UNIVERSITÀ DEGLI STUDI DI TRIESTE

Facoltà di Scienze Matematiche, Fisiche e Naturali
Corso di Laurea Magistrale in Fisica

Tesi di Laurea

MEASUREMENT OF B_S LIFETIME, DECAY WIDTH DIFFERENCE AND POLARIZATION AMPLITUDES OF THE $B_S \to J/\psi\phi$ DECAYS AT CDFII

Relatrice:
Dott.ssa ANNA MARIA ZANETTI

Correlatore:
Dott. MIRCO DORIGO

Laureanda:
LUCIA GRILLO

ANNO ACCADEMICO 2010-2011

Abstract

In this thesis the update of the measurement of the B_s^0 lifetime, the decay width difference between its heavy and light mass eigenstates and the polarization amplitudes of $B_s^0 \to J/\psi\phi$ decays of the B_s^0 meson is presented. About 9600 $B_s^0 \to J/\psi\phi$ decays have been reconstructed in the final state $[\mu^+\mu^-][K^+K^-]$ using a dataset of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV corresponding to $8.4\ \mathrm{fb}^{-1}$ integrated luminosity collected by the CDFII detector at the Fermilab Tevatron collider. The results are extracted from an analysis of the angular distributions of muons and kaons as a function of the decay time through an unbinned maximum likelihood fit which exploits identification of the quark content (b or $\bar{b}$) of the strange bottom meson at the time of production. Assuming the Standard Model prediction for the size of CP violation occurring in the B_s mixing, the estimated lifetime, decay width difference, polarization amplitudes and strong phase of the perpendicular amplitude are:

$$\begin{aligned}
\tau(B_s^0) &= 1.527 \pm 0.021(\text{stat.})\text{ps}, \\
\Delta\Gamma &= 0.063 \pm 0.029(\text{stat.})\text{ps}^{-1}, \\
|A_{\|}(0)|^2 &= 0.233 \pm 0.014(\text{stat.}), \\
|A_0(0)|^2 &= 0.514 \pm 0.012(\text{stat.}), \\
\delta_{\perp} &= 2.95 \pm 0.61(\text{stat.}).
\end{aligned}$$

These results are competitive with the world's best measurements. The analysis is part of the update of the measurement of the CP-violating mixing phase β_s by the CDF collaboration with the full Run II sample, and represents a necessary preliminary step to validate the tools used in the β_s estimation.

As of this writing, the whole Run II dataset ($\sim$10 fb^{-1}) has become available. The update of the results on the final CDF dataset is now in progress. Since all the features of the analysis have been well validated in this work, the final measurement is expected to be finalized for the Winter 2012 Conferences.

Sommario

Questa tesi presenta l'aggiornamento della misura di vita media del mesone B_s^0, della differenza tra le ampiezze di decadimento dei suoi autostati di massa e delle ampiezze di polarizzazione dei decadimenti $B_s^0 \to J/\psi\phi$. Circa 9600 decadimenti $B_s^0 \to J/\psi\phi \to [\mu^+\mu^-][K^+K^-]$ sono stati ricostruiti in un campione di collisioni $p\bar{p}$ all'energia $\sqrt{s} = 1.96$ TeV acquisito con il rivelatore CDFII e corrispondente alla luminositá integrata di 8.4 fb^{-1}. I risultati sono stati ottenuti dall'analisi delle distribuzioni angolari dei muoni e kaoni in funzione del tempo di decadimento del mesone B_s^0, mediante lo sviluppo di un fit di massima verosimiglianza, che si avvale degli algoritmi di riconoscimento del sapore dei mesoni B prodotti nelle interazioni $p\bar{p}$. Assumendo la predizione del modello standard per la violazione di CP nelle oscillazioni del B_s^0, i valori misurati della vita media, della differenza tra le ampiezze di decadimento, delle ampiezze di polarizzazione e della fase forte dell'ampiezza di polarizzazione perpendicolare, sono rispettivamente:

$$\begin{aligned}
\tau(B_s) &= 1.527 \pm 0.021(\text{stat.})\text{ps},\\
\Delta\Gamma &= 0.063 \pm 0.029(\text{stat.})\text{ps}^{-1},\\
|A_{\|}(0)|^2 &= 0.233 \pm 0.014(\text{stat.}),\\
|A_0(0)|^2 &= 0.514 \pm 0.012(\text{stat.}),\\
\delta_{\perp} &= 2.95 \pm 0.61(\text{stat.}).
\end{aligned}$$

Questi risultati presentano una risoluzione comparabile con le migliori misure effettuate per tali osservabili. L'analisi si inserisce nel contesto dell'aggiornamento delle misura della fase di violazione di CP nell'oscillazione del B_s^0 (β_s) che la collaborazione CDF sta affrontando effettuando su tutto il campione di dati acquisiti nel Run II. Questo lavoro costituisce un requisito essenziale per la validazione delle tecniche di fit che vengono utilizzate nella misura della fase β_s. Al momento della scrittura di questa tesi, l'intero campione ($\sim$10 fb^{-1}) di dati del Run II di CDF e‘ stato completato e reso pronto per le analisi. L'aggiornamento delle misure qui presentate e‘ dunque in corso utilizzando tutta la statistica disponibile. Dato che le componenti dell'analisi sono state testate e capite a fondo in questo lavoro di tesi, ci si aspetta di ottenere i nuovi risultati in breve tempo, e la loro approvazione ufficiale da parte della collaborazione e‘ attesa entro la fine dell'anno

in corso.

Contents

List of Figures

Introduction

Since the discovery of the simultaneous violation of charge and parity symmetries (CP violation) [1], CP violation has played a major role in the development of the Standard Model (SM) of particle physics. CP violation can be accommodated in the SM in the quark mixing process, but it demands the existence of at least three families of quarks. This led Kobayashi and Maskawa to the prediction of the third generation of quarks (b and t) before its experimental discovery. As a consequence of the extension of the Cabibbo mixing angle to the Cabibbo - Kobayashi - Maskawa (CKM) quark mixing matrix, in the SM all CP–violating effects are explained by a single irreducible complex phase of the CKM matrix. Therefore, CP violation represents a useful tool to probe New Physics (NP) beyond the SM by searching for processes that cannot be accounted for by the CKM mechanism. In this scenario, one of the most promising processes is the oscillation between B_s^0 and $\bar{B}_s^0$ mesons (B_s^0 mixing), where CP–violating effects are accurately predicted to be very small in the SM. In this case, the parameter of interest which accounts for CP violation is the phase β_s, that is expressed in terms of the elements of the CKM matrix as $\beta_s \equiv \arg\left(-V_{ts}V_{tb}^*/V_{cs}V_{cb}^*\right)$ and takes the value $\beta_s = 0.020 \pm 0.001$ in the SM [2, 3, 4]. A broad class of NP models introduces additional sources of CP violation resulting in a large value of β_s [5, 19]. The measurement of β_s is therefore a crucial test of the SM consistency.

Because of its clean experimental signature, the golden mode to perform such a test is the decay of the B_s^0 meson to the $J/\psi\phi$ final state: once the b quark content of the B_s^0 meson at its production is identified, the analysis of the time–evolution of $B_s^0 \to J/\psi\phi$ decays allows a direct measurement of β_s. Since $J/\psi\phi$ final state is common to both B_s^0 and $\bar{B}_s^0$, direct decays of the B_s^0 to $J/\psi\phi$ are possible as well as decays where the B_s^0 first oscillates to $\bar{B}_s^0$ and then decays. The CP violation in $B_s^0 \to J/\psi\phi$ decays arises from the interference of the two paths (with or without mixing) to the final state. In addition the $B_s^0 \to J/\psi\phi$ mode belongs to the class of decays where the final state consists of a pair of light spin–1 mesons. Three independent amplitudes govern such decays, corresponding to the allowed polarizations of the final state vector mesons: longitudinal polarization, and transverse polarization with spins parallel or perpendicular to each other. The first two states are CP–even, while the last one is CP–odd. This adds more complexity to the analysis which is based on the determination of the relative contribution of the amplitudes with definite CP-parity as a function of the decay time. Along with the estimation of β_s, such time–dependent analysis of $B_s^0 \to J/\psi\phi$ decays can provides the most precise measurements in a single channel of the B_s^0 lifetime and the decay width difference $\Delta\Gamma$ between the heavy and light B_s^0 mass eigenstates.

Analyses assuming the SM prediction for β_s to extract B_s^0 lifetime and $\Delta\Gamma$ along with polarization amplitudes were carried out as a necessary preliminary step prior to the measurement of β_s. While searches for NP usually focus on the β_s measurement, the experimental determination of $\Delta\Gamma$ is actually interesting by its own right. First of all, while a non–zero value of $\Delta\Gamma$ is clearly predicted in the SM [20], a firm experimental observation of a non-zero lifetime difference has not yet been given. Then, recent updates of the Tevatron's analyses [119, 97] and a new precise measurement by LHCb [40] suggest a value of β_s close to the SM value within 1σ. Although experimental uncertainties still leave room for physics beyond the SM, the strength of possible NP should be weaker than what could be expected from first observations [35]. If this is the case, it is potentially in disagreement with the CP–violating asymmetry of semileptonic decays of neutral B mesons (A_{SL}) recently seen by the DØ. The discrepancy of about 4 σ between the measured A_{SL} value and the SM expectation is very unlikely due to NP contributions from the B_d^0 sector, which is well constrained by B–factories measurements. Nevertheless, the large size of A_{SL} could be accounted for in models that allow NP effects in $\Delta\Gamma$ rather than in β_s. Roughly speaking, in some models it is possible to accommodate the A_{SL} discrepancy and the mild β_s tension properly tuning the value of $\Delta\Gamma$. Accurate measurements of $\Delta\Gamma$ could constrain the available parameter space and probe the reliability of either the SM prediction (assuming that the DØ measurement of A_{SL} is correct) or the DØ measurement (assuming β_s close to the SM expectation).

Concerning the polarization amplitudes, phenomenological models based on factorization, which provide predictions for B decays into light vector mesons, are not reliable in the case of $B_s^0 \to J/\psi\phi$, because of the comparable values of the J/ψ and B_s masses. However, a recent work which relies on the SU(3) symmetry derives the polarization amplitudes and the strong phases of $B_s^0 \to J/\psi\phi$ decays starting from the measured polarizations of $B^0 \to J/\psi K^\star$ [21]. Such a prediction gives the possibility to constrain the strong phases of $B_s \to J/\psi\phi$ decays in the β_s measurement, allowing for the resolution of its sign ambiguity. Indeed, previous measurements of $B_s^0 \to J/\psi\phi$ support SU(3) symmetry expectations.

The measurement of B_s^0 lifetime can test the predictions of Heavy Quark Expansion (HQE) models [127, 126]. HQE can predict the decay–width of B_s^0 mesons to final states common to B_s^0 and $\bar{B}_s^0$, as the case of $B_s^0 \to J/\psi\phi$. Knowing the lifetime of the B_d^0 meson, the measurement of B_s^0 lifetime provides an accurate way to verify the HQE framework as non SM effects are expected to be highly suppressed in lifetimes.

In this thesis, the update of the latest CDF measurements of the decay width difference $\Delta\Gamma$, the B_s^0 lifetime and the polarization amplitudes of $B_s^0 \to J/\psi\phi$ decays assuming the SM prediction for β_s are presented. The analysis has been carried out as part of the update of the analysis for the measurement of β_s by the CDF collaboration with the entire Run II data set, and constitutes the necessary preliminary step to validate the tools used in the estimation of β_s. The obtained results are competitive with the world's best measurements.

The thesis is organized as follow.

In **Chapter 1** some theoretical basics necessary to understand the measurement are briefly reported. The discussion is an introduction to the features of the decay mode considered in the measurement and it does not pretend to be exhaustive. Particular relevance has been given to the presentation of the $B_s^0 \to J/\psi\phi$ differential decay rates as a function of decay time and angles, since it represents the most important component of the likelihood function used in the analysis.

Chapter 2 describes the experimental apparatus, both the Tevatron collider and the CDF detector, focusing on the elements that are most relevant for this measurement. For this reason the tracking system and the muon detectors are described in greater details.

Chapter 3 presents the data sample used in the analysis along with the signal selection criteria. The latter mainly includes the trigger requirements (on–line selection) and the signal–to–background discriminator based on an Artificial Neural Network (off–line selection). Also explained in this chapter is the procedure used to identify the quark content of the mesons at the production, known as *flavor tagging*. The last section of the chapter presents the sample of simulated events used to compute the detector acceptance.

Chapter 4 introduces the unbinned maximum likelihood fit that is used in the measurement, and provides a detailed description of all the components of the likelihood function.

In **Chapter 5** the validation studies of the fit are presented and the fit results are discussed. The method to compute the systematic uncertainties is also reported.

The final part of the thesis is dedicated to the conclusions.

Chapter 1

Theoretical framework and analysis motivations

Contents

The purpose of this chapter is to provide some theoretical basics necessary in order to understand the object and the motivations of this measurement.
First of all it is important to recall some basics of the elementary particles and the fundamental forces that have been observed in nature and then of the SM of particle physics.

The SM is in fact the model that describe the fundamental interactions of particles including electroweak theory, that describes electro-magnetic and weak interactions, and the so called Quantum Cromo Dynamics (QCD), theory of the strong interactions (cfr. section 1.0.1).
Secondly, should be noticed that symmetries and conservation laws have played an important role in the development of the SM, and in order to understand this thesis is relevant to focus on the symmetry with respect to Charge-conjugation-Parity (CP) operator (cfr. section 1.0.2).
The following considerations will be about the weak interaction (cfr. section 1.0.3).
The violation of CP symmetry occurs in the interaction between quarks and the charged gauge boson $W^{\pm}$*, carrier of the weak force. The description of these interactions is provided in the SM by the CKM mechanism (cfr. section 1.0.4). In order to test the predictions of these model several processes, such as the production and decay of the neutral mesons K and B, can be exploited. Should be pointed out that parameters govering these interactions are not predicted by the SM , but it is possible to test the self-consistency of the model, and possibly measure the parameters of the theory from the data. While testing the self-consistency of the description, one can actually test whether there is physics beyond the SM mechanism (the so called NP).*
The study of the production, time-evolution and decay of the B_s *mesons yields a set of physical observables that are able to test the CKM description. In section 1.1 some concepts about the* B *mesons and their production at hadron colliders will be presented .The* B_s *meson, an unstable bound state of an anti-bottom quark and a strange quark (*$\bar{b}s$*) has been object of interest in particle physics research for at leas two reasons. The first relies in the fact that* B_s *mesons undergo oscillations between their matter and antimatter states (*B_s *and* $\bar{B}_s$*) at very high frequency (*$17.77 \pm 0.12\,\mathrm{ps}^{-1}$*), and the measurement of this frequency provides important constraints to the CKM model. The second reason is that it is possible to measure the degree of CP violation in* B_s *system,and verify whether it is consistent or not with the SM expectation (cfr. section 1.2).*
In this thesis the $B_s \to J/\psi\phi$ *decay mode has been taken into consideration because it represents the golden mode to measure the CP-violating phase of the* B_s *mixing process, due to the precise prediction in the SM and to the clean experimental signature that characterize this mode. Final remarks will be about the current experimental status of this analysis followed by purposes and strategy of this thesis.*

1.0.1. The Standard Model: particles and interactions

Particle physics is the study of the elementary constituents of the universe, their properties, and their interactions. Experiments in the field have so far revealed four fundamental forces, or interactions: gravity, electromagnetism, the weak interaction, and the strong interaction.
Theoretical insights and relative experimental proofs in this specific field, have been collected in the SM picture. More properly this is the name given to the electro-weak unified theory formulated by Glashow, Salam e Weinberg, whose meaning is then enlarged in order to include also the so called QCD. The SM is a so called Quantum Field Theory,

belonging to the class of gauge theories, providing a mathematical description of particles and interactions that incorporates quantum mechanics and special relativity by construction. According to this theoretical framework the interactions between fermions has been introduced by requiring the invariance of the Lagrangian with respect to an arbitrary transformation of the fields describing the particles according to the local symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. The first term of the direct product represents the symmetry of the QCD lagrangian while the other two terms describe the electro-weak sector.

The elementary particles known so far fall into two categories, defined by the interactions they experience: $quarks$, which are subject to all four, and $leptons$, which are not subject to the strong interaction. Quarks and leptons are fermions, particles with spin of 1/2. The interactions between them occur as the result of the exchange of bosonic (particle with intrinsic spin 1) mediating particles, called intermediate gauge bosons.
Known fermions are 12, divided in three families of leptons and three families of quarks as shown in Fig. 1.1. The overwhelming majority of the matter around us is made of just three of the fundamental particles listed in Fig. 1.1: up and down quarks, and electrons. The bosonic fields, associated to the symmetry group are 12: 8 different gluons (strong interaction carriers), and W^+, W^-, Z, γ for the electro-weak sector. Experiments have confirmed the existence of the $W^\pm$ and Z bosons, the photon, and the gluons, mediating respectively charged weak, neutral weak, electromagnetic, and strong interactions.
For each particle, there exists an anti-particle with identical properties except for opposite quantum numbers, e.g. same mass and intrinsic spin but opposite charge.
In order to give masses to the gauge bosons and fermions and preserve the renormalization in the SM a mechanism known as "spontaneous symmetry breaking" hes been introduced. In this way another bosonic field is added, the Higgs boson. Both gauge bosons and fermions acquire their masses through the interaction with this new particle.

An important difference between QCD and Electro-Weak sector relies in the fact that particles that the QCD deals with have never been observed singly, but only in bounds states of two (mesons) or three (baryons) quarks. The interaction between gluons provided by the non abelian summetry $SU(3)_C$, depends on the behavior of the strong coupling constant $\alpha_s(\mu)$. Where μ is the energy scale of the process. $\alpha_s(\mu)$ tends to zero as μ increases (that behavior is usually referred as $asymptotic\ freedom$) and increases with decreasing values of μ ($confinement$). The subsequent theoretical interpretation is that the $partons$ (elementary particles) are confined into the hadrons, bound states of more elementary particles.

Until now the SM has given good predictions of the observables measured by many experiments, and currently represents the best theoretical description of the elementary particle physics. Discrepancies between measurements and theoretical values are of order 1% or less in the electro-weak sector and of order 10% in QCD. For QCD this is a consequence of the $confinement$ of the partons in the hadrons, and the uncertainties arising from the theoretical calculations, because of the approximated description of such processes at an energy scale that cannot be treated making use of the perturbative theory. In spite of its success, the SM does not provide a satisfactory framework for some very fundamental theoretical and experimental results as the mechanism of symmetry breaking, the dark matter, the gravitational interaction and the barionic asymmetry seen in the universe.

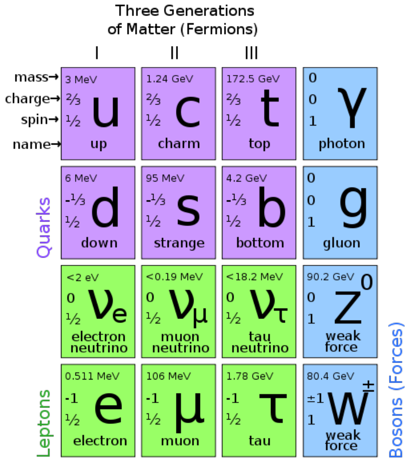


Figure 1.1.: Elementary particles predicted by the SM and experimentally observed.

1.0.2. Conservation Laws

This subsection will provide a brief overview of some invariance properties, and associated conservation laws, of the interactions provided by the Standard Model and that are relevant to the measurement presented in this thesis.

Parity

P is the parity operator that produces the inversion of the spatial coordinates:

$$P\psi(\vec{r}) = \psi(-\vec{r})$$

P is hermitian ($P = P^{\dagger}$), idempotent ($P^2 = 1$), and with eigenvalues ± 1. If $P\psi(\vec{r}) = \psi(-\vec{r}) = \psi(\vec{r})$. ψ is usually called a symmetric state, otherwise if $P\psi(\vec{r}) = \psi(-\vec{r}) = -\psi(\vec{r})$, ψ is defined as an antisymmetric state . For a system of two particle α and β whose intrinsic parity is ξ_α and ξ_β respectively, the intrinsic parity is defined as:

$$P|\alpha\beta\rangle = \xi_\alpha \xi_\beta (-1)^L |\alpha\beta\rangle$$

with L the angular momentum of the system.

The violation of the parity symmetry in weak interactions is most evident in the neutrino system. Neutrinos carry no electromagnetic or color charge, so they interact only via weak force. Treating neutrinos as massless particle, it can be stated that neutrinos with positive helicity (spin aligned with momentum) are right handed, and the ones with negative helicity (spin anti-parallel with respect to the momentum) are left handed. The parity inverts the handedness of the particle, so if parity was conserved, under parity a left handed neutrino should be transformed in a right handed neutrino, but it is experimentally tested that all observed neutrinos are left handed and all observed antineutrinos are right handed.

Charge conjugation

The action of the charge conjugation operator C is to reverse the sign of the change and magnetic moment of a particle. In other words, the operator C changes all the internal quantum numbers of a particle in the ones of the respective anti particle (For example $C|e^-\rangle = |e^+\rangle$) and leaves all the other coordinates (external degrees of freedom, such as the spatial coordinates for example) unchanged ($C|\vec{p}\rangle = |\vec{p}\rangle$). C is hermitian ($C = C^\dagger$) and idempotent ($C^2 = 1$).

Again in the case of neutrinos, the C operator transforms a particle in the relative antiparticle, so a neutrino (left handed) would be transformed in an antineutrino, left handed again, and this is in contradiction with the experimental evidence. Therefore weak interaction are not invariant under charge conjugation.

Charge-Parity conjugation

It was thought that CP symmetry was conserved in weak interactions, until 1964, when neutral Kaons were observed to decay in two different CP eigenstates, violating in this way CP invariance. A more accurate description and implications of CP violation are presented in section 1.2.

Charge-Parity and Time-reversal

CPT theorem states that all interactions are invariant under the operation if time reversal ($T\psi(t,\vec{r}) = \psi(-t,\vec{r})$), parity and charge conjugation applied together. This is based on the QFT principles: it is impossible to build a Quantun Field Theory where CPT invariance is violated. This theorem implies that if time reversal invariance or charge-parity (CP) invariance is violated, then also the other symmetry should be broken.

Summary

A description of the invariance properties of the interactions provided by the SM that are relevant for this thesis has been given. Tab. 1.1 has the purpose to summarize these conserved quantities under strong, EM and weak interactions in the SM.

Conserved quantity	Strong	Electromagnetic	Weak
Parity (P)	yes	yes	no
Charge conjugation (C)	yes	yes	no
CPT	yes	yes	yes
$CP(orT)$	yes	yes	violation $\mathcal{O}(10^{-5})$

Table 1.1.: Symmetries of the fundamental interactions.

1.0.3. Weak interactions

From experimental observations, lepton flavor seems to be largely conserved for charged leptons. Instead in charged current weak interactions ($W^{\pm}$ exchange) in the quark sector, transitions such as $s \to u$ are not uncommon, although they are rarer than transitions within one generation such as $u \to d$.
In order to maintain flavor conservation in the lepton sector, and yet allow cross-generational interactions between quarks, there has to be a significant difference between the actions of the weak force on leptons and quarks.

In the early 1960s, when this problem was first underlined, only u, d and s quarks had been discovered. At that time, Cabibbo suggested that the strength of the interaction governing the process $d \to u + W^-$carries a factor $\cos(\theta_c)$ and the process $s \to u + W^-$ a factor $\sin(\theta_c)$. The relative rates of these two examples implies that θ_c must be small. This solution was successful in resolving the rates of many interactions, but one significant problem indicated that this was not the full answer: the rate of $K^0 \to \mu^-\mu^+$ measured experimentally was far below the calculated rate under Cabibbo's theory to solve this issue. In 1970, Glashow, Iliopoulos and Maiani (GIM) proposed to introduce a fourth quark. This quark would couple more strongly to s than to d quarks, and importantly, introduced a new diagram in K^0 decays which cancelled the equivalent diagram for the u quark, therefore reducing the expected decay amplitude.

Combining the GIM mechanism with the Cabibbo theory leads to the conclusion that the quark states acted on by the charged weak current are not the physical states, but states rotated by the Cabibbo angle θc,

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

Eq. 1.0.3 shows that the d' and c' states which interact via the charged weak current are linear combinations of the physical d and s states. In 1974 the discovery of the J/ψ particle interpreted as a $c\bar{c}$ resonance gave strong evidence to this model. The extension by Kobayashi and Maskawa to three generations allowed $u\bar{s}$ for third generation mixing when the b quark was later discovered.

1.0.4. CKM mechanism

The Lagrangian describing the charged-current interaction for quarks is written:

$$L_{W^{\pm}} = \frac{g}{\sqrt{2}} \bar{u}_{L_i} (V_{CKM})_{ij} d_{L_j} W^{+}_{\mu} + h.c.$$

where the coupling constant g is associated with the $SU(2)_L$ group. The indices i and j represent the quark generations. Instead V_{CKM} is the Cabibbo-Kobayashi-Maskawa matrix (CKM), whose elements specify the strength of the coupling in each transition between quarks i and j. The CKM matrix relates the mass quarks eigenstates (d, s, b) to the quarks weak interaction eigenstates (d', s', b') in the following way:

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

A 3×3 complex matrix has 18 real parameters, but in the SM the CKM has only 4 free parameters, due to a couple of reasons:

1. The unitarity of V_{CKM} ($V^{\dagger}_{CKM} V_{CKM}$ = 1), that is required in order to guarantee the gauge invariance in the electro-weak interaction. This condition reduces the number of free parameters from 18 to 9: three angles and and six phases.

2. The quark fields can be redefined up to a phase by the transformation $u_i \to e^{i\xi} u_i$. This freedom allow us to eliminate five out of the six phases, leaving only one physical phase.

The four free parameters are traditionally interpreted as three rotational angles (θ_{ij}) and one phase (δ), and this interpretation leads to a certain parametrization of the CKM matrix:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}$$

where $c_{ij} = \cos(\theta_{ij})$ and $s_{ij} = \sin(\theta_{ij})$.

An alternative parametrization, proposed by Wolfenstein, which emphasize the relative amplitudes of the elements of the matrix, shows that transitions across two quark generations, or more generally transitions that involve further off-diagonal elements of the CKM matrix elements are suppressed with respect to transitions involving diagonal elements. In this parametrization, the 4 parameters used are: A, λ, ρ and η. The role of the CP-violating phase is played by η, and $\lambda \equiv \sin(\theta_{12})$ is the expansion parameter [1]. The Wolfenstein parametrization, ignoring terms of $\mathcal{O}(\lambda^5)$ or smaller, can be written as follows:

---

[1] The remaning 3 parameters are defined as follows: $\eta \equiv \sin(\theta_{13}) \sin\delta / A\lambda^3$, $\rho \equiv sin(\theta_{13}) \cos(\delta)/A\lambda^3$ and $A \equiv \sin(\theta_{13})/\lambda^2$.

$$V_{CKM} \sim \begin{pmatrix} 1-\frac{\lambda^2}{2}-\frac{\lambda^4}{8} & \lambda & A\lambda^3(\rho-i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2}-\frac{\lambda^4}{8}(1+4A^2) & A\lambda^2 \\ A\lambda^3(1-(1-\frac{\lambda^2}{2})(1-\rho-i\eta)) & -A\lambda^2+\frac{A\lambda^4}{2}(1-2(\rho+i\eta)) & 1-\frac{A^2\lambda^4}{2} \end{pmatrix}$$

At order of $\mathcal{O}(\lambda^4)$, it becomes explicit the imaginary part of the matrix element V_{ts}, that together with V_{tb}, V_{cb} and V_{cs} enters in the definition of β_s as will be explained later in this section (see eq. 1.0.5). The free parameters of the CKM matrix are not specified by the theory and they must be determined by the experiments. In the case of the CP-violating phase means that the SM can not predict its amount, but has sufficient complexity to accommodate the *CP* violation in its framework [2]. In the SM the unitary of the CKM matrix must hold:

$$\sum_{i=1}^{3} V_{ij}V_{ik}^{*} = \sum_{i=1}^{3} V_{ji}V_{ki}^{*} = \delta_{ij} \tag{1.0.1}$$

Expanding eq. 1.0.1 for any j and k yields nine equations, of which the six equations involving the off-diagonal elements of δ_{ij} describe triangles in the complex plane. These six triangles fall into two groups of three, differing only by their orientation in the complex plane: these are the so-called unitary triangles (see Fig. 1.2). By considering the triangles that arise out of the product of columns of V_{CKM} , the equations that have to be written are:

$$(V^{\dagger}V)_{31} : V_{ub}^{*}V_{ud} + V_{cb}^{*}V_{cd} + V_{tb}^{*}V_{td} = 0, \tag{1.0.2}$$

$$(V^{\dagger}V)_{32} : V_{ub}^{*}V_{us} + V_{cb}V_{cs} + V_{tb}^{*}V_{ts} = 0, \tag{1.0.3}$$

$$(V^{\dagger}V)_{21} : V_{us}^{*}V_{ud} + V_{cs}V_{cd} + V_{ts}^{*}V_{td} = 0. \tag{1.0.4}$$

The study of processes involving flavor-changing charged weak interaction (i.e. matter-antimatter oscillations of mesons or weak decays) allows the measurements of physical observables (oscillation frequencies, decay rates) that depend on real quantities such as the moduli of elements $|V_{ij}|$ in various combinations. These measurements can be converted into measurements of the length of the sides and interior angles of the unitary

---

[2] it can be shown that it is possible to have a CP-violating phase only in the case of having a number of quark generations equal to 3 or larger

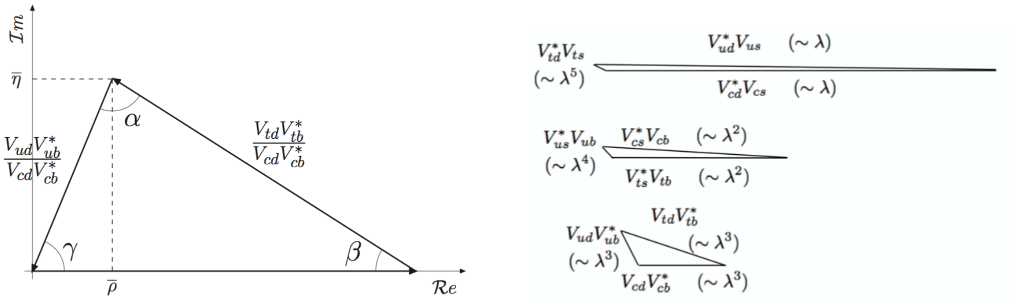
triangles. By measuring all the sides and interior angels, one can over-constrain the triangles and test whether unitary holds. Having larger interior angles makes easier the measurement of themselves, this is the reason why the most well studied triangle is the one arising from eq. 1.0.2. The interior angles of such a triangle are defined as:

$$\beta \equiv \arg\left(-\frac{V_{cb}^{*}V_{cd}}{V_{tb}^{*}V_{td}}\right); \quad \alpha \equiv \arg\left(-\frac{V_{tb}^{*}V_{td}}{V_{ub}^{*}V_{ud}}\right); \quad \gamma \equiv \arg\left(-\frac{V_{ub}^{*}V_{ud}}{V_{cb}^{*}V_{cd}}\right)$$

The triangle of interest in this thesis is the one arising from eq. 1.0.3. The value of its smallest angle, which is called β_s, by analogy to the canonical unitary triangle (the one formed by eq. 1.0.2), is defined as

$$\beta_s \equiv \arg\left(-\frac{V_{ts}^{*}V_{tb}}{V_{cs}^{*}V_{cb}}\right) \tag{1.0.5}$$

and it is proportional to the phase of the transition amplitude for $B_s - \bar{B}_s$ oscillation. The magnitude of such a transition is proportional to Δm_s. The oscillation frequency of Δm_s is related to CKM parameters according to $\Delta m_s \propto |V_{tb}^{*}V_{ts}|$, where the proportionality constant is a coefficient that includes hadronic matrix elements that must be calculated using lattice QCD.



(a) Normalized unitary triangle of the $B^0 - \bar{B}^0$ system (obtained from eq. 1.0.2).

(b) Unitary triangles of the CKM matrix, corresponding to eq. 1.0.4 (top), eq. 1.0.3 (middle), eq. 1.0.2, (bottom).

Figure 1.2.: Unitary Triangles.

1.1. The B mesons

B mesons are composed of a bottom antiquark and either an up (B^+), down (B^0), strange (B_s^0) or charm (B_c^+) quark [3]. Each B meson clearly ha its corresponding antiparticle, which is composed of a bottom quark and an up (B^-), down ($\bar{B}^0$), strange ($\bar{B}_s^0$) or

---

[3] The combination of $\bar{b}$ and t is not possible, because of the top quark's short lifetime. The $\bar{b}b$ combination is not a B meson, it is called "bottonium". In this thesis B_s^0 is denoted also as B_s and B^0 is also as B_d

charm (B_c^-) antiquark. Tab. 1.2 summarizes the main characteristics of the B mesons.

Meson	Quark content	Isospin	Mass(MeV/c^2)	Mean Lifetime (ps)
B^+	$u\bar{b}$	1/2	5279.17± 0.29	1.641 ± 0.008
B^0	$d\bar{b}$	1/2	5279.50 ±0.30	1.519 ± 0.007
B_s^0	$s\bar{b}$	0	5366.3 ± 0.6	$1.472^{+0.024}_{-0.026}$
B_c^+	$c\bar{b}$	0	6277± 6	0.4534 ± 0.041

Table 1.2.: B mesons: main characteristics. All B mesons have the samespin and parity $J^P = 0^-$.

As can be seen from the table, the B_s meson is composed by an heavy quark (the b quark) and a light one (the s quark). Reminding that, in the static quark model, the rest mass of the b quark is $4.20^{+0.17}_{-0.07}$ GeV/c ^{2}, and the s quark rest mass is only 104^{+26}_{-34} MeV/c ^{2}, what can be noticed is that one quark is much heavier than the other one. Mesons like this, consisting of one heavy (mass M) and one light (mass m) quark have some aspects in common with the hydrogen atom, where the nucleus ,i. e. the proton, (that is the analogue of the heavy quark in the present case) constitutes the most of the mass , and the electron (analogue of the light quark) is bound to the nucleus [22]. The atomic wave-function is, to the first order, independent of the nuclear mass; corrections are of the order (electron mass)/(nucleus mass) or (atomic binding energy)/(nucleus mass). Similarly, in a heavy-light quarks system, the effect of the heavy quark on the energy levels is of the order Λ_{QCD}/M, where $\Lambda_{QCD} \backsimeq 0.2$GeV is the strong interaction scale parameter, representing the light quark mass and the gluon binding potential. As a consequence, in the limit $M >> \Lambda_{QCD}$, the heavy quark acts approximately as a static color-triplet source, and its spin and flavor do not affect the light degrees of freedom. This is analogous to atomic physics, where isotopes with different nuclei are characterized by nearly the same properties. Therefore, properties of heavy-light quark hadrons are related by a symmetry called Heavy Quark Symmetry (HQS)[31].

1.1.1. Production

In order to understand the production of B mesons, two subsequent processes have to be presented: first the b (anti) quark production at the elementary interactions level, and then, due the quark confinement dynamics, its $fragmentation$ into a meson.

b quark production

There are mainly two efficient ways to produce b quarks, which are different for the kind of accelerator involved: the first method makes use of e^+ e^- colliders (e. g. the B factories at Stanford Linear Accelerator Center renamed SLAC National Accelerator Laboratory (SLAC) and KEK); in the second method hadron colliders are used (such as the Tevatron, where $p\bar{p}$ collisions are produced, or the LHC, where pp collisions are instead exploited). In both cases, the origin is a point-like interaction that results in a $b\bar{b}$ pair. The main difference of the two types of production is that in the $e^+ e^-$ machines the energy of the b quark generated is well known and is the same for all the events. This is due to the fact that e^+ and e^- are elementary particles, and so the beam energy is very close to the energy of the b quarks produced. Instead at hadron colliders the point-like interaction occurs between the elementary constituents of the hadrons (i.e. quarks and gluons): their energies span a continuous spectrum, because they are fractions of the energy of the colliding particles. In the case of hadron machines, the cross section for the production of a $b\bar{b}$ pair is calculable in perturbative QCD as much as the heavy quark mass m is larger than Λ_{QCD} [23]. The cross section in the QCD improved parton model as a function of the center of mass energy squared s is :

$$\sigma(s) = \sum_{i,j} = \int \frac{dx_1}{x_1}\frac{dx_2}{x_2}\hat{\sigma_{ij}}(\hat{s}, m^2, \mu^2)F_i(x_1, \mu^2)F_j(x_2, \mu^2) \tag{1.1.1}$$

where the F_i are the momentum densities of the partons in the incoming hadrons, x_i are the partons momentum fractions and $\hat{s} = x_1 x_2 s$ is the parton center of mass energy squared. The quantity $\hat{\sigma_{ij}}$ is the short distance cross section (i.e. the cross section of the point-like process involved).

The lowest order short distance cross section is calculated with the diagrams in Fig. 1.3

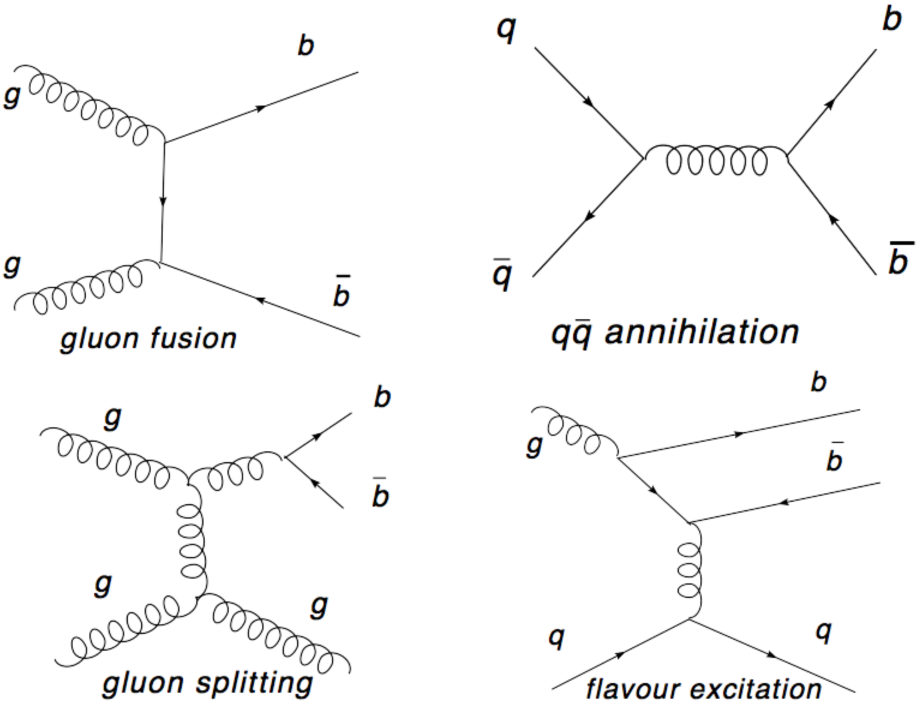


Figure 1.3.: Diagrams for b quark production at lowest order.

The majority of the $b\bar{b}$ pairs are produced in either of:

- *Flavor creation* can occur by two gluons from the colliding beam hadron undergoing hard scattering , or from the annihilation of a quark and an anti-quark. This is also called *direct production* and it is shown in the two upper diagrams in Fig. 1.3 .

- *Gluon splitting.* It takes place when a gluon produced in the hadron collision results in a $b\bar{b}$ pair in the fragmentation process

another process that allows the $b\bar{b}$ pair creation is the:

- *Flavor excitation.* A b quark from the sea of one of the interacting hadrons is scattered out in the strong interaction with a parton of the other colliding particle

All the three mechanisms of productions end in $b\bar{b}$ pairs which differ for their resulting angular distribution in the laboratory frame: in the flavor creation the two particle is back-to- back; in the gluon splitting the azimuthal angles of the two particles are very close, and the flavor excitation presents the intermediate configurations between the two previous situations. The cross section of $p\bar{p} \to \bar{b}X$ at the Tevatron, where $\sqrt{s} = 1.96\,\mathrm{TeV}$, is:

$$\sigma(p\bar{p} \to \bar{b}X, \sqrt{s} = 1.96\,\mathrm{TeV}, |Y_b| < 1) = 29.4^{+6.2}_{-5.4}\,\mu\mathrm{b} \tag{1.1.2}$$

where Y_b is the b quark rapidity (see Sect. 2.2.1) In spite of the high $b - \bar{b}$ cross section, about a hundred times greater than the $e^+\,e^-$ collider one, in hadron colliders it is a small percentage of the total cross section, which is of the order $\sigma_T \backsimeq$ 70 mb. This implies the need of an appropriate trigger selection to discriminate b-flavored events from the huge amount of background. Typical trigger selection requires a lepton from semileptonic b-decays, a dimuon signal from $b \to J/\psi$ decays, and often a signature based on the long lifetime of b-hadrons.

Fragmentation

Once the b ($\bar{b}$) is produced, the scale μ drops: at this point starts the fragmentation process. In fact, the fragmentation of quarks and gluons into hadrons involves confinement dynamics, and occurs at time scales that are long compared to those of the hard scattering that produced the quarks and gluons.

In single-particle inclusive hard-scattering processes, the fragmentation is factorized in a perturbativeQCD component (see [24] and references therein) from the hard interaction and a non-perturbative Fragmentation Function (FF), that here is called $D_i^H(x, \mu)$. The perturbative contribution can be calculated from theory, while the non perturbative FF must be determined from a fit to data itself. $D_i^H(x, \mu)$ is the probability density of a hadron H to form from parton i with momentum fraction x at factorization scale μ. Though non-perturbative, these FFs are universal and so, they may be determined for each hadron H in a few calibration experiments at some fixed scale μ_0, for subsequent use in other experiments and at other values of μ (see for istance [25], [26]). The fragmentation of heavy quarks is somewhat different. When the heavy quark is produced with an energy not much larger than its mass, the fragmentation process consists mainly of the non-perturbative transition of the heavy quark to the hadron H, which one assumes can be described by a non-perturbative FF. One may make a general ansatz for the functional form of this FF, whose parameters are to be fixed by fitting to experimental data. A heavy flavored meson should retain a large fraction of the momentum of the primordial heavy quark, and therefore its FF should be much harder than that of a light hadron. In the limit of a very heavy quark, one expects the FF for a heavy quark to go into any heavy hadron to be peaked near 1. This effect for the FF of a b quark in a B meson is used in the experimental identification of the heavy-flavored hadron.

The results of the fragmentation is to bind the quarks and gluons in colorless clusters of low relative momenta. They group in jets of hadrons, which flight in approximately the same directions of the quarks produced in the hard scattering. If one of the original quarks is a b, at least one of the jet's particles is a B meson or a b-flavored barion (e. g. Λ_b). This is the main difference between the heavy and light quarks fragmentation: a b-flavored hadron can be experimentally identified under certain conditions allowing us to identify b-jets

The most part of b-hadrons produced in the hadronization are excited states (e. g., B^* and B^{**}): they decay by electromagnetic and strong interactions into other b-mesons (or barions), and pions, kaons and photons. The time scale of these processes are so short that is not possible to distinguish them from the primary interaction.

1.1.2. Electro-Weak Decay: B mesons lifetimes

The measurement of the B_s lifetime play an important role in the analysis presented in this document. In the $B_s \to J/\psi\phi$ decay, we measure the average decay width Γ_s and the width difference $\Delta\Gamma_s$ between the two mass eigenstates (see Section 1.2). This measurement is an important test of our analysis framework and provides us with a precise measurement of $\Delta\Gamma_s$. In this section, we provide a brief theoretical introduction to B meson lifetimes. In the Standard Model, a b quark decays to a c or u quark through a virtual W boson. The decay width is given by:

$$\Gamma(b \to q) = \frac{3G_F^2 m_b^5}{192\pi^2} |V_{qb}|^2 F(\epsilon_q) \tag{1.1.3}$$

where G_F is the Fermi coupling constant and $F(\epsilon_q)$ is the factor associated with the decay phase space. Eq. 1.1.3 defines the relationship between fundamental SM parameters and the lifetime $\tau \equiv 1/\Gamma$ measurement. B meson lifetime are described to first order by the Spectator Model, where a heavy quark (here the b) is bound to a lighter spectator quark. In this model, the heavy quark dominates the lifetime, and the lifetimes of various B mesons should be the same. This is contradicted by the observations, which give the following hierarchy:

$$\tau_{B_c} < \tau_{B_s} \approx \tau_{B_d} < \tau_{B_u} \tag{1.1.4}$$

More precise predictions of hadron lifetimes must take into account additional considerations beyond eq. 1.1.3, since it makes a prediction only for free quarks, not for bound states which will be affected by quark-quark interactions.

1.2. CP violation in Mixing and Decay of neutral B mesons

The neutral B mesons, B_d and B_s, spontaneously transform into their own antiparticles and back. This phenomenon is called mixing (or flavor oscillation). The oscillation frequency has been measured in the $B_0 - \bar{B}_0$ system to be about $0.507 \pm 0.005\,\mathrm{ps}^{-1}$ and in the $B_s - \bar{B}_s$ system to be $17.77 \pm 0.10\mathrm{stat} \pm 0.07\mathrm{syst}\ \mathrm{ps}^{-1}$. The formalism that we are going to present for the mixing description can be applied both to the $B_0 - \bar{B}_0$ and to the $B_s - \bar{B}_s$ system. Thus, it can be chosen to use the notation B_q, where q = d or s, properly pointing out the differences of the parameters, where necessary. In the SM, $B_q - \bar{B}_q$ mixing is caused by the fourth order flavor-changing weak interaction described by the box diagrams in Fig. 1.4. Such transitions are called $|\Delta B| = 2$ transitions, because they change the bottom quantum number B by two units. In the SM $|\Delta B| = 2$ transition amplitudes are small, so measurements involving $B_q - \bar{B}_q$ mixing could easily be sensitive to new physics.

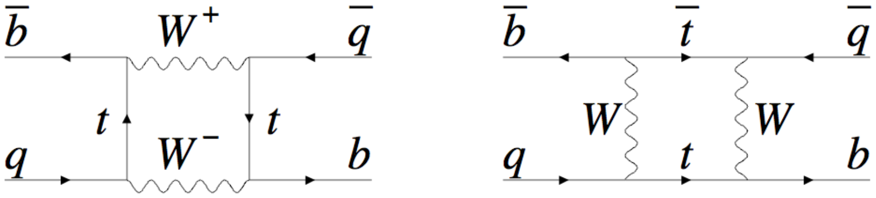


Figure 1.4.: Dominant box diagrams for the $B_q^0 \to \bar{B}_q^0$ transitions (q = d or s). Similar diagrams exist where one or both t quarks are replaced with c or u quarks.

1.2.1. Time developement

An initially produced B_q (or $\bar{B}_q$) evolves in time into a superposition of B_q and $\bar{B}_q$. Let $|B_q(t)\rangle$ denote the state at time t of a meson produced as B_q at $t = 0$, and analogously $|\bar{B}_q(t)\rangle$ the state at the time t of a meson produced as $\bar{B}_q$ at $t = 0$. The time evolution of these states is governed by the Schrodinger equation:

$$i\frac{t}{dt}\begin{pmatrix} |B_q(t)\rangle \\ |\bar{B}_q(t)\rangle \end{pmatrix} = \left(M - i\frac{\Gamma}{2}\right)\begin{pmatrix} |B_q(t)\rangle \\ |\bar{B}_q(t)\rangle \end{pmatrix} \tag{1.2.1}$$

with the mass matrix $M = M^\dagger$ and the decay matrix $\Gamma = \Gamma^\dagger$. The off-diagonal elements $M_{12} = M_{21}^*$ and $\Gamma_{12} = \Gamma_{21}^*$ change the flavor when acting on the meson, while the diagonal elements are flavor conserving. In the SM the leading contributions to M_{12} and Γ_{12} stem from the box diagram in Fig. 1.4. Γ_{12} originates from the real final states into which both B_q and B_q can decay. CPT invariance implies that $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$.

The mass eigenstates at time $t = 0$, $|B_L\rangle$ and $|B_H,\rangle$, are linear combinations of $|B_q\rangle$ and $|\bar{B}_q\rangle$:

$$\begin{array}{lll} lighter \quad eigenstate: & & |B_L\rangle = p|B_q\rangle + q|\bar{B}_q\rangle \\ heavier \quad eigenstate: & & |B_H,\rangle = p|B_q\rangle - q|\bar{B}_q\rangle \end{array} \tag{1.2.2}$$

with $|p|^2 + |q|^2 = 1$.
Let us denote masses and widths of the two eigenstates with $M_{L,H}$ and $\Gamma_{L,H}$ and define:

$$\Gamma = \frac{1}{\tau_b} = \frac{\Gamma_L + \Gamma_H}{2} \tag{1.2.3a}$$

$$\Delta m = M_H - M_L \tag{1.2.3b}$$

$$\Delta\Gamma = \Gamma_L - \Gamma_H. \tag{1.2.3c}$$

while $\Delta m > 0$ by definition, $\Delta\Gamma$ can have either sign, depending on the definition. The definition 1.2.3 implies $\Delta\Gamma > 0$ in the SM.

The SM predicts a very small value for $\Delta\Gamma_d/\Gamma_d$ the $B_d - \bar{B}_d$ system (below 1%), and for the $B_s - \bar{B}_s$ system a ratio $\Delta\Gamma_s/\Gamma_s$ considerably larger (order of 10%). These width differences are caused by the existence of final states to which both the B_q^0 and $\bar{B}_q^0$ mesons can decay. Such decays involve $b \to c\bar{c}q$ quark-level transitions, which are Cabibbo suppressed if $q = d$ and Cabibbo-allowed if $q = s$.

Thus, in all formulae, terms containing $\Delta\Gamma_q$ can be neglected for the $B_d - \bar{B}_d$ system, but not for the $B_s - \bar{B}_s$. For these reason, in the following we prefer to refer to B_s mesons only, putting $q = s$ everywhere.

Last update results [27] for B_s mesons are:

$$\Delta\Gamma_s = 0.062^{+0.034}_{-0.037}\,\mathrm{ps}^{-1} \tag{1.2.4a}$$

$$\Delta\Gamma_s/\Gamma_s = 0.092^{+0.051}_{-0.054} \tag{1.2.4b}$$

$$\tau_L = \frac{1}{\Gamma_L} = 1.408^{+0.033}_{-0.030}\,\mathrm{ps} \tag{1.2.4c}$$

$$\tau_H = \frac{1}{\Gamma_H} = 1.543^{+0.058}_{-0.060}\,\mathrm{ps} \tag{1.2.4d}$$

The time evolution of the mass eigenstates is governed by:

$$|B_{L,H}(t)\rangle = e^{-(iM_{L,H} - \Gamma_{L,H}/2)t}|B_{L,H}\rangle \tag{1.2.5}$$

with $|B_{L,H}\rangle = |B_{L,H}(t = 0)\rangle$. By using eq. 1.2.2 , together with eq. 1.2.5, one can find the time evolution for $|B_s\rangle$ and $|\bar{B}_s\rangle$:

$$|B_s(t)\rangle = g_+(t)|B_s\rangle - \frac{q}{p} g_-(t)|\bar{B}_s\rangle \tag{1.2.6a}$$

$$|\bar{B_s}(t)\rangle = \frac{p}{q} g_-(t)|B_s\rangle + g_+(t)|\bar{B}_s\rangle \tag{1.2.6b}$$

where

$$g_+(t) = e^{-imt} e^{-\Gamma t/2} \Big[\cosh \frac{\Delta\Gamma t}{4} \cos \frac{\Delta m t}{2} - i \sinh \frac{\Delta\Gamma t}{4} \sin \frac{\Delta m t}{2} \Big] \tag{1.2.7a}$$

$$g_-(t) = e^{-imt} e^{-\Gamma t/2} \Big[- \sinh \frac{\Delta\Gamma t}{4} \cos \frac{\Delta m t}{2} + i \cosh \frac{\Delta\Gamma t}{4} \sin \frac{\Delta m t}{2} \Big] \tag{1.2.7b}$$

1.2.2. CP violating variables

Going back to the eigenvalues problem in eq. 1.2.1, the first consideration that has to be made is that in general, the mass eigenstates are not the same as the CP eigenstates, which are defined as:

$$\begin{aligned} |B_s^{even}\rangle &= \frac{1}{\sqrt{2}}(|B_s\rangle - |\bar{B}_s\rangle) \\ |B_s^{odd}\rangle &= \frac{1}{\sqrt{2}}(|B_s\rangle + |\bar{B}_s\rangle) \end{aligned} \tag{1.2.8}$$

Then, by examining the eigenvalue problem in eq. 1.2.1 it can be found that the experimental information $\Delta m >> \Gamma$ model-independently implies $|\Gamma_{12}| << |M_{12}|$. By expanding the eigenvalues and q/p in Γ_{12}/M_{12}, one finds:

$$\Delta m = 2|M_{12}| \tag{1.2.9}$$

$$\Delta\Gamma = 2|\Gamma_{12}| cos\phi \tag{1.2.10}$$

$$\frac{q}{p} = -e^{-i\phi_M} \Big[1 - \frac{a}{2} \Big]. \tag{1.2.11}$$

Where the phase ϕ is defined as:

$$\frac{M_{12}}{\Gamma_{12}} = - \Big| \frac{M_{12}}{\Gamma_{12}} \Big| e^{i\phi} \tag{1.2.12}$$

The mixing phase ϕ_M is:

$$\phi_M = \arg M_{12} \tag{1.2.13}$$

and a small correction has been kept in eq. 1.2.11:

$$a = \left|\frac{\Gamma_{12}}{M_{12}}\right| \sin\phi \tag{1.2.14}$$

but all the terms of order Γ_{12}^2/M_{12}^2 have been neglected. The quantity a in eq. 1.2.14 is also known as A_{SL} and has been measured for instance as the asymmetry between positive and negative same sign muon pairs originating from semileptonic decays of pairs of B hadrons (see sec. 1.4.) The phase ϕ is physical and convention-independent; if $\phi = 0$, CP violation in mixing vanishes. Since Γ_{12} is dominated by the $b \to c\bar{c}s$ transition, that is a CKM-favoured tree-level decay, so it is practically insensitive to new physics. On the other hand, M_{12} is almost completely induced by short-distance physics. Within the Standard Model the top quarks give the dominant contribution to $B_s\bar{B}_s$ mixing.This contribution is suppressed by four powers of the weak coupling constant and two powers of the CKM matrix element $|V_{ts}| \simeq 0.04$. Hence new physics can easily compete with the Standard Model and possibly even dominate M_{12}. If the non-standard contributions to M_{12} are unrelated to the CKM mechanism of the three-generation Standard Model, they will affect the mixing phase ϕ_M. The SM prediction is $\phi_M = arg(V_{tb}V_{ts}^*)^2$. [4].

For the Bs decay into a final state f, let us introduce the $|\Delta B| = 1$ matrix elements:

$$A_f = \langle f|B_s\rangle \qquad \text{and} \qquad \bar{A}_f = \langle f|\bar{B}_s\rangle \tag{1.2.15}$$

and the key quantity for the CP violation:

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f} \tag{1.2.16}$$

The time-dependent decay rate $\Lambda(B_s(t) \to f)$ of an initially tagged B_s into some final state f is defined as:

$$\Lambda(B_s(t) \to f) = \frac{1}{N_B}\frac{dN(B_s \to f)}{dt} \tag{1.2.17}$$

where $dN(B_s \to f)$ denotes the number of decays of a B_s meson at time t into the final state f occurring within the time interval between t and $t + dt$. $N_B$is the total number of B_s mesons produced at time $t = 0$. An analogous definition holds for $\bar{B}_s$.
Using the time evolution expressions above, one can write the full expression for the

---

[4] In the $B_d\bar{B}_d$ system the large value of the analogue quantity $\phi_M = arg(V_{tb}V_{td}^*)^2$ allows the observation of the CP violation in the interference with the mixing by means of the measurement of the quantity $\sin 2\beta$

decay rate of B_s and $\bar{B}_s$ as follows:

$$\begin{aligned}\Lambda(B_s(t) \to f) = \mathcal{N}_f |A_f|^2 e^{-\Gamma t}\Big[\frac{1+|\lambda_f|^2}{2}\cosh\frac{\Delta\Gamma t}{2} + \frac{1-|\lambda_f|^2}{2}\cos(\Delta m t) + \\ \mathcal{R}\lambda_f \sinh\frac{\Delta\Gamma t}{2} + \mathcal{I}\lambda_f \sin(\Delta m t)\Big]\end{aligned} \tag{1.2.18}$$

$$\begin{aligned}\Lambda(\bar{B}_s(t) \to f) = \mathcal{N}_f |A_f|^2 e^{-\Gamma t}\Big(\frac{p}{q}\Big)^2\Big[\frac{1+|\lambda_f|^2}{2}\cosh\frac{\Delta\Gamma t}{2} - \frac{1-|\lambda_f|^2}{2}\cos(\Delta m t) + \\ \mathcal{R}\lambda_f \sinh\frac{\Delta\Gamma t}{2} + \mathcal{I}\lambda_f \sin(\Delta m t)\Big]\end{aligned} \tag{1.2.19}$$

Moreover one could write $\left(\frac{p}{q}\right)^2$ as $(1+a)$, and remember that a is a small correction factor. If B_s and $\bar{B}_s$ are produced in equal number, without making a distinction between the meson and its antiparticle, the decay rate is given by:

$$\begin{aligned}\Lambda(f,t) &= \Lambda(B_s(t) \to f) + \Lambda(\bar{B}_s(t) \to f) = \\ &= \mathcal{N}_f |A_f|^2 (1+|\lambda_f|^2) e^{\Gamma t}\Big[\cosh\frac{\Delta\Gamma t}{2} - \frac{2Re\lambda_f}{1+|\lambda_f|^2}\sinh\frac{\Delta\Gamma t}{2}\Big] + \mathcal{O}(a)\end{aligned} \tag{1.2.20}$$

The time evolution formulae and the CP asymmetries expressions in the forthcoming part of this section can be maybe conveniently expressed in terms of:

$$\mathcal{A}_{\mathcal{CP}}^{\text{dir}} = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}; \qquad \mathcal{A}_{\mathcal{CP}}^{\text{mix}} = -\frac{2Im\lambda_f}{1+|\lambda_f|^2}; \qquad \mathcal{A}_{\Delta\Gamma} = -\frac{2Re\lambda_f}{1+|\lambda_f|^2} \tag{1.2.21}$$

which obey the relation:

$$|\mathcal{A}_{\mathcal{CP}}^{\text{dir}}|^2 + |\mathcal{A}_{\mathcal{CP}}^{\text{mix}}|^2 + |\mathcal{A}_{\Delta\Gamma}|^2 = 1 \tag{1.2.22}$$

If f is a CP eigenstate, $CP|f\rangle = \pm|f\rangle$, then $\mathcal{A}_{\mathcal{CP}}^{\text{dir}} \neq 0$ or $\mathcal{A}_{\mathcal{CP}}^{\text{mix}} \neq 0$ signals CP violation: a non-vanishing $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$ implies $|A_f| \neq |\bar{A}_f|$, meaning direct CP violation; $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$ measures mixing-induced CP violation in the interference of $B_s \to f$ and $B_s \to f$. The third quantity, $\mathcal{A}_{\Delta\Gamma}$, plays a role if $\Delta\Gamma$ is sizeable.

Then CP-violation effects in decays $B_s \to J/\psi\phi$ can be classified into three phenomenological types: CP-violation in decay (or direct CP-violation), in mixing, and in the interference between decays with mixing and without mixing, i. e. :

- *direct CP violation:*

$$\Big|\frac{A_f}{\bar{A}_f}\Big| \neq 1 \tag{1.2.23}$$

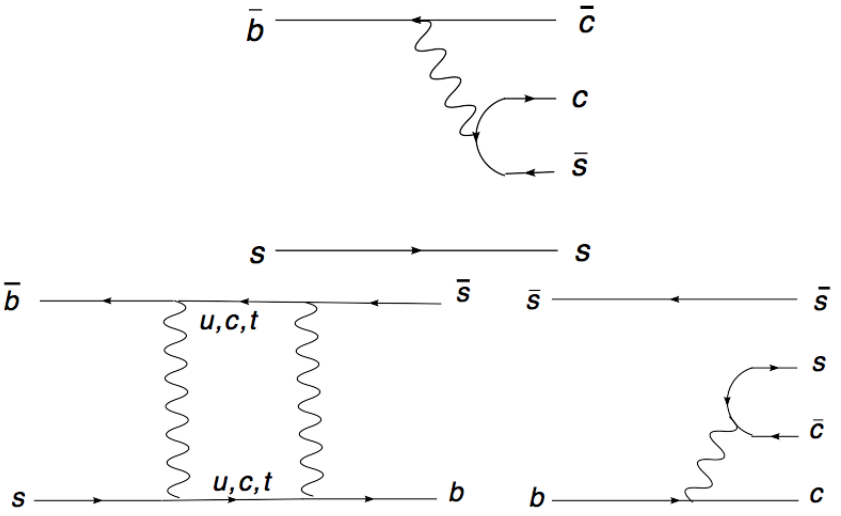


Figure 1.5.: CP violation in interference.

that corresponds to a non-vanishing $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$.
In SM the weak phase contributing to direct CP violation in B_s decays is Cabibbo-suppressed by λ^2, where λ is the Wolfenstein parameter. Furthermore, new physics contributions to this phase have to compete with a tree diagram, and therefore are not expected to provide a sufficiently large experimental signature. We therefore assume that $|A_f| = |\bar{A}_f|$ in Bs decays, ignoring direct CP violation.

- *CP violation in mixing:*

$$\left|\frac{q}{p}\right| \neq 1 \tag{1.2.24}$$

In the B_s system, $|q/p| - 1 < \mathcal{O}(10^{-2})$. The presence of CP violation in mixing results in a charge asymmetry in semileptonic B_s decays such as $B_s \to \mu + D_s^- \nu X$ and its conjugate B_s mode. The resultant time-integrated decay asymmetry is sensitive to deviations of $|q/p|$ from 1. The decay rate of $B_s \to J/\psi\phi$ does not isolate factors of $|q/p|$, and we therefore do not speak strictly of CP violation in mixing.

- *CP violation in the interference* between decay with and without mixing

$$\mathcal{I}m(\lambda_f) \neq 0, \qquad \lambda_f \equiv \frac{q}{p}\frac{\bar{A}_f}{A_f} \tag{1.2.25}$$

This type of CP violation is present only in modes for which the final state is available to both B_s and B_s, therefore including $B_s \to J\psi\phi$. Fig. 1.5 shows the Feynman diagrams demonstrating the interference in this decay mode.

For decays of B_s to a CP eigenstate (that is denoted here as f_{CP}), new physics can manifest itself in CP violation arising from the interference between decay amplitudes with and without mixing, which can be probed by measuring the time-dependent CP asymmetry:

$$a_{CP}(t) = \frac{\Lambda(B_s(t) \to f_{CP}) - \Lambda(\bar{B}_s(t) \to f_{CP})}{\Lambda(B_s(t) \to f_{CP}) + \Lambda(\bar{B}_s(t) \to f_{CP})} = \frac{\mathcal{A}_{CP}^{dir} \cos(\Delta m t) + \mathcal{A}_{CP}^{mix} \sin(\Delta m t)}{\cosh \frac{\Delta\Gamma t}{2} - \mathcal{A}_{\Delta\Gamma} \sinh \frac{\Delta\Gamma t}{2}} \tag{1.2.26}$$

Because of the value of $\Delta\Gamma_s \neq 0$ and the fact that the final state of the $B_s \to J/\psi\phi$ is not a CP-eigenstate, but an admixture of CP-even and CP-odd contributions, an observation of CP violation in $B_s \to J/\psi\phi$ requires a statistical determination of the odd and even contribution to final state decays. This statistical separation has been performed in this measurement by using an angular analysis (cfr. section 1.3).

1.2.3. Standard Model expectation and New Physics

As seen in sec. 1.0.4 the phase β_s , defined as the phase of λ_f in the $b \to c\bar{c}s$ transitions, is associated with the CP violation in the B_s system, and it is the angle $\beta_s = \arg(-V_{tb}V_{ts}^*/V_{cb}V_{cs}^*)$ of the B_s unitary triangle. The phase β_s is predicted by the Standard Model to be equal to $\beta^{SM} = 0.020 \pm 0.005$ [2, 3, 4]. We can relate β and ϕ in the following manner. First, we define SM and new physics (NP) contributions to ϕ. With the phase conventions we have chosen, the NP contribution to ϕ and β_s is:

$$\phi = \phi^{SM} + \phi^{NP} \qquad \text{and} \qquad 2\beta_{\mathrm{s}} = 2\beta_{\mathrm{s}}^{\mathrm{SM}} - \phi^{\mathrm{NP}} \tag{1.2.27}$$

Since β^{SM} is expected to be negligibly small, in the presence of new s physics, we can neglect the SM contribution, and use the shorthand:

$$2\beta_s \simeq -\phi^{NP} \qquad \text{and} \qquad \phi \simeq \phi^{\mathrm{NP}}. \tag{1.2.28}$$

1.3. Angular Distributions of $P \to VV$ decays

The notation $P \to VV$ stands for the class of the decays of a Pseudoscalar meson (with $J^P = 0^-$), like the B, into two Vector mesons ($J \neq 0$).

Most B-decay polarization analyses are limited to the case when the spin of the B-meson daughters is 1. In this case, there are only three independent amplitudes a, b and c, corresponding to the probability that the B_s meson decays in a state with one of the three possible relative angular momenta: $L = 0, 1, 2$ to obtain the spin 0 of the initial state, if L is the relative angular momentum of the two vector mesons. The overall decay amplitude would involve three complex terms proportional to the above amplitude. The exact angular dependence would depend on the quantum numbers of the B-meson daughters and of their decay products.

In the analysis here presented $P \to VV$ decays are of particular relevance since the final state with $L = 1$ corresponds to a CP-odd state, while the states with angular momentum $L = 0$ and $L = 2$ correspond to CP-even final states (which are linear combination of the parallel and of the longitudinally polarized final states). In the absence of CP violation the heavy long-lived mass eigenstate B_H is CP-odd and decays to CP-odd final state, as the light, short lived mass eigenstate B_L decay to CP-even final states. For this reason, in the hypothesis of CP-violation, the rate difference between the CP-even and CP-odd final states, provides a measurement of the CP-violation.

The starting point to obtain our observables is the definition of an appropriate reference frame. In this frame the number of variables necessary to the description of the process are identified as the kinematics degrees of freedom of the system. Let us refer to the B meson as the parent (or as the initial state) and to the two vector mesons as daughter particles (V_1 and V_2), and to their decay products as final state particles (P_1, P_2 from V_1, and P_3, P_4 from V_2).

Each of the four final state particles has four kinematic degrees of freedom (the four components of the particle four-momentum). There are seven particles of well defined mass in the decay, which provide seven constraints from the relation between the four momentum and the invariant mass. In addition, since the entire system is translation and rotation invariant, the vector momentum of the parent B meson is irrelevant to the decay process, as well as the orientation of the decay with respect to the laboratory frame is. The former removes three degrees of freedom in the form of the vector momentum components of the B meson, while the latter removes three more in the form of Euler rotation angles of the decay frame with respect to the laboratory one. This leaves us with 16 - 7 - 3 - 3 = 3 degrees of freedom in the decay. These remaining degrees of freedom are the angles at which the final state particles are produced. There are two standard choices we can make to define the three decay angles we wish to measure: the helicity and the transversity formalisms. In the present analysis the transversity basis has been chosen, in order to make easier the extraction the CP-odd and CP-even components.

In the transversity basis, the angular distribution is written in the *linear polarization* basis, which is defined in the following form:

$$\Lambda(B \to V_1 V_2) = A_0 \frac{\lambda_1^{*L} \lambda_1^{*L}}{u} - A_{\parallel} \frac{\lambda_1^{*T} \cdot \lambda_1^{*T}}{\sqrt{2}} - i A_{\perp} \frac{\lambda_1^{*} \times \lambda_2^{*}}{\sqrt{2}} \cdot \hat{p_2} \tag{1.3.1}$$

with $u = p_1 p_2 / m_1 m_2$, $\hat{p_2}$ is the unit vector along the V_2 motion direction in the V_1 rest frame and λ_1, λ_2 and m_1, m_2 represent the polarization vectors and masses of the two vector mesons, respectively. $\vec{p}$ indicates the three-momentum of the decaying particle, $\vec{k}$ and $\vec{q}$ are the three-momenta of the vector mesons V_1 and V_2 respectively. The symbols λ_L represent the longitudinal components of the polarization vectors ($\lambda_1^L = \lambda_1 \cdot \vec{k}/|\vec{k}|, \lambda_2^L = \lambda_2 \cdot \vec{q}/|\vec{q}|$), while λ^T are the transverse component, $\lambda_i^T = \lambda_i - \lambda_i^L$. The unitary relation holds for the transversity amplitudes:

$$|A_0|^2 + |A_{\parallel}|^2 + |A_{\perp}|^2 = 1 \tag{1.3.2}$$

These amplitudes are often referred to as *polarization amplitudes*.

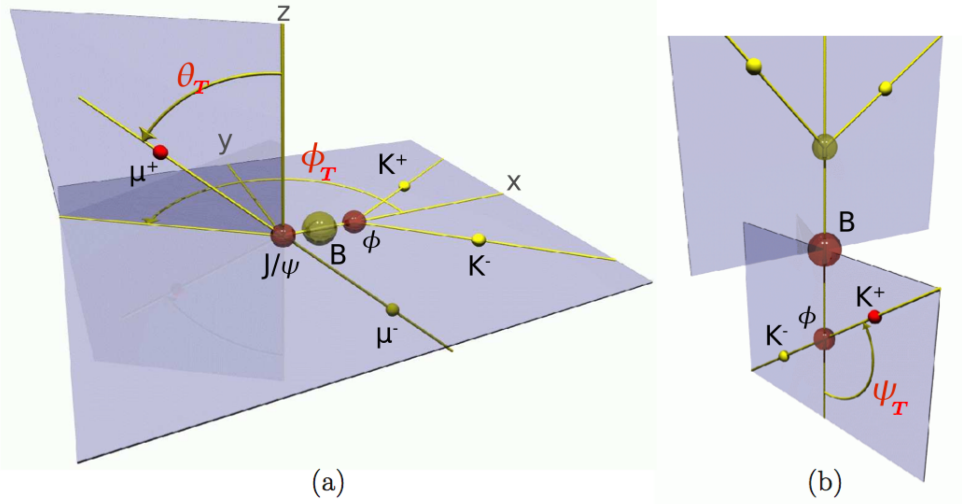


Figure 1.6.: Transversity basis definition.

The transversity basis, illustrated in Fig. 3.3.2, requires a change of angular variables used to describe the decay, defined as follows. The first two angles are calculated in the rest frame of the J/ψ, and the third in the rest frame of the ϕ. In the rest frame of the J/ψ, the ϕ meson direction defines the x axis. The plane of K^+K^- defines the xy plane with $p_y(K^+) > 0$. From there:

- θ_T : in the J/ψ meson rest frame, the angle between $p(\mu^+)$ and the xy plane
- ϕ_T : in the J/ψ meson rest frame, the angle between the x axis and $p_{xy}(\mu^+)$
- ψ_T : in the ϕ meson rest frame, the angle between $p(K^+)$ and $p(J/\psi)$

For brevity and convenience sometimes the symbol $\vec{\omega} = \{\cos\theta_T, \phi_T, \cos\psi_T\}$ is used to refer to the transversity variables together. The vector notation is a simply short-hand, as $\vec{\omega}$ does not transform as a vector under rotations in space. Throughout this thesis, the transversity subscripts are dropped on $\{\theta, \phi, \psi\}$ when there is no possible ambiguity.

The time and angle dependent $P \to VV$ decay rates for B_s and $\bar{B}_s$ expressed in terms of the polarization amplitudes $|A_0|$, $|A_{\|}|$, and $|A_{\perp}|$, and two relative strong phases of the amplitudes defined as: $\delta_{\|} = \arg(A_0 A_{\|}^*)$ and $\delta_{\perp} = \arg(A_0 A_{\perp}^*)$, are derived in a manner analogous to the derivation of eq. 1.2.18, and the result is:

$$\frac{d^3\Lambda(\vec{\omega},t)}{d\vec{\omega}} \propto \sum_k O^k(t) g^k(\vec{\omega}) \tag{1.3.3}$$

where the angular functions $g^k(\vec{\omega})$ contain the dependence on the transversity angles:

$$\begin{aligned}
g^1(\psi,\theta,\phi) &= 2\cos^2\psi(1-\sin^2\theta\cos^2\phi)\\
g^2(\psi,\theta,\phi) &= \sin^2\psi(1-\sin^2\theta\sin^2\phi)\\
g^3(\psi,\theta,\phi) &= \sin^2\psi\sin^2\theta\\
g^4(\psi,\theta,\phi) &= (-1)\sin^2\psi\sin 2\theta\sin\phi\\
g^5(\psi,\theta,\phi) &= \frac{1}{\sqrt{2}}\sin 2\psi\sin^2\theta\sin 2\phi\\
g^6(\psi,\theta,\phi) &= \frac{1}{\sqrt{2}}\sin 2\psi\sin 2\theta\cos\phi
\end{aligned} \tag{1.3.4}$$

The dependence on $2\beta_s$, $\Delta m, \Delta\Gamma, \Gamma, \delta_{\parallel}$ and $\delta_{\perp}$ is contained in the observables $O^k(t)$, which for the B_s ($\bar{B}_s$) taking the upper signs (lower), are given by:

$$\begin{aligned}
O^1(t) &= |A_0(t)|^2 =\\
&= |A_0|^2 e^{-\Gamma t}(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma| t}{2} \mp \sin 2\beta_s \sin\Delta m t)\\
O^2(t) &= |A_{\parallel}(t)|^2 =\\
&= |A_{\parallel}|^2 e^{-\Gamma t}(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma| t}{2} \mp \sin 2\beta_s \sin\Delta m t)\\
O^3(t) &= |A_{\perp}(t)|^2 =\\
&= |A_{\perp}|^2 e^{-\Gamma t}(\cosh\frac{\Delta\Gamma t}{2} + |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma| t}{2} \pm \sin 2\beta_s \sin\Delta m t)\\
O^4(t) &= \mathcal{I}m\{A_{\parallel}^*(t)A_{\perp}(t)\} =\\
&= |A_{\parallel}||A_{\perp}|e^{-\Gamma t}(\pm\sin(\delta_{\perp}-\delta_{\parallel})\cos\Delta m t \mp \cos(\delta_{\perp}-\delta_{\parallel})\cos 2\beta_s \sin\Delta m t\\
&+ \cos(\delta_{\perp}-\delta_{\parallel})\sin 2\beta_s \sinh\frac{\Delta\Gamma t}{2})\\
O^5(t) &= \mathcal{R}e\{A_0^*(t)A_{\parallel}(t)\} =\\
&= |A_0||A_{\parallel}|e^{-\Gamma t}\cos\delta_{\parallel}(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma| t}{2} \mp \sin 2\beta_s \sin\Delta m t)\\
O^6(t) &= \mathcal{I}m\{A_0^*(t)A_{\perp}(t)\} =\\
&= |A_0||A_{\perp}|e^{-\Gamma t}(\pm\sin\delta_{\perp}\cos\Delta m t \mp \cos\delta_{\perp}\cos 2\beta_s + \cos\delta_{\perp}\sin 2\beta_s \sinh\frac{\Delta\Gamma t}{2})
\end{aligned} \tag{1.3.5}$$

They can be rearranged in order to express the angular decay rate in terms of Γ_H and Γ_L.

The $O^k(t)$, functions, and therefore eq. 1.3.3 , require thee knowledge of the full time evolution of the B_s meson from production to decay, including the production flavor (B_s or $\bar{B}_s$). This requires the use of flavor- tagging algorithms explained in Chapter 3.

We obtain another useful expression retaining a dependence on $2\beta_s$ even if we do not know the production flavor of the B meson. Since B mesons are produced as a result of $p\bar{p} \to b\bar{b}$ interactions, and the probability for hadronization into B_s is equal for matter and antimatter, an equal number of B_s and B_s are created at the Tevatron. Summing the decay rates for upper and lower signs in 1.3.3 cancels out the dependence on Δm_s.

$$\frac{d^3\Lambda^{untag}(\vec{\omega},t)}{d\vec{\omega}} \propto \sum_k O^{k\ untag}(t) g^k(\vec{\omega}) \tag{1.3.6}$$

with

$$\begin{aligned}
O^{1\ untag}(t) &= |A_0|^2 e^{-\Gamma t} 2(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s| \sinh\frac{|\Delta\Gamma| t}{2}) \\
O^{2\ untag}(t) &= |A_\parallel|^2 e^{-\Gamma t} 2(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s| \sinh\frac{|\Delta\Gamma| t}{2}) \\
O^{3\ untag}(t) &= |A_\perp|^2 e^{-\Gamma t} 2(\cosh\frac{\Delta\Gamma t}{2} + |\cos 2\beta_s| \sinh\frac{|\Delta\Gamma| t}{2}) \\
O^{4\ untag}(t) &= |A_\parallel||A_\perp| e^{-\Gamma t} 2(\cos(\delta_\perp - \delta_\parallel) \sin 2\beta_s \sinh\frac{\Delta\Gamma t}{2}) \\
O^{5\ untag}(t) &= |A_0||A_\parallel| e^{-\Gamma t} cos\delta_\parallel 2(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s| \sinh\frac{|\Delta\Gamma| t}{2}) \\
O^{6\ untag}(t) &= |A_0||A_\perp| e^{-\Gamma t} 2(\cos\delta_\perp \sin 2\beta_s \sinh\frac{\Delta\Gamma t}{2})
\end{aligned} \tag{1.3.7}$$

1.4. Current experimental status

Once the theoretical framework has been introduced and it has be seen how parameters such as $\Delta\Gamma$, the lifetime and the polarization amplitudes enter the expressions of the differential decay rate of the $B_s \to J/\psi\phi$ mode, we focus the attention on their current experimental and theoretical status of knowledge.

1.4.1. β_s measurements

Nowadays the measurement of *CP* violation in $B_s \to J/\psi\phi$ is performed by three experiments: CDF and DØ, using $p\bar{p}$ collisions at the Tevatron, and LHCb, using pp collisions at the Large Hadron Collider (LHC). Details of those measurements are available for CDF in Refs. [35, 36, 37], D0 in Refs. [38, 39, 119] and LHCb in Ref. [40]. The selected sample comprises of 6504 ± 85 $B_s \to J/\psi\phi$ signal events in $5.2\ \text{fb}^{-1}$ of data at CDF, 5598 ± 113 signal events in $8.0\ \text{fb}^{-1}$ of data at DØ and 8276 ± 94 $B_s \to J/\psi\phi$ decays in $0.34\ \text{fb}^{-1}$ at LHCb. Non-trivial statistical issues are present in the estimation of β_s, because of the particular features that affect the likelihood function used. The latter is in fact symmetric under a certain simultaneous transformation of the polarization strong phases, $\Delta\Gamma$ and β_s. Such property can be traced in eq. 1.3.5 and they will be further discussed in Chapter 4. It follows that there is a danger of non–Gaussian behaviour of the likelihood, which to some extent depends on the statistics and the true values of the parameters. Given the importance of the measurement for putting bounds on NP, it is important to make sure that any non–Gaussian behavior is properly taken into account. In order to achieve this, the experiments resort to a frequentist treatment based on the Feldman and Cousins method ??. Therefore the results are usually quoted as confidence regions in the β_s–$\Delta\Gamma$ plane, or equivalently in the ϕ_s–$\Delta\Gamma$ plane (see eq. 1.2.28). The latest results in terms of such confidence regions obtained by CDF, DØ and LHCb experiments are compared in Fig. 1.7. As the SM is a special case, each experiment derives consistency between the data and the SM expectation. The consistency is characterized by the p-value, which is 44% (0.8σ) at both CDF and LHCb, and 22% (1.2σ) at DØ. Those tests provide an answer to the question of whether both $\Delta\Gamma$ and β_s are simultaneously consistent with the SM. It is interesting also the p-value of the β_s alone, rather than the allowed region in two–dimensional space. To obtain this, an analogous frequentist procedure is used, resulting in unidimensional confidence intervals. Also in this case, the compatibility of data with the SM prediction β_s is within $\sim 1\sigma$ for all experiments.

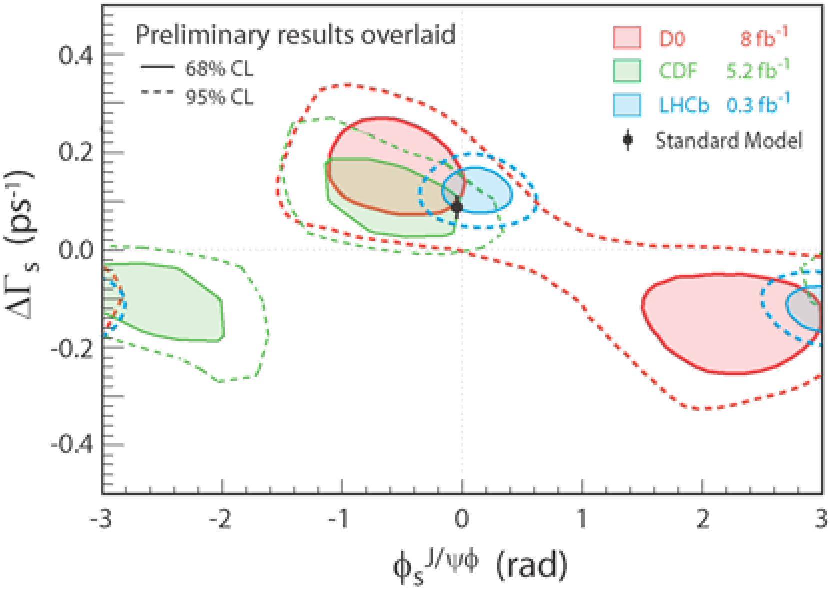


Figure 1.7.: Latest ϕ_s- $\Delta\Gamma$ confidence regions from CDF, DØ and LHCb experiments.

1.4.2. NP and $\Delta\Gamma$ measurement

While searches for NP usually focus on the β_s measurement, the experimental determination of $\Delta\Gamma$ is actually interesting by its own right. First of all, while a non–zero value of $\Delta\Gamma$ is clearly predicted in the SM [20], a firm experimental observation of a non–zero lifetime difference. Then, recently updates of Tevatron's analyses [119, 97] and a new precise measurement by LHCb [40] suggest a value of β_s close to the SM value within 1σ. Although experimental uncertainties still leave room for physics beyond the SM, the strength of possible NP should be weaker than what could be expected from first observations [ref]. If this is the case, it is potentially in disagreement with the CP–violating asymmetry of semileptonic decays of neutral B mesons (A_{SL}) recently seen by the DØ. The latter is defined as:

$$A_{SL} = C_d a^d_{sl} + C_s a^s_{sl}, \tag{1.4.1}$$

where coefficients C_d and C_s has been measured by LEP experiments [] and they depend on the mean mixing probability and the production rates of B^0 and B^0_s mesons. The quantities a^q_{sl} (with $q = d, s$) is actually the asymmetry defined in eq. 1.2.14 of Sec. 1.2.2, which can be expressed in terms of $\Delta\Gamma_q$, Δm_q and ϕ_q as follows:

$$a^q_{sl} = \frac{\Delta\Gamma_q}{\Delta m_q} \tan\phi_q. \tag{1.4.2}$$

The departure of the measured value of A_{SL} from SM expectation of about 4σ measured by the DØ collaboration is very unlikely due to NP contributions from B^0 sector, which is well constrained by B–factories measurements. The B^0_s mixing frequency is also measured [?, ?] with very high precision and found consistent with its SM prediction. Nevertheless, the large size of A_{SL} could be accounted for in models that allow NP effects in $\Delta\Gamma_s$ rather than in ϕ_s (and thus β_s). Roughly speaking, in some models it is possible to accommodate the A_{SL} discrepancy and the mild β_s tension properly tuning the value of $\Delta\Gamma_s$. Accurate measurements of $\Delta\Gamma_s$ could constrain the available parameter space and probe the reliability of either the SM prediction (assuming the measurement of A_{SL} is correct) or the DØ measurement (assuming β_s close to the SM expectation).

	SM prediction	CDF	DØ	LHCb
$\Delta\Gamma$ [ps^{-1}]	0.090 ± 0.024	$0.075 \pm 0.035 \pm 0.006$	$0.163^{+0.065}_{-0.063}$	$0.123 \pm 0.029 \pm 0.011$

Table 1.3.: SM prediction for $\Delta\Gamma$ compared with the world's best measurements. For each measurement, the first uncertainity is statistical, the second one (when present) is systematic.

1.4.3. Polarization amplitudes and strong phases

Concerning polarization amplitudes, phenomenological models based on factorization, which provide predictions for B decays into light vector mesons, are not reliable in the case of $B^0_s \to J/\psi\phi$, because of the comparable values of the J/ψ and B_s masses. However, a recent work which relies on the SU(3) symmetry has calculated values of polarization amplitudes and strong phases of $B^0_s \to J/\psi\phi$ decays starting from measured polarizations of $B^0 \to J/\psi K^\star$ [21]. Such a prediction gives the possibility to constrain the strong phases of $B^0_s \to J/\psi\phi$ decays in the β_s measurement, allowing for the resolution of its sign ambiguity. Indeed, previous measurements of $B^0_s \to J/\psi\phi$ support SU(3) symmetry expectations as reported in tab. 1.4.

	$B_s \to J/\psi\phi$	$B_d \to J/\psi K^*$
$\lvert A_0\rvert^2$	0.524± 0.013 ± 0.015	0.566 ± 0.012 ± 0.005
$\lvert A_\parallel\rvert^2$	0.231± 0.014±0.015	0.204 ± 0.015 ± 0.005
$\delta_\perp$	2.95± 0.64 ± 0.07	2.96 ± 0.07 ± 0.05

Table 1.4.: Polarization amplitudes and the strong phase $\delta_\perp$ comparison between $B_s \to J/\psi\phi$ [97] and $B_d \to J/\psi K^*$ [34] modes . For each measurement, the first uncertainity is statistical, the second one (when present) is systematic.

1.4.4. B_s lifetime

As mentioned in Sec. 1.1.2, a correction to Spectator Model predicitons has to be abplied to caluclation of B mesons decay width. This is provided by the Heavy Quark Expansion framework. The latter usually is adopt to predict low energy QCD effects in many flavor observables. HQE can predict the decay–width of B_s^0 mesons to final states common to B_s^0 and $\bar{B}_s^0$, as the case of $B_s^0 \to J/\psi\phi$. Knowning the lifetime of the B^0 meson, the measurement of B_s^0 lifetime provides an accurate way to test the HQE framework as non SM effects are expected to be highly suppressed in lifetimes. In Tab. 5.6 is reported the range predicted by HQE framework, the ratio $\tau(B_s^0)/\tau(B_s^0)$ is predicted and the $\tau(B^0)$ has been taken from the PDG [27].

	Theor. prediction	CDF [97]	DØ [119]	LHCb [40]
$\tau(B_s)[\,\mathrm{ps}]$	$[1.497 - 1.541]$	1.529 ± 0.025 ± 0.012	$1.443^{+0.038}_{-0.035}$	$1.524 \pm 0.020 \pm 0.018$

Table 1.5.: Theoretical prediction of τB_s compared with the world's best measurements. For each measurement, the first uncertainity is statistical, the second one (when present) is systematic.

1.5. Analysis purpose and strategy

The purpose of the analysis presented in this thesis is to update the latest CDF measurements of $\Delta\Gamma$, along with the B_s^0 lifetime and polarization amplitudes of $B_s^0 \to J/\psi\phi$ decays, assuming the SM prediction of β_s. The updated values will be competitive with the world's best measurements for such observables. The reasons of interest in these parameters have already been explained in the previous section.

In addition, this analysis represents a necessary preliminary step before the CDF update of the β_s measurement with the full Run II dataset. The data sample is indeed selected with same requirements as well as the off–line optimezed selection used to maximize the sensitivity on the parameters of interest.

As described in Sect. 1.3, the decay rate of $B_s^0 \to J/\psi\phi$ mode is decomposed in three independent amplitudes, corresponding to the polarizations of the final state vector mesons, which have different CP–parity. This adds more complexity to the analysis, since the extraction of $\Delta\Gamma$ (and β_s as well) is based on the determination of the relative contribution of amplitudes with definite CP–parity as a function of the B_s^0 decay time. In practice, the disentanglement of the polarization amplitudes, *i.e.* of the CP–admixture of the $J/\psi\phi$ final states, is statistically feasible by fitting the angular distribution of muons and kaons transversity angles. For this reason an unbinned maximum likelihood fit is developed. The transversity angles and the B_s^0 decay time are fitted together with the reconstructed B_s^0 mass, which provide a powerful discrimination between signal and backgorund candidates. Since, this maximum likelihood fit represents the starting point to the extension of the analysis for the β_s measurement, the algorithm developed for the identification of the flavor of the B meson at the production time (*flavor tagging*) are also implemented in the likelihood. This allows to get confidence in the flavor tagging technique which is of crucial importance in the next β_s analysis.

Chapter 2

Experimental apparatus

Contents

The measurement described in this thesis uses a data sample collected by the CDF II detector during the Run II of Fermilab's Tevatron collider. This chapter provides a general description of the experimental apparatus, both collider and detector, focusing on the elements most relevant for this analysis.

2.1. The Fermilab Tevatron collider

The Tevatron collider, located at the Fermi National Accelerator Laboratory (FNAL or Fermilab) in Batavia (Illinois, USA) is a proton-antiproton ($p\bar{p}$) superconducting synchrotron at a center-of-mass energy of 1.96 TeV. The Tevatron started its operations in 1985 at a center of mass energy of 1.8 TeV. At the beginning of Run II in 2001 the c.m.s. energy was increased to 1.96 TeV. The upgraded machine collides 36-bunch beams of protons and antiprotons every 396 ns. As shown in Fig. 2.1, the Tevatron complex consists of five major accelerators and storage rings used in successive steps, to produce, store and accelerate the particle beams. The main components are:

- the proton source, including a linear accelerator (LINAC) and a *Booster* synchrotron;
- the *Main Injector*, feeding the antiproton source and the Tevatron collider;
- the antiproton source, including the *Debuncher*, the *Accumulator* and the *Recycler* ring;
- the Tevatron collider;
- the extraction lines to fixed target experiments.

The Run II is the data acquisition period which started in 2001 and was terminated on September 30, 2011.

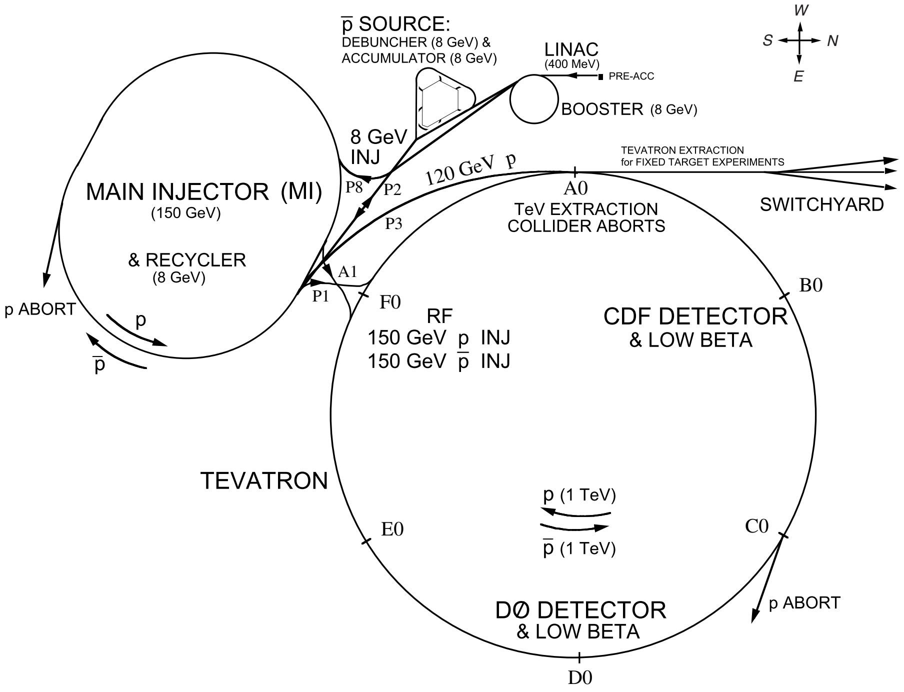


Figure 2.1.: The accelerator system operating at FNAL.

2.1.1. Proton source

The acceleration cycle of the proton beam begins with the production of negatively ionized hydrogen atoms, H^-, which are initially accelerated to a kinetic energy of 750 keV by a Cockroft-Walton electrostatic accelerator. Preaccelerated H^- ions are then injected into the LINAC, where they reach an energy of 400 MeV by traveling through a 150 m long chain of radio-frequency (RF) accelerating cavities. Prior to being injected into the *Booster*, the H^- ions pass through a carbon foil which strips their electrons off. In the *Booster* the protons are accelerated to 8 GeV by a number of RF cavities and then they are transferred to another synchrotron, called *Main Injector* [1], which brings their energy up to 150 GeV. This is the final step before protons are injected into the Tevatron.

2.1.2. Antiproton production

The production of the antiproton beam is significantly more complicated. The cycle starts with the extraction from the *Main Injector* of a 120 GeV proton beam, which is directed onto a Nickel alloy target. The collisions originates a variety of different particles, among which are $\bar{p}$, that are produced with an efficiency of about $18\,\bar{p}/10^6\,p$ through the following reaction:

$$p + \begin{pmatrix} n \\ p \end{pmatrix} \quad \rightarrow \quad p + p + \begin{pmatrix} n \\ p \end{pmatrix} + \bar{p}\,. \tag{2.1.1}$$

The particles, coming off the target at different angles, are focused into a beam line by means of a magnetic lithium collection lens. In order to select only the antiprotons, the beam is sent through a pulsed magnet which acts as a charge-mass spectrometer. The emerging antiprotons, which have a bunch structure similar to that of the incident protons and a large energy spread, are stored in the *Debuncher*, a storage ring where the $\bar{p}$ momentum spread is reduced via stochastic cooling [2] [43]. At the end of the debunching process, the bunch structure is destroyed resulting in a continuous beam of 8 GeV antiprotons which are successively transferred to the *Accumulator*. The *Accumulator* is a triangle-shaped storage ring, housed in the same tunnel as the *Debuncher*, where the antiprotons are further cooled down and stored until all the *Debuncher* cycles are completed. When the collected antiprotons saturate the *Accumulator* acceptance ($\sim 6\times 10^{11}$), they are transferred to the *Recycler* [3], a 8 GeV fixed energy storage ring with a larger ac-

---

[1] Completed in 1999 for Run II, it is located in a 3 km circumference tunnel, which houses also the antiproton *Recycler* and is approximately tangent to the Tevatron.

[2] Stochastic cooling is a technique used to reduce the transverse momentum and energy spread of a particle beam without any accompanying beam-loss. This is achieved by applying iteratively a feedback mechanism that senses the beam deviation from the ideal orbit with a set of electrostatic plates, processes and amplifies the signal, and transmits an adequately-sized synchronized correction pulse to another set of plates downstream.

[3] Antiproton availability is the most limiting factor at the Tevatron for attaining high luminosities: keeping a large antiproton beam inside the *Recycler* has been one of the most significant engineering challenges and the excellent performance of the *Recycler* is an achievement of prime importance for the good operation of the accelerator.

ceptance, made of permanent magnets and placed in the *Main Injector* enclosure. In the *Recycler* the size and spread of the antiproton beam is further shrunk by the electron cooling process: in one of the sections of the Recycler a beam of electrons travels close to the antiprotons at the same velocity, absorbing energy from the antiprotons. When a current sufficient to create 36 bunches with the required density is available, the $\bar{p}$ are injected into the *Main Injector* where they are accelerated to 150 GeV.

2.1.3. Tevatron collider

The Tevatron is 1 km-radius superconducting synchrotron, that accelerates particles from 150 GeV to 980 GeV. The proton and antiproton beams circulate in opposite directions in the same beam pipe. Electrostatic separators produce a strong electric field that keeps the two beams away form each other except at the collision point. The beam is steered by 774 super-conducting dipole magnets and focused by 240 quadrupole magnets with a maximum magnetic field of 4.2 Tesla. A cryogenic system based on liquid helium and nitrogen cools down the Tevatron magnets to 4.2 K, at which temperature the niobium-titanium alloy of the magnet coils becomes superconducting. The process of injecting particles into the machine, accelerating them, and initiating collisions is referred to as a "shot". It starts with the injection from the *Main Injector* of 150-GeV protons, two bunches at a time. Once the proton beam is in the machine, groups of four antiprotonsăbunches are mined from the *Recycler*, accelerated to 150 GeV in the *Main Injector* and injected into the Tevatron. The RF cavities accelerate the beams to 980 GeV, and then some electrostatic separators switch polarity to cause the beams to collide at two points. Each interaction point lies at the center of a particle detector: one named DØ, after its location in the Tevatron optics, and the other named the Collider Detector at Fermilab (CDF) and located at B0. Successively, beams are scraped with remotely-operated collimators to remove the beam halo and, as soon as the beam conditions are stable, the experiments begin to take data. A continuous period of collider operation with the same protons and antiprotons beams is called a "store".

Parameter	Run II value
number of bunches (N_b)	36
revolution frequency [MHz] (f_{bc})	1.7
bunch rms [m] σ_l	0.37
bunch spacing [ns]	396
protons/bunch (N_p)	2.7×10^{11}
antiprotons/bunch ($N_{\bar{p}}$)	3.0×10^{10}
total antiprotons	1.1×10^{12}
β^* $[cm]$	35

Table 2.1.: Accelerator nominal parameters for Run II configuration.

2.1.4. Luminosity and Tevatron performance

The performance of a collider is evaluated in terms of two key parameters: the available center-of-mass energy, $\sqrt{s}$, and the instantaneous luminosity, $\mathcal{L}$. The former defines the accessible phase-space for the production of final state particles. The latter is defined as the interaction rate per unit cross section of the colliding beams (collisions/(cm^{2}s)). In the absence of a crossing angle or position offset, the luminosity at the CDF or DØ is given by the expression:

$$\mathcal{L} = \frac{f_{bc} N_b N_p N_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} F\left(\frac{\sigma_l}{\beta^*}\right), \tag{2.1.2}$$

where f_{bc} is the revolution frequency, N_b is the number of bunches, $N_{p(\bar{p})}$ is the number of protons (antiprotons) per bunch, and $\sigma_{p(\bar{p})}$ is the transverse proton (antiproton) beam size at the interaction point. F is a form factor with a complicated dependence on the beta function value at the interaction point [4], β^*, and the bunch length, σ_l. Tab. 2.1 shows the design Run II accelerator parameters [52]. Fig. 2.2 shows the evolution of the integrated luminosity, defined as $L = \int \mathcal{L}dt$, and the instantaneous luminosity at the start of Tevatron stores during the Run II. The steady increase of the integrated luminosity and the continuous improvement of the instantaneous luminosity prove the outstanding performance of the accelerator. At the end of September 2011, the Tevatron had delivered 12 fb^{-1} per experiment, $\sim$10 of which were collected by the CDF and DØ detectors.

2.2. The CDF II detector

The CDF II detector, in operation since 2001, is an azimuthally and forward-backward symmetric apparatus designed to study the $p\bar{p}$ collisions at the B0 interaction point of the Tevatron. It is a general purpose, cylindrical-shaped detector (Fig. 2.3), which consists of:

- a **tracking system**, which comprises three silicon microstrip trackers (Layer 00, Silicon VerteX II (SVXII) and Intermediate Silicon Layers (ISL)) and an open-cell drift chamber (COT) inside a superconducting solenoid, that provides a constant 1.4 T magnetic field parallel to the beam direction, with the purpose of reconstructing the trajectories (helices) of charge particles and determining their momentum and charge;

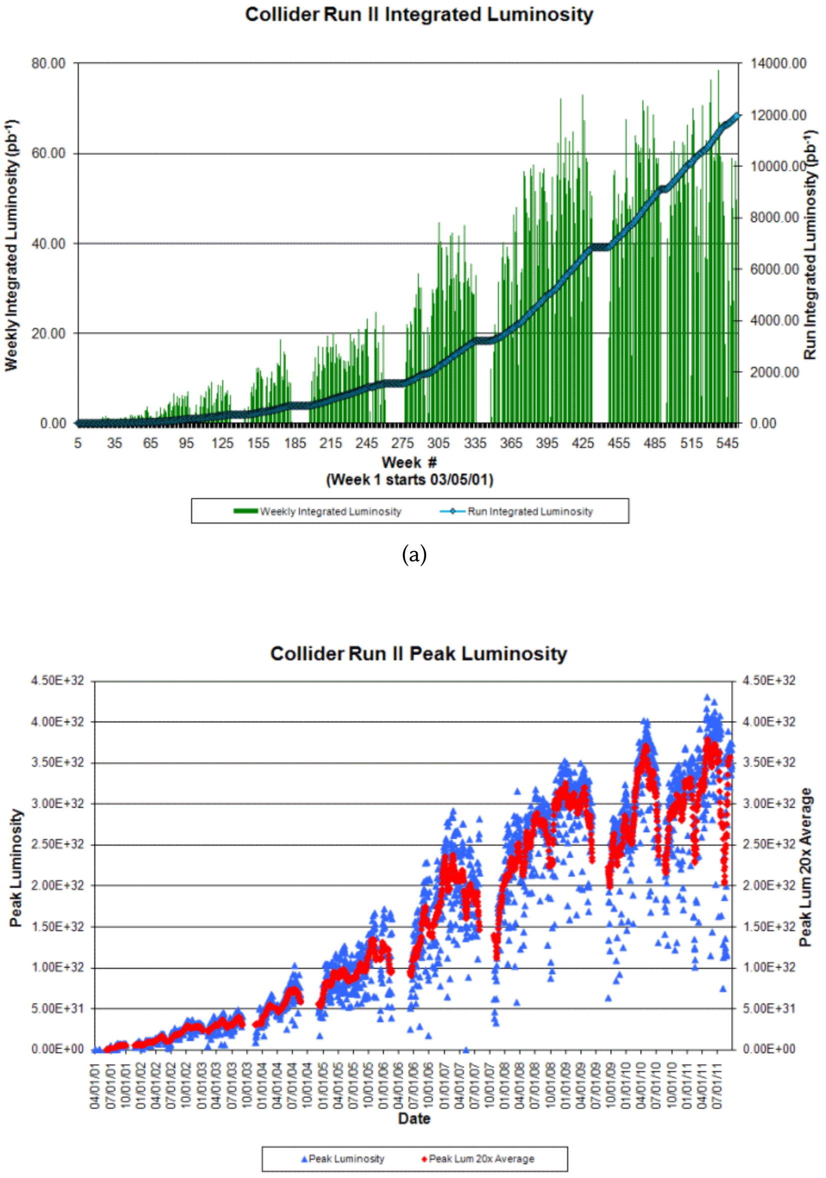
- a **Time of Flight** system (TOF), located outside the COT, for identification of charged particles with momenta up to 2 GeV/c;

- a **calorimeter system**, with the purpose of measuring the energy of charged and neutral particles;

- **muon chambers and scintillators**, used to track and identify muons, that pass through the calorimeters interacting as minimum-ionizing-particles (MIP);

---

[4] The beta function represents a measure of the transverse beam size along the accelerator ring. β^* is the value of this function at the collision point.



(a)

(b)

Figure 2.2.: Integrated luminosity as a function of the Run II weeks (2.2a) and Tevatron peak luminosity as a function of the calendar date (2.2b). Empty periods correspond to Tevatron shut-down periods.

- **luminosity monitors**, for the instantaneous luminosity measurement, necessary to predict event yields.

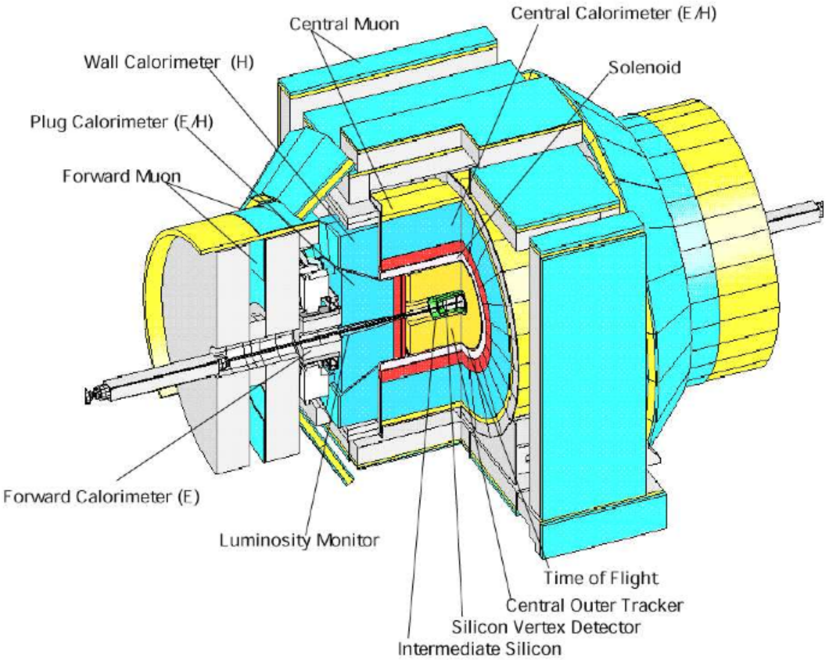


Figure 2.3.: The CDF II detector with a quadrant cut to show the different subdetectors.

2.2.1. Coordinates and Standard Definitions in CDF

CDF adopts a left handed Cartesian coordinate system with origin at the nominal B0 interaction point, coincident with the center of the drift chamber (see Sec. 2.3.2). The positive z-axis lies along the nominal beam-line and has the direction of the proton beam (eastwards). The x-y plane is therefore perpendicular to the beam-line, with the y-axis pointing upward and the x-axis in the horizontal plane, pointing radially outward with respect the center of the accelerator ring (Fig. 2.4). Since the colliding beams of the Tevatron are unpolarized, the resulting physical observations are invariant under rotations around the beam line axis. Thus, a cylindrical (r, ϕ, z) coordinate system is particularly convenient to describe the detector geometry, where

$$r = \sqrt{x^2 + y^2} \quad \text{and} \quad \phi = \tan^{-1}\frac{y}{x} \, .$$

A momentum-dependent particle coordinate, named *rapidity*, is also commonly used in particle physics for its transformation properties under Lorentz boosts. The rapidity is defined as

$$Y = \frac{1}{2}\ln\frac{E + p_z}{E - p_z} \, , \tag{2.2.1}$$

where E is the energy and p_z is the z component of the particle momentum. Rapidity intervals turn out to be Lorentz invariant. In the relativistic limit, or when the mass of the particle is negligible, rapidity depends only upon the production angle of the particle with

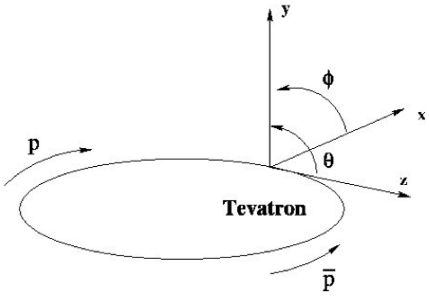


Figure 2.4.: CDF coordinate system.

respect to the beam axis, $\theta = \tan^{-1}\frac{\sqrt{x^2+y^2}}{z}$. This approximation is called *pseudorapidity* η and is defined by

$$Y \overset{p \gg m}{\rightarrow} \eta = -\ln\left(\tan\frac{\theta}{2}\right) .$$

A value of $\theta = 90°$, perpendicular to the beam axis, corresponds to $\eta = 0$. Since the event-by-event longitudinal position of the interaction is distributed around the nominal interaction point with a 30 cm rms width, sometimes a distinction between the detector pseudorapidity (usually indicated with η_{det}), measured with respect to the $(0, 0, 0)$ nominal interaction point, and the event pseudorapidity (η), which is measured with respect to the *z* position of the actual event vertex, is considered. The spatial separation between particles in the detector is commonly given in terms of a Lorentz invariant variable defined as:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} .$$

Other quantities, useful to describe the kinematics of $p\bar{p}$ interactions, are the transverse momentum and the transverse energy, defined as $p_T = p\sin\theta$ and $E_T = E\sin\theta$, respectively.

2.3. The tracking system

A three-dimensional tracking of charged particles is achieved through an integrated system consisting of three inner silicon subdetectors and a large outer drift-chamber, all immersed in the magnetic field of a superconducting solenoid. The silicon detectors provide a precise determination of the track impact parameter, the azimuthal angle and the *z* coordinate, whereas the COT has excellent resolution on the transverse momentum, ϕ and η. The combined information of the tracking detectors provides very accurate measurements of the helical paths of charged particles inside the detector. We will describe this system starting from the devices closest to the beam and moving outwards (see Fig. 2.5).

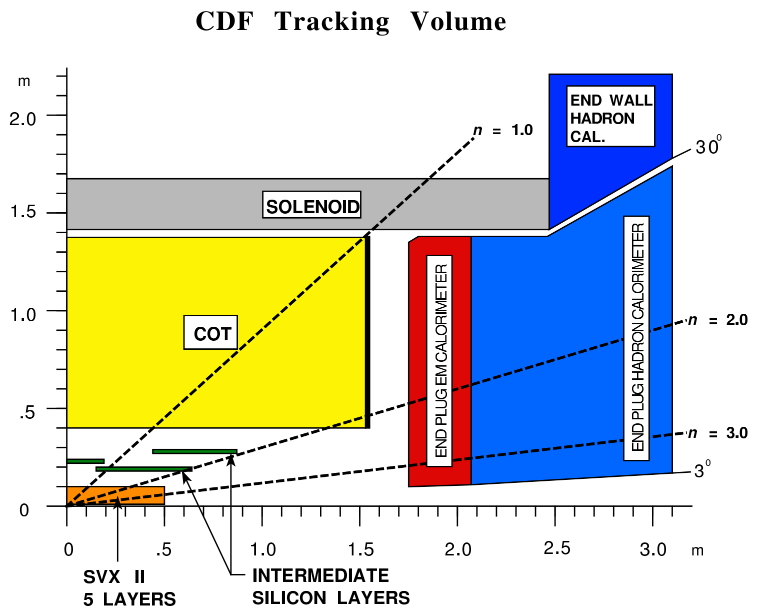


Figure 2.5.: Side view of one quarter of the CDF tracking system.

2.3.1. Silicon tracker

The full CDF silicon detector is composed of three approximately cylindrical coaxial subsystems: the Layer 00 (L00), the Silicon VerteX detector (SVX) and the Intermediate Silicon Layers (ISL). Silicon sensors operate as reverse-biased p-n junctions. When a charged particle passes through the detector, it releases energy and ionizes the material producing electron-hole pairs. If an electric filed is applied the electrons and holes drift to opposite electrodes. The amount of charge collected at the anode and the cathode is proportional to the path length traversed by the charged particle in the detector. By segmenting the p or n side of the junctions into "strips" and reading out the charge deposition separately on every strip we can measure the position of the charged particle. At CDF the typical distance between two strips is about 60μm. There are two types of microstrip detectors: single- and double-sided. In single-sided detectors only one side (p) of the junction is segmented into strips, double-sided detectors have both sides of the junction segmented into strips. In general single-sided sensors have strips parallel to the z direction and provide only r-ϕ position measurements, while double-sided detectors have strips at an angle (stereo angle) with respect to the z direction on one side and, therefore, provide also information on the particle position along z.

Layer 00 (L00) is a 90 cm long, radiation hard, assembly of single sided silicon detectors, structured in longitudinal strips. It is mounted directly on the beam pipe at 1.35-1.62 cm from the beam axis. The detector support structure is in carbon fiber with integrated cooling system. Being so close to the beam, L00 allows to reach a resolution of $\sim$25-30 μm on the impact parameter of tracks of moderate p_T

, providing a powerful handle to identify long-lived hadrons containing a b quark.

SVX is composed of three 29 cm-long cylindrical barrels, radially organized in five layers of double-sided silicon wafers extending from 2.5 cm to 10.7 cm. Each barrel is segmented into 12 wedges, each covering $\sim 30^\circ$ in ϕ. The double-side structure of the wafers allows a three dimensional position measurement: one side of the wafer has axial strips (parallel to the beam), the other one has either 90° strips (perpendicular to the beam) or 1.2° stereo strips (at small angle with respect to the beam). This detector provides position information with a 12 μm resolution on the single hit and some dE/dx ionization information.

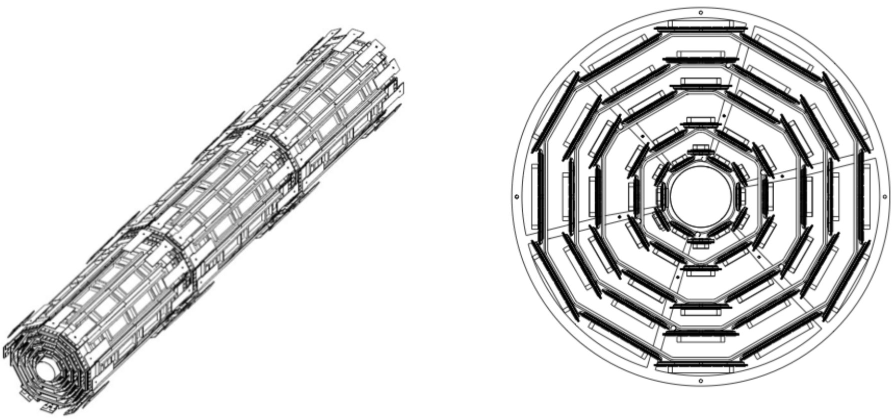
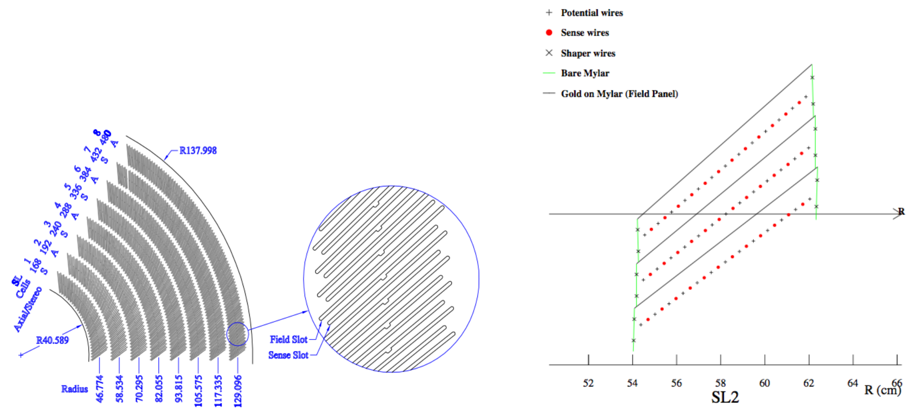


Figure 2.6.: The SVX silicon detector: on the left, a three-dimensional view of the detector allows to see the barrel structure along the beam axes; on the right, the transverse plane section shows in detail the layer sequence.

ISL consists of five layers of double sided silicon wafers, similar to those of SVX, four of which are assembled in a twofold telescopes with planes at a radial distance of 22 cm and 29 cm from the beam-line and covering $1 < |\eta| < 2$. One single central layer is located at r = 22 cm, covering $|\eta| < 1$. The two ISL layers are important to increase the tracking coverage in the forward region, where the COT coverage is limited, and to improve the matching between SVX and COT tracks.

The combined resolution of the CDF inner trackers for high momentum tracks is $\sim 40\mu$m in impact parameter and $\sim$70 μm along the *z* direction. All silicon detectors are used in the off-line track reconstruction algorithms, while SVX plays a crucial role also in the on-line track reconstruction of the trigger system. The CDF trigger employs an innovative processor, the Silicon Vertex Tracker (SVT) [56, 57], which uses the SVX information to measure the track impact parameter on-line with a precision that allows to resolve the secondary vertices, displaced from the primary interaction point, produced in B hadron decays.

2.3.2. COT



(a) COTsector (b) COT Cell

Figure 2.7.: A 1/6 section of the COT end-plate (2.7a): for each super-layer the total number of cells, the wire orientation (axial or stereo), and the average radius in cm are given. The enlargement shows in detail the slot where the wire planes (sense and field) are installed. Fig. 2.7 represents the cross-section of three axial cells in super-layer 2, the arrow indicates the radial direction.

Surrounding the silicon detector is the Central Outer Tracker (COT) [59]. It is a 3.1 m-long cylindrical drift chamber, coaxial with the beam, which covers the radial range from 40 to 137 cm for $|\eta| < 1$. The COT contains 96 sense wire layers, which are radially grouped into 8 superlayers, as inferred from the end plate slot structure shown in Fig. 2.7. Each superlayer is divided in ϕ into supercells, and each supercell has 12 alternated sense and field shaping wires. So within the supercell width the trajectory of a particle is sampled 12 times. The maximum drift distance is approximately the same for all superlayers. Therefore, the number of supercells in a given superlayer scales approximately with the radius of the superlayer. The entire COT contains 30240 sense wires. Approximately half the wires run along the *z* direction ($axial$), the other half are strung at a small stereo angles (2°) with respect to the *z* direction ($stereo$). The combination of the axial and stereo information allows to measure the *z* positions and a three-dimensional reconstruction of tracks. Particles originated from the interaction point, which have $|\eta| < 1$, pass through all the 8 COT superlayers.

The COT is filled with an Argon-Ethane gas mixture and Isopropyl alcohol (49.5:49.5:1). The mixture is chosen to have a constant drift velocity, approximately 50 μm/ns, across the cell width. When a charged particle passes through a chamber cell, the gas is ionized and pairs of free electrons and positive ions are created. The electric field inside the cell attracts the electrons towards the sense wires. As the electrons get closer to the wires the field intensity become more and more intense until, eventually, an avalanche multiplication of charge occurs due to electron-atom collisions, providing a gain of $\sim 10^4$. The movement of charges in the cell electric field induces a signal on the sense wire, a“hit”, which is read out by electronics. The maximum electron drift time is approximately $100\,\mathrm{ns}$. Due to the magnetic field that the COT is immersed in, electrons drift at a Lorentz angle of 35°. The supercells are tilted by 35° with respect to the radial direction to compensate for this effect and make the drift path perpendicular to the radial direction.

Signals on the sense wires are processed by the ASDQ (Amplifier, Shaper, Discriminator with charge encoding) chip, which provides input projection, amplification, pulse shaping, baseline restoration, discrimination and charge measurement [63]. The pulse is sent through $\sim 11\,\mathrm{m}$ of micro-coaxial cable, via repeater cards, to Time to Digital Converter (TDC) boards in the collision hall. The pulse leading edge gives the arrival time information and the pulse width, in nanosecond, is related to the amount of charge collected by the wire. After calibrating the width variations due to the COT geometry, to the path length of the associated track, to the gas gain differences for the 96 wires, the Landau associated to the track is determined, using the amount of the charge collected (in nanoseconds) for each hit along the track path length. From the Landau distribution the energy loss is measured and used for particle identification. A detailed description of the calibration is found in [64], [65]. The TDC boards contain also a buffer where the data are stored while waiting for the events to be accepted by the trigger.

The hit position resolution in the r-ϕ plane is about 140 μm. Tracking algorithms are utilized to reconstruct particle trajectories (helices) that best fit to the observed hits. The reconstructed trajectories are referred to as “tracks”. Particle momentum and charge are determined from the bending of tracks in the magnetic field. The COT hits are also processed on-line by the eXtremely Fast Tracker (XFT), which reconstructs the tracks used in the trigger system, (Sec. 2.7). The transverse momentum resolution of off-line tracks, estimated using cosmic ray events, is:

$$\frac{\sigma_{p_T}}{p_T^2} = 0.017\,[\mathrm{GeV/c}]^{-1} \tag{2.3.1}$$

for tracks with $p_T > 2$ GeV/c [60]. All the tracks that we use in this analysis are required to have the COT and the SVX information.

2.3.3. Track reconstruction

Charged particles traveling through a homogeneous solenoidal magnetic field along the z direction follow helical trajectories. Knowing that the projection of the helix on the x-y plane is a circle, to uniquely parametrize a helix in three dimensions, five parameters are needed:

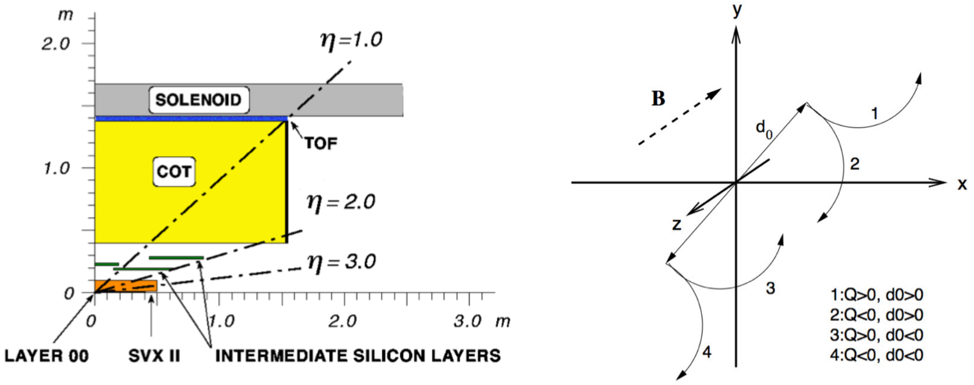
$\boldsymbol{C}$ – signed helix (half)-curvature, defined as $C = q/2R$, where R is the radius of the helix and q is the particle charge. This is directly related to the transverse momentum. When the magnetic field (B) is measured in Tesla, C in m^{-1} and p_T in $\mathrm{GeV/c}$: $p_T = 0.15\,qB/|C|$.

$\boldsymbol{\phi_0}$ – ϕ azimuthal angle of the particle trajectory at the point of closest approach to the z-axis.

d_0 – signed impact parameter, i.e. the radial distance of closest approach to the *z*-axis. defined as $d_0 = q(\sqrt{x_0^2 + y_0^2} - R)$, where (x_0, y_0) are the coordinates of the center. This is schematically drawn in Fig. 2.8b.

λ – helix pitch, i.e. $\cot(\theta)$, where θ is the polar angle of the particle at the point of its closest approach to the *z*-axis. This is directly related to the longitudinal component of the momentum: $p_z = p_T \cot\theta$.

z_0 – the *z* coordinate of the point of closest approach.



(a) The CDF II tracker layout showing the different subdetector systems.

(b) Schematic drawing of the impact parameter d_0.

Figure 2.8.: CDF tracking subdetectors and impact parameter d_0 definition.

Another useful quantity is the displacement of the secondary vertices of decaying particles in the transverse plane, L_{xy}:

$$L_{xy} = \frac{\widehat{x}_V \cdot \vec{p}_T}{|p_T|} \tag{2.3.2}$$

where $\widehat{x}_V$ is the decay vertex position in the transverse plane. In the following we will call L_{xy} "transverse decay length".

The trajectory of a charged particle satisfies the following equations [68]:

$$\begin{aligned} x &= r sin\phi - (r - d_0) sin\phi_0 \\ y &= -r cos\phi + (r + d_0) cos\phi_0 \\ z &= z_0 + s\lambda \end{aligned} \tag{2.3.3}$$

where s is the projected length along the track, $r = 1/2C$ and $\phi = 2Cs + \phi_0$. The reconstruction of a charged particle trajectory consists in determining the above parameters through an helical fit of a set of spatial measurements ("hits") reconstructed in the tracking detectors by clustering and pattern-recognition algorithms. The helical fit takes into account field non-uniformities and scattering by the detector material. All tracks are first fit in the COT and then extrapolated inward the silicon. This approach guarantees fast and efficient tracking with high tracks purities. The greater radial distance of the COT with respect to the silicon tracker results in a lower track density and consequent fewer accidental combination of hits in the track reconstruction. A brief overview of the tracking algorithms is given in the following, for more details see Ref. [69], [70].

Tracking algorithms

Using the hit positions in the tracking system, pattern recognition algorithms reconstruct the particle original trajectory measuring the five parameters of the helix that best match to the observed hits.

CDF employs several algorithms for track reconstruction, depending on which component of the detector a particle travels through. The principal one is the Outside-In (OI) reconstruction [66]. This algorithm, which exploits the information from both the central drift chamber and the silicon detectors, is used to track the particles in the central region ($|\eta| < 1$). It first reconstructs tracks in the COT and then extrapolates them inwards toward the beam.

The first step of pattern recognition in the COT looks for circular paths [5] in the axial superlayers. Cells in the axial superlayers are searched for sets of 4 or more hits that can be fit to a straight line. These sets are called “segments”. Once segments are found, there are two approaches to track finding [67] (“segment linking” and “histogram linking” algorithms). One approach is to link together the segments which are consistent with lying tangent to a common circle. The other approach is to constrain its circular fit to the “beam-line” (see Sec. 2.3.3). Once a circular path is found in the r-ϕ plane, segments and hits in the stereo superlayers are added depending on their proximity to the circular fit. This results in a three-dimensional track fit. Typically, if one algorithm fails to reconstruct a track, the other algorithm will not. This results in high track reconstruction efficiency in the COT for tracks passing through all 8 superlayers (97% for tracks with $p_T > 10\,\mathrm{GeV/c}$) [6].

Once a track is reconstructed in the COT, it is extrapolated inward to the silicon system. Based on the estimated errors on the track parameters, a three dimensional “road” is formed around the extrapolated track. Starting from the outermost layer, and working inwards, silicon hits found inside the road are added to the track. As hits are added, the road gets narrowed according to the knowledge of the updated track parameters and their covariance matrix. A reduction of the road width decreases the chance of adding wrong hits to the track, and also reduces the computation time. In the first pass of this algorithm, axial hits are added. In the second pass, hits with stereo information are added to the track. At the end, the track combination with the highest number of hits and lowest χ^2/ndf for the five parameters helix fit is kept.

---

[5] The helical track, when projected onto the r-ϕ plane, is a circle.

[6] The track reconstruction efficiency mostly depends on how many tracks are reconstructed in the event. If there are many tracks close to each other, hits from one track can shadow hits from the other track, resulting in efficiency losses.

Due to the limited COT coverage and the tight requirement on the hits, tracking in the forward region requires different algorithms [71], [72] that are not described here because the obtained tracks are not used in this analysis.

Primary vertex reconstruction

The primary vertex position for a given event is found by fitting high quality tracks to a common point of origin. At high luminosities, multiple collisions occur on a given bunch crossing. For a luminosity of $10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, there is an average of 2.3 interactions per bunch crossing. Typically, since the luminous region is sufficiently long (with $\sigma_z = 29\,\mathrm{cm}$), the primary vertices associated to the collisions are well separated in *z*. An iterative algorithm is used to find the vertex associated to the hardest collision: the first estimate of its position (x_V, y_V, z_V) is binned in the *z* coordinate, then the *z* position of each vertex is calculated from the weighted average of the *z* coordinates of all tracks within $1\,\mathrm{cm}$ of the first iteration vertex, with a typical resolution of $100\,\mu\mathrm{m}$; finally the vertex associated with the highest sum of the tracks p_T is defined as primary vertex of the event.

The locus of all primary vertices defines the beam-line, the position of the luminous region of the beam-beam collisions through the detector. The beam-line is used as a constraint to refine the knowledge of the primary vertex in a given event. Typically the beam transverse section is circular with a width of $\sim\ 30\ \mu$m at $z\ =\ 0$, rising to $\sim$ $50 - 60\ \mu m$ at $|z|\ =\ 40\,\mathrm{cm}$. The beam is not necessarily parallel nor centered in the detector and moves as a function of time.

2.4. Time-Of-Flight detector

Just outside the tracking system, supported on the inside of the solenoid, see Fig. 2.8a, CDF II has a Time of Flight (TOF) detector [61]. It is a barrel of scintillator bars almost 3 m long, located at 140 cm from the beam line with a total of 216 bars, each covering 1.7° in ϕ and the pseudorapidity range $|\eta| < 1$. Light is collected by photomultipliers at both ends of each scintillator bar. Single hit position in the Time Of Flight detector (TOF) is determined by the comparison of the signal times of the photomultipliers. Particle identification is achieved by measuring the time of arrival of a charged particle at the scintillators with respect to the collision time. Thus, combining the measured time-of-flight, the momentum and the path length the mass of the particle can be estimated:

$$m = \frac{p}{c}\sqrt{\frac{c^2t^2}{L^2} - 1} \qquad (2.4.1)$$

where the momentum p and the path length L are precisely measured by the tracking system. For the TOF measurement the collision time must be known and this is found with a 50 ps uncertainty by a best-fit process over all tracks in the event. The resolution in the time-of-flight measurement is ~120 ps and provides at least two standard deviation separation between $K^{\pm}$ and $\pi^{\pm}$ for momenta $p < 1.6$ GeV/c [7].

2.5. Muon detectors

Most of the particles produced in the primary interaction or in subsequent decays have a very high probability of being absorbed in the calorimeter system. Muons represent an exception. They are over 200 times more massive than electrons, so bremsstrahlung radiation, inversely proportional to the mass squared of the incident particle, is suppressed by a factor of $4 \cdot 10^4$ with respect to electrons. Muons do not interact via strong interaction with nuclei in matter either. Therefore, a muon with enough energy will pass through the calorimeter systems releasing only a small amount of its energy. At CDF the minimum muon energy required to reach the muon detectors, placed radially outside of the calorimeters, is $1.4\,\mathrm{GeV}$. In addition to the calorimeters, steel absorbers are placed upstream of the muon systems to reduce punch-through hadrons. The muon system is the outermost layer of the CDF II detector and consists of four layers of drift cells and scintillation counters which are used to reconstruct track segments (“stubs”) of minimum ionizing particles. These stubs are matched using dedicated algorithms with the COT information in order to reconstruct the full trajectory of the muons. Some additional steel shielding layers, in between the chambers and the calorimeters, reduce the probability for other particles to escape the calorimetric system. Four independent systems detect penetrating charged particles (muons) in the $|\eta| \lesssim 1.5$ pseudo-rapidity range reconstructing a small segment of their path (stub) sampled by the chambers, employing similar combinations of drift tubes, scintillation counters, and absorbers with differential azimuthal coverage [74], [75]. The momentum measurement is performed by pointing back the stub to the corresponding track in the COT. Scintillators serve as trigger and vetoes while the drift chambers measure the ϕ coordinate using the absolute difference of drift electrons arrival time between two cells, and the *z* coordinate by charge division. All type of muon detectors use a single wire, rectangular drift chambers, arranged in arrays with various azimuthal segmentation and coupled with scintillator counters. The chambers use a 50:50 gas admixture of Argon and Ethane, and operates in proportional regime. The four sub-detector systems are (see Fig. 2.9):

Central MUon detector (CMU): the CMU detector is located around the central hadronic calorimeter at a radius of $347\,\mathrm{cm}$ from the beam-line with coverage $0.03 \lesssim |\eta| \lesssim 0.63$. It is segmented into 24 wedges of $15°$, but only $12.6°$ in ϕ, with a gap of $2.4°$, of each wedge is active, resulting in an overall azimuthal acceptance of 84%. Each wedge is further segmented into three $4.2°$ modules each containing four layers of four drift cells.

Central Muon uPgrade (CMP): the CMP is a second set of muon drift chambers outside of CMU with an additional $60\,\mathrm{cm}$-thick steel absorbers between them. The material

---

[7] The expected separation power is defined as $\frac{TOF_i(p)-TOF_j(p)}{\sigma_{TOF}}$, where $TOF_i(p) = \frac{L}{c}\sqrt{\frac{m_i c^2}{p^2}+1}$ is the expected time of fight for the i particle of mass m_i and momentum p, σ_{TOF} is the time of flight resolution.

further reduces the probability of hadronic punch-through to the CMP. Muons need a transverse momentum of about $2.2\,\mathrm{GeV}$ to reach the CMP. The CMP system is arranged in a box shape of similar acceptance as the CMU and conventionally serves as a confirmation of CMU for higher momentum muons. A layer of scintillation counters (CSP) is mounted on the outer surfaces of the CMP. The CMP and CMU have a large overlap in coverage and are often used together. CMP helps to cover CMU ϕ gaps and the CMU covers the CMP η gaps. Muon candidates which have both CMU and CMP stubs are the least contaminated by fake muons.

Central Muon eXtension (CMX): the CMX consists of drift tubes and scintillation counters (CSX) assembled in conically arranged sections. The CMX extends the pseudo-rapidity coverage to $0.6 \lesssim |\eta| \lesssim 1$. There are 8 layers of drift chambers in total with a small stereo angle between layers.

Intermediate MUon System (IMU): the IMU extends the pseudo-rapidity coverage even further to $1.0 \lesssim |\eta| \lesssim 1.5$. The IMU is mounted on the toroid magnets which provide shielding and consists of Barrel MUon chambers (BMU), Barrel Scintillation counters (BSU) and Toroid Scintillation counters (TSU).

Tab. 2.2 summarizes a few of the relevant design parameters of there detectors.

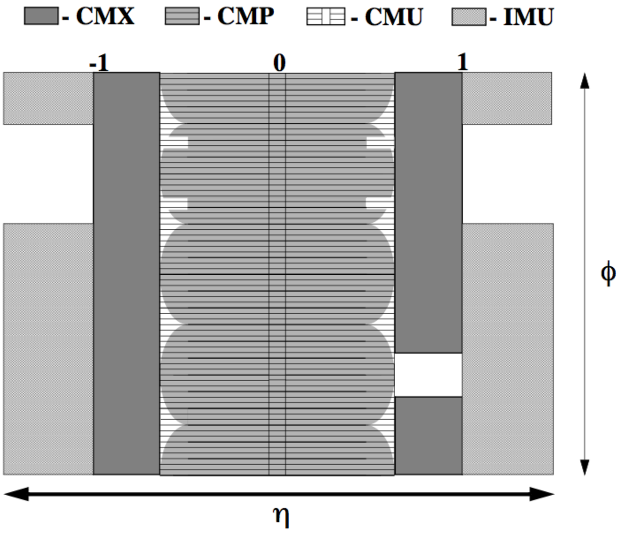


Figure 2.9.: Muon detectors coverage in the η-ϕ plane.

2.6. Other Detectors

This section will provide a brief description of the sub-detectors that have not been used in this analysis.

Parameter	CMU	CMP	CMX	IMU
Pseudorapidity range	$\|\eta\| < 0.6$	$\|\eta\| < 0.6$	$0.6 < \|\eta\| < 1.0$	$1.0 < \|\eta\| < 1.5$
Azimuthal coverage [°]	360	360	360	270
Maximum drift time [ns]	800	1400	1400	800
Drift tube cross section[cm]	2.68 × 6.35	2.5 × 15	2.5 ×15	2.5 × 8.4
Pion interaction length	5.5	7.8	6.2	62 - 20.0
Minimum $p_T(\mu)$ [GeV/c]	1.4	2.2	1.4	1.4-2.0

Table 2.2.: Design parameters of the muon detectors. Assembled from Ref. [74], [75]

2.6.1. Calorimetry system

Although it is not used in this analysis, the calorimeter system, together with the muon and tracking systems, represents one of the main sub-detector apparatuses of CDF II detector. A detailed description of this system can be found in the CDF II Technical Design Report [52]. The CDF II calorimetry system has been designed to measure energy and direction of neutral and charged particles leaving the tracking region. In particular, it is devoted to jet reconstruction and it is also used to measure the missing energy associated to neutrinos. Particles hitting the calorimeter can be divided in two classes, according to their main interaction with matter: electromagnetically interacting particles, such as electron and photon, and hadronically interacting particles, such as mesons or barions produced in hadronization processes. To detect these two classes of particles, two different calorimetric parts have been developed: an inner electromagnetic and an outer hadronic section, providing coverage up to $|\eta| < 3.64$. In order to supply information on particle position, the calorimeter is also segmented in towers, projected toward the geometrical center of the detector. Each tower consists of alternating layers of passive material and scintillator tiles. The signal is read out via WaveLength Shifters (WLS) embedded in the scintillator and light from WLS is then carried by light guides to photomultiplier tubes. The central sector of the calorimeter, covering the region $|\eta| < 1.1$, was recycled from Run I, while brand new calorimeters (called plug calorimeters) were built up to cover the forward and backward regions. Fig. 2.10b shows the plug calorimeter system while Fig. 2.10c shows an elevation view of the components of the CDF calorimeter: CEM, CHA, WHA, PEM and PHA.

The central calorimeter

Apart from upgrades on the readout electronics, needed to cope with the increased collision rate, the central calorimeter is almost the same as in Run I. The Central ElectroMagnetic calorimeter (CEM) is segmented in $\Delta\eta \times \Delta\phi = 0.11 \times 15°$ projective towers consisting of 31 alternate layers of lead and scintillator, for a total material depth of 19 X_0 [8]. The Central and end-Wall HAdronic calorimeters (CHA and WHA respectively),

---

[8] The radiation length X_0 describes the characteristic amount of matter transversed by high energy electrons to lose all but $1/e$ of their energy due to bremsstrahlung, which is equivalent to 7/9 of the length

whose geometry tower segmentation matches the CEM one, use 32 steel layers sampled each 2.5 cm by 1 cm thick acrylic scintillator. The total thickness of the hadronic section is approximately constant and corresponds to 4.5 interaction lengths (λ_0) [9]. A perspective view of a central electromagnetic calorimeter module (*wedge*) is shown in Fig. 2.10a, where both the arrangement in projected towers and the light-gatering system are visible. The projective geometry has been used in order to take advantage of the momentum conservation in the transverse plane: before the $p\bar{p}$ collision, the projection in the transverse plane w.r.t. the beam direction of the beam energy is zero, therefore this quantity have to be the same also after the collision took place. Thus, for each tower the transverse energy E_T is defined as $E_T = E \sin\theta$, where E is the energy detected by the tower and θ is the angle between the beam axis and the tower direction, in the CDF detector coordinates system. Two position detectors are embedded in each wedge of CEM:

- The Central Electromagnetic Strip chamber (CEntral Strip multiwire proportional chambers (CES)) is a two-dimensional stripwire chamber arranged in correspondence to maximum shower development ($\sim 5.9X_0$). It measures the charge deposit of the electromagnetic showers, providing information on their pulse-height and position with a finer azimuthal segmentation than calorimeter towers. This results in an increased purity on electromagnetic object reconstruction. The CES purpose is to measure the position and the shape of electromagnetic showers in both transverse plane and longitudinal direction, which is used to distinguish electrons and photons from hadrons.

- The Central Pre-Radiator (Central Pre-Radiator (CPR)) consists of two wire camber modules placed immediately in front of the calorimeter. It acts as pre-shower detector by using the tracker and the solenoid coil material as radiators, resulting to be a very useful tool in rejection of electron and photon background.

Calorimeter response is fast enough to match the time requirements imposed by Run II. However, wire chambers associated to CES and CPR may need to be integrated over several beam crossings; this will not be a problem since the high granularity of these devices guarantees a low detector occupancy. Tab. 2.3 summarize the basic quantities of calorimeter detectors. The energy resolution for each calorimeter section was measured in the test beam and, for a perpendicularly incident beam, it can be parametrized as:

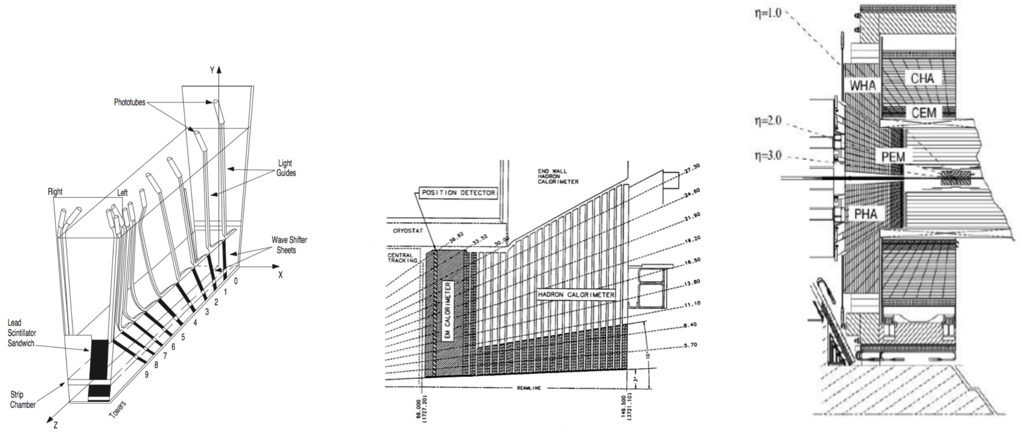
$$\frac{\sigma}{E} = \frac{\sigma_1}{\sqrt{E}} \oplus \sigma_2$$

where the first term comes from the sampling fluctuations and the photostatistics of the PMTs, and the second term comes from the non-uniform response of the calorimeter.

---

of the mean free path for e^+e^- pair production of high energy photons. The average energy loss due to bremsstrahlung for an electron of energy E is related to the radiation length by $\left(\frac{dE}{dx}\right)_{brem} = -\frac{E}{X_0}$

[9] An interaction length is the average distance that a particle will travel before interacting with a nucleus: $\lambda = \frac{A}{\rho\sigma N_A}$, where A is the atomic weight, ρ is the material density, σ the cross section and N_A the Avogadro number.



(a) CEM Wedge (b) Plug calorimeters (c) Whole calorimeter system

Figure 2.10.: One azimuthal electromagnetic calorimeter wedge 2.10a, the elevation view of one quarter of the plug calorimeter 2.10b. In 2.10c elevation view of the CDF detector showing the components of the CDF calorimeter: CEM, CHA, WHA, PEM and PHA.

The plug calorimeter

The plug calorimeter, covers the η region from 1.1 to 3.64. The new configuration, based on the same principles as the central calorimeter, allows the detector to operate in the Run II environment and makes experimental data more homogeneous. Both electromagnetic and hadronic sectors are divided in 12 concentric η regions, with $\Delta\eta$ ranging from 0.10 to 0.64, according to increasing pseudorapidity, each of them is segmented in 48 or 24 (for $|\eta| < 2.11$ or $|\eta| > 2.11$ respectively) projective towers. The actual size of these towers was chosen so that identification of electron in b-jets would be optimized. Projective towers consist in alternating layers of absorbing material (lead and iron for electromagnetic and hadronic sectors, respectively) and scintillator tiles. The first layer of the electromagnetic tile is thicker ($10\,\mathrm{mm}$ instead of $6\,\mathrm{mm}$) and made of material with higher photon yield. They act as a pre-shower detector.

Calorimeter	CEM	CHA	WHA	PEM	PHA
Absorber	Lead	Steel	Steel	Lead	Iron
Segmentation ($\eta \times \phi$)	0.1×15	0.1×15	0.1×15	(0.1 ÷ 0.6)×(7.5 ÷ 15)	(0.1 ÷ 0.6)×(7.5 ÷ 15)
Num. Towers ($\eta \times \phi$)	20×24	9×24	6×24	12×24(48)	11×24(48)
Thickness	$19\,\mathrm{X}_0$,$1\,\lambda_0$	$4.7\,\lambda_0$	$4.7\,\lambda_0$	$23\,\mathrm{X}_0$,$1\,\lambda_0$	$6.8\,\lambda_0$
Resolution (%)	$14/\sqrt{E_T} \oplus 2$	$50/\sqrt{E_T} \oplus 3$	$75/\sqrt{E_T} \oplus 4$	$16/\sqrt{E} \oplus 1$	$80/\sqrt{E} \oplus 5$

Table 2.3.: Summary of the main characteristics of the CDF II calorimeter system.

2.6.2. CLC detector and instantaneous luminosity measurement

Absolute measurements of the instantaneous luminosity by the machine, based on beam parameters measurements, have uncertainties of the order of 15-20%. For this reason in CDF, the beam luminosity is determined using gas Cherenkov counters (CLC) [76] located in the pseudorapidity region$3.7 < |\eta| < 4.7$, which measure the average number of inelastic interactions per bunch crossing. Each module consists of 48 thin, gas-filled, Cherenkov counters. The counters are arranged around the beam pipe in three concentric layers, with 16 counters each pointing to the center of the interaction region. The cones in the two outer layers are about 180 cm long and the inner layer counters, closer to the beam pipe, have a length of 110 cm. The Cherenkov light is detected with photomultiplier tubes and the momentum threshold for light emission is 9.3 MeV/c for electrons and 2.6 GeV/c for pions. The number of $p\bar{p}$ interactions in a bunch crossing follows a Poisson distribution with mean μ, where the probability of empty crossing is given by:

$$P_0(\mu) = e^{-\mu}$$

which is correct if the acceptance of the detector and its efficiency were 100%. In practice, there are some selection criteria, α, to define an “interaction.” An interaction is defined as a $p\bar{p}$ crossing with hits above a fixed threshold on both sides of the CLC detector. Therefore, an empty crossing is a $p\bar{p}$ crossing with no interactions. Given these selection criteria, the experimental quantity P_0 , called $P_0^{exp}\{\alpha\}$, is related to μ as:

$$P_0^{exp}\{\mu, \alpha\} = (e^{\epsilon_w \cdot \mu} + e^{-\epsilon_e \cdot \mu} - 1) \cdot e^{-(1-\epsilon_0)\cdot\mu}$$

where the acceptances ϵ_0 and $\epsilon_{w/e}$ are, respectively, the probability to have no hits in the combined east and west CLC modules and the probability to have at least one hit exclusively in west/east CLC module. The evaluation of these parameters is based on Monte Carlo simulations, and typical values are respectively 0.07 and 0.12. From the measurement of μ we can extract the luminosity. Since the CLC is not sensitive at all to the elastic component of the $p\bar{p}$ scattering, the rate of inelastic $p\bar{p}$ interactions is given by:

$$\mu \cdot f_{bc} = \sigma_{in} \cdot \mathcal{L}$$

where f_{bc} is the bunch-crossing frequency at the Tevatron and σ_{in} is the inelastic $p\bar{p}$ cross section. $\sigma_{in} = 60.7 \pm 2.0$ mb is obtained by extrapolating the combined results for the inelastic $p\bar{p}$ cross section of CDF at $\sqrt{s} = 1.8$ TeV and E811 measurements to $\sqrt{s} = 1.96 TeV$ [77]. Different sources of uncertainties are taken into account to evaluate the systematic uncertainties on the luminosity measurement [78]. The dominant contributions are related to the detector simulation and the event generator used, and have been evaluated to be about 3%. The total uncertainty in the CLC luminosity measurements is 5.8%, which includes uncertainties on the measurement (4.2%) and on the inelastic cross section value (4%).

2.7. Trigger and data acquisition

At the typical Tevatron instantaneous luminosity $\mathcal{L} \approx 4 \cdot 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, and with an inelastic $p\bar{p}$ cross-section of $\sigma_{p\bar{p}} \approx 60\,\mathrm{mb}$, approximately $2.5 \cdot 10^{7}$ inelastic collisions per second occur, corresponding to one inelastic $p\bar{p}$ interaction per bunch-crossing on average [10]. Since the read-out of the entire detector needs about $2\,\mathrm{ms}$ on average, after the acquisition of one event, another approximately 5000 interactions would remain unrecorded. When an event recording is prevented because the system is busy with a different event or a different task, this is called *dead-time*.

In this way the average size of information associated to each event is $140\,\mathrm{Kb}$. Even in case of deadtime-less read-out of the detector, in order to record all events, an approximate throughput and storage rate of $350\,\mathrm{Gb/s}$ would be needed, largely beyond the possibility of currently available technology [11].

The read-out system has to reduce the $2.3\,\mathrm{MHz}$ interaction-rate to the $100\,\mathrm{Hz}$ storage rate attainable at CDF. The challenge for the whole system is to cut-off events that don't have the minimal requirements to be reconstructed or seem to contain well-known processes, that don't need further study, focusing the acquisition system on the interesting processes. The Fig. 2.11 shows a scheme to explain how the information flows through the different parts.

To suppress unwanted events, the CDF trigger system is segmented in three levels, each level receiving the accepted events from the previous one, and, provided with detector information with increasing complexity and with more time for processing, determines if one of a set of existing criteria is verified by the event.

Prior to any trigger level, the bunched structure of the beam is exploited to reject cosmic-ray events by gating the front-end electronics of all sub-detectors in correspondence of the bunch crossing. The front-end electronics of each sub-detector, packaged in Vesa Modul Eurocard (VME) modules hosted in about 120 crates, has a 42-cells deep pipeline synchronized with the Tevatron clock-cycle set to $132\,\mathrm{ns}$. The Tevatron clock picks up a timing marker from the synchrotron Radio Frequency cavity (RF) and forwards this bunch-crossing signal to the trigger and to the front-end electronics. Since the inter-bunch time is $396\,\mathrm{ns}$, three times the Tevatron clock-cycle, the pipeline can collect data corresponding to a maximum of 14 bunch crossings. The pipeline depth gives the amount of time that Level-1 (L1) trigger has to decide to accept or reject an event otherwise the buffer content is overwritten: $396\,\mathrm{ns} \cdot 14 = 5.5\,\mu\mathrm{s}$. An event accepted by the Level-1 is then passed to the Level-2 (L2) buffer, where the number of buffers in the pipeline is 4, that gives $5.5\,\mu\mathrm{s} \cdot 4 = 22\,\mu\mathrm{s}$. This means that if an event is accepted by the L1 and the L2 doesn't have a free buffer deadtime will incur. Level-3 (L3) is composed by a computer farm, the L2 output rate is low enough to avoid in general deadtime problem in the connection between L2 and L3.

The following description will emphasize the aspect of the CDF Trigger that are related with the selection of rare events including b-hadrons with high purity.

---

[10] Abort gaps can be neglected for this estimate.

[11] The maximum current storage rate is approximately $250\,\mathrm{Kb/s}$

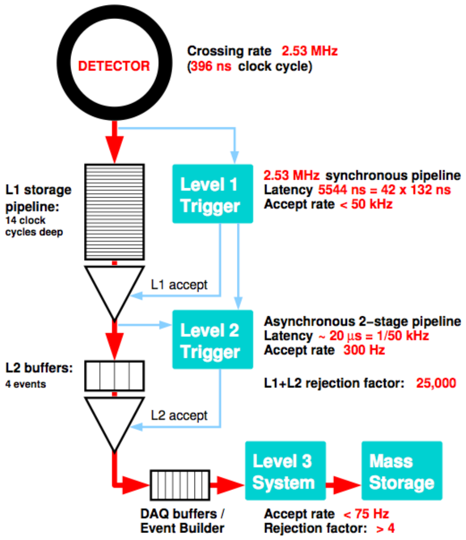


Figure 2.11.: Functional block diagram of the CDF II trigger and data acquisition systems.

LEVEL 1 is a synchronous system with event reading and a decision made every beam crossing. In a synchronous pipeline up to 42 subsequent events can be stored for $\sim 5.5\mu\,\mathrm{s}$ while the hardware is taking a decision. If no acceptance decision is made within that time the event is lost. L1 decision are made in average in about 4 μ s: no dead time is expected from this level. Level 1 rejects 97% of the events, it reduces the event rates from 2.53 MHz to less than 50 kHz. The L1 decision is generated by:

- XFT (extremely fast tracker), which reconstructs approximate tracks (pT >1.5 GeV) in the transverse plane by exploiting information from COT superlayers. These tracks are extrapolated to the other detector parts to contribute to all trig- ger levels.
- the calorimeter trigger, which indicates large energy releases in single electromagnetic or hadronic cells (these can be seed for electrons or jets identification).
- the muon trigger, which matches XFT tracks segments in the muon chambers.

The XFT is a custom processor used to identify two-dimensional tracks in the (r, ϕ) plane in the COT. The XFT is capable of reconstructing tracks with $p_T \gtrsim 1.5\,\mathrm{GeV}$ with an efficiency of about 95% and a fake rate of a few percent. The XFT has an angular segmentation of $1.25°$, and an angular resolution of $0.3°$. The momentum resolution is $\sigma_{p_T}/p_T^2 \approx 0.017\,[\mathrm{GeV/c}]^{-1}$. XFT sends the tracks to the extrapolation unit (XTRP) which feeds three L1 elements: L1 CAL, L1 TRACK, and L1 MUON. L1 CAL and L1 MUON use extrapolated tracks and information from the calorimetry and muon systems respectively to search for possible electron, photon, jets and muon candidates. A decision stage combines the information from these low-resolution physic objects, called “primitives”, into more sophisticated objects, e.g., track primitives are matched with muon stubs or tower primitives, to from muon, electron, or jet object, which are subject to basic selection. The accepted events are buffered for L2 analysis

LEVEL 2 is an asynchronous system which processes events that have received a L1 accept in FIFO (First In, First Out) manner. It is structured as a two stage pipeline with data buffering at the input of each stage. The first stage is based on dedicated hardware processor which assembles information from a particular section of the detector. The second stage consists of a programmable processors operating on lists of objects generated by the first stage. Each of the L2 stages is expected to take approximately 10 μ s with a latency of approximately 20 μ s. The L2 buffers can store up to four events. After the Level 2, the event rate is reduced to about 300 Hz (rejection factor ∼150). L2 purposes are:

- to add the energy deposited in the towers in small regions around L1 seeds, as an approximate measure of an electron or jet energy.
- to use calorimeter and CES chamber information to improve separation of $e^{\pm}$ from γ.
- to reconstruct a full COT track and associate it to an outer muon stub in order to improve muon signature.
- to feed startup information to the Silicon Vertex Trigger (SVT) [?] [?] which generates triggers on secondary vertexes from decay of long-lived heavy flavour hadrons.

Silicon Vertex Trigger (SVT) uses SVX $r - \phi$ hits to extend XFT track primitives inside the SVX volume, closer to beam-line. The SVT improves the XFT ϕ_0 and p_T resolutions and adds the measurement of the impact parameter d_0 (original XFT track primitives are beam-line constrained). Acting into the impact parameter it is a very useful handle in order to select decay modes of heavy b-hadrons. The impact parameter of decay products is strongly related to the decay length of the mother b-hadron, therefore a selection based on the tracks impact parameter turns directly in to a proper time requirement. This innovative system is the core of most of the trigger systems for B physics.

LEVEL 3 is a software trigger. L3 addresses event objects delivered by L2 to the Event Builder (EVB) [79], which reconstructs the entire event with the same accuracy as in the off-line analysis. The final decision to accept an event is made on the basis of a list of observables indicating candidate events of physical interest (top quark production events, W/Z events, Drell-Yan events, etc.). Events that satisfy the Level 3 trigger requirements are then transferred onward to the Consumer Server/Data Logger (Consumer Server/Data Logger (CSL)) system for storage first on disk and later on tape. The average processing time per event in Level 3 is on the order of one second. The Level 3 leads to a further reduction in the output rate, of about 50 Hz in average, with an accepted maximum of about 120 Hz.

A set of requirements that an event has to fulfill at Level 1, Level 2 and Level 3 constitutes a trigger path. The CDF II trigger system implements about 150 trigger paths, which are periodically adjusted depending on machine luminosity and physics needs. An event will be accepted if it passes the requirements of any one of these paths and, depending of the trigger path, and it will be stored in a trigger dataset. A typical list of the datasets at CDF Run II can be found in [80]. The trigger system described above exploits the information of all detector subsystem. Combining the measurements of the various subsystems it is possible to efficiently record, at the same time, events characterized by different signatures. Triggers which occupy most of the band width can be dynamically prescaled (DPS). A trigger path is said to be prescaled by a factor N if it is configured to accept only one event each N accepted events. Prescaling is dynamically implemented by luminosity-dependent factors during data taking. This is important in order to ensure that no trigger path reaches rates so high as to create unacceptable dead time to triggers on rare events of primary importance. During data taking the luminosity decreases with time, and consequently a number of prescale factors can be relaxed. The prescale factors decrease proportionally to the rate of triggered events, so as the number of recorded events is constant. Using dynamic prescaling ensures that optimal use is made for physics of the available luminosity.

The CDF II trigger system implements about 200 trigger paths. An event will be accepted if it passes the requirements of any one of these paths and, depending of the trigger path, it will be stored in a trigger dataset.

2.8. The di-muon trigger

All the analysis samples used in this thesis were acquired with the CDF di-muon trigger. The di-muon trigger relies on a clear signature of two muons coming from $J/\psi \to \mu^+\mu^-$ decays. In order to make trigger decisions, it uses the XFT tracking and muon system information available at Level 1. Levels 2 and 3 play a small role in the event selection decision. Level 2 is used to tighten any existing requirements of Level 1, e.g. on the transverse momentum, and Level 3 uses more precise determination of several event variables, such as the transverse momentum of tracks, better track-stub matching, di-muon mass, etc. Because the di-muon trigger plays such an important role in this analysis, we present a detailed description in the following. Although we refer to it as a single entity, the di-muon trigger is in fact a combination of two triggers: CMU-CMU, where both muons are found in the most central muon chamber, and CMU-CMX, where one muons is found in the CMU and one in CMX. We describe the CMU-CMU trigger, and then comment on the differences in the CMU-CMX. The following terminology is specific to triggering on CMU muons. A “stack” is a set of four drift cells stacked on top of each other. The CMU has 288 stacks in each of the East and West sides of the detector. A Level 1 “stub” is a track segment in a stack such that cells 1 and 3 or cells 2 and 4 have hits separated by no more than 396 ns (“or” is the mathematical $\vee$: the statement is true when one or both are true). A tower is a set of two neighboring stacks. A tower fires when one or both stacks have a Level 1 stub, and is empty otherwise. A muon tower is a fired tower matched with an XFT track. In order to keep the Level 1 decision time short enough to remain synchronous, only information about which towers have fired is used in triggering, rather than detailed hit positions and direction. The XFT provides the p_T and ϕ_6 (ϕ at COT superlayer 6), as well as the charge of the track to the XTRP. The XTRP extrapolates this track to the CMU radius and creates a footprint, a 3 σ window in ϕ (wide enough to account for multiple Coulomb scattering). If a tower is found within that footprint, it is a muon tower. The CMU-CMU trigger requires that at least two muon towers be found such that they are either on opposite sides of the detector or are separated by at least two other towers. The CMU-CMX trigger uses much of the same algorithm. The changes to the decision algorithm arise from the differences between the CMU and CMX detectors. In the CMU-CMX case, only XFT tracks with $p_T > 2.2$ GeV/c are used to match to the CMX as the extra material that muons pass through to reach the CMX limits further the momentum requirements on the muon, and no azimuthal separation is required because the muons are by definition in different subdetector volumes. Trigger algorithms are among the few elements of the experimental apparatus which are continuously improved and optimized, as this process does not require performing expensive and time-consuming hardware upgrades to the detector. The di-muon trigger has undergone constant revision in order to carry out such optimization. While the core logic outlined above is more or less constant, other parameters have often been changed to improve the trigger performance. Such parameters include requirements on the p_T of the XFT tracks, the difference in ϕ between the two muons $\Delta\phi$, and their transverse mass M_T . In addition, some of the triggers are prescaled (see Sec. 2.7). The various combinations of these requirements result in slightly different trigger requirements.

Chapter 3

Data sample and preparation of the analysis tools

Contents

In this chapter the focus will be on the data sample selection. Once the events have been reconstructed, the selection of signal and the rejection of background are performed by a loose preselection at first stage, and subsequently by a more refined procedure based on a Neural Network (NN). Variables involved in the selection process are kinematical variables and quantities based on Particle Identification (PID), therefore an overview of both of them *will be given. A section of this chapter describes the strategies used to identify the* B_s *meson flavor at the production, since this is, as will be shown in the next chapter, a relevant aspect of this measurement. Last section of the chapter will be dedicated to explain how is obtained the simulated events sample used to compute the detector acceptance.*

3.1. Data sample

The data used for for our analysis has been collected by the CDF experiment from March 2001 to March 2011, corresponding to an integrated luminosity of $\mathcal{L} = 8.3\,\mathrm{f}b^{-1}$. Two decay modes have been reconstructed, with two different purposes:

1. $B_s \to J/\psi[\to \mu^+\mu^-]\phi[\to K^+K^-]$ to measure β_s and used in this thesis
2. $B^+ \to J/\psi[\to \mu^+\mu^-]K^+$ used to calibrate the OST

The online selection of these decays has been performed by the di-muon triggers, which select the events through the identification of the $\mu^+\mu^-$ pair,as produced by the decay of the J/ψ particle. The triggers are able to reject a large portion of the background but at the offline level a further selection is applied and comprises three stages: decay reconstruction, preselection cuts and NNprocedure.

3.2. Trigger requirements

The first set of requirements imposed on the events of our analysis is enforced by the trigger system in real time as data is being collected at CDF. The di-muon triggers are a collection of more than 50 trigger paths [1] which have evolved over time to cope with the more demanding running conditions mainly driven by the increase in luminosity. The basics requirements of a basic di-muon trigger are summarized as follows:

1. Level 1
    - two XFT tracks with opposite charge
    - matching of each track with two muon stubs
    - each CMU (CMX) muon has $p_T^{XFT} > 1.5(2.2)\,\mathrm{GeV}/c^2$
    - $\Delta\phi_6(CMU, CMU) < 135°$ for some paths; no cut in $\Delta\phi_6(CMU, CMX)$
2. Level 3
    - $2.7 < M_{\mu\mu} < 4\,\mathrm{GeV}/c^2$

---

[1] Keeping into account all the trigger path currently used and all the trigger path that have been used in the past during the Run II

3.3. Event recontruction and variables definitions

3.3.1. Event reconstruction

Track reconstruction

The reconstruction of a charged-particle trajectory consists in the estimation of the parameter summarized in ?? through an helical fit of a set of spatial measurements along the track (the so called "hits") as identified in the tracking detectors by clustering and pattern-recognition algorithms. The helical fit takes into account magnetic field, detector non-uniformities and the particle scattering in the material. The trajectory parameters are extracted from a fit to hits of the COT and SVX.

J/ψ recontruction

In the di-muon triggers only the events with opposite sign dimuon pairs are retained. These pairs are fitted to a common vertex estimated vertex position, the vertex constrained tracks and a χ^2 value are obtained .The J/ψ invariant mass and $p_T(J/\psi)$ are then estimated from the refitted tracks.

ϕ recontruction

In order to find the ϕ candidates, oppositely charged pairs of non-muon tracks coming from a displaced vertex fitted by a kinematic fitting algorithm [86] are examined if they are within events containing a J/ψ. The two tracks are initially assumed to be kaons, at a subsequent stage a probability to be kaons is assigned. This probability con be computed by exploiting the Particle Identification method as dE/dx and TOF.

Decay recontruction

In our events the $primary\ vertex$ is the point of the $p\bar{p}$ collision, while the $secondary\ vertex$ is the decay point of a particle produced in the event. In the case of $B_s \to J/\psi\phi$ the secondary vertex is far away from the collision point, due to the long lifetime of the B_s meson. Actually the primary vertex is the point of the B_s meson production, while the secondary one is where the B_s meson decays.

Data Format

We use the `BStntuple` framework [89], which is an extension of the Stntuple [90] framework developed in CDF to minimize the computation time and storage space. The `BStntuple` contains structures that hold the reconstructed candidates informations, the stable and decaying objects, as well as information needed for flavor tagging (decision and raw dilution), and particle identification information (TOF, dE/dx, muon and electron quantities). The format allows several potential decay candidates in the same event to share links to common data blocks.

3.3.2. Variables definition

After the 4-track vertex fit has been performed, the best fit values of the B_s and daughter particles momenta are obtained. From these quantities, one can calculate variables that are relevant for our analysis the decay length (or the proper decay time) with the associated event-by-event uncertainty. Secondly the angular variables (described in 1.3) that are used in the separation of the CP eigenstates.

Proper decay time

Let us define $\vec{L_x y}$ as the displacement in the transverse plane of the secondary vertex from the primary vertex. ct, the proper decay time, is then defined as:

$$ct = c\frac{\vec{L_{xy}} \cdot \vec{p_T}}{|\vec{p_T}|^2}M \tag{3.3.1}$$

where $\vec{p_T}$ is the particle (in this case of the B meson) transverse momentum and M is its world average.

σ_{ct} uncertainty on the proper decay time

σ_{ct} is referred to as the event-by-event uncertainty associated to the proper decay time. To compute this uncertainty, only the measured error on $\vec{L_{xy}}$ is used, and we assume that the other quantities in eq. 3.3.1 give a negligible contribution.

Transversity angles

The transversity angles (see Fig.) are obtained by boosting the four-momenta of the decay particles into the B_s rest frame. The boost vector is computed by using the world average B_s mass, and the reconstructed $p_T(B_s)$. Then we need to boost into the J/ψ rest frame and the world average J/ψ mass is used, together with the reconstructed $p_T(J/\psi)$. In the case of the ϕ the reconstructed K^+K^- mass and momentum are used. In the case of the ϕ meson, the natural width of the particle is close to the resolution of the CDF detector; this is the reason why the reconstructed mass is used. In the case of the B_s and J/ψ, the particles have a natural width that is much smaller than the resolution, so the usage of the world average values for the masses is a more accurate method.

Let $\vec{p}_A^B$ be the three-momentum of the particle A in the rest frame of the particle B. Starting from the transversity angle ψ, $\cos\psi$ is defined as:

$$\cos\psi = -\frac{\vec{p}_{K+}^{\phi} \cdot \vec{p}_{J/\psi}^{\phi}}{|\vec{p}_{K+}^{\phi}| \cdot |\vec{p}_{J/\psi}^{\phi}|} \tag{3.3.2}$$

In order to calculate the other two angles, a reference system is defined:

$$\begin{aligned}
\hat{x} &= \frac{\vec{p}_{\phi}^{\,J/\psi}}{|\vec{p}_{\phi}^{\,J/\psi}|} \\
\hat{y} &= \frac{\vec{p}_{K+}^{\,J/\psi} - (\vec{p}_{K+}^{\,J/\psi} \cdot \hat{x})\hat{x}}{|\vec{p}_{K+}^{\,J/\psi} - (\vec{p}_{K+}^{\,J/\psi} \cdot \hat{x})\hat{x}|} \\
\hat{z} &= \hat{x} \times \hat{y}
\end{aligned} \tag{3.3.3}$$

Therefore the other two angles of the transversity basis are obtained as:

$$\begin{aligned}
\cos\theta &= \frac{\vec{p}_{J/\psi}^{\,\mu^+}}{|\vec{p}_{J/\psi}^{\,\mu^+}|} \cdot \hat{z} \\
\phi &= \mathsf{tan}^{-1}\Big(\Big(\frac{\vec{p}_{J/\psi}^{\,\mu^+}}{|\vec{p}_{J\psi}^{\,\mu^+}|} \cdot \hat{y}\Big)/\Big(\frac{\vec{p}_{J/\psi}^{\,\mu^+}}{|\vec{p}_{J/\psi}^{\,\mu^+}|} \cdot \hat{x}\Big)\Big)
\end{aligned} \tag{3.3.4}$$

3.4. Particle ID

Particle IDentification (PID) plays an important role in two different aspects of this analysis. The first is the signal sample selection obtained by means of a NN. PID provides us a component of the discriminating variables that are used in input for the NN. The second aspect is the the kaon particle identification necessary for the flavor tagging. Two quantities are used in order to separate kaons from pions: the Time of Flight, obtained from the dedicated TOF detector, and the the rate of energy loss through ionization (dE/dx) in the gas that fills the COTactive volume (see sec. 2.3.2).

3.4.1. Charged particle ionization energy loss (dE/dx)

A charged particle moving through matter loses energy due to ionization, according to the Bethe Bloch formula:

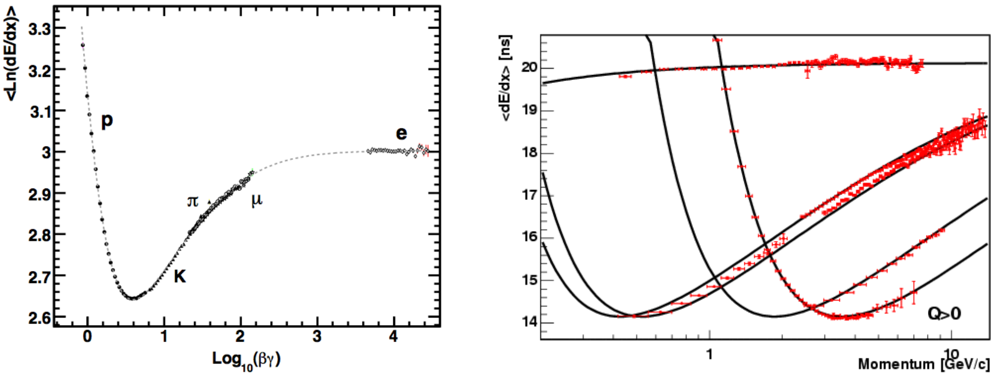
$$\left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi N e^4}{m_2 c^2 \beta^2} q^2 \left[\ln \frac{2 m_e c^2 (\beta\gamma)^2}{I^2} - \beta^2 \right] \tag{3.4.1}$$

where N is the electron density in the material considered, e is as usual the electron charge, m_e is the electron mass, q is the particle's charge, βc the particle speed and γ the Lorentz factor $\gamma = 1/\sqrt{1-\beta^2}$. I indicates the mean excitation energy of the material. SInce the gas of the COT is a material with known properties, eq. 3.4.1 , together with a momentum measurement, leads to the measurement of the incident particle's mass. Therefore, by comparison with the known masses of the particles, the type of the particle can be determinated. In order to model more realistically the CDF detector, the relation

3.4.1 have been empirically modified to give [87]

$$\left\langle \frac{dE}{dx} \right\rangle = \frac{1}{\beta^2}\left[c_1 \ln \frac{\beta\gamma}{b+\beta\gamma} + c_0\right] + a_1(\beta-1) + a_2(\beta-1)^2 + C \tag{3.4.2}$$

where a_i, b, c_j are parameters left free to float when fitting the data. From eq. 3.4.2 can be plotted the dE/dx behavior in function of $\beta\gamma$ (the so called $universal\ curve$). Moreover dE/dx can be plotted in function of the momentum of the incident particles (see for example Fig. 3.1), and this puts into evidence the ability to separate different particles by using the dE/dx.



(a) Universal curve: Log(dE/dx) versus Log($\beta\gamma$)

(b) Momentum dependence of the dE/dx

Figure 3.1.: PID by using dE/dx: universal curve and momentum dependence. 3.1a shows the relation between Log (dE/dx) and Log($\beta\gamma$) of reconstructed particles at CDF. Different ranges in $\beta\gamma$ are dominated by different particle species. Momentum dependence of the dE/dx for, in 3.1b from left, muons, pions, kaons, protons, and electrons (top) at CDF [88]. The amount of ionization charge produced by a charged particle near a COT sense wire is proportional to the signal strength on the wire. dE/dx is measured as the amount of charge, which is proportional to the width of the pulse from the read out chip Δt, whether this charge is above a certain threshold value. For this reason dE/dx values are given in time units instead o energy units.

PID making use of dE/dx is important because the identification of particles at $low - p_T$ at CDF is difficult, since the detector was mostly designed for high-pT physics measurements and not optimized for these capabilities. The TOF is the only subdetector entirely devoted to this function, but its performances are marginal for particles with momenta greater than$2\,\mathrm{Ge}V/c$ and in high-luminosity collisions. A part from the TOF information, only for particles with $p_T < 800\,\mathrm{Me}V/c$ some specific ionization from the silicon tracker has some effective identification power. For charged particles with $p_T > 2\,\mathrm{Ge}V/c$ from the dE/dx in the gas of the COT is effective.

The measurement capability of the COT for dE/dx is affected by effects due to both environmental and kinematic variables. These dependences reduce the identification power, as a consequence, in order to optimize the PID, the measurement of dE/dx must be calibrated, to remove or minimize these dependences. The calibration uses pure sample of

XFT-triggered kaons and pions from D^*- tagged D^0 decays and involves several stages [91]. Firstly all possible variables that can affect the dE/dx have to be explored (both environmental- and track-related quantities), ad out of these the set of x_i variables that have larger effect on the measurement have to be identified. The subsequent and most important stage is the removal of the dE/dx dependence on each x_j variable while taking into account dependencies between variables, which may require joint correction of multiple effects. This is done determining a set of "correction functions". One more step is the extraction of the ionization curve, i. e., the function that describes the expected average dE/dx for a charged particle as a function of its Lorentz boost ($\beta\gamma$), and the subsequent determination of the dE/dx resolution, on which depends the separation between pions and kaons. Finally the dE/dx correlations can be extracted.

3.4.2. Time of Flight

As mentioned in earlier sections, the Time of Flight is a complementary quantity to dE/dx, and it is the most powerful method to separate pions from kaons at low momenta ($p < 2\,\mathrm{GeV}/c$). It is useful to recall that PID with TOF exploits, together with a momentum measurement (obtained by using the COT) the relation:

$$m = \frac{p}{c}\sqrt{\frac{c^2 t^2_{flight} - 1}{L}} \tag{3.4.3}$$

where m is the mass of the particle that needs to be identified, L is the track length, p is the probed particle momentum. t_{flight} is the time of flight, obtained by subtracting from the time when the particles hits the TOF scintillators (that can be called t_{TOF}), the collision instant (t_0). Exactly as the dE/dx, also the TOF can be affected by the detector and the event kinematics. Therefore also the TOF measurement needs to be calibrated, in order to optimize the separation between particle types. The full calibration method is described in [92].

The expected separation power[2] for the various particle species that is achievable with TOF alone, assuming $\approx 110\,\mathrm{ps}$ for the time of flight resolution, as a function of momentum is shown in Fig. 3.2b. For comparison, the expected K/π separation from the COT dE/dx measurement is also shown to illustrate the complementary power of COT with respect to the TOF particle identification.

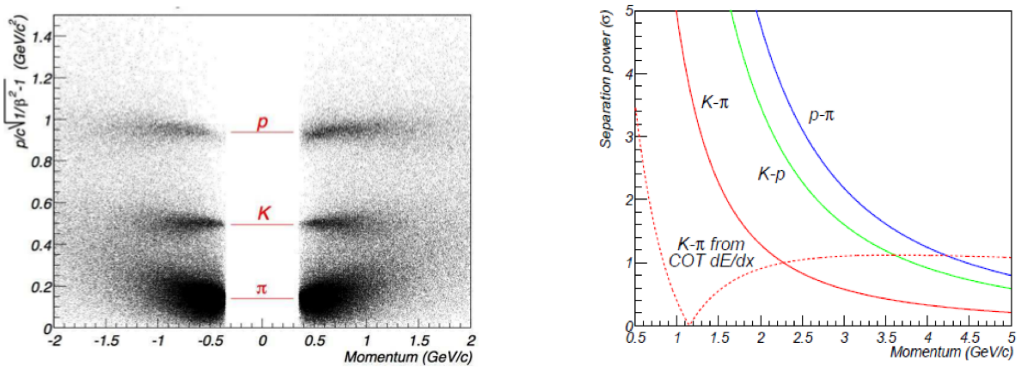
3.5. Event selection

3.5.1. Pre-selection requirements

Signal and background events are distinguishable by using kinematic variables characterizing the event. If a comparison between a certain variable distribution in the signal

---

[2] The expected separation power is defined as $\frac{TOF_i(p) - TOF_j(p)}{\sigma_{TOF}}$ where $TOF_i(p) = \frac{L}{c}\sqrt{\frac{m_i c^2}{p^2} + 1}$ is the expected time of fight of the i particle of mass m_i and momentum $p\sigma_{TOF}$ is the time of flight resolution.



(a) TOF distributions for different particles

(b) Separation power of TOF for different particles

Figure 3.2.: PID by using TOF: distribution and separation power.
3.2a TOF distributions for different particles 3.2bSeparation power of TOF for different particles at CDF, with dE/dx separation power for kaon and pion from COT superimposed.

and in the mass sidebands region is made, the discriminating power between signal and background of that variable is related with the differences between the signal and background distributions.
If we consider a mass plot as Fig. 3.8b, what we call $signal\ region$ is the peak, and the $sidebands$ are the two regions on the left and on the right of that peak, at a distance from the peak, where the contribution of the signal is estimated to be negligible. The aim is to select two classes of events enriched either of signal or background events. Using the sidebands of events, it is possible to study the behavior of the background events. With the assumption that the background present in the signal region has the same behavior of the background in the sidebands, it is possible to subtract the signal events from the data distributions and therefore obtain signal-only (or sideband-subtracted) distributions. In the comparisons, we adopt the following nomenclature: the side-band subtraction is the operation performed on the real data subtracting from the events, in a window around the signal region in the invariant mass histogram of the Bs candidate, [5.340; 5.393] GeV/c^2, those events which are located in the sidebands. The two sidebands are: the left one from 5.287 to 5.314 GeV/c^2 and the right one from 5.419 to 5.445 GeV/c^2.

Selections of the signal events and rejection of background ones can be implemented e. g. with the following strategies: the application of a set of cuts (that are called $rectangular\ cuts$) on different variables and the implementation of a procedure using a NN. According to the first method, the value of a single variable "decides" if the event is kept in the data sample or not. The second method instead is able to take into account also the discriminating power of the single variables, and the correlation between the variables. The first stage of event selection (called $pre-selection$) of the data sample used in this thesis, has been implemented making use of loose rectangular pre-selection cuts in order to reduce the sample size; afterwards the final selection is achieved by applying the NN. These cuts are defined as follows:

- Track quality: At least 10 axial and 10 stereo COT hits
  for kaon tracks + At least 3 Si hits for all tracks
- (5.1 $< M(B) <$ 5.6) GeV/c^2
- $Pt(k_1)$ >0.4 GeV/c and $Pt(k_2) >$ 0.4 GeV/c
- $Pt(\phi)$ >1 GeV/c
- (3.014 $< M(J/\psi) <$ 3.174) GeV/c^2
- (1.009 $< M(\phi) <$ 1.028) GeV/c^2
- $\chi^2_{r\phi} <$ 50[3]
- $Pt(B) >$ 4 GeV/c

The purpose of the pre-selection cuts is to eliminate most of the background events from the data sample, and at the same time to avoid rejecting signal events. In other words, one wants to improve the purity [4] of the sample, while keeping the same efficiency. This goal is achieved by keeping cuts "loose", meaning that some contamination of background event is accepted in our sample. An additional advantage that is gained is that the obtained sample is of a significant smaller size, improving the computational speed of the subsequent stage.

3.5.2. Neural Network selection

Neural Networks are more and more used in various fields for data analysis and classification, both by research and by commercial institutions. In particle physics they are used mainly for classification tasks, i.e. signal over background discrimination. NNs are non-linear statistical data modeling tools, used to model complex relationships between inputs and outputs or to find patterns in data [96]. A Neural Network is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information, using a connectionist approach to computation. The key element is the novel structure of the information processing system, which consists of an interconnected group of artificial neurons (this is why they are named networks), working in unison to solve specific problems.

A NN is comprised of an input layer, a hidden layer and an output layer. Each variable such as mass, p_T, etc. correspond to an *input node* (or *neuron*). Each input note is

---

[3] $\chi^2_{r\phi}$ indicates the value of the χ^2 for the fit on the vertex in the transverse plane. see 3.5.2

[4] Given a data sample where N_S is the number of signal events and N_B is the number of background events, any kind of possible selection produces a subsample of M_S signal events. The purity is defined as the number of events identified as the ratio between Signal events by the selection that really are signal events divided by M_S.

assigned a different weight in the NN's hidden layer, to produce an output variable O_{NN} that is used as discriminant. The output variable O_{NN} ranges between 0 (in our case meaning that the event is background-like) and 1 (signal-like). A node receives the inputs $(x_0, x_1 ... x_n)$ from the input layer, adds them in a weighted sum to generated an output $f(x_0, x_1 ... x_n)$. Where a standard form for $f(x)$ is $f(x_0, x_1 ... x_n) = \frac{1}{1+exp[-\sum_{i=0}^{i=n} w_i x_i]}$, but in the most general case is a monotonic function satisfying $\lim_{x\to-\infty} f(x) = 0$ and $\lim_{x\to+\infty} f(x) = 1$. In the case that a step function is used, the neural network reduces to a special case of rectangular cuts.

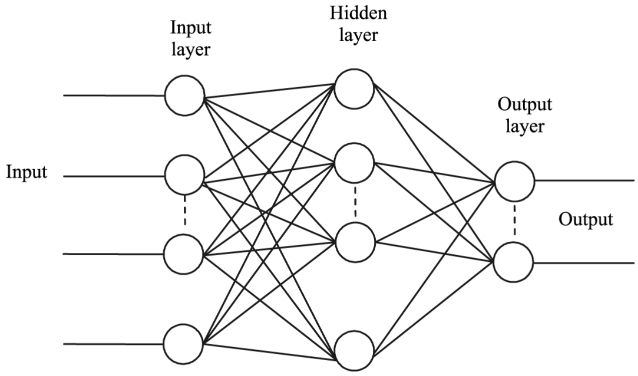


Figure 3.3.: Conceptual sketch of a NN with one input, one hidden, one output layer.

NNs, like people, learn by examples. The examples must be selected carefully otherwise time may be wasted or even worse the network might be not properly working. The disadvantage of the usage of NNs is that, because the network finds out how to solve the problem by itself, its operation can be unpredictable. The procedure used to "teach" to the NN how to discriminate between signal and background events is called $training$ and exploit training algorithms, and representative data samples used as training samples. The training of a NN consists in assigning weights w_i to each node such that they can be combined to form an output (decision) node that discriminates between signal and background. The training process consist of an iterative optimization to minimize the distance between a target vector of decision the NN should make and an output vector of decision that the NN does make at each stage of the optimization.[98]

The main disadvantages of using Rectangular Cuts originate from the rigidity of the requirements: if one single variable falls outside the accepted range, even by a small amount, the event has to be rejected. Moreover neither the discriminating power nor the correlations of the variables are kept into account, and this implies a reduction in the efficiency and in the purity of the data sample. Therefore choosing a NN method allows to limit these difficulties. A NN combines the information from all the kinematic distributions into a single output variable, that denotes whether an event is signal-like or background-like. This output variable, O_{NN}, assumes values between -1 and 1, where events with O_{NN} close to -1 are classified as background and events with O_{NN} close to 1 as signal (see Fig. 3.7). A weight is assigned to each kinematic variable in input to the NN and ite represents the magnitude of the variable contribution to the NN output. The weight associated to a certain variable is proportional to its discriminating power. In this way even a single variable is allowed to accept or reject an event, thus improving the efficiency of both the background rejection and the signal acceptance. In addition, rectangular cuts are more difficult to be optimized, while the usage of a NN allows to use a large of variety of optimization methods. This a method should be chosen focusing on the measurement purposes. A common criteria is to maximize a quantity like the Signal-to-Background ratio (S/B, with S number of signal events and B number of background ones) or analogous quantities (as $S/\sqrt{S+B}$), but if the aim of a measurement is to estimate a specific parameter with as good as possible resolution, an optimization on the NN output cut that minimizes the statistical error on that parameter measurement, could be the best option.

NN input variables

The following variables are used as input to the NN:

- $\chi^2_{r\phi}$ - the χ^2 of the two dimensional vertex fit in the transverse plane.
- $P(\chi^2, p)$ - χ^2 probability for the three dimensional vertex fit.
- $p_T(p)$ - Momentum component transverse to the beam direction for particle p.
- $LL_{\mu\mu}(p)$ - Value for a likelihood based quantity used for muon identification. [94]
- $LL_K(p)$ - Value of likelihood based discriminant for particle identification. It is constructed based on the dE/dx and TOF informations. [93]

They are listed in order of decreasing discriminating power and relevance to the final discriminant: the transverse momentum p_T of the ϕ meson, the kaon likelihood [93] based on TOF and dE/dx information, the muon likelihood [94] for the J/ψ muon daughters, $\chi^2_{r\phi}$ for the B_s decay vertex reconstruction, the transverse momentum p_T of the B_s meson, and the probabilities to reconstruct vertices from the B_s, ϕ, and J/ψ candidates. The muon and kaon likelihoods are quantities used for particle identification. The algorithm determining the muon likelihood is described in Ref. [94]. The kaon likelihood [93] is a combined discriminant constructed from the kaon track specific energy loss, dE/dx, and its time-of-flight information. Both likelihood variables have been calibrated on large data control samples.

NN training and cut optimization

The artificial neural network used to make the final candidate selection has been constructed using the `NeuroBayes` package [95] in the context of the previous iteration of this analysis [97], and trained with simulated events as signal sample and mass sidebands data as background training sample. The training sample consisted of about 350k signal events and about 300k background events. The sidebands region is defined as in sec. 3.5.1.

Once the NN have been trained, the value for the cut on O_{NN} needs to be chosen. In the previous analysis the cut was instead optimized by studying the sensitivity to β_s in terms of its statistical error. This has been done by investigating the size of the statistical errors on β_s in different samples of pseudo experiments. A pseudo experiment is a set of a number of simulated events generated according to the model that is supposed to describe the data behavior [5]. Several pseudo experiments have been generated, each corresponding to a different NN cut value that determines the associated signal and background events composition. Three different values for β_s have been used (0.02, 0.3, 0.5), in order to obtain an optimized cut independently on the β_s value. For each of these, the decay width difference $\Delta\Gamma$ is generated according to the theoretical s relationship between these parameters, and all other variables are generated according to their best values as determined from the previous $B_s \to J/\psi\phi$ analysis [97]. About 700 pseudo experiments have been generated and the fit procedure to measure β_s is applied for each case; the β_s statistical uncertainty is then checked for each NN cut value (Fig. 3.4). From this study, it was determined that the optimal NN cut value is 0.2. [6]

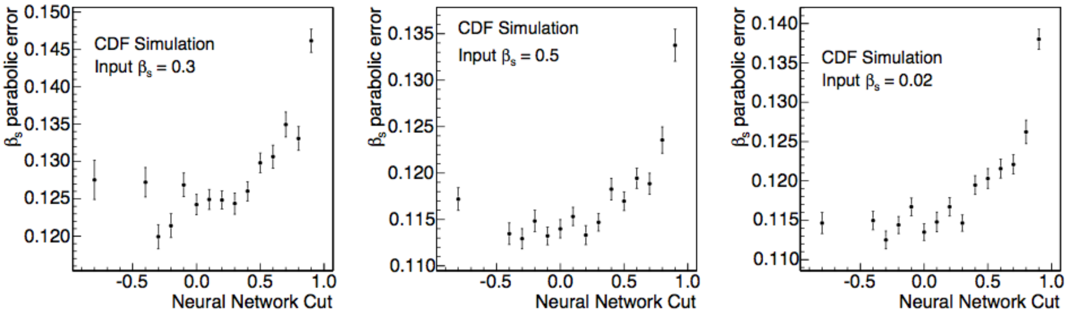


Figure 3.4.: NN cut optimization.

The first purpose of the current iteration of the "β_s $analysis$" is to obtain a measurement for β_s on a larger data sample ($\sim 8.4\,\text{fb}^{-1}$), by using the same procedure that has been performed in the previous analysis ($\sim 5.2\,\text{fb}^{-1}$ of data). For this reason the first step of the analysis is to make a comparison between the kinematic variables of the first $\sim 5.2\,\text{fb}^{-1}$ of data and the remaining ones. The CDF at Run II (CDFII) dataset is divided in periods according to the acquisition time. The first $\sim 5.2\,\text{fb}^{-1}$ of data corresponds to the data collected from the period 0 to period 25 included, while our dataset extends from

---

[5] see Chapter 4: the joint PDF that is used for the likelihood fit is the model that is supposed to describe the data. To generate pseudo experiments of simulated events with particular characteristics, a choice can be made on the parameters entering in the joint PDF. For instance if one wants to generate a data sample that would be well described by a huge size CP violation, the proper choice is to set β_s to a huge value in the joint PDF and then generate the simulated data according to that.

[6] this cut minimizes β_s errors by maximizing signal yield despite being a looser cut than would be selected by the traditional $S/\sqrt{S+B}$ with S.

period 0 to period 34, the whole Run II dataset include periods from 0 to 38. The comparison between the kinematic variables of the first $\sim 5.2\,fb^{-1}$ of data and the remaining ones, allows us to understand if and how much how much the data have changed over the time: e. g. the relative abundances of events given by a trigger path instead of another one can have changed during the time. This can happen for instance because of modifications to the trigger table (collection of all the trigger paths). Since each trigger path implements a set of requirements, often involving kinematic variables, the distribution of those variables can change depending on the used data sample. Furthermore the data comparison for different running periods can spot problems in the data sample sidebands subtracted distributions [7] have been compared at this stage and the data were required to satisfy the following set of rectangular cuts:

- $P_t(\mu_1) > 1.5\,\mathrm{GeV}/c$and $P_t(\mu_2) > 1.5\,\mathrm{GeV}/c$
- $P_t(k_1) > 0.6\,\mathrm{GeV}/c$ and $P_t(k_2) > 0.6\,\mathrm{GeV}/c$
- B vertex Prob > 0.001
- $P_t(B) > 5\,\mathrm{GeV}/c$

in substitution of the NN selection, since the NN performance is what we want to be probe. This comparison between the first $\sim 5.2\,fb^{-1}$ of data and the remaining ones has been made for all the NN input variables, in order to have a first information on whether these variables have a similar behavior in the two periods, justifying the NN use without the need to be trained again, maybe with a different MC for the signal region (or the same but weighted in such a way to recover the kinematic variables distributions of the data sample) and with the new sidebands. These plots can be found in Fig. 3.5 and Fig. 3.6. Another important check is the comparison between the NN outputs (see Fig. 3.7). The comparison in terms of the NN output has been done for the sideband subtracted sample, and for the sidebands separately. The statistical test used to quantify the goodness of the agreement between the two dataset is the Kolmogorov test, in both for the kinematic variables and the NN output. The distributions are in good agreement, this implying that the NN does not need to be trained again.

---

[7] see sec: 3.5.1

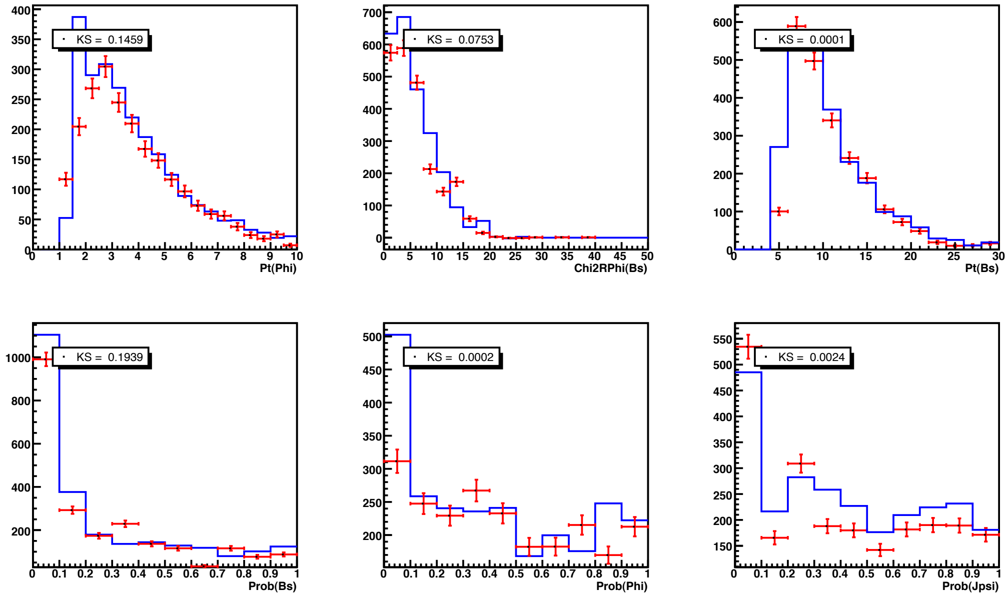


Figure 3.5.: NN input variables (kinematic varaibles): comparison between the first $5.2\,\mathrm{fb}^{-1}$ of data (BLUE line) and the full dataset (RED points). First raw from left to right $p_T(\phi)$, $\chi^2_{r\phi}$, $p_T(B_s)$, second raw $Prob(B_s), Prob(\phi), Prob(J/\psi)$.

3.6. The final data sample

Applying the pre-selection cuts and the cut on the NN output (O_{NN}) the invariant mass distribution $m_{\mu^+\mu^-K^+K^-}$ for the $B_s \to J/\psi\phi$ is obtained (see Fig. 3.9). The final data sample obtained has mainly three components:

Signal: the final yield obtained for $B_s \to J/\psi\phi$ decay events is of 9592 ± 104 events. This value is obtained by a binned likelihood fit on the mass histogram. The function used to model the signal is a single Gaussian. The standard deviation of the Gaussian turns to be ~ 10 MeV, that gives an indication about the goodness of the mass resolution that we can exploit.

Combinatorial background: these events are random combinations of charged tracks that satisfy accidentally the selection requirements, as well as events with real J/ψ reconstructed together with two random charged tracks. They produce a continuous invariant mass distribution and a smooth slowing decreasing distribution in the signal region is expected. It is the main source of background in our analysis.

Physics background: in our data sample there can be some contamination from $B_d \to J/\psi K^* \to [\mu^+\mu^-][K^\pm\pi^\mp]$ decay events mis- reconstructed as $B_s \to J/\psi\phi$ decays (defined as B^0 cross-feed); it occurs when in the reconstruction the daughter tracks of the K^* are assumed to be two kaons and an incorrect invariant mass is

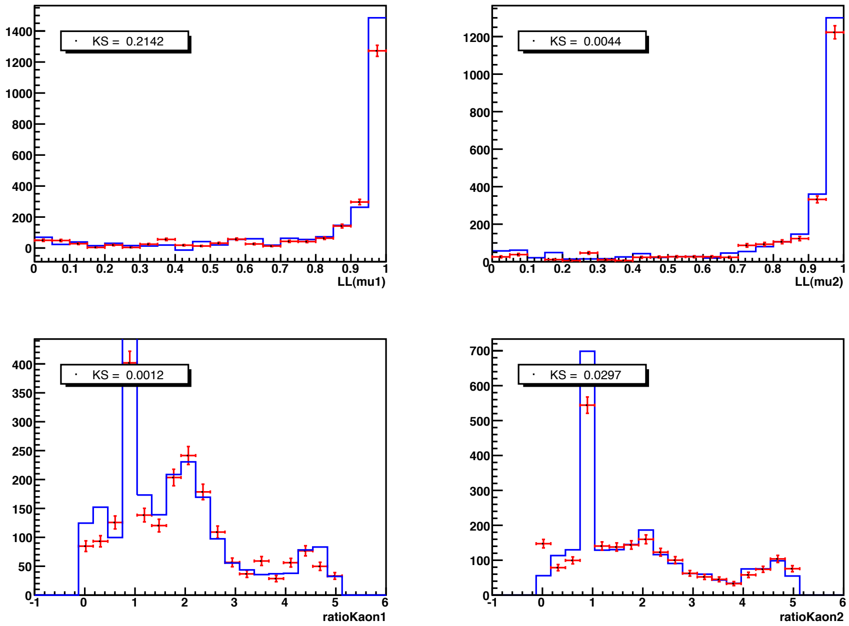


Figure 3.6.: NN input variables (identification varibles): comparison between the first $5.2\,fb^{-1}$ of data (BLUE line) and the full dataset (RED points). First raw from left to right $LL_{\mu}(\mu_1), LL_{\mu}(\mu_2)$, second raw $LL_K(K_1), LL_K(K_2)$.

computed. In this analysis there is a systematic error that accounts for this affect (see sec. 5.3). The fraction f of B^d cross-feed events in the B_s sample have been calculated to be (1.6 ± 0.6) into the signal sample of the previous analysis iteration. To estimate this fraction, production fractions of the B_s and B_d mesons need to be known as their relative decay rates to $J/\psi\phi$ and $J/\psi K^*$, respectively, and the efficiency of each type of event passing the final selection criteria established under the $B_s \to J/\psi\phi$ hypothesis. Both the production fractions and the branching fractions are taken from Ref. [27]. The efficiencies can be estimated using simulation, with both $B_s \to J/\psi\phi$ and $B_d \to J/\psi K^*$ modes reconstructed as $B_s \to J/\psi\phi$ decay. Eventually the fraction of B_d cross-feed is obtained as:

$$f(B_d\ in\ B_s\ sample) = \frac{f(\bar{b} \to B_d)\mathcal{B}(B_d \to J/\psi K^*)\epsilon(B_d)}{f(\bar{b} \to B_s)\mathcal{B}(B_s \to J/\psi\phi)\epsilon(B_s)}. \tag{3.6.1}$$

Another additional contributions from S-wave K^+K^- under the ϕ peak in $B_s \to J/\psi\phi$ decay can contribute up to few percents of the total rate. a normalized probability density for the decay $B_s \to J/\psi K^+K^-$ (kaons in an S-wave state) has been added to the likelihood function. These kaons can either be a non resonant pair of kaons, or the decay products of a scaler particle, the $f_0(980)$. In that case the final state of $B_s \to J/\psi f_0(980)$ can be only in S- wave, since the decay is a $P \to VS$

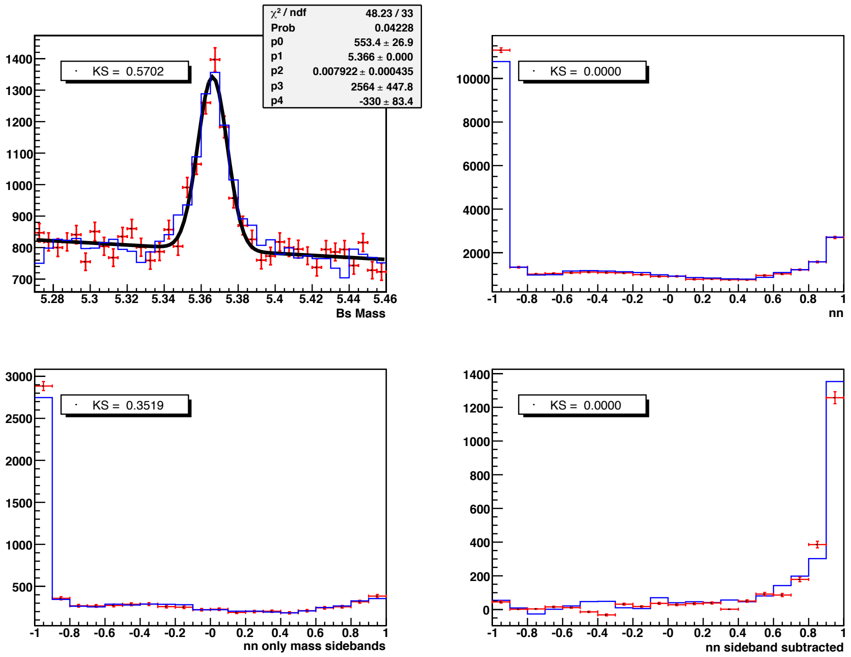
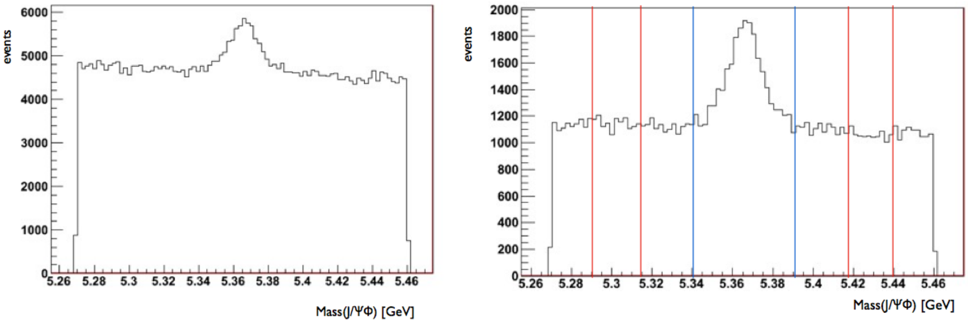


Figure 3.7.: NN output variable: comparison between the first $5.2\,fb^{-1}$ of data (BLUE line) and the full dataset (RED points).Top left the B_s mass distribution with a fit of mass overlayed. Top right the NN output variable (O_{NN}) for both signal (event with $O_{NN} \sim 1$) and background events ($O_{NN} \sim 1$). Bottom left: O_{NN} for the sidebands events. Bottom right: t O_{NN} signal events only (sideband subtracted).



(a) signal peak after the preselection cuts (b) signal peak with rectangular cuts applied

Figure 3.8.: Mass distribution of the events satisfying: 3.8a only preselection requrements, 3.8b preselection + rectangular cuts .

decay. We account for these two contributions by adding a normalized probability density to the likelihood in the full fit determining β_s, $\Delta\Gamma$ and all the parameters that we want to measure (see sec. 4.7).

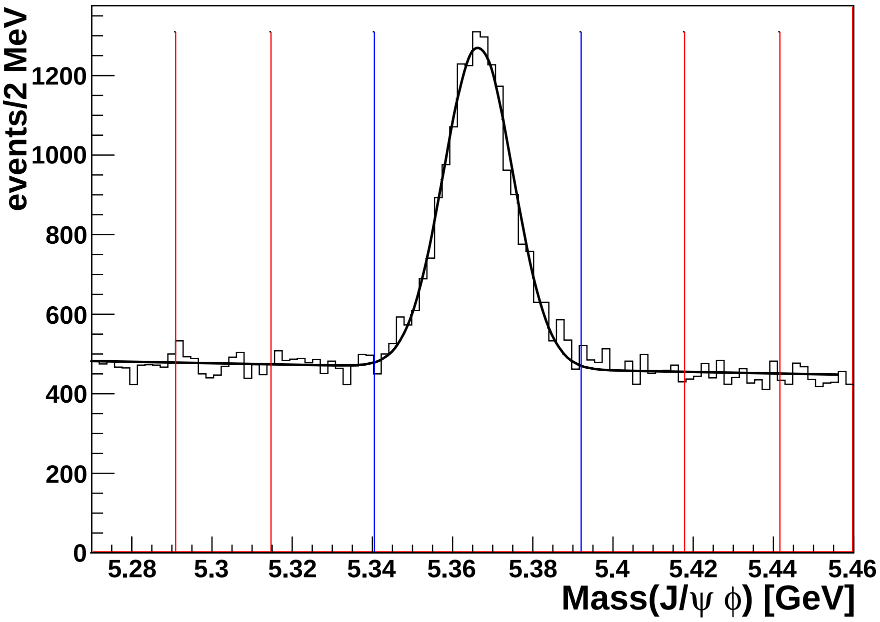


Figure 3.9.: Mass distribution of our final data sample: the events satisfying preselection requirements + NN. BLUE lines show the signal region, and RED lines the sidebands.

3.7. Flavor Tagging

3.7.1. Flavor tagging principles

Knowledge of the B_s meson flavor at the production is an important ingredient of this analysis. The distinction between B_s and $\bar{B}_s$ at the production is called *flavor tagging*. Properties of the $p\bar{p} \to b\bar{b}$ production process, and b quark hadronization and fragmentation are used in two flavour tagging algorithms, the Same Side Kaon and the Opposite Side Tagger (SSKT,OST) . Fig. 3.10 shows the kinematics exploited by the two tagger types.

The SSKT uses the fragmentation tracks of the $candidate$ B mesons (meaning the B meson of interest) to determine its flavor. The tagger is supposed to identify the flavor of the $s$trange quark (i.e. s or $\bar{s}$) of the B_s meson candidate in a kaon produced alongside the B_s (or $\bar{B}_s$). In the case of a B_s meson ($\bar{b}s$) at the production, the s quark of the kaon, that is the produced partner of the s of the B_s , should be a $\bar{s}$. If we have s $\bar{B}_s$ at the production, the kaon should be formed by an s quark. In other words if a K^+ is identified, the meson at the production was a B_s , instead in case of a K^- identification the initial meson was a $\bar{B}_s$.

The OST exploits the pair production of the b quarks, and uses the information from the hadronization and decay of the other b-quark, i.e. on the opposite side with respect to the B meson of interest. In practice this means determining the charge either of leptons

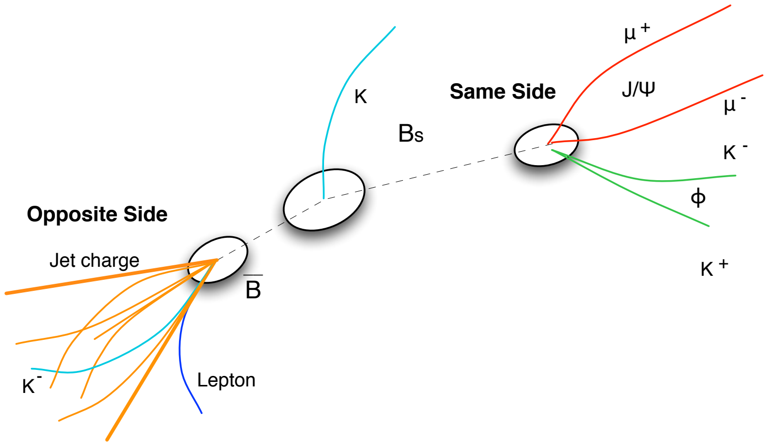


Figure 3.10.: Production and decay of a B_s meson.The main event features exploited by the SSKT and the OST are shown.

coming from semileptonic B decays or of b-jets on the opposite side, when dealing with hadronic decays. For example the identification of a negative (positive) charged lepton on the opposite side means that the meson of interest initially was a B_s ($\bar{B}_s$).

It is now useful to introduce some quantities commonly used to characterize flavor taggers:

- the *tag decision* ξ. It is a discrete variable, that can take the value -1, 0, or 1. $\xi = -1$ means that the meson at the production has been tagged as $\bar{B}_s$, $\xi = 1$ if the initial meson has been tagged as B_s, and $\xi = 0$ if the tagger could not make a decision.

- the *tagging efficiency* ϵ. It is the fraction of events for which a tag decision can be made. It is defined as follows:

$$\epsilon = \frac{N_{tagged}}{N_{untagged} + N_{tagged}} \tag{3.7.1}$$

  where N_{tagged} is the number of tagged events and $N_{untagged}$ is the number of events for which a tagging decision has not been taken.

- the $dilution$ $\mathcal{D}$, a quantity defined in order to characterize the rate of mis-tagging for a particular algorithm. It is defined as:

$$\mathcal{D} = 1 - 2P_W = \frac{N_R - N_W}{N_R + N_W} \tag{3.7.2}$$

  where N_R is the number of right-tagged events, N_W is the number of wrong-tagged events, and P_W is the probability for an event to be wrongly tagged. With this definition, when the tagger does not work properly, and it assigns randomly the tag decision , $P_W = 0.5$, so $\mathcal{D} = 0$. In the case of an algorithm that makes no mistakes in the tag decision assignment, the dilution is $\mathcal{D} = 1$. Moreover P_W is parametrized as a function of the calibration parameters, so that the dilution ca be predicted event by event. This gives us the advantages that, in a likelihood function, the events can be weighted according to the reliability of the tagger assignment, and that the tagger can be calibrated on a data subsample.

- the *effective tagging efficiency* defined as $\epsilon\mathcal{D}^2$, where $\mathcal{D}$ in this case represents the average dilution over the whole sample. This quantity is usually used to measure the performance of a flavor tagger.

The flavor taggers used in this analysis have been developed for the CDF B_s^0 mixing measurement, and have been re-calibrated and optimized for the previous iterations of the $B_s \to J/\psi\phi$ analysis. In the next Chapter will be shown how these quantities enter in the fitting likelihood function.

3.7.2. Opposite Side Tagger

As mentioned before, the OST can be performed in two ways:

- by *identifying the charge of the lepton from semileptonic B decays* (Soft Electron Tagger (SET) and Soft Muon Tagger (SMT)).
  The SET is described in detail in [99], and the SMT in [100]. The efficiency of these taggers is rather low, of order 20%, which is similar to the branching fraction of B to semileptonic decays. The tagging dilution is worsened by the mis-identification of leptons, and by the B_s mesons oscillating to $\bar{B}_s$ with the opposite flavor of their production and therefore giving an incorrect tag.

- by *identifying the charge of the opposite side b jet* (Jet charge Tagger (JQT)).
  The jet charge tagger [101] infers the flavour of the candidate B_s from the charge of the opposite side b jet. The jet charge is calculated as the momentum weighted sum on all the jet particles. Tracks are requested to be isolated from the candidate B meson, as it is important to look at jets only from the opposite side. Neural network algorithms are then used to find the jet most likely to come from a b quark. The dilution for this tagger is parametrized as a linear function of the jet charge and the probability that the jet contains a b (or $\bar{b}$ quark).

These three opposite side taggers are not independent, since they can share tracks, and this effect needs to be accounted for when using the three taggers together. They are combined to give a single opposite side tagging decision by means of a neural network procedure [105].

All these tagging processes are independent from the candidate side hadronization products, so it is possible to use the same opposite side tagging algorithm for sample of different B meson types. This means that the algorithms can be developed or calibrated with the high statistics light B meson samples, and then applied to tag B_s mesons. Events

from the l + SVT trigger sample are used for tagger development and parametrization of the predicted dilution in high statistics samples. The taggers are then calibrated on the di-muon samples, finding a global *scale factor* $S_{\mathcal{D}}$, which will be applied to the event by event dilution to account for the kinematic differences in the two samples.
In this analysis, in order to calibrate the OST tagger, the $B^+ \to J/\psi K^+$ mode has been used. It has kinematic quantities with similar behavior of the $B_s \to J/\psi\phi$ decay sample, but an higher statistics and the advantage that the flavor of the candidate meson is known, as it is tagged by the charge of the daughter kaon. This provides us with the true value to be compared with the tag decision, allowing the dilution to be measured. From this sample, a scale factor, S_D to apply to the predicted dilution is calculated. If the predicted dilution is suitable for this sample, the scale factor should be equal to 1.0 within errors. Two scale factors are estimates, for B^+ and B^- separately, to account for any charge related asymmetry in the tagging algorithms. In Fig. 3.12 the measured dilution is plotted against the predicted dilution for the b and $\bar{b}$ events. To validate the use of the OST developed on l+SVT data for the di-muon samples, the dependence should be consistent with a straight line of slope of order unity. The fitted slopes of these dependences are taken as the scale factors for the dilution. The fitted distributions of the measured dilution for the $B^+ \to J/\psi K^+$ sample as function of the predicted dilution are reported in Fig. 3.12.

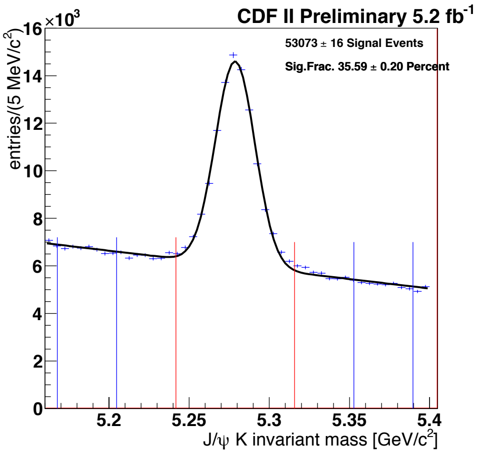


Figure 3.11.: Invariant mass for $B^+ \to J/\psi K^+$ sample, used to calibrate the OST.

OST calibration for this analysis uses $B^+ \to J/\psi K^+$ collected from February 2002 to March 2011 corresponding to an integrated luminosity of $\mathcal{L} = 8.5\,fb^{-1}$. The event selection, as in the case of $B_s \to J/\psi\phi$ decay sample is performed in two stages; after event reconstruction, a the following loose set of cuts is applied:

- $5.16 < M_{B^+} < 5.40\,\mathrm{GeV}/c^2$
- at least 3 axial hits per track in the silicon detector for muons

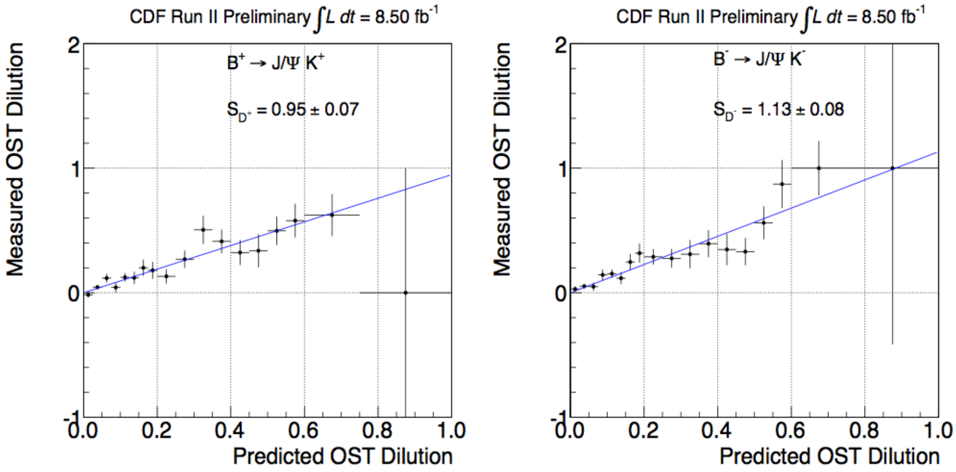


Figure 3.12.: Measured versus predicted dilution for B^+ (left) and B^- (right) for the entire data set used in this analysis.

- at least 3 axial hits per track in the silicon detector for the K
- XFT muons match
- $0.0 \leq c\tau \leq 0.1$

then the preselected data is fed into a Neural Network [8] to perform a more effective background suppression.

The measured scale factors in $\sim 8.5\,fb^{-1}$ of data are $S_{\mathcal{D}^+} = 0.95\pm0.07$ and $S_{\mathcal{D}^-=1.13}\pm 0.08$; the tagging efficiency is $93.4\pm0.1\%$ and the mean predicted dilution is $6.86\pm0.03\%$. Finally an effective tagging power $\epsilon\mathcal{D}^2 = 1.20 \pm 0.01\%$ is measured.

3.7.3. Same Side Kaon Tagger

The SSKT uses the fragmentation tracks which accompany the *candidate* B mesons to determine its flavor. Usually the charge of the kaon produced alongside the B_s is used (SSKT). The behavior of SSKT in different B species is expected to vary, eliminating the possibility of developing and calibrating the SSKT in the same high statistics light B samples as used for the OST. The SSKT used in this analysis is the tagger which was developed for the CDF B_s mixing measurement [102] Previously, the SSKT used in the CDF B-physics analyses has been calibrated on MC simulated data, Currently, the predicted dilution is still calculated in this way, however the dilution scale factor is computed by a data driven calibration using the measurement of the B_s mixing amplitude. The SSKT calibration is fully documented in [104].

---

[8] the NN output cut is chosen to 0.8 in order to maximize the ratio $S/\sqrt{S+B}$

The SSKT used in this analysis is the tagger which was developed for the CDF B_s mixing measurement [102], and has been calibrated using the most recent measurement of B_s mixing [103], made with $5.2\,\mathrm{f}b^{-1}$ of data where a single scale factor $S_{\mathcal{D}} = 0.94 \pm 0.15(\text{stat}) \pm 0.13(\text{syst})$is estimated . A tagging efficiency of 52.2 ± 0.7 and an average predicted dilution on signal of $21.8 \pm 0.3\%$ have been determined through the previous analysis. During last data taking period, the triggers which collect the calibration dataset have been highly suppressed. For this reason the SSKT calibration obtained with $5.2\,\mathrm{f}b^{-1}$ of data is used in our analysis, and the SSKT is applied only for the first $5.2\,\mathrm{f}b^{-1}$ of data.

3.8. Simulated events

In this analysis, simulation of B production and decay processes and of the subsequent detector response is used to determine the detector sculpting of the angles due to the non-hermeticity of the CDFII detector. The numerical simulation is the only method that can be used to achieve this purpose. The algorithms used all involve some type of random sampling to simulate processes, and are collectively called MC simulation. The simulation is divided into several steps, which reproduce in order the main physical processes and processing steps involved in collecting data from real $p\bar{p}$ interactions.

The first step in simulation is the treatment of the $p\bar{p}$ hard scattering, and the out-going quark and gluon collision products, followed by simulation of the fragmentation and hadronization processes which yield hadrons and associated jets. We used the `BGenerator` package [106]: it concentrates on producing only one B meson per event, which yields a great advantage in computational speed. On the other hand, by design it does not mimic the full collision environment, as the `PYTHIA` package could perform. For our purposes, `BGenerator` is sufficient since we wish to model single B decay samples. The second step is the simulation of the full decay chain of the B mesons under study. For this task, we use the `EVTGEN` package [107]. `EVTGEN` is specialized for heavy flavor decays and accounts correctly for quantum mechanical interference effects. In order to model the detector angular acceptance for $B_s \to J/\psi\phi$, we use the *phase-space* decay model of `EVTGEN` . Phase-space means that all spins of the particles in the final state are averaged. This yields flat distributions in the angular variables whose acceptance we wish to study. The third step in simulation incorporates the interaction of the decay products ($K^+K^-$and$\mu^+\mu^-$) with the detector material. For this task we use the `CDFSIM` package [108], which is a CDFII-specific full detector simulation based on the `GEANT` simulator [82]. The final step is the simulation of the triggering and event reconstruction that data events pass through. `CDFSIM` outputs simulated events with the same data banks as the raw real data events. The detector and trigger configurations have undergone several variations during Run II. The simulation can access the databases and thus allows us to simulate the detailed configuration of any set of real data-taking runs for modeling the realistic detector response in any given subset of data. On the other hand, the MC is not able to reproduce exactly the trigger behavior of the trigger-paths which have a built-in dependence on the instantaneous luminosity, because this information is not fully stored in the database. For this reason it is necessary to reweight MC events in a way suitable to reproduce the different trigger mixture observed in the dataset that we are using. The output of the MC simulation is then processed with the software package which reconstructs B decays and writes the output events in the `BStntuple` format, allowing us an easier comparison between MC and data.

In order to achieve our purpose, it is necessary that the simulated events have the same characteristics of the data sample, for this reason they are fed to the same pre-selection cuts and NN selection that has been used for the data (see Sec. 3.5.1, 3.5.2).

The Monte Carlo sample that is used in the current iteration of the "β_s analysis" has previously been used for the previous iterations of the same analysis, and for the untagged measurement of $\Delta\Gamma$ and β_s [110] and it corresponds to data conditions of the first $\sim$ 1 fb^{-1} of data. The quality of our determination of the transversity angle sculpting depends on the agreement between our data and the generated MC in variables that affect the angular decay features of the J/ψ and the ϕ. Previously it has be seen ([110], [111], [97]) a disagreement in the $p_T(B_s)$ spectrum between this MC sample and the data, and the same effect has been observed when considering our data sample (see Fig. 3.15a). Since the p_T spectra can affect the distributions of the transversity angles, the agreement between data and MC has been investigated in several variables. Since part of the original reweighting in Ref [111], [97] was depending on trigger prescales, the reweighting needs to be done again in order to match with the current dataset.

In order to weight the MC events according to the data sample, it is necessary to use a only-signal or sideband-subtracted data sample. The side-band subtraction is needed in this comparison, because the MC data reproduce only the signal events.

The MC reweighting procedure used involves three steps: the first takes into account the different trigger path mixture that characterizes our dataset; the second accounts for the agreement in the $p_T(B_s)$ spectrum, and the purpose of the third step is to account for the combined effect of both the different trigger paths mixture and and the $p_T(B_s)$ spectrum.

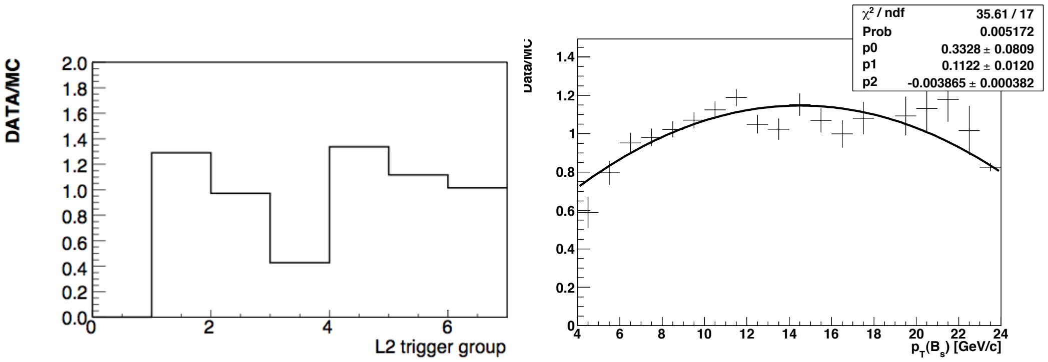
Trigger path mixture: the candidates are first of all split into two groups, depending whether the candidate triggers with a CMU-CMU or a CMU-CMX muon pair. After this, each of the two classes is split in three classes defined as:

- Both muons have $p_T > 3\ \text{GeV/c}$
- Both muons have $p_T > 2\ \text{GeV/c}$ and at least one muon has $p_T > 3\ \text{GeV/c}$
- all events left, not falling in the previous two classes

In this way 6 different classes have been obtained and their fraction in the simulated events has to be adjust in order to mach with the current data sample. This classes can be considered at first approximation mutual exclusive and such that their union gives the whole data sample. Fig 3.13a shows the weights found for the six classes. A weight is obtained by making the ratio between the number of real events belonging to one of the six classes defined above over the number of the simulated events belonging to the same class. The high of the first three bins of the histograms represents the weights for the three CMU-CMU trigger classes and the other bins involve the CMU-CMX triggers.

The $p_T(B_s)$ distribution is compared between data and MC events after the first step of reweighting procedure ("trigger classes" [9]). The number of the simulated events has previously been normalized to the number of the signal events in the data and the $p_T(B_s)$ region considered extends from 4 GeV/c to 24 GeV/c. This ratio has been fitted with a second order polynomial (see Fig 3.13b). That function will then be used to reweight the MC events.

Combined effect of trigger path admixture and $p_T(B_s)$ distribution: the weight factor associated to each simulated event is the product of the weight factor associated to the class at which the events belongs and a factor computed using the second order polynomial used to fit the ratio of the two $p_T(B_s)$ distributions (see the previous stage of the MC reweighting procedure). Fig. 3.14 and Fig. 3.15 show the good agreement between the data sample and the so reweight MC events both for the ”trigger classes" composition and for the $p_T(B_s)$ distributions.



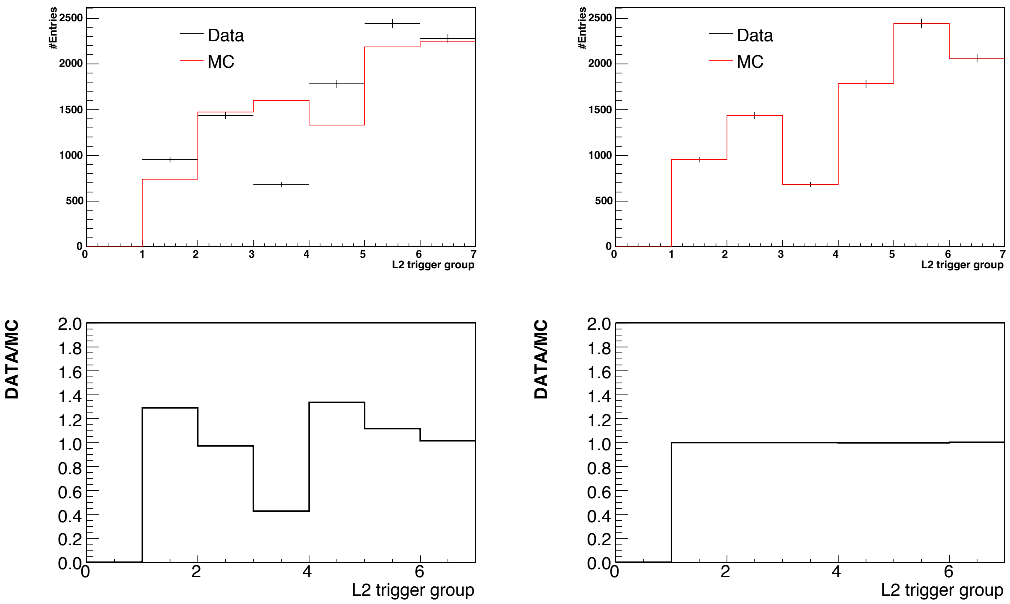
(a) Trigger group reweighting (stage 1)

(b) $p_T(B_s)$ reweighting (stage 2)

Figure 3.13.: MC weights. In 3.13a the weight according to trigger group is discrete (one different weight is associate to each trigger group defined in sec. 3.8. In 3.13b the weight according to the $p_T(B_s)$ distribution is according to the continuous function that fits the distribution in the plot, as a result for each different value of $p_T(B_s)$ there is a different weight.

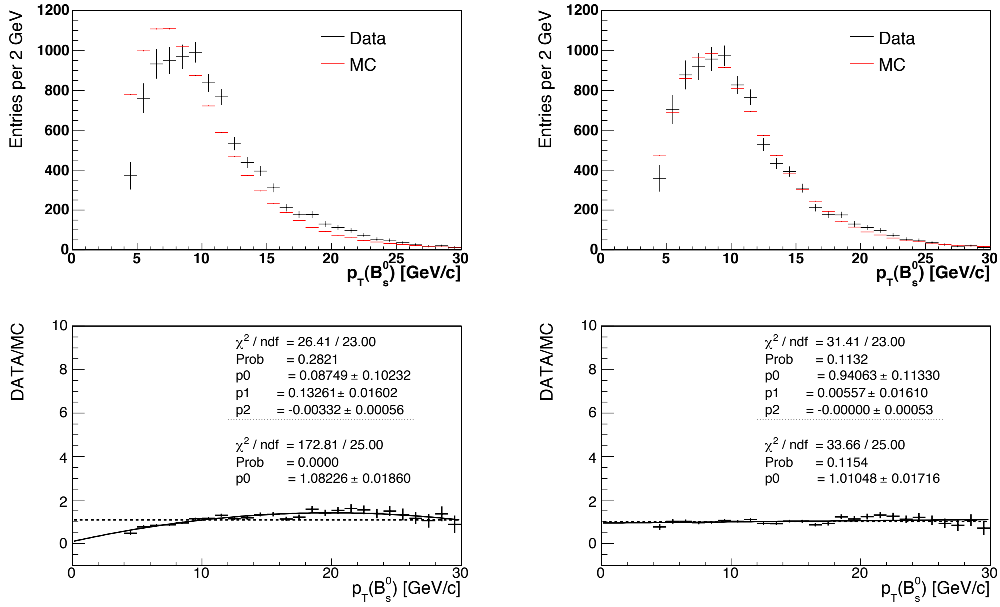
---

[9] "trigger classes" is written with quotation marks since we are not considering all the trigger paths individually, as in some analysis is done, and it would be more accurate. The sample has been divided in these six classes because we are using a large number of different trigger paths, around 50, and these six classes can have at first approximation the same properties that would have classes defined on individual trigger path basis: they are mutual exclusive and their sum gives the whole data sample.



(a) MC not reweighted (b) MC after the 3 reweighting stages

Figure 3.14.: Comparison between data and MC events samples: trigges classes.



(a) MC not reweighted (b) MC after the 3 reweighting stages

Figure 3.15.: Comparison between data and MC events samples: $p_T(B_s)$ distribution.

Chapter 4

Maximum Likelihood fit

Contents

In this chapter a description of the method used in this analysis to measure the physics parameters of interest will be given: an unbinned multi-dimensinal maximum likelihood fit. Each component of the likelihood function will be discussed, together with the inputs that are needed to the fit, in order to converge and describe properly the physics process involved. The last part of the chapter is in particular useful to understand which parameters enter in the likelihood function used, and how their values are related with the quantities that are intended to be measured.

4.1. Fitting method: unbinned maximum likelihood

A maximum likelihood fit has been used to extract informations on the parameters of interest in this analysis. Let us call $\vec{\theta}$ the vector of the unknown parameters that the fit is supposed to determine. Then a sample of N events is considered. For an event i, let $\vec{x_i}$ be the vector that collects all the variables characterizing that event (kinematic, identification and so on). The probability density function (Probability Density Function (PDF)) , denoted with $p(\vec{x_i}|\vec{\theta})$ describes the expected distribution of the events on the space of the event variables $\vec{x}$, given the vector of parameters $\vec{\theta}$. The notation $p(\vec{x_i}|\vec{\theta})$ can be understood as the probability to observed an event characterized by the observed quantities $\vec{x_i}$, in the hypothesis that such an event is belonging to the class described by the vector of parameters with values $\vec{\theta}$. This method is called "unbinned" since the likelihood function is baseg on event-by-event values of the parameters (different from the "binned" likelihood that exploits values of the parameters determined for each bin in which the events are collected.) The likelihood function, for a given dataset is constructed as follows:

$$\mathcal{L}(\vec{\theta}) = \prod_{i=1}^{N} p(\vec{x_i}|\vec{\theta}) \tag{4.1.1}$$

In practice what is usually done is the minimization of:

$$\log \mathcal{L}(\vec{\theta}) = \sum_{i=1}^{N} \log p(\vec{x_i}|\vec{\theta}) \tag{4.1.2}$$

In order to avoid numerical precision problems that can arise when multiplying many small numbers together. For the numerical minimization the `MINUIT` package [112] and the `ROOT` analysis framework [113] are used.

4.2. Components of the likelihood function

In order to construct the likelihood, one first needs to define the probability density function. Considering an unidimensional space of the event variables given by the variable x, the PDF $p(x|\theta)$ is defined in such a way that the probability of observe x in the interval [a, b] is given by:

$$P(x \in [a,b]) = \int_a^b p(x|\theta)dx \tag{4.2.1}$$

Now, in the case of a multidimensional variable space, the joint PDF can be built as a function of all event variables. The probability of observing variables ($\vec{x} = (x_1, x_2, ...x_n)$)within the n-dimensional volume V, with the vector of parameters characterizing the

PDF $\vec{\theta} = (\theta_1, \theta_2, ...\theta_n)$ is:

$$P(\vec{x} \in V) = \int_V p(\vec{x}|\vec{\theta})d\vec{x} = \int_V p((x_1, x_2, ...x_n)|(\theta_1, \theta_2, ...\theta_n))dx_1 dx_2...dx_n \quad (4.2.2)$$

If the PDF is separable, we can write it as the product of the individual PDFs in the variables subspaces:

$$p((x_1, x_2, ...x_n)|(\theta_1, \theta_2, ...\theta_n)) = p(x_1, \theta_1) \cdot p(x_1, \theta_1) \cdot ...p(x_n, \theta_n) \quad (4.2.3)$$

In the following sections the joint PDF used in this analysis will be presented, starting from the identification of its components. In order to make the explanation clearer, each independent variable subspace can be considered separately:

- **B_s mass PDFs for signal $P_s(m)$ and for the background $P_b(m)$.** The mass distribution is separable from lifetime and angular components of the joint PDF in the case of B_s
- **Signal decay time and angular PDF, $T(\theta, \phi, \psi, t)$.** In the time-dependent analysis of $B_s \to J/\psi\phi$ the proper decay length part of the joint PDF is not separable from the angular part. Its principal component consists of the differential decay rate for $B_s \to J/\psi\phi$, as will be discussed in the following sections.
- **background decay time PDF $P_b(t)$.**
- **background angular model PDFs $P(\theta), P(\psi), P(\phi)$.**
- **Lifetime error PDFs for signal $P_s(\sigma_{ct})$ and for the background $P_b(\sigma_{ct})$.**
- **PDFs for the flavor tagging variables $P_s(\xi), P_b(\xi)), P_s(\mathcal{D})$ and $P_b(\mathcal{D})$.** The signal time and angular dependence of the likelihood needs to be written in a different form when including the flavor tagging variables (i.e. the PDF for the temporal and angular dependence of the signal is different from the untagged to the tagged fit, see sec. 4.6.1).

4.3. B_s mass PDF

Since the peak of the B_s is well defined, the mass distribution of events is primary useful in order to separate signal candidates from background, while it cannot be used for separating CP-odd from CP-even decays of the B_s meson.

The signal mass distribution is modeled by a single gaussian function with central value M, smeared with an event-by-event mass resolution (σ) scaled using a scale factor (s_m) to account for a general mis-estimation on the mass errors (denoted as σ_m). The PDF is then constructed by normalizing the gaussian over the mass window determined by the preselection requirements; it forms is the following:

$$P_s(m|M,\sigma_m) = \frac{\frac{1}{\sqrt{2\pi}s_m\sigma}exp\Big[-\frac{1}{2}\Big(\frac{m-M}{s_m\sigma}\Big)\Big]}{\frac{1}{2}\Big[Erf\Big(\frac{M_{max}-M}{\sqrt{2}s_m\sigma}\Big) - Erf\Big(\frac{M_{min}-M}{\sqrt{2}s_m\sigma}\Big)\Big]} \tag{4.3.1}$$

where $Erf(z)$ is the error function and M_{min}, M_{max} are the limits of the mass window as defined by the preselection ($M_{min} = 5.1$ GeV/c^2 and $M_{max} = 5.6$ GeV/c^2, see sec. 3.5.1). Notice that the expressions for the single PDFs on the independent variable subspaces that are reported here, are meant event-by event quantities. For example, in eq. 4.3.1 $P_s(m|M,\sigma_m)$ should be written as $P_s(m_i|M,\sigma_m)$ meaning that the m is the mass of the candidate i .Here and in the following sections the index i will be dropped.

The background mass model is a first order polynomial function, normalized in the same interval used for the signal PDF gives the background mass PDF:

$$P_s(m|p_1) = p_1 \cdot m + \frac{1}{M_{max} - M_{min}}\Big[1 - \frac{p_1}{2}(M^2_{max} - M^2_{min})\Big] \tag{4.3.2}$$

where p_1 is the slope, the coefficient of the first order polynomial.

4.4. PDF in angular and time variables

The differential decay rate as function of time and angles derived in Chapter 1 (see sec. 1.3) describe the phenomenology of $B_s \to J\psi\phi$ decays without keeping into account effects that are not part of the physics concerning the decay considered, but are involved in the measurement as well. These kind of effects include the non-hermeticiity of a real detector, and the event selection used. The effect of the CDF detector need to be kept into account in the maximum likelihood fit to determine the parameters of interest. The detector affects mainly two aspects that are relevant for this measurement. The first are the angular distributions. Distributions of the events in the transversity angles ($\cos\theta, \phi, \cos\psi$) undergo some sculpting due to the detector angular acceptance and the event selection used. It is in principle important to account for this effect to obtain correctly the angular distributions, since they are used to distinguish CP-even from CP-odd final states, and therefore the CP violating asymmetry [1]. The second detector effect that needs to be account for is the time resolution. The latter affects both the PDF for the signal and for the background candidates in different manners.

---

[1] it is true that In practice we are performing a time dependent maximum likelihood fit instead of the computation of the asymmetry, but it exploits the same physical principle of identification of CP-even and odd states.

4.4.1. Detector Sculpting of Signal Angular Distributions

Let us recall the time and angular dependent description of the $B_s \to J\psi\phi$ signal component (see the decay rate eq. 1.3.3):

$$\begin{aligned} P(\theta,\phi,\psi,t) = P(\vec{\omega},t) \propto \frac{d^3\Lambda(\vec{\omega},t)}{d\vec{\omega}} \propto \sum_k O^k(t) g^k(\vec{\omega}) \\ \bar{P}(\theta,\phi,\psi,t) = \bar{P}(\vec{\omega},t) \propto \frac{d^3\bar{\Lambda}(\vec{\omega},t)}{d\vec{\omega}} \propto \sum_k \bar{O}^k(t) g^k(\vec{\omega}) \end{aligned} \tag{4.4.1}$$

where the first expression is referred to the B_s decay and the second to $\bar{B}_s$. The sum of $P(\theta,\phi,\psi,t)$ and $\bar{P}(\theta,\phi,\psi,t)$ needs to be normalized on the angles an time, in order to assume a probability meaning, and constitute a PDF:

$$\iiiint \Big[P(\theta,\phi,\psi,t) + \bar{P}(\theta,\phi,\psi,t)\Big] d(\cos\theta) d(\cos\psi) d\phi dt = 1 \tag{4.4.2}$$

The detector sculpting at this point can easily be inserted as an efficiency time-independent factor that multiplies the time and angular decay PDFs $P(\theta,\phi,\psi,t)$ and $\bar{P}(\theta,\phi,\psi,t)$. In this way new PDFs $P'(\theta,\phi,\psi,t)$ and $\bar{P}'(\theta,\phi,\psi,t)$ are obtained:

$$\begin{aligned} P'(\theta,\phi,\psi,t) = \frac{1}{N}\epsilon(\theta,\phi,\psi) P(\theta,\phi,\psi,t) \\ \bar{P}'(\theta,\phi,\psi,t) = \frac{1}{N}\epsilon(\theta,\phi,\psi) \bar{P}(\theta,\phi,\psi,t) \end{aligned} \tag{4.4.3}$$

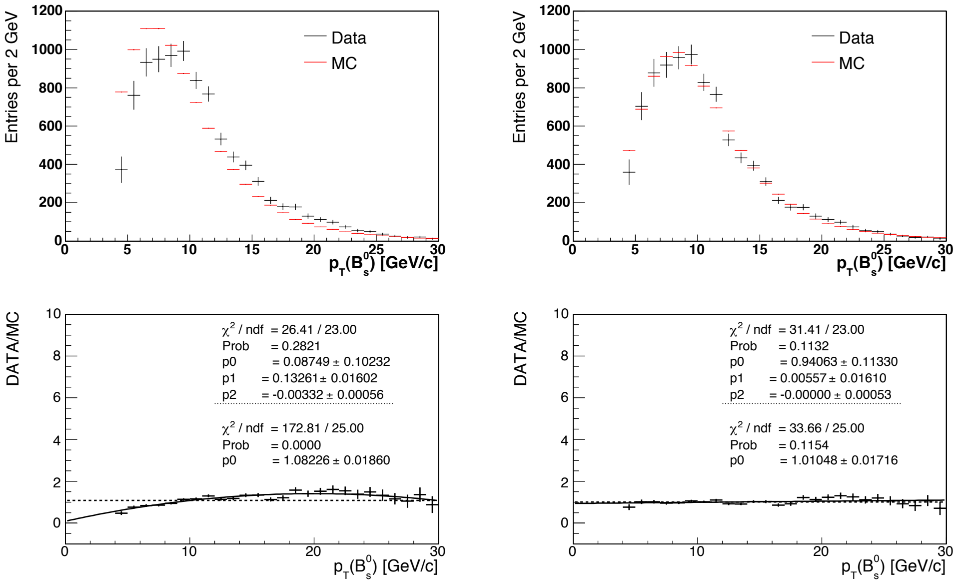
where the normalization factor N is given by:

$$N = \iiiint \Big[P(\theta,\phi,\psi,t) + \bar{P}(\theta,\phi,\psi,t)\Big] \epsilon(\theta,\phi,\psi) d(\cos\theta) d(\cos\psi) d\phi dt \tag{4.4.4}$$

The detector efficiency function is parametrized using an expansion of real spherical harmonics for the (θ,ϕ) angles, where spherical harmonic each term is expanded as a function of a Legendre polynomial used to fit ψ.

$$\epsilon(\theta,\phi,\psi) = \sum a^k_{lm} P_k(\theta,\phi,\psi) Y_{lm}(\theta,\phi) \tag{4.4.5}$$

The normalization factor N can be derive analytically, as shown in Appendix A , and the result is an expression involving the physics parameters of interest such as $|A_0|, |A_\perp|, |A_\parallel|, \beta_s, \Delta\Gamma$, and also the coefficients a^k_{lm} describing the detector efficiency. Those coefficient have been determined by a fit to a realistic Monte Carlo simulated data. Three samples of 100 milion $B_s \to J\psi\phi$ events each have been used. Those events correspond to data condition of the first 1.3 $\mathrm{fb}{-1}$ mostly in terms of trigger paths. The quality of the determination of the transversity angles sculpting depends on the agreement between data and MC in the distributions of the variables that affect the angular decay of the particles. In particular the first distribution to check is the $p_T(B_s)$ spectrum, since it affects the distributions of the transversity angles. Each trigger path includes a number of requirements on tracks and also on the p_T of particles, so if the Trigger Table [2] has been changed, also a difference in the p_T of the particles is expected. Moreover the trigger prescales are another factor that could affect the relative abundance of events coming from a trigger path with respect to another, and the shape of the p_T spectrum, as a consequence. For this reason, in order to have a good agreement between the dataset used and the MC events, a re-weighting of the MC has been done. Variables used to perform this re-weighting are the transverse momentum of the B_s and and the "class" of trigger paths to which the each event belongs. This full procedure is explained in Chapter 3.



(a) MC not re-weighted

(b) MC re-weighted

Figure 4.1.: Comparison of transverse momentum between data and realistic Monte Carlo used to determine angular efficiencies. Fig. 4.1a show the disagreement of the $p_T(B_s)$ spectrum between data and MC before the re-weighting procedure. Fig. 4.1b instead shows the agreement reached after the re-weighting.

The variables $\cos\theta, \phi$ and $\cos\psi$ have been generated flat in the MC events, and then they have been reconstructed using the same procedure used for the real data, thus the distribution found represent the detector efficiency in each angular variable. With the MC distributions of $\cos\theta, \phi$ and $\cos\psi$, a 3D histogram has been filled, and its analytical description used has been shown in eq. 4.4.5. To obtain the coefficients a^k_{lm}, the spherical

---

[2] the "list" of all the trigger paths used

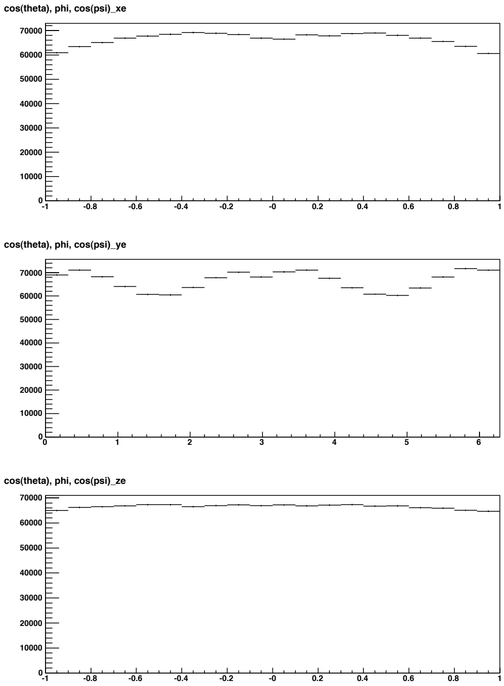


Figure 4.2.: Transversity angles distributions for Monte Carlo events. Each variable has been generated flat, therefore the shape shown by the plots are due only to the detector sculpting effect usinge the same reconstruction as for data.

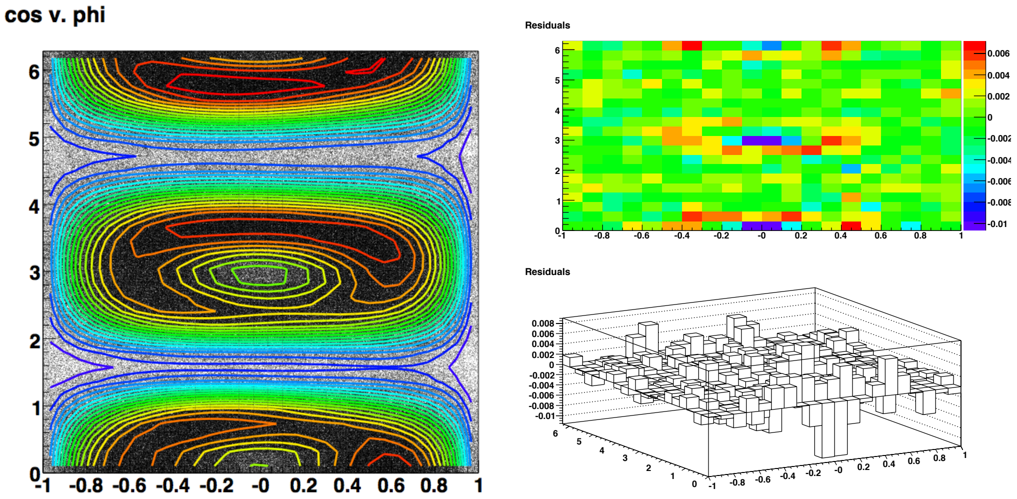
harmonics in (θ, ϕ) are expressed according to the Laplace series:

$$Y(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=0}^{l}[C_{lm}\cos(m\phi) + S_{lm}\sin(m\phi)]P_l^m(\cos\theta) \tag{4.4.6}$$

where the coefficients C_{lm} and S_{lm} are expanded in function of Legendre polynomial used to fit ψ:

$$\begin{aligned} C_{lm} &= \sum_{k=0}^{\infty} C_{lm}^{k}\sqrt{\frac{(2k+1)}{2}}P_k(cos\psi) \\ S_{lm} &= \sum_{k=0}^{\infty} S_{lm}^{k}\sqrt{\frac{(2k+1)}{2}}P_k(cos\psi) \end{aligned} \tag{4.4.7}$$

The series in eq. 4.4.5 can be related to the set of orthonormal basis function used in Appendix A for the calculation of the normalization N. In Fig. 4.3 is represented the two-dimensional fit to the (θ, ϕ) distribution, integrated over $\cos\psi$. with relative residuals.



(a) Fit to the (θ, ϕ) distribution (b) Residuals

Figure 4.3.: Two dimensional fit to $(\cos\theta, \phi)$ transversity angles integrated over $\cos\psi$.

4.4.2. Lifetime error PDFs

The unbinned Maximum Likelihood procedure that is used in this analysis, as explained in previous sections, makes use of the PDFs , that represents the probability distributions of the observables used in the fit, that behave like "templates". These templates can be completely determined y the the values assigned to the parameters of the fit, or can vary according to event-by-event quantities. The most common example of event-by-event quantity that plays such a role is the resolution of the measurement of a property of the events. For example, let the property of the events be the decay length ct and the event-by-event quantity that affects the PDF in ct be σ_{ct}, the event-by-event error on ct. In the sample considered there are two categories of events: signal and background events. In the case that the σ_{ct} distribution is different in case of background events and in signal events, then it is necessary to include a PDF for the separate error distributions [114]. In more detail, the Likelihood function, only considering the subspace of the variable ct can be written as :

$$L(f, c\tau) = \prod_i \Big[f p(x_i, \sigma_i|S) + (1-f) p(x_i, \sigma_i|B) \Big] \tag{4.4.8}$$

where f represent the fraction of signal events and the capital letters S and B correspond to the Signal hypothesis and Background hypothesis. Here in this particular moment the hypothesis is established only on the basis of the value of $c\tau$ parameter of the fit, so in instead of S and B in eq. 4.4.9 one could write $c\tau_S$ and $c\tau_B$[3]. Each probabil-

---

[3] Notice that this is not the real case: hypothesis "Signal" and "Background" depend on a large number of parameters, as is discussed in the whole chapter, but this was just to give a clearer explanation of the problem.

ity distribution $p(x_i, \sigma_i|X)$ (with $X = S, B$) depends on the couple of variables (x_i, σ_i), so they need to be treated as conditional probabilities: $p(x_i, \sigma_i|X) = p(x_i|\sigma_i X)p(\sigma_i|X)$. Therefore the correct, explicit expression for the Likelihood in this particular example is

$$L(f, c\tau) = \prod_i \Big[fp(x_i|\sigma_i S)p(\sigma_i|S) + (1-f)p(x_i|\sigma_i B)p(\sigma_i|B)\Big] \tag{4.4.9}$$

This is the general case, relevant especially when the event-by-event errors distributions are different from signal to background events. In the case they are the same for the two event categories ($p(\sigma_i|B) = p(\sigma_i|B) = p(\sigma_i)$) , the probability distribution for the errors $p(\sigma_i)$ can be factorized out, and since a factor does not affect the maximization, the $p(\sigma_i)$ can be omitted. The reason why the distribution of σ_{ct} are different in the case of signal or background, is that the combinatorial background events have a $c\tau$ that is fitted from a random combination of four tracks, this results in a worse resolution with respect to the signal events.

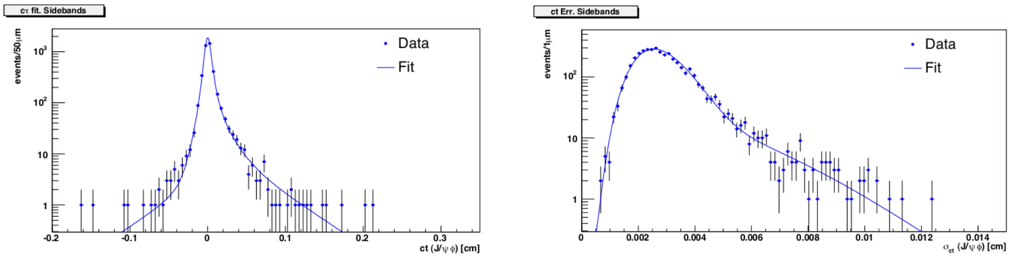
Once this feature of the temporal dependence of the Likelihood has been made clear, Both $p(x_i|\sigma_i X)$ and $p(\sigma_i|X)$ need to be defined. Let us start with the $p(\sigma_i|X)$ in case of signal and background, since they have the same kind of parametrization, only with different parameters. The explicit expression for $P_s(\sigma_{ct}|S) = P_s(\sigma_{ct}|a_1, a_2 b_1, b_2)$. For the decay time error, the PDF has been built using Gamma functions as follows:

$$P_s(\sigma_{ct}|f_{PS}, a_1, b_1, a_2, b_2) = f_{PS}\frac{(\sigma_{ct})^{a_1} e^{\frac{\sigma_{ct}}{b_1}}}{(b_1)^{a_1+1}\Gamma(a_1+1)} + (1-f_{PS})\frac{(\sigma_{ct})^{a_2} e^{\frac{\sigma_{ct}}{b_2}}}{(b_2)^{a_2+1}\Gamma(a_2+1)} \tag{4.4.10}$$

Where a_1, b_1, a_2, b_2 define the mean and the width of respectively the first and the second distribution, and f_{PS} define the fraction of the first distribution. Two distributions with different values for mean and lifetime have been used, and these parameters are found with a previous lifetime only fit on the data. Parameters determined with this method are subsequently used as input in the full likelihood used for the complete analysis. In Fig. 4.5 are shown the ct and the σ_{ct} projections of the only lifetime fit that is performed before the full. The same plots are shown for the signal events and for the background.

The r $P_s(\sigma_{ct}|B) = P_s(\sigma_{ct}|a_{B1}, a_{B2}, b_{B1}, b_{B2})$ for the background has the same form, it uses only different values for the parameters:

$$\begin{aligned} P_b(\sigma_{ct}|f_{PB}, a_{B1}, b_{B1}, a_{B2}, b_{B2}) = f_{PB}&\frac{(\sigma_{ct})^{a_{B1}} e^{\frac{\sigma_{ct}}{b_{B1}}}}{(b_{B1})^{a_{B1}+1}\Gamma(a_{B1}+1)} \\ &+ (1-f_{PB})\frac{(\sigma_{ct})^{a_{B2}} e^{\frac{\sigma_{ct}}{b_2}}}{(b_{B2})^{a_{B2}+1}\Gamma(a_{B2}+1)} \end{aligned} \tag{4.4.11}$$



(a) ct distribution for sidebands events events

(b) σ_{ct} distribution for sidebands events events

Figure 4.4.: ct projections of lifetime only fit.
This preliminary fit is used to determine parameters entering in eq. 4.4.10 .

4.4.3. Background lifetime PDF

The "template" for the background, in the ct variable is built the a strategy explained in the previous paragraph : $P_s(ct, \sigma_{ct}|S) = P_s(ct|\sigma_{ct}, c\tau, S) \cdot P_s(\sigma_{ct}|S)$, and with the $P_s(\sigma_{ct}|S)$ reported in eq. 4.4.11 . The proper decay time function is parametrized as a prompt Gaussian peak, two positive exponentials and a negative exponential. The components of the background description in ct are the following:

- The *prompt peak* models most of the combinatorial background events, that are expected to have no significant lifetime

- The *positive exponentials* are used to describe the longer lived background events such as kaons

- The *negative exponential* is needed to account for those events that present a negative decay length, due to the vertex reconstruction.

The background lifetime PDF turns therefor to be:

$$P_b(t, \sigma_{ct}) =$$
$$\Big[\Big\{ f_{gaus} + (1 - f_{gaus})\Big(f_{++}\frac{1}{\lambda_{++}} e^{-\frac{t}{\lambda_{++}}} + (1 - f_{++})\Big(f_{-}\frac{1}{\lambda_{-}} e^{-\frac{t}{\lambda_{-}}} + (1 - f_{-})\frac{1}{\lambda_{+}} e^{-\frac{1}{\lambda_{+}}} \Big)\Big\}\Big]$$
$$\cdot P_b(\sigma_{ct}|B) \tag{4.4.12}$$

The prompt peak of the background has a relevant meaning: it allows to determine the resolution function in ct. This is important since all the signal time-dependent components of the likelihood need to be convoluted with that function, in order to account for the detector resolution. The resolution function has been modeled using two gaussians, therefore three parameters are used to describe it: two scale factors for the two Gaussian distribution and the fraction of the first with respect to the other.

4.4.4. Smearing function for the signal time PDF components

The purpose of this section is to show how the detector resolution acts on the time-dependent components of the signal PDF. First of all, as mentioned in the previous section, the resolution function exploits two gaussian distribution, its complete expression is:

$$\mathcal{R}(f_1, s_{ct1}, s_{ct1}) = f_1 G_1(ct, \sigma_{ct}|s_{ct1}) + (1 - f_1) G_2(ct, \sigma_{ct}|s_{ct2}) \tag{4.4.13}$$

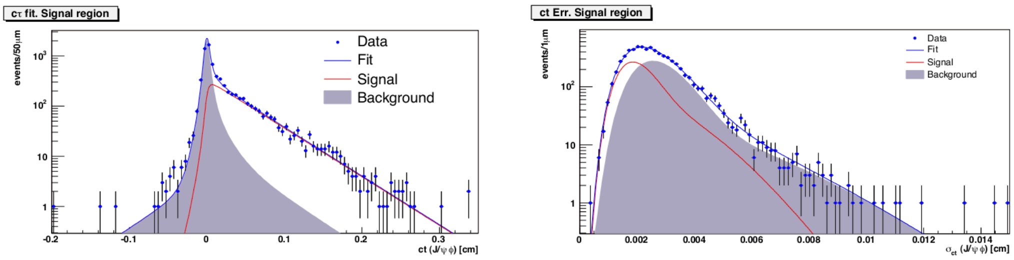
where f_1 is the fraction of the first gaussian, s_{ct1} and s_{ct2} represent the scale factor of the width of the first and of the second gaussian respectively. The σ_{ct} is of course the event by event ct-error , which is multiplied by the two different scale factors, that are left to float in the fit These scale factors account for the overall mis-estimation of the decay time resolution. The resolution function needs to be convoluted with all the time dependent terms, procedure that is known as detector resolution smearing. Ih the case of lifetime only fit and untagged fit, the term in the signal PDF that carries the time dependence has an exponential form, so the smeared exponential that is used in the likelihood is obtained as follows:

$$\begin{aligned}
&\frac{1}{c\tau} e^{-\frac{ct}{c\tau}} \otimes \mathcal{R}(f_1, s_{ct1}, s_{ct1}) = \\
&= \frac{1}{c\tau} e^{-\frac{ct}{c\tau}} \otimes \left[f_1 \frac{1}{\sqrt{2\pi} s_{ct1} \sigma_{ct}} e^{-\frac{c^2 t^2}{2 s_{ct1}^2 \sigma_{ct}^2}} + (1 - f_1) \frac{1}{\sqrt{2\pi} s_{ct2} \sigma_{ct}} e^{-\frac{c^2 t^2}{2 s_{ct2}^2 \sigma_{ct}^2}} \right] \\
&= \left[f_1 \frac{1}{c\tau} e^{-\frac{ct}{c\tau}} \otimes \frac{1}{\sqrt{2\pi} s_{ct1} \sigma_{ct}} e^{-\frac{c^2 t^2}{2 s_{ct1}^2 \sigma_{ct}^2}} + \right. \\
&\left. (1 - f_1) \frac{1}{c\tau} e^{-\frac{ct}{c\tau}} \otimes \frac{1}{\sqrt{2\pi} s_{ct2} \sigma_{ct}} e^{-\frac{c^2 t^2}{2 s_{ct2}^2 \sigma_{ct}^2}} \right]
\end{aligned} \tag{4.4.14}$$

where $\otimes$ represents the convolution.

In the likelihood exploited by the tagged fit, the differential decay rate for $B_s \to J/\psi\phi$ has two oscillating terms that carry the time-dependence, in addition to the exponential already present in the untagged likelihood case (see $\cos \Delta m t$ and $\sin \Delta m t$ dependences in eq. 4.6.7). In these terms relies the difference between the tagged and the untagged fit, for this reason they are important. These oscillating terms will be multiplied by the dilution factor in the full PDF (see sec.4.5, eq. 4.5.3), this will give a combined effect due to both the tagging performances and the detector resolution.
The detector smearing application on the functions $\cos \Delta m t$ and $\sin \Delta m t$ (i.e. the convolution of the functions) has been implemented using a class code from `CLHEP`, a class library for high energy physics [115].

In Fig. 4.5 are shown the ct and the σ_{ct} projections of the only lifetime fit that is performed before the full. The same plots are shown for the signal events and for the background.



(a) ct distribution for signal events

(b) σ_{ct} distribution for signal events

Figure 4.5.: ct projections of lifetime only fit.
This preliminary fit is used to determine parameters entering in eq. 4.4.10 .

4.4.5. Background angular PDF

There are no predictions for the shape of the background transversity angles. so these distributions have been parametrize empirically, from a fit performed on the data from the B_s sidebands . The following functions are found to give a good description of the sidebands data:

$$
\begin{aligned}
f(\cos\theta) &= \frac{a_0 - a_1\cos^2(\theta)}{2a_0 - 2a_1/3} \\
f(\phi) &= \frac{1 + b_1\cos(2\phi + b_0)}{2\pi} \\
f(\cos\psi) &= \frac{c_0 + c_1 cos^2(\theta)}{2c_0 + 2c_1/3}
\end{aligned}
\tag{4.4.15}
$$

The fit on the data from the mass sidebands is used in order to determine the starting values for the parameters $a_{0,1}$, $b_{0,1}$, $c_{0,1}$ that are left floating in the minimization of the full likelihood. In Fig. 4.6 is shown this preliminary fit.
The background angular PDFs are treated as independent from the rest of the likelihood, and each transversity angle distribution is reasonably considered uncorrelated with respect of the other two angles. For this reason each angle distribution is modeled separately: $P(\theta, \phi, \psi) = P(\theta)P(\phi)P(\psi)$

4.5. Flavor Tagging PDFs

Flavor tagging, as seen in Chapter 3 allows to distinguish between B_s and $\bar{B}_s$ at the production, this permits to follow separately the time evolution of B_s and $\bar{B}_s$. Thus it is necessary to include in the temporal and angular PDFs also the tagging decision (B_s or $\bar{B}_s$), the dilution, and the scale factor, for each tagger (OST or SSKT) used.

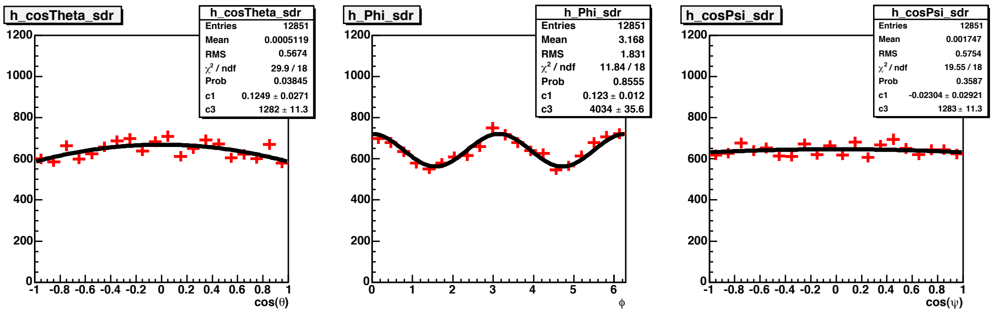


Figure 4.6.: Angular distributions describing the background events.

4.5.1. Tagging in signal decay PDF

For each tagger used, there are two additional variables and one gaussian constraint parameter that needs to be included in the joint PDF used in the Maximum Likelihood (ML) fit: tagging decision ξ, the dilutions $\mathcal{D}$ and the scale factors $S_{\mathcal{D}}$ that have been previously determined with the calibration. In this analysis, the OST calibration has been performed and the results are presented in sec, 3.7.2, while the SSKT has not been performed again and the results of the calibration of the analysis on 5.2 fb^{-1} have been used. The tag decision ξ gives the information on the initial state (B_s or $\bar{B}_s$) of the decaying particle, so it enters in the likelihood as a factor that "chooses" the term of the time and angular dependent decay PDF that must be used. The dilution $\mathcal{D}$ represents a weight associated to each tagged event meaning the quality of the tagging. It is an event-by-event quantity, and plays the same role that the σ_{ct} plays with respect to the ct variable. The scale factors are simply number that multiply the event-by-event dilution. The starting values for the scale factors have been determined with the calibrations, but in the main $B_s \to J/\psi\phi$ fit they are allowed to float within a Gaussian constraint, with σ the uncertainty on the calibration measurement. Form these considerations can be understood that for each tagger two PDFs are need in order to model the signal tagged events:

- one that accounts for the *tagging decision* ξ $P(\xi)$, that would be similar to the sum of two (three) Dirac-delta function, one for $\xi = 1$, one for $\xi = -1$ (and the third for $\xi = 0$)[4]

- a PDF $P(\mathcal{D})$ that accounts for the *dilution*, and has to be included in the full likelihood in a manner completely analogous to how the proper decay length error σ_{ct} is included (see sec. 4.4.2).

---

[4] The option $\xi = 0$ is between brackets since it actually represents the impossibility to establish a tagging decision

$P(\xi)$:

The probability density function that accounts for the tag decision has to take into account the efficiency ϵ of the tagger in making a tagging decision. Considering a single tagger, this PDF can be written as:

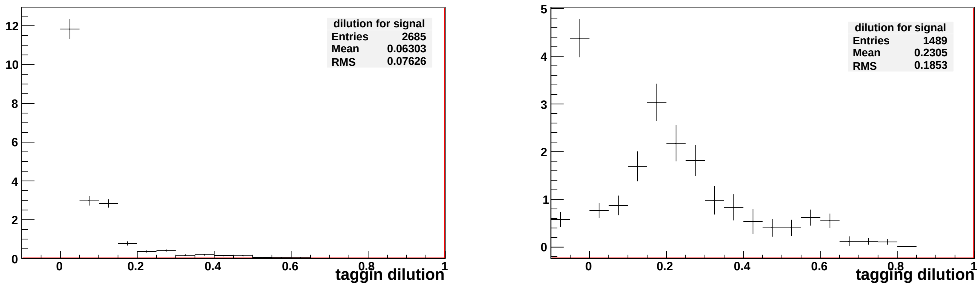
$$P(\xi) = \epsilon \cdot \delta(|\xi| - 1) + (1 - \epsilon) \cdot \delta(\xi - 0) \tag{4.5.1}$$

that can be generalized to two different taggers (indicated with the numbers 1 and 2)

$$P(\xi) \equiv P(\xi_1)P(\xi_2) = \sum_{j=1}^{2} \epsilon_j \cdot \delta(|\xi_j| - 1) + (1 - \sum_{j=1}^{2} \epsilon_j) \cdot \delta(\sum_{j=1}^{2} \xi - 0) \tag{4.5.2}$$

$P(\mathcal{D})$:

The probability density function that accounts for the dilution is modeled with a template that consists of an histogram, taken from the data itself. This kind of templates works exactly as the analytical PDFs that have been used in all the earlier sections, has the same probability meaning, and is used analogously preforming a maximum likelihood fit, the only difference is that it is a discrete object, non a continuous as it is an analytical function. Separate histograms are produced for the signal and the background, different histograms are produced for different taggers. The signal histograms are produced using background subtracted data (with the same procedure of sideband subtraction illustrated in Chapter 3), the background dilution histograms are complementary produced used mass sidebands region data. These distributions are reported in Fig. 4.7 and 4.8 for respectively signal and background and for both OST and SSKT.



(a) dilution histogram of OST tagger for signal

(b) dilution histogram of SST tagger for signal

Figure 4.7.: dilution histograms for signal events.

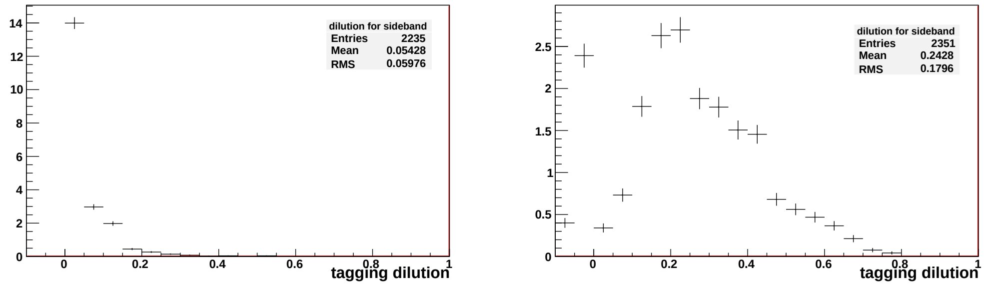
modified decay probability

The decay rate that is use to build the time and angular dependent signal PDF, can be modified to include the tagging variables. Let the starting point be eq. 4.4.3 i.e. the angular and lifetime dependent probability density function including the detector efficiency and the normalization factor $P'(\theta, \phi, \psi, t)$. Using a single tagger the tagging variables incorporation results as follows:

$$T(\theta, \phi, \psi, t, \mathcal{D}, \xi) = \begin{cases} \frac{1-s\mathcal{D}}{2} P'(\theta, \phi, \psi, t) + \frac{1+s\mathcal{D}}{2} \bar{P}'(\theta, \phi, \psi, t) & \text{for } \xi = -1 \\ P'(\theta, \phi, \psi, t) + \bar{P}'(\theta, \phi, \psi, t) & \text{for } \xi = 0 \\ \frac{1-s\mathcal{D}}{2} P'(\theta, \phi, \psi, t) + \frac{1+s\mathcal{D}}{2} \bar{P}'(\theta, \phi, \psi, t) & \text{for } \xi = -1 \end{cases}$$

Now it is needed to include both the tagger algorithms, and here it is chosen to indicate with index 1 the OST and with index 2 the $SSKT$. Moreover in the case of the OST, as presented in sec. 3.7.2 two different scale factors (for K^+ and K^- in the final state - or equivalently B^+ and B^- initial states) have been used. Thus a more compact notation that includes all these aspects is:

$$\begin{aligned} T(\theta, \phi, \psi, t, \mathcal{D}_1, \mathcal{D}_2, \xi_1, \xi_2) = & \frac{1 + \xi_1 s_1^+ \mathcal{D}_1}{1 + |\xi_1|} \frac{1 + \xi_2 s_2 \mathcal{D}_2}{1 + |\xi_2|} P'(\theta, \phi, \psi, t) \\ & + \frac{1 - \xi_1 s_1^- \mathcal{D}_1}{1 + |\xi_1|} \frac{1 - \xi_2 s_2 \mathcal{D}_2}{1 + |\xi_2|} \bar{P}'(\theta, \phi, \psi, t) \end{aligned} \tag{4.5.3}$$



(a) dilution histogram of OST tagger for sidebands (b) dilution histogram of SST tagger for sidebands

Figure 4.8.: Dilution histograms for background events.

$P_b(\xi)$ and $P_b(\mathcal{D})$: PDFs for background flavor fagging

Tagging information does not have a physical meaning for the background events, but it has a meaning in terms of the likelihood construction of the those events. Let us discuss briefly the form of the PDF for the tagging decision $P_b(\xi)$ and for the dilution $P_b(\mathcal{D})$

- $P_b(\xi)$
  it contains a term that accounts for the difference between the number of background events tagged as B_s and $\bar{B}_s$. The parameter that accounts for this contribution is notated as f_{b+} (events that are tagged to contain a $\bar{b}$ quark) and it is left floating in the full likelihood fit. The PDF resulting is:

$$P_b(\xi) = \delta(\xi - 0) + \delta(\xi - 1) \cdot f_{b+} + \delta(\xi + 1) \cdot (1 - f_{b+}) \tag{4.5.4}$$

- $P_b(\mathcal{D})$
  is determinate with histograms filled with data, exactly has it has been done for the signal.

4.6. Full fit likelihood

At this point it is useful to summarize the full likelihood function constructed for all the events, including all the PDFs explained until now and used in the fit to determine the values of the parameters. The likelihood including the tagging informations can be written as follows:

$$\begin{aligned} L = \prod f_s \cdot P_s(m) \cdot T(t,\psi,\theta,\phi) \cdot P_s(\sigma_t) \cdot P_s(\xi) \cdot P_s(D) + \\ (1 - f_s) \cdot P_b(m) \cdot P_b(t,\sigma_t) \cdot P_b(\psi) \cdot P_b(\theta) \cdot P_b(\phi) \cdot P_b(\sigma_t) \cdot P_b(\xi) \cdot P_b(D) \end{aligned} \tag{4.6.1}$$

with:

$$\begin{aligned} T(t,\psi,\theta,\phi) = \frac{1}{N} \cdot \epsilon(\psi,\theta,\phi) \cdot \Big(\frac{1+\xi_1 s_1 D_1}{1+|\xi_1|} \cdot \frac{1+\xi_2 s_2 D_2}{1+|\xi_2|} \cdot P(t,\psi,\theta,\phi) + \\ \frac{1-\xi_1 s_1 D_1}{1+|\xi_1|} \cdot \frac{1-\xi_2 s_2 D_2}{1+|\xi_2|} \cdot \bar{P}(t,\psi,\theta,\phi) \Big) \end{aligned} \tag{4.6.2}$$

And where

$$\begin{aligned} P(\theta,\phi,\psi,t) = P(\vec{\omega},t) \propto \frac{d^3\Lambda(\vec{\omega},t)}{d\vec{\omega}} \propto \sum_k O^k(t) g^k(\vec{\omega}) \\ \bar{P}(\theta,\phi,\psi,t) = \bar{P}(\vec{\omega},t) \propto \frac{d^3\bar{\Lambda}(\vec{\omega},t)}{d\vec{\omega}} \propto \sum_k \bar{O}^k(t) g^k(\vec{\omega}) \end{aligned} \tag{4.6.3}$$

for the expressions of O^k and $\bar{O}^k$ refer to Chapter 1, eq. 1.3.5.

As it has already been introduced in Chapter 1, also a measurement without having the knowledge of the B meson flavor at the production can be preformed. In this case, the likelihood function that has to be used has the following form:

$$\begin{aligned} L^{untag} = f_s \cdot P_s(m) \cdot T^{untag}(t,\psi,\theta,\phi) \cdot P_s(\sigma_t) + \\ (1 - f_s) \cdot P_b(m) \cdot P_b(t,\sigma_t) \cdot P_b(\psi) \cdot P_b(\theta) \cdot P_b(\phi) \cdot P_b(\sigma_t) \end{aligned} \tag{4.6.4}$$

with:

$$T^{untag}(t,\psi,\theta,\phi) = \frac{1}{N}\cdot\epsilon(\psi,\theta,\phi)\cdot P^{untag}(t,\psi,\theta,\phi) \tag{4.6.5}$$

And analogously to the tagged case:

$$P^{untag}(\theta,\phi,\psi,t) = P(\vec{\omega},t) + \bar{P}(\vec{\omega},t) \propto \frac{d^3\Lambda^{untag}(\vec{\omega},t)}{d\vec{\omega}} \propto \sum_k O^{k\ untag} g^k(\vec{\omega}) \tag{4.6.6}$$

for the $O^{k\ untag}$ refer to Chapter 1, eq. 1.3.7.

4.6.1. Symmetries of the likelihood function

Considering the decay probability density function for $B_s \to J/\psi\phi$, the likelihood presents some symmetry properties. In order to show such properties, it is useful to write again the observables O^k of the first Chapter:

$$\begin{aligned}
O^1(t) &= |A_0(t)|^2 = \\
&= |A_0|^2 e^{-\Gamma t}(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma|t}{2} \mp \sin 2\beta_s \sin\Delta m t) \\
O^2(t) &= |A_\parallel(t)|^2 = \\
&= |A_\parallel|^2 e^{-\Gamma t}(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma|t}{2} \mp \sin 2\beta_s \sin\Delta m t) \\
O^3(t) &= |A_\perp(t)|^2 = \\
&= |A_\perp|^2 e^{-\Gamma t}(\cosh\frac{\Delta\Gamma t}{2} + |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma|t}{2} \pm \sin 2\beta_s \sin\Delta m t) \\
O^4(t) &= \mathcal{I}m\{A_\parallel^*(t)A_\perp(t)\} = \\
&= |A_\parallel||A_\perp|e^{-\Gamma t}(\pm\sin(\delta_\perp-\delta_\parallel)\cos\Delta m t \mp \cos(\delta_\perp-\delta_\parallel)\cos 2\beta_s \sin\Delta m t \\
&+ \cos(\delta_\perp-\delta_\parallel)\sin 2\beta_s \sinh\frac{\Delta\Gamma t}{2}) \\
O^5(t) &= \mathcal{R}e\{A_0^*(t)A_\parallel(t)\} = \\
&= |A_0||A_\parallel|e^{-\Gamma t}\cos\delta_\parallel(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma|t}{2} \mp \sin 2\beta_s \sin\Delta m t) \\
O^6(t) &= \mathcal{I}m\{A_0^*(t)A_\perp(t)\} = \\
&= |A_0||A_\perp|e^{-\Gamma t}(\pm\sin\delta_\perp\cos\Delta m t \mp \cos\delta_\perp\cos 2\beta_s + \cos\delta_\perp \sin 2\beta_s \sinh\frac{\Delta\Gamma t}{2})
\end{aligned} \tag{4.6.7}$$

and recall that the decay function is written as $P(\psi,\theta,\phi,t) = \sum_k O^k(t) g^k(\psi,\theta,\phi)$.

This PDF is invariant under the simultaneous transformations:

$$\begin{cases} \beta_s \to \pi/2 - \beta_s \\ \Delta\Gamma \to -\Delta\Gamma \\ \delta_{\parallel} \to 2\pi - \delta_{\parallel} \\ \delta_{\perp} \to \pi - \delta_{\perp} \end{cases}$$

Where the last two transformations correspond to the transformations of the strong phases $\delta_{\perp} \to \pi - \delta_{\perp}$ and $\delta_{\parallel} \to 2\pi - \delta_{\parallel}$.[5]. One can easily seen that under these simultaneous transformations, the observables O^k written above do not change.

In the case of the fit without flavor tagging information, the symmetry under the transformation written above, holds when all the four parameters transform simultaneously, but also when they transform separately for the couples:

$$\begin{cases} \beta_s \to \pi/2 - \beta_s \\ \Delta\Gamma \to -\Delta\Gamma \end{cases}$$

and

$$\begin{cases} \delta_{\parallel} \to 2\pi - \delta_{\parallel} \\ \delta_{\perp} \to \pi - \delta_{\perp} \end{cases}$$

This can be verified by looking at the expressions of the $O^{k\ untag}(t)$ (as was already mentioned in Chapter 1 the decay rate can be written as $P^{untag}(\psi, \theta, \phi, t) = \sum_k O^{k\ untag}(t) g^k(\psi, \theta, \phi)$, similarly to the tagged case), that are reported just below:

$$\begin{aligned}
O^{1\ untag}(t) &= |A_0|^2 e^{-\Gamma t} 2(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s| \sinh\frac{|\Delta\Gamma| t}{2}) \\
O^{2\ untag}(t) &= |A_{\parallel}|^2 e^{-\Gamma t} 2(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s| \sinh\frac{|\Delta\Gamma| t}{2}) \\
O^{3\ untag}(t) &= |A_{\perp}|^2 e^{-\Gamma t} 2(\cosh\frac{\Delta\Gamma t}{2} + |\cos 2\beta_s| \sinh\frac{|\Delta\Gamma| t}{2}) \\
O^{4\ untag}(t) &= |A_{\parallel}||A_{\perp}| e^{-\Gamma t} 2(\cos(\delta_{\perp} - \delta_{\parallel}) \sin 2\beta_s \sinh\frac{\Delta\Gamma t}{2}) \\
O^{5\ untag}(t) &= |A_0||A_{\parallel}| e^{-\Gamma t} cos\delta_{\parallel} 2(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s| \sinh\frac{|\Delta\Gamma| t}{2}) \\
O^{6\ untag}(t) &= |A_0||A_{\perp}| e^{-\Gamma t} 2(\cos\delta_{\perp} \sin 2\beta_s \sinh\frac{\Delta\Gamma t}{2})
\end{aligned} \tag{4.6.8}$$

The use of the flavor tagging removes this symmetry, and this is due to the fact that in 4.6.8 only terms proportional to $\cos\delta_1$, $\cos\delta_2$, $\cos(\delta_2 - \delta_1)$ are left, instead of having also terms proportional to $\sin\delta_1$, $\sin\delta_2$, as it happens in the tagged likelihood case.

---

[5] Recall that $\delta_{\perp}$ and $\delta_{\parallel}$ are defined as $\arg(A_0 A_{\perp}^*)$ and $\arg(A_0 A_{\parallel}^*)$ respectively.

4.6.2. $P(\theta, \phi, \psi, t)$: formalism used to write the PDF

The purpose of this section is to write the decay rate $P(\theta, \phi, \psi, t) = P(\vec{\omega}, t) \propto \frac{d^3\Lambda(\vec{\omega},t)}{d\vec{\omega}}$, in an equivalent and more compact form, useful to be implemented in the likelihood [29]. Since the expressions that will be used in this section are in the form implemented in the joint PDF used for the fit, they are correctly normalized to the unity. For this reason their expressions are completed also with the correct normalization factors, unlike the expression reported in Chapter 1 and in the previous sections. Let us take as starting point the time dependent amplitudes for B_s and $\bar{B}_s$ to reach the states $\mathcal{P}_{\|}$, $\mathcal{P}_{\perp}$, $\mathcal{P}_0$, where the two vector mesons in the final state of our decay (J/ψ and ϕ), have their spins transversely polarized with respect to their momentum and parallel ($\mathcal{P}_{\|}$) or perpendicular ($\mathcal{P}_{\perp}$) to each other. And $\mathcal{P}_0$, they are both longitudinally polarized. When no CP violation is assumed, the heavy, long-lived mass eigenstate is CP-odd and decays to the CP-odd (orbital angular momentum $L = 1$) $\mathcal{P}_{\perp}$. The light-short lived mass eigenstate instead is CP-even and decays to both $L = 0$, and $L = 2$ CP-even final states, which are combinations of the states $\mathcal{P}_0$ and $\mathcal{P}_{\|}$. These time dependent amplitudes are here indicated as $\mathcal{A}_{\|}$ and $\mathcal{A}_{\perp}$, $\mathcal{A}_0$ and they represent the amplitudes for B_s and $\bar{B}_s$ to reach the states $\mathcal{P}_{\|}$, $\mathcal{P}_{\perp}$, $\mathcal{P}_0$ either with or without mixing. Their time-dependence is:

$$
\begin{aligned}
\mathcal{A}_i &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s(\tau_L - \tau_H)}}\Big[E_+(t) \pm e^{2i\beta_s} E_-(t)\Big] a_i \\
\bar{\mathcal{A}}_i &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s(\tau_L - \tau_H)}}\Big[\pm E_+(t) + e^{2i\beta_s} E_-(t)\Big] a_i
\end{aligned}
\tag{4.6.9}
$$

where $i = \{\|, \perp, 0\}$ and the upper signs are referred to a CP-even final state, and the lower to a CP-odd final state,and the a_i are complex amplitude parameters such that $\sum_i |a_i|^2 = 1$. $E_{\pm}$ are the eigenvalues of the problem (cfr ??) and can be written:

$$
E_{\pm}(t) = \Big[e^{+(\frac{-\Delta\Gamma}{4} + i\frac{\Delta m}{2})t} \pm e^{+(\frac{-\Delta\Gamma}{4} + i\frac{\Delta m}{2})t}\Big] \tag{4.6.10}
$$

At this point it is need to recall the system of coordinates that we are using: the transversity basis (see sec. 1.3). Within that reference system the muon momentum direction in the J/ψ rest frame is given by the unit vector:

$$
\hat{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \tag{4.6.11}
$$

And $\boldsymbol{A}(t)$, $\bar{\boldsymbol{A}}(t)$ complex vector functions can be defined as follows:

$$
\begin{aligned}
\boldsymbol{A}(t) &= \Big(\mathcal{A}_0\cos\psi, -\frac{\mathcal{A}_{\|}(t)\sin\psi}{\sqrt{2}}, i\frac{\mathcal{A}_{\perp}(t)\sin\psi}{\sqrt{2}}\Big) \\
\bar{\boldsymbol{A}}(t) &= \Big(\bar{\mathcal{A}}_0\cos\psi, -\frac{\bar{\mathcal{A}}_{\|}(t)\sin\psi}{\sqrt{2}}, i\frac{\bar{\mathcal{A}}_{\perp}(t)\sin\psi}{\sqrt{2}}\Big)
\end{aligned}
\tag{4.6.12}
$$

where $i = \{\|, \perp, 0\}$ and the upper signs are referred to a CP-even final state, and the lower to a CP-odd final state, the eigenvalues of the problem (cfr ??), and the $\mathcal{A}_i$ have been normalized. At this point the normalized probability density functions for B and $\bar{B}$ mesons can be written in the variables $t, \cos\psi, \cos\theta, \phi$:

$$\begin{aligned} P(\theta,\phi,\psi,t) &= \frac{9}{16\pi}|\boldsymbol{A}(t)\times\hat{n}|^2 \\ \bar{P}(\theta,\phi,\psi,t) &= \frac{9}{16\pi}|\bar{\boldsymbol{A}}(t)\times\hat{n}|^2 \end{aligned} \tag{4.6.13}$$

These expressions are completely equivalent to eq. 4.6.3 and all the formalism that has been used until now. As in the first chapter the observables O^k and the angular functions g^k have been made explicit and it was put in evidence how the O^ks carry the temporal dependence and the dependence on the parameters of interest as β_s and $\Delta\Gamma$, and the g^ks describe the angular dependence, analogously now eq. 4.6.13 will be clarified. First of all, $\mathcal{A}_i(t=0)$ will be indicated as $A_i(0)$, and since it is not ambiguous, also the indication (0) meaning $t=0$ is dropped. Then the quantities $|a_i|^2$ represent the time independent rate to each polarization state. See sec. 4.6.3 to understand how these two set of quantities are related. Now let us consider the time dependence. The time development of $\mathcal{A}_0(t)$ and $\mathcal{A}_\|(t)$ are identical, and they are different from that of $\mathcal{A}_\perp(t)$. Thus $A_+(t)$ and $A_-(t)$ are defined as:

$$\boldsymbol{A}(t) = \boldsymbol{A}_+(t) + \boldsymbol{A}_-(t) \qquad\qquad \bar{\boldsymbol{A}}(t) = \bar{\boldsymbol{A}}_+(t) + \bar{\boldsymbol{A}}_-(t) \tag{4.6.14}$$

And at this point it is possible to separate the angular dependence from the time dependence:

$$\begin{aligned} \boldsymbol{A}_+(t) &= \boldsymbol{A}_+ f_+(t) = (a_0\cos\psi, -\frac{a_\|\sin\psi}{\sqrt{2}}, 0)\cdot f_+(t) \\ \bar{\boldsymbol{A}}_+(t) &= \boldsymbol{A}_+ \bar{f}_+(t) = (a_0\cos\psi, -\frac{a_\|\sin\psi}{\sqrt{2}}, 0)\cdot \bar{f}_+(t) \end{aligned} \tag{4.6.15}$$

and

$$\begin{aligned} \boldsymbol{A}_-(t) &= \boldsymbol{A}_- f_-(t) = (0, 0, i\frac{a_\perp\sin\psi}{\sqrt{2}})\cdot f_-(t) \\ \bar{\boldsymbol{A}}_-(t) &= \boldsymbol{A}_+ \bar{f}_-(t) = (0, 0, i\frac{a_\perp\sin\psi}{\sqrt{2}})\cdot \bar{f}_-(t) \end{aligned} \tag{4.6.16}$$

where $f_\pm$ and $\bar{f}_\pm$ are defined:

$$\begin{aligned} f_\pm &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s(\tau_L - \tau_H)}}\Big[E_+ \pm e^{2i\beta_s}E_-(t)\Big] \\ \bar{f}_\pm &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s(\tau_L - \tau_H)}}\Big[\pm E_+ + e^{2i\beta_s}E_-(t)\Big] \end{aligned} \tag{4.6.17}$$

For the decay rate the following expression is obtained:

$$\begin{aligned} P(\theta,\phi,\psi,t) &= \frac{9}{16\pi}\Big\{|\boldsymbol{A}_+(t)\times\hat{n}|^2 + |\boldsymbol{A}_-(t)\times\hat{n}|^2 + 2\mathcal{R}e((\boldsymbol{A}_+(t)\times\hat{n})\cdot(\boldsymbol{A}_-(t)\times\hat{n}))\Big\} \\ &= \frac{9}{16\pi}\Big\{|\boldsymbol{A}_+\times\hat{n}|^2|f_+(t)|^2 + |\boldsymbol{A}_-\times\hat{n}|^2|f_-(t)|^2 + \\ &\mathcal{R}e((\boldsymbol{A}_+\times\hat{n})\cdot(\boldsymbol{A}_-\times\hat{n}))\cdot f_+(t)f_-^*(t)\Big\} \end{aligned} \tag{4.6.18}$$

and similarly for the $\bar{P}(\theta,\phi,\psi,t)$:

$$\begin{aligned} \bar{P}(\theta,\phi,\psi,t) = \frac{9}{16\pi}\Big\{&|\boldsymbol{A}_+\times\hat{n}|^2|\bar{f_+}(t)|^2 + |\boldsymbol{A}_-\times\hat{n}|^2|f_-\bar{}(t)|^2 + \\ &\mathcal{R}e((\boldsymbol{A}_+\times\hat{n})\cdot(\boldsymbol{A}_-\times\hat{n}))\cdot f_+\bar{}(t)\bar{f_-}^*(t)\Big\} \end{aligned} \tag{4.6.19}$$

Should be noticed how the angular dependence is carried by the vector products of the type "$\boldsymbol{A}\times\hat{n}$ " that play the same role of the g^k used in Chapter 1 , while the temporal dependence that here is included in the $f_\pm$, in Chapter 1 was included in the set of observables O^k. The difference is that the polarization amplitudes at $t = 0$ here are included in the "angular part" ("$\boldsymbol{A}\times\hat{n}$ ") while in the previous formalism they were appearing in the "temporal part" (O^k).

In order to give an idea of the consistency of the two formalisms used in this thesis, it could be useful to write explicitly the terms entering in eq. 4.6.18 4.6.19 [6].

---

[6] A substitution of the expressions in 4.6.20 and 4.6.21 in eq. 4.6.18 and eq. 4.6.19 will give $P(\theta,\phi,\psi,t) = \sum_k Q^k(t)g^k(\psi,\theta,\phi)$ completely analogous to the expression used in Chapter 1 for the differential decay rate. The expression for the angular functions will be the same, while for time-dependent observables a further step is needed in order to find $O^k = Q^k$. This step relies in the fact that the O^k of Chapter 1 are not written explicitly in terms of the mass eigenstates B_s^L and B_s^H, while here it happens for the Q^k.

$$
\begin{aligned}
|f_{\pm}|^2 &= \frac{1}{2}\frac{(1 \pm \cos 2\beta_s)e^{-\Gamma_L t} + (1 \mp \cos 2\beta_s)e^{-\Gamma_H t} \mp 2\sin 2\beta_s e^{-\Gamma t} sin\Delta mt}{\tau_L(1 \pm \cos 2\beta_s) + \tau_H(1 \mp \cos 2\beta_s)} \\
|\bar{f_{\pm}}|^2 &= \frac{1}{2}\frac{(1 \pm \cos 2\beta_s)e^{-\Gamma_L t} + (1 \mp \cos 2\beta_s)e^{-\Gamma_H t} \pm 2\sin 2\beta_s e^{-\Gamma t} sin\Delta mt}{\tau_L(1 \pm \cos 2\beta_s) + \tau_H(1 \mp \cos 2\beta_s)} \\
f_+(t)f_-^*(t) &= \frac{e^{-\Gamma t} cos\Delta mt + i\cos 2\beta_s e^{-\Gamma t} sin\Delta mt + i\sin 2\beta_s(e^{-\Gamma_L t} - e^{-\Gamma_H t})/2}{\sqrt{[(\tau_L - \tau_H)\sin 2\beta_s]^2 + 4\tau_L\tau_H}} \\
\bar{f_+}(t)\bar{f_-}^*(t) &= \frac{-e^{-\Gamma t} cos\Delta mt - i\cos 2\beta_s e^{-\Gamma t} sin\Delta mt + i\sin 2\beta_s(e^{-\Gamma_L t} - e^{-\Gamma_H t})/2}{\sqrt{[(\tau_L - \tau_H)\sin 2\beta_s]^2 + 4\tau_L\tau_H}}
\end{aligned}
\tag{4.6.20}
$$

And for the angular dependence:

$$
\begin{aligned}
|\boldsymbol{A}_+ \times \hat{n}| &= \Big(-i\frac{a_{\parallel}}{\sqrt{2}}\sin\psi\cos\theta, -a_0\cos\psi\cos\theta, a_0\cos\psi\sin\theta\sin\phi + i\frac{a_{\parallel}}{\sqrt{2}}\sin\psi\sin\theta\cos\phi\Big) \\
|\boldsymbol{A}_- \times \hat{n}| &= \Big(i\frac{a_{\perp}}{\sqrt{2}}\sin\psi\sin\theta\sin\phi, -\frac{a_{\perp}}{\sqrt{2}}\sin\psi\sin\theta\cos\phi, 0\Big)
\end{aligned}
\tag{4.6.21}
$$

The PDF $P(\theta,\phi,\psi,t)$ in the likelihood function used is written in the form that has just been presented, with the only difference that all the time-dependent components need to be replaced with the smeared equivalent as has been explained in sec. 4.4.2.

4.6.3. $|A_0|^2$and$|A_{\parallel}|^2$

The complete and explicit form of the $P(\theta,\phi,\psi,t)$ PDF has just been presented in order to clarify what are the parameters that the maximum likelihood fit is supposed to determine. The parameters entering in the fit that are giving the informations about the polarization amplitudes $\mathcal{A}_i$ are not directly $|A_i|^2$ but the a_i, more precisely two functions of the a_i presented in this paragraph. It is therefore useful to relate the the A_i to the a_i:

$$
\begin{aligned}
|A_{\perp}|^2 &= \frac{|a_{\perp}|^2 y}{1 + (y-1)|a_{\perp}|^2} \\
|A_{\parallel}|^2 &= \frac{|a_{\parallel}|^2}{1 + (y-1)|a_{\perp}|^2} \\
|A_0|^2 &= \frac{|a_0|^2}{1 + (y-1)|a_{\perp}|^2}
\end{aligned}
\tag{4.6.22}
$$

where $y \equiv (1+z)/(1-z)$ and $z \equiv \cos 2\beta_s\Delta\Gamma/(2\Gamma)$. The parameters really used in the fit are: α_{CPOdd} and $\alpha_{\|}$ defined in the following manner:

$$\begin{aligned} |a_{\perp}|^2 &= \alpha_{CPOdd} \\ |a_{\|}|^2 &= (1-\alpha_{CPOdd})\alpha_{\|} \\ |a_0|^2 &= (1-\alpha_{CPOdd})(1-\alpha_{\|}) \end{aligned} \tag{4.6.23}$$

Notice that α_{CPOdd} and $\alpha_{\|}$ have different meanings: the first is $|a_{\perp}|^2$ that behaves like and amplitude squared, while $\alpha_{\|}$ is like the fraction of the CP-even final state amplitude that can be attributed to the parallel ($L = 2$) polarization state. Using eq. 4.6.24 together with eq. 4.6.23, the relation between the polarization amplitudes and the fit parameters turns to be:

$$\begin{aligned} |A_{\perp}|^2 &= \frac{\alpha_{CPOdd}y}{1+(y-1)\alpha_{CPOdd}} \\ |A_{\|}|^2 &= \frac{(1-\alpha_{CPOdd})\alpha_{\|}}{1+(y-1)\alpha_{CPOdd}} \\ |A_0|^2 &= \frac{(1-\alpha_{CPOdd})(1-\alpha_{\|})}{1+(y-1)\alpha_{CPOdd}} \end{aligned} \tag{4.6.24}$$

with $y \equiv (1+z)/(1-z)$ and $z \equiv \cos 2\beta_s \Delta\Gamma/(2\Gamma)$ as in eq. 4.6.24.

4.7. S-wave

Since a contribution from S-wave K^+K^- under the ϕ peak in $B_s \to J/\psi\phi$ decay can contribute up to few percents of the total rate, a normalized probability density for the decay $B_s \to J/\psi K^+K^-$ (kaons in an S-wave state) has been added to the likelihood function. These kaons can either a non resonant pair of kaons, or the decay products of a scaler particle, the $f_0(980)$. In that case the final state $J/\psi f_0(980)$ can be only in S-wave, since the decay is a $P \to VS$ decay. The same formalism discussed in sec. 4.6.2 will be used. The starting point is again the polarization vector in the J/ψ rest frame, that is indicated as $\boldsymbol{B}(t)$ (analogous to $\boldsymbol{A}(t)$). The probability density function $Q(\theta, \phi, \psi, t)$ (analogous to $P(\theta, \phi, \psi, t)$) can be written as:

$$\begin{aligned} Q(\theta, \phi, \psi, t) &= \frac{3}{16\pi}|\boldsymbol{B}(t) \times \hat{n}|^2 \\ \bar{Q}(\theta, \phi, \psi, t) &= \frac{3}{16\pi}|\bar{\boldsymbol{B}}(t) \times \hat{n}|^2 \end{aligned} \tag{4.7.1}$$

That do not depend on the ψ angle. In the expressions above $\boldsymbol{B}(t) = (\mathcal{B}(t), 0, 0)$ and $\bar{\boldsymbol{B}}(t) = (\bar{\mathcal{B}}(t), 0, 0)$, with the time dependent amplitudes that can be written as:

$$
\begin{aligned}
\mathcal{B}_i &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s(\tau_L - \tau_H)}}\Big[E_+ - e^{2i\beta_s}E_-(t)\Big] \\
\bar{\mathcal{B}}_i &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s(\tau_L - \tau_H)}}\Big[-E_+ + e^{2i\beta_s}E_-(t)\Big]
\end{aligned}
\tag{4.7.2}
$$

From these expressions, the presence of the ” $-$ ” sign instead of the ” $\pm$ ” indicates the CP properties of the $J/\psi KK$ final state that is CP-odd. Both P-wave and S-wave are present, so both the amplitudes must be summed and squared in order to obtain the correct probability density function. The P-wave has a resonant structure (the ϕ) while the S-wave is flat, but can have any phase with respect to the P-wave. The S-wave contribution is modeled as flat even if it includes the contribution of the $f_0(980)$ decay product kaons, because in correspondence of the ϕ mass peak position, the mass distribution due to the $f_0(980)$ (a peak that can be modeled with an asymmetric Breit-Wigner, known as Flatte‘ distribution, that is usually exploit in experiments dealing with $f_0(980)$ [116]) can already be considered flat. In our analysis events for which the reconstructed mass $m(K^+K^-) = \mu$ in the window $[\mu_{min}, \mu_{max}]$ dictated by the preselection requrements. The normalized probability for the differential decay rate becomes a function of angels, time, and KK-mass. It is written:

$$
\begin{aligned}
\rho(\theta, \phi, \psi, t, \mu) &= \frac{9}{16\pi}\Big|\Big[\sqrt{1 - F_s}\, g(\mu)\boldsymbol{A}(t) + e^{i\delta_s}\sqrt{F_s}\frac{h(\mu)}{\sqrt{3}}\boldsymbol{B}(t)\Big] \times \hat{n}\Big|^2 \\
\bar{\rho}(\theta, \phi, \psi, t, \mu) &= \frac{9}{16\pi}\Big|\Big[\sqrt{1 - F_s}\, g(\mu)\bar{\boldsymbol{A}}(t) + e^{i\delta_s}\sqrt{F_s}\frac{h(\mu)}{\sqrt{3}}\bar{\boldsymbol{B}}(t)\Big] \times \hat{n}\Big|^2
\end{aligned}
\tag{4.7.3}
$$

where to model the ϕ resonance a relativistic asymmetric Breit-Wigner has been used $g(\mu)$, such that its squared module turns to be :

$$
|g(\mu)|^2 = \frac{\mu}{m_\phi} \cdot \Gamma_1 \cdot \frac{k^*(B_s, \mu, J/\psi)}{k^*(B_s, m_\phi, J/\psi)} \cdot \frac{1}{(m_\phi^2 - \mu^2)^2 + m_\phi^2 \cdot \Gamma_{t\phi}^2}
\tag{4.7.4}
$$

where the particle momentum is given by the k^* terms. This treatment assumes a two body decay, where the other daughter particle is the J/ψ, and the total decay width is $\Gamma_\phi = \Gamma_1 + \Gamma_2 + \Gamma_3$, where $\Gamma_{1,2,3}$ are the partial decay widths for the decays ϕ decays: $\phi \to K^+K^-(48.8 \pm 0.5\%)$, $\phi \to K_L^0K_S^0(34.2 \pm 0.4\%)$, and $\phi \to \rho\pi + \pi^+ + \pi^- + \pi^0(15.32 \pm 0.32\%)$ [27].

and for the S-wave a flat model has been used:

$$
h(\mu) = \frac{1}{\sqrt{\Delta\mu}}
\tag{4.7.5}
$$

with the following definitions:

$$\begin{aligned} \omega_{max} &= \tan^{-1}\frac{2(\mu_{max}-\mu_\phi)}{\Gamma_\phi}; \qquad & \omega_{min} &= \tan^{-1}\frac{2(\mu_{min}-\mu_\phi)}{\Gamma_\phi}; \\ \Delta\mu &= \mu_{max}-\mu_{min}; & \Delta\omega &= \omega_{max}-\omega_{min}. \end{aligned} \tag{4.7.6}$$

In all these relations F_s is the S-wave fraction and δ_s is the phase of the S-wave component relative to the P-wave component, μ_ϕ is the ϕ mass (1019 $\mathrm{MeV/c^2}$), $\Gamma\phi$ is the ϕ width (4.26 $\mathrm{MeV/c^2}$). The presence of the S-wave component, will obviously affect also the normalization. In order to calculate the normalization factor correctly and write more explicitly the decay probability density functions, the following quantities need to be defined:

$$\mathcal{F}(\mu) \equiv \sqrt{F_s(1-F_s)}h(\mu)g(\mu)e^{-i\delta_s} \tag{4.7.7}$$

and

$$\mathcal{I}(\mu) \equiv \int \mathcal{F}(\mu)d\mu \tag{4.7.8}$$

Now the probability density functions can be written, expanding eq. 4.7.3

$$\begin{aligned} \rho(\theta,\phi,\psi,t,\mu) = {} & (1-F_s)\frac{\Gamma_\phi/2}{\Delta\omega}\cdot\frac{1}{(\mu-\mu_\phi)^2+\Gamma_\phi^2/4}\cdot P(\theta,\phi,\psi,t) \\ & + F_s\frac{1}{\Delta\mu}Q(\theta,\phi,\psi,t) \\ & + 2\frac{\sqrt{27}}{16\pi}\mathcal{R}e\Big[\mathcal{F}_\mu((\boldsymbol{A}_-\times\hat{n})\cdot(\boldsymbol{B}\times\hat{n})\cdot|f_-(t)|^2 \\ & + (\boldsymbol{A}_+\times\hat{n})\cdot(\boldsymbol{B}\times\hat{n})\cdot f_+(t)\cdot f_-^*(t))\Big] \end{aligned} \tag{4.7.9}$$

$$\begin{aligned} \bar\rho(\theta,\phi,\psi,t,\mu) = {} & (1-F_s)\frac{\Gamma_\phi/2}{\Delta\omega}\cdot\frac{1}{(\mu-\mu_\phi)^2+\Gamma_\phi^2/4}\cdot \bar{P}(\theta,\phi,\psi,t) \\ & + F_s\frac{1}{\Delta\mu}\bar{Q}(\theta,\phi,\psi,t) \\ & + 2\frac{\sqrt{27}}{16\pi}\mathcal{R}e\Big[\mathcal{F}_\mu((\boldsymbol{A}_-\times\hat{n})\cdot(\boldsymbol{B}\times\hat{n})\cdot|\bar{f}_-(t)|^2 \\ & + (\boldsymbol{A}_+\times\hat{n})\cdot(\boldsymbol{B}\times\hat{n})\cdot \bar{f}_+(t)\cdot \bar{f}_-{}^*(t))\Big] \end{aligned} \tag{4.7.10}$$

Integrating out the ϕ-mass variable μ, the probability density functions results:

$$\begin{aligned}\rho(\theta,\phi,\psi,t) = (1-F_s)\cdot P(\theta,\phi,\psi,t) + F_s\frac{1}{\Delta\mu}Q(\theta,\phi,\psi,t)\\ + 2\frac{\sqrt{27}}{16\pi}\mathcal{R}e\Big[\mathcal{I}_\mu((\boldsymbol{A}_- \times \hat{n})\cdot(\boldsymbol{B}\times\hat{n})\cdot|f_-(t)|^2\\ + (\boldsymbol{A}_+ \times \hat{n})\cdot(\boldsymbol{B}\times\hat{n})\cdot f_+(t)\cdot f_-^*(t))\Big]\end{aligned} \tag{4.7.11}$$

$$\begin{aligned}\bar{\rho}(\theta,\phi,\psi,t) = (1-F_s)\cdot \bar{P}(\theta,\phi,\psi,t) + F_s\frac{1}{\Delta\mu}\bar{Q}(\theta,\phi,\psi,t)\\ + 2\frac{\sqrt{27}}{16\pi}\mathcal{R}e\Big[\mathcal{I}_\mu((\boldsymbol{A}_- \times \hat{n})\cdot(\boldsymbol{B}\times\hat{n})\cdot|\bar{f}_-(t)|^2\\ + (\boldsymbol{A}_+ \times \hat{n})\cdot(\boldsymbol{B}\times\hat{n})\cdot \bar{f}_+(t)\cdot \bar{f}_-{}^*(t))\Big]\end{aligned} \tag{4.7.12}$$

Due to the introduction of the S-wave contribution , the likelihood is not symmetric according to the transformation described in sec. 4.6.1 any more. But, if μ is integrated over a symmetric interval with respect to to the ϕ mass peak, the likelihood (that now is built in an enlarge parameter space including μ_ϕ and Γphi), is symmetric under the following simultaneous transformations:

$$\begin{cases}\beta_s \to \pi/2 - \beta_s\\ \Delta\Gamma \to -\Delta\Gamma\\ \delta_{\|} \to 2\pi - \delta_{\|}\\ \delta_\perp \to \pi - \delta_\perp\\ \delta_s \to \pi - \delta_s\end{cases}$$

Where the first four transformations are the same of sec. 4.6.1 . In order to give an idea of what happens: if the contribution to the integral coming from a slice in ϕ mass of one side of the peak and the one from the symmetrically-located slice are considered, while the contribution of either slice is not invariant under the transformation above, the contribution of both slices taken together turns to be invariant. [29].

The functions 4.7.11 and 4.7.12 can also be written with the that formalism including the S-wave contribution (and the interference term between S-wave and P-wave) the decay probability density function becomes:

$$\rho(\theta,\phi,\psi,t) \propto \sum_{k=1}^{10} O^k(t) g^k(\theta,\phi,\psi) \tag{4.7.13}$$

where the $g^{1-6}(\psi,\theta,\phi)$ have been defined in Chapter 1 (eq. 1.3.4) the four additional angular functions are:

$$\begin{aligned}
g^7(\psi,\theta,\phi) &= 2(1-\sin^2\theta\cos^2\phi)\\
g^8(\psi,\theta,\phi) &= 2\cos\psi(1-\sin^2\theta\cos^2\phi)\\
g^9(\psi,\theta,\phi) &= \frac{1}{\sqrt{2}}2\sin\psi\sin^2\theta\sin 2\phi\\
g^{10}(\psi,\theta,\phi) &= \frac{1}{\sqrt{2}}2\sin\psi\sin 2\theta\cos\phi
\end{aligned} \tag{4.7.14}$$

Let us conclude with the explicit expressions for the observables O^k, entering in eq. 4.7.13. Since a new decay amplitude has been introduced, should be pointed out that the phases of the decay amplitudes are defined by $A_j = |A_j|e^{-i\delta_j}$, with $j = \{0, \|, \perp, s\}$. In previous expressions, was kept into account the fact that only relative strong phases differences can be measured, so the common convention is to take $\delta_0 = 0$. Then in Chapter 1 and in previous sections also the replacements $\delta_2 = \delta_\perp$ and $\delta_1 = \delta_\perp - \delta_\|$ have been adopted. The following expressions instead are written with the strong phases δ_j

with the usual definition recalled here.

$$
\begin{aligned}
O^1(t) &= |A_0(t)|^2 = \\
&= |A_0|^2 e^{-\Gamma t}(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma|t}{2} \mp \sin 2\beta_s \sin\Delta m t) \\
O^2(t) &= |A_{\|}(t)|^2 = \\
&= |A_{\|}|^2 e^{-\Gamma t}(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma|t}{2} \mp \sin 2\beta_s \sin\Delta m t) \\
O^3(t) &= |A_{\perp}(t)|^2 = \\
&= |A_{\perp}|^2 e^{-\Gamma t}(\cosh\frac{\Delta\Gamma t}{2} + |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma|t}{2} \pm \sin 2\beta_s \sin\Delta m t) \\
O^4(t) &= \mathcal{R}e\{A_0^*(t)A_{\|}(t)\} = \\
&= |A_0||A_{\|}|e^{-\Gamma t}\cos(\delta_{\|}-\delta_0)(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma|t}{2} \mp \sin 2\beta_s \sin\Delta m t) \\
O^5(t) &= \mathcal{I}m\{A_{\|}^*(t)A_{\perp}(t)\} = \\
&= |A_{\|}||A_{\perp}|e^{-\Gamma t}(\pm\sin(\delta_{\perp}-\delta_{\|})\cos\Delta m t \mp \cos(\delta_{\perp}-\delta_{\|})\cos 2\beta_s \sin\Delta m t \\
&+ \cos(\delta_{\perp}-\delta_{\|})\sin 2\beta_s \sinh\frac{\Delta\Gamma t}{2}) \\
O^6(t) &= \mathcal{I}m\{A_0^*(t)A_{\perp}(t)\} = \\
&= |A_0||A_{\perp}|e^{-\Gamma t}(\pm\sin(\delta_{\perp}-\delta_0)\cos\Delta m t \mp \cos(\delta_{\perp}-\delta_0)\cos 2\beta_s \\
&+ \cos(\delta_{\perp}-\delta_0)\sin 2\beta_s \sinh\frac{\Delta\Gamma t}{2})t \\
O^7(t) &= |A_s(t)|^2 = \\
&= |A_s|^2 e^{-\Gamma t}(\cosh\frac{\Delta\Gamma t}{2} + |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma|t}{2} \pm \sin 2\beta_s \sin\Delta m t) \\
O^8(t) &= \mathcal{I}m\{A_s^*(t)A_{\perp}(t)\} = \\
&= |A_s||A_{\perp}|e^{-\Gamma t}\sin(\delta_{\perp}-\delta_s)(\cosh\frac{\Delta\Gamma t}{2} - |\cos 2\beta_s|\sinh\frac{|\Delta\Gamma|t}{2} \pm \sin 2\beta_s \sin\Delta m t) \\
O^9(t) &= \mathcal{R}e\{A_s^*(t)A_{\|}(t)\} = \\
&= |A_s||A_{\|}|e^{-\Gamma t}(\pm\cos(\delta_{\|}-\delta_s)\cos\Delta m t \mp \cos(\delta_{\|}-\delta_s)\cos 2\beta_s \sin\Delta m t \\
&+ \sin(\delta_{\|}-\delta_s)\sin 2\beta_s \sinh\frac{\Delta\Gamma t}{2}) \\
O^{10}(t) &= \mathcal{R}e\{A_s^*(t)A_0(t)\} = \\
&= |A_a||A_0|e^{-\Gamma t}(\pm\cos(\delta_0-\delta_s)\cos\Delta m t \mp \sin(\delta_0-\delta_s)\cos 2\beta_s \\
&+ \sin(\delta_0-\delta_s)\sin 2\beta_s \sinh\frac{\Delta\Gamma t}{2})t
\end{aligned}
\tag{4.7.15}
$$

where in all the relations the upper signs are referred to B_s and the lower signs to $\bar{B}_s$.

4.8. Summary of the fit variables

At this point it could be useful to summarize the parameters that are supposed to be determined by the maximum likelihood fit. Table 4.1 lists all these parameters with its very short description. The first raws of the table include the physics parameters of interest ,in particular in red are reported the quantities that are measured in this thesis. Follow the quantities involved in the mass PDF, the lifetime PDF, the model for the background events in ct and in the angular variables, and finally the tagging variables. Each of these categories of parameters are explained in detail in the previous sections. Some of the so called nuisance parameters are left free to float within a gaussian constrain, as it happen for the tagging scale factor. In this table are not listed the parameters used to model the error lifetime PDF, since they are not floating in the full fit, but they are determined with a previous only lifetime fit and are fixed in the full fit to determine the parameters of interest.

Parameter	Description
β_s	β_s CP -violating phase
$\Delta\Gamma$	$\Gamma_L - \Gamma_H$
$\alpha_\perp$	CP odd fraction
$\alpha_\parallel$	fraction in CP even states
$\delta_\perp$	$arg(A_\perp A_0)$
$\delta_\parallel$	$arg(A_\parallel A_0)$
$c\tau$	$\frac{1}{\Gamma_s} = \frac{2}{\Gamma_L+\Gamma_H}$
A_{SW}	fraction of S-wave KK component in the signal
δ_{SW}	phase of S-wave component
Δm_s	B_s^0 mixing frequency
f_s	Signal fraction
s_m	Mass error scale factor
m	B_s mass [GeV/c^2]
p_1	mass background slope
$s_{c\tau}1$	lifetime error scale factor 1
$s_{c\tau}2$	lifetime error scale factor 2
f_{sf_1}	fraction of first lifetime error scale factor
f_p	fraction of prompt background
f_-	fraction of bkg which decays with λ_-
f_{++}	fraction of bkg which decays with λ_{++}
λ_-	Effective bkg lifetime, neg. component
λ_+	Effective bkg lifetime, pos. component 1
λ_{++}	Effective bkg lifetime, pos. component 2
ϕ_1	first parameter in bkg fit to ϕ
$cos(\psi)_1$	first parameter in bkg fit to cos(ψ)
$cos(\theta)_1$	first parameter in bkg fit to cos(θ)
$S_{D^+}(OST+)$	OST + dilution scale factor
$S_{D^-}(OST-)$	OST - dilution scale factor
$S_D(SST)$	SST dilution scale factor
$\epsilon_b(OST)$	OST tagging efficiency for background
$\epsilon_b(SST)$	SST tagging efficiency for background
$A^+(OST)$	OST background positive tag asymmetry
$A^+(SST)$	SST background positive tag asymmetry
$\epsilon_s(OST)$	OST tagging efficiency for signal
$\epsilon_s(SST)$	SST tagging efficiency for signal

Table 4.1.: Fit parameters. Starting from the physics parameters going to the nuisance ones. RED: parameters that are supposed to be determined with the analysis presented in this thesis, GREEN: parameters describing the S-wave, PURPLE: mass PDF, BLUE: lifetime PDF, ORANGE: temporal model for the background, YELLOW: angular PDF for the background, PINK: tagging parameters.

Chapter 5

Results

Contents

The purposes of this chapter are: to show the investigations on the consistency of our likelihood fit that have been performed, to presents the results obtained by the fit to our data sample, together with a comparison between the results obtained by using the untagged fit and a consistency check with the previous analysis . The techniques used to verify the good behavior of the fit include pull studies, the probe on the sensibility of the fit towards small changed in the inputs values and distributions, and the examination of the fit projections. The consistency check with the previous analysis is carried out dividing the dataset in two subsamples and fitting them separately. Last section of the chapter describes the method to obtain the systematic uncertainties that need to be associated to the measured value of the physics parameters of interest.

5.1. Fitter validation

5.1.1. Pull studies

Once understood the likelihood function, before carrying out the fit to the data sample, we investigated its consistency to probe whether biases are present in values of the fitted parameters. A standard technique is used is the *pull studies*.

Usually for a given x which has a Gaussian distribution with mean μ and width σ, the pull quantity is defined as follows:

$$g = \frac{x - \mu}{\sigma} \tag{5.1.1}$$

and the pull distribution is expected to be Gaussian by definition. Also in case of non-Gaussian variables, the pull distribution is expected to be gaussian, due to the central limit theorem. We are interested in pull distributions of the fit parameters. in this case the expression for the pull is analogously written , with the difference that the mean μ is replaced with the generation value for the parameter considered, θ_g, and the measured value of the same parameter is θ_m. Thus the pull definition turns to be:

$$g = \frac{\theta_m - \theta_g}{\sigma_m} \tag{5.1.2}$$

where σ_m is the statistical uncertainty associated by the fit to the measured quantity. The same definition is used for each parameter that enters the likelihood. The idea for the pull studies is to generate a large number of samples (pseudo experiments) of simulated events whose variables have been chosen randomly polling the PDF in each variable subspace.

Those PDFs, as it has been explained in the previous chapter, are parametrized as functions of the parameters $\vec{\theta}$. About 1000 pseudo experiments have been generated with the same statistics of data ($\sim$ 55000 events). By looking at the distributions of the parameters of all the pseudo experiments, one wants to understand:

- if the mean value for the distribution of a parameter is equal or different from the generation value.
- if the shape of the distribution is Gaussian or presents strange behaviors. For instance it can be useful to understand if the shape of distribution is symmetric or not with respect to the mean value
- if the width of the distribution is unitary or not. This should happen because the pull are by definition the residuals divided by the standard deviation, so the distribution should be normalized to the Normal Standard distribution.

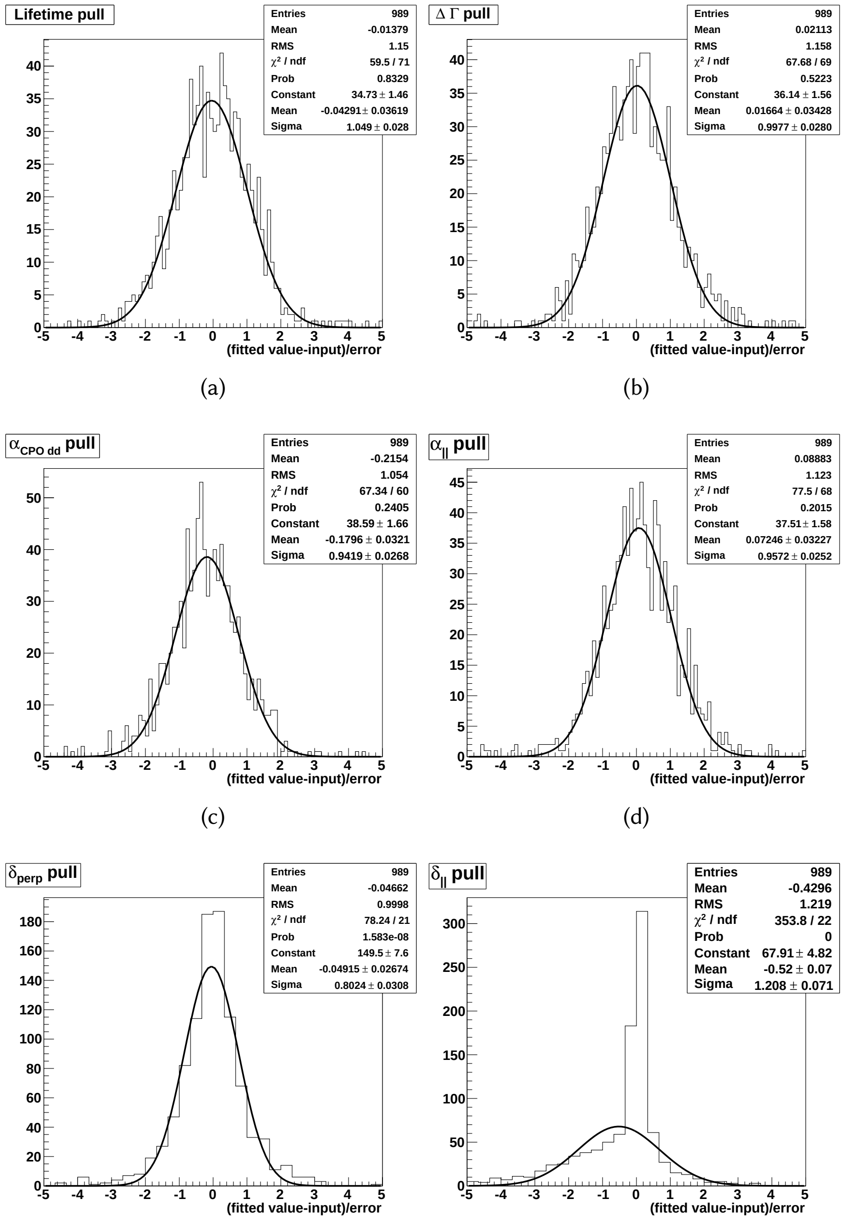
The first point of the itemize above provides the information on whether the fit has biases on the parameter estimation. Having a bias means that the value of a parameter is systematically measured larger or smaller than the generation value. This implies that in the real measurement the result of the fit for the parameter considered is systematically measured larger or smaller then the "true" value.

It is important to check whether biases are present or not in order to decide if the values for the parameters determined by the fit are reliable or not. In particular in case of biases it is important to quantify the amount of bias present for each variable. This allows to associate a proper systematic uncertainty to the value measured for the parameter.

The shape of the pull distribution gives also informations about the behavior of the fit with respect to the parameter considered. Usually if the pull distribution does not present a Normal distribution, or distribution that is close to be a Normal, the likelihood used is not considered a good estimator for that parameter. For this reason the choice to not quote the measured values for the parameters with a pathological pull distribution, is usually preferred.

Let us firstly examine the pull distribution obtained for the the parameters that are meant to be measured: $\Delta\Gamma$, $c\tau$, $\alpha_{\parallel}$, α_{CPOdd}, and the strong phases $\delta_{\perp}$ and $\delta_{\parallel}$ (see Fig. 5.1). The plots show a good behavior for all the quantities, a part from $\delta_{\parallel}$. The pulls for $\delta_{\parallel}$ show non-Gaussian behavior, for this reason is not possible to quote a value for that parameter with unbiased errors. The reason why the pull distribution for $\delta_{\parallel}$ is so badly behaving, probably relies in the likelihood symmetries. For $\delta_{\parallel}$ there is a reflected symmetry about π, and for values close to π the fit cannot always clearly determine the value between the two cases. In this situation, the fit has the tendency to return the value at the boundary of π as the fitted value for $\delta_{\parallel}$, since it in precisely in the middle of the two possible solutions.

It is useful to look also at the pull distribution for the S-wave amplitude A_{SW} and its phase with respect to the P-wave δ_s (see Fig. 5.2), focusing in particular on the S-wave amplitude. The S-wave amplitude parameter A_{SW} is left free to float in the full fit, but within the boundary at [0,1] that prevents A_{SW} assuming negative values and allows the overall normalization. It is not possible to take off such a boundary, it is evident in the likelihood parametrization that we are using (see sec. 4.7), but it is true also for any other parametrization.The point is that in the including S-wave likelihood (in eq. 4.7.3), the square root of the S-wave amplitude (or, fraction, the meaning is the same, A_{SW} is F_s in eq. 4.7.3) enters. For this reason a positive value for A_{SW} is required. The value obtained for the A_{SW} by the full fit to data is small, around $2\% \pm 2\%$, so very close to the boundary. This may cause problems with the convergence of the fit, as usually happens when a parameter is at limit. It has been noticed that when β_s is left floating, sometimes the fit shows convergence problems, because the minimizer(`MINUIT`) has been stuck at the limit for A_{SW} (see Ref. [121]). In the case of β_s fixed to zero, this problem is also present, the strategy to avoid it is to restart the minimization from the local minimum of the likelihood, moving the starting point of A_{SW}. This strategy has been applied for the fit on data, but not for the pull studies. This explains the strange behavior of the pull distribution for A_{SW} (Fig. 5.3c). As can be seen from Fig. 5.3a slightly more than half of the times the fit finds a value for A_{SW} that is either zero or in the interval between 0.0 and 0.02. The residual plot in Fig. 5.3b shows how the boundary at $A_{SW} = 0$ is responasble of the values found for A_{SW} and therefore the pull distribution behavior (Fig. 5.3c).
To support the hypothesis that the boundary was causing the bad behavior of the A_{SW} pull distribution, another set of pseudo experiments has been generated, with the generation value of $A_{SW} = 25\%$ far away from the boundary. We expect a gaussian distribution centered on 0.25 for the values of A_{SW} fitted on the pseudo experiment, and a Normal distribution for the pulls. Fig 5.4 shows the distribution obtained that in this case match quite well with those expeted.



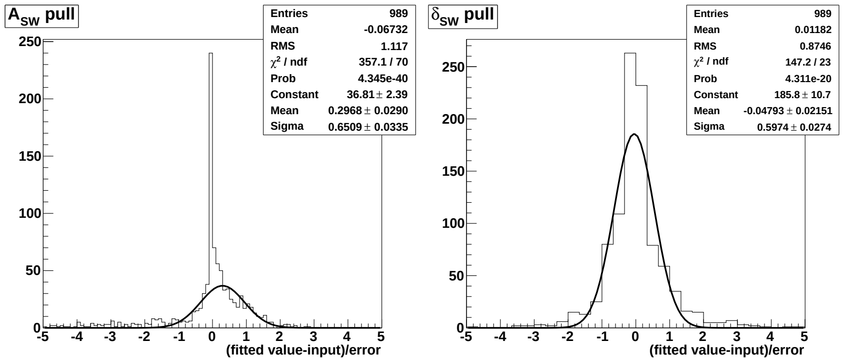
(a) (b)

(c) (d)

(e) (f)

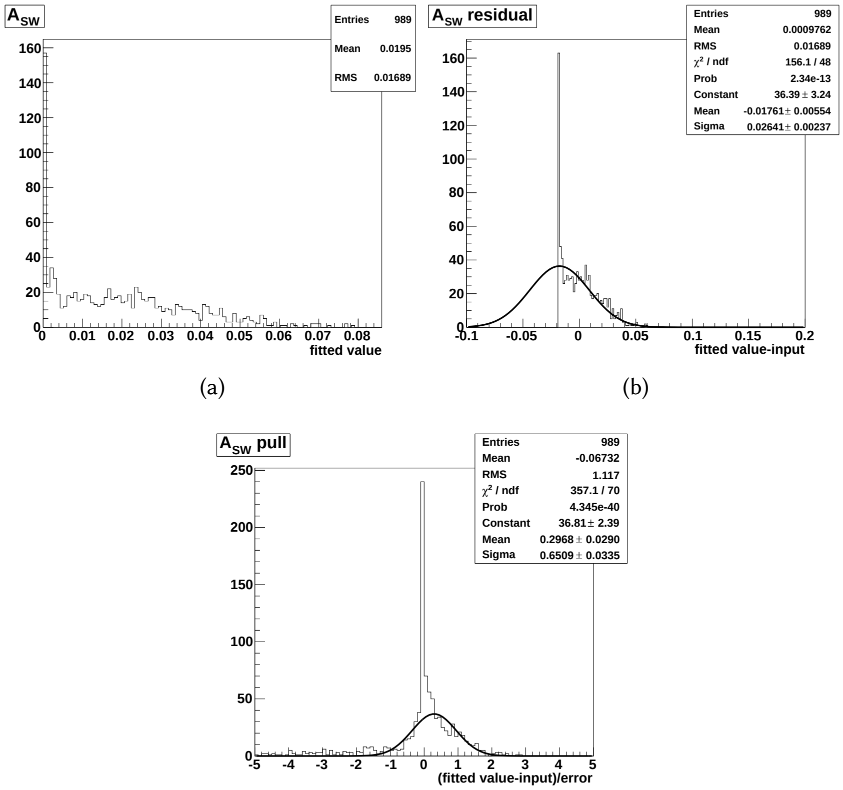
Figure 5.1.: Pull distributions for the physics parameters that are meant to be measured.

Another check that can be performed when having the pull distribution is to verify whether the mean error associated by the fit to a given parameter in the toy (or pseudo-experiment) studies, is comparable with the error associated by the fit to data. This comparison has been made for the physics parameters we are interested in, and is re-



(a) (b)

Figure 5.2.: Pull distributions for the S-wave amplitude A_{SW} and it phase relative to the P-wave δ_s.

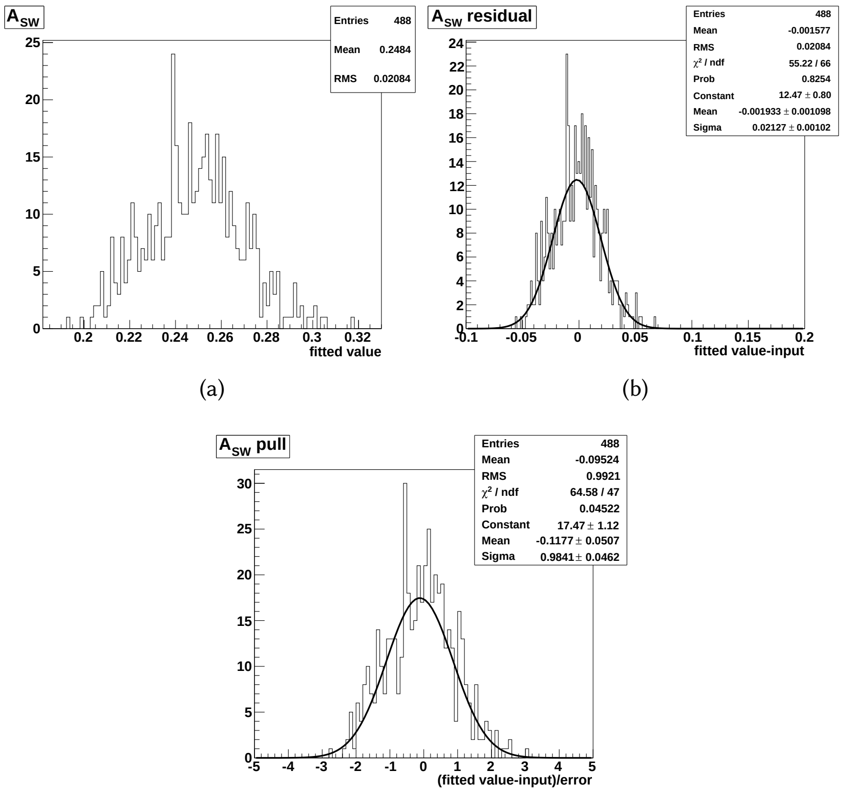


(a) (b)

(c)

Figure 5.3.: Value, residual and pull distributions for the S-wave amplitude A_{SW}, pseudo experiments generated with A_{SW} value from the CP conserving fit on data, 1.8 %

ported inTab. 5.1, together with the parameters describing the pull distributions (mean



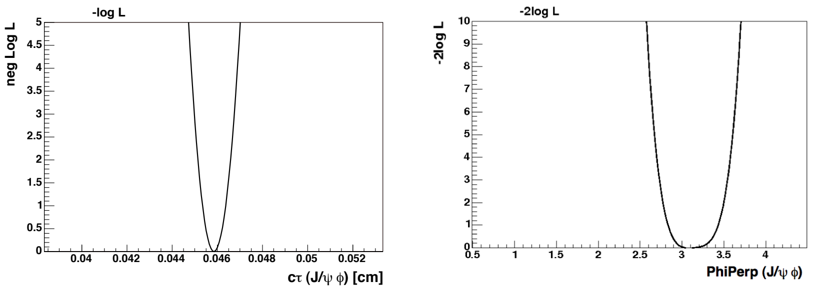
(a) (b)

(c)

Figure 5.4.: Value, residual and pull distributions for the S-wave amplitude A_{SW}, pseudo experiments generated with A_{SW} value of 25 %

and width). For the physical quantities that we want to measure in this thesis, the pull distributions features show that the values found by the fit, can be relied, and the comparison between the statistical error on the data with the average error on the pseudo experiment shows satisfactory agreement, with the exception of $\delta_\perp$. In this case the mean error is much larger than the error assigned by the fit to data, and if can be justified by looking at the likelihood profile in the neighborhood of the global minimum for the fit (see Fig. 5.5). Considering a large number of pseudo experiments, it can happen that the fit to some pseudo experiments converges in correspondence to a minimum that is not the generation value, due to the proximity of the two minima. This enlarges the mean error.

Usually the attention is focused on the quantities that are supposed to be measured, even if a general look at all the pull distribution represent a common test in order to understand if the fit is behaving properly. Pull distributions for all the variables of the likelihood function (summarized by Tab. 4.1) are reported in Appendix C, here Fig. 5.6 has the purpose to summarize mean and width of the pull distributions for all the parameters appearing in the full likelihood function.



(a) (b)

Figure 5.5.: Likelihood profile for $c\tau$ and $\delta_\perp$.This is a fit of a single dataset, it is repeated for each pseudo experiment generated. In 5.5b the two likelihood minima fall closer as it can happen that for some pseudo experiments. The consequence is that the fit can converge either in a minimum or in the other, and this spreads the mean error on the parameter $\delta_\perp$. In the other parameter that we are measuring the behavior of the likelihood profile is similar to that showed in 5.5a for $c\tau$.

Parameter	Pull mean	Pull σ	Mean Error	Error
$c\tau$	- 0.043±0.036	1.049±0.028	0.00061	0.00062
$\Delta\Gamma$	0.016±0.034	0.998±0.028	0.028	0.029
α_{CPOdd}	-0.180±0.032	0.942±0.028	0.012	0.012
$\alpha_{\|\|}$	0.072±0.032	0.957±0.027	0.013	0.012
$\phi_\perp$	-0.049±0.027	0.802±0.030	1.354	0.612

Table 5.1.: Pull values (mean and σ), mean error on the variable considered, in last column the value of the error on the fitted parameters.

5.1.2. Fit sensibility with respect to the input values and distributions

A good behaving fit is supposed to be not very sensible to small changes in the input parameters and distribution. In Chapter 4, all the joint PDF components have been presented and it has been pointed out that the parameters describing some of them, need to be fed in the fitter as input. Those input have been updated in order to perform correctly the fit with the new data sample. The parameters used to quantify the detector sculpting effects (see sec. 4.4.1), the parameters describing the σ_{ct} PDF (see sec. 4.4.2) , the parameters describing the background angular PDF (see sec. 4.4.5) and the tagging dilution scale factors are here indicated as "input parameters", while the term "input distributions" we refer to the tagging dilution histograms, that are used in the fit as PDF for the tagging (see sec. 4.5). In order to probe the sensibility of the fit to small variations of the inputs,

two different fit have been performed with the only difference relying in the input files (histogram describing the detector acceptance, the parameters entering in the σ_{ct} PDF, the parameters describing the background angular PDF and the parameters and dilution distributions describing the tagging performances.). In the first case the inputs are the

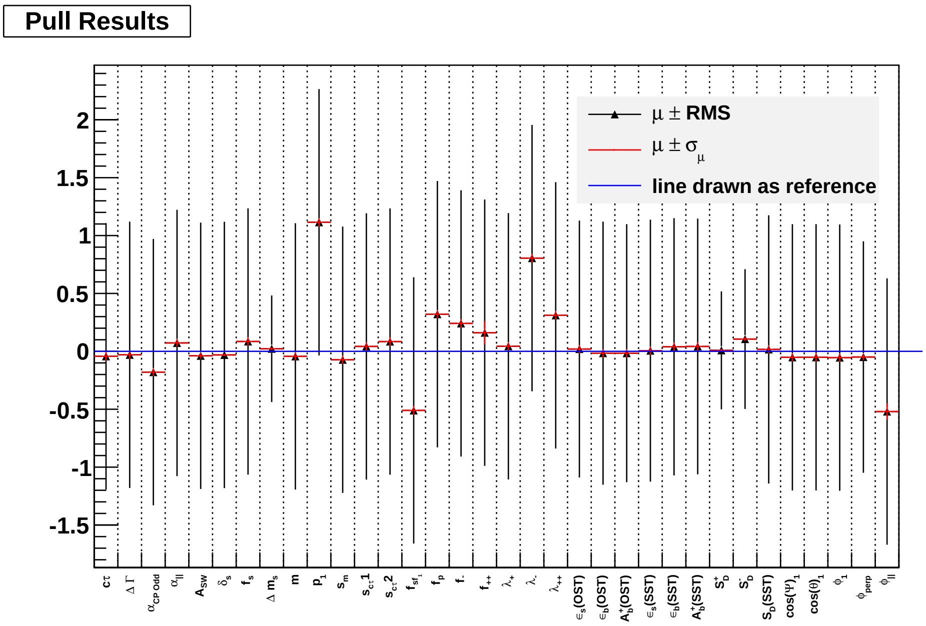


Figure 5.6.: Summary of the pull distributions for all the variables of the likelihood function.

updated inputs used for the full final fit, in the other case the inputs calculated for the previous iteration of the analysis have been used. Tab. 5.2 and Fig. 5.7 summarize the results obtained. This test could also be interpreted as an indication about the systematic error that needs to be associated to the values obtained for the parameters. As shows Fig. 5.7, if a systematic contribution can be associated to the variation of all these quantities together, for most of the parameters it turns to be less then 0.2σ in these case. It should kept in mind that this is only an indication, since the correct method to estimate the systematic uncertainties consists first of all in isolating the effect of the different effects that can contribute to the overall uncertainty. The correct method used to calculate the systematic uncertainties to be quoted together with the result will be explained in 5.3.

Parameter	NOT updated input	updated input
$c\tau$	0.04604 ± 0.00063	0.04580 ± 0.00062
$\Delta\Gamma$	0.062± 0.028	0.063 ± 0.029
$\alpha_\perp$	0.272± 0.011	0.272 ± 0.012
$\alpha_\parallel$	0.308 ± 0.012	0.311 ± 0.012
$\phi_\perp$	2.877 ± 0.522	2.949 ± 0.612

Table 5.2.: Fit parameters results in case of input updated or not.

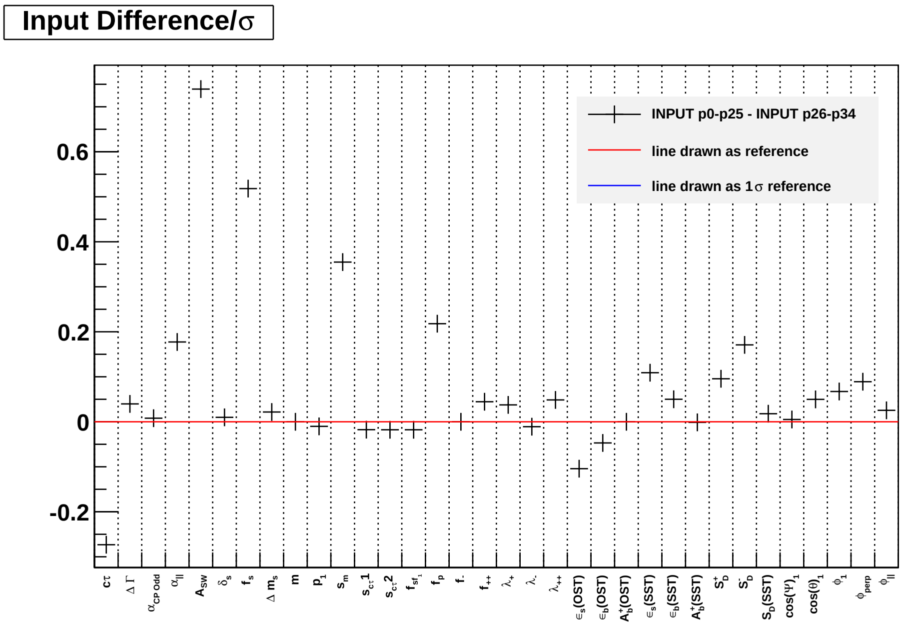


Figure 5.7.: Comparison between fit results obtained with fit inputs updated or keeping the inputs of the last iteration of this analysis.

5.1.3. Fit projections

Since the un-binned maximum likelihood method does not readily provide a goodness of fit estimator, fit projections onto the data can be used to support the quality of the fit. The PDF where all parameters are fixed to their best-fit values, is overlaid on top of data distributions. Such projections are performed for both signal and background events and separately in the subspaces of the B_s decay time, decay time uncertainty and transversity angles.

This correspond to project on a single variable the multidimensional fit that has been performed, and gives an information about the agreement between the fit projected on a single variable and the data. It doesn't give informations about the goodness of the full fit, but if a projection shows a particular disagreement, then maybe one or more of the PDFs used to model the events is written using a bad parametrization.

As first example in Fig. 5.8 the projections on the angular variables ($\cos\psi$, $\cos\theta$ and ϕ) for all data, signal events, sideband events, and sideband-subtracted data, are reported. In order to make this kind of plot, a sample of simulated events with statistics much higher than the data sample has to be generated. This simulated sample has to be generated according to the PDFs distribution described with the parameter values obtained by the fit. Histograms filled with the simulates data, normalized to the data histograms, will constitute the likelihood projection on the variable considered. A plot of real data histogram on the variable considered superposed with simulated data histogram normalized to data, will suggest the agreement between data and likelihood projection. In order to quantify the goodness of such agreement, a bin-per-bin χ^2 test can be used. In the case of the angular projections in Fig. 5.8, a χ^2 fit has been performed on the sideband-subtracted plots, and values of $\chi^2/ndf = 3.5/10$, $\chi^2/ndf = 9.8/10$, $\chi^2/ndf = 13.8/21$ have been found for the projection in $\cos\psi$, $\cos\theta$ and ϕ respectively.

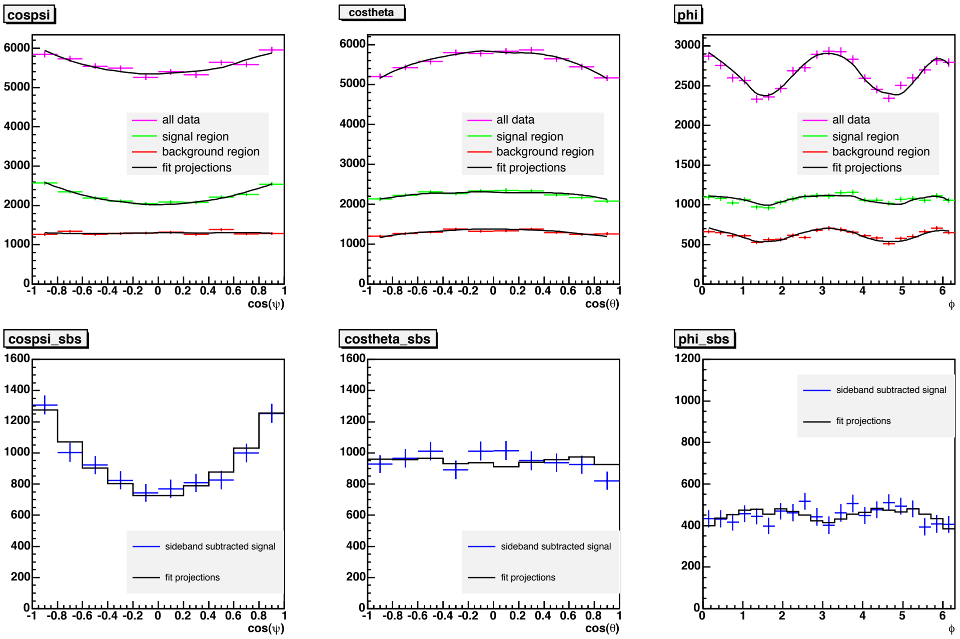
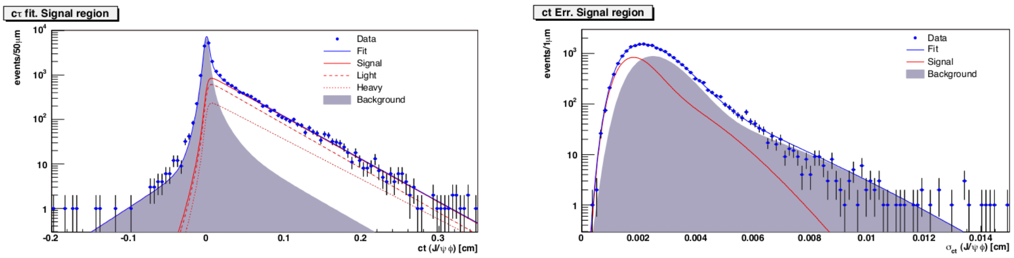


Figure 5.8.: Angular projection for the full fit. A χ^2 fit has been performed on the three lower plots, in order to quantify the agreement between data and interpolating function. Values of $\chi^2/ndf = 3.5/10$, $\chi^2/ndf = 9.8/10$, $\chi^2/ndf = 13.8/21$ have been found for the projection in $\cos\psi$, $\cos\theta$ and ϕ respectively.

One more meaningful choice of projections is represented by the ct and σ_{ct}, for both signal (Fig. 5.9) and background (5.10) events. These projections show how the ct and the σ_{ct} PDFs used are good in describing the data. Out of these plots, Fig. 5.9a for negative values of ct shows the experimental data points to be always above the likelihood projection. This can suggest to change the ct PDF, for instance a model with three instead of two gaussians can be used. Another possible strategy to reach an improvement could be to use a cut con ct on the data. This latter solution opens the problem that a resolution function (to be convoluted with the other components of the joint PDF) needs to be determined, so it is necessary to find a channel that can be used in order to do that. [1].

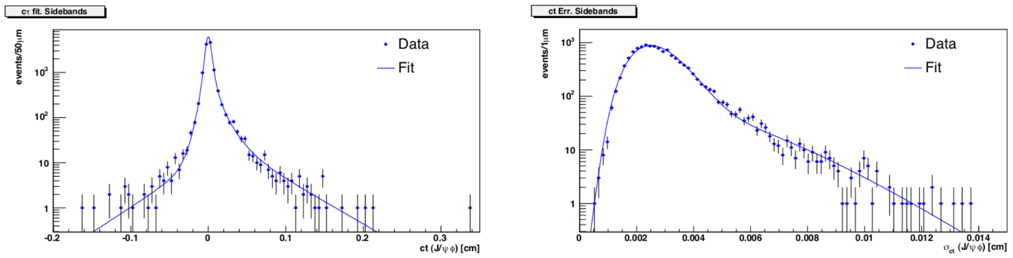
---

[1] also the J/ψ $prompt$ can be used (J/ψ that are not the decay products of the B_s, but are produced directly in the hard process of the $p\bar{p}$ collision), even if they do not represent the best solution, since it would always be a contamination coming from the J/ψ produced through B mesons decays.



(a) ct distribution for signal events (b) σ_{ct} distribution for signal events

Figure 5.9.: ct and σ_{ct} projections for signal events.



(a) ct distribution for background events (b) $\sigma_c t$ distribution for background events

Figure 5.10.: ct and σ_{ct} projections for the background events.

5.2. Tagged Fit results

The full multidimensional maximum likelihood fit on the dataset that was available when the analysis of this thesis has been realized ($\sim 8.3\ \mathrm{fb}^{-1}$, see Chapter 3 for the details of run, trigger and selection criteria used), gives the following results for the physics parameters of interest:

$$\begin{aligned} c\tau &= 458.0 \pm 6.2(\text{stat.})\mu\text{m} \\ \Delta\Gamma &= 0.063 \pm 0.029(\text{stat.})\text{ps}^{-1} \\ |A_{\parallel}(0)|^2 &= 0.233 \pm 0.014(\text{stat.}) \\ |A_0(0)|^2 &= 0.514 \pm 0.012(\text{stat.}) \\ \phi_{\perp} &= 2.95 \pm 0.61(\text{stat.}) \end{aligned} \tag{5.2.1}$$

The results in Eq. 5.2.1 show good agreement with previous measurements [97], [117], [118]. The B_s mean lifetime can be calculated as $\tau(B_s) = 1.527 \pm 0.021(\text{stat})$ps It can be compared to the most recent measurement from the DØ collaboration using a data sample based on $8\ \mathrm{fb}^{-1}$ of integrated luminosity [119] quoting $\tau(B_s) = 1.443^{+0.038}_{-0.035} ps$ and to the Particle Data Group (PDG) average of $\tau(B_s) = 1.472^{+0.024}_{-0.026} ps$??. The $\Delta\Gamma$ value is of comparable precision to the current world (PDG) average of $\Delta\Gamma = 0.062^{+0.034}_{-0.037}$[27]. Our central value is somewhat smaller than the most recent measurement of $\Delta\Gamma = 0.163 \pm 0.065$ from the D0 collaboration [119] and from the LHCb collaboration $\Delta\Gamma = 0.123 \pm 0.029(\text{stat}) \pm 0.011(\text{syst})$ but compares well to the PDG average as well as the SM prediction $\Delta\Gamma = 0.090 \pm 0.024$ [120].

5.2.1. Comparison with untagged results

As explained in previous chapter, also a measurement that does not exploit the tagging technique should in principle be possible. This is true, as it has been shown when the purpose is to measure β_s, but also whether one is interested in the measurement of $\Delta\Gamma$, $c\tau$, $|A_0|^2$ and $|A_\||^2$. Two are the reasons why we have chosen to use the tagged fit instead of the untagged one, that have their root in the two motivations that have led to this thesis.

- *The fit with $\beta_s = 0$ is necessary to the complete understanding of the fit with β_s floating.*
  For this reason it is useful to use the same likelihood, since some problems arising in the simpler case ($\beta_s = 0$) can be useful to avoid problems in the more complex case (β_s floating). For the β_s floating fit, there are three reasons to prefer the tagged fit with respect to the untagged one.
  The first relies in the fact that the use of the tagged fit removes two out of the four possible solutions (absolute minima of the likelihood) that can be found due to the likelihood symmetries (see sec. 4.6.1). The second reason consist in the fact that the measurement exploiting the tagging is expected to be more precise.
  Last motivation is that from a set of cross checks, the untagged likelihood seemed to behave less reliably than the tagged one, at least in the neibhorhood of $\beta_s = 0$, when β_s is floating in the fit. For the untagged likelihood , the complication arises from the fact that the strong phase $\delta_\perp$ appears always as a product with $\sin 2\beta_s$. As a result, in case of no *CP*-violation, there is no sentivity on $\delta_\perp$, but if the sensitivity to *CP*-violation is small, the fit tends to bias the result as by increasing CP-violation, the fit gains δ_{perp} as an additional parameter available to describe statistical fluctuation. It has been noticed that the bias is non-linear and decreases with increasing CP-violation. Since in the *CP*-conserving untagged fit the strong phase $\delta_\perp$ does not enter the likelihood, the fit with β_s fixed to zero does not suffer of the above problem and it is believed to be reliable, and the results that it gives have been compared with the tagged fit results, as will be explained later in this section.

- *Measurements of the parameters such as $\Delta\Gamma$, $c\tau$ and polarization amplitudes are interesting on their own right*
  The tagged fit, unlike the untagged one, gives the possibility to measure also the strong phase $\delta_\perp$ (see likelihood expressions for the tagged and the untagged case in Chapter 4: in the untagged case, only $\delta_\|$ enters the likelihood). With the likelihood that we are using, only the value obtained for $\delta_\perp$ can be quoted, because of the

not-Gaussian behavior of the pulls for $\delta_{\|}$. As a consequence the choice of using the tagged fit has been made also to measure and additional physics parameter, that is $\delta_{\perp}$, which could not be determined with the untagged fit.

The results for the case of *CP*-conservation obtained with the untagged fit for the parameters of interest are reported in Tab. 5.3, where are compared with the results of the tagged fit. The plot in Fig. 5.11 summarizes the comparison of the results with the two strategies for all the parameters involved in the untagged likelihood fit. The fact that with a different likelihood, that has been already tested, results compatible with the results reached with the tagged fit are obtained, supports the confidence in the tagged fit method used in this analysis, discussed and tested so far. In this specific case, the fit using the likelihood without flavor tagging has the purpose to check for any bias which could be introduced by the tagging.

Parameter	Tagged Fit	Untagged Fit
$c\tau$	0.04580 ± 0.00062	0.04556 ± 0.00063
$\Delta\Gamma$	0.063 ± 0.029	0.058 ± 0.029
$\alpha_{\perp}$	0.272 ± 0.012	0.276 ± 0.011
$\alpha_{\|}$	0.311 ± 0.012	0.312± 0.012

Table 5.3.: Comparison between tagged and un-tagged fit results for the parameters that are supposed to be measured.

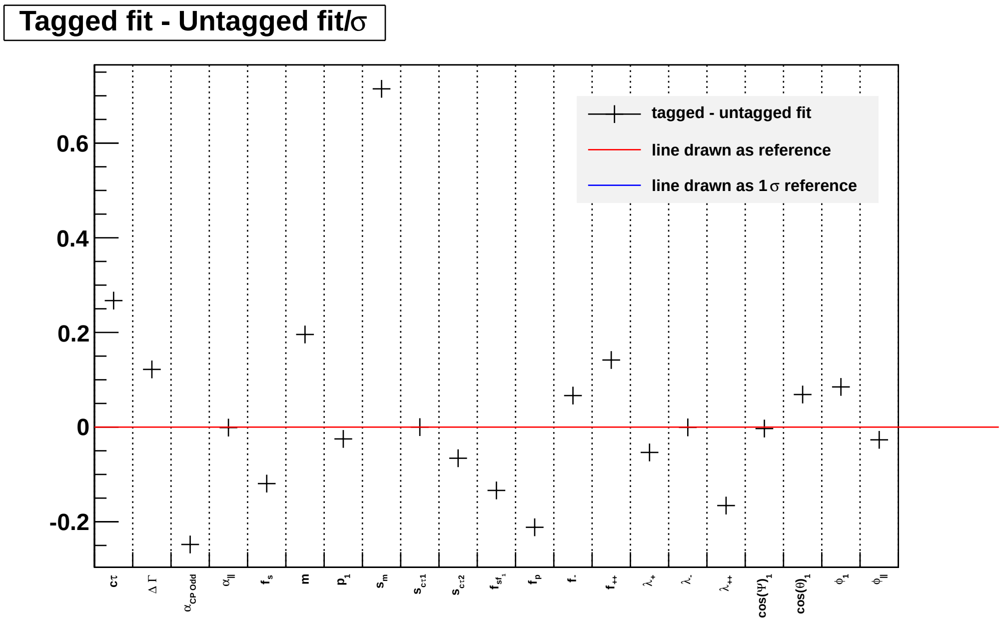


Figure 5.11.: Comparison between fit results in case of tagged fit and untagged fit

5.2.2. Consistency with previous analysis results

Another consistency test that can be interesting to implement, is to perform the fit on different subsections of the whole dataset. The results for the parameters in this case are expected to be compatible with each other. This test is usually done by changing the manner used to choose the subsample (temporal order of acquisition, random subsamples, can be used as criteria) different times, in order to show the data-independence of the likelihood fit used. In the present case, since the analysis presented had been already completed twice before (with respectively 1.3 fb^{-1} and 5.2 fb^{-1}) within this same experiment, we wanted to compare the results obtained with the last portion of data acquired (periods from 26 to 34 included, for an amount of 3.1 fb^{-1} of data approximately), with the results quoted from the group that took care of the previous iteration (5.2 fb^{-1}). These results are summarized in Tab. 5.4 and in Fig. 5.12.

Parameter	p0-p25	p26-p34
$c\tau$	0.0459 $\pm$ 0.00075	0.04564 $\pm$ 0.00114
$\Delta\Gamma$	0.075 $\pm$ 0.03	0.012 $\pm$ 0.053
$\alpha_\perp$	0.266 $\pm$ 0.014	0.280 $\pm$ 0.021
$\alpha_\parallel$	0.306 $\pm$ 0.015	0.315 $\pm$ 0.022
$\phi_\perp$	2.95 $\pm$ 0.637	2.885$\pm$ 1.212

Table 5.4.: Different sub-datasets: p0-p25 the dataset used for the previous iteration of this analysis and p26-p34 the sample used for this thesis.

5.3. Systematics

Until now only *statistical* uncertainties have been considered,that mainly depend on the data sample that has been considered. This kind of uncertainty cannot account for effects due to the modeling of the likelihood or to the detector itself, that can affect the measurement results. Uncertainties that account for these latter effects are called *systematics*. In order to understand what are the principal sources of systematical uncertainties, the documentation of the previous iteration of this analysis have been used.

Effects that are not accounted for in the likelihood fit, but are considered with the systematics include potential mis-parameterization in the fit model, impact of particular assumptions in the fit model, and physical effects which are not well known or fully incorporated into the model. To estimate the size of the systematic uncertainties, two sets of pseudo-experiments should generated. One set with each of the considered systematic variations and another set of default pseudo-experiments. Each pair of modified and not modified pseudo-experiments are generated with the same random seeds. The unbinned likelihood function used for the final full fit has to be maximized over the modified pseudo-experiments as well as over the corresponding default ones. For each systematic effect, the associated uncertainty is the difference between the mean of the best fit value

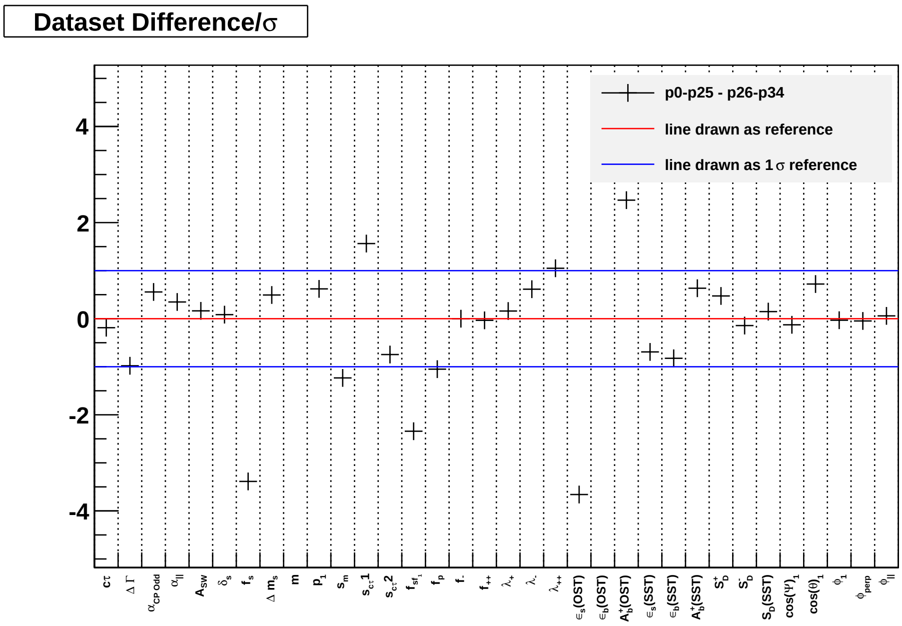


Figure 5.12.: Comparison between fit results in case of two independent data sample: p0-p25 and p26-p34, the same subsample presented in Tab. ??.

for the pseudo-experiments with the systematic alteration included, and the equivalent mean value for the reference set of pseudo-experiments generated with the default model. If the difference between the fitted value and the generation value divided by the error ($(\theta_m - \theta_g)/\sigma_m$) is indicated with the term *shift*, the single systematic error (indicated with *syst*) is calculated as follows:

$$\mathrm{syst} = \mathrm{shift}_{\mathrm{syst}} - \mathrm{shift}_{\mathrm{reference}} = ((\theta_{\mathrm{m}} - \theta_{\mathrm{g}})/\sigma_{\mathrm{m}}) - \mathrm{shift}_{\mathrm{reference}} \tag{5.3.1}$$

The individual systematic uncertainties are summed in quadrature to give the overall systematic uncertainty error. Let us now summarize the individual source of systematic uncertainty that must be kept into account:

1. **Signal Angular Efficiency**
   One source of systematic uncertainty is the modeling of the angular efficiency of the detector described in sec. 4.4.1 . The detector efficiency has been modeled with a linear combination of Legendre polynomials and spherical harmonics as has been described. The expansion coefficients of these functions are obtained by fitting a three-dimensional efficiency distribution obtained using simulated events. This simulated sample is re-weighted to match the p_T distributions observed in data, as explained in Chapter 3. If the modeling is inaccurate, or the pT re-weighting incorrect, a systematic uncertainty could be introduced. These effects can be investigated separately. The former effect is investigated by using the default fit model on pseudo-experiments, generated with angular efficiencies of the three-

dimensional histogram obtained using simulate events rather than the default parameterization. The latter effect is investigated by generating pseudo-experiments with non-reweighted MC events as input for the angular efficiencies. The second test is a rather extreme case, since it is very unlikely that the data behave like the non-reweighted MC, but in the previous analysis iteration, has shown only a small systematic effect. The two systematic shifts have to be summed in quadrature.

2. **Signal mass model**
   The B_s mass distribution is fitted by default with a single Gaussian distribution. If, for example multiple Gaussians would model the true distribution better, this could introduce a systematic uncertainty. To test the size of a potential systematic effect, pseudo-experiments with a double Gaussian signal mass model, extracted from data can be generated and fit with the usual single Gaussian parameterization.

3. **Background mass model**
   Similarly, the model used for the mass distribution of combinatorial background events could contribute a systematic uncertainty if it is insufficient to describe the data. To study this, pseudo-experiments with a second order polynomial background model instead of the default first order polynomial can be generated and fit with the default straight line.

4. **Lifetime resolution model**
   A particularly important effect to consider for the lifetime measurement is the lifetime resolution model. In order to account for the detector resolution, each lifetime component of the likelihood is convoluted with a two-Gaussian resolution function. To test the effect of a mis-parameterization of the resolution function, pseudo-experiments with a three-Gaussian resolution model, extracted from data, can be generated and fit with the default two Gaussian model.

5. **Background lifetime fit model**
   As well as the lifetime resolution, the modeling of the various components of the background lifetime can systematically affect the B_s lifetime measurement. To check this effect pseudo-experiments with the decay time of the background events taken from histograms of the B_s mass sidebands can be generated and fit with the default model.

6. **Angular background model and correlations**
   In the previous analysis iteration, three possible sources of systematic uncertainty related to the transversity angles of the background events have been considered: mis-modeling of the parameterization described in sec. 4.4.5, ignoring the observed small correlations between the three angles, and correlations between the angles and the expected proper decay time uncertainty, σ_{ct}. The effect of these sources of uncertainty have been checked using the actual data distributions from the mass sidebands to generate pseudo-experiments and test the difference between our model and the true distributions.

For the parameterization check, one could just use the data background angular distributions in the generation of the pseudo-experiments before fitting with the default model.
To check the effect of neglecting the small correlations between the angles, pseudo-experiments where two of the background angles are sampled randomly from the data distributions, and the third one from a two-dimensional histogram according to the sampled value of the second angle can be generated.
To check the effect of ignoring correlations between the transversity angles and σ_{ct}, the ϕ angle distribution, found to have the largest correlation with σ_{ct}, can be sampled using a two-dimensional histogram of ϕ versus σ_{ct} in order to generate the pseudo-experiments.
In principle the three systematic shifts presented have to be summed in quadrature, but the effect of ignoring these very small correlations has been found to result in an almost negligible systematic uncertainty on the measurements by the previous analysis.

7. B_d **crossfed**
In the default fit, we do not account for contamination from $B_d \rightarrow J/\psi K^*$ events mis- reconstructed as $B_s \rightarrow J/\psi\phi$ decays (B_d cross-feed). A small fraction of these events lies in the B_s mass signal region. The first step in identifying the size of the systematic effect is to estimate the size of this contribution by using measured production fractions of the B_s and B_d mesons, their relative decay rates to $J/\psi\phi$ and $J/\psi K^*$, respectively, and the efficiency of each type of event passing the final selection criteria established under the $B_s \rightarrow J/\psi\phi$ hypothesis. Both the production fractions and the branching fractions are taken from Ref. [27]. The efficiencies can be estimated using simulation, with both $B_s \rightarrow J/\psi\phi$ and $B_d \rightarrow J/\psi K^*$ modes reconstructed as $B_s \rightarrow J/\psi\phi$ decay. The fraction f of B^d cross-feed events in the B_s sample have been calculated to be (1.6 ± 0.6) into the signal sample of the previous analysis iteration. In that analysis a conservative estimate has been chosen by generating pseudo-experiments with a fraction of 2.2% B_d cross-feed, and fit with the default model which does not account for this component. The cross-feed component have been generated using values of the B_d lifetime, decay width and transversity amplitudes from the CDF angular analysis od $B_d \rightarrow J/\psi K^*$ decays [124].

8. **SVX alignment**
A systematic uncertainty can be introduced by the assumption that the silicon detector is perfectly aligned, when it could actually be mis-aligned by bowing of the detector layers of up to $50\,\mu m$. A study on the effect of the limited knowledge of the CDF silicon detector alignment concluded that a conservative estimation of the systematic uncertainty on the decay length $c\tau$ in CDF lifetime measurements is given by a $2\,\mu m$ systematic uncertainty on $c\tau$ [122] [123]. This study was done by fully reconstructing both data and simulation under different silicon alignment assumptions, including shifts of $\pm 50\,\mu m$in all silicon detector components. The

lifetime was fitted in several $B \to J/\psi X$ channels, and the worst shift was taken as the systematic uncertainty on the lifetime due to the assumption of perfect silicon alignment.
The value of $2\,\mu m$ systematic uncertainty on $c\tau(B_s)$ can also be used to assess secondary effects on the other parameters of interest. Due to correlations between the B_s lifetime and the other physics parameters, it is expected that an addiltional uncertainty on the lifetime measurement will also cause additional uncertainties in the measurement of the other parameters. To quantify this contribution to the uncertainties of the other parameters, pseudo-experiments in which the decay time in each event is randomly shifted $\pm 2\,\mu m$ can be generated and fit with the default fit.

9. **Mass error distribution**
   In the fit likelihood, we consider the mass resolution to be the same for signal and background events. The effect of any inaccuracy in this assumption can be tested by generating pseudo- experiments with mass uncertainty distributions modeled by histograms of B_s sideband data for background events and sideband subtracted signal region data for signal events separately, and then fitted with the default model.

10. **ct error distribution**
   To account for a possible mis-parametrization of the ct distributions effect, pseudo-experiments generated with the uncertainty distributions taken from data histograms rather than from the model described in sec. 4.4.2 can be generated and fitted with the default fit. This systematic check also accounts for any effect caused by small observed correlations between σ_{ct} and the invariant mass by sampling the background uncertainties from separate upper and lower sideband histograms according to the generated B_s mass.

From the previous analysis, the most relevant contributions to the overall systematic uncertainties are given by the background mass model, the detector acceptance parametrization, the $B^0 \to J/\psi K^{0*}$ crossfeed , signal mass model and ct error distribution. Tab. 5.5 summarized the systematic uncertainty associated to each of the sources listed above, by the previous iteration of this analysis. For the measurement reported in this thesis, we expect roughly the same order of magnitude for the systematic uncertainties, but a more evident change is expected in the contribution due to those effect that change more according to the statistics of the data sample used. For instance the contribution due to the detector efficiency parametrization is expected to change more than the other contributions, since the reweight of the MC depends on the data sample used.

Systematic	$\Delta\Gamma$ [ps $^{-1}$]	$c\tau$ [μ m]	$\|A_{\|\|}(0)\|^2$	$\|A_0(0)\|^2$	$\delta_\perp$
Signal efficiency					
Parameterization	0.0024	0.96	0.0076	0.008	0.016
MC re-weighting	0.0008	0.94	0.0129	0.0129	0.022
Signal mass model	0.0013	0.26	0.0009	0.0011	0.009
Background mass model	0.0009	1.4	0.0004	0.0005	0.004
Resolution model	0.0004	0.69	0.0002	0.0003	0.022
Background lifetime model	0.0036	2.0	0.0007	0.0011	0.058
Background angular distribution					
Parameterization	0.0002	0.02	0.0001	0.0001	0.001
σ_{ct} correlation	0.0002	0.14	0.0007	0.0007	0.006
Non-factorization	0.0001	0.06	0.0004	0.0004	0.003
$B_d \to J/\psi K^*$ cross-feed	0.0014	0.24	0.0007	0.0010	0.006
SVX alignment	0.0006	2.0	0.0001	0.0001	0.020
Mass resolution	0.0001	0.58	0.0004	0.0004	0.002
σ_{ct} modeling	0.0012	0.17	0.0005	0.0007	0.013
Pull bias	0.0012		0.0013	0.0021	
Totals	0.006	3.6	0.015	0.015	0.07

Table 5.5.: Summary of systematic uncertainties assigned by the previous analysis iteration to the five physics quantities that are supposed to be measured .

Conclusions

In this thesis the update of the measurement of the B_s^0 lifetime, the decay width difference between its heavy and light mass eigenstates and the polarization amplitudes of $B_s^0 \to J/\psi\phi$ decays of the B_s^0 meson has been presented.

About 9600 $B_s^0 \to J/\psi\phi$ decays have been reconstructed in the final state $[\mu^+\mu^-][K^+K^-]$ using a dataset of $8.4\ \mathrm{fb}^{-1}$ integrated luminosity collected by the CDFII detector at the Fermilab Tevatron collider.

The results are obtained by determining the relative contribution of the decay amplitudes with definite CP–parity as a function of the decay time, through an unbinned maximum likelihood fit which exploits flavor tagging of the produced B_s^0 meson.
Tables 5.6 and 5.7 compare my results with the theoretical predictions and the world's best measurements, respectively.

	Our measurement	Theor. prediction
$\tau[\ \mathrm{ps}]$	$1.527 \pm 0.021(\mathrm{stat})$	[1.497-1.541] [126] [127]
$\Delta\Gamma\ [\ \mathrm{ps}^{-1}]$	$0.063 \pm 0.029(\mathrm{stat.})$	$0.090 \pm 0.024(SM)$ [120]
$\lvert A_0\rvert^2$	$0.514 \pm 0.012(\mathrm{stat.})$	0.531 ± 0.022 [21]
$\lvert A_\parallel\rvert^2$	$0.233 \pm 0.014(\ \mathrm{stat.})$	0.230 ± 0.028 [21]
$\delta_\perp$	$2.95 \pm 0.61(\ \mathrm{stat.})$	2.97 ± 0.18 [21]

Table 5.6.: Comparison between my results and theoretical predictions.

The lifetime value is about 2σ greater than the recent measurement by the DØ collaboration, but it compares well with the LHCb estimation. In addition, it confirms the prediction of the HQE models [127, 126].

The $\Delta\Gamma$ result has a comparable resolution and is in good agreement with the current world (PDG) average, $\Delta\Gamma = 0.062^{+0.034}_{-0.037}$[27]. The central value is somewhat smaller than the most recent measurements by the DØ and LHCb collaborations [119, 40]. It must be pointed out that both DØ and LHCb measurements are performed in the same analysis as the β_s measurement, while our result is obtained under the SM assumption for that CP-violating phase. It is interesting to note that my value is in agreement with the SM

	Our measurement	DØ [119]	LHCb [40]
$\tau(B_s)[\text{ ps}]$	1.527 ± 0.021	$1.443^{+0.038}_{-0.035}$	$1.524 \pm 0.020 \pm 0.018$
$\Delta\Gamma\ [\text{ ps}^{-1}]$	0.063 ± 0.029	$0.163^{+0.065}_{-0.063}$	$0.123 \pm 0.029 \pm 0.011$
$\lvert A_0\rvert^2$	0.514 ± 0.012	$0.558^{+0.017}_{-0.019}$	$0.497 \pm 0.013 \pm 0.031$
$\lvert A_\parallel\rvert^2$	0.233 ± 0.014	$0.231^{+0.024}_{-0.030}$	$0.265 \pm 0.028 \pm 0.042$
$\delta_\perp$	2.95 ± 0.61	3.15 ± 0.22	$2.94 \pm 0.37 \pm 0.12$

Table 5.7.: My results compared with the world's best measurements. For each measurement, the first uncertainity is statistical, the second one (when present) is systematic.

prediction $\Delta\Gamma = 0.090 \pm 0.024$ [120] whitin 1σ.

The polarization amplitudes are compared with the theoretical predictions based on SU(3) symmetry [21] and confirm the possibility to constrain phases of such amplitudes to corresponding measured values in the $B^0 \to J/\psi K^*$ decay, allowing for the resolution of the β_s sign ambiguity in the analysis for its estimation. Indeed, the measurement of $\delta_\perp$ is in reasonable agreement with $\delta_\perp$ from $B^0 \to J/\psi K^*$ decays, although still has a large uncertainty.

All results are in agreement with the previous CDF measurements [97, 117], as should be expected. This work represents an intermediate update of the latest CDF measurement done with 5.2 fb^{-1} of data before performing the measurement with the final CDF dataset. Aspects concerning this task span from the control of the optimized selection of the data sample; the validation of the new data along with the study of their trigger composition; the correction to apply to the simulated events to properly model the detector acceptance in different acquisition periods; the test of the multidimensional maximum likelihood fit used to get the measurement of the observables of interest.

As of this writing, the whole Run II dataset ($\sim$10 fb^{-1}) has become available. We are now working on the update of the results using the final CDF dataset. Since all the features of the analysis has been well validated in this work, such update is expected in a very short time. Therefore, the systematic uncertainities presented at the end of Chapter 5 will be carried out for the final CDF measurement. The overall systematic uncertainities are expected to be of the same order as the ones quoted in the previous iteration of this analysis [97]. We consider to obtain the final results and their official approval by the CDF collaboration by the end of this year.

Appendix A

Normalization of $B_s \to J/\psi\phi$ transversity PDF

The differential time and angular dependent decay rate for $B_{\to}J/\psi\phi$ decays that has been used for the likelihood fit has been presented in Chapter 4. Here the purpose is to explain in greater detail the analytic normalization of this component of the PDF, and the formalism effectively used in the fit implementation, here is used as well (see Sec. 4.6.2). Starting from Eq. 4.4.4 the first step is to integrate over time the decay rates:

$$
\begin{aligned}
&\int |f_+(t)|^2 + |\bar{f}_+(t)|^2 = 1,\\
&\int |f_-(t)|^2 + |\bar{f}_-(t)|^2 = 1,\\
&\int \sum_{i=B,\bar{B}} f_+ f_-^* dt = \frac{i}{\sqrt{1+\frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta_s)^2}}},
\end{aligned}
\tag{A.0.1}
$$

giving factors which can be substituted in Eq. 4.6.13, to rewrite the normalization N as:

$$
\begin{aligned}
N = \frac{9}{16\pi}\int\int\int d(cos\psi)d(cos\theta)d\phi\\
\left[|\boldsymbol{A}_+ \times \hat{n}|^2 + |\boldsymbol{A}_- \times \hat{n}|^2 + 2\mathcal{R}e(\boldsymbol{A}_+ \times \hat{n})(\boldsymbol{A}_-^\star \times \hat{n})\frac{i}{\sqrt{1+\frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta_s)^2}}}\right]\\
\epsilon(\psi,\theta,\phi),
\end{aligned}
\tag{A.0.2}
$$

At this point we need t carry out the angular integration. In order to do that we will use the relations:

$$\begin{aligned}
\sin\theta\cos\phi &= \sqrt{\frac{2\pi}{3}}(Y_1^{-1} - Y_1^1),\\
\sin\theta\sin\phi &= \sqrt{\frac{2\pi}{3}}i(Y_1^{-1} + Y_1^1),\\
\cos\theta &= \sqrt{\frac{4\pi}{3}}Y_1^0.
\end{aligned} \tag{A.0.3}$$

that enter the PDF as products, then N can be written as:

$$\begin{aligned}
N = \frac{9}{16\pi}\int\int\int \epsilon(\psi,\theta,\phi)d(\cos\psi)d(\cos\theta)d\phi\Big\{\frac{|A_\perp|^2}{2}\Big[\sqrt{\frac{16\pi}{9}}Y_0^0 - \sqrt{\frac{16\pi}{45}}Y_2^0\Big]\sin^2\psi\\
+\frac{|A_\parallel|^2}{2}\Big[\sqrt{\frac{4\pi}{9}}Y_0^0 + \sqrt{\frac{16\pi}{45}}Y_2^0\Big]\sin^2\psi + |A_0|^2\Big[\sqrt{\frac{4\pi}{9}}Y_0^0 + \sqrt{\frac{16\pi}{45}}Y_2^0\Big]\cos^2\psi\\
+\frac{|A_\parallel|^2}{2}\Big[\sqrt{\frac{4\pi}{9}}Y_0^0 + \sqrt{\frac{4\pi}{45}}Y_2^0 + \sqrt{\frac{2\pi}{15}}(Y_2^{-2} + Y_2^2)\Big]\sin^2\psi\\
+\frac{|A_\parallel|^2}{2}\Big[\sqrt{\frac{4\pi}{9}}Y_0^0 - \sqrt{\frac{4\pi}{45}}Y_2^0 - \sqrt{\frac{2\pi}{15}}(Y_2^{-2} + Y_2^2)\Big]\cos^2\psi\\
+i\frac{A_\parallel A_0^\star + A_\parallel^\star A_0}{\sqrt{2}}\Big[\sqrt{\frac{2\pi}{15}}(Y_2^{-2} - Y_2^2)\Big]\sin\psi\cos\psi + 2\mathcal{R}e\Big[\frac{i}{\sqrt{1+\frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta_s)^2}}}\\
\times\Big\{\frac{A_\perp A_0^\star}{2}\sqrt{\frac{2\pi}{15}}(Y_2^{-1} + Y_2^1)\sin^2\psi + i\frac{A_0 A_\perp^\star}{\sqrt{2}}\sqrt{\frac{2\pi}{15}}(Y_2^{-1} - Y_2^{-1})\sin\psi\cos\psi\Big\}\Big]\Big\}
\end{aligned} \tag{A.0.4}$$

Let us recall now the expression of the detector efficiency in terms of spherical harmonics and Legendre polynomials (see Eq. 4.4.5).

$$\epsilon(\theta,\phi,\psi) = \sum a_{lm}^k P_k(\theta,\phi,\psi)Y_{lm}(\theta,\phi) \tag{A.0.5}$$

The integration of such a relation gives:

$$\begin{aligned}
\int\int Y_l^m(\theta,\phi)\bar{Y}_{l'}^{m'}(\theta,\phi)\sin\theta d\theta d\phi = \delta_{mm'}\delta_{ll'},\\
Y_l^{-m}(\theta,\phi)) = (-1)^m\bar{Y}_{l'}^{m'}(\theta,\phi).
\end{aligned} \tag{A.0.6}$$

As can be seen, the detector efficiency enters N through the coefficients a_{lm}^k as will be shown in the following relations. Using the definitions above in A.0.4, the resulting expression for N is:

$$
\begin{aligned}
N = \frac{9}{16\pi}\int d(\cos\psi)\Big\{&\frac{|A_\perp|^2}{2}\Big[\sqrt{\frac{16\pi}{9}}a_0^0(\psi) - \sqrt{\frac{16\pi}{45}}a_2^0(\psi)\Big]\sin^2\psi \\
&+ \frac{|A_\parallel|^2}{2}\Big[\sqrt{\frac{4\pi}{9}}a_0^0(\psi) + \sqrt{\frac{16\pi}{45}}a_2^0(\psi)\Big]\sin^2\psi + |A_0|^2\Big[\sqrt{\frac{4\pi}{9}}a_0^0(\psi) + \sqrt{\frac{16\pi}{45}}a_2^0(\psi)\Big]\cos^2\psi \\
&+ \frac{|A_\parallel|^2}{2}\Big[\sqrt{\frac{4\pi}{9}}a_0^0(\psi) + \sqrt{\frac{4\pi}{45}}a_2^0(\psi) + \sqrt{\frac{2\pi}{15}}(a_2^{-2}(\psi) + a_2^2(\psi))\Big]\sin^2\psi \\
&+ \frac{|A_\parallel|^2}{2}\Big[\sqrt{\frac{4\pi}{9}}a_0^0(\psi) - \sqrt{\frac{4\pi}{45}}a_2^0(\psi) - \sqrt{\frac{2\pi}{15}}(a_2^{-2}(\psi) + a_2^2(\psi))\Big]\cos^2\psi \\
&+ i\frac{A_\parallel A_0^\star + A_\parallel^\star A_0}{\sqrt{2}}\Big[\sqrt{\frac{2\pi}{15}}(a_2^{-2}(\psi) - a_2^2(\psi))\Big]\sin\psi\cos\psi + 2\mathcal{R}e\Big[\frac{i}{\sqrt{1 + \frac{4\tau_L\tau_H}{((\tau_L - \tau_H)\sin 2\beta_s)^2}}} \\
&\times\Big\{\frac{A_\perp A_0^\star}{2}\sqrt{\frac{2\pi}{15}}(a_2^{-1}(\psi) + a_2^1(\psi))\sin^2\psi + i\frac{A_0 A_\perp^\star}{\sqrt{2}}\sqrt{\frac{2\pi}{15}}(a_2^{-1} - a_2^{-1})\sin\psi\cos\psi\Big\}\Big]\Big\}
\end{aligned}
\tag{A.0.7}
$$

A set of orthonormal basis functions, in terms of spherical harmonics:

$$
\begin{aligned}
Y_{lm} &= Y_l^m (m = 0), \\
Y_{lm} &= \frac{1}{\sqrt{2}}(Y_l^m + (-1)^m Y_l^{-m}) m > 0, \\
Y_{lm} &= \frac{1}{i\sqrt{2}}(Y_l^{|m|} + (-1)^{|m|} Y_l^{-|m|}) m > 0.
\end{aligned}
\tag{A.0.8}
$$

are used to express the coefficients a_{lm} as:

$$
\begin{aligned}
a_{lm} &= a_l^m (m = 0), \\
a_{lm} &= \frac{1}{\sqrt{2}}(a_l^m + (-1)^m a_l^{-m}) m > 0, \\
a_{lm} &= \frac{1}{i\sqrt{2}}(a_l^{|m|} + (-1)^{|m|} a_l^{-|m|}) m > 0.
\end{aligned}
\tag{A.0.9}
$$

Substituting the explicit forms of Eq. A.0.9 into the expression for N, and simplifying terms, we are left with an integral that depends only on ψ:

$$\begin{aligned}
N = {} & \frac{3}{8\sqrt{\pi}}\Big[(|A_\perp|^2+|A_\parallel|^2)\int a_{00}\sin^2\psi d(\cos\psi) + 2|A_0|^2\int a_{00}\cos^2\psi d(\cos\psi)\Big] \\
& + \frac{3}{8\sqrt{5\pi}}\Big[(-|A_\perp|^2+\frac{1}{2}|A_\parallel|^2)\int a_{20}\sin^2\psi d(\cos\psi) + |A_0|^2\int a_{20}\cos^2\psi d(\cos\psi)\Big] \\
& - \frac{9}{16\sqrt{15\pi}}\frac{1}{\sqrt{1+\frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta_s)^2}}}\Big[(A_\parallel^\star A_\perp + A_\parallel A_\perp^\star)\int a_{2-1}(\psi)\sin^2\psi d(\cos\psi)\Big] \\
& + \frac{9\sqrt{2}}{16\sqrt{15\pi}}\frac{1}{\sqrt{1+\frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta_s)^2}}}\Big[(A_0^\star A_\perp + A_0 A_\perp^\star)\int a_{21}(\psi)\sin\psi\cos\psi d(\cos\psi)\Big] \\
& + \frac{9}{8\sqrt{15\pi}}\Big[\frac{|A_\parallel|^2}{2}\int a_{22}(\psi)\sin^2 d(\cos\psi) - |A_0|^2\int a_{22}(\psi)\cos^2 d(\cos\psi)\Big] \\
& + \frac{9\sqrt{2}}{16\sqrt{15\pi}}\Big[(A_0^\star A_\parallel + A_0 A_\parallel^\star)\int a_{2-2}(\psi)\sin\psi\cos\psi d(\cos\psi)\Big].
\end{aligned} \tag{A.0.10}$$

Finally N has to be integrated over ψ, the $a_{lm}(\psi)$ are then expressed as a Fourier-Legendre series:

$$a_{lm}(\psi) = a_{lm}^k P_k(\cos\psi) \tag{A.0.11}$$

And therefore the few a_{lm}^k terms after the integration read:

$$\begin{aligned}
\int a_{lm}(\psi)\sin^2\psi d(\cos\psi) &= \frac{4}{3}a_{lm}^0 - \frac{4}{15}a_{lm}^2, \\
\int a_{lm}(\psi)\cos^2\psi d(\cos\psi) &= \frac{2}{3}a_{lm}^0 + \frac{4}{15}a_{lm}^2, \\
\int a_{lm}(\psi)\sin\psi\cos\psi d(\cos\psi) &= \frac{\pi}{8}a_{lm}^1 + \frac{\pi}{32}a_{lm}^3 + \dots
\end{aligned} \tag{A.0.12}$$

which are used in Eq. A.0.10 in order to obtain the normalization for the $B_s \to J\psi\phi$

PDF, that is:

$$
\begin{aligned}
N = {} & \frac{3}{8\sqrt{\pi}}\Big[\frac{4a^0_{00}}{3}(|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2) + \frac{4a^2_{00}}{15}(|A_0|^2 - |A_\parallel|^2 - |A_\perp|^2)\Big] \\
& + \frac{3}{8\sqrt{5\pi}}\Big[\frac{2a^0_{20}}{3}(|A_0|^2 + |A_\parallel|^2 - 2|A_\perp|^2) + \frac{4a^2_{20}}{15}(|A_0|^2 - \frac{1}{2}|A_\parallel|^2 + |A_\perp|^2)\Big] \\
& - \frac{9}{16\sqrt{15\pi}}\frac{1}{\sqrt{1+\frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta_s)^2}}}\Big[(A^\star_\parallel A_\perp + A_\parallel A^\star_\perp)\Big(\frac{4}{3}a^0_{2-1} - \frac{4}{15}a^2_{2-1}\Big)\Big] \\
& + \frac{9\sqrt{2}}{16\sqrt{15\pi}}\frac{1}{\sqrt{1+\frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta_s)^2}}}\Big[(A^\star_0 A_\perp + A_0 A^\star_\perp)\Big(\frac{\pi}{8}a^1_{21} + \frac{\pi}{32}a^3_{21} + ...\Big)\Big] \\
& + \frac{9}{8\sqrt{15\pi}}\Big[\frac{2a^0_{22}}{3}(-|A_0|^2 + |A_\parallel|^2) - \frac{4a^2_{22}}{15}(|A_0|^2 + \frac{1}{2}|A_\parallel|^2)\Big] \\
& + \frac{9\sqrt{2}}{16\sqrt{15\pi}}\Big[(A^\star_0 A_\parallel + A_0 A^\star_\parallel)\Big(\frac{\pi}{8}a^1_{2-2} + \frac{\pi}{32}a^3_{2-2} + ...\Big)\Big].
\end{aligned}
\tag{A.0.13}
$$

Appendix **B**

Propagation of correlated errors

In sec. 4.6.3 the relations between the fit parameters $\alpha_{\|}$ and α_{CPOdd} and the initial transversity amplitudes squared $|A_0(0)|^2$ and $|A_{\|}(0)|^2$ have been reported. The statistical uncertainty to be associated to the measured values for $|A_0(0)|^2$ and $|A_{\|}(0)|^2$ has to be computed by using the usual error propagation:

$$\begin{aligned}
\sigma^2_{|A_0(0)|^2} &= \sum_{i=1}^{4} \frac{\partial|A_0(0)|^2}{\partial P_i}\sigma_i^2 + \sum_{i=1}^{4}\sum_{j=1}^{4} \frac{\partial|A_0(0)|^2}{\partial P_i}\frac{\partial|A_0(0)|^2}{\partial P_j}\rho(P_i,P_j) \\
\sigma^2_{|A_{\|}(0)|^2} &= \sum_{i=1}^{4} \frac{\partial|A_{\|}(0)|^2}{\partial P_i}\sigma_i^2 + \sum_{i=1}^{4}\sum_{j=1}^{4} \frac{\partial|A_{\|}(0)|^2}{\partial P_i}\frac{\partial|A_0(0)|^2}{\partial P_j}\rho(P_i,P_j)
\end{aligned} \tag{B.0.1}$$

where P_i and P_j represent the parameter entering in the expressions for the squared amplitudes ($\alpha_{CPOdd}, \alpha_{\|}, c\tau, \Delta\Gamma$), and $\rho(P_i, P_j)$ the correlation between each two of them. Let us report here the partial derivatives that are used to propagate the correlated errors according to the the usual error propagation is used. Those partial derivatives to be used have the following expressions, for $|A_0(0)|^2$:

$$\begin{aligned}
\frac{\partial|A_0(0)|^2}{\partial c\tau} &= \frac{4\frac{\Delta\Gamma}{c}(1-\alpha_{\|})(1-\alpha_{CPOdd})\alpha CPOdd}{(1+(y-1)\alpha_{CPOdd})^2(2+\frac{\Delta\Gamma}{c}c\tau)^2} \\
\frac{\partial|A_0(0)|^2}{\partial \Delta\Gamma} &= \frac{4c\tau(1-\alpha_{\|})(1-\alpha_{CPOdd})\alpha CPOdd}{(1+(y-1)\alpha_{CPOdd})^2(2+\frac{\Delta\Gamma}{c}c\tau)^2} \\
\frac{\partial|A_0(0)|^2}{\partial \alpha_{\|}} &= \frac{(1-\alpha_{CPOdd})}{(1+(y-1)\alpha_{CPOdd})} \\
\frac{\partial|A_0(0)|^2}{\partial \alpha_{CPOdd}} &= \frac{(-\alpha_{\|}(1+(y-1)\alpha_{CPOdd})-(y-1)(1-\alpha_{CPOdd})(1-\alpha_{\|}))}{(1+(y-1)\alpha_{CPOdd})^2}
\end{aligned} \tag{B.0.2}$$

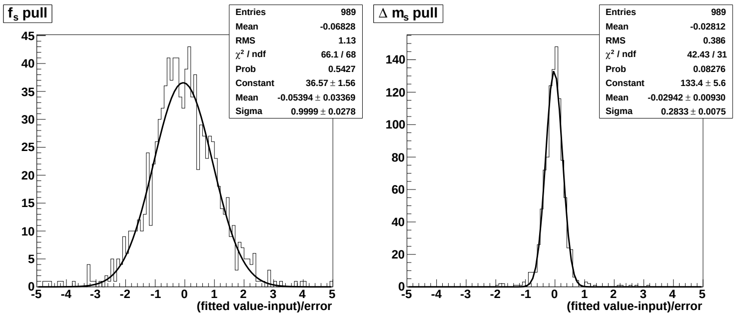
and for $|A_{\parallel}(0)|^2$:

$$\begin{aligned}
\frac{\partial|A_{\parallel}(0)|^2}{\partial c\tau} &= \frac{4\frac{\Delta\Gamma}{c}\alpha_{\parallel}(1-\alpha_{CPOdd})\alpha CPOdd}{(1+(y-1)\alpha_{CPOdd})^2(2+\frac{\Delta\Gamma}{c}c\tau)^2}\\
\frac{\partial|A_{\parallel}(0)|^2}{\partial \Delta\Gamma} &= \frac{4c\tau\alpha_{\parallel}(1-\alpha_{CPOdd})\alpha CPOdd}{(1+(y-1)\alpha_{CPOdd})^2(2+\frac{\Delta\Gamma}{c}c\tau)^2}\\
\frac{\partial|A_{\parallel}(0)|^2}{\partial \alpha_{\parallel}} &= \frac{-(1-\alpha_{CPOdd})}{(1+(y-1)\alpha_{CPOdd})}\\
\frac{\partial|A_{\parallel}(0)|^2}{\partial \alpha_{CPOdd}} &= \frac{-(1-\alpha_{\parallel}(1+(y-1)\alpha_{CPOdd})-(y-1)(1-\alpha_{CPOdd})(1-\alpha_{\parallel}))}{(1+(y-1)\alpha_{CPOdd})^2}
\end{aligned} \tag{B.0.3}$$

Appendix C

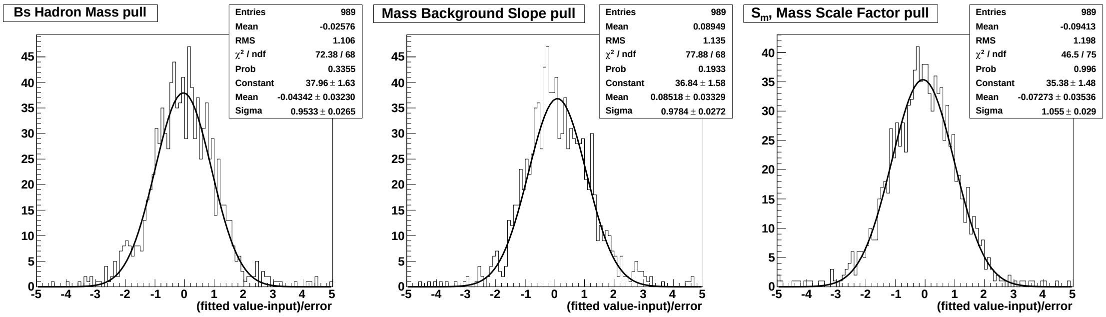
Pull distributions

pull distribution for all the parameters entering in the fit are reported here. Mean end width of the distributions are summarized in the plot in Fig. 5.6 of Chapter 5.



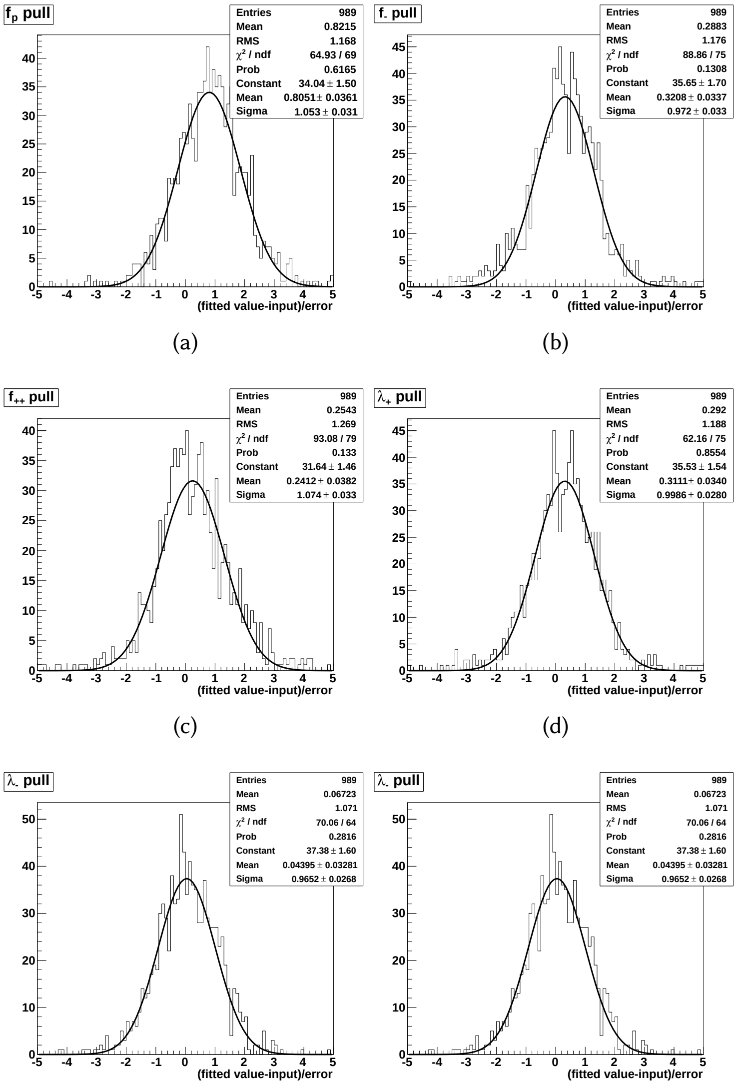
(a) (b)

Figure C.1.: Pull distributions for the signal fraction and Δm.



(a) (b) (c)

Figure C.2.: Pull distributions for the mass PDF parameters.

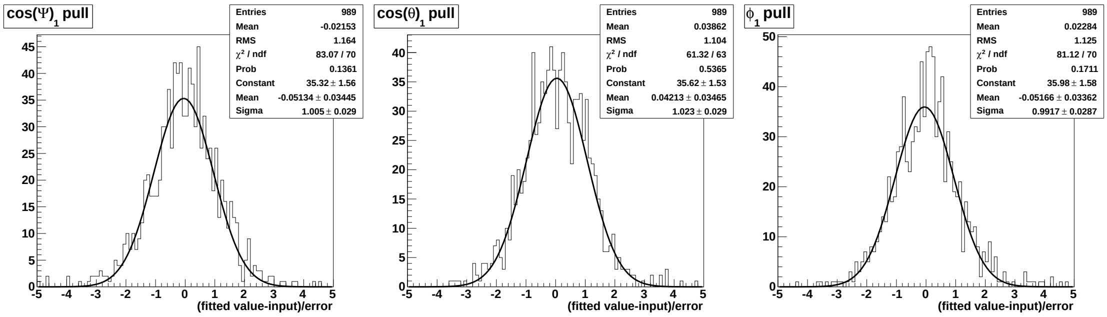


(a) (b)

(c) (d)

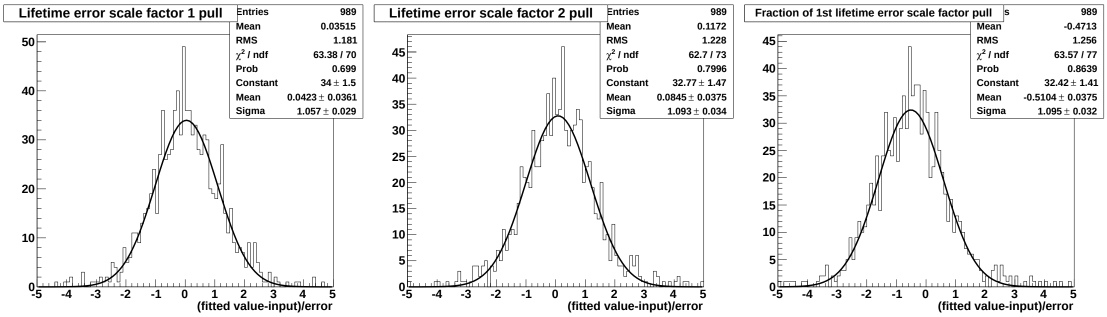
(e) (f)

Figure C.3.: Pulls distributions of the ct background distribution parameters.



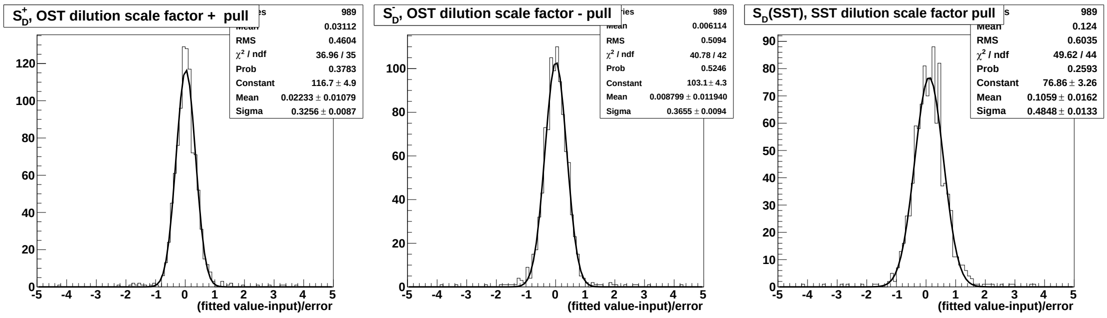
(a) (b) (c)

Figure C.4.: Pull distributions for the angular background distribution parameters.



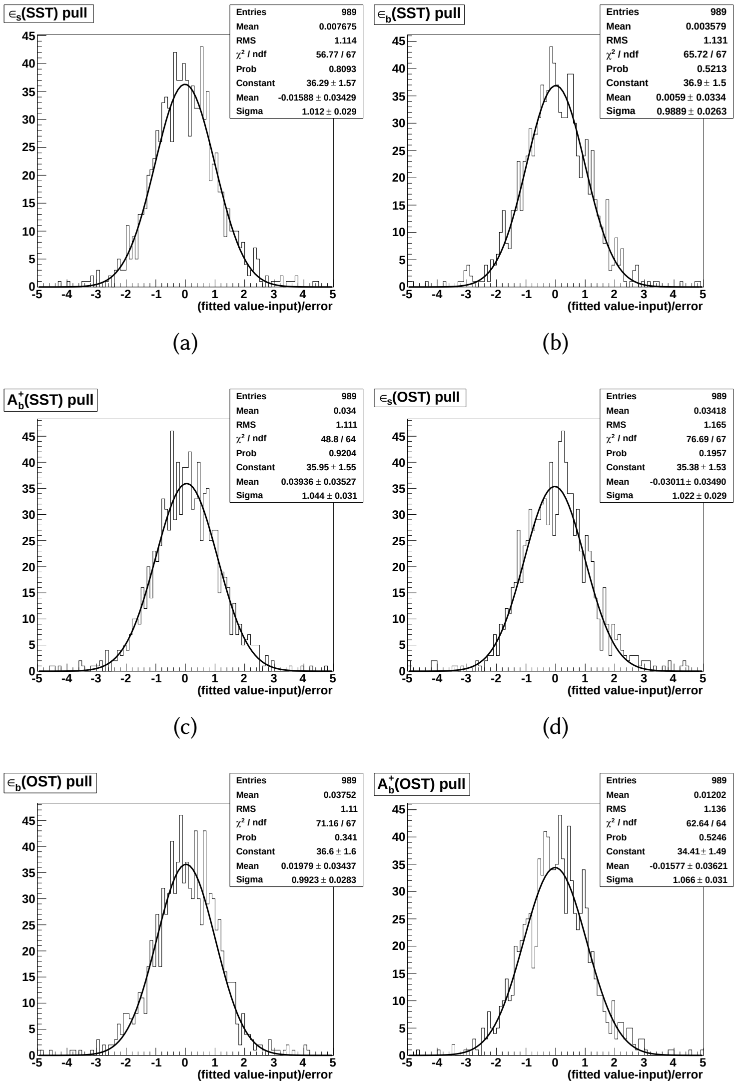
(a) (b) (c)

Figure C.5.: Pull distributions for the time resolution parameters.



(a) (b) (c)

Figure C.6.: Tagging parameters: pull distributions for the dilution scale factors.



(a) (b)

(c) (d)

(e) (f)

Figure C.7.: Tagging parameters: pull distributions for tagging efficiencies and tag asymmetry.

Acronyms

SM Standard Model

QCD Quantum Cromo Dynamics

CP Charge-conjugation-Parity

CKM Cabibbo - Kobayashi - Maskawa

NP New Physics

SLAC Stanford Linear Accelerator Center renamed SLAC National Accelerator Laboratory

HQS Heavy Quark Symmetry

HQE Heavy Quark Expansion

CDF Collider Detector at Fermilab

LHC Large Hadron Collider

FNAL Fermi National Accelerator Laboratory

RF Radio-Frequency cavitiy

CDF II CDF at Run II

CDFII CDF at Run II

DØ DØ experiment

LHCb LHCb experiment at CERN

CLC Cherenkov Luminosity Counters

COT Central Outer Tracker

SVXII Silicon VerteX II

SVX Silicon VerteX

L00 Layer 0 0

ISL Intermediate Silicon Layers

TOF Time Of Flight detector

ISL Intermediate Silicon Layers

XFT eXtremely Fast Tracker

TDC Time to Digital Converter

PID Particle IDentification

CEM Central ElectroMagnetic calorimeter

CES CEntral Strip multiwire proportional chambers

CHA Central HAdronic calorimeter

WHA Wall HAdronic calorimeter

PEM Plug ElectroMagnetic calorimeter

PHA Plug HAdronic calorimeter

CPR Central Pre-Radiator

CMU Central MUon detector

CMP Central Muon uPgrade

CMX Central Muon eXtension

BSU Barrel Scintillation counters

TSU Toroid Scintillation counters

BMU Barrel MUon chambers

IMU Intermediate MUon System

CLC Cherenkov Luminosity Counters

RF Radio Frequency cavity

VME Vesa Modul Eurocard

SVT Silicon Vertex Trigger

CSL Consumer Server/Data Logger

OST Opposite Side Tagger

NN Neural Network

SET Soft Electron Tagger

SMT Soft Muon Tagger

JQT Jet charge Tagger

SSKT Same Side Kaon Tagger

MC Monte Carlo

PDF Probability Density Function

ML Maximum Likelihood

Bibliography

[1] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Phys. Rev. Lett.* 13, 138 (1964).

[2] A. Lenz and U. Nierste, J. *High Energy Phys.* 06, 072 (2007).

[3] J. Charles et al. (CKMfitter Group), *Eur. Phys.* J. C 41, 1 (2005).

[4] M. Bona et al. (UTfit Collaboration), *J. High Energy Phys.* 10, 081 (2006).

[5] S. Nandi and A. Soni, *Phys. Rev.* D 83, 114510 (2011).

[6] J. Drobnak, S. Fajfer, and J. F. Kamenik, *Phys. Lett. B* 701, 234 (2011).

[7] R.-M. Wang, Y.-G. Xu, Q. Chang, and Y.-D. Yang, Phys. Rev. D 83, 095010 (2011).

[8] J. Shelton and K. M. Zurek, *Phys. Rev. D* 83, 091701 (2011).

[9] J. Girrbach et al., J. *High Energy Phys.* 06, 044 (2011), [Erratum-ibid.07:001,2011].

[10] A. J. Buras, M. Nagai, and P. Paradisi, *J. High Energy Phys.* 05, 005 (2011).

[11] Z. J. Ajaltouni and E. Di Salvo, *J. Phys. D* 37, 125001 (2010).

[12] B. A. Dobrescu, P. J. Fox, and A. Martin, *Phys. Rev. Lett.* 105, 041801 (2010).

[13] Y. Bai and A. E. Nelson, *Phys. Rev. D* 82, 114027 (2010).

[14] L. L. Everett, J. Jiang, P. G. Langacker, and T. Liu, *Phys. Rev. D* 82, 094024 (2010).

[15] J. P. Saha, B. Misra, and A. Kundu, *Phys. Rev. D* 81, 095011 (2010).

[16] G. Isidori, Y. Nir, and G. Perez, Ann. Rev. *Nucl. Part.* Sci. 60, 355 (2010).

[17] F. J. Botella, G. C. Branco, and M. N. Rebelo, *Phys. Lett. B* 687, 194 (2010).

[18] C.-W. Chiang et al., J. *High Energy Phys.* 04, 031 (2010).1135

[19] Z. Ligeti, M. Papucci, and G. Perez, *Phys. Rev. Lett.* 97, 101801 (2006).

[20] A. Lenz and U. Nierste, JHEP 0706 (2007) 072 [arXiv:hep-ph/0612167].

[21] M. Gronau, J. L. Rosner, *Phys. Lett. B* 669, 321 (2008).

[22] The BaBar *Physics Book* eds. P.F. Harrison and H.R. Quinn (SLAC report 504, October 1998)

[23] W. Beenakker, H. Kuijf, W. L. van Neerven, and J. Smith. QCD corrections to heavy-quark production in $p\bar{p}$ collisions. *Phys. Rev. D*, 40(1):54Ð82, Jul 1989

[24] J. C. Collins, D. E. Soper, and G. Sterman. Perturbative QCD, 1989.

[25] Mario Greco and Simona Rolli. Light mesons production at the Tevatron to next-to-leading order. *Physical Review D*, 52:3853, 1995.

[26] S. Kretzer. Fragmentation Functions from Flavour-inclusive and Flavour-tagged e^+e^- Annihilations. *Physical Review D*, 62:054001, 2000.

[27] K. Nakamura et al. (Particle Data Group), *J. Phys. G* 37, 075021 (2010) and 2011 partial update for the 2012 edition.

[28] Isard Dunietz, Robert Fleischer, and Ulrich Nierste. In Pursuit of New Physics with B_s Decays. *Physical Review D*, 63:114015, 2001.

[29] F. Azfar et al. Formulae for the Analysis of the Flavor-Tagged Decay $B_s \to J/\psi\phi$ JHEP 1011:158,2010 arXiv:1008.4283v2

[30] A.S. DIghe, I. Dunietz and R. Fleischer, *Eur. Phys. J.* **C6** (1999) 647.

[31] K. Anikeev et al. B Physics at the Tevatron: Run II and Beyond (2002). arXiv:hep-ph/0201071

[32] Michal Kreps *Mod. Phys. Lett.*A26:2007-2028 (2011). arXiv:1108.6257

[33] G. J. Feldman and R. D. Cousins, *Phys. Rev. D* 57, 3873Ð3889 (1998), [arXiv:physics/9711021].

[34] BABAR-PUB-04/030 SLAC-PUB-10808 hep-ex 0411016 *Physical Review D*, 2004

[35] T. Aaltonen et al. (CDF Collaboration) *Phys. Rev. Lett.* 100:161802 (2008) arXiv:0712.2397

[36] T. Aaltonen et al. (CDF Collaboration) CDF public note 9458, 2008 (Unpublished)

[37] T. Aaltonen et al. (CDF Collaboration) CDF public note 10206, 2010 (Unpublished)

[38] V. M. Abazov et al (D0 Collaboration) *Phys. Rev. Lett.*101:241801 (2008), arXiv:0802.2255.

[39] V. M. Abazov et al (D0 Collaboration), D0 Conference Note 6098-CONF, 2010 (Unpublished) arXiv:0802.2255.

[40] The LHCb Collaboration, LHCb Conference Note LHCb-CONF-2011-049, 2011 (Unpublished).

[41] G. Aubrecht and et al, A Teachers Guide To The Nuclear Science Wall Chart. Contemporary Physics Education Project, 2003.

[42] M. Popovic, L. Allen, and C. Schmidt, Fermilab linac injector, revisited, prepared for 16th IEEE Particle Accelerator Conference (PAC 95) and International Conference on High-energy Accelerators (IUPAP), Dallas, Texas, 1-5 May 1995.

[43] D. Mohl, Physics and technique of stochastic cooling, *Phys. Reports*, vol. 58, 1980.

[44] S. Nagaitsev and et al, Experimental demonstration of relativistic electron cooling, *Phys. Rev. Lett.*, vol. 96, 2006.

[45] F. B. Division, RUN II Handbook, 1998. [Online]. Available: http: //www-bd.fnal.gov/lug/runII handbook/RunII index.html

[46] D. Tonelli, First observation of the $B_s^0 \to K^+K^-$ decay mode, and measurement of the B^0 and B_s^0 mesons decay-rates into two-body, charmless final states at cdf, Ph.D. dissertation, Scuola Normale Superiore di Pisa, 2006.

[47] C. CDFII, The CDF IIb detector: Technical Design Report.

[48] The CDF II Detector Technical Design Report, October, 1996, sum- mary of the CDF RunII Detector Parameters.

[49] Proposal for Ehnancement of the CDF II Detector: an Inner Silicon Layer and a Time of Flight Detector, October 23, 1998.

[50] D. Acosta and et al, The performance of the CDF luminosity monitor, *Nucl. Instrum. Meth.*, vol. A494, pp. 57Ð62, 2002.

[51] , F. Abe and et al, The CDF detector: an overview,*Nucl. Instr. Meth.*, vol. A271, pp. 387Ð403, 1988, FERMILAB-PUB-94/024-E (1994).

[52] R. Blair and et al, The CDF-II detector: Technical Design Report, FERMILAB-PUB-96/390-E (1996).

[53] T. Nelson, The CDF Layer 00 detector, *Int. J. Mod. Phys.*, vol. A16S1C, pp. 1091Ð1093, 2001.

[54] A. Sill, CDF run II silicon tracking projects, *Nucl. Instrum. Meth.*, vol. A447, pp. 1Ð8, 2000.

[55] A. A. Affolder and et al, Status report of the intermediate silicon layers detector at CDFII, *Nucl. Instrum. Meth.*, vol. A485, pp. 6Ð9, 2002.

[56] I. Vila for the CDF Collaboration Performance and First Physics Results of the SVT Trigger at CDF II ECONF 2003 [arXiv:hep-ph/0307165].

[57] B. Ashmanskas et al. (CDF Collaboration) The CDF Silicon Vertex Trigger *Nucl. Instrum. Meth.* A518 (2004) 532 [arXiv:physics/0306169]

[58] K. T. Pitts, ÒThe CDF central outer tracker,Ó *Nucl. Phys. Proc. Suppl.*, vol. 61B, pp. 230Ð235, 1998.

[59] T. Aolder et al. (CDF Collaboration) COT Central Outer Tracker *Nucl. Instrum. Meth.* A526 (2004) 249.

[60] See CDF Run II Detector Parameters, COT Simulation Validation Plots.

[61] D. Acosta et al. (CDF Collaboration) A Time-Of-Flight Detector in CDF II *Nucl. Instrum. Meth.* A518 (2004) 605.

[62] C. S. Hill, Operational experience and performance of the CDFII silicon detector, *Nucl. Instrum. Meth.*, vol. A530, pp. 1Ð6, 2004.

[63] [Online]. Available: http://penn01.fnal.gov

[64] S. DÕAuria, D. Lucchesi, S. D. Ronco, R. Carosi, M. Ciocci, S. Donati, P. Catastini, M. Morello, G. Punzi, P. Squillacioti, D. Tonelli, S. Torre, M. Rescigno, S. DeCecco, S. Giagu, M. Casarsa, and A. Cerri, Track-based calibration of the cot specific ionization CDF-Note 6932, 2004.

[65] S. E. Yu, COT dE/dx measurement and corrections, CDF-Note 6361, 2004.

[66] P. Azzi, G. Busetto, P. Gatti, and A. Ribon, Histogram tracking in the cot, CDF-Note 5562, 2001.

[67] F. Snider, Tracking at CDF: Algorithms and experience from Run I and Run II, *Nucl. Instrum. Meth.*, vol. A566, pp. 133Ð141, 2006.

[68] J. Marriner Secondary vertex fit with mass and pointing constraints (CTVMFT) CDF internal note 1996 (1993) unpublished.

[69] P. Gatto. Performance of the new tracking system at CDF II. MasterÕs thesis, University of Padova, 2001. Ph.D. thesis, FERMILAB-THESIS-2001-23. (Cited on page 42.)

[70] S. Menzemer. TrackingCal - A tracking and alignment software package for the CDF II silicon detector. MasterÕs thesis, University of Karlsruhe, 2002. CDF Internal Note 5968, unpublished.

[71] C. P. Hays and et al, Inside-out tracking at CDF, *Nucl. Instrum. Meth.*, vol. A538, pp. 249Ð254, 2005.

[72] J. Goldstein, C. Issever, T. Nelson, R. Snider, and D. Stuart, Silicon tracking for plug electrons, CDF-Note 5970, 2002.

[73] D. Acosta and et al, A time-of-flight detector in CDF-II, *Nucl. Instrum. Meth.*, vol. A518, pp. 605Ð608, 2004.

[74] C. Ginsburg, CDF run 2 muon system, *Eur. Phys. J.*, vol. C33, pp. s1002Ð s1004, 2004.

[75] G. Ascoli and et al, CDF central muon detector, *Nucl. Instrum. Meth.*, vol. A268, p. 33, 1988.

[76] D. Acosta and et al, The CDF Cherenkov luminosity monitor, *Nucl. In- strum. Meth.*, vol. A461, pp. 540Ð544, 2001.

[77] S. Klimenko J. Konigsberg and T. Liss Averaging of the inelastic cross sections measured by the CDF and the E811 experiments Fermilab-FN-0741 (2003).

[78] S. Jindariani et al. Luminosity Uncertainty for Run 2 up until August 2004 CDF Note 7446.

[79] G. Gomez-Ceballos and et al, Event builder and level 3 at the cdf experiment, *Nucl. Instrum. Meth.*, vol. A518, pp. 522Ð524, 2004.

[80] The Trigger and Dataset Working Group, Run II trigger table and datasets plan, CDF Note 4718, 2001.

[81] E. Gerchtein and M. Paulini. CDF detector simulation framework and performance. In eConf. C0303241 TUMT005, 2003. [physics/0306031].

[82] R. Braun et al. GEANT: Simulation Program For Particle Physics Experiments. User Guide And Reference Manual, 1978. CERN-DD-78-2-REV.

[83] P. A. Movilla-Fernandez. Performance of the CDF Calorimeter Simulation in Tevatron Run II. In R. Yoshida S. Magill, editor, Proceedings of the XII International Conference on Calorimetry in High Energy Physics (CALOR 06), 2006. [physics/0608081].

[84] R. Veenhof. GARFIELD, recent developments. *Nucl. Instrum. Methods*, A419, 1998.

[85] M. Gold. Description of the Parameterized Charge Deposition Model. CDF Internal Note, 5871, 2002. unpublished.

[86] K. Burkett et al, CTVMFT, CDF note 6430.

[87] S-S Yu at al., COT dE/dx Measurement and corrections, CDF Note 6361

[88] COT Track Based Calibrations for Summer Conferences, (2004) http://www-cdf.fnal.gov/internal/physics/bottom/COTDEDX/COT_dedx_cali.html

[89] Ch. Paus et al., BottomMods, cdfcodebrowser.fnal.gov/CdfCode/source/BottomMods/.

[90] P. Murat et al., Stntuple, cdfcodebrowser.fnal.gov/CdfCode/source/Stntuple/.

[91] F. Azfar, L. Oakes, D. Tonelli , Extended and improved track-based dE/dx calibration for 3/fb analyses, CDF Note 9592

[92] M. Feindt, M. Kreps, Th. Kuhr, J. Morlock, Particle Dependent Time-of-Flight Calibration, CDF Note 9618

[93] T. Aaltonen et al. (CDF Collaboration), *Phys. Rev. Lett.* 100, 121803 (2008)

[94] G. Giurgiu, Ph.D. thesis, Carnegie Mellon University, FERMILAB-THESIS-2005-41 (2005)

[95] M. Feindt, A Neural Bayesian Estimator for Conditional Probability Densities, arXiv:physics/0402093.

[96] G. Cowan, Statistical Data Analysis, Oxford Science Publications (1998).

[97] F. Azfar et al, An updated measurement of the CP violating phase β_s in $B_s \to J/\psi\phi$ decays using $L = 5.2\,\mathrm{f}b^{-1}$, CDF note 10053.

[98] K. Makhoul CP violation in Flavor-Tagged $B_s \to J/\psi\phi$ decays, P.h.D. dissertation, 2009, MIT

[99] V.Tiwari et at, Likelihood based electron tagger, CDF note 7121

[100] D. Usynin et al, B flavor identification using opposite side muons, CDF note 6483

[101] Bs Mixing Group, A Neural Network based jet charge tagger, CDF note 7492

[102] A. Abulencia et al. Observation of B_s^0 - B_s^0 Oscillations, *Phys. Rev. Lett.*, 97 (2006) 242003.

[103] M. Feindt, T. Kuhr, M. Kreps, J. Morlock, A. Schmidt, Measurement of Mixing Frequency and Amplitude of Neutral B-Mesons, CDF note 10025

[104] M. Feindt, T. Kuhr, M. Kreps, J. Morlock, A. Schmidt, Calibration of the Same Side Kaon Tagger, CDF public note 10108.

[105] I. Furic et al, Combined opposite side flavour tagger, CDF note 8460.

[106] K. Anikeev, P. Murat, and Ch. Paus. Description of BGenerator II, 1999. CDF Internal Note 5092.

[107] D. J. Lange. The EvtGen Particle Decay Simulation Package. *Nucl. Instrum. Meth.*, 2001.

[108] CDF Simulation Group. CDF Detector Simulation, 2000. www-cdf.fnal.gov/cdfsim/cdfsim main.html.

[109] J. Boudreau, J.P. Fernandez, K. Gibson, G. Giurgiu, G. Gomez-Ceballos, L. Labarga, C. Liu, P. Maksimovic, M. Paulini, Measurement of the CP Violation Paramter betaÕ in $B_s to J/\psi\phi$, CDF note 8960.

[110] M. Feindt, M. Kreps, T. Kuhr, M. Milnik, Measurement of DeltaGamma and deltaphi in untagged $B_s \to J/\psi\phi$ and $B_d \to J/\psi K^*$ decays, CDF note 8753

[111] M. Milnik. Measurement of the lifetime difference and CP-violating phase in $B_s \to J/\psi\phi$ decays. 2007. FERMILAB-THESIS-2007-38.

[112] F. James and M. Roos, Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations, Comput. Phys. Commun., 10:343Ð367, 1975.

[113] R. Brun and F. Rademakers, ROOT: An Object Oriented Data Analysis Frame- work, *Nucl. Instrum. Meth.*, A389:81Ð86, 1997.

[114] G. Punzi, Comments on Likelihood fits with variable resolution, PHYSTAT2003, SLAC Sept 8-11 , 2003 arXiv:physics/0401045

[115] https://proj-clhep.web.cern.ch/proj-clhep

[116] M. Kreps, Observation of $B_s \to J/\psi f_0(980)$ decays, CDF note 10334.

[117] T. Aaltonen et al. (CDF Collaboration), *Phys. Rev. Lett.* 100, 161802 (2008). 1140.

[118] V. M. Abazov et al. (D0 Collaboration), *Phys. Rev. Lett.* 101, 241801 (2008).

[119] V. M. Abazov et al. (D0 Collaboration), arXiv:hep-ex/1109.3166 (2011).

[120] A. Lenz and U. Nierste, J. *High Energy Phys.* 06, 072 (2007).

[121] F. James, Minuit: Function minimization and Error analysis (Minuit reference manual)
http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minuit.ps

[122] T. Aaltonen et al. (CDF Collaboration), *Phys. Rev. Lett.* 104, 102002 (2010).

[123] T. Aaltonen et al. (CDF Collaboration), *Phys. Rev. Lett.* 106, 121804 (2011).

[124] K. Anikeev, Ph.D. thesis, MIT, FERMILAB-THESIS-2004-12 (2004).

[125] T. Aaltonen et al. (CDF Collaboration), *Phys. Rev. Lett.* 100, 121803 (2008) [arXive:0712.2348].

[126] T. Aaltonen et al. (CDF Collaboration), Measurement of b hadron lifetimes in exclusive decays containing a J/psi in $p-\bar{p}$ collisions at $\sqrt{(}s) = 1.96TeV$ *Phys.Rev.Lett.*106:121804,2011[arXiv:1012.3138v2]

[127] M. Beneke et al., Nucl. Phys. B 639, 389 (2002); E. Franco, V. Lubicz, F. Mescia, and C. Tarantino, Nucl. Phys. B 633, 212 (2002); C. Tarantino, Eur. Phys. J. C 33, S895 (2004); A. J. Lenz, AIP Conf. Proc. 1026, 36 (2008).

Ringraziamenti

Un semplice ringraziamento alla fine della tesi non potrà mai esaurire tutta la gratitudine che provo, alla fine di questi cinque anni di università, e all'inizio di tutto ciò che verrà dopo. Non rinuncio al tentativo di ringraziare:
La mia famiglia, mamma, papà e nonna in particolare, che prima di chiunque altro hanno saputo sostenermi nella vita di questi anni.
Gli amici che sono stati per me compagnia vera in questi anni e in questi ultimi mesi e giorni. Le amiche con cui ho abitato e sto abitando ora, perchè la vita quotidiana svela in modo inequivocabile ciò che il cuore desidera.
I miei compagni di corso di Trieste e di Copenhagen, e gli amici insieme ai quali ho imparato tutta la fisica che so.
I docenti che, insegnando con cura e dedizione, mi hanno fatto appassionare alle loro materie.
Desidero, inoltre, ringraziare la mia relatrice e il mio correlatore, per l'attenzione dedicatami in questi mesi di tirocinio e di tesi. Ringrazio anche tutte le persone del gruppo di CDF con cui ho potuto lavorare all'AREA di Ricerca di Trieste e al Fermilab, per la serenità e la disponibilità dell'ambiente che ho trovato, e per tutto cio' che ho potuto imparare.
Anna Maria per avermi permesso di vivere questa avventura, e per avermi seguita nel lavoro.
Mirco, per tutta la fisica che mi ha insegnato, per il tempo dedicatomi, la pazienza nel correggermi, per il sostegno, e l'aiuto concreto senza i quali, non avrei concluso nulla. Nessun altro dottorando ne sarebbe stato capace.
Marco e Diego, per i commenti e correzioni in corso d'opera.
Massimo, per il gli insegnamenti ed il sostegno in questi mesi, arrivato fino alla praticità della correzione di pezzi di tesi e codice, e per la settimana di shift a CDF.

Lucia Grillo, *18 Novembre 2011*