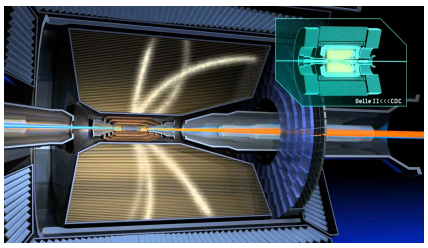
DEPARTMENT OF PHYSICS
INDIAN INSTITUTE OF TECHNOLOGY MADRAS
CHENNAI – 600036

Study of $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ at the Belle II experiment



A thesis

Submitted by

SHUBHAJIT SANA

For the award of the degree

Of

Dual Degree in Physics

May, 2023

THESIS CERTIFICATE

This is to certify that the thesis titled **STUDY OF $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ AT THE BELLE II EXPERIMENT**, submitted by **SHUBHAJIT SANA**, to the Indian Institute of Technology Madras, for the award of **Dual Degree in Physics**, is a bona fide record of the research work done by him under my supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Place: Chennai

Date: 13^{th} May 2023

Prof. Jim Libby
Research Guide
Professor
Department of Physics
IIT Madras, 600036

ACKNOWLEDGEMENTS

I would like to take this moment to express my deep appreciation to the individuals whose unwavering encouragement and support have made this thesis possible. I am particularly grateful to my project supervisor, Prof. Jim Libby, whose dedicated guidance and support throughout the course of this project have been invaluable. Without his advice and constructive criticisms, this thesis would not have come to fruition. It has been a privilege to work with him, and I am indebted to him for his encouragement and patient mentorship over the last one year.

In addition, I would like to extend my gratitude to all the teachers in the Department of Physics at IIT Madras, who have provided me with a solid foundation and understanding of many subjects in Physics, enabling me to carry out this project.

I would also like to extend my gratitude to my seniors, Neeraj Kumar, Ansu Johnson, and Gaurav Sharma, for their valuable assistance in providing me with a fundamental understanding of the analysis framework. Finally, I would like to express my heartfelt appreciation to my family members for their unwavering love and support throughout my academic journey.

ABSTRACT

This thesis presents a study of rare decay channel $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ using 200 fb^{-1} Monte Carlo simulated data generated by the Belle-II software framework. Topological analysis is being performed to identify the major sources of peaking background and appropriate veto is being applied to suppress them. A Machine Learning algorithm has been utilized to distinguish between the signal and the dominant continuum background based on their distinctive topological characteristics. The measured value of the branching ratio for the decay channel is $\mathbf{(11.1 \pm 0.1) \times 10^{-3}}$ which is consistent with the figure specified in the decay file used for the production of generic Monte Carlo (MC). However, this result encompasses the contribution from strong intermediate resonance states, such as a_1^- and ρ^0.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

This chapter serves as an introduction to the theoretical basis of the research presented in this thesis. Section 1.1 provides a concise overview of the Standard Model. Section 1.2 presents a brief historical account of the field of flavor physics. Section 1.3 discusses the motivation for studying the decay channel $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$.

1.1 The Standard Model of particle physics

The Standard Model (SM) [1] is a comprehensive framework for understanding three of the four fundamental interactions observed in nature (excluding gravity). It is an effective quantum field theory that unites quantum mechanics with special relativity, providing a foundational understanding of the physical world. In this framework, a field is a collection of values assigned to each point in space and time that represents certain physical properties. Quantum fields are pervasive throughout spacetime and are governed by the laws of quantum mechanics. Perturbing a quantum field results in oscillatory states, known as field excitations or particles, which carry more energy than the resting state. For example, the electron is a massive excitation of the corresponding electron field. Quantum fields interact with each other and the Standard Model is the theory that explains their dynamics at energy levels relevant to the subatomic world. The particles and their interactions are described using a Lagrangian formalism, which includes all combinations of fields and interaction operators that are not prohibited by the symmetries of the system. The key concept of the Lagrangian is local gauge symmetry, which means that the Lagrangian remains unchanged when space-time-dependent transformations are applied to the phases of the fields. Interaction terms appear in the free-field Lagrangian after requiring it to be invariant under local gauge symmetries.

The symmetry group $SU_U(3) \otimes SU_L(2) \otimes U_Y(1)$ forms the foundation of the Standard Model. $SU_U(3)$ is a standard unitary group that describes the strong interactions, also known as quantum chromodynamics (QCD), where C stands for the color charge. $SU_L(2) \otimes U_Y(1)$ is a product of groups that describes the combination of weak and electromagnetic interactions, with L representing left and Y representing hypercharge. The interactions are mediated by spin-1 particles called gauge bosons. The strong interactions are mediated by eight massless particles corresponding to the $SU_U(3)$ generators, referred to as gluons, which carry a charge that can be of three types, called color. Weak interactions are mediated by two charged massive bosons, $W\pm$, and a neutral massive boson, Z^0. Electromagnetic interactions occur between particles carrying electric charge and are mediated by a neutral massless boson called the photon γ. The physical electroweak bosons ($W\pm, Z^0, \gamma$) are formed from linear combinations of $SU_L(2) \otimes U_Y(1)$ generators.

Fermions are matter particles that correspond to excitations of spin-1/2 fields. They are associated with free parameters that determine their masses. Each fermion is also linked to an anti-particle that has the same mass but opposite internal quantum numbers. Fermions can be divided into two classes: quarks, which are the fundamental building blocks of nuclear matter, and leptons, which are organized into three weak-isospin doublets.

- Each quark doublet is composed of an up-type quark, which carries a charge of 3/2 e, and a down-type quark, which carries a charge of -1/3 e. Quarks interact with both the strong and electroweak interactions and possess a color and a "flavor" quantum number, which come in six different varieties. The flavor quantum number is conserved in electromagnetic and strong interactions, but not in weak interactions. Free quarks are not directly observable due to color confinement, and are only observed in their colorless bound states, such as mesons (composed of a quark and an anti-quark) and baryons (composed of three quarks). Baryons

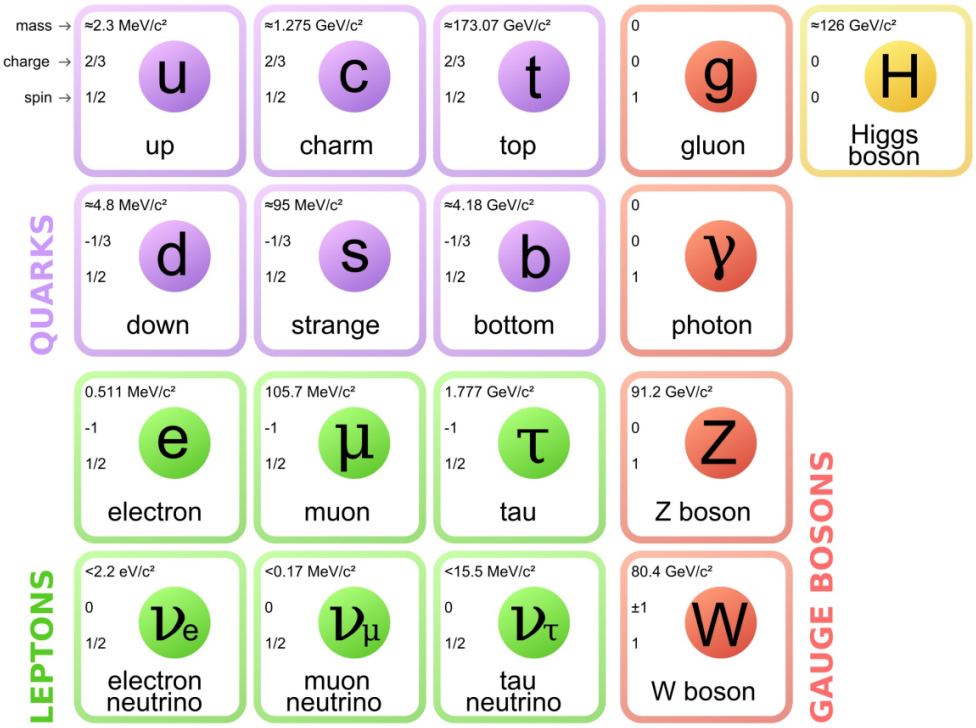


Figure 1.1: The Standard Model of fundamental particles. [2]

possess a quantum number known as baryon number, which is conserved even if no symmetry of the Lagrangian requires it.

- A lepton doublet consists of a massless, neutral neutrino and a massive particle with electric charge -e. These particles interact only through the electroweak force. Each lepton is associated with a lepton family quantum number. The sum of these quantum numbers, known as the global lepton number, is conserved in all interactions, despite the absence of any symmetry that requires it. However, individual lepton numbers may not be conserved in certain processes, such as neutrino oscillations.

Fig. 1.1 shows the elementary particles in the Standard Model. In addition to gauge symmetry, discrete symmetries, such as parity (P), charge conjugation (C), and time reversal (T), also play a crucial role in constraining the dynamics. The product of these three symmetry transformations is found to be conserved in all interactions as dictated

by fundamental principles of field theory, but they are not conserved individually [3]. Parity symmetry is violated most significantly in weak interactions, while the combined CP symmetry is violated at the 0.1% level in weak interactions. The strong interaction has the potential to violate CP symmetry, but no experimental evidence of such behavior has been found yet. To explain this phenomenon, axions have been hypothesized as yet-to-be-discovered particles.

1.2 History of Flavour Physics

Flavour physics has played a crucial role in the establishment and advancement of the Standard Model over the course of several decades. The study of flavour began in the 1950s and 1960s, when Feynman and Gell-Mann postulated the universality of weak interactions, which correctly predicted the close similarity between the effective Fermi coupling constants in the muon decay $\mu \rightarrow e^- \bar{\nu}_e \nu_\mu$ and in the neutron β-decay $n \rightarrow p e^- \bar{\nu}_e$ [4]. However, their theory failed to account for the observed difference in the effective Fermi constant in the decays of strange particles, such as the $\Lambda \rightarrow p e^- \bar{\nu}_e$ decay, which was observed to be smaller by a factor of 4-5. During the 1960s, Cabibbo proposed the Cabibbo angle, denoted by θ_C, which helped maintain the universality of the weak couplings. The angle accounted for $\cos\theta_C$ in $d \rightarrow u$ transitions and $\sin\theta_C$ in $s \rightarrow u$ transitions, with a value of $\theta_C \simeq 13^\circ$ that was consistent with experimental data and the same Fermi constant [5].

Despite its success, Cabibbo theory could not explain the absence of strangeness-changing neutral-current processes, including the $K^0 \rightarrow \mu^+\mu^-$ decay, which was expected to have a measurable rate but was not observed at that time. To solve this, the Glashow-Iliopoulos-Maiani (GIM) mechanism was introduced in 1970 [6]. The mechanism elegantly solved the puzzle by proposing the existence of the fourth quark, the charm quark c, which had already been theorised in 1964 by Bjorken and Glashow [7]. They predicted the mass of the charm quark to be around 1.5 GeV/c^2. The experimental discovery of the J/ψ meson

in 1974, a bound state of a charm and an anti-charm quark, confirmed the predictions of the GIM mechanism [8]. Furthermore, the discovery of the charm quark and the GIM mechanism allowed for the interpretation of the Cabibbo angle, θ_C, as a rotation angle in the flavour space between the weak eigenstates d', s' and the mass eigenstates d, s as follows

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

Even with the addition of the charm quark, the theory still lacked a coherent explanation for the observed violation of CP symmetry in the decays of neutral kaons, which was first noted by Cronin and Fitch in 1964 [9]. This explanation came about in 1973 when Japanese physicists Kobayashi and Maskawa proposed the existence of a third generation of quarks and extended the Cabibbo matrix to a 3×3 matrix that included a complex phase, δ, responsible for CP violation [10]. The discovery of the bottom quark, b, in 1977 by the E288 experiment at Fermilab led by Lederman provided the first experimental evidence for their theory [11]. The top quark, t, was not discovered until 1994 through measurements made by the D0 and CDF collaborations at Fermilab [12, 13]. However, there were indirect indications of the existence of a heavy top quark earlier on through the observation of $B^0\bar{B}^0$ mixing [14]. This phenomenon was first noticed in 1987 by the ARGUS experiment conducted at the DORIS II electron-positron collider located in DESY, Germany. Despite important contributions made by experiments like CLEO at Cornell (USA) in the 1990s towards the study of B physics, CP violation in B decays was not observed.

The field of particle physics was revolutionized with the emergence of B-factories, which gave rise to two experiments - BaBar (1999-2010) at the PEP-II accelerator located in SLAC, USA, and Belle (1999-2010) at the KEKB collider situated in KEK, Japan. The first observation of CP violation in B decays was made by the BaBar and Belle experiments in 2001, demonstrating that this phenomenon was not unique to kaon mixing

but was a genuine feature of weak interactions [15, 16]. In addition to this groundbreaking discovery, these experiments also conducted many other measurements related to the decays of B mesons, which helped to better understand flavour physics and to constrain the Kobayashi-Maskawa theory of CP violation. To achieve even higher luminosity, the concept of a super B-factory was introduced after the shutdown of Belle and BaBar, which led to the upgrade of KEKB to SuperKEKB and the establishment of the Belle II experiment, which began collecting data in March 2018. Hadronic collisions have also been successful in studying flavour physics since the early 2000s, with experiments like D0 and CDF at the Tevatron (Fermilab, USA), as well as LHCb, CMS, and ATLAS at the LHC (CERN, Switzerland).

Flavour physics has witnessed an exciting period in the recent years. Through precise comparisons of experimental measurements with predictions of the Standard Model (SM), a consistent pattern of flavour anomalies has emerged. These tensions with the SM have been observed in $b \to sll$ [17] and $b \to c\tau\nu$ [18] transitions, the muon anomalous magnetic moment [19], and other cases. The significance of such findings is that flavour observables can be sensitive to energy scales beyond the reach of colliders and suggest the potential contribution of new phenomena. Against this backdrop, Belle II's ability to measure $b \to s\nu\bar{\nu}$ transitions and other rare decays with unprecedented accuracy offers a unique opportunity to unravel the mystery of flavour anomalies and provide insights into the search for new physics.

1.3 Motivation to study $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ decay channel

The study of $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ decay channel is of great interest due to several reasons. In the study of particle physics, determining the CKM matrix elements $|V_{ub}|$ and $|V_{cb}|$ is of great importance, and inclusive and exclusive $b \to ul\nu$ and $b \to cl\nu$ transitions are crucial in achieving this goal. Full Event Interpretation (FEI) is a powerful technique used to reconstruct these decays with missing energy. Hadronic tag decays, such as

the $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ channel, which have relatively higher branching fractions, are particularly useful for this purpose. However, simulating these decays in Monte Carlo is challenging due to the large uncertainty in the measured branching fraction, which is $(5.6 \pm 2.1) \times 10^{-3}$, and was determined using a very small amount of data by the CLEO collaboration in 1992 [20]. Fortunately, we have access to 711 fb^{-1} data from Belle and 427.79 fb^{-1} data from Belle II, which can be used to decrease the statistical uncertainty in the significance label.

The decay channels $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ and $B^+ \to \bar{D}^0K^+\pi^-\pi^+$ were utilized to measure the weak phase γ with LHCb's limited $3fb^{-1}$ data, resulting in a measurement of $\left(74^{+20}_{-19}\right)^\circ$ [21]. However, with the vast amount of data available from Belle and Belle II, we can measure γ with less statistical uncertainty. Moreover, the decay $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ is a fascinating channel for studying the dynamics of four-body decays, which are complicated and not yet fully understood. Exploring these decays can offer valuable insights into the non-perturbative QCD effects that govern the hadronization process, and can help refine theoretical models of multi-body decays.

CHAPTER 2

The Belle II Experiment

This chapter provides a concise overview of the Belle II experiment. It begins with a brief discussion of B-factories and the SuperKEKB accelerator in Sections 2.1 and 2.2, respectively. Section 2.3 then describes the Belle II detector in detail, including separate paragraphs for each subdetector. The Monte Carlo simulation used in the experiment is discussed in Section 2.4, while Section 2.5 explains the object reconstruction process. Finally, Section 2.6 focuses on the Belle II analysis software framework, basf2.

2.1 B-factories

B-factories are particle accelerators designed to produce large numbers of B mesons and anti-B mesons for study. The first B-factory, the BaBar detector, was built at the Stanford Linear Accelerator Center (SLAC) in California in the late 1990s. It was designed to study the properties of B mesons produced in collisions between electrons and positrons. The BaBar detector operated from 1999 to 2008 and produced a wealth of data on B meson decays. In 1999, a similar B-factory called Belle was built in Japan at the KEK (High Energy Accelerator Research Organization) laboratory. Belle was designed to study B mesons produced in collisions between high-energy electrons and positrons. Belle has also produced a large amount of data on B meson decays. Both BaBar and Belle have made important contributions to our understanding of B mesons and their decays. These contributions include measurements of the B meson lifetime, the rate of B meson decays into different final states, and the CP violation parameters of B meson decays.

The BaBar detector has been decommissioned, but Belle has been upgraded to Belle II, which is designed to produce even larger numbers of B mesons for study. Belle II

began collecting data in 2018 and is expected to operate for at least a decade. Belle II is a significant upgrade over the original Belle detector. It includes a larger detector, more powerful magnets, and improved electronics. These upgrades allow Belle II to produce more B mesons and to detect a wider range of final states. Belle II has already produced important results [22].

2.2 SuperKEKB accelerator

SuperKEKB is a particle accelerator located in Tsukuba, Japan. It is a successor to the KEKB accelerator, which operated from 1999 to 2010. It is designed to produce large numbers of B mesons and anti-B mesons for study. A schematic diagram of it is structure is given in Fig 2.1. It is made up of two storage rings, the high-energy ring (HER) and low-energy ring (LER), which are approximately three kilometers in circumference and built side-by-side in the TRISTRAN tunnel. These two rings can store currents up to 2.6 A and 1.1 A, respectively, and use an RF of 508.9 MHz to accelerate the beams. The HER (LER) can accelerate electron (positron) beam upto 7 GeV (4 GeV) which result center-of-mass energy 10.58 GeV in the relativistic limit. They cross at the interaction point in the Tsukuba experimental hall at the center of the Belle II detector. It is designed to increase the luminosity to $6 \times 10^{35}\text{cm}^{-2}\text{s}^{-1}$, i.e. 60 $\text{pb}^{-1}\text{s}^{-1}$, which is 40 times higher than KEKB [23]. The increase in luminosity is achieved by decreasing the beam sizes at the interaction point, and maximizing the crossing angle as prescribed by P. Raimondi's nanobeam collision scheme [24]. The Lorentz boost due to the beam-energy asymmetry is $\beta\gamma \approx 0.28$, resulting in an average flight distance of about 130 μm for the B mesons. To decrease beam losses caused by Touschek scattering in the lower energy beam, the boost is somewhat decreased compared to KEKB. This change has the benefit of enhancing the acceptance of solid angle for decays with missing energy.

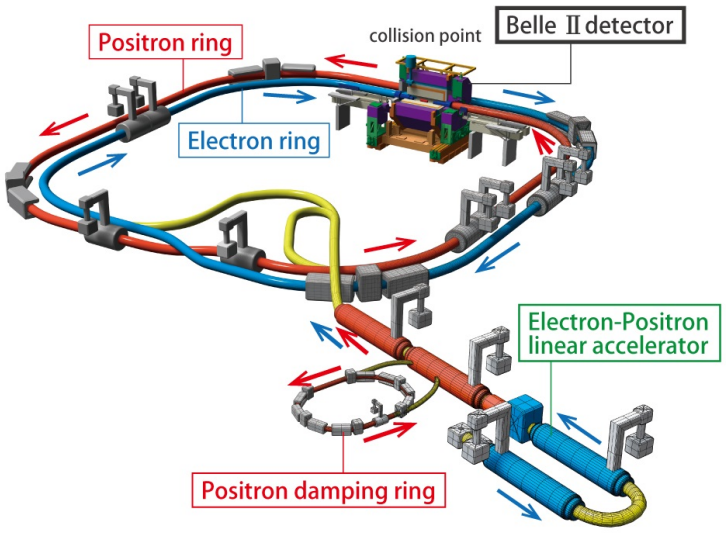


Figure 2.1: Schematic representation of SuperKEKB [25]

2.3 Belle II Detector

The Belle II detector [26] is composed of a cylindrical arrangement of subdetectors surrounding the beam pipe, similar to its predecessor Belle, as shown in Fig 2.2 The center of the detector contains three tracking subdetectors, namely the pixel and silicon vertex detector (PXD, SVD), and the central drift chamber (CDC). Recently added to the detector, the PXD consists of two layers positioned at 14 mm and 22 mm, and it provides improved resolution for the impact parameter of high-momentum tracks. The SVD has been upgraded from the previous Belle detector, with four layers of double-sided silicon strips, larger radial coverage, new electronics, and sensor design. One of the main roles of these two Vertex Detectors (VXD) is to reconstruct decay vertices and low-momentum tracks that do not reach the CDC. The CDC encloses the vertex detectors and determines charge, momentum, and energy loss by ionization of charged-particle tracks. It has additional layers, extends the detector to a larger radius, smaller drift cells, and new readout electronics compared to its predecessor. Two detectors, the time-of-propagation counter (TOP) and the aerogel ring-imaging Cherenkov counter (ARICH), cover the barrel and forward end cap regions of the detector, respectively, and are used for charged-particle identification (PID). The electromagnetic calorimeter (ECL)

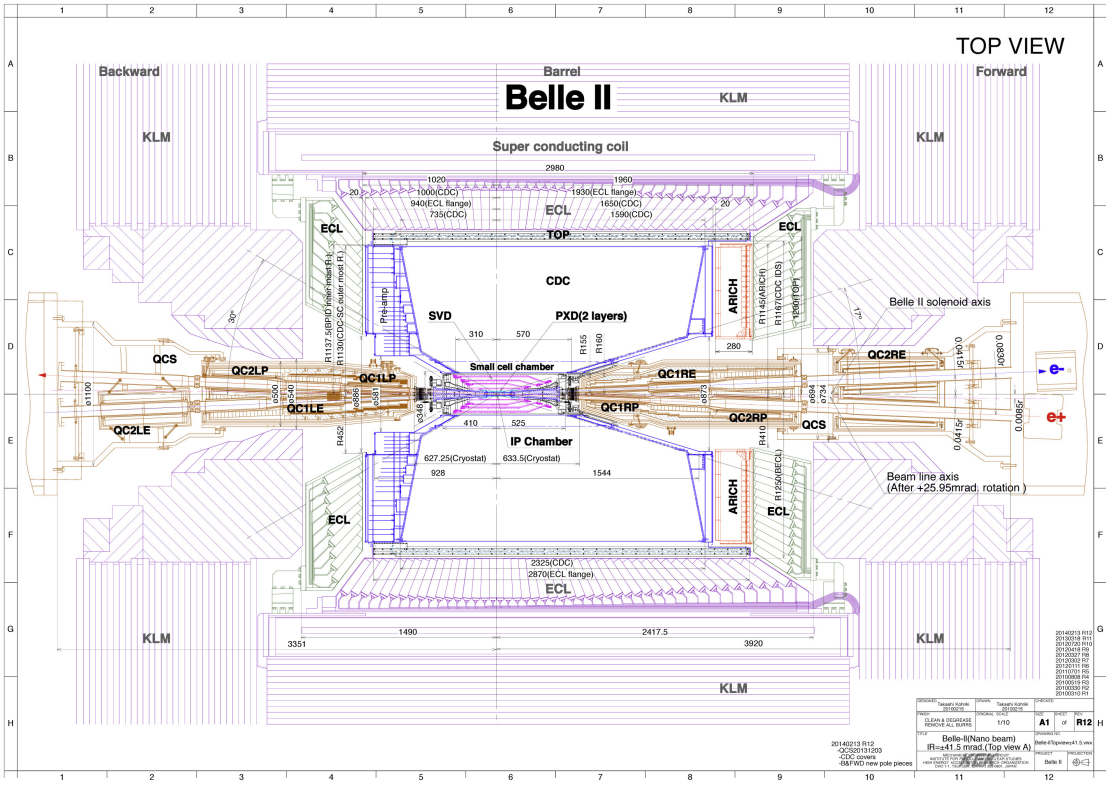


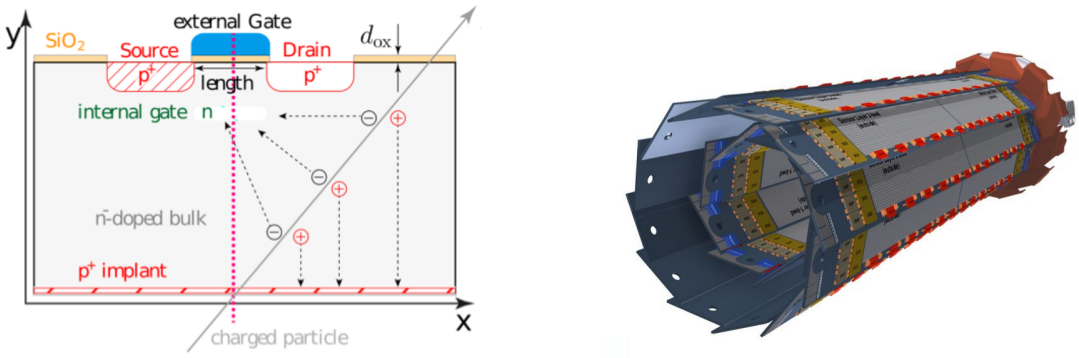
Figure 2.2: Longitudinal cross section of the Belle II detector. [27]

is located inside a superconducting solenoid that generates a magnetic field of 1.5 T and measures the energy of electromagnetically-interacting particles, including photons and electrons. The detector also includes an upgraded KLM detector in its outermost part, which is used to identify K_L^0 and μ. To cope with the challenging background conditions anticipated at SuperKEKB, Belle II is equipped with a high-efficiency trigger system. The following subsections outline the main modifications made to the subsystems compared to the Belle detector.

2.3.1 Vertex Detector

The VXD is positioned at the innermost part of the Belle II subsystem, encircling the beam pipe, and consists of six layers - two layers of PXD and four layers of SVD.

The PXD [28] has been developed utilizing the DEpleted P-channel Field Effect Transistor (DEPFET) technology, which offers low power consumption and intrinsic noise, as well



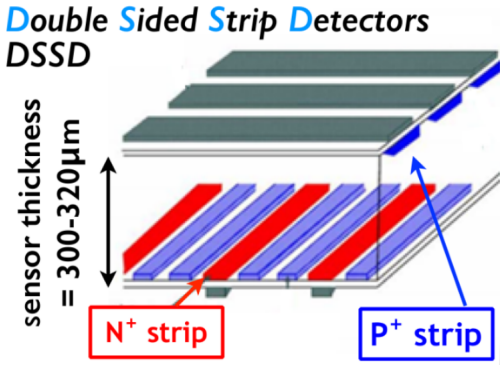
(a) Operating principle of DEPFET

(b) Schematic view of PXD

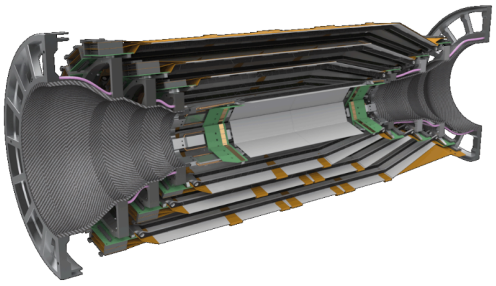
Figure 2.3: Schematic view of PXD and Operating principle of DEPFET [30]

as a small material budget. The setup and the operating principle of a DEPFET pixel are schematically illustrated in Fig. 2.3a. Each DEPFET pixel is a cell, which is a monolithic structure with internal amplification, making it significantly smaller than other devices that require external amplification. The two PXD layers are situated directly outside the beam pipe at radial distances of 14 mm (layer 1) and 22 mm (layer 2), and are composed of eight and 12 ladders, respectively. An illustration of this structure is provided in Fig. 2.3b. The layers are incredibly lightweight, with a thickness of less than 0.2% radiation length per layer. Each module's active area is segmented into 250 x 768 pixels, with sizes ranging from 50 x 55 μm^2 in the center of the inner layer to 50 x 85 μm^2 in the outer layer. The hit-efficiency of most of the modules, having readout time 20 μs, is greater than 98%, and they guarantee vertex reconstruction with an average spatial resolution of approximately 14 μm [29]. Due to its PXD position, it is exposed to significant QED background, and therefore the modules are designed to resist upto 20 Mrad radiation dose.

The SVD [31] consists of four layers of double-sided silicon strip detectors (DSSDs), depicted in Fig. 2.4a, with radii of 39 mm (layer 3), 80 mm (layer 4), 104 mm (layer 5), and 135 mm (layer 6), and has angular coverage of 17°< θ < 150°. An illustration of this structure is provided in Fig. 2.4b. The n and p side strips are arranged perpendicular and



(a) Setup of DSSD



(b) Schematic view of the geometrical arrangement of the SVD

Figure 2.4: Schematic view of the geometrical arrangement of the SVD and Setup of DSSD [33]

parallel to the beam direction, respectively, providing both x and y coordinates for the hits. When a charged particle traverses a sensor, electron-hole pairs are formed along its path by ionization. Electrons are collected by n-strips, and holes are by p+ strips, enabling two coordinates of the particle position to be read. In total, there are 72 SVD sensors and around 220 thousand strips. All the sensors are connected to fast APV25 front-end chips, which were originally designed for the CMS experiment, resulting in a good signal-to-noise ratio for all the layers. The optimized ladder design and support structures result in a material budget per ladder equivalent to only 0.7% radiation length at normal incidence. The average hit efficiency is greater than 99.5%, and the spatial resolution varies from 18 μm to 35 μm [32]. To cope up the high luminosity environment at Belle II, the SVD radius coverage has been increased to 135 mm, which allows the SVD to trace low-momentum particles that do not reach the central-drift-chamber (CDC), such as the pions from K^0_S decays or slow pions from D decays.

2.3.2 Central Drift Chamber

The Central Drift Chamber (CDC) [34] is a crucial instrument in the Belle II spectrometer, and it is composed of a large volume drift chamber with small drift cells. It has multiple functions, primarily serving as a tracking detector that reconstructs charged tracks and measures their momenta. It also identifies charged particles by measuring their energy

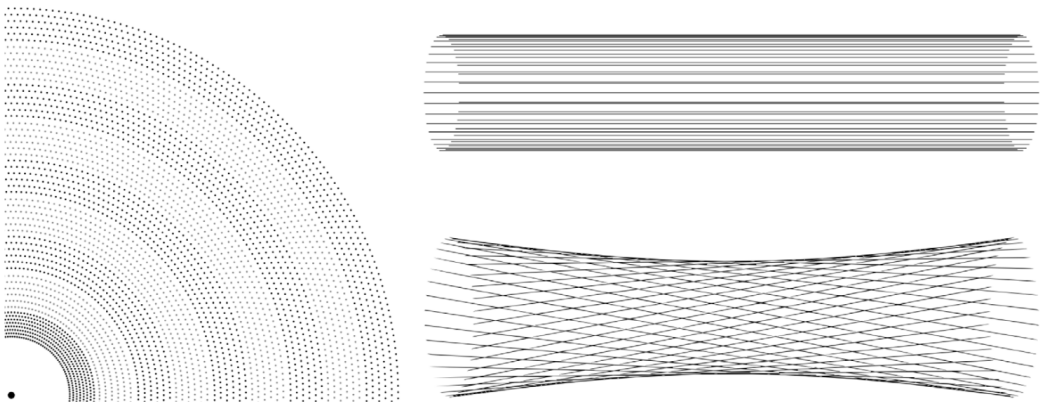


Figure 2.5: Left: A quadrant of a slice of the $r - \phi$ projection of the drift chamber. The innermost superlayer contains eight layers, all others contain six. Right: A visualization of stereo wires (bottom) relative to axial wires (top). The skew is exaggerated [36]

loss within the gas volume. The CDC generates reliable trigger signals for charged particles, and its volume is contained between an inner cylinder with a radius of 160 mm and an outer cylinder with a radius of 1130 mm. Within the chamber, there are about 50,000 sense and field wires within a gas mixture of 50% helium and 50% ethane that define drift cells of approximately 2 cm in size. The sense wires are arranged in layers, and six or eight adjacent layers correspond to a so-called superlayer. The CDC consists of eight outer superlayers, each composed of six layers of 160 to 384 wires, while the innermost superlayer consists of eight layers with 160 wires, defining smaller drift cells to handle the increase in background at smaller radial distances from the interaction point. The superlayers alternate between axial (A) and stereo (U, V) orientation to create an AUAVAUAVA configuration that can reconstruct a full three-dimensional helix track. An illustration of the CDC wire configuration is provided in Fig. 2.5. The average drift velocity is 3.3 cm/μs, and the maximum drift time is approximately 350 ns [35]. The polar angular coverage is between 17°and 150°. Lastly, it has an average spatial resolution of approximately 120 μm, with a relative accuracy of about 12% for dE/dx measurements for particles with an incidence angle of 90°[36].

2.3.3 Particle identification system

At Belle II, there are two subsystems dedicated to particle identification: the Time-of-Propagation (TOP) counter [37] in the barrel region and the Aerogel Ring-Imaging Cherenkov (ARICH) [38] counter in the end-cap. Although both sub-detectors use the Cherenkov effect to identify charged particles passing through the spectrometer, their operating principles differ.

The TOP counter is located on the CDC outer cover with angular acceptance range of $31° < \theta < 128°$. It is composed of 16 modules, each consisting of two fused quartz bars, a mirror at the forward end of the bars, and a 10 cm long prism that couples the bar with an array of micro-channel-plate photomultiplier (MCP-PMT) tubes [39]. Figure 2.6 illustrates a TOP module and it is operating principle. When a charged particle travels through the quartz radiator, it emits Cherenkov photons in a cone with an opening angle θ_C, which depends on the particle velocity. Part of the photons is internally reflected inside the quartz bar, and their time of propagation to the photomultiplier surface is related to θ_C. Photon arrival times are measured using the time of collision as a reference, and thus, they are further separated in time based on the time of flight of the corresponding charged particle. At a momentum of 2 GeV/c, the arrival-time differences between photons emitted by kaons and pions are approximately 100 ps, while the MCP PMT used in the detector has a time resolution of approximately 50 ps. The PID information is obtained by comparing the time distribution of the photons in one of the modules with the probability density functions (PDFs) describing the expected distributions for six particle hypotheses: e, μ, π, K, p, d. The performance of the TOP detector is evaluated in terms of K/π separation, and it exhibits a good 85% kaon identification efficiency at a 10% pion misidentification rate [40].

The ARICH detector is utilized in the forward endcap, with angular acceptance spans from 14°to 30°, to distinguish between kaons and pions across their momentum spectrum and to distinguish between pions, muons, and electrons below 1 GeV/c. It is a proximity-

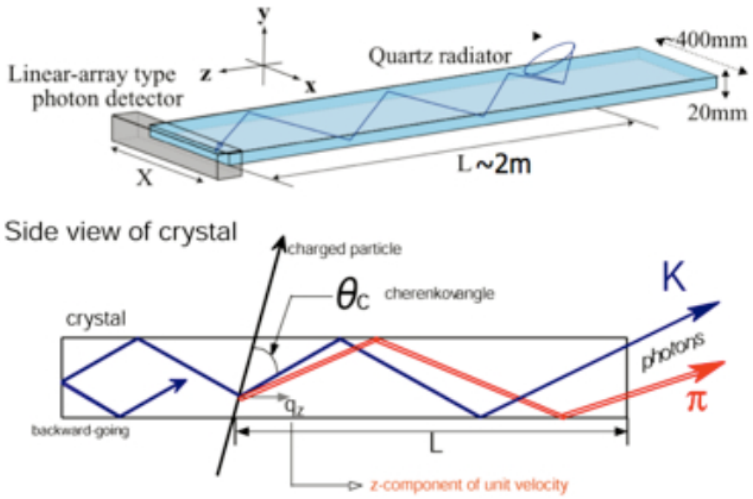
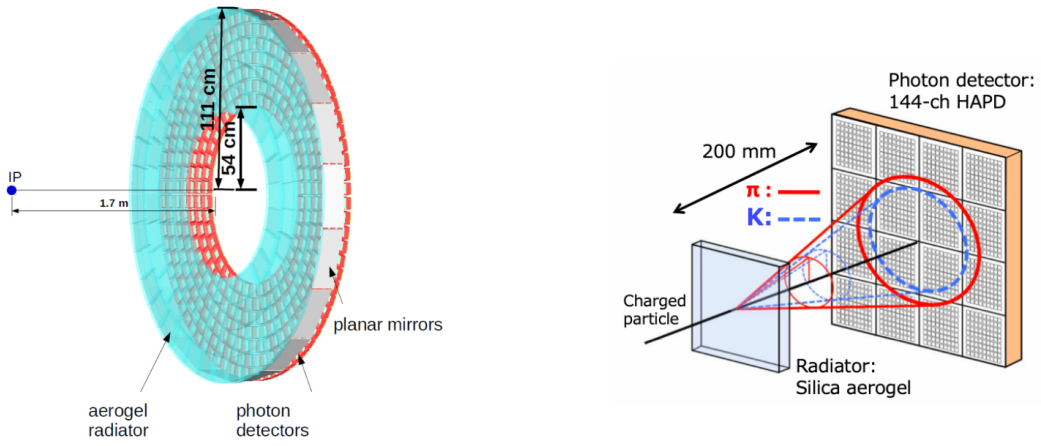


Figure 2.6: Schematic diagram of TOP module and operating principle of the top counter [41]

focusing ring-imaging Cherenkov (RICH) detector that employs aerogel as a radiator. An illustration of this subdetector is shown in Fig. 2.7a. It has a toroidal shape with an inner radius of 420 mm, an outer radius of 1145 mm, and a length of 280 mm. The components of the detector along the beamline, beginning with the nearest to the collision point, include a 40 mm thick radiator consisting of 248 silica aerogel tiles arranged in two layers, a 160 mm expansion space, and an 80 mm thick photon detection system consisting of 420 Hybrid Avalanche Photo Detectors (HAPD) [42]. The operating principle of the ARICH is shown in Fig. 2.7b. By measuring the emission angle of the Cherenkov photons, the type of charged particle traveling through the detector can be identified. The particle identification data is obtained by comparing the observed photon pattern with the PDFs that depict the expected distribution of Cherenkov photons on the photodetector plane for various particle hypotheses. The ARICH counter achieves in is capable of separating kaons from pions with an identification efficiency of 93% at 10% pion fake rate [43].



(a) (b)

Figure 2.7: a) Schematic diagram of ARICH detector and b) principle of π/K identification for the ARICH counter [44]

2.3.4 Electromagnetic Calorimeter

The ECL in Belle II is an upgraded version of the same calorimeter used in Belle, providing improved electronics for a variety of functions. Its primary use is for energy measurements of electromagnetic particles. Additionally, it efficiently detects photons, identifies and separates electrons from hadrons, detects K_L^0 mesons in conjunction with the KLM, generates signals for the trigger, and measures luminosity. Fig.2.8 elustrates the geometry of the electromagnetic calorimeter. It has a 3 m long barrel section with an inner radius of 1.25 meters, as well as forward and backward end caps located at z = 2.0 meters and z = -1.0 meters from the interaction point, respectively. It covers a polar angle region of 12.4°< θ < 155.1°, with two 1°gaps between the barrel section and the end caps. The barrel section contains 6,624 thallium-doped cesium iodide (CsI(Tl)) crystals, each being a truncated pyramid with an average cross-sectional size of 6 × 6 cm2 and a length of 30 cm (16.2 radiation lengths) [27]. The end caps consist of 2,112 cesium iodide crystals, each of which is wrapped with a layer of 200 μm thick Teflon and covered by 50 μm thick aluminized polyethylene. Light readout is performed by 10 × 200 mm2 photodiodes attached to the rear surface of the crystal. A preamplifier is connected to each photodiode, providing two independent output lines from each crystal. The two

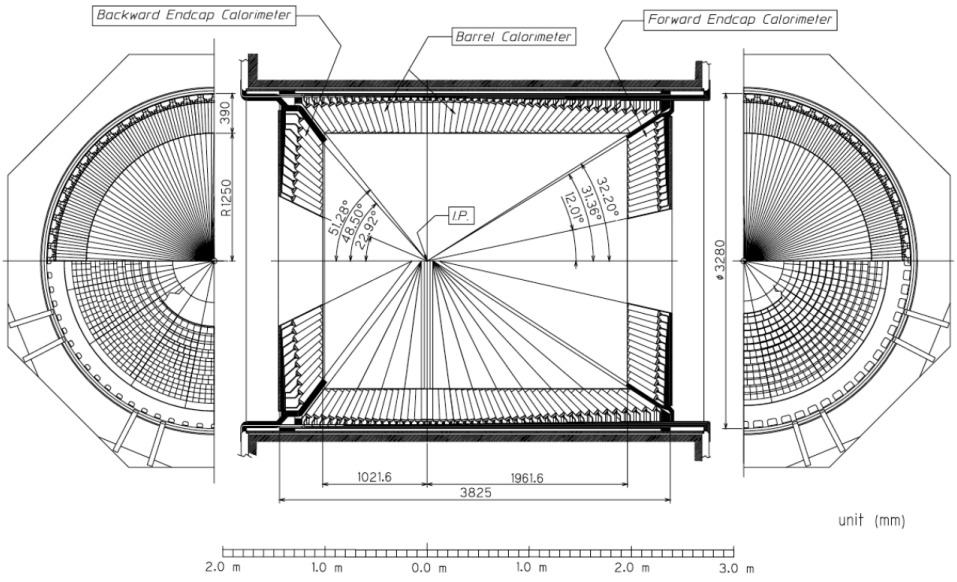


Figure 2.8: Schematic diagram of the ECL. [47]

pulses produced are then summed in the "Shaper-DSP board" which reduces the shaping time from 1 μs to 0.5 μs [45]. During calibrations with cosmic rays, the average output signal from the crystals is measured to be approximately 5,000 photoelectrons per 1 MeV, with a noise level of approximately 200 keV. The energy resolution for photons, σ_E/E, ranges from approximately 2.5% at 100 MeV to approximately 1.7% at 5 GeV [46].

2.3.5 K_L^0 and μ Detector

The K_L^0 and muon detector (KLM) [48] is made of iron plates and active detector elements arranged in an alternating sandwich pattern outside of the superconducting solenoid. The iron is required to serve as the solenoid's magnetic flux return and allows the K_L^0 to shower in its volume hadronically with a material budget of 3.9 interaction lengths [27]. The KLM has an angular acceptance of 20°< θ < 155°including the barrel region and two end caps (forward and backward). The outermost layers of the barrel section consist of modules made of two coupled Resistive Plate Chambers (RPC) with independent power supplies. In contrast to Belle, Belle II expects significant background rates in the endcaps and the innermost layers in the barrel area of the KLM detection system

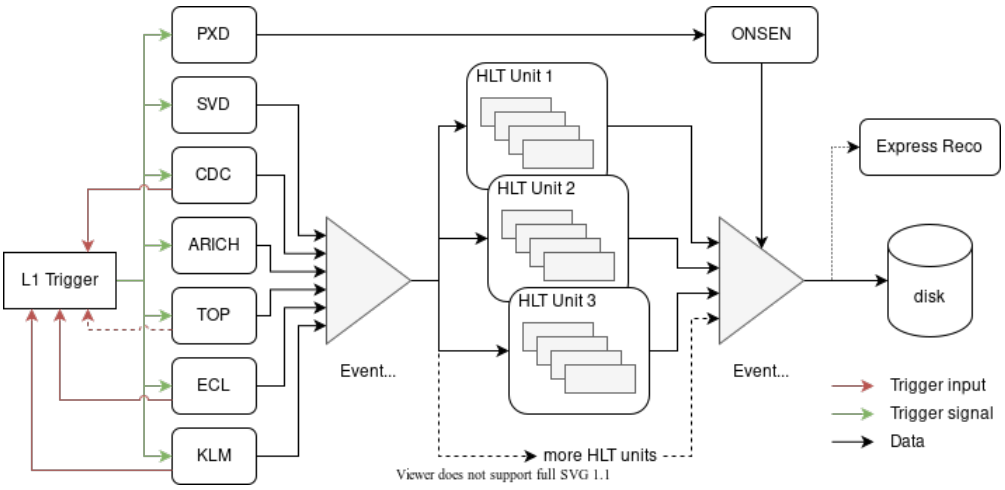


Figure 2.9: Simplified diagram of data flow in the Belle II trigger system. [50]

due to the neutrons generated mainly by the background processes in electromagnetic showers. Because of the large hit rate, the longer dead-time of RPCs increases the muon misidentification probability. That's why scintillator strips with silicon photomultipliers (SiPM) are installed in the inner two layers and endcaps to avoid the problem.

The KLM system achieves a muon detection efficiency of 89% for tracks with momentum above 1 GeV/c, at approximately 1.3% hadron misidentification rate. The K_L^0 detection efficiency increases almost linearly with momentum to a plateau of 80% at 3 GeV/c [26].

2.3.6 Trigger System

The efficient collection and analysis of large volumes of data from the SuperKEKB accelerator in Belle II heavily depend on the trigger system. It is required to have an efficiency of approximately 100% for hadronic and low multiplicity events, with an expected event rate of 15 kHz at the goal luminosity of $8 \times 1035\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$. The maximum acquisition frequency of the data acquisition system (DAQ) limits the trigger rate to 30 kHz. Additional requirements are a timing precision of around 10 ns and a minimum time separation between two events lower than 200 ns [49]. Overall this system comprises the hardware-based Level 1 (L1) trigger and the software-based High-Level Trigger (HLT). Figure 2.9 describes the overall data flow in the Belle II trigger system.

The CDC and the ECL sub-triggers provide the primary information to the L1 trigger. The CDC enables the reconstruction of the z-coordinate of the primary vertex of an event in a few microseconds, reducing background produced far from the interaction region [26]. The ECL generates fast trigger signals based on the total energy released in the calorimeter and the number of isolated clusters. The reconstructed information and the trigger signals from the CDC and the ECL are merged with the hit information from the TOP and KLM by the Global Reconstruction Logic (GRL), where a low-level reconstruction is performed. The results from the GRL are sent to the Global decision Logic (GDL), which issues a trigger based on the input information, and the L1 output is then sent to the HLT for further background rejection.

The HLT is a type of software trigger that operates on a server farm and is designed to reduce the background and limit the event rate from the L1 trigger, which can generate up to 30 kHz, to a maximum of 10 kHz for offline storage. This trigger determines whether an event should be stored or discarded based on the computed values of various event topology variables and particle properties. It uses all available subdetector information, except for the PXD, to reconstruct events from the L1 trigger, using the same software as in the offline reconstruction process. The track parameters obtained from the online reconstruction done by the HLT are extrapolated to the PXD layers, where a region-of-interest(ROI) is identified. Only pixels within this region are read, and the corresponding PXD hits are utilized to complete the event reconstruction before final storage.

2.4 Simulation

The Monte Carlo approach, used for simulation, involves using event generators, which are software packages that produce sets of four-vectors representing final states of $e^+ e^-$ collisions according to theoretical models of particle kinematics and interactions. The generated data is then subjected to detector simulation, where models of the detector geometry and material are combined with models of interactions of particles with matter

and signal formation to recreate expected observed quantities. The resulting simulated data contains information about reconstructed particles as well as the generated true particles. By matching these two, we can determine if the observed particles have been reconstructed accurately or if there are any frequent instances of misreconstruction or background. This truth matching procedure is crucial for optimizing selection requirements, calculating signal efficiency, classifying sample components, and ensuring consistency.

PYTHIA [51] is used to simulate the properties of virtual photons created in electron-positron annihilation and their subsequent splitting into a quark-antiquark pair that produces the observed hadrons. And generation of τ pairs is simulated by KKMC [52], while BABAYAGA.NLO [53] is used to simulate QED background processes such as $e^+ e^- \to e^+ e^- (\gamma)$ and $e^+ e^- \to \gamma\gamma(\gamma)$. $e^+ e^- \to e^+ e^- e^+ e^-$ and $e^+ e^- \to e^+ e^- \mu^+ \mu^-$ are simulated by AAFH [54]. EvtGen [55] models generate the decay of heavy hadrons and TAUOLA [56] models the τ decays. PHOTOS [57] generates photon emission by final-state charged particles. The generated particles are then passed through the GSIM, which is the standard Belle detector simulation model based on the CERN GEANT4 [58] package. GSIM simulates the interaction of the generated particles with the detector material and signal formation, resulting in simulated data in the same format as experimental data. This simulation samples take into account the impact of beam-induced background, which results from the scattering and loss of beam particles due to the Touschek effect and beam-gas scattering. Additionally, the backgrounds caused by Bhabha scattering and two-photon quantum electrodynamic processes, which are luminosity-dependent, are also considered.

2.5 Reconstruction

The reconstruction process is essential to convert the initial raw detector responses into usable objects at the analysis level. This process is performed on both simulated

and actual data, where the simulated data provides access to the true generator level information. Each subdetector's output information comprises raw objects that contain digitized signals, which are then processed to generate low-level objects like detector hits and digits. In the reconstruction process, specific algorithms are employed to produce higher-level information from the raw objects like ECL clusters, PID variables, and tracks. Afterward, the hits and cell-level information is discarded, thereby reducing the event size by roughly 40 times [35]. The reduced information, comprising high-level objects, is then used to determine particle-level information, such as four-momentum and event shape variables.

2.5.1 Clustering

Clustering is an initial step in the reconstruction process that involves combining detector responses. This technique is employed by various detectors in Belle II. The fundamental principle is to collect adjacent pixels, strips, or crystals (depending on the detector) with a readout signal surpassing the threshold and merge them into clusters from which position and other characteristic features can be measured.

2.5.2 Tracking

The process of Belle II track reconstruction involves two main components: track finding and track fitting, as shown in Fig. 2.10. The objective of track finding is to identify sets of hits or clusters in the tracking detectors (PXD, SVD, and CDC) that correspond to the same charged particle trajectory. Since the tracking detectors possess distinct properties, different algorithms are employed for each of them. Initially, the CDC signals are filtered and reconstructed by two independent algorithms: a global track-finding algorithm based on the Legendre algorithm [59] and a local algorithm utilizing a cellular automaton [60]. The outcomes of these two algorithms are merged, and the CDC-only tracks are fitted via a deterministic annealing filter (DAF) [61]. The CDC tracks are supplemented with SVD clusters using a combinatorial Kalman filter (CKF) [62]. In cases where high-curvature tracks fail to produce sufficient hits in the CDC, a standalone SVD track-finder that uses

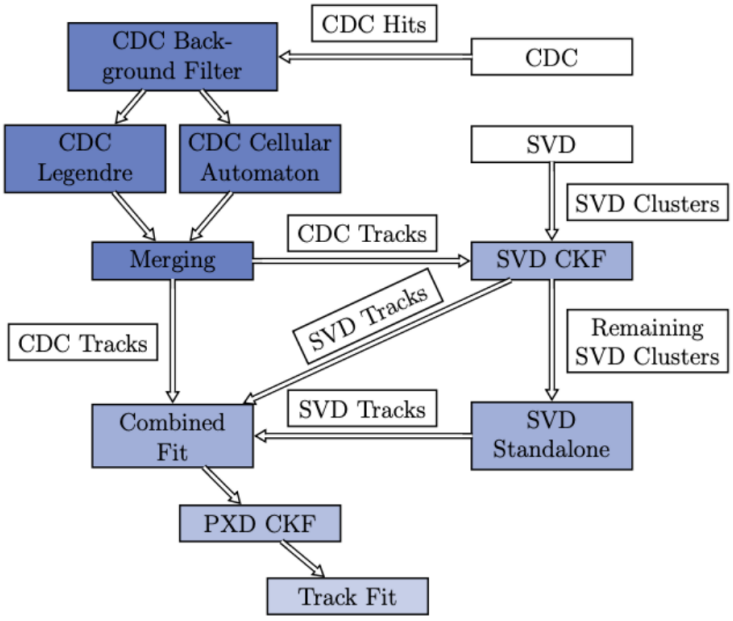


Figure 2.10: Simplified diagram of the steps of tracking workflow at Belle II [63]

a sophisticated filter concept known as Sector Map [36] and a cellular automaton is utilized. The outcomes are then combined, refitted using a deterministic annealing filter (DAF), and extrapolated to the PXD with a second CKF.

The last step of the track-reconstruction process involves the track fit, which is performed using the DAF provided by the GENFIT2 package [61]. In order to correctly calculate the energy loss and material effects, a specific particle hypothesis is assumed for the fit. During the track fit five helix parameters are estimated, which are:

- d_0: the signed distance of the point of closest approach (POCA) to the z axis.
- z_0: the longitudinal signed distance of the POCA from the origin.
- ϕ_0: the angle defined by the x axis and the track transverse momentum at the POCA.
- ω: the track curvature signed according to the particle charge.
- $\tan\lambda$: and the tangent of the angle between the momentum at the perigee and the transverse plane.

2.5.3 Charge Particle Identification

The detection of final-state charged particles depends on their interaction with the detector. In the case of charged hadrons, data collected from the TOP and ARICH are combined with specific ionization (dE/dx) measurements from the SVD and CDC. Similarly, the ECL provides primary information for electron identification, while the KLM is used for muon identification. To identify long-lived charged particle hypotheses, such as $e^+, \mu^+, \pi^+, K^+, p^+, d^+$, PID information from each subdetector is analyzed independently to determine an individual detector-based likelihood, $\mathcal{L}_i^{\text{det}}$. The likelihoods for each hypothesis are then combined to create a global likelihood for each hypothesis i,

$$\mathcal{L}_i = \prod_{\boldsymbol{det}} \mathcal{L}_i^{\text{det}}$$

Likelihood ratios are the PID indicators presented to the analyst as the final variables:

- global ratios,

$$\text{PID}_i = \frac{\mathcal{L}_i}{\sum_{j=e,\mu,\pi,K,p,d} \mathcal{L}_j} \tag{2.1}$$

- binary ratios,

$$\mathcal{L}(\alpha \mid \beta) = \frac{\mathcal{L}_\alpha}{\mathcal{L}_\alpha + \mathcal{L}_\beta} \tag{2.2}$$

If the value of $\mathcal{L}(\alpha \mid \beta)$ is greater than 0.5, it indicates that the track of a given particle is more likely to belong to the particle type α, rather than β.

2.5.4 Neutral Particle Identification

Neutral particles, such as neutral pions, photons, and K_L^0, are identified using different techniques. The identification of photons within the ECL involves analyzing the shower shape of clusters that do not match any reconstructed track. The method exploits the property that electromagnetic showers caused by photons have cylindrical symmetry around their direction and exhibit an exponential decrease in energy deposition from the incident axis. The identification process faces a significant challenge from the dominant background of hadron interactions, which result in asymmetric showers and multiple

ECL clusters not linked to any charged track.

Neutral pion reconstruction involves combining two-photon candidates. If the pion energy is below 1 GeV, the two photons create non-overlapping ECL clusters. When the energy of π^0 is between 1 GeV and 2.5 GeV, the two clusters overlap, but they can still be reconstructed as two separate photon candidates in the ECL. The energy of the π^0 can be reconstructed directly from the four-momenta. Additionally, the energy resolution is improved by mass constrained fit of the two-photon candidates to the nominal π^0 mass. When energy of π^0 is high(> 2.5 GeV), two photons are indistinguishable and are reconstructed as one photon [35].

The ECL and KLM detectors are used to identify K_L^0 mesons using boosted decision trees (BDT), described in Sec. 3.1. ECL and KLM clusters are independently classified as originating from K_L^0 based on their features. Neutrons and photons are the main sources of background, with the latter mostly originating from beam interactions with the detector and beam-pipe, and neutral particle products resulting from these interactions. The BDT input includes several features from the KLM detector, with the most prominent ones are listed below.

- For neutral clusters, the distance to the nearest track should be significant.
- Clusters originating from beam background should appear earlier than those from primary collisions.
- Hadronic clusters tend to have a larger radius than electromagnetic clusters.

The crucial characteristics for the ECL are as follows:

- Neutral clusters are expected to have few nearby tracks.
- The distribution shape of the energy ratio in the inner 3x3 and outer 5x5-4 clusters depends on the shower's origin, whether hadronic or electromagnetic.
- Normally, K_L^0 deposits less than 50 MeV of energy in the cluster.

2.6 Belle II Analysis Software Framework (basf2)

The basf2 [64] is a software framework used in Belle II for various tasks such as data taking, offline reprocessing, and data analysis. It consists of processing modules written in C++ [65] or Python [66], which are executed linearly within a defined path. Modules communicate with a common object store called DataStore, which keeps track of object relations in each event. The software operates on the principle of lazy computations, where the user specifies processing blocks, and the script starts the procedure only after explicitly specifying the processing command.

CHAPTER 3

Analysis techniques

In this chapter, the statistical tools and techniques employed in this thesis are the main focus. The presentation of multivariate classification techniques, along with the algorithms that have been specifically used in this study, can be found in Section 3.1. The concept of statistical inference is defined in Section 3.2, while Section 3.3delves into the fundamental aspects of the maximum likelihood method, providing the necessary context for the data fitting tools used in this research. Finally, Section 3.4 deals with the ${}_s\mathcal{P}lot$ technique.

3.1 Multivariate Analysis

The use of multivariate analysis (MVA) has become indispensable in analyzing data from colliders. Machine learning algorithms relying on multivariate classification provide greater insight into the collected data, particularly in measurements that are background-dominated and where the signal is hard to detect. Most often, MVA methods are utilized as binary classifiers. Here, a multidimensional feature vector is projected into a one-dimensional test-statistic, designed to separate the signal from the background and increase the signal-to-background ratio in an analysis. In contrast to simple cuts on independent features, MVA methods use information about correlations between features, thereby resulting in a better test-statistic in many cases. To adapt its model specific parameters to a given classification task, MVA methods must be trained on an independent dataset. Many MVA methods also require hyper-parameters, which are not automatically tuned by the training process and need to be chosen by the user.

3.1.1 Boosted Decision Trees

The process of a decision tree involves partitioning the feature vector into distinct rectangular regions to effectively separate the signal from the background. Figure 3.1 provides a schematic outline of this process, which involves ordering the binary cuts hierarchically in a tree and applying them successively. As a result, the feature space is divided into regions of varying sizes, each of which has its own signal fraction that serves as a classifier output when the data point is within that region. The nodes of a decision tree correspond to its cuts, while the regions are referred to as leaf nodes.

The generalization capability of a single decision tree is limited, meaning that it tends to memorize statistical features of the specific dataset used for training and performs poorly on new, independent data. This issue, known as over-training, can be addressed by restricting the depth of the tree, i.e., the number of consecutive cuts. However, such a shallow tree, referred to as a weak-learner, is not very effective in separating signal and background.

Although a single weak-learner may not be effective for classification tasks, combining many weak-learners can create a robust classifier that is resistant to over-training. One such method is the Boosting algorithm, such as Gradient Boosting [67], which assigns weights to each event in the training dataset to train a weak learner. The weights of incorrectly (correctly) classified events are increased (decreased) and used to train a new weak-learner. This process is repeated several hundred times. Finally, the classifier output is determined by the weighted sum of the outputs of the individual weak-learners. Boosting algorithms construct a "forest" of weak-learners and are categorized as ensemble methods.

3.2 Statistical inference

Experimental particle physics seeks to uncover natural phenomena by conducting precise measurements and analyzing the resulting data. In order to achieve this goal, experiments

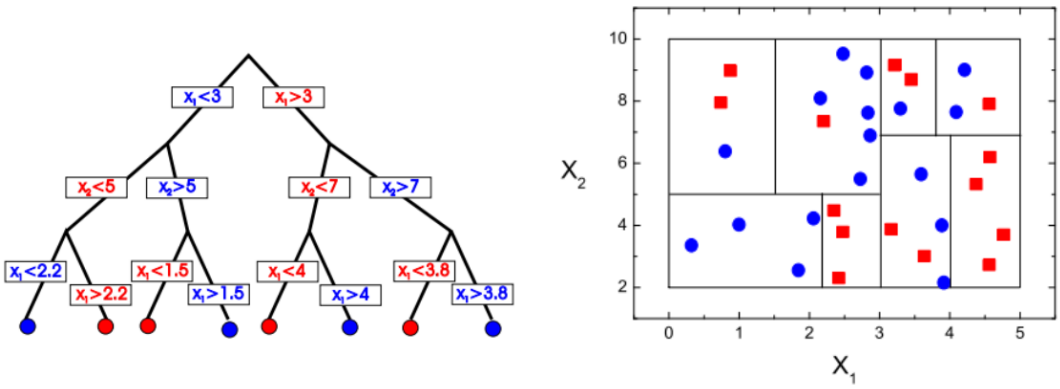


Figure 3.1: Schematic outline of a single Decision Tree (left) and its regions inside a two-dimensional feature vector resulting from the cuts in the tree (right) [68].

like Belle II collect millions of collision events, each containing vast amounts of unique data due to the inherent randomness of physics processes and detector responses. Theoretical predictions determine the expected distributions of the observables measured in the data, with these predictions being parameterized by particle couplings, branching fractions, and other similar parameters. Experimentalists aim to extract information on these parameters from the data distribution, which is known as statistical inference.

3.3 Maximum Likelihood Estimation

Maximum likelihood estimation (MLE) is, the most popular procedure for statistical inference used in High energy physics, a statistical method used to estimate the parameters of a probability distribution based on observed data. The method involves finding the parameter values that maximize the likelihood function, which is a function that measures how well the observed data fits the probability distribution.

Here are the key steps involved in maximum likelihood estimation:

1) **Choose a Probability Density Function (PDF):** The first step in MLE is to choose a PDF that describes the distribution of the data. The choice of the PDF is based on some assumptions about the distribution of the data. It could be a

simple function like a Gaussian or a more complex function that includes multiple components.

II) **Define the Likelihood Function:** The likelihood function is defined as the joint probability density function of the observed data, given the unknown parameters. In other words, the likelihood function is a function of the parameters and the observed data, and it measures how well the observed data fits the PDF.

III) **Maximize the Likelihood Function:** The next step is to maximize the likelihood function by finding the parameter values that make the observed data most probable under the chosen PDF. This is done by taking the derivative of the likelihood function with respect to each parameter and setting them equal to zero. The resulting equations are then solved to find the values of the parameters that maximize the likelihood function.

IV) **Evaluate the Fit:** Once the parameter values that maximize the likelihood function have been found, we evaluate the fit by comparing the observed data with the predicted PDF using these parameter values. The goodness of fit can be evaluated using various statistical tests, such as the chi-squared test.

V) **Interpret the Results:** Finally, we interpret the results of MLE by extracting the values of the parameters that maximize the likelihood function. These parameter values represent the best estimate of the true values of the parameters of the probability distribution.

For example, let's say we have a set of data points $\{x_1, x_2, ..., x_n\}$ that we believe are normally distributed with unknown mean (μ) and standard deviation (σ). To construct the likelihood function, we first write down the PDF of the normal distribution as

$$f\left(x_i \mid \mu, \sigma\right) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(\frac{-(x_i - \mu)^2}{2\sigma^2}\right)$$

where x_i is a single data point, μ is the mean, σ is the standard deviation, and $f(x_i \mid \mu, \sigma)$

is the probability density function of observing x_i given μ and σ.

To obtain the likelihood function for the entire dataset, we take the product of the PDF for each individual data point:

$$\mathcal{L}(\mu, \sigma) = \prod_{i=1}^{n} f\left(x_i \mid \mu, \sigma\right)$$

where $\mathcal{L}(\mu, \sigma)$ is the likelihood function, and the product is taken over all the data points in the dataset.

Now that we have defined the likelihood function, our goal is to find the values of μ and σ that maximize this function. This involves taking the derivative of the likelihood function with respect to each parameter (μ and σ) and setting them equal to zero:

$$\frac{\partial \mathcal{L}(\mu, \sigma)}{\partial \mu} = 0$$

$$\frac{\partial \mathcal{L}(\mu, \sigma)}{\partial \sigma} = 0$$

The resulting equations are then solved to find the values of the parameters that maximize the likelihood function. As the likelihood function is product of the PDF of the each data point, it is easier to work with the logarithm of the likelihood function, $\ln \mathcal{L}(\mu, \sigma)$.

The likelihood function is a crucial component of MLE because it is the function that we are trying to maximize to obtain the best estimates of the unknown parameters. The better the likelihood function fits the observed data, the more confident we can be in the estimates of the parameters that we obtain.

One of the key advantages of MLE is that it provides unbiased estimates of the parameters, meaning that the estimates are centered around the true values of the parameters. Another advantage of MLE is that it provides estimates that are efficient, meaning that they have the smallest variance compared to other unbiased estimators.

3.4 ${}_s\mathcal{P}lot$

Separating the signal from the background can be a challenging task, especially when the signal is rare or difficult to distinguish from the background. The ${}_s\mathcal{P}lot$ technique was developed to address this problem. It is based on the idea that the signal and background events have different properties that can be used to distinguish them. By assigning weights to each event based on its likelihood of being a signal or a background event, this technique helps to improve the statistical significance of the signal in the dataset by reducing the effect of the background events. This method involves a maximum likelihood fit to the data to determine the weights.

The ${}_s\mathcal{P}lot$ method involves the following steps:

I) **Calculate the Probability of Each Event Being a Signal or a Background Event:** The fitted PDF in Maximum likelihood estimation is used to calculate the probability of each event being a signal or a background event. This is done by calculating the probability density function for each event using the fitted PDF. The PDF is used to assign a probability to each event that it belongs to the signal or background category.

II) **Calculate the sWeights for Each Event:** The sWeights are the weights assigned to each event based on their likelihood of being a signal or a background event. It is calculated using the probability of each event being a signal or a background event. It is normalized to ensure that the sum of the signal sWeights and the sum of the background sWeights are equal to the total number of signal events and background events in the dataset, respectively.

III) **Reweight the Dataset:** Once the sWeights are calculated, the dataset is reweighted by multiplying the sWeights of each event by its original weight. This creates a new dataset where each event is weighted according to its likelihood of being a signal or a background event.

iv) **Perform a New Analysis:** Finally, a new analysis is performed using the reweighted dataset. This analysis should show an improvement in the statistical significance of the signal compared to the original dataset. The improvement in statistical significance is achieved by reducing the effect of the background events on the analysis results. The resulting distribution gives an estimate of the signal distribution, which is free from the effects of the background.

CHAPTER 4

Analysis of $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ signal

This chapter outlines the methodology for reconstructing the $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ decay, where the $\bar{D}^0$ decays to $K^+\pi^-$. The data samples used for this analysis are discussed in Section 4.1, while the object selections are summarized in Section 4.2. The selection of $\bar{D}^0$ and B^+ are outlined in Sections 4.3 and 4.4, respectively. The process of rejecting peaking and continuum background is described in Sections 4.5 and 4.6, respectively. Finally, Section 4.7 explains the single-candidate selection and provides an overview of the signal efficiencies.

4.1 Data Sample

In this thesis, all analyses have been performed using simulated data. Our future objective is to analyze experimental data after optimizing selection criteria, estimating signal efficiencies, and identifying various sources of background based on the analysis of simulated data.

4.1.1 Simulated Data

The samples of simulated Monte Carlo (MC) data used consist of events from seven background categories, including four categories of continuum events involving the production of light-quark pairs ($e^+\ e^- \to q\bar{q}$, where $q = u, d, c, s$) as well as two categories of generic decays of $B^0\bar{B}^0$ and B^+B^- meson pairs produced on resonance and one category of tau pairs ($e^+\ e^- \to \tau^+\ \tau^-$). For $B\bar{B}$ background studies, 200 fb^{-1} generic MC samples were used, while 1 ab^{-1} generic MC samples were used to investigate continuum background.

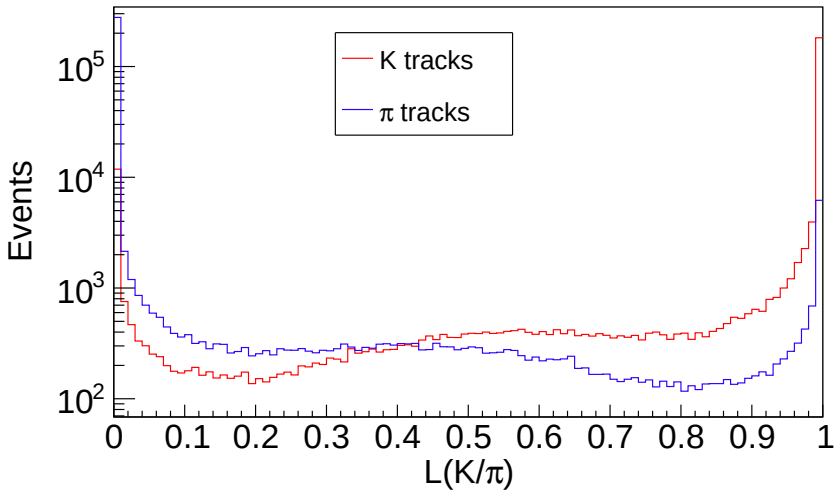


Figure 4.1: Distribution of $\mathcal{L}(K \mid \pi)$ for charged kaon and pion tracks in the signal MC sample

Simulated Signal

One million official signal samples were used to train multivariate discriminants during the continuum background analysis and to estimate signal efficiencies. The signal sample consisted of $\Upsilon(4S)$ decays to B^+B^- events, in which one of the B mesons decayed to $\bar{D}^0\pi^+\pi^-\pi^+$ and $\bar{D}^0$ decayed to $K^+\pi^-$. The EVTGEN model used to generate the $\Upsilon(4S) \rightarrow B^+B^-$, $B^+ \rightarrow \bar{D}^0\pi^+\pi^-\pi^+$, and $\bar{D}^0 \rightarrow K^+\pi^-$ decays and the B^- decays inclusively.

4.2 Object Selection

The selection of tracks involves the requirement of $|d0| < 0.2$ cm and $|z0| < 1$ cm, which correspond to the distance of the closest approach to the interaction point (IP) in the transverse plane to the beam direction and beam direction, respectively. Such criteria eliminate poorly-reconstructed tracks that do not originate from the IP region. Additionally, the criterion $\cos\theta > -0.6$ is applied to remove backward tracks (outside the PID sub-detectors: TOP or ARICH acceptance), reducing the K - π misidentification rate. Information from the sub-detectors CDC, TOF and ARICH is then used to identify these tracks as either a kaon or pion, constructing the binary likelihood, $\mathcal{L}(K \mid \pi)$ of the

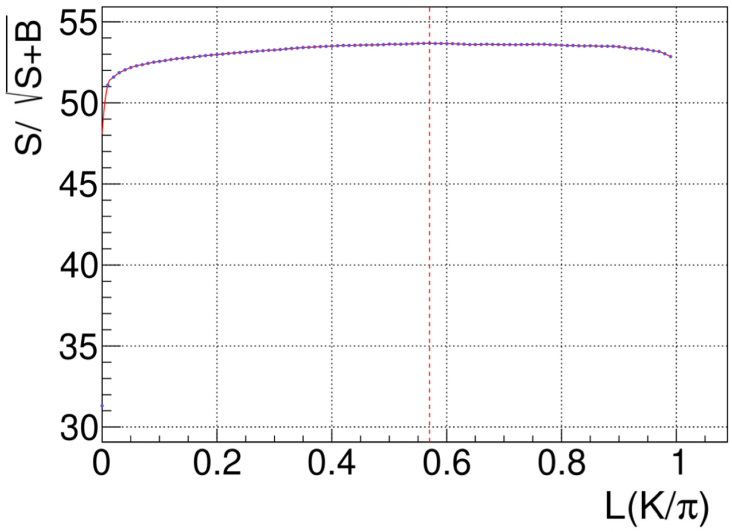


Figure 4.2: Distribution of $S/\sqrt{S+B}$ for PID selection optimization for kaon tracks originating from $\bar{D}^0$

particle being either a pion or a kaon using equation 2.2. The $\mathcal{L}(K \mid \pi)$ distributions for π^+ and K^+ tracks in the signal MC sample are depicted in Fig. 4.1. To distinguish the kaons from pions, a requirement of $\mathcal{L}(K \mid \pi) > 0.6$ is utilized and optimized using the metric $S/\sqrt{S+B}$, where S and B denote the total signal and background events present in the window $|\Delta E| < 0.07$ GeV and $5.266\ \text{GeV/c}^2 < M_{bc} < 5.29\ \text{GeV/c}^2$, respectively, with a specific condition on $\mathcal{L}(K \mid \pi)$. This metric corresponds to the statistically significance of the signal over background in a counting experiment within the defined region. The optimization of $S/\sqrt{S+B}$ is illustrated in Fig. 4.2. The red dashed line in the plot corresponds to the maximum value of $S/\sqrt{S+B}$, which occurs at $\mathcal{L}(K \mid \pi) = 0.57$. For simplicity, we rounded off this value to 0.6 and used it for our analysis.

4.3 Selection of $\bar{D}^0$ Candidates

The reconstruction of $\bar{D}^0$ mesons is carried out by selecting a pair of oppositely charged pion-kaon tracks. The invariant mass of the daughter particles is then fitted, as demonstrated in Figure 4.3. This fit model is not perfect, as seen in the residuals, but is adequate to define the cut range. The selection of the $\bar{D}^0$ mesons is made using the

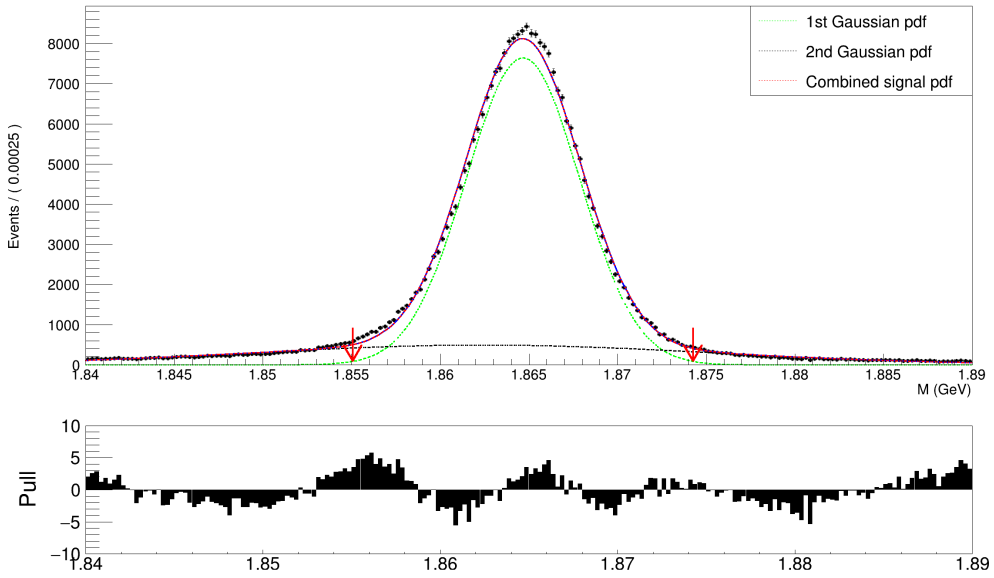


Figure 4.3: Fitting invariant mass of daughter particles (K^+, π^-) of $\bar{D}^0$

$\pm 3\sigma$ region around the mean, which results in a selection range of (1.855, 1.874) Gev/c^2. Additionally, to enhance the resolution of the selected $\bar{D}^0$ candidates' four-momentum, a kinematic fit is then applied to the selected $\bar{D}^0$ candidates, which constrains the reconstructed mass to the known $\bar{D}^0$ mass and the decay products to a common vertex point.

4.4 Selection of B Candidates

To reconstruct a B meson candidate, a D meson candidate is combined with three charged pion tracks. Discriminating B decays from combinatorial or partially reconstructed background involves using the beam-constrained mass

$$M_{\mathrm{bc}} = \sqrt{E_{\mathrm{beam}}^2 - \left(\sum_i^{N_{\mathrm{daughter}}} \vec{p}_i\right)^2} \tag{4.1}$$

and beam energy difference

$$\Delta E = \sum_i^{N_{\mathrm{daughter}}} E_i - E_{\mathrm{beam}} \tag{4.2}$$

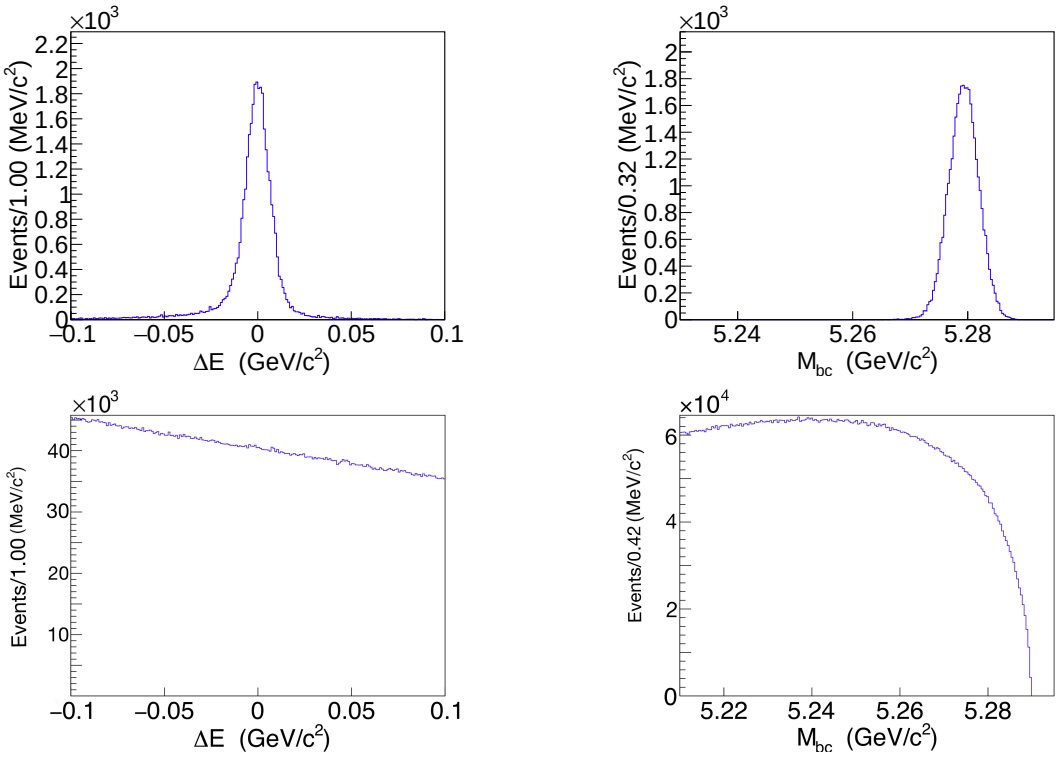


Figure 4.4: ΔE (top left) and M_{bc} (top right) distributions of signal MC sample. ΔE (bottom left) and M_{bc} (bottom right) distributions of generic background MC sample.

as kinematic variables. Here, E_{beam} refers to the beam energy in the center-of-mass (c.m.) frame, and E_i and $\vec{p}_i$ represent the energy and momenta of the daughter particles in the c.m. frame. In the case of correctly reconstructed signal events, the M_{bc} value peaks at the nominal B meson mass, and ΔE is zero. Candidates that fulfill the conditions $|\Delta E| < 0.1$ GeV and $5.27\ \text{GeV/c}^2 < M_{bc} < 5.29\ \text{GeV/c}^2$ are selected. The distributions of ΔE and M_{bc} in both signal and generic background MC events are shown in Fig. 4.4. To ensure that the B daughters originate from a common vertex, a kinematic vertex fit is conducted.

4.5 Suppression of peaking background

After performing topological analysis [69] on truth-matched (TM) events reconstructed from generic Monte Carlo (MC), it was observed that various intermediate resonance

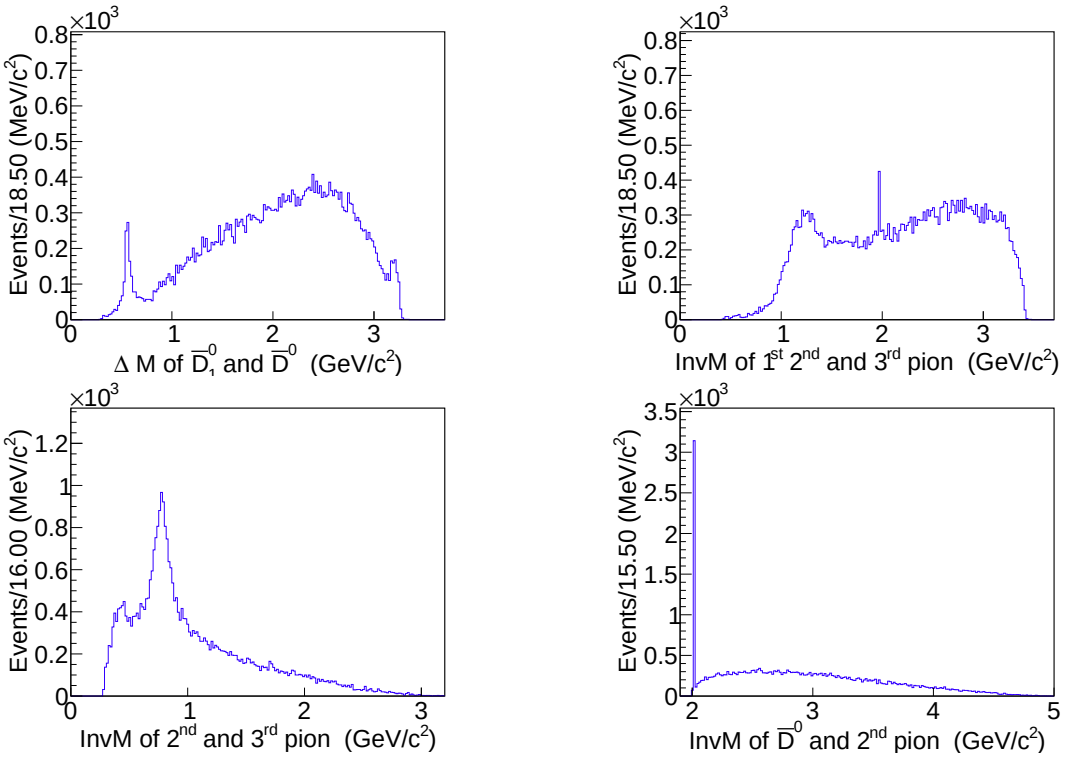


Figure 4.5: Distribution of the invariant mass of different pairs of daughter particles of the B meson

states, such as D^{*-}, $\bar{D}_1^0$, D_s^-, and K_S^0, contribute to the peaking background. The source of this background was confirmed by examining the invariant mass distribution of different combinations of daughter particles of the B meson. Some of the combinations where a peak was observed are shown in Fig. 4.5. Additionally, we observe a background contribution from K_S^0, which produces two oppositely charged π out of three daughter π of the B meson. While this contribution is not significant in $B\bar{B}$ events, it is noticeable in $s\bar{s}$ and $c\bar{c}$ profiles, as shown in Fig. 4.6.

In addition to the sources of peaking background mentioned earlier, there is also a correlated background source from $\bar{D}^0$ and π in the Rest of the Event (ROE), originating from D^{*-}. This background arises when attempting to reconstruct B^+B^- in an event where $B^0\bar{B}^0$ is produced and B^0 decays to $D^{*-}X^+$, with D^{*-} decaying to $\bar{D}^0\pi$ and X^+ decaying to final state particles of three charged π. During reconstruction, to conserve charge in such an event, one π from D^{*-} is neglected, resulting in this background. To

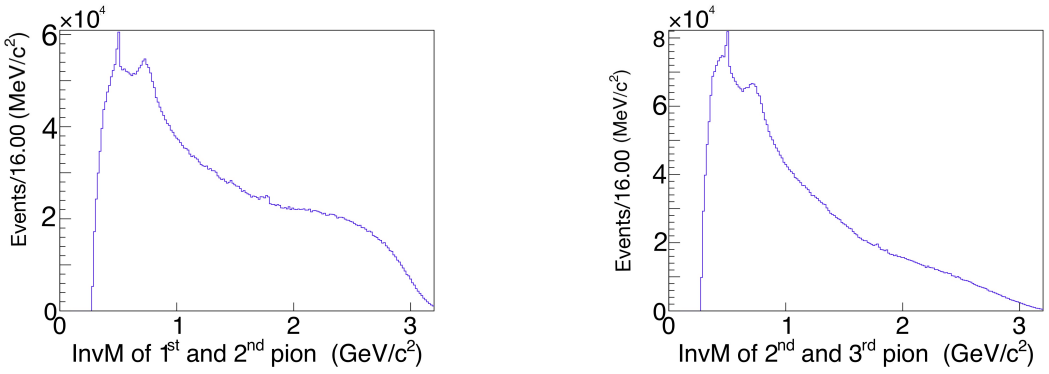


Figure 4.6: Invariant mass distributions of two combinations of oppositely charged π (out of three daughter π) in $s\bar{s}$ as well as $c\bar{c}$ profile.

address this issue, D^{*-} was reconstructed from $\bar{D}^0$ on the signal side and π on the ROE side. The mass difference between D^{*-} and $\bar{D}^0$ is shown in Figure 4.7.

As the process of suppressing background contribution from the strong intermediate resonance state involves much more complexity, we have opted to use a veto method to suppress the weak intermediate resonance state instead. Gaussian PDFs were used to fit these distributions, along with some background PDFs, and $\pm 3\sigma$ window around the mean was chosen for suppressing the peaking background. The veto cuts along with their efficiencies are mentioned in Table 4.1.

Table 4.1: Different veto cuts and their efficiencies.

Variable name	Veto region (GeV/c^2)	Efficiency (%)
Invariant mass of three charged π	(1.95769 - 1.97872)	98.52
Invariant mass of one charged π and $\bar{D}^0$	(2.00837 - 2.01217)	91.35
Mass difference of $\bar{D}_1^0$ and $\bar{D}^0$	(0.492200 - 0.632197)	96.75
Invariant mass of two oppositely charged $\pi(1^{st}$ and $2^{nd})$	(0.488814 - 0.506173)	99.08
Invariant mass of two oppositely charged $\pi(2^{nd}$ and $3^{rd})$	(0.48987 - 0.504896)	98.83
Mass difference of D^{*-} and $\bar{D}^0$	(0.143913 - 0.146967)	97.32

4.6 Continuum background suppression

The $e^+e^- \to q\bar{q}$ continuum events have a cross section that is approximately three times larger than that of $e^+e^- \to B\bar{B}$. Therefore, these events usually constitute the primary

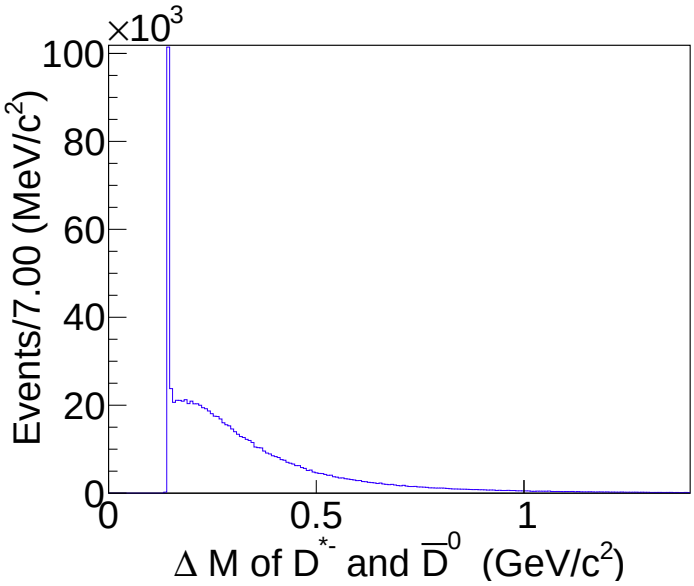


Figure 4.7: Distributin of mass difference between D^{*-}, reconstructed from $\bar{D}^0$ of signal side and π of ROE side, and $\bar{D}^0$

background for any B meson decay. To address this issue, the topological differences between the two processes are exploited. The decay products from $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ are uniformly distributed over the 4π solid angle since the B mesons do not have significant momentum in the c.m. frame. On the other hand, the events from $e^+e^- \rightarrow q\bar{q}$ are highly boosted and have larger momenta due to their small masses. As these events are produced back-to-back, they exhibit a jet-like structure, as illustrated in Fig. 4.8. The FBDT is fed with ten topological variables that are based on angular configuration, displaced vertices, and information associated with the flavor of the other B meson, to reject the continuum background. Initially, the training was performed with many input variables, but later, the less discriminative variables that had negligible impact on the FBDT performance were removed.

Variables used to suppress continuum

The FBDT is provided with following input variables:

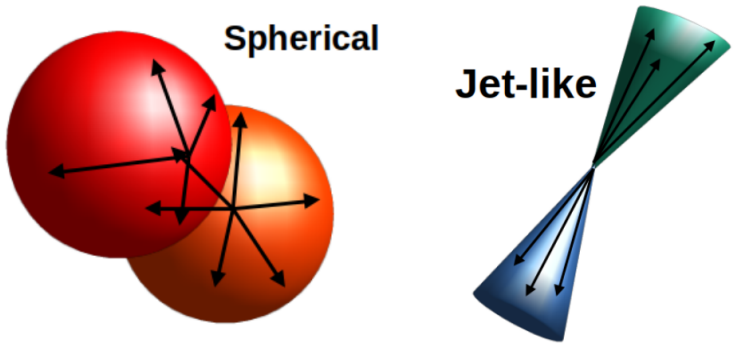


Figure 4.8: Event topology of $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ (left) and $e^+e^- \rightarrow q\bar{q}$ (right) events [70].

- **KSFW moments:** The Fox-Wolfram moments are defined by

$$H_l = \sum_{i,j} |p_i| \left|p_j\right| P_l \left(\cos\theta_{ij}\right) \text{ [71]} \tag{4.3}$$

where p_i and p_j are the momenta of reconstructed particle in an event, P_l is the Legendre Polynomial of order l and $\cos\theta_{ij}$ is the angle between momenta of particles i and j. The purpose of these moments is to characterize the event shape. To improve the discrimination between continuum and $B\bar{B}$ events, the reconstructed tracks are classified according to whether they belong to the "signal" B meson or the "opposite" (tagging side) B meson. Accordingly, the modified Super Fox-Wolfram sub-moments are defined separately for "signal-opposite" and "opposite-opposite" combinations as

$$H_{xl}^{so} = \sum_{i,jx} \left|p_{jx}\right| P_l \left(\cos\theta_{i,jx}\right) \tag{4.4}$$

$$H_l^{oo} = \sum_{j,k} \left|p_j\right| |p_k| P_l \left(\cos\theta_{jk}\right) \tag{4.5}$$

These sub-moments are calculated using all signal particles for i and all opposite particles for j and k. Among the available moments, KSFW E_t, H_0^{oo} and H_{12}^{so} are found to have better discriminating power. Also R2, defined as ratio of the 2^{nd} to the 0^{th} order moments, is a very good discriminating varaible. The distribution of

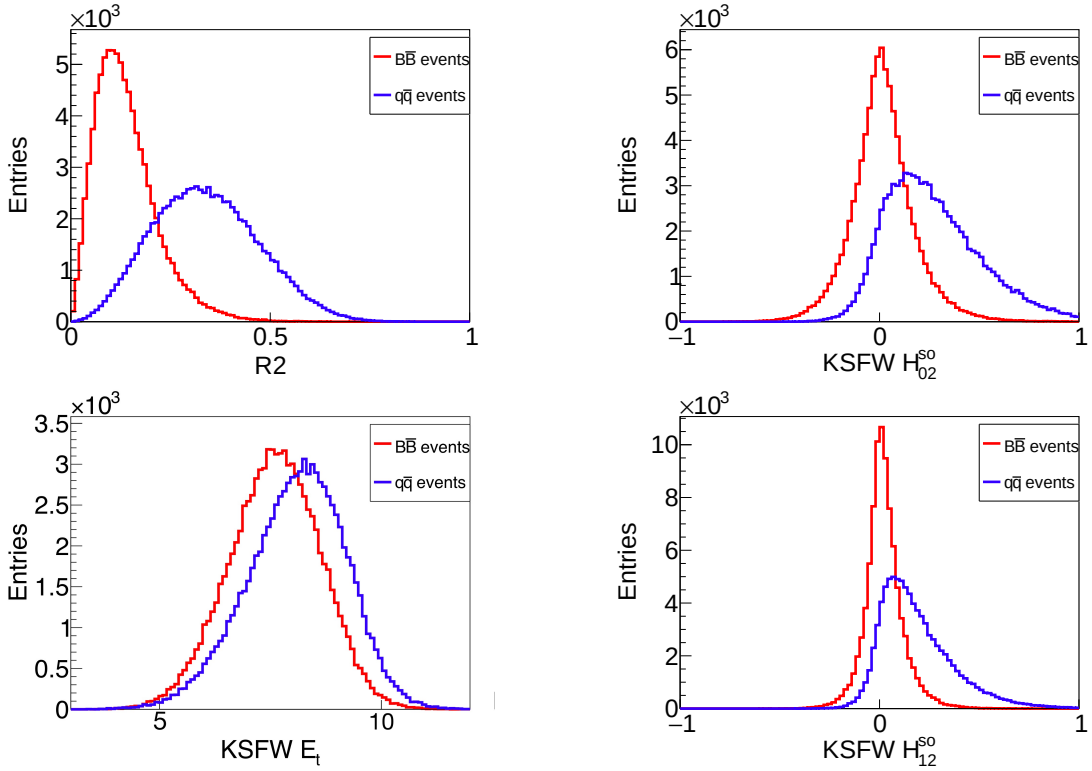


Figure 4.9: Distribution of R2(Left top), KSFW H_{02}^{so}(Right top), KSFW E_t(Left Bottom) and KSFW H_{12}^{so}(Right Bottom)

these variables is illustrated in Fig. 4.9.

- **B meson direction:** The center-of-mass frame's polar angle, θ_B, is used as an independent variable for discrimination. The decay of the spin-1 parent particle, $\Upsilon(4S)$, into two spin-0 daughter particle, B, has a (1 - $\cos^2\theta_B$) angular distribution to conserve the angular momentum with respect to the beam axis. However, the spin-1/2 $q\bar{q}$ and the resulting jets have a uniform distribution in $\cos\theta_B$. Thus, the variable $\cos\theta_B$ is used to distinguish between signal B decays and the continuum background by measuring the polar angle between the reconstructed momentum of the B candidates (calculated in the $\Upsilon(4S)$ reference frame) and the beam axis. The distribution of this variable is depicted in Fig. 4.10.

- **Thrust:** The unit vector indicating the maximum total projection of a group of N momenta p_i is called the thrust axis $\vec{T}$. Here, p_i represents the momentum of the

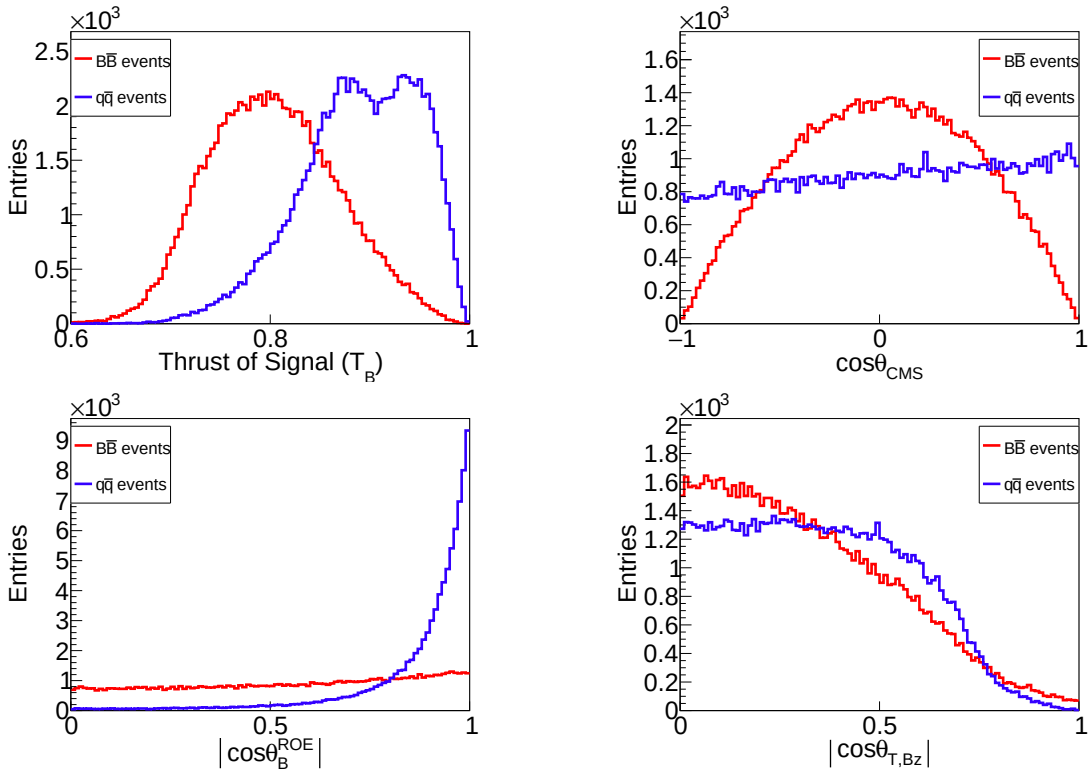


Figure 4.10: Distribution of T_{signal}(Left top), $\cos\theta_{\text{CMS}}$(Right top), $\left|\cos\theta_B^{\text{ROE}}\right|$(Left Bottom) and $\left|\cos\theta_{\text{T,Bz}}\right|$(Right Bottom)

ith particle the magnitude of the thrust axis is defined as

$$T = \frac{\sum_i^N \left|\vec{T}\cdot\vec{p}_i\right|}{\sum_i^N |\vec{p}_i|}. \tag{4.6}$$

Discrimination between signal and continuum events can be achieved using both the thrusts of the signal B and the rest-of-event (ROE). The magnitude of the signal B thrust, denoted as T_{signal}, is also used as input variables, and it's distribution is depicted in Fig. 4.10.

- **Thrust angle:** The angle between the thrust axis of the reconstructed momenta of B meson decay products and the thrust axis of the remaining particles in the event is a useful discriminant variable. In events where a $B\bar{B}$ pair is produced, both B mesons are nearly at rest in the $\Upsilon(4S)$ rest frame, causing their decay

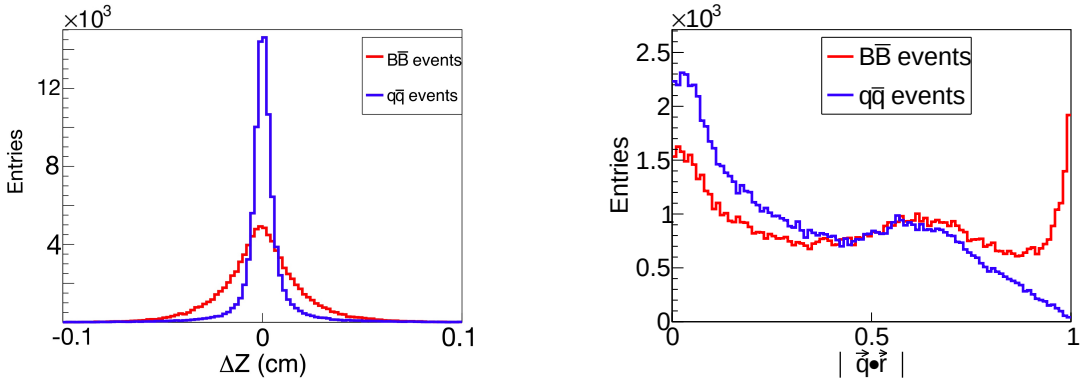


Figure 4.11: Distribution of Δz(Left), and $|q \cdot r|$(Right)

products to be isotropically scattered and uniformly distributed in the range (0, 1) of $|\cos\theta_B^{ROE}|$. On the other hand, in $q\bar{q}$ events, the momenta of particles follow the direction of the jets, resulting in highly collimated thrusts, with the distribution of $|\cos\theta_B^{ROE}|$ peaking towards higher values. Along with this $|\cos\theta_B^{\text{signal}}|$ is also good discriminating variable. The distributions of these variables are illustrated in Fig. 4.10.

- **Vertex position:** Due to their higher momentum, B mesons have longer lifetimes compared to lighter mesons, resulting in a larger average flight distance. To discriminate between $B\bar{B}$ events and continuum events, the difference in the longitudinal components of the decay vertices of the signal B and ROE vertices, $\Delta z = z_{B_{sig}} - z_{B_{tag}}$, is used as a variable in the FBDT classifier. This quantity is wider for $B\bar{B}$ events than for continuum events. The distribution of this variable is illustrated in Fig. 4.11.

- **Flavour tagging:** The flavor tagger (FT) employs a method that merges several multivariate classifiers into a single fast boosted decision tree, using data on the charge of leptons and kaons that are not connected to the signal B decay. This helps identify the flavor q of the signal B candidates, with the tagging efficiency expressed as $\sum \varepsilon_i \times (1-2w_i)^2$, where ε_i indicates the efficiency of the ith classifier,

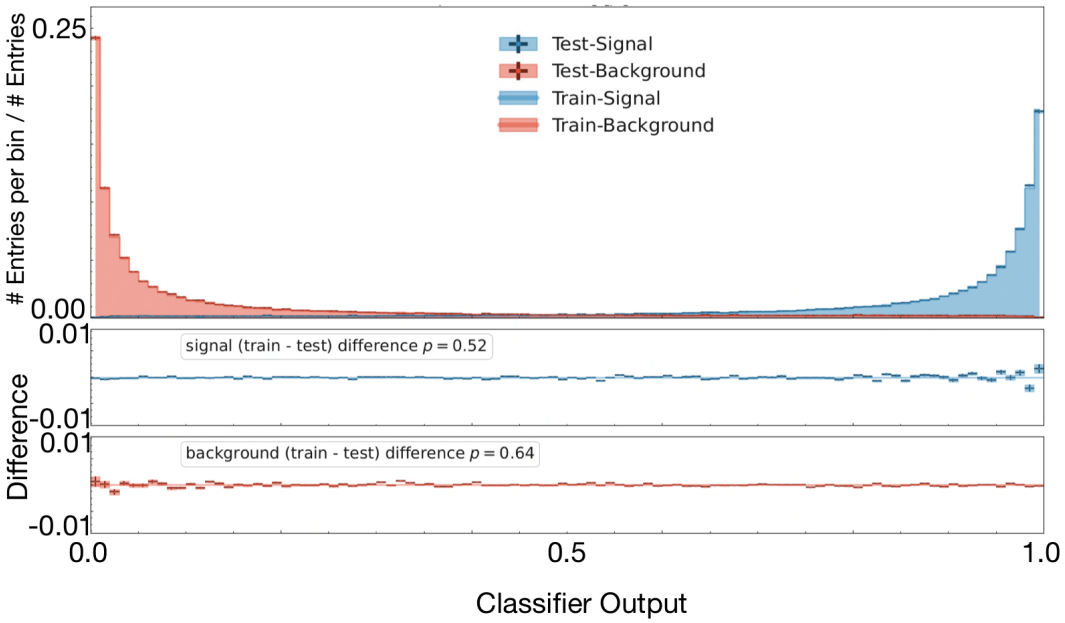


Figure 4.12: Overtraining plot

and w_i is the mis-tag fraction for the flavor. The resulting output of the algorithm is denoted by $|q \cdot r|$, where r represents the quality factor that ranges from zero (for cases where the B-tag flavor cannot be accurately determined) to one (for known B-tag events), and the flavor q is assigned +1 (if the signal and tagging B mesons are believed to have the same flavor) or -1 otherwise. The $|q \cdot r|$ distribution for signal and continuum events is illustrated in Fig. 4.11.

MVA output & Figure of Merit Analysis

The FBDT is trained using around 90,000 signal and continuum MC samples, and it's performance is assessed using independent MC samples to ensure that it is not overtrained, i.e., the FBDT does not solely rely on statistical fluctuations within the training sample. The Figure 4.12 shows the result of a Kolmogorov-Smirnov test, which confirms that the FBDT is not overtrained. As shown in Figure 4.13, the FBDT output (C) has a peak at zero for continuum events and at one for $B\bar{B}$ events. The ROC curve, presented in Figure 3.20, illustrates the signal efficiency vs. background rejection performance of the FBDT. The selection condition for C is chosen based on the maximum value of the Figure of

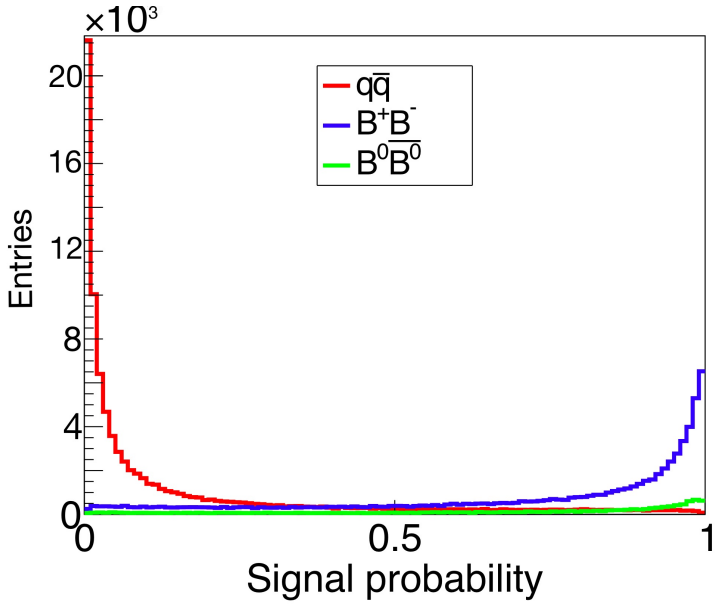


Figure 4.13: Classifier output (C)

Merit (FoM), which is $S/\sqrt{(S+B)}$, and is required to be greater than 0.36, shown in Fig. 4.14. This selection condition leads to a rejection of 57% of the background with 91% signal efficiency.

4.7 Multiplicity, Single-candidate selection and efficiency

In some events, it is possible to have more than one candidate. Misreconstructed tracks can result in a false candidate alongside the true ones. The multiplicity distributions for B candidates are shown in Fig. 4.15a after applying all optimized selection criteria, with an average multiplicity of 1.139.

To select the candidate closest to the generated event produced, the single-candidate selection (SCS) criteria are applied. SCS is performed after the event passes through the detector and reconstruction software steps. If the generator level information matches the reconstructed information, the generated Monte-Carlo event is translated into a truth-matched (TM) candidate. SCS is done to obtain the candidate that is closest to the TM candidate.

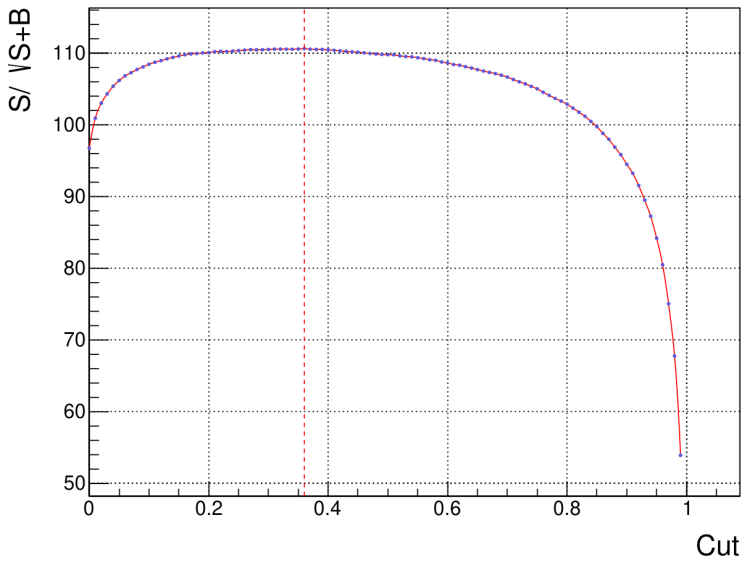
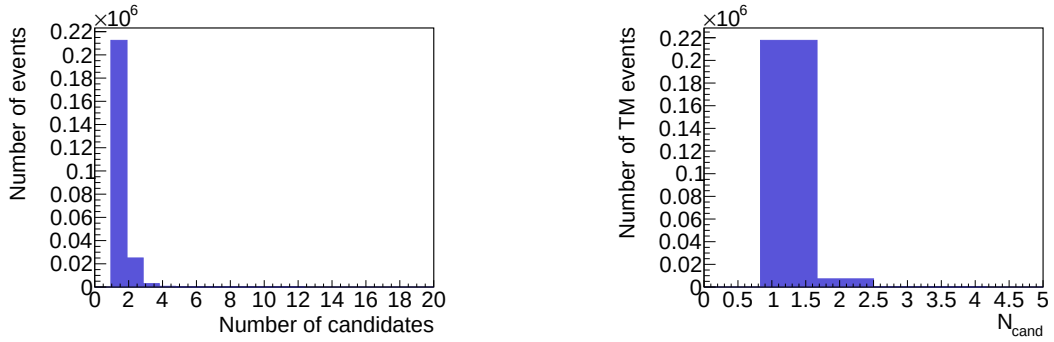


Figure 4.14: FoM analysis of FBDT output

The χ^2 metric of M_{bc} and M_{D^0}, defined as $\left(\frac{M_{bc}-M_B^{\text{PDF}}}{\sigma_{M_{bc}}}\right)^2 + \left(\frac{M_{D^0}-M_{D^0}^{\text{PDF}}}{\sigma_{M_{D^0}}}\right)^2$ where $\sigma_{M_{bc}}$ and $\sigma_{M_{D^0}}$ are the resolutions of M_{bc} and M_{D^0}, is used for the SCS. The resolutions of M_{bc} and M_{D^0} are obtained from fits in the signal MC samples. For M_{bc}, a single Gaussian PDF is used, while for M_{D^0}, a sum of two Gaussians is used. The resulting effective resolutions are 2.6 MeV/c^2 and 3.2 MeV/c^2 for M_{bc} and M_{D^0}, respectively.

The figure in 4.15b illustrates the distribution of TM events for various SCS ranked candidates(N$_{cand}$). The results indicate that 96.60% of TM events are attributed to the 1st ranked candidate, 3.20% of TM events to the 2nd ranked candidate, 0.16% of TM events to the 3rd ranked candidate, and 0.03% of TM events to the 4th ranked candidate. Therefore, it is evident that this metric is doing good job for SCS. The signal efficiencies obtained after applying all the selection criteria and single-candidate selection is 21.77%.



(a) (b)

Figure 4.15: a) Multiplicity distribution and b) distribution of TM events for various SCS ranked candidates.

CHAPTER 5

Yield Measurement

Once the background has been reduced through various analyses, we proceed to fit ΔE with a suitable probability density function to accurately calculate the mean and yield of both the signal and background from the total dataset. The quality of the fit is determined by certain parameters, such as the χ^2/ndf (where ndf represents the number of degrees of freedom) and the 'pull' distributions. The χ^2/ndf value should be close to one and the pull distribution should remain within ±3 across the entire data range if the selected probability density function (pdf) fits the data well.

5.1 ΔE fit

The parameters are derived through a one-dimensional fitting process of ΔE to the corresponding Monte Carlo (MC) samples. The signal component of ΔE is represented by a double Gaussian distribution, which has been modeled based on a fitting performed on large signal MC samples shown in Fig. 5.1. To account for the resolution difference between data and MC, a multiplicative factor $f_{\Delta E}$, commonly known as a "fudge factor", is introduced on the principal σ of the ΔE signal component. This factor indicates whether the resolution in data is better or worse than that in the MC. A Chebyshev polynomial ($P_n(x)$) of first order is found to be suitable for modeling the combinatorial and the continuum backgrounds. Table 5.1 provides a summary of all the obtained parameters, while Fig. 5.2 shows the overall fit result.

After performing the fitting process, the yield of signal events was determined to be **21632 ± 207**, which includes the contribution from strong intermediate resonance states such as a_1^- and ρ^0. To extract the yield of these contributions, a four-body amplitude analysis of $\bar{D}^0$ and the three pions is required. In Figure 5.3, the predicted distribution of

Table 5.1: Fit parameter information of ΔE distribution of MC sample. The parameters $\mu_{1,2}$ and $\sigma_{1,2}$ represent the mean and width of first and second Gaussians, respectively. The parameter f represents the weight of first Gaussian and $f_{\Delta E}$ represent the fudge factor. Parameter b_0 represents the free parameter of Chebyshev polynomial.

Component	Parameter	Value
Signal	μ_1	$(-0.175 \pm 0.054) \times 10^{-3}$ GeV
	μ_2	$(-12.61 \pm 0.68) \times 10^{-3}$ GeV
	σ_1	$(6.486 \pm 0.051) \times 10^{-3}$ GeV
	σ_2	$(35.79 \pm 0.60) \times 10^{-3}$ GeV
	f	0.836 ± 0.004
Background	b_0	-0.259 ± 0.010
Scale Factor	$f_{\Delta E}$	0.976 ± 0.009

the invariant mass of three charged π and two π of signal events, as determined by the ${}_s\mathcal{P}lot$ technique, is compared to the original distribution. This will enable us to measure the partial branching fraction as a function of the invariant mass of three pions or two pions, which will allow us to improve the simulation and perform a four-body amplitude analysis.

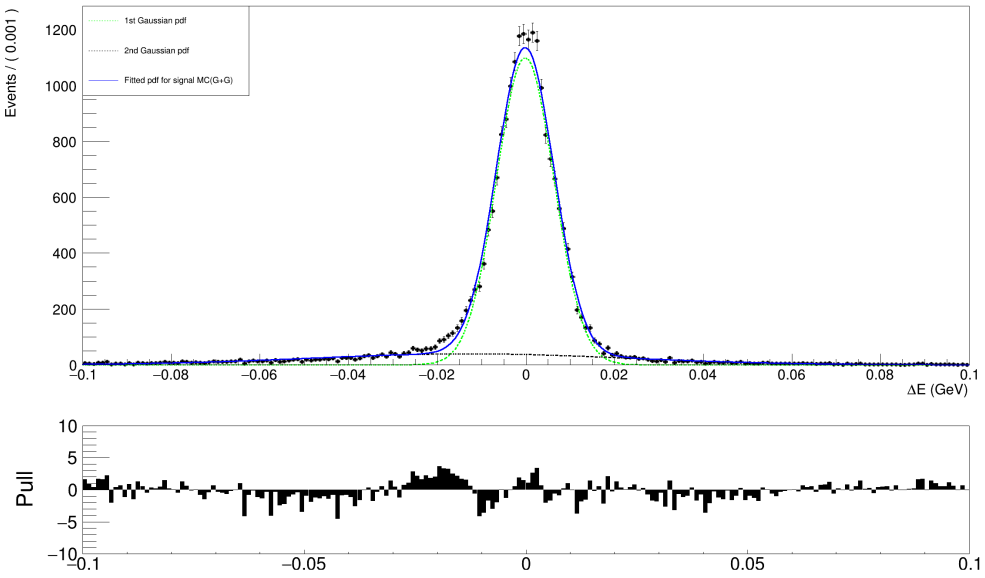


Figure 5.1: ΔE fit of Signal events.

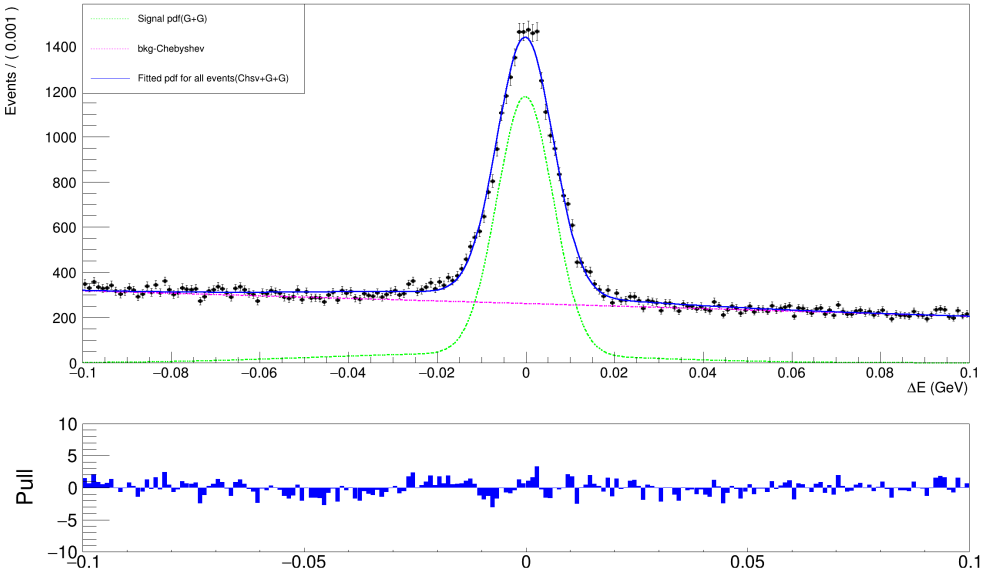
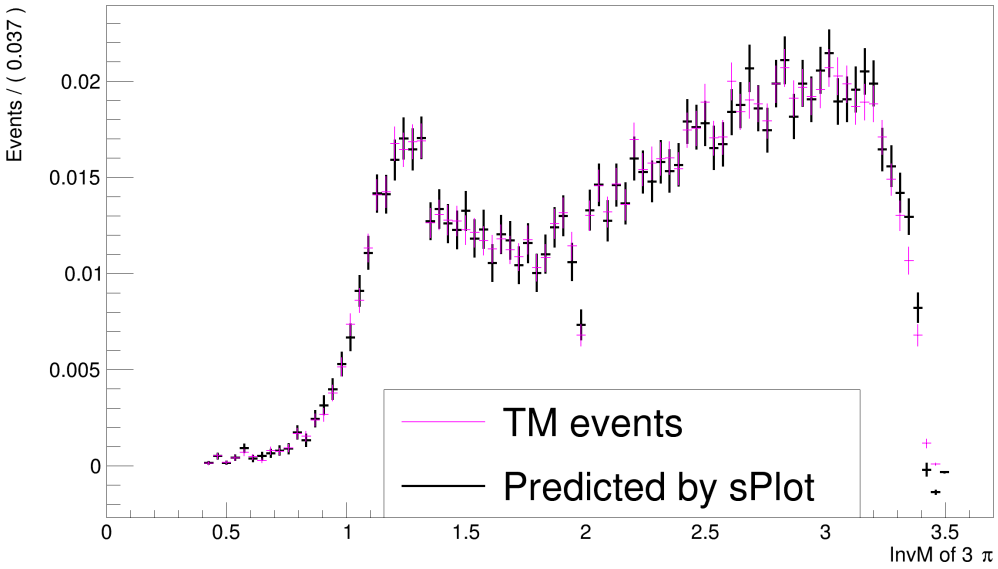
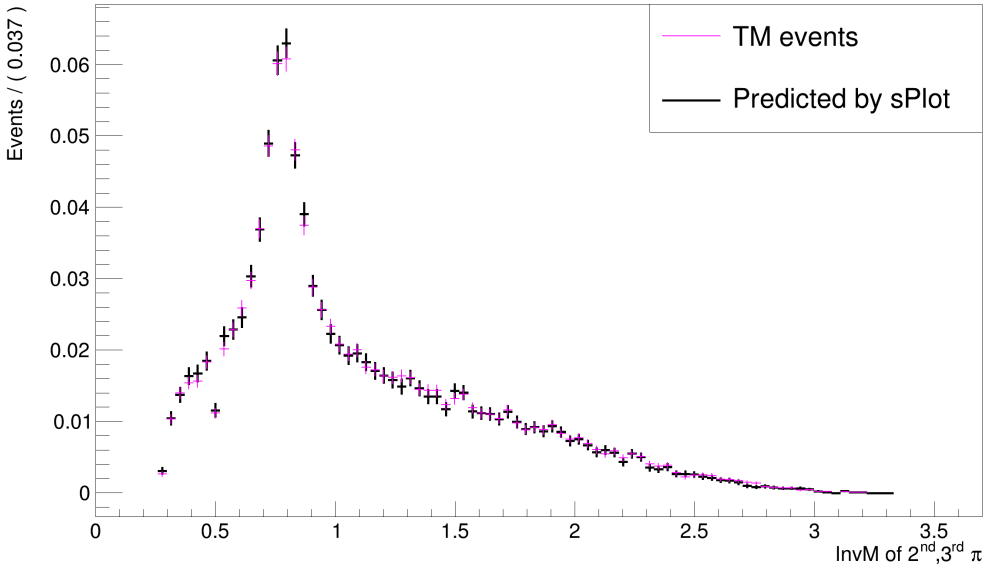


Figure 5.2: ΔE fit of all events.



(a)



(b)

Figure 5.3: Comparison of predicted invariant mass distribution of a) three charged π and b) two π of signal events with original distribution.

CHAPTER 6

Result and Conclusions

6.1 Result

The obtained signal yield from the combined signal and background fit is come out to be

$$N_{sig} = 21632 \pm 207.$$

The signal selection efficiency after all analyses is

$$\varepsilon = 21.77\%.$$

We will calculate the branching fraction using following relation

$$\mathcal{B}\left(\mathrm{B}^{+} \to \bar{\mathrm{D}}^{0}\pi^{+}\pi^{-}\pi^{+}\right) = \frac{N_{sig}}{2 \cdot \varepsilon \cdot \mathscr{L} \cdot \sigma_{B^{+}B^{-}} \cdot \mathcal{B}\left(\bar{\mathrm{D}}^{0} \to K^{+}\pi^{-}\right)} \tag{6.1}$$

where $\mathscr{L}$ refers to the integrated luminosity of the experiment (here we are using 200 fb^{-1} of simulated data) and σ refers to the cross section of the production of $B^{+}B^{-}$ pairs in the experiment (0.5654 nb in the Belle-II experiment). $\mathcal{B}\left(\bar{\mathrm{D}}^{0} \to K^{+}\pi^{-}\right)$ is the branching ratio for the decay $\bar{\mathrm{D}}^{0} \to K^{+}\pi^{-}$ as given in PDG which is (3.947 ± 0.030)% [72]. After substituting all the values in the equation 6.1, we get:

$$\mathcal{B}\left(\mathrm{B}^{+} \to \bar{\mathrm{D}}^{0}\pi^{+}\pi^{-}\pi^{+}\right) = 0.0111.$$

The statistical error is calculated as 207/21632 = 1.0%. The error in the absolute branching ratio of $\mathcal{B}\left(\bar{\mathrm{D}}^{0} \to K^{+}\pi^{-}\right)$ is determined to be 0.030/3.947 = 0.8%. The systematic uncertainty in counting B meson is approximately 1.3%, and the uncertainty in PID is 0.01% [35]. After adding all these uncertainties in quadrature, the total uncertainty is 1.83%. Therefore, the measured value of the branching ratio for the decay channel $\mathrm{B}^{+} \to \bar{\mathrm{D}}^{0}\pi^{+}\pi^{-}\pi^{+}$ is $(11.1 \pm 0.1) \times 10^{-3}$. This value includes the contribution from strong intermediate resonance states such as a_1^{-} and ρ^{0}.

6.2 Discussion of the result

The result obtained is consistent with the figure specified in the decay file used for the production of generic Monte Carlo (MC) data. By utilizing the ${}_s\mathcal{P}lot$ technique, it will be possible to measure the differential branching fraction, which can be utilized to enhance the simulation or perform a four-body amplitude analysis.

6.3 Conclusion

In conclusion, this study represents the first measurement of the branching fraction for the rare decay channel $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ by Belle II. With the vast dataset available at Belle II, the 40% uncertainty in the absolute measurement of the branching ratio by CLEO can be significantly reduced. Prior to the analysis of experimental data, simulated data was utilized to identify possible background sources, and a topological analysis was performed using generator level information to identify major sources of peaking background. A Machine Learning approach was employed to separate the signal from the dominant continuum background based on their topological differences, and the resulting yield predicted the number of signal events in the simulated data with approximately 1.83% uncertainty. Overall, this study provides a solid foundation for future experimental measurements of this decay channel at Belle II.

References

[1] S. Weinberg, *Phys. Rev. Lett.*, 1967, **19**, 1264–1266.

[2] Wikimedia Commons, The Standard Model of elementary particles, 2019.

[3] G. Lüders, *Annals of Physics*, 1957, **2**, 1–15.

[4] R. P. Feynman and M. Gell-Mann, *Phys. Rev.*, 1958, **109**, 193–198.

[5] N. Cabibbo, *Phys. Rev. Lett.*, 1963, **10**, 531–533.

[6] S. L. Glashow, J. Iliopoulos and L. Maiani, *Phys. Rev. D*, 1970, **2**, 1285–1292.

[7] B. Bjørken and S. Glashow, *Physics Letters*, 1964, **11**, 255–257.

[8] Aubert et al., *Phys. Rev. Lett.*, 1974, **33**, 1404–1406.

[9] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, *Phys. Rev. Lett.*, 1964, **13**, 138–140.

[10] M. Kobayashi and T. Maskawa, *Progress of Theoretical Physics*, 1973, **49**, 652–657.

[11] Herb et al., *Phys. Rev. Lett.*, 1977, **39**, 252–255.

[12] Abe et al., *Phys. Rev. Lett.*, 1995, **74**, 2626–2631.

[13] Abachi et al., *Phys. Rev. Lett.*, 1995, **74**, 2632–2637.

[14] H. Albrecht et al., *Physics Letters B*, 1987, **192**, 245–252.

[15] Abe et al., *Phys. Rev. Lett.*, 2001, **87**, 091802.

[16] Aubert et al., *Phys. Rev. Lett.*, 2001, **87**, 091801.

[17] Aaij et al., *Phys. Rev. Lett.*, 2020, **125**, 011802.

[18] Aaij et al., *Phys. Rev. Lett.*, 2015, **115**, 111803.

[19] Abi et al., *Phys. Rev. Lett.*, 2021, **126**, 141801.

[20] D. Bortoletto *et al.*, *Phys. Rev. D*, 1992, **45**, 21–35.

[21] Steven R. Blusk et al., *Study of* $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ *and* $B^+ \to \bar{D}^0K^+\pi^-\pi^+$ *decays and determination of the CKM angle* γ, Dec 2015.

[22] https://confluence.desy.de/display/BI/Journal+Publications.

[23] K. F. K. Akai and H. Koiso., *SuperKEKB collider, Nucl. Instrum. Methods Phys. Res. A, 907:188-199*, Nov. 2018.

[24] SuperB Collaboraion, *SuperB: A High-Luminosity Asymmetric e^+e^- Super Flavor Factory. Conceptual Design Report*, Mar, 2007.

[25] https://cds.cern.ch/record/2713605/plots.

[26] T. Abe et al., *Belle II Technical Design Report*, 2010.

[27] I. Adachi, T. Browder, P. Križan, S. Tanaka and Y. Ushiroda, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2018, **907**, 46–59.

[28] H. Ye et al. Commissioning and performance of the Belle II pixel detector, *Nucl. Instrum. Methods Phys. Res. A, 987:164875*, Jan. 2021.

[29] J. Kemmer and G. Lutz. New detector concepts, *Nucl. Instrum. Methods Phys. Res. A, 253:365-377*, Jan. 1987.

[30] https://indico.belle2.org/event/1307/sessions/378/attachments/3070/5651/B2SKW_PXD.pdf.

[31] K. Adamczyk et al. (Belle II SVD), *The Design, Construction, Operation and Performance of the Belle II Silicon Vertex Detector*, Jan. 2022.

[32] G. Rizzo et al. The Belle II Silicon Vertex Detector: Performance and Operational Experience in the First Year of Data Taking, *JPS Conf. Proc., 34:010003*, Dec. 2021.

[33] https://indico.cern.ch/event/775317/contributions/3227530/attachments/1773865/2883570/talk_SVD_Manish.pdf.

[34] N. Taniguchi (Belle II Collaboration). Central Drift Chamber for Belle-II, *J. Instrum., 2:C06014*, Jun. 2017.

[35] E. Kou et al. The Belle II Physics Book, *Prog. Theor. Exp. Phys., 2019*, Dec. 2019.

[36] V. Bertacchi, T. Bilka et al., *Computer Physics Communications*, 2021, **259**, 107610.

[37] J. Fast (Belle II TOP Group). The Belle II imaging Time-of-Propagation (iTOP) detector, *Nucl. Instrum. Meth. A, 876:145-148*, Dec. 2017.

[38] S. Nishida et al. Aerogel RICH for the Belle II forward PID, *Nucl. Instrum. Meth. A, 766:28-31*, Dec 2014.

[39] K. Inami. MCP-PMT development for Belle-II TOP counter, *Phys. Procedia, 37:683-690*, Oct. 2012.

[40] K. Kojima (Belle II TOP Group). The operation and performance of the TOP detector at the Belle II experiment, *Proc. Sci., EPS-HEP202:803*, May 2022.

[41] https://www2.kek.jp/proffice/archives/feature/2010/BelleIIBPID.html.

[42] S. Korpar (Belle II ARICH Group). A 144-channel HAPD for the Aerogel RICH at Belle II, *Nucl.Instrum. Meth. A, 766:145-147*, 2014.

[43] M Yonenaga et al. Performance evaluation of the aerogel RICH counter for the Belle II spectrometer using early beam collision data, *Progress of Theoretical and Experimental Physics*, Aug. 2020. 093H01.

[44] S. Iwata et al., *Progress of Theoretical and Experimental Physics*, 2016 033H01.

[45] V. Aulchenko et al., JINST 12 C08001 , 2017.

[46] V Aulchenko et al., *Journal of Physics: Conference Series*, 2015, **587**, 012045.

[47] https://www.phys.hawaii.edu/~teb/BelleII_NIM_special.pdf.

[48] T. Aushev et al., *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2015, **789**, 134–142.

[49] S. Yamada, R. Itoh, K. Nakamura, M. Nakao, S. Y. Suzuki, T. Konno, T. Higuchi, Z. Liu and J. Zhao, *IEEE Transactions on Nuclear Science*, 2015, **62**, 1175–1180.

[50] https://software.belle2.org/development/sphinx/online_book/fundamentals/02-datataking.html#triggers-and-filters.

[51] Stefan Ask et al., *Computer Physics Communications*, 2015, **191**, 159–177.

[52] S. Jadach, B. Ward and Z. Was, *Computer Physics Communications*, 2000, **130**, 260–325.

[53] C. Carloni Calame, G. Montagna, O. Nicrosini and F. Piccinini, *Nuclear Physics B - Proceedings Supplements*, 2004, **131**, 48–55.

[54] F. Berends, P. Daverveldt and R. Kleiss, *Computer Physics Communications*, 1986, **40**, 271–284.

[55] D. J. Lange, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2001, **462**,

152–155.

[56] S. Jadach, J. H. Kuhn and Z. Was, *Computer Physics Communications*, 1991, **64**, 275–299.

[57] E. Barbiero and Z. Was, *Comp. Phys. Commun.*, 291 (1994).

[58] S. Agostinelli et al., *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2003, **506**, 250–303.

[59] T. Alexopoulos, M. Bachtis, E. Gazis and G. Tsipolitis, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2008, **592**, 456–462.

[60] R. Fruhwirth et al., *Nucl. Instrum. Meth.*, A732, 95-98 (2013).

[61] T. Bilka et al., *Implementation of GENFIT2 as an experiment independent track-fitting framework*, Feb. 2019.

[62] N. Braun, *Combinatorial Kalman Filter*, 117-174, Springer, Cham, Aug. 2019.

[63] https://indico.cern.ch/event/773049/contributions/3474751/attachments/1940572/3217515/Belle2_TrackPerf_CHEP_7Nov2019.pdf.

[64] T. Kuhr et al., The Belle II Core Software., *Comput. Software Big Sci.*, 3:1, Nov. 2019.

[65] *Standard C++ Foundation*, https://isocpp.org/std/the-standard.

[66] G. Van Rossum and F. L. Drake Jr., *Python reference manual*, 1995.

[67] J. H. Friedman, *The Annals of Statistics*, 2001, **29**, 1189 – 1232.

[68] G. Bohm and G. Zech, *Einführung in Statistik und Messwertanalyse für Physiker*, DESY, 2006.

[69] X. Zhou, S. Du, G. Li and C. Shen, *Computer Physics Communications*, 2021, **258**, 107540.

[70] https://docs.belle2.org/record/2964/files/BELLE2-PTHESIS-2022-002.pdf.

[71] G. C. Fox and S. Wolfram, *Phys. Rev. Lett.*, 1978, **41**, 1581–1585.

[72] https://pdglive.lbl.gov/BranchingRatio.action?desig=1&parCode=

S032&home=MXXX035.

[73] S. L. Glashow, *Nuclear Physics*, 1961, **22**, 579–588.

[74] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C 680519 (1968) 367.

[75] F. Englert and R. Brout, *Phys. Rev. Lett.*, 1964, **13**, 321–323.

[76] P. W. Higgs, *Phys. Rev. Lett.*, 1964, **13**, 508–509.

[77] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Phys. Rev. Lett.*, 1964, **13**, 585–587.

[78] Steven R. Blusk et al., *Phys. Rev. Lett.*, 2012, **108**, 161801.

[79] T. Keck, *FastBDT: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification*, 2016.