Search for rare decays $B^0 \to K^+K^-\pi^0$ with Belle data

THESIS SUBMITTED FOR THE AWARD OF THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

PHYSICS

BY

VIPIN GAUR

Under the Supervision of

Prof. Mohammad Zafar

(Aligarh Muslim University, Aligarh)

and

Prof. Tariq Aziz

(Tata Institute of Fundamental Research, Mumbai)

DEPARTMENT OF PHYSICS, ALIGARH MUSLIM UNIVERSITY

ALIGARH 202002, INDIA

2013

PHYSICS DEPARTMENT, ALIGARH MUSLIM UNIVERSITY
ALIGARH – 202002, INDIA

Prof. Mohammad Zafar
Phone: 91-9359903848
email: mzafar.ph1@gmail.com

CERTIFICATE

Certified that this thesis entitled, "Search for rare decays $B^0 \to K^+K^-\pi^0$ with Belle data", embodies the original work of Mr. Vipin Gaur carried out under my supervision and co-supervision of Prof. Tariq Aziz of Tata Institute of Fundamental Research, Mumbai. This work is worthy of consideration for the award of Ph.D. degree.

Professor Mohammad Zafar

TATA INSTITUTE OF FUNDAMENTAL RESEARCH

National Centre of the Government of India for Nuclear Science and Mathematics

Homi Bhabha Road, Mumbai 400 005, India

Prof. Tariq Aziz

Phone: 91-22-2278 2358

Fax: 91-22-2280 4610

email: aziz@tifr.res.in

CERTIFICATE

Certified that this thesis entitled, "Search for rare decays $B^0 \to K^+K^-\pi^0$ with Belle data", embodies the original work of Mr. Vipin Gaur carried out under my co-supervision. This work is worthy of consideration for the award of Ph.D. degree.

Professor Tariq Aziz

Acknowledgements

All the work presented in this document would not have been possible without the help from many people. I would like to take this opportunity to convey my heartfelt thanks to them.

I owe a great debt of gratitude to my thesis supervisors, Prof. M. Zafar and Prof. T. Aziz. Prof. Zafar gave me the opportunity to get involved in the beautiful world of experimental high energy physics and subsequently be a part of Belle collaboration experiment through TIFR (Tata Institute of Fundamental Research), Mumbai, where I worked under the supervision of Prof. Aziz and Dr. G. Mohanty as a research fellow. Most of the thesis work has been carried out in meticulous guidance of Dr. Mohanty. I am grateful to him for his constant encouragements in carrying out a high quality research.

I am grateful to Prof. Y. Sakai, Prof. M. Nakao, Prof. T. Hara, Prof. S. Nishida and Dr. K. Trabelsi for providing me with critical insights during my discussions with them at KEK (High Energy Accelerator Research Organization), Japan. I would like to extend my gratitude to Prof. M. Nakao, Dr. Y. Unno and Prof. C. H. Wang for internally refereeing my analysis and helping us establish its correctness and authenticity. The analysis was refined and presented many times in physics group meetings chaired by Dr. Unno and Dr. Mohanty. I am especially thankful to Dr. Unno for several fruitful hours of discussions with him in Tsukuba hall of KEK. I would like to thank Prof. T. Browder, Prof. Y. Sakai, Prof. M. Nakao, Prof. L. Piilonen, Prof. S. Eydelman, Prof. V. Savinov, Prof. Y. Watanabe, Prof. M. Sevior, Prof. A. Schwartz, Prof. B. Yabsley and Dr. T. Aushev for their useful suggestions and comments during - collaboration wide review of my paper draft and presentations in Belle analysis meetings & Belle general meetings. Special thanks to my senior colleagues: Dr. Y. Horii, Dr. N. Joshi, Minori-San, Dr. H. Sahoo, Dr. V. Bhardwaj, A. Das and Dr. K. Sakai. I have learned methods of physics analysis from them. I am sure that this thesis did not exist without their great support.

I formally wish to thank Chairman, Department of Physics, AMU (Aligarh Muslim University), Aligarh, and Chairman, Department of High Energy Physics, TIFR, Mumbai, for providing me with an excellent work environment.

I would like to thank all my professors and colleagues at AMU ,TIFR and KEK for their support. Special thanks to Shakeel sahab, Anuj, Kausar bhai, Bibhuti, Sanjeev, Saurabh, Nisar, Minakshi, Puneet and Vijay for being there for me always.

It would not have been possible for me to achieve this milestone in my career without the love and support of my family. There are no words to express how grateful I am towards them. I thank all my family members for their support.

Last but not the least, without the financial support by TIFR, nothing of this would have been possible, hence a special thanks to TIFR.

Finally, I would like to thank all those, whose direct and indirect support helped me in completing my thesis in time.

Vipin Gaur

Contents

List of Tables

List of Figures

1 Introduction

1.1 The Standard Model

The Standard Model (SM) describes the world of elementary particles. The fundamental matter particles are spin-1/2 fermions while the interactions between them are mediated by spin-1 bosons. The elementary fermions are leptons and quarks. There are three types (generations) of lepton; each contains a charged particle e^-, μ^-, or τ^-, with electric charge $-e_0 = -1.6 \times 10^{-19}$ C, and a corresponding neutral partner ν_e, ν_μ, or ν_τ neutrino. Quarks fall into three generations and come in six different flavors: the up-type quarks, *u*, *c* and *t* have the electric charge $+2e_0/3$, whereas the charge of down-type quarks, *d*, *s* and *b* is $-e_0/3$. In addition to the electric charge, quarks also carry the color charge. In the SM, the interactions between matter particles result from matter particles exchanging force-mediating particles. In total, there are 12 particles of the latter type in the SM: eight gluons (*g*), photon (γ) and three weak bosons, $W^\pm$ and Z^0. The gluon is responsible for the most powerful force, the strong force, which binds together quarks inside protons and neutrons, and holds together protons and neutrons inside an atomic nucleus. The gluons themselves carry color charge, and therefore (like the quarks) do not exist as isolated particles. The photon is the electromagnetic force carrier. Lastly, the $W^\pm$ and Z^0 bosons are attributed to the weak force, playing a role in the radioactive decay. The weak force is very important in observing neutrino reactions, because the neutrinos are impervious to the electromagnetic force (due to their lack of charge) and unaffected by the strong force (which governs nuclear interactions), leaving only the weak force to characterize them. Particles that make up the ordinary matter, acquire their mass through interaction with Higgs boson (more specifically with the Higgs field). It is the only SM particle not yet experimentally observed. Recently, CMS and ATLAS experiments at CERN in their search for the SM Higgs boson, reported clear evidence for the production of a neutral boson (which can possibly be the SM Higgs) with a measured mass of 126 ± 0.4 GeV/c^2 and spin different from 1 [1, 2].

The weak interaction, mediated by $W^\pm$ and Z^0 bosons, is the only interaction in the SM that can change the flavor of the quarks. A down-type quark can emit or absorb a *W* boson and convert into an up-type quark, and vice versa. In this transformation, a quark is allowed only to change its charge by a unit amount e_0 (the charge of the electron). Because quarks can change flavor by the weak interaction, only the lightest quarks and leptons (namely, up and down quarks in the form of protons and neutrons, and electrons) are included in the stable matter of the world around us – all heavier ones decay to one or another of the lighter ones. The flavor-changing transitions occur because the quark mass eigenstates are not the same as the weak-interaction eigenstates. In 1963, Nicola

Cabibbo introduced an angle to preserve the universality of the weak interaction [3]. The so called Cabibbo angle (θ_c) is related to the relative probability of the down and strange quarks decaying into up quarks ($|V_{ud}|^2$ and $|V_{us}|^2$, respectively). In the language of particle physics, the object that couples to the up quark via charged-current weak interaction is a superposition of down-type quarks, here denoted by d'. Mathematically, this is:

$$d' = V_{us}s + V_{ud}d,$$

or, using the Cabbibo angle:

$$d' = \sin\theta_c s + \cos\theta_c d.$$

The Cabbibo angle can be calculated by using the relation:

$$\tan\theta_c = \frac{|V_{us}|}{|V_{ud}|} = \frac{0.2257}{0.97419} \rightarrow \theta_c = 13.04^o$$

When the charm quark was discovered in 1974 [4, 5], it was noticed that the down and strange quark could decay into either the up or charm quark, leading to two sets of equations:

$$\begin{aligned} d' &= V_{us}s + V_{ud}d; \\ s' &= V_{cs}s + V_{cd}d, \end{aligned}$$

This can also be written in the matrix notation as:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} \\ V_{cd} & V_{cs} \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix},$$

or using the Cabibbo angle

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix},$$

where the various $|V_{ij}|^2$ represent the probability that the quark of flavor i decays into a quark of flavor j. This 2×2 rotation matrix is called the Cabibbo matrix.

For two generations of quarks, there are no *CP* violating phases. To explain *CP* violation within the Cabibbo scheme a complex number in the rotation matrix is needed but such a term could always be eliminated by the suitable redefinition of the quark phases. In 1964 *CP* violation was first observed in the neutral kaon system [6]. Kobayashi and Maskawa extended Cabibbo's idea of flavor mixing to accommodate the phenomena of *CP* violation within the SM, by proposing a possibility of the third generation [7]. The

quark flavor mixing mechanism introduced by Kobayashi and Maskawa was certainly a bold step, considering the fact that not even the charm quark, the heavier member of the second quark family, was hinted from the experimental observations, so far. After the discovery of charm quark in 1974, bottom quark was discovered at Fermilab in 1976 [8] therefore immediately triggering the search for the missing third-generation quark, the top quark. Finally the top quark was discovered at Fermilab by CDF and D0 experiments [9, 10].

If we look at all the ways in which one quark can turn into another quark with a charge change of e_0, that's just all quarks with charge $+2e_0/3(\mathrm{u},\ \mathrm{c},\ \text{or}\ \mathrm{t})$ paired with quarks with charge $-e_0/3(\mathrm{d},\ \mathrm{s},\ \text{or}\ \mathrm{b})$. That's nine possible pairings. Kobayashi and Maskawa introduced a 3×3 unitary matrix (CKM matrix) that contains the information on the strength of flavor-changing transitions.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{td} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

In above equation, on the left side is the weak interaction partners of up-type quarks, and on the right is the CKM matrix along with a vector of mass eigenstates of down-type quarks. The mass spectra of the up- and down-quarks sectors and the CKM matrix are shown to have clear and distinctive hierarchical structures, as shown in Fig. 1. The magnitude of the CKM matrix elements decreases from the diagonal to the off-diagonal elements.

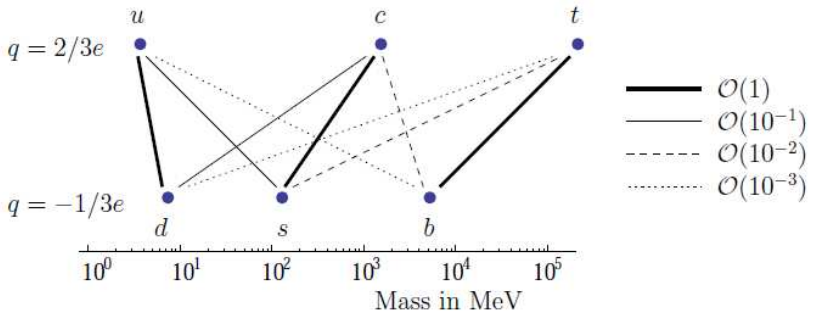


Figure 1: The spectrum of quark masses and possible charged-current transitions in which the quark charge changes by one unit. The thickness of the lines indicates the strength of the corresponding transitions.

The triumph of the SM is that it predicts a distinct set of relationships between the nine elements of the CKM matrix and that the matrix elements have certain properties that result in *CP* violation. In particular, the *CP* violation is related to the fact that the

matrix elements include imaginary numbers. If the CKM parameterisation is correct then all mixing and weak decays must give consistent results and the phase we extract from *CP* violation in the *K* system will predict the amount of *CP* violation in the *B* system. The CKM parameterisation allows us to makes many measurements and to consequently test whether the CKM formalism works. In 2008 half of the Nobel Prize in Physics was jointly awarded to Makoto Kobayashi and Toshihide Maskawa "for the discovery of the origin of the broken symmetry which predicts the existence of at least three families of quarks in nature". Kobayashi and Maskawa's explanation of the origin of *CP* symmetry violation predicted that large *CP* asymmetries could be observed in the decays of particles containing b-quarks. Verification of their hypothesis was one of the primary goals of the B-factory experiments, Belle at KEKB and BaBar at PEP-II. They did the job exceedingly well!

Currently, the best determination of the magnitudes of the CKM matrix elements [11] is:

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{td}| \end{pmatrix} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix}$$

In Wolfenstein parametrization [12], the CKM matrix is expanded in powers of a small parameter λ. The Wolfenstein parametrization of the CKM matrix, to order λ^3, is

$$\begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix}$$

To summarize the main features of the Wolfenstein parametrization,

- The only complex parameter, $i\eta$ present in the parametrization, resides in the most off-diagonal entries, representing V_{ub} and V_{td}
- $|\lambda|$ is equal to $|V_{us}| = 0.22$
- $A \sim 0.81$, $\rho \sim 0.14$ and $\eta \sim 0.35$ are real numbers.
- The more off-diagonal one goes, the λ dependence increases, and hence the strengths decrease. Flavors among different generations show mixing with different strengths. The relatively stronger ones are referred to as "Cabibbo-favored", while the weaker are termed "Cabibbo-suppressed" that appear in rare *B* decays

1.2 The rare decay $B^0 \to K^+K^-\pi^0$

The B-meson decay $B^0 \to K^+K^-\pi^0$ is suppressed in the SM and thus offers a useful probe for new physics beyond the SM. Figure 2 shows typical Feynman diagrams that contribute to this decay. The dominant one is the color- and Cabibbo-suppressed $b \to u$ tree transition, followed by the internal W exchange diagram leading to $B^0 \to K^{*\pm}K^{\mp}$ with $K^{*\pm} \to K^{\pm}\pi^0$. The latter diagram dominantly contributes to the decay $B^0 \to K^+K^-$, for which only upper limits on the branching fraction have been set [13, 14, 15, 16]. This is in contrast to the related decays (having two kaons in the final state) that are already observed such as $B^0 \to K^0\bar{K}^0$, $B^+ \to K^0K^+$ [16, 17] and $B^+ \to K^+K^-\pi^+$ [18, 19], where the $b \to d$ gluonic penguin amplitude can contribute as well [20].

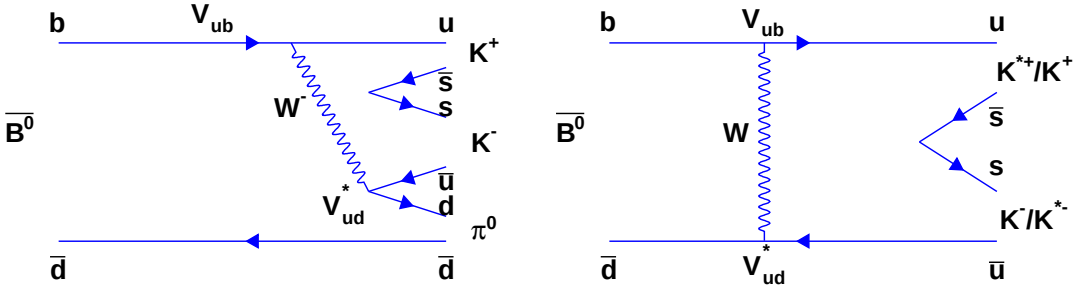


Figure 2: color- and Cabibbo-suppressed $b \to u$ tree (left) and internal W-exchange (right) diagrams contributing to the $K^+K^-\pi^0$ final state.

The three-body decay $B^0 \to K^+K^-\pi^0$ has not yet been observed with the only upper limit being available at 90 % confidence level, $\mathcal{B}(B^0 \to K^+K^-\pi^0) < 19 \times 10^{-6}$, from the CLEO Collaboration [21]. Intermediate resonant modes that preferentially decay to this final state have also not been seen. A search for a related channel by Belle has set an upper limit $\mathcal{B}(B^0 \to \phi\pi^0) < 1.5 \times 10^{-7}$ [22]. Practically, no experimental information is available for other potential resonance modes, *e.g.*, $K^*(892)^{\pm}K^{\mp}$, $K_0^*(1430)^{\pm}K^{\mp}$ and $f_0(980)\pi^0$.

The study of the $B^0 \to K^+K^-\pi^0$ decay is further motivated due to the importance of the aforementioned resonant modes. The decay $B^0 \to \phi\pi^0$, in particular, is sensitive to possible beyond-the-SM contributions. A measurement of branching fraction of $O(10^{-7})$ would be evidence for new physics, for example, the supersymmetry [23]. Study of the process $B^0 \to \phi\pi^0$ is also important to understand the theory uncertainties associated with the measurement of CP violation asymmetries in the $B^0 \to \phi K^0$ decays. At the subleading order, $B \to \phi\pi$ decay amplitudes are related to that of $B^0 \to \phi K^0$ [24], and can therefore provide stringent bounds on possible contribu-

tions to the time-dependent *CP* asymmetry in $B^0 \to \phi K^0$ [25]. Various other predictions available in the literature are summarized in Table 1.

Table 1: Predictions for various branching fractions relevant for the $K^+K^-\pi^0$ final state. We attempt here to give the state-of-art predictions that rely on experimental inputs.

Mode	Technique	Prediction	Ref.
$B^0 \to K^{*\pm}(892)K^{\mp}$	QCD factorization	1.9×10^{-8}	[26]
	QCD factorization	2.8×10^{-8}	[27]
	perturbative QCD	$(7.4^{+1.0}_{-1.3}) \times 10^{-8}$	[28]
$B^0 \to \phi(1020)\pi^0$	QCD factorization	1.0×10^{-9}	[26]
	QCD factorization	2.0×10^{-9}	[27]
	flavor SU(3)	$(1.0-2.0) \times 10^{-9}$	[29]
$B^0 \to f_0(980)\pi^0$	None known		

Another motivation for the study of $B^0 \to K^+K^-\pi^0$ comes from the observation of $B^+ \to K^+K^-\pi^+$ by the BaBar Collaboration [18]. In particular, an unexpected structure , we call $f_X(1500)$, is seen near $1.5\,\mathrm{GeV/c^2}$ in the K^+K^- invariant-mass spectrum [see Fig. 3 (left)], which accounts for about half of the total events. Similar structures have been observed in the Dalitz-plot analyses of $B^+ \to K^+K^-K^+$ [Fig. 3(right)] and $B^0 \to K^+K^-K^0_s$ [Fig. 4(left)] decays [30, 31]. Recent LHCb study of $B^\pm \to K^+K^-\pi^\pm$ decays [19] has also revealed an unidentified structure [Fig. 4(right)] in the same mass range; however, it is only present in B^+ events. If these structures are due to a K^+K^- resonant state, it should show up in $B^0 \to K^+K^-\pi^0$, while if it is a reflection from the $b \to d$ penguin, it would not contribute to $K^+K^-\pi^0$. In Fig. 5 we show a possible Feynman diagram for the production of the $f_X(1500)$ in the B decay, subsequently leading to the $K^+K^-\pi^0$ final state. Since the u and d quarks are spectators in the $b \to u$ tree diagram [Fig. 2(left)] for $B^+ \to K^+K^-\pi^+$ and $B^0 \to K^+K^-\pi^0$, respectively, one can estimate the branching fraction for the latter using the BaBar measurement of $\mathcal{B}(B^+ \to K^+K^-\pi^+)$ [32]. Let's assume for the time being that $f_X(1500)$, the single largest source of the measured branching fraction for $B^+ \to K^+K^-\pi^+$, solely contributes to $B^0 \to K^+K^-\pi^0$. In that case, assuming isospin symmetry (d and u interchange in Fig. 5) we find expected number of signal events [$N_{exp}(B^0 \to K^+K^-\pi^0)$] as

$$N_{exp}(B^0 \to K^+K^-\pi^0) \approx (1/2) \times \mathcal{B}(B^+ \to K^+K^-\pi^+) \times \varepsilon_{rec} \times N_{B\bar{B}} \approx\ 400, \qquad (1)$$

where the reconstruction efficiency (ε_{rec}) and $N_{B\bar{B}}$ are 20% and 772×10^6, respectively. Thus assuming isospin symmetry and that the $b \to u$ transition is the main contributor

to $B^0 \to K^+K^-\pi^0$, we expect its branching fraction at the level of 3×10^{-6}, which is well within Belle's reach.

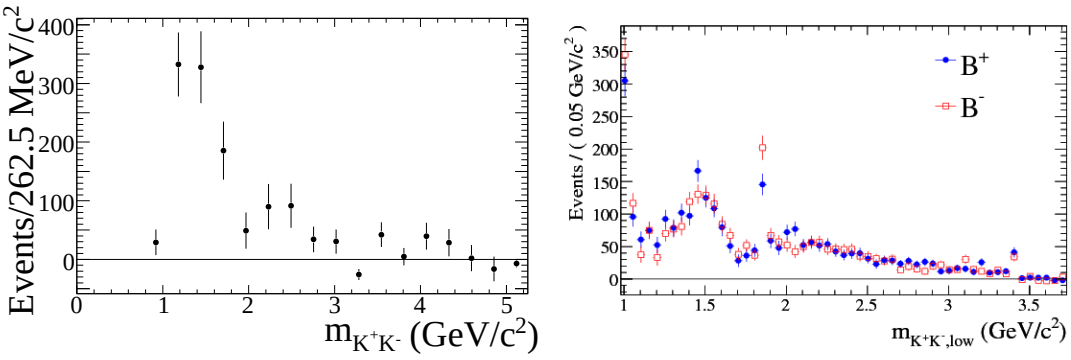


Figure 3: Efficiency-corrected distribution of the K^+K^- invariant mass in the decay $B^+ \to K^+K^-\pi^+$ (left) and signal-weighted K^+K^- distribution for $B^+ \to K^+K^-K^+$ candidates (right) from BaBar.

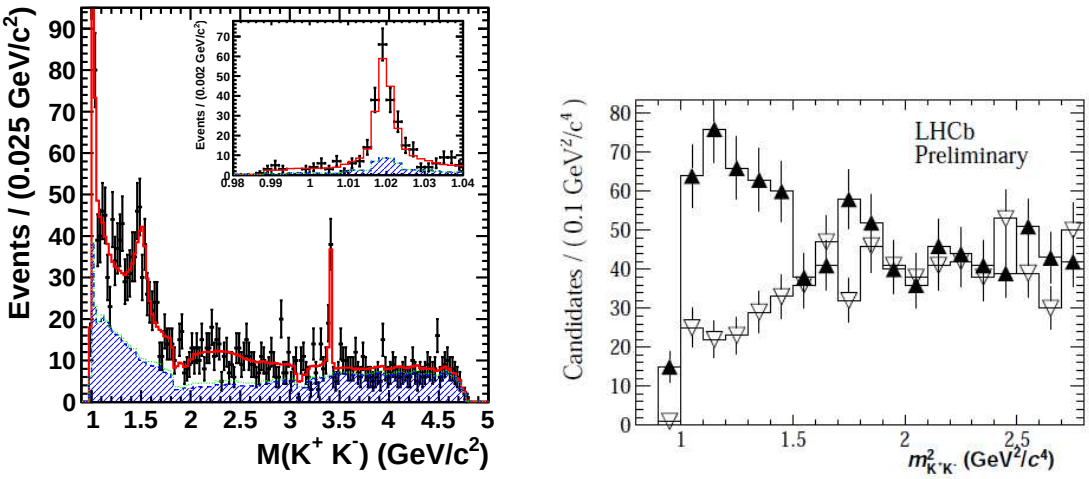


Figure 4: The K^+K^- invariant mass distribution in $B^0 \to K^+K^-K_s^0$ observed by Belle (left) and $m^2(K^+K^-)$ distribution in $B^\pm \to K^+K^-\pi^\pm$ observed by LHCb (right), where the empty triangles represent B^- and the filled triangles represent B^+ events.

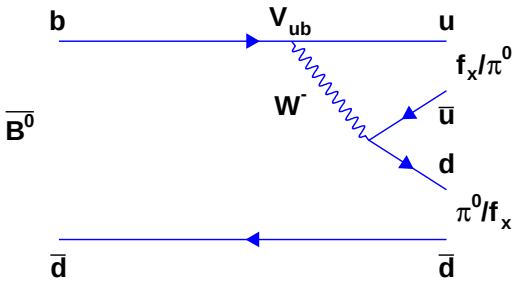


Figure 5: Possible $f_X(1500)$ production in the B decay, which subsequently leads to the $K^+K^-\pi^0$ final state.

2 KEK B-Factory

2.1 Introduction

The main goals of the B-factory are to test Kobayashi-Maskawa mechanism for *CP* violation in the *B* meson system, to precisely determine the Cabibbo-Kobayashi-Maskawa (CKM) [33, 7] matrix elements, and to search for rare *B* decays as an indirect probe for physics beyond the Standard Model (SM). To achieve all of these, the Belle experiment was commissioned at KEK (High Energy Accelerator Research Organization) B-factory in Japan. The KEK B-factory has two components: the accelerator called KEKB [34, 35] and the detector called Belle detector [36, 37, 38].

The first generation of B-factories, DORIS II at DESY in Germany with the ARGUS experiment [39], and CESR in Cornell, USA with the CLEO detector [40], were symmetric colliders, where the electron and positron beams had the same energy, corresponding to half of the $\Upsilon(4S)$ mass. On the other hand, the second generation of B-factories, KEKB at KEK in Tsukuba, Japan with the Belle experiment, and PEP-II at SLAC in Stanford, USA with the BaBar experiment [41] are asymmetric colliders.

The KEKB is an asymmetric energy e^+e^- collider with electron having energy 8.0 GeV and positron having energy 3.5 GeV. The large asymmetry of energy makes the flight length of the *B* mesons produced long enough that one can measure. The center-of-mass (CM) energy $\sqrt{s}$ is:

$$\sqrt{s} = \sqrt{4E_{e^+}E_{e^-}} = 10.58 \text{ GeV}, \tag{2}$$

which is equal to the mass of the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ resonance is a vector meson comprising of a b quark and its antiparticle $\bar{b}$. The $\Upsilon(4S)$ decays via the strong force almost instantly to a $B\bar{B}$ meson pair. The colliding beam energy asymmetry causes the $\Upsilon(4S)$ to have a non-zero velocity in the laboratory frame. So, the $B\bar{B}$ meson pair is created with a Lorentz boost factor of:

$$\gamma_\beta = \frac{E_{e^-} - E_{e^+}}{\sqrt{s}} = 0.425, \tag{3}$$

along the electron beam direction (z axis). Since the *B* mesons from the $\Upsilon(4S)$ decay have a very little kinetic energy in the CM frame, they have approximately the same boost γ_β in the laboratory frame. This boost is needed for the study of time-dependent CP asymmetries. *B* mesons fly about 0.2 mm before decaying and the decay-length difference between the two *B* mesons along the z axis (Δz) can be measured with the

silicon vertex detector. The B mesons decay-time difference (Δt) can then be estimated as:

$$\Delta t \approx \frac{\Delta z}{\gamma_\beta c} \tag{4}$$

The design luminosity of KEKB machine is $1 \times 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ which corresponds to production of 10^8 B mesons per year. The physics of B mesons is also investigated at hadron colliders, such as the Tevatron, a $p\bar{p}$ collider with $\sqrt{s} = 1.96$ TeV, located at Fermilab near Chicago, USA with the DZero [42] and CDF experiments [43]. The ATLAS [44] and CMS experiments [45], and especially the b-dedicated LHCb experiment [46] also perform the study of B mesons at the Large Hadron Collider (LHC), a pp collider currently operating at $\sqrt{s} = 8$ TeV, located at CERN near Geneva, Switzerland.

2.2 The $\Upsilon(4S)$

The $\Upsilon(nS)$ resonances are bound systems of b and $\bar{b}$ quarks having quantum numbers $J^{PC} = 1^{--}$ and can be produced directly in e^+e^- collisions. The cross section for e^+e^- annihilations in the region of the Υ resonances is shown Fig. 6. The decay widths of the three lowest $b\bar{b}$ states, $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$, are relatively small, in the range of a few ten $\mathrm{KeV/c^2}$. The narrow widths of these resonances are an antefact of the Okubo-Zweig-Iizuka (OZI) suppression of hadronic decays.

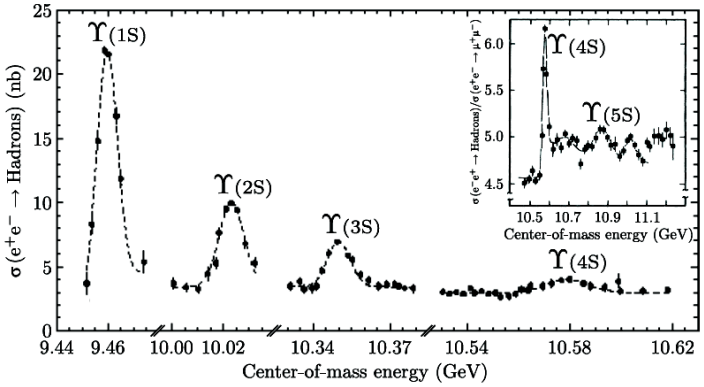


Figure 6: Cross section of e^+e^- annihilations measured by CUSB [47, 48] revealing the family of Υ resonances.

The mass of the $\Upsilon(4S)$ meson is just above the threshold of $B\bar{B}$ production, where B refers generically to a bound state of either $b\bar{u}$ or $b\bar{d}$ quarks. It decays above 96% of the time into such $B\bar{B}$ pairs. The cross section of the $\Upsilon(4S)$ production at its peak

position is about $1/3^{rd}$ of that of total hadronic production $e^+e^- \to q\bar{q}(q = u,d,s,c)$, which is often referred to as "continuum". Therefore, although KEKB is operating at the $\Upsilon(4S)$ resonance in order to produce $B\bar{B}$ pairs, it also produces three times more $q\bar{q}$ events that constitute a major source of background for the study of B-physics. Due to the low invariant mass of the $q\bar{q}$ pair, the jets from these light quarks are produced with a significant boost factor almost back-to-back. In contrast, the B and $\overline{B}$ mesons from the $\Upsilon(4S)$ decay are nearly at rest, and thus the decay products have a spherical topology. We will use this difference later to suppress continuum events with the so-called event shape variables. In order to study the contribution of $q\bar{q}$ events, KEKB is sometimes operated at 60 MeV below the $\Upsilon(4S)$ resonance. The data taken at the $\Upsilon(4S)$ resonance and 60 MeV below are referred to as "on-resonance" and "off-resonance", respectively. Around 10% of data recorded by the Belle detector is the off-resonance data.

2.3 KEKB Accelerator

The KEKB colliding-beam accelerator provides electron-positron collisions at the heart of the Belle detector. Electrons with energy 8.0 GeV and positrons with energy 3.5 GeV are stored in High Energy Ring (HER) and Low Energy Ring (LER), respectively. Since two beams have different energies, separate beam pipes are used, and rings are positioned 11 m underground. The circumference of each ring is 3016 m, having four straight sections. Figure 7 shows an overview of the KEKB collider. The electron and positron beams are created and accelerated to their final energy in the linear accelerator (Linac) and are then injected into the KEKB storage ring.

The beam pipes of KEKB accelerator cross at two points. One is the just crossing point and the other is the interaction point (IP), where the detector of the KEK B-factory (Belle) is located. At the IP, electrons and positrons collide with a finite crossing angle of ± 11 mrad. In order to compensate for the energy loss of the beams due to radiation as they circulate in the KEKB rings, two kinds of RF cavities: normal cavities called ARES and superconducting cavities (SCC) are installed [49].

The most important parameter that demonstrates ability of an accelerator is called the luminosity L, since it is directly connected to the event rate R with the relation $R = \sigma L$, where σ is the Cross section. The luminosity is expressed as:

$$L = \frac{N_{e^+} N_{e^-} f}{4\pi\sigma_x^* \sigma_y^*} R_L, \tag{5}$$

where e^+ and e^- denote positron and electron bunches, respectively, N is the number of particles per bunch, f is the collision frequency, R_L is the geometric reduction

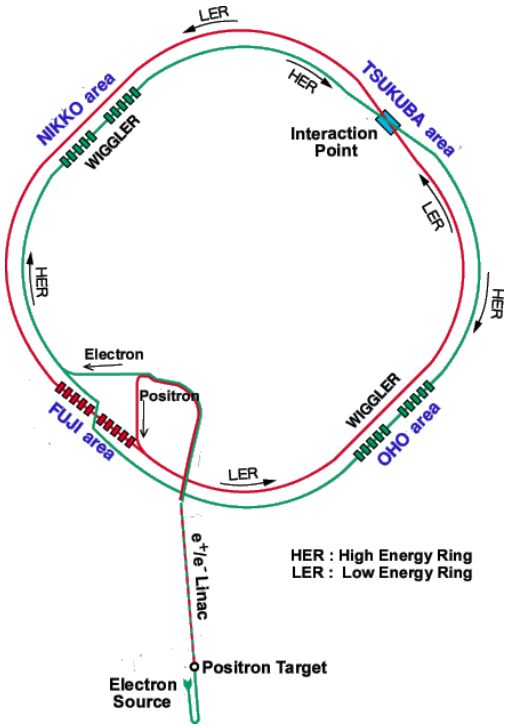


Figure 7: KEKB e^-e^+ collider configuration.

factor, and σ_x^* and σ_y^* are the horizontal and vertical beam size at IP, respectively. Beam-beam tune shift parameters are given by:

$$\xi_{x,y,e^\pm} = \frac{r_e}{2\pi\gamma_{e^\pm}\sigma_{x,y}^*}\frac{N_{e^\pm}.\beta_{x,y}^*}{\sigma_x^*+\sigma_y^*}R_{\xi_{x,y}}, \tag{6}$$

where $\xi_{y(x)}$ is the vertical (horizontal) beam-beam tune shift parameter, r_e is the classical electron radius, β_y^* is the vertical β function at IP, R_ξ the reduction factor for the beam-beam parameter, and γ is the Lorentz boost parameter. Combining equations(5) and (6), the luminosity can be written as:

$$L = \frac{\gamma_{e^\pm}}{2er_e}\left(1+\frac{\sigma_y^*}{\sigma_x^*}\right)\left(\frac{I_{e^\pm}.\xi_{y,e^\pm}}{\beta_y^*}\right)\left(\frac{R_L}{R_{\xi_y}}\right), \tag{7}$$

where e is the elementary electric charge and I is the beam current. The units of L, I, and β_y^* are given in $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, Ampere, and cm respectively. The equation 7 is derived by assuming that σ_x^*, σ_y^*, β_y^* and ξ_y are same for both the beams because unequal parameters cause incomplete overlap of both the beams during collision. This assumption requires $I_+E_+ = I_-E_-$, so the LER current should be higher than the HER current, ξ_y should be made larger, and β_y^* should be made smaller in order to achieve the higher luminosity.

Until 2007 the electron and positron bunches in the KEKB accelerator beams crossed at an angle of 22 mrad. The crossing angle, a unique feature of the KEKB design, provided an effective separation of the beams after collision, avoiding a high background in the detector. Its success was evident in the world-beating luminosities that the collider achieved previously. To boost the luminosity further, however, a scheme was required that would allow an effective head-on collision between the beams while still retaining the crossing angle. To accomplish this goal, KEKB designed and built special superconducting RF cavities called crab cavities [50] that kick each beam sideways in the horizontal plane so that the bunches collide head-on at the interaction point. With the introduction of these crab cavities, in June, 2009, the KEKB collider achieved the world record of the peak luminosity of 2.11×10^{34} cm^{-2}s^{-1} and the total integrated luminosity accumulated by the Belle detector reached 1000 fb^{-1} on November 29, 2009, which was one of the initial goals of the KEKB project. Figure 8 summarizes the integrated luminosity of the two B factories: Belle and BaBar.

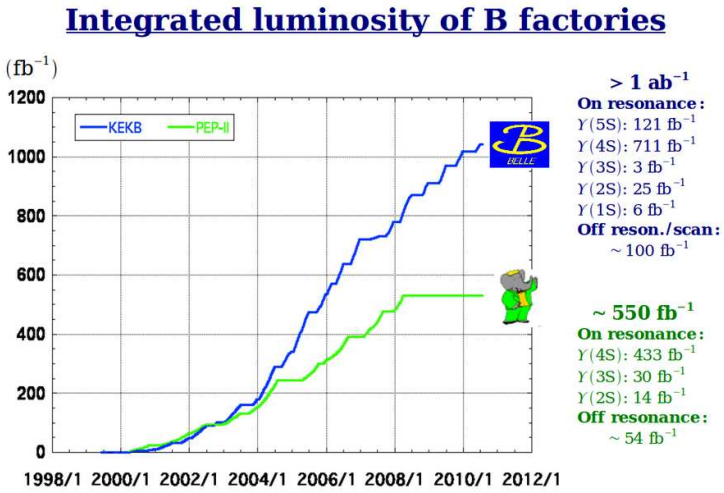


Figure 8: Integrated luminosity summary.

3 Belle Detector

3.1 Introduction

The Belle detector is a multi layered general purpose detector surrounding the interaction point (IP) to detect the particles produced in e^+e^- collisions. The Belle detector is configured within a 1.5 T superconducting solenoidal magnet and iron structure. It is located at the Tsukuba interaction region of the KEKB beams, and consists of following subdetectors - a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL) comprised of CsI(Tl) crystals located inside the superconducting solenoid coil. A multi-levels of resistive plate counters (KLM). Figure 9 shows the side-view of the Belle detector.

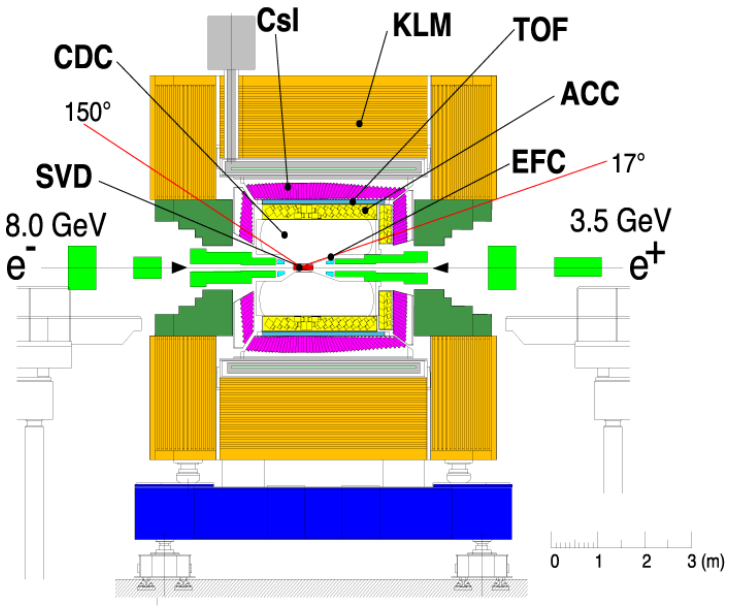


Figure 9: Side-view of the Belle detector.

The SVD is situated just outside of a cylindrical beryllium beam pipe. It measures B meson decay vertices and aids the CDC in providing charged particle tracking. Specific ionization energy loss measurements made with the CDC are combined with light yield readings from the ACC and time of flight information from the TOF to provide charged hadron identification. Electromagnetic shower measurements, crucial for electron and photon detection, are performed by the ECL. An iron flux return located outside the coil (KLM) is instrumented to detect K_L^0 mesons and to identify muons. The Belle detector is described in much detail elsewhere [51].

The coordinate system: The origin is defined as the position of the nominal IP. For reference, the z-axis is defined as the direction of the magnetic field within the solenoid, which is anti-parallel to the positron beam. The x- and y-axes are aligned horizontally and vertically respectively, and correspond to a right-handed coordinate system. The polar angle θ, is subtended from the positive z-axis. The azimuthal angle ϕ, subtended from the positive x-axis, lies in the xy plane. The radius, defined in a cylindrical coordinate system, is measured from the origin in the xy plane, $r = \sqrt{x^2 + y^2}$. The timing of the Belle detector relies on the KEKB Radio Frequency (RF), which governs the bunch crossing rate ($\sim$ 509 MHz, 2 ns interval).

The detector is divided into three regions: the barrel section, which is parallel to the beam axis, and two endcaps, which extend radially from the beam axis at the forward and backward ends of the detector. The polar angle coverage of each of the three sections is listed in Table 2. Details of each subdetector part is described in the following sections.

Table 2: Polar angle coverage of the Belle detector.

Region	Polar angle coverage
Barrel	$34^0 < \theta < 127^0$
Forward endcap	$17^0 < \theta < 34^0$
Backward endcap	$127^0 < \theta < 150^0$

3.2 Beam Pipe

The beam pipe is a double-wall beryllium cylinder enclosing the IP. It is the innermost component of the Belle detector. The role of the beam pipe is to maintain the vacuum and to cool down the heat induced by the beam. The determination of a *B* decay vertex is limited by multiple Coulomb scattering in the beam pipe and the distance from the IP to the first layer of the SVD. Moreover, since the vertex resolution improves inversely with the distance to the first detection layer, the vertex detector has to be placed as close as possible to the IP and thus to the beam pipe wall. Figure 10 shows the cross section of beam pipe at IP. The beam pipe consists of two concentric cylinders with different radii, 20.0 mm and 23.0 mm respectively. The wall thickness is 0.5 mm for both pipes and are made of beryllium (low Z material) to reduce the Coulomb scattering. The gap between the inner and outer walls is filled with helium gas to cool the beam pipe and shield the SVD from beam-induced heating. The outer beryllium cylinder is covered with a 20 μm thick gold sheet in order to reduce low energy X-rays from the high-energy ring. Helium gas is cycled through the gap between the inner and outer walls to provide

cooling.

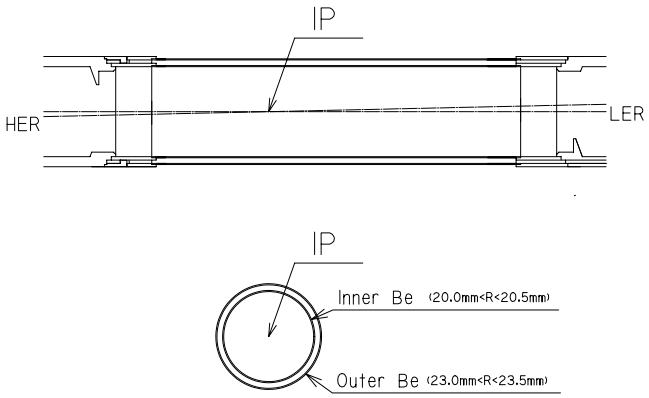


Figure 10: Cross-section of the beryllium beam pipe at the interaction point.

3.3 Silicon Vertex Detector (SVD)

The main goal of the Belle experiment is to study *CP* violation in *B*-decays, which requires a good vertex resolution to effectively measure the difference in *z*-vertex positions for the *B* meson pairs. The SVD is designed to provide the required *z*-resolution for *B*-vertices of 100 μm or better. The SVD also contributes to the reconstruction of charged particles and helps in improving the low momentum resolution of the particle.

Figure 11 shows the side and endcap view of the first version of the vertex detector, SVD1. It consists of three cylindrical detection layers of "ladders", where a ladder is composed of two, three or four double sided silicon strip detectors (DSSD). SVD1 covers a solid angle $23^0 < \theta < 139^0$. As shown in Fig. 11, the three layers of SVD1 consists of 8, 10 and 14 ladders for inner, middle and outer layers, respectively. The radii of each ladder is 3.0 cm, 4.55 cm and 6.05 cm respectively. Each ladder is made of two half-ladders that are electrically independent, and one half-ladder contains one or two DSSDs which are supported by boron-nitride (BN) ribs sandwiched with carbon-fiber reinforced plastic (CFRP). In total, there are 32 ladders and 102 DSSDs. A DSSD is essentially a depleted *pn* junction under reverse bias. A charged particle passing through the junction liberates electrons from the valence band into the conduction band creating electron-hole pairs. These pairs create currents in the p^+ and n^+ strips aligned along and perpendicular to the beam axis respectively. Two-dimensional hit positions are determined from the charge distributions on these orthogonally segmented strips. Each DSSD has 1280 sense strips and 640 readout pads on both side. The size of the

DSSDs is 57.5×33.5 mm^2. Signal from DSSDs are read out by 128 channel VA1 chips [52] placed on both sides of the ladder. The total number of readout channels are 81920.

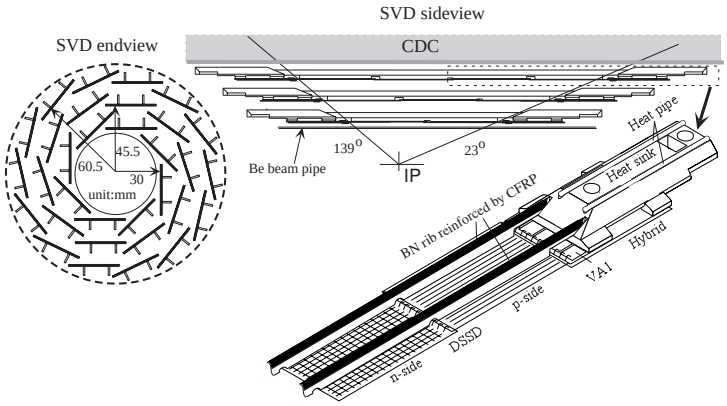


Figure 11: Detector configuration of SVD1.

As mentioned before, the main purpose of the SVD is to determine the B meson decay vertex and to improve the charged particle tracking. For studies of time-dependent CP asymmetries, the z-axis distance of the two B vertices for an $\Upsilon(4S)$ decay must be measured with the precision of about 100 μm. As shown in Fig. 12, the momentum and angular dependences of the impact parameter (closest approach of tracks to the IP) resolution behave like:

$$\sigma_{xy} = 19 \oplus 49(p\beta sin^{3/2}\theta)^{-1}\mu\text{m}, \tag{8}$$

$$\sigma_{x} = 36 \oplus 42(p\beta sin^{5/2}\theta)^{-1}\mu\text{m}, \tag{9}$$

where $\oplus$ indicates a quadratic sum and the momentum p is given in units of GeV/c. The impact parameter resolution for an 1 GeV/c normal track is around 55 μm. The SVD1 readout electronics had limited radiation tolerance and had to be upgraded four times to the latest radiation tolerant readout chips without changing the three-layer mechanical structure. In the summer of 2003, the SVD1 was replaced by a four-layer structure SVD2, covering a range of $17^0 < \theta < 150^0$ for improved tracking, better vertexing and momentum resolution. The innermost layer is closer to the IP at a distance of 2 cm instead of 3 cm in SVD1. This is possible with a significantly smaller beam pipe. The fourth layer is accommodated by a redesign of inner region of the CDC. The four layers of SVD2 contains 6, 12, 18 and 18 full ladders from inside to outside. Each half ladder consists of 1, 2 or 3 DSSDs. Both SVDs used a common double-sided

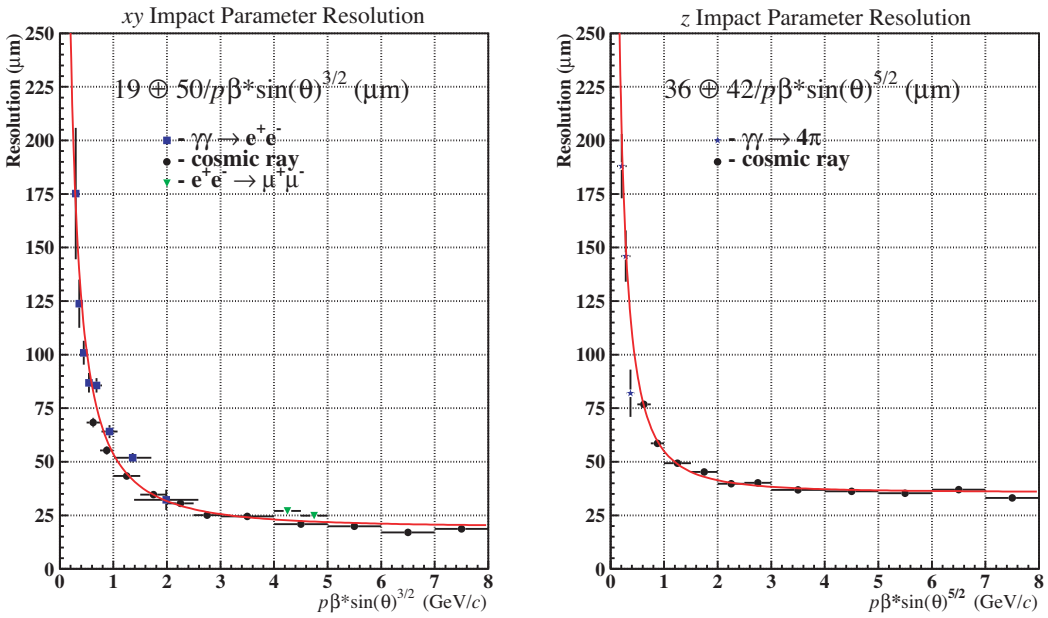


Figure 12: Impact parameter resolution of the SVD

silicon strip detector (DSSD) design. The DSSDs for SVD2 have 512 readout channels in both r - z and r - ϕ, or in total 110,592 readout channels.

3.4 Central Drift Chamber (CDC)

The CDC is a cylindrical wire drift chamber immersed in a 1.5 T magnetic field produced by a solenoidal coil. It is designed for efficient reconstruction of charged particle tracks and precise determination of their momenta. The magnetic field of superconducting solenoid bends the charged particles according to their momenta. The physics goal of the experiment requires a momentum resolution of $\sigma_{p_t}/p_t \sim 0.5\%\sqrt{(1+p_t^2)}$ (where p_t is in GeV/c) for all charged particles with $p_t \geq 100$ MeV/c in the polar angle region $17^0 < \theta < 150^0$. In addition, the CDC is used to measure the energy loss (dE/dx) of charged particles for their identification. The amount of dE/dx depends on $\beta = v/c$ of the charged particle (Bethe-Bloch formula).

The structure of CDC is shown in Fig. 13. It is a cylindrical wire drift chamber having 50 layers (32 axial and 18 small angle stereo layers) of anode wires and three cathode strip layers. The CDC is asymmetric in z-direction, the axial wires are configured parallel to z-axis while the stereo wires are slanted approximately ± 50 mrad. The stereo layers combined with axial layers provide z information of tracks. The cathode strips improve the z-measurement as well as produce a highly efficient fast z-trigger. An anode wire (sense wire) and field wires that surround the anode wire form a drift cell.

The CDC has a total of 8400 drift cells and each drift cell has a maximum drift distance between 8 mm to 10 mm. The sense wires are gold-plated tungsten wires of 30 μm diameter while the field wires are of unplated aluminum of 126 μm diameter. When a charged particle travels through a cell, it ionizes the gas atoms, thus released electrons cause release of more electrons while drifting to the sense wire. They are collected by the sense wires and a hit is recorded by the CDC electronics. Three z-coordinate measurements at the inner-most radii are provided by cathode strips as shown in Fig. 14 The cathode strip having width of 7.4 mm is divided into eight segments in the ϕ direction and has an 8.2 mm pitch in the z-direction. The total number of cathode channels are 1792.

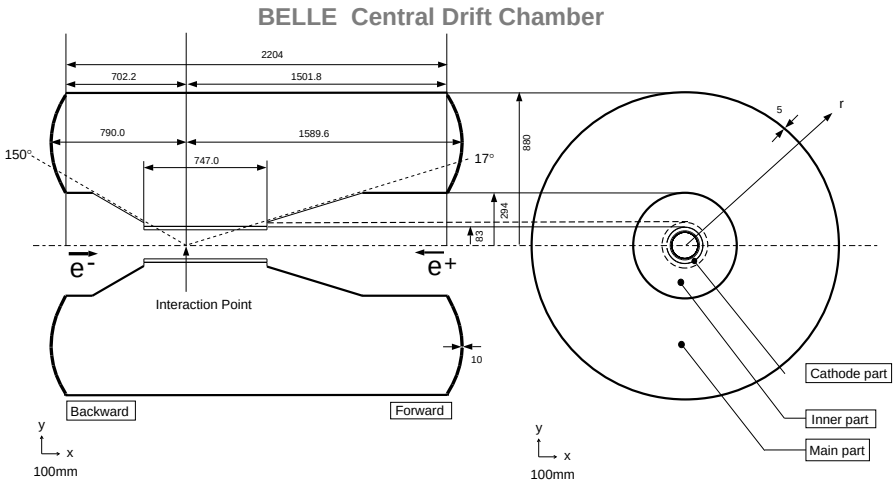


Figure 13: Structure of the CDC. Lengths in Fig. are in the unit of mm

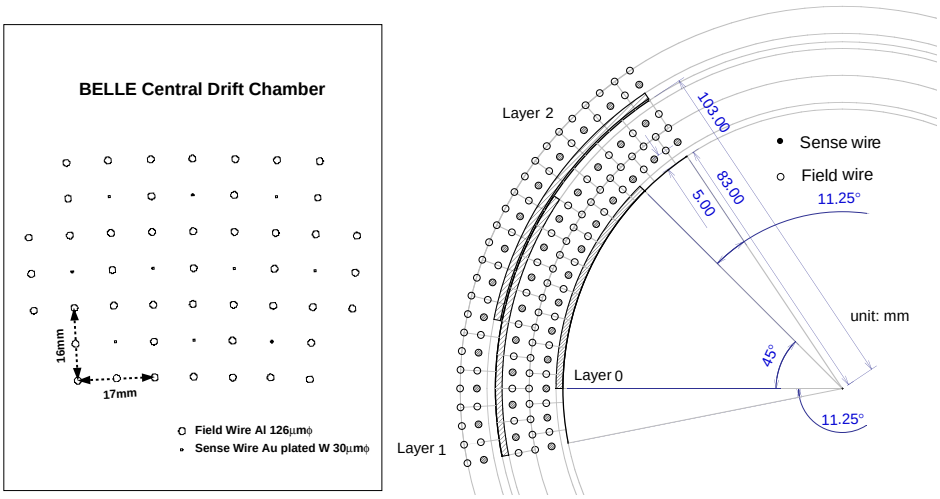


Figure 14: Cell structure of the CDC. Cathode sector configuration is also shown in the right figure.

The path of a charged particle in the constant magnetic field is a helix which is defined with five independent parameters measured by the CDC: the signed curvature of the helix, the slope of the helix, and the 3D coordinates of the helix reference point. The curvature radius is proportional to p_t, the slope is proportional to p_z, and particle charge is identified by the sign of the curvature. The reference point is called a pivot and chosen as the wire position of the innermost hit in the CDC.

To minimize multiple Coulomb scattering contributions to the momentum resolution of the charged particles, we use a gas having low Z. A mixture of 50% helium (He) and 50% ethane (C_2H_6) gas is used in our experiment. This mixture has a long radiation length (640 m) and a drift velocity that saturates at 4 cm$/\mu$s at a relatively low electric field. A good dE/dx resolution is provided by the large C_2H_6 component. The ethane component increases the electron density, which improves the ionization energy loss measurement resolution.

Figure 15 shows the spatial resolution as a function of the drift distance. The spatial resolution is approximately $\sigma_{r\phi} = 130\mu$m. The transverse momentum resolution measured from the cosmic ray data is:

$$\sigma_{p_t}/p_t = \sqrt{(0.20p_t)^2 + (0.29/\beta)^2}\%, \tag{10}$$

where p_t is the transverse momentum measured in GeV/c and β is velocity in units of the speed of light.

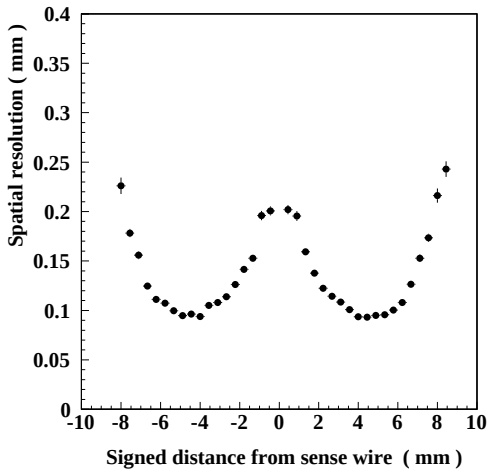


Figure 15: Spatial resolution as a function of the drift distance

The CDC is involved in particle identification for the tracks with p $<$ 0.8 GeV/c and p $>$ 2.0 GeV/c through the measurement of dE/dx. A charged particle's energy loss due to ionization (dE/dx) in the drift cell is determined using the hit amplitude

recorded on the sense wire. Since the energy loss depends on the particle velocity at a given momentum, dE/dx will vary according to the particle mass. Figure 16 shows the scatter plot of the measured dE/dx and the particle momentum (p). The expected relation for π, K, p and e are shown by the solid curves in Fig. 16. The separation between the different particles can be clearly seen.

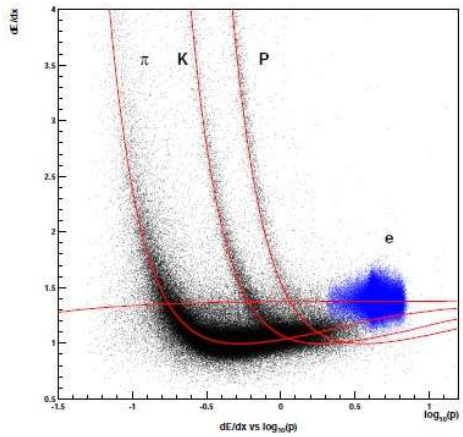


Figure 16: Scatter plot for momentum vs dE/dx. Expected relation for π, K, p and e are shown by the solid curves. The momenta are given in units of GeV/c

3.5 Time of Flight (TOF)

Time-of-flight measurements are performed with scintillating plastic counters with a design time resolution of 100 ps. A very good time resolution enables efficient particle identification for particle momentum below 1.2 GeV/c. This system also provides fast trigger signals. However, the trigger rate of the time-of-flight counter would be too high. Therefore, thin dedicated trigger scintillation counter (TSC) are added to the system (Fig. 17)to produce a fast trigger signal with a manageable rate (below 70 kHz).

The Time of Flight counter (TOF) measures the velocity of charged particles in an intermediate momentum range 0.8 GeV/c to 1.2 GeV/c. The velocity is measured by the particles's time of flight and the flight length. The latter is provided by the CDC's measurements of the track helix parameters. The velocity combined with momentum (as provided by the SVD and CDC) determines the particle's mass and therefore its type.

The TOF works on the principle of scintillation - the property of certain chemical compounds to emit short light pulses after excitation by the passage of charged particles or by photons of high energy. Scintillation is characterized by the light yield. The TOF

measures the time of flight between a particle originating at the IP and passing through the scintillator. The time of flight at p $=$ 1.2 GeV/c is 4.3 ns for $K^{\pm}$ and 4.0 ns for $\pi^{\pm}$.

The TOF system consists of 64 modules concentrically arranged at a radius of 1.2 m, covering a polar angle range of $33^0 < \theta < 121^0$. A module is made up of two trapezoidal shaped time-of-flight counter and Trigger Scintillation Counter (TSC) separated by a radial gap of 1.5 cm, as shown in Fig. 17. A total number of 128 TOF counters are placed in ϕ sectors. The total number of TSC counters are 64. The total number of readout channels are 256 for the TOF and 64 for the TSC.

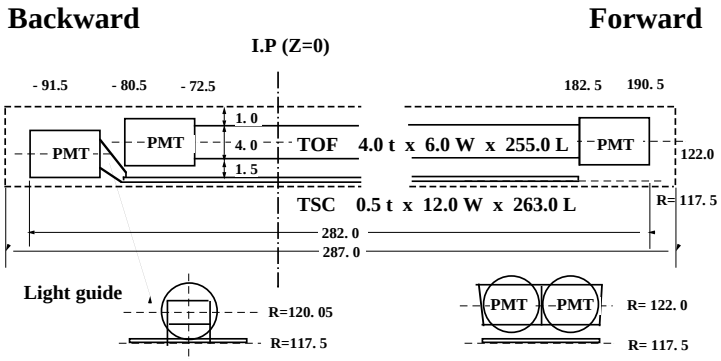


Figure 17: Configuration of a TOF module.

Figure 18 shows the mass distribution for each track in hadron events, calculated by using the equation:

$$m^2 = \left(\frac{1}{\beta^2} - 1\right) p^2 = \left(\left(\frac{cT_{twc}^{obs}}{L_{path}}\right)^2 - 1\right) p^2, \tag{11}$$

where m is mass of the particle, p is the momentum and L_{path} is the path length of the particle determined from CDC track fit assuming the muon mass. Figure 18 shows the clear peaks corresponding to $\pi^{\pm}$, $K^{\pm}$ and protons (p). The data points are in good agreement with MC predictions (histogram) obtained by assuming $\sigma_{TOF} = 100ps$. The TOF hit efficiency is 95% for single-end hits and 88% for both-end hits in $e^+e^- \to \mu^+\mu^-$ pair events.

3.6 Aerogel Cherenkov Counter (ACC)

In the momentum region below 1 GeV/c, the K/π separation is performed by dE/dx measurement from the CDC and the TOF measurements. The ACC extends the momentum coverage for K/π separation up to 3.5 GeV/c. It provides the separation in the momentum range of 1.2 GeV/c $<$ p $<$ 3.5 GeV/c by detecting the Cherenkov

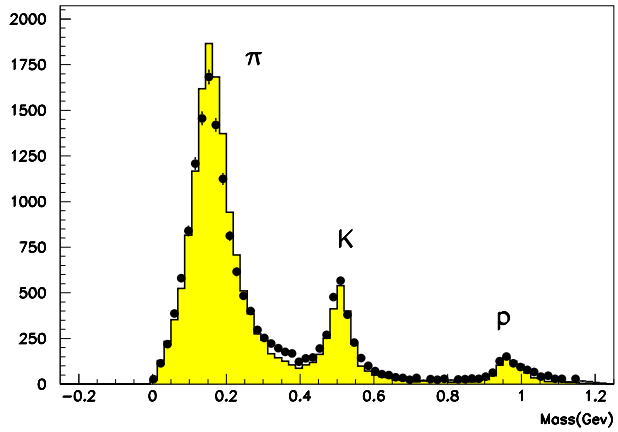


Figure 18: Mass distribution from TOF measurements for particles with momentum below 1.2 GeV/c. Points and histogram show the data and MC distributions, respectively.

light from particles that penetrate through silica aerogel radiator. When a high energy charged particle traverses dielectric media, part of the light emitted by excited atoms appear in the form of a coherent wavefront at a fixed angle with respect to the trajectory - a phenomenon known as Cherenkov effect. Such radiation is produced whenever the velocity of the particle exceeds the speed of light in the medium. The light is emitted if velocity of charged particle β satisfies:

$$\beta = \frac{p}{\sqrt{p^2 + m^2}} > 1/n \tag{12}$$

where m and p are particle mass and momentum respectively, and n is the refractive index. Therefore, there is a momentum region where pions emit the light while kaons do not, depending on the refractive index of the matter. For example, pions with momentum 2 GeV/c emit the light in the matter if $n > 1.002$ while $n > 1.030$ is necessary for kaons with the same momentum.

The ACC is divided into barrel and forward endcap regions. It spans a polar angle region of $17^0 < \theta < 127^0$. The barrel ACC consists of 960 counter modules segmented into 60 cells in the ϕ direction, and 228 modules arranged in 5 concentric layers for the forward endcap part of the detector. All the counters are arranged in a semi-tower geometry, pointing to the IP. The configuration of the ACC is shown in Fig. 19 In order to obtain good K/π separation for the whole kinematic range, the refractive indices of aerogels are selected between 1.01 and 1.03, depending on their polar angle region. A typical single ACC module is shown in Fig. 20 for barrel and endcap ACC respectively.

The Cherenkov light generated in the silica aerogel is fed into one or two fine mesh photomultipliers (FM-PMTs) attached to the aerogel radiator modules which are oper-

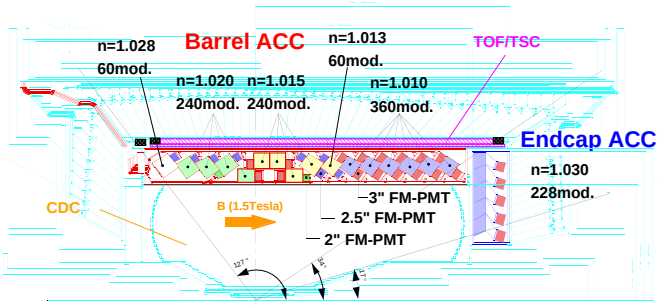


Figure 19: The arrangement of ACC at the central part of the Belle detector. Here n indicates the refractive index.

ated in the 1.5 T magnetic field. The total number of PMTs readout channels are 1560 in barrel ACC and 228 in the endcap ACC.

Figure 21 shows the measured pulse height distribution for the barrel ACC for $e^{\pm}$ tracks in Bhabha scattering events and $K^{\pm}$ candidates in hadronic events, which are selected by TOF and dE/dx measurements. This figure clearly demonstrates a clear separation between high energy-electrons and below-threshold particles. It also indicates good agreement between data and Monte Carlo (MC).

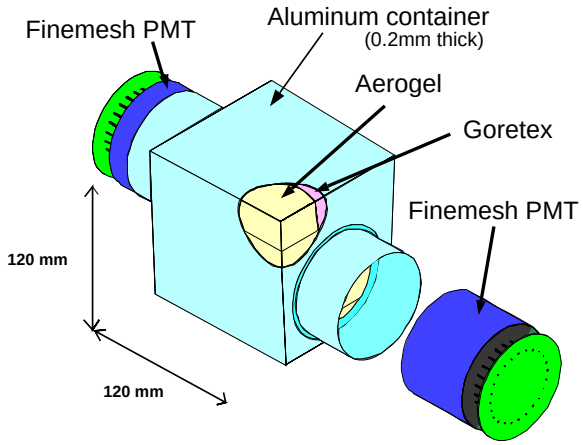
Information from the TOF combined with the ACC and the dE/dx measurements from the CDC give the Belle detector more than 3σ separation between charged kaons and pions over the whole momentum range up to 3.5 GeV/c.

3.7 Electromagnetic Calorimeter (ECL)

When a high-energy electron or photon is incident on a thick absorber, it initiates an electromagnetic cascade as pair production and bremsstrahlung processes generate more electrons and photons with lower energy. The longitudinal development of the electromagnetic shower scales with the radiation length X_0 of the matter, which is defined as the mean distance over which a high-energy electron loses all but $1/e$ of its energy by bremsstrahlung.

The Electromagnetic Calorimeter (ECL) is designed to measure the energy of photons and electrons produced in Belle. Since most of the photons are end products of the cascade decays and have relatively low energies, and thus having good performance below 500 MeV is especially important. Important two-body decay modes such as $B \to K^*\gamma$ and $B \to \pi^0\pi^0$ produce photon energies upto 4 GeV, so high resolution is

a) Barrel ACC Module



b) Endcap ACC Module

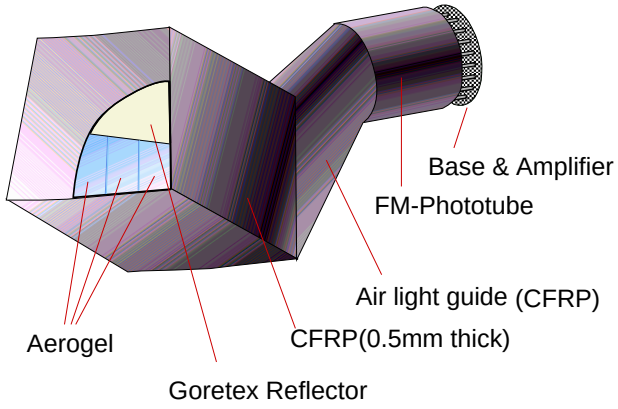


Figure 20: Schematic drawing of typical ACC module. (a) barrel and (b) endcap modules.

needed to reduce background for these modes. Electron identification in Belle relies primarily on a comparison of charged particle momentum and the energy deposited in ECL. Good electromagnetic energy resolution also results in a better hadron rejection. High momentum π^0 detection requires the separation of two nearby photons and a precise determination of their opening angle. This requires a fine-grained segmentation in the calorimeter.

Figure 22 and Table 3 show the ECL configuration. It consists of a highly segmented array of CsI (Tl) crystals with silicon photodiode readout, installed in a magnetic field of 1.5 T inside a superconducting solenoid. The crystal emits photons at a rate proportional to the energy loss, and thus the particle generates a shower of electromagnetic particles. The CsI crystals are doped with Tl in order to shift the wavelength of the photons in the

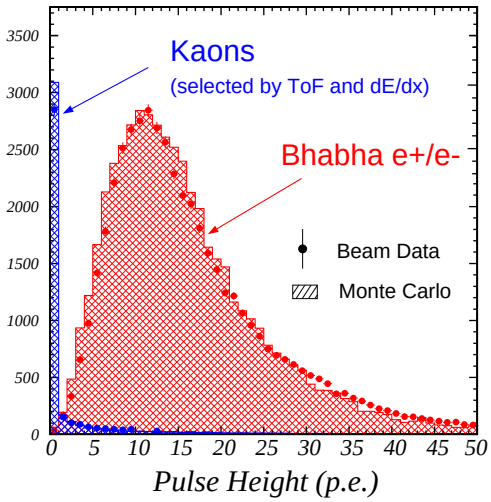


Figure 21: Pulse height spectrum for electrons and kaons in units of photoelectrons (p.e.) observed by the barrel ACC. Kaon candidates are obtained by dE/dx and TOF measurements. The MC expectation are superimposed.

electromagnetic showers into the visible spectrum so that the showers may be detected by the two photodiodes attached to the back of each crystal.

The barrel section of ECL is 3.0 m in length with inner radius of 1.25 m and annular endcaps at z $= +2.0$ m and -1.0 m from the IP. Each crystal has a tower like shape and is arranged such that it points almost to the IP. The calorimeter covers the polar angle region of $17^0 < \theta < 150^0$. Typical size of CsI (Tl) crystals for the barrel region is 5.5cm $\times$ 5.5cm in the front face and 6.5cm $\times$ 6.5cm in the back face. Each CsI (Tl) crystal length corresponds to $16.2X_0$, where X_0 is a radiation length.

Table 3: Geometrical parameters of ECL.

Item	θ coverage	θ seg.	ϕ seg.	No. of crystals
Forward endcap	$12.4^0 - 31.4^0$	13	48 – 144	1152
Barrel	$32.2^0 - 128.7^0$	46	144	6624
Backward endcap	$130.7^0 - 155.1^0$	10	64 – 144	960

Each CsI (Tl) crystal is read out by an independent pair of silicon PIN photodiodes and charge-sensitive pre-amplifiers attached at the end of the crystal. The total number of readout channels of ECL are 17472.

The energy resolution of ECL measured as a function of incident photon energy with 3×3 ECL matrices is given by:

$$\frac{\sigma_E}{E} = \frac{0.0066}{E} \oplus \frac{1.53}{E^{1/4}} \oplus 1.18(\%) \tag{13}$$

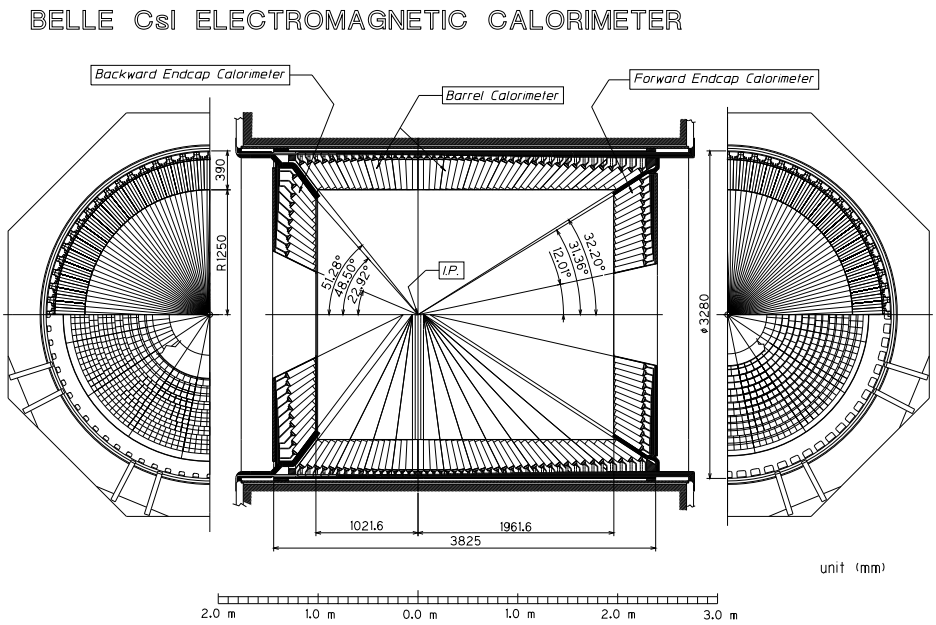


Figure 22: Configuration of the ECL.

where E is in GeV. Here, the first term is due to the contribution from electronic noise, and second and a part of the third term comes from the shower leakage fluctuations. The third term also incorporates systematic effects such as the uncertainty of the calibration on crystals. The spatial resolution measured by the photon beams is given by:

$$\sigma_x(mm) = 0.27 + \frac{3.4}{E^{1/2}} + \frac{1.8}{E^{1/4}} \tag{14}$$

where E is measured in the units of GeV. Two-photon invariant mass distributions in hadronic events for π^0 and η are shown in Fig. 23 The clear peaks of π^0 and η are seen at each nominal mass and the energy resolution has been achieved to be 4.8 MeV for π^0 and less than 12.1 MeV for η.

In addition to the measurement of energy of photons and electrons, the ECL plays an important role in the electron identification. The electron identification is performed combining the following information:

- Matching between the position of the charged track measured by the CDC and that of the energy cluster measured by the ECL,
- E/p, the ratio of energy measured by the ECL to momentum measured by the CDC,
- E_9/E_{25} at the ECL, the ratio of ECL shower energy in an array of 3×3 crystals to the energy in an array of 5×5 crystals,

- dE/dx in CDC,
- Light yield in the ACC,

The probability density functions (PDFs) for above parameters are made and then a likelihood ratio for every track is calculated.

The efficiency of electron identification is greater than 90% with a hadron fake rate (the probability to misidentify hadron as electron) of $\sim 0.3\%$ for a track with $p > 1\,\mathrm{GeV/c}$. The ECL also provides the trigger information and online luminosity information.

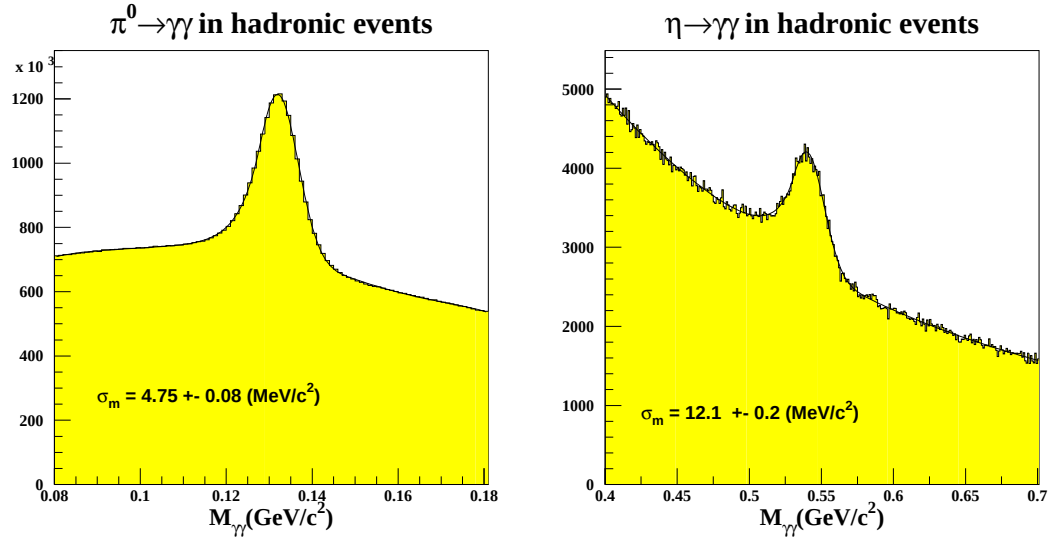


Figure 23: Two photon invariant mass distribution for hadronic events: $\pi^0 \to \gamma\gamma$ (left) and $\eta \to \gamma\gamma$ (right).

3.8 K_L and Muon Detector (KLM)

The K_L and muon detector (KLM) was designed for the detection of K_L mesons and muons with high efficiency and low fake rate over a broad momentum range above 600 MeV/c. Since muons have a relatively small interaction cross section, they penetrate further through the Belle detector than most of the particles coming from the IP. Any track that penetrates several layers of the KLM after leaving a track in the CDC is almost certainly a muon. In addition, the neutral K_L^0 meson, which does not interact with any of the subdetectors closer to the IP, is identified when it is stopped by the 3.9 nuclear interaction lengths of iron contained in the KLM. Hits in the KLM without an associated charged track in the CDC are assumed to be from K_L^0 mesons.

The detector is placed outside the solenoid magnetic field. The KLM contains 15 detector layers and 14 iron plates in the barrel region, and 14 detector layers in each of the forward and backward end-caps. A detector layer is a super-layer of two resistive-plate counters (RPC) surrounded by θ and ϕ cathode strips. Each RPC has two parallel-plate electrodes separated by a gas-filled gap as shown in Fig. 24. An ionizing particle passing through the gap initiates a streamer in the gas that results in a local discharge of the plates. The discharge induces a signal on the external cathode strips. The KLM covers the polar angle region of $20^0 < \theta < 155^0$. The barrel shaped region around the IP covers an angular range of $45^0 < \theta < 125^0$ and endcaps in the forward and backward directions extend this range to $20^0 < \theta < 155^0$.

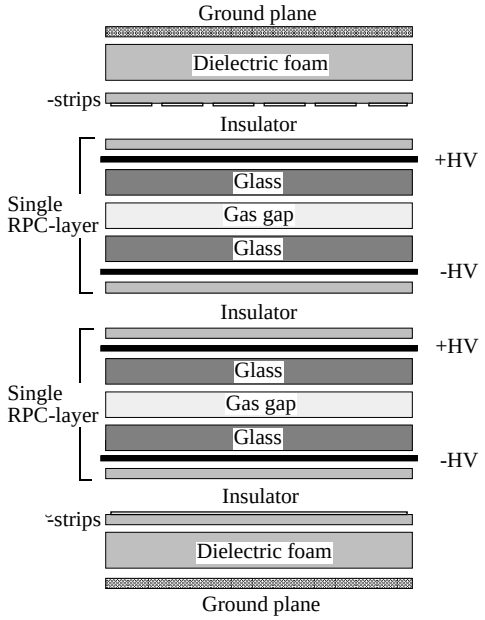


Figure 24: Configuration of the KLM.

3.9 Extreme Forward Calorimeter (EFC)

The Extreme Forward Calorimeter (EFC) extends the range of electron and photon calorimetry to the extreme forward $6.4^0 < \theta < 11.5^0$ and backward regions $163.3^0 < \theta < 171.2^0$ to detect electrons and photons very close to the beam pipe. The EFC is attached to front faces of cryostats of the KEKB accelerator compensation solenoid magnets surrounding the beam pipe. The EFC also acts as a beam mask to reduce the backgrounds for CDC. In addition, the EFC is used as a beam monitor in the KEKB accelerator and a luminosity monitor for the Belle experiment. A three dimensional view of the crystal

arrangement is shown in Fig. 25. Since the EFC is placed in the very high radiation region around the beam pipe near the IP, it is required to be radiation hard. So, the $Bi_4Ge_3O_{12}$ (BGO) crystal has been adopted which has the property of radiation hardness at a Mrad level and has an excellent e/γ energy resolution of $(0.3-1.0)\%/\mathrm{E\ GeV}$. Both forward and backward EFC consist of BGO crystals segmented into 5 regions in θ direction and 32 regions in the ϕ direction in order to provide better position resolution. Typical cross-section of a crystal is about 2×2 cm^2 with $12X_0$ for the forward and $10.5X_0$ in backward region, where X_0 is the radiation length.

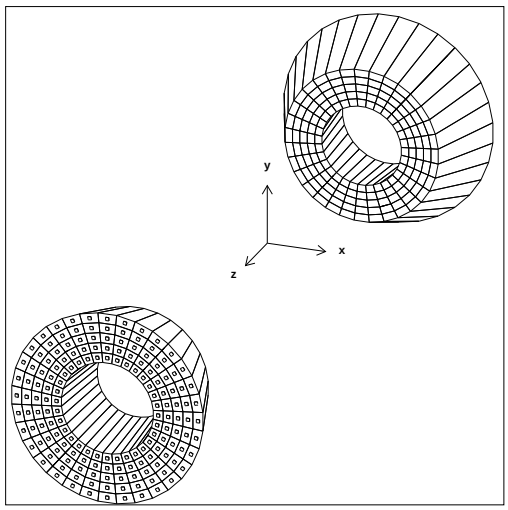


Figure 25: Configuration of the EFC

3.10 Detector Solenoid and Iron Structure

The superconducting solenoid provides a 1.5 T magnetic field parallel to the beam pipe for charged particle momentum measurement. The coil is wound round the inner surface of as an aluminum support cylinder of 3.4 m diameter and 4.4 m length. Indirect cooling is provided by a liquid helium circulation through a tube on the inner surface of the aluminum cylinder. The coil is surrounded by an iron structure, which serves as the return path of the magnetic flux and an absorber material for the KLM. The overall structure of the cryostat and the schematic drawing of the coil cross section are shown in Fig. 26.

3.11 Trigger and Data Acquisition System

The role of the trigger is to decide when the various subsystems of the Belle detector should record an event. Once a particular particle collision satisfies the trigger criteria,

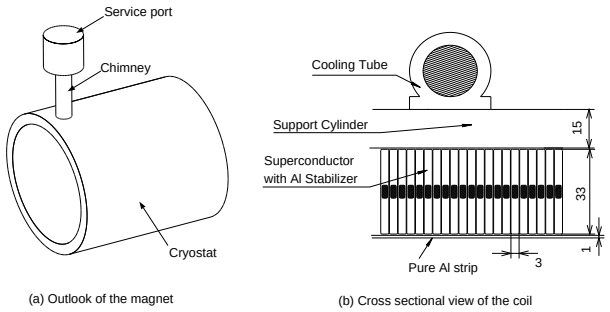


Figure 26: Configuration of the solenoid magnet.

data from all subsystems are read out and stored for further use. The word "event" is used interchangeably to represent the physical particle collision or the stored data representing the collision. The decision to read out is based on criteria carefully chosen to remove background events while retaining events of interest at a high efficiency. Once an event is triggered, the data acquisition (DAQ) system transfers the raw data from the detector to the data storage system. The main sources of background events are undesirable interactions between the electron-positron beams, collisions of a beam particle with residual gas molecule or with the beam-pipe, and synchrotron radiation from the beams. The total cross-sections and trigger rates at the design luminosity 10^{34} cm^{-2}s^{-1} for various physical processes of interest are listed in Table 4. Events of interest are primarily hadronic ($e^+e^- \to q\bar{q}$ or $e^+e^- \to \Upsilon(4S) \to B\bar{B}$) and QED events ($e^+e^- \to e^+e^-$ or $e^+e^- \to \mu^+\mu^-$ or $e^+e^- \to \tau^+\tau^-$), used for physics analyses as well as for detector calibration and luminosity measurements. For this analyses only hadronic events are of interest and thus the following discussion focuses on the hadronic trigger. Event rates for both physics and background events at the design luminosity are about 100 Hz each; but to accommodate higher backgrounds, the trigger is designed to operate up to 500 Hz.

The trigger system is composed of the Level-1 (L1) hardware trigger and the Level-3 (L3) and Level-4 (L4) software triggers. An overview of the system is shown in Fig. 27

The L1 trigger system consists of sub-detector triggers and a central trigger system called the Global Decision Logic (GDL) [53]. The sub-detectors provide trigger signals based on the track and energy information to the GDL after processing event signals in parallel. The GDL uses the information to characterize the event type within 0.35 μs Figure 28 shows the signal flow in the L1 trigger system.

The DAQ collects all sub-detector information for the event which passes the L1 trigger. An event builder, within DAQ system, converts parallel data from the sub-detectors to event-by-event data. The event builder output is transferred to an online

Table 4: Total cross-section and trigger rates with $L = 10^{34}/cm^2/s$ from various physics processes at $\Upsilon(4S)$

Physics process	Cross section (nb)	Rate (Hz)
$\Upsilon(4S) \to B\bar{B}$	1.2	12
Hadron production from continuum	2.8	28
$\mu^+\mu^- + \tau^+\tau^-$	1.6	16
Bhabha ($\theta_{lab} \geq 17^0$, *prescaled by* 100)	44	4.4
$\gamma\gamma(\theta_{lab} \geq 17^0$, *prescaled by* 100)	2.4	0.24
2γ *processes* ($\theta_{lab} \geq 17^0, p_t \geq 0.1$ GeV/c)	~ 15	~ 35
Total	~ 67	~ 96

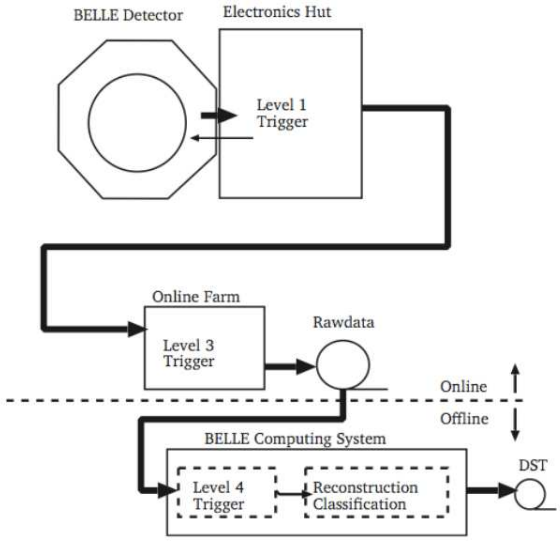


Figure 27: The Belle trigger system.

computer farm to be screened by the L3 trigger. The L3 trigger system is a software trigger that uses an ultra-fast track finder. By requiring at least one track with an impact parameter along the z axis less than 5.0 cm and total energy deposited in the ECL to be less than 3 GeV, the trigger reduces the event rate by $50 \sim 60\%$.

The L4 trigger is applied just before converting the raw data to fully reconstructed data. Events with at least one track having $p_T > 300$ GeV/c and impact parameters less than 1 cm in $r\phi$ plane and 4 cm along the z axis are stored in the DST files. The data size is further reduced for specific physics processes, and the skimmed data are stored Mini-DST (MDST) files.

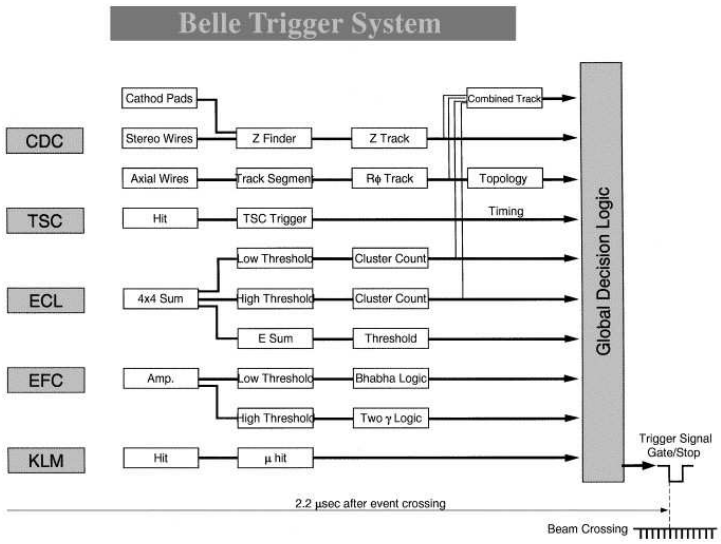


Figure 28: The Level-1 trigger system for the Belle detector.

3.12 Detector Simulation

To generate the events based on detector-independent physical processes involved in the decay we use EvtGen [54], an event generator specially developed for decays related to *B*-meson, likely to be occurring at a B-factory. In case of existence of different models for the same decay, the choice is made according to the need by the user and can be supplied as an input to the decay file, where evolution of a process is specified.

To include the detector performance effects, the EvtGen generated events are passed through a GEANT based MC simulator gsim [55]. In order to achieve near-to-the-real experimental conditions, during each experimental run, for collecting the real data, the detector configuration and performance is monitored and saved into data-files, which can later be fed to gsim to mimic run-dependent conditions while generating the MC samples. This reduces the risk of generating inconsistencies between the more idealistic MC samples and the realistic data sample. Generating run-dependent MC samples allows one to model the real data with high accuracies. A small disagreement between the real data and the MC sample has to be taken care of by either correcting with the help of some efficient control sample or by adding this effect into overall systematic uncertainty, whenever applicable.

Table 5: List of background MC samples.

Background MC	Size (times data)	Description
$q\bar{q}$	2	$e^+e^- \to q\bar{q}$ (q=u,d,s,c) events underneath the $\Upsilon(4S)$
Generic $B\bar{B}$	5	both B's decaying generically; doesn't include rare decays
Rare $B\bar{B}$	50	rare B decays

4 Event selection

4.1 Data and Monte Carlo samples

We generate an experiment- and run-number dependent signal MC sample. The total number of generated events is 2×10^6. In this analysis, we assume B^0 or $\overline{B^0}$ decays to the $K^+K^-\pi^0$ final state according to the decay model "PHOTOS PHSP" [54]. PHSP explains a generic phase space to N-bodies with all spins of particles in the initial and final states are averaged, while PHOTOS is used to include final state radiation in the decay. Table 5 lists various background MC samples used in the analysis.

4.2 Event selection

We reconstruct the B meson candidate from its decay particles (daughters and grand daughters). The particles used to reconstruct B meson in the analysis undertaken are two oppositely charged kaons and a pair of photons arising from the π^0 decay.

4.2.1 Charged track selection

Each track candidate must have a minimum transverse momentum (p_T) of 100 MeV/c, and a distance of closest approach with respect to the interaction point (IP) of less than 0.2 cm in the transverse r–ϕ plane (dr) and less than 5.0 cm along the z axis (dz), where the z axis is defined by the direction opposite to the e^+ beam. These criteria correspond to a $\pm 6.5\sigma$ window around the nominal IP (0,0,0).

4.2.2 Kaon identification

Charged kaon identification relies on the information from ACC, CDC and TOF detectors. In particular, it is based on a likelihood ratio $L_{K/\pi} = \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi}$, where $\mathcal{L}_K$ and $\mathcal{L}_\pi$ denote the individual likelihoods for kaons and pions, respectively, calculated using specific ionization in the CDC, time-of-flight information from the TOF, and the number of photoelectrons from the ACC [see Sec. 3.6]. A requirement $L_{K/\pi} > 0.6$ is applied to select the two kaon candidates. Figure 30 (left) shows the distribution of the

kaon-pion likelihood. The kaon identification efficiency is approximately 86% while the probability of misidentifying a pion as a kaon is 11%.

4.2.3 Photon selection

Photons are required to have an energy greater than 60 (100) MeV in the barrel (endcap) ECL. Figure 31 shows the photon energy distributions for the barrel and endcap regions. Photons detected in the barrel region have a better resolution than the one found in the endcap. Photons that are in the forward and backward regions are more likely to be the result of particles scattered by the beam pipe, and hence a tighter requirement to remove these backgrounds is imposed.

4.2.4 π^0 reconstruction

We reconstruct π^0 candidates from the photon pairs that have an invariant mass between 112 and 156 MeV/c^2 [see Fig. 30 (right)], corresponding to $\pm 3.5\sigma$ around the nominal π^0 mass [56]. In addition, a requirement on the π^0 decay helicity angle, $|\cos\theta_{\rm hel}| < 0.95$ [see Fig. 32 (right)], is imposed, where this requirement is based on the fact that correctly reconstructed π^0 candidates tend to have smaller values of $|\cos\theta_{\rm hel}|$ compared to misreconstructed events, where $\theta_{\rm hel}$ is the angle between one of the daughter photons and the B momentum in the π^0 rest frame. For the π^0 candidates, which satisfy all the selection criteria, the daughter photons are fitted to accommodate a kinematical constraint due to the precise π^0 mass, using the Kalman-fitting techniques [57]. To gain further improvements in purity, we impose a requirement on the π^0 mass-constrained fit chi-square, $\chi^2_{\rm mass} < 50$ [see Fig. 32 (left)]. This selection is pretty wide, and eliminates only the pion candidates that have masses that vary greatly from the invariant mass of the photon pair from which they have been created.

4.2.5 B meson reconstruction

B meson candidates are identified using two kinematic variables: beam-energy constrained mass, $M_{\rm bc} = \sqrt{E^2_{\rm beam} - |\sum_i \vec{p}_i|^2}$, and energy difference, $\Delta E = \sum_i E_i - E_{\rm beam}$, where $E_{\rm beam}$ is the beam energy, and $\vec{p}_i$ and E_i are the momentum and energy, respectively, of the i-th daughter of the reconstructed B in the center-of-mass (CM) frame. The advantage of using these variables lies in the following facts:(1) $E_{\rm beam}$ is known to a very good accuracy; it has a resolution of about 3 MeV. The B energy, $\sum_i E_i$, on the other hand, is derived from track momenta of the final state particles, which have resolutions of about 1-5 MeV/c each. As a result, the B energy resolution usually ranges from 10 MeV to 40 MeV, depending upon kinematical details of the decay being re-

constructed (2) The energy of the track is an unknown, unmeasured quantity and is rather determined from the track momentum after the mass-assignment. Therefore, the energy resolution of each track is dominated by the track momentum uncertainty. ΔE as defined above, has the dominant contribution from the track momenta and has no effect of $E_{\rm beam}$ measurements. As a result, ΔE is highly sensitive to the overall momentum of the B^0 candidate. We retain events with $5.271\,\mathrm{GeV/c^2} < M_{\rm bc} < 5.287\,\mathrm{GeV/c^2}$ and $-0.30\,\mathrm{GeV} < \Delta E < 0.15\,\mathrm{GeV}$ for further analysis. The $M_{\rm bc}$ requirement corresponds to approximately $\pm 3\sigma$ around the nominal B^0 mass [56], and we apply a looser $(-12\sigma, +6\sigma)$ requirement on ΔE because it is used in the fitter.

In the following, we summarize various requirements applied to select $B^0 \to K^+K^-\pi^0$ decay candidates.

- Track $p_T > 100$ MeV/c
- $|\mathrm{dr}| < 0.2$ cm and $|\mathrm{dz}| < 5.0$ cm
- Kaon likelihood against pion, $L(K/\pi) > 0.6$
- Energy of photons $> 60\ (100)$ MeV in the barrel (endcap) ECL
- Reconstructed π^0 mass is in the range $[112, 156]\,\mathrm{MeV/c^2}$
- π^0 mass-constrained $\chi^2 < 50$
- The decay helicity angle of π^0, $|\cos(\theta_{\rm hel})| < 0.95$
- $5.271\,\mathrm{GeV/c^2} < M_{\rm bc} < 5.287\,\mathrm{GeV/c^2}$
- $-0.30\,\mathrm{GeV} < \Delta E < 0.15\,\mathrm{GeV}$

Figures 29 to 32 show some of the selection variables with the cut position indicated by an arrow.

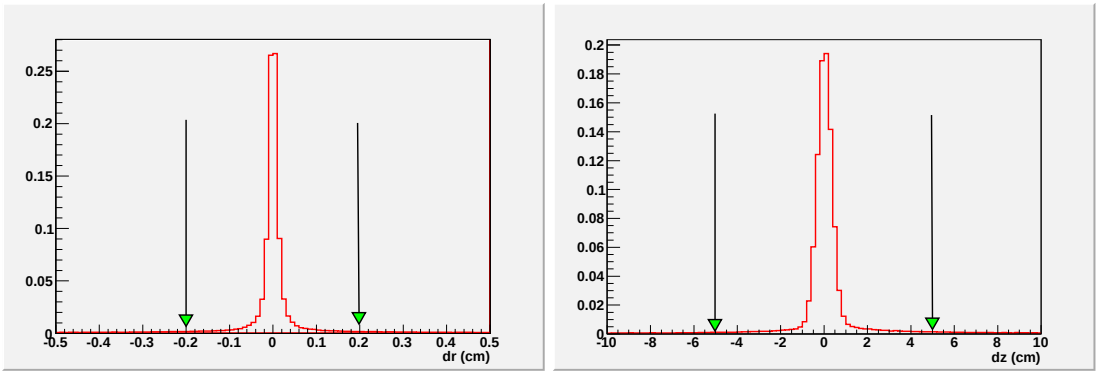


Figure 29: Distributions of distance of closest approach with respect to the IP for track candidates: dr (with $\sigma \sim 0.3$mm) [left] and dz (with $\sigma \sim 7.7$mm) [right].

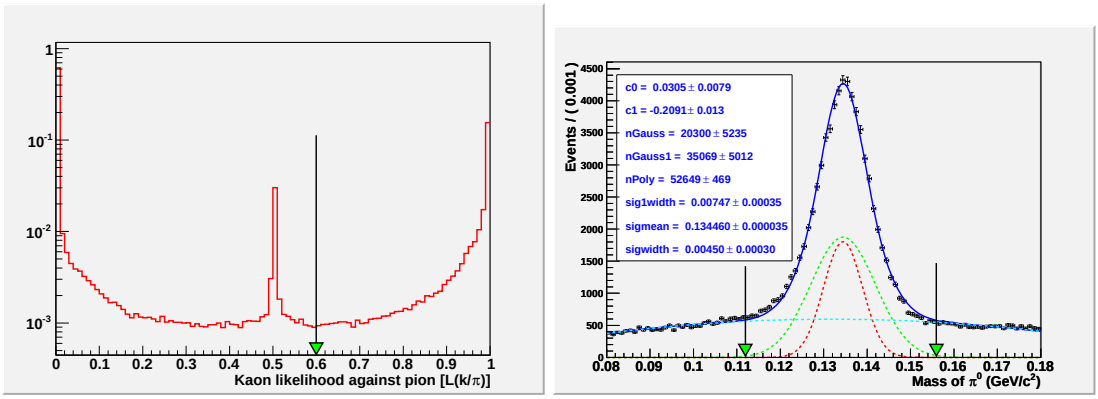


Figure 30: Distributions of kaon-pion likelihood (left) and reconstructed π^0 invariant mass (right).

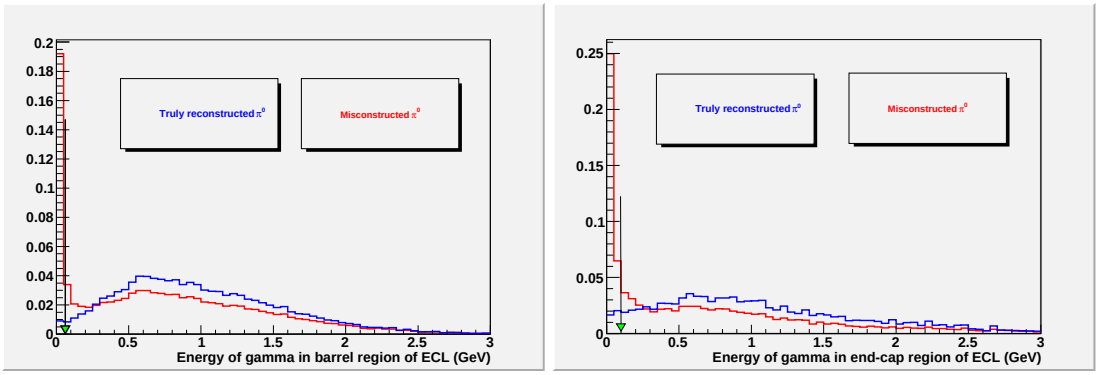


Figure 31: Energy of photons in the barrel (left) and endcap (right) ECL.

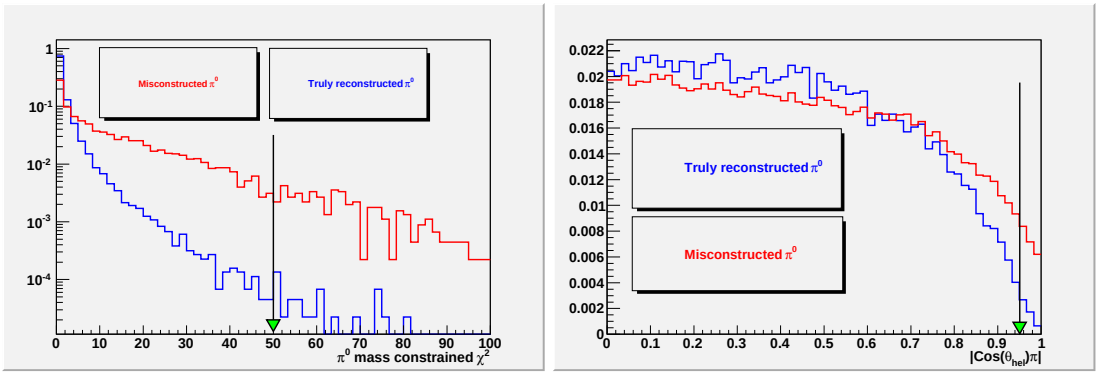


Figure 32: Mass-constrained χ^2 (left) and cosine of the helicity angle (right) of the π^0 candidate.

4.3 Best candidate selection

After applying the selection criteria mentioned in Sec 4.2, we find that some events have more than one B candidate. Figure 33 shows the so-called B candidate multiplicity. The average number of B candidates found per event is 1.3 in signal MC and 1.4 in real data samples. In events with multiple B candidates we choose the one(s) with the π^0 that has the lowest χ^2_{mass} value. If there are still more than one candidate having the same χ^2_{mass}, the candidate with the lowest B^0 vertex fit chi-square is chosen. With our best candidate selection, the number of times the selected B candidate corresponds to a true B candidate is 94%.

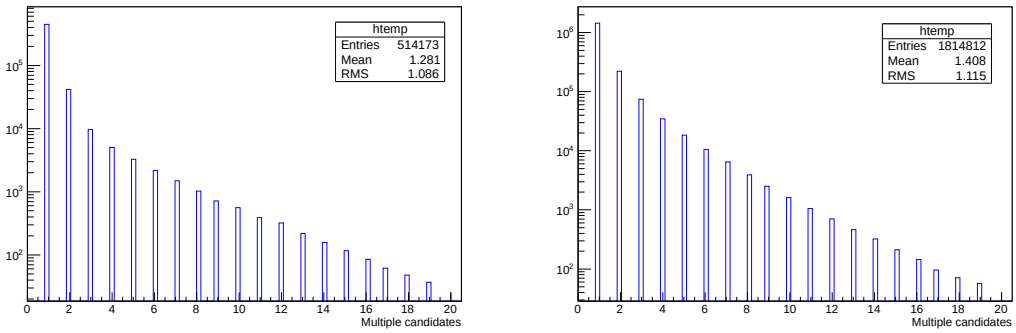


Figure 33: B candidate multiplicity in the signal MC (left) and real data (right) samples.

4.4 Efficiency and self-crossfeed

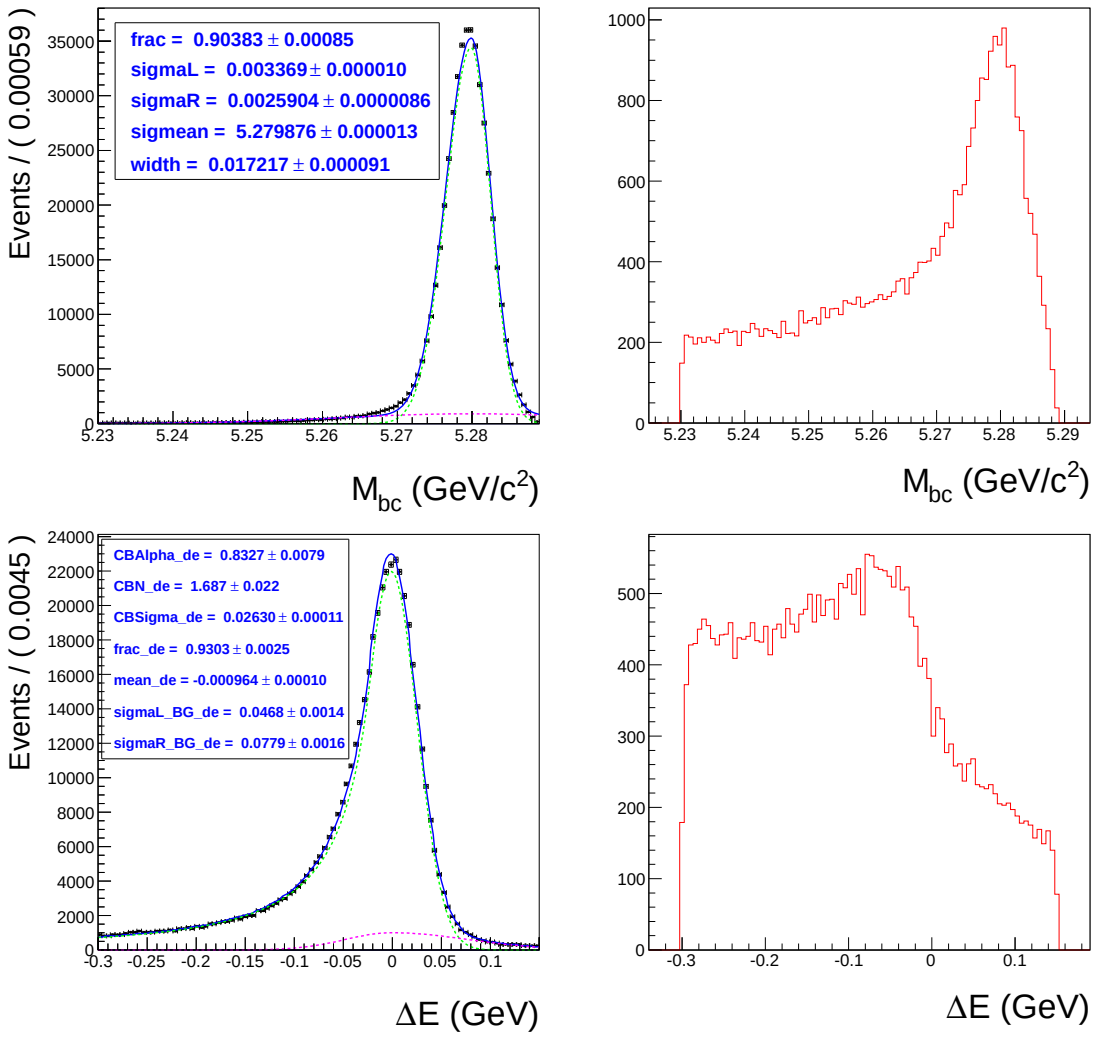


Figure 34: $M_{\rm bc}$ (top) and ΔE (bottom) distributions for the correctly reconstructed (left) and misreconstructed (right) signal events.

To identify correctly reconstructed signal candidates, we follow a truth-matching procedure where we perform a matching of reconstructed final state particles to generator-level final state particles. When a match is established, we call the mother of these final state particles to be correctly reconstructed. In Fig. 34 we show the distributions of the kinematic variables, $M_{\rm bc}$ and ΔE, for correctly reconstructed and mis-reconstructed signal events. For correctly reconstructed signal events: $M_{\rm bc}$ resolution is 2.98 MeV/c^2 and ΔE resolution is 26.3 MeV.

We calculate efficiency of correctly reconstructed (CR) signal events = $\frac{S}{G}$ = 24.08 ± 0.03% and fraction of self-crossfeed signal events (SCF) = $\frac{B}{S+B}$ = 6.37 ± 0.04%, where S (B) is the number of correctly reconstructed (misreconstructed) signal events and G is

the number of generated signal events. Once we apply a mild cut on the $q\bar{q}$ suppression variable [Section 5] and the charm veto [Section 6.1], the above two quantities reduce to $20.80 \pm 0.03\%$ and $5.82 \pm 0.04\%$. Finally after the signal-region requirement on $M_{\rm bc}$, we obtain the efficiency for CR signal events as $19.73 \pm 0.03\%$ and the SCF fraction as $3.18 \pm 0.01\%$.

5 Continuum suppression

The dominant background arises from the $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) continuum process. To suppress this background, observables based on the event topology are utilized. The event shape in the CM frame is expected to be spherical for $B\bar{B}$ events and jet-like for continuum events. We employ a neural network [58] to combine the following six input variables:

- LR[RooKSFW]: the likelihood ratio formed out of 16 modified Fox-Wolfram moments [59]. Figure 35 shows distributions of the missing mass squared (imm) that has 7 bins, the KSFW moments in each of the imm bins, followed by
- cosb: cosine of the angle (θ) between the B momentum and z axis. The B decays follow a $1-\cos^2\theta$ distribution, while the $q\bar{q}$ background is nearly flat in $\cos\theta$,
- costhr: the cosine of the angle between the thrust axis of the B candidate and that of the rest of the event, where the thrust axis is oriented in such a way that the sum of momentum projections along that direction is maximized,
- cosbt: cosine of the angle between the B thrust and z axis,
- R2: ratio of 2^{nd} and 0^{th} order Fox-Wolfram moments [59]. It is a measure of the jettiness of an event. For a more jet-like event the distribution shifts from 0 to 1,
- Δz: the vertex separation along the z axis between the B candidate and the remaining tracks. For the signal event, the absolute value tends to be larger because of the longer lifetime of the B meson.

Figures 36 to 38 show distributions of various continuum fighting variables for signal and background. All of them exhibit a clear separation between the two classes of events. The training and optimization of the neural network are accomplished with signal and $q\bar{q}$ MC simulated events. In Fig. 39 we show the resultant neural network output (C_{NB}) distribution obtained with the signal and continuum MC events. We require the neural network output (C_{NB}) to be greater than 0.2 to substantially reduce the continuum background. The relative signal loss due to this requirement is approximately 12%, whereas the achieved continuum suppression is close to 92%. The remainder of the C_{NB} distribution strongly peaks near 1.0 for signal (see the left plot in Fig. 39), and thus we find difficulty in modeling it with an analytic function. We instead translate it to $C'_{NB} = \log[(C_{NB} - 0.2)/(1.0 - C_{NB})]$. This translation is useful as the resultant distributions become smoother (see the right plot in Fig. 39) and hence can be modeled with some analytical function.

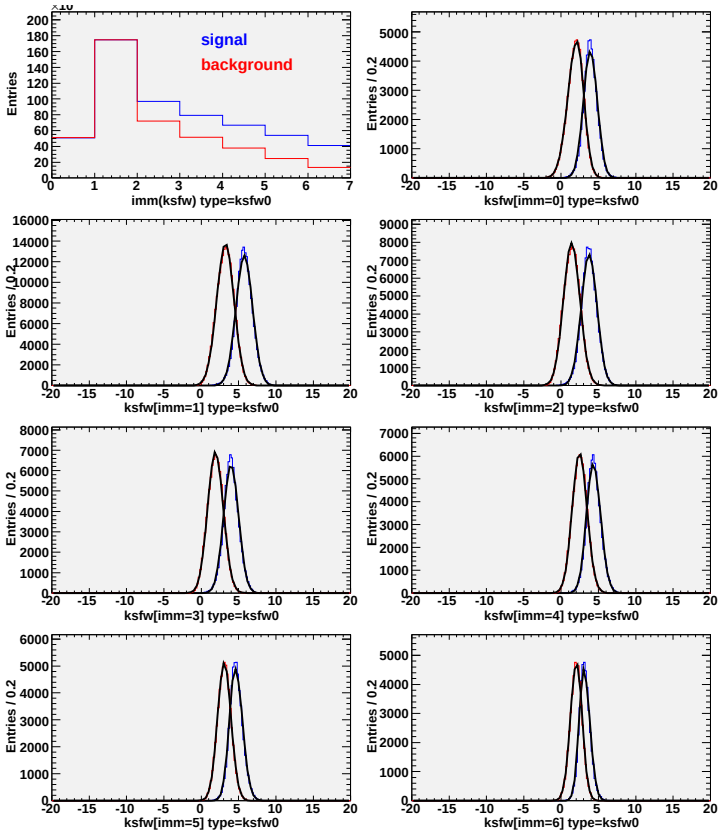


Figure 35: Distributions of imm (top left) and KSFW moments for each of the seven imm bins. Blue (red) histograms represent signal ($q\bar{q}$) MC events.

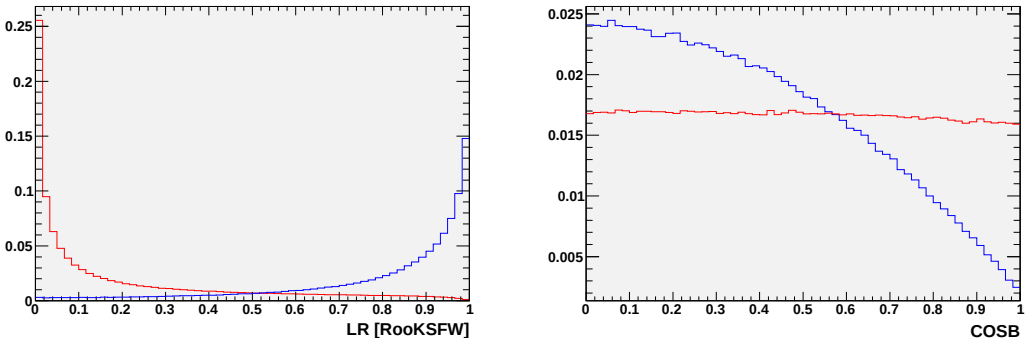


Figure 36: Distributions of LR (left) and cosb (right) for signal and continuum background.

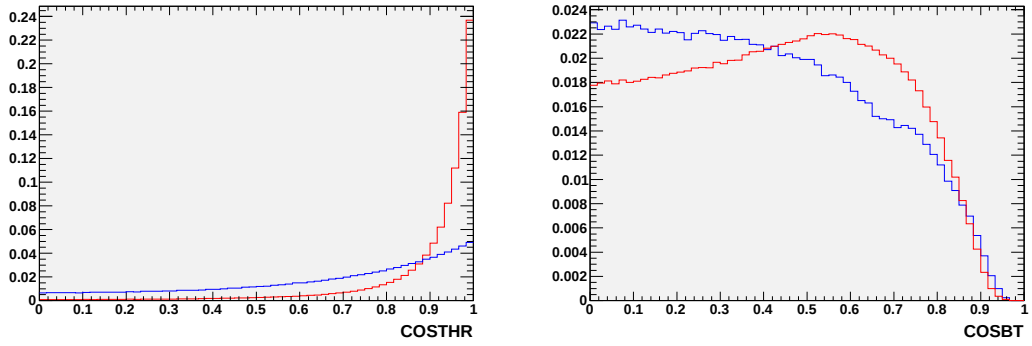


Figure 37: Distributions of costhr (left) and cosbt (right) for signal and continuum background.

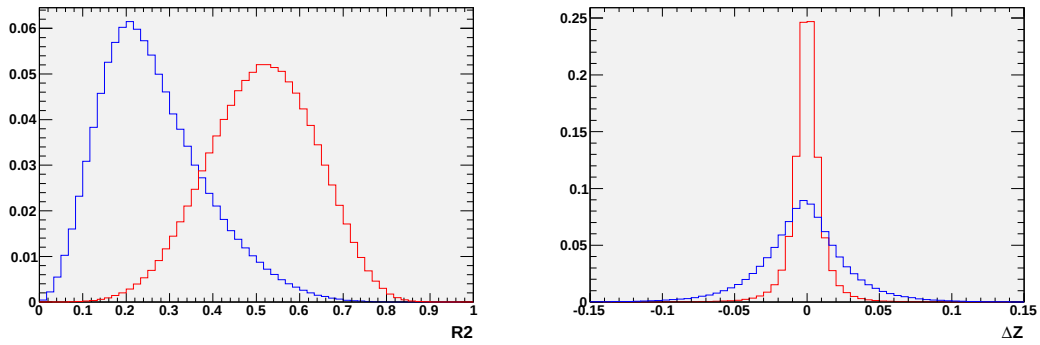


Figure 38: Distributions of R2 (left) and Δz (right) for signal and continuum background.

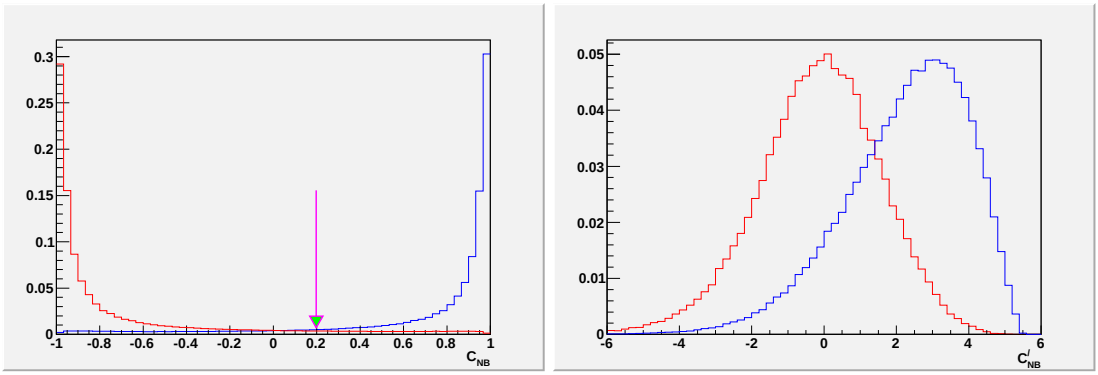


Figure 39: C_{NB} and C'_{NB} distributions obtained with signal (left) and $q\bar{q}$ (right) MC sample. The arrow indicates the applied cut position.

6 B background study

6.1 Generic B background

To find out the potential background from $b \to c$ decays, we process a sample of generic $B\bar{B}$ MC as described in Sec 4.1, and look at their $M_{\rm bc}$ and ΔE distributions (Fig. 40). The $M_{\rm bc}$ distribution is found to be strongly peaking in the signal region. In order to investigate what is the source of this peak, we study the invariant mass distributions of $K^+\pi^0$ and K^+K^- systems.

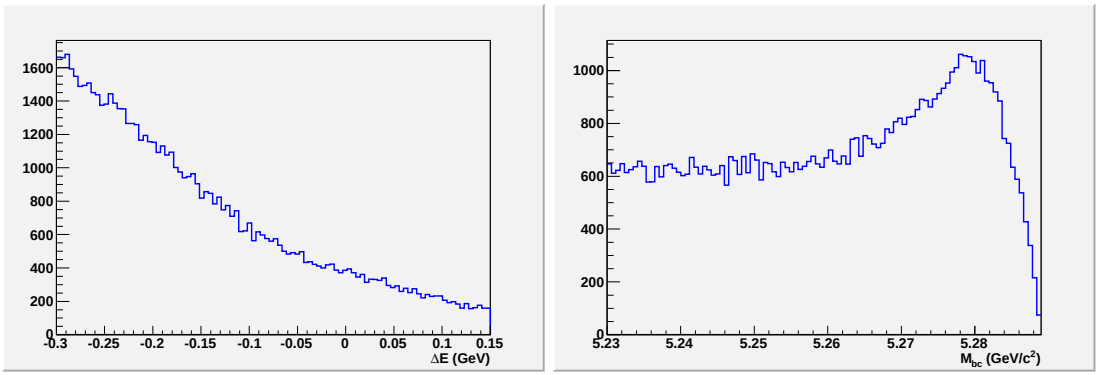


Figure 40: ΔE (left) and $M_{\rm bc}$ (right) distributions obtained from the generic $B\bar{B}$ MC sample.

We do not see any peak in $K^+\pi^0$ mass distribution [Fig. 41 (left)] but the K^+K^- invariant mass distribution [Fig. 41 (right)] depicts two peaks that correspond to the contributions from a charm meson: the left one peaking at the nominal D^0 mass is the decay $D^0 \to K^+K^-$, and the right one with the peak slightly shifted from the D^0 mass comes from $D^0 \to K^-\pi^+$ owing to K–π misidentification. (We confirm the $D^0 \to K^-\pi^+$ contribution by changing the mass assignment of the kaon tracks to pion, which subsequently shifts the invariant mass to the nominal D^0 position.) In order to reject the charm background we employ a veto around the two peaks. The veto window $[1846, 1884]\,{\rm MeV/c^2}$, which corresponds to approximately $\pm 5\sigma$ around the mean, is decided by fitting the K^+K^- invariant mass to a Gaussian function. Figure 42 shows results of the Gaussian fit along with the chosen veto position, indicated by two arrows. In Fig. 43 we show the $m(K^+\pi^-)$ vs $m(\pi^+K^-)$ distribution while in Fig. 44 we illustrate the utility of the veto requirement on the KK and $K\pi$ systems.

Figure 45 shows the $M_{\rm bc}$ and ΔE distributions before and after applying the charm veto. We find the resultant distributions have a linear and an ARGUS shapes, respectively, similar to the continuum $q\bar{q}$ background. Thus we introduce another dimension C'_{NB} [see Fig. 51 (top-right)] to avoid a large event flow between generic $B\bar{B}$ and con-

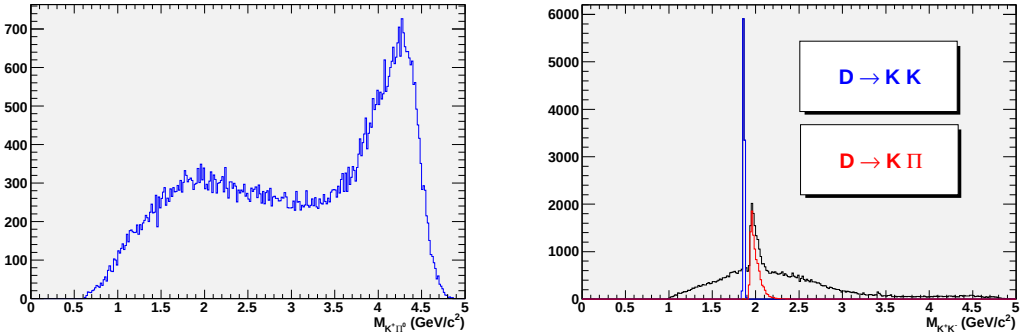


Figure 41: Invariant mass distributions of $K^+\pi^0$ (left) and K^+K^- (right) systems obtained from generic $B\bar{B}$ MC sample. The former shows no peaking structure, whereas the latter has two peaks that correspond to the contributions from $D^0 \to K^+K^-$ and $D^0 \to K^+\pi^-$.

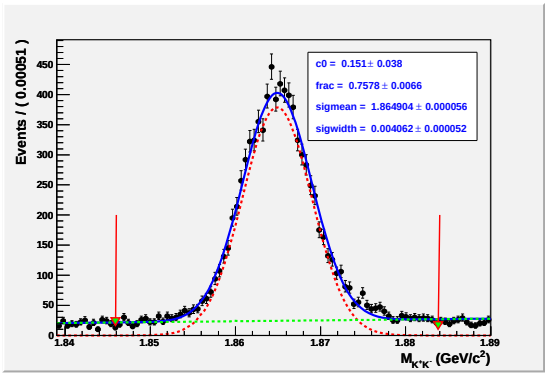


Figure 42: Fitted mass distribution of the K^+K^- system. Red arrows indicate the chosen veto region.

tinuum $q\bar{q}$ components owing to similar $M_{\rm bc}$ and ΔE shapes.

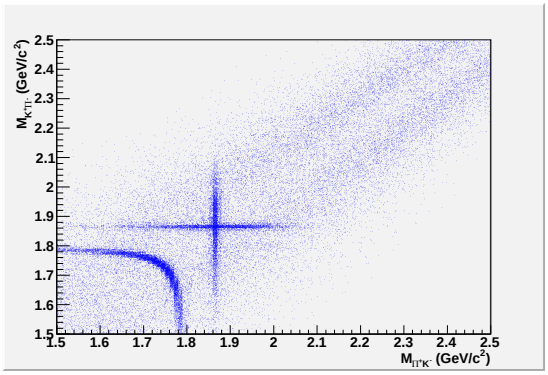


Figure 43: $m[K^+\pi^-]$ (obtained with mass assignment of K^- changed to π^-) vs $m[\pi^+K^-]$ (obtained with mass assignment of K^+ changed to π^+) - using generic $B\bar{B}$ MC.

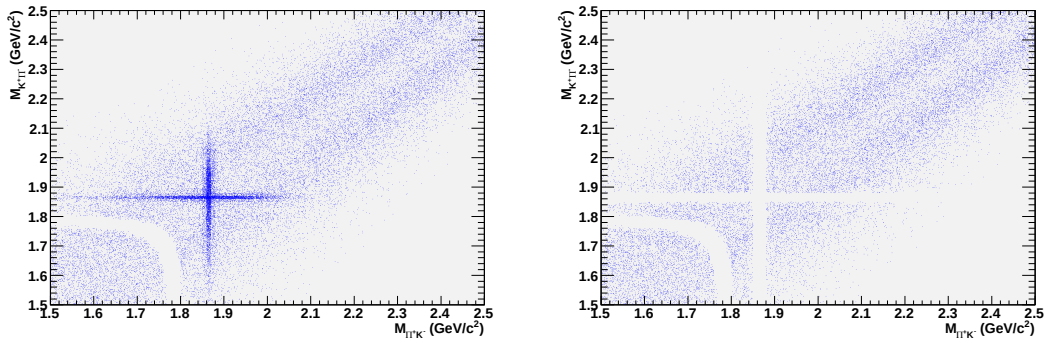


Figure 44: $m(K^+\pi^-)$ vs $m(\pi^+K^-)$ distributions with the charm veto applied on the K^+K^- system (left) only as well as on both KK and $K\pi$ systems (right).

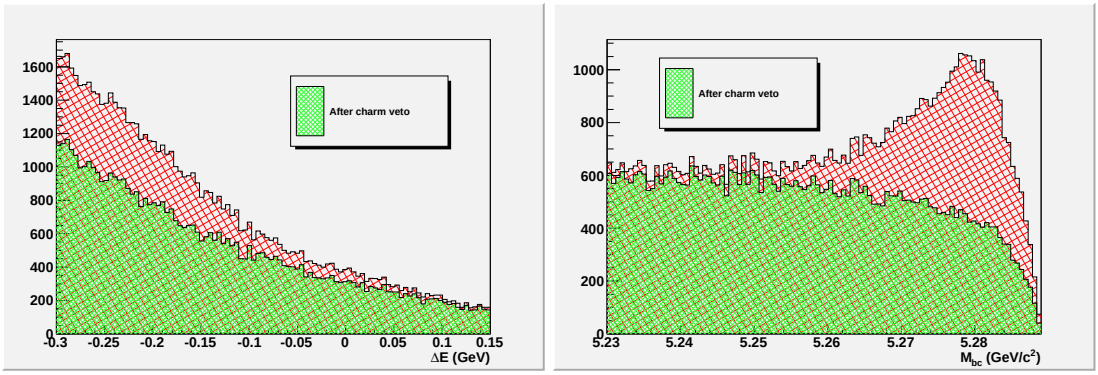


Figure 45: ΔE (left) and $M_{\rm bc}$ (right) distributions before and after the charm veto application.

6.2 Rare B background

There are a few background modes that contribute in the $M_{\rm bc}$ signal region having the ΔE peak shifted from zero on the positive side. The so-called “rare peaking”

background modes, resulting mostly due to K–π misidentification, are identified with a $B\bar{B}$ MC sample in which one of the B mesons decays via $b \to u,d,s$ transitions with known (where we have a measurement for the branching fraction) or estimated (where we have an upper limit) branching fractions. The rare peaking background includes the $B^0 \to K^+\pi^-\pi^0$ nonresonant decay as well as possible intermediate resonant modes that contribute to the $K^+\pi^-\pi^0$ Dalitz plot, such as $B^0 \to K^*(892)^0\pi^0$ and $B^0 \to K^*(892)^+\pi^-$. Table 6 lists various peaking background modes identified in the rare $B\bar{B}$ MC sample. We represent this component by green filled histograms in Fig. 46.

Table 6: List of possible peaking background modes identified in the rare MC sample. Inclusion of charge conjugate reactions is implicit.

Mode number	Mode
1.	$B^0 \to K^{*0}(1410)\pi^0$
2.	$B^0 \to K^{*0}(1680)\pi^0$
3.	$B^0 \to K^+\pi^-\pi^0$
4.	$B^0 \to \rho(770)^- K^+$
5.	$B^0 \to K_0^{*+}(1430)\pi^-$

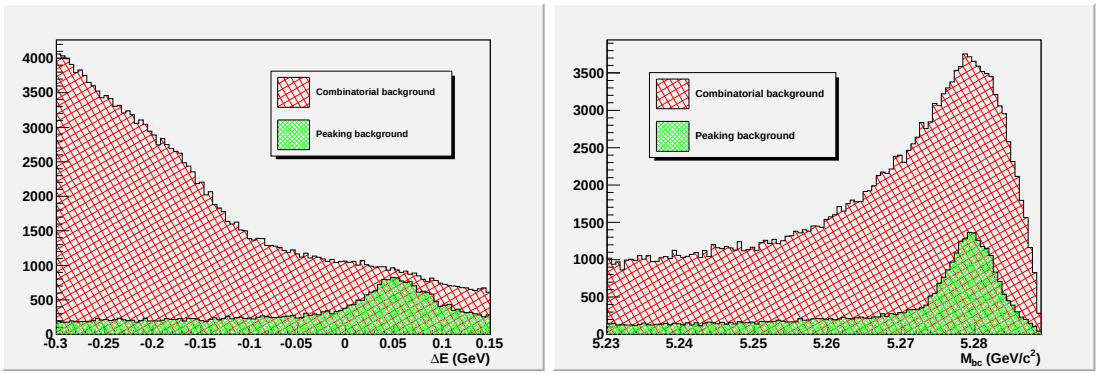


Figure 46: ΔE (left) and $M_{\rm bc}$ (right) distributions showing the contribution of rare combinatorial and rare peaking background.

The expected peaking background yield = $\mathcal{B}(B^0 \to K^+\pi^-\pi^0) \times N_{B\bar{B}} \times \varepsilon_{\rm rec}$ = 96, where, $\mathcal{B}(B^0 \to K^+\pi^-\pi^0) = (37.8 \pm 3.2) \times 10^{-6}$ [60], $N_{B\bar{B}}$ is the total number of $B\bar{B}$ pairs (772×10^6) and $\varepsilon_{\rm rec}$ (peaking component efficiency) = 0.3% estimated from the rare $B\bar{B}$ MC sample. We fix the yield of the rare peaking background in the fitter because we observe a bias in the signal yield if we allow it to float due to a small overlap between this and signal components.

The events that remain after removing the rare peaking component comprise the "rare combinatorial" background, shown as red filled histograms in Fig. 46. Figure 47 shows the dominant mode that contributes to the rare combinatorial background.

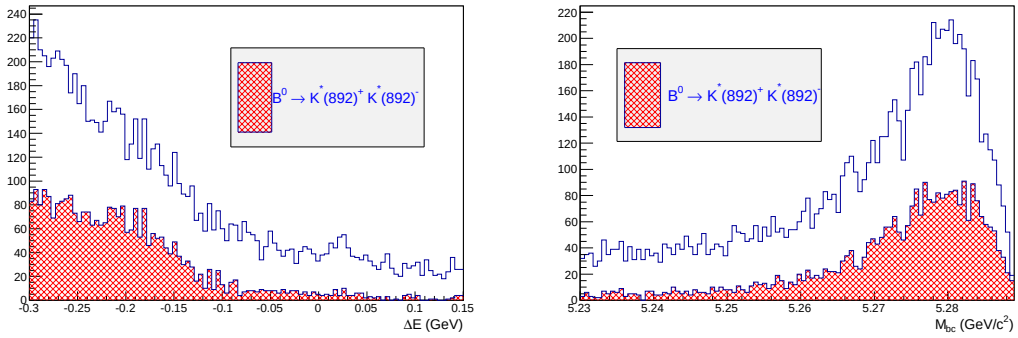


Figure 47: ΔE (left) and $M_{\rm bc}$ (right) distributions showing the dominant mode of the rare combinatorial background.

7 Extended maximum likelihood fit

7.1 Fit to MC samples

Our initial plan was to go with a 3D unbinned extended maximum likelihood fit comprising ΔE, M_{bc} and $C_{NB}^{'}$. However, we observed a bias of 12% on the fitted signal yield owing to a strong correlation between ΔE and M_{bc}. Table 7 shows Pearson's linear correlation coefficient between ΔE and M_{bc} for various event categories.

Table 7: Linear correlation coefficient between $M_{\rm bc}$ and ΔE for various MC samples.

Component	Correlation factor
CR signal	-7.19%
SCF signal	-4.70%
Continuum $q\bar{q}$	$+1.48\%$
Generic $B\bar{B}$	-3.09%
Rare peaking	$+4.07\%$
Rare combinatorial	$+2.05\%$

Therefore our strategy is to rather perform a 2D unbinned fit to the ΔE and $C_{NB}^{'}$ distributions for extracting the signal yield in the data. As a first step, we carry out a basic test by fitting various MC samples. We simultaneously fit the two distributions using different analytic functions; whenever we face difficulty with analytic functions we resort to nonparametric (histogram) shapes. Table 8 lists various PDF shapes used to model ΔE and $C_{NB}^{'}$ distributions for the MC samples. The PDF shapes are described in Appendix A.2. In Figs. 48 to 53 we present the obtained fit results and corresponding pull (data-fit) distributions, where analytic functions are used.

Table 8: List of PDFs used to model ΔE and $C_{NB}^{'}$ distributions of various MC samples. G, BG, CB and Poly2 denote Gaussian, bifurcated Gaussian, Crystal Ball [61] and second-order Chebyshev polynomial function, respectively.

Component	ΔE	$C_{NB}^{'}$
CR signal	CB+BG	3 BG
SCF signal	1D histogram	1D histogram
Continuum $q\bar{q}$	Poly2	BG
Generic $B\bar{B}$	Poly2	BG
Rare peaking	2G	BG
Rare combinatorial	1D histogram	3 BG

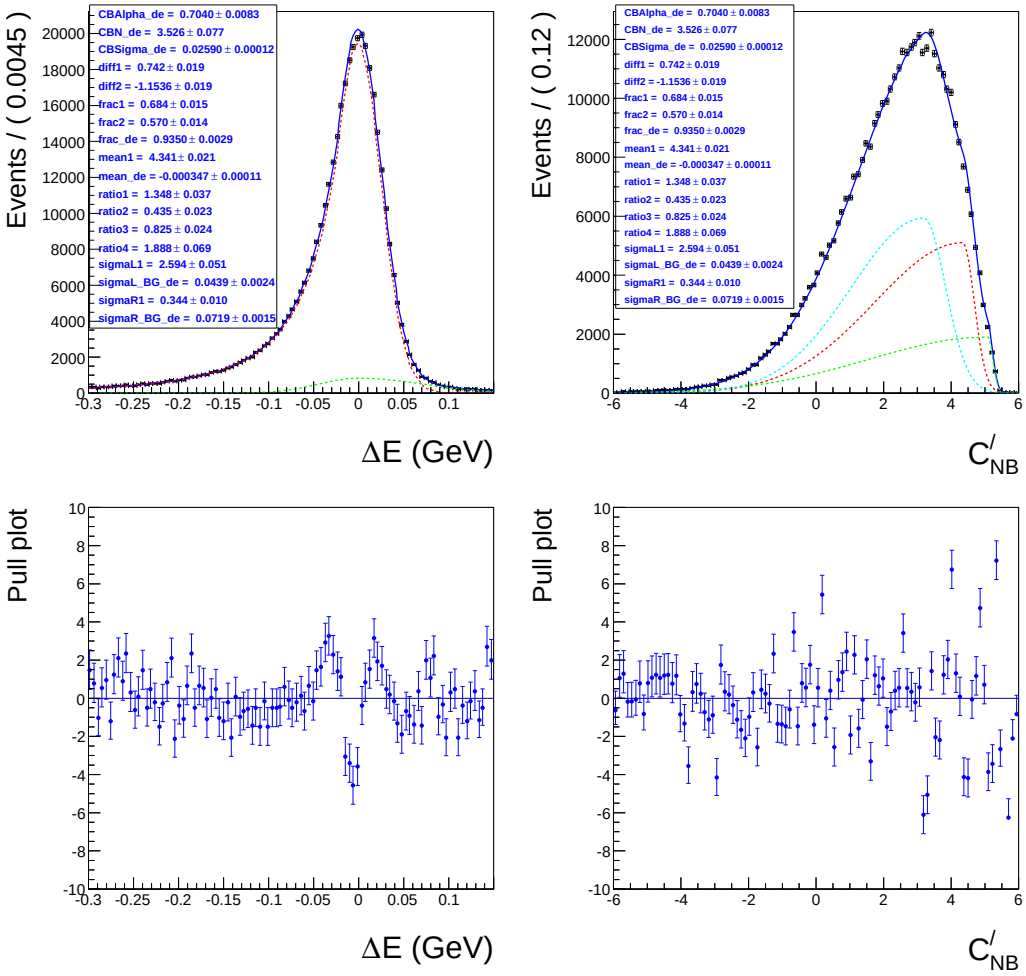


Figure 48: 2D fit obtained using the CR signal component.

For the CR signal component, C'_{NB} is parameterized in terms of:

diff1 $= \mu_2 - \mu_1$ and diff2 $= \mu_3 - \mu_1$

ratio1 $= \sigma_{L2}/\sigma_{L1}$ and ratio2 $= \sigma_{R2}/\sigma_{R1}$

ratio3 $= \sigma_{L3}/\sigma_{L1}$ and ratio4 $= \sigma_{R3}/\sigma_{R1}$

we define two μ differences and four σ ratios to minimize the correlations between these fit parameters.

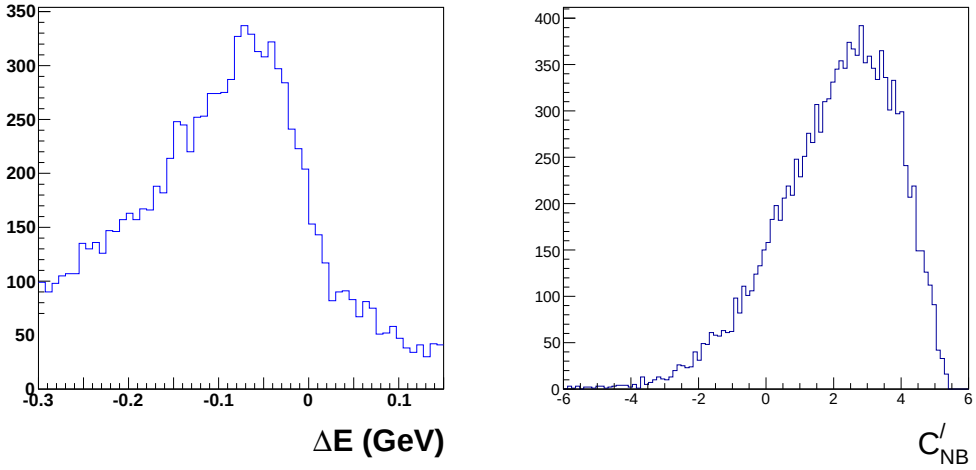


Figure 49: 1D histogram PDFs used for ΔE (left) and C'_{NB} (right) of the SCF signal component.

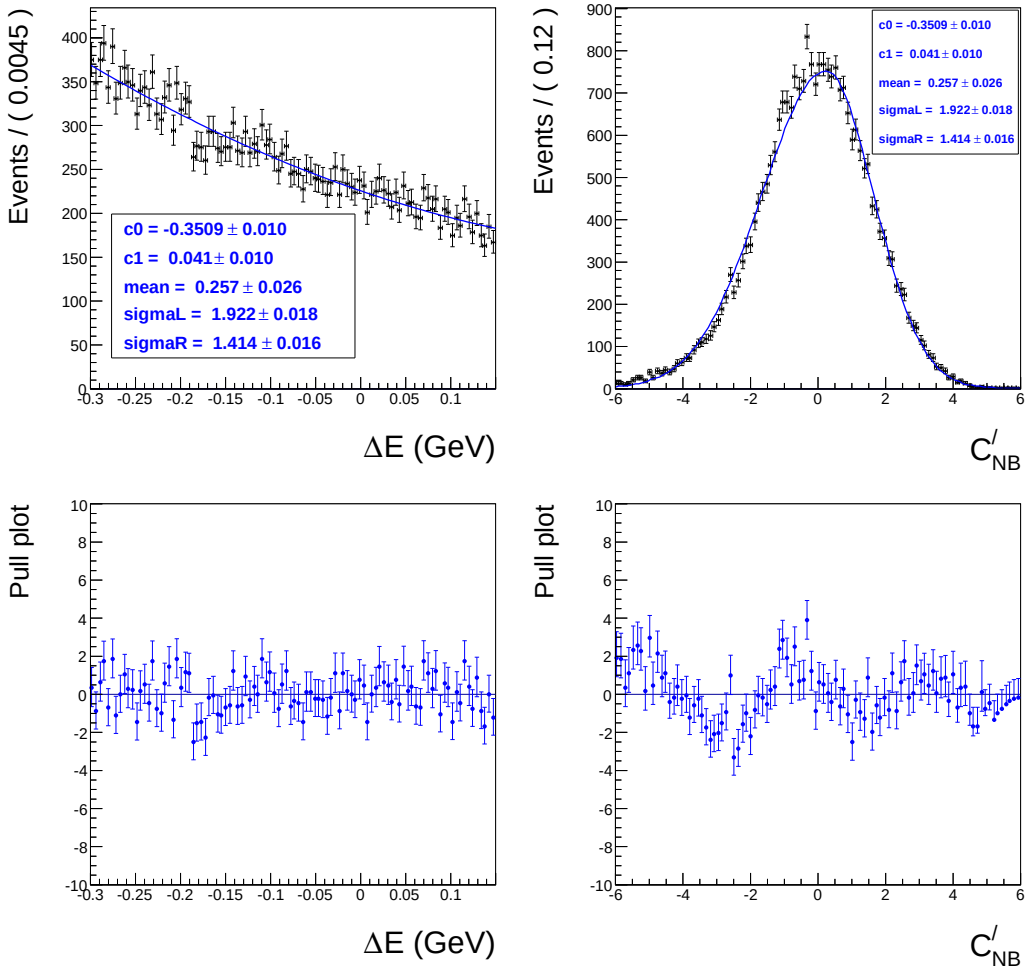


Figure 50: 2D fit obtained using the continuum $q\bar{q}$ MC sample.

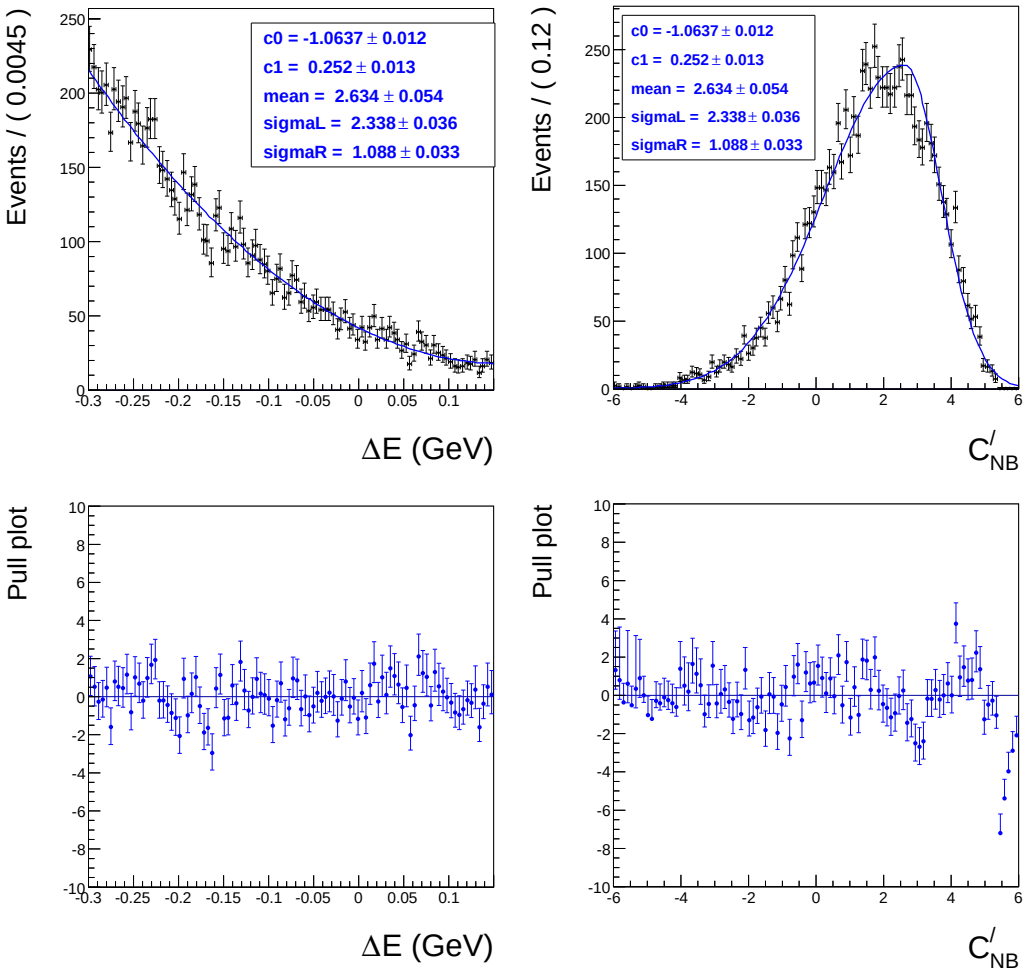


Figure 51: 2D fit obtained using the generic $B\bar{B}$ MC sample.

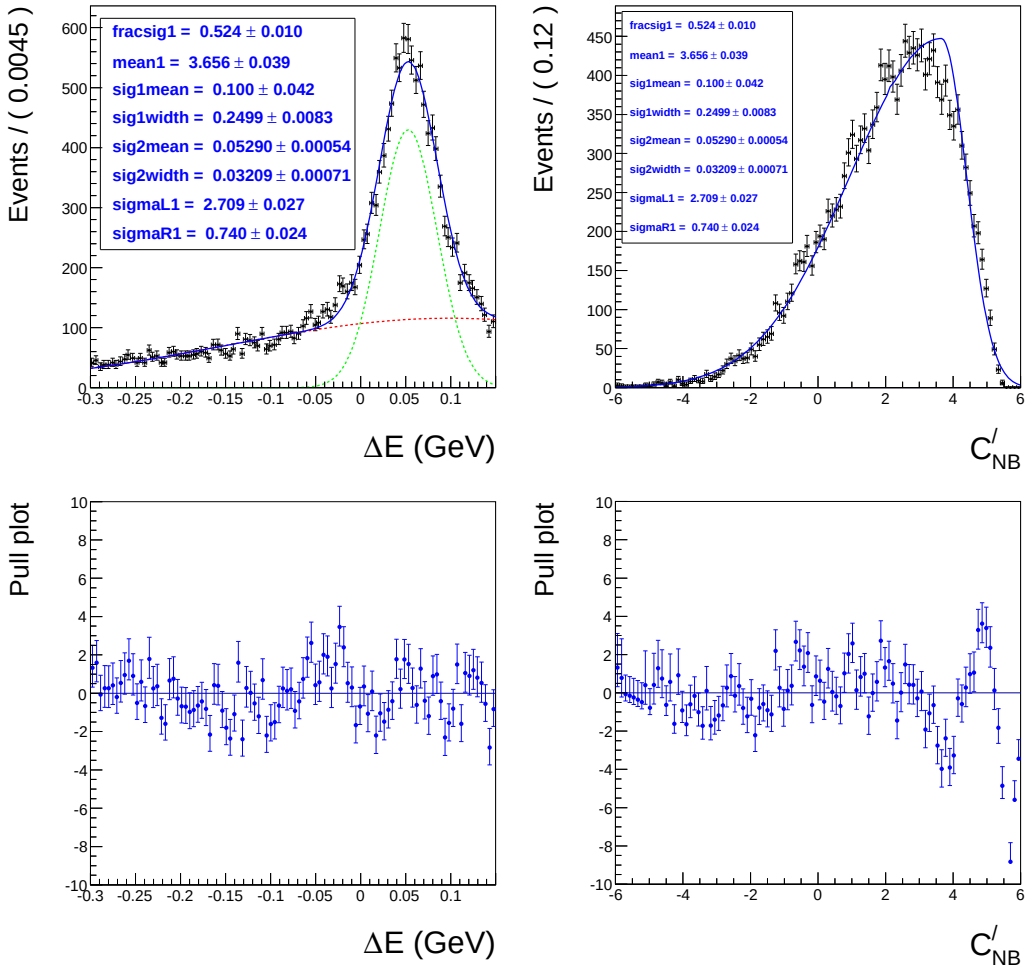


Figure 52: 2D fit obtained using the rare peaking $B\bar{B}$ MC sample.

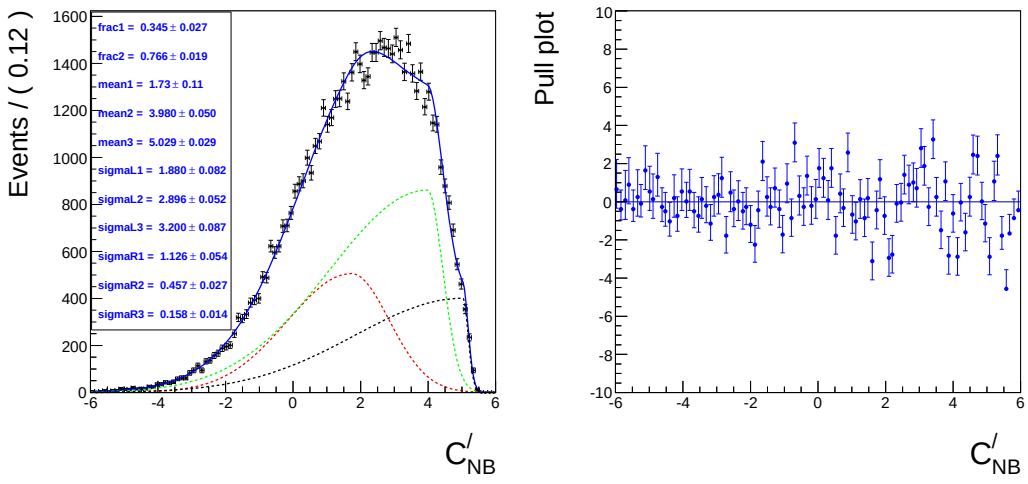


Figure 53: 1D fit obtained using the rare combinatorial $B\bar{B}$ MC sample.

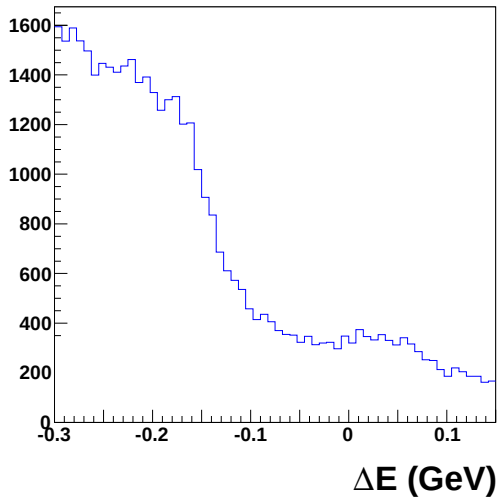


Figure 54: 1D histogram PDF used for ΔE of the rare combinatorial $B\bar{B}$ MC.

7.2 Correlations between different discriminating variables

Before going for a two-dimensional fit comprising ΔE and $C_{NB}^{'}$, it is important to find out the level of correlation between the two variables for all event categories. In Table 9 we present Pearson's linear correlation coefficient between ΔE and $C_{NB}^{'}$. In addition, we have also studied two-dimensional scatter plots as well as the variation of the mean value of $C_{NB}^{'}$ in bins of ΔE (see Appendix A.5). All these results point toward a negligible correlation between the two variables.

Table 9: Linear correlation coefficient between ΔE and $C_{NB}^{'}$ for various MC samples.

Component	Correlation factor
CR signal	$+0.00\%$
SCF signal	-1.42%
Continuum $q\bar{q}$	-0.42%
Generic $B\bar{B}$	-4.21%
Rare peaking	$+2.14\%$
Rare combinatorial	$+2.17\%$

We define a probability density function (PDF) for each event category j (signal, $q\bar{q}$, generic $B\bar{B}$, rare peaking, and rare combinatorial $B\bar{B}$ backgrounds):

$$\mathcal{P}_j^i \equiv \mathcal{P}_j(\Delta E^{\,i})\mathcal{P}_j(C_{NB}^{'\,i}), \tag{15}$$

where i denotes the event index. Since the correlation between ΔE and $C_{NB}^{'}$ is found to be negligible, the product of two individual PDFs is a good approximation for the combined PDF. The extended likelihood function is

$$\mathcal{L} = \exp\left(-\sum_j n_j\right) \times \prod_i \left[\sum_j n_j \mathcal{P}_j^i\right], \tag{16}$$

where n_j is the yield of event category j.

7.3 Pure toy study with the 2D fitter

In this section we present results of a two-dimensional pure toy ensemble test comprising ΔE and $C_{NB}^{'}$ after applying a signal-region requirement on $M_{\rm bc}$ ($5.271\,{\rm GeV/c^2} < M_{\rm bc} < 5.287\,{\rm GeV/c^2}$). We prepare an ensemble of 200 pseudo-experiments, each having a data set of similar size to what is expected in the full $\Upsilon(4S)$ sample. PDF shapes listed in Table 8 are used to generate these toy datasets. We then fit to the ensemble of pseudo-experiments to check for the error coverage and any preset bias. If none of

them were present, we would expect the fit to yield a Gaussian distribution with zero mean and unit width for each of the floated parameters. In the fit, we float the expected yield for the four event categories (signal, continuum, generic $B\bar{B}$ and rare combinatorial backgrounds) as well as the following PDF shape parameters for the continuum background:

- c0 and c1 of the second-order Chebyshev polynomial used for ΔE
- mean, σ_L and σ_R of the bifurcated Gaussian used for $C_{NB}^{'}$

The CR and SCF components of signal are treated as a single component in the fitter. Their combined PDF is given as $n_{sig} \times [(1-f)\,\mathrm{PDF_{CR}} + f\,\mathrm{PDF_{SCF}}]$, where f is the SCF fraction, fixed to the MC expected value 3%, and n_{sig} is the total signal yield.

Figure 55 presents the fitted and pull distribution of the signal yield (The number in red denotes the expected value of the parameter.) For all the floated parameters in the fit, we find the fitted values to be normally distributed around the expected value and the pull distributions to be distributed around zero with unit width (see Table 10). This exercise confirms a good error coverage for different fit parameters and that there is no preset bias.

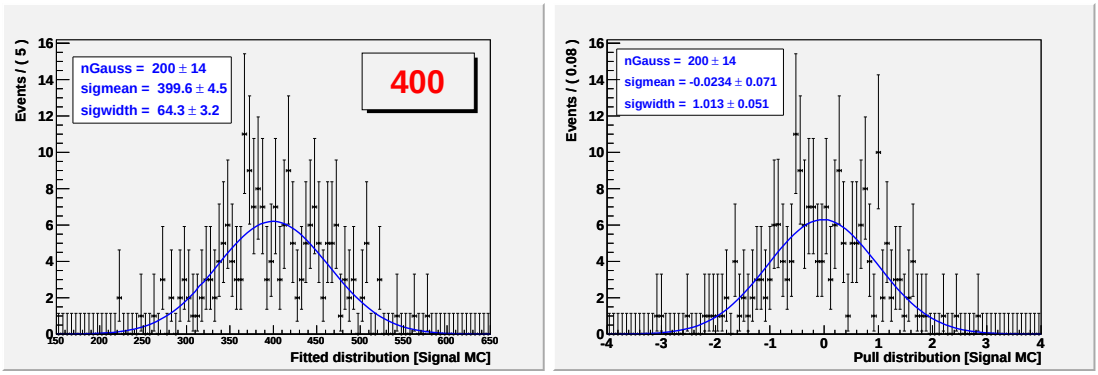


Figure 55: Fitted (left) and pull (right) distributions of the signal yield in the 2D pure toy test.

Table 10: Pure toy results.

Floated parameter	Expected value	Fitted value	Pull mean	Pull width
Signal yield	400	399.6 ± 4.5	-0.02 ± 0.07	1.01 ± 0.05
Continuum $q\bar{q}$ yield	26151	26126 ± 25	-0.11 ± 0.07	0.98 ± 0.05
Generic $B\bar{B}$ yield	1704	1734 ± 32	$+0.12 \pm 0.08$	1.08 ± 0.05
Rare combinatorial yield	1201	1180 ± 17	-0.14 ± 0.07	1.05 ± 0.05
$c0$ parameter	-0.35	-0.35 ± 0.00	$+0.03 \pm 0.08$	1.08 ± 0.05
$c1$ parameter	0.04	0.04 ± 0.00	$+0.05 \pm 0.07$	0.96 ± 0.05
mean parameter	0.26	0.26 ± 0.00	-0.02 ± 0.07	1.00 ± 0.05
σ_L parameter	1.92	1.92 ± 0.00	-0.03 ± 0.07	0.96 ± 0.05
σ_R parameter	1.41	1.41 ± 0.00	-0.02 ± 0.07	0.99 ± 0.05

7.4 GSIM ensemble test with the 2D fitter

Results of the pure toy test described in the previous subsection won't tell us if there is a bias inherent on the fit because of the unaccounted for correlation between the fit variables. Therefore, we perform an ensemble test comprising 200 pseudo-experiments where signal is embedded from the corresponding MC sample and PDF shapes are used to generate the dataset for all type of background events (continuum, generic or rare $B\bar{B}$). Similar to the pure toy case, we present the fitted and pull distributions for signal yield in Fig. 56.

As evident from the left plot of Fig. 56, we observe a bias of 14.5 events, which is 3.6% of the expected yield and about 22.5% of the expected statistical error. We plan to allocate this bias as part of the systematic uncertainty. There is a non-negligible overlap between the $C_{NB}^{'}$ distribution of signal and $q\bar{q}$ background, where the former component distributes towards the right while the later towards the left. A slight shift of the mean or width of the $q\bar{q}$ $C_{NB}^{'}$ towards the right can help in the migration of few events to the signal side. Of course, migration of few events from the background side won't have a big impact on its fitted yield because of its sheer size, however it causes a bias of 3.6% on the fitted signal yield. Thus the possible variation in the continuum $C_{NB}^{'}$ PDF shape is found out to be the major contributor here.

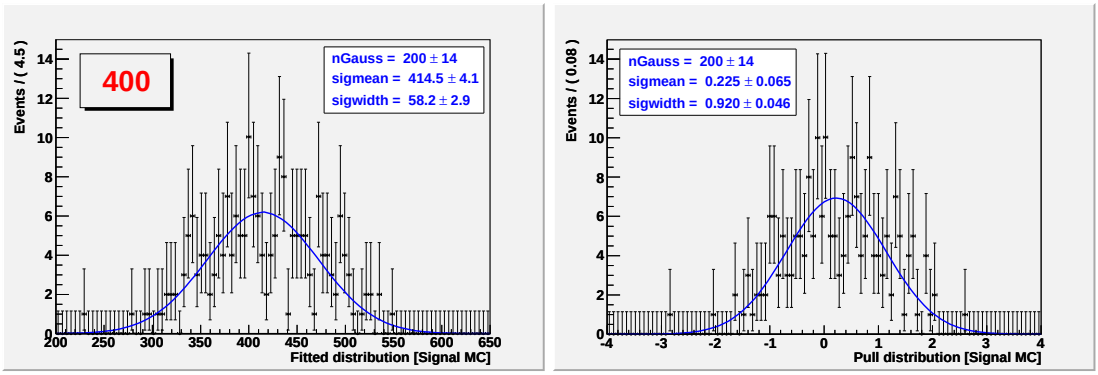


Figure 56: Fitted (left) and pull (right) distributions of the signal yield in the 2D GSIM test.

We perform a linearity test where several 2D GSIM ensemble tests are carried out with an assumed signal yield ranging from 0 to 400. This is particularly important as we do not know for sure whether our expectation of 400 would really hold in the data or not. Results of the linearity test are presented in Figs. 57 and 58.

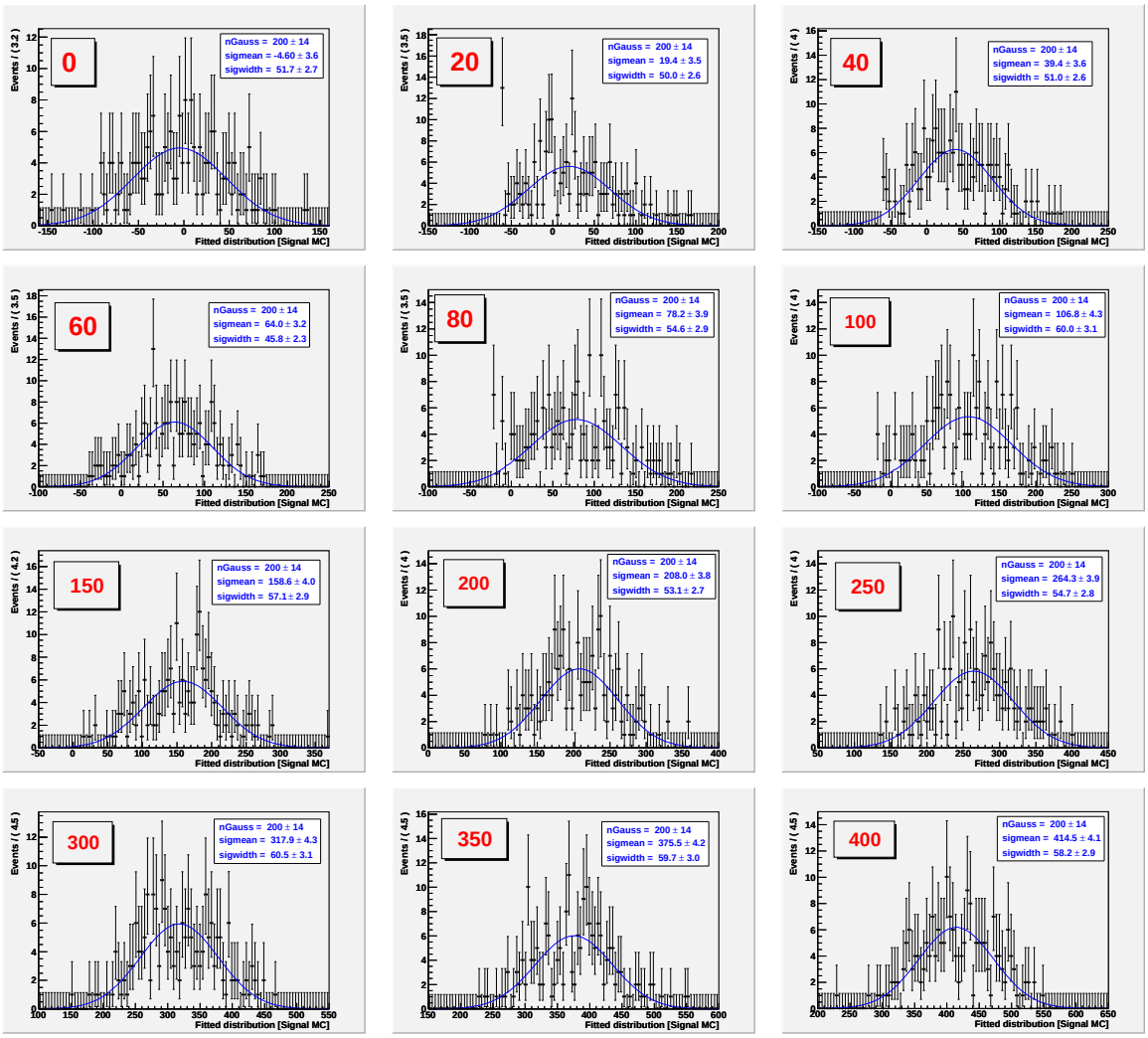


Figure 57: Distributions of fitted signal yields in a number of 2D GSIM test carried out with the input signal yield varying from 0 to 400.

Looking at the fitted yield as a function of the input yield, we expect to get a straight line of unit slope and zero intercept. From Fig. 58 we find slope (intercept) to be 1.06 ± 0.01 (-2.22 ± 1.69). Results of the GSIM linearity test are therefore reasonably consistent with expectations.

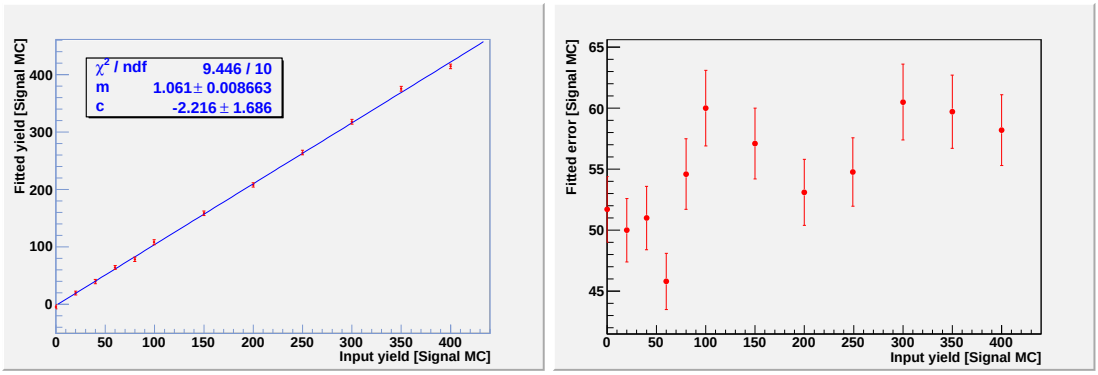


Figure 58: Fitted yield vs input yield (left) and fitted error vs input yield (right) obtained with the GSIM linearity test.

8 Control sample study

In order to rely less on Monte Carlo simulations but more on data (for suppressing possible systematic effects) we use a high-statistics control sample to determine the mean and width of ΔE and $C_{NB}^{'}$ used in the fit. We choose a decay process that has the advantage of a large branching fraction but at the same time does not suffer from backgrounds that much. Our chosen control channel is $B^+ \to \bar{D}^0(K^+\pi^-\pi^0)\pi^+$ which has a similar topology as our signal decay though with an extra pion. Two million signal MC events for the control channel are generated according to the D decay Dalitz plot [54].

8.1 Selection requirements

In the following, we list various requirements applied to select the control channel candidates. These requirements are identical to our signal selection, except for the last one on the reconstructed D mass.

- Track $p_T > 100$ MeV/c
- $|\mathrm{dr}| < 0.2$ cm and $|\mathrm{dz}| < 5.0$ cm
- $L(K/\pi) > 0.6$ for kaon and $L(K/\pi) < 0.4$ for pion
- Energy of photons $> 60\ (100)$ MeV in the barrel (endcap) ECL
- Reconstructed π^0 mass is in the range $[112, 156]\,\mathrm{MeV/c^2}$, corresponding to a $\pm 3.5\sigma$ window around the nominal mass
- The π^0 mass-constrained $\chi^2 < 50$
- The decay helicity angle of π^0, $|\cos(\theta_{hel})| < 0.95$
- $5.271\,\mathrm{GeV/c^2} < M_{\mathrm{bc}} < 5.287\,\mathrm{GeV/c^2}$
- $-0.30\,\mathrm{GeV} < \Delta E < 0.15\,\mathrm{GeV}$
- Reconstructed D mass is in the range $[1.814, 1.909]\,\mathrm{GeV/c^2}$, corresponding to a $\pm 3.5\sigma$ window around the nominal mass (Fig. 59)

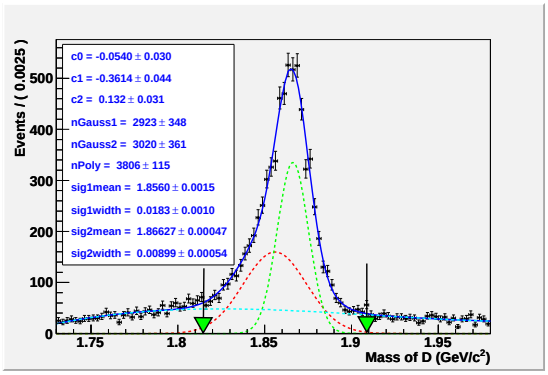


Figure 59: Reconstructed D mass for the control channel signal MC.

8.2 Fit to MC sample

For ΔE and $C_{NB}^{'}$ we use a Crystal-ball plus bifurcated Gaussian and a sum of three bifurcated Gaussians, respectively [Fig. 60]. Note that these functions are identical to what we have used for the ΔE-$C_{NB}^{'}$ 2D fit corresponding to $B^0 \to K^+K^-\pi^0$ signal. For $C_{NB}^{'}$, relative fractions between bifurcated Gaussians are tuned to the values from $B^0 \to K^+K^-\pi^0$ signal MC, so that the trend remains identical between the two cases.

8.3 ΔE-$C_{NB}^{'}$ distributions with the data

We study ΔE and $C_{NB}^{'}$ distributions for the control channel in the data, after applying the $M_{\rm bc}$ signal region cut. We observe two peaking structures in the ΔE. Before going ahead with the ΔE-$C_{NB}^{'}$ fit, we want to know the source of these structures. To investigate further we use charged and mixed type of generic $B\bar{B}$ MC samples. Most of the peaking contribution seems to arise from the charged generic sample. Therefore we look at the ΔE distribution for this sample in the $M_{\rm bc}$ signal region. Various modes identified are plotted separately in Fig. 61. The major contributor to the peaking structures is the decay $B^+ \to \bar{D}^{*0}(2007)\pi^+, \bar{D}^{*0} \to \bar{D}^0\pi^0/\bar{D}^0\gamma$.

We then apply a 1D ΔE fit (Fig. 62) to the data using a Crystal-ball plus bifurcated Gaussian function. The fraction between the Crystal-ball and the bifurcated Gaussian is fixed to the value obtained with the control channel signal MC. Two Crystal-ball function parameters (α and N) and two bifurcated Gaussian parameters (σ_L and σ_R), are also fixed to the MC values. For background we use two Gaussian functions to take care of the two peaking structures, and a second order Chebyshev polynomial to parameterize its combinatorial component.

We perform a 1D $C_{NB}^{'}$ fit (Fig. 63) using the data putting the ΔE signal-region cut. A small continuum contribution is taken into account using the ΔE sideband, $0.05\,\text{GeV} <$

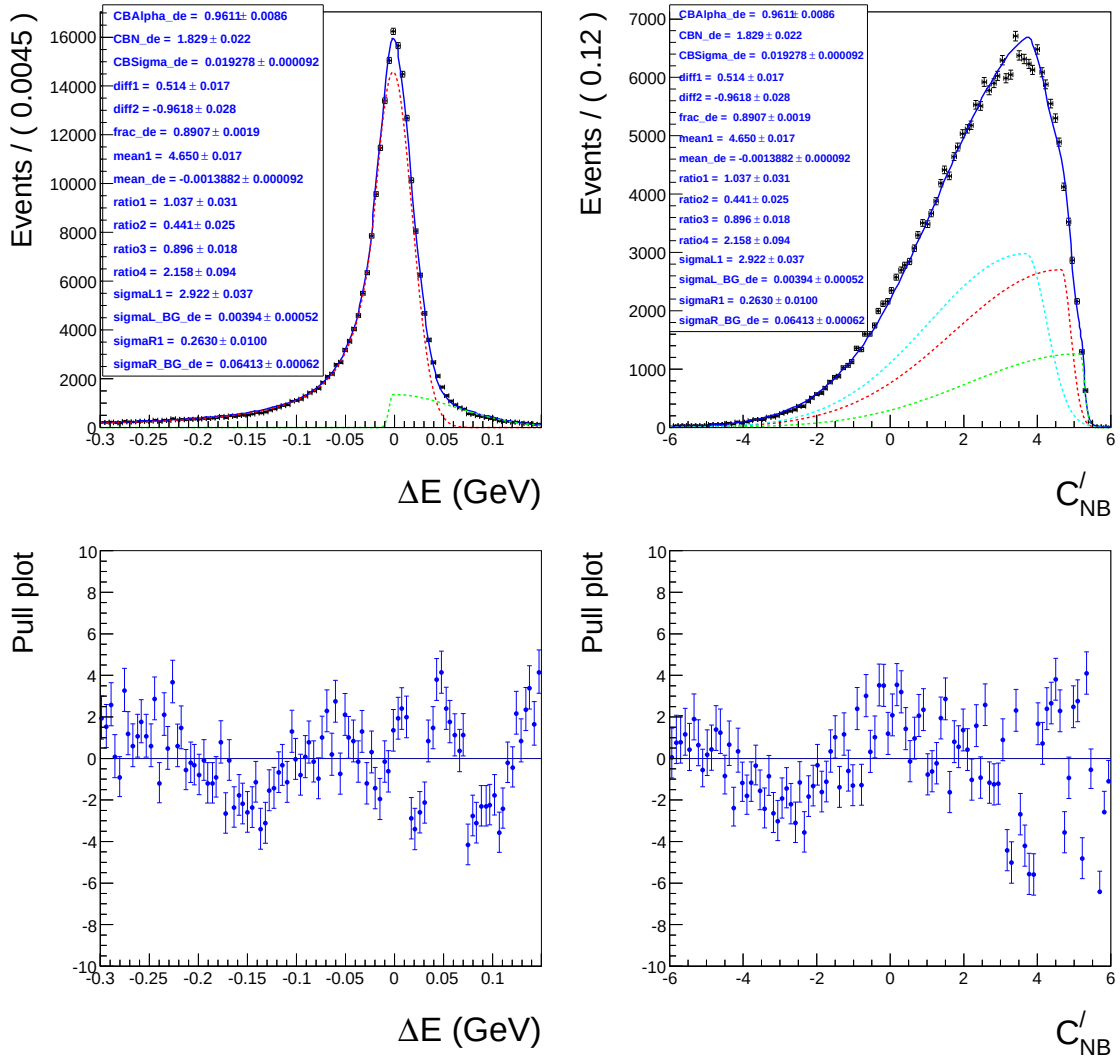


Figure 60: 2D fit obtained using the control channel signal MC.

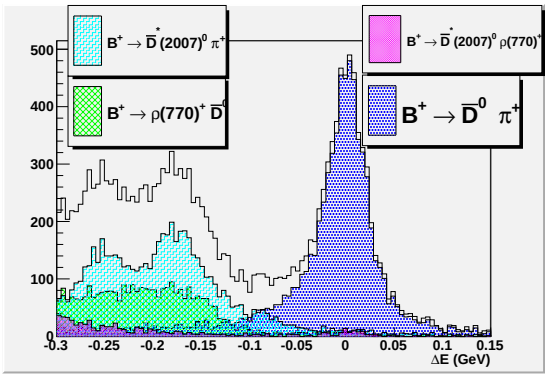


Figure 61: ΔE distribution in the M_{bc} signal region using charged type generic $B\bar{B}$ MC sample.

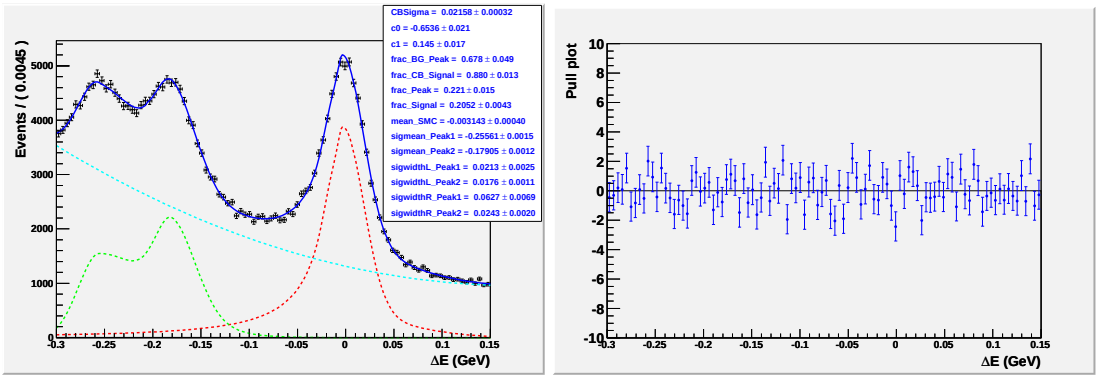


Figure 62: 1D ΔE fit (left) and pull plot (right) in the M_{bc} signal region using the data for the control channel.

$\Delta E < 0.15\,\text{GeV}$. We use the ΔE shape from the sideband and normalize it to the number of expected events, corresponding to the contribution of the second-order Chebyshev polynomial in $|\Delta E| < 50$ MeV. The continuum contribution is subtracted bin-by-bin from the total ΔE distribution, and the remainder is fitted to a sum of three bifurcated Gaussians similar to the $C_{NB}^{'}$ fit of the control sample signal MC. Two fractions between these functions as well as four sigma ratios (ratio1, ratio2, ratio3 and ratio4) and mean difference (diff1) are fixed to the signal MC values.

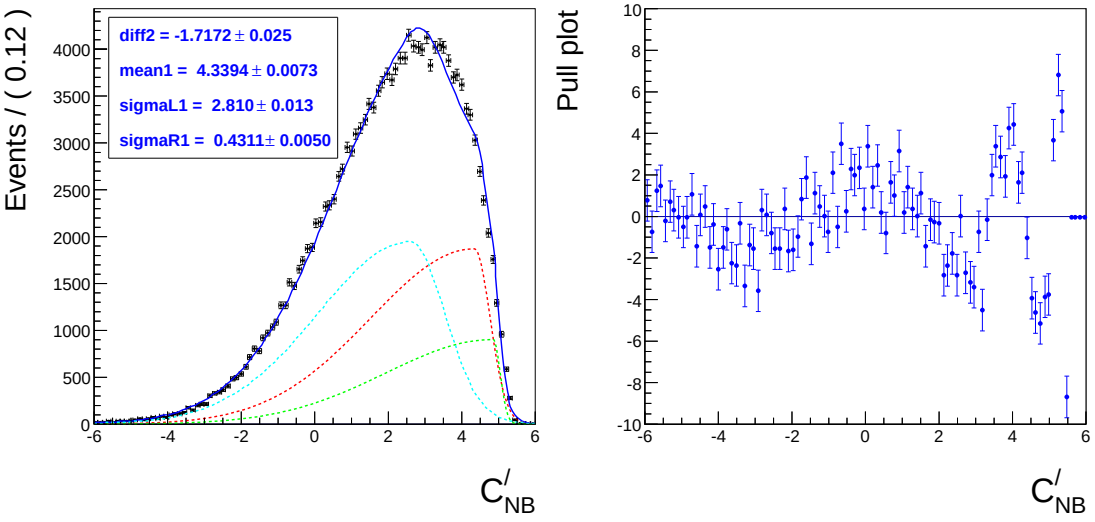


Figure 63: 1D $C_{NB}^{'}$ fit (left) and pull plot (right) in the ΔE signal region using the data for the control channel.

Now we are in a position to calculate the fudge factors, which will be applied as corrections at the time of fitting to the data in the signal region for $B^0 \to K^+K^-\pi^0$. Table 11 summarizes fudge factors for ΔE. These are obtained by comparing the left plot of Fig. 60 with that of Fig. 62. Table 12 shows fudge factors for $C_{NB}^{'}$. These are

obtained by comparing the right plot of Fig. 60 with the left plot of Fig. 63. The fudge factor on mean is the difference between the data and MC value, while the width fudge factor is the ratio of the corresponding data to MC value.

Table 11: Fudge factors for ΔE.

Parameter	Data	MC	Fudge factor
μ_{CB}/μ_{BG}(GeV)	-0.0031 ± 0.0004	-0.0014 ± 0.0001	-0.0018 ± 0.0005
σ_{CB}(GeV)	0.0216 ± 0.0003	0.0193 ± 0.0001	1.1194 ± 0.0219

Table 12: Fudge factors for C'_{NB}.

Parameter	Data	MC	Fudge factor
diff2	-1.7172 ± 0.0250	-0.9618 ± 0.0280	-0.7554 ± 0.0530
μ_1	4.3394 ± 0.0073	4.6500 ± 0.0170	-0.3106 ± 0.0243
σ_{L1}	2.8100 ± 0.0130	2.9220 ± 0.0370	0.9617 ± 0.0166
σ_{R1}	0.4311 ± 0.0050	0.2630 ± 0.0100	1.6392 ± 0.0814

We also measure the branching fraction for the control sample. The measured $\mathcal{B}(B^+ \to \bar{D}^0\ (K^+\ \pi^-\ \pi^0)\pi^+\ ) = [4.60 \pm 0.26(\text{stat})] \times 10^{-3}$ can be compared against $\mathcal{B}_{PDG} = [4.84 \pm 0.15] \times 10^{-3}$ [56]. Our result is therefore consistent with the world-average value.

9 Estimation of the signal yield

9.1 2D ΔE-$C_{NB}^{'}$ fit using the real data

We perform 2D ΔE-$C_{NB}^{'}$ fit using the real data. While making the projection plot for a variable, we have applied the following signal-region requirement for the other variable.

- $|\Delta E| < 30$ MeV, and
- $C_{NB}^{'} > 3$.

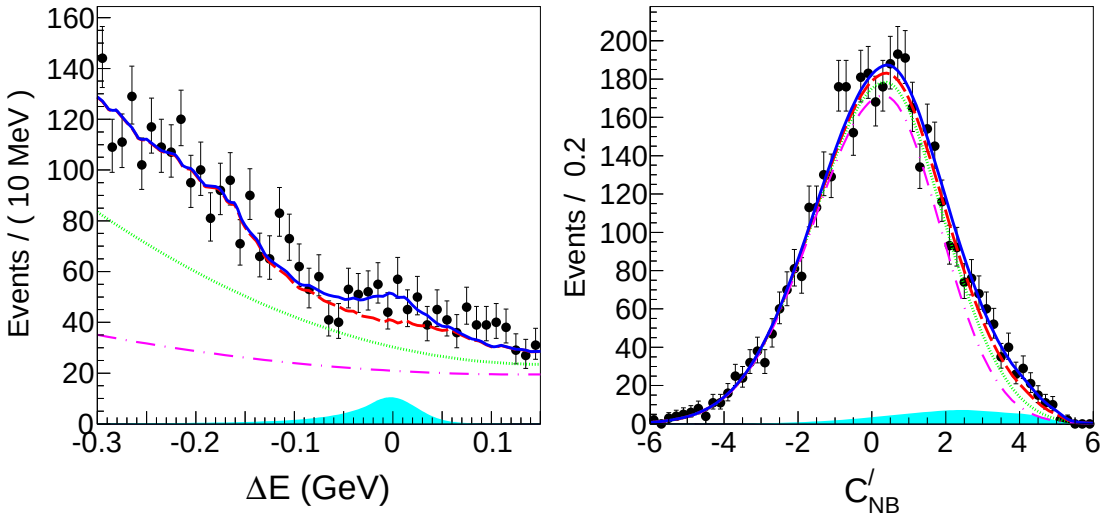


Figure 64: Projections of candidate events onto (left) ΔE for $C_{NB}^{'} > 3$ and (right) $C_{NB}^{'}$ for $|\Delta E| < 30$ MeV. Points with error bars are the data, solid (blue) curves are the total PDF, dashed (red) curves are the total background, dotted (green) curves are the sum of continuum $q\bar{q}$ and generic $B\bar{B}$ backgrounds, dash-dotted (magenta) curves are the continuum $q\bar{q}$ background, and filled (cyan) regions show the signal.

The results are presented in Fig. 64. A signal yield of 299 ± 83 is obtained with a statistical significance of 3.7σ. The significance is defined as $\sqrt{2\log(\mathcal{L}_0/\mathcal{L}_{\rm max})}$, where $\mathcal{L}_{\rm max}$ ($\mathcal{L}_0$) is the likelihood for the best fit with the signal branching fraction being allowed to vary (fixed to zero). Figure 70 (top-left) shows the projection plot of NLL.

9.2 Systematic error

Following are the possible sources of systematic error:
a)Fit bias: We need to assign a systematic error for the fit bias on signal yield (observed from the GSIM ensemble test).
Case A: We perform an ensemble test comprising 200 pseudo-experiments, where signal is embedded from the corresponding GSIM MC sample and PDF shapes are used to generate the dataset for all kind of background events (continuum, generic or rare $B\bar{B}$). In the test the central value of all floated parameters are picked from the 2D ΔE-C'_{NB} fit obtained using the real data. We present the fitted and pull distribution of signal MC in Fig. 65. As evident from the left plot of Fig. 65, we observe a bias of 3 events, which is 0.9% of the expected yield.

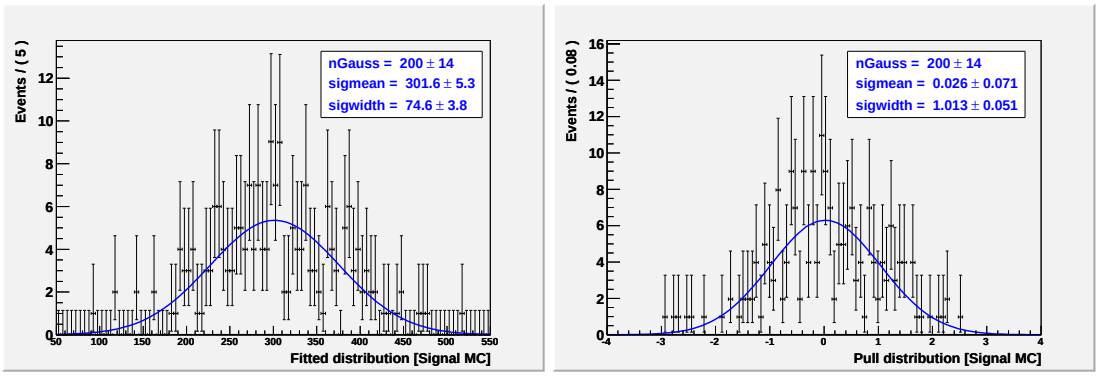


Figure 65: Fitted (left) and pull (right) distributions of the signal yield in the 2D GSIM test for case A.

Case B: We perform an ensemble test comprising 167 pseudo-experiments, where signal and rare peaking component are embedded from the corresponding GSIM MC samples and PDF shapes are used to generate the dataset for all remaining background events (continuum, generic $B\bar{B}$ and combinatorial component of rare $B\bar{B}$). The size of peaking component of rare MC limits us to work with 167 pseudo-experiments. In the test the central value of all floated parameters are picked from the 2D ΔE-C'_{NB} fit obtained using real data. We present the fitted and pull distribution in Fig. 66. We observe a bias of 4 events, which is 1.3% of the expected yield. We choose the case B to allocate systematic uncertainty for the fit bias. The fitted signal yield is 303.0 ± 6.1. We add the error on the fitted yield in quadrature to the bias and allocate a systematic error of 2.4% for the fit bias.

b) Systematic error from unbinned likelihood fit: The systematic error is obtained by changing each fixed PDF parameter by $\pm 1\sigma$. Table 13 shows the systematic uncertainties for various PDFs.

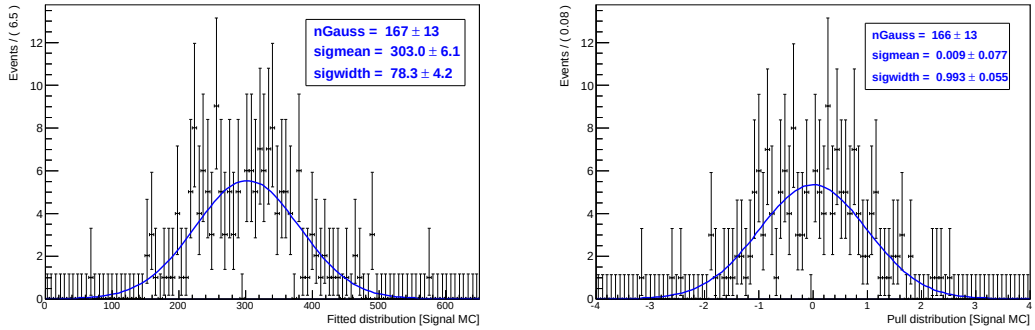


Figure 66: Fitted (left) and pull (right) distributions of the signal yield in the 2D GSIM test for case B.

Table 13: Systematic error obtained by changing each fixed PDF parameter by $\pm 1\sigma$.

Source	Uncertainties (%)	
	$+\sigma$	$-\sigma$
Signal PDF	$+3.4$	-2.9
Generic $B\bar{B}$ PDF	$+2.4$	-3.1
Combinatorial background PDF	$+1.3$	-2.0
Peaking background PDF	$+1.7$	-1.9

c) Systematic error due to the fixed SCF fraction: We also need to account for the self-crossfeed fraction that is fixed to the MC value. The uncertainty due to the fixed, small SCF fraction (3.2%) is estimated without knowing apriori how these SCF events vary across the Dalitz plot. We adopt a conservative approach to vary the SCF fraction by $\pm 50\%$ while calculating the associated systematic error. We studied two cases - in the first SCF fraction is 1.5% and in the second SCF fraction is 4.5%. In both the cases we observe an error of 1.7%.

d) Uncertainty on the signal $C_{NB}^{'}$ PDF shape: We use three bifurcated Gaussains to obtain $C_{NB}^{'}$ fit using the real data in the control mode. We study three cases - in all these cases two fractions between bifurcated Gaussians as well as four sigma ratios are fixed to the signal MC values.

- Case A: Both diff1 and diff2 are floated (Fig. 67),
- Case B: diff1 is fixed to the signal MC value while diff2 is floated (Fig. 63), and
- Case C: Both diff1 and diff2 are fixed to the signal MC values (Fig. 68).

Parameters in case B are used as nominal correction factors for $C_{NB}^{'}$ PDF, while we plan to use either case A or case C for the systematic error estimation by performing 2D ΔE - $C_{NB}^{'}$ fits. We observe an error of 1.7% and 2.3% for case A and case C, respectively. We decide to use the case C as a conservative estimate of the systematic error.

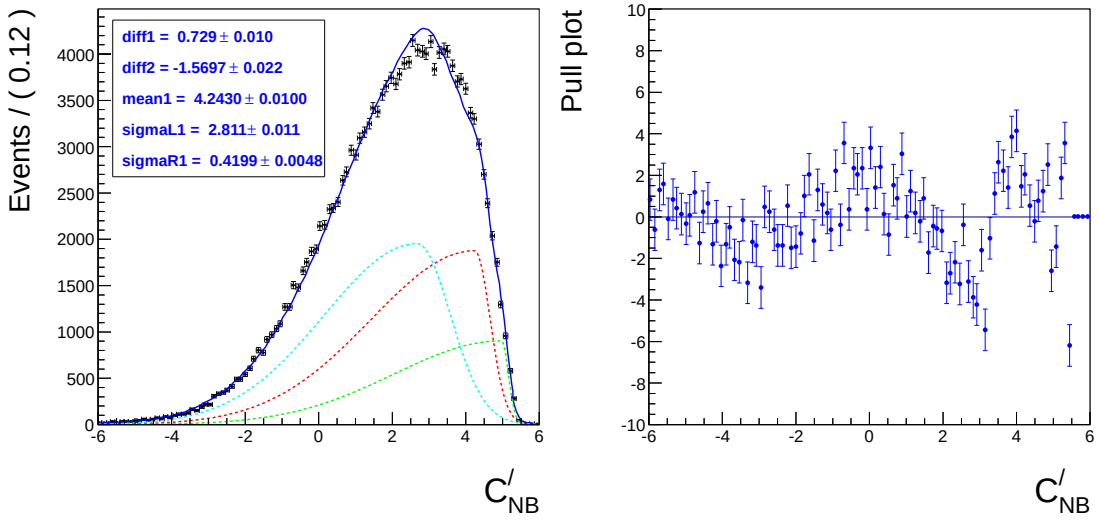


Figure 67: 1D $C_{NB}^{'}$ fit (left) and pull plot (right) in the ΔE signal region using the data with diff1 and diff2 floated.

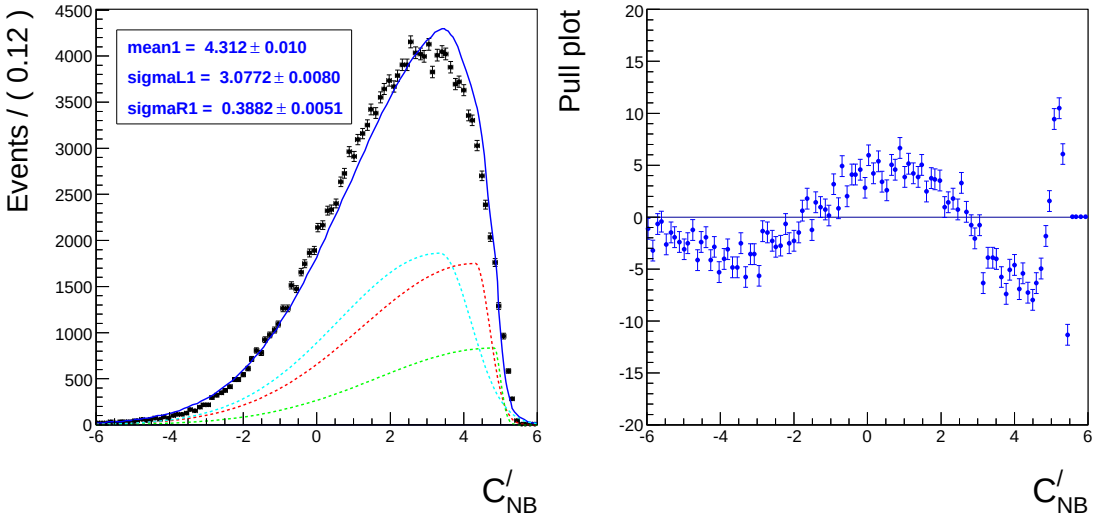


Figure 68: 1D $C_{NB}^{'}$ fit (left) and pull plot (right) in the ΔE signal region using the data with diff1 and diff2 fixed to the signal MC values.

e) Systematic error due to $q\bar{q}$ suppression requirement: We use the control sample to assign this systematic uncertainty. In our analysis a nominal cut of $C_{NB} > 0.2$ is implemented. For assigning systematic error, we compare two cases - case A: $C_{NB} > 0.2$ and case B: No C_{NB} cut. In both the cases, we reject ΔE region below $-0.02\,\text{GeV}$ to make the fit converge properly. This is particularly important when we fit on the data with no C_{NB} cut.

For case A: signal yield (A) and signal yield (A') are 49391 ± 368 and 224695 ± 1056, for the fit on data and MC, respectively

For case B: signal yield (B) and signal yield (B') are 53618 ± 536 and 239600 ± 1091, for the fit on data and MC, respectively

We calculate efficiencies, ε_{data}, ε_{MC} and the relative efficiency ε as:

- $\varepsilon_{data} = \frac{A}{B} = 0.9212 \pm 0.0115$
- $\varepsilon_{MC} = \frac{A'}{B'} = 0.9378 \pm 0.0061$
- $\varepsilon = \frac{\varepsilon_{data}}{\varepsilon_{MC}} = 0.98 \pm 0.01$

From the relative efficiency, we observe a small discrepancy of 2% between data and MC. We add the discrepancy and the error of 1% in quadrature and assign a systematic uncertainty of 2.2%.

f) Systematic error due to $M_{\rm bc}$ requirement: We use the control sample to assign this systematic uncertainty. In this analysis a nominal cut of $M_{\rm bc} > 5.271\,\text{GeV/c}^2$ is implemented. We first determine $\sigma(M_{\rm bc})$ using the control sample real data. The $\sigma(M_{\rm bc})$

value is $0.003\,\mathrm{GeV/c^2}$. We study two cases - case A: $M_{\mathrm{bc}} > 5.2695\,\mathrm{GeV/c^2}$ and case B: $M_{\mathrm{bc}} > 5.2725\,\mathrm{GeV/c^2}$. These two cases correspond to a M_{bc} nominal cut value $\pm$ half of $[\sigma(M_{\mathrm{bc}})]$.

For case A: signal yield, efficiency (from control MC) and branching fraction are 53856 ± 729, 11.73% and $(4.59 \pm 0.18) \times 10^{-3}$, respectively

For case B: signal yield, efficiency (from control MC) and branching fraction are 52501 ± 687, 11.25% and $(4.67 \pm 0.18) \times 10^{-3}$, respectively

We observe an error of 0.2% and 1.5%, respectively for case A and case B.

g) Systematic error due to histogram PDFs used in the fitter: We fluctuate bin contents by $\sqrt{N}$ and for every case we obtain the fitted signal yield and then we take the average. For rare combinatorial $B\bar{B}$ component, the yield of the largest contributing mode ($B^0 \to K^*(892)^+ K^*(892)^-$) is varied by changing its branching fraction by $\pm$50%. Finally we combine errors from the two source and assign a systematic error of $+1.7\%$ and -2.0%.

h) The systematic uncertainty due to π^0 reconstruction: It is evaluated by comparing data-MC differences of the yield ratio between $\eta \to \pi^0\pi^0\pi^0$ and $\eta \to \pi^+\pi^-\pi^0$. Systematic error due to π^0 detection efficiency is 4.0%.

i) Systematic error due to tracking: We use partially reconstructed $D^{*+} \to D^0(K^0_S\pi^+\pi^-)\pi^+$ decays to assign the systematic uncertainty due to charged-track reconstruction (0.35% per track). Thus for two tracks the uncertainty is 0.7%.

j) Systematic error from number of $B\bar{B}$ pairs is 1.37%

k) Systematic error due to efficiency variation over the Dalitz plot: We divide Dalitz plane into four regions:

- Region A: $[M^2_{K^+K^-} < 6\,\mathrm{GeV^2/c^4}]$ with fitted signal yield (N_1) and efficiency (ε_1): 139 ± 46 and 20.78%,
- Region B: $[M^2_{K^+K^-} > 6\,\mathrm{GeV^2/c^4} \& M^2_{K^+\pi^0} < 6\,\mathrm{GeV^2/c^4}]$ with fitted signal yield (N_2) and efficiency (ε_2): 95 ± 49 and 17.63%,

- Region C: $[M^2_{K^+K^-} > 6\,\mathrm{GeV}^2/\mathrm{c}^4 \& M^2_{K^+\pi^0} > 6\,\mathrm{GeV}^2/\mathrm{c}^4 \& M^2_{K^-\pi^0} < 6\,\mathrm{GeV}^2/\mathrm{c}^4]$ with fitted signal yield (N_3) and efficiency (ε_3): 18 ± 8 and 20.38%, and
- Region D: [Excluding above three regions] with zero fitted signal yield 21.30% efficiency

We use the information from the first three regions having a non-zero signal yield to obtain an average efficiency:

$$\varepsilon_{mean} = \frac{\sum_i N_i \times \varepsilon_i}{\sum_i N_i} = 19.6\% \tag{17}$$

and error on the average efficiency:

$$\Delta\varepsilon_{mean} = \frac{\sum_i N_i \times |\varepsilon_i - \varepsilon_{mean}|}{\sum_i N_i} = 1.5\% \tag{18}$$

where N_i and ε_i are the fitted signal yield and efficiency in the i-th cell of the Dalitz plot. From this we estimate a systematic error due to possible variation of efficiency across the Dalitz plot as $\Delta\varepsilon_{mean}/\varepsilon_{mean} = 7.5\%$. The average efficiency ($\varepsilon_{mean} = 19.6\%$) is used to calculate the branching fraction while the systematic error due to efficiency variation across the Dalitz plot is added in quadrature toward the total systematic error calculation.

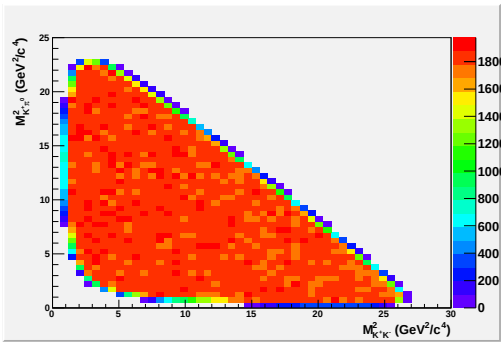


Figure 69: Dalitz plots using events at generation level.

l) PID: This uncertainty arises from kaon selection with the requirement $L(K/\pi) > 0.6$.

We divide the Dalitz plot in four regions as done for determining efficiency variation over the Dalitz plot.

For region A, PID correction and systematic error are 0.9199 and 1.9%. For region B, PID correction and systematic error are 0.9063 and 2.0%. For region C, PID correction and systematic error are 0.9099 and 1.7%. For region D, PID correction and systematic error are 0.9183 and 1.6%.

From Dalitz plot regions A, B and C we obtain following weighted averages (excluding the region D because the fitted signal yield in that region is zero):

PID correction (average) = 0.9141%, Systematic error (average) = 1.93%

The PID data-MC correction is almost constant over the Dalitz plane and we assign a systematic error of 1.9%.

We present a summary of various systematic errors in Table 14. In the branching fraction calculation, numerator has the signal yield while the denominator has terms related to number of $B\bar{B}$, efficiency and PID correction factor: the source of uncertainties listed in the first horizontal block influence the numerator part while the source of uncertainties listed in the second horizontal block influence the denominator part. Thus we have two horizontal blocks in Table 14.

Table 14: Summary of various systematic uncertainties.

Source	Uncertainities (%)	
Signal PDF	$+3.4$	-2.9
Generic $B\bar{B}$ PDF	$+2.4$	-3.1
Combinatorial background PDF	$+1.3$	-2.0
Peaking background PDFs	$+1.7$	-1.9
Fixed histogram PDF	$+1.7$	-2.0
Signal $C_{NB}^{'}$ shape	$+2.3$	-2.3
Fixed SCF fraction	$+1.7$	-1.7
Fit bias	$+2.4$	-2.4
Continuum suppression	$+2.2$	-2.2
Requirement on $M_{\rm bc}$	$+1.5$	-0.2
PID requirement	$+1.9$	-1.9
π^0 detection efficiency	$+4.0$	-4.0
Charged track reconstruction	$+0.7$	-0.7
Efficiency variation over the Dalitz plot	$+7.5$	-7.5
Number of $B\bar{B}$ pairs	$+1.4$	-1.4
Total	$+11.1$	-11.3

9.3 Signal significance after systematic uncertainty

In order to include the effect of systematics we convolve the statistical likelihood, presented in top-right Fig. 70, with a Gaussian function of width given by the additive systematic uncertainty. The resultant likelihood is presented in the bottom-left plot of Fig. 70. The bottom-right plot in Fig. 70 shows NLL (negative log of likelihood) distributions, with the red curve showing the convoluted histogram. We obtain a total significance of 3.5 standard deviations.

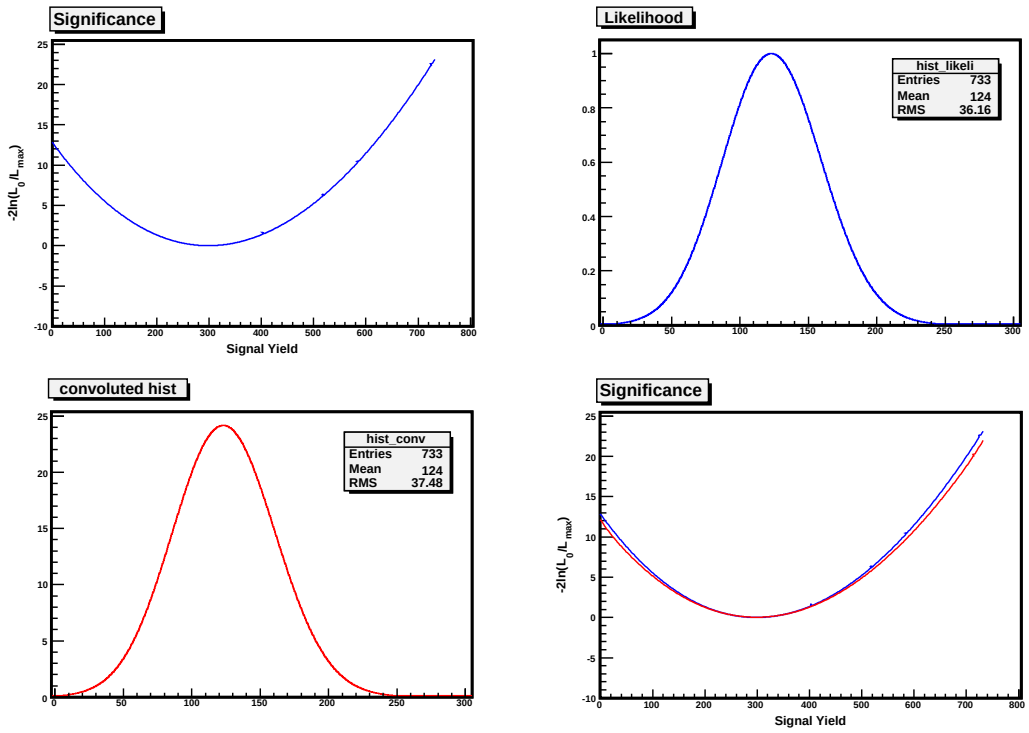


Figure 70: Projection plot of NLL and statistical likelihood; red curve shows convoluted histogram.

9.4 Branching fraction calculation

The decay branching fraction is calculated as:

$$\mathcal{B}(B^0 \to K^+K^-\pi^0) = \frac{n_{\rm sig}}{N_{B\bar{B}} \times \varepsilon_{\rm rec} \times r_{K/\pi}}, \tag{19}$$

where $N_{B\bar{B}}$ is the total number of $B\bar{B}$ pairs (772×10^6), $\varepsilon_{\rm rec}$ is the signal reconstruction efficiency (19.6%) and $r_{K/\pi}$ denotes the kaon-identification efficiency correction factor that accounts for a small data-MC difference. It is given by

$$r_{K/\pi} \equiv \varepsilon^{\rm data}_{K/\pi}/\varepsilon^{\rm MC}_{K/\pi}, \tag{20}$$

where $\varepsilon^{\rm data}_{K/\pi}$ ($\varepsilon^{\rm MC}_{K/\pi}$) is the efficiency of the $L(K/\pi)$ requirement in data (MC simulations). The $r_{K/\pi}$ value per kaon track is 0.95, resulting in a total $r_{K/\pi} = 0.95^2 = 0.90$ for two kaons.

The resulting branching fraction value is

$$\mathcal{B}(B^0 \to K^+K^-\pi^0) = [2.17 \pm 0.60 \pm 0.24] \times 10^{-6}, \tag{21}$$

where the uncertainties are statistical and systematic, respectively.

9.5 Study of Dalitz plot distributions

To elucidate the nature of the signal observed, especially whether there are contributions from the decays with intermediate resonant states, we study their K^+K^- and $K^+\pi^0$ invariant mass distributions. We perform the $[\Delta E, C'_{NB}]$ two-dimensional fit in various bins of $m(K^+K^-)$ and $m(K^+\pi^0)$ distributions without and after applying the requirements $m(K^+\pi^0) > 1.5\,{\rm GeV/c^2}$ and $m(K^+K^-) > 2.0\,{\rm GeV/c^2}$, respectively. These requirements are designed to suppress the $K^+K^-\pi^0$ nonresonant contribution. Figures 72 and 71 show the resulting signal yields with and without the requirements. It is evident that at the current level of statistics, we cannot make any definitive statement about possible resonance final states including the excess seen by BaBar near $1.4\,{\rm GeV/c^2}$.

9.6 Study of $B^0 \to \phi\pi^0$

We perform a 2D ΔE-C'_{NB} fit, with an additional requirement of $1.008\,{\rm GeV/c^2} < m_{K^+K^-} < 1.031\,{\rm GeV/c^2}$ on the K^+K^- invariant mass, using real data. The results are presented in Fig. 73.

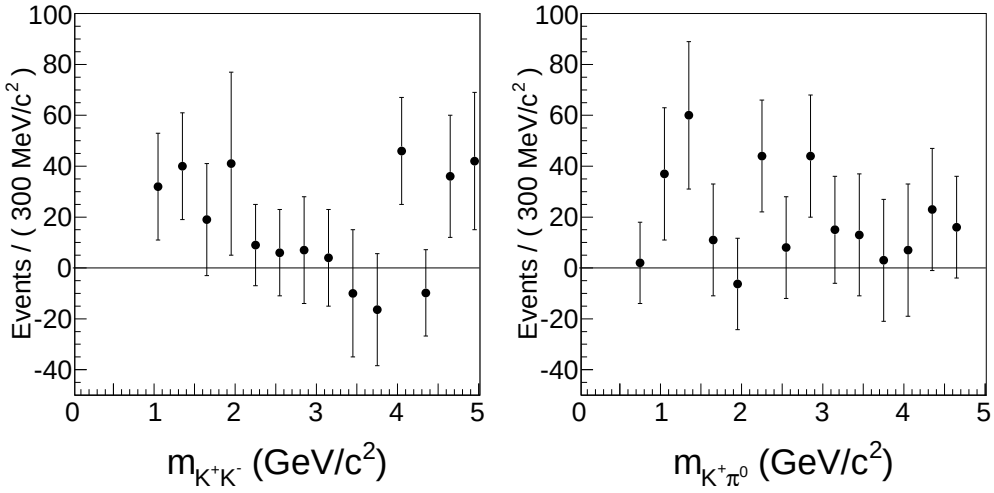


Figure 71: Distributions of the signal candidates in (left) $m(K^+K^-)$ and (right) $m(K^+\pi^0)$ distributions.

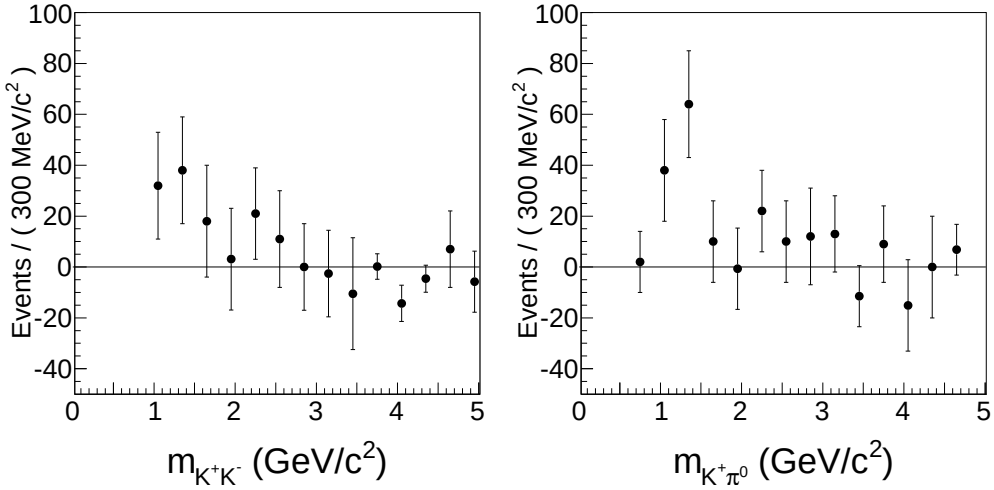


Figure 72: Distributions of the signal candidates in (left) $m(K^+K^-)$ with $m(K^+\pi^0) > 1.5\,\mathrm{GeV/c^2}$ and (right) $m(K^+\pi^0)$ with $m(K^+K^-) > 2.0\,\mathrm{GeV/c^2}$.

A signal yield of 12.0 ± 7.8 is obtained. In view of a small signal yield, we set an upper limit (UL) on $\mathcal{B}(B^0 \to \phi\pi^0)$. The signal efficiency is 10.3%. For obtaining the UL we convolve the statistical likelihood with a Gaussian function of width equal to the systematic error (that we know from $B^0 \to K^+K^-\pi^0$ analysis). Finally, we set a UL of 2.1×10^{-7} at 90% confidence level as shown in Fig. 74. We also calculate an expected UL (for a null signal hypothesis) to be 1.8×10^{-7}. Figure 75 shows the signal yield. The expected UL on the signal yield is 13.5 at 90% confidence level. Our UL results are comparable to the recent Belle results [22] on a dedicated analysis on $B^0 \to \phi\pi^0$.

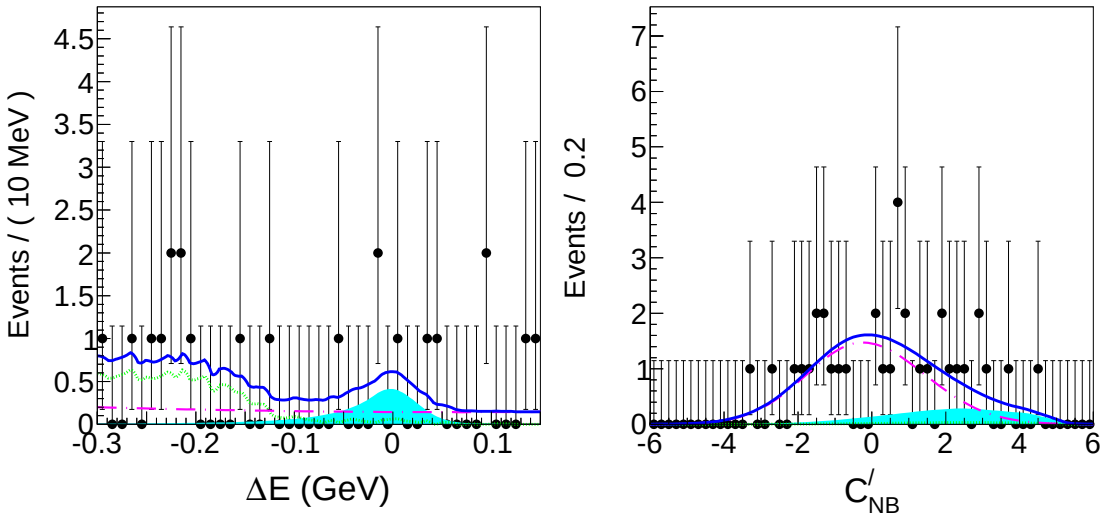


Figure 73: Projections of $B^0 \to \phi\pi^0$ candidate events onto (left) ΔE for $C'_{NB} > 3$ and (right) C'_{NB} for $|\Delta E| < 30\,\text{MeV}$. Points with error bars are the data, solid (blue) curves are the total PDF, dotted (green) curves are rare combinatorial backgrounds, dash-dotted (magenta) curves are the continuum $q\bar{q}$ background, and filled (cyan) regions show the signal.

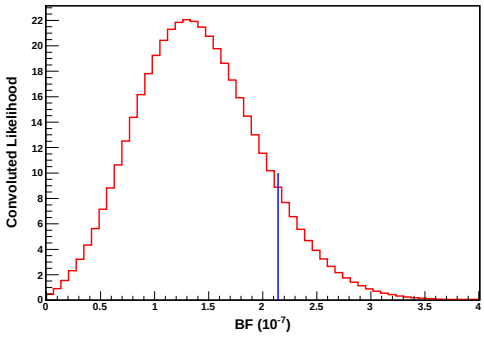


Figure 74: Convoluted likelihood vs branching fraction for $B^0 \to \phi\pi^0$ with the blue line showing the UL at 90% CL.

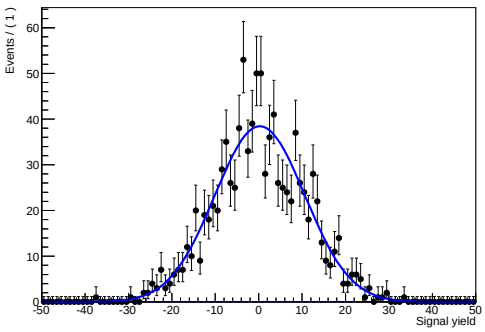


Figure 75: The fitted $B^0 \to \phi\pi^0$ signal yield obtained for a null signal hypothesis.

10 Conclusions

In summary, we report a measurement of the suppressed decay $B^0 \to K^+K^-\pi^0$ using the full $\Upsilon(4S)$ data sample collected with the Belle detector. We employ a two-dimensional fit for extracting the signal yield. Our measured branching fraction $\mathcal{B}(B^0 \to K^+K^-\pi^0) = [2.17 \pm 0.60(\text{stat}) \pm 0.24(\text{syst})] \times 10^{-6}$ constitutes the first evidence for the decay.

From the study of the K^+K^- and $K^+\pi^0$ invariant mass distributions we conclude that we cannot make any definitive statement about possible intermediate K^+K^- resonances, including the structure seen by BaBar near $1.5\,\text{GeV}/c^2$ [18]. It is worth noting here that the recent LHCb study of $B^\pm \to K^+K^-\pi^\pm$ decays [19] has revealed an unidentified structure in the same mass range; however, it is only present in B^+ events, giving rise to a large local *CP* asymmetry. Furthermore, we observe some excess of events around $1.4\,\text{GeV}/c^2$ in the $K^+\pi^0$ invariant-mass spectrum. A detailed interpretation will require an amplitude analysis with higher statistics that would be available at a next-generation flavor factory [62].

References

[1] S. Chatrchyan *et al.,* (CMS Collaboration) Phys. Lett. B **716**, 30-61 (2012).

[2] G. Aad *et al.,* (ATLAS Collaboration) Phys. Lett. B **716**, 1-29 (2012).

[3] N. Cabibbo Phys. Rev. Lett. **10**, 531-533 (1963).

[4] J. J. Aubert *et al.,* Phys. Rev. Lett. **33**, 1404-1406 (1974).

[5] J. E. Augustin *et al.,* Phys. Rev. Lett. **33**, 1406-1408 (1974).

[6] J. H. Christenson *et al.,* Phys. Rev. Lett. **13**, 138-140 (1964).

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

[8] S. W. Herb *et al.,* Phys. Rev. Lett. **39**, 252-255 (1977)

[9] F. Abe *et al.,* (CDF Collaboration) Phys. Rev. Lett. **74**, 2626-2631 (1995)

[10] S. Abachi *et al.,* (D0 Collaboration) Phys. Rev. Lett. **74**, 2422-2426 (1995)

[11] J. Beringer *et al.,* (Particle Data Group) Phys. Rev. D **86**, 010001 (2012)

[12] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945-1947 (1983).

[13] B. Aubert *et al.,* (BaBar Collaboration), Phys. Rev. D **75**, 012008 (2007).

[14] T. Aaltonen *et al.,* (CDF Collaboration), Phys. Rev. Lett. **108**, 211803 (2012).

[15] R. Aaij *et al.,* (LHCb Collaboration), J. High Energy Phys. **10**, 037 (2012).

[16] Y. T. Duh *et al.,* (Belle Collaboration), Phys. Rev. D **87**, 031103 (2013).

[17] B. Aubert *et al.,* (BaBar Collaboration), Phys. Rev. Lett. **97**, 171805 (2006).

[18] B. Aubert *et al.,* (BaBar Collaboration), Phys. Rev. Lett. **99**, 221801 (2007).

[19] R. Aaij *et al.,* (LHCb Collaboration), Presented at the HCP conference in Kyoto, Japan, LHCb-CONF-2012-028 (2012).

[20] Throughout this paper, inclusion of the charge-conjugate process is implied unless otherwise explicitly mentioned.

[21] E. Eckhart *et al.,* (CLEO Collaboration), Phys. Rev. Lett. **89**, 251801 (2002).

[22] J. H. Kim *et al.,* (Belle Collaboration), Phys. Rev. D **86**, 031101 (2012).

[23] S. Bar-Shalom, G. Eilam, and Y. D. Yang, Phys. Rev. D **67**, 014007 (2003).

[24] A. J. Buras and L. Silvestrini, Nucl. Phys. B **569** (2000) 3.

[25] Y. Grossman, Z. Ligeti, Y. Nir, and H. Quinn, Phys. Rev. D **68** 015004 (2003).

[26] D. S. Du, H. J. Gong, J. F. Sun, D. S. Yang, and G. H. Zhu, Phys. Rev. D **65**, 094025 (2002); [Erratum-ibid. D **66**, 079904 (2002)].

[27] M. Beneke and M. Neubert, Nucl. Phys. B **675**, 333 (2003).

[28] L. Guo, Q. G. Xu, and Z. J. Xiao, Phys. Rev. D **75**, 014019 (2007).

[29] C. W. Chiang, M. Gronau, Z. Luo, J. L. Rosner, and D. A. Suprun, Phys. Rev. D **69** 034001 (2004).

[30] J. P. Lees *et al.,* (BaBar Collaboration), Phys. Rev. D **85**, 112010 (2012).

[31] Y. Nakahama *et al.*, (Belle Collaboration), Phys. Rev. D **82**, 073011 (2010).

[32] B. Aubert *et al.,* [BaBar Collaboration], Phys. Rev. Lett. **99** 221801 (2007).

[33] N. Cabibbo, Phys. Rev. Lett. **10**, 531-533 (1963).

[34] S. Kurokawa and E. Kikutani, Nucl. Instrum. and Meth. A **499**, 1 (2003).

[35] N. Toge *et al., KEKB B− factory design report* KEK Report 95-7 (1995).

[36] M. T. Cheng *et al.*, (Belle Collaboration), Letter of Intent for a study of CP-violation in B meson decays, KEK Report 94-2 (1994).

[37] M. T. Cheng *et al.,* (Belle Collaboration), A study of CP-violation in B meson decays: Technical design report, KEK Report 95-1 (1995).

[38] A. Abashian *et al.,* (Belle Collaboration), Nucl. Instr. Meth. A **479**, 117 (2002).

[39] H. Albrecht *et al.,* (ARGUS Collaboration), Nucl. Instrum. and Meth. A **275**, 1-48 (1989).

[40] D. Andrews *et al.,* (CLEO Collaboration), Nucl. Instrum. and Meth. A **211**, 47-71 (1983).

[41] B. Aubert *et al.,* (BaBar Collaboration), Nucl. Instrum. and Meth. A **479**, 1-116 (2002).

[42] http://www-d0.fnal.gov/

[43] http://www-cdf.fnal.gov/

[44] http://home.web.cern.ch/about/experiments/atlas

[45] http://home.web.cern.ch/about/experiments/cms

[46] http://home.web.cern.ch/about/experiments/lhcb

[47] T. Bohringer *et al.,* (CUSB Collaboration) Phys. Rev. Lett. **44**, 1111-1114 (1980).

[48] G. Finocchiaro *et al.,* (CUSB Collaboration) Phys. Rev. Lett. **45**, 222-225 (1980).

[49] K. Akai *et al.,* RF systems for the KEK B-factory, KEK Preprint 2001-157C (2001).

[50] T. Abe *et al.,* (Belle Collaboration), arXiv:**0706.3248** [hep-ex].

[51] A. Abashian *et al.,* (BELLE Collaboration), Nucl. Instrum. and Meth. A **479**, 117 (2002).

[52] O. Toker *et al.,* Nucl. Instrum. and Meth. A **340**, 572 (1994).

[53] Y. Ushiroda *et al.,* (BELLE Collaboration), Nucl. Instrum. and Meth. A **438**, 460-471 (1999).

[54] D. J. Lange, Nucl. Instrum. and Meth. A **462**, 152 (2001).

[55] R. Brun *et al.,* CERN Report DD/EE/84-1, 1984.

[56] J. Beringer *et al.,* (Particle Data Group), Phys. Rev. D **86**, 010001 (2012).

[57] R. Fruhwirth Nucl. Instrum. Meth. A **262** (1987), 444-450

[58] M. Feindt and U. Kerzel, Nucl. Instrum. Methods Phys. Res., Sect. A **559**, 190 (2006).

[59] S. H. Lee *et al.,* (Belle Collaboration), Phys. Rev. Lett. **91**, 261801 (2003).

[60] Y. Amhis *et al.,* (Heavy Flavor Averaging Group), arXiv:1207.1158 [hep-ex].

[61] T. Skwarnicki, DESY Report No. F31-86-02 (1986), Appendix E.

[62] T. Abe *et al.,* (Belle II Collaboration), arXiv:1011.0352 [physics.ins-det].

A Appendix

A.1 Performance of NeuroBayes

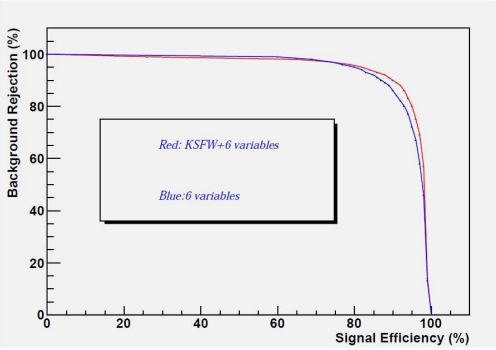


Figure 76: Background rejection vs signal efficiency

In addition to LR[RooKSFW] other six variables are:

- cosb: cosine of the angle (θ) between the B momentum and z axis,
- costhr: the cosine of the angle between the thrust axis of the B candidate and that of the rest of the event,
- cosbt: cosine of the angle between the B thrust and z axis,
- R2: ratio of 2^{nd} and 0^{th} order Fox-Wolfram moments,
- Δz: the vertex separation along the z axis between the B candidate and the remaining tracks, and
- qr: for flavor tagging information. It was finally excluded from the network to retain the possibility of performing A_{CP} measurement.

A.2 PDF shapes used in the analysis

A.2.1 Chebyshev polynomial of the first kind

The Chebyshev polynomials of the first kind are denoted $T_n(x)$. They are used as an approximation to a least squares fit. The first few polynomials are illustrated in Fig. 77 for $x \in [-1,1]$ and $n = 1,2,...,5$. The use of Chebychev polynomials over regular polynomials is recommended because of their superior stability in fits. Chebychev polynomials and regular polynomials can describe the same shapes, but a clever re-organization of power terms in Chebychev polynomials results in much lower correlations between the coefficients in a fit, and thus to a more stable fit behavior.

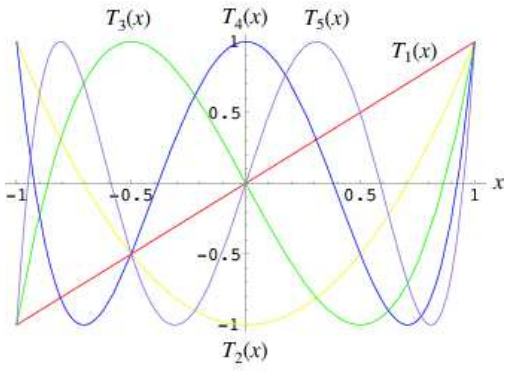


Figure 77: The first few Chebyshev polynomials.

The Chebyshev polynomial of the first kind $T_n(z)$ can be defined by the contour integral:

$$Tn(z) = \frac{1}{4\pi i} \oint \frac{(1-t^2)t^{-n-1}}{(1-2tz+t^2)} dt,$$

where the contour encloses the origin and is traversed in a counterclockwise direction.

The first few Chebyshev polynomials of the first kind are

$$T_0(x) = 1$$
$$T_1(x) = x$$
$$T_2(x) = 2x^2 - 1$$
$$T_3(x) = 4x^3 - 3x$$
$$T_4(x) = 8x^4 - 8x^2 + 1$$
$$T_5(x) = 16x^5 - 20x^3 + 5x$$

A.2.2 Crystal Ball function

It is a probability density function commonly used to model various lossy processes in high-energy physics. It consists of a Gaussian core portion and a power-law low-end tail, below a certain threshold. The function itself and its first derivative are both continuous. Figure 78 shows examples of the Crystal Ball function.

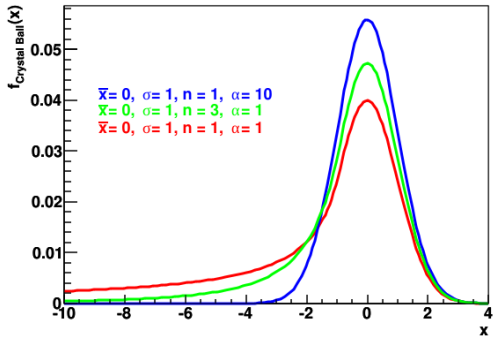


Figure 78: Examples of the Crystal Ball function.

The Crystal Ball function is given by:

$$f(x;\alpha,n,\bar{x},\sigma) = N.exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right), \quad for\ \frac{x-\bar{x}}{\sigma} > -\alpha$$
$$f(x;\alpha,n,\bar{x},\sigma) = N.A\left(B-\frac{x-\bar{x}}{\sigma}\right)^{-n}, \quad for\ \frac{x-\bar{x}}{\sigma} \leq -\alpha$$

where

$$A = \left(\frac{n}{|\alpha|}\right)^n .exp\left(-\frac{|\alpha|^2}{2}\right),$$
$$B = \frac{n}{|\alpha|} - |\alpha|,$$
$$N = \frac{1}{\sigma(C+D)},$$
$$C = \frac{n}{|\alpha|}.\frac{1}{n-1}.exp\left(-\frac{|\alpha|^2}{2}\right),$$
$$D = \sqrt{(\pi/2)}.\left(1+erf\left(\frac{|\alpha|}{\sqrt{2}}\right)\right)$$

N is a normalization factor and $\alpha, n, \bar{x}, \sigma$ are parameters which are fitted with the data. erf is the error function.

A.2.3 Asymmetric Gaussian

It is a variation on the Gaussian where the width of the Gaussian on the low and high side of the mean can be set independently. The asymmetric Gaussian is written as:

$$f(x) = \frac{C}{\sqrt{2}\pi\sigma}\left([1-\theta(x-<x>)]exp\left(\frac{-(x-<x>)^2}{2\sigma_L^2}\right)+\theta(x-<x>)exp\left(\frac{-(x-<x>)^2}{2\sigma_R^2}\right)\right)$$

where $\theta(u)$ is a step function:

$$\theta(u) = 0 \quad for \;\; u < 0$$
$$\theta(u) = 1 \quad for \;\; u > 0$$

and $<x>$ is the peak of asymmetric Gaussian, σ is the average RMS width: $\sigma = \frac{\sigma_L+\sigma_R}{2}$, and C is the area under curve (the distribution is normalized, see below).

The distribution is normalized:

$$\int_{-\infty}^{\infty} f(x)dx = \frac{C}{\sigma}\left(\frac{1}{\sqrt{2}\pi}\int_{-\infty}^{0} exp\left(\frac{-(x-<x>)^2}{2\sigma_L^2}\right)dx + \frac{1}{\sqrt{2}\pi}\int_{0}^{\infty} exp\left(\frac{-(x-<x>)^2}{2\sigma_H^2}\right)dx\right),$$

$$\int_{-\infty}^{\infty} f(x)dx == \frac{C}{\sigma}\left(\frac{\sigma_L}{2}+\frac{\sigma_H}{2}\right) = C$$

A.3 ΔE-M_{bc} correlations for various MC samples

Except for the signal MC, we do not observe a strong correlation between ΔE and M_{bc} for various MC samples.

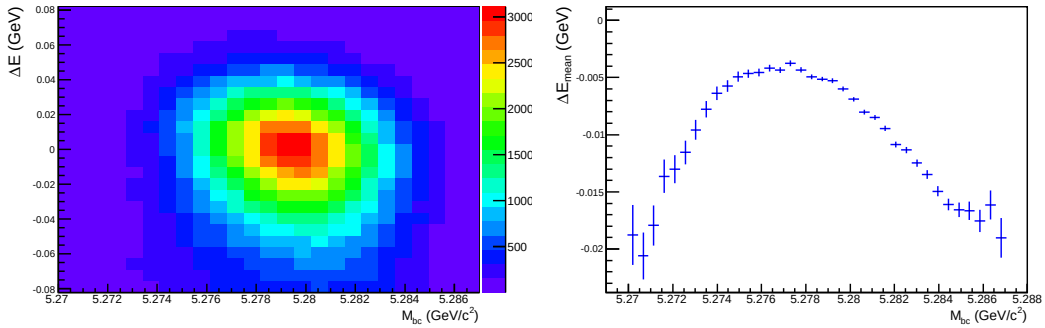


Figure 79: (left) ΔE vs M_{bc} and (right) the ΔE mean in bins of M_{bc} for signal MC. The correlation factor is -6.93%.

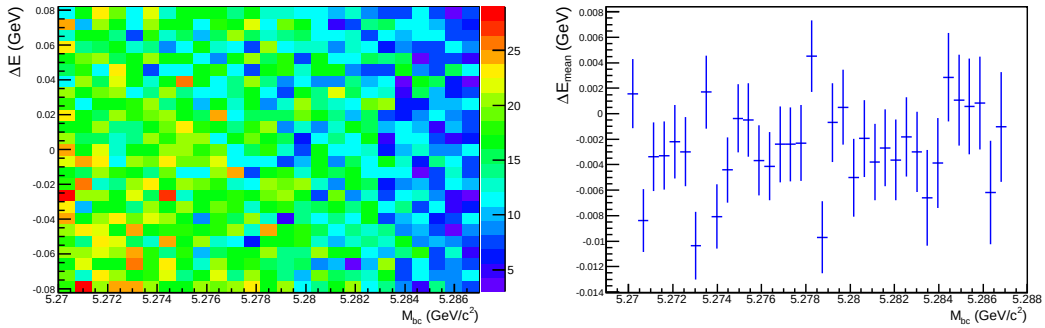


Figure 80: (left) ΔE vs M_{bc} and (right) the ΔE mean in bins of M_{bc} for $q\bar{q}$ MC. The correlation factor is $+1.48\%$.

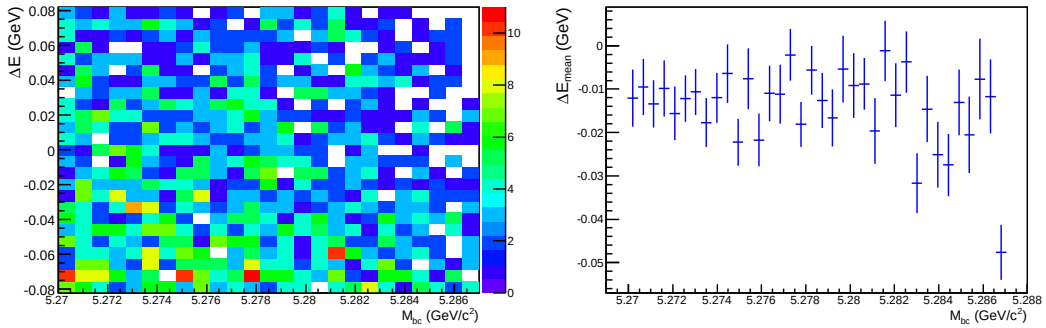


Figure 81: (left) ΔE vs M_{bc} and (right) the ΔE mean in bins of M_{bc} for generic $B\bar{B}$ MC. The correlation factor is -3.09%.

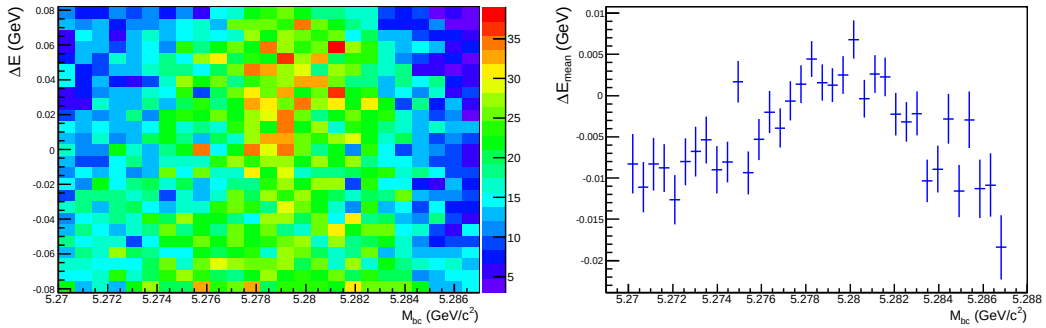


Figure 82: (left) ΔE vs M_{bc} and (right) the ΔE mean in bins of M_{bc} for the combinatorial component of rare $B\bar{B}$ MC. The correlation factor is $+2.05\%$.

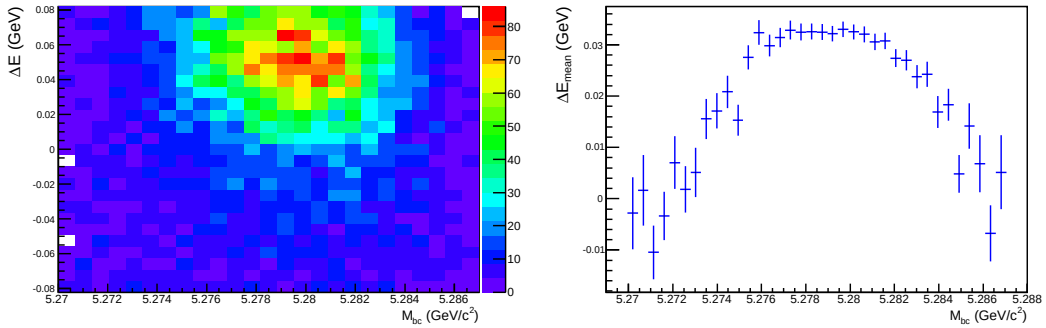


Figure 83: (left) ΔE vs M_{bc} and (right) the ΔE mean in bins of M_{bc} for the peaking component of rare $B\bar{B}$ MC. The correlation factor is $+4.07\%$.

A.4 ΔE distributions in different M_{bc} regions for various MC samples

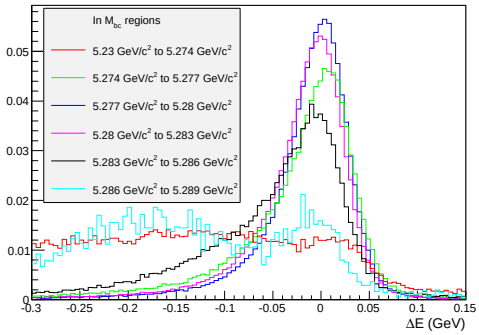


Figure 84: ΔE distribution in different M_{bc} regions for truth-matched component of signal MC

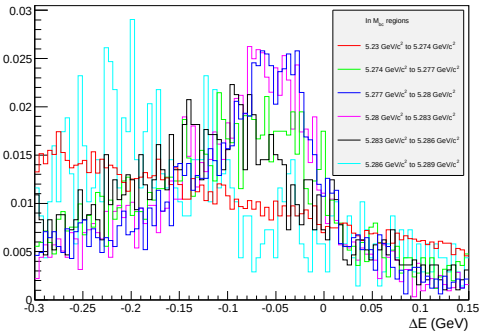


Figure 85: ΔE distribution in different M_{bc} regions for SCF component of signal MC

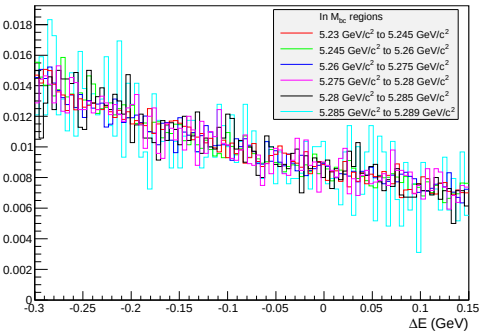


Figure 86: ΔE distribution in different M_{bc} regions for $q\bar{q}$ MC

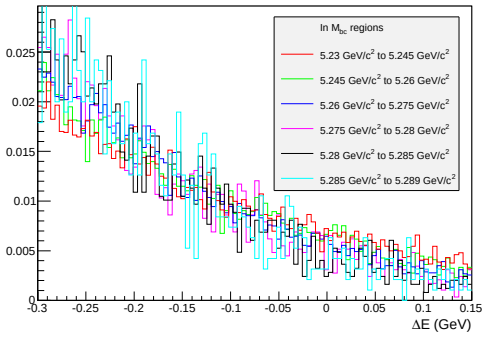


Figure 87: ΔE distribution in different M_{bc} regions for generic $B\bar{B}$ MC

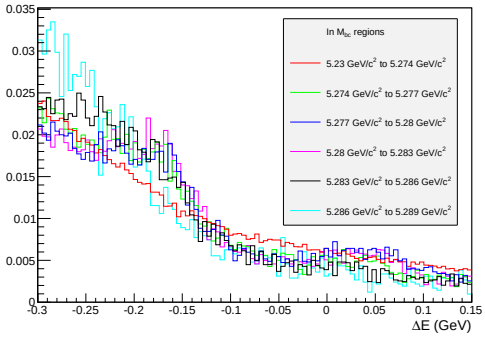


Figure 88: ΔE distribution in different M_{bc} regions for combinatorial component of rare MC

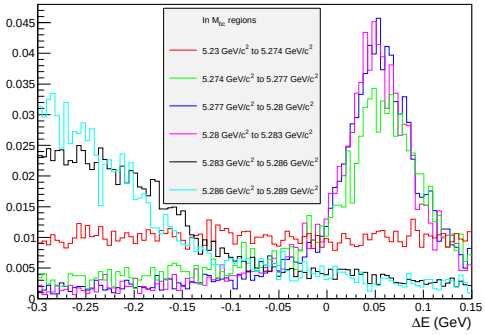


Figure 89: ΔE distribution in different M_{bc} regions for peaking component of rare MC

A.5 $C_{NB}^{'}$-ΔE correlations for various MC samples

We do not observe a strong correlation between $C_{NB}^{'}$ and ΔE for various MC samples.

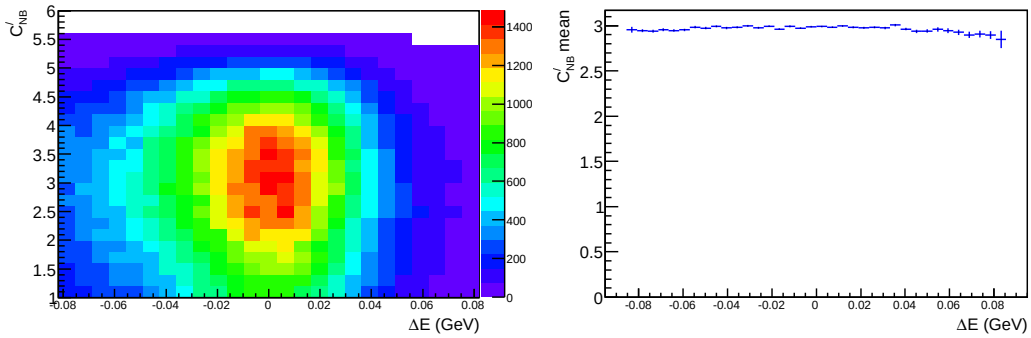


Figure 90: (left) $C_{NB}^{'}$ vs ΔE and (right) the $C_{NB}^{'}$ mean in bins of ΔE for signal MC. The correlation factor is $+0.06\%$.

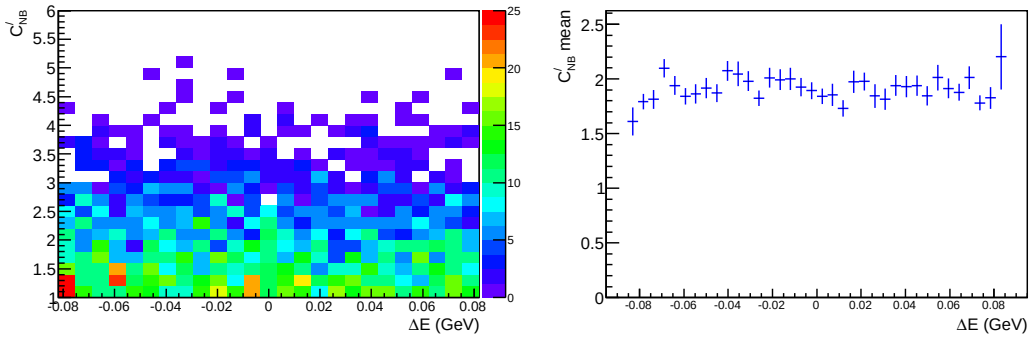


Figure 91: (left) $C_{NB}^{'}$ vs ΔE and (right) the $C_{NB}^{'}$ mean in bins of ΔE for $q\bar{q}$ MC. The correlation factor is -0.42%.

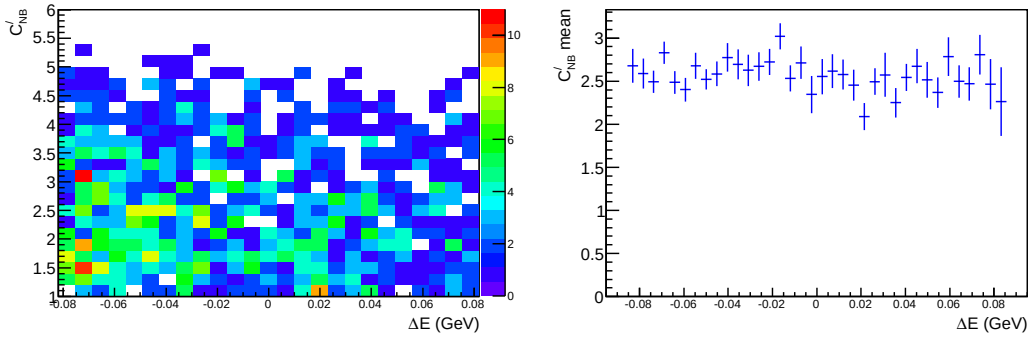


Figure 92: (left) $C_{NB}^{'}$ vs ΔE and (right) the $C_{NB}^{'}$ mean in bins of ΔE for generic $B\bar{B}$ MC. The correlation factor is -4.21%.

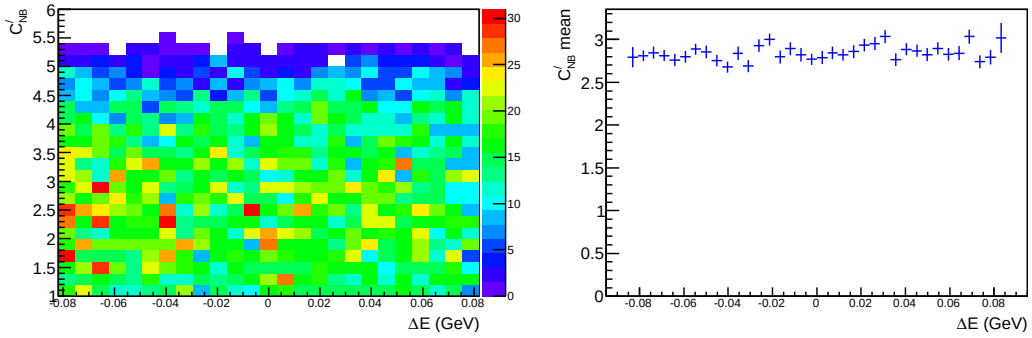


Figure 93: (left) $C_{NB}^{'}$ vs ΔE and (right) the $C_{NB}^{'}$ mean in bins of ΔE for the combinatorial component of rare $B\bar{B}$ MC. The correlation factor is $+2.17\%$.

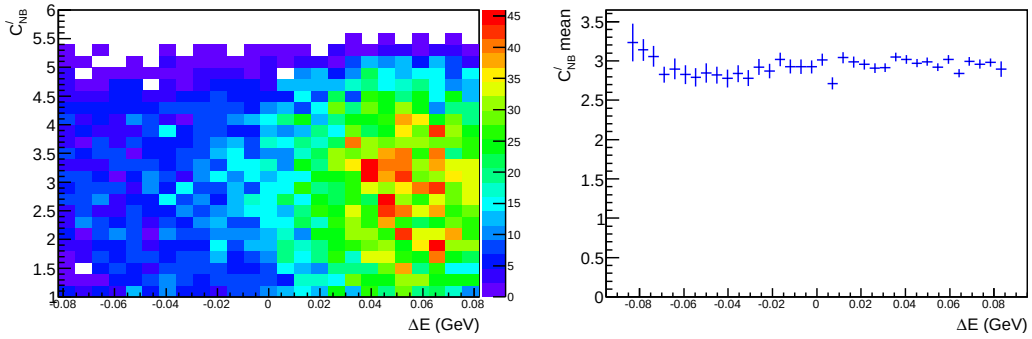


Figure 94: (left) $C_{NB}^{'}$ vs ΔE and (right) the $C_{NB}^{'}$ mean in bins of ΔE for the peaking component of rare $B\bar{B}$ MC. The correlation factor is $+2.14\%$

A.6 ΔE distributions in different $C_{NB}^{'}$ regions for various MC samples

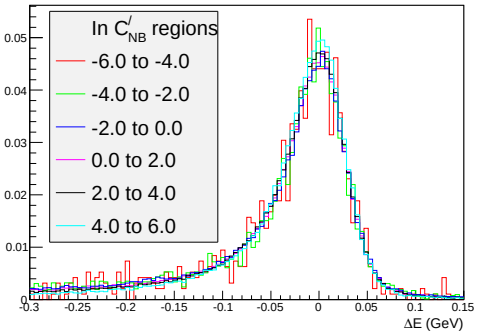


Figure 95: ΔE distribution in different $C_{NB}^{'}$ regions for truth-matched component of signal MC

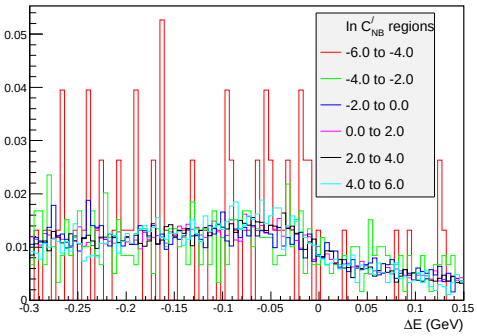


Figure 96: ΔE distribution in different $C_{NB}^{'}$ regions for SCF component of signal MC

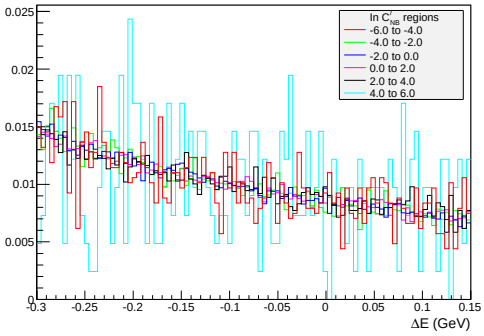


Figure 97: ΔE distribution in different $C_{NB}^{'}$ regions for $q\bar{q}$ MC

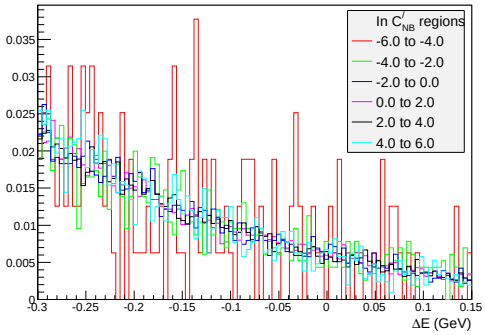


Figure 98: ΔE distribution in different $C_{NB}^{'}$ regions for generic $B\bar{B}$ MC

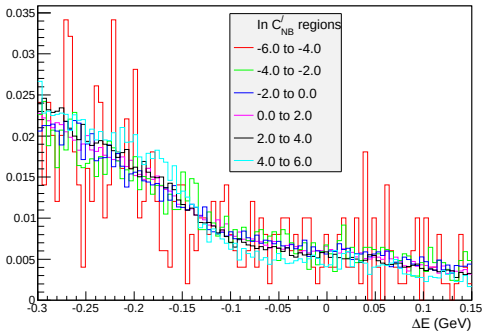


Figure 99: ΔE distribution in different $C_{NB}^{'}$ regions for combinatorial component of rare MC

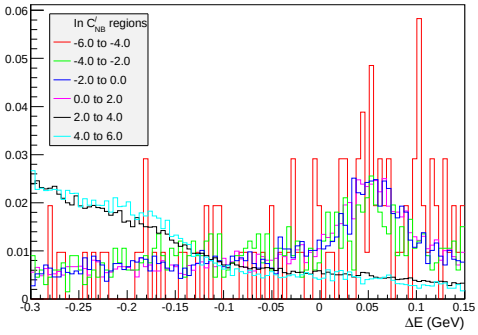


Figure 100: ΔE distribution in different $C_{NB}^{'}$ regions for peaking component of rare MC

A.7 $C_{NB}^{'}$-M_{bc} correlations for various MC samples

We did not observe a strong correlation between $C_{NB}^{'}$ and M_{bc} for various MC samples.

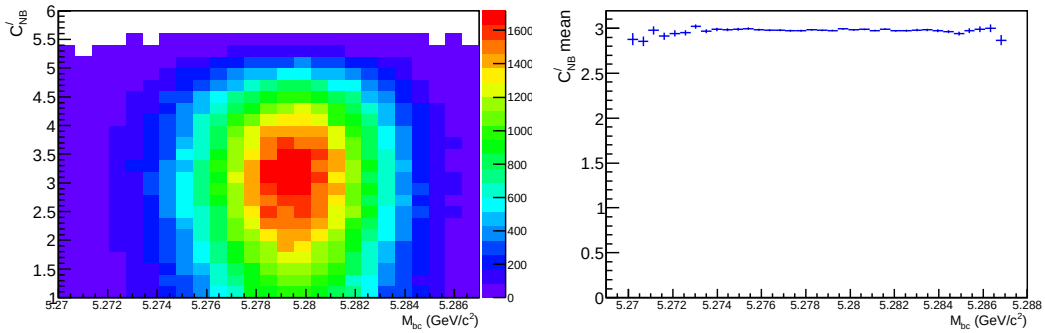


Figure 101: (left) $C_{NB}^{'}$ vs M_{bc} and (right) the $C_{NB}^{'}$ mean in bins of M_{bc} for signal MC. The correlation factor is -0.01%.

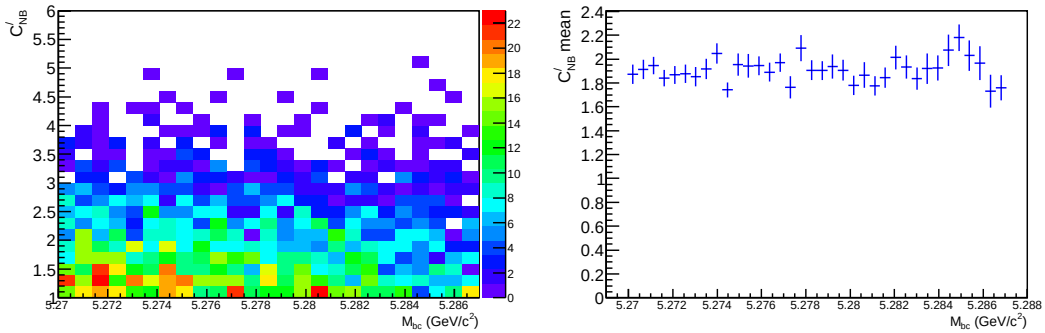


Figure 102: (left) $C_{NB}^{'}$ vs M_{bc} and (right) the $C_{NB}^{'}$ mean in bins of M_{bc} for $q\overline{q}$ MC. The correlation factor is $+1.30\%$.

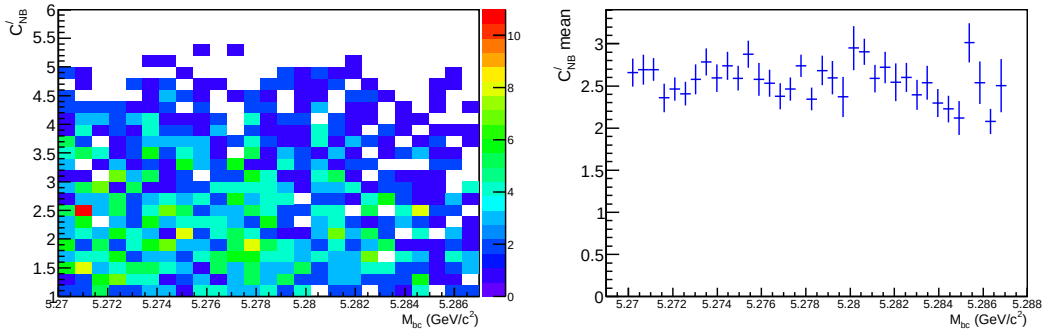


Figure 103: (left) $C_{NB}^{'}$ vs M_{bc} and (right) the $C_{NB}^{'}$ mean in bins of M_{bc} for generic $B\bar{B}$ MC. The correlation factor is -3.53%.

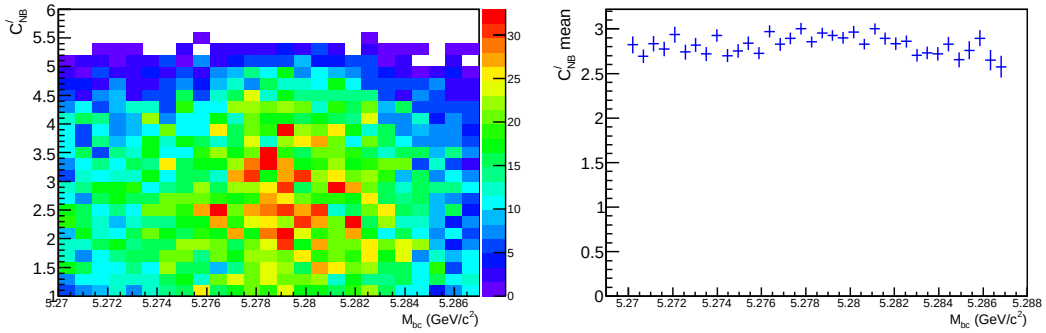


Figure 104: (left) $C_{NB}^{'}$ vs M_{bc} and (right) the $C_{NB}^{'}$ mean in bins of M_{bc} for the combinatorial component of rare $B\bar{B}$ MC. The correlation factor is -0.33%.

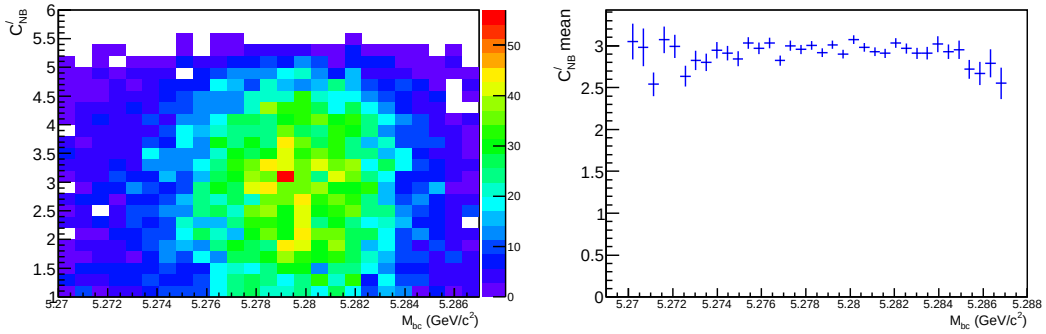


Figure 105: (left) $C_{NB}^{'}$ vs M_{bc} and (right) the $C_{NB}^{'}$ mean in bins of M_{bc} for the peaking component of rare $B\bar{B}$ MC. The correlation factor is $+0.42\%$.

A.8 M_{bc} distributions in different $C_{NB}^{'}$ regions for various MC samples

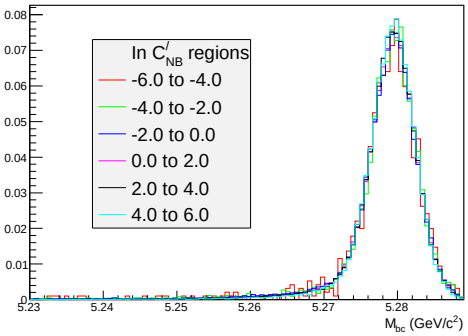


Figure 106: M_{bc} distribution in different $C_{NB}^{'}$ regions for truth-matched component of signal MC

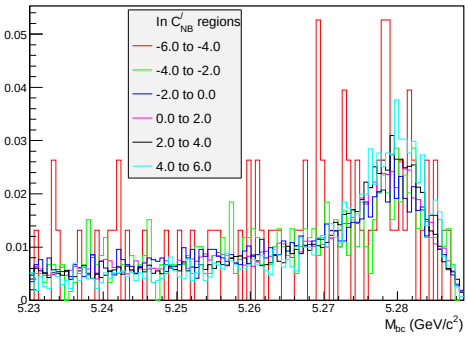


Figure 107: M_{bc} distribution in different $C_{NB}^{'}$ regions for SCF component of signal MC

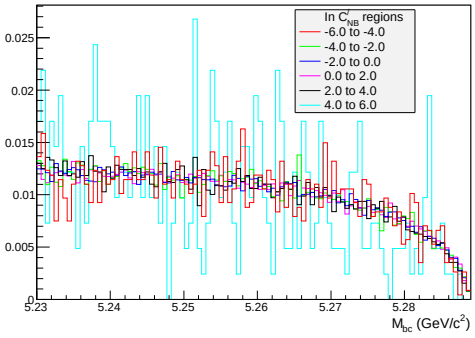


Figure 108: M_{bc} distribution in different $C_{NB}^{'}$ regions for $q\bar{q}$ MC

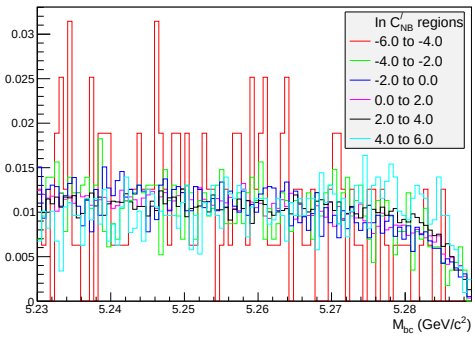


Figure 109: M_{bc} distribution in different C'_{NB} regions for generic $B\bar{B}$ MC

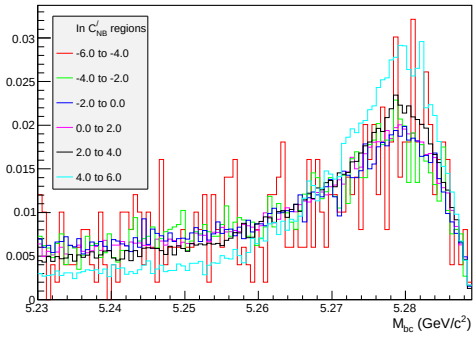


Figure 110: M_{bc} distribution in different C'_{NB} regions for combinatorial component of rare MC

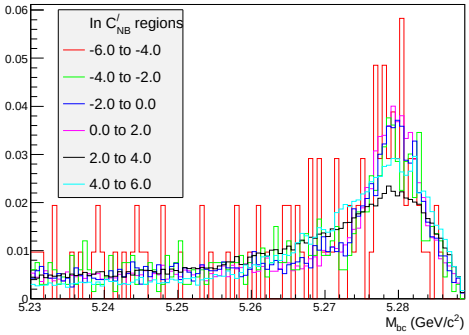


Figure 111: M_{bc} distribution in different C'_{NB} regions for peaking component of rare MC

A.9 Distribution of CR and SCF signal events over Dalitz plot

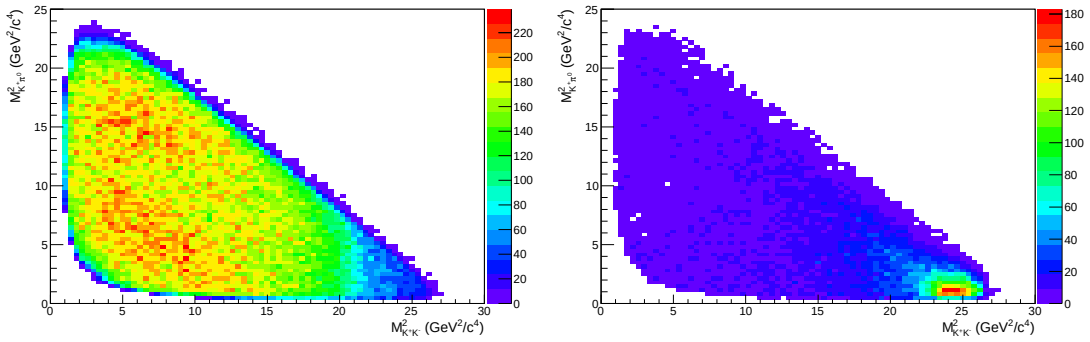


Figure 112: Distribution of CR (left) and SCF (right) signal events over Dalitz plot.

A.10 List of Publications

I Evidence for the decay $B^0 \to K^+K^-\pi^0$
V. Gaur *et al.,* (Belle Collaboration)
Phys. Rev. D **87**, 091101 (2013)

II Evidence for $B^- \to \tau^-\bar{\nu_\tau}$ with a Hadronic Tagging Method Using the Full Data Sample of Belle
K. Hara *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **110**, 131801 (2013)

III Evidence for the $\eta_b(2S)$ and Observation of $h_b(1P) \to \eta_b(1S)\gamma$ and $h_b(2P) \to \eta_b(1S)\gamma$
R. Mizuk *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **109**, 232002 (2012)

IV Evidence for CP Violation in the Decay $D^+ \to K^0_S\pi^+$
B. R. Ko *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **109**, 021601 (2012)

V Observation of New Resonant Structures in $\gamma\gamma \to \omega\phi, \phi\phi$, and $\omega\omega$
Z. Q. Liu *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **108**, 232001 (2012)

VI First Observation of $B^0_s \to J/\psi\,\eta$ and $B^0_s \to J/\psi\,\eta^/$
J. Li *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **108**, 181808 (2012)

VII Precise Measurement of the *CP* Violation Parameter $\sin2\phi_1$ in $B^0 \to (c\bar{c})K^0$ Decays
I. Adachi *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **108**, 171802 (2012)

VIII Measurement of the *CP*-Violation Parameter $\sin2\phi_1$ with a New Tagging Method at the $\Upsilon(5S)$ Resonance
Y. Sato *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **108**, 171801 (2012)

IX Observation of Two Charged Bottomoniumlike Resonances in $\Upsilon(5S)$ Decays
A. Bondar *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **108**, 122001 (2012)

X Search for *CP* Violation in $D^{\pm}$ Meson Decays to $\phi\pi^{\pm}$
M. Staric *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **108**, 071801 (2012)

XI Evidence for Direct *CP* Violation in $B^{\pm} \to \eta\, h^{\pm}$ and Observation of $B^0 \to \eta\, K^0$
C. T. Hoi *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **108**, 031801 (2012)

XII First Observation of the *P*-Wave Spin-Singlet Bottomonium States $h_b(1P)$ and $h_b(2P)$
I. Adachi *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **108**, 032001 (2012)

XIII Observation of $D^+ \to K^+\eta^{(/)}$ and Search for *CP* Violation in $D^+ \to \pi^+\eta^{(/)}$ Decays
E. Won *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **107**, 221801 (2011)

XIV Observation of $X(3872) \to J/\psi\,\gamma$ and Search for $X(3872) \to \psi^{/}\,\gamma$ in B Decays
V. Bhardwaj *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **107**, 091803 (2011)

XV Observation of Transverse Polarization Asymmetries of Charged Pion Pairs in e^+e^- Annihilation near $\sqrt{s} = 10.58$ GeV
A. Vossen *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **107**, 072004 (2011)

XVI Evidence for the Suppressed Decay $B^- \to DK^-$, $D \to K^+\pi^-$
Y. Horii *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **106**, 231803 (2011)

XVII Observation of $B_s^0 \to J/\psi\ f_0(980)$ and Evidence $B_s^0 \to J/\psi\ f_0(1370)$
J. Li *et al.,* (Belle Collaboration)
Phys. Rev. Lett. **106**, 121802 (2011)

XVIII Measurement of the inclusive semileptonic branching fraction $\mathcal{B}(B_s^0 \to X^- l^+ \nu_l)$ at Belle
C. Oswald *et al.,* (Belle Collaboration)
Phys. Rev. D **87**,072008 (2013)

XIX Search for heavy neutrinos at Belle
D. Liventsev *et al.,* (Belle Collaboration)
Phys. Rev. D **87**, 071102 (2013)

XX Observation of $\psi(4040)$ and $\psi(4160)$ decay into $\eta\, J/\psi$
X. L. Wang *et al.,* (Belle Collaboration)
Phys. Rev. D **87**, 051101 (2013)

XXI Measurements of branching fractions and direct *CP* asymmetries for $B \to K\pi$, $B \to \pi\pi$ and $B \to KK$ decays
Y. T. Duh *et al.,* (Belle Collaboration)
Phys. Rev. D **87**, 031103 (2013)

XXII Precise measurement of the branching fractions for $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ and first measurement of the $D_s^{*+} D_s^{*-}$ polarization using e^+e^- collisions
S. Esen *et al.,* (Belle Collaboration)
Phys. Rev. D **87**, 031101 (2013)

XXIII Study of the hadronic transitions $\Upsilon(2S) \to (\eta, \pi^0)\Upsilon(1S)$ at Belle
U. Tamponi *et al.,* (Belle Collaboration)
Phys. Rev. D **87**, 011104 (2013)

XXIV Measurement of branching fraction and first evidence of *CP* violation in $B^0 \to a_1^{\pm}(1260)\pi^{\mp}$ decays
J. Dalseno *et al.,* (Belle Collaboration)
Phys. Rev. D **86**, 092012 (2012)

XXV Measurement of $\gamma\gamma^* \to \pi^0$ transition form factor at Belle
S. Uehara *et al.,* (Belle Collaboration)
Phys. Rev. D **86**, 092007 (2012)

XXVI Evidence for $B^- \to D_s^+ K^- l^- \bar{\nu}_l$ and search for $B^- \to D_s^{*+} K^- l^- \bar{\nu}_l$
J. Stypula *et al.,* (Belle Collaboration)
Phys. Rev. D **86**, 072007 (2012)

XXVII First observation of *CP* violation and improved measurement of the branching fraction and polarization of $B^0 \to D^{*+} D^{*-}$ decays
B. Kronenbitter *et al.,* (Belle Collaboration)
Phys. Rev. D **86**, 071103 (2012)

XXVIII First study of $\eta_c(1S)$, $\eta(1760)$ and $X(1835)$ production via $\eta^{/}\pi^+\pi^-$ final states in two-photon collisions
C. C. Zhang *et al.,* (Belle Collaboration)
Phys. Rev. D **86**, 052002 (2012)

XXIX Observation of exclusive $\Upsilon(1S)$ and $\Upsilon(2S)$ decays into light hadrons
C. P. Shen *et al.*, (Belle Collaboration)
Phys. Rev. D **86**, 031102 (2012)

XXX Search for $B \to \phi\pi$ decays
J. H. Kim *et al.*, (Belle Collaboration)
Phys. Rev. D **86**, 031101 (2012)

XXXI Search for B^0 decays to invisible final states at Belle
C. L. Hsu *et al.*, (Belle Collaboration)
Phys. Rev. D **86**, 032002 (2012)

XXXII Search for the decay $B^0 \to DK^{*0}$ followed by $D \to K^-\pi^+$
K. Negishi *et al.*, (Belle Collaboration)
Phys. Rev. D **86**, 011101 (2012)

XXXIII First measurement of ϕ_3 with a model-independent Dalitz plot analysis of $B^\pm \to DK^\pm$, $D \to K^0_S\pi^+\pi^-$ decay
H. Aihara *et al.*, (Belle Collaboration)
Phys. Rev. D **85**, 112014 (2012)

XXXIV Measurements of branching fractions and time-dependent *CP* violating asymmetries in $B^0 \to D^{*\pm}D^\mp$ decays
M. Rohrken *et al.*, (Belle Collaboration)
Phys. Rev. D **85**, 091106 (2012)

XXXV Measurement of $B^0 \to J/\psi\,\eta^/$ and constraint on the $\eta - \eta^/$ mixing angle
M. C. Chang *et al.*, (Belle Collaboration)
Phys. Rev. D **85**, 091102 (2012)

XXXVI Search for time-dependent CPT violation in hadronic and semileptonic B decays
T. Higuchi *et al.*, (Belle Collaboration)
Phys. Rev. D **85**, 071105 (2012)

XXXVII Search for double charmonium decays of the P-wave spin-triplet bottomonium states
C. P. Shen *et al.*, (Belle Collaboration)
Phys. Rev. D **85**, 071102 (2012)

XXXVIII Search for charmonium and charmoniumlike states in $\Upsilon(2S)$ radiative decays
X. L. Wang *et al.*, (Belle Collaboration)
Phys. Rev. D **84**, 071107 (2011)

XXXIX Search for lepton-number-violating $B^+ \to D^- l^+ l^+$ decays
O. Seon *et al.,* (Belle Collaboration)
Phys. Rev. D **84**, 071106 (2011)

XL Observation of $B^- \to \bar{p}\Lambda D^0$ at Belle
P. Chen *et al.,* (Belle Collaboration)
Phys. Rev. D **84**, 071501 (2011)

XLI First observation of radiative $B^0 \to \phi K^0 \gamma$ decays and measurements of their time-dependent *CP* violation
H. Sahoo *et al.,* (Belle Collaboration)
Phys. Rev. D **84**, 071101 (2011)

XLII Bounds on the width, mass difference and other properties of $X(3872) \to \pi^+\pi^- J/\psi$ decays
S. K. Choi *et al.,* (Belle Collaboration)
Phys. Rev. D **84**, 052004 (2011)

XLIII Study of the decays $B \to D_{s1}(2536)^+ \bar{D}^*$
T. Aushev *et al.,* (Belle Collaboration)
Phys. Rev. D **83**, 051102 (2011)

XLIV Search for lepton-flavor and lepton-number-violating $\tau \to lhh$ decay modes
Y. Miyazaki *et al.,* (Belle Collaboration)
Phys. Lett. B **719**, 4-5 (2013), 346-353

XLV Study of $B^\pm \to K^\pm (K_s K \pi)^0$ decay and determination of η_c and $\eta_c(2S)$ parameters
A. Vinokurova *et al.,* (Belle Collaboration)
Phys. Lett. B **706**, 2-3 (2011), 139-149

XLVI Search for *CP* Violation in the Decay $D^+ \to K_S^0 K^+$
B. R. Ko *et al.,* (Belle Collaboration)
JHEP 02 (2013) 098