Ph.D. Thesis

Precision Measurements of Radiative B Meson Decay $B \to X_s\gamma$ with a Semi-inclusive Reconstruction Method

(準包括的再構成法を用いた $B \to X_s\gamma$ の精密測定)

Department of Physics, Tohoku University

Tomoyuki Saito

December 2013

Abstract

The $b \to s\gamma$ process is forbidden at tree level in the Standard Model(SM) and proceeds via loop radiative penguin diagrams. It is a sensitive probe of new physics(NP) beyond the SM because new particles can appear in the loop. We report the precise measurement of the branching fraction for the inclusive decay $B \to X_s\gamma$, where X_s is all the hadron combination that a strange quantum number of s quark. 38 final states which consist of $K^{\pm}, K^0_s, \pi^{\pm}, \pi^0$ and η are reconstructed with a semi-inclusive reconstruction method. We use a data sample that contains $771{\times}10^6$ $B\overline{B}$ pairs collected by the Belle detector at the KEKB collider.

The inclusive branching ratio in M_{X_s} <2.8 GeV/c^2 is measured to be

$$\mathcal{BR}(B \to X_s\gamma) \;=\; (3.51 \pm 0.17 \pm 0.33) \times 10^{-4},$$

where the first uncertainty is statistical and the second is systematic. The inclusive branching ratio with a minimum photon energy of 1.6 GeV is measured to be

$$\mathcal{BR}(B \to X_s\gamma) \;=\; (3.74 \pm 0.18 \pm 0.35) \times 10^{-4}.$$

This measurement is consistent with the world average, $(3.55{\pm}0.26){\times}10^{-4}$, within 0.4σ, and consistent with the SM prediction, $(3.15{\pm}0.23){\times}10^{-4}$, within 1.3σ. This result provides a constraint the NP. We evaluate a constraint to the two Higgs Doublet Model, and the charged Higgs mass region below 238 GeV/c^2 is excluded at 95% CL.

Acknowledgment

First of all, I would like to express my deep gratitude to my supervisor Prof. Hitoshi Yamamoto, for giving me a opportunity to research in high energy physics. It is my great fortune that I spent my Ph. D. course in his laboratory.

I deeply appreciate Dr. Akimasa Ishikawa for his insightful advices. He taught me so many things for this analysis and always led me to right direction. I learned a lot of things from his sincere attitude to physics.

I would like to express my gratitude to Dr. Mikihiko Nakao and Dr. Shohei Nishida of the KEK. They gave me important advices on my study and encouraged me.

I would like to thank to the internal referees in the Belle collaboration, Dr. Phillip Urquijo, Dr. Jamal Rorie and Prof. Yutaka Ushiroda, who gave me a lot of valuable suggestions and corrections to my thesis.

I am deeply thankful to all the members of the BelleII SVD group. When I addressed the development of the SVD readout system, they gave me many suggestions. I am specially grateful to Dr. Koji Hara and Dr. Mikihiko Nakao of the KEK for their supports.

I would like to thank Dr. Tomoyuki Sanuki and Dr. Tadashi Nagamine for their encouragements. I am deeply thankful to the secretary Ms. Kaori Kobayashi, for taking care of every official business.

It is my pleasure to express my biggest thanks to all the members in this laboratory. We were sharing many good and hard times. I'm proud of being a member of the laboratory.

I deeply appreciate my mother Midori, my father Yasushiro, who gave me so many supports. Also, my sister Rina gave me many encouragements. At last, I would like to express my sincere thankfulness to my wife Shiori. I could not have made up this thesis without her dedicated supports.

Sincerely,
Tomoyuki Saito

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The aim of particle physics is to understand the universe. A lot of theoretical and experimental physicists have made a great effort and led to construct the Standard Model(SM), which describes the basic structure of particles and interactions. This theory provides a description of the strong, weak and electromagnetic interactions, and contains quarks, leptons, force carriers. All the experimental results to date are consistent with the SM prediction basically. Furthermore, a Higgs boson was observed by the ATLAS and CMS at the Large Hadron Collider(LHC) in July, 2012(Figure 1). The discovery was a remarkable achievement and only the missing piece in the SM. The SM is very successful theory framework, however leaves some big questions unanswered; why there are so many fundamental particles and why they make generations. In addition, the SM simply fails to explain some phenomena, such as the matter-antimatter asymmetry in the universe, the existence of dark matter. These issues lead us to need a new physics(NP) beyond the SM.

A variety of approaches are essential for the search of the NP. We use three basic approaches, the Energy, Cosmic and Intensity Frontiers. Each approach uses different tools and techniques, but they have the same final goals. The Intensity Frontier explores fundamental physics with intense sources and ultra-sensitive detectors and covers searches for extremely rare processes and for tiny deviations from the SM. Intensity Frontier experiments use precision measurements to probe quantum effects. They can investigate the NP effects at higher energy than that directly accessible in collider experiments and provides an opportunity for substantial new discoveries complementary to Energy and Cosmic Frontier experiments.

The LHC experiments which represent the Energy Frontier accumulated data well at 7 and 8 TeV by the end of 2012, however no evidence for the NP

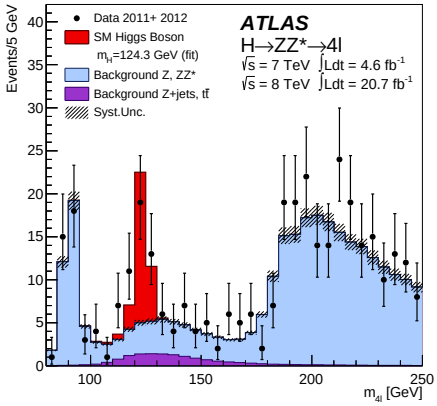


Figure 1.1: The distribution of the four-lepton invariant mass in the ATLAS. The estimated background, as well as the expected SM Higgs boson signal for $m_H = 124.3$ GeV, are also shown.

was discovered. We have a great expectation of the discovery in the running at 13 or 14 TeV from 2015. Once the NP is discovered, it is essential to measure the structure in the NP model by the indirect search, such as measurements of the Intensity Frontier. If the evidence of the NP is not found by the direct search at the LHC, the importance of the indirect search at the Intensity Frontier is more and more.

For the indirect search, rare processes that are forbidden or suppressed in the SM, but may be enhanced in the NP are significant. In the SM, the radiative b decay, $b \to s\gamma$, which is a Flavor Changing Neutral Current(FCNC) process, is forbidden at a tree-level and proceeds with loop diagrams. Thus, it has a good sensitivity to a new heavy particle in the loop and is a good probe to the NP.

In this dissertation, we measure a branching fraction for $B \to X_s\gamma$(X_s is all the hadron combinations that carry a strange quantum number of s quark) by a semi-inclusive technique with KEKB accelerator and Belle detector. The Belle experiment has a huge number of $B\overline{B}$ pairs, 770×10^6 and the expected number of $B \to X_s\gamma$ signal is about 5×10^5. Thus, the systematic uncertainty is dominant in this analysis. For precise measurement, it is essential to find a way to evaluate correctly and suppress the systematic uncertainty.

The outline of this thesis is as follows: the physics and measurement to date for $B \to X_s\gamma$ are given in Chapter 2. An overview of the KEKB accelerator and the Belle detector is presented in Chapter 3. Data sets in this analysis are explained in Chapter 4. Event selection and reconstruction are described in

Chapter 5 and background is studied in Chapter 6. The method to extract the signal yield is described in Chapter 7. The systematic uncertainties are studied in Chapter 8. The method to calculate the branching ratio and evaluate the hadronization model uncertainty is discussed with the partial data in Chapter 9. Finally, the branching ratio is calculated with the full data in Chapter 10 and the conclusion is given in Chapter 11.

Chapter 2

Radiative B meson Decay

2.1 Radiative B Meson Decay

The B meson system, which is a bound state that consists of a b quark and a light quark, provides an ideal laboratory for precise study of the SM, and thus facilitates the search for the NP. Because the b quark mass is much larger than the typical scale of the strong interaction, the troublesome strong interactions are generally less important and are under better control than in other lighter meson systems. Radiative decays of the B meson with emission of a photon are particularly important. These processes is the flavor changing neutral current(FCNC), that is the transition of a b quark with an electric charge of -1/3 into an s or a d quark of the same charge. The simplest transition is $b \to s/d\gamma$. The rate of the $b \to d\gamma$ is about 100 times smaller than that of $b \to s\gamma$ due to a suppression by $|V_{td}/V_{ts}|^2$ in the Cabbibo Kobayashi Maskawa(CKM) matrix. The diagram of the $b \to s\gamma$ is shown in Figure 2.1. This process is forbidden at the tree level in the SM and proceeds via radiative loop diagrams(Appendix A). Since the loop diagram is dominant, effects of new particles within the loop predicted by many NP models in Figure 2.2 which we have not been directly accessible in collider experiments, enhancing or suppressing this branching ratio, can be investigated by precision measurements. The $b \to s\gamma$ is highly sensitive to the NP because it is theoretically well-understood. The search for such NP effects complements a search for new particles directly by produced high energy collider experiments, such as the LHC.

Quark-level process cannot be directly measured because the strong interaction forms hadrons from underlying quarks, unlike $\mu \to e\gamma$ process. Instead, the experimentally measured and theoretically calculated process is a B meson decay into a photon plus an inclusive hadronic final state X_s which

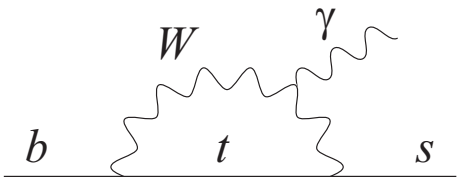
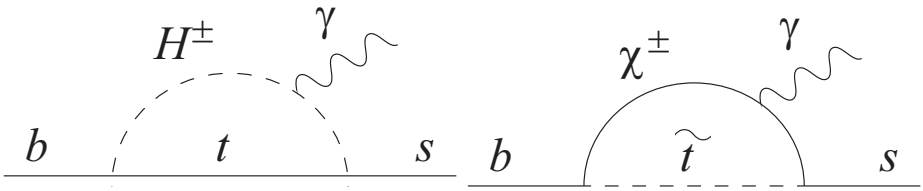


Figure 2.1: $b \to s\gamma$ diagram in the SM



(a) 2 Higgs Doublet Model (b) SUSY

Figure 2.2: $b \to s\gamma$ diagrams in NP model

includes all the hadron combinations that carry a strange quantum number of s quark(Inclusive decay, $B \to X_s\gamma$). In addition, exclusive decays have one or a few specific hadrons in the final state(e.g., $B \to K^*\gamma$)

2.2 Electroweak Effective Hamiltonian

This section provides a detail description of $b \to s\gamma$ branching fraction and shows that the branching fraction of inclusive decay is more useful for the NP search than that of exclusive decay.

Inclusive B decays are theoretically clean because they are dominated by partonic(perturbatively calculable) contributions. Non-perturbative corrections are generally rather small[1, 2, 3, 4]. This result can be helped by the heavy mass expansion(HME) of the inclusive decay rates in inverse powers of the b quark mass. Up-to-date predictions of exclusive decay B decays are based on the quantum chromodynamics(QCD)-improved factorization(QCDF) and soft collinear effective theory(SCET) methods. In general, exclusive decay rate has larger non-perturbative QCD corrections than inclusive decay rate.

Radiative B decays are governed by the weak and strong interactions. The QCD corrections that arise from hard gluon exchange bring in large logarithms of the form $\alpha_s^n(m_b)\log^m(m_b/M)$, where $M = m_t$ or $M = m_w$ and $m \leq n$(with n=0,1,2...). These large logarithms are a natural feature in any process in which two different mass scales are present. To obtain a reasonable result, we must resum at least the leading-log(LL) series, $n = m$, with the help of

renormalization-group techniques. Working for next-to-leading-log(NLL) or next-to-next-to-leading-log(NNLL) precision corresponds to resummation of all the terms with $n = m + 1$ or $n = m + 2$. A suitable framework to achieve the necessary resummations of the large logarithms is an effective low-energy theory with five quarks; this frameworks obtained by interacting out the heavy particles, which are the electroweak bosons and the top quark in the SM. This effective field theory approach provides a theoretical framework for both inclusive and exclusive decays. The standard method of the operator product expansion(OPE)[5, 6] allows a separation of the B meson decay amplitude into two distinct parts: the long-distance contributions contained in the operator matrix elements and the short-distance physics described by the so-called Wilson coefficients.

The electroweak effective Hamiltonian[7, 8, 9] can be written as

$$\mathcal{H}_{eff} = \frac{4G_F}{\sqrt{2}} \sum C_i(\mu, M)\mathcal{O}_i(\mu), \tag{2.1}$$

where $O_i(\mu)$ are the relevant operators and $C_i(\mu, M)$ are the corresponding Wilson coefficients. As the heavy fields are integrated out, the complete top and W mass dependence is contained in the Wilson coefficients. G_F denotes the Fermi coupling constant.

The effective electroweak Hamiltonian in the SM reads[1]

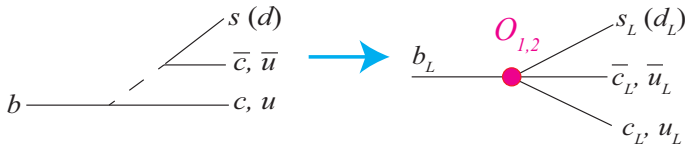
$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} \left(\lambda_q^t \sum_{i=1}^{10} C_i \mathcal{O}_i + \lambda_q^u \sum_{i=1}^{2} C_i \left(\mathcal{O}_i - \mathcal{O}_i^u \right) \right), \tag{2.2}$$

where the explicit CKM factors are $\lambda_q^t = V_{tb}V_{tq}^*$ and $\lambda_q^u = V_{ub}V_{uq}^*$. The unitary relations $\lambda_q^c = -\lambda_q^t - \lambda_q^u$ have already been used. The dimension-six operators

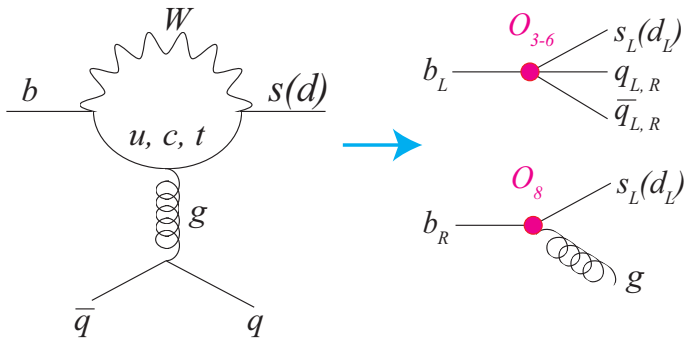
are

$$\mathcal{O}_1 = (\overline{s}_L\gamma_\mu T^a c_L)\,(\overline{c}_L\gamma^\mu T^a b_L)\,, \tag{2.3}$$
$$\mathcal{O}_1^u = (\overline{s}_L\gamma_\mu T^a u_L)\,(\overline{u}_L\gamma^\mu T^a b_L)\,, \tag{2.4}$$
$$\mathcal{O}_2 = (\overline{s}_L\gamma_\mu c_L)\,(\overline{c}_L\gamma^\mu b_L)\,, \tag{2.5}$$
$$\mathcal{O}_2^u = (\overline{s}_L\gamma_\mu u_L)\,(\overline{u}_L\gamma^\mu b_L)\,, \tag{2.6}$$
$$\mathcal{O}_3 = (\overline{s}_L\gamma_\mu b_L)\sum_q (\overline{q}\gamma^\mu q)\,, \tag{2.7}$$
$$\mathcal{O}_4 = (\overline{s}_L\gamma_\mu T^a b_L)\sum_q (\overline{q}\gamma^\mu T^a q)\,, \tag{2.8}$$
$$\mathcal{O}_5 = (\overline{s}_L\Gamma b_L)\sum_q (\overline{q}\Gamma' q)\,, \tag{2.9}$$
$$\mathcal{O}_6 = (\overline{s}_L\Gamma T^a b_L)\sum_q (\overline{q}\Gamma' T^a q)\,, \tag{2.10}$$
$$\mathcal{O}_7 = \frac{e^2}{16\pi^2} m_b\,(\overline{s}_L\sigma^{\mu\nu} b_R)\,F_{\mu\nu}, \tag{2.11}$$
$$\mathcal{O}_8 = \frac{g_s}{16\pi^2} m_b\,(\overline{s}_L\sigma^{\mu\nu} T^a b_R)\,G^a_{\mu\nu}, \tag{2.12}$$
$$\mathcal{O}_9 = \frac{e^2}{16\pi^2}\,(\overline{s}_L\gamma_\mu b_L)\sum_l \left(\bar{l}\gamma^\mu l\right), \tag{2.13}$$
$$\mathcal{O}_{10} = \frac{e^2}{16\pi^2}\,(\overline{s}_L\gamma_\mu b_L)\sum_l \left(\bar{l}\gamma^\mu\gamma_5 l\right), \tag{2.14}$$

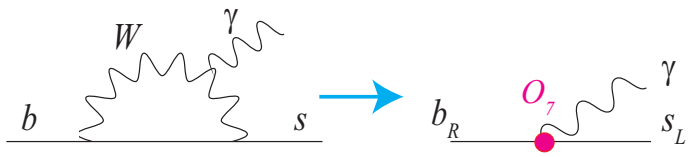
where T^a = are SU(3) color generator, $F_{\mu\nu}$ and $G_{\mu\nu}$ are electromagnetic and chromomagnetic fields, and $\Gamma = \gamma_\mu\gamma_\nu\gamma_\lambda$ and $\Gamma' = \gamma^\mu\gamma^\nu\gamma^\lambda$. The subscripts L and R refer to left- and right-handed components, respectively, of the fermion field. $\mathcal{O}_{1-2}$ are current-current operators in Figure 2.3(a), $\mathcal{O}_{3-6}$ are QCD penguin operators in Figure 2.3(b), $\mathcal{O}_{7-8}$ are electro- and chromo- operators in Figure 2.3(c) and 2.3(b), respectively, and $\mathcal{O}_{9-10}$ are semi-leptonic operators in Figure 2.3(d). In $b \to s$ transitions the contributions proportional to λ_s^u are rather small, whereas in $b \to d$ transitions where λ_d^u is of the same order as λ_d^t, these contributions play an important role in CP and isospin asymmetries. The $\mathcal{O}_{9-10}$ occur only in the semi-leptonic $b \to s/d l^+ l^-$ modes. Among the four-quark operators, only the effective couplings for i =1, 2 are large at the low scale, $C_{1,2}(m_b) \sim 1$, whereas the coupling of the other four-quark operators have almost negligible values. But the electromagnetic operator with $C_7(m_b)$ ∼-0.3 and the chromomagnetic operator with $C_8(m_b)$ ∼-0.15 play a significant role in both $b \to s(d)\gamma$ and $b \to s(d)l^+l^-$. The vector and axial-vector contributions to $b \to s(d)l^+l^-$ have $C_9(m_b)$ ∼4, C_{10} ∼-4.



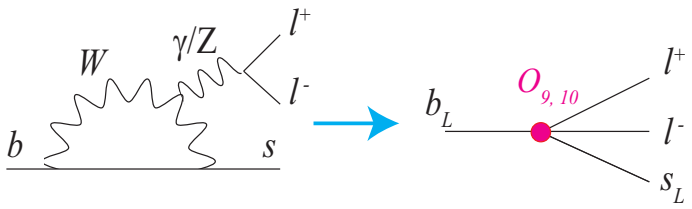
(a) Current-Current Operators



(b) Strong Penguin Operators



(c) Electromagnetic Operators



(d) Semi-leptonic Operators

Figure 2.3: Operators

Although the Wilson coefficients $C_i(\mu)$ enter both inclusive and exclusive processes and can be calculated with perturbative methods, the calculational approaches to the matrix elements of the operators differ between the two cases. In inclusive modes, we can use the quark-hadron duality to derive a well-defined HME of the decay rates in powers of Λ/m_b[10, 11, 12, 13, 14, 15]. In particular, the decay width of the $B \to X_s\gamma$ is well approximated by the partonic decay rate, which can be calculated in renormalization-group-improved perturbation

theory[16, 17]. On the other hand, in exclusive modes, we cannot rely on the quark-hadron duality, so we must estimate the matrix elements between meson states. Therefore, the exclusive final states have less predictive power theoretically, whereas the inclusive mode is a good probe for NP search due to the large prediction power.

The matrix element and decay width of $b \to s\gamma$ transition[18] are

$$\mathcal{M} = \frac{4G_F}{\sqrt{2}} \frac{e}{16\pi^2} V_{ts}^* V_{tb} m_b C_7 \left(\overline{s}\sigma_{\mu\nu} b_R\right) F^{\mu\nu}, \tag{2.15}$$

$$\Gamma = \frac{G_F^2 \alpha_{em} m_b^5}{32\pi^4} |C_7|^2 |V_{tb} V_{ts}^*|^2. \tag{2.16}$$

The $b \to s\gamma$, which is mainly generated through the electromagnitic diagram, are sensitive to the absolute value of C_7. In contrast, asymmetries on the $b \to s\gamma$ are sensitive to the imaginary part of the coefficients.

2.3 Experimental Techniques for Inclusive decay $B \to X_s\gamma$

Measurement of the inclusive decay $B \to X_s\gamma$ is difficult for the hadron collider, such as LHCb, due to a huge γ background. On the other hand, the Υ (4S) resonance produced by e^+e^- collision provides a clean sample of $B\overline{B}$ meson pairs. Therefore, the techniques at the e^+e^- collision are described in this section.

Fully inclusive measurement In a fully inclusive measurement, only photon is basically reconstructed and the X_s system is not reconstructed. In this method, a large uncertainty from the X_s decay model can be highly suppressed. But, the background rejection is challenging in this measurement because a measured object is only photon. The dominant background photon sources are copiously produced $\pi^0 \to \gamma\gamma$ decays.

Semi-inclusive measurement In a semi-inclusive measurement, we measure as many exclusive modes as possible and then calculate their sum to obtain the inclusive branching fraction. Exclusive branching fractions measured to date do not saturate the inclusive process, but we can still infer the total branching fraction by estimating the fraction of unmeasured modes using simulated hadronization processes. Therefore, the signal modeling of the X_s

decay in the simulation is significant. In this method, B mesons can be reconstructed by full reconstructions of final state particles. Thus, the signal peak of $B \to X_s\gamma$ are clearly obtained in background events in contrast to the fully-inclusive measurement. In addition, a background suppression is easier since the B meson information is available. This method also provides direct information about the B meson. For example, the charge and flavor information allows CP and isospin asymmetry measurements.

Fully-inclusive and B full-reconstruction measurement Other potentially technique is a full-inclusive and B full-reconstruction measurement, in which only a photon is reconstructed in the signal side in the same method of the fully inclusive technique and the other B meson of $B\overline{B}$ pair is fully reconstructed. Thus, this measurement allow the target B decay to be measured in a very clean environment. The reconstruction efficiency is very low and this method will be more important in future experiments, such as BelleII.

2.4 Previous Measurements

2.4.1 Branching Fraction

The branching fraction of $B \to X_s\gamma$ has been measured by the CLEO, BABAR and Belle summarized in Figure 2.4. The SM calculation has been performed at next-to-next leading order in the perturbative term and the result is $\mathcal{BR}(B \to Xs\gamma) = (3.15 \pm 0.23) \times 10^{-4}$ for a photon energy above 1.6 GeV in the B meson rest frame[21, 22, 23]. The current world averages by HFAG2012[20] and PDG2013 are in an agreement with the SM calculation within the uncertainties.

2.4.2 Constraints to the Two Higgs Doublet Model

The results of the $\mathcal{BR}(B \to Xs\gamma)$ imply very stringent constraints on NP models. Here, we describe the Two Higgs Doublet Models(2HDM)[24] which are simple extensions to the SM Higgs sector, only introducing an additional $SU(2)_L \times U(1)_Y$ Higgs doublet, leading to five physical Higgs bosons. Three Higgs bosons (A^0, h^0, H^0) are electrically neutral and the two remaining ones $(H^\pm)$ are electrically charged. The free parameters of the 2HDM are the Higgs boson masses $M_{A^0}, M_{h^0}, M_{H^0}$ and $M_{H^\pm}$, the ratio of the vacuum expectation values of the two Higgs doublets $\tan\beta = v_2/v_1$, occurring in the mixing of

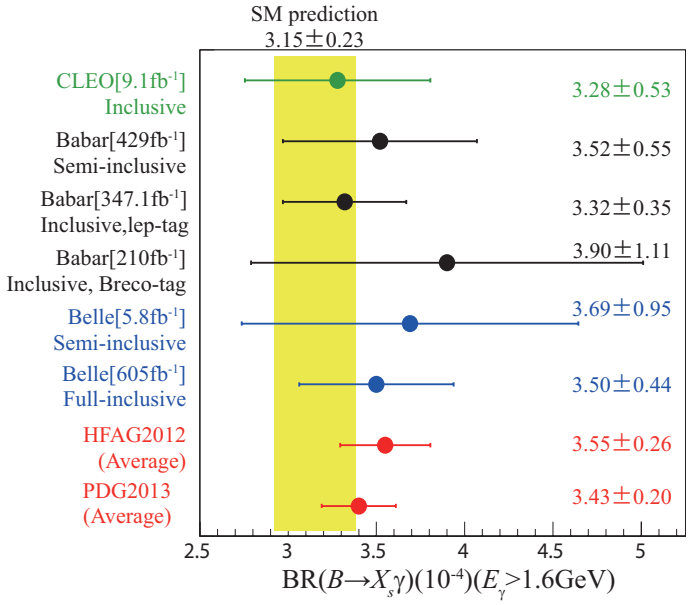
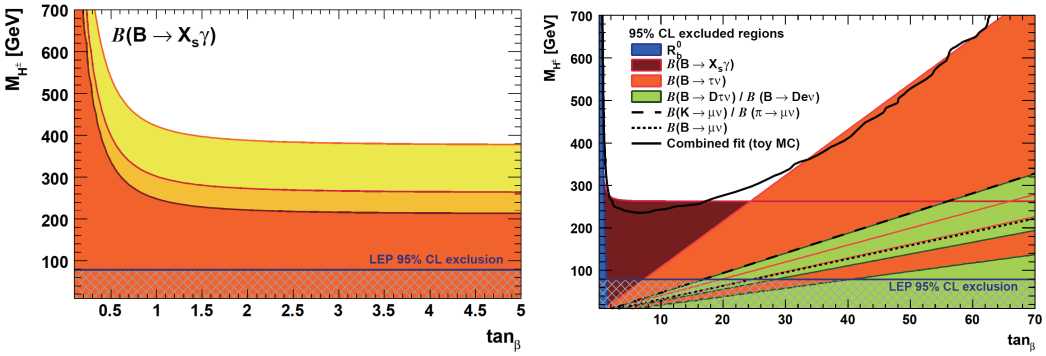


Figure 2.4: Measurement results and theoretical calculation of $B \to X_s\gamma$. The CLEO, BaBar and Belle measurements are shown, and the averages are shown by red markers. The yellow band means the SM prediction.

charged and neutral Higgs fields, and the angle α, governing the mixing of the neutral CP-even Higgs fields.

The branching fraction of the $B \to X_s\gamma$ can be changed with an appropriate choice of the Higgs-fermion couplings. For example, in the Type-I 2HDM only one Higgs doublet couples to the fermion sector. In the Type-Ⅱ 2HDM[25], one Higgs doublet couples to the up-type quarks and leptons only, while the other one couples only to the down-type quarks and leptons. The Type-Ⅱ 2HDM resembles the Higgs sector in the Minimal Supersymmetric Standard Model(MSSM). It fixes the basis of the Higgs fields and promotes $\tan\beta$ to a physical parameter[28].

By investigating observables that are sensitive to corrections from a charged Higgs exchange in Figure 2.2(a) we can search effects on the charged Higgs and provide constraints on the allowed charged-Higgs mass $M_{H^\pm}$ and $\tan\beta$. In direct searches, LEP has derived a lower limit of $M_{H^\pm} > 78.6$ GeV at 95 % CL[26], for any value of $\tan\beta$. The constraints on the charged Higgs are currently dominated by indirect measurements, as opposed to direct searches at high-energy accelerators. The most relevant observables for the search of Type-Ⅱ 2HDM signal are the electroweak precision variable R_b^0 which is the hadronic branching fraction of of Z to b quarks, branching fractions of rare



(a) Constraint from $\mathcal{BR}(B \to X_s\gamma)$. Two-sided 68%, 95% and 99% CL exclusion region.

(b) 95 % CL excluded regions from combined fit with R_b^0, B, D, K decays.

Figure 2.5: 2HDM paramter plane M_{H^+} versus tanβ[28]. These constraints are a bit old, but the differences are not large.

semileptonic B, D and K decays, and loop-induced radiative B decays.[1] The 2HDM contribution is always positive in the typeⅡ model. The constraint to the 2HDM from the branching fraction of the $B \to X_s\gamma$ shows in Figure 2.5(a)[28]. The $\mathcal{BR}(B \to X_s\gamma)$ is especially sensitive to tanβ for small values. For larger tanβ it provides an almost constraint area of exclusion of a charged Higgs lighter than $\simeq$270 GeV/c^2. This figure is a bit old, and the present constraint is close to 380 GeV/c^2[29]. The 95 % CL excluded regions from combined fit with R_b^0, and B, D, K decays are shown in Figure 2.5(b). Recently, the LHC experiments attempt to directly detect signals from charged Higgs production, using the τ+jets channel[30]. For light charged Higgs bosons($m_{H^\pm} < m_{top}$), the $t\bar{t} \to H^+bWb$ production mode is dominant, while for heavy charged Higgs bosons, associated production of $tH^\pm$ is dominant. No evidence for a charged Higgs boson is found. The constraint from the result are shown in Figure 2.6. This excluded region is already strongly disfavored by the current $\mathcal{BR}(B \to X_s\gamma)$ measurements.

[1]Decays of τ and μ leptons can also occur through charged Higgs tree diagrams giving anomalous contributions to the decay parameters measured in these decays. Their present sensitivity is however not competitive with the other observables.

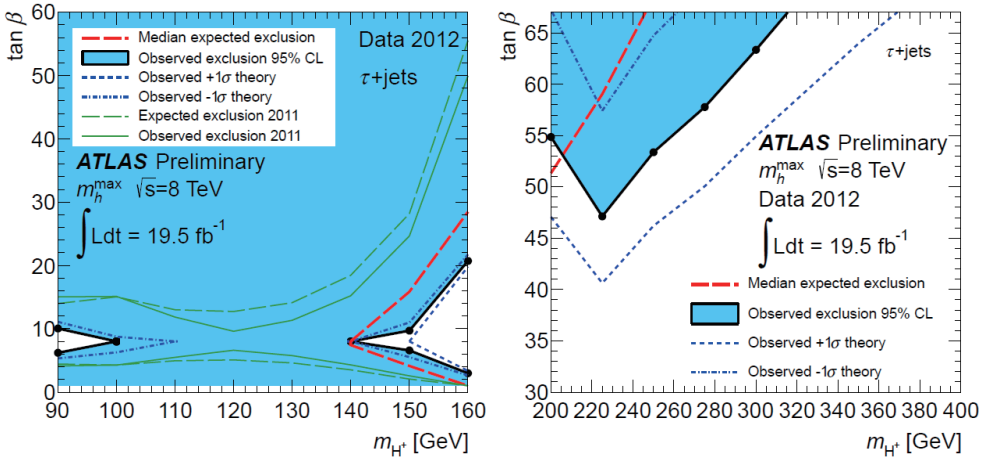


Figure 2.6: Constraints of the light $H^{\pm}$ and the heavy $H^{\pm}$, in the context of the MSSM m_h^{max} scenario with μ=200 GeV.

2.4.3 Other Observables

CP Asymmetry The direct CP asymmetry is important observable for NP search. It is defined as

$$A_{CP} = \frac{\Gamma(b \to s\gamma) - \Gamma(\bar{b} \to \bar{s}\gamma)}{\Gamma(b \to s\gamma) + \Gamma(\bar{b} \to \bar{s}\gamma)}. \tag{2.17}$$

The next-to-leading order expression for the asymmetry can be found in Reference [31].

$$\begin{aligned} A_{CP} &= \frac{\alpha_s(m_b)}{|C_7|^2} \left\{ \frac{40}{81} \mathrm{Im}[C_2 C_7^*] - \frac{8z}{9}[v(z) + b(z,\delta)] \mathrm{Im}[(1+\epsilon_s) C_2 C_7^*] \right. \\ &\quad \left. - \frac{4}{9} \mathrm{Im}[C_8 C_7^*] + \frac{8z}{27} b(z,\delta) \mathrm{Im}[(1+\epsilon_s) C_2 C_8^*] \right\}, \end{aligned} \tag{2.18}$$

where $z = \left(\frac{m_c}{m_b}\right)^2$ and we define a quatity δ by the requirement that $E_\gamma > (1-\delta) E_\gamma^{max}$, i.e. δ is the fraction of the spectrum above the cut.

$$\epsilon_s = \frac{v_u}{v_t} = \frac{V_{us}^* V_{ub}}{V_{ts}^* V_{tb}} \sim \lambda^2 (i\eta - \rho) = \mathcal{O}(10^{-2}). \tag{2.19}$$

In the last step, ϵ_s is expressed in terms of the Wolfenstein parameters, with $\lambda = \sin\theta_c \sim 0.22$ and $\rho, \eta = \mathcal{O}(1)$. The $v(z)$ and $b(z)$ are defined as

$$\begin{aligned} v(z) &= \left(5 + \ln z + \ln^2 z - \frac{\pi^2}{3}\right) + \left(\ln^2 z - \frac{\pi^2}{3}\right) z + \left(\frac{28}{9} - \frac{4}{3}\ln z\right) z^2 \\ &\quad + O(z^3), \end{aligned} \tag{2.20}$$

$$b(z) = g(z,1) - g(z,1-\delta), \tag{2.21}$$

$$\begin{aligned} g(z,y) &= \theta(y-4z)\left\{(y^2 - 4yz + 6z^2)\ln\left(\sqrt{\frac{y}{4z}} + \sqrt{\frac{y}{4z} - 1}\right)\right. \\ &\quad \left. - \frac{3y(y-2z)}{4}\sqrt{1 - \frac{4z}{y}}\right\} \end{aligned} \tag{2.22}$$

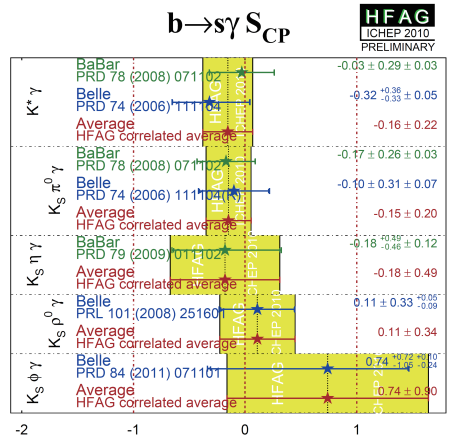
In the SM, the Wilson coefficients take the real values. The imaginary part of the small quantity ϵ_s is thus the only source of CP violation. Note that all terms involving this quantity are GIM suppressed by a power of the small ratio $z = \left(\frac{m_c}{m_b}\right)^2$. Hence, the SM prediction for the CP asymmetry is suppressed by three small factors: $\alpha_s(m_b)$ arising from the strong phase, $\sin\theta_c^2$ reflecting the CKM suppression and $\left(\frac{m_c}{m_b}\right)^2$ resulting from the GIM suppression. As a result, the SM predicts a tiny asymmetry of $\mathcal{O}(0.5\%)$ and is not very sensitive to the choice of the photon-energy cutoff[31]. In NP models which have contributions to C_7 and C_8 such that the ratio C_7/C_8 has a non-trivial weak phase, the third term in Eq.(2.18), there is the possibility of generating large CP asymmetries. The current results, based on 152 and 383 million $B\overline{B}$ samples by Belle and BABAR, are $0.002 \pm 0.050 \pm 0.030$[34] and $-0.011 \pm 0.030 \pm 0.014$[35], respectively, and have been averaged by PDG 2013 to be

$$\mathcal{A}_{\mathcal{CP}} = -0.008 \pm 0.029. \tag{2.23}$$

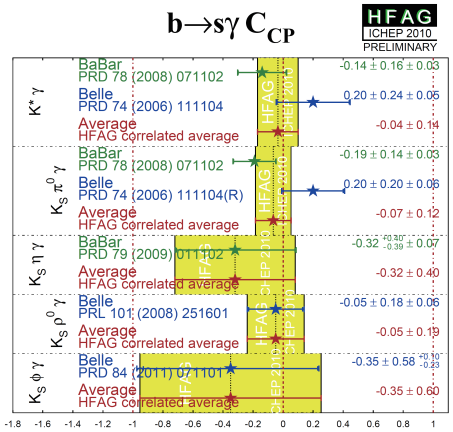
This result is consistent with the SM prediction.

Time-dependent CP Asymmetry In the SM, the emitted photon in $b \to s\gamma$ decays are predominantly left-handed, namely right-handed contribution is suppressed by $C_7' = \frac{m_s}{m_b} C_7$. Large contributions from right-handed photon indicate the NP. The branching fraction is not sensitive to C_7' since it measures only $|C_7|^2 + |C_7'|^2$. To extract the ratio C_7'/C_7 the time-dependent CP asymmetry for $b \to s\gamma$ is useful. The CP asymmetry is suppressed, thus, the expected mixing-induced CP asymmetry parameter(S_{CP}) is $\mathcal{O}(3\%)$[32, 33]. This measurement faces two experimental challenges. First, the modes and statistics that can be used for the measurement are rather limited. Second,

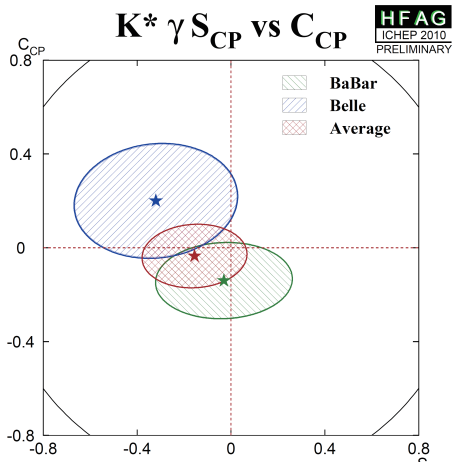
the B meson decay vertex position has to be extrapolated from the displaced $K_s \to \pi^+\pi^-$ vertex and the K_s^0 momentum vector. Therefore, the K_s decays inside the vertex detector volume(55% in Belle, 68% in BaBar) and the resulting vertex resolution is degraded. The measurements for the time-dependent CP asymmetry of $b \to s\gamma$ are summarized in Figure 2.7[20]. All results are compatible with null asymmetry with errors that are not still small enough to provide constraints on right-handed currents. This measurement will be a good probe for the NP search in future experiments, such as BelleII.



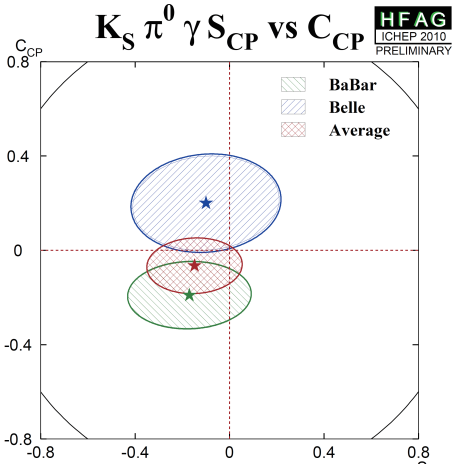
(a) Mixing-induced CP asymmetry S_{CP}



(b) Direct CP asymmetry C_{CP}



(c) S_{CP} vs C_{CP} ($B \to K^*\gamma$)



(d) S_{CP} vs C_{CP} ($B \to K_s\pi^0\gamma$)

Figure 2.7: Measurement for the time-dependent asymmetries

Isospin Asymmetry Another important observable is the isospin asymmetry given by

$$\Delta_{0+} = \frac{\Gamma(B^+ \to X_s^+\gamma) - \Gamma(\overline{B}^0 \to X_s^0\gamma)}{\Gamma(B^+ \to X_s^+\gamma) + \Gamma(\overline{B}^0 \to X_s^0\gamma)}, \tag{2.24}$$

where the partial decay rates are CP-averaged. In the SM, the prediction is as small as $\mathcal{O}(5\%)$ due to spectator quark dependent effects. The measurement is especially sensitive to NP effects in the penguin sector, namely to the ratio of the two effective coupling C_6/C_7. The current world averages by PDG2013 for $B \to K^*\gamma$ are

$$\Delta_{0+}(B \to K^*\gamma) = -0.052 \pm 0.026. \tag{2.25}$$

This result is in an agreement with the SM prediction. The measurement for the inclusive $B \to X_s\gamma$ with semi-inclusive method by BaBar is

$$\Delta_{0+}(B \to X_s\gamma) = -0.006 \pm 0.058 \pm 0.026, \tag{2.26}$$

which is consistent with null asymmetry but is not yet as precise.

2.5 Goal for this thesis

In this thesis, we measure the branching fraction of $B \to X_s\gamma$ with the semi-inclusive method as described in Section 2.3. This is updated from the measurements with only 5.8 fb^{-1}[39] at the Belle, in reconstruction of more final states, wider X_s mass region measurement and an improved analysis procedure, as well as more than 100 times integrated luminosity.

Chapter 3

Belle Experiment

The Belle experiment was designed to search the nature of CP violation, which is one of the key to the origin of the universe. At that time, CP violation in only K meson decay had been observed and everyone had great expectations on discovery of CP violation in B meson decay. To perform a precise measurement of B decay the huge amount of B mesons are essential because the typical decay rate is between 10^{-3} and 10^{-6}. Therefore, two high luminosity machines, which are called "B factory", were built at the High Energy Accelerator Research Organization(KEK) in Tsukuba, Japan and the Stanford Linear Accelerator Center(SLAC) in California, U.S.. The experiment in Japan was named as "Belle", while that in U.S. was named as "BaBar". B factories produce a copious numbers of B mesons like a factory in order to study on various aspects on B meson.

KEKB and PEP-Ⅱ are electron-positron colliders of Belle and BABAR, respectively, whose center of mass energy is adjusting at $\Upsilon(4s)$ resonance, 10.58 GeV. Both accelerators in Belle and BABAR, respectively, have an asymmetric beam energy to produce boosted B mesons for the measurement of the time-dependent CP violation.

The clean experimental condition of the Belle makes it possible to study the decay processes involving neutrinos and inclusive decays which can not be measured at hadronic machines: e.g. $b \to s\gamma, b \to sll, b \to s\nu\nu, B \to \tau\nu, B \to \overline{D}^{(*)}\tau\nu$.

The Belle started to collect data in June, 1999 and finished in June, 2010. We achieved the world record peak luminosity, 2.1×10^{34} cm^{-2}s^{-1} and have accumulated 1040 fb^{-1}.

3.1 The KEKB accelerator

KEKB(Figure 3.1) is a two-ring asymmetric energy e^+e^- collider to produce a huge number of B and $\overline{B}$ meson pairs. To produce largely boosted B meson for the time-dependent CP violation study, e^+ and e^- beam energies must be asymmetry. In that case these beams cannot have the same orbit under common magnetic field and thus KEKB is designed to have two separate rings for the e^+ and the e^- beams. The e^+ and the e^- beams are injected directly into the main rings at Fuji area from a linear accelerator. The e^+ beam , which is called the Low Eenergy Ring(LER), circulates anti-clockwise with energy $E^+ = 3.5$ GeV and the e^- beam, which is called the High Energy Ring(HER), circulates clockwise with energy $E^+ = 8.0$ GeV. The KEKB has two crossing points at Tsukuba and Fuji experimental hall. The beams collide at the interaction points(IP) in the Tsukuba hall, where the Belle detector is located. On the other hand two beams are displaced vertically and pass through each other at the Fuji hall.

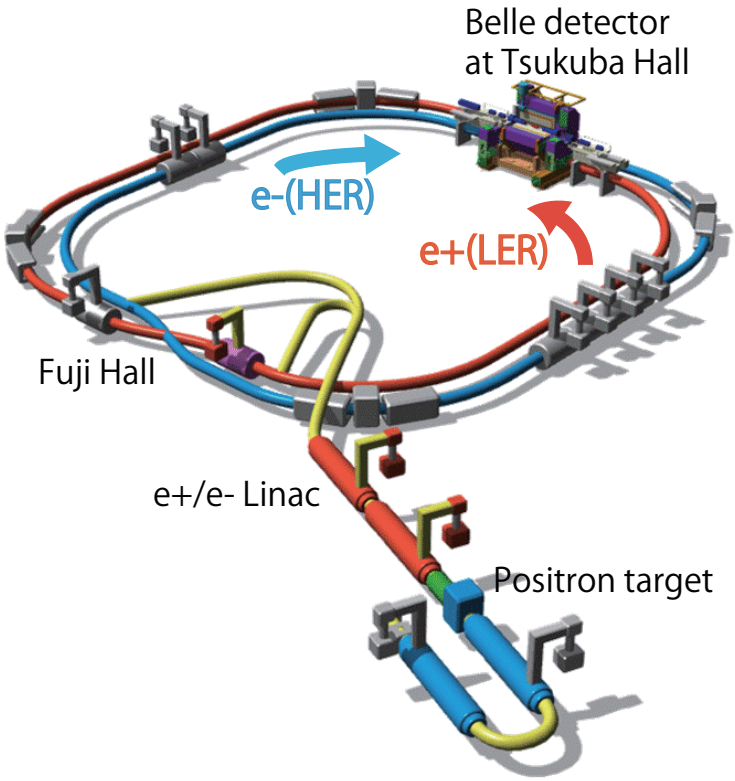


Figure 3.1: KEKB accelerator

The center of mass energy at the IP is $\sqrt{s} = \sqrt{4E^+E^-} = 10.58$GeV, which corresponds to the mass of $\Upsilon(4S)$ that decays into $B\overline{B}$ pair with more than

96% of the branching fraction. The cross-sections in e^+e^- collision at the $\Upsilon(4S)$ resonance are summarized in Table 3.1. The $B\overline{B}$ cross-section is about 1.1 nb. The $\Upsilon(4S)$ resonance stand on top of large continuum background coming from light-quark pair production($e^+e^- \to q\overline{q}(q = u, d, s)$).

Table 3.1: Cross-section in e^+e^- collision at $\sqrt{s} = 10.58$ GeV. QED refers to Bhabha and radiative Bhabha process

Process	σ[nb]
$B\overline{B}$	1.1
$c\overline{c}$	1.3
$q\overline{q}(q = u, d, s)$	2.1
$\tau\overline{\tau}$	0.93
QED($25.551^\circ < \theta < 159.94^\circ$)	37.8
$\gamma\gamma$	11.1

For measurement of time-dependent CP asymmetry, the distance of the decay vertices(Δz) of the B meson pairs is measured with asymmetric beam energy, instead of the difference of the decay time(Δt) from $\Delta z \sim c\beta\gamma\Delta t$. The Lorentz boost parameter is

$$\beta\gamma = \frac{E^- - E^+}{\sqrt{s}} = 0.425. \tag{3.1}$$

As a result a boosted B meson runs about $0.425 \times c\tau(B^0 : 455\mu\text{m}) \sim 200\mu\text{m}$ in average.

The design luminosity of KEKB is 1.0×10^{34} cm^{-2}s^{-1}. The collider achieved the goal in May 2003 and finally, the luminosity reached

$$L = 2.11 \times 10^{34}\text{cm}^{-2}\text{s}^{-1} \tag{3.2}$$

which is the world's highest luminosity in e^+e^- colliders and corresponds to an approximate production rate of 20 $B\overline{B}$ pairs per second.

In early 2004, a new method of operation at KEKB called "continuous injection mode" was successfully introduced which removes the dead time of the ordinary injection method. It allowed the data taking to avoid to stop every hour to replenish the beams in the storage ring. In early 2007, a new instrument called a "clab" cavity was installed. In the original design, the two beams do not collide head-on, but with a small crossing angle of ±11 mrad. It has the advantage, at some cost on the luminosity, of simplifying the design of the interaction region and reducing the background in the detector. To cope

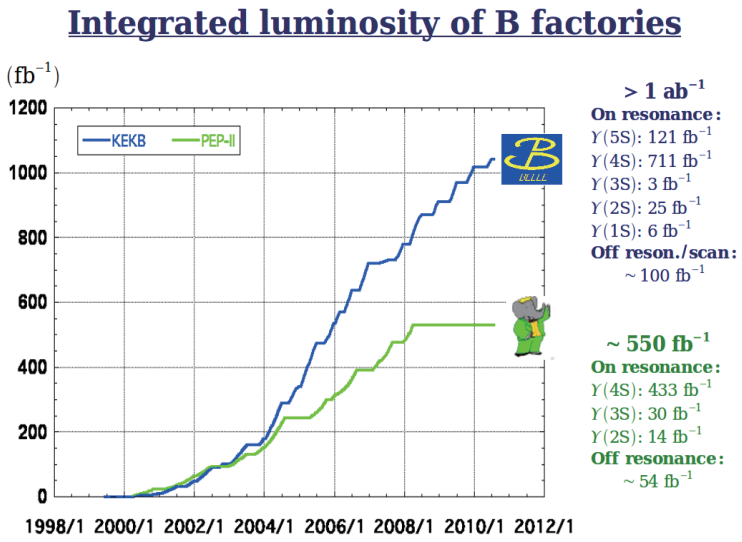


Figure 3.2: Integrated luminosity of Belle(blue) and BaBar(green)

with the luminosity loss, the bunches are tilted by a crab cavity in each ring, to the bunches with a maximum overlap.

Figure 3.1 shows the history of the integrated luminosity. The total integrated luminosity reached 1000 fb^{-1}, which is one of the primary targets of the KEKB projects, by finishing the data taking in June 2010.

3.2 The Belle detector

In the Belle detector(Figure 3.2)[40], B meson decay vertices are measured by a silicon vertex detector(SVD) located just outside of a beryllium beam pipe. A charged particle tracking from B decay is provided by a central drift chamber(CDC) and the SVD. Particle identifications can be achieved by the dE/dx measurements in the CDC and an aerogel cherenkov counter(ACC) and a time-of-flight counters(TOF) situated just outside of the CDC. An electromagnetic shower is detected in a electromagnetic calorimeter(CsI(Tl) crystals) located inside the solenoid coil. Muon and K_L mesons are identified by arrays of resistive plate counters interspersed in the iron yoke.

A major detector upgrade was performed in the summer of 2003. A 3-layer SVD with a 2 cm radius beam-pipe was used and a data sample corresponding to a integrated luminosity of 150 fb^{-1}(DS-I) was collected with this configuration. After the upgrade, a 4-layer SVD, a 1.5 cm radius beam-pipe and a small-cell inner drift chamber were installed. A data sample corresponding to the integrated luminosity of 850 fb^{-1}(DS-II) was collected with this configu-

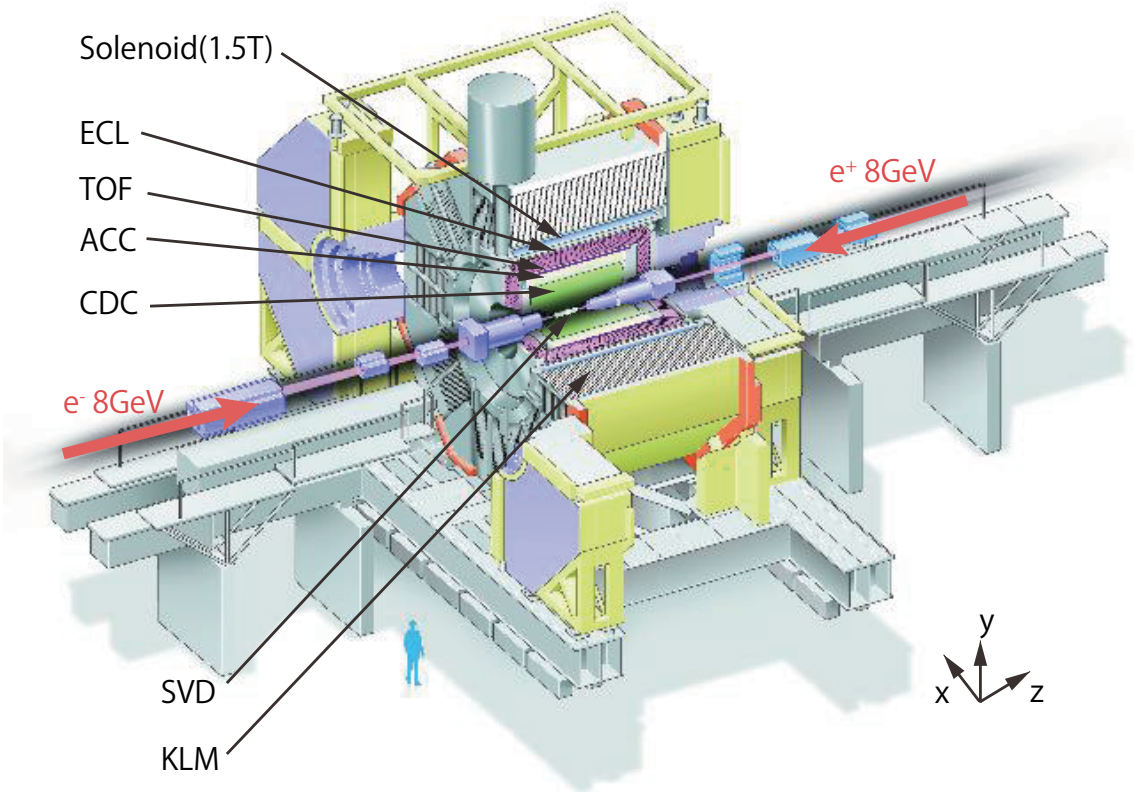


Figure 3.3: Belle detector

ration.

The origin of the coordinate system is defined as the position of the nominal IP. The z axis is aligned with the direction of the electron beam and is parallel to the direction of the magnetic field within the solenoid. The x axis is horizontal and points towards the outside of the storage ring and the y axis is vertical. The polar angle θ and azimuthal angle ϕ are measured relative to the positive z and x axes, respectively. The radial distance is defined with $r = \sqrt{x^2 + y^2}$.

The performance parameters of the Belle detector are summarized in Table 3.2. The following subsections provide a detailed description of sub-detectors.

Table 3.2:
Performance parameters for the Belle detector

Detector	Type	Configuration	Readout	Performance
Beam-pipe(DS-I)	Beryllium double wall	Cylindrical r=20mm 0.5/2.5/0.5 mm=Be/He/Be He gas for cooling		
Beam-pipe(DS-**II**)	Beryllium double wall	Cylindrical r=15mm 0.5/2.5/0.5 mm=Be/Paraffin/Be Paraffin gas for cooling		
SVD(DS-I)	Double-sided Si strip	3-layers:8/10/14 ladders Strip pitch:25(p)/50(n) μm	ϕ: 40.96k z:40.96k	
SVD(DS-**II**)	Double-sided Si strip	4-layers:6/12/18/18 ladders Strip pitch:75(p)/50(n)(layer1-3) μm 73(p)/65(n)(layer4) μm	z: 55.296k	
CDC(DS-I)	Small cell drift chamber	Anode: 50 layers Cathode:3 layers r=8.3-87.4cm -78.72cm$\leq$z$\leq$158.77cm	Anode:8.4k Cathode:1.8k	$\sigma_{r\phi} = 130\mu m$ $\sigma_z = 200 \sim 1400\mu m$ $\sigma_{pt}/p_t = 0.3\%\sqrt{p_t^2+1}$ $\sigma_{dE/dx} = 8\%\sqrt{p_t^2+1}$
CDC(DS-**II**)	Small cell drift chamber	Anode: 49 layers Cathode:no layer r=10.4-87.4cm -78.72cm$\leq$z$\leq$158.77cm	Anode:8.5k	$\sigma_{r\phi} = 130\mu m$ $\sigma_z = 200 \sim 1400\mu m$ $\sigma_{pt}/p_t = 0.3\%\sqrt{p_t^2+1}$ $\sigma_{dE/dx} = 8\%\sqrt{p_t^2+1}$
ACC	Silica aerogel (n=1.01-1.03)	$\sim \times 12 \times 12\text{cm}^3$blocks 960 barrel/228 endcap FM-PMT readout	1788	$N_{p.e.} \geq 6$ K/π separation at 1.2$<p<$3.5GeV
TOF	Plastic scintillator	128/64 ϕ segmentation r=120cm,L=3m	128$\times$2/64	$\sigma_t = 100ps$ $K/\pi separation upto 1.2GeV$
ECL	CsI	r=125-162cm(Barrel) z=-102cm(Backward Endcap) z=+196cm(Forward Endcap)	6624(B) 960(BE) 1152(FE)	$\sigma_E/E = 1.3\%/\sqrt{E}$ $\sigma_{pos} = 0.5cm/\sqrt{E}$ E in GeV
Magnet	Super-conducting	Inner radius=170cm		B=1.5T
KLM	Resistive plate counters(RPC)	14 layers(5cm Fe+4cm gap) 2RPCs for each gap θ and ϕ strips	θ:16k,ϕ:16k	$\Delta\phi = \Delta\theta = 30$mrad for K_L $\sigma_t = 1ns$ $\sim 1\% hadron fakes$
EFC	BGO	$\sim 12 \times 12 cm^3 blocks$ Photodiode readout Segmentation:32(ϕ),5(θ)	100$\times$2	Energy resolution(rms): 7.3% at 8GeV 5.8% at 8GeV

3.2.1 Silicon Vertex Detector(SVD)

The Silicon Vertex Detector(SVD) is placed at the most inner part of the Belle detector(just outside of the beam pipe) and provides very precise position measurement for B meson vertex reconstruction, which is crucially important for the time-dependent CP asymmetry in the B meson decay. In the analysis, it is essential to measure the difference in z-vertex positions for B meson pair with a precision of 100 μm. The SVD also provides the information to reconstruct D and τ decay vertices and contributes to the tracking.

Since most particles of interest in the Belle have momenta of about 1 GeV/c or less, the vertex resolution is dominated by the multiple-Coulomb scattering. This imposes strict constraints on the design of the detector. In particular, the innermost layer of the vertex detector must be placed as close to the interaction point as possible. A support structure must be low in mass, but rigid and readout electronics must be located outside of the tracking volume. The design must also withstand large beam background. With the high-luminosity operation of the KEKB, the radiation dose to the detector due to beam background was expected to be 30 krad/yr at the full design current.

A sensor of the SVD is Double-sided Silicon Strip Detector(DSSD). A charged particle passing through the junction liberates electrons from the valence band into the conduction band creating electron-hole pairs($e^- p^+$). These pairs create currents in the p^+ and n^+ strips located on the surface of the DSSD. The p^+ strips are aligned along the beam axis and measure the azimuthal angle. The n^+ strips are aligned perpendicularly to the beam axis and measure z position. The DSSD are originally designed for the DELPHI micro-vertex detector[41] and fabricated by Hamamatsu Photonics. The readout chain for DSSDs is based on the VA1 integrated circuit[42]. The VA1 has excellent noise performance and reasonably good radiation tolerance of 200 krad. The back-end electronics is a system of flash analog-to-digital converters(FADCs), digital signal processors(DSPs) which perform on-line common-mode noise subtraction, data sparsification and data formatting, and field programmable gate arrays(FPGAs), mounted on standard 6U VME boards.

SVD1 Figure 3.4 shows the side and end views of the SVD1[43]. The size fo the active region is 53.5×32.0 mm^2 on the z-side and 54.5×32.0 mm^2 on the ϕ-side. The overall DSSD size is 57.5×33.5 mm^2. SVD1 consists of 102 DSSDs in total and three concentric cylindrical layers arranged to cover $23^\circ < \theta < 139^\circ$. It covers with 86 % of full solid angle. The radii of three layers are 30.0, 45.5

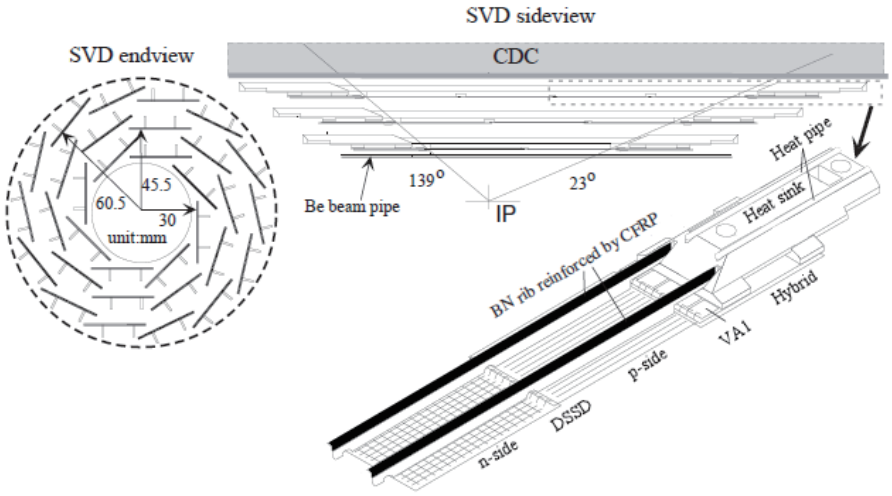


Figure 3.4: SVD1 configuration

and 60.5 mm, composed of 8,10 and 14 ladders, respectively.

SVD2 The SVD upgrade was performed in the summer of 2003 due to a limitation of SVD1, especially its radiation tolerance and non-negligible dead time. These limitations led to the design of a new detector called SVD2. The SVD2 consists of four concentric cylindrical layers and the polar angle acceptance is improved to cover $17° < \theta < 150°$ which is the same as CDC and corresponds to the 92 % of the full solid angle. The radii of four layers are 20.0, 43.5, 70.0 and 88.0 mm(the radius of beam pipe is 1.5 cm) and these layers are made up of 6, 12, 18 and 18 ladders, respectively. SVD2 consists of 138 DSSDs.

The impact parameter resolution of the SVD1 and the SVD2 for reconstructed tracks with cosmic ray events are

$$\sigma_{r\phi}(\mu m) = 19.2 \oplus 54.0/\tilde{p}, \sigma_z(\mu m) = 42.2 \oplus 44.3/\tilde{p} \quad (3.3)$$

$$\sigma_{r\phi}(\mu m) = 21.9 \oplus 35.5/\tilde{p}, \sigma_z(\mu m) = 27.8 \oplus 31.9/\tilde{p}, \quad (3.4)$$

respectively, where $\tilde{p}$ is defined as $\tilde{p} = p\beta\sin^{3/2}\theta$ for $r-\phi$ side and $\tilde{p} = p\beta\sin^{5/2}\theta$ for z side. The resolutions as a fucntion of the track momentum p are plotted in Figure 3.6. The resolutions of the SVD2 is better than them of the SVD1, mainly owing to the smaller radius of the innermost layer.

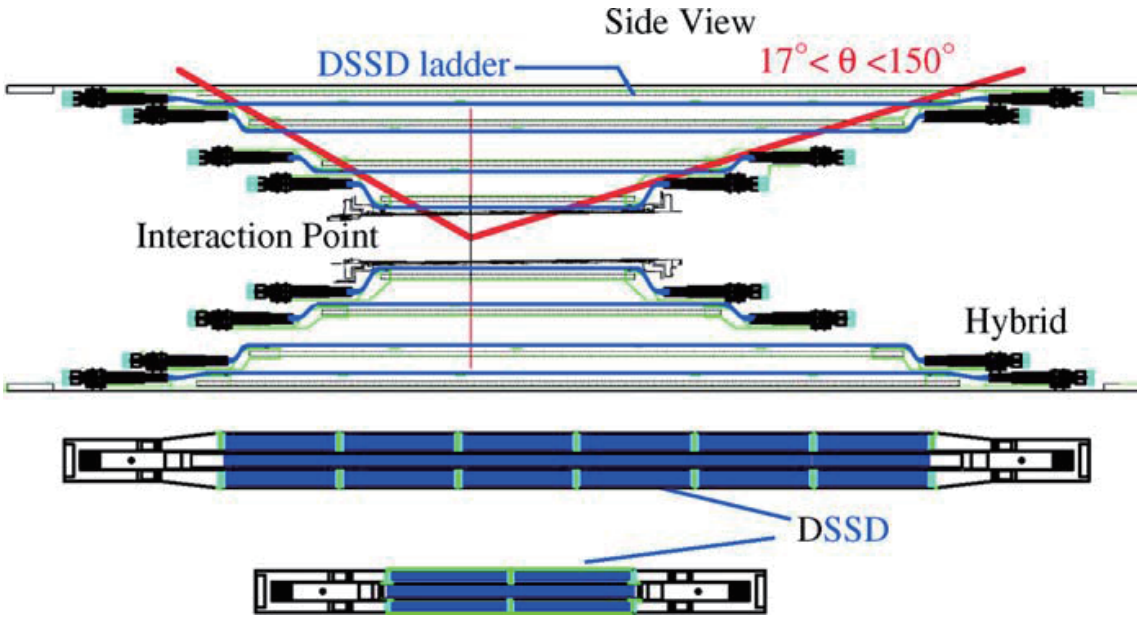


Figure 3.5: SVD2 configuration

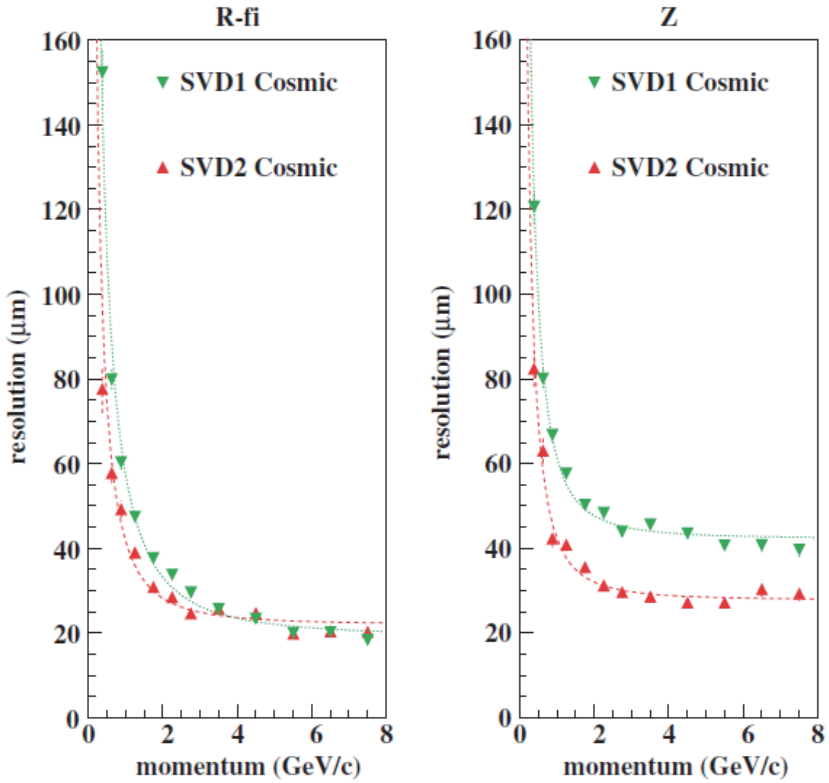


Figure 3.6: Impact parameter resolution of SVD2 and SVD1

3.2.2 Central Drift Chamber(CDC)

The Central Drift Chamber(CDC)[44] is a cylindrical wire drift chamber and one of the most important sub-detectors in the Belle detectors. The CDC is placed in a 1.5 T magnetic filed produced by the solenoid coil, therefore a charged track follows a helicoidal trajectory in the CDC. The CDC provides following three measurements.

- Precise determination of three-dimensional trajectories, providing measurements of charged particle momentum vectors
- Measurement of charged particle energy loss in the chamber gas (dE/dx) for particle identification
- Provision of fast-track information for discriminating interesting physics events at the trigger level

The structure of the CDC used to collect DS-I is shown in Figure 3.7. It is asymmetric in the z direction in order to accommodate the fact that the particles from $\Upsilon(4S)$ are boosted because of the asymmetric nature of the collider. It covers $17° < \theta < 150°$, which corresponds to 92 % of the full solid angle. The longest wires are 2400 mm long. The outer radius is 874 mm and the inner one is extended down to 83 mm. The chamber has a total of 8400 drift cells that are organized into superlayers of six axial and five small-angle-stereo which provide z position information. Each superlayers consists of between three and six radial layers, all with the same number of drift cells in azimuthal direction. The (almost) rectangular shaped cell consists of one sense wire and eight filed wires as shown in Figure 3.8. The sense wires are 30 μm diameter gold-plated tungsten. To reduce the material, the field wires are unplated aluminum. At the inner layers of the CDC, three cathode strip layers are made for higher precision z measurement

In the summer of 2003, the inner part structure of the CDC was jointly modified with upgrade of the SVD. The three inner layers with cathode strips were removed to make the space for the upgraded SVD with larger radius. Instead, two layers of smaller cells which we call small-cell CDC were installed. The inner radius after the modification is 104 mm, while the other geometry is unchanged. The small-cell CDC maintains the performance of the Level-1 trigger by keeping the number of inner layers used for the trigger to be five, which was six before the modification. In addition, we exploit the small drift time due to the smaller cell to provide additional information for the Level-0

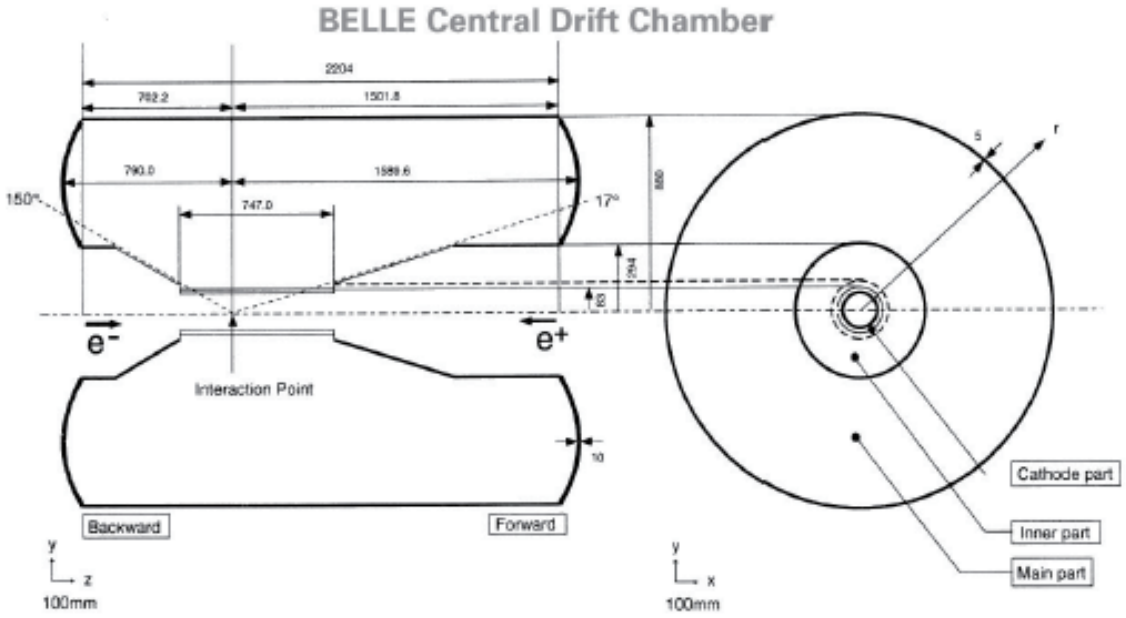


Figure 3.7: CDC configuration

trigger logic required by the SVD, which was provided by the information from the TOF alone before the upgrade.

The use of a low-Z gas is important for minimizing multiple coulomb scattering contributions to the momentum resolution. Since lowZ gases have a smaller photo-electric cross-section than argon based gases, they have an additional advantage of reducing backgound hits caused by low-energy photons from synchrotron radiation and spent particles. We use a 50% He-50% C_2H_6 gas mixture, which has a $\sim$640 m radiation length, and a drift velocity that saturates at $\sim$4 cm/μs for a $\sim$2 kV/cm electric field. This drift velocity saturation reduces the sensitivity of the distance-to-time relation function to the value of the applied high voltage and simplifies the calibration. In spite of the low-Z nature of the mixture, good dE/dx resolution is provided by the large ethane component of the gas.

The typical spatial resolution is measured to be 120 μm-150μm with dependence on the incident angles and layers. The transverse momentum, p_t, resolutions using cosmic ray data are

$$\sigma_{p_t}/p_t = \sqrt{(0.28p_t)^2 + (0.35/\beta)^2} \ (p_t \text{ in GeV/c}) \tag{3.5}$$

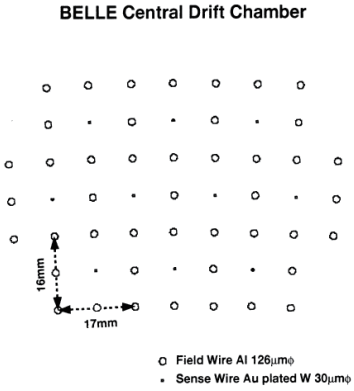


Figure 3.8: The cell structure of CDC

without the SVD information, and

$$\sigma_{p_t}/p_t = \sqrt{(0.19p_t)^2 + (0.30/\beta)^2} \ (p_t \text{ in GeV/c}) \tag{3.6}$$

with the SVD information as shown in Figure 3.9.

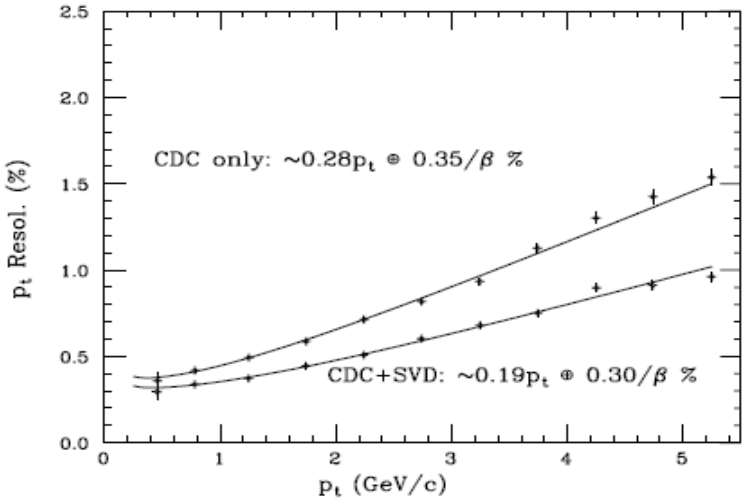


Figure 3.9: p_t resolution with cosmic ray

Using dE/dx information with a given momentum, the population of π, K, p and e are clearly separated. Figure 3.10 shows the measured dE/dx as a function of momentum, together with the expected mean energy loss for different particle species. The dE/dx resolution for minimum ionized pions from $K^0_s \to \pi^+\pi^-$ decays is measured to be 7.8 % with their momentum between 0.4 and 0.6 GeV/c.

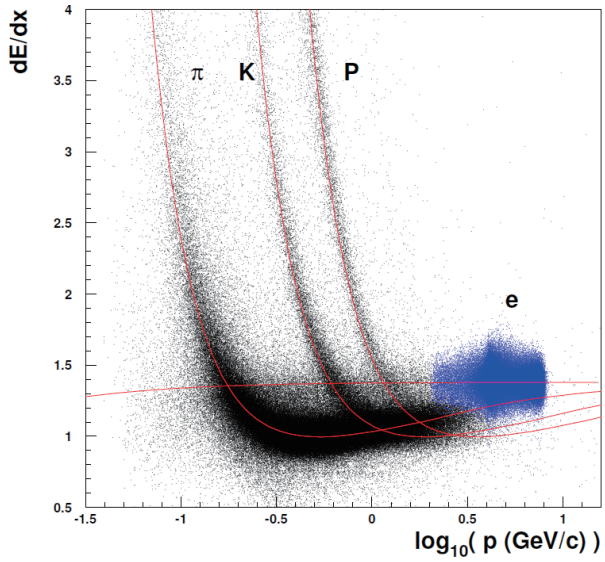


Figure 3.10: Measured dE/dx as a function of the charged track momentum with collision data

3.2.3 Aerogel Cerenkov Counter

Particle identification, in particular the identification of charged pions and kaons, plays an important role in many measurements of B decays. The momentum distribution of the final state kaons from the cascade decays ranges up to around 1.5 GeV/c. The K/π separation in this relatively low momentum region can be achieved by dE/dx measurement with the CDC in Section 3.2.2 together with a time-of-flight measurement in Section 3.2.4. On the other hand, the K/π separation up to ~4 GeV/c is required for reconstruction of the two-body decay from B meson and therefore the detector must be equipped with a device based on Cerenkov technologies. A threshold Aerogel Cerenkov Counter(ACC)[45] enable the Belle to extend the momentum coverage for the K/π separation up to 3.5 GeV/c.

The Cherenkov radiation is emitted when a charged particle passes through a material medium at a speed greater than the phase velocity of light in the medium, namely in case of $n < 1/\beta = \sqrt{1 + (m/p)^2}$, where β, m and p are the velocity, mass, and momentum of the charged particle, respectively, and n is the refractive index of the matter through which the particle is passing. Since $m_{K^\pm} > m_{\pi^\pm}$, there is a momentum region where pions emit Cherenkov light, but kaons and heavier particles do not. Thus, one can identify pions against by choosing the proper refractive index n for the momentum region of interest.

The ACC consists of blocks of silica aerogel. The silica aerogel is a transpar-

ent solid material with a colloidal form of glass that can provide a Cherenkov light. A typical module consists of five aerogel tiles in a thin aluminium box with the size of 12×12×12 cm^3. To detect the Cherenkov light, two(one) fine-mesh type photomultiplier tubes(FM-PMT) are attached to each module in the barrel(end-cap) part. The FM-PMTs are designed to operate in strong magnetic field of 1.5 T[47].

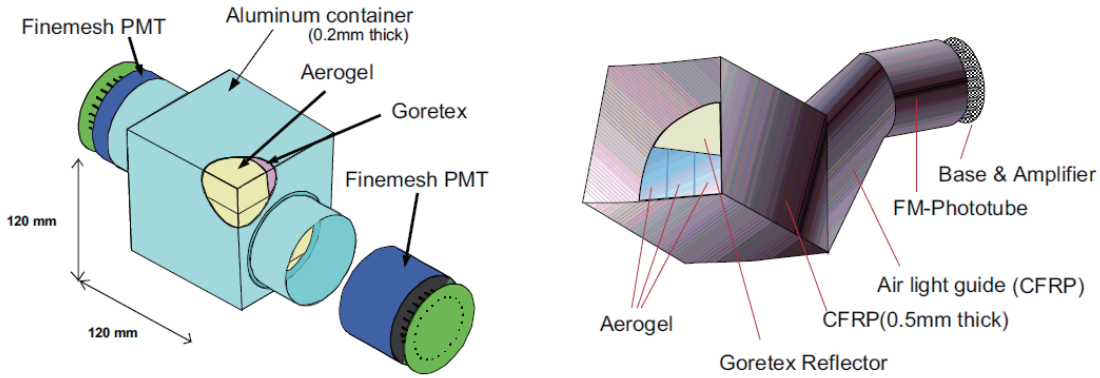


Figure 3.11: Side view of the ACC system together with other nearby detectors. The refractive indexes, n, are given for each ACC module

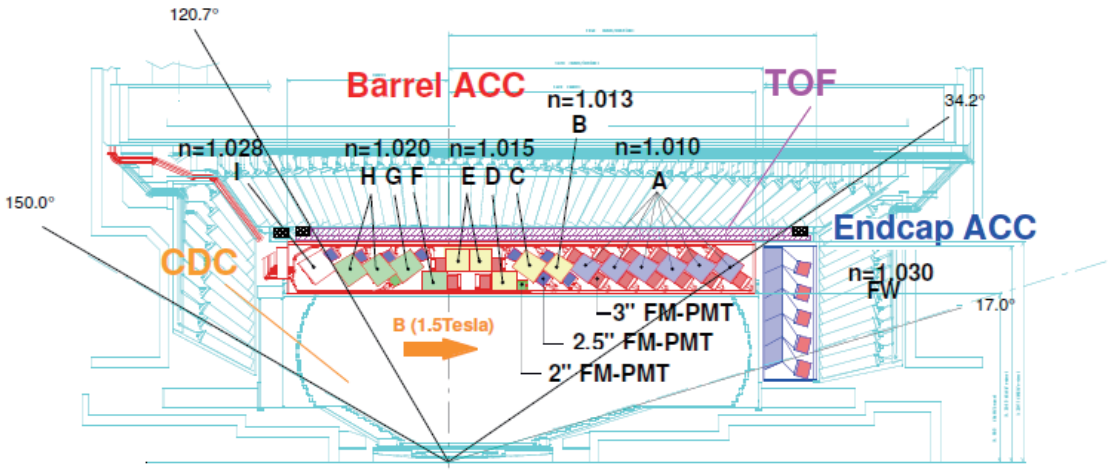


Figure 3.12: Side view of the ACC system together with other nearby detectors. The refractive indices, n, are given for each ACC module

In order to achieve a good K/π separation for the required kinematic range, the refractive indices of aerogels are selected to be between 1.01 and 1.03, depending on their polar angle region. In barrel region, they are optimized for the momentum of the daughter particles of B meson two-body decay, while in the end-cap region, they are optimized for the momentum of $K^{\pm}$ from B

cascade decay to improve a performance of B flavor tagging.

The performance of the ACC is evaluated using a decay process of $D^{*+} \to D^0\pi^+, D^0 \to K^-\pi^+$, where an identification of the charged particles from the D^0 decay can be determined without using the ACC information by the charge of π from D^{*+} decay. Figure shows numbers of photo-electron from $\pi^\pm$ and $K^\pm$ in this decay, where $\pi^\pm$ is well separated from $K^\pm$, being consistent with MC.

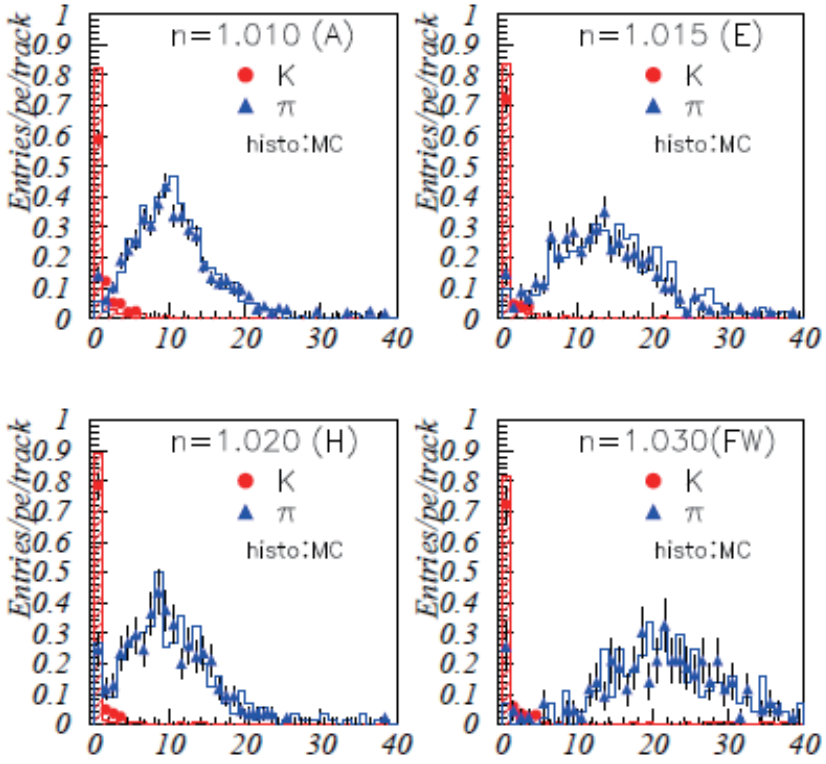


Figure 3.13: Distribution of photo-electron for $\pi^\pm$ and $K^\pm$ in D^{*+} decay

3.2.4 Time-Of-Flight Counter(TOF)

The Time-Of-Flight Counter(TOF)[46] provides $K^\pm/\pi^\pm$ separation for particle momentum below 1.2 GeV/c. Furthermore, it provides fast timing signals for the trigger system.

The mass of the particle m can be determined from the time-of-flight T measured with the TOF and the momentum p measured with the CDC as follows:

$$T = \frac{L}{c\beta} = \frac{L}{c}\sqrt{1 + (\frac{m}{p})^2}, \tag{3.7}$$

where L is a length of the flight.

For example, when L=120 cm and p=1.2 GeV/c, T=4.0 ns for a pion($m_{\pi^\pm}$ = 140 MeV/c^2), while T=4.3 ns for a kaon($m_{K^\pm}$ = 494MeV/c^2). The difference of T between pions and kaons is ∼300 ps, thus, $K^\pm/\pi^\pm$ separation with 3σ significance can be obtained with the time resolution of 100 ps.

The Belle TOF system consists of 64 modules and each module includes two trapezoidal TOF counters and one Trigger Scintillation Counters(TSC) counter(128 TOFs and 64 TSCs in total) as shown in Figure 3.14. TSC is a thin scintillation counter to provide the fast timing signal for the Belle trigger system. The TOF modules are located at a radius of 1.2 m from the IP covering a polar angle range from 34° to 120°. Fine-mesh photomultiplier tubes(FM-PMTs) are attached to both ends of the TOF counter with air gaps of 0.1 mm. As for the TSC counters, the FM-PMTs are glued to the light guides at the backward ends.

Figure 3.15(a) shows the timing resolution for forward and backward PMTs and for weighted average time as a function of the z positing on a TOF counter using $e^+e^- \to \mu^+\mu^-$ decay. The resolution for the weighted average time is about 100ps with a small z dependence, which satisfies the design goal. Figure 3.15(b) shows the mass distribution for each track in hadron events by 3.7, calculated using the momentum of the particle determined from the CDC track fit assuming muon mass. Clear peaks corresponding to $\pi^\pm$, $K^\pm$ and protons can be seen and the data points are in good agreement with a MC prediction(histogram) obtained by assuming σ_{TOF} =100 ps.

3.2.5 Electromagnetic Calorimeter(ECL)

The main purpose of the electromagnetic calorimeter(ECL)[48] is the detection of photons from B meson decays with high efficiency and good resolutions in energy and position. Since most of these photons are end products of cascade decays, they have relatively low energies and, thus, good performance below 500 MeV is especially important. However, important modes with a high energy photon, such as $b \to s\gamma$, produce energies up to 4 GeV and high resolution is needed to reduce backgrounds. Good electromagnetic energy resolution results in better hadron rejection. High momentum π^0 detection requires the separation of two nearby photons and a precise determination of their opening angle, therefore, a fine-grained segmentation is required for the ECL. Furthermore, the ECL is a main sub-detector for electron identification. In the ECL, energy from an photon or electron is deposited in electromagnetic showers produced by Bremsstrahlund and pair production, whild other charged

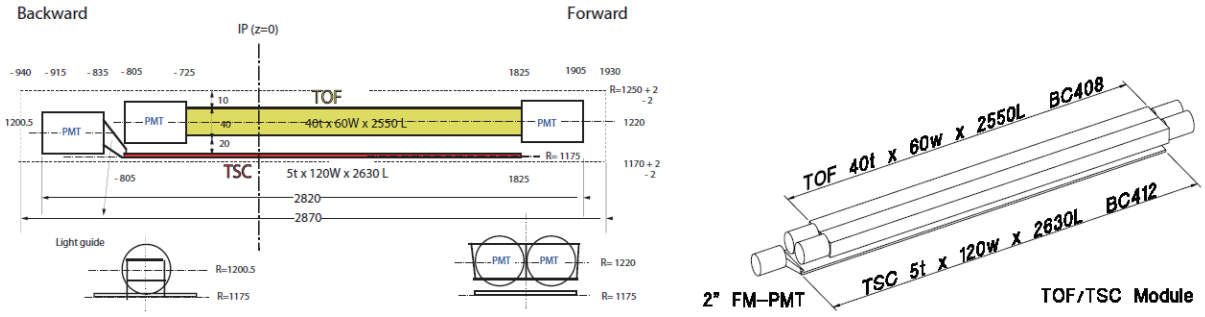
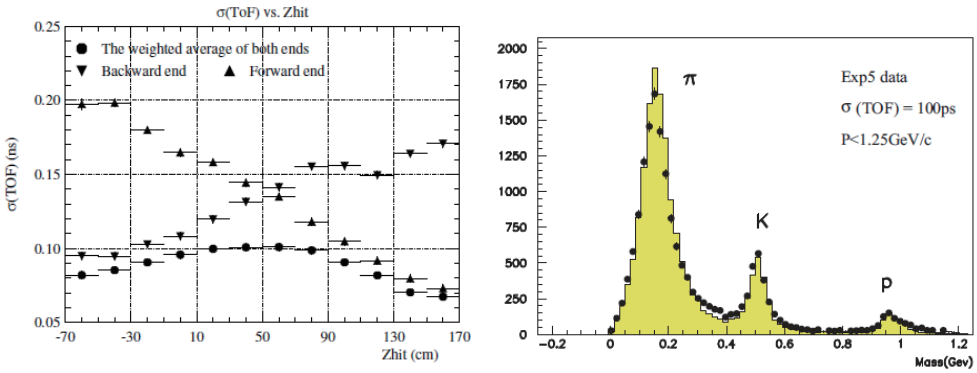


Figure 3.14: TOF module



(a) TOF timing resolution as a function of z position using $e^+e^- \to \mu^+\mu^-$ decay

(b) Mass distribution calculated from the measured time-of-flight and momentum for charged particles with momentum below 1.25 GeV/c using data.

Figure 3.15: TOF performance

particles deposit a small amount of energy by dE/dx ionization. Therefore, the ratio of the cluster energy measured by the ECL to the charged track momentum measured by the CDC, E/p is close to unity for a electron and smaller than unity for other particles.

In order to satisfy these requirements, a highly segmented array of CSI(Tl) crystals. CsI(Tl) crystals have various features such as a large photon yield, weak hygroscopicity, mechanical stability and moderate price.

Figure 3.16 shows the overall configuration of the ECL, which contains 8736 crystals. The ECL consists of three sections: the forward endcap section consists of 1152 crystals and cover $12.4° < \theta < 31.4°$, the barrel section has 6624 crystals and cover $32.2° < \theta < 128.7°$ and the backward section has 960 crystals and cover $130.7° < \theta < 155.1°$.

The size of a crystal in the $\theta - \phi$ direction is determined so that a crystal constrains approximately 80 % of the total energy deposit by a photon injected

at the center of its front face. The typical dimension of a crystal is 55 mm×55 mm at front face and 65 mm×65 mm at rear face for the barrel part. The thickness in r direction is 30 cm, which corresponds to 16.2 radiation length. This is long enough to avoid deterioration of the energy resolution at high energy due to the shower leakage. Total weight of the crystals is about 43 ton. The light of each crystal is readout by two PIN photodiodes mounted a preamplifier at the end of each crystal.

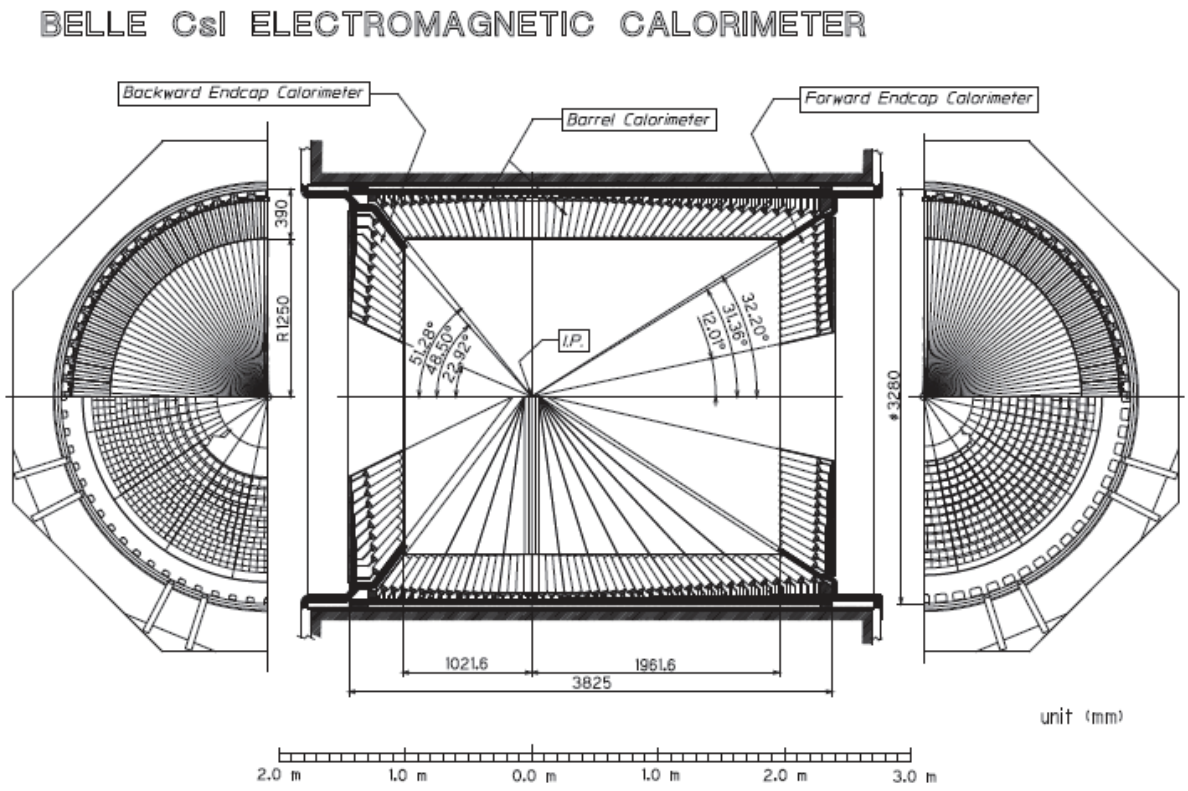


Figure 3.16: ECL configuration

The energy dependence of the average position resolution is estimated by MC and can be approximated by

$$\sigma(\text{mm}) = 0.27 + \frac{3.4}{\sqrt{E}} + \frac{1.8}{\sqrt{E}} \ (E \text{ in GeV}) \tag{3.8}$$

in Figure 3.17. This is in a good agreement with a result of beam test[49].

The energy resolution given by the beam test is

$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{0.066}{E}\right)^2 + \left(\frac{0.81}{\sqrt[4]{E}}\right)^2 + 1.34^2} \ (E \text{ in GeV}) \tag{3.9}$$

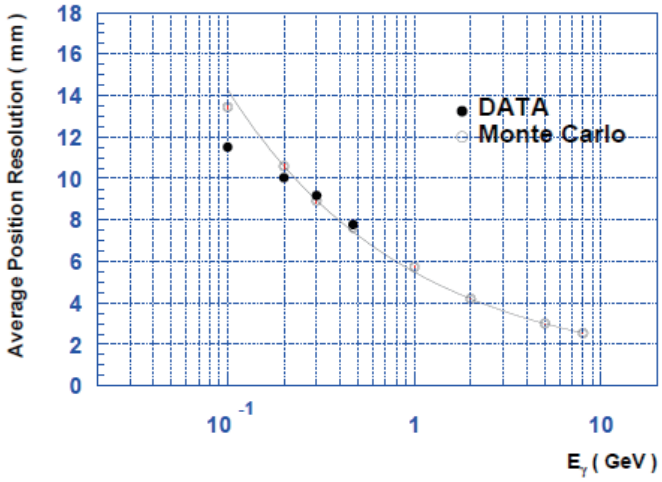


Figure 3.17: ECL position resolution as a function of energy(The solid curve is the result of fit to MC.)

3.2.6 K_L and Muon Detector(KLM)

K_L and Muon Detector(KLM)[50] is designed to identify K_L and μ with high efficiency over a broad momentum range greater than 600 MeV/c. The KLM consists alternating layers of glass-electrode resistive-plate counters(RPC) for charged particle detection and 4.7 cm -thick iron plates.

The RPCs have two parallel plate electrodes with high bulk resistivity($\geq 10^{10}\Omega$cm) separated by a gas(Argon:Butane:Freon=30:8:62)-filled gap. In the streamer mode, an ionizing particle transversing the gap initiates a streamer in the gas that results in a local discharge of the plates. This discharge is limited by the high resistivity of the plates and the quenching characteristics of the gas. The discharge induces a signal on external pickup strips, which can be used to record the location and the time of the ionization.

A K_L interacts in the iron or ECL and produces a shower of ionizing particles. The location of this shower detemines the direction of the K_L, but fluctuations in the size of the shower do not allow a useful measurement of the K_L energy. The multiple layers allow the discrimination between muons and charged hadrons($\pi^{\pm}$ and $K^{\pm}$) based on their range and transverse scattering. Muons travels much further with smaller deflections on average than strongly interacting hadrons.

There are 15 RPC layers and 14 iron layers in the octagonal barrel region and 14 layers in each of the forward and backward end-caps. The barrel region

covers angular range from 45° to 125° in the polar angle and the end caps extend the range to 20° and 155°, respectively.

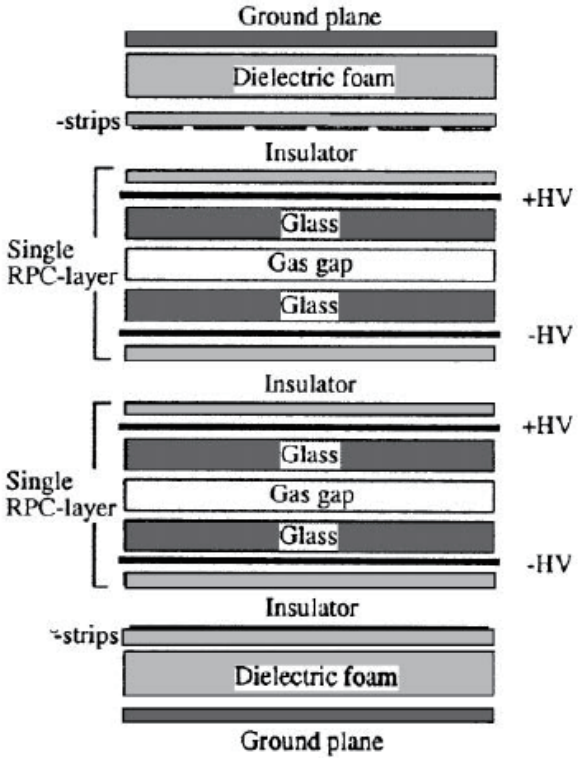


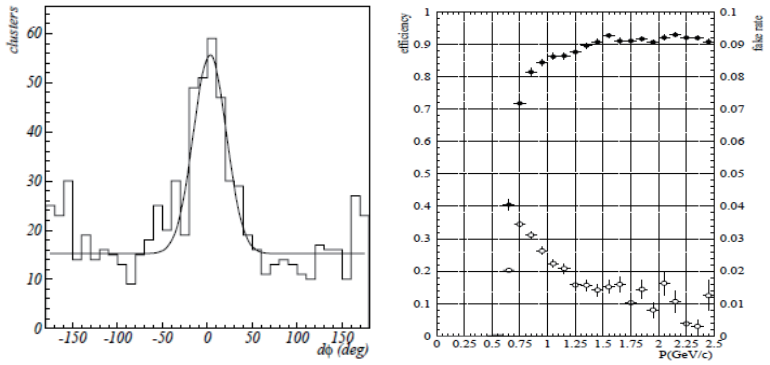
Figure 3.18: Cross section of a KLM module

Figure 3.19(a) shown a histogram of the difference between the direction of the K_L cluster candidates and the missing momentum direction which is obtained from the hadronic events. We can see a clear peak where the direction of the neutral cluster measured in the KLM is consistent with the missing momentum in the event. The angular resolution for K_L is estimated to be 0.03 mrad with the MC.

Figure 3.19(b) shows the muon detection efficiency as a function of momentum in the cosmic ray events. Below 500 MeV/c, the muon does not reach the KLM detectors. The likelihood being a muon is obtained from a comparison of the measured range of a particle with the predicted range for a muon. We have a muon identification efficiency of better than 90 % with a fake rate of less than 2 %.

3.2.7 Trigger and Data Acquisition

The total cross section shows at the design luminosity of 10^{34} cm^{-2}s^{-1} are listed in Table 3.1. Since the QED and $\gamma\gamma$ events are very large, the trigger rates



(a) Difference between the neutral cluster and the direction of missing momentum in the KLM

(b) Muon identification efficiency and fake rate as a function of momentum

Figure 3.19: KLM peformance

must be prescaled by a factor of $\sim 1/100$. In addition, high beam backgrounds are expected because of the high beam current. The trigger system is required to be robust against unexpectedly high beam background rates. The trigger conditions should be flexible so that background rates are kept within the tolerance of the data acquisition system, while the efficiency for physics events of interest is kept high.

Figure 3.20 shows the schematic view of the Belle trigger system. The trigger system consists of the sub-detector trigger systems and the central trigger system called the Global Decision Logic(GDL). The sub-detectors process signals in parallel and provide trigger information to the GDL. The GDL combines the trigger signal from each sub-detector and makes a final decision to initiate a Belle data acquisition within 2.2 μs from the event occurrence.

The global scheme of the Belle data acquisition system(DAQ) is shown in Figure 3.21. The entire system is segmented into seven subsystems running in parallel, each handling the data from a sub-detector. Charge-to-time convertesr(QTCs) and time-to-digital converters(TDCs) are used for all the sub-detectors except for the SVD and KLM. The KLM does not have the QTC since the pulse does not provide useful information. As for the SVD, the DSSDs are readout by on-board chips and processed in analog-digital converters(FADC). The readout sequence starts when the sequence controller receives a final trigger from the GDL and distributes a common stop signal to the TDCs.

Data from each subsystem are combined into a single event record by an

event builder, which converts "detector-by-detector" parallel data streams to an "event-by-event" data river. The event builder output is transferred to an online computer farm another level of event filtering is done after the fast event reconstruction. The data are then sent to a mass storage system located at the computer center.

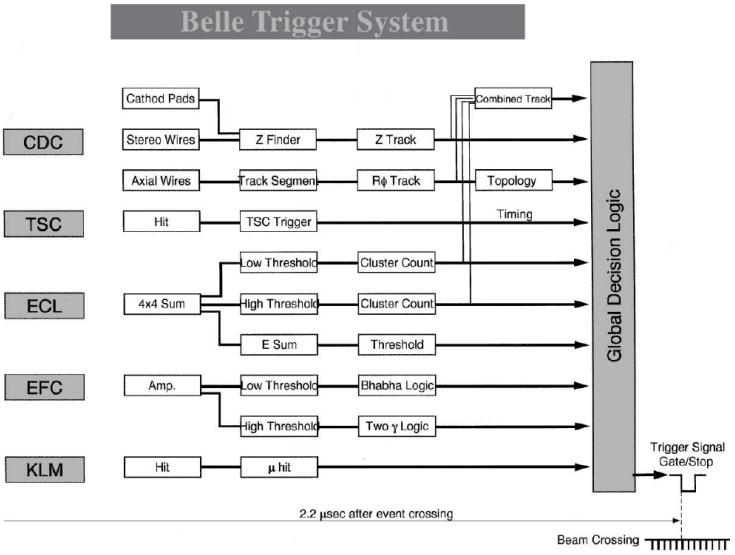


Figure 3.20: Overview of the software trigger

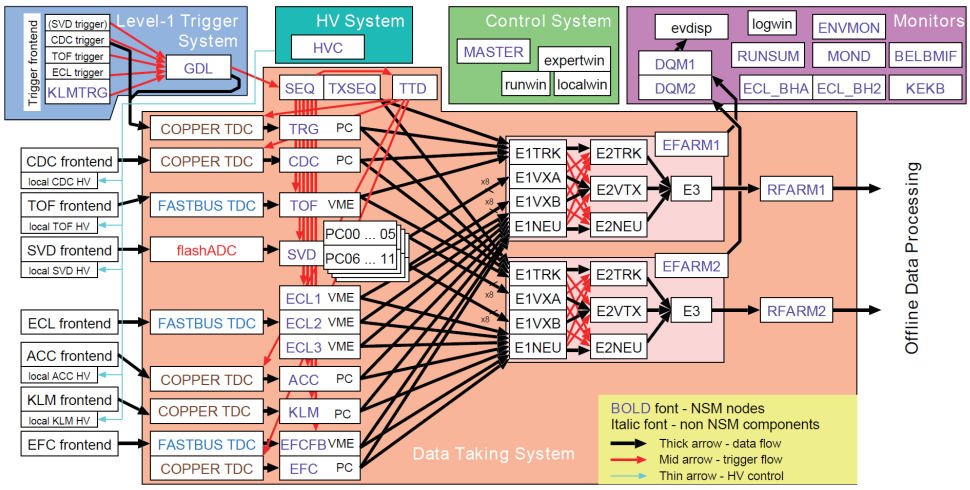


Figure 3.21: Overview of the Data Acquisition System(DAQ)

3.3 Analysis Tools

3.3.1 Software

The raw data obtained by the Belle detector are processed by the reconstruction tools, in which the tracking of the charged particles in the CDC, clustering in the ECL, and particle identifications. The output of the reconstruction is called a Data Summary Tape(DST) which is converted to a Mini Summary Tape(MDST) for analysis. In the Monte Carlo simulation, an event generator and full detector simulator are used to produce the data and the other parts in analysis are same as the real data.

An analysis for both the real data and Monte Carlo simulation in the Belle is performed in a Belle AnalysiS Framework(BASF) which was developed by the Belle collaboration and based on Framework for the Parallel Data Analysis(FPDA). The BASF is the main generic structure for the Belle analysis software and links different "module"s dynamically at run time. We provide an analysis code with a specific purpose as a module that is written as an object of a class of C++.

3.3.2 Monte Carlo Simulator

The event generator simulates physical processes of particle decay chains. The initial state is $\Upsilon(4S)$ for $B\overline{B}$ or $q\overline{q}$ and the final states consist of stable particles. We use the EvtGen[53], an event generator that is well suited for B physics and implemented many detail models. The B decay is performed by referring to the decay table that contains decay modes and branching ratios. The $q\overline{q}$ event generation uses the LUND(Pythia[19]) program, in which the subsequent hadronization process is based on the Lund string fragmentation model[54].

The full detector simulator is based on the GEANT3[55], which is a large library program developed at CERN to simulate reactions between particles an matters. This simulator takes data from the EvtGen as an input and traces the behavior of each particle in the detector, and simulates detector response.

3.3.3 $K^{\pm}/\pi^{\pm}$ separation

$K^{\pm}/\pi^{\pm}$ separation[51] in the Belle is based on three measurements;

- dE/dx measurement with CDC
- Identification by Cherenkov light with ACC
- Time of flight measurement with TOF.

These measurements cover with momentum of kaon and pion complementarily, as shown in Figure 3.22. The likelihood functions $\mathcal{L}_K$ and $\mathcal{L}_\pi$ are constructed

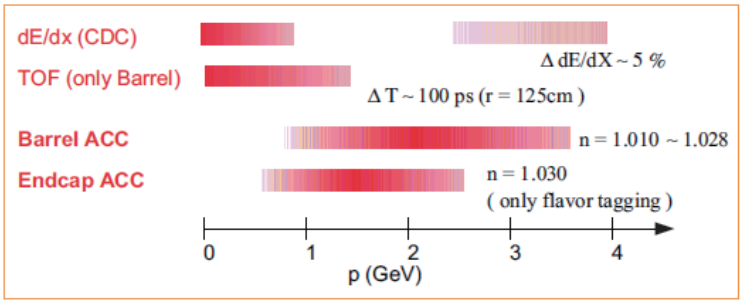


Figure 3.22: Momentum coverage of sub-detectors in $K^{\pm}/\pi^{\pm}$ separation

on the product of the likelihood function for three discriminants;

$$\mathcal{L}_i = \mathcal{L}_i^{dE/dx} \cdot \mathcal{L}_i^{ACC} \cdot \mathcal{L}_i^{TOF} \quad (i = K, \pi). \tag{3.10}$$

The likelihood ratio $\mathcal{P}_{K/\pi}$ is calculated as

$$\mathcal{P}_{K/\pi} = \frac{\mathcal{L}_\mathcal{K}}{\mathcal{L}_\mathcal{K} + \mathcal{L}_\pi} \tag{3.11}$$

As a result, $K^{\pm}$ and $\pi^{\pm}$ can be well separated with more than 3σ up to the momentum of 3.5 GeV/c. The performance of kaon identification is checked with a decay chain $D^{*+} \to D^0\pi^+, D^0 \to K^-\pi^+$. With $\mathcal{P}_{K/\pi}$ >0.6, an average kaon efficiency and π fake rate with $0.5 < p < 4.0$ GeV/c are about 88 % and 8.5 %, respectively.

3.3.4 Electron identification

The electron identification is based on a discriminant on two differences between electrons and other hadrons. First, we exploits the major difference in the electromagnetic showers induced by electrons and the hadronic showers induced by the pions and other hadrons. Second, we make use of the difference in velocity for electrons and hadrons of the same momentum. Specifically, there are following five discriminants.

- Ratio of cluster energy to track momentum
- dE/dx measurements with the CDC
- Matching between track and ECL cluster
- Cluster shape parameter

- ACC light yields

These are combined into a signal variable with a likelihood method. For each discriminants, the electron likelihood($\mathcal{L}_e$), and the non-electron likelihood($\mathcal{L}_{\bar{e}}$) are separately calculated. Each likelihood is combined using

$$\mathcal{P}_e = \frac{\mathcal{L}_e}{\mathcal{L}_e + \mathcal{L}_{\bar{e}}} \tag{3.12}$$

The electron identification efficiency in hadronic events is evaluated using single-electron MC tracks embedded in real hadronic events. With $\mathcal{P}_e$ >0.5, the efficiencies are 92 % (87 %) for the momentum region 1.0(0.5)< p <3.0 GeV and the whole polar angle range. The pion fake rate is evaluated using inclusive $K_s \to \pi^+\pi^-$ decays and 0.22 % for 0.5< p <3.0 GeV.

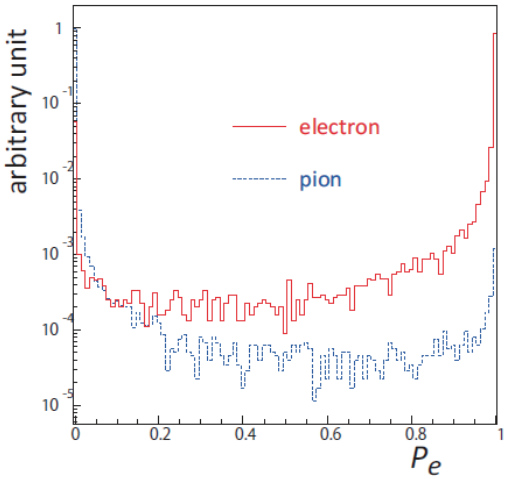


Figure 3.23: Likelihood ratio $\mathcal{P}_e$ for electron(red) and pion(blue)

3.3.5 Muon identification

Muons, which are a heavy lepton, lose the energy mainly by multiple scattering in the detector material. A muon with the momentum above 500 MeV can penetrate easily to the KLM. The KLM hits are associated to the reconstructed track by the CDC and SVD and the track is refitted. A likelihood function for the muon identification[52] is calculated based on the two discriminants;

- Difference between the expected and measured track in the KLM
- Goodness of the fit of the transverse deviations of all hits associated with the track.

The likelihood ratio of $\mathcal{P}_\mu$ is calculated as

$$\mathcal{P}_\mu = \frac{\mathcal{L}_\mu}{\mathcal{L}_\mu + \mathcal{L}_K + \mathcal{L}_\pi} \tag{3.13}$$

The performance of the muon identification is evaluated by using two-photon sample $e^+e^- \to e^+e^-\mu^+\mu^-$. The measured efficiency is 89 % for $\mathcal{P}_\mu$ >0.9 and 93 % for $\mathcal{P}_\mu$ >0.1 over 1.0< p < 3.0 GeV/c. The average fake rate is evaluated by using inclusive $K_s \to \pi^+\pi^-$ decays and is 1.4 % for $\mathcal{P}_\mu$ >0.9 and 2.8 % for $\mathcal{P}_\mu$ >0.1 over 1.5< p < 3.0 GeV/c.

Chapter 4

Data and Monte Carlo Samples

4.1 Data Sample

The data sample used in this analysis is a full data set at the Υ(4S) resonance collected by the Belle experiment, which corresponds to a integrated luminosity of 711 fb^{-1}.

Since $e^+e^- \to q\overline{q}$ background at the $\Upsilon(4S)$ resonance is three times larger than the $B\overline{B}$ decay, to study the $q\overline{q}$ background is essential. For the purpose the Belle takes data at about 60 MeV below at the $\Upsilon(4S)$ resonance. The off-resonance data collected by the Belle is totally 89.5 fb^{-1}. We use this data to evaluate a contribution from $q\overline{q}$ background.

4.1.1 Hadronic Event Selection

After data processing, events taken by the Belle are classified into several categories. Some of the categories such as Bhabha events, muon pair events and γ pair events are used for detector calibration, while hadronic events for analysis of B and charm mesons are applied a skim, called HadronBJ. The HadronBJ events are selected based on the track multiplicity and visible energy; the event must have at least three charged tracks with a transverse momentum greater than 0.1 GeV/c that originate from the vicinity of interaction point ($|\Delta r| <$ 2 cm and $|\Delta z| <$ 4 cm), and the sum of the energy of charged tracks and reconstructed photons(E_{vis}) must be greater than 20 % of $\sqrt{s}$. These selections remove the majority of beam gas background and two-photon events.

4.1.2 Number of $B\overline{B}$ pairs in Data

The number of $B\overline{B}$ pairs in the HadronBJ evcent sample is given by

$$N(B\overline{B}) = N(\text{On}) - acN(\text{Off}), \tag{4.1}$$

where N(On) is a number of events from Υ(4S)(On-resonance), and N(Off) is a number of $q\overline{q}$ background(Off-resonance). a is the scaling factor for On-resonance to Off-resonance data given by

$$\begin{aligned} a &= N(qq)(\text{On})/N(qq)(\text{Off}) && (4.2)\\ &= N(e^+e^-)(\text{On})/N(e^+e^-)(\text{Off}) && (4.3)\\ &= N(\mu^+\mu^-)(\text{On})/N(\mu^+\mu^-)(\text{Off}). && (4.4) \end{aligned}$$

We calculate 'a' using both barrel bhabha and di-muon events and take the average of the two,

$$a = 0.5(a(e^+e^-) + a(\mu^+\mu^-)). \tag{4.5}$$

The systematic uncertainty in 'a' is take as the difference

$$da = \pm 0.5|a(e^+e^-) - a(\mu^+\mu^-)|. \tag{4.6}$$

'c' is the ratio of the qq efficiency for On-resonance and Off-resonance data given by

$$c = \epsilon(\text{On})/\epsilon(\text{Off}). \tag{4.7}$$

We calculate using MC and determine the error in 'c' by comparing the event classification distributions between qq MC and off resonance data.

The totall number of $B\overline{B}$ pairs is $(771.1\pm10.6)\times10^6$.

4.2 Signal Monte Carlo Sample

In the semi-inclusive measurement, the signal modeling in the MC is significant for a precision measurement. Two types of signal MC are generated, one for the $K^*(892)$ region ($M_{X_s} < 1.15$GeV) and another for the inclusive X_s region ($M_{X_s} > 1.15$GeV). The $M_{X_s} < 1.15$GeV region is well-understood and the $K^*(892)\gamma$ is highly dominant. Therefore exclusive $K^*(892)\gamma$ signal MC is used

in this region. In the inclusive signal MC, various final states exist. The photon energy spectrum, X_s mass distribution, and breakdown of final states are not well-understood. Therefore, we take following measures to generate the inclusive MC sample in Section 4.2.1 and 4.2.2.

The exclusive $K^*(892)\gamma$ MC and the inclusive MC are mixed by the ratio of the measured branching ratio, 4.3×10^{-5} for $B \to K^*(892)\gamma$ and 3.1×10^{-4} ($\mathcal{BR}(B \to X_s\gamma) - \mathcal{BR}(B \to K^*(892)\gamma)$), respectively.

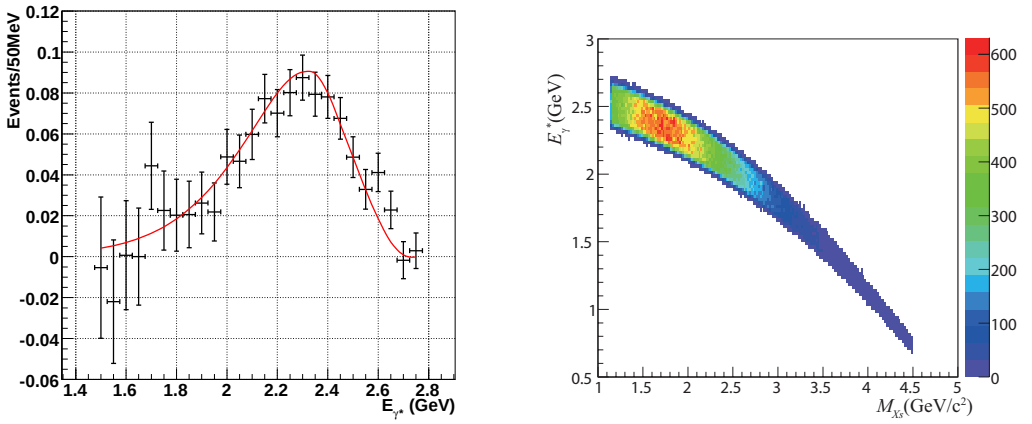
4.2.1 Photon energy and X_s mass distributions in the inclusive MC

The photon energy spectrum and X_s mass distribution are produced following a Kagan-Neubert model[56]. The Fermi motion of the b quark inside the B meson, which determines the characteristic shape to the photon energy spectrum, can be consistently described in this model. The distributions are evaluated in a next-to-leading order and has two parameters, the b quark mass, m_b, and the b quark momentum parameter in B meson, μ_π^2. We use the parameter setting which is the best fit with the photon energy spectrum in the previous Belle's result by the full-inclusive analysis(m_b=4.440 GeV/c^2, μ_π^2=0.750 GeV^2) in Figure 4.1 and the M_{X_s} distribution is shown in Figure 4.2(a).

Since a difference on the M_{X_s} shape between the MC and data occurs a large systematic uncertainty, a method to suppress the uncertainty is needed. We discuss this issue later.

4.2.2 Hadronization model in the inclusive MC

In the X_s decay of the inclusive region, the light quark pair is generated and final state hadrons are produced according to QCD theory. QCD perturbation theory, formulated in terms of quarks and gluons, is valid at short distances. At long distances, QCD becomes strongly interacting and perturbation theory breaks down. In this confinement regime, the colored partons are transformed into colorless hadrons, a process called either hadronization or fragmentation. The hadronization process is generated in Pythia[19], which is frequently used for event generation in high-energy physics. The Pythia has a huge numbers of parameters on the hadronization model. Default values in Pythia are used in the Belle, basically, but some parameters are changed according to the data information, which is summarized in Appendix B.



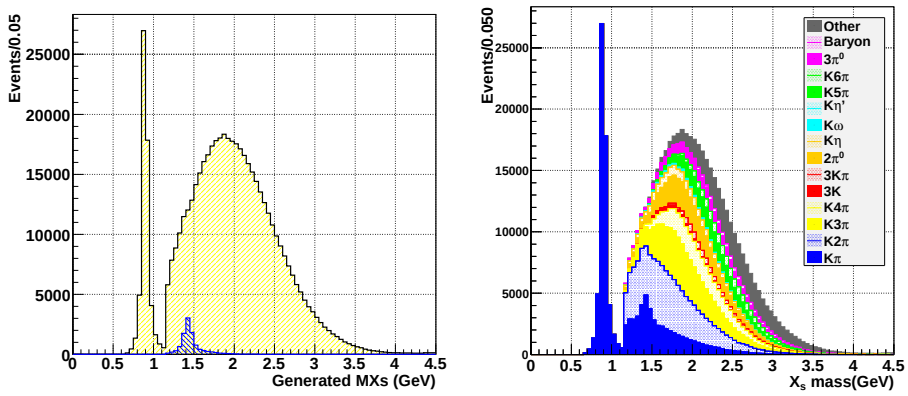
(a) The plot shows Belle's result with Full-inclusive method[71] and the line shows the best fit shape based on KN model.

(b) M_{X_s} vs E_γ in the $\Upsilon(4S)$ rest frame. Photon energy spectrum is smeared by nonzero momentum of the B meson in the $\Upsilon(4S)$ rest frame.

Figure 4.1: X_s mass and Photon energy in signal MC.

In Table 4.1, the breakdown in X_s final states are shown and the breakdown as a function of M_{X_s} is shown in Figure 4.2(b). Fractions of each mode in M_{X_s} bins is also shown in Figure 4.3. Since each quark in the X_s is hadronized separately the spin of X_s is treated as 0, however actually the spin is 1. The assumption does not have a large effect in higher multiplicity-decay than three-body, but in two-body decay, MC distributions are not necessarily correct. Therefore, $B \to K_2^*(1430)\gamma$ and $K^*(1680)\gamma$, which have a kaon with spin ≥ 1 that decay into two-body and measured or theoretically calculated branching ratio[57], are added to the inclusive MC as exclusive signal MC, in which has only two-body decay(Blue line in Figure 4.2). The inclusive MC is reweighted by the acceptance-rejection method so that the M_{X_s} distribution after adding $K_2^*(1430)\gamma$ and $K^*(1680)\gamma$ signal MC is consistent with the KN model.

A signal reconstruction efficiency depends on the particle contents in final states, and a difference on the hadronization model between the MC and data occurs a large systematic uncertainty. Thus, it is significant to understand X_s decay model in data and calibrate the hadronization model in the MC. After signal box opening, the hadronization model of the inclusive signal MC should be calibrated by the data. We discuss this issue later.



(a) X_s mass(Blue line : spin-1 kaon mode)

(b) Breakdown of X_s mass

Figure 4.2: X_s mass in signal MC.

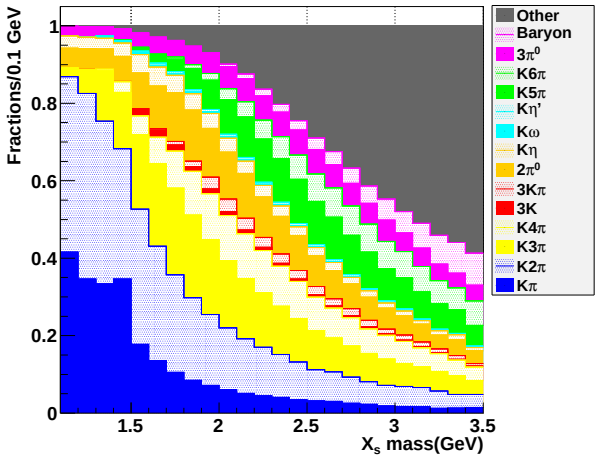


Figure 4.3: Fractions of each mode in M_{X_s} bins in signal MC(1.15< M_{X_s} <3.5 GeV/c^2).

4.3 Background Monte Carlo Sample

For the background study, we use $q\overline{q}$ and $B\overline{B}$ MC samples which corresponds to six times amount of real data. They are large enough comparing to the real data size. In the $B\overline{B}$ background samples, $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\overline{B}$ events are generated and the $B\overline{B}$ pairs decay according to branching fractions measured to date and covers most of the known channels.

Table 4.1: Fraction of each X_s final states(%)

Category	Fraction(%) (Total M_{X_s} region)	Fraction(%) ($M_{X_s} > 1.15$GeV)
$K\pi$ (at most one π^0)	15.1	7.3
$K2\pi$ (at most one π^0)	11.6	13.1
$K3\pi$ (at most one π^0)	9.0	10.1
$K4\pi$ (at most one π^0)	5.1	5.7
$3K$(at most $2K_s$)	0.5	0.6
$3K\pi$(at most $2K_s$)	0.8	0.9
$K2\pi^0$	1.0	1.2
$K\pi2\pi^0$	2.1	2.4
$K2\pi2\pi^0$	2.8	3.2
$K\eta$	0.5	0.6
$K\eta\pi$	0.6	0.6
$K\eta2\pi$	0.4	0.4
$K\eta\pi^0$	0.3	0.3
$K\eta\pi\pi^0$	0.3	0.4
$K\eta2\pi^0$	0.1	0.1
$K5\pi$(at most $2\pi^0$)	4.5	5.1
$K6\pi$(at most $2\pi^0$)	2.3	2.6
$K7\pi$(at most $2\pi^0$)	1.0	1.1
$K8\pi$(at most $2\pi^0$)	0.4	0.4
$K9\pi$(at most $2\pi^0$)	0.1	0.1
$K3\pi^0$(at most 6π)	4.3	4.9
$K4\pi^0$(at most 5π)	1.7	1.9
$K5\pi^0$(at most 4π)	0.5	0.6
$K6\pi^0$(at most 3π)	0.2	0.2
$K\omega$(with at most 4π)($\omega \to \pi^0\gamma$)	0.7	0.8
$K\eta'$(with at most 4π)($\eta' \to \rho^0\gamma$)	0.5	0.5
$K\eta3\pi$	0.7	0.8
$K\eta4\pi$	0.5	0.6
$3K2\pi$(at most $2K_s$)	0.7	0.8
$3K3\pi$(at most $2K_s$)	0.5	0.6
$3K_s$(at most 2π)	0.1	0.1
Baryon modes	1.6	1.8
K_L modes	27.2	27.4
Other	2.5	2.8

Chapter 5

Reconstruction of $B \to X_s\gamma$ with a Semi-inclusive Method

In this chapter, we use Monte Carlo (MC) simulations described in Section4.2 and subsec:bgMC to model signal and background events and to optimize the selection prior to opening the signal region in the data for ensuring no bias.

5.1 Particle Selection

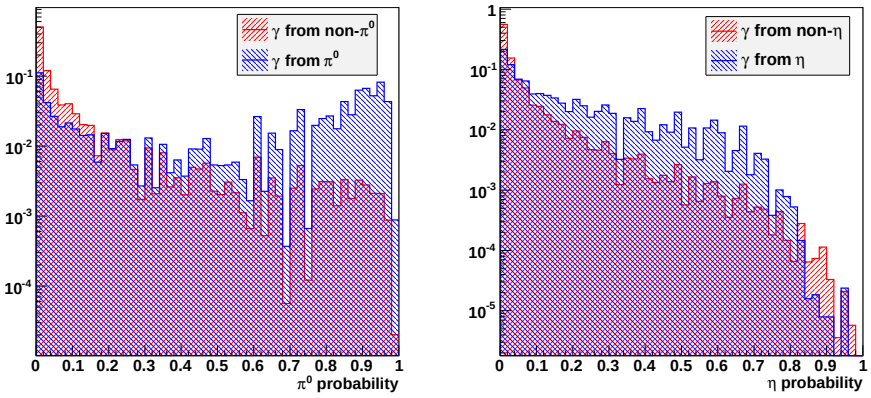
5.1.1 High-energy Photon Selection

The high energy photon is a prominent signature of the $B \to X_s\gamma$ final state. A photon is detected as an isolated energy cluster in the ECL not associated with charged tracks. We take the candidate with the energy in the CM frame between 1.8 and 3.4 GeV. The primary photon candidate is requested to be within the acceptance of the barrel ECL, $33^\circ < \theta < 132$ in order to avoid systematic uncertainty in the end cap region. This cut also suppresses a large initial state radiation background. The candidates must satisfy $E_9/E_{25} \geq 0.95$, which is the ratio of energy deposition within the central 3×3 cells to that in 5×5 around the maximum energy ECL cell of the cluster. This selection means the shower shape is consistent with a single isolated electromagnetic shower.

π^0 veto and η veto The main source of high-energy photon background is due to decays of high energy π^0, and η to a smaller extent. Because of the rapidly decreasing energy spectrum of π^0 from B decays, most high energy photons from π^0 are generated in an asymmetric decay. Consequently, the

candidate photon takes a high fraction of the π^0 energy, in contrast, the other photon(hereafter called "slow photon") is likely to be very low energy. Thus, the slow photon is difficult to disentangle from the electronic background in the ECL.

To veto these π^0 and η backgrounds, the high-energy photon candidate γ_1 is combined with any other photon γ_2 and assigned a π^0 and η probabilities"[58]. We use a veto procedure based on the π^0 and η probabilities, which are derived from a two-dimensional probability density function, invariant mass of γ_1 and γ_2, and E_{γ_2}. Figure 5.1 shows the π^0 and η probabilities. We apply a π^0 probability below 0.05 and η probability below 0.1. Furthermore, π^0 candidates with the two photon invariant mass between 117 and 153 MeV/c^2 are also rejected as an additional suppression.



(a) π^0 probability (b) η probability

Figure 5.1: π^0/η probability distributions. Red distributions are signal photons and Blue ones are backgrounds from π^0 and η.

5.1.2 Charged Particle Selection

Charged tracks are reconstructed by the CDC and SVD. The momentum of a track is calculated by using the curvature and the track is extrapolated to obtain the momentum at the closest point to the origin in the $x-y$ plane. A charged particle candidate is selected with requirements based on the distance of closest approach to the IP, $|dr| < 0.5$ cm and $|\Delta z| < 5$ cm. The track momentum p is required to be > 0.1 GeV/c in order to reduce low momentum combinatorial background. K/π separation(Section 3.3.3) is based on the technique based on the combined likelihood constructed in a way that a pion-like

track gives zero and a kaon-like track gives one. The charged tracks with the probability of greater than 0.6 are counted among kaon candidates, while the pion candidate is applied for < 0.6. We also removed electrons by rejecting tracks with electron ID >0.6(Section 3.3.4).

5.1.3 π^0 Selection

π^0 candidates are reconstructed from pairs of photons with more than 50 MeV energy in the laboratory frame to remove low energy beam background. The candidates must have an invariant mass between 125 and 145 MeV/c^2(Figure 5.2(a)). This selection of the invariant mass is tight since there are a huge low energy photon background. We also require a minimum momentum $p_{\pi^0} > 0.1$ GeV/c in the CM frame. Since the most of π^0 candidates are boosted largely, the photons from the π^0 go to the same direction as the π^0 and an angle between two photons should be small. Thus, the selection of the angle is applied, $cos\theta_{\gamma\gamma} > 0.4$.

5.1.4 η Selection

η candidates are also reconstructed from two photon candidates with more than 100 MeV energy and must have an invariant mass between 515 and 570 MeV/c^2(Figure 5.2(b)). We also require a minimum momentum $p_\eta > 0.2$ GeV/c in the CM frame. The candidate must have a helicity angle, θ_{hel}, less than 0.8 of cosθ_{hel}, which is the angle between the photon momentum and η boost direction from the laboratory frame in the η rest frame. The correct η candidate has a flat distribution of cosθ_{hel}, while a background from π^0 has a peak at cosθ_{hel}=1 since π^0 mass is smaller than that of η. Although we do not explicitly reconstruct $\eta \to \pi^+\pi^-\pi^0$ decay mode, it is implicitly included in the final states if there is at most one other pion in the event as shown in Section 5.2.

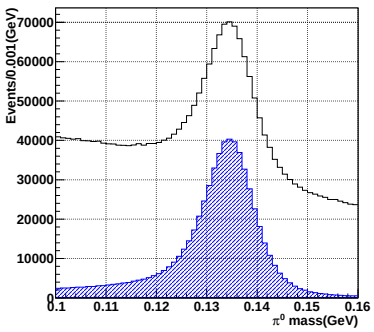
5.1.5 K_s Selection

We use a K_s selection method based on a neural network technique[65]. The following variables are used for the input;

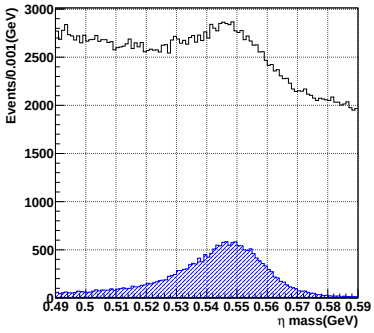
- Distance between two helices in z direction
- Flight length in $x - y$ plane
- Angle between K_s momentum and IP direction
- Shorter distance between interaction point and child helix

- K_s momentum in the lab frame
- Pion momentum from the K_s in the K_s frame
- SVD hit
- number of hit of axial wire in the CDC
- number of hit of stereo wire in the CDC.

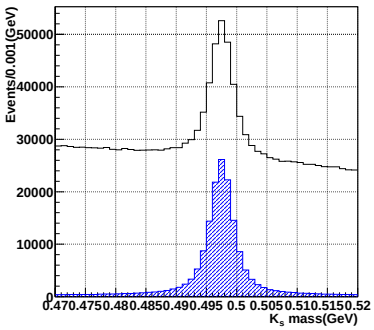
The performance is evaluated by MC and the efficiency is 87 % and the purity is 94 % for whole momentum region. Furthermore, K_s candidates are required to be $|M_{K_s}-\text{PDG mass}| < 10$ MeV/c^2(Figure 5.2(c)).



(a) π^0 mass



(b) η mass



(c) K_s mass

Figure 5.2: Mass distribution(Blue line shows correctly reconstructed particles, based on truth information.):Large tails at low side in π^0 and η mass distributions come from gamma energy resolution in the ECL.

5.2 X_s Reconstruction with a Semi-inclusive Method

5.2.1 X_s Reconstruction with a Semi-inclusive Method

We use a semi-inclusive method for X_s reconstruction, in which charged and neutral particles in an event are combined to form a X_s candidate. Ideally, all X_s final states in Table 4.1 should be reconstructed, however it is impossible since some decay rates are too small to measure(e.g. modes with ω and η') or the reconstruction efficiency is too low due to the high multiplicity in final state(e.g. $K5\pi, K6\pi$) and the amount of background is too large(e.g. modes with more than three π^0s).

It is important for the semi-inclusive method to measure as many modes as possible in order to minimize the systematic uncertainty from the hadronization model in the MC. Similarly, a measure as high X_s mass region as possible, in which a huge background from low energy photon exists, is also significant to understand the X_s decay model.

In Table 5.1, reconstructed X_s final states in this analysis are summarized. Actually, we tried to reconstruct ω modes($K\omega$, $K\omega\pi$, $K\omega 2\pi$), but found their branching fractions and reconstruction efficiencies were too small to measure. The total number of final states is 38, which covers 56% X_s final state. Assuming the isospin asymmetry between K_L and K_s and the amount of the K_L mode is same as that of the reconstructed K_s mode, the fraction is 69%, for example, the amount of $K_L\pi$ is assumed to be same as that of $K_S\pi$. We infer the total branching fraction by estimating the fraction of unmeasured modes using simulated fragmentation process.

5.2.2 $K4\pi$ Category Selection

The $K4\pi$ final state(Mode ID=13-16) has a large number of signal cross-feed, which is a background from the signal event by a mis-combination, and background due to the high multiplicity. To remove such background, additional selections of momentum in the CM frame are applied. The momentum of the fastest π and the second fastest π of 4π are useful for the background rejection. The selections are as follows.

- $p_\pi^{1st*} > 0.40$ GeV/c
- $p_\pi^{2nd*} > 0.25$ GeV/c

Table 5.1: Reconstructed X_s final states

Mode ID	Final state	Mode ID	Final state	Mode ID	Final state
1	$K^+\pi^-$	16	$K_s\pi^+\pi^+\pi^-\pi^0$	31	$K^+\eta\pi^-\pi^0$
2	$K_s\pi^+$	17	$K^+\pi^0\pi^0$	32	$K_s\eta\pi^+\pi^0$
3	$K^+\pi^0$	18	$K_s\pi^0\pi^0$	33	KKK
4	$K_s\pi^0$	19	$K^+\pi^-\pi^0\pi^0$	34	KKK_s
5	$K^+\pi^+\pi^-$	20	$K_s\pi^+\pi^0\pi^0$	35	KK_sK_s
6	$K_s\pi^+\pi^-$	21	$K^+\pi^+\pi^-\pi^0\pi^0$	36	$K^+K^+K^-\pi^-$
7	$K^+\pi^+\pi^0$	22	$K_s\pi^+\pi^-\pi^0\pi^0$	37	$K^+K^-K_s\pi^+$
8	$K_s\pi^+\pi^0$	23	$K^+\eta$	38	$K^+K^+K^-\pi^0$
9	$K^+\pi^+\pi^-\pi^-$	24	$K_s\eta$		
10	$K_s\pi^+\pi^+\pi^-$	25	$K^+\eta\pi^-$		
11	$K_s\pi^+\pi^0$	26	$K_s\eta\pi^+$		
12	$K_s\pi^+\pi^0$	27	$K^+\eta\pi^0$		
13	$K^+\pi^+\pi^+\pi^-\pi^-$	28	$K_s\eta\pi^0$		
14	$K_s\pi^+\pi^+\pi^-\pi^-$	29	$K^+\eta\pi^+\pi^-$		
15	$K_s\pi^+\pi^+\pi^-\pi^0$	30	$K_s\eta\pi^+\pi^-$		

We adopt the loose selection because the dependence on the hadronization model is needed to be small.

5.2.3 $K2\pi^0$ Category Selection

Since $K2\pi^0$ modes(Mode ID=17-22) have a large number of background originated from π^0 we apply an additional selection of π^0 momentum in the CM frame.

- $p_{\pi^0}^{1st*}$ >0.40 GeV/c
- $p_{\pi^0}^{2nd*}$ >0.25 GeV/c

5.3 B Meson Reconstruction

We combine the high energy photon candidate and the X_s candidate to form B meson candidates in the event. Two independent kinematic variables, the beam energy constrained B meson mass, M_{bc}, and the energy difference, ΔE, are calculated in the $\Upsilon(4S)$ rest frame.

$$M_{bc} = \sqrt{(E^*_{beam}/c^2)^2 - |\overrightarrow{p}^*_B/c|^2} \quad (5.1)$$

$$\Delta E = E^*_B - E^*_{beam} \quad (5.2)$$

E^*_{beam} is the beam energy, and E^*_B and p^*_B are the energy and momentum, respectively, of the B meson candidate in the CM frame. The energy E^*_B is calculated as $E^*_B = E^*_\gamma + E^*_{X_s}$. The momentum p^*_B is calculated without using the absolute value of the photon momentum according to

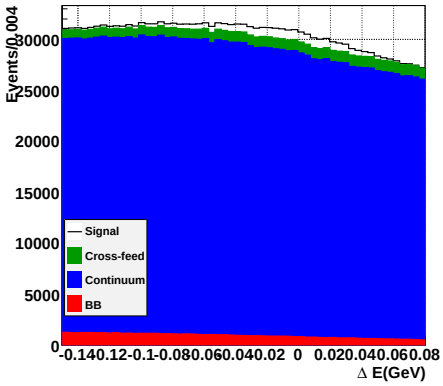
$$\overrightarrow{p^*_B} = \overrightarrow{p^*_{X_s}} + \frac{\overrightarrow{p^*_\gamma}}{|\overrightarrow{p^*_\gamma}|} \times (E^*_{beam} - E^*_{X_s}) \tag{5.3}$$

since the X_s momentum and the beam energy are determined with substantially better precision than that of the primary photon. The M_{bc} means a B meson mass obtained from the beam energy and B meson momentum, and a signal event has a peak at nominal B meson mass(5.279 GeV/c^2). The ΔE means a difference between the beam energy and B meson energy, and a signal event has a peak at 0 GeV. Finally, we fit the M_{bc} distribution to extract the signal yield. We select events with M_{bc} >5.24 GeV/c^2 and -0.15< ΔE <0.08, where for the final states with $2\pi^0$ and $\eta\pi^0$(Mode ID=17-22, 27, 28, 31, 32) which have a huge number of background, ΔE selection is tightened to $-0.10 < \Delta E < 0.05$ GeV.

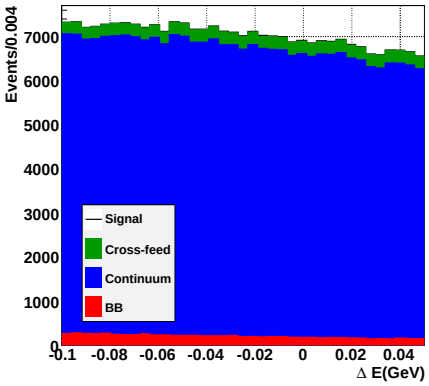
M_{bc} and ΔE distributions are shown in Figure 5.3 and numbers of signal and background are summarized in Table 5.2. $q\overline{q}$ BG is a background from $e^+e^- \to q\overline{q}$ decay, and $B\overline{B}$ BG is a background from $\Upsilon(4S) \to B\overline{B}$ decay. As you can see, a huge background still remains after the event selection, thus additional background suppressions are applied in next Chapter.

Table 5.2: Number of signal and background after Event selection(MC scaled to data size, $M_{bc} > 5.27$GeV, $M_{X_s} < 2.8$GeV)

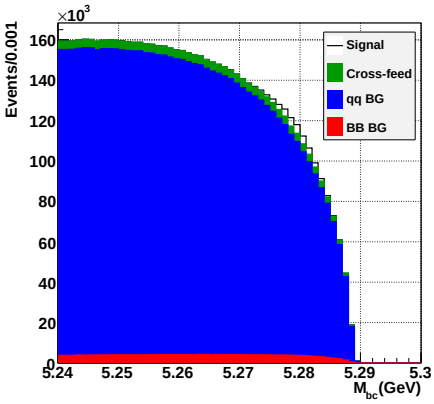
	Event selection
Signal	30356
Cross-feed	90940
$q\overline{q}$ BG	2545069
$B\overline{B}$ BG	231770
Significance	17.8



(a) ΔE distribution



(b) ΔE distribution(Mode ID=17-22, 27, 28, 31, 32)



(c) M_{bc} distribution

Figure 5.3: ΔE and M_{bc} distributions after Event selection

Chapter 6

Background Study

There are three types in the background. First one is the $B\overline{B}$ background, and the largest source is $B \to D^{(*)}\rho^{+}$. Since such background makes a peak in the signal region, it is necessary to reject as much as possible. This background is suppressed by a D veto in Section6.1. Second one is the $q\overline{q}$ background which is dominant one. This background is suppressed by a event shape in Section 6.2. Last one is the cross-feed background which comes from the signal events by a mis-combination. This background is suppressed by a method in which a B candidate is selected by requiring a most B meson like one in an event in Section 6.3.

6.1 $B\overline{B}$ Background Suppression :D veto

A lot of backgrounds from $B\overline{B}$ decay remain after the selection as shown in Table 5.2. The main origin are events with a $D^{(*)}$ meson, specifically $B \to D^{(*)}\rho^{+}$, which has 100 times branching ratio more than that of the signal. For example, $B \to D\rho$ looks like the signal when π^0 from the ρ emits a high energy photon. At first we attempted to veto on the ρ mass to suppress the $B\overline{B}$ background, however it is not effective due to the wide ρ mass width. Therefore, a suppression by D meson mass is applied for the $B\overline{B}$ background suppression.

6.1.1 D Meson Candidate Reconstruction and Selection

A D meson candidate is reconstructed as a combination of particles used in X_s reconstruction, where only combinations with the branching ratio >1% of D decay are allowed. We adopt a D veto without wrong sign constraint, in

which a combination is taken as a D candidate even if a sign of particle in D decay is different from correct one(for example, $D^+ \to K^+\pi^+\pi^-$) for a stronger background suppression. D veto without K/π particle ID selection, in which a combination is taken as a D candidate even if K/π particle ID is wrong, is also investigated, but it is not effective since a lot of the signal are rejected. The candidate whose D mass is the closest to the nominal D mass in an event is selected.

6.1.2 D Mass Veto

The region around the nominal D mass is vetoed. In order to take into account the mass difference between charged D and neutral D, different veto windows are applied to the D^+ and D^0. Since the mass of D candidate with π^0 or η in the children has a tail in the low side the D mass veto window should be enlarged. For the above reasons, the following 4 different veto windows are provided.

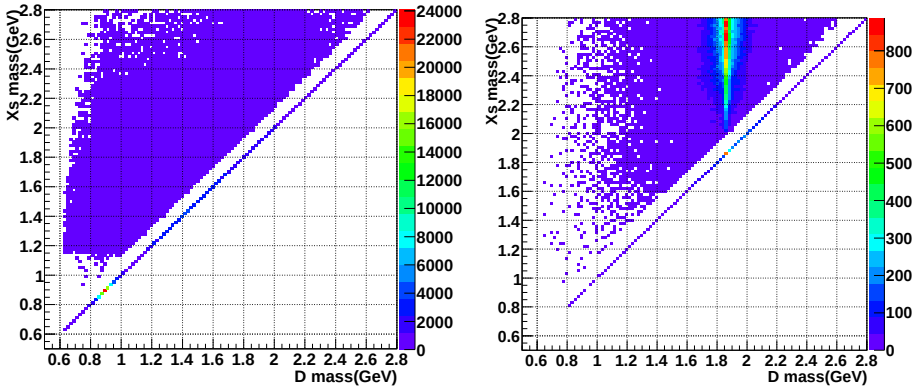
- D^0 without π^0/η : 1835< M_{D^0} <1895 MeV/c^2
- D^+ without π^0/η : 1840< M_{D^+} <1900 MeV/c^2
- D^0 with π^0/η: 1800< M_{D^0} <1905 MeV/c^2
- D^+ with π^0/η : 1805< M_{D^+} <1910 MeV/c^2

In Figure 6.1, the signal does not have a sharp peak around the nominal D mass, while the $B\overline{B}$ background peaks at the D mass at high X_s mass. Therefore, the veto window should be defined by the D mass and the X_s mass region. The D veto by the above mass windows are applied only to events with M_{Xs} >2.0 GeV/c^2. Consequently, 90 % of the signal is kept, while the background is reduces to 23%. We also attempt D veto without M_{X_s} >2.0 GeV/c^2 and evaluate two veto methods by using a ratio of the signal to the peaking background as a figure-of-merit. The figure-of-merits of two vetoes are almost same, however we adopt D veto with M_{X_s} >2.0 GeV/c^2 since it keeps more signal events.

In addition, we attempted to D^* meson veto, however it was not effective. Thus, we did not adopt D^* veto.

6.2 $q\overline{q}$ Background Suppression

The dominant background comes from $e^+e^- \to q\overline{q}$ events. The production cross-section from e^+e^- collisions at the $\Upsilon(4S)$ resonance receives sizable con-



(a) Signal(Peak at M_{X_s} =0.9 GeV/c^2 comes from events with K^*.)

(b) $B\overline{B}$ background

Figure 6.1: D mass and X_s mass

tributions other than $B\overline{B}$. To separate $q\overline{q}$ background, additional background rejection is ensured by exploiting the differences in the "event shape". In $B\overline{B}$ events both B mesons are produced almost at rest in the $\Upsilon(4S)$ frame. As a result, the B decay products are distributed isotropically. In contrast, for $q\overline{q}$ events, the quarks are produced with a large initial momentum, and yield a back-to-back fragmentation into two jets of light hadrons. For the same reason in $B\overline{B}$ events, the angular distribution of decay products from the two B mesons are uncorrelated, while for $q\overline{q}$ background a sizable correlation arises, thus the decay particles from each B candidate tend to align with the direction of its jet. Information based on the event shape of decay particles are quite useful for $q\overline{q}$ background suppression and can be quantified by many different ways. We adopt the variables in Section 6.2.1.

6.2.1 Variables for $q\overline{q}$ Background Suppression

In this analysis, we utilize following 10 event shape variables(1)-(4) and adopt other two variables, flavor tagging variable(5) and likelihood of ΔE(6).

(1) cosθ_B

(2) Thrust and related variables

- cosθ_T
- Thrust$_{other-side}$

(3) Sphericity and related variables

- Sphericity

- Aplanarity
- cos(v1-z)
- cos(v1-v1)
- cos(v2-v2)
- cos(v3-v3)

(4) Likelihood ratio of KSFW

(5) B flavor tagging

(6) likelihood of ΔE

Each variable for $q\overline{q}$ suppression are described in the following and the distributions are shown in Figure 6.2.

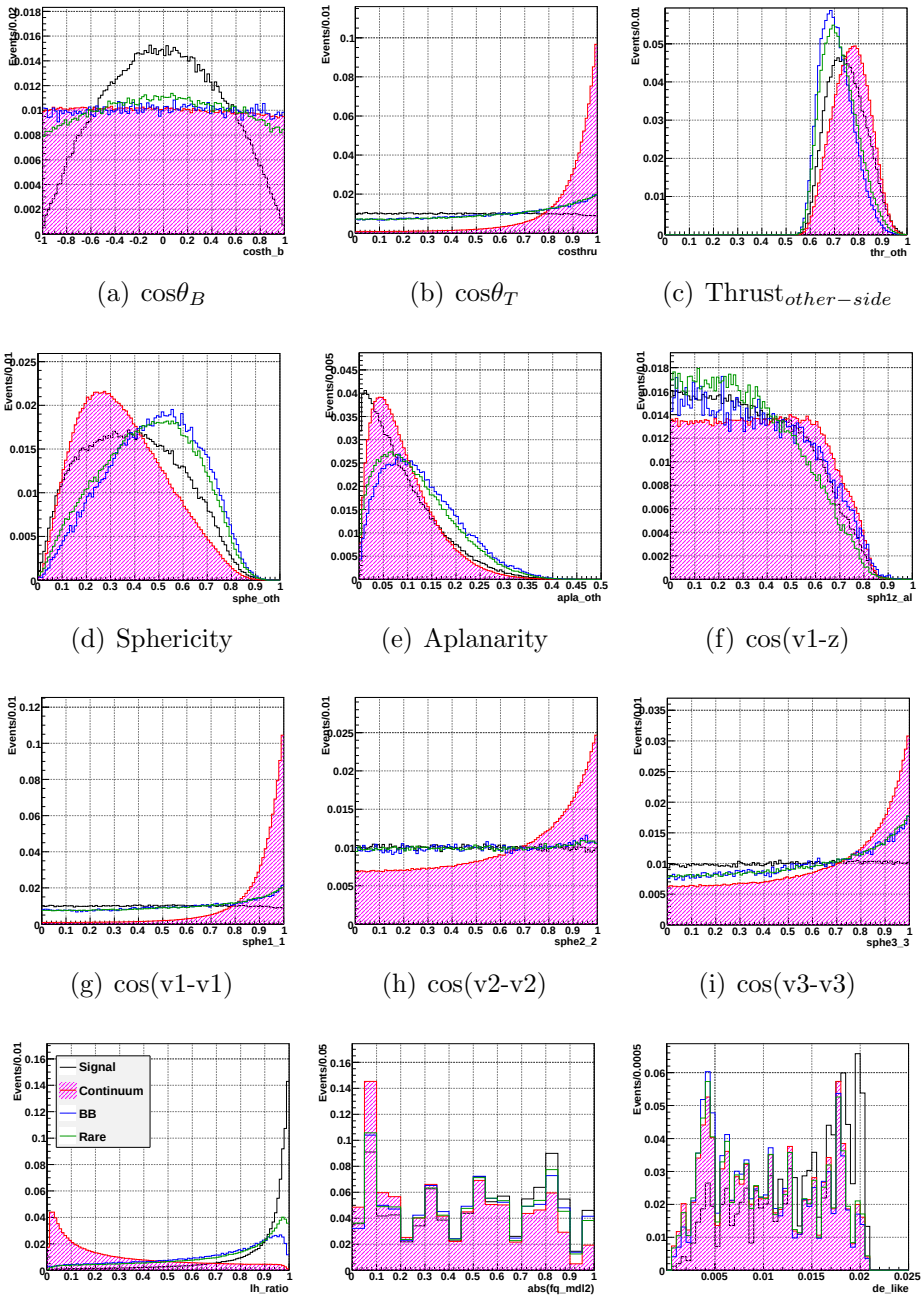
(1) cosθ_B The spin-1 $\Upsilon(4S)$ decaying into two spin-0 B mesons results in a 1-cos$^2\theta_B$ angular distribution with respect of the beam axis. On the other hand, the $q\overline{q}$ background has a flat cosθ_B distribution since the background is randomly reconstructed(Figure 6.2(a)).

(2) Thrust and related variable For a collection of N momenta p_i (i=1, $\cdots$, N), the thrust axis T is defined as the unit vector along which their total projection is maximal; the thrust is a derived quantity defined as

$$T = \frac{\Sigma_{i=1}^{N}|\mathbf{T}\cdot\mathbf{p_i}|}{\Sigma_{i=1}^{N}|\mathbf{p_i}|} \tag{6.1}$$

A useful related variable is cosθ_T, where θ_T is the angle between the thrust axis of the momenta of the B candidate decay particles and the thrust axis of all other particles in the event. For $B\overline{B}$ event, both B mesons are produced almost at rest in the $\Upsilon(4S)$ rest frame, so their decay particles are isotropically distributed, their thrust axes are randomly distributed. Thus the cosθ_T has a uniform distribution. In contrast, for $q\overline{q}$ events, the momenta of particles follow the direction of the jets in the event and as a consequence |cosθ_T | distribution has a strong peak at large values(Figure 6.2(b)). In this analysis, a thrust(Figure 6.2(c)) calculated by only other-side tracks which are not used in the signal reconstruction is adopted since a thrust calculated by all tracks has a correlation with M_{X_s}.

(3) Sphericity and related variables Sphericity and thrust are strongly correlated concepts. For a collection of momenta p_i, the sphericity tesor S is



(a) $\cos\theta_B$ (b) $\cos\theta_T$ (c) $\text{Thrust}_{other-side}$

(d) Sphericity (e) Aplanarity (f) cos(v1-z)

(g) cos(v1-v1) (h) cos(v2-v2) (i) cos(v3-v3)

(j) Likelihood ratio of KSFW (k) B flavor tagging (l) Likelihood of ΔE

Figure 6.2: Input variables for Neurobayes(Black:Signal, Red:$q\overline{q}$ background), Green:Signal cross-feed, Blue:$B\overline{B}$ background)

defined as

$$S^{\alpha,\beta} = \frac{\Sigma_{i=1}^{N} p_i^{\alpha} p_i^{\beta}}{\Sigma_{i=1}^{N} |\mathbf{p_i}|^2} \tag{6.2}$$

(with $\alpha, \beta = x, y, z$) and provides a three-dimensional representation of the spatial distribution of the p_i collection. For the isotropic distribution, its three eigenvalues λ_k have similar magnitude, while for a planar distribution, one of the eigenvalues is significantly smaller, with its eigenvector orthogonal to that plane. For a very directional distribution, the eigenvector oriented in that preferred direction has an eigenvalue considerably larger than the two others. Useful quantities derived from the sphericity are the sphericity scalar and the sphericity axis. The sphericity scalar S is defined as

$$S = \frac{3}{2}(\lambda_2 + \lambda_3), \tag{6.3}$$

where λ_2 and λ_3 are the 2nd and 3rd largest eigenvalues, respectively. The values of S close to 1 correspond to very isotropically distributed momentum collections, while very collimated distributions yield sphericity values close to zero(Figure 6.2(d)). In addition, aplanarity A, with definition $A \equiv \frac{3}{2}\lambda_3$, is useful. It is constrained to the range $0 \leq A \leq 1/2$ and measures the transverse momentum component out of the event plane; a planar event has $A \sim 0$ and an isotropic one $A \sim 1/2$(Figure 6.2(e)). In this analysis, a sphericity scalar calculated by only other-side tracks is adopted for $q\overline{q}$ suppression because a sphericity scalar with signal-side tracks has a correlation with M_{X_s}. Futhermore, 4 variables related with sphericity axes are used. The cos(v1-z) is an angle between 1st sphericity axis of all tracks and z direction(Figure 6.2(f)). The cos(v1,2,3-v1,2,3) is an angle between 1st, 2nd, 3rd sphericity axes of signal-side and of other side(Figure 6.2(g), 6.2(h),6.2(i)).

(4) KSFW The KSFW[59](Appendix C) is a Fisher discriminant extended from the Fox-Wolfram variables using information such as missing mass and is constructed from 17 variables. We make a Likelihood of KSFW and use a Likelihood ratio ($L = \frac{L_{Sig}}{L_{Sig}+L_{Bg}}$) for $q\overline{q}$ suppression(Figure 6.2(j)).

(5) B flavor tagging B flavor tagging[60] of the other-side B meson is useful for the additional background discrimination. The flavor can be determined from the charge of

1. high-momentum lepton from $B^0 \to Xl^+\nu$ decays,
2. kaons, since the majority of them originate from $B^0 \to K^+X$ decays through the cascade transition $\overline{b} \to \overline{c} \to \overline{s}$,
3. intermediate momentum leptons from $\overline{b} \to \overline{c} \to \overline{s}$,

4.) high momentum pions coming from $B^0 \to D^{(*)}\pi^+ X$ decays.

5. slow pions from $B^0 \to D^{*-}X, D^{*-} \to \overline{D^0}\pi^-$ decays, and

6. $\overline{\Lambda}$ baryons from the cascade decays $\overline{b} \to \overline{c} \to \overline{s}$.

We use two parameters, q and r, as the flavor tagging outputs. The parameter q is the flavor of the tag-side B. The parameter r is an expected flavor dilution factor that ranges from zero for no flavor information to unity for unambiguous flavor assignment. Using a binned multi-dimensional look-up table the signed probability, $q \cdot r$, is given by

$$q \cdot r = \frac{N(B^0) - N(\overline{B^0})}{N(B^0) + N(\overline{B^0})}, \tag{6.4}$$

where $N(B^0)$ and $N(\overline{B^0})$ are the numbers of B^0 and $\overline{B^0}$ in each bin of the look-up table prepared from a large statistics MC event sample. For the majority of the $q\overline{q}$ background events, $q \cdot r$ populates around 0, where the flavor information is poorly known. On the other hand, a sizable number of B decay events have $q \cdot r \sim \pm 1$(Figure 6.2(k)).

(6) Likelihood of ΔE ΔE(Equation 5.2) is useful for background suppression since the distribution of the signal has a peak at 0, while the background has a flat distribution. However ΔE has a large tail in negative side due to the shower leakage in the photon reconstruction, especially in modes with π^0 or η . Thus, a non-uniform efficiency occurs between modes with/without π^0 or η by using the distribution. To escape the non-uniform efficiency, we make a likelihood of ΔE and provide two different PDFs for modes with/without π^0 or η to eliminate a bias in the efficiency(Figure 6.3). We attempted constructing three PDF's for no π^0, $1\pi^0$ and $2\pi^0$ modes, but made little improvement. ΔE plays an important role in not only the background suppression, but also Best candidate selection(Section 6.3). In Figure 6.4, ΔE distributions of each modes are shown. An uneven distribution in Figure 6.2(l) is caused by the binning of Figure 6.3.

Correlations between the input variables, M_{X_s} and M_{bc} are investigated in Figure 6.5. Clearly, the correlations are small enough, thus the $q\overline{q}$ suppression with these variables have no effect on the shape of M_{X_s} and M_{bc}.

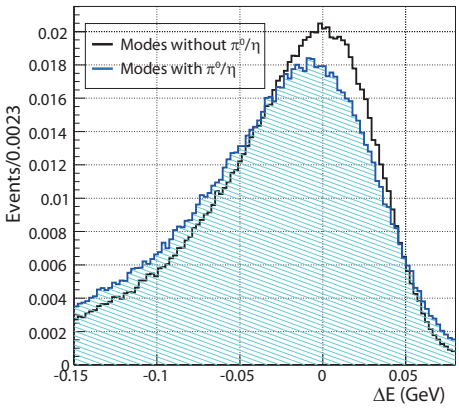
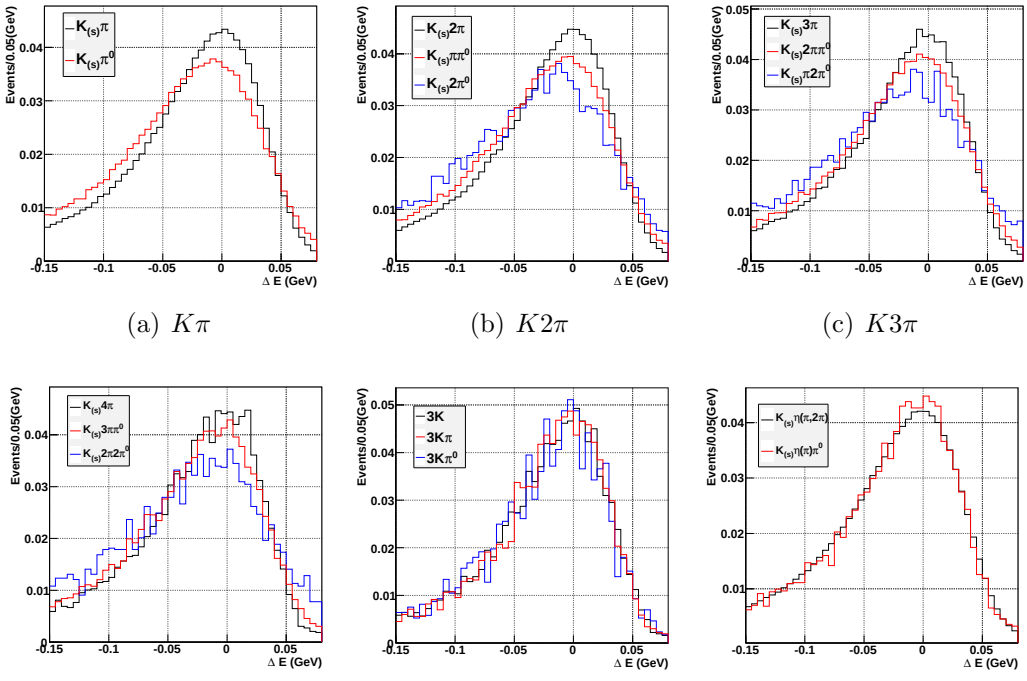


Figure 6.3: ΔE PDF with/without π^0



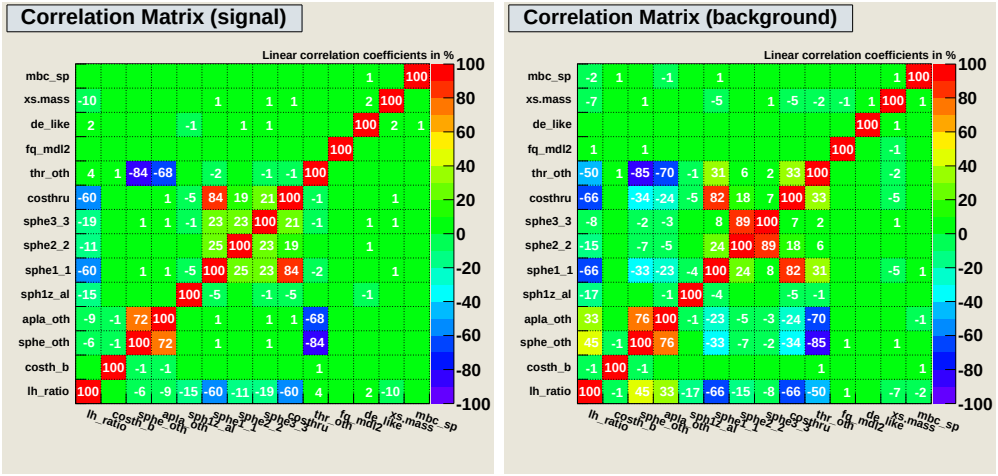
(a) $K\pi$ (b) $K2\pi$ (c) $K3\pi$

(d) $K4\pi$ (e) $3K$ (f) $K\eta$

Figure 6.4: ΔE distributions of signal

6.2.2 NeuroBayes Neural Network

For an effective background rejection we combined the above variables using the NeuroBayes package(NB)[61], which is a highly sophisticated tool for multivariate analysis based on Bayesian statistics. The advantage of a neural network technique is to be able to add more variables that may have correlations with existing ones and consider their non-linearly correlations. An auto-



(a) Signal (b) Background

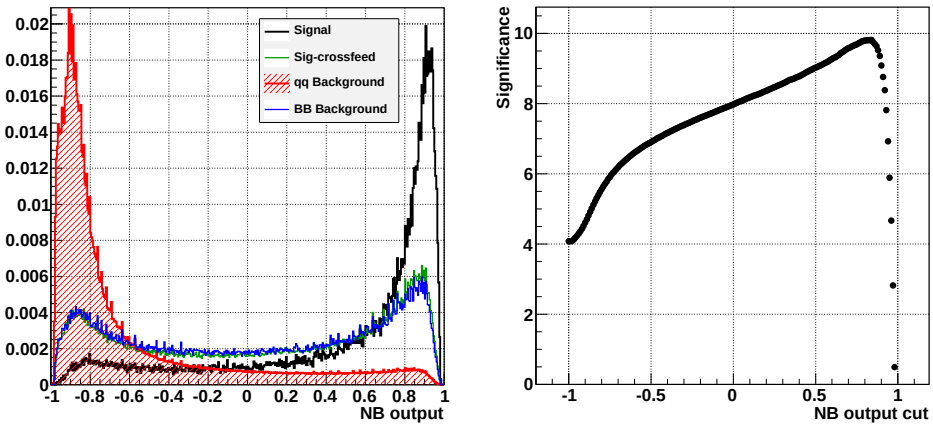
Figure 6.5: Correlation matrix

mated preprocessing of the input variables is followed by a three-layered(input, hidden and output layers) neural network is combined with an automated preprocessing of the input variable. The complex relationships between the input variables are learnt by using a provided dataset such as simulated data, and transformed into the output for analyzing the data of interest. The output can be utilized for classification.

6.2.3 $q\overline{q}$ Suppression with NeuroBayes

The NB is trained with MC samples. A measurement in low M_{X_s} region is relatively easy because the signal event is enough and the amount of background event is not large. On the other hand, a measurement in high M_{X_s} region(>2.2 GeV/c^2) is difficult due to the low signal statistics and huge background. Therefore, for the signal and the background training, the events with 2.2< M_{X_s} <2.8 GeV/c^2 in the signal MC and $q\overline{q}$ background MC sample, respectively, are used since we need for more events in high M_{X_s} region to suppress the systematic uncertainty. The output of NB is shown in Figure 6.6. The cut value of the NB output is optimized by significance in 2.2< M_{X_s} <2.8 GeV/c^2, which is defined as $= N_{sig}/\sqrt{(N_{sig}+N_{BG})}$, where N_{sig} means number of signal and N_{BG} means number of the signal cross-feed, the $q\overline{q}$ background and the $B\overline{B}$ background in 2.2< M_{X_s} <2.8 GeV/c^2. In Figure 8.4(b), the significance is plotted as a function of the cut value of the NB output and the cut value is 0.78 for the largest significance. As a result, 52 % of the signal keeps, on the other hand the $q\overline{q}$ background reduces to 2% in $M_{bc} > 5.27$ GeV/c^2 and

M_{X_s} <2.8 GeV/c^2.



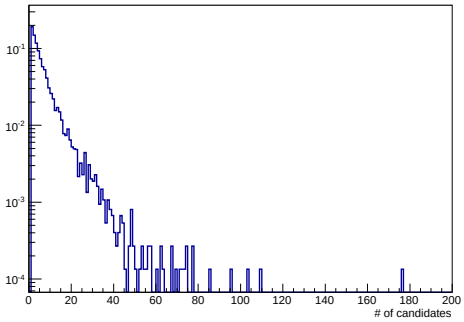
(a) NeuroBayes Output

(b) Significance at 2.2< M_{X_s} < 2.8 GeV/c^2 as a function of NB output cut value

Figure 6.6: NeuroBayes output and Significance

6.3 Best Candidate Selection of B Meson

One inherent problem in the inclusive analysis of $B \to X_s\gamma$ is that we can form a large number of possible B candidates, actually 6.4 candidates in an event on average(M_{bc} >5.24 GeV) before $q\overline{q}$ suppression in Figure 6.7(a), because 38 modes are reconstructed at the same time. The number of candidates



(a) Number of B candidates per event

Figure 6.7: Number of B candidates per event

depends on the multiplicity in final states, therefore, $K4\pi$, $K2\pi2\pi^0$ modes have a large number of candidates. To select "most B like" candidate, in previous

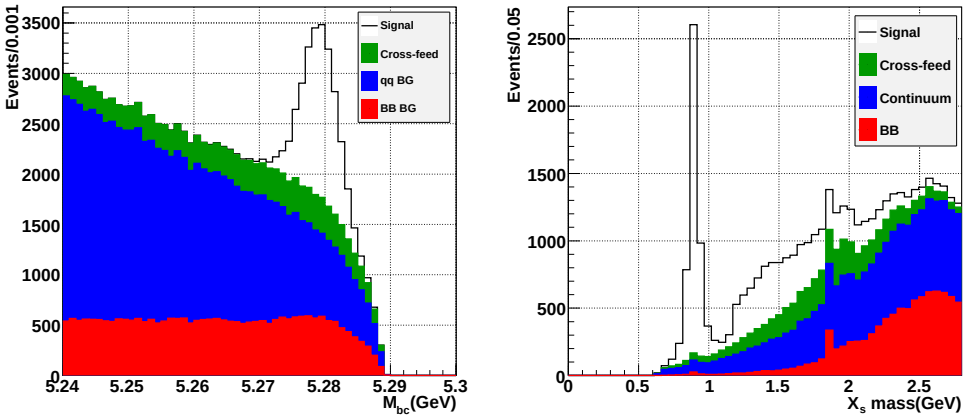
analysis[39], the vertex information of X_s was basically used and $|\Delta E|$ was used in mode without the vertex information, namely no charged particle mode(e.g. $K_s\pi^0\gamma$ etc). It has a problem that a difference in the efficiency between modes occurs, specifically modes with/without vertex information and π^0. Therefore, only ΔE information is used for all modes to escape the non-uniform efficiency. At first, we evaluated a Best candidate selection(BCS) with a likelihood of ΔE before the $q\overline{q}$ background suppression, but the efficiency was low(56 %). In order to improve it, the BCS is applied after the background suppression to reduce the multiple-candidate events, and the signal cross-feed is decreased to 19 %. Furthermore, the likelihood of ΔE is added to the variables for the NB input and the candidate with the largest output of the NB is selected as a B meson candidate. Finally, the efficiency on the BCS rises to 85 %.

A cut flow is shown in Figure 6.1 from the event selection to the BCS. A significance is improved from 18 to 58, and a significance in high M_{X_s} region($2.0 < M_{X_s} < 2.8$ GeV/c^2) is also improved from 4 to 10. Figure 6.8 shows M_{bc} and M_{X_s} distributions after the BCS. The background is reduced effectively, compared with Figure 5.3. In the high M_{X_s} region(> 2.2 GeV/c^2), however, the amount of the background is still large though the $q\overline{q}$ background suppression is optimized to this region.

M_{X_s} region is enlarged from 2.05 GeV/c^2 in the previous measurement[39] to 2.80 GeV/c^2 to suppress the systematic uncertainty by understanding a high M_{X_s} region, although a number of background is greatly increased (Figure 6.8(b)). A measurement above 2.8 GeV/c^2 is difficult due to the low signal statistics.

Table 6.1: Cut flow table(MC scaled to data size, $M_{bc} > 5.27$GeV, $M_{X_s} < 2.8$GeV)

	Event selection	D veto	$q\overline{q}$ suppression	BCS
Signal	30356	27137	14068	11824
Cross-feed	90940	64938	13096	5563
$q\overline{q}$ BG	2545069	1837720	42195	15226
$B\overline{B}$ BG	231770	118749	20023	8976
Significance	17.8	19.0	47.1	58.0



(a) M_{bc} at M_{X_s} <2.8 GeV/c^2

(b) X_s mass at M_{bc} >5.27 GeV/c^2. In M_{X_s} distribution, a peak of the $B\overline{B}$ background around 1.9 GeV/c^2 comes from events with D meson.

Figure 6.8: M_{bc} and X_s mass after BCS

6.4 Signal Efficiency

Finally, we fit M_{bc} distribution to extract the signal yield and calculate the branching ratio by using a signal efficiency. The signal efficiency obtained from the MC is 2.14 %. Since the efficiency depends on the particle content of the final states, this efficiency obtained from the MC is not necessarily correct. Thus, X_s decay model in the inclusive MC should be calibrated by using the data. For the calibration on the X_s decay model, the fragmentation study in data is needed. Therefore, we investigate expected numbers and significances for each mode in Table 6.2 and show the M_{bc} distributions in Figure 6.9. In Table 6.4, the signal efficiencies on each mode are included.

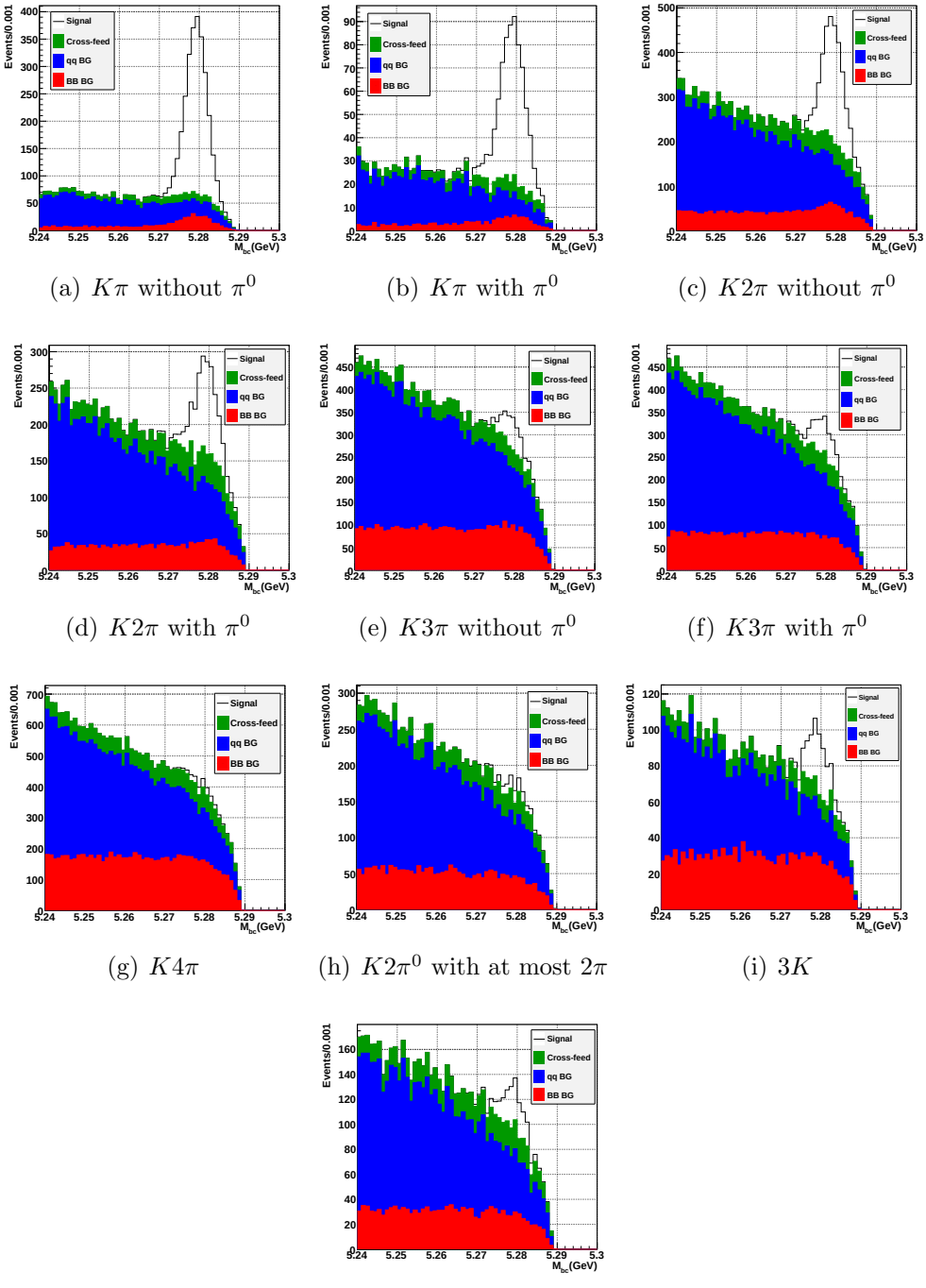
To measure partial branching ratios on M_{X_s}, the study on each M_{X_s} bin is needed. Therefore, we investigate expected numbers and significances for each M_{X_s} bin in Table 6.3 by MC and show the M_{bc} distributions in Figure 6.10 and 6.11. Table 6.5 include signal efficiencies on M_{X_s} bins. Figure 6.12 shows the signal efficiency as a function of M_{X_s}.

Table 6.2: Number of events for each final states (MC scaled to data size, 1.15< M_{X_s} <2.80 GeV/c^2, $M_{bc} > 5.27$GeV/c^2)

Mode	Definition	Mode	Signal	Cross-feed	$q\overline{q}$	$B\overline{B}$	Significance
1	$K\pi$ without π^0	1,2	2129	217	494	305	38
2	$K\pi$ with π^0	3,4	534	95	167	76	18
3	$K2\pi$ without π^0	5,6	1827	734	1830	803	25
4	$K2\pi$ with π^0	7,8	941	640	1393	594	16
5	$K3\pi$ without π^0	9,10	458	682	2441	1514	6.4
6	$K3\pi$ with π^0	11,12	564	774	2317	1227	8.1
7	$K4\pi$	13-16	193	988	2993	2681	2.3
8	$K2\pi^0$ with at most two π	17-22	142	501	1471	754	2.7
9	$K\eta$ with at most two π	23-32	236	343	848	457	5.4
10	$3K$ with at most oneπ	33-38	218	186	503	471	5.9

Table 6.3: Number of events for each M_{X_s} bin after BCS(MC scaled to data size, $M_{bc} > 5.27$GeV)

M_{X_s} bin(GeV)	Signal	Cross-feed	$q\overline{q}$	$B\overline{B}$	Significance
0.6-0.7	11	13	44	2	1.3
0.7-0.8	148	41	98	7	8.6
0.8-0.9	2427	81	140	28	47
0.9-1.0	1752	97	167	26	39
1.0-1.1	199	106	197	18	8.7
1.1-1.2	267	147	259	28	10
1.2-1.3	626	200	305	38	18
1.3-1.4	743	249	381	59	20
1.4-1.5	823	302	439	73	20
1.5-1.6	730	360	518	90	18
1.6-1.7	684	424	631	135	16
1.7-1.8	643	481	730	182	14
1.8-1.9	578	509	901	491	12
1.9-2.0	505	530	1019	429	10
2.0-2.1	430	401	920	479	9.1
2.1-2.2	323	363	1017	594	6.7
2.2-2.4	487	618	2311	1726	6.8
2.4-2.6	289	408	2547	2245	3.9
2.6-2.8	161	237	2605	2326	2.2



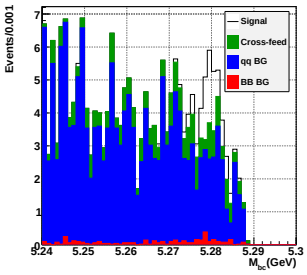
(a) $K\pi$ without π^0 (b) $K\pi$ with π^0 (c) $K2\pi$ without π^0

(d) $K2\pi$ with π^0 (e) $K3\pi$ without π^0 (f) $K3\pi$ with π^0

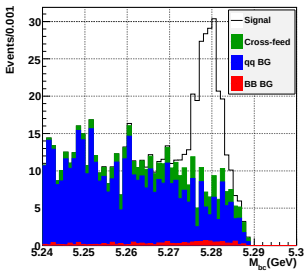
(g) $K4\pi$ (h) $K2\pi^0$ with at most 2π (i) $3K$

(j) $K\eta$

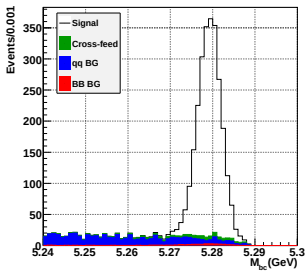
Figure 6.9: M_{bc} distributions for each final state at 1.15< M_{X_s} <2.80 GeV/c^2



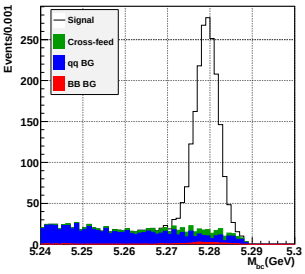
(a) 0.6< M_{X_s} <0.7 GeV/c^2



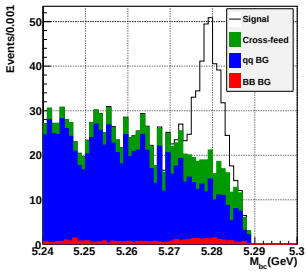
(b) 0.7< M_{X_s} <0.8 GeV/c^2



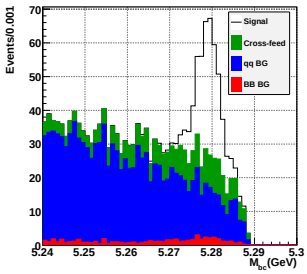
(c) 0.8< M_{X_s} <0.9 GeV/c^2



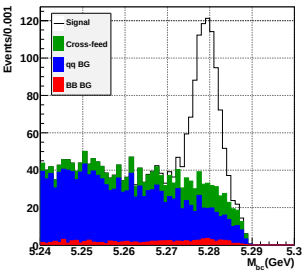
(d) 0.9< M_{X_s} <1.0 GeV/c^2



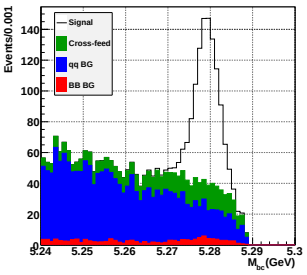
(e) 1.0< M_{X_s} <1.1 GeV/c^2



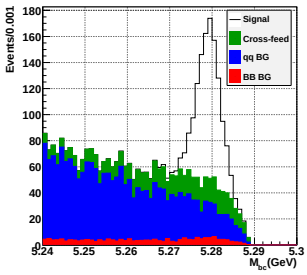
(f) 1.1< M_{X_s} <1.2 GeV/c^2



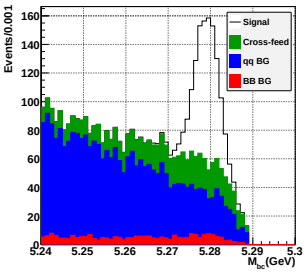
(g) 1.2< M_{X_s} <1.3 GeV/c^2



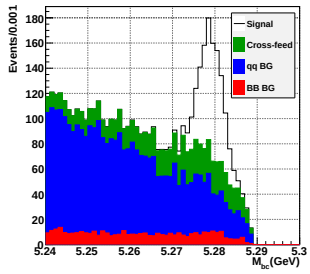
(h) 1.3< M_{X_s} <1.4 GeV/c^2



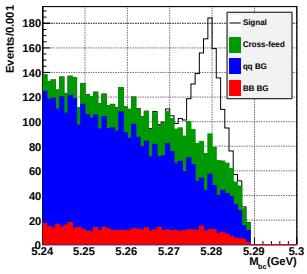
(i) 1.4< M_{X_s} <1.5 GeV/c^2



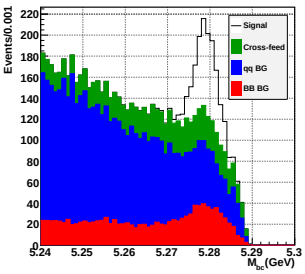
(j) 1.5< M_{X_s} <1.6 GeV/c^2



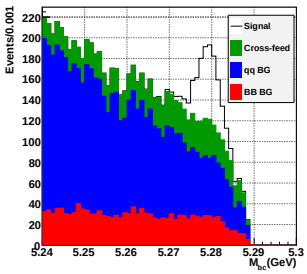
(k) 1.6< M_{X_s} <1.7 GeV/c^2



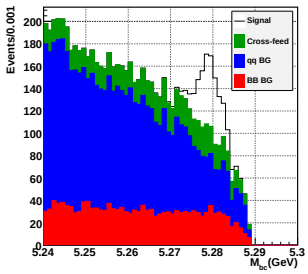
(l) 1.7< M_{X_s} <1.8 GeV/c^2



(m) 1.8< M_{X_s} <1.9 GeV/c^2

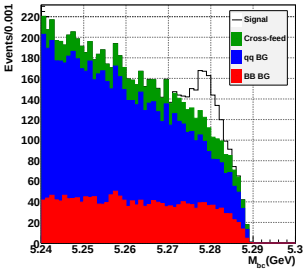


(n) 1.9< M_{X_s} <2.0 GeV/c^2

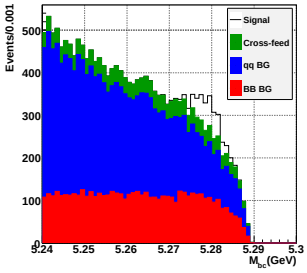


(o) 2.0< M_{X_s} <2.1 GeV/c^2

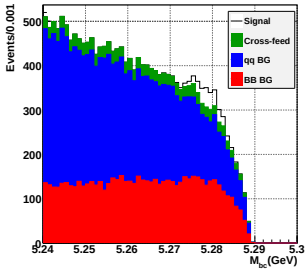
Figure 6.10: M_{bc} distributions for each M_{X_s} bin(0.6< M_{X_s} <2.1 GeV/c^2)



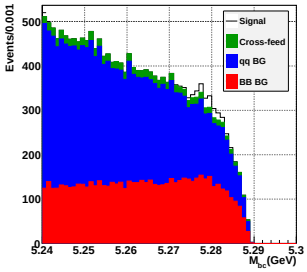
(a) 2.1< M_{X_s} <2.2 GeV/c^2



(b) 2.2< M_{X_s} <2.4 GeV/c^2



(c) 2.4< M_{X_s} <2.6 GeV/c^2



(d) 2.6< M_{X_s} <2.8 GeV/c^2

Figure 6.11: M_{bc} distributions for each M_{X_s} bin(2.1< M_{X_s} <2.8 GeV/c^2)

Table 6.4: Signal efficiency on each mode(M_{bc} > 5.24GeV, 1.15< M_{X_s} <2.80GeV/c^2)

Mode	Reconstruction efficiency(%)	Mode	Reconstruction efficiency(%)
1	8.84	6	1.74
2	4.33	7	0.83
3	6.42	8	0.52
4	2.86	9	2.35
5	3.74	10	3.43

Table 6.5: Signal efficiency on each M_{X_s} bin

M_{X_s} bin(GeV)	Signal efficiency(%)	M_{X_s} bin(GeV)	Signal efficiency(%)
0.6-0.7	7.32	1.6-1.7	2.43
0.7-0.8	7.59	1.7-1.8	2.10
0.8-0.9	7.03	1.8-1.9	1.80
0.9-1.0	7.36	1.9-2.0	1.52
1.0-1.1	7.10	2.0-2.1	1.28
1.1-1.2	4.58	2.1-2.2	0.99
1.2-1.3	4.22	2.2-2.4	0.80
1.3-1.4	3.98	2.4-2.6	0.56
1.4-1.5	3.78	2.6-2.8	0.41
1.5-1.6	2.91		

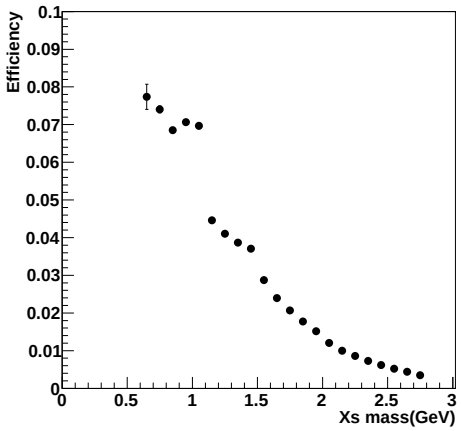


Figure 6.12: Signal efficiency function of M_{X_s}

Chapter 7

Maximum Likelihood Fit

In this chapter, the procedure to extract the signal event is described. The signal yield is extracted from M_{bc} distribution. At first, the unbinned maximum likelihood fit method is described, and then, PDFs used in the M_{bc} are reported. At last, the fit bias is checked.

7.1 Unbinned Maximum Likelihood Fit Method

In the M_{bc} fit, we use the extended unbinned maximum likelihood fit. In a assumption, a function $\mathcal{P}(\alpha, x)$ reproduce the experimental result x when a number of x is large enough, where α is a set of the parameters to vary the shape of $\mathcal{P}(\alpha, x)$. A function $\mathcal{P}(\alpha, x)$ is a normalized distribution:

$$\int \mathcal{P}(\alpha, x)dx = 1, \tag{7.1}$$

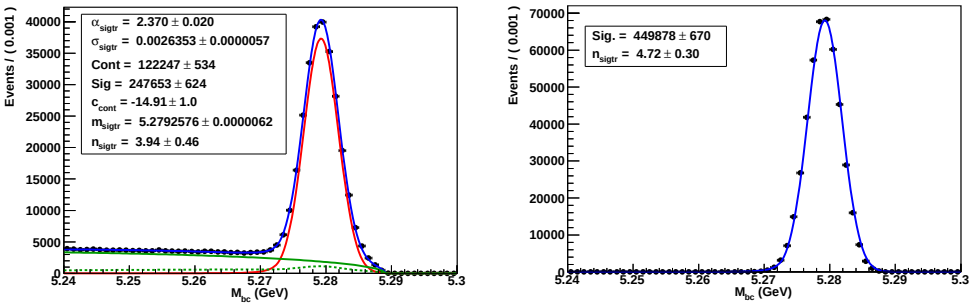
where $\mathcal{P}(\alpha, x)$ is called a probability density function(PDF).

The likelihood function $\mathcal{L}(\alpha)$ is the joint probability density of the experimental results $x_i (i = 1, 2, ..., N)$:

$$\int \mathcal{L}(\alpha)dx = \prod_{i=1}^{N} \mathcal{P}(\alpha, x_i), \tag{7.2}$$

where α is fitting parameters. We try to find the most probable value of α as the solution which gives the maximum value of $\mathcal{L}(\alpha)$, varying the value of α.

The signal makes a peak at the nominal B meson mass(5.279 GeV/c^2). The fit region is defined as above 5.24 GeV/c^2 in order to determine the background shape. In the likelihood function, we consider five components, signal, signal



(a) M_{bc} fit for $D\pi$ data to obtain signal PDF (Red solid line: Signal, Green solid line: $q\overline{q}$ BG, Green dashed line: $B\overline{B}$ BG)

(b) M_{bc} fit for signal MC to obtain a tail parameter

Figure 7.1: Signal PDF

cross-feed, peaking background, non-peaking part from $B\overline{B}$ background and non-peaking part from $q\overline{q}$ background. The fit method is summarized in Table 7.1.

7.2 Signal PDF

A Crystal ball function is adopted to model the signal PDF:

$$f_{CrystalBall}(x) = \begin{cases} exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right) & (|\frac{x-m}{\sigma}| < \alpha), \\ \frac{\left(\frac{n}{\alpha}\right)^n e^{-\frac{1}{2}\alpha^2}}{\left(\frac{n}{\alpha}-\alpha-\frac{x-m}{\sigma}\right)^n} & (|\frac{x-m}{\sigma}| > \alpha), \end{cases} \tag{7.3}$$

where m and σ are the peak position and width, respectively, and the parameters α and n take account of the non-Gaussian tail. Shape parameters are fixed to the values obtained from $B \to D\pi$ data(Figure 7.1(a)), where the shape and yield of $B\overline{B}$ background PDF are fixed to values obtained from MC, but the contribution is small enough. The tail parameters n and α is provided from the signal MC (Figure 7.1(b)) because the tail shape of $D\pi$ sample does not agree with that of the signal well(Appendix D). But, it should not be the large source of the uncertainty since the signal PDF is determined precisely. The floated parameter in the M_{bc} fit is the signal yield.

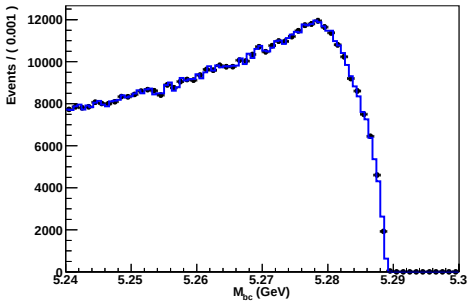


Figure 7.2: Signal cross-feed PDF

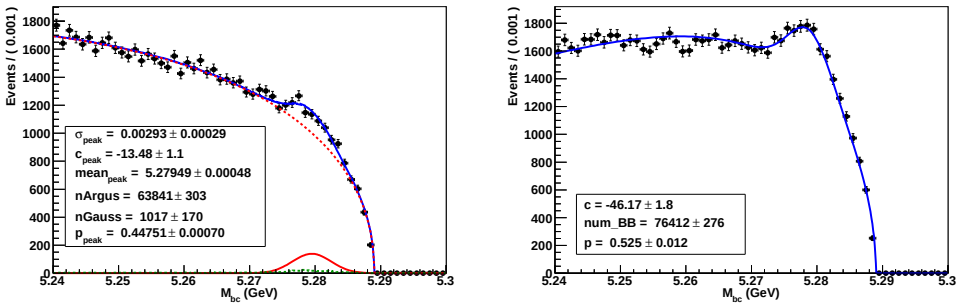
7.3 Signal Cross-feed PDF

For the signal cross-feed, we construct a histogram-PDF from the signal MC samples in Figure 7.2. A fraction of the signal cross-feed to the signal is fixed to the values obtained from the MC samples since the number of the cross-feed is directly proportional to that of the signal.

7.4 Peaking Background PDF

Gaussian function is adopted to model the peaking background. The shape parameters and yield are fixed to the values obtained from an anti-π^0/η veto sample. The anti-π^0/η veto is defined by requiring the π^0/η probabilities above 0.8. Figure 7.3(a) shows M_{bc} fit for data in the anti-π^0/η veto region by a gaussian and ARGUS functions. In Figure 7.3(b), M_{bc} of $B\overline{B}$ background MC in the signal region is fitted by the gaussian and the ARGUS, where the shape parameters and the yield of the gaussian are fixed to the values obtained from the M_{bc} fit in Figure 7.3(a) and the shape parameters and the yield of the ARGUS are floated. The shape and yield of the peaking background are in a agreement with them of the signal region.

In addition, the difference on the π^0/η veto efficiency in the signal region between the data and the MC are evaluated. For this study, we use a event sample with anti-D veto after $q\overline{q}$ suppression, in other words, events which are rejected by the D veto. M_{bc} distribustions of the $B\overline{B}$ background MC and data with anti-D veto are shown in Figure 7.4 and fitted to obtain yields of the peaking background component. As a result, the π^0/η veto efficiencies of the MC and data are 92.8 and 92.7 %, respectively. The MC is in a good agreement with the data as for the π^0/η veto.



(a) M_{bc} fit for data in anti-π^0/η veto region to obtain peaking BG PDF(Red solid line: Peaking component, Red dashed line: Non-peaking component, Green dashed line: Signal cross-feed obtained from signal MC)

(b) M_{bc} fit of $B\overline{B}$ BG(MC) in signal region by the peaking BG PDF(Figure 7.3(a))

Figure 7.3: Peaking BG PDF

Therefore, it is reasonable that the sample with anti-π^0/η veto are used to determine the peaking background PDF.

7.5 PDF for Non-peaking part from $B\overline{B}$ Background

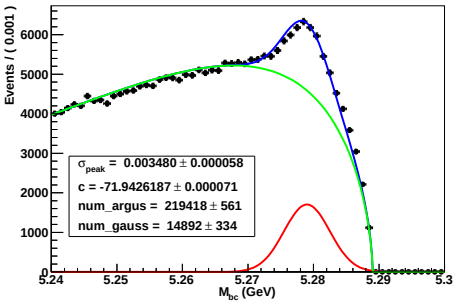
The non-peaking part from $B\overline{B}$ background is modeled by an ARGUS function. The ARGUS function is defined as

$$f_{ARGUS}(x) = x\left(1-\left(\frac{x}{E^*_{beam}}\right)^2\right)^p \cdot exp\left(c\left(1-\left(\frac{x}{E^*_{beam}}\right)^2\right)\right). \quad (7.4)$$

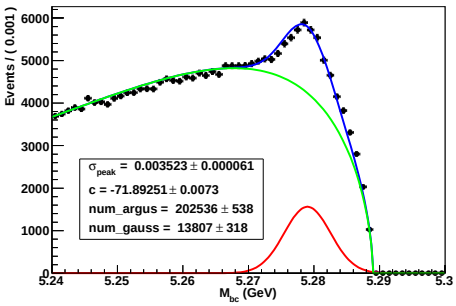
The endpoint of ARGUS function is fixed to the beam energy E^*_{beam} in the CM frame and other shape parameters and a yield are floated.

7.6 PDF for Non-peaking Part from $q\overline{q}$ Background

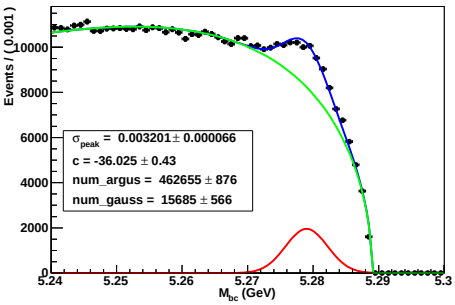
At first, we used a ARGUS PDF for non-peaking parts from $B\overline{B}$ and $q\overline{q}$ background. However, it incurred a negative bias on the signal yield since the $q\overline{q}$ background tended to take away the signal yield in the MC study. This is because a slope of the $q\overline{q}$ background is steep and not consistent with the



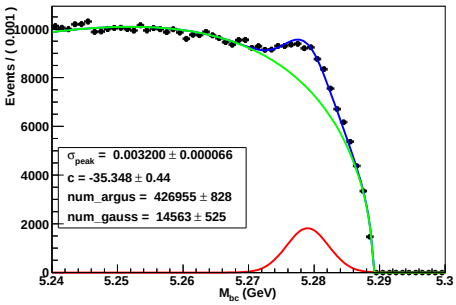
(a) $B\overline{B}$ background MC with anti-D veto before π^0/η veto



(b) $B\overline{B}$ background MC with anti-D veto after π^0/η veto



(c) Data with anti-D veto before π^0/η veto



(d) Data with anti-D veto after π^0/η veto

Figure 7.4: M_{bc} fit for anti-D veto sample to evaluate the π^0/η veto efficiency(Red solid line: Peaking component, Green line: Non-peaking component)

ARGUS shape. Thus, we adopt a modified ARGUS function defined as

$$f_{ARGUS}(x) = x\left(1-\left(\frac{x}{E^*_{beam}}\right)^q\right)^p \cdot exp\left(c\left(1-\left(\frac{x}{E^*_{beam}}\right)^2\right)\right), \qquad (7.5)$$

where a new floated parameter q is introduced to the original one(Equation 7.4) instead of the fixed value 2. The shape and yield are fixed to the values obtained from off-resonance data in Figure 7.5.

7.7 M_{bc} Fit with MC Sample

Figure 7.6 shows the M_{bc} fit by using the above PDFs with the MC sample. The signal yield of the fit result is 11426±217(Expected:11878, statistical uncertainty:1.90%). In Figures 7.7 and 7.8, M_{bc} fits are performed for each decay mode and the fit results are summarized in Table 7.2. In Figure 9.5, 9.6 and 9.7, M_{bc} fits are performed for each M_{X_s} bin and the fit results are summarized in Table 7.3.

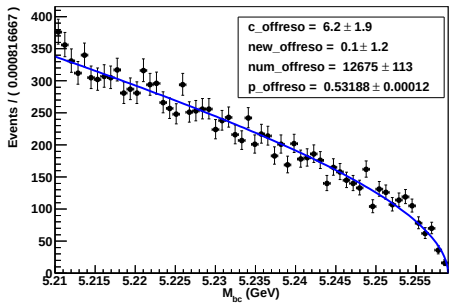


Figure 7.5: $q\overline{q}$ background PDF obtained from off-resonance data

Table 7.1: Fit method

PDF component	Function	Parameter (fixed/floated)
Signal	Crystal Ball	N_{sig} : floated All shape parameters : fixed
Signal cross-feed(scf)	histogram-PDF	N_{scf}/N_{sig} : fixed
Peaking background	Gaussian	$N_{peaking}$: fixed All shape parameters : fixed
Non-peaking background from $B\overline{B}$	Argus	$N_{B\overline{B}}$: floated Shape parameters : floated Endpoint : fixed
Non-peaking background from $q\overline{q}$	Modified Argus	$N_{q\overline{q}}$: fixed Shape parameters : fixed Endpoint : fixed

7.8 Fitter check

To investigate a bias for the signal yield, We generate the test event sample from the PDFs for all components(signal, cross-feed, peaking background, non-peaking background from $B\overline{B}$, and non-peaking background from $q\overline{q}$). The amount of each generated event are also fluctuated according to the Poisson distribution, where the signal and cross-feed events are generated separately, thus the fraction of the signal to the cross-feed are also fluctuated. Figure 7.12 shows a pull distribution for total fit. The pull is defined as

$$\text{Pull} = \frac{(\#\text{ of signal from fit result}) - (\text{Expected }\#\text{ of signal})}{\text{Error from fit result}}. \tag{7.6}$$

We obtain the pull distribution which has a mean of 0 and sigma of 1. Figure 7.13 and 7.14, 7.15 show pull distributions in each decay mode and M_{Xs} bin. In these distributions, a bias is not appeared and we can confirm the validity

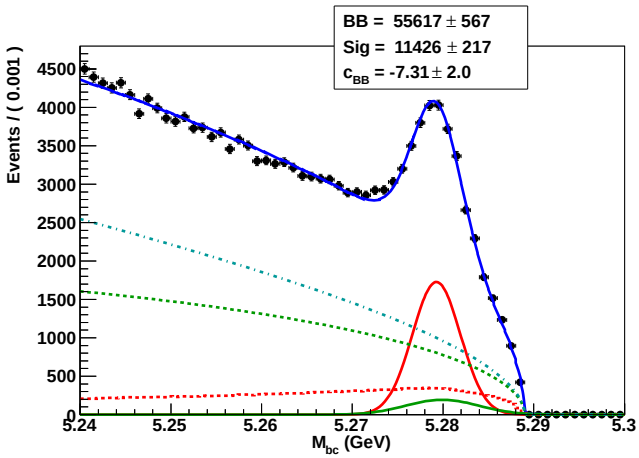


Figure 7.6: M_{bc} fit(MC)(Red solid line: Signal, Red dashed line: Signal cross-feed, Green solid line: Peaking BG, , Green dashed line: Non-peaking background from $B\overline{B}$, Cyan dashed line:Non-peaking background from $q\bar{q}$)

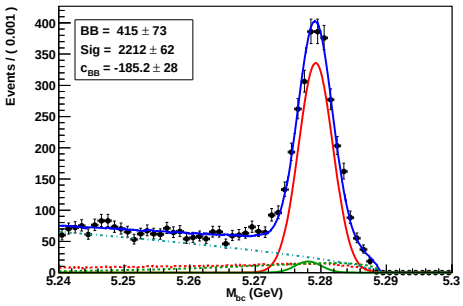
Table 7.2: Fit result for each final states(MC, 1.15< M_{X_s} <2.8 GeV/c^2)

Mode	Definition	Signal yield
1	$K\pi$ without π^0	2212±62
2	$K\pi$ with π^0	460±32
3	$K2\pi$ without π^0	1556±74
4	$K2\pi$ with π^0	884±64
5	$K3\pi$ without π^0	396±66
6	$K3\pi$ with π^0	457±63
7	$K4\pi$	194±59
8	$K2\pi^0$	158±40
9	$K\eta$	214±44
10	$3K$	198±38

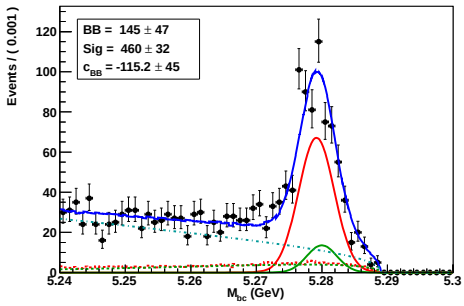
on our fitter.

Table 7.3: Fit result for each each M_{X_s} bin(MC)

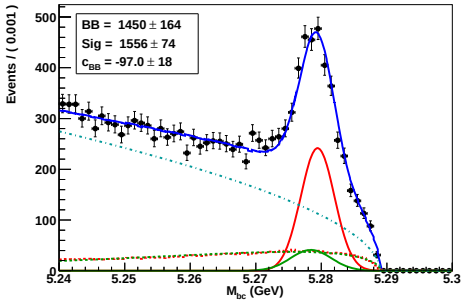
M_{X_s} bin(GeV)	Signal yield
0.6-0.7	8±4
0.7-0.8	141±17
0.8-0.9	2401±52
0.9-1.0	1673±54
1.0-1.1	170±21
1.1-1.2	262±21
1.2-1.3	521±34
1.3-1.4	724±41
1.4-1.5	855±41
1.5-1.6	655±37
1.6-1.7	630±38
1.7-1.8	642±43
1.8-1.9	578±46
1.9-2.0	490±48
2.0-2.1	390±45
2.1-2.2	318±44
2.2-2.4	538±68
2.4-2.6	331±68
2.6-2.8	128±64



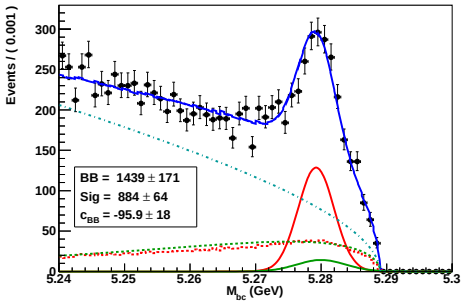
(a) $K\pi$ without π^0



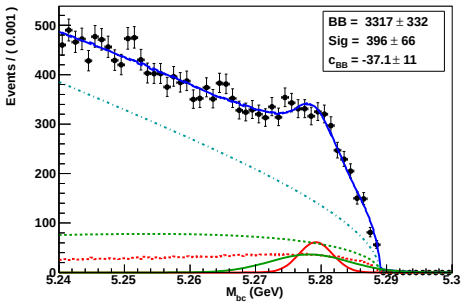
(b) $K\pi$ with π^0



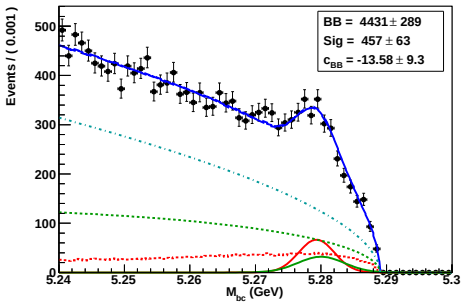
(c) $K2\pi$ without π^0



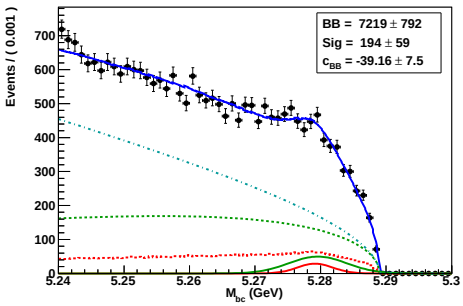
(d) $K2\pi$ with π^0



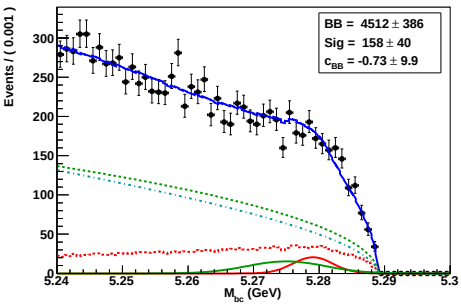
(e) $K3\pi$ without π^0



(f) $K3\pi$ with π^0

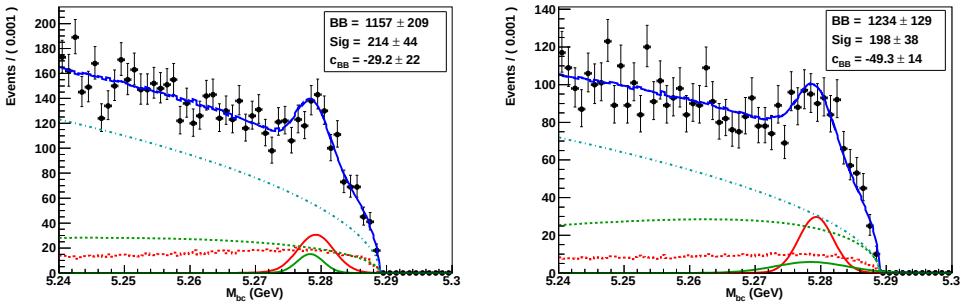


(g) $K4\pi$



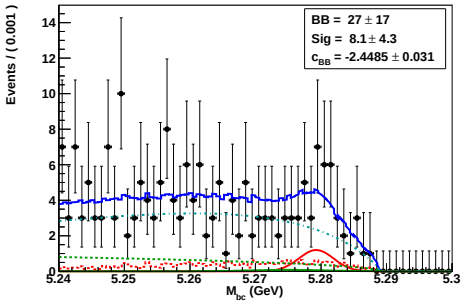
(h) $K2\pi^0$ with at most 2π

Figure 7.7: M_{bc} fit for each final state(MC) at 1.15< M_{X_s} <2.80 (Red solid line : Signal, Red dashed line: Signal cross-feed, Green solid line: Peaking BG)

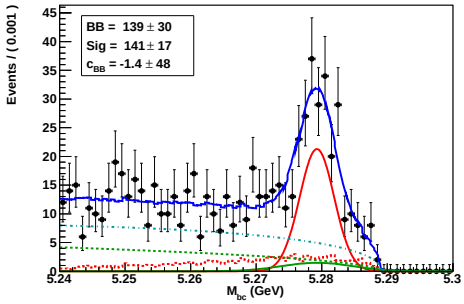


(a) $K\eta$ (b) $3K$

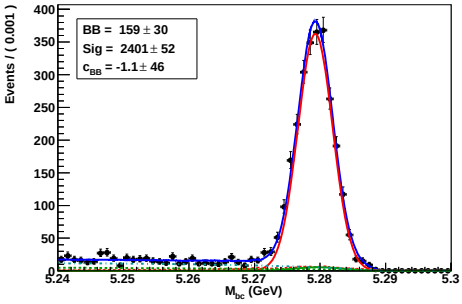
Figure 7.8: M_{bc} fit for each final state(MC) at 1.15< M_{X_s} <2.80 GeV/c^2 (Red solid line : Signal, Red dashed line: Signal cross-feed, Green solid line: Peaking BG)



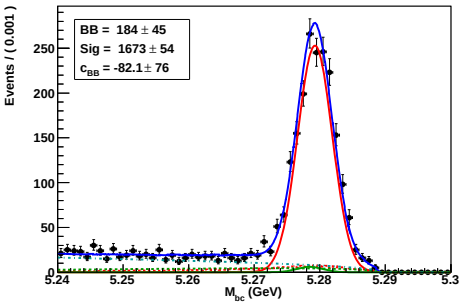
(a) 0.6< M_{X_s} <0.7 GeV



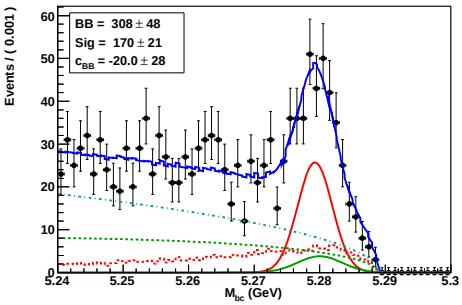
(b) 0.7< M_{X_s} <0.8 GeV



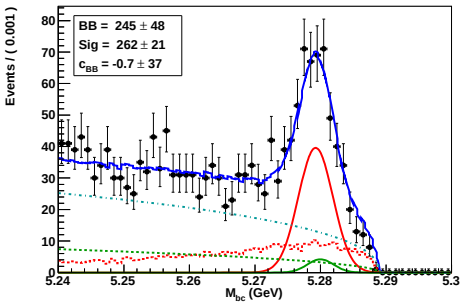
(c) 0.8< M_{X_s} <0.9 GeV



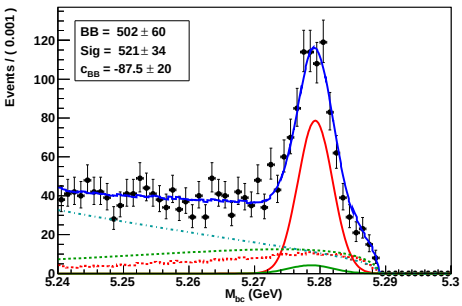
(d) 0.9< M_{X_s} <1.0 GeV



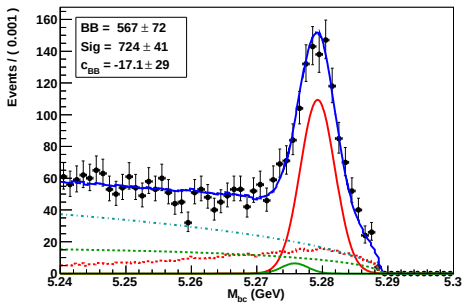
(e) 1.0< M_{X_s} <1.1 GeV



(f) 1.1< M_{X_s} <1.2 GeV

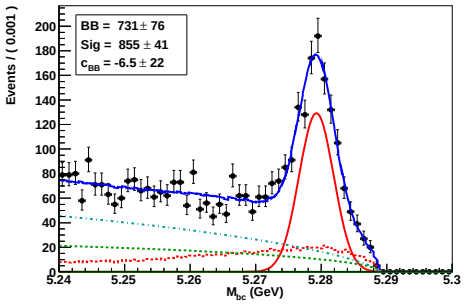


(g) 1.2< M_{X_s} <1.3 GeV

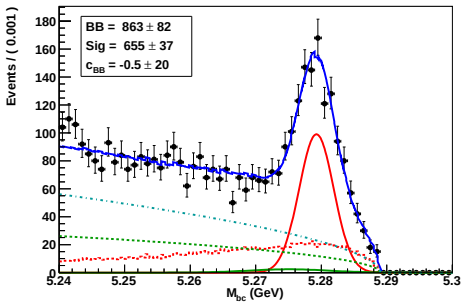


(h) 1.3< M_{X_s} <1.4 GeV

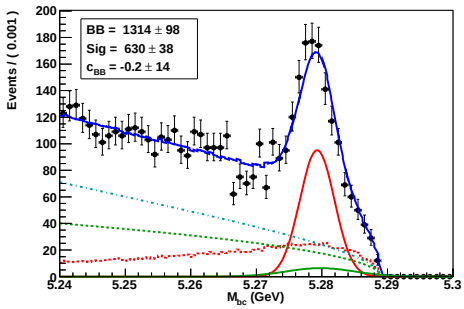
Figure 7.9: M_{bc} fit for each M_{X_s} bin(0.6-1.4 GeV, MC)



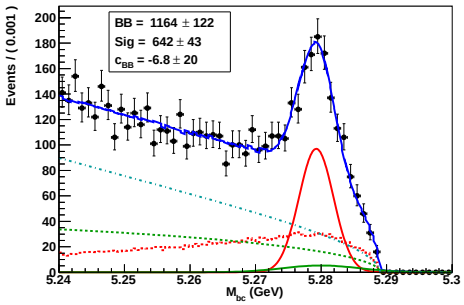
(a) 1.4< M_{X_s} <1.5 GeV



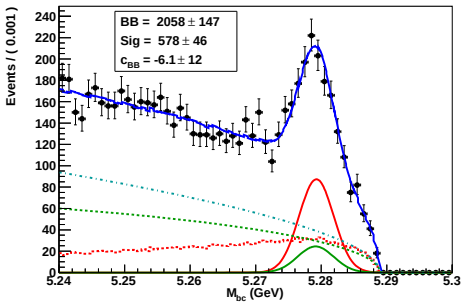
(b) 1.5< M_{X_s} <1.6 GeV



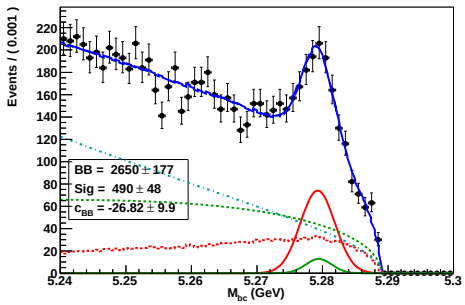
(c) 1.6< M_{X_s} <1.7 GeV



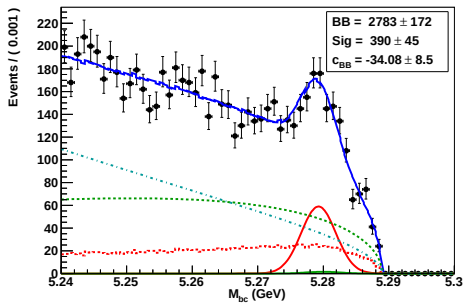
(d) 1.7< M_{X_s} <1.8 GeV



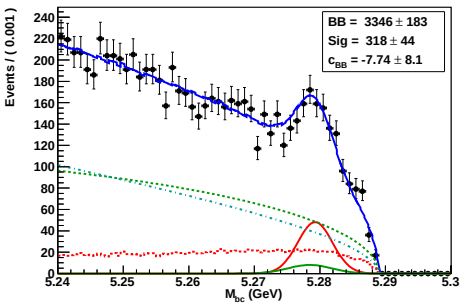
(e) 1.8< M_{X_s} <1.9 GeV



(f) 1.9< M_{X_s} <2.0 GeV

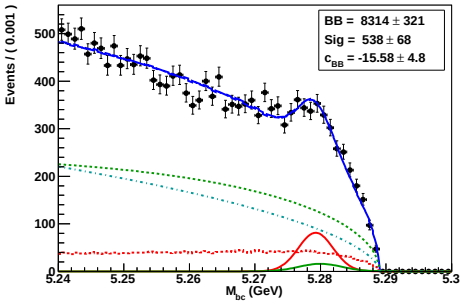


(g) 2.0< M_{X_s} <2.1 GeV

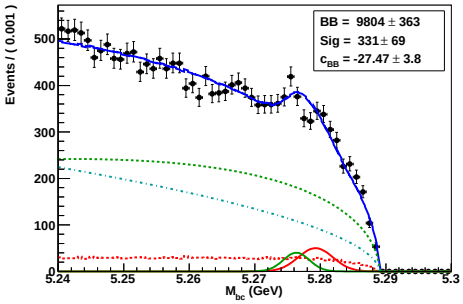


(h) 2.1< M_{X_s} <2.2 GeV

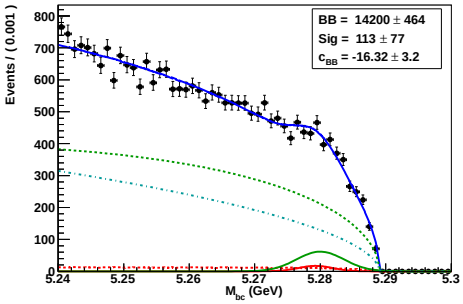
Figure 7.10: M_{bc} fit for each M_{X_s} bin(1.4-2.3 GeV, MC)



(a) 2.2< M_{X_s} <2.3 GeV



(b) 2.4< M_{X_s} <2.6 GeV



(c) 2.6< M_{X_s} <2.8 GeV

Figure 7.11: M_{bc} fit for each M_{X_s} bin(2.3-2.8 GeV, MC)

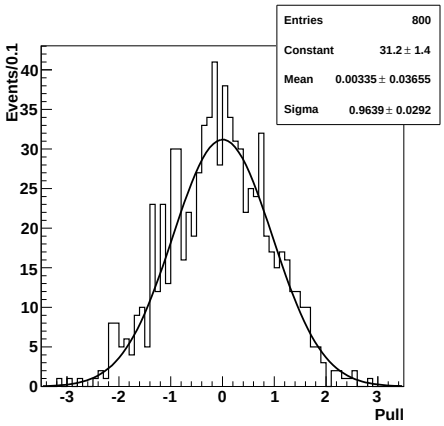
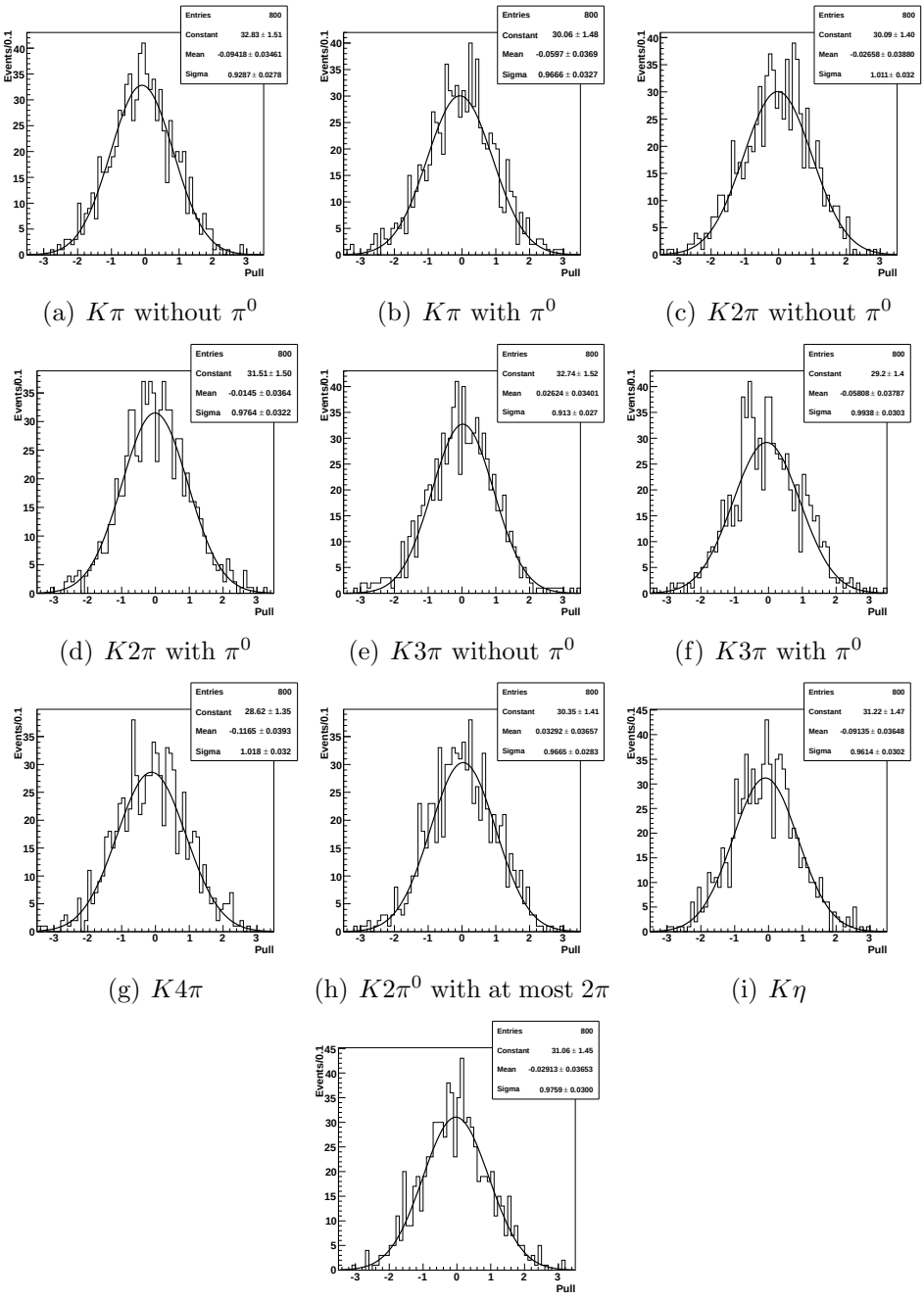


Figure 7.12: Pull distribution for total M_{bc} fit



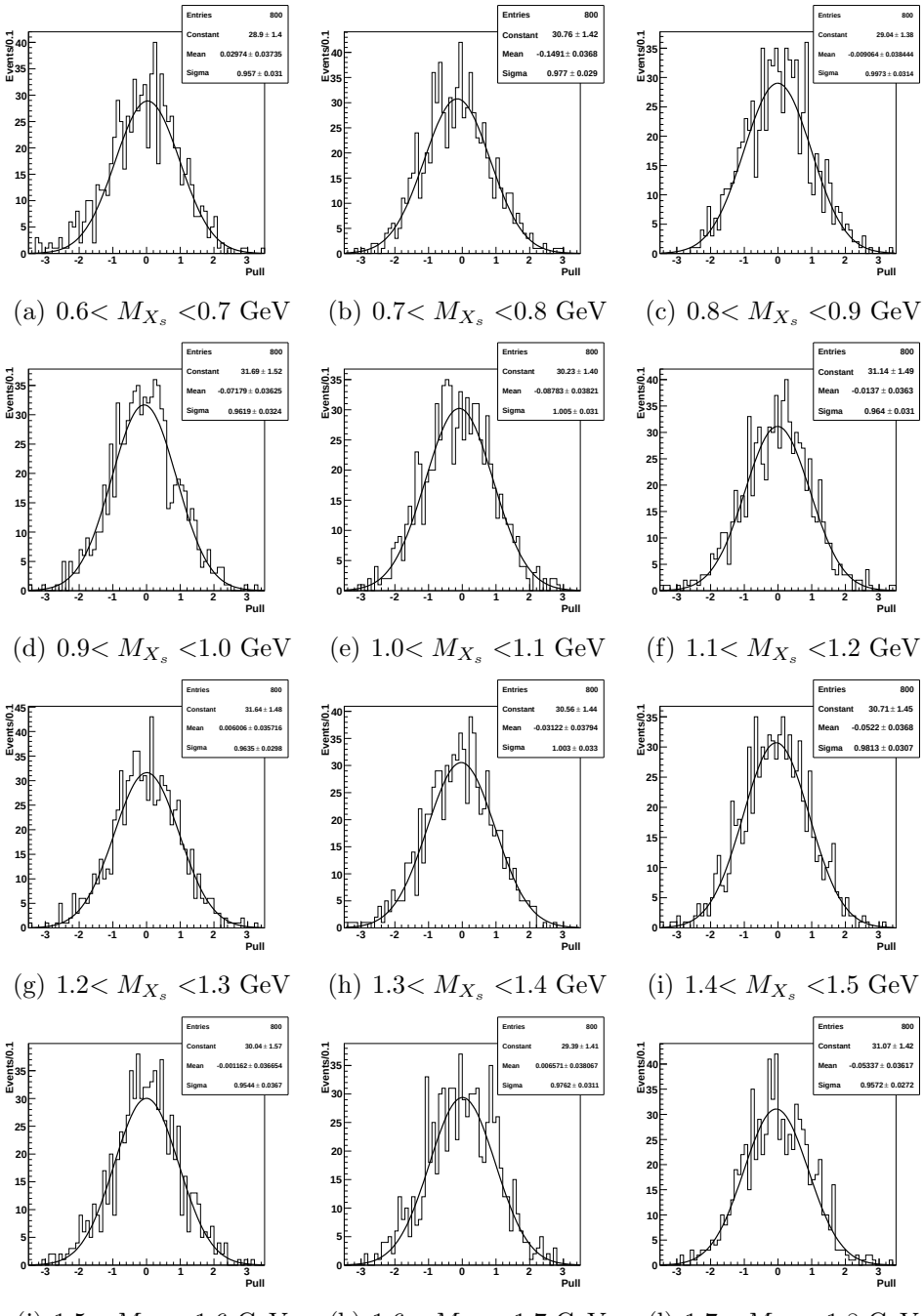
(a) $K\pi$ without π^0 (b) $K\pi$ with π^0 (c) $K2\pi$ without π^0

(d) $K2\pi$ with π^0 (e) $K3\pi$ without π^0 (f) $K3\pi$ with π^0

(g) $K4\pi$ (h) $K2\pi^0$ with at most 2π (i) $K\eta$

(j) $3K$

Figure 7.13: Pull distribution for each mode



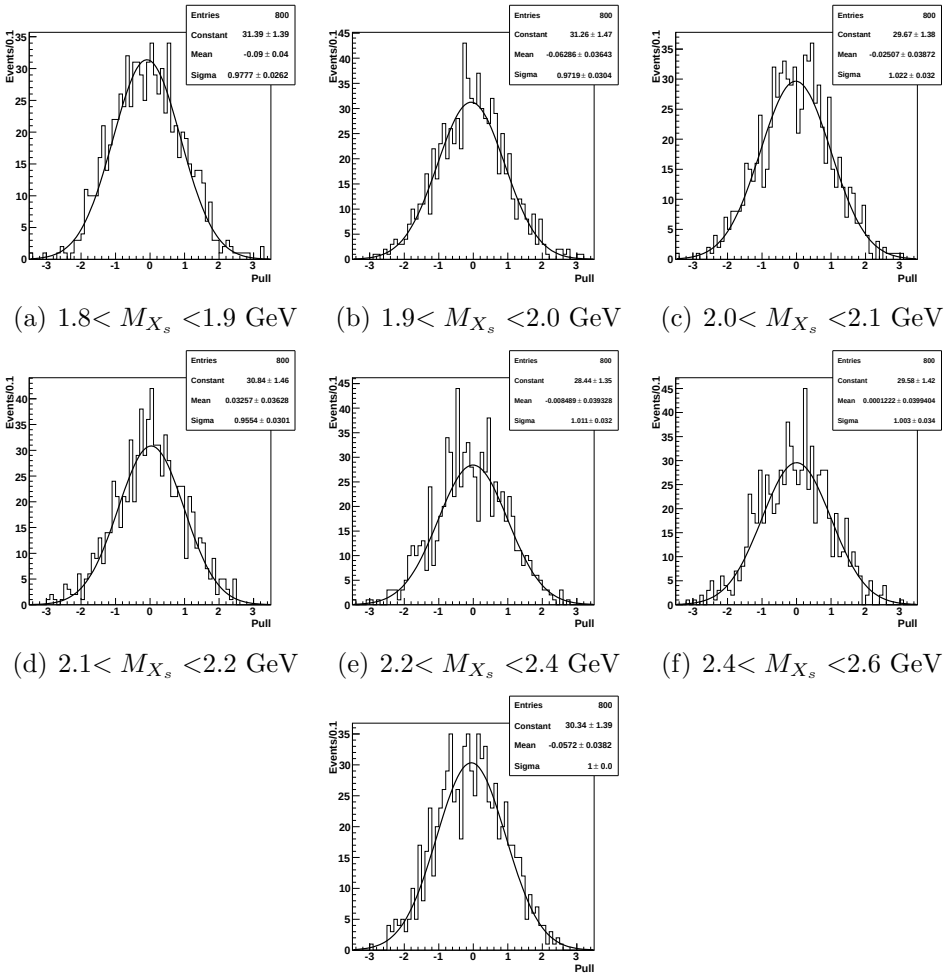
(a) 0.6< M_{X_s} <0.7 GeV (b) 0.7< M_{X_s} <0.8 GeV (c) 0.8< M_{X_s} <0.9 GeV

(d) 0.9< M_{X_s} <1.0 GeV (e) 1.0< M_{X_s} <1.1 GeV (f) 1.1< M_{X_s} <1.2 GeV

(g) 1.2< M_{X_s} <1.3 GeV (h) 1.3< M_{X_s} <1.4 GeV (i) 1.4< M_{X_s} <1.5 GeV

(j) 1.5< M_{X_s} <1.6 GeV (k) 1.6< M_{X_s} <1.7 GeV (l) 1.7< M_{X_s} <1.8 GeV

Figure 7.14: Pull distribution for each M_{X_s} bin(0.6-1.8 GeV)



(a) 1.8< M_{X_s} <1.9 GeV (b) 1.9< M_{X_s} <2.0 GeV (c) 2.0< M_{X_s} <2.1 GeV

(d) 2.1< M_{X_s} <2.2 GeV (e) 2.2< M_{X_s} <2.4 GeV (f) 2.4< M_{X_s} <2.6 GeV

(g) 2.6< M_{X_s} <2.8 GeV

Figure 7.15: Pull distribution for each M_{X_s} bin(1.8-2.8 GeV)

Chapter 8

Systematic Uncertainties

In this chapter, we investigate the systematic uncertainties which can be evaluated without the data in the signal region.

8.1 Uncertainty in Number of B Mesons

As we discussed in section 4.1.2, the number of $B\overline{B}$ pairs used in this analysis is (771.9 $\pm$ 10.57) $\times 10^6$. The relative uncertainty is 1.37 %.

8.2 Detector Response Uncertainties

8.2.1 High-energy Photon Reconstruction

The detection efficiency of high energy photons(with typical energy above $E_\gamma \sim$ 2 GeV) is measured using radiative Bhabha events : $e^+e^- \to e^+e^-\gamma$[62]. After requiring exactly two tracks in an event that are identified as an e^+e^- pair, the missing energy direction can be computed. The reconstruction efficiency is estimated from the fraction of events that have a reconstructed photon matching the magnitude and direction of the missing energy. The recontruction efficiency in MC agrees with that in data and the systematic uncertainty is 2.00 %. In this analysis, 2.00 % is assigned to the systematics of the high-energy photon reconstruction efficiency.

8.2.2 Reconstruction of Particles from X_s

The reconstruction efficiencies of particles from X_s, charged particles($K^\pm, \pi^\pm$), π^0, η, K_s are evaluated every final state and summed up with fractions of the final states.

◯ **Charged Particle Reconstruction** The systematic uncertainty on the charged particle tracking with high momentum ($p > 200$ MeV/c) is evaluated by using the decay chain of $D^* \to D^0\pi, D^0 \to \pi^+\pi^- K_s$ and $K_s \to \pi^+\pi^-$, which provides a very clean sample with sufficient statistics[63]. The decay can be reconstructed without actually detecting one of the pions from the K_s decay. The four-momentum of this pion can be inferred from the kinematic constraints of the decay chain. The ratio of the yield of such partially reconstructed D^* to those fully reconstructed with both pions from the K_s detected is the track reconstruction efficiency. As the result the data-MC ratio for charged-track reconstruction with high momentum track is (99.87±0.32)%.

The efficiency of low momentum tracks ($p < 200$ MeV/c) is evaluated by using the sample of $B^0 \to D^*\pi_s$ followed by $D^* \to D^0\pi$, which provides a large sample of slow pions [68]. The data-MC ratios on efficiency are (102.0±3.48)% (SVD1) and (98.6±1.36)% (SVD2).

◯ **Particle π^0 and η Reconstructions** The reconstruction efficiency on π^0 and $\eta \to \gamma\gamma$ can differ between data and MC mainly for the imperfect modeling of the material distribution in detector and the photon shower shape. π^0 and η reconstruction efficiency is evaluated by inclusive η decay[67]. The π^0 efficiency is (92.4± 1.42)% and η is (100± 2.00)%. The slow π^0 and $\eta(p < 200$ MeV/c) is (102± 3.48)%[SVD1] and (98.6± 1.36)%[SVD2][67].

◯ **Particle K_s Reconstruction** In this uncertainty study, a fully reconstructed D^* decay is used and it provides a clean sample for K_s efficiency[66]. The decay chain of interest is $D^* \to \pi_s D, D^0 \to K_s\pi^+\pi^-$ As a result, the data-MC ratios on efficiency are (98.96±1.03)% (SVD1) and (98.05±0.49)% (SVD2).

$\epsilon_{Data}/\epsilon_{MC}$ on the reconstruction efficiency is summarized in Table 8.1. Totally, the data-MC ratios on the reconstruction efficiency of particles from X_s is (97.1±1.29)%.

8.2.3 Kaon and Pion Identification Efficiency

The uncertainty in K and π identification efficiency(Partile ID;PID) is estimated using inclusive D^* decay sample, $D^{*+} \to D^0\pi^+$ followed by $D^0 \to K^-\pi^+$[69]. Figure 8.1 shows typical curves of the efficiencies and mis-identification rates for the kaon identification in the barrel region. Discrepancies between

data and MC can be seen, especially in the mis-identification distribution. For each bin, the efficiency and mis-identification rate for K and π is estimated both for the data and MC. These quantities are provided as a look-up table and the systematic uncertainty is can be calculated by using the table. In this analysis, K and π identification efficiencies are calculated every final state by the table and added with weight of a fraction of the final state after all selections.

$\epsilon_{Data}/\epsilon_{MC}$ on the PID efficiency is summarized in Table 8.1. Totally, the efficiency ratio of data to MC is $(96.6 \pm 1.79)\%$.

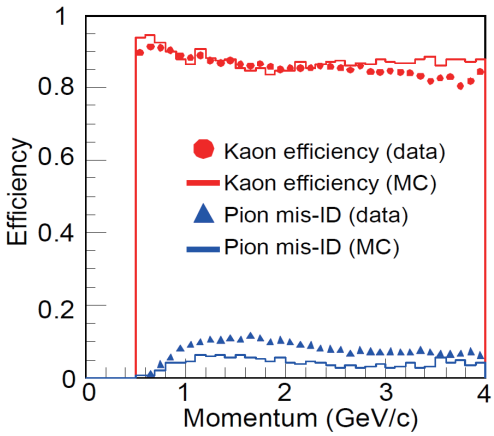


Figure 8.1: Kaon efficiency

8.3 Background Suppression Uncertainties

8.3.1 D Veto Uncertainty

The D veto uncertainty is evaluated by using a control sample, $B \to X_s J/\psi$ decay followed by $J/\psi \to ll(l = e, \mu)$.

Reconstruction of $B \to X_s J/\psi$ The electron(muon) candidates are required to have momentum above 0.40(0.80) GeV/c and the electron(muon) probability more than 0.80(0.97). The electron(muon) pair for J/ψ candidate must have an invariant mass between 3.02(3.05) and 3.12 (3.12) GeV/c^2. The X_s is reconstructed as same final states in the signal, where the X_s in $J/\psi X_s$ has a mass below about 2.0 GeV/c^2 since J/ψ mass is 3.097 GeV/c^2. However, we have to observe the D veto window, $M_{X_s} > 2.0$ GeV/c^2, using $J/\psi X_s$ sample. The issue can be resolved through adding a lepton from J/ψ in the

Table 8.1: Correction($\epsilon_{Data}/\epsilon_{MC}$) on Reconstruction and K/π PID efficiency

Mode ID	Reconstruction	PID	Mode ID	Reconstruction	PID
1	0.997±0.007	0.962±0.016	22	0.835±0.045	0.971±0.022
2	0.981±0.009	0.963±0.008	23	0.999±0.023	1.003±0.009
3	0.923±0.017	0.999±0.008	24	0.982±0.026	1±0
4	0.908±0.020	1±0	25	0.997±0.027	0.987±0.020
5	0.996±0.011	0.946±0.026	26	0.981±0.030	0.979±0.010
6	0.979±0.014	0.941±0.018	27	0.923±0.037	1.006±0.010
7	0.921±0.021	0.977±0.018	28	0.908±0.040	1±0
8	0.906±0.024	0.972±0.009	29	0.995±0.033	0.980±0.031
9	0.994±0.016	0.944±0.038	30	0.979±0.034	0.976±0.022
10	0.977±0.018	0.936±0.029	31	0.921±0.041	0.998±0.022
11	0.920±0.026	0.969±0.029	32	0.906±0.044	0.993±0.009
12	0.904±0.029	0.960±0.019	33	0.996±0.010	1.030±0.029
13	0.991±0.022	0.950±0.050	34	0.980±0.012	1.023±0.020
14	0.975±0.024	0.947±0.040	35	0.964±0.015	1.021±0.022
15	0.917±0.032	0.971±0.041	36	0.994±0.014	1.030±0.042
16	0.903±0.034	0.959±0.032	37	0.978±0.017	1.015±0.031
17	0.853±0.032	1.008±0.010	38	0.920±0.024	1.044±0.033
18	0.839±0.034	1±0			
19	0.851±0.036	0.990±0.020			
20	0.837±0.038	0.982±0.011			
21	0.849±0.043	0.989±0.031			
Total	0.971±0.0129	0.966±0.0179			

X_s children. If a X_s is charged, a lepton from J/ψ which has a different charge from that of the X_s is reconstructed in the X_s children, while if a X_s is neutral, a lepton with lower energy is reconstructed in X_s children. The lepton added in the X_s children is treated as a charged pion and another lepton from J/ψ is treated as a primary photon of the signal. ΔE are required to be between -0.06(-0.03) and 0.03(0.03) GeV for $X_s ee(X_s \mu\mu)$. In Figure 8.2, D mass distribution of $B \to X_s J/\psi$ have a broad peak at nominal D mass region as that of the signal, therefore $B \to X_s J/\psi$ is useful for the control sample of D veto.

Systematic uncertainty on D veto M_{bc} distributions before/after D veto are shown in Figure 8.3. The efficiencies on MC(ϵ_{MC}) and data(ϵ_{data}) are 91.54 % and 90.98 %, respectively. Since the number of $X_s J/\psi$ with M_{X_s} >2.0 GeV/c^2 is larger than that of $X_s\gamma$, the D veto efficiency of $X_s J/\psi$ is lower than that of $X_s\gamma$ (97 %). The efficiency discrepancy($\epsilon_{data}/\epsilon_{MC} - 1$) on D veto between the MC and data, which is taken as the systematic uncertainty, is 0.61 %.

8.3.2 $q\overline{q}$ Background Suppression Uncertainty

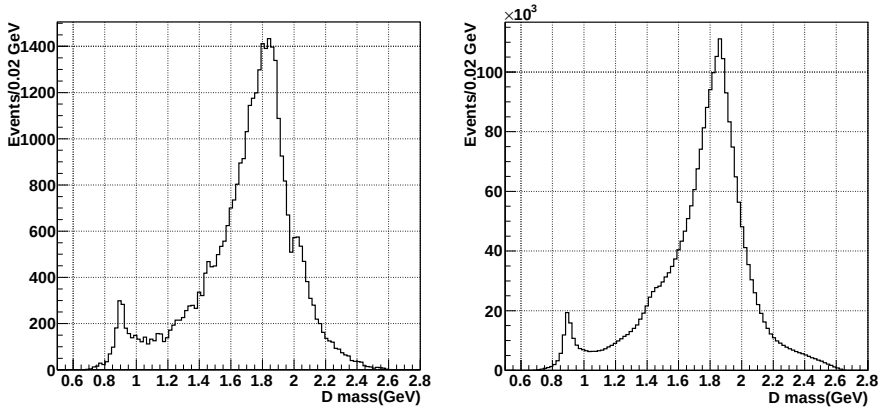
The $q\overline{q}$ suppression uncertainty is evaluated by using a control sample, $B \to D\pi$ decay, which provides a clean sample with sufficient statistics.

Reconstruction of $B \to D\pi$ In $B \to D\pi$ reconstruction, a pion from B is treated as the primary photon of the signal. We should note that the charge of the D is different from that of X_s, for instance, X_s^0 decays into $K^+\pi^-$ while D^0 decays into $K^-\pi^+$. Since D does not decay all of the final states in the signal, only 22 of the 38 signal final states are reconstructed.

There are a few differences on selection of the signal. In the primary photon(π) selection, the E_9/E_{25} cut and π^0/η vetoes are not applied. In $X_s(D)$ selection, D veto is not applied and $X_s(D)$ mass is required to be between 1850 and 1880 MeV/c^2. ΔE is required to be between -0.03 and 0.03 GeV.

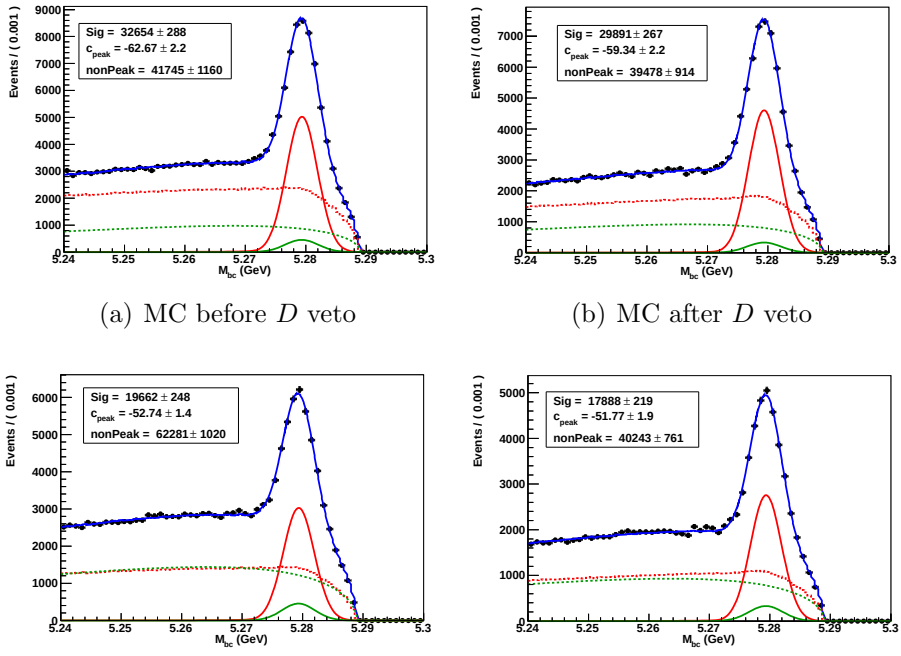
The NB is trained by the same method as $B \to X_s\gamma$ in Section 6.2.2 and the cut value of the NB output optimized by significance is -0.10 (Figure 8.4). In Figure 8.5, M_{bc} distributions of MC and data before/after $q\overline{q}$ suppression are shown.

Systematic Uncertainty on $q\overline{q}$ background suppression The efficiencies on MC(ϵ_{MC}) and data(ϵ_{data}) are 91.17 % and 88.40 %, respectively. The



(a) $B \to X_s\gamma$ (b) $B \to X_s J/\psi$

Figure 8.2: D mass distribution in $M_{X_s} > 2.0$ GeV/c^2



(a) MC before D veto (b) MC after D veto

(c) DATA before D veto (d) DATA after D veto

Figure 8.3: M_{bc} distributions before/after D veto(Red solid line: Signal, Red dashed line, cross-feed, Green solid line:Peaking background, Green dashed line:Non-peaking background)

efficiency discrepancy($\epsilon_{data}/\epsilon_{MC} - 1$) between the MC and data, which is taken as the systematic uncertainty, is 3.04% .

Since only 22 reconstruction modes are used, we investigate an effect on the other modes(38-22=17 modes) by using the signal MC. The fraction of the 22 modes is 87 %, which is large enough and the fraction is not changed before

and after $q\overline{q}$ suppression. Thus, we conclude that the other modes do not have a large effect on the efficiency.

8.3.3 Best Candidate Selection(BCS) Uncertainty

The Best candidate selction(BCS) uncertainty is evaluated by using a control sample, $B \to X_s J/\psi$ decay followed by $J/\psi \to ll(l = e, \mu)$.

Reconstruction of $B \to X_s J/\psi$ The reconstruction procedure and selection are same as them of the D veto uncertainty study(Section 8.3.1). In this uncertainty study, we use $B \to X_s J/\psi$ sample after $q\overline{q}$ background suppression, as in the signal. Since a number of $q\overline{q}$ background in the $X_s J/\psi$ is small highly, we do not have enough MC sample for NB training. Therefore, $q\overline{q}$ and $B\overline{B}$ background MC samples are used for the training. In a calculation of ΔE likelihood, which is one of the input variables, two ΔE PDFs for $X_s ee$ and $X_s \mu\mu$ are provided to remove a bias in efficiencies between the two final states.

Systematic uncertainty on BCS M_{bc} distributions before/after BCS are shown in Figure 8.6. The efficiencies on MC(ϵ_{MC}) and data(ϵ_{data}) are 94.83 % and 95.93 %, respectively. The efficiency discrepancy($\epsilon_{data}/\epsilon_{MC} - 1$) on the BCS between the MC and data, which is taken as the systematic uncertainty, is 1.16 %.

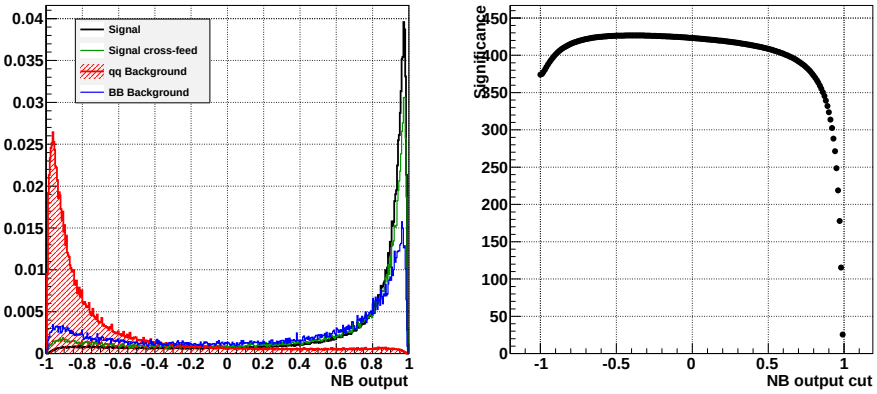
8.4 M_{bc} PDF Uncertainties

In study of M_{bc} PDF uncertainties, the data in the signal region are needed. We evaluate it after the signal box opening in Section 9.5.3 and 10.4.3.

8.5 Signal Modeling Uncertainties

8.5.1 M_{X_s} Shape Uncertainty

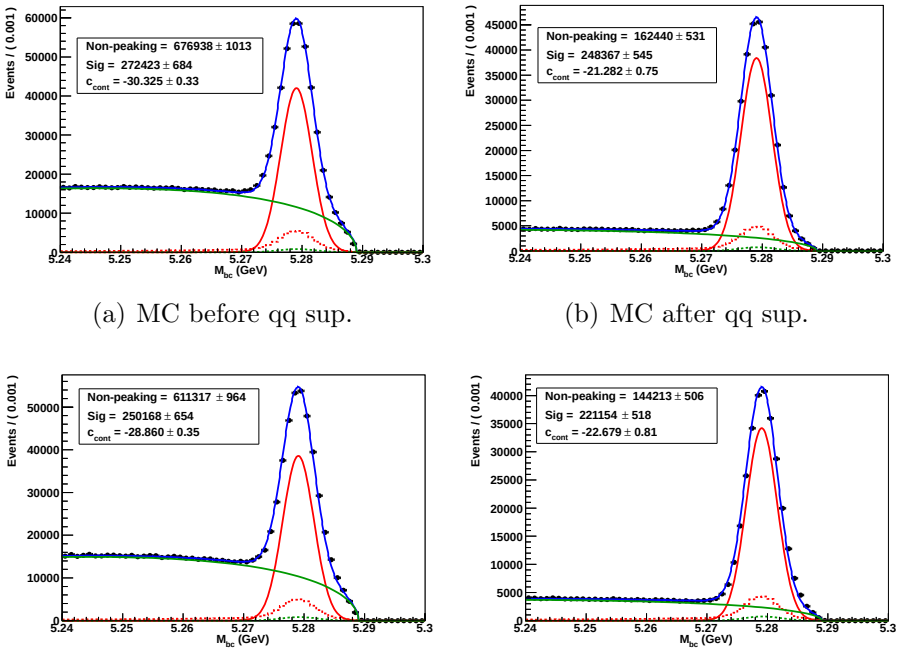
In the inclusive signal MC(M_{X_s} >1.15 GeV/c^2), the M_{X_s} shape is produced according to the Kagan-Neubert model(KN) [56], which has two parameters, the b quark mass, m_b, and the b quark momentum parameter in B meson, μ_π^2, as mentioned in Section 4.2.1. Their default values are selected to minimize χ^2 to the Belle's data as shown in Figure 4.1. The χ^2 are calculated in 1.75< E_γ^* <2.6 GeV because the region in $E_\gamma^* \leq$1.75 GeV has large uncertainties and does not



(a) NeuroBayes Output

(b) Significance as a function of NB output cut value

Figure 8.4: NeuroBayes output and Significance($B \to D\pi$ control sample)



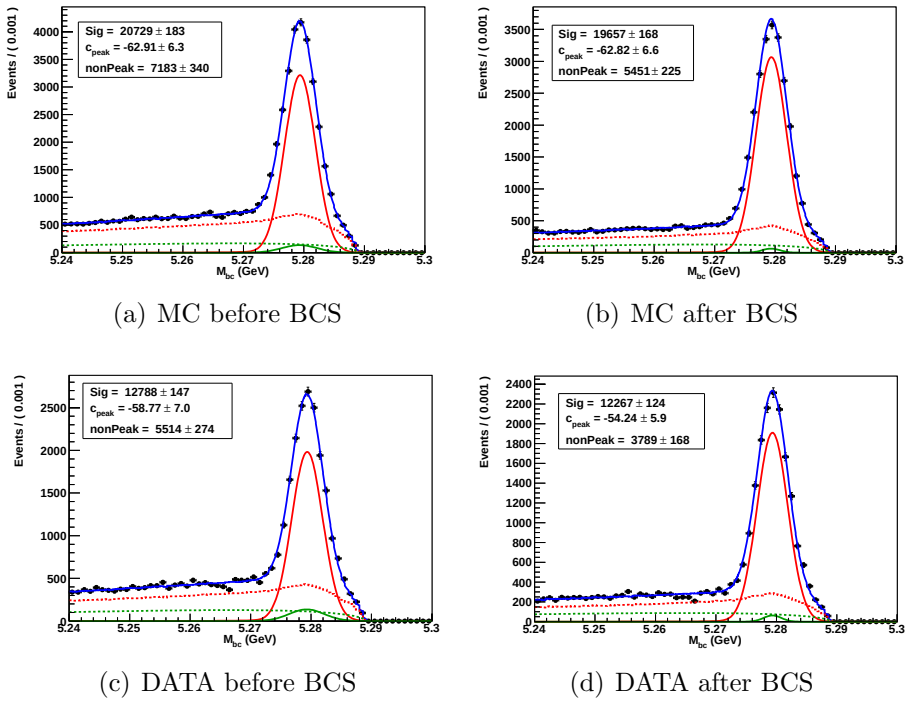
(a) MC before qq sup.

(b) MC after qq sup.

(c) DATA before qq sup.

(d) DATA after qq sup.

Figure 8.5: M_{bc} distributions before/after $q\overline{q}$ suppression(Red solid line: Signal, Red dashed line, cross-feed, Green solid line:Peaking background, Green dashed line:Non-peaking background)



(a) MC before BCS (b) MC after BCS

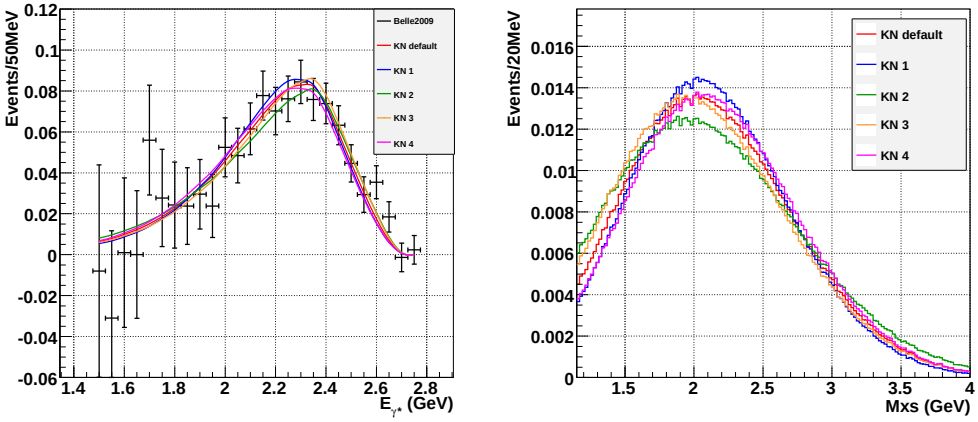
(c) DATA before BCS (d) DATA after BCS

Figure 8.6: M_{bc} distributions before/after BCS(Red solid line: Signal, Red dashed line, cross-feed, Green solid line:Peaking background, Green dashed line:Non-peaking background)

have large differences by changing the parameters, and the region in E^*_γ $\geq$2.6 GeV has an effect of $K^*\gamma$. The parameters(m_b and μ_π^2) in the KN model are fluctuated to evaluate the systematic uncertainty in Figure 8.7(a), where the parameters are fluctuated by $\sqrt{\chi^2 - \chi^2_{default}} = 1$, which corresponds to the σ. The parameter settings are summarized in Table 8.2. Figure 8.7(b) shows M_{X_s} distributions fluctuated the parameters. The signal efficiencies on the KN1-4 models are investigated and the deviations from the default are included in Table 8.2. The deviations are taken as the systematic uncertainty, (+3.26 -7.96)%.

This uncertainty is large, thus, it may be better to adopt M_{X_s} bin analysis, in which M_{bc} distributions on each M_{X_s} bin are fitted to extract signal yields and branching ratios on each M_{X_s} bin are summed to obtain the total one. In this method, the large M_{X_s} shape uncertainty is avoidable since we use signal efficiencies in each M_{bc} bin, where the shape information is not used. However, the statistical uncertainties for each M_{X_s} bin should be large. We discuss this issue with the partial data in Chapter 9.

We also investigate another model, Dressed Gluon Exponentiation(DGE)[70],



(a) Photon energy distribution at CM frame. Plot shows Belle's result with Full-inclusive method[71] and a red line shows the default signal MC. Other lines are fluctuated to evaluate the systematic uncertainty

(b) M_{X_s} distributions fluctuated along with the photon shape.

Figure 8.7: Systematics study on M_{X_s} shape

Table 8.2: Parameter setting in KN model for the systematics on M_{X_s}

	m_b (GeV/c^2)	μ_π^2 (GeV^2)	χ^2	$\sqrt{\chi^2 - \chi^2_{default}}$	Signal efficiency(%)	Deviation(%)
Default	4.440	0.750	2.947		2.098	
KN1	4.440	0.630	4.110	1.08	2.034	-3.04
KN2	4.440	1.000	4.114	1.08	2.131	1.59
KN3	4.480	0.750	4.034	1.04	2.166	3.26
KN4	4.405	0.750	4.082	1.07	1.931	-7.96

which has two parameters, $\alpha_s(M_Z), m_b(\overline{MS})$. In Figure 8.8, photon energy distributions with $M_{X_s} > 1.15$GeV in KN and DGE are shown together with that of the Belle's result[71]. DGE distributions in Figure 8.8 have 6 different parameter sets in Figure 8.9. As shown in Figure 8.8, DGE distributions are greatly different from it, therefore, we do not use DGE model for the systematics study.

8.5.2 Hadronization Model Uncertainty

The fragmentation of the hadronic system in the inclusive region, M_{X_s} >1.15 GeV/c^2, is modeled in Pythia. The signal efficiency depends on the particle

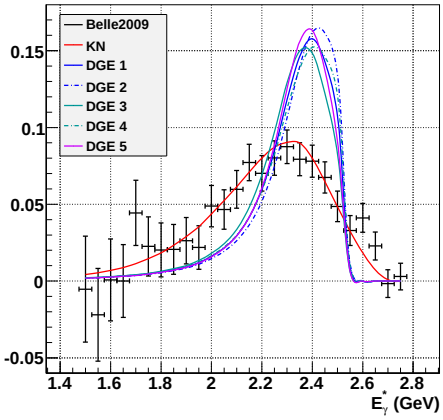


Figure 8.8: E^*_γ distributions

	$\alpha_s(M_Z)$	$m_b(\overline{MS})$
DGE 1	0.1176	4.20
DGE 2	0.1176	4.24
DGE 3	0.1176	4.16
DGE 4	0.1196	4.20
DGE 5	0.1156	4.20

Figure 8.9: Parameter sets in DGE

content of the final states and the uncertainty from the fragmentation model is large. In the previous Belle analysis[34], it was identified that the fragmentation in MC was greatly different from that of data. Thus, a calibration on the hadronization model in the MC is essential for the correct result with smaller uncertainty. The procedure of the calibration is discussed by using partial data(140 fb^{-1}) in Chapter 9.

8.5.3 Missing Final States Uncertainty

The fraction of missing final states that is not included in our reconstructed modes has also a dependence on the breakdown of the X_s decay. Thus, the differences on this fraction between the MC and the data should be evaluated as the systematic uncertainty. This study is performed by using the MC after the calibration on the hadronization model in Section 9.5.2 and 10.4.2.

8.5.4 $K^* - X_s$ Transition Uncertainty

The position of $K^* - X_s$ transition in the signal MC is fixed at 1.15 GeV/c^2 as mentioned in Section 4.2.1. The position will be fluctuated after the signal box opening to evaluate the systematic uncertainty.

8.6 Summary of Systematic Uncertainties

The systematic uncertainties until now are summarized in Table 8.3. Blank spaces in this table are evaluated by using real data later. The uncertainty from the M_{X_s} shape is dominant, thus, it should be suppressed for the precision

measurement. It can be avoidable by evaluating branching ratios in M_{X_s} bins which are divided finely since the M_{X_s} shape information in the MC is not used. In addition, the uncertainty from the hadronization model is expected to be large. The uncertainty comes from a difference on the model between the data and MC. Thus, the hadronization model in the MC should be calibrated to that of the data. We discuss methods to suppress these uncertainties with the data in next chapter.

Table 8.3: Systematic uncertainty(%). Blank spaces are evaluated by using real data later.

Source		Uncertainty(%)
Number of $B\overline{B}$		± 1.37
Detector response	γ detection	± 2.00
	X_s particles(K, K_s, π, π^0, η) reconstruction	± 1.29
	K/π separation	± 1.79
Background rejection	π^0 veto	± 0.30
	η veto	± 0.60
	D veto	± 0.61
	$q\overline{q}$ suppression	± 3.04
	Best candidate selection	± 1.16
M_{bc} PDF	Signal PDF	
	Signal Cross-feed PDF	
	Peaking Background PDF	
	Non-peaking part from $q\overline{q}$ background	
Signal modeling	M_{X_s} shape	+3.26 -7.96
	Hadronization model	
	Missing mode	
	$K^* - X_s$ transition	
Total		±9.2

Chapter 9

Partial Data Analysis

The dominant sources of the systematic uncertainty are expected to be the signal modeling in the MC, specifically, the hadronization model in the X_s decay and the M_{X_s} shape. To suppress these uncertainties, we discuss the calibration method on the hadronization model and M_{bc} fit procedure with a partial data in this chapter.

140 fb^{-1} data which corresponds to one-fifth to the full data and was already analyzed in CP study of the $B \to X_s\gamma$ at the Belle[34] is used.

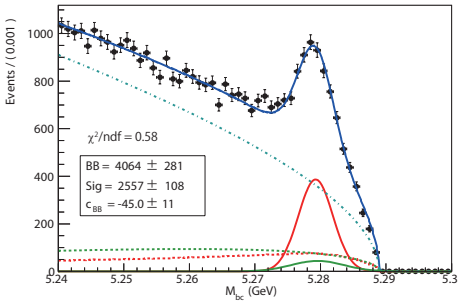
Before using data in the signal region, we study $\mathcal{BR}(B \to D\pi)$ to confirm a validity on this analysis and obtain a consistent result with the PDG in Appendix E.

9.1 Signal Yield in 140 fb^{-1} data

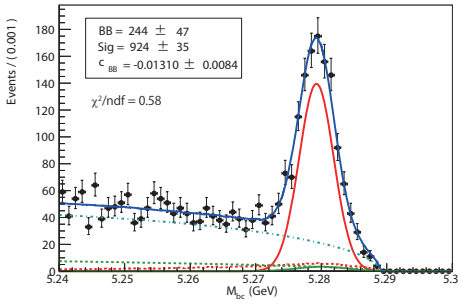
At first, we evaluate the signal yield in the 140 fb^{-1} data from the M_{bc} distribution. Figure 9.1(a) shows M_{bc} distribution in 0.0< M_{X_s} <2.8 GeV/c^2 and the signal yield obtained from the fit is 2557±108.

9.2 Branching Fraction of $B \to K^*\gamma$

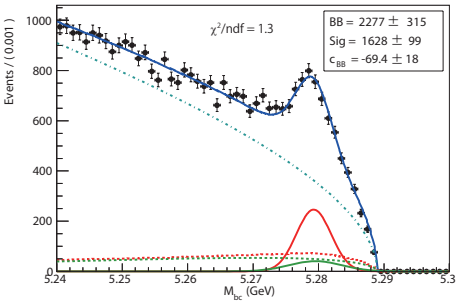
In Figure 9.1(b), M_{bc} fit in M_{X_s} <1.15 GeV/c^2 which corresponds to exclusive $K^*\gamma$ region is shown. Before evaluating the total branching fraction we try to obtain the branching fraction of $B \to K^*\gamma$ at first to confirm a validity on our measurement. The fit result is (924±35) and the signal efficiency obtained from the MC is 6.82 % which is corrected by the reconstruction and PID



(a) 0< M_{X_s} <2.8 GeV/c^2



(b) M_{X_s} <1.15 GeV/c^2



(c) 1.15< M_{X_s} <2.8 GeV/c^2

Figure 9.1: M_{bc} fit(140 fb^{-1} data)(Red solid line: Signal, Red dashed line: Signal cross-feed, Green solid line: Peaking BG, , Green dashed line: Non-peaking background from $B\overline{B}$, Cyan dashed line:Non-peaking background from $q\overline{q}$)

efficiencies. We can calculate the $\mathcal{BR}(B \to K^*\gamma)$ as

$$\mathcal{BR} = \frac{N_{sig}}{2N_{B\overline{B}}\epsilon}, \tag{9.1}$$

where N_{sig} is the number of the signal, 924±35, $N_{B\overline{B}} = 154.3\times10^6$ is the number of $B\overline{B}$ pairs in the 140 fb^{-1} data and ϵ is the signal efficiency. Thus, the branching fractionn is

$$\mathcal{BR}(B \to K^*\gamma) = (4.38 \pm 0.17) \times 10^{-5}, \tag{9.2}$$

where the uncertainty is only statistical one. According to the PDG,

- $\mathcal{BR}(B^+ \to K^{*+}\gamma) = (4.21\pm0.18)\times10^{-5}$
- $\mathcal{BR}(B^0 \to K^{*0}\gamma) = (4.33\pm0.15)\times10^{-5}$.

Thus,

$$\begin{aligned}\mathcal{BR}(B \to K^*\gamma) &= (4.21 \pm 0.18) \times 10^{-5} \times 0.513\ (\Upsilon(4S) \to B^+B^-) \\ &+ (4.33 \pm 0.15) \times 10^{-5} \times 0.487\ (\Upsilon(4S) \to B^0\overline{B}^0) \\ &= (4.26 \pm 0.17) \times 10^{-5}. \end{aligned} \quad (9.3)$$

We can obtain the branching fraction of the $B \to K^*\gamma$ which is consistent with the PDG value within 0.5σ and confirm the validity on our measurement.

9.3 Calibration on Hadronization Model

The final result depends on the contents in the final states, thus, the hadronization model(Section 4.2) in the inclusive MC is calibrated to data's one. Here, we investigate a procedure and precision for the calibration on the hadronization model in the inclusive signal MC with the 140 fb^{-1} data.

Ten groups of final states we compare data and the MC are given in Table 9.1. In Figure 9.2 and 9.3, M_{bc} distributions of each mode are shown. The fit results are summarized in Table 9.1, where efficiencies corrected by the reconstruction and PID efficiencies are used to calculate branching fractions. The statistical uncertainty in fitting each mode in data is used for the uncertainty on the fraction. Actually, we measure the branching fraction in M_{X_s} <2.8 GeV/c^2, thus, fractions should be obtained in the same region. But, the fractions in Table 9.1 are them in M_{X_s} <2.4 GeV/c^2 since the signal statistics and the signal to the background ratio in M_{X_s} <2.8 GeV/c^2 are low, especially $K4\pi$ and $K2\pi^0$ modes.

Table 9.1: Fit result for each final states(140 fb^{-1} data, 1.15< M_{X_s} <2.4 GeV/c^2)

Mode	Definition	Signal yield	BR(10^{-6})	Fraction(%)
1	$K\pi$ without π^0	211±23	7.72±0.84	5.06±0.89
2	$K\pi$ with π^0	37±13	2.88±1.01	2.53±0.44
3	$K2\pi$ without π^0	483±38	24.3±1.91	17.4±1.37
4	$K2\pi$ with π^0	372±29	44.2±3.44	31.6±2.47
5	$K3\pi$ without π^0	117±27	9.61±2.66	7.00±1.62
6	$K3\pi$ with π^0	110±29	21.2±5.60	15.2±4.01
7	$K4\pi$	43±14	15.9±3.40	11.6±3.80
8	$2\pi^0$	6±20	4.06±13.5	2.91±9.70
9	$K\eta$	50±17	6.53±2.22	4.68±1.59
10	$3K$	35±11	2.70±0.85	1.93±0.61

In Table 9.2, the fractions in the MC are compared to them of the 140 fb^{-1} data.The MC has much higher fractions of $K\pi$ and smaller fractions of $K2\pi$ than them of the data. To calibrate the fractions of the MC, parameters on the hadronization model in Pythia are investigated and we identify the following parameters which have the large impact on the breakdown of final states;

- PARJ(2) : (D=0.30) is the suppression of s quark pair production in the field compared with u or d quark pair production,
- PARJ(11) : (D=0.50) is the probability that a light meson(containing u and d quarks only) has spin 1,
- PARJ(15) : (D=0.05) is the probability that a spin = 1 meson is produced with an orbital angular momentum 1, for a total spin=1,
- PARJ(25) : (D=1.0) is the extra suppression factor for η production in fragmentation,

where D means a default value in the Belle. We try to tune these parameters and are able to correct the signal MC in Table 9.2, where PARJ(2)=0.10, PARJ(11)=0.95, PARJ(15)=0.25, PARJ(25)=0.03. Numbers in () in this table are deviations from the data, defined as (Difference from data)/(Error of the fraction). Total χ^2 is improved from 185 to 22 by this calibration. Next,

Table 9.2: Fractions(%) on each mode in the 140 fb^{-1} data and the MC. Number in () is a deviation to that of the data, defined as ([Fraction in MC]-[Fraction in data])/σ_{data}.

Mode	Fraction in Data	Fraction in MC before calibration	Fraction in MC after calibration
1	5.06±0.89	11.7 (+7.5)	4.76 (-0.3)
2	2.53±0.44	6.16 (+8.2)	2.44 (+0.5)
3	17.4±1.37	13.6 (-2.8)	14.7 (-2.0)
4	31.6±2.47	16.0 (-6.3)	22.4 (-3.7)
5	7.00±1.62	5.66 (-0.8)	5.98 (-0.6)
6	15.2±4.01	15.5 (+0.1)	21.5 (+1.6)
7	11.6±3.80	10.5 (-0.3)	9.36 (-0.6)
8	2.91±9.70	7.72 (+0.5)	7.72 (+0.5)
9	4.68±1.59	4.84 (+0.1)	4.90 (+0.1)
10	1.93±0.61	2.63 (-0.7)	1.76 (-1.3)

we investigate fractions of modes in three M_{X_s} regions, 1.15< M_{X_s} <1.5, 1.5< M_{X_s} <2.0 and 2.0< M_{X_s} <2.4 GeV/c^2, in Table 9.3. The most fractions in the MC are consistent with them of the data within 2σ. This fact means

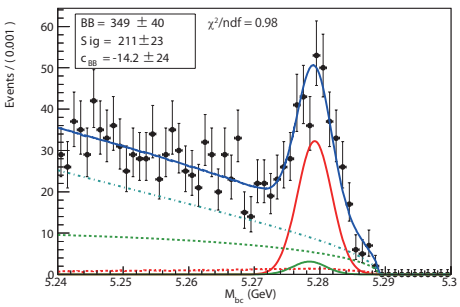
that the modeling by Pythia and the fractions in total M_{X_s} region is effective from the point of view of hadronization models in each M_{X_s} region. However, the fractions of $K\pi$(Mode=1, 2) in 1.15< M_{X_s} <1.5 GeV/c^2 and $K3\pi$ without π^0(Mode 5) in 1.5< M_{X_s} <2.0 GeV/c^2 have deviations more than 2σ, and these effects are included in the systematic uncertainty in Section 9.5. Figure 9.4 shows the hadronization model of the data in the three regions.

We conclude the calibration method by the parameters in Pythia works well.

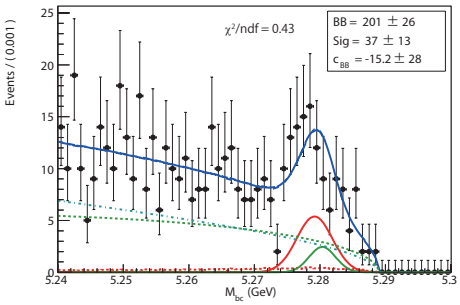
Table 9.3: Fractions(%) on each mode in each M_{X_s} region(140 fb^{-1} data)

1.15< M_{X_s} <1.5 GeV/c^2			1.5< M_{X_s} <2.0 GeV/c^2		
Mode	Data	MC	Mode	Data	MC
1	9.51±1.42	14.5 (+3.6)	1	2.21±0.97	2.91 (+0.7)
2	4.75±0.71	7.50 (+3.9)	2	1.11±0.48	1.49 (+0.7)
3	23.4±2.03	21.6 (-0.9)	3	16.3±2.11	15.0 (-0.6)
4	43.3±3.70	36.5 (-1.8)	4	27.7±3.10	22.0 (-1.8)
5	0.90±0.60	0.95 (-0.1)	5	13.7±2.73	6.58 (-2.6)
6	11.8±6.80	14.9 (-0.5)	6	20.1±4.35	23.7 (-0.8)
7	-1.00±1.00	0.52 (+1.5)	7	15.9±4.60	8.35 (-1.6)
8	5.81±3.32	1.85 (-1.2)	8	-5.14±14.4	8.20 (-0.9)
9	1.53±1.31	0.93 (-0.5)	9	5.84±3.21	5.78 (+0.0)
10	0.00±0.00	0.01 (+0.0)	10	2.33±0.71	1.29 (-1.5)

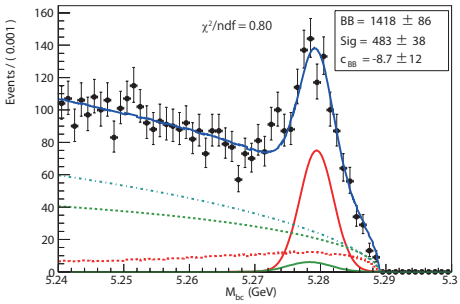
2.0< M_{X_s} <2.4 GeV/c^2		
Mode	Data	MC
1	1.51±1.15	2.91 (-0.3)
2	0.75±0.58	1.49 (-0.3)
3	10.6±3.87	15.0 (-0.3)
4	10.9±6.96	22.0 (+0.4)
5	4.30±6.45	6.58 (+0.6)
6	17.6±14.1	23.7 (+0.4)
7	36.4±18.2	8.35 (-1.1)
8	0.00±46.3	8.20 (+0.2)
9	15.9±9.74	5.78 (-1.0)
10	1.96±2.94	1.29 (-0.2)



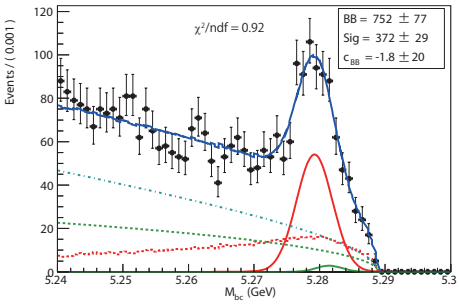
(a) $K\pi$ without π^0



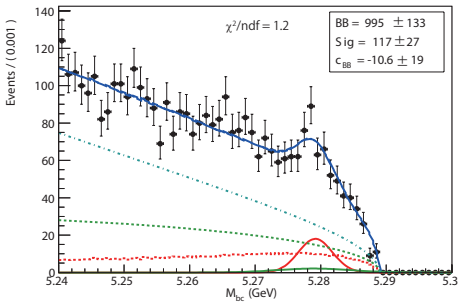
(b) $K\pi$ with π^0



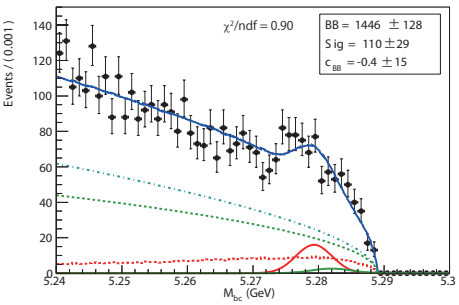
(c) $K2\pi$ without π^0



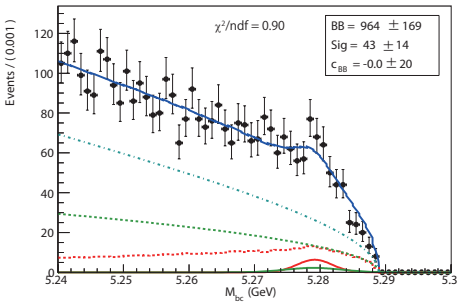
(d) $K2\pi$ with π^0



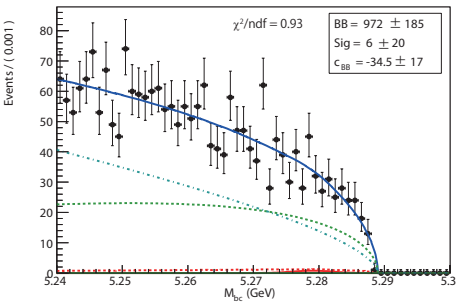
(e) $K3\pi$ without π^0



(f) $K3\pi$ with π^0

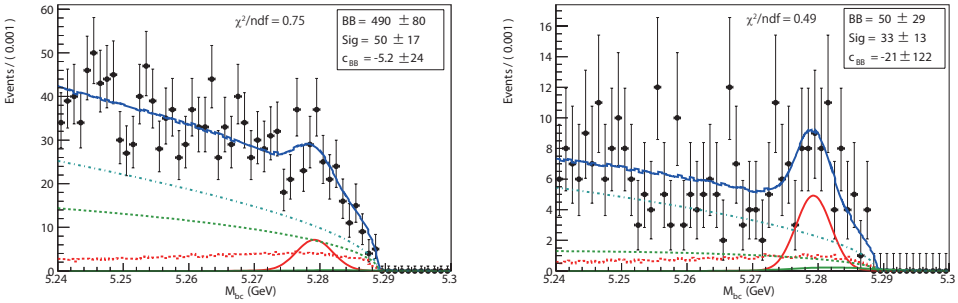


(g) $K4\pi$



(h) $K2\pi^0$ with at most 2π

Figure 9.2: M_{bc} fit for each final state(140 fb^{-1} data) at 1.15< M_{X_s} <2.40 (Red solid line : Signal, Red dashed line: Signal cross-feed, Green solid line: Peaking BG)



(a) $K\eta$ (b) $3K$

Figure 9.3: M_{bc} fit for each final state(140 fb^{-1} data) at 1.15< M_{X_s} <2.40 GeV/c^2 (Red solid line : Signal, Red dashed line: Signal cross-feed, Green solid line: Peaking BG)

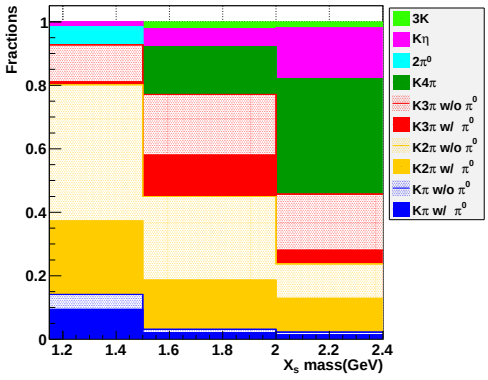


Figure 9.4: Fractions on each mode in each M_{X_s} region(140 fb^{-1} data)

9.4 Branching Fraction

We use two methods to obtain the branching fraction, the total M_{X_s} region fit and M_{X_s} bin fit.

9.4.1 Total M_{X_s} Region Fit

A signal efficiency in 0.0< M_{X_s} <2.8 GeV/c^2 obtained from the calibrated MC is 2.08 %. It should be corrected by the reconstruction and PID efficiencies in Table 8.1, and results 1.95 %. The branching fraction $\mathcal{BR}$ is calculated by using Equation 9.1, where N_{sig} is the number of the signal, 2557±107 in Figure 9.1(a), $N_{B\overline{B}} = 154.3\times10^6$ is the number of $B\overline{B}$ pairs in the 140 fb^{-1} data and ϵ is the signal efficiency. Thus, the branching fraction is

$$\mathcal{BR}(B \to X_s\gamma) = (4.25 \pm 0.18) \times 10^{-4}, \tag{9.4}$$

where the uncertainty is only statistical one. The fraction of M_{X_s} <2.8 GeV/c^2 is 86.9% which is obtained from the MC. Thus, we can obtain

$$\mathcal{BR}(B \to X_s\gamma) = (3.69 \pm 0.16) \times 10^{-4}(M_{X_s} < 2.8\mathrm{GeV/c^2}), \tag{9.5}$$

Usually, we compare the experimental result in photon energy in the B rest frame above 1.6 GeV with the theoretical prediction. Therefore, the result in Equation 9.5 should be extrapolated to E_γ >1.6 GeV. We use a following extrapolation factor on the photon energy spectrum in the standard method[73], in which the factors are calculated by results of the Belle, BaBar, CDF, CLEO and DELPHI in three theoretical models and they are averaged.

$$\begin{aligned} R(E_{1.9\mathrm{GeV}}) &= \frac{\mathcal{BR}(B \to X_s\gamma)_{1.9\mathrm{GeV}}}{\mathcal{BR}(B \to X_s\gamma)_{1.6\mathrm{GeV}}} \\ &= 0.936 \pm 0.010. \end{aligned} \tag{9.6}$$

Thus, at first, the result is interpolated to E_γ >1.9 GeV from M_{X_s} <2.8 GeV/c^2 by a factor, 0.998, obtained from the MC. Next, the branching fraction of the $(B \to X_s\gamma)$ at E_γ >1.6 GeV is calculated as

$$\begin{aligned} \mathcal{BR}(B \to X_s\gamma) &= (3.69 \pm 0.16) \times 10^{-4} \times 0.998/0.936 \\ &= (3.93 \pm 0.17) \times 10^{-4}(E_\gamma > 1.6\mathrm{GeV}). \end{aligned} \tag{9.7}$$

This result is consistent with the world average by the HFAG, (3.55±0.26), within 1.2σ.

9.4.2 M_{X_s} Bin Fit

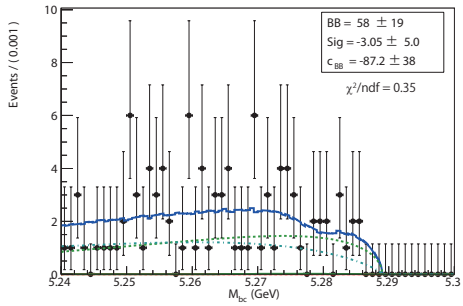
Next, we evaluate branching fractions on each M_{X_s} bin and sum them to obtain the total branching fraction. Figure 9.5, 9.6 and 9.7 show M_{bc} distributions on each M_{X_s} bin. Signal efficiencies on each M_{X_s} bin obtained from calibrated MC are summarized in Table 9.4 and the partial branching fractions are also shown in Table 9.4.The sum of the partial branching fraction is $(3.10\pm0.36)\times10^{-4}$. To calculate branching fraction in E_γ >1.6 GeV, we use the same extrapolation method in Equation 9.7. As a result, the extrapolated one is calculated as

$$\begin{aligned} \mathcal{BR}(B \to X_s\gamma) &= (3.10 \pm 0.36) \times 10^{-4} \times 0.998/0.936 \\ &= (3.31 \pm 0.38) \times 10^{-4} (E_\gamma > 1.6\text{GeV}). \end{aligned} \tag{9.8}$$

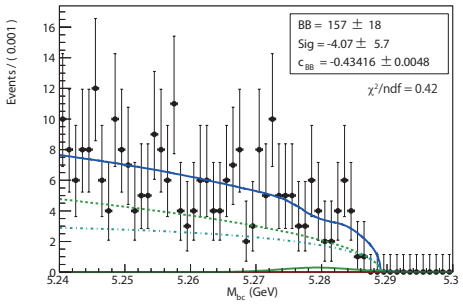
This result is consistent with the world average by the HFAG, (3.55±0.26), within 0.7σ.

9.4.3 Comparison Between Two Methods

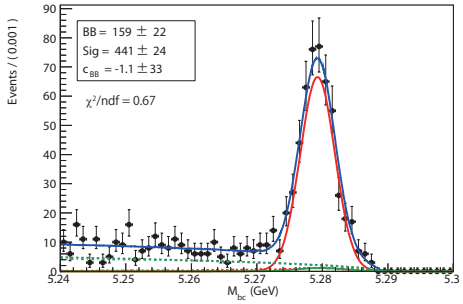
The difference on the results on the branching fraction between the total M_{X_s} region fit in Equation 9.7 and the M_{X_s} bin fit in Equation 9.8 is large and seems to come from the M_{X_s} distribution. The M_{X_s} distributions of the data and the MC are shown in Figure 9.8 and the shape of the MC in 1.2< M_{X_s} <1.5 GeV/c^2 does not match that of the data well. These partial branching fractions in the data are consistent with the BaBar's result[74], thus, the MC's distribution is seems to be wrong. We produced the M_{X_s} distribution in the MC which is consistent with the E_γ distribution at CM frame in the full-inclusive analysis at the Belle in Section 4.2. But, the E_γ distribution has a wide width in 1.2< M_{X_s} <1.5 GeV/c^2 by the B meson momentum in Figure 4.1(b). Therefore, we conclude it is difficult to reproduce M_{X_s} shape in the MC and the M_{X_s} bin fit is adopted.



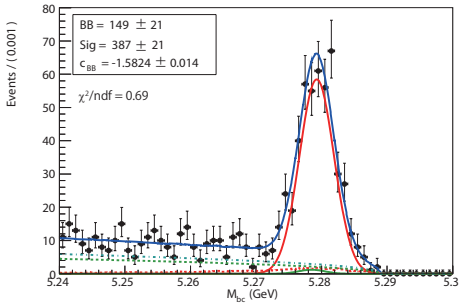
(a) 0.6< M_{X_s} <0.7 GeV



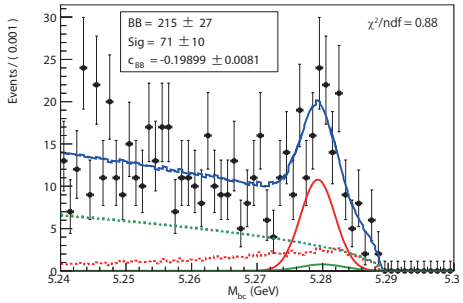
(b) 0.7< M_{X_s} <0.8 GeV



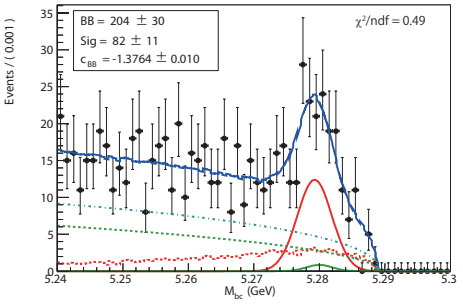
(c) 0.8< M_{X_s} <0.9 GeV



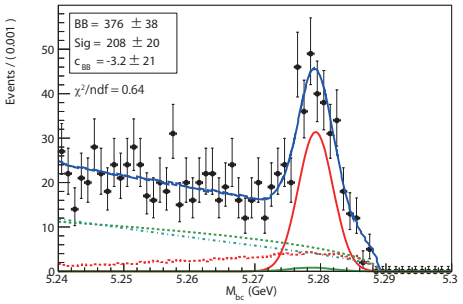
(d) 0.9< M_{X_s} <1.0 GeV



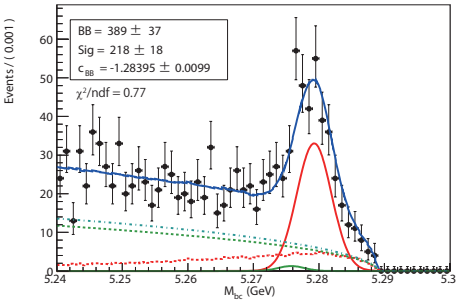
(e) 1.0< M_{X_s} <1.1 GeV



(f) 1.1< M_{X_s} <1.2 GeV

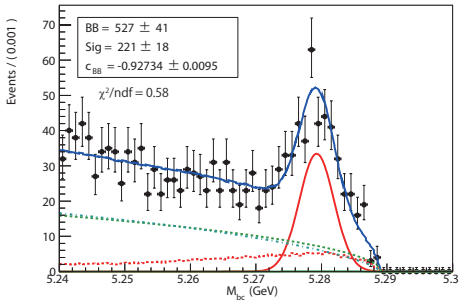


(g) 1.2< M_{X_s} <1.3 GeV

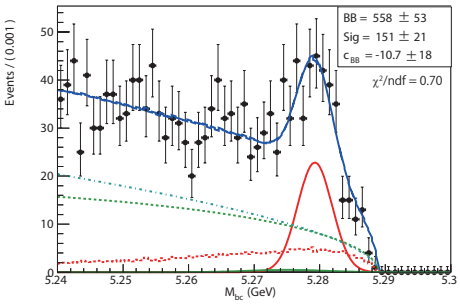


(h) 1.3< M_{X_s} <1.4 GeV

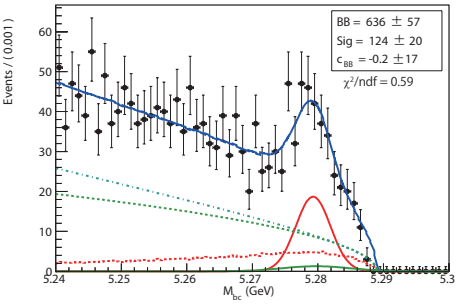
Figure 9.5: M_{bc} fit for each M_{X_s} bin(0.6-1.4 GeV, 140 fb^{-1} data)



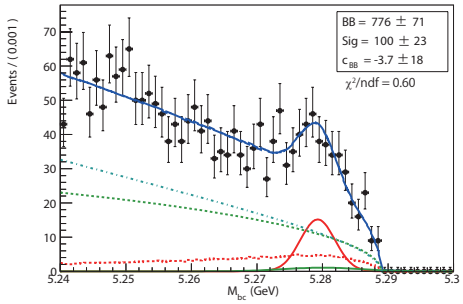
(a) 1.4< M_{X_s} <1.5 GeV



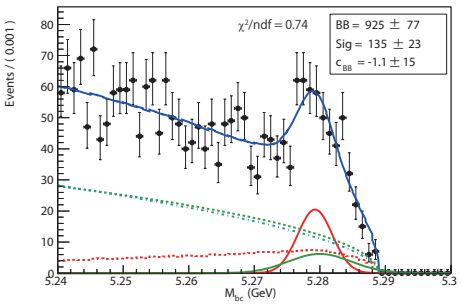
(b) 1.5< M_{X_s} <1.6 GeV



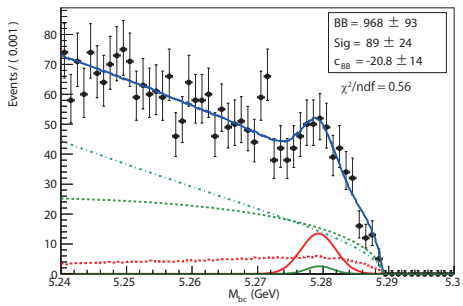
(c) 1.6< M_{X_s} <1.7 GeV



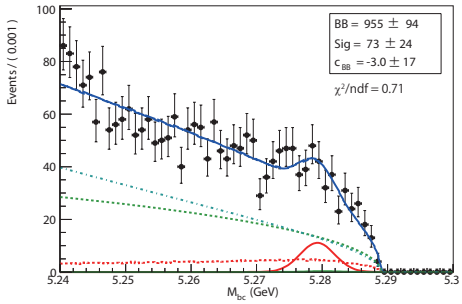
(d) 1.7< M_{X_s} <1.8 GeV



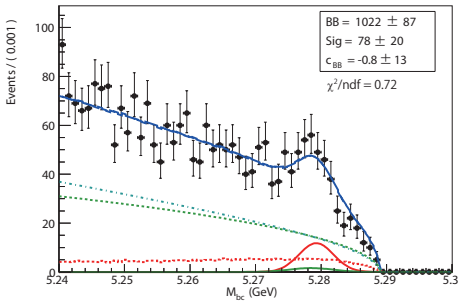
(e) 1.8< M_{X_s} <1.9 GeV



(f) 1.9< M_{X_s} <2.0 GeV

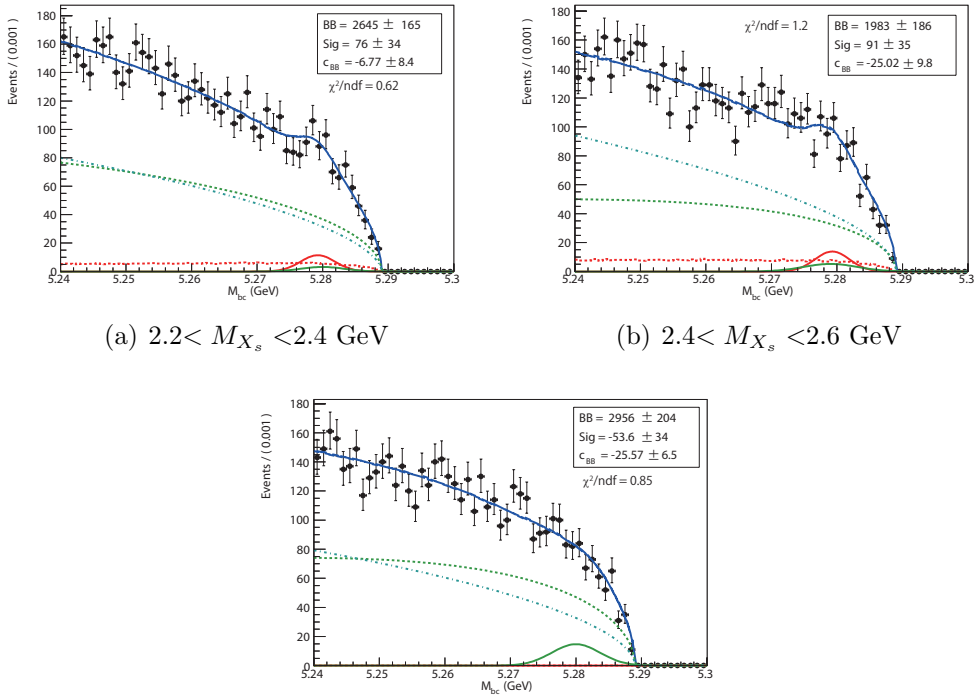


(g) 2.0< M_{X_s} <2.1 GeV



(h) 2.1< M_{X_s} <2.2 GeV

Figure 9.6: M_{bc} fit for each M_{X_s} bin(1.4-2.2 GeV, 140 fb^{-1} data)



(a) 2.2< M_{X_s} <2.4 GeV

(b) 2.4< M_{X_s} <2.6 GeV

(c) 2.6< M_{X_s} <2.8 GeV

Figure 9.7: M_{bc} fit for each M_{X_s} bin(2.2-2.8 GeV, 140 fb^{-1} data)

9.5 Systematic Uncertainties

9.5.1 Hadronization Model Uncertainty

To evaluate the systematic uncertainty, the fractions of the modes are fluctuated by $\pm 1\sigma$. Since it is difficult to realize the fluctuation by parameters in Pythia, we generate reweighted samples on the fractions by removing events of specific modes . The result is summarized in Table 9.5, where the deviation on the efficiency from the default MC is assigned to the systematic uncertainty. Total uncertainty is 12.1 %.

In addition, we evaluate the deviations more than 2σ in $K\pi$ with and without π^0(Mode 1 and 2) in 1.15< M_{X_s} <1.5 GeV/c^2, and $K3\pi$ without π^0(Mode 5) in 1.5< M_{X_s} <2.0 GeV/c^2. To take into account the deviations as the systematic uncertainty, we generate samples reweighted to these fractions in the data and check differences on the signal efficiencies. The results are reported in Table 9.6, which are included in the systematic uncertainty on the hadronization model. The total uncertainty is 2.34 %.

Table 9.4: The partial branching fraction on M_{X_s}(140 fb^{-1} data)

M_{X_s} bin(GeV/c^2)	Yield	Efficiency(%)	$\mathcal{BR}(10^{-6})$
0.6-0.7	-3±5	8.44	-0.1±0.2
0.7-0.8	-4±6	7.50	-0.2±0.3
0.8-0.9	441±24	7.26	19.7±1.1
0.9-1.0	387±21	7.11	17.6±1.0
1.0-1.1	71±10	6.81	3.4±0.5
1.1-1.2	82±11	4.21	6.3±0.8
1.2-1.3	208±20	3.54	19.0±1.8
1.3-1.4	218±18	3.33	21.2±1.8
1.4-1.5	221±18	3.33	21.5±1.8
1.5-1.6	151±21	2.37	20.6±2.9
1.6-1.7	124±20	2.09	19.2±3.1
1.7-1.8	100±23	1.74	18.6±4.3
1.8-1.9	135±23	1.58	27.6±4.7
1.9-2.0	89±24	1.34	21.6±5.8
2.0-2.1	73±24	1.12	21.1±6.9
2.1-2.2	78±20	0.88	28.7±7.4
2.2-2.4	76±34	0.72	34.1±15
2.4-2.6	91±35	0.63	46.4±18
2.6-2.8	-54±34	0.49	-35.8±23
Total			310±36

9.5.2 Missing Mode Uncertainty

We have to consider an uncertainty from missing modes which are not reconstructed in this analysis. The fraction of the reconstructed modes in the MC after the calibration is 84.4 % (1.15< M_{X_s} <2.4 GeV/c^2), where fractions of the K_L modes which corresponds to the reconstructed K_s modes are included in this value. In order to evaluate the uncertainty, the parameters in Pythia are shifted within parameter region which is consistent with the model of the data. 7 examples of the parameter settings are shown in Table 9.7. The maximum and minimum of the fraction of the reconstructed mode are 86.8 and 81.1, respectively. Therefore, the differences from that of the calibrated MC are (86.8-84.4)/84.4 = +2.84 % and (81.1-84.4)/84.4 = -3.91 %, respectively. The uncertainty is needed in M_{X_s} >1.15 GeV/c^2, thus, the fraction is considered. As a result, we assign ±3.36 % as the systematic uncertainty on the missing modes. In M_{X_s} bin analysis, the uncertainties of the missing mode on each M_{X_s} bin are needed. Therefore, the fraction of the reconstructed mode in every mass bin are investigated and the maximum and minimum values are

Table 9.5: Systematic uncertainty(%) on hadronization model(140 fb^{-1} data).

M_{X_s} bin	Fluctuated mode										
(GeV/c^2)	1	2	3	4	5	6	7	8	9	10	Total
1.1-1.2	6.26	2.88	2.23	2.73	3.64	4.01	1.40	3.09	2.57	1.41	10.5
1.2-1.3	2.29	1.18	0.73	1.28	2.44	1.47	3.20	1.89	1.90	2.10	6.23
1.3-1.4	1.68	1.80	2.74	1.51	0.62	2.74	0.65	3.07	1.23	1.12	6.54
1.4-1.5	4.59	2.87	2.69	2.35	1.36	4.66	1.93	3.74	4.40	3.76	10.7
1.5-1.6	1.13	0.91	2.49	2.54	0.48	3.43	2.07	1.82	3.05	4.36	8.03
1.6-1.7	2.56	3.49	3.74	3.93	3.38	1.88	1.09	3.37	4.78	3.28	10.5
1.7-1.8	3.00	0.27	3.30	3.38	3.17	2.45	1.10	2.77	3.22	4.58	9.16
1.8-1.9	1.27	2.16	2.22	2.22	2.97	3.64	3.42	5.70	3.85	2.00	10.4
1.9-2.0	4.50	3.50	6.90	6.90	4.09	4.71	2.18	6.22	6.22	5.96	16.2
2.0-2.1	1.90	2.00	2.37	1.82	2.67	3.73	2.87	4.81	3.66	0.88	9.14
2.1-2.2	2.35	1.65	2.66	3.18	1.16	3.06	2.51	8.56	1.86	2.13	12.1
2.2-2.4	7.27	6.71	9.09	6.90	9.37	11.6	6.53	7.69	9.22	6.27	26.5
2.4-2.6	8.56	9.49	5.64	8.63	7.98	7.79	8.86	9.36	9.21	8.12	26.7
2.6-2.8	11.4	12.7	9.41	15.6	13.1	11.1	10.4	11.2	12.8	11.4	27.6
Total	4.27	3.86	4.53	4.87	5.24	3.75	3.80	5.49	4.41	3.64	12.1

Table 9.6: Systematic uncertainty(%) on hadronization model 2(140 fb^{-1} data).

M_{X_s} bin	Reweighted mode			
(GeV/c^2)	1	2	5	Total
1.1-1.2	9.44	9.86		13.7
1.2-1.3	2.08	0.53		2.15
1.3-1.4	4.32	4.89		6.53
1.4-1.5	6.38	4.71		7.93
1.5-1.6			0.82	0.82
1.6-1.7			0.03	0.03
1.7-1.8			7.20	7.20
1.8-1.9			8.81	8.81
1.9-2.0			9.08	9.08
Total	1.07	0.91	1.86	2.34

reported in Table 9.8, which are assigned to the systematic uncertainty for each M_{X_s} bin.

9.5.3 M_{bc} PDF

◯ **Signal PDF** The signal shape parameters in the signal PDF are fixed by values obtained from a $B \to D\pi$ data sample. To evaluate the uncertainty,

Table 9.7: Fractions(%) of reconstructed and missing modes. In calculation of the fraction of Reco mode, the K_L modes are taken into account. PARJ(12) (D=0.6) is the probability that a strange meson has spin 1.

Mode	calibrated MC	1	2	3	4	5	6	7
	PARJ(2)=0.10	0.30	0.10	0.10	0.10	0.10	0.10	0.15
	PARJ(11)=0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
	PARJ(15)=0.25	0.05	0.05	0.30	0.30	0.30	0.40	0.40
	PARJ(25)=0.03	1.00	1.00	1.00	0.03	0.03	0.03	0.03
	PARJ(12)=0.60	0.60	0.60	0.60	0.60	0.70	0.70	0.70
1	4.76	4.20	4.24	4.88	5.05	4.19	6.43	6.81
2	2.44	2.18	2.15	2.49	2.60	2.13	3.33	3.48
3	14.7	13.0	13.4	14.7	14.9	11.2	14.6	14.8
4	22.4	22.3	23.1	21.6	19.7	15.4	25.3	23.2
5	5.98	6.01	6.30	5.82	6.68	8.64	3.68	4.50
6	21.5	24.1	24.6	20.6	20.3	25.4	20.5	19.6
7	9.36	9.54	10.1	9.21	9.43	11.8	7.96	8.73
8	12.2	7.68	8.33	7.48	7.89	10.7	5.58	6.52
9	4.90	3.04	1.97	7.00	6.77	2.06	4.62	5.86
10	1.76	2.27	0.90	1.05	1.06	1.80	2.67	1.50
Reco mode	84.4	84.7	86.8	82.8	81.2	86.4	81.1	82.1
Missing mode	15.6	15.3	13.2	17.2	18.8	13.6	18.9	17.9

Table 9.8: The fractions of the reconstructed mode in every M_{X_s} mass bin(140 fb^{-1} data).

M_{X_s} bin(GeV/c^2)	Default	Maximum	Minimum	Systematics(%)
1.1-1.2	0.996	0.996	0.926	+0.00 -7.08
1.2-1.3	0.992	0.992	0.945	+0.00 -4.71
1.3-1.4	0.980	0.980	0.956	+0.00 -2.45
1.4-1.5	0.968	0.975	0.956	+0.68 -1.29
1.5-1.6	0.923	0.952	0.898	+3.10 -2.74
1.6-1.7	0.918	0.949	0.890	+3.39 -3.06
1.7-1.8	0.914	0.946	0.888	+3.50 -2.88
1.8-1.9	0.891	0.924	0.856	+3.61 -3.96
1.9-2.0	0.850	0.890	0.810	+4.67 -4.75
2.0-2.1	0.798	0.831	0.760	+4.10 -4.75
2.1-2.2	0.742	0.775	0.695	+4.45 -6.26
2.2-2.4	0.661	0.686	0.618	+3.75 -6.52
2.4-2.6	0.548	0.557	0.509	+1.68 -7.06
2.6-2.8	0.456	0.456	0.424	+0.00 -7.08

the shape parameters of the PDF are fluctuated according to the Gaussian distribution whose width is the statistical error of the $D\pi$ sample and 500 times of the fits are performed to extract signal yield from data. The width of the yield distribution is taken as the systematic uncertainty and the result is summarized in Table 9.9.

◯ **Signal cross-feed PDF** The histogram PDF of the signal cross-feed is obtained from the signal MC sample. To evaluate the uncertainty, the entries in the bins are fluctuated according to the Gaussian distribution whose widths are the statistical errors of each bin in data. The ratio of the cross-feed to the signal is fixed in the fit and is also fluctuated in a statistical uncertainty of the cross-feed in data. 500 times such fits are repeated to extract the signal yield from data. The width of the yield distribution is taken as the systematic uncertainty. The result is reported in Table 9.9.

◯ **Peaking background PDF** The shape parameters and the yield of the peaking background, which are obtained from anti-π^0/η veto data sample, are fixed in the M_{bc} distribution fit. To evaluate the uncertainty, they are fluctuated according to the Gaussian distribution whose width is the statistical error of the anti-π^0/η veto data sample. 500 times such fits are repeated to extract the signal yield from data. The width of the yield distribution is taken as the systematic uncertainty. The result is summarized in Table 9.9.

◯ **PDF for non-peaking part from $q\overline{q}$ background** The shape parameters and the yield of the peaking part from $q\overline{q}$ background, which are obtained from the off-resonance data sample which is collected at energy below 60 MeV from $\Upsilon(4S)$ resonance, are fixed in the M_{bc} fit. To evaluate the uncertainty, they are fluctuated according to the Gaussian distribution whose width is the statistical error. 500 times fits are repeated to extract the signal yield from data. The width of the yield distribution is taken as the systematic uncertainty. The result is reported in Table 9.9.

9.5.4 $K^* - X_s$ Transition Position

In the signal MC, the exclusive $K^*\gamma$ and the inclusive $X_s\gamma$ MC are included, and the transition position is fixed at 1.15 GeV/c^2. In M_{X_s} distribution in data(Figure 9.8), we can see a rising from the non-resonance modes in 1.2 GeV/c^2 at least. Thus, the position is fluctuated by $\pm$50 MeV to evaluate the

Table 9.9: Systematic uncertainties(%) on M_{bc} PDF(140 fb^{-1} data).

M_{X_s} bin (GeV/c^2)	Signal	Cross-feed	Peaking BG	$q\overline{q}$ BG	Total
0.6-0.7	0.00	0.00	0.00	0.00	0.00
0.7-0.8	0.00	0.00	0.00	0.00	0.00
0.8-0.9	0.09	0.86	0.86	0.05	1.22
0.9-1.0	0.31	1.13	0.64	0.05	1.34
1.0-1.1	0.51	6.96	2.04	0.00	7.27
1.1-1.2	0.73	8.10	2.35	0.37	8.47
1.2-1.3	0.17	3.52	1.20	0.05	3.73
1.3-1.4	0.15	3.72	1.04	0.05	3.87
1.4-1.5	0.16	2.74	0.14	0.27	2.76
1.5-1.6	0.09	5.70	0.13	0.13	5.70
1.6-1.7	0.48	5.78	3.01	0.48	6.55
1.7-1.8	0.10	10.2	2.94	0.30	10.6
1.8-1.9	0.57	6.43	6.01	0.33	9.47
1.9-2.0	0.11	11.1	6.11	0.22	12.7
2.0-2.1	0.14	15.5	1.89	0.82	15.6
2.1-2.2	0.31	14.4	5.72	0.38	15.5
2.2-2.4	0.43	18.7	10.1	0.26	21.4
2.4-2.6	0.14	22.4	12.5	0.33	25.7
2.6-2.8	0.00	0.00	0.00	0.00	0.00

uncertainty to the efficiency in 1.1< M_{X_s} <1.2 GeV/c^2. The signal efficiencies on the transition position at 1.1 and 1.2 GeV are 4.50 and 6.29 %, respectively(defalut:4.79%). The deviations are -6.10 and +31.4 %, and are included in the systematic uncertainty on the hadronization model in 1.1< M_{X_s} <1.2 GeV/c^2.

9.5.5 Extrapolation Factor to E_γ >1.6 GeV

The systematic uncertainty in the extrapolation factor to E_γ(B rest frame)>1.6 GeV is evaluated.

At first, the M_{X_s} shape is fluctuated according to the method in Section 8.5.1 to estimate the interpolation uncertainty to E_γ >1.9 GeV from M_{X_s} <2.8 GeV/c^2. We obtain the factors of 0.998, 0.996, 0.998, and 0.998 for KN1, 2, 3, 4 parameter settings, respectively. We assign ±0.19 % to the systematic uncertainty.

As for the extrapolation factor to E_γ >1.6 GeV from E_γ >1.9 GeV, we use the uncertainty in this reference[73], 0.936±0.010.

9.5.6 Summary of Systematic Uncertainty

Systematic uncertainties in each M_{X_s} mass bin are summarized in Table 9.11. Total systematic uncertainty is reported in Table 9.10.

Table 9.10: Systematic uncertainty(%) (140 fb^{-1} data)

Source	Systematic uncertainty (%)
$B\overline{B}$ counting	1.37
Detector response	2.98
Background rejection	3.38
M_{bc} PDF	4.84
Hadronization model	12.3
Missing mode	3.36
Total	14.4

Table 9.11: Systematic uncertainties(%) in every M_{X_s} mass bin(140 fb^{-1} data).

M_{X_s} bin (GeV/c^2)	# of $B\overline{B}$	Detector response	Background rejection	M_{bc} PDF	Hadronization	Missing mode	Total
0.6-0.7	1.37	2.68	3.38	0.00	-	-	4.52
0.7-0.8	1.37	2.59	3.38	0.00	-	-	4.47
0.8-0.9	1.37	2.57	3.38	1.22	-	-	4.99
0.9-1.0	1.37	2.58	3.38	1.34	-	-	5.16
1.0-1.1	1.37	2.57	3.38	7.27	-	-	15.7
1.1-1.2	1.37	2.99	3.38	8.47	35.9	7.07	38.8
1.2-1.3	1.37	3.15	3.38	3.73	6.59	4.71	10.5
1.3-1.4	1.37	3.17	3.38	3.87	9.24	2.45	11.5
1.4-1.5	1.37	3.10	3.38	2.76	13.3	1.29	14.7
1.5-1.6	1.37	3.34	3.38	5.70	8.07	3.10	11.3
1.6-1.7	1.37	3.49	3.38	6.55	10.5	3.39	13.7
1.7-1.8	1.37	3.59	3.38	10.6	11.7	3.50	15.5
1.8-1.9	1.37	3.69	3.38	9.47	13.6	3.96	16.2
1.9-2.0	1.37	3.74	3.38	12.7	18.6	4.75	22.4
2.0-2.1	1.37	3.78	3.38	15.6	9.14	4.75	14.5
2.1-2.2	1.37	3.78	3.38	15.5	12.1	6.26	16.7
2.2-2.4	1.37	3.77	3.38	21.2	26.5	6.52	30.6
2.4-2.6	1.37	3.79	3.38	25.7	26.7	7.06	38.1
2.6-2.8	1.37	3.82	3.38	0.00	37.6	7.08	38.6

9.6 Results of 140 fb^{-1} Data

By using the 140 fb^{-1} data, we evaluate the precision on M_{bc} fit in total M_{bc} region and each M_{bc} bin, and the M_{bc} bin fit is adopted. The calibration on the hadronization model by Pythia is performed and we found it worked well.

The partial branching fractions on the M_{X_s} is summarized in Table 9.12. The total branching fraction in M_{X_s} <2.8 GeV/c^2 with 140 fb^{-1} data is

$$\mathcal{BR}(B \to X_s\gamma) = (3.10 \pm 0.36 \pm 0.45) \times 10^{-4} (M_{X_s} < 2.8\mathrm{GeV/c^2}). \quad (9.9)$$

The extrapolated one to E_γ(B rest frame)>1.6 GeV is

$$\mathcal{BR}(B \to X_s\gamma) = (3.31 \pm 0.38 \pm 0.48) \times 10^{-4} (E_\gamma > 1.6\mathrm{GeV}), \quad (9.10)$$

where the extrapolation factor of the systematic uncertainty is added.

Table 9.12: The partial branching fraction on M_{X_s} with 140 fb^{-1} data

M_{X_s} bin(GeV/c^2)	$\mathcal{BR}(10^{-6})$
0.6-0.7	-0.1±0.2±0.0
0.7-0.8	-0.2±0.3±0.0
0.8-0.9	19.7±1.1±1.0
0.9-1.0	17.6±1.0±0.9
1.0-1.1	3.4±0.5±0.5
1.1-1.2	6.3±0.8±2.3
1.2-1.3	19.0±1.8±2.0
1.3-1.4	21.2±1.8±2.0
1.4-1.5	21.5±1.8±2.7
1.5-1.6	20.6±2.9±2.3
1.6-1.7	19.2±3.1±2.6
1.7-1.8	18.6±4.3±2.5
1.8-1.9	27.6±4.7±4.0
1.9-2.0	21.6±5.8±4.4
2.0-2.1	21.1±6.9±3.0
2.1-2.2	28.7±7.4±4.8
2.2-2.4	34.1±15±10
2.4-2.6	46.4±18±18
2.6-2.8	-35.8±23±14

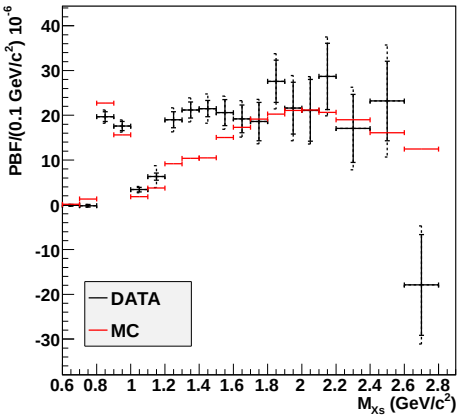


Figure 9.8: Partial branching fraction(140 fb^{-1} data). The first solid error is the statistical one and the second dashed error is a quadratic sum of the statistical and systematic errors.

Chapter 10

Results

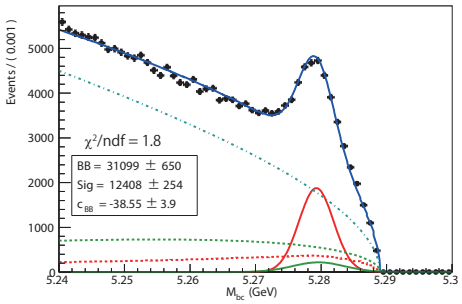
In this chapter, we obtain the branching fraction of $B \to X_s\gamma$ with the Belle's full data. Before evaluating the total branching fraction, we discuss the calibration method. Next, the branching fractions in M_{X_s} bins are obtained from the M_{bc} fits and they are summed to get the total branching fraction. Then, the systematic uncertainty is evaluated and the final result is obtained. At last, the total branching fraction is compared with the SM prediction and provides a constraint to the two Higgs Doublet Model.

10.1 Branching Fraction of $B \to K^*\gamma$

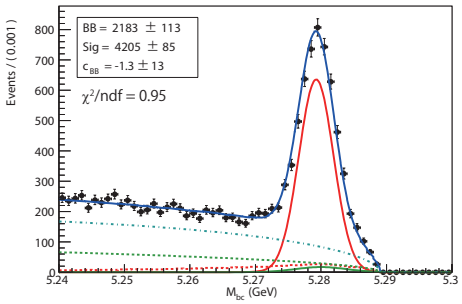
Figure 10.1 shows M_{bc} distribution in 0.0< M_{X_s} <2.8 GeV/c^2 and the signal yield obtained from the fit is 12408±254. In Figure 10.1(b), M_{bc} fit in M_{X_s} <1.15 GeV/c^2 which corresponds to exclusive $K^*\gamma$ region is shown. The signal yield is (4205±85) and the signal efficiency obtained from the MC is 6.82 % which is corrected by the reconstruction and PID efficiencies. The $\mathcal{BR}(B \to K^*\gamma)$ is calculated by the Equation 9.1. N_{sig} is the number of the signal, 4205±85, $N_{B\overline{B}} = 771.9\times10^6$ is the number of $B\overline{B}$ pairs in the full data and ϵ is the signal efficiency, 6.82 %. The branching fraction is

$$\mathcal{BR}(K^*\gamma) = (3.99 \pm 0.08) \times 10^{-5}, \tag{10.1}$$

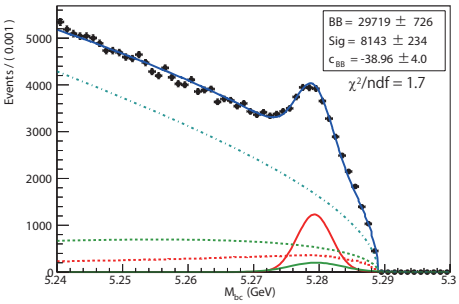
where the uncertainty includes only statistical one. This result is consistent with the world average, (4.26±0.17)×10^{-5}, within 1.4σ. We can obtain a validity on our analysis.



(a) 0< M_{X_s} <2.8 GeV/c^2



(b) M_{X_s} <1.15 GeV/c^2



(c) 1.15< M_{X_s} <2.8 GeV/c^2

Figure 10.1: M_{bc} fit with the full data(Red solid line: Signal, Red dashed line: Signal cross-feed, Green solid line: Peaking BG, , Green dashed line: Non-peaking background from $B\overline{B}$, Cyan dashed line:Non-peaking background from $q\overline{q}$)

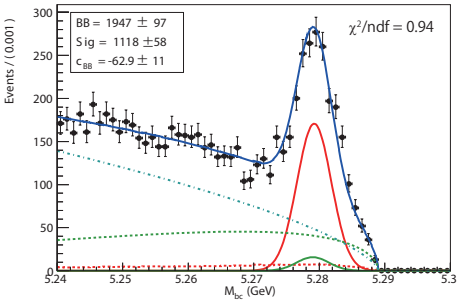
10.2 Calibration on Hadronization model

10.2.1 Hadronization Model in Data

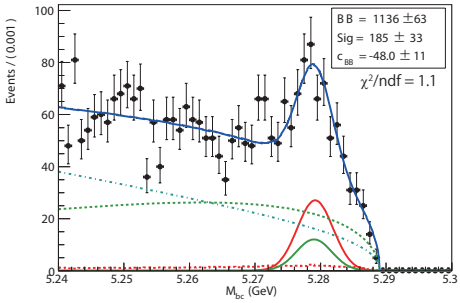
The hadronization model in 1.15< M_{X_s} <2.8 GeV/c^2 is investigated with the same method in the partial data analysis. In Figure 10.2 and 10.3, M_{bc} distributions of each decay mode are fitted to obtain the yield. The fit result, branching fractions and fractions of every decay modes are summarized in Table 10.1. The fractions of the default MC are compared with them of the data in total M_{X_s} region in Table 10.2. Fractions of $K\pi$ and $K2\pi$ have large deviations from them of the data, especially. To calibrate these fractions, we use parameters in Pythia in next section.

10.2.2 Calibration by Pythia

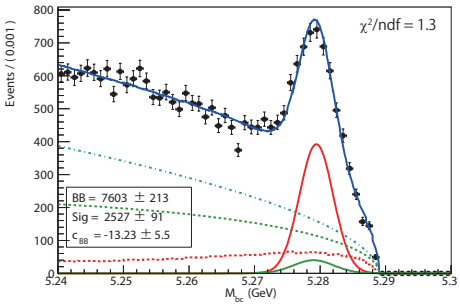
The hadronization model in the MC is calibrated by parameters in Pythia in the same method as 140 fb^{-1} data study. The result of the calibration is



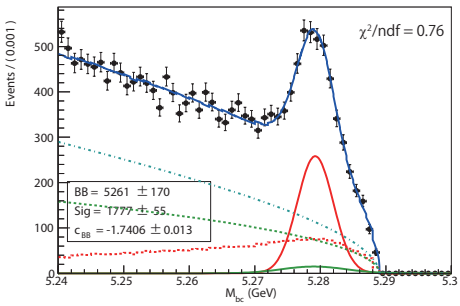
(a) $K\pi$ without π^0



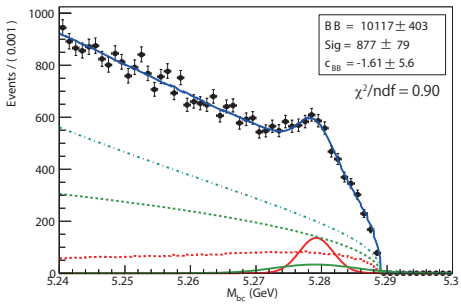
(b) $K\pi$ with π^0



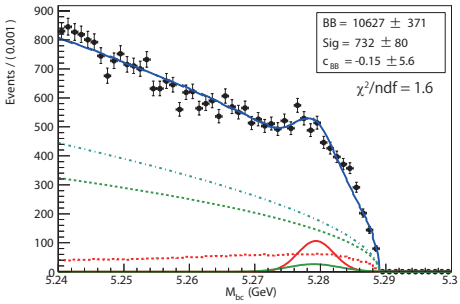
(c) $K2\pi$ without π^0



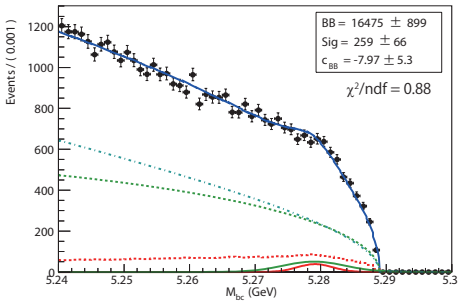
(d) $K2\pi$ with π^0



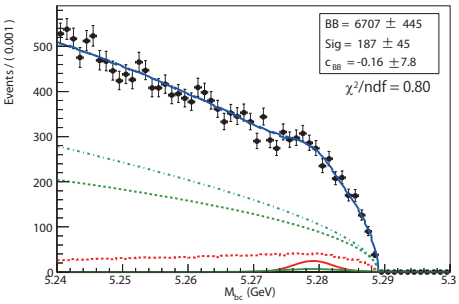
(e) $K3\pi$ without π^0



(f) $K3\pi$ with π^0

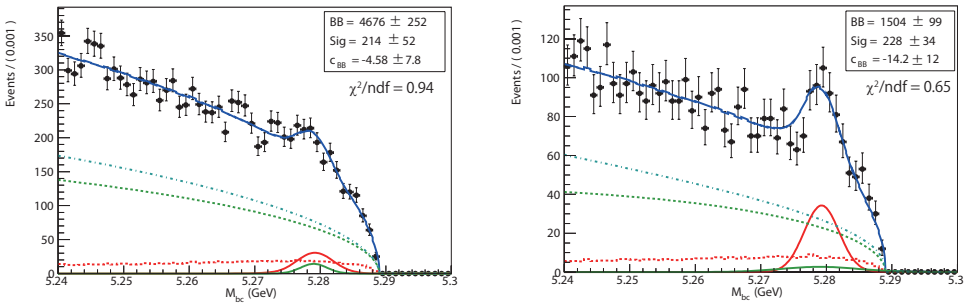


(g) $K4\pi$



(h) $K2\pi^0$ with at most 2π

Figure 10.2: M_{bc} fit for each final state(Full data) at 1.15< M_{X_s} <2.80 (Red solid line : Signal, Red dashed line: Signal cross-feed, Green solid line: Peaking BG)



(a) $K\eta$ (b) $3K$

Figure 10.3: M_{bc} fit for each final state(Full data) at 1.15< M_{X_s} <2.80 GeV/c^2 (Red solid line : Signal, Red dashed line: Signal cross-feed, Green solid line: Peaking BG)

Table 10.1: Fit result for each final states(Full data, 1.15< M_{X_s} <2.8 GeV/c^2)

Mode	Definition	Signal yield	BR(10^{-6})	Fraction(%)
1	$K\pi$ without π^0	1118±58	8.19±0.42	4.66±0.24
2	$K\pi$ with π^0	185±33	2.89±0.51	1.64±0.29
3	$K2\pi$ without π^0	2527±91	25.5±0.92	14.5±0.52
4	$K2\pi$ with π^0	1777±55	42.3±1.31	24.0±0.74
5	$K3\pi$ without π^0	877±79	14.7±1.32	8.34±0.75
6	$K3\pi$ with π^0	732±80	28.3±3.09	16.1±1.76
7	$K4\pi$	259±66	11.1±2.80	11.1±2.80
8	$2\pi^0$	187±45	25.4±6.10	14.2±3.47
9	$K\eta$	214±52	5.60±1.36	3.18±0.77
10	$3K$	228±34	3.52±0.52	2.00±0.30

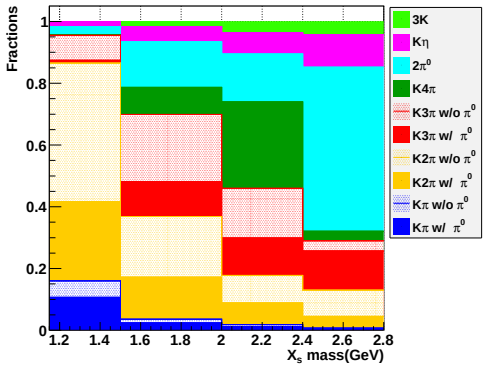


Figure 10.4: Fractions on each mode in each M_{X_s} region(Full data)

Table 10.2: Fractions(%) on each mode in the full data and the MC. Number in () is a deviation to that of the data, defined as ([Fraction in MC]-[Fraction in data])/σ_{data}.

Mode	Partial Data ($1.15< M_{X_s} <2.4$)	Full Data ($1.15< M_{X_s} <2.4$)	Full Data ($1.15< M_{X_s} <2.8$)	Default MC	Calibrated MC
1	5.06±0.89	4.72±0.26	4.20±0.25	10.3 (+17)	4.61 (+1.2)
2	2.24±0.44	2.36±0.13	2.10±0.13	5.42 (+19)	2.38 (+1.6)
3	17.4±1.37	16.0±0.53	14.5±0.52	12.9 (-3.1)	15.7 (+2.4)
4	31.6±2.47	25.9±1.03	24.0±0.74	15.2 (-12)	24.0 (-0.0)
5	7.00±1.62	8.02±0.67	8.34±0.75	5.90 (-3.3)	4.58 (-5.0)
6	15.2±4.01	15.8±1.21	16.1±1.76	15.7 (-0.2)	19.2 (+1.8)
7	11.6±3.80	7.90±1.90	11.1±2.80	12.3 (+0.4)	10.2 (-0.3)
8	2.91±9.70	13.9±3.37	14.4±3.47	14.4 (-0.0)	11.6 (-0.8)
9	4.68±1.59	3.63±0.65	3.18±0.77	4.92 (+2.3)	5.35 (+2.8)
10	1.93±0.61	1.86±0.26	2.00±0.30	2.98 (-3.3)	2.31 (-1.0)

Table 10.3: Fractions(%) on each mode in each M_{X_s} region(Full data)

$1.15< M_{X_s} <1.5$ GeV/c^2			$1.5< M_{X_s} <2.0$ GeV/c^2		
Mode	Data	MC	Mode	Data	MC
1	9.51±1.42	14.5 (+6.4)	1	2.39±0.35	2.91 (+1.5)
2	5.32±0.31	7.50 (+7.1)	2	1.19±0.18	1.49 (+1.7)
3	25.7±0.82	21.6 (-5.0)	3	13.6±0.76	15.0 (+1.9)
4	44.8±1.51	36.5 (-5.5)	4	19.7±1.06	22.0 (+2.2)
5	0.91±0.52	0.95 (+0.1)	5	11.3±0.94	6.58 (-5.0)
6	8.06±2.17	14.9 (+3.1)	6	21.7±2.39	23.7 (+0.8)
7	0.30±0.50	0.52 (+0.5)	7	8.80±2.70	12.2 (-1.2)
8	2.52±2.52	2.51 (+0.0)	8	14.7±2.08	8.20 (-3.1)
9	1.71±0.43	0.93 (-1.8)	9	5.00±1.27	5.78 (+0.6)
10	0.00±0.00	0.01 (+0.0)	10	1.64±0.24	1.29 (-1.5)
$2.0< M_{X_s} <2.4$ GeV/c^2			$2.4< M_{X_s} <2.8$ GeV/c^2		
Mode	Data	MC	Mode	Data	MC
1	1.21±0.64	1.15 (-0.1)	1	0.46±0.65	0.90 (+0.7)
2	0.60±0.32	0.60 (+0.0)	2	0.23±0.32	0.49 (+0.8)
3	7.06±1.37	9.64 (+1.9)	3	3.84±2.15	8.20 (+2.0)
4	8.93±2.63	13.9 (+1.9)	4	8.49±4.03	11.8 (+0.8)
5	12.1±2.53	8.33 (-1.5)	5	12.7±5.20	8.18 (-0.9)
6	16.1±5.65	22.6 (+1.1)	6	3.27±12.8	21.2 (+1.4)
7	28.0±9.10	16.5 (-1.3)	7	3.10±26.7	20.4 (-0.7)
8	15.5±15.5	18.5 (+0.4)	8	53.1±28.7	20.2 (-1.2)
9	6.82±3.69	6.16 (-0.2)	9	10.6±8.19	5.89 (-0.6)
10	3.61±1.10	1.42 (-2.0)	10	4.13±2.84	1.04 (-1.1)

Table 10.4: Scale factors in the direct calibration

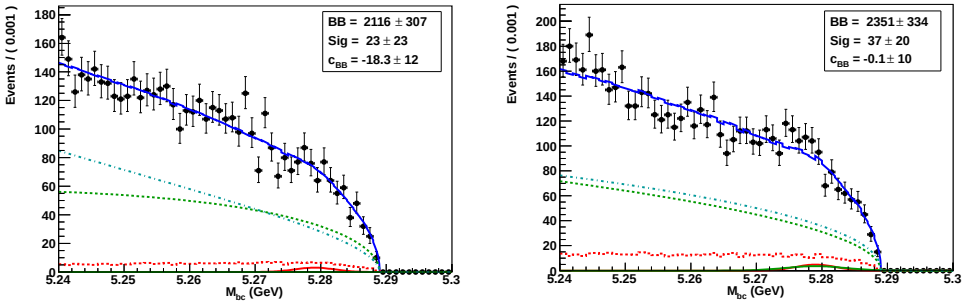
Mode	1.15< M_{X_s} <1.5	1.5< M_{X_s} <2.0	2.0< M_{X_s} <2.4	2.4< M_{X_s} <2.8
1	0.66±0.10	0.82±0.12	1.05±0.56	0.51±0.72
2	0.71±0.04	0.80±0.12	1.00±0.53	0.47±0.65
3	1.19±0.04	0.91±0.05	0.73±0.14	0.47±0.26
4	1.23±0.04	0.90±0.05	0.64±0.19	0.72±0.34
5	0.96±0.55	1.72±0.14	1.45±0.30	1.55±0.64
6	0.54±0.15	0.92±0.10	0.71±0.25	0.15±0.60
7	0.58±0.96	0.72±0.22	1.70±0.55	0.15±1.30
8	1.00±1.00	1.79±0.25	0.84±0.84	2.63±14.2
9	1.84±0.46	0.87±0.22	1.11±0.60	1.80±1.39
10	0.00±0.00	1.27±0.19	2.54±0.77	3.97±2.73

reported in last column in Table 10.2. Total χ^2 is improved from 831 to 52 by this calibration, but it is still large and some fractions have large deviations, particularly. In Table 10.3, the hadronization model in four M_{X_s} regions, 1.15< M_{X_s} <1.5, 1.5< M_{X_s} <2.0, 2.0< M_{X_s} <2.4 and 2.4< M_{X_s} <2.8 GeV/c^2, are compared with that of the data. Similarly, many fractions have large deviations. We conclude the fine-tuning by parameters in Pythia is difficult and attempt to a second calibration in next section.

10.2.3 Direct Calibration

The fine-tuning on the hadronization model by parameters in Pythia is difficult as mentioned in last section. Consequently, we attempt a direct calibration as a second one, in which fractions are directly reweighted to them of the data. The fractions in the MC are corrected by scale factors, defined as (fraction in data)/(fraction in MC). The scale factors obtained from Table 10.3 are summarized in Table 10.4. We note that $K2\pi^0$ mode(Mode=8) fractions in M_{X_s} >2.0 GeV/c^2 have very large uncertainties. In Figure 10.5(a) and 10.5(b), M_{bc} distributions of $K2\pi^0$ modes in 2.0< M_{X_s} <2.4 and 2.4< M_{X_s} <2.8 GeV/c^2 are shown. We can not observe signal peaks around 5.28 GeV/c^2 clearly. Therefore, these fractions should not be used in the direct calibration. For the $K2\pi^0$ mode in M_{X_s} >2.0 GeV/c^2, fractions in the MC are used in Table 10.3. +100%-50% fluctuations are given for these fractions in the systematic uncertainty, where the fluctuation in the low side is -50% since -100% means that $K2\pi^0$ modes have a null fraction and the situation is highly improbable.

The signal efficiencies in each M_{X_s} bin before and after the calibration are reported in Table 10.5.



(a) 2.0< M_{X_s} <2.4 GeV/c^2 (b) 2.4< M_{X_s} <2.8 GeV/c^2

Figure 10.5: M_{bc} fit of $K2\pi^0$ modes in M_{X_s} >2.0 GeV/c^2

Table 10.5: The efficiencies before and after the fragmentation calibration. These fractions are corrected by the reconstruction and $K^{\pm}/\pi^{\pm}$ identification efficiencies in Table 8.1.

M_{X_s} bin (GeV/c^2)	Efficiency(%) before calibration	Efficiency(%) after calibration
0.6-0.7	7.0	7.0
0.7-0.8	7.2	7.2
0.8-0.9	6.7	6.7
0.9-1.0	7.0	7.0
1.0-1.1	6.7	6.7
1.1-1.2	4.3	4.2
1.2-1.3	4.0	3.5
1.3-1.4	3.7	3.3
1.4-1.5	3.6	3.3
1.5-1.6	2.7	2.4
1.6-1.7	2.3	2.1
1.7-1.8	2.0	1.7
1.8-1.9	1.7	1.6
1.9-2.0	1.4	1.3
2.0-2.1	1.2	1.1
2.1-2.2	0.9	0.9
2.2-2.4	0.8	0.7
2.4-2.6	0.5	0.6
2.6-2.8	0.4	0.5

10.3 Branching Fraction

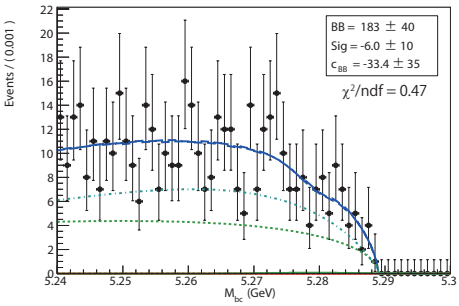
M_{bc} distributions on each M_{X_s} mass bin are fitted to obtain signal yields in Figure 10.6, 10.7 and 10.8. The fit result is shown in Table 10.6 and the branching fractions are calculated by using the efficiencies obtained from the

calibrated MC. The branching fraction at M_{X_s} <2.8 GeV/c^2 is

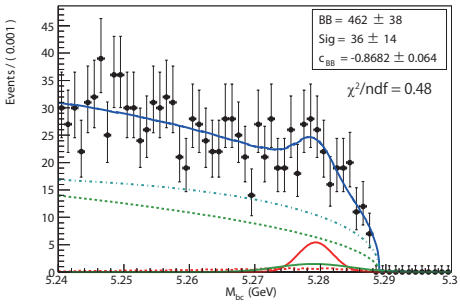
$$\mathcal{BR}(B \to X_s\gamma) \;=\; (3.51 \pm 0.17) \times 10^{-4} (M_{X_s} < 2.8\mathrm{GeV/c^2}). \quad (10.2)$$

Table 10.6: The partial branching fraction in each M_{X_s} mass bin(Full data)

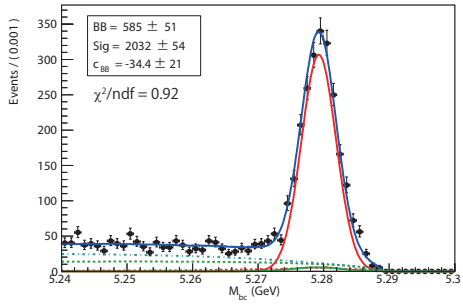
M_{X_s} bin(GeV/c^2)	Yield	Efficiency(%)	$\mathcal{BR}(10^{-6})$
0.6-0.7	-6±10	6.96	-0.1±0.1
0.7-0.8	36±14	7.19	0.3±0.1
0.8-0.9	2032±54	6.65	19.8±0.5
0.9-1.0	1689±49	6.98	15.7±0.5
1.0-1.1	301±27	6.72	2.9±0.3
1.1-1.2	310±31	4.21	4.8±0.5
1.2-1.3	1019±46	3.54	18.7±0.8
1.3-1.4	1117±50	3.33	21.8±1.0
1.4-1.5	1090±52	3.33	21.2±1.0
1.5-1.6	806±50	2.37	22.0±1.4
1.6-1.7	723±37	2.09	22.4±1.1
1.7-1.8	664±37	1.74	24.8±1.4
1.8-1.9	652±54	1.58	26.7±2.2
1.9-2.0	542±60	1.34	26.3±2.9
2.0-2.1	403±54	1.12	23.3±3.1
2.1-2.2	285±35	0.88	21.0±2.6
2.2-2.4	449±80	0.72	40.3±7.2
2.4-2.6	273±84	0.63	27.9±8.6
2.6-2.8	87±82	0.49	11.5±11
Total			351±17



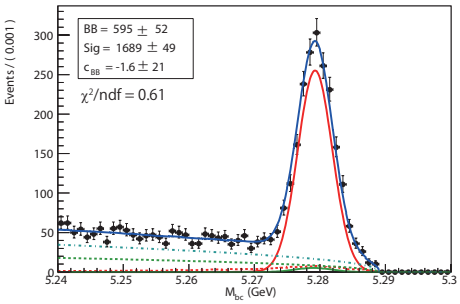
(a) 0.6< M_{X_s} <0.7 GeV



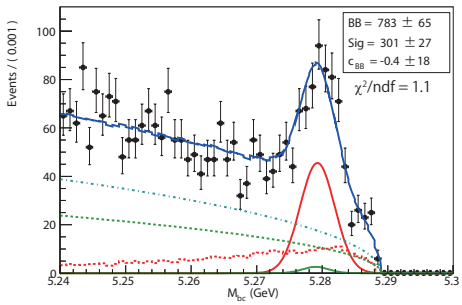
(b) 0.7< M_{X_s} <0.8 GeV



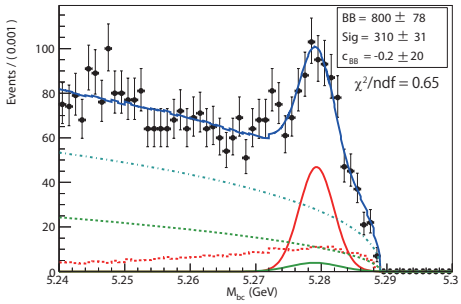
(c) 0.8< M_{X_s} <0.9 GeV



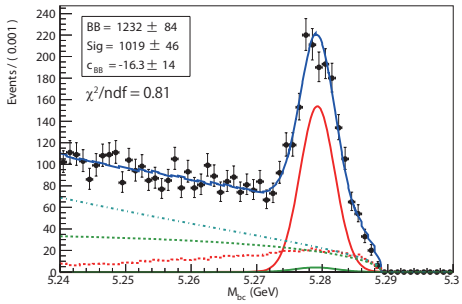
(d) 0.9< M_{X_s} <1.0 GeV



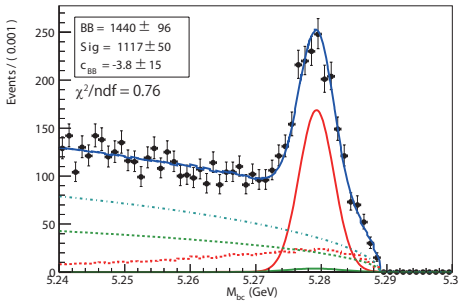
(e) 1.0< M_{X_s} <1.1 GeV



(f) 1.1< M_{X_s} <1.2 GeV

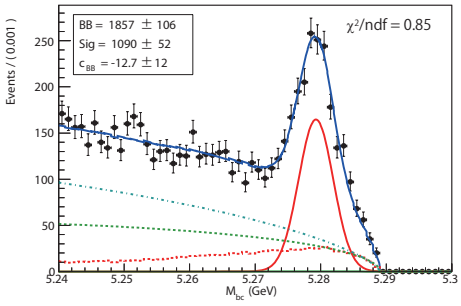


(g) 1.2< M_{X_s} <1.3 GeV

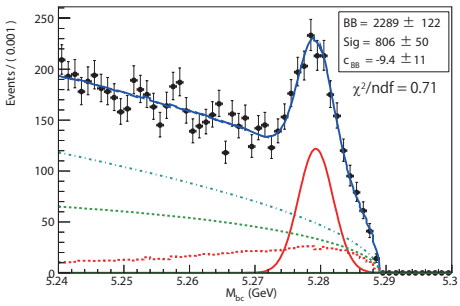


(h) 1.3< M_{X_s} <1.4 GeV

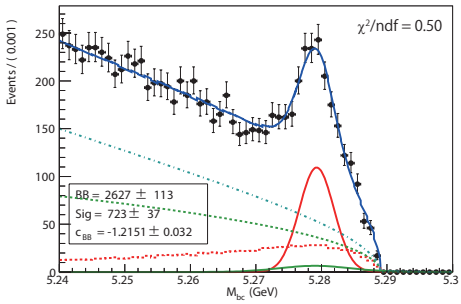
Figure 10.6: M_{bc} fit for each M_{X_s} bin(0.6-1.4 GeV, Full data)



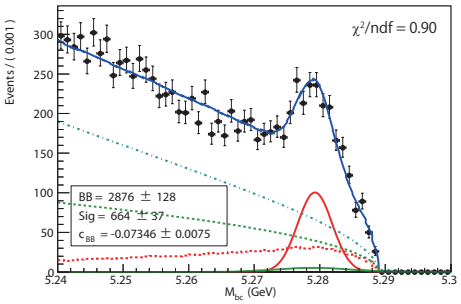
(a) 1.4< M_{X_s} <1.5 GeV



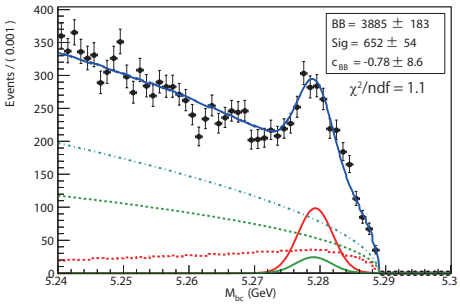
(b) 1.5< M_{X_s} <1.6 GeV



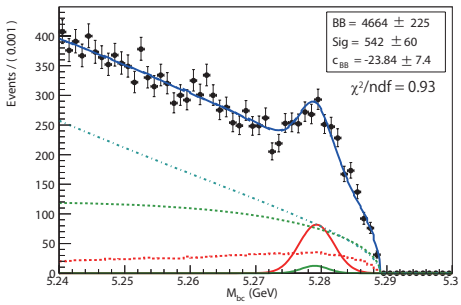
(c) 1.6< M_{X_s} <1.7 GeV



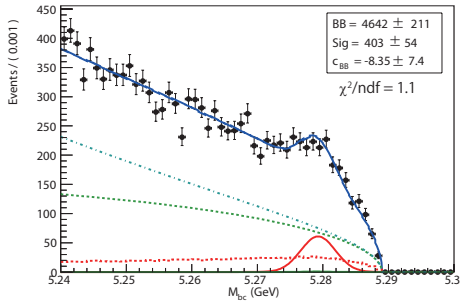
(d) 1.7< M_{X_s} <1.8 GeV



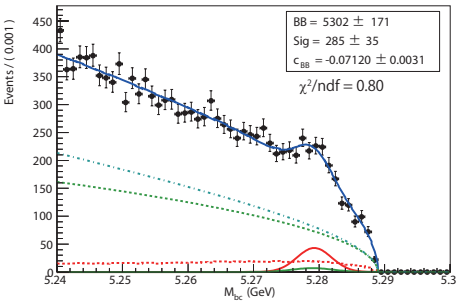
(e) 1.8< M_{X_s} <1.9 GeV



(f) 1.9< M_{X_s} <2.0 GeV

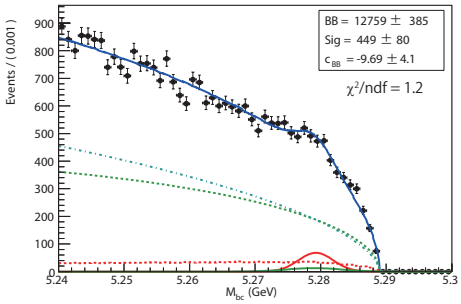


(g) 2.0< M_{X_s} <2.1 GeV

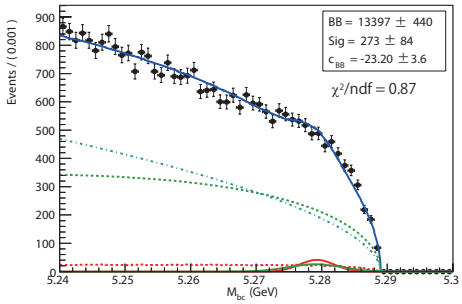


(h) 2.1< M_{X_s} <2.2 GeV

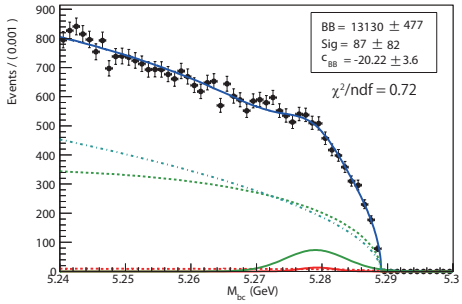
Figure 10.7: M_{bc} fit for each M_{X_s} bin(1.4-2.2 GeV, Full data)



(a) 2.2< M_{X_s} <2.4 GeV



(b) 2.4< M_{X_s} <2.6 GeV



(c) 2.6< M_{X_s} <2.8 GeV

Figure 10.8: M_{bc} fit for each M_{X_s} bin(2.2-2.8 GeV, Full data)

10.4 Systematic Uncertainties

10.4.1 Hadronization Model Uncertainty

To evaluate an uncertainty on the hadronization model, the fractions of the modes in Table 10.3 are fluctuated by $\pm 1\sigma$ and deviations from the default efficiency are assigned to the systematic uncertainties. +100%-50% fluctuation is given to only $K2\pi^0$ mode fractions in M_{X_s} >2.0 GeV/c^2 as mentioned in Section 10.2.3. The result is summarized in Table 10.7. Uncertainties in M_{X_s} >2.0 GeV/c^2 are large, especially $K4\pi$(Mode=7) and $K2\pi^0$(Mode=8). In calculating a total uncertainty, correlations between M_{X_s} bins are taken into account and it is 6.65 %. In addition, the systematic uncertainty on the $K^* - X_s$ transition position in Section 9.5.4 are added in the hadronization model uncertainty in 1.1< M_{X_s} <1.2 GeV/c$^2 s$.

Table 10.7: Systematic uncertainty(%) on hadronization model by reweighting the fraction in 1.15< M_{X_s} <2.8 GeV/c^2(Full data).

M_{X_s} bin	Fluctuated mode										
(GeV/c^2)	1	2	3	4	5	6	7	8	9	10	Total
1.1-1.2	0.97	0.25	0.30	0.59	0.01	0.04	0.01	2.17	0.00	0.00	2.48
1.2-1.3	0.65	0.16	0.44	0.56	0.03	0.55	0.02	1.72	0.03	0.00	2.05
1.3-1.4	0.76	0.23	0.42	0.46	0.10	1.95	0.06	1.28	0.02	0.00	2.55
1.4-1.5	1.19	0.20	0.34	0.37	0.12	1.58	0.92	3.26	0.42	0.00	3.98
1.5-1.6	0.91	0.40	0.85	0.17	0.01	0.97	0.57	1.52	0.40	0.20	2.35
1.6-1.7	0.49	0.29	0.95	0.01	0.22	0.95	0.84	2.10	0.43	0.35	2.75
1.7-1.8	0.46	0.24	0.88	0.07	0.53	0.69	1.35	2.81	0.22	0.33	3.41
1.8-1.9	0.44	0.21	0.86	0.06	0.77	0.67	1.32	2.98	0.43	0.31	3.60
1.9-2.0	0.33	0.19	0.88	0.13	0.81	0.56	1.49	3.00	0.39	0.35	3.66
2.0-2.1	1.42	10.7	2.72	0.67	1.98	1.79	4.97	12.5	0.92	1.83	17.8
2.1-2.2	1.87	13.2	2.90	0.95	2.17	1.20	5.70	15.9	1.65	1.40	21.9
2.2-2.4	1.77	16.8	2.95	1.11	1.96	0.81	6.54	17.4	0.75	1.68	25.5
2.4-2.6	2.56	1.12	6.88	2.02	5.29	5.22	18.2	19.8	3.81	3.67	29.6
2.6-2.8	1.99	1.27	5.81	1.79	4.72	5.04	18.8	20.2	2.79	3.58	29.4

10.4.2 Missing Mode

The fraction of the reconstructed modes in the MC after calibration is 76.6 % (1.15< M_{X_s} <2.8 GeV/c^2), where fractions of the K_L modes which corresponds to the reconstructed K_s modes are included. In order to evaluate the

uncertainty, we use the MC after the Pythia calibration. The parameters in Pythia are shifted within parameter region which is consistent with the model of the data. The result is shown in Table 10.8. The maximum and minimum of the fraction of the reconstructed mode are 78.0 and 76.3, respectively. Therefore, the differences from that of the calibrated MC are (78.0-76.6)/76.6 = +1.79 % and (76.3-76.6)/76.6 = -0.39 %, respectively. The uncertainty is needed in $< M_{X_s} >$1.15 GeV/c^2, thus, the fraction is considered. As a result, we assign ±1.59 % as the systematic uncertainty on the missing modes.

Table 10.8: Fractions(%) of reconstructed and missing modes. In calculation of the fraction of Reco mode, the K_L modes are taken into account.

Mode	calibrated MC	1	2	3	4	5	6	7
	PARJ(2)=0.10	0.20	0.10	0.10	0.10	0.05	0.10	0.10
	PARJ(11)=0.95	0.95	0.93	0.95	0.95	0.95	0.95	0.96
	PARJ(15)=0.25	0.25	0.25	0.20	0.25	0.25	0.30	0.25
	PARJ(25)=0.03	0.03	0.03	0.03	0.50	0.03	0.03	0.03
1	4.61	4.25	4.89	4.09	4.18	4.22	4.38	3.91
2	2.38	2.18	2.52	2.11	2.16	2.17	2.23	1.99
3	15.7	13.7	13.6	13.6	13.7	14.0	13.9	13.7
4	24.0	20.8	20.4	21.3	20.8	21.1	20.4	21.2
5	4.58	6.15	6.28	6.36	6.28	6.34	6.14	6.25
6	19.2	21.2	21.2	22.3	21.4	21.5	20.4	21.4
7	10.2	10.6	10.9	11.1	11.0	11.0	10.6	10.8
8	6.65	7.48	7.68	7.89	7.62	7.80	7.41	7.51
9	5.35	4.99	4.96	3.91	5.51	5.11	6.20	5.07
10	2.31	3.15	1.86	1.80	1.80	1.20	2.81	2.78
Reco mode	76.6	76.6	77.2	78.0	76.9	77.6	76.3	77.0
Missing mode	23.4	23.4	22.8	22.0	23.1	22.4	23.7	23.0

The fraction of the reconstructed mode in every mass bin are investigated, and the maximum and minimum values are reported in Table 10.9, which are assigned to the systematic uncertainty.

10.4.3 M_{bc} PDF

The systematic uncertainty is evaluated by the same method in the partial data analysis. The result is summarized in Table 10.10. Uncertainties from the PDFs of the peaking background and cross-feed in M_{X_s} >2.0 GeV/c^2 are large.

Table 10.9: The fractions of the reconstructed mode in every M_{X_s} mass bin

M_{X_s} bin(GeV/c^2)	Default	Maximum	Minimum	Systematics(%)
1.1-1.2	0.996	1.00	0.984	+0.38 -1.21
1.2-1.3	0.992	0.994	0.982	+0.24 -0.98
1.3-1.4	0.980	0.980	0.961	+0.00 -1.90
1.4-1.5	0.968	0.975	0.956	+0.68 -1.29
1.5-1.6	0.923	0.935	0.911	+1.22 -1.33
1.6-1.7	0.918	0.929	0.900	+1.21 -1.94
1.7-1.8	0.914	0.920	0.905	+0.60 -0.97
1.8-1.9	0.891	0.901	0.873	+1.11 -2.06
1.9-2.0	0.850	0.864	0.840	+1.58 -1.58
2.0-2.1	0.798	0.806	0.781	+0.99 -2.17
2.1-2.2	0.742	0.747	0.727	+0.78 -1.91
2.2-2.4	0.661	0.669	0.650	+1.17 -1.57
2.4-2.6	0.548	0.553	0.542	+0.97 -1.00
2.6-2.8	0.456	0.460	0.447	+0.70 -2.04

Table 10.10: Systematic uncertainties(%) on M_{bc} PDF(Full data).

M_{X_s} bin (GeV/c^2)	Signal	Scf	Peaking BG	$q\overline{q}$ BG	Total
0.6-0.7	0.00	0.00	0.00	0.00	0.00
0.7-0.8	0.14	12.2	7.81	0.00	14.6
0.8-0.9	0.17	0.43	0.51	0.01	0.68
0.9-1.0	0.05	0.45	0.38	0.03	0.59
1.0-1.1	0.05	2.87	1.14	0.28	3.10
1.1-1.2	0.39	3.08	1.67	0.22	3.54
1.2-1.3	0.21	1.58	0.92	0.02	1.84
1.3-1.4	0.21	1.60	0.20	0.00	1.63
1.4-1.5	0.23	1.97	0.06	0.02	1.99
1.5-1.6	0.64	2.21	0.07	0.02	2.30
1.6-1.7	0.08	1.72	2.14	0.20	2.75
1.7-1.8	0.08	2.17	1.70	0.24	2.77
1.8-1.9	0.02	1.92	1.98	0.10	2.76
1.9-2.0	0.13	4.23	4.04	0.13	5.85
2.0-2.1	0.05	5.63	0.61	0.15	5.67
2.1-2.2	0.26	3.65	2.46	0.37	4.42
2.2-2.4	0.03	7.37	7.13	0.04	10.3
2.4-2.6	0.05	11.5	21.8	0.33	24.6
2.6-2.8	0.15	44.7	100	0.89	101

10.4.4 Summary of Systematic Uncertainty

Systematic uncertainties in each M_{X_s} mass bin are summarized in Table 10.11. We take each of these systematic uncertainties to be uncorrelated within an

M_{X_s} bin. The $B\overline{B}$ counting, detector response and background rejection systematics uncertainties are taken to be completely correlated between all mass bins. We take the uncertainties on the M_{bc} PDFs except for the cross-feed to be uncorrelated between all mass bins, and the uncertainty on the cross-feed PDF to be completely correlated. As for the fragmentation and missing fraction uncertainties are evaluated in different mass regions. We take the uncertainty on these mass regions to be uncorrelated with one another, but completely correlated between the mass bins within a mass region. Total systematic uncertainty is reported in Table 10.12.

Table 10.11: Systematic uncertainties(%) in every M_{X_s} mass bin(Full data).

M_{X_s} bin (GeV/c^2)	$B\overline{B}$ counting	Detector response	Background rejection	M_{bc} PDF	Hadronization	Missing mode	Total
0.6-0.7	1.37	2.68	3.38	0.00	-	-	4.52
0.7-0.8	1.37	2.59	3.38	14.6	-	-	15.3
0.8-0.9	1.37	2.57	3.38	0.68	-	-	4.51
0.9-1.0	1.37	2.58	3.38	0.59	-	-	4.51
1.0-1.1	1.37	2.57	3.38	3.10	-	-	5.43
1.1-1.2	1.37	2.99	3.38	3.54	32.1	1.21	32.1
1.2-1.3	1.37	3.15	3.38	1.84	2.05	0.98	5.64
1.3-1.4	1.37	3.17	3.38	1.63	2.55	1.90	6.01
1.4-1.5	1.37	3.10	3.38	1.99	3.98	1.29	6.66
1.5-1.6	1.37	3.34	3.38	2.30	2.35	1.33	6.09
1.6-1.7	1.37	3.49	3.38	2.75	2.75	1.94	6.66
1.7-1.8	1.37	3.59	3.38	2.77	3.41	0.97	6.82
1.8-1.9	1.37	3.69	3.38	2.76	3.60	2.06	7.19
1.9-2.0	1.37	3.74	3.38	5.85	3.66	1.58	8.80
2.0-2.1	1.37	3.78	3.38	5.67	17.8	2.17	19.5
2.1-2.2	1.37	3.78	3.38	4.42	21.9	1.91	23.1
2.2-2.4	1.37	3.77	3.38	10.3	25.5	1.57	28.0
2.4-2.6	1.37	3.79	3.38	24.6	29.6	1.00	38.9
2.6-2.8	1.37	3.82	3.38	101	29.4	2.04	114

Table 10.12: Systematic uncertainty(%) (Full data)

Source	Systematic uncertainty (%)
$B\overline{B}$ counting	1.37
Detector response	2.98
Background rejection	3.38
M_{bc} PDF	5.06
Hadronization model	6.66
Missing mode	1.59
Total	9.3

10.5 Results

10.5.1 Partial Branching Fraction

The partial branching fractions on M_{X_s} are summarized in Table 10.13, and is plotted in Figure 10.9. In Figure 10.9(b), the partial branching fractions are compared with the result in the BaBar's measurement[74]. Most of the branching fractions in this result are in a good agreement with the BaBar's one, but the branching fractions in 2.0< M_{X_s} <2.2 and 2.2< M_{X_s} <2.4 GeV/c^2 have deviations $+2.2\sigma$ and $+1.6\sigma$, respectively. We guess these deviations come from a estimation of the peaking background since effects of the peaking background become large in M_{X_s} >1.8 GeV/c^2. In BaBar's measurement, the peaking background is estimated by the MC, while we estimate it by the data(Section 7.4). In order to investigate the validity of our result, we check an efficiency of the D veto for the peaking background since the D veto is applied in M_{X_s} >2.0 GeV/c^2. If there is a difference on the efficiency of the D veto for the peaking background between the MC and data, the deviations from the BaBar's result in M_{X_s} >2.0 GeV/c^2 may occur. The efficiency of the D veto for the peaking background is investigated with the anti-π^0/η veto sample(same as Section 7.4) and we do not observe a large difference between the MC and data in Appendix F. Thus, we conclude that our measurement is correct.

10.5.2 Total Branching Fraction

The total branching fraction is

$$\mathcal{BR}(B \to X_s\gamma) = (3.51 \pm 0.17 \pm 0.33) \times 10^{-4} (M_{X_s} < 2.8 \mathrm{GeV/c^2}), \quad (10.3)$$

where the first uncertainty is statistical and the second is the systematic.

The extrapolated one to E_γ(B rest frame)>1.6 GeV in the same method as Section 9.4 is

$$\mathcal{BR}(B \to X_s\gamma) = (3.74 \pm 0.18 \pm 0.35) \times 10^{-4} (E_\gamma > 1.6 \mathrm{GeV}), \quad (10.4)$$

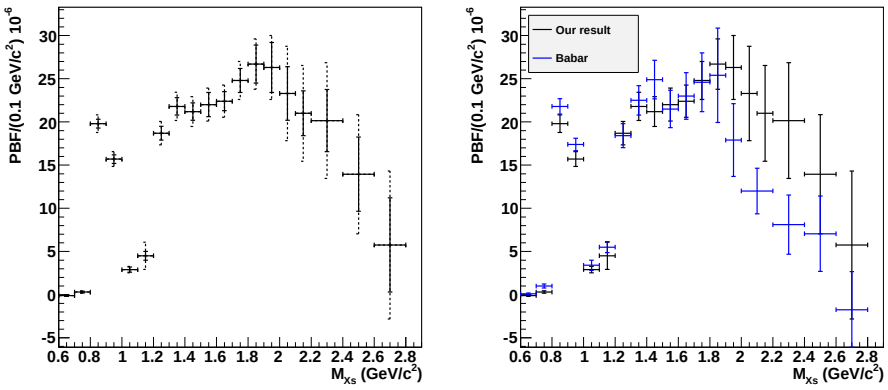
where the extrapolation factor of the systematic uncertainty is added.

This measurement result is compared with other measurements in Figure 10.10. Our result is the best measurement in the semi-inclusive method, and is same level as the full-inclusive method which is the best result. This is also in

Table 10.13: The partial branching fraction on M_{X_s}

M_{X_s} bin(GeV/c^2)	$\mathcal{BR}(10^{-6})$
0.6-0.7	-0.1±0.1±0.0
0.7-0.8	0.3±0.1±0.1
0.8-0.9	19.8±0.5±0.9
0.9-1.0	15.7±0.5±0.7
1.0-1.1	2.9±0.3±0.2
1.1-1.2	4.8±0.5±1.5
1.2-1.3	18.7±0.8±1.1
1.3-1.4	21.8±1.0±1.3
1.4-1.5	21.2±1.0±1.4
1.5-1.6	22.0±1.4±1.3
1.6-1.7	22.4±1.1±1.5
1.7-1.8	24.8±1.4±1.7
1.8-1.9	26.7±2.2±1.9
1.9-2.0	26.3±2.9±2.3
2.0-2.1	23.3±3.1±4.5
2.1-2.2	21.0±2.6±4.9
2.2-2.4	40.3±7.2±11
2.4-2.6	27.9±8.6±11
2.6-2.8	11.5±11±13

a good agreement with the world average, $(3.55\pm0.26)\times10^{-4}$, within 0.4σ. In addition, this result is consistent with the SM prediction, $(3.15\pm0.23)\times10^{-4}$, within 1.3σ, and provides a constraint to the NP model.



(a) Partial branching fraction.The first solid error is the statistical one and the second dashed error is a quadratic sum of the statistical and systematic errors.

(b) Partial branching fraction compared with the BaBar's result[74]

Figure 10.9: Partial branching fraction

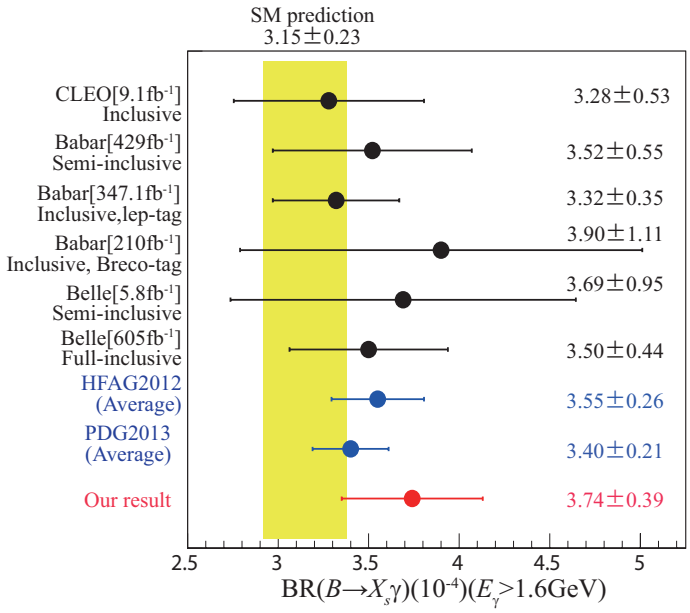


Figure 10.10: Measurement results and theoretical calculation of $B \to X_s\gamma$

10.5.3 Constraint to the 2HDM

We evaluate a constraint to the 2HDM as an example, and show two-sided 68%, 95% and 99% CL exclusion regions in $M_{H^\pm}$ versus tanβ in Figure 10.11. The charged Higgs mass region below 238 GeV/c^2 is excluded at 95% CL.

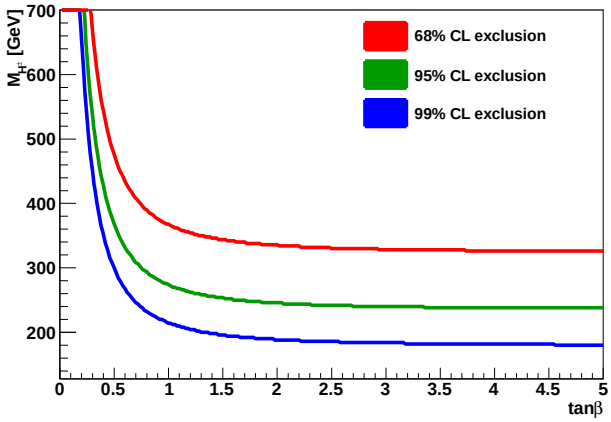


Figure 10.11: Constraints in the 2HDM parameter plane. The regions below the lines are excluded.

Chapter 11

Conclusions

11.1 Summary of Results

We measured the inclusive branching fraction for the radiative B meson decay $B \to X_s\gamma$ with a semi-inclusive reconstruction method. The measured branching ratio in $M_{X_s} < 2.8\mathrm{GeV/c^2}$ is

$$\mathcal{BR}(B \to X_s\gamma) = (3.51 \pm 0.17 \pm 0.33) \times 10^{-4} (M_{X_s} < 2.8\mathrm{GeV/c^2}), \quad (11.1)$$

where the first uncertainty is statistical and the second is the systematic. The extrapolated branching ratio to E_γ(B rest frame)>1.6 GeV is

$$\mathcal{BR}(B \to X_s\gamma) = (3.74 \pm 0.18 \pm 0.35) \times 10^{-4} (E_\gamma > 1.6\mathrm{GeV}). \quad (11.2)$$

This result is in a good agreement with the world average, (3.55±0.26)×10^{-4}, within 0.4σ. In addition, this result is consistent with the SM prediction, (3.15±0.23)×10^{-4}, within 1.3σ. Our result provides a constraint the NP model. We evaluate a constraint to the 2HDM as an example, and the charged Higgs mass region below 238 GeV/c^2 is excluded at 95% CL.

11.2 Perspectives

Improvements on the $\mathcal{BR}(B \to X_s\gamma)$ are expected from the KEKB and Belle upgrade project(Belle II) with an final target of 50 ab^{-1} integrated luminosity. The precision is dominated by the systematic uncertainty, especially on the hadronization model. The hadronization model uncertainty, however, depends on the data statistics, thus, is expected to become small at the Belle II by more

precise measurements of the each mode fraction. A prominent improvement can not be achieved easily, but we can expect to be an improvement above a factor two in my view. If the uncertainty on the current average of the inclusive branching fraction is reduced by half, not changing the the central value, the deviation from the SM prediction is just 1.5σ, where the theoretical uncertainty is much larger. If the uncertainty is reduced by half and the central value becomes same as the SM prediction, the constraint to the charged Higgs mass in the 2HDM is $\sim$600 GeV/c in Figure 11.1[21].

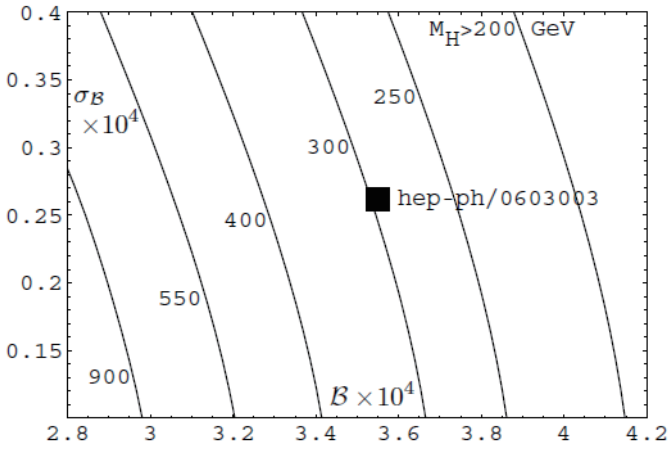


Figure 11.1: The 95% CL lower bound on $M_{H^\pm}$ as a function of the experimental central value(horizontal axis) and uncertainty(vertical axis).

In Belle II, the fully-inclusive and B full-reconstruction measurement as mentioned in Section 2.3 will be important. In addition, the theoretical uncertainty should be improved for the NP search with a higher precision.

Appendix A

Flavor Changing Neutral Current

In the SM, the charged current interacts as

$$\mathcal{L} = -\frac{g}{\sqrt{2}}\overline{U}_L\gamma^\mu D_L W_\mu^\dagger, \tag{A.1}$$

where U_L and D_L are gauge interaction eigenstates of the up-type and down-type quarks. This is transformed by using the mass eigenstates,

$$\mathcal{L} = -\frac{g}{\sqrt{2}}\overline{u}_L\gamma^\mu V_{uL}V_{dL}^\dagger d_L W_\mu^\dagger, \tag{A.2}$$

where u_L and d_L are mass eigenstates of the up-type and down-type quarks, V_{uL} and V_{dL} are unitary matrices to transform the mass matrix to a diagonal ones. Therefore, this can not keep a diagonal one. The matrix shown in Equation A.2,

$$V \equiv V_{uL}V_{dL}^\dagger, \tag{A.3}$$

is the Cabbibo-Kobayashi-Masukawa matrix, in which the flavor-mixing is occurred.

On the other hand, the electromagnetic and the neutral current have couplings which are same between the generations. The interaction of left-handed uptype-quark and Z is described as

$$\mathcal{L} = -g_Z\left(\frac{1}{2} - \frac{2}{3}\sin^2\theta_W\right)\overline{U}_L\gamma^\mu U_L Z_\mu^\dagger. \tag{A.4}$$

This can keep a diagonal one by transforming to mass eigenstates with $V_{uL}V_{uL}^\dagger =$

1.

In the SM, FCNC is forbidden since quarks with same charge have same descriptions of the gauge symmetry.

Appendix B

Parameter Setting for Pythia in the Belle

Pythia[19] has a huge parameters on the hadronization model. In the Belle, default values in Pythia are used for most parameters, but there are some parameters changed from the default values in the Belle. Changed parameters are shown in Table B.1.

Table B.1: Pythia parameters in the Belle

Parameter	Description	Default Value	Belle Value
MSTJ(107)	Radiative correction to continuum events	0(No)	2(Yes)
PARJ(14)	Probability that a spin=0 meson is produced with an orbital angular momentum 1, for a total spin=1.	0.0	0.05
PARJ(15)	Probability that a spin=1 meson is produced with an orbital angular momentum 1, for a total spin=0.	0.0	0.05
PARJ(16)	Probability that a spin=1 meson is produced with an orbital angular momentum 1, for a total spin=1.	0.0	0.05
PARJ(17)	Probability that a spin=1 meson is produced with an orbital angular momentum 1, for a total spin=2.	0.0	0.05
PARJ(33)	Energy below which the fragmentation of a parton system is stopped and two final hadron is formed	0.80	0.30

Appendix C

KSFW

The KSFW[59] is a Fisher discriminant[75] extended from the Fox-Wolfram variable[76] using information such as missing mass calculated from the daughter particles of the signal candidates and all the other particles in this event. The KSFW is constructed from 17 varialbes.

C.1 Fox-Wolfram Momenta

In most of B meson decay studies, large background comes from the $e^+e^- \to q\overline{q}(q = u, d, s, c)$. A difference of the event topology between B decay and $e^+e^- \to q\overline{q}$ can be characterized on the basis of the Fox-Wolfram momenta. The l-th moment is defined in CM the frame as

$$H_l = \sum_{ij} |\overrightarrow{p_i}||\overrightarrow{p_j}| P_l(cos\theta_{ij}), \tag{C.1}$$

where P_l is the l-th Legendre polynomial, $\overrightarrow{p_i}$ and $\overrightarrow{p_j}$ are the momenta of the i-th and j-th particles, respectively, θ_{ij} is the angle between the two momentum vectors. The sum is over the particles in the final state. Note that the overall constant is ignored here for simplicity.

C.2 SFW

SFW is devised by modifying the Fox-Wolfram moment H_l. In the SFW, the H_l is divided into three components and categorize the particles to the two

types; B signal candidate particles and remaining particles.

$$H_l = H_l^{SS} + H_l^{SO} + H_l^{OO} \quad \text{(C.2)}$$
$$H_l^{SS} = \sum_{ij} |p_i||p_j| P_l(cos\theta_{ij}) \quad \text{(C.3)}$$
$$H_l^{SO} = \sum_{jk} |p_j||p_k| P_l(cos\theta_{jk}) \quad \text{(C.4)}$$
$$H_l^{OO} = \sum_{kl} |p_k||p_l| P_l(cos\theta_{kl}) \quad \text{(C.5)}$$

where i and j iterate over B signal candidate particles(denoted by S for the signal) and the indices k and l iterate over the remaining particles(denoted by O for other) in the event. Extended Fisher discriminant names SFW(Super Fox-Wolfram) by the divided Fox-Wolfram moments is defined as

$$SFW \equiv \sum_l \alpha_l \left(\frac{H_l^{SO}}{H_O^{SO}} \right) + \sum_l \beta_l \left(\frac{H_0^{OO}}{H_0^{OO}} \right) \quad \text{(C.6)}$$

where α_l and β_l are Fisher coefficients.

C.3 KSFW

To increase the discrimination power, the SFW is modified to Kakuno Super Fox-Wolfram(KSFW) by taking into account charges of the particles, the missing mass of the event and normalization factor. The KSFW is defined as

$$KSFW \equiv \sum_{l=0}^{4} R_l^{SO} + \sum_{l=0}^{4} R_l^{OO} + \gamma \sum_{n=1}^{N_i} |(p_t)_n| \quad \text{(C.7)}$$

where N_i is a number of particles and γ is a Fisher coefficients to be optimized. Descriptions on these terms are provided in the following.

$\boxed{R_l^{SO}}$ A missing pseudo-particle is introduced as one particle that has the event's missing energy and momentum p_{miss}. Furthermore, the remaining particles int the event are categorized into three, "charged", "neutral" and "missing".

$$R_l{}^{SO} = \frac{\alpha_l{}^{charged} H_l{}^{SO_{charged}} + \alpha_l{}^{neutral} H_l{}^{SO_{neutral}} + \alpha_l{}^{missing} H_l{}^{SO_{missing}}}{E_{beam} - \Delta E}, \quad \text{(C.8)}$$

where $\alpha_l^i (i =$charged, neutral, missing) are Fisher coefficients. For signal and other remaining charged particles,

$$H_l^{SO_{charged}} = \begin{cases} \sum_i \sum_j |p_j| P_l(cos\theta_{ij}) & (\mathrm{l}: \mathrm{even}) \\ \sum_i \sum_j |p_j| Q_i Q_j P_l(cos\theta_{ij}) & (\mathrm{l}: \mathrm{odd}) \end{cases} \tag{C.9}$$

where the index i iterates over the particles in the B signal candidates and the index j iterates over all other remaining charged particles. The $Q_{i,j}$ are the charge of the particles i and j.

For signal and other remaining neutral particles,

$$H_l^{SO_{neutral}} = \begin{cases} \sum_i \sum_j |p_j| P_l(cos\theta_{ij}) & (\mathrm{l}: \mathrm{even}) \\ 0 & (\mathrm{l}: \mathrm{odd}). \end{cases} \tag{C.10}$$

For signal missing particles,

$$H_l^{SO_{missing}} = \begin{cases} \sum_i \sum_j |p_j| P_l(cos\theta_{iM}) & (\mathrm{l}: \mathrm{even}) \\ 0 & (\mathrm{l}: \mathrm{odd}). \end{cases} \tag{C.11}$$

where θ_{iM} is the opening angle between p_i and p_{miss}.

$\boxed{R_l^{OO}}$

$$R_l^{OO} = \frac{\beta_l H_l^O O}{(E_{beam} - \Delta E)^2} \tag{C.12}$$

$$H_l^{OO} = \begin{cases} \sum_i \sum_j |p_j||p_k| P_l(cos\theta_{jk}) & (\mathrm{l}: \mathrm{even}) \\ \sum_i \sum_j |p_j||p_k| Q_i Q_j P_l(cos\theta_{jk}) & (\mathrm{l}: \mathrm{odd}) \end{cases} \tag{C.13}$$

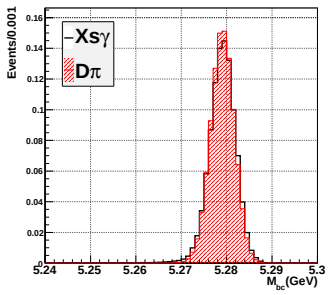
where the indices j and k iterates over all other particles.

$\boxed{\sum_{n=1}^{N_i} |(p_t)_n|}$ $\sum_{n=1}^{N_i} |(p_t)_n|$ is the scalar sum of the transverse momenta p_t of all the particles in the signal candidates and all the other remaining particles.

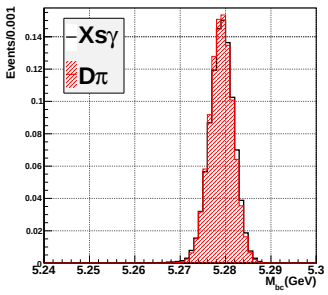
Appendix D

Signal PDF Study

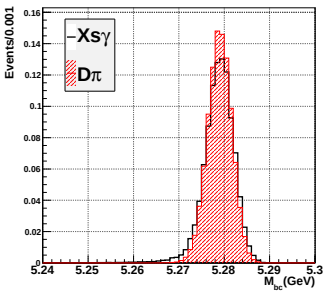
We check M_{bc} distributions of $D\pi$ to make the signal PDF. In Figure D.1, M_{bc} distributions of $D\pi$ are compared with the signal ones by MC. Total distributions of $D\pi$ are consistent with that of the signal, and we also investigate M_{bc} distributions on no π^0/η, one π^0/η and two π^0/η modes. M_{bc} distributions of $D\pi$ on no π^0/η mode is in a good agreement with that of the signal, but distributions of $D\pi$ on one and two π^0/η modes are different, particularly tail regions in low side.



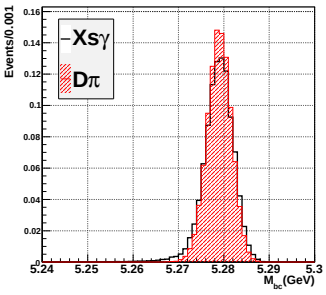
(a) Total



(b) No π^0/η mode



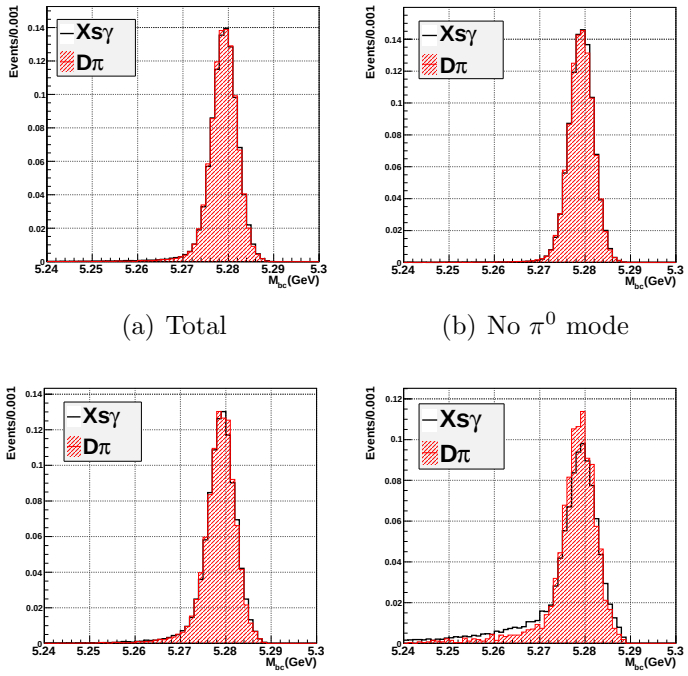
(c) $1\pi^0/\eta$ mode



(d) $2\pi^0/\eta$ mode

Figure D.1: M_{bc} distributions of $X_s\gamma$ and $D\pi$(MC)

In Figure D.2, M_{bc} distributions without photon momentum correction in Equation 5.3 of $X_s\gamma$ and $D\pi$ are shown. These $D\pi$ distributions are better consistent with them of the signal, especially one π^0 mode. We conclude this photon momentum correction works a bit differently between $D\pi$ and the signal because a charged pion is corrected in $D\pi$ event instead of a photon.



(a) Total (b) No π^0 mode

(c) $1\pi^0$ mode (d) $2\pi^0$ mode

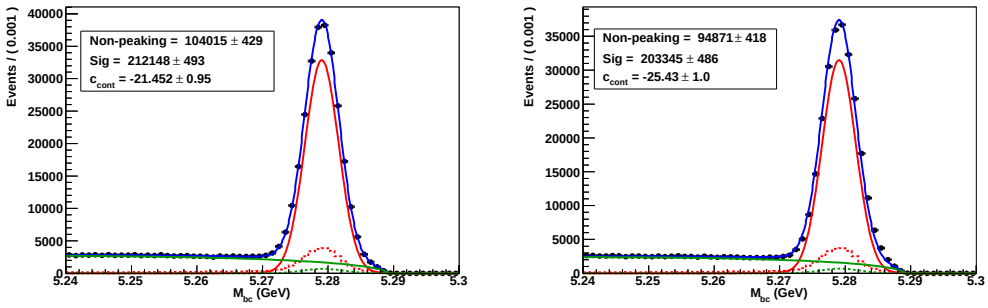
Figure D.2: M_{bc} distributions without photon momentum correction in Equation 5.3 of $X_s\gamma$ and $D\pi$(MC)

Appendix E

Control Sample Study

Before opening the signal box of the data, a control sample, $B \to D\pi(B^+ \to D^0\pi^+, B^0 \to D^-\pi^+)$ is studied in order to confirm the validity on this analysis. The $B \to D\pi$ reconstruction method is same as that of the systematic study on the $q\bar{q}$ background suppression in Section 8.3.2.

Figure E.1 is the M_{bc} distributions of $B \to D\pi$ after all selections, $q\bar{q}$suppression and BCS applied the signal. The efficiency obtained from the $D\pi$ MC is 3.395



(a) MC (b) Data

Figure E.1: M_{bc} distributions on $B \to D\pi$(Red solid line: Signal, Red dashed line: Cross-feed, Green Solid line: Non-peaking background, Green Dashed line:Peaking background)

%, thus the branching fraction is calculated as

$$\mathcal{BR}(B \to D\pi) = \frac{(203345 \pm 486)}{0.03395 \times 2 \times 771 \times 10^6} = (3.88 \pm 0.01) \times 10^{-3}, \quad \text{(E.1)}$$

where the uncertainty is only statistical one. According to PDG,

- $\mathcal{BR}(B^+ \to D^0\pi^+) = (4.81 \pm 0.15) \times 10^{-3}$
- $\mathcal{BR}(B^0 \to D^-\pi^+) = (2.68 \pm 0.13) \times 10^{-3}$

Thus,

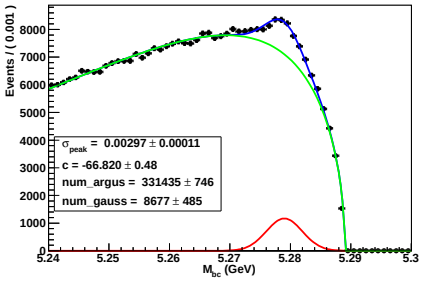
$$\begin{aligned}\mathcal{BR}(B \to D\pi) &= (4.81 \pm 0.15) \times 10^{-3} \times 0.513\ (\Upsilon(4S) \to B^+B^-) \\ &+ (2.68 \pm 0.13) \times 10^{-3} \times 0.487\ (\Upsilon(4S) \to B^0\overline{B}^0) \\ &= (3.77 \pm 0.14) \times 10^{-3}. \end{aligned} \tag{E.2}$$

Therefore, the result(Equation E.1) in our analysis is consistent with the PDG value (Equation E.2) and we can confirm the validity on our analysis.

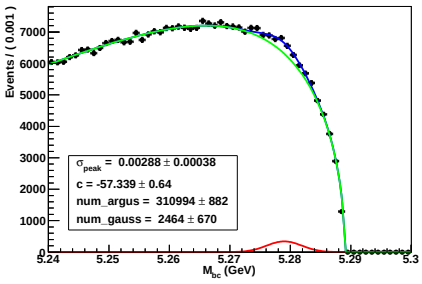
Appendix F

Efficiency of the D veto for the peaking background

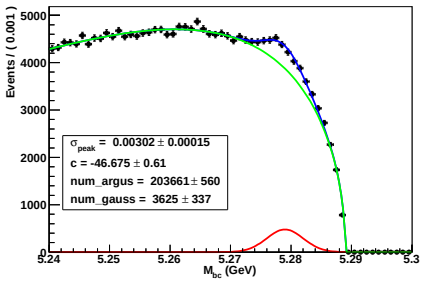
We investigate efficiencies of the D veto for MC and data to check the validity of the branching ratios in M_{X_s} >2.0 GeV/c^2. In this study, the anti-π^0/η veto sample which is defined by requiring the π^0/η probabilities above 0.8 are used in the same method as Section 7.4. In Figure F.1, M_{bc} distributios of the $B\overline{B}$ background MC and the data in M_{X_s} >2.0 GeV/c^2 before and after the D veto are shown. Efficiencies of the MC and data are 28.4 % and 26.1 %, respectively and are not large difference and consistent within the errors. Therefore, the D veto can not be the origin of the deviations from the BaBar's result in M_{X_s} >2.0 GeV/cc^2.



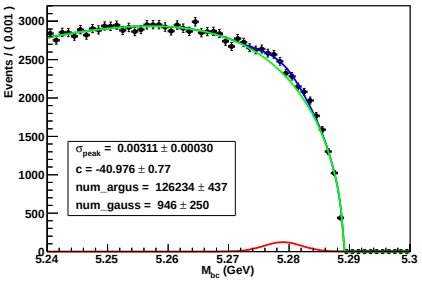
(a) $B\overline{B}$ background MC before D veto



(b) $B\overline{B}$ background MC after D veto



(c) DATA before D veto



(d) DATA after D veto

Figure F.1: M_{bc} distributions before/after D veto(Red solid line is the peaking background)

Bibliography

[1] Tobias Hurth and Mikihiko Nakao, Ann.Re.Nucl.Part.Sci.60,645-677(2010), arXiv:hep-ph/1005.1224

[2] Tobias Hurth, Rev.Mod.Phys.75,1159-1199(2003), arXiv:hep-ph/0212304

[3] Hurth T, Lunghi E, In the Proceedings of 2nd Workshop on the CKM Unitarity Triangle, Durham, England, 5-9, Apr 2003, arXiv:hep-ph/0703226

[4] Hurth T, Int.J.Mod.Phys.A 22:1781(2003), arXiv:hep-ph/0307142

[5] K.G. Wilson, Phys.Rev.,179:1499(1969)

[6] K.G. Wilson, Phys.Rev.D,3,1818-1846(1971)

[7] Gaillard MK and Lee BW, Phys.Rev.Lett.33:108(1974)

[8] Altarelli G and Maiani L, Phys.Lett.B 247:399(1990)

[9] Witten E, Nucl.Phys.Lett. B 52:351(1974)

[10] Chay J, George H, Grinstein B, Phys.Lett. B 247:399(1990)

[11] Bigi II,Uraltsev NG, Vainshtein AI, Phys.Lett.B, 293:430(1992)

[12] Bigi II, et al, arXiv:hep-ph/9703290

[13] Bigi II, Shifman MA and Uraltsev N.Ann, Rev.Nucl.Part.Sci., 47:591(1997)

[14] Manohar AV and Wise MB, Phys. Rev. D, 49:1310(1994)

[15] Manohar AV and Wise MB, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol, 10:1(2000)

[16] Falk AF, Luke ME, Savage MJ, Phys. Rev. D 49:3367(1994), arXiv:hep-ph/9308288

[17] Ali A, Hiller G, Handoko LT, Morozumi T, Phys. Rev. D 55:4105(1997), arXiv:hep-ph/0006124

[18] N. G. Deshpande, OITS-530

[19] Torbjorn Sjostrand, Stephen Mrenna and Peter Skands, JHEP 0605:026(2006), arXiv:hep-ph/0603175

[20] Heavy Flavor Averaging Group(HFAG), arXiv:hep-ex/12071158, 2012

[21] M. Misiak, M. Steinhauser, Nucl. Phys. B 764 (2007), 62-84

[22] M. Misiak, M. Steinhauser, Nucl. Phys. B 840 (2010), 271-283.

[23] M. Misiak, M. Poradzinski, Phys. Rev. D 83, 014024 (2011).

[24] H. E. Haber, G. L. Kane and T. Sterling, Nucl. Phys. B161, 493(1979)

[25] L. F. Abbott, P. Sikivie and M. B. Wise, Phys. Rev. D21, 1393(1980)

[26] LEP Higgs Working Group for Higgs boson searches, hep-ex/0107031

[27] Misiak M, et al., Phys.Rev.Lett. 98:022002(2007), arXiv:hep-ph/0609232

[28] The Gfitter Group, Eur.Phys.J.C60,543-583(2009)

[29] T. Hermann, M. Misiak, M. Steinhauser, JHEP, 1211(2012), 036, arXiv:hep-ph1208.2788

[30] ATLAS-CONF-2013-090 (2013)

[31] A. L. Kagan and M. Neubert, Phys.Rev. D 58(1998)094012, arXiv:hep-ph/9803368

[32] D. Atwood, M. Gronau and A. Soni, Phys. Rev. Lett. 79, 185(1997)

[33] D. Atwood, M. Hazumi and A. Soni, Phys. Rev. D 71, 076003(2005)

[34] Shohei Nishida and Belle Collab., Phys.Rev.Lett., 93, 031803(2004)

[35] Aubert B and Babar Collab., Phys.Rev.Lett., 101, 171804(2008)

[36] M.Misiak et, al., Phys. Rev. Lett. 98, 022002(2007)

[37] The LHCb collaboration, JHEP07 (2012) 133

[38] M.Benzke, S.J.Lee, M.Neubert and G.Paz, Phys.Rev.Lett, 106, 141801(2011)

[39] Belle, Phys.Lett.B511:151-158(2001)

[40] Belle collaboration, Nucl. Instrum. Meth. A 479, 117 (2002)

[41] V. Chabaud, et al., Nucl. Instr. and Meth. A 368(1996)

[42] E. Nygard, et al., Nucl. Instr. and Meth. A 301 (1991)506
O. Toker, et al., Nucl. Instr. and Meth. A 340 (1994) 572

[43] G. Alimonti et al. [Belle Collab.], Nucl. Instrum. Meth. A 453, 71 (2000)

[44] H. Hirano, et al. [Belle Collab.], Nucl. Instrum. Meth. A 455(2000) 294-304

[45] T. Iijima, et al. [Belle Collab.], Nucl. Instrum. Meth. A 453(2000) 321-325

[46] H. Kichimi, et al. [Belle Collab.], Nucl. Instrum. Meth. A 453(2000), 315-320

[47] T. Iijima, et al. [Belle Collab.], Nucl. Instrum. Meth. A 387(2000) 64-68

[48] H. Sagawa, [Belle Collab.], Nucl. Instrum. Meth. A 453(2000) 259-261

[49] H. Ikeda, et. al, [Belle Collab.], Nucl. Instrum. Meth. A 441(2000) 401-426

[50] A. Abashian et al. [Belle Collab.], Nucl. Instrum. Meth. A 449(2000) 112-124

[51] E. Nkano, Nucl. Instrum. Meth. A 494, 402 (2002)

[52] A. Abashian et al. [Belle Collab.], Nucl. Instrum. Meth. A 491(2002) 69-82

[53] http://robbep.web.cern.ch/robbep/EvtGen/GuideEvtGen.pdf

[54] B. Anderson, et al, Phys. Rep. 97, 31 (1983)

[55] R. Brun et al., GEANT3.21, CERN Report No. DD/EE/84-1(1987)

[56] A. L. Kagan, M. Neubert, arXiv:hep-ph/9805303

[57] S. Veseli and M. G. Olsson, Phys.Lett. B367, 309(1996)

[58] P. Koppenburg and Belle Collaboration, Phys, Rev, Lett, 93, 6(2004)

[59] S. H. Lee et al. [Belle Collab.], Phys. Rev. Lett. 91, 261801 (2003)

[60] M. Feindt and U. Kerzel, Nucl. Instrum. Meth Phys. Res., A 559 (2006) 190-194

[61] M. Feindt and U. Kerzel, Nucl. Instrum. Meth. A 559(2006) 190-194

[62] Bellenote 711

[63] Bellenote 1165

[64] Bellenote 1176

[65] Bellenote 1253

[66] Bellenote 1207

[67] Bellenote 645

[68] Bellenote 1176

[69] Bellenote 779

[70] Jeppe R. Andersen and Einan Gardi, JHEP01(2007)029

[71] A.Limosani et al.(The Belle collaboration), Phys.Rev.Lett. 103, 241801, 2009

[72] M.Neubert, Phys.Lett.B, 612, 2005

[73] O. Buchmuller and H. Flacher, Phys. Rev. D73, 073008, (2006)

[74] The BABAR Collaboration, Phys. Rev. D. 86, 052012 (2012)

[75] R. A. Fisher, Annals of Human genetics 7, 179 (1936)

[76] G. C. Fox an S. Wolfram, Phys. Rev. Lett. 41, 1581 (1978)