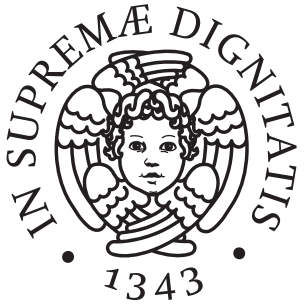


UNIVERSITÀ DI PISA

Dipartimento di Fisica *E. Fermi*
Corso di Laurea Magistrale in Fisica

TESI DI LAUREA MAGISTRALE IN FISICA

D^0 Lifetime Measurement with Belle II Early Data

CANDIDATO:
Gaetano de Marino

RELATORI:
Prof. Francesco Forti
Dott.ssa Giulia Casarosa

Anno Accademico 2018/2019

Contents

Abstract

Belle II is an international particle physics experiment located at the KEK Laboratory in Tsukuba, Japan, where the SuperKEKB accelerator operates. SuperKEKB is an asymmetric electron-positron B-factory, designed to operate at a center-of-mass energy of 10.58 GeV, corresponding to the mass of the $\Upsilon(4S)$ vector meson resonance, which decays into B meson-antimeson pairs nearly 100% of the times. The asymmetry in the beam energies results in a boost of the center of mass in the laboratory frame, necessary to study time-dependent decay rates and CP asymmetries.
While the previous experiments (*BABAR* and Belle) at first-generation B-factories (KEKB and PEP-II) were mainly dedicated to the measurement of CP violation in the B system, Belle II will focus on precision measurements of decays of bottom and charm mesons and τ leptons and on the search for rare or forbidden processes that may provide indications of physics beyond the Standard Model.
A significant Belle detector upgrade was necessary since the KEKB accelerator, already holding the world record for achieved luminosity, has gone through a substantial redesign. It will allow to obtain an even higher peak luminosity $\mathscr{L}$, providing the Belle II detector with a 50 times larger integrated luminosity $\int \mathscr{L}\,\mathrm{d}t = 50\,\mathrm{ab}^{-1}$. This will enhance the probability of finding hints of new physics in the flavour sector and will surely improve the precision on measurements of the Standard Model parameters.

My research project aims to measure the D^0 lifetime with the first processed data of the Belle II experiment, collected during the period of data taking between March 2019 and June 2019 and corresponding to an integrated luminosity of $2.6\,\mathrm{fb}^{-1}$. The target channel of the analysis is

$$D^{*+} \to D^0(\to K^-\pi^+)\pi_s^+ + \text{c.c.},$$

whose cross section at $\sqrt{s} = m_{\Upsilon(4S)}c^2$ is $\sim$9 pb. The objective is to demonstrate the vertexing capability of Belle II as an essential element for time-dependent measurements and for a large fraction of the physics program. This measurement requires both high resolution on track momenta and the ability of resolving the production and decay vertices of the D^0's, which is rather challenging since the average flight distance of D^0's from D^* decays is $\sim 220\,\mu$m. D^* candidates are selected from $c\bar{c}$ events and not from B decays because they add a non-negligible flight length to those of the D^0's and would induce e bias in the measurement. The momentum of D^* 's coming from the decay of B mesons has a maximum value that can be used as a threshold to discard the processes involving B decays.
The pion from D^* decay is traditionally referred to as "slow" pion since its phase space is very limited due to the small mass difference between the D^* and the D^0. This results in a very narrow distribution for the released energy of the decay, defined as $Q = m(D^0\pi_s) - m(K\pi) - m(\pi_s)$.

The main elements of the analysis examined in this thesis are: the calibrations, the analysis

tools and the reconstruction performance, mainly related to tracking (efficiency and impact parameter resolution), vertexing and particle identification.
A precise knowledge of the position of the Interaction Point (IP) and the size of the *beam spot*, i.e the region where the two colliding beams overlap, is needed. My task was to provide the run-dependent beam spot parameters and contribute to make them available to the analysis tools (e.g. fitters) through the access to the database. First, the IP position was reconstructed *online* in the context of the Data Quality Monitoring (DQM) of the experiment. Then, the same information stored for the DQM was also used for the beam spot calibration, which was as useful to the experiment as it was to our analysis.
The selection for the lifetime measurement takes place in two steps: the pre-selection, centrally made by the Belle II production to provide only the candidates passing basic requirements, and the final selection, which constitutes an important part of the work for background suppression. It also required to investigate the performance of the detector since the efficiency and purity corresponding to each of the considered variables needed to be evaluated. It is important to emphasize that the selection has been studied directly on data and not on Monte Carlo (MC) since the main goal of the thesis work was to prove the good performance of the detector. Furthermore, the agreement between data and MC is still non optimal due to the early phase of the experiment.
The fits to the m_{D^0} and the Q distributions have been performed in order to define the *signal region* where to extract the final candidates from. In the (m_{D^0}, Q) space, a *sideband region* is also identified, with the purpose of estimating the level of background. Characterisation of signal and background was made on MC by means of the Belle II analysis tools. MC samples used for the analysis consist of 10^6 signal events and the equivalent of $80\,\text{fb}^{-1}$ for the main processes at B-factories: $B\overline{B}$ decays, continuum processes ($e^+e^- \to q\bar{q},\ \ q = u,d,s,c$) and $e^+e^- \to \tau^+\tau^-$ events (collectively called *generic* MC). The signal was treated separately to study the resolution function and test the probability density function (PDF) for the unbinned maximum likelihood fit to the proper time distribution. The proper time t_{D^0} is obtained from the reconstructed momentum $\vec{p}$ of the D^0, the nominal mass m_{D^0} and the measured distance between the production and decay vertices.
The identified background sources are $c\bar{c}$ processes with $D^0 \to K\pi\gamma X$ decays and mis-reconstructed or fake π_s's attached to true D^0's to make the D^* candidates, and they are both treated as signal in the final model. Combinatorial background contamination, despite being very small due to the low statistics, is included in the final fit as a separate and fixed component extracted from the sidebands. The model has been validated on the full MC sample, obtained merging the various samples. The result obtained on the $2.6\,\text{fb}^{-1}$ sample of early data is

$$\tau_{D^0} = (400 \pm 6\,(\text{stat}) \pm 9\,(\text{syst}))\,\text{fs}\,.$$

Among the possible sources of systematic uncertainties, the largest one is given by the residual misalignment of VXD or biased beam spot position calibration. These two points can have the same effect of shortening or lengthening the measured proper time. The final value of 9 fs is obtained by summing in quadrature this term with the other evaluated contributions, related to the choices made for the selection and the PDF for the fit to the proper time.
The uncertainty on the measurement is dominated by the systematic error, although the available statistics is still very low.
The natural next step of the analysis would be to repeat the reconstruction once improved calibrations are available, both for alignment and for the beam spot. Further studies on MC can reveal all the possible effects that such non-optimal calibrations can induce on the measurement.

In addition, the analysis can be extended to other decay channels (multi-body charged and 3-body involving neutral states) but a deeper understanding of the detector performance is indispensable.
The progress of the analysis presented in this manuscript is also documented in a Belle II note, being reviewed by an internal committee for the approval of the plots and the entirety of the work and has been presented to the 2019 summer conferences.

Introduction

The Standard Model of particle physics is supported by a formidable level of confirmation by experiment, but it cannot be seen as a complete, satisfactory theory. Despite its tremendous success in describing the fundamental particles and their interactions, it does not provide answers to many fundamental questions. Furthermore, there are some experimental results showing discrepancies with the actual frame of the Standard Model (SM). The main goal of particle physics is to investigate these discrepancies, which could point to possible "new physics" scenarios, and try to explain the unresolved puzzles, briefly summarised here.
Since the Higgs boson discovery at LHC, a great effort is devoted to measure the Higgs dynamics as it is the responsible of the masses of the particles. Even though the Higgs mass generation mechanism is well understood in the frame of the SM, there still is the hierarchy problem, which raises the question of why the Higgs mass is so small compared to the Planck scale and why there is an observed hierarchy in the fermion masses. Also the dynamical origin of the tiny scale of neutrino masses remains a mystery: the detection of neutrino oscillations has established that neutrinos have small, but nonzero, masses and that the flavour and mass eigenstates do not coincide. Since it is not clear whether the Higgs boson can account for neutrino masses, any discovery in neutrino physics may have a great impact on the flavour structure of the SM.
Flavour physics is also struggling with the origin of CP violation and how do matter and antimatter particles differ. The Universe today contains a relatively small amount of matter and a lot of radiation, but no significant quantities of antimatter. One condition for generating such matter-antimatter asymmetry is, according to the Sakharov interpretation (1967), CP violation. However, the CP violation observed so far, as described within the SM via one single phase in the CKM matrix, is insufficient to generate the matter-antimatter asymmetry seen in the Universe today, which requires explanations beyond the SM.
Another cosmological puzzle is the nature of dark matter, that makes up 80 % of the matter in the Universe. Astrophysics assures that the formation of structures in the Universe and their persistence today is possible only with the help of additional gravitational attraction provided by some form of invisible non-relativistic matter, which does not bind to ordinary matter and has neither electric charge nor strong interactions.
Finally, any quantum field theory of the fourth known fundamental force, gravity, is absent in the present SM. There is a well-founded belief that there must be some more profound theory than SM where gravity can also be unified with electroweak and strong interactions.
In addition to these conceptual problems in the SM, a few experimental results are in tension with the SM expectations, pointing to possible New Physics (NP). In particular, the muon anomalous magnetic moment and the anomalies in flavour [1], involving both neutral and charged current processes. In perturbative quantum field theory, the magnetic moment of particles as predicted by classical field theory may receive corrections due to quantum effects. Such corrections are also called anomalous magnetic moments, traditionally denoted by "g-2". The anomalous magnetic moment of the muon $g_\mu - 2$ has become notorious for showing

a noticeable discrepancy between theoretical prediction from the SM and its measured value up to 3.6σ.
In the heavy flavour sector, various tensions with respect to SM theoretical expectations are currently present. In the SM particles belonging to the three generations of leptons have identical properties from one generation to the next, except for their masses; this is referred to as lepton universality. A possible violation of lepton universality is called a flavour anomaly, which would be a sign of NP beyond the SM. Indications of flavour anomalies have been observed consistently and with increasing statistical significance around $2.5-3\sigma$ in B meson processes in various channels by a number of independent experiments. Of particular interest are the branching ratios involving flavour changing $\Delta F = 1$ processes such as $R_{K^{(*)}}$ and $R_{D^{(*)}}$:

$$R_{K^{(*)}} = \frac{\Gamma(B \to K^{(*)}\mu^+\mu^-)}{\Gamma(B \to K^{(*)}e^+e^-)}, \quad R_{D^{(*)}} = \frac{\Gamma(B \to D^{(*)}\tau\,\nu_\tau)}{\Gamma(B \to D^{(*)}l\,\nu_l))} \qquad (l \in \{e,\mu\}).$$

Belle II is a next generation experiment of flavour physics whose goals are to search for NP in the flavour sector at the intensity frontier, and to improve the precision of measurements of Standard Model (SM) parameters. The research program follows the legacy of the previous experiments (Belle and *BABAR*) at B-factories (KEKB and PEP-II), in competition and complementarity with that of LHCb at LHC. Belle II installation was completed in December 2018 and a first data taking period took place in March-June 2019.
The aim of this work is to demonstrate the capabilities of Belle II through the analysis of its data, collected during the first three months of operations. In particular, the vertexing performance are tested for the first measurement of the D^0 meson lifetime in the channel $D^{*+} \to D^0(K^-\pi^+)\pi^+$. Working in the context of the charm analysis working group, I personally developed the analysis methodology and code, optimized the event selection and PDF modeling, and performed the final fit to data. Given the importance of the interaction point determination for this analysis, I developed the tools to provide the beam spot information during data taking and store it in the conditions database.
The present thesis is structured as follows.
Chapter 1 briefly introduces to the SM frame and and CP violation formalism. An overview of Belle II physics program is given, with particular attention to charm physics and D^0-$\overline{D}^0$ mixing and CP violation in the charm sector, which was observed for the first time by the LHCb experiment in March, 2019. A more detailed description of the discovery is reported in this chapter.
In Chapter 2 SuperKEKB and the Belle II Experiment are described, mainly focusing on the increase in luminosity and the detector upgrades required to maintain or improve the performance with respect to Belle. The main phases of the experiment are also outlined, presenting the performance achieved so far and the problems encountered during the first months of data taking.
In Chapter 3 the main elements for the analysis are explored: the calibrations, the analysis tools and the reconstruction performance, which strongly involves tracking, vertexing and particle identification.
A precise knowledge of the position of Interaction Point (IP) and the size of the interaction region or *beam spot*, where the two colliding beams overlap, is vital for several analyses related, for example, to charm and beauty physics. In Sec. 3.5 the main algorithm implemented for the time-dependent beam spot parameters determination is described and the effectiveness of such calibrations for the success of the measurement is demonstrated throughout the work.
Vertexing has a crucial role in this analysis since the production and decay vertices of D^0 are

few hundreds of microns from each other and require excellent resolution on the impact parameter of tracks. The Belle II VerteX Detector (VXD) efficiency has been evaluated in relation to the channel studied. Moreover, Particle IDentification (PID) is fundamental to obtain high-purity signal sample.
In Chapter 4 the selection criteria optimisation to reduce the background and enhance the purity of the sample are described. The final selection for Monte Carlo (MC) an data are both reported in this chapter. The official MC samples used for the analysis consist of 10^6 signal events and the equivalent of $80\,\text{fb}^{-1}$ for the main processes that can constitute a background at a B-factory.
MC signal has been mainly used for the characterisation of signal and background while the selection has been directly improved on data since the main goal was not to demonstrate perfect agreement between data and MC but to prove the good performance of the detector during its first months of operations.
In Chapter 5 the studies related to extraction of the D^0 lifetime are presented. The procedure is conducted on MC first, in order to study the resolution function and test the fit model on the various candidate categories identified in Chapter 4. Then the full available statistics collected at the $\Upsilon(4S)$ resonance was used to obtain the final measurement with its statistical error. The last section is devoted to the estimation of the systematic uncertainty, which reflects the present limits on the measurement and points to the improvement directions.
The precise measurement of primary and secondary vertices of short-lived particles such as D mesons is the first step towards the measurement of mixing and CP violation parameters. B-factories have proven to be excellent tools for charm mixing study and SuperKEKB will provide more than 6×10^{10} charm events (against the 10^9 collected at with Belle). The Belle II Vertex detector is the main element for D-mixing sensitivity measurement since it provides an improvement by a factor two on the D^0 decay time resolution with respect to Belle and *BABAR*, which subsequently improves the precision of mixing and CP violation parameters. For example, the extrapolated precisions at $50\,\text{ab}^{-1}$ (expected by 2027) are $\leq 0.1\%$ on x and y parameters, with an improvement of approximately 40% and 70% over the current respective measurements [2].

1 Belle II and flavour physics

This chapter discusses the Belle II physics program in the broader context of research in particle physics. The main measurements planned for the next years of data taking are briefly outlined. Particular attention is dedicated to charm physics, as the core of the present study is meant to evaluate the capabilities of Belle II experiment in preparation for time-dependent measurements on D mesons (as discussed in Sec. 1.5). The basic nomenclature and the formalism necessary to describe the phenomena of mixing and CP violation are exposed. Finally, the experimental status in flavour physics is briefly reviewed in section 1.4, including the recent observation of CP violation in the charm sector by the LHCB experiment.

1.1 Brief introduction to the SM

At the basis of the Lagrangian Field Theory there is the concept of symmetry, which denotes an invariance of the Lagrangian $\mathscr{L}$, and thus of the associated equations of motion, under some transformations. Such transformations can be local (coordinate-dependent) or global and they can be continuous or discrete. The invariance may be with respect to coordinate redefinitions, as in the case of Lorentz invariance, or field redefinitions, as in the case of gauge invariance.
Of particular interest are the global discrete invariances known as C or charge conjugation (replacement of a field by its particle-antiparticle conjugate), P or parity (sign reversal of all spatial coordinates), and T or time reversal (sign reversal of the time coordinate, which reverses the role of *in* and *out* states). When constructing a field theory, locality, Lorentz invariance, and hermiticity of $\mathscr{L}$ are always required. That is sufficient to make any field theory automatically also invariant under the product of operations CPT. In many theories the combination CP, and thus also T, are separately automatic symmetries. This is the reason why the experimental discovery that CP is not an exact symmetry of nature was such a breakthrough event.
The Standard Model is the unified gauge theory for electroweak and strong interactions, the gauge group being $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$[1]. In this model a single theory includes the electromagnetic and weak interactions — both neutral current (NC) and charged current (CC) — and QCD. In particular, electroweak theory has the symmetry group $SU(2) \otimes U(1)$ and electromagnetic and weak coupling constants are correlated by the theory, while QCD and electroweak theories are unified by the theoretical framework keeping their coupling constants independent [3]. The current SM frame describes the interactions between the fermion fields (quarks and leptons in Fig. 1.1) and the bosons as gauge fields:

- the foton γ and the weak bosons (W^+, W^-, Z^0) for the electroweak field (vector bosons);
- the gluons, mediators of the strong interactions;

[1] here c denotes the colour, L the chiral component and Y the hyper-charge.

- the Higgs boson, related to the Higgs scalar field, responsible for the mass generation mechanism.

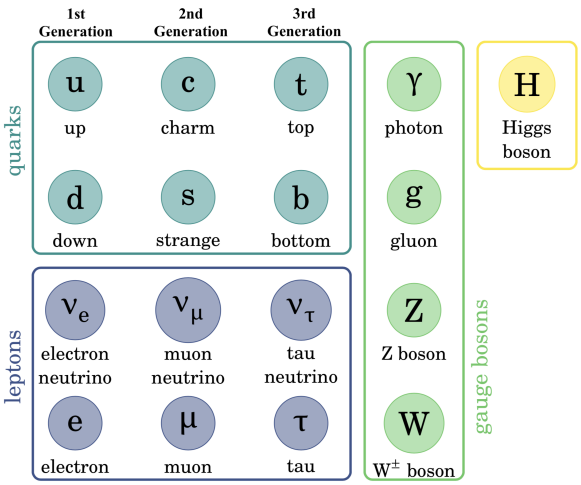


Figure 1.1: SM particles.

The fermion fields of the SM (quark and leptons in Fig. 1.1), written in terms of Weyl spinors in a $(c,L)_Y$ notation, are:

$$\begin{array}{lll} Q_i & (3,2)_{1/6} & \text{left}-\text{handed quark doublet} \\ U_i^c & (3,1)_{2/3} & \text{right}-\text{handed quarkup} \\ D_i^c & (3,1)_{-1/3} & \text{right}-\text{handed quarkdown} \\ L_i & (1,2)_{-1/2} & \text{left}-\text{handed lepton doublet} \\ E_i & (1,1)_{-1} & \text{right}-\text{handed charged lepton}\,, \end{array}$$

where $i = 1,2,3$ denotes the generation index. For example, $Q_i(3,2)_{1/6}$ denotes the left-handed (chiral) component of a field which is a triplet under SU(3) colour group and weak isospin doublet, whose hyper-charge is 1/3. The weak hyper-charge is defined as $Y_W = (Q - T_3)$, being Q the electric charge and T_3 the third component of weak isospin (SU(2) component).
The most general renormalizable Lagrangian can be written as:

$$\mathscr{L}_{\text{SM}} = \mathscr{L}_{\text{kin}} + \mathscr{L}_{\text{Higgs}} + \mathscr{L}_{\text{Yuk}}\,.$$

Here $\mathscr{L}_{kin}$ describes the free propagation in the spacetime and contains the covariant derivatives and the fields' strenghts; $\mathscr{L}_{\text{Higgs}}$ gives the Higgs scalar potential:

$$\mathscr{L}_{\text{Higgs}} = \mu^2|\Phi|^2 - \lambda|\Phi|^4 \tag{1.1}$$

where $\Phi = \begin{pmatrix} \phi_+ \\ \phi_0 \end{pmatrix}$.
Finally, the interaction of the quarks and leptons with the Higgs field, introduced to give masses to them and to the gauge bosons, is described by

$$\mathscr{L}_{\text{Yuk}} = y^d_{ij}\overline{Q}^i \Phi D^j + y^u_{ij}\overline{Q}^i \Phi^\dagger U^j + \text{lepton term} + \text{h.c.}\,.$$

Choosing $\mu^2 < 0$ and $\lambda > 0$ and defining $v^2 = -\frac{\mu^2}{\lambda}$, the scalar potential 1.1 is (up to a constant term):

$$\mathscr{L}_{\text{Higgs}} = -\lambda\left(\Phi^\dagger\Phi - \frac{v}{2}\right)^2$$

which implies that the field acquires a *vacuum expectation value*, or *vev*, $\langle\phi\rangle = v/\sqrt{2}$ for at least one direction of Φ, for example in the real direction of the down component:

$$\langle\phi\rangle = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}.$$

The Higgs *vev* breaks the $SU(2)\otimes U(1)$ symmetry down to a $U(1)$ subgroup and, writing the doublet fields out explicitly,

$$Q_L^i = \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix},$$

gives mass terms of the form

$$\mathscr{L}_{\text{mass,q}} = m_{ij}^d \overline{d}_L^i d_R^j + m_{ij}^u \overline{u}_L^i u_R^j. \tag{1.2}$$

The mass matrices are related to the Yukawa couplings and the Higgs *vev* $\langle\phi\rangle = v/\sqrt{2}$ by

$$m_{ij}^q = \frac{v}{\sqrt{2}} y_{ij}^q .$$

In general the m^q matrices are not diagonal; one can move to the mass basis and have the m^q's diagonal through a unitary transformation

$$\hat{m}_{ij}^q = (V_L^q)_{ik} (m^q)_{kl} (V_R^{q\dagger})_{lj},$$

with the notation $\hat{m}_{ij}^q$ indicating the diagonalised matrix. In principle the above L, R indices are distinct from the L, R of the previous notation (1.2), but the i index is associated with the left-chiral quark while the j index with the right-chiral quark (see Eq. 1.2). Thus rotate the left-chiral and right-chiral fields by V_L and V_R accordingly:

$$\begin{aligned} q_L^i &= (V_L^q)_j^i q_L^{'j} \\ q_R^i &= (V_R^q)_j^i q_R^{'j}, \end{aligned}$$

having on the left the interaction-basis field and the linear combination of mass-basis fields on the right. From here on the mass basis is used and the prime is dropped. In the mass basis the Yukawa interactions are diagonal since mass matrices are proportional to the Yukawa matrices but now the coupling of the W bosons contains off-diagonal terms. Indeed this change of the basis leads to:

$$\mathscr{L}_{Wqq} \propto \frac{g}{\sqrt{2}} \overline{u}_L i\gamma_\mu d_L W^\mu \rightarrow \frac{g}{\sqrt{2}} \overline{u}_L i\gamma_\mu (V_{uL} V_{dL}^\dagger) d_L W^\mu . \tag{1.3}$$

The element $(V_{uL} V_{dL}^\dagger)$ indicates the Cabibbo-Kobayashi-Maskawa unitary matrix:

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

where the quarks are ordered by increasing mass and the mass eigenstates are named after their dominant flavour. As an $N \times N$ unitary matrix, V_{CKM} has N^2 real parameters, $2N-1$ of which can be reabsorbed by the quark fields which multiply the right and left sides of the matrix in Eq. 1.3, by means of a global redefinition of the arbitrary phases. Of the remaining $(N-1)^2$ free parameters, $N(N-1)/2$ are the Euler angles while the others $\frac{N}{2}(N-3)+1$ are

unremovable complex phases - just one, for $N = 3$. The standard way to parametrize the matrix in terms of mixing angles and the irreducible phase is the following

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

given that $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$.
The measured values of the modules of the elements are reported below [4]:

$$V_{\rm CKM} = \begin{pmatrix} 0.97446 \pm 0.00010 & 0.22452 \pm 0.00044 & 0.00365 \pm 0.00012 \\ 0.22438 \pm 0.00044 & 0.97359^{+0.00010}_{-0.00011} & 0.04214 \pm 0.00076 \\ 0.0096^{+0.00024}_{-0.00023} & 0.04133 \pm 0.00074 & 0.999105 \pm 0.000032 \end{pmatrix}.$$

At the price of loosing unitarity, Wolfenstein introduced a different parametrization guided by the fact that the matrix elements can be expressed in terms of the four parameters λ, A, ρ, η, defined as:

$$\lambda \equiv s_{12} = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}$$
$$A\lambda^2 \equiv s_{23} = \lambda \left| \frac{V_{cb}}{V_{us}} \right|$$
$$A\lambda^3(\rho + i\eta) \equiv s_{13}e^{i\delta} = V_{ub}^* = \frac{A\lambda^3(\bar{\rho} + i\bar{\eta})\sqrt{1 - A^2\lambda^4}}{\sqrt{1-\lambda^2}[1 - A^2\lambda^4(\bar{\rho} + i\bar{\eta})]}$$

which leads to:

$$V_{\rm CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{1.5}$$

Experimental values of the four parameters are [4]:

$$\lambda = 0.22453 \pm 0.00044, \qquad A = 0.836 \pm 0.015, \qquad \bar{\rho} = 0.122^{+0.018}_{-0.017}, \qquad \bar{\eta} = 0.355^{+0.012}_{-0.011}.$$

The power of Wolfenstein representation is that it captures the relevant physics of CKM matrix: the upper left 2×2 matrix is the Taylor expansion in θ_{12} of Cabibbo matrix[2]. Furthermore, at the first order in λ (which is < 1) the first two generations are decoupled from the third one and only at third order in λ the complex phase shows up.
The relations 1.5 ensure that $\overline{\rho} + i\overline{\eta} = -(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)$ is phase convention independent and the CKM matrix written in tems of λ, A, $\bar{\rho}$ and $\bar{\eta}$ is unitary to all orders in λ.
A theory can be CP violating if and only if it has some set of couplings for which rephasing of all fields cannot remove all the phases, while CP conservation is automatic for any theory for which the most general form of the Lagrangian allows all complex phases to be removed. If the Lagrangian involves a sufficient number of fields, either fermions or scalars, so that there are more couplings than the possible phase redefinitions, not all couplings can be made real by rephasing the fields. The three generation Standard Model with a single Higgs doublet has only one CP-violating parameter. Thus CP violation per se is no longer a puzzle, but rather a natural part of the Standard Model. However, as already discussed in the Introduction, further

[2] the first form in which the matrix was postulated took into account only the first two generations. This proposal, made by Cabibbo, was the basis of the anticipation of the existence of the charm quark.

sources of CP violation are needed to explain the matter-antimatter asymmetry.
Any relationship of the form of a sum of three complex numbers equal to zero can be drawn as a closed triangle in the complex plane. Hence the unitarity relationships for the CKM matrix

$$V_{CKM}V_{CKM}^{\dagger} = V_{CKM}^{\dagger}V_{CKM} = \mathbb{1} \Longleftrightarrow \begin{cases} V_{ud}V_{us}^{*} + V_{cd}V_{cs}^{*} + V_{td}V_{ts}^{*} = 0 & (a) \\ V_{ud}V_{ub}^{*} + V_{cd}V_{cb}^{*} + V_{td}V_{tb}^{*} = 0 & (b) \\ V_{us}V_{ub}^{*} + V_{cs}V_{cb}^{*} + V_{ts}V_{tb}^{*} = 0 & (c) \\ V_{ud}V_{td}^{*} + V_{us}V_{ts}^{*} + V_{ub}V_{tb}^{*} = 0 & (d) \\ V_{ud}V_{cd}^{*} + V_{us}V_{cs}^{*} + V_{ub}V_{cb}^{*} = 0 & (e) \\ V_{cd}V_{td}^{*} + V_{cs}V_{ts}^{*} + V_{cb}V_{tb}^{*} = 0 & (f) \end{cases}$$

are referred to as the Unitarity Triangle relationships.
The fact that there is only one independent CP-violating quantity in the CKM matrix can be expressed in phase-convention-invariant form by defining the quantity J, called the Jarlskog invariant for Cecilia Jarlskog who first pointed out this form [5]:

$$\mathrm{Im}(V_{ij}V_{kl}V_{il}^{*}V_{kj}^{*}) = J\sum_{m,n=1}^{3}\varepsilon_{ikm}\varepsilon_{jln} \;\; (i,j,k,l = 1,2,3)\,;$$

V_{ij} are the CKM matrix elements and ε_{ikm} is the total antisymmetric tensor. All the unitarity triangles have the same area, J/2, being the measured value for $\mathrm{J} = (3.18 \pm 0.15) \times 10^{-5}$ [4]. This area shrinks to zero if the CP-violating phase differences in the matrix vanish.
Notice however that, while the triangles have the same area, they have very different shapes. Triangle (e), the so called "charm unitarity triangle", has two sides of order λ and one of order λ^5 (as is sketched in Fig. 1.2). Therefore measuring J using such a triangle is more difficult.

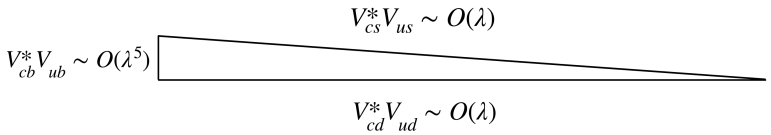


Figure 1.2: Charm Unitarity Triangle.

Triangle (b) is the most interesting, because it has all three sides of order λ^3 so all three angles are of comparable magnitude. The price one pays is that all the sides are small, but this is not as serious as the problem of measuring an asymmetry proportional to a very small angle. This triangle, the most often discussed in relation to B-meson decays as the main "Unitary Triangle" is conventionally drawn by dividing all sides by $V_{cd}V_{cb}^{*}$, which gives a triangle with base of unit length whose apex is the point $(\overline{\rho}, \overline{\eta})$ in the complex plane (Fig. 1.3a). It is the least flat triangle, hence the B_d meson system is where the highest sensibility to CPV is expected.

1.2 Mixing and CPV parameters in neutral meson systems

Flavoured neutral mesons, those which have non-zero strangeness S, charm C or beauty B content (neutral Kaons K^0, D^0 , B_d and B_s mesons) are an interesting physics laboratory, since particle and antiparticle are distinguished only by the flavour quantum number, which is not conserved by the weak interaction. This means that weak force can couple a neutral meson $|M^0\rangle$ and also the anti-particle $|\overline{M}^0\rangle$ to the same state I, resulting in the transitions:

$$M^0 \to I \to \overline{M}^0 \quad \text{and equivalently} \quad \overline{M}^0 \to I \to M^0\,,$$

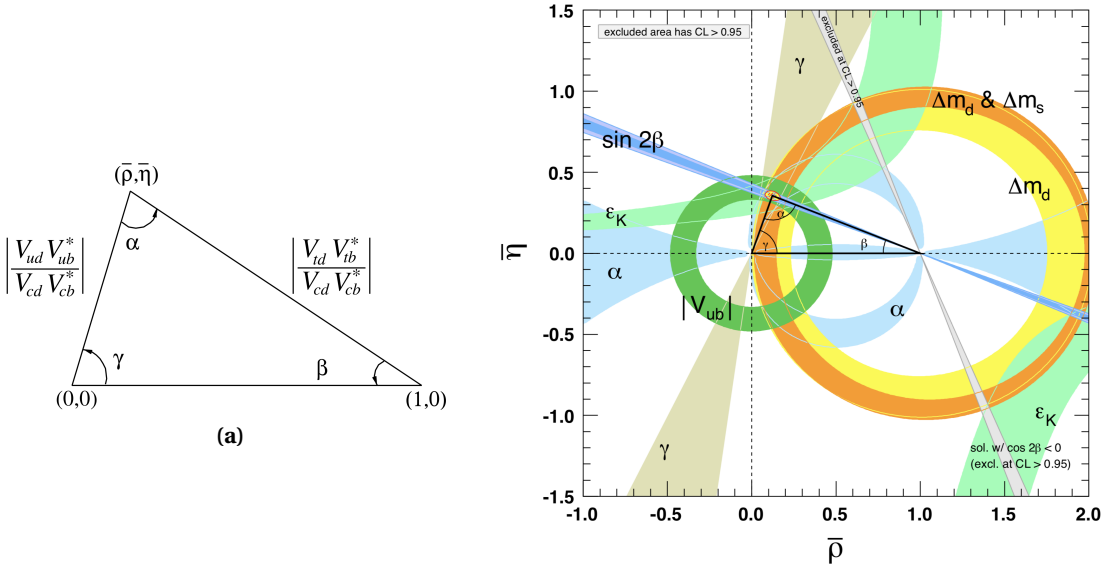


Figure 1.3: **(a)** Unitarity Triangle. **(b)** Constraints on the $\bar{\rho},\bar{\eta}$ plane. The shaded areas have 95% CL. Pictures from [4].

which combined generate *mixing*:

$$M^0 \to \overline{M}^0 \to M^0 \ldots .$$

Without loss of generality, we can refer to charmed mesons D^0's (with $c = \pm 1$) while introducing the mixing and oscillation formalism. Oscillations occur when the mixing process, violating the flavour number, leads to the mass eigenstates not being flavour eigenstates. To explain this, let's consider the time evolution of neutral D mesons solving the free Schrödinger equation

$$i\frac{d}{dt}\begin{pmatrix} D^0 \\ \overline{D}^0 \end{pmatrix} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}\begin{pmatrix} D^0 \\ \overline{D}^0 \end{pmatrix} . \tag{1.6}$$

CPT invariance imposes

$$M_{11} = M_{22} \qquad \Gamma_{11} = \Gamma_{22} .$$

The mass eigenstates, obtained through diagonalising this matrix, are given by (more details in [6])

$$|D_a\rangle = \frac{1}{\sqrt{|p|^2+|q|^2}}(p|D^0\rangle + q|\overline{D}^0\rangle), \qquad |D_b\rangle = \frac{1}{\sqrt{|p|^2+|q|^2}}(p|D^0\rangle - q|\overline{D}^0\rangle) . \tag{1.7}$$

with

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} .$$

They have different masses and widths, defined as follows:

$$\Delta m \equiv m_b - m_a \qquad \Delta\Gamma \equiv \Gamma_b - \Gamma_a . \tag{1.8}$$

The time evolution of the mass eigenstates is thus:

$$|D_{a,b}(t)\rangle = e^{-\frac{1}{2}\Gamma_{a,b}t}\, e^{-iM_{a,b}t}|D_{a,b}\rangle$$

and the probability to find a $\overline{D}^0$ at a time t in an initially pure D^0 beam is given by

$$P(D^0 \to \overline{D}^0, t) \propto \left|\langle \overline{D}^0 | D^0(t)\rangle\right|^2 = \frac{1}{4}\left|\frac{q}{p}\right| e^{-\Gamma_a t}(1 + e^{-\Delta\Gamma t} - 2e^{-\frac{1}{2}\Delta\Gamma t}\cos(\Delta m t)) \tag{1.9}$$

while the probability to find a D^0 at a time t in an initially pure $\overline{D}^0$ beam is given by

$$P(\overline{D}^0 \to D^0, t) \propto \left|\langle D^0 | \overline{D}^0(t)\rangle\right|^2 = \frac{1}{4}\left|\frac{p}{q}\right| e^{-\Gamma_a t}(1 + e^{-\Delta\Gamma t} - 2e^{-\frac{1}{2}\Delta\Gamma t}\cos(\Delta m t))\,. \tag{1.10}$$

We now introduce two dimensionless ratios, called *mixing parameters*:

$$x \equiv \frac{\Delta m}{\overline{\Gamma}}\,,\; y \equiv \frac{\Delta\Gamma}{2\overline{\Gamma}}\,, \qquad \overline{\Gamma} = \frac{1}{2}(\Gamma_a + \Gamma_b). \tag{1.11}$$

There is evidence of oscillation in the D system if at least one of the mixing parameters is different from zero. For the K system $x_K \sim -0.95$ and $y_K \simeq 1$; for the B_d system $x_d = 0.770 \pm 0.004$ but $y_d \sim 5 \times 10^{-3}$; the B_s system has $x_s = 26.72 \pm 0.09$ and $y_s \sim 0.15$ [7], while for charmed neutral mesons $x_D, y_D < 10^{-2}$.

As already mentioned in Sec. 1.1, CP violation is introduced in the SM with the complex phase in the CKM matrix; it can manifest through a rich phenomenology, classified in three cases [8]:

1. **Direct violation in decay**, when the probability of the state M^0 going into a final state f is different from the probability that the CP conjugate state $\overline{M}^0$ goes into $\overline{f}$. This form of CP violation is independent from mixing and can occur also for charged mesons. Defining $A_f = \langle f|\mathscr{H}_w|M^0\rangle$ and $\bar{A}_{\bar{f}} = \langle f|\mathscr{H}_w|\overline{M}^0\rangle$, where $\mathscr{H}_w$ is the Hamiltonian describing the weak process, CPV implies $\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| \neq 1$. In fact, if CP were a symmetry for $\mathscr{H}_w$, CP would commute with $\mathscr{H}_w$ and $|A_f| = |\bar{A}_{\bar{f}}|$.

   It is importanto to note that, in order to exhibit this type of CPV, it is necessary that the process $P \to f$ proceeds through at least two amplitudes with different phases:

$$\begin{aligned} A_f &= a_1 e^{i(\delta_1+\phi_1)} + a_2 e^{i(\delta_2+\phi_2)} \\ \bar{A}_{\bar{f}} &= a_1 e^{i(\delta_1-\phi_1)} + a_2 e^{i(\delta_2-\phi_2)}\,, \end{aligned}$$

   wehere the δ's are the CP-invariant strong phases and the ϕ's are CP-violating weak phases (they change sign under CP). In this case

$$|A_f|^2 - |\bar{A}_{\bar{f}}|^2 = -4a_1 a_2 \sin(\delta_1 - \delta_2)\sin(\phi_1 - \phi_2).$$

   Therefore the sensitivity to the uneliminable weak phase in the amplitudes A_f and $\bar{A}_{\bar{f}}$ only if $\delta_1 \neq \delta_2$. The experimentally accessible quantity is

$$A_{\rm CP} = \frac{\Gamma(M^0 \to f) - \Gamma(\overline{M}^0 \to \bar{f})}{\Gamma(M^0 \to f) + \Gamma(\overline{M}^0 \to \bar{f})}\,,$$

   which can possibly be measured as a function of time.

2. **Violation in the mixing**, when $\left|\frac{q}{p}\right| \neq 1$ (see Eq. 1.7). This would imply, from Eqs. 1.9 and 1.10 that the probability for a D^0 to go into a $\overline{D}^0$ is different from the inverse process.

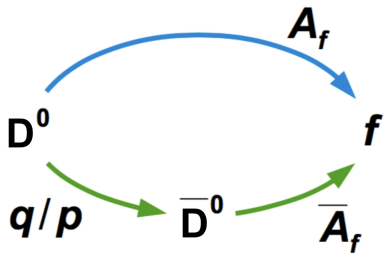


Figure 1.4: Interference between decay and mixing for a general final state f.

3. **Violation in the interference between mixing and decay amplitudes**, if both D^0 and $\overline{D}^0$ can decay into the same final state - which is thus possible only for neutral mesons. This means that the final state f can be reached both through a direct decay $D^0 \to f$, and through mixing, followed by decay: $D^0 \to \overline{D}^0 \to f$ (Fig. 1.4). This process can be parametrised by the quantity $\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f}$ and CP violation holds if $\mathrm{Im}(\lambda_f) \neq 0$.

The heavy-flavoured neutral mesons have quite different phenomenology [9] with respect to K systems, mainly due to the fact that the heavier mass leads to many more decay modes with comparable branching ratios, and this introduces new complexity to mixing studies. The D mesons' system is the most peculiar because of very small mixing parameters x, y (Sec. 1.4); furthermore, it is the only one made of up-type quarks, hence the only one for which the virtual quarks are down-type (as shown in Fig. 1.5). Therefore mixing in the D^0- $\overline{D}^0$ system brings complementary information and could be, in principle, sensitive to a different New Physics sector, and this is true also for what concerns CP violation.

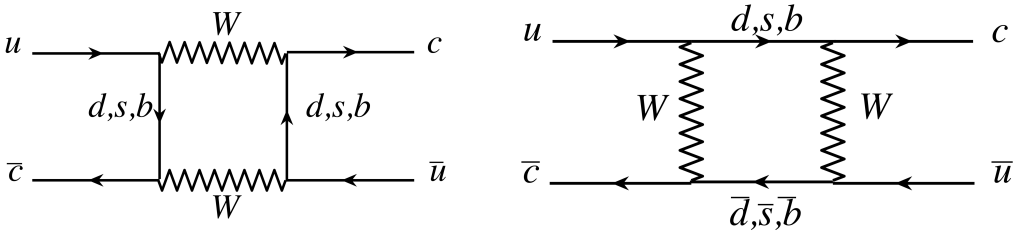
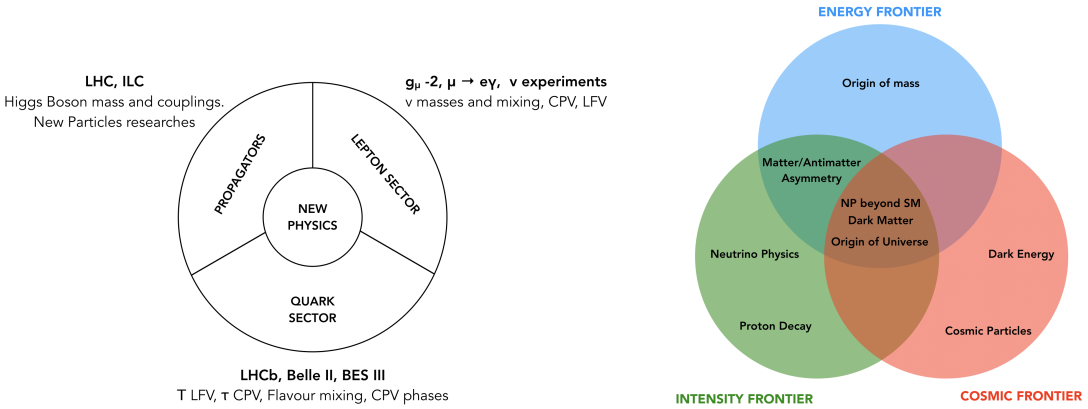


Figure 1.5: Box diagrams for the mixing of D mesons. Each contribution to the mixing is matched by a diagram where the quark triplet and the W boson are interchanged.

Values of the mixing parameters for the D^0 meson system can differ significantly from SM estimates in several NP models. Due to the large uncertainties in SM calculations, the values are obtained for specific NP contributions alone. For example the value of x can be sensible to the mass and CKM elements of a possible fourth generation of quarks.

1.3 Belle II physics program

Research in flavour physics is an essential component of the current program in particle physics and the Belle II experiment at SuperKEKB is a crucial part of it. The search for New Physics has a plenty of experiments playing different roles in order to cover the widest possible range of topics, such as the SM propagator particles couplings, neutrino physics and all the processes involving loops that can give access to high mass scales and new particles.
The history of particle physics supports the idea that fundamental discoveries may come from the examination of the quark sector and in this field Belle II has a major role: it can measure



(a) **(b)**

Figure 1.6: **(a)**: The experiments and the domains to be explored. **(b)**: The different frontiers that must be pushed forward to explain the major puzzles in fundamental physics.

the current CKM matrix observables with unprecedented precision and allow a large number of observables and modes that are relevant to NP in the quark sector. The sensitivity to NP depends on the strength of the flavour violating couplings of the NP. The mass reach for new particle/process effects can be as high as $\mathcal{O}(100\,\text{TeV})$ if the couplings are not as suppressed as in the SM. As leading exponent of the intensity frontier research (Fig. 1.6b), Belle II effort will move towards the observation of NP effects and their characterisation through constraining measurements in several related flavour physics processes.
In the intensity frontier approach, the channels used to look for New Physics are the ones with reduced theoretical uncertainties, in such a way that the precise measurements can be compared with the SM calculation with the chance to find discrepancies which can be interpreted as New Physics signatures.
The main objectives of the experiment are [2]:

- Study the processes that occur at 1-loop in the SM but may be $\mathcal{O}(1)$ in NP: flavour-changing neutral currents (FCNC), neutral meson mixing, CP violation. These loops probe energy scales that cannot be accessed directly, even at the LHC.

- Search for NP by phases, CP asymmetries, inclusive decay processes, rare leptonic decays, absolute branching fractions.

- Search for new invisible particles: Belle II has a unique sensitivity via missing entry decays at the MeV/GeV scale for dark matter and Axion-like particles searches.

- Measure the CP violation in D^0 mixing.

- Elucidate the nature of strong force in binding hadron and QCD processes in the low-energy regime (quarkonia).

One of the main goal of Belle II is to examine the difference between B^0 and $\overline{B}^0$ decay rates, namely via measurements of time-dependent CP violation in penguin transitions of $b \to s$ and $b \to d$ (Fig. 1.7) quarks, such as $B \to \phi K^0$ and $B \to \eta' K^0$. CP violation in charm mixing, which is negligible in the SM, will also provide information on new phenomena in the up-type quark sector (more details in the next Sec. 1.5). Another key area will be to understand the mechanisms that produce large amounts of CP violation in the time integrated rates of

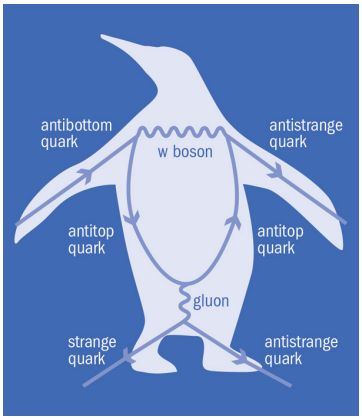


Figure 1.7: FCNC processes proceed to lowest order via one-loop diagrams, called penguins. This one shows a process for $\bar{b} \to \bar{s}$ transition.

charmless hadronic B decays, such as $B \to K\pi$ and $B \to K\pi\pi$, as observed by the B-factories and LHCb.
Within the program of precision measurements, CKM matrix elements, such as $|V_{ub}|$ and $|V_{cb}|$, are of primary importance. For example, leptonic and semileptonic decays of charm involve both well-understood weak interactions physics and non-perturbative strong-interaction effects. Leptonic decays of charm mesons are used to extract the product of a form factor, that can now be calculated with good precision via lattice QCD (LQCD), and a CKM matrix element: $|V_{cd}|f_D$ or $|V_{cs}|f_{D_s}$. The three possible procedures are:

1. inputting CKM matrix elements to yield a measurement of form factors and comparing these to theoretical calculations tests LQCD.

2. taking ratios of branching fractions such that CKM matrix elements cancel; this can provide a high precision test of LQCD.

3. inputting lattice QCD values for decay constants or form factors to yield a measurement of CKM matrix elements $|V_{cd}|$ and $|V_{cs}|$.

The e^+e^- facilities produce large samples of τ leptons, allowing for several measurements in Tau sector. For example, lepton flavour violating τ decays (which will be improved by orders of magnitude by Belle II), searches of charged Higgs-like coupling to τ leptons and CP violation in τ decays, possible both in the measurement of the cross section difference in τ as well as of various angular observables.
Finally, the prospects for quarkonium(like) physics, providing a detailed theoretical overview of perturbative QCD computation, lattice QCD as well as models for unconventional states (Tetraquark, Hybrid mesons and Hadronic molecule).

1.4 Current experimental status

D mesons are studied both at e^+e^- colliders and hadronic machines. The discovery of charm itself happened almost simultaneously in 1974 by two different experiments, one at the Brookhaven fixed target $p+$Be machine and the other one at SLAC's SPEAR e^+e^- collider. The resonance at 3.1 GeV was identified in the $c\bar{c}$ meson called J/ψ. Since then, an intense hunt for finding charm hadrons at accelerators was begun: the MARK I collaboration discovered the D^0 [10] and D^+ [11] mesons through narrow mass peaks in $K^-\pi^+$, $K^-\pi^+\pi^+$,

$K^-\pi^+\pi^+\pi^-$ channels. D mesons were soon thereafter detected also in neutrino-, hadron- and photon- induced reactions. Concerning charm lifetimes measurements,the first experimental evidence for weakly decaying charm hadrons was obtained in an emulsion experiment exposed to cosmic rays, whose resolving power was about $1\,\mu$m. Finally, thanks to the advent of microstrip vertex detectors the primary (production) and the secondary (decay) vertices of short living particles could be reconstructed with high precision, opening the field to search and study specific decay channels. Charm physics at hadron colliders was born with the success obtained by CDF at Tevatron and then continued by LHCb experiment at LHC, which has obtained unprecedented huge samples of charm decays, allowing mixing and CPV measurements. In this field LHCb is particularly sensitive to modes with all charged particles in the final state, while Belle II will be dominant in measurements involving neutral final states. Recent values for charm mixing and CPV, obtained combining Belle, *BABAR* and LHCb measurements, are [12]:

$$
\begin{aligned}
x_D &= 0.39^{+0.11}_{-0.12}\ \% \\
y_D &= 0.651^{+0.063}_{-0.069}\ \% \\
\left|\frac{q}{p}\right| &= 0.969^{+0.050}_{-0.045} \\
\phi &= \arg\left(\frac{q}{p}\right) = (-3.9^{+4.5}_{-4.6})^\circ\,.
\end{aligned}
$$

More generally, B-factories have a cleaner experimental environment with respect to a hadron collider, where high rates of QCD backgrounds and pile-up demand for complex trigger schemes and lead to worse reconstruction efficiency. The complex nature of hadrons implies that the initial state at every interaction point is unknown, both in terms of CM energy and of quark content. On the other hand, electrons are point-like, which allows to have a well-defined initial state in terms of $\sqrt{s}$. This makes e^+e^- colliders better suited for measurement of final states with missing energy and neutrals, whereas hadron machines are fare more competitive on rare decays with good signatures.
Large amounts of D meson pairs can be produced at e^+e^- colliders running as 'charm factories'; for example, CLEO-c at CESR, Cornell and BESIII at BEPC, Beijing, running at the $\psi(3770)$ resonance. This energy is just above the $D\overline{D}$ threshold but below the D^+D^- one, where the disadvantage of not being able to perform time-dependent analyses due to the low momenta is offset by the fact that the initial state is very clean (no fragmentation products accompanying the D mesons) and well constrained kinematically. B-factories however also produce very large samples of D mesons ($\sim 120\cdot 10^6\ c\,\bar{c}$ pairs per $100\,\text{fb}^{-1}$), both from continuum and from B decays, but hadronic colliders currently provide the most precise results.

LHCb measurement

Until LHCb observation all measurements in the D system were consistent with CP symmetry. Since its discovery, CPV measurements have been a crucial part of HEP research; indeed, through the precise measurements of the CKM matrix elements and the CPV parameters the SM robustness can be challenged and the explanations on many unresolved questions approached. This year, LHCb experiment has provided the first observation of CP violation in the decay of charm mesons [13]. Due to their very slow mixing rate ($x, y \sim 10^{-2}$) this kind of measurement requires very high statistics, favoured by the large production charm cross section at LHC: $\sigma(pp \to c\,\bar{c}\,X) \sim 2.8\,\text{mb}$ at $\sqrt{s} = 13\,\text{TeV}$.

LHCb is a dedicated experiment for B physics at the LHC: its detector is on one side only of the interaction region and has a large angular acceptance, resulting in a large effective bb cross section ($\sim 230\,\mu$b). Vertices are reconstructed by a silicon micro-strip detector VELO (VErtex LOcator) consisting of twenty-one stations with alternating r - φ strips. An aerogel ring imaging Čerenkov counter (RICH) is used for large-angle particle identification, and is followed by more silicon micro-strip detectors. The spectrometer consists of a 4 T m dipole field integral, followed by straw tracking chambers and a second small angle RICH. Additional particle identification capability over the large range of track momenta (1-100 GeV/c) is provided by a lead-scintillator electromagnetic calorimeter used to reconstruct photons with the help of a pre-radiator, and to identify electrons; a hadron calorimeter and muon chambers complete the set-up. As with other LHC experiments, the very high rates requires most of the detectors and front-end electronics to be radiation tolerant.
The LHCb trigger selects high transverse momentum muons and events with high transverse energy in the calorimeters, while rejecting events with multiple identified interaction vertices; the vertex tracker is used in the early stages of the trigger, providing a primary vertex resolution of about 60 μm along the beam axis and 20 μm transverse to it with a fast tracking algorithm; tracks originating from secondary vertices displaced by 0.1-3 mm from the primary one are then selected as B meson candidates. Running at a luminosity of $4\times10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, smaller than the typical figure for LHC, about 10^{12} $b\bar{b}$ pairs and D^*'s are produced each year. The large yield of charmed mesons and the huge number of decay modes of both B_d and B_s mesons makes LHCb research complementing the ones pursued at B-factories. Indeed, B-factories are optimized to run at the $\Upsilon(4S)$ and they do not produce any B_s mesons. For this reasons experiments at hadron colliders are the only ones that can do measurements of B_s decays and baryons containing b-quarks.
The approach followed at LHCb for the

$$A_{\mathrm{CP}} = \frac{\Gamma(D^0\to f)-\Gamma(\bar{D}^0\to \bar{f})}{\Gamma(D^0\to f)+\Gamma(\bar{D}^0\to \bar{f})} \tag{1.12}$$

measurement consists in measuring quantities where both production and detection asymmetries perfectly cancel out. In fact any asymmetry in the production of $D^0/\bar{D}^0$'s[3] or in the detection and reconstruction between positive and negative tracks can in principle induce a fake asymmetry in the final measurement.
The cancellation of all the asymmetries happens for $\Delta A_{\mathrm{CP}} \equiv A_{\mathrm{raw}}(f_1) - A_{\mathrm{raw}}(f_2)$ at very high level precision, where $A_{\mathrm{raw}}(f) = \frac{N(D^0\to f)-N(\bar{D}^0\to \bar{f})}{N(D^0\to f)+N(\bar{D}^0\to \bar{f})}$. The selected channels are $D^0\to h^+h^-$ with $h=\{K,\pi\}$, which are CP eigenstates and clean modes to search for direct CPV.
In general the flavour assignment of a candidate is made looking for decay channels that provide unambiguous flavour signatures. Flavour signatures correspond to signed characteristics correlated with the charge sign of the c-quark in the D meson at production. These characteristics can be the parent or decay products, or particular decay topologies. Discriminating variables are calculated for each particle candidate and dedicated algorithms based on multivariate methods provide a flavour tag q for the D^0 meson together with the corresponding flavour dilution factor r. The output $y\in[-1,1]$ of the flavour tagger is equivalent to $y = q\cdot r$, where $y=-1(+1)$ corresponds to a perfectly tagged $\bar{D}^0$ (D^0). For this analysis the flavour of the D^0 is inferred from the production mechanism, since the final states are CP eigenstates, not flavour ones. In order to measure CPV it is crucial to determine the flavour

[3] At $p-p$ colliders the production rate of D^0's is different from that of $\bar{D}^0$'s because the initial state is not symmetric under CP.

of the D^0 or $\overline{D}^0$ at production. This is commonly achieved selecting the D^0 coming from the $D^{*+} \to D^0\pi^+$ with the charge of the pion determining the charm-quark flavour of the neutral meson (Fig. 1.8a). The other production channel, shown in Fig. 1.8b, involves the (semi) leptonic decay of a B meson into a D, providing the flavour of the candidate.

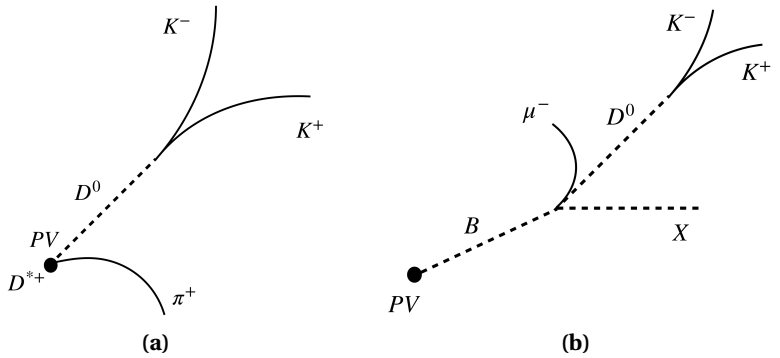


Figure 1.8: The two processes used to label the flavour of the D^0 candidate: **(a)** D^* - and **(b)** μ - tagged D^0's. PV stands for primary vertex; in the case **(b)** the production vertex of the D^0 is displaced from the PV since the finite decay length of the B is added.

The two processes under study are

I. D^***(or** π**)-tagged:** $D^{*+} \to [D^0 \to h^+h^-]\pi_s^+$

II. μ**-tagged:** $\overline{B} \to \mu^-[D^0 \to h^+h^-]\bar{\nu}_\mu X$

and the *raw* asymmetries measured counting for each flavour the number of decays in the given final state f:

I. $A_{\rm raw}^{\pi-{\rm tagged}}(f) = \frac{N_{D^{*+}}(f)-N_{D^{*-}}(f)}{N_{D^{*+}}(f)+N_{D^{*-}}(f)} \approx A_{\rm CP}(f) + A_D(\pi) + A_P(D^*)$

II. $A_{\rm raw}^{\mu-{\rm tagged}}(f) = \frac{N_{D^0}(f)-N_{\overline{D}^0}(f)}{N_{D^0}(f)+N_{\overline{D}^0}(f)} \approx A_{\rm CP}(f) + A_D(\mu) + A_{\rm P}(B)$

where $A_{\rm CP}$ is the CP asymmetry and A_D and A_P refer to, respectively, charge-dependent asymmetry in the reconstruction of the tagging particle (π_s or μ) and the D^0 production asymmetry. Reconstruction is made online with Turbo Stream [14] and candidates are filtered with MVA techniques. The approach to the measurement consists of

I.
- ⋆ Fit to $m(D^0\pi)$ mass
- ⋆ $A_{\rm raw}$ measured from simultaneous fit to D^{*+} and D^{*-} candidates

II.
- ⋆ Fit to $m(D^0)$ mass
- ⋆ $A_{\rm raw}$ measured from simultaneous fit to D^0 and $\overline{D}^0$ candidates

The effects of direct CPV can be isolated by taking the difference

$$\Delta A_{\rm CP} \equiv A_{\rm CP}(K^+K^-) - A_{\rm CP}(\pi^+\pi^-) \approx \Delta a_{\rm CP}^{\rm dir}\left(1 + \frac{\overline{\langle t\rangle}}{\tau_D} y_{\rm CP}\right) + \frac{\Delta\langle t\rangle}{\tau_{D^0}} a_{\rm CP}^{\rm ind}$$

where $\Delta a_{\rm CP}^{\rm dir} = a_{\rm CP}^{\rm dir}(K^-K^+) - a_{\rm CP}^{\rm dir}(\pi^-\pi^+)$, $a_{\rm CP}^{\rm ind}$ is a possible indirect CPV source in mixing or interference between mixing and decay, $\overline{\langle t\rangle}$ is the arithmetic average between the mean decay times for the two channels and $\Delta\langle t\rangle$ is their difference.

Indeed, this allow to eliminate possible instrumental and production asymmetries and also residual indirect CPV contributions due to possible decay time dependent acceptance function, which can be different for the two channels.
The main systematic uncertainties evaluated are related to the fit model, the bias due to residual contamination of D^{*+} decay from B's, mis-reconstructed background and the different efficiency between the two final states, as well as B^0 oscillations. Furthermore, extensive robustness studies have been performed, leading to the final result [13]

$$\Delta A_{\rm CP} = (-15.4 \pm 2.9) \times 10^{-4} \,,$$

which is consistent with SM expectations, although theoretical predictions still have large uncertainties. This observation triggers a new chapter in flavour research and encourages further theoretical studies as to give even tighter constraints to SM.
After this discovery of direct CPV in charm, the goal is to measure CPV in as many modes as possible; LHCb is expected to be the strongest player for charged final states, while Belle II having a unique role for final states with neutrals (more details are give in Sec. 1.5).
The second approach towards $A_{\rm CP}$ measurements would be to obtain direct $A_{\rm CP}$, as to complete the sketched picture. This needs excellent detector calibration and simulation and many Cabibbo-favoured normalisation channels.

1.5 Belle II charm physics program

Charm quark occupies a unique place among up-type quarks since it is the only one whose hadronization and subsequent weak decay can be studied[4]. Indeed studies of transitions involving the charm quark play an important role in both searches for NP and in understanding QCD. Decays of charmed mesons and baryons probe a variety of NP scenarios, e.g., couplings to intermediate charged Higgs states, and decays to light dark matter [15]. Decays of the down-type quarks s and b are very promising to reveal new physics since their CKM couplings are greatly suppressed, especially for b. This is not the case for up-type quarks. Yet NP could quite conceivably induce flavour changing neutral currents that are larger for the up-type than the down-type quarks. In that case charm decays would be best suited to reveal such non-standard dynamics. For all these reasons, charm transitions are a unique portal for obtaining a novel access to flavour dynamics.
The Belle II experiment is ideal for searching for time-integrated CP-violating effects in a variety of final states, and will reach precisions of the order of 0.01% level. The LHCb experiment has provided extremely precise measurements of CP asymmetries in decays with charged particles in the final state, and also the first observation of CPV in charm (as described in Sec 1.4). The excellent γ and π^0 reconstruction (and thus η, $\eta^{'}$, etc.) will allow Belle II to search for CP violation in complementary final states that contain neutral particles. The high flavour-tagging efficiency, the numerous control samples with which to study systematics, in addition to the excellent reconstruction of charged particles will allow Belle II to compete with LHCb in measurements of time-integrated CP violation. For example, studying CP eigenstates like $D^{*+} \to D^0[K_S\pi^0]\pi^+$, which need a pair of pions with a displaced vertex and two photons measured with good resolution and low background, is beyond LHCb capabilities. The Belle II potential for charm physics lies in the PID performance, vertex resolution and neutral states reconstruction capability. Reconstruction efficiency for charm is expected to be higher at

[4] for example, top quark decays before it can hadronize.

Belle II than it was at Belle or *BABAR*. The improved vertex resolution should result, for example, in improved rejection of random combinations of tracks and improved resolution on decay lengths and flight times. The expected precision on $A_{\rm CP}$ for the final Belle II dataset are collected in Tab. 1.1, where it is possible to make a direct comparison with the results of Belle. The expected precision for Belle II is estimated by scaling the Belle statistical error $\sigma_{\rm stat}$ by the

Mode	L (fb^{-1})	Belle $A_{\rm CP}$ Measurement (%)	Belle II extrapolated error at 50 ab^{-1} (%)
$D^0 \to K^+K^-$	976	$-0.32 \pm 0.21 \pm 0.09$	0.03
$D^0 \to \pi^+\pi^-$	976	$+0.55 \pm 0.36 \pm 0.09$	0.05
$D^0 \to \pi^0\pi^0$	966	$-0.03 \pm 0.64 \pm 0.10$	0.09
$D^0 \to K^0_S\pi^0$	966	$-0.21 \pm 0.16 \pm 0.07$	0.02
$D^0 \to K^0_S K^0_S$	921	$-0.02 \pm 1.53 \pm 0.02 \pm 0.17$	0.23
$D^0 \to K^0_S\eta$	791	$+0.54 \pm 0.51 \pm 0.16$	0.07
$D^0 \to K^0_S\eta'$	791	$+0.98 \pm 0.67 \pm 0.14$	0.09
$D^0 \to \pi^+\pi^-\pi^0$	532	$+0.43 \pm 1.30$	0.13
$D^0 \to K^+\pi^-\pi^0$	281	-0.60 ± 5.30	0.40
$D^0 \to K^+\pi^-\pi^+\pi^-$	281	-1.80 ± 4.40	0.33
$D^+ \to \phi\pi^+$	955	$+0.51 \pm 0.28 \pm 0.05$	0.04
$D^+ \to \pi^+\pi^0$	921	$+2.31 \pm 1.24 \pm 0.23$	0.17
$D^+ \to \eta\pi^+$	791	$+1.74 \pm 1.13 \pm 0.19$	0.14
$D^+ \to \eta'\pi^+$	791	$-0.12 \pm 1.12 \pm 0.17$	0.14
$D^+ \to K^0_S\pi^+$	977	$-0.36 \pm 0.09 \pm 0.07$	0.02
$D^+ \to K^0_S K^+$	977	$-0.25 \pm 0.28 \pm 0.14$	0.04
$D^+_s \to K^0_S\pi^+$	673	$+5.45 \pm 2.50 \pm 0.33$	0.29
$D^+_s \to K^0_S K^+$	673	$+0.12 \pm 0.36 \pm 0.22$	0.05

Table 1.1: Time-integrated CP asymmetries measured by Belle, and the precision expected for Belle II in 50 ab^{-1} of data.

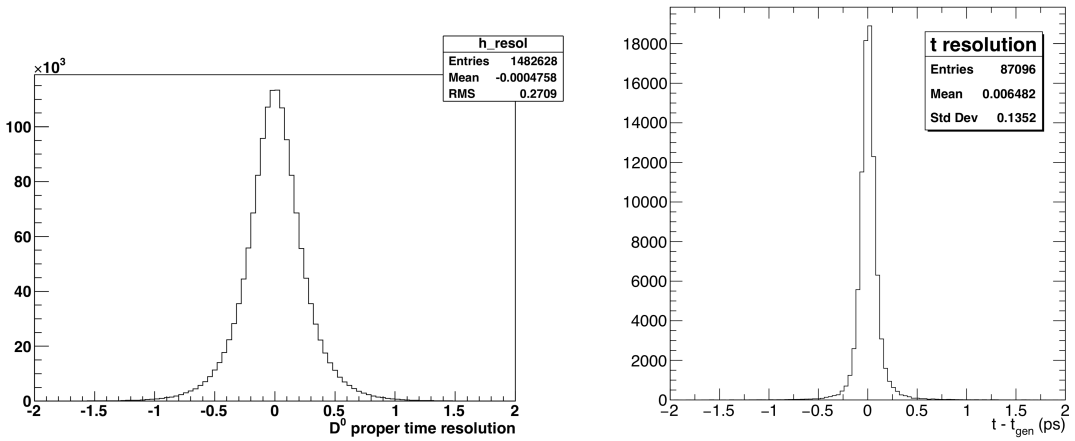
ratio of integrated luminosities, and dividing the systematic error into those that scale with luminosity — such as background shapes measured with control samples ($\sigma_{\rm syst}$), and those that do not scale with luminosity — such as effects due to detector misalignment ($\sigma_{\rm irred}$). The overall error estimate is calculated as

$$\sigma_{\rm Belle\,II} = \sqrt{(\sigma^2_{\rm stat} + \sigma^2_{\rm syst}) \cdot (L_{\rm Belle}/50\,{\rm ab}^{-1}) + \sigma^2_{\rm irred}}\,.$$

However, this estimate does not take into account the expected improvements in reconstruction at Belle II.

Beside time-integrated measurements, Belle II will produce a large number of time-dependent measurements sensitive to D^0 mixing and CP violation. The Belle II vertex detector allows to reconstruct the D^0 decay vertex with a precision of $\sim 40\,\mu$m, significantly improving Belle and *BABAR* performance. This ability is due to the reduced distance between the first vertex detector layer and the IP. The resolution on the D^0 decay time is improved by a factor two, and this should greatly improve the precision of time-dependent measurements of D^0- $\overline{D}^0$ mixing and searches for CP violation. From simulations the average proper time error σ_t is expected to be a factor of three smaller at Belle II, and the proper time resolution a factor of two smaller with respect to *BABAR* (Fig. 1.9).

Improvements and the enormous statistics expected for Belle II give very good prospects for measuring mixing and CPV parameters. The predictions about the accuracies that can be



(a) **(b)**

Figure 1.9: Resolution on D^* -tagged D^0 proper time at **(a)** *BABAR* and **(b)** Belle II (simulation). The same plot as **(a)** was not available for Belle, but the comparison holds as the performance on this channel was comparable.

obtained have been estimated with Toy MC, plugging the proper time resolution estimated from the simulation. This study indicates that the better precision achievable by Belle II is not only due to larger statistics, but it will benefit from a better-performing vertex detector, and tracking. The results of the study is reported in Tab. 1.2, where $x' = x\cos\delta + y\sin\delta$, $y' = y\cos\delta - x\sin\delta$, and δ is the strong phase difference between $D^0 \to K^+\pi^-$ and $\overline{D}^0 \to K^-\pi^+$ amplitudes. These estimations need to be confirmed on real data and this work is the first

Error on	**scaled at** $50\,\text{ab}^{-1}$ **from Belle**	**Toy MC with improved** σ_t**, CPV**
$x'(\%)$	0.45	0.15
$y'(\%)$	0.16	0.10

Table 1.2: From [2], uncertainties on mixing parameters resulting from fitting Belle II samples of $D^0 \to K^+\pi^-$ and $\overline{D}^0 \to K^-\pi^+$ decays, allowing for CP violation.

step on that direction (as discussed in the next chapters).
In summary, the large yields of charm produced at SuperKEKB make the charm physics program of Belle II quite broad. It will not be limited to mixing and CPV searches, but also studies on leptonic and semileptonic decays, hadronic decays, and provide useful inputs for QCD and measurements of the CKM parameters. Finally, the cleanliness of the event will allow to search for rare and forbidden decays, also with missing energy; as an example, a search of $D \to$ invisible can be pursued at Belle II and any signal observation would be NP or a further constraint to be put int the SM.
For what concerns CP violation measurements, Belle II will measure A_{CP} for different final states (as shown in Tab. 1.1) as the initial state is symmetric under CP and there are no production asymmetries. This will provide additional information with respect to the ΔA_{CP}, which is important for the understanding of the nature of CP violation.

2 SuperKEKB and Belle II

This chapter intends to describe the main features of the SuperKEKB accelerator as a new generation B-factory. The machine upgrade and the physics goals have required a profound transformation to the detector, which had to overcome several technological challenges. A general picture of the experiment is presented, from the sub-detectors architecture to the software and analysis tools, giving more details on the topics which will be useful for the rest of the dissertation.

2.1 B-Factories

In an electron-positron collider the most efficient way to produce B^0 mesons is to tune the energy to the $\Upsilon(4S)$, the lightest strong resonance - the third among the radial excitations of the Υ - with a mass sufficient to decay in b-flavoured mesons ($B^0\overline{B}^0$ and B^+B^- pairs). Indeed the beauty meson threshold is at $2m(B) \simeq 10.56\,\text{GeV}/c^2$ and $m_{\Upsilon(4S)} \simeq 10.58\,\text{GeV}/c^2$.
Experimentally, an e^+e^- collider experiment is ideal for studying rare or forbidden decays, since backgrounds are much lower than that at a hadron machine and trigger efficiencies are higher. Furthermore, the average track multiplicity is low[1], as well as the detector occupancy, thus making the B, D and τ reconstruction efficiency high and the trigger bias low. This reduces systematic uncertainties in many measurements, such as Dalitz plot analyses. There are usually numerous control samples available with which to study backgrounds and estimate systematic uncertainties. Since the recorded luminosity can be determined by measuring Bhabha scattering, absolute (in addition to relative) branching fractions can be measured.

The precisely known center-of-mass energy and the excellent detector hermeticity are indispensable to perform "missing mass" analyses, where the existence of new particles can be inferred via energy/momentum conservation.
The branching fractions for the decay of the $\Upsilon(4S)$ into BB pair is over 96%. The pair of B mesons results entangled in the decay and the $\Upsilon(4S)$ in unflavoured, which implies that from the knowledge of the flavour of one B it is possible to assign the flavour of the second B too. The two B mesons are produced with small momenta and the current tracking resolutions do not allow to accurately measure the difference in distance from the collision point of the two decays. Furthermore, the size of the beam-beam interaction region contributes to make this difference irresolvable. For CPV measurements, where the time difference between the decays is needed, the two storage rings must have different energies so that the $\Upsilon(4S)$, and likewise the pair of B's to which it decays, are produced with a significant relativistic boost along the beam axis, allowing precision measurements of lifetimes and mixing and CPV parameters.

[1] a typical $\Upsilon(4S)$ event has 11 charged tracks, 5 neutral pions and 1 neutral kaon, for an average composition of: $\pi^\pm$ (72.8%), $K^\pm$ (14.9%), $e^\pm$ (5.8%), $\mu^\pm$ (4.7%), $p^\pm$ (1.8%).

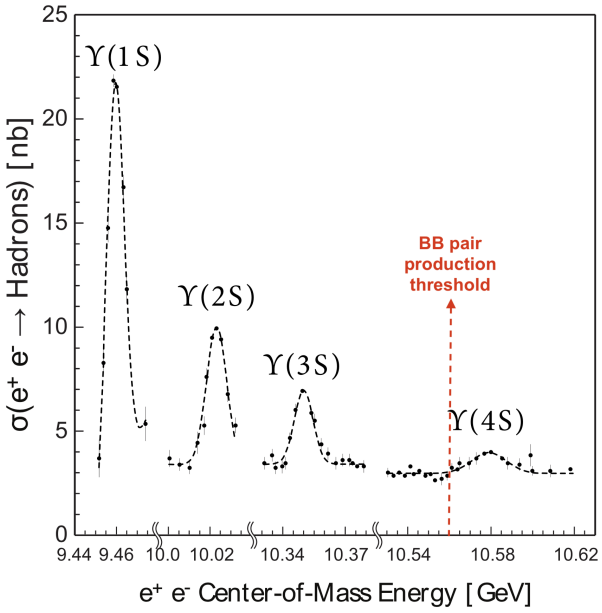


Figure 2.1: Hadronic cross section as a function of the center-of-mass energy/c^2 in the region of the first four S-wave Υ resonances.

$e^+e^- \to$	Cross Section
$\Upsilon(4S)$	1.110 ± 0.008 [16]
$c\bar{c}$	1.30
$u\bar{u}$	1.61
$d\bar{d}$	0.40
$s\bar{s}$	0.38
$\tau^+\tau^-(\gamma)$	0.919
$\mu^+\mu^-(\gamma)$	1.148
e^+e^-	300 ± 3

Table 2.1: Production cross sections for e^+e^- processes at $\Upsilon(4S)$ resonance: $\sqrt{s} = 10.58\,\text{GeV}$. The cross section for Bhabha $e^+e^- \to e^+e^-$ is computed cutting on the following variables, in the center-of-mass frame:$E_e^* > 0.15\,\text{GeV}$ and $10° < \theta_e^* < 170°$. Adapted from [17].

2.2 SuperKEKB accelerator

The Belle II experiment is installed at the SuperKEKB accelerator in Tsukuba, Japan. The machine is currently the only running e^+e^- B-factory in the world with a design luminosity of $8 \times 10^{35}\,\text{cm}^{-2}\text{s}^{-1}$ [17]. This luminosity, about 40 times larger than that of KEKB, will allow to produce approximately 5×10^{10} b, c and τ pairs by 2027. The increase in luminosity is obtained through a combination of drastic reduction in the beam dimensions (*nano-beam scheme* [18]) and a moderate increase in beam currents. This design requires a larger crossing angle and a smaller energy asymmetry compared to KEKB. The increase in the beam energy of the LER (Low Energy Ring) from 3.5 to 4.0 GeV helps mitigate the emittance growth due to intra-beam scattering and the short beam lifetime due to Touschek effect. In addition, the decrease in the beam energy of HER (High Energy Ring) from 8.0 to 7.0 GeV is beneficial in obtaining a lower power consumption. The resulting boost

$$\beta\gamma = \frac{\left|\vec{p}_{e^-} + \vec{p}_{e^+}\right| c}{\sqrt{s}} \simeq \frac{E_{e^-} - E_{e^+}}{\sqrt{4E_{e^-}E_{e^+}}} \simeq 0.28$$

is around two-thirds of that at KEKB. This condition is advantageous for analyses with neutrinos in the final state that require good detector hermeticity, but lead also to less separation between the B vertices (average distance between the two B mesons of about 130 µm, 35% smaller than the 200 µm of KEKB). However, the beam pipe at the interaction region has a radius of 10 mm, two-thirds of that in KEKB, so that the innermost layer of Belle II's vertex detector can be positioned at approximately half the distance with respect to Belle, leading to an improvement of a factor close to two on the impact parameter resolution. This is described in Chapter 3, where several indications of the performance achieved by the detector are given.

The luminosity of a collider is expressed by the following formula:

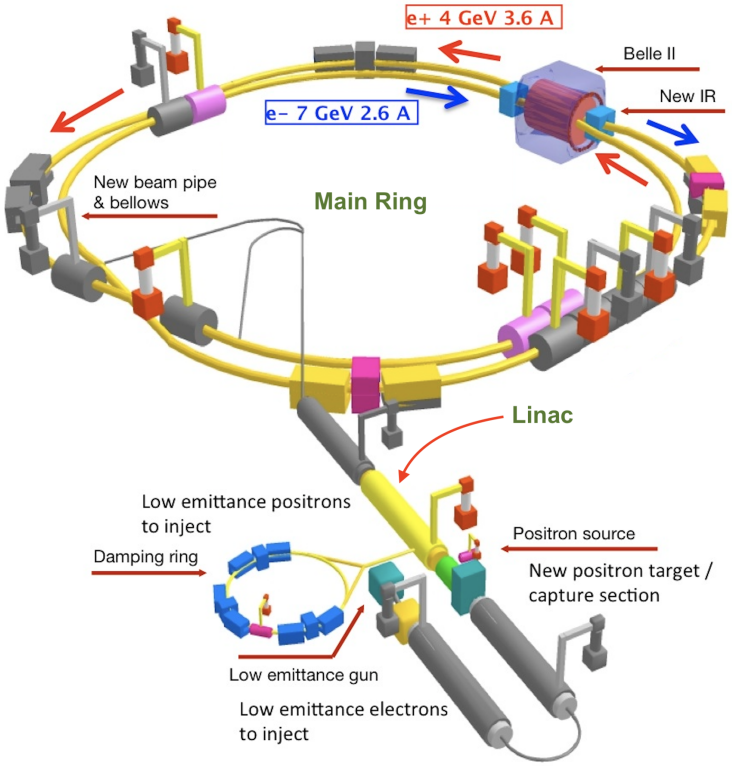


Figure 2.2: SuperKEKB accelerator complex with the upgraded and the new components. The linear accelerator (Linac) provides the 7-GeV e^- and the 4-GeV e^+ beams to the main ring. Electrons destined to the main ring (injection e^-'s) are produced by a photocathode RF gun in a new pre-injector and accelerated by the Linac. Electrons for positron production (primary e^-'s), instead, are supplied by the old pre-injector. A thermionic electron gun is used for the production of the primary electrons because a high-current electron beam is required, whereas the injection e^-'s must have a low emittance. The primary e^-'s are accelerated up to 3.5 GeV and irradiate a tungsten target located in the middle of the Linac to generate the positrons from the capture section. This process generates e^+'s with a larger emittance than that needed in the nano-beam scheme. To compensate for this, a 1.1 GeV damping ring (DR) is introduced; the e^+'s are circulated there and once the emittance is optimised, they are re-injected in the Linac, where they are accelerated up to 4.0 GeV.

$$L = \frac{\gamma_\pm}{2e\,r_e}\left(1+\frac{\sigma_y^*}{\sigma_x^*}\right)\left(\frac{I_\pm \xi_{y_\pm}}{\beta_y^*}\right)\left(\frac{R_L}{R_{\xi_{y_\pm}}}\right), \qquad \xi_{y\pm} \propto \frac{N_\mp \beta_y^*}{\sigma_x \sigma_y} \tag{2.1}$$

where "+" is for positrons and "-" for electrons. The $\sigma_{x,y}^*$ parameters denote the beam sizes at the IP in the horizontal and vertical plane, I indicates the total beam current, ξ_y is the vertical beam-beam parameter and β_y^* the vertical beta function at the IP. R_L and $R_{\xi_{y\pm}}$ are the reduction factors for the luminosity and the beam-beam parameter, r_e the classical electron radius, e its charge and γ is the Lorentz factor. For a more detailed description of the parameters, and their interplay in determining the machine performance, please refer to [19]. It is assumed that the vertical beta function and the transverse beam size of the positron beam are the same as the electron beam. Therefore, the formula tells us that the luminosity is proportional to the beam-beam parameter and the beam current, and inverse of the vertical beta function at the IP. The vertical beam-beam parameter $\xi_{y_\pm}$ (Eq. 2.1) describes the focusing force exerted on a bunch by the electromagnetic field of the opposite bunch and depends on $N_\mp$, the number of particles (e^-/e^+) in a bunch. The beam-beam parameter is limited by the nonlinear force in a circular collider and a maximum value of about 0.02-0.1 has been

obtained from various colliders in 40 years. In order to increase the luminosity one can increase I and lower the beam sizes. Indeed, at SuperKEKB the currents are increased from 1.6A/1.2 A to 3.6/2.6 A, while the bunches vertical size σ_y is reduced from 1.1 µm to ∼ 60 nm. However, these manipulations lead to an increase of the non-linear beam-beam interactions which cause lifetime reduction, beams blowup and other instabilities. Besides, the transverse beam size varies along the bunch length: the minimum is at the collision focal point and this value is kept only over a short distance, of the order of β_y^*, then it rapidly grows towards the tails. Since the bunch length (σ_z, of the order of mm) is longer than β_y^*, the actual luminosity is usually smaller than the nominal one; this effect is known as the *hourglass effect*.
The main new feature in the upgrade of KEKB is the *nano-beam* scheme, where the β_y^* is squeezed down to 0.3 mm (20 times smaller than at KEKB), leading to a vertical beam size of only $\simeq 60$ nm.

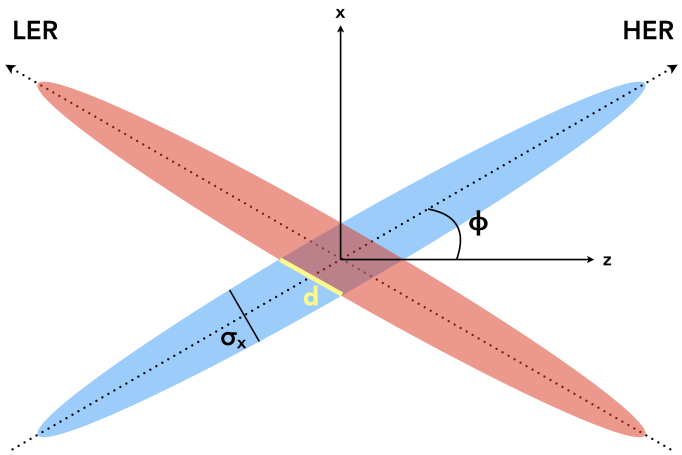


Figure 2.3: Pictorial representation and not to scale of the crossing of the two beams, the LER (Low Energy Ring) and the HER (High Energy Ring). d is the dimension of the overlap region, in relation with the σ_x and the half crossing angle according to Eq. 2.2.

However the effective bunch length for the *nano-beam scheme* corresponds to the overlap region of the beams at the IP, which is much smaller than the bunch length σ_z. The length d (refer to Fig 2.3) is determined by the horizontal half crossing angle (ϕ) and the horizontal beam size at the IP (σ_x^*) via the following equation:

$$d = \frac{\sigma_x^*}{\sin 2\phi} \,. \tag{2.2}$$

The design parameters of SuperKEKB allow to fulfil the *hourglass* requirement $d \leq \beta_y^*$: the length d is shortened using a relatively large horizontal crossing angle and extremely small horizontal emittances and horizontal beta functions at the IP for both beams. The half crossing angle ϕ is 41.5 mrad, which is about 4 times larger than that of KEKB. This choice, in addition to contributing to lower the effective bunch length, is mainly determined by considerations related to the optics of the interaction region (IR), magnet design, and the detector background. With a large crossing angle, the final focus quadrupole magnets can be independent for the two beams, which has the merit of reducing the detector background due to synchrotron radiation (see next section).
A three-dimensional rendering of the nano-beam scheme is proposed in Fig. 2.4,in comparison with previous configurations as to give an insight on the impact on the crossing region dimension.

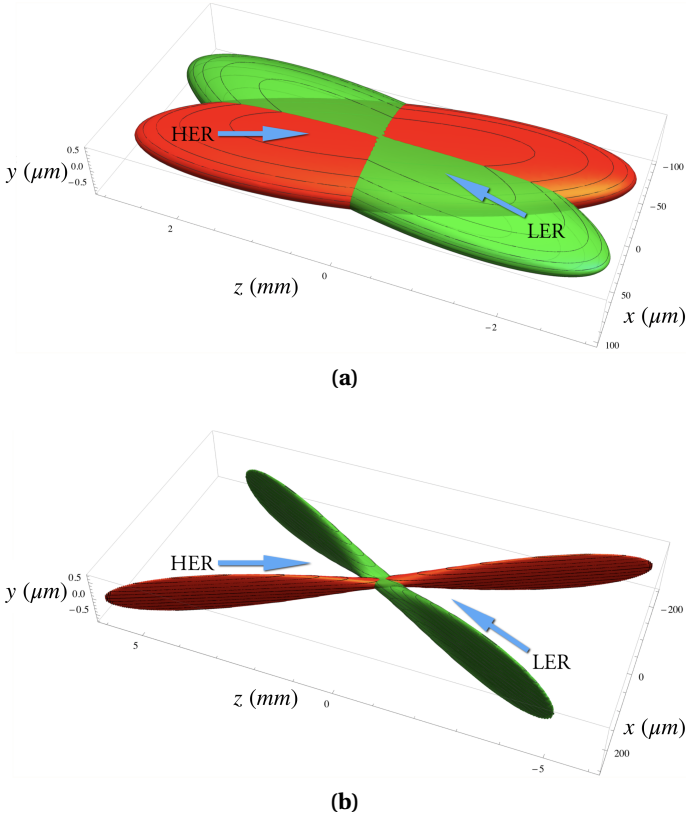


Figure 2.4: Bunches constant density surfaces at the IP in first-generation B-factories **(a)** and those obtained with a nano-beam **(b)** configuration.

As mentioned before, $\xi_{y\pm}$ is a crucial parameter for the luminosity. For low beam currents, the beam-beam parameter increases with the number of particles in a bunch. However, for high beam currents, the emittance increases as the two beams disrupt each other, and the increasing particle number is compensated by the increasing beam size. In that limit, the beam-beam parameter saturates and the luminosity depends only linearly on the beam currents. With vertical beta functions at the IP smaller by almost by a factor of 20 and the currents doubled with respect to those at KEKB, the total luminosity increase is expected to reach a factor of 40 compared to KEKB, assuming that the same beam-beam parameter value (0.09) can be achieved. Some of the main parameters mentioned so far are shown in Tab. 2.2, usefully comparing the values obtained during KEKB operations with those SuperKEKB aims to reach. In the third column a feedback on the performance achieved during the first months of the experiment is given.

2.2.1 Beam losses and background sources in e^+e^-storage rings

There are two types of background present at SuperKEKB [20, 21]: single-beam backgrounds and luminosity-dependent ones. The first class of processes can originate from the interaction between a particle of a bunch with residual gas in the beam pipe or other particles within the same bunch, whereas luminosity-dependent backgrounds are induced by beam collisions and include QED (Quantum Electrodynamics) processes, such as (radiative) Bhabha scattering $e^+e^- \rightarrow e^+e^-(\gamma)$ or two-photon processes $e^+e^- \rightarrow e^+e^-e^+e^-$. These various effects, which combined determine the total lifetime of the beams and the amount of background, are briefly described here.

	KEKB Achieved	**SuperKEKB Design**	**SuperKEKB Achieved**$^{(a)}$
Energy (GeV) (LER/HER)	3.5/8.0	4.0/7.0	4.0/7.0
ξ_y	0.129/0.090	0.090/0.088	0.0266/0.0165
β_y^*(mm)	5.9/5.9	0.27/0.41	3/3
I (A)	1.64/1.19	3.60/2.62	0.65/0.6
Luminosity ($10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$)	2.11	80	0.5

Table 2.2: Fundamental design parameters of SuperKEKB and the values achieved by KEKB. $^{(a)}$ The achieved parameters at SuperKEKB refer to Phase 3 Spring Run and are not necessarily obtained at the same time. A brief history of Belle II operations is reported in Sec.2.4.

1. SINGLE-BEAM BACKGROUND

   1.1. *Touschek effect* consists in a collision between two particles of the same bunch, which can transfer momentum from transverse to longitudinal motion. After a certain number of turns of the beams, the particles can exceed the longitudinal acceptance: in this case they are lost and can hit the beam pipe, eventually producing electromagnetic showers that might reach the detector. The high density of particles into a single bunch clearly increases the probability of such effect to happen. Moreover Touscheck effect is inversely proportional to the first power of σ_y and to the third power of the energy of the beam; for this reason it is expected to be a non-negligible background source at SuperKEKB, due to the small σ_y adopted in the nano-beam scheme, even more for LER than for HER because of the lower energy.

   1.2. The beam particles can undergo Coulomb scattering with residual gas molecules inside the beam pipe, leading to immediate beam losses. This effect depends on the pressure over the beam pipe, which in turn depends on the current of the beam. Indeed the pressure can be approximated as a linear function of the beam current, with a constant corresponding to the pressure where no beam is stored and a linear positive term resulting in a *dynamic pressure*.

   1.3. Synchrotron radiation (SR): Particles of the LER and HER beams are transversely accelerated by the bending and focusing magnets of the accelerator, which leads to the emission of synchrotron radiation. Due to the beam correction magnets and the QCS (Compensating Superconducting Quadrupoles), located at about 1 m from the IP, the emitted photons can reach the Belle II vertex detector. This is not enough for e^+e^- pair production (photons' energy is from few keV to tens of keV), but the photons can still create hits in the vertex detector. Dedicated masking structures are installed inside the beam pipe to absorb such radiation. In addition, a thin layer of gold is deposited on the beam pipe surface around the IP to absorb the low-energy residual photons. Being the SR power proportional to the beam energy squared and magnetic field strength squared, HER mainly contributes to this type of background.

2. LUMINOSITY BACKGROUND

   2.1. Two-photon process $e^+e^- \rightarrow e^+e^-e^+e^-$: it creates four low-momentum electrons and positrons at the IP, which are bent by the magnetic field into spirals inside the vertex detector volume, leaving multiple hits on its components.

2.2. Radiative Bhabha scattering (RBB) $e^+e^- \to e^+e^-\gamma$: electron and positrons can scatter and produce photons at small angles. Such photons along the beam axis can hit materials and produce showers. However the scattering angle is very small, and therefore most particles that have been scattered leave the beam pipe far from the vertex detector, making this background source less effective than the previous one. In any case the rate of RBB and two-photon process depends on the luminosity and becomes predominant at the design luminosity of SuperKEKB.

A radiation monitoring and a Beam Abort system based on diamond sensors is installed at the interaction region, with the purpose of measuring the background level and providing protection for the detector and accelerator components.

2.3 Belle II Detector, a universal spectrometer

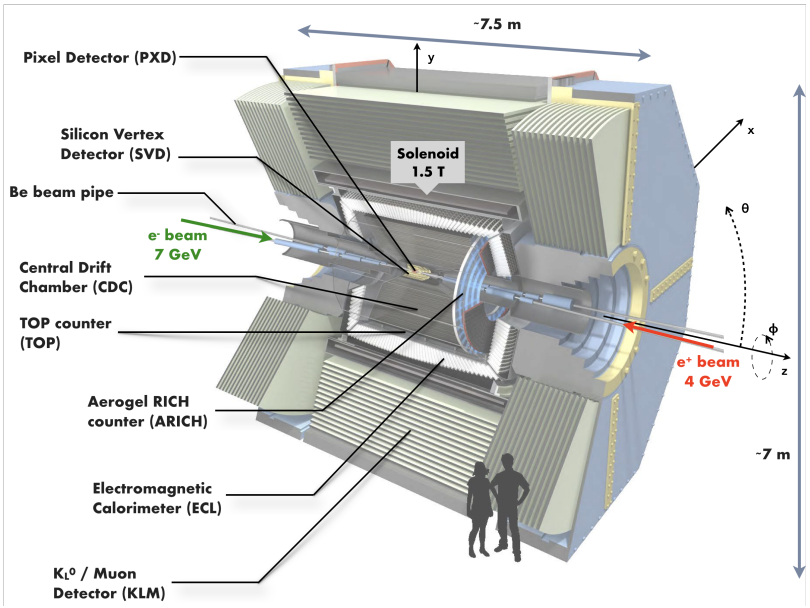


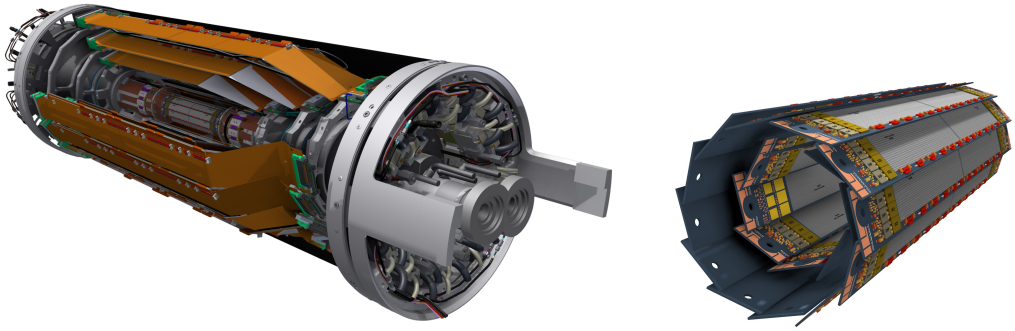
Figure 2.5: Belle II detector. Reported are the axes of the reference frame: the x axis points along the radial direction towards the outside the ring while the y axis indicates the vertical upward direction . The positive z axis points towards the direction of the magnetic field of the solenoid; positive z coordinates correspond to the *forward region*, those with negative z belong to the *backward region*. The polar angle $\theta \in [0,\pi]$ in the $z-y$ plane and the azimuthal angle $\phi \in [-\pi,\pi]$ in the $x-y$ plane.

The upgrade of KEKB to SuperKEKB has required a substantial renewal of Belle II detector, whose performance are challenged by radiation damage and occupancy issues, higher event rate (×50 for nominal luminosity) and higher background (i.e. ×20 for Touschek background). This requires fast and efficient trigger system, as well as data acquisition system capable of processing large quantities of data. The aim of such an upgrade is to guarantee in SuperKEKB environment equal or better performance than those achieved by Belle with KEKB [22]. A picture of the experiment is given in Fig. 2.5.
Like the previous detectors at B-factories (Belle and *BABAR*), Belle II keeps the typical design features dictated by the physics requirements [9]. It must provide a large acceptance and a high efficiency for B decays to maximise the number of collected events, good momentum and energy resolutions to separate the small signal from the backgrounds, very good vertex position resolution, efficient and clean particle identification capabilities for hadrons. Since

the average charged particle momentum is below 1 GeV/c, the minimisation of the amount of material producing multiple scattering is vital. The same necessity holds for the amount of dead material in front of the electromagnetic calorimeter, which has to detect showers with energies as low as 20 MeV. The detector is offset with respect to the IP, reflecting the asymmetric energy configuration, and comprises: silicon vertex detectors close to the beam pipe for secondary vertex measurement and to provide the first tracking points, drift chamber with helium-based gas for charged particle tracking and particle identification through ionisation and energy loss, Čerenkov detectors for particle identification and the CsI(Tl) crystal calorimeter for the measurement of neutral particles. The high light yield and small Molière radius of CsI crystals allow excellent energy and angular resolutions; moreover the high yield permits the use of silicon photodiodes, which can work in magnetic fields; all the above detectors are inside a 1.5 T cryogenic superconducting solenoid at a temperature of 4.5 K. The magnetic flux return yokes are used to absorb hadrons and contain scintillator bars and resistive plate chambers (RPCs) to perform as muon and neutral hadron detectors. Two-level hardware-software trigger systems cope with the full rate of e^+e^- interactions. Belle II reuses the structure, the solenoid, the CsI(Tl) crystals and part of the barrel RPCs from the original Belle detector, while most other components of the sub-detectors are new and their development has involved the collaboration among many nations.
This brief introduction is followed by an overview of the various sub-detectors, focusing on the necessary upgrades and the expected performance.

2.3.1 Vertex Detector (VXD)

The new VXD is comprised of two devices, the silicon Pixel Detector (PXD) and a Silicon Vertex Detector (SVD), with altogether six layers around a 10 mm radius Be beam pipe. The outermost layer is at 135 mm, against the 80 mm of the past detector, while the beam pipe and the first 2 detector layers are closer to the IP (respectively at 14 mm and 22 mm). As a result, a significant improvement is expected in the vertex resolution, as well as in the reconstruction efficiency for $K_s \to \pi^+\pi^-$ decays with hits in the vertex detector. Due to the higher beam current and luminosity, the detectors are required to operate at up to 30 times larger beam background and up to 30 kHz trigger rate. More details on the detector performance can be found in Sec. 3.3.1.



(a) SVD **(b)** Magnified view of PXD

Figure 2.6: The Belle II Vertex Detector volume. The four outer layers that build up the SVD **(a)**; some modules are taken away in this artistic rendering to make visible the PXD in the center, which is magnified in **(b)**.

The PXD, the closest sub-detector to the IP, has to face extremely high hit rates, caused by beam-related background; for example, low-momentum-transfer QED processes. The 10-

mm beam pipe radius has a good impact on the physics analysis related to vertex reconstruction, but is a challenge for the vertex detector itself because the background increases roughly with the inverse square of the radius. The innermost layers of a high precision vertex detector can no longer be realized by strip detectors due to the large occupancy, defined as the fraction of channels hit in each triggered event. The solution is to use pixel sensors rather than strips for the innermost layers, which have a much larger number of channels and therefore a much smaller occupancy. For the luminosity envisaged at SuperKEKB, strip detectors are safe beyond a radius of 30 mm. This scheme — pixels followed by strips — has been successfully applied for the detectors at the LHC. However, at the lower energies of the SuperKEKB machine, the hybrid pixel LHC silicon detectors are too thick: the relatively large amount of material of these pixel sensors would cause too much multiple scattering for a precise reconstruction of B-decay vertices at the energies of few GeV. The PXD is based on the DEPFET (DEPleted Field Effect Transistor) technology which allows for very thin (50 µm) sensors. The sensitive area of the PXD is assembled in 20 ladders and 40 modules, where each module consists of a 250×768 pixel matrix of DEPFET pixel sensors.The readout electronics, which needs active cooling, realized with the evaporative CO_2 cooling technique, is located outside the acceptance region and therefore does not contribute to the multiple-scattering material budget. The sensors themselves consume very little power so that air cooling is sufficient. In addition, these sensors can be engineered in an inherently radiation-hard technology.

The outermost detector of VXD, the SVD, covers the full Belle II angular acceptance of $17° < \theta < 150°$, with the inner radius at 39 mm (Layer 3) and outer radius at 135 mm (Layer 6); these are determined by the radii of the PXD and CDC. It inherits the good characteristics of the Belle vertex detector: low mass, high precision, immunity to background hits, radiation tolerance and long-term stability. It is designed as a 4 layers of double-sided silicon strip sensors to avoid the huge channel count of pixels without compromising the vertex-detection capability of Belle II. These devices simultaneously measure two coordinates of the traversing charged particle. A depleted silicon bulk is sandwiched by a p—type and a n—type doped layer, which is segmented into strips. The strips are implanted into the bulk and oriented orthogonal to each other. A voltage is applied between the two strip layers in order to deplete the silicon. The sensors used for the Belle II SVD have rectangular or trapezoidal shape (Tab. 2.3).

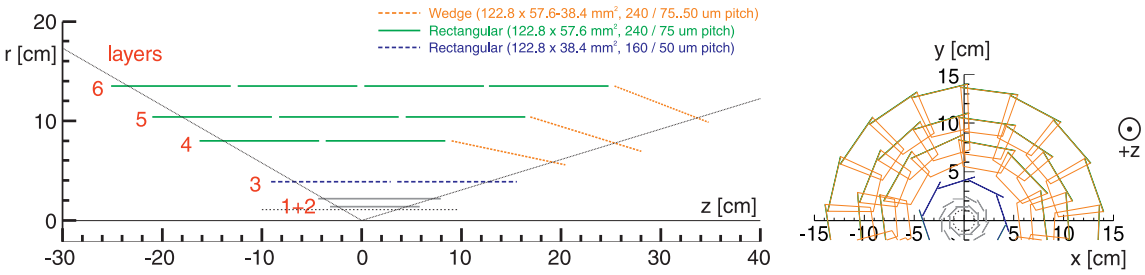


Figure 2.7: View of the SVD in the $r-z$ (left) and $x-y$ (right) planes. The picture, adapted from [17], shows the ladder layout and resumes some characteristics of the used sensors (shape, active area, pitch).

The rectangular ones have the long strips on the p—side, parallel to the detector's z—axis. The short n—side strips along $r-\phi$ face towards the outside of the SVD (Fig. 2.7). The slanted sensors are similar, with a variable strip pitch on the p—side, giving them a wedge-like shape. The strips on the n—side, though, are parallel with a constant strip pitch. They are arranged such that they are perpendicular to the central strip of the p—side. There are two types of rectangular sensors: while layer 3 consists of smaller sensors, layer 4 to 6 are equipped with larger sensors, both having a thickness of 320 µm against the 300 µm of the forward sensors.

	Layer 3	Layer 4	Layer 5	Layer 6
Distance from IP (mm)	39	80	104	135
Number of ladders	7	10	12	16
Number of sensors per ladder	2	3	4	5
Kind of sensors	2 small	wedge + 2 large	wedge + 3 large	wedge + 4 large
Slant angle θ_s (°)	0	11.9	17.2	21.1
Material budget $[X_0]$ (%)	0.6	0.6	0.6	0.6
Overlap (%)	7.1	15.4	3.8	9.0

Table 2.3: Design parameters for the 4 Layers of SVD. The overlap percentage refers to the $r-\phi$ area where neighbouring ladders overlap each other.

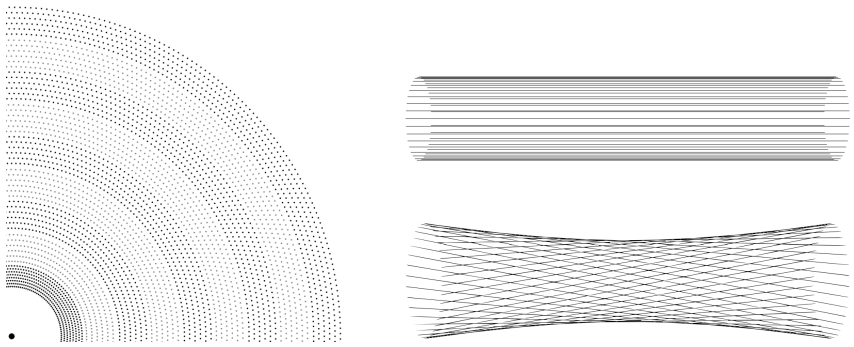
Layer	Rectangular sensor			Wedge sensor		
	strips $p \times n$	pitch (μm) n−side	pitch (μm) p−side	strips $p \times n$	pitch (μm) n−side	pitch (μm) p−side
3	768×768	160	50	-	-	-
4	768×512	240	75	768×512	240	75 to 50
5	768×512	240	75	768×512	240	75 to 50
6	768×512	240	75	768×512	240	75 to 50

Table 2.4: Number of strips and pitches for rectangular and wedge sensors.

The slanted structure for forward ladders of SVD has made the ladders assembly difficult but contributes in suppressing the multiple scattering. In fact, if the slant angle is θ_s, the path length inside the sensor is proportional to the inverse of $\sin(\theta+\theta_s)$, where θ is the angle of the track with respect to the z-axis. Since the θ of the track at the edge of the forward side is 17° and the thickness of the sensor is 300 µm the flight length of the particle inside the FWD sensor of Layer 6 is $\frac{300\,\mu\text{m}}{\sin(17°+21.1°)} \sim 490\,\mu\text{m}$. Without the slant angle, the path length would be 1030 µm, resulting in a multiple scattering increased by a factor 1.7. The SVD design parameters are summarized in Tabs. 2.3-2.4, while a schematic picture of the detector, showing the layers and the ladders layout, is given in Fig. 2.7.

2.3.2 Central Drift Chamber (CDC)

The central tracking device is a large-volume drift gas chamber with small drift cells and filled with He(50%):C_2H_6(50%) gas mixture. The CDC contains 14336 sense wires arranged in 56 layers (Fig. 2.8a), either in axial or stereo orientation (Fig. 2.8b), which allows to reconstruct a full 3D helix track via tracking in the magnetic field. The CDC reconstructs the trajectories of charged particles to precisely determine their momenta and provides particle identification in the low- momentum region using energy loss in the CDC volume. In addition, it generates trigger information for charged particles. The main improvements compared to its predecessor in Belle are the longer lever arm (larger radius, see Tab. 2.5), which yield better momentum resolution over the whole kinematic range of the experiment, fast electronics and smaller drift cells. Even though the larger beam background in Belle II degrades the CDC tracking performance, the new software and the standalone tracking in the PXD and SVD improves the overall reconstructions efficiency so that even better performance with respect to Belle can be obtained.



(a) Quadrant of an $r-\phi$ projection of a portion of the chamber. The innermost superlayer has 8 layers, whereas the others contain six layers each.

(b) Visualization of stereo wires relative to axial wires (the skew is amplified).

Figure 2.8: CDC wires layout in $r-\phi$ and axial orientation.

	Belle	Belle II
Radius of inner cylinder (mm)	77	160
Radius of outer cylinder (mm)	880	1130
Radius of innermost sense wire (mm)	88	168
Radius of outermost sense wire (mm)	863	1111.4
Number of layers	50	56
Number of sense wires	8400	14336
Gas	He–C_2H_6	He–C_2H_6
Diamater of sense wires (μm)	30	30

Table 2.5: Main parameters of the Belle and Belle II CDC [17].

2.3.3 Particle identification: ARICH and TOP

The particle ID has been improved in both barrel and endcaps, in order to build up an efficient system to separate pions, kaons, protons, electrons and muons over the full kinematic range of the experiment. The detectors used for the PID are: the time-of-propagation (TOP) counter, used in the the barrel region, and the ARICH, for the forward end cap region. The first is a special kind of Čerenkov detector in which two-dimensional information about a Čerenkov ring image is given by the time of arrival and impact position of Čerenkov photons at pixelated photo-detectors at one end of a 2.6 m long quartz bar (Fig. 2.9). Each detector module (16 in total) consists of a 45 cm wide and 2 cm thick quartz bar with a small expansion volume (about 10 cm long) at the sensor end of the bar. An important component of the module is also a spherical mirror at the far end of the quartz bar, essential for reducing the chromatic error. The TOP counter requires photo-sensors with a single photon time resolution of about 100 ps, which can be achieved with a 16-channel MCP (Micro-Channel Plate) PMT specially developed for this purpose. For precision timing required in this type of counter, custom-made pipelined waveform sampling read-out electronics is used. For this identification method the starting (particle production) time has to be known with a precision of better than 50 ps; it is provided by the precision radiofrequency (RF) clock of the SuperKEKB accelerator.

In the forward endcap, the proximity focusing Čerenkov ring imaging detector (ARICH) has been designed to separate kaons from pions over most of their momentum spectrum and

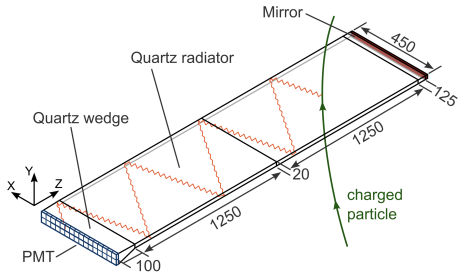
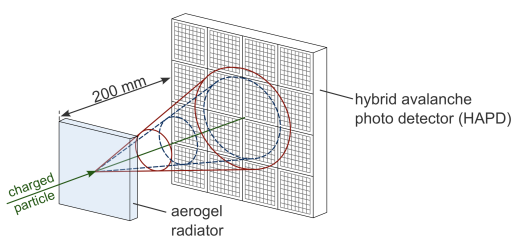
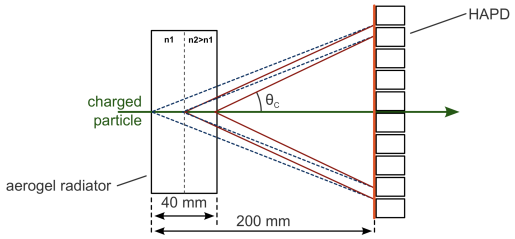


Figure 2.9: Single TOP module, with two blocks fused together (dimensions are reported in mm). Picture taken from [23].



(a) $K^{\pm}\pi^{\pm}$ separation. Blue dashed lines indicate kaons, while red solid lines the pions.



(b) Resolution improvement obtained by splittinf the aerogel radiator of ARICH. The first Čerenkov (blue) starts in the first layer, the second (red) starts at the second layer.

Figure 2.10: ARICH working principle. Pictures adapted from [23].

to provide discrimination between pions, muons and electrons below 1 GeV/c. The ARICH (Fig. 2.10) is composed by an aerogel radiator, where Čerenkov photons are produced by charged particles, an expansion volume to allow the photons to form rings on the photon detector surface, an array of position-sensitive photon detectors, that is capable of detecting single photons in a high magnetic field with high efficiency and with good resolution in two dimensions, and a read-out system for the photon detectors.
A key parameter of the Čerenkov detectors is the number of detected photons ($N_{p.e.}$, see Tab. 2.6): the resolution on the Čerenkov ring σ_{ring}, and thus the separating power between pions and kaons, is inversely proportional to $\sqrt{N_{p.e.}}$. For this reason the two 2 cm thick layers of aerogel have different refractive indices (n = 1.045 upstream, n = 1.055 downstream) in order to improve the light yield without degrading the Čerenkov angle resolution. A more detailed description on PID performance follows in Sec. 3.6.

2.3.4 Electromagnetic Calorimeter (ECL)

Since one third of B-decay products are π^0's or other neutral particles that provide photons in a wide energy range from 20 MeV to 4 GeV, a high resolution electromagnetic calorimeter is a very important part of the Belle II detector. CsI(Tl) was chosen as the scintillation crystal material for the Belle calorimeter due to its high light output, relatively short radiation length, good mechanical properties and moderate price. The main tasks of the calorimeter are:

- detection of photons with high efficiency,
- precise determination of the photon energy and angular coordinates,
- electron identification,

- generation of the proper signal for trigger,
- on-line and off-line luminosity measurement,
- K_L detection together with the KLM.

All the three ECL regions, barrel and forward and backward end-caps, are instrumented with a total of 8736 crystals, covering about the 90% of the solid angle in the center-of-mass system (Fig. 2.11). The CsI(Tl) crystals, preamplifiers and support structures have been reused from Belle, whereas the readout electronics and reconstruction software have been upgraded.

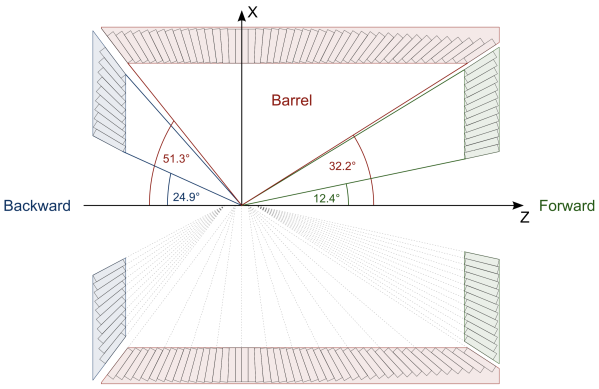


Figure 2.11: Sketch of the ECL, with the forward (green), barrel (red) and backward (blue) regions with their polar acceptance. Picture taken from [23].

In the Belle experiment, the energy resolution observed with the same calorimeter was σ_E / E = 4% at 100 MeV, 1.6 % at 8 GeV, and the angular resolution was 13 mrad (3 mrad) at low (high) energies; π^0 mass resolution was 4.5 MeV/c^2; in absence of backgrounds a very similar performance would also be expected in Belle II. In the presence of considerably elevated background levels as compared to the operation in Belle, the relatively long decay time of scintillations in CsI(Tl) crystals will considerably increase the overlapping of pulses from neighbouring (background) events. To mitigate the resulting large pile-up noise, scintillator photosensors were equipped with waveform-sampling read-out electronics.

2.3.5 K_L and Muon Detector (KLM)

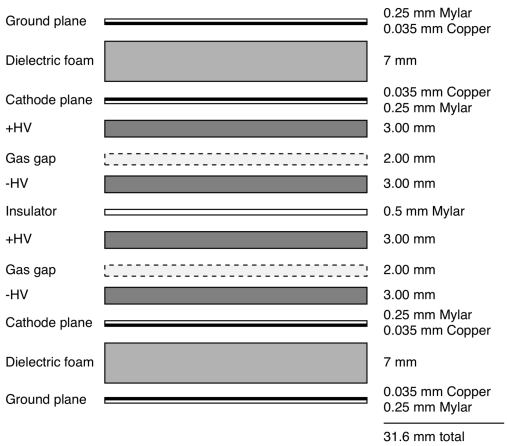


Figure 2.12: Exploded cross-section of an RPC superlayer. Picture taken from [17].

The K_L and muon detector (KLM) consists of an alternating sandwich of 4.7 cm thick iron plates and active detector elements located outside the superconducting solenoid. The iron plates serve as the magnetic flux return for the solenoid. They also provide 3.9 interaction lengths or more of material, in addition to the 0.8 interaction lengths of the calorimeter, in which K_L^0 mesons can shower hadronically. The Belle KLM system, based on glass-electrode resistive plate chambers (RPCs) (Fig. 2.12), demonstrated good performance during the entire data taking period of the Belle experiment. In contrast to Belle, in Belle II larger backgrounds are expected in both endcaps and the innermost layers in the barrel region due to neutrons that are mainly produced in EM shower from background reactions (e.g. radiative Bhabha scattering). To mitigate the long dead time of the RPCs due to the recovery of the electric field after a discharge — which would significantly reduce the detection efficiency if the hit rate is large — in the critical regions the RPCs have been replaced by layers of scintillator strips with wavelength shifting fibers , which are tolerant to higher rates.

Component	Performance
PXD SVD	Impact parameter resolution (PXD and SVD) $\sigma_{z_0} \sim 20\,\mu\text{m}$, $\sigma_{d_0} \sim 10\,\mu\text{m}$ Tracking down to $p_t \simeq 50\,\text{MeV}/c$
CDC	$\sigma_{r\phi} = 100\,\mu\text{m}$, $\sigma_z = 2\,\text{mm}$ $\sigma_{p_t}/p_t = \sqrt{(0.2\% p_t\,[\text{GeV}/c])^2 + (0.3\%/\beta)^2}$ $\sigma_{p_t}/p_t = \sqrt{(0.1\% p_t\,[\text{GeV}/c])^2 + (0.3\%/\beta)^2}$ (with SVD) $\sigma_{\text{dE/dx}} = 5\%$
TOP	$N_{p.e.} \sim 20$, $\sigma_t = 40\,\text{ps}$ K/π separation: 99% at 0.5% of π mis-ID
ARICH	$N_{p.e.} \sim 13$ K/π separation at 4 GeV/c: 96% at 1% of π mis-ID
ECL	$\frac{\sigma_E}{E} = \frac{0.2\%}{E} \oplus \frac{1.6\%}{\sqrt[4]{E}} \oplus 1.2\%$ $\sigma_{\text{pos}} = 0.5\,\text{cm}/\sqrt{E}$ (E in GeV)
KLM	$\Delta\phi = \Delta\theta = 20\,\text{mrad}$ for K_L $\sim$1% hadron fake for muons

Table 2.6: Expected performance for the components of the Belle II detector [17]. More information about sub-detector perfomance measured during the early phase of the experiment, especially on tracking and PID, will be given in the next chapter.

2.3.6 Trigger system

The trigger system has the role to identify and record the events of interest during data taking, unlocking a broad variety of topics not probed in the previous B-factories (e.g. single photon trigger for dark sector searches and 2-or-3 photon triggers for axion-like particle searches) but mantaining an efficiency close to 100 % for most B-decays, for events reconstructed offline. The scheme of Belle II trigger system is composed of two levels: hardware based low level trigger (L1, latency of 5 µs, and maximum trigger output rate of 30 kHz, limited by the read-in rate of DAQ) and software based High Level Trigger (HLT). As a key component of DAQ, the HLT must suppress online event rates to 10 kHz for offline storage, and it must identify track regions of interest for PXD readout to reduce data flux. The HLT reconstructs the event with

offline reconstruction algorithms, allowing access to full granularity event reconstruction using all detectors except for the PXD. Trigger rates for some common processes are reported in Tab. 2.7.

Physics process ($e^+e^- \to$)	Cross section (nb)	Trigger Rate (Hz)
$\Upsilon(4S) \to B\bar{B}$	1.2	960
Hadron prod. from continuum	2.8	2200
$\mu^+\mu^-$	0.8	640
$\tau^+\tau^-$	0.8	640
Bhabha($\theta_{\text{lab}} > 17°$)	44	$350^{(a)}$
$\gamma\gamma$ ($\theta_{\text{lab}} > 17°$)	2.4	$19^{(a)}$

Table 2.7: The total cross sections and trigger rates at the goal luminosity of $8 \times 10^{35}\,\text{cm}^{-2}\text{s}^{-1}$ for several physical processes of interest. Samples of Bhabha and $\gamma\gamma$ events are used to measure the luminosity and to calibrate the detector responses. $^{(a)}$ Since the Bhabha and $\gamma\gamma$ cross sections are very large, these triggers are pre-scaled by a factor of 100; this is straightforward due to their distinct signatures [17]. The maximum average trigger rate for Belle II is 30 kHz.

2.4 SuperKEKB and Belle II operations

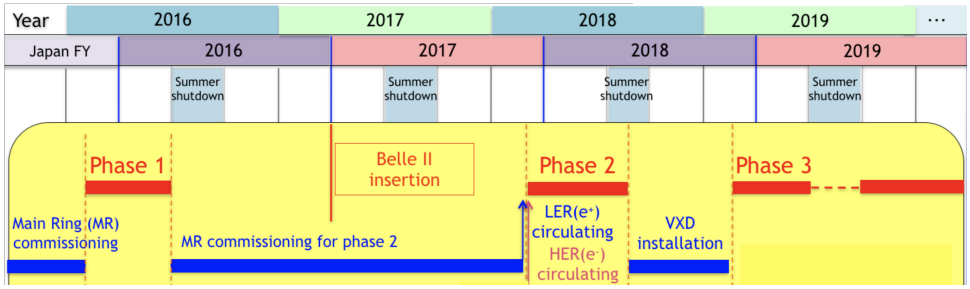


Figure 2.13: Timeline of the Belle II experiment. Adapted from [24]

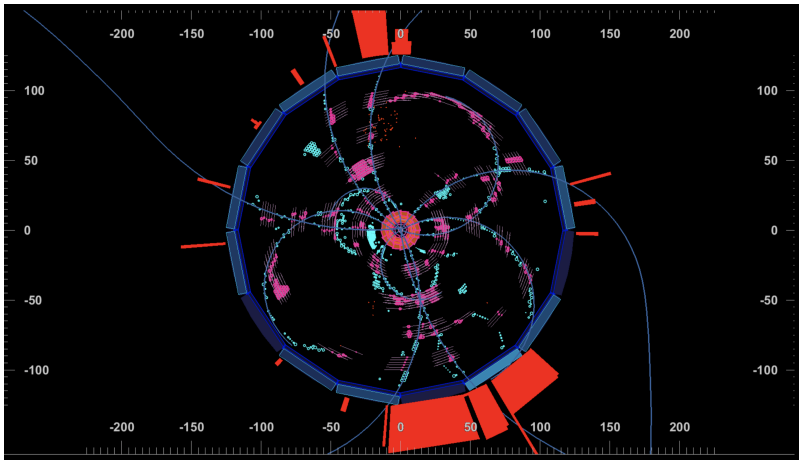
- Phase 1: Successfully completed in 2016. SuperKEKB was commissioned without beam collisions. Single-beam backgrounds were studied with several background monitors in the absence of the Belle II detector.

- Phase 2 : April 26$^{\text{th}}$ 2018 - July 2018. Beam collisions were realized and some physics data ($\sim 0.5\,\text{fb}^{-1}$) were collected without the full vertex detector installed. Only two ladders of the PXD and four ladders of the SVD were installed, one for each layer. This VXD subset was located at $\phi = 0$, i.e. in opposite direction compared to center of the ring. Background monitors were installed at the core of the Belle II detector in order to measure the beam backgrounds. The lack of the VXD elements had a large impact on physics during Phase 2. In fact, track reconstruction was entirely dependent on the CDC: tracks had to be able to reach this detector and this lead to efficiency losses at low p_T.

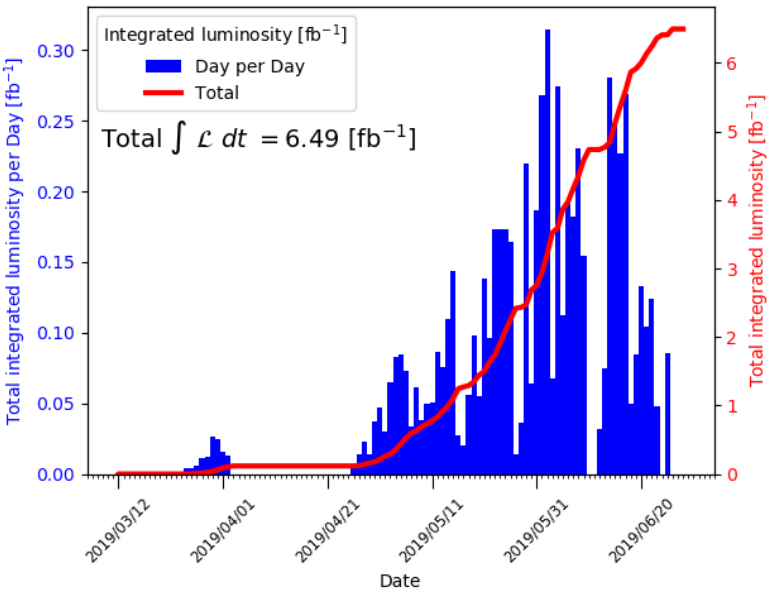
- Phase 3

  1. Spring Run: March 25$^{\text{th}}$ 2019 - July 2019. This phase has seen a quasi-full detector, since PXD is not in its two-layer form: the inner layer has one broken module and

the outer one has just two ladders, in such a way that the whole azimuthal angle is covered. The total integrated luminosity is 6.5 fb^{-1}, collected with different machine paramters, adjusted during intensive tuning studies. Between May 7th and May 14th, continuous injections for LER and HER started. Indeed two different injection modes have been used during the Run: the “normal injection" and the “continuous injection". The continuous injection is necessary to maintain constant luminosity against a short beam lifetime (the lifetime of the main rings will be $\sim$ 600 s at design beam currents), especially for HER (Figs. 2.15a–2.15b). Moreover, each charge injection is small and this allows to avoid turning-off DAQ, which instead is done in the normal injection scheme where almost the entire beam current must be restored.



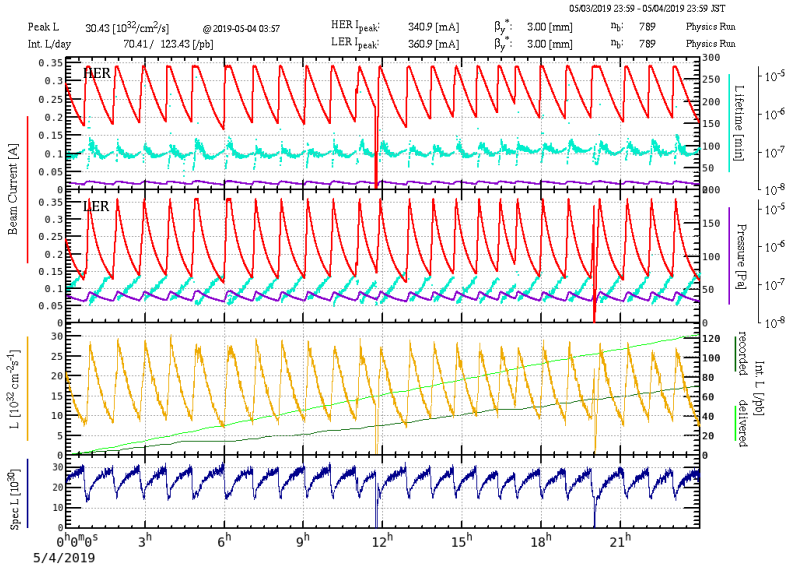
(a) First candidate of hadronic event of the Belle II experiment, recorded on March, 25th.



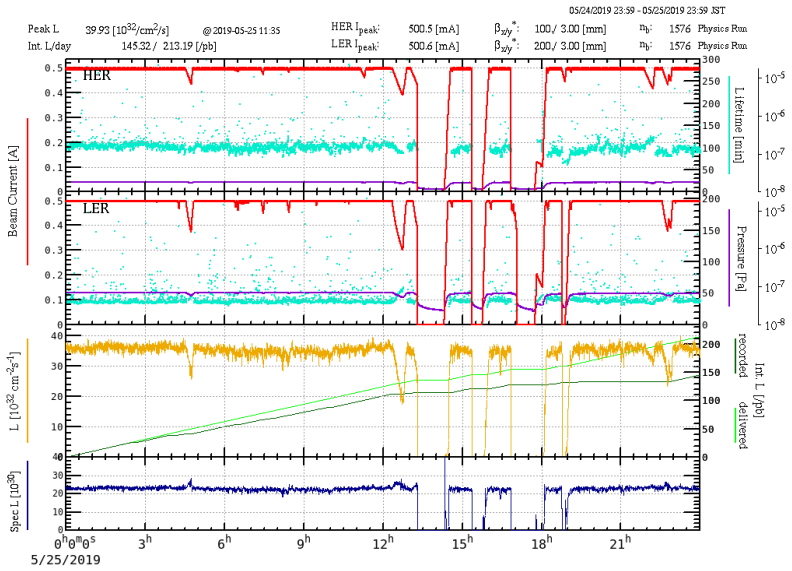
(b) Integrated luminosity per day and the total one.

Figure 2.14: Early Phase 3 - 2019 Spring Run highlights.

The performance achieved during the Spring Run are summarized both in Tab. 2.2



(a) Normal Injection, May 4$^{\text{th}}$. Currents have the typical sawtooth fashion due to finite beam lifetime.



(b) Continuous injection, from 00:00 to around 12:00, May 25$^{\text{th}}$. This technique allows to keep the beam currents and the luminosity almost constant.

Figure 2.15: Snapshots of monitoring plots shown in the SuperKEKB control room. The parameters of the beams are reported: HER and LER currents (red lines) and lifetimes (light blue), pressure (violet), instantaneous luminosity, delivered and recorder luminosity (lighter and darker green, respectively). "Spec L" denotes the specific luminosity, defined as $\mathscr{L}_{\text{spec}} = \frac{\mathscr{L}}{n_b I_b^+ I_b^-}$, where the $I_b^{\pm}$ are the bunch currents and n_b is the number of bunches.

and Tab. 2.8 and can be considered a successful result, considering the numerous issues encountered. The major troubles were:

1.1. April 15$^{\text{th}}$ - A fire broke out in a building adjacent to the KEK Linac. Several sectors of the Linac were badly contaminated with soot and plastic by-products. Components in these sectors had to be cleaned and tested in a major mobilization of the KEK accelerator team. This accident did not cause any human injuries or radiation leakage. On April 22$^{\text{nd}}$, the components of the Linac were

carefully turned on again and recovery began. By the afternoon of April 26th, beams were circulating again in the HER and LER of SuperKEKB and Belle II began taking physics data again.

1.2. Several QCS quenches, mainly caused by power supply failures (for example, that of May 25th) or by steered beams with huge vertical beam size. These events were supposedly caused by beam-dust collision; the most severe episode happened on June 9th, when PXD, SVD and a vertical collimator were affected by a LER beam lost of $\sim$ 150 mA, for an accumulated dose of 3 rad within 40 µs. Fortunately SVD did not encountered severe issue, while PXD suffered some damages and one module became inoperable (see Fig. 3.7, Chapter 3).

2. Autumn Run: October 15th - December 12th. The plan for the upcoming run is to integrate $\sim 14\,\text{fb}^{-1}$ - considering for beam currents a realistic value of 770 mA.

Performance	Value	Notes
Higher $\mathscr{L}_{\text{peak}}$ ($\text{cm}^{-2}\text{s}^{-1}$)	5.495×10^{33} 1.2×10^{34}	During continuous injection mode Last day before shut down
Best Efficiency ($\mathscr{L}_{\text{recorded}}/\mathscr{L}_{\text{delivered}}$)(%)	97	with HER/LER $\sim$ 500/500 mA June 16th
Best daily $\mathscr{L}_{\text{integrated}}$ (pb^{-1})	300	-
Dominant background source	-	LER beam-gas

Table 2.8: Machine performance during Spring Run.

2.5 Belle II software analysis framework

The Belle II Software Framework (`basf2`) is a C++/Python framework to process the events recorded by the Belle II detector [25]. `basf2` has a core written in C++ and features a Python 3 based user interface which offers a simple way to use the software. The work flow, shown in Fig. 2.16 for different tasks such as MC generation, detector simulation and recontruction, can be defined in Python *steering files*. Through these scripts the analysts indicate the desired path of chained modules, each of them designed to perform a specific task on event-by-event basis. These modules, which usually start and with a data input and output, respectively, can access the `basf2` libraries on the C++ level of the software where the actual operations on the data are performed. The exchange of the data-objects between the different modules occurs via the common `DataStore`. Certain libraries such as the MC generation or the detector simulation rely on additional external libraries, e.g. `EvtGen` (B-physics experiments' generator) [26] or Geant4 [27], and CERN's particle physics data analysis framework ROOT [28] for data handling.

The analysis package offers a collection of tools to extract physically relevant information from reconstructed particle tracks, particle identification (PID) likelihoods, calorimeter clusters and informations from KLM which are stored in so-called mini Data Summary Tables — mDST for short. The size of each mDST file is reduced by a *skimming* process to produce different sets of μDSTs files, where each *skim* is designed to meet the preliminary criteria of a specific analysis. At the same time, the skimmed mDSTs should not be produced with tight cuts on any of the associated variables, since this could constrain, and potentially alter, the

course of a specific analysis.

All physics analyses at Belle II require the generation and analysis of Monte Carlo events in addition to the analysis of the real detector data. This involves the creation of physics events by means of a set of decay tables listing all the possible decay processes, their branching ratios and amplitudes. Those events are then passed to a full detector simulation, simulating the passage of the particles through the detector material of the various sub-detectors. Each sub-detector has a digitizer assigned, that uses the input of the full detector simulation in order to perform a realistic detector response simulation. The digitized data is then passed to the reconstruction stage, where the input data for the physics analyses is created. Among the various reconstruction procedures are, for example, the clustering of pixels and strips, track finding and track fitting. The last step in the processing chain is then the physics analysis. The real data processing chain, however, skips the Monte Carlo and simulation steps and feeds the real detector data directly to the reconstruction stage. The Monte Carlo chain and the real data chain share the same reconstruction and analysis procedures.
The geometry handling system of `basf2`, which contains also the magnetic fielsd of Belle II, stores the parameter values needed to fully describe the Belle II detector in a database. Storing parameter values instead of concrete geometry objects allows for a simple and generic way of handling time varying geometry. The full detector simulation in `basf2` is based on Geant4, the standard toolkit for the simulation of particles traversing matter in the high energy physics community.

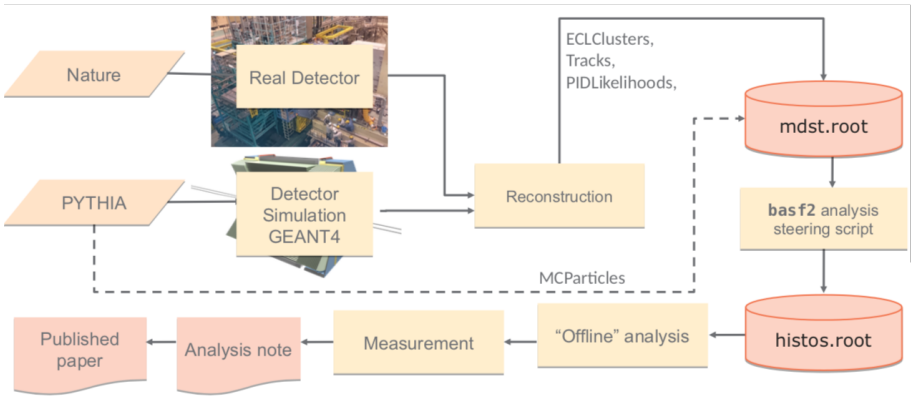


Figure 2.16: Data reconstruction for both MC and experiment sources. Raw data, coming from the real experiment or the simulated one, go through the same reconstruction, whose output, the mDSTs, are the basis for the analysis work. The analysis part can be performed with the Python interface. `PYTHIA` is an auxiliary tool to `EvtGen`, which handles the fragmentation into all the possible final states states not included in the decay file.

2.5.1 Calibrations

For the success of the experiment it is essential to have detailed information on calibrations and varying detector conditions to be available when reconstructing and analyzing the data. "Conditions" indicate a number of settings or calibrations, stored in a central Conditions Database, that can evolve over the time but not on a per-event rate. For example, the noise on SVD strips must be evaluated frequently as it depends on many environmental and operating variables, such as radiation levels and HV settings. Another example of conditions is the alignment of the VXD, which is quantified by a large number of constants that need

to be monitored in order to have knowledge of the systematic deformation of the detector or random displacements of the single pixel and silicon sensors. In this way the mapping between the actual positions and orientations of sensors and the nominal ones can be corrected and the tracking reconstruction performance is optimised. The bulk conditions data is the payload representing the set of sensible parameters consumed in the various workflows of the physics data processing. The payloads contain the information related to a wide range of subsystems, including detectors, triggers, beam monitoring and physics algorithms. For example the information on SVD noise corresponding to a single calibration is stored in a dedicated payload, as well as the alignment constants extracted for a defined dataset. The time information for the validity of the payloads is specified with a parameter called Interval of Validity (IoV). In this context, time is represented by a run number, where a run is a data taking period with stable operating conditions and an experiment number. "Experiment" indicated a period typically lasting for several weeks, during which it is expected that there will be no major changes in Belle II or SuperKEKB operating conditions. The Database interface guarantees the user to get the correct payload for a given run.
A fully qualified set of conditions consists of a set of payloads and their associated IoVs covering the time span required by the workload. A label called Global Tag (GT) identifies the version of the set. When a new global tag is created, it is open for modifications so that assignments of IoVs to payloads can be added or removed. Once a GT is published, it becomes immutable. For the reconstruction the user needs to specify the GT with the most up-to-date calibrations.
The calibrations of Belle II data consists of three major procedures.
The *prompt* calibration provides the first version of the conditions and is used to perform the first processing of raw data into mDST format. In order to enable efficient prompt mDSTs production, the calibrations contained in the prompt GT should be ready on the order of a week following data collection. These constants are not optimal and cannot be used for public results. More precise estimates are needed and thus a second offline stage is performed.
The *offline* calibrations contain the best versions of those produced during prompt calibration. The benefit of the offline procedure is that larger samples may be used and constants can be more effectively applied. Following the offline calibration, the data must be reprocessed with updated calibration constants. Most likely this will occur every 6-12 months but indeed it may happen more often during the early stages of data taking. It should be emphasized that if calibration algorithms are modified, all the data produced up to that point must be reprocessed.
Finally, *online* calibrations are very different from the others since they are used during data taking. For the online reconstruction the HLT retrieves the available payloads, whose IoVs are still open. Once a new calibration is taken, such IoVs can be closed and the online reconstruction switches to the new payloads. The online GT is also the main Global Tag bacause it is where all the conditions used during the `basf2` processing on HLT are stored. Unlike prompt and offline calibrations, where the IoVs indicate the set of runs used to produce a certain calibration, for the online case the validity of a payload ends only when a new calibration is produced and takes over the previous one for the reconstruction.

3 Reconstruction for the D^0 lifetime measurement

This chapter introduces the goal of the thesis work and the methods adopted. The performance of the key reconstruction tools are reported, in particular: tracking and vertexing for the precision on the D^0 flight time measurement and particle identification for the background rejection. A further central analysis tool is the decay fitter, whose main features are briefly described in Sec. 3.4.
A considerable part of the work has been devoted to determine the Interaction Point (Sec. 3.5) and make the information available for the analysis purposes. This requires, from the software side, to correctly put the conditions in the DataStore and ensure that the analysis tools, such as the geometrical fitter, are able to retrieve them.

3.1 Goal and signal channel

Measuring the D^0 lifetime is a powerful tool to evaluate the performance of the Belle II experiment, which has started its physics program after years of development and commissioning. Even though the first collisions of SuperKEKB took place on 2018, only at the end of that year the VXD was finally installed and in January 2019 it joined the rest of the experiment for DAQ tests and Cosmic Runs. The lifetime measurement aims to demonstrate the vertexing capabilities of the Belle II detector, highly related to the tracking and vertex detector, with its two-technology design in the extreme proximity of the interaction point.
There are two experimental ingredients necessary for the lifetime measurement, and more generally to study the time dependent decay rates for the determination of the mixing parameters. The first one is the identification of stable charged particles (PID). Very good PID capabilities are required for separating hadronic final states of D decays such as $\pi^+\pi^-$, $K^\pm\pi^\mp$, K^+K^-, and thus reducing the backgrounds arising from final states which differ from the one under study. The second ingredient is the determination of the decay time of the neutral D meson, based on the measurement of its decay length between the production and decay points. The accuracy on the D meson decay length depends critically on the vertex detectors of the experiment. Furhtermore, an excellent tracking performance (high resolution on the impact parameter) is the basis for a good vertex reconstruction.
The channel studied for the present analysis is:

$$D^{*+} \to (D^0 \to K^-\pi^+)\pi_s^+ + \text{c.c.},$$

whose cross section at $\sqrt{s} = \Upsilon(4S)$ is $\sim 9\,\text{pb}$ (cross sections and branching fractions are reported in Tab. 3.1). The choice of the $K^-\pi^+$ final state of the D^0 (Fig. 3.1) is due to the relatively high branching fraction, in order to increase statistics[1], the low multiplicity and absence of neutrals particles for better precision in vertex position determination. The D^0's

[1] $\sim 9 \times 10^3$ signal events are expected to be produced for each integrated fb^{-1}.

from D^* decays are called D^*- *tagged* D^0's, to be distinguished from those being directly produced in e^+e^- interactions (*prompt* D^0's) or produced in B decays. This latter configuration degrades the D^0 flight time estimation, since the B's have a finite decay length, while the D^*'s decay "immediately" in the interaction region.
The daughter pion of the D^* is usually denoted as "slow pion" π_s, since its average laboratory momentum is low, around 250 MeV/c.
The charge of the π_s is used to determine (or *tag*) the flavour of the neutral D meson; indeed the quark composition of the meson determines the only possible process to be $D^{*+} \to D^0\pi^+$ or the charge conjugate $D^{*-} \to \overline{D}^0\pi^-$.

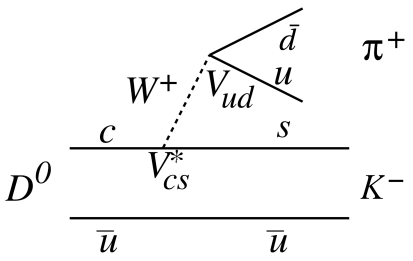


Figure 3.1: Diagram showing at quark level the D^0 decay process, where a W boson is exchanged. The V_{ij} are the CKM matrix elements discussed in Ch. 1.

$\sigma(e^+e^- \to c\bar{c})$	$1.3\,\text{nb}^{(a)}$
$\frac{\Gamma(c\to D^*(2010)^+X)}{\Gamma(c\to X)}$	0.255 ± 0.017
b.r.$(D^{*+} \to D^0\pi^+)$	$(67.7 \pm 0.5)\%$
b.r.$(D^0 \to K^-\pi^+)$	$(3.89 \pm 0.04)\%$

Table 3.1: Branching fractions and cross sections taken from [29] and related to the channel under study. Given the integrated luminosity, the abundance of D^0's can be estimated. $^{(a)}$ The number refers to the cross-section at B-factories, running at $\sqrt{s} = \Upsilon(4S)$.

The difference between the invariant masses of the D^* and the D^0 mesons, ΔM, or the equivalent Q = ΔM - m_{π_s} are powerful selection variables that considerably restrict the amount of combinatorial background. The effectiveness is further enhanced by the excellent experimental resolution in ΔM arising from the cancellation of experimental uncertainty in the determination of the momenta of particles comprising the D^0.

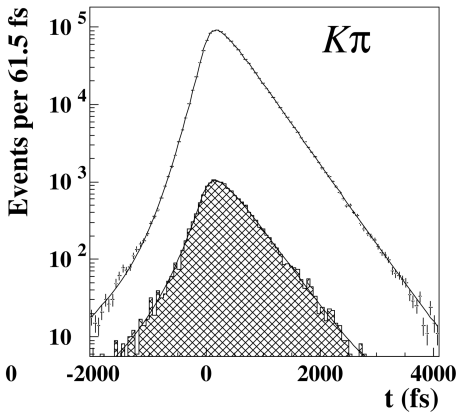
3.2 Method description

The D^0 lifetime τ_{D^0} is extracted with an unbinned maximum likelihood fit to the distribution of proper time of the D^0, t_{D^0}, defined as:

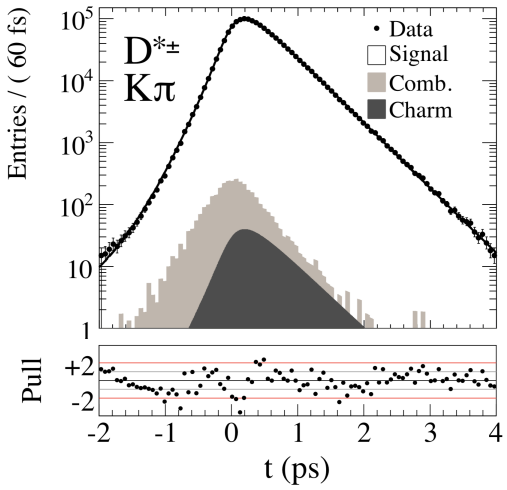
$$t_{\mathrm{D}^0} = m_{\mathrm{D}^0}\frac{l_{\mathrm{D}^0}}{p} = m_{\mathrm{D}^0}\vec{d}\cdot\frac{\vec{p}}{p^2}\,. \tag{3.1}$$

The proper decay time is computed using the vector $\vec{d}$ connecting the D^0 production and decay vertices, the reconstructed momentum $\vec{p}$ and the nominal mass of the D meson m_{D^0}. The decay point ($\vec{V}_{D^0}$, in Fig. 3.3) is obtained from fitting the D meson decay products to a single vertex, whereas the D^0 production point $\vec{V}_{D^*}$ is obtained by intersecting the D^0 flight direction, determined by its momentum vector, and $\vec{V}_{D^0}$, with the e^+e^- interaction region (beam spot, see Sec. 3.5). The slow pion from the D^* decay is fitted to a common space point together with the D^0 momentum with the spatial constraint of the interaction region. This improves the momentum accuracy of the slow pion and, thereby, the experimental resolution in Q.
The average D^0 decay length at the B-factories is about 200 µm while the average D^0 momentum is 3.4 GeV/c. The precision by which the interaction region is known is given in Sec. 3.5. The resulting resolution on the decay length is approximately 45 µm. This suggests that the resolution on vertices' positions and the resolution on the reconstructed momentum are crucial values for lifetime measurements. Between the two, the reconstruction of the vertices is

the most critical as it is strongly tied to the resolution on the impact parameter (defined in Sec. 3.3).
The best estimate of the decay length l_{D^0} is obtained by projecting the distance vector $\vec{d}$ onto the resulting momentum of the D^0's daughters' tracks. The uncertainties on the positions of the two vertices (qualitatively represented by the grey blobs in Fig. 3.3) affect the precision on the measurement of $\vec{d}$ and the projection is meant to mitigate this effect. The proper time distribution obtained at Belle on the same channel studied here is reported in Fig. 3.2a and compared with the *BABAR* result (Fig. 3.2b).



(a) From [30]. Belle result of the simultaneous fit to proper time distributions (here the $D^0 \to K^-\pi^+$ decay is reported). The cross-hatched area represents background contributions, the shape of which was fitted using M sideband events. The obtained lifetime is $\tau = (408.7 \pm 0.6(\mathrm{stat}))$ fs.



(b) From [31, 32]. *BABAR* proper time distribution for the $K\pi$ decay mode with the fit result overlaid. The combinatorial distribution (indicated as 'Comb.' in light gray) is stacked on top of the misreconstructed-charm distribution (indicated as 'Charm' in dark grey). The result on 468 fb^{-1} is $\tau = (408.97 \pm 0.24(\mathrm{stat}))$ fs.

Figure 3.2: Proper time distributions obtained by the Belle **(a)** and *BABAR* **(b)** experiments. Both histograms are related to D^*-tagged candidates decaying into the $K\pi$ final state.

In the measurement of the D^0 decay times, it is assumed that the D^0's are produced at the primary production vertex. For D^0's from B meson decays, this assumptions results in a biased measurement due to the finite decay time of the B mesons and the corresponding D^0 decay time distribution has a more difficult parametrisation. This complication (and the related systematic uncertainties that would result from it) is avoided by requiring the center-of-mass momentum of the tagging D^* mesons to exceed the kinematic limit for $B \to D^*X$ decays. The minimal value required for the D^* momentum[2] is 2.5 GeV/c.
To measure the D^0 lifetime one needs to accurately describe the resolution function, to be convolved with the decay time distribution, assumed to be an exponential. This is true under the approximation of negligible effect of mixing on the decay rate, being the mixing rate small ($\leq 1\%$), as exposed in Sec. 1.2. Furthermore, the mass eigenstates D_a and D_b have very similar lifetimes ($\mathrm{y} \sim 0.5\%$), making possible to define a "lifetime" for the D^0 (unlike the neutral kaons).

[2]The largest momentum of a D^*, in a two body $B \to D^*\pi$ decay in the center-of-mass system, where the B meson is approximately at rest, is about 2.3 GeV/c.

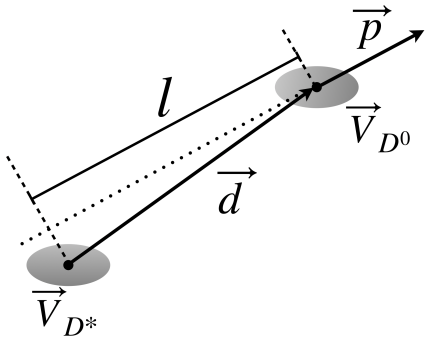


Figure 3.3: Pictorial representation of the decay length measurement, using the production and decay vertices and the reconstructed momentum of the D^0 candidate.

3.3 Impact parameter resolution

The impact parameter is a fundamental quantity for the track parametrisation and the achievable resolution on that parameter is a central performance benchmark for the D^0 lifetime measurement.
A track propagating in vacuum in a constant magnetic field moves along a helix described by five parameters defined at a point $\vec{P}$ of the trajectory. In Belle II this point is identified with the perigee, or point of closest approach (POCA) to the origin in the $x-y$ ($r-\phi$) plane. If the stochastic processes, like energy loss and multiple scattering in detector material, cannot be ignored, the track can still be parameterized locally as a helix. The Belle II tracking software employs the following parametrisation:

- d_0: called *transverse impact parameter*, it is the signed distance between the origin and the POCA in the transverse plane. The sign convention is defined as follows: moving along the track into the direction of the particle's momentum, the sign is positive (negative) if the origin is to the right (left) of the track at the POCA, looking from the positive $z-$axis direction. Ultimately the sign depends on the direction of the angular momentum of the track at the POCA with respect to the magnetic field and it is such that $\vec{p}_t$, $\vec{d}_0$ and $\hat{e}_z$ form a right-handed system (Figs. 3.4a–3.4b).

- z_0: called *longitudinal impact parameter*, it is the signed distance along the z axis from the perigee to the helix (Fig. 3.4c).

- ϕ_0: the angle between the transverse momentum at the POCA and the x axis.

- $\tan\lambda$: the tangent of the angle between the momentum at the POCA and the transverse plane.

- ω: proportional the curvature, whose sign corresponds to the track's charge. $\omega = 1/2R$, where R is the radius of the circle corresponding to the projection of the helix on the $x-y$ plane.

The accuracy that can be achieved for d_0 and z_0 consists of two independent components: the first originates from the intrinsic detector resolution, while the second describes the effect of multiple scattering that a particle experiences while traversing the material in the beam pipe and the sub-detectors. In order not to affect the impact parameter resolution, Belle II has been conceived to minimise material budget and multiple scattering.
The resolution achieved during the early phase of the experiment is measured reconstructing Bhabha events from a sample of $20\,\text{pb}^{-1}$. Assuming that the two e^+e^- tracks come from the

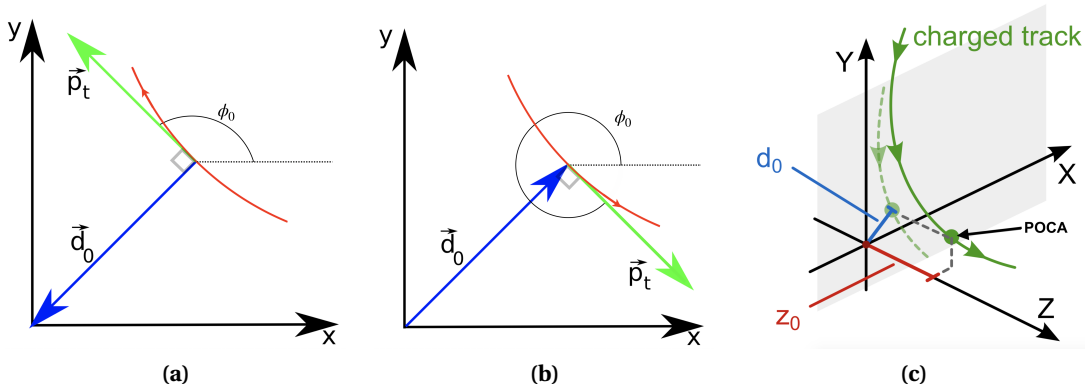


Figure 3.4: Perigee parameters of a particle track (red) in the x-y plane. In each quadrant two configurations can occur so that $\vec{p}_t$, $\vec{d}_0$ and $\hat{e}_z$ form a right-handed system. For example, in the quadrant with $x, y > 0$, the norm of the transverse impact parameter $|\vec{d}_0|$ gets a negative **(a)** or positive sign **(b)** depending on the direction of the transverse momentum of the particle. **(c)**: 3-D wiew of the track with POCA and its projections onto transverse plane and longitudinal direction (taken from [23]).

origin, the width of the difference $\Delta d_0/\sqrt{2} \equiv (d_0(t_-) + d_0(t_+))/\sqrt{2}$ is an estimate of the d_0 resolution; the same holds for the z_0 resolution $\Delta z_0/\sqrt{2} \equiv (z_0(t_-) - z_0(t_+))/\sqrt{2}$. The difference of signs in Δd_0 and Δz_0 is due to the fact that by definition the two tracks t_- and t_+ must have opposite signed d_0's but have the same z_0. The d_0 and z_0 resolution estimates are obtained using the half of the symmetric range around the median containing 68 % of the d_0 and z_0 distributions, respectively. The statistical error is is evaluated via a statistical resampling technique and the systematic uncertainty is estimated by comparing the width obtained with different iterations of the alignment corrections. The values obtained are

$$\sigma(d_0) = (14.2 \pm 0.1\,(\text{stat}) \pm 0.1\,(\text{syst}))\,\mu\text{m}\,,$$
$$\sigma(z_0) = (16.1 \pm 0.1\,(\text{stat}) \pm 0.1\,(\text{syst}))\,\mu\text{m}\,.$$

There is a slight discrepancy with the MC, around 15%, which comes from imperfections in the material description and detector simulation.

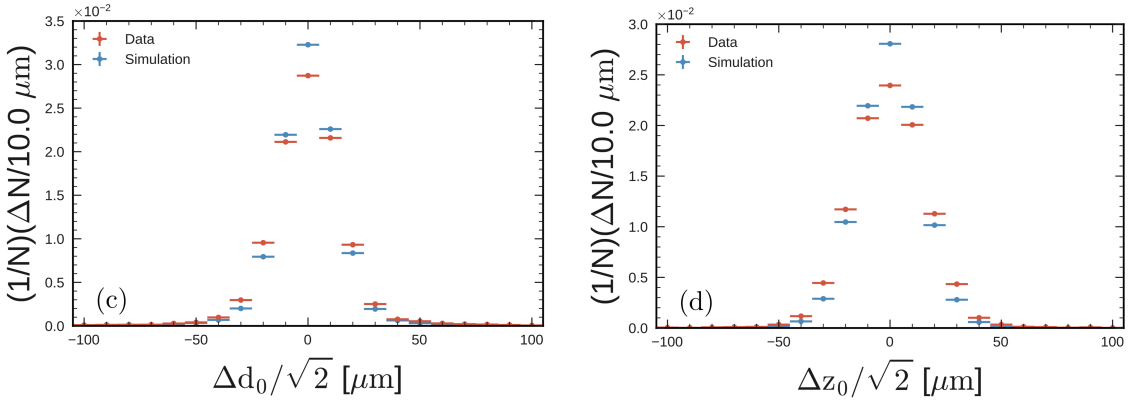


Figure 3.5: d_0 distributions of selected tracks reconstructed via the chain (CDC+SVD+PXD), requiring a certain threshold number of hits for each sub-detector. Other cuts are applied in order to ensure that 2-track events are selected and that the particles experience low multiple scattering. The distributions are used to extract the resolution estimates $\sigma(d_0)$ and $\sigma(z_0)$ as the half of the symmetric range around the median containing 68% of the area.

3.3.1 Tracking performance

The Belle II track reconstruction software consists of multiple independent algorithms to process the measurements of each tracking sub-detector and integrate all of them into one final set of tracks available for physics analysis. This allows to employ algorithms which properties are especially suited for the different sub-detectors.
In order to perform the combination of sub-detector tracks without increasing the fake rate of the output tracks, the Combinatorial Kalman Filter (CKF) is used to link hits and tracks across the sub-detector boundaries.
Track reconstruction progresses through the following steps:

- Hit clustering: the hits within each of the tracking detectors are selected and merged to remove noise hits and obtain a more precise estimate of the point of intersection between the track and the detector, called *space-point*.

- Pattern recognition: dedicated algorithms identify the radial sequences of clusters more likely to correspond to real trajectories, taking into account the constraints due to the production topology (IP position) and the detector geometry.

- Track fitting: the final estimation of track parameters is achieved with a fit of the cluster positions after removing the clusters that do not belong to the identified patterns. For the fit, a specific particle hypothesis must be assumed to calculate correctly the energy loss and the material effects. In Belle II, all reconstructed tracks are fitted with the π, K and p hypothesis. The fitted tracks are extrapolated to the point of closest approach (POCA) to the interaction point (IP) afterwards to extract their helix parameters.

As a first step, the measured signal in the CDC is filtered and reconstructed by two independent algorithms, whose results are merged in the CDC-only tracks, which are then fitted. A CKF is used to enrich the CDC tracks with SVD hits. Moreover, high-curvature tracks that do not produce enough hits in the CDC (low momentum particles with $p_t < 100\,\mathrm{MeV}/c$) are reconstructed with a standalone SVD track finder . The results are combined, fitted again and extrapolated to the PXD with a second CKF. After a final fit, the results can be used in a physics analysis.
The tracking efficiency is evaluated with a MC simulation by comparing the number of reconstructed tracks to the generated ones. A good track finding algorithm should collect in a track all the hits of one and only one MC particle. Two figures of merit are then defined for each pair of MC-track and track found by the pattern recognition (PR-track):

1. The hit efficiency quantifies how efficient is the pattern recognition in identifying all the hits belonging to a single particle. It is defined as the fraction of hits of a given MC-track contained in a given PR-track.

2. The hit purity quantifies how precise is the pattern recognition in identifying the hits belonging to only one particle. It is defined as the fraction of hits of a given PR-track contained in a given MC-track.

A PR-track is defined as matched to a given MC-track if the hit purity exceeds 66% and the hit efficiency exceeds 5%. The low hit efficiency requirement accounts for low momentum tracks curling in the tracking volumes which may leave several hundreds of hits.
If there are two or more PR-tracks that are matched to the same MC-track, the PR-track having the highest efficiency and purity is defined as correctly identified and the remaining ones are

defined as clones. If the PR-track fails the purity requirement, e.g. the PR-track is made up of hits from two MC-tracks, each one with a hit purity below 66 % or the PR-track is made of background hits, it is defined as *fake*.
According to the strategy employed to find a track, SVD provides necessary information to extrapolate the tracks reconstructed in the CDC to the PXD with high efficiency. The efficiency of SVD has been monitored during the first Run of Phase 3, showing excellent performance (efficiencies $\gtrsim 99\%$ for all the layers, as shown in Fig. 3.6 and Tab. 3.2). The efficiency is measured as the fraction of times a cluster is found within ± 0.5 mm from the extrapolated position of the track on the sensor, for both sides $p-n$. The lower efficiency of Layer 3 is due to the fact that one readout chip was masked. This high efficiency is also confirmed by a study conducted for the candidate selection, which is presented in Chapter 4. Having established the high performance of the SVD, we can now focus on the PXD efficiency.

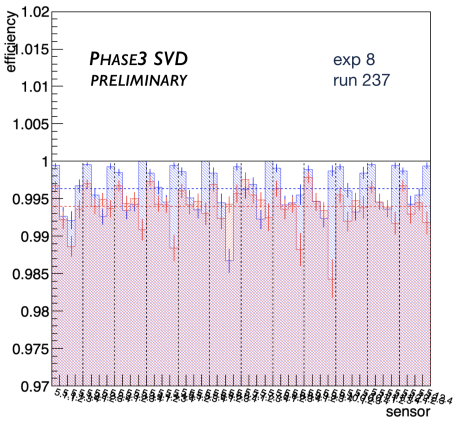


Figure 3.6: The efficiency per each sensor of Layer 5 is reported.

Efficiency	P side	N side
Layer 3	$(99.72 \pm 0.04)\%$	$(98.47 \pm 0.06)\%$
Layer 4	$(99.66 \pm 0.04)\%$	$(99.34 \pm 0.09)\%$
Layer 5	$(99.61 \pm 0.08)\%$	$(99.4 \pm 0.1)\%$
Layer 6	$(99.2 \pm 0.2)\%$	$(99.3 \pm 0.2)\%$

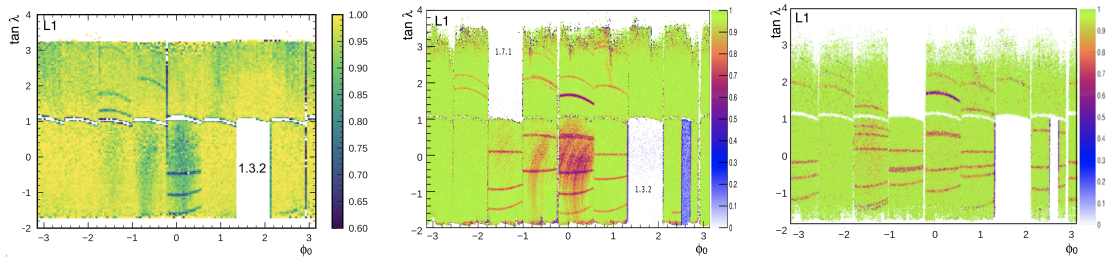
Table 3.2: The preliminary average hit efficiency of each sensor measured on run 237 of Experiment 8 with both beams in continuous injection mode. The efficiencies for each layer, on both p and n sides are summarised.

As introduced in Sec. 2.4, the PXD experienced critical moments due to the beam instabilities. Moreover, given that Belle II is in its initial phases, PXD had to deal with unpredictable backgrounds.
The efficiency of attaching the PXD hits during the pattern recognition for each module is reported in Figs. 3.7a–3.7c. The selected snapshots show the PXD conditions before and after the main accident of May 28$^{\text{th}}$ and how it recovered.

3.4 Vertex fitter

A fundamental tool for the measurement is the vertex fitter, to reconstruct the signal chain with the two vertices V_{D^*} and V_{D^0}. Among the various fitters provided in the `basf2` environment, we choose `TreeFitter` [33] because it is conceived to perform simultaneous vertex fitting of a decay tree [34]. `TreeFitter` attempts to perform the fit accordingly to the decay structure defined by the user when reconstructing the particles (*hypothesis-based fitting*). Three options can be switched on during the vertexing process:

- Mass Constraint: Require the mass of selected composite particles to match the PDG mass. Energy-momentum free parameters of the fit drop from 4 to 3.
- IP Constraint: The production vertex of the head of the decay is constrained to the beam spot region, i.e. the region of space where the e^+e^- collision is more likely to happen.



(a) May 21$^{\text{th}}$. The empty area corresponds to the bad module 1.3.2, whose acceptance is covered by the 2.4.2 module of Layer2.

(b) After May 28$^{\text{th}}$. Efficiency has further worsen after the accident, leaving a module (1.7.1) over until June 7$^{\text{th}}$.

(c) After June 9$^{\text{th}}$. Module 1.7.1 and few gates recovered.

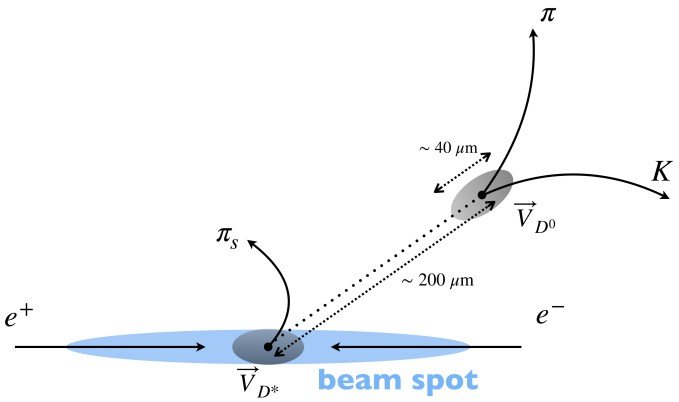
Figure 3.7: PXD efficiency map for the Layer 1. On y-axis the $\tan\lambda = \frac{p_z}{p_t}$ ($\lambda = \frac{\pi}{2} - \theta$) is shown. The gaps between the forward and backward modules are visible (white regions at $\tan\lambda \approx 1$), as well as lower efficiency regions due to dead gates (darker stripes). The damages due to late-May accident (Sec. 2.4) and the subsequent recovery are visible.

This constraint is treated as the IP was a particle mother of the tree specified by the analyst. The IP position and beam spot size are specified in the dedicated payloads (calibration are introduced in Sec. 2.5), which are created during data taking. The algorithm used to store the information about beam spot is described in the next section. In vertex reconstruction the average beam spot can be used as a constraint on the production vertex of the D meson. The values ofthe central position and size of the beam spot are used in the χ^2 minimisation for the fit.

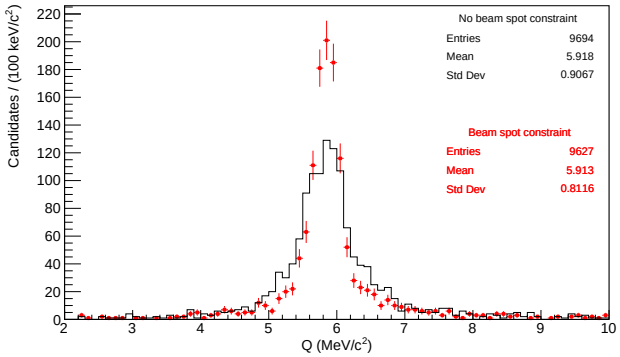
- Update daughters momenta: If this option is not set, only the 4-momenta for the head of the tree are updated after the fit. Otherwise the entire tree is updated with the fitted values of momenta.

Among the other fitters provided in the `basf2` environment, `TreeFitter` is preferred in complex topologies and modes with neutrals. Furthermore, it is fully interfaced with the analyst work; for example it allows to directly extract the lifetimes and the decay lengths of a candidate particle. For the D^0 lifetime measurement, we apply the mass constraint on the D^0 mass of the candidate, the beam spot constraint, as it allows to have a run-dependent information on position and size of interaction region, and we also apply a cut on the confidence level to accept fitted decay trees. This means that candidates with a confidence level lower than the chosen threshold are removed from the particle list. We set the conservative cut `pValue`>0, which removes only the fits that fail. The `pValue` variable indicates, in the `basf2` environment, a context-dependent χ^2 probability associated to the fit related to a certain particle. If the particle is track-based (final-state charged one), then this is the p-value of the track fit. If the particle is composite and a vertex fit has been performed, as in the case of D^0 and D^* , then it indicates the χ^2 probability of the vertex fit result.

Figure 3.8b shows the energy released $Q \equiv M(D^0\pi^+) - M(D^0) - M^{\text{PDG}}_{\pi^+_s}$ in $e^+e^- \to D^{*+}X$ continuum events where $D^{*+} \to D^0\pi^+$ decays with $D^0 \to K^-\pi^+$ have been selected and the reconstruction is performed both with and without the constraint that the D^{*+} originates from the beam spot. Due to its low momentum the direction of the slow pion is very sensitive to multiple scattering and requiring it to originate from the interaction region substantially improves the resolution on Q.



(a) Pictorial image and not to scale of the process under study, with the orders of magnitude of the lengths to be measured. Adapted from [16].



(b) Q distribution in $e^+e^- \to D^{*+}X$ continuum events where we have selected $D^{*+} \to D^0\pi^+$ decays with $D^0 \to K^-\pi^+$ with and without the constraint that the D^{*+} originates from the beam spot. Distributions are obtained from a 0.34 fb^{-1} data subsample.

Figure 3.8: Use of `TreeFitter` to perform the reconstruction on the chain $D^{*+} \to (D^0 \to K^-\pi^+)\pi^+$.

3.5 Beam spot calibrations

The beam-beam collisions occur in a tiny region in the center of the detector, called interaction region or *beam spot*, whose size is determined by the beam optics. As discussed above, run-dependent knowledge of the spatial distribution of beam spot can be used as a powerful constraint when fitting decay chains and this technique is indispensable for time-dependent measurements.

The position and shape of the interaction region vary with time and needs to be carefully calibrated and monitored. The calibration is based on the spatial distribution of reconstructed Interaction Points (IPs). In the production of a $B\overline{B}$ pair at the $\Upsilon(4S)$ resonance there are no particles originating from the primary collision point other than the B mesons themselves. Consequently, the primary vertex cannot be directly reconstructed in these decays and the beam spot calibration instead relies on continuum events. Bhabha and di-muon events have the advantage that there are only two tracks in the event, that have both relatively high momentum and are guaranteed to originate from the IP. Hadronic events have more tracks and consequently a smaller statistical per-event uncertainty on the vertex position, but they are polluted by a $b\bar{b}$ contribution, whereas Bhabha and di-muon events benefit from the higher

cross section and better resolution on the vertex determination.
During the 2019 Spring Run, the calibrations of the beam spot region were tightly related to the determination of the IP position for the Data Quality Monitoring (DQM). I have personally developed the system to provide in real time several histograms showing the main performance-related variables for all the sub-detectors. This is a vital tool to determine, both *online* and *offline*, the quality of the stored data and to promptly diagnose issues on hardware and DAQ sides. The *online* monitoring of the IP position has proved to be a straightforward feedback during data taking and machine tuning. The IP position is determined from the reconstruction of di-muon events, where both the candidate tracks must have

1. $|d_0| < 0.5\,\text{cm}$ and $|z_0| < 2\,\text{cm}$
2. $p > 1\,\text{GeV}/c$

and the candidate di-muon event must have an invariant mass $M \in [9.5, 11.5]\,\text{GeV}/c^2$ in order to avoid radiative processes. The vertex fit is performed using the `kFit` package, which is the fastest among the fitters provided in the Belle II framework. The necessary quantities for the beam spot calibration are computed from the DQM histograms and stored for each run. The algorithm implemented is described below:

I. The **IP vector** $\vec{x}_{IP}$, is a vector that we define as:

$$\vec{x}_{\text{IP}} = (\text{median}(x), \text{median}(y), \text{median}(z)) \ .$$

The coordinates are the medians of the histograms of the fitted vertices. An example of reconstructed primary vertices in di-muon events in a typical Belle II run is shown in Figure 3.9. In the y direction the width of the distribution is dominated by the vertex resolution. In the z direction it is dominated by the beam spot size, while in the x direction it is a combination of both.
We choose the median as a good robust alternative to the sample mean because the distributions cannot be assumed as simple Gaussians due to the increased-tail contaminations. This is also related to the fact that the machine parameters can slightly change during a single run, especially for those that last several hours. On the other side, for short runs the effect of outliers can significantly affect the estimation of the central tendency of the population. Indeed the sample median fits the bulk of the data and is not much influenced by the outlier, unlike the mean.

II. The **variance matrix** of the IP vector, whose elements are defined as:

$$V_{ii} = \frac{\sigma^2_{\text{i,trim.}}}{N}, \qquad V_{ij} = V_{ji} = 0\,,$$

where $\sigma_{\text{i,trim.}}$ is the standard deviation of the i-th coordinate histogram evaluated in a $6-\sigma$ range around the median and N its the number of candidates. The cut is made to discard possible outliers and the use of the variance of the mean is a good approximation due to the dependence to the inverse of the number of events.

III. The **beam spot size matrix**, estimated as:

$$\Sigma_{ii} = \sigma^2_{\text{i,trim.}} - \mu_i, \qquad \Sigma_{ij} = \Sigma_{ji} = 0\,,$$

being μ_i the mean of the histogram of the variances returned by the fitter (an example is reported in Fig. 3.10) for the i-th coordinate.

The event selection for the IP position determination can be refined in order to achieve a better resolution on the vertices. For example, requirements on the number of hits produced by the two tracks in the vertex detector can be applied, or the kinematic range of the muon candidates can be restricted. Furthermore, the topology of the di-muon event can affect the quality of the reconstructed vertex and thus the error associated to it. The mild selection was aimed at not losing efficiency during the data taking for the purpose of having an immediate online information.
The computation of the parameters for the calibration should also be revised; in particular, robust estimators for all the beam spot size matrix elements must be established. The choice to set to zero the non-diagonal terms during this first run period is a good assumption for $x-y$ and and $y-z$ elements but not for the $x-z$ due to the crossing-angle of the beams. Furthermore, the x and y size are both overestimated and very similar to each other. This suggests that the measurements are dominated by the resolution effects, especially for the size on the y direction, which is actually around $1.5\,\mu$m, and that subtracting the average error is not enough to remove the effect. Although the machine parameters are still far from the design values, the vertical beam spot is already too small to be measured from vertexing and could be in principle taken as a fixed parameter from SuperKEKB side.

Within the scope of the D^0 lifetime measurement, the IP position calibration is the most critical part since an erroneously positioned IP could introduce a bias in the extracted D^0 decay length and thus on the flight time. On the other side, the overestimation on the beam spot size is expected to degrade the resolution on proper time but not affect the central value of the measurement.

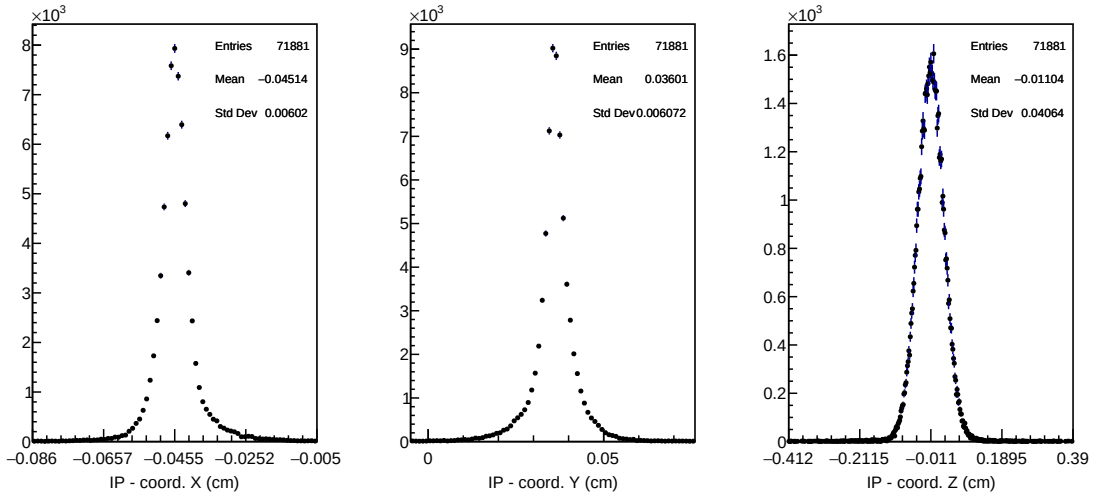


Figure 3.9: Reconstructed Interaction Points for di-muon events for the x, y and z coordinates of the IP. The distributions refer to a single run and are manipulated to obtain the information for the calibrations.

The average positions and the beam spot sizes (square roots of the Σ_{ii} elements) as functions of run number is shown in Figure 3.11 for the period of time called Experiment 8.

In order to take into account the uncertainty on the center of the beam spot, as well as its size, the V_{ij} and Σ_{ij} are summed for the constraint to be applied by `TreeFitter` once the V_{D^*} vertex is created during the fit.

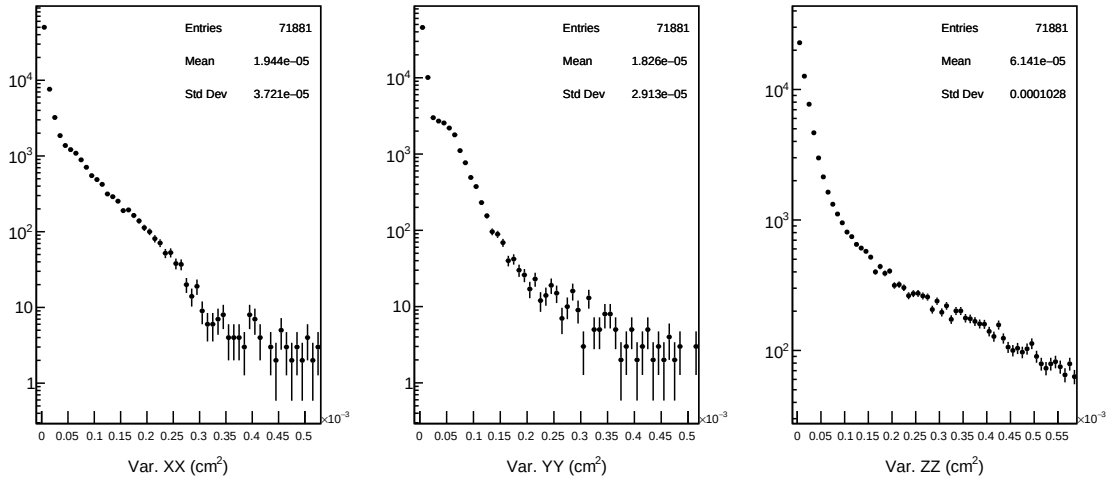


Figure 3.10: Distribution of diagonal elements of the variance matrix returned by the fitter for the same run of Fig 3.9. Their means are used to compute the beam spot size matrix elements.

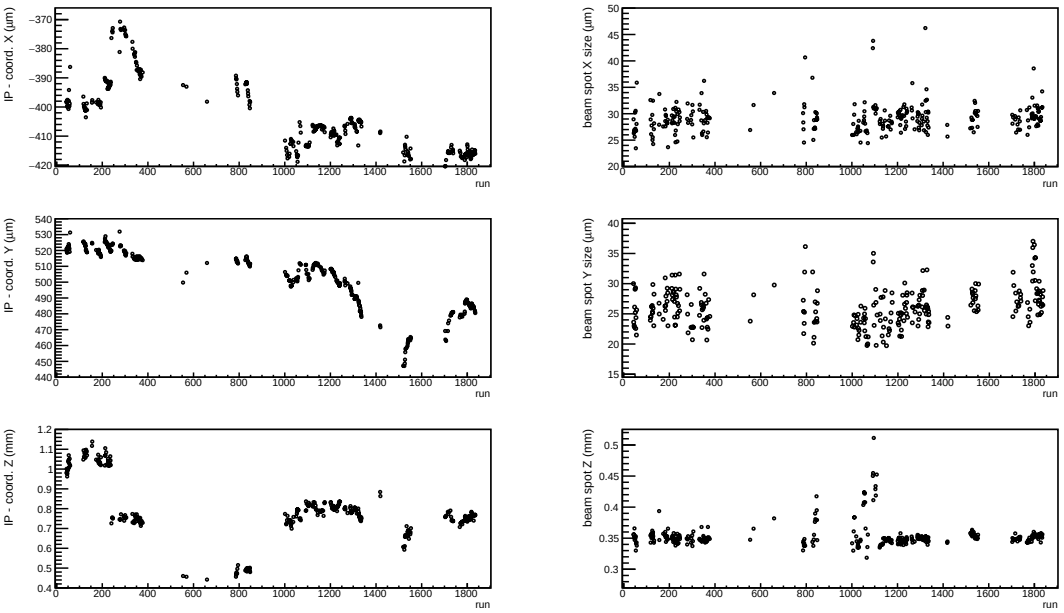


Figure 3.11: On the left, the coordinates of the Interaction Points reconstructed for each run. On the right, the respective estimated sizes. The trend over time of IP coordinates shows the tuning work during SuperKEKB operation. The beam spot sizes are approximately the same in x and y, due to resolution effects.

3.6 Particle identification performance

Effective and efficient particle identification (PID) is vital to isolate an hadronic final state and suppress the background in the D^0 lifetime measurement. As described in Sec. 2.3.3 the Belle II detector contains an upgraded PID system, including a Time-Of-Propagation (TOP) counter in the barrel region of the detector and a proximity-focusing Aerogel Ring-Imaging

Cherenkov (ARICH) detector in the forward endcap region, to provide information on charged particles over the full kinematic range. The information from these detector systems is combined with that from specific ionization (dE/dx) measurements from the SVD and CDC in order to act as the primary sources of information for charged hadron PID. The particle ID in the ECL is mainly related to charged particles which do not reach the KLM (particles with low transverse momentum $p_t < 0.5\,\mathrm{GeV}/c$) and neutral particles (photons and K_L^0's).
Charged particle identification at Belle II relies on likelihood based selectors. Information from each PID system is analyzed independently to determine a likelihood for each charged particle hypothesis. Then, the *binary* likelihood ratios are used to construct a PID value $\mathscr{L}(\alpha:\beta)$ computed as:

$$\mathscr{L}(\alpha:\beta) = R_{\alpha,\beta} = \frac{1}{1+e^{\ln\mathscr{L}_\beta - \ln\mathscr{L}_\alpha}} = \frac{\mathscr{L}_\alpha}{\mathscr{L}_\alpha + \mathscr{L}_\beta}$$

where $\alpha,\beta \in P = \{e,\mu,\pi,K,p,d\}$ and $\mathscr{L}_\alpha$ is the product over the detectors providing PID information for the particle type of interest: $\prod_{\mathrm{det}} \mathscr{L}_\alpha^{\mathrm{det}}$. Using the log-likelihoods, this corresponds to:

$$\log\mathscr{L}_\alpha = \log\mathscr{L}_\alpha^{\mathrm{SVD}} + \log\mathscr{L}_\alpha^{\mathrm{CDC}} + \log\mathscr{L}_\alpha^{\mathrm{TOP}} + \log\mathscr{L}_\alpha^{\mathrm{ARICH}} + \log\mathscr{L}_\alpha^{\mathrm{ECL}} + \log\mathscr{L}_\alpha^{\mathrm{KLM}}$$

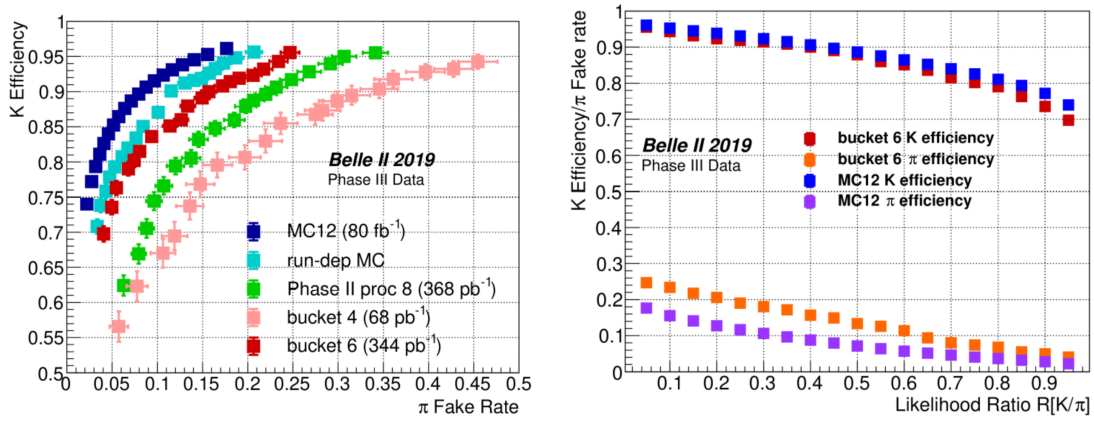
for the generic species α. The binary ratio is greater than 0.5 for a charged track that resembles the most to the particle of type α than one of type β.
Finally, the different hypotheses can also be compared through a ratio of all the likelihoods, usually called "Global Probability" for a fixed species, p, and defined as:

$$\mathrm{Prob(p)} = \frac{\mathscr{L}_p}{\sum_{\alpha\in P}\mathscr{L}_\alpha}\,.$$

The use of the ratio reflects the Bayesian nature of the problem of assigning a mass hypothesis: the likelihood value is actually proportional to the conditional probability associated to a certain species.
The performance of the PID algorithm are evaluated by three parameters:

- Efficiency, or the ability to correctly assign the ID:
  $\varepsilon(K) = N(\mathrm{K\ identified\ as\ K})/N(\mathrm{real\ K})$, equal, by definition, to the "probability of a kaon to identified as a kaon";
- Mis-ID probability, the ability not to assign the incorrect ID:
  $\mathrm{Mis\text{-}ID}(K) = N(\mathrm{non\text{-}K\ identified\ as\ K})/N(\mathrm{non\ K})$, which is the "probability of a non-kaon to be called kaon";
- Fake-rate, fraction of particle with the wrong ID:
  $F(K) = N(\mathrm{non\text{-}K\ identified\ as\ K})/N(\mathrm{identified\ as\ K})$, equal to the "fraction of non-kaons in the kaon collection".

PID performance related to the Spring Run data is summarised in Fig. 3.12 and compared to MC. The kaon efficiency versus the pion fake rate for data has good agreement with MC but still does not match run-dependent MC expectations. This is not totally unexpected as we are in an early phase of the experiment and more time is needed for the simulation to be completely reliable.
The analysis presented in this manuscript was carried out during the data taking phase of the experiment while the studies on PID performance – both for hadrons and leptons – were still ongoing, and PID algorithms still needed tuning. The initial ignorance of PID performance



(a) **(b)**

Figure 3.12: PID performance for all sub-detectors evaluated from two 2019 data subsamples (called `bucket4` and `bucket6`, see Tab. 4.1, Cap. 4) and comparison with MC expectations. In **(a)** Kaon efficiency vs. π-fake rate. For `bucket6` a significant improvement is obtained, with better agreement with MC, with respect to Phase II. PID performance is even closer to random triggered run dependent MC samples. In **(b)** the Kaon efficiency and the pion fake rate vs. $R_{K,\pi}$ are shown. Kaon efficiency has good agreement between data and MC, and disagreement is mostly in the pion fake rates.

and then the measured discrepancies observed between data and MC led to the choice of refining on data the selection for the measurement. The whole analysis process should be done again once a more profound knowledge of the various sub-detectors is available and increased agreement with MC is obtained.

4 Candidate selection and background characterization

This chapter describes the candidate selection procedure for the signal channel, presented in Sec. 3.1. The dataset used for the selection refinement is a subset of the final one. The strategy requires studying the impact of some critical variables on the efficiency and also the purity of the signal sample, where the signal candidates are identified from the peak in the D^0 reconstructed invariant mass distribution. The selections applied on data and on MC are listed throughout the chapter, as well as the models used for the fits to the m_{D^0} and Q distributions, which are also reported in Appendix B. These fits allow to complete the set of cuts for the selection, defining a signal region in the (m_{D^0}, Q) plane. The background composition is studied on MC in view of the fit to the proper time distribution. Comparison between data and MC is performed and some notable distributions are shown in Sec. 4.7.1, while additional fits to MC are collected in Appendix C.

4.1 Datasets

The analysis is performed on the data collected from the start of Spring Run on March 25$^{\text{th}}$ 2019 up to June 3$^{\text{rd}}$. Our work has progressed in parallel with the experiment and various studies have been repeated once a new subset of data was available. The datasets and the nomenclature used within the collaboration are shown in Tab. 4.1. The results presented are obtained with the datasets `bucket6`, used to shape and refine the selection and test the fitter codes, and `proc9`, used to deepen the understanding of the detector, as more features were visible at that point, having accumulated a ~ 7.5 larger statistics. The final measurement is obtained with `proc9` data. We use also the official generic and signal MC samples to study the selection, the reconstruction performance and the proper time fit. Generic MC indicates simulated $B\overline{B}$ decays (both B^+B^- and $B^0\overline{B}^0$), continuum processes ($e^+e^- \to q\bar{q},\ q = u,d,s,c$) and $e^+e^- \to \tau^+\tau^-(\gamma)$ events. The sample consists of 80 fb^{-1} equivalent data for each component and the combination of all the generic samples is called `cocktail`. The dedicated signal MC sample of 10^6 $D^* \to (D^0 \to K\pi)\pi$ events corresponds to a luminosity of ~ 115 fb^{-1}. For the data-MC comparison the `cocktail` is used, as it contains both signal events (coming from $c\bar{c}$ processes) and generic background events.

4.2 D^* skim selection

After data processing, events taken by Belle II are classified into several categories through the skimming process. Some of them, such as Bhabha, muon or gamma - pair events or *hadronic* events are used either for detector calibration or physics analyses. The skim for hadronic events, basically identified using track multiplicity and a set of variables that reject

	Dataset name	Period	Int. luminosity (fb^{-1})
DATA	`bucket4`	March 25th - April 4th	0.068
	`bucket6`	April 26th - May 5th	0.344
	`proc9`	March 25th - June 3rd	2.6
MC	Signal	-	115
	`cocktail`	-	80

Table 4.1: Datasets used and the equivalent sizes for data and MC. Data subsamples are all collected at $\Upsilon(4S)$ resonance. The time gap between `bucket4` and `bucket6` is due to the fire accident occurred in the area next to the SuperKEKB Linac.

Bhabha events, is mainly used for analyses of B and charm mesons. Further skims that contain smaller categories of physics events are made from these basic skims and provided to individual analyses. For example the D^* calibration *skim*, produced from the hadronic skim, looks for $D^* \to (D^0 \to f)\pi_s$ processes, where f indicates the main hadronic final states of the D^0, including the $K\pi$ one. This latter is the channel we are studying to measure the D^0 lifetime and for this reason the D^* calibration *skim* is the starting point of our analysis. The cuts applied during the skimming process are listed below and grouped for the three main classes of particle candidates: the final-state charged pions and kaons, the D^0's and the D^*'s.

1. $\boldsymbol{\pi_s^{\pm}, K^{\mp}, \pi^{\pm}}$ **candidates**
   - 1.1. the absolute value of the transverse impact parameter (defined in 3.3): $|d_0| < 0.5\,\text{cm}$
   - 1.2. the absolute value of the longitudinal impact parameter: $|z_0| < 3\,\text{cm}$
2. $\boldsymbol{D^0}$ **candidates**
   - 2.1. the invariant mass of the ($K^{\mp}$, $\pi^{\pm}$) pair: $1.7 < m_{K\pi}(\text{GeV}/c^2) < 2.1$
3. $\boldsymbol{D^*}$ **candidates**
   - 3.1. the center-of-mass frame momentum: $p_{\text{CM}}(D^*) > 2.5\,\text{GeV}/c$
   - 3.2. $m(D^0\,\pi_s) - m(K\pi) = \Delta M < 0.16\,\text{GeV}/c^2$.

The cuts 1.1 and 1.2 are quite standard for the analysis in Belle II and are intended to keep only the tracks coming from the IP, while $m_{K\pi}$ is selected in an interval of reasonable width around the the true value of D^0 mass — (1864.84 ± 0.05) MeV$/c^2$ [29]. D^*'s coming from the decay of B mesons are rejected using p_{CM} as discriminating variable; as explained in Sec. 3.2, D^*'s from $B \to D^*X$ processes have a maximum momentum of 2.5 GeV$/c$ in the center-of-mass frame, given the available energy at B-factories. This cut removes a potential bias on the measurement of the position of the production vertex of the D^0, $\vec{V}_{D^*}$, as the finite decay length of the B meson displaces that vertex far from the IP (a few hundred microns). Such threshold of 2.5 GeV$/c$ is quite high considering that it is meant for a pre-selection phase. Official skims for analyses make use of more conservative cuts, but they are made available much later than the calibration skims. The last cut on ΔM is to retain only the kinematically compatible $D^* \to D^0\,\pi_s$ decays; in fact the D^* mass (2010.26 MeV$/c^2$) is very close to that of D^0, resulting in a small available phase space for the π_s. This can be equivalently expressed using the variable Q, the released energy $Q = \Delta M - m_{\pi_s}$, which is just few MeV (5.85 MeV$/c^2$). In the next section the D^0 reconstructed mass and Q distribution obtained performing the reconstruction on the D^* skim data are shown (Figs. 4.1a–4.1b).

4.3 Selection criteria

The *skimmed* data go through the `basf2` reconstruction, which consists of searching for the D^* and D^0 candidates meeting the kinematical requirements, and reconstructing the two vertices, $\vec{V}_{D^*}$ and $\vec{V}_{D^0}$ (Sec. 3.2). In addition to the above criteria (1.1 and 1.2 for kaons and pions, and 3.1 for D^*'s), which are explicitly included in the reconstruction, further ones are applied:

1. $\boldsymbol{\pi_s^+, K^-, \pi^+}$ **candidates**

   1.1. must fall into the CDC and VXD polar angle θ acceptance $\theta \in [17°, 150°]$ (Sec. 2.3.1)

2. $\boldsymbol{D^0}$ **candidates**

   2.1. $1.80 < m_{K\pi}(\mathrm{GeV}/c^2) < 1.95$

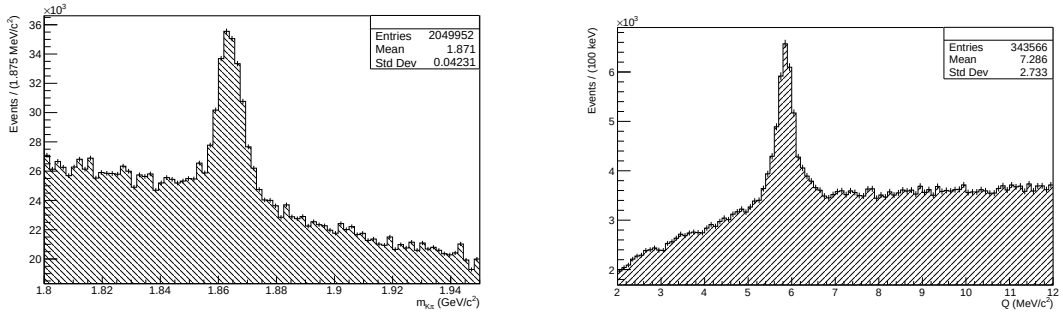
Then the geometrical fit to the chain with `TreeFitter` (presented in Sec. 3.4), is set up with the following constraints:

- D^0 mass constraint
- IP constraint (using the run-dependent beam spot calibrations, introduced in Sec. 3.5).

As described in Sec 3.4, the mass constraint consists in requiring the mass of selected composite particle to match the PDG mass. The momenta of D^0 daughters are thus adjusted in such a way that the invariant mass $m_{K\pi}$ matches $m_{D^0}^{\mathrm{PDG}}$ [29]. This can in principle introduce a bias in the lifetime measurement, as the magnitude of the track's momentum affects the coordinates of the impact parameter (d_0, z_0), propagating to the vertices determination. Even though the selection refinement has been finalized using the D^0 mass constraint, the study of background composition has raised the need to repeat the reconstruction, this time disabling that constraint. In Sec. 4.6 we explain the reasons behind the choice to continue with the second method, i.e. without the mass constraint.
While performing the reconstruction of the decay chain, `TreeFitter` constrains the D^0 production vertex to the beam spot region, whose position and shape can change over time and therefore need to be revised on a run-time scale basis. Among the D^* candidates thus obtained, i.e. the fitted chains, we remove all the fits that do not converge. The distributions resulting from the selection and the reconstruction described here and in the previous section are reported in Fig. 4.1.

4.3.1 Selection refinement

For the lifetime measurement an optimisation of the selection is required in order to minimise the sources of background. The refinement has been studied on `bucket6` data, while data taking was still ongoing. Optimisation is finalized on data and not on MC because efficiencies and detector performance are not perfectly reproduced by MC yet. For example, the purity and the efficiency related to a cut on the PID of tracks still shows discrepancies with MC (as discussed in Sec. 3.6) and the detector efficiencies are 100% for MC. Furthermore, the main purpose of the thesis is to give feedback on detector and analysis tools capabilities by studying the early data.
We study the impact on purity and relative efficiency of cuts on the hits on VXD and the PID of the D^0 daughters and the `pValue` returned by `TreeFitter`. The relative efficiency is computed taking the ratio of the signal yields obtained with and without the cut. The signal and



(a) **(b)**

Figure 4.1: Reconstructed D^0 mass **(a)** and Q **(b)** distributions obtained from reconstructing the D^* *skim* with the criteria presented on page 63.

the background yields are extracted as parameters from extended unbinned maximum likelihood (EUML) fits to the m_{D^0} distribution. The model used is

$$M_{\mathrm{PDF}}(m|\mu,\sigma,c_1,c_2) = N_{\mathrm{sig}} \times \mathrm{G}(m|\mu,\sigma) + N_{\mathrm{bkg}} \times p_2(m|c_1,c_2)\,. \tag{4.1}$$

The Gaussian G defines the signal while the second-order polynomial p_2 defines the background. For each cut, which is applied singularly and independently from the others, we measure how this affects efficiency and purity on the mass distribution. To evaluate the signal yield and the purity, the strategy adopted is the following:

1. Fit the distribution of D^0 mass using the model in Eq. 4.1.

2. Compute the yield $N_{\mathrm{sig}}^{\mathrm{NO\,cut}}$ from the integral of the Gaussian over an interval $[\mu-3\sigma,\mu+3\sigma]$.

3. Apply the cut on the chosen variable . Compute the yield $N_{\mathrm{sig}}^{\mathrm{cut}}$ from the integral of the Gaussian over an interval $[\mu'-3\sigma',\mu'+3\sigma']$ (μ',σ' are the parameters after the cut) and obtain

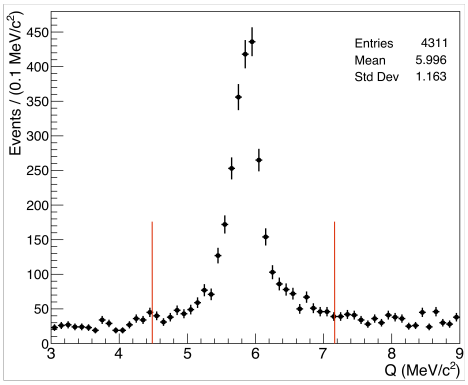
   - the efficiency, defined as
   $$\varepsilon_{\mathrm{cut}} = \frac{N_{\mathrm{sig}}^{\mathrm{cut}}}{N_{\mathrm{sig}}^{\mathrm{NO\,cut}}}\,, \tag{4.2}$$
   - the purity, defined as
   $$\Pi_{\mathrm{cut}} = \frac{N_{\mathrm{sig}}^{\mathrm{cut}}}{N_{\mathrm{sig}}^{\mathrm{cut}} + N_{\mathrm{bkg}}^{\mathrm{cut}}}\,. \tag{4.3}$$

Past analyses on D^0 coming from D^* used to select also a region in the (m_{D^0},Q) space to discard the majority of background. In fact, using $Q = \Delta M - m_\pi$ as selection variable considerably restricts the amount of combinatorial background. The effectiveness is justified by the the good experimental resolution arising from the cancellation of experimental uncertainty in the determination of the momenta of particles that make up the D^0 and D^* mesons.
Therefore, before proceeding with the selection refinement, we study a preliminar cut on the Q variable, while m_{D^0} is more critical and we prefer to not perform any cut on it in this phase, since m_{D^0} is correlated with the proper time. The correlation is due to the fact that decays

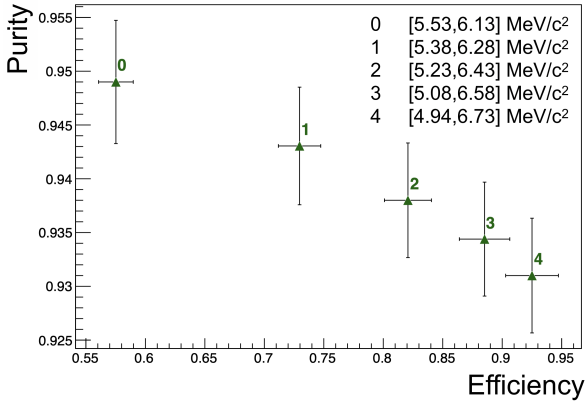
with too high reconstructed proper time might have a vertex further away from the interaction point than the true one, resulting in a too large reconstructed opening angle θ_{dau} between the D^0 daughters, and thus mass, since:

$$m^2_{D^0} = m^2_{K\pi} = m^2_K + m^2_\pi + 2(E^2_K E^2_\pi - p_K p_\pi \cos\theta_{\text{dau}}) \,. \tag{4.4}$$

Similarly, events with too low proper time have smaller opening angles and therefore too low mass.
We define some Q value intervals with increasing widths and evaluate relative efficiency and purity (Fig. 4.2b).



(a) Q distribution from `bucket6` data. It is obtained with only the cuts listed in 4.3.



(b) Efficiency and purity evaluated for five different intervals on the meaured Q.

Figure 4.2: First look at the Q distribution and study on efficiency and purity related to the cut on the Q variable.

For the refinement we choose a reasonable region with high efficiency $Q \in [4.54, 7.18]$ GeV$/c^2$ (illustrated in Fig. 4.2a); this cut is beyond the point 4 of Fig. 4.2b as it corresponds to a purity of $\sim 92\%$. The final selection cut on Q, used for the signal region definition, will be determined in Sec. 4.4 and is positioned between point 1 and 2 of Fig. 4.2b.

VXD hits cut

The high precision on vertex reconstruction is an essential element for lifetime measurements. PXD improves the resolution on the impact parameter and this has a strong effect on the detector capabilities of reconstructing short-lived particles decay vertices. The improvement in resolution brought by PXD is accompained by a loss in efficiency, especially during the first weeks of the data taking. For this reason, we evaluated the efficiency of requiring that both D^0 daughters have an hit in the PXD, and how this affects the lifetime determination. We distinguish among three cases for each daughter:

1. PXD: the track has a hit in at least one of the PXD layers
2. L3: the track has a hit in the SVD innermost layer
3. SVD: the tracks has hits in any SVD layer.

For example, for `bucket6` we observe that selecting D^0 candidates whose daughters have both hits in the PXD has an efficiency $\varepsilon_{\text{TOT}} \sim 80\%$ (Tab. 4.2). Instead, SVD has an overall very high efficiency, as confirmed by the aforementioned studies (Sec. 3.3.1).
All the cases considered do not show significative variation in purity, which is around 93%.

	1st daughter	2nd daughter	Logic	$\varepsilon_{\mathrm{TOT}}(\%)$		
				`bucket4`	`bucket6`	MC
1.	PXD	PXD	OR	90	97	99
2.	SVD	SVD	OR	99	100	100
3.	**PXD**	**PXD**	**AND**	**54**	**81**	**89**
4.	SVD	SVD	AND	100	100	100
5.	PXD	L3	AND	61	83	91
6.	L3	PXD	AND	64	84	93
7.	L3	L3	OR	99	100	100
8.	L3	L3	AND	83	88	95

Table 4.2: Summary of the efficiencies for `bucket4`, `bucket6` and MC related to each request made to the daughters of the D^0. During the course of the data taking the VXD performance has improved and the measured values have approached those of MC.

The inefficiency for the single track can be extracted combining the measured efficiencies for the cases 1. and 3., if equal probabilities to have a hit in the PXD for the two D^0 daughters are assumed. The inefficiency, or probability of not to find the hit on the PXD, is 6% on MC, 11% for `bucket6` and 28% for `bucket4`.
We can understand the 6% on MC by writing the contributions for $\varepsilon_{\mathrm{TOT}}$:

$$\varepsilon_{\mathrm{TOT}}^{\mathrm{AND}}(\mathrm{PXD}) = \varepsilon_{\mathrm{PXD}}^{(\mathrm{dau.\,1})} \cdot \varepsilon_{\mathrm{PXD}}^{(\mathrm{dau.\,2})} \cdot \varepsilon_{\substack{\mathrm{PAT.}\\ \mathrm{REC.}}}^{2} \cdot \varepsilon_{\substack{\mathrm{TRACK}\\ \mathrm{FIT.}}}^{2},$$

where $\varepsilon_{\mathrm{PXD}}$ is the detector efficiency, $\varepsilon_{\substack{\mathrm{PAT.}\\ \mathrm{REC.}}}$ is the efficiency of finding the PXD hit due to pattern recognition and $\varepsilon_{\substack{\mathrm{TRACK}\\ \mathrm{FIT.}}}$ the efficiency of fitting the track with that hit. For MC $\varepsilon_{\mathrm{PXD}}$ is set to 100%, as well as the fitting efficiency. This means that the $\varepsilon_{\mathrm{TOT}}^{\mathrm{AND}}(\mathrm{PXD})$ is determined by $\varepsilon_{\substack{\mathrm{PAT.}\\ \mathrm{REC.}}} \simeq 94\%$, which is compatible with the aforementioned 6% inefficiency.
Even though requiring both tracks having hits in PXD leads to an efficiency loss, it guarantees an improved impact parameter resolution and thus a better reconstruction of the D^0 decay vertex. In order to verify if this also has an impact on the statistical error of measured D^0 lifetime, different efficiencies, including those measured for `bucket4` and `bucket6`, are simulated on MC.
Five samples of $N = 10^4$ candidates are generated, each one with a different fraction (f_{AND}) of D^0 candidates having both daughters with hits in the PXD, thus with a different $N_{\mathrm{AND}} = f_{\mathrm{AND}} \cdot N$ candidates. For each sample two fits to the proper time distribution are performed to extract the statistical error on the lifetime (σ_τ):

- No cut: using the $N = 10^4$ candidates with no further requirement ($\sigma_\tau^{\mathrm{no\,cut}}$);
- Cut: using only the N_{AND} candidates, explicitly requiring that both D^0 daughters have hits in the PXD ($\sigma_\tau^{\mathrm{cut}}$).

The results, reported in Tab. 4.3, show the improvement on the error $\sigma_\tau^{\mathrm{no\,cut}}$ for higher AND fractions and that the difference between $\sigma_\tau^{\mathrm{cut}}$ and $\sigma_\tau^{\mathrm{no\,cut}}$ becomes smaller as the fraction increases: from ~ 1 fs to ~ 0.3 fs. The errors in the third column show the expected values of $\sigma_\tau^{\mathrm{cut}}$ assuming a $\varepsilon^{-\frac{1}{2}}$ scaling with respect to $\sigma_\tau^{\mathrm{cut}}(f_{\mathrm{AND}} = 100\%)$
We choose to maintain the requirement of D^0 daughters having hits in the PXD because the goal of our work is to verify the performance of the vertex detector. Moreover, the resulting increase on the statistical error is not decisive for the final measurement of τ_{D^0} ($\lesssim 0.3$ fs for

f_{AND} (%)	σ_τ (fs)		
	No cut	Cut	Scaled
54	**5.37**	**6.44**	**6.49**
70	5.15	5.63	5.70
81	**4.99**	**5.28**	**5.30**
90	4.89	5.02	5.03
100	—	4.77	-

Table 4.3: Statistical errors on τ_{D^0} obtained fitting the proper time distribution. 10^4 MC events are used and different f_{AND}'s are tested (the two corresponding to `bucket4` and `bucket6` efficiencies are highlighted in bold). In the "No cut" column the errors refer to the whole sample having a certain f_{AND}, while in the "Cut" column the requirement of having hits in the PXD for the D^0 daughters is applied. The third column shows the expected σ_τ considering the scaling $\propto \frac{1}{\sqrt{f_{\text{AND}}}}$ with respect to σ_τ^{cut} obtained with the 100% AND sample.

efficiencies $\geq 81\%$ on the whole `proc9` dataset) as will be discussed in the section 5.5 dedicated to the evaluation of the systematic error.
It is also worth noting that the estimated error on the proper time is significantly improved by the contribution of PXD. As shown in Fig. 4.3, the proper time error distribution related to D^0 candidates with both daughters having hits in the PXD is narrower and its mean is $\sim 37\%$ smaller with respect to the NAND candidates (in which at most one daughter has hits in the PXD).

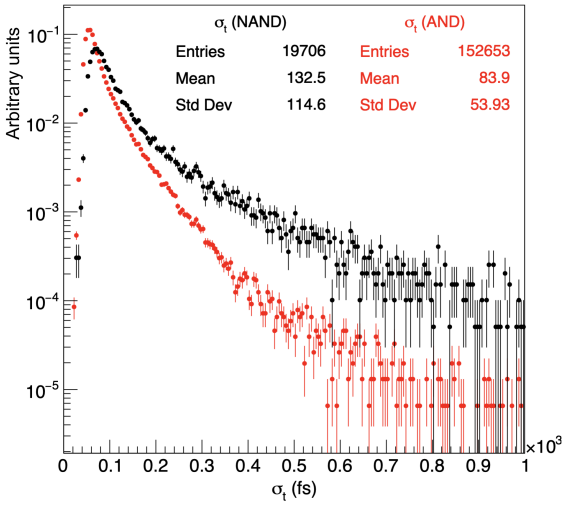


Figure 4.3: Estimated error on proper time for the AND and NAND candidates. The effect of requiring the hits in PXD is the lowering of the mean and the standard deviation of the distribution.

PID cut

We apply the same procedure used for the Q variable on other possible discriminating variables, in order to improve the selection. For example, requiring that the two daughters of the D^0 are identified as a K and a π with a PID$> \alpha$, with $\alpha = \{0.0, 0.2, 0.4, 0.6, 0.8\}$, the purity-efficiency trend is evaluated, as reported in Fig. 4.4. The result is that one can achieve very high purities but with the price of loosing a large amount of events, up to 60% for a 97% purity. Our choice falls on the point 3 (PID > 0.2), whose purity is still high ($> 95\%$) and also efficiency.

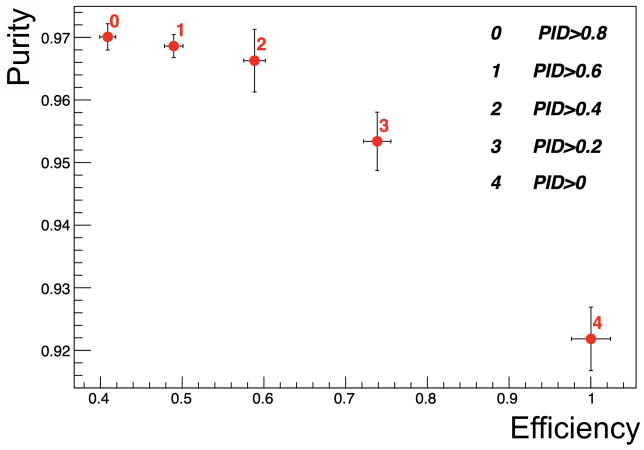


Figure 4.4: Efficiency and purity evaluated for five different requests on PID for D^0 daughters.

`pValue` cut

We also investigate on the possibility to cut on the fit quality. An index of quality is the χ^2 probability of the fit to the whole decay tree. We retrieve the value from `TreeFitter` and measure the correlation between purity and efficiency for different cuts on χ^2 probability. As shown in 4.5 we do not measure any significative improvement in purity for the selected configurations, as expected.

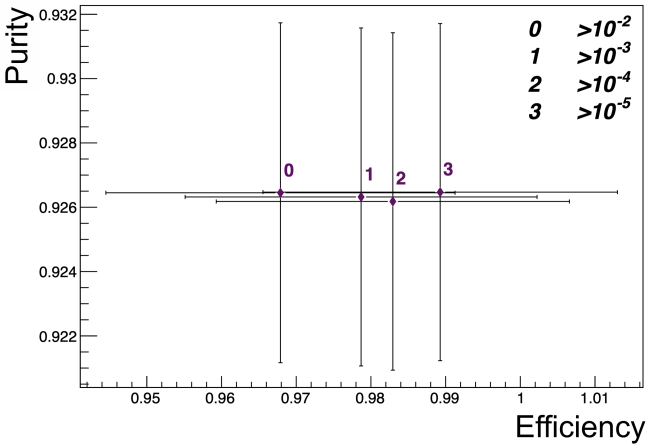


Figure 4.5: Efficiency and purity evaluated for four different cuts on the χ^2 probability of the fit.

As mentioned in Sec. 4.3, our initial cut consisted of rejecting only failed fits, i.e. selecting `pValue` > 0, which would correspond in this graph to a relative efficiency equal to 1 and a purity that, following the observed trend, is unchanged at $\sim 93\%$.

The study presented in this chapter has proved that additional requests may be applied to further reduce the background - hence increase purity - and to improve the measurement of the production and decay vertices, leading to a smaller error on proper time. Therefore, from now on, the following cuts to both D^0 daughters are added to those presented in Sec. 4.2 and 4.3:

- PID > 0.2;
- they should have at least one hit in the PXD detector;

The cut on Q is treated separately, as a preliminary fit to the Q distribution is necessary. This will be studied in the next section.

4.4 Signal region definition

The D^0 mass and Q distributions have been fitted to define a region in the (m_{D^0}, Q) plane, called *signal region*, where most of the signal candidates are located. This region finally defines the candidates whose proper time must be considered. This has been done with `bucket6`; once `proc9` data were provided, the fits have been re-done but the signal region has remained unchanged. Since the D^0 mass distribution needs to be obtained with a cut on the Q variable and vice versa, we proceed with an iterative method. The mass distribution is initially obtained with a loose cut on Q, then the estimated 3-σ interval centered on the mean is used for fitting the Q distribution, from which we extract the interval for Q. Once we apply the cut on Q, a new interval for m_{D^0} is produced, which returns an interval for Q compatible with the previous one. In this way a pair of consistent intervals is set and a signal region is defined.
We fit the reconstructed D^0 mass, after having applied all the cuts discussed in the previous section. Since the D^0 mass constraint is applied (Sec. 3.4), the fitted variable is the one obtained before performing the `TreeFitter`.
The reconstructed D^0 mass distribution is obtained with the further cut $Q \in [5.348, 6.353]$ MeV$/c^2$, determined by the fit in Fig. 4.7. The model has a Gaussian for signal and a second-order polynomial background, already defined in Eq. 4.1. The results of the fit are shown in Fig. 4.6 and Tab. 4.4.

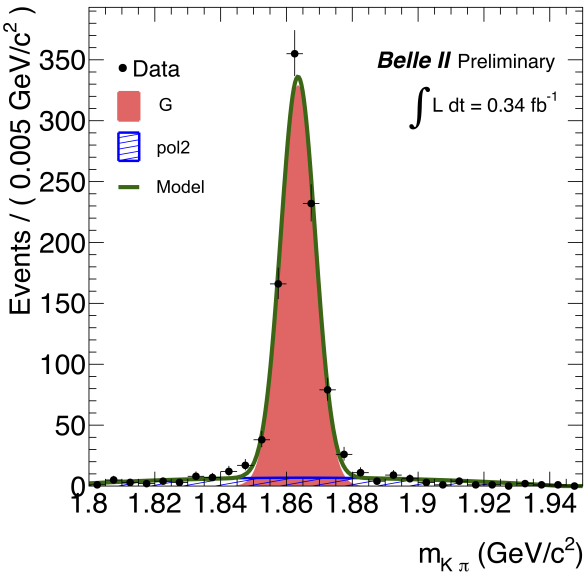


Figure 4.6: EUML fit to the reconstructed D^0 mass for candidates with $5.348 < Q(\text{MeV}/c^2) < 6.353$. The shaded region represents the signal candidates, while the blue region represents the background candidates.

parameter	extracted value
N_{sig}	860 ± 30
$\mu\ (\text{MeV}/c^2)$	1863.5 ± 0.2
$\sigma\ (\text{MeV}/c^2)$	5.22 ± 0.17
N_{bkg}	140 ± 15
$c_1\ (\text{MeV}/c^2)^{-2}$	-0.46 ± 0.11
$c_2\ (\text{MeV}/c^2)^{-3}$	-0.85 ± 0.17

Table 4.4: Parameters extracted from the EUML fit to the reconstructed D^0 mass.

The extracted value of the D^0 mass is slightly lower than the PDG average (1864.84 ± 0.05) MeV$/c^2$ [29], as reported in other analysis conducted on early data. This shift, probably due to non-optimal calibrations, is not an issue for the measurement of the D^0 lifetime at

this stage, as the related systematic error is negligible with respect to the statistical error and the systematic errors from other sources (which will be discussed in Sec. 5.5). As a feedback on Belle II performance during this early phase, it is interesting to note that resolution of 5.2 MeV/c^2 is better than those achieved by Belle and *BABAR* (around 7 MeV/c^2), while the purity in the signal region is lower ($\sim$ 95% against the *BABAR*'s 99% [32]).

We also fit the Q value distribution and define an interval to either enclose most of the peaking component and reject the majority of the background. The PDF used for the EUML fit is defined as:

$$Q_{\mathrm{PDF}}(q|\mu_1,\sigma_1,\mu_2,\sigma_2,q_0,a,b,c) = N_1 \times \mathrm{G}(q|\mu_1,\sigma_1) + N_2 \times \mathrm{G}(q|\mu_2,\sigma_2) + \\ + N_{\mathrm{bkg}} \times \mathrm{Thr}(q|q_0,a,b,c), \tag{4.5}$$

with two Gaussian contributions to allow different widths and the Threshold function, defined in A.7 (appendix A) and adopted to model the background in order to constrain it to zero for $q = q_0 = 0$. All functions (G and Thr) are normalized in the fit range. The result of the fit for the candidates belonging to the interval $m_{D^0} \in [1.848, 1.879]$ GeV/c^2 is reported in Fig. 4.7 and the extracted values of PDF parameters are reported in Tab. 4.5. The mean value $\mu_1 = (5.849 \pm 0.004)$ MeV/c^2 is compatible with the true value[1] $Q^{\mathrm{TRUE}} = \Delta M^{\mathrm{PDG}} - m_\pi^{\mathrm{PDG}} = (5855.7 \pm 0.4(\mathrm{stat}) \pm 1.7(\mathrm{syst}))$ keV/c^2 and very good resolution is observed: 170 keV/c^2 for the G_1.

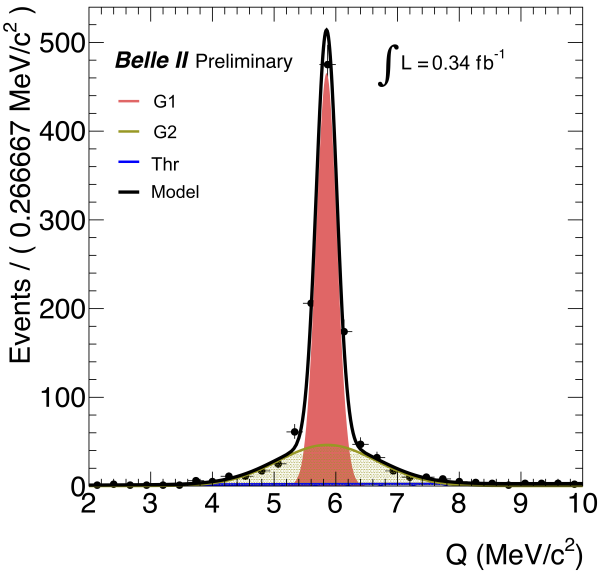


Figure 4.7: EUML fit to the reconstructed Q value for candidates with $1.848 < M(\mathrm{GeV}/c^2) < 1.879$. The shaded regions represent the signal candidates, while the blue region represent the background candidates.

parameter	extracted value
N_1	732 ± 18
μ_1 (MeV/c^2)	5.849 ± 0.004
σ_1 (MeV/c^2)	0.167 ± 0.004
N_2	351 ± 40
μ_2 (MeV/c^2)	5.86 ± 0.04
σ_2 (MeV/c^2)	0.8 ± 0.1
N_{bkg}	63 ± 50
q_0 (MeV/c^2)	(0 ± 6)
a	(0.0 ± 0.6)
b	$(0 \pm 8) \cdot 10^{-5}$
c (MeV/c^2)	(5 ± 24)

Table 4.5: Parameters extracted from the EUML fit to the reconstructed Q distribution.

The intervals on the two variables are chosen according to the following criteria:

- for the cut on M, the $[\mu - 3\sigma, \mu + 3\sigma]$ interval (values from Tab. 4.4) is taken;
- similarly, for the cut on Q the $[\mu_1 - 3\sigma_1, \mu_1 + 3\sigma_1]$ interval (values from Tab. 4.5).

[1] The m_π is the PDG average without error and ΔM is taken from [35].

Finally the signal region is defined as:

$$1.848 < m_{D^0}(\mathrm{GeV}/c^2) < 1.879 \tag{4.6}$$

$$5.348 < Q(\mathrm{MeV}/c^2) < 6.353\,. \tag{4.7}$$

4.5 Selected D^0 candidates in MC generic

For MC studies we apply where possible the same selection and reconstruction as on data. The differences between data and MC reconstruction are:

1. `TreeFitter` uses the MC information on the beams to constrain the D^0 production vertex. This means that *design* parameters, not measured ones, are used for reconstructing MC. For example, the y-size is the nominal one, which amounts to tens of nm, while in the run-dependent beam spot calibrations made from data, it is overestimated by at least one order of magnitude, for the reasons stated in Sec. 3.5.

2. For MC selection, we use high purity lists for D^0 daughters:

   2.1. p-value on tracks fit >0.001

   2.2. $|z_0| < 2\,\mathrm{cm}$

   2.3. PID>0.5

   and no cuts for the slow pion. Note that these cuts are different from those on data (for historical reasons) and therefore candidate numbers cannot be compared directly.

In Tab. 4.6 we report the number of D^0 candidates surviving the reconstruction and the selection, and the expected number of candidates for `proc9` luminosity. We also computed the number of candidates in our (m_{D^0}, Q) signal region.

		Mis-reconstructed D^0 candidates				
	signal	$c\bar{c}$	$u\bar{u}+d\bar{d}+s\bar{s}$	τ	B^+B^-	$B^0\overline{B}^0$
tot. cand.	147059	41575	1539+319+710	25	353	179
scaled ($2.6\,\mathrm{fb}^{-1}$)	4780	1351	50+10+23	8.1×10^{-1}	11	6
in signal region	3980	774	1.4+0.5+0.6	3×10^{-2}	0.4	0.2

Table 4.6: Number of selected signal and mis-reconstructed D^0 candidates from each generic MC sample. The signal column contains the number of truth-matched D^0 candidates, while for the other columns (mis-reconstructed D^0 candidates) we explicitly require that the truth-matching failed. In the first line (tot. cand.) we report the total number of candidates surviving the selection, while in the line below we rescale this number to `proc9` luminosity using the MC sample luminosity. The scaling is not reliable as the selection criteria between data and MC are different, and thus the selection efficiencies. The last line additionally requires that the D^0 candidates fall in the signal region, defined in the (m_{D^0}, Q) plane, described in section 4.4. The numbers are extracted from the reconstruction without the D^0 mass constraint – see next section.

The information relative to all the generated particles are referred to as "MC truth" and are available to the analysts. During the analysis this information can be used to study the selection, or a particular decay chain, matching the reconstructed tracks with the MC truth (*truth-matching*). A `basf2` module performs the matching on reconstructed composite particles on

a simulated sample. It maps, when applicable, reconstructed particles (`Particles`) to the generated ones (`MCParticles`). The algorithm searches for the first common mother (ancestor) of all daughter particles that form it and makes a `basf2` relation between `Particle` and `MCParticle`. For example, consider a BB event where one might want to reconstruct a D^0 candidate via $D^0 \to K^-\pi^+$ decay mode. If a given $D^0 \to K^-\pi^+$ candidate is reconstructed by combining a K^- and a π^+ candidates that are linked to the generated kaon and pion and originate from a $D^0 \to K^-\pi^+$ as illustrated in Fig. 4.8a, then the MC matching algorithm will make a `basf2` relation between reconstructed D^0 candidate (`Particle`) and generated D^0 (`MCParticle`). If, on the other hand, a given $D^0 \to K^-\pi^+$ candidate is reconstructed by combining a K^- and π^+ candidates that are linked to generated kaon and pion but do not originate from a $D^0 \to K^-\pi^+$ as illustrated in Fig. 4.8b, then the MC matching algorithm will make a `basf2` relation between reconstructed D^0 and the first common mother of the K^- and π^+, which is in this case a $\Upsilon(4S)$. The MC matching module loops over all particles within the specified particle list and performs the matching on them as well as all daughter particles.

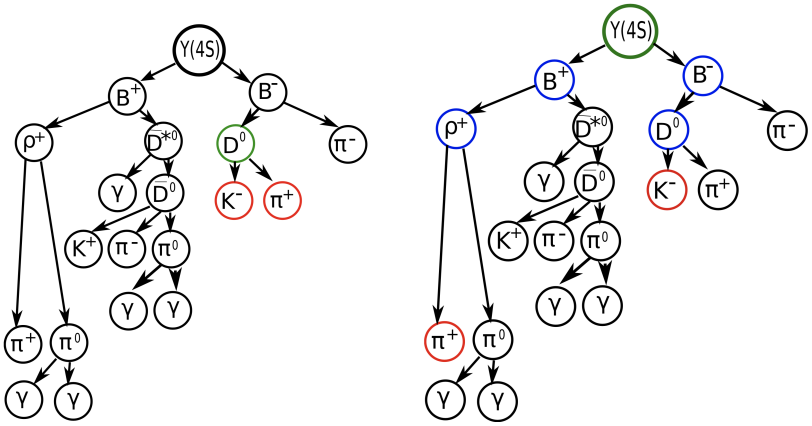


Figure 4.8: Generic B^+B^- event, with the whole genealogy. In the case **(a)**, the selected kaon and pion come from the D^0, hence the truth-matching is positive (*flag* 1). In the case **(b)** the pion is a random-one, that is, coming from another process. This produces a negative matching and a 0 *flag*.

The MC matching allows to characterize all the events that pass our selection and verify how many of these are actually signal events and, for those that are not signal, to identify the real generated process.

4.6 Signal and background studies on simulated events

In this section a survey of signal and background components is done through the characterization of their main variables: D^0 mass, Q and proper time. Distributions are obtained with the selection described in the sections above and the samples used are from MC, explicitly requiring positive truth-matching for studying signal candidates and negative for background. The expected sources of background are mainly the three below:

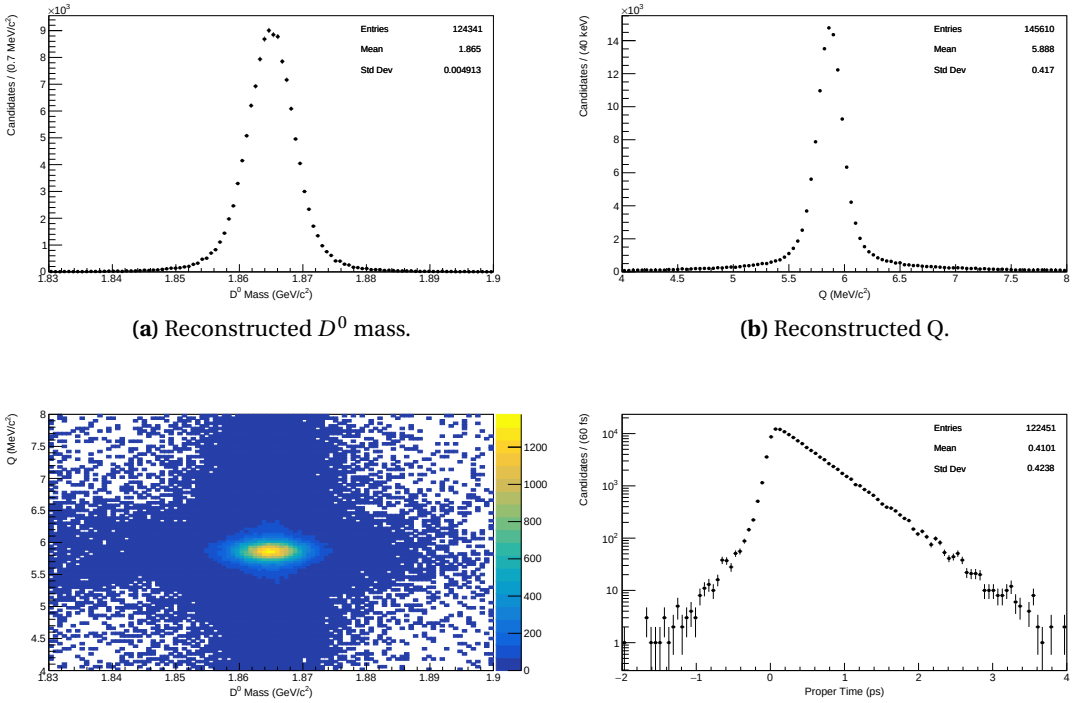
- D^*-tagged D^0's decaying into a different final state. The associated proper time is poorly measured since the missed particles in the decay degrade the reconstruction of the $\vec{V}_{D^0}$ vertices.
- True D^0's not coming from D^* 's, which are still reconstructed because a fake pion is attached to the D^0's.

- Random tracks combined into a D^0. This source is expected to not peak at the right value of m_{D^0} and has an undefined lifetime.

The qualitative description given here will be discussed during the next paragraphs; each of the generic MC samples gives a certain contribution to the classes listed above

Signal

The distributions of m_{D^0}, Q, and proper time for MC signal candidates are presented in Fig. 4.9. Clear peaks in m_{D^0} and Q are visible. From Fig. 4.9d the mean $\langle t \rangle$ and the standard deviation are very similar to $\tau_{D^0} = (410.1 \pm 1.5)$ fs (current PDG average [29]); the two would coincide in case of infinite precision as the distribution would be an exponential. This is a promising indication that the proper time resolution is *small* with respect to τ_{D^0}.



(a) Reconstructed D^0 mass.

(b) Reconstructed Q.

(c) Q vs. m_{D^0}.

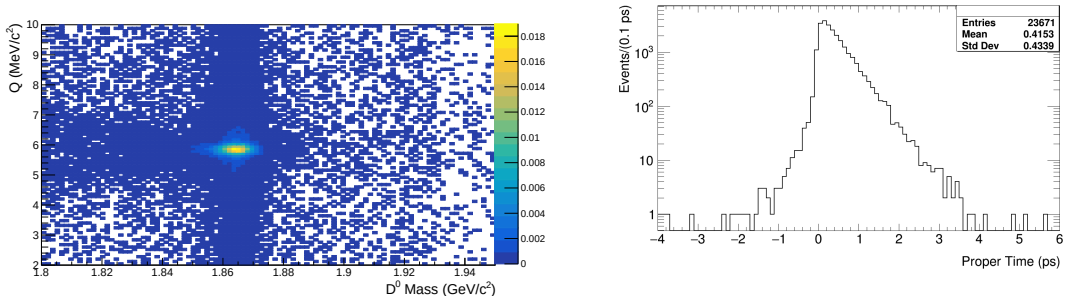
(d) Reconstructed proper time.

Figure 4.9: Distributions for truth-matched signal candidates.

$c\bar{c}$ background

From the MC study, whose results are summarised in Tab. 4.6, the only generic background that represents a significant fraction of the candidates in the signal region is the generic $c\bar{c}$ sample: these candidates will be addressed as *mis-reconstructed charm candidates*. The Q vs. m_{D^0} distribution for mis-reconstructed charm candidates (Fig. 4.10a) has the typical peak at the true m_{D^0} and Q, and also the proper time distribution (Fig. 4.10b) corresponds to that of D^0 candidates with the correct expected lifetime. The mis-reconstructed charm candidates are real D^0's that decay in a different final state, and therefore have different values of the reconstructed D^0 mass, but the decay and production vertices of the D^0's are almost correctly reconstructed as well as the proper times, as suggested by the mean of the distribution of

Fig. 4.10b. Hence, the difference between signal and mis-reconstructed charm candidates should consist, as a first approximation, in the resolution on proper time, but not the lifetime itself.



(a) Q vs. m_{D^0}. **(b)** Reconstructed proper time.

Figure 4.10: $c\bar{c}$ non truth-matched candidates. Both proper time ((**a**)) and Q-mass scatter plot ((**b**)) show that truth-matching categorize as background candidates that clearly peak in (m_{D^0},Q) plane and in proper time.

We therefore study in more detail the composition of mis-reconstructed charm candidates. In Tab. 4.7 we report the fraction of reconstructed candidates with a D^0 and/or a D^* correctly reconstructed, for candidates in the signal region. In the signal region we obtain:

1. a purity $N^{\text{true}}_{D^0}/N^{\text{tot}}_{D^0}$ of 84% ;
2. 15% of candidates with *fake* D^0's (flag 0) and thus also D^* has flag 0;
3. only 1% of candidates with a correctly reconstructed D^0 but a wrong D^* , i.e. a random or mis-reconstructed slow pion is attached to the signal D^0.

	True D^0	*Fake* D^0
True D^*	83.7%	-
Fake D^*	1.5%	14.8%

Table 4.7: Composition of reconstructed MC candidates according to the truth-matching. The case with a *fake* D^0 and a *true* D^* cannot exist, given that a mis-reconstructed D^0 (that here is called *fake*), automatically makes the D^* candidate as a fake one as well.

The purity obtained counting the fraction of signal candidates (84%) is lower with respect to those measured in sections 4.3.1 and 4.4, around 95%; this is explained by the fact that the simplified model 4.1 takes as signal the majority of mis-reconstructed D^0's, thus enhancing purity. Checking the truth information about candidates with flag 0, we find out that, in the signal region:

A. Candidates with mis-reconstructed D^0's:

   A.1 the 99.3% are real D^0's with (at least) a third daughter whose PDG code corresponds to that of the photon ($D^0 \to K\pi\gamma X$);

   A.2 the rest are *fake* D^0's, namely tracks coming from continuum (or fake tracks) combined to make a D^0. They can be defined as a combinatorial background source.

B. Candidates with correctly reconstructed D^0's but fake D^* 's with:

B.1 reconstructed π_s's corresponding to *fake tracks* (54%);

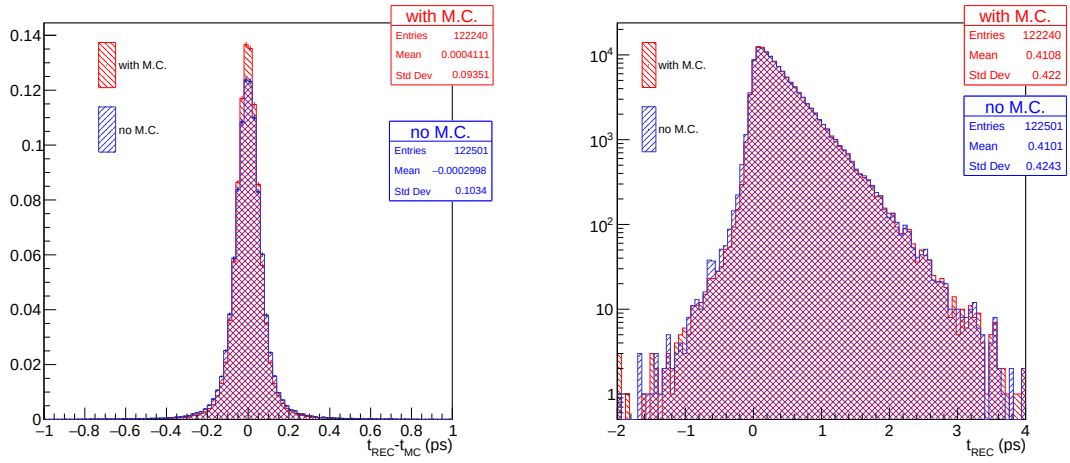
B.2 π_s's decaying in flight (27%);

B.3 *prompt* pions from a different mother (7%);

B.4 different chain, i.e. $D^* \to D^0 X \pi$ (6%);

B.5 wrong sign assignment of the charge of the π_s's (6%).

In Chapter 5 we discuss whether the two components A. and B. can be considered as signal, or treated as background. The first case holds if the related resolution functions on the proper time are compatible with that on signal candidates, implying that the purity of our selection on MC would be 100%.
The process described at point A.1 is generated by `EvtGen` and the MC matching interprets the K^-, π^+ and also γ as D^0 daughters, which is then classified as a non truth-matched process. Unfortunately the current version of `basf2` software does not allow to make the truth-matching algorithm pass also these events. After the release of the new software, scheduled for October 2019, it will be possible to study more thoroughly these candidates.
The presence of the non-negligible background component of real D^0 candidates radiating a photon makes our choice of using the D^0 mass constraint a weak point of our reconstruction. Indeed forcing the invariant mass of the two D^0 daughters to match $m_{D^0}^{\mathrm{PDG}}$, even when a third, non detected daughter is present, can induce a bias in the decay length measurement, and thus in proper time (bearing in mind the eq. 3.1). For example, looking at the equation 4.4, manipulating the K and π momenta can change the angle θ_{dau} and displace the D^0 decay vertex $\vec{V}_{D^0}$ from its true position. For this reason, we decide to perform the reconstruction without the mass constraint and compare some critical variables, to check if significant differences are present. Shown in figure 4.11 are the resolution on reconstructed proper time, and the proper time itself, for both methods.



(a) Resolution ($t_{\mathrm{REC}} - t_{\mathrm{MC}}$). **(b)** Proper time.

Figure 4.11: Comparison between the two different methods: D^0 mass constraint enabled in `TreeFitter` ('with M.C.', in red) and disabled ('no M.C', in blue). Even though the mass constraint produces a better resolution, it also results in a possible bias for processes like $D^0 \to K^-\pi^+\gamma$.

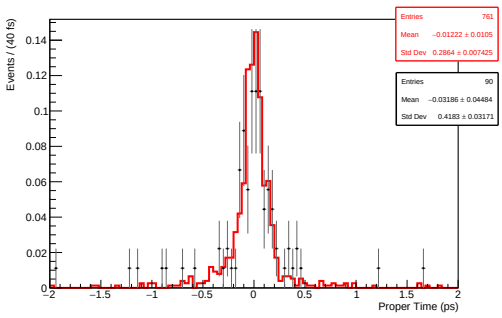
At distribution level, without performing any fit, we only observe a $\sim 10\%$ worse proper time resolution for the reconstruction without the D^0 mass constraint; nevertheless, we prefer this method as it is more robust for proper time determination.

u, d, s background

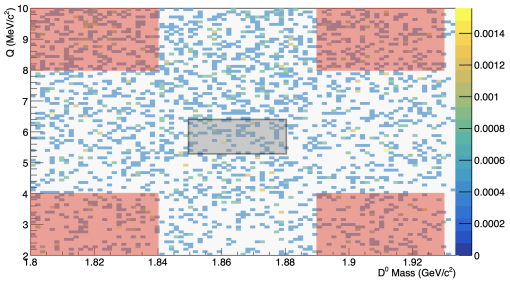
We expect the u,d,s processes to be sources of combinatorial background for our channel. Since normalisation with luminosity is not reliable, given the different selection with respect to data, we decide to estimate the combinatorial background yield in the *signal region* directly from data. A common technique consists of counting the number of events in the *sideband region*, and then scaling that number on the actual dimension of the *signal region*. This is justified under the hypothesis that the background uniformly populates the (m_{D^0}, Q) plane. The *sideband region* is taken sufficiently far from the *signal region* (Fig. 4.12b) and defined as:

$$\begin{aligned} &Q \in [2,4] \cup [8,10]\ \mathrm{MeV}/c^2 \\ &m_{D^0} \in [1.800, 1.840] \cup [1.885, 1.925]\ \mathrm{GeV}/c^2\,. \end{aligned} \tag{4.8}$$

The choice is to not investigate a more optimized strategy to locate the *sideband region* since its definition has a negligible impact on the systematic error. The consistency of the choice to infer the number of combinatorial background from the sidebands can be checked by comparing in the MC the distributions in proper time for the candidates belonging to the sidebands and to the signal region (Fig. 4.12a). Data from u,d,s samples are used and the Kolmogorov-Smirnov test on the two samples returns a probability of 16%.



(a) Comparison between the reconstructed proper time for the candidates selected from the sidebands, superimposed to that of candidates belonging to the signal region.



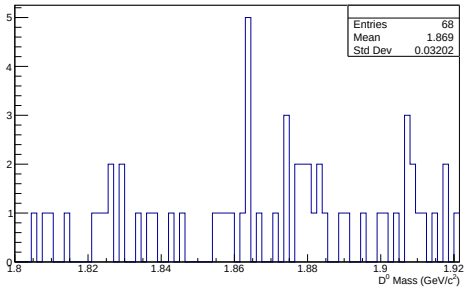
(b) *Sidebands* (red) and *signal region* (grey) from which the proper time distributions are taken to provide an estimate of combinatorial background.

Figure 4.12: Study on combined $u\bar{u}$, $d\bar{d}$, $s\bar{s}$ non truth-matched candidates.

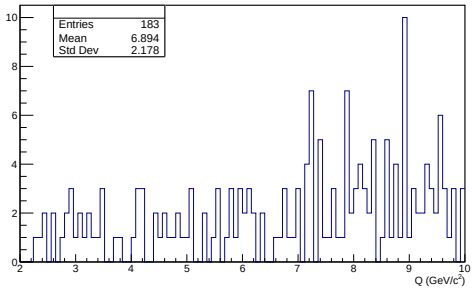
τ and $B\bar{B}$ background

The background associated with D^* production from B decays is almost completely removed with the cut on the center-of-mass frame momentum of the D^* . Fig. 4.13 shows the resulting distributions of the reconstructed candidates: from the mass and (m_{D^0}, Q) distributions it can be seen that still some D^0 candidates pass the selection and that the proper time is not centered on zero, as has instead been found for u,d,s processes. However, only 19 candidates are found in the signal region, which is a negligible number for the luminosity achieved in `proc9`.

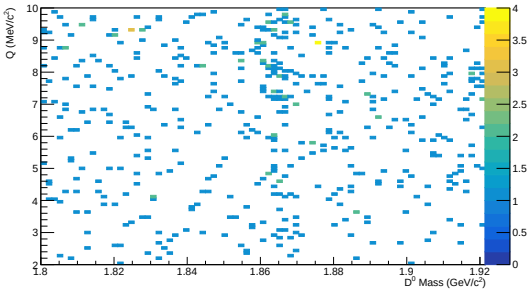
Due to the effectiveness of the selection, the number of final candidates is very small; hence the choice is to not use any dedicated PDF for D^0 candidates from $B\overline{B}$ samples. The same applies for candidates from τ events.



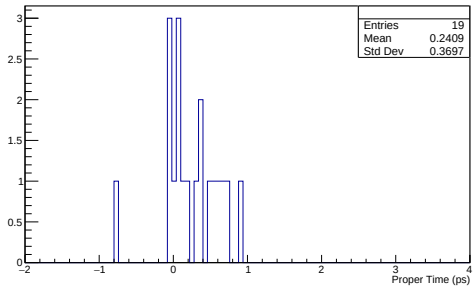
(a) Reconstructed D^0 mass.



(b) Reconstructed Q.



(c) Q vs. m_{D^0}



(d) Reconstructed proper time.

Figure 4.13: Distributions for non truth-matched $B\overline{B}$ candidates.

Summary

The study presented in this section has given an overview of the possible background sources. The seven generic samples are combined into the cocktail and the final components which need to be studied, for the fit to the proper time, are summarized in Tab. 4.8.

Component	Fraction (%)	Scaled to 2.6 fb^{-1}
Signal	83.7	3980
$D^0 \to K\pi\gamma$	14.6	695
Mis-reconstructed π_s	1.5	74
Combinatorial	0.2	8

Table 4.8: The sources of background identified from studies on MC generic.

4.7 Selected D^0 candidates in proc9 dataset

The distributions of reconstructed D^0 mass and Q value, reported in Fig. 4.14, result from the application of the selection criteria described in Sections 4.2 and 4.3. Clear signal peaks are visible on top of a quite flat distribution of background. In the previous sections the cuts on the reconstructed D^0 mass and the Q value are described. The *signal region* definition finally allows to reject a large fraction of the combinatorics while retaining most of the signal candidates, as demonstrated by comparing the distributions in Fig. 4.14 and those of Sec. 4.7.2.

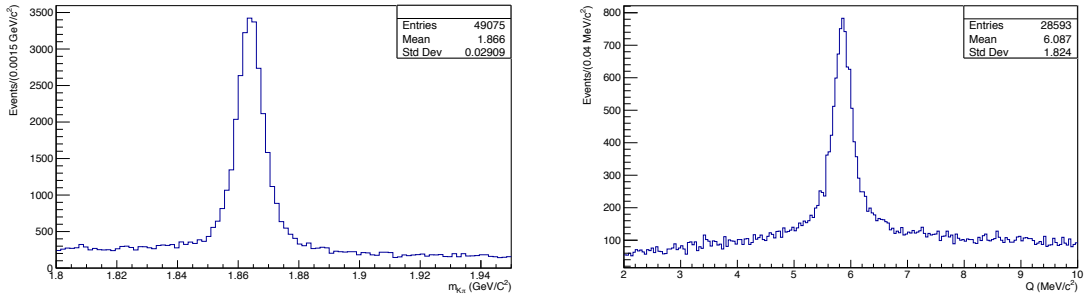


Figure 4.14: Reconstructed D^0 mass (left) and Q value (right) distributions after the selection described in 4.2 and 4.3.1.

4.7.1 Data - MC comparison

In Figs. 4.15–4.17 we report the comparison of the shapes obtained in data and MC for the following variables:

- reconstructed D^0 mass and Q value (Fig. 4.15)
- reconstructed transverse momentum of the slow pion and χ^2 probability of the fit, provided by `TreeFitter` (Fig. 4.16)
- reconstructed lab momentum of the D^0 daughters, and center-of-mass momentum of the D^0 (Fig. 4.17).

The normalization is based on the number of entries of the histogram. The momenta distributions on data show excellent agreement with MC. The resolution of D^0 mass and Q value are clearly better in MC than in data, probaly due to a non-perfect description of the detector material. The agreement for the χ^2 probability of the D^* decay chain fit is quite bad.

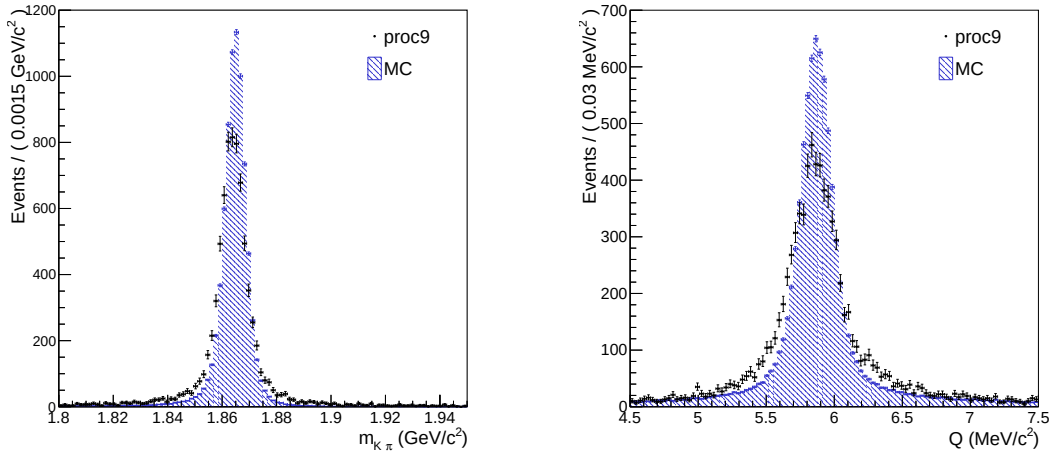


Figure 4.15: Reconstructed D^0 mass (left) and Q value (right) distributions after the selection described in 4.2 for data and MC. The plots are normalized to the same number of entries.

The fact that on MC the `pValue` is less populated for values close to 1 could be due to the stronger constraint associated with the smaller beam spot, which reduces the probability of the fit.

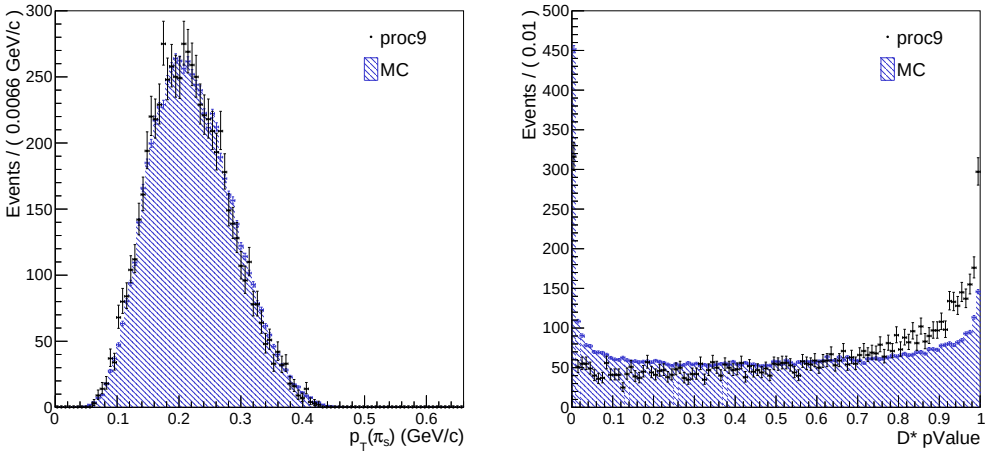


Figure 4.16: Reconstructed π_s transverse momentum (left) and D^* χ^2 probability (right) distributions after the selection described in 4.2 for data and MC. The plots are normalized to the same number of entries.

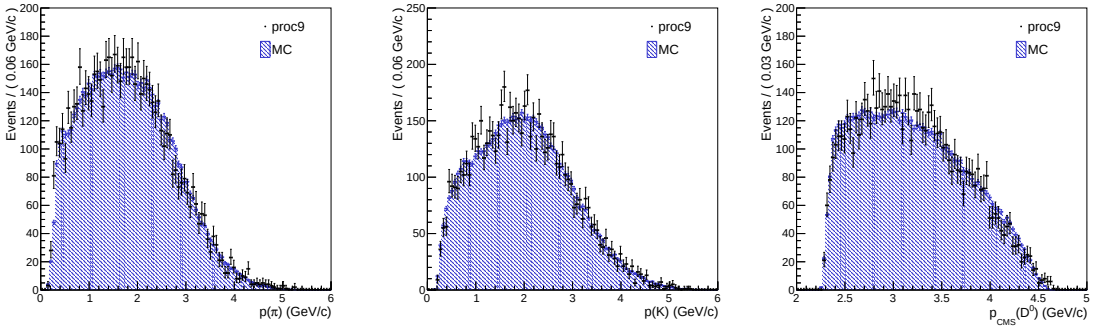


Figure 4.17: Reconstructed π (left) and K (middle) lab momentum, and D^0 momentum in the center-of-mass frame (right) distributions after the selection described in 4.2 for data and MC. The plots are normalized to the same number of entries.

4.7.2 Q and M fits for proc9 data

We repeat on proc9 data the same procedure applied for bucket6, in order to verify that compatible results are obtained for M and Q distributions with a higher statistics.
We fit the Q distribution with the model of Eq. 4.9:

$$\begin{aligned} Q_{\mathrm{PDF}}(q|\mu_1,\sigma_1,\mu_2,\sigma_2,\delta,\gamma,c_1) = & N_1 \times \mathrm{G}(q|\mu_1,\sigma_1) + N_2 \times J_{\mathrm{SU}}(q|\mu_2,\sigma_2,\delta,\gamma) + \\ & + N_{\mathrm{bkg}} \times p_1(q|c_1) \end{aligned} \tag{4.9}$$

which uses for the background a first-order polynomial. In this way a possible positive slope is still allowed, as for the Threshold function, but less parameters are used, considering that many of those of the Thr were compatible with zero (see Tab. 4.5). Moreover the Johnson's S_U distribution (defined in Eq. A.8, Appendix A) is used in place of the second Gaussian to have a better agreement along the tails. The fit results reported in Figure 4.18 and in Table 4.9.
We choose not to revise the Q interval for proc9 data since the obtained values of μ_1 and σ_1

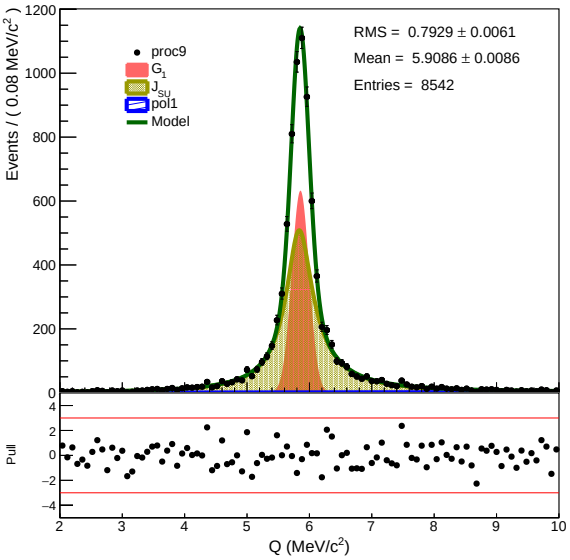


Figure 4.18: EUML fit to the reconstructed Q value for candidates with $1.848 < M(\mathrm{GeV}/c^2) < 1.879$. The shaded regions represent the signal candidates, while the blue region represent the background candidates.

parameter	extracted value
N_1	2700 ± 500
$\mu_1\,(\mathrm{MeV}/c^2)$	5.86 ± 0.01
$\sigma_1\,(\mathrm{MeV}/c^2)$	0.140 ± 0.006
N_2	5700 ± 500
$\mu_2\,(\mathrm{MeV}/c^2)$	5.83 ± 0.02
$\sigma_2\,(\mathrm{MeV}/c^2)$	0.210 ± 0.005
δ	0.57 ± 0.07
γ	-0.09 ± 0.05
N_{bkg}	130 ± 120
$c_1(\mathrm{MeV}/c^2)^{-2}$	0.7 ± 0.8

Table 4.9: Parameters extracted from the EUML fit to the reconstructed Q distribution.

are compatible with those of `bucket6` within the errors.

The increased statistics of `proc9` has required a revision also for the fit to the mass distribution. Indeed, the new PDF chosen to better shape data is the 4.10:

$$M_{\mathrm{PDF}}(m|\mu_1,\sigma_1,\mu_2,\sigma_2,\alpha,n) = N_1 \times \mathrm{G}(m|\mu_1,\sigma_1) + N_2 \times \mathrm{CB}(x|\mu_2,\sigma_2,\alpha,n) + \\ + N_{\mathrm{bkg}} \times p_0(m). \tag{4.10}$$

As before, the Gaussian G represents the signal, but this time a second signal function is used for the wider component with possible asymmetry due to final state radiation processes. For these reasons the Crystal Ball function (CB) (Eq. A.3, Appendix A) has been added to the PDF. The polynomial p_0, used for the flat residual combinatorial, fits only a small fraction of events; this model is more robust and better justified, since we expect the background to be flat in M. Even though the new CB component is considered as signal, we choose to maintain the selection interval for M as defined in 4.4, due to the difficulty of isolating those events to study their distribution in proper time. Besides, keeping the M interval as before allows us to reject both the poorly reconstructed events and the background ones. The fit and the parameters extracted are shown in Fig. 4.19 and in Table 4.10.

4.8 Selection summary

This chapter provides all the information necessary for the last part of the work, dedicated to fitting the proper time distribution. The composition of the signal and background is evaluated through the use of MC samples. However, the yields found are not reliable due to the slightly different selection between MC and data, as well as the known discrepancies on the efficiencies. For this reason, the possibility of treating the two main background components,

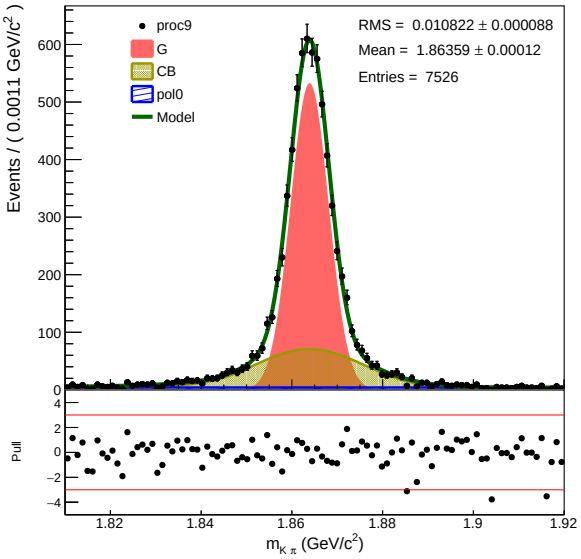


Figure 4.19: EUML fit to the reconstructed D^0 mass for candidates with $5.348 < Q(\mathrm{MeV}/c^2) < 6.353$.

parameter	extracted value
N_1	5100 ± 200
$\mu_1\,(\mathrm{MeV}/c^2)$	1863.90 ± 0.08
$\sigma_1\,(\mathrm{MeV}/c^2)$	4.2 ± 0.1
N_2	2000 ± 100
$\mu_2\,(\mathrm{MeV}/c^2)$	1863.9 ± 0.4
$\sigma_2\,(\mathrm{MeV}/c^2)$	12.1 ± 0.8
α	1.5 ± 0.1
n	3
N_{bkg}	400 ± 50

Table 4.10: Parameters extracted from the EUML fit to the reconstructed D^0 mass.

both found in the $c\bar{c}$ sample, the same way as the signal is evaluated in the next chapter. Instead, the amount of combinatorial background is estimated directly from the sidebands and will be taken as a fixed part of the final PDF.

5 D^0 lifetime extraction

This chapter presents the results of the analysis, during which a considerable effort was devoted to identify the model to extract the D^0 lifetime from the early Belle II data. Monte Carlo data were used to evaluate the resolution function of proper time and to characterize the signal and the background components (defined in Sec. 4.6). The fits to the proper time distributions are reported here, while those to D^0 mass and Q, which are preparatory to the measurement, are illustrated in Chapter 4 and in Appendix C. Finally, a first estimate of the systematic error is given, considering the main sources that could bring a bias to the measurement.

The strategy adopted for the lifetime extraction takes place in five steps:

1. Study the resolution function on a signal MC sample with size comparable to data, in order to find a sufficiently accurate PDF for the low statistics available on data.
2. Test the model on MC signal; the full statistics is used to check with low statistical uncertainty if the fit returns the expected value of τ_{D^0}.
3. Evaluate the two $c\bar{c}$ background components on the entire sample available, given the low measured fractions.
4. Quantify from the sidebands the yield of combinatorial background and determine its PDF.
5. Fit the proper time distribution with the final PDF describing the signal, the $c\bar{c}$ backgrounds and the combinatorial background. This is done both on `cocktail` and on data.

5.1 Signal MC characterization

The truth-matched candidates from signal MC sample are characterized by studying their resolution function for the fit to the proper time distribution. This part is followed by studies on background, which completes the parametrization for the final fit to data.

5.1.1 Resolution function

The resolution function for truth-matched signal candidates can be studied using the distribution of the proper time residuals, i.e. the difference between the reconstructed proper time and the generated one: $t_{\mathrm{REC}} - t_{\mathrm{MC}}$.
At Belle II the typical flight length of a D^0 is of the order of 200 µm, which is small compared to the distance between the beam spot and the first layer of the PXD (placed at 1.4 cm from the origin). As a consequence, the resolution function is independent of the proper time of the D^0 candidate, as shown in Fig. 5.1, where the distributions corresponding to three intervals of the generated proper time ($t_{\mathrm{MC}} \leq \tau_{D^0}$, $\tau_{D^0} \leq t_{\mathrm{MC}} \leq 2\tau_{D^0}$, $t_{\mathrm{MC}} \geq 2\tau_{D^0}$) are superimposed.

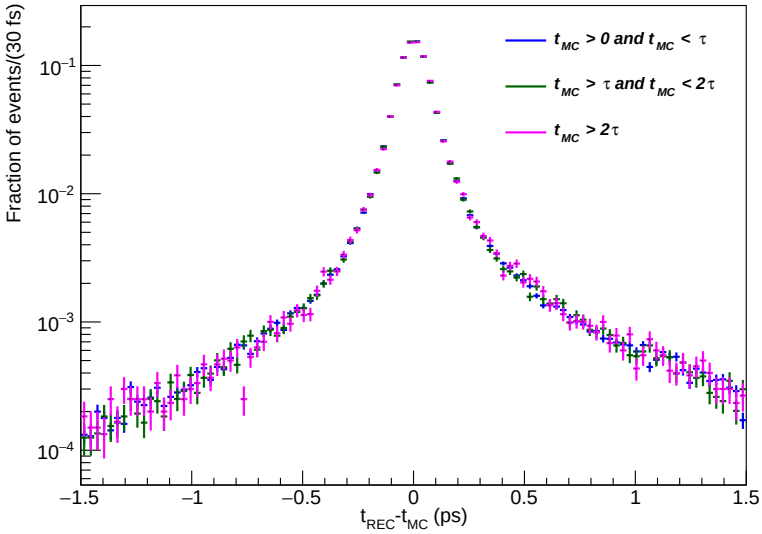


Figure 5.1: Proper time residuals for D^0 truth-matched candidates with different generated proper times t_{MC}. The three classes are taken from the full available MC statistics and normalized to their number of candidates.

We reduce the signal MC sample to a subset of the same number of D^0 candidates expected for `proc9` data - 6764, in order to find a model which is sufficiently accurate for the available statistics. Then we fit the $t_{\mathrm{REC}} - t_{\mathrm{MC}}$ distribution with two or three Gaussians. The fits are reported in Fig. 5.2 and the extracted values are in Tab. 5.1.

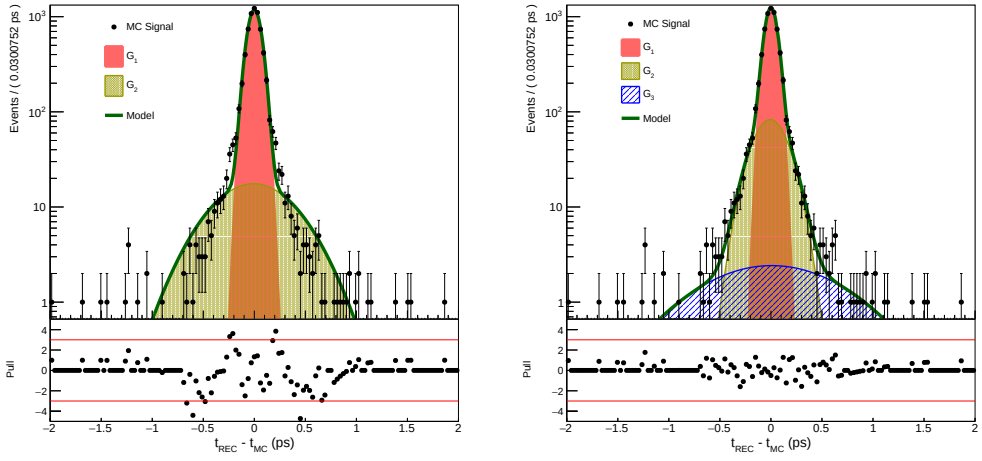


Figure 5.2: EUML fit to the proper time residuals with the sum of two (left) or three (right) Gaussians. The number of events is reduced to the statistics of `proc9` data.

As expected, all the Gaussians have means compatible with zero. For the data, instead, it may happen that there is some misalignment of the vertex detector or a systematic shift of the beam spot. These two factors can result in one of the means at least to be different from zero. The resolution of the best measured candidates, fitted by the first Gaussian, is of the order of 60 fs. The wider Gaussians (both G_2 or G_3) describe candidates that are truth-matched

parameter	2 Gaussians	3 Gaussians
N_1	6230 ± 80	5540 ± 140
μ_1 (fs)	0 ± 1	1 ± 1
σ_1 (fs)	64.4 ± 0.8	57 ± 1
N_2	570 ± 40	1130 ± 110
μ_2 (fs)	-8 ± 17	-7 ± 6
σ_2 (fs)	386 ± 14	162 ± 11
N_3	-	137 ± 26
μ_3 (fs)	-	10 ± 60
σ_3(fs)	-	683 ± 70

Table 5.1: Extracted parameters of the models with two and three Gaussians for the proper time resolution.

but with poorly measured vertices positions and thus have lower sensitivity to misalignment. Since we have not further investigated this class of candidates and for the above considerations, it is safe to fix the mean of the second Gaussian either to zero, or constrain it to the average of the narrower Gaussian and let the μ_1 float to reveal such possible shifts of the proper time.

From this study the best resolution function is the one with three Gaussians as it better describes the core of the distribution (see the pulls for $-0.5 \leq t_{\rm REC} - t_{\rm MC} \leq 0.5$). This is expected because we are using Gaussians with fixed widths: we basically assume the resolution to be a sum of delta-functions placed at two or three fixed σ_i's. However, the proper time error σ_t for truth-matched signal candidates (Fig. 5.3) has a broad distribution with a mean of ~ 60 fs and a standard deviation of 40 fs. The more Gaussians are used, the more populations of the distribution are taken into account and the fit improves. Past analyses conducted at *BABAR* and Belle with full statistics used a sum of two or three Gaussians and the width of each Gaussian was given by the per-event proper time error σ_t, multiplied by a scaling factor (each Gaussian had its own) to take into account the mis-evaluation of the proper time error. Then the signal PDF was multiplied by the proper time error PDF, i.e. the probability to have a given error σ_t. Weighing each measured t_{D^0} by its estimated error gives a better description of the data, as this allows to exploit the full measurement information. However the two simple models proposed so far are accurate enough considering the available statistics.

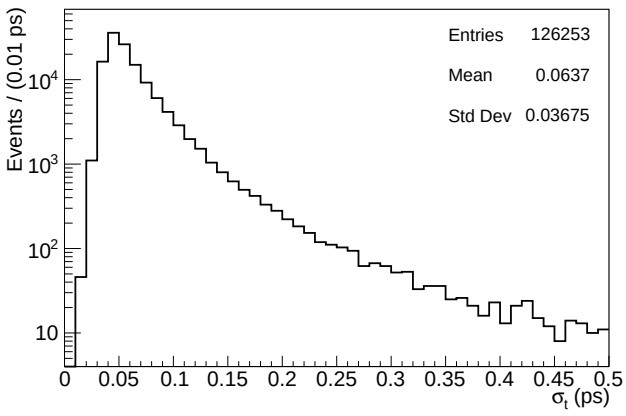


Figure 5.3: Proper time error distribution for signal truth-matched candidates.

Although the pulls of the fit with 3 Gaussians are better than the one with 2, the description of the resolution would really improve only if the per-candidate proper time error was implemented in the PDF. Hence the choice is to proceed with a two Gaussian resolution function given the fact that:

1. The fit to the truth-matched proper time distribution, using this model with the constraint $\mu_2 = 0$, yields the expected value for the D^0 lifetime — as will be demonstrated in the next section.

2. Given the low statistics, we prefer to use a PDF with as few parameters as possible. This arbitrary choice brings a systematic uncertainty which will be quantified in Sec. 5.5 to be rather small.

5.1.2 Signal PDF

The model for signal candidates is composed of two Gaussians convolved with the exponential function. The use of an exponential for the decay of the D^0's has been justified in section 3.2, while the choice of two Gaussians is discussed in the previous study on the resolution distribution. The initial PDF is

$$T_{\mathrm{PDF}}(t|\mu_1,\sigma_1,\mu_2,\sigma_2,\tau) = N_1 \times R(t|\mu_1,\sigma_1,\tau) + N_2 \times R(t|\mu_2,\sigma_2,\tau) \tag{5.1}$$

where R indicates the result of the convolution between the Gaussian and the exponential:

$$\mathrm{R}(x|\mu,\sigma,\tau) = \mathrm{Exp}(y|\tau) * G(y|\mu,\sigma). \tag{5.2}$$

The weight of each component is determined from the extended unbinned maximum likelihood fit (EUML), which therefore treats the two factors N_1 and N_2 as floating parameters as well.
The model 5.1 with two Gaussians and all the parameters floating is applied to the MC signal and the fit result is given in Fig. 5.4. Among the values reported in Tab. 5.2, the τ_{D^0} differs from the true one by $\sim 4.8\sigma$ and μ_2 is significantly different from zero. This suggests the non-optimality of the model since it should at least guarantee to reproduce the τ_{D^0} used for the event generation (410.1 fs) and give no-bias for the resolution functions. Furthermore, the measured widths are higher than those obtained fitting the resolution distribution (Tab. 5.1); this is due to the fact that using the entire statistics the tails are more populated and events with higher errors are also taken into account. These findings confirm the choice made in Sec. 5.1.1 to avoide μ_2 to be a further floating parameter; this can be done, for example, setting it to zero or equal to μ_1.
Indeed, these two slightly alternative models are tried to verify the dependence of τ_{D^0} from μ_2. Shown in Fig. 5.5 and Tab. 5.3 is the model with $\mu_1 = 0$, and it can be verified that it returns compatible results with respect to the model $\mu_2 = \mu_1$ (which is reported in Fig. C.3, Appendix C).
Both models yield to a τ_{D^0} compatible with the true one and μ_1 still compatible with zero. We therefore decide to proceed with the model fixing μ_2 to zero, taking care to estimate the error associated with this choice with respect to the other ($\mu_2 = \mu_1$). Moreover, the two fit results just mentioned are still to be preferred to the one with both means floating, which shows higher correlation between the μ's and τ_{D^0} and leads to an underestimation of τ_{D^0}.

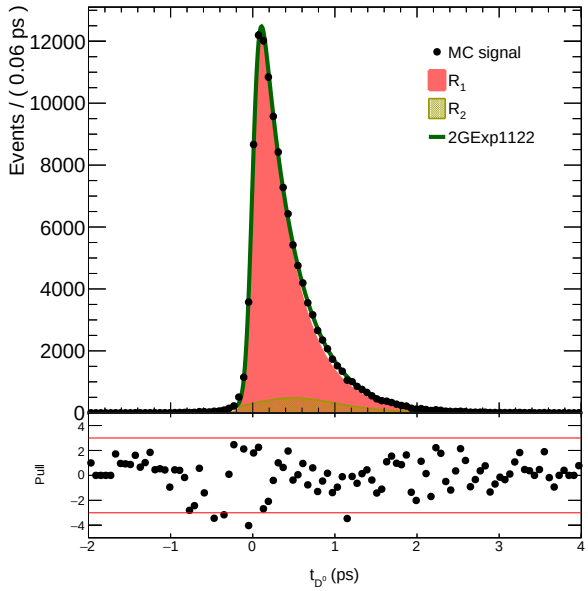


Figure 5.4: EUML fit to the reconstructed proper time for truth-matched MC candidates belonging to the signal region. The model used is the one with two Gaussians with all the parameters left floating.

parameter	extracted value
N_1	111500 ± 900
μ_1 (fs)	1 ± 1
σ_1 (fs)	68 ± 1
N_2	11000 ± 800
μ_2 (fs)	210 ± 30
σ_2 (fs)	450 ± 10
τ (fs)	391 ± 4

Table 5.2: Fit parameters for the model with 2 Gaussians with all the parameters left floating. The extracted central value of τ_{D^0} is $\sim 5\sigma$ from the true one and μ_2 is not compatible with zero.

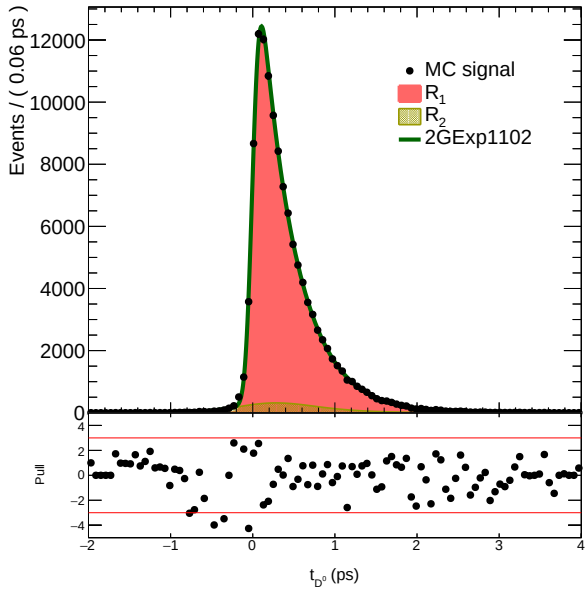


Figure 5.5: EUML fit to the reconstructed proper time for truth-matched candidates belonging to the signal region. The model used is the one with the second Gaussian having the mean fixed to zero.

parameter	extracted value
N_1	115700 ± 500
μ_1 (fs)	0 ± 1
σ_1 (fs)	69 ± 1
N_2	6800 ± 300
μ_2 (fs)	0 fixed
σ_2 (fs)	400 ± 9
τ (fs)	410 ± 1

Table 5.3: Fit parameters for the model with $\mu_2 = 0$. The alternative model, with the second Gaussian having $\mu_2 = \mu_1$ is reported in fig. C.3, appendix C. In both cases the central value of τ_{D^0} is compatible with the true one and μ_1 is compatible with zero.

5.2 $c\bar{c}$ backgrounds

The model with the second Gaussian having the mean fixed to zero is applied also to the $c\bar{c}$ sample, limited to non truth-matched candidates and divided for the two background components. As shown in previous Chapter (Sec. 4.6), the $c\bar{c}$ background is composed by two main classes of candidates: the mis-reconstructed D^0's, related to the presence of a final

state radiation, and the mis-reconstructed D^* 's, due to errors in the reconstruction of the π_s's. We need to ascertain whether these two components can be treated as signal, and thus use the same resolution function as the "well reconstructed" signal component. First we fit the $K\pi\gamma$ candidates proper time with the two Gaussian model with $\mu_2 = 0$ and obtain the PDF shown in Fig. 5.6 and the values of Tab. 5.4.

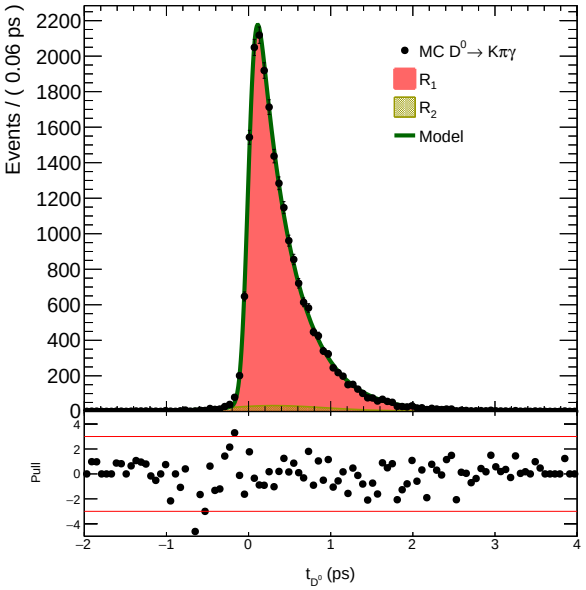


Figure 5.6: EUML fit to the reconstructed proper time for non truth-matched candidates belonging to $c\bar{c}$ subsample of $K\pi\gamma$ candidates.

parameter	extracted value
N_1	20550 ± 170
μ_1 (fs)	1.2 ± 1.7
σ_1 (fs)	73 ± 2
N_2	850 ± 90
μ_2 (fs)	0 fixed
σ_2 (fs)	540 ± 30
τ (fs)	405 ± 3

Table 5.4: Fit parameters for the model with the second Gaussian having $\mu_2 = 0$.

The same model is applied to the candidates with mis-reconstructed π_s's (Fig. 5.7 and Tab. 5.5).

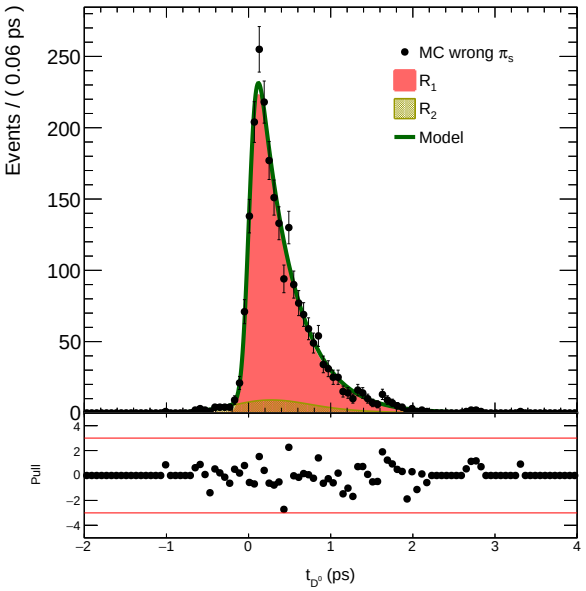


Figure 5.7: EUML fit to the reconstructed proper time for non truth-matched candidates belonging to the $c\bar{c}$ subsample having mis-reconstructed π_s's.

parameter	extracted value
N_1	2090 ± 60
μ_1 (fs)	11 ± 5
σ_1 (fs)	72 ± 6
N_2	200 ± 50
μ_2 (fs)	0 fixed
σ_2 (fs)	410 ± 50
τ (fs)	395 ± 10

Table 5.5: Fit parameters for the model with the second Gaussian having the mean fixed to zero.

The two measured σ_1's are compatible with each other and with that of the signal, as well as the μ_1 of mis-reconstructed D^0's, while the μ_1 of mis-reconstructed D^* 's differs from zero by more than 2σ; this shift should be checked by using a larger MC dataset and verify if a true bias exists. The two σ_2's are still compatible and we assume that the second Gaussian in the final fit can incorporate both of them, especially since both $c\bar{c}$ background components are estimated from MC to be respectively 15% and < 2% of the total candidates (Tab. 4.8). Therefore we decide to treat these candidates with the same resolution function as for signal; a possible bias introduced by this choice is evaluated in the next Section and in Sec. 5.5, where the systematic error induced by the PDF choice is evaluated.

5.3 Fit including combinatorial background

The fit to all the generic samples merged in the `cocktail` is performed using the PDF with 2 Gaussians for the resolution (the wider one having $\mu_2 = 0$) – that include both signal and $c\bar{c}$ backgrounds – and with the combinatorial background component PDF. This latter is obtained following the procedure described in Chapter 4: the proper time distribution for the candidates belonging to the sidebands is fitted and the corresponding PDF (Fig. 5.8) is inserted as a fixed part of the total PDF for the candidates in the signal region. The model used is a sum of two Gaussians, G_a and G_b, each described by the yield B, the mean μ and the standard deviation σ. The fraction $\alpha = \frac{B_a}{B_a+B_b}$ is then used in the fit to the `cocktail` as the relative weight of the two Gaussians combined for the third term of the PDF in Eq. 5.3

$$\begin{aligned} T_{\mathrm{PDF}}(t|\mu_1,\sigma_1,\sigma_2,\tau) = & N_1 \times R(t|\mu_1,\sigma_1,\tau) + N_2 \times R(t|\mu_2=0,\sigma_2,\tau) \\ & N_3 \times \left[\alpha\, \mathrm{G}(t|\mu_{3,a}^{\mathrm{fixed}},\sigma_{3,a}^{\mathrm{fixed}}) + (1-\alpha)\mathrm{G}(t|\mu_{3,b}^{\mathrm{fixed}},\sigma_{3,b}^{\mathrm{fixed}})\right]. \end{aligned} \tag{5.3}$$

The means and widths of G_a and G_a, extracted from the sidebands, are taken as fixed parameters, as well as the yield of the whole combinatorial background PDF (N_3 in the Table 5.7), obtained by scaling $B_a + B_b$ to the ratio between the areas of the sidebands and the signal region (parameters are summarized in Tab. 5.6). The possible bias induced by fixing N_3 is discussed and quantified in Sec. 5.5: a variation of 100% of the number of combinatorial candidates has an effect on the lifetime measurement. However the procedure is rather consistent since the number of events is not predicted, but measured by the sidebands.
The PDF used for the fit to the `cocktail` consists of the two Gaussian model 5.1 (with $\mu_2 = 0$) for the signal and the $c\bar{c}$ backgrounds, as described in 5.1.1 and confirmed in 5.2, and the background PDF obtained above (Tab. 5.6). The fit is shown in Fig. 5.9 for the sample rescaled to the statistics of `proc9` and the parameters are listed in Tab. 5.7.
The estimated D^0 lifetime is compatible with the true one but is 5 fs lower than the τ_{D^0} obtained for MC signal (Tab. 5.3). This is an indication of systematic error related to the absence of a dedicated PDF for the $c\bar{c}$ background component, which must be taken into account in the final estimation of Sec. 5.5. The extracted μ_1 and σ_1 are compatible with the ones measured for MC signal; the larger σ_2 is due to the presence of background and to the fact that the G_2 yield is smaller on `cocktail` with respect to the signal.

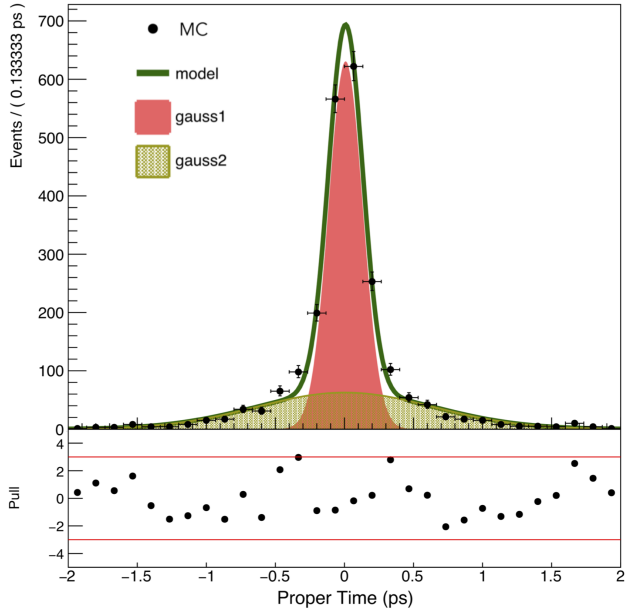


Figure 5.8: EUML fit to the proper time distribution of MC candidates belonging to the sidebands (defined in 4.8).

parameter	extracted value
B_a	1470 ± 50
μ_a (fs)	8 ± 4
σ_a (fs)	123 ± 5
B_b	750 ± 50
μ_b (fs)	3 ± 2
σ_b (fs)	640 ± 20

Table 5.6: The parameters obtained for the candidates from the sidebands are used for the final fit to the `cocktail`.

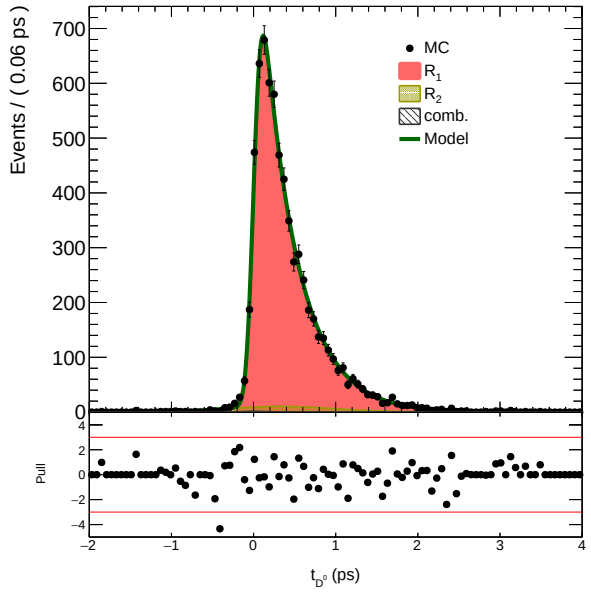


Figure 5.9: EUML fit to the proper time of candidates from the `cocktail` dataset, reduced to `proc9` statistics. The model with 2 Gaussians convolved with the exponential has $\mu_2 = 0$ and all the components for the combinatorial background are fixed.

parameter	extracted value
N_1	6500 ± 100
μ_1 (fs)	6 ± 3
σ_1 (fs)	73 ± 4
N_2	290 ± 60
μ_2 (fs)	0 fixed
σ_2 (fs)	560 ± 60
N_3	10 fixed
$\mu_{3,a}$ (fs)	8 fixed
$\sigma_{3,a}$ (fs)	123 fixed
$\mu_{3,b}$ (fs)	3 fixed
$\sigma_{3,b}$ (fs)	640 fixed
τ (fs)	405 ± 6

Table 5.7: The parameters obtained from the fit to the `cocktail`, MC.

5.4 Fit to Data

5.4.1 Combinatorial background model

The estimation on u, d, s and $c\bar{c}$ background studies conducted in Chapter 4 and in the previous sections, justify the choice to consider for the fit to the D^0 candidates in the signal region the four components as was done for the `cocktail`. The total yield of the combinatorial background is given by $B_a + B_b$, both extracted from the fit to the candidates of the sidebands, where a two-Gaussian PDF is used (Fig. 5.10 and Tab. 5.8).

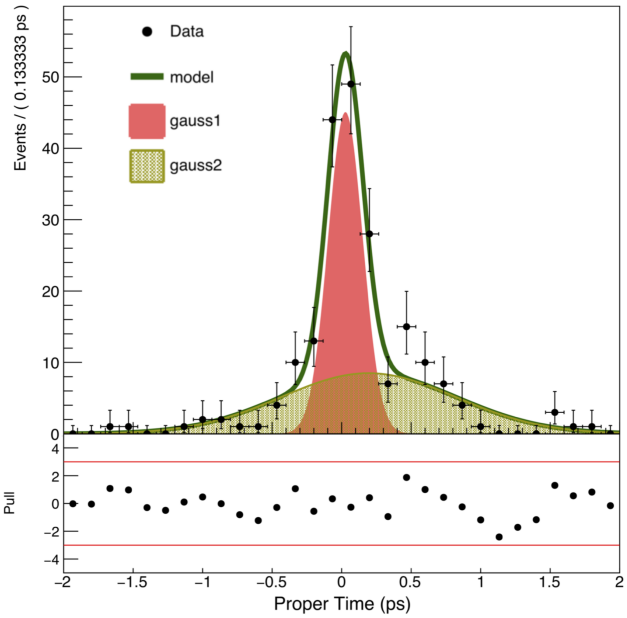


Figure 5.10: EUML fit to the proper time distribution of candidates belonging to the sidebands (defined in Eqs. 4.6).

parameter	extracted value
B_a	106 ± 15
μ_a (fs)	27 ± 16
σ_a (fs)	125 ± 17
B_b	100 ± 14
μ_b (fs)	190 ± 70
σ_b (fs)	630 ± 60

Table 5.8: The parameters obtained for the sidebands candidates are used as fixed values for the final fit to `proc9` data.

Due to the low statistics, parameter have large uncertainties. For what concerns the broader Gaussian, there seems to be a higher right-tail, which cannot be investigated for the small population of the sidebands. Nevertheless, the PDF is inserted in the final fit and rescaled for the expected number of combinatorial background candidates in the signal region, which are only 20. The widths of the two Gaussians are compatible with those measured on `cocktail` sidebands but are very different with respect to signal: σ_a is nearly a factor two larger. This would confirm that candidates from sidebands, obtained from random combination of tracks or being very poorly reconstructed signal events, have worst resolution.

5.4.2 Final fit

The same procedure described in Sec. 5.3 is performed for the fit to data, taking into account all the components (signal, $c\bar{c}$ background and combinatorial one). The proper time distribution from `proc9` data is obtained using the definition of signal region given in Sec. 4.4. Figure 5.11 shows the final fit to data and Table 5.9 resumes the floating parameters and the fixed ones.
The extracted value of the D^0 lifetime is

$$\tau^*_{D^0} = (400 \pm 6)\ \text{fs} \tag{5.4}$$

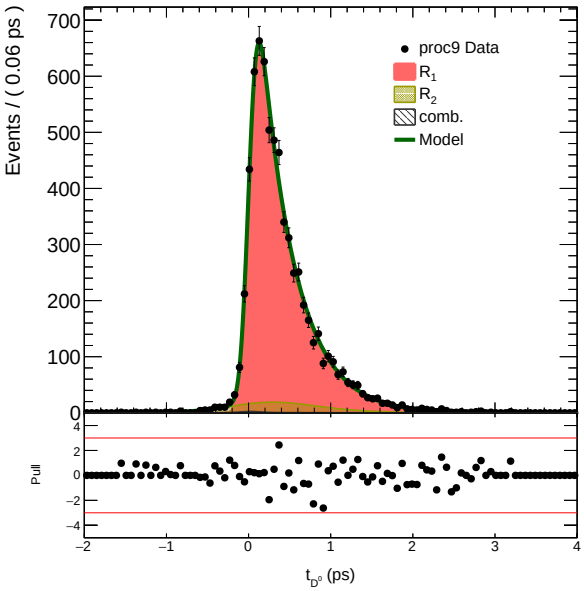


Figure 5.11: EUML fit to the proper time of candidates from `proc9` data. Candidates belong to the signal region (defined in Eqs. 4.6, 4.7) and the model used was tested on `cocktail`.

parameter	extracted value
N_1	6300 ± 100
μ_1 (fs)	9 ± 4
σ_1 (fs)	83 ± 4
N_2	440 ± 80
μ_2 (fs)	0 fixed
σ_2 (fs)	450 ± 50
N_3	20 fixed
$\mu_{3,a}$ (fs)	27 fixed
$\sigma_{3,a}$ (fs)	125 fixed
$\mu_{3,b}$ (fs)	193 fixed
$\sigma_{3,b}$ (fs)	628 fixed
τ (fs)	400 ± 6

Table 5.9: Parameters extracted from the fit to `proc9` data. The model has 2 Gaussians convolved with the exponential and a fixed combinatorial background component.

where the statistical error is the same as the one obtained on the reduced `cocktail` sample (Tab. 5.7).
The measured resolution σ_1 is worse than that estimated from the fit to MC `cocktail`, but this was expected, since the estimated errors on measured proper times are larger with respect to MC (Fig. 5.12).

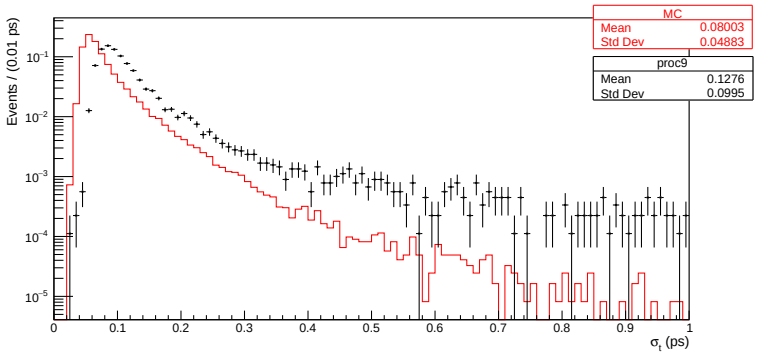


Figure 5.12: Estimated proper time error for truth-matched MC (red line) and proc9 data (black dots) candidates passing the reconstruction and belonging to the signal region.

The discrepancy of the measured D^0 lifetime is higher than that obtained for the `cocktail` but the central value is still compatible, within less than 2σ, with the PDG average [29]

$$\tau_{D^0}^{\mathrm{PDG}} = (410.1 \pm 1.5)\ \mathrm{fs}.$$

In figure 5.13 the same plot is reported in logarithmic scale in order to put in evidence the tails and the small combinatorial PDF.

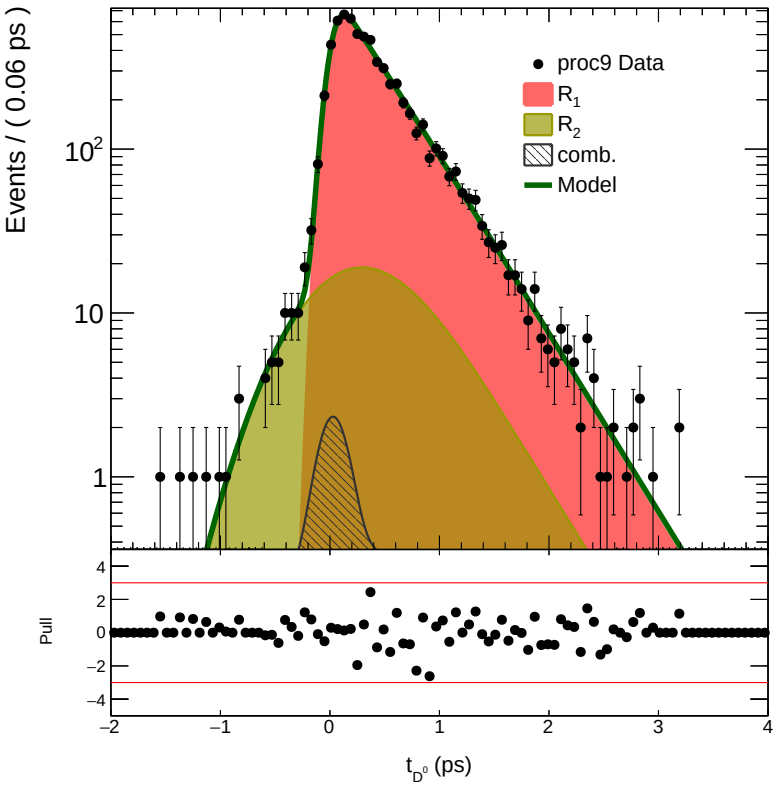


Figure 5.13: EUML fit to `proc9` data, logarithmic scale. The R_1 and R_2 components result from the convolution of the Gaussian resolution with the exponential, while the combinatorial PDF is a sum of two Gaussians and its yield is extracted from the number of candidates in the *sidebands*.

5.5 Systematic uncertainties estimation

In this section the main sources of systematic uncertainties are identified and their magnitudes evaluated. This is done, for example, by considering alternative choices to the ones we made during the analysis process and measuring the resulting differences in lifetime $\Delta[\tau_{D^0}] = \left|\tau^*_{D^0} - \tau^{'}_{D^0}\right|$. Here $\tau^*_{D^0}$ denotes the value obtained in the nominal fit shown in Tab. 5.9, while $\tau^{'}_{D^0}$ corresponds to one of the alternatives taken into consideration. The variations under study are:

1. *Selection cut on the PID of D^0 daughters*
   A tighter cut is performed — PID>0.6, against the 0.2 used for the analysis. A looser cut has not been considered since purity is required to be high.

2. *Signal model*
   We quantify the impact of the arbitrary choice of using $\mu_2 = 0$ instead of $\mu_2 = \mu_1$ for the G_2 component of the resolution function in Eq. 5.3. We also examine the impact of adding a third Gaussian to the signal PDF with $\mu_2 = \mu_3 = 0$.

3. *Signal region definition*
   The D^0 mass interval is either broadened or shortened by 1/3 of its initial width. The model used for the final fit is the 5.3 and the yield of the combinatorial background N_3 is calculated according to the new width of the mass interval defining the signal region.

Another effect than can enter in our systematics review is the error on the estimated yield of the combinatorial background. In fact the extrapolated number of combinatorial events from sidebands has its own uncertainty, which propagates once it is rescaled to the size of the signal region. For this reason the background yield is either doubled or set to zero, to evaluate the two extreme cases.
For each variation considered so far, if more than one possibility is tried (e.g. broader and shorter ΔM), the largest $\Delta[\tau_{D^0}]$ is taken for the final systematic error evaluation. The total systematic error is the sum in quadrature of the various uncertainty contributions.
As already mentioned in Sec. 5.3, the non-optimal modeling of the background with the same PDF of the signal introduces a systematic error of ~ 5 fs. In fact, the PDF that returns the correct value of τ_{D^0} for the signal, gives in turn a value of 5 fs smaller for the `cocktail`. The fact that relative fractions of twe two resolution components are not fixed and that the two Gaussian must incorporate the different resolutions measured for signal and $c\bar{c}$ backgrounds, as well as different means, leads to a difference in the measured τ_{D^0}'s. This term is also added in the sum shown in Tab. 5.11.

Beam spot and alignment calibrations

A longer discussion is required for possible biases introduced in the measurement by the presence of a residual misalignment or by an incorrect determination of the beam spot parameters. These two points are qualitatively discussed below and some quantitative considerations are also made.
The innermost layer of the vertex detector dominates the measurement of the vertices by virtue of its proximity to the IP, and hence the D^0 production and decay vertices. The first measured points of tracks originating from the IP are recorded in PXD and thus the position of the PXD ladders must be known precisely. This position can change over time, and if not corrected for, this degrades the knowledge of each hit position, and hence fitted tracks

and computed vertices. The purpose of the VXD calibration is to correct for variations in the alignment as a function of run period. While the detector calibration is extremely effective at correcting for variations in the detector position as a whole, there is an uncertainty arising from any residual lack of knowledge in the position and orientation of each ladder (r, ϕ, z coordinates etc.). VXD alignment procedure is able to determine 6 rigid body parameters: 3 for shifts (P_u, P_v, P_w for $u(r-\phi)$, v(z) and w(r) directions, respectively) and 3 rotations α, β and γ per each VXD sensor. If necessary, it can be extended to further 7 parameters for the elimination of surface deformations. An indication of the magnitude of the change in the parameters during data taking is given in Fig. 5.14, where the residual displacements for the directions r, z, $r-\phi$ are reported for the entire duration of the early Phase 3. These estimates on the residual misalignments are still preliminary since a better selection of the tracks should be made and because the calibrations have often been performed only locally (on the vertex detector) and not on the entire detector.

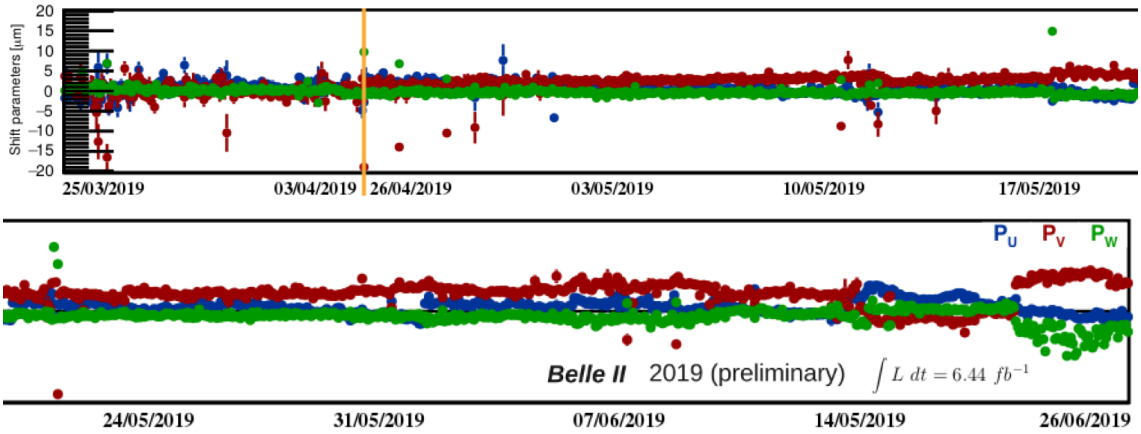
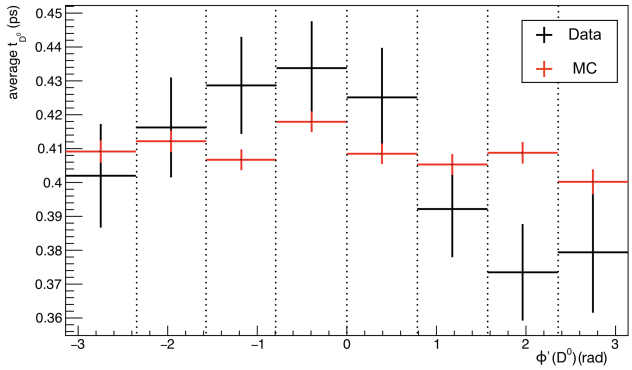
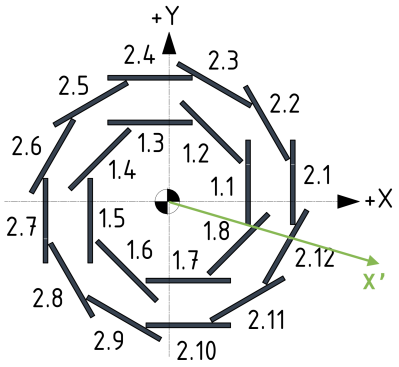


Figure 5.14: Preliminary plot showing residual displacements measured for the three shift parameters P_u, P_v, P_w with respect to the alignment calibration. The yellow line indicates the period of shut-down due to the accident in the Linac area.

The magnitude of uncertainty resulting in our measurement due to misalignment could be estimated by applying to MC different sets of alignment constants in a plausible range (for both global displacement and rotation as well as random misalignment for each ladder), and verify how this affects the measurement. This was not done due to the time available but still some quantitative indication can be obtained. Using the average proper time $\langle t_{D^0} \rangle$ as an estimator of τ_{D^0} (being the proper time distribution very similar to the exponential one) for each bin of D^0 azimuthal angle ϕ, one can actually verify that this number is not constant. This variation is made more expressive if the division into bins is made according to an azimuthal angle ϕ' aligned to the Layer 1 ladders of the PXD, in such a way that each bin can represent approximately a ladder (Fig. 5.15).
The candidates belonging to the bins (II, III, IV, V) and (I, VI, VII, VIII) are taken separately and fitted with the model with 2 Gaussians and the second mean fixed to zero. The combinatorial background is considered negligible for these two halves of the total statistics. The fit results are shown in Tab. 5.10 and the observed difference in τ_{D^0} is of $\sim$ 13 fs, corresponding to a maximum $\Delta[\tau_{D^0}]$ of 7 fs.
A larger deviation occurs instead for the averages $\langle t_{D^0} \rangle$, where the gap between the highest value at $\phi' \sim -0.5$ rad and the lower at $\phi' \sim 2$ rad is almost 60 fs. This is due to the fact that the μ_1 partially absorbs the effect of misalignment but still the slope is sensitive to the depen-



(a) Average proper time for each of the eight bins in which the range for the modified azimuthal angle ϕ' of the D^0 is divided.



(b) The 2 PXD layers in the transverse plane, with the ladder numbering for each layer. The second layer is the design one: only the 2.4 and 2.5 ladders are actually installed.

Figure 5.15: Average proper time for bins of D^0 modified azimuthal angle **(a)**, where the binning follows the position of the eight ladders in which the innermost layer of the PXD is divided **(b)**. The profile in red indicates the MC, which has a flat trend centered on the true τ_{D^0}. The black profile corresponds to data, where a modulation over the ϕ angle is visible. This can be associated to a residual misalignment of PXD, in this case in the $r-$direction.

parameter	Bins I,VI,VII,VIII	Bins II,III,IV,V
N_1	4900 ± 90	3480 ± 80
μ_1 (fs)	-11 ± 4	20 ± 5
σ_1 (fs)	92 ± 5	77 ± 5
N_2	440 ± 60	290 ± 60
μ_2 (fs)	0 fixed	0 fixed
σ_2 (fs)	620 ± 50	420 ± 50
τ (fs)	393 ± 7	406 ± 8

Table 5.10: Fit results for the candidates belonging to the two set of bins.

dence on ϕ angle. Taking fewer bins would give a larger $\Delta[\tau_{D^0}]$ but also larger statistical errors on extracted τ_{D^0}'s; with the actual division the statistics is still reasonable and the resulting $\Delta[\tau_{D^0}]$ is still a conservative estimate of the systematic error since our final measurement does average over the whole azimuthal angle acceptance.

A bias on τ_{D^0} is expected for a residual displacement of PXD ladders along the radial direction, with respect to that determined by calibrations. Such a shift in the $r-$direction implies a biased measurement of the flight distance l_{D^0} (Eq. 3.1) and thus of the proper time t_{D^0}; for example a ladder having a shorter radial distance with respect to the nominal one can overestimate the propagation time of the D^0.

Since the beam spot is used as a constraint during vertices reconstruction via adding an additional term in the χ^2 of the track fit, a mis-calibration can introduce a systematic error. The error on the beam spot location and the poorly measured beam spot size together translate into a systematic uncertainty on the reconstructed value of the proper time. The limited knowledge of the vertical position of the beam spot depends on the resolution with which the y coordinates of the interaction points are measured: as seen in Sec. 3.5, the beam spot sizes in x and y are of the same order of magnitude even though the vertical size achieved by SuperKEKB is far smaller ($\sigma_y \simeq 1.5\,\mu$m). One improvement could be to use machine parameters

for the σ_y as to have an even more stringent constraint during the fit and better calibrations of the shape of the beam spot. In relation to the IP position determination, the calibration is run-dependent and intended to follow the machine tuning during SuperKEKB operations; the effects brought by a possible bias in the IP determination or the overestimation of σ_y can be investigated on MC through a complete reconstruction with modified beam spot parameters, but again this goes beyond the scope of the present work. Finally, this brief study on systematic uncertainty is summarized in Table 5.11, where for each error source only the largest measured effect is reported.

Source	$\Delta[\tau_{D^0}]$ (fs)
Signal region definition	3.4
Signal PDF definition	0.7
Combinatorial background yield	1.1
Selection (PID)	2.0
MC signal vs. `cocktail`	4.8
Alignment/ beam spot	6.7
TOTAL	**9.2**

Table 5.11: Source of systematic uncertainties on D^0 lifetime and their estimates, reported as differences in fs with the central value obtained by the fit in Sec. 5.4.2.

The overall systematic uncertainty of 9 fs is obtained from the sum in quadrature of the six terms; the major contributions are given by the lack of a dedicate model for the background component and by the non-optimal calibrations.
By comparing this error with the statistical one (Eq. 5.4), we find that the uncertainty on our measurement is dominated by the systematic source, although the available statistics is still very low. The achievable precision on this measurements depends on the reduction of the main sources of uncertainty, related to background parametrization and calibrations.

Conclusion

In this work the feasibility of measuring the D^0 lifetime with the early data of the Belle II experiment is demonstrated. Despite the short average flight length travelled, the tracking system can resolve the production and the decay vertices of the D^0's. The new vertex detector, with its double technology and the six-layer design, where the innermost layer is at only 1.4 cm from the IP, significantly improves the resolution and the error on the measured proper time.

The dataset used for the measurement corresponds to $2.6\,\text{fb}^{-1}$ collected at the $\Upsilon(4S)$ resonance during the period between the start of the experiment (March 25^{th}) and June 3^{rd}. During that period, and until the end of the Run on July 1^{st}, the operations did not go always smooth, due to the various accidents occurred, and the maximum peak luminosity achieved during data taking was $\sim 5\times 10^{33}\,\text{cm}^{-2}\text{s}^{-1}$.

The analysis used the process $D^{*+}\to D^0(K^-\pi^+)\pi_s^+$, where the D^*'s come directly from the interaction region and not from the decays of B mesons. This choice allows to constrain the production vertex of the D^0 to lie in the beam spot region, which thus needs to be properly calibrated. Indeed the overlap region of the two beams can change over time due to the tuning of the machine parameters. My task was to provide the run-dependent beam spot parameters and contribute to make them available to the analysis tools (e.g. fitters) through the access to the database. First, the IP position was reconstructed *online* in the context of the Data Quality Monitoring (DQM) of the experiment. Then, the same information stored for the DQM was also used for the beam spot calibration, which was as useful to the experiment as it was to our analysis.

The measured D^0 lifetime is

$$\tau_{D^0} = \big(400 \pm 6(\text{stat}) \pm 9(\text{syst})\big)\,\text{fs}\,.$$

Besides giving important feedback on the detector status, this measurement is also an important first step towards many measurements in the charm sector, concerning mixing parameters and CP violation.
Among the possible sources of systematic uncertainties, the largest one is given by the residual misalignment of VXD or biased beam spot position calibration. These two points can have the same effect of shortening or lengthening the measured proper time. The final value of 9 fs is obtained by summing in quadrature this term with the other evaluated contributions, related to the choices made for the selection and the PDF for the fit to the proper time.
The uncertainty on the measurement is dominated by the systematic error, although the available statistics is still very low.
The natural next step of the analysis would be to repeat the same analysis with the full statistics collected during the Spring Run ($6.5\,\text{fb}^{-1}$). Then, the reconstruction should be performed once the new Belle II software version is released and improved calibrations are available,

both for alignment and for the beam spot. Further studies on MC can reveal all the possible effects that such non-optimal calibrations can induce on the measurement.
In addition, the analysis can be extended to other decay channels (multi-body charged and 3-body involving neutral states) but a deeper understanding of the performance of all the sub-detectors is indispensable.
I conducted this analysis from the beginning, constantly updating on its progress during the weekly calibration, tracking, performance and physics meetings. The progress of the analysis presented in this manuscript is also documented in a Belle II note, being reviewed by an internal committee for the approval of the plots and the entirety of the work and has been presented to the 2019 summer conferences.
After a few months of data acquisition, this measurement has already proved to be a useful tool to demonstrate the capabilities of the detector and could be used as a regular test of the performance of the experiment throughout its course. In order for this measurement to become competitive, instead, much higher statistics and better control of systematic effects are needed.

A PDFs

The PDFs used to build the models for the unbinned maximum likelihood fits are shown here.

⋆ **Gaussian PDF**

$$G(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{A.1}$$

⋆ **Exponential PDF**

$$\mathrm{Exp}(x|\tau) = \frac{1}{\tau}e^{-\frac{x}{\tau}} \tag{A.2}$$

⋆ **Crystal Ball Function**

$$\mathrm{CB}(x|\mu,\sigma,\alpha,n) = \begin{cases} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] & \text{if } \frac{x-\mu}{\sigma} > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n \frac{\exp\left(-\frac{\alpha^2}{2}\right)}{\left(\frac{n}{|\alpha|}-|\alpha|\pm\frac{x-\mu}{\sigma}\right)^n} & \text{otherwise} \end{cases} \tag{A.3}$$

where the + sign is chosen for $\alpha \geq 0$, the - sign otherwise.

⋆ **Chebyshev Polynomials**

$$p_0(x) = 1 \tag{A.4}$$

$$p_1(x|c_1) = 1 + c_1 x \tag{A.5}$$

$$p_2(x|c_1,c_2) = 1 + c_1 x + c_2(2x^2-1) \tag{A.6}$$

⋆ **Threshold function**

$$\mathrm{Thr}(x|q_0,a,b,c) = \theta(x-q_0)\left[1-\exp\left(-\frac{x-q_0}{c}\right)\right]\cdot\left(\frac{x}{q_0}\right)^a \cdot b\left(\frac{x}{q_0}-1\right) \tag{A.7}$$

⋆ **Johnson function**

$$J_{\mathrm{SU}}(x|\mu,\sigma,\delta,\gamma) = \frac{\delta}{\sqrt{2\pi\sigma^2}\sqrt{1+(\frac{x-\mu}{\sigma})^2}}\exp\left\{-\frac{1}{2}\left[\gamma+\delta\sinh^{-1}\left(\frac{x-\mu}{\sigma}\right)\right]\right\}, \text{ where}$$
$$\sinh^{-1}(z) = \log\left(z+\sqrt{(1+z^2)}\right) \tag{A.8}$$

⋆ **Exponential + Gaussian Convolution**

$$
\begin{aligned}
\mathrm{R}(x|\mu,\sigma,\tau) &= \mathrm{Exp}(y|\tau) * G(y|\mu,\sigma) \\
&= \frac{1}{\tau\sqrt{2\pi\sigma^2}} \int_0^\infty e^{-\frac{y}{\tau}} e^{-\frac{(x-y)^2}{2\sigma^2}} \mathrm{d}y \\
&= \frac{1}{2\tau} e^{-\frac{1}{2\tau^2}[2\tau(x-\mu)-\sigma^2]} \left[1 + \mathrm{erf}\left(\frac{\tau(x-\mu)-\sigma^2}{\sqrt{2}\sigma\tau}\right)\right], \text{ where} \\
\mathrm{erf}(x) &= \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} \mathrm{d}z
\end{aligned}
\tag{A.9}
$$

B Models

The PDFs listed in Appendix A are combined for the fit models exploited during the analysis. All the PDFs, already reported in the manuscript body, are listed here below.

⋆ **Fit to D^0 Mass - pol2**

$$M_{\rm PDF}(m|\mu,\sigma,c_1,c_2) = N_{\rm sig}\times {\rm G}(m|\mu,\sigma) + N_{\rm bkg}\times p_2(m|c_1,c_2) \tag{B.1}$$

⋆ **Fit to D^0 Mass - CB**

$$\begin{aligned} M_{\rm PDF}(m|\mu_1,\sigma_1,\mu_2,\sigma_2,\alpha,n) =& N_1\times {\rm G}(m|\mu_1,\sigma_1) + N_2\times {\rm CB}(x|\mu_2,\sigma_2,\alpha,n) + \\ &+ N_{\rm bkg} p_0(m) \end{aligned} \tag{B.2}$$

⋆ **Fit to reconstructed Q - Thr**

$$\begin{aligned} Q_{\rm PDF}(q|\mu_1,\sigma_1,\mu_2,\sigma_2,q_0,a,b,c) =& N_1\times {\rm G}(q|\mu_1,\sigma_1) + N_2\times {\rm G}(q|\mu_2,\sigma_2) + \\ &+ N_{\rm bkg}\times {\rm Thr}(q|q_0,a,b,c) \end{aligned} \tag{B.3}$$

⋆ **Fit to reconstructed Q - Johnson**

$$\begin{aligned} Q_{\rm PDF}(q|\mu_1,\sigma_1,\mu_2,\sigma_2,\delta,\gamma,c_1) =& N_1\times {\rm G}(q|\mu_1,\sigma_1) + N_2\times J_{\rm SU}(q|\mu_2,\sigma_2,\delta,\gamma) + \\ &+ N_{\rm bkg}\times p_1(q|c_1) \end{aligned} \tag{B.4}$$

⋆ **Fit to proper time -** `2GExp1122` **model**

$$T_{\rm PDF}(t|\mu_1,\sigma_1,\mu_2,\sigma_2,\tau) = N_1\times R(t|\mu_1,\sigma_1,\tau) + N_2\times R(t|\mu_2,\sigma_2,\tau) \tag{B.5}$$

⋆ **Fit to proper time -** `2GExp112233` **model**

$$\begin{aligned} T_{\rm PDF}(t|\mu_1,\sigma_1,\mu_2,\sigma_2,\mu_3,\sigma_3,\tau) =& N_1\times R(t|\mu_1,\sigma_1,\tau) + \\ &+ N_2\times R(t|\mu_2,\sigma_2,\tau) + N_3\times R(t|\mu_3,\sigma_3,\tau) \end{aligned} \tag{B.6}$$

⋆ **Fit to proper time -** `3GExp11f2Gff` **model**

$$\begin{aligned} T_{\rm PDF}(t|\mu_1,\sigma_1,\sigma_2,\tau) =& N_1\times R(t|\mu_1,\sigma_1,\tau) + N_2\times R(t|\mu_2=0,\sigma_2,\tau) \\ & N_3\times\left[\alpha\,{\rm G}(t|\mu^{\rm fixed}_{3,a},\sigma^{\rm fixed}_{3,a}) + (1-\alpha){\rm G}(t|\mu^{\rm fixed}_{3,b},\sigma^{\rm fixed}_{3,b})\right] \end{aligned} \tag{B.7}$$

C Fits to MC

In this Appendix some additional fits to MC are reported: those to the Q and M distributions (Figs. C.1–C.2) with the 80 fb^{-1} `cocktail` sample and another one to the proper time for signal events, for testing the alternative condition $\mu_1 = \mu_2$ on the model B.6.

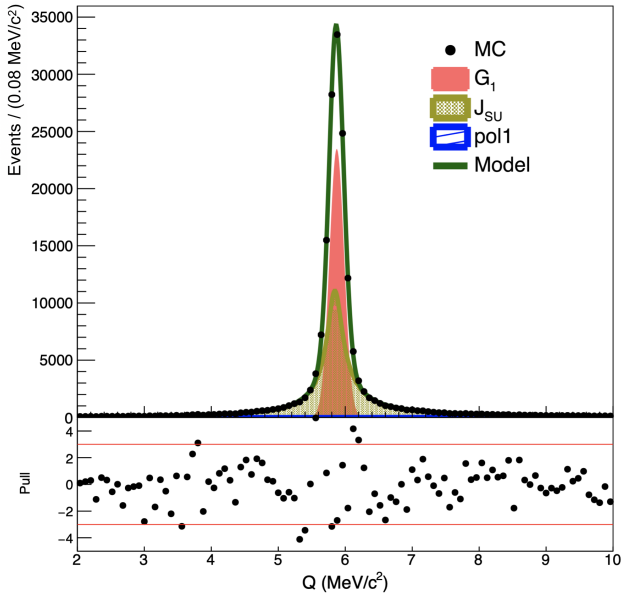


Figure C.1: Fit to the reconstructed Q for candidates of the entire `cocktail` sample.

parameter	extracted value
N_1	80000 ± 1500
$\mu_1\,(\mathrm{MeV}/c^2)$	5.8742 ± 0.0015
$\sigma_1\,(\mathrm{MeV}/c^2)$	0.109 ± 0.001
N_2	91000 ± 1500
$\mu_2\,(\mathrm{MeV}/c^2)$	5.841 ± 0.006
$\sigma_2\,(\mathrm{MeV}/c^2)$	0.112 ± 0.006
δ	0.387 ± 0.010
γ	-0.101 ± 0.012
N_{bkg}	1200 ± 500

Table C.1: Parameters extracted from the EUML fit of the reconstructed Q value.

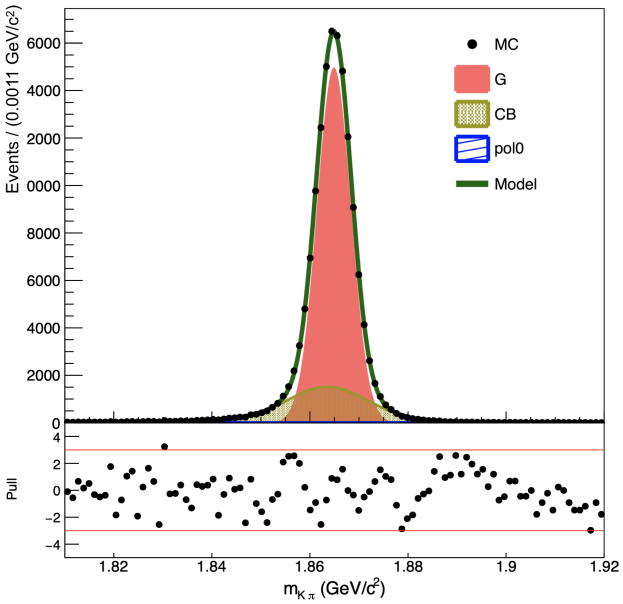


Figure C.2: Fit to the reconstructed D^0 mass for candidates of the entire `cocktail` sample.

parameter	extracted value
N_1	119000 ± 1000
$\mu_1\,(\mathrm{MeV}/c^2)$	1864.900 ± 0.012
$\sigma_1\,(\mathrm{MeV}/c^2)$	3.475 ± 0.018
N_2	30700 ± 900
$\mu_2\,(\mathrm{MeV}/c^2)$	1863.6 ± 0.1
$\sigma_2\,(\mathrm{MeV}/c^2)$	8.50 ± 0.15
α	1.503 ± 0.021
n	3
N_{bkg}	1500 ± 80

Table C.2: Parameters extracted from the EUML fit of the reconstructed D^0 mass.

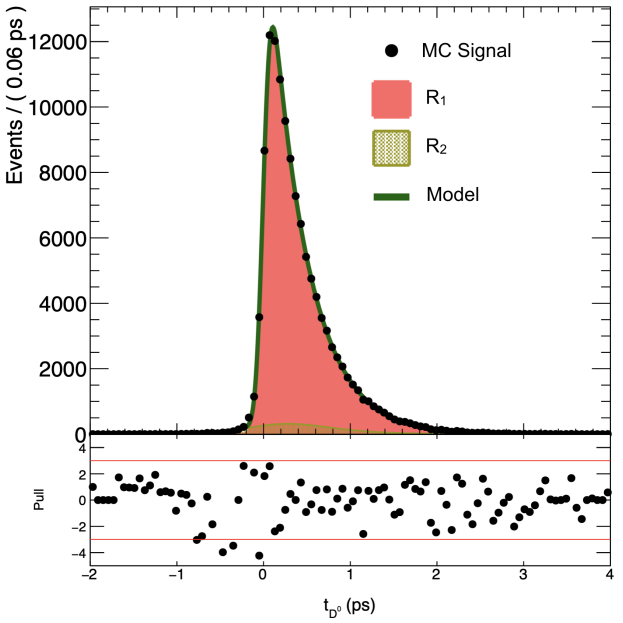


Figure C.3: Fit to the reconstructed proper time for truth-matched MC candidates belonging to the signal region. The model used is the B.6 with $\mu_1 = \mu_2$.

parameter	extracted value
N_1	115700 ± 500
$\mu_1\,(\mathrm{fs})$	0 ± 1
$\sigma_1\,(\mathrm{fs})$	69.6 ± 0.9
N_2	6800 ± 300
$\mu_2\,(\mathrm{ps})$	$= \mu_1$
$\sigma_2\,(\mathrm{ps})$	0.40 ± 0.01
$\tau\,(\mathrm{fs})$	410 ± 1

Table C.3: Parameters extracted from the EUML fit to the proper time distribution.

Bibliography

[1] I. Bigaran, J. Gargalionis, and R. R. Volkas. "A near-minimal leptoquark model for reconciling flavour anomalies and generating radiative neutrino masses" (2019). arXiv: 1906.01870 [hep-ph] (cit. on p. 7).

[2] W. Altmannshofer et al. "The Belle II Physics Book" (2018). arXiv: 1808.10567 [hep-ex] (cit. on pp. 9, 19, 26).

[3] Y. Grossman and P. Tanedo. "Just a Taste: Lectures on Flavor Physics". *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics : Anticipating the Next Discoveries in Particle Physics*. 2018, pp. 109–295. DOI: 10.1142/9789813233348_0004 (cit. on p. 11).

[4] M. Tanabashi et al. "Review of Particle Physics". *Phys. Rev. D* **98** (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: http://pdg.lbl.gov/2019/reviews/rpp2018-rev-ckm-matrix.pdf (cit. on pp. 14–16).

[5] H. Quinn. "B physics and CP violation" (2001). [ICTP Lect. Notes Ser.10,1(2002)], pp. 1–54. arXiv: hep-ph/0111177 [hep-ph] (cit. on p. 15).

[6] I. Bigi and A. I. Sanda. "CP violation" (2000). [Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.9,1(2009)]. DOI: 10.1017/CBO9780511581014 (cit. on p. 16).

[7] M. Tanabashi et al. "Review of Particle Physics". *Phys. Rev. D* **98** (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: http://pdg.lbl.gov/2018/tables/contents_tables_mesons.html (cit. on p. 17).

[8] M. A. Giorgi. "Highlights from BABAR". *Subnucl. Ser.* **41** (2005), pp. 232–364. DOI: 10.1142/9789812701794_0012 (cit. on p. 17).

[9] M. Sozzi. *Discrete Symmetries and CP Violation: From Experiment to Theory*. Oxford Graduate Texts. OUP Oxford, 2012. ISBN: 9780199655427 (cit. on pp. 18, 33).

[10] G. Goldhaber et al. "Observation in e^+e^- Annihilation of a Narrow State at 1865-MeV/c^2 Decaying to Kπ and K$\pi\pi\pi$". *Phys. Rev. Lett.* **37** (1976), pp. 255–259. DOI: 10.1103/PhysRevLett.37.255 (cit. on p. 20).

[11] I. Peruzzi et al. "Observation of a Narrow Charged State at 1876-MeV/c^2 Decaying to an Exotic Combination of K$\pi\pi$". *Phys. Rev. Lett.* **37** (1976), pp. 569–571. DOI: 10.1103/PhysRevLett.37.569 (cit. on p. 20).

[12] *Heavy Flavor Averaging Group (HFLAV)*. https://hflav-eos.web.cern.ch/hflav-eos/charm/Moriond19/results_mix_cpv.html. Accessed: 2019-08-07 (cit. on p. 21).

[13] R. Aaij et al. "Observation of CP Violation in Charm Decays". *Phys. Rev. Lett.* **122**.21 (2019), p. 211803. DOI: 10.1103/PhysRevLett.122.211803 (cit. on pp. 21, 24).

[14] S. Benson et al. "The LHCb Turbo Stream". *Journal of Physics: Conference Series* **664**.8 (Dec. 2015), p. 082004. DOI: 10.1088/1742-6596/664/8/082004 (cit. on p. 23).

[15] S. Bianco, F. L. Fabbri, et al. "A Cicerone for the physics of charm". *Riv. Nuovo Cim.* **26N7** (2003), pp. 1–200. DOI: 10.1393/ncr/i2003-10003-1 (cit. on p. 24).

[16] A. J. Bevan et al. "The Physics of the B Factories". *Eur. Phys. J.* **C74** (2014), p. 3026. DOI: 10.1140/epjc/s10052-014-3026-9 (cit. on pp. 28, 55).

[17] T. Abe et al. "Belle II Technical Design Report" (2010). arXiv: 1011.0352 [physics.ins-det] (cit. on pp. 28, 35, 37, 39–41).

[18] P. Raimondi. *Status of the SuperB Effort*. http://www.lnf.infn.it/conference/superb06/talks/raimondi1.ppt. Presentation at the 2nd Workshop on SuperB Factory, LNF-INFN, Frascati. 2006 (cit. on p. 28).

[19] Ohnishi et al. "Accelerator design at SuperKEKB". *Progress of Theoretical and Experimental Physics* (Mar. 2013). DOI: 10.1093/ptep/pts083 (cit. on p. 29).

[20] J. Lettenbichler. "Real-time Pattern Recognition in the Central Tracking Detector of the Belle II Experiment". PhD thesis. Technische Universität Wien, 2016 (cit. on p. 31).

[21] H. Tanigawa. "A study of beam background from SuperKEKB on Belle II Silicon Vertex Detector". MA thesis. University of Tokyo, 2018 (cit. on p. 31).

[22] I. Adachi et al. "Detectors for extreme luminosity: Belle II". *Nucl. Instrum. Meth.* **A907** (2018), pp. 46–59. DOI: 10.1016/j.nima.2018.03.068 (cit. on p. 33).

[23] A. Moll. "Comprehensive study of the background for the Pixel Vertex Detector at Belle II". PhD thesis. LMU Münich, 2015 (cit. on pp. 38, 39, 51).

[24] A. Martini. "The Belle II experiment: status and prospects". *Journal of Physics: Conference Series* **1137** (Jan. 2019), p. 012035. DOI: 10.1088/1742-6596/1137/1/012035 (cit. on p. 41).

[25] T. Kuhr et al. "The Belle II Core Software". *Computing and Software for Big Science* **3** (Sept. 2018). DOI: 10.1007/s41781-018-0017-9 (cit. on p. 44).

[26] D. J. Lange. "The EvtGen particle decay simulation package". *Nucl. Instrum. Meth.* **A462** (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4 (cit. on p. 44).

[27] S. Agostinelli et al. "GEANT4: A Simulation toolkit". *Nucl. Instrum. Meth.* **A506** (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8 (cit. on p. 44).

[28] R. Brun and F. Rademakers. "ROOT: An object oriented data analysis framework". *Nucl. Instrum. Meth.* **A389** (1997), pp. 81–86. DOI: 10.1016/S0168-9002(97)00048-X (cit. on p. 44).

[29] M. Tanabashi et al. "Review of Particle Physics". *Phys. Rev. D* **98** (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: http://pdg.lbl.gov/2016/tables/rpp2016-tab-mesons-charm.pdf (cit. on pp. 48, 62, 63, 69, 73, 92).

[30] M. Staric et al. "Evidence for D^0 - $\bar{D}^0$ Mixing". *Phys. Rev. Lett.* **98** (2007). [,65(2007)], p. 211803. DOI: 10.1103/PhysRevLett.98.211803 (cit. on p. 49).

[31] J. P. Lees et al. "Measurement of $D^0-\bar{D}^0$ Mixing and CP Violation in Two-Body D^0 Decays". *Phys. Rev.* **D87**.1 (2013), p. 012004. DOI: 10.1103/PhysRevD.87.012004 (cit. on p. 49).

[32] G. Casarosa. "Measurement of D^0-$\overline{D}^0$ Mixing and CP Violation at *BABAR*". *Proceedings, 5th International Workshop on Charm Physics (Charm 2012): Honolulu, Hawaii, USA, May 14-17, 2012*. 2012. arXiv: 1208.6436 [hep-ex] (cit. on pp. 49, 70).

[33] J. F. Krohn et al. "Global Decay Chain Vertex Fitting at B-Factories" (2019). arXiv: 1901.11198 [hep-ex] (cit. on p. 53).

[34] W. D. Hulsbergen. "Decay chain fitting with a Kalman filter". *Nucl. Instrum. Meth.* **A552** (2005), pp. 566–575. DOI: 10.1016/j.nima.2005.06.078 (cit. on p. 53).

[35] J. P. Lees and others (*BABAR* Collaboration). "Measurement of the $D^*(2010)^+$ Meson Width and the $D^*(2010)^+ - D^0$ Mass Difference". *Phys. Rev. Lett.* **111** (11 Jan. 2013), p. 111801. DOI: 10.1103/PhysRevLett.111.111801 (cit. on p. 70).