31, 19]:

$$\begin{aligned}
\gamma \equiv \phi_3 &\equiv \arg\left[-\frac{V_{\rm ud}V_{\rm ub}^*}{V_{\rm cd}V_{\rm cb}^*}\right], \\
\beta \equiv \phi_1 &\equiv \arg\left[-\frac{V_{\rm cd}V_{\rm cb}^*}{V_{\rm td}V_{\rm tb}^*}\right], \\
\beta_{\rm s} \equiv \chi &\equiv \arg\left[-\frac{V_{\rm cb}V_{\rm cs}^*}{V_{\rm tb}V_{\rm ts}^*}\right], \\
\beta_{\rm K} \equiv \chi' &\equiv \arg\left[-\frac{V_{\rm us}V_{\rm ud}^*}{V_{\rm cs}V_{\rm cd}^*}\right].
\end{aligned} \tag{1.8}$$

The angles β and γ belong to the (db) triangle, and are the only two large independent phases in the CKM matrix. We may introduce yet another angle:

$$\alpha \equiv \phi_2 \equiv \arg\left[-\frac{V_{\rm td}V_{\rm tb}^*}{V_{\rm ud}V_{\rm ub}^*}\right].$$

Nevertheless, we have by construction $\alpha + \beta + \gamma = \pi$, such that α is linearly dependent on β and γ. The other two angles in (1.8), $\beta_{\rm s}$ and $\beta_{\rm K}$, respectively appear in the (sb) and

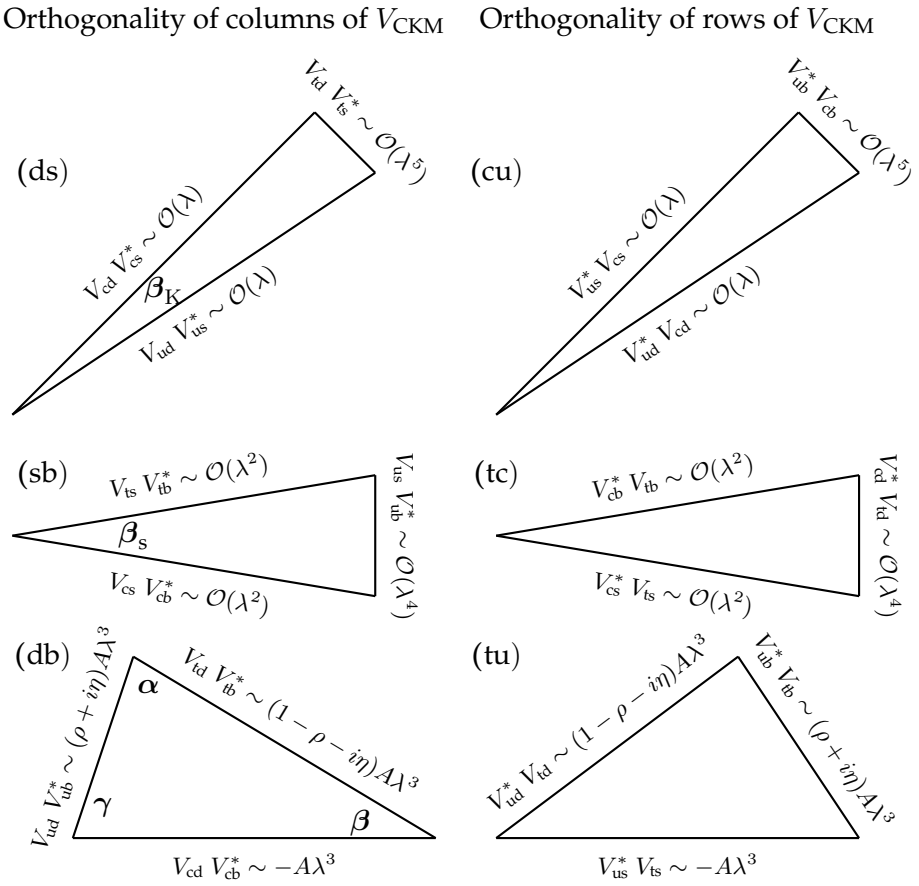


Figure 1.2: The six unitarity triangles in the complex plane, with the leading non-vanishing contributions in Wolfenstein's parameterization.

(ds) triangles. As we can see from Figure 1.2, the $\beta_{\rm s}$ angle is doubly Cabibbo-suppressed, whereas $\beta_{\rm K}$ is of $\mathcal{O}(\lambda^4)$.

Looking at (1.8), it is obvious that the angles from the (db) triangle will be mainly determined from the $\rm B_d$ system. The (sb) triangle will be useful for the $\rm B_s$ system, since $V_{\rm tb}V_{\rm ts}^*$ will control the $\rm B_s$–$\overline{\rm B}_{\rm s}$ oscillations. Finally, the angle $\beta_{\rm K}$ will describe the K system. Moreover, the interesting feature of (1.8) is that these weak phases are all related to the single CP-violating parameter η. Indeed, using Wolfenstein's parameterization, we get:

$$\begin{aligned}
\left|\frac{V_{\rm td}V_{\rm tb}^*}{V_{\rm cd}V_{\rm cb}^*}\right| e^{-i\beta} \equiv {\rm R}_{\rm t}^{\rm (db)} e^{-i\beta} &\approx 1-\rho-i\eta\,,\\
\left|\frac{V_{\rm ud}V_{\rm ub}^*}{V_{\rm cd}V_{\rm cb}^*}\right| e^{-i\gamma} \equiv {\rm R}_{\rm b}^{\rm (db)} e^{-i\gamma} &\approx \rho-i\eta\,,\\
\beta_{\rm s} = \arg\left[1-\lambda^2\left(\frac{1}{2}-\rho-i\eta\right)+\mathcal{O}(\lambda^4)\right] &\approx \lambda^2\eta\,,\\
\beta_{\rm K} = \arg\left[1-A^2\lambda^4\left(\frac{1}{2}-\rho-i\eta\right)+\mathcal{O}(\lambda^6)\right] &\approx A^2\lambda^4\eta.
\end{aligned} \tag{1.9}$$

Furthermore, we have $\beta \approx -\arg(V_{\rm td})$, $\gamma \approx -\arg(V_{\rm ub})$, and $\beta_{\rm s} \approx \arg(V_{\rm ts}) - \pi$.

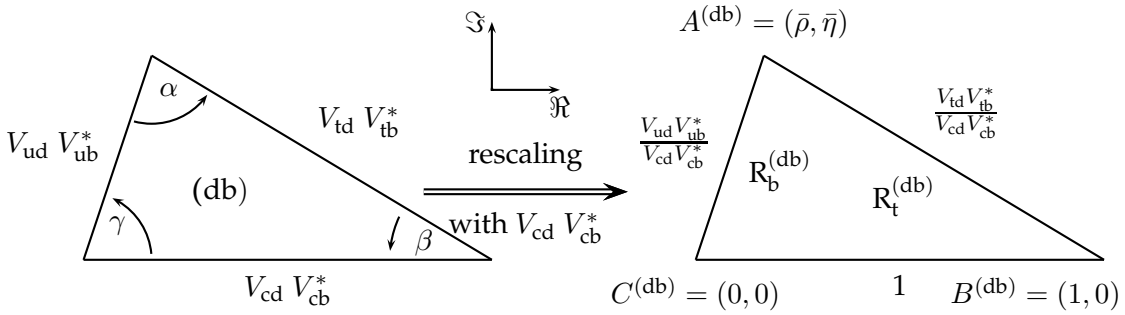


Figure 1.3: Rescaled (db) unitary triangle and the $(\bar{\rho}, \bar{\eta})$ plane.

The unitary triangle (db) plays a central role in the test of the CKM picture, because given the current experimental accuracy, it is the easiest to constrain. Looking at the (db) orthogonality relation in (1.7), and using Wolfenstein's parameterization to leading and next-to-leading orders we respectively get:

$$\begin{aligned} [(\rho + i\eta) + (-1) + (1 - \rho - i\eta)]\, A\lambda^3 + \mathcal{O}(\lambda^4) &= 0\,, \\ [(\bar{\rho} + i\bar{\eta}) + (-1) + (1 - \bar{\rho} - i\bar{\eta})]\, A\lambda^3 + \mathcal{O}(\lambda^7) &= 0\,, \end{aligned} \tag{1.10}$$

where $\bar{\rho}$ and $\bar{\eta}$ are defined in (1.5). We can thus represent the (db) triangle in the (ρ, η) plane or in the $(\bar{\rho}, \bar{\eta})$ plane, if we divide by the overall normalization factor $V_{cd}V_{cb}^* = -A\lambda^3 + \mathcal{O}(\lambda^7)$, which is real to an excellent approximation. The rescaled (db) triangle in the $(\bar{\rho}, \bar{\eta})$ plane is shown in Figure 1.3. It is obtained by aligning $V_{cd}V_{cb}^*$ with the real axis and dividing by $|V_{cd}V_{cb}^*| = A\lambda^3$, such that the length on the real axis is unity, and the unitarity relation now reads:

$$\mathrm{R}_b^{(db)} e^{i\gamma} + \mathrm{R}_t^{(db)} e^{-i\beta} = 1\,, \quad \text{rescaled with } A\lambda^3\,,$$

where the lengths of the (db) triangle are given by:

$$\mathrm{R}_b^{(db)} \equiv \left| \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right| = \sqrt{\overline{\rho}^2 + \overline{\eta}^2} = \left(1 - \frac{\lambda^2}{2}\right) \frac{1}{\lambda} \left| \frac{V_{ub}}{V_{cb}} \right|\,, \tag{1.11}$$

$$\mathrm{R}_t^{(db)} \equiv \left| \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} \right| = \sqrt{(1 - \overline{\rho})^2 + \overline{\eta}^2} = \frac{1}{\lambda} \left| \frac{V_{td}}{V_{cb}} \right|\,. \tag{1.12}$$

The apex coordinates are given by $A^{(db)} = (\bar{\rho}, \bar{\eta}) \approx (0.2, 0.4)$. The other two angular summits are given by $B^{(db)} = (1, 0)$ and $C^{(db)} = (0, 0)$. Thus, $\mathrm{R}_b^{(db)}$ is the length $C^{(db)}A^{(db)}$, whereas $\mathrm{R}_t^{(db)}$ corresponds to the size of $B^{(db)}A^{(db)}$. Anticipating on the theory of B decays, we will briefly describe how to determine the different rephasing invariant quantities in the next subsection.

1.2.3 Constraining the CKM Picture

The combination of the experimental results and the constraints on the $(\bar{\rho}, \bar{\eta})$ plane are summarized in Figure 1.4, obtained by the CKMfitter group (frequentist approach) [33].

The plot shows the current (Spring 2006) experimental status, as of the European Physical Society 2005 (HEP EPS 05) and the Flavor Physics and CP Violation (FPCP 2006) conferences, including CDF's latest measurement of ΔM_s [14].

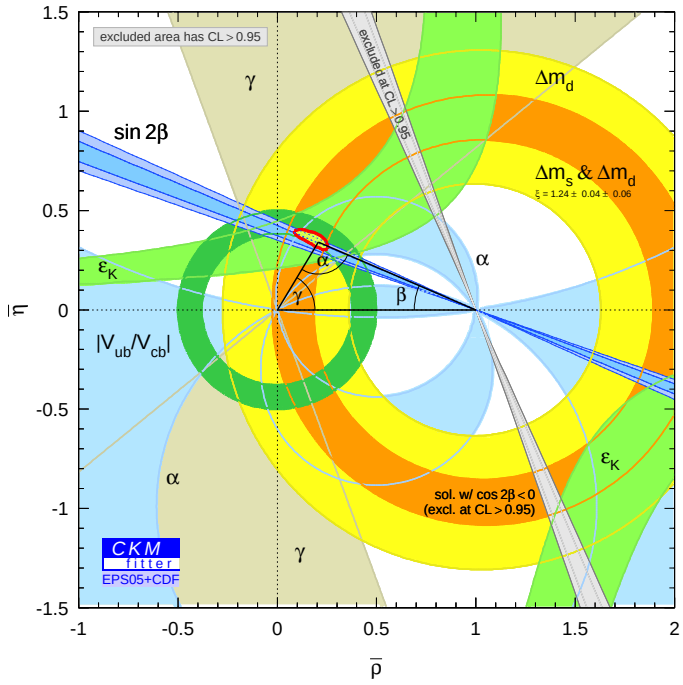


Figure 1.4: Experimental constraints on the $(\bar{\rho}, \bar{\eta})$ plane from the CMKfitter group [33], as of the EPS05 and FPCP 2006 conferences, including CDF's ΔM_s measurement.

We outline hereafter the main features allowing us to constrain the CKM picture:

$R_b^{(db)}$ This parameter involves the ratio $|V_{ub}/V_{cb}|$. It can thus be determined from b $\to$ u and b $\to$ c decays, e.g. in semi-leptonic B decays. Note that $|V_{cb}|$ determines the parameter A. The corresponding constraints are defined by a circle of radius $R_b^{(db)}$ and centered at the origin $C^{(db)} = (0, 0)$. From global fits to the CKM parameters, we typically have $R_b^{(db)} \approx 0.4$.

$R_t^{(db)}$ This parameter involves the ratio $|V_{td}/V_{cb}|$. Since the mass difference in the B_d–$\overline{B}_d$ system is $\Delta M_d \propto |V_{td}V_{tb}^*|^2$, we can constraint $R_t^{(db)}$. Furthermore, we can improve $R_t^{(db)}$ by using the measurement of $\Delta M_s \propto |V_{ts}V_{tb}^*|^2$. Indeed, we have

$$\frac{\Delta M_s}{\Delta M_d} \propto \frac{M_{B_s}}{M_{B_d}} \xi^2 \frac{|V_{ts}|^2}{|V_{td}|^2}, \quad \xi = \mathcal{O}(1),$$

where ξ is a SU(3)-breaking parameter obtained from lattice QCD calculations, accounting for corrections to the matrix elements [34]. The theoretical errors partly cancel in the ratio of the ΔM_q and we can thus further constrain $|V_{td}|$ with the measure of ΔM_s. These bounds correspond to the circles centered at $B^{(db)} = (1, 0)$ with radius $R_t^{(db)}$. Note we can also extract $|V_{ts}/V_{td}|^2$ from b $\to \ell\ell\{s, d\}$ transitions. From global fits to the CKM parameters, we typically have $R_t^{(db)} \approx 0.9$.

ε_K This parameter describes CP violation in the neutral kaon system arising from a box diagram involving all up quarks as intermediate lines. These bounds correspond to a hyperbola in the $(\bar{\rho}, \bar{\eta})$ plane. From global fits to the CKM parameters, we typically have $\varepsilon_K \approx 2 \cdot 10^{-3}$.

α The measurement of this phase can be done through $\bar{b} \to \bar{u}u\bar{d}$ quark-level transitions, though α is not directly accessed. For instance using $B_d \to \pi^+\pi^-$ decays, neglecting the penguin pollution for the illustration, we can determine $\sin(\phi_d + 2\gamma)$ from the corresponding CP asymmetry, where $\phi_d \equiv 2\arg[V_{td}^* V_{tb}]$ is the B_d–$\overline{B}_d$ mixing phase. In the SM we have $\phi_d \approx 2\beta$, such that the CP asymmetry measures $\sin(2\alpha)$. In practice, we will have to deal with penguins, thus we will effectively measure γ. From global fits to the CKM parameters, we typically have $\sin(2\alpha) \approx -0.2$.

β This angle is best determined from $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions in B_d decays, where the phase of the dominant tree-level amplitude is approximatively real. Consequently the CP asymmetry will determine $\sin(\phi_d)$, which reduces to $\sin(2\beta)$. The gold-plated channel for this measurement is $B_d \to J/\psi\ K_S^0$, which is a decay to a pure CP-odd eigenstate. From global fits to the CKM parameters, we typically have $\sin(2\beta) \approx 0.7$.

β_s The determination of β_s is the strange counterpart of the β measurement, using $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions in B_s decays. We can use the $B_s \to \eta_c\phi$, $B_s \to D_sD_s$, $B_s \to J/\psi\eta^{(\prime)}$, and $B_s \to J/\psi\phi$ decays, as we will extensively discuss in this thesis. The CP asymmetry for these channels will probe the B_s–$\overline{B}_s$ mixing phase $\phi_s \equiv 2\arg[V_{ts}^* V_{tb}]$, which corresponds to $\phi_s \approx -2\beta_s$. From global fits to the CKM parameters, we typically have $\sin(2\beta_s) \approx 0.04$.

γ This angle can independently be determined in the $B_s \to D_s^{(*)\pm}\ K^\mp$ and the counterparts $B_d \to D^{(*)\pm}\ \pi^\mp$, $B_d \to D\ K$ decays. These decays actually enable the determination of $\gamma + \phi_s - \beta_K$ and $\gamma + \phi_d$ respectively for the B_s and B_d decays. Making use of (1.9), we can thus determine $\gamma - 2\beta_s$ and $\gamma + 2\beta$. From global fits to the CKM parameters, we typically have $\sin(\gamma + 2\beta) \approx 0.9$.

Note that in the present experimental situation, the constraints are dominated by the well measured $\sin(2\beta)$. As a summary, $R_t^{(db)}$ and $\beta_{(s)}$ are determined from processes involving B_q–$\overline{B}_q$ mixing, whereas $R_b^{(db)}$ and γ are obtained from processes which do not involve mixing, and thus purely from (tree) decays.

1.3 The Neutral B–$\overline{\text{B}}$ System

We derive in this section the quantum formalism to describe the particle–antiparticle system, directly applied to the B system, and we introduce the observables and quantities related to CP violation. We generically denote by B_q and $\overline{B}_q$ the eigenstates of the strong and electromagnetic interactions, with well-defined flavors, where $q \equiv s$ for the B_s system and $q \equiv d$ for the B_d system.

1.3.1 Particle–Antiparticle Formalism

Let $|B_q\rangle$ and $|\overline{B}_q\rangle$ be the flavor eigenstates, corresponding to the B_q and $\overline{B}_q$ particles with the valence quark contents $B_d = \bar{b}d$, $\overline{B}_d = b\bar{d}$, $B_s = \bar{b}s$, $\overline{B}_s = b\bar{s}$. CPT invariance of the

strong and electromagnetic interactions implies that the masses of the $\mathrm{B_q}$ and of the $\overline{\mathrm{B}}_\mathrm{q}$ are identical, which we denote by $M_{\mathrm{B_q}}$.

Assuming CPT invariance, we have the freedom to rephase the meson states with arbitrary phases φ_{CPT} and $\bar{\varphi}_{\mathrm{CPT}}$. We next introduce the CP transformation ($CP^2 = \mathbb{1}$) of the $\mathrm{B_q}$ and $\overline{\mathrm{B}}_\mathrm{q}$, which involves an *arbitrary* phase φ_{CP} resulting from the invariance under rephasing of the quark fields:

$$CP \left|\mathrm{B_q}\right\rangle = e^{i\varphi_{\mathrm{CP}}} \left|\overline{\mathrm{B}}_\mathrm{q}\right\rangle \,, \quad CP \left|\overline{\mathrm{B}}_\mathrm{q}\right\rangle = e^{-i\varphi_{\mathrm{CP}}} \left|\mathrm{B_q}\right\rangle \,. \tag{1.13}$$

Note that φ_{CP} transforms as $\varphi_{\mathrm{CP}} \to \varphi_{\mathrm{CP}} + (\varphi_{\mathrm{CPT}} - \bar{\varphi}_{\mathrm{CPT}})$ under a rephasing of the meson states.

At a time $t = 0$ in the rest frame of the $\mathrm{B_q}$–$\overline{\mathrm{B}}_\mathrm{q}$ system, the particle and the antiparticle are stable under the strong and electromagnetic interactions, and flavor conservation implies that $\left\langle \overline{\mathrm{B}}_\mathrm{q} \left|\mathrm{B_q}\right\rangle = 0$. At a time $t > 0$ we turn on the weak interaction, and the $\mathrm{B_q}$ and $\overline{\mathrm{B}}_\mathrm{q}$ start to mix ($|\Delta F| = 2$ transitions) or decay into final states f ($|\Delta F| = 1$). We can write the general state $|\psi(t)\rangle$ as a superposition of the initial states and all final states:

$$|\psi(t)\rangle = a(t) \left|\mathrm{B_q}\right\rangle + b(t) \left|\overline{\mathrm{B}}_\mathrm{q}\right\rangle + \sum_i c_i(t) \left|f_i\right\rangle \,, \tag{1.14}$$

satisfying the Schrödinger equation:

$$i\partial_t \left|\psi(t)\right\rangle = H_{\mathrm{tot}} \left|\psi(t)\right\rangle \,, \quad H_{\mathrm{tot}} \equiv H_{\mathrm{s}} + H_{\mathrm{em}} + H_{\mathrm{w}} \,, \tag{1.15}$$

where the total Hamiltonian is the sum of the strong, electromagnetic and weak contributions. We assume that the total Hamiltonian is Hermitian ($H_{\mathrm{tot}} = H_{\mathrm{tot}}^\dagger$), such that we have conservation of the transition probabilities for the coefficients:

$$|a(t)|^2 + |b(t)|^2 + \sum_i |c_i(t)|^2 = 1 \,.$$

We will now solve (1.15), with the following initial conditions at $t = 0$:

$$|a(0)|^2 + |b(0)|^2 = 1 \,, \quad c_i(0) = 0 \,, \forall i \,. \tag{1.16}$$

We go to the Dirac picture and use second-order perturbation theory. We introduce:

$$|\psi(t)\rangle = \exp[-i(H_{\mathrm{s}} + H_{\mathrm{em}})t] \left|\psi(t)\right\rangle_{\mathrm{D}} \,, \tag{1.17}$$

such that (1.15) becomes:

$$i\partial_t \left|\psi(t)\right\rangle_{\mathrm{D}} = V(t) \left|\psi(t)\right\rangle_{\mathrm{D}} \,, \quad V(t) = \exp[i(H_{\mathrm{s}} + H_{\mathrm{em}})t] H_{\mathrm{w}} \exp[-i(H_{\mathrm{s}} + H_{\mathrm{em}})t] \,, \tag{1.18}$$

where $V(t)$ is the new *time-dependent* Hamiltonian. The coefficients in (1.14) become *operators*, redefined through:

$$\hat{q}(t) = q(t) \exp[i(H_{\mathrm{s}} + H_{\mathrm{em}})t] \,, \quad \hat{q} = \{\hat{a}, \hat{b}, \hat{c}_i\} \,, q = \{a, b, c_i\} \,. \tag{1.19}$$

Inserting (1.19) into (1.18), operating on the left with the bras, and using the fact that $\left|\mathrm{B_q}\right\rangle$, $\left|\overline{\mathrm{B}}_\mathrm{q}\right\rangle$ and $|f_i\rangle$ are eigenstates of $H_{\mathrm{s}} + H_{\mathrm{em}}$, we get the following set of operator differential

equations:

$$\begin{aligned}
i\partial_t \hat{a}(t) &= \left\langle \mathrm{B_q} \right| H_\mathrm{w} \left| \mathrm{B_q} \right\rangle \hat{a}(t) + \left\langle \mathrm{B_q} \right| H_\mathrm{w} \left| \overline{\mathrm{B}}_\mathrm{q} \right\rangle \hat{b}(t) + \sum_i e^{i(E_i - M_{\mathrm{B_q}})t} \left\langle \mathrm{B_q} \right| H_\mathrm{w} \left| f_i \right\rangle \hat{c}_i(t)\,, \\
i\partial_t \hat{b}(t) &= \left\langle \overline{\mathrm{B}}_\mathrm{q} \right| H_\mathrm{w} \left| \mathrm{B_q} \right\rangle \hat{a}(t) + \left\langle \overline{\mathrm{B}}_\mathrm{q} \right| H_\mathrm{w} \left| \overline{\mathrm{B}}_\mathrm{q} \right\rangle \hat{b}(t) + \sum_i e^{i(E_i - M_{\mathrm{B_q}})t} \left\langle \overline{\mathrm{B}}_\mathrm{q} \right| H_\mathrm{w} \left| f_i \right\rangle \hat{c}_i(t)\,, \\
i\partial_t \hat{c}_i(t) &= e^{i(M_{\mathrm{B_q}} - E_i)t} \left\langle f_i \right| H_\mathrm{w} \left| \mathrm{B_q} \right\rangle \hat{a}(t) + e^{i(M_{\mathrm{B_q}} - E_i)t} \left\langle f_i \right| H_\mathrm{w} \left| \overline{\mathrm{B}}_\mathrm{q} \right\rangle \hat{b}(t) \\
&\quad + \sum_{i'} e^{i(E_{i'} - E_i)t} \left\langle f_i \right| H_\mathrm{w} \left| f_{i'} \right\rangle \hat{c}_{i'}(t)\,, \quad \forall i\,. \qquad (1.20)
\end{aligned}$$

We can make use of the Wigner–Weisskopf approximation [35], namely given the times t in which we are interested are much larger than the typical strong interaction scale, we can neglect the weak interactions between the final states, i.e. we simply set $\left\langle f_i \right| H_\mathrm{w} \left| f_{i'} \right\rangle = 0$. With this approximation we can now integrate by parts the last relation of (1.20), using the initial conditions (1.16), and neglecting second order terms in H_w:[4]

$$\hat{c}_i(t) = \lim_{\varepsilon \to 0+} \frac{e^{i(M_{\mathrm{B_q}} - E_i)t}}{(E_i - M_{\mathrm{B_q}}) + i\varepsilon} \left[\left\langle f_i \right| H_\mathrm{w} \left| \mathrm{B_q} \right\rangle \hat{a}(t) + \left\langle f_i \right| H_\mathrm{w} \left| \overline{\mathrm{B}}_\mathrm{q} \right\rangle \hat{b}(t) \right] \qquad (1.21)$$

where we introduced a small imaginary part which tells us the way to go around the pole at $M_{\mathrm{B_q}} = E_i$. The above expression enables to decouple the equations for $\hat{a}(t)$ and $\hat{b}(t)$, if we insert (1.21) into (1.20). Before doing that, we need the following functional identity [36]:

$$\begin{aligned}
\lim_{\varepsilon \to 0+} \frac{1}{x \pm i\varepsilon} &= \mathcal{P}\left(\frac{1}{x}\right) \mp i\pi\delta(x)\,, \qquad (1.22) \\
\mathcal{P}\left(\int_{-A}^{B} \frac{\mathrm{d}x}{x} f(x) \right) &= \lim_{\eta \to 0+} \left[\int_{-A}^{-\eta} + \int_{+\eta}^{B} \right] \frac{\mathrm{d}x}{x} f(x)\,,
\end{aligned}$$

with $A, B > 0$, $f(x)$ regular at $x = 0$, where $\delta(x)$ is Dirac's (even) distribution, and $\mathcal{P}$ is Cauchy's principal value. Using the above relations we can rewrite (1.20) with (1.19), (1.21), (1.22), and we get the effective Schrödinger equation, after elimination of the $\hat{q}(t)$:

$$i\partial_t \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = H_\mathrm{eff} \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \left(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma} \right) \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}, \qquad (1.23)$$

where H_eff is the effective Hamiltonian, and $\mathbf{M}$ and $\mathbf{\Gamma}$ are respectively the mass and decay 2×2 matrices. We therefore see that we can describe the $\mathrm{B_q}$–$\overline{\mathrm{B}}_\mathrm{q}$ system in its rest frame by the two-component state solution of (1.23):

$$\left| \psi(t) \right\rangle = a(t) \left| \mathrm{B_q} \right\rangle + b(t) \left| \overline{\mathrm{B}}_\mathrm{q} \right\rangle . \qquad (1.24)$$

The matrix elements in (1.23) are given by:

$$\begin{aligned}
M_{kl}^{(\mathrm{q})} &= M_{\mathrm{B_q}} \delta_{kl} + \left\langle k \right| H_\mathrm{w} \left| l \right\rangle + \sum_i \mathcal{P}\left(\frac{\left\langle k \right| H_\mathrm{w} \left| i \right\rangle \left\langle i \right| H_\mathrm{w} \left| l \right\rangle}{M_{\mathrm{B_q}} - E_i} \right), \qquad (1.25) \\
\Gamma_{kl}^{(\mathrm{q})} &= 2\pi \sum_i \delta(M_{\mathrm{B_q}} - E_i) \left\langle k \right| H_\mathrm{w} \left| i \right\rangle \left\langle i \right| H_\mathrm{w} \left| l \right\rangle , \qquad (1.26)
\end{aligned}$$

[4] We neglect terms proportional to $\left\langle f_i \right| H_\mathrm{w} \left| \mathrm{B_q} \right\rangle \partial_{t'} \hat{a}(t')$ and $\left\langle f_i \right| H_\mathrm{w} \left| \overline{\mathrm{B}}_\mathrm{q} \right\rangle \partial_{t'} \hat{b}(t')$, given that the weak interaction is much weaker than the electromagnetic and strong interactions.

where $k, l = 1, 2$ denote $\mathrm{B_q}$ and $\overline{\mathrm{B}}_\mathrm{q}$ respectively, and δ_{kl} is the Kronecker symbol. The delta distribution $\delta(M_{\mathrm{B_q}} - E_i)$ in $\mathbf{\Gamma}$ ensures the energy conservation, therefore the states to which $\mathrm{B_q}$ and $\overline{\mathrm{B}}_\mathrm{q}$ may decay are *physical* states (i.e. real, on their massshell). On the other hand, the states appearing in (1.25) are *virtual* (i.e. off their massshell), such that the sum must be taken over *all* possible intermediate states.

The effective potential H_{eff} is no longer Hermitian, as a consequence of the Wigner–Weisskopf approximation. However, the $\mathbf{M}$ and $\mathbf{\Gamma}$ matrices are Hermitian, i.e. $M_{kl}^{(\mathrm{q})} = M_{lk}^{(\mathrm{q})*}$ and $\Gamma_{kl}^{(\mathrm{q})} = \Gamma_{lk}^{(\mathrm{q})*}$. Note that CPT (and CP) is violated if $M_{11}^{(\mathrm{q})} \neq M_{22}^{(\mathrm{q})}$ or $\Gamma_{11}^{(\mathrm{q})} \neq \Gamma_{22}^{(\mathrm{q})}$. CP (and T) is violated if $H_{12}^{(\mathrm{q})} \neq H_{21}^{(\mathrm{q})}$. In the following we assume CPT conservation. We have:

$$H_{\text{eff}} = \begin{pmatrix} H_0^{(\mathrm{q})} & H_{12}^{(\mathrm{q})} \\ H_{21}^{(\mathrm{q})} & H_0^{(\mathrm{q})} \end{pmatrix} \equiv \left(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma} \right) = \begin{pmatrix} M^{(\mathrm{q})} & M_{12}^{(\mathrm{q})} \\ M_{12}^{(\mathrm{q})*} & M^{(\mathrm{q})} \end{pmatrix} - \frac{i}{2} \begin{pmatrix} \Gamma^{(\mathrm{q})} & \Gamma_{12}^{(\mathrm{q})} \\ \Gamma_{12}^{(\mathrm{q})*} & \Gamma^{(\mathrm{q})} \end{pmatrix} . \tag{1.27}$$

The eigenvalues of H_{eff} are easily obtained by solving (1.23) using the Ansatz $a(t) = C_+ e^{-i\lambda_+ t} + C_- e^{-i\lambda_- t}$. The initial conditions (1.16) define the constants $C_+ = C_- = 1/2$, assuming a $\mathrm{B_q}$ is created at $t = 0$. The eigenvalue equation reads :

$$\lambda_\pm^2 - 2\lambda_\pm H_0^{(\mathrm{q})} - \left(H_{21}^{(\mathrm{q})} H_{12}^{(\mathrm{q})} - H_0^{(\mathrm{q})2} \right) = 0 \,, \tag{1.28}$$

with solutions:

$$\lambda_\pm = H_0^{(\mathrm{q})} \pm \sqrt{H_{21}^{(\mathrm{q})} H_{12}^{(\mathrm{q})}} \equiv H_0^{(\mathrm{q})} \pm H_{12}^{(\mathrm{q})} \frac{q}{p} \,, \tag{1.29}$$

where we introduced the complex number q/p given by:

$$\frac{q}{p} \equiv e^{in\pi} \sqrt{\frac{H_{21}^{(\mathrm{q})}}{H_{12}^{(\mathrm{q})}}} = e^{in\pi} \sqrt{\frac{2M_{12}^{(\mathrm{q})*} - i\Gamma_{12}^{(\mathrm{q})*}}{2M_{12}^{(\mathrm{q})} - i\Gamma_{12}^{(\mathrm{q})}}} \,, \tag{1.30}$$

where the quantity $n \in \{0, 1\}$ is introduced to parameterize the sign of the square root. It is important to stress here that q/p is *not* a physical observable, since there is a two-fold ambiguity when taking the square root of a complex number. In the same way, $M_{12}^{(\mathrm{q})}$ and $\Gamma_{12}^{(\mathrm{q})}$ (and their phases) depend on the choice of the CP transformation phase. The ratio q/p is fixed by the diagonalization of (1.27):

$$H_{\text{eff}} \left| \mathrm{B}_\pm \right\rangle = \lambda_\pm \left| \mathrm{B}_\pm \right\rangle \,, \tag{1.31}$$

where the physical eigenstates B_+ with eigenvalue λ_+ and B_- with eigenvalue λ_- are given by:

$$\left| \mathrm{B}_\pm \right\rangle = \frac{1}{\sqrt{1 + \left| \frac{q}{p} \right|^2}} \left(\left| \mathrm{B_q} \right\rangle \pm \frac{q}{p} \left| \overline{\mathrm{B}}_\mathrm{q} \right\rangle \right) \,, \quad |q|^2 + |p|^2 = 1 \,. \tag{1.32}$$

The coefficients entering (1.14) can be expressed in term of the eigenvalues as $a(t) = g_+(t)$ and $b(t) = (q/p) g_-(t)$, where the $g_\pm$ functions are defined through:

$$g_\pm(t) = \frac{1}{2} \left(e^{-i\lambda_+ t} \pm e^{-i\lambda_- t} \right) \,, \tag{1.33}$$

We can now write the general solution of (1.15) for an initially ($t = 0$) pure $\mathrm{B_q}$, respectively $\overline{\mathrm{B}}_\mathrm{q}$ meson, as:

$$\begin{aligned}
\left|\mathrm{B_q}(t)\right\rangle &= g_+(t)\left|\mathrm{B_q}\right\rangle + \frac{q}{p} g_-(t)\left|\overline{\mathrm{B}}_\mathrm{q}\right\rangle \\
&= \frac{1}{2}\sqrt{1+\left|\frac{q}{p}\right|^2}\left(e^{-i\lambda_+ t}\left|\mathrm{B}_+\right\rangle + e^{-i\lambda_- t}\left|\mathrm{B}_-\right\rangle\right), \qquad (1.34)\\
\left|\overline{\mathrm{B}}_\mathrm{q}(t)\right\rangle &= g_+(t)\left|\overline{\mathrm{B}}_\mathrm{q}\right\rangle + \frac{p}{q} g_-(t)\left|\mathrm{B_q}\right\rangle \\
&= \frac{1}{2\left(\frac{q}{p}\right)}\sqrt{1+\left|\frac{q}{p}\right|^2}\left(e^{-i\lambda_+ t}\left|\mathrm{B}_+\right\rangle - e^{-i\lambda_- t}\left|\mathrm{B}_-\right\rangle\right). \qquad (1.35)
\end{aligned}$$

We thus see that the flavor states can either remain unchanged or oscillate into each other. This feature is named particle-antiparticle mixing. Moreover, the general states expressed in terms of the physical eigenstates $\mathrm{B}_\pm$ show us that the $\mathrm{B}_\pm$ have exponential evolution laws with well defined masses and decay widths:

$$\left|\mathrm{B}_\pm(t)\right\rangle = e^{-i\lambda_\pm t}\left|\mathrm{B}_\pm\right\rangle ,$$

where t is the time measured in the rest frame of the decaying particle. These physical eigenstates are thus called mass eigenstates. We can introduce the corresponding masses $M_\pm$ and decay widths $\Gamma_\pm$, which are defined by:[5]

$$\lambda_\pm \equiv M_\pm - \frac{i}{2}\Gamma_\pm \equiv H_0^{(\mathrm{q})} \pm H_{12}^{(\mathrm{q})}\frac{q}{p}, \qquad (1.36)$$

where we can identify:

$$\begin{aligned}
M_\pm &\equiv \Re(\lambda_\pm) = M^{(\mathrm{q})} \pm \Re(H_{12}^{(\mathrm{q})}\frac{q}{p}), \qquad (1.37)\\
\Gamma_\pm &\equiv -2\Im(\lambda_\pm) = \Gamma^{(\mathrm{q})} \mp 2\Im(H_{12}^{(\mathrm{q})}\frac{q}{p}). \qquad (1.38)
\end{aligned}$$

1.3.2 $\mathrm{B_q}$–$\overline{\mathrm{B}}_\mathrm{q}$ Mixing

The $\mathrm{B_q}$–$\overline{\mathrm{B}}_\mathrm{q}$ mixing refers to the transitions between the two flavor eigenstates, which are caused by the off-diagonal terms in the effective Hamiltonian defined in (1.27). As we mentioned in the previous subsection, $\Gamma_{12}^{(\mathrm{q})}$ originates from the real final states to which $\mathrm{B_q}$ and $\overline{\mathrm{B}}_\mathrm{q}$ may decay. Since $\Gamma_{12}^{(\mathrm{q})}$ is dominated by tree-level decays, it is almost not sensitive to NP. The intermediate states appearing in the determination of $M_{12}^{(\mathrm{q})}$ are on the other hand virtual, thus NP could easily contribute to it. In the SM, the leading contributions to $M_{12}^{(\mathrm{q})}$ are induced by fourth-order flavor changing weak transitions. These $|\Delta F| = 2$ transitions correspond to the box diagrams depicted in Figure 1.5. Their dispersive parts are dominated by internal top-quark exchanges, such that the short-range contribution yields [20]:

$$M_{12}^{\mathrm{SM}\,(\mathrm{q})} \propto (V_{\mathrm{tq}}^* V_{\mathrm{tb}})^2 e^{i(\pi-\varphi_{\mathrm{CP}})} \qquad (1.39)$$

[5] These $M_\pm$ masses and $\Gamma_\pm$ widths are *not* the eigenvalues of $\mathbf{M}$ and $\mathbf{\Gamma}$, respectively.

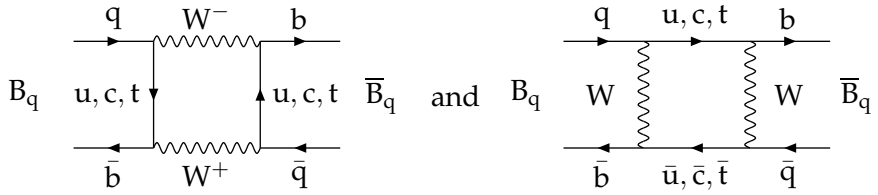


Figure 1.5: The box diagrams contributing to the B_q–$\overline{B}_q$ mixing in the SM ($q \in \{d, s\}$).

where we hid the QCD corrections, the dependence on the masses (top, B_q, W) in the proportionality factor. Note that the phase of $M_{12}^{(q)}$ depends on the CP phase convention through φ_{CP}, which is defined in (1.13). Moreover, it also depends on the CKM phase convention $\xi_q = (\xi_{q^j} - \xi_{q^i})$ when rephasing the quark spinor fields:

$$(CP)q(CP)^\dagger = e^{i\xi_q}\gamma^0 C\bar{q}^T\,, \quad (CP)\bar{q}(CP)^\dagger = -e^{-i\xi_q}\bar{q}^T C^{-1}\gamma^0\,,$$
$$(CP)\left((\bar{q})^i_L\gamma^\mu(q)^j_L\right)(CP)^\dagger = -e^{i(\xi_{q^j}-\xi_{q^i})}\left((\bar{q})^j_L\gamma^\mu(q)^i_L\right)\,.$$

The phase of $M_{12}^{(q)}$, $\theta_M^{(q)}$, is defined by:

$$\begin{aligned} M_{12}^{(q)} &\equiv e^{i\theta_M^{(q)}}\left|M_{12}^{(q)}\right|\,, \\ \theta_M^{(q)} &\equiv \arg\left[M_{12}^{(q)}\right] = \phi_q + \pi - \varphi_{CP}\,, \end{aligned} \tag{1.40}$$

where the weak phase ϕ_q is the observable phase of the effective Hamiltonian describing the mixing. In the SM:

$$\phi_q^{SM} \equiv 2\arg[V_{tq}^* V_{tb}] = \begin{cases} \phi_d^{SM} \approx 2\beta & \text{for q = d,} \\ \phi_s^{SM} \approx -2\beta_s & \text{for q = s.} \end{cases} \tag{1.41}$$

In the B_s case, the contribution (1.39) is suppressed by the weak coupling constant, and by two powers of $|V_{ts}| \approx 0.04$, such that NP could easily contribute to $M_{12}^{(q)}$. It is worth noting here that since in the SM the B_s–$\overline{B}_s$ mixing phase ϕ_s^{SM} is $\mathcal{O}(-0.04)$ rad, any deviation from that expectation will be a clear sign for NP. In this case, we would actually effectively observe the contributions from the SM and NP, $\phi_s = \phi_s^{SM} + \phi_s^{NP}$. Finally, as we will see below, we will be able to probe the phase difference between $M_{12}^{(q)}$ and $\Gamma_{12}^{(q)}$ which constitutes a physical observable. Without loss of generality, we can adopt the phase convention in which $\varphi_{CP} = \pi$ and $\xi_q = 0$, as used in [22], such that we can identify ϕ_s with $\theta_M^{(q)}$.

The calculation of the absorptive parts of the box diagrams in Figure 1.5 gives [20]:

$$\frac{\Gamma_{12}^{(q)}}{M_{12}^{(q)}} \approx \mathcal{O}\left(m_b^2/m_t^2\right) \ll 1\,. \tag{1.42}$$

We therefore see that the deviation of $|q/p|$ from unity is really small, and will be neglected in our calculations. As we will see later, we ignore in this way the so-called CP violation in the mixing.

1.3.3 Mixing Parameters

The sign choice in (1.30) is fixed by introducing the *positive* mass difference $M_- - M_+ > 0$. We can now change our notations ($\mathrm{B}_- \sim \mathrm{B_H}$ and $\mathrm{B}_+ \sim \mathrm{B_L}$) and we redefine the physical mass eigenstates as:

$$\begin{aligned} \text{Lighter eigenstate:} \quad |\mathrm{B_L}\rangle &= p\,\big|\mathrm{B_q}\big\rangle + q\,\big|\overline{\mathrm{B}}_\mathrm{q}\big\rangle\,, \\ \text{Heavier eigenstate:} \quad |\mathrm{B_H}\rangle &= p\,\big|\mathrm{B_q}\big\rangle - q\,\big|\overline{\mathrm{B}}_\mathrm{q}\big\rangle\,. \end{aligned} \tag{1.43}$$

The relations defining the average mass $M_{\mathrm{B_q}}$ and width Γ_q of the $\mathrm{B_q}$-meson eigenstates, and the difference of their mass ΔM_q and width $\Delta\Gamma_\mathrm{q}$ are:

$$\begin{aligned} M_{\mathrm{B_q}} &= \frac{M_\mathrm{H}+M_\mathrm{L}}{2} = M_{11}\,, \qquad & \Delta M_\mathrm{q} &= M_\mathrm{H} - M_\mathrm{L} > 0\,, \\ \Gamma_\mathrm{q} &= \frac{\Gamma_\mathrm{H}+\Gamma_\mathrm{L}}{2} = \Gamma_{11}\,, \qquad & \Delta\Gamma_\mathrm{q} &= \Gamma_\mathrm{L} - \Gamma_\mathrm{H}\,. \end{aligned} \tag{1.44}$$

In this convention, the SM prediction for $\Delta\Gamma_\mathrm{q}$ is positive. The SM predictions [20] for the $\mathrm{B_d}$ and $\mathrm{B_s}$ systems are respectively $\Delta\Gamma_\mathrm{d}/\Gamma_\mathrm{d} \sim 10^{-2}$ and $\Delta\Gamma_\mathrm{s}/\Gamma_\mathrm{s} \sim 10^{-1}$, and $\Delta M_\mathrm{d} \sim 0.5$ ps^{-1} and $\Delta M_\mathrm{s} \sim 20$ ps^{-1}. The mixing parameters ΔM_q and $\Delta\Gamma_\mathrm{q}$ are the physical observables describing the $\mathrm{B_q}$–$\overline{\mathrm{B}}_\mathrm{q}$ oscillations.

If we square (1.29) and separate the real and the imaginary parts we get:

$$\Re\,\text{part} \quad \rightarrow \quad (\Delta M_\mathrm{q})^2 - \frac{1}{4}(\Delta\Gamma_\mathrm{q})^2 = 4\left|M_{12}^{(\mathrm{q})}\right|^2 - \left|\Gamma_{12}^{(\mathrm{q})}\right|^2\,, \tag{1.45}$$

$$\Im\,\text{part} \quad \rightarrow \quad \Delta M_\mathrm{q}\Delta\Gamma_\mathrm{q} = -4\Re\left(M_{12}^{(\mathrm{q})}\Gamma_{12}^{(\mathrm{q})*}\right)\,. \tag{1.46}$$

We can express the complex parameter q/p as:

$$\frac{q}{p} = -\sqrt{\frac{H_{21}^{(\mathrm{q})}}{H_{12}^{(\mathrm{q})}}} = -\frac{\Delta M_\mathrm{q} + \frac{i}{2}\Delta\Gamma_\mathrm{q}}{2M_{12}^{(\mathrm{q})} - i\Gamma_{12}^{(\mathrm{q})}} = -\frac{2M_{12}^{(\mathrm{q})*} - i\Gamma_{12}^{(\mathrm{q})*}}{\Delta M_\mathrm{q} + \frac{i}{2}\Delta\Gamma_\mathrm{q}}\,, \tag{1.47}$$

where the minus signs results from our ΔM_q convention, independently of the sign choice of $\Delta\Gamma_\mathrm{q}$. As already stressed previously, q/p is not a phase invariant observable.

The functions defined in (1.33) can be rewritten in terms of the mass and width differences:

$$g_\pm(t) = e^{-iM_{\mathrm{B_q}}t}\,e^{-\Gamma^{(\mathrm{q})}t/2}\left[\pm\cosh\frac{\Delta\Gamma_\mathrm{q}\,t}{4}\,\cos\frac{\Delta M_\mathrm{q}\,t}{2} \mp i\sinh\frac{\Delta\Gamma_\mathrm{q}\,t}{4}\,\sin\frac{\Delta M_\mathrm{q}\,t}{2}\right]\,. \tag{1.48}$$

We also introduce the following relations which will be useful in the determination of the transition rates:

$$\begin{aligned} |g_\pm(t)|^2 &= \frac{e^{-\Gamma_\mathrm{q}t}}{2}\left[\cosh\frac{\Delta\Gamma_\mathrm{q}\,t}{2} \pm \cos\Delta M_\mathrm{q}t\right]\,, \\ g_+^*(t)\,g_-(t) &= \frac{e^{-\Gamma_\mathrm{q}t}}{2}\left[-\sinh\frac{\Delta\Gamma_\mathrm{q}\,t}{2} + i\sin\Delta M_\mathrm{q}t\right]\,. \end{aligned} \tag{1.49}$$

Note that these expressions depend on the sign choice of $\Delta\Gamma_\mathrm{q}$.

We next define the phase difference between $M_{12}^{(\mathrm{q})}$ and $\Gamma_{12}^{(\mathrm{q})}$ which appears in many observables related to the $\mathrm{B_q}$–$\overline{\mathrm{B}}_\mathrm{q}$ mixing :

$$\frac{M_{12}^{(\mathrm{q})}}{\Gamma_{12}^{(\mathrm{q})}} = -\left|\frac{M_{12}^{(\mathrm{q})}}{\Gamma_{12}^{(\mathrm{q})}}\right| e^{i\phi_{M/\Gamma}} \leftrightarrow \phi_{M/\Gamma} = \arg\left[-\frac{M_{12}^{(\mathrm{q})}}{\Gamma_{12}^{(\mathrm{q})}}\right], \tag{1.50}$$

where the minus sign means that $M_{12}^{(\mathrm{q})}$ and $\Gamma_{12}^{(\mathrm{q})}$ are antiparallel. When $\Gamma_{12}^{(\mathrm{q})}/M_{12}^{(\mathrm{q})} \ll 1$, the complex quantity q/p reduces to:

$$\frac{q}{p} = -\frac{M_{12}^{(\mathrm{q})*}}{\left|M_{12}^{(\mathrm{q})}\right|} = -e^{-i\theta_M^{(\mathrm{q})}}, \quad \frac{\Gamma_{12}^{(\mathrm{q})}}{M_{12}^{(\mathrm{q})}} \lll 1,$$

and consistent with our positive mass difference convention. The phase $\phi_{M/\Gamma}$, which is independent of phase convention, is an observable CP-violating phase. As shown in (1.42), neglecting CP violation in the mixing is a good approximations for the $\mathrm{B_q}$–$\overline{\mathrm{B}}_\mathrm{q}$ system. In this case, we may solve the equations (1.45), (1.46) and (1.47) by using the relations $M_{12}^{(\mathrm{q})} \gg \Gamma_{12}^{(\mathrm{q})}$ and $\Delta M_\mathrm{q} \gg \Delta\Gamma_\mathrm{q}$, satisfied by both the $\mathrm{B_{d,s}}$ systems. We get [22]:

$$\begin{aligned}
\Delta M_\mathrm{q} &= 2\left|M_{12}^{(\mathrm{q})}\right|\left[1+\mathcal{O}\left(\left|\frac{\Gamma_{12}^{(\mathrm{q})}}{M_{12}^{(\mathrm{q})}}\right|^2\right)\right], \\
\Delta\Gamma_\mathrm{q} &= 2\left|\Gamma_{12}^{(\mathrm{q})}\right|\cos\phi_{M/\Gamma}\left[1+\mathcal{O}\left(\left|\frac{\Gamma_{12}^{(\mathrm{q})}}{M_{12}^{(\mathrm{q})}}\right|^2\right)\right], \\
\frac{q}{p} &= -e^{-i\theta_M^{(\mathrm{q})}}\left[1-\frac{a}{2}\right]+\mathcal{O}\left(\left|\frac{\Gamma_{12}^{(\mathrm{q})}}{M_{12}^{(\mathrm{q})}}\right|^2\right),
\end{aligned} \tag{1.51}$$

where the small quantity a_q is defined by:

$$a_\mathrm{q} = \Im\left(\frac{\Gamma_{12}^{(\mathrm{q})}}{M_{12}^{(\mathrm{q})}}\right) = \left|\frac{\Gamma_{12}^{(\mathrm{q})}}{M_{12}^{(\mathrm{q})}}\right|\sin\phi_{M/\Gamma} = \frac{\left|\Delta\Gamma_\mathrm{q}\right|}{\Delta M_\mathrm{q}}\frac{\sin\phi_{M/\Gamma}}{\left|\cos\phi_{M/\Gamma}\right|}. \tag{1.52}$$

The key observables for the study of CP violation are given by the following complex quantities:

$$\begin{aligned}
\lambda_f^{(\mathrm{q})} &\equiv \frac{q}{p}\frac{\overline{A}_f^{(\mathrm{q})}}{A_f^{(\mathrm{q})}} \approx -e^{-i\theta_M^{(\mathrm{q})}}\frac{\overline{A}_f^{(\mathrm{q})}}{A_f^{(\mathrm{q})}}\left[1-\frac{a_\mathrm{q}}{2}\right], \\
\lambda_{\bar{f}}^{(\mathrm{q})} &\equiv \frac{q}{p}\frac{\overline{A}_{\bar{f}}^{(\mathrm{q})}}{A_{\bar{f}}^{(\mathrm{q})}} \approx -e^{-i\theta_M^{(\mathrm{q})}}\frac{\overline{A}_{\bar{f}}^{(\mathrm{q})}}{A_{\bar{f}}^{(\mathrm{q})}}\left[1-\frac{a_\mathrm{q}}{2}\right].
\end{aligned} \tag{1.53}$$

These quantities play a crucial role in the CP asymmetries in the following sections. The instantaneous (non-evolved) decay amplitudes to the final states f, $\bar{f}$ in (1.53) are defined through the $|\Delta F| = 1$ matrix elements:

$$\begin{aligned}
A_f^{(\mathrm{q})} &\equiv A\left(\mathrm{B_q}\to f\right) = \langle f|\,\mathcal{H}_{\mathrm{eff}}^\dagger\,|\mathrm{B_q}\rangle, \quad & \overline{A}_f^{(\mathrm{q})} &\equiv A\left(\overline{\mathrm{B}}_\mathrm{q}\to f\right) = \langle f|\,\mathcal{H}_{\mathrm{eff}}\,|\overline{\mathrm{B}}_\mathrm{q}\rangle, \\
A_{\bar{f}}^{(\mathrm{q})} &\equiv A\left(\mathrm{B_q}\to \bar{f}\right) = \langle \bar{f}|\,\mathcal{H}_{\mathrm{eff}}^\dagger\,|\mathrm{B_q}\rangle, \quad & \overline{A}_{\bar{f}}^{(\mathrm{q})} &\equiv A\left(\overline{\mathrm{B}}_\mathrm{q}\to \bar{f}\right) = \langle \bar{f}|\,\mathcal{H}_{\mathrm{eff}}\,|\overline{\mathrm{B}}_\mathrm{q}\rangle,
\end{aligned} \tag{1.54}$$

where the effective Hamiltonian density $\mathcal{H}_{\rm eff}$ describes the low energy $|\Delta F| = 1$ transitions and it will be introduced in Section 1.3.6 when discussing non-leptonic B decays. The observable λ_f does not depend on any phase convention: the spurious phases in $\theta_M^{(\rm q)}$ are canceled out through the amplitudes ratios in (1.53). This is the case since the quark rephasing phases in $\mathcal{H}_{\rm eff}$ and in $\theta_M^{(\rm q)}$ are the same. The CP transformation phase cancels as we get a $e^{i\varphi_{\rm CP}}$ factor from the amplitude in the denominator $A_f^{(\rm q)}$ (thus $e^{-i\varphi_{\rm CP}}$ in the numerator) when changing $\mathcal{H}_{\rm eff}^\dagger \to \mathcal{H}_{\rm eff}$ and using the CP transformation of the operators in $\mathcal{H}_{\rm eff}$.

1.3.4 Time-Dependent Decay Rates

The evolution of the flavor eigenstates, $\mathrm{B_q}$ and $\overline{\mathrm{B}}_{\rm q}$, is given by (1.34) and (1.35). Thus the amplitudes for the $\mathrm{B_q}$ and $\overline{\mathrm{B}}_{\rm q}$ mesons to decay to some final state f at a time t are given by:

$$A\left(\mathrm{B_q}(t) \to f\right) = \langle f|\,\mathcal{H}_{\rm eff}^\dagger\,\big|\mathrm{B_q}(t)\big\rangle = g_+(t)A_f^{(\rm q)} + \frac{q}{p}g_-(t)\overline{A}_f^{(\rm q)}\,, \tag{1.55}$$

$$A\left(\overline{\mathrm{B}}_{\rm q}(t) \to f\right) = \langle f|\,\mathcal{H}_{\rm eff}\,\big|\overline{\mathrm{B}}_{\rm q}(t)\big\rangle = g_+(t)\overline{A}_f^{(\rm q)} + \frac{p}{q}g_-(t)A_f^{(\rm q)}\,. \tag{1.56}$$

Using (1.49) with the modulus squared of the above amplitudes we find the corresponding decay rates of the $\mathrm{B_q}$ and $\overline{\mathrm{B}}_{\rm q}$ mesons into the final state f and the CP conjugated $\bar{f}$ state:

$$\begin{aligned}
\Gamma[\mathrm{B_q}(t) \to f] &= N_f|A_f^{(\rm q)}|^2\left\{|g_+(t)|^2 + |\lambda_f^{(\rm q)}|^2|g_-(t)|^2 + 2\Re[\lambda_f^{(\rm q)}g_+^*(t)g_-(t)]\right\},\\
\Gamma[\mathrm{B_q}(t) \to \bar{f}] &= N_f|\overline{A}_{\bar{f}}^{(\rm q)}|^2\left|\frac{q}{p}\right|^2\left\{|g_-(t)|^2 + |\lambda_{\bar{f}}^{(\rm q)-1}|^2|g_+(t)|^2 + 2\Re\left[\lambda_{\bar{f}}^{(\rm q)-1}g_+(t)g_-^*(t)\right]\right\},\\
\Gamma[\overline{\mathrm{B}}_{\rm q}(t) \to f] &= N_f|A_f^{(\rm q)}|^2\left|\frac{q}{p}\right|^2\left\{|g_-(t)|^2 + |\lambda_f^{(\rm q)}|^2|g_+(t)|^2 + 2\Re[\lambda_f^{(\rm q)}g_+(t)g_-^*(t)]\right\},\\
\Gamma[\overline{\mathrm{B}}_{\rm q}(t) \to \bar{f}] &= N_f|\overline{A}_{\bar{f}}^{(\rm q)}|^2\left\{|g_+(t)|^2 + |\lambda_{\bar{f}}^{(\rm q)-1}|^2|g_-(t)|^2 + 2\Re\left[\lambda_{\bar{f}}^{(\rm q)-1}g_+^*(t)g_-(t)\right]\right\},
\end{aligned} \tag{1.57}$$

where $N_f = N_{\bar{f}}$ is a normalization factor arising from kinematics. The transition probabilities are simplified when using the approximation $|p/q|^2 = (1+a_{\rm q})$ and ignoring higher corrections in $|\Gamma_{12}^{(\rm q)}/M_{12}^{(\rm q)}|$:

$$\begin{aligned}
\Gamma[\mathrm{B_q}(t) \to f] &= N_f\,|A_f^{(\rm q)}|^2\,e^{-\Gamma_{\rm q}t}\left\{\frac{1+\left|\lambda_f^{(\rm q)}\right|^2}{2}\,\cosh\frac{\Delta\Gamma_{\rm q}\,t}{2}\right. \\
&\left. +\frac{1-\left|\lambda_f^{(\rm q)}\right|^2}{2}\,\cos(\Delta M_{\rm q}\,t) - \Re\lambda_f^{(\rm q)}\,\sinh\frac{\Delta\Gamma_{\rm q}\,t}{2} - \Im\lambda_f^{(\rm q)}\,\sin\left(\Delta M_{\rm q}\,t\right)\right\},
\end{aligned} \tag{1.58}$$

$$\begin{aligned}
\Gamma[\overline{\mathrm{B}}_{\rm q}(t) \to f] &= N_f\,|A_f^{(\rm q)}|^2\,(1+a_{\rm q})\,e^{-\Gamma_{\rm q}t}\left\{\frac{1+\left|\lambda_f^{(\rm q)}\right|^2}{2}\,\cosh\frac{\Delta\Gamma_{\rm q}\,t}{2}\right. \\
&\left. -\frac{1-\left|\lambda_f^{(\rm q)}\right|^2}{2}\,\cos(\Delta M_{\rm q}\,t) - \Re\lambda_f^{(\rm q)}\,\sinh\frac{\Delta\Gamma_{\rm q}\,t}{2} + \Im\lambda_f^{(\rm q)}\,\sin(\Delta M_{\rm q}\,t)\right\},
\end{aligned} \tag{1.59}$$

$$\Gamma[\mathrm{B_q}(t)\to\bar f] = N_f\left|\overline{A}_{\bar f}^{(\mathrm{q})}\right|^2(1-a_\mathrm{q})\,e^{-\Gamma_\mathrm{q}t}\left\{\frac{1+|\lambda_{\bar f}^{(\mathrm{q})}|^{-2}}{2}\cosh\frac{\Delta\Gamma_\mathrm{q}\,t}{2}\right. \tag{1.60}$$
$$\left.-\frac{1-|\lambda_{\bar f}^{(\mathrm{q})}|^{-2}}{2}\cos(\Delta M_\mathrm{q}\,t)-\Re\frac{1}{\lambda_{\bar f}^{(\mathrm{q})}}\sinh\frac{\Delta\Gamma_\mathrm{q}\,t}{2}+\Im\frac{1}{\lambda_{\bar f}^{(\mathrm{q})}}\sin(\Delta M_\mathrm{q}\,t)\right\},$$

$$\Gamma[\overline{\mathrm{B}}_\mathrm{q}(t)\to\bar f] = N_f\left|\overline{A}_{\bar f}^{(\mathrm{q})}\right|^2e^{-\Gamma_\mathrm{q}t}\left\{\frac{1+|\lambda_{\bar f}^{(\mathrm{q})}|^{-2}}{2}\cosh\frac{\Delta\Gamma_\mathrm{q}\,t}{2}\right. \tag{1.61}$$
$$\left.+\frac{1-|\lambda_{\bar f}^{(\mathrm{q})}|^{-2}}{2}\cos(\Delta M_\mathrm{q}\,t)-\Re\frac{1}{\lambda_{\bar f}^{(\mathrm{q})}}\sinh\frac{\Delta\Gamma_\mathrm{q}\,t}{2}-\Im\frac{1}{\lambda_{\bar f}^{(\mathrm{q})}}\sin(\Delta M_\mathrm{q}\,t)\right\}.$$

The above decay rates can be further simplified for decays for which no CP violation is expected. For instance flavor-specific decays have by definition $\overline{A}_f^{(\mathrm{q})}=A_{\bar f}^{(\mathrm{q})}=0$, and hence $\lambda_f^{(\mathrm{q})}=1/\lambda_{\bar f}^{(\mathrm{q})}=0$. An example of flavor-specific decay is $\mathrm{B_s}\to\mathrm{D_s}\pi$, where $\mathrm{B_s}$ can decay to the final state $f=\mathrm{D_s^-}\pi^+$, while $\overline{\mathrm{B}}_\mathrm{s}$ cannot. The flavor-specific decay $\mathrm{B_s}\to\mathrm{D_s}\pi$ can be used to measure ΔM_s through the mixing (or flavor) asymmetry:

$$\mathcal{A}_{\mathrm{fs}}(t)=\frac{\Gamma[\overline{\mathrm{B}}_\mathrm{s}(t)\to f]-\Gamma[\mathrm{B_s}(t)\to\bar f]}{\Gamma[\overline{\mathrm{B}}_\mathrm{s}(t)\to f]+\Gamma[\mathrm{B_s}(t)\to\bar f]}=-\frac{\cos(\Delta M_\mathrm{q}\,t)}{\cosh\frac{\Delta\Gamma_\mathrm{q}\,t}{2}}, \tag{1.62}$$

where we assumed $a_\mathrm{q}=0$ (i.e. no CP violation in the mixing) and $|A_f^{(\mathrm{q})}|=|\overline{A}_{\bar f}^{(\mathrm{q})}|$ (i.e. no CP violation in the decay amplitudes), and we used (1.64) and (1.65) given hereafter:

$$\Gamma[\mathrm{B_q}(t)\to f] = N_f\,|A_f^{(\mathrm{q})}|^2\,\frac{e^{-\Gamma_\mathrm{q}t}}{2}\left[\cosh\frac{\Delta\Gamma_\mathrm{q}\,t}{2}+\cos(\Delta M_\mathrm{q}\,t)\right], \tag{1.63}$$

$$\Gamma[\overline{\mathrm{B}}_\mathrm{q}(t)\to f] = N_f\,|A_f^{(\mathrm{q})}|^2\,(1+a_\mathrm{q})\,\frac{e^{-\Gamma_\mathrm{q}t}}{2}\left[\cosh\frac{\Delta\Gamma_\mathrm{q}\,t}{2}-\cos(\Delta M_\mathrm{q}\,t)\right], \tag{1.64}$$

$$\Gamma[\mathrm{B_q}(t)\to\bar f] = N_f\,|\overline{A}_{\bar f}^{(\mathrm{q})}|^2\,(1-a_\mathrm{q})\,\frac{e^{-\Gamma_\mathrm{q}t}}{2}\left[\cosh\frac{\Delta\Gamma_\mathrm{q}\,t}{2}-\cos(\Delta M_\mathrm{q}\,t)\right], \tag{1.65}$$

$$\Gamma[\overline{\mathrm{B}}_\mathrm{q}(t)\to\bar f] = N_f\,|\overline{A}_{\bar f}^{(\mathrm{q})}|^2\,\frac{e^{-\Gamma_\mathrm{q}t}}{2}\left[\cosh\frac{\Delta\Gamma_\mathrm{q}\,t}{2}+\cos(\Delta M_\mathrm{q}\,t)\right]. \tag{1.66}$$

Note that flavor-specific decays can be used to look for CP violation in the mixing. Indeed, the corrections to the flavor asymmetry (1.62), still assuming equal amplitudes for the conjugated decays, will be proportional to the parameter a_q. The $\mathrm{B_q}$–$\overline{\mathrm{B}}_\mathrm{q}$ oscillations are shown in Figure 1.6 for both the $\mathrm{B_d}$ and the $\mathrm{B_s}$ systems.

1.3.5 CP Violation in the $\mathrm{B_q}$–$\overline{\mathrm{B}}_\mathrm{q}$ System

There are three phase-convention independent physical CP-violating observables, each exhibiting CP violation whenever we have:

$$\left|\frac{q}{p}\right|\neq1, \quad \text{or} \quad \left|\frac{\overline{A}_{\bar f}^{(\mathrm{q})}}{A_f^{(\mathrm{q})}}\right|\neq1, \quad \text{or} \quad \arg\lambda_f^{(\mathrm{q})}+\arg\lambda_{\bar f}^{(\mathrm{q})}\neq0.$$

To each of these three observables corresponds a type of CP violation, that we briefly discuss below. All types of CP violation have been measured.

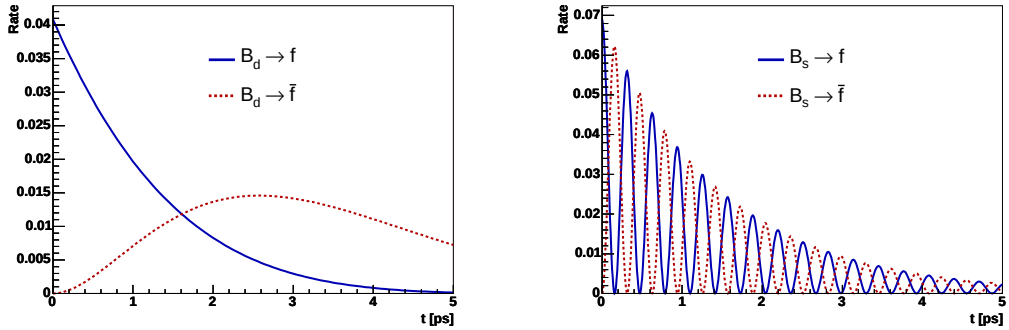


Figure 1.6: Probability for a produced B_d (left) or B_s (right) meson to decay to a flavor-specific final state f or $\bar{f}$ after a proper time t [ps]. In this example $\Delta M_d = 0.5\ \text{ps}^{-1}$, $\Delta\Gamma_d = 0$, $a_d = 0$, and $\Delta M_s = 20\ \text{ps}^{-1}$, $\Delta\Gamma_s/\Gamma_s = 10\%$, $a_s = 0$, and $1/\Gamma_q = 1.5\ \text{ps}^{-1}$.

CP violation in the mixing

CP violation in the mixing (or indirect CP violation) occurs when the relative phase $\phi_{M/\Gamma}$ between $M_{12}^{(q)}$ and $\Gamma_{12}^{(q)}$ does not vanish. The deviation of $|q/p|$ from unity is characterized by the parameter a_q given in (1.52). If CP in the mixing is violated, then the mass eigenstates are different from the CP eigenstates, as it can be seen from (1.43). Indeed, if CP is violated, then the two physical eigenstates are not orthogonal since $\langle B_H\ |B_L\rangle = |p|^2 - |q|^2$.

CP violation in the mixing is small for the B mesons, of the order of 10^{-2}. This type of CP violation can be tested with semi-leptonic B decays, and other flavor-specific decays.

CP violation in the decay

The quantity $|\overline{A}_{\bar{f}}^{(q)}/A_f^{(q)}|$ is phase convention independent, and when it differs from one we have the so-called CP violation in the decay amplitudes, also named direct CP violation. There are two kind of complex phases involved in an amplitude, namely:

Strong phases These phases are CP even (or CP violating) and arise from absorptive parts of the decay amplitude, e.g. in final states interactions dominated by the strong interaction.

Weak phases These phases are CP odd (or CP conserving) and originate from complex couplings in the Lagrangian density describing the CKM matrix.

We can thus write the decay amplitude as:

$$A_f^{(q)} = \sum_i A_i e^{i(\delta_i+\phi_i)}, \quad \overline{A}_{\bar{f}}^{(q)} = \sum_i A_i e^{i(\delta_i-\phi_i)},$$

where i labels the contributions, A_i are the magnitudes of each contribution, and δ_i and ϕ_i denote the strong and weak phases, respectively. Note that only the differences between two weak or strong phases are physical. From the above definitions we obviously have CP violation when $|\overline{A}_{\bar{f}}^{(q)}/A_f^{(q)}| \neq 1$, i.e. direct CP violation arises from the clash between the phases of two interfering decay amplitudes in the total decay amplitude.

Mixing-induced CP violation

The so-called mixing-induced CP violation arises from a clash between the phase of q/p and the phases of the decay amplitudes. Even when CP is conserved in the mixing and in the decay amplitudes, we may still have CP violation in this interference.

A particularly interesting situation where mixing-induced CP violation occurs is that of decays to CP eigenstates, i.e. $f = \bar{f}$ such that $\lambda_f^{(q)} = \lambda_{\bar{f}}^{(q)}$. In that case, the condition for the presence of this kind of CP violation is:

$$\arg \lambda_f^{(q)} + \arg \lambda_{\bar{f}}^{(q)} \neq 0 \quad \rightarrow \quad \Im\lambda_f^{(q)} \neq 0\,. \tag{1.67}$$

The interplay between the B_q–$\overline{B}_q$ mixing and the subsequent decay of the B_q meson to the final state $f = \bar{f}$ is illustrated in Figure 1.7. We define the CP eigenvalue $\eta_f = \pm 1$ of the CP eigenstate final state:

$$\left|\bar{f}\right\rangle = (CP)\,|f\rangle = \eta_f\,|f\rangle\,. \tag{1.68}$$

For decays to CP eigenstates dominated by a single CKM phase, e.g. $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions, then there is no CP violation in the decay amplitudes and we can directly probe the B_q–$\overline{B}_q$ mixing phase. This will be detailed in the Section 1.4 when looking at $B_s \to \eta_c\phi$, $B_s \to D_sD_s$, $B_s \to J/\psi\eta$, and $B_s \to J/\psi\phi$ decays.

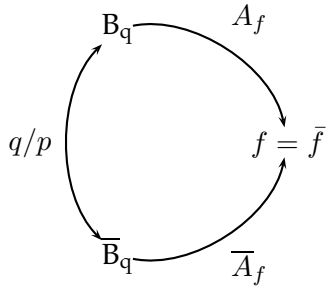


Figure 1.7: Mixing-induced CP violation.

CP asymmetry

The time-dependent CP asymmetry into CP eigenstates is defined by:

$$\mathcal{A}_{CP}(t) = \frac{\Gamma[\overline{B}_q(t) \to f] - \Gamma[B_q(t) \to f]}{\Gamma[\overline{B}_q(t) \to f] + \Gamma[B_q(t) \to f]}\,. \tag{1.69}$$

We can then use the decay rates (1.58) and (1.59) to obtain:

$$\mathcal{A}_{CP}(t) = -\frac{\left(1 - \left|\lambda_f^{(q)}\right|^2\right)\cos(\Delta M_q\, t) - 2\,\Im\lambda_f^{(q)}\sin(\Delta M_q\, t)}{\left(1 + \left|\lambda_f^{(q)}\right|^2\right)\cosh(\Delta\Gamma_q\, t/2) - 2\,\Re\lambda_f^{(q)}\sinh(\Delta\Gamma_q\, t/2)} + \mathcal{O}(a_q)\,. \tag{1.70}$$

This asymmetry is non-zero if any type of CP violation occurs. In particular we can have a non-vanishing mixing-induced CP asymmetry with $\Im\lambda_f^{(q)} \neq 0$ and $\left|\lambda_f^{(q)}\right|^2 = 1$.

In order to explicitly show the origin of each term in the asymmetry (1.70), we can express $\lambda_f^{(q)}$ in terms of time-independent "asymmetries":

$$\lambda_f^{(q)} \equiv \frac{-1}{1 + \mathcal{A}_{CP}^{dir}}\left(\mathcal{A}_{\Delta\Gamma_q} + i\mathcal{A}_{CP}^{mix\text{-}ind}\right)\,, \quad \left|\mathcal{A}_{\Delta\Gamma_q}\right|^2 + \left|\mathcal{A}_{CP}^{mix\text{-}ind}\right|^2 + \left|\mathcal{A}_{CP}^{dir}\right|^2 = 1\,, \tag{1.71}$$

where we defined:

$$\mathcal{A}_{\Delta\Gamma_q} \equiv -\frac{2\Re\lambda_f^{(q)}}{1+\left|\lambda_f^{(q)}\right|^2}\,, \quad \mathcal{A}_{CP}^{\text{mix-ind}} \equiv -\frac{2\Im\lambda_f^{(q)}}{1+\left|\lambda_f^{(q)}\right|^2}\,, \quad \mathcal{A}_{CP}^{\text{dir}} \equiv \frac{1-\left|\lambda_f^{(q)}\right|^2}{1+\left|\lambda_f^{(q)}\right|^2}\,. \tag{1.72}$$

A non-vanishing $\mathcal{A}_{CP}^{\text{dir}}$ means direct CP violation, and $\mathcal{A}_{CP}^{\text{mix-ind}}$ measures mixing-induced CP violation. The third asymmetry, $\mathcal{A}_{\Delta\Gamma_q}$, is related to the presence of a non-negligible $\Delta\Gamma_q$.

Using these definitions, the asymmetry given in (1.70) now reads:

$$\mathcal{A}_{CP}(t) = -\frac{\mathcal{A}_{CP}^{\text{dir}}\cos(\Delta M_q\, t) + \mathcal{A}_{CP}^{\text{mix-ind}}\sin(\Delta M_q\, t)}{\cosh(\Delta\Gamma_q\, t/2) + \mathcal{A}_{\Delta\Gamma_q}\sinh(\Delta\Gamma_q\, t/2)} + \mathcal{O}(a_q)\,. \tag{1.73}$$

If we consider decays into CP eigenstates whose decay amplitudes are dominated by only one CKM phase, such as $\bar{b} \to \bar{c}c\bar{s}$ transitions, then the term $\mathcal{A}_{CP}^{\text{dir}}$ vanishes and we get for the mixing-induced time-dependent CP asymmetry:

$$\mathcal{A}_{CP\,B_q}^{\text{mix-ind}}(t) = -\frac{\eta_f \sin\phi_q \sin(\Delta M_q\, t)}{\cosh(\Delta\Gamma_q\, t/2) - \eta_f \cos\phi_q \sinh(\Delta\Gamma_q\, t/2)}\,. \tag{1.74}$$

where we have identified ϕ_q and $\theta_M^{(q)}$ through (1.40) and (1.41), with the approximation (1.53), and the cancellation of the spurious phases with the ratio of the amplitudes. Furthermore, we used the parameterization given in Section 1.2.1 in which the tree phase of the $b \to c\bar{c}s$ transition is real to a good approximation, since we have $\phi_D \equiv \arg[V_{cb}V_{cs}^*] = \mathcal{O}(\lambda^6)$ such that $\phi_D \approx 0$. We thus see that we can directly probe the B_q–$\overline{B}_q$ mixing phase ϕ_q with these particular decays. Note that there is a priori a four-fold ambiguity in $\phi_{M/\Gamma} \sim \phi_q$. The determination of the sign of a_q, see (1.52), would reduce the ambiguity to a two-fold one by determining the sign of $\sin\phi_{M/\Gamma}$. Moreover, since $\Delta\Gamma_q$ and $\cos\phi_{M/\Gamma}$ have the same sign, see (1.51), the measurement of the asymmetry $\mathcal{A}_{CP}^{\text{mix-ind}}$ will not allow to resolve the two-fold ambiguity in the sign of $\cos\phi_{M/\Gamma} \approx \pm 1$ in the SM (i.e. ambiguity at $\pi - \phi_s$). Different methods to unambiguously determine $\phi_{M/\Gamma}$ can be found in [22] for the weak mixing phase of B_s mesons.

In the case of the B_d-$\overline{B}_d$ mixing, we have to a good approximation $\Delta\Gamma_d \sim 0$, and (1.74) simplifies to:

$$\mathcal{A}_{CP\,B_d}^{\text{mix-ind}}(t) = -\eta_f \sin\phi_d \sin(\Delta M_d\, t)\,, \tag{1.75}$$

corresponding to the famous "$\sin 2\beta$ measurement" ($\phi_d \approx 2\beta$).

1.3.6 Effective Hamiltonians

The exploration of CP violation requires the knowledge of hadronic matrix elements. Purely hadronic decays consist of final states involving only quarks such that QCD comes into play. The evaluation of the different hadronic parameters is difficult, and bears large uncertainties. The use of low-energy effective Hamiltonians offers a theoretical framework allowing the calculation of the different transition amplitudes. We will hereafter only outline the structure of the effective Hamiltonians. A complete review can be found in [37], which we use here as baseline.

Classification

The most complicated b-hadron decays are the non-leptonic transitions, which are mediated by $b \to q_1\bar{q}_2 d(s)$ quark-level processes, with $q_1, q_2 \in \{u, c\}$ or $q_1, q_2 \in \{d, s\}$. We may distinguish two kinds of topologies contributing to such decays: "tree" and "penguin" topologies. The latter consist of gluonic (QCD) and electroweak (EW) penguins. We show in Figures 1.8 – 1.10 the corresponding leading-order Feynman diagrams. We may classify the $b \to q_1\bar{q}_2 d(s)$ decays depending on the flavor content of their final states:

- $q_1, q_2 \in \{u, c\}$ and $q_1 \neq q_2$: *only* tree diagrams contribute (at first leading orders).
- $q_1, q_2 \in \{u, c\}$ and $q_1 = q_2$: tree *and* penguin competing diagrams contribute.
- $q_1, q_2 \in \{d, s\}$ ($q_1 = q_2$ or $q_1 \neq q_2$): *only* penguin diagrams contribute.

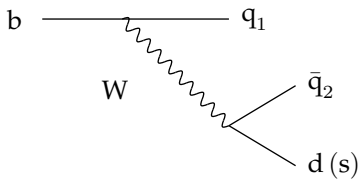


Figure 1.8: Tree diagrams ($q_1, q_2 \in \{u, c\}$).

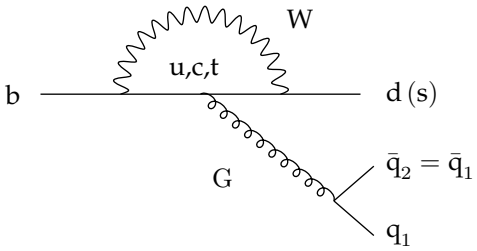


Figure 1.9: QCD penguin diagrams ($q_1 = q_2 \in \{u, d, c, s\}$).

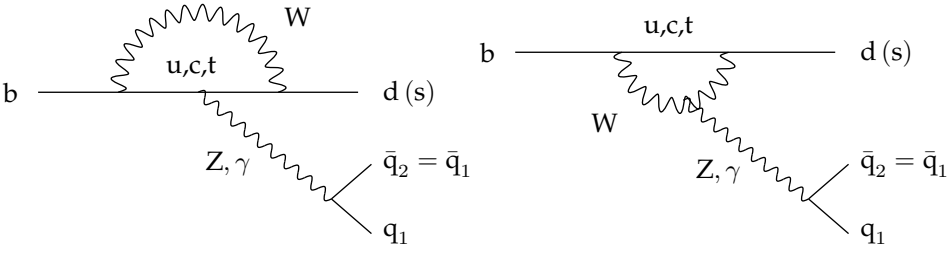


Figure 1.10: Electroweak penguin diagrams ($q_1 = q_2 \in \{u, d, c, s\}$).

From the above classification, we will thus have both penguin and tree contributions for $\bar{\mathrm{b}} \to \bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}$ quark-level transitions. However, we will show in Section 1.4.1 that the penguin pollution is doubly Cabibbo-suppressed and can therefore be neglected.

Operator product expansion

The transition amplitudes for non-leptonic decays are calculated using low-energy effective Hamiltonians which are expressed using the Operator Product Expansion (OPE), which factorizes QCD and weak effects. The transition amplitudes have the following structure:

$$\langle f|\mathcal{H}_{\text{eff}}|i\rangle = \frac{G_{\text{F}}}{\sqrt{2}}\lambda_{\text{CKM}}\sum_k C_k(\mu)\langle f|Q_k(\mu)|i\rangle\,, \tag{1.76}$$

where λ_{CKM} carries the CKM matrix elements. The Fermi coupling constant G_{F} describes the point-like interaction after "integrating out" the W boson ($G_{\text{F}}/\sqrt{2} = g^2/8m_{\text{W}}^2$) at the renormalization scale μ. The OPE technique allows to separate the short-distance contributions (C_k) to this amplitude from the long-distance contributions (Q_k). The Wilson coefficients $C_k(\mu)$ are perturbative quantities, while the non-perturbative Q_k are local operators which govern "effectively" the decay in question. Informally, the $C_k(\mu)$ are the scale-dependent couplings of the hadronic matrix elements $\langle f|Q_k(\mu)|i\rangle$. The short-distance part contains the information on the integrated heavy fields (given the scale μ), which are treated as dynamical degrees of freedom.

Local four-quark operators

The exploration of CP violation requires different local operators in order to describe both tree and penguin decays. For the latter class, with quark flavor $\mathrm{r} \in \{\mathrm{d},\mathrm{s}\}$, only two independent weak amplitudes contribute to any given decay since the unitarity of the CKM matrix relates the CKM factors through:

$$V_{\text{ur}}^* V_{\text{ub}} + V_{\text{cr}}^* V_{\text{cb}} + V_{\text{tr}}^* V_{\text{tb}} = 0 \quad \mathrm{r} \in \{\mathrm{d},\mathrm{s}\}\,, \tag{1.77}$$

The effective Hamiltonian for $|\Delta F| = |\Delta B| = 1$ transitions has thus the following structure:

$$\mathcal{H}_{\text{eff}} = \frac{G_{\text{F}}}{\sqrt{2}}\sum_{\mathrm{j}=\mathrm{u},\mathrm{c}} V_{\text{jr}}^* V_{\text{jb}} \left\{ \sum_{k=1}^{2} C_k(\mu)\, Q_k^{\text{jr}} + \sum_{k=3}^{10} C_k(\mu)\, Q_k^{\text{r}} \right\}\,, \tag{1.78}$$

with $\mathrm{j} \in \{\mathrm{u},\mathrm{c}\}$, and where the top quark and the W boson have been "integrated out". The different local four-quark operators Q_k^{jr} are combinations of the quark spinors involved in the processes, where a list is given in [37, 20].

Off-diagonal term $\boldsymbol{\Gamma_{12}}$

The off-diagonal term $\Gamma_{12}^{(\mathrm{q})}$ arises from final states f which are common to both $\mathrm{B_q}$ and $\overline{\mathrm{B}}_{\mathrm{q}}$, including all pairs of CP-conjugate decay modes. Even though $\Gamma_{12}^{(\mathrm{q})}$ receives contributions from loops involving light u and c quarks, it is dominated by CKM-favored tree diagrams. Schematically $\Gamma_{12}^{(\mathrm{q})}$ is given by $|\Delta B| = 1$ flavor transitions:

$$\begin{aligned} \Gamma_{12}^{(\mathrm{q})} &\sim \sum_f \left\langle \mathrm{B_q}\right| \mathcal{H}_{\text{eff}}^{|\Delta B|=1} \left|f\right\rangle \left\langle f\right| \mathcal{H}_{\text{eff}}^{|\Delta B|=1} \left|\overline{\mathrm{B}}_{\mathrm{q}}\right\rangle \\ &= A_f^{(\mathrm{q})*}\overline{A}_f^{(\mathrm{q})} + A_{\bar{f}}^{(\mathrm{q})*}\overline{A}_{\bar{f}}^{(\mathrm{q})} + A_g^{(\mathrm{q})*}\overline{A}_g^{(\mathrm{q})} + \dots\,. \end{aligned}$$

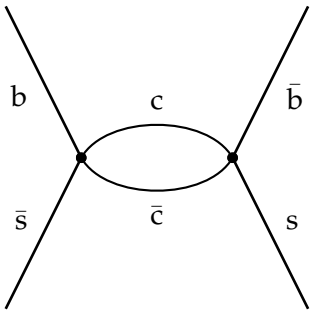


Figure 1.11: OPE diagram for the B_s width difference.

The sum extends over all decay modes f common to the B_q and thus involves spectator quarks. The width is dominated by the decay $b \to c\bar{c}s$ part of the weak Hamiltonian. The small correction involving $V_{ub}V_{us}^*$ is further doubly Cabibbo-suppressed by $\lambda = |V_{us}| \approx 0.22$ [38]. The OPE for the B_s width difference $\Delta\Gamma_s \approx 2|\Gamma_{12}^{(s)}|\cos\phi_{M/\Gamma}$ is illustrated in Figure 1.11.

The CP-violating phase defined in (1.50) becomes for the B_s system:

$$\phi_{M/\Gamma} = \phi_s - \arg\left[-\Gamma_{12}^{(s)}\right] = \phi_s - (2\arg[V_{cb}V_{cs}^*] + \mathcal{O}(\lambda^2)) \approx \phi_s - \mathcal{O}(\lambda^6) \approx \phi_s\,, \qquad (1.79)$$

where we used the parameterization of Section 1.2.1.

1.4 B_s Decays through $\bar{b} \to \bar{c}c\bar{s}$ Quark-Level Transitions

The study of B_s decays to CP eigenstates proceeding through $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions will enable the determination of the B_s–$\overline{B}_s$ mixing phase ϕ_s, as motivated in Section 1.3. These decays are dominated by the tree-diagram weak phase $\arg[V_{cb}^*V_{cs}]$, and receive negligible – doubly Cabibbo-suppressed – contributions from penguin diagrams. Therefore there is no CP violation in the decay amplitudes, and neglecting CP violation in the mixing directly probes the mixing phase. The phase mismatch between the decay and mixing weak phases thus allows to perform a time-dependent mixing-induced CP measurement.

There are several decays of interest for the determination of ϕ_s:

- Pure CP-even eigenstates: $B_s \to \eta_c\phi$, $B_s \to D_sD_s$, $B_s \to J/\psi\eta^{(\prime)}$ ... . Note that $B_s \to D_sD_s$ is color-allowed, and we may have to deal with final state interactions effects, as discussed in Subsection 1.4.1.

- Admixture of CP eigenstates: $B_s \to J/\psi\phi$. This is a decay to two vectors such that we need to consider different polarization states, with CP eigenvalues $\eta_f = +1, -1, +1$. The corresponding components will be denoted by $0, \perp, \parallel$, respectively.

The decays to pure CP eigenstates have the advantage that they do not require any angular analysis to separate the CP components. However, the expected annual yields at LHCb are one order of magnitude smaller compared to $B_s \to J/\psi\phi$, as presented in Chapter 6.

The decay $B_s \to J/\psi\phi$ requires an angular analysis in order to disentangle the CP eigenstates, at a cost of statistical power. This analysis can be carried out in the so-called transversity basis, as described in Subsection 1.4.3.

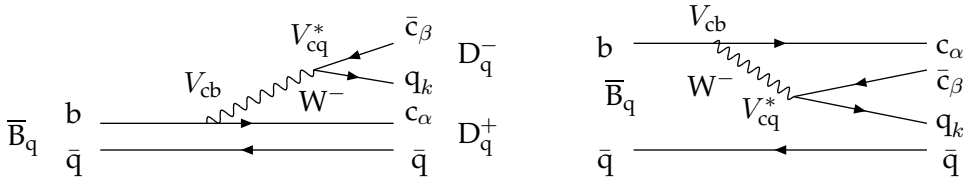


Figure 1.12: Tree-level $b \to c\bar{c}s$ diagrams ($\bar{q} \in \{\bar{d}, \bar{s}\}$). Left: color-allowed diagram, with two charged mesons. Right: color-suppressed diagram, with two neutral mesons.

1.4.1 Penguin Pollution

As presented in Subsection 1.3.6, we have to deal with both tree and penguin diagrams when considering the decay amplitudes for $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions. Considering first the tree-level diagrams for $b \to c\bar{c}s$ quark-level transitions, the initial b hadron can hadronize into two mesons in two different ways, as shown in Figure 1.12. The two possibilities have a different color structure, where the subscripts denote the color indices:

Color-allowed The two final mesons D^+_q and D^-_q are charged with a quark content $c_\alpha \bar{q}$ and $\bar{c}_\beta q_k$ respectively. The $\bar{c}_\beta$ and q_k quarks arise in a color singlet from the charged current forming a color singlet meson. This quark transition is denoted by $b \to c_\alpha[\bar{c}_\beta q_k]$.

Color-suppressed The two final mesons (e.g. $\eta_c\phi$, $J/\psi\eta^{(\prime)}$, $J/\psi\phi$) are neutral mesons with quark content $c_\alpha \bar{c}_\beta$ and $q_k \bar{q}$. In this case the colors of the c_α and $\bar{c}_\beta$, or q_k and $\bar{q}$, are initially independent and then form a color singlet meson. This quark transition is denoted by $b \to [c_\alpha \bar{c}_\beta] q_k$.

If we hinder rescattering effects, the second possibility is suppressed compared to the first one. Roughly speaking the open charm combination is favored because of the extra degree of freedom given by the different quark types entering the final bound mesons. The penguins topologies may in the color-allowed case contain important contributions from the final state interaction effects. Nevertheless, we will ignore this in our discussion and consider the $B_s \to D_s D_s$ decay as a candidate for the determination of ϕ_s through the CP violating mixing-induced effects, as considered in [22, 21]. Note that if the final state interactions are indeed not negligible, then we can use $B_s \to D_s D_s$ together with $B_d \to DD$ to determine the angle γ, assuming the U-spin symmetry[6] of the strong interactions [38].

Following [38], we can express the $\bar{b} \to \bar{c}c\bar{s}$ amplitude as a combination of the tree and the EW and QCD penguins contributions. The amplitude reads:

$$A(\bar{b} \to \bar{c}c\bar{s}) \;=\; V_{cs}V^*_{cb}(A_T + P_c) + V_{us}V^*_{ub}P_u + V_{ts}V^*_{tb}P_t\,, \tag{1.80}$$

where the P_i denote the penguin amplitudes with internal $i \in \{u, c, t\}$ quarks, and A_T stands for the tree contribution. We next make use of the unitarity triangle (sb), see (1.7), to eliminate $V_{ts}V^*_{tb}$ with $V_{ts}V^*_{tb} = -V_{us}V^*_{ub} - V_{cs}V^*_{cb}$, and we get:

$$A(\bar{b} \to \bar{c}c\bar{s}) \;=\; V_{cs}V^*_{cb}(A_T + P_c - P_t) + V_{us}V^*_{ub}(P_u - P_t)\,. \tag{1.81}$$

[6]The U-spin symmetry corresponds to the replacement of the d by the s quark, such that U-spin symmetry (i.e. a rotation) is the subgroup of SU(3) relating the d and s quarks. This approximation is fair for the B_d and B_s as their masses are well above the mass of the spectator quarks.

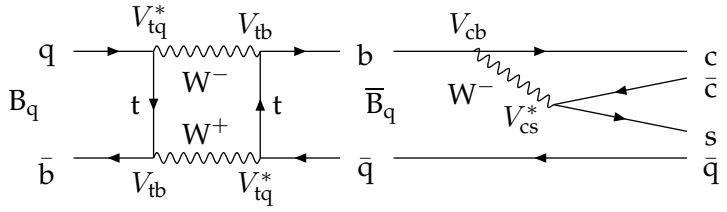


Figure 1.13: Mixing-induced CP violation and $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions.

The contribution $(P_u - P_t)$ is highly suppressed with respect to the contribution with the tree amplitude $(A_T + P_c - P_t)$. Indeed, using the parameterization of Section 1.2.1, and the relations (1.9) and (1.11) we have $V_{cs}V_{cb}^* \sim A\lambda^2(1-\lambda^2/2)$ and $V_{us}V_{ub}^* \sim A\lambda^4(\rho + i\eta)$ such that the ratio of contributions $(P_u - P_t)/(A_T + P_c - P_t)$ is suppressed by two powers of λ. We therefore have that the $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions are to a good approximation dominated by a single weak phase given in terms of the CKM factors $V_{cs}V_{cb}^*$.

1.4.2 B_s Decay Rates in $\bar{b} \to \bar{c}c\bar{s}$ Transitions

In the mixing-induced CP violation, the B_s (or $\overline{B}_s$) must first oscillate, and then decay into the final states, as shown in the Feynman diagrams of Figure 1.13. In order to derive the decay rates for the $\bar{b} \to \bar{c}c\bar{s}$ channels using the results of Section 1.3, we first need to determine the complex observable $\lambda_f^{(q)}$ for the B_s system and the $\bar{b} \to \bar{c}c\bar{s}$ transitions. We summarize hereafter the quantities and definitions that we need in order to determine $\lambda_f^{(s)}$, assuming no CP violation in the mixing and the dominance of a single tree phase for the $\bar{b} \to \bar{c}c\bar{s}$ transitions.

CP complex observable We define $\lambda_f^{(s)}$ with:

$$\lambda_f^{(s)} \equiv \frac{q}{p}\frac{\overline{A}_f^{(s)}}{A_f^{(s)}}\,. \tag{1.82}$$

B_s–$\overline{B}_s$ mixing phase We define the SM value of ϕ_s of the Hamiltonian describing the mixing with:

$$\phi_s^{SM} \equiv 2\arg[V_{ts}^* V_{tb}] \sim \mathcal{O}(-0.04)\ \text{rad}\,. \tag{1.83}$$

We have $\phi_s^{SM} \approx -2\beta_s$ ($\beta_s \equiv \chi$), and $\beta_s \approx \arg(V_{ts}) - \pi$ where β_s is defined in (1.8), and corresponds to one of the angles of the squashed (sb) triangle, see (1.7).

Tree decay phase The $b \to c\bar{c}s$ transitions are dominated by a single tree phase ϕ_D:

$$\phi_D \equiv \arg[V_{cb}V_{cs}^*] = \mathcal{O}(\lambda^6) \sim 0\,. \tag{1.84}$$

The CP eigenvalues of the final states ($\eta_f = \pm 1$) are defined through

$$\left|\bar{f}\right\rangle = (CP)\left|f\right\rangle = \eta_f \left|f\right\rangle\,.$$

The decay amplitudes controlled by an effective operator $\mathcal{O}$ with weak phase ϕ_D are then given by:

$$\begin{aligned} \overline{A}_f^{(s)} &= e^{i\phi_D}\left\langle f\right|\mathcal{O}\left|\overline{B}_q\right\rangle\,, \quad A_f^{(s)} = e^{-i\phi_D}\left\langle f\right|\mathcal{O}^\dagger\left|B_q\right\rangle\,, \\ \Rightarrow \overline{A}_f^{(s)} &= \eta_f e^{i(\phi_D - \varphi_{CP})}\left\langle f\right|\mathcal{O}^\dagger\left|B_q\right\rangle = \eta_f e^{i(2\phi_D - \varphi_{CP})}A_f^{(s)}\,, \end{aligned} \tag{1.85}$$

where we explicitly included the CP transformation phase to illustrate its cancellation in $\lambda_f^{(\mathrm{s})}$.

Amplitudes ratio The ratio of amplitudes is given by

$$\frac{\overline{A}_f^{(\mathrm{s})}}{A_f^{(\mathrm{s})}} = \eta_f e^{i(2\phi_\mathrm{D} - \varphi_\mathrm{CP})} \,. \tag{1.86}$$

q/p ratio Assuming no CP violation in the mixing we have:

$$\frac{q}{p} = e^{-i(\phi_\mathrm{s} - \varphi_\mathrm{CP})} \,. \tag{1.87}$$

Combining all the above results, we get the final result for $\lambda_f^{(\mathrm{s})}$, and its imaginary part and real parts:

$$\begin{aligned} \lambda_f^{(\mathrm{s})} &= e^{-i(\phi_\mathrm{s} - \varphi_\mathrm{CP})} \eta_f e^{i(2\phi_\mathrm{D} - \varphi_\mathrm{CP})} = \eta_f e^{-i(\phi_\mathrm{s} - 2\phi_\mathrm{D})} \,, \quad \phi_\mathrm{D} \sim 0 \,, \\ \Im\lambda_f^{(\mathrm{s})} &= -\eta_f \sin\phi_\mathrm{s} \,, \quad \Re\lambda_f^{(\mathrm{s})} = \eta_f \cos\phi_\mathrm{s} \,. \end{aligned} \tag{1.88}$$

The corresponding time-dependent CP asymmetry was already given in (1.74), arising from the phase mismatch between ϕ_s and ϕ_D.

The corresponding decay rates are obtained from (1.58) and (1.59), and using (1.88). We get:

$$\begin{aligned} \Gamma[\mathrm{B_s}(t) \to f] &= N_f \, |A_f^{(\mathrm{s})}|^2 \, e^{-\Gamma_\mathrm{s} t} \Bigg\{ \cosh\frac{\Delta\Gamma_\mathrm{s}\, t}{2} - \eta_f \cos\phi_\mathrm{s} \, \sinh\frac{\Delta\Gamma_\mathrm{s}\, t}{2} \\ &\qquad + \eta_f \sin\phi_\mathrm{s} \, \sin\left(\Delta M_\mathrm{s}\, t\right) \Bigg\} \,, \end{aligned} \tag{1.89}$$

$$\begin{aligned} \Gamma[\overline{\mathrm{B}}_\mathrm{s}(t) \to f] &= N_f \, |A_f^{(\mathrm{s})}|^2 \, e^{-\Gamma_\mathrm{s} t} \Bigg\{ \cosh\frac{\Delta\Gamma_\mathrm{s}\, t}{2} - \eta_f \cos\phi_\mathrm{s} \, \sinh\frac{\Delta\Gamma_\mathrm{s}\, t}{2} \\ &\qquad - \eta_f \sin\phi_\mathrm{s} \, \sin(\Delta M_\mathrm{s}\, t) \Bigg\} \,. \end{aligned} \tag{1.90}$$

For a SM $\mathrm{B_s}$–$\overline{\mathrm{B}}_\mathrm{s}$ mixing phase, and if we neglect quadratic terms in ϕ_s (i.e. $\cos\phi_\mathrm{s} \approx 1$), the above time evolutions correspond to exponential decays with lifetimes $1/\Gamma_\mathrm{H}$ and $1/\Gamma_\mathrm{L}$, where $\Gamma_\mathrm{s} = (\Gamma_\mathrm{H} + \Gamma_\mathrm{L})/2$. This can be seen from the following relations:

$$\Gamma[\mathrm{B_s}(t) \to f_\mathrm{e}] = N_{f_\mathrm{e}} \, |A_{f_\mathrm{e}}^{(\mathrm{s})}|^2 \, \left\{ e^{-\Gamma_\mathrm{L} t} + e^{-\Gamma_\mathrm{s} t} \, \sin\phi_\mathrm{s} \, \sin\left(\Delta M_\mathrm{s}\, t\right) \right\} \,, \tag{1.91}$$

$$\Gamma[\overline{\mathrm{B}}_\mathrm{s}(t) \to f_\mathrm{e}] = N_{f_\mathrm{e}} \, |A_{f_\mathrm{e}}^{(\mathrm{s})}|^2 \, \left\{ e^{-\Gamma_\mathrm{L} t} - e^{-\Gamma_\mathrm{s} t} \, \sin\phi_\mathrm{s} \, \sin\left(\Delta M_\mathrm{s}\, t\right) \right\} \,, \tag{1.92}$$

$$\Gamma[\mathrm{B_s}(t) \to f_\mathrm{o}] = N_{f_\mathrm{o}} \, |A_{f_\mathrm{o}}^{(\mathrm{s})}|^2 \, \left\{ e^{-\Gamma_\mathrm{H} t} - e^{-\Gamma_\mathrm{s} t} \, \sin\phi_\mathrm{s} \, \sin\left(\Delta M_\mathrm{s}\, t\right) \right\} \,, \tag{1.93}$$

$$\Gamma[\overline{\mathrm{B}}_\mathrm{s}(t) \to f_\mathrm{o}] = N_{f_\mathrm{o}} \, |A_{f_\mathrm{o}}^{(\mathrm{s})}|^2 \, \left\{ e^{-\Gamma_\mathrm{H} t} + e^{-\Gamma_\mathrm{s} t} \, \sin\phi_\mathrm{s} \, \sin\left(\Delta M_\mathrm{s}\, t\right) \right\} \,, \tag{1.94}$$

where f_e (e.g. $\mathrm{D_sD_s}$) and f_o (e.g. $\mathrm{J}/\psi\mathrm{f_0}$) denote a CP-even and CP-odd final state, respectively. Note that $\Gamma_\mathrm{s} = 1/\tau_\mathrm{s}$ corresponds to the average lifetime as measured from hadronic decays.

We can then introduce the effect of flavor tagging, i.e. the identification of the original flavor of the b or $\bar{\text{b}}$ quark in the detected B_s meson. The tagging procedure does not always a give correct answer, which results in a dilution of the CP asymmetry through the $\sin\phi_\text{s}$ term with a dilution factor $D = (1 - 2\omega_\text{tag})$, where ω_tag is the probability of having a wrong identification ($\omega_\text{tag} = 1/2$ in case there is no tag). We introduce the *observed* decay rates R, which are defined by:

$$\begin{aligned} R\left[\text{B}_\text{s}(t) \to f\right] &= (1-\omega_\text{tag})\,\Gamma\left[\text{B}_\text{s}(t) \to f\right] + \omega_\text{tag}\,\Gamma\left[\overline{\text{B}}_\text{s}(t) \to f\right] , \\ R\left[\overline{\text{B}}_\text{s}(t) \to f\right] &= \omega_\text{tag}\,\Gamma\left[\text{B}_\text{s}(t) \to f\right] + (1-\omega_\text{tag})\,\Gamma\left[\overline{\text{B}}_\text{s}(t) \to f\right] , \end{aligned}$$

leading to:

$$\begin{aligned} R[\text{B}_\text{s}(t) \to f] &= N_f\,|A_f^{(\text{s})}|^2\,e^{-\Gamma_\text{s} t}\left\{\cosh\frac{\Delta\Gamma_\text{s}\,t}{2} - \eta_f\cos\phi_\text{s}\,\sinh\frac{\Delta\Gamma_\text{s}\,t}{2}\right. \\ &\qquad \left. +\eta_f\,D\,\sin\phi_\text{s}\,\sin\left(\Delta M_\text{s}\,t\right)\right\}, \end{aligned} \tag{1.95}$$

$$\begin{aligned} R[\overline{\text{B}}_\text{s}(t) \to f] &= N_f\,|A_f^{(\text{s})}|^2\,e^{-\Gamma_\text{s} t}\left\{\cosh\frac{\Delta\Gamma_\text{s}\,t}{2} - \eta_f\cos\phi_\text{s}\,\sinh\frac{\Delta\Gamma_\text{s}\,t}{2}\right. \\ &\qquad \left. -\eta_f\,D\,\sin\phi_\text{s}\,\sin(\Delta M_\text{s}\,t)\right\}. \end{aligned} \tag{1.96}$$

We see in particular that even when $D = 0$, we can still have access to ϕ_s through the cosine term. Thus we can obtain useful information from untagged events as well. For a SM value of ϕ_s, i.e. a small value, untagged events are expected to yield a small sensitivity to ϕ_s. However, if there is New Physics with larger values of the phase of $M_{12}^{(\text{s})}$, then the cosine term contribution can become important. Note that transforming $\phi_\text{s} \to -\phi_\text{s}$ is equivalent to interchanging B_s and $\overline{\text{B}}_\text{s}$.

1.4.3 Angular Analysis of the CP Components

In the $\text{B}_\text{s} \to \text{J}/\psi\phi$ decay the spinless B_s decays into two pseudo-vectors with $J^{PC} = 1^{--}$. In the B_s rest frame, the final states will thus have a relative orbital momentum with allowed values of $l = 0, 1, 2$ as a result of total spin conservation. The CP eigenvalues of the $\text{J}/\psi\phi$ final state are then given by $CP(\text{J}/\psi\phi) = CP(\text{J}/\psi)CP(\phi)(-1)^l = +1, -1, +1$. The final state is therefore an admixture of CP-even and CP-odd eigenstates.

In order to perform a lifetime measurement, and moreover a CP measurement, we need to separate the different CP eigenstates. The relevant decay mode we consider is B_s $\to \text{J}/\psi(\to \ell^+\ell^-)\phi(\to \text{K}^+\text{K}^-)$ with four final decay products. Consequently, the directions of their momenta enable to distinguish the different angular states. This can be done in terms of the helicity formalism, where the J/ψ and ϕ have the same $(2s+1)$ helicity ($\lambda = \vec{s}\cdot\vec{p}/|\vec{s}||\vec{p}|$) states ($\lambda = -1, 0, +1$) since the B_s has no spin. Using the Wigner matrices $\mathcal{D}^j_{\lambda m'}(R)$ (matrix representation of the rotation operator R [36]), we can go to the helicity frame in which the transition amplitude can be written in terms of the different helicity states.

The general decay amplitude for the $\text{B}_\text{s} \to \text{J}/\psi\phi$ decay can be expressed in terms of linear polarization states of the J/ψ and ϕ vector mesons [39, 40]:

$$A(B_s \to \text{J}/\psi\,\phi) = A_0\,(m_\phi/E_\phi)\;\epsilon^{*L}_{\text{J}/\psi}\,\epsilon^{*L}_\phi - A_\parallel\,\epsilon^{*T}_{\text{J}/\psi}\cdot\epsilon^{*T}_\phi/\sqrt{2} - iA_\perp\,\epsilon^*_{\text{J}/\psi}\times\epsilon^*_\phi\cdot\hat{p}_\phi/\sqrt{2}\,, \tag{1.97}$$

where E_ϕ is the energy of the ϕ in the J/ψ rest frame, and $\hat{p}_\phi$ is the unit vector along the direction of motion of ϕ in the J/ψ rest frame, . The polarization three-vectors $\epsilon^*_{J/\psi}$ and ϵ^*_ϕ are in the J/ψ rest frame. The rotationally invariant quantities are linear in $\epsilon^*_{J/\psi}$ and ϵ^*_ϕ, involving possible powers of $\hat{p}_\phi$. The two CP-even decay amplitudes are the terms with A_0 and $A_\parallel$, whereas the term with $A_\perp$ corresponds to the CP-odd component. The superscripts T refer to the projections perpendicular to $\hat{p}_\phi$, and the superscripts L refer to the longitudinal projected states $\epsilon^{*L} \equiv \hat{p}_\phi \cdot \epsilon^*$, such that $\epsilon^{*L}_{J/\psi} \cdot \hat{p}_\phi \, \epsilon^{*L}_\phi \cdot \hat{p}_\phi = \epsilon^{*L}_{J/\psi} \, \epsilon^{*L}_\phi$. This is in reference to the photon polarization states. Note that we can see from (1.97) that A_0 and $A_\parallel$ are CP even by counting powers of $\hat{p}_\phi$.

The helicity amplitudes $H_-(-1,-1)$, $H_0(0,0)$ and $H_+(+1,+1)$, normalized to the trace of the helicity density matrix, are related to the polarized amplitudes through:

$$A_0 = H_0\,, \quad A_\parallel = \frac{1}{\sqrt{2}}(H_+ + H_-)\,, \quad A_\perp = \frac{1}{\sqrt{2}}(H_+ - H_-)\,,$$

With no CP violation in the decay amplitudes we have, at $t = 0$:

$$\overline{A}_0 = A_0\,, \qquad \overline{A}_\parallel = A_\parallel\,, \qquad \overline{A}_\perp = -A_\perp\,.$$

The final state is then an admixture of CP eigenstates with three independent polarization states normalized such that the decay rate is given by:

$$\Gamma(t) \propto |A_0(t)|^2 + |A_\parallel(t)|^2 + |A_\perp(t)|^2\,, \tag{1.98}$$

where the individual decay amplitudes are given by (1.89) (or (1.95) in case of wrong tag) with the corresponding eigenvalue for the CP even ($\eta_{f_e} = +1$) and CP odd ($\eta_{f_o} = -1$) components. We thus have for the observed decay amplitudes for a $B_s \to f$ transition, with $f_e \in \{0, \parallel\}$ and $f_o \in \{\perp\}$:

$$\begin{aligned} R[B_s(t) \to f_e] &= N_{f_e} \, |A_{f_e}|^2 \, e^{-\Gamma_s t} \left\{ \cosh \frac{\Delta\Gamma_s \, t}{2} - \cos\phi_s \, \sinh \frac{\Delta\Gamma_s \, t}{2} \right. \\ &\qquad \left. + D \, \sin\phi_s \, \sin(\Delta M_s \, t) \right\}, \end{aligned} \tag{1.99}$$

$$\begin{aligned} R[B_s(t) \to f_o] &= N_{f_o} \, |A_{f_o}|^2 \, e^{-\Gamma_s t} \left\{ \cosh \frac{\Delta\Gamma_s \, t}{2} + \cos\phi_s \, \sinh \frac{\Delta\Gamma_s \, t}{2} \right. \\ &\qquad \left. - D \, \sin\phi_s \, \sin(\Delta M_s \, t) \right\}, \end{aligned} \tag{1.100}$$

from which we get the analytical rates by setting $\omega_{tag} = 0$ in $D = (1 - 2\omega_{tag})$:

$$\begin{aligned} R[B_s(t) \to f_e]|_{\omega_{tag}=0} &\equiv \Gamma[B_s(t) \to f_e]| \equiv N_{f_e} \, |A_{f_e}(t)|^2\,, \quad \text{CP even: } f_e \in \{0, \parallel\}\,, \\ R[B_s(t) \to f_o]|_{\omega_{tag}=0} &\equiv \Gamma[B_s(t) \to f_o]| \equiv N_{f_o} \, |A_{f_o}(t)|^2\,, \quad \text{CP odd: } f_o \in \{\perp\}\,. \end{aligned}$$

In the same fashion we get the observed decay amplitudes for a $\overline{\mathrm{B}}_\mathrm{s} \to f$ transition:

$$R[\overline{\mathrm{B}}_\mathrm{s}(t) \to f_\mathrm{e}] = N_{f_\mathrm{e}} |A_{f_\mathrm{e}}|^2 \, e^{-\Gamma_\mathrm{s} t} \left\{ \cosh \frac{\Delta\Gamma_\mathrm{s}\, t}{2} - \cos\phi_\mathrm{s} \, \sinh \frac{\Delta\Gamma_\mathrm{s}\, t}{2} - D \, \sin\phi_\mathrm{s} \, \sin\left(\Delta M_\mathrm{s}\, t\right) \right\}, \tag{1.101}$$

$$R[\overline{\mathrm{B}}_\mathrm{s}(t) \to f_\mathrm{o}] = N_{f_\mathrm{o}} |A_{f_\mathrm{o}}|^2 \, e^{-\Gamma_\mathrm{s} t} \left\{ \cosh \frac{\Delta\Gamma_\mathrm{s}\, t}{2} + \cos\phi_\mathrm{s} \, \sinh \frac{\Delta\Gamma_\mathrm{s}\, t}{2} + D \, \sin\phi_\mathrm{s} \, \sin\left(\Delta M_\mathrm{s}\, t\right) \right\}, \tag{1.102}$$

from which we get the analytical rates by setting $\omega_\mathrm{tag} = 0$ in $D = (1 - 2\omega_\mathrm{tag})$:

$$\begin{aligned} R[\overline{\mathrm{B}}_\mathrm{s}(t) \to f_\mathrm{e}]|_{\omega_\mathrm{tag}=0} &\equiv \Gamma[\overline{\mathrm{B}}_\mathrm{s}(t) \to f_\mathrm{e}]| \equiv N_{f_\mathrm{e}} |\overline{A}_{f_\mathrm{e}}(t)|^2 \,, \quad \text{CP even: } f_\mathrm{e} \in \{0, \|\} \,, \\ R[\overline{\mathrm{B}}_\mathrm{s}(t) \to f_\mathrm{o}]|_{\omega_\mathrm{tag}=0} &\equiv \Gamma[\overline{\mathrm{B}}_\mathrm{s}(t) \to f_\mathrm{o}]| \equiv N_{f_\mathrm{o}} |\overline{A}_{f_\mathrm{o}}(t)|^2 \,, \quad \text{CP odd: } f_\mathrm{o} \in \{\perp\} \,. \end{aligned}$$

Since the lifetimes are really close, the inclusion of an angular analysis will improve the accuracy on $\Delta\Gamma_\mathrm{s}$, and furthermore the determination of ϕ_s. In particular, for $\Gamma_\mathrm{H} = \Gamma_\mathrm{L}$ it is impossible to disentangle the CP eigenstates without the angular information. The advantage of (1.98), and as we shall see of the so-called one-angle transversity distribution, is that we have to deal with amplitudes, i.e. moduli squared. Even though a full angular analysis will introduce interference terms (and strong phases), the access to the magnitudes of the amplitudes could improve the measurements.

The transversity basis

The transversity basis eases the angular analysis, using the direction of the momenta of the four final states in $\mathrm{B_s} \to \mathrm{J}/\psi(\to \ell^+\ell^-)\phi(\to \mathrm{K}^+\mathrm{K}^-)$ to define three independent physical angles. The transversity basis corresponds to a circular permutation of the helicity axis, using the correlation between the spatial components of the J/ψ and the ϕ polarizations, since the $\mathrm{B_s}$ has zero spin. The J/ψ has thus a single linear polarization vector state ϵ^*, such that in the J/ψ rest frame we have [40]:

$$A_0 : \quad \epsilon^* = \hat{x} \,, \quad A_\| : \quad \epsilon^* = \hat{y} \,, \quad A_\perp : \quad \epsilon^* = \hat{z} \,. \tag{1.103}$$

We can then define the three angles used to describe the angular distribution following [40, 41], and as displayed on Figure 1.14:

- The x axis is the direction of the ϕ vector in the J/ψ rest frame, and the direction of the ϕ polarization lies in the $x - y$ plane.
- The z axis is perpendicular to the $\phi \to \mathrm{K}^+\mathrm{K}^-$ decay plane, in the J/ψ rest frame.
- The y axis lies in the plane spanned by the two kaons, with $p_y(\mathrm{K}^+) \geq 0$, in the rest frame of the J/ψ. This defines the right-handedness of the Cartesian system. The direction of the z axis is thus given by $\hat{p}_{\mathrm{K}^-} \times \hat{p}_{\mathrm{K}^+}$ (in this order), where $\hat{p}_{\mathrm{K}^\pm}$ are the unit momentum vectors of the two kaons in the J/ψ rest frame.

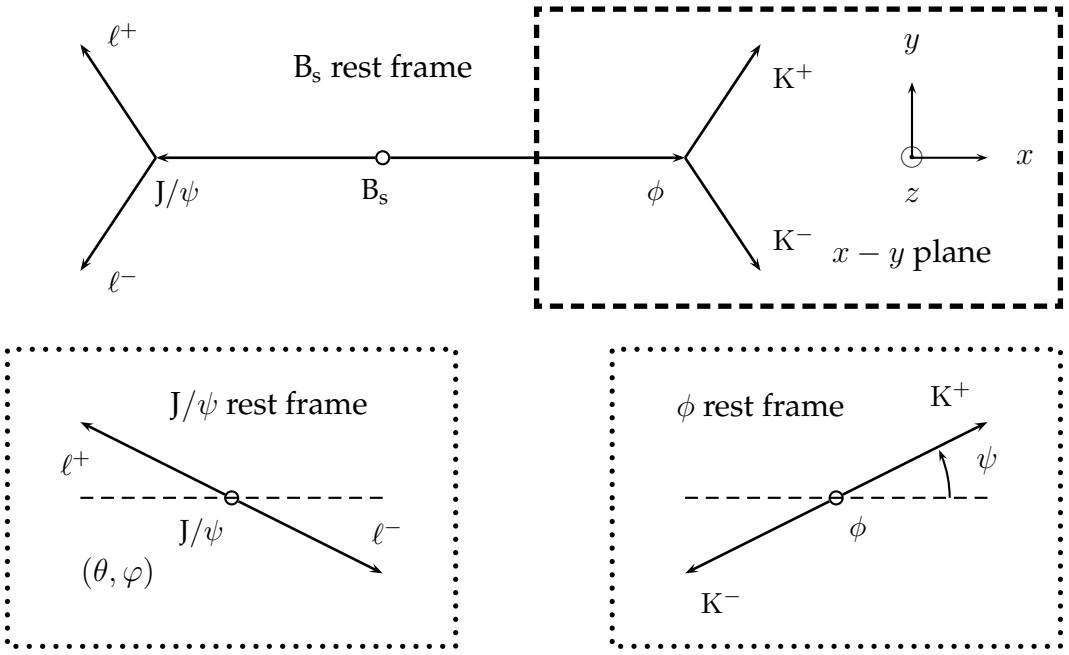


Figure 1.14: The angles (θ, φ, ψ) of the transversity basis.

- The angle ψ is the angle between the momentum $\hat{p}'_{K^+}$ in the ϕ rest frame and the helicity axis x. Note that the x axis direction is opposite to the direction of the J/ψ (denoted by $\hat{p}'_{J/\psi}$) in the ϕ rest frame.

- The other two transversity angles (θ, φ) describe the direction of the positive lepton ℓ^+ from the J/ψ in the J/ψ rest frame.

A unit vector $\hat{n}$ in the direction of the ℓ^+ may be defined:

$$\hat{n} = (n_x, n_y, n_z) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)\,, \tag{1.104}$$

such that with the above conventions we have [40, 41]:

$$\hat{x} = \hat{p}_\phi, \qquad \hat{y} = \frac{\hat{p}_{K^+} - \hat{p}_\phi(\hat{p}_\phi \cdot \hat{p}_{K^+})}{|\hat{p}_{K^+} - \hat{p}_\phi(\hat{p}_\phi \cdot \hat{p}_{K^+})|}, \qquad \hat{z} = \hat{x} \times \hat{y},$$
$$\sin\theta\cos\varphi = \hat{p}_{\ell^+} \cdot \hat{x}, \qquad \sin\theta\sin\varphi = \hat{p}_{\ell^+} \cdot \hat{y}, \qquad \cos\theta = \hat{p}_{\ell^+} \cdot \hat{z}\,. \tag{1.105}$$

The vectors are all *unit* three-vectors and everything is measured in the rest frame of the J/ψ. Also:

$$\cos\psi = -\hat{p}'_{K^+} \cdot \hat{p}'_{J/\psi}, \tag{1.106}$$

where the primed quantities are *unit* vectors measured in the rest frame of the ϕ.

Consider two particles a and b in the laboratory frame described by their four-vectors $p_a = (E_a, \vec{p}_a)$ and $p_b = (E_b, \vec{p}_b)$, respectively. If we consider a Lorentz boost[7] to the rest

[7]We use $m^2 = E^2 - \vec{p}^{\,2}$ and we set the speed of light $c = 1$.

frame of b, with $\gamma_b = 1/\sqrt{1-\beta_b^2} = E_b/m_b$ and $\vec{\beta}_b = \vec{p}_b/E_b$, then a will be described in the rest frame of b by the four-vector $p'_a = (E'_a, \vec{p}_a{}')$ satisfying the following relations:

$$E'_a = \frac{1}{m_b}\left(E_b\, E_a - \vec{p}_b \cdot \vec{p}_a\right) , \tag{1.107}$$

$$\vec{p}_a{}' = \frac{1}{m_b}\left(-E_a\, \vec{p}_b + E_b\, \vec{p}_a\right) = \vec{p}_a - \left(\frac{E'_a + E_a}{E_b + m_b}\right)\vec{p}_b\,, \tag{1.108}$$

where m_b is the mass of the particle b. The above relations are useful when computing the different quantities in the transversity basis.

The three-angle transversity distribution

The three-angle distribution of an initially produced $\mathrm{B_s}$ is obtained from (1.97) in terms of the transversity angles defined in (1.105) and (1.106). The result is [40, 41]:

$$\begin{aligned}\frac{\mathrm{d}^3\Gamma[\mathrm{B_s}(t)\to \mathrm{J}/\psi(\to \ell^+\ell^-)\phi(\to \mathrm{K}^+\mathrm{K}^-)]}{\mathrm{d}\cos\theta\,\mathrm{d}\varphi\,\mathrm{d}\cos\psi} &\propto \frac{9}{32\pi}\Big\{2|A_0(t)|^2\cos^2\psi(1-\sin^2\theta\cos^2\varphi)\\ &+\sin^2\psi\left[|A_\parallel(t)|^2(1-\sin^2\theta\sin^2\varphi)+|A_\perp(t)|^2\sin^2\theta-\Im\left(A_\parallel^*(t)A_\perp(t)\right)\sin 2\theta\sin\varphi\right]\\ &+\frac{1}{\sqrt{2}}\sin 2\psi\left[\Re\left(A_0^*(t)A_\parallel(t)\right)\sin^2\theta\sin 2\varphi+\Im\left(A_0^*(t)A_\perp(t)\right)\sin 2\theta\cos\varphi\right]\Big\}\,. \end{aligned}\tag{1.109}$$

The angular distribution for the CP conjugate decay is obtained by simply replacing A's with $\bar{A}$'s [41].

Using all the information from the three angles (θ, φ, ψ) will improve the lifetime difference measurement, and the ϕ_s determination, in comparison to using only one angle. For untagged events, and besides the $\cos\phi_\mathrm{s}$ term , we have access to ϕ_s through the interference terms $\Im(A_\parallel^*(t)A_\perp(t))$ and $\Im(A_0^*(t)A_\perp(t))$.

The interference terms can be found in [22], with the explicit formulae in the case of a non-negligible $\cos\phi_\mathrm{s}$ contribution. Let us introduce the strong phases δ_1 and δ_2, defined as follows [40, 41]:

$$\delta_1 \equiv \arg\Big\{A_\parallel^* A_\perp\Big\}, \quad \delta_2 \equiv \arg\Big\{A_0^* A_\perp\Big\}, \tag{1.110}$$

The non-vanishing parts containing ϕ_s in the case of untagged mesons in the interference terms are such that:

$$\begin{aligned}\Re\{A_0^*(t)A_\parallel(t)\} &\propto \cos(\delta_2-\delta_1)\,\cos\phi_\mathrm{s}\,,\\ \Im(A_\parallel^*(t)A_\perp(t)) &\propto \cos\delta_1\,\sin\phi_\mathrm{s}\,,\\ \Im(A_0^*(t)A_\perp(t)) &\propto \cos\delta_2\,\sin\phi_\mathrm{s}\,,\end{aligned}$$

with similar expressions for the interference terms with $\bar{A}$'s. We can therefore expect the largest sensitivity from the imaginary interference terms, in case of a SM value of ϕ_s. In the case of the three-angle distribution, we expect different sensitivities to ϕ_s depending on its sign, due to the imaginary interference terms. Note that the likelihood fit to the full three-angle distribution is a complicated task, in which one may need to separate the likelihood function in smaller independent parts, as suggested in [41], using the method of angular moments.

The one-angle transversity distribution

Even though we will lose sensitivity, the analysis is considerably simplified in the case where we integrate over (φ, ψ), assuming we have a flat selection acceptance over these angles. In this case we obtain the angular distribution in terms of one angle, the so-called transversity angle θ [42], separating the CP-even and CP-odd components. The angular differential distribution for $\mathrm{B_s} \to f$ in terms of θ is given by [40]:

$$\frac{\mathrm{d}\,\Gamma[\mathrm{B_s(t)} \to \mathrm{f}]}{\mathrm{d}\cos\theta} \propto (|A_0(t)|^2 + |A_\parallel(t)|^2)\,\frac{3}{8}\,(1+\cos^2\theta) + |A_\perp(t)|^2\,\frac{3}{4}\sin^2\theta\,, \tag{1.111}$$

where the time evolution of the terms are defined through (1.99) and (1.100) by setting $\omega = 0$. From this expression we have that the CP components are separated by their lifetime, and by their distinct angular distributions. In the latter case we will have a good access to ϕ_s, with a smaller sensitivity for untagged decays and a SM mixing phase.

The fraction of CP-odd eigenstates, R_T, is defined as

$$R_\mathrm{T} \equiv \frac{|A_\perp|^2}{\sum_{f=0,\parallel,\perp}|A_f|^2}\,, \tag{1.112}$$

where the A_f are taken at $t = 0$. Note that the value of R_T has been measured [43, 44], with an averaged value of $R_\mathrm{T} \sim 0.16$, as shown in Section 1.5. Interestingly, we get the CP-even limit when $R_\mathrm{T} \to 0$, and the maximal dilution of the CP asymmetry when $R_\mathrm{T} = 0.5$. Indeed, assuming that both the CP-even and CP-odd components *have the same angular distributions* (which is absolutely not true), then the terms involving ϕ_s vanish when $R_\mathrm{T} = 0.5$. This can be seen from:

$$\begin{aligned}
\Gamma[\mathrm{B_s}(t) \to f] &\propto (1-R_\mathrm{T})\,\Gamma[\mathrm{B_s}(t) \to f_\mathrm{e}] + R_\mathrm{T}\,\Gamma[\mathrm{B_s}(t) \to f_\mathrm{o}]\,, \quad \text{same } f(\theta)\,, \\
&\propto e^{-\Gamma_\mathrm{s} t}\left[\cosh\frac{\Delta\Gamma_\mathrm{s}\,t}{2} - D_\mathrm{T}\,\cos\phi_\mathrm{s}\,\sinh\frac{\Delta\Gamma_\mathrm{s}\,t}{2} + D_\mathrm{T}\,\sin\phi_\mathrm{s}\,\sin\left(\Delta M_\mathrm{s}\,t\right)\right]\,,
\end{aligned}$$

where $D_\mathrm{T} \equiv (1-2R_\mathrm{T})$ acts as a dilution factor.

The one-angle angular distribution (1.111) will be used in Chapter 6 to assess the sensitivity of LHCb to ϕ_s with $\mathrm{B_s} \to \mathrm{J}/\psi\phi$ decays.

1.4.4 Control Channel

The determination of ϕ_s requires the knowledge of ΔM_s, as the better sensitivity will be obtained from the $\sin(\Delta M_\mathrm{s}\,t)$ term. For $\bar{\mathrm{b}} \to \bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}$ quark-level transitions, ΔM_s is accessible. If the oscillation wiggles are not significant enough, we would need an external ΔM_s. There is a large correlation between the mistag and ϕ_s. We therefore need a control channel to extract the tagging dilution and ΔM_s by simultaneously fitting all the samples. To this end, the flavor-specific $\mathrm{B_s} \to \mathrm{D_s}\pi$ decay channel can be used. The relevant decay rates were given in (1.63) and (1.64). Accounting for a possible mistag, we get the following observed decay rates:

$$R[\mathrm{B_s}(t) \to f] = N_f\,|A_f^{(\mathrm{s})}|^2\,\frac{e^{-\Gamma_\mathrm{s} t}}{2}\left[\cosh\frac{\Delta\Gamma_\mathrm{s}\,t}{2} + D\,\cos(\Delta M_\mathrm{s}\,t)\right]\,, \tag{1.113}$$

$$R[\overline{\mathrm{B}}_\mathrm{q}(t) \to f] = N_f\,|A_f^{(\mathrm{s})}|^2\,\frac{e^{-\Gamma_\mathrm{s} t}}{2}\left[\cosh\frac{\Delta\Gamma_\mathrm{s}\,t}{2} - D\,\cos(\Delta M_\mathrm{s}\,t)\right]\,. \tag{1.114}$$

1.5 Experimental Tests of the CKM Picture

We summarize in this section the experimental status and results of CKM fits as of April 2006. The different parameters of the B mesons are listed in Table 1.1. Note that the theoretical expectations for $\Delta\Gamma_s/\Gamma_s$ are $10-12\%$ [45, 46].

Table 1.1: Neutral B-mesons parameters. The values are taken from [27, 47]. The ΔM_s value corresponds to CDF's measurement [14]. For $1/\Gamma_d$, we assumed $\Delta\Gamma_d = 0$.

B_q	Mass [MeV/c]	Lifetime [ps]	ΔM_q [ps^{-1}]	$\Delta\Gamma_q/\Gamma_q$	$1/\Gamma_q$ [ps]
B_d	5279.4 ± 0.5	1.527 ± 0.008	0.508 ± 0.004	0.009 ± 0.037	1.527 ± 0.008
B_s	5369.6 ± 2.4	1.461 ± 0.040	$17.33^{+0.42}_{-0.21} \pm 0.07$	$0.31^{+0.10}_{-0.11}$	$1.396^{+0.044}_{-0.046}$

The results of the Unitarity Triangle fit (UTfit) [48, 49] using a Bayesian method to constrain the allowed region in the $(\bar{\rho}, \bar{\eta})$ plane are shown in the left plot of Figure 1.15. The fit gives $\bar{\rho} = 0.193 \pm 0.029$ and $\bar{\eta} = 0.355 \pm 0.019$. On the same figure the result of the full CKM fit for the angle $\chi \equiv \beta_s$ is drawn, where the fit yielded $\sin 2\chi = 0.038 \pm 0.002$. All the UTfit results include CDF's measurement of ΔM_s.

The results of the generalization of the UTfit beyond the Standard Model to the Universal Unitary Triangle [50], the UUTfit, can be found in [51]. The results of the fit for the $(\bar{\rho}, \bar{\eta})$ plane are shown in Figure 1.16, together with the shift induced in the B_s–$\overline{B}_s$ mixing phase with a NP parameter ϕ_{B_s}. For the latter, the UUTfit takes into account the measurements of $\Delta\Gamma_q/\Gamma_q$, and the semi-leptonic and di-leptons asymmetries.

The fraction of the CP-odd component R_T in $B_s \to J/\psi\phi$ has been measured [43, 44], assuming the CP-violating phase is negligible. The results are respectively:

1. Results from CDF [43]:
   - $R_T = 0.125 \pm 0.069$.
2. Results from DØ [44]:
   - $R_T = 0.16 \pm 0.10\ (\text{stat.}) \pm 0.02\ (\text{syst.}) = 0.16 \pm 0.10$. Note that the largest systematic uncertainty comes from the integration over two of the angles of the transversity basis φ and ψ. The acceptance versus θ has a four times smaller systematic uncertainty compared to the $\varphi - \psi$ integration.

The measurement of ΔM_s has constrained the magnitude of the B_s–$\overline{B}_s$ oscillations to a SM model value, still the modulus of $M_{12}^{(s)}$ is *not* very well known [15]. The next crucial step is to determine its phase. As ϕ_s is expected to be small in the SM, we should be able to probe any NP contribution, such that $\phi_s = \phi_s^{SM} + \phi_s^{NP}$. This has become one of LHCb's primary goals. The expected performance on the determination of ϕ_s will be assessed in Chapter 6.

1.5.1 En Route Towards New Physics with B_s Decays?

The presence of New Physics may affect the low-energy effective Hamiltonian governing the B_s–$\overline{B}_s$ mixing in different ways:

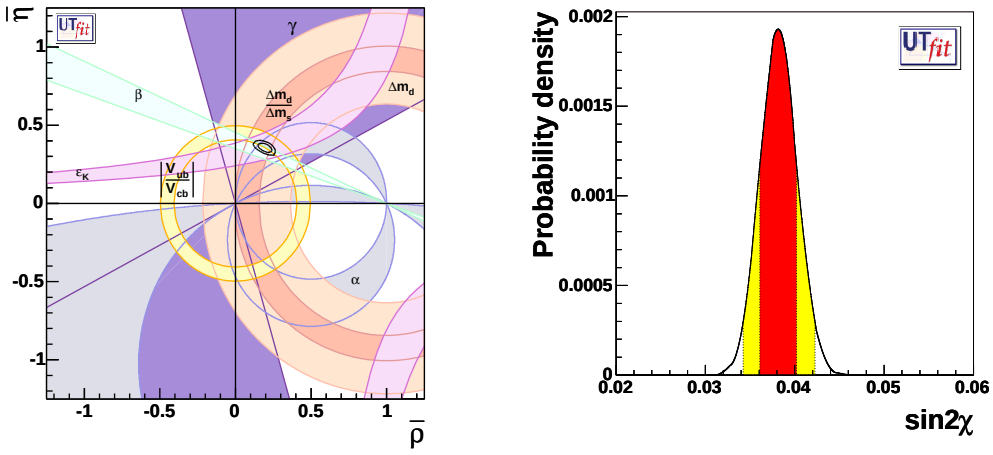


Figure 1.15: Left: allowed regions in the $(\bar{\rho}, \bar{\eta})$ plane as obtained by the UTfit [49]. The closed contours at 68% and 95% probability are shown. The colored regions correspond to 95% probability regions from each constraint. Right: result of the full CKM fit for $\sin 2\chi = 0.038 \pm 0.002$.

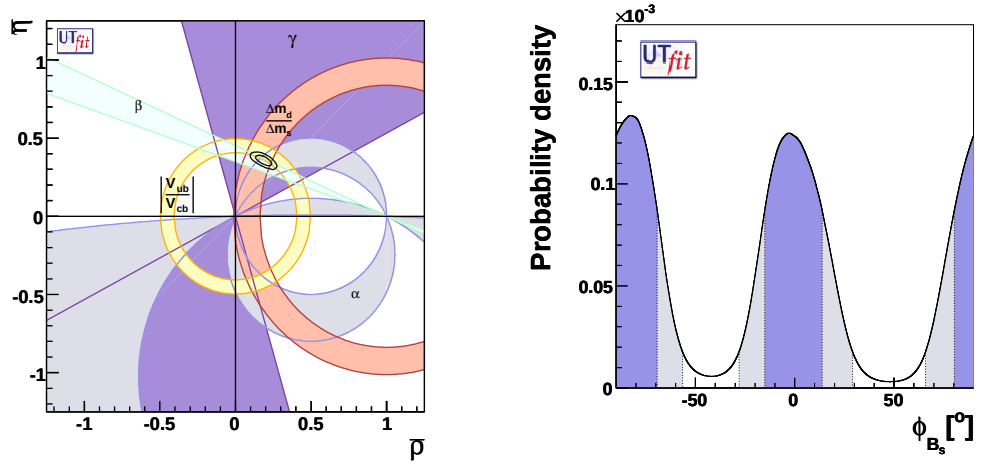


Figure 1.16: Left: determination of $\bar{\rho}$ and $\bar{\eta}$ from the UUTFit [51]. Right: constraints on $\phi_{\mathrm{B_s}}$ coming from the generalized NP UUTFit analysis.

- new contributions to the Wilson coefficients;
- new local operators;
- new CP-violating phases.

The existence of NP may originate at the tree level from new interactions, or at the loop level through virtual exchange of new particles. All these effects may have a role in the B_s–$\overline{B}_s$ mixing, in particular for observables which are suppressed in the Standard model, such as the B_s–$\overline{B}_s$ mixing phase.

The measurement of ΔM_s compatible with the SM expectation has called forth the measurement of the B_s–$\overline{B}_s$ mixing phase. The B_s–$\overline{B}_s$ mixing is due in the SM to the box diagrams with W boson and up-type quark exchanges, inducing an effective Hamiltonian responsible for the B_s–$\overline{B}_s$ oscillations. As described in Section 1.3, the B_s–$\overline{B}_s$ mixing is characterized by the frequency, $\Delta M_s = M_H - M_L = 2|M_{12}^{(s)}|^2$ and the CP-violating phase ϕ_s which enters the mixing-induced CP asymmetry, as shown for $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions. In the SM, we have:

$$\begin{aligned} M_{12}^{(s)SM} &\propto (V_{ts}^* V_{tb})^2 \,, \\ \phi_s^{SM} &= 2\arg[V_{ts}^* V_{tb}] \approx -2\beta_s \end{aligned}$$

We may parameterize, in a model independent way, the matrix element $M_{12}^{(s)}$ by taking into account a possible NP contribution [15, 52]:

$$M_{12}^{(s)} = M_{12}^{(s)SM}(1 + h_s e^{i\sigma_s}) \,,$$

where $h_s \geq 0$ measures the relative strength of the NP contribution with respect to the SM, whereas σ_s is a new CP-violating phase. We then have for the B_s mixing parameters:

$$\begin{aligned} \Delta M_s &= \Delta M_s^{SM} \left[1 + h_s e^{i\sigma_s}\right] \,, \\ \phi_s &= \phi_s^{SM} + \phi_s^{NP} = \phi_s^{SM} + \arg\left[1 + h_s e^{i\sigma_s}\right] \,. \end{aligned}$$

For the B_s system, the SM CP-violating phase $\phi_s^{SM} \approx -2\beta_s$ is small, and appears in the squashed (sb) triangle. Thus it is practically independent of the constraints in the $(\bar{\rho}, \bar{\eta})$ plane. We can use the known constraints from the B_d system to in turn constrain the NP contributions to the B_s system. A theoretical clean way of doing this is through the ratio of the oscillation frequencies:

$$\frac{\Delta M_s}{\Delta M_d} \propto \frac{M_{B_s}}{M_{B_d}} \xi^2 \frac{|V_{ts}|^2}{|V_{td}|^2} \frac{|1 + h_s e^{i\sigma_s}|}{|1 + h_d e^{i\sigma_d}|} \,,$$

where h_d and σ_d are the equivalent relative strength and CP-violating phase of NP for the B_d system. These parameters are strongly correlated to the $(\bar{\rho}, \bar{\eta})$ plane. On the other hand, the (sb) triangle is highly degenerate and therefore the determination of h_s and σ_s is almost independent from $\bar{\rho}, \bar{\eta}$. Even though the effects of NP could cancel in this ratio, it enables to exclude regions in the h_s and σ_s parameter space. In particular, a measurement of CP violation in $B_s \to J/\psi\phi$ decays will be a very sensitive probe to NP.

A scan of the allowed range for (h_s, σ_s), with CDF's measurement of ΔM_s [14] can be found in [52], and the results are shown in Figure 1.17. The scan assumes a statistical precision on $\sigma(\sin 2\beta_s) \sim 0.03$ at LHCb with $2\,\text{fb}^{-1}$, and for a SM value of $\sin 2\beta_s \sim 0.038$. We see that already with one year of data the bound on h_s will be better than ~ 0.1.

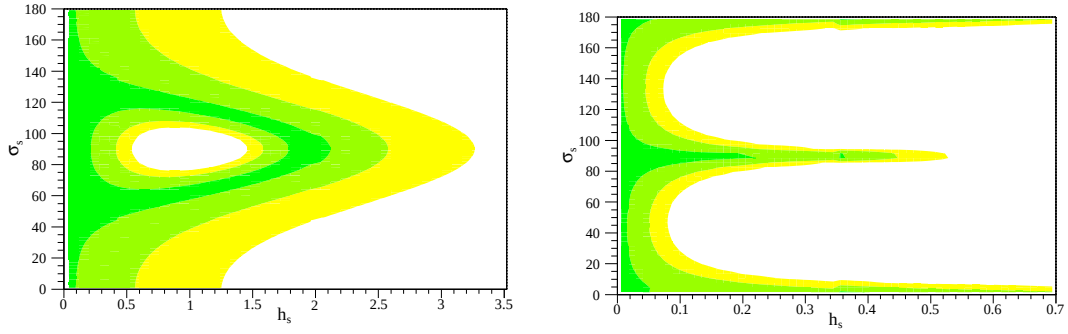


Figure 1.17: The allowed range for h_s and σ_s assuming the SM prediction as central value [52], without LHCb (left) and using one year LHCb sensitivity (right). The dark, medium, and light shaded areas have confidence levels larger than $0.90, 0.32$ and 0.05, respectively.

Finally, it is shown in [15] that even with the measurement of ΔM_s, there is ample space for the NP parameters (h_s, σ_s), in a model independent way. In particular, the new CP-violating phase σ_s is essentially unconstrained.

We can thus conclude that New Physics may be hiding in the B_s–$\overline{B}_s$ mixing, obscured by parameter uncertainties, and the smoking gun for the presence of New Physics will be the measurement of a sizable ϕ_s^{NP}.

Chapter 2

The LHCb Detector

We introduce the Large Hadron Collider and the Large Hadron Collider beauty experiment. The distinctive features of the latter are presented, as well as an overview of the experimental setup, including the trigger system.

THE European Organization for Nuclear Research (CERN) sitting astride the border between Switzerland and France, near Geneva, is the world's largest particle physics laboratory [53]. It was founded in 1954 and was created primarily to provide physicists with the facilities (computing centers, accelerators) needed to investigate elementary high-energy particles. After the numerous successes and discoveries achieved at the Large Electron Positron collider (LEP), physicists were eager to pursue testing the limits of the Standard Model and beyond. The Large Hadron Collider (LHC) is currently being built for these purposes. This huge accelerator will operate at a ground-breaking energy scale, thus also opening a gate to New Physics. The LHC will host several large experiments, the Large Hadron Collider beauty (LHCb) being one of them.

2.1 The Large Hadron Collider

The primary goals of the LHC are to test the Standard Model (SM) description of particle physics and to look for New Physics. The LHC is a proton-proton (pp) collider at a center-of-mass energy of $\sqrt{s} = 14$ TeV. The yet unobserved and only missing particle in the SM, namely the Higgs boson, is expected to be detected at the LHC. At these very high energies, new particles beyond the SM's predictions are also likely to be produced. For instance, particles arising from supersymmetry (SUSY) are hoped to be seen. The LHC will also enable high-precision b-physics measurements. Finally, the LHC will provide in dedicated runs heavy ions collisions (e.g. Pb-Pb, Ca-Ca) instead of pp collisions to study the behavior of nuclear matter in extreme conditions, and the formation of quark-gluon plasma.

The choice of having a circular pp collider instead of an electron-positron ($\mathrm{e^+e^-}$) machine is motivated by the high-energy scale, $\mathcal{O}(\mathrm{TeV})$, required to study new phenomena. These energies cannot be reached with circular $\mathrm{e^+e^-}$ machines as we are limited by the synchrotron radiation loss. The LHC will operate at the highest energy ever achieved, e.g. much larger than the $\mathrm{p\overline{p}}$ Tevatron collider at Fermilab (USA) which has $\sqrt{s} = 2$ TeV.

The LHC is being installed underground in the old LEP tunnel and it is expected to start its first pilot runs after the Summer 2007. The ring's circumference is about 27 km in which the protons are accelerated in opposite directions to an energy of 7 TeV. The Super Proton Synchrotron (SPS) is used as injector for the LHC, process which is preceded by several acceleration phases. The two proton beams enter the LHC at the TI2 and TI8 injection points, as shown in Figure 2.1.

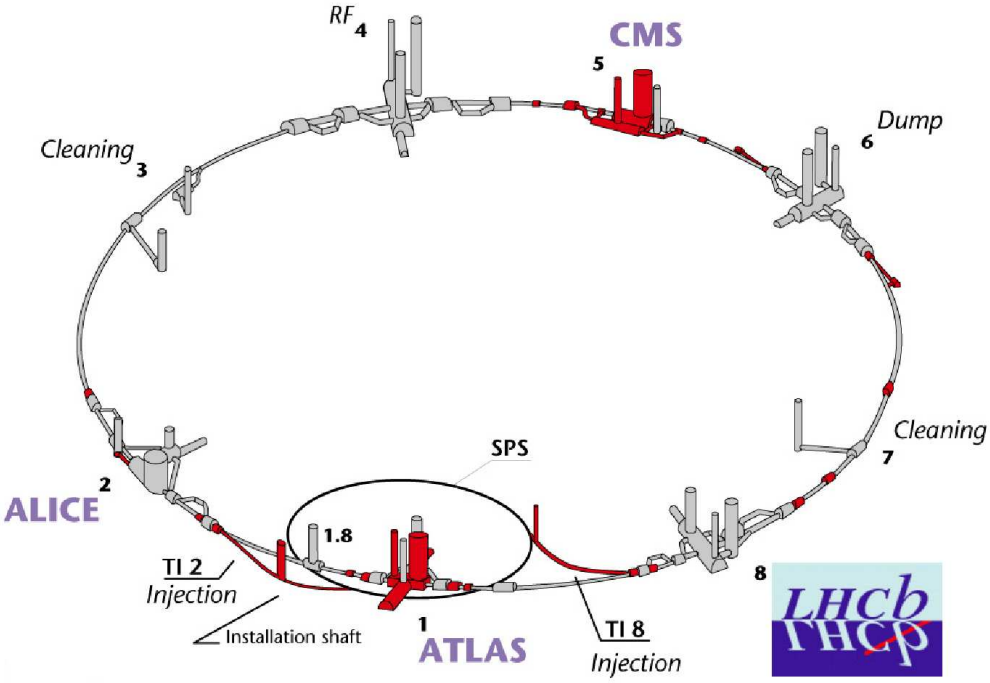


Figure 2.1: The LHC complex with the main experiments [54].

Four main experiments will be housed at the LHC, as shown in Figure 2.1:

- LHCb is located at IP8 and is dedicated to the study of CP violation and b physics.
- ALICE is situated at IP2 and will study heavy-ion collisions.
- ATLAS and CMS are central detectors (i.e. with full angular coverage) and are installed opposite to each other, at IP1 and IP5 respectively, both having similar general physics programs. In particular, they will compete for the search of Higgs and SUSY particles. They also have some b physics in their program, though being limited to the early stage of LHC running.

2.1.1 LHC Parameters

At the LHC energy, protons of 7 TeV need a strong magnetic field to be maintained in their orbit. This is accomplished by means of superconducting magnets using a Nb-Ti conductor placed into a cryostat containing superfluid helium and cooled down to 1.9 K. Given the Lorentz force the protons undergo, the displacement between magnets and the

radius of the LHC ring, a bending field of 8.34 T is chosen. The two protons beams will then travel in opposite directions in separated and different beam pipes.

The number of proton bunches per LHC beam is $N_{\text{bx}} \sim 2800$, each bunch consisting of $\sim 10^{11}$ protons. The bunch crossing frequency is 40 MHz, corresponding to a time of 25 ns (or ~ 7.5 m) spacing between two consecutive bunches. This frequency sets the basic clock cycle for the detector Front-End (FE) electronics. Due to the bunch structure in particular having empty bunches as a result of the filling procedure, the average crossing frequency ν_{bx} is somehow smaller, of the order of ~ 30 MHz at IP8 [17].

The number of pp collisions at each interaction point depends on two quantities:

- The total pp cross-section $\sigma_{\text{pp}}^{\text{tot}} = \sigma_{\text{pp}}^{\text{inel}} + \sigma_{\text{pp}}^{\text{el}}$. The part due to elastic scattering leaves the protons intact which in general are not observed in the detector. The inelastic part gives rise to a large multiplicity of interactions passing through the detector setup. At $\sqrt{s} = 14$ TeV the expected inelastic cross-section is $\sigma_{\text{pp}}^{\text{inel}} = 80$ mb [55].

- The luminosity $\mathcal{L}$. It determines the number of pp interactions at a crossing point. The luminosity depends on the compactness of the beam (emittance), the ability of the magnets to focus the beam at the interaction point (betatron function), the crossing angle and the bunch crossing frequency. The design luminosity of LHC (and the number of protons per bunch) yields $\mathcal{L} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$. This instantaneous luminosity will however not be reached before a few years, and the LHC will start with a low-luminosity phase at $\mathcal{L} = 10^{33}$ $\text{cm}^{-2}\text{s}^{-1}$.

The number of inelastic pp collisions (N_{pp}) over a given time interval is given by:

$$N_{\text{pp}} = \sigma_{\text{pp}}^{\text{inel}} \int \mathcal{L} \text{d}t \,. \tag{2.1}$$

The (Poisson) mean number of interactions per bunch crossing is then given by:

$$\langle \text{n}_{\text{pp}} \rangle = \frac{n_{\text{pp}}}{N_{\text{bx}}} = \frac{\sigma_{\text{pp}}^{\text{inel}} \mathcal{L}}{\nu_{\text{bx}}} \,. \tag{2.2}$$

At a luminosity of 10^{34} $\text{cm}^{-2}\text{s}^{-1}$ this "pileup effect" corresponds to $\langle \text{n}_{\text{pp}} \rangle \sim 25$ interactions.

2.1.2 LHCb Luminosity

LHCb needs to identify the b-decay vertices and the corresponding primary interaction they origin from. Obviously pileup is a limiting factor in the correct identification of vertices. For this reason LHCb will run at a lower luminosity by adjusting the beam focusing at its interaction.

The choice of luminosity for LHCb relies on the number of visible (inelastic) interactions we wish to have. The number of pp interactions n_{pp} in a given bunch crossing follows a Poisson distribution of mean $\langle \text{n}_{\text{pp}} \rangle$ as defined in (2.2), where the probability is given by:

$$P\left(\langle \text{n}_{\text{pp}} \rangle, n_{\text{pp}}\right) = \frac{\langle \text{n}_{\text{pp}} \rangle^{n_{\text{pp}}}}{n_{\text{pp}}!} e^{-\langle \text{n}_{\text{pp}} \rangle} \,. \tag{2.3}$$

The probability for observing $n = 0, 1, 2, 3$ or 4 inelastic pp collisions as a function of the luminosity is shown in Figure 2.2.

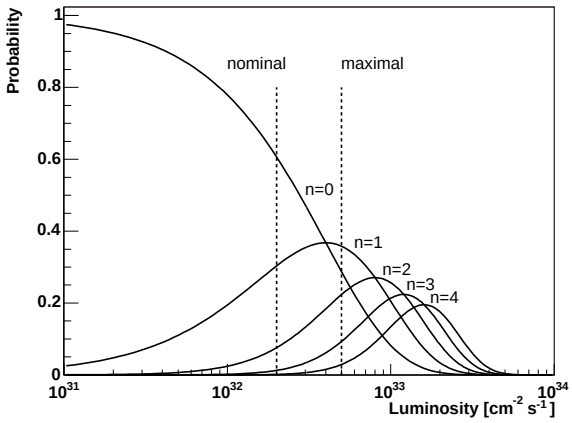


Figure 2.2: Probability of having $n = 0, 1, 2, 3$ or 4 inelastic interactions as a function of the luminosity. The vertical dashed lines represent LHCb's nominal and maximal luminosities.

For LHCb purposes, the optimum choice of luminosity is for single interaction events ($n = 1$). The chosen nominal luminosity is 2×10^{32} cm^{-2}s^{-1} and the maximal designed luminosity of LHCb is 5×10^{32} cm^{-2}s^{-1}. This leaves room for higher luminosity upgrades. Note that the probability of having a single interaction reaches a maximum at 4×10^{32} cm^{-2}s^{-1}. Thus the nominal luminosity is an educated choice, since the number of multiple interactions rapidly increases with the luminosity.

2.1.3 Production Cross-Sections

The physics processes involved in pp collisions, and in particular the $\mathrm{b\overline{b}}$ production mechanisms will be discussed in Chapter 3. We give in Table 2.1, the assumed production cross-sections and those calculated with LHCb's tuning of the PYTHIA event generator [56]. The PYTHIA results are from [57]. The $\mathrm{b\overline{b}}$ cross-section is badly known and may be between 175 and 950 μm [58]. We will assume the $\mathrm{b\overline{b}}$ cross-section at $\sqrt{s} = 14\,\mathrm{TeV}$ in pp collisions to be $\sigma_{\mathrm{b\bar{b}}} = 500\ \mu\mathrm{b}$ in the calculation of annual yields and background levels. This corresponds to an annual (i.e. 10^7 s) production of 10^{12} $\mathrm{b\overline{b}}$ pairs at LHCb's nominal luminosity.

As it will be explained in Chapter 3, the production of $\mathrm{b\overline{b}}$ pairs at the LHC is such that the b hadrons are very correlated. Their angular distribution is peaked at low polar angles such that b hadrons from the same $\mathrm{b\overline{b}}$ pair are likely to fly in the same backward or forward direction. This has motivated the design and angular coverage of the LHCb detector.

The annual (i.e. 10^7 s) number of expected $\mathrm{X_{\bar{b}}}$ hadrons ($\mathrm{X_{\bar{b}}}$ = $\mathrm{B_u}$, $\mathrm{B_d}$, $\mathrm{B_s}$, Λ_b) in 4π at a nominal luminosity $\mathcal{L} = 2 \times 10^{32}\mathrm{cm^{-2}s^{-1}}$ is given by:

$$N_{\mathrm{X_{\bar{b}}}} = \sigma_{\mathrm{b\bar{b}}} \times \int \mathcal{L}\,\mathrm{d}t \times 2 \times \mathrm{BR}\left(\mathrm{\bar{b}} \to \mathrm{X_{\bar{b}}}\right) ,$$

where the factor 2 accounts for the production of both b and $\mathrm{\bar{b}}$ hadrons, with a production fraction of BR $\left(\mathrm{\bar{b}} \to \mathrm{X_{\bar{b}}}\right)$. The production fractions are given in Table 2.2. Note that the number of $\mathrm{X_{\bar{b}}}$ hadrons at LHCb will be lower, given the acceptance of the detector.

Table 2.1: Assumed and PYTHIA cross-sections used throughout this document.

σ_{pp}^{tot}	100 mb
σ_{pp}^{inel}	80 mb
$\sigma_{b\bar{b}}$	$500\ \mu$b
$\sigma_{b\bar{b}}^{Py}$	$627\ \mu$b
$\sigma_{D_s}^{Py}$	1.004 mb

Table 2.2: $X_{\bar{b}}$ production fractions BR $\left(\bar{b} \to X_{\bar{b}}\right)$ [27]. Λ_b denotes any b baryon.

BR $\left(\bar{b} \to B_u\right)$ [%]	BR $\left(\bar{b} \to B_d\right)$ [%]	BR $\left(\bar{b} \to B_s\right)$ [%]	BR $\left(\bar{b} \to \Lambda_b\right)$ [%]
39.8 ± 1.2	39.8 ± 1.2	10.3 ± 1.4	10.0 ± 2.0

2.2 The Large Hadron Collider Beauty Experiment

The Large Hadron Collider beauty experiment (LHCb) is a single-arm forward spectrometer dedicated to the precise measurements of CP violation and rare decays [55]. The study of b hadrons is the main goal, though other topics will be investigated to a lesser extend, e.g. charm physics, light Higgs.

The LHCb detector, housed at the interaction point IP8 where the Delphi experiment used to be, is presently under construction. The available physical space in the cavern and the properties of the $b\overline{b}$ pairs production have driven the design of the experimental setup, shown in Figure 2.3. The right-handed coordinate system is defined such that the z axis points towards the Muon Stations and the vertical axis y points upwards. The main magnetic field component is along the y axis, thus defining the horizontal or bending plane $x - z$ and the vertical or non-bending plane $y - z$.

The angular coverage is $10 - 300$ mrad in the horizontal plane and $10 - 250$ mrad in the vertical plane, where the acceptance is given in terms of the polar angle θ with respect to the z axis. Equivalently, this acceptance corresponds to a range in pseudo-rapidity of $1.9 < \eta < 4.9$, where $\eta = -\ln\left(\tan\left(\theta/2\right)\right)$. The dimensions of the detector are approximatively of $(x = 6\,\text{m}) \times (y = 5\,\text{m}) \times (z = 20\,\text{m})$.

The LHCb detector is composed of several subdetectors, see Figure 2.3, listed hereafter from left (upstream) to right (downstream):

- Vertex Locator (VELO);
- First Ring Imaging Cherenkov counter (RICH1);
- Trigger Tracker (TT);
- Magnet;

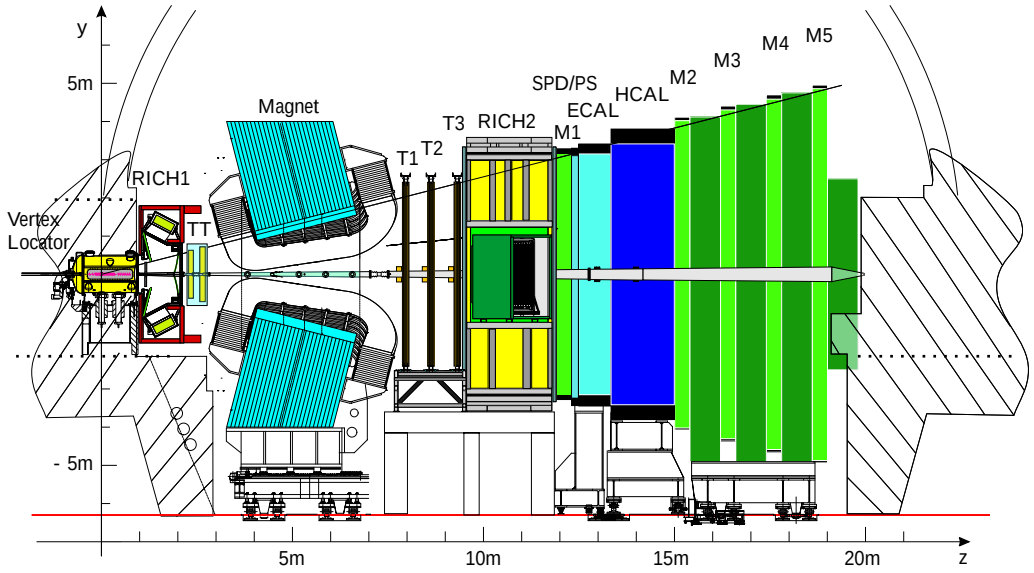


Figure 2.3: The LHCb detector in the vertical (non-bending) plane [17].

- Tracking stations (T1, T2, T3), consisting of an inner part, the Inner Tracker (IT) and of an outer part, the Outer Tracker (OT);
- Second Ring Imaging Cherenkov counter (RICH2);
- First Muon Station (M1);
- Scintillating Pad/Pre-Shower Detector (SPD/PS);
- Electromagnetic Calorimeter (ECAL);
- Hadronic Calorimeter (HCAL);
- Remaining Muon Stations (M2, M3, M4, M5).

The beam pipe traverses the whole detector setup except the VELO which is enclosed in a vacuum vessel connected to the beam pipe.

The subdetectors can be categorized into tracking devices, used to determine the track trajectories and vertices, and particle identification (PID) detectors, used to identify the various particle species:

Tracking: VELO, TT, Magnet, IT, OT.

PID: RICH1, RICH2, ECAL, HCAL.

The magnet is used to bend the trajectory of charged particles and thus determine their momentum from the deflection in the magnetic field. All the different parts of the detector will be concisely discussed in the next subsections.

The trigger system is divided in three levels: the Level-0 (L0), the Level-1 (L1) and the High-Level Trigger (HLT), and will be presented in Section 2.2.8. The role of the trigger is to select the events that are relevant to LHCb's physics goals. As we cannot afford to record every event, the retention rate is successively reduced by each level before storage, while trying to keep the highest efficiency on signal events.

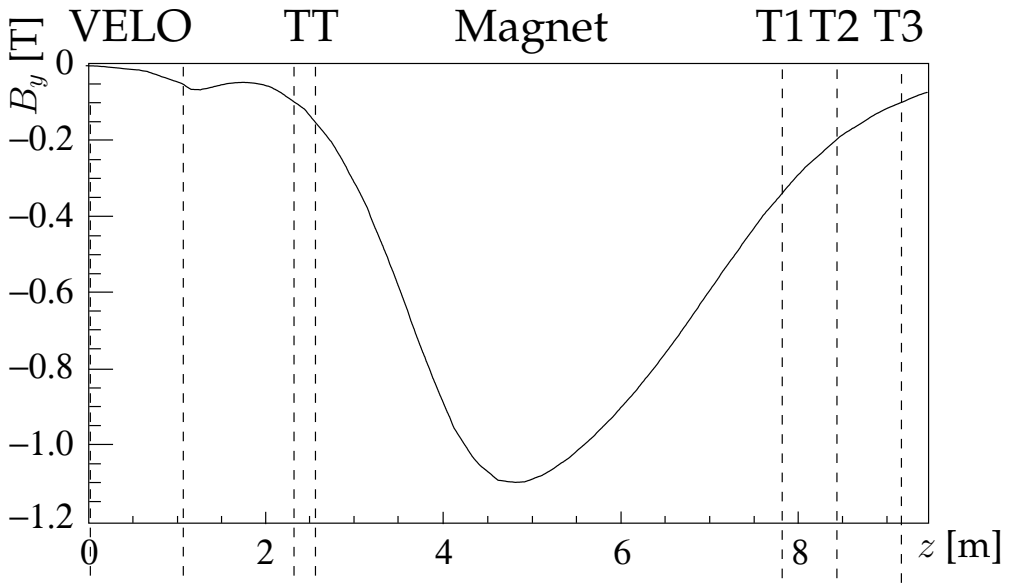


Figure 2.4: The main component of the magnetic field strength (B_y) along the z axis [17].

2.2.1 Magnet

The spectrometer's dipole [59] is located near the interaction region and its aperture defines the detector's acceptance. The momentum of tracks is determined from their curvature in the magnetic field, which has a bending power characterized by the total integrated field $\int B \, \mathrm{d}l \sim 4\,\mathrm{Tm}$. The main component of the field is oriented along the vertical y axis; it is shown as a function of z in Figure 2.4, where the area of the curve represents the integrated field for B_y. The magnet's bending power enables precise measurements of the momentum. For instance a momentum resolution of $\delta p/p \sim 0.4\%$ is achieved for high-momentum long tracks with $p \sim 40\,\mathrm{GeV/c}$.

The polarity of the magnet can be reversed in order to study the systematics induced by possible left-right asymmetries in the detector. This motivated the choice of a warm magnet rather than superconducting.

2.2.2 Vertex Locator

The role of the Vertex Locator [17, 60] is to provide precise measurements of the charged particles close to the interaction region. These measurements are used to reconstruct primary vertices and decay vertices. The VELO provides information on backward tracks used together with forward tracks to identify primary tracks. The forward VELO tracks determine the direction of charged particles to be later used for the matching in the other tracking stations. This subdetector is also used in the trigger to help identifying signal events.

An accurate measurement of track parameters is required to determine the time of flight of long-lived particles. To this end, the VELO uses silicon strip sensors placed very close to the beam, perpendicular to the beam direction. The whole system is contained inside a vacuum vessel with a Roman pot system to be able to move away from the beam the two detector halves during beam injection, and to avoid radiation damage. The

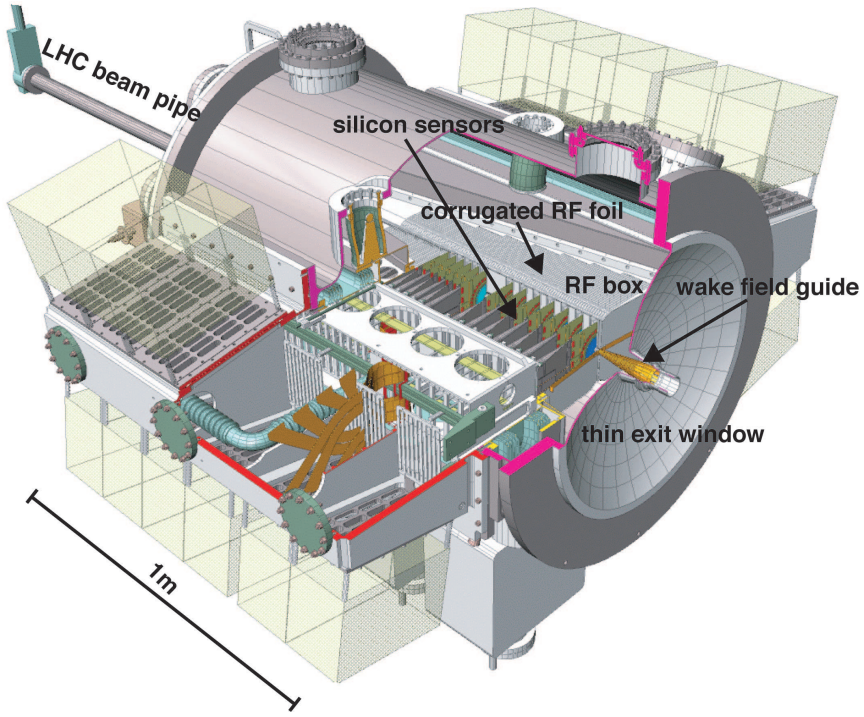


Figure 2.5: The VELO vacuum vessel with its main components [17], where the sensor layout is that of [60].

VELO sensors are placed in a secondary vacuum by a thin aluminum corrugated foil. This so-called RF foil acts as a radio frequency (RF) shield and protects the electronics from electromagnetic pickup from the LHC beams. The VELO vacuum vessel together with the main VELO components are shown in Figure 2.5, with the old sensor layout from [60].

The VELO consists of 21 stations spread along the beam line each with two types of silicon sensors: one measures the radial r coordinate with circular strips centered around the beam axis and the second measures the azimuthal φ coordinate with radial strips. The half-disc sensors, shown in Figure 2.6, are arranged in pairs mounted astride the beam pipe. The sensors are 220 μm thick n-to-n single-sided silicon strips.

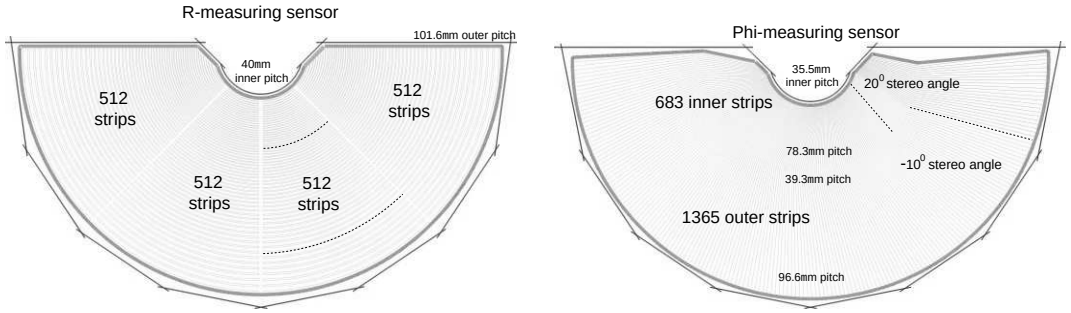


Figure 2.6: Layout of the r and φ sensors [17]. Some strips are indicated with dotted lines for illustration.

The trigger exploits the $r - \varphi$ geometry by first reconstructing tracks in the $r - z$ projection where high impact parameter tracks with respect to the primary vertices are easily identified. Then the space tracking is done on a subset of tracks reconstructed in two dimensions. This strategy allows to save processing time.

The VELO also has a pileup counter which consists of two dedicated r sensors stations upstream of the interaction point. The pileup system only detects backward tracks; it determines the track multiplicity in the backward region as well as the number of primary interactions within a given bunch crossing. This information is used in the very first trigger level to select clean events.

2.2.3 Trigger Tracker

The Trigger Tracker (TT) [17] is located downstream of the RICH 1 and in front of the entrance of the magnet, at ~ 2.5 m from the interaction region. Its main purpose is to assign or add momentum information to tracks. As it can be seen from Figure 2.4, there is a low integrated field between VELO and TT of ~ 0.15 Tm. The deflection of the tracks induced by the fringe magnetic field in TT can thus be used to roughly estimate the momentum of particles. For high-momentum particles, e.g. a few GeV/c, the deviation from a straight trajectory in the middle of TT will be a few millimeters in the bending plane. This is sufficient to determine the momentum with 10 – 40% accuracy without the use of the other tracking stations T. The required spatial resolution is provided by silicon strip detectors.

The TT is first used in the trigger, to assign momentum information to the different trigger tracks. Depending on the trigger stage, not all the subdetector information is available and the processing time for the reconstruction of triggering objects and for the trigger decision is limited. For instance TT is used to determine the transverse momentum only for the high impact parameter tracks. TT is also used in the on-line tracking to improve the speed performance. The VELO-TT tracking in the trigger will be explained in detail in Chapter 4. However, we give in this subsection the relevant information used in (the development of) the tracking code.

Besides its use in the trigger, TT also serves to assign a momentum estimate to low momentum tracks, e.g. the slow pion from $D^* \to D^0\pi$. These tracks seldom reach the T stations since they are bent out of the acceptance. Furthermore, the information from TT is used to improve the resolution of the long tracks. Finally, TT enables the reconstruction of the decay products of long-lived particles, e.g. K^0_S, that could decay outside the VELO acceptance.

The Trigger Tracker consists of four detection layers grouped in two stations of two layers each, TTa and TTb, separated by 27 cm. The readout strips are arranged in stereo views to allow the spatial reconstruction, with a vertical orientation to precisely measure the position in the horizontal bending-plane. The first and the fourth layers have vertical strips (X layers, 0° angle with the y axis) whereas the second and the third have strips rotated by a stereo angle of -5° (U layer) and $+5^\circ$ (V layer), respectively.[1] The layout of the two layers in TTb is shown in Figure 2.7, which corresponds to the layout of the simulation used in this dissertation. The detailed description of the TT station can be found in [61].

[1]The stereo angle is defined here as the angle between the strip's direction and the y axis, in the $x - y$ plane. The angle is *positive* when rotating from the positive x axis to the positive y axis in the transverse plane.

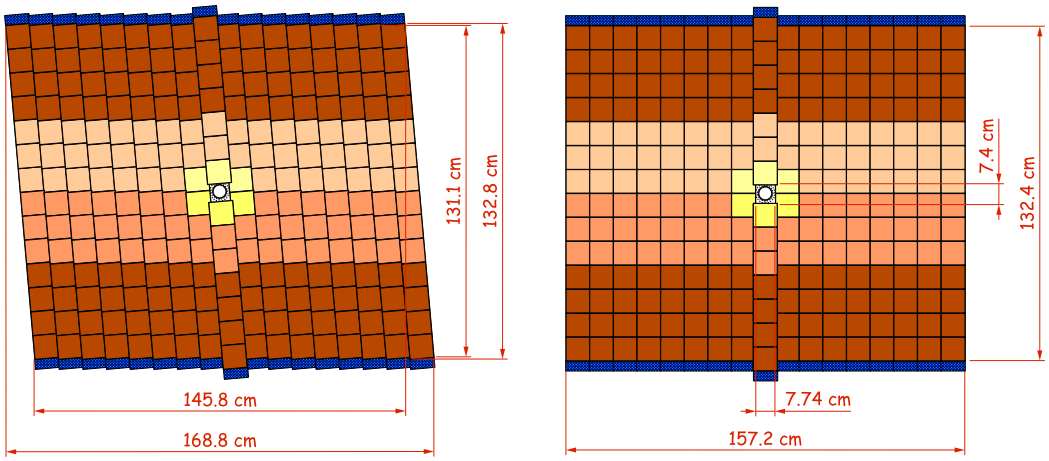


Figure 2.7: Layout of the first $+5^\circ$ stereo (left) and second vertical (right) layers of TTb [61]. The different readout sectors along a silicon ladder are indicated by different colors.

The first two layers (TTa) are centered around $z = 235$ cm (X layer: $z = 2332$ mm, U layer: $z = 2368$ mm) downstream of the nominal interaction point, and the last two layers (TTb) are centered around $z = 262$ cm (V layer: $z = 2602$ mm, X layer: $z = 2638$ mm). The layout of the detection layer is such that the areas above and below the beam pipe are each covered by a single seven-sensor long silicon ladder, the areas to the left and to the right of the beam pipe are covered by seven (TTa) or eight (TTb) 14-sensor (2×7) long ladders. The sensors are staggered, i.e. the position of each sensor is shifted in z with respect to their nominal detection plane such that sensors overlap in x between consecutive ladders. Each ladder is vertically split into several readout sectors as indicated by different colors in Figure 2.7. A finer segmentation in the central region of the detection layers giving an improved trigger performance, there are two different partitioning of the readout sectors:

- 4-2-1-1-2-4: The three ladders closest to the beam pipe are split in two-sensor three readout sectors per half: an outer four-sensor readout sector, an intermediate readout sector and an inner single-sensor readout sector.

- 4-3-3-4: All the other ladders have a two readout sectors per half.

The number of sensors is 420 in TTa (2 layers times 15 ladders times 14 sensors) and 476 in TTb (2 layers times 17 ladders times 14 sensors), representing a total active area of $\sim 8.4\,\mathrm{m}^2$. The silicon readout strip pitch is $183\ \mu$m.

The acceptance of the TT station is limited towards small polar angles by the beam pipe. The square-shaped beam pipe hole has a width of ~ 7.7 cm and a height of ~ 7.4 cm. The approximative width and height of the detection layers are ~ 160 cm and ~ 130 cm, respectively. The acceptance coverage is reduced by the dead area between two consecutive sensors on a ladder. These gaps lead to vertical dead zones of ~ 2.8 mm.

2.2.4 Tracking Stations

The main purpose of the tracking stations (T1, T2, T3), situated right after the magnet, is to provide the different clusters to be used in the reconstruction of long charged tracks in

order to measure their momentum. Moreover, the T stations have to provide the direction of the tracks to be used for the reconstruction of Cherenkov rings by the RICH detectors. The T measurements are also used as seeds for the reconstruction in the calorimeters and muon chambers. Each T station consists of two parts: the Inner Tracker (IT) and the Outer Tracker (OT). This separation is the result of the variation in particle density for different polar angles, hence the use of two detectors with different granularity. The track multiplicity and density is expected to be the highest at small angles thus requiring the use of silicon micro-strip sensors. This is the technology used for the IT. At larger polar angles and away from the beam, the particle density will be moderate. This region will be covered by the OT using gas drift chambers, where this choice of detector is based on the optimization of the cost and the spatial resolution.

Inner Tracker

The Inner Tracker (IT) [62][2] covers the innermost region of the T stations with a cross-shaped area around the beam pipe. An IT station contains four boxes with four layers each. Similarly to the TT station, the layers are arranged in X-U-V-X stereo views with silicon sensors. The area covered by the IT is approximatively 120 cm in width and 40 cm in height.

Outer Tracker

The Outer Tracker (OT) [63] covers the rest of the area of the T stations, outside of the IT acceptance. As for the IT, the configuration of the detection layers is two stereo layers embraced between the external X layers. The detection layers are made of straw tubes filled with a gas mixture containing 75% Ar, 15% CF_4 and 10% CO_2. This composition was determined to get a total drift time (or total signal collection time) well below the maximum 50 ns, i.e. within the time of two LHC bunch crossings (2×25 ns). In this way the number overlapping events is under control.[3]

2.2.5 RICH Counters

The Ring Imaging Cherenkov counters (RICH) [17, 63] provide particle identification, i.e. enable the discrimination between different particle hypotheses. This information is required in b-physics analyses for the separation of the decay channels with the same topology, such as in $\mathrm{B_q} \to \mathrm{hh}$ ($\mathrm{h} = \pi, \mathrm{K}, \mathrm{p}$). Moreover a good $\pi - \mathrm{K}$ separation is needed for the flavor tagging.

The particle identification by the RICH counters must cover the largest momentum spectrum. To this end two RICH detectors are used at LHCb : RICH1 is located between VELO and TT, and RICH2 is situated between the T stations and the calorimeters. RICH1 identifies particles in the momentum range between ~ 1 and ~ 60 GeV/c and RICH2 provides identification up to ~ 150 GeV/c.

The RICH detectors use the Cherenkov effect: when a charged particle traverses a medium with a velocity greater than the speed of light in that medium, then an electromagnetic radiation is emitted. By measuring the angle θ_C between this Cherenkov light and the direction of the particle, one can determine the particles velocity β and thus

[2] Many changes in design occurred since this TDR, which is somehow obsolete.

[3] Note that the gas mixture has recently changed to a composition of 70% Ar and 30% CO_2, for a drift time below 75 ns.

determine its mass when the momentum is known. The Cherenkov angle $\theta_{\rm C}$ satisfies $\cos\theta_{\rm C} = 1/(n\beta)$, where n is the refraction index of the radiator medium. In order to observe the Cherenkov light, the condition $n\beta \geq 1$ must be satisfied. The different radiators are then chosen depending on the desired momentum coverage. For instance, in the large momentum range a small refractive index will be used. The radiators used for RICH1 are silica aerogel ($n = 1.03$) and C_4F_{10} ($n = 1.0014$). For RICH2 the radiator is CF_4 ($n = 1.0005$).

In the RICH detectors mirrors project the light cones as rings onto a plane of photon detectors (Hybrid Photon Detectors, HPDs), where the radius of each ring gives a measure of the corresponding Cherenkov angle $\theta_{\rm C}$. A global likelihood analysis is performed to fit at best the rings on the observed patterns, and thus to assign to each track probabilities associated with each mass hypothesis. A typical event in RICH2 is shown in Figure 2.8.

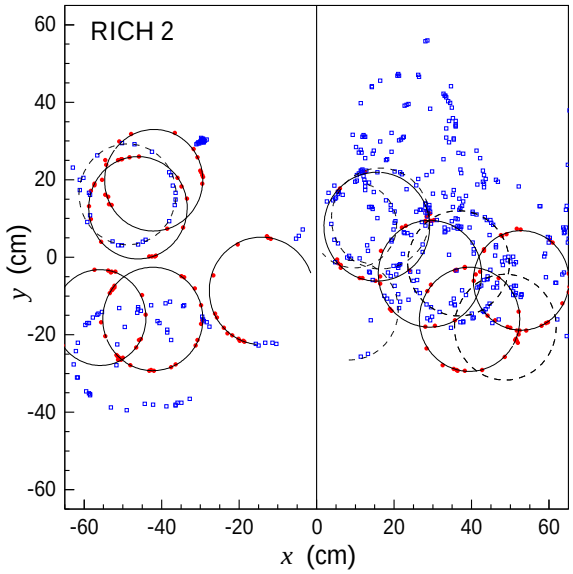


Figure 2.8: A typical RICH2 event display of the detected photoelectrons in the two detection planes [17]. The reconstructed rings are superimposed.

2.2.6 Calorimeters

The calorimeters [64] are located in the middle of the LHCb detector, between the first and the second Muon Stations. They are used to absorb almost all particles, measuring their energies and positions. The only exception is that of the weakly interacting muons, which are detected in the Muon Stations, and of course the neutrinos. The calorimeter clusters are determined by measuring the energy deposited by a particle as a result of the ionization process induced by the cascade of secondary particles produced by this particle, until absorption. All the light emitted by the shower in the scintillating material is collected and represents a measure of the particle energy.

The calorimeter clusters are used in the Level-0 trigger to identify high transverse mass objects as signatures of b decays. They are also used to identify electrons and to reconstruct neutral particles (γ, π^0).

Scintillating Pad Detector (SPD)

The SPD identifies charged particles (as opposed to neutral particles) before the calorimeter shower through their ionization. This detector thus distinguishes between photons and electrons. It is made of 15 mm thick scintillator pads. The scintillation light is directed to the multi-anode photo-multipliers with a wavelength shifting (WLS) fiber.

Pre-Shower (PS)

A 12 mm thick lead wall initiates the electromagnetic shower then detected by the PS, which shares the SPD features. The PS distinguishes between electrons and hadrons.

Electromagnetic Calorimeter (ECAL)

The ECAL uses Shashlik-type modules[4], which consist of alternating 4 mm think scintillating tiles with 2 mm thick slices. The ECAL detects electrons and photons via their electromagnetic showers of $\mathrm{e}^+\mathrm{e}^-$ and γ. The segmentation of the ECAL is that of the SPD and PS. The design energy resolution is $\sigma(E)/E = 10\%/\sqrt{E} \oplus 1.5\%$ (E in GeV), where the first stochastic term describes the statistical fluctuations in the showers, the second determinist term represents the systematics induced by the detector, and both terms are quadratically added.

Hadronic Calorimeter (HCAL)

The HCAL is placed after the ECAL and identifies hadrons via their interactions with the detector's material. The HCAL consists of 16 mm thick iron plates and 4 mm thick scintillating tiles which are parallel to the beam direction. The expected energy resolution is $\sigma(E)/E = 80\%/\sqrt{E} \oplus 10\%$.

2.2.7 Muon System

Muons are the only charged particles likely to traverse the calorimeters. They are thus identified by a dedicated Muon System, placed behind the calorimeters (Stations M2, M3, M4, M5) with the exception of Station M1 which is located before the SPD.

The muon detector plays an important role in the trigger. In the off-line reconstruction the Muon System is used to identify muons, starting from the tracks found in the T stations and extrapolating them to the Muon Stations to confirm the muon hypothesis.

The Muon Stations are for the greatest part equipped with multi-wire proportional chambers (MWPC), divided into four regions with different pad granularity. The innermost part of M1 uses gaseous electron multiplier detectors (triple-GEM). The muon shield consists of the calorimeters and of four layers of 80 cm thick iron walls.

2.2.8 Trigger System

The LHC bunch crossing frequency of 40 MHz basically means that every 25 ns a pp collision can take place. Recording events at this rate is not possible and would lead to a huge data size that cannot be reconstructed in a reasonable time. Furthermore, most events are uninteresting for physics. The event rate must therefore be reduced before sending

[4] A Shashlik is a popular form of Shish kebab from the former Soviet Union, usually alternating pieces of meat and fat.

events to permanent storage and making them available for a full off-line reconstruction. This is the task of the trigger system [65], which thus represents an essential component of the LHCb experiment. The trigger must be fast to decide whether or not to keep an event as early as possible in the trigger chain, and it must be efficiency on the "signal" events we look for. All the difficulty in designing the trigger selections stems from these two key constraints: speed and efficiency.

At the nominal luminosity of $\mathcal{L} = 2 \times 10^{32}$ cm^{-2}s^{-1}, an inelastic pp collision takes place at an average rate of $\sigma_{\rm pp}^{\rm inel}\mathcal{L} = 16$ MHz, which reduces to 12.4 MHz minimum-bias event rate including the effect of the pileup.[5] The probability of having a $b\overline{b}$ pair in each event is $\sim 1/160$, hence a large part of the trigger input rate is actually made of "background" events, as opposed to "signal" b-events. The rate of events with a $b\overline{b}$ pair is correspondingly of $\sigma_{b\bar{b}}\mathcal{L} = 100$ kHz. Only a small fraction of the $b\overline{b}$ pairs will have their decay products in LHCb's acceptance such that the effective rate we could detect in LHCb is of the order of a few kHz. The role of the trigger is thus to select the interesting events while rejecting the background, and lower the event rate to an acceptable level. The trigger selection is based on typical b signatures: large impact parameter tracks and large transverse momentum particles.[6]

The trigger is divided into different levels. In the initial stage the selection of events is based on simple requirements using partial detector information since there is not enough time for a full readout reconstruction. Then the complexity of the trigger selection increases as the retention rate is successively reduced and more tracking information becomes available. The LHCb trigger system has three trigger levels:

Level	Input rate	Output rate	Reduction factor
Level-0 Trigger (L0)	12.4 MHz	1 MHz	16
Level-1 Trigger (L1)	1 MHz	40 kHz	25
High-Level Trigger (HLT)	40 kHz	2 kHz	20

While L0 is implemented in custom electronics ("hardware" trigger), L1 and the HLT are software triggers to be run in the on-line CPU farm. Note that in the next generation of the trigger, currently being implemented, the Level-1 trigger no longer exists, and a 1 MHz readout scheme is adopted. This enables a more uniform trigger avoiding double buffer decoding and reconstruction, as well as a better treatment of the triggering objects.

As the requirement for the trigger is to provide the largest number of b events to be used off-line, the trigger efficiencies are determined on off-line selected events. As an illustration of the trigger performance, we consider hereafter two important channels, namely $B_s \to D_s\pi$ and $B_s \to J/\psi\phi$. The different cut values given correspond to the Data Challenge 04 version of the code used for this thesis.

The Level-0 trigger

The L0 trigger is the first trigger stage used to reduce the event rate from 16 MHz to 1 MHz. The latency of L0, which is the time elapsed between a pp interaction and the arrival of the L0 trigger decision to the FE electronics, is fixed to 4 μs. This implies a buffering of the data in the pipe-line memory of the FE chips.

The L0 involves four subsystems: the pileup system, the calorimeter trigger, the muon trigger, and the decision unit that compiles the global decision:

[5] The average number of pp interactions in bunch crossings with at least one interaction is obtained from (for a constant luminosity) $< n_{\rm pp} > /(1 - e^{-<n_{\rm pp}>}) = 1.29$, with (2.2): $< n_{\rm pp} >= (80 \times 2)/300$.

[6] Large refers here to the kind of particles encountered in LHCb's environment.

Pileup system This system detects the multiple primary vertices, as explained in Section 2.2.2. Its purpose is to veto bunch crossings with multiple interactions as a decision only based on large transverse energy is not sufficient. An event is vetoed if the total track pileup multiplicity is ≥ 112, or if the number of tracks used to reconstruct a second pp interaction is ≥ 3. The di-muon trigger overrules the pileup veto.

Calorimeter trigger An event is accepted when it contains a calorimeter cluster with $E_{\mathrm{T}} > 2.6$ GeV for electrons, $E_{\mathrm{T}} > 2.3$ GeV for photons, $E_{\mathrm{T}} > 3.5$ GeV for hadrons or $E_{\mathrm{T}} > 4.0$ GeV for π^0's.[7] There is veto on the event if the sum of the total energy in the calorimeters is below 5.0 GeV. Moreover, the event is also rejected if the SPD multiplicity is ≥ 280.

Muon trigger An event is accepted if there is a muon with a transverse momentum $p_{\mathrm{T}} > 1.3$ GeV/c or when the sum of the p_{T} of the two largest p_{T} muons is above 1.5 GeV/c. In the latter case, the event is accepted irrespective of the vetoes.

Decision Unit It combines the information from the pileup system, the calorimeter and muon triggers to yield the L0 decision.

The L0 efficiency is $\sim 43\%$ for $\mathrm{B_s} \to \mathrm{D_s}\pi$ and $\sim 94\%$ for $\mathrm{B_s} \to \mathrm{J}/\psi\phi$, on off-line selected events.

The Level-1 trigger

The L1 trigger will reduce the event rate from 1 MHz to a maximum output rate of 40 kHz. It is a software trigger sharing the on-line farm with the High-Level Trigger. Assuming 400 dedicated CPUs, the processing time is ~ 0.4 ms. The latency is variable, with values up to 55 ms.

The L1 trigger decision is based on data from the VELO, TT, and the L0 Decision Unit. The selection strategy is based on detached tracks and large transverse energy or momentum particles. Before taking any decision, the VELO and VELO-TT reconstructions are performed to determine the primary vertices, track impact parameters with respect to those primary vertices (IP), and transverse momenta (p_{T}). The reconstruction steps are summarized hereafter:

VELO reconstruction The L1 tracks are reconstructed in the $r-z$ projection and used to determine the 2D primary vertices. Only tracks with $0.15 <$ IP < 3 mm or the ones matching a L0 object are reconstructed in 3D using the φ clusters.

VELO-TT matching The matching of VELO space tracks to the TT station determines the momentum of particles from their deflection in the magnetic field. The VELO-TT code used is the same as in the HLT, with a few less features. This will be explained in great detail in Chapter 4.

The events passing the L1 trigger can be selected by different parallel trigger lines:

- **Generic line:** The event is accepted if the logarithmic sum of the transverse momenta (in MeV/c) is $\sum \ln p_{\mathrm{T}} > 14.34$ for the two largest p_{T} tracks, with no veto on the number of reconstructed primary vertices.

[7] There is actually a distinction between a "local" π^0 where the two photons are expected on the same FE card ($E_{\mathrm{T}} > 4.3$ GeV) and a "global" π^0 where the two photons are on neighboring FE cards ($E_{\mathrm{T}} > 3.7$ GeV).

- **Single muon line:** The event is accepted if there is a muon with $p_{\mathrm{T}} > 2.3$ GeV/c.
- **Di-muon line:** The invariant mass of two muon candidates must satisfy $m_{\mu\mu} > 500\,\mathrm{MeV/c^2}$ with a positive impact parameter $\mathrm{IP}_{\mu\mu} > 0.075$ mm in order to trigger an event.
- **J/ψ line:** A di-muon with an invariant mass close to or above the J/ψ mass, $m_{\mu\mu} > m_{\mathrm{J}/\psi} - 500\,\mathrm{MeV/c^2}$, triggers the "J/$\psi$ line".
- **Electron line:** The event is selected if the largest E_{T} electron has $E_{\mathrm{T}} > 3.44$ GeV and $\sum \ln E_{\mathrm{T}} > 13.2$ (E_{T} in MeV).
- **Photon line:** The event is triggered if the largest E_{T} photon has $E_{\mathrm{T}} > 3.06$ GeV and $\sum \ln E_{\mathrm{T}} > 13.2$ (E_{T} in MeV).

The L1 efficiency is $\sim 82\%$ for $\mathrm{B_s} \to \mathrm{D_s}\pi$ and $\sim 94\%$ for $\mathrm{B_s} \to \mathrm{J}/\psi\phi$, on events passing the off-line selection and the L0 trigger.

The High-Level Trigger

The High-Level Trigger (HLT) is the last trigger stage before sending events to storage. Its task is to reduce the rate from 40 kHz to 2 kHz. The HLT will run with 400 CPUs in the on-line farm, such that the total processing time available is 10 ms per event.[8] At this stage all the detector information is available and a fast on-line reconstruction is run to obtain all the objects required to take the final decision. The on-line reconstruction is first run on a limited number of tracks and the full reconstruction is then executed provided the event is accepted by the generic HLT. The generic HLT is a pre-HLT trigger confirming the L1 decision and looking for potential triggering muons. When an event is selected by the generic HLT, it is passed onto different HLT streams for the HLT decision.

The HLT is divided into four streams with the following preliminary output rates:

Exclusive b ($\sim$ 200 Hz): The core physics stream with exclusively reconstructed decays including sidebands and control channels. A decay mode of interest for the study of CP violation typically has a visible branching ratio ($\mathrm{BR_{vis}}$), i.e. a product of all the branching fractions involved in the decay chain, smaller than 10^{-4}. Furthermore, not all the decays will be in LHCb's acceptance, approximatively $\varepsilon_{\mathrm{acc}} \sim 10\%$. Then we have to account for the total trigger efficiency $\varepsilon_{\mathrm{trg}} = \varepsilon_{\mathrm{L0}} \times \varepsilon_{\mathrm{L1}} \times \varepsilon_{\mathrm{HLT}}$, which is typically 30% for a 4-prong purely hadronic channel. Given $\sigma_{\mathrm{b\bar{b}}}\mathcal{L} = 100$ kHz and considering that ~ 100 decay channels will be looked at, the event rate for b-physics studies is:

$$\sigma_{\mathrm{b\bar{b}}}\mathcal{L} \times \sum_{i \in \mathrm{decays}} \mathrm{BR}^i_{\mathrm{vis}} \times \varepsilon_{\mathrm{trg}} \times \varepsilon_{\mathrm{acc}} \lesssim \mathcal{O}(30\,\mathrm{Hz}) < 200\,\mathrm{Hz}\,.$$

This simple calculation does not take background into account, but shows that many channels can be accommodated within the dedicated bandwidth.

D* ($\sim$ 300 Hz): PID-blind $\mathrm{D}^* \to \mathrm{D}^0\mathrm{h}$ events with $\mathrm{D}^0 \to \mathrm{hh}$ and loose D^0 mass cut. Each final state hadron is assigned the pion mass ($\mathrm{h} = \pi$) which allows to measure the PID efficiency and misidentification rate. For instance, we can benefit from the fact

[8] We consider the readout scheme with the L1 trigger.

that the particle with the smallest momentum corresponds to the slow pion, or we can use the two same sign decay products to separate pions from kaons. This stream can also be used for CP measurements in D decays.

Di-muon ($\sim$ 600 Hz): Lifetime unbiased di-muons with a mass above a threshold of 2.5 $\mathrm{GeV/c^2}$. These events are used to measure the uncertainty on lifetime measurements. For instance, the negative tail in the proper time distribution will enable the extraction of the proper time resolution.

Inclusive b $\rightarrow \mu$ ($\sim$ 900 Hz): Events with a high p_{T} and high IP muon, used for systematic studies of the trigger efficiency and for data-mining. Because of the muon, this sample is highly tagging-enriched.

Additional inclusive triggers, e.g. inclusive ϕ or $\mathrm{D_s}$ could also be added. Note that in the HLT version presented in this dissertation, the RICH information is not used on-line. The whole HLT reconstruction and selection sequence will be extensively described in Chapter 4, which is devoted to the exclusive HLT selections.

The combined (generic HLT and HLT streams) HLT efficiency is $\sim 78\%$ for $\mathrm{B_s} \rightarrow \mathrm{D_s}\pi$ and $\sim 90\%$ for $\mathrm{B_s} \rightarrow \mathrm{J}/\psi\phi$, on events passing the off-line selection and the L0 and L1 triggers.

Chapter 3

LHCb Monte Carlo Simulation

The full Monte Carlo simulation of the detector is introduced. First the software framework is explained with each component of the event generation, detector's simulation, the reconstruction of tracks, and the analysis program. Then the flavor tagging methods are discussed.

"La Scapiliata", Leonardo da Vinci, circa 1508.

THE complexity of the experiments in high-energy particle physics requires many years of research and development to design the detector and all the computing technologies. The large amount of data to be collected demands an efficient and reliable computing framework and a stable software. In order to assess the detector's performance and to test the software, the use of so-called Monte Carlo simulations is essential.

A Monte Carlo (MC) method is a numerical technique for calculating probabilities and their related quantities using sequences of random numbers.[1] The (pseudo-) random numbers are uniformly distributed values in the interval $[0, 1]$, and are generated by a *generator*. This sequence is in turn used to generate another sequence distributed according to some probability density functions (pdf's) which describe or *emulate* the observables.[2] These values are what we call the simulated MC data or simply the MC truth. Note that the use of random numbers as seeds of the simulation makes the MC technique a stochastic (i.e. non-deterministic) method.

There are several steps in LHCb's Monte Carlo simulation, which will be described in this chapter. Firstly, we need to generate the events, that is the primary pp collisions and the successive decays. Secondly, the simulation part will emulate the passage of particles through the apparatus and the detector response. Thirdly, the simulated data is digitized for the last phase of the simulation. At this point we can ignore the fact that we used a MC simulation and treat the data as if it were coming directly from the detector and DAQ system. Finally, the reconstruction of the data is performed, that is clusters and tracks are reconstructed, to yield the necessary objects and make them available to the physics analyses.

We will thus introduce how the event generation is done, describe how the relevant

[1] The Monte Carlo appellation is in reference to the famous casino in the Principality of Monaco.

[2] Given a measurement x, the pdf must be normalized such that $\int \mathrm{pdf(x)dx} = 1$: x must be somewhere in the sample space. Formally, the MC calculations are integrations of the pdf's over the sample space.

data samples are generated, and how the results can be validated using the knowledge of the MC truth. Furthermore, we describe the applications used to perform the physics analyses and the underlying software framework.

3.1 Simulation Framework

The LHCb software [66] is based on the object-oriented (OO) architecture of GAUDI [67, 68]. This framework is written in C++, allowing objects to transform themselves through their methods. The distinctive feature of GAUDI is that the algorithmic part of data processing is also considered as a set of OO objects. There is thus a clear distinction between data objects (e.g. Event Model) and the algorithms or specialized tools that manipulate these objects with a well defined input and output data. The data flow between algorithms proceeds via the so-called Transient Store. For instance, the average user performing a physics event selection will use the Transient Event Store (TES) to communicate between algorithms the particles (objects) he is interested in. Note that these objects are *transient*, that is they will be available in the store for the lifetime of the job. For the TES, this lifetime is one event.

The GAUDI framework is decomposed into independent services providing the basic software tasks needed in an application. These services have the possibility of using third-party components. All the LHCb applications are embedded in the GAUDI framework where they perform their tasks, e.g. event generation, detector simulation, digitization, reconstruction, trigger, physics analysis, and the event and detector visualization. An important example of service provided by GAUDI is the Job Options Service. It is used to configure the applications at run-time through the named properties associated to the data members of the components, e.g. algorithms. This is for instance extensively used in the exclusive High-Level Trigger selections, where each dedicated selection is an instance of the same algorithm configured by options and sharing standard tools for the selection criteria [69, 70].

The different components of the LHCb software and simulation are listed hereafter, together with their GAUDI applications:

Event generation The pp collisions are generated with the PYTHIA program [56]. The output is a collection of particles described by their momentum four-vector. The pileup effects, i.e. multiple pp collisions in the same bunch crossing, are simulated in PYTHIA for each collision. The decay of the b hadrons is delegated to a specialized decay package, EVTGEN [71]. This package was originally designed by the BABAR collaboration and has been adapted by LHCb for the decays of all b hadrons [72]. Both programs are controlled by GAUSS [73], a GAUDI-based application. More details on the event generation will be given in Section 3.2.

Detector simulation The second phase, also steered by GAUSS, consists of the tracking in the LHCb detector of the particles produced by the generator phase. The evolution of the particles in the detector including the effect of the magnetic field, the interactions with the detector material and environment, and many other processes which the particles undergo when traversing the experimental setup are taken into account. These tasks are done by the GEANT 4 toolkit [74].

Digitization The final phase in the simulation is the digitization, controlled by yet another GAUDI application named BOOLE [75]. BOOLE applies the detector response

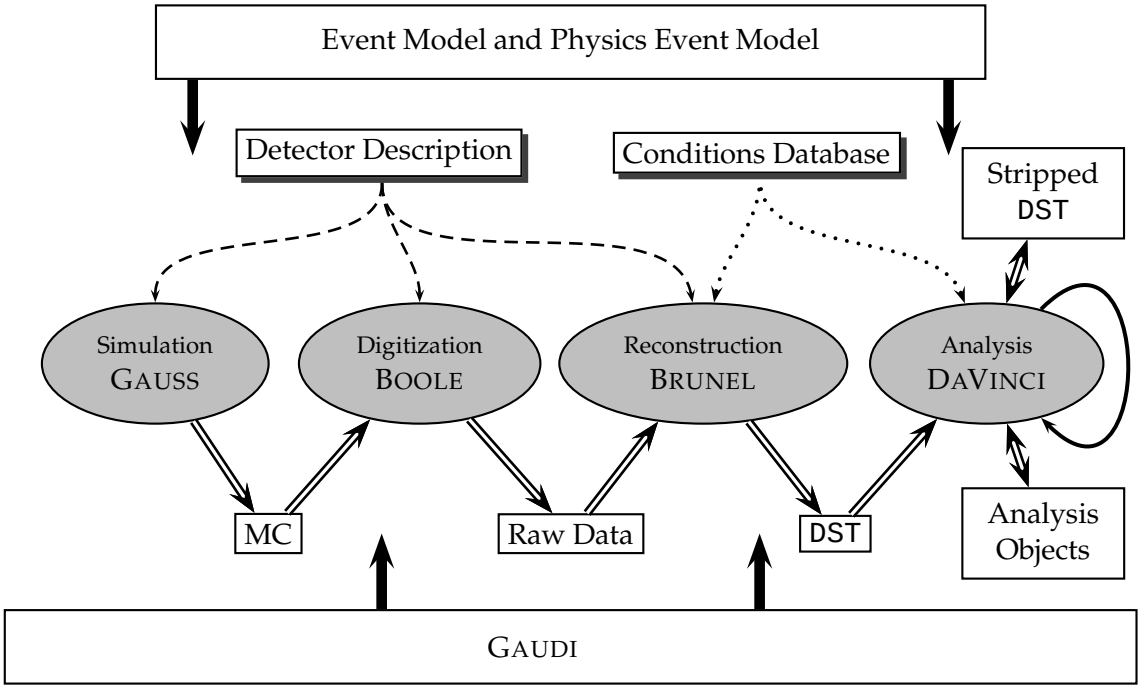


Figure 3.1: The LHCb data processing applications and data flow. The Event Model describes the data expected by the applications embedded in GAUDI. From the Raw Data box, there is no distinction between MC data and data coming from the detector.

to the hits deposited by the particles in the sensitive materials of the detector and these hits are digitized into an electric signal. The simulated response is calibrated using test-beam data. The digitization procedure includes the simulation of the detector response, of the readout electronics, and of the L0 trigger hardware. Furthermore, the spill-over effects are simulated by adding bunch-to-bunch collisions. Spill-over can occur since slow particles from a previous collision may leave measurements in the detector, or because of the readout time of some subdetectors being larger than the time between two consecutive bunch crossings (25 ns). BOOLE uses the instantaneous luminosity of the current event to determine the probability of one or more interactions occurring in the two preceding ($t = -25$ ns, -50 ns) and in the following ($t = 25$ ns) beam crossings. The output of BOOLE (L1 buffer, raw buffer) is identical to that of the real data coming from the detector, such that the rest of the simulation can ignore the MC history of the data.

On-line tracking and trigger software As presented in Section 2.2.8, the L1 and HLT trigger software is run on a dedicated processor farm, together with the on-line reconstruction needed to produce the required information for the trigger decision. This intermediate step before sending an event to storage and to the off-line reconstruction is foreseen to run in the same application as the final reconstruction. Note that the on-line tracking starts from the raw buffer. This stage will be detailed in Chapter 4, for the HLT sequence.

Reconstruction After a positive trigger decision, all the data coming from the detector is processed by the BRUNEL [76] application. The (off-line) reconstruction of tracks

and the identification of particles is done using information from all the subdetectors. As the reconstruction starts from the raw buffers, BRUNEL can process data independently of the MC simulation. The output of BRUNEL is saved to DST's (Data Summary Tapes) and made available for physics analyses. The reconstruction with possible misalignments of the subdetectors is currently under investigation. The reconstruction of tracks will be presented in Section 3.3.

Physics analyses The final stage of the data processing is the event selection of the decays of interest, within the DAVINCI application [77]. The assignment of particle identification hypotheses to tracks and calorimeter clusters is performed, yielding particle objects. The reconstruction of primary vertices is then done. Finally, a series of algorithms is executed to combine particles until the whole decay of interest is selected. A set of generic algorithms to select a decay to n decay products and applying selection filters is provided [69], and used in several off-line selections, preselections, and for all the exclusive HLT selections [70, 78]. In addition, a user-friendly toolkit, LOKI [79], is available to facilitate the physics analyses. It provides high-level utilities with physics-oriented semantics and making use of the modern technique of generic template programing. The event selection of $B_s \to \eta_c \phi$, performed within DAVINCI, is presented in Chapter 5.

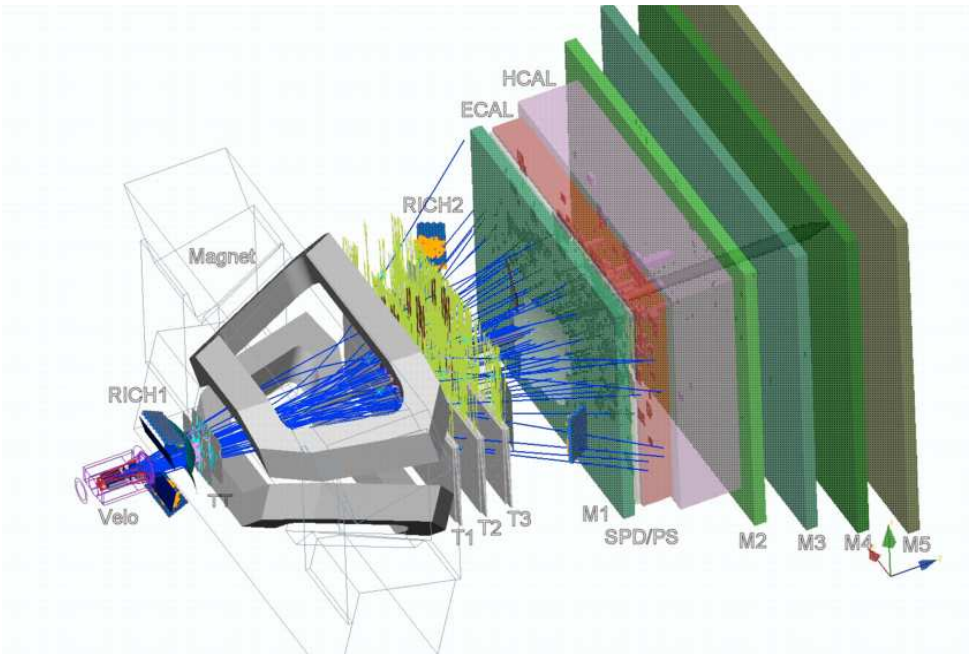


Figure 3.2: The LHCb spectrometer displayed using PANORAMIX. Some detectors are only shown partially to allow visualization of their measurements [66].

Event display The graphical display of the detector geometry and of the event data objects is provided by the visualization application PANORAMIX [80]. Being based on the GAUDI framework, it can be run with any of the LHCb applications, where the event data can be read from files or be produced on the fly. A picture of the LHCb spectrometer displayed with PANORAMIX is shown in Figure 3.2.

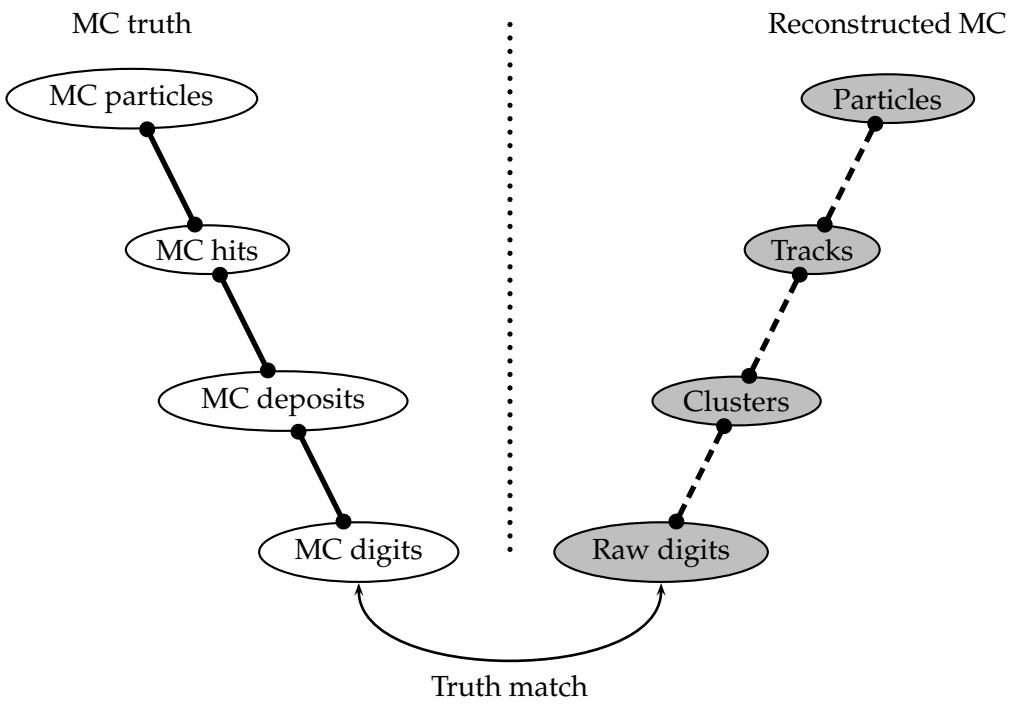


Figure 3.3: Relationships: MC truth data and reconstructed data.

3.1.1 MC Truth Data and Reconstructed Data

There is a clear separation between the generated data (MC truth data) and the reconstructed data in the LHCb Event Model. The information about the MC truth is however preserved in the simulation as it enables to determine efficiencies, and helps in the code development and debugging. The relationship between objects (see Figure 3.3) can only occur between adjacent classes. As a result of the constraint induced by the TES convention which does not allow to modify already published objects, the references between classes are direct and always from the class further in the processing towards the preceding object. The navigation from the MC truth to the reconstructed data proceeds via the matching of unique channel identifiers (`channelID`).

The MC matching between distant objects is done through associators that use a relationship table between the objects. The creation of the relationship tables is done by two different implementations, used in different contexts. These two implementations are called `Linkers` [81] and `Relations` [82].

Additionally, the use of associators helper classes or tools allows to retrieve the association between high-level analysis objects (tracks, `Particles`) and Monte Carlo particles (`MCParticles`) in a transparent way. A reconstructed track or particle is said to be associated to a MC particle if the clusters used to form the underlying track are matched to a certain fraction of hits coming from the same MC particle. If no association is found, then the track or particle is said to be a ghost. Two or more tracks associated to the same MC particle are called clones.

In this thesis, both implementations creating the relationship tables are used depend-

ing on the environment: for off-line quantities the `Relations` are used whereas for on-line and trigger objects the `Linkers` are chosen. As the association to the MC truth is a crucial component for the performance studies, we will further discuss the truth matching in Section 3.5.3 for our use-cases.

3.2 Event Generation

The task of the event generation is to generate the physics processes that take place in pp collisions at a center-of-mass energy of $\sqrt{s} = 14$ TeV and the different production mechanisms. We briefly discuss the phenomenology of pp collisions, the bottom production and the generation of signal samples.

3.2.1 Proton-Proton Collisions

The PYTHIA program simulates the processes occurring during the pp collisions, which are dominated by QCD (Quantum Chromo Dynamics) effects. The factorization hypothesis separates two different regimes: the long-distance non-perturbative QCD effects between the incoming protons (soft scattering) and the short-distance effects (hard scattering, depicted in Figure 3.4) describing the scattering of the partons (gluons and quarks) inside the protons.

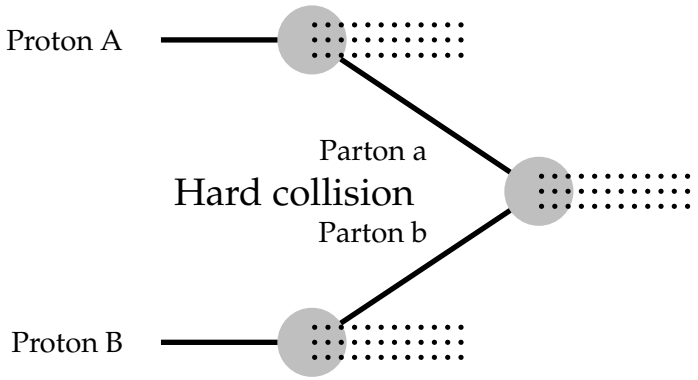


Figure 3.4: Schematic picture of a hard pp collision.

Soft scattering: the protons interact at large distance, hence they effectively see each other as particles without any structure. The momentum transfer is small such that the resulting particles will be produced at small polar angles, along the beam direction.

Hard scattering: two partons from the incoming protons will interact at small distances with a large momentum transfer. This will give rise to large transverse momentum particles with respect to the beam axis, and thus to the creation of massive particles. The hard scattering determines the event characteristics that are relevant to the study of b hadrons.

The complex phase describing the initial pp collision is followed by the hadronization process:

- fragmentation of the partons to colorless[3] hadrons;
- decay of the unstable hadrons.

3.2.2 Bottom Production

In PYTHIA the dominant mechanisms contributing to the $\mathrm{b\bar{b}}$ production and displayed in Figure 3.5 are:

Pair creation ($\sim 16\%$) At leading-order the heavy quark creation is produced through the hard processes of gluon fusion ($\mathrm{gg} \rightarrow \mathrm{\bar{b}b}$) and quark–antiquark annihilation ($\mathrm{q\bar{q}} \rightarrow \mathrm{\bar{b}b}$). The dominant first order contribution is from gluon fusion.

Flavor excitation ($\sim 57\%$) A virtual heavy $\mathrm{\bar{b}}$ quark from the sea quarks of one of the incoming proton scatters with a parton from the other proton ($\mathrm{\bar{b}q} \rightarrow \mathrm{\bar{b}q}$ or $\mathrm{\bar{b}g} \rightarrow \mathrm{\bar{b}g}$) and ends up on its massshell.

Gluon splitting ($\sim 27\%$) In this mechanism no heavy quark is involved in the hard scattering. Instead, the $\mathrm{b\bar{b}}$ pair is produced in the parton shower from a $\mathrm{g} \rightarrow \mathrm{\bar{b}b}$ transition.

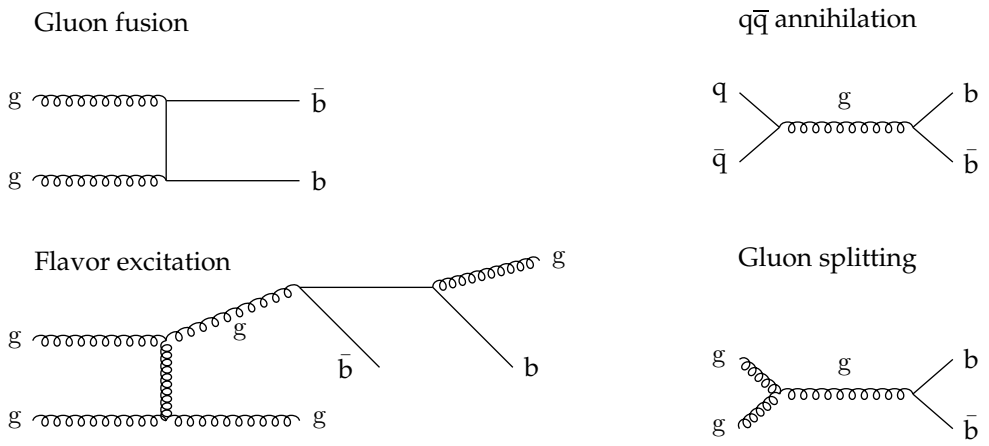


Figure 3.5: Diagrams of the dominant $\mathrm{b\bar{b}}$ production mechanisms, at QCD leading order.

The contributions to the $\mathrm{b\bar{b}}$ production as generated by PYTHIA are indicated within the parentheses, for events with at least one b hadron with all its decay products in LHCb acceptance [83]. The true contributions suffer from large uncertainties, in particular from the phenomenological description of the next-to-leading orders. Furthermore, if the production mechanisms are wrongly simulated, the kinematic properties of the resulting particles may be (dangerously) biased and may not reproduce the real data.

[3]The color here refers to the $\mathrm{SU}(3)$ gauge group of QCD.

The partons involved in the pp hard scattering interact with a large momentum transfer. Since this difference in momentum increases with the center-of-mass energy, then the produced $b\bar{b}$ pair will be boosted along the direction of the higher momentum parton and the direction of the beam. The b hadrons originating from the $b\bar{b}$ pair are therefore predominantly produced in the same forward or backward cone. This is shown in Figure 3.6, where the angular correlation of the b and $\bar{b}$ hadrons produced in pp collisions generated by PYTHIA is plotted. The $b\bar{b}$ angular correlation motivated the design of the LHCb detector as a single-arm forward spectrometer.

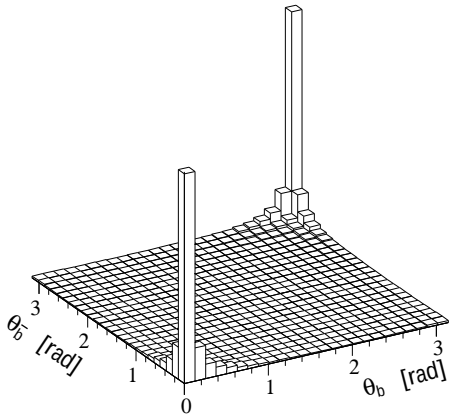


Figure 3.6: Correlation of the polar angles of the b and $\bar{b}$ hadrons produced in pp collisions at $\sqrt{s} = 14$ TeV, as generated by PYTHIA [55].

3.2.3 Samples Generation

The final phase of the event generation after the hadronization of the partons by PYTHIA is the decay of the outgoing hadrons. This step is taken care of by the EVTGEN generator. We distinguish three classes of data samples, keeping different fractions of the events generated by the PYTHIA generator:

Minimum-bias events For this sample all the events generated by PYTHIA are kept, and all the particles are decayed with EVTGEN.

Inclusive events The events containing at least one b or c hadron in 400 mrad with respect to the beam axis are selected. If an event is not selected, the whole interaction is reversed by changing $z \rightarrow -z$. This allows to save processing time as events outside the acceptance cannot be reconstructed. All the particles are decayed with EVTGEN.

Signal events The events containing the b hadron (or any other signal) of interest within 400 mrad are selected. If the event contains several candidates with the required flavor, then one is chosen randomly. The directions of particles are again reversed if the signal hadron is in the backward direction. If the signal hadron does not have the correct flavor (including b hadron excited states), then the hadronization process of PYTHIA is repeated until the interaction contains the required hadron type. Finally the signal candidate is forced to decay in EVTGEN according to the decay chain specified by a decay file, and all the other underlying particles from the event are decayed.

Except for the minimum-bias sample, all the other samples have a generator level cut of 400 mrad. LHCb not being a central detector, generating events in 4π would just be a waste of processing time. The efficiency of this cut depends on the hadron type and must be taken into account for the yields and background levels estimations. The acceptance of the 400 mrad cut is given in Table 3.1 for different samples in the Data Challenge 04. The efficiency for the inclusive $b\bar{b}$ sample is higher compared to the signal samples as either one of the two b hadrons can be in the acceptance.

Table 3.1: Generator-level angular cut efficiencies in the Data Challenge 04 [57]. The uncertainties are statistical.

$\varepsilon^{\theta}_{\mathrm{sig\,B_u}}$ [%]	$\varepsilon^{\theta}_{\mathrm{sig\,B_d}}$ [%]	$\varepsilon^{\theta}_{\mathrm{sig\,B_s}}$ [%]	$\varepsilon^{\theta}_{\mathrm{sig}\,\Lambda_b}$ [%]	$\varepsilon^{\theta}_{\mathrm{D_s}}$ [%]	$\varepsilon^{\theta}_{\mathrm{b\overline{b}}}$ [%]
34.9 ± 0.3	34.9 ± 0.3	34.8 ± 0.3	34.7 ± 0.3	36.5 ± 0.3	43.4 ± 0.3

The EVTGEN package is especially designed to handle b hadron decays. In particular the $\mathrm{B_q}$–$\overline{\mathrm{B}}_\mathrm{q}$ mixing is simulated with the following parameters [57]:

$$\begin{array}{lllllllll} \mathrm{B_d}: & \Delta M_\mathrm{d} & = & 0.502 \cdot 10^{12}\,\mathrm{s}^{-1}\,, & \Delta\Gamma_\mathrm{d} & = & 0\,\mathrm{s}^{-1}\,, & \tau_\mathrm{d} & = & 1.536 \cdot 10^{-12}\,\mathrm{s}\,. \\ \mathrm{B_s}: & \Delta M_\mathrm{s} & = & 20 \cdot 10^{12}\,\mathrm{s}^{-1}\,, & \Delta\Gamma_\mathrm{s} & = & 6.852 \cdot 10^{10}\,\mathrm{s}^{-1}\,, & \tau_\mathrm{s} & = & 1.461 \cdot 10^{-12}\,\mathrm{s}\,. \end{array}$$

The decay rates for the $\mathrm{B_s} \to \mathrm{B_s}$ and $\mathrm{B_s} \to \overline{\mathrm{B}}_\mathrm{s}$ oscillations as generated in the simulation, together with the corresponding flavor asymmetry as a function of the true generated proper time are shown in Figure 3.7. The plots are for $\sim$ 500 k true $\mathrm{B_s} \to \mathrm{D_s}\pi$ events, with all final states reconstructed as long tracks and without any selection cuts (cheated selection).

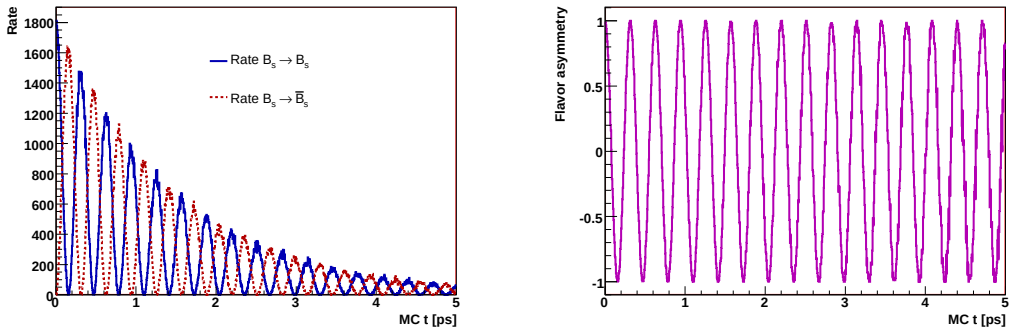


Figure 3.7: Generated decay rates in $\mathrm{B_s} \to \mathrm{D_s}\pi$ events for the $\mathrm{B_s} \to \mathrm{B_s}$ and $\mathrm{B_s} \to \overline{\mathrm{B}}_\mathrm{s}$ oscillations (left) and the flavor asymmetry (right), versus the true proper time [ps]. The events have all the signal final states reconstructed as off-line long tracks and without any selection bias.

3.3 Off-line Track Reconstruction

In a hadron machine environment the reconstruction of tracks is an arduous enterprise, given the large number of tracks originating from the primary interaction ($\sim 10^2$). The track reconstruction aims at finding all possible tracks, and not only the signal decay products. For instance, the reconstruction of primary vertices is vital for the determination of the decay flight of b hadrons. A good accuracy of the track measurements (momentum, impact parameter, errors) is furthermore mandatory to achieve the required sensitivity for LHCb's physics goals.

The strategy for the off-line track reconstruction can be divided in two steps:

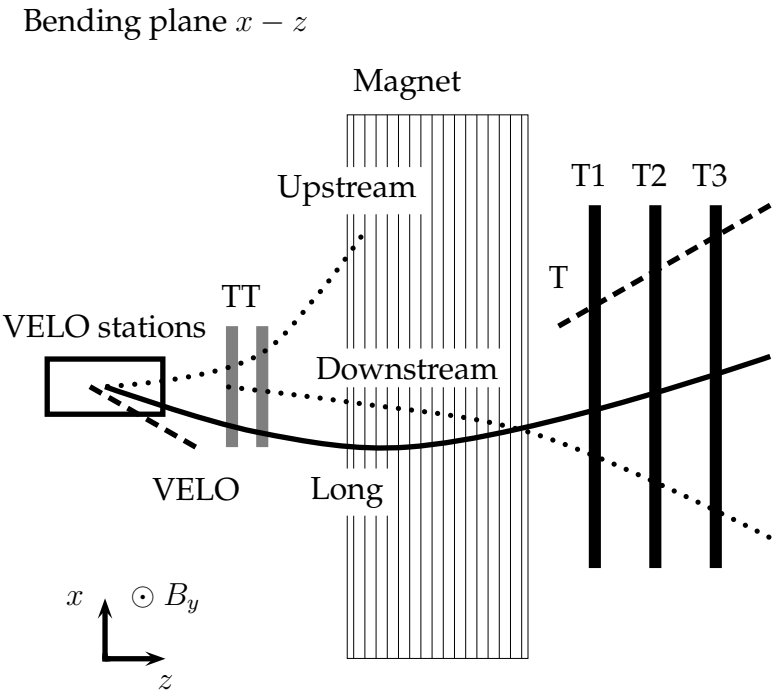


Figure 3.8: A schematic illustration in the bending plane of the five track types: VELO, upstream, downstream, T and long tracks. The illustration is not to scale.

- **Pattern recognition:** efficiently assign the correct clusters to the tracks while keeping a low ghost rate (combinatorics or noise).
- **Track fitting:** accurately determine the track parameters and their errors.

We present in this section the off-line pattern recognition and fitting strategies, and give a few aspects of their performance. Further details can be found in [17], and an excellent and extensive review of the track reconstruction is presented in [84].

3.3.1 Track Types

The charged tracks have different trajectories in the detector depending on their origin and momentum, and they can be classified in the following types, illustrated in Figure 3.8:

VELO tracks: only traverse the VELO and are typically large polar angle or backward tracks. They allow an accurate determination of the primary vertices.

Upstream tracks: traverse only the VELO and the TT stations. They are generally low momentum tracks bent out of the acceptance by the magnet before reaching the T stations. Their poor momentum resolution makes them seldom used in physics analyses.

Downstream tracks: traverse only the TT and T stations, without any clusters in the VELO. They are mainly used for the reconstruction of particles decaying outside the VELO acceptance (e.g. K^0_S, Λ).

T tracks: are only measured in the T stations, typically produced in secondary interactions. They are useful in the RICH2 reconstruction.

Long tracks: traverse the full detector setup, from the VELO down to the T stations. They have accurate track parameters and are thus the most important set of tracks for the selection of b-hadron decays.

3.3.2 Pattern Recognition and Track Fitting

The first step in track reconstruction is to build track segments in the different tracking detectors. Then the track segments from each tracking detector are matched with each other or used as seeds for the tracks' extensions. After the pattern recognition, the trajectories of the tracks are refined by applying an iterative fit.

The track finding procedure consists of the following algorithms, each producing (segments of) tracks with a collection of measurements from the tracking detectors, i.e. VELO, TT and T stations:

VELO tracking In the VELO region the magnetic field is low enough, see Figure 2.4, so that tracks can be approximated as straight lines. The tracks are reconstructed using the radial r and azimuthal φ coordinates of the VELO clusters, combined in triplets. A VELO seed provides an initial track segment for other track finding algorithms, without any momentum information. The same VELO tracking will be used in the on-line and off-line environments [85].

Forward tracking This first method of reconstructing long tracks [86] is also used in the High-Level Trigger [87]. The forward tracking starts with VELO segments and tries to extend the track using a cluster in the T stations, since the whole trajectory can be determined from this information, when neglecting multiple scattering. Using a parameterization of the trajectory constraining the track parameters, clusters in the others tracking stations are searched for. The fraction of long tracks reconstructed using this method amounts to about 90%. Clusters used in this method are ignored by the subsequent tracking algorithms in order to save processing time.

T seeding T segments are reconstructed from the IT and OT measurements in a standalone algorithm [88]. First solutions are looked for in the bending plane using x measurements, assuming straight trajectories as the field is moderate in this region. In a second pass the trajectory is parameterized as a parabola to account for the fringe field. Finally, the stereo information is added to confirm the tracks.

Track matching This second method of reconstructing long tracks [84] starts from the T seeds and extrapolates the track states to the VELO region using a Runge–Kutta method, where a match with a VELO seed is looked for. This method uses the momentum-kick of a track assuming it originates from the interaction point. A χ^2 criterion is used to select the candidates, and TT clusters are searched for and assigned to the track. This method reconstructs an additional 5% of the long tracks.

Upstream and downstream tracking At this stage all already used VELO and T clusters are discarded. Upstream tracks are obtained by matching VELO seeds to at least three clusters in the TT stations. For downstream tracks, the search starts from the T seeds with an initial momentum estimate and tries to match them to clusters in TT.

VELO and T track search The remaining of the VELO and T seeds that have not been used so far are made as VELO and T tracks, with no or a poor momentum resolution, respectively.

Note that the off-line tracking has many similarities with the on-line tracking that will be described in Chapter 4. Ultimately, both pattern recognitions should run the same algorithms, without the refinements of the off-line for the HLT use. This is already the case of the VELO tracking, and moreover the on-line forward tracking is being improved to be able to ignore the TT clusters in its first pass, which is presently required for timing reasons. Finally, the off-line VELO-TT tracking should also be merged with its on-line equivalent.

The average number of off-line reconstructed tracks in a $b\bar{b}$ event is ~ 101: 34 VELO, 6 upstream, 14 downstream, 19 T, and 26 long tracks. The latter type is the most important for physics studies. An example of a reconstructed event is displayed in Figure 3.9.

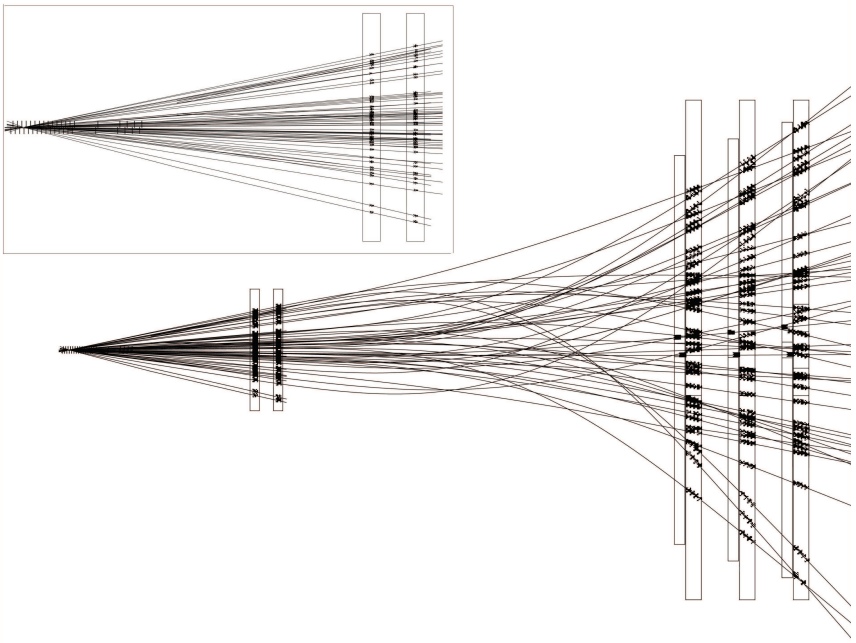


Figure 3.9: Display of the reconstructed tracks and corresponding assigned clusters in a busy event [17]. This event contains 50% more hits than an average $b\bar{b}$ event.

Having found the tracks, the trajectories are refitted with a Kalman filter. The trajectory of a particle depends on six initial values and can be described with a collection of track states depending on five track parameters at a reference plane. The natural choice of parameterization in LHCb is along the beam (z coordinate), given its forward geometry. The trajectories are therefore represented as a collection of track states at different z, together with their covariance matrices, called state vectors. The purpose of the Kalman filter is to update every state vector at each measurement plane starting from the most downstream state and taking into account the interactions with the material, such as energy losses or multiple scattering. At the most upstream measurement, the fit reverses direction to update all the states with the full information.

Important criteria for the performance of the tracking are the pull distributions of the track parameters, i.e. their residuals weighted by the corresponding errors obtained from

the track fits. As an example, the errors on the track parameters for the long tracks used in the $B_s \to \eta_c \phi$ off-line selection are overestimated by $3 - 11\%$ for the positions and slopes, whereas the momentum errors are too small by 19%. The latter discrepancy is mainly due to multiple scattering effects.

3.3.3 Long Tracks Performance

The following quantities characterize the off-line tracking performance:

1. The track finding efficiencies.
2. The ghost rates.
3. The quality of the track parameters (positions and pulls).

For the first items, we need to define the set of tracks used to assess the performance. To this end we can make use of the MC truth and we define the conditions for a track to be *reconstructible* as long:

- the particle must give at least 3 r and 3 φ hits in the VELO (reconstructible as a VELO track);
- the particle must have at least 1 x and 1 stereo hit in each of the T stations (reconstructible as a T track).

Note that *no* requirement on the TT hits is demanded. A reconstructible long track is then defined as successfully *reconstructed* if both the VELO and T tracks are associated to the same MC particle, with each at least 70% of their clusters originating from this MC particle. A ghost is a reconstructed track failing these association criteria. The track finding efficiency is defined as the fraction of reconstructible particles that are successfully reconstructed, and the ghost rate is defined as the fraction of reconstructed tracks found that are ghosts.

The long track finding efficiency depends on the momentum of the particles, as shown in Figure 3.10. For tracks above 10 GeV/c, the efficiency is about 94%. The ghost rate for interesting tracks, i.e. with a transverse momentum larger than 300 MeV/c, is approximatively 3%.

The momentum resolution and the impact parameter resolution are measures of the accuracy of the track parameters. The momentum resolution for b tracks is $\sim 0.3\%$ on average. The impact parameter resolution exhibits a linear dependence on the inverse of the transverse momentum ($1/p_{\mathrm{T}}$) as illustrated in Figure 3.10. This dependence is fully exploited in the High-Level Trigger to parameterize the on-line tracking errors.

The tracking performance will be detailed for a specific decay in Chapter 5, when discussing the $B_s \to \eta_c \phi$ selection.

3.4 Flavor Tagging

Flavor tagging is the identification of the initial flavor of the reconstructed neutral B meson, namely if it originally contained a b or a $\bar{\mathrm{b}}$ quark. The knowledge of the original flavor at production is needed, though not always required, in order to study CP asymmetries and flavor oscillations.

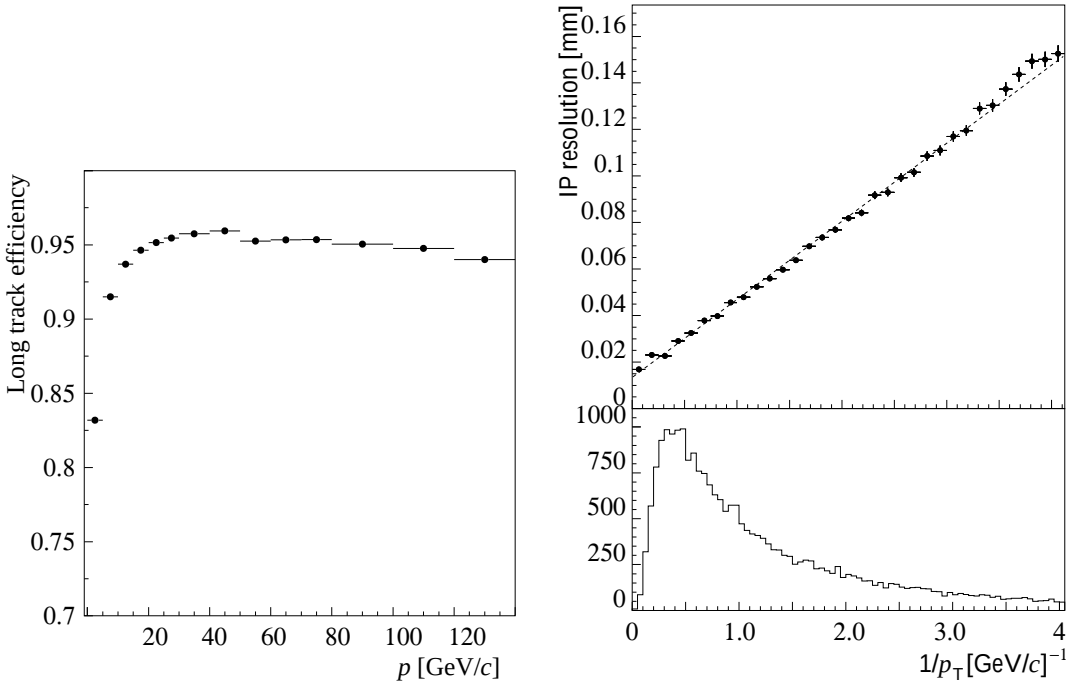


Figure 3.10: Left plot: long track finding efficiency versus the true momentum of the generated particle [GeV/c]. Right plot: impact parameter resolution [mm] of long tracks at the production vertex of the track as a function of $1/p_{\mathrm{T}}$ [$(\mathrm{GeV/c})^{-1}$]. Both plots are from [17].

The output from the flavor tagging algorithms (or tag) is three fold: the flavor of the B meson was b, or $\bar{\mathrm{b}}$, or unknown (i.e. untagged). The performance of the tagging is represented by the following quantities:

- The tagging efficiency $\varepsilon_{\mathrm{tag}}$: probability that the tagging procedure gives an answer.
- The wrong-tag fraction ω_{tag}: probability for the answer to be incorrect when a tag is present (the wrong tag is often also denoted by the letter w).

Given the numbers N_{R}, N_{W}, and N_{U} of correctly tagged, wrongly tagged, and untagged candidates respectively, the probabilities $\varepsilon_{\mathrm{tag}}$ and ω_{tag} are defined as:

$$\varepsilon_{\mathrm{tag}} = \frac{N_{\mathrm{R}} + N_{\mathrm{W}}}{N_{\mathrm{R}} + N_{\mathrm{W}} + N_{\mathrm{U}}}\,, \quad \omega_{\mathrm{tag}} = \frac{N_{\mathrm{W}}}{N_{\mathrm{R}} + N_{\mathrm{W}}}\,. \tag{3.1}$$

The effect of the wrong-tag fraction is to dilute the CP and flavor asymmetries, thus reducing the amplitude of the $\mathrm{B_q}$–$\overline{\mathrm{B}}_{\mathrm{q}}$ oscillations. To illustrate this, we can consider the decay of a $\mathrm{B_q}$ meson into a self-conjugated final state $f = \overline{f}$ (decay $\mathrm{B_q} \to f$) and the charge-conjugate decay (decay $\overline{\mathrm{B}}_{\mathrm{q}} \to f$). Introducing the theoretical asymmetry $\mathcal{A}_{\mathrm{th}}$:

$$\mathcal{A}_{\mathrm{th}}(t) \;=\; \frac{\Gamma\left(\overline{\mathrm{B}}_{\mathrm{q}} \to f\right)(t) - \Gamma\left(\mathrm{B_q} \to f\right)(t)}{\Gamma\left(\overline{\mathrm{B}}_{\mathrm{q}} \to f\right)(t) + \Gamma\left(\mathrm{B_q} \to f\right)(t)}\,,$$

we get that the observed asymmetry $\mathcal{A}_{\mathrm{obs}}$ is diluted:

$$\mathcal{A}_{\mathrm{obs}}(t) \;=\; \frac{R\left(\overline{\mathrm{B}}_{\mathrm{q}} \to f\right)(t) - R\left(\mathrm{B_q} \to f\right)(t)}{R\left(\overline{\mathrm{B}}_{\mathrm{q}} \to f\right)(t) + R\left(\mathrm{B_q} \to f\right)(t)} = D \times \mathcal{A}_{\mathrm{th}}(t)\,, \tag{3.2}$$

where $D = (1-2\omega_{\text{tag}})$ is the dilution factor. We see that the observed asymmetry vanishes when the wrong tag is 50% or when there is no tag. The flavor tagging is therefore extremely important in the study of these asymmetries. However, a vanishing asymmetry does *not* mean that untagged events are useless. Indeed, instead of fitting the asymmetry directly, a more judicious approach is to look at the decay rates. In this case, we can gain in sensitivity for the mixing parameters. As we will see in Chapter 6, untagged events in $\bar{\text{b}} \to \bar{\text{c}}\text{c}\bar{\text{s}}$ transitions do give access to both $\Delta\Gamma_{\text{s}}$ and ϕ_{s} (but not to ΔM_{s}). Note that the only advantage in fitting the asymmetry would be that the proper time acceptance cancels out at first order.

We define the effective tagging efficiency ε_{eff} or simply tagging power as:

$$\varepsilon_{\text{eff}} \;=\; \varepsilon_{\text{tag}} \times (1-2\omega_{\text{tag}})^2 = \varepsilon_{\text{tag}} \times D^2 \,. \tag{3.3}$$

From (3.2) and propagating the asymmetry errors quadratically we get the statistical uncertainties of the observed and analytical asymmetries, respectively $\sigma_{\mathcal{A}_{\text{obs}}}$ and $\sigma_{\mathcal{A}_{\text{th}}}$, given by:

$$\begin{aligned}
\sigma^2_{\mathcal{A}_{\text{obs}}} &= D^2 \times \sigma^2_{\mathcal{A}_{\text{th}}} \,, \\
\sigma^2_{\mathcal{A}_{\text{obs}}} &= \frac{1-{\mathcal{A}_{\text{obs}}}^2}{\varepsilon_{\text{tag}} N_{\text{phys}}} \,, \\
\sigma^2_{\mathcal{A}_{\text{th}}} &= \frac{1-{\mathcal{A}_{\text{obs}}}^2}{(\varepsilon_{\text{tag}} \times D^2) N_{\text{phys}}} \,,
\end{aligned} \tag{3.4}$$

where N_{phys} is the number of reconstructed B_{q} candidates. We thus get:

$$\begin{aligned}
\sigma_{\mathcal{A}_{\text{obs}}} &\propto \frac{1}{\sqrt{\varepsilon_{\text{tag}} N_{\text{phys}}}} \,, \\
\sigma_{\mathcal{A}_{\text{th}}} &\propto \frac{1}{\sqrt{\varepsilon_{\text{eff}} N_{\text{phys}}}} \,.
\end{aligned}$$

The effective tagging efficiency ε_{eff} is the quantity to maximize in order to minimize the statistical error on the corrected asymmetry. The tagging power combines the knowledge of ε_{tag} and ω_{tag} in one parameter to be optimized in the tagging procedure.

Table 3.2: Tagging performance for a few decay channels, after off-line selection and High-Level Trigger. The performance for the B_{d} channels is taken from [89].

Decay	ε_{tag} [%]	ω_{tag} [%]	ε_{eff} [%]
$\text{B}_{\text{d}} \to \pi^+\pi^-$	51.0 ± 0.2	35.0 ± 0.3	4.5 ± 0.2
$\text{B}_{\text{d}} \to \text{J}/\psi\text{K}^0_{\text{S}}$	50.6 ± 0.1	33.9 ± 0.1	5.2 ± 0.1
$\text{B}_{\text{s}} \to \eta_{\text{c}}\phi$	65.7 ± 1.7	30.9 ± 2.0	9.6 ± 2.9
$\text{B}_{\text{s}} \to \text{D}_{\text{s}}\pi$	62.8 ± 0.2	31.0 ± 0.3	9.1 ± 0.3
$\text{B}_{\text{s}} \to \text{J}/\psi\phi$	57.0 ± 0.2	33.0 ± 0.2	6.6 ± 0.2

We give in Table 3.2 the tagging performance for a few channels, after off-line selection and the full trigger chain. The tagging power is much better for B_{s} mesons than for B_{d}

mesons because of the same-side kaon tag. The performance for $B_s \to J/\psi\phi$ is worse than that for hadronic B_s decays, due to the presence of muons in the final states. The correlations with the trigger, mainly the muon streams, reduce the performance for $B_s \to J/\psi\phi$ since we mainly trigger on the signal. The $b\overline{b}$ correlations induce a softer opposite B_q meson which worsens the tagging performance.

3.4.1 Tagging Categories

The tagging information for a B_q candidate must be extracted from the rest of the event. After the fragmentation of the $b\overline{b}$ pair and the b-quark hadronization, the newly formed hadrons evolve and decay. Based on the signature of these decays, it is possible to extract the nature of the signal B_q at production.

The object carrying the tagging information is called the tagger. As it will be presented below, there are several tagging categories, each yielding a decision. The combination of the different tags is based on a neural network approach, combining each non-exclusive categories into exclusive probability categories. Several ideas for new taggers are currently under study; we will only present here the ones used in the simulations of Chapters 5 and 6. The different tagging methods, illustrated in Figure 3.11, are [90]:

Opposite-side tagging This method determines the flavor of the b hadron accompanying the signal B_q meson. There is an additional intrinsic dilution when the tagging b hadron is a neutral meson since it can oscillate.

- **Opposite-side lepton tag:** the charge of the lepton from a semi-leptonic b-hadron decay is used.
- **Opposite-side kaon tag:** the charge of the kaon from the $b \to c \to s$ decay chain is used.
- **Vertex charge tag:** the charge of the inclusively reconstructed decay vertex of the accompanying b hadron is used.

Same-side tagging This method directly determines the flavor of the signal B_q meson exploiting the correlation in the fragmentation process.

- **Same-side kaon tag:** if a $B_s = \bar{b}s$ is produced in the fragmentation of a $\bar{b}$ quark, an extra $\bar{s}$ is available to form a kaon, which is charged in $\sim 50\%$ of the cases. The positive charge of the kaon reveals the flavor of the B_s. The tagging kaons emerge from the same primary vertex as the B_s and are correlated in phase space with it.
- **Same-side pion tag:** the same method as for the same-side kaon tag can be applied to the B_d, though the large pion combinatorics spoil the performance. B resonances, e.g. $B^{**\pm} \to B^{*0}\pi^{\pm}$, also provide same-side pion tags.

All these methods have natural limitations due to the production fractions of the different b-hadron species, and the branching ratios of b hadrons to kaons or leptons. Moreover, the mixing of neutral B_q mesons, wrong-sign kaons, leptons from charm decays and background tracks all tend to increase the wrong-tag fraction.

There are also systematic errors that can spoil the performance of the flavor tagging, as well as that of any CP measurement. For instance a charge-dependent detection efficiency, or trigger and off-line selection biases. Whereas the first effect can be studied

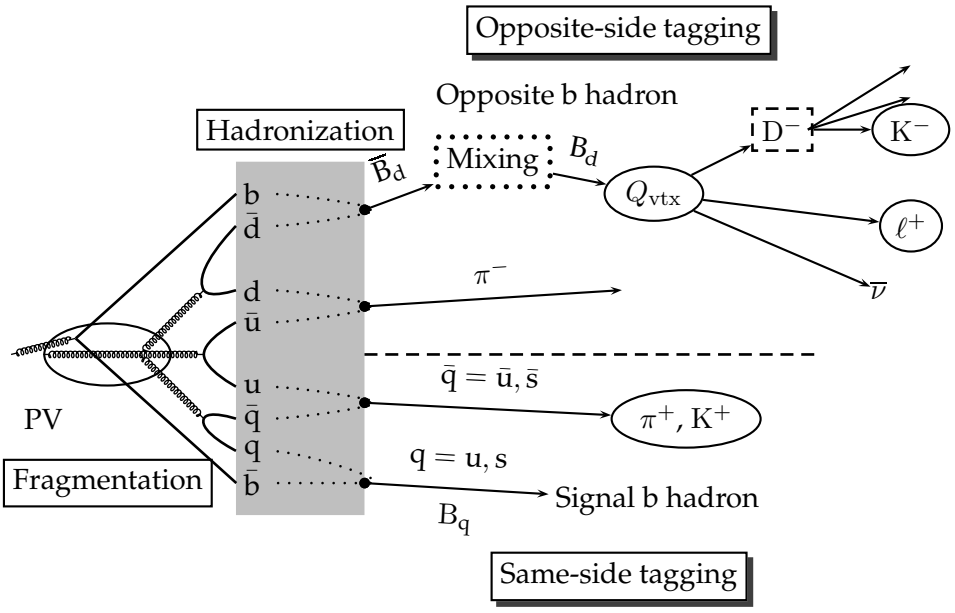


Figure 3.11: Cartoon representation of the flavor tagging procedure, showing the two tagging sides. In this example, the opposite B meson first oscillates and then semi-leptonically decays via $B_d \to (D^- \to K^- \ldots)\, \ell^+\, \overline{\nu}$.

by looking at asymmetries in the b sidebands or reversing the polarity of the magnet, the last effect is much more delicate and currently under study (so-called "buffer tampering", "trigger source").

The wrong-tag fraction can be directly extracted from the real data by means of flavor specific decays, e.g. with $B_s \to D_s\pi$, acting as control channels for other CP measurements, like ϕ_s, for which the determination of the wrong tag is not possible. As in general signal and control samples are triggered in a different way, this yields different tagging performances for the two samples, as shown for example in Table 3.2 for $B_s \to D_s\pi$ and $B_s \to J/\psi\phi$. This must be corrected for by equalizing the phase space of the signal and control channels in order to extract the signal wrong-tag fraction. This is again a source of systematic uncertainty, which will be ignored in the study of Chapter 6.

3.5 Data Samples and MC Technicalities

In this section we detail the versions of the programs used in this thesis, the data samples, and we give some additional information and definitions for the HLT studies. All the software and data samples are part of the Data Challenge 04 (DC04). This section being quite technical, it is not required for the understanding of the rest of this dissertation. The reader may thus directly move onto the next chapter.

3.5.1 Software Versions

Both the trigger and off-line selections algorithms were run in DAVINCI `v12r15`, with several updated packages. For all the studies performed on the full MC, this formally

corresponds to DAVINCI `v12r17`, with the exception that `Trg/TrgForward v4r0` was used. The most relevant packages are:

- `Hlt/HltSelections v4r1p1`, `Hlt/HltSelChecker v4r1` and `Hlt/HltInclusive v1r1` for the HLT studies of Chapter 4;
- `PhysSel/Bs2PhiEtac v3r8` and LOKI `v3r14p1` for the $B_s \to \eta_c \phi$ selection presented in Chapter 5;
- All analysis n-tuples (HLT and off-line) for the study of the selections were generated with `Phys/DaVinciMCTools v5r18`, using the generic algorithm `DecayChainNTuple` [69].

3.5.2 Data Samples

We give hereafter the data samples used in the trigger studies of Chapter 4 and in the selection and sensitivity studies of Chapters 5 and 6. The description of the EVTGEN decay files can be found in [72], with the numbering scheme of [91].

Trigger studies samples

- Minimum-bias DC04-v1: 32'667 stripped events (with an obsolete DAVINCI version) at 40 kHz. This sample was used to tune all the selection criteria, however all the results (e.g. rates, plots) presented in Chapter 4 are based on the DC04-v2 sample.
- Minimum-bias DC04-v2: 131'320 L0 and L1 stripped events (done with DAVINCI v12r15) from an original sample of 48'254'400 events.The stripped sample corresponds to ~ 3.28 s of LHCb running. This is the minimum-bias sample used to assess the HLT performance presented in Chapter 4. Applying the generic HLT to this sample leaves 28'458 events which is equivalent to an output rate of (8668.3 ± 45.5) Hz.
- Signal DC04-v1/-v2 data: several signal types, after the L0 and L1 triggers, and the corresponding specific off-line selection (here "TDR selections" were applied, see [17]). The description of the data samples used can be found in [78].

Off-line studies samples

- DC04-v1 450'501 events of event type: 13336000.
- DC04-v2r3 998'500 events of event type: 11900000.
- DC04-v1 607'500 events of event type: 11902000.
- DC04-v2r3 50'000 events of event type: 13491400.
- DC04-v1 383'000 events of event type: 13296000.
- DC04-v1 170'000 events of event type: 13296200.
- DC04-v1 326'000 events of event type: 11296000.

- DC04-v2r3 668′500 events of event type: 13266000.
- DC04-v2r3 9′826′209 events of event type: 23263000.
- DC04-v2r3 3′990′568 events of event type: 13264000.
- DC04-v1 1′659′000 events of event type: 13144000.
- DC04-v1 18′060′757 stripped $b\overline{b}$ events (used to tune the $B_s \to \eta_c\phi$ selections cuts).
- DC04-v2 27′291′931 stripped $b\overline{b}$ events (used to determine the background levels).

MC random seed

The random seed used in the MC simulation is generated as a 32-bit number. The first round of the generation, DC04-v1, only used the 24 least significant bits of the seed, corresponding to a maximum of $2^{24} = 16'777'216$ events if all the seeds are used at least once. Assuming that the fraction of independent events is the *same* in every subsample of the initial sample, then the number of independent events N_{ind} and the corresponding error $\sigma_{N_{\text{ind}}}$ are given by:

$$\begin{aligned} N_{\text{ind}} &= M - M \times \left(1 - \frac{1}{M}\right)^N \approx M \times \left(1 - e^{-\frac{N}{M}}\right) , \\ \sigma^2_{N_{\text{ind}}} &= M \times (M-1) \times \left(1 - \frac{2}{M}\right)^N + M \times \left(1 - \frac{1}{M}\right)^N - M^2 \times \left(1 - \frac{1}{M}\right)^{2N} \\ &\approx N_{\text{ind}} \times e^{-\frac{N}{M}} , \end{aligned} \tag{3.5}$$

where $M = 2^{24}$ and N is the size of the initial sample. This bug affects all the DC04-v1 samples, whereas for the rest this was fixed using effectively $2^{32} = 4'294'967'296$ independent seeds.

The different efficiencies are not affected, while their errors have to be corrected, for the DC04-v1 data. The binomial errors will thus be scaled with the (inverse of the square root of the) fraction of independent events when computing efficiencies errors for instance in Chapter 5, and ignoring $\sigma_{N_{\text{ind}}}$. The number and the fraction of independent (f_{ind}) events for the $B_s \to \eta_c\phi$ and $B_s \to J/\psi\phi$ DC04-v1 samples are:

- $B_s \to \eta_c\phi$ ($450'501$ generated events): $N_{\text{ind}} \sim 444'506$, $f_{\text{ind}} \sim 98.6\%$;
- $B_s \to J/\psi\phi$ ($1'659'000$ generated events): $N_{\text{ind}} \sim 1'579'610$, $f_{\text{ind}} \sim 95.2\%$.

3.5.3 HLT Technicalities

Timing measurement

The time measurements are performed with the equivalent of a 2.8 GHz Xeon processor by means of the `SequencerTimerTool` used by the `GaudiSequencer`. Two different measures of the time are used in Chapter 4:

CPU time It gives the CPU cycles used by the process with a millisecond sampling, independently of the machine load.

Clock time It measures the elapsed time between the beginning and the end of an algorithm. This time is affected by the load of the machine. In order to have a reliable time measurement of the full HLT sequence, a "clean" replica of a lxbatch node with a 2.8 GHz Xeon processor was used (pclbonsrv05), only running the HLT code.

All the timing estimates are obtained on minimum-bias events and applying the triggers prior to the HLT.

Efficiency measurement

The criteria to associate an on-line track depend on the part of the track we consider, which translate to the track types as:

- A requirement of 70% of the VELO measurements to be associated to the same MC particle is asked for a track to be reconstructed as a VELO track.
- A VELO-TT track is associated if its VELO part is and if the TT part has at most one wrong measurement. A track accepted by the VELO-TT algorithm without any momentum is counted as reconstructed even if there are no TT clusters attached to it, provided the VELO part is correctly reconstructed.
- For long tracks, 70% of the measurements for both and independently the VELO and the T stations (inner and outer trackers) need to be associated to the same MC particle and the TT measurements are ignored. In particular a track could be reconstructed as long even if the TT part has more than one wrong cluster.
- A track is a ghost if it is not associated to any MC particle. Ghosts are reconstructed tracks made of clusters from different MC particles, and/or from noise.

The details of the associators used in the HLT studies are given in [78].

The tracking efficiencies are computed on events where all the charged signal final states are off-line reconstructed as long tracks. The tracking efficiencies are defined by:

- Per-track efficiency: $\varepsilon^{\rm tr}_{\rm rec} = N^{\rm tr}_{\rm rec}/(N^{\rm tr}_{\rm fin} \times N^{\rm evt}_{\rm off})$;
- Per-event efficiency: $\varepsilon^{\rm evt}_{\rm rec} = N^{\rm rec}_{\rm evt}/N^{\rm evt}_{\rm off}$;

with the following definitions:

- $N^{\rm tr}_{\rm fin}$: number of charged final states in the decay;
- $N^{\rm evt}_{\rm off}$: number of events after generic HLT and off-line selections, for events with all signal final states off-line reconstructed;
- $N^{\rm tr}_{\rm rec}$: number of on-line reconstructed tracks;
- $N^{\rm rec}_{\rm evt}$: number of events with *all* final states on-line reconstructed.

Note that for the per-event efficiency $\varepsilon^{\rm evt}_{\rm rec}$ we require all the final states to be on-line reconstructed and hence it is not determined as $(\varepsilon^{\rm tr}_{\rm rec})^{N^{\rm tr}_{\rm fin}}$. The ghost rates are determined on minimum-bias events, after the generic HLT selection.

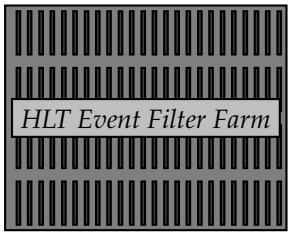
The HLT selection efficiencies on off-line selected signal events are computed after the generic HLT, independently if the off-line selected event contains a correctly reconstructed candidate. The selection efficiencies include the track finding efficiencies. In particular, we will consider the following definitions:

- $\varepsilon_{\mathrm{HLT}}^{\mathrm{specific}}$: the efficiency for a given decay with the dedicated specific HLT selection;
- $\varepsilon_{\mathrm{HLT}}^{\mathrm{exb}}$: the HLT efficiency of the exclusive b stream;
- $\varepsilon_{\mathrm{HLT}}^{\mathrm{b}\to\mu}$: the HLT efficiency of the inclusive b $\to \mu$ single muon stream;
- $\varepsilon_{\mathrm{HLT}}^{\mu\mu}$: the HLT efficiency of the di-muon stream;
- $\varepsilon_{\mathrm{HLT}}^{\mathrm{D}^*}$: the HLT efficiency of the D^* stream;
- $\varepsilon_{\mathrm{HLT}}^{\mathrm{tot}}$: the total HLT efficiency of all HLT streams.

A measure of the total track reconstruction efficiency is given by the cheated HLT selection efficiency $\varepsilon_{\mathrm{HLT}}^{\mathrm{cheated}}$, which is equivalent to running the HLT selections without any cut and only with the tracks associated to the signal.

Chapter 4

Exclusive Trigger Selections



This chapter describes the performance of the exclusive High-Level Trigger (HLT) selections for the exclusive b and D streams. A comparison between the on-line and off-line reconstructions is made and the HLT efficiencies for the core physics channels are given, as well as the corresponding minimum-bias output rates and timing performance. Different sources of inefficiency are discussed and identified in oder to motivate future improvements.*

THE High-Level trigger (HLT) will have an output rate of ~ 2 kHz and it is divided in four streams, as introduced in Section 2.2.8:

- Exclusive b (~ 200 Hz);
- D^* (~ 300 Hz);
- Di-muon (~ 600 Hz);
- Inclusive b $\rightarrow \mu$ (~ 900 Hz).

Each of these streams will irrespectively of each other's decision send their selected events to storage. In this thesis we assume that the entire HLT chain, on-line reconstruction and selection, will be running on the on-line farm at a rate of 40 kHz and using about 400 CPUs in the initial strategy with a L1 trigger. The average HLT processing time should thus not exceed 10 ms per event. The change of scheme to the 1 MHz readout, yet to be implemented, will nevertheless not affect much the exclusive b and D^* streams as they will anyway be executed after pre-HLT algorithms.

The exclusive b HLT should efficiently select all final states relevant to the LHCb physics program. For the first implementation a list of ten core channels has been defined:

1. $B_s \rightarrow D_s h$
2. $B_d \rightarrow J/\psi K^0_S$
3. $B_s \rightarrow J/\psi\phi$
4. $B_q \rightarrow hh$

5. $B_d \to D^0 K^{*0}$
6. $B_s \to \phi\phi$
7. $B_d \to D^* \pi$
8. $B_d \to \mu^+\mu^- K^{*0}$
9. $B_s \to \mu^+\mu^-$
10. $B_s \to \phi\gamma$

The above list is meant to evolve and many more channels could be added. For instance new HLT selections involving electrons (e.g. $B \to K^{(*)}ee$) or π^0 ($B \to \rho\pi$) or K^0_S ($B \to \phi K^0_S$). Also the use of additional inclusive streams (ϕ, D_s, ...) is under study.

The present chapter describes the performance of the HLT exclusive b selections and the D^* stream, based on the design and implementation choices of [70], in the framework described in [69]. This chapter is an extract from the results of [78], with additional information related to the on-line tracking, especially for the VELO-TT pattern recognition.

4.1 On-line and HLT Environment

The design and the implementation choices of the HLT selections are described in [70], based on the HLT implementation for the Data Challenge 04 (DC04) compatible software. The picture for the HLT selections will not drastically change with the 1 MHz readout scheme and the additional data available at this higher rate. However, many improvements in terms of tracking strategy and performance are expected as well as a new design for the generic HLT, in particular its merge with the present L1 trigger, hence removing the current 40 kHz boundary. In the new design alleys are foreseen, i.e. the confirmation of previously found triggering objects, with as spin-off an even more "vertical" trigger between the L0 and the generic HLT lines. The 1 MHz scheme also avoids the repetition of identical tasks, such as the decoding of VELO or TT buffers.

There should be no conceptual difference in the way any off-line or HLT selections are performed: both work with particles provided as inputs, which are then combined together to form composite particles and applying several selection criteria until a whole decay chain is reconstructed. The HLT selections therefore use for both the exclusive b and D^* streams the same generic algorithms and tools allowing the complete selection of events steered by options [69]. The use of common tools throughout the HLT selections ensures maximal correlations with a limited amount of code used. This implementation enables the quick addition or removal of streams (e.g. inclusive ϕ stream) from the configuration database.

4.1.1 From Tracks to Particles

The first stage in the HLT processing is the on-line reconstruction [87] in parallel with the generic HLT [92]. Note that there were some improvements with respect to the on-line pattern recognition described in [87], and they are described in the text.

The main part of the pattern recognition is actually done in the generic HLT sequence. The full tracking is first run on a limited set of tracks without performing the rest of the reconstruction unless the event is accepted by the generic HLT trigger. The combined

tracking and generic HLT procedure is briefly described hereafter. Each tracking step implicitly includes a data preparation and raw buffer decoding.

VELO tracking and primary vertex search

VELO-RZ tracking [87, 85]:
VELO-RZ tracks are created from 2D tracking in the $r - z$ projection using aligned r-cluster triplets in consecutive r sensors.

VELO-space tracking [87, 85]:
VELO-space tracks are reconstructed by extending to 3D the VELO-RZ tracks using information from the VELO φ sensors.

Primary vertex reconstruction [93]:
The primary vertex search has largely been improved and consequently the fraction of non-reconstructed primary vertices identified as secondary b vertices is better under control. The new implementation is a two-step procedure using VELO-space tracks where first the vertex seeds are looked for, and second the iterative vertex fit is performed. The fit now can merge (split) close (large) primary vertices to avoid the misidentification of primary and secondary vertices. The borderline between a primary and a secondary vertex is currently set by requiring a minimum of six tracks to form a primary vertex.

Generic HLT and full tracking

Stand-alone muon pattern recognition [94]:
The muon identification has a dedicated pattern recognition in the HLT. Muon track segments are first looked for using information from the muon stations M5 to M2 and then matched to VELO tracks. Finally, the differences in x and y at M2 between the track extrapolation and the muon segments are compared to identify muons. Using this approach the muon reconstruction for the HLT does not directly depend on the standard on-line pattern recognition, and the muons can be reconstructed very fast.

VELO-TT tracking [95]:
The matching of VELO-space tracks to TT clusters is done assuming that the effect of the magnetic field can be replaced by a kick of the track at a fixed z. All VELO tracks are then extrapolated to the z position of the measurements and the distances perpendicular to the strips between the extrapolation and the measurement are projected onto a reference plane. The solutions are looked for among the accumulations in the projection plane. When the track extrapolation is outside TT sensitive area (e.g. beam pipe, TT gaps) and under some special conditions, VELO tracks might be passed directly to the forward tracking without any TT clusters (or momentum). The VELO-TT tracking will be detailed in Section 4.2.1.

Forward tracking [87]:
The long tracks search is performed by matching VELO-space tracks to clusters in the T stations (inner and outer trackers), using the momentum estimate (if any) obtained from VELO-TT tracks to possibly restrict the search window for the T clusters, at a cost of some inefficiencies due to wrong TT clusters or to the acceptance.

Using a seed in the T stations, the trajectory is parameterized depending on the track parameters, deflection and the field. All measurements compatible with this trajectory are collected and an iterative procedure removes the worse clusters along the trajectory.

Track errors parameterization [96]:
The errors from the tracking are not used and the covariance matrix of all long tracks is replaced by a simplified and parameterized version as a function of the transverse momentum. The upstream tracks errors are not overwritten.

HLT muon streams [92]:
A single muon and di-muons are searched for the HLT decision of the inclusive $b \to \mu$ single muon ($p_T(\mu) > 3$GeV/c, sIPS(μ) > 3) and di-muon ($m_{\mu\mu} > 2.5$GeV/c^2, $\chi^2_{\mu\mu} < 5$, IP$_{\mu\mu} > 0.1$ mm) streams.

Generic HLT [92]:
The generic HLT redoes the L1 generic line with better momentum determination. The decision is based on high transverse momentum. The event is accepted if the logarithmic sum of the transverse momenta (in MeV/c) is $\sum \ln p_T > 14.7$ for the two largest p_T tracks satisfying $0.1 <$ IP < 3 mm. The decision is also positive if either of the muon lines is triggered.

This first part of the HLT processing is followed by the creation of particles to be later used by the HLT selections. Di-muons are directly retrieved from the generic HLT, and all the upstream and forward tracks are made as both pions and kaons. Finally, photons are made from on-line electromagnetic calorimeter (ECAL) clusters.

The possibility of using particle identification provided by the on-line RICH is at present not used. The latter allows to save CPU time by reducing combinatorics. However, the lower combinatorics do not compensate for the time lost in building the RICH information.

4.1.2 Dedicated On-line Software

The HLT will be running in the on-line farm with a limited processing time to reconstruct all the necessary information and to decide decision whether or not an event will be sent to storage. The corresponding off-line algorithms, for which an "unlimited" time is available, can therefore not be run. In order to cope with this time restriction, we need fast dedicated on-line software which will unavoidably give inefficiencies when comparing the on-line performance to the off-line one.

We list hereafter the main sources of inefficiency or differences that were identified for the HLT selections:

- On-line and off-line pattern recognitions are different, yielding on-line track finding inefficiencies on off-line reconstructed tracks.

- The differences in the primary vertex reconstruction strategy may give events with a different number of on-line and off-line reconstructed primary vertices. Moreover, a missed primary vertex can fake the presence of a secondary vertex.

- The track parameters are not identical to the off-line ones (e.g. worse momentum resolution) and their covariance matrices are not comparable. In the HLT a simplified model for the track uncertainties is used [96], whereas in the off-line they are obtained from the track fitting.

- The HLT geometrical and vertexing tools assume straight-line trajectories for the tracks and use a constant parameterized covariance matrix [97]. The latter yields a few percent inefficiency in the vertexing for the decay products of long-lived particles.

- The choice of selection criteria also degrades the HLT performance. For instance transverse momentum and impact parameter cuts are needed for timing reasons and applied to all long tracks in the HLT, whereas several off-line selections do not apply any or very soft requirements on these variables.

4.1.3 Selection Criteria Definitions

Ideally, in order to ensure a maximal correlation and efficiency with the off-line selections, the HLT selections should mimic what is done in the off-line with looser requirements whenever allowed by the output rate. Since the computing time is limited, we must avoid the proliferation of selection-specific criteria, and so a limited set of simple cuts is used in the HLT. This allows for instance the introduction of shared composite particles with soft requirements to be later reused and refined to form the b-hadron candidates in the specific streams. In this way the reconstruction of the same particle several times per event can be avoided. The selection criteria are applied through a series of standard and reproducible filters, as described in [69].

The list of selection criteria used in the HLT is defined hereafter, where some are illustrated in Figure 4.1:

- Momentum p or transverse momentum p_{T} of a particle.

- Impact parameter IP or impact parameter significance IPS $\equiv \mathrm{IP}/\sigma_{\mathrm{IP}} = \sqrt{\chi^2}$, where σ_{IP} denotes the error on IP. Impact parameter cuts are unsigned and always applied with respect to all reconstructed primary vertices, which is equivalent to cutting on the smallest impact parameter significance denoted by sIPS.

- Difference in mass δm between the reconstructed particle and the mass from the `ParticlePropertySvc` (generated nominal Monte Carlo value).

- $\chi^2 \times$ ndf of the mass-unconstrained vertex fit.

- Flight distance FD or flight distance significance FS (ratio of FD to its computed error). The flight distance is defined as the distance between the decay vertex of a composite particle and the reconstructed primary vertices.

- Pointing angle $\theta_{p,\mathrm{F}}$ defined as the angle between the reconstructed momentum $\vec{p}$ and the direction of flight $\vec{\mathrm{F}}$ for particles originating from a primary vertex.

An identified source of background corresponds to events where the primary vertex reconstruction fails to find a true non-elastic primary interaction. A simple way for getting rid of theses cases is to apply vertex isolation cuts [78, 98, 99].

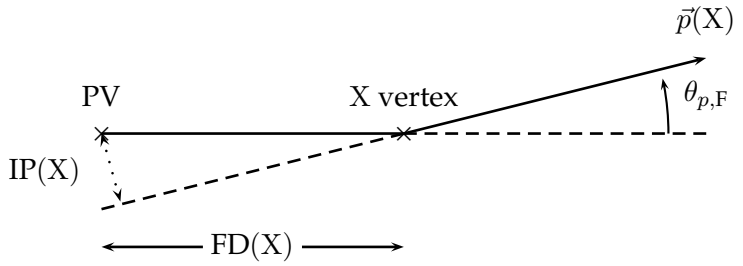


Figure 4.1: Pointing angle ($\theta_{p,\mathrm{F}}$), flight distance (FD) and impact parameter (IP) definitions for a particle X.

4.2 On-line Reconstruction Performance

The performance of the HLT selections strongly relies on the on-line tracking. The main tasks of the reconstruction are:

- Find all necessary tracks for the HLT decision, especially the signal tracks that will be used in an off-line analysis;
- Provide accurate track parameters and reliable errors;
- Find *all* possible primary vertices.

All these items will then be used to reconstruct the largest number of interesting events within the output rate specifications by reducing the minimum-bias rate.

4.2.1 On-line VELO-TT Tracking

The investigation of the HLT inefficiencies has shown that they were dominated by track-finding problems. In particular, the VELO-TT tracking has large inefficiencies because of the lack of constraints one can impose on the TT segments. Indeed, we have only four TT planes to look for compatible TT clusters and match to the numerous VELO-space tracks. Moreover, many tracks are not in TT sensitive area: they are either not in the acceptance (in the central hole or outside the detection planes) or traverse the dead regions (vertical gaps between consecutive bonded sensors). The layout and specifications of the Trigger Tracker are given in Section 2.2.3. Whereas TT clusters are *not* required in the off-line tracking, in the HLT the VELO-TT tracks are the "seeds" of the forward tracking: the momentum estimate obtained from TT is primordial for the speed performance of the long tracking. Based on these observations, and with a compromise between speed and performance, a new strategy for the VELO-TT tracking has been developed for both the L1 and the HLT. This new implementation is an improvement of the VELO-TT tracking described in [87]. The strategy is a bit different for L1 for timing reasons: a single pass is done, and all created tracks have a defined momentum estimate.

The data is first decoded from the raw buffer for the HLT, or from the L1 buffer, as explained in [87]. A filter is applied to ignore clusters without the L1 high-threshold bit (for the L1 version) or with a total ADC signal content less than 6 (high threshold for the HLT version).

The TT clusters (data) are grouped in horizontal bands with a similar vertical range. These horizontal bands, or `TTPlanes`, contain the sensors (wafers) such that the number of sensors in a band is the number of ladders of the corresponding TT layer. As described in Section 2.2.3, TT consists of two stations:

- TTa: X layer with 0° stereo angle (vertical strips) and U layer with -5° stereo angle;
- TTb: V layer with $+5^\circ$ stereo angle and X layer with 0° stereo angle (vertical strips).

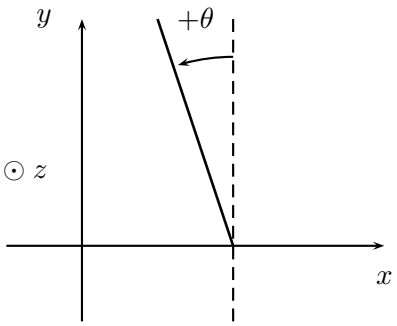


Figure 4.2: Stereo angle θ definition.

The orientation of the stereo angle is defined in Figure 4.2. The measurements in TT are in the sensor's frame, with stereo coordinates (u, v). The stereo coordinates are related to the Cartesian coordinates (x, y) by a rotation with the stereo angle θ:

$$\begin{aligned} u &= x\cos\theta + y\sin\theta\,, \\ v &= -x\sin\theta + y\cos\theta\,, \end{aligned}$$

such that for a X layer $(u, v) \equiv (x, y)$.

VELO-TT matching strategy

The principle of the VELO-TT matching relies on the assumption that the effect of the magnetic field can be replaced by the momentum kick of the track at a fixed $z = z_{\rm m}$. Then the larger the momentum of the track is, the smaller the kick will be, as high-momentum tracks will be less deflected by the field. The plane $z = z_{\rm m}$ corresponds to the "middle of the field" defined as the plane at which the integrated field equals half of the total magnetic field between VELO and TT: $\frac{1}{2}\int_{\rm VELO}^{TT} \vec{B}\cdot {\rm d}\vec{l}$. The value of $z_{\rm m} = 1620$ mm was determined by fitting MC hits in a stand-alone program [87].

The momentum kick is proportional to the mean horizontal deviation Δx (perpendicular to the strips) between the TT measurements and the linear extrapolation of a VELO track, all projected onto a reference plane at $z = z_{\rm proj}$. The reference plane is chosen to be in the middle of the TT stations at $z_{\rm proj} = 2485$ mm, between TTa and TTb. The mean distance Δx is obtained by first, for each cluster, scaling to the reference plane the distance Δx_i between the measurement i in the detection layer located at z_i and the linear extrapolation of the VELO track, using as scaling factor $S_i = (z_{\rm proj} - z_{\rm m})/(z_{\rm i} - z_{\rm m})$, which will be larger for measurements closer to the middle of the field. Then Δx is computed as the mean distance of the n compatible individual measurements: $\Delta x = \frac{1}{n}\sum_i S_i \Delta x_i$. Finally, the momentum of the VELO-TT track is estimated as $p = 1/C \times |\Delta x|$, where $C = 0.000031$ $({\rm MeV/c})^{-1}{\rm mm}^{-1}$ is a calibration coefficient which is related to the field integral between the VELO and TT. The VELO-TT matching is schematically illustrated in Figure 4.3.

The matched tracks are then created as VELO-TT tracks by adding the TT states to the VELO-space tracks. The track parameters are given at $z = z_{\rm proj}$ with the positions and slopes given by the linear extrapolation. The TT state is updated with the VELO x position corrected with Δx, and with the VELO slope $t_x = {\rm d}x/{\rm d}z$ corrected with

$\Delta x/(z_{\text{proj}} - z_{\text{m}})$. Note that there is no kick correction in the vertical plane. Then the momentum information is added by setting $q/p = C \times \Delta x$.

The final step of the matching is a refit of the VELO-TT tracks to correct for wrong momentum estimates [100]. A global χ^2 is built with two contributions. The first is the χ^2_{TT} of a straight TT segment and the second χ^2_{VELO} contains the contribution from the horizontal VELO slope variation. From the track refit we obtain new values for the track parameters x and t_x at TT, imposing the constraint that the two track segments have to pass through the same point (x, y) at z equal to half of the integrated field between VELO and TT for the track under study. The field integral and the position of the middle of the field are retrieved from look-up tables, depending on the vertical slope and on the origin z coordinate of the VELO track. Note that the z coordinate of the midfield used in this fit is different for each track (and does not correspond to z_{m}). The refit modifies the track parameter q/p and assigns a χ^2 to the tracks, without modifying the rest of the state.

The VELO-TT tracking also produces tracks without any momentum information, when the track extrapolation is outside TT sensitive area. An attempt to recover these tracks will be performed by the forward tracking by opening the search windows.

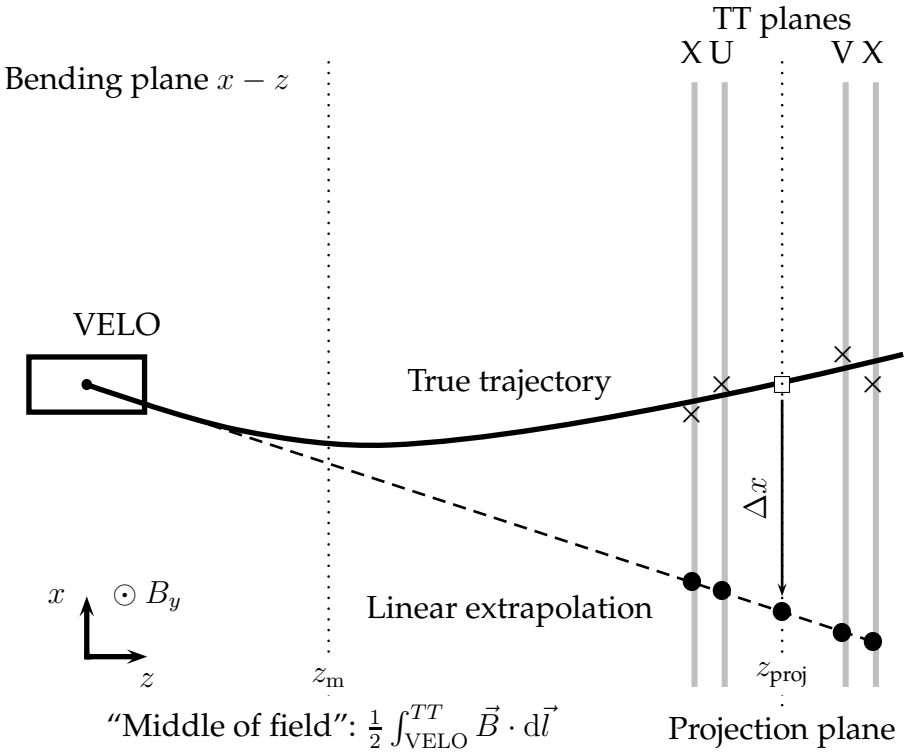


Figure 4.3: Schematic picture of the VELO-TT matching (not to scale).

VELO-TT pattern recognition

The improved VELO-TT reconstruction for the HLT has two passes: the first with strong matching tolerances and global cleanup, and the second with looser tolerances using only TT clusters left over from the first pass. This strategy provides significant improvement in efficiency for a relatively small extra time penalty. This improvement translates to

$\sim 1\%$ increase in efficiency per long track and has a larger impact for higher multiplicity decays. For instance the track-finding efficiency for having all final states in $B_s \to D_s\pi$ is improved by $\sim 5\%$, compared to the former implementation. For L1 the strategy is a single pass with only unused TT clusters (i.e. ignoring the clusters used by a previous track in the same pass). The performance of the on-line tracking is described in Section 4.2.2. We explain hereafter the VELO-TT pattern recognition, i.e. how the track candidates are chosen.

VELO-TT candidates The search starts from forward VELO-3D tracks. From the slopes of the first VELO track state, a first filter rejects tracks outside an enlarged acceptance for TT, i.e. if $t_x > 350$ mrad or if $t_y > 300$ mrad. All the horizontal bands of sensors (`TTPlane`s) are then looked at, and the planes without any clusters are ignored. The track is linearly extrapolated to the z position of the `TTPlane` under consideration to get the predicted $(x_0^{\rm e}, y_0^{\rm e})$ of the track. The $y_0^{\rm e}$ coordinate is used to check if the track can have a measurement in this region of the detector, given some tolerance depending on the momentum of the track (e.g. for multiple scattering). As no momentum is available yet, a term inversely proportional to the minimal momentum track (cut on $p_{\rm min} = 1.5$ GeV/c) is added. Thus the prediction is compatible with a `TTPlane` provided that:

$$(-x_0^{\rm e}\sin\theta + y_0^{\rm e}\cos\theta) + \left(y^{\rm tol} + t_y^{\rm tol}\frac{1}{Cp^{\rm min}}\right) > y_{\rm min}^{\rm TTP}\,,$$
$$(-x_0^{\rm e}\sin\theta + y_0^{\rm e}\cos\theta) - \left(y^{\rm tol} + t_y^{\rm tol}\frac{1}{Cp^{\rm min}}\right) < y_{\rm max}^{\rm TTP}\,,$$

where $y_{\rm min}^{\rm TTP}$ and $y_{\rm max}^{\rm TTP}$ define the height of a band as retrieved from the geometry, and θ is the stereo angle of this band. The tolerances are $y^{\rm tol} = 0.8$ mm and $t_y^{\rm tol} = 200$ mrad. This extremely loose filter will be refined once Δx is known.

The extrapolated track must be in TT sensitive area, namely outside the central hole square size ($d^{\rm hole} = 38$ mm), and inside the overall TT planes ($X_{\rm min} = 660$ mm, $Y_{\rm min} = 580$ mm). In order for the track to cross all TT detection layers, the extrapolated coordinates in the sensor's frame $(u_0^{\rm e}, v_0^{\rm e})$ must thus satisfy:

$$(|u_0^{\rm e}| > d^{\rm hole} \quad \text{or} \quad |v_0^{\rm e}| > d^{\rm hole}) \quad \text{and} \quad |u_0^{\rm e}| < X_{\rm min} \quad \text{and} \quad |v_0^{\rm e}| < Y_{\rm min}\,.$$

This verification will be used later to pass a track, under certain conditions, onto the forward tracking without any momentum estimate.

The track is reextrapolated to the z position of each measurement, as this coordinate varies from sensor to sensor. For each cluster i in the plane the distance Δx_i is computed and scaled to the projection plane, as explained earlier. The momentum of a track is larger at small polar angle with respect to the beam axis. In order to take this into account, the tolerance of the horizontal search window $X^{\rm tol}$ is defined as a constant term ($x^{\rm tol} = 0.35$ mm) and a linear dependence ($s_x^{\rm tol} = 350$ mm) with the polar angle Θ, and all scaled with the ratio of the z distances of the actual ($z_{\rm TTP}$) and projection ($z_{\rm proj}$) planes to the position of the middle of the field ($z_{\rm m}$):

$$\begin{aligned} X^{\rm tol} &= \frac{\left(x^{\rm tol} + s_x^{\rm tol}\sin\Theta\right)}{S}\,, \\ S &= (z_{\rm proj} - z_{\rm m})/(z_{\rm TTP} - z_{\rm m})\,, \\ \sin\Theta &= \sqrt{\frac{p_{\rm T}^2}{p^2}} = \sqrt{\frac{t_x^2 + t_y^2}{1 + t_x^2 + t_y^2}}\,, \end{aligned}$$

where $t_x = p_x/p_z$ and $t_y = p_y/p_z$ are the slopes of the VELO track. The width of the search window is limited to the deflection of the minimal momentum track ($p_{\text{min}} = 1.5$ GeV/c), such that the maximal tolerance is $X^{\text{tol}}_{\text{max}} = 1/Cp^{\text{min}}$. Finally, only the clusters for which $|\Delta x_i| < X^{\text{tol}}$ are considered.

The vertical tolerance is refined for the clusters passing the horizontal tolerance, requiring that:

$$v_i^{\text{e}} + \left(y^{\text{tol}} + t_y^{\text{tol}}|\Delta x_i|\right) > y^{\text{m}}_{\text{min}}\,,$$
$$v_i^{\text{e}} - \left(y^{\text{tol}} + t_y^{\text{tol}}|\Delta x_i|\right) < y^{\text{m}}_{\text{max}}\,,$$

where $(u_i^{\text{e}}, v_i^{\text{e}})$ are the coordinates of the extrapolation at the z of the measurement, and $x^{\text{m}}_{\text{min}}$, $y^{\text{m}}_{\text{min}}$ define the height of the detection area that provided the measurement.

All the clusters compatible with the track under study are collected, thus attaching a list of TT clusters to each VELO track. Additionally, and only for the first pass and in the HLT mode, we check if the linear extrapolation traverses vertical dead zones between consecutive sensors. This is done for each horizontal band of sensors and for each sensor, and ignoring the fact that the sensors are staggered, thus considering the extrapolation at the average z of the TT band. The track is flagged as having traversed a dead zone in a given TT layer if the vertical extrapolation is not inside a sensor sensitive area given some tolerance of 1.4 mm.

At this point we have a list of TT clusters for a given VELO track. These measurements are sorted by increasing ordinate and TT clusters with compatible Δx_i are grouped together, in order to build one or more VELO-TT candidates for a given VELO track. The grouping tolerances are $\Delta x^{\text{tol}}_{\text{1st}} = 0.8$ mm for the first pass, and $\Delta x^{\text{tol}}_{\text{2nd}} = 2.4$ mm for the second pass. In this way the first pass will select the high-momentum tracks, whereas the second pass, if any, will recover tracks with a larger deflection. The compatible groups of clusters are formed by requiring at least three clusters in three different TT layers, and allowing for a maximum of two clusters per layer, since sensors overlap in x. The solutions with more than two compatible clusters in the same detection layer are discarded, as they are mainly due to busy regions. Obviously the number of compatible clusters within one layer is directly related to the choice of Δx^{tol}. The choices of tolerances will be motivated later. In order to have a compatible group of clusters assigned to the VELO track, we require that $|\Delta x_i - \Delta x_j| < \Delta x^{\text{tol}}$ for any pair of clusters, independently of the TT layer. This global tolerance is refined for clusters in the same station (TTa or TTb) with $0.5 \times \Delta x^{\text{tol}}$, and for clusters in the same layer with $0.25 \times \Delta x^{\text{tol}}$. The search of compatible clusters in a group is performed until one cluster fails any of the grouping tolerances. The clusters already used in a solution with four layers fired are ignored in the following groups in order to avoid having solutions that share measurements in three layer when one of them has four layers. The output of this part is a list of groups of compatible TT clusters for the VELO track under consideration.

The creation of the VELO-TT tracks depends on the running mode. For L1, a local cleanup of the VELO-TT candidates is directly applied for each VELO track. The criteria for choosing the best candidate will be explained later, and only one VELO-TT track is created per VELO track, all having a definite momentum. Moreover, only unused clusters are considered as one loops over the VELO tracks.

For the HLT mode, a global cleanup of the different VELO tracks sharing clusters is done in the first pass, before creating the VELO-TT tracks. As we do not want to depend on the order in which VELO tracks are processed, there is no unused clusters requirement for this first pass. If a VELO track has no compatible TT clusters, a VELO-TT track is still created without momentum information if the extrapolation of the VELO track does not cross all four TT layers, or if it traverses at least two dead regions in any different TT layers. Finally, the second pass is similar to the L1 pass, only considering the VELO tracks that were not successfully extended in the first pass. For this second pass the grouping tolerance is looser, and only unused clusters are considered without the global cleanup and ignoring tracks without any momentum. Note that the global cleanup of candidates in the HLT first pass is applied *before* the local cleanup. The purpose of the global cleanup is to reject candidates based on different VELO tracks and sharing clusters. The final decision is always taken by the local cleanup, which only outputs one VELO-TT track per input VELO track.

Global cleanup This cleanup considers VELO-TT candidates based on different VELO tracks. The VELO-TT candidates with a unique combination of TT clusters are ignored, as well as the tracks without any momentum. We only consider VELO-TT candidates with the same number of layers fired, and sharing at least two TT clusters. In the latter case, the chosen candidate is then the one for which the variance of its Δx is the smallest, using the definition:

$$\mathrm{Var}[\Delta x] = \frac{1}{n}\sum_i (S_i \Delta x_i)^2 - \left(\frac{1}{n}\sum_i (S_i \Delta x_i)\right)^2 ,$$

where n is the number of TT clusters used and Δx_i is the horizontal distance between the measurement of the cluster i and the predication.

Local cleanup This cleanup is applied to get only one VELO-TT track per VELO track, by comparing solutions based on the same VELO track:

- If two candidates have a different number of layers fired, then the solution with the largest number of layers fired is kept.
- If two candidates have the same number of layers fired, then the solution with the smallest Var$[\Delta x]$ is kept. An exception is made for the first pass in the HLT when one (and only one) of the two candidates has clusters in three different layers and a missing cluster in the fourth layer due to a dead region: in this case this candidate is selected regardless of its Var$[\Delta x]$.

The final step of the pattern recognition is the creation of the VELO-TT tracks by adding the TT state to the VELO track, and the refit of the track to correct for momentum mistakes. Note that some of the VELO-TT tracks do not have any momentum estimate when running in the HLT mode; these tracks may be recovered by the forward tracking.

The different tolerances were tuned on three different decay topologies, namely $B_d \to \pi^+\pi^-$, $B_s \to \phi\,(K^+K^-)\phi\,(K^+K^-)$ and $B_s \to D_s\,(K^+K^-\pi)\pi$, considering both the speed performance and the efficiency after the forward tracking. The use of the momentum obtained from VELO-TT tracks reduces the processing time of the forward tracking by a factor two. Clearly, the speed performance depends on the quality of the VELO-TT momentum determination. The tolerances in VELO-TT should be tight enough not to

provide the forward tracking with a wrong momentum estimate, which is used to define the search windows in the long tracking. Moreover, providing more than one VELO-TT track per VELO track significantly worsens the overall timing performance. The use of TT has however limitations. The first is when the tracks do not traverse the sensitive areas. The recovery of these tracks is delegated to the forward tracking, and represents an important part of its processing time. The second limitation is due to the few constraints one can impose on the TT clusters, given that there are only four detection planes.

In Table 4.1 we show the number of signal tracks that were correctly reconstructed as VELO-space tracks and failed to be extended to the tracking stations as a function of the number of TT hits left by the corresponding signal MC particle. We see that the majority of the failures corresponds to tracks that do not have enough clusters to be reconstructed: approximatively 60% of the missed signal tracks that were reconstructed in 3D in the VELO have less than three TT hits. For each of the signal samples, the largest fraction of missed tracks is for signal tracks with two TT hits. This is in most cases a consequence of the vertical dead regions between silicon sensors: given the large enough separation between TTa and TTb, a track crossing a gap in one of the two TT stations usually leaves two hits in the other station. Failures to reconstruct tracks when the signal MC particle left 3 or 4 MC hits are either due to wrong choices in the cleanup phase ($\sim 2/3$ of the cases for signal tracks from 4-prong B decays) or deviations Δx_i outside the tolerances.

Table 4.1: Performance of the VELO-TT pattern recognition for tracks from different signal B decays. $N^{\mathrm{tr}}_{\mathrm{VELO-space}}$ is the number of VELO-space tracks, $N^{\mathrm{tr}}_{\mathrm{TT}}$ is the number of VELO-TT tracks, and $N^{\mathrm{hits}}_{\mathrm{TT}}$ is the number of TT clusters left by the signal MC particle.

Signal channel	$N^{\mathrm{tr}}_{\mathrm{VELO-space}}$	$N^{\mathrm{tr}}_{\mathrm{TT}}$	Number of failures as a function of $N^{\mathrm{hits}}_{\mathrm{TT}}$				
			0	1	2	3	4
$B_d \to \pi^+\pi^-$	2438	2427	2	1	4	2	2
$B_s \to \phi\phi$	3715	3665	5	12	18	10	14
$B_s \to D_s\pi$	10237	10114	3	15	54	22	29

Without giving the full detail of each feature, we describe hereafter a few aspects that are critical to the VELO-TT and forward tracking performances. Moreover, we ignore the correlations between the different features (e.g. when a track traverses a dead zone in TTa and the central hole in TTb). A few key points of the current on-line pattern recognition are listed hereafter:

- TT has a limited acceptance (TT planes, central hole) which can make the extrapolation of the VELO track be outside of the sensitive area, and the correct solution will have at least a missing cluster in one of the TT layers. The partial recovery of these tracks is achieved by passing them to the forward tracking, but without any momentum estimate.

- Consecutive TT sensors have gaps $\Delta y = 2 \times 1.4$ mm between them which are periodic in the sensor's height. Consequently a loss of $(2 \times 1.4\,\mathrm{mm})/108\,\mathrm{mm} \sim 2\%$ per track is expected, neglecting the effect of stereo layers.

- The study of the performance when looking for two cluster solutions or several solutions per VELO track shows that these strategies considerably increase the forward tracking timing and the chances for wrong combinations. Moreover, it turns out that it is more efficient to pass a track without any TT clusters at all compared to passing many solutions with wrong momentum estimates.

- Given the fact that a wrong momentum assignment in TT significantly affects the performance of the forward tracking, the matching tolerances applied in TT are meant to favor high-momentum tracks. Even though the tolerances have a dependence on the momentum, a worse performance is expected for low-momentum tracks or for high-multiplicity decays. The second pass in VELO-TT was introduced to recover some low-momentum tracks.

4.2.2 On-line Track-Finding

The track-finding efficiencies are determined for three purely hadronic channels with distinct topologies: $B_d \to \pi^+\pi^-$, $B_s \to \phi(K^+K^-)\phi(K^+K^-)$ and $B_s \to D_s(K^+K^-\pi)\pi$. The timing and ghost rates are measured with the whole minimum-bias sample. For the timing we assume in this section that we redo all the VELO-TT and forward tracking after the generic HLT and ignore the fact that part of the forward tracking was already made. Moreover the quoted times were not determined on dedicated machines, hence they just represent an order of magnitude. The detailed and accurate speed performance of each step will be given in Section 4.4.2.2.

A measure of the track-finding performance is given by the ghost rates, since ghost tracks have many wrong clusters. The fractions of ghosts in minimum-bias events for each track type are given in Table 4.2, including the result for long tracks without any TT information. As expected, VELO-TT tracks have the highest ghost rate since TT has only four planes and therefore there are not many constraints to assign the correct clusters to a VELO track. Nevertheless, the use of TT actually allows to reduce the ghosts in the long tracks by almost a factor two.

The reduction of the long track ghost rate when using TT not only reflects the fact that the momentum information obtained from VELO-TT tracks can be used in the forward tracking to narrow its search regions, but also considerably reduces its processing time. This can be seen in Table 4.3 where the timing for the VELO-TT pattern recognition and the forward tracking are shown. The quoted values are for both the buffer decoding and the tracking parts, and the equivalent time when skipping TT is shown for comparison. We see that the use of TT gives a total time reduction of at least a factor two.

Table 4.2: Ghost rates determined from ~ 3.28 s minimum-bias data, for each type of on-line track. The samples contain over 1 M tracks each.

	Minimum-bias				
Track types	VELO-RZ	VELO-space	VELO-TT	Long	Long (no TT)
Ghost rate [%]	12.3	9.4	16.6	5.8	11.0

Note that the quoted efficiencies for VELO-TT tracks include the effect of a track being later found by the forward tracking without any TT clusters. This is relevant if one wants

Table 4.3: Times for the VELO-TT and forward tracking algorithms as measured on minimum-bias events with a 2.8 GHz Xeon processor. These measurements are affected by the machine load, and include the times for buffer decoding and full tracking.

	Average speed per minimum-bias event	
Algorithm	Standard on-line [ms]	On-line without TT [ms]
VELO-TT	1.9	-
Forward	8.4	22.3
Combined	10.3	22.3

to use upstream tracks in the HLT since the efficiency for finding the signal tracks with momentum information is a bit lower than the quoted ones. Moreover, in the association to the MC truth of the long tracks, the fraction of wrong TT hits is ignored for the decision.

The track-finding efficiencies for the different signals are given in the Tables 4.4, 4.5, and 4.6 for each of the track types. Note that the on-line VELO tracking *should* be 100% efficient on off-line reconstructed tracks, as the same version of code will be run in the future. The per-track efficiency decreases with the number of tracks involved in the b decay. Tracks from low-multiplicity decays will have in average a higher momentum making them easier to reconstruct in VELO-TT. Higher multiplicity decays will see their final states undergo more multiple scattering resulting in a worse VELO slope determination, strongly affecting the forward tracking performance.

For decays to purely hadronic charged final states involving only long tracks the fraction of events supposedly lost due to tracking inefficiencies and before applying any specific HLT selection is:

$\sim 2\%$ per track, $\sim 3-4\%$ for a 2-prong decay ($B_d \to \pi^+\pi^-$);

$\sim 4\%$ per track, $\sim 13\%$ for a 4-prong decay ($B_s \to \phi\phi$, $B_s \to D_s\pi$);

$\sim 5\%$ per track, $\sim 26\%$ for a 6-prong decay ($B_s \to \phi(K^+K^-)\eta_c(4h)$).

A larger loss is expected for channels using upstream tracks, such as the slow π from D^*, since these tracks have in average lower momentum.

The use of TT in the on-line tracking improves the timing by at least a factor two, however it leads to inefficiencies that are difficult to recover and that would mean an additional time cost. If we want to have in the HLT all the tracks relevant to an off-line selection, then the on-line and off-line codes should have a better correlation and the use of TT should probably be avoided as an intermediate step before the extrapolation after the magnet.

4.3 HLT Selections

We describe in this section the selection procedure for the exclusive b and D^* streams. The HLT selection algorithms are executed on all events giving a positive generic HLT decision. The current simplified data-flow of the HLT sequence is summarized in Figure 4.4. We present hereafter the steps involved in the exclusive b and the D^* streams selections:

Table 4.4: On-line tracking efficiencies on signal $B_d \to \pi^+\pi^-$ signal tracks. Results are for $N^{evt}_{off} = 1225$ events after generic HLT and off-line selection, for events with all signal final states ($N^{tr}_{fin} = 2$) off-line reconstructed. The uncertainties are statistical.

$B_d \to \pi^+\pi^-$	Per track		Per event	
Track type	N^{tr}_{rec}	ε^{tr}_{rec} [%]	N^{evt}_{rec}	ε^{evt}_{rec} [%]
VELO-RZ	2447	99.9 ± 0.1	1222	99.8 ± 0.1
VELO-space	2438	99.5 ± 0.1	1213	99.0 ± 0.3
VELO-TT	2427	99.1 ± 0.2	1202	98.1 ± 0.4
Long	2411	98.4 ± 0.3	1186	96.8 ± 0.5
Long (no TT)	2415	98.6 ± 0.2	1191	97.2 ± 0.5

Table 4.5: On-line tracking efficiencies on signal $B_s \to \phi\phi$ signal tracks. Results are for $N^{evt}_{off} = 943$ events after generic HLT and off-line selection, for events with all signal final states ($N^{tr}_{fin} = 4$) off-line reconstructed. The uncertainties are statistical.

$B_s \to \phi\phi$	Per track		Per event	
Track type	N^{tr}_{rec}	ε^{tr}_{rec} [%]	N^{evt}_{rec}	ε^{evt}_{rec} [%]
VELO-RZ	3754	99.5 ± 0.1	926	98.2 ± 0.4
VELO-space	3715	98.5 ± 0.2	895	94.9 ± 0.7
VELO-TT	3656	96.9 ± 0.3	845	89.6 ± 1.0
Long	3609	95.7 ± 0.3	809	85.8 ± 1.1
Long (no TT)	3638	96.4 ± 0.3	829	87.9 ± 1.1

Table 4.6: On-line tracking efficiencies on signal $B_s \to D_s\pi$ signal tracks. Results are for $N^{evt}_{off} = 2580$ events after generic HLT and off-line selection, for events with all signal final states ($N^{tr}_{fin} = 4$) off-line reconstructed. The uncertainties are statistical.

$B_s \to D_s\pi$	Per track		Per event	
Track type	N^{tr}_{rec}	ε^{tr}_{rec} [%]	N^{evt}_{rec}	ε^{evt}_{rec} [%]
VELO-RZ	10306	99.9 ± 0.0	2568	99.5 ± 0.1
VELO-space	10237	99.2 ± 0.1	2502	97.0 ± 0.3
VELO-TT	10114	98.0 ± 0.1	2386	92.5 ± 0.5
Long	9955	96.5 ± 0.2	2241	86.9 ± 0.7
Long (no TT)	10042	97.3 ± 0.2	2317	89.8 ± 0.6

- Preparation of the final state particles;
- Creation of the shared composite particles;
- HLT selections to fully reconstruct the different decay chains.

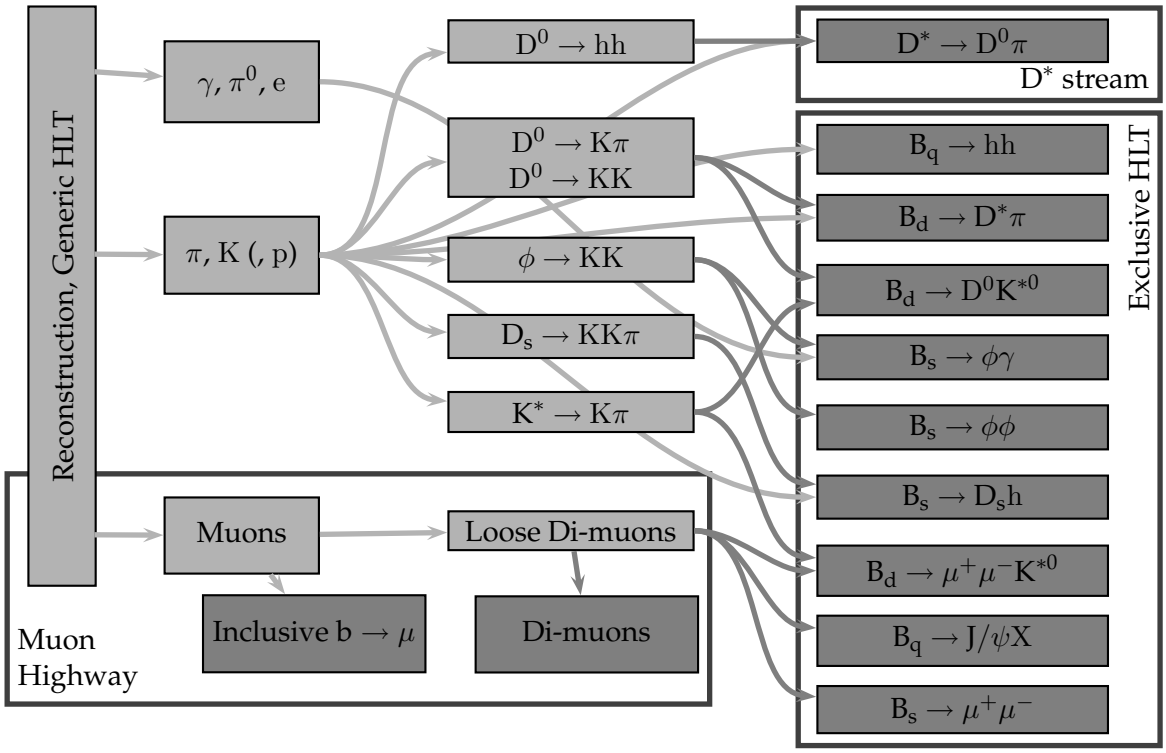


Figure 4.4: Simplified data flow in the HLT. Each gray box is an algorithm (or a set of algorithms).

The first two steps are motivated by important requirements for the HLT, namely speed performance and output rate. Given the limited available time budget to process the whole HLT, we cannot afford to perform several times identical operations unless it is absolutely necessary. Thus to avoid duplication the creation of final states and of composite particles is executed only once per event, for each particle type. A few selection criteria are already applied at this stage, to limit the combinatorics and therefore save some additional processing time. Finally, a set of common cuts is also applied to all final states made from long tracks to control the output rate.

These shared particles will then be reused and refined according to the needs of the dedicated HLT exclusive selections, combining them all the way to the signal b-hadron or D^* candidates. The modes used to reconstruct the different exclusive b and D^* streams are presented in Section 4.3.3. Every exclusive selection is executed independently of others on all events passing the generic HLT. The final decision is then a logical *OR* of all the individual exclusive selections. Note that the D^* stream can be considered as an exclusive selection since it is *always* reconstructed with the same signature (slow pion and $D^0 \to hh$) and with all tracks made as pions. From the bandwidth point of view, the D^* selection nevertheless represents a separate stream with respect to the exclusive b selections.

4.3.1 Final States

The HLT selections use final state particles as inputs and then combine them to form composite particles until the whole decay chain is reconstructed. The initial step is therefore to create the final states, using particle makers that assign particle identification (PID hypotheses) to the end products of the reconstruction. We distinguish three different classes of particles used in the HLT:

Hadrons Charged hadron particles (π, K, p) are made from VELO-TT or long tracks and reconstructed by the standard on-line pattern recognition. All tracks are made as both pion and kaon candidates. The pions should not be affected by PID requirements to ensure that the flavor-blind selections work for any particle hypothesis.

Muons Muon particles are made from muon segments and compatible with the μ hypothesis as identified by the stand-alone muon reconstruction. Muons are directly retrieved from the generic HLT sequence.

ECAL ECAL particles are built from neutral on-line electromagnetic calorimeter clusters (photons γ, resolved π^0's) or from long tracks matched to charged ECAL clusters that were not already identified as muons (electrons e, including bremsstrahlung correction). The ECAL reconstruction is only run after the generic HLT.

Once all the final state particles are created, they are distributed to different pools where a preliminary selection is applied in order to reduce combinatorics as early as possible. At this point electrons and π^0's are lost since they are not used in the current implementation of the HLT.

The choice of preliminary selection criteria for the final states is based on simple b-decay features (or c decays to a lesser extend) and used as signatures for signal tracks:

- Some transverse momentum is required with respect to the beam axis, as a result of the large mass difference between b hadrons and their decay products.
- Some impact parameter is required with respect to the interaction vertex, as a consequence of the presence of significantly displaced secondary vertices with respect to the production vertex due to the long b hadron lifetime.

The preselection only affects hadron long tracks, and the list of cuts is given in Table 4.7. The choice of cuts will be justified in Section 4.3.5.

4.3.2 Shared Composite Particles

All the intermediate composite particles are built only once per event and using the filtered final states from the different pools as discussed in Section 4.3.1. This avoids the duplication of computing and the resulting composite particles are then used and refined by the different exclusive streams. Since these particles are meant to be shared among all possible exclusive selections, the selection criteria are loose enough to satisfy all use cases. Besides the obvious time gain, using shared composite particles has the advantage of uniforming and standardizing the HLT selections.

The shared composite particles are, for the time being, reconstructed in the following modes (including charge conjugates when applicable):

- $\mathrm{K}^{*0} \to \mathrm{K}^+\pi^-$;

Table 4.7: Preliminary selection criteria defining the different pools of particles used in the HLT selections.

π, K long tracks		
p [MeV/c]	$>$	2000.
p_{T} [MeV/c]	$>$	300.
sIPS	$>$	2.
Loose π (upstream, long)		
sIPS	$>$	1.
Photons γ		
p_{T} [MeV/c]	$>$	300.
Muons		
No cuts		

- $\phi \to \mathrm{K}^+\mathrm{K}^-$;
- $\mathrm{D}^0 \to \pi^+\pi^-$, $\mathrm{K}^+\pi^-$, $\pi^+\mathrm{K}^-$, $\mathrm{K}^+\mathrm{K}^-$ (always labeled as "D^0" and never as "$\overline{\mathrm{D}}^0$");
- $\mathrm{D}_\mathrm{s}^- \to \mathrm{K}^+\mathrm{K}^-\pi^-$;
- $\mathrm{J}/\psi \to \mu^+\mu^-$.

All these shared particles are formed using particles made from long tracks (or muons from the generic HLT), with the cuts given in Table 4.7. The list of preliminary cuts applied to these shared particles is given in Table 4.8 and will be justified later when discussing all the HLT selection criteria.

Table 4.8: HLT selection criteria for shared composite particles.

Shared K^{*0}		
δm [MeV/c^2]	$\pm$	160.
p_{T} [MeV/c]	$>$	500.
χ^2 vertex	$<$	100.
Shared ϕ		
δm [MeV/c^2]	$\pm$	50.
χ^2 vertex	$<$	100.
Shared D_s		
δm [MeV/c^2]	$\pm$	45.
p_{T} [MeV/c]	$>$	1750.
χ^2 vertex	$<$	36.
FS	$>$	4.

Shared $\mathrm{D}^0 \to \mathrm{K}^+\pi^-$, $\pi^+\mathrm{K}^-$, $\mathrm{K}^+\mathrm{K}^-$		
δm [MeV/c^2]	$\pm$	100.
p_{T} [MeV/c]	$>$	500.
χ^2 vertex	$<$	100.
Shared $\mathrm{D}^0 \to \pi^+\pi^-$		
Upper δm [MeV/c^2]	$+$	50.
Lower δm [MeV/c^2]	$-$	700.
p_{T} [MeV/c]	$>$	1250.
χ^2 vertex	$<$	16.
FS	$>$	6.

The HLT mass cuts should be loose enough in order to keep sidebands and account for the worse on-line mass resolutions. The resolutions will justify the mass cuts applied when using the shared composite particles in the different exclusive streams. On-line mass resolutions are determined using off-line selected events passing the generic HLT and for which all signal tracks are reconstructed in the trigger.

A comparison between on-line and off-line resolutions is given in Table 4.9. The HLT mass resolutions are roughly a factor two worse when compared to the off-line. This is a consequence of the worse track parameters and momentum resolution.

Table 4.9: On-line (left) and off-line (right) mass resolutions of shared composite particles. The quoted values are from single Gaussian fits: $\sigma_{\rm reso}$ is the purely experimental resolution, while σ includes the natural width of the resonance.

On-line			Off-line		
Shared	σ [MeV/c^2]	$\sigma_{\rm reso}$ [MeV/c^2]	Shared	σ [MeV/c^2]	$\sigma_{\rm reso}$ [MeV/c^2]
$\rm K^{*0}$	32.3	5.4	$\rm K^{*0}$	31.7	3.0
ϕ	3.7	2.0	ϕ	3.2	0.9
$\rm D^0$	12.9	12.9	$\rm D^0$	7	7
$\rm D_s$	10.5	10.5	$\rm D_s$	5	5

4.3.3 Exclusive HLT Selection Channels

The exclusive b and $\rm D^*$ events are reconstructed in the modes described hereafter:

$\rm B_s \to D_s h$ is always built as $\rm B_s \to D_s\pi$ using a shared $\rm D_s \to K^+K^-\pi$ and a bachelor π made from a long track and with $p_{\rm T}$ and sIPS cuts. The $\rm B_s$ mass window is enlarged to accommodate $\rm B_s \to D_sK$.

$\rm B_q \to J/\psi X$ is reconstructed as a muon pair from the generic HLT with large sidebands around the $\rm J/\psi$ nominal value. No extra tracks are required. As the rate of true $\rm J/\psi$ is much larger than the few hertz we would like to grant to this exclusive selection, tight cuts are applied to the $\rm J/\psi$ and the muons, such that this exclusive selection effectively becomes a lifetime biased di-muon stream. As most of the di-muons will anyway be already triggered by the muon streams, we do not worry about the efficiency for this selection.

$\rm B_q \to hh$ is always built as a PID-blind $\rm B_d \to \pi\pi$ using opposite sign pions made from long tracks passing $p_{\rm T}$ and sIPS cuts. The mass window is wide enough to accommodate all $\rm B_q \to hh$ modes, namely $\rm B_d \to \pi^+\pi^-$, $\rm B_d \to K^+\pi^-$, $\rm B_s \to K^+K^-$, and $\rm B_s \to \pi^+K^-$.

$\rm B_d \to D^0K^{*0}$ is formed with a shared $\rm D^0 \to K^+\pi^-$, $\pi^+\rm K^-$, or $\rm K^+K^-$, and a shared $\rm K^{*0} \to K^+\pi^-$, including wrong sign combinations;

$\rm B_s \to \phi\phi$ is selected with shared $\phi \to \rm K^+K^-$ particles.

$B_d \to D^* \pi$ is built using a shared $D^0 \to K^+\pi^-$, π^+K^-, K^+K^-, a slow pion made from an upstream or long track and without any cuts for the $D^* \to D^0\pi$, and a fast pion made from a long track passing p_T and sIPS cuts.

$B_d \to \mu^+\mu^- K^{*0}$ is built using a muon pair from the generic HLT and a shared $K^{*0} \to K^+\pi^-$.

$B_s \to \mu^+\mu^-$ is built using a muon pair from the generic HLT.

$B_s \to \phi\gamma$ is formed using a shared $\phi \to K^+K^-$ and a high transverse energy (E_T) photon. Note that the photon does not participate in the vertex determination and it is just used to compute the resulting four-momentum and its covariance matrix.

D^* stream is selected using a shared $D^0 \to \pi^+\pi^-$ and a detached slow pion made from an upstream or long track. This stream is PID-blind and use asymmetric mass windows to accept the other two-body D^0 decays.

4.3.4 Exclusive HLT Selection Criteria

The HLT selections have a bandwidth of $\sim$ 200 Hz for the exclusive b stream and of $\sim$ 300 Hz for the D^* stream. In order to lower the minimum-bias rate down to these specifications a series of filters is applied to all the channels. The goal is of course to have the lowest possible rate while loosing as few interesting events as possible. We imposed the following constraints for the cuts determination:

- Only selection criteria from the list given in Section 4.1.3 are used, with round cut values to avoid any excessive and too specific tuning of the cuts.
- Unless specified, the preliminary requirements introduced in Section 4.3.1 are applied to all selections and the standard composite particles of Section 4.3.2 are used. According to the needs of the different selections, these shared particles may be further refined.
- The mass cuts, though extremely useful in terms of background rejection, are loose enough to include sidebands and generic PID-blind selections. The mass cuts for the intermediate particles correspond to at least three standard deviations and the b hadron mass windows are ± 500 MeV/c^2 or more. Finally, for PID-blind selections the mass windows are enlarged sufficiently to accommodate the different particle hypotheses.
- The cuts are tuned such that the output rate of each of the dedicated exclusive b streams is at most of the order of 10 Hz, ignoring possible correlations between these streams. All the available bandwidth is intentionally *not* filled thus leaving room for additional new exclusive selections. For the D^* stream, all the bandwidth is used.

The procedure to determine the selection criteria can be summarized as follows:

1. We use signal events after off-line selection and generic HLT to study the effect of the cuts. A set of loose preselection cuts is then defined for each stream, ignoring the PID when required, namely for the PID-blind selections such as $B_q \to h\,h$. The values of the preselection cuts are obvious choices when considering off-line and on-line distributions.

2. The preselection cuts are then applied on minimum-bias events to limit the combinatorics. Based on the distributions of the minimum-bias events versus the HLT-preselected and associated signal candidates, the final cuts are chosen such that the output rate is lowered to the required level while only cutting the tails of the signal distributions. There is no fine tuning of the selection cuts and we impose cuts such as to get rid of signal candidates with abnormal parameters due to the on-line environment or simply not in the core of the distributions. As the HLT selections can tolerate to have some background within the output rate, there is no optimization of the significance of the signal.

The selection cuts were determined on the DC04-v1 minimum-bias sample and then blindly applied to the DC04-v2 minimum-bias sample to assess the performance. The final selection criteria and the preselection cuts used for their determination are given in the Tables 4.10 – 4.19.

4.3.5 Final State Filtering

The preliminary filtering of the final states introduced in Section 4.3.1 and in particular the IP cuts do not affect muons, as these do not have any significant impact on combinatorics. We therefore only consider pions and kaons for the determination of the preselection cuts. The effect of cutting on the transverse momentum (p_{T}) and on the (unsigned) impact parameter significance with respect to *all* reconstructed primary vertices (sIPS) has been investigated in terms of:

1. Processing time on minimum-bias events for all the exclusive b and D^* selections.
2. Output rate on minimum-bias events of the exclusive b stream.
3. Efficiency of the dedicated selections for each of the studied signal channels involving hadrons.

A scan for different values of the preliminary selection criteria p_{T} and sIPS applied to all π and K made from long tracks is done for each of the above items. Besides the obvious gain in timing and retention rate achieved when applying such a selection at the particle creation level, the choices for the cuts values were motivated by the fact that in the longer term many more exclusive channels, $\sim \mathcal{O}(100)$, will be considered in addition to the core signal channels studied here. The final preliminary selection cut values thus represent a conservative choice in terms of output rate, and are in some cases to the prejudice of signal efficiencies. This is actually one place where we are forced to cut harder compared to several off-line selections. However, the performance could be improved in a next iteration (except for lifetime unbiased selections) by retuning the off-line selections that presently apply very loose or no p_{T} and sIPS requirements .

For the p_{T} and sIPS scan, the final selection criteria as presented in Section 4.3.4 are applied, except for the p_{T} and sIPS cuts on all hadrons. For exclusive HLT selections with stronger requirements of the final states, such as $\mathrm{B_q} \to \mathrm{h\ h}$ or the bachelor in $\mathrm{B_s} \to \mathrm{D_s\ h}$, the preliminary selection criteria on the final states have no effect. Moreover, only long hadron tracks are affected, i.e. particles made from upstream tracks are not considered in the scan.

The impact on the processing time of the different p_{T} and sIPS cuts in minimum-bias events is shown in Figure 4.5. We consider the HLT selections CPU time normalized to a reference algorithm to avoid side effects due to the machine load. The chosen reference is

Table 4.10: $B_s \to D_s h$ final HLT selection criteria. The rightmost values in parentheses indicate the preselection used in the final cut determination.

D_s products			
p [MeV/c]	$>$	2000.	(1500.)
p_T [MeV/c]	$>$	300.	(200.)
sIPS	$>$	2.	(1.)

D_s			
δm [MeV/c^2]	$\pm$	45.	(80.)
p_T [MeV/c]	$>$	1750.	(1000.)
χ^2 vertex	$<$	36.	(100.)
FS	$>$	4.	$(-)$

Bachelor h			
p [MeV/c]	$>$	2000.	(1000.)
p_T [MeV/c]	$>$	700.	(500.)
sIPS	$>$	3.	(1.)

B_s			
δm [MeV/c^2]	$\pm$	600.	(600.)
χ^2 vertex	$<$	9.	(100.)
sIPS	$<$	3.	(9.)
FS	$>$	1.	(1.)
FD [mm]	$>$	1.	$(-)$
$\cos\theta_{p,F}$	$>$	0.9998	(0.99)

Table 4.11: $B_q \to J/\psi X$ final HLT selection criteria. The rightmost values in parentheses indicate the preselection used in the final cut determination.

J/ψ products			
p_T [MeV/c]	$>$	500.	$(-)$
sIPS	$>$	2.	$(-)$

J/ψ			
δm [MeV/c^2]	$\pm$	120.	$(-)$
FS	$>$	2.	$(-)$
FD [mm]	$>$	2.5	$(-)$
sIPS	$>$	3.	$(-)$
χ^2 vertex	$<$	25.	$(-)$

Table 4.12: $B_q \to hh$ final HLT selection criteria. The rightmost values in parentheses indicate the preselection used in the final cut determination.

B_q products			
p [MeV/c]	$>$	3000.	(2000.)
p_T [MeV/c]	$>$	800.	(600.)
sIPS	$>$	4.	(2.)

B_q			
δm [MeV/c^2]	$\pm$	600.	(600.)
p_T [MeV/c]	$>$	1000.	(800.)
χ^2 vertex	$<$	9.	(50.)
FS	$>$	8.	(4.)
FD [mm]	$>$	1.	$(-)$
sIPS	$<$	3.	(9.)
$\cos\theta_{p,F}$	$>$	0.9995	(0.999)

Table 4.13: $B_d \to D^0 K^{*0}$ final HLT selection criteria. The rightmost values in parentheses indicate the preselection used in the final cut determination.

D^0 and K^{*0} products		
p [MeV/c]	$> 2000.$	$(1000.)$
p_T [MeV/c]	$> 300.$	$(200.)$
sIPS	$> 2.$	$(1.)$

D^0		
δm [MeV/c^2]	$\pm 50.$	$(100.)$
p_T [MeV/c]	$> 1000.$	$(500.)$
χ^2 vertex	$< 36.$	$(100.)$
sIPS	$> 4.$	$(-)$

K^{*0}		
δm [MeV/c^2]	$\pm 100.$	$(160.)$
p_T [MeV/c]	$> 1000.$	$(500.)$
χ^2 vertex	$< 36.$	$(100.)$
sIPS	$> 4.$	$(-)$

B_d		
δm [MeV/c^2]	$\pm 500.$	$(500.)$
χ^2 vertex	$< 16.$	$(100.)$
FS	$> 4.$	$(-)$
FD [mm]	$> 1.$	$(-)$
sIPS	$< 4.$	$(9.)$
$\cos\theta_{p,F}$	> 0.9999	(0.999)

Table 4.14: $B_s \to \phi\phi$ final HLT selection criteria. The rightmost values in parentheses indicate the preselection used in the final cut determination.

ϕ products		
p [MeV/c]	$> 2000.$	$(1000.)$
p_T [MeV/c]	$> 300.$	$(200.)$
sIPS	$> 2.$	$(1.)$

ϕ		
δm [MeV/c^2]	$\pm 20.$	$(50.)$
χ^2 vertex	$< 49.$	$(50.)$

B_s		
δm [MeV/c^2]	$\pm 500.$	$(500.)$
χ^2 vertex	$< 49.$	$(100.)$
FS	$> 6.$	$(4.)$
FD [mm]	$> 1.$	$(0.)$
sIPS	$< 6.$	$(9.)$
$\cos\theta_{p,F}$	> 0.9995	(0.999)

Table 4.15: $B_d \to D^*\pi$ final HLT selection criteria. The rightmost values in parentheses indicate the preselection used in the final cut determination.

D^0 products			
p [MeV/c]	$>$	2000.	(1000.)
p_T [MeV/c]	$>$	300.	(200.)
sIPS	$>$	2.	(1.)

D^0			
δm [MeV/c^2]	$\pm$	50.	(100.)
p_T [MeV/c]	$>$	1000.	(500.)
χ^2 vertex	$<$	36.	(100.)

Slow π (upstream and long tracks)
No cuts

D^*			
δm [MeV/c^2]	$\pm$	50.	(100.)
$m(D^* - D^0)$ [MeV/c^2]	$<$	6.	(60.)
p_T [MeV/c]	$>$	1000.	(500.)
χ^2 vertex	$<$	36.	(100.)

Fast π			
p [MeV/c]	$>$	2000.	(1000.)
p_T [MeV/c]	$>$	1300.	(1000.)
sIPS	$>$	2.	(1.)

B_d			
δm [MeV/c^2]	$\pm$	500.	(500.)
χ^2 vertex	$<$	25.	(100.)
FS	$>$	3.5	(2.)
FD [mm]	$>$	1.	(0.)
sIPS	$<$	4.	(9.)
$\cos\theta_{p,F}$	$>$	0.9995	(0.995)

Table 4.16: $B_d \to \mu^+\mu^- K^{*0}$ final HLT selection criteria. The rightmost values in parentheses indicate the preselection used in the final cut determination.

K^{*0} products			
p [MeV/c]	$>$	2000.	(1000.)
p_T [MeV/c]	$>$	300.	(200.)
sIPS	$>$	2.	(1.)

K^{*0}			
δm [MeV/c^2]	$\pm$	120.	(160.)
p_T [MeV/c]	$>$	850.	(500.)
χ^2 vertex	$<$	36.	(100.)

μ			
p_T [MeV/c]	$>$	900.	(500.)

B_d			
δm [MeV/c^2]	$\pm$	500.	(500.)
χ^2 vertex	$<$	100.	(150.)
FS	$>$	5.	$(-)$
FD [mm]	$>$	1.	$(-)$
sIPS	$<$	5.	(9.)
$\cos\theta_{p,F}$	$>$	0.9995	(0.998)

Table 4.17: $B_s \to \mu^+\mu^-$ final HLT selection criteria. The rightmost values in parentheses indicate the preselection used in the final cut determination.

μ		
p_T [MeV/c]	$> 1200.$	$(500.)$

B_s		
δm [MeV/c^2]	$\pm$ 500.	$(500.)$
p_T [MeV/c]	$> 2900.$	$(1000.)$
χ^2 vertex	$<$ 49.	$(100.)$
FS	$>$ 12.	$(-)$
sIPS	$<$ 5.	$(9.)$
$\cos\theta_{p,F}$	$>$ 0.999	$(-)$

Table 4.18: $B_s \to \phi\gamma$ final HLT selection criteria. The rightmost values in parentheses indicate the preselection used in the final cut determination.

ϕ products		
p [MeV/c]	$> 2000.$	$(1000.)$
p_T [MeV/c]	$>$ 300.	$(200.)$
sIPS	$>$ 2.	$(1.)$

ϕ		
δm [MeV/c^2]	$\pm$ 20.	$(50.)$
p_T [MeV/c]	$> 1000.$	$(-)$
χ^2 vertex	$<$ 49.	$(100.)$

γ		
E_T [MeV]	$> 2800.$	$(2000.)$

B_s		
δm [MeV/c^2]	$\pm$ 500.	$(500.)$
FD [mm]	$>$ 1.	$(0.)$
sIPS	$<$ 3.5	$(9.)$
$\cos\theta_{p,F}$	$>$ 0.999	(0.995)

Table 4.19: D^* final HLT selection criteria. The rightmost values in parentheses indicate the preselection used in the final cut determination.

D^0 products		
p [MeV/c]	$> 2000.$	$(1000.)$
p_T [MeV/c]	$>$ 300.	$(200.)$
sIPS	$>$ 2.	$(1.)$

D^0		
δm [MeV/c^2]	$+$ 50.	$(100.)$
δm [MeV/c^2]	$-$ 700.	$(1300.)$
p_T [MeV/c]	$> 1250.$	$(500.)$
χ^2 vertex	$<$ 16.	$(100.)$
FS	$>$ 6.	$(-)$

Slow π (upstream and long tracks)		
sIPS	$>$ 1.	$(-)$

D^*		
δm [MeV/c^2]	$+$ 50.	$(100.)$
δm [MeV/c^2]	$-$ 600.	$(1200.)$
$m(D^* - D^0)$ [MeV/c^2]	$<$ 10.	$(60.)$
p_T [MeV/c]	$> 1250.$	$(500.)$
χ^2 vertex	$<$ 16.	$(100.)$

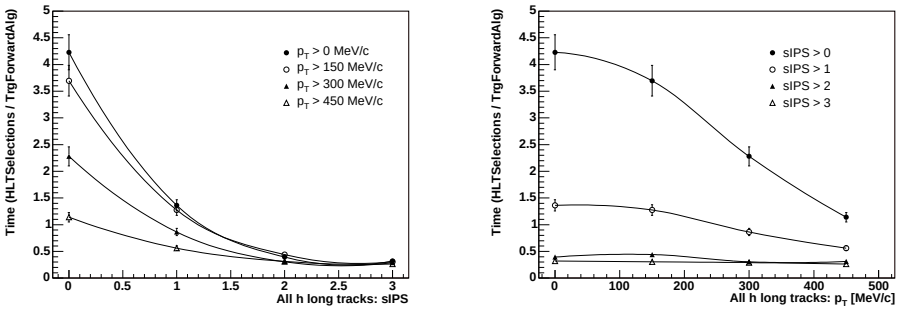


Figure 4.5: HLT selections (`HLTSelections`) execution time normalized to the forward tracking algorithm (`TrgForwardAlg`) CPU time on minimum-bias events. Left: time as a function of the sIPS cut on all long hadron tracks. Right: time as a function of the p_{T} [MeV/c] cut on all long hadron tracks. The errors $\sim 7.8\%$ represent the spread of the normalized time distribution weighted by the mean in a series of identical jobs.

the forward tracking algorithm since it is the most time consuming part of the HLT. The measured CPU time includes the creation and refinement of the final states and shared composite particles, as well as all the exclusive b and D^* selections. The normalized time is not representative of the relative times between the HLT selections and the forward tracking algorithm since the efficiency of the generic HLT on minimum-bias events ($\sim 21.7\%$) is not taken into account. Moreover, we consider the time to make the forward tracking on all tracks. The relevant observable is therefore the relative normalized times between different p_{T} and sIPS cuts. A detailed comparison of the timing of all the HLT parts will be given in Section 4.4.2.2.

The combined HLT exclusive b output rates as a function of the p_{T} and sIPS cuts are shown in Figure 4.6. The different shapes are similar to what is obtained for the timing.

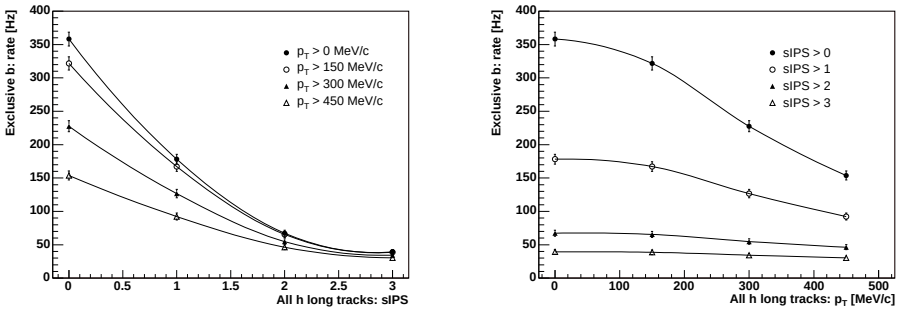


Figure 4.6: Overall HLT exclusive b selections retention rate on minimum-bias events. Left: rate as a function of the sIPS cut on all long hadron tracks. Right: rate as a function of the p_{T} [MeV/c] cut on all long hadron tracks. The uncertainties are statistical.

The performance of the selection criteria of Table 4.7 on the signal efficiencies strongly depends on the channel, and on whether or not these kinds of cuts are applied off-line. This is discussed in detail in [78]. We summarize in Table 4.20 the additional loss in the exclusive dedicated efficiency for the HLT channels involving hadrons. These ineffi-

ciencies represent the losses between the case where we apply all HLT selection criteria except for the p_{T} and sIPS requirements, and the case where we add these p_{T} and sIPS cuts. The efficiencies are with respect to off-line selected events and after the generic HLT. The large loss for some channels is due to the off-line selection keeping regions of phase-space not accessible in the trigger, typically low momentum tracks. Note that the tracking performance has not been factorized out. Finally, it is worth mentioning that the p_{T} knob is probably the safest parameter to tune to control the output rate, given the good momentum resolution and the fact that the on-line errors are not perfectly estimated.

Table 4.20: Exclusive HLT efficiency loss when adding to the HLT cuts the $p_{\mathrm{T}} > 300$ MeV/c and sIPS > 2 preliminary selection criteria. The efficiencies are from the dedicated exclusive selection only and on off-line selected events passing the generic HLT.

Channel	Loss [%]
$B_s \to D_s h$	5.5 ± 0.3
$B_d \to D^0 K^{*0}$	10.0 ± 1.0
$B_s \to \phi\phi$	3.6 ± 0.6
$B_d \to D^* \pi$	5.5 ± 0.6
$B_d \to \mu^+\mu^- K^{*0}$	9.4 ± 0.9
$B_s \to \phi\gamma$	1.1 ± 0.4

4.4 HLT Selection Performance

The exclusive selections should reproduce as much as possible the off-line selections, with softer requirements whenever doable. The choice of standard and limited cuts to ensure the stability of the HLT selections, as described in Section 4.3, will inevitably lead to some inefficiencies as the off-line selections use very different types of cuts. Moreover, as we have seen in Section 4.3 the preliminary filtering of the final states required for timing and output rate is an important source of inefficiency.

The HLT efficiencies for all the core channels are shown in Section 4.4.1 and the different sources of inefficiency are further described for each HLT selection. In Section 4.4.2 we present the results obtained on minimum-bias, such as retention rates, the quark contents and the timing performance.

4.4.1 HLT Selection Performance on Signal Events

The selection efficiencies are determined on events selected by one or more off-line selections relevant to the HLT selection under study, and passing the generic HLT.

The number of events considered and all the HLT efficiencies on signal are summarized in Tables 4.21 and 4.22. The different definitions are given in Section 3.5.3. Note that the quoted HLT efficiencies include the on-line track-finding efficiencies.

Table 4.21: The number of events selected by the dedicated HLT selections and the corresponding specific efficiency on $N_{\rm off}^{\rm evt}$ signal events after the generic HLT. The uncertainties are statistical.

Channel	$N_{\rm off}^{\rm evt}$	$N_{\rm HLT}^{\rm specific}$	$\varepsilon_{\rm HLT}^{\rm specific}$ [%]	$N_{\rm HLT}^{\rm exb}$	$N_{\rm HLT}^{\rm b\to\mu}$	$N_{\rm HLT}^{\mu\mu}$	$N_{\rm HLT}^{\rm D^*}$	$N_{\rm HLT}^{\rm tot}$
$B_s \to D_s h$	5701	4400	77.2 ± 0.6	4407	206	26	196	4496
$B_q \to J/\psi X$	3008	1853	61.6 ± 0.9	2197	1389	2449	84	2796
$B_q \to hh$	4957	4637	93.5 ± 0.4	4638	134	11	59	4652
$B_d \to D^0 K^{*0}$	996	651	65.4 ± 1.5	653	21	0	50	675
$B_s \to \phi\phi$	986	782	79.3 ± 1.3	787	37	3	26	794
$B_d \to D^*\pi$	1568	1060	67.6 ± 1.2	1063	59	7	996	1161
$B_d \to \mu^+\mu^- K^{*0}$	1074	775	72.2 ± 1.4	784	671	619	33	1015
$B_s \to \mu^+\mu^-$	1645	1486	90.3 ± 0.7	1573	1567	1459	26	1639
$B_s \to \phi\gamma$	855	762	89.1 ± 1.1	762	29	0	15	771

Table 4.22: HLT efficiencies on generic HLT triggered events and after the dedicated offline selections. The uncertainties are statistical.

Channel	$\varepsilon_{\rm HLT}^{\rm exb}$ [%]	$\varepsilon_{\rm HLT}^{\rm b\to\mu}$ [%]	$\varepsilon_{\rm HLT}^{\mu\mu}$ [%]	$\varepsilon_{\rm HLT}^{\rm D^*}$ [%]	$\varepsilon_{\rm HLT}^{\rm tot}$ [%]
$B_s \to D_s h$	77.3 ± 0.6	3.6 ± 0.3	0.5 ± 0.1	3.4 ± 0.2	78.9 ± 0.5
$B_q \to J/\psi X$	73.0 ± 0.8	46.2 ± 0.9	81.4 ± 0.7	2.8 ± 0.3	93.0 ± 0.5
$B_q \to hh$	93.6 ± 0.4	2.7 ± 0.2	0.2 ± 0.1	1.2 ± 0.2	93.9 ± 0.3
$B_d \to D^0 K^{*0}$	65.6 ± 1.5	2.1 ± 0.5	0.0 ± 0.0	5.0 ± 0.7	67.8 ± 1.5
$B_s \to \phi\phi$	79.8 ± 1.3	3.8 ± 0.6	0.3 ± 0.2	2.6 ± 0.5	80.5 ± 1.3
$B_d \to D^*\pi$	67.8 ± 1.2	3.8 ± 0.5	0.5 ± 0.2	63.5 ± 1.2	74.0 ± 1.1
$B_d \to \mu^+\mu^- K^{*0}$	73.0 ± 1.4	62.5 ± 1.5	57.6 ± 1.5	3.1 ± 0.5	94.5 ± 0.7
$B_s \to \mu^+\mu^-$	95.6 ± 0.5	95.3 ± 0.5	88.7 ± 0.8	1.6 ± 0.3	99.6 ± 0.2
$B_s \to \phi\gamma$	89.1 ± 1.1	3.4 ± 0.6	0.0 ± 0.0	1.8 ± 0.5	90.2 ± 1.0

Exclusive $B_s \to D_s h$

As we always reconstruct $B_s \to D_s h$ with $h = \pi$ the performance is checked on both $B_s \to D_s \pi$ and $B_s \to D_s K$ events with samples of approximatively the same size not to have any flavor bias. As given in Table 4.10 the mass window compared to other channels is enlarged to $\pm 600\,\mathrm{MeV/c^2}$ around the nominal B_s mass to account for the mass shift induced by the wrong PID hypothesis. This effect can be seen in Figure 4.7 where the mass distribution on $B_s \to D_s K$ is moved to lower values due to the underestimation of the bachelor mass with the appearance of a long negative tail. For $B_s \to D_s \pi$ data the on-line mass resolution is $\sim 29.4\mathrm{MeV/c^2}$ to be compared to the off-line average $B_s \to D_s h$ resolution of $\sim 13.2\,\mathrm{MeV/c^2}$. Given the π–K mass difference and this resolution, the mass window is appropriate.

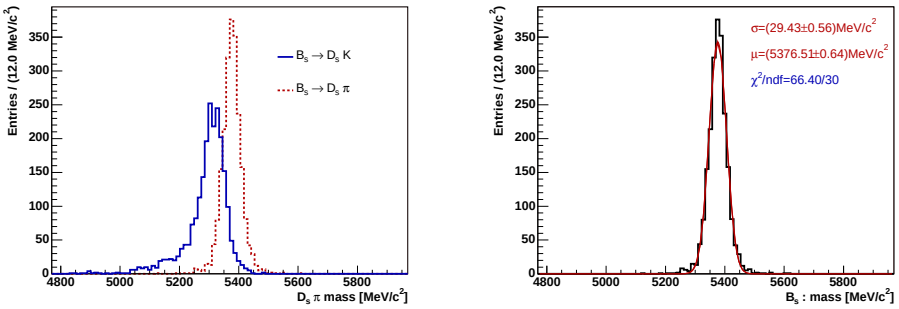


Figure 4.7: On-line reconstructed mass [$\mathrm{MeV/c^2}$] of B_s signal candidates after the dedicated $B_s \to D_s h$ exclusive HLT selection applied on off-line selected events passing the generic HLT. Left: reconstructed mass on true $B_s \to D_s \pi$ (dashed red line) and $B_s \to D_s K$ (solid blue line) events. The histograms are not normalized to the relative branching fractions. Right: on-line reconstructed mass on true $B_s \to D_s \pi$ events with a Gaussian fit.

The efficiency of the cheated selection after the generic HLT is $\varepsilon_{\mathrm{HLT}}^{\mathrm{cheated}} \sim 86\%$ giving a measure of the efficiency of having all final states on-line reconstructed, which is a bit smaller than the value given in Section 4.2.2. This is due to the fact that for the computation of $\varepsilon_{\mathrm{HLT}}^{\mathrm{cheated}}$ we do not require all the final states to be off-line reconstructed (i.e. the off-line selection may select fake combinations in the signal sample).

We notice that the other exclusive selections do not contribute significantly to the overall $\varepsilon_{\mathrm{HLT}}^{\mathrm{exb}}$ efficiency. This is not surprising as $B_s \to D_s h$ is the only core channel with a displaced tertiary vertex with three tracks. There is nevertheless a small correlation with the exclusive $B_d \to D^0 K^{*0}$ stream. The latter selects $\sim 3.6\%$ of the events triggered by the exclusive $B_s \to D_s h$, partly due to the non-negligible D^0 lifetime.

Exclusive $B_q \to J/\psi X$

For this channel we do not actually look for the X but directly rely on the presence of a di-muon pair, as provided by the generic HLT. The efficiency is expected to be far from 100% since we are limited by the rate of true di-muons and the available bandwidth. Strict cuts are thus applied keeping however very large sidebands with a mass cut of $\pm 120\,\mathrm{MeV/c^2}$ around the J/ψ central mass. The on-line and off-line reconstructed masses are shown in Figure 4.8. The on-line J/ψ mass resolution is $\sim 19.2\,\mathrm{MeV/c^2}$ whereas the off-line mass

resolution is $\sim 10.8\,\mathrm{MeV/c^2}$.

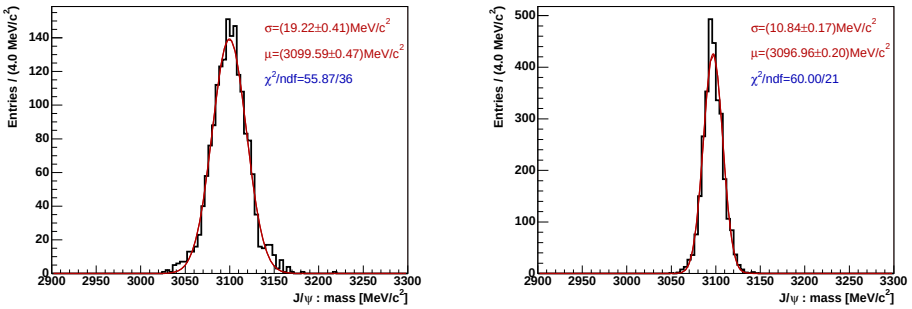


Figure 4.8: Reconstructed mass [$\mathrm{MeV/c^2}$] of $\mathrm{J/\psi}$ signal candidates after the dedicated exclusive HLT selection applied on off-line selected $\mathrm{B_s \to J/\psi\phi}$ events passing the generic HLT. Right: off-line reconstructed mass . Left: on-line reconstructed mass [$\mathrm{MeV/c^2}$]. The histograms are fitted with a single Gaussian.

The efficiency of the generic HLT for $\mathrm{B_s \to J/\psi\phi}$ is $\varepsilon_{\mathrm{HLTGen/L1}} \sim 95.7\%$, higher than for purely hadronic channels thanks to the muon streams.

We see that the other exclusive selections contribute significantly to the $\mathrm{B_q \to J/\psi X}$ efficiency, essentially due to $\mathrm{B_d \to \mu^+\mu^- K^{*0}}$. Note that the di-muon efficiency is really close to the cheated selection efficiency meaning that we basically always trigger when we have a true di-muon. $\mathrm{B_q \to J/\psi X}$ is much less affected by the tracking issues since it can benefit from the single muon line, that is not necessarily requiring the two muons to be reconstructed. The total HLT efficiency is therefore better compared to other large multiplicity channels.

Exclusive $\mathrm{B_q \to hh}$

The exclusive $\mathrm{B_q \to hh}$ is meant to select all $\mathrm{B_d \to \pi^+\pi^-}$, $\mathrm{B_d \to K^+\pi^-}$, $\mathrm{B_s \to K^+K^-}$, and $\mathrm{B_s \to \pi^+K^-}$ modes in one selection by only looking at $\mathrm{h = \pi}$, and always reconstructing a $\mathrm{B_d}$ ($\mathrm{q = d}$). We must therefore be careful in the choice of the mass window since we must consider the effect of the π–K *and* the $\mathrm{B_d}$–$\mathrm{B_s}$ mass differences. The choice of the mass window is $\pm 600\ \mathrm{MeV/c^2}$ around the nominal $\mathrm{B_d}$ mass based on the plot of Figure 4.9. We can see how is the reconstructed mass affected when using the false flavor hypothesis for each type of $\mathrm{B_q \to hh}$ decay. For $\mathrm{B_d \to \pi^+\pi^-}$ there is no effect and the on-line mass resolution is $\sim 33.4\ \mathrm{MeV/c^2}$, compared to $\sim 16.5\ \mathrm{MeV/c^2}$ in the off-line case. The lowest invariant masses are for $\mathrm{B_d \to K^+\pi^-}$ as we underestimate the kaon mass. The largest invariant masses correspond to $\mathrm{B_s \to \pi^+K^-}$ when we underestimate the kaon mass compensated by the $\mathrm{B_d}$–$\mathrm{B_s}$ mass difference, whereas for $\mathrm{B_s \to K^+K^-}$ the two misidentifications and the wrong $\mathrm{B_q}$ flavor results in a value close to the $\mathrm{B_d}$ mass.

For $\mathrm{B_q \to hh}$ practically all the events are triggered by the exclusive selection. The HLT efficiency is high, as these low multiplicity channels are much less affected by the tracking inefficiencies and as the off-line requires high IP and p_{T} tracks. Only a few percent are lost compared to the cheated selection, meaning that the HLT selection is quite optimal.

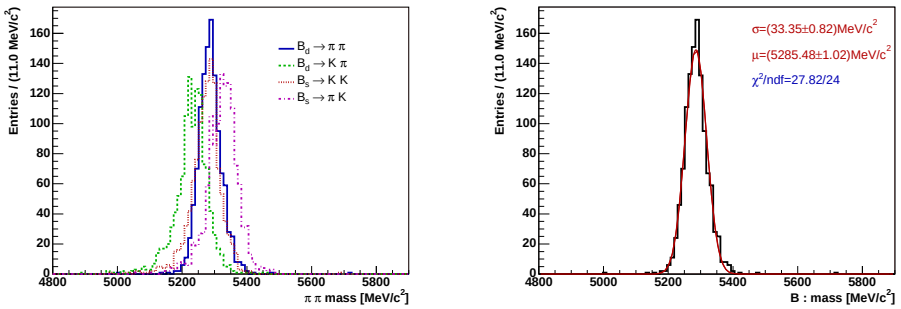


Figure 4.9: On-line reconstructed mass [MeV/c^2] of B_q signal candidates after the dedicated $B_q \to hh$ exclusive HLT selection applied on off-line selected events passing the generic HLT. Left: contributions from true $B_d \to \pi^+\pi^-$ (blue solid line), $B_d \to K^+\pi^-$ (green dotted line), $B_s \to K^+K^-$ (dashed red line), and $B_s \to \pi^+K^-$ (dashed-dotted magenta line) events. The histograms are not normalized to the relative branching fractions. Right: true $B_d \to \pi^+\pi^-$ events with a Gaussian fit.

Exclusive $B_d \to D^0K^{*0}$

The performance for $B_d \to D^0K^{*0}$ is checked against off-line selected $B_d \to D^0(\to K^+\pi^-)$ $K^{*0}(\to K^+\pi^-)$ events. The on-line B_d mass resolution is ~ 24.8 MeV/c^2 and the off-line one is ~ 11.3 MeV/c^2.

This channel has the worse exclusive selection efficiency among all the core channels. The difference compared to $B_s \to D_sh$ is more than 10% and it is mostly due to the loss caused by the preliminary selection on the final states, see Table 4.7.

Exclusive $B_s \to \phi\phi$

The mass resolutions in the $B_s \to \phi\phi$ selection are ~ 21.1 MeV/c^2 and ~ 11.5 MeV/c^2 respectively for the on-line and off-line cases. The cheated selection efficiency is $\varepsilon_{\rm HLT}^{\rm cheated} \sim 83.1\%$, smaller than the tracking efficiency quoted previously due to the purity of the off-line events, of the order of $\sim 95.6\%$ (i.e. the off-line selection does not always select true off-line reconstructed tracks).

The exclusive HLT efficiency for this channel can hardly be improved by a smarter choice of selection cuts, since we can see that only a few percent of the events are lost because of the selection cuts when compared to the cheated selection efficiency. The main loss is due to the $p_{\rm T}$ cuts on the ϕ daughters.

Exclusive $B_d \to D^*\pi$

The selection of $B_d \to D^*\pi$ makes use of both upstream and long tracks for the slow pion from the D^*. As upstream tracks get their rough momentum estimate from the VELO-TT matching, the resulting momentum resolution is at least an order of magnitude worse compared to what is achieved for long tracks. The contribution of VELO-TT tracks translates to non-Gaussian tails in the invariant D^* mass distribution shown in Figure 4.10. The resulting mass including the two type of tracks is ~ 12.5 MeV/c^2 and ~ 5.0 MeV/c^2 respectively for the on-line and off-line cases. The contribution of upstream and long

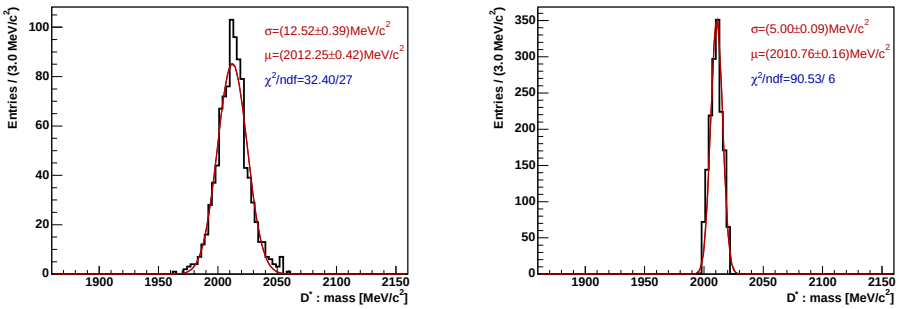


Figure 4.10: Reconstructed mass [$\mathrm{MeV/c^2}$] of $\mathrm{D^*}$ signal candidates after the dedicated exclusive HLT selection applied on off-line selected $\mathrm{B_d} \to \mathrm{D^*}\pi$ events passing the generic HLT. Right: off-line reconstructed mass. Left: on-line reconstructed mass. The histograms are fitted with a single Gaussian.

tracks can be seen on Figure 4.11 where the reconstructed $\mathrm{B_d}$ mass is plotted. Though it is small compared to the long track case, we see a non-negligible fraction of the candidates formed with upstream tracks. These tracks clearly contribute to the mass with a larger resolution compared to the forward matched tracks. The core resolution assuming a single contribution is for the on-line $\mathrm{B_d}$ mass $\sim 29.8\,\mathrm{MeV/c^2}$ and $\sim 14.7\,\mathrm{MeV/c^2}$ for the off-line. These resolutions are obtained on $\mathrm{B_d} \to \mathrm{D^*}(\to \mathrm{D^0}(\to \mathrm{K^+}\pi^-)\pi)\pi$ events.

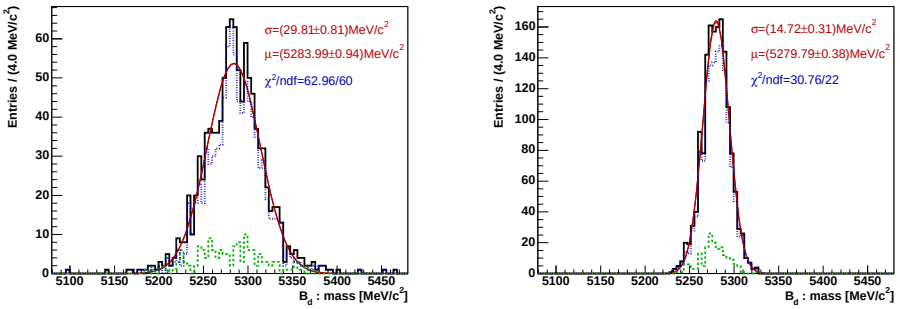


Figure 4.11: Reconstructed mass [$\mathrm{MeV/c^2}$] of $\mathrm{B_d}$ signal candidates after the dedicated exclusive HLT selection applied on off-line selected $\mathrm{B_d} \to \mathrm{D^*}\pi$ events passing the generic HLT. The contributions from the slow pions made from long (dashed blue line) and from upstream (dotted green line) tracks are superimposed. Right: off-line reconstructed mass. Left: on-line reconstructed mass. The histograms are fitted with a single Gaussian.

This channel clearly benefits from the $\mathrm{D^*}$ stream. Besides the tracking inefficiency, this channel suffers from large combinatorics introduced by the slow pion on which no cuts are applied. Indeed, cutting on this particle would reduce significantly the signal. The $\mathrm{D^*}$–$\mathrm{D^0}$ mass window was chosen to approximatively correspond to the mass difference between the $\mathrm{D^*}$ and $\mathrm{D^0}\pi$.

Exclusive $B_d \to \mu^+\mu^- K^{*0}$

The performance for $B_d \to \mu^+\mu^- K^{*0}$ is checked against $B_d \to \mu^+\mu^- K^{*0}(\to K^+\pi^-)$ events. The on-line B_d mass resolution is $\sim 25.2\,\mathrm{MeV/c^2}$ and the off-line one is $\sim 14.4\,\mathrm{MeV/c^2}$.

For this channel the preliminary cuts on the hadrons appear to be important sources of inefficiency. Even though the HLT streams are far from being optimal, a pretty high total HLT efficiency can be achieved, thanks to the combination of all HLT streams.

Exclusive $B_s \to \mu^+\mu^-$

This channel is expected to be the easiest one regarding the trigger selection, with no combinatorics at all. Indeed, requiring two high p_T muons with invariant mass close to the B_s mass and well above the J/ψ mass will easily kill all the background. The mass resolutions for the on-line and off-line B_s are respectively $\sim 29.1\,\mathrm{MeV/c^2}$ and $\sim 18.4\,\mathrm{MeV/c^2}$ as shown in Figure 4.12.

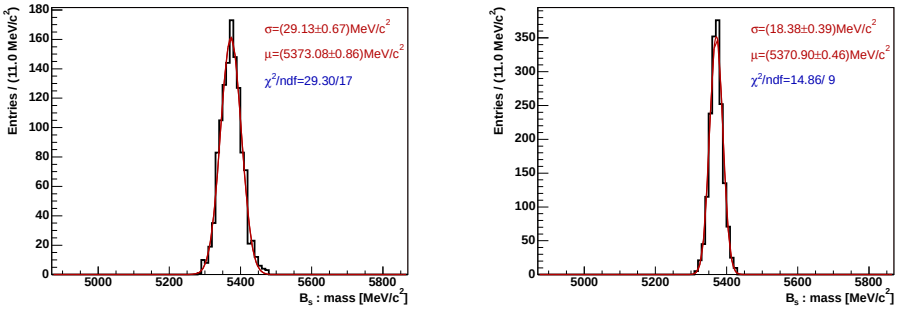


Figure 4.12: Reconstructed mass [$\mathrm{MeV/c^2}$] of B_s signal candidates after the dedicated exclusive HLT selection applied on off-line selected $B_s \to \mu^+\mu^-$ events passing the generic HLT. Right: off-line reconstructed mass. Left: on-line reconstructed mass. The histograms are fitted with a single Gaussian.

When combining all the streams we see that the total HLT efficiency $\approx 100\%$. Looking at the dedicated exclusive selection $\varepsilon_{\mathrm{HLT}}^{\mathrm{specific}}$ we notice that we actually loose a large fraction of the events. The total exclusive efficiency $\varepsilon_{\mathrm{HLT}}^{\mathrm{exb}}$ is much higher than the dedicated one as we add the result of the $B_q \to hh$ exclusive selection that triggers a significant fraction of the signal events, approximatively $\sim 88.5\%$.

Exclusive $B_s \to \phi\gamma$

The quality of the photon reconstruction in the HLT is within errors compatible with the off-line one as it can be seen from Figure 4.13. Including the stochastic and calibration contributions the photon energy resolutions are 2.58% and 2.50% for the on-line and off-line, respectively. We note the better calibration in the on-line parameters (as it can be seen from the mean of the mass distributions). These photon resolutions translate to the observed invariant B_s mass resolutions shown in Figure 4.14. For the on-line we have $\sim 69.3\,\mathrm{MeV/c^2}$ and for the off-line it is slightly better, $\sim 64.5\,\mathrm{MeV/c^2}$, due to the charged tracks, but we are in both cases clearly dominated by the photon contribution.

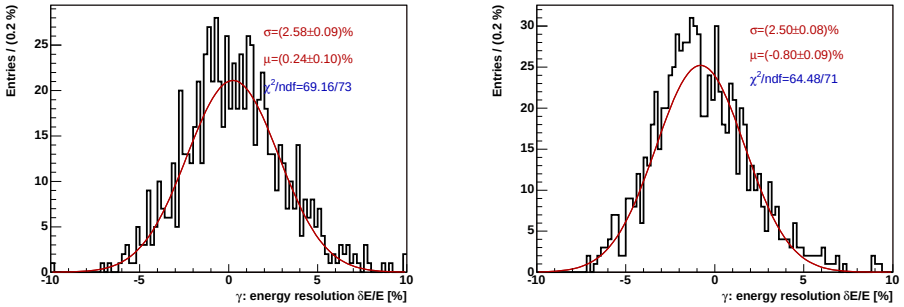


Figure 4.13: Residual of the photon energy from the B_s signal candidates after the dedicated exclusive HLT selection applied on off-line selected $B_s \rightarrow \phi\gamma$ events passing the generic HLT. Right: off-line γ relative energy resolution [%]. Left: on-line γ relative energy resolution [%]. The histograms are fitted with a single Gaussian.

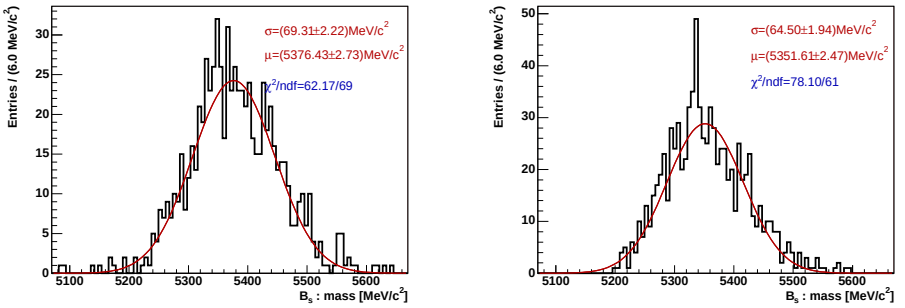


Figure 4.14: Reconstructed mass [MeV/c^2] of B_s signal candidates after the dedicated exclusive HLT selection applied on off-line selected $B_s \rightarrow \phi\gamma$ events passing the generic HLT. Right: off-line reconstructed mass. Left: on-line reconstructed mass. The histograms are fitted with a single Gaussian.

The efficiency of the generic HLT is only $\varepsilon_{\mathrm{HLTGen/L1}} \sim 73.4\%$. This is merely due to the absence of ECAL trigger in the generic HLT and hence we lose all the bandwidth from the L1 photon line. The cheated selection efficiency is $\varepsilon_{\mathrm{HLT}}^{\mathrm{cheated}} \sim 95.1\%$, including the photon reconstruction. Basically all the photons found in the off-line are found in the on-line as well, hence leaving us with the usual tracking inefficiencies. Note that the presence of a ϕ further worsen the track-finding efficiency, as opposed to the $\mathrm{B_q} \to \mathrm{hh}$ channels.

The $\mathrm{B_s} \to \phi\gamma$ HLT selection is challenging as we have a loosely determined ϕ vertex, we are overwhelmed by photons that do not contribute to the position determination, and we cannot apply strict mass cuts to reduce combinatorics. Requiring large transverse momentum and energy for the ϕ and γ enables to reduce the bandwidth while keeping a high efficiency, a few percent smaller than the $\mathrm{B_q} \to \mathrm{hh}$ one.

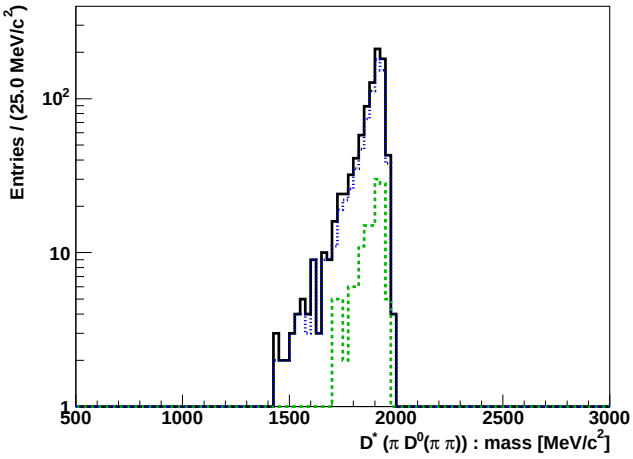


Figure 4.15: On-line reconstructed mass [MeV/c^2] of the selected associated D^* signal candidates after the dedicated D^* stream HLT selection and applied off-line selected $\mathrm{B_d} \to \mathrm{D}^*\pi$ events passing the generic HLT. The contributions from the slow pions made from long (dashed blue line) and from upstream (dotted green line) tracks are superimposed.

Exclusive $\mathrm{D}^* \to \mathrm{D}^0\pi$ stream

We use the $\mathrm{B_d} \to \mathrm{D}^*\pi$ decay channel, with $\mathrm{D}^0 \to \mathrm{K}^+\pi^-$ to check the performance on signal of the D^* stream. In this stream the D^* is always reconstructed assuming the pion hypothesis for all the stable decay products, using the $\mathrm{D}^0 \to \pi^+\pi^-$ mode. There are loose asymmetric mass cuts on both the D^* and the D^0 as given in Table 4.19. Given their mass resolutions, the upper bound corresponds to a cut of more than four standard deviations. The lower limit is set to remove the long negative tails, though we have to be careful since we underestimate the mass for the $\mathrm{D}^0 \to \mathrm{K}^+\pi^-$ and $\mathrm{D}^0 \to \mathrm{K}^+\mathrm{K}^-$ cases as we only consider the D^0 to two pions decays in the D^* stream. The reconstructed D^* mass distribution on signal with $\mathrm{D}^0 \to \mathrm{K}^+\pi^-$ is shown in Figure 4.15, where the contributions from long and upstream tracks are also drawn. We see that almost all signal candidates fall in the asymmetric mass window, which justifies our choice. The D^0 mass cuts have also been tested for the $\mathrm{D}^0 \to \mathrm{K}^+\mathrm{K}^-$ case, using $\mathrm{B_d} \to \mathrm{D}^0\mathrm{K}^{*0}$ data. The signal candidates

falling in the negative tails have been found to be badly reconstructed. Moreover, these candidates also fall in the upper tails of the distribution of the D^*–D^0 mass difference because of the strong correlation induced by the small available phase space for the slow pion.

The efficiency of the D^* stream is 63.5% on $B_d \to D^*\pi$ signal events, see Table 4.22, which is a few percent smaller than the dedicated counterpart. Though we have a much larger available bandwidth for the D^* stream, we need to cut much harder if we want to keep loose mass windows. Moreover, we do not have the extra constraints of the fast pion and of forming a B_d to suppress the background. In particular, we have to impose a sIPS cut on the slow pion to reject the large combinatorics introduced by the pions. Finally, the exclusive and D^* streams are not completely correlated as the combination of the two results in a significant gain for the total efficiency on the $B_d \to D^*\pi$ signal channel.

Summary

As we can see from all the signal efficiencies in Table 4.22, the performance of the HLT can be very different depending on the decay channel we consider. The five dominant sources of performance differences identified are:

1. The on-line track-finding. A large fraction of the signal events is lost because the tracks are not reconstructed. We also observe inefficiencies in the reconstruction and identification of muons.

2. The type of decay. Having at least a muon in the decay chain helps to cover for the tracking inefficiencies as we benefit from the generous bandwidth granted to muons. Channels with large combinatorics, such as those involving a D_s, will undergo larger losses as we need to cut harder to suppress the background.

3. The on-line–off-line reconstruction quality differences. The momentum and IP resolutions, the vertexing, and more importantly the difference in tracking errors lead to inefficiencies that increase with the decay multiplicity.

4. The primary vertex reconstruction. A missed primary vertex gives rise to large combinatorics possibly faking the presence of a secondary vertex.

5. The on-line–off-line selection differences. The preliminary cuts imposed to all long tracks result in a loss of events due to the quality difference and to the choice of off-line selection cuts. Selections applying very loose such cuts, in particular the hadronic channels, will inevitably lose events since we anyways apply these cuts in the HLT. The use of common standard composite particles with momentum cuts on the decay products will certainly indirectly improve the HLT selection efficiencies.

For completeness we give in Table 4.23 a comparison between the on-line and off-line mass resolutions involved in the HLT selections, and that were not already given in Table 4.9.

4.4.2 HLT Selections Performance on Minimum-Bias Events

The HLT performance on minimum-bias events can now be assessed using the whole trigger chain. For that we consider 131'320 minimum-bias events (DC04-v2) at 40 kHz after L0 and L1, which is equivalent to ~ 3.28 s of data taking. The HLT selection criteria

Table 4.23: Mass resolutions σ_m of the on-line and off-line composite particles and b hadrons. The quoted values are from single Gaussian fits.

Channel	On-line σ_m [MeV/c^2]	Off-line σ_m [MeV/c^2]
$B_s \to D_s h$	29.4	13.2
$J/\psi(B_q \to J/\psi X)$	19.2	10.8
$B_q \to hh$	33.4	16.5
$B_d \to D^0 K^{*0}$	24.8	11.3
$B_s \to \phi\phi$	21.1	11.5
D^*	12.5	5.0
$B_d \to D^* \pi$	29.8	14.7
$B_d \to \mu^+\mu^- K^{*0}$	25.2	14.4
$B_s \to \mu^+\mu^-$	29.1	18.4
$B_s \to \phi\gamma$	69.3	64.5

have been tuned on a different sample (DC04-v1) and are blindly applied. The generic HLT already removes an important part of the minimum-bias events, as we are left with about 9 kHz after it, including the 1.5 kHz from the muons streams. Hence we see that we still need to reduce the rate by at least a factor 15 to bring the rate for the HLT selections down to 500 Hz. This is actually the tough part given the charm and beauty contents after the generic HLT.

4.4.2.1 HLT Output Rates

The selection criteria were tuned to have at most ~ 10 Hz for the exclusive selections individually, whereas all the bandwidth for the D^* stream is used. The number of selected events in the minimum-bias sample and the corresponding output rates are given in Tables 4.24 and 4.25. The rates of the other HLT streams ($b \to \mu$, $\mu\mu$ and D^*) for events selected by the dedicated HLT selection are given to show the correlations. We see from these tables that the maximal exclusive b output rate is (57.0 ± 4.2) Hz, i.e. when adding all exclusive b selected events and ignoring the correlations between them. The exclusive b rate is thus well within specification. For the D^* stream we have an output rate just below the nominal 300 Hz.

The exclusive selections involving muons obviously have large correlations with the muon streams. The overlap is very large for the $B_q \to J/\psi X$ due to the high rate of true J/ψ. The largest bandwidth among the purely hadronic channels is taken by $B_s \to D_s h$, because of the numerous D_s combinations. In general there is a very small correlation between the exclusive b selections and the muons streams.

We give in Table 4.26 the total output rates for the whole HLT, assuming 40 kHz L1 output rate. The overall HLT retention rate on minimum-bias events is below the design 2 kHz. The exclusive b output rate is obtained by taking the logical *OR* of the different output rates given in Table 4.25. All the selection efficiencies on minimum-bias events for each of the HLT streams, channels and their correlations can be found in [78].

Table 4.24: Number of selected events $N^{specific}$ by each of the specific HLT selections on 131'320 L0 + L1 stripped minimum-bias events, with the result of the other streams $N^{specific}_{b\to\mu}$, $N^{specific}_{\mu\mu}$ and $N^{specific}_{D^*}$ on these $N^{specific}$ events.

Channel	$N^{specific}$	$N^{specific}_{b\to\mu}$	$N^{specific}_{\mu\mu}$	$N^{specific}_{D^*}$
$B_s \to D_s h$	36	2	2	4
$B_q \to J/\psi X$	55	18	48	1
$B_q \to hh$	12	2	0	1
$B_d \to D^0 K^{*0}$	15	0	0	3
$B_s \to \phi\phi$	0	0	0	0
$B_d \to D^*\pi$	34	1	1	6
$B_d \to \mu^+\mu^- K^{*0}$	7	2	7	0
$B_s \to \mu^+\mu^-$	0	0	0	0
$B_s \to \phi\gamma$	28	1	2	3
D^* stream	964	18	71	964

Table 4.25: Output rates $R^{specific}$ of each of the specific HLT selections on 131'320 L0 + L1 stripped minimum-bias events, with the shared rates of the other streams $R^{specific}_{b\to\mu}$, $R^{specific}_{\mu\mu}$ and $R^{specific}_{D^*}$ on the $R^{specific}$. The uncertainties are statistical.

Channel	$R^{specific}$ [Hz]	$R^{specific}_{b\to\mu}$ [Hz]	$R^{specific}_{\mu\mu}$ [Hz]	$R^{specific}_{D^*}$ [Hz]
$B_s \to D_s h$	11.0 ± 1.8	0.6 ± 0.4	0.6 ± 0.4	1.2 ± 0.6
$B_q \to J/\psi X$	16.8 ± 2.3	5.5 ± 1.3	14.6 ± 2.1	0.3 ± 0.3
$B_q \to hh$	3.7 ± 1.1	0.6 ± 0.4	0.0 ± 0.0	0.3 ± 0.3
$B_d \to D^0 K^{*0}$	4.6 ± 1.2	0.0 ± 0.0	0.0 ± 0.0	0.9 ± 0.5
$B_s \to \phi\phi$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
$B_d \to D^*\pi$	10.4 ± 1.8	0.3 ± 0.3	0.3 ± 0.3	1.8 ± 0.8
$B_d \to \mu^+\mu^- K^{*0}$	2.1 ± 0.8	0.6 ± 0.4	2.1 ± 0.8	0.0 ± 0.0
$B_s \to \mu^+\mu^-$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
$B_s \to \phi\gamma$	8.5 ± 1.6	0.3 ± 0.3	0.6 ± 0.4	0.9 ± 0.5
D^* stream	293.6 ± 9.4	21.6 ± 2.6	9.3 ± 1.7	293.6 ± 9.4

Table 4.26: Minimum-bias trigger output rates. The uncertainties are statistical.

Trigger	$N^{\rm evt}$	Rate [Hz]
Level-1	131320	40 kHz reference
Generic HLT	28458	8668.3 ± 45.5
HLT	6355	1935.7 ± 23.7
Exclusive b stream	180	54.8 ± 4.1
b $\to \mu$ stream	2547	775.8 ± 15.2
Di-muon stream	2981	908.0 ± 16.4
D* stream	964	293.6 ± 9.4

An important criterion of the trigger performance is the fraction of interesting events in the selected bandwidth. Each step of the trigger is meant to successively reduce the minimum-bias output rate while enriching its content of b and c quarks as these are the events we wish to reconstruct at LHCb. The quark content of the selected events after each of the HLT streams is given in Table 4.27 (we first look for a b quark, and then for a b or c). The charm component includes all charmed hadrons and not necessarily only the open charm particles. We first notice that the di-muon stream has the lowest "purity" (from b), which is intended as we want all J/$\psi \to \mu^+\mu^-$. The HLT selections, i.e. the exclusive b and the D*, have comparable fractions of events with a b and a b or c quark. A large fraction of the charm component is given by the $B_q \to J/\psi X$ selecting true J/ψ. The detail of the quark content for each exclusive selection can be found in [78]. As we can see from the quark content table, there is still a significantly large fraction of events selected by the HLT selections that do not contain any b or c quarks, of the order of $\sim 30\%$. The selected events without any b or c quark correspond in most cases, and especially for the hadronic modes, to events where the primary vertex finder failed to reconstruct a true collision, and to candidates that are in fact formed with tracks from a missed primary vertex [78].

Table 4.27: Quark contents, b and b or c, of the HLT triggered minimum-bias events.

Trigger	b [%]	b or c [%]
Generic HLT	34.1	56.2
HLT	39.9	66.6
Exclusive b stream	52.2	71.7
b $\to \mu$ stream	70.6	90.5
Di-muon stream	11.9	46.0
D* stream	52.8	73.4

The on-line reconstructed mass of all the selected candidates in minimum-bias events is shown in Figure 4.16 which illustrates the mass windows used in the HLT selections.

The peak centered at the J/ψ mass correspond to the $B_q \rightarrow J/\psi X$ selected candidates, which are all true J/ψ. The exclusive b mass spectrum shows that the selected candidates have an evenly distributed mass due to the combinatorics. Nevertheless, there is a distinctive peaking background at low mass. These candidates are presumably due to partially reconstructed b decays, as we do select more than 50% of the events with a b quark.

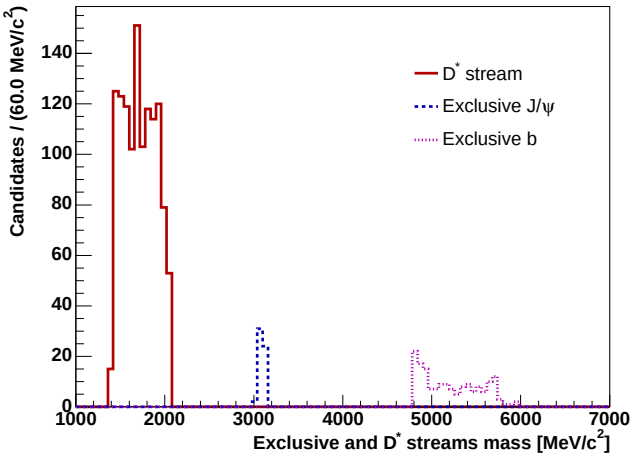


Figure 4.16: Reconstructed mass [MeV/c^2] of the HLT selected candidates with the exclusive b and D^* streams.

4.4.2.2 HLT Timing Performance

We give in this section the timing performance for the entire HLT sequence as measured on minimum-bias events. We assume an input rate after L1 of 40 kHz, with 400 CPUs of the on-line farm dedicated to the on-line reconstruction and the HLT. The average computing time per event should thus not exceed 10 ms. The clock times were determined on a dedicated machine only running the HLT code and are therefore unaffected by the machine load.

In the timing measurement given below, we have taken into account the efficiency of the generic HLT and the fact that the full tracking is not executed for each event after the generic HLT. To be conservative we ignore the fact that some of the tracks were already reconstructed in the generic HLT and we count the time to reconstruct all the tracks after a positive generic HLT decision.

The per-event execution time for the full HLT is measured to be 30.04 ms on average. The error on this average is approximatively 0.05 ms, determined in a series of identical jobs. The contributions from the main components are given below, always normalized to the total time such that the time equivalent value can directly be obtained.

HLT VELO tracking $\sim 20.2\%$. It includes the times for the raw VELO buffer decoding, the VELO $r-z$ and 3D tracking, and the primary vertex reconstruction. The main contribution to this part is:

- VELO-space tracking: $\sim 14.5\%$.

Generic HLT sequence $\sim 52.3\%$. It includes the times for the VELO-TT decoding and tracking, the partial forward tracking, the error parameterization, the muon reconstruction, the generic HLT, and the muon streams decisions. The main contributions to this part are:

- TT decoding: $\sim 3.2\%$;
- VELO-TT tracking: $\sim 5.5\%$;
- Forward decoding: $\sim 5.3\%$;
- Forward tracking: $\sim 34.0\%$.

The following parts are only executed on generic HLT triggered events.

Full tracking $\sim 18.0\%$. It corresponds to the time for the rest of the forward tracking.

HLT particles $\sim 3.4\%$. It represents the time for the final state particles making. The main contributions to this part are:

- ECAL reconstruction: $\sim 2.1\%$;
- Particle making: $\sim 1.3\%$.

HLT selections $\sim 6.1\%$. It includes the times for the final states filtering, the shared composite particles making, the exclusive b and the D^* selections. The relevant contributions to this part are:

- Final states copying and filtering: $\sim 2.1\%$;
- Shared particles: $\sim 2.0\%$;
- D^* selection: $\sim 0.3\%$;
- Exclusive b selections: $\sim 1.6\%$.

The time taken by the HLT selections including particle making amounts to less than 10% of the total time. Adding extra exclusive selections is not an issue as far as timing is concerned as most of the time spent in the HLT selections is in the creation and filtering of the shared particles. These particles will be in most cases already available to the new channels. The most consuming part is by far the forward tracking which represents more than 50% of the total time.

The overall time of 30.04 ms corresponds to the performance for a single process. In the real experiment we will have at least two CPUs running on the *same* machine, i.e. 2 out of 4 cores for each box in the on-line farm dedicated to the HLT. If we compare this time with what will be required during data taking we therefore see that the overall timing is practically in the good range. The performance must however be reevaluated once the new 1 MHz scheme will be implemented.

4.5 Outlook of the Exclusive Trigger Selections

The exclusive b and D^* HLT selections are up and running. The selection criteria have been tuned to achieve the design output rates. These criteria will be revisited by the different physics working groups as new selections are added to the HLT, to optimize the physics goals. The purity of the selected bandwidth in terms of beauty and charm

quarks is above 70%, where the remainder is due to reconstruction problems, such as the primary vertex search. The timing performance is approximatively in the required range.

The total HLT efficiencies on the core signal channels vary in a large range, $70-100\%$, and strongly depend on the type and multiplicity of the decay. The specific efficiencies are essentially limited by the on-line–off-line differences, rather than by a fine tuning of the selection requirements. The restricted processing time demands a fast dedicated on-line software that inevitably leads to HLT inefficiencies. The observed and identified main sources of inefficiency are summarized hereafter:

- On-line and off-line pattern recognitions are different yielding track-finding inefficiencies for the tracks used in the off-line analyses, where the loss becomes important for high-multiplicity decays. The use of TT is fundamental in the current implementation for the speed performance, but yields inefficiencies due to the lack of constraints and its geometry. Without TT and regardless of the timing, the forward tracking still misses a few percent of the off-line reconstructed tracks. The forward tracking is currently under revision for speed and efficiency. Tracking inefficiencies in the muon reconstruction are also observed, though these are covered by the muon high rate streams.

- There was a significant improvement in the primary vertex reconstruction strategy allowing a better discrimination of the minimum-bias background. This was taken into consideration for the tuning of the selection criteria presented in this chapter. However, part of the HLT bandwidth is due to a reconstructed signal candidate in place of a primary vertex.

- The use of a simplified covariance matrix for the track parameters gives a few percent inefficiency for the HLT selections as the errors are different and systematically overestimated compared to the off-line. A more complete parameterization could be implemented (e.g. using a look-up table), but the performance is satisfactory for the HLT needs, where anyway some inefficiency should be tolerated. The geometrical and vertexing tools using this constant parameterized covariance matrix also give reliable results, with a few limitations for vertices significantly displaced with respect to the beam axis (few percent effect).

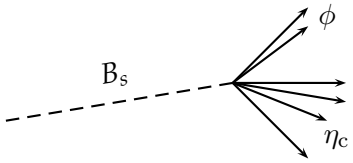
- The HLT uses preliminary sIPS and p_{T} cuts on all the final states long tracks, introduced for speed and output rate reasons.

- The choice of final selection criteria indirectly degrades the performance as the HLT applies a series of standardized cuts in order to form shared composite particles. Thus the use of these kinds of standard particles by the off-line selections would make them more uniform and in turn improve the HLT selection efficiencies.

The first complete HLT prototype is fully implemented yielding reasonable results in terms of efficiency, timing and output rates. The development and study of the HLT selections performance has emphasized a few weaknesses of the on-line software. These issues should be considered and taken into account in the next iteration of the HLT related software, especially now that the L1 boundary has been removed.

Chapter 5

The $B_s \to \eta_c \phi$ Event Selection



This chapter describes the reconstruction and the selection of $B_s \to \eta_c \phi$ events using the full Monte Carlo simulation. The event selection is presented, and the annual yield and background levels are estimated. The trigger and flavor tagging performances on $B_s \to \eta_c \phi$ off-line selected events are determined. Finally, a detailed analysis of the resolutions and of the proper time is provided.

BEFORE performing any CP measurement, we need to reconstruct and select the decays of interest. The first step in the event selection is the trigger selection, followed by the off-line reconstruction of tracks to finally be able to look for the different channels. To illustrate this procedure, we will present the event selection of the $B_s \to \eta_c \phi$ decay channel using the full Monte Carlo simulation described in Chapter 3.

The $B_s \to \eta_c \phi$ decays are caused by $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions to pure CP eigenstates and they may therefore be used to probe the B_s–$\overline{B}_s$ mixing phase. These kinds of quark transitions are of particular interest as they are to a very good approximation dominated by just one CKM tree phase, as the penguin contributions are doubly Cabbibo-suppressed. Their CP asymmetry therefore directly measures the weak mixing phase ϕ_s.

Though being more difficult experimentally compared to $B_s \to J/\psi\phi$, the $B_s \to \eta_c \phi$ decays offer an alternative determination of ϕ_s that does not require a CP angular analysis as opposed to $B_s \to J/\psi\phi$. Indeed, since η_c is a pseudo-scalar ($J^{PC} = 0^{-+}$) and ϕ is a pseudo-vector ($J^{PC} = 1^{--}$), these mesons must have a relative orbital momentum $l = 1$, given that $J^P(B_s) = 0^-$ and because of angular momentum conservation. The resulting final state is then a pure CP-even eigenstate.

The ability of LHCb in reconstructing $B_s \to \eta_c \phi$ events will be determined in terms of yield, background level, resolutions and tagging performance, and all these ingredients will then be used in the determination of ϕ_s in Chapter 6.

5.1 Annual Production

As most of the B_s decay modes, the $B_s \to \eta_c \phi$ branching ratio (BR) has not yet been measured. We can however make an estimate using its counterpart decay in the B_d system,

namely $B_d \to \eta_c K^0$. Both decays originate from the same $\bar{b} \to \bar{c}c\bar{s}$ quark-level tree diagram and are related to each other by interchanging the strange and down quarks. In the spectator quark approximation and rescaling the branching ratios by means of similar decays involving J/ψ, we can write:

$$\frac{BR(B_s \to \eta_c\phi)}{BR(B_s \to J/\psi\phi)} = \frac{BR(B_d \to \eta_c K^0)}{BR(B_d \to J/\psi K^0)}. \quad (5.1)$$

Using the above approximation and the known branching ratios of Table 5.1, we get the following estimate:

$$BR(B_s \to \eta_c\phi) = (1.24 \pm 0.52) \times 10^{-3}, \quad (5.2)$$

which is compatible with the $B_s \to J/\psi\phi$ branching ratio. However, the absence of a di-muon will considerably spoil the trigger and selection performances, given the high-background rejection that needs to be achieved. In order to compensate for these losses and to fully benefit from its pure CP eigenstate feature, the signature of the $B_s \to \eta_c\phi$ decay must be chosen as clean as possible.

Table 5.1: Measured branching ratios used in the estimation of $BR(B_s \to \eta_c\phi)$ [27].

Channel	Branching ratio BR
$B_s \to J/\psi\phi$	$(9.3 \pm 3.3) \times 10^{-4}$
$B_d \to \eta_c K^0$	$(1.16 \pm 0.26) \times 10^{-3}$
$B_d \to J/\psi K^0$	$(8.72 \pm 0.33) \times 10^{-4}$

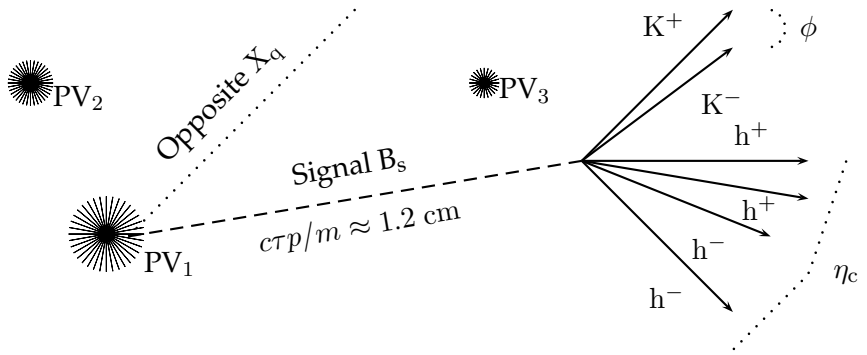


Figure 5.1: Artistic cartoon representation of the $B_s \to \eta_c(h^+h^+h^-h^-)\phi(K^+K^-)$ decay topology, where $h^\pm$ stands for pions or kaons. The scale is arbitrary.

The choice of final states in $B_s \to \eta_c\phi$, that is the decay products of the final mesons, is motivated by the prevailing role played by the proper time resolution in the determination of the B_s–$\overline{B}_s$ mixing phase. As the proper time resolution is dominated by the

secondary vertex resolution, a large multiplicity secondary vertex is chosen. Moreover, the most abundant decays are picked in order to have the highest possible production yields and therefore the largest statistical sensitivity to ϕ_s.

The signature used to reconstruct the $B_s \to \eta_c\phi$ decay is depicted in Figure 5.1. The decay modes used for the final mesons are $\phi \to K^+K^-$ and $\eta_c \to \pi^+\pi^-\pi^+\pi^-$, $\pi^+\pi^-K^+K^-$, $K^+K^-K^+K^-$. Since ϕ and η_c are short-lived resonances, their decay vertices overlap when compared to the experimental position resolution. The corresponding topology is consequently that of a 6-prong. The most distinctive feature is the presence of a detached secondary vertex. As the B_s has $c\tau \approx 418.5\ \mu$m and given its average momentum at LHCb $p \approx 150\,\text{GeV/c}$, the B_s will in average travel $c\tau p/m \approx 1.2\ \text{cm}$ before decaying.

The relevant branching ratios for the ϕ and η_c are given in Table 5.2, where the contributions from both resonant and non-resonant decays are taken into account. Note that the $\eta_c \to K^+K^+K^-K^-$ decay mode is now one order of magnitude smaller compared to the other $\eta_c \to h^+h^+h^-h^-$ modes. For historical reasons we will still consider this decay mode, though its contribution to the total event yield should be marginal given the current branching ratio average.

Table 5.2: Measured and computed branching ratios for ϕ and η_c [27], including resonant decays.

Decay	BR (10^{-2})
$\phi \to K^+K^-$	49.2 ± 0.6
$\eta_c \to \pi^+\pi^-\pi^+\pi^-$	1.2 ± 0.3
$\eta_c \to K^+K^-\pi^+\pi^-$	1.5 ± 0.6
$\eta_c \to K^+K^+K^-K^-$	0.15 ± 0.07
$\eta_c \to h^+h^+h^-h^-$	2.85 ± 0.67

The visible branching ratio BR_{vis} for $B_s \to \eta_c\phi$ using the decay modes presented above are given in Table 5.3, where BR_{vis} corresponds to the total branching ratio:

$$\text{BR}_{\text{vis}} \equiv \text{BR}\left(B_s \to \eta_c\phi\right) \times \text{BR}\left(\eta_c \to h^+h^+h^-h^-\right) \times \text{BR}\left(\phi \to K^+K^-\right) .$$

Assuming a $b\bar{b}$ production cross section of $\sigma_{b\bar{b}} = 500\ \mu$b and an average nominal luminosity $\mathcal{L} = 2 \times 10^{32}\text{cm}^{-2}\text{s}^{-1}$, we get the annual number of expected $b\bar{b}$ *pairs* (in 4π):

$$N_{b\bar{b}} = \sigma_{b\bar{b}} \times \int \mathcal{L}\, dt = 1 \times 10^{12} ,$$

with an annual integrated luminosity of 2 fb^{-1} (i.e. for 10^7 s). Consequently the number of expected B_s mesons (N_{B_s}) per year and of $B_s \to \eta_c\phi$ signal events (N_{sig}) are given by:

$$\begin{aligned} N_{B_s} &= N_{b\bar{b}} \times 2 \times \text{BR}\left(\bar{b} \to B_s\right) , \\ N_{\text{sig}} &= N_{B_s} \times \text{BR}_{\text{vis}} , \end{aligned} \tag{5.3}$$

where the factor 2 takes into account the production of both b and $\bar{b}$ hadrons, with a production fraction of $\text{BR}\left(\bar{b} \to B_s\right) = (10.3 \pm 1.4)\%$ [27]. The resulting $B_s \to \eta_c\phi$ signal annual production yields for each of the considered decay modes are listed in Table 5.3.

Table 5.3: Assumed visible branching fractions for $B_s \to \eta_c\phi$ (BR_{vis}) and number of annual signal decays (N_{sig}), in 4π.

Decay	BR_{vis} (10^{-6})	N_{sig} (10^6)
$B_s \to \eta_c(\pi^+\pi^-\pi^+\pi^-)\phi$	7.3 ± 3.6	1.5 ± 0.8
$B_s \to \eta_c(K^+K^-\pi^+\pi^-)\phi$	9.1 ± 5.3	1.9 ± 1.1
$B_s \to \eta_c(K^+K^+K^-K^-)\phi$	0.9 ± 0.6	0.2 ± 0.1
$B_s \to \eta_c(h^+h^+h^-h^-)\phi$	17.3 ± 8.4	3.6 ± 1.8

5.1.1 Signal Sample Generation

The signal sample used to determine the performance of the $B_s \to \eta_c\phi$ selection corresponds to the event type 13336000 [91]. In order to take into account the effect of the possible intermediate resonances in the η_c decays, the dedicated EVTGEN decay file includes these contributions. Table 5.4 gives the relative fractions of the η_c decays used in the forced B_s decay, based on the different measured $\eta_c \to h^+h^+h^-h^-$ branching ratios, and assuming isospin conservation. The intermediate resonances are assumed to decay in their main hadronic modes which final products yield charged pions and/or kaons. All the branching ratios used are from Reference [27].

Table 5.4: Relative fractions of the η_c decay modes used in the $B_s \to \eta_c\phi$ signal sample.

Channel	Fraction [%]
$\eta_c \to \rho^0\rho^0$	30.4
$\eta_c \to \pi^+\pi^-\pi^+\pi^-$	11.7
$\eta_c \to K^{*0}\overline{K}^{*0}$	6.6
$\eta_c \to K^{*0}K^-\pi^+, \overline{K}^{*0}K^+\pi^-$	40.2
$\eta_c \to K^+K^-\pi^+\pi^-$	5.8
$\eta_c \to \phi\phi$	2.2
$\eta_c \to \phi K^+K^-$	2.8
$\eta_c \to K^+K^+K^-K^-$	0.3

We summarize hereafter a few properties of the signal particles used in the generation[1]. Note that the masses of the ϕ and η_c resonances were generated with Breit–Wigner distributions. We shall refer to these values as the true generated Monte Carlo values, e.g. when quoting mass windows or resolutions.

- B_s : mass m_{MC} = 5369.6 MeV/c^2 ;
- ϕ : central mass m_{MC} = 1019.456 MeV/c^2, width $\Gamma_{MC} \approx$ 4.260 MeV/c^2 ;

[1]The particle properties used by the `ParticlePropertySvc` and EVTGEN can be found in the relevant version of `$PARAMFILESROOT/data/ParticleTable.txt`.

- η_c : central mass $m_{\rm MC}$ = 2979.7 MeV/c^2, width $\Gamma_{\rm MC} \approx$ 16.0 MeV/c^2.

In the MC, the lifetimes are used instead of the widths. Note that electromagnetic radiative corrections are not simulated.

5.2 Particle Identification

Particles are provided by the reconstruction as protoparticles, that is tracks with no definite particle identification (PID). In the background rejection, PID is a powerful handle to reject fake combinations. Moreover, it enables to tighten the mass cuts around some composite particle, assuming a correct identification of the particles used to form it.

The PID relies on the information from various subdetectors. The RICH system identifies hadrons (i.e. $\pi^\pm$, $\mathrm{K}^\pm$ and p), also offering some discrimination for leptons. The identification of photons and π^0's is exclusively made by the ECAL system which also provides electron identification. Finally, the Muon System gives the best separation between the muon hypothesis and the rest.

The PID information for each charged particle is expressed as a likelihood ratio between the given PID hypothesis and the pion hypothesis:

$$\Delta \ln \mathcal{L}_{a\pi} = \ln \mathcal{L}(a) - \ln \mathcal{L}(\pi) = \ln \left[\frac{\mathcal{L}(a)}{\mathcal{L}(\pi)} \right] , \tag{5.4}$$

where $\mathcal{L}(a)$ is the combined likelihood assuming the particle is of type a. This is obtained as the product of likelihoods for PID estimators from the different subdetectors. For instance for a hadron h:

$$\mathcal{L}(\mathrm{h}) \;\; = \;\; \mathcal{L}^{\rm RICH}(\mathrm{h}) \times \mathcal{L}^{\rm ECAL}(\text{not e}) \times \mathcal{L}^{\rm MUON}(\text{not } \mu) \, .$$

As all the likelihood ratios are given with respect to the pion hypothesis, we can obtain any $\Delta \ln \mathcal{L}$ between two particles a and b using:

$$\Delta \ln \mathcal{L}_{ab} = \ln \mathcal{L}_{a\pi} - \ln \mathcal{L}_{b\pi} \, .$$

A particle with true type a will then tend to have a positive $\Delta \ln \mathcal{L}_{ab}$.

The choice of PID selection criteria strongly depends on the properties of the decay under study, such as topology, the particle types and the background levels. The abundance of the different particle species in long tracks in $\mathrm{b\overline{b}}$ events gives an idea of the combinatorics in the absence of PID. These fractions are approximatively of 73% for pions, 16% for kaons, 6% for electrons, 4% for protons and less than 1% for muons. Obviously, requiring a di-muon in an event already represents a powerful cut, whereas a large multiplicity decay involving kaons and pions such as $\mathrm{B_s} \to \eta_c \phi$ will be flooded with combinatorics.

For the $\mathrm{B_s} \to \eta_c \phi$ selection only the RICH information is used in the $\Delta \ln \mathcal{L}$. The selection of pions and kaons is non-exclusive, i.e. tracks can be selected as more than one particle type. As a result, the large number of final states in $\mathrm{B_s} \to \eta_c \phi$ renders the PID discrimination globally less efficient than applying geometrical criteria. Thus only mild PID cuts will be applied in the final selection.

5.3 Algorithmic Sequence

The selection of a given decay channel is generally a sequence of the type B → C S, where the particles C and S can either be final states or composite particles[2] to which a series of selection criteria is applied. The core parts of most selections can thus be reused if embedded in some generic algorithms and tools. This is the approach used in the HLT where the cuts and what decays to reconstruct are defined and configured in options using generic code, see [69].

The above approach is well suited for the selection of $B_s \to \eta_c\phi$ and has therefore replaced the previous selection algorithms for this channel [16]. The algorithmic sequence for the $B_s \to \eta_c\phi$ selection and analysis is shown in Figure 5.2. The left side of this figure contains the algorithms and the right side shows the objects in the Transient Event Store (TES). These transient objects are the inputs/outputs of the different algorithms.

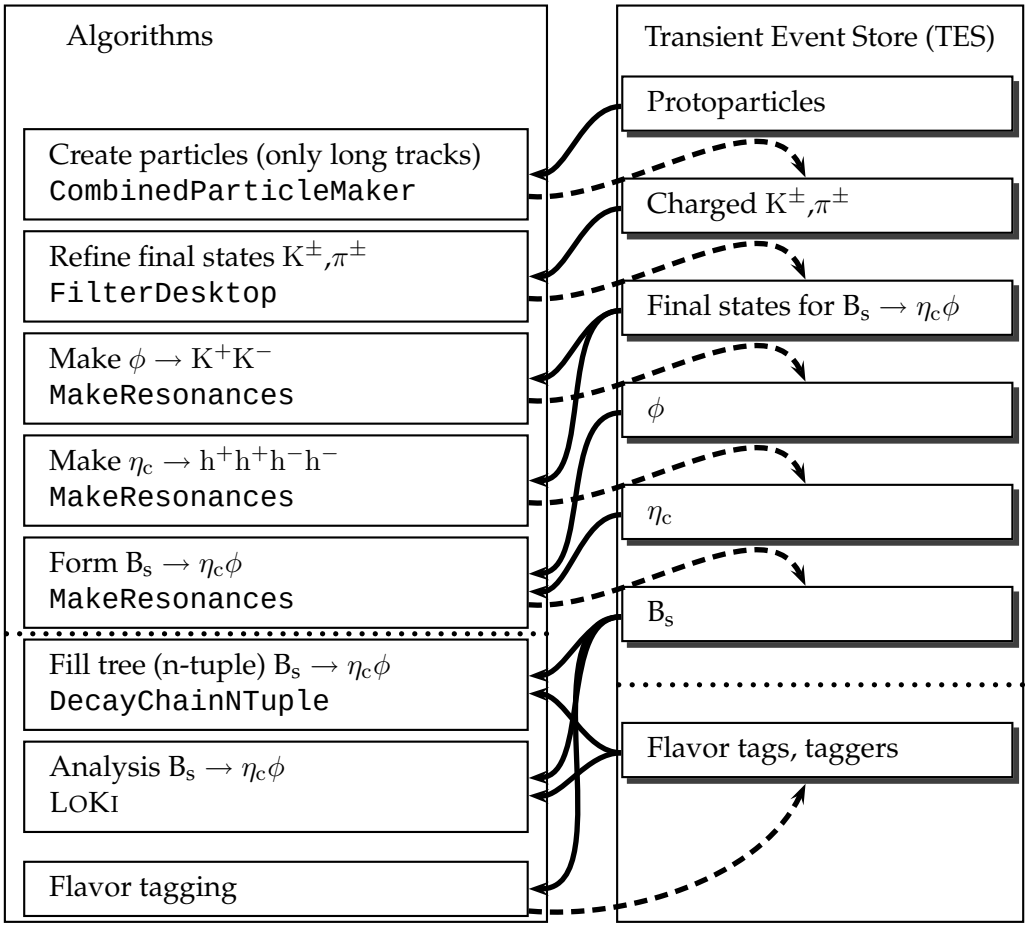


Figure 5.2: Design of the algorithmic sequence for the $B_s \to \eta_c\phi$ selection and analysis.

The stage prior to the event selection is the reconstruction of tracks and the primary vertex search. At this point the end-products of the reconstruction, namely protoparticles, can be fed to the first algorithm executed in the selection sequence which is the creation of

[2]We call final states the particles that originate from protoparticles and that are thus not composite particles. The latter can be formed with either final states or other composites, or both.

the final state particles. Based on some PID requirements and requiring only long tracks, all the charged $\pi^{\pm}$ and $K^{\pm}$ that will be used for the $B_s \to \eta_c\phi$ reconstruction are created and saved to the TES. These $\pi^{\pm}$ and $K^{\pm}$ are then refined by applying common cuts and they are saved to a pool of particles which will be considered as the final states candidates for $B_s \to \eta_c\phi$. These particles are then distributed to the ϕ and η_c selection algorithms, where they will be combined to form the different resonances by applying selection cuts. Finally, all the ϕ and η_c are combined to make the B_s candidates and further refining the previous selection criteria or applying new requirements. This ends the selection sequence, where each step is *not* executed unless the previous algorithms saved a non-empty result to the TES.

After the selection sequence several other algorithms are executed in parallel for the analysis and monitoring purposes. For instance:

- Creation of a $B_s \to \eta_c\phi$ tree (n-tuple) for the selection optimization;
- Analysis algorithm (within LOKI), e.g. for lifetime studies;
- Flavor tagging algorithms.

Note that all these examples of analysis or monitoring algorithms retrieve the selected B_s candidates from the TES.

5.4 Selection Variables

The reconstruction of $B_s \to \eta_c\phi$ candidates involves large combinatorics. The selection variables should enable the separation of wrong combinations, referred to as background, from the true signal. To this end a series of filter criteria is applied to every particle involved in the decay chain.

The choice of these variables is based on the cuts used in the HLT, see Section 4.1.3. Moreover, using the same set of cuts ensures a maximal correlation with the HLT and should improve the trigger efficiency for $B_s \to \eta_c\phi$ signal events. As we will see in Section 5.6.2, the exclusive HLT selection for $B_s \to \eta_c\phi$ will just be a mimic of the off-line selection with looser requirements, and without RICH particle identification. We apply in the off-line selection the exact same cuts for each of the η_c decay modes, except for the PID requirements. This limits biases induced by too specific selections depending on the PID as in the end we would like to use all selected events in the CP violation measurement. As we shall see, this approach nevertheless represents an additional source of selection inefficiency as we will have to cut harder whenever pions are involved, as a result of the dominant π abundance.

The selection of $B_s \to \eta_c\phi$ is based on the presence of a detached secondary vertex with respect to the primary vertices, given the B_s long lifetime and the large boost in the beam direction. The B_s daughters should therefore have a significant impact parameter with respect to all the primary vertices. The large mass difference between the B_s and its daughters gives rise to particles with larger transverse momentum on average than tracks coming from a primary vertex (PV).

We give hereafter the track types used, we summarize and motivate the selection variables used in the off-line $B_s \to \eta_c\phi$ selection. We have the following requirements:

Track types

- Only particles made from tracks reconstructed as long are used, i.e. reconstructed in the VELO and extrapolated to the T stations.

Kinematic cuts

- A cut on the momentum p is applied to all final states to reduce combinatorics at the particle creation level.
- All the B_s daughters and final states are required to have some transverse momentum p_T.
- The invariant mass of the daughter particles should be compatible with the mass of the mother. The cuts on the reconstructed masses are symmetric around the true nominal MC value (i.e. used in the generation) and denoted by $\pm\delta m$, therefore corresponding to full mass windows of $2\delta m$.

Geometrical criteria

- Unsigned impact parameter significances IPS $\equiv$ IP$/\sigma_{IP}$, where σ_{IP} denotes the uncertainty on IP, are computed with respect to each primary vertex. For the B_s daughters and granddaughters a cut on the smallest IPS is applied.
- The selected particles should come from a common point in space, which quality is characterized by the χ^2 of a simple vertex fit.
- The flight distance FD (or flight distance significance FS $\equiv$ FD$/\sigma_{FD}$) is defined as the distance between the B_s decay vertex and its associated primary vertex.
- A cut is applied on the angle $\theta_{p,F}$ of the reconstructed B_s momentum $\vec{p}$ and the direction of flight $\vec{F}$.

PID requirements

- The $\Delta \ln \mathcal{L}$ variables are used to limit the combinatorics. The assigned PID, based upon the RICH information, should correspond to the required particle type. For kaons $\Delta \ln \mathcal{L}_{K\pi}$ and $\Delta \ln \mathcal{L}_{Kp}$ requirements are applied, whereas for pions only a $\Delta \ln \mathcal{L}_{\pi K}$ separation is required.

Topological cut

- There is yet another topological requirement applied after all the final selection criteria, which is a vertex isolation cut. This cut was found to be extremely useful and efficient against events where a primary vertex was missed by the reconstruction and where we actually reconstruct this vertex as the B_s one. We require that only a limited number of particles from the pool of selected pions and kaons point to the B_s vertex within some maximum IPS cut.

5.4.1 Plotting Selection Variable Distributions

In the study and optimization of the selection criteria, see Section 5.5, different kind of plots are produced, depending on the data under study, on the kind of cuts applied and on the environment (on- or off-line). The level of cuts are listed hereafter:

No cuts When running on signal we ask for the reconstructed B_s candidate to be fully associated to the MC truth without applying any cuts, except for the preselection $\Delta \ln \mathcal{L}$ cuts in the off-line case. This enables to visualize the possible (pre)selection values and estimate the loss caused by each individual cut, with the effect of the other variables factorized out. Anticipating the next section, we show in the Figures 5.3 – 5.5 all the selection variables on signal data, when applying no cuts and for associated candidates.

Preselection The selection criteria will be optimized considering inclusive $b\overline{b}$ events as the dominant source of background. In order to ease the optimization and reduce the $b\overline{b}$ combinatorics, a set of preselection cuts was determined and applied to both the signal and $b\overline{b}$ samples.

Selection The selection performance (efficiencies, resolutions, ...) will be quoted after applying all the final tight values of the selection criteria.

When plotting signal candidates we will always require the entire B_s decay chain of the selected B_s candidate to be fully associated to the true generated MC decay, in order not to bias the distributions with combinatorial background from the signal sample. When plotting background, all the candidates found by the selection algorithms are considered such that histograms entries represent candidates and not events. This is also the case for the associated candidates in the signal sample as the association is in a few percent of the cases not unique.

The number of entries in the B_s distributions corresponds to the number of selected B_s candidates. As the same particle could enter different B_s combinations for the same event, we only plot the particle once since it could obviously distort the distributions. The number of entries in the plots for the B_s *daughters* therefore represents the number of *different* particles used in the selection of all the B_s candidates.

When comparing signal data with $b\overline{b}$ background we arbitrarily normalize both the signal and the $b\overline{b}$ histograms to unity, considering all the histograms entries. All $b\overline{b}$ distributions are obtained from the DC04-v2 $b\overline{b}$ sample.

5.5 Event Selection Tunes

Many random combinations of particles can fake our signal if no selection criteria are applied. Moreover, these fake combinations will distort the time-dependent measurements as they do not exhibit the same characteristics as the signal. We therefore need to apply requirements on our combinations to select as many correct associations as possible while controlling the level of wrong combinations.

The different sources of background are not necessarily obvious. The distinctive features of b decays allow to discard events with only light flavor quark states since for instance they will not give sufficiently displaced secondary vertices with large invariant mass.

Inclusive $b\overline{b}$ events share the same properties as our signal and thus represent a potential background. Moreover, since the production fraction of $b\overline{b}$ is huge compared to that of the signal, this can be considered as the dominant background. It was decided to tune all the selection cuts considering $b\overline{b}$ events as the only source of background, irrespective of any other type of background. Another potential source of background are charm decays, though there is a large mass gap between c and b hadrons.

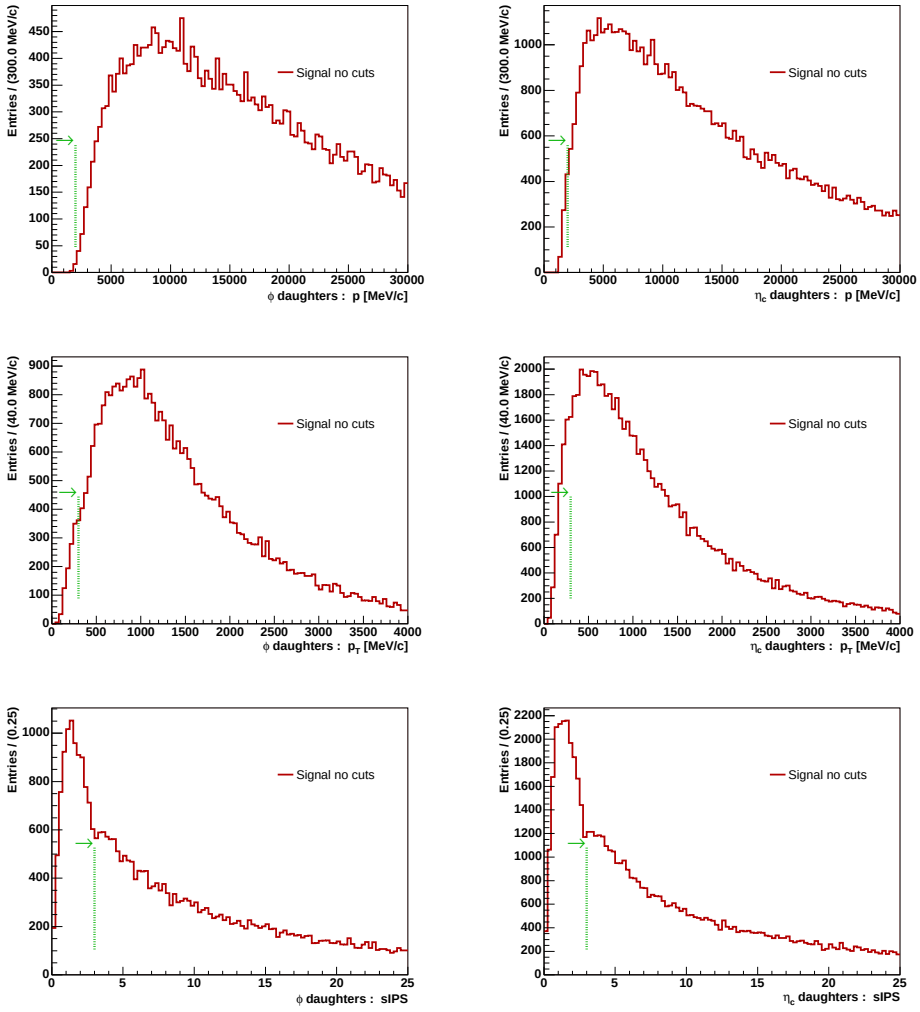


Figure 5.3: Distributions of p [MeV/c] (top), p_{T} [MeV/c] (middle), and sIPS (bottom) for the off-line associated daughters of the ϕ (left) and of the η_{c} (right) from $\mathrm{B_s} \rightarrow \eta_{\mathrm{c}}\phi$ signal decays. *No cuts* are applied. The vertical lines and arrows indicate the *final selection cuts*.

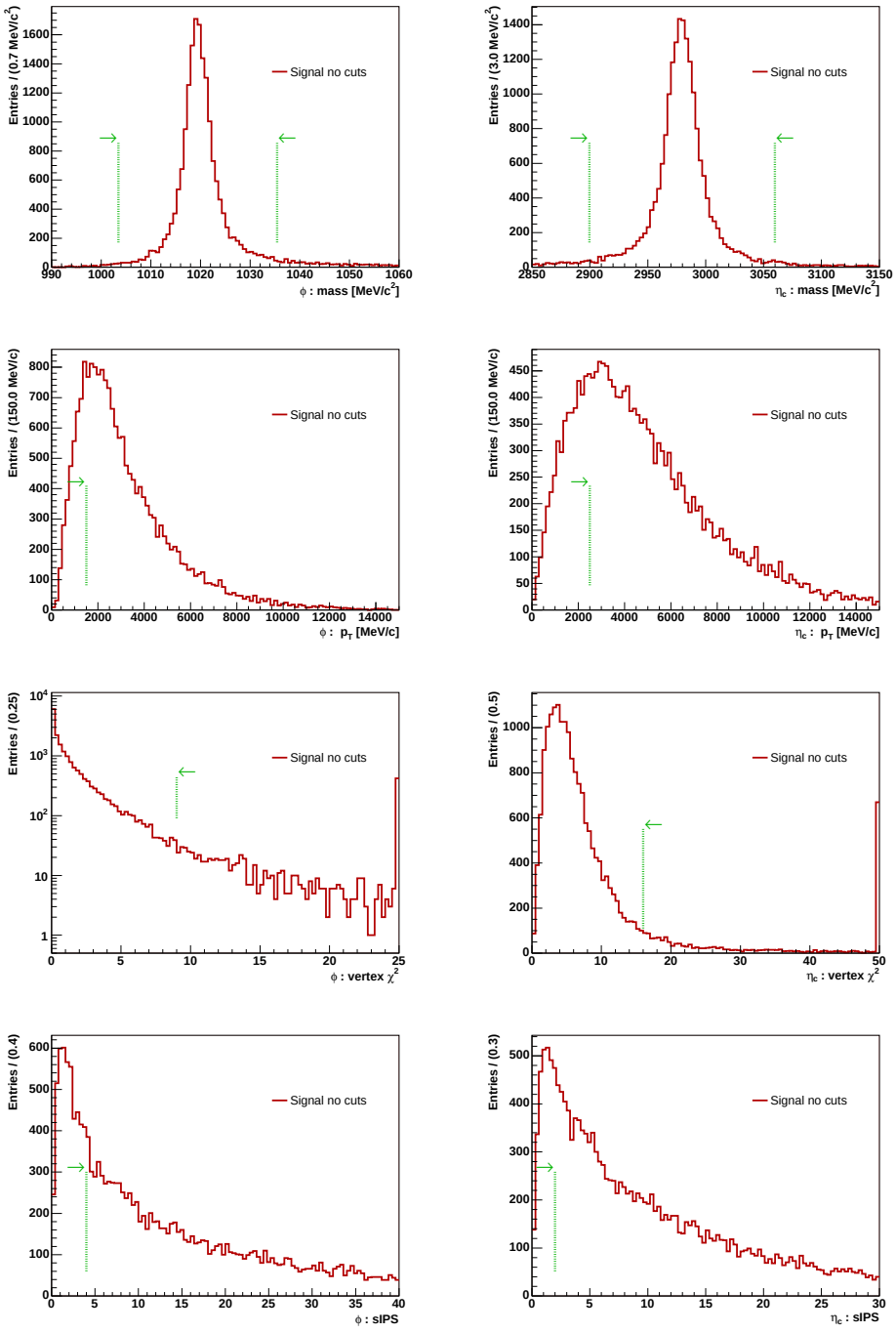


Figure 5.4: Distributions of m [MeV/c^2] (top), $p_{\rm T}$ [MeV/c] (middle top), vertex χ^2 (middle bottom), and sIPS (bottom) for the off-line associated ϕ (left) and of the $\eta_{\rm c}$ (right) from $\rm B_s \to \eta_c \phi$ signal decays. *No cuts* are applied. The vertical lines and arrows indicate the *final selection cuts*.

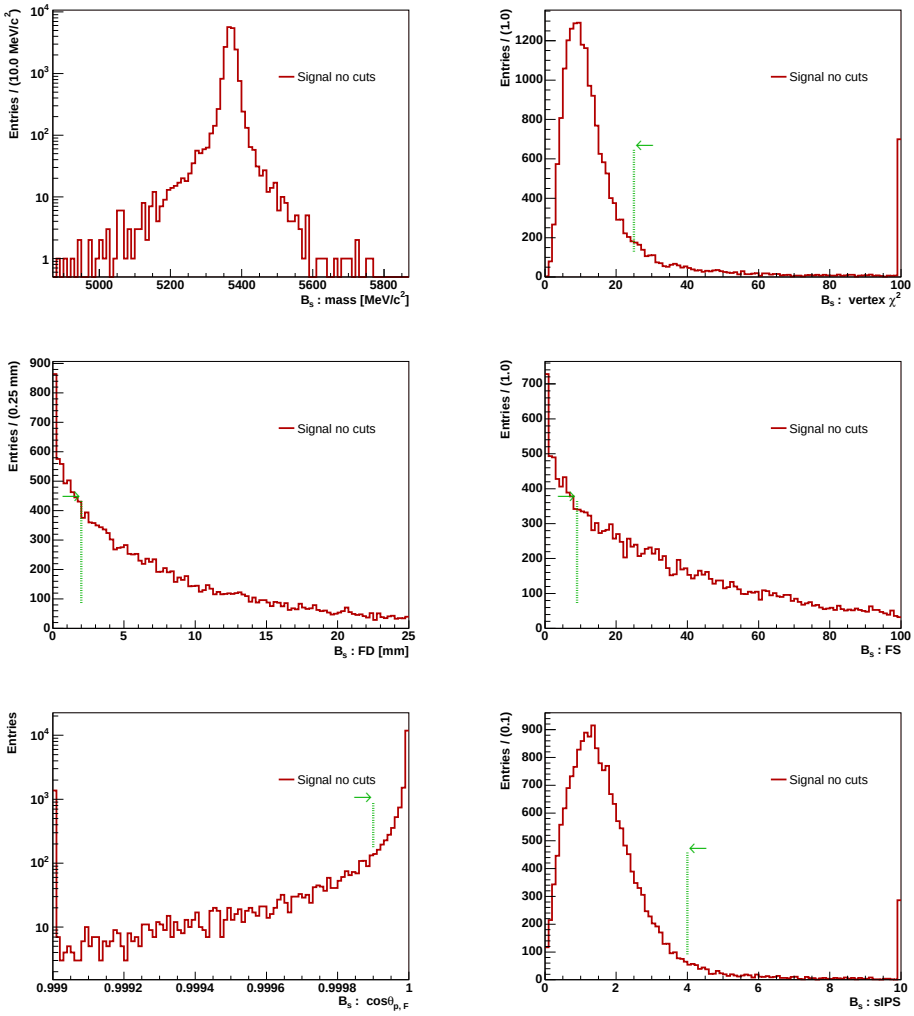


Figure 5.5: Distributions of m [MeV/c^2] (top left), vertex χ^2 (top right), FD [mm] (middle left), FS (middle right), $\cos\theta_{p,\mathrm{F}}$ (bottom, left), sIPS (bottom right), for the off-line associated B_s from $B_s \rightarrow \eta_c \phi$ signal decays. *No cuts* are applied. The vertical lines and arrows indicate the *final selection cuts*.

The available statistics of inclusive $b\bar{b}$ data are limited as a few millions of events only correspond to a few minutes of LHCb running. In order to cope with this limitation a trick is used to artificially increase the statistics. We consider a very loose mass window around the B_s mass ($\pm 500\,\mathrm{MeV/c^2}$) and assume that the background level in this enlarged mass window depends linearly on the reconstructed mass. Then an estimate of the background in the tight final mass window ($\pm 50\,\mathrm{MeV/c^2}$) is obtained by dividing the number of candidates in the loose mass window by the ratio of the two windows widths (10 in our case). This mass window trick is equivalent to simulating ten times more background, provided we have a linear distribution for the background mass. The limitations of this assumption and how we deal with them will be discussed in the Section 5.7.2.

A preselection tuned on the DC04-v1 $b\bar{b}$ data was designed to achieve a rejection factor of at least 1000. The preselection cuts are listed in Table 5.5. The signal and background distributions after preselection obtained with DC04-v2 $b\bar{b}$ data are given in Figures 5.6 – 5.9. This preselection is then applied to the signal as well and used as a starting point for the determination of the final selection. The number of preselected events on DC04-v2 $b\bar{b}$ is 15'568 out of 27'291'931 events, yielding a rejection factor of $f = (1753 \pm 14)$.

The procedure adopted to tune the off-line selection cuts is the following:

- Impose a HLT-like selection for the final states.
- Require the masses to be at most $4-5\sigma$ from the nominal central values, taking into account the intrinsic widths of the resonances.
- Choose the final cuts on some variables (e.g. χ^2 of vertices, FD, IPS of the B_s, ...) based on the signal and $b\bar{b}$ distributions after preselection (with DC04-v1 data).
- Vary the remaining cuts with large steps in order to kill *all* $b\bar{b}$ events using the full statistics of the DC04-v1 sample, while keeping the highest efficiency on signal.
- Once all the background events are killed, we further tighten some of the cuts without loosing additional signal, profiting from the cuts correlations to have an extra safety factor.

As we will see in Section 5.7, an important source of background corresponds to events where a true primary vertex failed to be reconstructed. These events are potentially dangerous since they offer an important collection of tracks of high momentum with presumably large impact parameter with respect to all other reconstructed primary vertices. Due to the large combinatorics involved, these missed primary vertices may be identified as the B_s decay vertex. Moreover, the standard filter criteria are in this case not discriminant enough.

A vertex isolation cut is thus applied requiring that only a limited number of tracks be compatible with the B_s vertex. For each $\pi^{\pm}$, $K^{\pm}$ passing the $B_s \to \eta_c\phi$ selection cuts for the final states, we count the number of non-signal tracks pointing to the reconstructed B_s vertex with IPS < 4. If this number is larger than 12, the B_s candidate is discarded. This cut was also studied with the DC04-v1 $b\bar{b}$ sample by relaxing a few cuts in order to study the nature of the $b\bar{b}$ background. This last cut is set such that no signal event passing the other selection criteria is lost. Figure 5.10 shows the number of B_s candidates as a function of the number of particles from the pool of selected kaons and pions compatible with the B_s decay vertex, i.e. with IPS < 4. The distributions are after all selections cuts, with a loose mass window for the DC04-v2 $b\bar{b}$ selected candidates and with the tight mass

Table 5.5: $B_s \to \eta_c\phi$ off-line selection criteria. The rightmost values in parentheses indicate the preselection cuts.

Final states $\pi^\pm$ and $K^\pm$			
Track type		long	
$\pi : \Delta \ln \mathcal{L}_{\pi K}$	$>$	-2.0	(-5.0)
$K : \Delta \ln \mathcal{L}_{K\pi}$	$>$	-2.0	(-5.0)
$K : \Delta \ln \mathcal{L}_{Kp}$	$>$	-5.0	(-5.0)
p [MeV/c]	$>$	2000.0	(2000.0)
p_T [MeV/c]	$>$	300.0	(300.0)
sIPS	$>$	3.0	(1.0)
		ϕ	
δm [MeV/c^2]	$\pm$	16.0	(20.0)
p_T [MeV/c]	$>$	1500.0	(1000.0)
χ^2 vertex	$<$	9.0	(16.0)
sIPS	$>$	4.0	(2.0)
		η_c	
δm [MeV/c^2]	$\pm$	80.0	(100.0)
p_T [MeV/c]	$>$	2500.0	(1000.0)
χ^2 vertex	$<$	16.0	(25.0)
sIPS	$>$	2.0	(2.0)
		B_s	
δm [MeV/c^2]	$\pm$	50.0	(500.0)
χ^2 vertex	$<$	25.0	(60.0)
FS	$>$	9.0	(1.0)
FD [mm]	$>$	2.0	(1.5)
sIPS	$<$	4.0	(6.0)
$\cos\theta_{p,F}$	$>$	0.9999	(0.0)
B_s vertex isolation cut (final selection)			
Pool: $\pi^\pm$, $K^\pm$ final states cuts			
Accept if $n \leq 12$ particles with IPS < 4 w.r.t. B_s vertex			

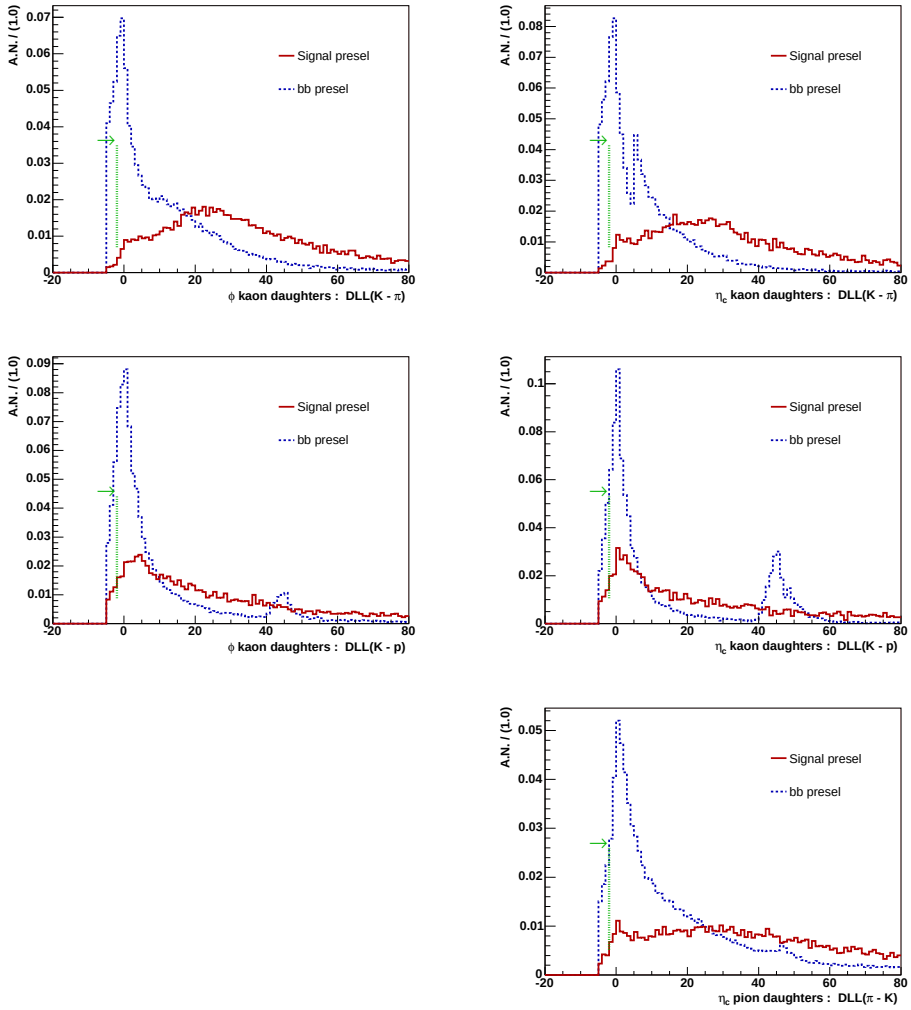


Figure 5.6: Distributions of the kaons $\Delta \ln \mathcal{L}_{\mathrm{K}\pi}$ (top) and $\Delta \ln \mathcal{L}_{\mathrm{Kp}}$ (middle), and of the pions $\Delta \ln \mathcal{L}_{\pi\mathrm{K}}$ (bottom), for the off-line associated $\mathrm{B_s} \to \eta_\mathrm{c}\phi$ signal final states (red solid lines) and $\mathrm{b\overline{b}}$ (blue dotted lines), after preselection. Left plots: ϕ daughters. Right plots: η_c daughters. The vertical lines and arrows indicate the *final selection cuts*.

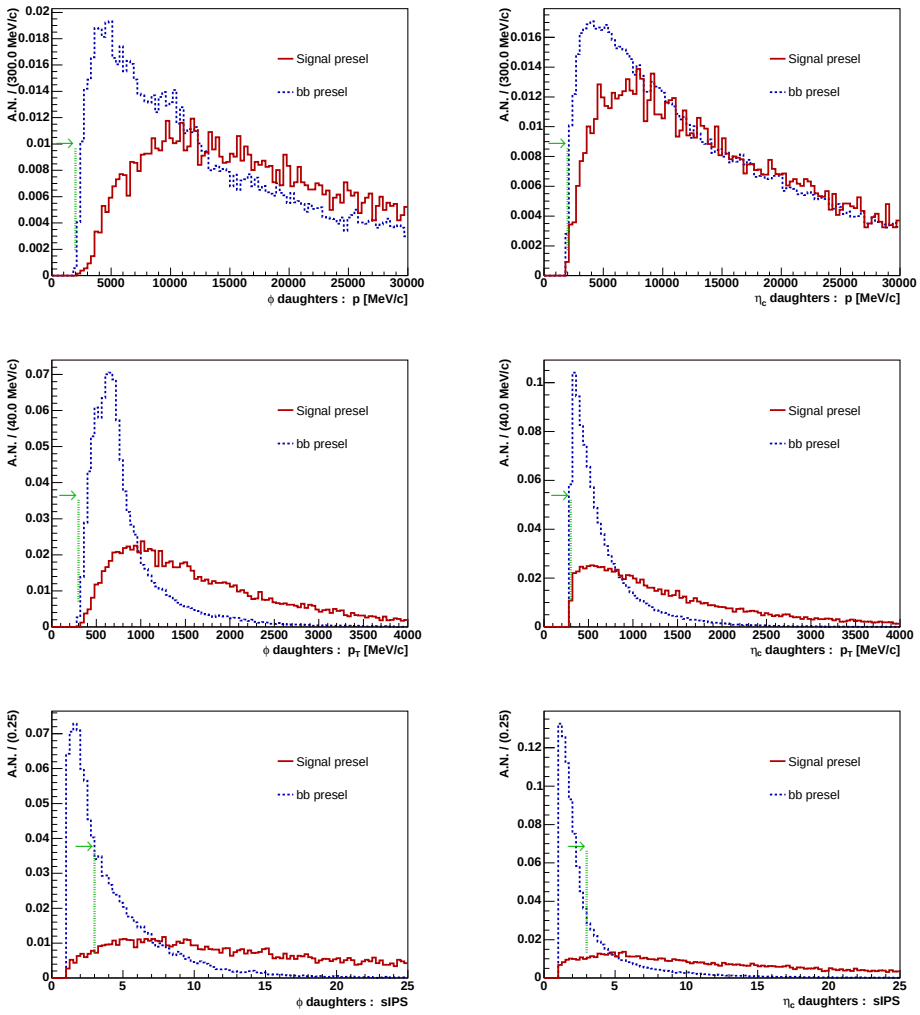


Figure 5.7: Distributions of p [MeV/c] (top), p_{T} [MeV/c] (middle), and sIPS (bottom) for the off-line associated $\mathrm{B_s} \rightarrow \eta_c \phi$ signal final states (red solid lines) and $\mathrm{b\overline{b}}$ (blue dotted lines), after preselection. Left plots: ϕ daughters. Right plots: η_c daughters. The vertical lines and arrows indicate the *final selection cuts*.

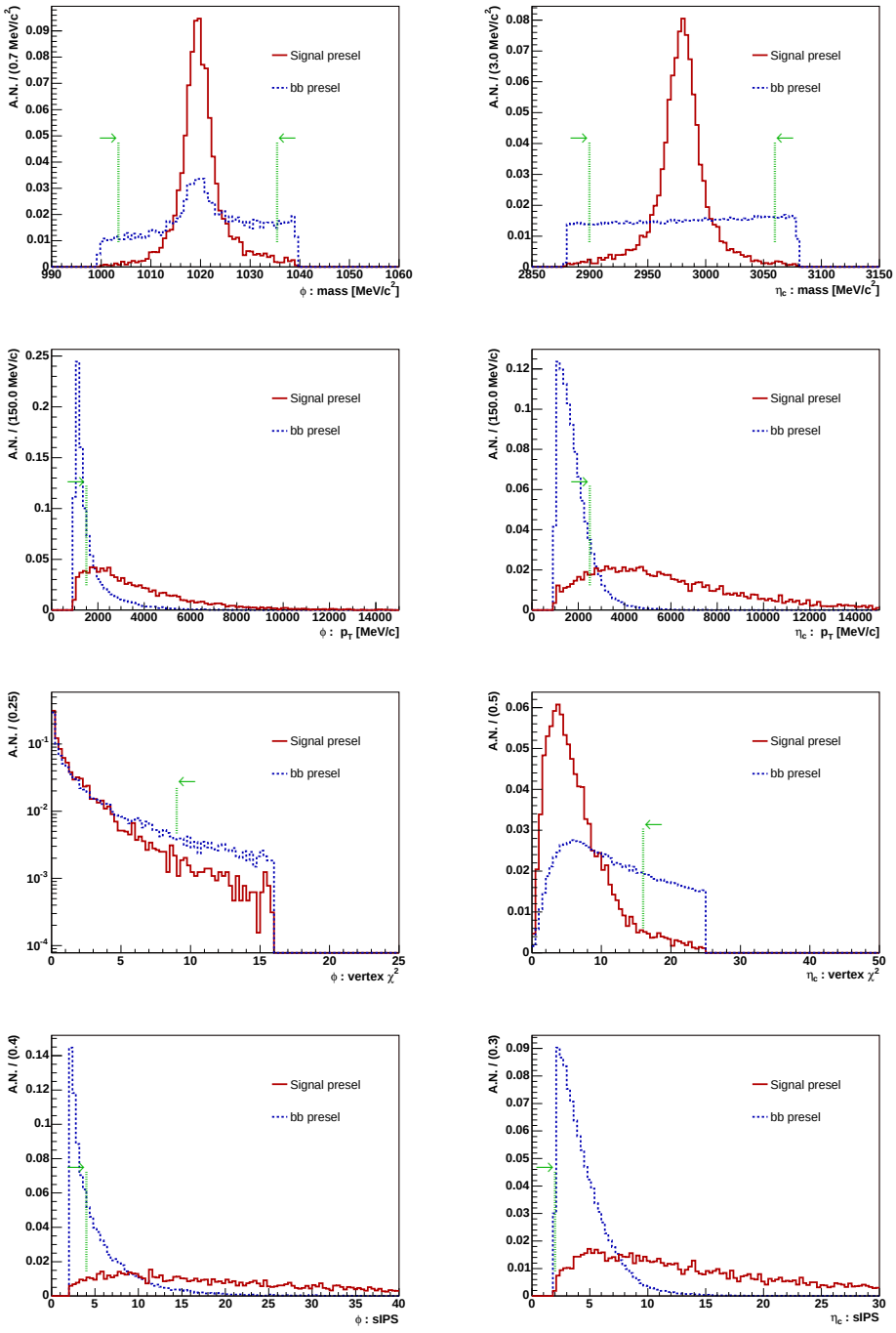


Figure 5.8: Distributions of m [MeV/c^2] (top), $p_{\rm T}$ [MeV/c] (middle top), vertex χ^2 (middle bottom), and sIPS (bottom) for the off-line associated $B_s \to \eta_c \phi$ signal composite daughters (red solid lines) and $b\overline{b}$ (blue dotted lines), after preselection. Left plots: ϕ. Right plots: η_c. The vertical lines and arrows indicate the *final selection cuts*.

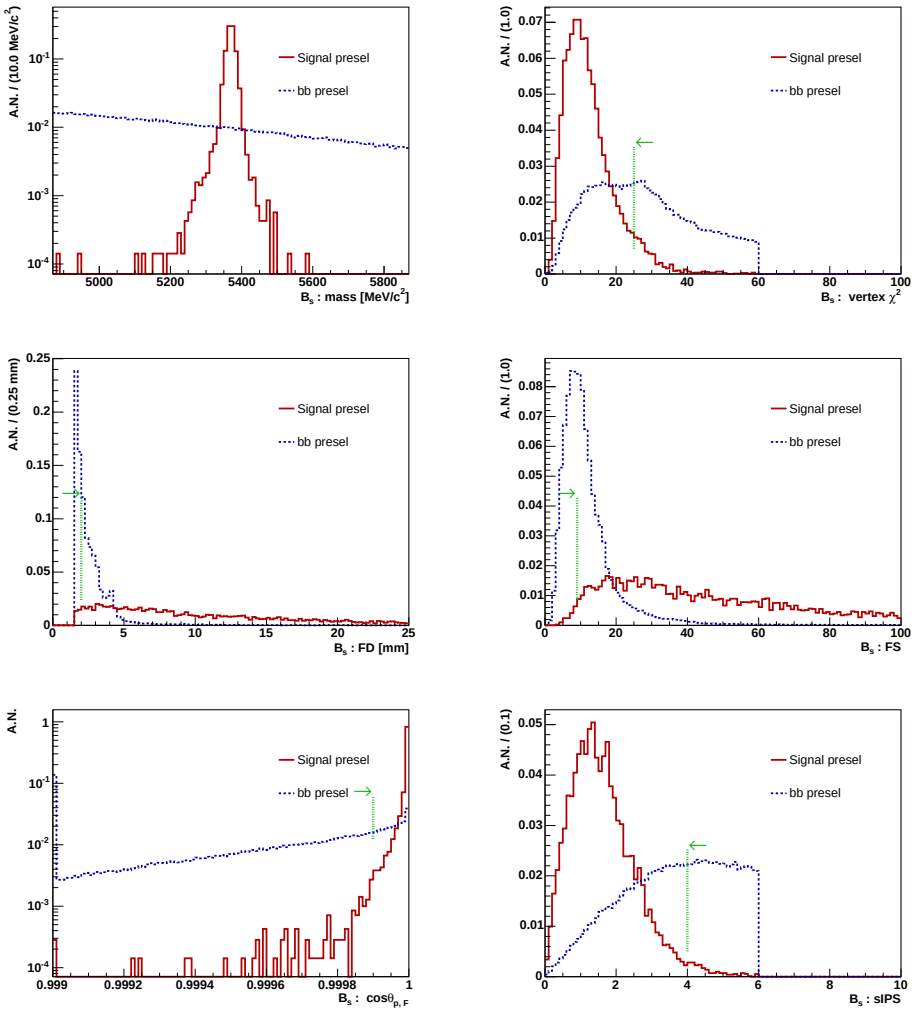


Figure 5.9: Distributions of m [MeV/c^2] (top left), vertex χ^2 (top right), FD [mm] (middle left), FS (middle right), $\cos\theta_{p,\mathrm{F}}$ (bottom left), sIPS (bottom right), for the off-line associated $\mathrm{B_s} \rightarrow \eta_c \phi$ signal $\mathrm{B_s}$ (red solid lines) and $\mathrm{b\overline{b}}$ (blue dotted lines), after preselection. The vertical lines and arrows indicate the *final selection cuts*.

window for the selected signal. We see that this cut could be further refined without any loss in signal efficiency. The nature of the $b\overline{b}$ candidates and the final results of the selection will be discussed in the next sections.

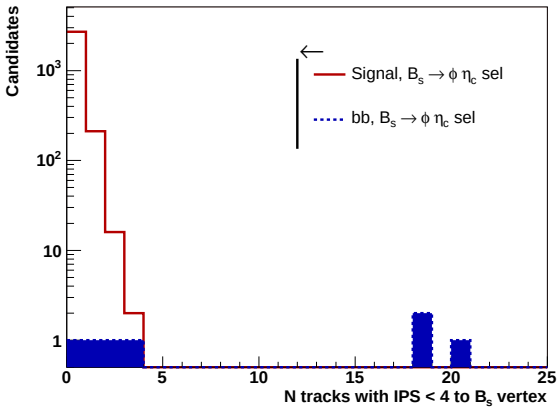


Figure 5.10: Effect of the isolation vertex filter after all other selection cuts, on signal (red solid line) and $b\overline{b}$ (blue dashed line).

The final selection criteria are summarized in Table 5.5. We will use these cuts and the inclusive DC04-v2 $b\overline{b}$ sample to test the selection and get an unbiased background-to-signal ratio (B/S). In particular, we will completely ignore the DC04-v1 $b\overline{b}$ sample for the B/S determination.

5.6 Signal Event Yield

In this section we give the performance of the final $B_s \to \eta_c \phi$ selection on signal. As mentioned in Section 3.5.2, the generation of the signal sample is affected by the generation random seed bug. We assume that the number of independent events is the same in each subset of the initial sample. The quoted efficiencies are therefore unaffected while we scale their binomial errors with the fraction of independent obtained from Equation (3.5), and the error on the number of independent events is neglected.

5.6.1 Reconstruction and Selection Efficiencies

In order to disentangle the performance of the event selection from the track finding efficiencies we introduce the following definitions for the reconstruction on all events:

- N_{gen}: total number of generated Monte Carlo signal events considered in the analysis, i.e. the number of events on tape;
- N_{ble}: number of reconstructible events, considering only long tracks;
- $N_{\text{'ted}}$: number of reconstructed events, with all final state particles reconstructed as long tracks;
- $N_{\text{rec'\&}}$: number of reconstructible and reconstructed events;

- N_{sel}: number of off-line selected events (and not candidates), after all final cuts.

All the above numbers are listed in Table 5.6.

Table 5.6: Reconstruction and off-line selection numbers for $B_s \to \eta_c\phi$.

Channel	N_{gen}	$N_{\text{'ble}}$	$N_{\text{'ted}}$	$N_{\text{rec'\&}}$	N_{sel}
$B_s \to \eta_c(\pi^+\pi^-\pi^+\pi^-)\phi$	190139	14140	11197	8811	1344
$B_s \to \eta_c(K^+K^-\pi^+\pi^-)\phi$	236955	16900	13920	11086	1439
$B_s \to \eta_c(K^+K^+K^-K^-)\phi$	23407	1645	1378	1100	126
$B_s \to \eta_c(h^+h^+h^-h^-)\phi$	450501	32685	26495	20997	2909

We can now define the total selection efficiency before trigger:

$$\varepsilon_{\text{tot}} \quad = \quad \varepsilon^{\theta}_{\text{sig}\,B_s} \times \frac{N_{\text{sel}}}{N_{\text{gen}}} = \varepsilon_{\text{det}} \times \varepsilon_{\text{rec/det}} \times \varepsilon_{\text{sel/rec}}\,, \tag{5.5}$$

where $\varepsilon^{\theta}_{\text{sig}\,B_s} = (34.7 \pm 0.3)\%$ is the acceptance of the 400 mrad generator-level cut on the signal B_s, see Section 3.2.3. The second equality is introduced to separate the selection, the reconstruction and the detection efficiencies using the following definitions:

- $\varepsilon_{\text{det}} = \varepsilon^{\theta}_{\text{sig}\,B_s} \times \frac{N_{\text{sel}}}{N_{\text{gen}}} \times \frac{1}{\varepsilon_{\text{rec/det}}} \times \frac{1}{\varepsilon_{\text{sel/rec}}}$ is the detection efficiency;
- $\varepsilon_{\text{rec/det}} = \frac{N_{\text{rec'\&}}}{N_{\text{'ble}}}$ is the reconstruction efficiency on reconstructible (i.e. detected) events;
- $\varepsilon_{\text{sel/rec}} = \frac{N_{\text{sel}}}{N_{\text{'ted}}}$ is the off-line selection efficiency on the reconstructed events.

All the above efficiencies are given in Table 5.7, where the uncertainties are statistical and take into account the fraction of independent events.[3] As we can see from this table, the efficiency ε_{det} for having all the decay products reconstructible is small due to the high multiplicity of the $B_s \to \eta_c\phi$ decay. The reconstruction efficiency for having all six final states reconstructed as long tracks is low again because of the multiplicity of the decay as in a first approximation $\varepsilon_{\text{rec/det}} \sim (\varepsilon_{\text{tr}})^n$ where ε_{tr} is the per-track reconstruction efficiency and $n = 6$ in our case. The reconstruction efficiency tends to increase with the number of kaons in the final states, as it can be seen from the $\sim 6\sigma$ deviation in $\varepsilon_{\text{rec/det}}$ between the $B_s \to \eta_c(\pi^+\pi^-\pi^+\pi^-)\phi$ and $B_s \to \eta_c(K^+K^-\pi^+\pi^-)\phi$ decay modes. This is due to the higher momentum of the $\eta_c \to \pi^+\pi^-\pi^+\pi^-$ decay products. Indeed, the reconstruction efficiency of long tracks reaches a plateau for tracks with $p > 20\,\text{GeV/c}$, and then slightly decreases with increasing momentum, see Figure 3.10.

The selection efficiency is significantly enhanced whenever more pions are involved. This is due to our cuts on p_{T} and IPS since for pions these two variables are in average larger, as a result of the mass difference between η_c and its daughters. In the selection procedure we purposely ignored this fact to have a better alignment with the HLT, as in

[3] The error σ_ε on an efficiency $\varepsilon = n/N$ with a fraction of independent events f is $\sigma_\varepsilon = \sqrt{(\varepsilon(1-\varepsilon))/fN}$, when ignoring the error on f. A quadratic propagation of the errors has been used for Table 5.7 and the statistical error on ε_{tot} is computed using the first equality in (5.5), as the factorization in the second term leads to some numerical imprecision.

the exclusive HLT selections the particle identification is ignored. Finally, the choice of final cuts is dominated by the $B_s \to \eta_c(\pi^+\pi^-\pi^+\pi^-)\phi$ mode because of the abundance of pions in comparison to kaons and therefore the larger combinatorics induced. For instance, the selection cuts for $B_s \to \eta_c(K^+K^+K^-K^-)\phi$ could be relaxed as requiring six kaons forming a common vertex is a strong constraint.

Table 5.7: Reconstruction and off-line selection efficiencies for $B_s \to \eta_c\phi$, before the trigger. The uncertainties are statistical.

Channel	ε_{det} [%]	$\varepsilon_{\text{rec/det}}$ [%]	$\varepsilon_{\text{sel/rec}}$ [%]	ε_{tot} [%]
$B_s \to \eta_c(\pi^+\pi^-\pi^+\pi^-)\phi$	3.3 ± 0.1	62.3 ± 0.4	12.0 ± 0.3	0.246 ± 0.007
$B_s \to \eta_c(K^+K^-\pi^+\pi^-)\phi$	3.1 ± 0.1	65.6 ± 0.4	10.3 ± 0.3	0.211 ± 0.006
$B_s \to \eta_c(K^+K^+K^-K^-)\phi$	3.1 ± 0.4	66.9 ± 1.2	9.1 ± 0.8	0.187 ± 0.017
$B_s \to \eta_c(h^+h^+h^-h^-)\phi$	3.2 ± 0.1	64.2 ± 0.3	11.0 ± 0.2	0.225 ± 0.005

5.6.2 Trigger Performance on Signal

In this section we detail the performance of each of the trigger levels on off-line selected signal events, with a particular attention to the HLT. Indeed, $B_s \to \eta_c\phi$ cannot profit from the generous bandwidths granted to the muons streams, and we therefore need a dedicated exclusive HLT selection. It should be noted that many changes are expected in the future for the trigger performance on this channel as a result of the 1 MHz readout scheme.

L0 and L1 efficiencies

We give in Table 5.8 the L0 efficiencies on off-line selected signal events. The different lines are explained in Section 2.2.8. As expected most of the bandwidth is taken by the hadronic line: $\sim 82\%$ of the L0 accepted are triggered by hadrons.

Table 5.9 summarizes the L1 efficiencies on off-line selected and L0-accepted signal events. The different lines are explained in Section 2.2.8. Here again the dominant contribution is from the L1-generic line: $\sim 99\%$ of the L1-accepted events are triggered by this line. This trigger level could have a better efficiency if it were not for the recent change in bandwidth division due to the adjustments made for muons. Indeed, the generic part only takes $\sim 53\%$ of the output rate on mimimum-bias events, when running on the sample used in Chapter 4. Since there no longer will be any boundary between the L1 and the HLT with the 1 MHz strategy, a better treatment of hadronic modes can be expected.

Exclusive HLT selection

The HLT version used here is identical to that of Chapter 4. We will therefore only briefly describe the exclusive HLT selection for $B_s \to \eta_c\phi$ and give the main results.

The processing of the HLT starts with the on-line reconstruction and the execution of the generic HLT. For $B_s \to \eta_c\phi$ the performance of the generic HLT is:

Table 5.8: Number of L0-triggered events N_{L0} and corresponding efficiency $\varepsilon_{\mathrm{L0/sel}}$ on $N_{\mathrm{sel}} = 2909$ off-line selected $\mathrm{B_s} \to \eta_c \phi$ events, with the contribution from each L0 (possibly overlapping) line. The uncertainties are statistical.

L0 line	N_{L0}	$\varepsilon_{\mathrm{L0/sel}}$ [%]
h	1149	39.5 ± 0.9
e	172	5.9 ± 0.4
γ	120	4.1 ± 0.4
π^0	337	11.6 ± 0.6
μ	204	7.0 ± 0.5
$\mu\mu$	280	9.6 ± 0.5
Total	1400	48.1 ± 0.9

Table 5.9: Number of L1-triggered events N_{L1} and corresponding efficiency $\varepsilon_{\mathrm{L1/L0}}$ on $N_{\mathrm{L0}} = 1400$ off-line selected $\mathrm{B_s} \to \eta_c \phi$ and L0 accepted events, with the contribution from each L1 (possibly overlapping) line. The uncertainties are statistical.

L1 line	N_{L1}	$\varepsilon_{\mathrm{L1/L0}}$ [%]
$\sum \ln p_{\mathrm{T}}$	1159	82.8 ± 1.0
e	87	6.2 ± 0.7
γ	61	4.4 ± 0.6
μ	63	4.5 ± 0.6
$\mu\mu$	19	1.4 ± 0.3
J/ψ	23	1.6 ± 0.3
Total	1175	83.9 ± 1.0

- $N_{\rm HLTGen} = 1126$ events, $\varepsilon_{\rm HLTGen/L1} = (95.8 \pm 0.6)\%$.

The explanation for this loss is twofold:

1. The better momentum determination in the HLT corrects for L1 $p_{\rm T}$ mistakes, e.g. when the decision tracks have too many wrong TT clusters assigned. Moreover, the hadronic L1-confirmation part of the generic HLT has a stronger requirement on $\sum \ln p_{\rm T}$.

2. The ECAL lines are not yet reproduced in the generic HLT. This effect is however marginal for $B_s \to \eta_c\phi$, given that the generic L1 line triggers basically all the events.

After a positive generic HLT decision, the full on-line tracking is performed, and the HLT selections are run. In order to exclusively reconstruct $B_s \to \eta_c\phi$, we first need to find all the on-line signal tracks, see Chapter 4. The corresponding efficiencies for $B_s \to \eta_c\phi$, given in Table 5.10, are obtained after the generic HLT and off-line selections, on events where all the signal MC particles are associated to protoparticles.

Table 5.10: On-line tracking efficiencies for $B_s \to \eta_c\phi$ signal determined on $N^{\rm evt}_{\rm off} = 1079$ events after generic HLT and off-line selection. The uncertainties are statistical.

$B_s \to \eta_c\phi$	Per track		Per event	
Track type	$N^{\rm tr}_{\rm rec}$	$\varepsilon^{\rm tr}_{\rm rec}$ [%]	$N^{\rm evt}_{\rm rec}$	$\varepsilon^{\rm evt}_{\rm rec}$ [%]
VELO-RZ	6454	99.7 ± 0.1	1061	98.3 ± 0.4
VELO-space	6373	98.4 ± 0.2	992	91.9 ± 0.8
VELO-TT	6270	96.8 ± 0.2	902	83.6 ± 1.1
Long	6135	94.8 ± 0.3	803	74.4 ± 1.3
Long (no TT)	6199	95.8 ± 0.3	857	79.4 ± 1.2

The HLT performance for $B_s \to \eta_c\phi$ will enormously depend on the on-line tracking inefficiencies because of the large multiplicity of the decay, as it can be seen from Table 5.10. The on-line tracking efficiencies were already discussed in Chapter 4.

The design of an exclusive HLT selection for $B_s \to \eta_c\phi$ is trivial as the off-line selection uses the generic algorithms developed for the HLT. Thus only a simple reshuffling of the options is needed, in particular we feed the selection sequence with pions and kaons made from on-line tracks, without any PID. The on-line selection variables are the same as the off-line but with looser cuts and their determination follows the procedure explained in Section 4.3.4. A few points specific to the HLT exclusive selection are listed hereafter:

- The final states filtering on $p_{\rm T}$ and IPS is imposed.

- The mass windows are enlarged by roughly one off-line standard deviation (combined resolution and intrinsic width), to account for the slightly worse on-line resolution, and to accommodate sidebands. The $\pm 500\,{\rm MeV/c^2}$ loose B_s mass window is used.

- The selection criteria *not* depending on the tracking errors are set to their off-line value.

- The vertex χ^2 cuts are doubled compared to the off-line, to allow for some additional indetermination due to the on-line environment.
- The significance cuts are relaxed by a few σ units, to take into account the overestimation of the on-line errors with respect to the off-line.
- These cuts were then blindly applied to the ~ 3.28 s of minimum-bias events at 40 kHz.

The list of cuts applied to the exclusive HLT $B_s \to \eta_c \phi$ selection is given in Table 5.11, with the corresponding off-line values for comparison.

Table 5.11: Exclusive HLT $B_s \to \eta_c \phi$ selection criteria. The rightmost values in parentheses indicate the off-line selection cuts.

Final states $\pi^\pm$ and $K^\pm$			
Track type			long
p [MeV/c]	$>$	2000.0	(2000.0)
p_T [MeV/c]	$>$	300.0	(300.0)
sIPS	$>$	2.0	(3.0)
		ϕ	
δm [MeV/c^2]	$\pm$	20.0	(16.0)
p_T [MeV/c]	$>$	1500.0	(1500.0)
χ^2 vertex	$<$	18.0	(9.0)
sIPS	$>$	3.0	(4.0)
		η_c	
δm [MeV/c^2]	$\pm$	100.0	(80.0)
p_T [MeV/c]	$>$	2500.0	(2500.0)
χ^2 vertex	$<$	32.0	(16.0)
sIPS	$>$	2.0	(2.0)
		B_s	
δm [MeV/c^2]	$\pm$	500.0	(50.0)
χ^2 vertex	$<$	50.0	(25.0)
FS	$>$	6.0	(9.0)
FD [mm]	$>$	2.0	(2.0)
sIPS	$<$	6.0	(4.0)
$\cos\theta_{p,F}$	$>$	0.9999	(0.9999)

When running with a cheated exclusive HLT selection, that is only considering events where the signal final states are on-line reconstructed and applying no cuts, the corresponding efficiency is $\sim 72.3\%$ on off-line selected events passing the generic HLT. Note that this efficiency is a bit lower compared to the tracking efficiency of Table 5.10 as this time we do not require the association of the off-line events (mainly due to ghosts and association inefficiency).

The number of events selected by the dedicated HLT selection and the corresponding efficiency on the N_{HLTGen} signal events after the generic HLT are given below, together with the detailed number of events of all the HLT streams. The results are:

Channel	N_{HLTGen}	$N_{\mathrm{HLT}}^{\mathrm{specific}}$	$\varepsilon_{\mathrm{HLT}}^{\mathrm{specific}}$ [%]	$N_{\mathrm{HLT}}^{\mathrm{exb}}$	$N_{\mathrm{HLT}}^{\mathrm{b}\to\mu}$	$N_{\mathrm{HLT}}^{\mu\mu}$	$N_{\mathrm{HLT}}^{\mathrm{D}^*}$	$N_{\mathrm{HLT}}^{\mathrm{tot}}$
$\mathrm{B_s} \to \eta_c\phi$	1126	767	68.1 ± 1.4	770	41	4	111	819

The corresponding full breakdown of the HLT efficiencies for the different streams is listed hereafter, including possible overlaps between streams:

Channel	$\varepsilon_{\mathrm{HLT}}^{\mathrm{exb}}$ [%]	$\varepsilon_{\mathrm{HLT}}^{\mathrm{b}\to\mu}$ [%]	$\varepsilon_{\mathrm{HLT}}^{\mu\mu}$ [%]	$\varepsilon_{\mathrm{HLT}}^{\mathrm{D}^*}$ [%]	$\varepsilon_{\mathrm{HLT}}^{\mathrm{tot}}$ [%]
$\mathrm{B_s} \to \eta_c\phi$	68.4 ± 1.4	3.6 ± 0.6	0.4 ± 0.2	9.9 ± 0.9	72.7 ± 1.3

We notice that the other exclusive selections do not contribute significantly to the overall $\varepsilon_{\mathrm{HLT}}^{\mathrm{exb}}$ efficiency. This is not surprising as $\mathrm{B_s} \to \eta_c\phi$ is the only 6-prong channel in the HLT. The other streams help in selecting this channel at the level of $\sim 5\%$. The quoted efficiency of the dedicated specific selection, $\varepsilon_{\mathrm{HLT}}^{\mathrm{specific}}$, includes the track finding efficiencies, which clearly limit the performance. Factorizing out the tracking component, the exclusive efficiency amounts to $\sim 94\%$.

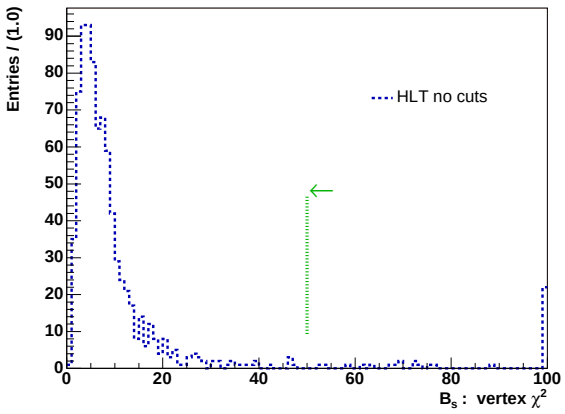


Figure 5.11: Vertex χ^2 of the $\mathrm{B_s}$ after on-line cheated selection without any cuts, on off-line selected $\mathrm{B_s} \to \eta_c\phi$ events passing the generic HLT. The vertical line and the arrow indicate the *exclusive HLT selection cut*.

The drop in efficiency due to the selection cuts is caused by the difference of tracking errors between the on-line and off-line. Actually the loss due to most of the HLT selection criteria is marginal, except for the χ^2 cuts. This can be seen on Figure 5.11, for the on-line vertex of the $\mathrm{B_s}$ after cheated selection. The histogram entries correspond to candidates, nevertheless as we mostly have only one true associated candidate per event, the fraction of events lost can be directly read off the plot where the overflow bin is explicitly drawn. Essentially all the events where the six final states are present are lost because of the χ^2 cut, even though its value was enlarged by a factor two compared to the off-line selection. As mentioned in Chapter 4, the vertex fitter gives abnormal χ^2 results for very long-lived $\mathrm{B_s}$ candidates.

Finally, the output rate of the $B_s \to \eta_c\phi$ exclusive HLT selection is well under control, as only two events pass when running on ~ 3.28 s of minimum-bias events at 40 kHz, one of them being also triggered by the inclusive D^* stream. This corresponds to an output rate of only (0.6 ± 0.4) Hz.

Trigger efficiency breakdown

We summarize in Table 5.12 the trigger efficiencies as presented previously, for each of the $B_s \to \eta_c\phi$ modes.

Table 5.12: Trigger efficiency breakdown, on $B_s \to \eta_c\phi$ off-line selected events.

Modes	$\varepsilon_{\text{L0/sel}}$ [%]	$\varepsilon_{\text{L1/L0}}$ [%]	$\varepsilon_{\text{HLTGen/L1}}$ [%]	$\varepsilon_{\text{HLT/HLTGen}}$ [%]	$\varepsilon_{\text{HLT/sel}}$ [%]
$\eta_c(4\pi)\phi$	48.9 ± 1.4	82.2 ± 1.5	95.4 ± 0.9	72.4 ± 2.0	27.8 ± 1.2
$\eta_c(2\pi 2\text{K})\phi$	48.0 ± 1.3	85.7 ± 1.3	96.5 ± 0.8	73.6 ± 1.9	29.2 ± 1.1
$\eta_c(4\text{K})\phi$	41.3 ± 4.4	82.7 ± 5.3	93.0 ± 3.9	65.0 ± 7.5	20.6 ± 3.6
$\eta_c(4\text{h})\phi$	48.1 ± 0.9	83.9 ± 1.0	95.8 ± 0.6	72.7 ± 1.3	28.2 ± 0.8

As argued above, the trigger performance for $B_s \to \eta_c\phi$ is expected to be considerably improved with the next iteration of the software. In particular, a better treatment of hadronic lines is foreseen with the removal of the L1 trigger, presumably recovering the loss in the generic HLT. Also as the on-line and off-line tracking should be identical, a large improvement in the tracking efficiencies can be assumed. Furthermore, an improvement in the forward tracking is to be expected, without the use of TT as an intermediate step. Finally, the possibility of having some inclusive ϕ stream will also profit to $B_s \to \eta_c\phi$.

5.6.3 Annual Yield

The annual yield corresponds to the number of signal events we expect to reconstruct at LHCb in one year of data taking (10^7 s), taking into account the signal production rate, the geometrical acceptance, the reconstruction and selection efficiencies as well as the trigger performance. The annual yield thus englobes the following factors:

- Annual integrated luminosity: 2 fb^{-1} and the corresponding number of produced B_s mesons (N_{B_s}), see (5.3).
- The number of expected $B_s \to \eta_c\phi$ events in 4π (N_{sig}), see Table 5.3.
- The total selection efficiency ε_{tot} before trigger, including the detector's acceptance, given in Table 5.7.
- The global trigger efficiency:

$$\varepsilon_{\text{HLT/sel}} = \varepsilon_{\text{L0/sel}} \times \varepsilon_{\text{L1/L0}} \times \varepsilon_{\text{HLTGen/L1}} \times \varepsilon_{\text{HLT/HLTGen}} ,$$

  with the breakdown of Table 5.12.

The number of $B_s \to \eta_c\phi$ events available for physics studies is then given by:

$$S = N_{B_s} \times BR_{vis} \times \varepsilon_{tot} \times \varepsilon_{HLT/sel} = N_{sig} \times \varepsilon^{\theta}_{sig\,B_s} \times \frac{N_{sel\&trg}}{N_{gen}} = (2.26 \pm 0.08) \times 10^3\,, \quad (5.6)$$

where $N_{sel\&trg}$ is the number of off-line selected events and after the trigger selection. The annual event yield after all triggers is $\mathcal{O}(2-3)$ k events and strongly depends on the HLT strategy and the on-line performance.

Signal purity

The selected events in the $B_s \to \eta_c\phi$ signal sample are not necessarily true decays, as we can also select combinatorial candidates. For instance a candidate could be reconstructed with five out of six tracks from the signal, and an extra track taken from somewhere else. The results are:

- 2'928 reconstructed candidates in 2'909 signal events:
  - $\sim 96\%$ of the candidates are associated to signal.
  - $\sim 4\%$ are ghosts. In general we have five tracks correctly associated to the signal while the last one has no MC particle associated. This is almost always due to an inefficiency of the associators, as the reconstructed particle and the non-matched true particle have very similar four-vectors, within resolution effects.
  - $< 1\%$ are background-like candidates. These are:
    - **signal reflection**: 6 candidates with a double π–K misidentification, leading to a B_s mass slightly under/overestimated, since we have more than two final states.
    - **signal partial reconstruction**: 1 candidate where the muon from the decay in flight of the signal pion is used, $\pi \to \mu\nu_\mu$. The resulting B_s mass is practically unaffected.
    - **from the same primary vertex**: 2 candidates where five tracks come from the signal, and the missing track is a primary track from the same primary as the signal. This is due to wrongly reconstructed impact parameters.

5.6.4 Flavor Tagging

In a CP-asymmetry measurement with neutral B mesons, the flavor tagging, i.e. the identification of the flavor at production, plays a crucial role. Indeed, as seen in Section 3.4, the wrong-tag fraction will attenuate the B_s–$\overline{B}_s$ oscillations resulting in a dilution of the observed CP asymmetry. The flavor tagging has a direct impact on the signal statistics, as it *cannot* be fully efficient unless we reconstruct exclusively all the possible b decays in the event! Moreover, flavor tagging represents an important source of systematic error.

The flavor tagging results for $B_s \to \eta_c\phi$ after off-line selection and each trigger level are quoted in Table 5.13. The tagging techniques are described in Section 3.4.1. The detail of each tagging category is given in Table 5.14, after the full HLT. Note that the annual yields of Section 5.6.3 were before tagging, as even untagged events carry some information on the B_s–$\overline{B}_s$ mixing phase.

Table 5.13: Tagging efficiency ε_{tag}, wrong-tag fraction ω_{tag} and effective tagging efficiency ε_{eff} after $B_s \to \eta_c\phi$ selection before trigger, and after each trigger level on off-line selected signal events. The uncertainties are statistical.

$B_s \to \eta_c\phi$	Selection [%]	L0 [%]	L1 [%]	HLT [%]
ε_{tag}	59.7 ± 0.9	63.2 ± 1.3	64.9 ± 1.4	65.7 ± 1.7
ω_{tag}	31.0 ± 1.1	29.6 ± 1.6	29.2 ± 1.7	30.9 ± 2.0
ε_{eff}	8.6 ± 1.4	10.6 ± 2.2	11.3 ± 2.5	9.6 ± 2.9

Table 5.14: Performance of the different tags after $B_s \to \eta_c\phi$ selection and HLT. The combined performance is obtained from a neural network, which also takes into account that an event may have several tags. The uncertainties are statistical.

$B_s \to \eta_c\phi$	ε_{tag} [%]	ω_{tag} [%]	ε_{eff} [%]
Opposite-side μ	11.9 ± 1.1	29.6 ± 4.6	2.0 ± 0.9
Opposite-side K	30.9 ± 1.6	36.6 ± 3.0	2.2 ± 1.0
Opposite-side vertex charge	25.3 ± 1.5	37.5 ± 3.4	1.6 ± 0.9
Same-side K	34.1 ± 1.7	33.1 ± 2.8	3.9 ± 1.3
Total	65.7 ± 1.7	30.9 ± 2.0	9.6 ± 2.9

The effective tagging efficiency is as for most of the B_s decays a few percents larger compared to the B_d mesons. This is essentially due to the same-side kaon tag. Besides ε_{tag} which will reduce the available statistics for the determination of the $\sin\phi_s$ term of the decay rates, the ω_{tag} from the control channel used to simultaneously fit the signal and extract ΔM_s should be as close as possible to that of the signal channel. Any difference will have to be corrected for by weighting the control channel's wrong-tag fraction. This also represents a cost in statistics and is a source of systematic uncertainty.

As we will see in Chapter 6, the control channel used for the signal B_s is $B_s \to D_s\pi$, for which $\omega_{tag} \sim 31\%$ and $\varepsilon_{tag} \sim 63\%$, after the HLT. The tagging performances for $B_s \to \eta_c\phi$ and $B_s \to D_s\pi$ are the same within statistical errors, and a possible difference is expected to be very small since both are hadronic decays. Note that another decay channel that could serve as control channel for $B_s \to \eta_c\phi$ with possibly better correlation is $B_s \to D_s\pi\pi\pi$, since this channel has the same multiplicity as the signal.

5.7 Background Studies

We give in this section the results of the selection on different backgrounds. As explained in Section 5.5, the assumed dominant source of background is $b\bar{b}$ events. We will determine the background level using the DC04-v2 inclusive $b\bar{b}$ sample, and using the mass window trick. Then other potential sources of background will also be investigated, in order to have a broader overview of the different possible contributions.

5.7.1 Preselection on $b\overline{b}$

In order to understand the sources of background in the $b\overline{b}$ sample, it is instructive to examine the preselected candidates. To this end we introduce the following categorization[4], observed in the $B_s \to \eta_c\phi$ preselection:

b-hadron background: all the selected final states come from the same true particle.

- Signal (code 0).
- Reflection (code 30): at least one track misidentified.
- Partially reconstructed decay (code 40): at least one missing charged particle when compared to the MC truth.
- Low-mass background (code 50): same as above, with possibly missing neutrals.

Background due to reconstruction problems: tracks from a primary vertex, badly reconstructed tracks, non-reconstructed primary vertex.

- Ghost (code 60): at least one selected track used has no MC particle associated to it.
- From a PV (code 70): at least one track originating from a primary vertex.
- From same PV (code 80): all the tracks used come from the same primary vertex.
- From different PV (code 100): at least two tracks come from different primary vertices.

Combinatorial background: pure combinatorics.

- combinatorial $b\overline{b}$ (code 110): all tracks come from different b hadrons.

The distribution of these categories is displayed in Figure 5.12.

The first observation when looking at Figure 5.12 is that we are flooded with candidates from categories 60, 70, 80 and 100. Even though the number of preselected events is ~ 15 k, the actual number of combinations is huge. This is a consequence of the high-multiplicity decay we look for together with the large number of primary-vertex tracks.

The category 70 (60) typically happens when we add a primary track (ghost) to a partially reconstructed b decay. These combinations should be wiped out by imposing large IPS cuts. However, badly reconstructed tracks could still be considered as detached tracks. The category 100 candidates should also disappear when imposing impact parameter cuts. The most dangerous source of background is clearly category 80, since if a primary is missed by the reconstruction, then it is likely to be reconstructed as the secondary B_s vertex simply because of combinatorics. The isolation criterion, see Section 5.4, was precisely designed to reject these events since in this particular case standard criteria are not efficient enough.

The other categories contain 90 events distributed as follows:

- Category 0: 2 candidates in 2 events.
- Category 40: 4 candidates in 4 events.

[4]This classification is 2based on the `BackgroundCategory` tool.

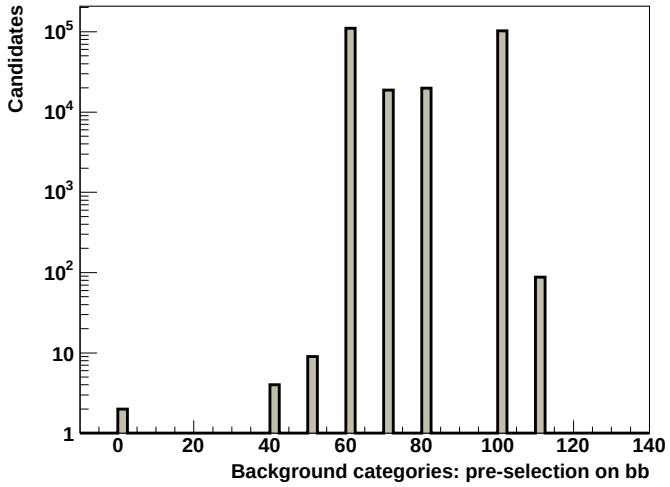


Figure 5.12: Background categories of $B_s \to \eta_c\phi$ candidates in $b\overline{b}$ events after preselection.

- Category 50: 9 candidates in 9 events.
- Category 110: 88 candidates in 76 events.

It is worth pointing out that no candidate is formed solely with tracks not coming from b-hadron decays, consistent with our assumption that non $b\overline{b}$ events are not the dominant background.

The events from category 40 are all b-hadron decays to at least six charged final states and with additional γ, π^0. The invariant mass is always below 5.2 GeV/c^2. These will always be outside the B_s signal mass region and can therefore be ignored. The only relevant features of all these specific backgrounds is that most of them are $B_s \to D_s^{\mp(*)}X$ high-multiplicity decays. We will thus consider these kinds of decays in the specific background studies of Section 5.7.3.

5.7.2 $b\overline{b}$ Background Level

In order to determine the background level from $b\overline{b}$ we make use of the mass window trick explained in Section 5.5. This is valid as long as we do not have any peaking background, such as partially reconstructed decays with all the tracks coming from the same b hadron. We will therefore ignore any peaking background falling outside the tight B_s mass window for the purpose of estimating the B/S ratio in this window.

There are 7 events out of $\sim$ 27 M DC04-v2 inclusive $b\overline{b}$ events passing the $B_s \to \eta_c\phi$ final selection with the enlarged mass window, and before applying the isolation cut.

1. Signal event:

$$B_s \quad \to \quad (\eta_c \to \pi^+\pi^-\pi^+\pi^-)(\phi \to K^+K^-)\,.$$

This event is not counted as background. Note that given the $B_s \to \eta_c\phi$ selection efficiency, and the total BR used in EvtGen for this channel (1.61×10^{-5}), we expect to select ~ 3 signal events in this inclusive $b\overline{b}$ sample.

2. Partially reconstructed b decay:

$$\mathrm{B_u^-} \quad \rightarrow \quad \pi^+(\eta \rightarrow \pi^+\pi^-\gamma)(\mathrm{D}^{*0} \rightarrow \pi^0(\mathrm{D}^0 \rightarrow \mathrm{K}^-\pi^+\pi^+\pi^-)) \,.$$

   In this event the K is misidentified as a π and the neutral γ and π^0 not reconstructed. The reconstructed invariant $\mathrm{B_s}$ mass is 5081 MeV/c^2 and this event is not counted in the B/S determination.

3. $\mathrm{b\overline{b}}$ combinatorial background event:

$$\begin{aligned} \overline{\mathrm{B}}_\mathrm{s} \quad &\rightarrow \quad (\phi \rightarrow \mathrm{K}^+\mathrm{K}^-)\pi^+\pi^-(\rho^0 \rightarrow \pi^+\pi^-) \,, \\ \pi^0 \quad &\rightarrow \quad \gamma(\gamma \rightarrow \mathrm{e}^-\mathrm{e}^+) \,. \end{aligned}$$

   The above $\overline{\mathrm{B}}_\mathrm{s}$ decay has the same signature as the signal, however only five tracks are taken from it while the last particle used in the fake combination is an e from a γ conversion, originating from a primary π^0, and the e is misidentified as a π. This event must be counted as background.

4. Ghost and missed primary events: two candidates are reconstructed each with five tracks coming from a non-reconstructed primary vertex whereas the last particle used is a ghost. These two candidates do not pass the isolation filter and are therefore ignored for the B/S calculation.

5. All tracks from the same non-reconstructed primary vertex: this event is killed by the isolation filter and is thus not counted as background.

6. All the tracks from the same non-reconstructed primary vertex: this event is not killed by the isolation filter and is counted as background.

The relevant feature in the $\mathrm{b\overline{b}}$ events surviving the $\mathrm{B_s} \rightarrow \eta_\mathrm{c}\phi$ final selection is the dominance of reconstruction issues such as primary vertex search, and badly reconstructed tracks faking high IP. Any improvement in the primary vertex search will directly benefit to the $\mathrm{B_s} \rightarrow \eta_\mathrm{c}\phi$ selection and its background level. Nevertheless, the primary vertex finding efficiency is somehow limited by the number of tracks required to form it and we should have very special care with this requirement as we do not want to identify large multiplicity b-hadron decay vertices as primary vertices.

As described above, two events are to be counted as background for the B/S estimate in the signal window from the inclusive $\mathrm{b\overline{b}}$ sample. The assumption that the background in the enlarged mass window depends linearly on the reconstructed mass allows us to set the limits on the expected background levels from $\mathrm{b\overline{b}}$ events. Furthermore, we assume that the trigger efficiency is the same for the signal and background, and therefore we will quote the background level before trigger. Computing the B/S after trigger would just result in a loss of statistics. The background-to-signal ratio is thus computed as:

$$\left(\frac{B}{S}\right)^{\mathrm{b\overline{b}}} = \left(\frac{\varepsilon^{\theta}_{\mathrm{b\overline{b}}}}{\varepsilon^{\theta}_{\mathrm{sig\,B_s}}}\right) \times \left(\frac{1}{2 \times \mathrm{BR}\left(\mathrm{\bar{b} \rightarrow B_s}\right) \times \mathrm{BR_{vis}}}\right) \times \left(\frac{N^{\mathrm{b\overline{b}}}_{\mathrm{sel}} / \left(F_m \times N^{\mathrm{b\overline{b}}}_{\mathrm{gen}}\right)}{N_{\mathrm{sel}}/N_{\mathrm{gen}}}\right) \,, \quad (5.7)$$

where:

- $\varepsilon^{\theta}_{\mathrm{b\overline{b}}} = 43.4\%$ is the acceptance of 400 mrad requirement for the $\mathrm{b\overline{b}}$ sample, see Section 3.2.3;

- $\varepsilon^{\theta}_{\text{sig}\,B_s} = 34.7\%$ is the efficiency of the 400 mrad cut for the signal B_s sample, see Section 3.2.3;

- BR $\left(\bar{b} \to B_s\right) = 10.3\%$ is the $b \to B_s$ production fraction, see Section 2.1.3;

- BR_{vis} is the $B_s \to \eta_c\phi$ visible branching ratio given in Table 5.3;

- $N^{b\bar{b}}_{sel} = 2$ is the number of events passing the final $B_s \to \eta_c\phi$ selection and counted as background;

- $N^{b\bar{b}}_{gen} = 27'291'931$ is the number of number of $b\bar{b}$ events analyzed, as given in Section 3.5.2;

- $F_m = 500/10 = 10$ is the mass window enlargement factor when going from the tight (± 50 MeV/c^2) to the loose (± 500 MeV/c^2) mass window;

- N_{sel} is the number of off-line selected signal events passing the final selection and given in Table 5.6;

- N_{gen} is the number of generated signal events, see Table 5.6.

The number of $b\bar{b}$ events passing the selection being close to zero, we use the unified 90% confidence levels (CL) $[\nu_1, \nu_2]$ for the mean of a Poisson distributed variable given $N^{b\bar{b}}_{sel}$ observed events in the absence of background. The method of Feldman–Cousins [101] for the construction of confidence intervals is used. The 90% unified CL for the B/S from $b\bar{b}$ events for $B_s \to \eta_c\phi$, together with the central value are respectively given by:

$$B/S \in [0.10, 1.17]\,, \quad B/S = 0.40 \pm 0.28\,, \tag{5.8}$$

where the error is statistical only.

5.7.3 Specific Backgrounds

We give here the results of the selection on different specific backgrounds. Whenever central values are given for the specific B/S estimates, we correct the statistical errors with the fraction of independent events if the sample used was affected by the random bug seed. All event types are described in Section 3.5.2.

$B_d \to \phi X$ background

Looking at the mass distribution of the preselected ϕ candidates in the $b\bar{b}$ sample, see Figure 5.8, we observed a significant ϕ mass peak. We therefore looked at an inclusive $B_d \to \ldots \to \phi X$ sample, with an estimated total BR of 1.63%, where the ϕ is forced to decay to two kaons. The results are:

N^{spec}_{gen} of events analyzed	N selected in loose δm	N selected in tight δm
998500	0	0

This background source is not dangerous.

Charmless B_d background

A possible background arises from large multiplicity b decays without any charm intermediate states such that all the final states originate from the same vertex. We thus tested the $B_s \to \eta_c\phi$ selection on a B_d charmless decays sample, with a total estimated BR of 0.58%. The results are:

N_{gen}^{spec} of events analyzed	N_{sel}^{spec} selected in loose δm	N_{sel}^{spec} selected in tight δm
607500	0	0

This background source seems to be negligible. Unfortunately, only four-body decays or less were generated.

$B_s \to D_sD_s$ cocktail background

Inclusive D_s decays represent a potential background as previously mentioned in Section 5.7.1. A sample of $B_s \to D_sD_s$ decays is then used to check the $B_s \to \eta_c\phi$ selection. Many decay modes are generated, involving hadrons and neutral particles. The results are:

N_{gen}^{spec} of events analyzed	N_{sel}^{spec} selected in loose δm	N_{sel}^{spec} selected in tight δm
50000	0	0

The low available statistics for this sample do not allows us to draw a definite conclusion. We thus study hereafter specific b-hadron decays involving D_s mesons.

$B_s \to D_sD_s$ background

The most dangerous source of specific background are B_s decays with the same signature and multiplicity as $B_s \to \eta_c\phi$. The $B_s \to D_sD_s$ events where $D_s \to KK\pi$ are likely to be selected, in particular for events where both D_s do not fly much. In this case the two tertiary D_s vertices could be close enough to be reconstructed as a single B_s-decay candidate vertex. The results are:

N_{gen}^{spec} of events analyzed	N_{sel}^{spec} selected in loose δm	N_{sel}^{spec} selected in tight δm
383000	12	12

The corresponding B/S contribution is calculated as:

$$\left(\frac{B}{S}\right)^{B_s\to D_sD_s} = \left(\frac{BR_{vis}^{B_s\to D_sD_s}}{BR_{vis}}\right) \times \left(\frac{N_{sel}^{spec}/N_{gen}^{spec}}{N_{sel}/N_{gen}}\right), \tag{5.9}$$

where $BR_{vis}^{B_s\to D_sD_s}$ is the estimated $B_s \to D_sD_s$ visible branching ratio calculated as:

$$\begin{aligned} BR_{vis}^{B_s\to D_sD_s} &= BR(B_d \to D^-D_s^+) \times BR(D_s \to KK\pi)^2 = (17.6 \pm 4.3) \times 10^{-6}, \\ BR(B_d \to D^-D_s^+) &= (6.5 \pm 0.21) \times 10^{-3}, \\ BR(D_s \to KK\pi) &= (5.2 \pm 0.9) \times 10^{-2}. \end{aligned}$$

The above branching ratios are taken from Reference [27]. Note that $BR_{vis}^{B_s\to D_sD_s}$ is similar to that of $B_s \to \eta_c\phi$, as it can be seen from Table 5.3. The corresponding background level

from this specific source is $B/S = (0.005 \pm 0.002)$, and can therefore be considered as negligible.

Rather than imposing a D_s mass veto which would possibly result in a significant loss of signal statistics, we could test these specific events with a dedicated $B_s \to D_sD_s$ selection. This possibility has however not been investigated as the specific B/S contribution is rather low.

The $B_s \to D_sD_s$ decay channel could actually be considered as signal for the determination of ϕ_s since it originates from color-allowed $\bar{b} \to \bar{c}c\bar{s}$ transitions, and provided final state interactions are negligible.

$B_s \to D_s^*D_s$ background

The $B_s \to D_s^*D_s$ decay mode, with $D_s^* \to D_s\gamma$ or $D_s\pi^0$ and $D_s \to KK\pi$, can for the same reasons as $B_s \to D_sD_s$ be considered as a source of background. However, since we will miss the neutral particle it is unlikely that these events end up in the signal region. Note that in the generated sample there is only *one* D_s^*. The results of the $B_s \to \eta_c\phi$ selection on this sample are:

$N_{\text{gen}}^{\text{spec}}$ of events analyzed	$N_{\text{sel}}^{\text{spec}}$ selected in loose δm	$N_{\text{sel}}^{\text{spec}}$ selected in tight δm
170000	17	1

All the reconstructed combinations using all charged final states from the specific decay end up outside the tight mass region. The only event passing the tight mass cut uses five tracks from the specific background, and one from the other b hadron. This event is counted as combinatorial background such that $B/S < 0.011$ at 90% CL. This contribution can be ignored. The branching ratios used are from [27].

$B_s \to D_s\pi\pi\pi$ background

This channel is the specific background resembling the most to the $B_s \to \eta_c\phi$ signal. Events with a short-lived D_s are perfectly suited to pass the final signal selection. The results are:

$N_{\text{gen}}^{\text{spec}}$ of events analyzed	$N_{\text{sel}}^{\text{spec}}$ selected in loose δm	$N_{\text{sel}}^{\text{spec}}$ selected in tight δm
668500	37	36

All the selected events use the specific background final states. $B_s \to D_s\pi\pi\pi$ has a high visible branching ratio in comparison to the signal and represents the dominant specific background for $B_s \to \eta_c\phi$ as it can be seen from the resulting $B/S = (0.20 \pm 0.08)$. The branching ratios used [27] in this estimate are:

$$\begin{aligned} \text{BR}_{\text{vis}}^{B_s \to D_s\pi\pi\pi} &= \text{BR}(B_d \to D\pi\pi\pi) \times \text{BR}(D_s \to KK\pi) = (4.16 \pm 1.49) \times 10^{-4}\,, \\ \text{BR}(B_d \to D\pi\pi\pi) &= (8.0 \pm 2.5) \times 10^{-3}\,. \end{aligned}$$

The B/S contribution from this specific background is in principle already included in the background level intervals from $b\bar{b}$ given in (5.8). Nevertheless, no event of this type was found after the final signal selection on the inclusive $b\bar{b}$ sample due to the limited statistics. The total fraction of $B_s \to D_s\pi\pi\pi$ events in the generic sample is 2.91×10^{-4} (from EVTGEN) such that we would expect to select ~ 0.4 of such events when running on

~ 27 M $\mathrm{b\bar{b}}$ events, given the $\mathrm{B_s} \to \eta_c\phi$ selection efficiency of $\sim 5.4 \times 10^{-5}$ on this specific sample. This is compatible with the fact that no event was found in the $\mathrm{b\bar{b}}$ sample.

We can also note that $\mathrm{B_s} \to \mathrm{D_s}\pi\pi\pi$ may serve as a control channel for $\mathrm{B_s} \to \eta_c\phi$. This control channel being a flavor-specific channel its background contamination for $\mathrm{B_s} \to \eta_c\phi$ could spoil the signal CP-asymmetry measurement.

$\mathrm{B_d} \to$ DD background

As for its strange counterpart, $\mathrm{B_d} \to \mathrm{DD}$ with $\mathrm{D} \to \pi\pi\mathrm{K}$ could possibly be reconstructed as a signal candidate. However, we would need to have a double misidentification of the final states to select it as $\mathrm{B_s} \to \eta_c\phi$. We can thus test our PID cuts with this specific channel. Given the large pion-kaon mass difference, it is basically impossible to select events in the signal region. The results are:

$N_{\mathrm{gen}}^{\mathrm{spec}}$ of events analyzed	$N_{\mathrm{sel}}^{\mathrm{spec}}$ selected in loose δm	$N_{\mathrm{sel}}^{\mathrm{spec}}$ selected in tight δm
326000	0	0

The B/S contribution can be calculated from:

$$\left(\frac{B}{S}\right)^{\mathrm{B_d}\to\mathrm{DD}} = \left(\frac{\varepsilon^{\theta}_{\mathrm{sig\,B_d}}}{\varepsilon^{\theta}_{\mathrm{sig\,B_s}}}\right) \times \left(\frac{\mathrm{BR}\left(\bar{\mathrm{b}} \to \mathrm{B_d}\right)}{\mathrm{BR}\left(\bar{\mathrm{b}} \to \mathrm{B_s}\right)}\right) \times \left(\frac{\mathrm{BR}_{\mathrm{vis}}^{\mathrm{B_d}\to\mathrm{DD}}}{\mathrm{BR}_{\mathrm{vis}}}\right) \times \left(\frac{N_{\mathrm{sel}}^{\mathrm{spec}}/N_{\mathrm{gen}}^{\mathrm{spec}}}{N_{\mathrm{sel}}/N_{\mathrm{gen}}}\right), \quad (5.10)$$

where $\varepsilon^{\theta}_{\mathrm{sig\,B_d}} = 34.9\%$ is the efficiency of the 400 mrad cut for the signal $\mathrm{B_d}$ sample, see Section 3.2.3, and $\mathrm{BR}\left(\bar{\mathrm{b}} \to \mathrm{B_d}\right) = 39.8\%$ [27]. As no event is selected, we are only able to set a 90% confidence upper limit $B/S < 4 \times 10^{-4}$. This specific background is therefore not dangerous. The branching ratios used are from [27].

5.7.4 Inclusive $\mathrm{D_s}$ background

As we have seen from the previously studied backgrounds, large multiplicity b-hadron decays with a $\mathrm{D_s}$ in their decay chain are likely to be selected by the $\mathrm{B_s} \to \eta_c\phi$ selection. Yet another source of background to be analyzed is prompt $\mathrm{D_s}$ decays, i.e. not originating from a b hadron. These decays will also produce detached vertices and could be reconstructed as fake $\mathrm{B_s} \to \eta_c\phi$ combinations.

The inclusive $\mathrm{D_s}$ sample was generated with an admixture of prompt $\mathrm{D_s}$ and $\mathrm{D_s}$ coming from b hadrons. The contribution from $\mathrm{D_s}$ originating from b hadrons is already included in the B/S ratio we obtained from the $\mathrm{b\bar{b}}$ sample, and thus we are only interested in the prompt component of the inclusive $\mathrm{D_s}$ sample. Given that the fraction of inclusive $\mathrm{D_s}$ from b decays in the inclusive $\mathrm{b\bar{b}}$ sample and in the inclusive $\mathrm{D_s}$ sample are identical, we can estimate the fraction of prompt $\mathrm{D_s}$ ($f_{\mathrm{pr\,D_s}}$) using the following back-of-the-envelope calculation:

$$\begin{aligned}
\sigma^{\mathrm{Py}}_{\mathrm{b\bar{b}}} &= 0.627\ \mathrm{mb}\,, \qquad \sigma^{\mathrm{Py}}_{\mathrm{D_s}} = \sigma^{\mathrm{Py}}_{\mathrm{D_s^+}} + \sigma^{\mathrm{Py}}_{\mathrm{D_s^-}} = (0.508 + 0.496)\ \mathrm{mb} = 1.004\ \mathrm{mb}\,, \\
\mathrm{BR}(\mathrm{b} \to \mathrm{D_s X}) &= 28.1\%\,, \\
\sigma(\mathrm{b\bar{b}} \to \mathrm{D_s X}) &= 2 \times \sigma^{\mathrm{Py}}_{\mathrm{b\bar{b}}} \times \mathrm{BR}(\mathrm{b} \to \mathrm{D_s X}) = 0.352\ \mathrm{mb}\,, \\
&\Rightarrow\ \sigma(\mathrm{c\bar{c}} \to \mathrm{D_s X}) = \sigma^{\mathrm{Py}}_{\mathrm{D_s}} - \sigma(\mathrm{b\bar{b}} \to \mathrm{D_s X}) = 0.652\ \mathrm{mb}\,, \\
&\Rightarrow\ f_{\mathrm{pr\,D_s}} = \frac{\sigma(\mathrm{c\bar{c}} \to \mathrm{D_s X})}{\sigma^{\mathrm{Py}}_{\mathrm{D_s}}} = 65\%\,.
\end{aligned}$$

In the above calculation we used the measured PYTHIA cross sections from the generation [57] and the branching ratios are from [27].

The $B_s \to \eta_c\phi$ does not select any candidate in events with a prompt D_s in the loose mass window ± 500 MeV/c^2. Making use of the mass window trick and with the fraction $f_{\mathrm{pr\,D_s}}$ estimated above, the background level from prompt D_s is calculated as:

$$\left(\frac{B}{S}\right)^{\mathrm{pr\,D_s}} = \left(\frac{\varepsilon^{\theta}_{\mathrm{D_s}}}{\varepsilon^{\theta}_{\mathrm{sig\,B_s}}}\right) \times \left(\frac{\sigma^{\mathrm{Py}}_{\mathrm{D_s}} \times f_{\mathrm{pr\,D_s}} \times \mathrm{BR}(\mathrm{D_s} \to \mathrm{KK}\pi)}{\sigma_{\mathrm{b\bar{b}}} \times 2 \times \mathrm{BR}\left(\bar{\mathrm{b}} \to \mathrm{B_s}\right) \times \mathrm{BR_{vis}}}\right) \times \left(\frac{N^{\mathrm{pr\,D_s}}_{\mathrm{sel}} / \left(F_m \times N^{\mathrm{pr\,D_s}}_{\mathrm{gen}}\right)}{N_{\mathrm{sel}}/N_{\mathrm{gen}}}\right), \tag{5.11}$$

where:

- $\varepsilon^{\theta}_{\mathrm{D_s}} = 36.5\%$ is the acceptance of 400 mrad requirement for the D_s sample, see Section 3.2.3;
- $N^{\mathrm{pr\,D_s}}_{\mathrm{sel}} = 0$ is the number of events passing the final $B_s \to \eta_c\phi$ selection and not originating from a b hadron;
- $N^{\mathrm{D_s}}_{\mathrm{gen}} = 9'826'209$ is the number of inclusive D_s events analyzed.

Using the unified confidence intervals we get a 90% upper limit of $B/S < 0.077$. We can therefore conclude that prompt D_s do not represent a sizable source of background for $B_s \to \eta_c\phi$. This is not surprising as we would need to add to the detached D_s three extra random tracks, which is very unlikely given the tight final selection cuts applied.

For completeness we show in Figure 5.13 the distributions of the mass and of the fitted proper time (see Section 5.8.3) of the reconstructed B_s candidates for a few specific background channels. Note that all the selected candidates are shown. For instance for the $B_s \to D_s\pi\pi\pi$ sample, some events have one candidate from the specific decay and one candidate with a random track. Finally, the $B_s \to \eta_c\phi$ off-line selection was also tested on ~ 3.28 s of minimum-bias events at 40 kHz (i.e. after L0 and L1) and no event was selected. Given the suppression factor of 160 between minimum-bias and $b\bar{b}$ events, the statistics are much too small to set any meaningful background level. Nevertheless, it represents a nice verification of the selection on a light quark-content background.

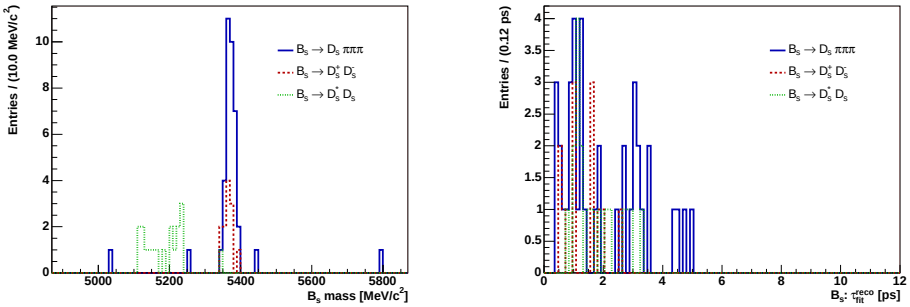


Figure 5.13: B_s mass [MeV/c^2] (left plot) and fitted proper time [ps] (right plot) of the reconstructed $B_s \to \eta_c\phi$ candidates in some specific background samples. The histograms are not normalized.

5.8 Reconstruction Quality and Selection Bias

We present in this section the different resolutions on the selected $B_s \to \eta_c\phi$ signal. As the physics purpose is the determination of the weak B_s–$\overline{B}_s$ mixing phase, an excellent proper time resolution is required in order to resolve the fast B_s–$\overline{B}_s$ oscillations. We will thus determine the quality of the observables used in determination of the proper time, as well as model the biases induced by the off-line and trigger selections.

We introduce the following definitions used to discuss the reconstruction quality:

Residual δx Given an observable x, the residual is defined as the difference between the direct measurement x_{rec} (reconstructed or fitted quantity) and the true Monte Carlo value x_{MC}: $\delta x = x_{\text{rec}} - x_{\text{MC}}$. The residual thus represents a measure of the error on x.

Resolution σ_x^{res} The resolution is defined as the standard deviation of a Gaussian fit to the residual δx.

Pull The pull (or stretch) value of a variable x is the residual δx normalized to the estimated error σ_x^{err} of this difference. The error is generally determined from the covariance matrix of x_{rec}. Assuming Gaussian errors, the pull $(x_{\text{rec}} - x_{\text{MC}})/\sigma_x^{\text{err}}$ should exhibit a Normal Gaussian distribution with zero mean, i.e. no bias, and unit width (denoted by Σ_x). The deviations from a normal distribution will allow us to identify wrong error assignments and/or incorrect assumptions.

Note that we will abusively use the same symbol σ for resolutions, errors or pulls in many of the plots shown in this chapter. Unless otherwise specified, all the fits to the distributions shown use the χ^2 method as estimator. We will perform in some cases unbinned fits using the negative-log-likelihood method ($-\ln L$). Finally, note that all histograms entries for the plots shown in the next sections are for candidates associated to the $B_s \to \eta_c\phi$ signal and passing the corresponding off-line selection, and in general before applying the trigger selections.

5.8.1 Mass and Momentum Resolutions

The mass cuts imposed in the event selection represent the most powerful criteria as they suppress most of the fake combinations. The cuts should take into account the momentum reconstruction quality as we cannot reconstruct particle trajectories with an "infinite" precision. A measure of this quality is given by the mass and momentum resolutions. Moreover, the intrinsic width of resonances should be considered in the determination of the mass cuts.

The distributions of the reconstructed ϕ and η_c masses are shown in the left plots of Figure 5.14, which we fitted with single Gaussian functions in order to quantify our mass cuts in terms of standard deviations (σ). Note that these distributions are of course not Gaussians, but this enables us to discuss the mass cuts in a simple way. The mean of the fits are within errors perfectly compatible with the generated values, see Section 5.1.1. The widths of the Gaussians show us that the mass cuts correspond to $\sim 5\sigma$ on both sides of the central values.

The mass resolutions are obtained by fitting the residual mass distributions, see the right plots of Figure 5.14, where the resolutions are $\sigma_m^{\text{res}}(\phi) = (0.95 \pm 0.02)\ \text{MeV/c}^2$ and $\sigma_m^{\text{res}}(\eta_c) = (8.68 \pm 0.17)\ \text{MeV/c}^2$ (the true generated `HepMC` mass is *not* a constant). Note

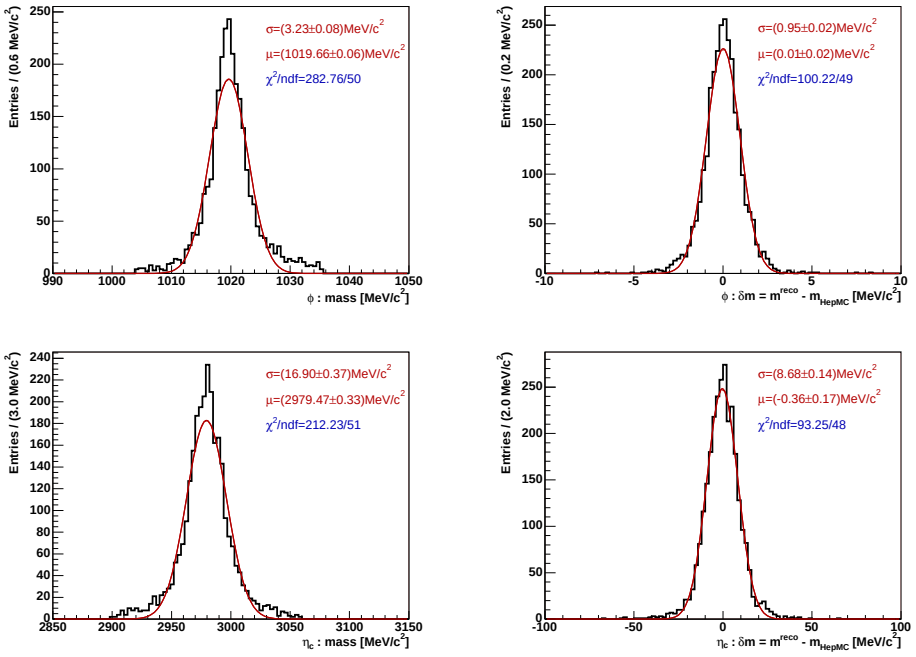


Figure 5.14: ϕ (top) and η_c (bottom) reconstructed masses [MeV/c^2] (left) and mass residuals [MeV/c^2] (right). The superimposed curves are the result of single Gaussian fits.

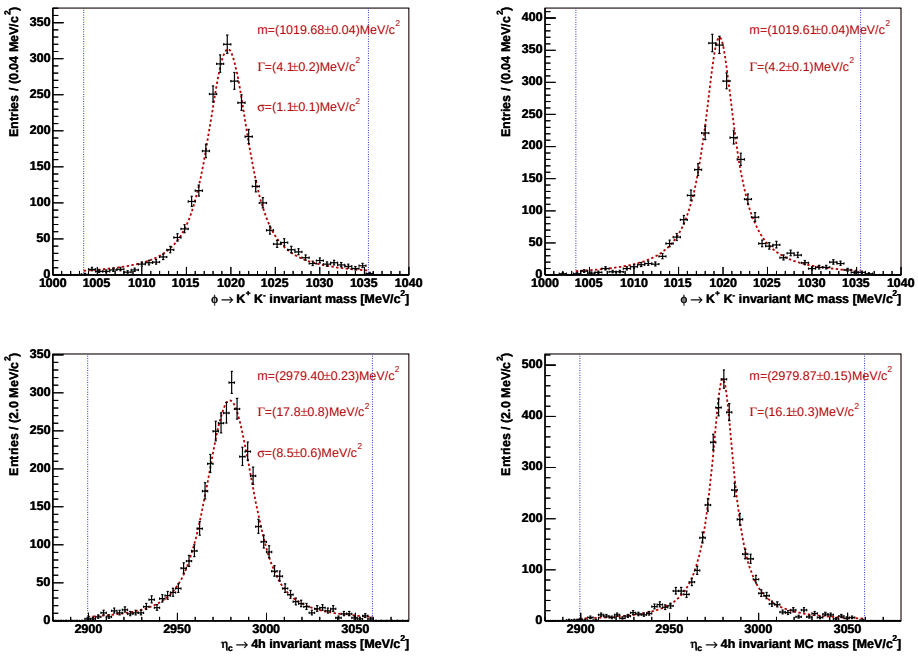


Figure 5.15: ϕ (top) and η_c (bottom) invariant masses $m_{\rm rec}$ [MeV/c^2] (left) and true generated (HepMC) masses $m_{\rm MC}$ [MeV/c^2] (right). The superimposed curves are the result of a Voigtian and Breit–Wigner likelihood fits to $m_{\rm rec}$ and $m_{\rm MC}$ respectively.

a small $\sim 2\sigma$ bias for the reconstructed η_c mass. As it will be done with real data, we can also extract the mass resolution of these resonances by means of a Voigtian[5] fit to the reconstructed mass, as shown on the left plots of Figure 5.15. Using an unbinned $(-\ln L)$ fit, we get:

$$\begin{aligned}
\phi: &\quad m = (1019.68 \pm 0.04)\,\mathrm{MeV/c^2}, \quad \Gamma = (4.1 \pm 0.2)\,\mathrm{MeV/c^2}, \quad \sigma = (1.1 \pm 0.1)\,\mathrm{MeV/c^2};\\
\eta_c: &\quad m = (2979.40 \pm 0.23)\,\mathrm{MeV/c^2}, \quad \Gamma = (17.8 \pm 0.8)\,\mathrm{MeV/c^2}, \quad \sigma = (8.5 \pm 0.6)\,\mathrm{MeV/c^2}.
\end{aligned}$$

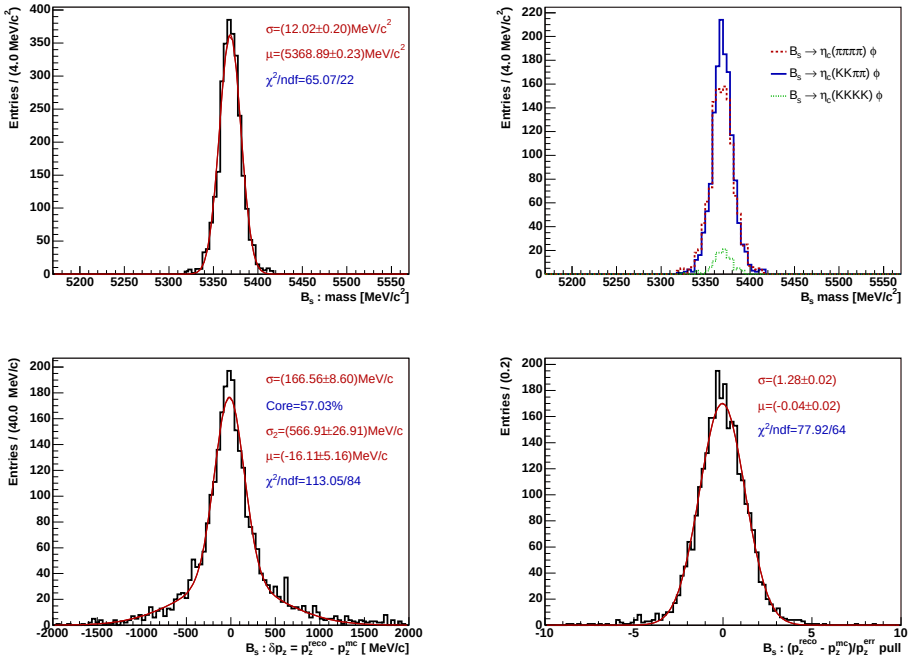


Figure 5.16: Top: invariant B_s reconstructed mass [$\mathrm{MeV/c^2}$] on the left, with the different contributions on the right. Bottom: p_z residual [MeV/c] (left) and pull (right). The superimposed curves are the result of Gaussian fits.

The fit for η_c is not perfect as it gives a too large Γ in comparison to the generated value. This is visible on the plot where the superimposed curve is off the data around the central mass value. A more precise value is obtained when fitting the true generated MC mass with a Breit–Wigner as shown on the right plots of Figure 5.15. The fit results are in this case $\Gamma(\phi) = (4.2 \pm 0.1)\,\mathrm{MeV/c^2}$ and $\Gamma(\eta_c) = (16.1 \pm 0.3)\,\mathrm{MeV/c^2}$, compatible with the values used in the generation given in Section 5.1.1.

The mass and momentum resolutions of the B_s are important quantities for the proper time determination. A badly reconstructed momentum can significantly deteriorate the proper time resolution as it is the case for decays involving photons. The reconstructed B_s mass is shown in the top plots of Figure 5.16, where we also show the contributions from each of the signal decay modes. The mass resolution is $\sigma_m^{\mathrm{res}}(B_s) = (12.0 \pm 0.2)\,\mathrm{MeV/c^2}$, with no bias. The tight mass cut thus corresponds to $\sim 4\sigma$ on both sides of the central mass. The momentum resolution is very precise: $\sigma_p^{\mathrm{res}}(B_s) = 0.22\%$. Due to the

[5] A Voigtian (V) is the convolution of a Breit–Wigner (BW) distribution with a Gaussian: $V(x, m, \Gamma, \sigma) \propto BW(x, m, \Gamma) \otimes \exp\left(-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right)$, with $BW(x, m, \Gamma) \propto \frac{1}{(x-m)^2 + \frac{1}{4}\Gamma^2}$.

large boost in the beam direction z, the dominant contribution to the momentum resolution is induced by the z component of the momentum. The p_z residual is shown in Figure 5.16 together with its pull. The resolution obtained with a double Gaussian fit yields: $\sigma_{p_z}^{\text{res}}(\text{B}_\text{s}) = (167\,[57\%] + 567\,[43\%])\,\text{MeV/c}$. From the corresponding pull, we note that the errors on p_z are underestimated by $\sim 28\%$. For completeness, we give the momentum resolutions and pulls for each of the components and with single Gaussian fits for simplicity:[6]

$$\begin{aligned}
p_x(\text{B}_\text{s}): &\quad \sigma_{p_x}^{\text{res}} = 17.2\,\text{MeV/c}, \quad \Sigma_{p_x} = 1.20 \pm 0.02\,,\\
p_y(\text{B}_\text{s}): &\quad \sigma_{p_y}^{\text{res}} = 16.5\,\text{MeV/c}, \quad \Sigma_{p_y} = 1.20 \pm 0.02\,,\\
p_z(\text{B}_\text{s}): &\quad \sigma_{p_z}^{\text{res}} = 393\,\text{MeV/c}, \quad \Sigma_{p_z} = 1.28 \pm 0.02\,.
\end{aligned}$$

From the above results we see that the transverse components have much better resolutions in comparison to the z direction.

5.8.2 Vertex Resolutions

The precise determination of the origin and decay vertices of the B_s is essential for a good proper time resolution. As we will see later, it is actually the secondary vertex determination which is critical for lifetime studies. As we could naively expect, the larger the multiplicity of a decay is, the better its vertex determination will be. This is confirmed with the $\text{B}_\text{s} \to \eta_\text{c}\phi$ channel, where its six legs are very helpful.

The B_s decay vertex residual and pull in the z direction are shown in Figure 5.17, where the corresponding quantities for the primary vertex assigned[7] to the B_s are also given. The resolution of the primary vertex is $\sigma_{\text{P}_z}^{\text{res}} \sim 45\,\mu\text{m}$. This good resolution is achieved thanks to the large variety of tracks used to determine the PV, especially with large angle. The significant positive bias is induced by bottom and charm decay products possibly used in the vertex fit. For the B_s vertex the core (69%) resolution is $\sigma_{\text{S}_z}^{\text{res}} \sim 145\,\mu\text{m}$. This resolution is worse compared to that of the primary vertex because signal tracks are mainly produced at low angles. The detail of the resolution and pull for each components of the primary and secondary vertices is given hereafter:

$$\begin{aligned}
\text{P}_x: &\quad \sigma_{\text{P}_x}^{\text{res}} = 8.6\,\mu\text{m}, \quad \Sigma_{\text{P}_x} = 1.32 \pm 0.02\,,\\
\text{P}_y: &\quad \sigma_{\text{P}_y}^{\text{res}} = 7.9\,\mu\text{m}, \quad \Sigma_{\text{P}_y} = 1.26 \pm 0.02\,,\\
\text{P}_z: &\quad \sigma_{\text{P}_z}^{\text{res}} = 45.1\,\mu\text{m}, \quad \Sigma_{\text{P}_z} = 1.16 \pm 0.02\,.
\end{aligned}$$

$$\begin{aligned}
\text{S}_x(\text{B}_\text{s}): &\quad \sigma_{\text{S}_x}^{\text{res}} = 14.2\,\mu\text{m}, \quad \Sigma_{\text{S}_x} = 1.06 \pm 0.01\,,\\
\text{S}_y(\text{B}_\text{s}): &\quad \sigma_{\text{S}_y}^{\text{res}} = 13.8\,\mu\text{m}, \quad \Sigma_{\text{S}_y} = 1.11 \pm 0.02\,,\\
\text{S}_z(\text{B}_\text{s}): &\quad \sigma_{\text{S}_z}^{\text{res}} = 261\,\mu\text{m}, \quad \Sigma_{\text{S}_z} = 1.06 \pm 0.02\,.
\end{aligned}$$

The transverse directions have much better resolutions simply because of the lower momentum perpendicular to the beam.

[6] For the z component we simply compute the resolution with $\sigma = \left(f_1 \times \sigma_1^2 + f_2 \times \sigma_2^2\right)^{1/2}$ using the result of the previous double Gaussian fit.

[7] The primary vertex assigned to the B_s is the primary vertex for which the IPS of the B_s is the smallest. In our discussion we ignore wrong primary vertex assignments.

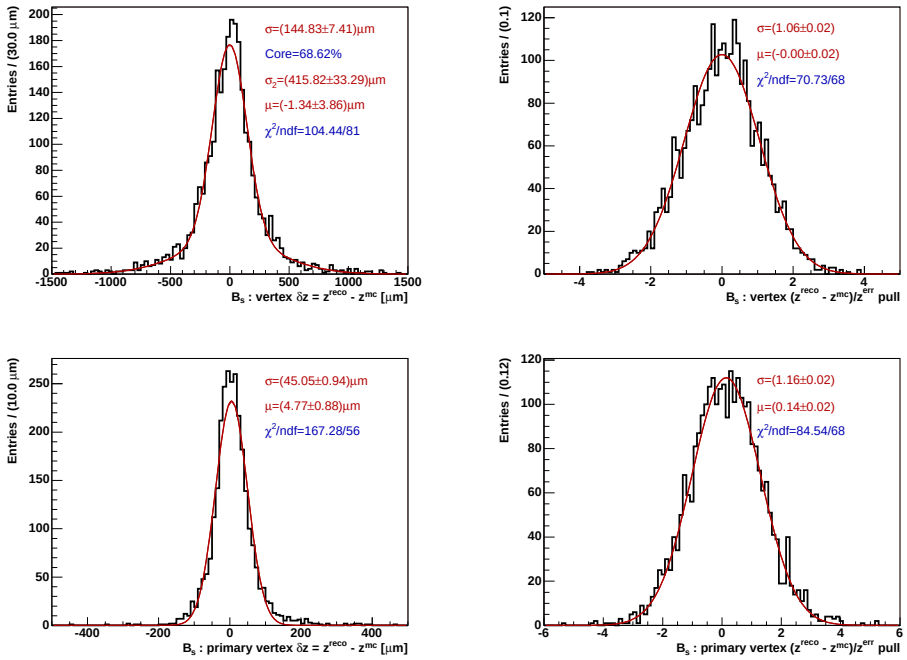


Figure 5.17: Left plots: secondary (top) and primary vertex (bottom) residuals in the z direction [μm]. Right plots: pulls of the z position of the secondary (top) and of the primary (bottom) vertices. The superimposed curves are the result of Gaussian fits.

5.8.3 B_s Proper Time

The key experimental observable in the time-dependent CP asymmetry is the proper time. Its measurement should be precise enough to be able to resolve the fast oscillating B_s meson.

The proper time τ is linked to the physical observable t in the laboratory through $t = \gamma\tau = (1 - \vec{\beta}^{\,2})^{-1/2}$, where γ is the relativistic Lorentz factor of a B_s traveling at a velocity $\vec{\beta}c$. The B_s flight distance is given by $\vec{\mathrm{FD}} = c\vec{\beta}t = \gamma\vec{\beta}c\tau = (\vec{p}/m)\tau$, where $\vec{p}$ and m are the reconstructed B_s momentum and mass, respectively. We therefore see that the proper time can be determined from:

$$\vec{\mathrm{FD}} = \vec{\mathrm{S}} - \vec{\mathrm{P}} = \tau\frac{\vec{p}}{m} \quad \leftrightarrow \quad \tau = m\frac{\vec{p}\cdot\vec{\mathrm{FD}}}{|\vec{p}|^2}\,, \tag{5.12}$$

where $\vec{\mathrm{P}}$ and $\vec{\mathrm{S}}$ are respectively the positions of the primary and secondary vertices.

Proper time fit

In order to extract the proper time and its error on an event-by-event basis, a fit[8] is used constraining the kinematics of the decay through (5.12). We will use the method of least-squares, following the notations of [27].

[8] We use the `LifetimeFitter` tool for this fit, introduced in Reference [102].

We have nine observables which we can write as a vector of measurements $\vec{y}$:

$$\vec{y} = \left(\vec{\mathrm{S}}, \quad \vec{p}, \quad \vec{\mathrm{P}}\right)^T \qquad \dim(\vec{y}) = 9 \times 1\,.$$

where $\vec{y}$ is understood as a column vector, and T denotes the transposition. Since in general the measurements are not independent, we must introduce their covariance matrix V:

$$V_{ij} = \mathrm{cov}[y_i, y_j]\,, \quad i,j = 1,2,\ldots,N = 9\,, \quad \dim(V) = 9 \times 9\,.$$

After the B_s selection it is unlikely that any of the B_s decay products could have been used in the primary vertex fit. We can therefore assume that the primary vertex determination and the B_s observables are *independent*. This simplifies the (symmetric) covariance matrix as $V_{7k} = 0 = V_{k7}$ for $k = 7,8,9$. Furthermore, this assumption enables us to have a non-degenerate χ^2 function. Note that even though the correlations between the B_s observables and the primary vertex measurements are ignored, the primary vertex errors are taken into account in the χ^2 fit. Moreover, there is *no* refit of the primary vertex after a B_s candidate is found. We can now introduce the seven unknown parameters $\vec{\theta}$ describing the flight of the B_s:

$$\vec{\theta} = \left(\vec{\mathrm{S}}_{\mathrm{fit}}, \quad \vec{p}_{\mathrm{fit}}, \quad \tau_{\mathrm{fit}}\right)^T\,, \quad \dim(\vec{\theta}) = 7 \times 1\,.$$

We have $N = 9$ observables $\vec{y}$, $n = 7$ unknown parameters $\vec{\theta}$ and one constraint, (5.12). The set of parameters $\vec{\theta}$ is then determined by finding the least-square estimators $\vec{\theta}^*$ that minimize the following χ^2:

$$\begin{aligned} \chi^2(\vec{\theta}) &= \left(\vec{y} - \vec{F}(\vec{\theta})\right)^T V^{-1} \left(\vec{y} - \vec{F}(\vec{\theta})\right)\,, \\ &\to \vec{\nabla}_{\vec{\theta}}\chi^2(\vec{\theta})\Big|_{\vec{\theta}=\vec{\theta}^*} = \vec{0}\,, \end{aligned} \tag{5.13}$$

where $\vec{F}(\vec{\theta})$ is the vector of predicted values (9×1 column vector). The χ^2 is hence related to the measurement residuals and their weight matrix. In our problem $\vec{F}$ is a linear function of the parameters θ_i, and it is given by:

$$\vec{F}(\vec{\theta}) \equiv H\vec{\theta} = \left(\vec{\mathrm{S}}_{\mathrm{fit}}, \quad \vec{p}_{\mathrm{fit}}, \quad \vec{\mathrm{S}}_{\mathrm{fit}} - \tau_{\mathrm{fit}}\frac{\vec{p}_{\mathrm{fit}}}{m}\right)^T\,,$$

where we used the constraint (5.12) for the last component of $\vec{F}$. Note that the mass m is the *reconstructed* mass and it is not considered as a parameter. The linear transformation of the parameters is denoted by H, and it is understood as a 9×7 matrix.

The minimization of (5.13) therefore reduces to solving a system of $n = 7$ linear equations yielding the least-squares estimators $\vec{\theta}^*$:

$$\vec{\theta}^* = \left[U\, H^T\, V^{-1}\right]\vec{y}\,,$$

where $U_{ij} = \mathrm{cov}(\theta_i^*, \theta_j^*)$ is the 7×7 covariance matrix for the estimators given by:

$$U = \left(H^T\, V^{-1}\, H\right)^{-1} \quad \leftrightarrow \quad (U)^{-1}_{ij} = \frac{1}{2}\,\frac{\partial^2\chi^2}{\partial\theta_i\partial\theta_j}\bigg|_{\vec{\theta}=\vec{\theta}^*}\,.$$

The errors on our seven parameters are thus given by the second order derivatives of the χ^2 with respect to the least-squares estimators.

Proper time fit results

The B_s proper time is obtained from the lifetime fit described above. The fitted proper time $\tau_{\rm fit}^{\rm rec}$ and its per-event error $\tau_{\rm fit}^{\rm err}$ are displayed in Figure 5.18, together with the residual $\delta\tau = \tau_{\rm fit}^{\rm rec} - \tau_{\rm MC}$ and pull $(\tau_{\rm fit}^{\rm rec} - \tau_{\rm MC})/\tau_{\rm fit}^{\rm err}$ (with width Σ_τ) distributions.

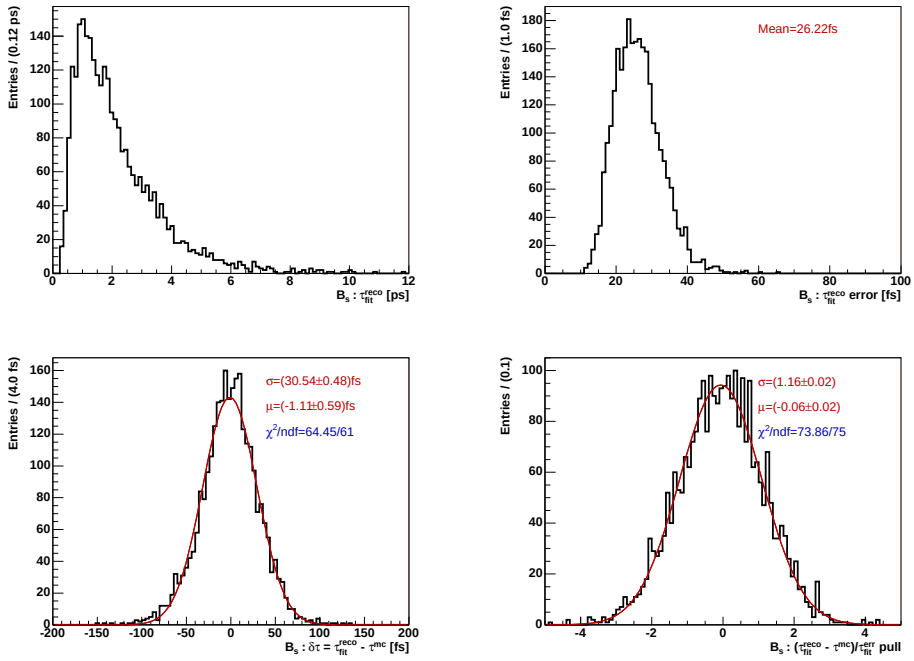


Figure 5.18: Results of the proper time fit: $\tau_{\rm fit}^{\rm rec}$ [ps] (top left), $\tau_{\rm fit}^{\rm err}$ [fs] (top right), $\delta\tau$ [fs] (bottom left) and pull (bottom right). The superimposed curves are the result of Gaussian fits.

The depletion at low proper times, which is due to our selection criteria biasing the proper time, e.g. impact parameter and flight distance cuts, will be discussed later on. The errors $\tau_{\rm fit}^{\rm err}$ vary in a large range that can be explained by the spread in the kinematic distributions of the momentum and decay distance. The distribution of the errors start at ~ 10 fs which sets a lower bound on the resolution that can be achieved. The mean error $< \tau_{\rm fit}^{\rm err} >= 26.2$ fs is a sign that $B_s \to \eta_c\phi$ reaches an excellent proper time performance, provided the errors are well estimated. For correctly estimated errors, we should have that $< \tau_{\rm fit}^{\rm err} > \approx \sigma_\tau$, where $\sigma_\tau = 30.5$ fs is the proper time resolution obtained from the width of a Gaussian fit to $\delta\tau$. This is not the case as there is a $\sim 14\%$ underestimation of the errors. This is confirmed by the proper time pull which displays a standard deviation of 1.16, indicating that the errors are underestimated by $\sim 16\%$. The $B_s \to \eta_c\phi$ proper time performance is summarized hereafter:

$$\begin{aligned}
\text{Resolution}: &\quad \sigma_\tau = (30.5 \pm 0.5)\,\text{fs}, \quad \text{mean} = (-1.1 \pm 0.6)\,\text{fs}\,; \\
\text{Mean error}: &\quad \tau_{\rm fit}^{\rm err} = 26.2\,\text{fs}\,; \\
\text{Scale factor}: &\quad \Sigma_\tau = 1.16 \pm 0.02\,, \quad \text{mean} = (-0.06 \pm 0.02)\,.
\end{aligned}$$

A small $\sim 2\sigma$ bias is observed in the proper time resolution and will be explained in the next subsection. As we will see in Chapter 6, the $B_s \to \eta_c\phi$ channel is the one with the

best proper time determination. All channels presented in the next chapter have both a larger average proper time error and scale factor. Thus the proper time performance of $B_s \to \eta_c \phi$ represents a significant advantage with respect to other $\bar{b} \to \bar{c}c\bar{s}$ transitions in the determination of ϕ_s.

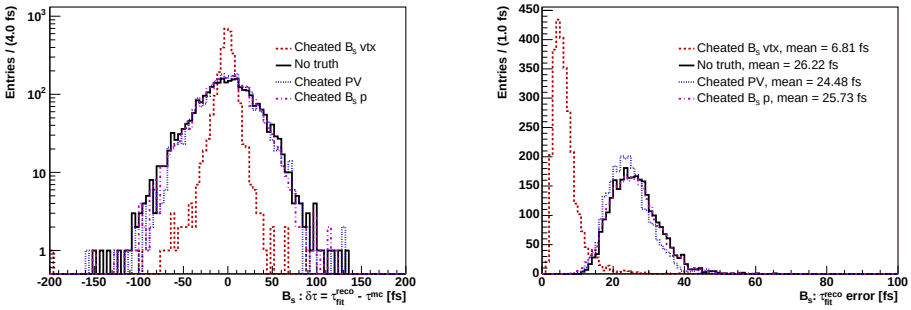


Figure 5.19: Results of the proper time fit: $\delta\tau$ [fs] (left) and $\tau_{\text{fit}}^{\text{err}}$ [fs] (right). The different fits shown are: no MC truth (solid black line), cheated secondary vertex (dotted red line), cheated primary vertex (dashed blue line) and cheated B_s momentum (dashed-dotted magenta line).

Proper time contributions

The proper time pull represents a probe of the error assignment, however it does not tell us what are the dominant contributions but just that on "average" our proper time errors are underestimated. From the definition of the proper time (5.12), we have 9 observables contributing in different ways to the overall error. In order to disentangle the contributions, we can perform several lifetime fits each time fixing one of the observables $\vec{S}$, $\vec{p}$ and $\vec{P}$ to its true MC value. The results of the cheated lifetime fits are shown in Figure 5.19, for the $\delta\tau$ residual and corresponding per-event error $\tau_{\text{fit}}^{\text{err}}$.

The proper time resolution is not affected much by either the primary vertex nor the B_s momentum. However, it is clear that it is the secondary vertex that spoils the measurement, as the residual when using the true B_s vertex position has a much thiner width compared to all other cases. Moreover, the proper time errors are in average ~ 4 times smaller with the cheated secondary vertex. The $\tau_{\text{fit}}^{\text{err}}$ distribution is much more peaked and the core is well below 20 fs. We give in Table 5.15 the full detail of the different contributions.

We observe from the different resolutions given in Table 5.15 that the key parameter is the secondary vertex position. A proper time resolution of ~ 7 fs is obtained when setting this observable to its true value. The biases in the proper time resolution and pull disappear when using the true primary vertex position. The primary vertex is thus the main cause of this bias, as already observed in Section 5.8.2. One could possibly get rid of this bias by refitting the primary vertex and trying to identify and ignore decay products from c and b hadrons.

Regarding the scale factors, the variations are not significant enough to draw any definite conclusion. Each observable seems to equivalently yield larger pulls. In the determination of the B_s–$\overline{B}_s$ mixing phase, see Chapter 6, the per-event errors will be used as

Table 5.15: Proper time fit results.

	Resolution		Error	Pull	
$B_s \to \eta_c\phi$	σ_τ [fs]	Mean [fs]	$< \tau_{\text{fit}}^{\text{err}} >$ [fs]	Σ_τ	Mean
No MC truth	30.5 ± 0.5	-1.1 ± 0.6	26.2	1.16 ± 0.02	-0.06 ± 0.02
True $\vec{p}$	28.3 ± 0.5	-1.0 ± 0.6	25.7	1.09 ± 0.02	-0.05 ± 0.02
True $\vec{P}$	27.9 ± 0.5	0.0 ± 0.5	24.5	1.13 ± 0.02	-0.00 ± 0.02
True $\vec{S}$	7.0 ± 0.2	-0.6 ± 0.1	6.8	1.12 ± 0.02	-0.13 ± 0.02

inputs of the simulation taking into account the scale factors. Thus the relevant quantity modelling the detector's response in the CP-asymmetry determination is the proper time pull, and not the resolution.

The proper time determination is extremely important for the determination of ϕ_s. Its effect is similar to that of the tagging dilution, as the B_s–$\overline{B}_s$ wiggles will be attenuated by the proper time resolution, naively with a factor $\exp[-\frac{1}{2}(\Delta M_s\,\sigma_\tau)^2]$.

Proper time acceptance

The trigger and off-line selections of $B_s \to \eta_c\phi$ both require the presence of a detached secondary vertex and relatively large impact parameter decay products. The proper time distribution is thus modified such that the probability of a true event to be detected is no longer uniform and depends on the true proper time. This probability is a time-dependent selection efficiency described by a so-called acceptance function.

The determination of the acceptance function on real data is a challenging task. This is currently under study by the LHCb collaboration, and will not be discussed here. In the Monte Carlo world this is however much simpler. The acceptance function can be parameterized by fitting the bin-to-bin ratio of the histogram of the true proper time after all selection criteria to the histogram before any selection. As the latter is a subset of the former, this ratio represents an efficiency as a function of the true proper time. For the histogram before any selection, we use the result of a cheated selection requiring all signal final states to be reconstructed and only applying the $B_s \to \eta_c\phi$ PID preselection cuts. In principle one should use events at generator level without *any* cuts. We thus assume that the tracking does not induce any bias. The proper time selection efficiency plots, after the off-line $B_s \to \eta_c\phi$ selection and the HLT are shown in Figure 5.20.

The acceptance function plotted in Figure 5.20 is the result of a χ^2 fit to the data using the following parameterization:

$$A(\tau_{\text{MC}}) \propto \frac{(s_{\text{low}} \times \tau_{\text{MC}})^3}{1 + (s_{\text{low}} \times \tau_{\text{MC}})^3}\,, \tag{5.14}$$

where s_{low} represents the slope of the raise at low proper times and the proportionality term corresponds to the selection efficiency with respect to the cheated selection at large proper times. The bias at small proper times is clearly visible and it is mainly due to the impact parameter, flight distance and pointing angle requirements.

The acceptance function can be used to describe the reconstructed proper time behavior at small proper times, whereas an exponential decay will dominate the larger proper

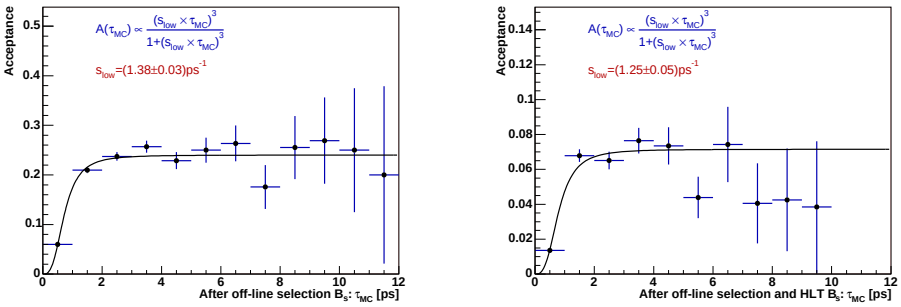


Figure 5.20: Proper time selection efficiency after off-line selection (left) and after the HLT (right), as a function of the true proper time $\tau_{\rm MC}$ [ps]. The superimposed curves are the result of the fit with the parameterization (5.14). The error bars correspond to binomial errors.

times. This is shown in Figure 5.21 where the $\tau_{\rm fit}^{\rm rec}$ distribution is fitted with the product of an exponential with decay constant τ together with the parameterization (5.14), after the off-line selection. The fit yields $\tau = (1.40 \pm 0.04)$ ps, which is consistent with the true MC value $\tau = 1.461$ ps.

The simple parameterization (5.14) does not seem to reproduce correctly the proper time efficiency for $\tau_{\rm MC} \gtrsim 6$ ps, though the available statistics in that regime are limited. Even though most of the events have a proper time below ~ 6 ps and hence the majority of the physics will take place there, we may need to properly describe the acceptance at large proper times to account for all possible effects.

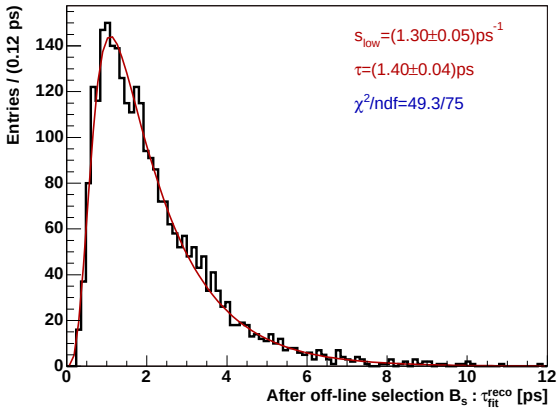


Figure 5.21: Reconstructed proper time $\tau_{\rm fit}^{\rm rec}$ [ps] fitted with the acceptance function (5.14) times an exponential.

In order to characterize the behavior of the time-dependent efficiency at large proper times, we introduce the following parameterization:

$$A(\tau_{\rm MC}) \propto \frac{(s_{\rm low} \times \tau_{\rm MC})^n}{1 + (s_{\rm low} \times \tau_{\rm MC})^n} \left(1 + s_{\rm high} \times \tau_{\rm MC}\right) , \tag{5.15}$$

where $s_{\rm low}$ and n govern the time dependence at small proper times and $s_{\rm high}$ describes the tendency of the acceptance function at large proper times. We can then use this parameterization to fit the proper time plots. The floated parameters are $s_{\rm low}$, n, $s_{\rm high}$ and the normalization. The results are shown in Figure 5.22, after each step of the $B_s \to \eta_c \phi$ event selection namely after off-line, L0, L1 and HLT selections.

We see that the parameterization (5.15) describes very well the behavior at small proper times, but we do not observe any significant deviation from a flat acceptance for large proper times. There seems to be a drop in efficiency above $\tau_{\rm MC} \sim 6$ ps, especially after the L1 and the HLT. For the latter $s_{\rm high}$ exhibits a $\sim 2\sigma$ deviation from zero, with a negative slope. This bias is presumably due to the trigger (L1 and HLT generic) maximum IP cuts for the selection of tracks to start triggering an event.

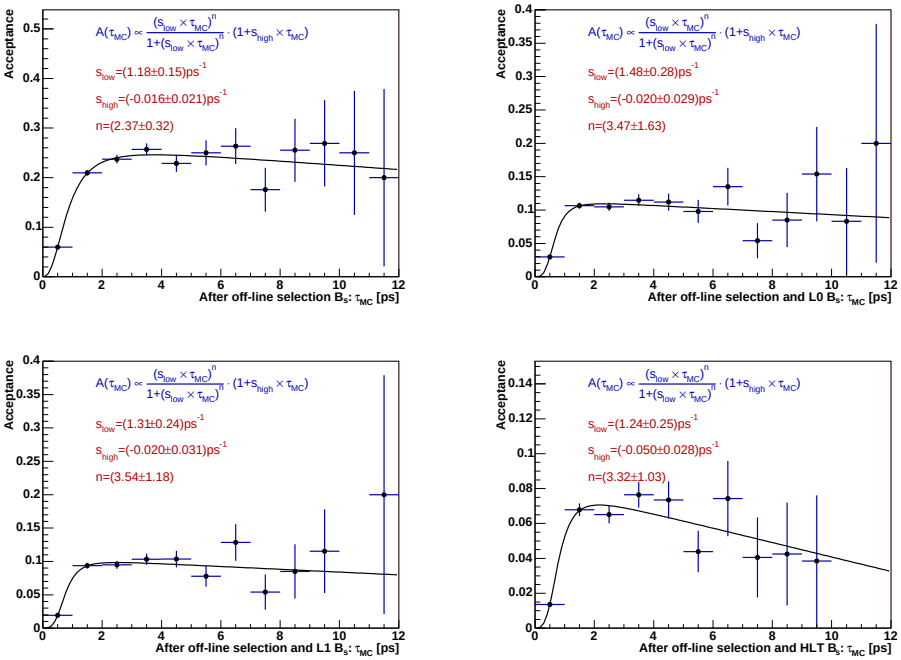


Figure 5.22: Proper time selection efficiency after off-line selection (top left), L0 (top right), L1 (bottom left) and HLT (bottom right), as a function of the true proper time $\tau_{\rm MC}$ [ps]. The superimposed curves are the result of the fit with the parameterization (5.15). The error bars correspond to binomial errors.

The acceptance will be taken into account in the studies of Chapter 6, in order to describe at best what the situation with real data will be. The choice of acceptance function is the parameterization (5.14), as the behavior at large proper times is not well defined. However, if the efficiency really does drop for large proper times, then assuming a flat acceptance function in that regime is a conservative choice. Indeed, a negative slope for large proper times would bring back events to lower proper time values, thus increasing our statistics there.

A non-uniform proper time efficiency has the consequence of removing events at small proper time and hence the negative tail of the distribution. This makes the resolution determination impossible for lifetime-biasing selections, and it must be extracted from somewhere else, e.g. using prompt di-muons or control samples.

5.9 Full Monte Carlo Simulation Results for $B_s \to \eta_c \phi$

We summarize in this section the main results of the $B_s \to \eta_c \phi$ selection from the full Monte Carlo simulation, in particular those that will be used as inputs for the determination of the B_s–$\overline{B}_s$ mixing parameters in Chapter 6.

B_s mass resolution: $\sigma_m^{\rm res}(B_s) = 12.0\,{\rm MeV/c^2}$.

Per-event proper time errors: $< \tau_{\rm fit}^{\rm err} >= 26.2$ fs. These errors will be provided as a data sample for the likelihood fitting. The proper time resolution is $\sigma_\tau = 30.5$ fs.

Scale factor: the standard deviation of the pull distribution $\Sigma_\tau = 1.16$. The effect of $\pm 10\%$ deviations on the fit results will be investigated.

Annual event yield: 3k, after full trigger selection (for 2 fb^{-1}).

Tagging performance: tagging efficiency $\varepsilon_{\rm tag} = 66\%$, wrong-tag fraction $\omega_{\rm tag} = 31\%$ after trigger, corresponding to a tagging power of $\varepsilon_{\rm eff} = 9.6\%$.

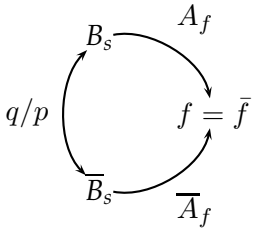
Background level: $B/S = B/S|_{b\overline{b}} + B/S|_{B_s \to D_s \pi\pi\pi} + \cdots = 0.4 + 0.2 = 0.6$ as nominal background level, using the central values from the $b\overline{b}$ and $B_s \to D_s \pi\pi\pi$ contributions. Note that the latter is added to the $b\overline{b}$ contribution to be conservative. The 90% confidence interval from $b\overline{b}$ is $B/S \in [0.10, 1.17]_{b\overline{b}}$. In addition to using the nominal estimate, the importance of the B/S in the sensitivity to ϕ_s will also be investigated using the upper 90% limit, i.e. 1.2, hence a B/S twice larger compared to the central value.

Acceptance parameterization: using (5.14), the acceptance parameter is $s_{\rm low} = 1.25\,{\rm ps}^{-1}$ after selection and trigger.

The remaining point to elucidate given the above $B_s \to \eta_c \phi$ performance is to what extend this decay can contribute to the determination of the B_s–$\overline{B}_s$ mixing phase. This is presented in the next chapter.

Chapter 6

Sensitivity to B_s–$\overline{B}_s$ Mixing Parameters



We give an overview of the sensitivity to the B_s–$\overline{B}_s$ mixing parameters at LHCb, using decays proceeding through $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions. The performance is assessed by means of toy Monte Carlo simulations, using as inputs the results of the full Monte Carlo studies. Both the $\bar{b} \to \bar{c}c\bar{s}$ decays to pure and to an admixture of CP eigenstates are considered, and used to probe the B_s–$\overline{B}_s$ mixing phase through a time-dependent mixing-induced CP measurement.

We describe in this chapter the toy (i.e. fast) Monte Carlo (MC) simulation used to determine the performance of the LHCb experiment to the B_s–$\overline{B}_s$ mixing observables. First a comparison between the different channels used is discussed in Section 6.1, considering the performance obtained from the full MC simulation for the $B_s \to \eta_c\phi$, $B_s \to D_sD_s$, $B_s \to J/\psi\eta(\gamma\gamma)$, $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0)$, and $B_s \to J/\psi\phi$ signal decay channels, as well as the control channel used, namely the flavor-specific $B_s \to D_s\pi$ decay channel. The method for the determination of the physical observables is outlined in Section 6.2, and the physics models and corresponding likelihoods are presented in Section 6.3. Finally, the adopted strategy to extract the relevant observables and the corresponding results are detailed in Section 6.4, with an outlook and suggestions for future improvements in Section 6.5.

The theoretical support for this channel was presented in Chapter 1. In particular, the necessary formulae for the description of the transition rates for $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions were derived in Section 1.4. All the results presented in this chapter are based on the outputs from the full Data Challenge 04 MC simulation, with updated yields and background estimates, and the improved tagging performance. This work is a continuation of the results of [103] for the $B_s \to J/\psi\phi$ decay channel and the former full MC studies. In particular, several aspects were added to the model, e.g. the acceptance and higher terms in ϕ_s are considered. For the latter, we fit for ϕ_s and not for the amplitude of $\sin\Delta M_s$, to be able to investigate larger values of ϕ_s with both the $\sin\phi_s$ and $\cos\phi_s$ contributions to the decay rates. Yet another change is the distinction between the proper

time uncertainties assigned to the events for the signal and control samples. In addition, the study is extended to the decays to pure CP eigenstates in order to evaluate their contribution in the determination of the B_s–$\overline{B}_s$ mixing phase at LHCb.

6.1 Experimental Aspects of the B_s Decays of Interest

The access to the B_s–$\overline{B}_s$ mixing parameters requires an excellent proper time resolution in order to resolve the fast B_s–$\overline{B}_s$ oscillations. As illustrated in Section 5.8.3, the proper time τ depends on the reconstructed parameters involved in $\tau = m\vec{p}\cdot(\vec{S}-\vec{P})/|\vec{p}|^2$: the mass m and the three-momentum of the B_s meson $\vec{p}$, and its origin $\vec{P}$ and decay $\vec{S}$ positions.

The determination of the B_s vertex is the crucial measurement for a good proper time resolution. Naively, the largest multiplicity decays will have the best τ resolution. This is true as long as we do not have long-lived B_s daughters, such as the D_s meson. Indeed, in the latter case, and for instance for $B_s \to D_s\pi$, the B_s vertex is actually determined from two particles. Moreover, the quality of the direction of the D_s momentum is worse than that of a non-composite final state particle. Another key parameter is the momentum resolution. For decays involving only charged tracks the contribution to the proper time resolution from the momentum is equivalent to that of the primary vertex. For decays involving photons or π^0's, the momentum resolution spoils the proper time determination in a significant way. This is dealt with by using a Kalman filter to refine the whole decay chain.

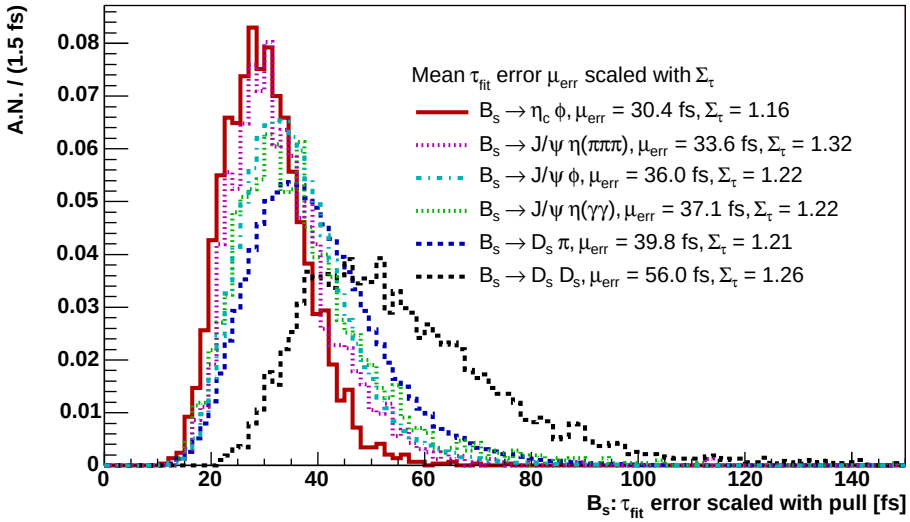


Figure 6.1: Distribution of the proper time errors [fs] for the different B_s channels, as obtained from the full MC. The errors are *scaled* with the corresponding Σ_τ. The normalization is arbitrary.

The proper time performance is very different for the channels considered in this chapter. As it will be explained later, the toy MC studies will use the proper time errors $\tau_{\rm fit}^{\rm err}$ obtained from the full MC, together with the scale factor Σ_τ corresponding to the standard deviation of the Gaussian fit of the proper time pull distribution (i.e. proper

time residual divided by the proper time error). In order to take into account the underestimation of the errors as obtained from the lifetime fit, the scale factor is used to model the proper time resolution. The distributions of the per-event proper time errors $\tau_{\rm fit}^{\rm err}$ for the different channels are shown in Figure 6.1, after *scaling* with the corresponding Σ_τ. As expected, the best performance is for the $B_s \to \eta_c\phi$ decay. The worse performance is for $B_s \to D_sD_s$, which is to be considered as a 2-prong decay for the B_s vertex determination, even though it has six final states. Moreover, the large spread of the errors is enhanced for $B_s \to D_sD_s$ since the two D_s can fly different distances.

We give in Table 6.1 the inputs from the full DC04 MC simulation used in this chapter. Additional information can be found in Section 5.9, and in Appendix A.

Table 6.1: Inputs from the full MC simulation for the B_s signal samples and the $B_s \to D_s\pi$ control sample.

Parameters	$J/\psi\phi$	$\eta_c\phi$	D_sD_s	$J/\psi\eta(\gamma\gamma)$	$J/\psi\eta(\pi\pi\pi)$	$D_s\pi$
2 fb^{-1} yield [k events]	131	3	4	8.5	3	120
Background level B/S	0.12	0.6	0.3	2.0	3.0	0.4
Mass σ_{B_s} [MeV/c]	14	12	6	34	20	14
Acceptance $s_{\rm low}$ [ps^{-1}]	2.81	1.25	1.6	1.86	1.54	1.36
Mean $< \tau_{\rm fit}^{\rm err} >$ [fs]	29.5	26.2	44.4	30.4	25.5	32.9
Scale factor Σ_τ	1.22	1.16	1.26	1.22	1.32	1.21
Wrong tag $\omega_{\rm tag}$ [%]	33	31	34	35	30	31
Tagging $\varepsilon_{\rm tag}$ [%]	57	66	57	63	62	63

The statistics available also play an important role in the determination of the B_s–$\overline{B}_s$ mixing phase ϕ_s. The $B_s \to J/\psi\phi$ decay channel has an order of magnitude larger expected event yield at LHCb compared to all the decays to pure CP eigenstates considered ($B_s \to \eta_c\phi$, $B_s \to D_sD_s$, $B_s \to J/\psi\eta(\gamma\gamma)$ and $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0)$). However, $B_s \to J/\psi\phi$ decaying to an admixture of CP eigenstates, we need to perform an angular analysis in order to disentangle them. This is not the case for the pure CP eigenstates decay channels, where the smaller statistics could be compensated by the absence of the need for an angular analysis.

The background levels represent either a central value or a 90% confidence level (upper limit), and are mainly obtained from the inclusive $b\overline{b}$ sample. The values for $B_s \to \eta_c\phi$ and $B_s \to D_s\pi$ include additional specific sources. For $B_s \to J/\psi\eta(\gamma\gamma)$ and $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0)$, the prompt J/ψ background has been studied. For these channels, the background levels are much higher due to the presence of photons or π^0's, which also results in a worse B_s mass resolution.

The effect of the selection cuts on the true proper time is modelled by the acceptance function defined in (5.14), and characterized by the parameter $s_{\rm low}$. This parameter represents the slope of the raise of the proper time efficiency at small proper times. As we can see from Table 6.1, only the $B_s \to J/\psi\phi$ selection can afford to have an almost lifetime unbiased selection. Note that in the yield of $B_s \to J/\psi\phi$ are included all the triggered events, i.e. not necessarily the events only triggered by the lifetime unbiased streams.

Supposedly the first measurement of the mixing phase ϕ_s will be performed using lifetime unbiased events and the $B_s \to J/\psi\phi$ sample.

The largest sensitivity to ϕ_s, for a Standard Model value, will come from tagged events. Since the wrong-tag fraction and ϕ_s modulate the B_s–$\overline{B}_s$ oscillations, a control sample is necessary to determine the wrong-tag fraction ω_{tag}. The same mistag rate and tagging efficiency are assumed for both the signal and the control sample in the toy MC presented in this chapter, taking the performance of the signal as input. In the real experiment, we will need to proceed to a reweighting of the $B_s \to D_s\pi$ phase space and to use different tagging categories in order to correct for the different tagging performance. This will introduce systematic errors in our measurements.

The primary purpose of this chapter is to determine LHCb sensitivity to ϕ_s using benchmark channels, and to compare the results for $B_s \to J/\psi\phi$ and for the pure CP eigenstates. The background shapes are assumed to be the same for each channel, without trying to have a dedicated model for each sample. The refinements of the background description will have to be adjusted with the real data. We could also add several other channels, such as $B_s \to J/\psi\eta'$, though the contribution is negligible as the performance is expected to be at the level of $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0)$, or worse. We also included the color-allowed $B_s \to D_sD_s$ channel, which could be affected by final state interactions, as explained in Section 1.4.1. Nevertheless, from an experimental point of view the validity of this assumption must be checked by performing the CP measurement. The study of different channels probing ϕ_s will enable to cross-check the results obtained with $B_s \to J/\psi\phi$. Finally, a promising contribution to the ϕ_s mixing phase will come from $B_s \to J/\psi\phi$ decays, with $J/\psi \to e^+e^-$. These decays have not been included, still they would add $\sim 1/6$ of statistics for a ~ 2 times larger background level, when compared to $B_s \to J/\psi\phi$, with $J/\psi \to \mu^+\mu^-$. Thus, their contribution will be at the level of the best pure CP eigenstates decay channels.

6.2 Toy Monte Carlo Simulation

The sensitivity of LHCb to the B_s–$\overline{B}_s$ mixing parameters is determined by performing toy Monte Carlo simulations. These toys are fast parameterized simulations using the results of the full MC simulation as inputs. The generation and the fitting of the events are based on the RooFit toolkit for data modelling [104], integrated with the object-oriented data analysis framework of Root [105].[1] The fitting part is performed by MINUIT [106], the CERN physics analysis tool for function minimization.

The toy MC generates events for each sample according to the corresponding expected event yield and the background level, as given in Table 6.1. Each event is generated by randomizing the following observables: the mass m, the proper time t, the error on t, and the cosine of the transversity angle $\cos\theta$ in the case of $B_s \to J/\psi\phi$. The description of the probability density functions (pdf's) used for the generation and the fitting will be given in Section 6.3. They include the background contribution, as well as the tagging and wrong-tag efficiencies. Moreover, the pdf's include the experimental resolution effects, the proper time acceptance, and each signal event has a different per-event proper time error.

The total (signal and background) likelihood function used to both generate and fit

[1] We used a modified RooFit v2.03 version to include non-existing features, together with Root 4.02.

the events is given by the product over all events i:

$$\mathcal{L}^{\bar{\mathrm{b}}\to\bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}} = \frac{(N_{\mathrm{sig}}+N_{\mathrm{bkg}})^{N_{\mathrm{obs}}}}{N_{\mathrm{obs}}!}e^{-(N_{\mathrm{sig}}+N_{\mathrm{bkg}})} \times \prod_{i\in \mathrm{B_s}\to f}^{N_{\mathrm{obs}}}\left[f^{\mathrm{sig}}\mathcal{L}_m^{\mathrm{sig}}(m_i)\mathcal{L}_i^{\mathrm{sig}}(t_i^{\mathrm{rec}},\tau_i^{\mathrm{err}},q_i(,\theta_i)) + f^{\mathrm{bkg}}\mathcal{L}_m^{\mathrm{bkg}}(m_i)\mathcal{L}_i^{\mathrm{bkg}}(t_i^{\mathrm{rec}}(,\theta_i))\right],$$

where N_{obs} is the actual number of observed events in the sample (Poisson distributed) such that the number of expected events by the model is $N_{\mathrm{exp}} = N_{\mathrm{sig}} + N_{\mathrm{bkg}}$. The signal and background probabilities are given by $f^{\mathrm{sig}} = N_{\mathrm{sig}}/(N_{\mathrm{sig}} + N_{\mathrm{bkg}})$ and $f^{\mathrm{bkg}} = N_{\mathrm{bkg}}/(N_{\mathrm{sig}} + N_{\mathrm{bkg}})$, respectively. The mass pdf's for the signal and the background are respectively denoted by $\mathcal{L}_m^{\mathrm{sig}}(m_i)$ and $\mathcal{L}_m^{\mathrm{bkg}}(m_i)$, and are based on the reconstructed mass m_i. The signal pdf $\mathcal{L}_i^{\mathrm{sig}}$ describes the decay rates for the proper time t_i^{rec}. It depends on the per-event error τ_i^{err} in order to assign an uncertainty on the reconstructed proper time. The effect of the tagging is included in q_i ($q_i = +1$ if the signal meson is tagged as a $\mathrm{B_s}$, $q_i = -1$ if it is tagged as a $\overline{\mathrm{B}}_{\mathrm{s}}$, and $q_i = 0$ if the event is untagged). For $\mathrm{B_s} \to \mathrm{J}/\psi\phi$, the term $\mathcal{L}_i^{\mathrm{sig}}$ depends in addition on the transversity angle θ_i. The background pdf $\mathcal{L}_i^{\mathrm{bkg}}$ does not depend on τ_i^{err} since the per-event error does not represent an accurate estimate of the proper time resolution of the combinatorial background events, which do not have any defined true proper time. We assume the same tagging performance for the signal and the background. However, the background decay rates do not depend on the tagging result.

The extraction of the different parameters is obtained by performing a likelihood fit to the mass and the proper time distributions of the signal $\bar{\mathrm{b}} \to \bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}$ quark-level transitions, and to the transversity angle for $\mathrm{B_s} \to \mathrm{J}/\psi\phi$. The $\bar{\mathrm{b}} \to \bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}$ likelihood is simultaneously maximized with a similar likelihood for the control sample, i.e. $\mathrm{B_s} \to \mathrm{D_s}\pi$, by minimizing the total negative log-likelihood function ($-\ln\mathcal{L}^{\mathrm{tot}} = -\ln\mathcal{L}^{\bar{\mathrm{b}}\to\bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}} - \ln\mathcal{L}^{\mathrm{control}}$). In particular, the wrong-tag fraction and the tagging efficiency are assumed to be identical for the signal and the control samples. Thus, the statistical uncertainty due to the control sample is included in the fit results. As previously mentioned, the biases introduced by the trigger and the selection lead to different tagging performances between the signal and the control samples, that will add systematic uncertainties. This is ignored in our study, and should be corrected for when performing the analysis with real data.

We generate ~ 225 experiments for each signal sample with a nominal set of parameters. One experiment corresponds to 10^7 s (one year) data taking at the nominal luminosity $\mathcal{L} = 2\times10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, i.e. with the expected statistics for the signal and control samples with $2\ \mathrm{fb}^{-1}$. Additionally, for $\mathrm{B_s} \to \mathrm{J}/\psi\phi$ and $\mathrm{B_s} \to \eta_{\mathrm{c}}\phi$ we will scan several values of the parameter space to study the dependence on the parameterization. The likelihood fit is performed for each of the experiments, and the corresponding output values of the fitted variables and their errors are retrieved.

For the $\mathrm{B_s} \to \mathrm{J}/\psi\phi$ generation and fitting, each experiment takes about 2 normalized CPU hours, on the CERN batch facilities consisting of a huge farm equipped with dual-processor machines. The normalized CPU time for a single experiment is about 30 minutes for the pure CP eigenstate channels. Note that the fitting strategy is crucial, see Section 6.4, due to the complexity of the likelihood function. It also important for the timing optimization to clearly separate as much as possible each component, trying to factorize out each dependence on the randomized observables. A particularly important example is that of the acceptance function which must weight the decay rates *before* mul-

tiplying the pdf's with the angular parts. In this way we can decouple the angular and proper time parts of the likelihood for each of the CP components in $B_s \to J/\psi\phi$. Note that this is no longer possible when performing a full angular analysis, unless a different approach is used, such as the method of angular moments [41].

The likelihood fit converges in 100% of the experiments. However, for some experiments the covariance matrix is not positive-definite, which results in correlation coefficients close to unity. These experiments are ignored as they clearly do not yield meaningful results. For the nominal settings, this happens in $\sim 2\%$ of the experiments for the $B_s \to J/\psi\phi$ simulation, and basically never for pure CP eigenstates experiments. We observe an increase of the experiments with an ill-defined covariance matrix only for values of ϕ_s close to $-\pi/2$, that is when the $\cos\phi_s$ term of the signal decay rates vanishes. In this case the fraction of events discarded is $\sim 10\%$ for both the $B_s \to J/\psi\phi$ and the pure CP eigenstates simulations. In this extreme case, the fit does not produce reliable results, and we therefore ignore the $\phi_s = -\pi/2$ value from our scanned parameters.

In our toy studies, each experiment is repeated a large number of times and it is generated independently with identically distributed parameters all having finite variances. Thus, the average of the fit outputs for each individual variable should converge to the mean value of the fitted parameter. For a given parameter, the spread of the distribution of the central value represents the statistical error on the parameter. In this chapter, we define the sensitivity to a parameter as the root-mean-square (rms) of the distribution of the fit outputs for this parameter.

In order to test the performance of the fit, we will compare the rms of the fitted variable to the average of the errors returned by the fit. The latter represents the mean statistical error on the fitted parameter, and should be compatible with the rms of the fit outputs for properly estimated errors. Moreover, the pull distribution of the parameter should tend to a Normal Gaussian distribution for correctly estimated errors, assuming Gaussian errors. We will in particular show that the average of the fitted errors scaled with the width of the corresponding pull distribution are in general compatible with the rms of the fitted value. We can thus in this way account for the under/overestimation of the errors for the different parameters, and look for biases in the fitted values. With real data, we will have to adjust our fit errors to what is observed in the Monte Carlo.

6.3 Likelihood Modelling

The likelihood describing each sample can be decomposed in two parts with distinctive physics models, namely the background contribution and the $\bar{b} \to \bar{c}c\bar{s}$ signal or the $B_s \to D_s\pi$ control contributions. The total likelihood of a given sample depends on the following observables:

B_s mass The mass pdf's tell us if an event is signal or background. We distinguish three regions in the mass spectrum:

- The total mass region defines the mass range in which all events are generated. The window corresponds to $M_{B_s} \pm m_{\text{total}}$, where $M_{B_s} = 5369.6\,\text{MeV/c}^2$ is the B_s mass used in the generation, and where m_{total} is $150\,\text{MeV/c}^2$ or $250\,\text{MeV/c}^2$. The latter value is for the $B_s \to J/\psi\eta$ decays, as they have a worse mass resolution.
- The signal window is defined as $|m_i - M_{B_s}| < 50\,\text{MeV/c}^2$, or as $|m_i - M_{B_s}| <$

100 MeV/c^2 for $B_s \to J/\psi\eta$. This region is used to extract the physics observables, and the background level in this region is given by the B/S.[2]

- The sideband region is defined as $|m_i - M_{B_s}| > 75$ MeV/c^2, or as $|m_i - M_{B_s}| > 150$ MeV/c^2, for $B_s \to J/\psi\eta$. This region is used to fit the background properties, and the acceptance function. We thus assume that these parameters are independent of the reconstructed mass.

Proper time The analytical proper time decay rates for the control sample are those presented in Section 1.4.4, and the decay rates for a given CP eigenstate are given in Section 1.4.2. For $B_s \to J/\psi\phi$ we need to account for the angular distribution as discussed in Section 1.4.3, and the fraction of CP-odd eigenstates is described by the observable R_T defined in (1.112). For the signal and the control decay rates we include the effect of the wrong tag, and the signal events are assigned a proper time uncertainty depending on the per-event proper time errors obtained from the proper time fit in the full MC simulation. The resolution is then taken into account by smearing these errors with a Gaussian pull resolution function. The background decay rates are assumed to be exponentials and independent of the per-event proper time errors. All the decays rates are weighted by the parameterization of the acceptance function given in (5.14).

Transversity angle This variable is only defined for the $B_s \to J/\psi\phi$ sample. The angular distribution considered is the one-angle transversity distribution given in Section 1.4.3. The distribution is expressed as a function of $\cos\theta$.

In the following, we describe each of the terms involved in the total likelihood, and discuss the models used.

6.3.1 Description of the Mass Model

The signal mass pdf is described by a Gaussian (G) and the background pdf by an exponential (E), respectively corresponding to the following expressions:

$$\begin{aligned} \mathcal{L}_m^{\rm sig}(m_i; M_{B_s}, \sigma_{B_s}) &\propto G(m_i; M_{B_s}, \sigma_{B_s}) \,, \\ \mathcal{L}_m^{\rm bkg}(m_i; \kappa_{\rm bkg}) &\propto E(m_i; \kappa_{\rm bkg}) \,, \end{aligned} \tag{6.1}$$

where:

$$\begin{aligned} G(m_i; M_{B_s}, \sigma_{B_s}) &= \exp\left(-\frac{1}{2}\left(\frac{m_i - M_{B_s}}{\sigma_{B_s}}\right)^2\right) \,, \\ E(m_i; \kappa_{\rm bkg}) &= \exp(\kappa_{\rm bkg}\, m_i) \,. \end{aligned}$$

We have the following parameters entering the above pdf's:

- m_i is the reconstructed mass for the event i;
- $M_{B_s} = 5369.6$ MeV/c^2 is the nominal B_s mass;
- σ_{B_s} is the B_s mass resolution, given in Table 6.1;

[2] In the current implementation of the toy, only one signal window can be defined for both the signal and control samples; we thus scale linearly the B/S of the control sample by a factor two when considering the $B_s \to J/\psi\eta$ channels.

- $\kappa_{\rm bkg}$ describes the shape of the background "mass". In the real experiment, the background may have peaking backgrounds, thus a more complex model should be used. For simplicity, we take $\kappa_{\rm bkg} = -1.0\ ({\rm MeV/c})^{-1}$.

The projections of the likelihoods onto the mass distribution are shown for a random experiment in Figure 6.2. The values used for the generation are those of Table 6.1, with the relevant mass regions for the signal under study.

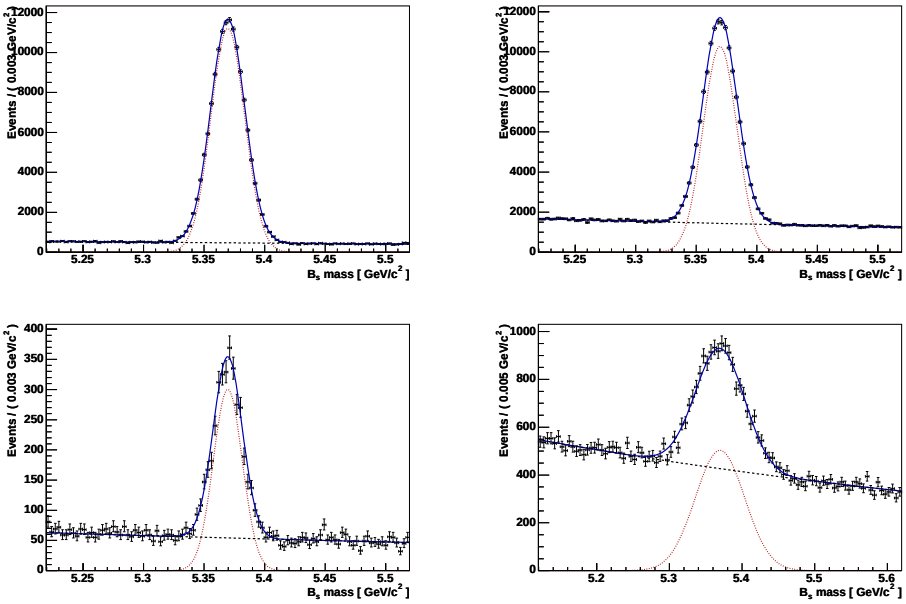


Figure 6.2: Projection of the likelihood onto the mass distribution [GeV/c^2] for the $B_s \to J/\psi\phi$ (top left), the $B_s \to D_s\pi$ (top right), the $B_s \to \eta_c\phi$ (bottom left), and the $B_s \to J/\psi\eta(\gamma\gamma)$ (bottom right) samples. The solid blue curve is the projection of all contributions, i.e. signal and background. The dotted red line is the signal contribution, whereas the dashed black line corresponds to the background.

The combined (signal and background) extended mass likelihood describes the actual number of observed events $N_{\rm obs}$ in the full mass region, and satisfies:

$$\mathcal{L}_m^{\rm comb} = \frac{e^{-\left(N_{\rm sig}+N_{\rm bkg}\right)}}{N_{\rm obs}!} \times \prod_{i\in {\rm B_s}\to f}^{N_{\rm obs}} \left[N_{\rm sig}\, \mathcal{L}_m^{\rm sig}(m_i; M_{\rm B_s}, \sigma_{\rm B_s}) + N_{\rm bkg}\, \mathcal{L}_m^{\rm bkg}(m_i; \kappa_{\rm bkg}) \right] .$$

Note that $N_{\rm obs}$ is a Poisson distributed variable with a mean corresponding to the number of expected events in the sample, to account for the variable size of the sample. The expected number of events by the model is such that the number of events in the signal window corresponds to the signal event yield $N_{\rm sig}^{\rm sig,reg}$, with a number of background events in this region given by $N_{\rm bkg}^{\rm sig,reg} = N_{\rm sig}^{\rm sig,reg} \times B/S$, where the B/S ratio and the signal yield are given in Table 6.1.

The total number of events ($N_{\rm sig} + N_{\rm bkg}$) expected by the model in each experiment, and for each sample ($\bar{\rm b} \to \bar{\rm c}{\rm c}\bar{\rm s}$ or $B_s \to D_s\pi$), in the full mass window are given by:

- $B_s \to J/\psi\phi$ generation: 178'363 events for the $B_s \to J/\psi\phi$ sample and 264'523 events for the $B_s \to D_s\pi$;

- $B_s \to \eta_c\phi$ generation: 8'418 events for the $B_s \to \eta_c\phi$ sample and 264'523 events for the $B_s \to D_s\pi$;
- $B_s \to D_sD_s$ generation: 7'612 events for the $B_s \to D_sD_s$ sample and 264'523 events for the $B_s \to D_s\pi$;
- $B_s \to J/\psi\eta(\gamma\gamma)$ generation: 51'400 events for the $B_s \to J/\psi\eta(\gamma\gamma)$ sample and 362'104 events for the $B_s \to D_s\pi$;
- $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0)$ generation: 25'697 events for the $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0)$ sample and 362'104 events for the $B_s \to D_s\pi$.

6.3.2 Description of the Angular Model

The one-angle transversity angular distribution of the CP eigenstates in the $B_s \to J/\psi\phi$ decay is modelled using (1.111). The transversity angle θ is defined in Section 1.4.3, and it is illustrated in Figure 6.3. As we show below, we can safely integrate over the other two transversity angles (φ and ψ), after checking the selection acceptance is uniform as a function of these angles.

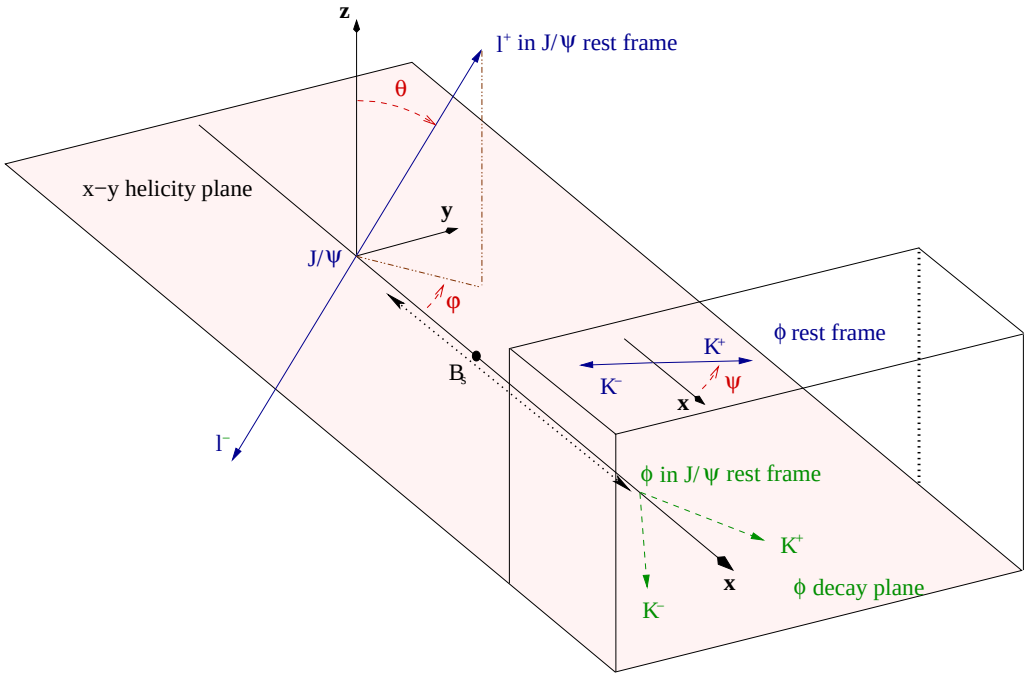


Figure 6.3: The definition of the transversity angle θ, and the illustration of the other two transversity angles φ and ψ.

The angular distributions of $\cos\theta$, φ and $\cos\psi$ in $B_s \to J/\psi\phi$ events, at the generator level (EvtGen), are shown in Figure 6.4. All CP components are included, with a CP-odd fraction of $\sim 10\%$. The distributions are fitted with $a_0(a_1 + x^2)$ for $\cos\theta$ and $\cos\psi$, and with $a_0 \cos x \sin x + a_1 \cos^2 x + a_2 \sin^2 x$ for φ.

The acceptance as a function of each the transversity angle after off-line selection, reconstruction and full trigger is shown in Figure 6.5 for events within 400 mrad, using a cheated off-line selection without any selection for the denominator. It has been checked

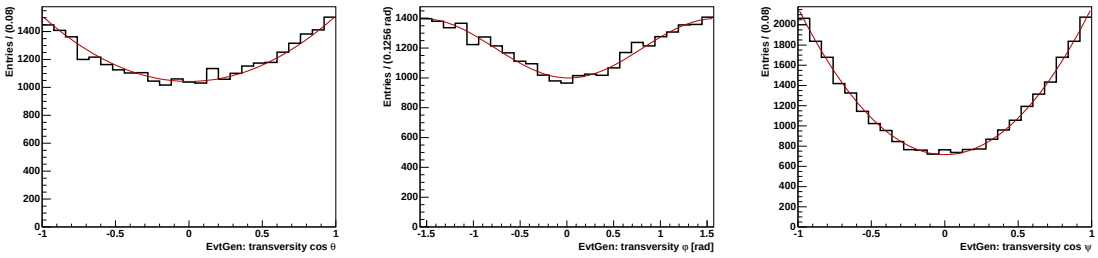


Figure 6.4: Transversity angles distributions in $B_s \to J/\psi\phi$ events at the generator level (EVTGEN). Left: $\cos\theta$. Middle: φ. Right: $\cos\psi$.

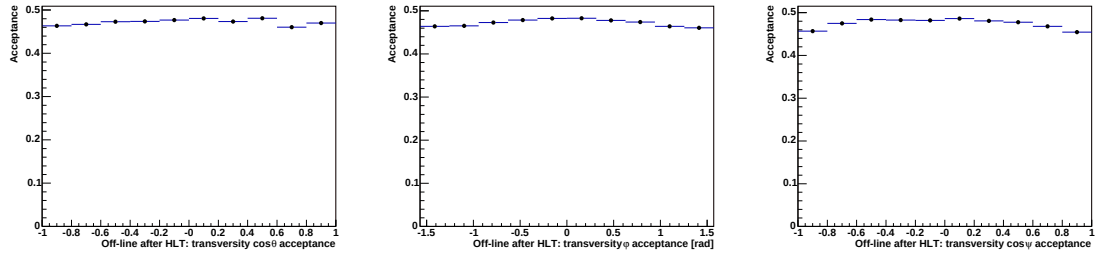


Figure 6.5: Acceptance as a function of the three transversity angles after full trigger, reconstruction and off-line selections in $B_s \to J/\psi\phi$ events. Left: $\cos\theta$. Middle: φ. Right: $\cos\psi$. The error bars correspond to binomial errors.

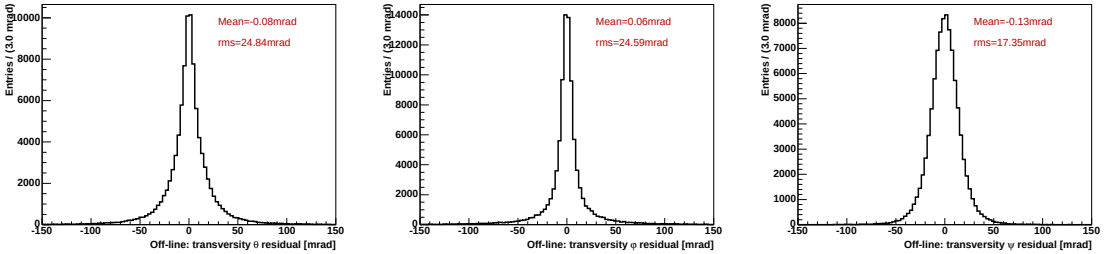


Figure 6.6: Transversity angles residuals after full trigger, reconstruction and off-line selections in $B_s \to J/\psi\phi$ events. Left: $\delta\theta$ [mrad]. Middle: $\delta\varphi$ [mrad]. Right: $\delta\psi$ [mrad].

that the 400 mrad generator-level cut has no effect on the acceptance of the transversity angles. We observe that the acceptance is approximatively flat for each of the angles, which justifies our initial assumption. The small non uniformities should in principle be accounted for in a real analysis, but we ignore them for this sensitivity study. For the angle θ, these will lead to a slight overestimation of R_{T}, and hence of the statistical uncertainty on ϕ_s.

The residuals of the transversity angles are shown in Figure 6.6. The corresponding rms values are respectively given by $\mathrm{rms}(\delta\theta) = 24.8$ mrad, $\mathrm{rms}(\delta\varphi) = 24.6$ mrad, and $\mathrm{rms}(\delta\psi) = 17.4$ mrad. The effect of the θ resolution is introduced in the generation of the events by smearing the θ distribution with a Normal Gaussian variable multiplied by a constant factor of 20 mrad. This resolution corresponds to the rms of the θ residual as

observed in the full MC simulation of a previous study [103]. However, the likelihoods used to fit the generated data do not include this resolution, as we expect the variations in the θ distributions to be small compared to this resolution.

The angular pdf's terms for the signal CP-even, the signal CP-odd, and the background components are respectively given by:

$$\begin{aligned} \mathcal{L}^{\rm sig}_{\theta,{\rm even}}(\theta_i) &\propto (1+\cos\theta_i{}^2)/2\,, \\ \mathcal{L}^{\rm sig}_{\theta,{\rm odd}}(\theta_i) &\propto (1-\cos\theta_i{}^2)\,, \\ \mathcal{L}^{\rm bkg}_{\theta}(\theta_i) &\propto (1+\alpha_{\rm bkg}\cos\theta_i{}^2)\,, \end{aligned} \tag{6.2}$$

where $\alpha_{\rm bkg}$ parameterizes the background angular component. We assume that the background is flat in $\cos^2\theta$, and we thus set $\alpha_{\rm bkg}=0$ as nominal value. The relative weight of the CP-even and CP-odd components is controlled by the fraction of CP odd, namely the observable $R_{\rm T}$. The projections of the likelihoods onto the $\cos\theta$ distribution are shown in Figure 6.7.

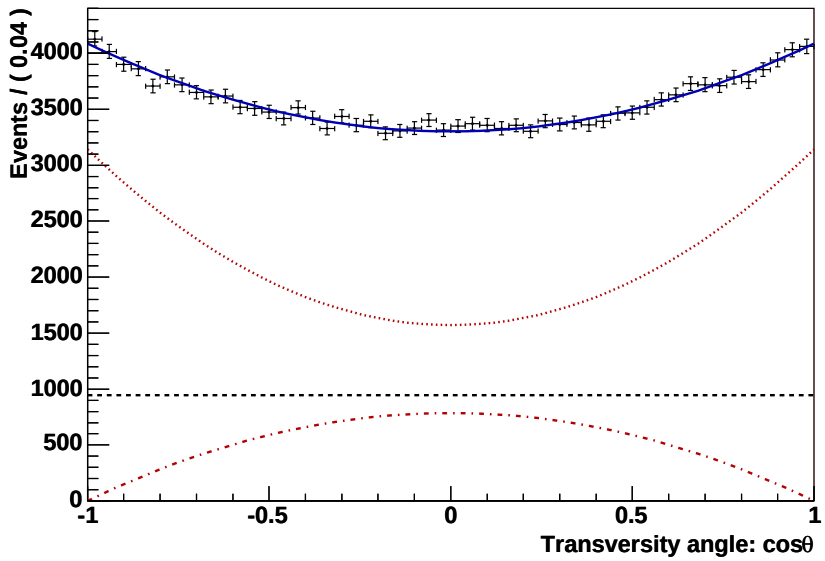


Figure 6.7: Projection of the likelihood onto the $\cos\theta$ distribution for the $B_s \to J/\psi\phi$ decay sample with $R_{\rm T}=0.2$, in the total mass window. The solid blue curve is the projection of all contributions, i.e. signal and background. The dotted red line is the CP-even signal contribution, the red dashed-dotted line is the CP-odd signal contribution, and the dashed black line corresponds to the background.

6.3.3 Modelling of the Time-Dependent Decay Rates

The last component of the total likelihood is the one describing the proper time. The physics models to be used for the $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions and for the control $B_s \to D_s\pi$ channel were discussed at length in Chapter 1, including the effect of the tagging. We now need to take into account the effect of the proper time acceptance $A(t_i^{\rm rec}; s_{\rm low})$ and the effect of the errors $\tau_i^{\rm err}$ in the likelihood description.

As argued in Chapter 1, we assume for the $\bar{\mathrm{b}} \to \bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}$ quark-level transitions that there is no production asymmetry for the $\mathrm{B_s}$ and $\overline{\mathrm{B}}_\mathrm{s}$ meson. The background is described as a simple exponential lifetime component. Respectively for the CP-even, the CP-odd, and the background contributions, we have that the proper time pdf's are given by:

$$\begin{aligned}
\mathcal{L}^{\mathrm{sig}}_{t,\mathrm{even}}(t_i^{\mathrm{rec}}, \tau_i^{\mathrm{err}}, q_i; s_{\mathrm{low}}, \Sigma_\tau, \mu_\tau, \omega_{\mathrm{tag}}, \vec{\alpha}) &\propto A(t_i^{\mathrm{rec}}; s_{\mathrm{low}}) \times R_{\mathrm{even}}\left(t_i^{\mathrm{true}}, q_i; \omega_{\mathrm{tag}}, \vec{\alpha}\right) \\
&\quad \otimes G(t_i^{\mathrm{rec}} - t_i^{\mathrm{true}}; \mu_\tau \tau_i^{\mathrm{err}}, \Sigma_\tau \tau_i^{\mathrm{err}}) \,, \\
\mathcal{L}^{\mathrm{sig}}_{t,\mathrm{odd}}(t_i^{\mathrm{rec}}, \tau_i^{\mathrm{err}}, q_i; s_{\mathrm{low}}, \Sigma_\tau, \mu_\tau, \omega_{\mathrm{tag}}, \vec{\alpha}) &\propto A(t_i^{\mathrm{rec}}; s_{\mathrm{low}}) \times R_{\mathrm{odd}}\left(t_i^{\mathrm{true}}, q_i; \omega_{\mathrm{tag}}, \vec{\alpha}\right) \\
&\quad \otimes G(t_i^{\mathrm{rec}} - t_i^{\mathrm{true}}; \mu_\tau \tau_i^{\mathrm{err}}, \Sigma_\tau \tau_i^{\mathrm{err}}) \,, \\
\mathcal{L}^{\mathrm{bkg}}_{t}(t_i^{\mathrm{rec}}; s_{\mathrm{low}}, \tau_{\mathrm{bkg}}) &\propto A(t_i^{\mathrm{rec}}; s_{\mathrm{low}}) \times E(t_i^{\mathrm{true}}; \tau_{\mathrm{bkg}}) \\
&\quad \otimes \delta(t_i^{\mathrm{rec}} - t_i^{\mathrm{true}}) \,. \qquad (6.3)
\end{aligned}$$

The signal decay rates for the $\bar{\mathrm{b}} \to \bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}$ transitions are given by the following expressions for the CP-even (R_{even}) and the CP-odd (R_{odd}) components:

$$\begin{aligned}
R_f\left(t_i^{\mathrm{true}}, q_i; \omega_{\mathrm{tag}}, \vec{\alpha}\right) \;\propto\; e^{-\Gamma_\mathrm{s} t_i^{\mathrm{true}}} \Bigg\{ & \cosh\frac{\Delta\Gamma_\mathrm{s}\, t_i^{\mathrm{true}}}{2} - \eta_f \cos\phi_\mathrm{s} \sinh\frac{\Delta\Gamma_\mathrm{s}\, t_i^{\mathrm{true}}}{2} \qquad (6.4) \\
& + \eta_f\, q_i\, D \sin\phi_\mathrm{s} \sin\left(\Delta M_\mathrm{s}\, t_i^{\mathrm{true}}\right) \Bigg\} \,,
\end{aligned}$$

where the final state can be $f \in \{\mathrm{even}\,, \mathrm{odd}\}$ with the eigenvalues $\eta_f = +1$ for the CP even and $\eta_f = -1$ for the CP odd. The tagging result is described by q_i, which takes the values of $q_i = +1$ if the signal meson is tagged as a $\mathrm{B_s}$ at production time, $q_i = -1$ if it is tagged as a $\overline{\mathrm{B}}_\mathrm{s}$, and $q_i = 0$ if the meson is untagged. As we can see from the decay rates, untagged events still give access to $\cos\phi_\mathrm{s}$ and $\Delta\Gamma_\mathrm{s}$. Yet another important fact is that for a small value of ϕ_s, basically all the sensitivity will come from the $\sin\phi_\mathrm{s}$ term, which is multiplied by the tagging dilution D. Given that both terms modulate the ΔM_s oscillations, we must therefore introduce a control sample, e.g. $\mathrm{B_s} \to \mathrm{D_s}\pi$, to help in the determination of ΔM_s and of the wrong tag ω_{tag}.

In the above proper time pdf's and decay rates we have introduced the following definitions:

- τ_i^{err} is the per-event proper time error.
- q_i is the tagging category: $+1$ for $\mathrm{B_s}$, -1 for $\overline{\mathrm{B}}_\mathrm{s}$, and $q_i = 0$ if there is no tag. We assume that the background decay rates are independent of the tagging result.
- $A(t_i^{\mathrm{rec}}; s_{\mathrm{low}})$ is the acceptance function depending on the reconstructed proper time t_i^{rec}, and s_{low} describes the acceptance parameterization. Note that in Chapter 5 the acceptance was given as a function of the *true* proper time. In the fast simulation of this chapter we express it as a function of the *reconstructed* proper time such that this acceptance can be factorized out of the true proper time convolution.
- G is the Gaussian resolution model, depending on the per-event proper time errors. The width of the resolution is scaled with the scale factor Σ_τ, which corresponds to the standard deviation of the pull as obtained in the full MC. The bias of the resolution function is assumed to be $\mu_\tau = 0$.
- E is the exponential model for the background proper time, characterized by the parameter τ_{bkg}, which we chose to be $\tau_{\mathrm{bkg}} = 1.0\ \mathrm{ps}^{-1}$ for simplicity.

- $D = (1 - 2\omega_{\text{tag}})$ is the tagging dilution factor, given in terms of the wrong-tag fraction ω_{tag}.

- $\vec{\alpha} = (\Delta\Gamma_s, \Gamma_s, \phi_s, \Delta M_s)$ is the vector of physics parameters describing the B_s–$\overline{B}_s$ mixing. The average decay width $\Gamma_s = 1/\tau_s$ is given by $\Gamma_s = (\Gamma_H + \Gamma_L)/2$. The mass difference is $\Delta M_s = M_H - M_L$, and the decay width difference is $\Delta\Gamma_s = \Gamma_L - \Gamma_H$.

Note that the convolution signs $\otimes$ in the likelihood expressions mean that we integrate over the proper time used for the generation t_i^{true}, from zero to infinity.

We will now illustrate the effect of the acceptance and of the proper time resolution on the signal $\bar{b} \to \bar{c}c\bar{s}$ decay rates. As an example, the plots in Figure 6.8 show the effect on the analytical decay rates for the $\bar{b} \to \bar{c}c\bar{s}$ decays to pure CP-even eigenstates when we successively add a wrong-tag fraction ($\omega_{\text{tag}} = 30\%$), a constant Gaussian resolution (35fs), and an acceptance function ($s_{\text{low}} = 1.3\text{ps}^{-1}$). The parameters used in the figure are $\phi_s = -0.2$ rad, $\Delta M_s = 17.5$ ps^{-1}, $1/\Gamma_s = 1.45$ ps, and $\Delta\Gamma_s/\Gamma_s = 15\%$. Note that the value of $\phi_s = -0.2$ rad used in the plots is five times the SM expectation, in order to accentuate the wiggles in this example. As we can see from these plots, the fast oscillations and their reduced amplitudes make of the determination of ϕ_s a challenging measurement.

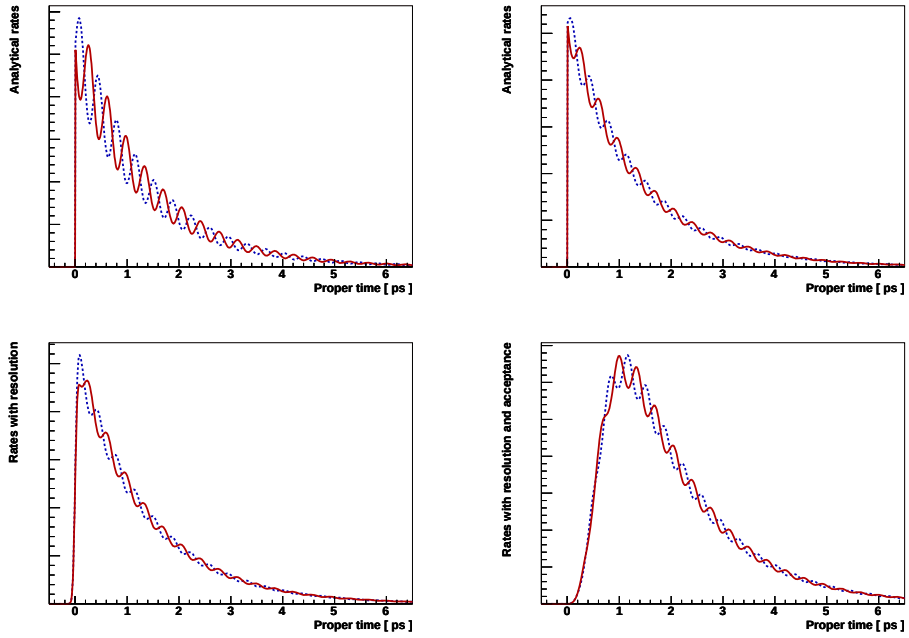


Figure 6.8: Signal decay rates [ps] of a $\bar{b} \to \bar{c}c\bar{s}$ transition to pure CP-even eigenstates. The red solid line is for an initially tagged B_s, whereas the blue dashed line is for an initially tagged $\overline{B}_s$. The top left plot shows the analytical decay rates, the top right one shows the effect of the wrong-tag, the bottom left plot shows the effect of a constant proper time resolution, and the bottom right one shows the combination of all the effects together with an acceptance function.

We already know that the tagging will dilute the oscillations. A similar effect is ascribed to the proper time resolution, where the amplitude of the wiggles will be attenuated by a non-accurate proper time measurement. The acceptance function has the effect

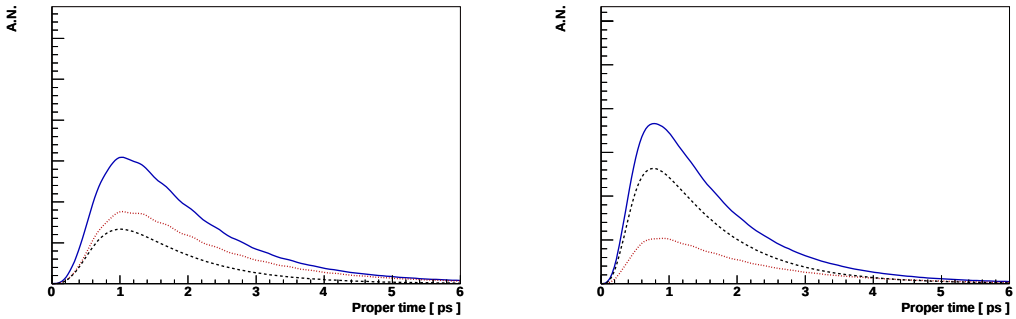


Figure 6.9: Projection of the likelihood onto the proper time distribution [ps], and for events initially tagged as B_s, in the signal region. The mixing parameters are the nominal ones. The left plot is for the $B_s \to \eta_c\phi$ sample, and the right one for the $B_s \to J/\psi\eta(\gamma\gamma)$ sample. The solid blue curve is the projection of all contributions, i.e. signal and background. The dotted red line is the CP-even signal contribution, and the dashed black line corresponds to the background.

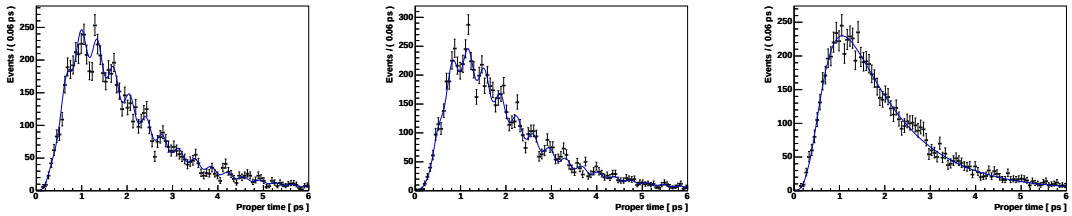


Figure 6.10: Projection of the likelihood onto the proper time distribution [ps], with $10\ \text{fb}^{-1}$, for $B_s \to \eta_c\phi$, in the signal region. The mixing parameters are the nominal ones, with a NP $\phi_s = -\pi/4$ rad. The left plot is for events initially tagged as B_s, the middle one for events initially tagged as $\overline{B}_s$, and the right plot is for untagged events. The solid blue curve is the projection of all contributions, i.e. signal and background.

of removing events at low proper times, due to the selection and trigger cuts. This has the important consequence of making the determination of the proper time resolution from the negative tail of the signal proper time distribution impossible. Thus, the proper time resolution has to be extracted from somewhere else. This can be done from lifetime unbiased di-muons, or from the ΔM_s oscillations by exploiting the different uncertainties assigned to the signal events. In our study, the determination of the proper time scale factor from the likelihood fit is not possible due to the acceptance function, combined with the fact that we need to measure the amplitude of the diluted oscillations.

In order to further demonstrate the difficulty of the ϕ_s measurement, we can consider the $B_s \to \eta_c\phi$ and the $B_s \to J/\psi\eta(\gamma\gamma)$ channels. The projection of the CP-even likelihood $\mathcal{L}_{t,\text{even}}$ and of the background one $\mathcal{L}_t^{\text{bkg}}$ onto the proper time distribution are shown in Figure 6.9. The plots are for events initially tagged as B_s, and in the signal region. The input values are those of Table 6.1. The mixing parameters used are $\phi_s = -0.04$ rad, $\Delta M_s = 17.5\ \text{ps}^{-1}$, $1/\Gamma_s = 1.45$ ps, and $\Delta\Gamma_s/\Gamma_s = 15\%$, defining our nominal choice of parameters.

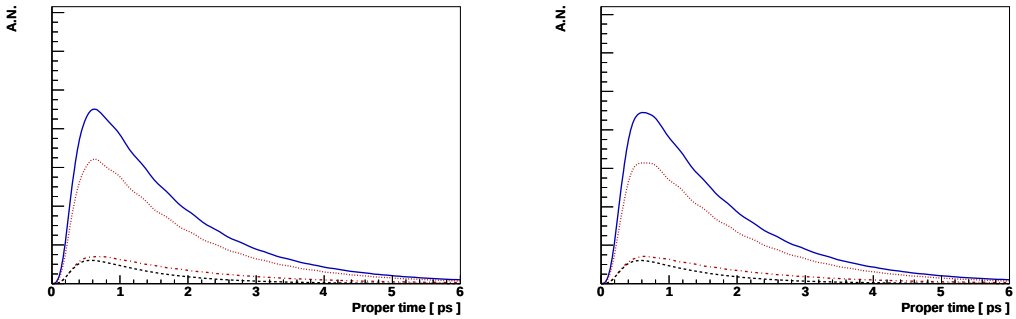


Figure 6.11: Projection of the likelihood onto the proper time distribution [ps], for $B_s \to J/\psi\phi$, in the signal region. The mixing parameters are the nominal ones. The left plot is for events initially tagged as B_s, and the left plot is for events initially tagged as $\overline{B}_s$. The solid blue curve is the projection of all contributions, i.e. signal and background. The dotted red line is the CP-even signal contribution, the dashed-dotted red line is the CP-odd signal contribution, and the dashed black line corresponds to the background.

There are two main observations to be noted from Figure 6.9. Firstly, we can see the effect of the worse proper time errors for the $B_s \to J/\psi\eta(\gamma\gamma)$, as we can barely distinguish the wiggles in the signal pdf projection. On the other hand, we clearly see the oscillations for $B_s \to \eta_c\phi$ in the signal projection. The second observation is the role of the background, which tends to flatten even more the wiggles as its contribution is added to the signal one. This can be seen from the projection of the total proper time likelihood for $B_s \to J/\psi\eta(\gamma\gamma)$, where the oscillations are invisible, whereas we can still observe them for $B_s \to \eta_c\phi$ in the combined background-signal likelihood projection. These two examples show that for decays to pure CP eigenstates the key parameters for the determination of ϕ_s are small proper time errors and an acceptable background level. Nevertheless, larger statistics and a good tagging performance are also important as they can compensate for a worse performance in the key parameters. We give in Figure 6.10 what would be the situation after five nominal years LHCb running ($10\,\text{fb}^{-1}$) for $B_s \to \eta_c\phi$, with the nominal mixing parameters, and with a significant New Physics contribution to the mixing phase $\phi_s = -\pi/4 \approx -0.79$ rad.

The last dilution of the wiggles happens when we have an admixture of CP eigenstates, as in $B_s \to J/\psi\phi$. Fortunately, a total dilution cannot occur, even when the fractions of CP odd and CP even are equal, i.e. $R_T = 0.5$. This cannot happen because we still have very distinctive angular distributions, as we can see from (6.2). Even though we must perform an angular analysis for $B_s \to J/\psi\phi$, its large statistics enable to compensate this and provide the best accuracy for the ϕ_s measurement. Figure 6.11 shows the different contributions to the proper time likelihood, with the nominal parameters.

Decay rates of the control sample

The proper time likelihoods for the flavor-specific control sample $B_s \to D_s^-\pi^+$ are similar to those of (6.3), but with the following decay rates:

$$R_f\left(t_i^{\text{true}}, r_i; \omega_{\text{tag}}, \vec{\beta}\right) \;\propto\; \frac{e^{-\Gamma_s t_i^{\text{true}}}}{2}\left\{\cosh\frac{\Delta\Gamma_s t_i^{\text{true}}}{2} + r_i D\cos\left(\Delta M_s t_i^{\text{true}}\right)\right\}, \qquad (6.5)$$

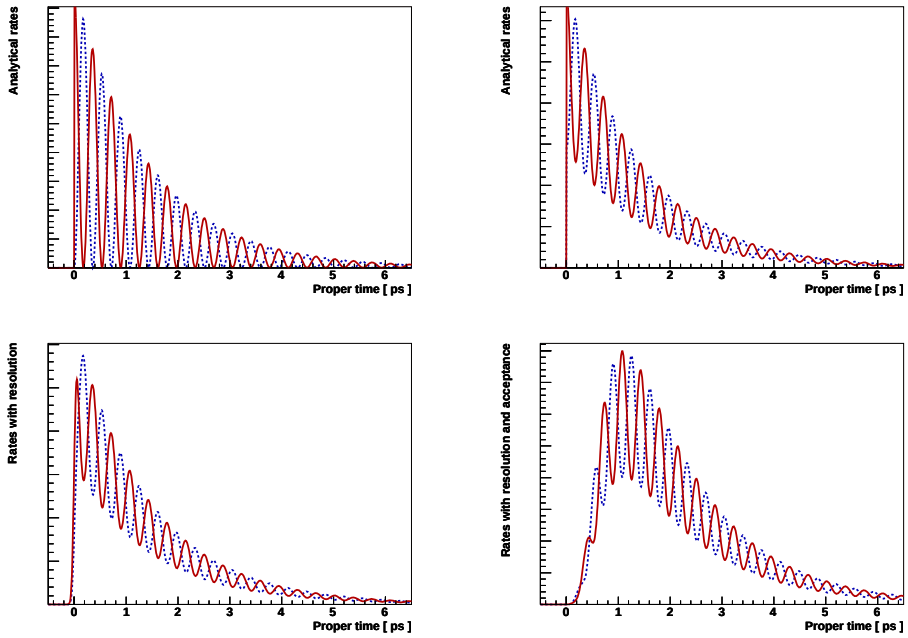


Figure 6.12: Signal decay rates [ps] for $B_s \to D_s\pi$. The red solid line is for an initially tagged B_s, whereas the blue dashed line is for an initially tagged $\overline{B}_s$. The top left plot shows the analytical decay rates, the top right one shows the effect of the wrong tag, the bottom left plot shows the effect of a constant proper time resolution, and the bottom right one shows the combination of all the effects together with an acceptance function.

where the results are:

- $r_i = +1$ for a B_s candidate tagged as unmixed (i.e. tagged as having the same flavor at production and decay);
- $r_i = -1$ for a B_s tagged as mixed (i.e. tagged as having different flavors at production and decay);
- $r_i = 0$ for an untagged B_s candidate.

The vector of physics parameters is given by $\vec{\beta} = (\Delta\Gamma_s, \Gamma_s, \Delta M_s)$. Note that there is only one term modulating the oscillations, which depends on the wrong tag ω_{tag}, such that the amplitude of the oscillations is much larger compared to those of the $\bar{b} \to \bar{c}c\bar{s}$ transitions. The effect of successively adding a wrong tag, a proper time resolution, and an acceptance function on the $B_s \to D_s\pi$ oscillations can be seen in Figure 6.12, where we used the same settings as in Figure 6.8.

6.3.4 Final Likelihood Functions

The final likelihood for the $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions used for the generation and to fit the events is given by:

$$\mathcal{L}^{\bar{b}\to\bar{c}c\bar{s}} = \frac{e^{-\left(N_{\text{sig}}+N_{\text{bkg}}\right)}}{N_{\text{obs}}!} \prod_{i \epsilon B_s \to f}^{N_{\text{obs}}} \mathcal{L}_i^{\bar{b}\to\bar{c}c\bar{s}}(m_i, \theta_i, t_i^{\text{rec}}, \tau_i^{\text{err}}, q_i)\,,$$

with:

$$\begin{aligned}\mathcal{L}_i^{\bar{\mathrm{b}}\to\bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}}(m_i,\theta_i,t_i,\tau_i^{\mathrm{err}},q_i) &= N_{\mathrm{sig}}\,\mathcal{L}_m^{\mathrm{sig}}(m_i)\Big[R_{\mathrm{T}}\mathcal{L}_{\theta,\mathrm{odd}}^{\mathrm{sig}}(\theta_i)\mathcal{L}_{t,\mathrm{odd}}^{\mathrm{sig}}(t_i^{\mathrm{rec}},\tau_i^{\mathrm{err}},q_i)\\ &\qquad + (1-R_{\mathrm{T}})\mathcal{L}_{\theta,\mathrm{even}}^{\mathrm{sig}}(\theta_i)\mathcal{L}_{t,\mathrm{even}}^{\mathrm{sig}}(t_i^{\mathrm{rec}},\tau_i^{\mathrm{err}},q_i)\Big]\\ &\quad + N_{\mathrm{bkg}}\,\mathcal{L}_m^{\mathrm{bkg}}(m_i)\mathcal{L}_\theta^{\mathrm{bkg}}(\theta_i)\mathcal{L}_t^{\mathrm{bkg}}(t_i^{\mathrm{rec}})\,. \end{aligned} \tag{6.6}$$

The different likelihood terms are given in (6.1) for the mass components, in (6.2) for the angular components, and in (6.3) for the proper time components. For the decays to pure CP-even eigenstates, we set the CP-odd fraction $R_{\mathrm{T}} = 0$, and we ignore the angular terms.

The $\bar{\mathrm{b}} \to \bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}$ likelihood is multiplied by the corresponding final likelihood for the $\mathrm{B_s} \to \mathrm{D_s}\pi$ control sample, and the total likelihood is thus simultaneously maximized. The final likelihood for the $\mathrm{B_s} \to \mathrm{D_s}\pi$ sample is similar to that of the $\bar{\mathrm{b}} \to \bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}$ sample, with the decay rates given in (6.5). In the generation, we assume that half of the $\mathrm{B_s} \to \mathrm{D_s}\pi$ events decay through $\mathrm{B_s} \to \mathrm{D_s^-}\pi^+$, and the other half through $\overline{\mathrm{B}}_{\mathrm{s}} \to \mathrm{D_s^+}\pi^-$. Moreover, the same tagging efficiency is assumed for the $\bar{\mathrm{b}} \to \bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}$ transitions and the control sample, but with different proper time errors.

6.4 Extracted Parameters

We report in this section the results of the likelihood fits. First the fit strategy is presented, then the results for the nominal parameters are shown. Finally, we also present the results for the different sets of parameters.

The inputs from the full MC were already given in Table 6.1. The nominal set of physics parameters are chosen to be compatible with the results of Chapter 1, and in particular with Section 1.5. We adopted the following nominal choice of parameters, to be considered as SM parameters:

- $M_{\mathrm{B_s}} = 5369.6\ \mathrm{MeV/c^2}$;
- $\Delta M_{\mathrm{s}} = 17.5\ \mathrm{ps^{-1}}$;
- $\phi_{\mathrm{s}} = -0.04$ rad;
- $\Delta\Gamma_{\mathrm{s}}/\Gamma_{\mathrm{s}} = 0.15$;
- $\tau_{\mathrm{s}} = 1/\Gamma_{\mathrm{s}} = 1.45$ ps;
- $R_{\mathrm{T}} = 0.2$, for $\mathrm{B_s} \to \mathrm{J}/\psi\phi$.

6.4.1 Fit Strategy

The likelihood fit is rather complex due to the large number of free parameters, and the large number of events involved. Thus, the likelihood fit is performed in three different steps, successively for each of the distributions of the observables: the mass m_i, the transversity angle θ_i (only for $\mathrm{B_s} \to \mathrm{J}/\psi\phi$), and the proper time t_i^{rec}. The only shared parameters between the $\bar{\mathrm{b}} \to \bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}$ sample and the control $\mathrm{B_s} \to \mathrm{D_s}\pi$ sample are $M_{\mathrm{B_s}}$, ΔM_{s}, τ_{s}, $\Delta\Gamma_{\mathrm{s}}/\Gamma_{\mathrm{s}}$, and ω_{tag}. We implicitly duplicate all the other parameters for the $\bar{\mathrm{b}} \to \bar{\mathrm{c}}\mathrm{c}\bar{\mathrm{s}}$ and the $\mathrm{B_s} \to \mathrm{D_s}\pi$ samples. The fit procedure is the following:

Mass distributions fit

The mass distributions are fitted in the total mass window. This yields the signal and background probabilities. The free parameters are:

- $N_{\rm sig}$;
- B/S;
- $M_{\rm B_s}$;
- $\sigma_{\rm B_s}$;
- $\kappa_{\rm bkg}$.

The above fitted parameters are then fixed.

Sidebands fit

The sidebands are used to determine the background parameters, assuming we can extrapolate its properties to the signal region. This may lead to some systematic uncertainties. Moreover, we assume that the acceptance function is the same for the signal and background events, and independent of the reconstructed mass. The free parameters in this second step are:

- $s_{\rm low}$;
- $\tau_{\rm bkg}$;
- $\alpha_{\rm bkg}$, for $\rm B_s \to J/\psi\phi$.

The above fitted parameters are then fixed. Note that the strategy with real data will certainly change as we will determine the acceptance function on signal rather than on background.

Signal window fit

The extraction of the physics parameters is done in the signal region, with all the parameters describing the background distributions, the mass distributions and the acceptance function fixed. The parameters left free in this last step are $\phi_{\rm s}$, $\Delta M_{\rm s}$, $\Delta\Gamma_{\rm s}/\Gamma_{\rm s}$, $\tau_{\rm s} = 1/\Gamma_{\rm s}$, and $\omega_{\rm tag}$. For the $\rm B_s \to J/\psi\phi$ sample, $R_{\rm T}$ is also determined. Only $\phi_{\rm s}$ and $R_{\rm T}$ are solely determined from the $\rm \bar{b} \to \bar{c}c\bar{s}$ sample.

6.4.2 Fit Results

We first present the sensitivity to $\Delta M_{\rm s}$ using the $\rm B_s \to D_s\pi$ likelihood alone, and then the results of the full simultaneous likelihood fit for each of the signal channels.

Sensitivity to $\Delta M_{\rm s}$ from $\rm B_s \to D_s\pi$

The flavor-specific channel $\rm B_s \to D_s\pi$ enables the determination of the wrong-tag fraction $\omega_{\rm tag}$ and the $\rm B_s$–$\overline{\rm B}_{\rm s}$ oscillation frequency $\Delta M_{\rm s}$. Before performing the full simultaneous fit, we can verify with what statistical precision we can extract $\omega_{\rm tag}$ and $\Delta M_{\rm s}$. To this end, we use as inputs the nominal mixing parameters and the inputs from the full MC simulation. Furthermore, we use the $\rm B_s \to D_s\pi$ tagging performance, i.e. $\omega_{\rm tag} = 31\%$ and $\varepsilon_{\rm tag} = 63\%$. This performance is obtained from off-line selected events passing the full

trigger chain, see the Appendix A. Note that we let free $\vec{\beta} = (\Delta\Gamma_s, \Gamma_s, \Delta M_s)$ and ω_{tag} in the fit.

The results are shown in Figures 6.13 and 6.14. We can see that the errors returned by the fit correctly reproduce the variance of the fitted parameters, as the rms values of the fit outputs are in good agreement with the mean errors. Moreover, the width of the pull distributions are compatible with unity, and there is no bias on ΔM_s. The expected statistical sensitivities for 2 fb^{-1} are:

- $\sigma(\Delta M_s) = \pm 0.007$ ps^{-1}, for $\Delta M_s = 17.5$ ps^{-1};
- $\sigma(\omega_{tag}) = \pm 0.36\%$, i.e. $\sigma(\omega_{tag})/\omega_{tag} = 0.012$.

We will thus have easily access to ΔM_s at LHCb, with a very small statistical uncertainty. With the available statistics at LHCb and the excellent proper time resolution, the accuracy on the present measurement of ΔM_s will be significantly improved after one year LHCb data taking.

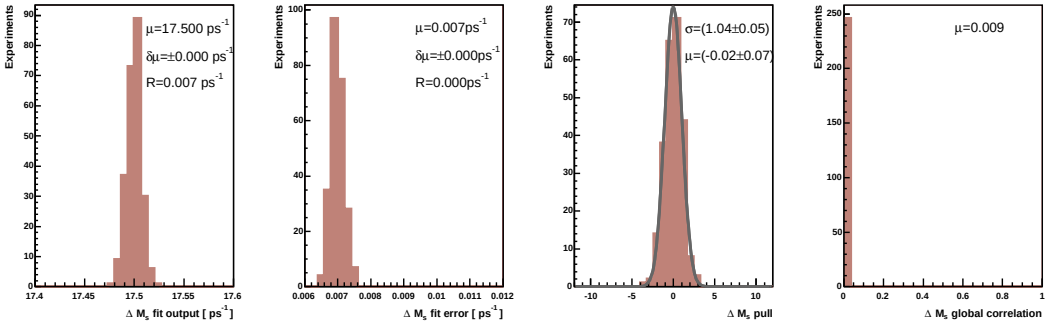


Figure 6.13: Outputs from the likelihood fit for ΔM_s from the $B_s \to D_s\pi$ sample alone, with the nominal parameters. The inputs for the tagging performance are $\omega_{tag} = 31\%$ and $\varepsilon_{tag} = 63\%$. Respectively, the plots shown from left to right are: the fitted outputs, the errors estimated by the fit, the pull distribution, and the global correlation coefficient.

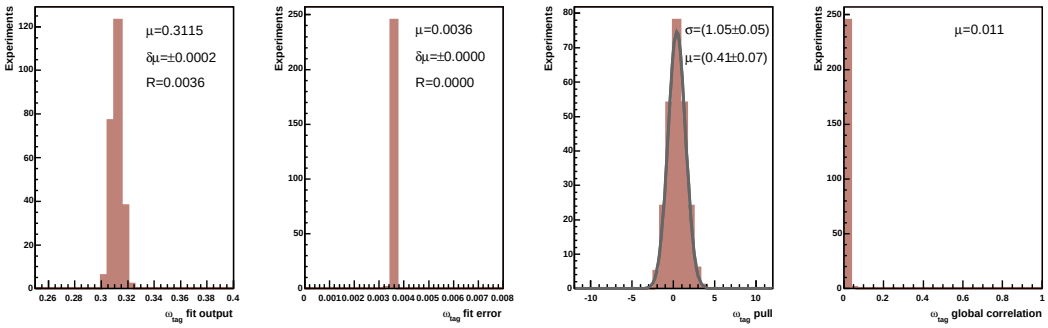


Figure 6.14: Outputs from the likelihood fit for ω_{tag} from the $B_s \to D_s\pi$ alone, with the nominal parameters. The inputs for the tagging performance are $\omega_{tag} = 31\%$ and $\varepsilon_{tag} = 63\%$. Respectively, the plots shown from left to right are: the fitted outputs, the errors estimated by the fit, the pull distribution, and the global correlation coefficient.

We can see from the global correlation coefficient in Figures 6.13 and 6.14 that ω_{tag} and ΔM_{s} are totally uncorrelated. This is expected, as ω_{tag} determines the amplitude and ΔM_{s} the frequency of the flavor oscillations.

Simultaneous fit

The full detail on the results of the simultaneous fits to the control $\text{B}_\text{s} \to \text{D}_\text{s}\pi$ and to the signal $\bar{\text{b}} \to \bar{\text{c}}\text{c}\bar{\text{s}}$ samples are given in the Appendix B.1. The fit results for the nominal inputs and $\text{B}_\text{s} \to \text{J}/\psi\phi$ are shown in Figures 6.15 – 6.20. In general, and depending on the parameter, the errors returned by the fit are not perfectly estimated. However, the rms of the fitted values is compatible with the mean error as returned by the fit when scaling with the width of the pull distribution. We will quote hereafter the sensitivities to the physics parameters as the rms of the central values. In the real experiment, only the error returned by the fit is accessible, such that we will need to actually rely on the MC to calibrate it.

We observe the following features from the $\text{B}_\text{s} \to \text{J}/\psi\phi$ fit results:

ϕ_s The errors returned by the fit are overestimated by $\sim 11\%$ and there is a $\sim 3\sigma$ bias on the fitted values. We observe a small average global correlation coefficient, with a second little peak at ~ 0.7. This has been checked to be due to the positive correlation coefficient between ϕ_s and ω_{tag}. The wrong tag exhibits the same feature in its global correlation distribution. Some correlation is to be expected between these two parameters as they both modulate the B_s–$\overline{\text{B}}_\text{s}$ oscillations.

ω_{tag} and ΔM_s These two parameters are mainly determined from the control sample. The results for ΔM_{s} are in excellent agreement with the ones obtained by fitting $\text{B}_\text{s} \to \text{D}_\text{s}\pi$ alone. Note that as we use the same tagging performance for the signal and the control sample when performing the simultaneous fit, we do have small variations.

$\tau_\text{s} = 1/\Gamma_\text{s}$ and $\Delta\Gamma_\text{s}/\Gamma_\text{s}$ These parameters are obviously strongly correlated, as a result of their definitions, and this is confirmed by the global correlation coefficient. The errors for $\Delta\Gamma_{\text{s}}/\Gamma_{\text{s}}$ are underestimated by $\sim 20\%$, whereas the effect is larger for $1/\Gamma_{\text{s}}$. The rms values are however compatible with the mean of the errors when taking into account the pull width. The width difference is hard to determine for small values of ϕ_s as we basically measure individually the lifetimes of the short-lived (Γ_{L}) CP-even eigenstate and of the long-lived (Γ_{H}) CP-odd eigenstate. Note that it was checked that the bias on the fitted values for $\Delta\Gamma_{\text{s}}/\Gamma_{\text{s}}$ is reduced when increasing significantly the statistics.

R_T The determination of the CP-odd fraction is rather accurate, with however a significant negative bias in the pull distribution.

The detail of the parameter correlations is given in Table 6.2. Each correlation coefficient is the mean of all the toy experiments used to assess the sensitivities with $\text{B}_\text{s} \to \text{J}/\psi\phi$, and with the nominal parameters. As we see from this table, ϕ_s and ω_{tag} are anticorrelated to τ_s and $\Delta\Gamma_{\text{s}}/\Gamma_{\text{s}}$. The largest correlation for ϕ_s, and however still small, is with ω_{tag} and R_{T}. As R_{T} determines the fraction between the CP-odd and the CP-even components, a negative bias will increase the CP-even contribution which is the most sensitive part to ϕ_s.

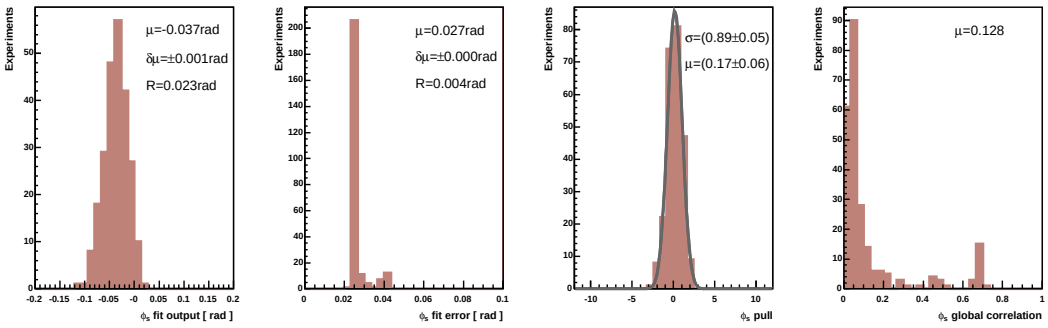


Figure 6.15: Outputs from the likelihood fit for ϕ_s from $B_s \rightarrow J/\psi\phi$, with the nominal parameters. Respectively, the plots shown from left to right are: the fitted outputs, the errors estimated by the fit, the pull distribution, and the global correlation coefficient.

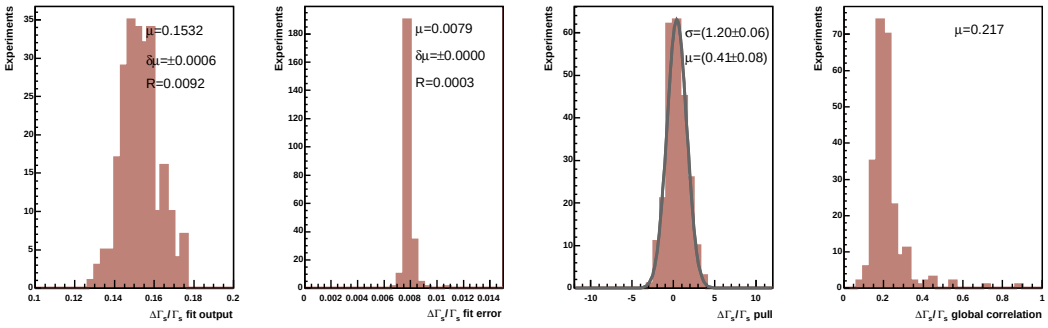


Figure 6.16: Outputs from the likelihood fit for $\Delta\Gamma_s/\Gamma_s$ from $B_s \rightarrow J/\psi\phi$, with the nominal parameters. Respectively, the plots shown from left to right are: the fitted outputs, the errors estimated by the fit, the pull distribution, and the global correlation coefficient.

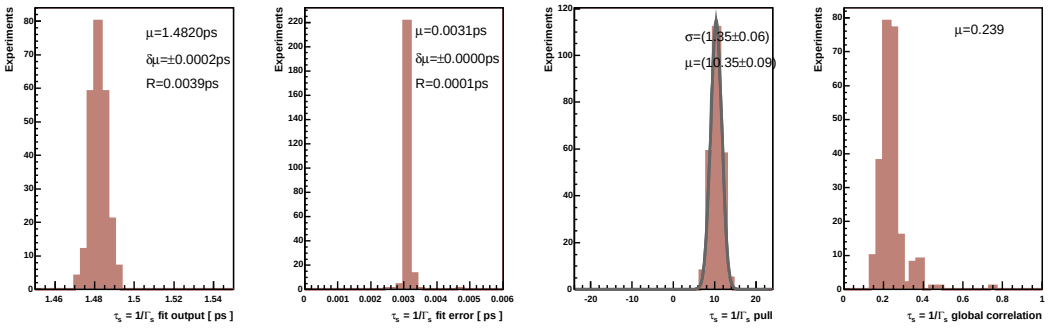


Figure 6.17: Outputs from the likelihood fit for $\tau_s = 1/\Gamma_s$ from $B_s \rightarrow J/\psi\phi$, with the nominal parameters. Respectively, the plots shown from left to right are: the fitted outputs, the errors estimated by the fit, the pull distribution, and the global correlation coefficient.

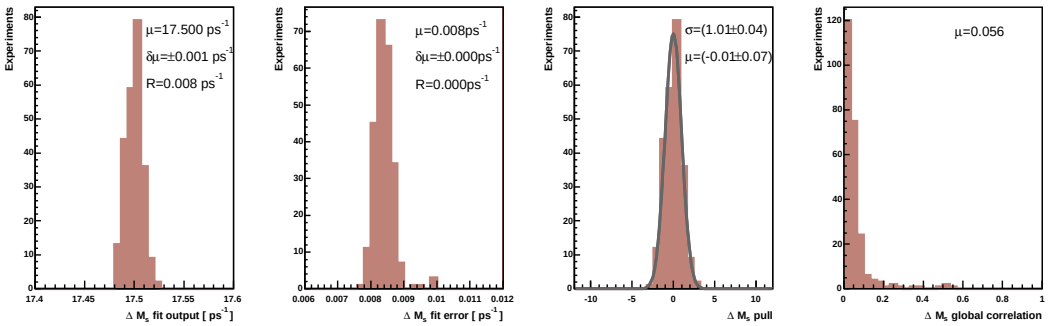


Figure 6.18: Outputs from the likelihood fit for ΔM_{s} from $\mathrm{B_s} \rightarrow \mathrm{J}/\psi\phi$, with the nominal parameters. Respectively, the plots shown from left to right are: the fitted outputs, the errors estimated by the fit, the pull distribution, and the global correlation coefficient.

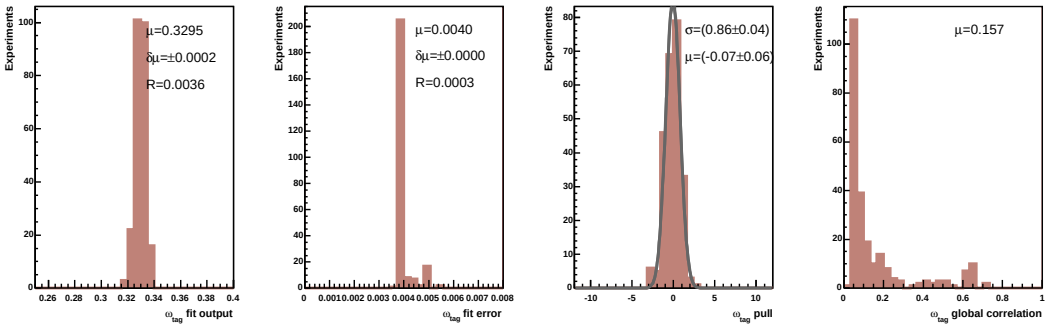


Figure 6.19: Outputs from the likelihood fit for ω_{tag} from $\mathrm{B_s} \rightarrow \mathrm{J}/\psi\phi$, with the nominal parameters. Respectively, the plots shown from left to right are: the fitted outputs, the errors estimated by the fit, the pull distribution, and the global correlation coefficient.

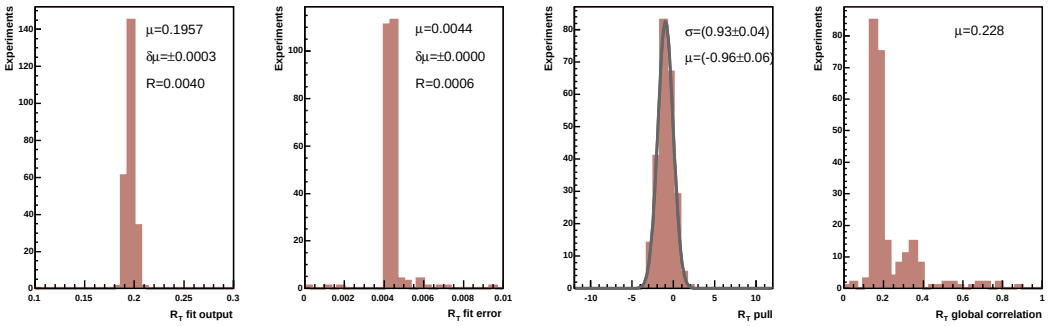


Figure 6.20: Outputs from the likelihood fit for R_{T} from $\mathrm{B_s} \rightarrow \mathrm{J}/\psi\phi$, with the nominal parameters. Respectively, the plots shown from left to right are: the fitted outputs, the errors estimated by the fit, the pull distribution, and the global correlation coefficient.

Table 6.2: Correlation matrix $(\rho)_{ij}$ of the six fitted parameters in the last step of the $B_s \to D_s\pi$ and $B_s \to J/\psi\phi$ likelihood fit. Each matrix element ρ_{ij} is the mean value of the corresponding correlation coefficient on all the toy experiments with the nominal parameters. The vector of fitted parameters is denoted by $\vec{x}$ in the table.

$$\vec{x} = \left(\Delta\Gamma_s/\Gamma_s\,,\quad R_T\,,\quad \Delta M_s\,,\quad \phi_s\,,\quad \tau_s = 1/\Gamma_s\,,\quad \omega_{tag}\right)$$

$$(\rho)_{ij} = \begin{pmatrix} 1.0 & 0.016 & 0.008 & -0.011 & 0.177 & -0.045 \\ 0.016 & 1.0 & 0.043 & 0.068 & -0.112 & 0.088 \\ 0.008 & 0.043 & 1.0 & 0.008 & -0.001 & 0.0005 \\ -0.011 & 0.068 & 0.008 & 1.0 & -0.013 & 0.052 \\ 0.177 & -0.112 & -0.001 & -0.013 & 1.0 & -0.025 \\ -0.045 & 0.088 & 0.005 & 0.052 & -0.025 & 1.0 \end{pmatrix}$$

The study of the results for the pure CP eigenstates shows similar features, as it can be seen from the tables in Appendix B.1. In general, the sensitivities to ω_{tag} and ΔM_s are compatible, showing that their determination is mainly from $B_s \to D_s\pi$. The pulls for ϕ_s are close to unity, with a small underestimation of the errors returned by the fit. There are large biases for both τ_s and $\Delta\Gamma_s/\Gamma_s$, which are due to their strong correlation and the fact that we have only the CP-even component. Indeed, for small values of ϕ_s we can *almost* only determine the short lifetime: for a small ϕ_s the signal decay rates depend only on Γ_L and Γ_s independently, and not on Γ_s and $\Delta\Gamma_s$. We still let free τ_s and $\Delta\Gamma_s/\Gamma_s$ in the fit as they do *not* affect the determination of ϕ_s. It has been checked that the same sensitivity to ϕ_s for $B_s \to \eta_c\phi$ is reached when fixing $\Delta\Gamma_s/\Gamma_s$. Note that with real data we will fit all the pure CP eigenstates channels together with $B_s \to J/\psi\phi$. Therefore, floating $\Delta\Gamma_s/\Gamma_s$ will not be an issue.

The sensitivities to ϕ_s using the pure CP eigenstates $B_s \to \eta_c\phi$, $B_s \to D_sD_s$, $B_s \to J/\psi\eta$, as well as the admixture of CP eigenstates in $B_s \to J/\psi\phi$ are summarized in Table 6.3. The best performance for the pure CP eigenstates is from $B_s \to \eta_c\phi$ even though it has the smallest statistics. This demonstrates the importance of an excellent proper time resolution. The performance for $B_s \to J/\psi\eta(\gamma\gamma)$ is similar. This channel unfortunately suffers from a worse mass resolution and (therefore) a large background level, due to the presence of two photons. The other pure CP eigenstates have much worse performances, due to their poorer statistics or worse proper time resolution (e.g. $B_s \to D_sD_s$). In comparison to the pure CP eigenstates channels, the sensitivity to ϕ_s from $B_s \to J/\psi\phi$ is much better, thanks to its large statistics. The results are combined in quadrature, using:

$$\sigma_{stat}(\phi_s) = \frac{1}{\sqrt{\sum_i \frac{1}{\sigma_i^2}}}\,,$$

where the sum runs over the different channels. The correlations arising from the use of the same control sample for each signal fit are ignored. Such issue will disappear once a

simultaneous fit to all the channels is performed. If we scale the expected statistical sensitivity to $10\ \text{fb}^{-1}$ (corresponding to five years of data taking at the nominal luminosity), we get a combined statistical sensitivity of $\sigma(\phi_\text{s}) = \pm 0.0098$ rad. This is equivalent to a $\sim 4.1\sigma$ measurement, assuming a SM value of ϕ_s. We will therefore be able to detect possible deviations from the SM expectations with relatively good accuracy, and to probe New Physics. Note that we have checked that the statistical error correctly scales with $\sqrt{5}$ by explicitly running our toys simulations with 5 times larger statistics. For $\text{B}_\text{s} \to \eta_\text{c}\phi$, we get $\sigma(\phi_\text{s}) = \pm 0.0048$ rad from the rms distribution, which is perfectly compatible with $\sigma(\phi_\text{s}) = \pm(0.108/\sqrt{5})$ rad. Finally, we note that the contribution of the pure CP eigenstates is small, as their combined statistical weight represents only about $\sim 13\%$ of the combined sensitivity to ϕ_s. This contribution is however not negligible, and the decays to pure CP eigenstates offer additional constraints in the determination of the B_s–$\overline{\text{B}}_\text{s}$ mixing phase.

Table 6.3: Combined expected statistical errors on ϕ_s for the different $\bar{\text{b}} \to \bar{\text{c}}\text{c}\bar{\text{s}}$ quark-level transitions to CP eigenstates, with $2\ \text{fb}^{-1}$ and the nominal parameters.

Channels	$\sigma(\phi_s)$ [rad]	Weight $(\sigma/\sigma_i)^2$ [%]
$\text{B}_\text{s} \to \text{J}/\psi\eta(\pi^+\pi^-\pi^0)$	0.142	2.3
$\text{B}_\text{s} \to \text{D}_\text{s}\text{D}_\text{s}$	0.133	2.6
$\text{B}_\text{s} \to \text{J}/\psi\eta(\gamma\gamma)$	0.109	3.9
$\text{B}_\text{s} \to \eta_\text{c}\phi$	0.108	3.9
Combined sensitivity for pure CP eigenstates	**0.060**	**12.7**
$\text{B}_\text{s} \to \text{J}/\psi\phi$	0.023	87.3
Combined sensitivity for all CP eigenstates	**0.022**	**100.0**

The expected sensitivities of the other physics parameters from $\text{B}_\text{s} \to \text{J}/\psi\phi$ and with $2\ \text{fb}^{-1}$ are:

- $\sigma(\Delta\Gamma_\text{s}/\Gamma_\text{s}) = \pm 0.92\%$, for $\Delta\Gamma_\text{s}/\Gamma_\text{s} = 15\%$;
- $\sigma(\tau_\text{s}) = \pm 0.0039$ ps, for $\tau_\text{s} = 1/\Gamma_\text{s} = 1.45$ ps;
- $\sigma(R_\text{T}) = \pm 0.0040$, for $R_\text{T} = 0.2$.

For each parameter, we show in Figures 6.21 – 6.23 the projections of $-\ln\mathcal{L}^{\bar{\text{b}}\to\bar{\text{c}}\text{c}\bar{\text{s}}}$ alone in a random $\text{B}_\text{s} \to \text{J}/\psi\phi$ experiment, simultaneously fitted with the control sample. In these plots, each physics parameters is fixed to the result of the simultaneous, except for the variable being plotted. We can note that the $\cos\phi_\text{s}$ term in the signal decay rates actually discriminates the ambiguity at $\pi - \phi_\text{s}$. This has been checked by running dedicated experiments and explicitly setting $\cos\phi_\text{s} = 1$ in the signal decay rates. The likelihood curve in the ϕ_s projection has a well defined quadratic-like shape. We can also see an illustration of the bias for $\Delta\Gamma_\text{s}/\Gamma_\text{s}$.

The conclusion from these likelihood projections is that the $\bar{\text{b}} \to \bar{\text{c}}\text{c}\bar{\text{s}}$ sample alone cannot determine ΔM_s, such that we need a control sample. Moreover, we can extract ϕ_s

from the $\bar{b} \to \bar{c}c\bar{s}$ sample when ω_{tag} is known, but we cannot determine both ω_{tag} and ϕ_s without the help of a control sample.

6.4.3 Scanned Parameters

We scan a few values of the parameter space in order to explore the sensitivities to the mixing parameters in different configurations. The input parameters are varied one at a time, while still letting free the other variables with their nominal values. The scan is performed for $B_s \to J/\psi\phi$ as it yields the most sensitive results, and for $B_s \to \eta_c\phi$ as it best represents the decays to pure CP eigenstates. All the results are summarized in the tables of Appendix B.2. We briefly discuss here the different behaviors observed.

The scan of the mixing parameters is the following:

- ϕ_s: -0.04rad (nominal), -0.2rad, $-\pi/4$rad. The sensitivity to ϕ_s versus ϕ_s is plotted in Figure 6.24. We checked for $B_s \to J/\psi\phi$ that the sensitivity does not depend on the sign of ϕ_s. This is valid for the pure CP eigenstates, and for $B_s \to J/\psi\phi$ without the full angular analysis.
- $\Delta\Gamma_s/\Gamma_s$: 0.05, 0.15 (nominal), 0.25. The sensitivity to ϕ_s is plotted versus $\Delta\Gamma_s/\Gamma_s$ in Figure 6.25.
- R_T: 0.0, 0.15, 0.2 (nominal), 0.25, 0.5. The sensitivity to ϕ_s is plotted versus R_T in Figure 6.26.

The most distinctive behaviors in the scans are listed hereafter:

ϕ_s For $B_s \to J/\psi\phi$ the sensitivity decreases steadily with $|\phi_s|$. The significance is already of at least 7σ for $\phi_s = -0.2$ rad. For $B_s \to \eta_c\phi$, there is a sharper raise of the statistical error for a large ϕ_s. Nevertheless, the significance with $B_s \to \eta_c\phi$ is about 5σ for $\phi_s = -\pi/4$ rad.

$\Delta\Gamma_s/\Gamma_s$ The statistical error on ϕ_s gently decreases with larger values of $\Delta\Gamma_s/\Gamma_s$. A larger width difference enables to better distinguish the mass eigenstates, thus increasing the sensitivity to ϕ_s.

R_T The pure CP-even limit, i.e. $R_T = 0$ gives as expected the best sensitivity to ϕ_s. The maximal dilution case, i.e. $R_T = 0.5$, gives the largest statistical error on ϕ_s. In particular, we see the effect of the distinctive angular distributions between the CP components, as if the latter had the same shape, we would not have any sensitivity at all. Even in the case of $R_T = 0.5$, we still have a better sensitivity with $B_s \to J/\psi\phi$ than with all the other pure CP eigenstates combined. We also note that the sensitivity to $\Delta\Gamma_s/\Gamma_s$ increases with smaller values of R_T, as a result of the increasing "purity" in the CP components (thus we expect the same effect for a CP-odd fraction approaching unity).

As expected, the sensitivities to ΔM_s are practically unaffected by the above scans, meaning that the oscillation frequency is uncorrelated with these parameters. For a large ϕ_s the statistical precision to ω_{tag} is slightly better, due to the correlations mentioned previously.

We do not show any result of a ΔM_s scan as it is now relatively well measured. Nevertheless, we know that larger values of ΔM_s will spoil the performance on ϕ_s as it becomes harder to resolve the B_s–$\overline{B}_s$ oscillations. In order to evaluate the effect of the proper time on our results, we performed a scan of the proper time scale factor, and assuming that the

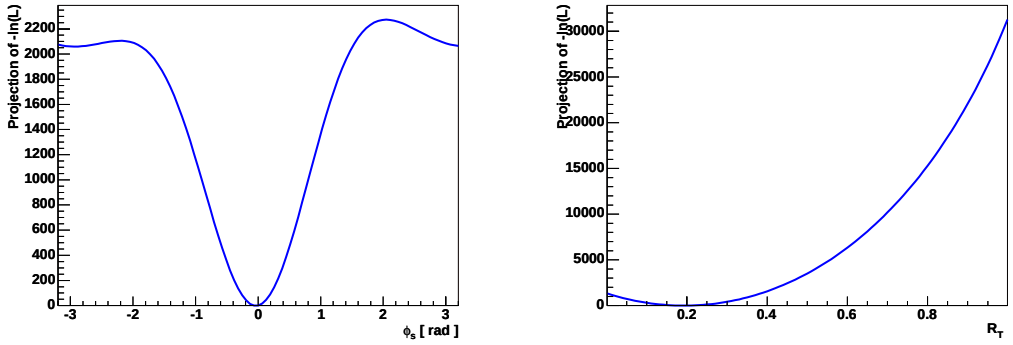


Figure 6.21: Projection of $-\ln\mathcal{L}^{\bar{b}\to\bar{c}c\bar{s}}$ for ϕ_s (left) and R_T (right), corresponding to a random experiment of $B_s \to J/\psi\phi$ with the nominal parameters, and simultaneously fitted with the control sample.

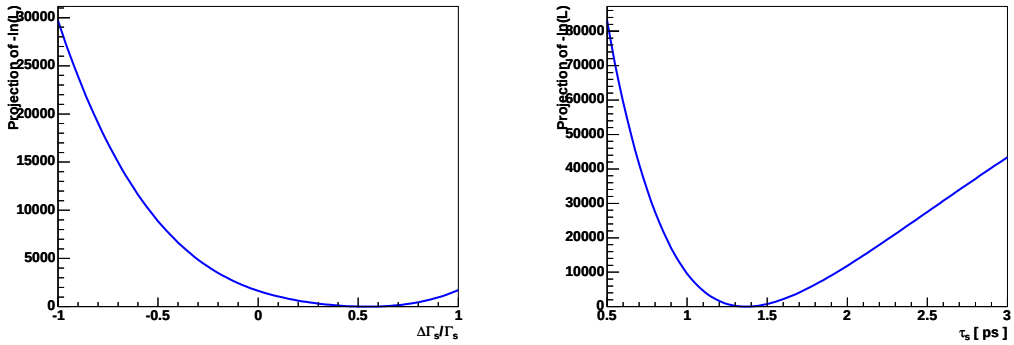


Figure 6.22: Projection of $-\ln\mathcal{L}^{\bar{b}\to\bar{c}c\bar{s}}$ for $\Delta\Gamma_s/\Gamma_s$ (left) and $\tau_s = 1/\Gamma_s$ (right), corresponding to a random experiment of $B_s \to J/\psi\phi$ with the nominal parameters, and simultaneously fitted with the control sample.

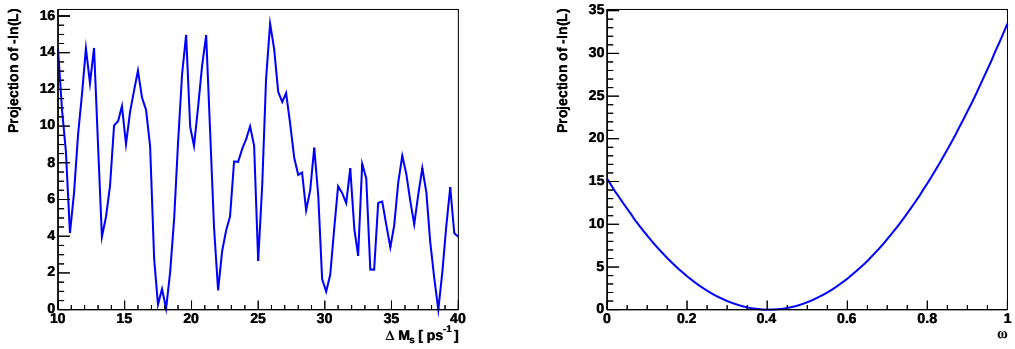


Figure 6.23: Projection of $-\ln\mathcal{L}^{\bar{b}\to\bar{c}c\bar{s}}$ for ΔM_s (left) and ω_{tag} (right), corresponding to a random experiment of $B_s \to J/\psi\phi$ with the nominal parameters, and simultaneously fitted with the control sample.

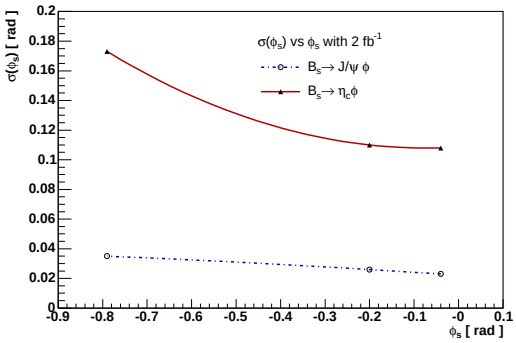


Figure 6.24: Sensitivity to $\sigma(\phi_s)$ [rad] versus ϕ_s [rad] for $B_s \to J/\psi\phi$ and $B_s \to \eta_c\phi$ with $2\,fb^{-1}$. All the other parameters are kept at their nominal values.

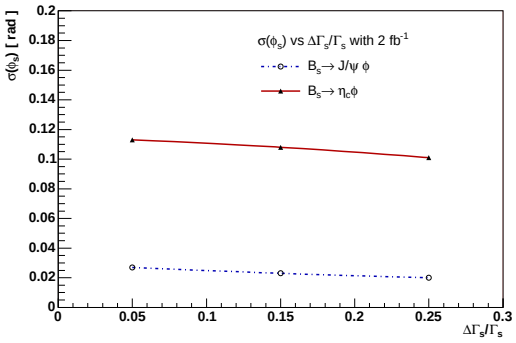


Figure 6.25: Sensitivity to $\sigma(\phi_s)$ [rad] versus $\Delta\Gamma_s/\Gamma_s$ for $B_s \to J/\psi\phi$ and $B_s \to \eta_c\phi$ with $2\,fb^{-1}$. All the other parameters are kept at their nominal values.

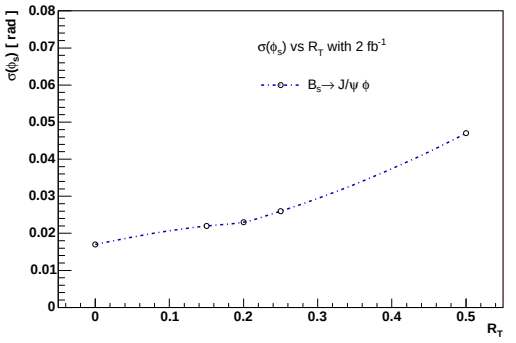


Figure 6.26: Sensitivity to $\sigma(\phi_s)$ [rad] versus R_T for $B_s \to J/\psi\phi$ with $2\,fb^{-1}$. All the other parameters are kept at their nominal values.

errors are overestimated or underestimated by 10%, when compared to the nominal scale factor obtained from the full MC. We also performed the fit for a twice larger background level. The results of these scans are given in Table 6.4. Note that we implicitly scan the $B_s \to D_s\pi$ parameters in the same way as the signal sample.

Table 6.4: Expected statistical errors on ϕ_s for a few special settings, and the nominal parameters. The $B_s \to D_s\pi$ parameterization is changed in the same way as the signal sample. The results are for 2 fb^{-1}.

	$\sigma(\phi_s)$ [rad]	
Scan	$B_s \to J/\psi\phi$	$B_s \to \eta_c\phi$
Nominal	0.023	0.108
$\Sigma_\tau + 10\,\%$	0.025	0.108
$\Sigma_\tau - 10\,\%$	0.023	0.103
$B/S \times 2$	0.025	0.118

The impact of a scale factor changing by $\pm 10\%$ is not too drastic on the sensitivity to ϕ_s. The effect of a twice larger background level shows us that the B/S ratio needs to be under control, but it does not spoil the results too much for $B_s \to J/\psi\phi$. For $B_s \to \eta_c\phi$, a B/S of the order of one seems a good compromise.

Throughout this chapter, we assumed a given shape for the background. This will have to be determined from the data. Nevertheless, we do not expect any significant changes in the performance. Yet another issue is the modelling of the proper time, where we used a single Gaussian pull resolution function. For lifetime unbiased selections using a J/ψ, we would need to add a prompt component as well in order to better characterize the background shape. Finally, as a test of the dependence on the proper time resolution model, we performed a scan for $B_s \to J/\psi\phi$ with a double Gaussian proper time resolution function. The inputs used are determined by fitting the proper time pull from the full MC, yielding $\Sigma_\tau^1 = 1.13$ for the core Gaussian (86%), and $\Sigma_\tau^2 = 1.80$ (14%) for the second one. We also modify the $B_s \to D_s\pi$ resolution model in the same fashion. The result of the likelihood maximization for a large number of toys yields $\sigma(\phi_s) = \pm 0.021$ rad, to be compared with the nominal result of $\sigma(\phi_s) = \pm 0.023$ rad. We get a better sensitivity, which is due to the core resolution which has a thinner width. We can anyhow build more complex resolution functions, e.g. adding a bias scaled with the proper time scale factor. However, we do not expect this to have an important impact on the statistical error on the weak mixing phase ϕ_s.

6.5 Outlook and Future Improvements

A general likelihood fitting code has been developed for the study of the flavor-specific decays (e.g. $B_s \to D_s\pi$), and for the study of $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions to pure CP eigenstates (e.g. $B_s \to \eta_c\phi$) and to an admixture of CP eigenstates (e.g. $B_s \to J/\psi\phi$). The fit uses the parameterizations obtained from the full MC simulation. Moreover, the per-event proper time errors from the full MC are used to assign a proper time uncertainty to

the different events, and we take into account the underestimation of these errors by the off-line tracking using a resolution scale factor. The likelihood description is realistic, and accounts for background, detector resolution effects, as well as the tagging performance and the wrong-tag dilution.

The expected sensitivities to the B_s–$\overline{B}_s$ mixing parameters at LHCb after one year data taking are listed in Table 6.5, for our nominal set of parameters. The statistical uncertainty on ΔM_s obtained from $B_s \to D_s\pi$ is extremely small. The combined sensitivity to ϕ_s is such that we can expect a $\sim 4.1\sigma$ measurement for a SM model ϕ_s after five years data taking, only considering the statistical uncertainties. The contribution from the pure CP eigenstates is at the level of $\sim 13\%$, serving as an extra constraint for the measurement dominated by the $B_s \to J/\psi\phi$ performance.

Table 6.5: Expected statistical errors to the B_s–$\overline{B}_s$ mixing parameters at LHCb with $2\,\text{fb}^{-1}$, and with the nominal parameters. For completeness, the sensitivities to ω_{tag} and R_T are also quoted.

Parameter	Sensitivity	Channel
ϕ_s [rad]	0.022	$J/\psi\phi$, $\eta_c\phi$, $J/\psi\eta(\gamma\gamma)$, $J/\psi\eta(\pi\pi\pi)$, D_sD_s
$\Delta\Gamma_s/\Gamma_s$	0.0092	$J/\psi\phi$
ΔM_s [ps^{-1}]	0.007	$D_s\pi$
ω_{tag}	0.0036	$D_s\pi$
R_T	0.00040	$J/\psi\phi$

The study of the correlations shows that the observables are in general uncorrelated, allowing in this way a proper determination of the parameters. A parameter scan has also been performed to underline the dependencies on larger – non SM – values of the mixing parameters. Remarkably, we will have a significant sensitivity to ϕ_s for larger values of this parameter.

Finally, we draw a non-exhaustive list of the possible future improvements or strategies:

- Improve the proper time resolution description, and define a strategy to use and extract the proper time resolution directly from the data.

- Model and determine the proper time acceptance function directly from the reconstructed data. We should also push forward the study of a lifetime unbiased measurement with $B_s \to J/\psi\phi$. The latter will probably yield the first results on the weak B_s–$\overline{B}_s$ mixing phase determination once LHCb will have its first data.

- Understand and evaluate the systematic uncertainties arising from the use of a control sample to extract the mistag probability. This goes through phase space reweighting of the samples, and the differentiation of the tagging categories.

- Study the possibility of a full angular analysis for $B_s \to J/\psi\phi$. In order to ease the analysis with real data, we could in a first step integrate over all angles and

perform the CP measurement without any angular information. Alternatively, we could integrate over the time to extract information on R_{T}.

- Perform a combined likelihood fit to all the signal samples, in order to account for uncertainties due to the use of the same control sample. Moreover, this enables the determination of the width difference from $\mathrm{B_s} \to \mathrm{J}/\psi\phi$ to be used for the decays to pure CP eigenstates.
- Add new channels, in particular the decay mode $\mathrm{B_s} \to \mathrm{J}/\psi\phi$ with electrons.

The goal is of course to be ready for the measurement of ϕ_s, and the possibility to probe New Physics.

Conclusion

THE studies presented in this dissertation can be divided into three parts: the development of the exclusive High-Level Trigger (HLT), the $B_s \to \eta_c \phi$ event selection, and the study of the sensitivity to the B_s–$\overline{B}_s$ mixing parameters using several decays mediated by $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions. We outline hereafter the main results obtained in this thesis.

The first prototype of the exclusive b and the D^* HLT selections has been successfully implemented in the full Monte Carlo simulation. The design and implementation choices have been motivated, and we demonstrated the feasibility of this trigger stage. The use of common tools and algorithms for all the exclusive HLT selections ensures their maximal correlations, while limiting the amount of code used. Moreover, the selection criteria chosen are based on typical b-hadron signatures, thus avoiding the use of too specific selection cuts. The minimum-bias output rates are shown to be under control and reduced to the design level, while keeping relatively large efficiency on the decays of interest. The timing performance has been shown to be in the required range. The purity of the minimum-bias sample after the HLT exhibits a b and c contents of $\sim 70\%$.

In the HLT study, we have emphasized a few weaknesses of the on-line reconstruction, to be considered for future improvements. The dominant source of inefficiency for the exclusive HLT selections is due to events with at least a missing signal track, when this track was found by the off-line reconstruction. Thus, the total HLT efficiencies on the core signal channels vary in a large range, 70–100%, and strongly depend on the type and multiplicity of the decay under study. Furthermore, the off-line selections should revise their selection criteria in order to have a better correlation with the HLT. The non-signal component of the selected bandwidth is primarily due to badly reconstructed tracks, thus faking large impact parameter particles, or to events with a non-reconstructed primary vertex.

The $B_s \to \eta_c \phi$ event selection has been performed using the full Monte Carlo simulation. The selection algorithms have been designed to select the largest number of signal events, while keeping the background level under control. An annual event yield of $\sim 3'000$ events is expected after the trigger and the off-line selections, with LHCb's nominal running conditions. However, this channel being a 6-prong, it is limited by the track finding efficiencies as a result of its large multiplicity. Any substantial improvement in the tracking performance will thus fully benefit to $B_s \to \eta_c \phi$. A detailed study of several sources of background has been done, where the dominant contribution was found to be from inclusive $b\overline{b}$ events yielding a 90% unified confidence interval of $[0.10, 1.17]$ for the background level, and a central value of $B/S = 0.4$. In addition, the study of specific D_s decays has shown a significant contribution from $B_s \to D_s \pi\pi\pi$ decays, with a ratio

$B/S = 0.2$.

A detailed study of the $B_s \to \eta_c\phi$ signal characteristics is provided. In particular, an excellent B_s mass resolution is obtained, with a value of 12 MeV/c^2. The different contributions to the B_s proper time have been pointed out, and this channel possibly has the best proper time performance of all decay channels studied at LHCb, with a proper time resolution of ~ 31 fs. Finally, the tagging power reaches $\sim 10\%$. Even though it is limited by statistics, the $B_s \to \eta_c\phi$ decay channel offers an interesting potential for the determination of the B_s–$\overline{B}_s$ mixing phase, given its excellent reconstruction features.

The sensitivity to the B_s–$\overline{B}_s$ mixing parameters has been assessed based on a fast parameterized Monte Carlo simulation. This simulation uses a realistic description of the resolutions, the per-event proper time errors, the event yields, the background levels, the tagging performance, as obtained from the full Monte Carlo of the different decay channels used. A likelihood fitting code has been developed to extract the relevant observables for the $\bar{b} \to \bar{c}c\bar{s}$ quark-level transitions to CP eigenstates, as well as to the variables of flavor-specific decay channels. The sensitivities to the different observables are obtained by generating and fitting many experiments, each corresponding to one year LHCb data taking, and for each of the channels considered. The rms of the fitted parameters are then used to quote their statistical errors.

The sensitivity to the B_s–$\overline{B}_s$ oscillation frequency has been determined to be $\sigma(\Delta M_s) = \pm 0.007$ ps^{-1}, from a flavor-specific sample taken to be the decay $B_s \to D_s\pi$. For the determination of the mixing phase ϕ_s, the $B_s \to \eta_c\phi$, $B_s \to D_sD_s$, $B_s \to J/\psi\eta(\gamma\gamma)$, $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0)$, and $B_s \to J/\psi\phi$ decays were considered in a simultaneous fit with the control sample. The dominant contribution is found to be that of $B_s \to J/\psi\phi$, a decay to an admixture of CP eigenstates. Using the one-angle angular analysis in the transversity basis, we thus found a statistical sensitivity of $\sigma(\phi_s) = \pm 0.023$ rad, with 2 fb^{-1}, and for a Standard Model ϕ_s. The contribution of the pure CP eigenstates amounts to a substantial $\sim 13\%$, yielding a combined sensitivity to ϕ_s at LHCb of $\sigma(\phi_s) = \pm 0.022$ rad for all the channels considered. The largest contributions from the decays to pure CP eigenstates come from the $B_s \to \eta_c\phi$ and $B_s \to J/\psi\eta(\gamma\gamma)$ decay modes. The natural continuation of this study would be to account for systematic uncertainties induced by the control sample, and perform a full angular analysis of $B_s \to J/\psi\phi$.

The determination of the B_s–$\overline{B}_s$ is one of the most exciting measurement in flavor physics, since the B_s meson is totally to be explored. After a few years, LHCb will be able to measure ϕ_s assuming a Standard Model value, and any sizable deviation from the Standard Model expectation would immediately open the gate towards New Physics, that may be hiding in the B_s–$\overline{B}_s$ mixing. The prospects are that LHCb really does have the potential to perform the first measurement of the yet unknown mixing phase ϕ_s, and to possibly exclude several New Physics models already in 2008!

Appendix A

Full MC Results

We briefly describe the performance of the reconstruction and selection of $B_s \to J/\psi\phi$, $B_s \to D_s\pi$, $B_s \to J/\psi\eta$ and $B_s \to D_sD_s$ decays, based on the full MC simulation. The results are used in the sensitivity studies of Chapter 6. All the results are obtained with DC04 MC data, and, unless otherwise specified, the branching ratios used are from [27]. The different residuals and pulls are fitted assuming single Gaussian distributions.

A.1 $B_s \to J/\psi\phi$ Event Selection

The $B_s \to J/\psi\phi$ off-line selection is the same as in the Reoptimization TDR [17], and the selection cuts are given in [102]. The signal window is $\pm 50\,\mathrm{MeV/c^2}$ around the B_s mass, and the loose mass window used for the $b\overline{b}$ estimate is $\pm 600\,\mathrm{MeV/c^2}$.

The visible branching ratio $BR_{vis}^{B_s \to J/\psi\phi}$ is given by:

$$\begin{aligned}
BR_{vis}^{B_s \to J/\psi\phi} &= BR(B_s \to J/\psi\phi) \times BR(\phi \to K^+K^-) \times BR(J/\psi \to \mu^+\mu^-, \gamma) \\
&= (30.9 \pm 11.0) \cdot 10^{-6}\,, \\
BR(B_s \to J/\psi\phi) &= (9.3 \pm 3.3) \cdot 10^{-3}\,, \\
BR(\phi \to K^+K^-) &= (49.2 \pm 0.6) \cdot 10^{-2}\,, \\
BR(J/\psi \to \mu^+\mu^-) &= (5.88 \pm 0.1) \cdot 10^{-2}\,, \\
BR(J/\psi \to \mu^+\mu^-\gamma) &= (0.88 \pm 0.14) \cdot 10^{-2}\,, \\
BR(J/\psi \to \mu^+\mu^-, \gamma) &= BR(J/\psi \to \mu^+\mu^-) + BR(J/\psi \to \mu^+\mu^-, \gamma) = (6.76 \pm 0.17) \cdot 10^{-2}\,.
\end{aligned}$$

The results of the off-line selection and the trigger efficiencies on off-line selected $B_s \to J/\psi\phi$ events in the tight mass window are:

Channel	N_{gen}	N_{sel}	N_{L0}	N_{L1}	N_{HLTGen}	N_{HLT}
$B_s \to J/\psi\phi$	1659000	123575	115594	109006	104622	98156

The trigger efficiency breakdown (with statistical uncertainties) is:

Channel	$\varepsilon_{L0/sel}$ [%]	$\varepsilon_{L1/L0}$ [%]	$\varepsilon_{HLTGen/L1}$ [%]	$\varepsilon_{HLT/HLTGen}$ [%]
$B_s \to J/\psi\phi$	93.5 ± 0.1	94.3 ± 0.1	96.0 ± 0.1	93.8 ± 0.1

The tagging efficiency ε_{tag}, the wrong-tag fraction ω_{tag}, and the effective tagging efficiency ε_{eff} after $B_s \to J/\psi\phi$ selection and each trigger level (with statistical uncertainties) are:

$B_s \to J/\psi\phi$	Selection [%]	L0 [%]	L1 [%]	HLT [%]
ε_{tag}	56.1 ± 0.1	56.7 ± 0.2	57.0 ± 0.2	57.0 ± 0.2
ω_{tag}	33.3 ± 0.2	33.2 ± 0.2	33.1 ± 0.2	33.0 ± 0.2
ε_{eff}	6.3 ± 0.2	6.4 ± 0.2	6.5 ± 0.2	6.6 ± 0.2

The summary of the $B_s \to J/\psi\phi$ performance is:

B_s mass resolution: $\sigma_m^{\text{res}}(B_s) \sim 14\,\text{MeV}/c^2$.

Per-event proper time errors: $< \tau_{\text{fit}}^{\text{err}} >= 29.5$ fs. The proper time resolution is $\sigma_\tau \sim 35$ fs.

Scale factor: the standard deviation of the pull distribution is $\Sigma_\tau = 1.22$.

Annual event yield: (131193 ± 1205) k ~ 131 k, after full trigger selection ($2\,\text{fb}^{-1}$, MC statistical errors only).

Tagging performance: tagging efficiency $\varepsilon_{\text{tag}} = 57.0\%$, wrong-tag fraction $\omega_{\text{tag}} = 33.0\%$ after trigger, corresponding to a tagging power of $\varepsilon_{\text{eff}} = 6.6\%$.

Background level: The selection is applied blindly to the DC04-v2 stripped inclusive $b\overline{b}$ data ($\sim$ 27M events). The results in the enlarged $\pm 600\,\text{MeV}/c^2$ mass window before applying the trigger are:

- **12 low-mass background** candidates divided as:
  - 1 $B_s \to J/\psi\phi$ with an additional γ (from the B_s): not counted as combinatorial background (outside the signal region);
  - 2 $B_s \to J/\psi\phi$ with two additional γ (from the B_s): not counted as combinatorial background (outside the signal region);
  - 7 $B_s \to J/\psi\phi$ with an additional π^0 (from the B_s): not counted as combinatorial background (outside the signal region);
  - 2 $B_s \to J/\psi\phi$ with two additional charged π (from the B_s): not counted as combinatorial background (outside the signal region).

- **1 partially reconstructed decay** of the type:
  - 1 $B_d \to (J/\psi \to \mu^+\mu^-)(K_1(1270)^0 \to K^-\pi^+)\pi^0$, with the pion misidentified as a kaon, not counted as combinatorial background (outside the signal region).

- 10 **signal** candidates.

- 4 **ghost** candidates:
  - 1 $\overline{B}_d \to (J/\psi \to \mu^+\mu^-)\pi^0 K^-$ with an additional ghost K^+, counted as combinatorial background;
  - 1 $B_s \to (D_s^{-*} \to \gamma(D_s^- \Rightarrow K^+K^-\pi^-\pi^0))\mu^+\nu_\mu$ with an additional ghost μ^-, counted as combinatorial background;

  - 1 $B_d \to (D^- \to K^+\pi^-\pi^-)\mu^+\nu_\mu$ using the μ^+, a primary K^+, a ghost K^-, and a μ^- from the decay in flight of a primary pion, counted as combinatorial background;
  - 1 signal decay with a ghost due to the inefficiency of the associator, not counted as combinatorial background.

- **From primary vertex (same collision)**: 9 events with at least one of the final states originating from the same primary vertex as the b hadron in the partially reconstructed event, all events are counted as combinatorial background.

- **From different primary vertices**: 1 event reconstructing partially a b decay, with one track taken from a different collision, counted as combinatorial background.

- **$b\overline{b}$**: 2 events with tracks from different b hadrons:
  - 1 event with the ϕ from $B_u^- \to (D^{*0} \to \pi^0(D^0 \to K^0\phi))D^{*-}D^0$ and the J/ψ from $B_u^+ \to J/\psi\pi^0\pi^+K^{*0}$, counted as combinatorial background;
  - 1 event taking tracks from a 11 charged tracks B_u^- decay and from another b hadron, counted as combinatorial background.

After removal of the non-dangerous backgrounds, the total number of events to be considered for the combinatorial background level estimate is 15. Using the mass window trick presented in Section 5.7, a central value of $B/S|_{b\overline{b}} = (0.12 \pm 0.03)$ is obtained, where the error is from statistics only, and we assumed the same trigger efficiency on background and signal events.

For the sensitivity studies of Chapter 6, we will use a nominal $B/S = 0.12$ and scan with a twice larger background level to account for other sources of background. This is motivated by the 50% fraction of non-b events after the exclusive HLT, see Chapter 4.

Acceptance parameterization: using (5.14), the acceptance parameter is $s_{\text{low}} = 2.81\ \text{ps}^{-1}$ after selection and trigger.

A.2 $B_s \to D_s\pi$ Event Selection

The results of the $B_s \to D_s\pi$ off-line selection are from [107], with the updated branching ratios from [27]. The signal window is $\pm 50\ \text{MeV}/c^2$ around the B_s mass, and the loose mass window used for the $b\overline{b}$ estimate is $\pm 500\ \text{MeV}/c^2$.

The estimated visible branching ratio $BR_{\text{vis}}^{B_s \to D_s\pi}$ is given by:

$$\begin{aligned} BR_{\text{vis}}^{B_s \to D_s\pi} &= BR(B_s \to D_s\pi) \times BR(D_s \to KK\pi) = (150.3 \pm 30.1) \cdot 10^{-6}\,, \\ BR(B_s \to D_s\pi) &= BR(B_d \to D^-\pi) = (2.89 \pm 0.29) \cdot 10^{-3}\,, \\ BR(D_s \to KK\pi) &= (5.2 \pm 0.9) \cdot 10^{-2}\,. \end{aligned}$$

The results of the off-line selection selection and the trigger efficiencies on off-line selected $B_s \to D_s\pi$ events in the tight mass window are:

Channel	N_{gen}	N_{sel}	N_{L0}	N_{L1}	N_{HLTGen}	N_{HLT}
$B_s \to D_s\pi$	3990570	158889	68835	56740	53692	44508

The trigger efficiency breakdown (statistical uncertainties) is:

Channel	$\varepsilon_{\text{L0/sel}}$ [%]	$\varepsilon_{\text{L1/L0}}$ [%]	$\varepsilon_{\text{HLTGen/L1}}$ [%]	$\varepsilon_{\text{HLT/HLTGen}}$ [%]
$B_s \to D_s\pi$	43.3 ± 0.1	82.4 ± 0.2	94.6 ± 0.1	82.9 ± 0.2

The tagging efficiency ε_{tag}, the wrong-tag fraction ω_{tag}, and the effective tagging efficiency ε_{eff} after $B_s \to D_s\pi$ selection and each trigger level (with statistical uncertainties) are:

$B_s \to D_s\pi$	Selection [%]	L0 [%]	L1 [%]	HLT [%]
ε_{tag}	58.1 ± 0.1	61.4 ± 0.2	62.9 ± 0.2	62.8 ± 0.2
ω_{tag}	32.8 ± 0.2	31.4 ± 0.2	31.3 ± 0.2	31.0 ± 0.3
ε_{eff}	6.9 ± 0.2	8.5 ± 0.3	8.8 ± 0.3	9.1 ± 0.3

The summary of the $B_s \to D_s\pi$ performance is:

B_s mass resolution: $\sigma_m^{\text{res}}(B_s) \sim 14\,\text{MeV/c}^2$.

Per-event proper time errors: $< \tau_{\text{fit}}^{\text{err}} >= 32.9$ fs. The proper time resolution is $\sigma_\tau \sim 39$ fs.

Scale factor: the standard deviation of the pull distribution is $\Sigma_\tau = 1.21$.

Annual event yield: (120158 ± 1181) k ~ 120 k, after full trigger selection (2 fb^{-1}, MC statistical errors only).

Tagging performance: tagging efficiency $\varepsilon_{\text{tag}} = 62.8\%$, wrong-tag fraction $\omega_{\text{tag}} = 31.0\%$ after trigger, corresponding to a tagging power of $\varepsilon_{\text{eff}} = 9.1\%$. In the sensitivity studies, we assume the same tagging efficiency for $B_s \to D_s\pi$ and the signal channel. The tagging performance used for $B_s \to D_s\pi$ is that of the signal it controls.

Background level: The selection is applied blindly to the DC04-v2 stripped inclusive $b\bar{b}$ data ($\sim$ 27 M events), using the mass window trick. In addition a few specific backgrounds were studied. The central value of the background level is $B/S = 0.4$.

Acceptance parameterization: using (5.14), the acceptance parameter is $s_{\text{low}} = 1.36\,\text{ps}^{-1}$ after selection and trigger.

A.3 $B_s \to D_sD_s$ Event Selection

The results of the $B_s \to D_sD_s$ off-line selection (with $D_s \to KK\pi$) are from [108] and before the HLT, using DC04 data. An educated guess based on the experience of Chapter 4 is used to extrapolate the yield after the full trigger chain.

The summary of the $B_s \to D_sD_s$ performance is:

B_s mass resolution: $\sigma_m^{\text{res}}(B_s) \sim 6\,\text{MeV/c}^2$. The D_s vertices are fitted with a mass constraint.

Per-event proper time errors: $< \tau_{\text{fit}}^{\text{err}} >= 44.4$ fs. The proper time resolution is $\sigma_\tau \sim 40 - 50$ fs.

Scale factor: the standard deviation of the pull distribution is $\Sigma_\tau = 1.26$.

Annual event yield: ~ 4 k, after full trigger selection (2 fb^{-1} guess).

Tagging performance: tagging efficiency $\varepsilon_{\text{tag}} \sim 57\%$, wrong-tag fraction $\omega_{\text{tag}} \sim 34\%$ after the Level-1 trigger, corresponding to a tagging power of $\varepsilon_{\text{eff}} \sim 6\%$.

Background level: 90% CL upper limit $B/S = 0.3$, from inclusive $b\overline{b}$ and some specific backgrounds.

Acceptance parameterization: using a similar parameterization as (5.14), the acceptance parameter is $s_{\text{low}} = 1.6$ ps^{-1} after selection and Level-1 trigger.

A.4 $B_s \to J/\psi\eta$ Event Selection

The results of the $B_s \to J/\psi\eta$ off-line selection are from [109] and after the HLT, using DC04 data. The $B_s \to J/\psi\eta$ decay channel is reconstructed in the $B_s \to J/\psi(\mu^+\mu^-)\eta(\gamma\gamma)$ and $B_s \to J/\psi(\mu^+\mu^-)\eta(\pi^+\pi^-\pi^0 \to (\gamma\gamma))$ decay modes. Note that a Kalman filter is applied to the $B_s \to J/\psi\eta$ decay chains to improve the resolutions. This results in a large proper time pull indicating a possible problem with the calibration of neutrals.

The summary of the $B_s \to J/\psi\eta$ performance is:

B_s mass resolution:

- $B_s \to J/\psi\eta(\gamma\gamma) : \sigma_m^{\text{res}}(B_s) \sim 34$ MeV/c^2.
- $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0) : \sigma_m^{\text{res}}(B_s) \sim 20$ MeV/c^2.

Per-event proper time errors:

- $B_s \to J/\psi\eta(\gamma\gamma) : < \tau_{\text{fit}}^{\text{err}} >= 30.4$ fs. The proper time resolution is $\sigma_\tau \sim 36$ fs.
- $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0) : < \tau_{\text{fit}}^{\text{err}} >= 25.5$ fs. The proper time resolution is $\sigma_\tau \sim 33$ fs.

Scale factor: the standard deviation of the pull distribution

- $B_s \to J/\psi\eta(\gamma\gamma) : \Sigma_\tau = 1.22$.
- $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0) : \Sigma_\tau = 1.32$.

Annual event yield:

- $B_s \to J/\psi\eta(\gamma\gamma) : \sim 8.5$ k, after full trigger selection (2 fb^{-1}).
- $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0) : \sim 3$ k, after full trigger selection (2 fb^{-1}).

Tagging performance:

- $B_s \to J/\psi\eta(\gamma\gamma)$: tagging efficiency $\varepsilon_{\text{tag}} \sim 63\%$, wrong-tag fraction $\omega_{\text{tag}} \sim 35\%$ after trigger, corresponding to a tagging power of $\varepsilon_{\text{eff}} \sim 6\%$.

- $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0)$: tagging efficiency $\varepsilon_{tag} \sim 62\%$, wrong-tag fraction $\omega_{tag} \sim 30\%$ after trigger, corresponding to a tagging power of $\varepsilon_{eff} \sim 10\%$.

Background level:

- $B_s \to J/\psi\eta(\gamma\gamma)$: central value $B/S = 2.0$, from inclusive $b\overline{b}$ and some specific backgrounds.
- $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0)$: $B/S < 3.0$ at 90% CL from inclusive $b\overline{b}$.

Acceptance parameterization: using the parameterization (5.14)

- $B_s \to J/\psi\eta(\gamma\gamma)$: the acceptance parameter is $s_{low} = 1.86\,\text{ps}^{-1}$ after selection and trigger.
- $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0)$: the acceptance parameter is $s_{low} = 1.54\,\text{ps}^{-1}$ after selection and trigger.

Appendix B

Full Sensitivity Results

We present in this appendix the full detail of the fit results of Chapter 6. The entries of the tables in this appendix are:

- Fit output: the mean of the distribution of the fit output and the rms.
- Fit error: the mean of the error as returned by the fit.
- Pull distribution: mean and standard deviation (Σ) of a Gaussian fit to the residual $\delta x = x_{\rm out} - x_{\rm in}$ normalized to the fitted error $\sigma_x^{\rm err}$.
- Global correlation coefficient (G.C.): the mean of the G.C. distribution.

The global correlation coefficient ρ_i is a measure of the largest correlation between the i-th variable x_i and every possible linear combination of all the other variables. It is defined by:

$$\rho_i = \sqrt{1 - \frac{1}{(V)_{ii}(V^{-1})_{ii}}}\,, \quad 0 \leq \rho_i \leq 1\,,$$

where $(V)_{ii}$ and its inverse $(V^{-1})_{ii}$ are the diagonal elements of the covariance matrix defined by:

$$(V)_{ij} = E[(x_i - E[x_i])(x_j - E[x_j])] = E[x_i x_j] - E[x_i]E[x_j]\,.$$

Here $E[x_i] = \int x_i f(x_i) \mathrm{d}x_i$ is the expectation value of x_i with pdf f. For the discussion of the correlations in Chapter 6, we also introduce here the correlation matrix:

$$\rho_{ij} = \frac{(V)_{ij}}{\sigma(x_i)\sigma(x_j)}\,, \quad -1 \leq \rho_{ij} \leq 1\,,$$

where the standard deviation $\sigma(x_i)$ is given in terms of the variance $V[x_i] = (V)_{ii} = \sigma^2(x_i)$.

B.1 Results with the Nominal Parameters

Table B.1: Fit outputs for $B_s \to J/\psi\phi$ with the nominal parameters.

$B_s \to J/\psi\phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.037 ± 0.001	0.023	0.027	0.17 ± 0.06	0.89 ± 0.05	0.13
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.1532 ± 0.0006	0.0092	0.0079	0.41 ± 0.08	1.20 ± 0.06	0.22
τ_s [ps]	1.4820 ± 0.0002	0.0039	0.0031	10.35 ± 0.09	1.35 ± 0.06	0.24
ΔM_s [ps^{-1}]	17.500 ± 0.001	0.008	0.008	-0.01 ± 0.07	1.01 ± 0.04	0.06
ω_{tag} [$-$]	0.3295 ± 0.0002	0.0036	0.0040	-0.07 ± 0.06	0.86 ± 0.04	0.16
R_T [$-$]	0.1957 ± 0.0003	0.0040	0.0044	-0.96 ± 0.06	0.93 ± 0.04	0.23

Table B.2: Fit outputs for $B_s \to \eta_c\phi$ with the nominal parameters.

$B_s \to \eta_c\phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.042 ± 0.007	0.108	0.115	-0.04 ± 0.06	0.90 ± 0.04	0.05
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.0353 ± 0.0007	0.0110	0.0182	-6.33 ± 0.04	0.68 ± 0.04	0.20
τ_s [ps]	1.5170 ± 0.0003	0.0053	0.0045	14.88 ± 0.08	1.27 ± 0.06	0.19
ΔM_s [ps^{-1}]	17.499 ± 0.000	0.007	0.007	-0.07 ± 0.07	1.09 ± 0.06	0.02
ω_{tag} [$-$]	0.3112 ± 0.0002	0.0036	0.0035	0.30 ± 0.07	0.94 ± 0.05	0.05

Table B.3: Fit outputs for $B_s \to J/\psi\eta(\gamma\gamma)$ with the nominal parameters.

$B_s \to J/\psi\eta(\gamma\gamma)$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.043 ± 0.007	0.109	0.113	-0.03 ± 0.07	0.97 ± 0.04	0.05
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.0576 ± 0.0007	0.0112	0.0159	-5.87 ± 0.04	0.66 ± 0.03	0.24
τ_s [ps]	1.5132 ± 0.0003	0.0049	0.0045	14.11 ± 0.08	1.24 ± 0.06	0.23
ΔM_s [ps^{-1}]	17.499 ± 0.001	0.009	0.009	-0.07 ± 0.07	1.07 ± 0.06	0.02
ω_{tag} [$-$]	0.3511 ± 0.0002	0.0036	0.0036	0.29 ± 0.07	1.00 ± 0.05	0.05

Table B.4: Fit outputs for $B_s \to J/\psi\eta(\pi^+\pi^-\pi^0)$ with the nominal parameters.

$B_s \to J/\psi\eta(\pi\pi\pi)$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.041 ± 0.009	0.142	0.146	0.04 ± 0.06	0.95 ± 0.04	0.04
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.0236 ± 0.0005	0.0078	0.0186	-6.79 ± 0.04	0.54 ± 0.02	0.14
τ_s [ps]	1.5182 ± 0.0003	0.0053	0.0045	15.27 ± 0.08	1.29 ± 0.07	0.13
ΔM_s [ps^{-1}]	17.500 ± 0.000	0.007	0.007	-0.01 ± 0.07	1.01 ± 0.06	0.02
ω_{tag} [$-$]	0.3012 ± 0.0002	0.0033	0.0036	0.35 ± 0.06	0.91 ± 0.05	0.04

Table B.5: Fit outputs for $B_s \to D_sD_s$ with the nominal parameters.

$B_s \to D_sD_s$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.039 ± 0.009	0.133	0.140	-0.02 ± 0.06	0.94 ± 0.05	0.05
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.0511 ± 0.0008	0.0123	0.0171	-5.82 ± 0.05	0.78 ± 0.03	0.26
τ_s [ps]	1.5142 ± 0.0004	0.0057	0.0045	14.22 ± 0.09	1.37 ± 0.07	0.25
ΔM_s [ps^{-1}]	17.498 ± 0.001	0.009	0.009	-0.20 ± 0.07	1.00 ± 0.05	0.02
ω_{tag} [$-$]	0.3410 ± 0.0002	0.0035	0.0038	0.23 ± 0.06	0.94 ± 0.04	0.04

B.2 Results with other Parameters

$\mathbf{B_s \to J/\psi\phi}$

Table B.6: Fit outputs for $B_s \to J/\psi\phi$ with the nominal parameters, but $\phi_s = -0.2$ rad.

$B_s \to J/\psi\phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.179 ± 0.002	0.026	0.027	0.76 ± 0.07	0.99 ± 0.06	0.21
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.1511 ± 0.0006	0.0091	0.0080	0.17 ± 0.08	1.17 ± 0.07	0.21
τ_s [ps]	1.4818 ± 0.0003	0.0039	0.0031	10.42 ± 0.09	1.35 ± 0.07	0.23
ΔM_s [ps^{-1}]	17.500 ± 0.001	0.009	0.008	-0.08 ± 0.09	1.15 ± 0.07	0.06
ω_{tag} [$-$]	0.3286 ± 0.0003	0.0042	0.0039	-0.34 ± 0.08	1.13 ± 0.06	0.19
R_T [$-$]	0.1957 ± 0.0003	0.0045	0.0044	-1.01 ± 0.07	1.04 ± 0.05	0.22

Table B.7: Fit outputs for $B_s \to J/\psi\phi$ with the nominal parameters, but $\phi_s = -0.79$ rad.

$B_s \to J/\psi\phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.686 ± 0.002	0.035	0.036	2.93 ± 0.07	1.12 ± 0.05	0.47
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.1322 ± 0.0006	0.0100	0.0097	-1.76 ± 0.07	1.01 ± 0.04	0.16
τ_s [ps]	1.4848 ± 0.0003	0.0040	0.0032	10.99 ± 0.09	1.38 ± 0.07	0.32
ΔM_s [ps^{-1}]	17.500 ± 0.000	0.008	0.007	0.04 ± 0.07	1.04 ± 0.04	0.06
ω_{tag} [$-$]	0.3255 ± 0.0002	0.0036	0.0037	-1.26 ± 0.07	0.99 ± 0.06	0.38
R_T [$-$]	0.1953 ± 0.0003	0.0046	0.0042	-1.13 ± 0.08	1.15 ± 0.06	0.20

Table B.8: Fit outputs for $B_s \to J/\psi\phi$ with the nominal parameters, but $\Delta\Gamma_s/\Gamma_s = 0.05$.

$B_s \to J/\psi\phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.036 ± 0.002	0.027	0.031	0.20 ± 0.06	0.91 ± 0.04	0.17
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.0775 ± 0.0006	0.0090	0.0089	3.09 ± 0.07	1.07 ± 0.06	0.42
τ_s [ps]	1.4846 ± 0.0003	0.0042	0.0033	10.47 ± 0.10	1.49 ± 0.07	0.42
ΔM_s [ps^{-1}]	17.500 ± 0.001	0.008	0.008	-0.03 ± 0.07	1.01 ± 0.04	0.07
ω_{tag} [$-$]	0.3291 ± 0.0002	0.0036	0.0041	-0.26 ± 0.06	0.85 ± 0.05	0.20
R_T [$-$]	0.1987 ± 0.0003	0.0045	0.0045	-0.42 ± 0.06	0.94 ± 0.04	0.21

Table B.9: Fit outputs for $B_s \to J/\psi\phi$ with the nominal parameters, but $\Delta\Gamma_s/\Gamma_s = 0.25$.

$B_s \to J/\psi\phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.036 ± 0.001	0.020	0.024	0.24 ± 0.06	0.83 ± 0.03	0.15
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.2378 ± 0.0005	0.0080	0.0070	1.18 ± 0.07	-1.74 ± 0.08	0.14
τ_s [ps]	1.4839 ± 0.0003	0.0040	0.0031	10.76 ± 0.09	1.32 ± 0.06	0.25
ΔM_s [ps^{-1}]	17.500 ± 0.001	0.009	0.009	-0.02 ± 0.07	1.01 ± 0.05	0.07
ω_{tag} [$-$]	0.3297 ± 0.0003	0.0041	0.0041	-0.07 ± 0.07	1.03 ± 0.05	0.20
R_T [$-$]	0.1922 ± 0.0003	0.0042	0.0043	-1.92 ± 0.06	0.89 ± 0.05	0.29

Table B.10: Fit outputs for $B_s \to J/\psi\phi$ with the nominal parameters, but $R_T = 0$.

$B_s \to J/\psi\phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.039 ± 0.001	0.017	0.018	0.09 ± 0.07	0.99 ± 0.05	0.17
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.1664 ± 0.0005	0.0071	0.0063	2.64 ± 0.09	1.25 ± 0.07	0.47
τ_s [ps]	1.4856 ± 0.0003	0.0040	0.0033	10.69 ± 0.10	1.33 ± 0.08	0.44
ΔM_s [ps^{-1}]	17.500 ± 0.001	0.008	0.008	-0.01 ± 0.07	1.01 ± 0.04	0.06
ω_{tag} [$-$]	0.3292 ± 0.0003	0.0038	0.0042	-0.16 ± 0.07	0.94 ± 0.05	0.20
R_T [$-$]	0.0015 ± 0.0001	0.0011	0.0010	0.35 (mean)	1.09 (rms)	0.23

Table B.11: Fit outputs for $B_s \to J/\psi\phi$ with the nominal parameters, but $R_T = 0.15$.

$B_s \to J/\psi\phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.035 ± 0.001	0.022	0.024	0.25 ± 0.07	0.99 ± 0.06	0.12
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.1570 ± 0.0006	0.0087	0.0074	0.96 ± 0.09	1.24 ± 0.07	0.28
τ_s [ps]	1.4825 ± 0.0003	0.0039	0.0031	10.24 ± 0.08	1.24 ± 0.06	0.30
ΔM_s [ps^{-1}]	17.500 ± 0.001	0.009	0.008	0.04 ± 0.07	1.06 ± 0.05	0.06
ω_{tag} [$-$]	0.3291 ± 0.0003	0.0039	0.0040	-0.24 ± 0.07	1.05 ± 0.05	0.16
R_T [$-$]	0.1460 ± 0.0003	0.0044	0.0044	-0.98 ± 0.08	1.08 ± 0.06	0.24

Table B.12: Fit outputs for $B_s \to J/\psi\phi$ with the nominal parameters, but $R_T = 0.25$.

$B_s \to J/\psi\phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.035 ± 0.002	0.026	0.032	0.25 ± 0.06	0.88 ± 0.04	0.16
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.1458 ± 0.0006	0.0086	0.0085	-0.45 ± 0.07	1.04 ± 0.05	0.17
τ_s [ps]	1.4818 ± 0.0003	0.0041	0.0031	10.41 ± 0.09	1.37 ± 0.07	0.19
ΔM_s [ps^{-1}]	17.499 ± 0.001	0.008	0.008	-0.09 ± 0.07	1.03 ± 0.05	0.06
ω_{tag} [$-$]	0.3297 ± 0.0003	0.0040	0.0040	-0.09 ± 0.08	1.06 ± 0.06	0.19
R_T [$-$]	0.2460 ± 0.0003	0.0044	0.0044	-0.90 ± 0.07	1.05 ± 0.05	0.23

Table B.13: Fit outputs for $B_s \to J/\psi\phi$ with the nominal parameters, but $R_T = 0.5$.

$B_s \to J/\psi\phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.033 ± 0.003	0.047	0.057	0.10 ± 0.06	0.88 ± 0.04	0.12
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.0931 ± 0.0006	0.0098	0.0121	-4.73 ± 0.06	0.86 ± 0.04	0.44
τ_s [ps]	1.4885 ± 0.0002	0.0036	0.0033	11.64 ± 0.09	1.33 ± 0.06	0.42
ΔM_s [ps^{-1}]	17.500 ± 0.001	0.009	0.008	-0.04 ± 0.07	0.97 ± 0.05	0.05
ω_{tag} [$-$]	0.3308 ± 0.0002	0.0035	0.0039	0.21 ± 0.06	0.91 ± 0.04	0.13
R_T [$-$]	0.4983 ± 0.0003	0.0040	0.0041	-0.43 ± 0.07	0.98 ± 0.06	0.18

$\mathbf{B_s \to \eta_c \phi}$

Table B.14: Fit outputs for $B_s \to \eta_c \phi$ with the nominal parameters, but $\phi_s = -0.2$ rad.

$B_s \to \eta_c \phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.214 ± 0.007	0.110	0.118	-0.13 ± 0.06	0.95 ± 0.05	0.08
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.0344 ± 0.0007	0.0103	0.0183	-6.35 ± 0.04	0.69 ± 0.03	0.21
τ_s [ps]	1.5168 ± 0.0003	0.0052	0.0045	14.80 ± 0.09	1.29 ± 0.08	0.19
ΔM_s [ps^{-1}]	17.500 ± 0.000	0.007	0.007	0.02 ± 0.08	1.14 ± 0.05	0.02
ω_{tag} [$-$]	0.3112 ± 0.0002	0.0038	0.0035	0.34 ± 0.07	1.08 ± 0.05	0.06

Table B.15: Fit outputs for $B_s \to \eta_c \phi$ with the nominal parameters, but $\phi_s = -0.79$ rad.

$B_s \to \eta_c \phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.851 ± 0.011	0.173	0.178	-0.16 ± 0.06	0.86 ± 0.05	0.23
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.0173 ± 0.0005	0.0081	0.0195	-6.81 ± 0.05	0.67 ± 0.03	0.25
τ_s [ps]	1.5178 ± 0.0004	0.0053	0.0045	15.28 ± 0.09	1.32 ± 0.06	0.13
ΔM_s [ps^{-1}]	17.500 ± 0.000	0.007	0.007	-0.02 ± 0.07	0.98 ± 0.05	0.02
ω_{tag} [$-$]	0.3112 ± 0.0002	0.0034	0.0035	0.33 ± 0.07	1.03 ± 0.06	0.12

Table B.16: Fit outputs for $B_s \to \eta_c\phi$ with the nominal parameters, but $\Delta\Gamma_s/\Gamma_s = 0.05$.

$B_s \to \eta_c\phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.027 ± 0.007	0.113	0.121	0.14 ± 0.06	0.94 ± 0.07	0.05
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.0088 ± 0.0005	0.0085	0.0170	-2.42 ± 0.04	0.54 ± 0.03	0.07
τ_s [ps]	1.5015 ± 0.0004	0.0056	0.0044	11.73 ± 0.09	1.40 ± 0.07	0.06
ΔM_s [ps^{-1}]	17.499 ± 0.000	0.007	0.007	-0.04 ± 0.07	1.05 ± 0.06	0.02
ω_{tag} [$-$]	0.3096 ± 0.0002	0.0037	0.0035	-0.08 ± 0.07	1.07 ± 0.05	0.04

Table B.17: Fit outputs for $B_s \to \eta_c\phi$ with the nominal parameters, but $\Delta\Gamma_s/\Gamma_s = 0.25$.

$B_s \to \eta_c\phi$	Fit output		Fit error	Pull		G.C.
Parameters	Mean	rms	Mean	Mean [$-$]	Σ [$-$]	Mean
ϕ_s [rad]	-0.040 ± 0.006	0.101	0.104	-0.04 ± 0.06	0.98 ± 0.04	0.04
$\Delta\Gamma_s/\Gamma_s$ [$-$]	0.0824 ± 0.0009	0.0137	0.0196	-8.55 ± 0.05	0.74 ± 0.04	0.51
τ_s [ps]	1.5460 ± 0.0003	0.0055	0.0052	18.39 ± 0.11	1.58 ± 0.08	0.5
ΔM_s [ps^{-1}]	17.500 ± 0.000	0.007	0.007	0.05 ± 0.06	0.94 ± 0.05	0.02
ω_{tag} [$-$]	0.3142 ± 0.0002	0.0038	0.0035	1.19 ± 0.07	1.09 ± 0.05	0.08

Bibliography

[1] EMMY NOETHER. Invariante Variationsprobleme. *Nachr. v. d. Ges. d. Wiss. zu Göttingen*, 2:235–237, 1918.

[2] GERHART LÜDERS. On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories. *Dan. Mat. Phys. Medd.*, 28:5, 1954.

[3] WOLFGANG PAULI. Niels Bohr and the Development of Physics. McGraw-Hill, New York, 1955.

[4] GERHART LÜDERS. Proof of the TCP theorem. *Ann.Phys.*, 2:1–15, 1957.

[5] C. S. WU ET AL. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.*, 105:1413–1415, 1957.

[6] J. H. CHRISTENSON ET AL. Evidence for the 2 π Decay of the K_2^0 Meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[7] B. AUBERT AND OTHERS [BABAR COLLABORATION]. Observation of CP Violation in the B^0 Meson System. *Phys. Rev. Lett.*, 87, 2001, 091801.

[8] K. ABE AND OTHERS [BELLE COLLABORATION]. Observation of Large CP Violation in the Neutral B Meson System. *Phys. Rev. Lett.*, 87, 2001, 091802.

[9] MAKOTO KOBAYASHI AND TOSHIHIDE MASKAWA. CP-Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[10] NICOLA CABIBBO. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531–533, 1963.

[11] ANDREI. D. SAKHAROV. Violation of CP invariance, C asymmetry, and baryon asymmetry of the Universe. *JETP Lett.*, 5:24–27, 1967.

[12] VALERY. A. RUBAKOV AND MIKHAIL E. SHAPOSHNIKOV. Electroweak baryon number non-conservation in the early Universe and in high energy collisions. *Phys. Usp.*, 39:461–502, 1996.

[13] V. ABAZOV AND OTHERS [DØ COLLABORATION]. First Direct Two-Sided Bound on the B_s^0 Oscillation Frequency. 2006, hep-ex/0603029.

[14] G. GÓMEZ-CEBALLOS [CDF COLLABORATION]. Measurement of the B_s^0–$\overline{B}_s^0$ Oscilation Frequency, Talk at the 2006 FPCP. 2006, `http://fpcp2006.triumf.ca`.

[15] PATRICIA BALL AND ROBERT FLEISCHER. Probing New Physics through B Mixing: Status, Benchmarks and Prospects. 2006, hep-ph/0604249.

[16] LUIS FERNÁNDEZ. The $B_s \to \eta_c(4h)\phi(K^+K^-)$ Reconstruction at LHCb. CERN-LHCb/2003-101, IPHE-2003-015, 2003.

[17] LHCb collaboration. LHCb Technical Design Report: Reoptimization detector design and performance. CERN-LHCC/ 2003-030, 2003.

[18] LUIS FERNÁNDEZ AND THOMAS SCHIETINGER. Decision Tracks in Minimum-Bias Events Selected by the Level-1 Trigger. CERN-LHCb-2004/047, LPHE-2004-006, 2004.

[19] GUSTAVO C. BRANCO, LUÍS LAVOURA, AND JOÃO P. SILVA. *CP Violation*. Oxford University Press, New York, 1999.

[20] ROBERT FLEISCHER. Flavour physics and CP violation. 2004, hep-ph/0405091.

[21] K. ANIKEEV ET AL. B Physics at the Tevatron: Run II and Beyond. 2002, hep-ph/0201071.

[22] ISARD DUNIETZ, ROBERT FLEISCHER, AND ULRICH NIERSTE. In Pursuit of New Physics with B_s Decays. *Phys. Rev.*, D63:114015, 2001, hep-ph/0012219.

[23] SHELDON L. GLASHOW. Partial-Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

[24] ABDUS SALAM. Weak and Electromagnetic Interactions. In N. SVARTHOLM, editor, *Elementary Particle Theory*. Almqvist and Wiksell, Stockholm, 1968.

[25] STEVEN WEINBERG. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[26] LINCOLN WOLFENSTEIN. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945–1947, 1983.

[27] S. EIDELMAN ET AL. Review of Particle Physics. *Physics Letters C*, 592:1+, 2004 and 2005-2006 partial updates for edition 2006, `http://pdg.lbl.gov`.

[28] LING-LIE CHAU AND WAI-YEE KEUNG. Comments on the Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 53:1802–1805, 1984.

[29] A. J. BURAS, M. E. LAUTENBACHER, AND G. OSTERMAIER. Waiting for the Top Quark Mass, $K^+ \to \pi^+\nu\bar{\nu}$, $B_s^0 - \bar{B}_s^0$ Mixing and CP Asymmetries in B-Decays. *Phys. Rev.*, D50:3433–3446, 1994, hep-ph/9403384.

[30] CECILIA JARLSKOG. Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Nonconservation. *Phys. Rev. Lett.*, 55:1039–1042, 1985.

[31] R. ALEKSAN, B. KAYSER, AND D. LONDON. Determining the Quark Mixing Matrix from CP-Violating Asymmetries. *Phys. Rev. Lett.*, 73:18–20, 1994, hep-ph/9403341.

[32] CECILIA JARLSKOG AND RAYMOND STORA. Unitarity polygons and CP violation areas and phases in the standard electroweak model. *Phys. Lett.*, B208:268–274, 1988.

[33] CKMFITTER GROUP (J. CHARLES ET AL.). CP Violation and the CKM Matrix: Assessing the Impact of the Asymmetric B Factories. *Eur.Phys.J.*, C41:1–131, 2005, hep-ph/0406184. The updated results and plots are available at the address: `http://ckmfitter.in2p3.fr`.

[34] OLIVIER SCHNEIDER. B^0–$\bar{B}^0$ mixing. Review in [27], hep-ex/0405012.

[35] VICTOR F. WEISSKOPF AND EUGENE P. WIGNER. Calculation of the natural brightness of spectral lines on the basis of Dirac's theory. *Z. Phys.*, 63:54–73, 1930.

[36] JUN JOHN SAKURAI. *Modern Quantum Mechanics*. Revised edition, Addison-Wesley, 1994.

[37] ANDRZEJ J. BURAS AND ROBERT FLEISCHER. Quark mixing, CP violation and rare decays after the top quark discovery. *Adv. Ser. Direct. High Energy Phys.*, 15:65–238, 1998, hep-ph/9704376.

[38] ROBERT FLEISCHER. Extracting γ from $B_{s(d)} \to J/\psi K_S$ and $B_{d(s)} \to D^+_{d(s)} D^-_{d(s)}$. *Eur. Phys. J.*, C10:299–306, 1999, hep-ph/9903455.

[39] JONATHAN L. ROSNER. Determination of pseudoscalar charmed meson decay constants from B-meson decays. *Phys. Rev.*, D42:3732–3740, 1990.

[40] AMOL S. DIGHE, ISARD DUNIETZ, HARRY J. LIPKIN, AND JONATHAN L. ROSNER. Angular distributions and lifetime differences in $B_s \to \mathrm{J}/\psi\phi$ decays. *Phys. Lett.*, B369:144–150, 1996, hep-ph/9511363.

[41] AMOL S. DIGHE, ISARD DUNIETZ, AND ROBERT FLEISCHER. Extracting CKM Phases and B_s–$\overline{B}_s$ Mixing Parameters from Angular Distributions of Non-Leptonic B Decays. *Eur. Phys. J.*, C6:647–662, 1999, hep-ph/9804253.

[42] ISARD DUNIETZ, HELEN QUINN, ART SNYDER, WALTER TOKI, AND HARRY J. LIPKIN. How to extract CP-violating asymmetries from angular correlations. *Phys. Rev.*, D43:2193–2208, 1991.

[43] CDF, D. ACOSTA ET AL. Measurement of the Lifetime Difference Between B_s Mass Eigenstates. *Phys. Rev. Lett.*, 94:101803, 2005, hep-ex/0412057.

[44] DØ, V. M. ABAZOV ET AL. Measurement of the Lifetime Difference in the B^0_s System. *Phys. Rev. Lett.*, 95:171801, 2005, hep-ex/0507084.

[45] MARTIN BENEKE AND ALEXANDER LENZ. Lifetime difference of B_s mesons: Theory status. *J. Phys.*, G27:1219–1224, 2001, hep-ph/0012222.

[46] ALEXANDER LENZ. Decay Rate Difference in the Neutral B-System: $\Delta\Gamma_{B_s}$ and $\Delta\Gamma_{B_d}$. 2004, hep-ph/0412007.

[47] E. BARBERIO AND OTHERS [HFAG]. Averages of b-hadron Properties at the End of 2005. 2006, online update at `http://www.slac.stanford.edu/xorg/hfag`, hep-ex/0603003.

[48] M. CIUCHINI ET AL. 2000 CKM-Triangle Analysis A Critical Review with Updated Experimental Inputs and Theoretical Parameters. *JHEP*, 0107:013, 2001, hep-ph/0012308.

[49] Unitarity Triangle fit group. The recent results and plots are available on the world wide web. `http://utfit.roma1.infn.it`.

[50] ANDRZEJ J. BURAS ET AL. Universal Unitarity Triangle and Physics Beyond the Standard Model. *Phys. Lett.*, B500:161–167, 2001, hep-ph/0007085.

[51] THE UTFIT COLLABORATION. THE UTFIT COLLABORATION REPORT ON THE UNITARITY TRIANGLE BEYOND THE STANDARD MODEL: SPRING 2006. 2006, hep-ph/0605213.

[52] ZOLTAN LIGETI, MICHELE PAPUCCI, AND GILAD PEREZ. Implications of the measurement of the B_s^0–$\bar{B}_s^0$ mass difference. 2006, hep-ph/0604112.

[53] CERN public web pages. `http://public.web.cern.ch/`.

[54] CERN document server. `http://cdsweb.cern.ch/`.

[55] LHCb collaboration. LHCb Technical Proposal. CERN-LHCC/ 98-4, 1998.

[56] TORBJÖRN SJÖSTRAND ET AL. High-energy-physics event generation with PYTHIA 6.1. *Comput. Phys. Commun.*, 135:238–259, 2001, hep-ph/0010017.

[57] LHCb collaboration. LHCb - Parameters of DC04 Monte Carlo productions. `https://uimon.cern.ch/ twiki/bin/view/LHCb/SettingsDc04`.

[58] P. NASON ET AL. Bottom production. In G. ALTARELLI AND M.L. MANGANO, editors, *Proceedings of the 1999 Workshop on Standard Model Physics (and more) at the LHC*, volume CERN-2000-004. CERN Organisation européenne pour la recherche nucléaire, 1999, hep-ph/0003142.

[59] LHCb collaboration. LHCb Magnet Technical Design Report. CERN-LHCC/ 2000-007, 2000.

[60] LHCb collaboration. LHCb VELO Technical Design Report. CERN-LHCC/ 2001-011, 2001.

[61] J. GASSNER, M. NEEDHAM, AND O. STEINKAMP. Layout and Expected Performance of the LHCb TT Station. CERN-LHCb/2003-140, 2003.

[62] LHCb collaboration. LHCb Inner Tracker Technical Design Report. CERN-LHCC/ 2002-029, 2002.

[63] LHCb collaboration. LHCb Outer Tracker Technical Design Report. CERN-LHCC/ 2001-024, 2001.

[64] LHCb collaboration. LHCb Calorimeters Technical Design Report. CERN-LHCC/ 2000-036, 2000.

[65] LHCb collaboration. LHCb Trigger System Technical Design Report. CERN-LHCC/ 2003-031, 2003.

[66] LHCb collaboration. LHCb Computing Technical Design Report. CERN-LHCC/ 2005-019, 2005.

[67] G. BARRAND ET AL. GAUDI- A software architecture and framework for building hep data processing applications. *Comput. Phys. Commun.*, 140:45, 2001.

[68] GAUDI project. `http://proj-gaudi.web.cern.ch/proj-gaudi/welcome.html`.

[69] PATRICK KOPPENBURG AND LUIS FERNÁNDEZ. DAVINCI for Busy People. CERN-LHCb-2005/016, LPHE-2005-012, 2005.

[70] PATRICK KOPPENBURG AND LUIS FERNÁNDEZ. HLT Exclusive Selections Design and Implementation. CERN-LHCb-2005/015, LPHE-2005-011, 2005.

[71] D.J. LANGE. The EVTGEN particle decay simulation package. *Nucl. Inst. Meth.*, A 462:152–155, 2001.

[72] LHCb collaboration. EVTGEN. `http://lhcb-comp.web.cern.ch/lhcb-comp/Simulation/evtgen.htm`.

[73] LHCb collaboration. GAUSS - The LHCb Simulation Program. `http://lhcb-comp.web.cern.ch/lhcb-comp/Simulation`.

[74] GEANT 4 collaboration, S. AGOSTINELLI ET AL. GEANT 4 - a simulation toolkit. *Nucl. Inst. Meth.*, A 140:250, 2003.

[75] LHCb collaboration. BOOLE - The LHCb digitization program. `http://lhcb-comp.web.cern.ch/lhcb-comp/Digitization`.

[76] LHCb collaboration. BRUNEL - The LHCb reconstruction program. `http://lhcb-comp.web.cern.ch/lhcb-comp/Reconstruction`.

[77] LHCb collaboration. DAVINCI - The LHCb analysis program. `http://lhcb-comp.web.cern.ch/lhcb-comp/Analysis`.

[78] LUIS FERNÁNDEZ AND PATRICK KOPPENBURG. Exclusive HLT Performance. CERN-LHCb-2005/047, LPHE-2005-015, 2005.

[79] IVAN (VANYA) BELYAEV. LoKi: Smart & Friendly C++ Physics Analysis Toolkit. CERN-LHCb/2004-023, 2004.

[80] LHCb collaboration. Data Visualization. `http://lhcb-comp.web.cern.ch/lhcb-comp/Frameworks/Visualization`.

[81] OLIVIER CALLOT. A new implementation of the Relations: The Linker objects. CERN-LHCb/2004-007, 2004.

[82] IVAN BELYAEV. Relations - Generic external relations in GAUDI. CERN-LHCb/2005-005, 2005.

[83] H. DIJKSTRA, N. TUNING, AND N. BROOK. Some Remarks on Systematics Effects of the Trigger and Event Generator Studies. CERN-LHCb/2003-157, 2003.

[84] JEROEN VAN TILBURG. Track simulation and reconstruction in LHCb. CERN-THESIS/2005-040, 2005.

[85] OLIVIER CALLOT. VELO tracking for the High Level Trigger. CERN-LHCb/2003-027, 2003.

[86] MAURICE BENAYOUN AND OLIVIER CALLOT. The Forward Tracking, an Optical Model Method. CERN-LHCb/2002-008, 2002.

[87] OLIVIER CALLOT. Online Pattern Recognition. CERN-LHCb/2004-094, 2004.

[88] ROGER FORTY. Track seeding. CERN-LHCb/2001-109, 2001.

[89] LHCb collaboration. LHCb flavour tagging working group - Physics Performances. `https://uimon.cern.ch/twiki/bin/view/LHCb/FlavourTagging`.

[90] M. CALVI, O. DORMOND, AND M. MUSY. LHCb flavour tagging performance. CERN-LHCb-2003/115, 2003.

[91] G. CORTI, L. FERNÁNDEZ, P. ROBBE, AND O. SCHNEIDER. Monte Carlo Event Type Definition Rules. CERN-LHCb/2005-034, 2005.

[92] JOSÉ-ANGEL HERNANDO ET AL. HLT-generic package: `Hlt/HltGeneric v2r8`. LHCb Software, 2006.

[93] MARIUSZ WITEK. Improved vertex finding for HLT. T-Rec Meeting, CERN, December 12, 2005.

[94] ALESSIA SATTA. Muon identification in the LHCb High Level Trigger. CERN-LHCb/2005-071, 2005.

[95] LUIS FERNÁNDEZ. New Velo-TT tracking implementation. T-Rec Meeting, CERN, June 13, 2005.

[96] HUGO RUIZ. Parameterization of track uncertainties for the HLT. CERN-LHCb/2005-012, 2005.

[97] HUGO RUIZ. Fast tools for vertexing and geometry calculations for the HLT. CERN-LHCb/2005-013, 2005.

[98] HANS DIJKSTRA. Isolation of HLT signal candidates. T-Rec Meeting, CERN, December 12, 2005.

[99] LUIS FERNÁNDEZ. HLT rates: old HLT PV vs new HLT PV. T-Rec Meeting, CERN, December 19, 2005.

[100] MARIUSZ WITEK. VELO-TT matching and momentum determination at Level-1 trigger. CERN-LHCb/2003-060, 2003.

[101] GARY J. FELDMAN AND ROBERT D. COUSINS. A Unified approach to the classical statistical analysis of small signals. *Phys. Rev.*, D57:3873–3889, 1998, physics/9711021.

[102] GERHARD RAVEN. Selection of $B_s \to J/\psi\phi$ and $B^+ \to J/\psi K^+$. CERN-LHCb/2003-118, 2003.

[103] GERHARD RAVEN. Sensitivity studies of χ and $\Delta\Gamma$ with $B_s \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$. CERN-LHCb/2003-119, 2003.

[104] WOUTER VERKERKE AND DAVID KIRKBY. The ROOFIT Toolkit for Data Modelling. `http://roofit.sourceforge.net`.

[105] RENÉ BRUN AND FONS RADEMAKERS. ROOT: An Object-Oriented Data Analysis Framework. `http://root.cern.ch`.

[106] FRED JAMES. MINUIT reference Manual, Function minimization and error analysis. CERN Program Library Long Writeup D506, `http://wwwinfo.cern.ch/asdoc/minuit`.

[107] JÉRÉMIE BOREL. Results of a DC04 updated $B_s \to D_s\pi$ selection. Private communication, 2006.

[108] DAVID BORRAS AND STÉPHANE MONTEIL. Results of a DC04 $B_s \to D_sD_s$ selection. Private communication, 2006.

[109] BENJAMIN CARRON. $B_s^0 \to J/\psi\eta$ decays and sensitivity to the B_s^0 mixing phase at LHCb. CERN-THESIS/2005-059, 2005.

Luis Fernández

Laboratory for High-Energy Physics (LPHE)
Swiss Federal Institute of Technology (EPFL)
CH-1015 Lausanne-Dorigny

Born on February 16th, 1977
Countries of citizenship: Spain, Switzerland

Education and degrees

August 2006	Ph. D. in High-Energy Physics on *Exclusive Trigger Selections and Sensitivity to the B_s–$\overline{B}_s$ Mixing Phase at LHCb* , Swiss Federal Institute of Technology (EPFL), Switzerland
2004	CERN-JINR *2004 European School of High-Energy Physics*, Sant Feliu de Guíxols, Spain
2001 – 2006	Ph. D. studies in High-Energy Physics at the IPHE (UNIL) and the LPHE (EPFL), Switzerland
March 2001	Physics diploma thesis on *Scalar Field Propagator in a Brane World*, Institute for Theoretical Physics (IPT), University of Lausanne (UNIL), Switzerland
1998 – 1999	Exchange student, Uppsala Universitet, Sverige
1996 – 2001	Physics studies, University of Lausanne (UNIL), Switzerland
1995 – 1996	"Cours de Mathématiques Spéciales" (CMS), EPFL, Switzerland