Development of proper-time resolution function for improved measurement of time evolution of B mesons in the Belle II experiment

ZHANG Tingyu

Department of Physics
the University of Tokyo

This dissertation is submitted for the degree of
Master of Science

August 2020

I would like to dedicate this thesis to my loving parents ...

Acknowledgements

First of all, I would like to express my gratitude to my supervisor Prof. Hiroaki Aihara for giving me the opportunity to participate the Belle II experiment and the excellent supervision throughout my master's study. I am also grateful to Prof. Yoshiyuki Onuki for his patient guidance and constant support from the enthusiasm and rich knowledge in the field.

I wish to thank all the experts in the Japan ICPV group and TDCPV group in the Belle II collaboration, in particular Dr. Takeo Higuchi, Dr. Humair Thibaud, Dr. Yosuke Yusa, Dr. Chiara La Licata, Hikaru Tanigawa, Dr. Vladimir Chekelian and Dr. Alessandro Gaz for their invaluable help and inputs to the analysis work in this thesis.

I also wish to thank all the members in Aihara-Yokoyama group in the University of Tokyo for the wonderful time spent together. I would like to pay my special regards to Kun Wan for the countless number of interesting and rewarding discussions, to Johnny Alejandro Mora Grimaldo for his expert skills in computer science, to Dr. Yifan Jin for the valuable advices in career building, to Prof. Masashi Yokoyama for the insights he gave from the broad experience in Belle during the group meetings and to Kuniko Kono for her patience and efficiency to support our group.

Last but not least, I would like to express deepest appreciation to my parents, who have supported and encouraged me as always.

Abstract

In the Standard Model, the single irreducible complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix causes the effects of Charge-Parity (CP) violation. The Belle and BaBar Collaborations have experimentally confirmed the existence of CP violation in the B meson system. Starting from 2009, as the next-generation B-factory experiment located at the SuperKEKB accelerator complex in Tsukuba, Japan, Belle II aims to reach 40 times larger luminosity and to collect 50 times more data than its predecessor for measurements of the Standard Model parameters with better precision and searching for new physics phenomena over the 8 years operation period.

In the B factory, most electrons and positrons collide at center-of-mass energy of $\Upsilon(4s)$ resonance to produce clean B-$\bar{B}$ pairs with no fragmentation particles. The design of asymmetric beam energies provides a boost to the center-of-mass in lab frame, thus allowing time dependent decay rate and CP violation measurements to be performed. In the time dependent measurements, the proper-time resolution function is a necessary piece that refers to the resolution between measured and true evolution of time.

At the very early stage of physics run in Belle II starting from 2019, several prompt attempts to the parameterization of proper-time resolution function has been made. I took part in the research project to develop an artificial proper-time resolution function that is comparable with Belle. In this thesis, we present the development of such optimized event-by-event model being constructed and validated by lifetime fits on Monte Carlo simulations and 8.7 fb^{-1} experimental data collected in Belle II. The goal of such study is to provide the necessary tools for improved measurements of time evolution of B mesons.

Table of contents

Chapter 1

Introduction

The Standard Model (SM) is the best tested theory of fundamental particle physics at the current level of experiment precision. While tremendous success was made in explaining the fundamental particles and their interactions excluding gravity, there are still fundamental questions that cannot be explained. One of the unsolved problems is the matter-antimatter imbalance at the cosmological scale [1]. The existence of Charge-Parity (CP) violation is a necessary condition for the evolution of the matter dominated universe now. While mechanism of CP violation is allowed within SM in the quark sector originating from the complex phase of Cabibbo-Kobayashi-Maskawa (CKM) matrix [2], the observed magnitude is many orders too small to explain the dominance in universe. Hence new sources of CP violation must be undiscovered yet.

Significantly upgraded from Belle, the Belle II experiment is the next generation B factory that provides possibilities to improve the current measurement of SM parameters and search for new physics (NP) at the intensity frontier. The Belle II project started in 2016 and just began its physics measurement from early 2019. In other words, Belle II is still operating at its very early phase. In the B factory, one of the methods to precisely measure CP violation is through the time evolution of B mesons, also known as the time-dependent CP violation (TDCPV) analysis. In this analysis, the proper-time Δt resolution function plays an essential role as the connection between physics model and measured decay time of B mesons:

$$P(\Delta t) = \int d(\Delta t') P_{sig}(\Delta t') R(\Delta t - \Delta t'). \tag{1.1}$$

Here R is the Δt resolution function, P is the measured and P_{sig} is the probability density function of proper-time difference based on physics model. At the current stage of Belle II experiment, several prompt attempts have been made to model the resolution function,

including analytical function constituting sum of three Gaussian distributions used in the measurement of B^0 lifetime in the 2019 Belle II dataset [3], and advanced BaBar-style analytical resolution function with reweighting techniques [4]. In Belle, the resolution function [5] is more artificial and complicated, and has proven to have good performance over the whole operation period. This motivates the study to develop a Belle-style Δt resolution function that is compatiable to the Belle II environment. The goal is to provide a reliable resolution function that not only is analytically consistent with Monte Carlo simulations, but also utilizes parameters and components from different sources in a physics explainable way, to achieve the more precise measurements of the time evolutions in Belle II.

The decay $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ is a flavor specific mode where the initial flavor of B meson is known from final state particles, making it extremely useful for the TDCPV analysis. In addition, the unambiguous reconstruction the kinematics and the relatively large branching fraction $B(B^0 \to D^-\pi^+) = 2.68 \times 10^{-3}$ makes it the perfect candidate for the study of resolution function. This thesis presents the development of the proper-time resolution function as follows:

In Chapter 2, a theoretical background of CP violation in the Standard Model and how time evolutions are measured in the Belle II experiment is given.

In Chapter 3, the SuperKEKB accelerator and the Belle II detectors are described, with a focus on the upgrades from Belle that lead to measurements with better precision. The commissioning phases of Belle II project are also introduced.

In Chapter 4, the essential procedures to time-dependent measurements in data processing that lead to the reconstruction of B events are presented, including track reconstruction, vertex reconstruction and selection criteria. The optimal set of fitting options is studied and chosen that result in better resolution compared with Belle.

In Chapter 5, the modelling and parameterization of Δt resolution function is presented in detail. The main elements are signal side resolution dominated by detector resolution, tag side resolution that is a combinational effect of detector resolution and non-primary tracks smearing, and the kinematic smearing resolution. Modules dedicated to the separation of effects from different sources are developed and studied.

In Chapter 6, the performance of the developed Δt resolution function is tested by performing lifetime fits to Monte Carlo samples and experimental data.

In Chapter 7, the Δt resolution function is further validated by linearity test and comparison between parameters in MC and experimental data. It follows by the discussions of future improvements and systematic uncertainties.

Chapter 2

Phenomenology of B physics

2.1 CP violation in the Standard Model

As one of the most successful theory in modern particle physics, the Standard Model incorporates fundamental particles and fundamental forces in the universe and has made great predictions to experiments with great accuracy.

The full Standard Model Lagrangian is in the form of:

$$\mathscr{L}_{SM} = \mathscr{L}_{kinetic} + \mathscr{L}_{Higgs} + \mathscr{L}_{Yukawa}. \tag{2.1}$$

Consider the column vectors of left-handed up-type quarks $U' \equiv (u\ c\ t)^T$ and left-handed down-type quarks $D' \equiv (d\ s\ b)^T$ in the weak interaction eigenstate basis, the charged current part derived from kinetic term is given by:

$$\mathscr{L}_{CC} = \frac{g}{\sqrt{2}} \bar{U}'_L \gamma^\mu D'_L W^+_\mu + h.c., \tag{2.2}$$

where L indicates left-handed chirality. The quark mass is acquired through their interactions with Higgs bosons in the Yukawa term of Lagrangian

$$\mathscr{L}_{mass} = \bar{U}'_L M'_U U'_R + \bar{D}'_L M'_D D'_R + h.c., \tag{2.3}$$

where M'_U and M'_D are 3×3 matrices, which in general are not diagonal in flavor basis. To diagonalize the mass matrices, it is possible to introduce unitary matrices $V_{UL}, V_{DL}, V_{UR}, V_{DR}$

such that

$$U'_L = V_{UL}U_L, \quad D'_L = V_{DL}D_L,$$
$$U'_R = V_{UR}U_R \quad D'_R = V_{DR}D_R.$$

Here, U_L, D_L, U_R, D_R are vectors of hadrons in mass basis. With this, one can rewrite the charged current term in mass basis

$$\mathscr{L}_{CC} = \frac{g}{\sqrt{2}}\bar{U_L}\gamma^\mu D_L(V_{UL}^\dagger V_{DL})W_\mu^+ + h.c.. \tag{2.4}$$

Thus, in mass basis, up-type quarks couple with down-type quarks through a unitary matrix

$$V_{CKM} \equiv V_{UL}^\dagger V_{DL}.$$

V_{CKM} is also known as the Cabibbo-Kobayashi-Maskawa matrix:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{2.5}$$

If we examine the charged current coupling in mass basis again and impose CP transformation:

$$\mathscr{L}_{CC} = \frac{g}{\sqrt{2}}\bar{U_L}\gamma^\mu D_L V_{CKM} W_\mu^+ + \frac{g}{\sqrt{2}}\bar{D_L}\gamma^\mu U_L V_{CKM}^* W_\mu^-, \tag{2.6}$$
$$\mathscr{L}_{CC}^{CP} = \frac{g}{\sqrt{2}}\bar{D_L}\gamma^\mu U_L V_{CKM} W_\mu^- + \frac{g}{\sqrt{2}}\bar{U_L}\gamma^\mu D_L V_{CKM}^* W_\mu^+. \tag{2.7}$$

It is observed that if all elements of V_{CKM} are real in some basis, the Lagrangian is unchanged, thus CP-symmetry is conserved. In other words, the complex nature of the CKM matrix is the origin of CP violation.

2.2 CKM matrix and the Unitary Triangle

We start by counting the number of free parameters for the CKM matrix introduced in the previous section. For a general $n \times n$ complex matrix, there are $2n^2$ real parameters. Unitarity provides constraints on n^2 parameters. The phases of quarks can be rotated freely and $2n-1$ relative phases can be removed. The remaining $(n-1)^2$ parameters can be divided to $n(n-1)/2$ Euler angles and $(n-1)(n-2)/2$ complex phases. In the Standard Model

with $n = 3$ (three generations), we have 3 Euler angles and 1 complex phase. The standard parameterization of CKM matrix is given by:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & -c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \tag{2.8}$$

where $s_{ij} = sin\theta_{ij}, c_{ij} = cos\theta_{ij}$. $\theta_{12}, \theta_{23}, \theta_{13}$ are the Euler angles and δ_{13} refers to the complex phase.

One popular way to the parameterization of CKM matrix is the Wolfenstein parameterization [6]. It approximates by expanding CKM matrix as power series of $\lambda = sin\theta_C \cong 0.22$ where θ_C is the Cabibbo angle. To order λ^3 it is:

$$\begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4). \tag{2.9}$$

The three Euler angles and one complex phase are replaced by λ, A, ρ and η where

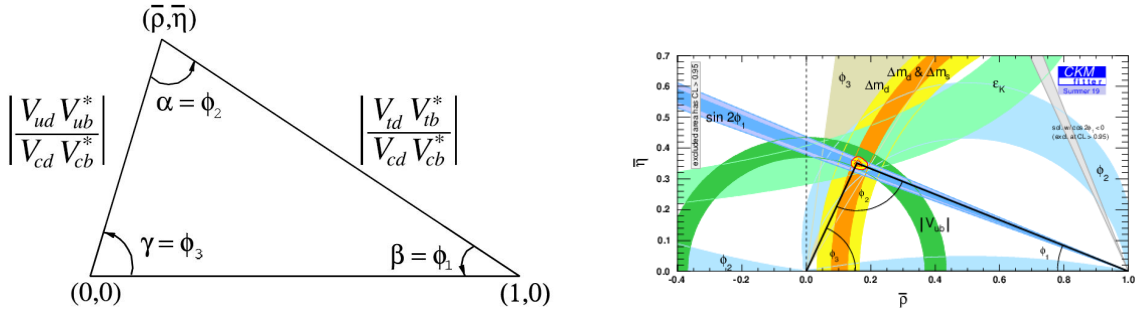
$$\begin{aligned} \lambda &= s_{12}, \\ A\lambda^2 &= s_{23}, \\ A\lambda^3(\rho - i\eta) &= s_{13}e^{-i\delta_{13}}. \end{aligned}$$

The best determination of the Wolfenstein parameters is [7]:

$$\lambda = 0.2257^{+0.0009}_{-0.0010},\ A = 0.814^{+0.021}_{-0.022},\ \rho = 0.135^{+0.031}_{-0.016} \text{ and } \eta = 0.349^{+0.015}_{-0.017}.$$

The unitarity of the CKM matrix $V^\dagger V = VV^\dagger = \mathbb{1}$ for off-diagonal elements leads to:

$$V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0, \tag{2.11a}$$
$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0, \tag{2.11b}$$
$$V_{td}V_{cd}^* + V_{ts}V_{cs}^* + V_{tb}V_{cb}^* = 0, \tag{2.11c}$$
$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0, \tag{2.11d}$$
$$V_{td}V_{ud}^* + V_{ts}V_{us}^* + V_{tb}V_{ub}^* = 0, \tag{2.11e}$$
$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{2.11f}$$



(a) (b)

Fig. 2.1 the Unitary Triangle in the complex ρ-η plane.

These relations can be visualized as closed triangles in the complex plane with combinations of CKM elements. The triangles in (2.11a) and (2.11b) are the flattest, with two sides being $O(\lambda)$ and one side being $O(\lambda^5)$. The triangles in (2.11c) and (2.11d) are also flat, with two sides being $O(\lambda^2)$ and one side being $O(\lambda^4)$. The remaining triangles in (2.11e) and (2.11f)) are the standard ones as all sides are $O(\lambda^3)$, at comparable lengths. The triangle derived from the last equation is usually referred to as the "Unitary Triangle" (Fig 2.1). The three angles are defined as

$$\phi_1 \equiv \pi - arg(\frac{-V_{tb}V_{td}^*}{-V_{cb}V_{cd}^*}), \tag{2.12}$$

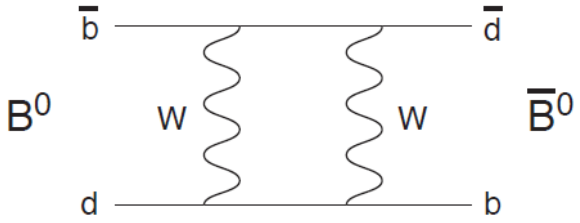
$$\phi_2 \equiv \pi - arg(\frac{V_{tb}V_{td}^*}{-V_{ub}V_{ud}^*}), \tag{2.13}$$

$$\phi_3 \equiv \pi - arg(\frac{V_{ub}V_{ud}^*}{-V_{cb}V_{cd}^*}). \tag{2.14}$$

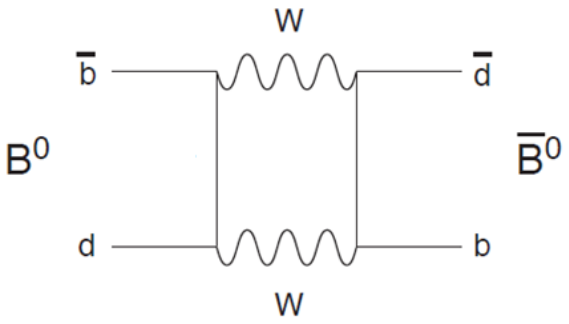
With the use of unitarity of CKM matrix, it can be shown that a quantity

$$J \equiv Im(V_{\alpha i}^* V_{\alpha j} V_{\beta i} V_{\beta j}^*) \tag{2.15}$$

is equivalent for all combinations of $\alpha \neq \beta$ and $i \neq j$. This quantity is known as the Jarlskog invariant, which is twice the area of the corresponding unitary triangle. As a result, all triangles from (2.11) have the same area. In fact, the area of the unitary triangles can vanish only when all elements in the CKM matrix are real, making it an important indicator of CP violation in the KM sector.



(a)



(b)

Fig. 2.2 Lowest order box diagrams responsible for B^0-$\bar{B}^0$ mixing.

2.3 Mixing and CP violation in the neutral B meson system

An important phenomenon in the B physics is the mixing of neutral B mesons through second order weak interactions described in the box diagrams in Fig 2.2.

Let state $|B^0\rangle$ and $|\bar{B}^0\rangle$ be the neutral B meson flavor eigenstates, an arbitrary time-dependent superposition of the states can be written as:

$$\psi(t) = a(t)\,|B^0\rangle + b(t)\,|\bar{B}^0\rangle = \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}. \tag{2.16}$$

The wavefunction obeys the time-dependent Schrodinger equation:

$$i\frac{\partial \psi}{\partial t} = \mathbf{H}\psi, \tag{2.17}$$

where in the basis of $(B^0, \bar{B}^0)$, $\mathbf{H}$ can be written as the a sum of the mass term $\mathbf{M}$ and the decay term $\mathbf{\Gamma}$:

$$\mathbf{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}. \tag{2.18}$$

The Hermitian properties of $\mathbf{M}$ and $\mathbf{\Gamma}$ implied that

$$M_{21} = M_{12}^* \quad \text{and} \quad \Gamma_{21} = \Gamma_{12}^*, \tag{2.19}$$

In addition, we assume CPT symmetry, requiring that

$$M_{11} = M_{22} \quad \text{and} \quad \Gamma_{11} = \Gamma_{22}. \tag{2.20}$$

With the constraints above, one can obtain the eigenvalues of $\mathbf{H}$ to be

$$\mu_H = M_{11} - \frac{i}{2}\Gamma_{11} + \frac{1}{2}(\Delta m - \frac{i}{2}\Delta\Gamma), \tag{2.21}$$

$$\mu_L = M_{11} - \frac{i}{2}\Gamma_{11} - \frac{1}{2}(\Delta m - \frac{i}{2}\Delta\Gamma), \tag{2.22}$$

where ΔM and $\Delta\Gamma$ satisfy

$$(\Delta m)^2 - (\frac{\Delta\Gamma}{2})^2 = 4|M_{12}|^2 - |\Gamma_{12}|^2, \tag{2.23}$$

$$\Delta m \cdot \Delta\Gamma = 4\mathrm{Re}(M_{12}\Gamma_{12}^*) \tag{2.24}$$

Note that here we choose $\Delta m > 0$ so that labels H and L stand for "heavy" and "light" respectively.

The mass eigenstates of the Hamiltonian are given by:

$$|B_L\rangle = p\,|B^0\rangle + q\,|\bar{B}^0\rangle\,, \tag{2.25a}$$

$$|B_H\rangle = p\,|B^0\rangle - q\,|\bar{B}^0\rangle\,, \tag{2.25b}$$

where p and q satisfy:

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{2.26}$$

Next, we study the time evolution of the flavor eigenstates of neutral B mesons. One can rewrite (2.25) to

$$|B^0\rangle = \frac{1}{2p}(|B_H\rangle + |B_L\rangle\,, \tag{2.27a}$$

$$|\bar{B}^0\rangle = \frac{1}{2q}(|B_H\rangle - |B_L\rangle\,. \tag{2.27b}$$

We know the time evolution of mass eigenstates $|B_H\rangle$ and $|B_L\rangle$ from the Schrodinger equations:

$$|B_H(t)\rangle = e^{-im_H t - \frac{1}{2}\Gamma_H t}\,|B_H(0)\rangle\,, \tag{2.28a}$$

$$|B_L(t)\rangle = e^{-im_L t - \frac{1}{2}\Gamma_L t}\,|B_L(0)\rangle\,. \tag{2.28b}$$

Thus, the time evolution of $|B^0\rangle$ and $|\bar{B}^0\rangle$ can be calculated as

$$|B^0(t)\rangle = g_+(t)\,|B^0\rangle + (\frac{q}{p})g_-(t)\,|\bar{B}^0\rangle\,, \tag{2.29a}$$

$$|\bar{B}^0(t)\rangle = g_-(t)(\frac{p}{q})\,|B^0\rangle + g_+(t)\,|\bar{B}^0\rangle\,, \tag{2.29b}$$

where we have defined that

$$g_\pm(t) \equiv \frac{1}{2}e^{-iMt}(e^{-im_H t - \frac{1}{2}\Gamma_H t} \pm\ e^{-im_L t - \frac{1}{2}\Gamma_L t}). \tag{2.30}$$

Consider the processes where B^0 or $\bar{B}^0$ decays to some final CP eigenstate f_{CP}. The decay rates can be written as:

$$A_{f_{CP}} \equiv \left\langle f_{CP}\right|\, H\,\left|\, B^0\right\rangle, \quad \bar{A}_{f_{CP}} \equiv \left\langle f_{CP}\right|\, H\left|\bar{B}^0\right\rangle. \tag{2.31}$$

Combining them with the time evolution of $|B^0\rangle$ and $|\bar{B}^0\rangle$ in (2.29), it is straightforward to work out the expressions for the neutral B meson decay rates, also known as the **master equations** [8]:

$$\begin{aligned}
&\Gamma\left(B^0 \to f_{CP};t\right) \propto \left|\left\langle f_{CP}\right| H \left|B^0\right\rangle\right|^2 \\
&= e^{-\Gamma t}\left|A_{f_{CP}}\right|^2 \left[\frac{1+\left|\lambda_{f_{CP}}\right|^2}{2} cosh\left(\frac{\Delta\Gamma t}{2}\right) - Re(\lambda_{f_{CP}}) sinh(\frac{\Delta\Gamma t}{2})\right] \\
&\quad + e^{-\Gamma t}\left|A_{f_{CP}}\right|^2 \left[\frac{1-\left|\lambda_{f_{CP}}\right|^2}{2} cos(\Delta m t) - Im(\lambda_{f_{CP}}) sin(\Delta m t)\right],
\end{aligned} \tag{2.32a}$$

$$\begin{aligned}
&\Gamma\left(\bar{B}^0 \to f_{CP};t\right) \propto \left|\left\langle f_{CP}\right| H \left|\bar{B}^0\right\rangle\right|^2 \\
&= e^{-\Gamma t}\left|A_{f_{CP}}\right|^2 \left|\frac{p}{q}\right|^2 \left[\frac{1+\left|\lambda_{f_{CP}}\right|^2}{2} cosh\left(\frac{\Delta\Gamma t}{2}\right) - Re(\lambda_{f_{CP}}) sinh(\frac{\Delta\Gamma t}{2})\right] \\
&\quad - e^{-\Gamma t}\left|A_{f_{CP}}\right|^2 \left|\frac{p}{q}\right|^2 \left[\frac{1-\left|\lambda_{f_{CP}}\right|^2}{2} cos(\Delta m t) - Im(\lambda_{f_{CP}}) sin(\Delta m t)\right],
\end{aligned} \tag{2.32b}$$

where $\Gamma \equiv (\Gamma_H + \Gamma_L)/2$ and $\lambda_{f_{CP}} \equiv \frac{q}{p}\frac{\bar{A}_{f_{CP}}}{A_{f_{CP}}}$.

The observation of CP violation effects can be classified into three different types:

1) **CP violation through decay**

This implies that decay rate of a B to final state is different from the anti-B to the CP-conjugate final state. This type of CP violation occurs if and only if

$$\left|\bar{A}_{\bar{f}}/A_f\right| \neq 1.$$

An example of such type of CP violation is decay $B^0 \to K^+\pi^-$, where the CP violating observable is observed to be

$$A_{CP} = \frac{\Gamma(\bar{B}^0 \to K^+\pi^-) - \Gamma(B^0 \to K^-\pi^+)}{\Gamma(\bar{B}^0 \to K^+\pi^-) + \Gamma(B^0 \to K^-\pi^+)} < 0.$$

2) **CP violation through mixing**

This implies that decay rate of B^0 to $\bar{B}^0$ is different for the reverse process. Experimental searches are usually conducted in the semi-leptonic decays for neutral B mesons. By comparing numbers of equal charge final states, we can define the CP violating observable in mixing:

$$A_{CP} = \frac{N_{++} - N_{--}}{N_{++} + N_{--}} = \frac{(p/q)^2 - (q/p)^2}{(p/q)^2 + (q/p)^2}.$$

We find that it is violated when $|p/q| \neq 1$. However, in the case of $B^0 - \bar{B}^0$ system $|p/q| \approx 1$ is good approximation within experimental accuracy and theoretical expectation.

3) **CP violation through mixing-decay interference** This type of CP violation is usually measured when B^0 and $\bar{B}^0$ decay to the same final CP eigenstate. We can define CP asymmetry as:

$$A_{CP}(t) = \frac{\Gamma(\bar{B}^0 \to f_{CP}) - \Gamma(B^0 \to f_{CP})}{\Gamma(\bar{B}^0 \to f_{CP}) + \Gamma(B^0 \to f_{CP})}.$$

With the use of master equations in (2.32), along with the approximation $|p/q| \approx 1$ and $\Delta\Gamma/\Gamma \ll 1$, we rewrite it in the form of

$$A_{CP}(t) = S_{CP} sin(\Delta mt) + C_{CP} cos(\Delta mt),$$

where

$$S_{CP} \equiv \frac{2Im(\lambda_{f_{CP}})}{1 + |\lambda_{f_{CP}}|^2}, \quad C_{CP} \equiv \frac{|\lambda_{f_{CP}}|^2 - 1}{|\lambda_{f_{CP}}|^2 + 1}.$$

S_{CP} is often regarded as **indirect** CP violation while C_{CP} is regarded as **direct** CP violation observables.

The classification of different types of CP violation is illustrated in Fig. 2.3.

2.4 Measurements of time evolution in Belle II

The experimental designs to measure time evolution of B mesons in Belle II are inherited from its predecessor [10]. In the B factories, the e^+e^- collider is running on the resonance energy of $\Upsilon(4S)$. The branching fraction for $\Upsilon(4S)$ to decay to $B\bar{B}$ pairs is more than 96%, producing a rather clean sample of neutral $B^0\bar{B}^0$ pairs for the CP violation analysis. The $B^0\bar{B}^0$ pairs are in a coherent quantum state, meaning that at the same certain time, if one meson is detected to be B^0, the other must be $\bar{B}^0$. With the considerations above, the $B^0\bar{B}^0$ system at $t = 0$ can be expressed as

$$|B^0\bar{B}^0(0,0)\rangle = \frac{1}{\sqrt{2}}(\ |B^0\bar{B}^0\rangle + \ |\bar{B}^0B^0\rangle). \tag{2.33}$$

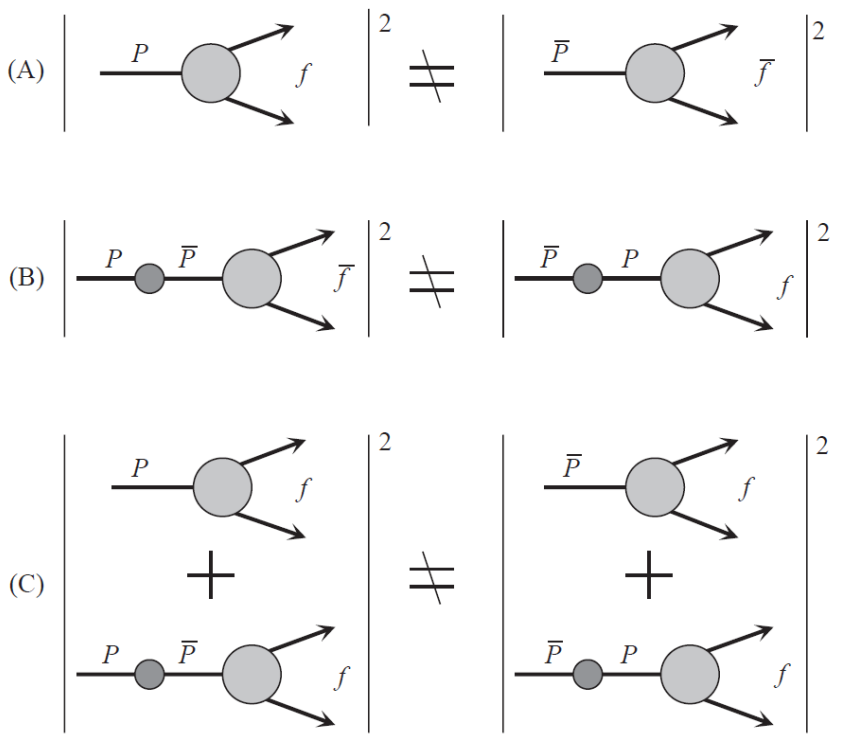


Fig. 2.3 Illustration of three types of CP violation in observation for general neutral meson P [9]. (A) CP violation through decay without mixing. (B) CP violation through mixing. (C) CP violation through interference of decay chains for $P \to f$ and $\bar{P} \to f$ decays.

Using (2.27), the time evolution of the system can be expressed as

$$
\begin{aligned}
|B^0\bar{B}^0(t_1,t_2)\rangle &= \frac{1}{\sqrt{2}}\big(\ |B^0(t_1)\bar{B}^0(t_2)\rangle + \ |\bar{B}^0(t_1)B^0(t_2)\rangle\big) \\
&= -\frac{1}{2\sqrt{2}} e^{(-iM+\frac{1}{2}\Gamma)(t_1+t_2)} \left[e^{(i\Delta m-\frac{1}{2}\Delta\Gamma)\frac{t_1+t_2}{2}} - e^{(-i\Delta m+\frac{1}{2}\Delta\Gamma)\frac{t_1+t_2}{2}} \right] \left(\frac{p}{q}|B^0B^0\rangle - \frac{q}{p}|\bar{B}^0\bar{B}^0\rangle \right) \\
&\quad + \frac{1}{2\sqrt{2}} e^{(-iM+\frac{1}{2}\Gamma)(t_1+t_2)} \left[e^{(i\Delta m-\frac{1}{2}\Delta\Gamma)\frac{t_1+t_2}{2}} - e^{(-i\Delta m+\frac{1}{2}\Delta\Gamma)\frac{t_1+t_2}{2}} \right] \left(|B^0\bar{B}^0\rangle - |\bar{B}^0B^0\rangle \right).
\end{aligned}
\tag{2.34}
$$

Now let us consider the more realistic case in the measurement where one neutral B meson decays to the CP eigenstate such as $J/\psi K_s$, and its pair partner $\bar{B}$ decays to some flavor specific final state X (for example semi-leptonic decays). We label the B meson decaying to CP eigenstate B_{sig} as it is the signal decay mode of interest. The other B meson is labeled B_{tag} as its decay is used to determine the flavor at time of decay. Their proper time of decay

are t_{sig} and t_{tag} respectively. The decay rate of such process is:

$$\begin{aligned}
&\Gamma\left(B^0 \to f_{CP}, \bar{B}^0 \to f_X; t_{sig}, t_{tag}\right) \propto \left|\left\langle f_{CP} f_X \right| H \left| B^0 \bar{B}^0 \right\rangle\right|^2 \\
&= \frac{1}{2} e^{-\Gamma(t_{sig}+t_{tag})} \left|A_{f_{CP}}\right|^2 \left[\frac{1+\left|\lambda_{f_{CP}}\right|^2}{2} cosh\left(\frac{\Delta\Gamma\Delta t}{2}\right) - Re(\lambda_{f_{CP}}) sinh(\frac{\Delta\Gamma\Delta t}{2})\right] \\
&\quad + \frac{1}{2} e^{-\Gamma(t_{sig}+t_{tag})} \left|A_{f_{CP}}\right|^2 \left[\frac{1-\left|\lambda_{f_{CP}}\right|^2}{2} cos\left(\Delta m \Delta t\right) - Im(\lambda_{f_{CP}}) sin(\Delta m \Delta t)\right],
\end{aligned} \tag{2.35a}$$

$$\begin{aligned}
&\Gamma\left(\bar{B}^0 \to f_{CP}, B^0 \to f_X; t_{sig}, t_{tag}\right) \propto \left|\left\langle f_X f_{CP} \right| H \left| B^0 \bar{B}^0 \right\rangle\right|^2 \\
&= \frac{1}{2} e^{-\Gamma(t_{sig}+t_{tag})} \left|A_{f_{CP}}\right|^2 \left[\frac{1+\left|\lambda_{f_{CP}}\right|^2}{2} cosh\left(\frac{\Delta\Gamma\Delta t}{2}\right) - Re(\lambda_{f_{CP}}) sinh(\frac{\Delta\Gamma\Delta t}{2})\right] \\
&\quad - \frac{1}{2} e^{-\Gamma(t_{sig}+t_{tag})} \left|A_{f_{CP}}\right|^2 \left[\frac{1-\left|\lambda_{f_{CP}}\right|^2}{2} cos\left(\Delta m \Delta t\right) - Im(\lambda_{f_{CP}}) sin(\Delta m \Delta t)\right].
\end{aligned} \tag{2.35b}$$

Here we have defined that $\Delta t = t_{sig} - t_{tag}$. In real experiment, at the rest frame of $\Upsilon(4S)$, B mesons have almost 0 momentum so that the decay point is very close to production point. Thus, the measurement of proper time difference Δt is required to be at the scale of B lifetime ($\sim$1.5 ps). This is made possible by the design of energy-asymmetric e^+e^- collider. It gives a boost of $(\beta\gamma)_{\Upsilon(4S)} \approx 0.28$ along beam direction z in the lab frame so that B mesons fly for an average 125 μm before decaying. Therefore, the proper time difference Δt is measured by using

$$\Delta t \approx \frac{\Delta z}{c \cdot (\beta\gamma)_{\Upsilon(4S)}}. \tag{2.36}$$

A topological illustration is shown in Fig. 2.4. Furthermore, to measure proper lifetime, the Δt distribution of probability density is

$$P\left(\Delta t; \tau_{B^0}\right) = \frac{1}{2\tau_{B^0}} e^{-|\Delta t|/\tau_{B^0}}. \tag{2.37}$$

To measure CP violation, the Δt distribution becomes

$$P\left(\Delta t; \xi, S_{CP}, C_{CP}\right) = \frac{1}{4\tau_{B^0}} \left\{1 + \xi \left[C_{CP} cos(\Delta m \Delta t) + S_{CP} sin(\Delta m \Delta t)\right]\right\}, \tag{2.38}$$

where $\xi = +1(-1)$ in cases where B_{tag} is tagged as $B^0(\bar{B}^0)$.

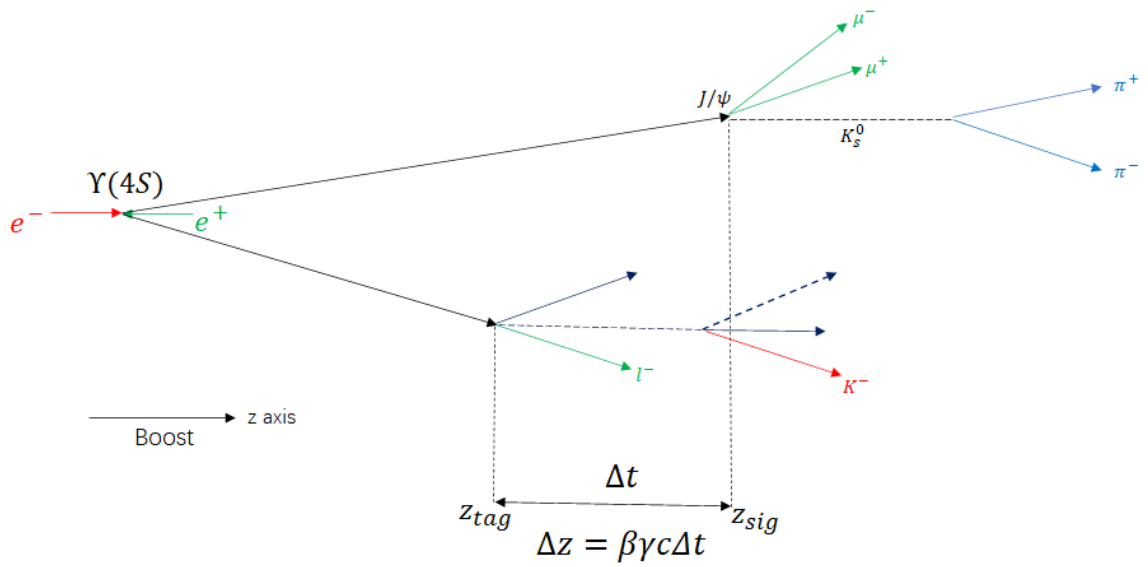


Fig. 2.4 Topological illustration of time dependent CP violation analysis to calculate proper decay time difference Δt.

Chapter 3

SuperKEKB and Belle II

3.1 Overview of SuperKEKB

SuperKEKB is a particle collider located at KEK, with electrons and positrons colliding at the energy close to Υ resonances. The layout of SuperKEKB main ring is shown in Fig. 3.1. Most data are taken at the $\Upsilon(4S)$ resonance where large numbers of B meson pairs are produced, making it perfect to study exciting physics of B mesons. SuperKEKB leads its way in the intensity frontier, with a design luminosity of 8×10^{35} cm^{-2}s^{-1}, 40 times larger than its predecessor KEKB. The following content includes some descriptions on how this can be realized.

The luminosity of a particle collider can be expressed as [12]:

$$L = \frac{\gamma_{\pm}}{2er_e}\left(\frac{I_{\pm}\xi_{y\pm}}{\beta^*_{y\pm}}\right)\left(\frac{R_L}{R_{\xi_y}}\right), \tag{3.1}$$

where $\pm$ specifies values for positrons (+) and electrons (-), γ is the Lorentz factor, e is the elementary electron charge, r_e is the electron radius, I is the beam current, ξ_y is the vertical beam-beam parameter and β^*_y is the vertical beta function. R_L and R_{ξ_y} are reduction factors for luminosity and the vertical beam-beam parameter, and their ratio is usually close to unity. Table 3.1 summarizes the fundamental parameters accounting for the luminosity upgrades from KEKB to SuperKEKB. The key factor is that beta function at the interaction point (IP) is squeezed down 20 times and beam current doubles while beam-beam parameter remains unchanged, leading to approximately 40 times luminosity to KEKB. This squeezed vertical beta function can be realized owing to a novel "Nanobeam" scheme. Fig. 3.2 gives a schematic view of beam collision in Nanobeam scheme.

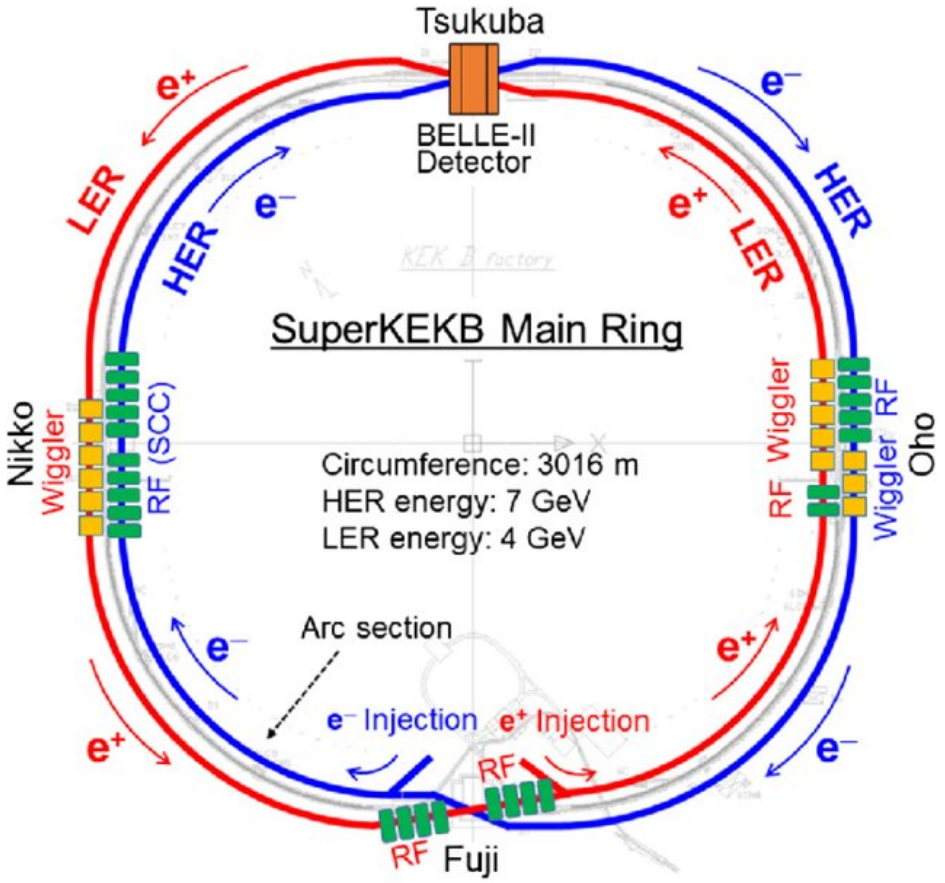


Fig. 3.1 Layout of the SuperKEKB Main Ring (MR) [11].

	KEKB	SuperKEKB
Energy (GeV) (LER/HER)	3.5/8.0	4.0/7.0
ξ_y	0.129/0.090	0.090/0.088
β_y (mm)	5.9/5.9	0.27/0.41
I (A)	1.64/1.19	3.60/2.62
Luminosity ($10^{34}\ cm^{-2}s^{-1}$)	2.11	80

Table 3.1 Fundamental parameters accounting for the luminosity upgrades from KEKB to SuperKEKB.

In accelerator physics, there is an hourglass effect [14] that restricts the minimum value for beta function. That is when bunch length gets larger than β, luminosity is adversely affected. The hourglass condition for usual head-on collision is

$$\beta_y^* > \sigma_z, \tag{3.2}$$

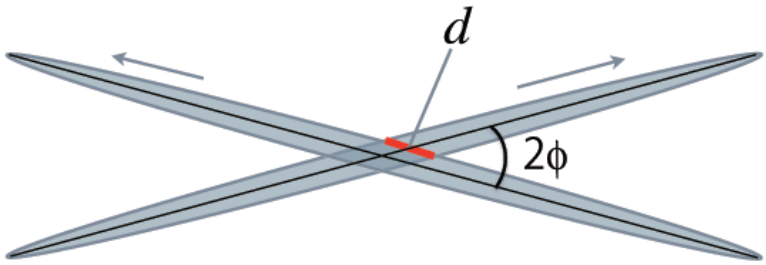


Fig. 3.2 Schematic view of beam collision in Nanobeam scheme [13].

while in the nanobeam scheme, the bunch length is replaced with the effective length d

$$\beta_y^* > d \cong \frac{\sigma_x^*}{\phi}. \tag{3.3}$$

The effective bunch length depends on the horizontal half crossing angle ϕ and the horizontal beam size σ_x^* at IP. Therefore, it requires extremely small horizontal beam size with a relatively large crossing angle. In SuperKEKB ϕ is 41.5 mrad, approximately 4 times larger than that of KEKB [15]. This machine parameter was decided not just in favor for hourglass condition, but also based on various factors from optics of interaction region, magnet design and the detector background.

With the nanobeam scheme, it also comes with various challenges from emittance growth due to intra-beam scattering and shortened beam lifetime due to Touschek effect. To mitigate such effects, the beam energies in SuperKEKB are changed accordingly: low-energy ring (LER) energy increases from 3.5 to 4.0 GeV. In addition, high-energy ring (HER) ring energy decreases from 8.0 to 7.0 GeV is advantageous for a lower beam emittance.

3.2 Belle II detector

3.2.1 Vertex Detector VXD

In Belle II, the inner most detector is the new VXD comprised of two parts, the silicon Pixel Detector (PXD) and the Silicon Vertex Detector (SVD). In the full configuration of VXD, there are altogether 6 layers, where the first two layers are at $r = 14$ mm and $r = 22$ mm using DEPFET type pixelated sensors, the following four layers are at the radius of 38 mm, 80 mm, 115 mm and 140 mm using the double-sided silicon strip sensors. In comparison with vertex detector in Belle, the first two layers are closer to the interaction point, and the outermost vertex detector at considerably larger radius.

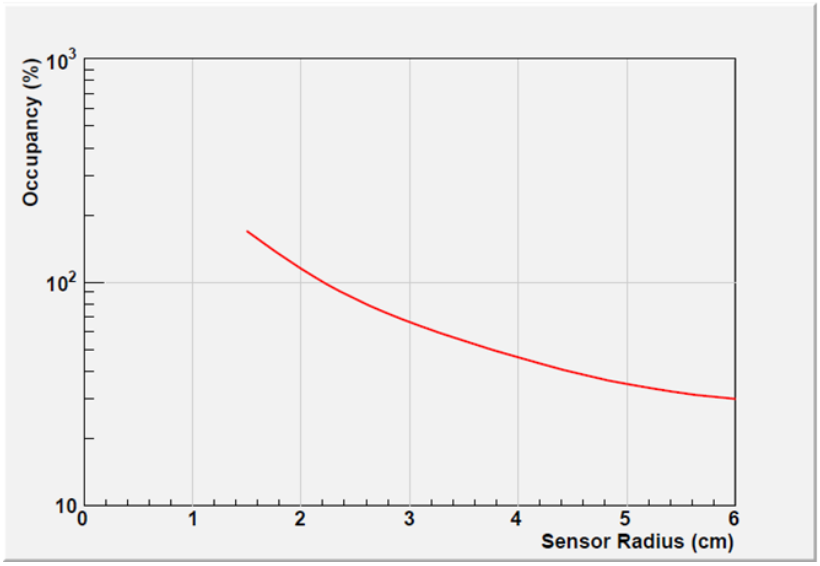


Fig. 3.3 SVD occupancy in the background of SuperKEKB, assuming the same hardware with Belle.

In the nanobeam scheme of SuperKEKB, the beampipe radius in the interaction region is only about 10 mm. While this is good news for physics analysis related to vertex reconstruction, it poses challenges as background increase with inverse square of radius. Because of this, strip detectors suffer from large occupancy, making it almost impossible to reconstruct B-decay vertices. The solution is to use pixel sensors rather than strips for the innermost layers, as the much larger number of channels in pixel sensors mitigates the occupancy problem. As shown in Fig. 3.3, only at radius of about 40 mm away, strip detectors are safe for use at the SuperKEKB luminosity.

The concept of pixel detector in Belle II, the PXD, is based on DEPFET (Depleted Field Effect Transistors) technology, a combination of detection and amplification within one device invented by Josef Kemmer and Gerhard Lutz in 1987. The PXD is incorporated to Belle II requirements through two main changes. First one is in the thickness, as relatively large amount of material in the pixel sensors would cause too much multiple scattering for precise reconstruction of B-decay. The PXD in Belle II was developed to allow for very thin sensors (50 μm). Another optimization is the radiation-hard technology to cope with the harsh backgrounds near the Belle II beamline. The readout electronics are located outside the acceptance region and will not contribute to multiple scattering thickness. The energy consumption of sensors are very little so that it is sufficient to use air cooling.

As shown in Fig. 3.4, the design of PXD in Belle II consists of two layers sensors just outside the interaction region. The inner layer at 14 mm has 8 planar sensors (“ladders”), each with a width of 15 mm and sensitive length of 90 mm. The outer layer at 22 mm contains 12

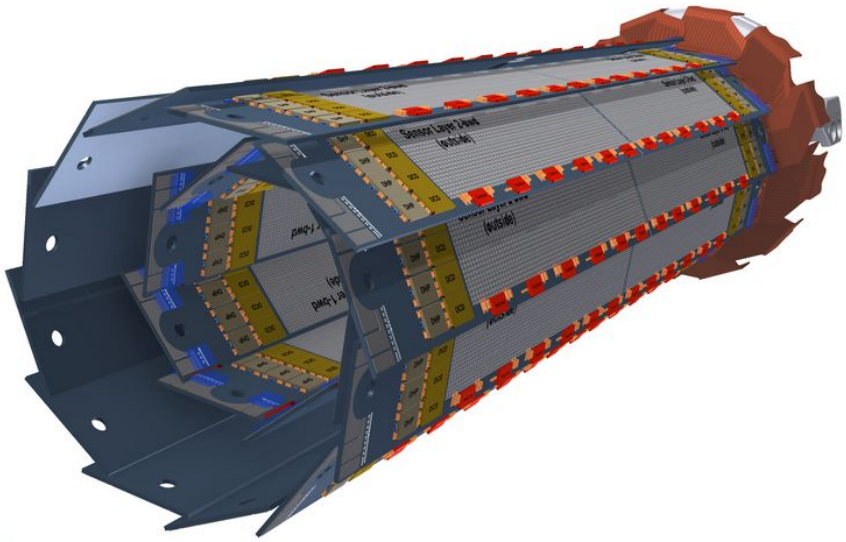


Fig. 3.4 Geometrical arrangement of the PXD. The grey surfaces are the DEPFET pixels.

ladders with width of 15 mm and sensitive length of 123 mm, to cover the required angular acceptance of tracker ($17° < \theta < 150°$).

The main purpose of the outer four layers of SVD is towards the precise measurement of mixing-induced CP asymmetry in B decay vertices. The Belle II SVD inherits the good characteristics of the Belle vertex detector: low mass, high precision, immunity to background hits, radiation tolerance and long-term stability. The design of Belle II SVD consists double-sided silicon strip detectors (DSSDs) fabricated from six-inch wafers, on which the APV25 (Fig. 3.5) readout chip hybrids with fast shaping time of 50 ns are mounted in order to suppress the background hits in the 40-fold higher luminosity in SuperKEKB. The arrays of DSSD modules form different ladders for SVD. Layer-3, 4, 5 and 6 has 7, 10, 12 and 16 ladders. The configuration is shown in Fig. 3.6. Ladders on layer-3 are flat, while the last ladder in the forward region of layer-4, 5 and 6 are slanted with angle degrees of 11.9, 16.0 and 21.1, respectively. The slanted shape was designed to reduce overall material budget as well as number of sensors, while improving hit quality by avoiding shallow hits with large cluster widths. P-side strips face towards the longitudinal direction (same as the beam direction) except for the slanted sensors in the forward region of layer-3, 4 and 5 which almost intersect with beam axis, and n-side strips face towards transverse direction. The fast shaping of APV25 readout chips is associated with the increase to capacitance noise. In order to minimize the signal-to-noise ratio, APV25 chips needs to be located as close to sensor

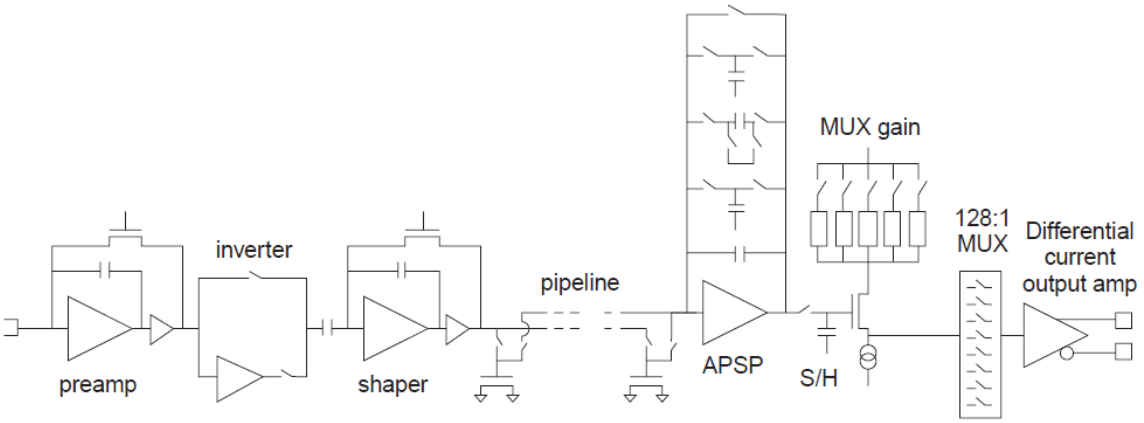


Fig. 3.5 Building blocks of the 128 channels of the APV25 front-end readout chip.

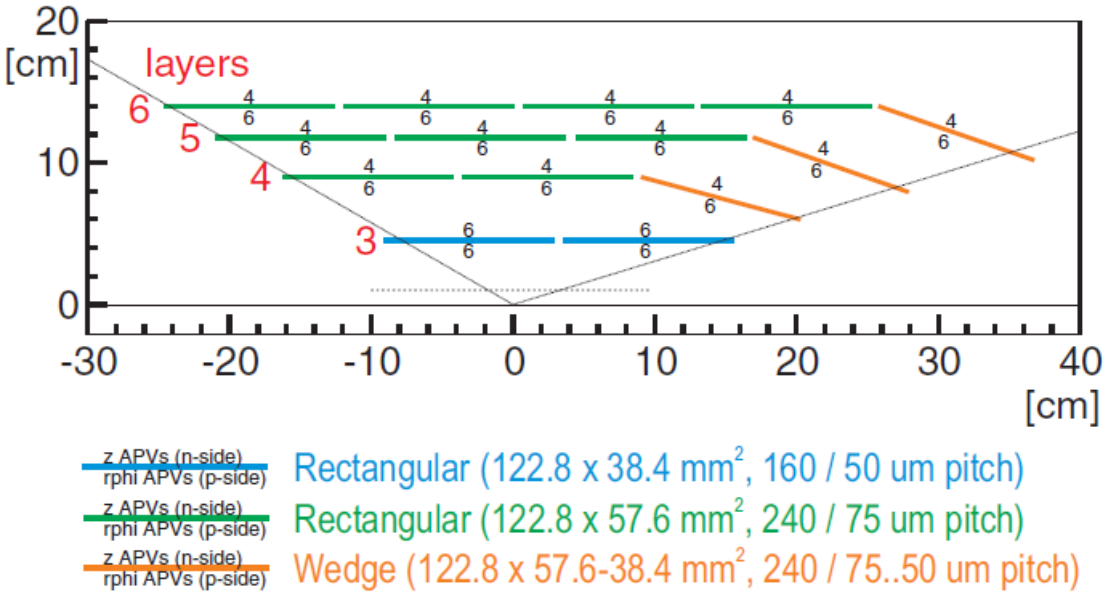


Fig. 3.6 Configuration of four strip layers of SVD, with slanted sensors in the forward region.

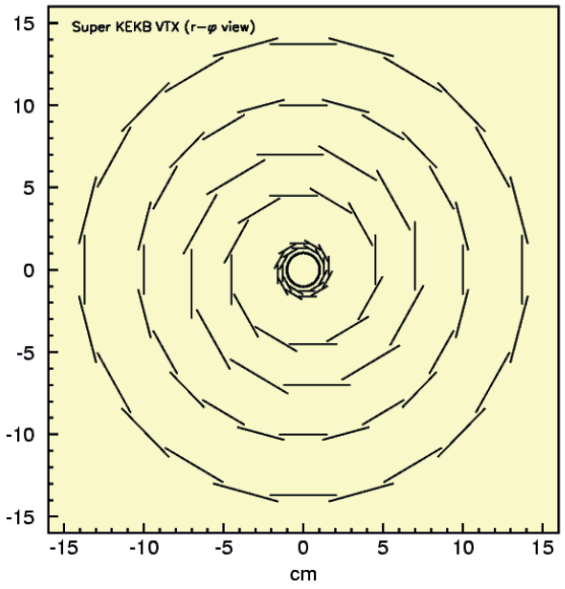


Fig. 3.7 A schematic view of the Belle II vertex detector.

strips as possible. In SVD of Belle II, the concept of Origami chip-on-sensor was adopted. With the use of pitch adapters bent around the edge, APV25 chips are connected to DSSDs. The bent fan-out pieces are placed in a windmill shape as in Fig. 3.7

In summary, VXD is one of the most important components of Belle II detector especially for measurements of time evolutions. With the significant upgrades to vertex detector from Belle, improvement in the vertex resolution for more precise measurements can be expected.

3.2.2 Central Drift Chamber CDC

CDC is the central device for tracking and identification of charged particles in the Belle II experiment. CDC reconstructs trajectories of charged particles to make precise measurement of their momentum and it provides particle identification information in the low momentum region through energy loss within the gas volume. It also provides trigger information of charged particles. The global structure of CDC in Belle II follows from CDC in Belle, which had a stable operation for more than 10 years. One of the main updates of CDC come from larger radius, faster electronics and smaller cells. The extension to CDC outer radius was made possible because of the more compact barrel particle identification device. The readout electronics are also updated to handle the higher trigger rates with less deadtime. The smaller cell size is required due to the large beam background in SuperKEKB. The wire configuration upgrade is shown in Fig. 3.8.

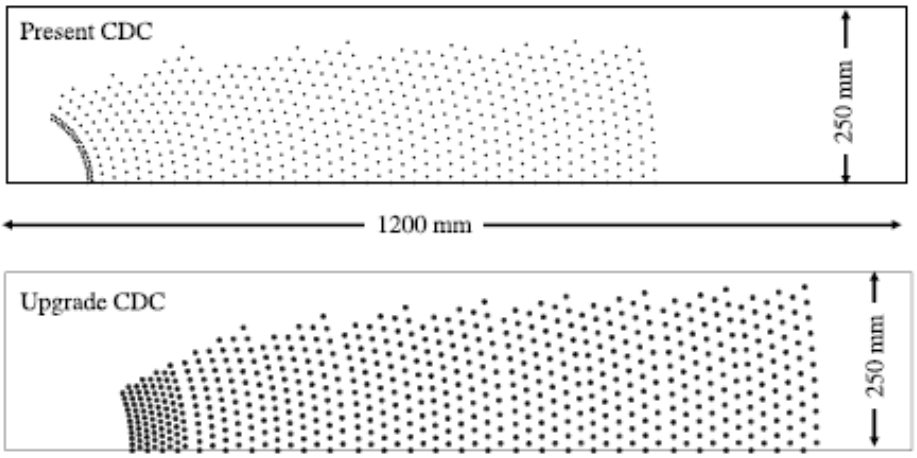


Fig. 3.8 Comparison of wire configuration between Belle (top) and Belle II (below) [16].

With the 20 larger background levels than Belle expected, CDC tracking performance might have some degrades, for example in the loss of efficiency when reconstructing some decays. But luckily, the new software and standalone tracking in PXD and SVD gives an overall improved performance in reconstruction efficiency.

3.2.3 Particle identification system TOP and ARICH

The higher background environment requires an update in the particle identification system.

In the barrel region for particle identification, the time-of-flight and aerogel Cherenkov counters are replaced with time-of-propagation (TOP) counter. On the outer wall of the CDC, 16 detector modules of TOP counters are located in different ϕ in the baseline geometry. Each module of TOP counter (Fig. 3.9) measures 45 cm wide, 2 cm thick and comprises of 125 cm long quartz bar, 10 cm quartz bar with mirror at one end of bar, and a wedge-shaped quartz block at the other end of bar [17]. 32 micro-channel plate (MCP) PMTs are installed on the wedge-shaped block. TOP counter identifies particles by measuring the precise time of Cherenkov photons spent reflecting inside the quartz radiator, and together with 2-d impact position information, the Cherenkov image is reconstructed. Combining with momentum measured in the inner trackers, the mass of particle is then identified.

In the forward end-cap region, the proximity-focusing Aerogel Ring-Imaging Cherenkov detector (ARICH) is used for identification of charged particles. It is designed to separation of kaons and pions over most of their momentum spectrum and to provide discrimination

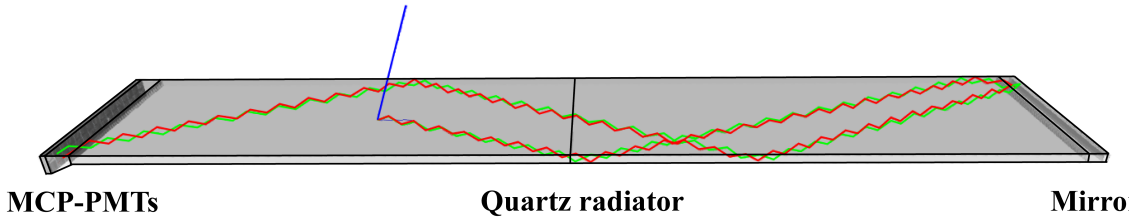


Fig. 3.9 Conceptual overview of TOP counter.

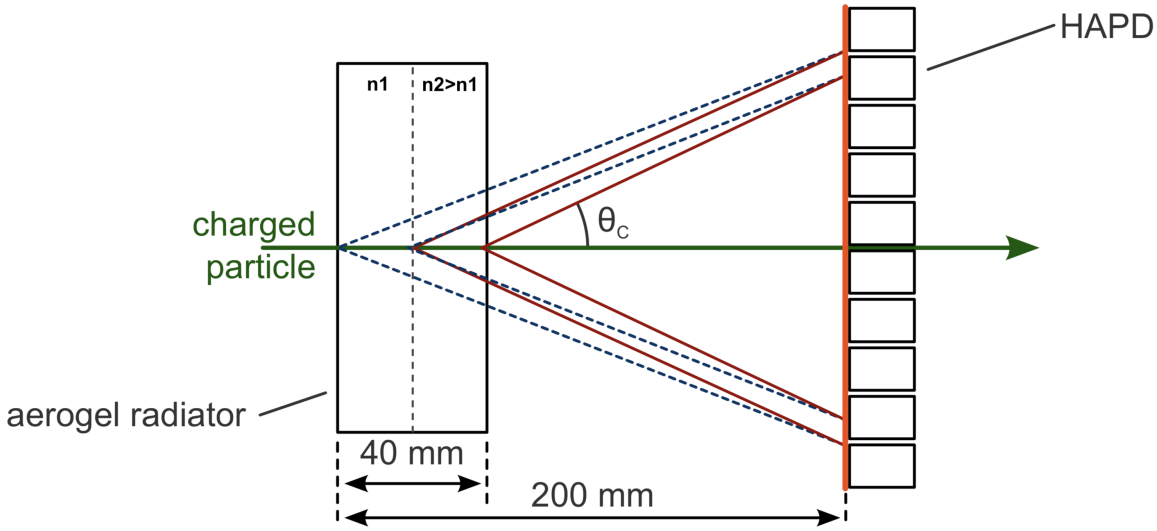


Fig. 3.10 Proximity focusing ARICH, with an inhomogeneous aerogel radiator ($n_1 = 1.045$ upstream, $n_2 = 1.055$ downstream).

between pions, muons and electrons below 1 GeV/c, in order to provide extension to our physics reach. Main elements of ARICH (Fig. 3.10) include an aerogel radiator where Cherenkov photons are produced by charged particles, an expansion volume to form rings on photon detector, an array of position sensitive photon detectors capable of detecting single photons in a high magnetic field with high efficiency and good resolution.

The Cherenkov angle resolution per charged particle can be used as indicator of performance goodness, $\sigma_{track} = \sigma_\theta/\sqrt{N}$. One limitation to the Cherenkov angle resolution is the emission point uncertainty. In ARICH in Belle II experiment, this was optimized with a non-homogeneous radiator, the 2 cm thick aerogel is divided to two layers with different refractive indices ($n_1 = 1.045$ upstream, $n_2 = 1.055$ downstream as in Fig. 3.10). This is equivalent to focusing the photons within the radiator, thus yield can be increased without much compromise to Cherenkov angle resolution due to emission point uncertainty.

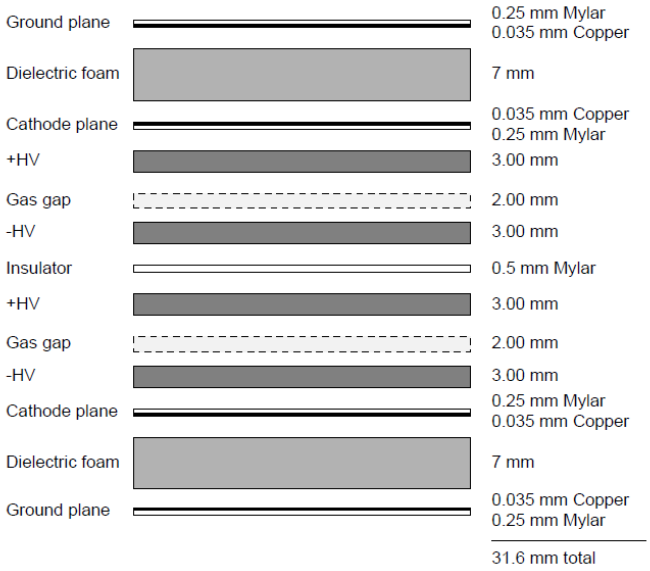


Fig. 3.11 Cross section for RPC superlayer in KLM.

3.2.4 Electromagnetic Calorimeter ECL

ECL is used to detect gamma rays and to identify electrons from hadrons like pions. CsI(TI) crystals are chosen as the scintillation crystal material due to its high light output, relatively short radiation length, good mechanical properties and moderate price. The CsI(TI) crystals, preamplifiers and support structures have been reused from Belle, while readout electronics and reconstruction software are updated. ECL consists of 3 parts, a 3 m long barrel section with an inner radius of 1.25 m, a forward endcap section at $z = 1.96$ m and a backward endcap section at $z = -1.02$ m from the interaction point. A total number of 8736 crystals are instrumented, covering about 90% of the solid angle in the center-of-mass system.

While the performance of ECL in Belle II is expected to have very similar performance in the absence of backgrounds, the pile up noise caused by overlapping of pulses from background events will degrade the performance of ECL, even with the upgraded electronics. The problem becomes more severe in the future planned much higher luminosity, as a result, possible solutions are under study (for example, replacing CsI(TI) with considerably faster and radiation tolerant pure CsI).

3.2.5 K_L-Muon Detector KLM

The main purpose of KLM in Belle II is to identify K_L and muons. It is located outside the superconducting solenoid, consisted of an alternating sandwich of iron places and active detector parts. The 4.7 cm iron plates serve as the magnetic flux return for the solenoid.

In Belle, KLM is based on glass-electrode resistive plate chambers (RPC, Fig. 3.11) and has demonstrated good performance during the entire data taking period. Compared with Belle, large background rates are expected due to neutrons produced in electromagnetic showers from background reactions. The long dead time of the RPCs results in serious fake muon identification probability, especially in the region of the endcaps and inner layers of the barrel. To mitigate this problem, in such regions scintillator strips with wavelength shifting fibers are instrumented. The readout is done by silicon photomultipliers (SiPMs) at Geiger mode.

3.2.6 Trigger and Data Acquisition (DAQ) Systems

The trigger system in Belle II plays the role of identifying and recording the events of interest during data taking. It must be robust and flexible in the much higher background rates in SuperKEKB, and also capable to unlock topics not probed in the previous generation of B factory, like single-photon trigger for searches in the dark sector, and two-photon and three-photon triggers for axion-like particle searches.

In Belle, the trigger system consists of sub-trigger systems and a final-decision logic. The Belle II trigger adapts this concept but replace all components with new technology of Field Programmable Gate Array (FPGA) so that the trigger logic is configurable rather than hard-wired. The nominal trigger system in Belle II can achieve a latency of about 5 μs, a timing precision of less than 10 ns, a minimum two-event separation of 200 ns, a maximum trigger rate of 30 kHz and a high efficiency of physics processes.

The goal of the data acquisition (DAQ) system is to read out detector signals upon the trigger decision given by the trigger system. The data from front-end electronics transfers through several steps of data processing, including a unified data link called Belle2Link, the common readout platform called COPPER, the event builder system, and the high level trigger (HLT) system, and finally to the storage system. For offline storage, the HLT system first suppresses the online event rates to 15 kHz using the information from CDC track finding and ECL reconstruction. The passing events are further considered for full event reconstruction to reject residual beam background not found by the hardware-based trigger system. The event rate is typically reduced to 10 kHz after the full event reconstruction.

The total cross sections and trigger rates at the goal luminosity of 8×10^{35} cm^{-2}s^{-1} for some typical physics processes of interest are listed in Table 3.2.

Physics process	Cross section (nb)	Rate (Hz)
$\Upsilon(4S) \to BB$	1.2	960
Hadron production form continuum	2.8	2200
$\mu^+\mu^-$	0.8	640
$\tau^+\tau^-$	0.8	640
Bhabha ($\theta_{lab} \geq 17°$)	44	350 [(a)]
$\gamma\gamma$ ($\theta_{lab} \geq 17°$)	2.4	19 [(a)]
2γ processes ($\theta_{lab} \geq 17°$, $p_t \geq 0.1$ GeV/c)	~ 80	~ 15000
Total	~ 130	~ 20000

[(a)] rate is pre-scaled by a factor of 1/100

Table 3.2 Total cross sections and trigger rates at the luminosity $L = 8 \times 10^{35}$ cm^{-2}s^{-1} from different physics processes at the CMS energy of $\Upsilon(4S)$.

3.3 Belle II software framework

The Belle II experiment creates huge amounts of data, and sophisticated algorithms for simulation, reconstruction and many analyses are required to fully exploit these data. The Belle II Analysis Software Framework (basf2) [18] is the fundamental tools developed for this purpose (Fig. 3.12). This Belle II-specific code is organized into many packages for base-level framework, different detector components, track reconstruction and post-reconstruction analysis tools. The packages are written in C++ to build particles from primitive objects (like tracks and calorimeter clusters). While the C++ based codes do the heavy lifting, the Python based codes help provide a more friendly user interface. The main use of Python in basf2 is to load and configure C++ modules through steering files, along with some high-level analysis. The basf2 also has a dependence on various high energy physics specific third-party packages like ROOT, Geant4 and EvtGen to produce Monte Carlo (MC) datasets.

The data processing strategy in Belle II is the use of data-summary table (dst), a special ROOT file. The raw data recorded has the size of $\sim$100 kB/event, translating to $\sim$10 PB over the next few years. Such large dataset goes through a process of unpacking and reconstruction performed by the data production group and is saved in a well-defined format mini data-summary table (mdst). An mdst file contains the essential particle information for physics analysis, for example, trajectory of tracks, fitting results of tracks to given particle hypotheses, particle identification (PID) likelihoods, reconstructed clusters in ECL and KLM, trigger information and MC generated information of each particle. The typical size of mdst file is reduced to $\sim$10 kB/event. The size is further reduced by a loose pre-selection (skimming) based on the target physics analysis of interest to skimmed user data-summary table (udst).

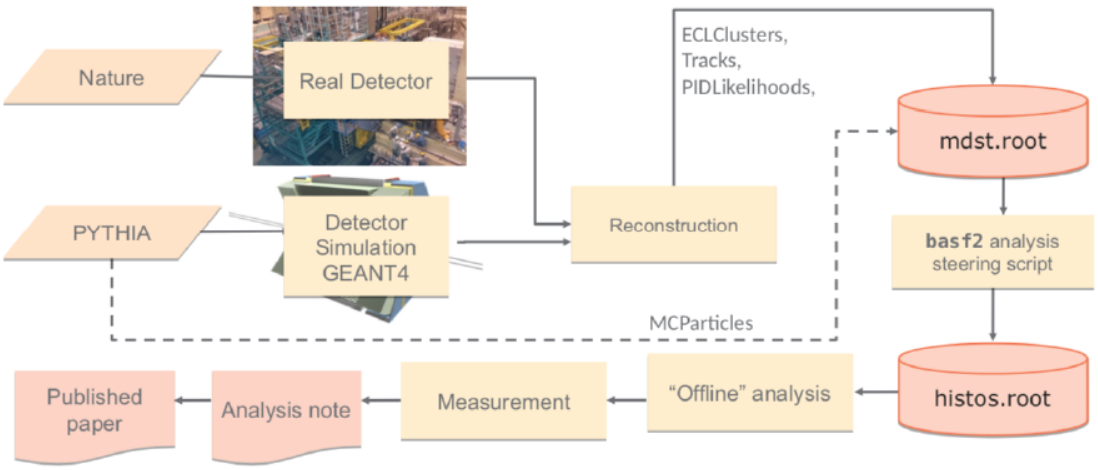


Fig. 3.12 The experiment's flow in the Belle II Analysis Software Framework.

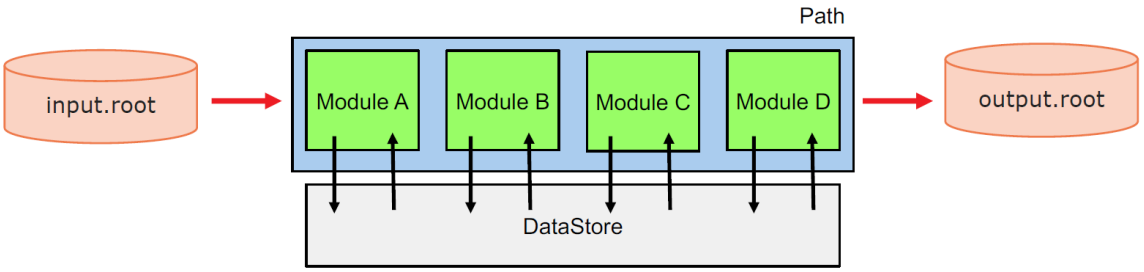


Fig. 3.13 User configured data process using basf2 analysis steering script.

The dst files are passed to user analysts for configured data processing (Fig. 3.13). The process contains three parts

- Module: A set of classes (modules) that process the data
- DataStore: A set of classes (dataobjects) that hold the data and are allowed to pass from one to another module.
- Path: An order in which the modules are executed.

Scripts written in Python codes play the role of steering in the data processing, and the particle information in mdst are converted into more friendly quantities like particle lists and vertex positions.

3.4 Belle II operations

In the Belle II experiment, the detector has been in three commissioning phases (Fig. 3.14):

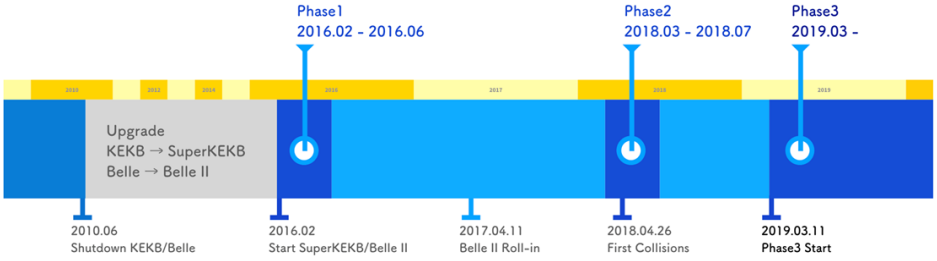


Fig. 3.14 Timetable for the commissioning of the Belle II experiment.

1) **Phase 1** Starts in February 2016 and ends in June 2016, during this phase, low-emittance Ampere-level beams were circulated in both rings, but no collisions took place. The solenoid was not active, and the Belle II detector was not yet installed.

2) **Phase 2** From April 2018 to July 2018, this period marks the pilot run of Belle II. During this phase, the final magnetic field configuration was realized for charged track reconstruction. All subsystems of detector except for VXD was ready. The partially installed VXD contained only one octant of PXD and SVD, consisting of 2 and 4 ladders, respectively. They were placed in the highest beam background region in the $+X$ direction, in order to guarantee the safe operation of the full VXD in the future. Background sensors only for Phase 2 was installed to judge background conditions for VXD in Phase 3. The nanobeam scheme was also verified by reaching a luminosity larger than 10^{34} cm^{-2}s^{-1}.

3) **Phase 3** From March 2019, Belle II begins its "physics" run period. In this phase, the final detector configuration will be realized. It is planned to take data for 9 month every year in the future. The first B-$\bar{B}$ like event in Belle II Phase 3 physics run is shown in Fig. 3.15. Now, SVD has been fully installed and in operation. PXD, however, is not in the final form yet. Only 2 out of the 12 ladders in the outer layer was installed as shown in Fig. 3.16. The full installation is scheduled in 2022.

Up untill now in June, 2020, Belle II has recorded an integrated luminosity of 74.10 fb^{-1} as shown in Fig. 3.17.

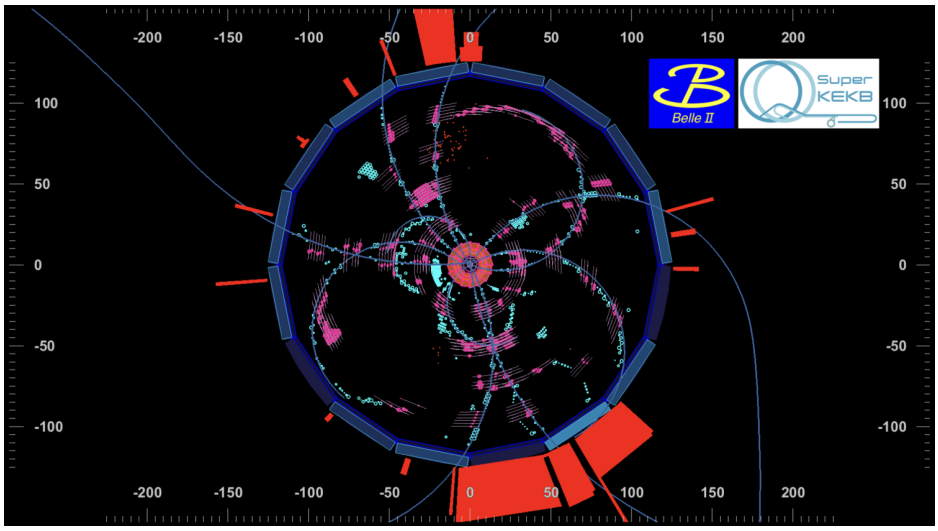


Fig. 3.15 First observed B-$\bar{B}$ like event in the Belle II Phase 3 physics run.

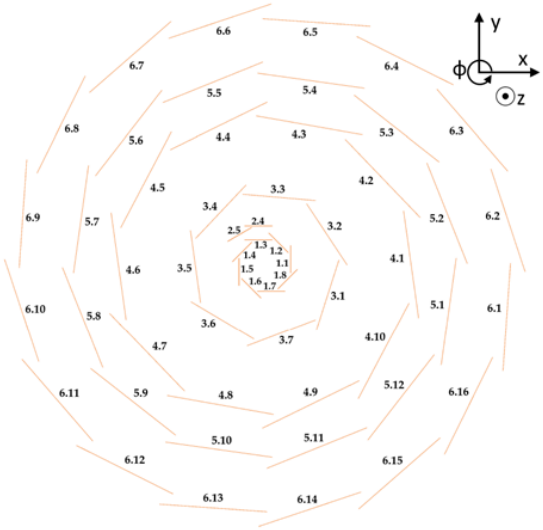


Fig. 3.16 Current configuration of the VXD in Phase 3 [19]. The outer layer of PXD has only 2 ladders installed.

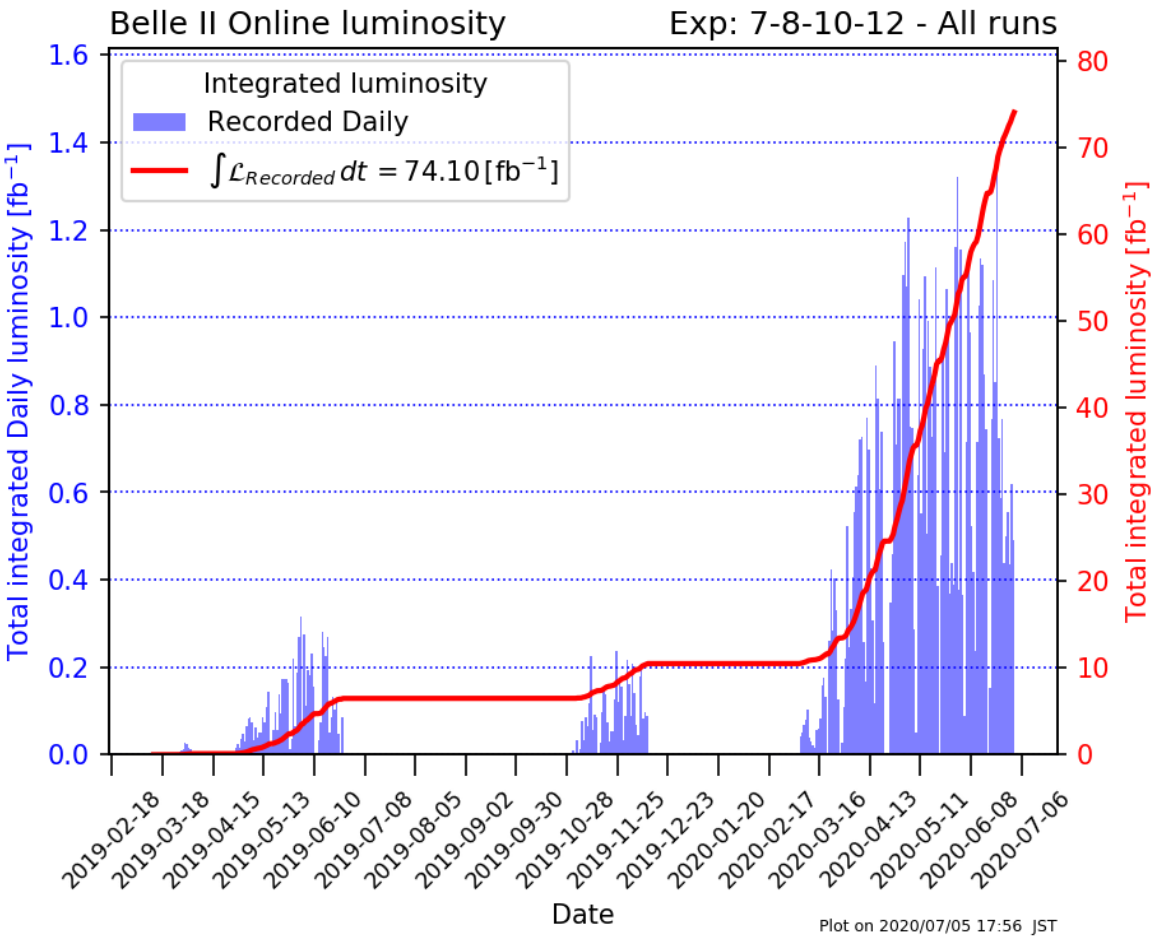


Fig. 3.17 Online luminosity in the Belle II Phase 3 physics run.

Chapter 4

Reconstruction of Events

4.1 Track reconstruction

The task of track reconstruction is to reconstruct charged particles originated from primary or secondary vertices from the detector hit information (PXD, SVD and CDC). The performance of track reconstruction is an important benchmark to physics analysis like time-dependent CP asymmetry measurements in Belle II.

The design of track reconstruction in Belle II is summarized in Fig. 4.1. The workflow contains two main parts , track finding and track fitting.

In the **track finding** part, detector hits that belong to a single track are collected into a track candidate. In SVD and CDC, independent pattern recognition algorithms are used. In SVD, a cellular-automaton (CA) algorithm is used to reduce the large number of combinatorial track candidates, especially in the presence of background. Compatible pairs of hits that satisfy the requirement from the SVD simulated sector map in the adjacent layers are combined to a cell. The shared hit in the neighbor cells are identified as threads if it is geometrically allowed. This results in a set of possible combinations of hits where same hits may appear in several combinations. Then a Monte Carlo trained Hopfield network is used to obtain the best non-overlapping set of track candidates and feed to track fitting modules. In CDC, other than the adapted track finding algorithms used in Belle, two newly developed complementary algorithms are employed. The global track finder looks for intersections of tracks in the Legendre space. It copes with missing hits and is used in the vicinity of interaction point. The other local track finder uses CA algorithm, making it robust to energy losses through the wire chamber volume. The combination of the two track finders results in good efficiency in CDC. The SVD and CDC reconstructed tracks are merged through a Combinational Kalman Filter (CKF). The standard approach runs CKF before SVD track

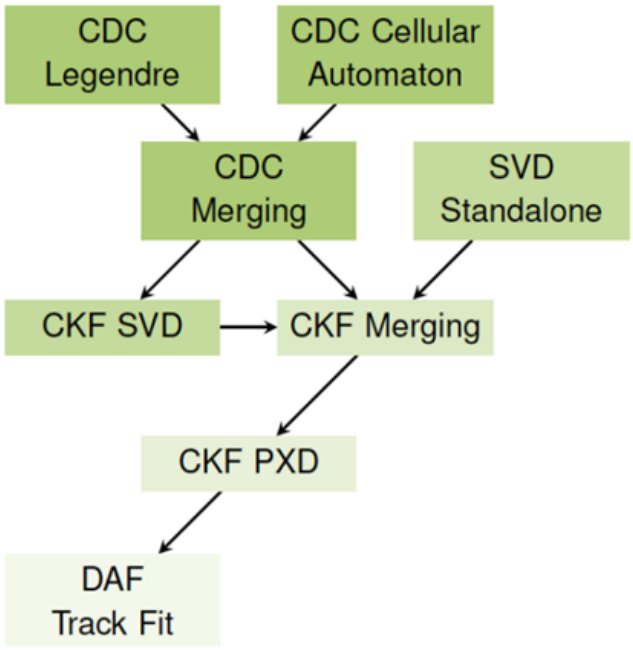


Fig. 4.1 Illustration of procedures for track reconstruction in Belle II [20].

finder to reduce the possible number of combinations. VXD hits are also attached using CKF as the last step of track finding.

In the **track fitting** part, a point $\vec{P}$ is defined as the closest approach of a trajectory to the origin in the $r-\phi$ plane. Five parameters are used to describe the helix formed by charged particle propagating in the presence of magnetic field at $\vec{P}$:

1) d_0: the transverse signed distance of the perigee from the origin. The sign is determined by the direction of angular momentum with respect to the magnetic field.

2) z_0: the longitudinal signed distance of the perigee from the origin.

3) ϕ_0: the angle between the transverse momentum at the perigee and the x axis.

4) $tan\lambda$: the tangent of the angle between the momentum at perigee and the transverse plane.

5) ω: the curvature with sign corresponds to charge of track.

The main track fitting algorithm is the deterministic annealing filter (DAF) which is based on Kalman filter. DAF deals with problems like background-induced hit assignments and left/right ambiguities in the wire chamber measurements by iterating with weighting and annealing to find the best fit result. In Belle II, effects of non-ideal helix due to interaction

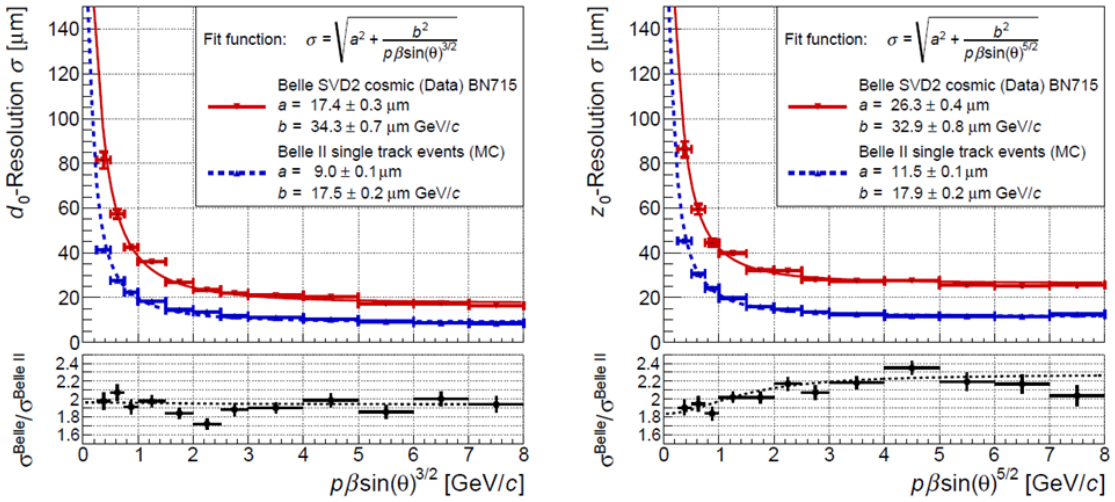


Fig. 4.2 Resolution of the transverse d_0 and longitudinal z_0 impact parameters. The resolution is estimated using the standard deviations of the single Gaussian function fits. Red line stands for fits to cosmic data in Belle, blue line stands for fits to single tracks events in MC in Belle II.

with detectors, as well as non-uniformity of magnetic field are taken into account. To correctly identify different particles, different mass hypotheses are used during track fitting.

As described in the previous chapter, the designed Lorentz boost in Belle II is $\beta\gamma \approx 0.284$, about two-thirds of that in its predecessor Belle experiment. Such lowered boost results in smaller separation of B mesons, thus a more precise track reconstruction is needed to measure the time evolution. Luckily, as we can see from Fig. 4.2., the resolutions of transverse d_0 and longitudinal z_0 impact parameters in comparison with Belle have improved by a factor of two.

4.2 Vertex fitting

Vertex fitting [21] provides tools to determine the momentum, invariant mass and decay vertex position from the reconstructed parameters of the final state particles. The vertex fitting for B_{sig} and B_{tag} from pair productions use very different methods. For B_{sig} reconstruction where all decay products are specified, the standard fitting algorithm is TreeFitter.

TreeFitter [22] is a type of decay chain fitting algorithm, originally developed by BaBar collaboration and adapted to Belle II experiment. Unlike the traditional approach to start fitting from final states and then build the tree from bottom up (also known as “leaf to leaf”), TreeFitter uses a global technique where the full tree is fitted simultaneously. Such global

approach poses challenge in the computation in the inversion of large matrices. It is solved by the Kalman Filter where inversion of multiple small matrices is calculated and combined for faster calculations. Neutral particles are not associated with tracks and are identified only by energy deposit in the Belle II calorimeter. Such bias due to displaced neutral rich vertex can be corrected by TreeFitter by using information from the rest of the decay chain.

Different from the signal side, B_{tag} vertex fitting is inclusive one where final state is not known. Tracks that originate from B_{sig} are removed, along with a veto to exclude combinations of two tracks with opposite charge and invariant mass in the range of K_S^0 to ensure tag side reconstruction efficiency. The remaining tracks are the input to vertex fitters. In Belle II analysis software framework (BASF2), two fitting algorithms are available:

RAVE Adaptive Fitter [23], originating from the CMS libraries, is a robust generalization of the Kalman filter, iteratively downweighting the outlier track contribution through soft assignment. Each track associated to a specific vertex is assigned with a weight based on the square of the standardized residual χ^2:

$$\omega_i(\chi_i^2) = \frac{e^{-\chi_i^2/2T}}{e^{-\chi_i^2/2T} + e^{-\sigma_{cut}^2/2T}}, \tag{4.1}$$

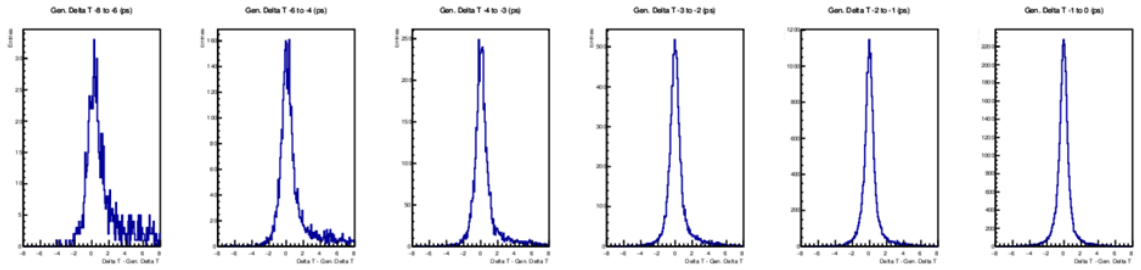
where σ_{cut} is defined as the standardized residual at $\omega_i = 0.5$, and the annealing temperature parameter T which is lowered geometrically in the iterative fitting process controls the shape of functional dependence to avoid falling into local minima. The fitter proceeds by minimizing the weighted least sum of squares.

Kinematic Fitter [24] which was inherited from the Belle experiment uses the Lagrangian multipliers to impose the kinematic constraints to the fit. The function to be minimized is in the form of:

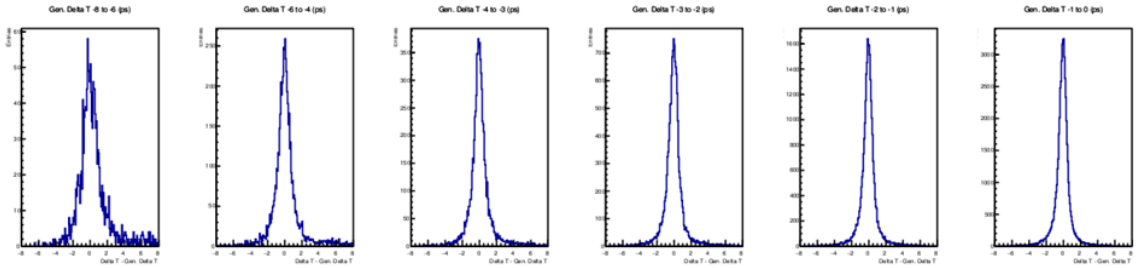
$$\chi^2 = (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0)^T V_{\boldsymbol{\alpha}_0}^{-1} \boldsymbol{\alpha} - \boldsymbol{\alpha}_0) + 2\boldsymbol{\lambda}^T \boldsymbol{H}(\boldsymbol{\alpha}), \tag{4.2}$$

where $\boldsymbol{\alpha}$ is the measurement to be improved, V is the corresponded covariance matrix, $\boldsymbol{H}$ is the kinematic constraints and $\boldsymbol{\lambda}$ is the set of Lagrangian multipliers.

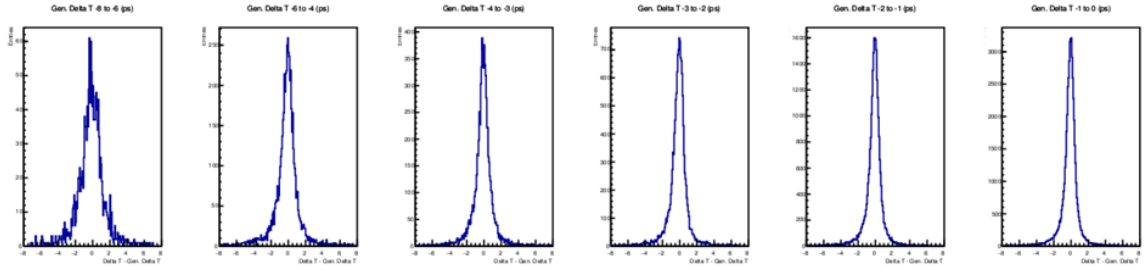
For the tag side vertex reconstruction, the constraints on position is important. In Belle, the beam spot size ($120 \times 5 \times 8000\ \mu\text{m}^3$) was used as interaction-point constraint. In Belle II, thanks to the nanobeam scheme the average spot size is much smaller ($6 \times 0.06 \times 150\ \mu\text{m}^3$). However, there comes some evidence showing possible negative consequences to the bias of decay vertex position, as in Fig. 4.3 In order to use the small beam spot size in a better way, a new idea of constraint Btube was introduced to Belle II.



(a) Vertex reconstruction with nominal beam spot constraint



(b) Vertex reconstruction with 2× beam spot constraint



(c) Vertex reconstruction with 20× beam spot constraint

Fig. 4.3 Δt resolution function of generated $B^0 \to [J/\psi \to \mu^+\mu^-][K_S^0 \to \pi^+\pi^-]$ sample plotted in different slices of negative Δt^{gen}. Bias at most negative slice of Δt^{gen} is possibly caused by beam spot constraint being too tight.

In the new Btube constraint, the fully reconstructed B_{sig} tracks are propagated to the beam spot and applied with a vertex fit. Such fit result is assumed to be the origin of both B mesons in pair production. The flight direction of B_{tag} can be obtained from the four-momentum conservation. The constraint is based on the reconstructed beam spot and constructed with infinite size in the flight direction of B_{tag}, hence the tube-like shape (Fig. 4.4).

The performance of vertex fitting can be examined by looking at the Δt resolution function. Fig. 4.5 shows the proper time difference distribution using $B^0 \to [J/\psi \to \mu^+\mu^-][K_S^0 \to \pi^+\pi^-]$, the major Belle II updates on both hardware and software translate to a resolution of 0.695 ps, providing a superior separation capability compared with Belle (resolution=0.92 ps).

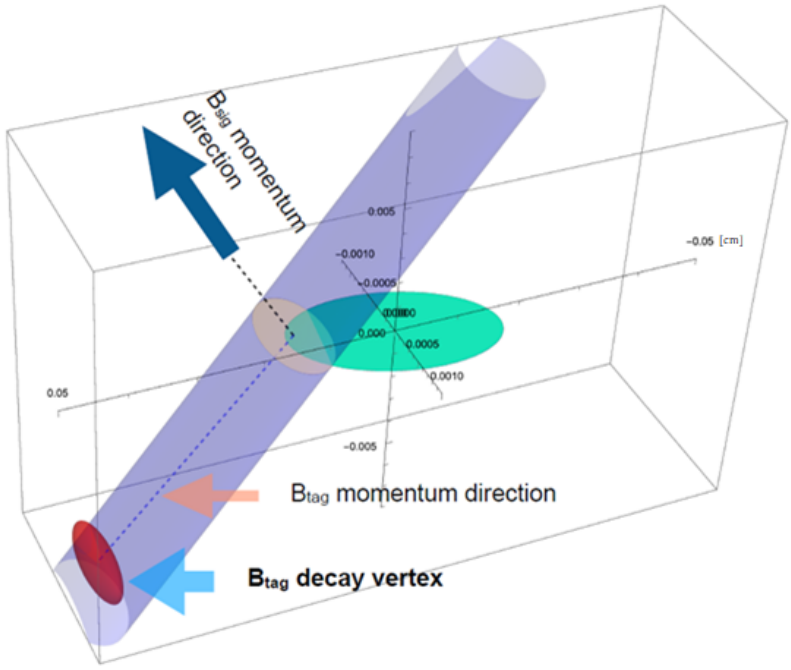


Fig. 4.4 Illustration of concept of the tube-like constraint Btube [25], the scale is in cm.

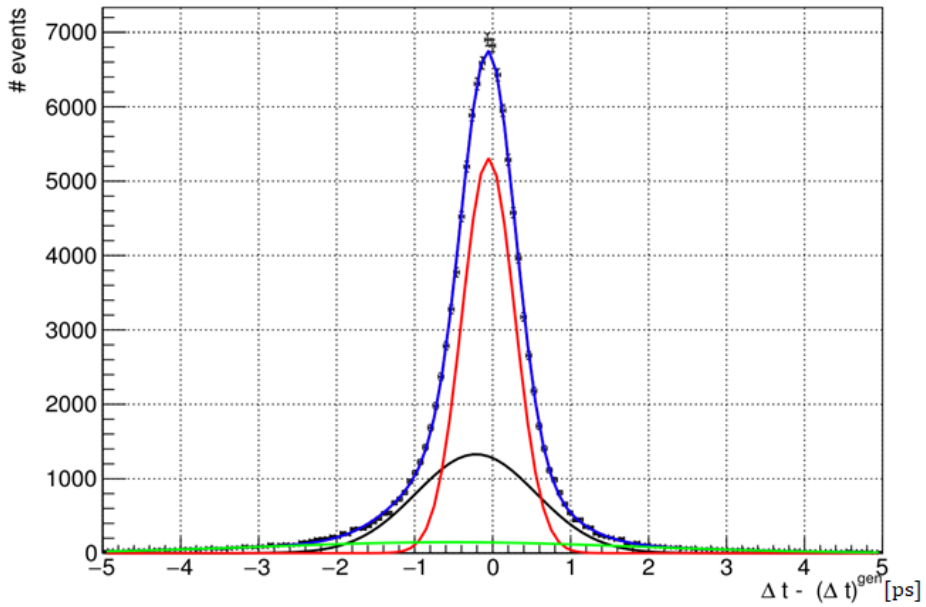


Fig. 4.5 Δt resolution function of generated $B^0 \to [J/\psi \to \mu^+\mu^-][K^0_S \to \pi^+\pi^-]$ sample fitted with three Gaussian functions. The resolution is estimated using the weighted standard deviation of the three Gaussian functions.

4.3 Event Selections

The decay channel studied for the analysis in this thesis is the hadronic control channel

$$B^0 \to \left[D^- \to K^+\pi^-\pi^-\right]\pi^+.$$

For charged particle final states, the tracks are reconstructed from PXD, SVD and CDC hits information. The track position is required to be close enough to the interaction point with the cut

$$|d_0| < 2 \text{ cm},\ |\mathrm{z}_0| < 5 \text{ cm}.$$

Here d_0 and z_0 are the transverse and longitudinal distance to the interaction point. In order to ensure signal reconstruction efficiency, a loose particle identification (PID) cut is applied

$$PID_K > 0.1,\ PID_\pi > 0.1.$$

D^- is reconstructed from tracks of K^+ and π^- using a particle combiner with a cut of mass range

$$1.85\ \mathrm{GeV/c^2} < \mathrm{m}_{\mathrm{K^+\pi^-\pi^-}} < 1.88\ \mathrm{GeV/c^2}.$$

In the case when multiple candidates of fit results exist, a best candidate selection with the smallest difference with nominal mass $|m_{rec} - m_{nom}|$ is used.
For the reconstruction of B^0, beam energy constraints are used for better background rejection

$$M_{bc} > 5.2\ \mathrm{GeV/c^2}, \Delta\mathrm{E} < 0.2\ \mathrm{GeV},$$

where $M_{bc} = \sqrt{E^{*2}_{beam} - \boldsymbol{p}^{*2}_B c^2}/c^2$ is the beam constrained mass and $\Delta E = E^*_B - E^*_{beam}$ is the energy difference.

For the vertex position fitting to the signal side, TreeFitter is used with minimum confidence level set to 0 to only accept tracks that converge during fit. An IP constraint forces the fit to converge at the region of vertex production. In the tag side, Btube is used as the position constraint for B_{tag} and kinematic fitter is chosen over Rave Adaptive Fitter in the consideration for the construction of Δt resolution function (more details will be shown in Chapter 5).

The total reconstruction efficiency with the event selection procedures described above is studied to 500k generated $B^0 - \bar{B}^0$ pairs to be 23.9%.

Chapter 5

Proper-time resolution function

5.1 Introduction

The proper-time resolution function $R(\Delta t)$ is a special concept in the context of time dependent analysis in Belle II. Defined as the experimental resolution for measurement of decay time difference between the two B mesons, it is represented as the number-of-events distribution of $\Delta t - \Delta t^{gen}$, the difference between measured and the true (generated) proper-time difference. Fig. 5.1 plots the simulated proper-time resolution function, from which we can see the distribution is of the same order as flight time of B mesons ($\sim$ 1.5 ps). Therefore, it is essential to understand and parameterize this experimental resolution very well for precise measurements.

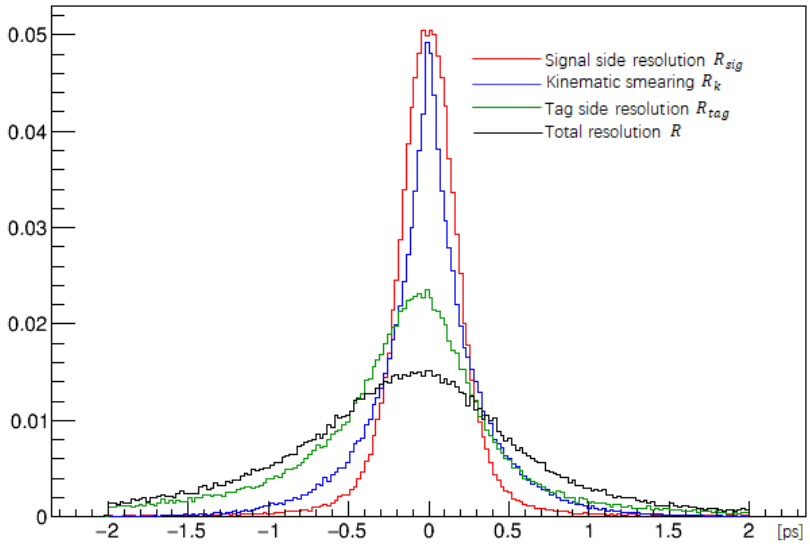


Fig. 5.1 Proper-time resolution function from simulated $B^0 \to [D^- \to K^+\pi^-\pi^-]\pi^+$ MC sample. The dominant contribution comes from tag side.

In the B factory experiments, such quantity is obtained from the vertex position due to the Lorentz boost:

$$\begin{aligned}\Delta t - \Delta t^{gen} &= (1/\beta\gamma c)\cdot(z_{sig}-z_{tag}) - \left(t_{sig}^{gen}-t_{tag}^{gen}\right)\\ &\approx (1/\beta\gamma c)\cdot(z_{sig}-z_{tag}) - (1/\beta\gamma c)\cdot\left(z_{sig}^{gen}-z_{tag}^{gen}\right)\\ &= (1/\beta\gamma c)\cdot\left(z_{sig}-z_{sig}^{gen}\right) - (1/\beta\gamma c)\cdot\left(z_{tag}-z_{tag}^{gen}\right).\end{aligned} \tag{5.1}$$

We observe three effects that constitute the total resolution function: signal side B meson vertex position resolution by $z_{sig}-z_{sig}^{gen}$, tag side B meson vertex position resolution by $z_{tag}-z_{tag}^{gen}$, and the kinematic smearing resolution due to the approximation made in the middle of calculation. As a result, we can write

$$R(\Delta t) = \iint d(\Delta t')d(\Delta t'')R_{sig}(\Delta t-\Delta t')R_{tag}(\Delta t'-\Delta t'')R_k(\Delta t''). \tag{5.2}$$

In the remaining parts of this chapter, the development of modelling and parameterization of Δt resolution function is presented using 500,000 simulated $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ MC events.

5.2 Signal side resolution

In the reconstruction of B_{sig}, the finite resolution of detector is the dominant cause of the broadened distribution between the measured and the true vertex position, namely

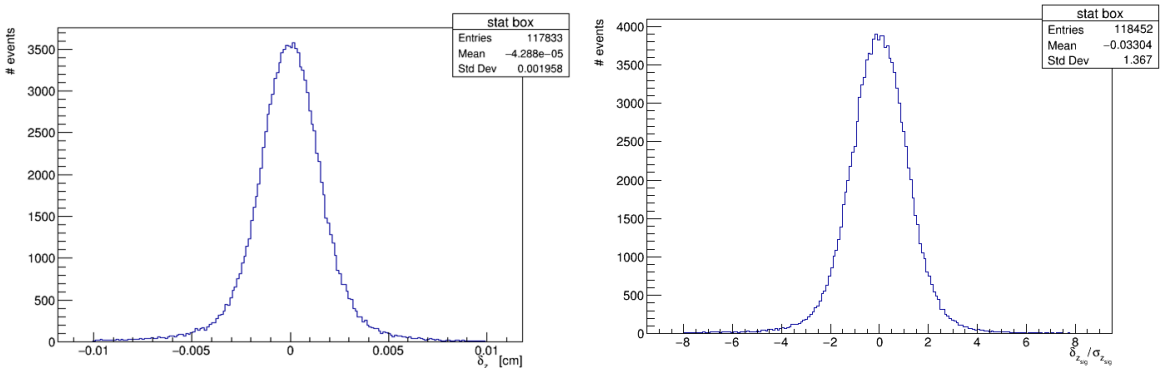
$$R_{sig} = R_{sig}^{det}. \tag{5.3}$$

The residual distribution of signal side vertex in boost direction $\delta_{z_{sig}} \stackrel{\text{def}}{=} z_{sig}-z_{sig}^{gen}$ is shown in Fig 5.2a.

In order to formulate a model that describes the resolution well, with the flexibility to possible disagreements between Monte Carlo simulation and experimental data, the event-by-event information implemented model is introduced as followed:

For each reconstructed vertex, the expected measurement uncertainty can be extracted from the vertex fitter. In particular, $\sigma_{z_{sig}}$, the uncertainty in the boost direction is to our interest. If $\sigma_{z_{sig}}$ describes the uncertainty perfectly, the distribution of $\delta_{z_{sig}}/\sigma_{z_{sig}}$ (also known as "the pull distribution") should be Gaussian with a standard deviation of 1:

$$P\left(\delta_{z_{sig}}/\sigma_{z_{sig}}\right) = G\left(\delta_{z_{sig}}/\sigma_{z_{sig}};\sigma\approx 1\right), \tag{5.4}$$



(a) Signal side boost direction $\delta_{z_{sig}}$ residual distribution

(b) Signal side boost direction $\delta_{z_{sig}}/\sigma_{z_{sig}}$ pull distribution

Fig. 5.2 Signal side resolution distribution from simulated MC sample.

where the Gaussian distribution is defined as

$$G(x;\sigma) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{x^2}{2\sigma^2}\right), \tag{5.5}$$

Fig. 5.2b shows the pull distribution of $\delta_{z_{sig}}/\sigma_{z_{sig}}$.

In addition, we consider the reduced chi-square of vertex fitting $(\chi^2/ndf)_{sig}$. The ndf here refers to the number of degrees of freedom, for signal side vertex reconstruction it is constant 6, 4 of which comes from charged tracks and the other 2 comes from interaction point constraint. We have found that if $\delta_{z_{sig}}/\sigma_{z_{sig}}$ pull is divided into slices with different values of $(\chi^2/ndf)_{sig}$, they can be fitted with single Gaussian distributions as in Fig. 5.3a. And furthermore, the standard deviations of the fitted Gaussian functions show a linear dependence on $(\chi^2/ndf)_{sig}$ as in Fig. 5.3b

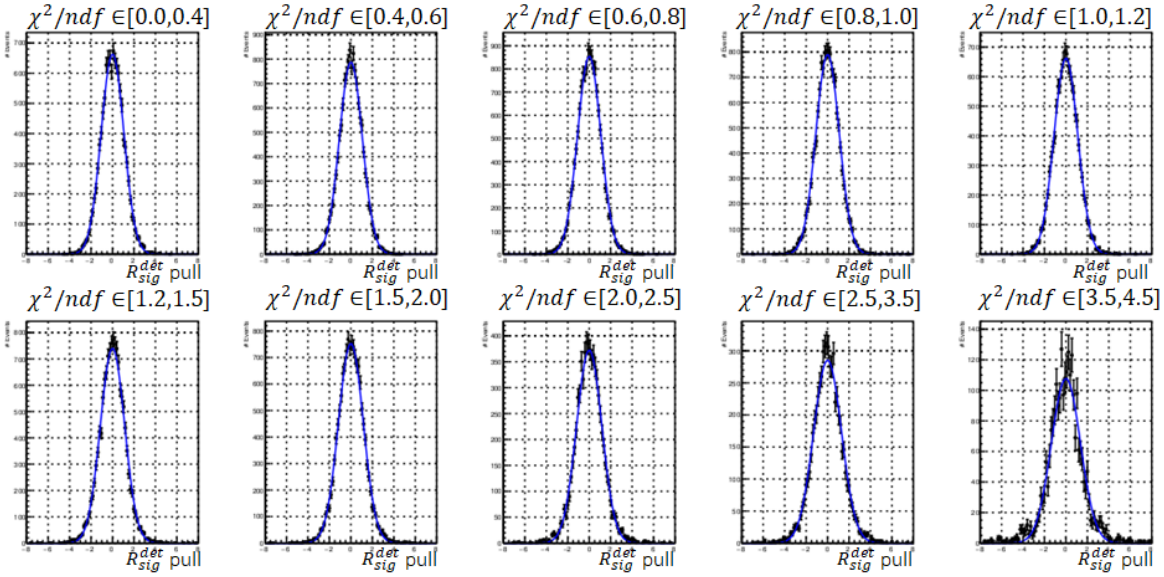
Based on the observations made above, we construct the signal side detector resolution to be the sum of two Gaussian fit functions

$$R^{det}_{sig}(\delta_{z_{sig}}) = (1-f^{tail}_{sig})G\left(\delta_{z_{sig}}; s^{main}_{sig}\cdot\sigma_{z_{sig}}\right) + f^{tail}_{sig}G\left(\delta_{z_{sig}}; s^{tail}_{sig}\cdot\sigma_{z_{sig}}\right), \tag{5.6}$$

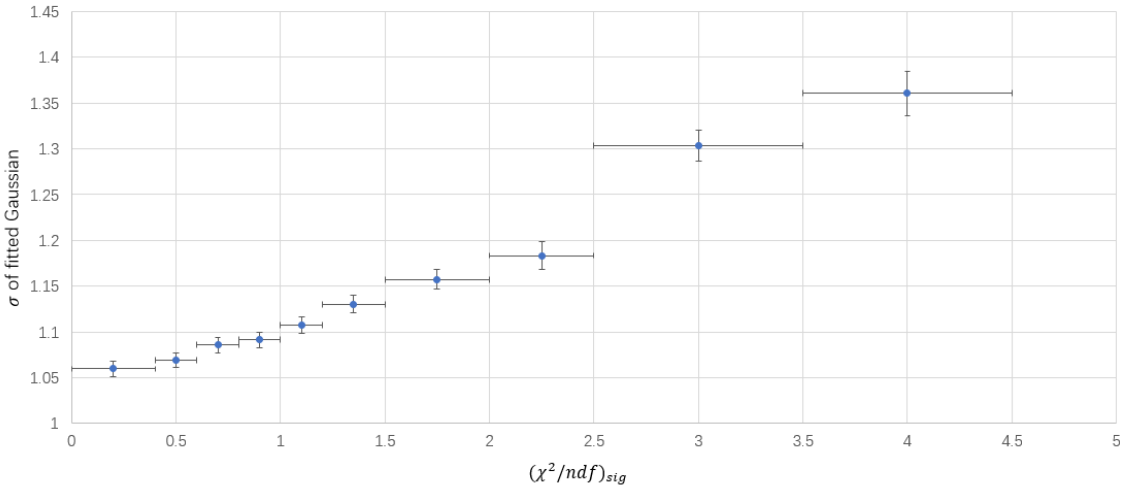
where s_{sig} is the first order polynomial of $(\chi^2/ndf)_{sig}$ to account for the linear dependence

$$s_{sig} \stackrel{\text{def}}{=} s^0_{sig} + s^1_{sig}\cdot(\chi^2/ndf)_{sig}. \tag{5.7}$$

The fit result for $\delta_{z_{sig}}$ residual distribution using such event-by-event model is shown in Fig. 5.4.



(a) Single Gaussian function fits to $\delta_{z_{sig}}/\sigma_{z_{sig}}$ pull in different slices of $(\chi^2/ndf)_{sig}$.



(b) Linear dependence of the standard deviations σ of the fitted Gaussian functions on $(\chi^2/ndf)_{sig}$.

Fig. 5.3 Introduction of the event-by-event vertex quality indicator $(\chi^2/ndf)_{sig}$ to signal side resolution function.

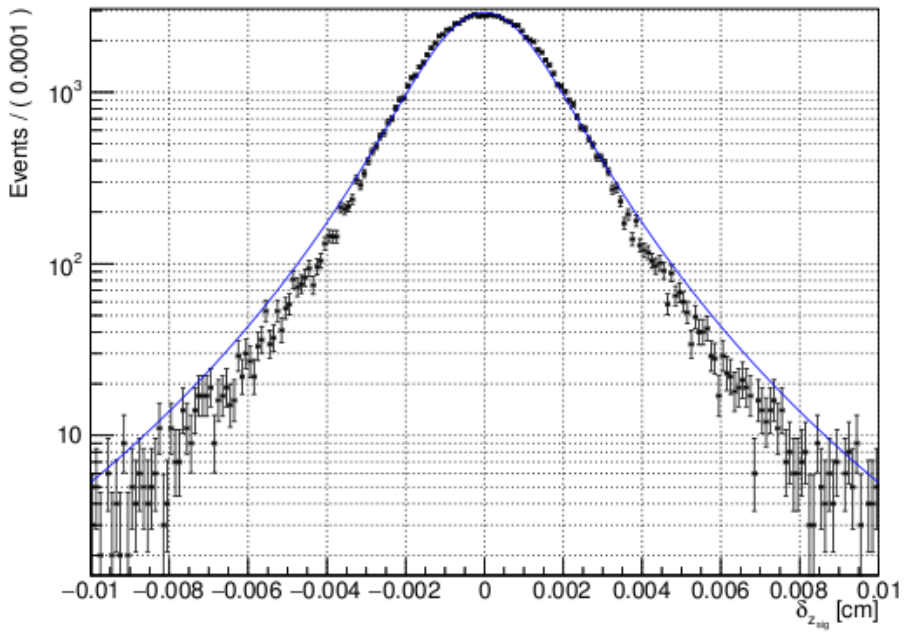


Fig. 5.4 $\delta_{z_{sig}}$ residual distribution with fit curve of signal side resolution function R_{sig}^{det} superimposed.

5.3 Tag side resolution

Modelling for tag side resolution has more complications. In addition to the detector resolution, smearing effects due to inclusion of non-primary (NP) charged tracks that originate from the decaying of long-lived particles, mainly charmed meson daughters also need to be considered. Thus, we model the tag side resolution as a convolution of contributions from detector and non-primary smearing effects:

$$R_{tag} = R_{tag}^{det} \otimes R_{tag}^{np}. \tag{5.8}$$

It is then an essential step to separate R_{tag}^{det} from R_{tag}^{np}. In order to serve this purpose, two dedicated modules with different approaches have been developed:

1) Tag vertex fitted with **truth** helix information: In this approach, the perigee or point of closest approach (POCA) of tracks are replaced with the generated (true) point in the Monte Carlo simulation, the momentum 4-vector is also replaced by truth information. The modified tracks are passed on to vertex fitter and in principle, detector resolution effects are removed in the fitted decay vertex. In another words, non-primary smearing resolution can be calculated as $R_{tag}^{np} = F(z_{tag}^{truth} - z_{tag}^{gen})$ and the

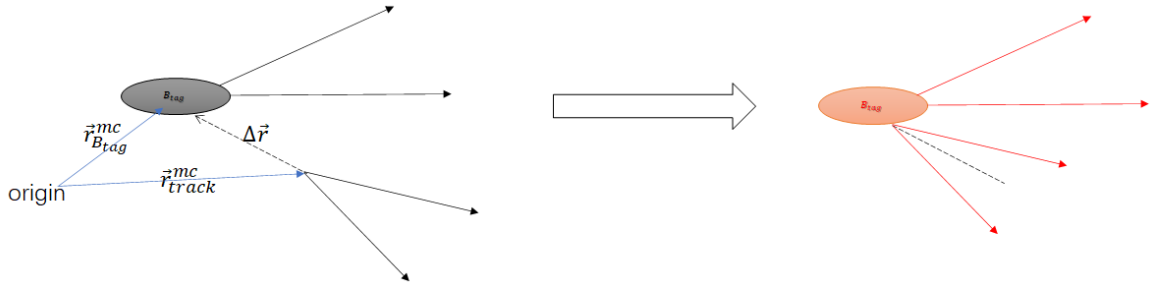


Fig. 5.5 A pictorial representation of roll-back module for tag side vertex fitting.

remaining $R^{det}_{tag} = F(z_{tag} - z^{truth}_{tag})$. Here z^{truth}_{tag} is vertex position obtained from tracks using truth helix information.

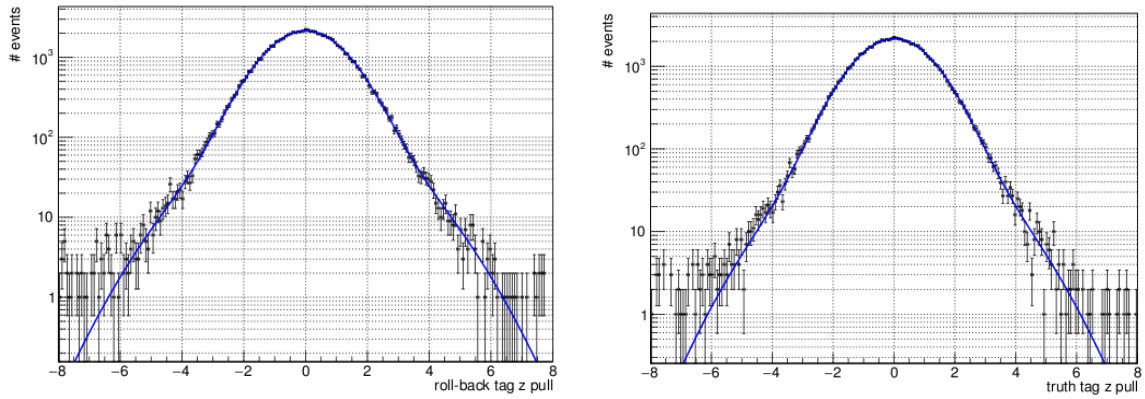
2) Tag vertex fitted with **rolled-back** tracks: Compared with the one mentioned above, this is the "opposite" approach as it removes the non-primary smearing effects in the decay vertex. In this approach, all the secondary tracks are rolled-back to the mother B with a shift

$$\Delta\vec{r}_i \overset{\text{def}}{=} \vec{r}^{\,mc}_{B_{tag,i}} - \vec{r}^{\,mc}_{track,i} \tag{5.9}$$

where $\vec{r}^{\,mc}_{track,\,i}$ is the MC generated track position in the i-th non-primary track, and $\vec{r}^{\,mc}_{B_{tag,\,i}}$ is the position vector of the MC generated decay vertex position of the corresponded B_{tag}. A pictorial representation is shown in Fig. 5.5. As flight length of non-primary particles are removed, detector resolution can be calculated as $R^{det}_{tag} = F(z^{roll-back}_{tag} - z^{gen}_{tag})$ and the remaining $R^{np}_{tag} = F(z_{tag} - z^{roll-back}_{tag})$. Here $z^{roll-back}_{tag}$ is vertex position obtained from rolled-back tracks.

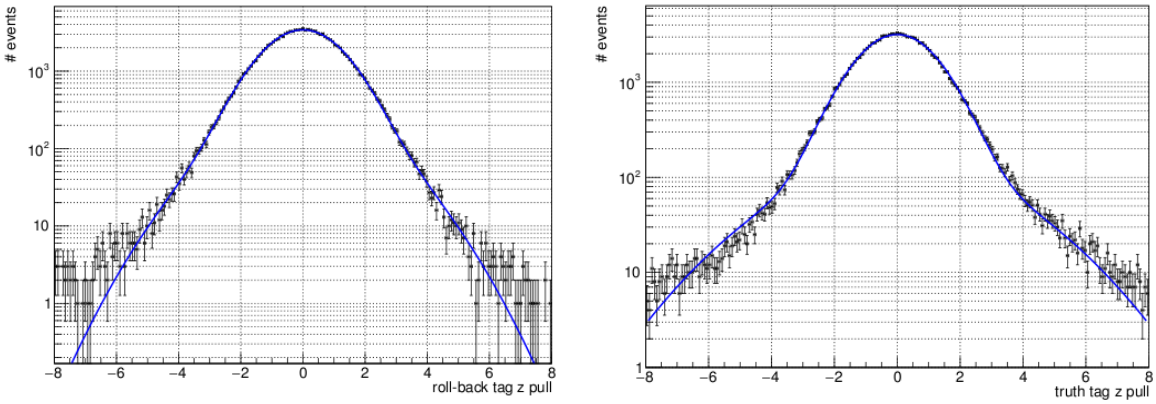
From Chapter 4, we know there are two different fitting algorithms implemented for tag vertex reconstruction. After dedicated studies, the kinematic fitter is chosen for the following reasons: The separation performance of R^{det}_{tag} is cross checked using roll-back and truth helix approaches with Kinematic fitter in Fig. 5.6 and Rave adaptive fitter in Fig. 5.7. It is observed that the fitted parameters for the fit functions are in agreement when Kinematic fitter is used, while different R^{det}_{tag} distributions are shown when Rave adaptive fitter is used. The possible explanation is that Rave assigns weight to all the tracks and in general, the weight assignment will be different when truth tracks and rolled-back tracks are used, leading to the inconsistent separation of tag side resolutions. Furthermore, in search for external parameter dependence, Kinematic fitter is more favorable with the linearly parameterized dependence

on reduced chi square as shown later in Fig. 5.8b, while Rave adaptive fitter distributes in a more complicated way, making it difficult to shape in the resolution function model.



(a) (b)

Fig. 5.6 Distribution of separated R_{tag}^{det} by (a) roll-back and (b) truth helix module using **Kinematic fitter**. Sum of two Gaussian fit functions centered at 0 are superimposed for each case.



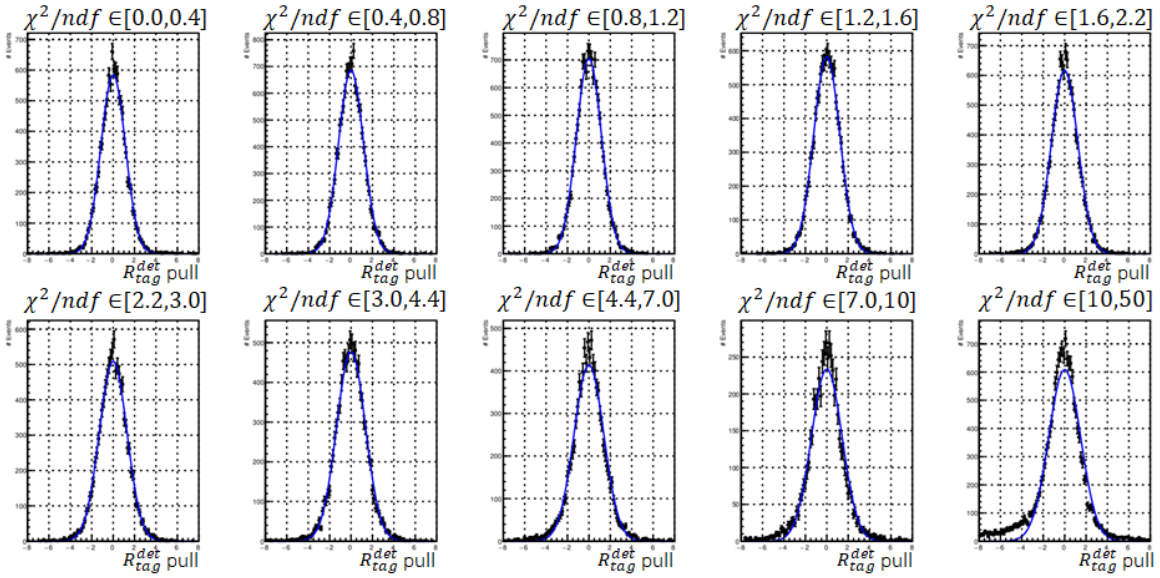
(a) (b)

Fig. 5.7 Distribution of separated R_{tag}^{det} by (a) roll-back and (b) truth helix module using **Rave adaptive fitter**. Sum of two Gaussian fit functions centered at 0 are superimposed for each case.

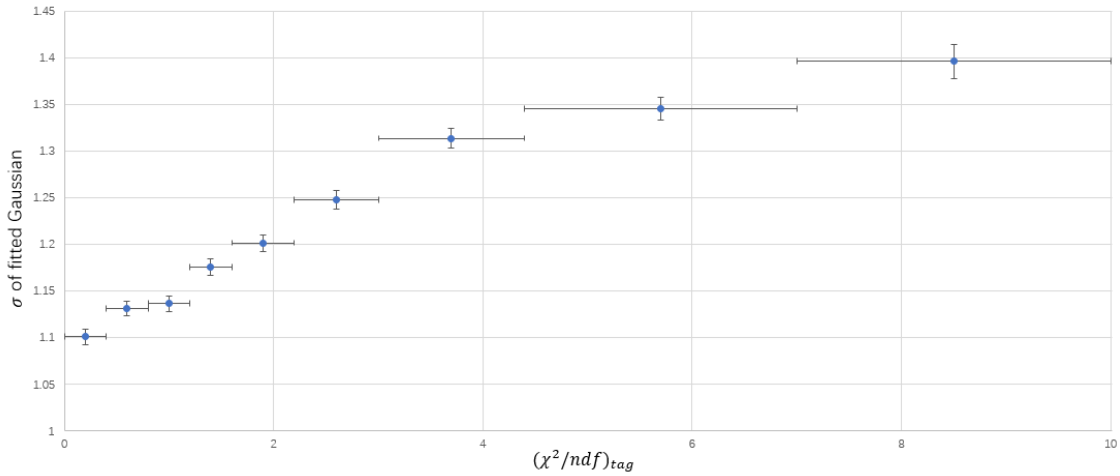
Following the similar method as signal side modelling, we define residual distribution of tag side vertex in boost direction $\delta_{z_{tag}} \stackrel{\text{def}}{=} z_{tag} - z_{tag}^{gen}$ and introduce $\sigma_{z_{tag}}$ as the uncertainty in the boost direction. Then we investigate the reduced chi-square $(\chi^2/ndf)_{tag}$ (here different from signal side, number of degree of freedom is not fixed as decay chain is general in

tag side) and search for its dependence on the standard deviation of Gaussian pull fits to $(\delta_{z_{sig}}/\sigma_{z_{sig}})$. Fig. 5.8a shows the single Gaussian fits to different slices of $(\chi^2/ndf)_{tag}$.

It is observed that when $(\chi^2/ndf)_{tag}$ is larger than some value around 8, the asymmetries between negative and positive sides in the R^{det}_{tag} pull distribution become too large to be suitable for Gaussian fits centered at 0. As a result, a cut $(\chi^2/ndf)_{tag} < 8$ is applied for tag side resolution function fitting (more details about this cut are discussed in Chapter 7). The linear dependence of standard deviation on $(\chi^2/ndf)_{tag}$ is shown in Fig. 5.8b.



(a) Single Gaussian function fits to $\delta_{z_{tag}}/\sigma_{z_{tag}}$ pull in different slices of $(\chi^2/ndf)_{tag}$.



(b) Linear dependence of the standard deviations σ of the fitted Gaussian functions on $(\chi^2/ndf)_{tag}$.

Fig. 5.8 Introduction of the event-by-event vertex fitting reduced chi-square $(\chi^2/ndf)_{tag}$ to tag side resolution function.

With the observation made above, we model the tag side detector resolution with a sum of two Gaussian fit functions

$$R_{tag}^{det}(\delta_{z_{tag}}) = (1 - f_{tag}^{tail})G\left(\delta_{z_{tag}}; s_{tag}^{main} \cdot \sigma_{z_{tag}}\right) + f_{tag}^{tail} G\left(\delta_{z_{tag}}; s_{tag}^{tail} \cdot \sigma_{z_{tag}}\right), \tag{5.10}$$

$$s_{tag} \stackrel{\text{def}}{=} s_{tag}^{0} + s_{tag}^{1} \cdot (\chi^2/ndf)_{tag}. \tag{5.11}$$

Fig. 5.9 shows results of the fit.

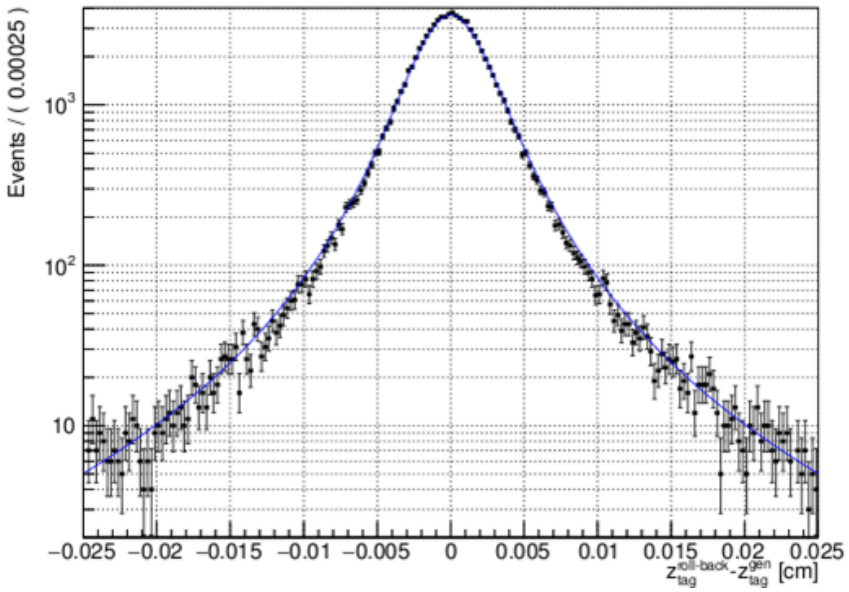


Fig. 5.9 $\delta_{z_{tag}}$ residual distribution obtained by roll-back module, with fit curve of tag side detector resolution function R_{tag}^{det} superimposed.

Now we consider the non-primary smearing effects. In the actual practice, non-primary smearing effects cannot be fully deconvoluted from detector resolution. So instead of fitting to R_{tag}^{np} separately, we observe the simulated distribution from MC (Fig. 5.10) and determine the shape to be in the form of:

$$R_{tag}^{np}(\delta z) = f_\delta \delta(\delta z) + (1 - f_\delta)\left[f_p E_p\left(\delta z; \tau_p \cdot \sigma_{z_{tag}}\right) + (1 - f_p) E_n\left(\delta z; \tau_n \cdot \sigma_{z_{tag}}\right)\right], \tag{5.12}$$

where δ is the Dirac delta function, E_p and E_n are one-sided exponential distributions defined as

$$E_p(x,\tau) \stackrel{\text{def}}{=} \frac{1}{\tau} exp(-\frac{x}{\tau})\ for\ x > 0,\ \text{otherwise } 0, \tag{5.13}$$

$$E_n(x,\tau) \stackrel{\text{def}}{=} \frac{1}{\tau} exp(+\frac{x}{\tau})\ for\ x \leq 0,\ \text{otherwise } 0. \tag{5.14}$$

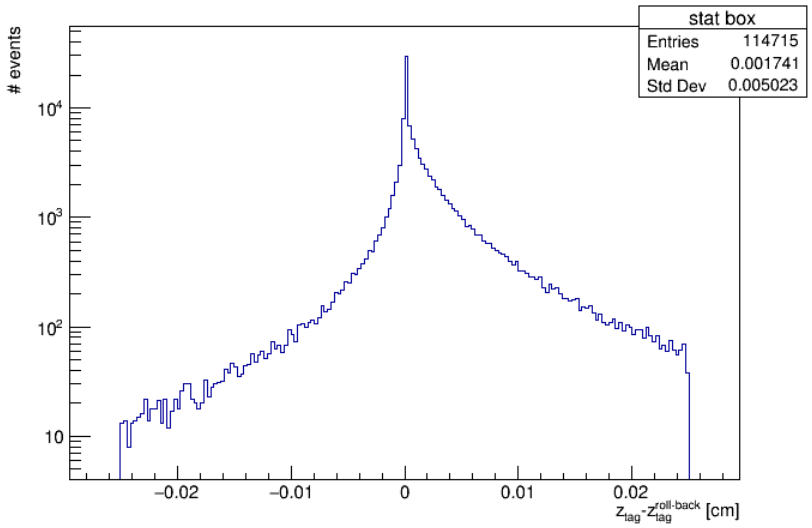


Fig. 5.10 R_{tag}^{np} distribution obtained by roll-back module.

As the non-primary smearing effects mainly come from the finite lifetime of charm daughters, the lifetime (exponential) form in R_{tag}^{np} comes naturally.

The parameters of R_{tag}^{np} are determined from the fit to total $R_{tag}^{det} \otimes R_{tag}^{np}$ with the model and parameters of R_{tag}^{det} fixed from truth/rolled-back module. Fig. 5.11 shows the result of the combined tag side residual fit.

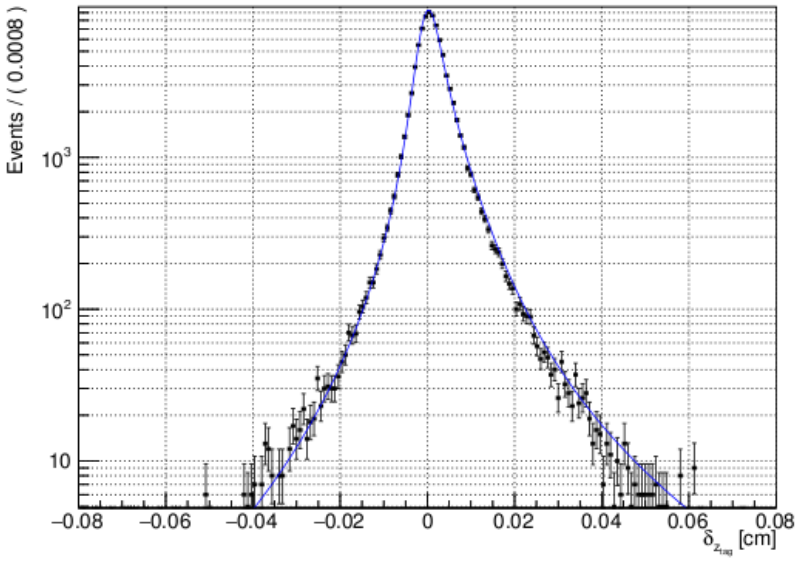


Fig. 5.11 $\delta_{z_{tag}}$ residual distribution with fit curve of convolution of tag side detector and non-primary smearing resolution function $R_{tag}^{det} \otimes R_{tag}^{np}$ superimposed.

5.4 Kinematic smearing

During the calculation of decay time difference from reconstructed decay length difference, the approximation $\Delta t \approx \left(1/c(\beta\gamma)_{\Upsilon(4S)}\right) \cdot (z_{sig} - z_{tag})$ was used. The approximation is valid based on the assumption that momentum of B mesons in the center-of-mass frame is small compared with momentum of $\Upsilon(4S)$. In reality, for better precision of parameterizing the Δt resolution function, we consider the low but non-negligible momentum of B mesons which gives rise to the kinematic smearing resolution R_k.

We denote Δt^{nom} to be the nominal time difference calculated from the true decay length by

$$\begin{aligned}\Delta t^{nom} &= \left(1/c(\beta\gamma)_{\Upsilon(4S)}\right) \cdot (z_{sig} - z_{tag}) \\ &= \frac{c(\beta\gamma)_{B_{sig}} t^{gen}_{sig} - c(\beta\gamma)_{B_{tag}} t^{gen}_{tag}}{c(\beta\gamma)_{\Upsilon(4S)}} \\ &= \frac{(\beta\gamma)_{B_{sig}}}{(\beta\gamma)_{\Upsilon(4S)}} t^{gen}_{sig} - \frac{(\beta\gamma)_{B_{tag}}}{(\beta\gamma)_{\Upsilon(4S)}} t^{gen}_{tag}.\end{aligned} \tag{5.15}$$

Here $(\beta\gamma)_{B_{sig}}$ and $(\beta\gamma)_{B_{tag}}$ are the Lorentz boost factors generated by the signal and tag side B meson. They are related to $(\beta\gamma)_{\Upsilon(4S)}$ by a change of basis from lab frame to center-of-mass frame of energy-momentum four vector:

$$\begin{aligned}(\beta\gamma)_{B_{sig}} &= \frac{(\beta\gamma)_{\Upsilon(4S)} E^{cms}_{B_{sig}}}{m_B c^2} + \frac{(\gamma)_{\Upsilon(4S)} p^{cms}_{B_{sig}} cos\ \theta^{cms}_{B_{sig}}}{m_B c} \\ &\stackrel{\text{def}}{=} (\beta\gamma)_{\Upsilon(4S)}(a_k + c_k),\end{aligned} \tag{5.16}$$

$$\begin{aligned}(\beta\gamma)_{B_{tag}} &= \frac{(\beta\gamma)_{\Upsilon(4S)} E^{cms}_{B_{sig}}}{m_B c^2} - \frac{(\gamma)_{\Upsilon(4S)} p^{cms}_{B_{sig}} cos\ \theta^{cms}_{B_{sig}}}{m_B c} \\ &\stackrel{\text{def}}{=} (\beta\gamma)_{\Upsilon(4S)}(a_k - c_k).\end{aligned} \tag{5.17}$$

Here, $E^{cms}_{B_{sig}}$, $p^{cms}_{B_{sig}}$ and $\theta^{cms}_{B_{sig}}$ are the energy, momentum and angle with respect to the beam direction of the signal side B meson in the center-of-mass frame.

Then we can express the difference between Δt^{nom} and Δt^{gen} as

$$\Delta t^{nom} - \Delta t^{gen} = (a_k + c_k - 1) t^{gen}_{sig} - (a_k - c_k - 1) t^{gen}_{tag}. \tag{5.18}$$

The probability density function to obtain Δt^{gen} around certain value is

$$F(\Delta t^{gen}) = \iint dt_{sig}^{gen} dt_{tag}^{gen} F_{sig}(t_{sig}^{gen}) F_{tag}(t_{tag}^{gen}) \delta\left(\Delta t^{gen} - \left(t_{sig}^{gen} - t_{tag}^{gen}\right)\right) \tag{5.19}$$

The probability density function to simultaneously obtain $\Delta t^{nom} - \Delta t^{gen}$ and Δt^{gen} is

$$\begin{aligned} F(\Delta t^{nom} - \Delta t^{gen}, \Delta t^{gen}) = \iint & dt_{sig}^{gen} dt_{tag}^{gen} F_{sig}(t_{sig}^{gen}) F_{tag}(t_{tag}^{gen}) \delta\left(\Delta t^{gen} - \left(t_{sig}^{gen} - t_{tag}^{gen}\right)\right) \\ & \delta\left((\Delta t^{nom} - \Delta t^{gen}) - ((a_k + c_k - 1) t_{sig}^{gen} - (a_k - c_k - 1) t_{tag}^{gen})\right). \end{aligned} \tag{5.20}$$

Here δ is the Dirac delta function and F_{sig} and F_{tag} are the probability density functions for t_{sig}^{gen} and t_{tag}^{gen}:

$$P_{sig}\left(t_{sig}^{gen}\right) = E_p(t_{sig}^{gen}, \tau_B), \tag{5.21}$$

$$P_{tag}\left(t_{tag}^{gen}\right) = E_p(t_{tag}^{gen}, \tau_B). \tag{5.22}$$

The kinematic smearing resolution generated by $\Delta t^{nom} - \Delta t^{gen}$ is the conditional probability to obtain $\Delta t^{nom} - \Delta t^{gen}$ at given Δt^{gen}:

$$\begin{aligned} R_k(\Delta t^{nom} - \Delta t^{gen}) &= \frac{P(\Delta t^{nom} - \Delta t^{gen}, \Delta t^{gen})}{P(\Delta t^{gen})} \\ &= \begin{cases} E_p\{(\Delta t^{nom} - \Delta t^{gen}) - [(a_k - 1)\Delta t^{gen} + c_k|\Delta t^{gen}|]; |c_k|\tau_B\}, & \text{if } c_k > 0 \\ \delta[(\Delta t^{nom} - \Delta t^{gen}) - (a_k - 1)\Delta t^{gen}], & \text{if } c_k = 0 \\ E_n\{(\Delta t^{nom} - \Delta t^{gen}) - [(a_k - 1)\Delta t^{gen} + c_k|\Delta t^{gen}|]; |c_k|\tau_B\}, & \text{if } c_k < 0 \end{cases} \end{aligned} \tag{5.23}$$

The distribution of R_k is plotted in Fig. 5.12 and its angular dependence is shown in Fig. 5.13.

In the actual measurement, the true decay time difference Δt^{gen} is not available. In order to account for kinematic smearing, one can perform a convolution of the true signal probability density function $P_{sig}(\Delta t^{gen})$ with $R_k(\Delta t - \Delta t^{gen})$ to obtain the smeared theoretical Δt distribution. For example, in the lifetime measurement $P_{sig}(\Delta t^{gen})$ given by (2.37), the resulting probability density function is

$$P_{sig \circledast k}(\Delta t) \stackrel{\text{def}}{=} \int d(\Delta t') P_{sig}(\Delta t') R_k(\Delta t - \Delta t') = \begin{cases} \frac{1}{2a_k\tau_B} exp\left[-\frac{|\Delta t|}{(a_k + c_k)\tau_B}\right], & \text{if } \Delta t \geq 0 \\ \frac{1}{2a_k\tau_B} exp\left[-\frac{|\Delta t|}{(a_k - c_k)\tau_B}\right], & \text{if } \Delta t < 0 \end{cases} \tag{5.24}$$

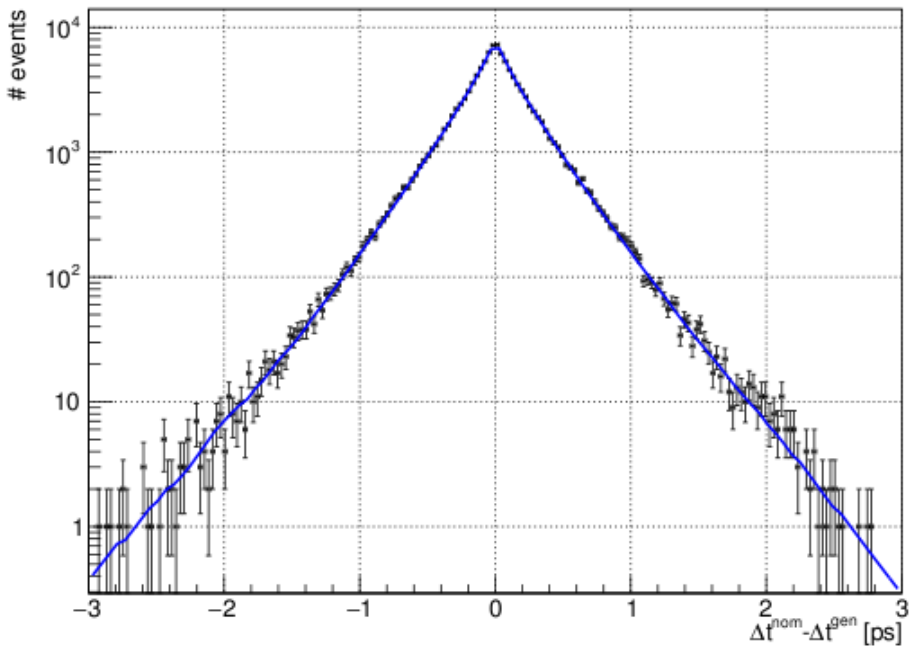
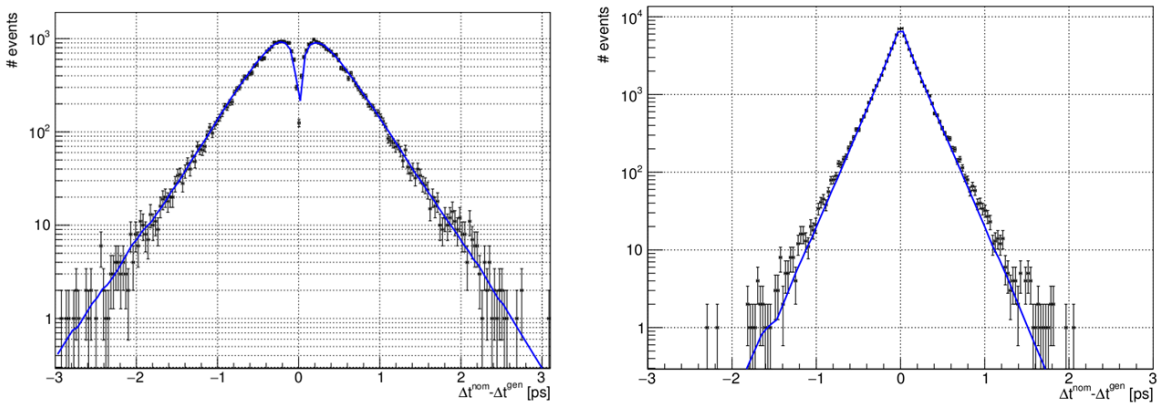


Fig. 5.12 R_k distribution with exact fit curve in (5.23) superimposed.



(a) (b)

Fig. 5.13 R_k distribution in the regions when flight direction of signal B meson and boost direction in center-of-mass frame are (a) colinear, $0° \leq \theta^{cms}_{B_{sig}} < 60°$ or $120° < \theta^{cms}_{B_{sig}} \leq 180°$ (b) perpendicular, $60° \leq \theta^{cms}_{B_{sig}} \leq 120°$, with exact fit curves superimposed.

As kinematic smearing resolution is purely analytical, no extra parameters is introduced to the resolution function model. In summary, a model with a total number of 14 floating parameters is constructed in this chapter. The fitted values are summarized in Table 5.1.

Parameters	Values
f_{sig}^{tail}	0.113 ± 0.012
$(s_{sig}^{main})^0$	0.976 ± 0.007
$(s_{sig}^{main})^1$	0.064 ± 0.005
$(s_{sig}^{tail})^0$	1.368 ± 0.034
$(s_{sig}^{tail})^1$	0.376 ± 0.028
f_{tag}^{tail}	0.101 ± 0.008
$(s_{tag}^{main})^0$	1.079 ± 0.006
$(s_{tag}^{main})^1$	0.039 ± 0.002
$(s_{tag}^{tail})^0$	1.712 ± 0.038
$(s_{tag}^{tail})^1$	0.276 ± 0.016
f_δ	0.595 ± 0.005
f_p	0.880 ± 0.005
τ_n^0	2.270 ± 0.071
τ_p^0	1.971 ± 0.020

Table 5.1 Fitted parameters of the resolution function using 500k simulated $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ signal samples.

Chapter 6

Extraction of proper lifetime

6.1 Lifetime fit to signal MC sample

In this section, lifetime fit to large number of simulated $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ signal events is performed, as validation of the performance of the constructed Δt resolution function. The number of simulated B^0-$\bar{B}^0$ pairs to the signal decays is 500k, which corresponds to 1.6 ab^{-1} integrated luminosity.

In the absence of background events, the measured probability density function of Δt distribution is the convolution of theoretical Δt distribution and the resolution function:

$$P(\Delta t) = \int d(\Delta t') P_{sig}(\Delta t') R(\Delta t - \Delta t'). \tag{6.1}$$

Here P_{sig} is described in (2.37). The time-dependent lifetime fit is performed by maximizing the following unbinned likelihood function constructed in RooFit [26]:

$$\begin{aligned}
\mathscr{L}(\Delta t; \tau_B) &= \prod_{i=1}^{N} P^i(\Delta t) \\
&= \prod_{i=1}^{N} \iiint d(\Delta t') d(\Delta t'') d(\Delta t''') P^i_{sig}(\Delta t - \Delta t') R^i_k(\Delta t' - \Delta t'') R^i_{sig}(\Delta t'' - \Delta t''') R^i_{tag}(\Delta t''') \\
&= \prod_{i=1}^{N} \iint d(\Delta t') d(\Delta t'') P^i_{sig\otimes k}(\Delta t - \Delta t') R^i_{sig}(\Delta t' - \Delta t'') R^i_{tag}(\Delta t'').
\end{aligned} \tag{6.2}$$

Here N is the total number of measurements. $P^i_{sig\otimes k}$, R^i_{sig}, R^i_{tag} are kinematic smeared theoretical Δt distribution, signal side resolution function and tag side resolution function for each event, respectively.

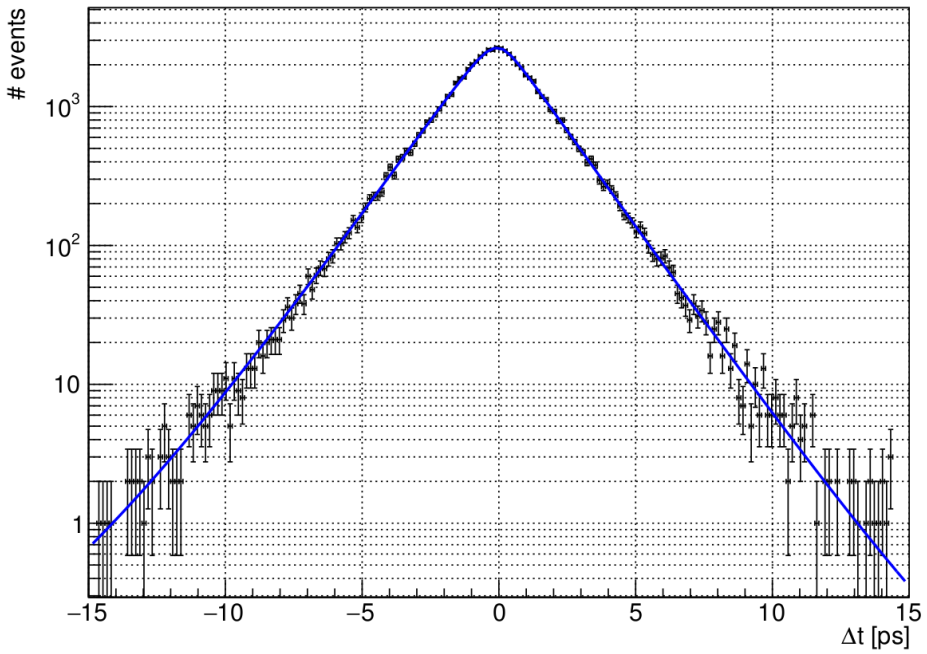


Fig. 6.1 Distribution of proper-time difference Δt of neutral B mesons in 500k simulated $B^0 \to [D^- \to K^+\pi^-\pi^-]\pi^+$ signal samples. The fit curve corresponding to lifetime fit is superimposed.

The proper lifetime fit result of τ_B to the signal MC is 1.5274 ± 0.0059 ps, approximately 0.4σ deviated from input value (1.525 ps). The fitted proper-time difference distribution is shown in Fig. 6.1.

6.2 Lifetime fit to generic MC sample

In this section, we perform lifetime fit to 400 fb^{-1} generic MC sample using the reconstructed $B^0 \to [D^- \to K^+\pi^-\pi^-]\pi^+$ events. The main difference with signal only MC sample is the inclusion of backgrounds originating from both beam collisions and products of various other decay chains. The measured Δt probability density function is modified into the extended form:

$$P_{tot}(\Delta t) = f_{sig} \int d(\Delta t') P_{sig}(\Delta t') R(\Delta t - \Delta t') + (1 - f_{sig}) P_{bkg}(\Delta t). \tag{6.3}$$

Here the additional components are the signal probability f_{sig} and background probability density function P_{bkg}. The extended likelihood function can then be constructed as

$$\mathscr{L}_{tot}(\Delta t;\tau_B) = \prod_{i=1}^{N} P^i_{tot}(\Delta t). \tag{6.4}$$

6.2.1 Signal probability

The signal probability f_{sig} is calculated based on the measured M_{bc} for each event. The M_{bc} distribution (Fig. 6.2) is fitted with the sum of 1 signal and 4 background components:

1) **Signal $\mathbf{F_{sig}}$**: The signal events are correctly reconstructed events matching the MC truth information. The signal function is described by a single Gaussian centered around 5.28 GeV/c^2.

$$F_{sig} = G(M_{bc} - \mu_{M_{bc}};\sigma_{M_{bc}}). \tag{6.5}$$

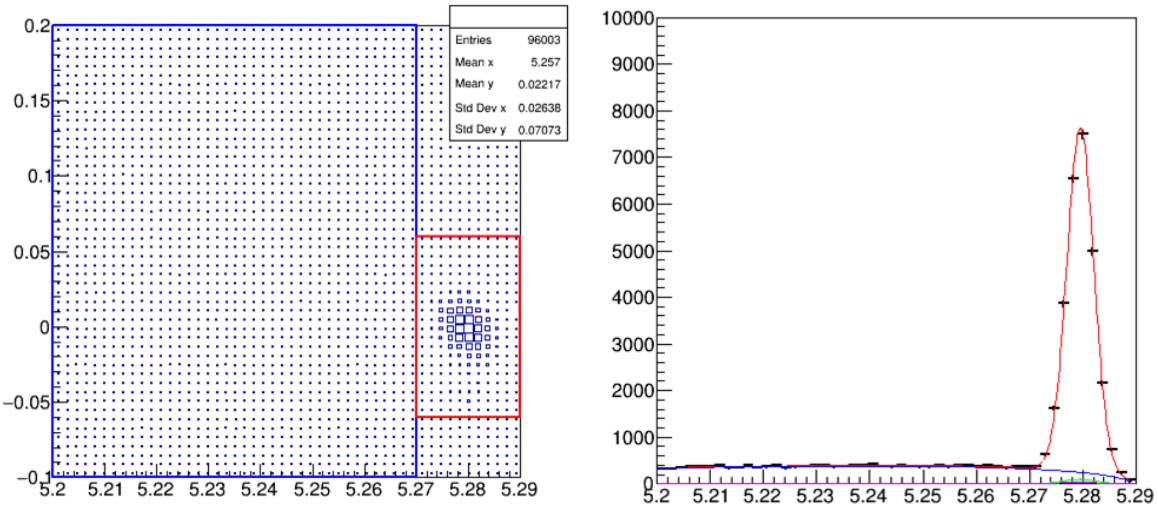
2) **Self-cross-feed (SCF) background $\mathbf{F_{SCF}}$**: SCF background refers to events reconstructed from final state particles of the MC truth matched target decay chain. But some of particles correspond with the other neutral B meson in the pair production. The SCF background distribution is parameterized by smoothed histograms, a quadric radial function fitting.

3) **Mixed background $\mathbf{F_{mixed}}$**: Mixed background refers to the rest of background events originated from neutral $B^0 - \bar{B}^0$ pairs but reconstructed from some final state particles in other decay chains. The mixed background distribution is also parameterized by smoothed histograms fitting.

4) **Charged background $\mathbf{F_{charged}}$**: Charged background refers to events reconstructed from charged B^+B^- pairs. $F_{charged}$ is similarly parameterized by smoothed histograms fitting.

5) **$\mathbf{q\bar{q}}$ background $\mathbf{F_{q\bar{q}}}$**: $q\bar{q}$ background refers to background from light quarks $e^+e^- \to q\bar{q}$ (uds). The shape is described by ARGUS function:

$$F_{q\bar{q}} = Argus(M_{bc};c) \stackrel{\text{def}}{=} N \cdot M_{bc} \cdot \sqrt{1-(\frac{M_{bc}}{E^{cms}_{beam}})^2}\, exp\left\{c \cdot \left[1-\left(\frac{M_{bc}}{E^{cms}_{beam}}\right)^2\right]\right\}. \tag{6.6}$$



(a) (b)

Fig. 6.2 (a) M_{bc}-ΔE distribution and the definition of signal (red) and background dominant (blue) region. (b) Distribution for M_{bc}, fit function for each component is superimposed.

Using the shape determined separately for each component, the signal probability for each event is calculated based on the measured M_{bc}:

$$f_{sig}(M_{bc}) = \frac{F_{sig}(M_{bc})}{F_{sig}(M_{bc}) + F_{SCF}(M_{bc}) + F_{mixed}(M_{bc}) + F_{charged}(M_{bc}) + F_{q\bar{q}}(M_{bc})} \tag{6.7}$$

The average signal probability inside the signal dominant region is ~ 0.9147 in the generic MC data set.

6.2.2 Background shape

The proper-time difference distribution of background events is obtained from the Mbc background dominant region (sideband region in Fig. 6.2a, $M_{bc} < 5.27$ GeV/c^2). The background shape is fitted with double Gaussian and the convolution with exponential as followed:

$$P_{bkg}(\Delta t) = \int d(\Delta t') \mathscr{P}_{bkg}(\Delta t') R_{bkg}(\Delta t' - \Delta t) \tag{6.8}$$

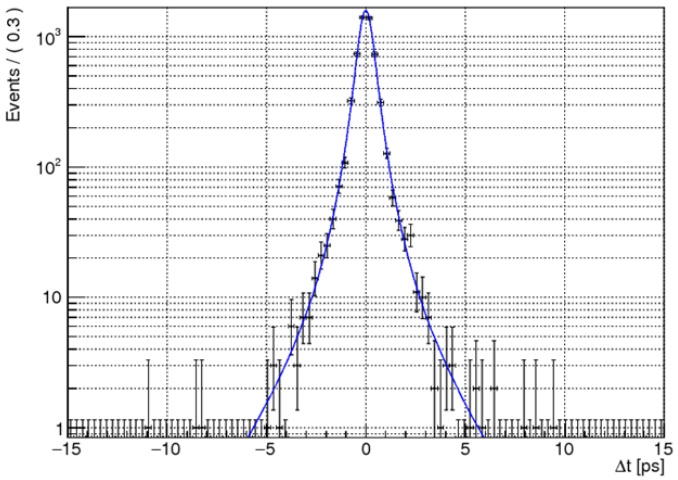


Fig. 6.3 The Δt distribution in background dominant region using $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ events reconstructed from 400 fb^{-1} simulated generic sample. The fit curve corresponding to background shape is superimposed.

where the probability density function $\mathscr{P}_{bkg}$ and background resolution function $R_{bkg}(\Delta t)$ have the form:

$$\mathscr{P}_{bkg} = f^{\delta}_{bkg}\delta(\Delta t - \mu_{bkg}) + (1 - f^{\delta}_{bkg})\frac{1}{2\tau_{bkg}} exp\left(-\frac{|\Delta t|}{\tau_{bkg}}\right), \tag{6.9}$$

$$R_{bkg}(\Delta t) = (1 - f^{tail}_{bkg})G\left(\Delta t; s^{main}_{bkg}\frac{1}{\beta\gamma c}\sqrt{\sigma^2_{z_{sig}} + \sigma^2_{z_{tag}}}\right) + f^{tail}_{bkg}G\left(\Delta t; s^{tail}_{bkg}\frac{1}{\beta\gamma c}\sqrt{\sigma^2_{z_{sig}} + \sigma^2_{z_{tag}}}\right). \tag{6.10}$$

The Δt distribution, superimposed with the fitted curve is shown in Fig. 6.3. The parameters after fitting are listed in Table 6.1.

Parameters	Values
f^{δ}_{bkg}	0.978 ± 0.009
μ_{bkg}	-0.001 ± 0.006
τ_{bkg}	1.741 ± 0.356
f^{tail}_{bkg}	0.195 ± 0.052
s^{main}_{bkg}	1.171 ± 0.045
s^{tail}_{bkg}	2.550 ± 0.318

Table 6.1 Fitted parameters of the background shape using $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ events reconstructed from 400 fb^{-1} simulated generic sample.

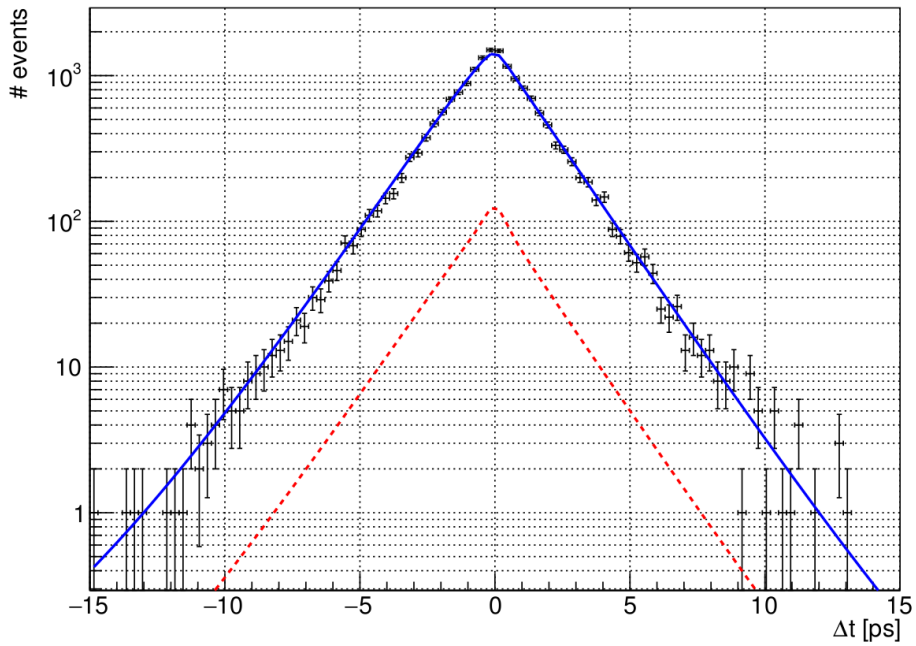


Fig. 6.4 Distribution of proper-time difference Δt of neutral B mesons in 400 fb^{-1} simulated $B^0 \to [D^- \to K^+\pi^-\pi^-]\pi^+$ generic samples. The fit curve corresponding to lifetime fit is superimposed. The red dashed line corresponds to background contribution.

With the signal and background components determined from above, the proper lifetime fit result of τ_B to the generic MC is 1.5339 ± 0.0141 ps, approximately 0.6σ deviated from input value (1.525 ps). The fitted proper-time difference distribution is shown in Fig. 6.4.

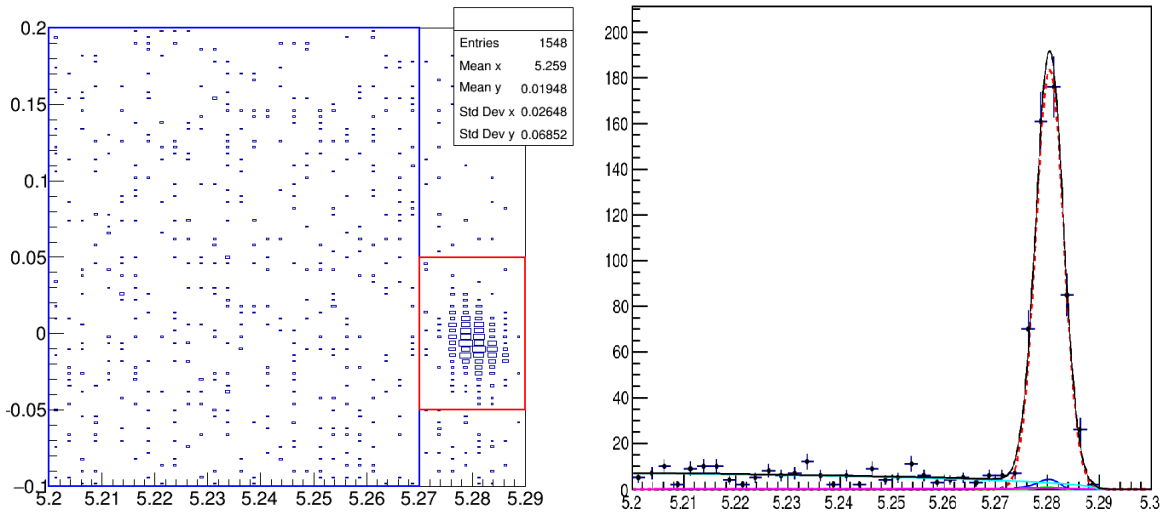
6.3 Lifetime fit to experimental data

In this section, lifetime fit to 8.7 fb^{-1} experimental data collected in Belle II from physics run in 2019 is demonstrated.

To calculate the signal fraction, the M_{bc} shape of signal, self-cross-feed background, mixed background and charged background components are fixed to the parameterization determined from generic MC study, while $q\bar{q}$ background parameters are left floating and fitted with experimental data (shown in Fig. 6.5).

The background Δt shape is determined from the defined background dominant region using the same model described in Section 6.2.2 fitted with experimental data. The fitted curve is shown in Fig. 6.6 and the fitted parameters are listed in Table 6.2.

To reduce the systematic uncertainty from the dependence on MC, the detector resolution components on both signal and tag side are fitted with B^0 lifetime simultaneously instead of



(a) (b)

Fig. 6.5 (a) M_{bc}-ΔE distribution and the definition of signal (red) and background dominant (blue) region. (b) Distribution for M_{bc}, fit function for each components is superimposed in the reconstructed $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ events from 8.7 fb^{-1} experimental data.

Parameters	Values
f_{bkg}^{δ}	0.938 ± 0.065
μ_{bkg}	-0.070 ± 0.047
τ_{bkg}	1.442 ± 1.010
f_{bkg}^{tail}	0.256 ± 0.603
s_{bkg}^{main}	1.451 ± 2.424
s_{bkg}^{tail}	1.333 ± 1.754

Table 6.2 Fitted parameters of the background shape using $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ events reconstructed from 8.7 fb^{-1} experimental data.

fixing from MC determined values. Furthermore, the small statistics of events in early phase data taking ($\sim$ 8.7 fb^{-1}) compared with large number MC simulation ($\sim$ 1.6 ab^{-1}) could cause effects of overfitting, making some components of the resolution function model lose their physics meaning. Thus, a simplified model for detector resolution in (5.6) and (5.10) is applied here

$$R_{sig}^{det}(\delta_{z_{sig}}) = G\left(\delta_{z_{sig}}; s_{sig} \cdot \sigma_{z_{sig}}\right), \tag{6.11}$$

$$R_{tag}^{det}(\delta_{z_{tag}}) = G\left(\delta_{z_{tag}}; s_{tag} \cdot \sigma_{z_{sig}}\right). \tag{6.12}$$

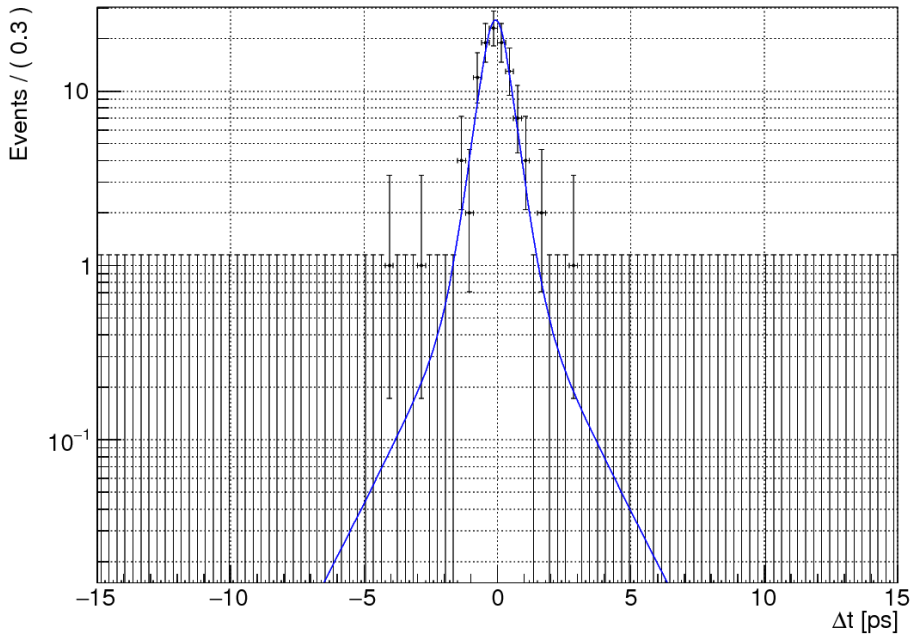


Fig. 6.6 The Δt distribution in background dominant region using $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ events reconstructed from 8.7 fb^{-1} experimental data. The fit curve corresponding to background shape is superimposed.

Parameters	Values
$\tau_B[ps]$	1.454 ± 0.097
s^0_{sig}	2.140 ± 1.814
s^1_{sig}	0.000 ± 0.345
s^0_{tag}	0.454 ± 0.707
s^1_{tag}	0.150 ± 0.141

Table 6.3 Fitted parameters of the lifetime fit using $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ events reconstructed from 8.7 fb^{-1} experimental data.

Finally, with all the pieces, the lifetime extraction is performed by fitting the extended likelihood function as in (6.4) to Δt distribution. The fitted lifetime along with parameters in the detector resolution is listed in Table 6.3. The result 1.454 ± 0.097 ps is consistent with the current world average measurement [27] of B^0 lifetime (1.519 ± 0.004 ps). The Δt distribution, superimposed with the fitted curve is shown in Fig. 6.7.

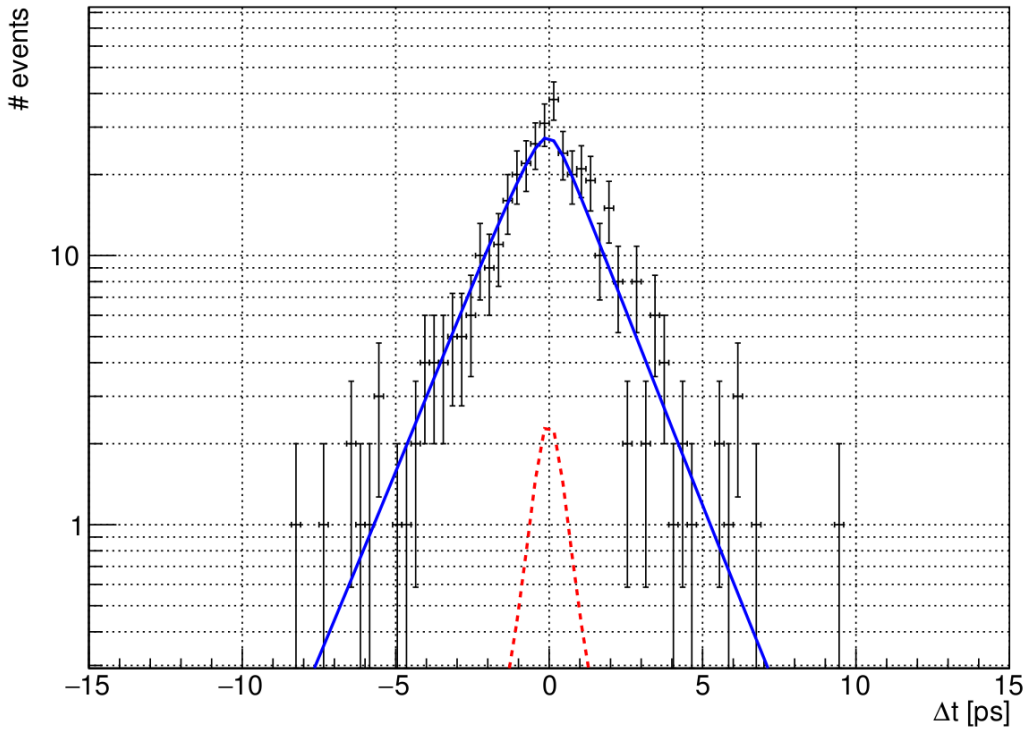


Fig. 6.7 Distribution of proper-time difference Δt of neutral B mesons in 8.7 fb^{-1} $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ experimental data. The fit curve corresponding to lifetime fit is superimposed. The red dashed line corresponds to background contribution.

Chapter 7

Discussions and conclusion

7.1 Linearity test

One of the ways to validate the fitting procedures to obtain lifetime is through linearity test, by investigating the response to different generator input values. In this test, different MC samples with varied input B^0 proper lifetime and mass values are generated. Each sample consists of 10,000 $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ pure signal events.

By performing the fitting procedures described in detail before, the dependence of deviation of fit and input values on generator values of B^0 lifetime and mass is plotted in Fig. 7.1 and Fig. 7.2.

This test confirms the linear response of time dependent lifetime fit.

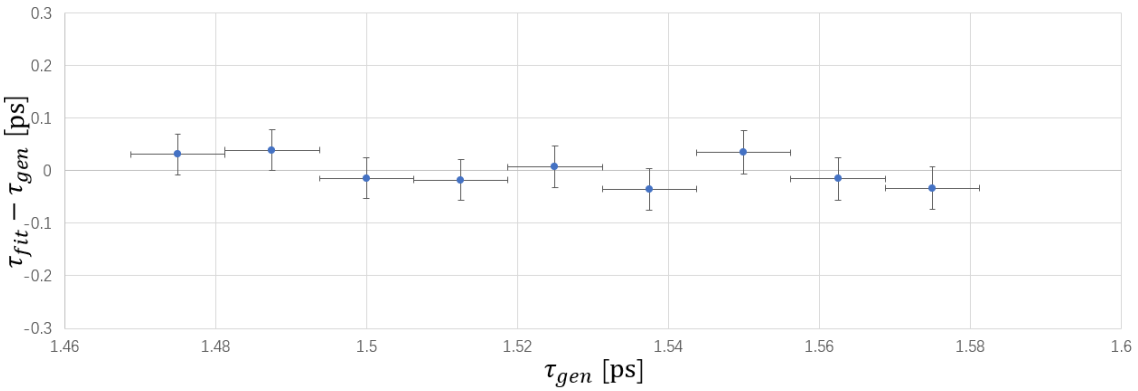


Fig. 7.1 Linearity test of fitted proper B^0 lifetime with shifted input B^0 lifetime.

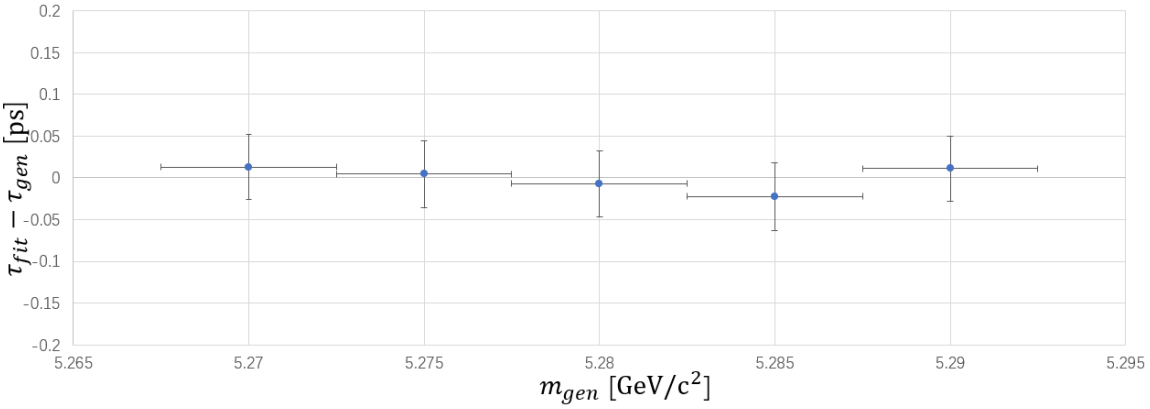


Fig. 7.2 Linearity test of fitted proper B^0 lifetime with shifted input B^0 mass.

7.2 Validation of reduced χ^2 in the resolution function model

In the model for resolution function we have developed, extra variables are introduced based on the observed linear dependence of standard deviation of detector resolution. These variables are the sig side and tag side reduced chi square χ^2/ndf.

It is necessary to check there is no correlation between χ^2/ndf and decay time. By extracting the generator information we can plot Fig. 7.3 that shows the fitted proper lifetime distributions in different subsets of $(\chi^2/ndf)_{sig}$ and Fig. 7.4 that shows distributions in different subsets of $(\chi^2/ndf)_{tag}$. We conclude that no visible correlation is observed between χ^2/ndf and decay time.

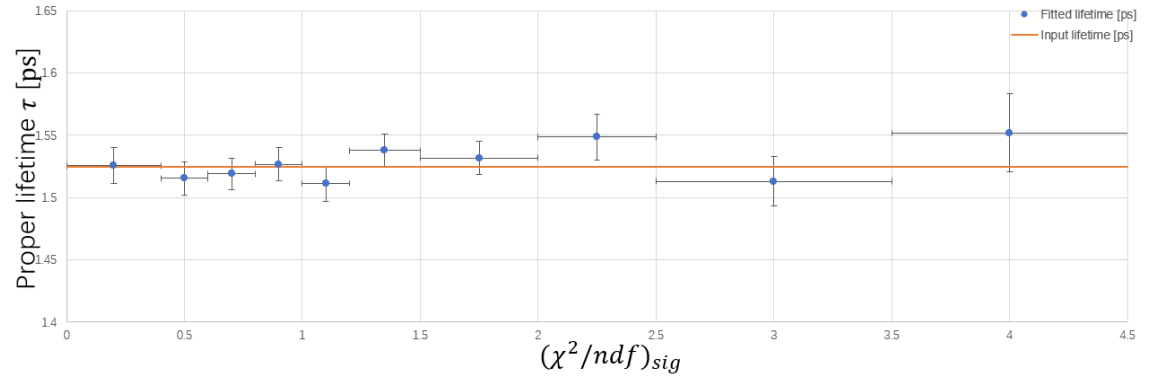


Fig. 7.3 The fitted proper lifetime distributions in different subsets of $(\chi^2/ndf)_{sig}$.

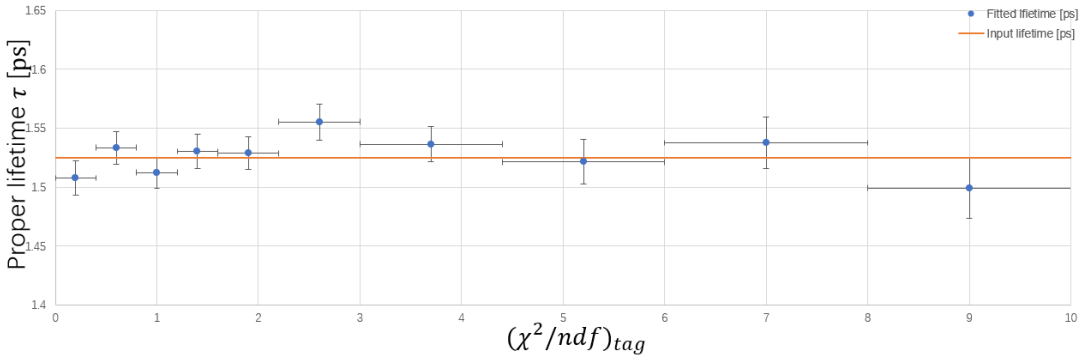
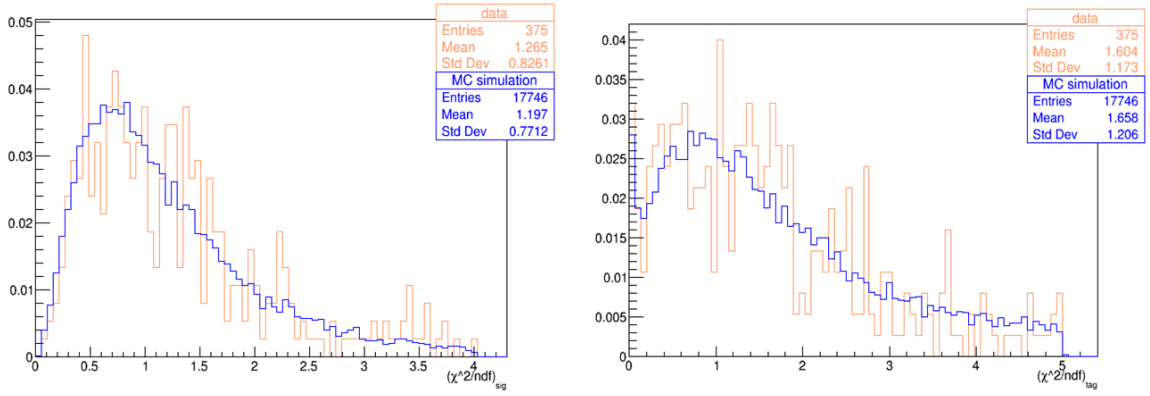


Fig. 7.4 The fitted proper lifetime distributions in different subsets of $(\chi^2/ndf)_{tag}$.

In the next step, the data-MC discrepancy in χ^2/ndf is investigated. Fig. 7.5 shows the event count distributions of χ^2/ndf from 400 fb^{-1} generic MC sample and 8.7 fb^{-1} Belle II data collected in 2019. Consistency in distribution is shown between data and MC.



(a) (b)

Fig. 7.5 Data-MC comparison of (a) $(\chi^2/ndf)_{sig}$ and (b) $(\chi^2/ndf)_{tag}$ using 8.7 fb^{-1} Belle II data collected in 2019.

In Chapter 5, we have seen that Gaussian fit function cannot describe tag side detector resolution well for events with poor fitting quality. Here we plot the distribution of R_{tag}^{det} in different subsets of fitting quality indicator $(\chi^2/ndf)_{tag}$ in Fig. 7.6.

As such poorly reconstructed events are not favorable for the fitting model, an extra selection $(\chi^2/ndf)_{tag} < 8$ is chosen and applied to all events. However, by doing so, the total number of useful events is compromised. From the study of simulated MC samples, approximately 77.5% of the events pass this quality selection.

The impact of this cut on signal-to-background ratio is also studied. The comparison of M_{bc} distribution before and after cut is plotted in Fig. 7.7. We observe that for each signal or background component, the rejection rate of the extra selection is similar.

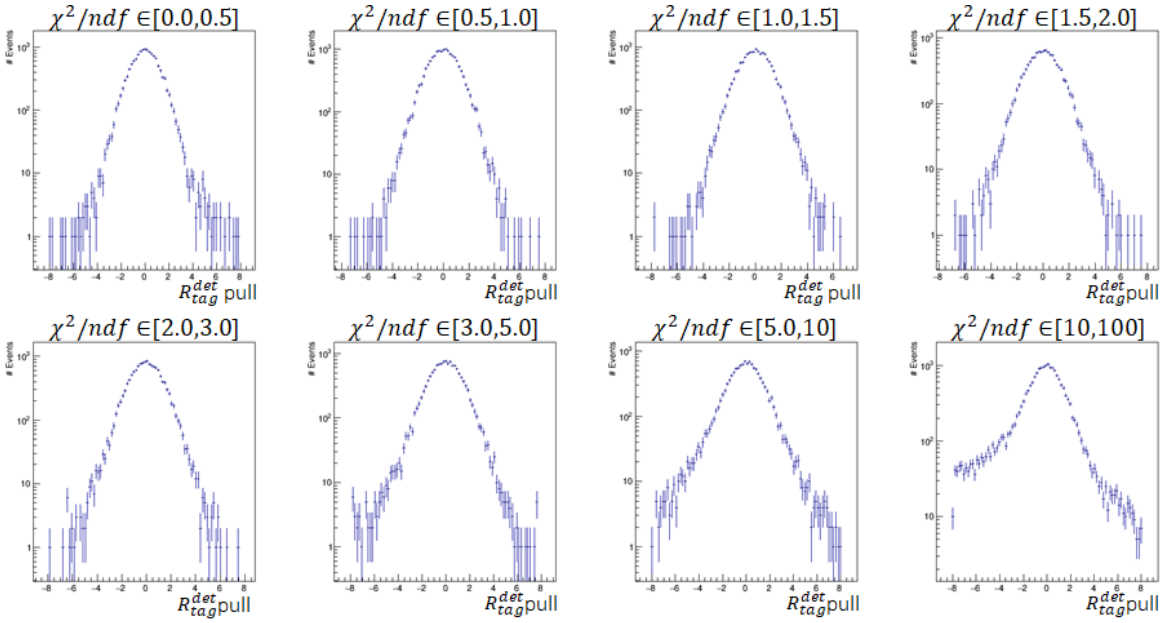
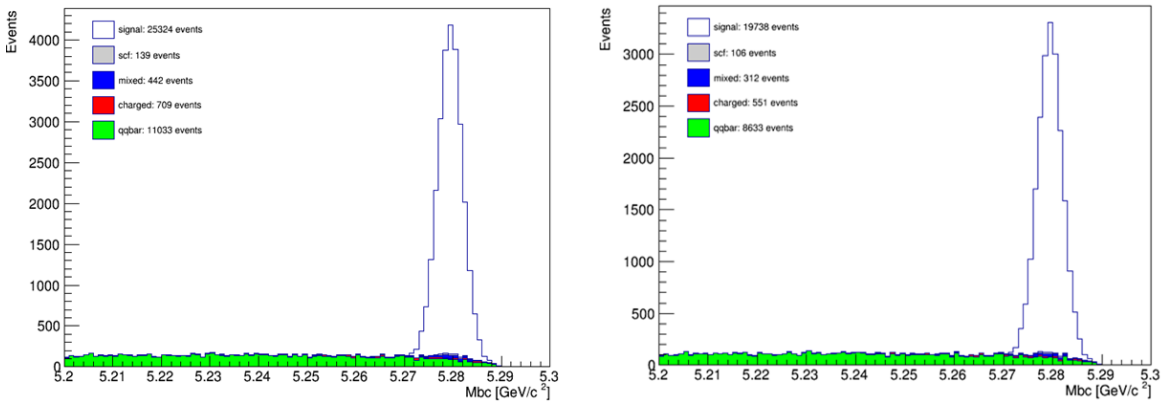


Fig. 7.6 R^{det}_{tag} distributions in different subsets of $(\chi^2/ndf)_{tag}$.



(a) (b)

Fig. 7.7 M_{bc} distribution (a) before applying cut $(\chi^2/ndf)_{tag} < 8$ and (b) after applying cut $(\chi^2/ndf)_{tag} < 8$.

In summary, χ^2/ndf shows consistent distribution between data-MC and is valid to be introduced as extra event-by-event dependence for the decay time resolution function model. However, at the current level of MC study, a tight cut on tag side reduced χ^2 is required for the success of parameterization. Consequently, the total number of useful events for fitting decreases by 22.5% while signal-to-background ratio remains unchanged. It means that comparing with measurement with all potentially useable events, the statistical uncertainty

with the current selecting criteria will be approximately 13.6% larger. This matter becomes more important as the experiment proceeds to the point where statistical uncertainty becomes comparable to world average value.

7.3 Future developments

Signal side resolution

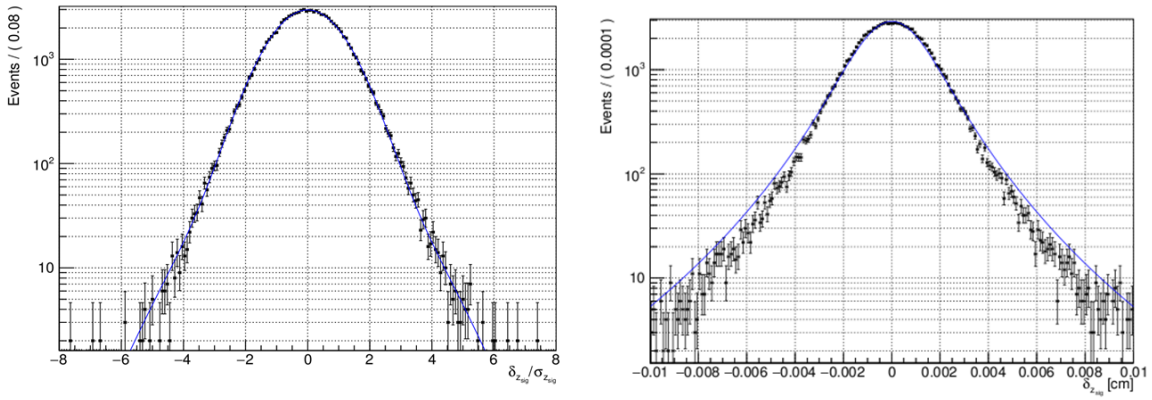
From the previous contents, signal side resolution is described by the sum of two Gaussian. The fit function used for residual distribution $P\left(\delta_{z_{sig}}\right)$ and pull distribution $P\left(\delta_{z_{sig}}/\sigma_{z_{sig}}\right)$ are

$$P\left(\delta_{z_{sig}}\right) = G\left(\delta_{z_{sig}}; s_{sig}\cdot\sigma_{z_{sig}}\right), \tag{7.1}$$

$$P\left(\delta_{z_{sig}}/\sigma_{z_{sig}}\right) = G\left(\delta_{z_{sig}}/\sigma_{z_{sig}}; s_{sig}\right), \tag{7.2}$$

where $s_{sig} \overset{\text{def}}{=} s^0_{sig} + s^1_{sig}\cdot\left(\chi^2/ndf\right)_{sig}$.

The pull fit was used to find the linear relation between $\left(\chi^2/ndf\right)_{sig}$ and s_{sig}, while the residual fit was used for the R^{det}_{sig} determination, which is to be convoluted with other components of the resolution function later. The two fits are highly related in a way that if the signal side residual distributes perfectly by the (7.1), the pull is automatically distributed by the (7.2). However as is shown in Fig. 7.8, the pull fit result agrees well with MC distribution but residual fit shows slight disagreement.



(a) (b)

Fig. 7.8 Comparison of signal side (a) $\delta_{z_{sig}}/\sigma_{z_{sig}}$ pull distribution and (b) $\delta_{z_{sig}}$ residual distribution, with fit functions superimposed.

The likely explanation is, for each event, in the residual fit different value sets of $\sigma_{z_{sig}}$ and $\left(\chi^2/ndf\right)_{sig}$ are used, while in the pull fit such combinational information is hidden and

only $(\chi^2/ndf)_{sig}$ is used. The $\sigma_{z_{sig}}$ information is combined with $\delta_{z_{sig}}$ to the pull distribution. Thus, the successful fit of pull distribution does not necessarily mean the function describes the signal side resolution well. It is also worth mentioning that there is no such unusual behavior in the tag side resolution function or in Belle.

In the future improvement, relations between vertex measurement uncertainty $\sigma_{z_{sig}}$ and quality indicator χ^2/ndf for the resolution function model should be better understood, especially for the signal side. For example, one can generate dedicated MC samples that distribute exactly according to the fit function and make comparisons with the more realistic simulated distributions to further study this behavior.

Tag side resolution

One of the major items for improvement in the tag side resolution model is the tight cut on $(\chi^2/ndf)_{tag}$. The situation becomes better when software analysis tools are upgraded in the future for better reconstruction efficiency to account for the event rejections. It could also be worth trying to include the effects of poorly reconstructed tracks and parameterize it into the resolution function model.

The other practical challenge for the tag side resolution is the correlation between R^{det}_{tag} and R^{np}_{tag}. Fig. 7.9 shows the R^{det}_{tag} in different slices of R^{np}_{tag}. The pragmatic solution now is to determine R^{np}_{tag} by fixing the shape of R^{det}_{tag} from the assumed fit function. In the future, the improvement could be towards the investigation of origin of such correlation. With such study, development of analysis modules with better separation power of effects in the tag side resolution is possible, thus reducing the systematic uncertainties related with R^{np}_{tag}.

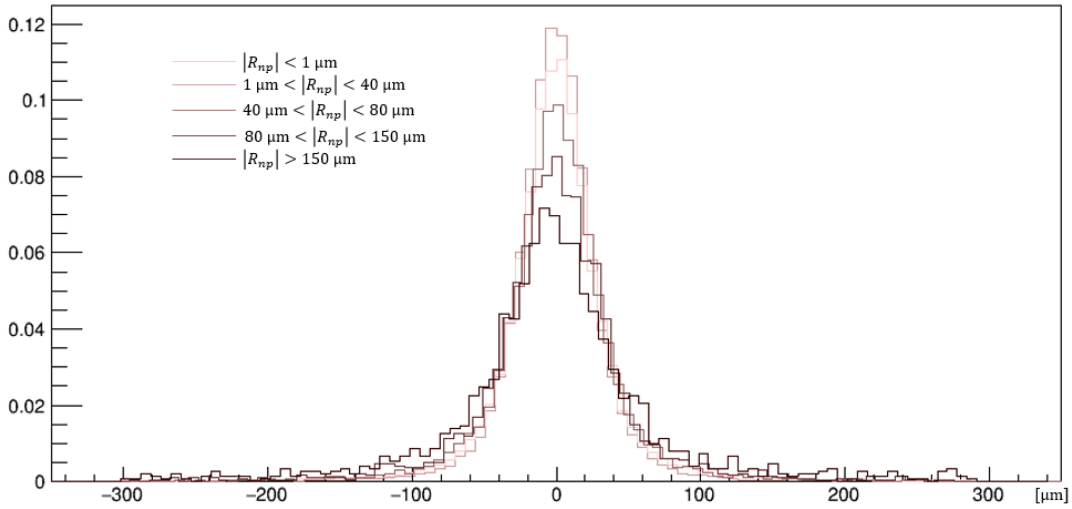
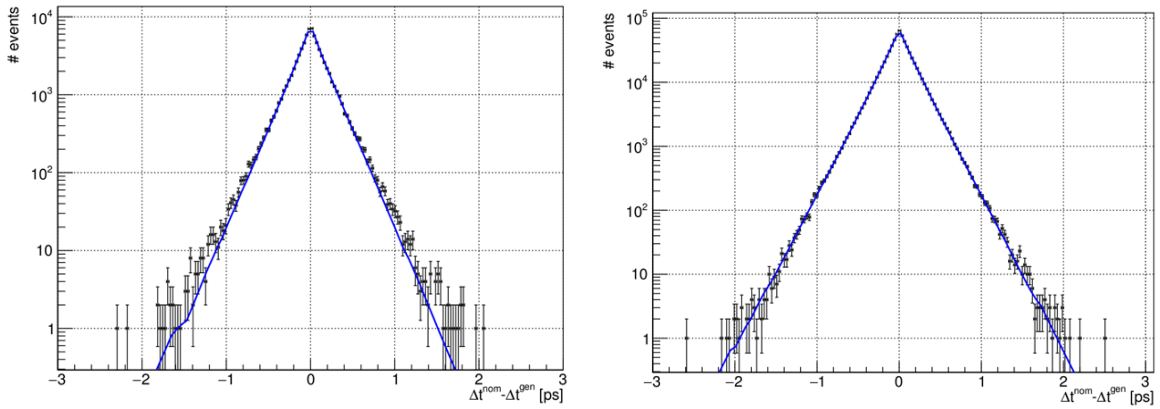


Fig. 7.9 R^{det}_{tag} distributions in different subsets of R^{np}_{tag}.

Kinematic smearing

The kinematic smearing effect is purely analytical, meaning that if the four-momentum is correctly measured, the kinematic smearing fit should be exact and agree with data sample used. However, it is observed that fit to events where B decays perpendicular to beam direction ($60° < \theta^{cms}_{B_{sig}} < 120°$) shows excess tail as in Fig. 5.13.

To investigate this problem, instead of fitting to successfully reconstructed events after selection (120k events, Fig. 7.10a), kinematic smearing fit is performed on all generated events (500k events, Fig. 7.10b). We observe that excess tail only appears in the fit to events after track and vertex reconstruction. So, the likely explanation is, as the fit result is plotted by the average function value summing $\theta^{cms}_{B_{sig}}$ in all events, the nonuniform reconstruction efficiency in $\theta^{cms}_{B_{sig}}$ could cause the small deviation at the tail between exact distribution and fit result. The amount of systematic uncertainty introduced from this behavior is remained to be further studied.



(a) (b)

Fig. 7.10 Distributions of R_k to (a) successfully reconstructed events and (b) all generated events. The exact fit curve is superimposed to each case.

7.4 Systematic uncertainties

The main sources of systematic uncertainties in the measurements of proper B^0 lifetime are listed as below:

Systematic uncertainties associated with event reconstruction

- Δt dependence of reconstruction efficiency

The systematic uncertainty from Δt dependence of reconstruction efficiency can be evaluated by performing fits to MC generated Δt^{gen} distribution in all generated events and reconstructed events separately with pure exponential function. The difference in the fitted lifetime is considered as bias due to reconstruction efficiency.

- Particle identification

  A loose particle identification cut $PID > 0.1$ for $K^{\pm}$ and $\pi^{\pm}$ was applied. The effects of tighter cuts on fitted lifetime should be considered.

- Vertex fitting

  In this study, the converging conditions for vertex fitting can be shifted. By applying the same lifetime extraction procedures, the systematic uncertainty from vertex fitting is obtained.

- IP constraint

  Different constraints on the IP for vertex fitting on both signal side and tag side can be tested to study the systematic uncertainty due to IP constraint.

- Poor quality vertex rejection

  The systematic uncertainty due to the selection requirement cuts on χ^2/ndf can be estimated by varying the cutting criteria.

- Signal region

  The signal region is determined from intervals of M_{bc} and ΔE. The effects can be investigated by comparing with fits to regions with broadened or shortened interval.

- Alignment and Calibration effects

  The effects of misalignment and miscalibration can be estimated by repeating the analysis to subsets of experimental data which was processed with different alignment and calibration constants.

Systematic uncertainties associated with fit model

- Fit bias

  The possible fit bias is studied from large number of simulated MC sample, where the deviation from input value is assigned as systematic uncertainty from bias in fitting.

- R_{det} parameterization

  The R_{det} parameters are fitted simultaneously with proper lifetime. We study the effect of R_{det} parameterization is by varying the model, for example, from sum of two Gaussians to single Gaussian for detector resolution.

- Non-primary smearing effects

  The effects from non-primary smearing effects on tag side can be estimated by shifting the R_{np} parameters by 1σ.

- Signal fraction

  Lifetime fits are repeated with the event-by-event signal fraction shifted by 1σ.

- Background Δt shape

  The parameters of background Δt shape are determined from fits to background region in experimental data. The systematic uncertainty due to background parameterization can be studied by shifting the parameters by 1σ.

7.5 Conclusion

In conclusion, we have constructed an event-by-event proper-time resolution function in the early phase of Belle II experiment that is comparable with Belle experiment. Compared with the existing analytical proper-time resolution functions developed in Belle II, this study has a more complicated structure and provides better physics implications to the sources of resolution components. The expected performance in time-dependent measurements is validated with demonstrations of lifetime fits to reconstructed $B^0 \to [D^- \to K^+\pi^-\pi^-]\,\pi^+$ control channel events from simulated MC and experimental data:

To the large number of signal only MC sample which corresponds to 1.6 ab^{-1} integrated luminosity, the fit result is

$$\tau_{B^0} = 1.5274 \pm 0.0059 \text{ ps},$$

and is consistent with input value 1.525 ps up to 0.4σ.

To the more realistic generic MC sample which corresponds to 400 fb^{-1} integrated luminosity, background components are taken into account and the fit result is

$$\tau_{B^0} = 1.5339 \pm 0.0141 \text{ ps},$$

which is consistent with input value 1.525 ps up to 0.6σ.

To the Belle II experimental data collected from physics run in 2019, which corresponds to 8.7 fb^{-1} integrated luminosity, the obtained lifetime is

$$\tau_{B^0} = 1.454 \pm 0.097 \text{ ps},$$

which is consistent with the current world average measurement 1.519 ± 0.004 ps.

Here the uncertainties correspond to statistical uncertainties, and the main sources of systematic uncertainties have been listed in Section 7.4. The evaluation of exact values for systematic uncertainties is still in development.

While consistent results have been shown in the extraction of B^0 lifetime using the resolution function model we constructed, there is still room for future improvements. With the Belle II experimental data accumulating in time, the final goal of this study is to provide a reliable analysis tool and aim for better precision in the measurements of time evolutions of B mesons. The study presented in this thesis is included as contribution to the measurement of CP violation parameter $sin2\phi_1$ in $B^0 \to J/\psi K_S^0$ decay for the ICHEP 2020 conference performed by the TDCPV working group.

References

[1] Planck Collaboration, N. Aghanim, Y. Akrami, et al. Planck 2018 results. VI. Cosmological parameters, 2018.

[2] M. Kobayashi and T. Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973. doi: 10.1143/PTP.49.652.

[3] Belle II Collaboration. Measurement of the B^0 lifetime using fully reconstructed hadronic decays in the 2019 Belle II dataset. *arXiv: High Energy Physics - Experiment*, 2020.

[4] C. Vladimir. The MPI Concept of Time-Dependent Fits at Belle II. Mar 2019. Talk given at the xFitter Workshop, Minsk, 18-20 March 2019, https://indico.desy.de/indico/event/22011/overview.

[5] H. Tajima, H. Aihara, et al. Proper-time resolution function for measurement of time evolution of B mesons at the KEK B-Factory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 533(3):370–386, Nov 2004. ISSN 0168-9002. doi: 10.1016/j.nima.2004.07.199. URL http://dx.doi.org/10.1016/j.nima.2004.07.199.

[6] L. Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945–1947, Nov 1983. doi: 10.1103/PhysRevLett.51.1945. URL https://link.aps.org/doi/10.1103/PhysRevLett.51.1945.

[7] Particle Data Group, M. Tanabashi, K. Hagiwara, et al. Review of Particle Physics. *Phys. Rev. D*, 98:030001, Aug 2018. doi: 10.1103/PhysRevD.98.030001. URL https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[8] A. Dighe. B Physics and CP Violation: an Introduction. In *Surveys in Theoretical High Energy Physics 1*, pages 1–21. Springer, 2013.

[9] R. Markus. *Time-Dependent CP Violation Measurements in Neutral B Meson to Double-Charm Decays at the Japanese Belle Experiment*. PhD thesis, Karlsruhe U., 2012.

[10] A. J. Bevan, B. Golob, et al. The Physics of the B Factories. *The European Physical Journal C*, 74(11), Nov 2014. ISSN 1434-6052. doi: 10.1140/epjc/s10052-014-3026-9. URL http://dx.doi.org/10.1140/epjc/s10052-014-3026-9.

[11] Y. Suetsugu, H. Fukuma, et al. Mitigating the electron cloud effect in the SuperKEKB positron ring. *Physical Review Accelerators and Beams*, 22, 02 2019. doi: 10.1103/PhysRevAccelBeams.22.023201.

[12] Belle II Collaboration, E. Kou, P. Urquijo, et al. The Belle II Physics Book. *Progress of Theoretical and Experimental Physics*, 2019(12), Dec 2019. ISSN 2050-3911. doi: 10.1093/ptep/ptz106. URL http://dx.doi.org/10.1093/ptep/ptz106.

[13] SuperB Collaboration. Superb: A high-luminosity asymmetric e+ e- super flavor factory. conceptual design report, 2007.

[14] M. Venturini and W. Kozanecki. The hourglass effect and the measurement of the transverse size of colliding beams by luminosity scans. In *PACS2001. Proceedings of the 2001 Particle Accelerator Conference (Cat. No.01CH37268)*, volume 5, pages 3573–3575 vol.5, 2001.

[15] Belle II Collaboration, T. Abe, I. Adachi, et al. Belle II Technical Design Report. *arXiv: High Energy Physics - Experiment*, 11 2010.

[16] A. Soffer. Belle II: Status and physics prospects. In *Intensity Frontier in Particle Physics: Flavor, CP Violation and Dark Physics*, Oct 2019.

[17] Y. Horii. TOP detector for particle identification at the Belle II experiment. In *The European Physical Society Conference on High Energy Physics*, volume 180, page 500. SISSA Medialab, 2014.

[18] T. Kuhr, C. Pulvermacher, et al. The Belle II Core Software. *Computing and Software for Big Science*, 3(1), Nov 2018. ISSN 2510-2044. doi: 10.1007/s41781-018-0017-9. URL http://dx.doi.org/10.1007/s41781-018-0017-9.

[19] H. Tanigawa. A study of beam background from SuperKEKB on Belle II Silicon Vertex Detector. Master's thesis, University of Tokyo, 2019.

[20] T. Hauth. Belle II Track Reconstruction and Results from first Collisions. *EPJ Web of Conferences*, 214:02032, 01 2019. doi: 10.1051/epjconf/201921402032.

[21] S. Tobias. Vertexing and Tracking Software at Belle II, 2014.

[22] F. Tenchini and J. Krohn. Decay Chain Reconstruction at Belle II. *EPJ Web of Conferences*, 214:06023, 01 2019. doi: 10.1051/epjconf/201921406023.

[23] W. Waltenberger. RAVE—A Detector-Independent Toolkit to Reconstruct Vertices. *Nuclear Science, IEEE Transactions on*, 58:434 – 444, 05 2011. doi: 10.1109/TNS.2011.2119492.

[24] J. Tanaka. Development of the Low Momentum Track Reconstruction Program and the Kinematic Fitter for the Belle Experiment. Master's thesis, University of Tokyo, 1999.

[25] S. Dey. Beam-constrained vertexing for B physics at the Belle II Experiment. In *Workshop on Frontiers in High Energy Physics*, 2019.

[26] W. Verkerke and D. Kirkby. The RooFit toolkit for data modeling, 2003.

[27] Particle Data Group, P.A. Zyla, R.M. Barnett, et al. The Review of Particle Physics. *to be published in Progress of Theoretical and Experimental Physics*, 2020.