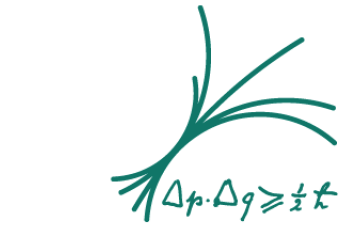


Master's Thesis

Optimization of the tagged B meson vertex resolution for the Belle II experiment

Christian Roca Catala

September, 2015 München

Supervisor: Prof. Dr. Christian Kiesling

LUDWIG-MAXIMILIAN UNIVERSITÄT - MAX PLANCK INSTITUT FÜR PHYSIK
DEPARTMENT OF PHYSICS

BELLE2-MTHESIS-2016-001
30/09/2015

Abstract

This thesis consists of two correlated but differentiated parts, both of them regarding the optimization of the vertexing procedure of the Belle II experiment. The time dependent analysis of the decay mode:

$$\mathrm{B}^0(\bar{\mathrm{B}}^0) \to J/\psi K_S$$
$$\bar{\mathrm{B}}^0(\mathrm{B}^0) \to \text{generic}$$

has been used as benchmark for the Belle II vertexing performance. In order to perform this analysis the vertex of the two exclusively produced B mesons must be determined. The first neutral B meson, usually called $\mathrm{B_{CP}}$, is fully reconstructed then its decay vertex can be calculated with a kinematic fit of the two tracks corresponding to the two muons of the J/ψ decay. The vertex of the other B is determined without reconstructing it. The complete reconstruction, in fact, would produce a very low efficiency. The usual procedure is to take all the tracks remaining after the signal reconstruction and fit them in one single vertex. For this purpose, a spatial constraint has been designed so that contributions of tracks pointing to secondary vertices of the tag side weight less during the fit. This prevents potential bias on the $\mathrm{B_{tag}}$ vertex. This procedure is called *Standard Algorithm*.

The first part of this study treats about the *Standard Algorithm*. More concretely, a deep analysis about the stability of the spatial constraint used in the algorithm is made. Shifts in different directions are applied to the center of the constraint, allowing the stability study. The results are very conclusive: a part from extreme cases, the shifts applied had little to no effect in the Tag Side vertex resolution.

The inclusion of the newly developed Pixel Vertex Detector in the Belle II detector simulation has improved the full reconstucted B vertex resolution with respect to the Belle results by almost a factor 3. Nevertheless, this vertexing improvement cannot be translated to the tag side as its resolution depends heavily on the fitting algorithm. The second part of this thesis explores the idea of doing a selection of the tracks, that are more suitable for the $\mathrm{B_{tag}}$ vertex fit. The track's selection is performed by analysing different properties of the tracks than could indicate whether the track comes either directly from the $\mathrm{B_{tag}}$ (primary track) or from a secondary vertex (secondary track). The main intention is to substitute or complement the *Standard Algorithm*. Two options are studied, called *Secondary Track Rejection* and *Single Track Fit*:

- **Secondary Track Rejection**: selection criteria are applied to reject secondary tracks. This method ends up being unsuccessful, leading to a worse resolution compared to the one obtained with the *Standard Algorithm*.

- **Single Track Fit**: this algorithm selects only one track to perform the fit, ideally a primary track. The results are very promising since the resolution is improved by a factor 1.4 with respect to the results obtained via *Standard Algorithm*. The total improvement with respect to the Belle collaboration's results is of a factor 2.2. As the efficiency of this algorithm is about a 15%, it needs to be combined with the *Standard Algorithm*.

Contents

Introduction

During the last two decades B-meson physics have become an important branch of modern particle physics. The two B-factories, the Belle experiment at the asymmetric electron positron collider KEKB [1] in Tsukuba, Japan, and the BaBar experiment at the PEP-II [2] collider at SLAC laboratory in California, USA, yielded a rich harvest of results. Among those, it's included the first observation of CP violation outside of the kaon system, measurements of the CKM parameters V_{ub} and V_{cb}, and measurements of purely leptonic B meson decays until the ending of data collection in 2010 and 2008 respectively. These experimental tests culminated in the 2008 Nobel Prize for physics awarded to M. Kobayashi [3] and T. Maskawa [4] for their theory of CP violation.

Nevertheless, this was only the first step of many towards a full understanding of flavour physics, both at this side or Beyond the Standard Model (BSM). Despite the good agreement of the results obtained with the expectations of the SM, further measurements of CP violation in B decays are needed to over constrain the Unitary Triangle.

Although the Standard Model has been very successful in predictions and interpretations of current measurements, it has become clear that it cannot answer all the outstanding questions. To resolve the remaining issues many new theories going BSM have been developed. These New Physics (NP) scenarios introduce new particles and processes. Tests of the NP processes in the existing data face the problem of large statistical errors for the crucial observables, hence only new high statistics experiments like Belle II would bring new insights into the field [5].

With this purpose in mind, SuperKEKB, the massive upgrade of the collider KEKB, aims at an integrated luminosity in excess 50 ab^{-1}. It will deliver an instantaneous luminosity of $8 \cdot 10^{35}$ cm^{-2} s^{-1}, which is 40 times higher than the world record set by its predecessor. This high luminosity is expected to shed light into all those unknowns presented above. In order to deal with this increase of statistics, the Belle II detector require an entirely new tracking, consisting of a Pixel

Vertex Detector (PXD), a Silicon strip Vertex Detector (SVD) and a Central Drift Chamber (CDC) [6].

The SuperKEKB factory's electron positron beam energies are chosen to be 7 GeV and 4 GeV respectively. The center-of-mass energy is approximately 10.58 GeV corresponding to the rest mass of the bottomonium resonance $\Upsilon(4S)$ which decays almost exclusively into a pair of B-mesons (B^0-$\bar{B}^0$and B^+-B^- with practically the same probability). The beams are chosen to be energetically asymmetric so that the centre-of-mass is moving relatively to the detector with a boost of $\gamma\beta = 0.283$ in the electron beam direction, and thus the decay products are detected at different space points simplifying the event reconstruction. Under this setup, time-dependent CP violating effects can be studied in Belle II by using the distribution's shape of the difference of the decay times of the neutral B mesons, called Δt [6]. Moreover, given that $\Delta t = \Delta z/\gamma\beta c$, the precision on Δt relies completely on the decay position difference of both B mesons, Δz, and consequently on the resolution of their vertices.

The Belle II Pixel Vertex Detector based on DEPFET technology [7, 8] will be installed on the beam pipe very close to the electron-positron interaction point and will allow Belle II to reconstruct B-meson vertices with an unprecedented resolution, meaning that time-dependent CP violating effects will be measured with very high precision. Nevertheless, the advanced technology of the Pixel Vertex Detectors could result helpless given a poor performance of the vertexing software. This is the reason why a new vertexing software is needed for Belle II, and the motivation for the present work.

In this thesis only the neutral meson decay mode $B^0(\bar{B}^0) \to [J/\psi \to \mu^+\mu^-][K_S \to \pi^+\pi^-]$ and $\bar{B}^0(B^0) \to$ generic is considered. The election of this channel is not arbitrary. Experimentally, the two muons from $J/\psi \to \mu^+\mu^-$ give one of the cleanest signals that can be observed which makes it ideal for a vertex measurement. Theoretically, the $B^0(\bar{B}^0) \to J/\psi K_S$ decay corresponds to a tree level process to CP eigenstate. Therefore it provides a clean measurement of the neutral B mixing parameter. For all this reasons, this mode is usually called *golden mode* [9] and allows to precisely test the Belle II performance in time dependent analysis.

Given that $B^0(\bar{B}^0) \to [J/\psi \to \mu^+\mu^-][K_S \to \pi^+\pi^-]$ holds for both $B^0$and $\bar{B}^0(B_{CP})$, the flavor of the meson can not be determined by the reconstruction of this decay, and thus a flavor identification needs to be carried out by using the other produced B meson. Both mesons are created in a quantum-entangled state which forbids them to oscillate freely before they decay [10], and then the

identification of the flavor of one of the B mesons (tagged B or B_{tag}) at its decay determines the flavor of the other B meson at the same time. The charge of the primary tracks coming from the B meson unequivocally relates to the flavor of the decay and thus determines whether B_{tag} (B_{CP}) is a $\bar{B}^0(B^0)$ or a $B^0(\bar{B}^0)$. A good example of this could be the decay mode $B^0 \to \mu^+\nu_\mu D^{*-}$, $\bar{B}^0 \to \mu^-\bar{\nu}_\mu D^{*+}$, where the (anti)muon acts like a clear signature of the flavor of the B meson.

CHAPTER 1

Physics Motivation

1.1 C,P symmetries

In physics, a symmetry is a physical feature of the system that is preserved or remains unchanged under some transformation. From Noether's theorem we know that each continuous symmetry in a given system leads to a conserved charge. Some well known examples are spatial or time translation invariance leading to momentum or energy conservation, or invariance under rotations leading angular momentum conservation. But besides continuous symmetries there are plenty of discrete symmetries which are very relevant in particle physics, like parity, charge conjugation or the combination of both of them. The correspondent transformations are:

Parity transformation:
A parity transformation can be performed as a reflection (mirror transformation) followed by a rotation of 180 degrees along the axis we used for mirroring. It can be interpreted as an inversion of the three spatial axes:

$$x \xrightarrow{P} -x \quad y \xrightarrow{P} -y \quad z \xrightarrow{P} -z$$

The idea is that the mirror image of a reaction occurs at the same rate as the original reaction, fact that appears to be valid for all reactions involving electromagnetism and strong interactions. Until 1956, parity conservation was believed to be one of the fundamental geometric conservation laws (along with conservation of energy and conservation of momentum). However, in 1956 a careful critical review of the existing experimental data by theoretical physicists Tsung-Dao Lee

and Chen Ning Yang [11] revealed parity symmetry was untested in the weak interaction, followed by a proposal of several possible direct experimental tests. The first test based on p decay of Cobalt-60 nuclei was carried out in 1956 by a group led by Chien-Shiung Wu [12], and demonstrated conclusively that weak interactions violate the P-symmetry.

Charge conjugation:
The charge conjugation inverts the sign of all the additive quantum numbers of a system, like charge, baryon or lepton number, but it does not change mass, energy, momentum or spin of particles involved. In other words, a charge conjugation operation transforms a particle into its own antiparticle. If a system is symmetric under a C-parity transformation it means the same physics apply for its antimatter homologous system.
After the experimental verification of the P-symmetry violation in the weak interactions, in 1958 Maurice Goldhaber [13] performed a measurement of the neutrino's helicity which lead to the experimental proof of C-symmetry violation. The results were clear, only left (right) handed neutrinos (antineutrinos) appear on nature, which is the ultimate proof of an asymmetry between matter and antimatter. Namely, the C-conjugate of LH neutrinos (LH antineutrinos) does not exist, and neither does the C-conjugate of RH antineutrinos (RH neutrinos).

1.2 CP Violation in SM

For a long time, the physics community thought that C, P and T symmetries were inviolable in the particle physics world. This changed once the acclaimed $\theta - \tau$ puzzle was solved. The existence of the kaon as parity violating particle through its decay changed the mind of the physicists in 1954 [11].

After finding that both C and P were not fundamental symmetries of the nature, CP-symmetry was proposed in 1957 by Lev Landau as the true symmetry between matter and antimatter [14]. This kind of transformation corresponds to applying a parity transformation followed by a charge conjugation, like for example when going from chiral left-handed electron e_L^- to chiral right-handed positron e_R^+.

Nevertheless, ten years afterwards (in 1964), James Cronin, Val Fitch and coworkers provided clear evidence [15] that CP-symmetry could also be broken via kaon rare decays, and its experimental verification shocked the particle physics world and opened the door to questions still at the core of particle physics and of cosmology. At that point, the Standard Model needed some updates to accommo-

date such a big change. A satisfactory explanation for this phenomena would have to wait until the 1970s to be given by two young postdocs called M. Kobayashi and T. Maskawa. Based on Cabibbo's previous work on quark mixing, Kobayashi and Maskawa predicted the existence of a third family of quarks so that CP violation could be understood as a quark mixing under the three dimensional CKM complex matrix. Their work was rewarded in 2008 with a Nobel Prize.

Nowadays there are three possible sources of CP violation: quark mixing via CKM matrix (measured), netrino mixing via PMNS matrix (expected to be measured soon [16]) and the QCD intrinsic CP violation (measurement failure - strong CP problem).

1.2.1 Cabibbo Mixing

Before going to CP violation problem, is worth to understand the previous work on quark mixing that lead to Maskawa and Kobayashi to postulate the mixing via CKM matrix. This previous work relies on Cabibbo's quark mixing studies [17]. In the decade of the 60s, the different coupling constant in the kaon decay was puzzling the physicist community. In fact the observed existence of at least three different Fermi coupling constants G_μ, G_n and G_K was in complete contradiction with the concept of a universality in weak interactions, opposite to what was found for electromagnetic and strong interactions. Thus, a roundabout was needed in order to keep only one weak constant G_F (so called Fermi's constant from then on) and to preserve universality. The solution to this problem finally came with the hypothesis of flavor mixing, implying that the quantum eigenstates of the weak interactions are rotated in quark flavour space with respect to the mass eigenstates. In other words, the weak interaction eigenstates can be written as a linear combination of the mass (or propagation) eigenstates, as follows:

$$d' = \cos\theta_C d + \sin\theta_C s \tag{1.1}$$

$$s' = -\sin\theta_C d + \cos\theta_C s \tag{1.2}$$

This concept was first postulated by Nicola Cabibbo in 1963 following an earlier suggestion by Murray Gell-Mann and Maurice Levy. For this reason the mixing is parametrized in terms of the so called Cabibbo angle θ_C.

1.2.2 Cabibbo-Kobayashi-Maskawa Matrix

Soon after Cabibbo's idea and almost ten years before the discovery of charm quark, CP violation was observed in the study of rare kaon decays by Christenson

and company. This effect is difficult to accommodate for two families of quarks. However an extension to three families allows it to be taken into account naturally. The six-quark model was proposed by Kobayashi and Maskawa in 1973, extending Cabibbo's 2x2 quark mixing matrix into the 3x3 Cabibbo-KobayashiMaskawa (CKM) matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_{weak} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{mass} \tag{1.3}$$

In order to understand why three families are needed, it is necessary to count the number of parameters in this matrix, $V_{CKM} = V$.

An N × N unitary matrix [1] requires N^2 real parameters to be specified. $2N-1$ of these parameters are not physically significant, because one phase can be absorbed into each quark field (both of the mass eigenstates, and of the weak eigenstates), in the following way:

$$u \to ue^{i\phi} \quad d \to de^{i\phi}$$

Hence, the total number of free variables independent of the choice of the phases of basis vectors is $N^2 - (2N-1) = (N-1)^2$. This amount of variables can be subdivided as:

$$(N-1)^2 = N(N-1)/2 + (N-1)(N-2)/2$$

Of these, $N(N-1)/2$ are rotation angles called quark mixing angles. The remaining $(N-1)(N-2)/2$ are complex phases, which cause CP violation.

For the case N = 2 families, there is only one parameter which is a mixing angle between two generations of quarks, the Cabibbo angle θ_C that was mentioned before. It is clear that no complex phase arises from this prescription and thus is incompatible with the observation of CP violation during the quark mixing.

For the Standard Model case with N = 3 families, there are three mixing angles θ_{12}, θ_{13}, θ_{23} and one irreducible CP-violating complex phase δ, hence is the first representation of quark mixing matrices that allows CP violation. This is how Kobayashi and Maskawa, extending Cabibbo's model, predicted a third quark generation before the c, b and t quarks were discovered.

[1] that is, a matrix v such that $VV^\dagger = I$, where $V^\dagger$ is the conjugate transpose of V and I is the identity matrix

Using the so-called *standard* representation of the three mixing angles θ_{ij} and the complex phase δ, the CKM matrix becomes [18]:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{1.4}$$

Where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ with $i, j = 1, 2, 3$. Indeed, the nullity of the angle θ_{ij} means no mixing among the families i and j. Therefore, when assuming only mixing between the families 1 and 2, the CKM matrix adopts the form of the Cabibbo matrix. Currently the best known values for the standard parameters are [18]:

$\theta_{12} = 13.04 \pm 0.05^\circ \quad \theta_{13} = 0.201 \pm 0.011^\circ \quad \theta_{23} = 2.38 \pm 0.06^\circ \quad \delta = 1.20 \pm 0.08\text{rad}$

1.2.3 Wolfenstein Parametrization and Unitary Triangle

Experimentally it is obtained that $s_{13} \ll s_{23} \ll s_{12} \ll 1$, and therefore is specially interesting to use a different parametrization proposed by L. Wolfenstein in 1983 [19]. This new parametrization introduces four new parameters λ, A, η, ρ that can be related to the standard mixing angles as follows:

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}} \rightarrow c_{12} = \sqrt{1 - \lambda^2} \tag{1.5}$$

$$s_{23} = A\lambda^2 = \lambda|\frac{V_{cd}}{V_{us}}| \rightarrow c_{23} = \sqrt{1 - A^2\lambda^4} \tag{1.6}$$

$$s_{13}e^{i\delta} = A\lambda^3(\rho + i\eta) \rightarrow c_{13} = \sqrt{1 - A^2\lambda^3(\rho^2 + \eta^2)} \tag{1.7}$$

The main idea behind is to express the CKM matrix as a low order expansion for small λ. After applying the relations in 1.5, this matrix becomes:

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4) \tag{1.8}$$

In Eq. 1.8 the sine $s_{12} = \lambda$ plays a role of an expansion parameter. The parameters A, ρ and η turn out to come from the first order contribution. From there it is also obvious that matrix elements V_{ub} and V_{td} contain an irreducible complex phase which is responsible for the CP violation.

From the unitarity condition of the CKM matrix, $V_{CKM}V_{CKM}^{\dagger} = \mathbf{1}$, one can extract the following set of conditions:

$$\sum_{k,i\neq j} V_{ik}V_{jk}^{*} = 0 \tag{1.9}$$

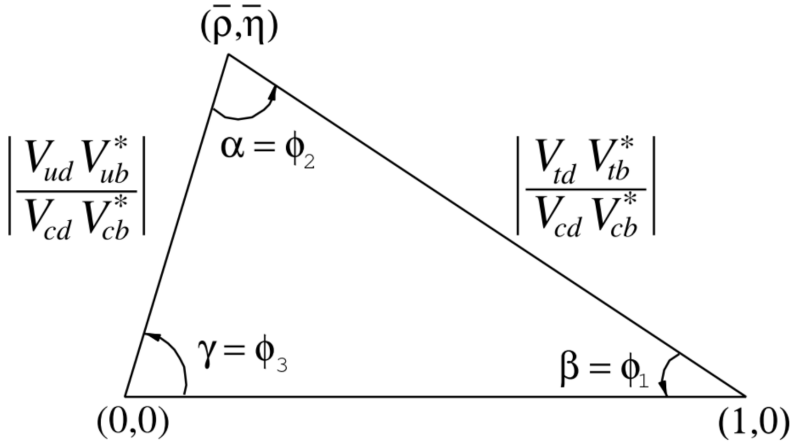


Figure 1.1: Representation of one of the unitary triangles in the complex plane that form the set of the remaining conditions on the CKM matrix. Note that the angles can be represented in two different forms: either in the BaBar notation (α, β, γ) or in the Belle notation (ϕ_1, ϕ_2, ϕ_3) [20]

.

For any fixed and different i and j, this is a constraint on three complex numbers, one for each k, which says that these numbers form the sides of a triangle in the complex plane, represented in the Fig. 1.1. There are six choices of i and j (three independent), and hence six such triangles, each of which is called a unitary triangle. Their shapes can be very different, but they all have the same area, which can be related to the CP violating phase. The area vanishes for the specific parameters in the Standard Model for which there would be no CP violation. The orientation of the triangles depend on the phases of the quark fields. Within the scope of this work the triangle described by Fig. 1.1 is more relevant, since it is related to processes involving B Mesons. Moreover, it corresponds to one of the two unitary triangles that have terms of equal order and thus sides of comparable length. Therefore, the interior angles are large enough to be precisely measured [21].

Measuring the decay of particular modes then contributes to the measurement of corresponding sides or angles of such a triangle. In case unitarity of V_{CKM} is fulfilled, meaning the SM's description of CP violation is correct and complete, then the triangle closes. Otherwise, deviations from unitarity would be a hint for New Physics. Since for defining a triangle it is sufficient to fix the value of either

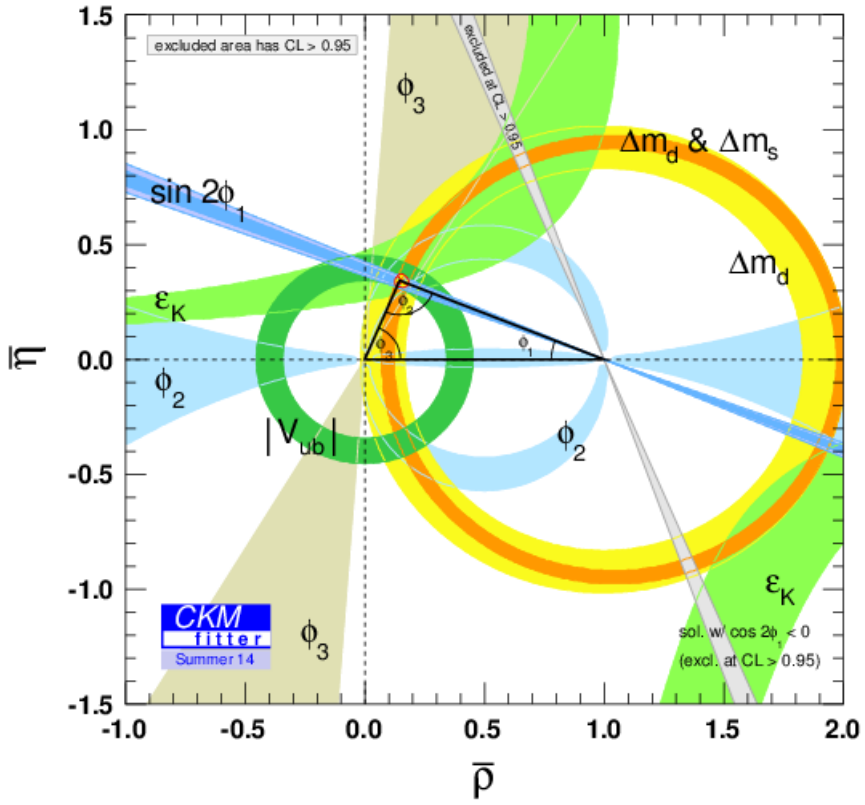


Figure 1.2: Exclusion plot for the unitary triangle of the CKM matrix in the $\bar{\rho}, \bar{\eta}$ plane [22]

.

two sides and an angle between them or two angles and one side or three sides, this problem is over-constrained from the experimental point of view as there are two sides and three angles, which can be measured. The results combined from many measurements are shown in Fig. 1.2 [22]. So far, they are well in agreement with the SM, but as the statistical uncertainties are big, there is still room for New Physics. Therefore, high precision measurements with very high statistics are needed. Belle II is expected to reach possible New Physics in the upcoming years.

1.2.4 Time Evolution of Neutral Mesons

Neutral meson have a characteristic quantum number depending on its quark composition, and that is called flavor of the meson. For example, K_0 contain *strangeness*, D_0 contain *charm* and B^0contain *beauty*. These quantum numbers

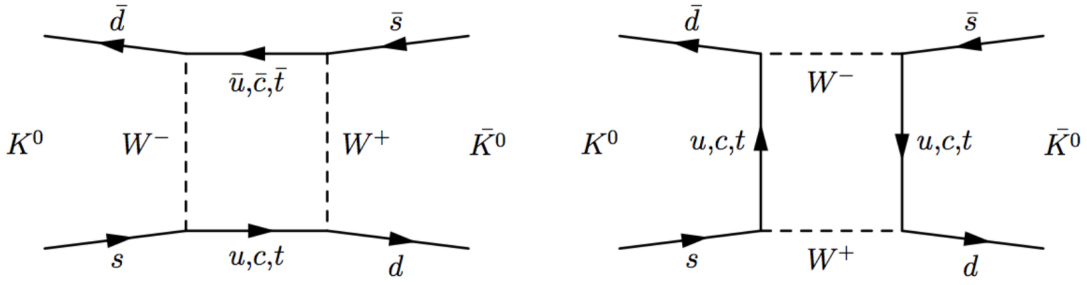


Figure 1.3: Example of neutral meson oscillation $\mathrm{K}^0 \leftrightarrow \bar{\mathrm{K}}^0$

.

are conserved in strong interactions and thus the flavor eigenstates of mesons are simultaneously eigenstates of strong interactions.

However this is not true for weak interactions, and therefore those mesons can oscillate between their particle and antiparticle states due to the flavor-changing weak interactions shown in the Feynman diagram in Fig. 1.3. This oscillation from matter to antimatter can be used to measure fundamental parameters of the SM and in addition to have even more striking effects, such as breaking the matter-antimatter symmetry in the Universe.

Mathematically, an oscillating neutral meson system can be described as a superposition of both particle and antiparticle states, i.e $\mathrm{B}^0$and $\bar{\mathrm{B}}^0$. The amplitudes of those states are time-dependant in an oscillating manner:

$$|\Psi\rangle(t) = \psi_1(t)|\mathrm{P}^0\rangle + \psi_2(t)|\bar{\mathrm{P}}^0\rangle = \begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix} \tag{1.10}$$

$$i\hbar\frac{\partial}{\partial t}\Psi = \mathcal{H}\Psi = (M - \frac{i}{2}\Gamma)\Psi \tag{1.11}$$

This matrix is not hermitian, otherwise the mesons would only oscillate but not decay [23]. M is a 2×2 mass matrix and Γ a 2×2 decay matrix. Both M and Γ are hermitian and therefore they satisfy $M_{12} = M_{21}^*$ and $\Gamma_{12} = \Gamma_{21}^*$.

The off-diagonal elements of the Hamiltonian represent flavour changing transitions $\mathrm{P}^0 \leftrightarrow \bar{\mathrm{P}}^0$, and the diagonal elements represent flavour conserving transitions. If CP is conserved, then $H_{12} = H_{21}^*$ and the oscillation rate is the same for both the particle and antiparticle states. In addition, the so called CPT theorem [24] requires that $M_{11} = M_{22} = M$ and $\Gamma_{11} = \Gamma_{22} = \Gamma$. This diagonal elements are determined by quark masses together with strong and electromagnetic (EM) interactions. Subsequently, the diagonal elements of $\mathcal{H}$ can be identified as the strong and EM Hamiltonian $\mathcal{H}_0$. Thus, $\mathcal{H}$ can be expressed in matricial form as:

$$\mathcal{H} = M - \frac{i}{2}\Gamma = \begin{pmatrix} M - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M^*_{12} - \frac{i}{2}\Gamma^*_{12} & M - \frac{i}{2}\Gamma \end{pmatrix} \tag{1.12}$$

One can diagonalise the matrix in 1.12 to obtain the CP eigenstates and its correspondent eigenvalues:

$$\omega_\pm = M - \frac{i}{2}\Gamma \pm \sqrt{(M_{12} - \frac{i}{2}\Gamma_{12})(M^*_{12} - \frac{i}{2}\Gamma^*_{12})} = M - \frac{i}{2}\Gamma \pm qp \tag{1.13}$$

And,

$$|P_{CP1}\rangle = q|\mathrm{P}^0\rangle + p|\bar{\mathrm{P}}^0\rangle \tag{1.14}$$

$$|P_{CP2}\rangle = q|\mathrm{P}^0\rangle - p|\bar{\mathrm{P}}^0\rangle \tag{1.15}$$

$$\tag{1.16}$$

Are the CP eigenstates. Here the parameters p,q are complex numbers satisfying:

$$\frac{q}{p} = \sqrt{\frac{M^*_{12} - \frac{i}{2}\Gamma^*_{12}}{M_{12} - \frac{i}{2}\Gamma_{12}}} \tag{1.17}$$

And, of course, the normalization condition $|q|^2 + |p|^2 = 1$. In case of CP invariance, the probability of one flavor being transformed into the other flavor is the same for both the neutral meson and its anti-meson. The conditions in the parameters is that both q, p are real and therefore $M_{12} = M^*_{12}$ and $\Gamma_{12} = \Gamma^*_{12}$

The masses $m_\pm$ and the lifetimes $\Gamma_\pm$ of the mass eigenstates can be obteined from the eigenvalues 1.13:

$$m_+ = \Re(\omega_+) \tag{1.18}$$

$$m_- = \Re(\omega_-) \tag{1.19}$$

$$\Gamma_+ = -2\Im(\omega_+) \tag{1.20}$$

$$\Gamma_- = -2\Im(\omega_-) \tag{1.21}$$

The time evolution of the neutral system can be easily written in terms of the mass eigenstates, that is, the ones diagonalizing the Hamiltonian, and its eigenvalues:

$$|P_{CP1}(t)\rangle = e^{-i\omega_+ t}(q|\mathrm{P}^0\rangle + p|\bar{\mathrm{P}}^0\rangle)$$
$$|P_{CP2}(t)\rangle = e^{-i\omega_- t}(q|\mathrm{P}^0\rangle - p|\bar{\mathrm{P}}^0\rangle) \tag{1.22}$$

But what one can observe in the production point is a well-defined flavor state. Thus, they should be described in terms of the mass eigenstates in order to observe its evolution:

$$|\mathrm{P}^0(t)\rangle = \frac{1}{2q}(|P_{CP1}(t)\rangle + |P_{CP2}(t)\rangle)$$
$$|\bar{\mathrm{P}}^0(t)\rangle = \frac{1}{2p}(|P_{CP1}(t)\rangle - |P_{CP2}(t)\rangle) \tag{1.23}$$

Now, one can introduce 1.22 into 1.23 to obtain the time-evolved flavor state in terms of the flavor basis:

$$|\mathrm{P}^0(t)\rangle = \frac{1}{2}\left((e^{-i\omega_+ t} + e^{-i\omega_- t})|\mathrm{P}^0(t)\rangle + \frac{q}{p}(e^{-i\omega_+ t} - e^{-i\omega_- t})|\bar{\mathrm{P}}^0(t)\rangle\right] \tag{1.24}$$
$$= \frac{1}{2}e^{-i\omega_+}\left[\left(1 + e^{(\Delta m - \frac{i}{2}\Delta\Gamma)t}\right)|\mathrm{P}^0\rangle + \frac{q}{p}\left(1 - e^{(\Delta m - \frac{i}{2}\Delta\Gamma)t}\right)|\mathrm{P}^0\rangle\right]$$

$$|\bar{\mathrm{P}}^0(t)\rangle = \frac{1}{2}\left(\frac{q}{p}(e^{-i\omega_+ t} - e^{-i\omega_- t})|\mathrm{P}^0(t)\rangle + (e^{-i\omega_+ t} + e^{-i\omega_- t})|\bar{\mathrm{P}}^0(t)\rangle\right] \tag{1.25}$$
$$= \frac{1}{2}e^{-i\omega_+}\left[\frac{q}{p}\left(1 - e^{(\Delta m - \frac{i}{2}\Delta\Gamma)t}\right)|\mathrm{P}^0\rangle + \left(1 + e^{(\Delta m - \frac{i}{2}\Delta\Gamma)t}\right)|\mathrm{P}^0\rangle\right]$$

Here $\Delta m = m_- - m_+$ and $\Delta\Gamma = \Gamma_- - \Gamma_+$ have been introduced as a measure of the decay and the interference of both flavor eigenstates throughout the time. At this point one can calculate the probability to observe $|\bar{\mathrm{P}}^0\rangle$ at a time t for a $|\mathrm{P}^0\rangle$:

$$|\langle\bar{\mathrm{P}}^0|\mathrm{P}^0(t)\rangle|^2 = \frac{1}{4}e^{-2i\Gamma_+ t}\left|\frac{q}{p}\right|^2 |1 - e^{-i(\Delta m - \frac{i}{2}\Delta\Gamma)t}|^2 \tag{1.26}$$

$$= \frac{1}{4}e^{-2i\Gamma_+ t}\left|\frac{q}{p}\right|^2 \left(1 - \cos\left(\left[\Delta m - \frac{i}{2}\Delta\Gamma\right]t\right)\right)^2 \tag{1.27}$$

If CP is conserved, both probabilities $P_{\mathrm{P}^0\to\bar{\mathrm{P}}^0}$ and $P_{\bar{\mathrm{P}}^0\to\mathrm{P}^0}$ would be the same since $q = p$. It is common to call $|P_{CP1,CP2}\rangle$ as $|P_{H,L}\rangle$ where H and L stand for the heavier and lighter state.

1.2.5 Classification of CP Violation

Consider now the amplitudes $\mathcal{A}_f$,$\mathcal{A}_{\bar{f}}$ ($\bar{\mathcal{A}}_f$,$\bar{\mathcal{A}}_{\bar{f}}$) of a initial state $\mathrm{P}^0(\bar{\mathrm{P}^0})$ decaying into a final state $f, \bar{f}$, given by:

$$\mathcal{A}_f = \langle f|\mathcal{H}\mathrm{P}^0\rangle \quad \bar{\mathcal{A}}_f = \langle f|\mathcal{H}\bar{\mathrm{P}}^0\rangle \quad \mathcal{A}_{\bar{f}} = \langle \bar{f}|\mathcal{H}\mathrm{P}^0\rangle \quad \bar{\mathcal{A}}_{\bar{f}} = \langle \bar{f}|\mathcal{H}\bar{\mathrm{P}}^0\rangle \tag{1.28}$$

where the time-evolution is directed by the Hamiltonian. In order to discuss the types of CP violations, a new parameter needs to be introduced to acquaint for the ratio between decay amplitudes $\mathcal{A}_f, \bar{\mathcal{A}}_f$ and mixing coefficients q, p:

$$\lambda_{CP} = \frac{q}{p}\frac{\bar{\mathcal{A}}_f}{\mathcal{A}_f} = e^{-i\phi_D}\left|\frac{q}{p}\right|\left|\frac{\bar{\mathcal{A}}_f}{\mathcal{A}_f}\right|e^{-i\phi_M} \tag{1.29}$$

with a phase difference for the mixing, ϕ_M and decay amplitude ϕ_D. From a measurement of time-dependent asymmetries, both the phase and the absolute value of λ_{CP} can be determined.

CP violation in decay

In the case when CP violation in decay occurs, the decay amplitudes for a particle and its antiparticle in a certain final state or conjugated final state differ from one another. An example of this type of CP violation, also known as direct CP violation, is shown in Fig.1.4 and is defined by:

$$\left|\frac{\bar{\mathcal{A}}_{\bar{f}}}{\mathcal{A}_f}\right| \neq 1 \tag{1.30}$$

This means that the two CP conjugate states have different absolute values for their decay amplitudes. The phase difference between the decay amplitudes entailed in 1.30 can be quantified by the complex phase ϕ_D introduced in 1.29. This kind of process is possible for charged and for neutral mesons. In charged meson decays where no mixing effects are involved (see next paragraph), this type of CP violation is the only possible source of CP asymmetries and is defined as:

$$\mathcal{A}_{CP} = \frac{\Gamma(P^- \to f) - \Gamma(\bar{P}^+ \to \bar{f})}{\Gamma(P^- \to f) + \Gamma(\bar{P}^+ \to \bar{f})}$$

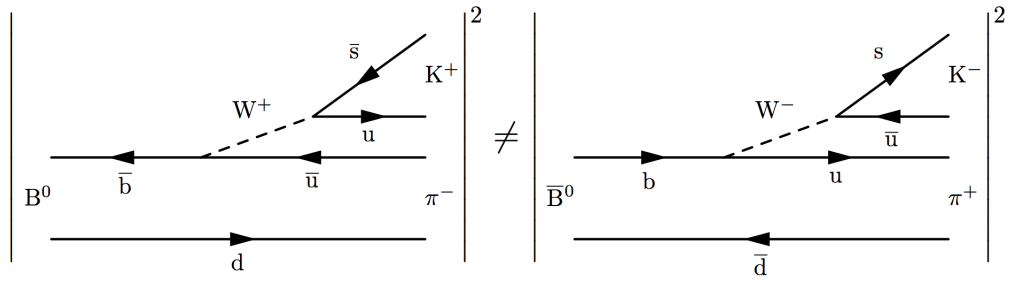


Figure 1.4: Example of CPV in decay of the B meson into a Kaon. The amplitude of the Feynman diagram to the right and the corresponding to its antiparticle decay to the left are not the same.

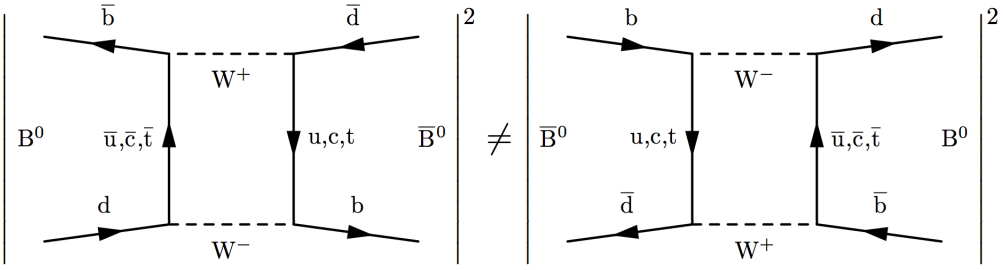


Figure 1.5: Example of CPV in mixing of $B^0 \leftrightarrow \bar{B}^0$. The amplitude of $B^0 \rightarrow \bar{B}^0$ is not equivalent to the amplitude of the reverse process $B^0 \leftarrow \bar{B}^0$.

CP violation in mixing

In the case of neutral mesons, flavor specific final states can reveal CP violation via $P^0 \leftrightarrow \bar{P}^0$ oscillations. Flavor specific decays are those that can come either from the neutral meson or from its antiparticle, but not from both:

The violation of CP in the neutral meson system mixing is depicted in Fig.1.5 and is described by

$$\left|\frac{q}{p}\right|$$

This kind of CP violation is only observable for neutral mesons. It originates from the fact that no such choice of phase convention exist in order to make the mass eigenstates identical to the CP eigenstates. Thus, there is an asymmetry in the flavor oscillation $|P^0\rangle \rightarrow |\bar{P}^0\rangle \rightarrow |P^0\rangle$ Regarding the charged mesons, mixing between charged particles is not allowed due to charge conservation.

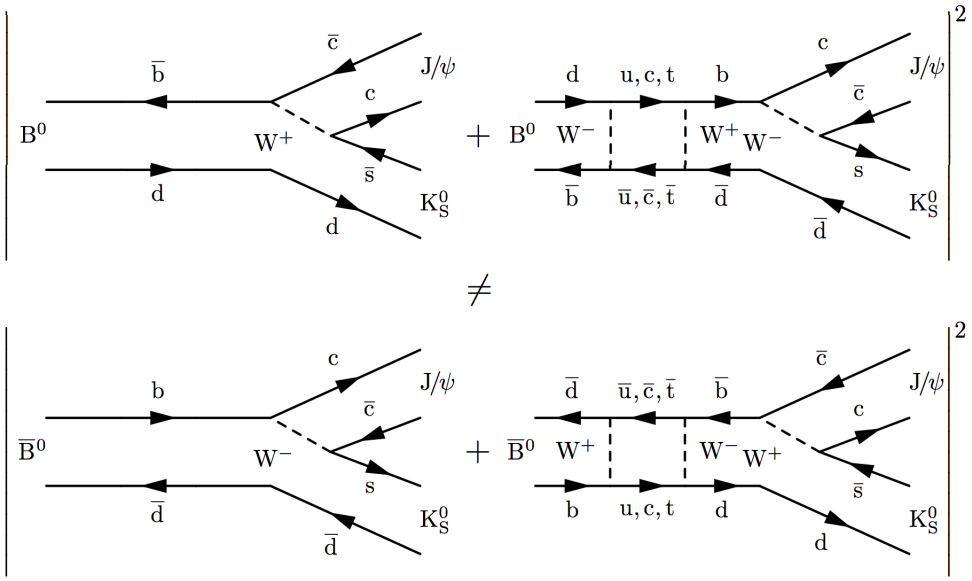


Figure 1.6: Example of CPV by interference of mixing and decay. In the Feynman diagrams to the right it is shown the direct decay into a CP eigenstate. To the left, the diagrams show a transition $B^0 \leftrightarrow \bar{B}^0$ before decaying to the very same CP eigenstate. The interference between the two diagrams is different for $B^0$and $\bar{B}^0$as initial state.

CP violation by interference of mixing and decay

CP violation can also appear when the particle and the antiparticle decay into a common final state, $P^0 \to f$ and $\bar{P}^0 \to f$. In order for this to happen the final state must be a CP eigenstate, $f = f_{CP}$. Even if there is no CP violation in mixing and decay respectively, i.e. if $\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| = \left|\frac{q}{p}\right| = 1$, the superposition of ϕ_D and ϕ_M can produce a phase difference and thus an interference between these two processes, generating consequently a violation of CP symmetry. Hence, the CP violation in interference between decay $P^0 \to f$ and mixing $P^0 \to \bar{P}^0 \to f$ where different states to common final CP eigenstate is defined by:

$$\Im\lambda_{CP} \neq 0 \to \phi_D + \phi_M \neq 0 \tag{1.31}$$

An example of the processes involving this type of CP violation is showed in Fig. 1.6.

This phenomenon was observed for the first time by the Belle [25] as well as by the BaBar [26] collaboration within the so called golden channel $B \to J/\psi K_S$ and related decay modes. This exact channel will be the one used for the studies performed in this thesis.

1.3 CP Violation in the B sector

As explained in the Sec.1.2.2, the CKM mechanism predicts the existence of a third generation of quarks if the CP symmetry is to be violated. The bottom quark was discovered in 1977 by the Fermilab E288 experiment team led by Leon M. Lederman [27].
Out from the bottom flavored quarks, B mesons can be created. Those mesons are composed of a bottom antiquark and either an up (B^+), down (B^0), strange (B_s^0) or charm quark (B_c^+), and in the same manner can be obtained the anti B mesons just changing the quarks involved for their own antiparticles. The bottom quark-antiquark pair is not treated as a B meson but rather as quark resonance called bottomnium. The combination of a bottom antiquark and a top quark is thought to be impossible because of the top quark's short lifetime. B mesons are produced via the strong interaction and decay weakly. The fact that the strong eigenstates do not correspond to the weak eigenstates is the reason of the quark mixing induced by the CKM matrix. Their mass $m_B \sim 5.279\,\mathrm{GeV}/c^2$ [28] is large due to the large mass of the b quark ($m_b \sim 4.18\,\mathrm{GeV}/c^2$ [28]).

The B meson system is very suitable for the study of CP violation. One reason for this is that CP violation is large in B decays. Another reason is that the B meson's lifetime is large because it cannot decay to a top quark due to its huge mass $m_t \sim 170\,\mathrm{MeV}$ and because all other decays are suppressed with at least the modulus of the matrix element $|V_{cb}| \sim \lambda^2$ (cf. Eq.1.8). With the exception of the heavier $B_{c\pm}$ mesons with a lifetime of $\tau \sim 0.5\,\mathrm{ps}$, all other B mesons have lifetimes around 1.5 ps, e.g. $\tau_{B^0} = (1.525 \pm 0.009) \times 10^{-12}$ s. This offers good conditions for the study of $B^0 \leftrightarrow \bar{B}^0$ oscillations as well as the B lifetime itself. The dominant decays occur via b $\rightarrow$ c transitions (such as $B^0 \rightarrow J/\psi K_S$), whereas those over b $\rightarrow$ u, d, s (such as $B^0 \rightarrow \omega K_S$) are smaller and referred to as rare decays. In the CKM matrix this is described by their small corresponding matrix element.

1.3.1 Time Evolution of Neutral B mesons

The neutral B mesons show some basic differences from the neutral kaons. Firstly, unlike the two CP eigenstates of the kaon K_S^0 and K_L^0, the eigenstates of the B meson B_H and B_L have almost equal lifetimes:

$$\frac{|\Delta\Gamma_B|}{\Gamma_B} = -0.001 \pm 0.014 \tag{1.32}$$

Therefore, we can simply take $\Delta\Gamma = 0$ in eqs. 1.25 and 1.24. Secondly, Another important difference is that the box diagram of $B^0 \leftrightarrow \bar{B}^0$ mixing (Fig. 1.5) is dominated by virtual t-quark exchange because $V_{tb} \simeq 1$ is the dominant CKM

element , i.e. $V_{tb} \simeq 1 \gg V_{cb} \gg V_{ub}$. Therefore, the eigenstates B_H and B_L can be considered as pure flavour eigenstates in a good approximation, and it follows from Eq. 1.17 that:

$$\left.\frac{q}{p}\right|_{B_d} = \sqrt{\frac{M_{12}^*}{M_{12}}} + \mathcal{O}\left(\frac{\Gamma_{12}}{M_{12}}\right)\Bigg|_{B_d} \simeq \frac{V_{tb}^* V_{td}}{V_{td}^* V_{tb}} \tag{1.33}$$

The term depending on $\dfrac{\Gamma_{12}}{M_{12}} \sim \dfrac{m_b^2}{m_t^2} \approx 10^{-2}$ [29] and can be neglected. By inserting this term into Eq. 1.29 one obtains:

$$\lambda_{CP} = e^{i\phi_M}\frac{\bar{\mathcal{A}}}{\mathcal{A}_f} = \frac{V_{tb}^* V_{td}}{V_{td}^* V_{tb}}\frac{\bar{\mathcal{A}}}{\mathcal{A}_f} \tag{1.34}$$

hence λ_{CP} can still have a non-trivial complex phase and mixing-induced CP violation is still possible. Thus, the time evolution of neutral B mesons, starting as a pure flavour state, is given by,

$$|\mathrm{B}^0(t)\rangle = \frac{1}{2}e^{-\Gamma t/2}\left[\left(1+e^{-i\Delta m t}\right)|\mathrm{B}^0\rangle + \frac{q}{p}\left(1-e^{-i\Delta m t}\right)|\mathrm{B}^0\rangle\right] \tag{1.35}$$

$$|\bar{\mathrm{B}}^0(t)\rangle = \frac{1}{2}e^{-\Gamma t/2}\left[\frac{q}{p}\left(1-e^{-i\Delta m t}\right)|\mathrm{B}^0\rangle + \left(1+e^{-i\Delta m t}\right)|\mathrm{B}^0\rangle\right] \tag{1.36}$$

where $e^{i\omega_+}$ is removed by a phase convention. When the initial state is a pure B^0, then the probabilities of observing a B^0 or a $\bar{\mathrm{B}}^0$ after some time t, is (compare with Eq. 1.26)

$$|\langle\bar{\mathrm{B}}^0|\mathrm{B}^0(t)\rangle|^2 = \frac{1}{2}e^{-\Gamma_+ t}\left(1+\cos\left(\Delta m t\right)\right)^2 \tag{1.37}$$

$$|\langle\mathrm{B}^0|\bar{\mathrm{B}}^0(t)\rangle|^2 = \frac{1}{2}e^{-\Gamma_+ t}\left(1-\cos\left(\Delta m t\right)\right)^2 \tag{1.38}$$

In Fig. 1.7, the probabilities above have been plotted using values from experimental results of the Belle experiment [30], $\Delta m_{\mathrm{B}^0}/\tau_{\mathrm{B}^0} = 0.774 \pm 0.008$ [28]. As it can be seen there, the most B^0 mesons decay before they change flavour, but there is a non vanishing probability that an initial B^0 develops into a pure $\bar{\mathrm{B}}^0$. In the neutral K meson system an initial K^0 will never convert into a pure $\bar{\mathrm{K}}^0$.

1.3.2 Time-Dependent CP Violation

In this section, we want to introduce the time-dependent CP rate asymmetry a_{CP} as well as the two CP observables $\mathcal{A}_{CP}$ and $\mathcal{S}_{CP}$.
The amplitudes of a $\mathrm{B}^0$and a $\bar{\mathrm{B}}^0$decay into a given final state f_{CP} are given by,

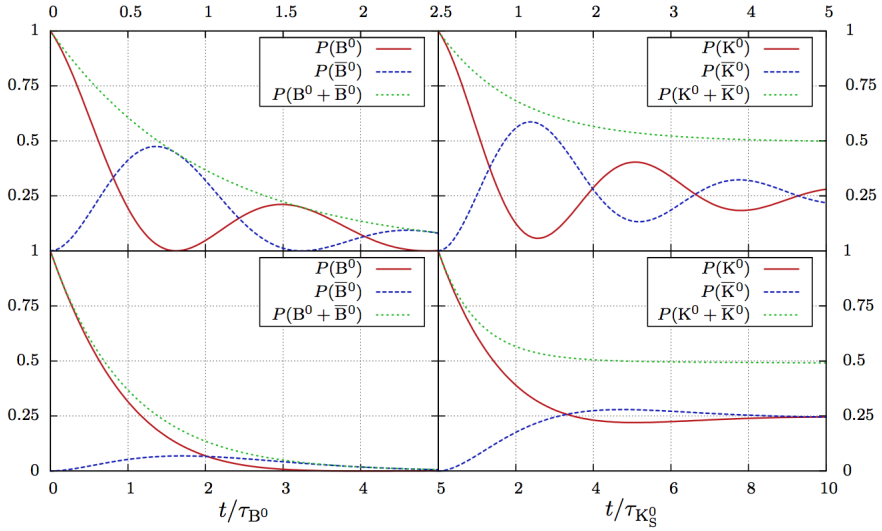


Figure 1.7: Mixing probabilities for initial pure B^0(left) and K^0 (right) mesons as functions of their lifetimes. The bottom row shows the probabilities using the values $\Delta m/\tau$ from experimental results. The top row shows the probabilities for a Δm 5 times the measured value [30].

$$A_{CP} = \langle f_{CP}|\mathcal{H}\mathrm{B}^0\rangle, \quad \bar{A}_{CP} = \langle f_{CP}|\mathcal{H}\bar{\mathrm{B}}^0\rangle \tag{1.39}$$

Using the time evolution of a neutral meson eigenstate in Eqs. 1.24 and 1.25, the time dependent decay amplitudes can be expressed as:

$$\begin{aligned} A_{CP}(t) = \langle f_{CP}|\mathcal{H}\mathrm{B}^0(t)\rangle &= \frac{1}{2}e^{-\Gamma t/2}\left[\left(1+e^{-i\Delta m t}\right)A_{CP} + \frac{q}{p}\left(1-e^{-i\Delta m t}\right)\bar{A}_{CP}\right] \\ &= \frac{1}{2}e^{-\Gamma t/2}A_{CP}\left[\left(1+e^{-i\Delta m t}\right) + \lambda_{CP}\left(1-e^{-i\Delta m t}\right)\right], \end{aligned} \tag{1.40}$$

$$\begin{aligned} \bar{A}_{CP}(t) = \langle f_{CP}|\mathcal{H}\bar{\mathrm{B}}^0(t)\rangle &= \frac{1}{2}e^{-\Gamma t/2}\left[\frac{q}{p}\left(1-e^{-i\Delta m t}\right)A_{CP} + \left(1+e^{-i\Delta m t}\right)\bar{A}_{CP}\right] \\ &= \frac{1}{2}e^{-\Gamma t/2}A_{CP}\left[\left(1-e^{-i\Delta m t}\right) + \lambda_{CP}\left(1+e^{-i\Delta m t}\right)\right] \end{aligned} \tag{1.41}$$

Where $\langle f_{CP}|\mathrm{B}^0\rangle = 0$ and $\langle f_{CP}|\bar{\mathrm{B}}^0\rangle = 0$ is assumed. The squares of the decay amplitudes give the time dependent rate:

$$\Gamma(\mathrm{B}^0 \to f_{CP}) = |\langle f_{CP}|\mathcal{H}\mathrm{B}^0(t)\rangle|^2 \tag{1.42}$$
$$= \frac{1}{2}e^{-\Gamma t}|A_{CP}|[1+\cos(\Delta mt) - 2\Im(\lambda_{CP}\sin(\Delta mt) + (1-\cos(\Delta mt))|\lambda_{CP}|^2],$$

$$\Gamma(\bar{\mathrm{B}}^0 \to f_{CP}) = |\langle f_{CP}|\mathcal{H}\mathrm{B}^0\rangle|^2(t) \tag{1.43}$$
$$= \frac{1}{2}e^{-\Gamma t}|A_{CP}|[1+\cos(\Delta mt) - 2\Im(\lambda_{CP}\sin(\Delta mt) + (1-\cos(\Delta mt))|\lambda_{CP}|^2]$$

The time dependent CP rate asymmetry is defined as:

$$\begin{aligned} a_{CP}(t) &\equiv \frac{\Gamma(\bar{\mathrm{B}}^0 \to f_{CP}) - \Gamma(\mathrm{B}^0 \to f_{CP})}{\Gamma(\bar{\mathrm{B}}^0 \to f_{CP}) + \Gamma(\mathrm{B}^0 \to f_{CP})} \\ &= \frac{(|\lambda_{CP}|^2 - 1)\cos(\Delta mt) + 2\Im\lambda_{CP})\sin(\Delta mt)}{1+|\lambda_{CP}|^2} \\ &= \mathcal{A}_{CP}\cos(\Delta mt) + \mathcal{S}_{CP}\sin(\Delta mt) \end{aligned} \tag{1.44}$$

$$\mathcal{A}_{CP} \equiv \frac{|\lambda_{CP}|^2 - 1}{|\lambda_{CP}|^2 - 1} \quad \text{and} \quad \mathcal{S}_{CP} \equiv \frac{2\Im(\lambda_{CP})}{|\lambda_{CP}|^2 + 1} \tag{1.45}$$

In case that $|\lambda_{CP}|^2 \neq 1, \mathcal{A}_{CP} \neq 0$. This means that the observable $\mathcal{A}_{CP}$ is a measure amount of direct CP violation. $\mathcal{S}_{CP} \neq 0$ reveals the presence of mixing-induced CP violation.
To access those parameters and the physics behind them, one needs to measure the time-dependent rates of $\mathrm{B}^0$and also $\bar{\mathrm{B}}^0$decaying into f_{CP}. In the case of rare decays, this is quite a challenge for experimentalists and this is why high precision measurements are of an essential importance.

1.3.3 Time-Dependent CP Violation Measurement

Due to the heaviness of the B mesons, there are many decay modes available for them, and the branching fractions for the modes usable for CP violation are generally small. Therefore, a great amount of B meson is necessary to obtain significant measurements of CP violation. Such high precision measurements were provided by the machines PEP II at SLAC [2] and KEKB at KEK [1] [2]. These machines enable the required conditions for performing an experimental determination of

[2]After the massive upgrade of the KEKB facilities is called SuperKEKB. The new accelerator will bring 40 times more luminosity than the record established by its predecessor

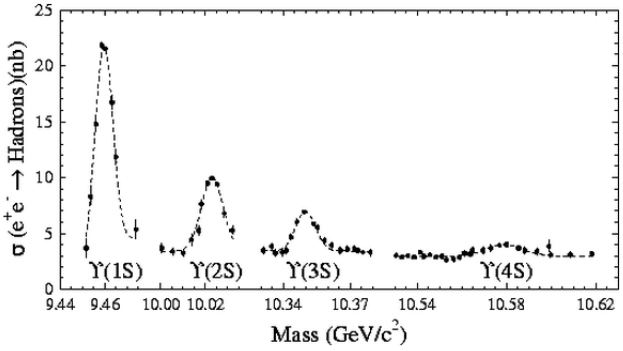


Figure 1.8: Hadronic cross section for e^+e^- collisions as a function of the e^+e^- center-of-mass energy in the region at and above the Υ(1S) resonance. At the center-of-mass energy of 10.56 GeV is located the production threshold of the pair $B\bar{B}$, right before the Υ(4S) resonance creation point [31].

the angles and sides of the unitarity triangle using B meson decays.

At the B meson factories $B\bar{B}$ pairs are produced at the Υ(4S) resonance, roughly at 10.58 GeV (Fig. 1.8). This the lowest energy at which B mesons can be produced in an e^+e^- collider. Since the mass of both B mesons approximately corresponds to the mass of the resonance, the formers are essentially produced at rest in the center-of- mass system (CMS). This fact forces the B factories to create the beams energetically asymmetric so that an effective boost is created in the CMS. The Υ(4S) decays immediately into $B\bar{B}$ pairs more that 95% of the times. The fractions of produced charged and neutral pairs is almost identical, namely B^+B^- $(51.3 \pm 0.6)\%$ and $B^0\bar{B}^0$ (48.7 ± 0.6) % , respectively [28].

About the quantum states created after the decaying of the resonance, there is some discussion that must be done. On one hand, the Υ(4S) has the quantum numbers $J^{PC} = 1^{--}$, while the produced B mesons are pseudo-scalars ($J^P = 0^-$). It is clear that in order to conserve angular momentum, the two pseudo scalar B mesons must be produced in p-wave, that is $\ell = 1$. Therefore they are not independent particles.

On the other hand, as it was explained before, the B mesons are produced via strong interactions, which means that C and P quantum numbers must be conserved. This leaves some conditions on the form of the wavefunction of the $B\bar{B}$ pair. The wave function of Υ(4S) is anti-symmetric (negative parity) and, accordingly, the state of the $B\bar{B}$ system has to be anti-symmetric too. This condition forbids the totally symmetric states B^0B^0 and $\bar{B}^0\bar{B}^0$, which must be symmetric by

reason of Bose-Einstein statistics. If the two B mesons in the system are allowed to oscillate independently, there is a non vanishing probability to find the forbidden $B\bar{B}$ and $\bar{\mathrm{B}}^0\bar{\mathrm{B}}^0$ at a given time. In consequence, the $\mathrm{B}^0\bar{\mathrm{B}}^0$, pair must oscillate coherently.

Finally, the quantum mechanical entangled state is given by:

$$|B_1(t_1), B_2(t_2)\rangle = \frac{1}{2}\left[|\mathrm{B}_1^0(t_1); \bar{\mathrm{B}}_2^0(t_2)\rangle - |\bar{\mathrm{B}}_1^0(t_1); \mathrm{B}_2^0(t_2)\rangle\right] \tag{1.46}$$

Due to the quantum entanglement between both B mesons, If one of them, called $\mathrm{B_{tag}}$, decays into a flavour specific mode at a time t_1, the other B's flavour is fixed to the opposite at this time. Once $\mathrm{B_{tag}}$ has decayed, the other meson is free to evolve until it decays at a time t_2. Consequently, during the time between the two decays $\Delta\mathrm{t} = t_2 - t_1$ the B meson is treated as an independent particle. This parameter, together with the information of the flavour, can then be used for the extraction of the CP asymmetry parameters. One can reparametrize the time evolution of the B meson states described in Eqs. 1.25, 1.24 so that the CP asymmetry (cf. Eq. 1.44) incorpores now a dependece in Δt:

$$a_{CP} = \mathcal{A}_{CP}\cos(\Delta m \Delta\mathrm{t})\mathcal{S}_{CP}\sin(\Delta m \Delta\mathrm{t}) \tag{1.47}$$

By normalizing Eqs. 1.42 and 1.43, the probability of observing a given final state f_{CP} after observing flavour q, of $\mathrm{B_{tag}}$ at time Δt, is given by,

$$\mathcal{P}(\Delta t, q) = \frac{e^{\frac{-|\Delta\mathrm{t}|}{\tau_{\mathrm{B}^0}}}}{4\tau_{\mathrm{B}^0}}\left[1 + q(\mathcal{A}_{CP}\cos(\Delta m \Delta\mathrm{t})\mathcal{S}_{CP}\sin(\Delta m \Delta\mathrm{t}))\right] \tag{1.48}$$

where $\mathrm{q} = +1(-1)$ for $\mathrm{B_{tag}} = \mathrm{B}^0(\bar{\mathrm{B}}^0)$ and τ_{B^0} is the B^0 lifetime.

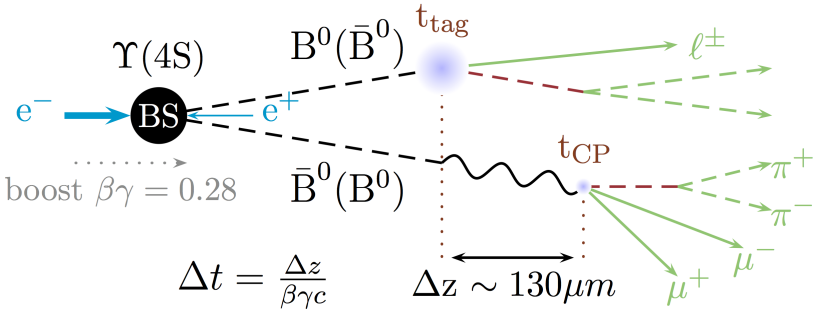


Figure 1.9: Decay scheme of both B mesons in the time-dependent analysis scheme.

Although the lifetime of the B mesons is relatively large, it is impossible to measure its value directly as precise as needed for the study of CP violation,

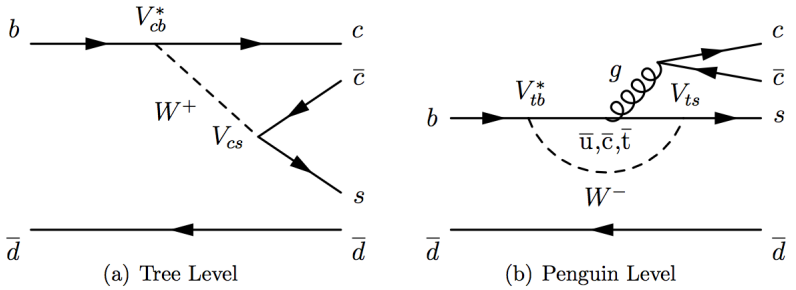


Figure 1.10: Tree-level (a) and penguin diagrams (b) contributing to $B^0 \to J/\psi K_S$ decays.

since a resolution in order of some pico seconds cannot be reached with current technology. A clever experimental trick in B-factories gets rid of this drawback by using asymmetric beam energies: the $\Upsilon(4S)$ system is produced with a Lorentz boost $\beta\gamma$ in the direction of the high energy beam, i.e. in positive z-direction if the convention of Belle is used. Because of the low momentum of the B mesons in the $\Upsilon(4S)$ center-of-mass frame, they fly practically parallel to the boost direction. Thus, the Lorentz boost amplifies the separation Δz between decay vertices and makes it detectable. The proper decay time difference is then,

$$\Delta t = \frac{\Delta z}{\beta\gamma c} \tag{1.49}$$

For a boost of $\beta\gamma \simeq 0.284$ in SuperKEKB, the mean flight distance of the B^0amounts only 130 µm. Such a spatial resolution for vertexing can be only achieved using high precision vertex detectors like those integrated in Belle II.

1.3.4 $\bar{b} \to \bar{c}cs$ Transitions and the Golden Mode

In the B meson system CP violation has been observed for the first time in $b \to c\bar{c}s$ transitions such as in $B^0 \to J/\psi K_S$ decay [25], [26]. This decay is commonly referred to as the *golden mode* for different reasons [9]. Experimentally, its signature comes from the decay of the J/ψ into two muons, which makes it one of the cleanest signals possible. Theoretically, this decay corresponds to a CP eigenstate which allows measurements of CP violation as explained in Sec. 1.3.3. More concretely allows to the determination of the angle ϕ_1 of the Unitarity Triangle (cf. Fig. 1.1). In $B^0 \to J/\psi K_S$ decay, the J/ψ and the K_S are produced with orbital angular momentum of $\ell = 1$ in a p-wave configuration. Neglecting the small CP violation in $\bar{K}^0 \leftrightarrow K^0$ mixing, the K_S has like the J/ψ a CP eigenvalue of +1. In combination with the $(-1)^{\ell=1} = -1$ contribution from the angular momentum, the $J/\psi K_S$ final state has the CP eigenvalue $\eta_{CP} = -1$.

The tree-level and penguin diagrams contributing to $B^0 \to J/\psi K_S$ decays are

shown in Fig. 1.10. On tree-level, the decay is mediated by $b \to c\bar{c}s$ transitions that are color-suppressed due to the internal emission of a W boson. The penguin diagrams have an internal loop with virtual up-type quarks. The loops involving virtual $\bar{c}$ and $\bar{t}$ quarks have CKM elements of the same order as the transitions in the tree diagram and carry the same weak phase. In contrast, the loop involving virtual $\bar{u}$ quarks, that could introduce a different weak phase via V_{ub}, is suppressed by a factor of λ^2, where $\lambda \approx 0.23$ refers to the expansion parameter in the Wolfenstein parameterisation (cf. Eq. 1.5). In the penguin diagrams, the $c\bar{c}$ pair forming the J/ψ is created from gluons. Since the $c\bar{c}$ pair has to be created in a color singlet state, it cannot be created from a single gluon. All above effects result in a high suppression of penguin amplitudes. No so-called penguin pollution and hence no direct CP violation are expected in $b \to c\bar{c}s$ transitions.

The mixing-induced CP violation in $\mathrm{B}^0 \to J/\psi K_S$ decays can be estimated from the involved decay amplitudes and mixing-phases. The parameter $\lambda_{CP}^{c\bar{c}\mathrm{K}^0}$ defined in Eq. 1.29 characterises the mixing-induced CP violation explained in Sec.2.4.3. It depends on the ratio q/p which introduces the B mixing phase and the amplitudes of the CP conjugated decays. Since B^0mesons decay to $J/\psi\mathrm{K}^0$, but $\bar{\mathrm{B}}^0$mesons decay to $J/\psi\bar{\mathrm{K}}^0$, the $J/\psi K_S$ final state involves $\bar{\mathrm{K}}^0 \leftrightarrow \mathrm{K}^0$ mixing and $\lambda_{CP}^{c\bar{c}\mathrm{K}^0}$ includes also the neutral kaon mixing phase. In consequence, the CP violating complex quantity $\lambda_{CP}^{c\bar{c}\mathrm{K}^0}$ becomes a product of three phases:

$$\begin{aligned}\lambda_{CP}^{c\bar{c}\mathrm{K}^0} &= -\left(\frac{q}{p}\right)_{\mathrm{K}^0} \cdot \frac{A_{(c\bar{c})\mathrm{K}^0}}{\bar{A}_{(c\bar{c})\mathrm{K}^0}} \cdot \left(\frac{q}{p}\right)_{\mathrm{B}^0} \simeq (-1)\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right)^{-1}\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right)^* \\ &= -(e^{i\phi_i})^{-1}(e^{i\phi_i})^* = e^{2i\phi_1} \end{aligned} \tag{1.50}$$

It follows then for the CP violating parameters that,

$$\mathcal{A}_{CP} = 0, \quad \mathcal{S}_{CP} = -\sin(2\phi_1) \tag{1.51}$$

In 2002 Belle observed this kind of time-dependent CP violation [25] after reconstruction of neutral B decays into the odd ($\eta_{CP} = -1$) CP eigenstates $J/\psi K_S$, $\psi(2S)K_S$, $\chi_{c1}K_S$, $\eta_c K_S$ and the even ($\eta_{CP} = +1$) $J/\psi K_L$. The actual values obtained by the Belle collaboration for the CP violating parameters are [32]:

$$\mathcal{A}_{CP}^{c\bar{c}\mathrm{K}^0} = 0.006 \pm 0.016(\mathrm{stat}) \pm 0.012(\mathrm{sys}) \tag{1.52}$$

$$\tag{1.53}$$

$$\sin(2\phi_1)^{c\bar{c}\mathrm{K}^0} = 0.667 \pm 0.023(\mathrm{stat}) \pm 0.012(\mathrm{sys}) \tag{1.54}$$

CHAPTER 2

Belle II Experiment

This chapter will give an overview of the Belle II experiment, especially of the tracking system. Belle II is an upgrade of the first Belle experiment, which was shut down in 2010. At this point, Belle II is still under construction, with a supposed delayed beginning of operation in 2018. It might be that the final experimental setup differs from what is specified here, however these changes will for sure be small. Most of the information contained in this chapter was taken from the Belle II Technical Design Report [6].
In the first place the massive upgrade Firstly, a review of the massive upgrade made to the KEKB accelerator will be done. Some details and specifications of the machine will be given together with a description of the new Nano-beam scheme. Secondly the Belle II detector will be treated. More concretely, the newly developed tracking system and the reasons behind its development.

2.1 SuperKEKB Accelerator and IR design

As argued in Sec. 1.3.3, the SuperKEKB factory uses asymmetric e^+e^- beam energies. The accelerator will be placed in the same tunnel as KEKB in the locality of Tsukuba, Japan. The positrons will be accelerated up to 4 GeV in the so called Low Energy Ring (LER) and the electrons up to 7 GeV in the High Energy Ring (HER), creating an asymmetry on the energies that will lead to an effective boost in the lab frame. The two beams intersect each other at the so-called Interaction Point (IP) by a relatively large crossing angle $2\phi = 83$ mrad (about 4 times larger than in KEKB). The 2m region around the IP is denominated as the interaction region (IR) (see Fig. 2.1). SuperKEKB is designed to have a 40 times higher luminosity than its predecessor, with $\mathcal{L} = 8 \times 10^{35}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. This will be achieved by adopting the so-called Nano-Beam scheme.

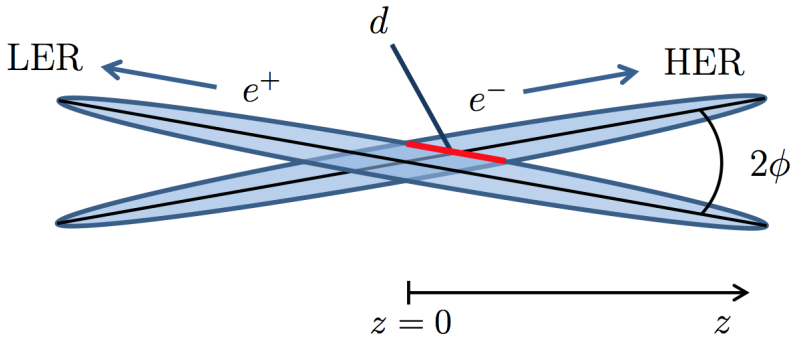


Figure 2.1: Schematic view of the beam intersection in the Nano Beam scheme [33].

2.1.1 Nano-beam scheme

The basic idea of this scheme is to squeeze the vertical beta function at the IP (β_y^*)[1] by minimizing the longitudinal size of the overlap region of the two beams at the IP. Fig. 2.1 shows a schematic view of the beam collision, which is a plane figure, in the Nano-Beam scheme. The size of the overlap region d, which is considered to be the effective bunch length for the Nano-Beam scheme, is much smaller than the bunch length (σ_z).

To shorten the length d, a relatively large horizontal crossing angle and extremely small horizontal emittances and horizontal beta functions at the IP for both beams are required. The luminosity of a collider is expressed by the following formula, assuming flat beams and equal horizontal and vertical beam sizes for two beams at the IP:

$$\mathcal{L} = \frac{\gamma_\pm}{2er_e}\frac{I_\pm \xi_{y\pm}}{\beta^*_{y\pm}}\frac{R_L}{R_{\xi_y}} \tag{2.1}$$

where γ, e and r_e are the Lorentz factor, the elementary electric charge and the electron classical radius, respectively. The suffix $\pm$ specifies the positron (+) or the electron (-). The parameters R_L and R_{ξ_y} represent reduction factors for the luminosity and the vertical beam-beam parameter, which arise from the crossing angle. The ratio of these parameters is usually not far from unity, and therefore, the luminosity is mainly determined by the three fundamental parameters; i.e. the total beam current (I), the vertical beam-beam parameter (ξ_y)) and the vertical beta function at the IP (β_y^*).

[1]The beta function in accelerator physics is a function related to the transverse size of the particle beam at the location s along the nominal beam trajectory.

The choice of these three parameters, the beam energy and the luminosity is shown in Table 2.1 together with those of present KEKB. For the vertical beam-beam parameter ξ_y, it is assumed that the same value of 0.09 as has been achieved at KEKB. The vertical beta functions at the IP for SuperKEKB are smaller by almost by a factor of 20 than those of the present KEKB owing to the adoption of the Nano-Beam scheme. Assuming these parameters, one needs to double the total beam currents compared with those of the present KEKB to achieve the luminosity goal of SuperKEKB, $8 \times 10^{35}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The machine parameters of SuperKEKB including the three fundamental parameters are shown in Table 2.1.

Table 2.1: Summary of the most relevant machine parameters of the B-factory SuperKEKB

	KEKB Achieved	SuperKEKB
Energy(GeV) (LER/HER)	3.5/8.0	4.0/7.0
ξ_y	0.129/0.090	0.090/0.088
β_y^* (mm)	5.9/5.9	0.27/0.41
I(A)	1.64/1.19	3.6/2.62
$\mathcal{L}$ $(8 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1})$	2.11	80

2.2 Belle II detector

With the advent of the new SuperKEKB collider, also the Belle detector needs a thorough overhaul. The 40-fold increase in luminosity poses considerable challenges especially to the innermost detector systems. The goal is to maintain the detector performance achieved by the former Belle experiment under tightened conditions in terms of background radiation and occupancy of the innermost layers.

The most remarkable change in the upgrade process, will be the introduction of a completely new pixel vertex detector (PXD) in the innermost part of the vertex detector. The preexisting silicon vertex detector (SVD) will be also renewed, in

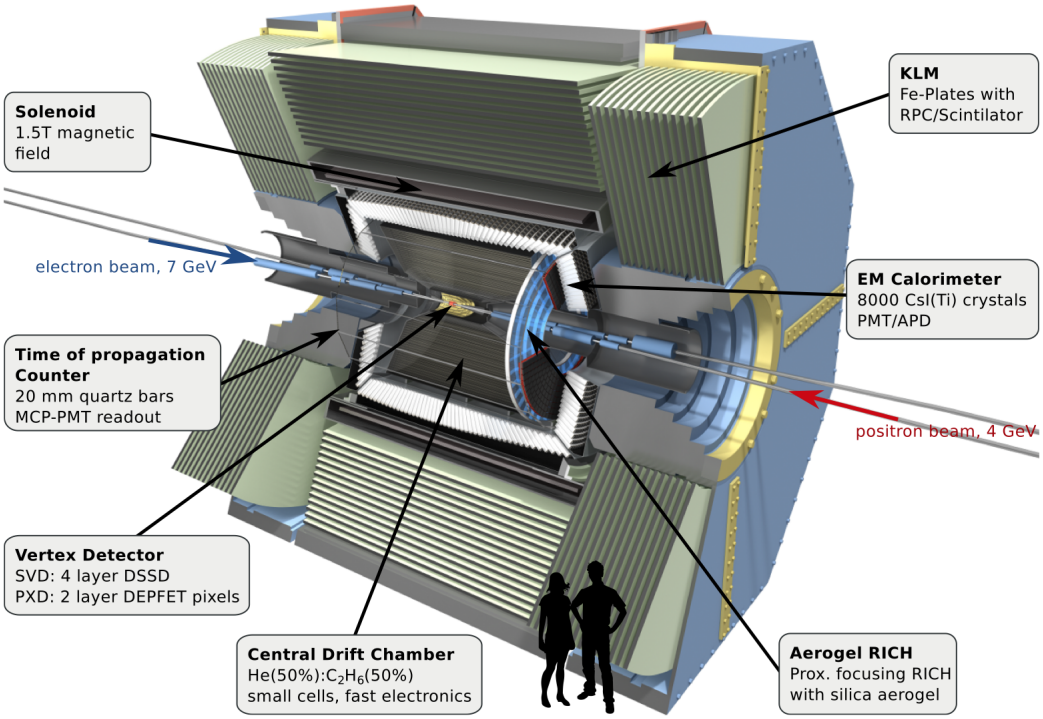


Figure 2.2: Overview of the Belle II detector [30].

order to free space for the new PXD. In consequence, the adjacent central drift chamber (CDC), responsible for the tracking of charged particles, will be extended to a larger radius while being subjected to other changes for the sake of a higher momentum resolution. The following sections will treat more carefully the design of the new tracking plane.

Behind the tracking detectors, the components for particle identification (PID) are situated. The PID devices are renewed as well and yield higher performance while being more compact, with very fast readout electronics. They consist of the Time-of-Propagation Counter (TOP), which is seated around the barrel, covering the whole φ-angle, and of the Aerogel Ring-Imaging Cherenkov Detector (ARICH), located at the end caps. Between the TOP and the solenoid sits the Electromagnetic Calorimeter (ECL), detecting photons and electrons; at the outermost, the barrel and the endcap K-Long and Muon Detectors (KLM) detect these two longer living particles.

2.2.1 Tracking Plane

Pixel Vertex Detector - PXD

As mentioned before, the vertex pixel detector is the completely new detector component in Belle II. It consists of two layers of pixel sensors based on Depleted Field-Effect Transistor (DEPFET) technology [7, 8].

Due to the new Nano-beam scheme and the new beam parameters in SuperKEKB, the luminosity of Belle II will be 40 times higher than its predecessor. Indeed this means a notable increase of statistics, but also carries with it the undeniable challenge to treat the enormous background. The background due to the new features is expected to cause very high hit rates, and consequently, there is no possibility to build the innermost layers of the vertex detector based on silicon strips. Thus, the innermost detector layers are closer to the interaction point (IP), i.e. at radii between 14-40 mm. Although this leads to a higher vertex resolution, it also increases the radiation levels. The large strip occupancy because of the high hit rates would make it impossible to reconstruct B-decays. The new PXD is the solution to this problem: pixel sensors have a notably larger number of channels and thus a much smaller occupancy.

The new PXD will provide an improved vertex resolution. As discussed in Sec. 1.3.3, a good vertexing capabilities are fundamental in order to distinguish the vertices of both B-mesons, and therefore, to measure CP violation.

The inner PXD layer has a radius of 14 mm and the outer layer 22 mm. Fig. 2.3 [6] shows a scheme of the geometrical sensor arrangement together with the external layers corresponding to the SVD.

The PXD is made of 40 sensors with a total amount of 7.68 million pixels. The DEPFET pixel sensors can be built very thin, due to their internal signal amplification. In $r\phi$ direction, the pixel size is 50 µm for both layers. In z direction, the sensors are segmented into two pieces. The size of the pixels in z direction varies from 55 µm to 85 µm; smaller pixels are placed closer to the detector center.

Due to the large number of readout channels, it's not feasible to read out all pixels for each event. Therefore, the SVD will perform quick particle tracking to formulate so-called ïegions of interest" in the PXD. Only the channels defined by these regions are read out.

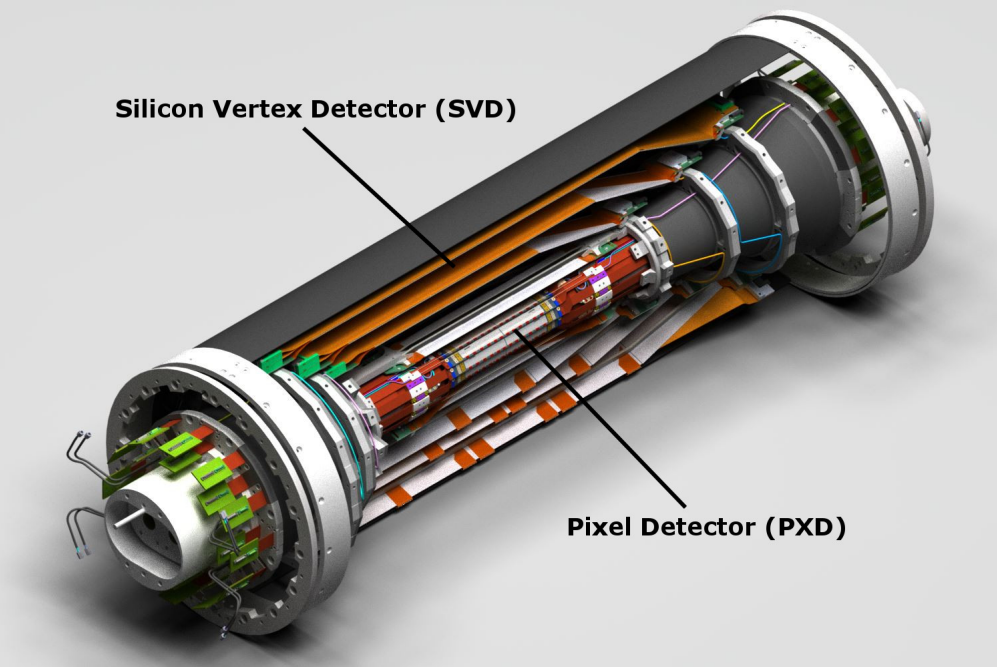


Figure 2.3: Schematic view of all the layers of the Vertex Detector, including the inner part (PXD) and the outer part (SVD).

Silicon Vertex Detector - SVD

The Belle II SVD consists of four layers equipped with double-sided silicon microstrip sensors (DSSD31s) at the nominal radii of 38 mm, 80 mm, 104 mm and 135 mm. The SVD provides the data needed to extrapolate the tracks detected in the CDC to the PXD, in order to improve the vertex reconstruction. Furthermore it helps in measuring the deflection of charged particles in the solenoidal magnetic field, determining their momenta.

The PXD and the SVD team up symbiotically. Using a simple and fast tracking algorithm, the SVD will provide the ẗegions of interest" in the PXD, which makes readout of the PXD possible in the first place. In return, the PXD provides very precise and unambiguous 2D position measurements with low occupancy, which would not be possible with silicon microstrip sensors.

Central Drift Chamber - CDC

The CDC is a large gaseous detector, which fulfils three major tasks: First, it reconstructs charged tracks and measures their momenta precisely from their deflection in the magnetic field. Second, it provides particle identification information using

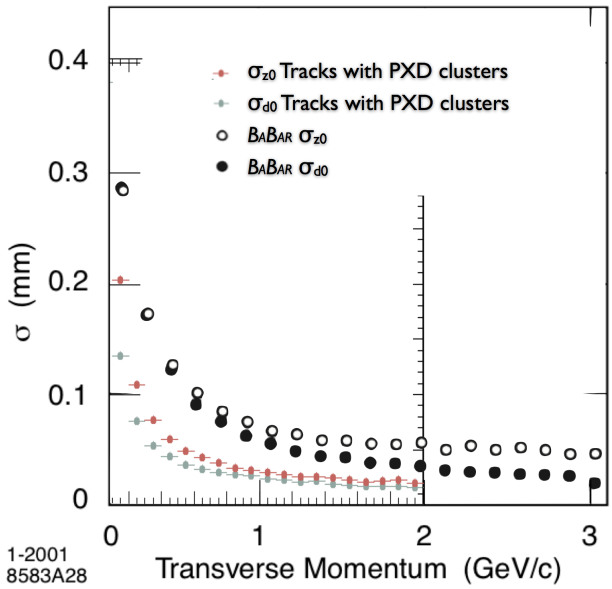


Figure 2.4: Comparison of Impact Parameter's resolution of BaBar and Belle II [35].

measurements of energy loss within its gas volume. Third, it provides efficient and reliable trigger signals for charged particles, since the PXD and SVD are not available for this purpose.

The CDC of Belle worked very reliably, so the Belle II CDC will see only minor modifications. Mainly, an increased radii and wire density, in order to improve the momentum resolution.

The Belle II CDC consists of 56 cylindrical layers and a total number of 14336 sense wires. The length of the CDC is 1960 mm and its inner and outer radii are 160 and 1130 mm, respectively. In contrast to Belle, the inner part of the Belle II CDC will feature a reduced spacing between the wires to increase the granularity. The performance is assumed to be similar or better to what was achieved in Belle: The CDC provides a position resolution of 2 mm in z and about 100 µm in $r\phi$ direction.

2.2.2 Impact Parameter

The expected improvement after the update of the tracking plane can be observed in the different impact parameters D_0, Z_0 resolution of BaBar and Belle II. represented in Fig. 2.4.

The impact parameters D_0 and Z_0 defined for each track, are the projections

of the distance from the point of closest approach (POCA) to the origin. They are a good measure of the overall performance of the tracking system and as such are used to find the optimal tracker configuration: they test the interplay between the geometry setup of individual sensitive detectors and physics effects.
With the excellent spatial resolution of the PXD, an impact parameter resolution in beam direction of $\sim 10\,\mu m$, which corresponds almost a factor two of improvement with respect to BaBar results (which were already better than Belle's).
The introduction of a new vertex fitter together with the improvement of the alignment procedure will thus sensibly improve the systematic error of any time dependent measurements.

CHAPTER 3

Analysis in Belle II

3.1 Reconstruction of B mesons

As explained in previous sections, SuperKEKB will operate at the center-of-mass energy corresponding to the $\Upsilon(4S)$. Since the $\Upsilon(4S)$ decays almost always to B^+B^- or $B^0\bar{B}^0$ (with only minor contamination from beam or electronics background) they enable the use of reconstruction techniques that would not be available at a hadron collider.

The principle in which the reconstruction of the B mesons in Belle II relies is illustrated in Fig. 3.1. Basically it reconstructs decays of particles into their immediate daughters and combines the output until reaching the B meson level. It uses the data recorded in terms of tracks and clusters in the electromagnetic calorimeter to perform a particle reconstruction level by level. To reduce the computational burden imposed by the combinatorics of the reconstruction, soft cuts based on the network output remove background at each stage without overly compromising efficiency [36].
Another reconstruction technique used in Belle II is the Full Event Interpretation (FEI), briefly explained in the following section.

3.1.1 Full Event Interpretation

As an analysis technique unique to B factories, the Full Event Interpretation will play an important role in the measurement of invisible decays involving neutrinos. This technique reconstructs one of the B mesons and infers strong constraints for the remaining B meson in the event using the precisely known initial state of the $\Upsilon(4S)$. The actual analysis is then performed on the second B meson. The two

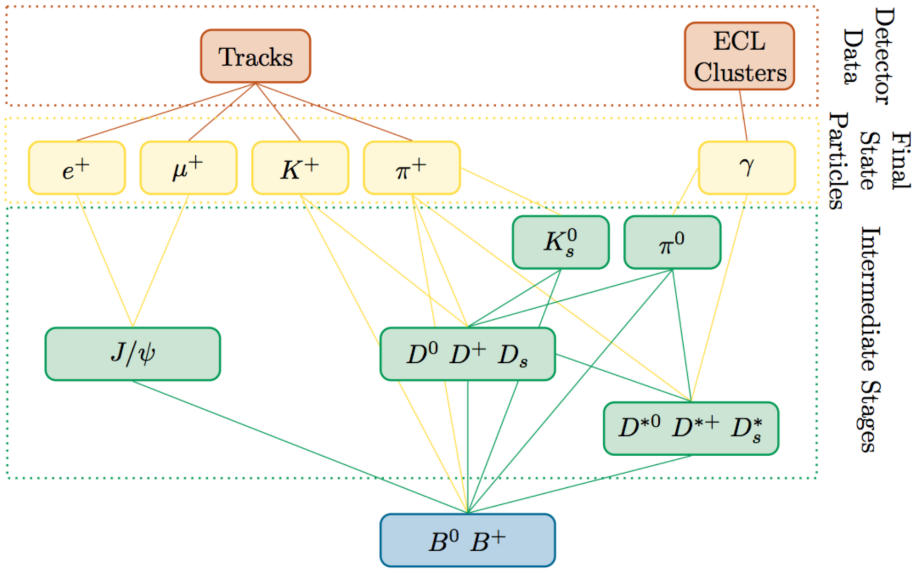


Figure 3.1: Hierarchical structure of the B meson reconstruction in Belle II [36].

mesons are called tag-side B_{tag} and signal-side B_{sig}, respectively. This situation is depicted in Fig. 3.2.

In effect the FEI allows one to reconstruct the initial $\Upsilon(4S)$ resonance, and thereby recovering the kinematic and flavour information of B_{sig}. Furthermore, the background can be drastically reduced by discarding $\Upsilon(4S)$ candidates with remaining tracks or energy clusters in the rest of event. [37]

This type of analysis is specially useful when dealing with neutrino channels, and even invisible ($B^0 \to \nu\bar{\nu}$). The signal side needs the kinematics to be resolved by the Tag Side reconstruction before performing any further analysis. Indeed, for theses analyses one needs to obtain with high precision the energy at the interaction point from the beam's energy.

However, there are other situations where this kind of analysis is not as efficient as needed. The time-dependent CP violating analysis case will be treated in the next section.

3.1.2 Time dependent analysis via B_{CP} → Golden Channel and B_{tag} → Generic

The best decays schemes to perform a time-dependent CP violation analysis (c.f Sec. 1.3.3) differ from the ones used for the FEI. In this case, one of the B mesons

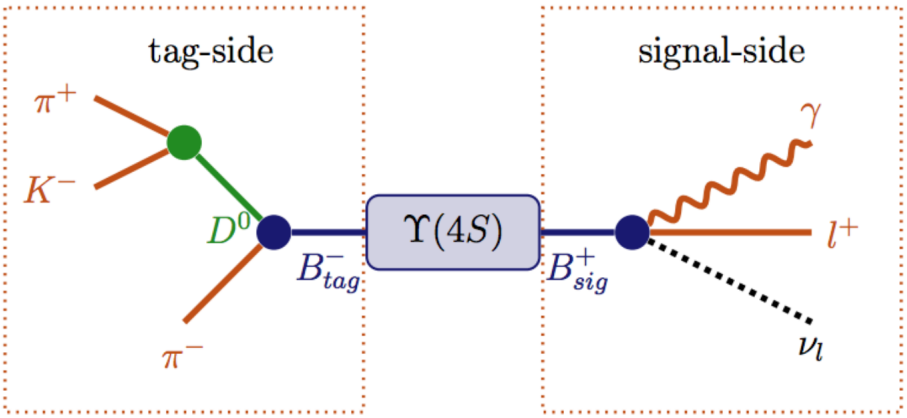


Figure 3.2: Decay of the Υ(4S) into a charged B meson pair. The signal-side and an important tag-side are shown for the analysis of the $B \to \ell\nu_\ell\gamma$ decay [37].

is chosen to decay into a CP eigenstate f_{CP}, being usually by convention B_{sig} that from now on will be called B_{CP} for being more intuitive. The remaining B meson, B_{tag} , is used for the flavor tagging as before, with the singularity that now it is not reconstructed.

The CP Side of the decay is then reconstructed for a specific CP eigenstate decay mode, i.e the golden channel $B^0 \to J/\psi K_S$. As discussed several times before, this channel allows to observe CP violation in the B sector with a decay which signature is hardly mis-understandable. As the B_{CP} is completely reconstructed, the kinematics of the decay are known. This implies that a second reconstruction for the B_{tag} on the Tag Side is then unnecessary. This carries the benefit that since the B_{tag} is not reconstructed from a determined decay mode, all of them are available and thus there is no loss on statistics.

There are, however, some problems that arise from this decay scheme. Since no specific decay mode is chosen for B_{tag} , there is no information about the tracks expected to come out directly after its decay. The solution for this problem comes usually from taking all the tracks that have not been used during the B_{CP} reconstruction and assign them to the Tag Side. For this reason, neither the flavor tagging nor the vertex fit on the Tag Side are performed trivially and both need a special treatment.

In the following sections the algorithms used to fit both vertices B_{CP} and B_{tag} together with the flavor tagging method will be treated in detail.

3.2 B_{CP} Vertex Reconstruction - RAVE: Kalman

A vertex reconstruction algorithm basically takes a set of reconstructed tracks as its input and produces reconstructed interaction vertices as its output. It deals both with finding (pattern recognition) and with fitting (statistical estimation) of the interaction vertices. In the case of the B_{CP} vertex, the input are both muon tracks. Those tracks can be filtered easily. One only needs to perform a kinematical cut to assure their invariant mass correspond to the invariant mass of the J/ψ. Moreover, a muon particle identification (PID) is performed for the tracks.

The classical vertex fitting algorithm used is the Kalman filter, which is a linear, least-squares method. The pattern recognition problem of finding a consistent subset of tracks is in the following assumed to be solved. Put it simple, the program gets as a input a given set of N tracks with three-momentum p_i, that need to be fitted in a certain point in the space, the vertex v. The measurements are the estimated track parameters $\bar{q}_i$ along with their covariance matrices $V_i = G_i^{-1}$ for i = 1, ... , n, at some reference surface. If one takes $q_i = f_i(v, p_i)$ [1] as the deterministic track model of track i, the function to be minimized in terms of the most suitable vertex is:

$$F(v, p_1, ..., p_N) = \sum_i^N e_i^T G_i e_i \tag{3.1}$$

Where the three-vector $e_i = q_i - \bar{q}_i$. The minimization of this function, which is function of all tracks' momenta, can be minimized in several ways that the reader may want to look up in [38]. For now the important point to take into account is that this algorithm takes all the tracks and finds the single point, the vertex, that better suits all of them by means of a least squares method.

This method is specially convenient in the absence of mis-measured or outlying tracks. In that case in particular, the vertex fit will be biased towards the outlying track, as can be seen in Fig. 3.3. In the case of the B_{CP} vertex reconstruction, the fit is performed by simply finding the J/ψ vertex, since it decays immediately after the B. This meson is reconstructed in its $\mu^+\mu^-$ decay, giving usually a set of two clear tracks that can be fitted easily by using the Kalman algorithm. That is, there is hardly any mis-measured or outlying track that could possibly worsen the fit. This is not the case for B_{tag} , which is chosen to decay into any possible decay channel, and that's why the algorithm needs to be refined, as will be explained in the next section.

[1] There are several types of track's model, depending on the analysis performed.

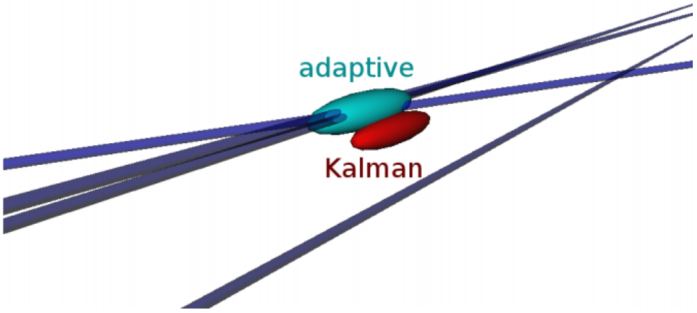


Figure 3.3: Reconstruction of a CMS Monte-Carlo$B_s \rightarrow J/\psi\phi \rightarrow K^+K^-\mu^+\mu^-$ decay vertex, where one of the tracks is mis-measured. The estimate of the Kalman filter is drawn towards the mis-measured track, the estimate of the adaptive method is less influenced by it [39].

3.3 B_{tag} Vertex Reconstruction - *Standard Algorithm*

In contrast with the reconstruction of a given channel for B_{CP} , the tagged B meson may decay into any possible final state. This means that, in principle, there is no invariant mass criteria that can be successfully applied without losing the majority of events, and therefore all tracks not used during the CP Side reconstruction need to be taken.

If the Kalman method were applied for performing the B_{tag} vertex fit, the vertex would be biased along the boost direction and the resolution would be degradated. In fact, there is no way to know, a priori, which tracks come from the B_{tag} vertex and which ones corresponds to secondary particle's decays.

For this reason, a new algorithm born from the Kalman method needs to be used. This algorithm do not discard tracks, but it rather down-weigths them depending on their position compared with the most probable vertex. In order words, outlying or isolated tracks contribute way less than the tracks that are packed and pointing towards the same vertex. This iterative weighting of the tracks is explained in detail in the next section.

3.3.1 Adaptive Vertex Fit

This set of algorithms [38, 40] introduces the concept of soft assignment; a track is associated to a specific vertex with an assignment probability, or weight w. The fitter is then implemented as an iterated, reweighted Kalman filter: in every iteration step new track weights are computed and subsequently for the determination of a vertex candidate. If χ_i^2 denotes the square of the standardized residual, then

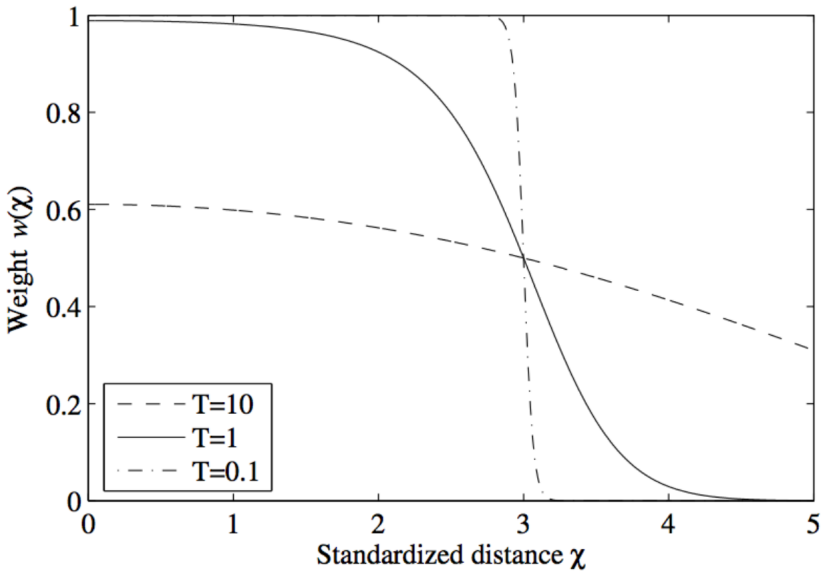


Figure 3.4: The weights of the AVF, as a function of the standardized residual $\chi \equiv p\chi^2$ and the annealing temperature T . For this plot, σ_{cut} has been chosen to be 3.0. This corresponds to $\approx 1\%$ of all inlier track being assigned a weight of less than 0.5, in the case of ideal data [40].

the weight w_i equals:

$$w_i(\chi_i^2) = \frac{e^{-\chi_i^2/2T}}{e^{-\chi_i^2/2T} + e^{\sigma_{cut}^2/2T}} \tag{3.2}$$

A cutoff parameter σ_{cut}^2 cut has been introduced which is defined as the standardized residual for which $w_i = 0.5$, and a *temperature* parameter T , which defines *softness* of the weight function. This weight can be interpreted as a track-to-vertex assignment probability. Instead of minimizing the least sum of squares, as was expected from a Kalman fitting method, this algorithm now minimizes the weighted least sum of squares. In order to avoid falling prematurely into local minima, a deterministic annealing schedule is introduced: in each iteration step the temperature parameter T is lowered. Typically, a *quasi-geometric* annealing schedule that converges towards 1 is employed:

$$T_i = 1 + r \cdot (T_{i-1} - 1) \tag{3.3}$$

Here, T_i refers to the temperature parameter T at iteration i and r denotes the annealing ratio. For convergence, $0 < r < 1$ is needed. Fig. 3.4 shows the weight as function of χ_i^2 and T.

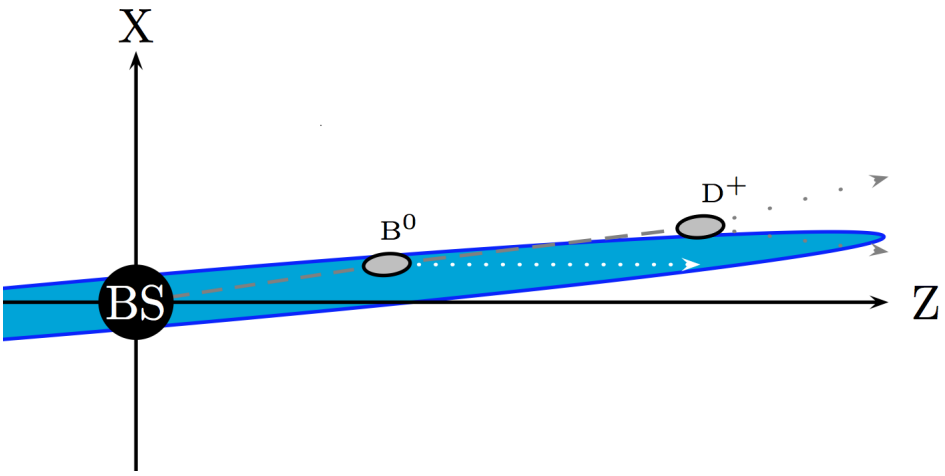


Figure 3.5: Fitting constraint implemented for the Standard Algorithm. It is an ellipsoid oriented in the boost direction of 400 µm for the largest semi-axis. It is centered in the Beam Spot. The decay depicted is $B^0 \rightarrow D^+\ell^-\bar{\nu}_\ell$. It can be observed the existence of two vertices, corresponding either to the B or to the D meson decay.

3.3.2 Fitting Constraint

However, although this method may help to reject badly reconstructed tracks, in the case of a generic B decay, it could reject an isolated track that precisely correspond to the only primary track in the event. For example, consider the following decay $B^0 \rightarrow D^-\mu^+\nu_\mu$. In this case there is only one possible primary track, the muon. Since the D meson may take more than one hundred of micrometres to decay, the tracks coming from its decay will often point to the wrong vertex. The algorithm takes the D vertex as the B_{tag} vertex, down-weighting the isolated muon track.

For this reason, a spatial constraint for the vertex fit has been implemented [41]. This constraint displays a three-dimensional Gaussian probability distribution around its center, modifying the position of the most probable vertex obtained via the Adaptive Vertex Fit, and therefore, refining the weighting of each track involved in the vertex fit. This spatial constraint is shaped as an ellipsoid of 500 µm long for the main axis. It is directed in the boost direction (41.5 mrad from the z axis) and centered in the beam spot (BS). A pictorial representation of the constraint can be observed in Fig. 3.5.

Due to the effective boost in the laboratory system, the B meson is expected to decay closer than any other secondary particle to the BS. This means that centring the constraint exactly in this point will guarantee that the track or tracks coming from either directly the B_{tag} or immediately decaying daughters will be

up-weighted with respect to other tracks that may originate further away from the center of the constraint.

3.4 Flavor Tagging

For the events in which it is reconstructed $B \to f_{CP}$ candidates, the neutral B flavor is identified from the decay products of B_{tag} . The charge of the available tracks after the CP Side reconstruction are usually a clear signature of the flavor of the B. The available information is obtained from leptons, kaons, Λ baryons, and pions. The different signatures of flavor specific decay channels can be grouped into categories. Some of the most relevant categories, and the ones that will be used for the rest of this thesis are the following [42]:

Categories	Targets
Electron	e^-
(Intermediate Electron)	e^+
Muon	μ^-
(Intermediate Muon)	μ^+
KinLepton	e^-
Kaon	K^-
KaonPion	K^-, π^+
SlowPion	π^+
FastPion	π^-
MaximumP	ℓ^-, π^-
FSC	ℓ^-, π^+
Lambda	Λ
Total= 10 (12)	

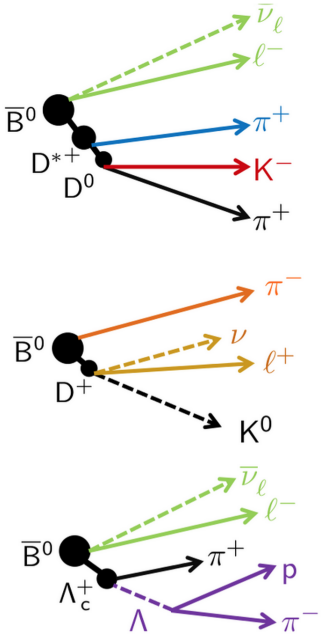


Figure 3.6: List of all available categories in the FlavorTagger currently. Three decays schemes have been included where one can observe the different targets corresponding to each category [42].

- Electron: In the decay $\bar{B}^0 \to Xe^-\bar{\nu}_e(b \to ce^-\bar{\nu}_e)$ the charge of the electron tags unambiguously the flavor of the B Meson.
- Muon: In the decay $\bar{B}^0 \to X\mu^-\bar{\nu}_\mu(b \to c\mu^-\bar{\nu}_\mu)$ the charge of the muon tags the flavor of the B Meson.

- Kaon: In the decay $\bar{\mathrm{B}}^0 \to XK^-(b \to c \to s)$ the charge of the Kaon is the searched flavour signature.
- FastPion: In the decay $\bar{\mathrm{B}}^0 \to D^+\pi^-(b \to c)$ the charge of the pion tags the flavor of the B Meson.
- SlowPion: In the decay $\bar{\mathrm{B}}^0 \to XD^{*+} \to XD^0\pi^+(b \to c)$ the charge of the slow pion tags the flavor. This is a secondary particle.
- MaximumP*: Here the particle with the highest CMS momentum is assumed to be a primary daughter of the B. Therefore, its charge is considered as flavor signature.
- Lambda: In the decay $\bar{\mathrm{B}}^0 \to \Lambda_c * X \to \Lambda X \to Xp\pi^-(b \to c \to s)$ the flavor of the Lambda tags the flavor of the B. For this, a proton and a pion are reconstructed to a Lambda.

In the following the particles providing the flavor tag information, i.e. the flavor signatures, are denoted as target. Fig. 3.6 shows the scheme of the different categories implemented in the FlavorTagger.

3.4.1 FlavorTagger Algorithm

The process of the FlavorTagger is inspired by the Flavor Tagging concept developed by Belle and BaBar. It proceeds in three steps or levels: *TrackLevel*, *EventLevel* and *CombinerLevel*. Each step relies on trained multivariate methods. An schematic overview of the different steps can be observed in Fig. 3.7.

At the starting point the available information consists mainly of: the tracks that have not been used on the $\mathrm{B_{CP}}$ reconstruction, the remaining clusters in the EM calorimeter (ECL) and kaon-muon detector (KLM). In the first level, the *TrackLevel*, the tracks pass through a mass hypothesis fit for each category, e.g electron, muon and so on. The different variables calculated from the fit are taken as inputs for a category specific multivariate method. To each category a *TrackLevel* multivariate method is assigned which is trained to give as output the probability of being the *RightTrack*, i.e. the searched signature, for the corresponding category. In other words, the output after the first level must be a probability for each track to be the signature or target of every category.

In the next level, the *EventLevel*, the probability for each track of belonging to the specific decay mode or category is calculated. It is called *RightCategory* probability and it is calculated using again the flavor tagging variables as input

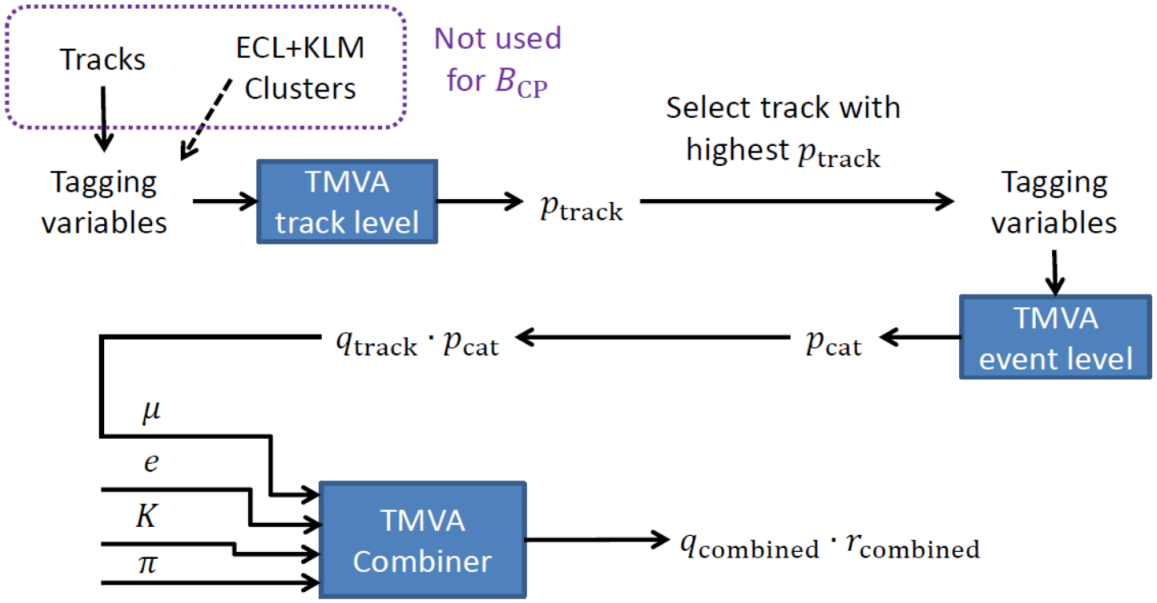


Figure 3.7: Organization chart of the FlavorTagger. It can be seen the different stages of the analysis, starting with the TMVA training, going through the track level, the event level and the combiner level [42].

for each track. For this purpose, during the calculation, the correspondent track is assumed to be the right track, regardless the *RightTrack* probability obtained in the *TrackLevel*.

Therefore, after the *EventLevel* each track in each one of the categories has associated two probabilities: the probability of being the *RightTrack* in its category and, assuming that it is the *RightTrack*, the probaility for its category to be the *RightCategory*.

For the final level, the *CombinerLevel*, the track with the highest *RightTrack* probability is selected as target for each category. The product $q \cdot p$ of charge and *RightCategory* probability of the target is an input value for the combiner. The multivariate method of the *CombinerLevel* is trained to give the qr value belonging to the tagged B-meson as output. Here q means flavor, and r is the so-called dilution factor.

So basically, the main result of the *FlavorTagger* is the parameter $qr \in [-1, 1]$. When it is closer to -1 $(+1)$, B_{tag} is more probable to be $\bar{B}^0(B^0)$. Accordingly, if $qr \approx 0$, the information about the flavor is null.

This information together with *RightCategory* and *RightTrack* probabilities can

and will be used in the following chapters in order to create a new B_{tag} vertex fitting algorithm.

3.5 B_{CP} Vertex Fit

For completeness of this thesis, a preliminar study of the PXD vertexing capabilities has been performed both on the CP Side and on the Tag Side. The results shown in the following lines are a reproduction of tests with the *Standard Algorithm*s that have been already performed in Belle II [41].

More concretely, a Geant 4 simulation of 200000 Υ (4S) $\rightarrow B^0\bar{B}^0$ events has been performed, with $B_{CP} \rightarrow [J/\psi \rightarrow \mu^+\mu^-][K_S \rightarrow \pi^+\pi^-]$ and $B_{tag} \rightarrow$ generic. The simulation includes the entire Belle II detector but not the beam background, as it is considered non reliable.

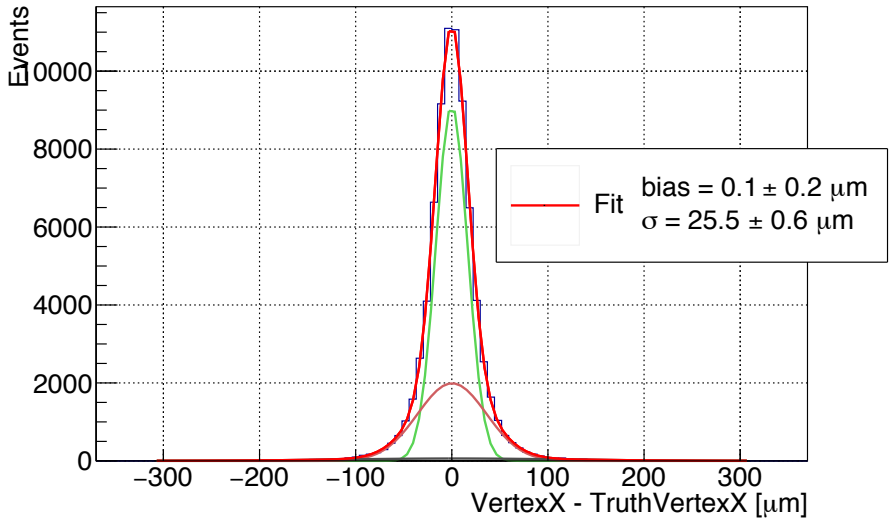
To obtain the most of the inclusion of the new Pixel Vertex Detector, during the vertex fit performed via Kalman algorithm the tracks have been forced to leave at least one hit in the PXD. This assures that the muons used for the fit of the $J/\psi \rightarrow \mu^+\mu^-$ are correctly reconstructed. Figs. 3.8a, 3.8b show the resolution of the B_{CP} vertex in the X and Y axis. The resolution is obtained as the difference between the fitted vertex using the reconstructed tracks and the generated vertex from the Monte Carlo simulation.

In Fig. 3.9a one can observe the CP vertex resolution on the Z axis. The particularity of the Z axis is that it corresponds to good extend to the direction of the boost, and therefore is the most sensitive axis to the algorithm features.

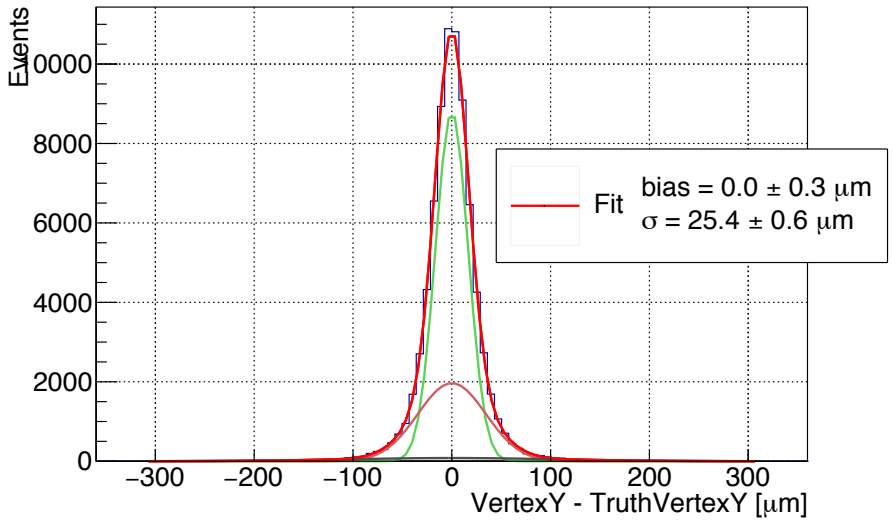
The fit shown in Fig. 3.9a results in a resolution of 22.9± 0.5 µm, almost three times better than the best resolution obtained by the Belle experiment for the same analysis, Fig. 3.9b. Besides the improvement in the resolution, a small bias of 1.8± 0.2 µm from the Monte Carlo generated vertex is observed. This bias, whose origin is still under investigation, could have some implications in the determination of the beam spot that will be treated in Sec. 4.1.

3.6 B_{tag} Standard Algorithm Vertex Fit

As it has been done for the CP Side, in this section the resolution of the tag vertex has been recalculated for completeness of the thesis [41]. This time the algorithm used to perform the vertex fit has been the Adaptive Vertex Fit complemented with the spatial constraint described in Sec. 3.3.2. The tracks used for the fit are



(a) $J/\psi \to \mu^+\mu^-$ vertex resolution in the X axis.

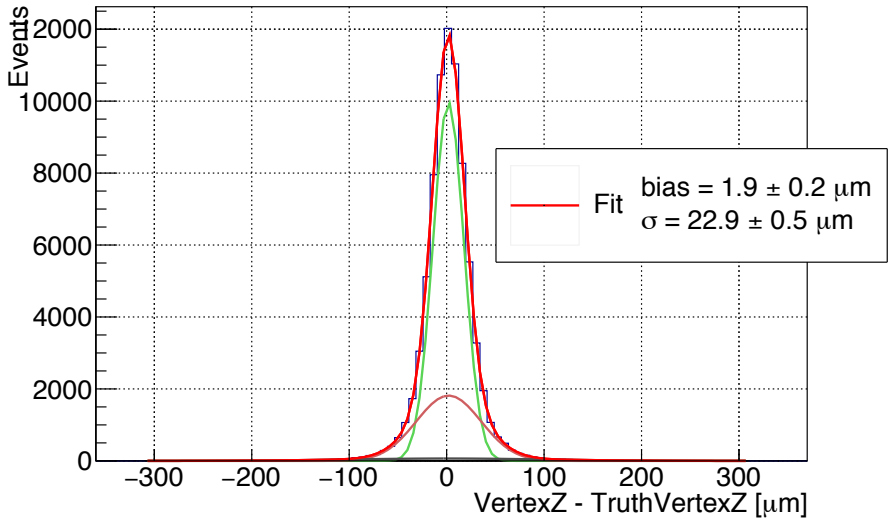


(b) $J/\psi \to \mu^+\mu^-$ vertex resolution in the Y axis.

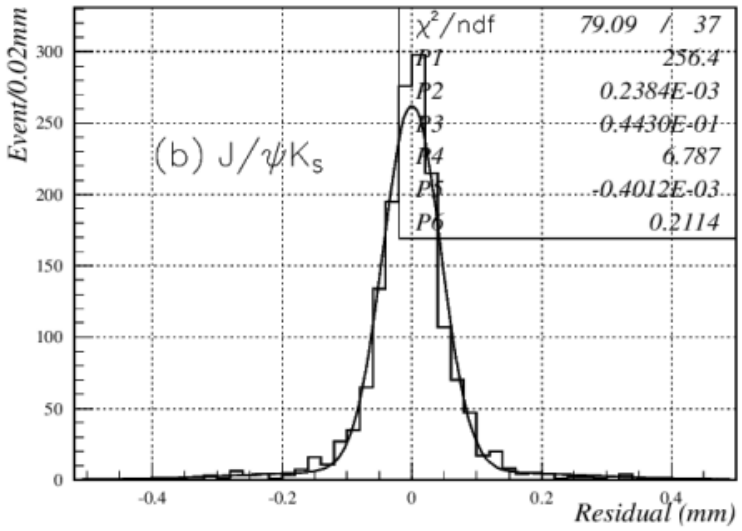
Figure 3.8: B_{CP} vertex resolution in the X,Y axis. The fit has been performed by fitting the distribution to three gaussians. Both the center and the resolution are a weighted average of those values for the three gaussians.

all the ones unused during the B_{CP} reconstruction. In a similar way as done for the CP Side, the tracks need to have at least one hit in the Pixel Vertex Detector.

The result of the simulation can be observed in Fig. 3.10a, and can be compared with the old results of Belle [43] in Fig. 3.10b. The improvement in the resolution is notorious, a factor 1.7, but not as good as in the CP Side. Indeed, the input track selection for the fit plays a crucial role on the final resolution. The very precise Pixel Detector is not as effective as it could be because of the non-optimal



(a) $J/\psi \to \mu^+\mu^-$ vertex resolution in the Z axis.

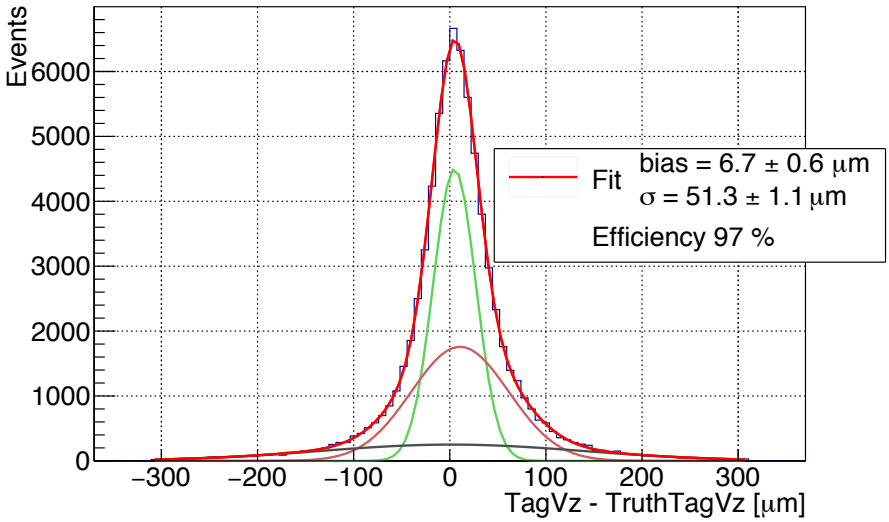


(b) Latests results regarding B_{CP} vertex resolution in Belle experiment. The result of the fit is bias = 0.2 µm, resolution = 63 µm [43].

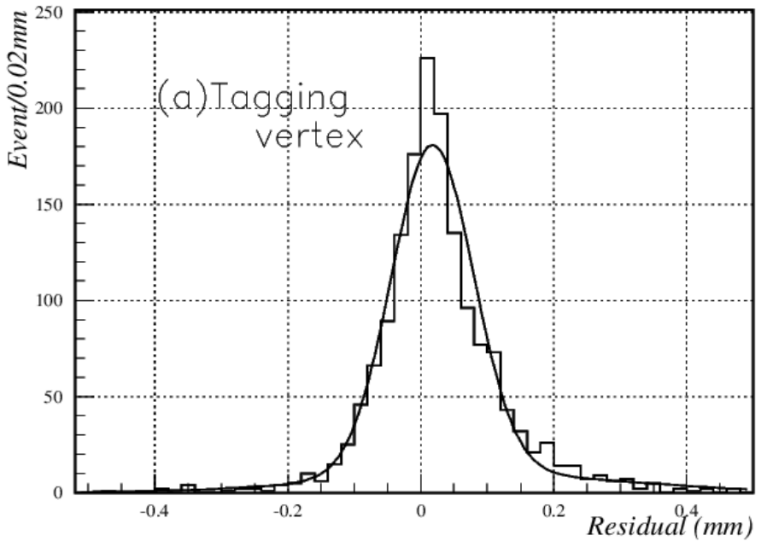
Figure 3.9: Comparison of the Belle and Belle II results for the B_{CP} vertex resolution in the Z axis.

algorithm for track selection. All the tracks unused for the CP Side reconstruction may not come directly from the B_{tag} vertex. In fact, most of them probably come from secondary daughters like D^+ that do not decay immediately and have a vertex far from the B_{tag} vertex.

This is the reason why an improvement of the track selection algorithm needs



(a) Z component of the B_{tag} vertex resolution. For the fit the *Standard Algorithm* has been used.



(b) Latests results regarding B_{tag} vertex resolution in Belle experiment. The result of the fit is bias = 29 μm, resolution = 89 μm [43].

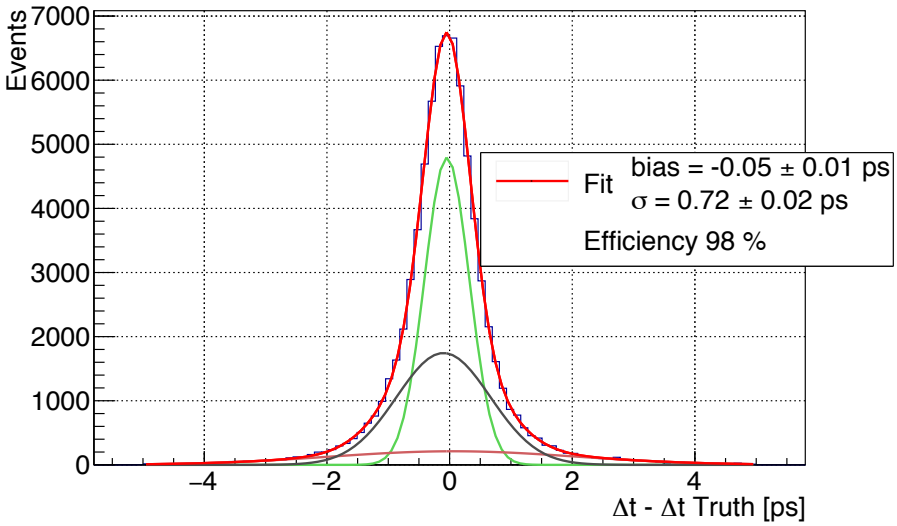
Figure 3.10: Comparison of the of the Belle and Belle II results for B_{tag} vertex resolution in the Z axis

to be performed for the Tag Side of the time-dependant analysis. The main objective of this thesis is, in fact, to design a new algorithm that improves the vertex resolution. The ideal case would be to find an algorithm that gave the same improvement in the Tag Side as in the CP Side, that is, almost a factor 3. By the end of this thesis it will be clear how, after using this new algorithm, the improvement

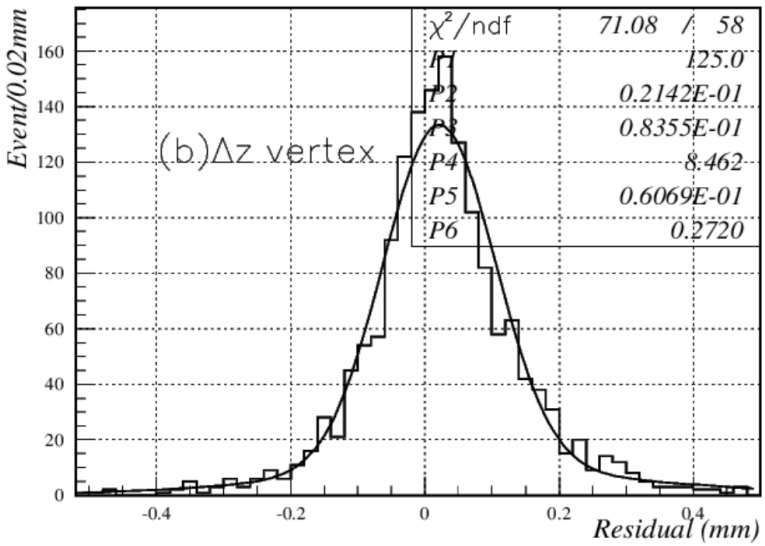
obtained is not yet but way closer to the ideal.

In Fig. 3.11 is portrayed the improvement in the Δt resolution. More concretely, Fig. 3.11a shows this resolution after using the *Standard Algorithm* for the B_{tag} vertex fit. In Fig. 3.11b, however, is shown the Δz resolution that can be correlated with Δt using Eq. 1.49.

The improvement of 3.11a with respect to the Belle results is about a factor 1.2. It is not as pronounced as the vertex resolution both in the CP and Tag Sides. This has a very simple explanation: the change of the boost. Even if the resolution of the vertices has improved, i.e Δz resolution, the boost plays an important role being in the denominator in Eq. 1.49. As explained in Sec. 2.1.1, the boost has been reduced from almost $\beta\gamma = 0.48$ to $\beta\gamma = 0.26$. Therefore, the uncertainties in Δz are amplified by this boost and the Δt resolution is worsened.



(a) Δt resolution in the most recent Belle II simulations. For the B_{tag} vertex fit the *Standard Algorithm* has been used.



(b) Latests results regarding Δz resolution in Belle experiment. The result of the fit is bias = 0.2 ps, resolution = 0.92 ps [43].

Figure 3.11: Comparison of Δt and Δz resolution of Belle II and Belle.

CHAPTER 4

Study about the constraint in the B_{tag} vertex Standard Algorithm

As detailed in Sec. 3.3.2, the constraint is a key feature during the B_{tag} vertex fit. It provides an effective down-weighting of secondary tracks, even if they are all bunched together and would be initially up-weighted by the plain Adaptive Vertex Fit algorithm. The position where the constraint is to be centered is therefore highly determinant on the weighting of the tracks. This location correspond to the Beam Spot (BS), where both beams intersect.

The present chapter will be devoted to the analysis of the effect of a shift of the center of the constraint and its impact on the B_{tag} vertex reconstruction. In other words, it is investigated the possible effect of a uncertainty or mistake on the determination of the Beam Spot, where the constraint is centered.

Throughout the following section a preliminary study about the possible uncertainty on the BS position will be discussed. This will be followed up by an intensive study of the constraint position's impact on the Tag Side vertexing.

4.1 Knowledge of the beam spot position - center of the constraint

In the Monte Carlo simulation used in this thesis, the BS is not represented by a single point in the space but more like a gaussian distribution around the origin. When data collection will start, the center of the BS woll be determined using $e^+e^- \to \mu^+\mu^-$ events fitting both muon tracks to one vertex, in a similar way to the vertex fit of $J/\psi \to \mu^+\mu^-$. In Fig. 3.9a (cf. Sec. 3.5) has been observed an unexpected shift on the Z-axis vertex distribution of J/ψ with respect to the Monte

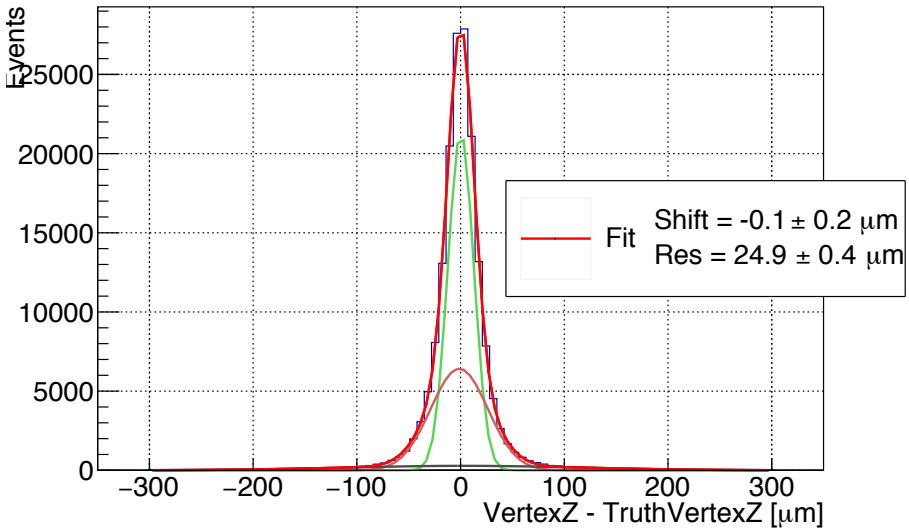


Figure 4.1: Z component of the $e^+e^- \to \mu^+\mu^-$ vertex resolution.

Carlo information. This unexpected bias could appear during the reconstruction of $e^+e^- \to \mu^+\mu^-$ vertex, and therefore would imply a shift on the BS.

A simulation of 500000 $e^+e^- \to \mu^+\mu^-$ have been performed in order to study the vertex reconstruction directly in the BS. The procedure is very similar to the one used to perform the J/ψ vertex reconstruction, but this time the invariant mass of both muons need to correspond to the energy carried by the virtual photon. Both muons are requested to have at least one hit on the Pixel Detector. In Fig. 4.1 the results of this analysis are shown. This time there is no apparent shift on the z-axis of the vertex resolutions. Instead, the small value is in perfect agreement with zero. The resolution is $\sigma = 23.5 \pm 0.3\,\mu m$, almost the same as the one obtained for J/ψ.

Although the pair of muon production may seem similar in both processes, i.e $e^+e^- \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$, there are some subtleties that may explain why the bias in the vertex reconstruction is not the same. The explanation could come from the angular distribution of the pair of muons. Regarding this topic, in Appendix A a deep study of the kinematics of the production is performed. The final conclusion is depicted in Table 4.1.

Indeed, the most visible difference between both vertices is that J/ψ decays into a pair of muons that have lower momentum and are produced with a higher boost than the ones from the e^+e^- vertex. This means that in the former, the amount of events with muons going in the same direction is much higher than in the latter, as can be observed in Table 4.1.

Table 4.1: Percentage of events for different emission directions of the pair of muons. Legend: FF forward-forward, BB backward-backward, FB forward-backward. All relative to the Z-axis

	FF	BF	BB
$e^+e^- \to \mu^+\mu^-$	15.3 %	84.7%	0%
$J/\psi \to \mu^+\mu^-$	26.7%	64.8%	8.5%

This would not be a problem by itself if it were not for some technical issues with the measurement of the impact parameter Z_0 of each track. The vertex is obtained as the intersection of both tracks, being Z_0 the closest point to the production vertex in the Z axis. If there were a small bias on the impact parameter, backward-forward pairs of tracks would cancel each other's biases in the Z axis. But events where both tracks are emitted in the same direction, the biases instead of cancel each other would sum up and give an effective shift on the vertex position. This possible cause is not confirmed yet and needs further studies by the tracking group of Belle II. However, it could explain the unexpected shift obtained for the J/ψ vertex, being at the same time in agreement with the results obtained in this thesis for the $e^+e^- \to \mu^+\mu^-$ vertex.

4.2 Shift on the center of the constraint

Although in the last section it is shown that, in principle, any possible shift in the BS is compatible with zero, it is still important to understand the effect of the constraint position during the vertex fit. Along this section, several shifts will be applied to the constraint and their effects will be analysed in terms of B_{tag} vertex resolution. Table 4.2 shows the different shifts applied.

The shifts are applied in a very simple way. Instead of centring the constraint in the origin of the coordinates, it is centred at some distance from it. In Fig. 4.2 is represented schematically an unspecified shift in the Z direction. As can be seen in Table 4.2, the shifts are performed in the three cartesian directions plus one more shift in the direction orthogonal to the boost.

As it has been explained in Sec. 2.1.1, both B mesons are produced in the lab frame with a non-zero boost almost along the Z axis. For this reason, the constraint needs to be tested with a higher shift in this direction. The shift goes up to 200 µm although that amount of uncertainty is orders of magnitude larger

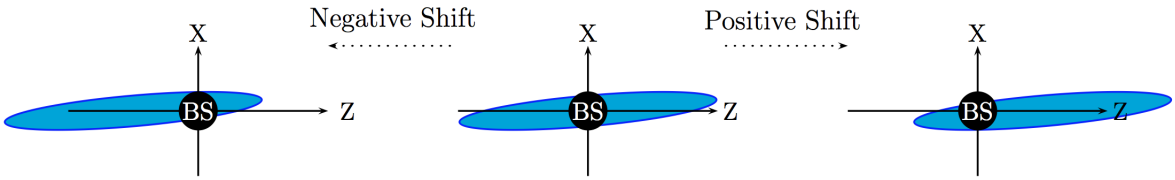


Figure 4.2: Scheme of the shift applied to the center of the constraint in the Z axis. The blue ellipse corresponds to the projection of the constraint in the plane XZ. The shift is performed in both the positive and the negative sides of the system. This scheme is representative of the shift applied in other directions as well.

than the expected one. Actually, that amount of shift is larger than the expected resolution of the BS (cf. Fig. 4.1). In any case, this extreme provides a check of the consistency of the constraint and to test its limits.

In the case of the remaining directions, a shift greater than a few micrometers would be undoubtedly non realistic. Both B mesons are produced practically at rest in the center-of-mass frame, and thus their momentum in the transversal plane is relatively reduced in comparison with the momentum gained in the Z-axis due to the boost. This means that their effective displacement before decaying in any direction but the Z-axis is very small. A shift in that direction of the order of magnitude of a possible shift in the Z-axis would bias the distribution heavily, as can be observed in Fig. 4.3

Table 4.2: Summary of the ranges chosen for the constraint's shift analysis. Boost means the direction of the boost, i.e 41.5 mrad with respect to the Z axis; Orthogonal means the direction perpendicular to the boost in the XZ plane, i.e 41.5 mrad with respect to the X axis.

Direction	Min. shift	Max. shift
X - axis	$\pm$ 1 µm	$\pm$ 4 µm
Y - axis	$\pm$ 1 µm	$\pm$ 4 µm
Boost	$\pm$ 50 µm	$\pm$ 200 µm
Orthogonal	$\pm$ 1 µm	$\pm$ 4 µm

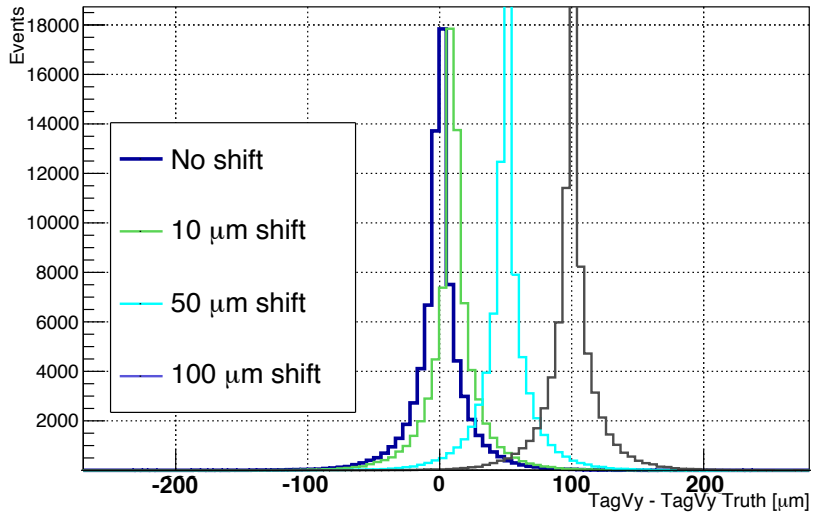


Figure 4.3: Y component of the B_{tag} Vertex after applying some excessive shifts on the fitting constraint center along the Y axis.

4.2.1 B_{tag} Vertex position and resolution

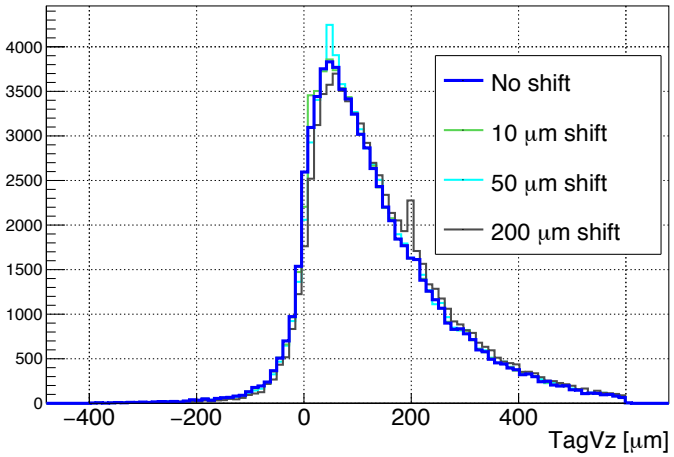
In this section the results of the B_{tag} vertex fit after applying the aforementioned shift of the constraint's center will be shown. More concretely, the Z component will be the one treated in the following paragraphs. In Fig. 4.4 it is shown the Z component of the vertex position, and in Fig. 4.5 can be seen instead the resolution of the vertex in that direction.

In Fig. 4.4 it can be observed a very similar vertex distributions for the different shifts. Basically there is no noticeable change up to 50 µm, and the first signs of an effective shift on the distribution happens around 50 µm.

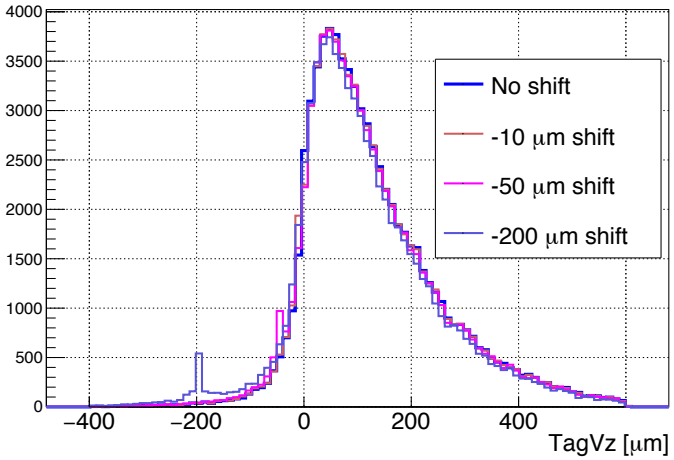
However, there are some noticeable and completely unexpected peaks in the distributions, more concretely at 10, 50 and 200 µm. Those peaks are not coming from physics, but from wrong computing. The algorithm used to perform the fit (RAVE: Adaptive Vertex Fit, see 3.2) has one subtlety when a bad reconstruction happens. When the pValue of the vertex fit is equal to zero the algorithm auto-assigns a fixed value for the vertex that corresponds to the origin of coordinates of the fit, i.e the center of the constraint. That's exactly why there is such a peak on the distribution, located each time in a different point depending on the shift applied to the constraint.

In Fig. 4.5, the resolution of the Z component of the B_{tag} vertex can be observed.

In this case it is more evident the effect of a shifted constraint of the fit. As in Fig. 4.4, up to 50 µm the effect is barely noticeable, being completely unobservable



(a) Z component of the B_{tag} Vertex after positive shifts

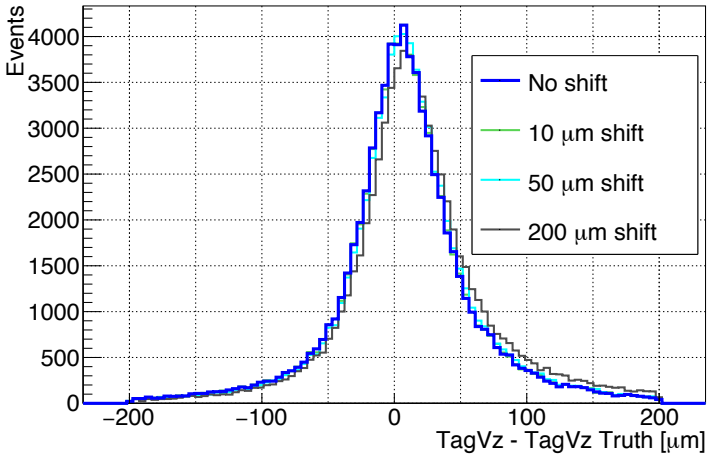


(b) Z component of the B_{tag} Vertex after negative shifts

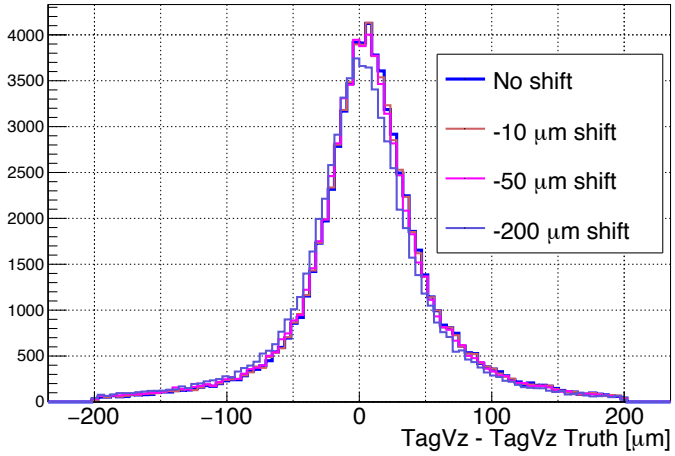
Figure 4.4: Z component of theB_{tag} Vertex position after applying different shifts on the fit constraint. Efficiency of 98%

for 10 µm. In contrast, for 200 µm the effect is evident, as it could be expected for such large shift.

It is important to point out that in Fig. 4.5, the unexpected peaks observed in Fig.4.4 are gone. During this analysis, the events that brought a pValue = 0 were discarded for illustrative purposes. After this cleaning it is clear that the peaks disappear, thus confirming the previous argumentation about the cause of such peaks. More concretely, from now on the pValue > 0.001 is equested in order to avoid events performing an incorrect vertex fit. This loss of statistics is acceptable for this kind of analysis, since a wrong values of the reconstructed vertex lead to wrongly calculated Δt values. For other types of analysis there is no reason why



(a) Z component of the B_{tag} Vertex resolution after positive shifts



(b) Z component of the B_{tag} Vertex resolution after negative shifts

Figure 4.5: Z component of theB_{tag} Vertex resolution after applying different shifts on the fit constraint. Efficiency of 96%

those events must be eliminated.

4.2.2 Δt distribution

In this section, the Δt distributions of the B meson system is obtained, for different shifts applied to the fit's constraint. The shifts are performed in several directions listed in Table 4.2. The following plots show the Δt distribution for B_{tag} corresponding to both $B^0$and $\bar{B}^0$. The fact that both distributions are not the same is a clear indicator of CP violation, and has been discussed previously in Sec. 2.1.1, the resolution of this distribution is key.

In Figs. 4.6 and 4.7 it can be observed the effect of the shift in the X and

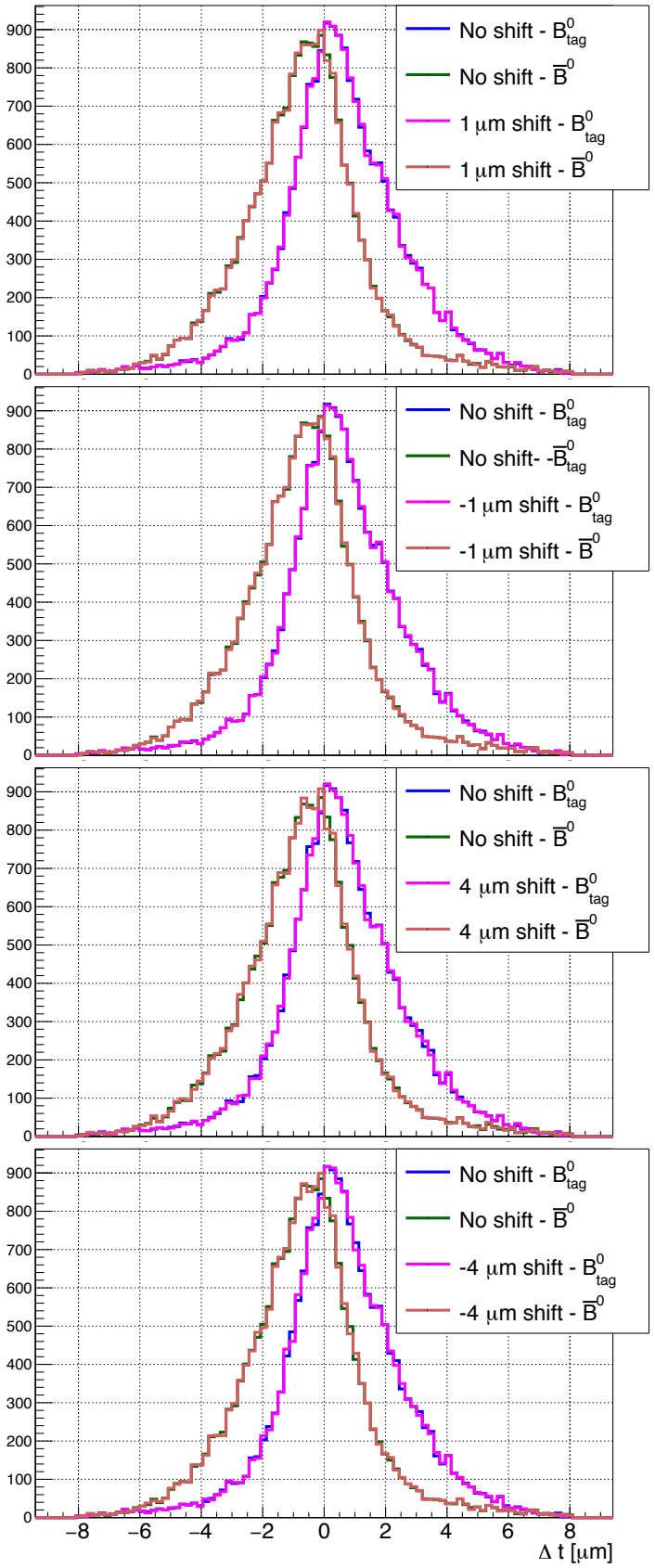


Figure 4.6: Δt distribution of the B-meson system after a shift on the fitting constraint of the Tag Side in the X-axis. (a) Small positive shift. (b) Small negative shift. (c) Large positive shift. (d) Large negative shift.

Y directions respectively. As expected, for 1 µm there is absolutely no difference. The shift is still too small. Nevertheless, for 4 µm a small bias can be observed in the Δt distribution, being positive (negative) for a positive (negative) shift on the constraint. The reason for this effect is simple: the shift on the constraint implies a shift on the reconstruction of the vertex. A vertex reconstructed further away from the interaction point implies that the travelling time must have been higher, and therefore the time it has taken for B_{tag} to decay longer. This forces Δt distribution to be shifted in the positive side. An analogous behaviour can be observed when the shift is negative, and so is the bias in the Δt distribution.

In Fig. 4.8 one can observe the result for the Δt distribution of the system after a shift on the fit's constraint in the Z-axis. This time, as it has been done in the previous section, the shift taken is larger than the one taken for the X,Y axis. 50 µm is a very large shift to be applied to the constraint, or in other words, a huge uncertainty or error in the BS position far from any possible scenario. The fact that for this shift the distributions are basically untouched is very interesting because that kind of shift outside the simulations will be, therefore, highly unexpected. Therefore, the distribution is safe from possible uncertainties. Indeed, when the shift goes up to 50 µm, the bias in the distribution is visible, as expected.

Finally, in Fig. 4.9 can be observed the same plots as before, but this time for a constraint's shift along the direction orthogonal to the boost, that is, 84mrad with respect to the X axis. The results are very similar to the one obtained in Figs. 4.6 and 4.7. No noticeable change on the distribution, just some signs of biasing when the shift goes up to 4 µm.

In general has been observed that the results are very safe from possible errors or uncertainties on the BS. The constraint's objective is to help down-weighting secondary tracks when the vertex fit is performed. This task is still manageable even if the center of the constraint is shifted slightly. The reason is because the primary tracks are still closer to the center of the constraint than the secondary tracks, and therefore they are still up-weighted. When the shift is too large, the constraint interprets that, although the primary tracks are still the correct ones to be up-weighted, the vertex is a little bit moved in the direction of the shift. In other words, the probability distribution held by the constraint would have the highest probability point for the vertex slightly deviated due to the large shift suffered by the constraint. This is however not a problem, since such large uncertainties are completely out of the question. As it was pointed out before, the main purpose if this study was to check the stability of the algorithm with respect to possible uncertainties. Indeed, the algorithm passed the test.

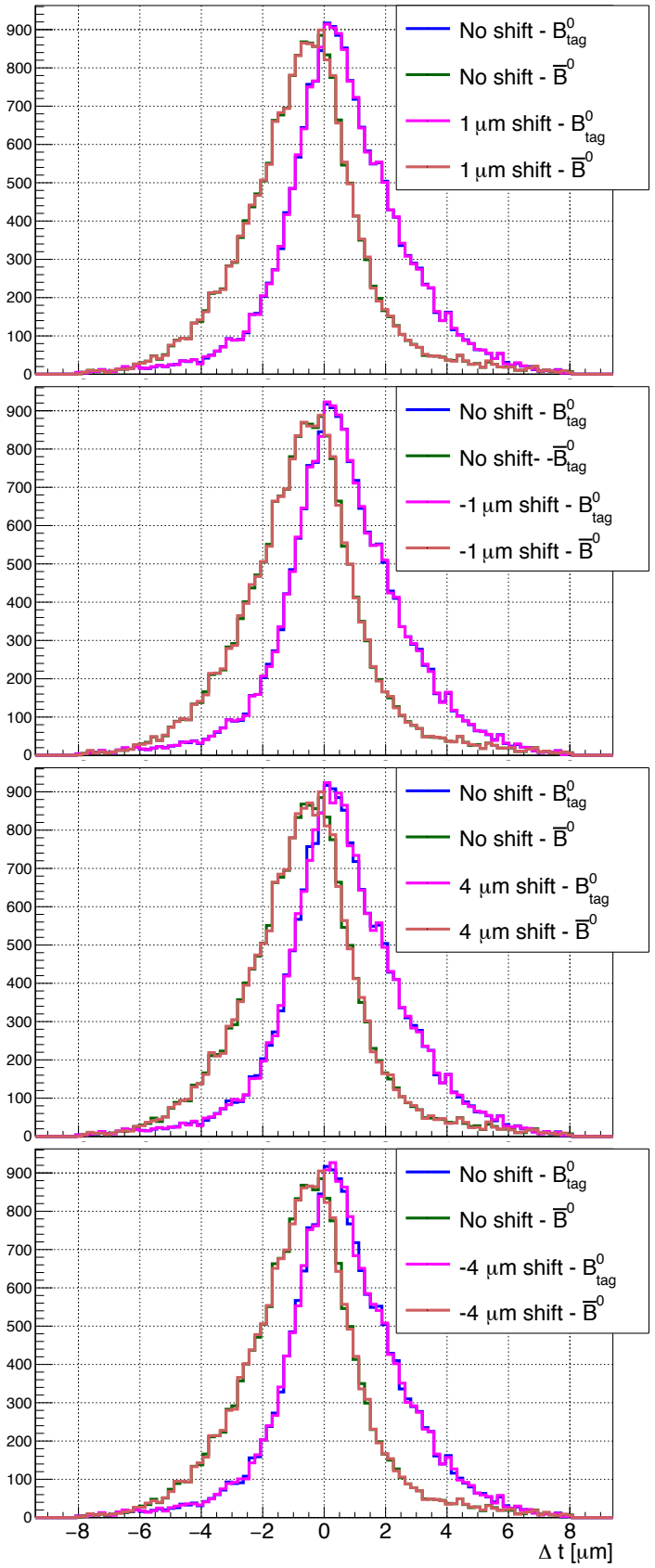


Figure 4.7: Δt distribution of the B-meson system after a shift on the fitting constraint of the Tag Side in the Y-axis. (a) Small positive shift. (b) Small negative shift. (c) Large positive shift. (d) Large negative shift.

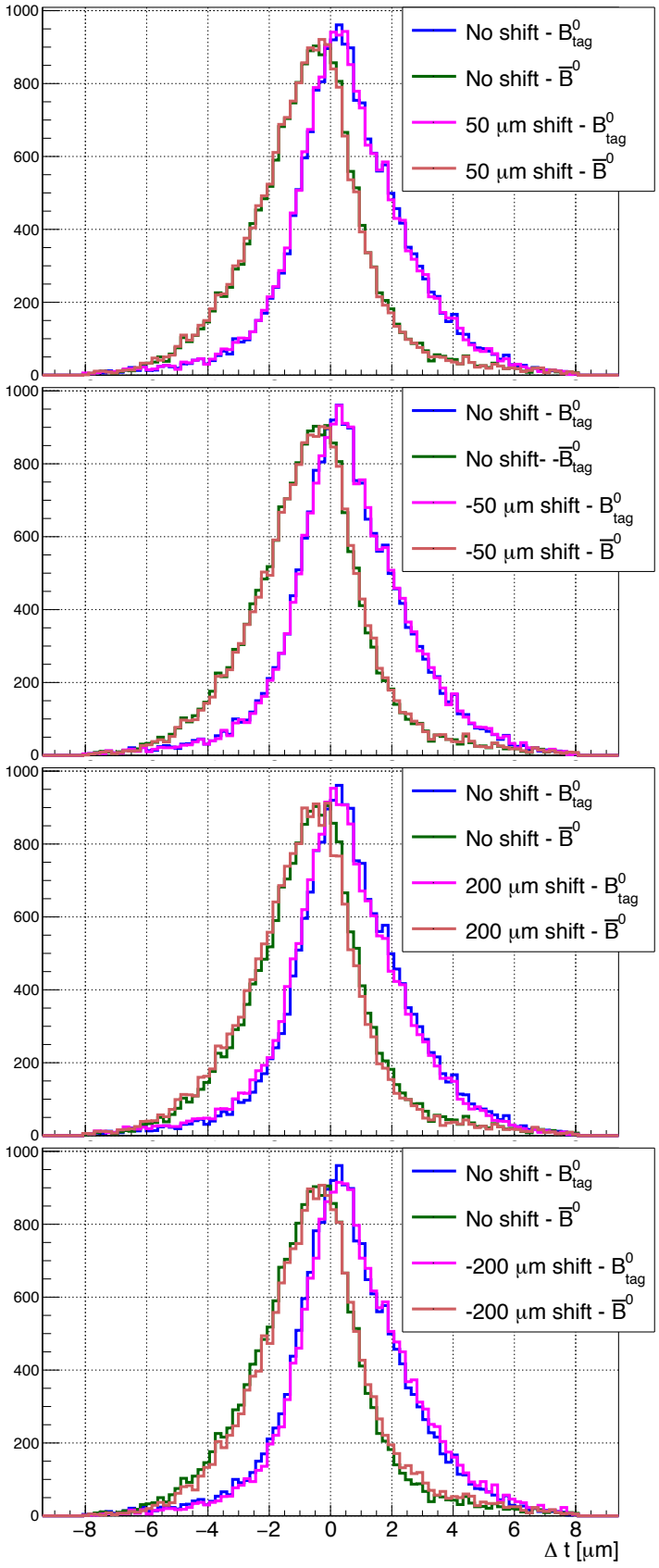


Figure 4.8: Δt distribution of the B-meson system after a shift on the fitting constraint of the Tag Side in the Y-axis. (a) Small positive shift. (b) Small negative shift. (c) Large positive shift. (d) Large negative shift.

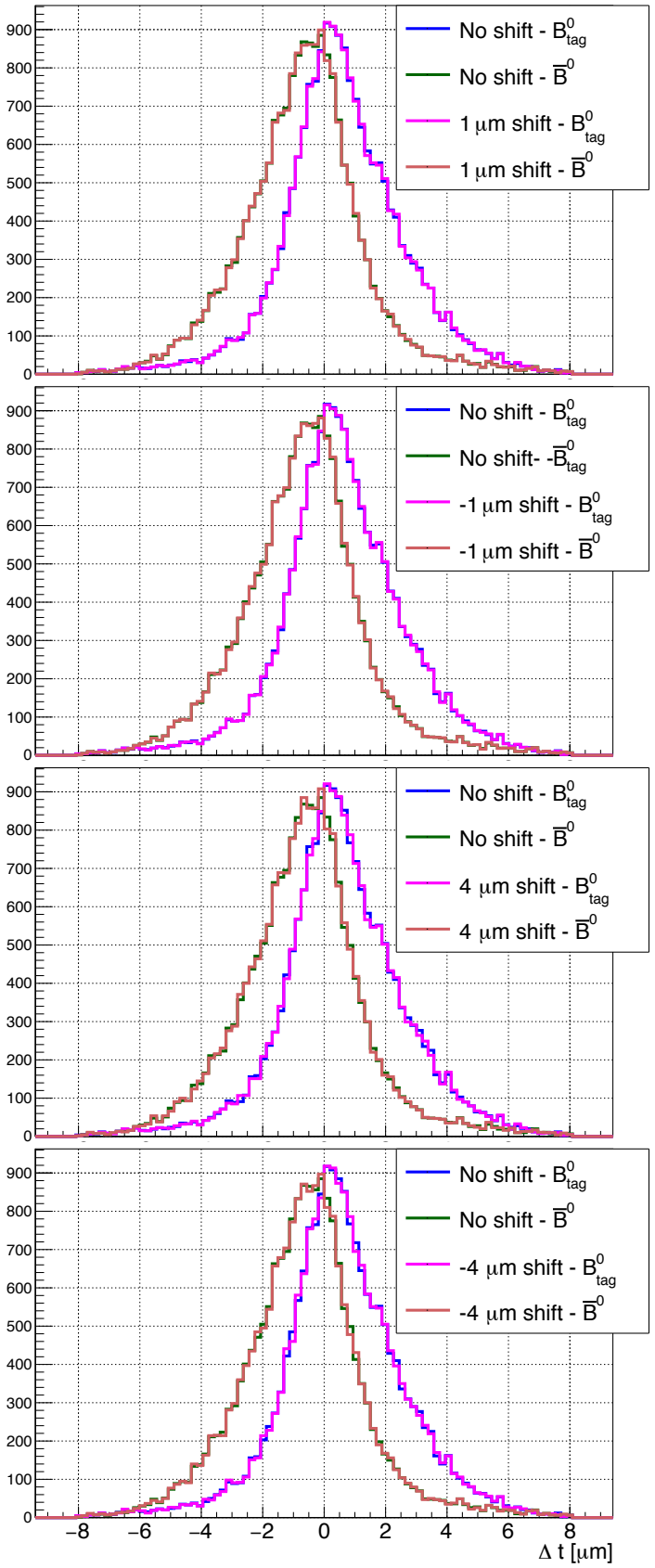


Figure 4.9: Δt distribution of the B-meson system after a shift on the fitting constraint of the Tag Side in the direction orthogonal to the boost. (a) Small positive shift. (b) Small negative shift. (c) Large positive shift. (d) Large negative shift.

CHAPTER 5

New B_{tag} vertex algorithm

In Sec. 3.6 was discussed the most recent results regarding B_{tag} vertex reconstruction. In BelleII the vertexing capabilities has been improved by a long run with respect to Belle. In Figs. 3.9a and 3.9b could be observed the almost factor 3 improvement in the CP Side. The Tag Side, however, has not been improved in the same way. Barely a factor 1.5 is shown in Figs. 3.10a and 3.10b for the B_{tag} vertex resolution. The algorithm responsible of choosing the best tracks for the fit is crucial in order to perform the best fit, and thus taking the maximum advantage from the high resolution capabilities of the PXD.

Throughout this chapter new ideas for the track's selection algorithm will be discussed with the aim of improving the B_{tag} vertex resolution. Firstly, a Monte Carlo study will be made by using the semileptonic decay of B_{tag} in order to ascertain a possible improvement on the vertex resolution by choosing the primary muon as the only track to perform the fit.

Once it is evident the benefits of performing a track selection on the fitting algorithm, a more complete test will be made. This time, $B_{tag} \rightarrow$ generic will be taken with the aim of developing an algorithm that can reject secondary tracks from the fitting. The main idea is to keep as many primary tracks as possible by removing the maximum amount possible of secondary tracks from the fit. For this reason it has been called *Secondary Track Rejection*. This scheme, however, will not work as expected and for that reason new paths need to be taken.

The last section of this chapter will describe the definite algorithm that will be finally implemented in the Belle II vertexing code. The main goal of this algorithm will be to choose only one track (a very clear primary track), and to perform the fit by using only that track. For this reason, it is called *Single Track Fit*.

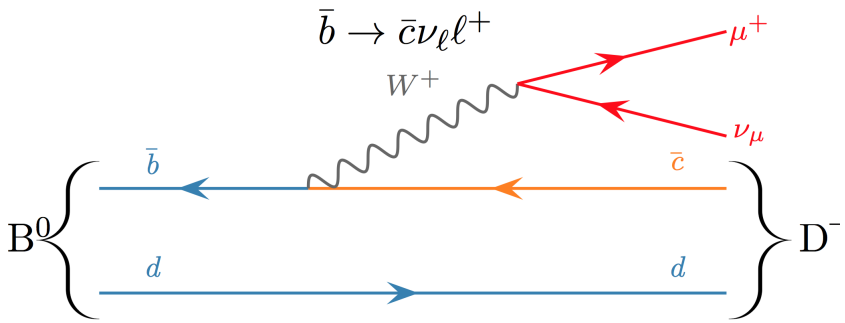


Figure 5.1: Feynman diagram of the semileptonic decay of the B^0.

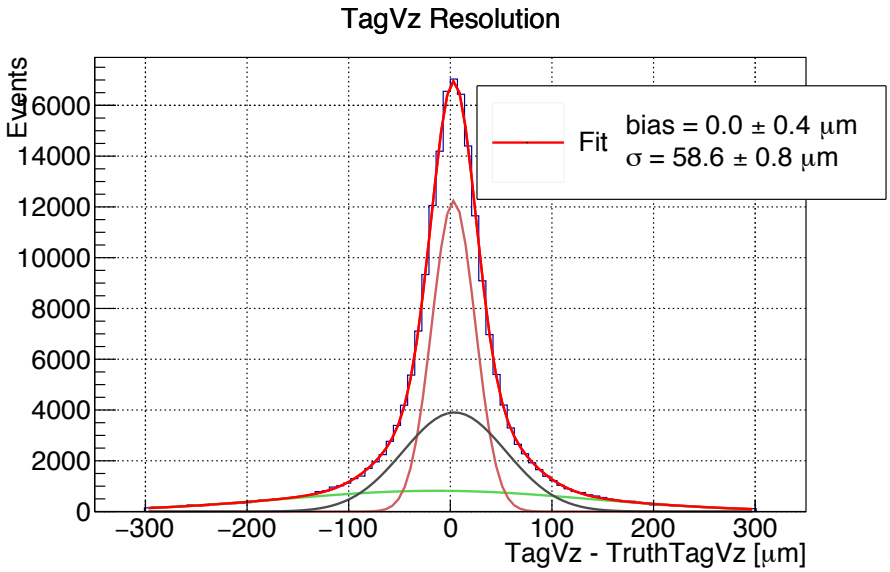
5.1 Preliminary study: semileptonic decay

Before start making changes on the vertexing software, one needs to be sure that a real improvement on the vertex resolution can be obtained. The Monte Carlo study performed along this section provides such evidence.

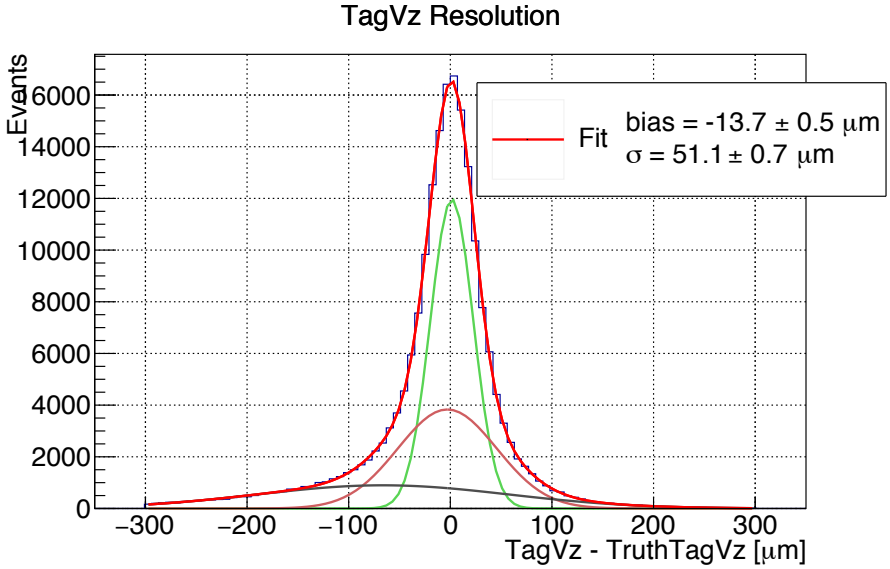
The channels chosen for this analysis are $B_{tag} \to \mu^+ D^- \nu_\mu$ and $B_{tag} \to \mu^- D^+ \bar{\nu}_\mu$. The semileptonic decay is specially convenient for this preliminar analysis. It is a very simple decay where one can find very clear primary tracks, the muons, and secondary tracks coming from the decaying of the D meson. The corresponding Feynman diagram is sketched in Fig. 5.1.

Therefore, the idea is to perform the vertex fit by using only a single track, the muon track. In order to do so, using Monte Carlo information is mandatory. Once the fit is performed in this way, it can be compared with the vertex resolution obtained by using all the tracks remaining after the CP Side reconstruction, as explained in Sec. 3.6.

In Fig. 5.2 it can be observed the results of the study. In 5.2a it is displayed the vertex resolution obtained after using the standard fit algorithm. In 5.2b, instead it is showed the vertex resolution obtained after using only Monte Carlo matched muons for the fit, as explained above. It is clear that there is a noticeable improvement in the resolution. In fact, it has been improved by almost a 20%. However, it can be observed ,together with the improved resolution, a clear negative bias of the center of the distribution. This bias derives from the constraint used for the fit. The constraint (cf. Sec. 3.3.2) is prepared to deal with secondary tracks, and its center is chosen so that the these are specially down-weighted. This process is a success in the *Standard Algorithm*, but when using only primary tracks it gives problems. Indeed, there are no secondary tracks in this kind of fit, and thus there is nothing to be down-weighted. The result is an *over-compensation* during the weighting of the tracks due to the probability distribution generated



(a) After using the *Standard Algorithm*.



(b) After using the only Monte Carlo matched muon tracks.

Figure 5.2: Z component of the B_{tag} vertex resolution for semileptonic events.

by the constraint. This can be observed in Fig. 5.2b.
This is a minor issue that will be solved in what follows, more concretely during the optimization of the *Single Track Fit*. The big picture to be understood from this analysis is that being careful about the tracks chosen to perform the fit pays off during the vertex fitting. The improvement of the vertex resolution in the semileptonic decay motivates the next studies about the generic decay of B_{tag} .

5.2 Secondary Track Rejection

At this stage of the thesis, it is time to start discussing about possible improvements in the track's selection algorithm for the B_{tag} vertex fit. As discussed above, the idea behind *Secondary Track Rejection* is to perform the vertex fit by using a set of tracks where the secondary tracks have been removed from. Indeed, for this method to work, an great amount of cuts that filter the secondary tracks from the rest need to be developed.

In Sec. 3.6 was presented the *Standard Algorithm* used for the B_{tag} Vertex Fit. This algorithm uses all the tracks unused during the CP Side reconstruction. From now on, for both the *Secondary Track Rejection* and the *Single Track Fit*, the initial set of tracks that the algorithm starts working with are not the same as the *Standard Algorithm*. Instead it uses the tracks selected by the *Flavor Tagger* module, described in Sec. 3.4.1. After passing some selection criteria only a few amount of tracks contribute in the vertex fit, ideally the primary ones.

Being a primary track it usually means that it originates in the primary vertex, i.e a direct daughter of the B_{tag} . During this work the meaning of primary has been extended to the tracks coming from immediately decaying daughters of the B_{tag} , e.g $J/\psi, D^*, \omega, \eta...$ Those particles have a very short lifetime because they decay via EM or strong interaction. For all practical purposes, the tracks coming from those particles are produced where the B_{tag} vertex is, as if they were direct daughters of the B meson.

Secondary tracks are defined as the ones coming from a secondary vertex that can be distinguished from the primary vertex. From the possible decays of the B meson[28], only a few meson fly enough time to generate secondary vertices and thus secondary tracks from them. Those mesons are:

$$K^0_L, D^0, D^{\pm}, D^{\pm}_s \tag{5.1}$$

5.2.1 Primary Track fit

A good starting point for the track's selection analysis is to perform a fit by using the Monte Carlo information to use only primary tracks. Indeed, this fit is merely informative for no real data analysis will count on Monte Carlo information. In any case, it is still a valuable piece of information since it will result in the best fit that can be achieved. There is no way that by using this type of track selection a better resolution is obtained.

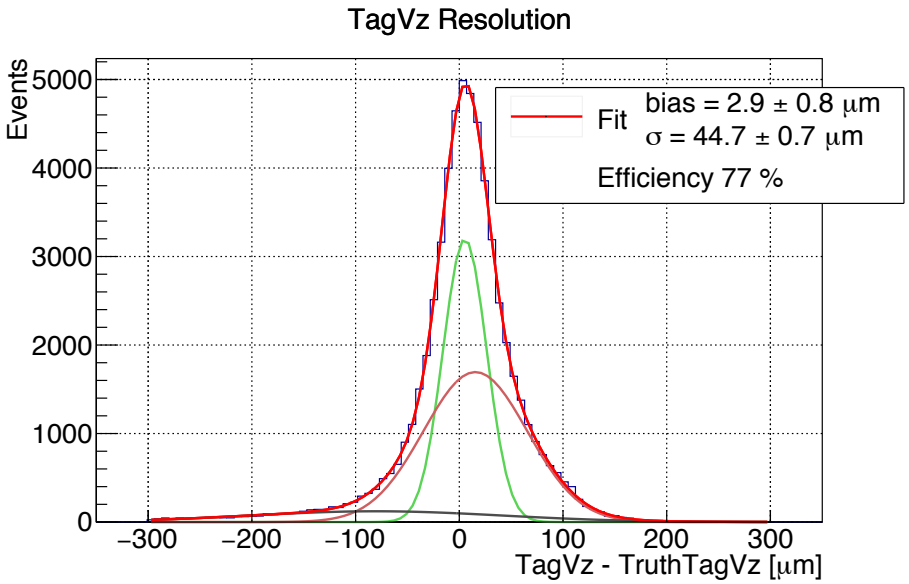


Figure 5.3: Z component of the B_{tag} vertex resolution. The fit has been performed by using only Monte Carlo Matched primary tracks, that is, tracks coming directly from the B_{tag} .

In Fig. 5.3 has been plotted the resolution of the B_{tag} vertex when the fit only uses Monte Carlo matched primary tracks. The result is a much better resolution that the last value obtained in Belle II by using the *Standard Algorithm* (cf. Fig. 3.10a). The efficiency is not very high because the *Flavor Tagger* does not always contain primary tracks, since is not specially its purpose. Although the resolution has improved notably, the fact that the efficiency is so low implies an enormous loss of statistics that makes this result unacceptable.

Nevertheless, this is not the end of the analysis. If the selection criteria is good enough, the *Secondary Track Rejection* can be combined with the *Standard Algorithm* in order to improve the efficiency. A similar problem is faced with the *Single Track Fit* regarding the efficiency of the algorithm, and it is solved easily by combining it with the *Standard Algorithm*. This will be treated in following sections.

5.2.2 Selection Criteria

The selection criteria for the tracks revolves around their properties carried by the *FlavorTagger*, like *TargetProbability* or *CategoryProbability*, all of them described in Sec. 3.4.1. An exhaustive Monte Carlo analysis of the tracks contained in the *FlavorTagger* shows that the most reliable categories to work with are Electron and Muon. Generally, a good *Category / TargetProbability* is a clear indicator of

the tracks (most of the cases electrons and muons, as expected) being primary. Therefore, a great part of the cuts are explicitly designed to either leave good muon or electron candidates or remove tracks that are surely not containing a primary muon or electron. The parameters most used are the momentum, the Target Probability and the Category Probability, together with the Impact Parameters of each track.

At the same time, the Monte Carlo study revealed that after the first cuts, the tracks contained inside other categories like Kaon or Fast Pion carried poor or insufficient information about whether they are primary or secondary. This means that distinguishing between primary Fast Pions and secondary Fast Pions, for example, is not an easy task at all, being in some cases even impossible.

The developed cuts are the following:

1. **Impact Parameter Cut**: D_0, Z_0 for each track must be at least less than 400 µm to be considered as primary track. By that distance the majority of B mesons must have decayed already.

2. **Lambda Cut**: The Monte Carlo study revealed that all the tracks associated with the Lambda category given by the *Flavor Tagger* correspond to secondary tracks. Therefore, removing those is an essential step.

3. **Very clear primary track**: there are several decays, like the semileptonic, where there is only a primary track which is very clear. If the track associated to Muon or Electron categories has a high Target and Category Probability (> 0.7), remove all the other tracks and keep only that one.

4. **Muon Momentum Cut**: In the semileptonic decay, the correspondent lepton has a relatively high momentum, since it comes directly from the decay of the B meson. For that reason, forcing the tracks in the Muon Category to have a certain amount of momentum is a good idea. In the cases where the track within the Muon Category does not correspond to the track within MaximumP: if $p(\mu) - p(MaxP) - p(e) < 3\,\mathrm{GeV}/c^2$, discard the Muon track.

5. **Electron momentum Cut**: analogous to the Muon Momentum Cut, but using the Electron Category instead.

6. **Muon Category Cut**: primary muon give a very clear signal. Thus, if the Category Probability of Muon is not higher than the sum of Category Probability of Kaon and Slow Pion summed, discard the muon.

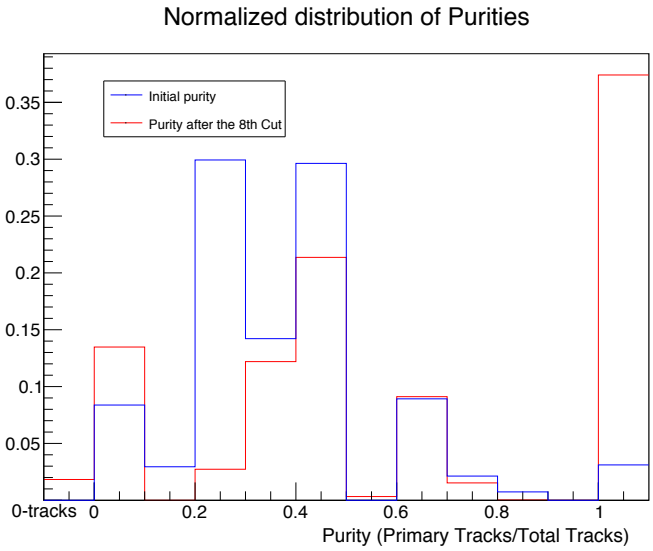


Figure 5.4: Normalized distribution of primary tracks' purity. The purity is defined as the ratio between primary tracks and the total amount of tracks available for the fit. Purity 1 means all the tracks are primary, and purity 0 means that the tracks are instead secondary. Anything inbetween corresponds to a mixture of primary and secondary tracks. The bin located at the extreme left side of the plot corresponds to the events that after the selection criteria end up without any track, nor primary nor secondary.

7. **Electron Category Cut**: similar to Muon Category cut, but for the Electron.

8. **Impact Parameter Comparison Cut**: compare D_0, Z_0 values among the tracks. If one track has an impact parameter much higher ($> 200\,\mu m$) with respect to another track, it means that the former track is probably secondary and the second is primary. Thus, remove the first track.

To understand the effect of the aforementioned cuts, it is necessary to define what has been called *Purity* of the tracks. This concept represents the quotient between the MC matched primary tracks within the group of available tracks and the total amount of tracks in such a group. In Fig. 5.4 it is portrayed the comparison between the purity before and after applying the selection criteria. There can be observed the great improvement regarding the purity of the tracks after the cuts are applied. It is specially remarkable how the bin corresponding to purity 1 has passed from less than 0.05 to almost 0.4. However, the bin corresponding to purity 0 has increased as well. This last bin corresponds to the case were there are no more primary tracks remaining, and that the fit will be done using all secondary tracks.

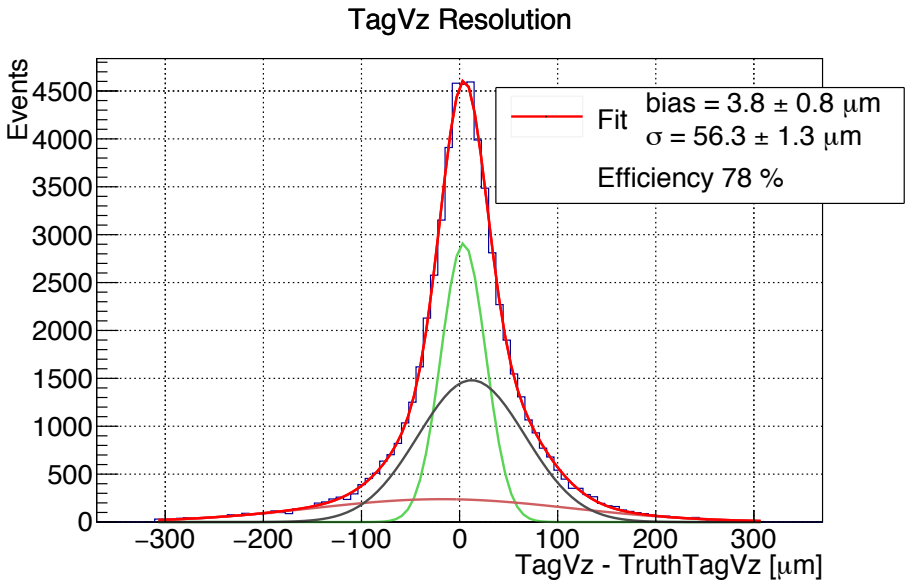


Figure 5.5: Z component of the B_{tag} vertex resolution. The fit has been performed by using the tracks obtained via *Secondary Track Rejection* algorithm, after performing the eight cuts. During this fit, events with no primary tracks left after the reconstruction of the B_{CP} have been ignored, which implies a loss of efficiency and an improvement of the fit.

After the eighth cut, keep removing secondary tracks becomes not trivial. As commented before, the secondary tracks and primary tracks lying within Kaon, Fast and Slow Pion Categories are not easily distinguished at all. They have very similar momentum and probabilities, since those categories created for the *Flavor Tagger* are not focused on finding whether the tracks are primary or not, but instead they extract the flavor information. For this reason, creating new cuts is a futile task. The results of the fit after using the cuts can be observed in Fig. 5.5.

It can be appreciated that the fit has not improved. In fact, the resolution has worsened with respect to the *Standard Algorithm* fit shown in Fig. 3.10a. There are some reasons that explain this, a priori, unexpected result. All the events where the purity reaches the 1, those are in the best case scenario contributing as depicted in Fig. 5.3. The events with a purity $\in (0, 1)$ are fitted by using some primary and some secondary tracks. Those cases end up fitting the vertex as the *Standard Algorithm* does, with the difference that now there are much less tracks performing the fit. Therefore, one can expect some results similar to Fig. 3.10a. The remaining events correspond to the purity 0 bin. Those events contain only secondary tracks, which is the worse possible case. The constraint is not anymore capable of down-weight secondary tracks because there are no primary to be compared with. The main question now could be, what would happen if the fit was performed only with secondary tracks? The answer is shown in Fig. 5.6. In

this fit no primary tracks have been used, and one can see how the resolution is dramatically worsened.

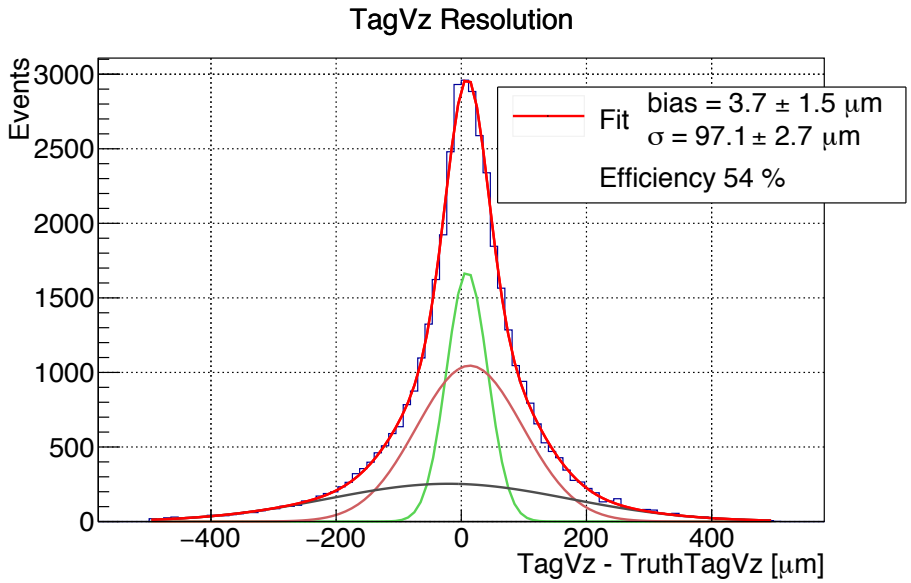


Figure 5.6: Z component of the B_{tag} vertex resolution. The fit has been performed by using only Monte Carlo matched secondary particles, that is, with tracks that do not origin in the B_{tag} vertex.

Those contributions can be weighted in terms of the percentage of events that are representing. A rough estimation could be:

$$35\% \times 45\,\mu\text{m} + 15\% \times 97\,\mu\text{m} + 50\% \times 53\,\mu\text{m} \approx 57\,\mu\text{m} \tag{5.2}$$

Which agrees totally with the results obtained. Therefore, the events containing no primary track makes the fit unacceptable, even if a high percentage of the remaining events have purity 1. It should be noted that, as shown in Fig. 5.4, even without applying any cut, because of the way the *Flavor Tagger* assigns the tracks to the categories, a lot of events have no primary tracks. Thus, a *Secondary Track Rejection* algorithm will never work in a scenario where the initial set of tracks is obtained via the *Flavor Tagger*.

5.3 Single Track Fit

The *Secondary Track Rejection* algorithm has not worked as expected. Nevertheless, It can be observed in Fig. 5.3 that it is still possible to perform a better fit than the *Standard Algorithm*. The *Standard Algorithm* has its strength on the

constraint used for the fit. Even if there is only a single primary track among all the available tracks, the algorithm will most probably up-weight this track. But what happens with the *Secondary Track Rejection* is that some events carrying a few primary tracks get those erased by mistake. When no primary tracks are present, the constraint is useless and the result of the fit is degraded (cf. Fig. 5.6).

The solution comes from carefully choosing the primary tracks, instead of trying to discard the secondary ones. Moreover, to select the primary track a very strict selection cut is performed to the tracks. This way of proceeding has the advantage that either it finds a very clear primary track to perform the fit, or it just leaves the whole set of tracks for a Standard Fit.

It is worth to note that, as the name of the algorithm says, the fit is performed by using only a single track: the primary track. The initial set of tracks where the single track is going to be extracted from comes from the *Flavor Tagger*, in the same manner as it has been done for the *Secondary Track Rejection* algorithm.

5.3.1 Selection Criteria

As has been discussed in Sec. 5.2, the best candidates to perform the fit are the tracks within the Muon and Electron categories. For this reason, the *Single Track Algorithm* will rely basically on finding those muons and electrons coming from the semileptonic decay, which give a very clear signal and are easily distinguished from the other tracks provided by the *Flavor Tagger*.

In Fig. 5.7 it is depicted a schema of the algorithm. As it will be discussed in the following section, to prevent a loss of efficiency the *Standard Algorithm* is going to be used in all the cases that the *Single Track Fit* cannot be performed. The selection criteria for the primary tracks is simpler but more restrictive than the cuts applied in the *Secondary Track Rejection*. Those criteria are the following:

1. **Pixel Vertex Detector Hits condition**: to be sure that the good vertexing capabilities of the Belle II detector are used, the tracks need to leave at least one hit in the PXD.

2. **Impact Parameter Condition**: $D_0, Z_0 < 500\,\mu m$ is a necessary condition for the track to be primary. This condition was also included in the previous section.

3. **Momentum Condition**: although is not always the case, the primary lepton in a semiletonic decay usually carries the highest momentum. Therefore,

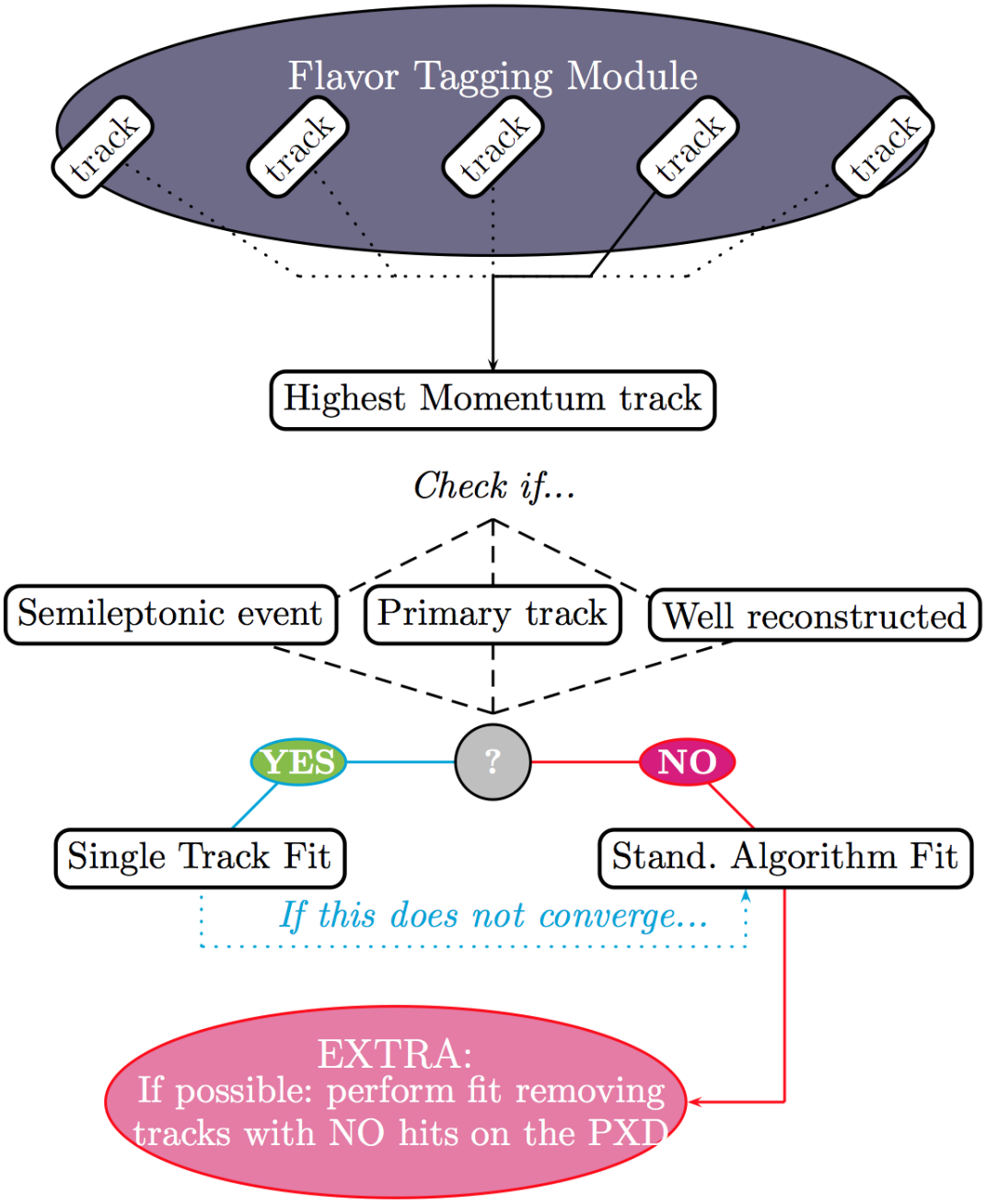


Figure 5.7: Z component of the B_{tag} vertex resolution. The fit has been performed by using only Monte Carlo matched secondary particles, that is, with tracks that do not origin in the B_{tag} vertex.

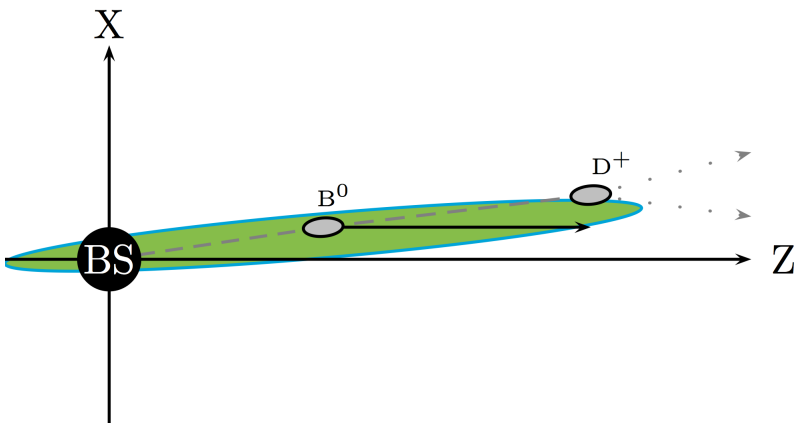


Figure 5.8: Redefinition of the parameters of the fit's constraint. It has been shortened and shifted 120 µm in the boost direction. This is the average distance that the B meson travels before decaying.

the track within Electron/Muon category have to be the same as the one within MaximumP.

4. **Category Condition**: the Category Probability needs to be > 0.7 in order to assure that the event is semileptonic.

5. **Target Condition**: together with the previous condition, it is necessary for the track to have a high probability to be the primary lepton in the semileptonic event, i.e > 0.5.

The minimum values for the probabilities have been chosen by trial and error. A multivariate analysis would have been an interesting way to find out the best combination for the fit. However, the fact that the fit needs to be performed to train the multivariate method implies a lot of computing time, and therefore has been left as an interesting continuation for the present work.

In Fig. 5.8 shows the changes that have been implemented regarding to the constraint of the fit. Since during the *Single Track Fit* there are no secondary tracks to be down weighted, the center of the constraint needs to be exactly where the B is expected to decay in average. It should be noted that for this type of fit involving only one track, the fitting constraint is not only advisable, but completely necessary. Without a probability distribution for the vertex in the space, a single track is not able to reconstruct it.

The shift in the constraint center shown in Fig. 5.8 corresponds to the length travelled by the B meson before decay. This average length can be calculated

easily from its $c\tau$ value [28] and the boost provided by SuperKEKB, $\beta \approx 0.27$ (cf. Sec. 1.3.3):

$$\ell_{\mathrm{B}^0} = c\tau_{\mathrm{B}^0} \cdot \beta = 455.4\,\mu\mathrm{m} \cdot 0.27 \approx 120\,\mu\mathrm{m} \tag{5.3}$$

5.3.2 $\mathrm{B_{tag}}$ Vertex Resolution

In Fig. 5.9 the results of the $\mathrm{B}_{\mathrm{tag}}$ vertex resolution after the *Single Track Fit* can be observed. Fig. 5.9a shows the results of the fit when only the *Single Track Fit* is used. Indeed, the efficiency is low, but this is completely understandable since only a few percentage of events fulfil the conditions for the muon/electron track selection.

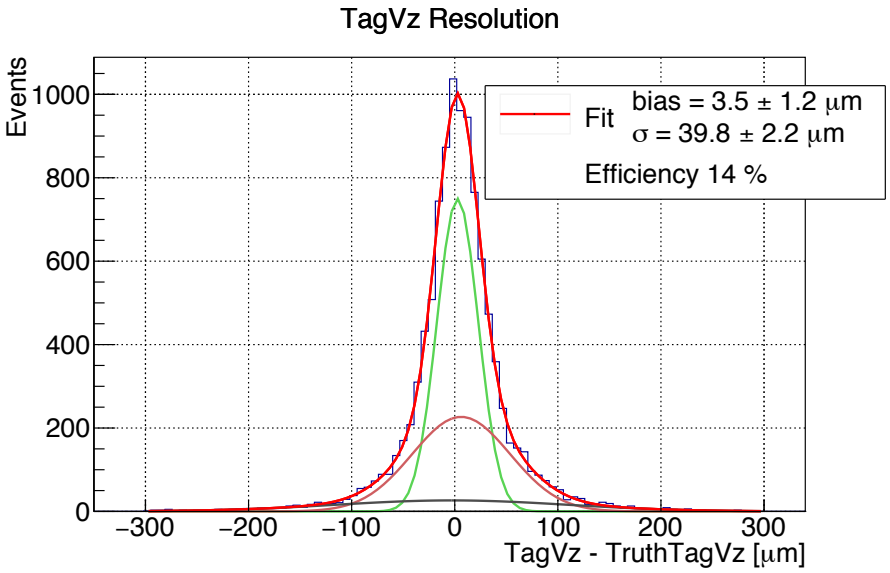
In order to deal with the low efficiency, as suggested in Sec. 5.2, a good idea is to combine this algorithm with the *Standard Algorithm*. That is, whenever the *Single Track Fit* fails to find a primary track, the *Standard Algorithm* is responsible for the fit. Moreover, even if the *Single Track Fit* manages to find the track, a non converging fit would trigger the *Standard Algorithm*.

In Fig. 5.9b it can be observed the final result of the resolution after combining both algorithms. It is clear that the resolution has improved notably at the same time that the efficiency has remained high enough. In fact, the improvement is of a factor 2.2 when using only the *Single Track Fit*, and 1.9 when using the combined algorithm, with respect to the Belle results. This improvement now is closer than before to the one obtained for the CP Side (cf. Fig. 3.9a), which was about a factor 2.7. In the next section, this improvement is translated to the Δt resolution.

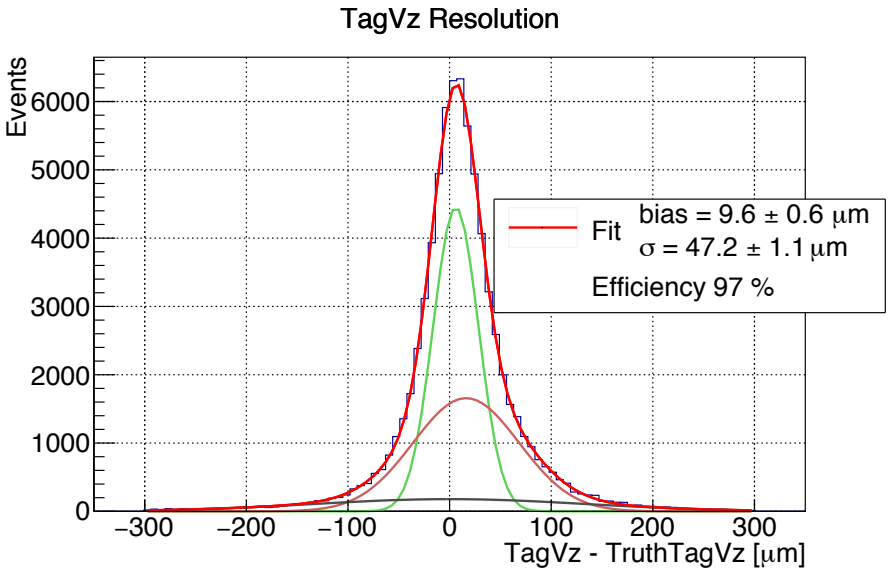
5.3.3 Δt Resolution

In Fig. 5.10 is shown the Δt resolution after the fit.

Again, the improvement is quite notable if one compares with the results in Belle. The fit performed by using the *Single Track Fit* alone depicted in Fig. 5.10a has an improvement of a factor 1.6. In Fig. 5.10b it can be seen the case when this algorithm is combined with the *Standard Algorithm*, leading to an improvement of a factor 1.4. Must be noted that the gain in terms of Δt is not comparable to the one obtained for the vertex resolution. The reason is the different boost applied in Belle and Belle II, and has already been discussed in Sec. 3.6.

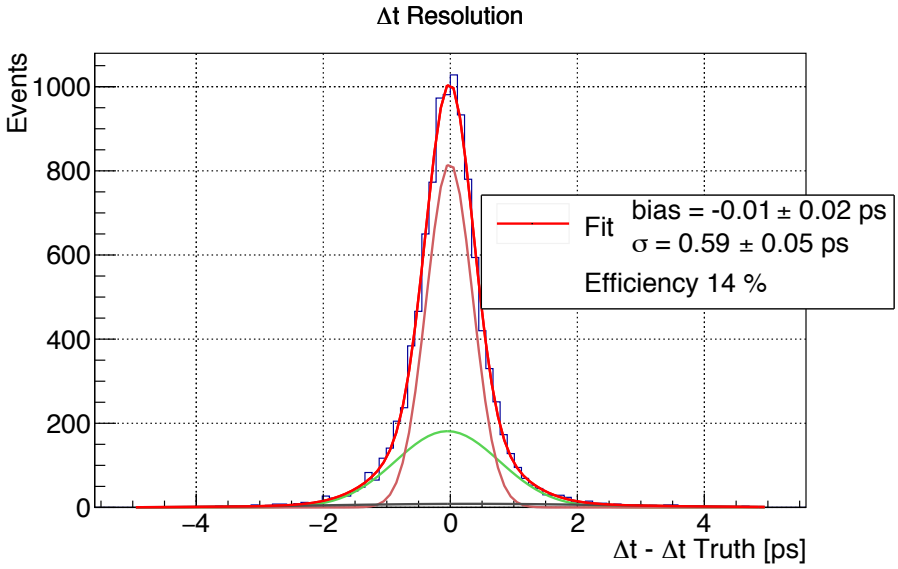


(a) After using the *Single Track Fit* alone.

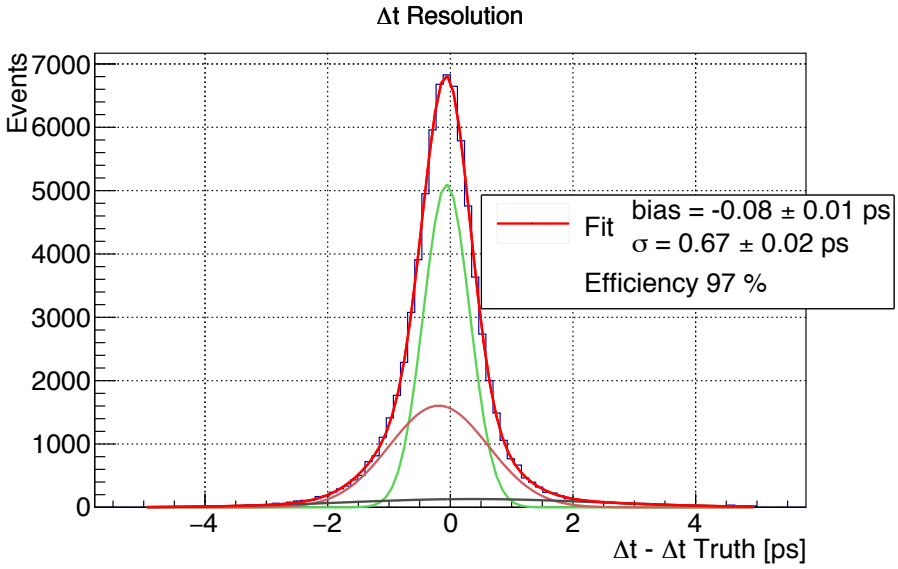


(b) After using the *Single Track Fit* combined with the *Standard Algorithm*

Figure 5.9: Z component of the B_{tag} vertex resolution. Results obtained when using *Single Track Fit*.



(a) After using the *Single Track Fit* alone.



(b) After using the *Single Track Fit* combined with the *Standard Algorithm*

Figure 5.10: Δt resolution obtained when using the *Single Track Fit* for the B_{tag} vertex fit.

Conclusions

The main objective pursued throughout the present thesis has been to analyse the actual performance of the Tag Side vertexing software of Belle II, and to try different ways to improve it. For the time-dependent analysis scheme presented in this work, one of the B mesons is chosen to decay into the Golden Channel $B^0 \to J/\psi K_S$, while the other B meson is chosen to decay into any possible channel. This second meson is usually called B_{tag} because it is used for flavor tagging. Two main tasks have been performed along the present work regarding the B_{tag} vertexing optimization.

First of all, a deep study about the B_{tag} vertex fit's constraint has been done. The main idea behind this study has been to understand the effect of possible uncertainties on the knowledge of the Beam Spot, i.e where the beams intersect and where the constraint is centred. Moreover, it is interesting to understand the capabilities of the *Standard Algorithm* used for fitting, to know where it fails and thus what are its limitations. This has been accomplished by performing the vertex fit for different constraint's center values. Apart from extreme cases, the shift has not perturbed either the B_{tag} vertex or Δt resolutions. Only with the most extreme shifts, like the one of 200 µm, a noticeable bias in the distributions can be observed, although an almost negligible one. Regarding the fit's constraint, it has been concluded therefore that the *Standard Algorithm* and its constraint are very stable and end up being more reliable than what was expected initially.

Secondly, a new algorithm to select the tracks that are used for the B_{tag} vertex fit has been implemented. The main idea behind the new algorithms developed is to select only tracks coming directly from the B_{tag} , i.e primary tracks. Initially, the track's selector algorithm was aimed towards discarding the secondary tracks, so that only primary tracks were left after applying a determined set of selection cuts. This idea however, which is called *Secondary Track Rejection*, turned out to be unsuccessful. The selection criteria used several parameters of each track given by the *Flavor Tagger*. This module is designed to find the most suitable tracks to perform the flavor tagging, which do not correspond necessarily with the

tracks that are primary. Therefore, using the information provided by the *Flavor Tagger* to reject secondary tracks ended up with some false positives that leaded to events with no primary track at the moment of the fit. The B_{tag} vertex resolution has worsened when using the *Secondary Track Rejection* algorithm, forcing the exploration of other ideas. The *Single Track Fit*, instead has been very successful regarding the improvement of the vertex resolution. As can be understood from its name, this new algorithm performs the fit by using only one track. Instead of discarding questionable secondary tracks, very clear primary tracks such as the leptons in the semileptonic decay are selected. Table 5.1 shows the summary of the best results in the time-dependent analysis achieved so far in the simulations of Belle II, including the improvements obtained with the new algorithm with respect to the Belle experiment.

Measurable (Res)	Result	Improv. Factor (wrt. Belle)	Algorithm
B_{CP} Vertex	22.3± 0.1 µm	2.7	Standard
B_{tag} Vertex	52.8± 0.3 µm	1.5	Standard
Δt	0.78±0.1 ps	1.2	Standard
B_{tag} Vertex	39.8/47.3± 0.3 µm	2.2/1.9	Single Track Fit
Δt	0.59/0.67±0.2 ps	1.6/1.4	Single Track Fit

Table 5.1: List of results obtained after the vertex fitting for the time-dependent simulation analysis in Belle II. Both the results of this thesis and the results obtained previously have been included.

The improvement obtained with the *Single Track Fit* is a very promising result. Nevertheless there are still several issues that need to be addressed:

1. **Improvement factor**: the increased Tag Side vertex resolution is still not at the same level as the one obtained for the CP Side. Belle II is still not taking full advantage of the newly developed Pixel Vertex Detector because of the high dependence on the fitting algorithm, and this needs to be addressed (see next point).

2. **Low efficiency**: the *Single Track Fit* is being triggered only around a 15% of the times. This is expectable because the track's selection is only looking

for primary leptons, most probably coming from semileptonic events. The solution would be to extend the search of primary tracks to other possible decays that are clear enough. However, extracting the information needed for each track will be difficult because they come from the *Flavor Tagger*, a problem discussed above.

3. **Large bias**: the fit observed in Fig. 5.9 has a non-zero bias in the positive side. This bias will need to be fixed at some point, although is not the top priority at this stage. The solution could come from editing the constraint's parameters, probably enlarging it on the boost direction so it deals better with exceptionally long-lived B mesons.

Appendices

APPENDIX A

Kinematic study

In this appendix, the kinematics of the $\mu^+\mu^-$ production will be treated. It will be focused on the angular distribution of the pairs produced after the collision of e^+e^- and the decay of the J/ψ. The interest of this appendix is to gain some insight on the possible differences arising between both aforementioned cases that could possibly explain the different vertex reconstructions discussed in Secs. 3.5 and 4.1.

As explained in Sec. 2.1.1, the electron and positron beams have an energy of 7 and 4 GeV respectively. The center of mass energy is thus $\sqrt{s} = E^* = 11\,\text{GeV}$. The effective momentum that generates the boost in the laboratory frame is $|\mathbf{P}| = |\mathbf{p}_{e^-} + \mathbf{p}_{e^+}| = 3\,\text{GeV}$[1] in the electron direction. This means that the boost parameters are:

$$\beta\gamma = 0.28 \rightarrow \beta = \frac{E^*}{P} \approx 0.27, \quad \gamma \approx 1.03 \tag{A.1}$$

A.1 First Scenario: $e^+e^- \rightarrow \mu^+\mu^-$

A.1.1 Step 1: Center of Mass frame of e^+e^-

From simple energy and momentum conservation in the center of mass frame:

$$E^* = E^*_{\mu^-} + E^*_{\mu^+} \tag{A.2}$$

$$\mathbf{p}^* = \mathbf{p}^*_{\mu^-} + \mathbf{p}^*_{\mu^+} = 0 \tag{A.3}$$

And assuming $m_\mu = 0$ for this analysis, the result is:

[1] For this appendix the usual notation of $c = 1$ has been used.

$$
\begin{aligned}
E^*_{\mu^-} = E^*_{\mu^+} = E^*/2 = 5.5\,\mathrm{GeV} = |\mathbf{p}_{\mu^-}{}^*| = |\mathbf{p}_{\mu^+}{}^*| \\
\mathbf{p}_{\mu^-}{}^* = -\mathbf{p}_{\mu^+}{}^*
\end{aligned}
\tag{A.4}
$$

This result is the standard result obtained for a two-body decay with massless particles. From now on, the absolute magnitude of the momentum will be represented as $p = |\mathbf{p}|$ to avoid overcharging the nomenclature.

A.1.2 Step 2: Laboratory frame

In order to apply the effective boost determined in Eq. A.1, one needs to apply a Lorentz transformation:

$$
\Lambda = \begin{pmatrix} \gamma & 0 & 0 & -\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\beta & 0 & 0 & \gamma \end{pmatrix}
$$

where it has been assumed that the boost is performed in the Z-axis for simplicity. The transformation is applied in the following way:

$$
\mathcal{P}^{\sigma}_{lab} = (\Lambda^{\sigma}_{\delta})^{-1}\mathcal{P}^{\delta}_{cm} \tag{A.5}
$$

where $\mathcal{P}_{lab} = (E'_P, \mathbf{p}'_{\mathbf{P}})$ corresponds to the transformed 4-momentum of the produced P particle. In an homologous way happens for $\mathcal{P}_{cm}$. It has been used the inverse Lorentz transformation because one needs to change from the moving frame to the steady frame. The projection of the momentum in the Z-axis and the energy become:

$$
\begin{aligned}
p_{\mu^-}\cos\theta_{\mu^-} = p_{z,\mu^-} = \gamma(E_{\mu^-}\beta + p^*_{\mu^-}\cos\theta^*) \\
p_{\mu^+}\cos\theta_{\mu^+} = p_{z,\mu^+} = \gamma(E_{\mu^+}\beta - p^*_{\mu^+}\cos\theta^*)
\end{aligned}
\tag{A.6}
$$

And,

$$
\begin{aligned}
E_{\mu^-} = \gamma(E^*_{\mu^-} + \beta p^*_{\mu^-}\cos\theta^*) \\
E_{\mu^+} = \gamma(E^*_{\mu^+} - \beta p^*_{\mu^+}\cos\theta^*)
\end{aligned}
\tag{A.7}
$$

Here $\theta^* \in [-\pi, \pi]$ is the angle in the center-of-mass frame subtended with respect to the Z-axis. For the μ^+ it is assumed that the subtended angle goes like $\cos(\theta^* + \pi) = -\cos\theta^*$ for momenum conservation. The primed variables are taken

to be the ones corresponding to the lab frame, where the angle with respect to the Z axis is not necessarily related by a simple addition of π. From now on, only the case of the particle (in the case μ^-) and not the antiparticle will be treated for simplicity. The results, however, are equivalent. For this purpose, $\cos\theta_\mu = \cos\theta$ for simplicity.

Combining Eqs. A.7, A.6 together with the massless assumption leads to a very simple formula for the angular distribution in the plane XZ:

$$\cos\theta = \frac{\cos\theta^* + \beta}{1 + \beta\cos\theta^*} \tag{A.8}$$

Some information about the direction of both muons can be extracted from this expression. Indeed, the first thing one notices is that the laboratory angle depends on the CoM angle and the boost applied to the system. Moreover the formula shows that necessarily there are events where both muons are emitted in the same direction. The condition for this to happen is:

$$\cos\theta^* < \beta \tag{A.9}$$

This condition is perfectly achievable, since $\beta = 0.27$ is high enough. Thus, the higher the boost is, the more are the chances for an event to have both muons going in the same direction. This can be observed in Fig. A.1, where some events (although not the majority), are in the positive-positive region of $\cos\theta_1 - \cos\theta_2$. Those cosines correspond to both muon angles with respect to the Z-axis in the laboratory frame.

There are some things to point out about Fig. A.1. First of all, besides from the "Bow" form on the plot, it can be seen an "Arrow" crossing the graph. This arrow corresponds to $\cos\theta_1 = \cos\theta_2$, which is indeed impossible. Those are events where both muons have been associated with the same track, most probably because the other track has not been well reconstructed. Another thing that needs to be pointed out is the fact that there is a cloud of dots scattered around the whole graph, including the negative-negative region. Indeed, this region is forbidden since the boost is positive and at least one muon, the one going in the positive Z direction in the CoM frame, has to be emitted forward. This cloud exists because the track reconstruction is not always as good as desired. Sometimes the momentum reconstructed simply do not match with momentum expected from momentum conservation. Those tracks can be removed easily when performing the reconstruction of the decay. Finally, it can be noted that the "Bow" is not a very defined curve, but has a certain width. This phenomena comes from the fact that some radiative photons can be emitted from the muon while it travels, taking some momentum from it.

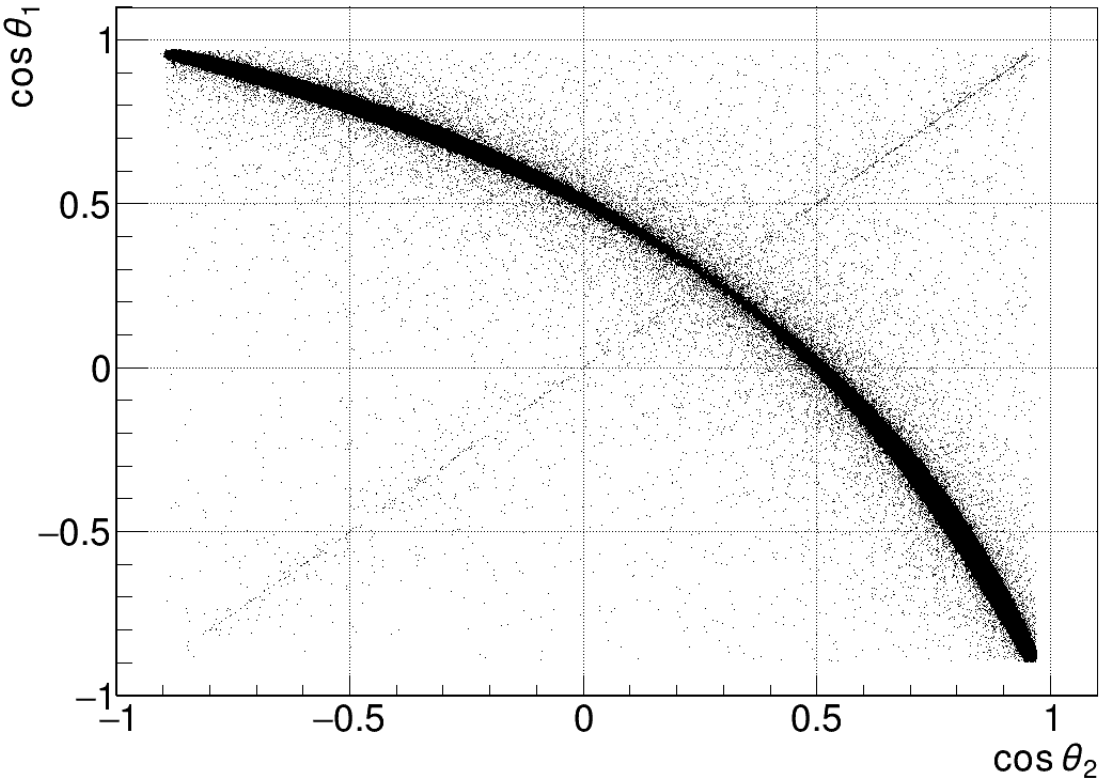


Figure A.1: Graph showing the directions in the Z axis of both muons coming from the process $e^+e^- \to \mu^+\mu^-$, by plotting $\cos\theta_1$ vs $\cos\theta_2$. Labels 1,2 stand for each muon.

A.2 Second scenario: $J/\psi \to \mu^+\mu^-$

A.2.1 Step 1: Center Of Mass $\Upsilon(4S) \to B^0\bar{B}^0$

In this scenario, there are two main differences with respect to the one presented in Sec. A.1.1. First of all, the B mesons cannot be considered massless as it has been assumed for the muons. Secondly, the $B\bar{B}$ as a consequence of being massive, are produced practically at rest. Again, from simple energy conservation:

$$E^* = m_\Upsilon = 10.58\,\text{GeV} \quad E_B^* = m_\Upsilon/2 = 5.29\,\text{GeV} \tag{A.10}$$

Therefore, from the mass-energy-momentum relation:

$$p_B^* = \sqrt{E_B^{*2} - m_B^2} \approx 0.34\,\text{GeV} \tag{A.11}$$

Which is barely three times higher than the mass of the muon, and thus can be neglected in the same way. This will simplify greatly future calculations without changing the physics.

A.2.2 Step 2: Laboratory frame $\Upsilon(4S) \to B^0\bar{B}^0$

At this point, the Lorentz Transformation in Eq. A.1.2 can be used again to move from the center of mass frame to the laboratory frame. Since we are assuming the B mesons are at rest in the CoM frame, the energy and momentum in the laboratory frame acquire a specially simple form:

$$E_B = \gamma E_B^* = 5.61\,\text{GeV}, \quad p_B \cos\theta_B^* = \gamma\beta E_B^* = 1.51\,\text{GeV} \tag{A.12}$$

Which basically means that no B mesons are ever emitted backwards, since $\cos\theta_B^*$ will always be positive. This is very obvious since the B mesons are produced at rest in the CoM frame and thus there is no negative contribution to the final momenta after the boost. Since θ does not depend on any CoM angle, in principle both B mesons would be emitted parallel in the Z axis. Indeed, this is not true in Belle II. In this analysis both B mesons are assumed to decay completely at rest in the CoM which is essentially not true in reality. Although the momentum is rather low, is not zero and therefore both B meson paths will be separated by a given angle in the laboratory frame. Otherwise it would not be possible to distinguish one from the other. Nevertheless, for the present analysis it is assumed that both B are emitted along the Z axis without any angular distribution.

A.2.3 Step 3: Center Of Mass $B^0 \to J/\psi K_S$

At this point, the scheme is rather similar to the other CoM analysis that have been done. This time the B meson decays into a pair of particles that do not have the same mass, i.e $J\psi$ with 3.096 GeV and K_S with 0.497 GeV [28]. This time the CoM energy is much less than the previous case, and for this reason the Kaon mass cannot be neglected. The general formula for a two body decay goes as follows [45]:

$$p_{J\psi}^* = p_{K_S}^* = \frac{\sqrt{(m_B^2 - (m_{J\psi} + m_{K_S})^2)((m_B^2 - (m_{J\psi} - m_{K_S})^2)}}{2M_B} = 1.68\,\text{GeV} \tag{A.13}$$

With this momentum the energy of the J/ψ can be obtained as well from Eq. A.11. It is $E_{J/\psi} = 3.52\,\text{GeV}$. The angular distribution of the 2-body decay is not isotropic anymore, since this time both particles cannot be considered massless. In fact, the decaying ratio needs to be integrated over $\cos\theta^*$, taking into account possible dependencies of $\mathcal{M}$ in θ^* which depend on Feynman diagram calculations:

$$d\Gamma(B^0 \to J/\psi K_S) = 2\pi \frac{p_{J/\psi}}{32\pi m_B^2} |\mathcal{M}|^2 d\cos\theta^* \tag{A.14}$$

This calculation is outside the scope of the present appendix and will be skipped. In any case, it is not of great relevance for the following sections nor for the final results.

A.2.4 Step 4: Laboratory frame $B^0 \to J/\psi K_S$

The B decays with a non-zero momentum along the Z axis, which ends up giving a boost to the decaying particles. From the values shown in Eq. A.12, the boost corresponds to $\beta \approx 0.27$ again. Following the same scheme as before by using the Lorentz Transformation in Eq. A.1.2, the energy and momentum in the laboratory frame can be obtained (cf. Eqs. A.7, A.6). The big difference between the decay of $\Upsilon(4S)$ or the production of $e^+e^- \to \mu^+\mu^-$ is that the J/ψ can be emitted backwards. From Eq. A.6, one can extract the exact condition for this to happen:

$$p^*_{J/\psi} \cos\theta^*_{J/\psi} > \beta E^*_{J/\psi} \quad \text{for } \theta^*_{J/\psi} < 0 \tag{A.15}$$

This will not be the majority of the cases but it definitely can happen. The implications of this fact for the angular distribution of the muons are crucial. The CoM frame of the decaying J/ψ is similar to what is discussed in Sec. A.1.1, with the particularity that the boost now depends on $\theta^*_{J/\psi}$ and can be much higher than before. From the values of the momentum and energy of the J/ψ it is easy to find that the boost will range from $\beta \approx 0.7$ to $\beta \approx 0.3$. From Eq. A.8 for higher β is more probable that both muons are emitted together (i.e the possible angles θ^* for which both muons can be emitted in the same directions increase, cf. Eq. A.9). Therefore, either going backwards or forward, the muons coming from the decay of the J/ψ are much more sensible to be emitted in the same direction. This can be observed in the Monte Carlo simulation depicted in Fig. A.2.

It can be observed that there is a group of events for which both muons are emitted backwards, in the same direction as J/ψ goes. Some other correspond to events where both muons go forward, and the rest of the events (the majority) correspond to evens with muon going in opposite directions in the z-axis.

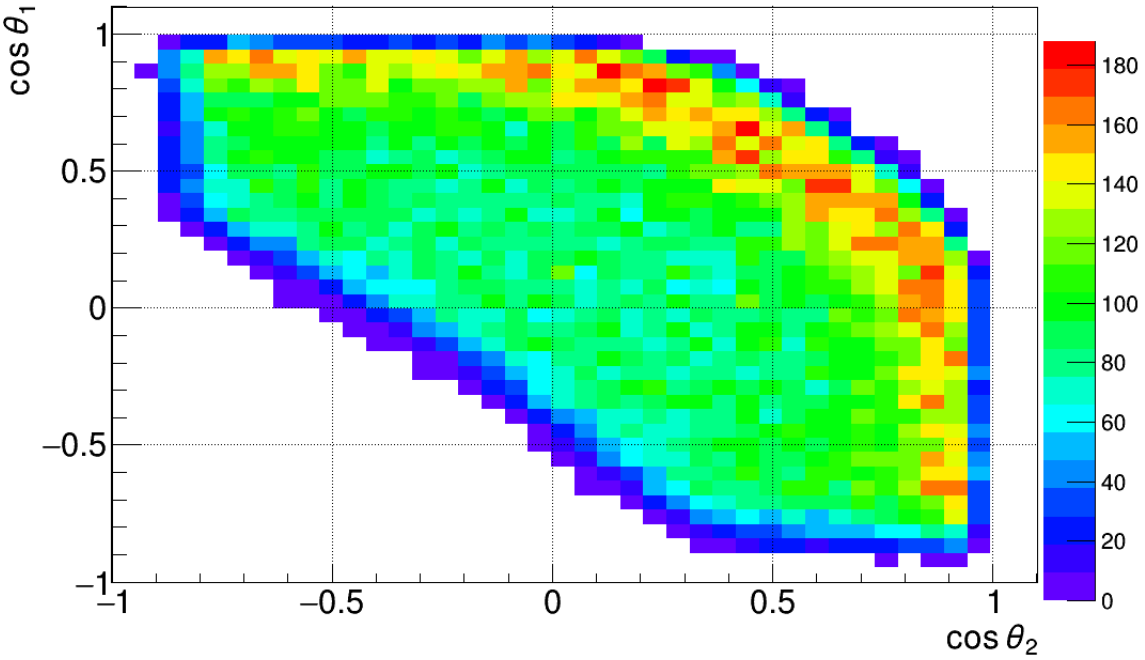


Figure A.2: Graph showing the directions in the Z axis of both muons coming from the decay of the J/ψ, by plotting $\cos\theta_1$ vs $\cos\theta_2$. Labels 1,2 stand for each muon.

Bibliography

[1] *F. Takasaki.*: **Status of KEKB accelerator and detector**. Int.J.Mod.Phys.A, 15S1:12–21. (Dec 1999) arXiv:hep-ex/9912004

[2] *A. Hutton.*: **PEP-II: an Asymmetric B Factory Based on PEP**. IEEE, 1:84–86, (May 1991) doi:10.1109/PAC.1991.164208

[3] *M. Kobayashi*: **Nobel Lecture: CP violation and flavor mixing**. Review of Modern Physics 81 (July 2009), pp. 1019–1025. doi: 10.1103/RevModPhys.81.1019

[4] *T. Maskawa*: **Nobel Lecture: What does CP violation tell us?**. Review of Modern Physics 81.3 (July 2009), pp. 1027. doi: 10.1103/RevModPhys.81.1027

[5] *K. Nishimura*: **New Physics Prospects in Mixing and CP Violation at Belle II** (Dec 2012). arXiv:hep-ex/1212.4112

[6] *T. Abe et al.* **Belle 2 Report** Tech. rep. High Energy Accelerator Research Organization (KEK), 2010. pp. 81–219. issn: 0146-6410. doi: 10.1016/0146-6410(95) 00042-H.

[7] *J. Kemmer and G. Lutz.* **New detector concepts**. Nucl. Instrum. Methods Phys. Res. Sect. A 253.3 (1987), pp. 365–377. issn: 0168-9002. doi: 10.1016/0168-9002(87) 90518-3.

[8] *R. Richter et al.* **Design and technology of DEPFET pixel sensors for linear collider applications**. Nucl. Instrum. Methods Phys. Res. Sect. A 511.1–2 (2003), pp. 250– 256. issn: 0168-9002. doi: 10.1016/S0168-9002(03)01802-3.

[9] *K.-F. Chen et al.* **Observation of Time-Dependent CP Violation in $\mathbf{B}^0 \to \eta' K0$ Decays and Improved Measurements of CP Asymmetries in $B0 \to \phi\mathbf{K}^0, K_S^0 K_S^0 K_S 0$ and $\mathbf{B}^0 \to J/\psi K_S$ Decays** PPhys. Rev.

[10] *A. Go et al.*: **Measurement of Einstein-Podolsky-Rosen-Type Flavor Entanglement in Υ(4S)$\rightarrow$ $B^0\bar{B}^0$**. Physical Review Letters, (Sept 2007) doi:10.1103/PhysRevLett.99.131802

[11] *T. D. Lee, C. N. Yang*: **Question of Parity Conservation in Weak Interactions**. In: Physical Review. 104, 1956, S. 254-258. doi:10.1103/PhysRev.104.254

[12] *C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, R. P. Hudson*: **Experimental Test of Parity Conservation in Beta Decay**. In: Physical Review. 105, 1957, S. 1413-1415. doi:10.1103/PhysRev.105.1413

[13] *M. Goldhaber, L. Grodzins, and A. W. Sunyar*: **Helicity of neutrinos**. In: Physical Review. 109, 1958 doi:10.1103/PhysRev.109.1015

[14] *L.D Landau* ZhETF 32 (1957) 405 [Sov. Phys. JETP 5 (1957) 336]; Nucl. Phys. 3 (1957) 127.

[15] *J. H. Christenson , J. W. Cronin, V. L. Fitch, and R. Turlay*: **Evidence for the 2π Decay of the K_2^0 Meson**. In Physical Review Letters 13.4 (July 1964), pp. 138–140. doi: 10.1103/PhysRevLett.13.138

[16] *NOνA Collaboration* **The NOνA Experiment** (21 March 2005) `http://nova-docdb.fnal.gov/0005/000593/001/NOvA_P929_March21_2005.pdf`

[17] *N. Cabibbo*: **Unitary Symmetry and Leptonic Decays**. Physical Review Letters 10 (June 1963), pp. 531–533. doi: 10.1103/PhysRevLett.10.531.

[18] *L. L. Chau and W. Y. Keung*: **Comments on the Parameterization of the Kobayashi-Maskawa Matrix**. Physical Review Letters, 53:1802–1805, (1984) doi: 10.1103/PhysRevLett.53.1802

[19] *L. Wolfenstein*: **Parameterization of the Kobayashi-Maskawa Matrix**. Physical Review Letters, 51:1945–1947, (1983) doi: 10.1103/PhysRevLett.51.1945

[20] *A. Martín Sánchez, for the LHCb Collaboration* **Measurement of the gamma angle from tree decays at the LHCb experiment** arXiv:1201.4736 [hep-ex]

[21] *K. Nakamura et al.* **Review of particle physics**. Journal of Physics G: Nuclear and PARTICLE Physics 37.7A (2010), p. 075021.

[22] *J. Charles et al.* **CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories**. Eur.Phys.J. C41 (2005). The CKMfitter group Webpage

[23] *J. P. Silva G. C. Branco, L. Lavoura.* **CP Violation. The International Series of Monografs on Physics**, 103, Oxford University Press, Oxford.

[24] *E. A. Paschos.* **Electroweak Theory**. Cambridge University Press, (2007).

[25] *K. Abe et al.*: **Observation of Large CP Violation in the Neutral Meson System**. Physical Review Letters, (Aug 2001) doi:10.1103/PhysRevLett.87.091802

[26] *B. Aubert et al*: **Observation of CP violation in the B^0meson system**. Physical Review Letters, (Aug 2001) doi:10.1103/PhysRevLett.87.091801

[27] *Fermilab* **Discoveries at Fermilab – Discovery of the Bottom Quark**, (Press release). Fermilab. 7 August 1977. Retrieved 2009-07-24.

[28] *J. Beringer et al.* **Review of Particle Physics** Particle Data Group. Retrieved (2015)

[29] *P. Vanhoefer*. **Measurement of the Branching Fraction of the Decay $B^0 \to \rho^0\rho^0$**. MA thesis. Max-Planck Institut für Physik München, 2010.

[30] *M. Ritter*. **Measurement of the Branching Fraction and Time Dependent CP Asymmetry in $B^0 \to D^+D^-K_S$ Decays at the Belle Experiment**. PhD thesis. Max-Planck Institut für Physik München, 2013.

[31] *T. E. Browder and K. Honscheid* **B mesons** Progress in PARTICLE and Nuclear Physics 35 (1995), Lett. 98 (3 Jan. 2007), p. 031802. doi: 10. 1103/PhysRevLett.98.031802

[32] *I. Adachi et al.* **Precise Measurement of the CP Violation Parameter $\sin^2\phi_1$ in $B^0 \to (c\bar{c})K^0$ Decays** Phys. Rev. Lett. 108 (17 Apr. 2012), p. 171802. doi: 10. 1103/PhysRevLett.108.171802

[33] *F. Abudinen*. **MeasurStudies on the neural z-Vertex-Trigger for the Belle II Particle Detector**. Master thesis. Max-Planck Institut für Physik München, 2014.

[34] *T. Bilka et al.* **Demonstrator of the Belle II Online Tracking and Pixel Data Reduction on the High Level Trigger System** (2015) arXiv:1406.4955v2

[35] *G. Casarosa* **Tracking Performance** 20th Belle II General Meeting (2015)

[36] *C. Pulvermacher, T. Keck, M. Feindt, M. Heck and T. Kuhr* **An automated framework for hierarchical reconstruction of B mesons at the Belle II experiment** Institut fur Experimentelle Kernphysik, Karlsruhe Institute of Technology (KIT), Journal of Physics: Conference Series 608 (2015) 012048

[37] *T. Keck* **The Full Event Interpretation for Belle II**, MSc. Thesis (2014), Institut fur Experimentelle Kernphysik, Karlsruhe Institute of Technology (KIT) doi:10.1088/1742-6596/608/1/012048

[38] *W. Waltenberger, W. Mitaroff, F. Moser, B. Pflugfelder and H. V. Riedel*: **The RAVE/VERTIGO vertex reconstruction toolkit and framework**. Journal of Physics: Conference Series 119 (2008) doi: 10.1088/1742-6596/119/3/032037

[39] *Wolfgang Waltenberger*: **RAVE – a detector-independent toolkit to reconstruct vertices**.Institute for High Energy Physics, Austrian Academy of Sciences, Vienna, Austria. Transactions on Nuclear Science `http://www.hephy.at/fileadmin/user_upload/Publikationen/rave.pdf`

[40] *W. Waltenberger, R. Frühwirth and P. Vanlaer*: **Adaptive Vertex Fitting**. CERN-CMS-NOTE-2008-033 (Jul 2008)

[41] *L. Li Gioi on behalf of the DEPFET collaboration* **Physics Benchmarks for the Belle II Pixel Detector** Max-Planck Institute für Physik München (2015)

[42] F. Abudinen and L. Li Gioi: **FlavorTagger Twiki Page - Belle II Internal Webpage** `https://belle2.cc.kek.jp/~twiki/bin/view/Physics/FlavorTagger`

[43] *T. Kawasaki*, **Simulation Study for Measurement of CP violation in B $\rightarrow$ charmonium + K_S at KEK B-factory**, Ph.D. Thesis.

[44] *Particle Data Group* **42. MONTE CARLO PARTICLE NUMBERING SCHEME** `http://pdg.lbl.gov/2014/reviews/rpp2014-rev-monte-carlo-numbering.pdf`

[45] *Wikipedia* **Particle Decay: Two-body decay** `https://en.wikipedia.org/wiki/Particle_decay#Two-body_decay`