Master's thesis

A search for $b \to u\ D_s$ decay at Belle experiment

(Belle実検における b → u D_s 崩壊の探索)

Graduate School of Tohoku University
Department of Physics

Yoshinari Mikami

February 2002

Acknowledgment

I'm grateful to Prof. A.Yamaguchi for leading me to high energy physics and his hearty aid of my study in long time. I would like to express special thanks to Prof. H.Yamamoto, I owe much of this analysis to his intelligence and kindness. I could not write this thesis without his advice. I'd like to express my gratitude to Dr T.Nagamine, his practical knowledge solved my problems many times. Staffs in Research Center of Neutrino Science, staffs in KEK and BELLE collaborators helped me so much, especially I appreciate their kindness of Mrs K.Tamae, Prof. Koya Abe, Prof. Kazuo Abe, Prof. M.Yamauchi, Prof. Y.J.Kwon, Dr T.Iijima, Dr Y.Iwasaki and Dr M.Yamaga. Students who I could get acquainted with gave me much interest and make me rise. I'm thankful to be surrounded by superior scientists and familiar people. Lastly, I want to thank my family for continuous support of my life.

Yoshinari Mikami
February 18, 2002.

Contents

List of Figures

List of Tables

abstract

In order to measure the Kobayashi-Maskawa matrix element $|V_{ub}|$, we analyzed the $b \to u\ D_s$ decay. We used 29.1 fb^{-1} on-$\Upsilon(4S)$-resonance data and $3.0 fb^{-1}$ off-resonance data which were accumulated with BELLE detector. We mesured branching fraction of inclusive $B \to D_s$ decay as $(11.0 \pm 0.3 \pm 2.9)\%$ (first error is statistics and second error is systematic). And, we measured the $|\frac{Vub}{Vcb}|^2$ as 0.0059 ± 0.0054 (statistical error only), then the upper limit only for $|V_{ub}|$ is calculated as $|V_{ub}| \leq 4.6 \times 10^{-3}$ (90% CL). We could approch the $|V_{ub}|$ with new method of $b \to u\ D_s$ decay.

Chapter 1

Introduction

When it was beginning of our universe, particles and anti-particles would exist with equal ratio. The problem is that we can see almost only particles in the present universe. The reason why such a unbalance exist can be considered as a CP violation. CP violation was found with weak interaction decay of K meson in 1964. In 1973, Kobayashi-Maskawa proved that if quark have three generations, it would contain CP violation as a natural result [1]. In 1980s, Sanda and Carter showed the decay of B meson would have large CP asymmetries [2]. Therefore, in order to get certain proof that CP violation is due to Kobayashi-Maskawa mechanism, experiment of B factory was started in 1995. If the probability amplitude of weak interaction have irreducible complex phase, quarks and anti-quarks could behave differently. BELLE experiment proved it with B meson decay.

KM matrix has the probability amplitude of weak interaction as each matrix element.

$$V_{KM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.1}$$

The matrix has four degrees of freedom for three generations of quarks. The four degrees consist of three degrees of rotation and one degree of complex phase, as shown below:

In general, n $\times$ n matrix have $2n^2$ degrees of freedom. Unitarity requires to reduce n^2 degrees of freedom, relative phase of 2n quarkes reduce (2n-1) degrees of freedom, and rotation need n(n-1)/2 degrees of freedom. So, the remained are

$$2n^2 \quad - \quad n^2 \quad - \quad (2n-1) \quad - \quad n(n-1)/2 \quad = (n-1)(n-2)/2. \tag{1.2}$$

((real+complex) $\times n^2$) - (Unitarity) - (relative phase) - (rotation) = (complex phase).

This (n-1)(n-2)/2 are degrees of freedom for complex phase. Therefore, if quarks have greater than 3 generations, they have complex phases. In particular, 3 generations of quarks have n(n-1)/2 = 3 rotation parameters, and (n-1)(n-2)/2 = 1 complex phase parameter.

The KM matrix can be written with 4 parameters θ_1, θ_2, θ_3, δ.

$$V_{KM} = \begin{pmatrix} cos\theta_1 & sin\theta_1 & 0 \\ -sin\theta_1 & cos\theta_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\theta_2 & sin\theta_2 \\ 0 & -sin\theta_2 & cos\theta_2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\delta} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\theta_3 & sin\theta_3 \\ 0 & -sin\theta_3 & cos\theta_3 \end{pmatrix} \tag{1.3}$$

$$= \begin{pmatrix} cos\theta_1 & sin\theta_1 cos\theta_3 & sin\theta_1 sin\theta_3 \\ -sin\theta_1 cos\theta_2 & cos\theta_1 cos\theta_2 cos\theta_3 - sin\theta_2 sin\theta_3 e^{i\delta} & cos\theta_1 cos\theta_2 sin\theta_3 + sin\theta_2 cos\theta_3 e^{i\delta} \\ sin\theta_1 sin\theta_2 & -cos\theta_1 sin\theta_2 cos\theta_3 - cos\theta_2 sin\theta_3 e^{i\delta} & -cos\theta_1 sin\theta_2 sin\theta_3 + cos\theta_2 cos\theta_3 e^{i\delta} \end{pmatrix}. \tag{1.4}$$

in order to compare each matrix elements easily, replace the 4 parameters θ_1, θ_2, θ_3, δ with λ, A, ρ, η. Then, λ and A are defined by $\lambda \equiv |V_{us}| \sim 0.22$ and by $A\lambda^2 \equiv |V_{cb}| \sim 0.04$ (A is of order unity). If only second order of λ is taken into account, V_{KM} is

$$V_{KM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & 0 \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ 0 & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^3). \tag{1.5}$$

If $|V_{ub}|$ and $|V_{td}|$ are replaced with arbitrary third order of λ, they can be written as $|V_{ub}| = A\lambda^3(\rho - i\eta)$ and as $|V_{td}| = A\lambda^3(\alpha - i\beta)$, then

$$V_{KM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(\alpha - i\beta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4), \tag{1.6}$$

where, unitrarity ($V_{KM}V_{KM}^{\dagger} = I$) requires $\alpha = 1 - \rho$, and $\beta = \eta$. Eventually, the KM matrix to the third order in λ becomes

$$V_{KM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4). \tag{1.7}$$

(Wolfenstein expression [5])

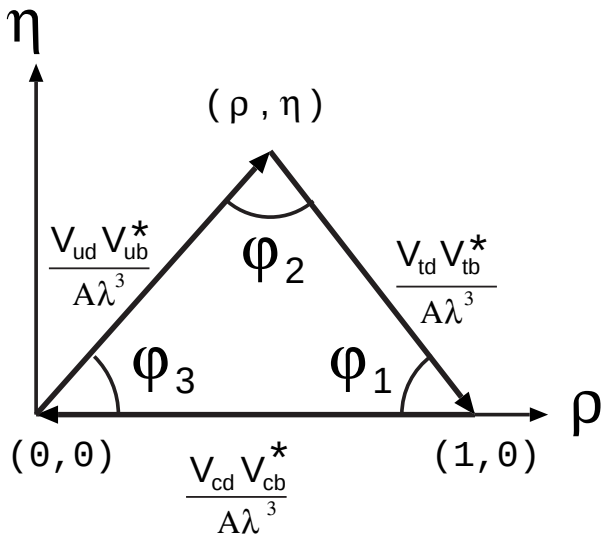
In order to prove CP violation, whether KM matrix have irreducible complex phase or not is investigated. Then it was used that V_{KM} must satisfy unitarity,

$$V_{KM}^{\dagger} V_{KM} = I, \tag{1.8}$$

the orthogonality of the d-column and the b-column leads

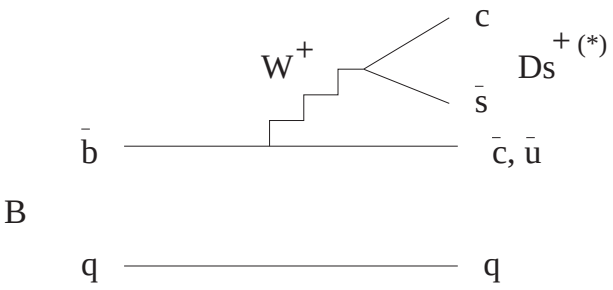
$$V_{ud} V_{ub}^{*} + V_{cd} V_{cb}^{*} + V_{td} V_{tb}^{*} = 0. \tag{1.9}$$

The three component of above equation can be written in complex plane, and they form a triangle. If it is drawn in (ρ, η) plane, the unitarity triangle becomes as shown below. Then, three angle ϕ_1, ϕ_2, and ϕ_3 are defined.



If the angle ϕ_1 or ϕ_3 is not zero, KM matrix should have complex phase. Already, BaBar collaboration and Belle collaboration showed that $sin2\phi_1$ is not zero [3] [4]. It means CP violation, but whether Kobayashi-Maskawa theory is right or not depends on whether unitarity triangle is closed or not. Parameter λ and A are well known, and $|V_{ud}|$ can be written with λ, therefore the measurement of $|V_{ub}|$ gives one side of the triangle which is opposite to ϕ_1. That is why, measurement of $|V_{ub}|$ is important because this measurement make it possible to confirm KM theory is correct or not.

One way of measuring $|V_{ub}|$ is to use the inclusive D_s productions $b \to u\ D_s$ which involves a $b \to u$ transition. Then, $b \to c\ D_s$ decay becomes serious background. Fynmman diagram which contribute this analysis is like below



Maximum momentum of D_s that come from B meson primary decay is different between $b \to u\ D_s$ decay and $b \to c\ D_s$ decay. Therefore, D_s which come from $b \to u\ D_s$ decay can be extracted in momentum spectrum as the D_s above the end point of $b \to c\ D_s$ decay. The mass difference between u-quark and c-quark appears as D_s momentum difference. The branching fraction for $B \to X_c\ D_s$ decay is 10% in PDG [6], and the estimated branching fraction for $B \to X_u\ D_s$ decay is 0.15%.

So far, $|V_{ub}|$ had been measured by the end point of lepton momentum spectrum in semi-leptonic decay. The recent result from CLEO collaboration (2000) [7] is,

$$|V_{ub}| = \{3.25 \pm 0.14\ (stat)\ {}^{+0.21}_{-0.29}\ (sys) \pm 0.55\ (theory)\} \times 10^{-3}.$$

The merit of using $b \to u\ D_s$ decay is that this decay can be regarded as a two-body decay in quark level, and don't contain neutrino in the final state. Therefore, essentially, momentum of the D_s is monochromatic in quark level, and about 50% of the spectrum of $b \to u\ D_s$ is above the endpoint of $b \to c\ D_s$ [8]. It means the signal region can be calculated more clearly and signal can be expected with less background than $b \to u\ l\nu$ mode.

Chapter 2

BELLE experiment

For the purpose of studying CP violation, BELLE experiment was started.

In order to prove the CP asymmetry in B meson system, $B\overline{B}$ pair is produced with a e^+e^- collider. Then, e^+ and e^- are collided on the $\Upsilon(4S)$ resonanse which is 10.58 GeV in the center of mass frame, and more than 96% of $\Upsilon(4S)$ decay into $B\overline{B}$ pair. BELLE detactor accumulated 29.1 fb^{-1} on-$\Upsilon(4S)$-resonance data and 3.0 fb^{-1} off-resonance data by August, 2001. The high statistics of $B\overline{B}$ data enables to measure both large CP asymmetry and precision test of Kobayashi-Maskawa theory.

2.1 KEKB accelerator

KEKB is a storage ring of 8.0 GeV electron and 3.5 GeV positron which are collided at one point. Each ring is designed to have 5000 bunches which are stored in 3 km rings, and they collide every 2ns with the angle of $\pm$11mrad. High energy ring (HER) have 1.4×10^{10} electron per bunch and low energy ring (LER) have 3.3×10^{10} positron with design beam current 1.1A and 2.6A for HER and LER. They have asymmetry energy because $B\overline{B}$ are required to be boosted in laboratory frame in order to get large difference of decay length between B and $\overline{B}$ for studying CP asymmetry. The boost factor can be calculated as

$$\beta\gamma = (E_{e^-} - E_{e^+})/\sqrt{s} = (\ 8.0 - 3.5\ )/10.58 = 0.425$$

The general view of KEKB accelerator is shown on the next page.

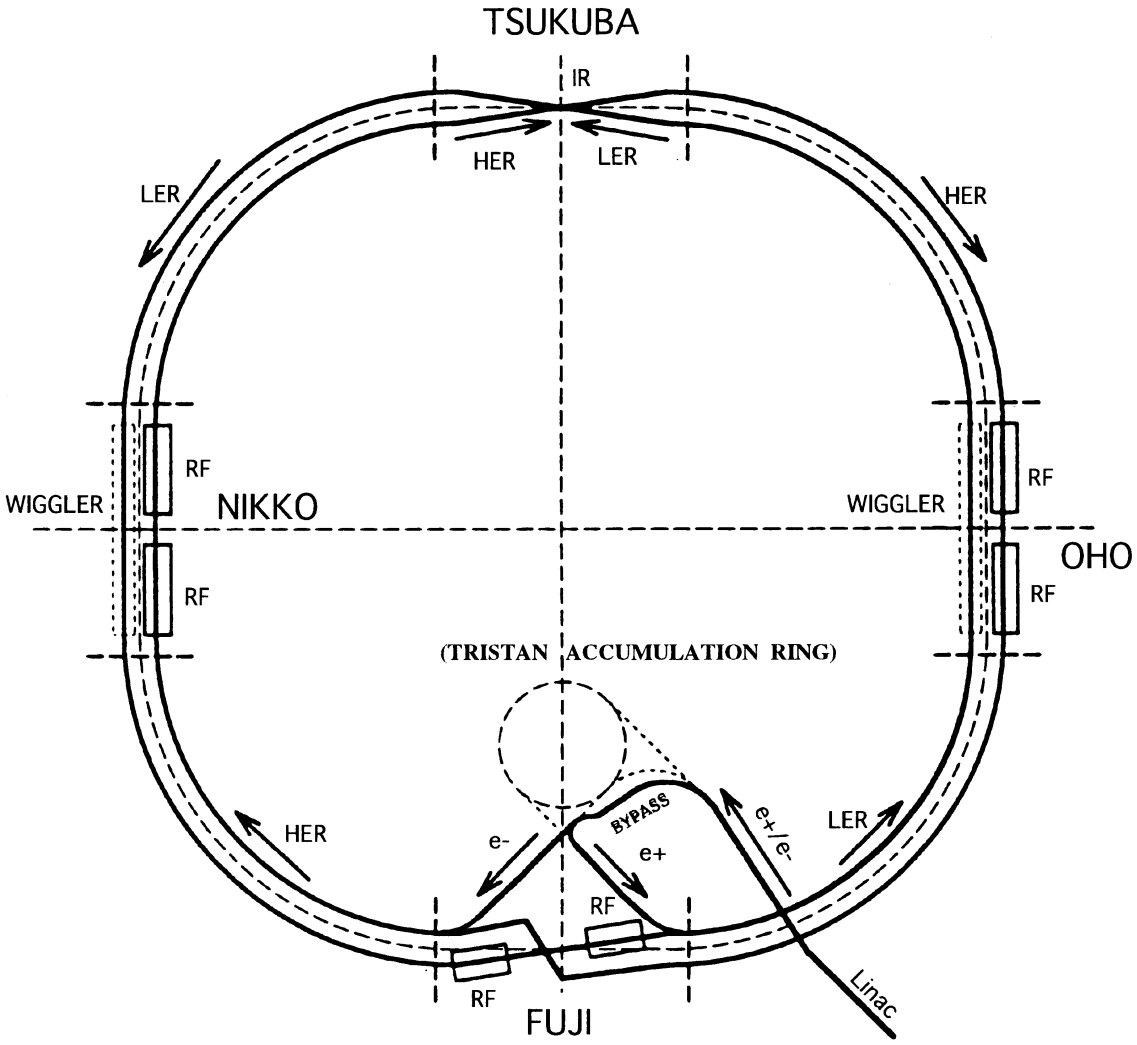


Figure 2.1: KEKB accelerator

2.2 BELLE detector

The aim of BELLE detector are measuring momentum or energy of particles, particle identification, for example, charged π, K, and lepton, precise measurement of B meson vertex, large acceptance and high efficiency, and less background and smooth data taking. Belle detector consists of beam-pipe, sub-detectors, solenoid magnet, among others. In order to detect various particles, each sub-detector was designed and have been developed continuously. Every sub-detector except for K_L/μ detector is located inside the superconducting solenoid with 1.5T magnet field. Features and performances of these devices are described in following subsections.

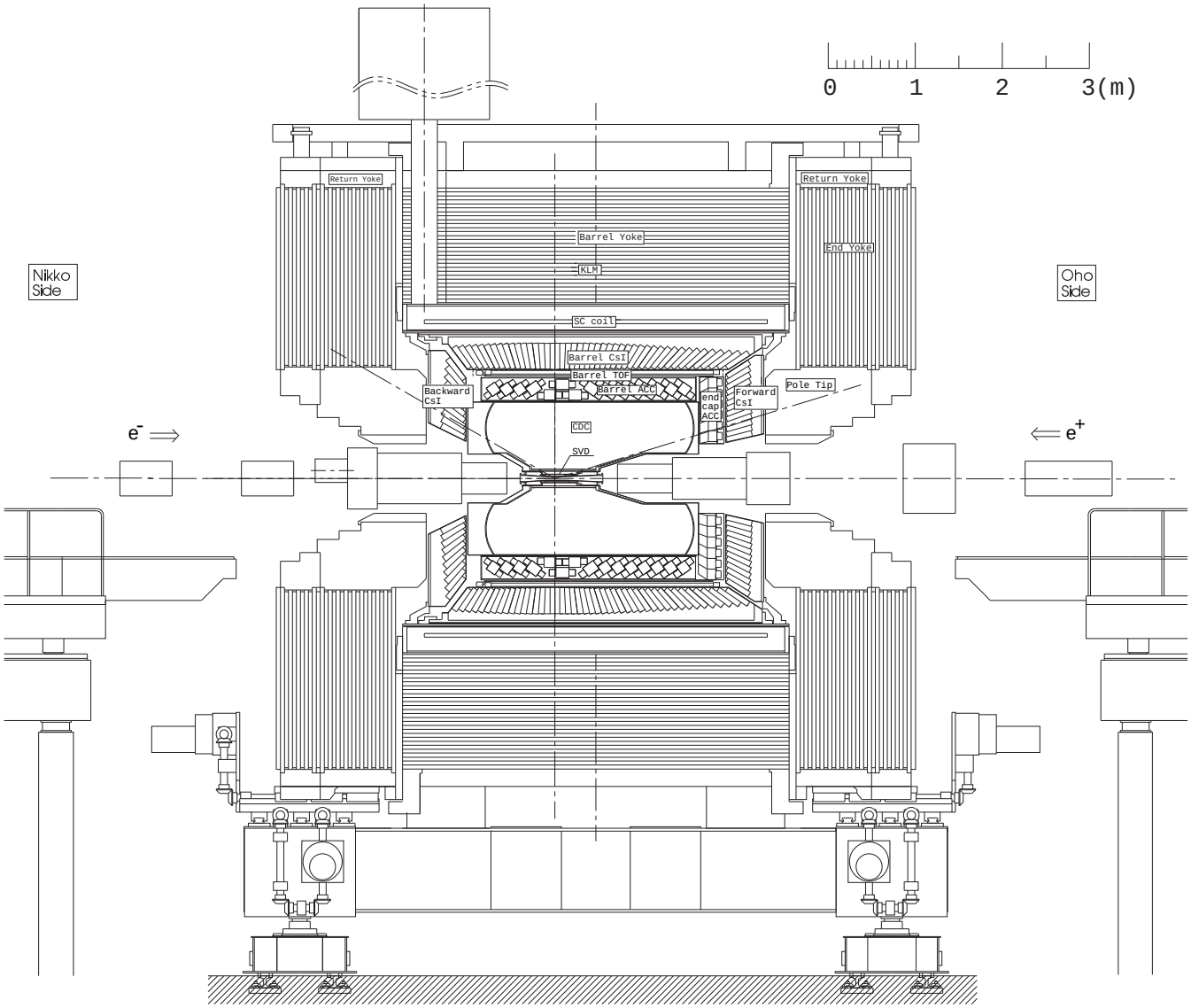


Figure 2.2: Side view of BELLE detector

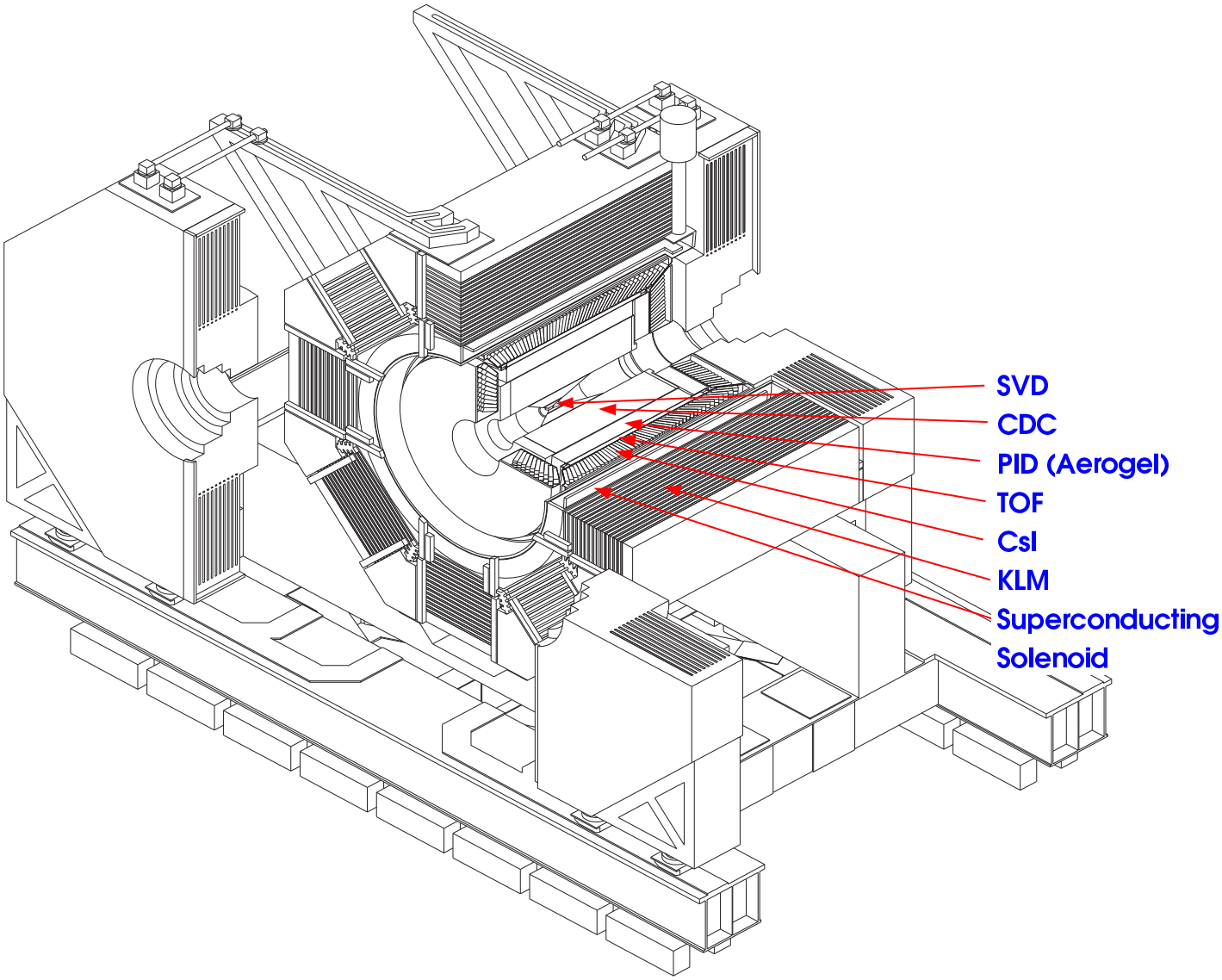


Figure 2.3: BELLE detector

2.2.1 Silicon Vertex Detector (SVD)

Many studies of CP violation that involves $B\overline{B}$ mixing requires measuring the flight-length difference of $B\overline{B}$ pair. In this aim, Silicon Vertex Detector(SVD) is located just around the beam pipe with 3.3cm to 5.8cm in radius and 22cm to 34cm in length, and it has three layers of $300\mu m$ thick Double Sided Silicon Detector (DSSD). The required resolution of vertex point is about $100\mu m$ on the beam axis. The resolution of impact parameter in data which depends on momentum and poler angle are as follows; $\sigma_{xy} = 19 \oplus 50/(p\beta sin^{3/2}\theta)\mu m$ and $\sigma_z = 36 \oplus 42/(p\beta sin^{5/2}\theta)\mu m$. The total number of readout channels are about 82,000 channels.

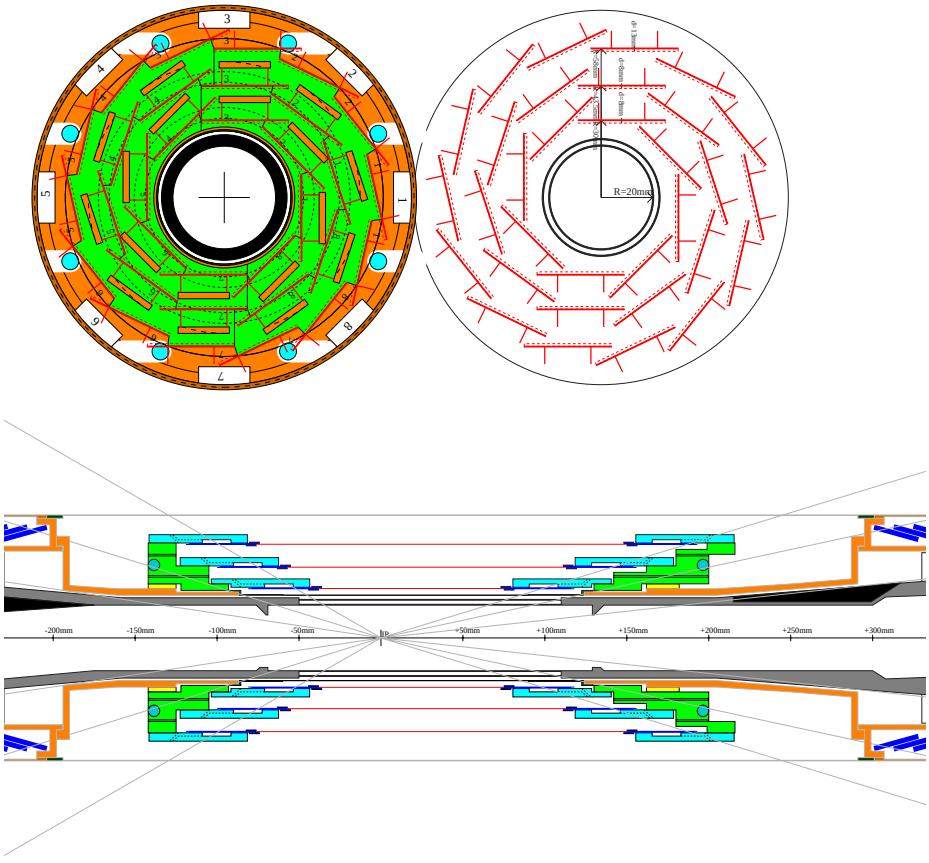


Figure 2.4: Silicon Vertex Detector

2.2.2 Central Drift Chamber (CDC)

Drift chamber is located around the SVD in order to measure track, momentum and dE/dx of charged particles. Central Drift Chamber (CDC) is filled with low material gas mixture (50% He and 50% C_2H_6) which can ionize charged particles. This CDC consists of small cell chamber that have 50 anode wires and 3 cathode strip layers, then, the 50 anode wires are divided into 32 axial wires and 18 stereo wires to reconstruct 3-dimensional track. Momentum of charged particles can be measured from curvature radius of track in the 1.5T magnet field. Measurement of specific ionization (dE/dx) makes it possible to identify particles. This chamber covers 77mm to 880mm in radius and 17 °to 150 °in polar angle, and it have 8,400 readout channels for anode and 1,792 channels for cathode.

The spatial resolution are 130μm in the transverse plane of the beam axis, and less than 2mm in the beam axis, and that means transverse momentum resolution σ/p_t is $\sqrt{((0.19p_t)^2 + 0.34^2)\%}$ with p_t is GeV/c. And, the resolution of dE/dx is 6.9%.

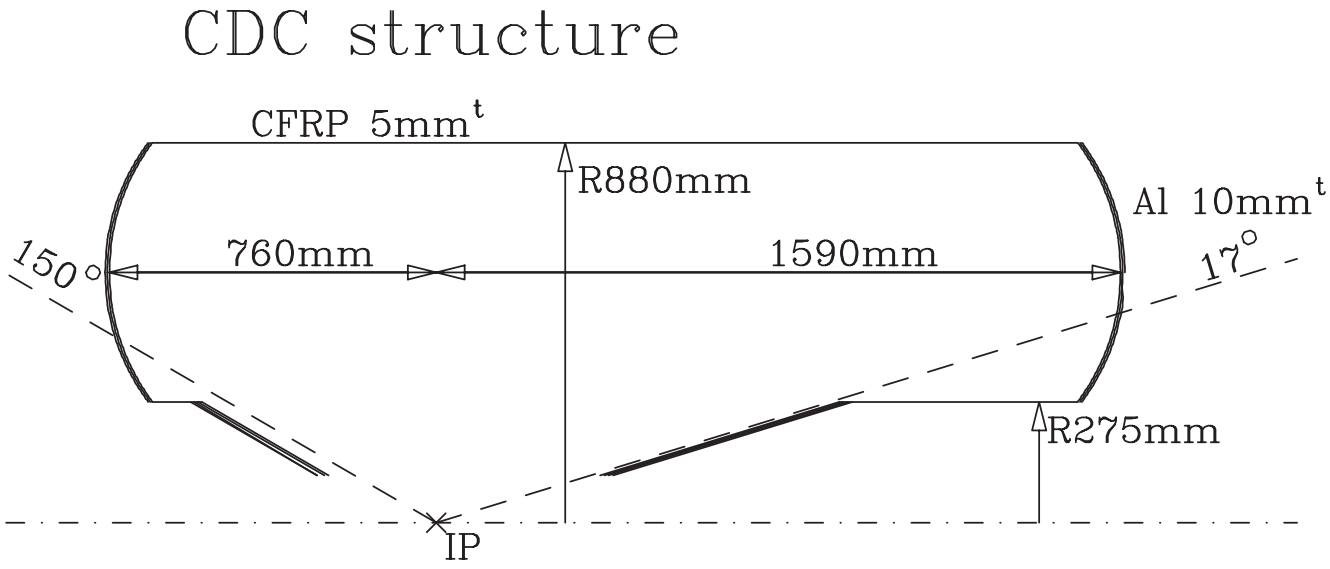


Figure 2.5: Central Drift Chamber (CDC)

2.2.3 Aerogel Cherenkov Counter (ACC)

In the aim of particle identification, Aerogel Cherenkov Counter (ACC) is installed. This counter can distinguish pion from Kaon in the high momentum region above 1.2 GeV/c because ACC detects pion with Cherenkov light which is emitted in Aerogel whose reflective index is 1.01∼ 1.03. Cherenkov light can be measured when particle velocity exceeds light velocity in the material, then the reflective index n is chosen as Kaon don't emit Cherenkov light.

$$\beta > 1/n$$

Aerogel block is typically 12 × 12 × 12 cm^3, and equip with fine-mesh photo-multiplier tubes (FMPMT). ACC has 960 counter in barrel region and 228 counter in endcap region.

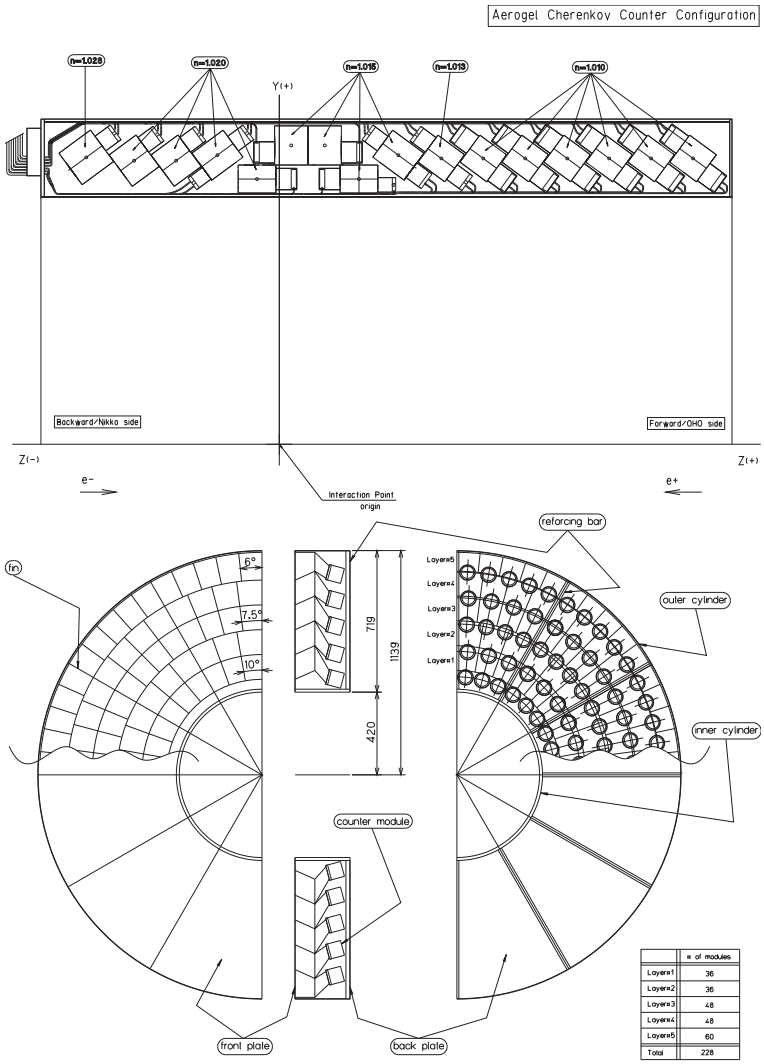


Figure 2.6: Aerogel Cherenkov Counter (ACC)

2.2.4 Time Of Flight counter (TOF)

The role of Time Of Flight counter (TOF) is to measure the velocity of particles for the particle identification of Kaon or pion in the low momentum region below 1.2 GeV/c. The required resolution is 100ps for the flight pass length of 1.2m. The length of this counter is 3m, and 128 counters is located. Each pair of two TOF have one trigger Scintillation Counter (TSC).

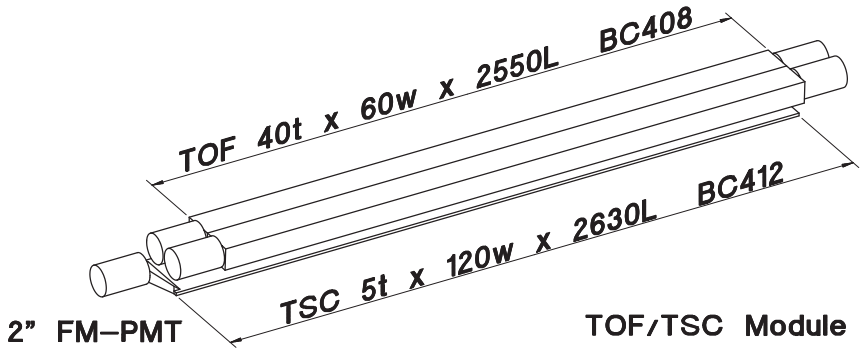


Figure 2.7: Time Of Flight counter (TOF)

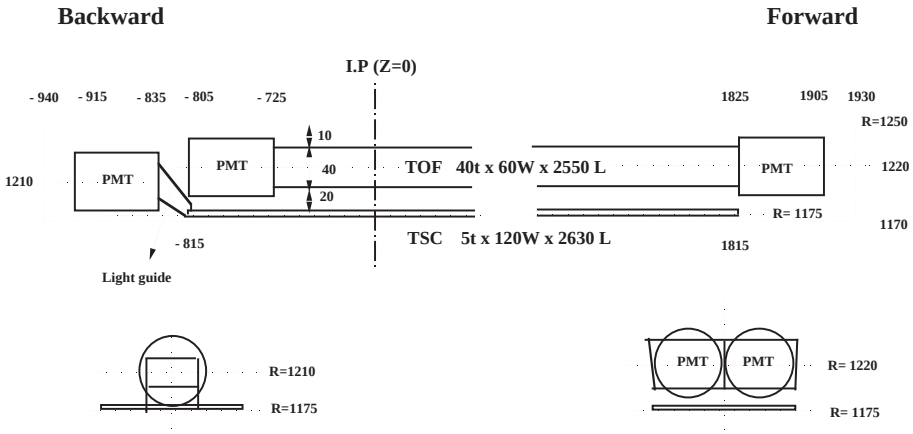


Figure 2.8: Trigger Scintillation Counter (TSC)

K/π separation Particles can be identified with combining the information of CDC, ACC, and TOF. Then, the value of Kaon identification (KID) and pion identification (πID) are defined with likelihood ratio of Kaon and pion as

$$KID = \frac{L(K)}{L(K) + L(\pi)}, \quad \pi ID = \frac{L(\pi)}{L(K) + L(\pi)} = 1 - KID$$

2.2.5 Electromagnetic CaLorimeter (ECL)

In order to measure the energy deposit of photon or electron by the scintillation light, CsI crystals are located. The number of crystal are 8,736 and typically they are 5.5 $\times$ 5.5 $\times$ 30 cm^3 that is correspond to 16 radiation length. Shower of electromagnetic interaction can be measured with 3 $\times$ 3 or 5 $\times$ 5 crystal set, and the ratio of total energy in these two region is used for knowing shower shape. This calorimeter covers 12.01 $^\circ$to 31.36 $^\circ$with the poler angle in barrel region and 32.2 $^\circ$to 128.7 $^\circ$in endcap region. The resolution with 5 $\times$ 5 crystal is $\sigma_E/E = 0.066\ \%/\mathrm{E} \oplus 0.81\ \%/E^{1/4} \oplus 1.34\ \%$ and the position resolution is $0.5\mathrm{cm}/\sqrt{E}$ with E is GeV/c.

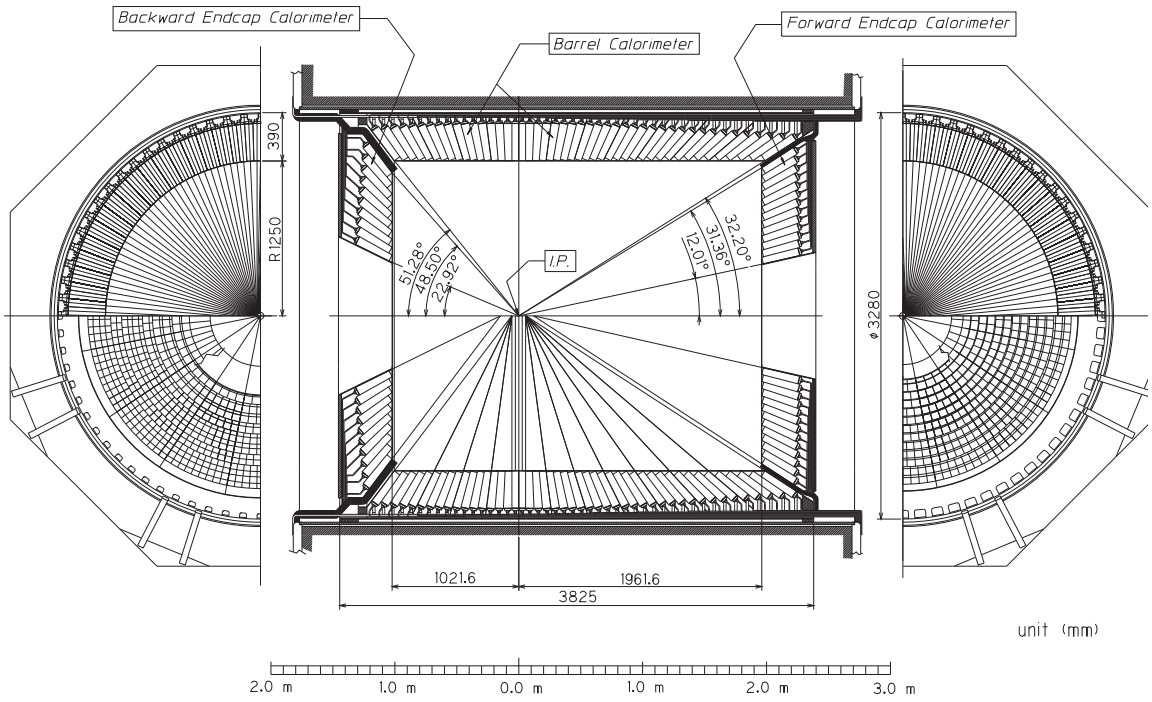


Figure 2.9: Electromagnetic CaLorimeter (ECL)

Electron identification Electrons are identified with the ratio of energy to momentum (E/p). It is close to 1 for electrons. And, we use the ratio of enrgy sum in 3 $\times$ 3 crystal set to enrgy sum in 5 $\times$ 5 crystal set (E9/E25). This use the difference between electromagnetic shower shape and hadronic shower shape. In addition to E/p and E9/E25, we use matching of an extrapolated track position and a cluster position at ECL, dE/dx in CDC, light yeild in ACC, and time of flight in TOF for electron identification. We calculate the likelihood value with pobability density function of these values in Monte Carlo.

2.2.6 Extreme Forward Calorimeter (EFC)

EFC is the detector for electron or photon that is located in extreme forward and backward which covers around the beam axis. EFC is made from 2cm × 2cm BGO ($Bi_4Ge_3O_{12}$) crystals because EFC is exposed in high radiation of photons from synchrotoron radiation or spent electron ($\sim$ 5MRad per year). Typical cross-section is 2cm × 2cm with $12X_0$ for forward and $10.5X_0$ for backward. They cover 6.4 $^\circ\sim$ 11.5 $^\circ$ with the poler angle in forward, and 163.3 $^\circ\sim$ 171.2 $^\circ$in backward.

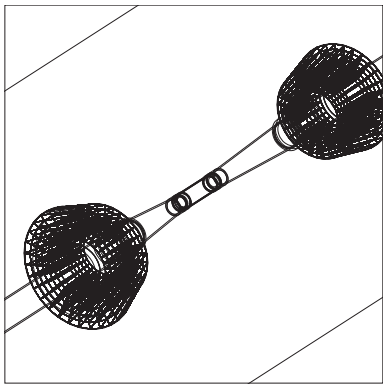


Figure 2.10: Extreme Forward Calorimeter (EFC)

2.2.7 K_L^0 • μ detector (KLM)

In order to detect muon or K_L meson, KLM detector is instolled. The KLM consists of galass Resistive Plate Counter (RPC) and 4.7cm-thick iron that have 14 to 15 layers. The RPC is one of spark chamber that is filled with gas mixture (30% Ar, 62% HFC134a, 8% C_3H_{10}) and it is applied with high voltage (+4.5kV for anode plane nad -3.5kV for cathode plane). Then, charged particle are measured in the RPC, and particles that penetrated through irons are identified as muon. The secondary particles of the interaction between K_L and the iron are measured as a cluster for the direction of K_L. Simoltaniously, the iron plays a role of return yoke. Each layer has two RPC (that is called super-layer) due to detect particles with high efficiency. This KLM detector covers 20 $^\circ\sim$ 155 $^\circ$in polar angle.

Muon identification We identfy muons with charged particles which penetrated through iron layers in KLM. Then, charged tracks which are measured in SVD and CDC are extrapolated into KLM and they are assosiated with KLM hits. We use two quantities for muon identification; one is ΔR which is defined as the difference between expected range of track and measured range of track in KLM. Another is χ_r^2 which is normalized transverce deviation of all hits assosiated with the track. We use probability density function of these two value with tracks of muons, pions and kaons in Monte Carlo to calculate the likelihood.

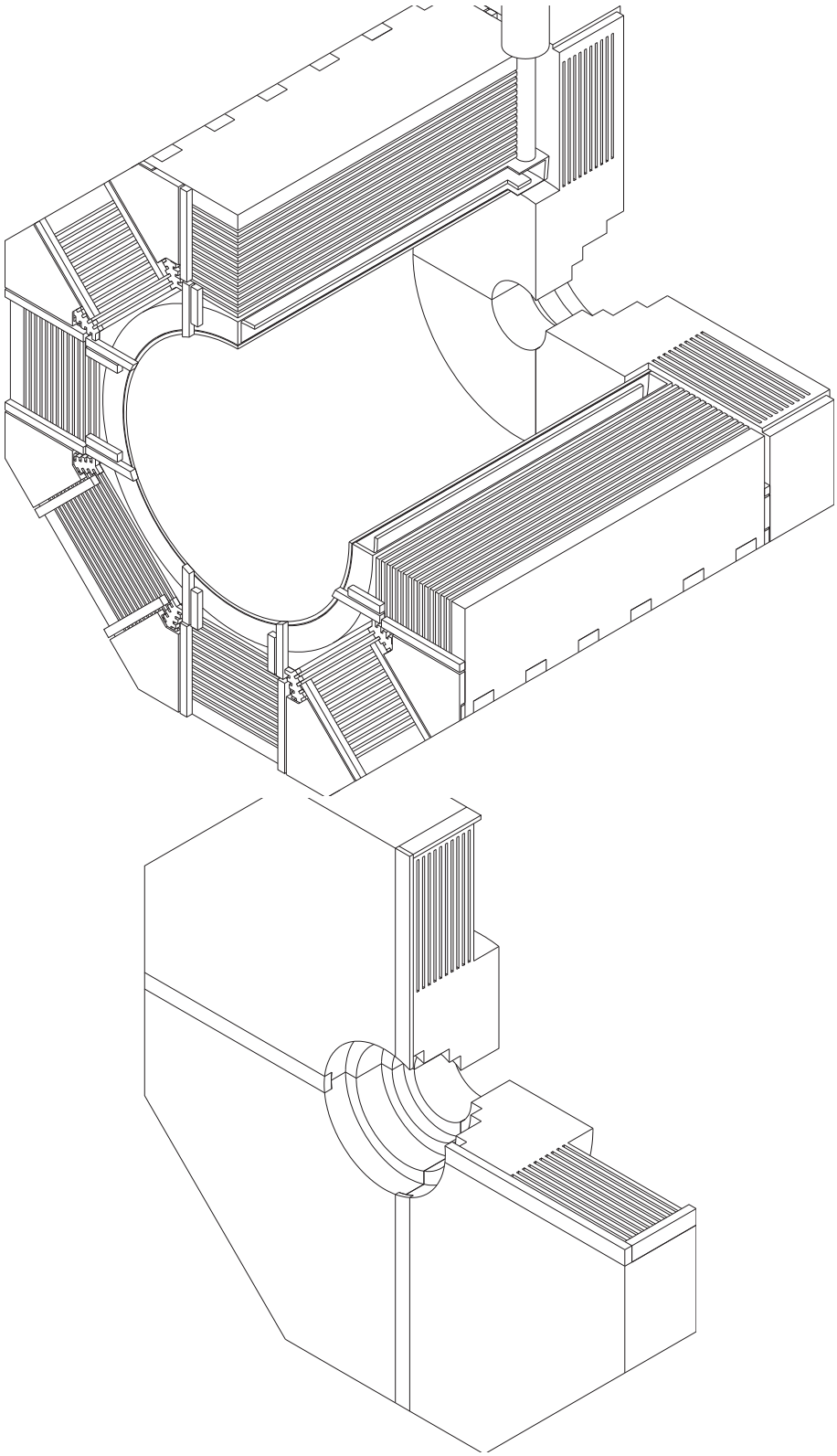


Figure 2.11: K_L^0 • μ detector (KLM)

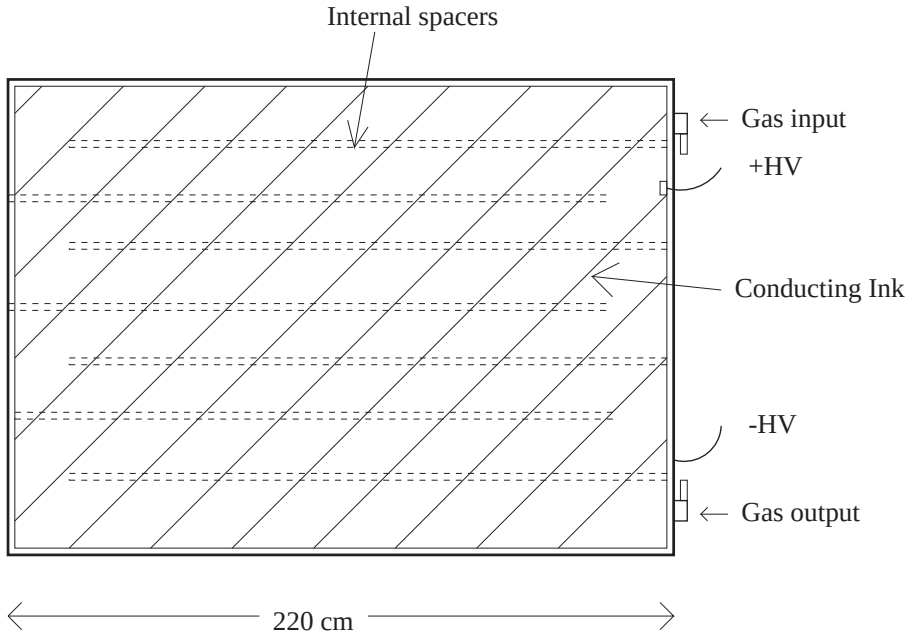


Figure 2.12: Barrel RPC

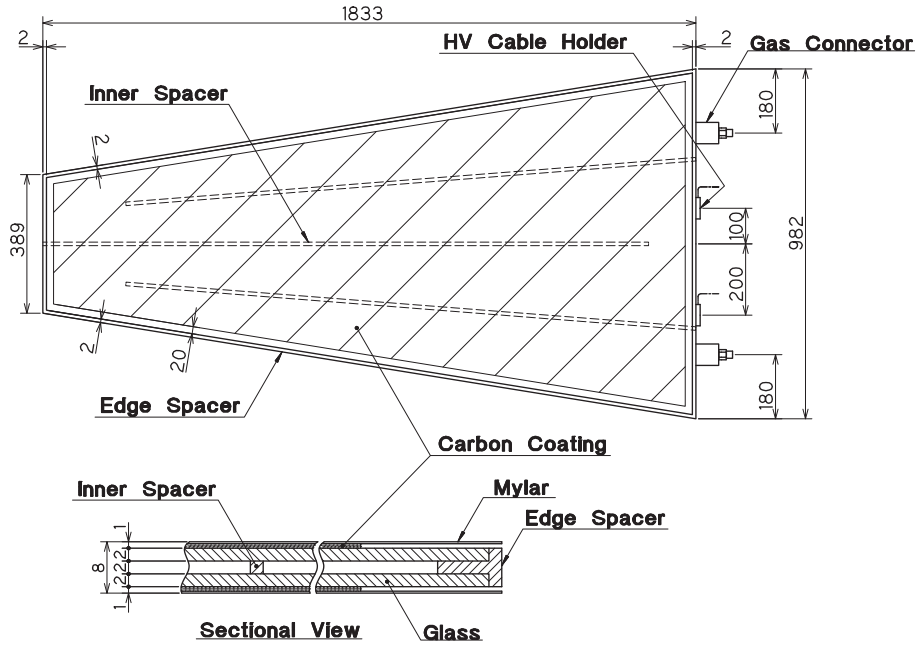


Figure 2.13: Endcap RPC

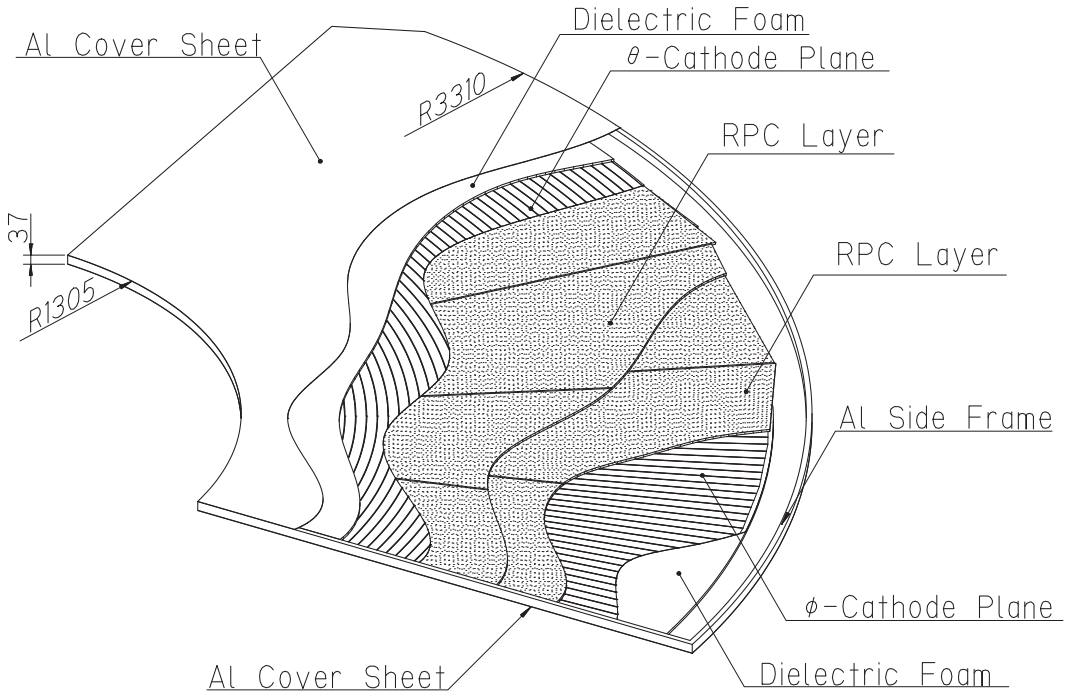


Figure 2.14: Layer structure (KLM)

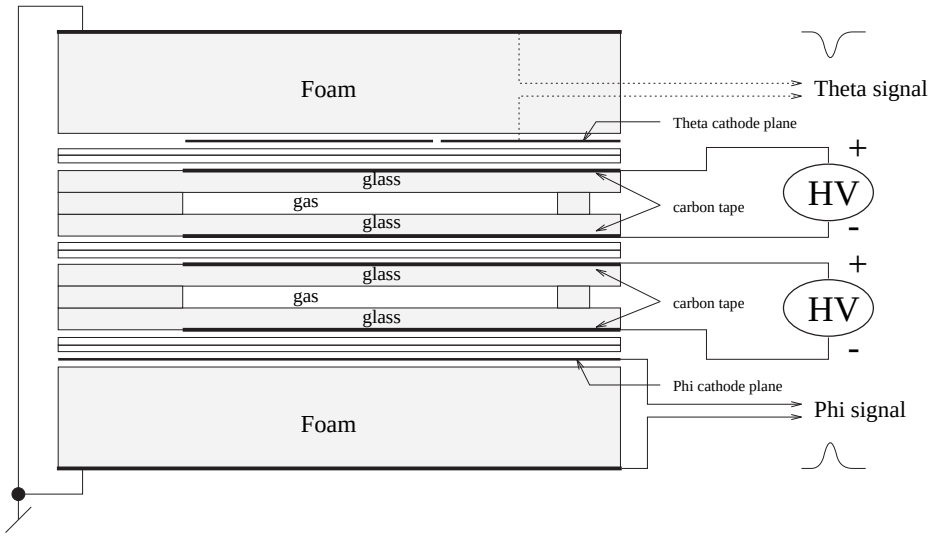


Figure 2.15: Super-layer

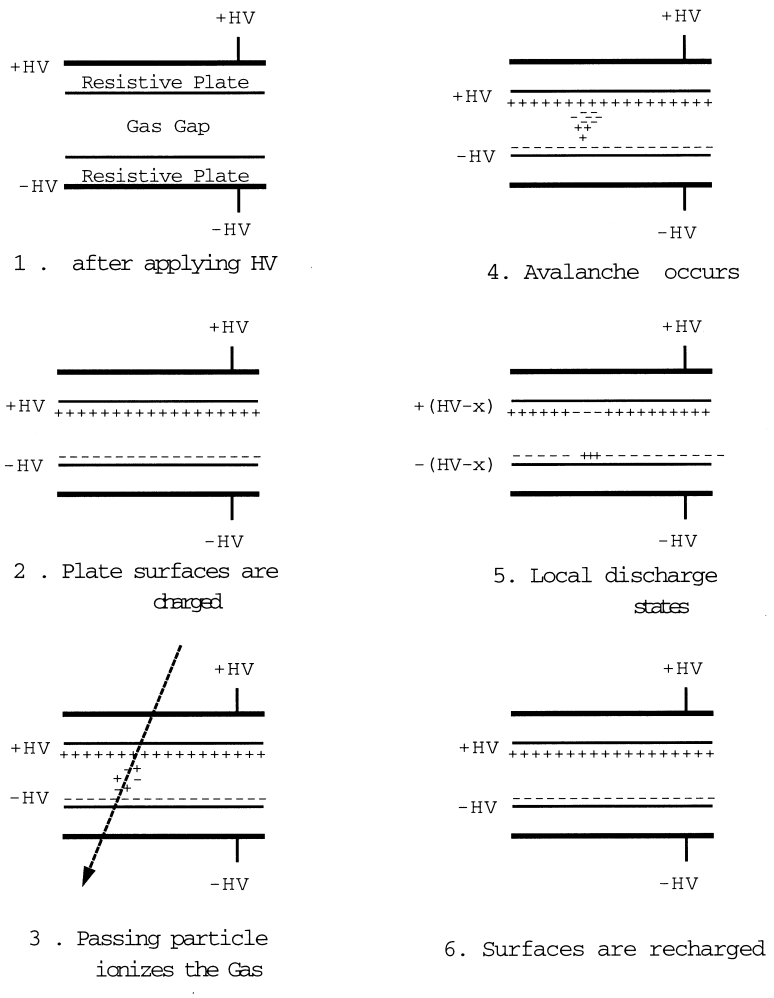


Figure 2.16: Principle of RPC operation

Figure 2.17: Efficiency vs. High voltage

2.3 Trigger and Data AcQuisition (DAQ)

The design values of trigger and data acquision are 200Hz for typical trigger, 500Hz for maximum trigger, 30kB/event for data size, and 15MB/s for data transfer speed.

The trigger consists of sub-detector trigger and central trigger which is called Global Decision Logic (GDL). The sub-detector trigger is combined into GDL. Then two independent trigger exist, one is track trigger and another is energy trigger. The track trigger consists of CDC r-ϕ track, TOF trigger, and isolated ECL cluster trigger. The energy trigger is based on ECL energy sum. The basic hadronic skim is a logical OR of next 4 qualification: tight 2 track trigger, loose 3 track trigger, number of isolated cluster is greater than 4, and energy sum is larger than 1GeV. The exact definition and condition of triggers are in the BELLE note [10]. The efficiency of each trigger are greater than 97%, therefore the final trigger efficiency is greater then 99.9%.

The timing signal is decided from TOF or ECL trigger that are adjusted to 1.85μs from the event crossing, and 0.35μs is used for GDL processing after timing trigger, totally trigger timing is adjusted to 2.2μs from the event crossing.

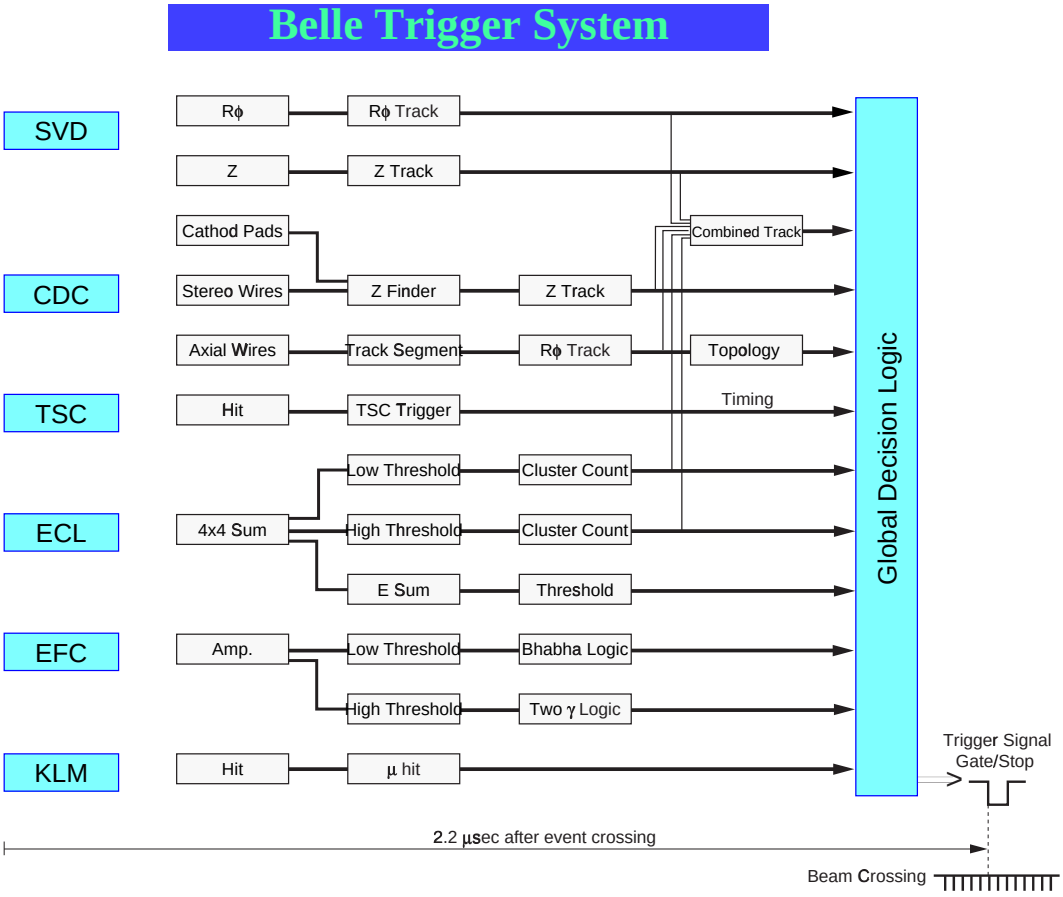


Figure 2.18: Trigger system

The readouts from sub-detector and trigger are running in parallel. Charge-to-Time (Q-to-T) and Time-to-digital-converter (TDC) are used to read out signals from most of sub-detectors. The data from each subdetector are combined in event builder, and the output is transfered to the online computer farm. Finally the data are sent to the storage system and stored in the tapes.

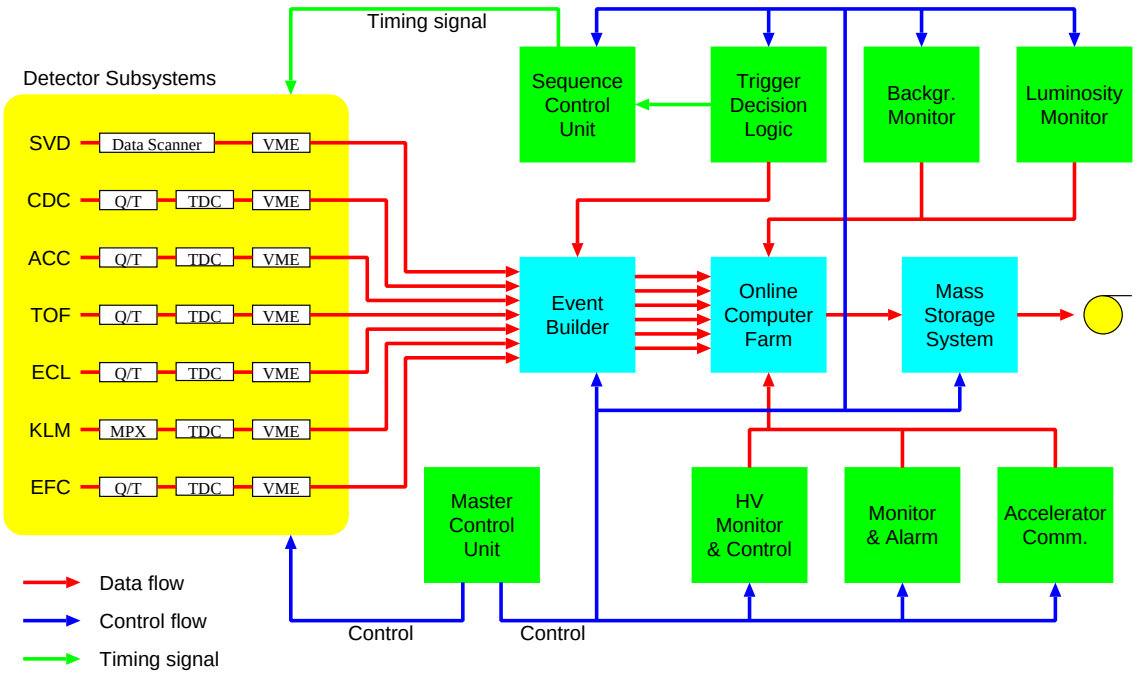


Figure 2.19: Daq system

After the software trigger which have almost 100% efficiency for hadronic event, each event is classified into some categories. In this analysis, we use the data set that called "HadronB" which is categolized as standerd hadronic event. The conditions of "HadronB" are as follows,

- Number of "good" track multiplicity is greater than 3.
  ("good" means $p_t > 0.1$GeV, $|dr| < 2.0$cm, and $|dz| < 4.0$cm)
- The visible energy of tracks and photons (E_{vis}), is greater than 20 % of $\sqrt{s}$.
- Momentum balance of z compornet; $|\Sigma p_z| \leq 0.5\sqrt{s}$.
- Primary event vertex is around interraction point; $|r| < 1.5$cm and $|z| < 3.5$cm.
- Energy sum in ECL; $0.18\text{GeV} < E_{sum}/\sqrt{s} < 0.8\text{GeV}$
- Number of ECL clusters in the region, $-0.7 < \cos\theta < 0.9$, is greater than 2
- Heavy jet mass; $M_{jet} > 0.25E_{vis}$ or $M_{jet} > 1.8$GeV.
  (Heavy jet mass means essencially invariant mass in one hemispere [15])
- Average of ECL clusters's energy is smaller than 1GeV.

2.4 Software tool

The overview of the data analysis and Monte Carlo simulation is shown on 2.20. The raw data which are acquited with BELLE detector are processed by reconstruction tools such as tracking of charged particle, measuring of energy and particle identification. Outputs of reconstruction tools is called Data Summary Tape (DST) and it is converted to more compact data set that is called Mini Data Summary Tape (MDST) for user analysis. When we simulate with BELLE detector, we use QQ generator [11] as event generator which is originally developed by CLEO collaboration and modified for BELLE detector [12], then we use two detector simulation; one is full detector simulator (GSIM) that generates detector response in the same form as real data. Another is fast detector simulator (FSIM) that use parametarized performances as the detector performances for fast simulation.

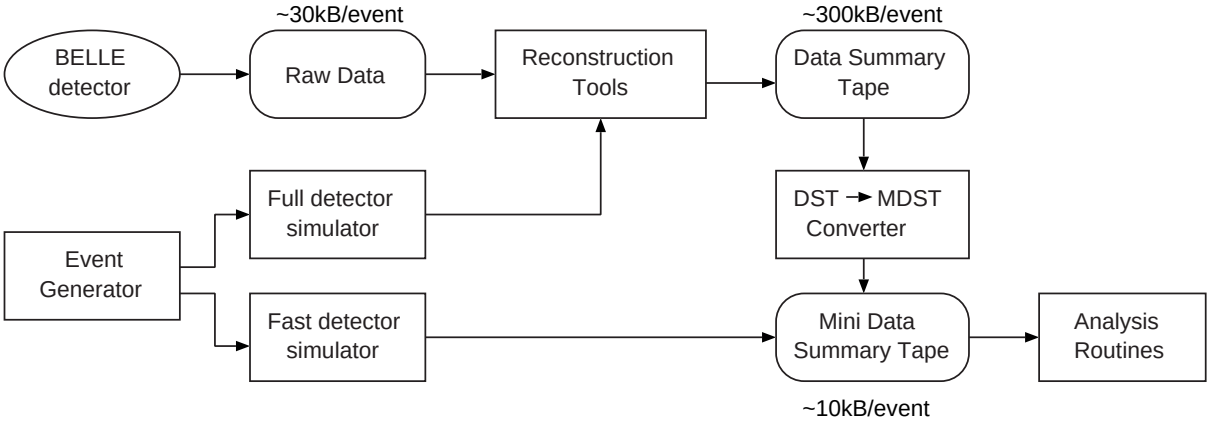


Figure 2.20: Flow of analysis

Event generator Event generator (QQ) simulates physical process of decay chains. Then, we use each decay mode and branching fraction in the result of CLEO group which are modified for BELLE detector. Decay modes and branching fractions are recorded in a decay table, and user can controll them by chainging the decay table. The outputs of decay information with QQ are stred in HEPEVT table [13]. There are two main states as initial state; one is $\Upsilon(4S)$ and another is $q\overline{q}$ (continuum prosess). Most of $\Upsilon(4S)$ decay into $B\overline{B}$ pair, and on the energy of $\Upsilon(4S)$ resonance, the main background is $e^+e^- \to q\overline{q}$ event ($\sigma(e^+e^- \to q\overline{q}) \simeq 3\sigma(e^+e^- \to \Upsilon(4S))$).

Geant (full) detector SIMulator (GSIM) The full detector simulator (GSIM) is based on GEANT [14] which is developed in CERN for the simulation of reaction between particles and materials in the detector. The data from HEPEVT table are inputed into GSIM, and GSIM traces the behavior of each particles in the detector and generates detector responce which is the same form with real data output. GSIM takes much time ($\sim$ 30 sec/event) because of tracing each particle one by one for the precise reactions with materials.

Fast detector SIMulator (FSIM) For the aim of fast simulation, fast detector simulator (FSIM) is developed. FSIM uses parametarized performance as detector performance that is based on the result of GSIM. The output of FSIM is format of MDST, and it means that user can analyze the outputs directly.

Chapter 3

Analysis of $B \rightarrow inclusive\ D_s$ decay

This analysis use 29.1 fb^{-1} on-$\Upsilon(4S)$-resonance data and $3.0fb^{-1}$ off-resonance data (which is taken 60MeV below the $\Upsilon(4S)$ peak in order to analyze for the background of $B\overline{B}$ events). The D_s momentum spectrum is used to extract $b \rightarrow u\ D_s$ decay. Then, we use a scaled momentum x which is defined as

$$x = \frac{P^*_{D_s}}{P^*_{max}} = \frac{P^*_{D_s}}{\sqrt{E^2_{beam} - M^2_{D_s}}},$$

where $P^*_{D_s}$ is the momentum of D_s in the e^+e^- center of mass system and P^*_{max} is its kinematic maximum, E_{beam} is the beam energy (5.29GeV for on resonance and 5.26GeV for off-resonance) in the same system, and M_{D_s} is the nominal D_s mass (1.968GeV).

Although the decay $b \rightarrow u\ D_s$ is two-body decay in quark level, the monochromatic D_s momentum spectrum is smeared by the b-quark Fermi motion and the B meson velocity ($\beta \sim$0.065). The endpoint of $b \rightarrow c\ D_s$ and that of $b \rightarrow u\ D_s$ are calculated to be 0.41 or 0.50 for $B \rightarrow D\ D_s$ decay and $B \rightarrow \pi\ D_s$ decay respectively. The detail of the calculation follows.

Assumption of $B \rightarrow \pi\ D_s$ decay means

$$M^2_\pi = (E_B - E_{D_s})^2 - (\vec{P_B} - \vec{P_{D_s}})^2.$$

In the B meson rest frame, we have

$$E_{D_s} = \frac{M^2_B + M^2_{D_s} - M^2_\pi}{2M_B}.$$

And, correction of $\beta_B \sim$0.065 introduce,

$$P^*_{D_s} = \gamma\beta E_{D_s} + \gamma P_{D_s} \sim 2.47GeV,$$

$$x_{end} = \frac{P^*_{D_s}}{\sqrt{E^2_{beam} - M^2_{D_s}}} \sim \frac{2.47GeV}{4.91GeV} \sim 0.503.$$

In the same way, the end point of x for $B \rightarrow D\ D_s$ decay is calculated as 0.406.

We define the $b \rightarrow u\ D_s$ transition signal region as $0.42 \leq x < 0.50$, then, only 2% of signal number is smeared by $b \rightarrow c\ D_s$ spillover in Monte Carlo study (as shown on the page 43). The main background source is continuum process ($e^+e^- \rightarrow q\bar{q}$), and off-resonance data is used as continuum events.

3.1 Event selection

The $D_s^{\pm}$ is reconstructed by the decay chain $D_s^+ \rightarrow \phi\ \pi^+$, $\phi \rightarrow K^+\ K^-$. Then, we use cuts of the particle identification for $K^{\pm}$ meson, cut of the invariant mass for ϕ meson, a cut of particle identification for $\pi^{\pm}$ meson and two angle cuts (defined in Figure 3.1). Each cut criteria are as follows,

Cut	Selection criteria
M_{K+K-}	at least one K ID ≥ 0.5 and another K ID ≥ 0.2 $\lvert M_{K^+K^-} - M_\phi \rvert \leq 10MeV$
$M_{\phi\pi}$	π ID ≥ 0.1
D_s helicity angle	$\lvert cos\theta_h \rvert \geq 0.35$
B decay angle	$cos\theta_d \leq 0.85$

Table 3.1: Selection criteria

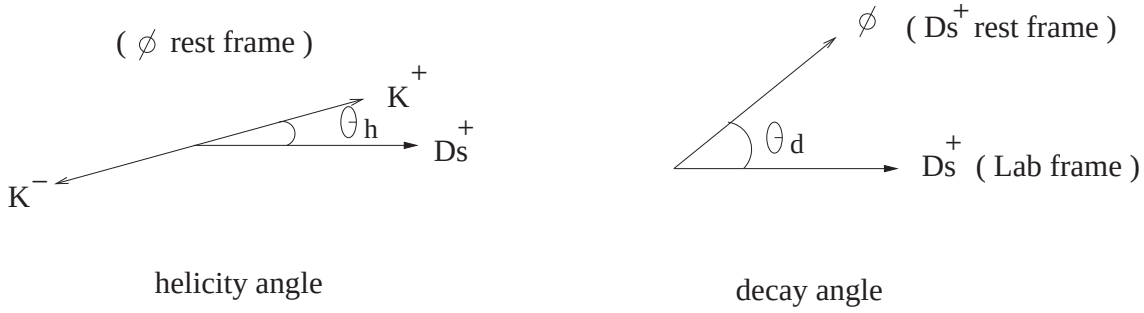


Figure 3.1: Angle definitions

The distribution of $cos\theta_h$ for the combinatorial background is flat, while that for the signal is proportional to $cos^2\theta_h$. The distribution of $cos\theta_d$ for the combinatorial with slow pion background peaks near 1, while that for the signal is flat.

We choose cut values as to maximize the figure of merit (F.O.M.) that is defined as

$$F.O.M. = \frac{S}{\sigma_N},$$

where S is the number of D_s with the signal $B \rightarrow \pi\ D_s$ Monte Carlo, and σ_N is the error of D_s number with background continuum Monte Carlo, for the region of $0.42 \leq$ x $<$ 0.54. Each $M_{\phi\pi}$ distribution is fitted with a double Gaussian which is liner combination of two Gaussian as the D_s signal shape and liner function as the combinatorial background shape, respectively. The statistics are 10,000 events for the decay chain; $B^0 \rightarrow \pi^-\ D_s^+$, $D_s^+ \rightarrow \phi\ \pi^+$ and $\phi \rightarrow K^+\ K^-$, and 15,000,000 events for the background continuum events. Each distribution for F.O.M. vs. cut-value are shown on Figure 3.2, 3.3 and 3.5. We use the values which maximize F.O.M. as shown already in the Table 3.1.

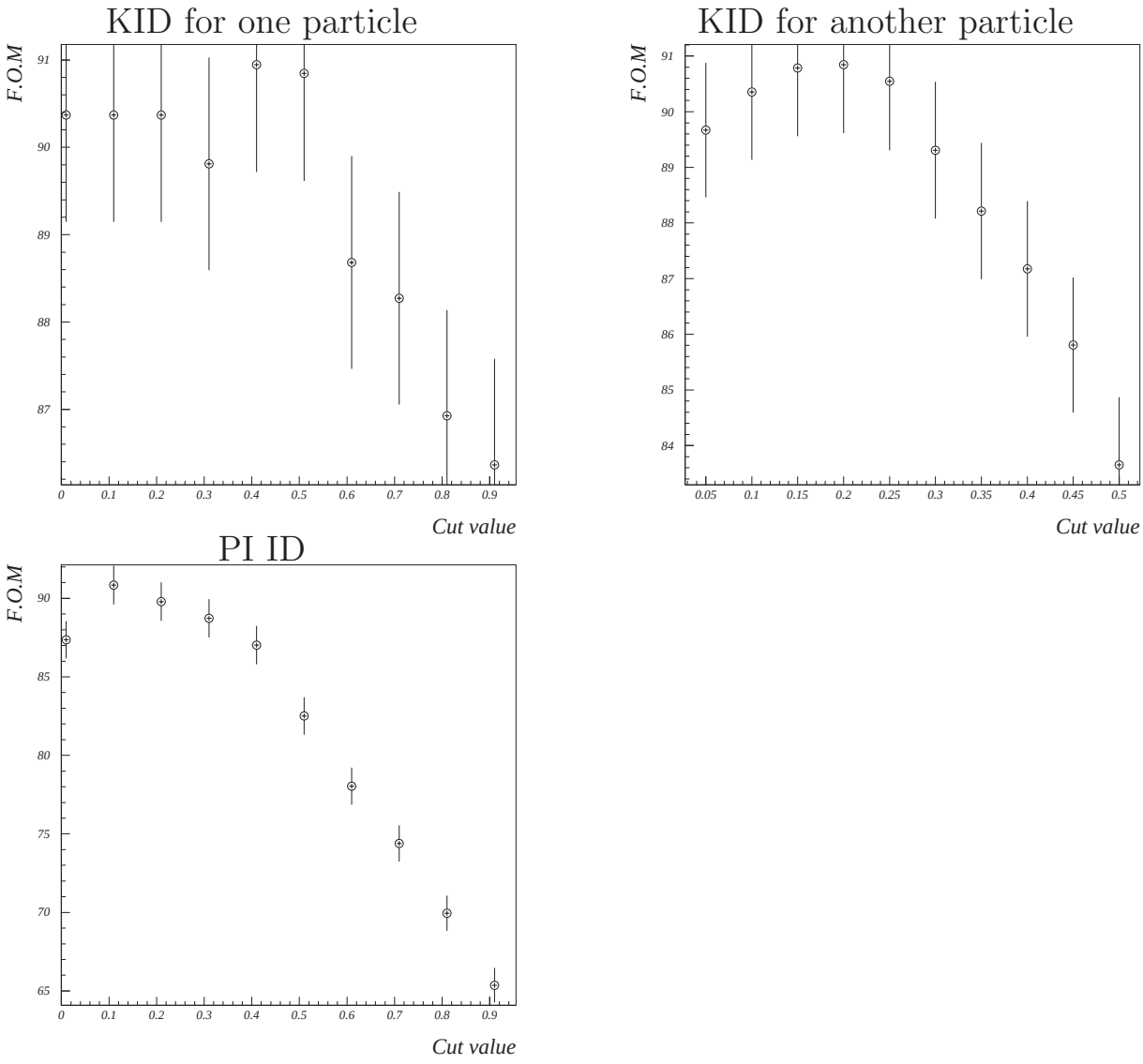


Figure 3.2: F.O.M. for particle identification

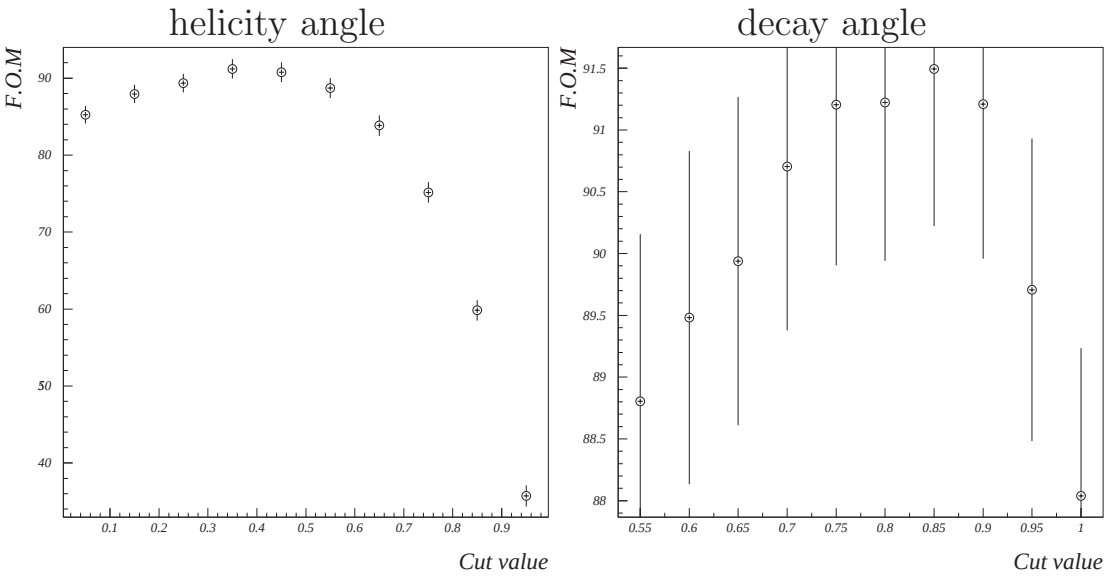


Figure 3.3: F.O.M. for angle cut

We use the event shape variable R_2 for the suppression of continuum background. R_2 is the Fox-Walfram [16] second moment normalized to the zeroth moment. This variable can distinguish jet like events such as continuum event from spherical events such as $B\overline{B}$ event. Figure 3.4 is the distributions of R_2 for the signal MC and the continuum MC. We require R_2 is less than 0.4 for the branching-fraction analysis because of maximizing F.O.M., and require less than 0.3 for the $|V_{ub}|$ measurement because this measurement is sensitive for continuum background. Efficiencies of only R2 cuts are 94.0 *and*98.5in 400,000 events of signal Monte Carlo which are the decay chain; $B^0 \rightarrow X_c\, D_s^+$, $D_s^+ \rightarrow \phi\, \pi^+$ and $\phi \rightarrow K^+\, K^-$. Table 3.2 is the list of decay chains and branching ratio of each decay which is contained in this signal Monte Carlo sample.

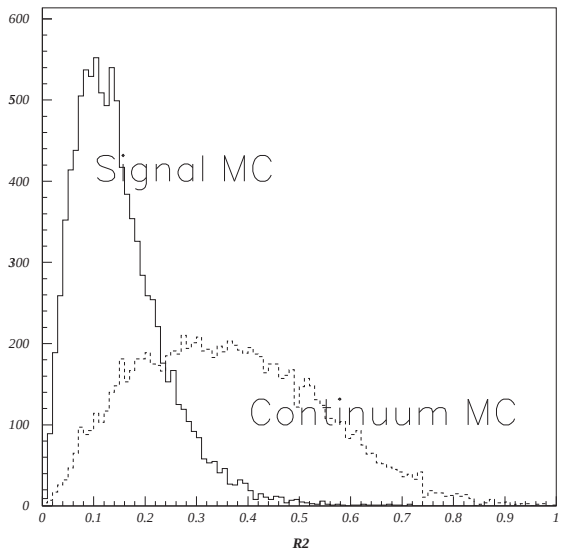


Figure 3.4: R_2 distribution

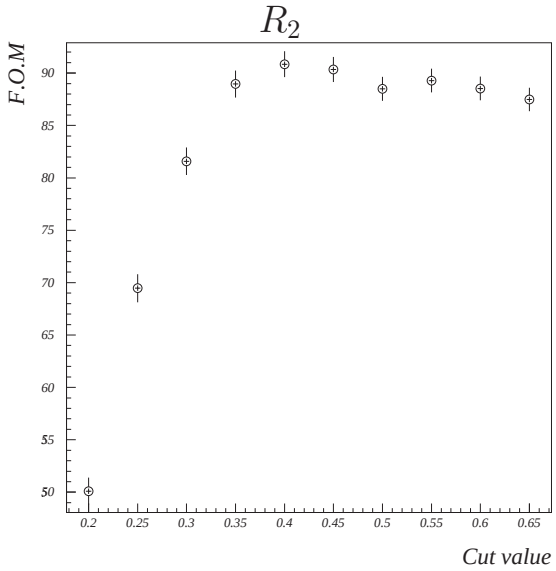


Figure 3.5: F.O.M. for R2

Decay chain	Branching ratio
$B^0 \rightarrow D_s^+ D^-$	9.4%
$B^0 \rightarrow D_s^+ D^{*-}$	9.0%
$B^0 \rightarrow D_s^{*+} D^-$	9.4%
$B^0 \rightarrow D_s^{*+} D^{*-}$	17.4%
$B^0 \rightarrow D_s^+ D^- \pi^0$	5.0%
$B^0 \rightarrow D_s^+ \overline{D^0} \pi^-$	10.0%
$B^0 \rightarrow D_s^{*+} D^- \pi^0$	5.0%
$B^0 \rightarrow D_s^{*+} \overline{D^0} \pi^-$	10.0%
$B^0 \rightarrow D_s^+ D^- \pi^+\pi^-$	5.0%
$B^0 \rightarrow D_s^+ D^- \pi^0\pi^0$	3.7%
$B^0 \rightarrow D_s^+ \overline{D^0} \pi^-\pi^0$	3.7%
$B^0 \rightarrow D_s^{*+} D^- \pi^+\pi^-$	5.0%
$B^0 \rightarrow D_s^{*+} D^- \pi^0\pi^0$	3.7%
$B^0 \rightarrow D_s^{*+} \overline{D^0} \pi^-\pi^0$	3.7%
$B^0 \rightarrow D_s^+ X_c$	100.0%

Table 3.2: Branching ratio of inclusive Ds decay

3.1.1 Result of Real Data

With the on-resonance data of 29.1 fb^{-1}, ϕ mass distribution which is reconstructed from K^+K^- is shown on Figure 3.6, and ϕ mass distribution with the off-resonance data of 3.0 fb^{-1} is shown on Figure 3.7. Then, we chose cut values of Kaon identification as to maximize the F.O.M. that mean at least one KID is greater than 0.5 and another KID is greater than 0.2. We fit these distribution with double Gaussian as signal shape and 5th polynomial as background shape for the region of 1.0 $GeV/c^2 \leq M_{kk} \leq 1.035$ GeV/c^2. At first, a mean parameter of this double Gaussian is given with free parameter, and after fitting, we fix the mean with 1.0195 Gev/c^2 of the nominal mass and we fix the parameters of two width with 0.0026 GeV/c^2 and 0.01 GeV/c^2 from Monte Carlo study as shown later. Figure 3.8 is the same distribution which only focused on the result of two KID for on-resonance data, and Figure 3.9 is also focusing on the distribution of off-resonance data. We choose the mass window as 1.01 $GeV/c^2 \leq M_{kk} \leq 1.03$ GeV/c^2 that is correspond to 1.8σ with the weighted deviation of this double Gaussian in Monte Carlo study. The number of ϕ with each cuts are summarized in the Table 3.3, then these cuts are applied in the turn of this list, and the efficiencies of only one cut which is also applied in the same turn are summarized in the Table 3.4. At that time, we calculate the number of ϕ for $B\overline{B}$ event with subtracting off-resonance data from on-resonance data. Then, we use 9.6 as the value of the scale factor to normalize off-resonance data to on-resonance data. It is calculated with the luminosity ratio and square inverse of the beam energy in the center of mass system as a cross section ratio like below,

$$\frac{29.1}{3.0} \times \frac{10.52 \times 10.52}{10.58 \times 10.58} \sim 9.6 \ .$$

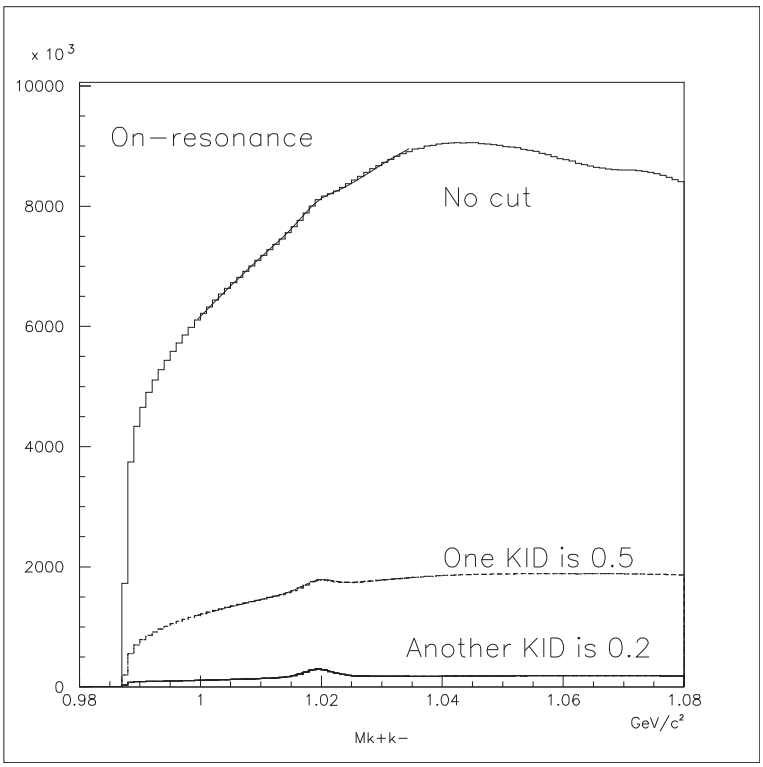


Figure 3.6: $M_{K^+K^-}$ for on-resonance (Real Data)

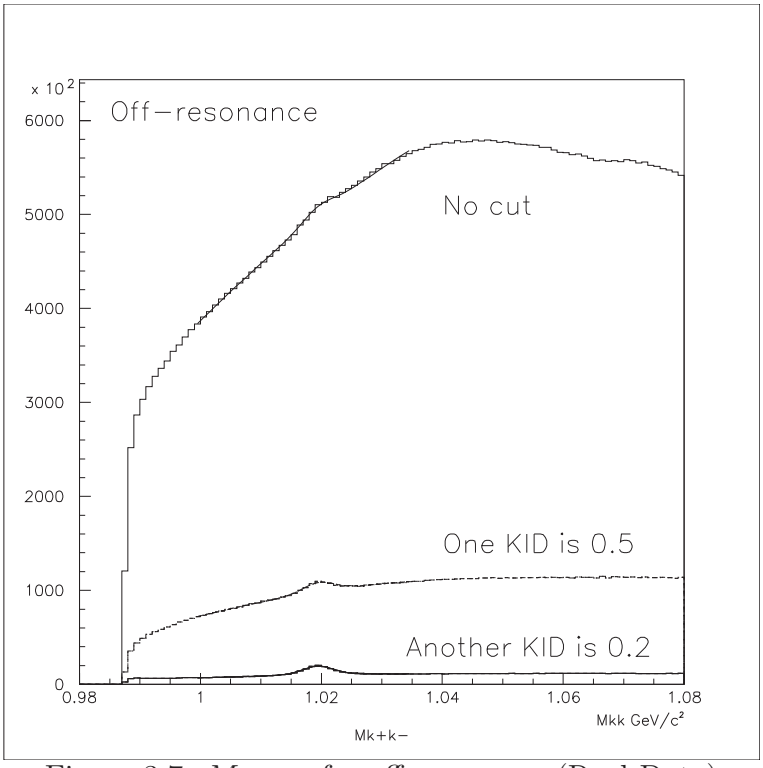


Figure 3.7: $M_{K^+K^-}$ for off-resonance (Real Data)

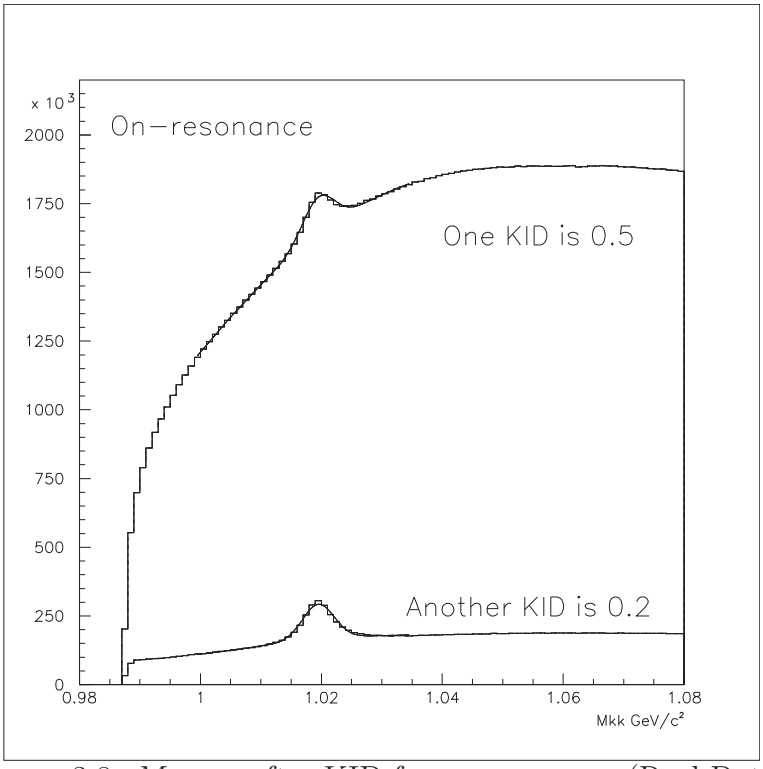


Figure 3.8: $M_{K^+K^-}$ after KID for on-resonance (Real Data)

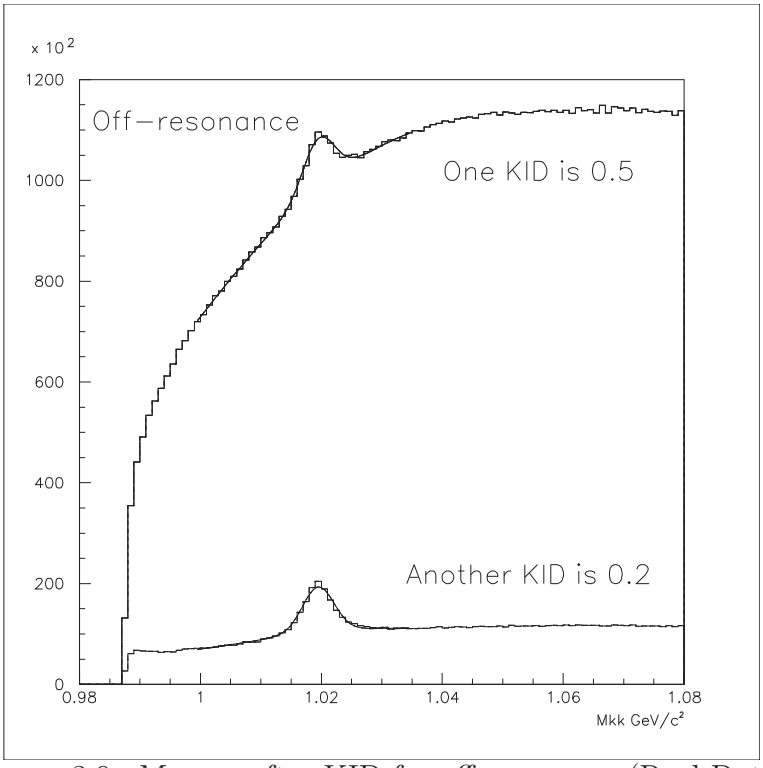


Figure 3.9: $M_{K^+K^-}$ after KID for off-resonance (Real Data)

cut	cut criteria	on-resonance Number of ϕ	off-resonance Number of ϕ	$B\overline{B}$ Number of ϕ
No KID		1,554,000 ± 8,244 (0.5%)	107,900 ± 3,928 (3.6%)	526,800 ± 38,600 (7.3%)
one KID	KID $\geq$ 0.5	1,490,000 ± 7,136 (0.5%)	108,900 ± 1,734 (1.6%)	442,600 ± 18,110 (4.1%)
another KID	KID $\geq$ 0.2	1,383,000 ± 2,565 (0.2%)	96,820 ± 664 (0.7%)	452,700 ± 6,871 (1.5%)

();relative error.

Table 3.3: Number of ϕ with each cut for M_{kk} (Real Data)

cut	cut criteria	on-resonance ϕ efficiency	off-resonance ϕ efficiency	$B\overline{B}$ ϕ efficiency
one KID	KID $\geq$ 0.5	95.8% ± 0.7%	100.9% ± 4.0%	84.0% ± 7.1%
another KID	KID $\geq$ 0.2	92.8% ± 0.5%	88.9% ± 1.5%	102.3% ± 4.5%
Total of 2KID	2KID	88.9% ± 0.5%	89.7% ± 3.3%	85.9% ± 6.4%

Table 3.4: Efficiency of each cut for M_{kk} (Real Data)

The invariant mass distribution of D_s which are reconstructed from the $D_s^+ \rightarrow \phi\ \pi^+$ decay chain with on-resonance data of 29.1 fb^{-1} or off-resonance data of 3.0 fb^{-1} are Figure 3.10 or Figure 3.11. Then, we choose each cut value as to maximize the F.O.M. as shown before and summarized already in the Table 3.1. We count the number of D_s by fitting with double Gaussian for the signal shape and liner function for the background shape. At that time, we fix the mean of two Gaussian with nominal mass of D_s, one of widths in double Gaussian as $0.0036 Gev/c^2$, the ratio of 2 area as 0.48 and the ratio of two width as 2.9, this means only one parameter is free parameter for signal shape. The number of D_s with each cut are summarized in the Table 3.5 The efficiencies of only each cut which is applied in turns after other cuts are summarized in the Table 3.6 which are for $29.1 fb^{-1}$ of on-resonance data, for off-resonance data of 3.0 fb^{-1}, and the calculated efficiency for $B\overline{B}$ pair. We use the error of area in double Gaussian as the error of D_s number. As a systematic error of each cut except for R2 cut , we use error of efficiencies for on-resonance data because of high statistics and comparability of the efficiencies between on-resonance and off-resonance data.

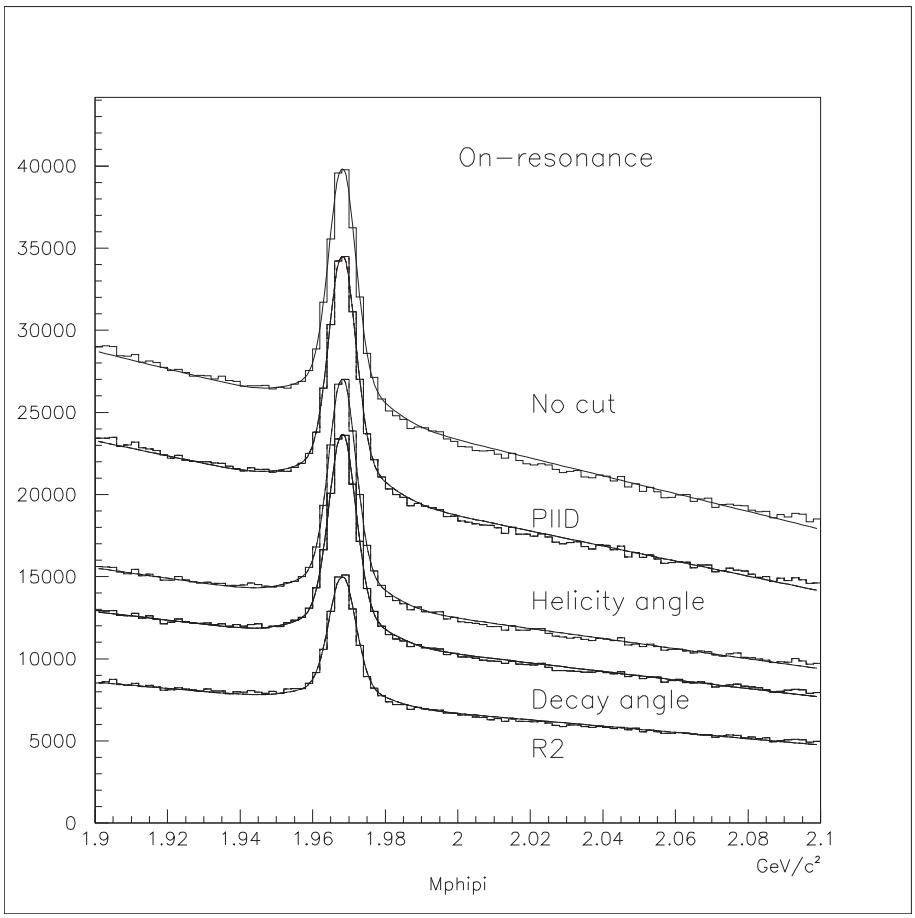


Figure 3.10: $M_{\phi\pi}$ of on-resonance for all momentum (Real Data)

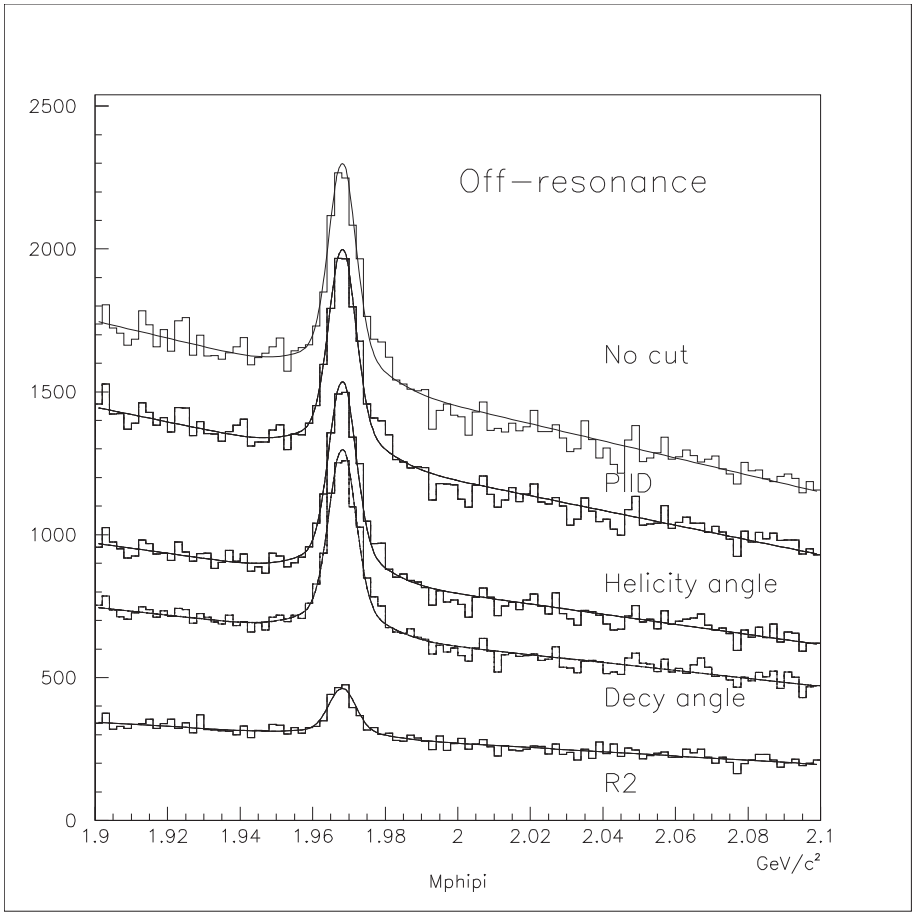


Figure 3.11: $M_{\phi\pi}$ of off-resonance for all momentum (Real Data)

cut	cut criteria	on-resonance Number of D_s	off-resonance Number of D_s	$B\overline{B}$ Number of D_s
No cut		86,180 ± 601 (0.7%)	4,403 ± 155 (3.5%)	43,910 ± 1,605 (3.7%)
π ID	π ID ≥ 0.1	83,690 ± 571 (0.7%)	4,241 ± 131 (3.1%)	42,970 ± 1,381 (3.2%)
D_s helicity angle	$\lvert cos\theta_h \rvert \geq 0.35$	79,210 ± 494 (0.6%)	4,000 ± 117 (2.9%)	40,840 ± 1,166 (2.9%)
B decay angle	$cos\theta_d \leq 0.85$	73,250 ± 456 (0.6%)	3,760 ± 110 (2.9%)	37,140 ± 1,150 (3.1%)
event shape	R2 ≤ 0.3	44,910 ± 367 (0.8%)	978 ± 66 (6.7%)	35,520 ± 735 (2.1%)

();relative error.

Table 3.5: Number of D_s with each cut for $M_{\phi\pi}$ (Real Data)

cut	cut criteria	on-resonance D_s efficiency	off-resonance D_s efficiency	$B\overline{B}$ D_s efficiency
π ID	π ID ≥ 0.1	97.2% $\pm$ 0.9%	96.3% $\pm$ 4.5%	97.9% $\pm$ 4.8%
D_s helicity angle	$\lvert cos\theta_h \rvert \geq 0.35$	94.6% $\pm$ 0.9%	94.2% $\pm$ 4.0%	95.0% $\pm$ 4.1%
B decay angle	$cos\theta_d \leq 0.85$	92.5% $\pm$ 0.8%	94.1% $\pm$ 3.9%	90.9% $\pm$ 3.8%
event shape	R2 ≤ 0.3	61.3% $\pm$ 0.6%	26.3% $\pm$ 1.9%	95.6% $\pm$ 3.6%
Total of $M_{\phi\pi}$ cuts	all	52.1% $\pm$ 0.6%	22.2% $\pm$ 1.7%	80.9% $\pm$ 3.3%

Table 3.6: Efficiency of each cut for $M_{\phi\pi}$ (Real Data)

The D_s helicity angle ($cos\theta_h$) distribution in 21.3 fb^{-1} on-resonance data is Figure 3.12. At that time, we count the number of D_s by fitting for each $M_{\phi\pi}$ distribution that is reconstructed in the region of each bin. This helicity angle shows $cos^2\theta$ distribution as explain below, and the arrows in this distribution are cut value.

The spin of Ds, ϕ, and π which are 0, 1, and 0 introduce the orbital angular momentum of $\phi\pi$ as 1, two body decay of $D_s \rightarrow \phi\pi$ means the orbital angular momentum of $\phi\pi$ with ϕ direction as 0, the spin of ϕ, K^+, and K^- which are 1, 0, and 0 introduce the orbital angular momentum of K^+K^- as 1, and the spin of ϕ with the last ϕ direction and the spin of $K^\pm$ introduce the orbital angular momentum of K^+K^- with the same direction as 0, therefore, spherical harmonic function become $Y_1^0 = \sqrt{3/4\pi}\ \cos\theta$, and it means that the amplitude of this decay is proportional to $cos^2\theta$ distribution.

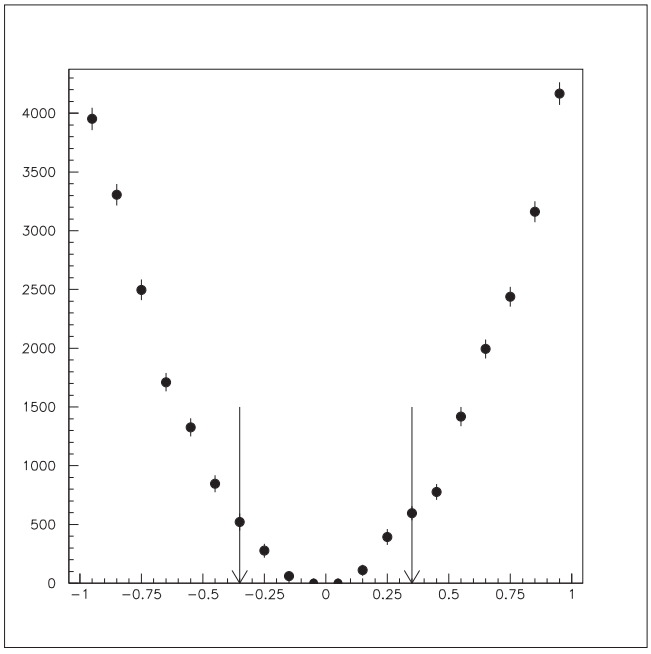


Figure 3.12: D_s helicity distribution (Real Data)

3.1.2 Monte Carlo study

In the signal Monte Carlo of 100,000 events for the decay chain; $B^0 \rightarrow X_c\ D_s^+$, $D_s^+ \rightarrow \phi\,\pi^+$ and $\phi \rightarrow K^+\,K^-$, the ϕ mass distribution becomes like Figure 3.13. Each histograms are distributions for without KID cut, after the cut of one KID is greater than 0.5, and after the cut of another KID is greater than 0.2. The number of ϕ with each cut and the efficiencies of each cut which are applied in turns after other cuts are summarized in the Table 3.7. Then, we use double Gaussian as signal shape which have $0.0026 GeV/c^2$ and $0.01 GeV/c^2$ as two widths, and ARGUS function which defined as below as background shape.

$$ARGUS(x) = P_1 \times \sqrt{\frac{x^2}{0.987^2} - 1} \times exp[P_2(1 - \frac{x^2}{0.987^2})]$$

where P_1 and P_2 are the free parameters of fitting and 0.987 is a sum of daughter particle's mass (in the case of this analysis, twice of charged Kaon mass), the typical values of parameters are about 4 times of background height for P_1 and 2 to 4 for P_2.

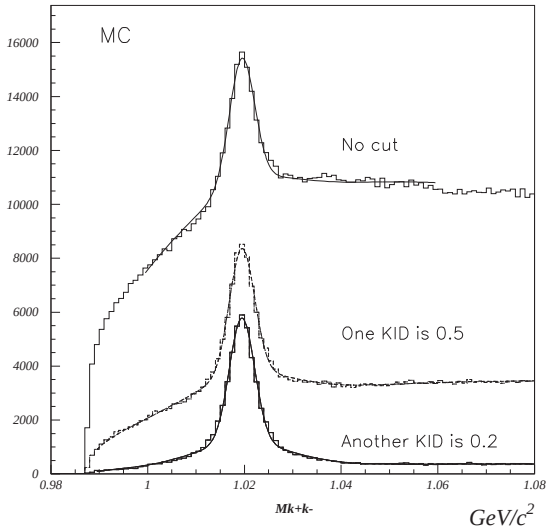


Figure 3.13: $M_{K^+K^-}$ (Monte Carlo)

cut	cut criteria	Number of ϕ after cut	ϕ efficiency
		100,000 events	
No KID	acceptance& tracking	54,287 ± 1,302 (2.4%)	54.3% ± 1.3%
one KID	KID $\geq$ 0.5	53,253 ± 679 (1.3%)	98.1% ± 2.7%
another KID	KID $\geq$ 0.2	51,194 ± 370 (0.7%)	96.1% ± 1.4%
$M_{K^+K^-}$	$\lvert M_{K^+K^-} - M_\phi \rvert \leq 10 MeV$	47,510 ± 343 (0.7%)	92.8% ± 0.9%
Total of 2 KID	2KID		94.3% ± 2.4%

();relative error.

Table 3.7: Efficiency of each cut for $M_{K^+K^-}$ (Monte Carlo)

In the same signal Monte Carlo sample, the D_s mass distribution becomes like Figure 3.30. Each distributions are histograms after each cuts. The number of D_s with each cut and the efficiencies of each cut which is applied in turns after other cuts are summarized in the Table 3.8. Then, we use double Gaussian as signal shape and liner function as background shape. This double Gaussian has $0.0042 GeV/c^2$ and $0.0129 GeV/c^2$ as two width, and 1.968 GeV/c^2 as nominal mass, and the ratio of two area is 0.53 from the result of fitting with all free parameter for $M_{\phi\pi}$ distribution without any cut. The efficiencies are well agree with real data.

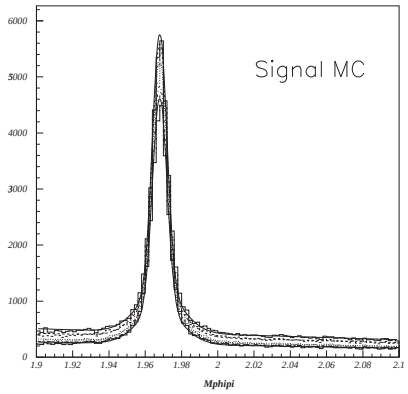


Figure 3.14: $M_{\phi\pi}$ (Monte Carlo)

cut	cut criteria	Number of D_s after cut	D_s efficiency
before		35,020 ± 213 (0.6%)	
πID	πID ≥ 0.1	34,360 ± 208 (0.6%)	98.1% ± 0.8%
D_s helicity angle	$\lvert cos\theta_h \rvert \geq 0.35$	32,950 ± 200 (0.6%)	95.9% ± 0.8%
B decay angle	$cos\theta_d \leq 0.85$	30,160 ± 191 (0.6%)	91.5% ± 0.8%
Event shape	R2 ≤ 0.3	28,780 ± 187 (0.7%)	95.4% ± 0.9%
Total of $M_{\phi\pi}$ cut	all		82.2% ± 0.8%

();relative error.

Table 3.8: Efficiency of each cut for $M_{\phi\pi}$ (Monte Carlo)

Table 3.9 is comparison of the efficiencies between real data and signal Monte Carlo sample. The values are agree with each other within those error, therefore we use Monte Carlo sample when estimate the momentum dependence of efficiency.

cut	on-resonance data	$B\overline{B}$ data	Signal Monte Carlo
2KID	89.0%±0.5%	85.9%±6.4%	94.3%±2.4%
$M_{\phi\pi}$ without R2 cut	85.0%±0.8%	84.6%±4.1%	86.1%±0.7%
R2 cut		95.6%±3.6%	95.4%±0.9%

Table 3.9: Comparison of efficiency between real data and MC

3.2 Measurement of $B \rightarrow X_c\ D_s$ braching fraction

As a first step, we measure the branching fraction of $B \rightarrow X_c\ D_s$ decay. Figure 3.15 is D_s momentum spectrum with 0.03 of bin width for the scaled momentum x. Black circle is on-resonance data of 29.1 fb^{-1}, and white circle is off-resonance data of 3.0 fb^{-1} which is normalized to on-resonance data with scale facter 9.6. We count each D_s number with bin by bin fitting for $M_{\phi\pi}$ distribution. We use double Gaussian as signal shape, and liner function as background shape with all bin.

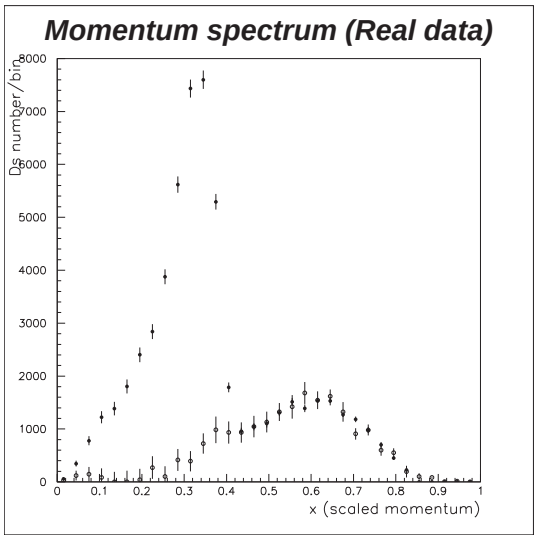


Figure 3.15: D_s momentum spectrum with $R_2 \geq 0.4$ (Real Data)

Figure 3.16 is the D_s momentum spectrum which subtracted off-resonance data from on-resonance data with 9.6 as the scale factor. The yield are 38,210 events in the region of $0 \leq$ x < 0.42 for the $b \rightarrow c\ D_s$ decay.

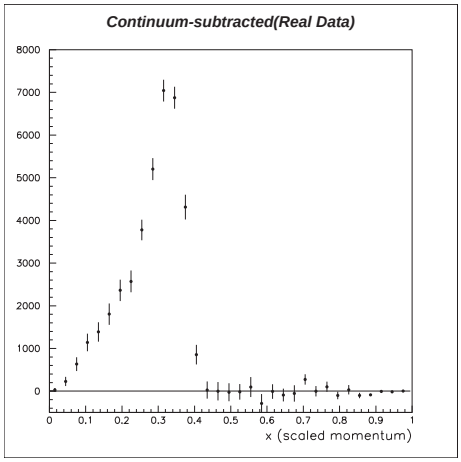


Figure 3.16: D_s momentum spectrum for $B\overline{B}$ (Real Data)

In order to estimate the efficiency of D_s reconstruction for each bin of momentum, we use four Monte Carlo event samples : 400,000 events of $B^0 \rightarrow D_s^+ X_c$(black circle), 10,000 events of $B^0 \rightarrow D_s^+ a_1^-$(white circle), 10,000 events of $B^0 \rightarrow D_s^+ \rho^-$(black triangle), and 10,000 events of $B^0 \rightarrow D_s^+ \pi^-$(black square). Then, all D_s^+ is decayed into $\phi\pi^+$, and ϕ is decayed into K^+K^-. Figure 3.17. is the plots for Efficiency vs. Scaled momentum. We fit efficiency curve with second polynomial, and use the fitting result to correct momentum spectrum. Momentum spectrum with efficiency correction became like Figure 3.18, and the yield are 124,200 events in the signal region of $0 \leq x < 0.42$.

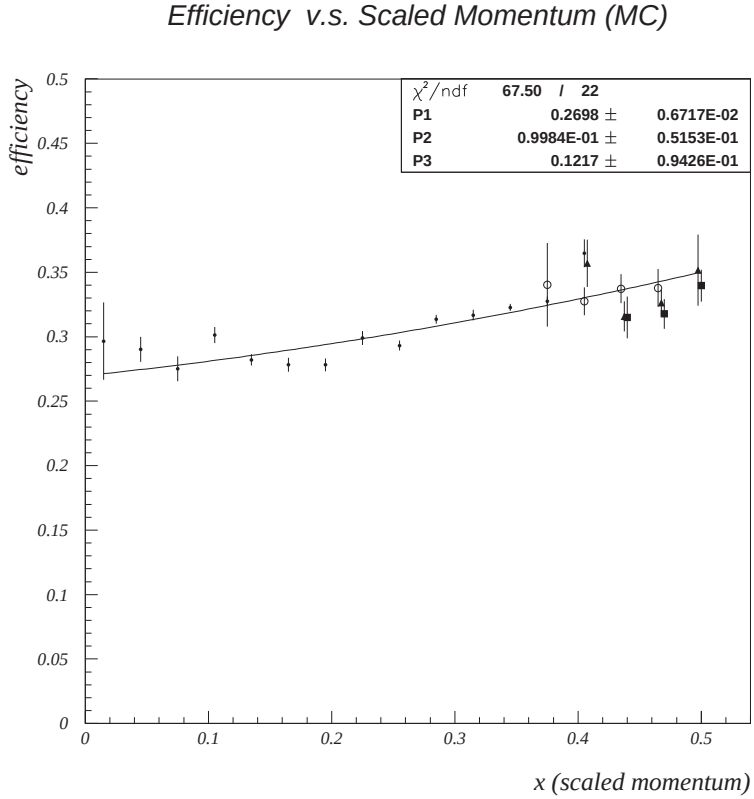


Figure 3.17: Efficiency vs. Scaled momentum (Monte Carlo)

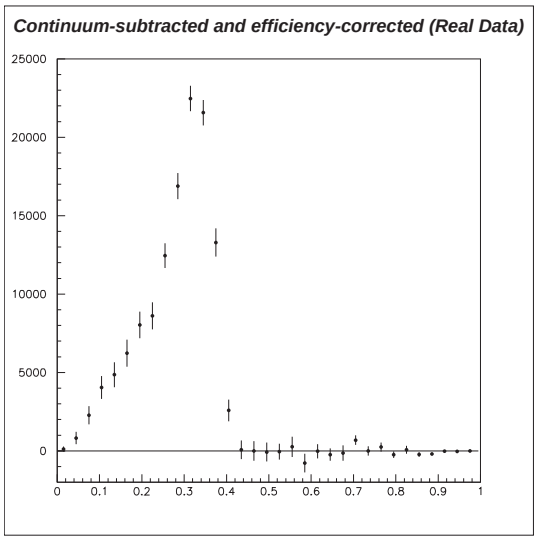


Figure 3.18: D_s momentum spectrum with efficiency correction (Real Data)

The branching fraction of $B \to X_c\ D_s$ decay is calculated with next number; number of $B\overline{B}$ events is 3.19×10^6 , branching fraction of $D_s^{\pm} \to \phi\ \pi^{\pm}$ decay is 0.036, and branching fraction of $\phi \to K^+\ K^-$ decay is 0.49. The branching fraction of $B \to X_c\ D_s$ decay is measured as

$$\frac{124,200(D_s\ number)}{3.17 \times 10^6(B\overline{B}) \times 2(2B\ pair) \times 0.036(Br(D_s^{\pm} \to \phi\ \pi^{\pm})) \times 0.49(Br(\phi \to K^+\ K^-))}$$

$\sim$ (11.0 $\pm$ 0.3 (stat) $\pm$ 2.9 (syst)) %

where, statistical error is calculated with adding error of each bin in quadrature, and list of systematic error is shown on Table 3.10.

Systematics	error
Br($D_s^{\pm} \to \phi\ \pi^{\pm}$)	25%
Br($\phi \to K^+\ K^-$)	1.4%
Number of $B\overline{B}$	1%
3 tracking	3.5%
M_{KK} cut	1.0%
2 KID	0.5%
πID	1.0%
D_s helicity angle	0.9%
B decay angle	0.8%
R2	0.9%
Fitting for efficiency curve	6.2%
$B \to X_u\ D_s$ decay	0.8%
Other than Br($D_s^{\pm} \to \phi\ \pi^{\pm}$)	7.7%
Total	26.2%

Table 3.10: Summary of systematic error

Most dominant systematic error is 25% for branching fraction of $D_s^{\pm} \to \phi\ \pi^{\pm}$. The error of $B\overline{B}$ number is estimated as 1% [17]. The error of tracking is estimated as 2% per high momentum particle. We use the result of Monte Carlo as the error of $M_{K^+K^-}$ mass window because the shape is decided from Monte Carlo study. We use the relative error of the efficiency for KID, for πID and for 2 angle cuts in on-resonance data as systematic error because of high statistics. For the R2 cut, we use the relative error of the efficiency in Monte Carlo because it's different between on-resonance and off-resonance data. The error of fitting for efficiency curve is estimated with the error of the used parameter and error matrix of fitting result. The contamination of $B \to X_u\ D_s$ decay is estimated as 1.0% as explain later (Appendix). Total systematic error become 26.2%, and other than the error of $D_s^{\pm} \to \phi\ \pi^{\pm}$ decay is 7.7%.

3.3 Search for $b \rightarrow u\ D_s$ decay

3.3.1 D_s momentum spectrum

The momentum spectrum of D_s without event shape cut is shown on Figure 3.19. Black circle is on-resonance data of 29.1 fb^{-1}, and white circle is off-resonance data of 3.0 fb^{-1} that is normalized to on-resonance data with the number of D_s which is over 0.5 with x (the endpoint of D_s in $b \rightarrow u\ D_s$ decay). We count the number of D_s with bin by bin fitting for each $M_{\phi\pi}$ distribution that is correspond to each width (x=0.01). We use the double Gaussian as signal shape and liner function as background shape. At that time, first, we make use of 7 all parameter with free parameter that consists of 2 area, 2 width and 1 mean of double Gaussian, and 2 degree of liner function, after fitting with free parameter, we fix one width as $0.0036 GeV/c^2$, another width as $0.0104 GeV/c^2$, and the ratio of 2 area as 0.48. That means only one area is fitted as free parameter in signal shape. Errors of total area are plotted with error bar in the distribution.

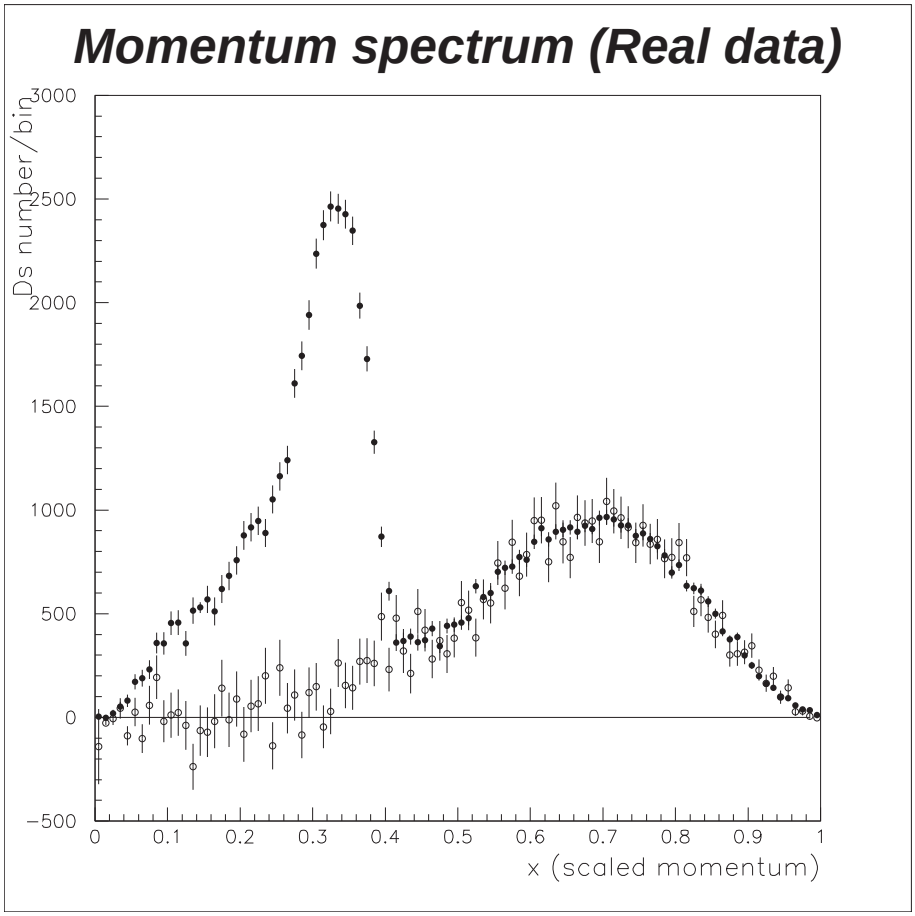


Figure 3.19: D_s momentum spectrum without event shape cut (Real Data)

The resolution of x is shown on Figure 3.20. This distribution is the x difference between generated x of D_s and reconstructed x of D_s in signal Monte Carlo ($B^0 \rightarrow X_c\ D_s^+$, $D_s^+ \rightarrow \phi\ \pi^+$ and $\phi \rightarrow K^+\ K^-$). The resolution is estimated as σ_x equals 0.0022 by the weighted mean of 2 width after fitting with double Gaussian.

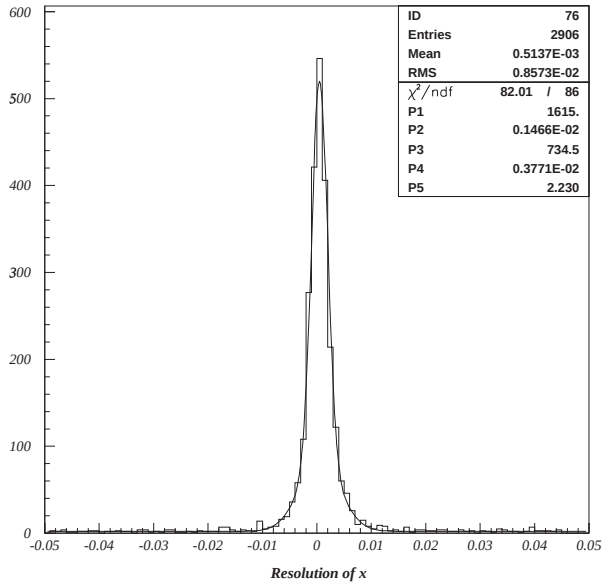


Figure 3.20: Resolution of x (Monte Carlo)

The distributions of $M_{\phi\pi}$ are shown on Figure 3.21 or Figure 3.22, those are for representative 8 bins such as, $0.1 \leq$ x ≤ 0.11, $0.2 \leq$ x ≤ 0.21, $\cdots$, $0.8 \leq$ x ≤ 0.81 with on-resonance data of 29.1 fb^{-1} or off-resonance data of 3.0 fb^{-1}.

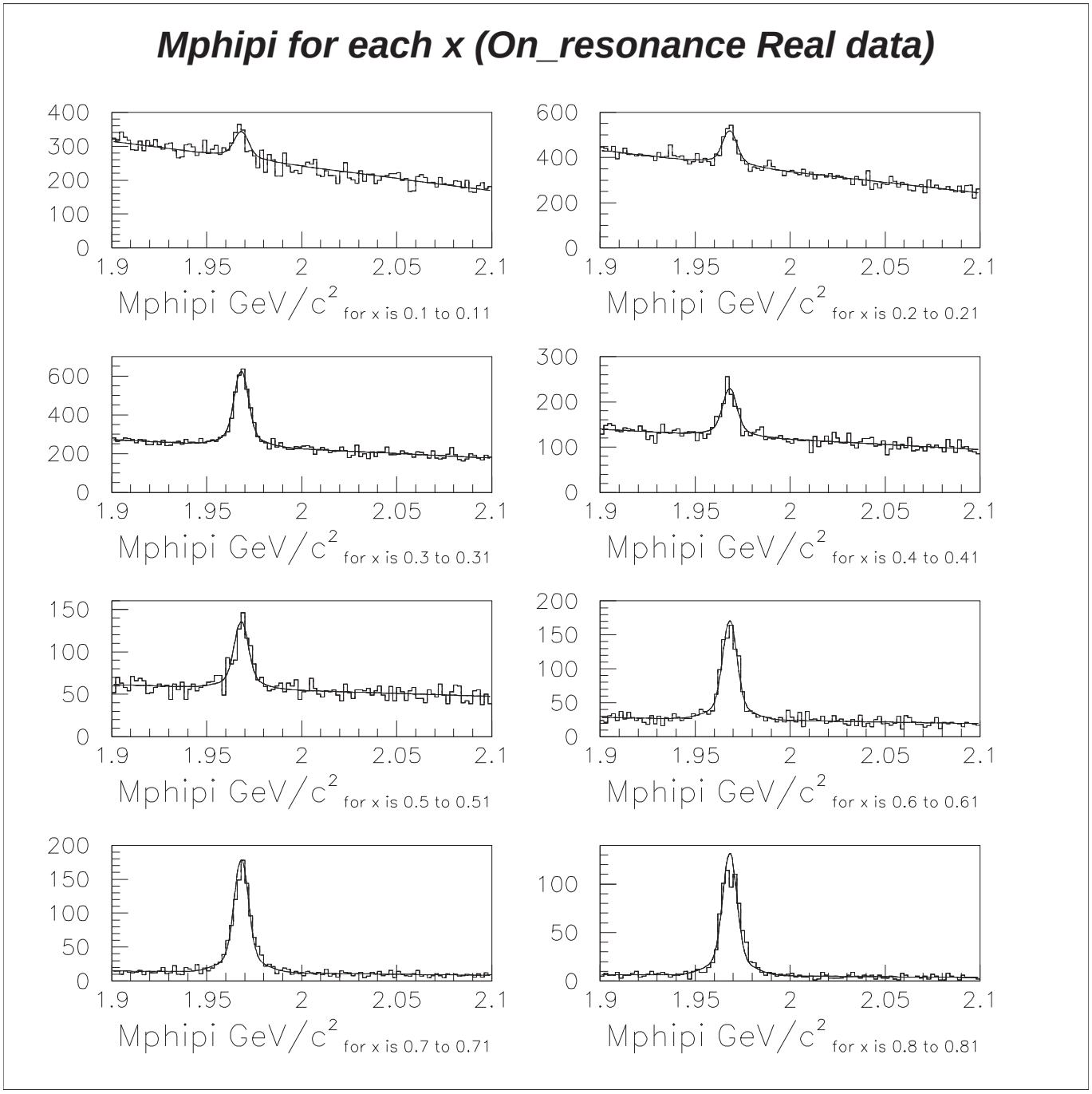


Figure 3.21: Representative bin's $M_{\phi\pi}$ distributions with on-resonance data (Real Data)

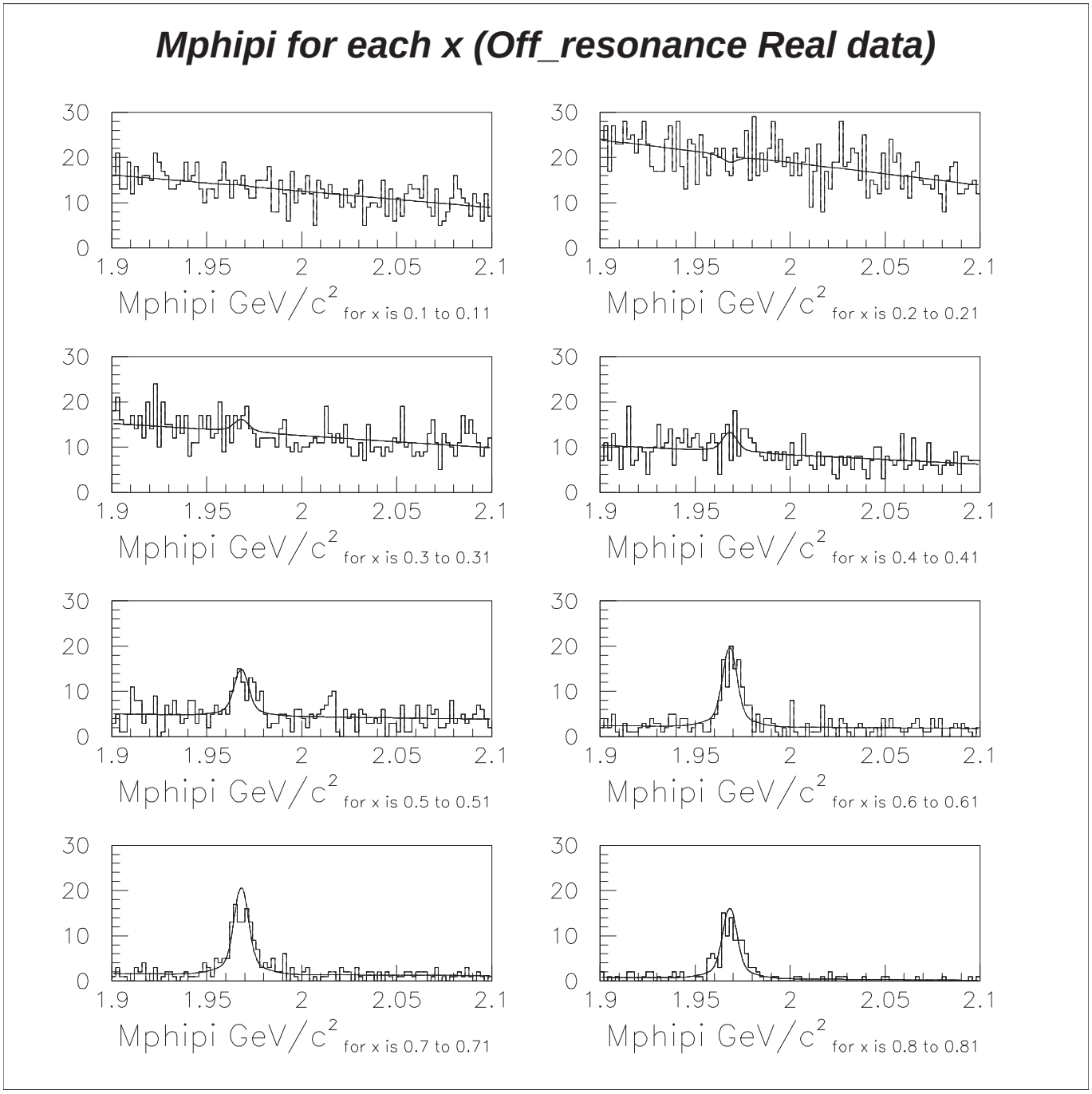


Figure 3.22: Representative bin's $M_{\phi\pi}$ distributions with off-resonance data (Real Data)

Figure 3.23 is the distribution after subtraction of off-resonance data from on-resonance data. Histogram is Monte Carlo sample of the decay chain; $B^0 \rightarrow X_c\ D_s^+$, $D_s^+ \rightarrow \phi\ \pi^+$ and $\phi \rightarrow K^+\ K^-$, which has 400,000 event that is correspond to $100 fb^{-1}$ and it is normalized to real data. The discrepancy between real data and Monte Carlo around $0.15 \leq x \leq 0.3$ can be considered as that this MC sample don't contain $B \rightarrow D^{**}\ D_s$ decay. The region of $0.42 \leq x < 0.50$ is defined as signal D_s region which come from $b \rightarrow u\ D_s$ transition. In this distribution, the D_s spillover from $b \rightarrow c\ D_s$ transition is estimated as 5.6 ± 6.5 events with Monte Carlo sample. The number of D_s in real data are 37,497 $\pm$ 810 (stat) for the $b \rightarrow c\ D_s$ region of $0 \leq x < 0.42$. The number of signal D_s are 351 $\pm$ 297(stat) for the signal region of $0.42 \leq x < 0.50$, then, statistical error is calculated with adding error of each bin in quadrature.

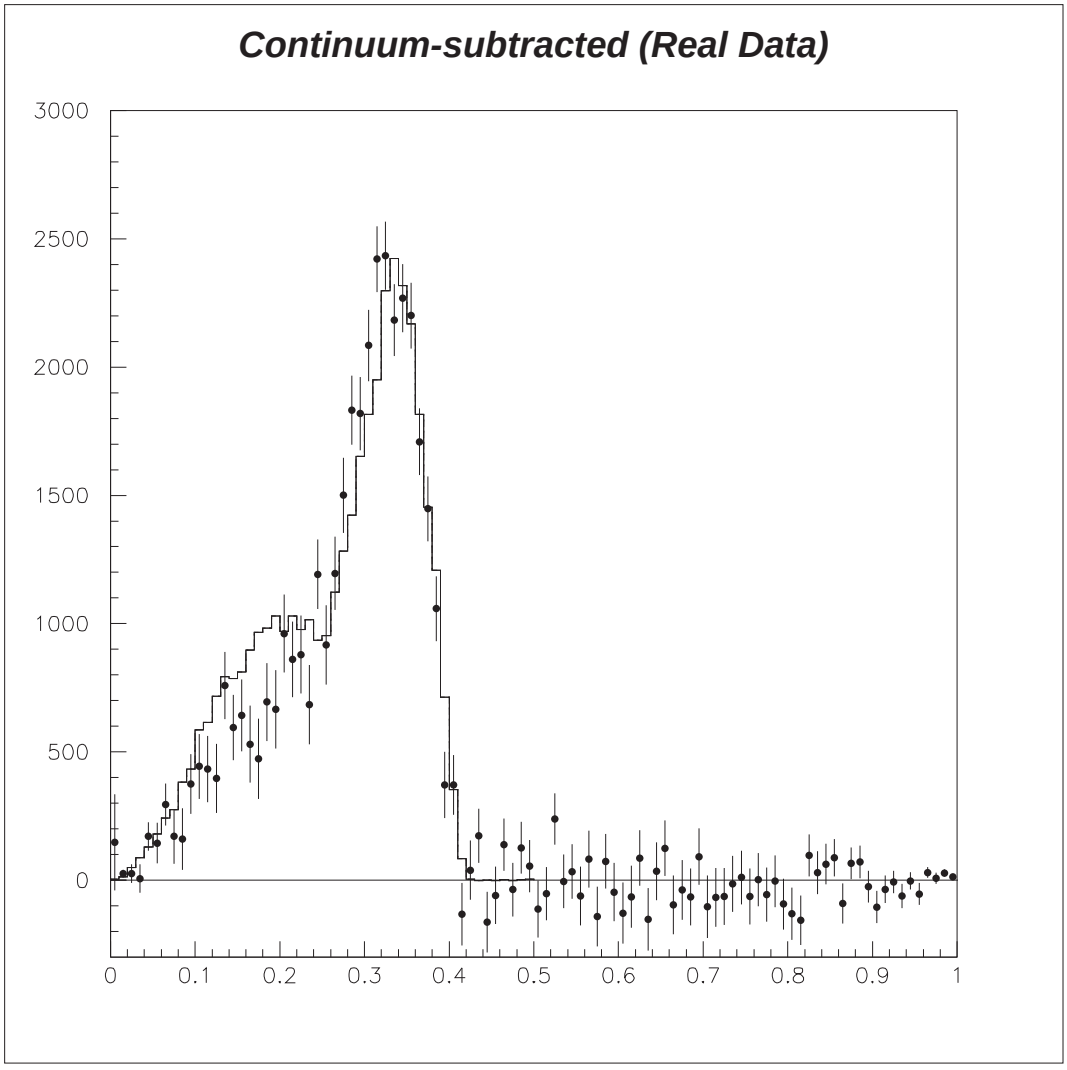


Figure 3.23: D_s momentum spectrum after subtraction of off-resonance data without event shape cut (Real Data)

3.3.2 Continuum suppression

Continuum background is largest background source. And after subtraction of off-resonance data from on-resonance data, the error of off-resonance becomes dominant error source because of less statistics. Therefore, we need suppress continuum events, and first, we try to use Monte Carlo continuum events for the subtraction. Histogram in Figure 3.24 is 26,000,000 events of continuum process in Monte Carlo which are correspond to $8fb^{-1}$ and normalized to real data: They consist of 16,000,000 events of $e^+e^- \rightarrow u\overline{u}$, $d\overline{d}$ or $s\overline{s}$, and 10,000,000 events of $e^+e^- \rightarrow c\overline{c}$. D_s spectrum of continuum Monte Carlo doesn't match in high momentum region. Therefore, we don't use continuum Monte Carlo.

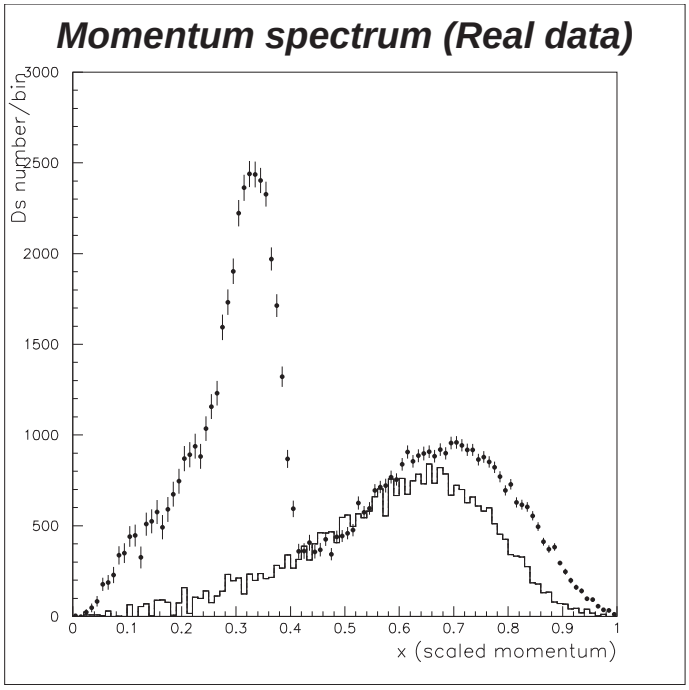


Figure 3.24: D_s momentum spectrum with continuum Monte Carlo

R_2 cut Second, we try the event shape variable cut: R_2 is less than 0.3. Then, the efficiency for signal events is 87%, otherwise 57% of continuum background is suppressed. The momentum spectrum is like below. Black circle is $29.1 fb^{-1}$ of on-resonance data, and white circle is off-resonance data of 3.0 fb^{-1} which are normalized to on-resonance data with D_s number whose x is over 0.5.

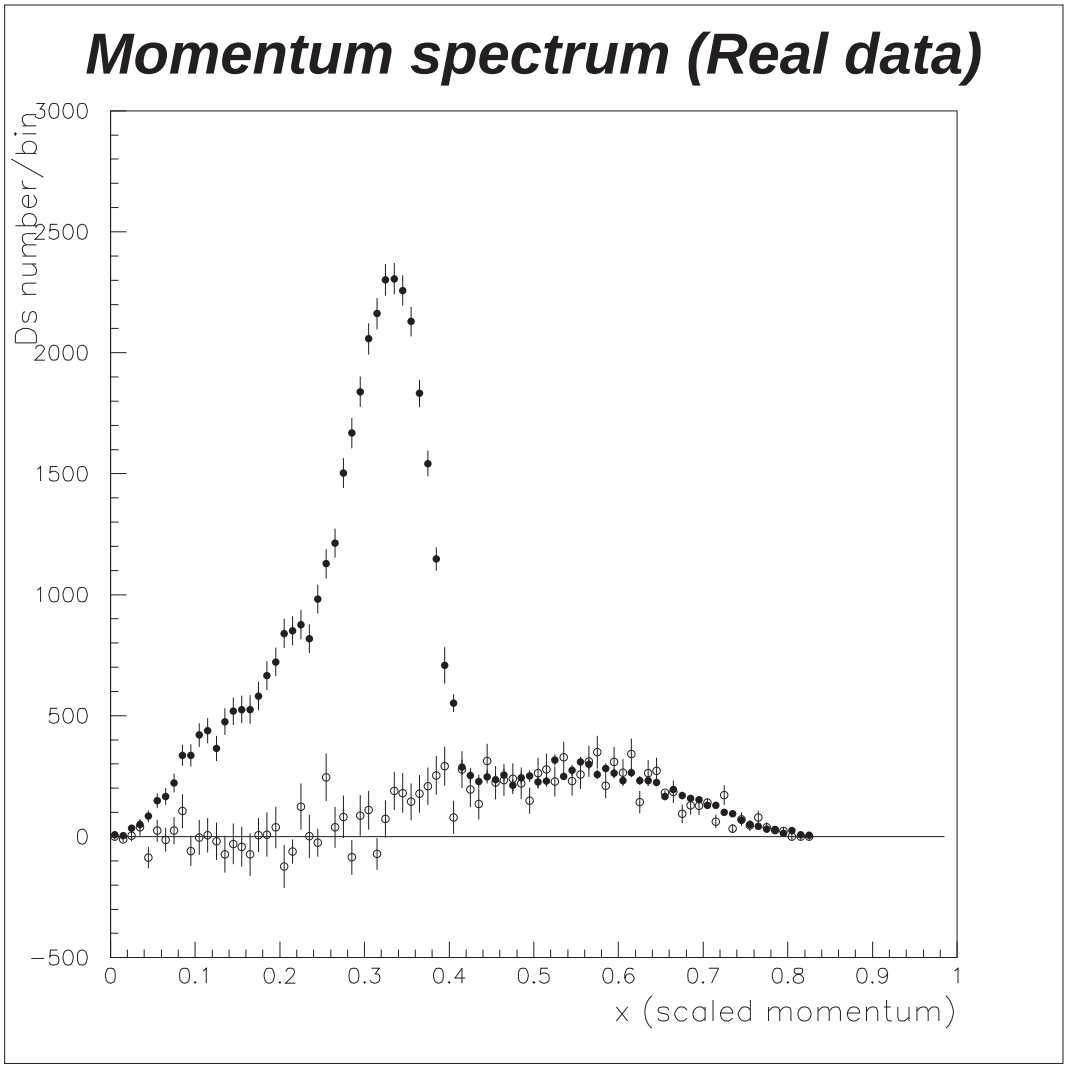


Figure 3.25: D_s momentum spectrum with R_2 cut (Real Data)

Figure 3.26 is momentum spectrum after subtraction of off-resonance data from on-resonance data with R_2 cut. The number of D_s is 35,583 $\pm$ 580 (stat) events for the $b \rightarrow c\ D_s$ region of $0 \leq$ x < 0.42, and 221 $\pm$ 201 (stat) events for the signal region of $0.42 \leq$ x < 0.50. We use these number when we calculate $|V_{ub}|$ value.

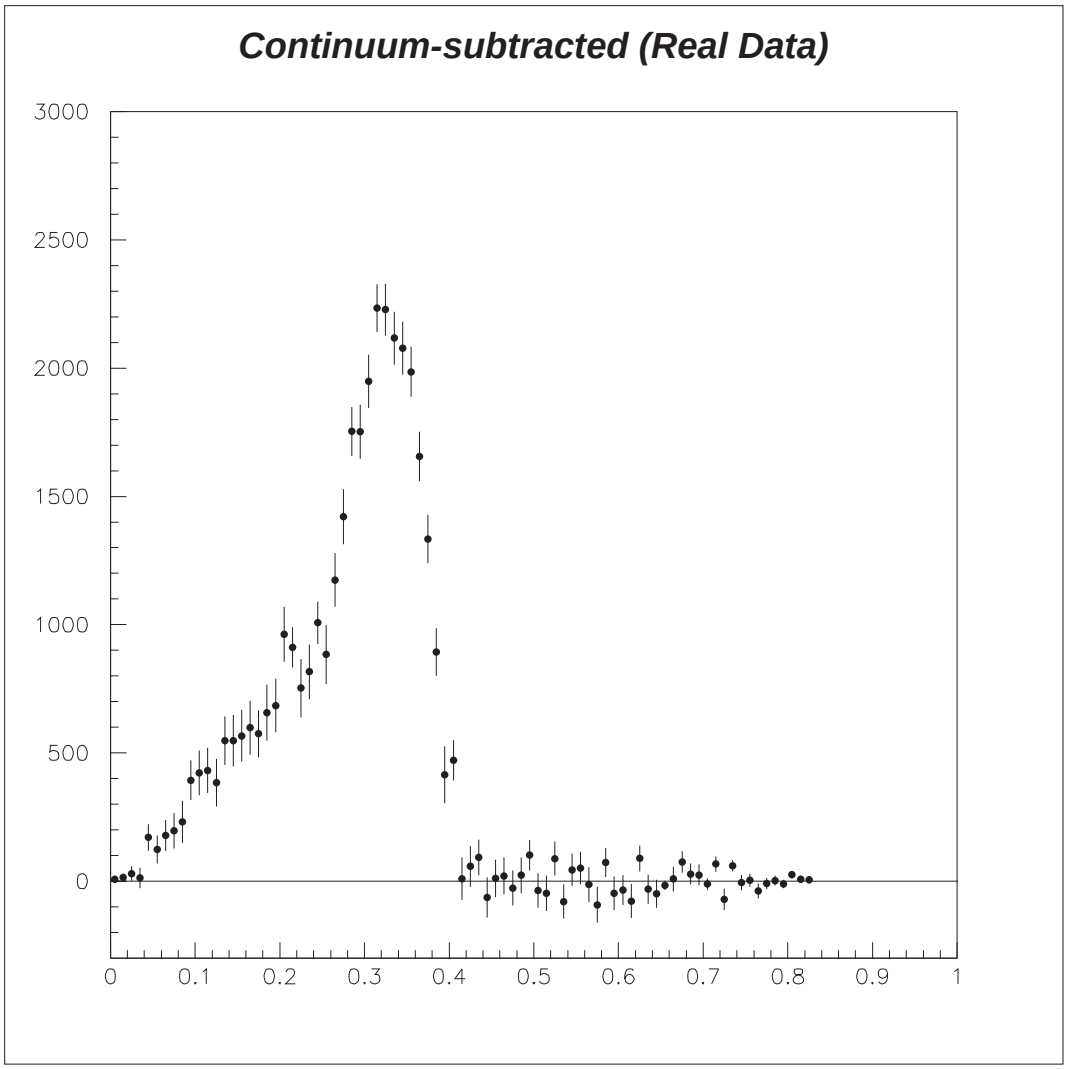


Figure 3.26: D_s momentum spectrum after subtraction of off-resonance data with R_2 cut (Real Data)

Next, we try to subtract off-resonance data with fitting result. Then, we use a generalized Peterson function F_{gen}(x) which is described like below [18]

$$F_{gen}(x) = \frac{P_1}{x[1 - P_3/x - P_2/(P_4 - x)]^2}$$

where, P_1, P_2, P_2, and P_4 are free parameter. And, original Peterson function F_{ori}(x) is a function of one parameter as

$$F_{ori}(x) = \frac{1}{x[1 - 1/x - P_1/(1 - x)]^2}$$

The fitting result is shown on Figure 3.27. Then, we use 116 points which consist of all 83 data points of off-resonance, and 33 on-resonance data points whose x are over 0.5.

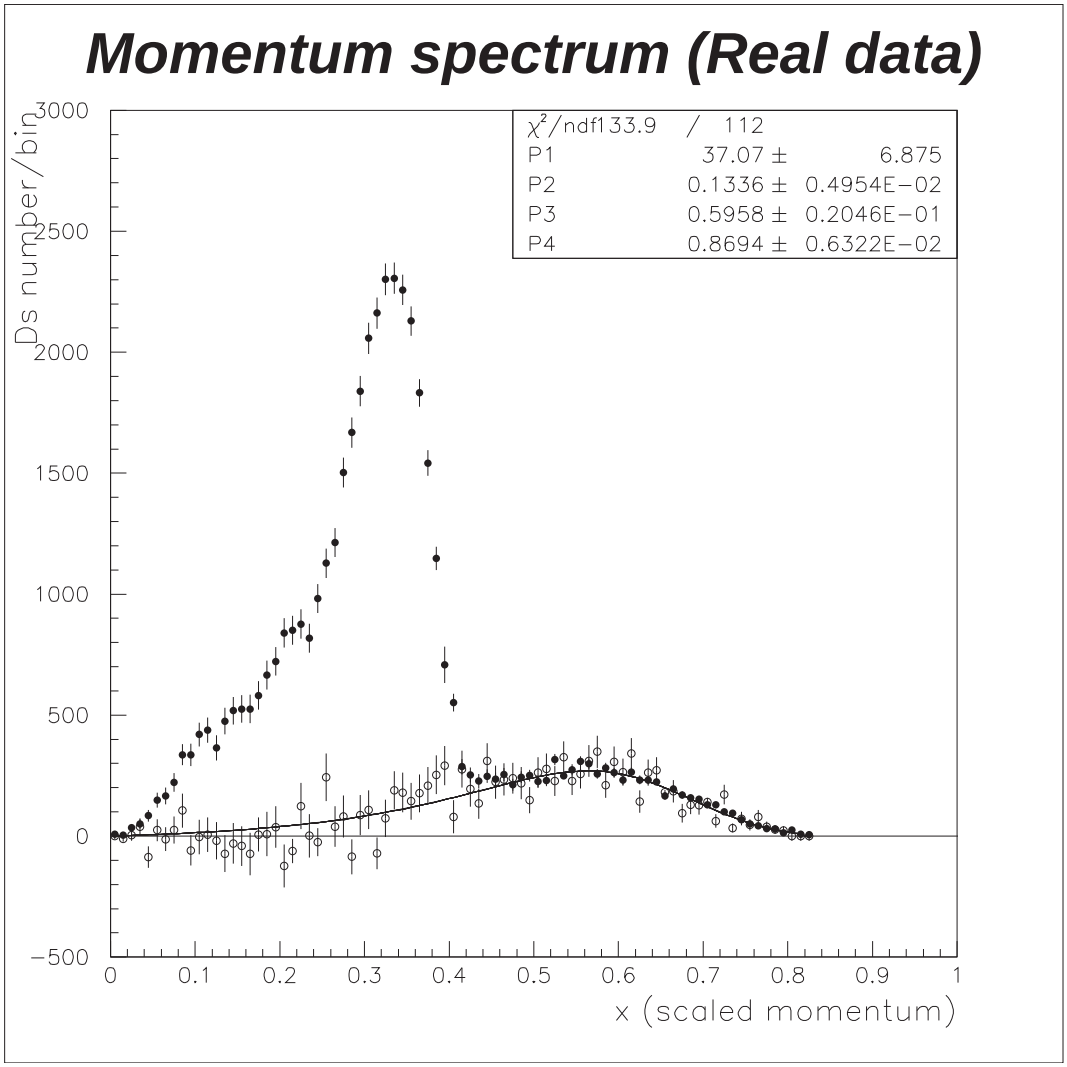


Figure 3.27: D_s momentum spectrum with fitting for off-resonance (Real Data)

Figure 3.28 is the momentum spectrum after subtraction of off-resonance data from on-resonance data with using the fitting result. The number of D_s is 35,163 $\pm$ 350 (stat) events for the $b \rightarrow c$ D_s region of $0 \leq$ x < 0.42, and 215 $\pm$ 136 events for the signal region of $0.42 \leq$ x < 0.50. The error is calculated from statistical error of on-resonance and fitting error of P_1 parameter. We consider that these number have large systematics error because some bias looks to exist in off-resonance data around 0.38 with x that is near the signal region.

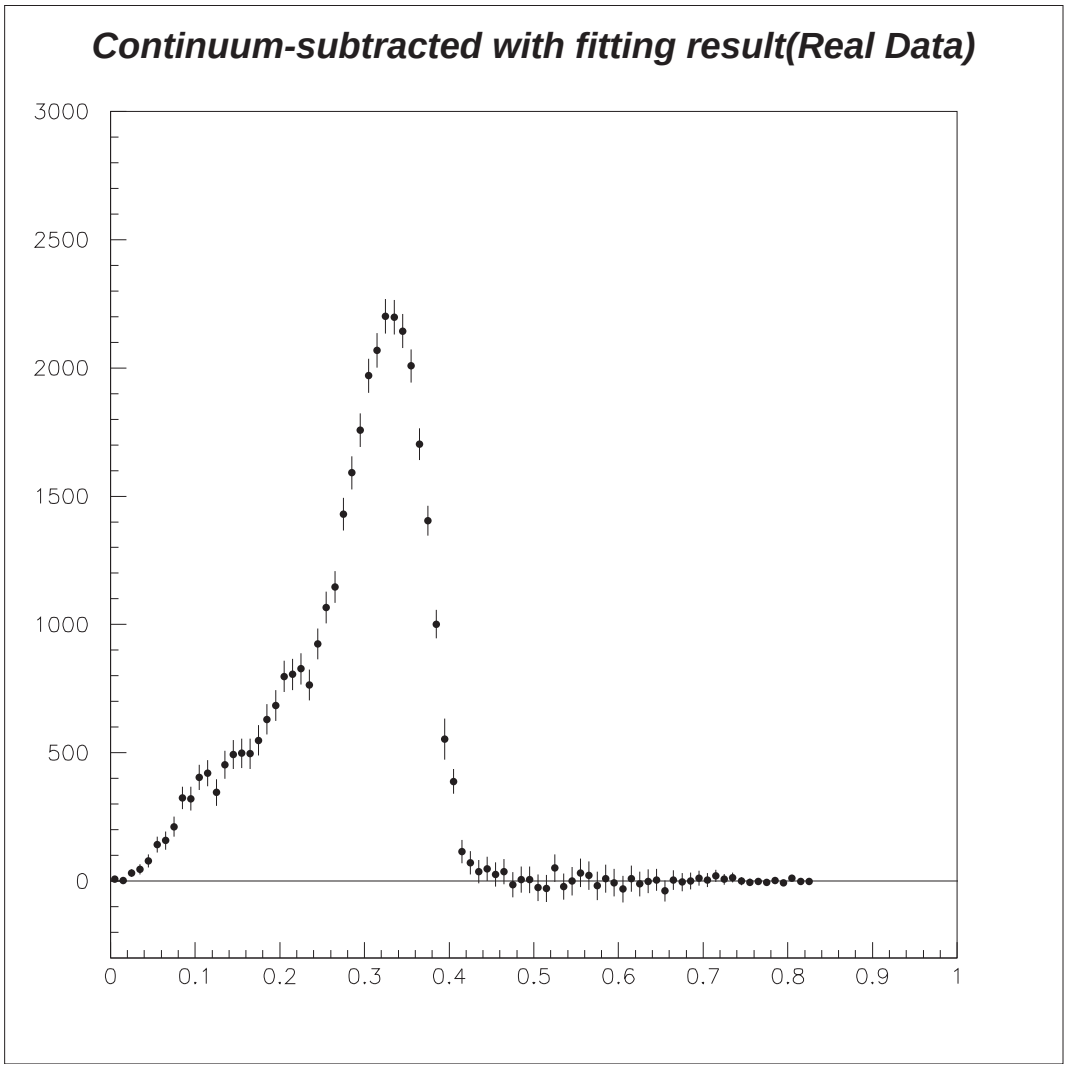


Figure 3.28: D_s momentum spectrum after subtraction of off-resonance data with fitting for off-resonance (Real Data)

Fisher discriminant We try to discriminate continuum events with Fisher discriminant method [19]. Fisher discriminant finds most discriminative line in the space of dimensions for parameter's number. In other wards, events with some correlative parameters are projected into one parameter to maximize discrimination. In this analysis, Fisher discriminant use R_2 and other 12 parameters; $|cos\theta_{D_s-beam}|$ is the angle between beam axis and D_s direction in $\Upsilon(4S)$ frame, $|cos\theta^*_{D_s-otherThrust}|$ is the angle between thrust axis of event (defined as below) and D_s direction in $\Upsilon(4S)$ frame, and other 10 parameters are energy sum of charged particle and photon energy around D_s direction or D_s opposite direction in $\Upsilon(4S)$ frame with the cone angle, 10^o, 30^o, 45^o, 60^o, or 90^o after removing D_s daughter particles (that is called virtual calorimeter).

Thrust axis of event is defined to maximize thrust function,

$$T(n) = \sum_{i=1}^{N} |P_{i_\parallel}| / \sum_{i=1}^{N} |P_i|,$$

where n is one direction, N is the number of particle (in this analysis, we use charged particles and photon) in the event, $|P_i|$ is the momentum of particle, $P_{i_\parallel}$ is the momentum of parallel component to the n direction.

Parameters that is used for Fisher discriminant are as follows

- R_2
- $|cos\theta_{D_s-beam}|$
- $|cos\theta^*_{D_s-otherThrust}|$
- $E_{10^o}, E_{30^o}, E_{45^o}, E_{60^o}, E_{90^o}$, $E'_{10^o}, E'_{30^o}, E'_{45^o}, E'_{60^o}$, and E'_{90^o} (Virtual calorimeter)

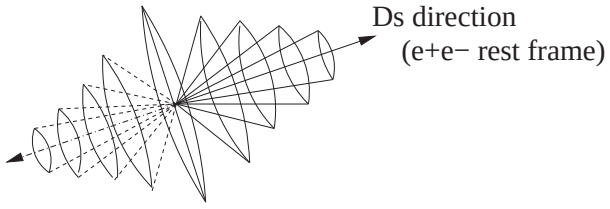


Figure 3.29: Virtual calorimeter

Each parameter's distribution with Monte Carlo background continuum event (hatched histogram) and Monte Carlo signal decay chain, $B^0 \rightarrow a_1^-\ D_s^+$, $D_s^+ \rightarrow \phi\ \pi^+$ and $\phi \rightarrow K^+\ K^-$ (open histogram) is Figure 3.31, Figure 3.32 and Figure 3.33. Then, each parameter is calculated when D_s are reconstructed within the mass window 1.95 $GeV/c^2 \leq M_{\phi\pi} \leq 1.99\ GeV/c^2$ as Figure 3.30.

Finally, all parameters are combined into Fisher discriminant

$$F = \alpha R_2 + \beta|cos\theta_{D_s-beam}| + \gamma|cos\theta^*_{D_s-otherThrust}| + \sum_{i=1}^{10} \delta_i E_i$$

where, E_i(i=1,2,3,...,10) are 10 parameters of $E_{10^o}, E_{30^o}, E_{45^o}, E_{60^o}, E_{90^o}, E'_{10^o}, E'_{30^o}, E'_{45^o}, E'_{60^o}$, and E'_{90^o}. And, α, β, γ, and δ_i are coefficient of Fisher discriminant which are optimized to maximize discrimination.

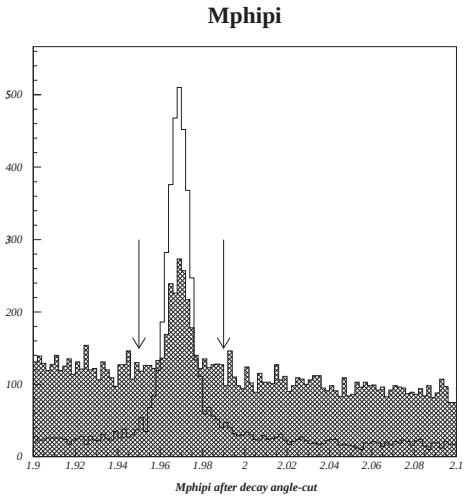


Figure 3.30: D_s mass window (Monte Carlo)

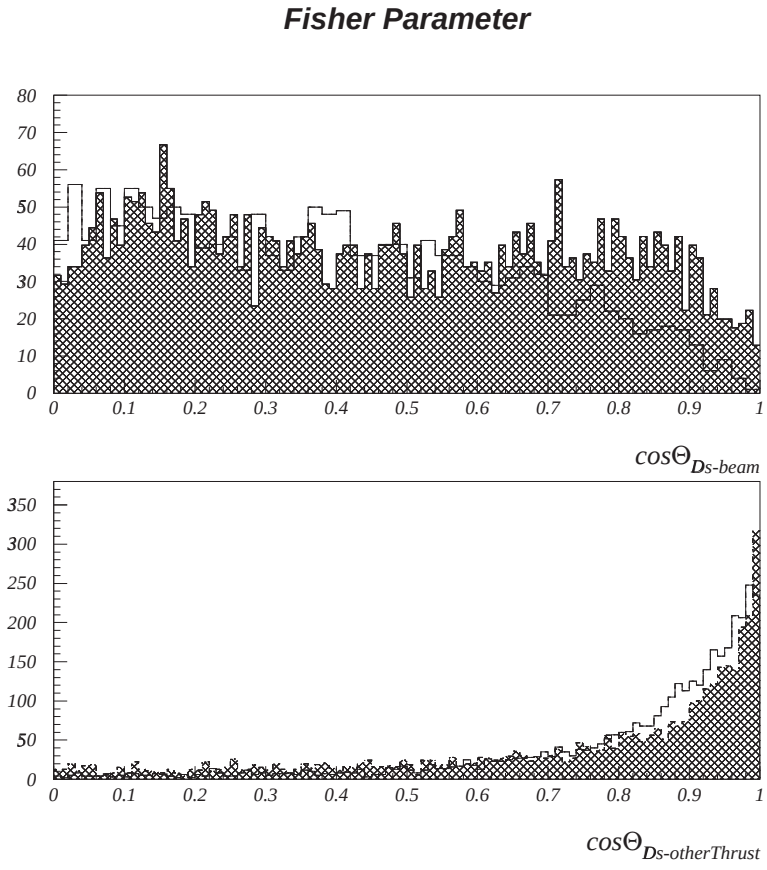


Figure 3.31: Fisher's angle parameter (Monte Carlo)

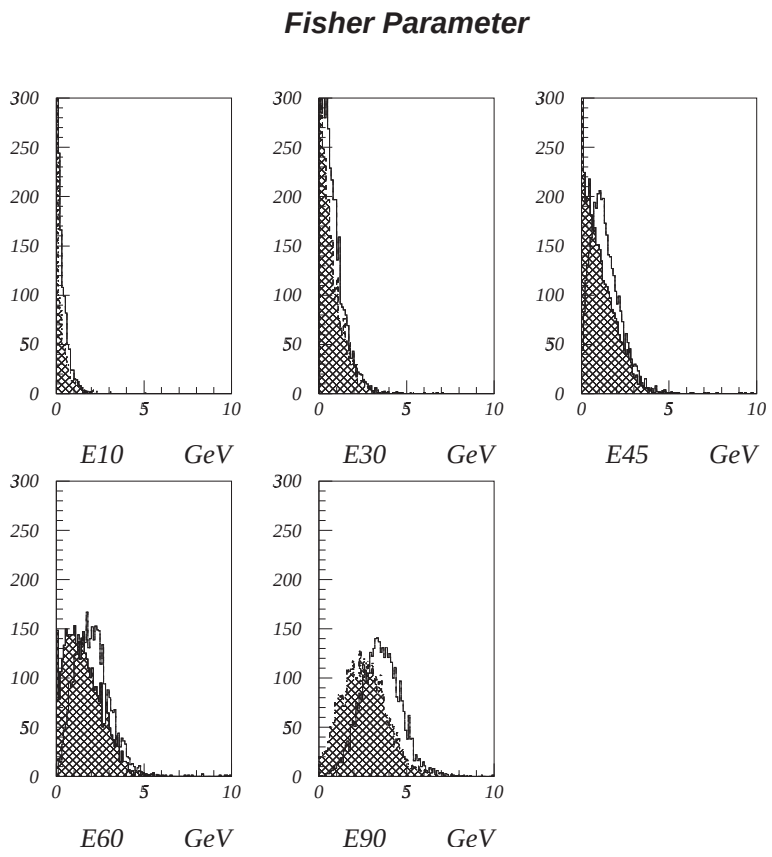


Figure 3.32: Fisher's energy parameter around D_s direction (Monte Carlo)

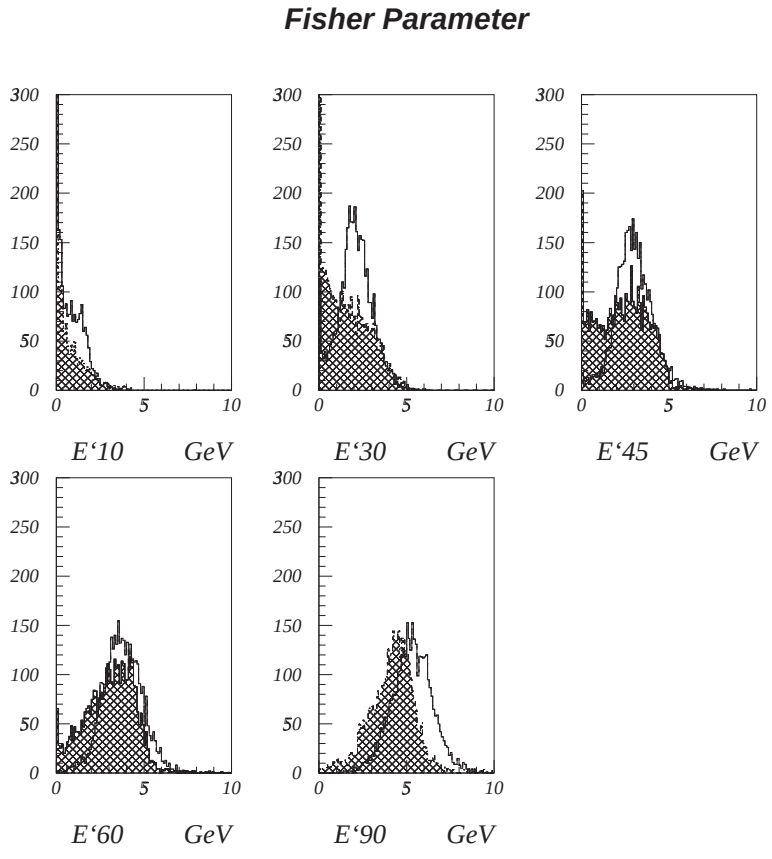


Figure 3.33: Fisher's energy parameter around D_s opposite direction (Monte Carlo)

Figure 3.34 is the distribution of Fisher Variable for signal Monte Carlo (dashed histogram), for continuum background Monte Carlo (solid histogram) and for off-resonance data (filled circle). We use that Fisher variable is greater than 0 as cut value. Then, the signal efficiency is 96%, otherwise 57% of background continuum events is suppressed .

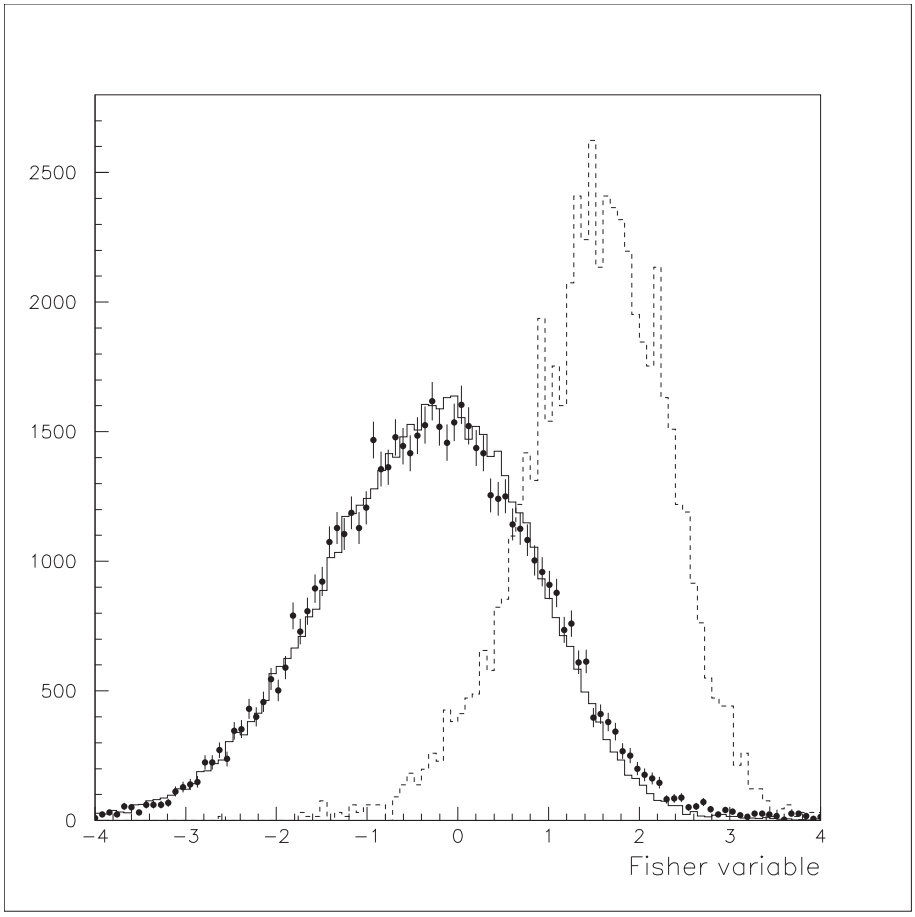


Figure 3.34: Fisher variable

Momentum spectrum with this cut is Figure 3.35. Black circle is $29.1 fb^{-1}$ on-resonance data, and white circle is $3.0 fb^{-1}$ off-resonance data which are normalized to on-resonance data with D_s number whose x is over 0.5. Figure 3.36 is momentum spectrum after subtraction of off-resonance data from on-resonance data with fisher variable cut. The number of D_s is 33,295 $\pm$ 630 (stat) events for the $b \rightarrow c\ D_s$ region of $0 \leq$ x < 0.42, and -5 $\pm$ 187 (stat) events for the signal region of $0.42 \leq$ x < 0.50.

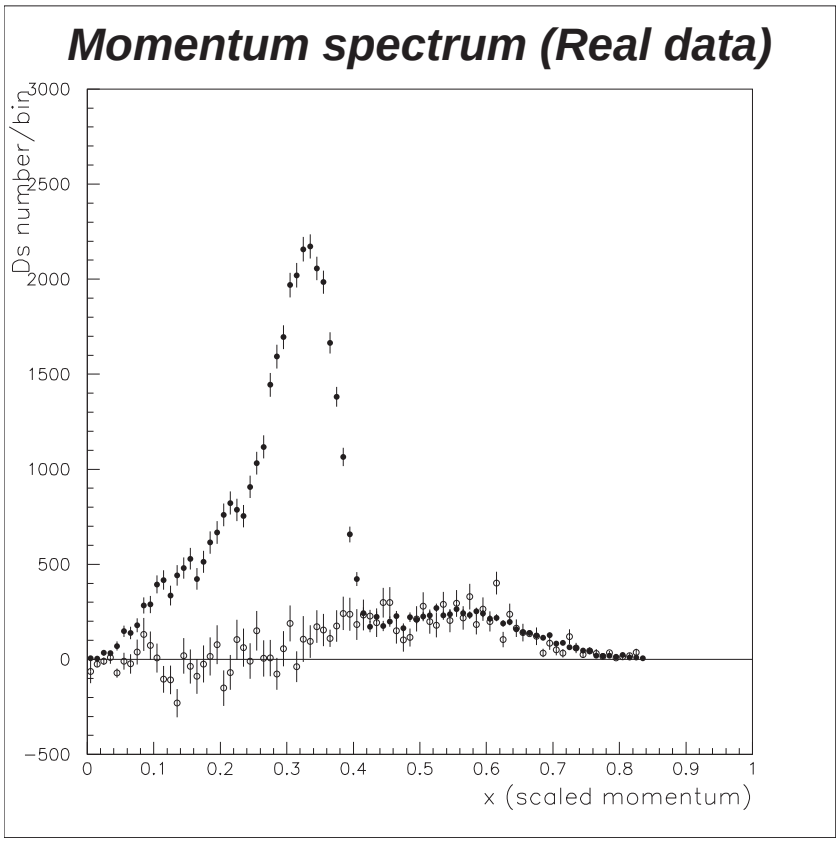


Figure 3.35: D_s momentum spectrum with Fisher variable ≥ 0 (Real Data)

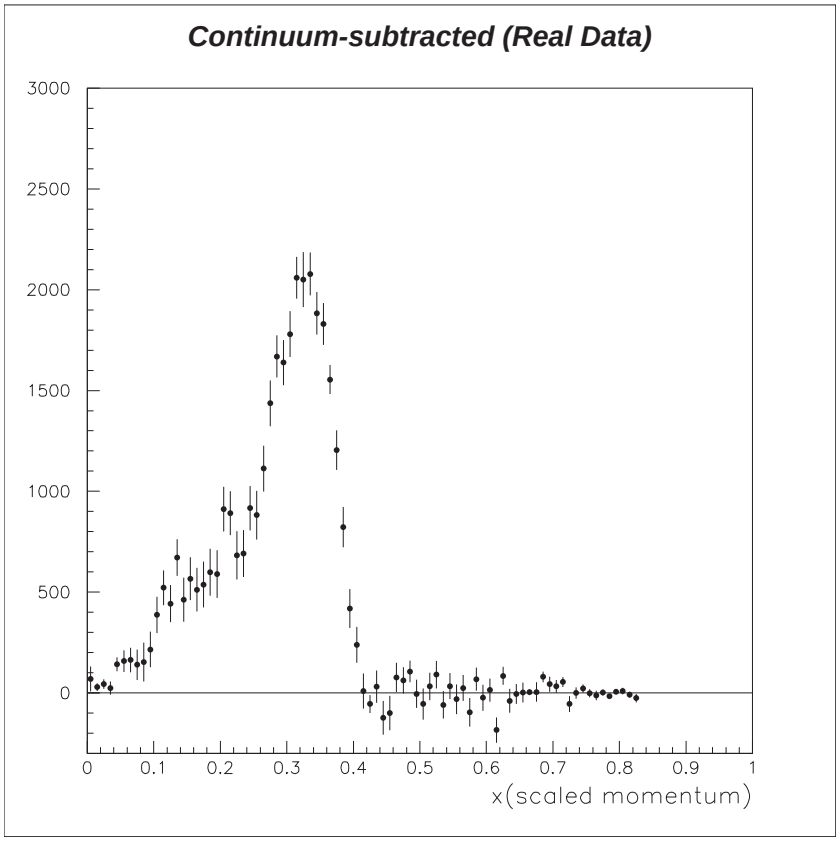


Figure 3.36: After subtraction of off-resonance data with Fisher variable ≥ 0 (Real Data)

We try another cut value that is greater than 1.0. Then, the signal efficiency is 73%, otherwise 89% of background continuum events is suppressed. Figure 3.37 is the momentum spectrum of $29.1 fb^{-1}$ on-resonance data with this cut (bin width is 0.03). The yield for on-resonance data is 389 ± 41 events in the signal region of $0.42 \leq x < 0.51$.

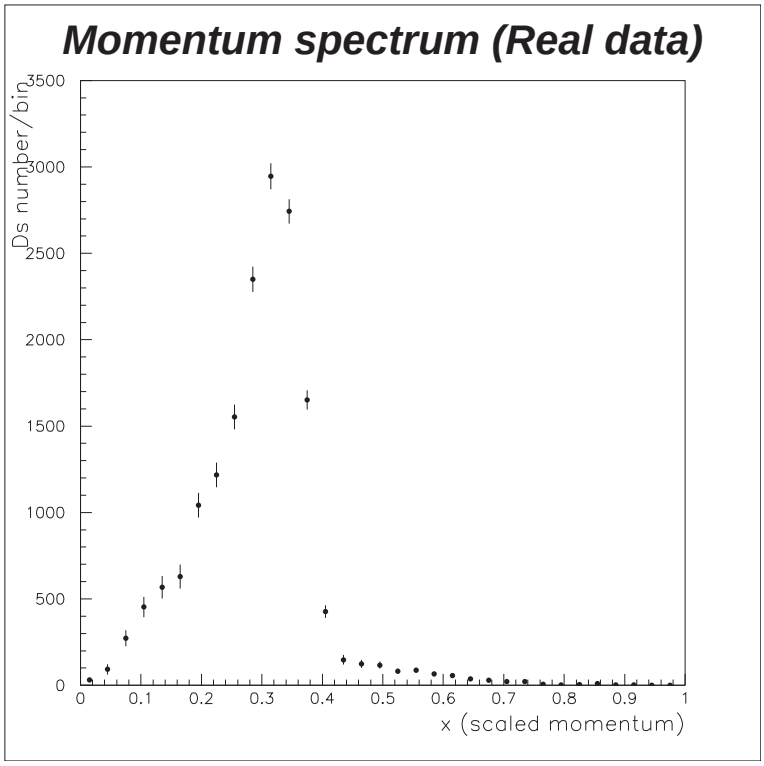


Figure 3.37: D_s momentum spectrum of on-resonance with Fisher variable ≥ 1.0 (Real Data)

The yield with the same cut for off-resonance data is 36 ± 10 events in the signal region of $0.42 \leq x < 0.51$, then we use single bin of $M_{\phi\pi}$ distribution for counting D_s number of off-resonance data because of less statistics. Figure 3.38 is the $M_{\phi\pi}$ distribution in the momentum region of $0.42 \leq x < 0.51$ for off-resonance data. We use double Gaussian as signal shape, and liner function as background shape. Number of signal events that means after subtraction of off-resonance data from on-resonance data is calculated as 43 ± 108 events.

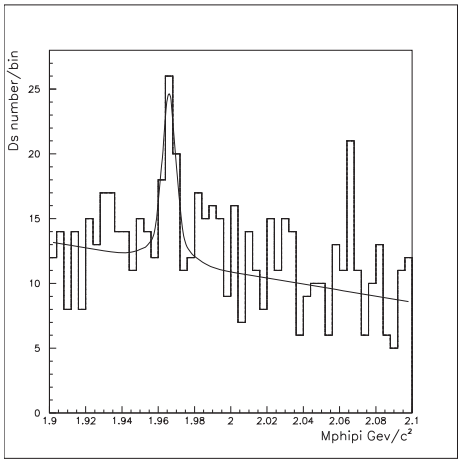


Figure 3.38: $M_{\phi\pi}$ in the signal region with Fisher variable ≥ 1.0 for $3.0 fb^{-1}$ of off-resonance (Real Data)

Lepton tagging Next method to suppress continuum events is lepton tagging. In the aim of tagging signal D_s with primary high momentum lepton from another B meson, opposite charge of D_s is required. We use particle identification that is greater than 0.9 for both electron or muon which means likelihood ratio. The requirement for momentum is greater than 1GeV in $\Upsilon(4S)$ frame. In addition to these cuts, we use the angle between D_s and lepton in $\Upsilon(4S)$ frame with the aim of continuum suppression. This $|cos\theta^*_{D_s-l}|$ become to be close 1 for jet like events, thus we require it to be less than 0.8. Each criteria are as follows

- $D_s^{\pm}$ and $l^{\mp}$ has opposite charge
- eID ≥ 0.9 or muID ≥ 0.9
- $|P_l^*| \geq 1$ GeV
- $|cos\theta^*_{D_s-l}| \leq 0.8$

Figure 3.39 is $M_{\phi\pi}$ distribution before tagging (without Fisher's cut) and after tagging for $4.4fb^{-1}$ of on-resonance data and $3.0fb^{-1}$ of off-resonance data (Left distribution is before tagging, right one is after tagging, upper is on-resonance data, and lower is off-resonance data).

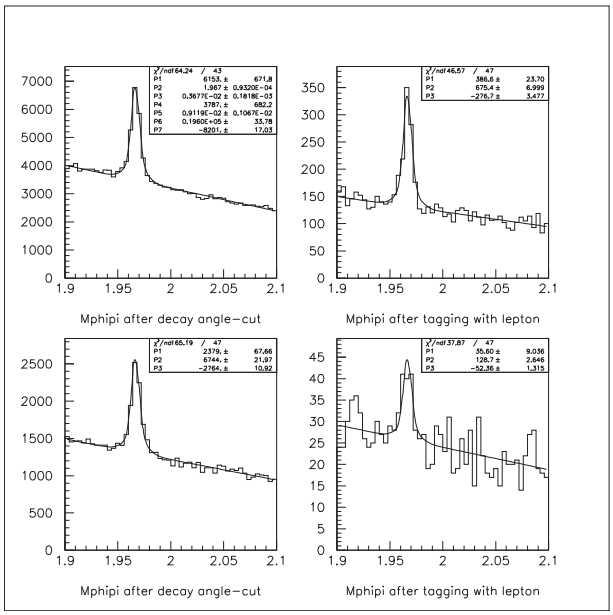


Figure 3.39: $M_{\phi\pi}$ comparison with lepton tagging (Real Data)

The efficiency only for this tagging in on-resonance of $4.4fb^{-1}$ and off-resonance of $3.0fb^{-1}$ are as below

data	efficiency
on-resonance	6.1% ± 0.4%
off-resonance	1.5% ± 0.4%
$B\overline{B}$	12.1% ± 1.2%

The momentum spectrum with lepton tagging after Fisher variable cut that this variable is greater than 0 for the on-resonance data of $28.3 fb^{-1}$ is shown on Figure 3.40. The yield for on-resonance data is 47.8 $\pm$ 12.6 events in the signal region of $0.42 \leq x < 0.51$.

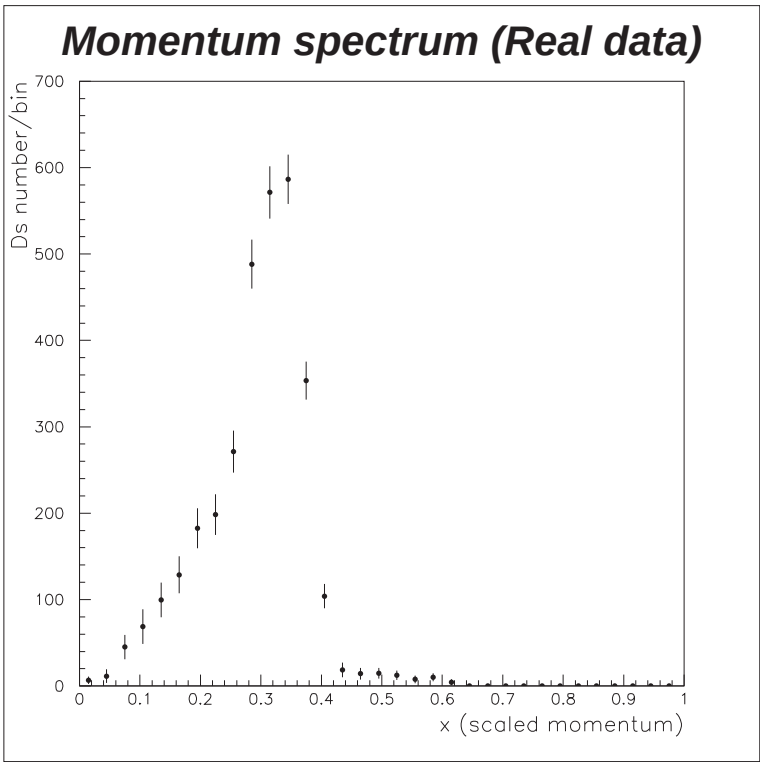


Figure 3.40: Momentum spectrum of on-resonance data with lepton tagging (Real Data)

Distribution of $M_{\phi\pi}$ with the same cut for $3.0 fb^{-1}$ off-resonance data is Figure 3.41. Solid histogram is invariant mass distribution for total momentum, and dashed line is invariant mass distribution for the momentum of the region of $0.42 \leq x < 0.51$. We use double Gaussian as signal shape, and constant as background shape. The yield for off-resonance is 3.0 $\pm$ 3.2 events.

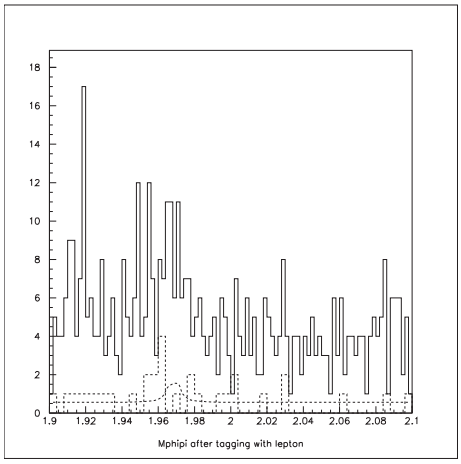


Figure 3.41: $M_{\phi\pi}$ in the signal region with lepton tagging for $3.0 fb^{-1}$ off-resonance (Real Data)

3.3.3 Prospect for future sensitivities

Branching ratio of the $b \rightarrow u\ D_s$ decay to the $b \rightarrow c\ D_s$ decay can be written as

$$\frac{\Gamma(b \rightarrow u\ D_s)}{\Gamma(b \rightarrow c\ D_s)} = |\frac{V_{ub}}{V_{cb}}|^2 \times (Ratio\ of\ phase\ space)$$

Then, the ratio of phase space factor can be calculated as 1.83 (Appendix), therefore, the branching ratio is

$$\frac{\Gamma(b \rightarrow u\ D_s)}{\Gamma(b \rightarrow c\ D_s)} = (0.09)^2 \times 1.83 = 1.5\%$$

where, we use 0.09 as the value of $|\frac{V_{ub}}{V_{cb}}|$ that is shown by Particle Data Group [20]. Using these values, we estimate sensitivities in future.

The significance is calculated with the number of D_s divided by the statistical error of this number in the signal region of $0.42 \leq \mathrm{x} < 0.50$ as

$$Sig = \frac{N_{D_s}^{B\overline{B}}}{\sqrt{\sigma_{on}^2 + (\alpha\sigma_{off})^2}}$$

where, $N_{D_s}^{B\overline{B}}$ is the number of D_s which come from B meson decay, σ_{on} or σ_{off} are error of D_s number with on-resonance or off-resonance data, and α is the scale factor that is the ratio of on-resonance luminosity to off-resonance luminosity. Then, we assume that σ_{off} equals $\sigma_{on}/\sqrt{\alpha}$ because continuum events is dominant in this region as a source of the D_s, then the significance is

$$Sig = \frac{N_{D_s}^{B\overline{B}}}{\sigma_{on}\sqrt{1+\alpha}}$$

The significance with present data set, and future significances with $100fb^{-1}$ on-resonance data and $10fb^{-1}$, $20fb^{-1}$, or $30fb^{-1}$ off-resonance data are summarized in Table 3.11. Then, the significances are calculated with expected values, and the expected values are defined as

$$N_{b\rightarrow cD_s} \times 0.09^2((|V_{ub}|/|V_{cb}|)^2) \times 1.83(phase\ space\ ratio) \times 0.48(over\ end\ point)$$

Where $N_{b\rightarrow cD_s}$ is the number of D_s in the region of $0 \leq \mathrm{x} < 0.42$ for $b \rightarrow cD_s$ transition. And we use 0.48 as the ratio of the over-endpoint event to total event in $b \rightarrow uD_s$ decay [8].

	b→u D_s yield ();expect	Significance on;29.1fb^{-1} off;3.0fb^{-1}	Significance on;100fb^{-1} off;10fb^{-1}	Significance on;100fb^{-1} off;20fb^{-1}	Significance on;100fb^{-1} off;30fb^{-1}
No event shape cut	351±297 (278)	0.9	1.7	2.3	2.6
R2 $\leq$ 0.3 (without fitting)	221±201 (252)	1.3	2.3	3.1	3.5
R2 $\leq$ 0.3 (with fitting)	215±136 (250)	1.8*	3.4*	4.6*	5.1*
Fisher variable $\geq$ 0	-5±200 (237)	1.2	2.3	3.0	3.3
Fisher variable $\geq$ 1.0	43±108 (113)	1.1	1.9	2.6	2.9
Lepton tagging	20±33 (21)	0.6	1.2	1.6	1.8

$*$; systematic error is large

Table 3.11: Prospect for future sensitivities

3.3.4 Measurement of $|V_{ub}|$

One of theory [8] shows that about 48% of the $b \rightarrow u\ D_s$ decay occurs in the region of x is greater than 2.05 GeV/c. This region is correspond to the over-endpoint of the $b \rightarrow c\ D_s$ decay (the momentum 2.05 GeV/c is 0.42 with x) with the assumption of 300MeV for theoretical parameter P_F in the ACCMM model [9] and 10MeV for the mass of spectator quark. Therefore, we calculate $|V_{ub}|$ with the expectation that 48% of the $b \rightarrow u\ D_s$ decay occurs in the signal region of $0.42 \leq \mathrm{x} < 0.50$. Then, we choose R2 cut to be conservative for the upper limit of Vub which is the main result of this analysis. At that time, $|\frac{V_{ub}}{V_{cb}}|^2$ can be extracted with the number of D_s for the result of R2 cut; $N_{b\rightarrow cD_s}$ is 35583 $\pm$ 580 events and $N_{b\rightarrow uD_s}$ is 221 $\pm$ 201 events as.

$$\frac{N_{b\rightarrow uD_s}}{N_{b\rightarrow cD_s}}\frac{\epsilon_{b\rightarrow c}}{\epsilon_{b\rightarrow u}} = |\frac{V_{ub}}{V_{cb}}|^2 \times 1.83 \times 0.48$$

where, $\epsilon_{b\rightarrow u}$ is 30% that is efficiency for the decay chain; $B^0 \rightarrow a_1^-\ D_s^+$, $D_s^+ \rightarrow \phi\ \pi^+$ and $\phi \rightarrow K^+\ K^-$, and $\epsilon_{b\rightarrow c}$ is 25% that is efficiency for the decay chain; $B^0 \rightarrow X_c\ D_s^+$, $D_s^+ \rightarrow \phi\ \pi^+$ and $\phi \rightarrow K^+\ K^-$ in Monte Carlo.

$$|\frac{V_{ub}}{V_{cb}}|^2 = 0.00589 \pm 0.00536\ (stat)$$

we still don't have signal evidence, therefore we assign the upper limit of $|\frac{V_{ub}}{V_{cb}}|^2$ and $|\frac{V_{ub}}{V_{cb}}|$ as

$$|\frac{V_{ub}}{V_{cb}}|^2 \leq 0.0132 \quad (90\%\ CL)$$

$$|\frac{V_{ub}}{V_{cb}}| \leq 0.114 \quad (90\%\ CL)$$

then, we calculate the upper limit of $|\frac{V_{ub}}{V_{cb}}|^2$ with assuming Gaussian and assign the number which is 90% of total positive area for the Gaussian. And, the PDG value;$|V_{cb}| = 0.0402 \pm 0.0019$ [20] introduce

$$|V_{ub}| \ \leq 4.6 \times 10^{-3} \quad (90\%\ CL)$$

Chapter 4

Conclusion

We measured the branching fraction of $B \to X_c\ D_s$ decay with using on-$\Upsilon(4S)$-resonance data of $29.1 fb^{-1}$ that is accumulated by BELLE detector as

$$\Gamma(B \to X_c\ D_s) = (\ 11.0\ \pm 0.3\ (stat)\ \pm 2.9\ (sys)\ )\%$$

where the values of Particle Data Group or CLEO collaboration are (10.0 ± 2.5) % in 2000 PDG, and (11.7 ± 0.9 (stat) ± 2.8 (sys))% in 1996 CLEO. We measured the value $|\frac{Vub}{Vcb}|^2$ with on-resonance data of $29.1 fb^{-1}$ and off-resonance data of $3.0 fb^{-1}$ as

$$|\frac{Vub}{Vcb}|^2 = 0.0059 \pm 0.0054\ (\text{stat})$$

we still don't have signal evidence, therefore the upper limit for $|\frac{V_{ub}}{V_{cb}}|^2$ and $|\frac{V_{ub}}{V_{cb}}|$ is

$$|\frac{V_{ub}}{V_{cb}}|^2 \leq 0.013 \quad (90\%\ CL)$$

$$|\frac{V_{ub}}{V_{cb}}| \leq 0.11 \quad (90\%\ CL)$$

where, the value of $|\frac{V_{ub}}{V_{cb}}|$ in PDG is 0.090 ± 0.025. And, we extracted the Kobayashi-Maskawa Matrix element $|V_{ub}|$ as

$$|V_{ub}|\ \leq 4.6 \times 10^{-3} \quad (90\%\ CL)$$

where, the value by the end point of lepton momentum spectrum of semi-leptonic decay in CLEO 2000 is $|V_{ub}| = \{\ 3.25\ \pm\ 0.14\ (stat)\ ^{+0.21}_{-0.29}\ (sys)\ \pm\ 0.55\ (theory)\ \} \times 10^{-3}$. We could approach the $|V_{ub}|$ with new method of $b \to uD_s$ decay.

Appendix A

Ratio of phase space

In the decay of $b \to q\ D_s$, the invariant amplitude $\mathcal{M}$ becomes

$$\mathcal{M} = (\overline{u'}\sqrt{2}G\gamma^{\mu}P_L u) f q_{\mu} \tag{A.1}$$
$$\tag{A.2}$$

where u is the four-component spinor of b-quark, $\overline{u'}$ is the four-component spinor of u-quark or c-quark, f is the decay constant of D_s, and q_μ is the four momentum of D_s.

$$= \sqrt{2}Gf\overline{u'}\not{q}\ P_L u \tag{A.3}$$
$$= \sqrt{2}Gf\overline{u'}(\not{P}\ - \not{P'})P_L u \tag{A.4}$$

where P is the four momentum of b-quark, and P' is the four momentum of u-quark or c-quark.

$$= \sqrt{2}Gf(\overline{u'}\not{P}\,P_L u - \overline{u'}\not{P'}P_L u) \tag{A.5}$$
$$= \sqrt{2}Gf(\overline{u'}P_R\not{P}\,u - \overline{u'}\not{P'}P_L u) \tag{A.6}$$
$$= \sqrt{2}Gf(\overline{u'}P_R\,m\,u - m'\ \overline{u'}P_L u) \tag{A.7}$$
$$= \sqrt{2}Gf\overline{u'}V u \tag{A.8}$$

where m is the mass of b-quark, and m' is the mass of u-quark or c-quark. And, replaced $(m\,P_R - m'P_L)$ with V, then $\overline{V}$ becomes $(m\,P_L - m'P_R)$, therefore, spin sum becomes

$$\sum_{spin}|\mathcal{M}|^2 = 2G^2f^2\sum_{spin}|\overline{u'}Vu|^2 \tag{A.9}$$
$$= 2G^2f^2\sum_{spin}\overline{u'}Vu\overline{u}\overline{V}u' \tag{A.10}$$
$$= 2G^2f^2\,Tr\{(\not{P'}+m')V(\not{P}+m)\overline{V}\} \tag{A.11}$$
$$= 2G^2f^2\,(Tr\not{P'}V\not{P}\,\overline{V} + m\,m'\,Tr\,V\overline{V}) \tag{A.12}$$

then, $\not{P}' V \not{P} \overline{V}$ and $V\overline{V}$ are

$$\begin{aligned} \not{P}' V \not{P} \overline{V} &= \not{P}'(m\,P_R - m' P_L)\not{P}\,(m\,P_L - m' P_R) && \text{(A.13)} \\ &= \not{P}'\not{P}\,(m\,P_L - m' P_R)\,(m\,P_L - m' P_R) && \text{(A.14)} \\ &= \not{P}'\not{P}\,(m^2 P_L + m'^2 P_R) && \text{(A.15)} \\ &= \frac{\not{P}'\not{P}}{2}\{(m^2 + m'^2) - (m^2 - m'^2)\gamma_5\} && \text{(A.16)} \end{aligned}$$

and

$$\begin{aligned} V\overline{V} &= (m\,P_R - m' P_L)\,(m\,P_L - m' P_R) && \text{(A.17)} \\ &= -m\,m'(P_L + P_R) && \text{(A.18)} \\ &= -m\,m' && \text{(A.19)} \end{aligned}$$

therefore, equations of (A.16), (A.19) and (A.12) introduce

$$\begin{aligned} \sum_{spin} |\mathcal{M}|^2 &= 2G^2 f^2\,\{(m^2 + m'^2)\frac{Tr\not{P}'\not{P}}{2} - (m^2 - m'^2)Tr\not{P}'\not{P}\,\gamma_5 - m^2\,m'^2 Tr1\} && \text{(A.20)} \\ &= 2G^2 f^2\,\{(m^2 + m'^2)(2\not{P}'\cdot\not{P}) - 4m^2\,m'^2\} && \text{(A.21)} \\ &= 4G^2 f^2\,\{(m^2 + m'^2)\,m\,E' - 2m^2\,m'^2\} && \text{(A.22)} \\ & && \text{(A.23)} \end{aligned}$$

because $P = (m, 0)$ and $P' \equiv (E', \overrightarrow{P'})$ in the b-quark frame, means $\not{P}'\cdot\not{P} = mE'$, so the two body decay lead,

$$\begin{aligned} \Gamma &= \frac{|\overrightarrow{P'}|}{8\pi\,m^2}\sum_{spin}|\mathcal{M}|^2 && \text{(A.24)} \\ &= \frac{G^2 f^2}{2\pi}\,|\overrightarrow{P'}|\,m^2\{(1 + \frac{m'^2}{m^2})\frac{E'}{m} - 2\frac{m'^2}{m^2}\} && \text{(A.25)} \end{aligned}$$

then, E' and $|\overrightarrow{P'}|$ can be described with M_{D_s} (which is the mass of D_s), x $\equiv M_{D_s}/m$, and y $\equiv m'/m$ as follows

$$\begin{aligned} E' &= \frac{m^2 + m'^2 - M_{Ds}^2}{2m} && \text{(A.26)} \\ &= m\,(\frac{1 + y^2 - x^2}{2}) && \text{(A.27)} \end{aligned}$$

and

$$\begin{aligned} P'^2 &= E'^2 - m'^2 && \text{(A.28)} \\ &= (\frac{m^2 + m'^2 - M_{Ds}^2}{2m})^2 - m'^2 && \text{(A.29)} \\ &= m^2(\frac{1 + x^4 + y^4 - 2x^2 - 2x^2y^2 - 2y^2}{4}) && \text{(A.30)} \\ &= m^2[\frac{\{1 - (x+y)^2\}\{1 - (x-y)^2\}}{4}] && \text{(A.31)} \end{aligned}$$

therefore, equations of (A.27), (A.31) and (A.25) introduce

$$\Gamma = \frac{G^2 f^2}{8\pi} m^3 \sqrt{\{1-(x+y)^2\}\{1-(x-y)^2\}} \times \{(1-y^2)(1+y^2-x^2)-4y^2\} \quad \text{(A.32)}$$

$$= \frac{G^2 f^2}{8\pi} m^3 \sqrt{\{1-(x+y)^2\}\{1-(x-y)^2\}} \times \{1-x^2-2y^2-x^2y^2+y^4\} \quad \text{(A.33)}$$

replacement of the equation (A.33) with mass of each particles again, we gain

$$\Gamma \propto \sqrt{\{1-(\frac{M_{D_s}}{m}+\frac{m'}{m})^2\}\{1-(\frac{M_{D_s}}{m}-\frac{m'}{m})^2\}} \times \{1-(\frac{M_{D_s}}{m})^2-2(\frac{m'}{m})^2-(\frac{M_{D_s}}{m})^2(\frac{m'}{m})^2+(\frac{m'}{m})^4\}$$

where, M_{D_s} means mass of Ds meson 1.968GeV, m means mass of bottom quark 4.2GeV, m' means mass of charm quark 1.25GeV, or means mass of u quark 3MeV.

We can calculate the ratio of phase space for $b \to u\ D_s$ decay to $b \to c\ D_s$ as 1.83 with above assumption of each mass. One of theory [8] shows that about 48% of the $b \to u\ D_s$ decay occurs in the region x $\geq$ 2.05 GeV/c that is over the end point of the $b \to c\ D_s$ decay (This momentum correspond to 0.42 with x), so we use the value 0.09$\times$0.09$\times$ 1.83$\times$0.52 $\sim$ 0.8% as the systematic error of spillover contamination of $B \to X_u\ D_s$ for $B \to X_c\ D_s$ decay when we measure the branching fraction.

Bibliography

[1] M.Kobayashi and T.Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[2] A.B.Carter and A.I.Sanda, Phys. Rev. Lett. **45**, 952 (1980);
Phys. Rev. **D 23**, 1567 (1981).

[3] B. Aubert et al., Phys. Rev. Lett. **87**, 091801 (2001)

[4] K.Abe et al., Phys. Rev. Lett. **87**, 091802 (2001).

[5] L.Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983)

[6] Particle Data Group, D.E.Groom et al., Eur. Phys. J. C **15**, 41 (2000).

[7] CLEO Collaboration, B.H. Behrens et al., Phys. Rev. **D 61**, 052001 (2000).

[8] R.Aleksan, et al., Phys. Rev. **D 62**, 093017 (2000).

[9] G.ALTARELLI, et al., Nucl. Phys. **B 208**, 365 (1982).

[10] Y.Ushiroda et al., Belle note **273** (1999); Belle note **280** (2000).
S.Nishida et al., Belle note **350** (2000); Belle note **381** (2001); Belle note **423** (2001).

[11] CLEO software document homepage :
http://www.lns.cornell.edu/public/CLEO/soft/QQ/index.html (1998).

[12] R. Itoh, *QQ quick reference for BELLE* (1995).

[13] T.Sjöstrand, "PYTHIA 5.7 and JETSET 7.4 Physics and Manual", CERN-TH.7112/93 (1993).

[14] *"GEANT" Detector Description and Simulation Tool*, CERN program Library Long Writeup W5013 (1993).

[15] Brendan Casey, Belle note **390** (2001)

[16] G.C.Fox and S.Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

[17] Belle HOME PAGE, <http://belle.kek.jp/>.

[18] C.Peterson et al, Phys. Rev. **D 27**, 105 (1983).

[19] R.A. FISHER, Annals of Eugenics, **7**, 179 (1936).

[20] Particle Data Group, D.E.Groom et al., Eur. Phys. J. C **15**, 111 (2000).