Ph.D. Thesis

First Evidence of the Suppressed B Meson Decay $B^- \to DK^-$ Followed by $D \to K^+\pi^-$ and Extraction of the CP-Violating Angle ϕ_3

（B 中間子の稀崩壊 $B^- \to DK^-$, $D \to K^+\pi^-$ の初の証拠と CP 非保存角 ϕ_3 の測定）

Department of Physics, Tohoku University

Horii Yasuyuki

December 2010

Abstract

We report the first evidence for the suppressed decay chain $B^- \to DK^-$ followed by $D \to K^+\pi^-$, where D indicates a $\bar{D}^0$ or D^0 state, with a significance of 4.1σ. An interference of the two decay paths, corresponding to the two intermediate D states, provides an important information on the CP-violating angle ϕ_3. We measure the ratio $\mathcal{R}_{DK}$ of the suppressed decay rate to the favored decay rate, obtained in the chain $B^- \to DK^-$, $D \to K^-\pi^+$, and the asymmetry $\mathcal{A}_{DK}$ between the charge-conjugate decays of the suppressed mode as

$$\begin{aligned} \mathcal{R}_{DK} &= [1.63^{+0.44}_{-0.41}(\text{stat})^{+0.07}_{-0.13}(\text{syst})] \times 10^{-2}, \\ \mathcal{A}_{DK} &= -0.39^{+0.26}_{-0.28}(\text{stat})^{+0.04}_{-0.03}(\text{syst}). \end{aligned}$$

The results are based on a data sample recorded with the Belle detector at the KEKB e^+e^- storage ring, corresponding to 772×10^6 $B\bar{B}$ pairs produced at the $\Upsilon(4S)$ resonance.

A method based on Bayesian statistics including the observables for the chains $B^- \to DK^-$ followed by $D \to CP$-eigenstates provides

$$\phi_3 = 89^\circ\ {}^{+13^\circ}_{-16^\circ}$$

(modulo 180°), which is consistent with the expectation obtained from the Cabibbo-Kobayashi-Maskawa framework in the standard model. Our study strongly encourages the precise measurements of ϕ_3 at the next-generation B factories, which will play an important role for investigating new physics.

Acknowledgements

I would like to express my deep appreciation to my supervisor, Prof. H. Yamamoto, for his great guidance and supervision. His teachings are definitely essential for my understanding of quantum field theory, phenomenology of the B mesons, and accelerator-based experiments.

I am heartily thankful to an analysis coordinator of the Belle experiment, Dr. K. Trabelsi, for his bright ideas and suggestions for the analyses, the presentations, and the publications. This thesis would not have been possible without his supports.

I appreciate the members of Tohoku University. Dr. T. Sanuki, Dr. T. Nagamine, Dr. Y. Onuki, and Dr. Y. Takubo, for their valuable assistance and encouragement. Prof. A. Yamaguchi and Mr. N. Kikuchi, for their great introduction on particle physics. Mr. K. Kousai, Ms. Y. Minekawa, Mr. H. Tabata, Mr. K. Ito, Mr. T. Kusano, Mr. R. Sasaki, Mr. K. Itagaki, Mr. D. Okamoto, Mr. Y. Sato, Mr. Z. Suzuki, Mr. T. Honda, Mr. H. Nakano, Mr. K. Negishi, Mr. T. Saito, Mr. D. Kamai, Ms. E. Kato, Mr. H. Katsurayama, Ms. M. Kikuchi, and Mr. Y. Ono, as very nice colleagues.

I am grateful to the members of the Belle collaboration. Prof. Y. Sakai, for his vital guidance from the initial to the final level. Prof. T. E. Browder, Dr. Z. Dolezal, Dr. N. Muramatsu, Prof. G. Mohanty, Dr. A. Zupanc, Mr. M. Prim, Dr. M. Nakao, Dr. H. Ozaki, Dr. K. Sakai, Ms. M. Watanabe, Dr. N. Taniguchi, and Dr. M. Iwabuchi, for their great contributions to the analysis. All the other members, for keeping high quality of the detector and the software framework. I also appreciate the KEKB accelerator group, for the high luminosity provided by the excellent operation.

Last but not the least, I would like to express my gratitude to my parents, brother, and wife. The project would not have been completed without their moral support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

For a long time, the science has been taken as an empirical and experimental activity into the nature. On the other hand, the metaphysics is one of non-empirical inquiries and known as "natural philosophy." Recently, the development of philosophy has provided the idea that the science is in fact a part of metaphysics. This concept is based on the standpoint that the science intends to connect the scientific knowledge to the reality of the nature whereas this connection can not be confirmed even empirically or experimentally.

However, the resulting knowledge of the science has actually revolutionized our view of the world, and transformed our society. Even if we can not reach the truth, the development of science provides effective insights into the nature and fertilizes our lives.

The particle physics is a primary means of inquiry into the basic working of the nature. In this field, every phenomena are ruled by universal laws and principles whose natural realm is at scales of time and distance far removed from our direct experience. In the twentieth century, a standard model has been developed with various confirmations of many of its aspects. In this framework, the matters are constructed from quarks and leptons, and the strong and electroweak interactions are mediated by gauge bosons.

One of the great constitutions of the standard model is the Cabibbo-Kobayashi-Maskawa (CKM) picture [1, 2], which was proposed by M. Kobayashi and T. Maskawa in 1973 by extending the idea of the quark mixing by N. Cabibbo. This picture naturally explains the CP violation in K^0 decays observed in 1964 [3], by introducing the third generation of quarks. In 1977 and 1995, bottom and top quarks, known as third-generation quarks, are observed, respectively [4, 5, 6]. A conclusive experimental result is the CP violation in B^0 decays observed in 2001 [7, 8]. Up to now, plenty of experimental outputs agree with the CKM picture in high precisions.

Despite the great success, there are various reasons for the standard model not to be an ultimate theory. An example is the fact that there are several absences of explanations on such as dark matter, gravitation, etc. Nowadays, precise measurements of the CKM parameters are considered to be ones of methods for revealing the extended physics called “new physics.” Any significant discrepancy between the expected and measured values would be a signature of new physics.

The parameter ϕ_3 (also known as γ) is one of the worst-determined CKM parameters and further constraints are strongly demanded. Since the measurement of ϕ_3 is obtained theoretically cleanly from the tree-dominated decays, it provides a unique basis for searching for new physics.

In this thesis, we report the first evidence of the suppressed decay chain[1] $B^- \to DK^-$ followed by $D \to K^+\pi^-$, where D indicates a $\bar{D}^0$ or D^0 state. The two decay paths, corresponding to the two intermediate D states, interfere and provide important information on ϕ_3. We use a data sample, that contains 772×10^6 $B\bar{B}$ pairs, recorded at the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB e^+e^- collider. For discriminating the large backgrounds from $e^+e^- \to q\bar{q}$ $(q = u, d, s, c)$ processes, we employ a new method based on a neural network technique.

In Chapter 2, we introduce the Kobayashi-Maskawa (KM) mechanism and the methods for measuring ϕ_3. In Chapter 3, we describe the experimental apparatus and analysis tools. In Chapter 4, 5, and 6, we show the event selection, the discrimination of the $q\bar{q}$ backgrounds, and the signal extraction, respectively. In Chapter 7, we estimate the contribution from the backgrounds remaining after the rejections. In Chapter 8, we measure the observables relative to ϕ_3, and in Chapter 9, we extract ϕ_3.

---

[1] Charge-conjugate modes are implicitly included throughout the thesis unless otherwise stated.

Chapter 2

KM Mechanism and CP-Violating Angle ϕ_3

2.1 Interactions between Quarks and W Bosons

In the standard model, the masses of quarks arise from the Yukawa interactions with the Higgs fields. The Yukawa terms are described by left-handed quark doublet q'_L and right-handed down-type (up-type) quark singlet d'_R (u'_R) both in the weak-interaction basis, as

$$\mathcal{L}_{\text{Yukawa,quark}} = -y^d_{ij}(\bar{q}'_{Li}\phi)d'_{Rj} - y^u_{ij}(\bar{q}'_{Li}\epsilon\phi^*)u'_{Rj} + \text{h.c.}, \tag{2.1}$$

where y^d and y^u are complex constants, i and j are the generation labels, ϕ is the Higgs doublet, and ϵ is the 2×2 antisymmetric tensor. The mass terms are produced by introducing a spontaneous symmetry breaking, where the Higgs field acquires a vacuum expectation value $\langle\phi\rangle = (0, v/\sqrt{2})$, as

$$\mathcal{L}_{\text{mass,quark}} = -\frac{vy^d_{ij}}{\sqrt{2}}(\bar{d}'_{Li}d'_{Rj}) - \frac{vy^u_{ij}}{\sqrt{2}}(\bar{u}'_{Li}u'_{Rj}) + \text{h.c.} \tag{2.2}$$

The quark states can be transformed by unitary matrices so that the terms for $i \neq j$ drop out, leading the mass eigenstates.

The interactions of physical quarks and W bosons are then given by

$$\mathcal{L}_{\text{int},qW} = -\frac{g}{\sqrt{2}}\left[\left(\overline{U}_L\gamma^\mu \boldsymbol{V} D_L\right)W^+_\mu + \left(\overline{D}_L\gamma^\mu \boldsymbol{V}^\dagger U_L\right)W^-_\mu\right], \tag{2.3}$$

where g is a dimensionless constant, γ^μ is a Dirac matrix, U shows up-type quarks (u, c, t), D shows down-type quarks (d, s, b), $W^\pm_\mu$ are the W-boson fields, and the $\boldsymbol{V}$ is a unitary matrix connecting the mass eigenstates and the weak-interaction states. The matrix $\boldsymbol{V}$ is called as the CKM matrix [1, 2], and usually expressed as

$$\boldsymbol{V} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{2.4}$$

which parameterizes the mixing of quarks in the weak interactions with W bosons.

2.2 KM Mechanism

Under the CP transformation, the $\mathcal{L}_{\text{int},qW}$ is invariant only when $\boldsymbol{V} = \boldsymbol{V}^*$, since

$$(CP)\mathcal{L}_{\text{int},qW}(CP)^\dagger = -\frac{g}{\sqrt{2}}\left[\left(\overline{U}_L\gamma^\mu\boldsymbol{V}^*D_L\right)W_\mu^+ + \left(\overline{D}_L\gamma^\mu(\boldsymbol{V}^*)^\dagger U_L\right)W_\mu^-\right], \tag{2.5}$$

which is supported by $(CP)W_\mu^+(CP)^\dagger = -W_\mu^-$ and $(CP)\bar{\psi}_\text{a}\gamma^\mu\psi_\text{b}(CP)^\dagger = -\bar{\psi}_\text{b}\gamma^\mu\psi_\text{a}$ with $\psi_\text{a,b}$ to be general fermion states. Therefore, the CP violation occurs when one or more elements of $\boldsymbol{V}$ are complex numbers.

Assuming n generations of quarks, the n-by-n complex matrix $\boldsymbol{V}$ generally has $2n^2$ of degrees of freedom. Since $\boldsymbol{V}$ is unitary, n^2 of them are removed by the unitarity conditions:

$$\sum_{j=1}^{n} V_{ij}V_{kj}^* = \delta_{ik} \quad (i,k = 1,2,...,n). \tag{2.6}$$

Other sources for reducing the degrees of freedom are the phases of quarks, which are meaningless in the quantum mechanics. Excluding the overall phase, which does not change $\boldsymbol{V}$, $2n-1$ degrees of freedom consequently disappear. Then, the degrees of freedom in $\boldsymbol{V}$ are

$$2n^2 - n^2 - (2n-1) = (n-1)^2. \tag{2.7}$$

Considering the case without complex elements, the unitary condition of $\boldsymbol{V}$ is equivalent to the orthogonal conditions:

$$\sum_{j=1}^{n} V_{ij}V_{kj} = \delta_{ik} \quad (i \le k). \tag{2.8}$$

The maximum number of real elements in $\boldsymbol{V}$ is thus

$$n^2 - \frac{(n+1)n}{2} = \frac{n(n-1)}{2}. \tag{2.9}$$

Therefore, the number of possible complex components in $\boldsymbol{V}$ is obtained from Eq. (2.7) by subtracting Eq. (2.9) as

$$(n-1)^2 - \frac{n(n-1)}{2} = \frac{(n-1)(n-2)}{2}. \tag{2.10}$$

When $n = 3$, $\boldsymbol{V}$ has at least one complex element, which reveals the CP violation in the interactions of quarks and W bosons. M. Kobayashi and T. Maskawa won the Nobel prize in physics 2008 for the discovery of the origin of the CP violation which predicts the existence of at least three families of quarks in nature.

2.3 Unitarity Triangle and CP-Violating Angles

The CKM matrix $\boldsymbol{V}$ can be parameterized as an expansion on $\lambda = \sin\theta_{\mathrm{C}} \sim 0.2$ (θ_{C} is the Cabibbo angle),

$$\boldsymbol{V} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{2.11}$$

where A, η, and ρ are real parameters having magnitudes $\mathcal{O}(1)$ [9]. Of the unitary conditions in Eq. (2.6), six vanishing combinations indicate the triangles in complex planes. The most commonly used unitarity triangle (Figure 2.1) arises from

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{2.12}$$

for which the three terms have comparable magnitudes $\mathcal{O}(\lambda^3)$.

The angles of the unitarity triangle are defined as

$$\phi_1 \equiv \arg\left(\frac{V_{cd}V_{cb}^*}{-V_{td}V_{tb}^*}\right), \tag{2.13}$$

$$\phi_2 \equiv \arg\left(\frac{V_{td}V_{tb}^*}{-V_{ud}V_{ub}^*}\right), \tag{2.14}$$

$$\phi_3 \equiv \arg\left(\frac{V_{ud}V_{ub}^*}{-V_{cd}V_{cb}^*}\right), \tag{2.15}$$

where ϕ_1, ϕ_2, and ϕ_3 are also known as β, α, and γ, respectively. Since these phases are fundamental parameters of the standard model, the determinations are of primary importance, and any significant discrepancy between the expected and measured values would be a signature of new physics.

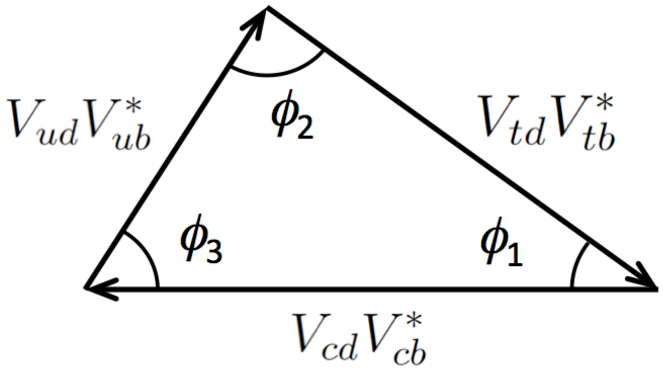


Figure 2.1: Unitarity triangle.

2.4 Methods for Measuring ϕ_3

In the usual quark phase convention in Eq. (2.11), large complex phases appear only in V_{ub} and V_{td}. Therefore, the angle ϕ_3 is expressed as

$$\phi_3 \sim -\arg(V_{ub}). \tag{2.16}$$

The measurement of ϕ_3 is equivalent to the extraction of the phase of V_{ub} relative to the phases of other CKM matrix elements except for V_{td}.

Several proposed methods for measuring ϕ_3 exploit the interference in the decay $B^- \to DK^-$ ($D = \bar{D}^0$ or D^0), which occurs when $\bar{D}^0$ and D^0 decay to common final state [10, 11, 12, 13]. Figure 2.2 shows the diagrams for $B^- \to DK^-$. The "tree" contributions enable theoretically clean determinations. The ϕ_3 is extracted with the magnitude of ratio of amplitudes r_B and the strong phase difference δ_B defined by

$$r_B = \left| \frac{A(B^- \to \bar{D}^0 K^-)}{A(B^- \to D^0 K^-)} \right|, \ \delta_B = \delta(B^- \to \bar{D}^0 K^-) - \delta(B^- \to D^0 K^-). \tag{2.17}$$

These are crucial parameters in the measurement of ϕ_3. The value of r_B is thought to be around 0.1, by taking a product of the ratio of the CKM matrix elements $|V_{ub}V_{cs}^*/V_{cb}V_{us}^*|$ and the color suppression factor.

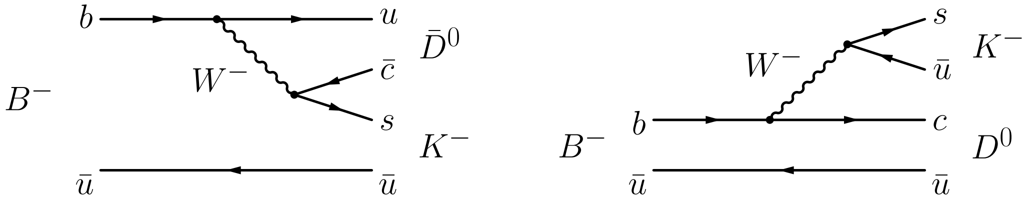


Figure 2.2: Diagrams for the $B^- \to \bar{D}^0 K^-$ and $B^- \to D^0 K^-$ decays, where the information of ϕ_3 is included in the transition $b \to u$.

The methods of measuring ϕ_3 using the decay $B^- \to DK^-$ can be categorized by the following D decays. In the sections below, we describe several techniques. Note that the D^0–$\bar{D}^0$ mixings can safely be neglected for current precision of ϕ_3, while it is possible to take them into account explicitly without a need of time-dependent analysis of the B decay [14].

2.4.1 GLW Method

The Gronau-London-Wyler (GLW) method to extract ϕ_3 use the D decays to the CP eigenstates [11]. In this method, the branching ratios $\mathcal{B}(B^- \to \bar{D}^0 K^-)$, $\mathcal{B}(B^- \to D^0 K^-)$, and $\mathcal{B}(B^- \to D_{CP+} K^-)$ are separately determined, where D_{CP+} is the CP eigenstate $D_{CP+} = (D^0 + \bar{D}^0)/\sqrt{2}$ and can be reconstructed from K^+K^-, $\pi^+\pi^-$, etc.

The phase difference between the amplitudes $A(B^- \to \bar{D}^0 K^-)$ and $A(B^- \to D^0 K^-)$ is $\delta_B - \phi_3$, while the one for B^+ decays is $\delta_B + \phi_3$. Thus, the phase $2\phi_3$ could be extracted by measuring three sides of the triangles in the complex plane as shown in Figure 2.3.

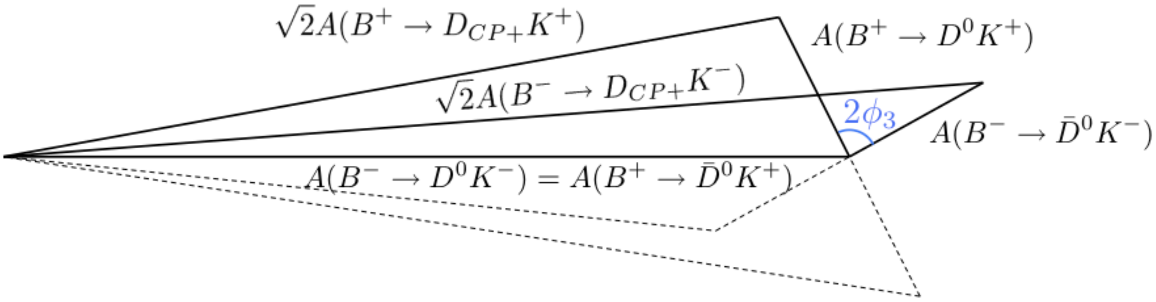


Figure 2.3: Complex triangles of amplitudes in the GLW method.

Unfortunately, a problem arises from the small magnitude of $A(B^- \to \bar{D}^0 K^-)$ expected by small r_B. Possible two ways of tagging $\bar{D}^0$ flavor, through semi-leptonic decays and through hadronic decays, are both likely to be impractical. However, the availability of using $\mathcal{B}(B^- \to D_{CP-}K^-)$, where $D_{CP-} = (D^0 - \bar{D}^0)/\sqrt{2}$ can be reconstructed from $K_S\pi^0$, $K_S\phi$, etc., overcomes the problem (dashed lines in Figure 2.3). The additional constraints enable the numerical extraction of ϕ_3 rather than geometrical method. Then, the observables

$$\begin{aligned}
\mathcal{R}_{CP\pm} &\equiv \frac{\mathcal{B}(B^- \to D_{CP\pm}K^-) + \mathcal{B}(B^+ \to D_{CP\pm}K^+)}{\mathcal{B}(B^- \to D^0K^-) + \mathcal{B}(B^+ \to \bar{D}^0K^+)} \\
&= 1 + r_B^2 \pm 2r_B \cos\delta_B \cos\phi_3, \qquad (2.18) \\
\mathcal{A}_{CP\pm} &\equiv \frac{\mathcal{B}(B^- \to D_{CP\pm}K^-) - \mathcal{B}(B^+ \to D_{CP\pm}K^+)}{\mathcal{B}(B^- \to D_{CP\pm}K^-) + \mathcal{B}(B^+ \to D_{CP\pm}K^+)} \\
&= \pm 2r_B \sin\delta_B \sin\phi_3 / R_{CP\pm}, \qquad (2.19)
\end{aligned}$$

are measured and taken as simultaneous equations for solving for ϕ_3. Possible four equations could constrain the three unknowns r_B, δ_B and ϕ_3.

Note that the effects of the CP violation to the observables are limited because of small r_B, and thus precise measurements of the observables are needed for obtaining effective constraint on ϕ_3. Also, the observables in Eqs. (2.18) and (2.19) are invariant under the three transformations

$$\begin{array}{lll}
\text{(i)} & \phi_3 \to \delta_B, & \delta_B \to \phi_3, \\
\text{(ii)} & \phi_3 \to -\phi_3, & \delta_B \to -\delta_B, \\
\text{(iii)} & \phi_3 \to \phi_3 + \pi, & \delta_B \to \delta_B + \pi,
\end{array}$$

and the ϕ_3 and δ_B have eight solutions in the region 0°–360°.

2.4.2 ADS Method

The effects of the CP violation could be enhanced if the final state is chosen so that the interfering amplitudes have comparable magnitudes [12]. In the Atwood-Dunietz-Soni (ADS) method, Cabibbo-favored $\bar{D}^0$ decay and doubly Cabibbo-suppressed D^0 decay are chosen as the following D decays to $B^- \to DK^-$. Denoting the final state of the D decay as f, the comparable magnitudes of $A(B^- \to [f]_{\bar{D}^0}K^-)$ and $A(B^- \to [f]_{D^0}K^-)$ provide a relatively large interfering effect on $A(B^- \to [f]_D K^-)$ as shown in Figure 2.4. Therefore, significant information of ϕ_3 is included in the branching ratio $\mathcal{B}(B^- \to [f]_D K^-)$.

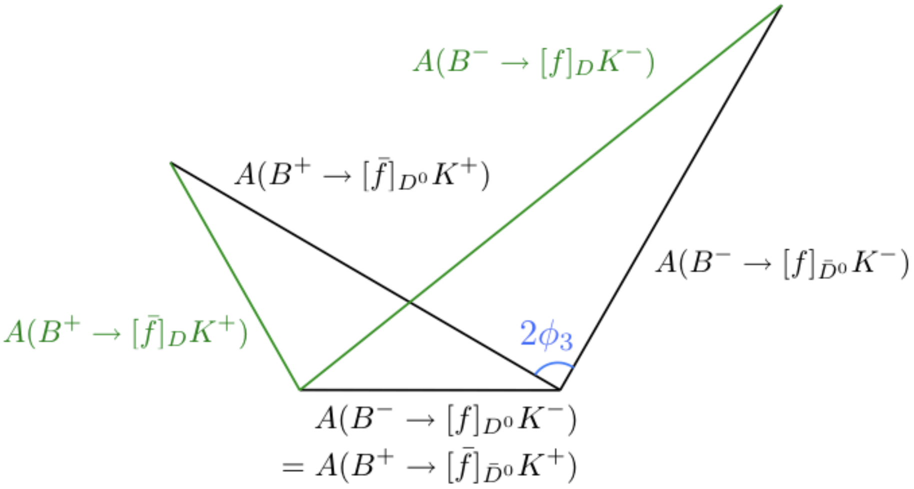


Figure 2.4: Complex triangles of amplitudes in the ADS method.

The usual observables are described as

$$\begin{aligned}
\mathcal{R}_{\mathrm{ADS}} &\equiv \frac{\mathcal{B}(B^- \to [f]_D K^-) + \mathcal{B}(B^+ \to [\bar{f}]_D K^+)}{\mathcal{B}(B^- \to [\bar{f}]_D K^-) + \mathcal{B}(B^+ \to [f]_D K^+)} \\
&= r_B^2 + r_D^2 + 2r_B r_D \cos(\delta_B + \delta_D)\cos\phi_3, \qquad (2.20) \\
\mathcal{A}_{\mathrm{ADS}} &\equiv \frac{\mathcal{B}(B^- \to [f]_D K^-) - \mathcal{B}(B^+ \to [\bar{f}]_D K^+)}{\mathcal{B}(B^- \to [f]_D K^-) + \mathcal{B}(B^+ \to [\bar{f}]_D K^+)} \\
&= 2r_B r_D \sin(\delta_B + \delta_D)\sin\phi_3 / R_{\mathrm{ADS}}, \qquad (2.21)
\end{aligned}$$

where $r_D = |A(D^0 \to f)/A(\bar{D}^0 \to f)|$ and $\delta_D = \delta(\bar{D}^0 \to f) - \delta(D^0 \to f)$. By measuring the observables for two final states for which r_D and δ_D are known, sufficient information are obtained to solve for ϕ_3. The most important final state f is $K^+\pi^-$, while other possibilities are $K^+\rho^-$, $K^+a_1^-$, and $K^{*+}\pi^-$. Important note is that the observables in the GLW method, Eqs. (2.18) and (2.19), are possible additional inputs; the combined approach is sometimes called as the GLW+ADS method.

The value of δ_D for $f = K^+\pi^-$ is measured with no fold ambiguity by a fit including several experimental results relative to the D meson decays [15]. Therefore, the number of solutions of ϕ_3 in 0°–360° when we include the observables for $f = K^+\pi^-$ could be reduced to two on

$$\phi_3 \rightarrow \phi_3 + \pi,\ \delta_B \rightarrow \delta_B + \pi,\ \delta_D \rightarrow \delta_D.$$

2.4.3 Dalitz Method

The above methods can be generalized to the cases for three-body D decays [13]. The CP-violating effects appear in the interfering amplitudes in the Dalitz plane. Denoting the particles coming from the D decay d_1, d_2, and d_3, the amplitude of $B^- \rightarrow [d_1 d_2 d_3]_D K^-$ is proportional to

$$M_- = f(s_{12}, s_{13}) + r_B e^{i(\delta_B - \phi_3)} \bar{f}(s_{12}, s_{13}), \tag{2.22}$$

where s_{ij} are defined as Dalitz plot variables $s_{ij} \equiv (p_i + p_j)^2$ with the momenta p_1, p_2, and p_3 for d_1, d_2, and d_3, respectively, and $f(s_{12}, s_{13})$ and $\bar{f}(s_{12}, s_{13})$ are the amplitudes of the decays $D^0 \rightarrow d_1 d_2 d_3$ and $\bar{D}^0 \rightarrow d_1 d_2 d_3$, respectively. The value of ϕ_3 is extracted by a fit on s_{12} and s_{13} or on the Dalitz plot. The method is sensitive to ϕ_3 in parts of the Dalitz plot including the resonances. The most important D decay is the Cabibbo-favored $D \rightarrow K_S \pi^+ \pi^-$, which has $D \rightarrow K_S \rho^0$, $D \rightarrow K^{*\pm}(892)\pi^{\mp}$, and higher resonances. The decay $D \rightarrow K_S K^+ K^-$ is also an effective mode.

The magnitudes of $f(s_{12}, s_{13})$ and $\bar{f}(s_{12}, s_{13})$ can be determined from the Dalitz plot of the flavor specific D decays, where the flavor is tagged by the charge of the soft pion in $D^{*-} \rightarrow \bar{D}^0 \pi^-$. However, the phases are not measurable without further model-dependent assumptions. A possibility for solving the problem is to assume the Breit-Wigner functions for the resonances, and take a sum with non-resonant term. The phases are obtained simultaneously with the magnitudes by fitting the Dalitz plot of the D decays. However, the model assumption causes nontrivial uncertainty on ϕ_3.

A model-independent approach is to obtain the phases of D-decay amplitudes simultaneously with ϕ_3, r_B, and δ_B by the fits on binned regions of the Dalitz plot. Enough constraints to solve for ϕ_3, r_B, δ_B, and D-decay parameters should be obtained by taking sufficiently large number of binned regions. Also, larger number of bins is preferred to avoid that the interference terms average out by integrating in the regions. Since smaller number of bins is primarily favored to obtain more statistics for each fit, the number of bins should be determined in the trade-off of the above situations.

Another possibility of model-independent method is to obtain the phases of D-decay amplitudes independently at charm factories by using the $\psi(3770)$ resonance. The particle $\psi(3770)$ decays into a pair of D mesons. By measuring the decay rate of this coherent state, the phases in the interfering amplitudes can be extracted.

2.5 Measurements of ϕ_3 as of Summer 2010

2.5.1 Result of Dalitz Method on $B^- \to [K_S\pi^+\pi^-]_D K^-$

The measurement of ϕ_3 as of summer 2010 is dominated by the Dalitz method on the decay $B^- \to [K_S\pi^+\pi^-]_D K^-$ [16, 17]. The most precise measurement on this mode is obtained by the Belle collaboration using a data sample that contains 657×10^6 $B\bar{B}$ pairs. The decay $B^- \to D^*K^-$ is also analyzed similarly by reconstructing the D^* from $D\pi^0$ or $D\gamma$ to improve the sensitivity, where the parameters r_B^* and δ_B^* are introduced. The amplitudes $f(s_{12}, s_{13})$ and $\bar{f}(s_{12}, s_{13})$ are obtained by a large sample of $D \to K_S\pi^+\pi^-$ decays produced in continuum e^+e^- annihilation, where the isobar model is assumed with Breit-Wigner functions for resonances. By obtaining the background fractions in several kinematic variables, the fit on Dalitz plot is performed with the parameters $x_\pm = r_\pm \cos(\delta_B \pm \phi_3)$ and $y_\pm = r_\pm \sin(\delta_B \pm \phi_3)$, where the r_B is separated for $B^\pm$ as $r_\pm$. The results are shown in Figure 2.5 for $B^- \to DK^-$ and $B^- \to D^*K^-$. The separations with respect to the charges of $B^\pm$ indicate an evidence of the CP violation. From the results of the fits, the value of ϕ_3 is measured to be

$$\phi_3 = 78.4^\circ\ {}^{+10.8^\circ}_{-11.6^\circ}(\text{stat}) \pm 3.6^\circ(\text{syst}) \pm 8.9^\circ(\text{model}), \tag{2.23}$$

where $r_B = 0.161^{+0.040}_{-0.038} \pm 0.011^{+0.050}_{-0.010}$, $r_B^* = 0.196^{+0.073}_{-0.072} \pm 0.013^{+0.062}_{-0.012}$, $\delta_B = 137.4^\circ\ {}^{+13.0^\circ}_{-15.7^\circ} \pm 4.0^\circ \pm 22.9^\circ$, and $\delta_B^* = 341.7^\circ\ {}^{+18.6^\circ}_{-20.9^\circ} \pm 3.2^\circ \pm 22.9^\circ$. The model error is due to the uncertainty in determining $f(s_{12}, s_{13})$ and $\bar{f}(s_{12}, s_{13})$. Model-independent measurements are strongly demanded for further constraints on ϕ_3.

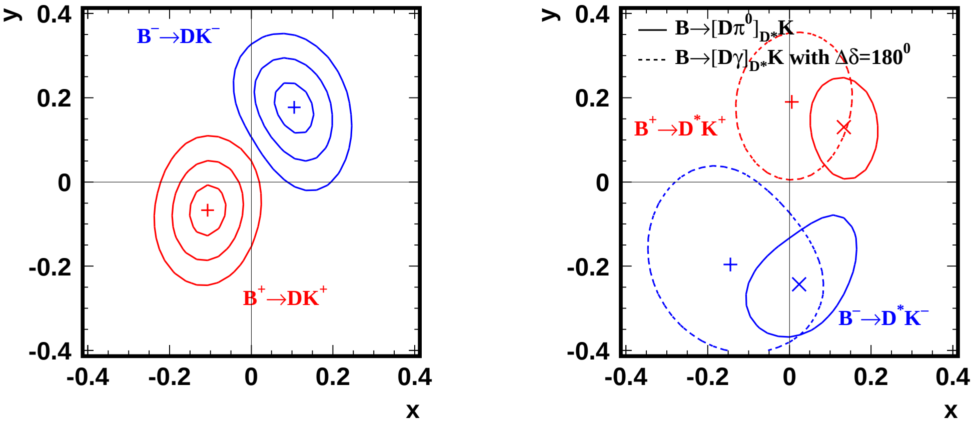


Figure 2.5: Results of the fits for $B^- \to DK^-$ (left) and $B^- \to D^*K^-$ (right) samples on the Dalitz analyses for $D \to K_S\pi^+\pi^-$, where the contours indicate 1, 2, and 3 (left) and 1 (right) standard-deviation regions.

2.5.2 Comparison of Constraints on CKM Parameters

The CKMfitter group provide a global analysis of determining the CKM parameters [18]. The result as of summer 2010 is shown on $\bar{\rho}$, $\bar{\eta}$ plane in Figure 2.6, where $\bar{\rho} + i\bar{\eta} = -(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)$; $\bar{\rho} = \rho(1 - \lambda^2/2\ ...)$ and $\bar{\eta} = \eta(1 - \lambda^2/2\ ...)$. All the measurements are consistent with each other.

The numerical results on direct measurements of the CKM angles are given as

$$\phi_1 = 21.15^\circ\ {}^{+0.90^\circ}_{-0.88^\circ}, \tag{2.24}$$

$$\phi_2 = 89.0^\circ\ {}^{+4.4^\circ}_{-4.2^\circ}, \tag{2.25}$$

$$\phi_3 = 71^\circ\ {}^{+21^\circ}_{-25^\circ}. \tag{2.26}$$

The error of ϕ_3 is larger than the one in Eq. (2.23), an important reason of which is the resulting smaller value of $r_B = 0.103^{+0.015}_{-0.024}$ enhancing generally the uncertainty on ϕ_3. An indirect measurement of ϕ_3, or an expectation of ϕ_3 from other measurements assuming the CKM picture, gives $\phi_3 = 67.2^\circ \pm 3.9^\circ$, which is consistent with the direct measurement. High-precision direct measurements of ϕ_3 are strongly demanded to test further the CKM picture. Any significant discrepancy between the measurements of the parameters would be a signature of new physics.

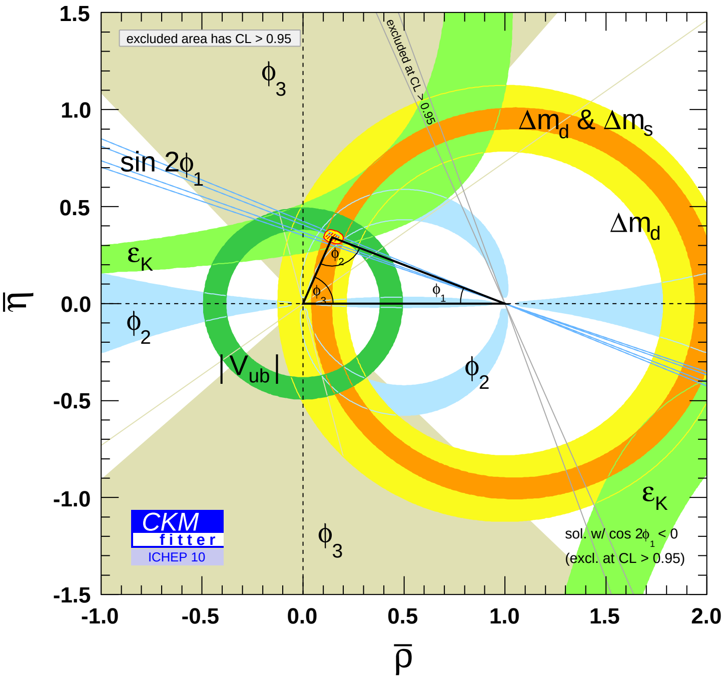


Figure 2.6: Unitarity triangle and constraints on it from various measurements, as of summer of 2010, provided by the CKMfitter group.

2.6 Motivation to Analyze $B^- \to [K^+\pi^-]_D K^-$

As shown in the previous section, it is fundamentally important to improve the sensitivity for ϕ_3 in a model-independent manner. A reliable possibility is to apply the GLW+ADS method, while another one is to apply the Dalitz method in a model-independent way by using the binned analysis or the $\psi(3770)$ information. Of the possible two ways, the subject in this thesis is the former.

The decay $B^- \to [K^+\pi^-]_D K^-$ (Figure 2.7) plays a quite important role in a model-independent measurement of ϕ_3 via the GLW+ADS method. The observables, denoted as $\mathcal{R}_{DK}$ and $\mathcal{A}_{DK}$, are obtained from Eqs. (2.20) and (2.21) with $f = K^+\pi^-$ as

$$\begin{aligned}
\mathcal{R}_{DK} &\equiv \frac{\mathcal{B}(B^- \to [K^+\pi^-]_D K^-) + \mathcal{B}(B^+ \to [K^-\pi^+]_D K^+)}{\mathcal{B}(B^- \to [K^-\pi^+]_D K^-) + \mathcal{B}(B^+ \to [K^+\pi^-]_D K^+)} \\
&= r_B^2 + r_D^2 + 2 r_B r_D \cos(\delta_B + \delta_D) \cos\phi_3, \qquad (2.27) \\
\mathcal{A}_{DK} &\equiv \frac{\mathcal{B}(B^- \to [K^+\pi^-]_D K^-) - \mathcal{B}(B^+ \to [K^-\pi^+]_D K^+)}{\mathcal{B}(B^- \to [K^+\pi^-]_D K^-) + \mathcal{B}(B^+ \to [K^-\pi^+]_D K^+)} \\
&= 2 r_B r_D \sin(\delta_B + \delta_D) \sin\phi_3 / R_{DK}. \qquad (2.28)
\end{aligned}$$

For the parameters r_D and δ_D, experimental inputs can be used [15]. The comparable magnitudes of $r_B \sim 0.1$ and $r_D = 0.058 \pm 0.001$ enhance the relative effects of the terms including ϕ_3 in both $\mathcal{R}_{DK}$ and $\mathcal{A}_{DK}$. Previous studies of this decay mode have not found a significant signal yield [19, 20]. It is very important to find an evidence of the signal, and measure the observables as precisely as possible.

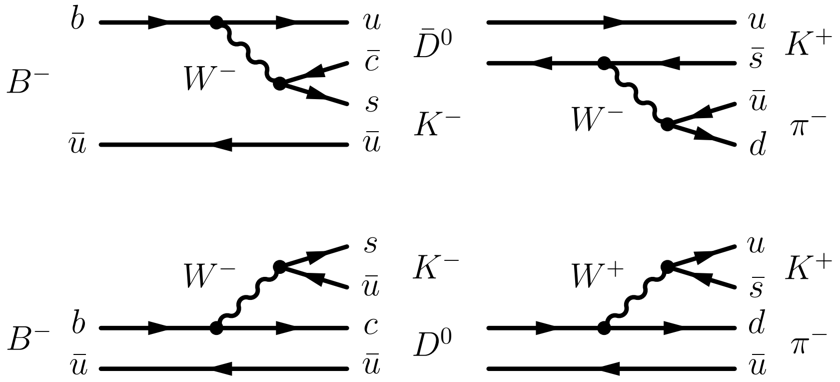


Figure 2.7: Diagrams for the decay $B^- \to [K^+\pi^-]_D K^-$. The color-suppressed B decay followed by the Cabibbo-favored D decay interferes with the color-favored B decay followed by the doubly Cabibbo-suppressed D decay. Comparable magnitudes of the amplitudes enhance the effects of CP violation.

Chapter 3

Experimental Apparatus and Analysis Tools

The data set is collected with the Belle detector at the KEKB accelerator located at the High Energy Accelerator Research Organization (KEK) in Japan. In this chapter, we describe the experimental apparatus and analysis tools. All the developments and the constructions shown in this chapter have been done by the collaborators. The author has only contributed to the detector operation and the management of the database.

3.1 KEKB Accelerator

Figure 3.1 shows a schematic view of the KEKB accelerator [21]. The electrons emitted from the thermionic gun are collected to make the bunches, parts of which are injected into a tungsten target causing the positron bunches. After accelerated, the electron and positron beams are transferred to the separated storage rings named High Energy Ring (HER) and Low Energy Ring (LER), respectively. The radio-frequency cavities accelerate the particles, while the dipole (quadrupole) magnets bend (focus) them. The beams are collided at the interaction point (IP), making variety of interesting reactions.

Thanks to the continuous efforts by the KEKB accelerator group, many advantageous features are realized. A steady operation is performed with the current of more than one ampere and the bunch size of $\mathcal{O}(100)$ μm for the horizontal and $\mathcal{O}(1)$ μm for the vertical. To avoid parasitic collisions and reduce beam-related backgrounds, the beam lines have a crossing angle of 22 mrad. A continuous injection mode, to add particles to the bunches without significant dead time at the detector, had been introduced in early 2004. A special superconducting radio-frequency cavities ("crab" cavities), which kick the beams in the horizontal plane, had been installed to make the head-on collisions while retaining the crossing angle of beams ("crab" crossing) in February 2007. By controlling the off-energy particles using skew sextupole magnets, world's highest luminosity[1] of $\mathcal{L} = 2.11 \times 10^{34}$ $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ had been achieved in June 2009. The total integrated luminosity had reached 1000 fb^{-1}, which is one of the primary targets of the KEKB project, by finishing the data taking in June 2010. Figure 3.2 shows the history of the integrated luminosity.

Most of the data are taken with the e^+e^- center-of-mass (CM) energy that corresponds to the Υ resonances from $e^+e^- \to b\bar{b}$ (Figure 3.3), while several parts are taken with sideband energy for the studies of non-resonant components mainly from $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$). The integrated luminosity of 711 fb^{-1} is recorded at the $\Upsilon(4S)$ resonance with the energy 8.0 GeV and 3.5 GeV for HER and LER, respectively, where 772×10^6 $B\bar{B}$ pairs are produced by the decays of the $\Upsilon(4S)$ mesons.

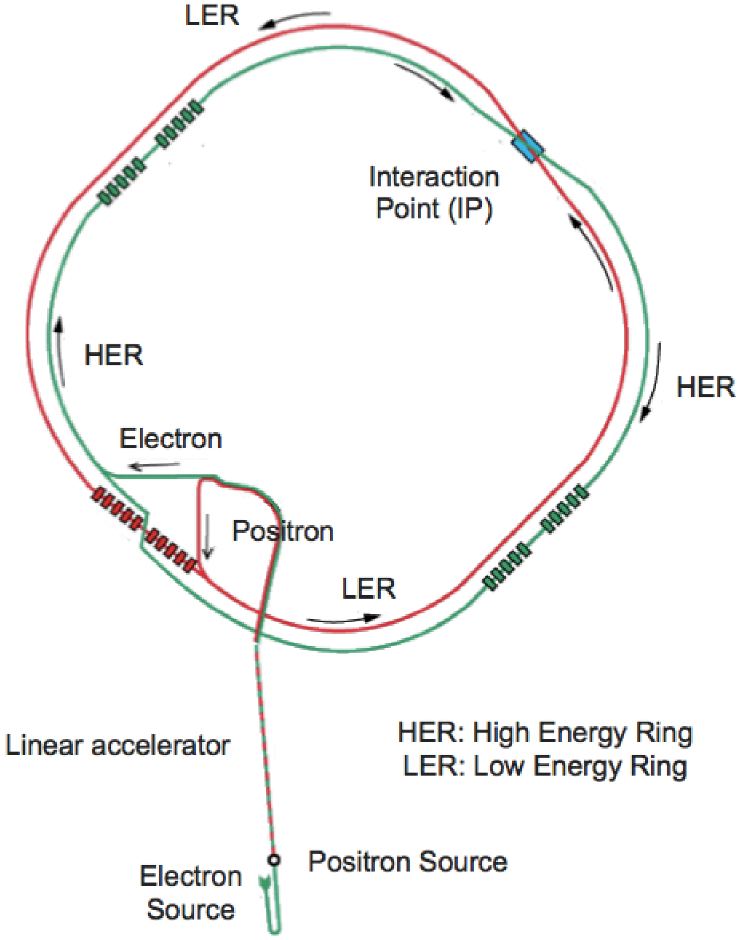


Figure 3.1: KEKB accelerator. The green and red boxes show the radio-frequency cavities for supplying missing energy caused by the radiations.

---

[1] The luminosity is described as $\mathcal{L} = N_+N_-f/4\pi\sigma_x^*\sigma_y^*$, where $N_\pm$ is the number of particles $e^\pm$ per bunch, f is the frequency of the collision, and $\sigma_{x,y}^*$ is the beam size at IP in x or y direction.

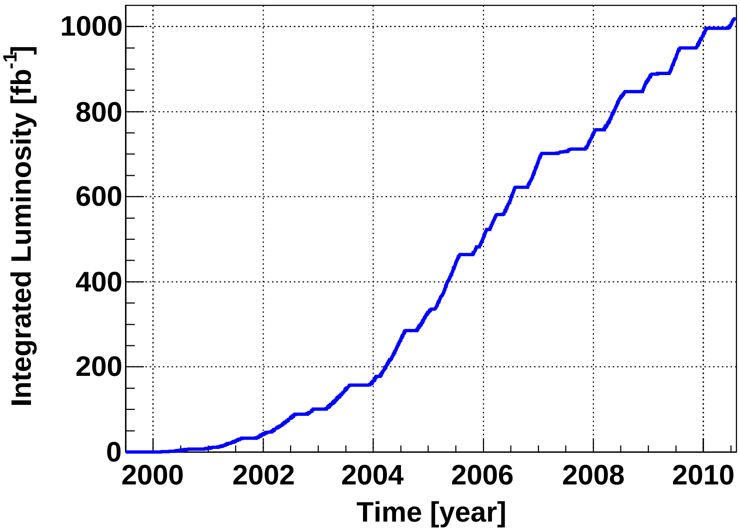


Figure 3.2: History of the integrated luminosity.

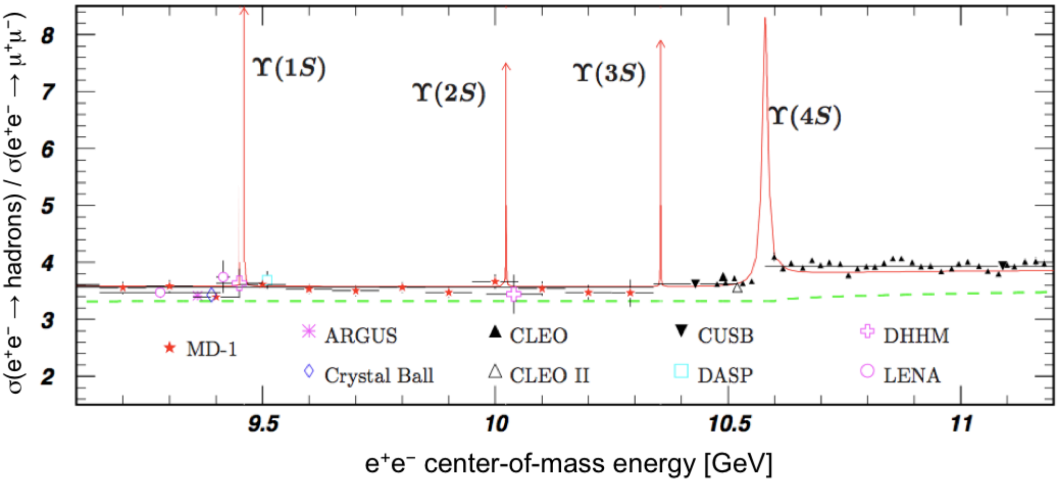


Figure 3.3: Cross section for $e^+e^- \to$ hadrons normalized by $e^+e^- \to \mu^+\mu^-$ around the CM energy to be 10 GeV, where the figure is taken from Ref. [22]. The Υ resonances from $e^+e^- \to b\bar{b}$ are placed on the continuum processes mainly from $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$).

3.2 Belle Detector

Belle detector [23] is a general-purpose detector consists of a barrel, forward, and backward components. Figure 3.4 shows the configuration of the Belle detector. A silicon vertex detector (SVD) and a central drift chamber (CDC) provide precision tracking and vertex measurements. The dE/dx measurement by CDC and the information from an array of aerogel-threshold Cherenkov counters (ACC) and a time-of-flight scintillation counters (TOF) enable the identification of charged pions and kaons. An electromagnetic calorimeter (ECL) composed of CsI(Tl) crystals detect the electromagnetic particles. The above detectors are located inside a superconducting solenoid coil that maintains 1.5 T magnetic field. The outermost subsystem is a K_L and muon detector (KLM).

Two inner detector configurations were used. A 2.0 cm beam pipe and a 3-layer silicon vertex detector were used up to summer of 2003 for the first sample of $152 \times 10^6 B\bar{B}$ pairs, while a 1.5 cm beam pipe, a 4-layer silicon detector, and a small-cell inner drift chamber were used from summer of 2003 to record the remaining $620 \times 10^6 B\bar{B}$ pairs. Almost all descriptions including figures in this section are based on Ref. [23].

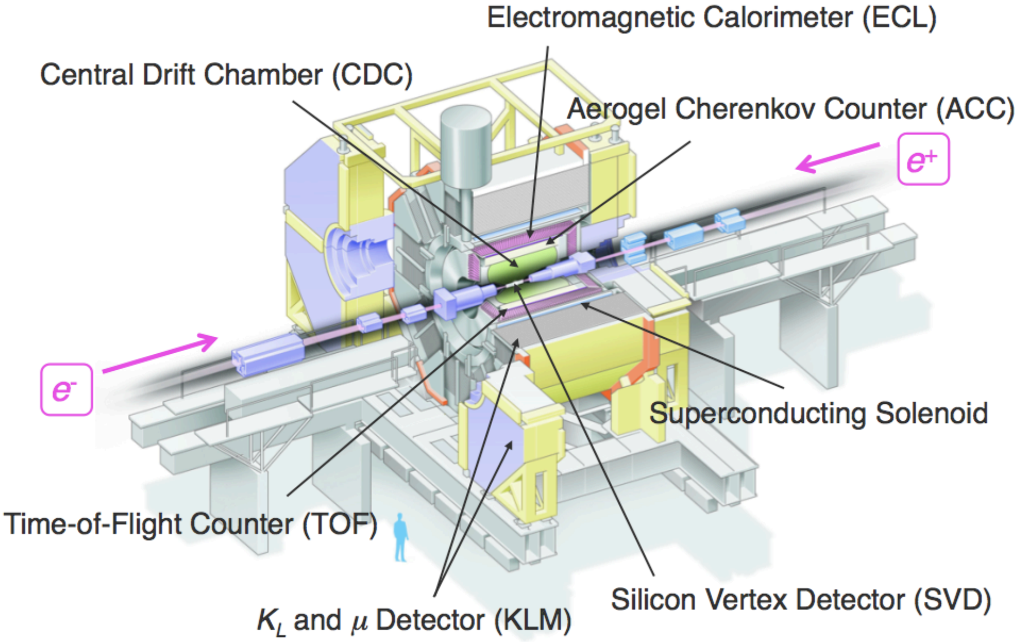


Figure 3.4: Overview of the Belle detector.

3.2.1 Silicon Vertex Detector (SVD)

Silicon detectors are based on the p-n junction diodes operated at reverse bias. The ionization currents caused by particles passing through the depleted region are detected and measured. We use double-sided silicon strip detectors (DSSDs) manufactured by Hamamatsu Photonics.[2] One side, called n-side, of the bulk has highly doped strips oriented perpendicular to the beam direction to measure the z coordinate, where z is defined as the opposite of the positron beam direction. The other side, called p-side, has longitudinal strips to measure the ϕ coordinate, where ϕ is the azimuthal angle around the z axis. The DSSDs reinforced by boron-nitride support ribs constitute a ladder to cover large range of polar angle θ from z axis, and the ladders construct a layer to cover all angle of ϕ. Multiple layers enable the tracking and vertex measurements. The readout chain of DSSDs is based on CMOS-integrated circuit placed outside of the tracking volume. The design is constrained to reduce multiple Coulomb scattering.

Figure 3.5 shows the side and end views of the SVD, SVD1, used up to the summer of 2003. There are three layers, where 8, 10, and 14 ladders are comprised in the inner, middle, and outer layers, respectively. Covered angle range is $23^\circ < \theta < 139^\circ$, which corresponds to 86% of the full solid angle. The readout electronics are based on the VA1 integrated circuits, which have excellent noise performance and reasonably good radiation tolerance of 0.2–1 Mrad depending on the version. The impact parameter resolutions $\sigma_{r\phi}$ and σ_z, where r is the distance from z axis, are measured using cosmic rays. Obtained values are

$$\sigma_{r\phi} = \sqrt{19.2^2 + \left(\frac{54.0}{p\beta \sin^{3/2}\theta}\right)^2}\ \mu\mathrm{m},\ \ \sigma_z = \sqrt{42.2^2 + \left(\frac{44.3}{p\beta \sin^{5/2}\theta}\right)^2}\ \mu\mathrm{m}, \tag{3.1}$$

where p and β are the momentum in GeV/c and the velocity divided by c of the particle, respectively.

A new vertex detector, SVD2, was installed in summer of 2003 [24]. Figure 3.6 shows the configuration of SVD2. There are four layers, where 6, 12, 18, and 18 ladders are comprised in the first, second, third, and fourth layers, respectively. It has larger coverage of $17^\circ < \theta < 150^\circ$, which corresponds to 92% of the full solid angle. The readout electronics are based on the VA1TA integrated circuits, which have excellent radiation tolerance of more than 20 Mrad. The impact parameter resolutions are

$$\sigma_{r\phi} = \sqrt{21.9^2 + \left(\frac{35.5}{p\beta \sin^{3/2}\theta}\right)^2}\ \mu\mathrm{m},\ \ \sigma_z = \sqrt{27.8^2 + \left(\frac{31.9}{p\beta \sin^{5/2}\theta}\right)^2}\ \mu\mathrm{m}. \tag{3.2}$$

The performance is slightly better than the one of SVD1, mainly owing to the smaller radius of the first layer.

---

[2] http://www.hamamatsu.com/

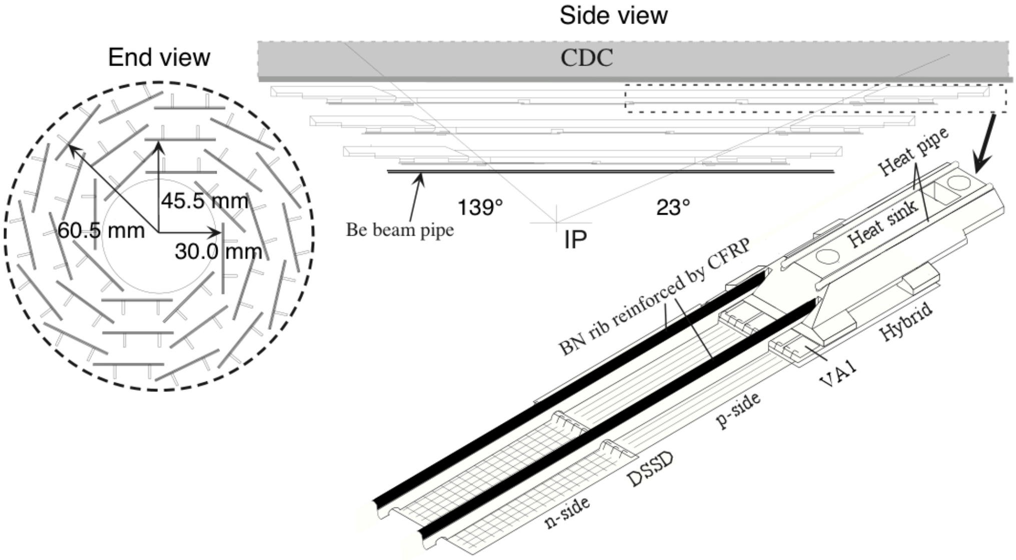


Figure 3.5: Configuration of SVD1.

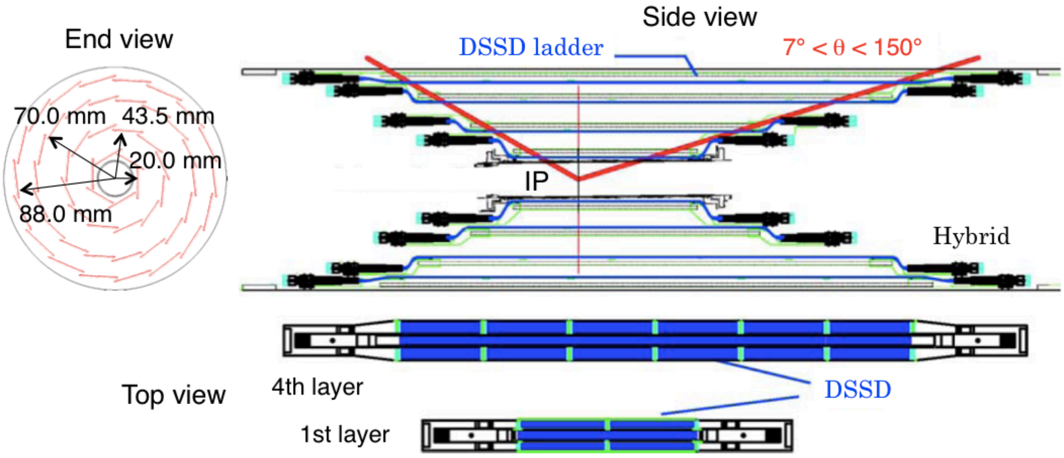


Figure 3.6: Configuration of SVD2.

3.2.2 Central Drift Chamber (CDC)

Gas detectors localize the ionization generated by charged particles. In a drift chamber, electrons released in the gas volume drift towards the anodes, producing avalanches in the increasing field. The drift time of the electrons can be used to measure the longitudinal position of a trajectory. Multiple wires enable the reconstruction of a track, the curvature of which provides the momentum measurement. Additional important information is the energy deposit dE/dx usable for the particle identification.

The overview of the CDC structure up to the summer of 2003 is shown in Figure 3.7. The angular coverage of $17° < \theta < 150°$ is provided by an asymmetric configuration in z direction. The chamber has 50 cylindrical layers of anode wires and 8400 drift cells. There are 32 wire layers parallel to the z axis ("axial wires") for the r–ϕ measurement and 18 wire layers slanted off the z axis ("stereo wires") for the z measurement. The most inner part includes three cathode layers for higher precision of low-momentum particles, which is beneficial for the trigger. In summer of 2003, the inner three layers have been replaced by two small-cell layers for making a space of SVD2, maintaining the performance of the trigger. The sense wires are gold-plated tungsten with diameters of 30 μm to maximize the electric drift field, while the field wires are made of unplated aluminum to reduce the material. A low-Z gas mixture of 50% helium and 50% ethane is employed to minimize the multiple-Coulomb scattering, where the ethane component permits a good dE/dx resolution.

The spacial resolution for particles passing near the center of the drift space is about 100 μm. The resolution of transverse momentum p_T in GeV/c is measured using cosmic rays as

$$\frac{\sigma_{p_T}}{p_T}(\%) = \sqrt{(0.20p_T)^2 + (0.29/\beta)^2}, \tag{3.3}$$

for which the p_T-dependent distribution is shown in Figure 3.8 (a). No apparent systematic effects due to the particle charge are observed. A truncated-mean method is employed to estimate the most probable dE/dx, where the largest values are discarded to minimize occasional fluctuations of the Landau tail. The resolution of dE/dx is measured to be 7.8% for the momentum range from 0.4 GeV/c to 0.6 GeV/c. Figure 3.8 (b) shows a scatter plot on dE/dx and momentum obtained using collision data, as well as the expected mean for different particles. Clear separations between pions, kaons, and protons are obtained for a momentum range up to around 1 GeV/c.

3.2.3 Aerogel Cherenkov Counter (ACC)

Cherenkov radiation is emitted by a charged particle when the velocity is greater than the speed of propagation of light in a material medium. Denoting the refractive index of the matter n, and the velocity, mass, and momentum of the particle β, m, and p, respectively, the condition for producing the radiation is given as

$$n > \frac{1}{\beta} = \sqrt{1 + \left(\frac{m}{p}\right)^2}. \tag{3.4}$$

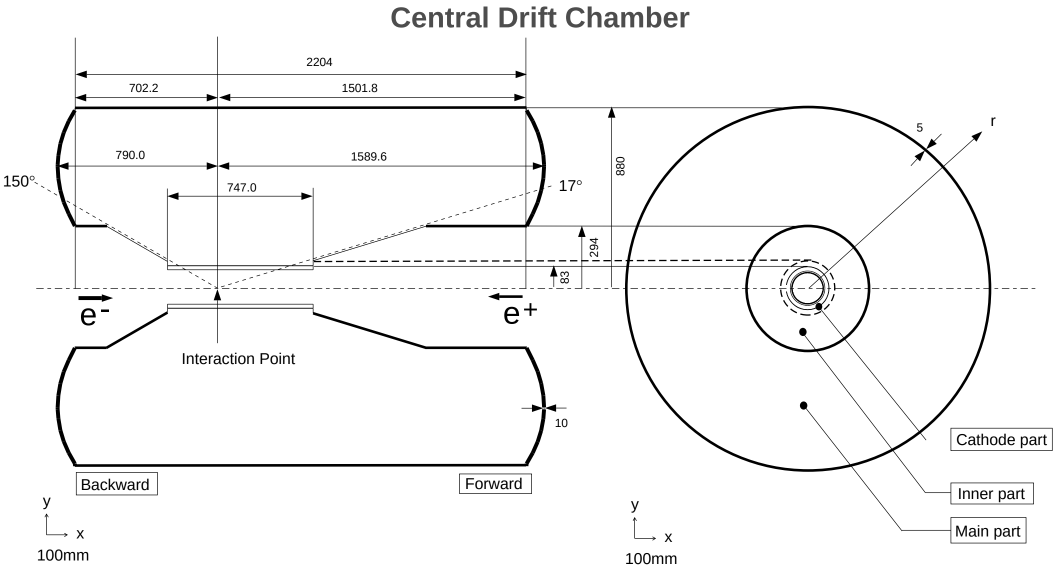
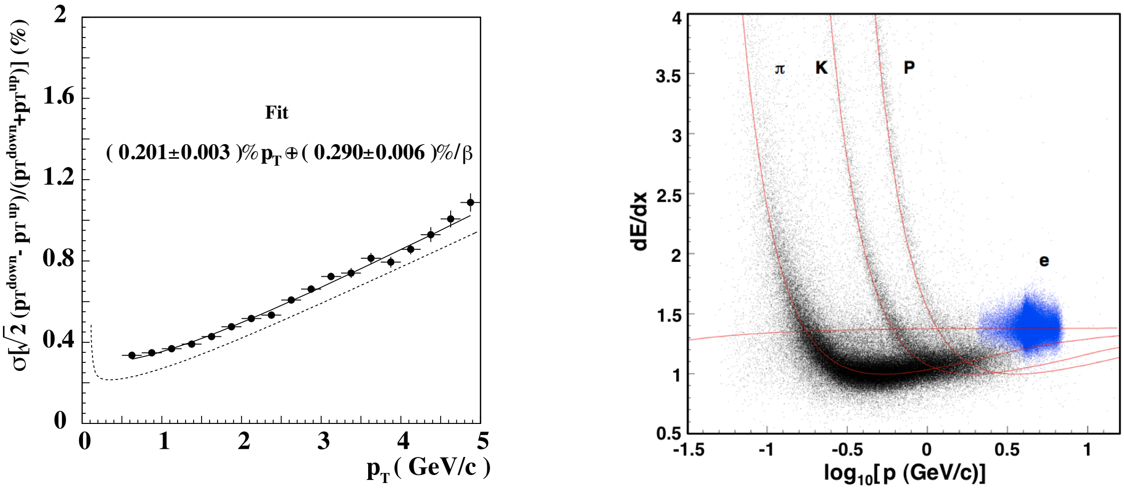


Figure 3.7: Overview of the CDC structure. The lengths are in units of mm.



(a) The p_t resolution for cosmic rays. The fitted result (solid curve) and the ideal expectation for $\beta = 1$ particles (dotted curve) are shown.

(b) Plot on dE/dx and momenta of particles obtained from collision data. Separations between pions, kaons, protons, and electrons are seen.

Figure 3.8: The performances of CDC.

There is a momentum region where a particle emits Cherenkov light while another heavier particle does not. Thus, the Cherenkov detector can identify charged particles having different masses by choosing the index n for the interested momentum region.

The identifications of $K^{\pm}$ and $\pi^{\pm}$ are very important for studying the B meson decays. The ACC is constructed to compensate the momentum range of 1.5–3.5 GeV$/c$, which is not covered by the dE/dx measurement by CDC and time-of-flight information from TOF. We use aerogels with the refractive indices from 1.01 to 1.03 depending on the polar-angle region. A highly hydrophobic aerogel was developed and produced to keep the transparency. The fine-mesh photomultiplier tubes (FM PMTs) are attached to the aerogels, since the ACC is operated in a high magnetic field. Figure 3.9 shows the configuration of ACC. It consists of 960 counter modules segmented into 60 cells for the barrel part and 228 modules arranged in 5 concentric layers for the end-cap part. All the counters are arranged in a semi-tower geometry, pointing to IP.

The performance of the ACC is checked by using the decay chain $D^{*-} \to \bar{D}^0\pi^-$ followed by $\bar{D}^0 \to K^+\pi^-$. Since the mass difference between D^{*-} and $\bar{D}^0$ is small, the momentum of π^- from the D^{*-} decay is characteristically low. Then, the K^+ and π^- from the $\bar{D}^0$ decay can be identified by the charges of the tracks, safely neglecting the doubly Cabibbo-suppressed decay $\bar{D}^0 \to K^-\pi^+$. Figure 3.10 shows the distributions of the number of photoelectrons, where good K/π separation is obtained for each module.

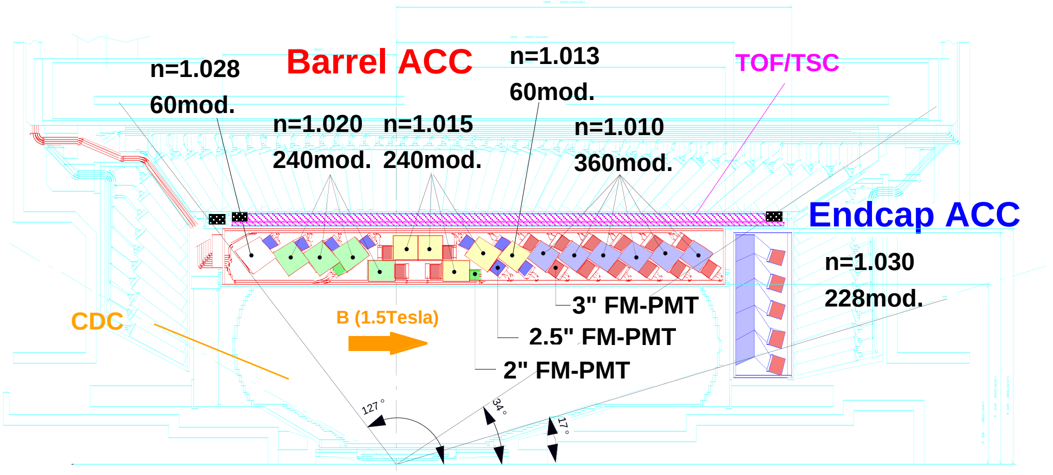


Figure 3.9: The arrangement of ACC.

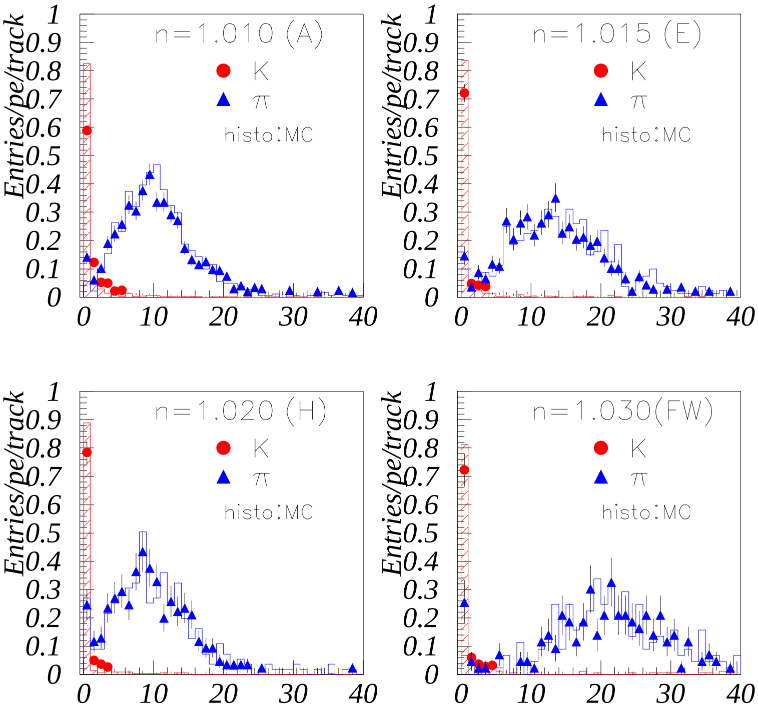


Figure 3.10: Distributions of the number of photoelectrons for $K^{\pm}$ and $\pi^{\mp}$ in $D^{*\mp}$ decays. Each plot corresponds to the different set of modules with a different refractive index.

3.2.4 Time-of-Flight Counter (TOF)

Scintillation counters utilize the ionization produced by charged particles to generate optical photons. We employ plastic scintillators with FM PMTs for improving the particle identification. The time of flight t of a particle is given by

$$t = \frac{l}{c\beta} = \frac{l}{c}\sqrt{1 + \left(\frac{m}{p}\right)^2}, \tag{3.5}$$

where l, β, p, and m are the path length, the velocity, the momentum, and the mass, respectively. Given the values of l and p, the measurement of t provides an identification of m. Figure 3.11 shows the dimension of a module, which contains two trapezoidally-shaped TOF counters and one thin trigger scintillation counter (TSC). The latter is used for keeping the fast trigger rate below 70 kHz. There are 64 modules located at a radius 1.2 m from IP, covering the barrel part $34^\circ < \theta < 120^\circ$.

By combining the measurements on the forward and backward FM PMTs, the time resolution of 100 ps is achieved, which provides an effective identification of particles of momenta below 1.2 GeV/c. Figure 3.12 shows the mass distribution obtained from TOF measurements for the particles with momenta below 1.2 GeV/c. Clear peaks corresponding to pions, kaons, and protons are seen.

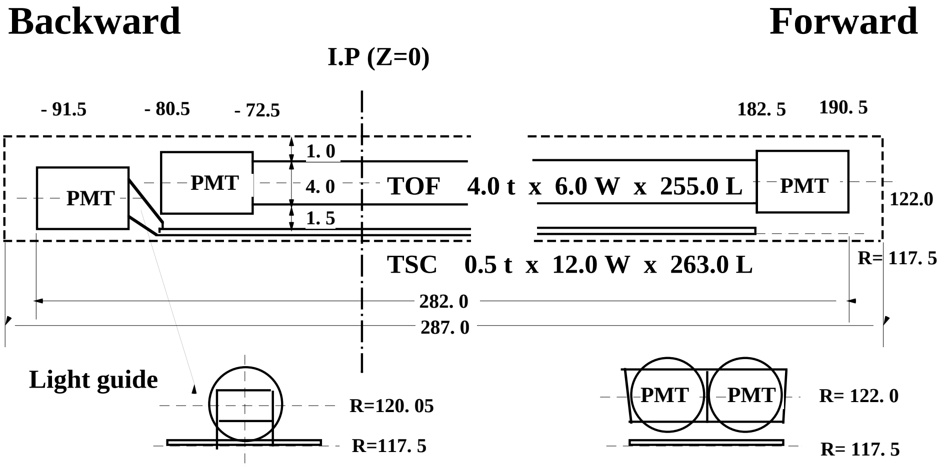


Figure 3.11: Dimension of a TOF/TSC Module

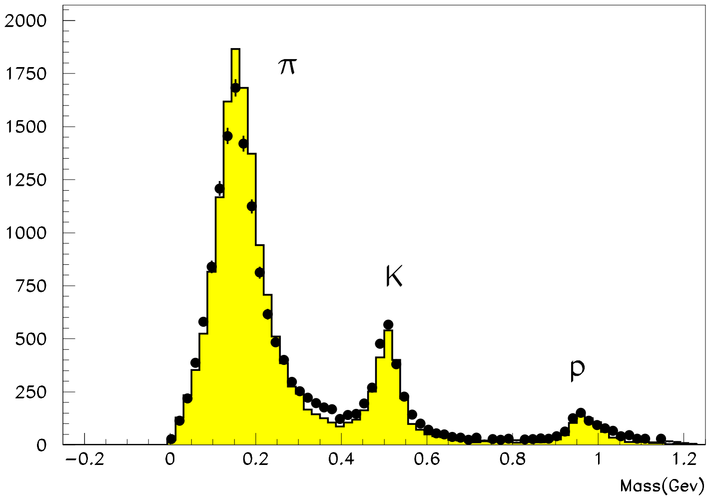


Figure 3.12: Mass distribution from TOF measurements for momenta below 1.2 GeV/c. The data points are in good agreement with a prediction for $\sigma_{\rm TOF} = 100$ ps (histogram).

3.2.5 Electromagnetic Calorimeter (ECL)

When an electron or a photon is incident on a thick absorber, an electromagnetic cascade is induced by pair production and bremsstrahlung. An electromagnetic calorimeter utilizes the generated shower for measuring the energy deposition and the position. A comparison with the momentum measurement provides the identification of electrons. The overall configuration of the ECL is shown in Figure 3.13. The ECL is based on an array of tower-shaped CsI(Tl) crystals that project to the vicinity of IP. The CsI(Tl) crystals have various nice features such as a large photon yield, weak hygroscopicity, and mechanical stability. The ECL consists of the barrel section of 3.0 m in length and the annular end-caps at $z = +2.0$ and $z = -1.0$ m from IP, covering the angular range $17° < \theta < 150°$. The entire system contains 8736 counters and weighs 43 tons. The readout is based on an independent pair of silicon PIN photodiodes and charge-sensitive preamplifiers attached at the end of each crystal.

The energy resolution is measured by a beam test to be

$$\frac{\sigma_E}{E}\ (\%) = \sqrt{\left(\frac{0.066}{E}\right)^2 + \left(\frac{0.81}{\sqrt[4]{E}}\right)^2 + 1.34^2} \quad (E \text{ in GeV}), \tag{3.6}$$

where the value is affected by the electronic noise (1st term), the shower leakage fluctuation (2nd and 3rd terms), and the systematic effect such as the uncertainty of calibration (3rd term). The spacial resolution is approximately found to be 0.5 cm/$\sqrt{E}$ (E in GeV).

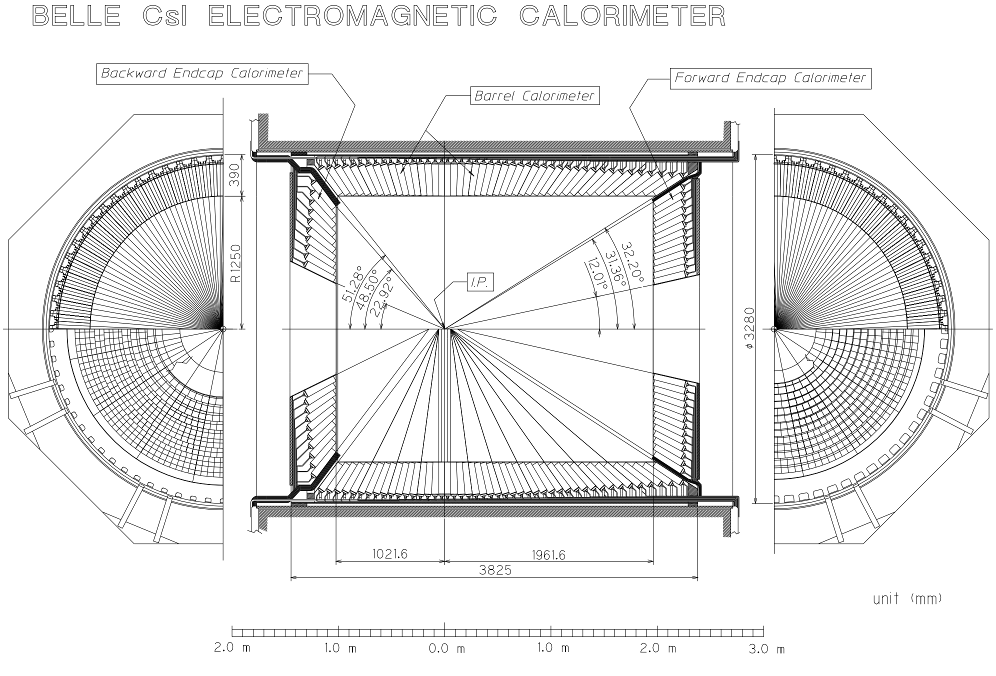
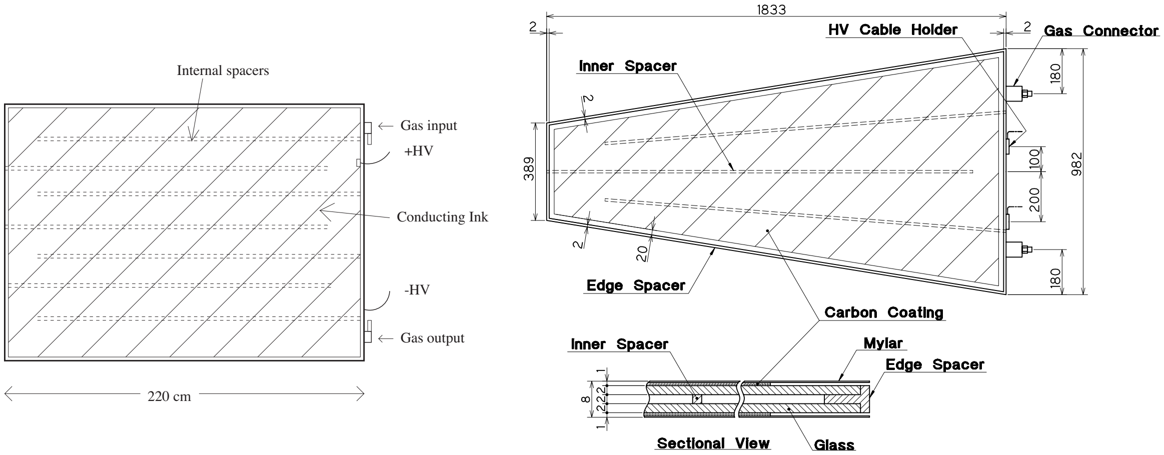


Figure 3.13: Overall configuration of ECL.

3.2.6 K_L and Muon Detector (KLM)

The KLM system makes it possible to identify K_L's and muons by the alternating layers of 4.7 cm-thick iron plates and charged particle detectors. There are 14 iron layers and 15 (14) detector layers in the barrel (end-cap) region. The iron plates provide 3.9 interaction length of material, in addition to 0.8 interaction length of ECL. The K_L that interacts in the iron plates or ECL produces a shower of ionizing particles, the location of which with respect to IP determines the flight direction of K_L. The muons travel much farther with smaller deflections on average than strongly interacting hadrons, which enables the discrimination between muons and other charged hadrons.

The charged particle detectors are based on glass-electrode-resistive plate counters (RPCs). The schematic diagrams of the barrel and end-cap modules are shown in Figure 3.14. The counters have two parallel plate electrodes with high bulk resistivity ($\geq 10^{10}$ Ωcm) separated by a gas-filled gap. An ionizing particle traversing the gap induces a streamer in the gas that results in a local discharge. The discharge generates a signal on external pickup strips, and the location and the time are recorded. The number of K_L clusters per event is in good agreement with the prediction. Typical muon identification efficiency is 90% with a fake rate around 2%.



(a) Barrel RPC.

(b) End-cap RPC.

Figure 3.14: Schematic diagrams of the internal spacer arrangement for (a) barrel and (b) end-cap RPCs.

3.2.7 Trigger and Data Acquisition

The events of our interest are smaller in the cross section than the background processes including the beam backgrounds, and they have to be triggered by a restrictive condition. Our trigger system consists of a hardware trigger and a software trigger. Figure 3.15 shows the schematic view of the hardware trigger system. The CDC and TSC are used to obtain trigger signals for charged particles. The ECL provides total energy deposit and cluster counting of crystal hits. The KLM gives additional information on muons. An extreme forward calorimeter (EFC), which is a radiation-hard BGO (Bismuth Germanate, $Bi_4Ge_3O_{12}$) crystal calorimeter covering 6.4°–11.5° and 163.3°–171.2°, is used for tagging two photon events as well as Bhabha events. The trigger signals of the sub-triggers are combined to characterize the event type in the central trigger system called the global decision logic. The software trigger is designed to be implemented in the online computer farm. The efficiency for multi-hadronic data samples is more than 99.5%.

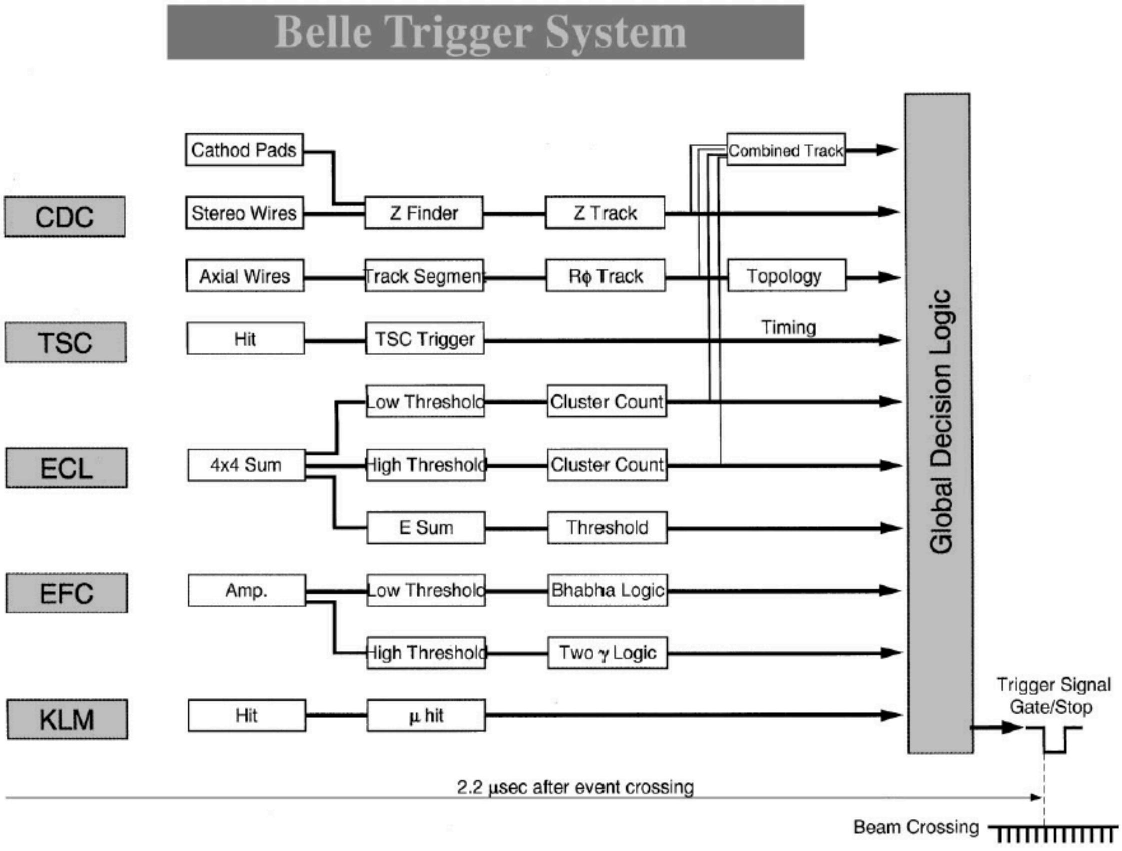


Figure 3.15: Schematic view of the trigger system.

The data acquisition is based on a distributed-parallel system working at 500 Hz with a deadtime fraction of less than 10%. Figure 3.16 shows the overview of the data acquisition system. Information from the seven sub-detectors are combined into a single event record by an event builder. The output of the event builder is transferred to an online computer farm, where an event filtering is applied after the fast event reconstruction. Then, the data are sent to a mass storage system via optical fibers.

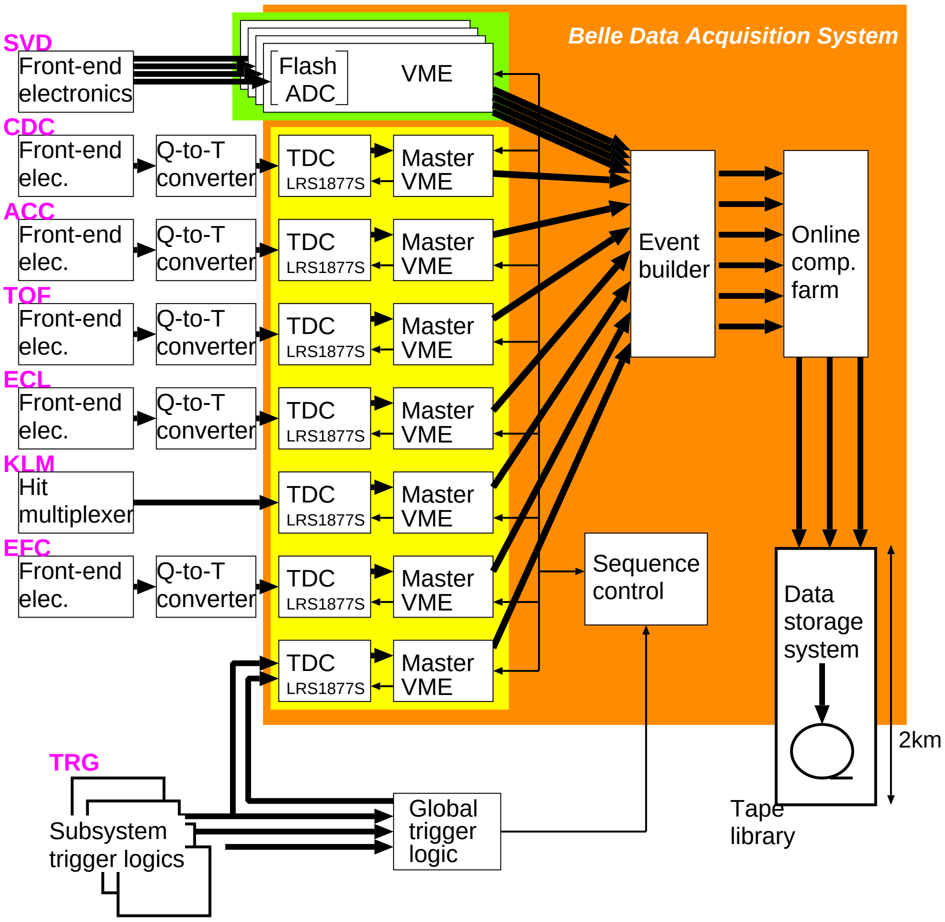


Figure 3.16: Overview of the data acquisition system.

3.3 Analysis Tools

3.3.1 $K^{\pm}/\pi^{\pm}$ Identification

The $K^{\pm}/\pi^{\pm}$ identification is based on the dE/dx measurement by CDC, the Cherenkov-light yield in ACC, and the time-of-flight information from TOF. The likelihood functions of $K^\pm$ ($\pi^\pm$) for the three detector responses are combined as a product to obtain the kaon (pion) likelihood $\mathcal{L}_K$ ($\mathcal{L}_\pi$). The likelihood ratio $P(K/\pi)$ is then calculated as

$$P(K/\pi) = \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi}, \tag{3.7}$$

from which we discriminate between $K^\pm$ and $\pi^\pm$.

The decay $D^{*-} \to \bar{D}^0\pi^-$ followed by $\bar{D}^0 \to K^+\pi^-$ can be reconstructed without particle identification requirements by utilizing the soft-pion charge as in Section 3.2.3. This feature is exploited to validate the method of the $K^\pm/\pi^\pm$ identification. Figure 3.17 shows a scatter plot of the track momentum and the likelihood ratio $P(K/\pi)$ for $K^\pm$ and $\pi^\pm$. Clear separation in a momentum range up to around 4 GeV/c is obtained, which is achieved by the different momentum coverages of the detectors shown in Figure 3.18.

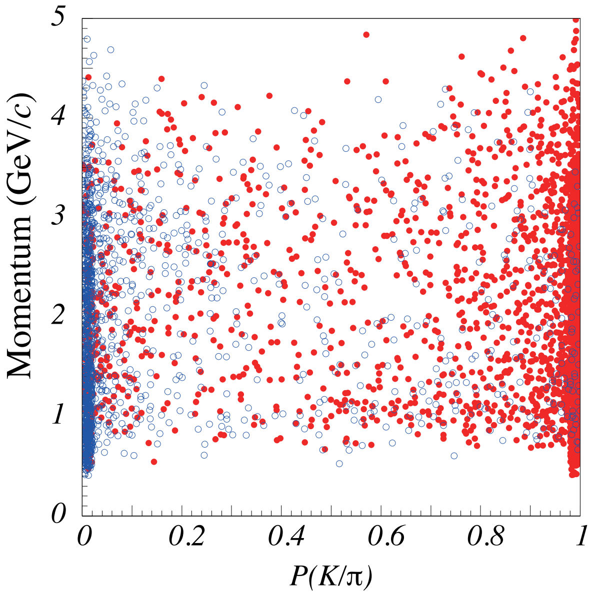


Figure 3.17: A scatter plot of the track momentum and the likelihood ratio $P(K/\pi)$ for $K^\pm$ (red circle) and $\pi^\pm$ (blue circle) obtained from the data of the decay chain $D^{*-} \to \bar{D}^0\pi^-$ followed by $\bar{D}^0 \to K^+\pi^-$. Strong concentration in the region of $P(K/\pi) \sim 1$ (~ 0) is observed for $K^\pm$ ($\pi^\pm$) over a wide momentum region up to about 4 GeV/c.

3.3.2 Continuum Suppression

In many analyses of the B meson decays, large backgrounds arise from the continuum processes $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$). The cross section of the transition $e^+e^- \to q\bar{q}$ is similar to the one of $e^+e^- \to B\bar{B}$ at the $\Upsilon(4S)$ resonance. Several methods to reduce the background events exploit the event topology, where a $q\bar{q}$ ($B\bar{B}$) event has a jet-like

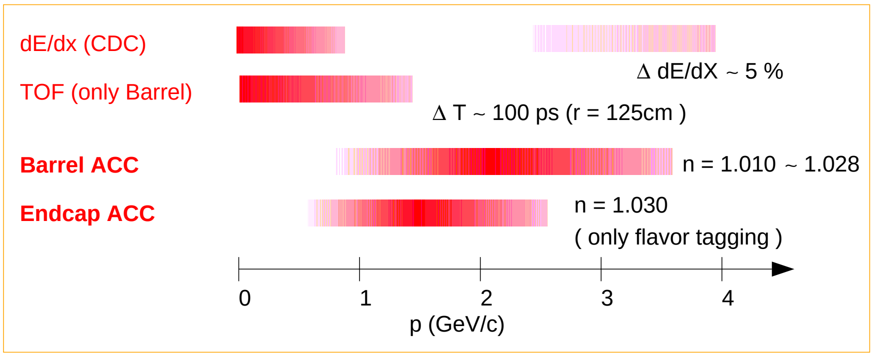


Figure 3.18: Momentum coverage of the detectors used for $K^{\pm}/\pi^{\pm}$ separation.

(spherical) shape. The event topology is characterized on the basis of the Fox-Wolfram moments [25]. The l-th moment is defined by the l-th Legendre polynomial P_l as

$$H_l \equiv \sum_{i,j} |\vec{p}_i||\vec{p}_j| P_l(\cos\theta_{ij}), \tag{3.8}$$

where $\vec{p}_i$ and $\vec{p}_j$ are the momenta in the CM frame of the i-th and j-th particles, respectively, θ_{ij} is the angle between $\vec{p}_i$ and $\vec{p}_j$, and the sum is over the particles in the final state. Note that the overall constant is ignored here for simplicity.

SFW

The Fox-Wolfram moments can be divided as

$$H_l = H_l^{\mathrm{ss}} + H_l^{\mathrm{so}} + H_l^{\mathrm{oo}}, \tag{3.9}$$

where H_l^{ss} is the moment on combinations of two particles "s" from the reconstructed B, H_l^{so} is the moment on combinations of a particle "s" from the reconstructed B and a particle "o" not from the reconstructed B, and H_l^{oo} is the moment on combinations of two particles "o" not from the reconstructed B. We define a Fisher discriminant [26] named SFW (Super Fox-Wolfram) by the divided Fox-Wolfram moments as

$$\mathrm{SFW} \equiv \sum_{l=2,4} \alpha_l \left(\frac{H_l^{\mathrm{so}}}{H_0^{\mathrm{so}}}\right) + \sum_{l=1}^{4} \beta_l \left(\frac{H_l^{\mathrm{oo}}}{H_0^{\mathrm{oo}}}\right), \tag{3.10}$$

where the moments with negligible correlations to the four-momentum of the B candidate are employed. The α_l and β_l are the Fisher coefficients, which are selected to maximize the power of separation. The SFW is an effective discriminant for suppressing the continuum background.

KSFW

In order to improve the discrimination, the SFW is modified to derive so-called KSFW (Kakuno Super Fox-Wolfram) [27, 28] by considering the charges of the particles, the missing mass of the event, and the normalization factors. The KSFW is defined as

$$\mathrm{KSFW} \equiv \sum_{l=0}^{4} R_l^{\mathrm{so}} + \sum_{l=0}^{4} R_l^{\mathrm{oo}} + \gamma \sum_{n=1}^{N_t} |p_{t,n}|, \tag{3.11}$$

where the explanations of the terms are the following.

- The R_l^{so} is obtained by dividing the particles "o" not from the reconstructed B into the charged particles "c" and the neutral particles "n" and adding the missing mass "m" as
  $$R_l^{\mathrm{so}} = \frac{\alpha_l^{\mathrm{c}} H_l^{\prime\mathrm{sc}} + \alpha_l^{\mathrm{n}} H_l^{\prime\mathrm{sn}} + \alpha_l^{\mathrm{m}} H_l^{\prime\mathrm{sm}}}{E_{e^+e^-} - E_B}, \tag{3.12}$$
  where $E_{e^+e^-}$ and E_B are the energy in the CM frame of e^+e^- and B, respectively, and α_l^{c}, α_l^{n}, and α_l^{m} are the Fisher coefficients for the charged, neutral, and missing-mass categories, respectively. The $H_l^{\prime\mathrm{sc}}$, $H_l^{\prime\mathrm{sn}}$, and $H_l^{\prime\mathrm{sm}}$ are the modified Fox-Wolfram moments with respect to the charges of particles. The normalization $E_{e^+e^-} - E_B$ is included to remove the small correlation of H_0^{so} to the B energy. The terms for "n" and "m" are zero for $l = 1, 3$, and we totally have 11 terms for $l = 0$–4.

- The R_l^{oo} is similar to the corresponding term in SFW, while the moment is modified with respect to the charges, noted as $H_l^{\prime\prime\mathrm{oo}}$, and the normalization is replaced to remove the correlation to the B energy as
  $$R_l^{\mathrm{oo}} = \frac{\beta_l H_l^{\prime\prime\mathrm{oo}}}{(E_{e^+e^-} - E_B)^2}. \tag{3.13}$$
  There are totally 5 terms corresponding to $l = 0$–4.

- The $\gamma \sum_{n=1}^{N_t} |p_{t,n}|$ is the scalar sum of the transverse momenta p_t of all particles in the event multiplied by a Fisher coefficient γ. The n represents particle index, and the N_t is the number of particles in the event.

The KSFW has totally 17 (= 11+5+1) Fisher coefficients. Since the KSFW is strongly correlated with the missing mass, the values of the coefficients are optimized depending on the divided missing-mass regions. The discriminants for the divided regions are combined by the likelihood method.

The definition of KSFW was obtained primarily for the analysis of $B^0 \to \pi\pi$, while it is also very efficient for other modes and widely used. The distributions of KSFW will be shown later for $B^- \to DK^-$.

Chapter 4

Event Selection

We detect the suppressed B meson decay $B^- \to [K^+\pi^-]_D K^-$ as well as the favored decay $B^- \to [K^-\pi^+]_D K^-$. The same selection criteria are used for both decays whenever possible in order to cancel systematic uncertainties. The decays $B^- \to [K^+\pi^-]_D \pi^-$ and $B^- \to [K^-\pi^+]_D \pi^-$ are also analyzed as reference modes. The ratio $\mathcal{R}_{D\pi}$ and the asymmetry $\mathcal{A}_{D\pi}$ are introduced similarly to Eqs. (2.27) and (2.28). The $D\pi^-$ mode have an analogous event topology and is Cabibbo-enhanced relative to the DK^- mode, while the effects of CP-violation are smaller.

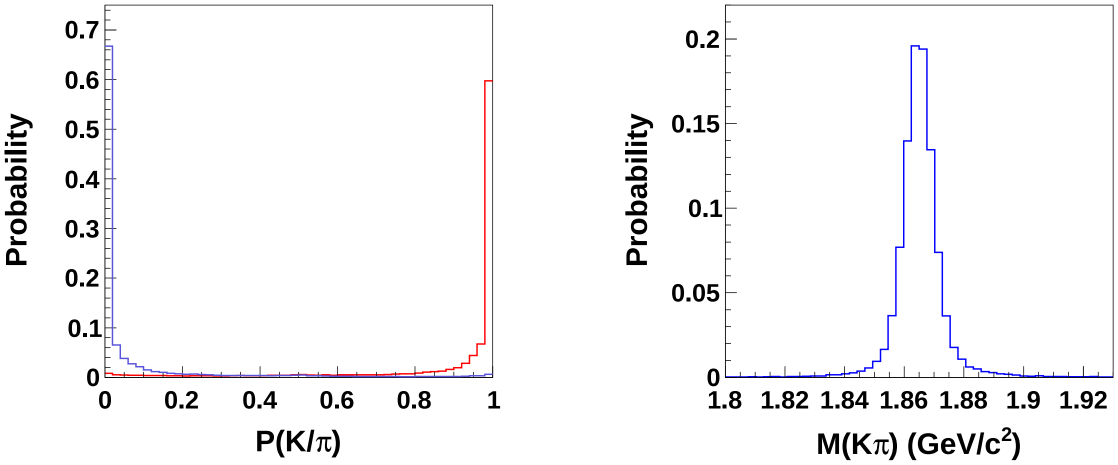
A kaon or a pion is indicated by a "h" meson. Also, a particle that originates directly from a B meson is referred to as a "prompt" particle.

4.1 Charged Tracks

All candidate tracks are required to have a point of closest approach to the beam line within ± 5 mm of IP in the direction perpendicular to the beam axis and ± 5 cm in the direction parallel to the beam axis.

4.2 Reconstruction of D Meson

The D meson candidates are reconstructed from pairs of oppositely charged tracks. For each track, we apply a particle identification requirement based on the K/π likelihood ratio $P(K/\pi) = \mathcal{L}_K/(\mathcal{L}_K + \mathcal{L}_\pi)$, where the distribution for a signal Monte Carlo (MC) sample is shown in Figure 4.1 (a). We use the requirements $P(K/\pi) > 0.4$ and $P(K/\pi) < 0.7$ for the kaon and pion candidates, respectively. The efficiency to identify a kaon or a pion is 90–95%, while the probability of a pion (kaon) being misidentified as a kaon (pion) is around 15%. The invariant mass of the $K\pi$ pair must be within $\sim 3\sigma$ of the nominal D mass [22]: $1.850\ \text{GeV}/c^2 < M(K\pi) < 1.880\ \text{GeV}/c^2$. The $M(K\pi)$ distribution for a signal MC sample is shown in Figure 4.1 (b). After the $M(K\pi)$ requirement, tracks from the D candidate are refitted with their invariant mass constrained to the nominal D mass to improve the momentum determinations.



(a) $P(K/\pi)$ for kaon (red) and pion (blue).

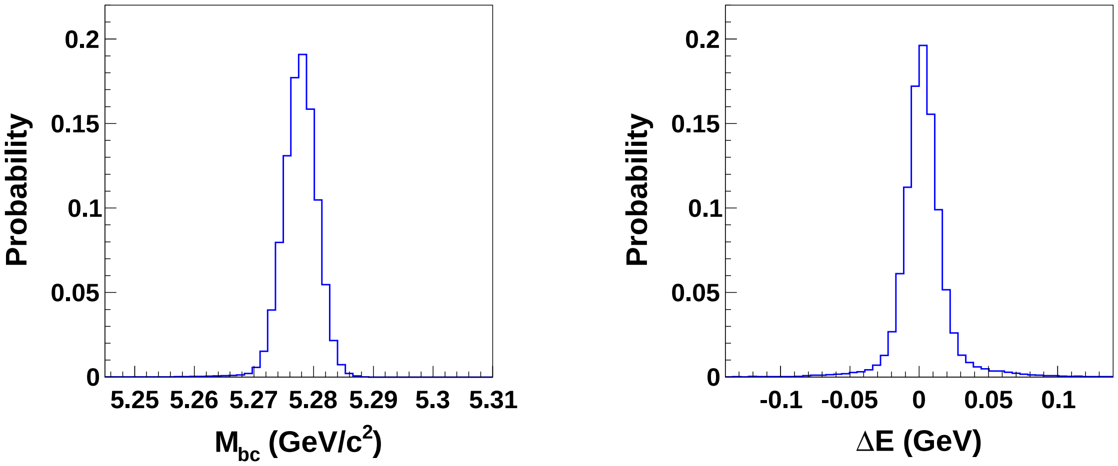
(b) $M(K\pi)$ distribution.

Figure 4.1: Distributions of $P(K/\pi)$ and $M(K\pi)$ for a signal MC sample.

4.3 Reconstruction of B Meson

The B meson candidates are reconstructed by combining a D candidate with a prompt charged hadron candidate. We apply particle identification requirements $P(K/\pi) > 0.6$ and $P(K/\pi) < 0.4$ for $B^- \to DK^-$ and $B^- \to D\pi^-$, respectively. The efficiency to identify a kaon or a pion is 85–90%, while the probability of a pion (kaon) being misidentified as a kaon (pion) is around 10%. Tighter criteria are employed to suppress the contamination between DK^- and $D\pi^-$. The signal is identified using the beam-energy-constrained mass ($M_{\rm bc}$) and the energy difference (ΔE) defined, in the CM frame, as $M_{\rm bc} = \sqrt{E_{\rm beam}^2 - |\vec{p}_B|^2}$ and $\Delta E = E_B - E_{\rm beam}$, where $E_{\rm beam}$ is the beam energy and $\vec{p}_B$ and E_B are the momentum and energy of the B meson candidate, respectively. The distributions for a signal MC sample are shown in Figure 4.2. We require 5.271 GeV/$c^2 < M_{\rm bc} <$ 5.287 GeV/c^2 ($\pm 3\sigma$), and fit the ΔE distribution in -0.1 GeV $< \Delta E <$ 0.3 GeV to extract the signal yield.

In the rare cases where there are more than one candidate in an event (0.3% for $B^- \to DK^-$ and 0.8% for $B^- \to D\pi^-$), we select the best candidate on the basis of a χ^2 determined from the difference between the measured and nominal values of $M(K\pi)$ and $M_{\rm bc}$. The former is the value without the D-mass constraint on the tracks. For the calculation of χ^2, the values 5.0 MeV/c^2 and 2.7 MeV/c^2 are used as the errors of $M(K\pi)$ and $M_{\rm bc}$, respectively.



(a) M_{bc} distribution. (b) ΔE distribution.

Figure 4.2: Distributions of M_{bc} and ΔE for a signal MC sample.

4.4 Rejection of Backgrounds

4.4.1 $B^- \to [K^+K^-]_D\pi^-$

For $B^- \to [K^+\pi^-]_D K^-$, a possible background arises from the decay $B^- \to [K^+K^-]_D\pi^-$, which has the same final state and can peak in the signal window. To reject this background, we veto the events that satisfy $\mid M(KK) - 1.865\ \mathrm{GeV}/c^2 \mid < 0.025\ \mathrm{GeV}/c^2$. Figure 4.3 shows the distributions of $M(KK)$ for the background and signal MC samples. Relative loss of signal efficiency by the requirement is 0.8%. The number of events remaining after the veto will be estimated later by using the $M(K\pi)$ sideband sample of data.

4.4.2 $B^- \to [K^-\pi^+]_D h^-$

The favored decay $B^- \to [K^-\pi^+]_D h^-$ can also produce a peaking background if both the kaon and the pion from the D decay are misidentified and the particle assignments are interchanged. We thus veto the events for which the invariant mass of the $K\pi$ pair is inside the D mass window when the mass assignments are exchanged: $|M(K\pi)_{\mathrm{exchanged}} - 1.865\ \mathrm{GeV}/c^2| < 0.020\ \mathrm{GeV}/c^2$. Figures 4.4 and 4.5 show the distributions of $M(K\pi)_{\mathrm{exchanged}}$ for the background and signal MC samples. Relative loss of signal efficiency by the requirement is 11.3%. The number of events remaining after the veto will be estimated later by using the $M(K\pi)$ sideband sample of data.

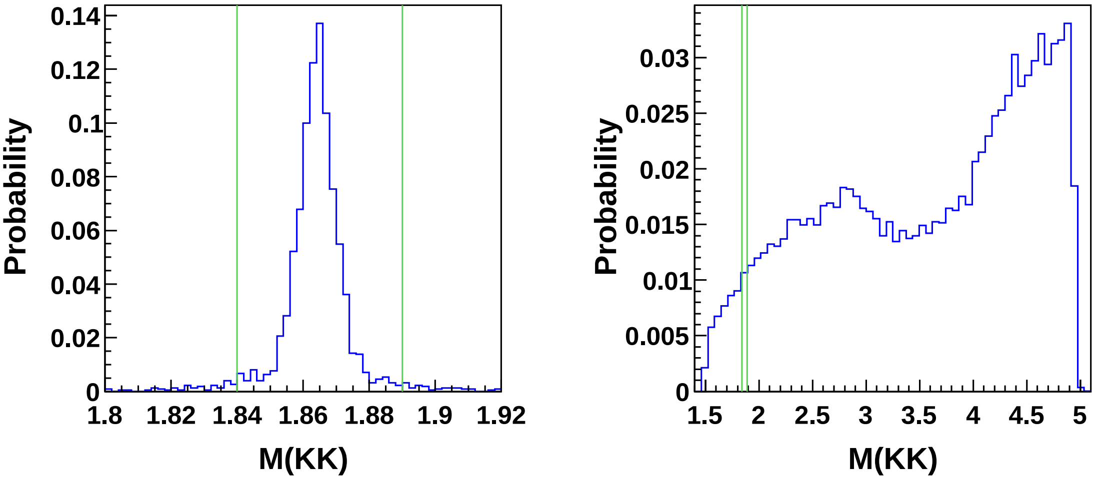


Figure 4.3: The distributions of $M(KK)$ for MC samples of the background $B^- \to [K^+K^-]_D\pi^-$ (left) and the signal $B^- \to [K^+\pi^-]_D K^-$ (right). The vertical lines show the veto region.

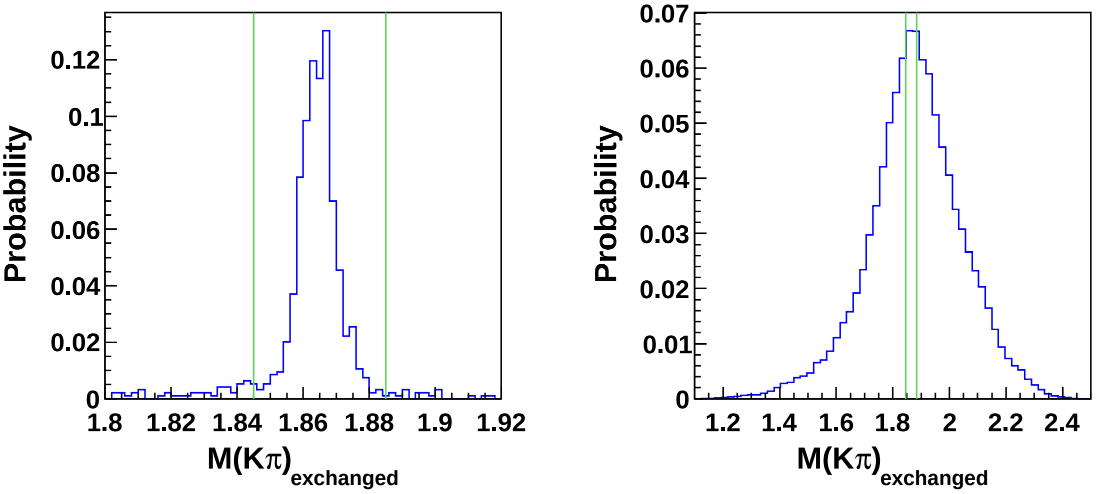


Figure 4.4: The distributions of $M(K\pi)_{\text{exchanged}}$ for MC samples of the background $B^- \to [K^-\pi^+]_D K^-$ (left) and the signal $B^- \to [K^+\pi^-]_D K^-$ (right). The solid vertical lines show the veto region.

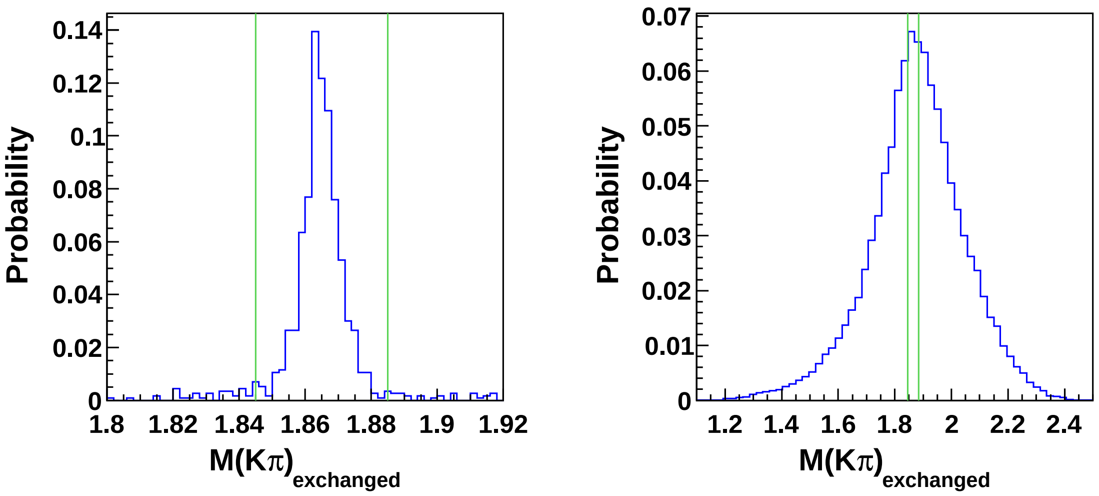


Figure 4.5: The distributions of $M(K\pi)_{\rm exchanged}$ for MC samples of the background $B^- \to [K^-\pi^+]_D\pi^-$ (left) and the signal $B^- \to [K^+\pi^-]_D\pi^-$ (right). The solid vertical lines show the veto region.

4.4.3 Continuum Background

A large background arises from the continuum processes $e^+e^- \to q\bar{q}$ $(q = u, d, s, c)$. In order to remove $D^{*-} \to \bar{D}^0\pi^-$ decays included in such a process, we employ a variable ΔM defined as the mass difference between the D^{*-} and D candidates, where the D^{*-} candidate is reconstructed from the D candidate used in the B reconstruction and a π^- candidate not used in the B reconstruction. When there are multiple charged tracks not used in the B reconstruction, we select the one for which the value of ΔM is closest to 0.14 GeV/c^2. Figure 4.6 shows the ΔM distributions for $q\bar{q}$ background and signal, where a good agreement between MC and data samples is seen. We require $\Delta M > 0.15$ GeV/c^2, which removes 20% of $q\bar{q}$ background with relative loss of signal efficiency of 3.9%. Note that the requirement on ΔM also removes the $D^{*-} \to \bar{D}^0\pi^-$ from $B\bar{B}$ events.

After the ΔM requirement, the continuum background is discriminated by using a neural network technique based on ten variables. For the output of neural network, named NB, we apply a loose requirement $NB > -0.6$, which retains 96% of the signal and removes 74% of the background according to MC samples. The variable NB is fitted with ΔE to extract the signal yield. The details about NB are described in the next chapter.

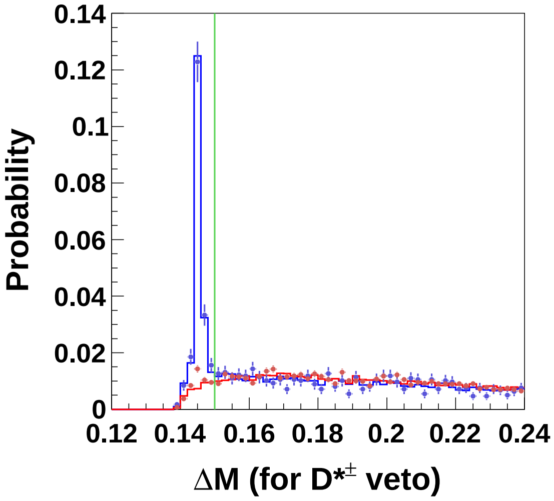


Figure 4.6: The ΔM distributions for the $q\bar{q}$ background (blue) and the signal (red). The histograms are obtained from MC samples, while the dots with error bars are obtained from data samples. For data, we use the sideband sample indicated by 0.15 GeV $< \Delta E <$ 0.30 GeV for the $q\bar{q}$ background and the $B^- \rightarrow [K^-\pi^+]_D\pi^-$ sample in $|\Delta E| <$ 0.05 GeV for the signal. The consistencies in different ΔE regions and in different signal samples are checked by MC samples. The solid vertical line shows the value for the veto.

Chapter 5

Discrimination of Continuum Background

A large background arises from the continuum processes $e^+e^- \to q\bar{q}$ $(q = u, d, s, c)$. In this chapter, we describe a new approach to discriminate between the signal and the continuum background using a neural network method.

5.1 Likelihood Method

In this section, we review the previous method [19] of the discrimination between the signal and the continuum background. We employed two variables, KSFW and $|\cos\theta_B|$, where θ_B is the polar angle of the B candidate in the CM frame. Figures 5.1 and 5.2 show the distributions of KSFW and $|\cos\theta_B|$ for MC samples. The KSFW was obtained for seven missing-mass regions. In the variable $|\cos\theta_B|$, the signal follows a $\sin^2\theta_B$ distribution, while the background is approximately uniform.

To combine the information from KSFW and $|\cos\theta_B|$, a likelihood ratio $\mathcal{LR}$ was obtained by the likelihoods for KSFW ($\mathcal{L}_{\text{KSFW}}$) and $|\cos\theta_B|$ ($\mathcal{L}_{|\cos\theta_B|}$) as

$$\mathcal{LR} = \frac{\mathcal{L}^{\text{sig}}_{\text{KSFW}}\mathcal{L}^{\text{sig}}_{|\cos\theta_B|}}{\mathcal{L}^{\text{sig}}_{\text{KSFW}}\mathcal{L}^{\text{sig}}_{|\cos\theta_B|} + \mathcal{L}^{\text{bg}}_{\text{KSFW}}\mathcal{L}^{\text{bg}}_{|\cos\theta_B|}}, \tag{5.1}$$

where the superscript “sig” (“bg”) indicates the likelihood for signal (background). Figure 5.3 (a) shows the $\mathcal{LR}$ distributions for the signal and background MC samples. A plot on the rate of background rejection and the signal efficiency for each x in the requirement $\mathcal{LR} > x$ is shown in Figure 5.3 (b). The optimized requirement in the previous analysis was $\mathcal{LR} > 0.9$, which rejects 99% of background retaining 45% of signal.

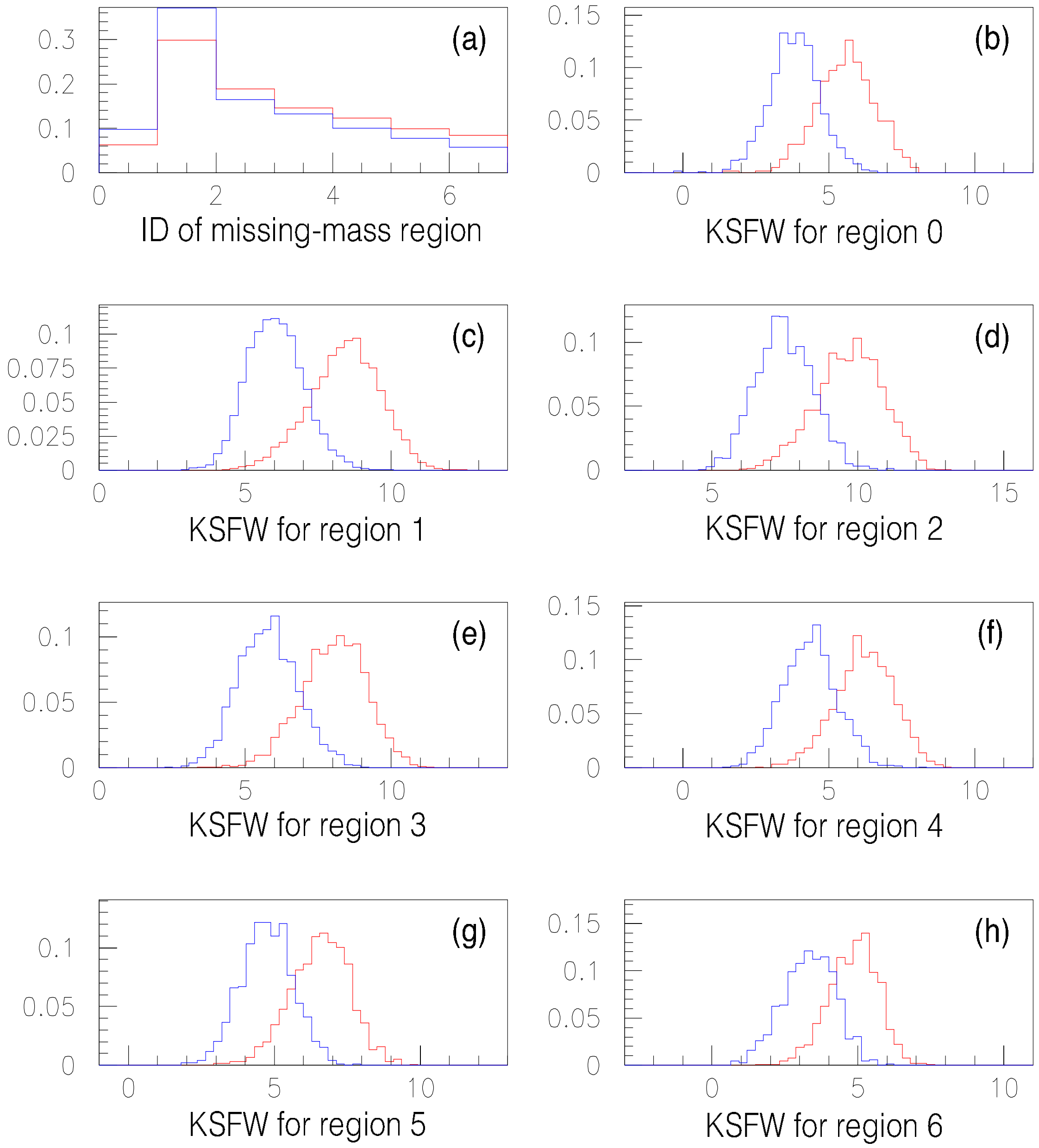


Figure 5.1: The distributions of (a) index of missing-mass region and (b)–(h) KSFW for each missing-mass region. The regions are defined by the missing-mass square mm^2 as $mm^2 < -0.5$ GeV/c^2 (region 0), -0.5 GeV/$c^2 < mm^2 < 0.3$ GeV/c^2 (region 1), 0.3 GeV/$c^2 < mm^2 < 1.0$ GeV/c^2 (region 2), 1.0 GeV/$c^2 < mm^2 < 2.0$ GeV/c^2 (region 3), 2.0 GeV/$c^2 < mm^2 < 3.5$ GeV/c^2 (region 4), 3.5 GeV/$c^2 < mm^2 < 6.0$ GeV/c^2 (region 5), and $mm^2 > 6.0$ GeV/c^2 (region 6). The red and blue histograms correspond to the signal and background MC samples, respectively.

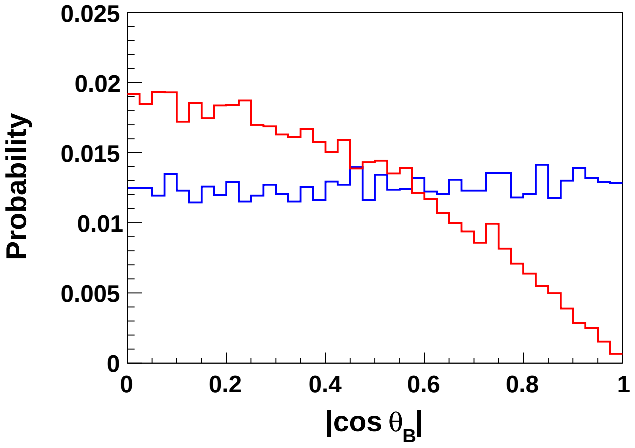
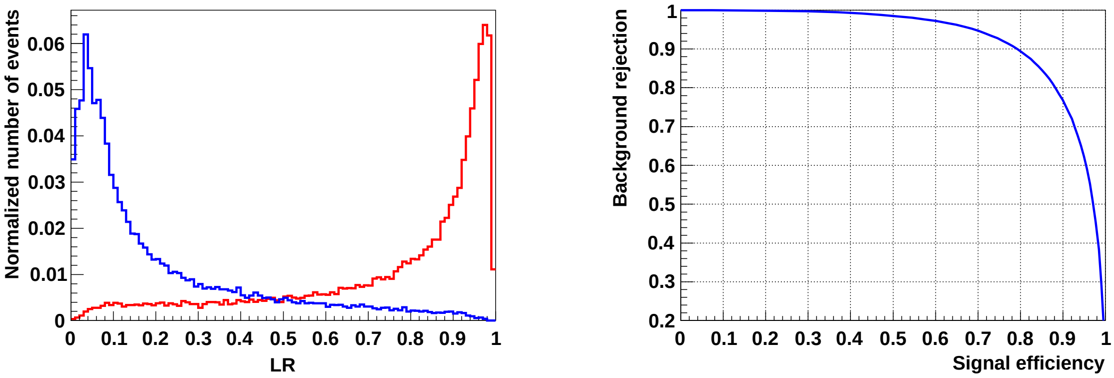


Figure 5.2: The $|\cos\theta_B|$ distributions for the signal (red) and background (blue) MC samples.



(a) The $\mathcal{LR}$ distributions for the signal (red) and background (blue) MC samples.

(b) Plot of the background rejection and the signal efficiency for each requirement on $\mathcal{LR}$.

Figure 5.3: Discrimination between signal and continuum background by the likelihood method based on KSFW and $|\cos\theta_B|$.

5.2 Introduction of Variables

To improve the discrimination between the signal and the continuum background, we introduce ten variables listed in the following. For each variable, the distributions for MC and data samples are shown in Figure 5.4.

- $\mathcal{LR}$(KSFW): a likelihood ratio of KSFW defined by $\mathcal{L}^{\rm sig}_{\rm KSFW}/(\mathcal{L}^{\rm sig}_{\rm KSFW}+\mathcal{L}^{\rm bg}_{\rm KSFW})$. The distribution is peaked at one for the signal and zero for the background.

- $|\cos\theta_{\rm T}|$: absolute value of the cosine of the angle between the thrust axes of the B candidate and the rest of the event, where the thrust axis is oriented in such a way that the sum of momentum projections has a maximum value. For the rest of the event, we use all charged tracks, for which the pion mass is assigned, and photons which have the energies larger than 0.1 GeV. The distribution is approximately uniform for the signal and strongly peaked at one for the background.

- Δz: vertex separation between the B candidate and the remaining tracks [29]. The absolute value tends to be larger for the signal than the background, since the B meson has larger life time.

- $\cos\theta^K_D$: cosine of the angle between the direction of the K from the D decay and the opposite direction to the B in the rest frame of D. The distribution for the signal is about flat because of the zero spins of the particles, while the one for the background is peaked mostly at one.

- $r_{\rm tag}$: expected B flavor dilution factor [30] that ranges from zero for no flavor information to unity for unambiguous flavor assignment. The signal events tend to have higher values than the background.

- $|\cos\theta_B|$: absolute value of the cosine of the polar angle of the B candidate in the CM frame. The signal follows a $\sin^2\theta_B$ distribution, while the background is approximately uniform.

- ΔQ: difference of charge sums in the D hemisphere and the opposite hemisphere after removing the B candidate, where we apply opposite signs for B^- and B^+ candidates. The value is around zero for the signal and slightly shifted for the $c\bar{c}$ background.

- Q_BQ_K: product of the charge of the B candidate and the charge sum of kaons not used in the B reconstruction. Since the B^- (B^+) meson tends to produce the K^- (K^+) meson, the value for the signal is likely to be lower than zero. The value is zero for most background events.

- d_{Dh}: distance of closest approach between trajectories of the D and h candidates. The value is close to zero for the signal, while the value can be larger for the $c\bar{c}$ background for which the h candidate can be secondary from a decay of a particle with longer life time.

- $\cos\theta_B^D$: cosine of the angle between the D direction and the opposite direction to the $\Upsilon(4S)$ in the rest frame of B. The distribution is approximately flat for the signal, while the one slopes slightly for the $c\bar{c}$ background.

5.3 NeuroBayes Neural Network

For combining the variables, we employ the NeuroBayes package [31], which is a highly sophisticated tool for multivariate analysis of correlated data on the basis of Bayesian statistics. An automated preprocessing of the input variables is followed by a three-layered (input, hidden, and output layers) neural network. 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 both for classification and shape reconstruction.

The preprocessing enables to find the optimal starting point for the subsequent network training. All variables can be normalized and linearly decorrelated such that the covariance matrix of new set of input variables becomes the unit matrix. Binary or discrete variables are automatically recognized and treated accordingly. A very important option is to handle the variables with a default value or delta-function, which can e.g. occur if the vertex fitting is failed for obtaining Δz. The preprocessing is very robust and efficient for powerful classification.

In the neural network, it is of vital importance to improve the generalization, which is realized by an advanced regularization and pruning techniques. Employing regularization based on Bayesian statistics practically eliminates the risk of overtraining and enhances the generalization abilities. Insignificant network connections and even entire nodes are removed to ensure that the network learns only the physical features of data. The resulting network represents the minimal topology needed to correctly reproduce the characteristics of the data while being insensitive to statistical fluctuations.

5.4 Neural Network Training

The neural network is trained with MC samples for signal and continuum background. For the latter, we loosen the requirement on $M_{\rm bc}$ as $5.20\ {\rm GeV}/c^2 < M_{\rm bc} < 5.29\ {\rm GeV}/c^2$ to obtain larger number of events.[1] We use 200,000 events for each sample after the event-selection requirements. Figure 5.5 (a) shows the result of the training, where the output is denoted as NB. A plot on the rate of background rejection and the signal efficiency for each x in the requirement $NB > x$ is shown in Figure 5.5 (b). In Table 5.1, we numerically summarize the performance with a comparison to the likelihood method. The output NB is fitted with ΔE to extract the signal yield. The expected significance of the signal is increased by about a factor two compared to the previous analysis. The details about the training are described in Appendix A.

---

[1] The correlation factors to $M_{\rm bc}$ for all input variables are less than 5%.

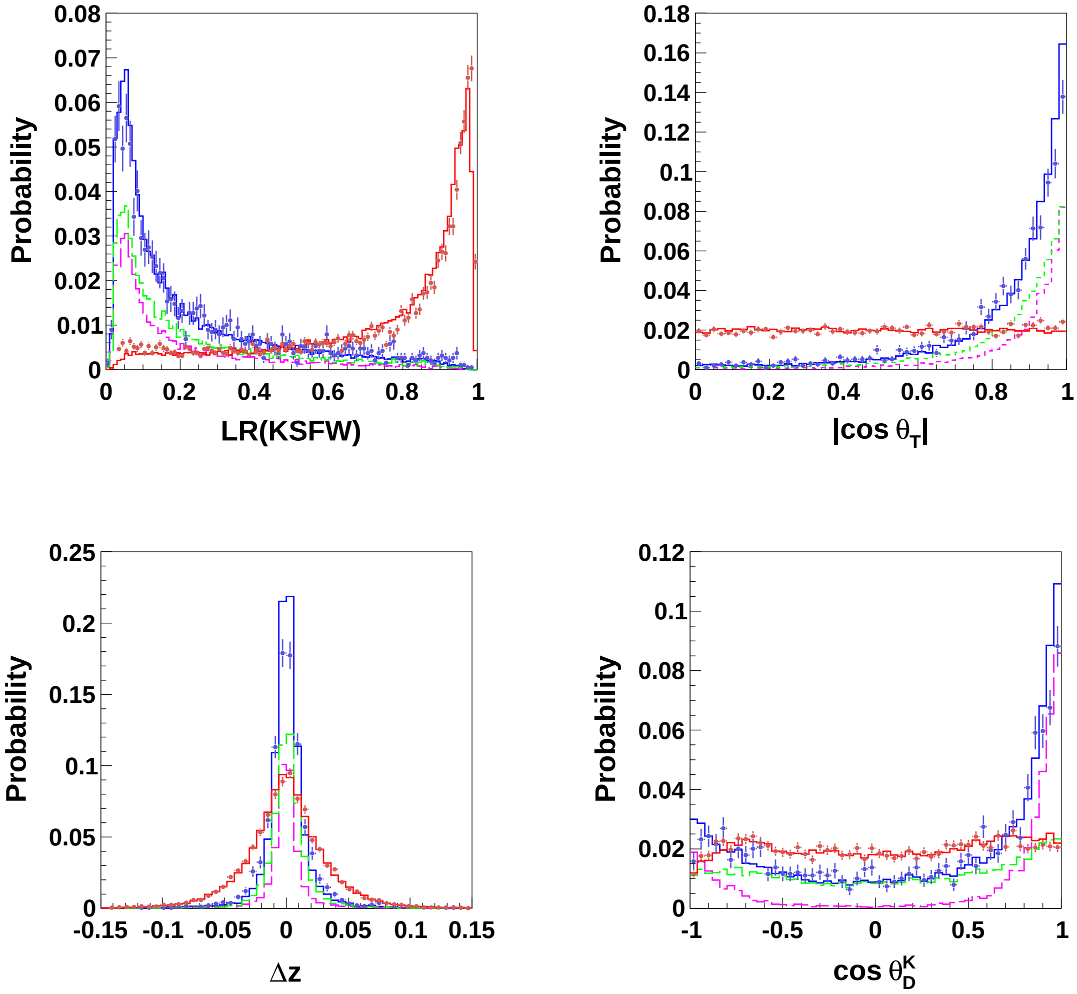


Figure 5.4: The distributions of the variables used for the discrimination between the signal and the continuum background. The histograms show the distributions for MC samples of signal (red) and background (blue), where the background is divided into $c\bar{c}$ (green) and $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ (magenta). The dots with error bars show the distributions for data samples of the $B^- \to [K^-\pi^+]_D\pi^-$ in $|\Delta E| < 0.05$ GeV (red) and the sideband $\Delta E > 0.15$ GeV for the $B^- \to [K^+\pi^-]_D K^-$ (blue). The MC and data samples are in good agreement, while there are small discrepancies in several variables. All the requirements except for the one on NB are applied before showing the plots.

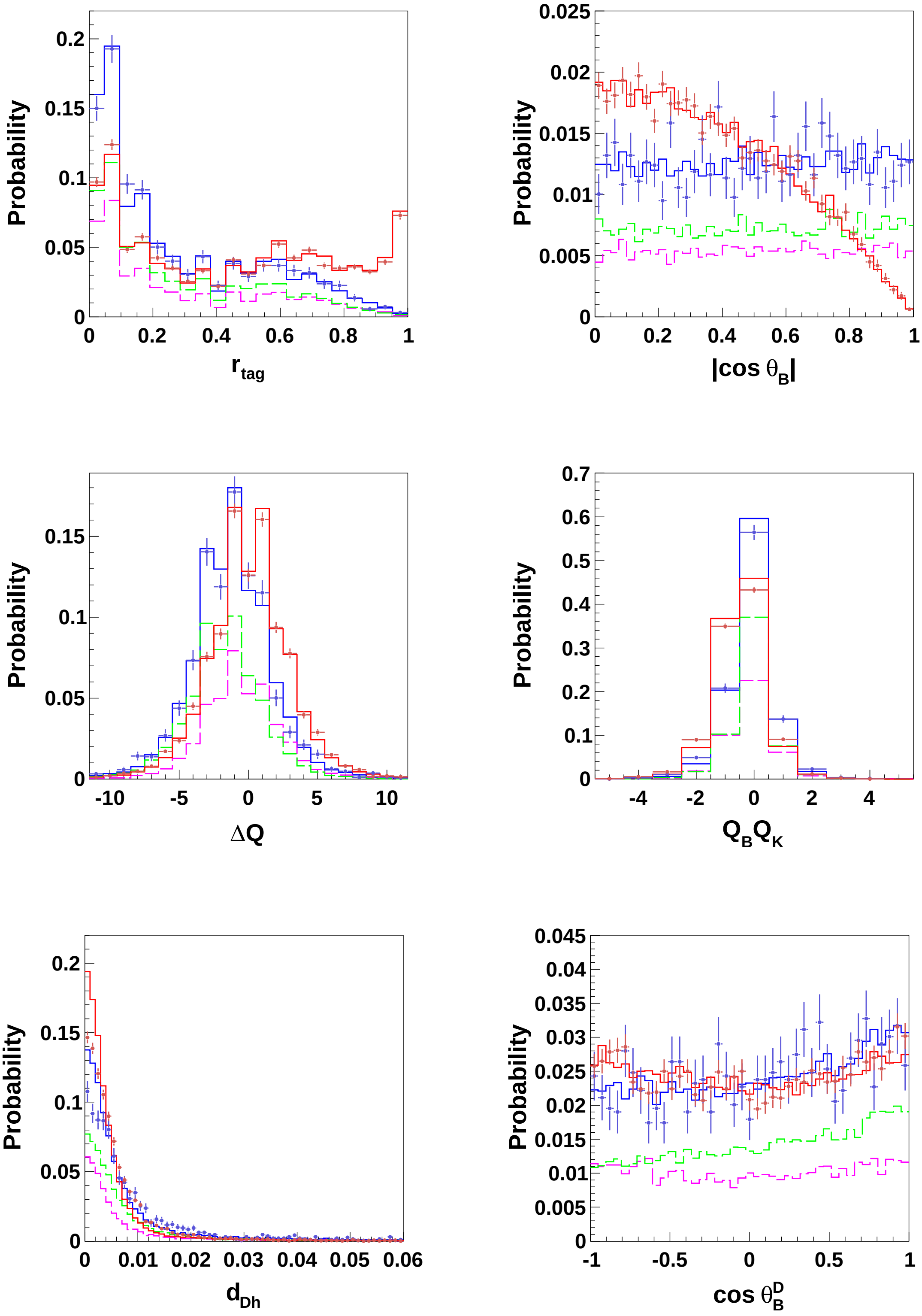
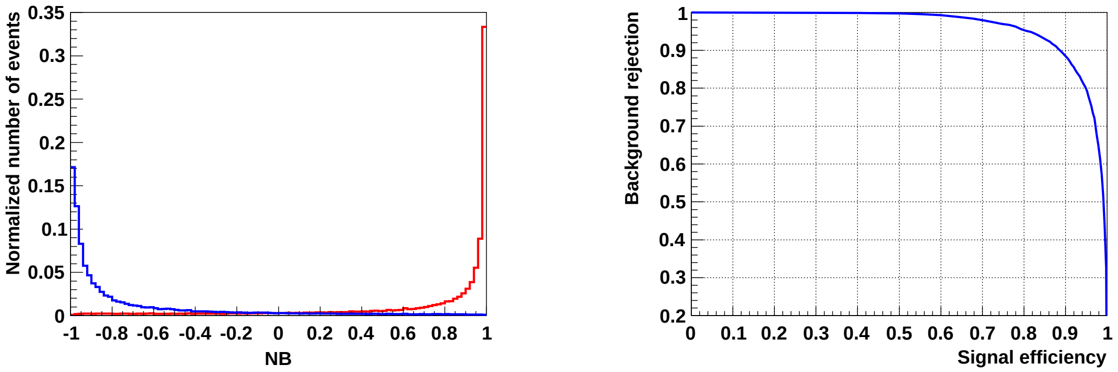


Figure 5.4: Continued.



(a) The *NB* distributions for the signal (red) and background (blue) MC samples.

(b) Plot of the background rejection and the signal efficiency for each requirement on *NB*.

Figure 5.5: Discrimination between the signal and the continuum background by the NeuroBayes method based on ten variables.

Background rejection (%)	Signal efficiency (%)		Rate for signal efficiency
	Likelihood Meth.	NeuroBayes Meth.	
99	42	60	1.43
95	69	81	1.17
90	79	88	1.11
80	88	94	1.07

Table 5.1: Summary table of the performance of the discrimination between the signal and the continuum background. The error of the efficiency is about 1%. The NeuroBayes method significantly improves the discrimination compared to the likelihood method.

5.4.1 Contribution of Each Variable

As a check of the training, we obtain the powers of individual variables as well as the correlation factors as in Table 5.2, where the explanations of contents are the following.

- Only this: the significance for single variable. The quantity is obtained as the correlation of a variable to the output multiplied by $\sqrt{n}$, where n is the sample size. The computation does not take into account other variables.

- Without this: the significance loss when a variable is removed. This is the loss of correlation multiplied by $\sqrt{n}$ when only one variable is removed from the input set and the total correlation to the output is recomputed.

- Corr. to others: the correlation of a variable to all the others, computed with the complete correlation matrix.

The values of "Only this" are higher for the variables which have the distributions clearly different between the signal and the background. The value of "Without this" is based on the "Only this", becoming lower if the value of "Corr. to others" is higher. A large correlation exists for the pair of $\mathcal{LR}$(KSFW) and $|\cos\theta_{\rm T}|$, since both variables utilize the event topology. Some correlations exist for several variables, all of which are reasonable since common particles are used in the definitions. All the values in Table 5.2 make sense according to the distributions in Figure 5.4 and the definitions.

Variable	Only this	Without this	Corr. to others
$\mathcal{LR}$(KSFW)	409	136	0.80
$\|\cos\theta_{\rm T}\|$	355	63	0.76
Δz	229	110	0.29
$\cos\theta_D^K$	189	76	0.27
$r_{\rm tag}$	186	56	0.38
$\|\cos\theta_B\|$	150	67	0.19
ΔQ	143	39	0.28
$Q_B Q_K$	142	46	0.30
d_{Dh}	141	60	0.21
$\cos\theta_B^D$	87	40	0.14

Table 5.2: The powers of individual variables and the correlation factors.

5.4.2 Check of Stability

To check whether the overtraining occurs or not, we apply the trained network to the MC samples independent of the training. Both the signal and background MC samples contain 100,000 events, where the same selection criteria as the training are applied. Figure 5.6 shows the distributions of resulting outputs as well as the ones from the training. The consistent results indicate that there is no significant overtraining.

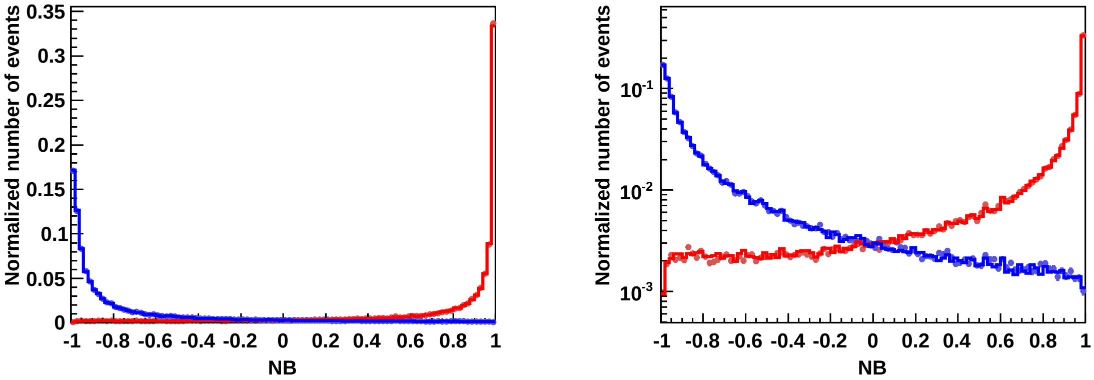
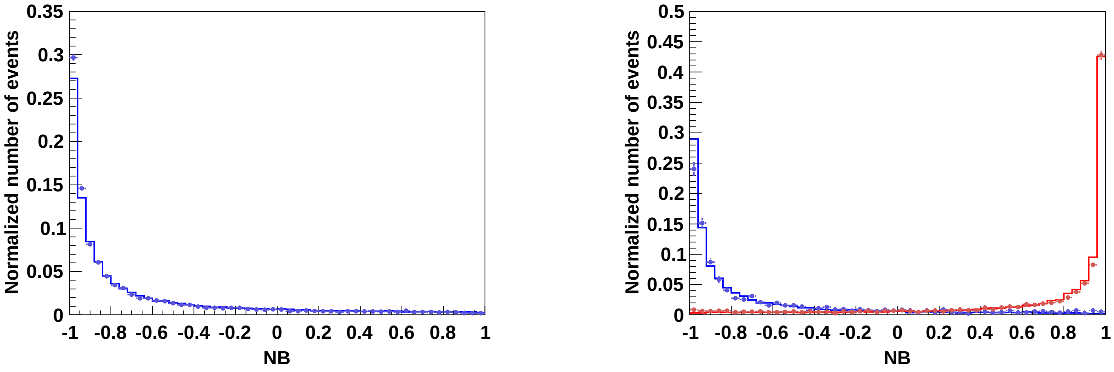


Figure 5.6: Check of the stability of the neural network. The *NB* distributions are shown for the training (histograms) and the test (dots with error bars) applied on the MC samples of the signal (red) and the background (blue). Linear (left) and logarithmic (right) scales are used for the vertical axes. The consistent results indicate that there is no significant overtraining.

5.5 Tests on Several Samples

In this section, we apply the trained network to several different samples. Figure 5.7 (a) shows the results for different $M_{\rm bc}$ requirements on a background MC sample. We apply $5.20\ {\rm GeV}/c^2 < M_{\rm bc} < 5.29\ {\rm GeV}/c^2$ and $5.271\ {\rm GeV}/c^2 < M_{\rm bc} < 5.287\ {\rm GeV}/c^2$, which correspond to the selections in the training and in the analysis, respectively. Similar distributions are obtained, validity of which is supported by the small correlations between the input variables and $M_{\rm bc}$. The result provides a confidence to use the wide $M_{\rm bc}$ region in the training.

Figure 5.7 (b) shows the results for the MC and data samples. For the latter, we use the $B^- \to [K^-\pi^+]_D\pi^-$ in $|\Delta E| < 0.05$ GeV for signal and the sideband $\Delta E > 0.15$ GeV of the $B^- \to [K^+\pi^-]_D K^-$ for background. Obtained *NB* distributions for MC and data samples are in good agreement. The small discrepancy is not due to the difference of the signal modes or the ΔE regions but due to the difference in MC and data. We will overcome this difficulty by using the data sample to obtain the functions for the fit.



(a) The *NB* distributions for 5.20 GeV/c^2 < M_{bc} < 5.29 GeV/c^2 (histogram) and 5.271 GeV/c^2 < M_{bc} < 5.287 GeV/c^2 (dots with error bars) for the background MC.

(b) The *NB* distributions for the signal (red) and the background (blue) obtained from MC (histograms) and data (dots with error bars) samples.

Figure 5.7: Comparison of the neural network outputs for several samples.

5.6 Test by Removing Three Variables

The variables Δz, $\cos\theta_D^K$, and d_{Dh} have some discrepancies between MC and data samples for the background (Figure 5.4). Figure 5.8 shows the results when the three variables are removed from the input set. The MC and data samples are in better agreement, which encourages us to be confident of the NeuroBayes method.

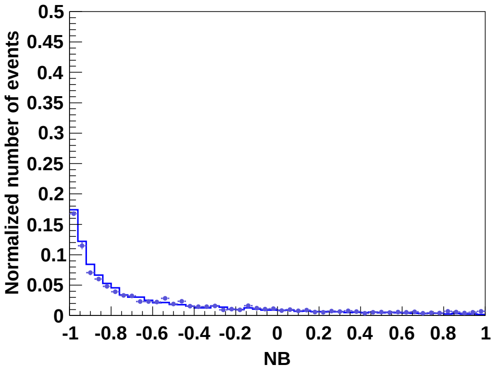


Figure 5.8: The output of the neural network for a MC sample of the background (histogram) and the data sample in the sideband $\Delta E > 0.15$ GeV (dots with error bars), where the variables Δz, $\cos\theta_D^K$, and d_{Dh} are removed from the input set.

Chapter 6

Signal Extraction

The signal yield is extracted from the two-dimensional distribution of ΔE and NB using the extended unbinned maximum likelihood fit. The fit is applied simultaneously for B^- and B^+ samples using the parameters $N_{B^-} + N_{B^+}$ and $(N_{B^-} - N_{B^+})/(N_{B^-} + N_{B^+})$, where $N_{B^\pm}$ is the yield of $B^\pm$, to obtain $\mathcal{R}_{Dh}$ and $\mathcal{A}_{Dh}$ directly. Also, the fit is applied simultaneously for DK and $D\pi$ samples, where the $D\pi$ (DK) "feed-across" in DK ($D\pi$) sample caused by the misidentification is parameterized by the efficiency difference in the two samples. We describe the probability density function (PDF) in Section 6.1 followed by the signal extraction in Section 6.2. Note that the fit on the data sample is not applied before all the strategy is fixed.

6.1 Probability Density Function (PDF)

The components of the fit are divided into the Dh signal, the Dh feed-across, the background from $e^+e^- \to B\bar{B}$ (denoted as $B\bar{B}$ background), and the continuum background from $e^+e^- \to q\bar{q}$ (denoted as $q\bar{q}$ background). For each component, the correlation between ΔE and NB is found to be small as in Figure 6.1. We thus obtain the PDF by taking a product of individual PDFs for ΔE and NB. The functions are obtained from data samples as many as possible to reduce the systematic uncertainties. A demonstration of the fit on MC samples is shown in Appendix B.

6.1.1 Functions for ΔE

The functions for ΔE are explained in the following.

- Dh signal: we use a sum of two Gaussians with a common mean. The parameters are floated in the fit to the favored modes, and fixed to the results in the fit to the suppressed modes. Common parameters are used for the DK and $D\pi$ signals, the validity of which is checked by MC samples as shown in Figure 6.2.

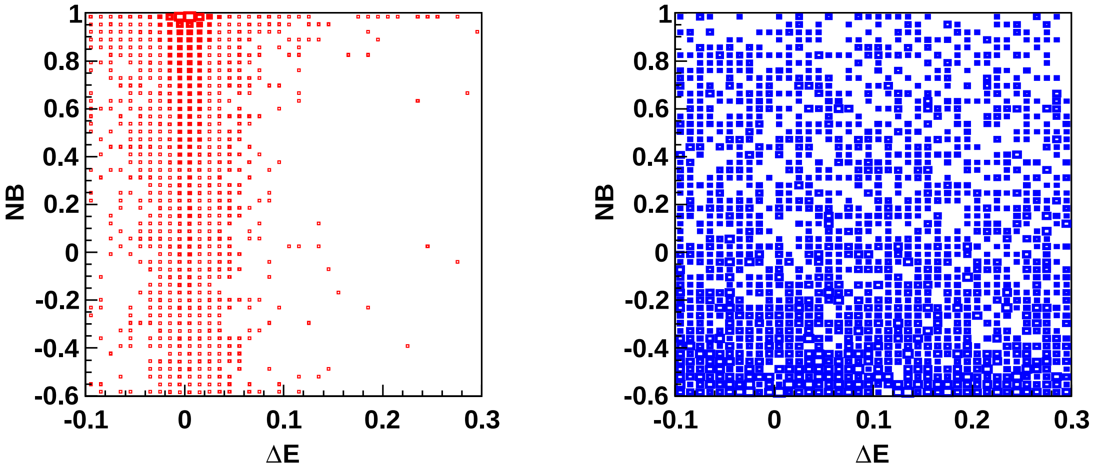


Figure 6.1: Two-dimensional distributions on ΔE and NB for MC samples of the Dh signal (left) and the $q\bar{q}$ background (right), where the correlation factors are -1.0% and 0.3%, respectively. The correlations for the Dh feed-across and the $B\bar{B}$ background are also very small, since the correlations to ΔE for main inputs in NB are less than 1%.

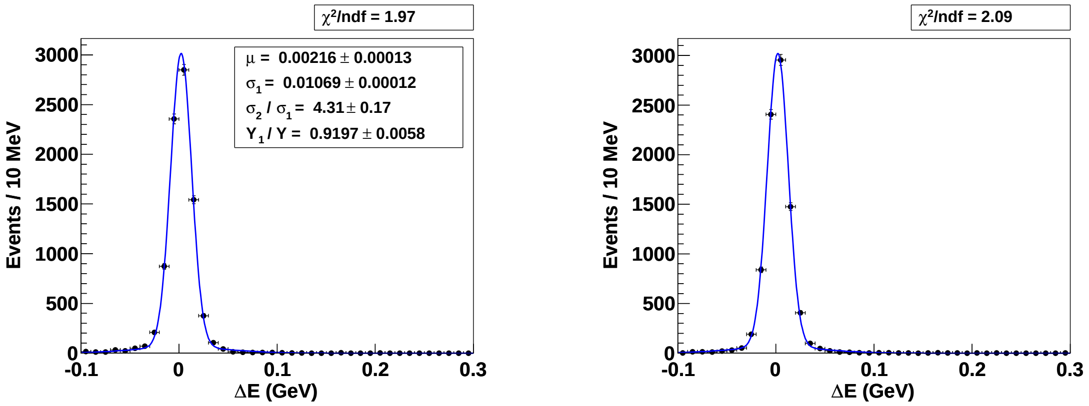
- Dh feed-across: we use a sum of an asymmetric Gaussian[1] and a Gaussian with a common mean. For $h = \pi$, the mean and the right-side width of the asymmetric Gaussian are floated in the fit to the favored modes, and fixed to the results in the fit to the suppressed modes. The remaining parameters are obtained from the fit to the favored $D\pi$ signal in the data sample where the kaon mass is assigned to the prompt pion, as shown in Figure 6.3 (a). For $h = K$, the parameters are obtained from a MC sample as shown in Figure 6.3 (b), since the DK signal is rather smaller than the $D\pi$ signal and the data sample cannot be used.

- $B\bar{B}$ background: we use a free exponential function, the validity of which is checked by MC samples as shown in Figure 6.4. In this component, there are contributions from $B^- \to D^*h^-$, $B^- \to D\rho^-$, and so on, which populate the negative ΔE region. An exponential function is suited to model their tails as well as the combinatorial contribution located in whole region.

- $q\bar{q}$ background: we use a free linear function, the validity of which is checked by MC samples as shown in Figure 6.5. The validity of using a linear function is also checked by a data sample of the NB sideband: $NB < 0$.

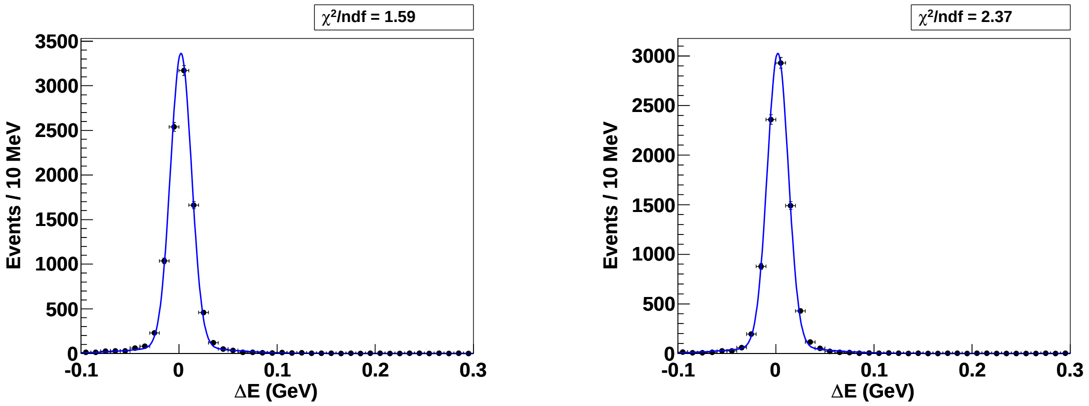
---

[1] The shift caused by the incorrect mass assignment makes the shape of the ΔE distribution asymmetric.



(a) The ΔE distribution for a signal MC sample of $B^- \rightarrow [K^-\pi^+]_D\pi^-$.

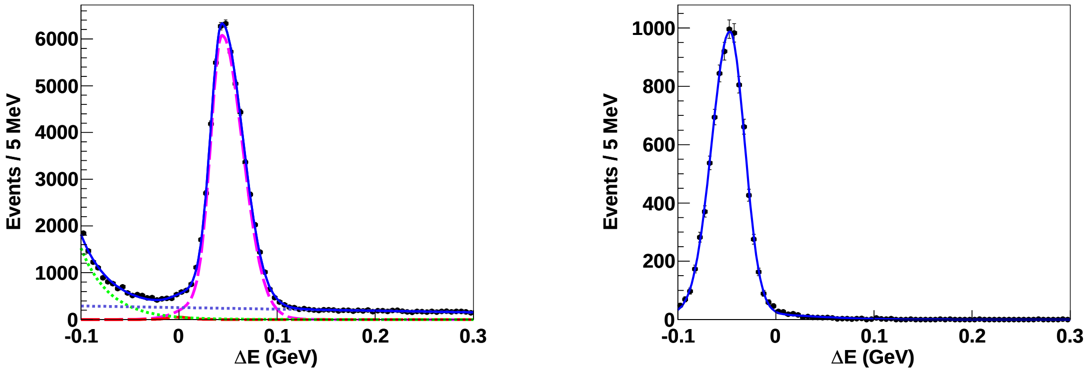
(b) The ΔE distribution for a signal MC sample of $B^- \rightarrow [K^-\pi^+]_D K^-$.



(c) The ΔE distribution for a signal MC sample of $B^- \rightarrow [K^+\pi^-]_D\pi^-$.

(d) The ΔE distribution for a signal MC sample of $B^- \rightarrow [K^+\pi^-]_D K^-$.

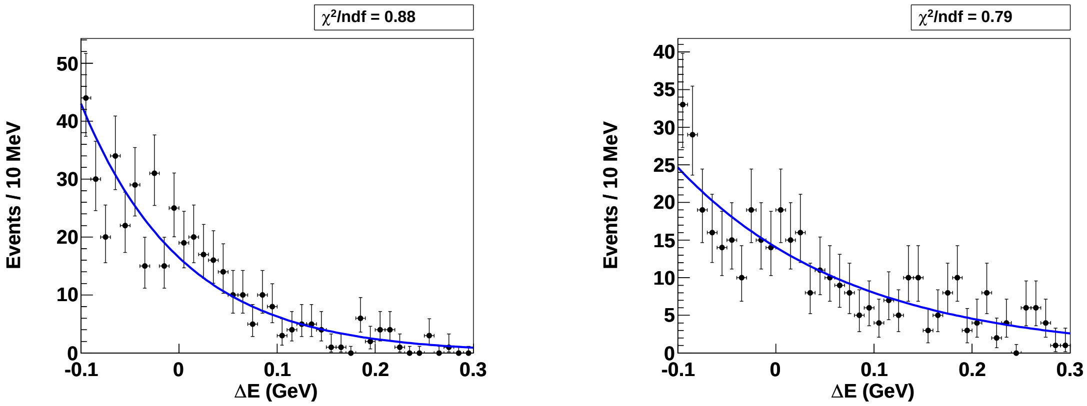
Figure 6.2: The ΔE distributions for signal MC samples fitted with a sum of two Gaussians. The shape parameters are floated in the fit to the mode $B^- \rightarrow [K^-\pi^+]_D\pi^-$, and fixed to the results for the other modes.



(a) The ΔE distribution for the data sample of $B^- \to [K^-\pi^+]_D\pi^-$, where the kaon-mass hypothesis is assigned for the prompt pion. The dashed curves show $D\pi$ (magenta), DK (red), $q\bar{q}$ (blue), and $B\bar{B}$ (green).

(b) The ΔE distribution for a MC sample of $B^- \to [K^-\pi^+]_D K^-$, where the pion-mass hypothesis is assigned for the prompt kaon.

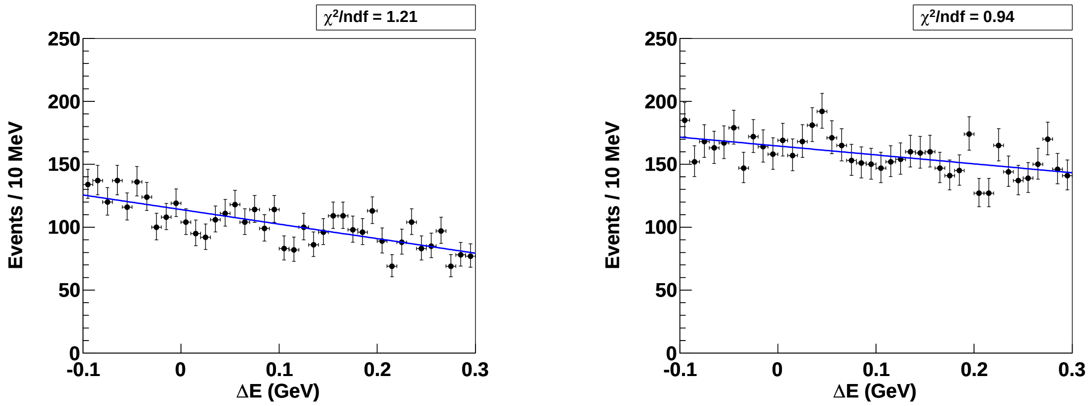
Figure 6.3: The ΔE distributions for the Dh feed-across samples. A fit is applied to obtain the values of the shape parameters.



(a) The ΔE distribution for a MC sample of the $B\bar{B}$ background for $B^- \to [K^+\pi^-]_D\pi^-$.

(b) The ΔE distribution for a MC sample of the $B\bar{B}$ background for $B^- \to [K^+\pi^-]_D K^-$.

Figure 6.4: The ΔE distributions for MC samples of the $B\bar{B}$ background. Each sample has two times larger size than the data sample. A fit is applied using free exponential function.



(a) The ΔE distribution for a MC sample of the $q\bar{q}$ background for $B^- \to [K^+\pi^-]_D\pi^-$.

(b) The ΔE distribution for a MC sample of the $q\bar{q}$ background for $B^- \to [K^+\pi^-]_D K^-$.

Figure 6.5: The ΔE distributions for MC samples of the $q\bar{q}$ background. Each sample has two times larger size than the data sample. A fit is applied using free linear function. The smaller yield for $B^- \to [K^+\pi^-]_D\pi^-$ is due to the two same-sign pions needed for the reconstruction.

6.1.2 Functions for NB

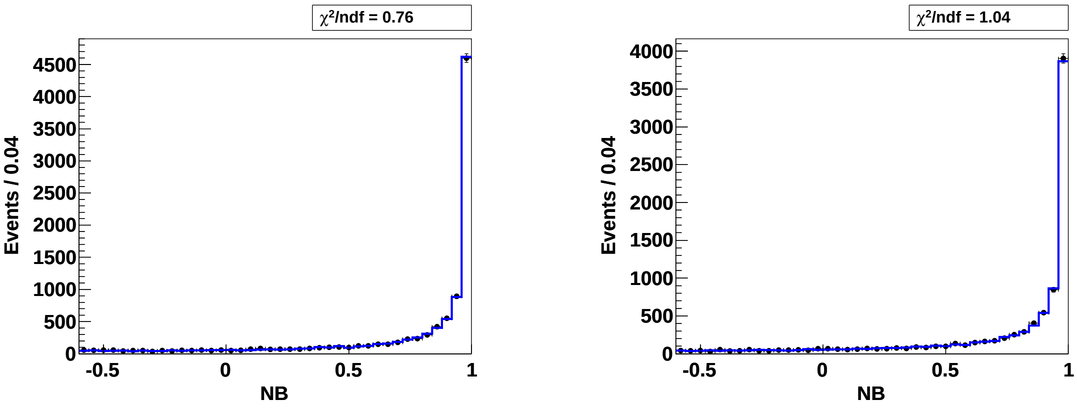
The functions for NB are explained in the following. The histogram functions are employed for all components.

- Dh signal: we obtain the shape for the favored signal from the data sample. The histogram in the region $|\Delta E| < 0.01$ GeV is used by subtracting the one for the region 0.15 GeV $< \Delta E <$ 0.30 GeV for which we multiply the weight 0.02/0.15. The same shape is used for the suppressed signal, validity of which is checked by MC samples as shown in Figure 6.6.

- Dh feed-across: we obtain the shape for $h = \pi$ by a similar way as the one for the Dh signal, where the region $|\Delta E - 0.05\ \text{GeV}| < 0.01$ GeV is used. We obtain the shape for $h = K$ from a MC sample, since the DK signal is rather smaller than the $D\pi$ signal and the data sample cannot be used.

- $B\bar{B}$ background: we obtain the shape from MC samples.

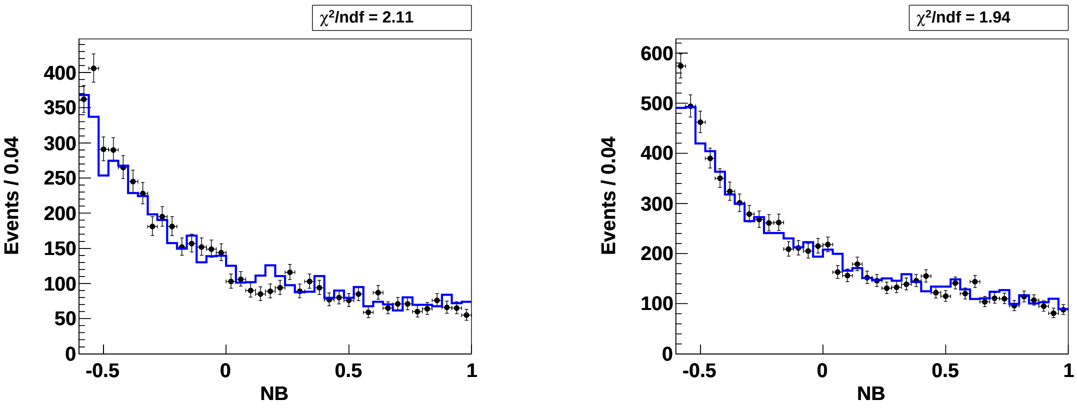
- $q\bar{q}$ background: we obtain the shape from the data sample of the sideband for which we require 5.20 GeV/c^2 $< M_{\text{bc}} <$ 5.26 GeV/c^2 and 0.15 GeV $< \Delta E <$ 0.30 GeV. Figure 6.7 shows the NB distributions for MC samples. There is small contribution from the $B\bar{B}$ background for the data sample (0.5%–1.0% depending on the modes), which is subtracted according to a MC study.



(a) The NB distribution for $B^- \to D\pi^-$.

(b) The NB distribution for $B^- \to DK^-$.

Figure 6.6: The NB distributions for MC samples of the suppressed modes fitted with histogram functions obtained from MC samples of the favored modes.



(a) The NB distribution for a MC sample of the $q\bar{q}$ background for $B^- \to [K^+\pi^-]_D\pi^-$.

(b) The NB distribution for a MC sample of the $q\bar{q}$ background for $B^- \to [K^+\pi^-]_D K^-$.

Figure 6.7: The NB distributions for MC samples of the $q\bar{q}$ background. Each sample has two times larger size than the data sample. The fit is applied using a histogram function obtained from the sideband of the MC sample.

6.2 Signal Extraction

The signal yield is extracted with the PDF described in the previous section. The fit is applied separately for the favored modes and the suppressed modes.

6.2.1 Fit to Favored Modes

Figure 6.8 shows the result of the fit to the favored modes. In this fit, the ratio of the DK yields in the $D\pi$ and DK samples is fixed to the value obtained from the decay $D^{*-} \to \bar{D}^0\pi^-$ followed by $\bar{D}^0 \to K^+\pi^-$, and the asymmetry for the $q\bar{q}$ background is fixed to zero. Figures 6.9 and 6.10 show the projections for several regions, where good quality of the fit is obtained for each region. Table 6.1 shows the list of the parameters in the fit.

The yield for the $D\pi$ (DK) signal is consistent with the expectation from the previous analysis [19]; the increase of the number of events is consistent with the factor 1.8 (2.6) which is obtained by multiplying the increases of the efficiency and the data.[2] The asymmetry for each component is very close to zero, as expected. Note that the ratio of the $D\pi$ yields in the DK and $D\pi$ samples is 1.70 ± 0.04 times larger than the one for the MC samples (Table B.1), which is consistent with the value 1.69 ± 0.10 obtained from the decay $D^{*-} \to \bar{D}^0\pi^-$ followed by $\bar{D}^0 \to K^+\pi^-$.

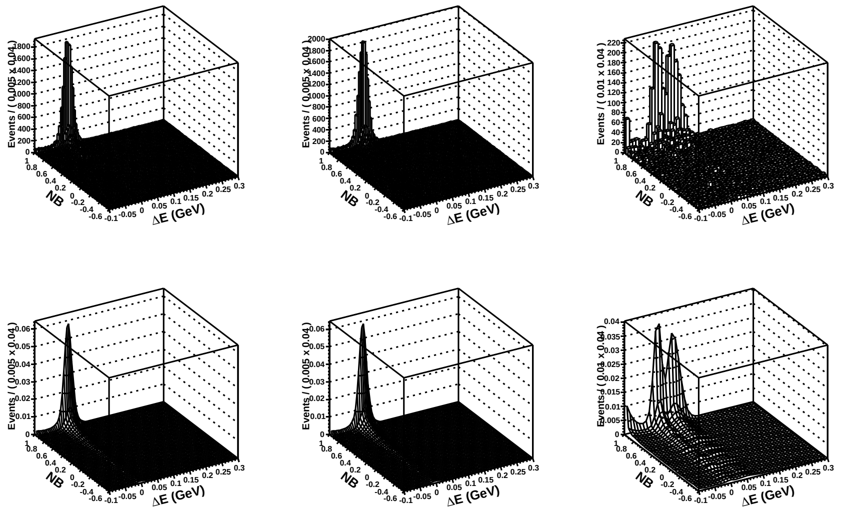
6.2.2 Fit to Suppressed Modes

Figure 6.11 shows the result of the fit to the suppressed modes. In this fit, all shape parameters for the Dh signal and the ratio of the $D\pi$ yields in the DK and $D\pi$ samples are fixed to the results for the favored modes. Figures 6.12 and 6.13 show the projections for several regions, where good quality of the fit is obtained for each region. Table 6.2 shows the list of the parameters in the fit.

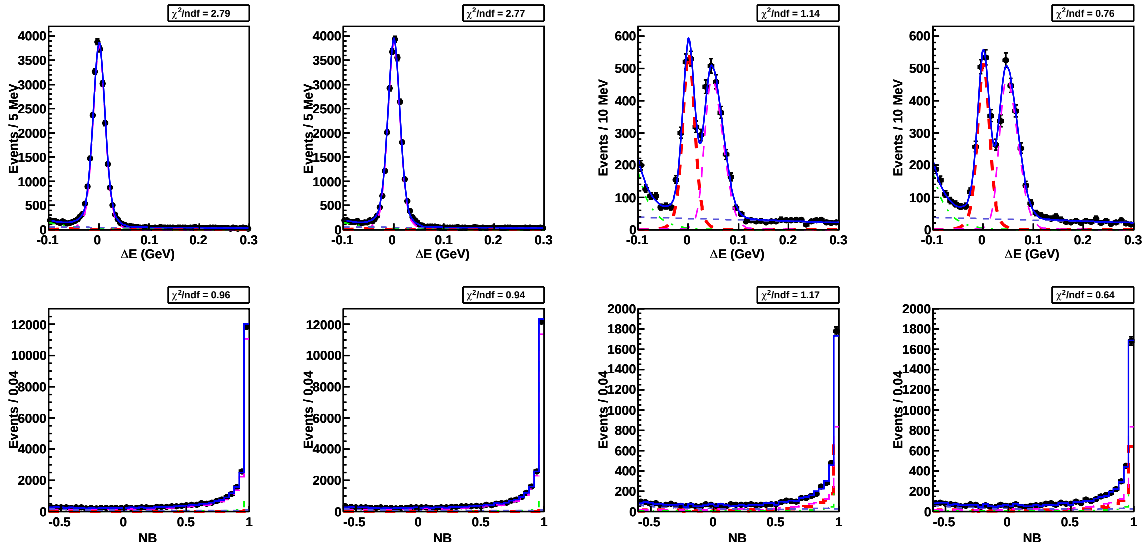
The yield and asymmetry for the $D\pi$ signal are 165^{+19}_{-18} and $(-4 \pm 11)\%$, respectively, which are consistent with the expected values 169 and -2%, respectively, obtained from the previous analysis [19]. The yield for the DK signal is $56.0^{+15.1}_{-14.2}$, for which the statistical significance is 4.4σ. The significance including the systematic error is obtained in Section 8.2. The asymmetry for the DK signal is $(-39^{+26}_{-28})\%$, which is consistent with the value $(-86^{+49}_{-50})\%$ obtained by BaBar [20]. Assuming $r_B = 0.1$, the yield (asymmetry) for the DK signal varies from 5 to 87 (varies in most of the region from -100% to 100%) depending on the values of ϕ_3 and the strong phases. Therefore, our measurement provides important information for measuring ϕ_3. Note that the asymmetry for the $B\bar{B}$ background in the DK sample, for which non-zero value can be generated by $B^- \to D^*K^-$, is $(30 \pm 17)\%$.

---

[2] The difference of the efficiency increase between the $D\pi$ and DK analyses is mainly due to the different requirements for rejecting the $q\bar{q}$ background in the previous analysis.

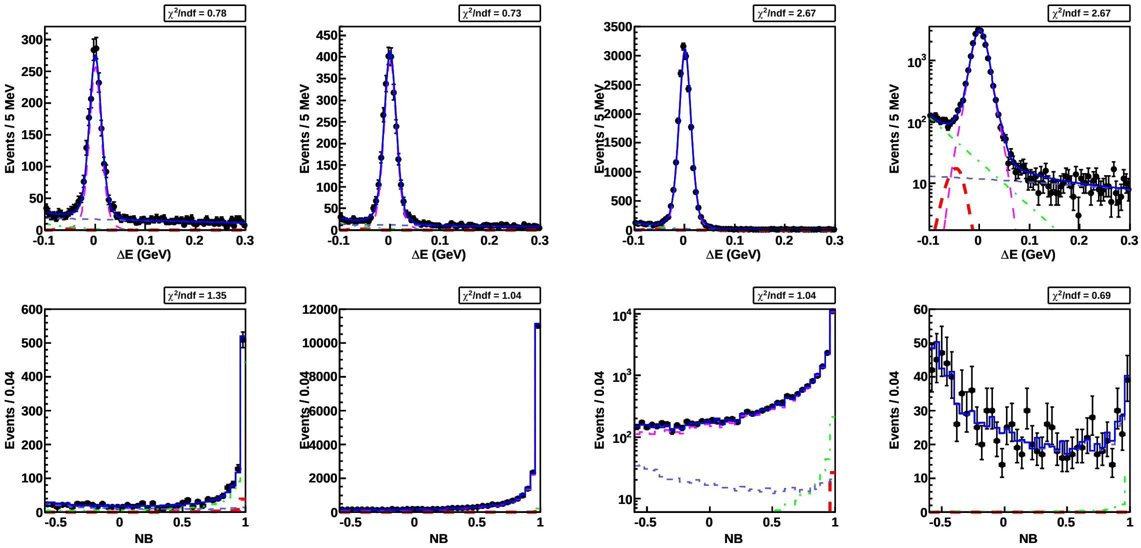


(a) The histograms of the data samples (upper) and the fitted PDFs (lower).

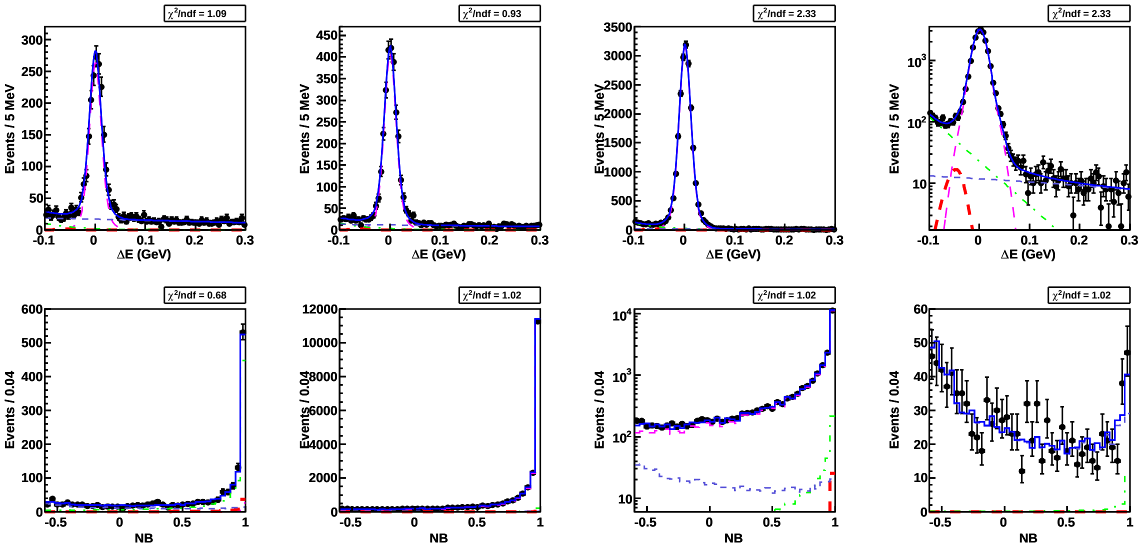


(b) The ΔE distributions (upper) and the NB distributions (lower), both of which are obtained by projecting all fitted regions. The fitted sample is shown with dots with error bars and the PDF is shown with the solid blue curve, for which the components are shown with thicker long-dashed red (DK), thinner long-dashed magenta ($D\pi$), dash-dotted green ($B\bar{B}$ background), and dashed blue ($q\bar{q}$ background).

Figure 6.8: The result of the fit to the favored modes. The distributions for $D\pi^-$, $D\pi^+$, DK^-, and DK^+ are shown from left to right.

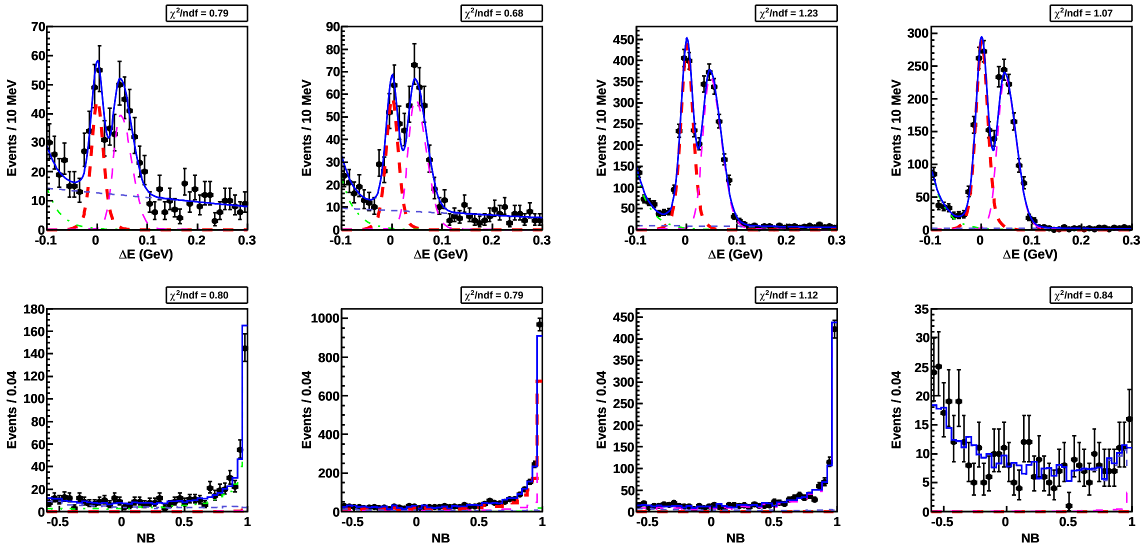


(a) The projections for the $D\pi^-$ mode.

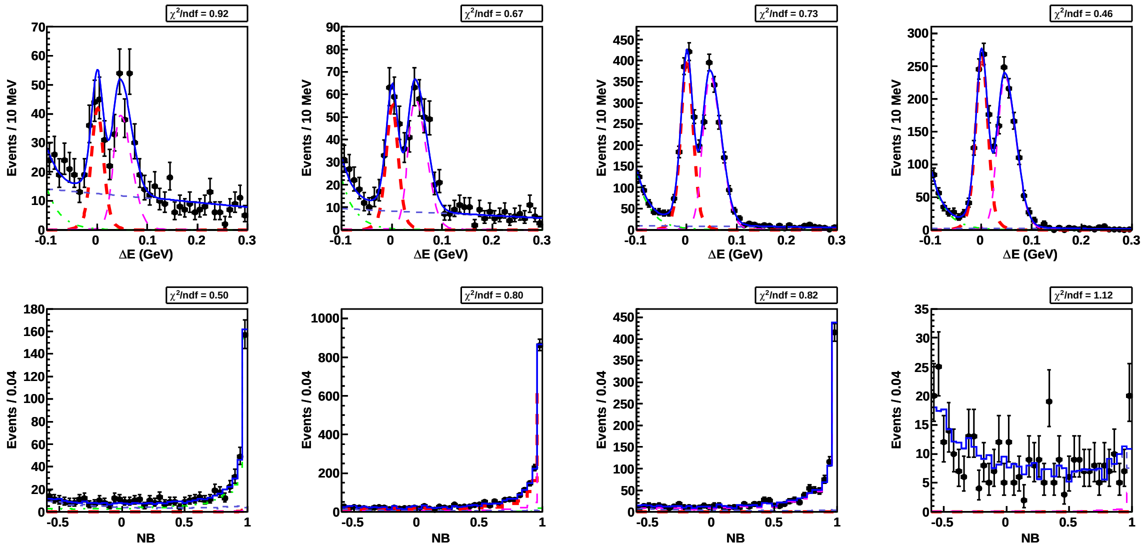


(b) The projections for the $D\pi^+$ mode.

Figure 6.9: The projections for the favored $D\pi$ sample. The ΔE distributions for $NB < 0$, $0 < NB < 0.5$, $NB > 0.5$, and $NB > 0.5$ (log-scaled) are shown from left to right (upper). The NB distributions for $\Delta E < -0.05$ GeV, $|\Delta E| < 0.04$ GeV, $|\Delta E| < 0.04$ GeV (log-scaled), and $\Delta E > 0.15$ GeV are shown from left to right (lower).



(a) The projections for the DK^- mode.

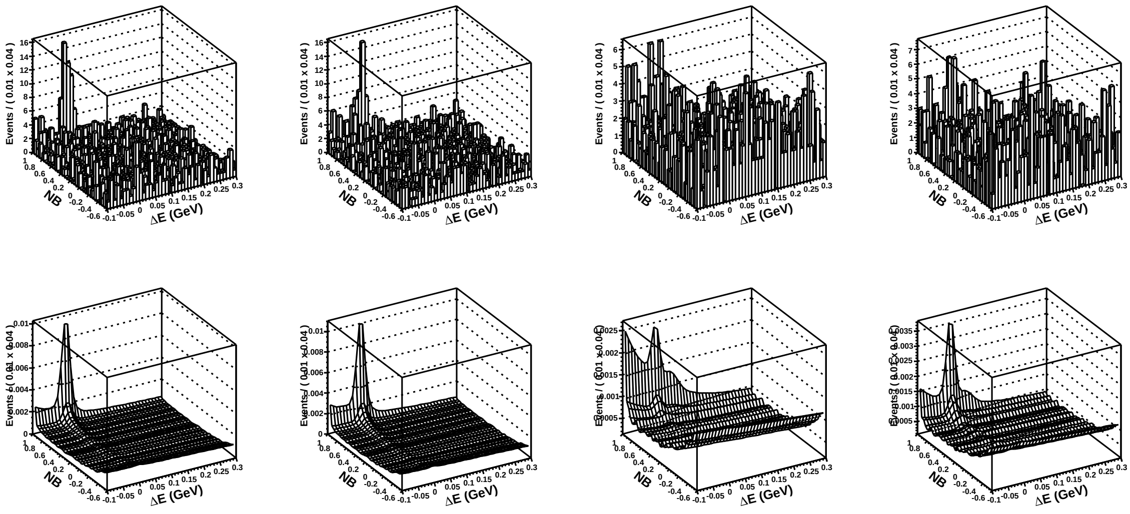


(b) The projections for the DK^+ mode.

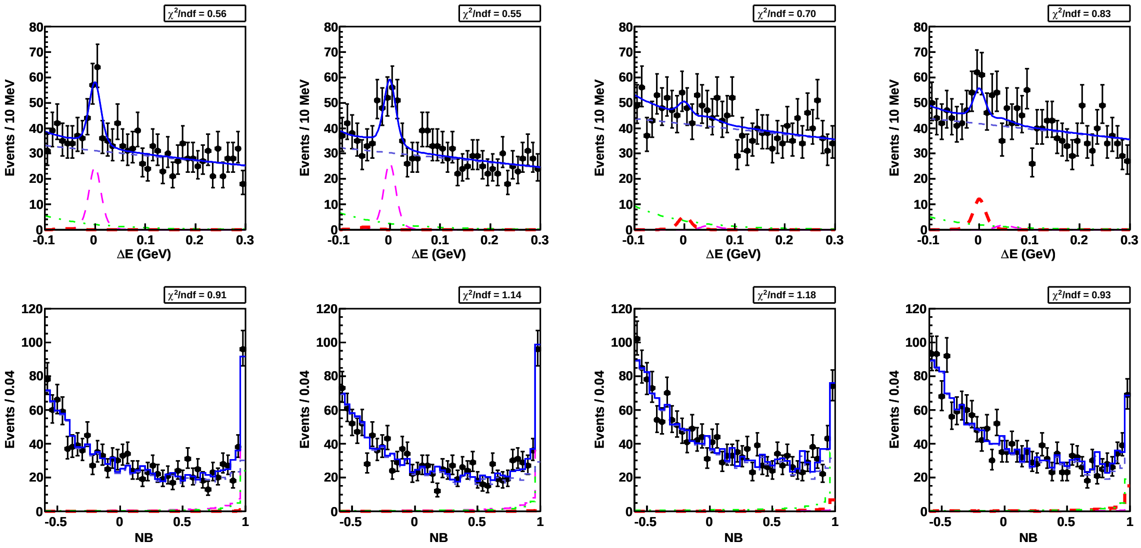
Figure 6.10: The projections for the favored DK sample. The ΔE distributions for $NB < 0$, $0 < NB < 0.5$, $NB > 0.5$, and $NB > 0.9$ are shown from left to right (upper). The NB distributions for $\Delta E < -0.05$ GeV, $|\Delta E| < 0.04$ GeV, 0.05 GeV $< \Delta E < 0.1$ GeV, and $\Delta E > 0.15$ GeV are shown from left to right (lower).

Component	Type	Parameter	Value
$D\pi$	General	Yield	49164^{+245}_{-244}
		Asymmetry	$(-0.97 \pm 0.45)\%$
	D. Gaussian (ΔE)	Mean	0.0008 ± 0.0001
		σ_1	0.0111 ± 0.0002
		σ_2/σ_1	$1.92^{+0.05}_{-0.04}$
		$\text{Yield}_1/\text{Yield}$	$(72.6^{+2.5}_{-2.6})\%$
DK	General	Yield	3394^{+68}_{-69}
		Asymmetry	$(2.3 \pm 1.9)\%$
$D\pi$ in DK	General	Yield in DK/Yield in $D\pi$	$(8.93 \pm 0.16)\%$
	D. Bif. Gaussian (ΔE)	Mean	0.0466 ± 0.0004
		$\sigma_{1,\text{right}}$	0.0223 ± 0.0004
		$\sigma_{1,\text{left}}/\sigma_{1,\text{right}}$	0.631 (fixed)
		$\sigma_2/\sigma_{1,\text{right}}$	3.49 (fixed)
		$\text{Yield}_1/\text{Yield}$	95.8% (fixed)
DK in $D\pi$	General	Yield in $D\pi$/Yield in DK	11.2% (fixed)
	D. Bif. Gaussian (ΔE)	Mean	-0.0463 (fixed)
		$\sigma_{1,\text{right}}$	0.0139 (fixed)
		$\sigma_{1,\text{left}}/\sigma_{1,\text{right}}$	1.29 (fixed)
		$\sigma_2/\sigma_{1,\text{right}}$	4.77 (fixed)
		$\text{Yield}_1/\text{Yield}$	85.1% (fixed)
$B\bar{B}$ in $D\pi$	General	Yield	3525^{+131}_{-133}
		Asymmetry	$(-0.4 \pm 2.5)\%$
	Exponential (ΔE)	Coefficient	$-16.6^{+0.6}_{-0.7}$
$B\bar{B}$ in DK	General	Yield	973^{+49}_{-47}
		Asymmetry	$(0.1 \pm 3.9)\%$
	Exponential (ΔE)	Coefficient	$-36.4^{+2.6}_{-2.7}$
$q\bar{q}$ in $D\pi$	General	Yield	6442^{+115}_{-113}
		Asymmetry	0 (fixed)
	Linear (ΔE)	Slope	-0.25 ± 0.03
$q\bar{q}$ in DK	General	Yield	2379^{+72}_{-71}
		Asymmetry	0 (fixed)
	Linear (ΔE)	Slope	-0.28 ± 0.04

Table 6.1: The list of parameters in the fit to the favored modes.

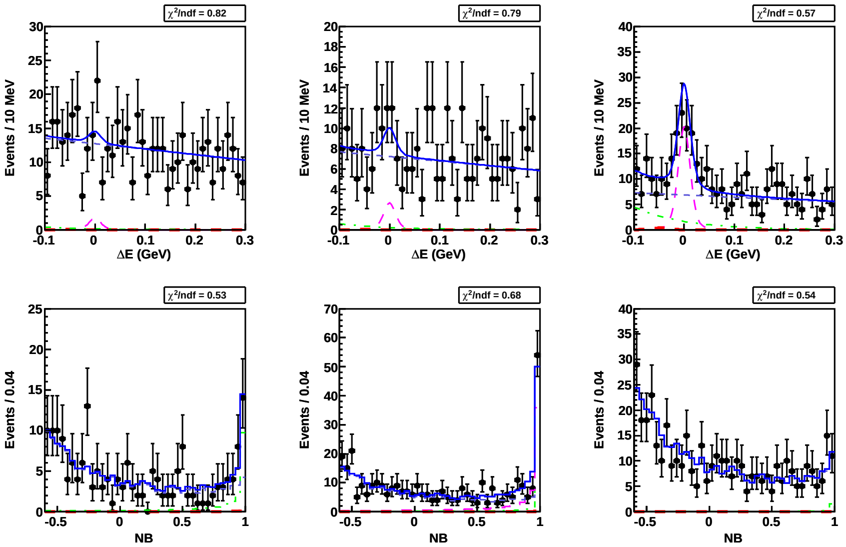


(a) The histograms of the data samples (upper) and the fitted PDFs (lower).

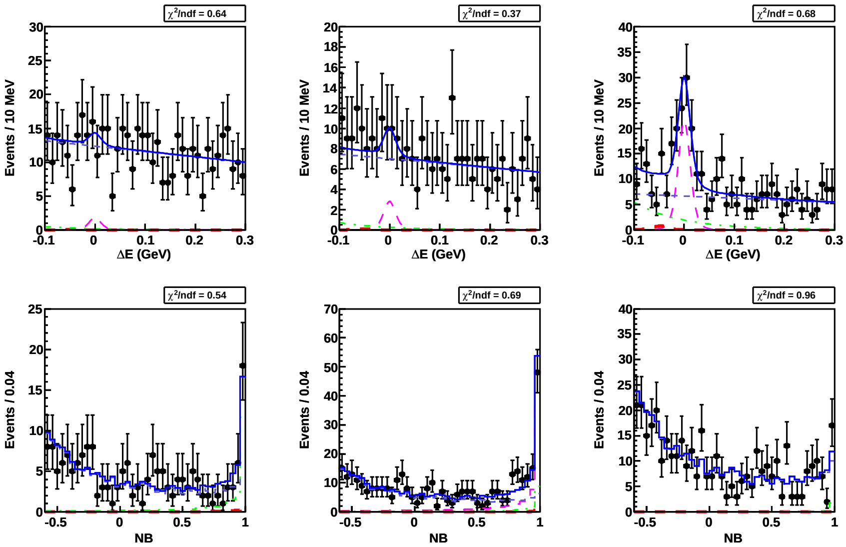


(b) The ΔE distributions (upper) and the NB distributions (lower), both of which are obtained by projecting all fitted regions. The fitted sample is shown with dots with error bars and the PDF is shown with the solid blue curve, for which the components are shown with thicker long-dashed red (DK), thinner long-dashed magenta ($D\pi$), dash-dotted green ($B\bar{B}$ background), and dashed blue ($q\bar{q}$ background).

Figure 6.11: The result of the fit to the suppressed modes. The distributions for $D\pi^-$, $D\pi^+$, DK^-, and DK^+ are shown from left to right.

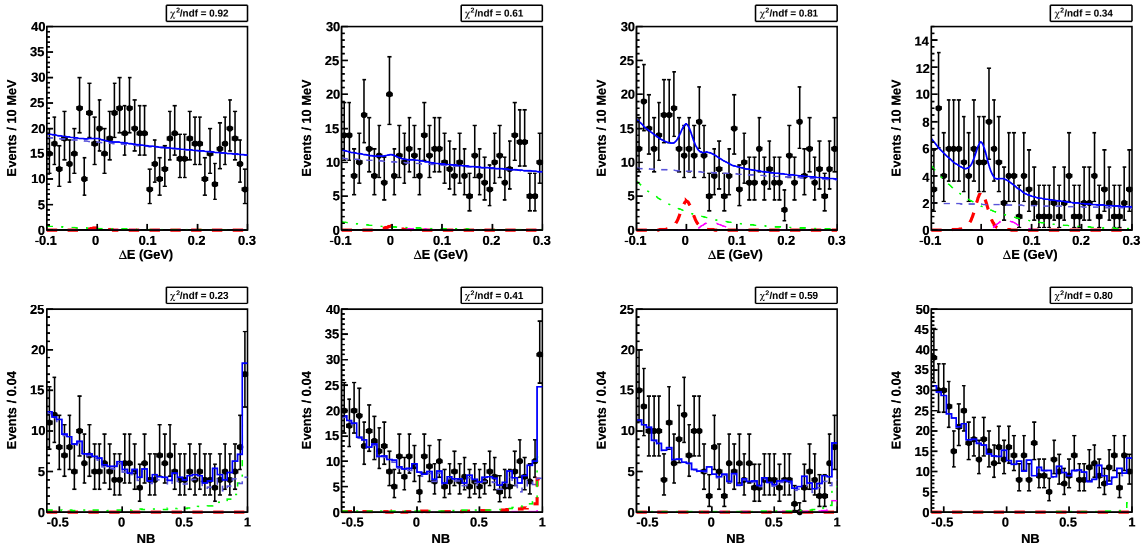


(a) The projections for the $D\pi^-$ mode.

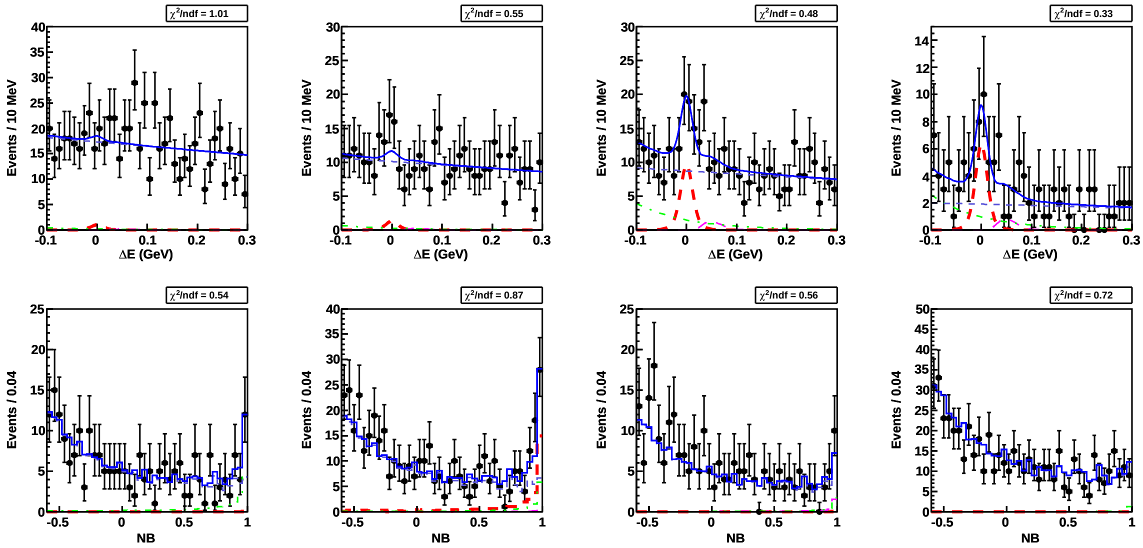


(b) The projections for the $D\pi^+$ mode.

Figure 6.12: The projections for the suppressed $D\pi$ sample. The ΔE distributions for $NB < 0$, $0 < NB < 0.5$, and $NB > 0.5$ are shown from left to right (upper). The NB distributions for $\Delta E < -0.05$ GeV, $|\Delta E| < 0.04$ GeV, and $\Delta E > 0.15$ GeV are shown from left to right (lower).



(a) The projections for the DK^- mode.



(b) The projections for the DK^+ mode.

Figure 6.13: The projections for the suppressed DK sample. The ΔE distributions for $NB < 0$, $0 < NB < 0.5$, $NB > 0.5$, and $NB > 0.9$ are shown from left to right (upper). The NB distributions for $\Delta E < -0.05$ GeV, $|\Delta E| < 0.04$ GeV, 0.05 GeV $< \Delta E <$ 0.1 GeV, and $\Delta E > 0.15$ GeV are shown from left to right (lower).

Component	Type	Parameter	Value
$D\pi$	General	Yield	$165.0^{+19.1}_{-18.1}$
		Asymmetry	$(-4.3^{+11.0}_{-10.9})\%$
	D. Gaussian (ΔE)	Mean	0.0008 (fixed)
		σ_1	0.0111 (fixed)
		σ_2/σ_1	1.92 (fixed)
		$\text{Yield}_1/\text{Yield}$	72.6% (fixed)
DK	General	Yield	$56.0^{+15.1}_{-14.2}$
		Asymmetry	$(-38.6^{+25.5}_{-27.8})\%$
$D\pi$ in DK	General	Yield in DK/Yield in $D\pi$	8.93% (fixed)
	D. Bif. Gaussian (ΔE)	Mean	0.0466 (fixed)
		$\sigma_{1,\text{right}}$	0.0223 (fixed)
		$\sigma_{1,\text{left}}/\sigma_{1,\text{right}}$	0.631 (fixed)
		$\sigma_2/\sigma_{1,\text{right}}$	3.49 (fixed)
		$\text{Yield}_1/\text{Yield}$	95.8% (fixed)
DK in $D\pi$	General	Yield in $D\pi$/Yield in DK	11.2% (fixed)
	D. Bif. Gaussian (ΔE)	Mean	-0.0463 (fixed)
		$\sigma_{1,\text{right}}$	0.0139 (fixed)
		$\sigma_{1,\text{left}}/\sigma_{1,\text{right}}$	1.29 (fixed)
		$\sigma_2/\sigma_{1,\text{right}}$	4.77 (fixed)
		$\text{Yield}_1/\text{Yield}$	85.1% (fixed)
$B\bar{B}$ in $D\pi$	General	Yield	$118.0^{+28.1}_{-27.8}$
		Asymmetry	$(-10.2^{+18.4}_{-18.5})\%$
	Exponential (ΔE)	Coefficient	$-10.2^{+2.8}_{-4.5}$
$B\bar{B}$ in DK	General	Yield	$142.7^{+30.1}_{-28.3}$
		Asymmetry	$(29.7^{+17.4}_{-17.1})\%$
	Exponential (ΔE)	Coefficient	$-9.9^{+2.3}_{-2.9}$
$q\bar{q}$ in $D\pi$	General	Yield	2301^{+54}_{-53}
		Asymmetry	0 (fixed)
	Linear (ΔE)	Slope	-0.14 ± 0.04
$q\bar{q}$ in DK	General	Yield	3171 ± 61
		Asymmetry	0 (fixed)
	Linear (ΔE)	Slope	-0.11 ± 0.03

Table 6.2: The list of parameters in the fit to the suppressed modes.

Chapter 7

Estimation of Peaking Backgrounds

In this chapter, we describe the backgrounds which peak inside the signal window. Since all the backgrounds have approximately flat $M(K\pi)$ distributions, the yields of the backgrounds are estimated by using $M(K\pi)$ sidebands of data. We use 1.815 GeV/c^2 < $M(K\pi)$ < 1.845 GeV/c^2 and 1.885 GeV/c^2 < $M(K\pi)$ < 2.005 (1.915) GeV/c^2 for DK^- ($D\pi^-$). The regions are taken to avoid the effects of the misidentification backgrounds from $B^- \to [K^+K^-]_D h^-$ and $B^- \to [\pi^+\pi^-]_D \pi^-$.

7.1 Peaking Backgrounds

In this section, we list the peaking backgrounds, where the $M(K\pi)$ distribution is shown for each component.

- The decay $B^- \to [K^+K^-]_D \pi^-$ remaining after the veto on $M(KK)$, described in Section 4.4, can have a peak under the signal $B^- \to [K^+\pi^-]_D K^-$. In Figure 7.1, the $M(K\pi)$ distribution is shown, where one million MC events are analyzed without the $M(K\pi)$ requirement. A constant function fits well, which indicates the validity of using the $M(K\pi)$ sidebands to estimate the background contribution.

- The favored decay $B^- \to [K^-\pi^+]_D h^-$ remaining after the veto on $M(K\pi)_{\mathrm{exchanged}}$, described in Section 4.4, can have a peak under the signal $B^- \to [K^+\pi^-]_D h^-$. In Figure 7.2, the $M(K\pi)$ distribution for $h = K$ (similar for $h = \pi$) is shown, where one million MC events are analyzed without the $M(K\pi)$ requirement. A constant function fits well, which indicates the validity of using the $M(K\pi)$ sidebands to estimate the background contribution.

- The charmless decay $B^- \to K^+K^-\pi^-$ can also be a peaking background for the signal $B^- \to [K^+\pi^-]_D K^-$.[1] In Figure 7.3, the $M(K\pi)$ distribution is shown, where one million MC events are analyzed without the $M(K\pi)$ requirement. For the MC generation, we take no interference and use a sum of $f_0\pi^-$ (14.9%), $K^{*0}K^-$

---

[1] The decay rate of $B^- \to K^+\pi^-\pi^-$ is thought to be much smaller and the signal has not seen yet.

(6.9%), $K_0^{*0}K^-$ (24.5%), and non-resonant $K^+K^-\pi^-$ (53.7%). This model is based on a theoretical calculation [32], which is consistent with the observation by BaBar [33]. A constant function fits well, which indicates the validity of using the $M(K\pi)$ sidebands to estimate the background contribution.

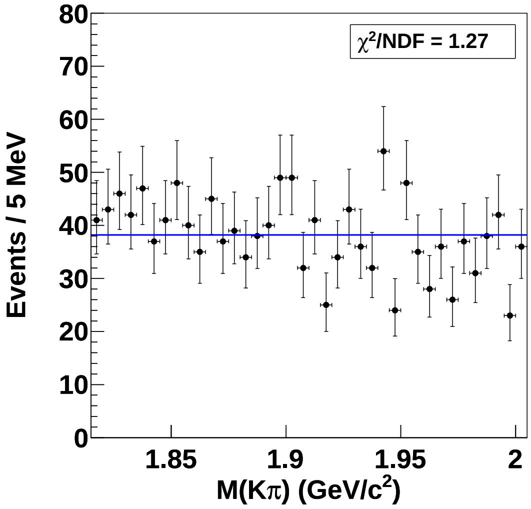


Figure 7.1: The $M(K\pi)$ distribution fitted by a constant function for a MC sample of the background $B^- \to [K^+K^-]_D\pi^-$ in the analysis of $B^- \to [K^+\pi^-]_DK^-$.

7.2 Estimation Using the $M(K\pi)$ Sidebands

For estimating the contributions from the peaking backgrounds, we use the events in the $M(K\pi)$ sidebands of data. Figures 7.4 and 7.5 show the results of the fits, where the same method as the one for the signal extraction is applied. By taking into account that the sideband region is five (two) times larger than the signal region, we obtain an expected background yield of $-1.9^{+3.7}_{-3.5}$ ($-3.2^{+7.0}_{-6.4}$) events for DK^- ($D\pi^-$). Since the yields are consistent with zero, we do not subtract these backgrounds from the signal yields but instead include the uncertainties in the systematic errors.

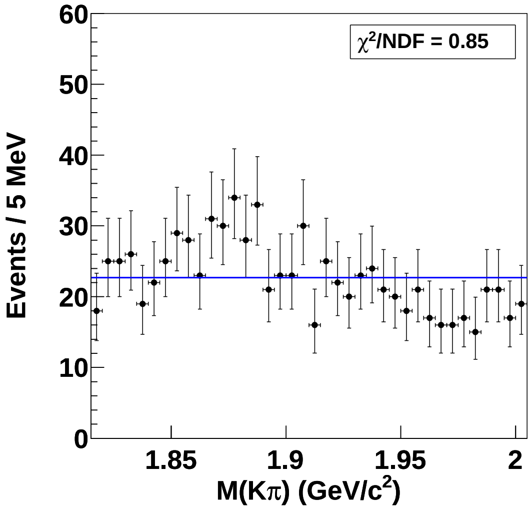


Figure 7.2: The $M(K\pi)$ distribution fitted by a constant function for a MC sample of the background $B^- \to [K^-\pi^+]_D K^-$ in the analysis of $B^- \to [K^+\pi^-]_D K^-$. It is quite similar for the background $B^- \to [K^-\pi^+]_D \pi^-$ in the analysis of $B^- \to [K^+\pi^-]_D \pi^-$.

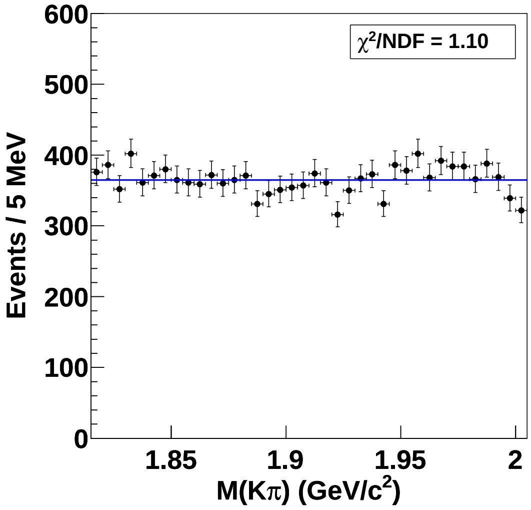


Figure 7.3: The $M(K\pi)$ distribution fitted by a constant function for a MC sample of the charmless background $B^- \to K^+K^-\pi^-$ in the analysis of $B^- \to [K^+\pi^-]_D K^-$.

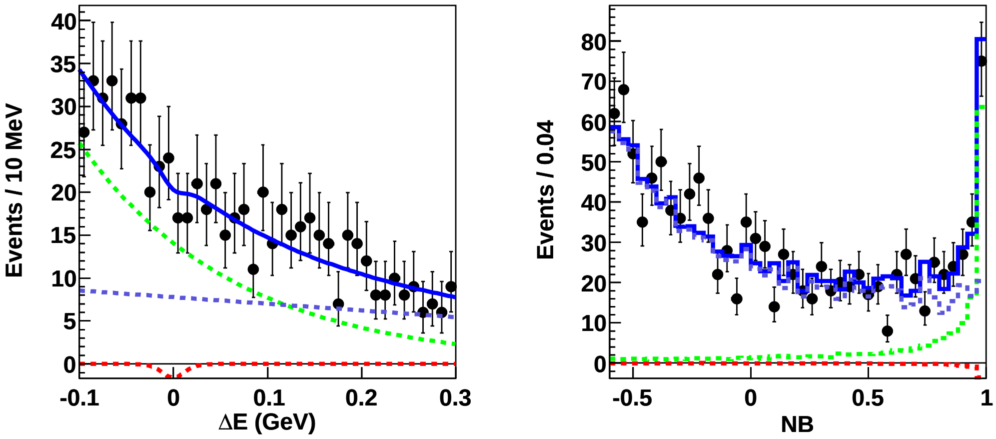


Figure 7.4: ΔE ($NB > 0.9$) and NB ($|\Delta E| < 0.03$ GeV) distributions for the $M(K\pi)$ sidebands of data in the analysis of $B^- \to [K^+\pi^-]_D K^-$. The fitted sample is shown with dots with error bars and the PDF is shown with the solid blue curve, for which the components are shown with dashed red (signal), dashed green ($B\bar{B}$ background), and dashed blue ($q\bar{q}$ background).

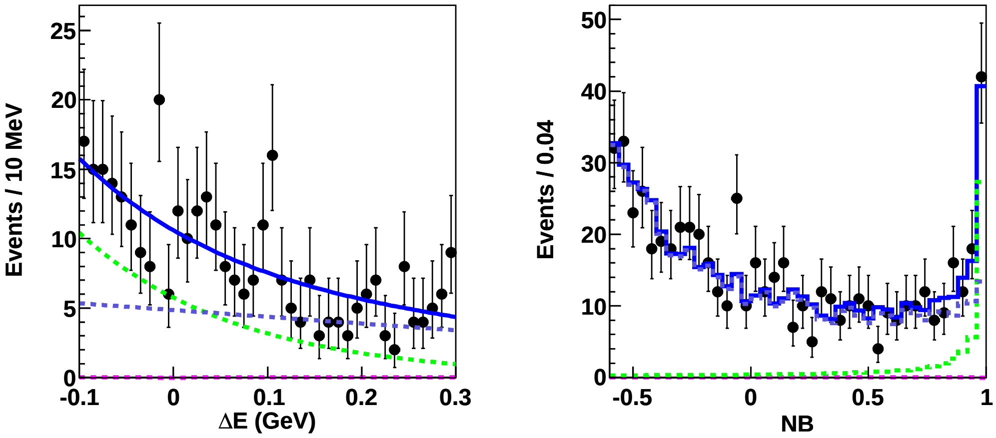


Figure 7.5: ΔE ($NB > 0.9$) and NB ($|\Delta E| < 0.03$ GeV) distributions for the $M(K\pi)$ sidebands of data in the analysis of $B^- \to [K^+\pi^-]_D \pi^-$. The fitted sample is shown with dots with error bars and the PDF is shown with the solid blue curve, for which the components are shown with dashed magenta (signal), dashed green ($B\bar{B}$ background), and dashed blue ($q\bar{q}$ background).

7.3 Check Using MC Samples

For a check of our study based on the $M(K\pi)$ sidebands of data, we estimate the background contributions using MC samples. We use the same MC samples as the ones in Section 7.1, and apply a fit for the events in the signal region of $M(K\pi)$. The results are consistent with the estimation in Section 7.2.

- Figure 7.6 shows the result of the fit for the background $B^- \to [K^+K^-]_D\pi^-$ in the analysis of $B^- \to [K^+\pi^-]_D K^-$. Estimated number of events which contribute to the signal extraction is

  $$\begin{aligned} N_{B\bar{B}} \times \mathcal{BR} \times \text{Eff.} &= 772 \text{ million} \times (1.92 \times 10^{-5}) \times (35.4 \pm 9.4)/1 \text{ million} \\ &= 0.5 \pm 0.1, \end{aligned} \tag{7.1}$$

  where $N_{B\bar{B}}$ is the number of $B\bar{B}$ pairs in data, $\mathcal{BR}$ is the branching ratio for the background, and Eff. is the efficiency in the analysis of $B^- \to [K^+\pi^-]_D K^-$ obtained by the fit. The errors on $N_{B\bar{B}}$ and $\mathcal{BR}$ are negligible compared to the error on Eff.

- Figures 7.7 and 7.8 show the results of the fits for the background $B^- \to [K^-\pi^+]_D h^-$ in the analysis of $B^- \to [K^+\pi^-]_D h^-$. Estimated number of events which contribute to the signal extraction is

  $$\begin{aligned} N_{B\bar{B}} \times \mathcal{BR} \times \text{Eff.} &= 772 \text{ million} \times (1.44 \times 10^{-5}) \times (28.9 \pm 7.6)/1 \text{ million} \times 1.6 \\ &= 0.5 \pm 0.1 \text{ for } h = K \text{ and} \qquad (7.2) \\ N_{B\bar{B}} \times \mathcal{BR} \times \text{Eff.} &= 772 \text{ million} \times (1.89 \times 10^{-4}) \times (14.5 \pm 6.6)/1 \text{ million} \times 1.6 \\ &= 3.4 \pm 1.5 \text{ for } h = \pi. \qquad (7.3) \end{aligned}$$

  Since the K/π misidentification probability is different between MC and data samples, we multiply the calibration factor 1.6 ± 0.1 obtained from the decay $D^{*-} \to \bar{D}^0\pi^-$ followed by $\bar{D}^0 \to K^+\pi^-$, where the error 0.1 is safely neglected in the calculation.

- For the charmless decay $B^- \to K^+K^-\pi^-$, we count the events which satisfy the selection criteria and obtain 1790 over one million generated events. The estimated contribution is calculated to be 6.9 ± 1.0 using the branching ratio obtained by BaBar [33]: $(5.0 \pm 0.7) \times 10^{-6}$. Note that the value is conservatively taken to be larger by the counting, since the ΔE distribution is slightly wider and the number obtained by the fit should become smaller.[2] Also, it is conservative to take no interferences of the resonances in the MC generation. In any case, the result is close to the value obtained in Section 7.2.

---

[2] If we apply a fit, the number 5.7 ± 0.8 is obtained, while the quality of the fit is slightly low.

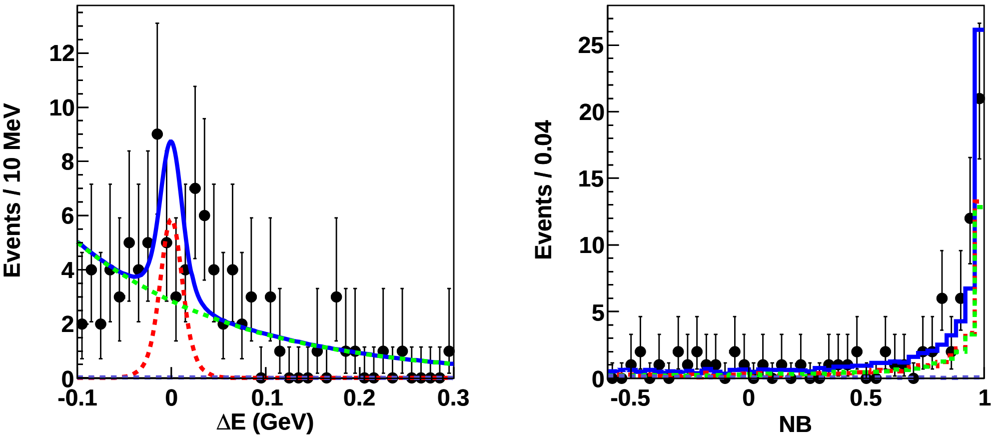


Figure 7.6: ΔE ($NB > 0.9$) and NB ($|\Delta E| < 0.03$ GeV) distributions for a MC sample of $B^- \to [K^+K^-]_D\pi^-$ in the analysis of $B^- \to [K^+\pi^-]_D K^-$. The fitted sample is shown with dots with error bars and the PDF is shown with the solid blue curve, for which the components are shown with dashed red (signal), dashed green ($B\bar{B}$ background), and dashed blue ($q\bar{q}$ background).

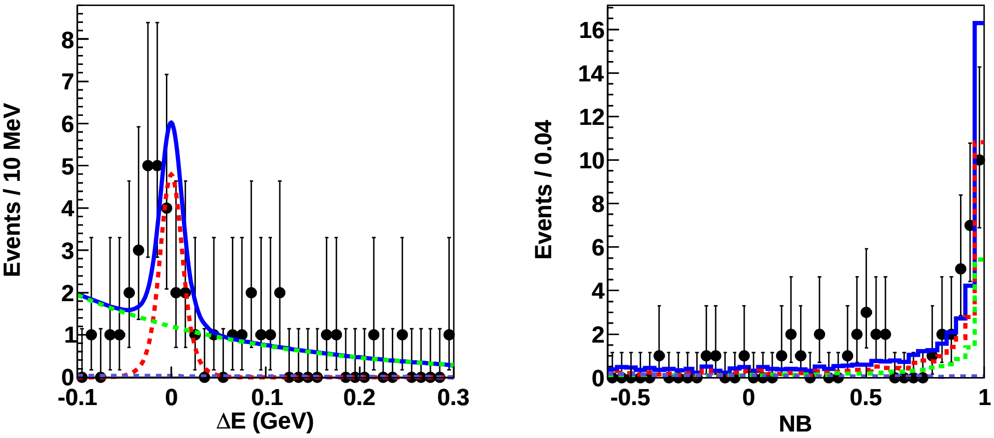


Figure 7.7: ΔE ($NB > 0.9$) and NB ($|\Delta E| < 0.03$ GeV) distributions for a MC sample of $B^- \to [K^-\pi^+]_D K^-$ in the analysis of $B^- \to [K^+\pi^-]_D K^-$. The fitted sample is shown with dots with error bars and the PDF is shown with the solid blue curve, for which the components are shown with dashed red (signal), dashed green ($B\bar{B}$ background), and dashed blue ($q\bar{q}$ background).

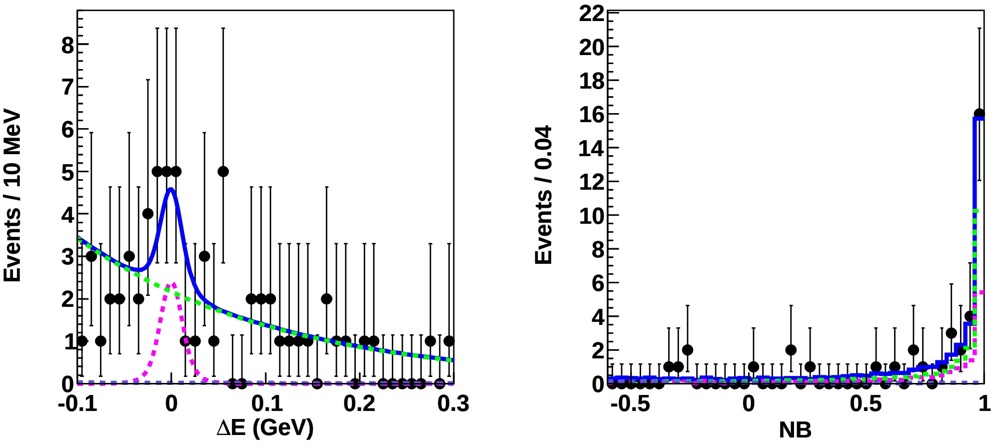


Figure 7.8: ΔE ($NB > 0.9$) and NB ($|\Delta E| < 0.03$ GeV) distributions for a MC sample of $B^- \to [K^-\pi^+]_D\pi^-$ in the analysis of $B^- \to [K^+\pi^-]_D\pi^-$. The fitted sample is shown with dots with error bars and the PDF is shown with the solid blue curve, for which the components are shown with dashed magenta (signal), dashed green ($B\bar{B}$ background), and dashed blue ($q\bar{q}$ background).

Chapter 8

Observables Relative to Angle ϕ_3

In this chapter, we evaluate the ratio $\mathcal{R}_{DK}$ and the asymmetry $\mathcal{A}_{DK}$ defined in Eqs. (2.27) and (2.28), respectively. The observables for $D\pi$, $\mathcal{R}_{D\pi}$ and $\mathcal{A}_{D\pi}$, are similarly obtained.

8.1 Detection Efficiency

The detection efficiencies are obtained from signal MC samples, where 100,000 events are generated for each mode. A calibration factor for the requirements of particle identifications, 0.96 ± 0.01 (0.94 ± 0.01) for DK ($D\pi$), is multiplied, where the decay $D^{*-} \to \bar{D}^0\pi^-$ followed by $\bar{D}^0 \to K^+\pi^-$ is used for the estimation. The resulting values are listed in Table 8.1.

8.2 Ratio $\mathcal{R}_{Dh}$

From the detection efficiency (ϵ) and the signal yield (N), we evaluate the ratio $\mathcal{R}_{Dh}$ as

$$\mathcal{R}_{Dh} = \frac{N_{[K^+\pi^-]_D h^-}/\epsilon_{[K^+\pi^-]_D h^-}}{N_{[K^-\pi^+]_D h^-}/\epsilon_{[K^-\pi^+]_D h^-}}, \tag{8.1}$$

for which the charge-conjugate modes are implicitly included. We obtain

$$\mathcal{R}_{DK} = [1.63^{+0.44}_{-0.41}(\text{stat})^{+0.07}_{-0.13}(\text{syst})] \times 10^{-2}, \tag{8.2}$$

$$\mathcal{R}_{D\pi} = [3.28^{+0.38}_{-0.36}(\text{stat})^{+0.12}_{-0.18}(\text{syst})] \times 10^{-3}, \tag{8.3}$$

where the values are summarized in Table 8.1.

The systematic errors (Table 8.2) are subdivided as follows.

(i) ΔE PDFs: The uncertainties due to the PDFs of ΔE for the DK signal, the $D\pi$ signal, and the $D\pi$ feed-across are obtained by varying the shape parameters by $\pm 1\sigma$. Those due to the DK feed-across are obtained by varying the width and the mean by $\pm 10\%$, which is the difference observed in the data and MC samples for the $D\pi$ feed-across. The total uncertainty is the quadratic sum of them: $^{+2.1}_{-1.8}\%$ ($^{+1.3}_{-1.2}\%$) for $\mathcal{R}_{DK}$ ($\mathcal{R}_{D\pi}$).

(ii) *NB* PDFs: The uncertainties from the PDFs of NB for the DK and $D\pi$ signals (the $D\pi$ feed-across) are estimated by obtaining PDFs from 0.01 GeV $< |\Delta E| <$ 0.02 GeV (0.01 GeV $< |\Delta E - 0.05$ GeV$| < 0.02$ GeV) instead of $|\Delta E| < 0.01$ GeV ($|\Delta E - 0.05$ GeV$| < 0.01$ GeV). Those due to the DK feed-across and the $B\bar{B}$ background are estimated by applying the PDF of the DK signal instead of MC-based PDF. Those due to the $q\bar{q}$ background are estimated by using different sideband regions in the space of $M_{\rm bc}$ and ΔE: 5.20 GeV$/c^2 < M_{\rm bc} < 5.26$ GeV$/c^2$ with -0.1 GeV $< \Delta E < 0.15$ GeV and $M_{\rm bc} > 5.26$ GeV$/c^2$ with 0.15 GeV $<$ $\Delta E < 0.3$ GeV. The total uncertainty is the quadratic sum of them: $^{+3.4}_{-3.0}\%$ ($\pm 3.1\%$) for $\mathcal{R}_{DK}$ ($\mathcal{R}_{D\pi}$).

(iii) Yields and asymmetries in the fit: The uncertainty due to the rate of the Dh yields in the DK and $D\pi$ samples is obtained by varying the value by $\pm 1\sigma$.[1] The uncertainty due to the asymmetry of the $q\bar{q}$ background is obtained by varying the value by ± 0.02 (± 0.00) for DK ($D\pi$), where the error of the asymmetry for the favored DK ($D\pi$) signal is used.[2] The total uncertainty is the quadratic sum of them: $^{+0.8}_{-0.9}\%$ ($\pm 0.1\%$) for $\mathcal{R}_{DK}$ ($\mathcal{R}_{D\pi}$).

(iv) Fit bias: Potential fit bias is checked by generating 10,000 pseudo experiments. We obtain almost standard-Gaussian pull distributions, and take the product of the mean of pull and the error of yield, providing -1.1% (-0.5%) for $\mathcal{R}_{DK}$ ($\mathcal{R}_{D\pi}$).

(v) Peaking backgrounds: The uncertainties of the yields of the backgrounds which peak under the signal were described in Section 7.2. The corresponding uncertainty in $\mathcal{R}_{DK}$ ($\mathcal{R}_{D\pi}$) is -6.6% (-4.2%).[3]

(vi) Efficiency: The uncertainties in detection efficiencies mainly arise from MC statistics and the uncertainties in the efficiencies of particle identifications. The estimated value is $\pm 1.7\%$ ($\pm 1.5\%$) for $\mathcal{R}_{DK}$ ($\mathcal{R}_{D\pi}$).

The total systematic error is the sum in quadrature of the above uncertainties.

The significances of $\mathcal{R}_{DK}$ and $\mathcal{R}_{D\pi}$ are estimated using $\sqrt{-2\ln(\mathcal{L}_0/\mathcal{L}_{\rm max})}$, where $\mathcal{L}_{\rm max}$ is the maximum likelihood and $\mathcal{L}_0$ is the likelihood when the signal yield is constrained to be zero. The distribution of the likelihood $\mathcal{L}$ is obtained by convolving the likelihood in the fit and an asymmetric Gaussian for taking the systematic error into account. Figure 8.1 shows the distributions of $\sqrt{-2\ln(\mathcal{L}/\mathcal{L}_{\rm max})}$. The significance for $\mathcal{R}_{DK}$ ($\mathcal{R}_{D\pi}$) is estimated to be 4.1σ (9.2σ), indicating the first evidence for the suppressed DK signal.

---

[1] The σ is the error in the estimation by the decay $D^{*-} \to \bar{D}^0\pi^-$ followed by $\bar{D}^0 \to K^+\pi^-$ for $h = K$ and the error in the fit to the favored modes for $h = \pi$.

[2] Momentum distributions of the particles used for the B reconstruction are similar in the Dh signal and in the $q\bar{q}$ background.

[3] The interference between the peaking background and the signal is negligible, since the "physical" width of the D meson is $\mathcal{O}$(meV).

Table 8.1: Summary of the results. The two errors for the ratio $\mathcal{R}_{Dh}$ are statistical and systematic, respectively.

Mode	Efficiency (%)	Yield	$\mathcal{R}_{Dh}$ [Significance]
$B^- \to [K^+\pi^-]_D K^-$	33.6 ± 0.4	$56.0^{+15.1}_{-14.2}$	$(1.63^{+0.44+0.07}_{-0.41-0.13}) \times 10^{-2}$ [4.1σ]
$B^- \to [K^-\pi^+]_D K^-$	33.2 ± 0.4	3394^{+68}_{-69}	
$B^- \to [K^+\pi^-]_D \pi^-$	36.5 ± 0.4	$165.0^{+19.1}_{-18.1}$	$(3.28^{+0.38+0.12}_{-0.36-0.18}) \times 10^{-3}$ [9.2σ]
$B^- \to [K^-\pi^+]_D \pi^-$	35.7 ± 0.4	49164^{+245}_{-244}	

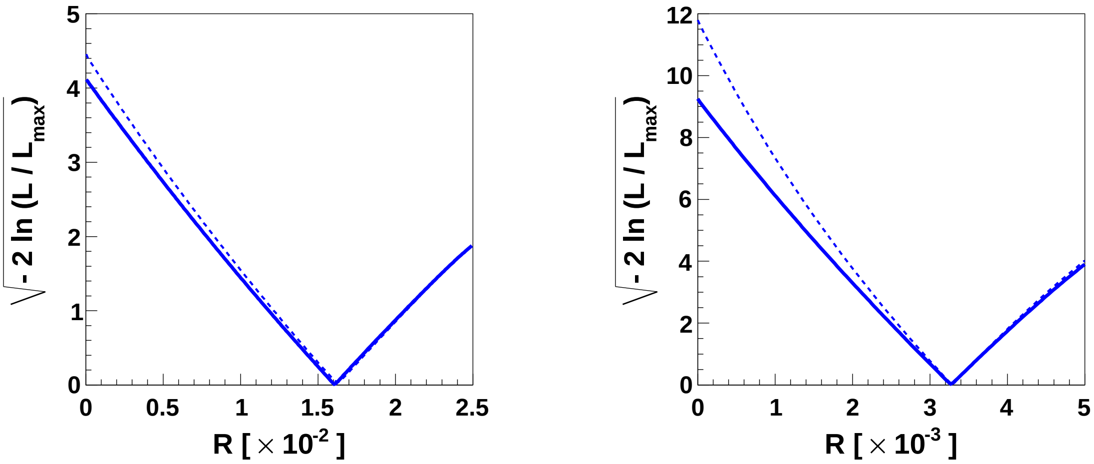


Figure 8.1: The distributions of $\sqrt{-2\ln(\mathcal{L}/\mathcal{L}_{\max})}$ for DK (left) and $D\pi$ (right). The dashed curve shows the distribution for the likelihood in the fit, while the solid curve shows the distribution obtained by including the systematic error.

8.3 Asymmetry $\mathcal{A}_{Dh}$

We obtain the asymmetry $\mathcal{A}_{Dh}$ in the suppressed decays as

$$\mathcal{A}_{DK} = -0.39^{+0.26}_{-0.28}(\text{stat})^{+0.04}_{-0.03}(\text{syst}), \tag{8.4}$$

$$\mathcal{A}_{D\pi} = -0.04 \pm 0.11(\text{stat})^{+0.02}_{-0.01}(\text{syst}). \tag{8.5}$$

The systematic errors (Table 8.2) are subdivided as follows. The uncertainties related to the fit are obtained by the same ways as the ones for $\mathcal{R}_{Dh}$. The uncertainty due to the yield of the peaking background is obtained to be +0.03 (+0.00) for $\mathcal{A}_{DK}$ ($\mathcal{A}_{D\pi}$) by varying the signal yield (denominator of the asymmetry). The uncertainty due to the asymmetry of the peaking background, where the value is 0.00 ± 0.10 for the charmless decay [33], is negligible. Possible bias due to the charge asymmetry of the detector is estimated to be the error of the asymmetry for the favored signal, ±0.02 (±0.00) for $\mathcal{A}_{DK}$ ($\mathcal{A}_{D\pi}$). The total systematic error is the sum in quadrature of the above uncertainties.

Table 8.2: Summary of the systematic uncertainties for $\mathcal{R}_{Dh}$ and $\mathcal{A}_{Dh}$. We use "$\cdots$" for the sources which do not contribute.

Source	$\mathcal{R}_{DK}$	$\mathcal{R}_{D\pi}$	$\mathcal{A}_{DK}$	$\mathcal{A}_{D\pi}$
ΔE PDFs	${}^{+2.1}_{-1.8}\%$	${}^{+1.3}_{-1.2}\%$	± 0.01	± 0.00
NB PDFs	${}^{+3.4}_{-3.0}\%$	$\pm 3.1\%$	${}^{+0.02}_{-0.01}$	± 0.01
Yields and asymmetries in fit	${}^{+0.8}_{-0.9}\%$	$\pm 0.1\%$	± 0.01	± 0.00
Fit bias	-1.1%	-0.5%	-0.01	-0.00
Peaking backgrounds	-6.6%	-4.2%	$+0.03$	$+0.00$
Efficiency	$\pm 1.7\%$	$\pm 1.5\%$	$\cdots$	$\cdots$
Detector asymmetry	$\cdots$	$\cdots$	± 0.02	± 0.00
Combined	${}^{+4.4}_{-7.8}\%$	${}^{+3.7}_{-5.6}\%$	${}^{+0.04}_{-0.03}$	${}^{+0.02}_{-0.01}$

8.4 Comparison of Results

In this section, we perform a comparison of the results between different analyses. Table 8.3 shows the previous values by Belle [19] and BaBar [20] as well as the results in this analysis. All the results for $\mathcal{R}_{D\pi}$ and $\mathcal{A}_{D\pi}$ are consistent with the value $(3.31 \pm 0.08) \times 10^{-3}$ [15] for which the decay $B^- \to \bar{D}^0\pi^-$, being small relative to the decay $B^- \to D^0\pi^-$ in the standard model, is neglected. The value of $\mathcal{R}_{DK}$ in this analysis is $\sim 1\sigma$ higher than the previous value by Belle. The reason can be statistical, since more than 60% events are independent from the ones in the previous analysis.[4] The values of $\mathcal{A}_{DK}$ are consistent with each other. Note that the contents of the systematic uncertainties in the three analyses are quite similar, while the values for this analysis are reduced by statistical advantages and careful studies. Additional uncertainties are included in the results by BaBar since they fit $M_{\rm bc}$, the value of which is not changed even when a misidentification is applied.

---

[4] The increase of the data sample is only 18% (657×10^6 $B\bar{B}$ pairs to 772×10^6 $B\bar{B}$ pairs), while many events are added by removing tight requirement for the continuum suppression.

Table 8.3: Comparison of the results for $\mathcal{R}_{Dh}$ and $\mathcal{A}_{Dh}$.

	This Analysis	Previous Result by Belle	Latest Result by BaBar
$\mathcal{R}_{DK}$ [$\times 10^{-2}$]	$1.63\ ^{+0.44}_{-0.41}\ ^{+0.07}_{-0.13}$	$0.78\ ^{+0.62}_{-0.57}\ ^{+0.20}_{-0.28}$	$1.1 \pm 0.6 \pm 0.2$
$\mathcal{R}_{D\pi}$ [$\times 10^{-3}$]	$3.28\ ^{+0.38}_{-0.36}\ ^{+0.12}_{-0.18}$	$3.40\ ^{+0.55}_{-0.53}\ ^{+0.15}_{-0.22}$	$3.3 \pm 0.6 \pm 0.4$
$\mathcal{A}_{DK}$	$-0.39\ ^{+0.26}_{-0.28}\ ^{+0.04}_{-0.03}$	$-0.1\ ^{+0.8}_{-1.0} \pm 0.4$	$-0.86 \pm 0.47\ ^{+0.12}_{-0.16}$
$\mathcal{A}_{D\pi}$	$-0.04 \pm 0.11\ ^{+0.02}_{-0.01}$	$-0.02\ ^{+0.15}_{-0.16} \pm 0.04$	$0.03 \pm 0.17 \pm 0.04$

Chapter 9

Extraction of Angle ϕ_3

The measurement of ϕ_3 is obtained by including the observables $\mathcal{R}_{CP\pm}$ and $\mathcal{A}_{CP\pm}$ and the strong phase δ_D. The list of the inputs is the following:

$$
\begin{aligned}
\mathcal{R}_{DK} &= 0.0163 \pm 0.0045, && (9.1)\\
\mathcal{A}_{DK} &= -0.39 \pm 0.28, && (9.2)\\
\mathcal{R}_{CP+} &= 1.18 \pm 0.08, && (9.3)\\
\mathcal{A}_{CP+} &= 0.24 \pm 0.06, && (9.4)\\
\mathcal{R}_{CP-} &= 1.09 \pm 0.08, && (9.5)\\
\mathcal{A}_{CP-} &= -0.10 \pm 0.07, && (9.6)\\
\delta_D &= 338.0^\circ \pm 11.2^\circ, && (9.7)
\end{aligned}
$$

where the values for $\mathcal{R}_{CP\pm}$, $\mathcal{A}_{CP\pm}$, and δ_D are taken from Ref. [15]. For each error, we take the sum in quadrature of the statistical and systematic errors. We have seven equations for four unknown parameters ϕ_3, r_B, δ_B, and δ_D.

The method for extracting ϕ_3 is based on Bayesian statistics with Gaussian assumptions for the observables. We minimize the $\chi^2(\phi_3, r_B, \delta_B, \delta_D)$ given by

$$
\chi^2(\phi_3, r_B, \delta_B, \delta_D) = \sum_i \frac{(\mathcal{O}^i_{\text{measured}} - \mathcal{O}^i(\phi_3, r_B, \delta_B, \delta_D))^2}{{\sigma_{\mathcal{O}^i}}^2}, \tag{9.8}
$$

where $\mathcal{O}^i_{\text{measured}}$ and $\sigma_{\mathcal{O}^i}$ are the measured values and the errors in Eqs. (9.1)–(9.7), $\mathcal{O}^i(\phi_3, r_B, \delta_B, \delta_D)$ are the observables as the functions of ϕ_3, r_B, δ_B, and δ_D, and the index i runs over the seven observables.

The result (Figure 9.1) is

$$
\begin{aligned}
\phi_3 &= 89^\circ\ {}^{+13^\circ}_{-16^\circ}, && (9.9)\\
r_B &= 0.12 \pm 0.02, && (9.10)\\
\delta_B &= 34^\circ\ {}^{+12^\circ}_{-10^\circ}, && (9.11)\\
\delta_D &= 329^\circ \pm 10^\circ, && (9.12)
\end{aligned}
$$

modulo 180° for ϕ_3 and δ_B. The projections of $\chi^2(\phi_3, r_B, \delta_B, \delta_D)$ for ϕ_3 and δ_B are asymmetric, and two standard deviation regions are

$$39^\circ < \quad \phi_3 \quad < 115^\circ, \tag{9.13}$$

$$14^\circ < \quad \delta_B \quad < 78^\circ. \tag{9.14}$$

No significant deviations from other measurements (Section 2.5 or Refs. [16, 17, 18]), including an indirect measurement assuming the standard model, are found.[1]

---

[1] The deviations for δ_B are slightly higher but less than 3σ, being estimated by using the likelihood distributions of the results.

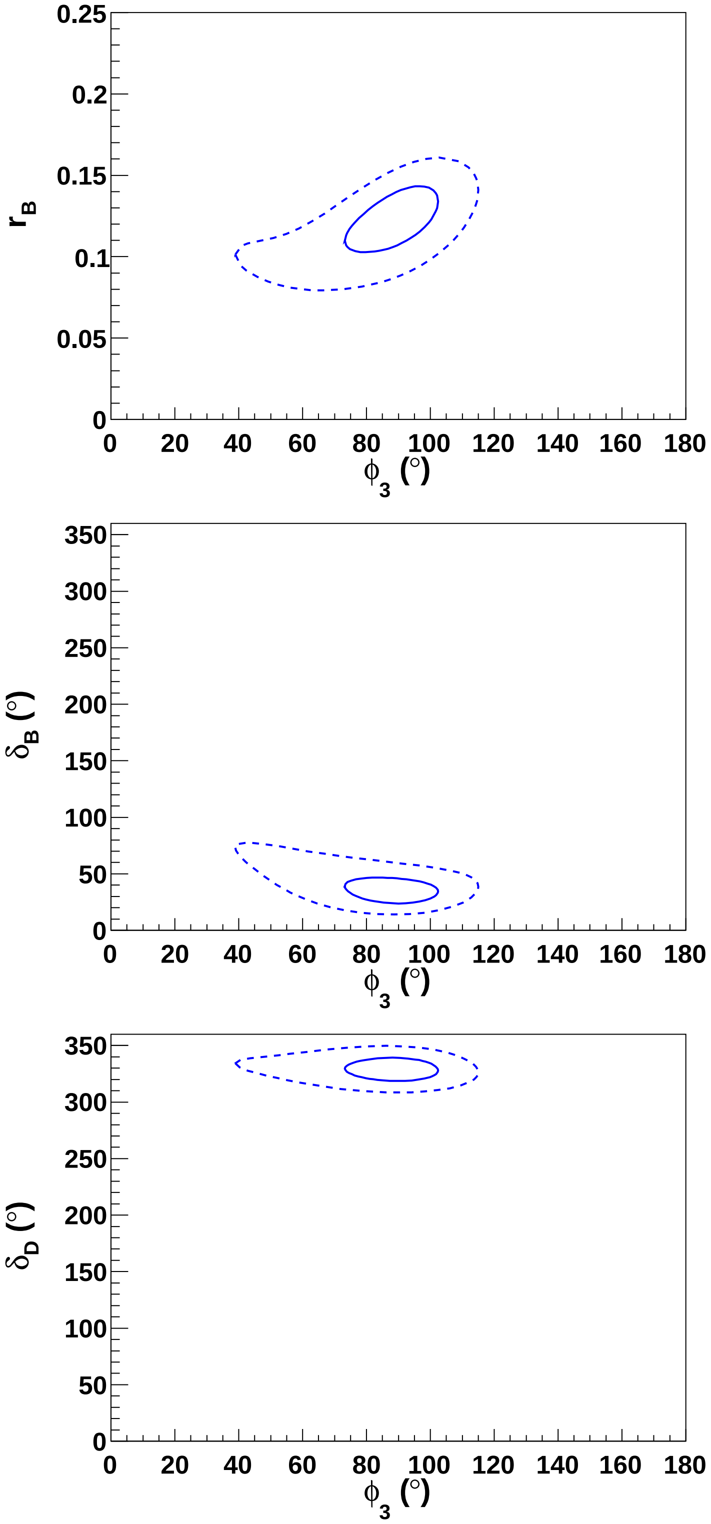


Figure 9.1: Result of the ϕ_3 fit. The contours indicate 1 and 2 standard deviation regions. Of the two solutions for ϕ_3 in 0°–360°, a solution in 0°–180° is selected.

Chapter 10

Conclusion

In summary, we report study of the suppressed decay $B^- \to [K^+\pi^-]_D K^-$ using 772×10^6 $B\bar{B}$ pairs collected with the Belle detector. A new method to discriminate the $q\bar{q}$ background combining ten variables with a neural network [31], a D^{*-} veto, a two-dimensional fit to extract the signal, as well as a 20% increase of the data sample enable a significant improvement on the measurement compared to the previous analysis [19]. We obtain the first evidence of the signal with a significance of 4.1σ, resulting in the most precise measurements of the ratio $\mathcal{R}_{DK}$ of the suppressed decay rate to the favored decay rate, obtained in the decay $B^- \to [K^-\pi^+]_D K^-$, and the asymmetry $\mathcal{A}_{DK}$ between the charge-conjugate decays of the suppressed mode as

$$\begin{aligned} \mathcal{R}_{DK} &= [1.63^{+0.44}_{-0.41}(\text{stat})^{+0.07}_{-0.13}(\text{syst})] \times 10^{-2}, \\ \mathcal{A}_{DK} &= -0.39^{+0.26}_{-0.28}(\text{stat})^{+0.04}_{-0.03}(\text{syst}). \end{aligned}$$

The decay $B^- \to D\pi^-$ is also analyzed as a reference, for which the decay rate is relatively large whereas the CP-violating effect is small. We obtain

$$\begin{aligned} \mathcal{R}_{D\pi} &= [3.28^{+0.38}_{-0.36}(\text{stat})^{+0.12}_{-0.18}(\text{syst})] \times 10^{-3}, \\ \mathcal{A}_{D\pi} &= -0.04 \pm 0.11(\text{stat})^{+0.02}_{-0.01}(\text{syst}), \end{aligned}$$

which are consistent with the expectations of the standard model.

The results for DK constitute important ingredients in a model-independent extraction of ϕ_3 based on a minimum χ^2 method including the observables for the decays $B^- \to D_{CP\pm}K^-$ [15]. We obtain

$$\phi_3 = 89^\circ \ {}^{+13^\circ}_{-16^\circ},$$

which is consistent with the expectation from the CKM picture in the standard model [18]. Our study strongly encourages the precision measurements of ϕ_3 at the next-generation B factories, which will play an important role for investigating new physics.

Appendix A

Details on Neural Network Training

This appendix presents the details about the neural network training. Table A.1 shows the summary table for input variables. The rank of importance ("rank"), the identification number ("node"), the name ("name"), and the preprocessing flag ("preprocess") are listed for all variables. The explanations on the preprocessing flags are the following.

- 12: transform to Gaussian with no delta-function
- 14: regularized spline fit with no delta-function
- 15: regularized monotonous spline fit with no delta-function
- 18: use map for unordered classes with no delta-function
- 34: regularized spline fit with delta-function

rank	node	name	preprocess
1	2	$\mathcal{LR}$(KSFW)	12
2	4	Δz	34
3	7	$\cos\theta_D^K$	14
4	3	$\lvert\cos\theta_B\rvert$	15
5	5	r_{tag}	18
6	6	$\lvert\cos\theta_{\text{T}}\rvert$	14
7	11	d_{Dh}	15
8	10	$Q_B Q_K$	18
9	8	$\cos\theta_B^D$	14
10	9	ΔQ	14

Table A.1: List of the input variables.

In the following pages, we show the figures related to the training. The explanations are included in the captions.

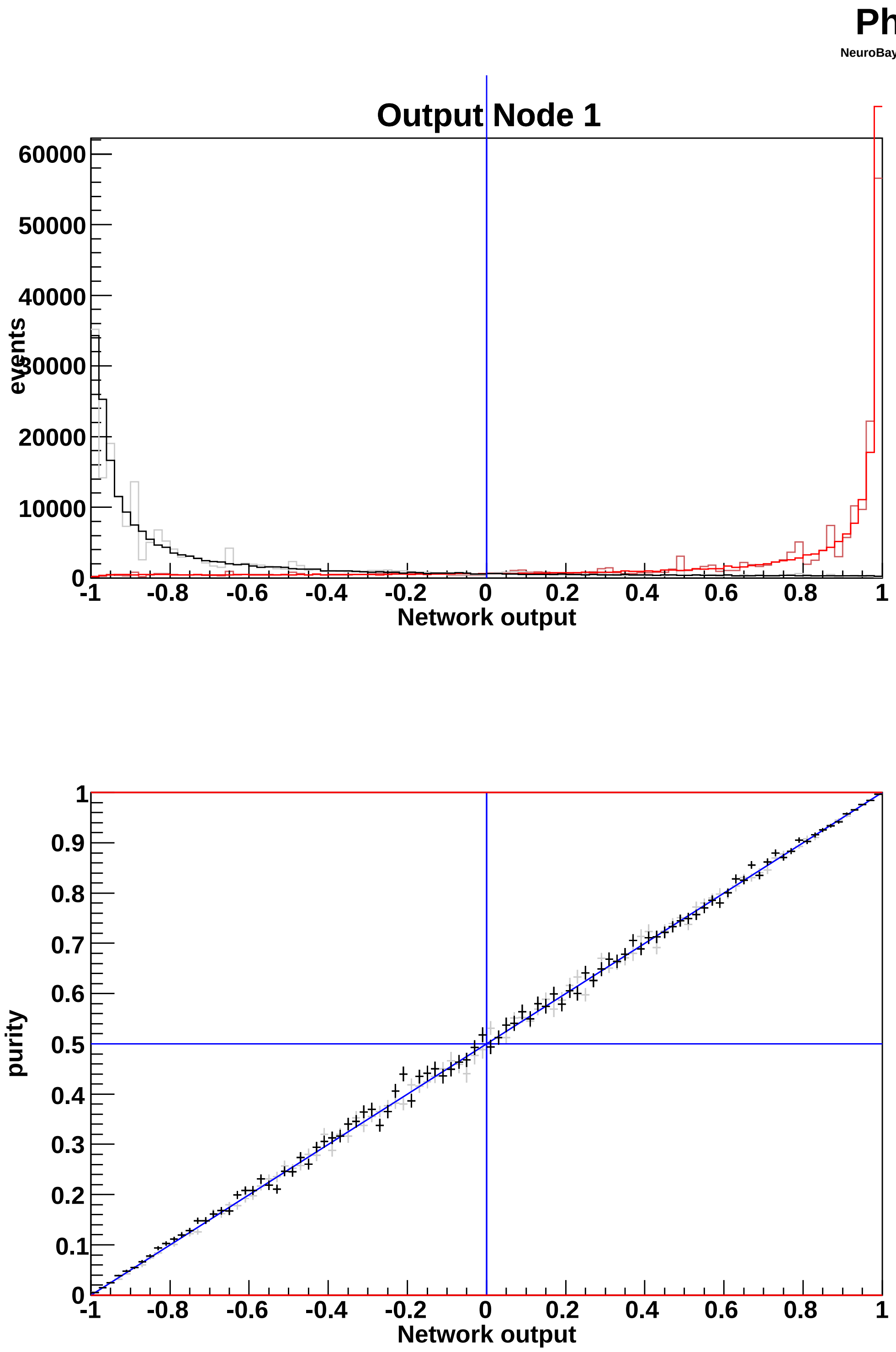


Figure A.1: The distributions of network output for MC samples of signal (red) and background (black) are shown (upper). The purity is plotted over the network output (lower). The figures given behind correspond to the output before the iteration training.

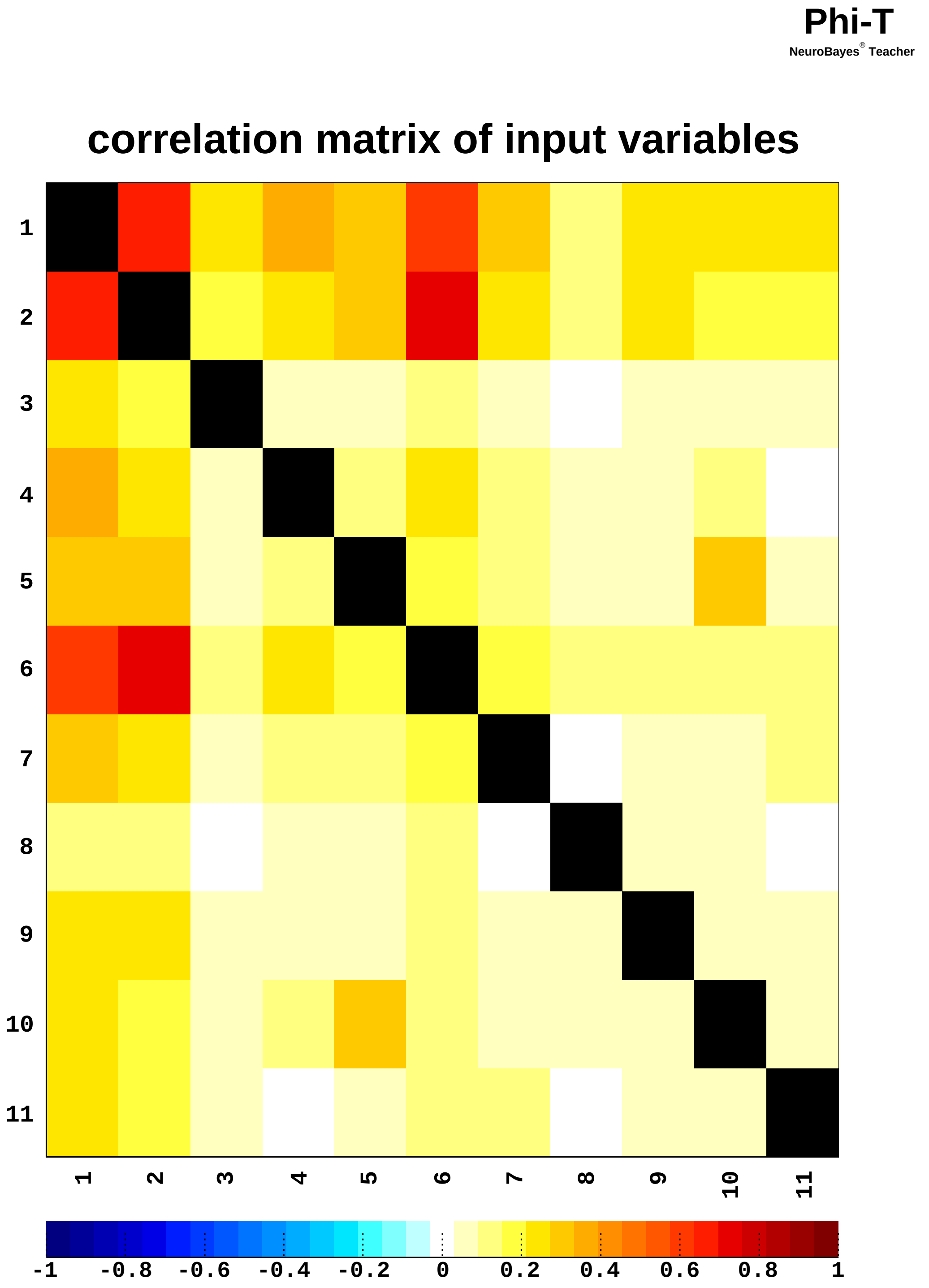


Figure A.2: Correlation matrix of input variables listed in the order of node numbers.

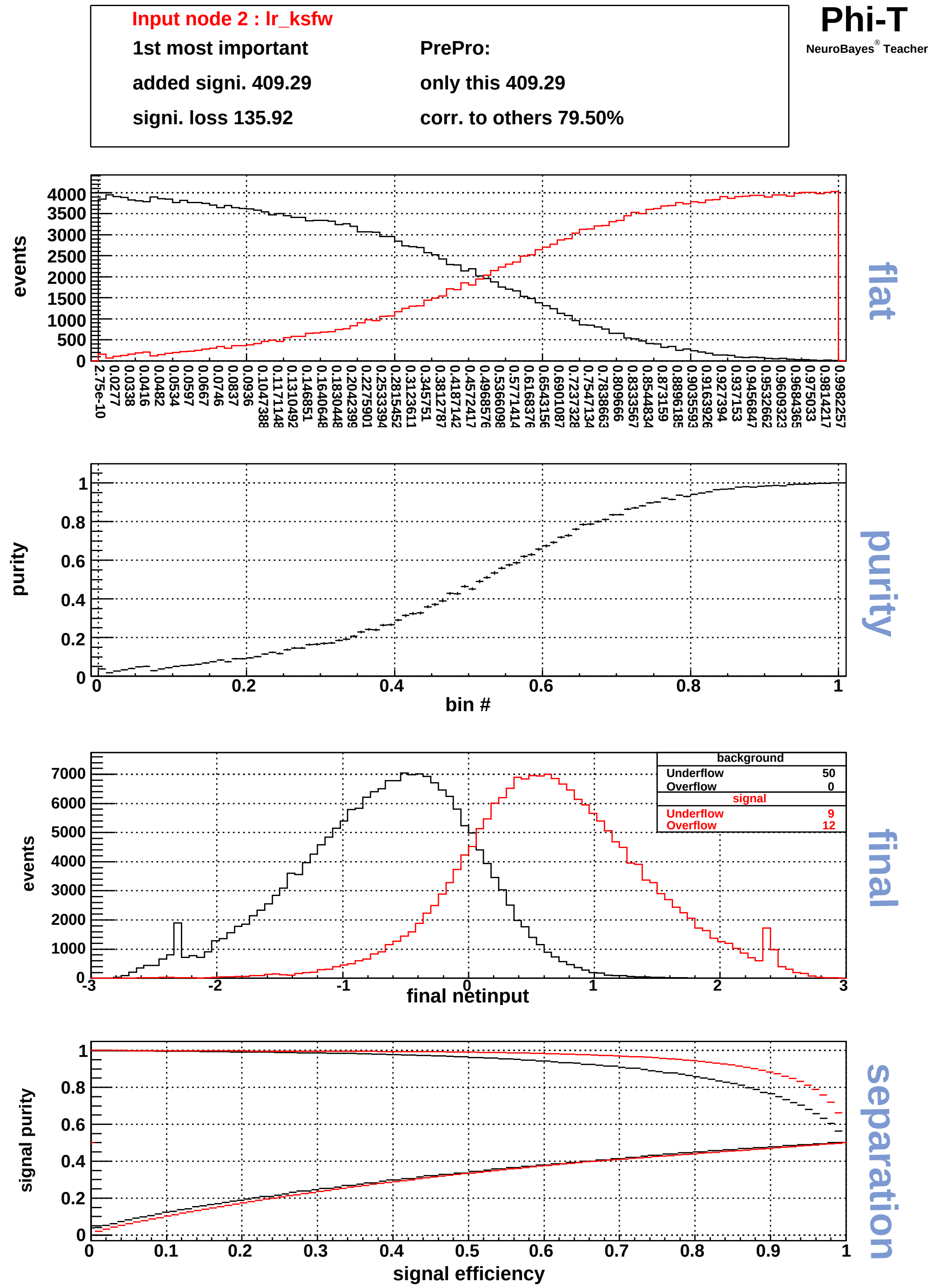


Figure A.3: The equalized distributions for the input variable (1st), the signal purity (2nd), the distributions after the preprocessing (3rd), and the purity curve on the efficiency (4th) for $\mathcal{LR}$(KSFW). In the 1st and 3rd figures, the red corresponds to the signal while the black corresponds to the background. In the 4th figure, the red corresponds to the result after the training using all variables and the black corresponds to the one for single variable.

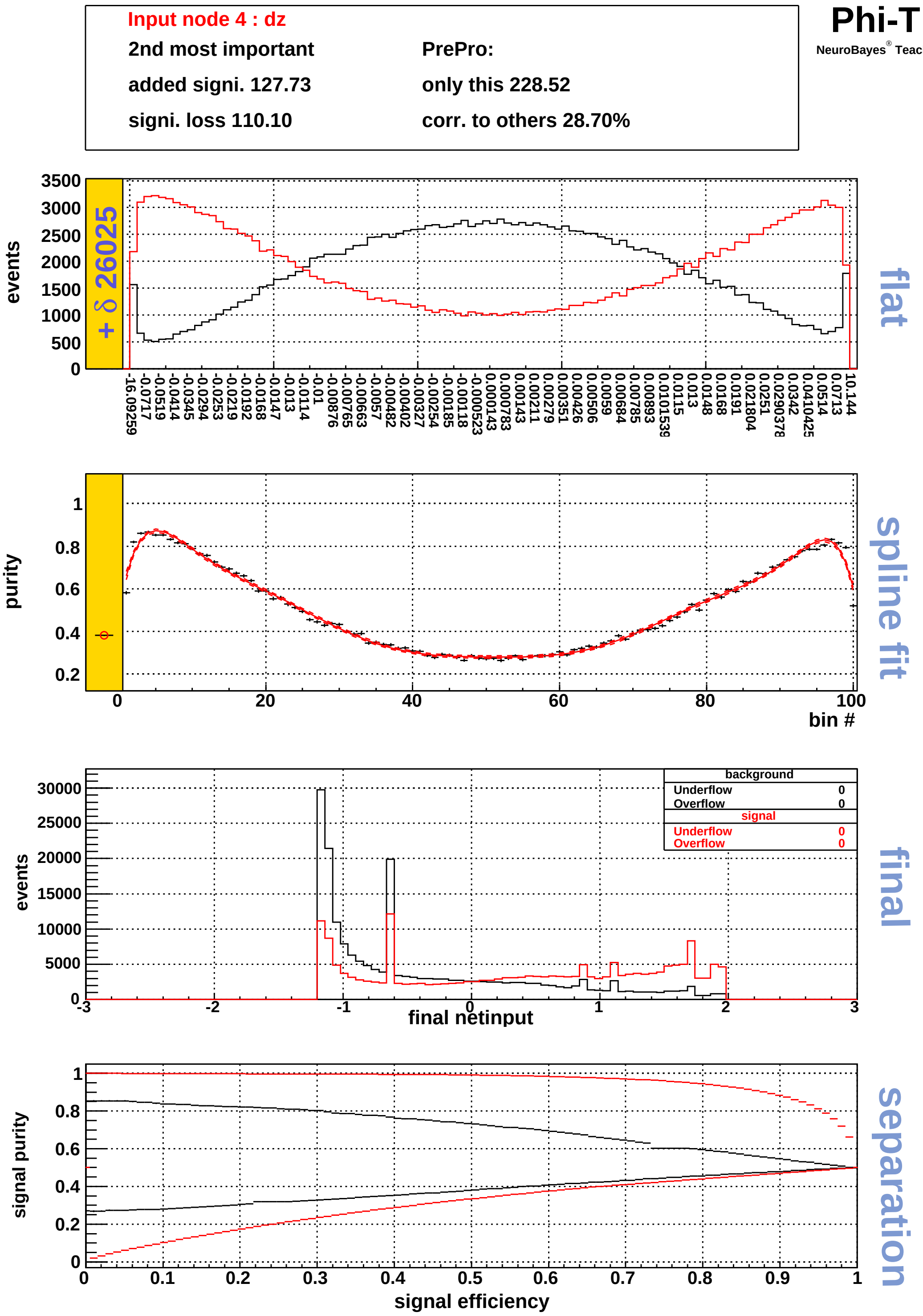


Figure A.4: The equalized distributions for the input variable (1st), the signal purity with a spline fit (2nd), the distributions after the preprocessing (3rd), and the purity curve on the efficiency (4th) for Δz. In the 1st and 3rd figures, the red corresponds to the signal while the black corresponds to the background. In the 4th figure, the red corresponds to the result after the training using all variables and the black corresponds to the one for single variable. We use a delta-function for the cases when the the vertex fitting is failed.

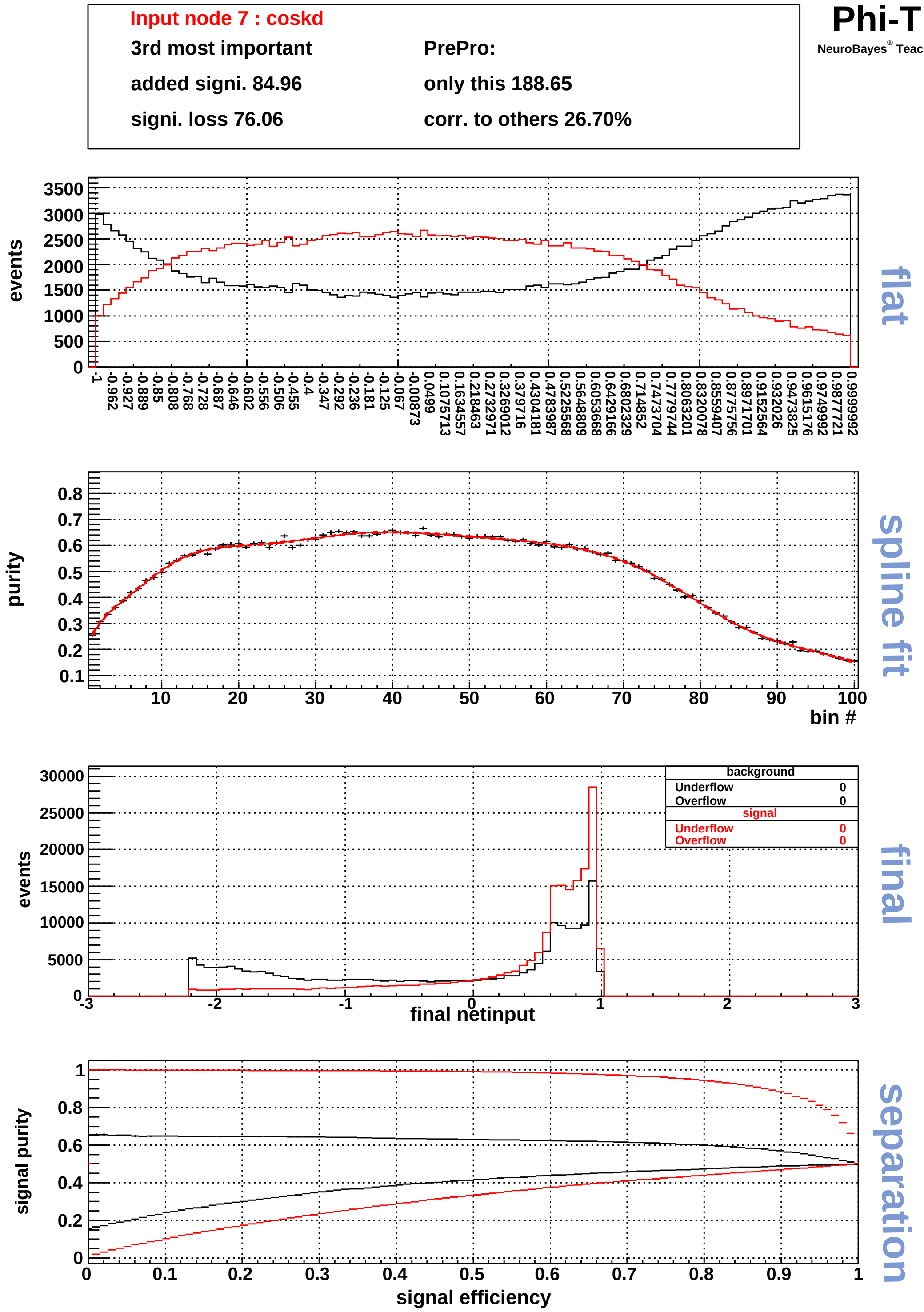


Figure A.5: The equalized distributions for the input variable (1st), the signal purity with a spline fit (2nd), the distributions after the preprocessing (3rd), and the purity curve on the efficiency (4th) for $\cos\theta_D^K$. In the 1st and 3rd figures, the red corresponds to the signal while the black corresponds to the background. In the 4th figure, the red corresponds to the result after the training using all variables and the black corresponds to the one for single variable.

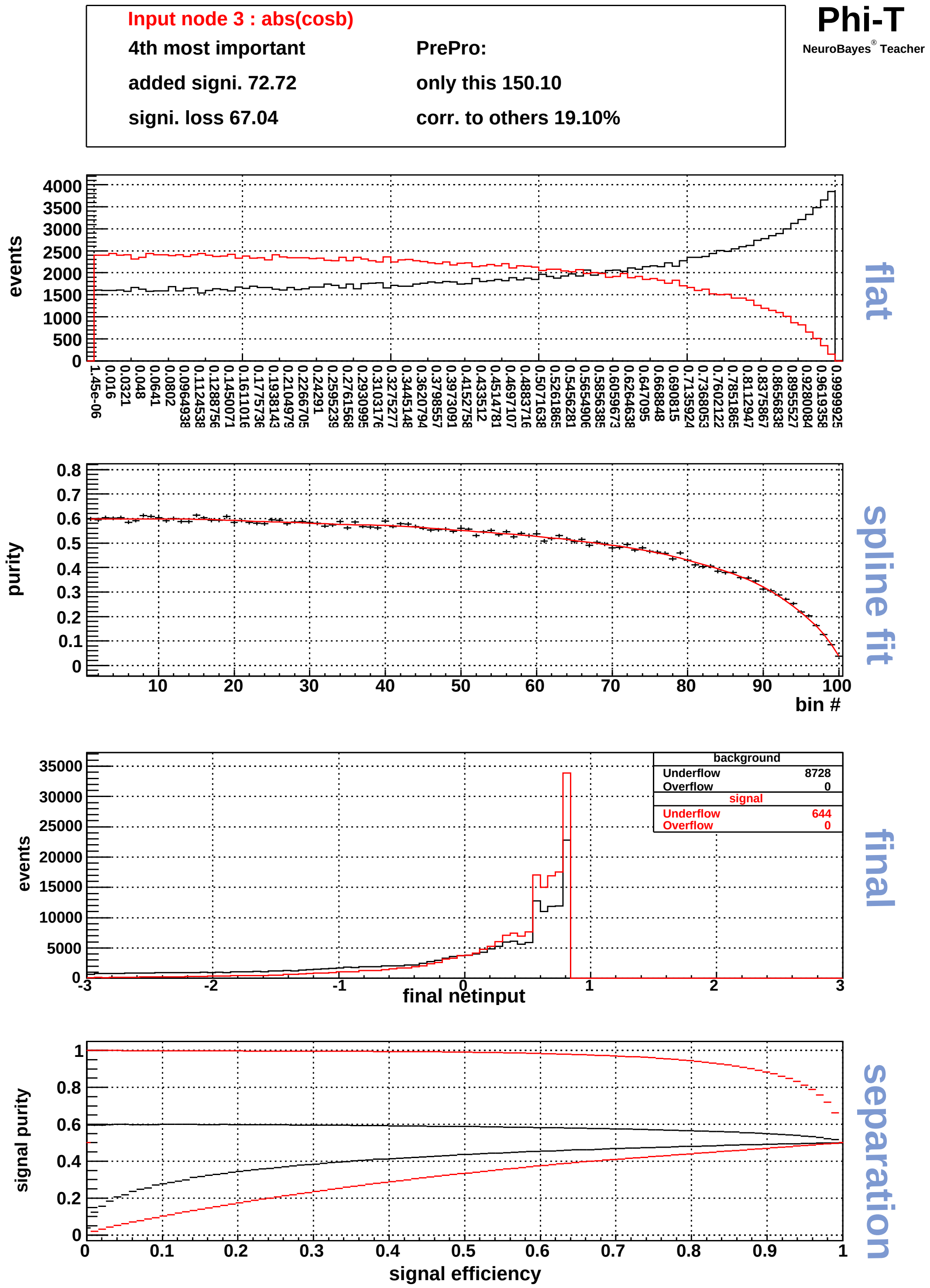


Figure A.6: The equalized distributions for the input variable (1st), the signal purity with a spline fit (2nd), the distributions after the preprocessing (3rd), and the purity curve on the efficiency (4th) for $|\cos\theta_B|$. In the 1st and 3rd figures, the red corresponds to the signal while the black corresponds to the background. In the 4th figure, the red corresponds to the result after the training using all variables and the black corresponds to the one for single variable.

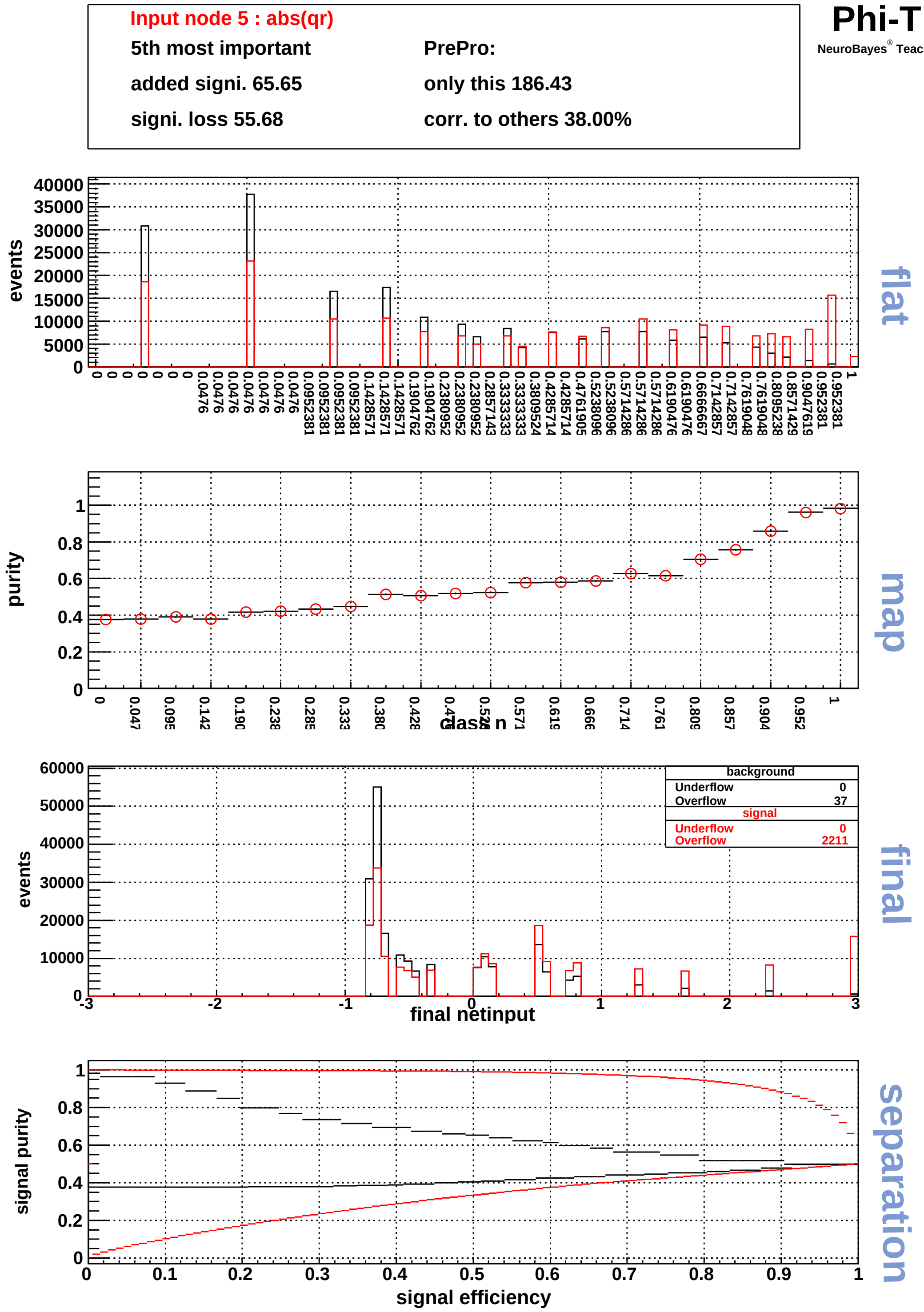


Figure A.7: The equalized distributions for the input variable (1st), the signal purity (2nd), the distributions after the preprocessing (3rd), and the purity curve on the efficiency (4th) for r_{tag}. In the 1st and 3rd figures, the red corresponds to the signal while the black corresponds to the background. In the 4th figure, the red corresponds to the result after the training using all variables and the black corresponds to the one for single variable.

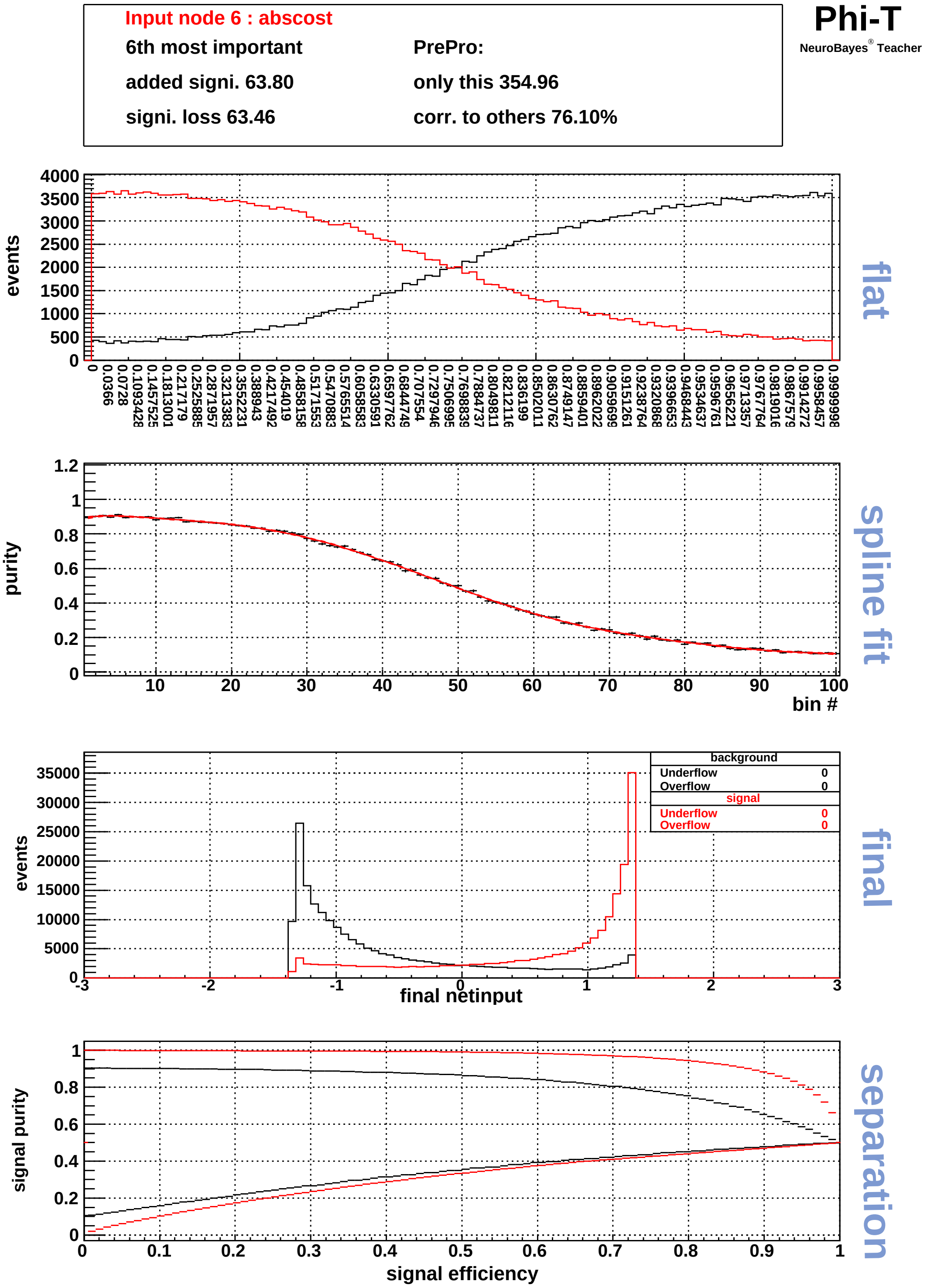


Figure A.8: The equalized distributions for the input variable (1st), the signal purity with a spline fit (2nd), the distributions after the preprocessing (3rd), and the purity curve on the efficiency (4th) for $|\cos\theta_{\rm T}|$. In the 1st and 3rd figures, the red corresponds to the signal while the black corresponds to the background. In the 4th figure, the red corresponds to the result after the training using all variables and the black corresponds to the one for single variable.

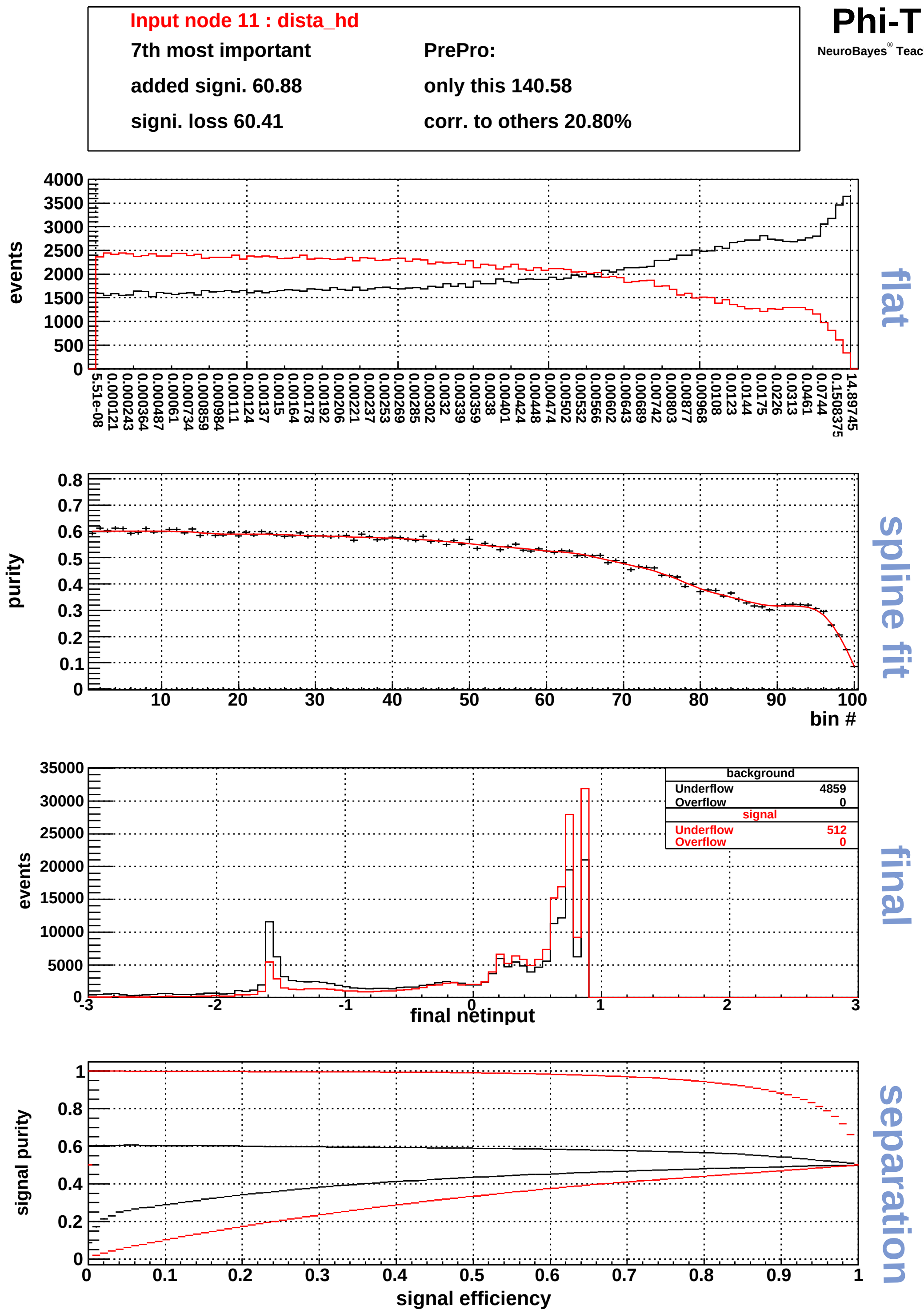


Figure A.9: The equalized distributions for the input variable (1st), the signal purity with a spline fit (2nd), the distributions after the preprocessing (3rd), and the purity curve on the efficiency (4th) for d_{Dh}. In the 1st and 3rd figures, the red corresponds to the signal while the black corresponds to the background. In the 4th figure, the red corresponds to the result after the training using all variables and the black corresponds to the one for single variable.

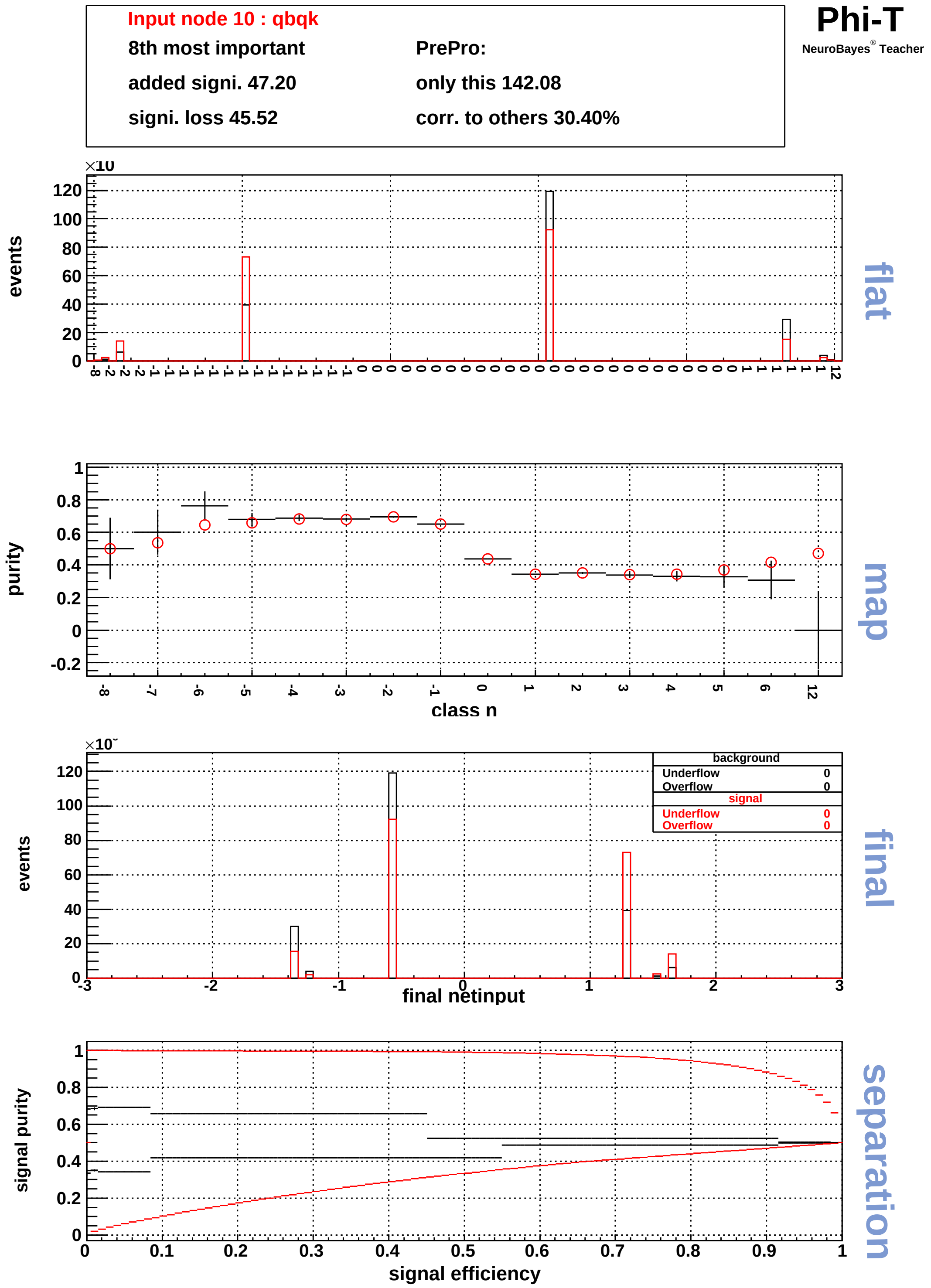


Figure A.10: The equalized distributions for the input variable (1st), the signal purity (2nd), the distributions after the preprocessing (3rd), and the purity curve on the efficiency (4th) for $Q_B Q_K$. In the 1st and 3rd figures, the red corresponds to the signal while the black corresponds to the background. In the 4th figure, the red corresponds to the result after the training using all variables and the black corresponds to the one for single variable.

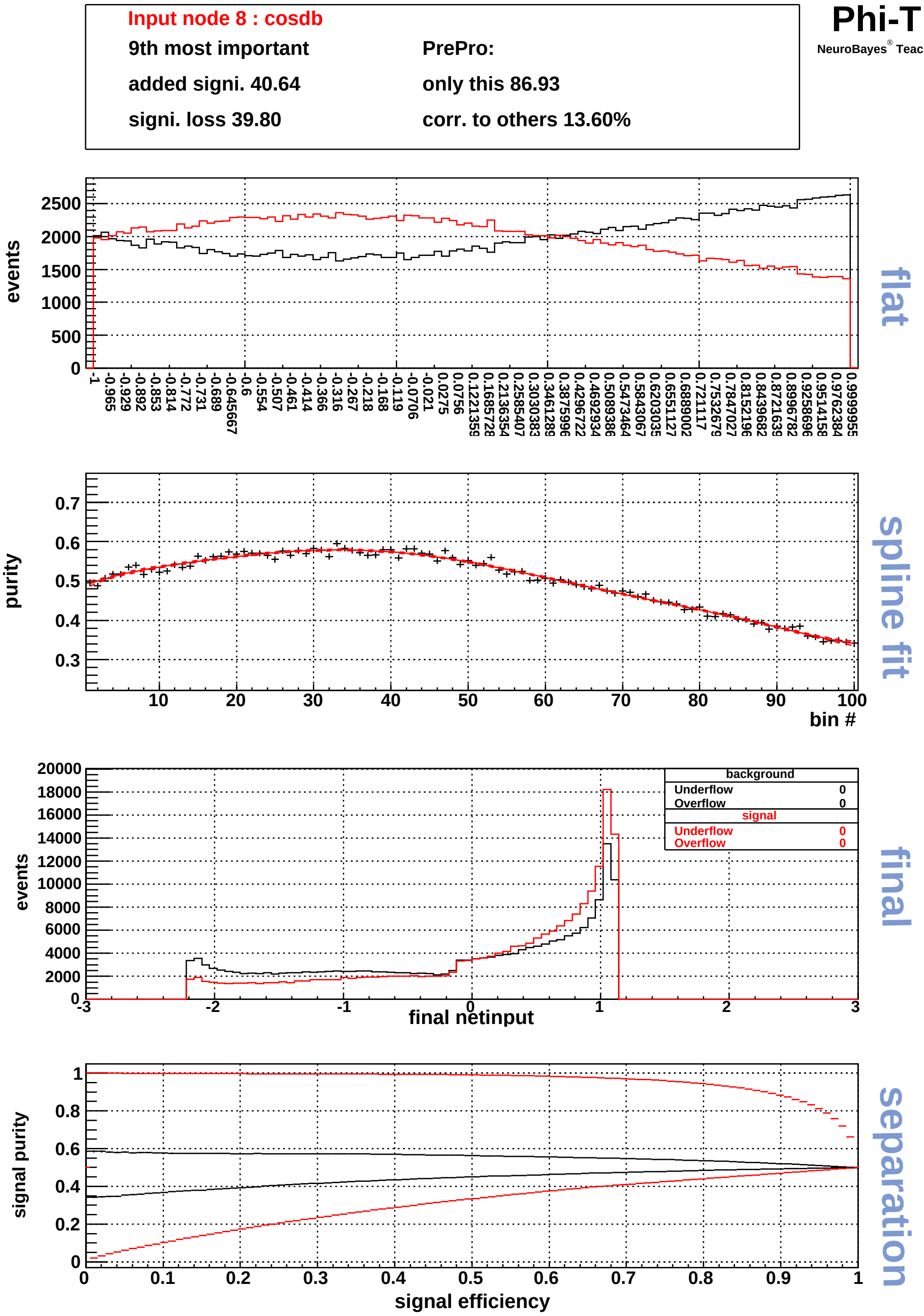


Figure A.11: The equalized distributions for the input variable (1st), the signal purity with a spline fit (2nd), the distributions after the preprocessing (3rd), and the purity curve on the efficiency (4th) for $\cos\theta_B^D$. In the 1st and 3rd figures, the red corresponds to the signal while the black corresponds to the background. In the 4th figure, the red corresponds to the result after the training using all variables and the black corresponds to the one for single variable.
]

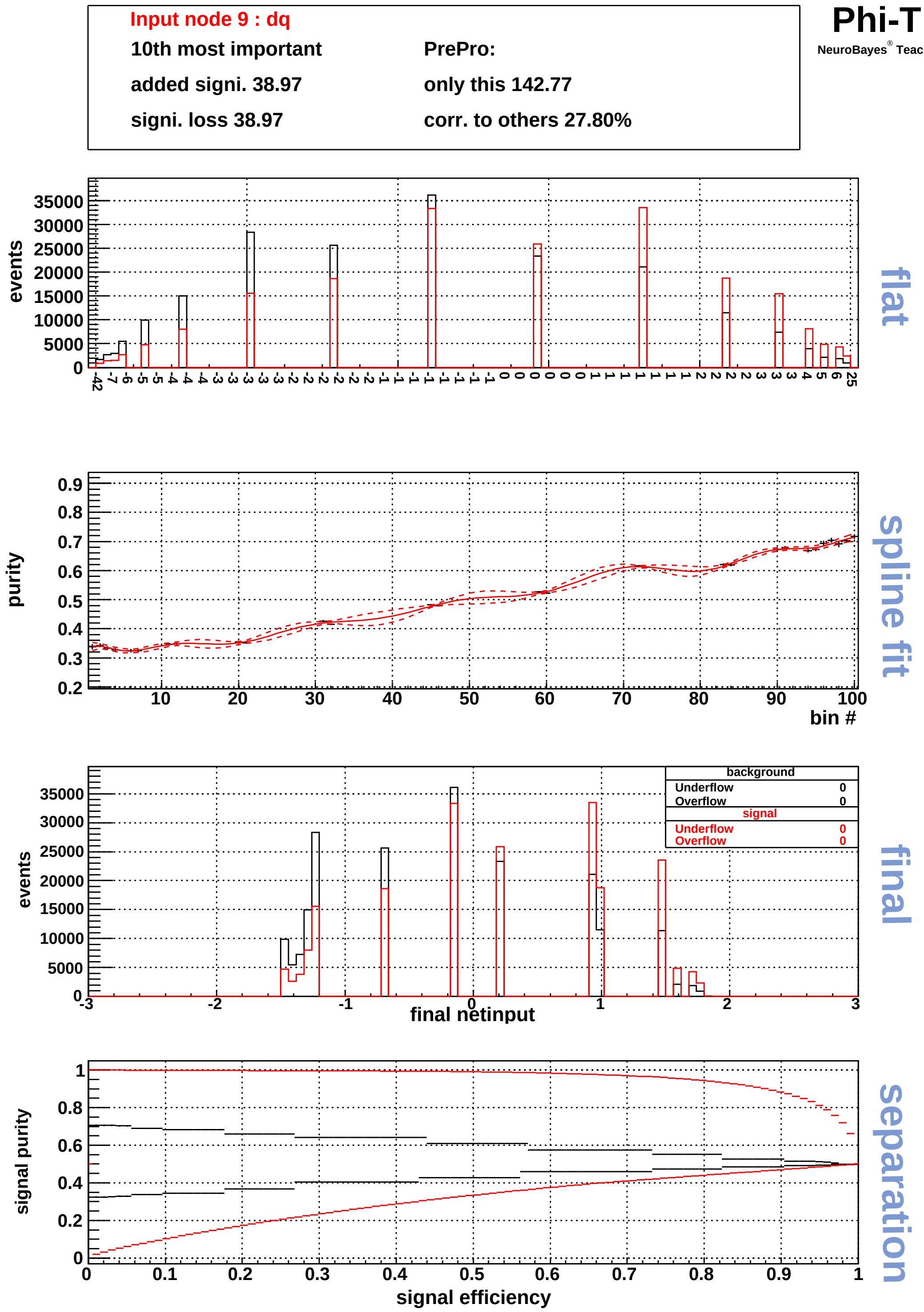


Figure A.12: The equalized distributions for the input variable (1st), the signal purity with a spline fit (2nd), the distributions after the preprocessing (3rd), and the purity curve on the efficiency (4th) for ΔQ. In the 1st and 3rd figures, the red corresponds to the signal while the black corresponds to the background. In the 4th figure, the red corresponds to the result after the training using all variables and the black corresponds to the one for single variable.

Appendix B

Demonstration of the Fit to MC Samples

Using the method of configuring PDF described in Section 6.1, we demonstrate the fit to MC samples, in which all known decays are allowed. Each sample has two times larger size than the data sample.

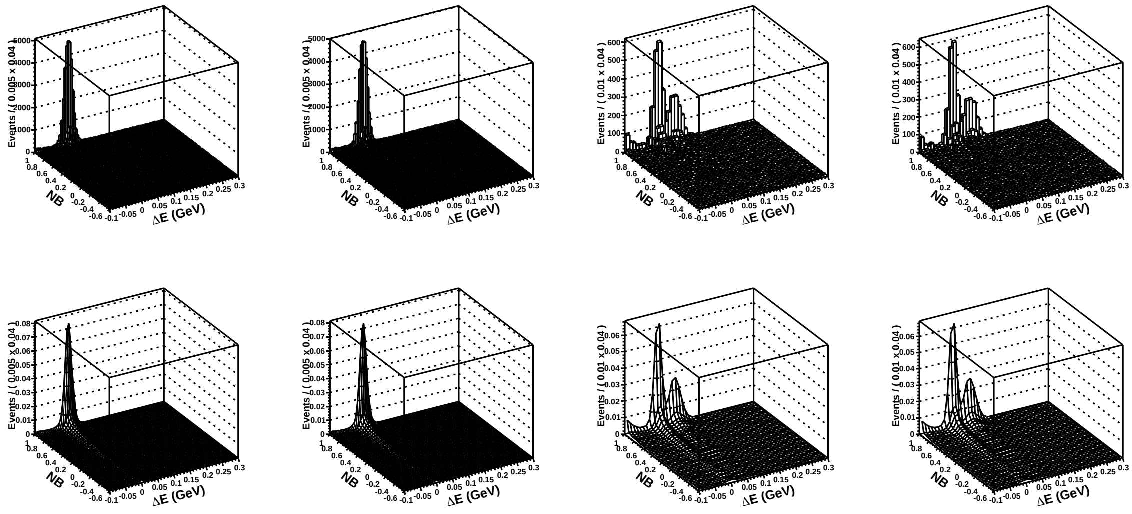
B.1 Fit to Favored Modes

Figure B.1 shows the result of the fit to the favored modes. Figures B.2 and B.3 show the projections for several regions, where good quality of the fit is obtained for each region. Table B.1 shows the list of the parameters in the fit. All resulting values are consistent with the inputs.

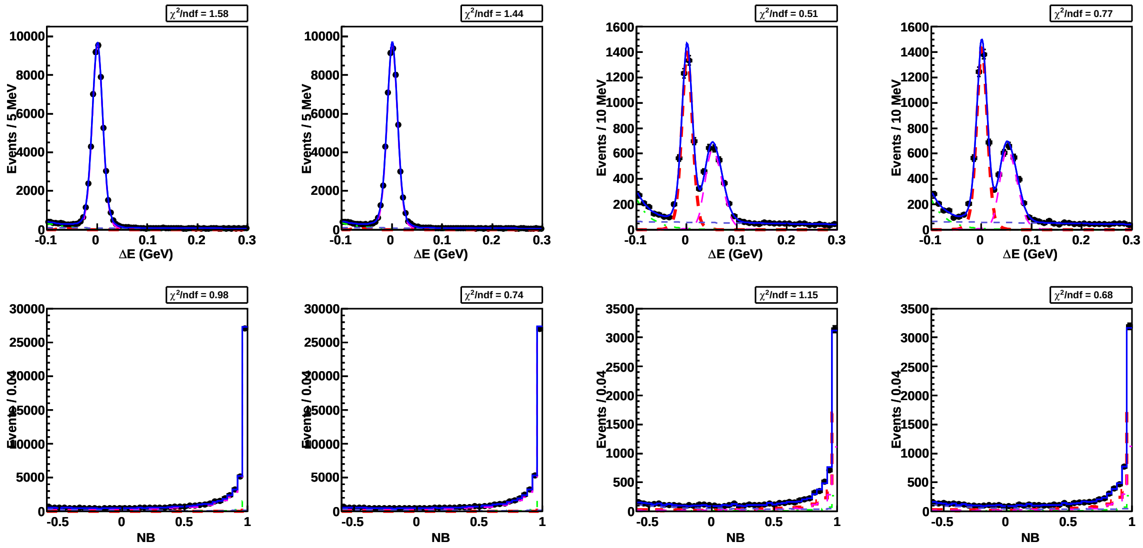
B.2 Fit to Suppressed Modes

Figure B.4 shows the result of the fit to the suppressed modes. Figures B.5 and B.6 show the projections for several regions, where good quality of the fit is obtained for each region. Table B.2 shows the list of the parameters in the fit.

The yields for the $D\pi$ and DK signals are obtained to be 428^{+26}_{-25} and 81^{+16}_{-15}, respectively, which are consistent with the included numbers 416 and 80, respectively. The asymmetry of the DK signal is obtained to be $(43 \pm 19)\%$, which is consistent with the input 38%. The asymmetries for the other components are consistent with the inputs 0. The fits are also applied for several other possible inputs. All the results are consistent with the inputs.

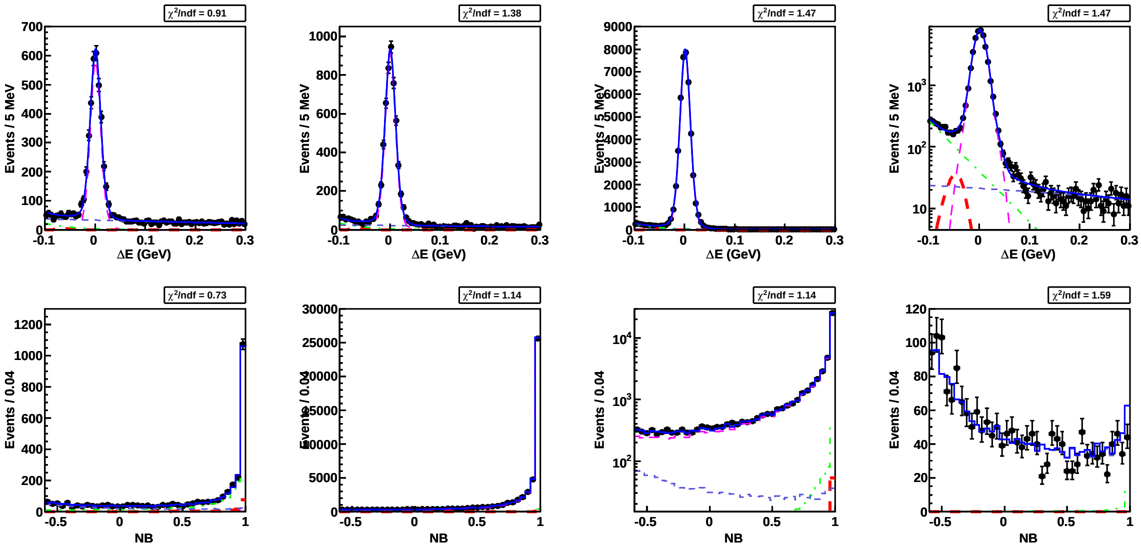


(a) The histograms of the data samples (upper) and the fitted PDFs (lower).

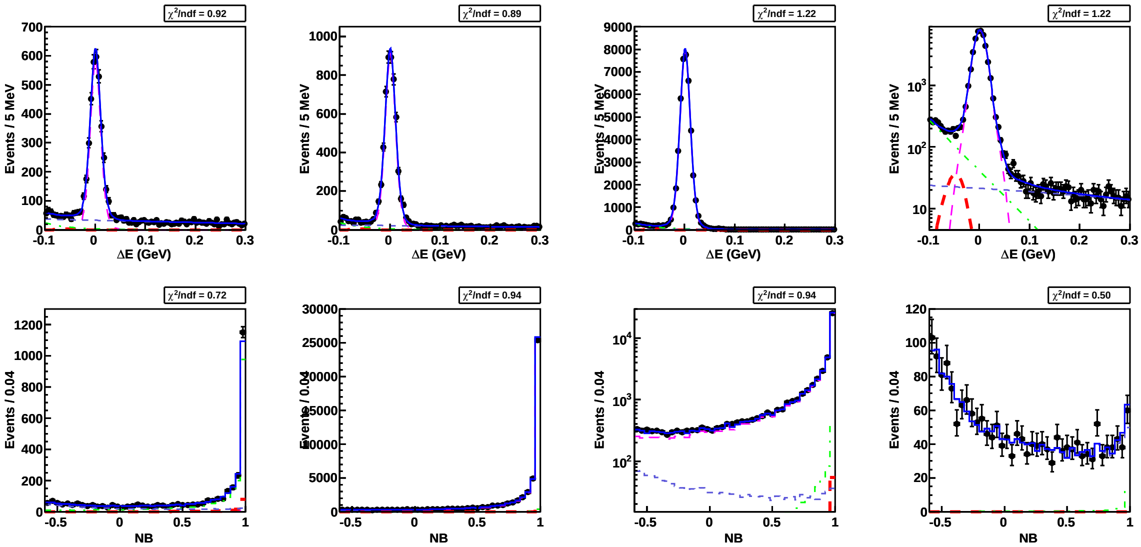


(b) The ΔE distributions (upper) and the NB distributions (lower), both of which are obtained by projecting all fitted regions. The fitted sample is shown with dots with error bars and the PDF is shown with the solid blue curve, for which the components are shown with thicker long-dashed red (DK), thinner long-dashed magenta ($D\pi$), dash-dotted green ($B\bar{B}$ background), and dashed blue ($q\bar{q}$ background).

Figure B.1: The result of the fit to the favored modes on MC. The distributions for $D\pi^-$, $D\pi^+$, DK^-, and DK^+ are shown from left to right.

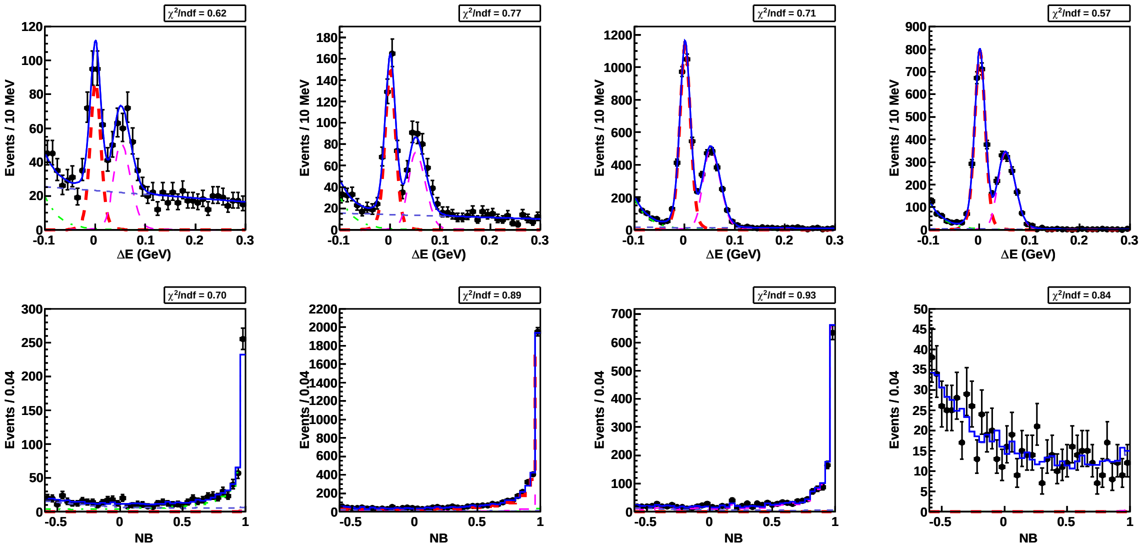


(a) The projections for the $D\pi^-$ mode.

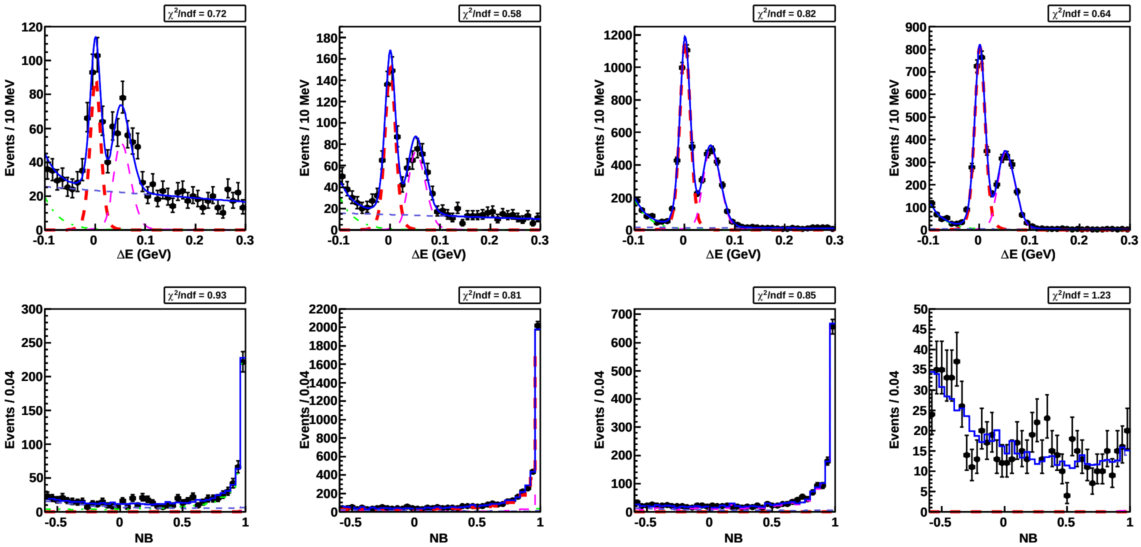


(b) The projections for the $D\pi^+$ mode.

Figure B.2: The projections for the favored $D\pi$ sample on MC. The ΔE distributions for $NB < 0$, $0 < NB < 0.5$, $NB > 0.5$, and $NB > 0.5$ (log-scaled) are shown from left to right (upper). The NB distributions for $\Delta E < -0.05$ GeV, $|\Delta E| < 0.04$ GeV, $|\Delta E| < 0.04$ GeV (log-scaled), and $\Delta E > 0.15$ GeV are shown from left to right (lower).



(a) The projections for the DK^- mode.

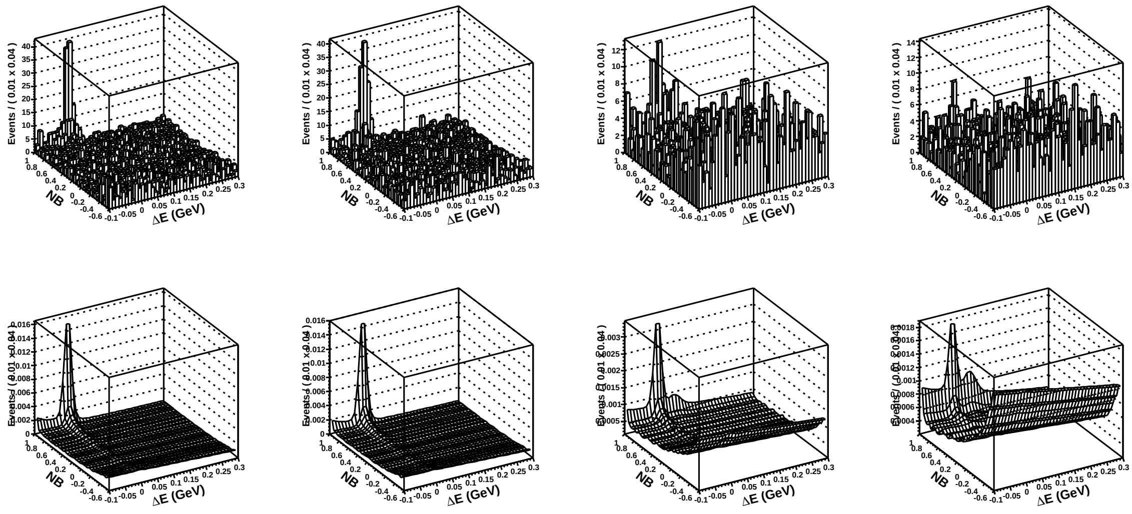


(b) The projections for the DK^+ mode.

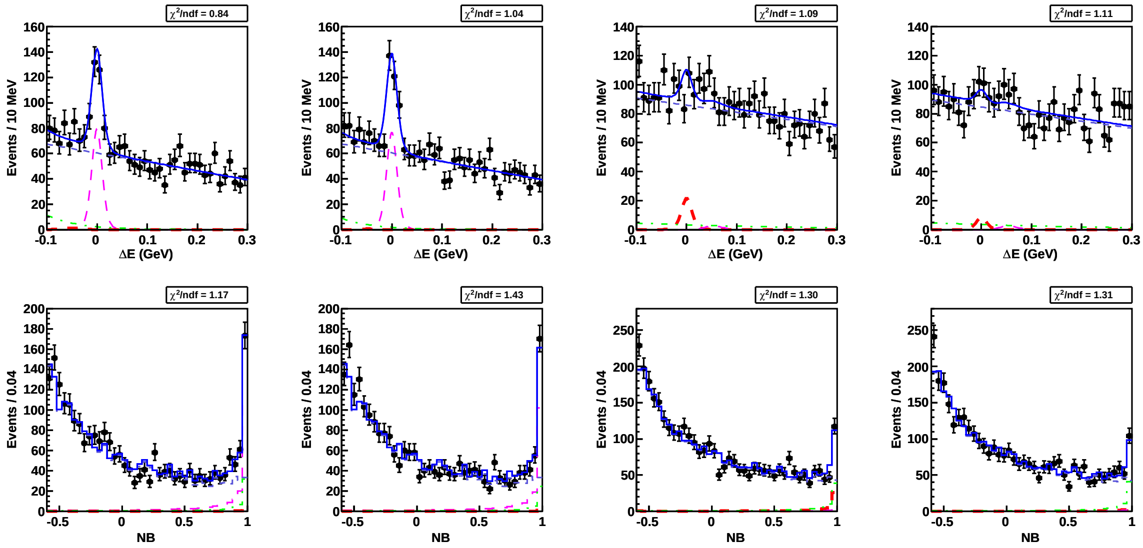
Figure B.3: The projections for the favored DK sample on MC. The ΔE distributions for $NB < 0$, $0 < NB < 0.5$, $NB > 0.5$, and $NB > 0.9$ are shown from left to right (upper). The NB distributions for $\Delta E < -0.05$ GeV, $|\Delta E| < 0.04$ GeV, $0.05 \text{ GeV} < \Delta E < 0.1$ GeV, and $\Delta E > 0.15$ GeV are shown from left to right (lower).

Component	Type	Parameter	Value
$D\pi$	General	Yield	104666 ± 349
		Asymmetry	$(-0.10 \pm 0.31)\%$
	D. Gaussian (ΔE)	Mean	0.0011 ± 0.0000
		σ_1	0.0100 ± 0.0001
		σ_2/σ_1	1.86 ± 0.05
		Yield$_1$/Yield	$(80.8^{+1.9}_{-2.1})\%$
DK	General	Yield	7715 ± 96
		Asymmetry	$(-0.8 \pm 1.2)\%$
$D\pi$ in DK	General	Yield in DK/Yield in $D\pi$	$(5.24 \pm 0.08)\%$
	D. Bif. Gaussian (ΔE)	Mean	0.0516 ± 0.0003
		$\sigma_{1,\mathrm{right}}$	0.0191 ± 0.0003
		$\sigma_{1,\mathrm{left}}/\sigma_{1,\mathrm{right}}$	0.754 (fixed)
		$\sigma_2/\sigma_{1,\mathrm{right}}$	2.48 (fixed)
		Yield$_1$/Yield	95.0% (fixed)
DK in $D\pi$	General	Yield in $D\pi$/Yield in DK	10.3% (fixed)
	D. Bif. Gaussian (ΔE)	Mean	-0.0463 (fixed)
		$\sigma_{1,\mathrm{right}}$	0.0139 (fixed)
		$\sigma_{1,\mathrm{left}}/\sigma_{1,\mathrm{right}}$	1.29 (fixed)
		$\sigma_2/\sigma_{1,\mathrm{right}}$	4.77 (fixed)
		Yield$_1$/Yield	85.1% (fixed)
$B\bar{B}$ in $D\pi$	General	Yield	7008^{+176}_{-177}
		Asymmetry	$(-1.7 \pm 1.6)\%$
	Exponential (ΔE)	Coefficient	$-18.9^{+0.5}_{-0.6}$
$B\bar{B}$ in DK	General	Yield	1420^{+59}_{-57}
		Asymmetry	$(1.5 \pm 3.2)\%$
	Exponential (ΔE)	Coefficient	$-34.2^{+2.0}_{-2.1}$
$q\bar{q}$ in $D\pi$	General	Yield	12239^{+152}_{-151}
		Asymmetry	0 (fixed)
	Linear (ΔE)	Slope	-0.27 ± 0.02
$q\bar{q}$ in DK	General	Yield	4179^{+89}_{-88}
		Asymmetry	0 (fixed)
	Linear (ΔE)	Slope	-0.21 ± 0.03

Table B.1: The list of parameters in the fit to the favored modes on MC.

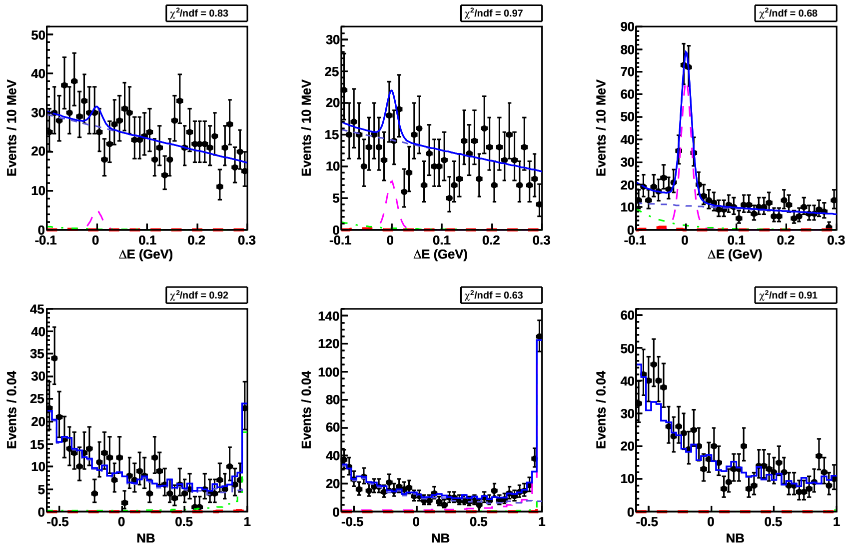


(a) The histograms of the data samples (upper) and the fitted PDFs (lower).

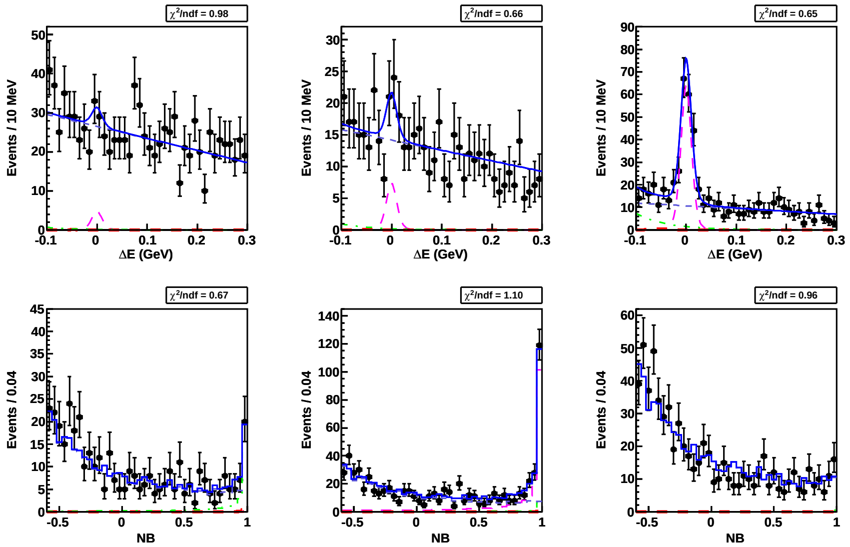


(b) The ΔE distributions (upper) and the NB distributions (lower), both of which are obtained by projecting all fitted regions. The fitted sample is shown with dots with error bars and the PDF is shown with the solid blue curve, for which the components are shown with thicker long-dashed red (DK), thinner long-dashed magenta ($D\pi$), dash-dotted green ($B\bar{B}$ background), and dashed blue ($q\bar{q}$ background).

Figure B.4: The result of the fit to the suppressed modes on MC. The distributions for $D\pi^-$, $D\pi^+$, DK^-, and DK^+ are shown from left to right.

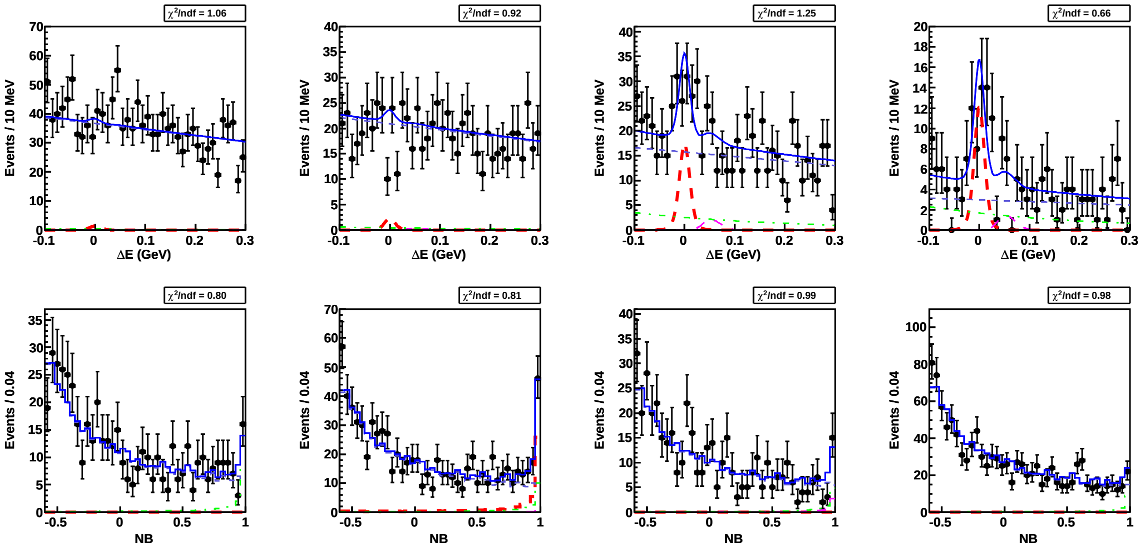


(a) The projections for the $D\pi^-$ mode.

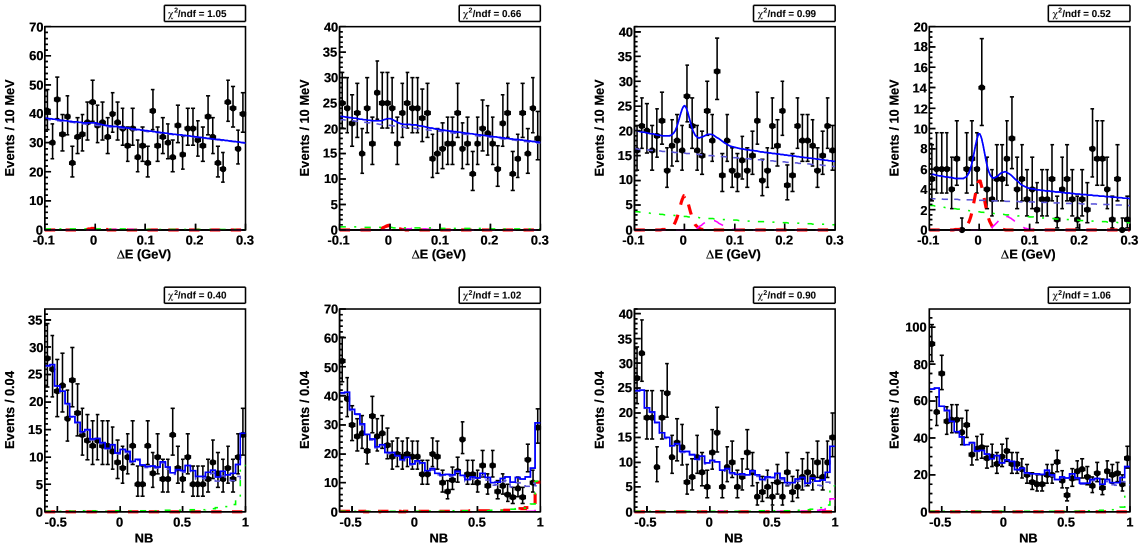


(b) The projections for the $D\pi^+$ mode.

Figure B.5: The projections for the suppressed $D\pi$ sample on MC. The ΔE distributions for $NB < 0$, $0 < NB < 0.5$, and $NB > 0.5$ are shown from left to right (upper). The NB distributions for $\Delta E < -0.05$ GeV, $|\Delta E| < 0.04$ GeV, and $\Delta E > 0.15$ GeV are shown from left to right (lower).



(a) The projections for the DK^- mode.



(b) The projections for the DK^+ mode.

Figure B.6: The projections for the suppressed DK sample on MC. The ΔE distributions for $NB < 0$, $0 < NB < 0.5$, $NB > 0.5$, and $NB > 0.9$ are shown from left to right (upper). The NB distributions for $\Delta E < -0.05$ GeV, $|\Delta E| < 0.04$ GeV, 0.05 GeV $< \Delta E <$ 0.1 GeV, and $\Delta E >$ 0.15 GeV are shown from left to right (lower).

Component	Type	Parameter	Value
$D\pi$	General	Yield	428^{+26}_{-25}
		Asymmetry	$(2.0^{+5.8}_{-5.9})\%$
	D. Gaussian (ΔE)	Mean	0.0011 (fixed)
		σ_1	0.0100 (fixed)
		σ_2/σ_1	1.86 (fixed)
		$\text{Yield}_1/\text{Yield}$	80.8% (fixed)
DK	General	Yield	$81.3^{+16.3}_{-15.4}$
		Asymmetry	$(43 \pm 19)\%$
$D\pi$ in DK	General	Yield in DK/Yield in $D\pi$	5.24% (fixed)
	D. Bif. Gaussian (ΔE)	Mean	0.0516 (fixed)
		$\sigma_{1,\text{right}}$	0.0191 (fixed)
		$\sigma_{1,\text{left}}/\sigma_{1,\text{right}}$	0.754 (fixed)
		$\sigma_2/\sigma_{1,\text{right}}$	2.48 (fixed)
		$\text{Yield}_1/\text{Yield}$	95.0% (fixed)
DK in $D\pi$	General	Yield in $D\pi$/Yield in DK	10.3% (fixed)
	D. Bif. Gaussian (ΔE)	Mean	−0.0463 (fixed)
		$\sigma_{1,\text{right}}$	0.0139 (fixed)
		$\sigma_{1,\text{left}}/\sigma_{1,\text{right}}$	1.29 (fixed)
		$\sigma_2/\sigma_{1,\text{right}}$	4.77 (fixed)
		$\text{Yield}_1/\text{Yield}$	85.1% (fixed)
$B\bar{B}$ in $D\pi$	General	Yield	127^{+27}_{-24}
		Asymmetry	$(12 \pm 17)\%$
	Exponential (ΔE)	Coefficient	$-16.0^{+3.4}_{-4.5}$
$B\bar{B}$ in DK	General	Yield	209^{+35}_{-34}
		Asymmetry	$(-4 \pm 16)\%$
	Exponential (ΔE)	Coefficient	$-3.2^{+1.5}_{-1.6}$
$q\bar{q}$ in $D\pi$	General	Yield	4273^{+69}_{-68}
		Asymmetry	0 (fixed)
	Linear (ΔE)	Slope	-0.26 ± 0.03
$q\bar{q}$ in DK	General	Yield	6425^{+86}_{-85}
		Asymmetry	0 (fixed)
	Linear (ΔE)	Slope	-0.12 ± 0.02

Table B.2: The list of parameters in the fit to the suppressed modes on MC.

Bibliography

[1] N. Cabibbo, "Unitary Symmetry and Leptonic Decays," Phys. Rev. Lett. **10**, 531–533 (1963).

[2] M. Kobayashi and T. Maskawa, "CP-Violation in the Renormalizable Theory of Weak Interaction," Prog. Theor. Phys. **49**, 652–657 (1973).

[3] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2π Decay of the K_2^0 Meson," Phys. Rev. Lett. **13**, 138–140 (1964).

[4] S. W. Herb *et al.*, "Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions," Phys. Rev. Lett. **39**, 252–255 (1977).

[5] F. Abe *et al.* (CDF Collaboration), "Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab," Phys. Rev. Lett. **74**, 2626–2631 (1995).

[6] S. Abachi *et al.* (D0 Collaboration), "Observation of the Top Quark," Phys. Rev. Lett. **74**, 2632–2637 (1995).

[7] B. Aubert *et al.* (BaBar Collaboration), "Observation of CP Violation in the B^0 Meson System," Phys. Rev. Lett. **87**, 091801 (2001).

[8] K. Abe *et al.* (Belle Collaboration), "Observation of Large CP Violation in the Neutral B Meson System," Phys. Rev. Lett. **87**, 091802 (2001).

[9] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa Matrix," Phys. Rev. Lett. **51**, 1945–1947 (1983).

[10] I. I. Bigi and A. I. Sanda, "On Direct CP Violation in $B \to \overset{(-)}{D}{}^0 K\pi$'s versus $\bar{B} \to \overset{(-)}{D}{}^0 \bar{K}\pi$'s Decays," Phys. Lett. B **211**, 213–220 (1988).

[11] M. Gronau and D. London, "How to Determine All the Angles of the Unitarity Triangle from $B_d^0 \to DK_S$ and $B_s^0 \to D\phi$," Phys. Lett. B **253**, 483–488 (1991); M. Gronau and D. Wyler, "On Determining a Weak Phase from Charged B Decay Asymmetries," Phys. Lett. B **265**, 172–176 (1991).

[12] D. Atwood, I. Dunietz, and A. Soni, "Enhanced CP Violation with $B \to KD^0(\bar{D}^0)$ Modes and Extraction of the Cabibbo-Kobayashi-Maskawa Angle γ," Phys. Rev. Lett. **78**, 3257–3260 (1997); D. Atwood, I. Dunietz, and A. Soni, "Improved Methods for Observing CP Violation in $B^\pm \to KD$ and Measuring the CKM Phase γ," Phys. Rev. D **63**, 036005 (2001).

[13] A. Giri, Yu. Grossman, A. Soffer, and J. Zupan, "Determining γ Using $B^\pm \to DK^\pm$ with Multibody D Decays," Phys. Rev. D **68**, 054018 (2003); A. Bondar, in Proceedings of BINP Special Analysis Meeting on Dalitz Analysis, 2002 (unpublished).

[14] Y. Grossman, A. Soffer, and J. Zupan, "Effect of D–$\bar{D}$ Mixing on the Measurement of γ in $B \to DK$ Decays," Phys. Rev. D **72**, 031501(R) (2005).

[15] D. Asner *et al.* (Heavy Flavor Averaging Group), "Averages of b-Hadron, c-Hadron, and τ-Lepton Properties," arXiv:1010.1589, and online update at http://www.slac.stanford.edu/xorg/hfag.

[16] A. Poluektov *et al.* (Belle Collaboration), "Evidence for Direct CP Violation in the Decay $B^\pm \to D^{(*)}K^\pm, D \to K^0_S\pi^+\pi^-$ and Measurement of the CKM Phase ϕ_3," Phys. Rev. D **81**, 112002 (2010).

[17] P. del Amo Sanchez *et al.* (BaBar Collaboration), "Evidence for Direct CP Violation in the Measurement of the Cabibbo-Kobayashi-Maskawa Angle γ with $B^\mp \to D^{(*)}K^{(*)\mp}$ Decays," Phys. Rev. Lett. **105**, 121801 (2010).

[18] J. Charles *et al.* (CKMfitter Group), "CP Violation and the CKM Matrix: Assessing the Impact of the Asymmetric B Factories," Eur. Phys. J. C **41**, 1–131 (2005), and updated results and plots at http://ckmfitter.in2p3.fr.

[19] Y. Horii *et al.* (Belle Collaboration), "Study of the Suppressed B Meson Decay $B^- \to DK^-$, $D \to K^+\pi^-$," Phys. Rev. D **78**, 071901(R) (2008).

[20] P. del Amo Sanchez *et al.* (BaBar Collaboration), "Search for $b \to u$ Transitions in $B^- \to DK^-$ and D^*K^- Decays," Phys. Rev. D **82**, 072006 (2010).

[21] S. Kurokawa and E. Kikutani, "Overview of the KEKB Accelerators," Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1–7 (2003), and other papers included in this volume.

[22] K. Nakamua *et al.* (Particle Data Group), "Review of Particle Physics," J. Phys. G **37**, 075021 (2010).

[23] A. Abeshian *et al.* (Belle Collaboration), "The Belle Detector," Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117–232 (2002).

[24] Z. Natkaniec *et al.* (Belle SVD2 Group), "Status of the Belle Silicon Vertex Detector," Nucl. Instrum. Methods Phys. Res., Sect. A **560**, 1–4 (2006).

[25] G. C. Fox and S. Wolfram, “Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes,” Phys. Rev. Lett. **41**, 1581–1585 (1978).

[26] R. A. Fisher, “The Use of Multiple Measurements in Taxonomic Problems,” Ann. Eugenics **7**, 179–188 (1936).

[27] S. H. Lee *et al.* (Belle Collaboration), “Evidence for $B^0 \to \pi^0\pi^0$,” Phys. Rev. Lett. **91**, 261801 (2003).

[28] N. Abe, “Measurement of a CP Violation Parameter ϕ_2 through B Meson Decay into $\pi^+\pi^-$,” Ph.D. thesis (2003).

[29] H. Tajima *et al.*, “Proper-Time Resolution Function for Measurement of Time Evolution of B Mesons at the KEK B-Factory,” Nucl. Instrum. Methods Phys. Res., Sect. A **533**, 370–386 (2004).

[30] H. Kakuno *et al.*, “Neutral B Flavor Tagging for the Measurement of Mixing-Induced CP Violation at Belle,” Nucl. Instrum. Methods Phys. Res., Sect. A **533**, 516–531 (2004).

[31] M. Feindt and U. Kerzel, “The NeuroBayes Neural Network Package,” Nucl. Instrum. Methods Phys. Res., Sect. A **559**, 190–194 (2006).

[32] C.-K. Chua, “Charmless Three-Body Decays of B Mesons,” J. Phys.: Conf. Ser. **110**, 052008 (2008).

[33] B. Aubert *et al.* (BaBar Collaboration), “Observation of the Decay $B^+ \to K^+K^-\pi^+$,” Phys. Rev. Lett. **99**, 221801 (2007).