Aligarh Muslim University

Search for CP violation and rare decays in neutral charm mesons at Belle

---

Thesis

Submitted for the award of the degree of

Doctor of Philosophy

By

Nisar NK

Supervisor:
Dr. Shakeel Ahmad
Department of Physics
AMU Aligarh

Co-supervisor:
Prof. Tariq Aziz
Department of High Energy Physics
Tata Institute of Fundamental Research
Mumbai

According to UGC (minimum standard for the award of Ph.D degree)
Regulations, 2009

In the name of God most gracious most merciful

CANDIDATE'S DECLARATION

I, **Nisar NK**, Department of Physics, certify that the work embodied in this Ph. D. thesis is my own bonafide work carried out by me under the supervision of **Dr. Shakeel Ahmad** at Aligarh Muslim University, Aligarh. The matter embodied in this Ph. D. thesis has not been submitted for the award of any other degree.

I declare that I have faithfully acknowledged, given credit to and referred to the research workers wherever their works have been cited in the text and the body of the thesis. I further certify that I have not willfully lifted up some other's work, para, text, data, result, etc. reported in the journals, books, magazines, reports, dissertations, theses, etc., or available at web-sites and included them in this Ph. D. thesis and cited as my own work.

Date:

Signature of the candidate
Nisar NK

CERTIFICATE FROM THE SUPERVISOR/CO-SUPERVISOR

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Signature of the Co-supervisor
Prof. Tariq Aziz
Department of High Energy Physics
Tata Institute of Fundamental Research
Mumbai, Maharashtra, 400005

Signature of the Supervisor
Dr. Shakeel Ahmad
Department of Physics
Aligarh Muslim University
Aligarh, Uttar Pradesh, 202002

(Signature of the Chairman of the Department with seal)

COURSE WORK/COMPREHENSIVE EXAMINATION/PRE-SUBMISSION SEMINAR COMPLETION CERTIFICATE

This is to certify that **Mr. Nisar NK**, Department of Physics, has satisfactorily completed the course work/ comprehensive examination and presubmission seminar requirement which is part of his Ph. D. programme.

Date: **(Signature of the Chairman of the Department)**

COPYRIGHT TRANSFER CERTIFICATE

Title of the Thesis: Search for CP violation and rare decays in neutral charm mesons at Belle

Candidate's Name: **Nisar NK**

COPYRIGHT TRANSFER

The undersigned hereby assigns to the Aligarh Muslim University, Aligarh copyright that may exist in and for the above thesis submitted for the award of the Ph. D. degree.

Signature of the candidate

Note: However, the author may reproduce or authorize others to reproduce material extracted verbatim from the thesis or derivative of the thesis for author's personal use provide that the source and the University's copyright notice are indicated.

Acknowledgements

First and foremost praise is to Almighty God for giving the strengths and blessing in completing this thesis.

First of all I thank Prof. Mohd. Zafar, with whom I started working for my Ph.D degree, for providing this wounderful opportunity for me. It is my great pleasure to express my gratitude to my supervisor Dr. Shakeel Ahmad for his guidance, support and encouragement in pursuing this work. I pay my gratitude to both of them for providing necessary infrastructure and resources to accomplish my research work.

I thank Prof. Mohd. Afzal Ansari, Chairman of our department, for the academic support and the facilities provided for me to carry out research work at department of physics, AMU.

I am deeply grateful to my co-supervisor Prof. Tariq Aziz for for making it possible to carry out this work in TIFR Mumbai. He has made available his support in a number of ways like, introducing me to the international high energy physics community and arrange necessary funding. This thesis would not have been possible without the guidance the hard work of Prof. Gagan B. Mohanty and Prof. Karim Trabelsi. Their guidance helped me in all the time of research and writing of this thesis.

I am thankful to Prof.Yoshihide Sakai, KEK-IPNS, for the financial and logistical support. I would like to express my sincere thanks towards Dr. Anze Zupanc and Prof. Alan J. Schwartz for devoting their time and knowledge in the completion of my analyses. I formally wish to thank Profs. Thomas E. Browder , Marko Starič, Byeong Rok Ko and Takanori Hara for their valuable suggestions on my work.

I express my gratitude toward my parents, siblings and wife for their encouragement, support and prayers which help me in completion of my Ph.D. This thesis has been kept on track and been seen through to completion with the support and encouragement of numerous people including my well wishers, my friends, colleagues and various institutions. Therefore I would like to extend my sincere gratitude to all of them.

I express my sincere thanks to my Belle colleagues and seniors Anu, Basith, Deepanwita, Minakshi, Nibedita, Prasanth, Saurabh, Subhashree, Tara, Varghese, Vipin and Vishal. I am also thankful to my friends Anand, Aslam, Bajarang, Bibhu, Haris, Hashim, Nizam, Saranya, Salih, Salim, Shafeeq, Shahistha and Zoheb. Department staff members Meerajudden bhai, Mushraf bhai,and Naseem bhai have always been available to help me in technical and administrative works.

Abstract

The Belle detector is a general purpose magnetic spectrometer with a large solid angle coverage. It is designed and optimized to carry out studies of the CP violation asymmetries as well as rare decays of B and D mesons. Operating at the world's highest luminosity machine (KEKB), Belle could accumulate 1 ab^{-1} of data at different center of mass energies. The detector is configured around a 1.5 T superconducting solenoid and iron structure surrounding the collision point at the Tsukuba interaction region.

Within the standard model (SM) CP violation in charm decays are expected to be very small. The $D^0 \to \pi^0\pi^0$ decay proceeds via singly cabibbo-suppressed amplitude which could interfere with new physics (NP) amplitude, eventually giving rise to an enhanced CP violation effect. In the SM CP violation arises due to interference between the tree and loop amplitudes and is suppressed by a factor of order $\sim 10^{-3}$. In 2012, LHCb and CDF Collaborations reported a large value of the difference of direct CP asymmetries (ΔA_{CP}) between $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ decays, and the world average for ΔA_{CP} was 4.6σ away from zero. Although a large ΔA_{CP} could be explained by non-SM physics, it may be simply due to an unexpectedly enhanced SM $c \to u$ penguin amplitude. In the latter case, one expects fractional-percent CP asymmetries in other decays such as $D^0 \to \pi^0\pi^0$.

We measure time-integrated CP violating asymmetry in the $D^0 \to \pi^0\pi^0$ decay and updated the result on $D^0 \to K^0_S\pi^0$ using Belle data of an integrated luminosity of 966 fb^{-1}. The charge of the accompanying low-momentum or "slow" pion π_s^+ in the decay $D^{*+} \to D^0\pi_s^+$ identifies the flavor of the neutral charm meson (whether it is a D^0 or $\bar{D}^0$) at its production. The measured asymmetry (A_{rec}) in the yield of D^{*+} and D^{*-} includes three contributions: the underlying CP asymmetry A_{CP} ,

the forward-backward asymmetry (A_{FB}) due to the production mechanism of D^{*+} mesons, and the detection asymmetry between positively and negatively charged pions ($A_{\epsilon}^{\pi_s}$). To estimate $A_{\epsilon}^{\pi_s}$, we use samples of the Cabibbo-favored decays $D^0 \to K^-\pi^+$ ("untagged") and $D^{*+} \to D^0\pi_s^+ \to K^-\pi^+\pi_s^+$ ("tagged"), and we assume the same A_{FB} for D^0 and $\bar{D}^0$ mesons. By subtracting the measured asymmetries in these two decay modes, we obtain the $A_{\epsilon}^{\pi_s}$ correction factor [$\mathcal{O}(0.1)\%$]. After $A_{\rm rec}$ is corrected for $A_{\epsilon}^{\pi_s}$, one is left with $A_{\rm rec}^{\rm corr} = A_{CP} + A_{FB}(\cos\theta^*)$. While A_{CP} is independent of all kinematic variables, A_{FB} is an odd function of the cosine of the D^{*+} polar angle θ^* in the center of mass frame. We thus extract A_{CP} and A_{FB} by estimating $A_{\rm rec}^{\rm corr}$ in 10 bins of $\cos\theta^*$, followed by the addition and subtraction of the asymmetry belonging to the bins with the same value of $|\cos\theta^*|$.

We use Monte Carlo (MC) simulated events to devise selection criteria and investigate possible sources of background. The selection optimization is performed by minimizing the expected statistical error on $A_{\rm rec}$. The level of background is obtained by appropriately scaling the number of events observed in a data sideband of the reconstructed D^* mass. After applying all selection criteria, we find that about 6% of the total $D^* \to D(\pi^0\pi^0)\pi_s$ events contain multiple candidates of which we select a single-best candidate per event. The asymmetry $A_{\rm rec}$ and the sum of the D^{*+} and D^{*-} yields are obtained from a simultaneous fit to their ΔM distributions, where ΔM is the difference in the reconstructed mass of D^{*+} and D^0. We obtain a signal yield of 34460 ± 273 events for $D^0 \to \pi^0\pi^0$ decay and 466814 ± 773 events for $D^0 \to K_S^0\pi^0$ mode. We estimate various systematics uncertainties associated with the measurements such as signal shape, slow pion correction, and A_{CP} extraction method. Each of the individual errors are added in quadrature and, the resulting values are 0.10% for $D^0 \to \pi^0\pi^0$ and 0.07% for $D^0 \to K_S^0\pi^0$.

In summary, we have measured the time-integrated CP asymmetry in the $D^0 \to \pi^0\pi^0$ decay by correcting for the detector-induced asymmetries to be $[-0.03 \pm$

$0.64(\text{stat}) \pm 0.10(\text{syst})]\%$, which is consistent with no CP violation. This constitutes an order of magnitude improvement over the existing result. We have also updated the CP asymmetry result in the $D^0 \to K^0_S\pi^0$ decay: $A_{CP}(D^0 \to K^0_S\pi^0) = [-0.21 \pm 0.16(\text{stat}) \pm 0.07(\text{syst})]\%$.

Flavor changing neutral current (FCNC) processes are forbidden at tree level in the SM, although they can occur at higher orders. In contrast, there are several NP models, which allow FCNC even at tree level and can thus substantially enhance the branching fraction of the related decay processes. The decay $D^0 \to \gamma\gamma$ is one such example that provides a sensitive NP probe. Calculations within the SM predict a branching fraction in the range $(1\text{–}3) \times 10^{-8}$. Using the framework of minimal supersymmetric standard model, theorists have suggested that exchange of gluinos, supersymmetric partner of gluons, can raise the $D^0 \to \gamma\gamma$ decay rate up to 6×10^{-6}. Previously CLEO, *BABAR*, and BESIII collaborations have searched for this decay without any signal. The *BABAR* limit, 2.2×10^{-6}, continues to be the world's best upper limit.

We search for the rare radiative decay $D^0 \to \gamma\gamma$ using the Belle data sample with an integrated luminosity of 832.4 fb^{-1}. We reconstruct a D^0 candidate from two energetic photons. To reduce large combinatorial backgrounds arising from random photon combinations, we require that the D^0 be produced in the decay $D^{*+} \to D^0\pi^+$. The D^{*+} production is associated with a large uncertainty (12.5%). Thus we measure the $D^0 \to \gamma\gamma$ branching fraction with respect to a well measured mode $D^0 \to K^0_S\pi^0$. The relative measurement also helps cancel out systematic uncertainties common to both the signal and normalization channels. We use two kinematic variables to identify signal: the reconstructed invariant mass of the D^0 candidate,(MD^0), and ΔM.

We use MC simulated data to identify various background processes and to optimize selection criteria. We study various backgrounds that can be broadly classified into peaking and combinatorial. To extract the signal yield, we perform an unbinned extended maximum likelihood fit to the two-dimensional (2D) distributions of $M(D^0)$ and ΔM. For the signal and combinatorial background component, the correlation between the two observables is found to be small. In case of the peaking background, there is a significant correlation between ΔM and $M(D^0)$, which we account for via a joint PDF. Signal and peaking background are fixed to the corresponding MC values, and we use different control samples to take possible data-MC difference into account.

Using the 2D fit we find 4 ± 15 signal, 210 ± 32 peaking and 2934 ± 59 combinatorial background events. In absence of a statistically significant signal, we derive an upper limit at 90% confidence level (CL) on the signal yield following a frequentist method using an ensemble of pseudoexperiments. For a given signal yield, we generate 5000 sets of signal and background events according to their PDFs, and perform the fit. The CL is obtained as the fraction of samples, which gives a fit yield larger than that observed in data. We estimate various systematics errors namely PDF shape, selection criteria, and photon reconstruction efficiency. We obtain a total systematics uncertainty of 4 events and it is accounted for in UL calculation by smearing the fit yield. Including systematic uncertainties the upper limit on the branching fraction of $D^0 \to \gamma\gamma$ is 8.4×10^{-7}. Our result constitutes the most restrictive limit on $D^0 \to \gamma\gamma$ to date and can be used to constrain the NP parameter space.

List of Publications

1. N.K. Nisar *et al.* (Belle Collaboration), (2014), **'Search for CP violation in $D^0 \rightarrow \pi^0\pi^0$ decays'**, Phys. Rev. Lett. **112**, 211601.

2. N.K. Nisar, (2014), **'D^0-$\bar{D}^0$ mixing and CP violation results from Belle'**, Nucl. Phys. B (Proc. Suppl.), arXiv:1411.2083[hep-ex]

3. N.K. Nisar *et al.* (Belle Collaboration), (2015), **'Search for the rare decay $D^0 \rightarrow \gamma\gamma$ at Belle'** Phys. Rev. D (communicated).

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1 Standard Model

Everything in our universe is made of a handful of basic building blocks called fundamental particles that are governed by four fundamental forces. The standard model (SM) of particle physics is the theory concerning these fundamental particles and their interactions. According to the SM, matter is built from a set of twelve fermions: six quarks and six leptons. Out of the four fundamental forces, electromagnetic, strong, weak and gravitational, all but for the last one are part of the SM. The quark content is given as

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix} \tag{1.1}$$

The quarks in the upper row have electric charge $+2/3$ while those in the lower row have charge $-1/3$ in unit of proton charge. The quarks can interact through all fundamental forces. The lepton content is

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix} \tag{1.2}$$

The upper row consists of neutrinos, massless neutral particles that interact solely through weak interaction. The leptons in the lower row have integral charges and participate in all interactions except the strong.

In addition to this, there are five bosons in the SM; four of them are force carriers while the remaining one is the Higgs boson that gives mass to all fermions and bosons in the SM. Among the force carriers gluons (g) give rise to the strong interaction, photons (γ) are the carrier of the electromagnetic interaction, and $W^{\pm}$, Z^0 bosons are the quanta of the weak interaction.

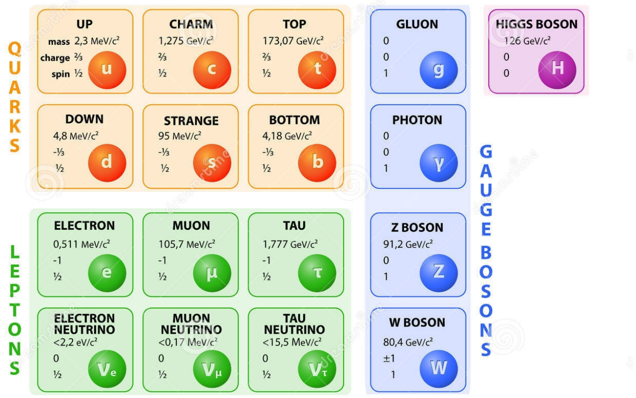


FIGURE 1.1: Elementary particles of the standard model, with the three generations of quarks and leptons in the first three columns, gauge bosons in the fourth column, and the Higgs boson in the fifth.

All stable matter in the universe is made from particles that belong to the first generation fermions (first column in Fig. 1.1) together with photons. Any heavier particles, built from quarks of the second and third generations, quickly can decay to the next most stable level via the weak interaction. Putting it differently, all quarks mix among themselves through the weak interaction. A historic progression in this field is given in the following sections.

1.1.1 Cabibbo Mechanism

Nicola Cabibbo, an Italian physicist, introduced an angle named after him (θ_c) [1] to preserve the universality of the weak interaction. An important observation by him was that the coupling constants of the following three flavor-changing decay modes were related:

(i) muon decay $\mu^- \to \nu_\mu e^- \overline{\nu}_e$: coupling constant $g_{e\mu}$,

(ii) neutron decay $n \to pe^- \bar{\nu}_e$ (at quark level $d \to ue^- \bar{\nu}_e$): coupling constant g_{ud}, and

(iii) kaon decay $K^- \to \pi^0 e^- \bar{\nu}_e$ (at quark level $s \to ue^- \bar{\nu}_e$): coupling constant g_{us}.

Estimations of coupling constants from above decay modes are consistent with the relation $|g_{e\mu}|^2 = |g_{ud}|^2 + |g_{us}|^2$. It implies that there is only one coupling constant $g \equiv g_{e\mu}$, and the u quark simply couples to one particular combination of d and s, given by $d' \equiv d \cos\theta_c + s \sin\theta_c$. Here θ_c is Cabibbo quark-mixing angle. In other words, the weak coupling strength between electron and muon are equal to the the coupling strength of u quark and the rotated doublet d'.

Only three quarks were known when Cabibbo introduced the rotated doublets. As it turned out, the Cabibbo angle was not enough to account for the suppression of $K_L \to \mu^+\mu^-$ compared to $K^+ \to \mu^+\nu_\mu$. This suppression of flavor-changing neutral current (FCNC) transition could be explained if another quark “c" with charge +2/3 were postulated, which would couple to the combination $s' = -d \sin\theta_c + s \cos\theta_c$. The suppression of the $s \to d$ process is due to the cancellation of the $s \to u \to d$ contribution to total amplitude by the $s \to c \to d$ contribution, and is known as the Glashow-Iliopoulos-Maiani (GIM) mechanism [2]. Now the rotated quark doublets are:

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \tag{1.3}$$

Here $|d\rangle$ and $|s\rangle$ are the mass eigenstates, while $|d'\rangle$ and $|s'\rangle$ are the weak eigenstates. The rotation of mass eigenstates through the Cabibbo angle gives rise to the weak eigenstates.

$$\begin{pmatrix} d' \\ s' \end{pmatrix}_{\text{weak}} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}_{\text{mass}} \tag{1.4}$$

1.1.2 CKM Matrix

The Cabibbo-Kobayashi-Maskawa (CKM) matrix is an extension of the Cabibbo mechanism to six quarks in three generations. It is worthwhile to note that Kobayashi and Maskawa predicted the existence of six quarks when only three quarks were known to exist. The CKM matrix not only parametrizes quark mixing in three generations but also explains CP (“CP" standing for the combination of charge conjugation and parity) violation as a natural consequence of quark mixing. In the SM, CP violation is incorporated via an irreducible complex phase appearing in the 3×3 unitary CKM matrix and all observed phenomena of CP violation can be explained by it.

Let U' and D' represent the weak (flavor) eigenstates, whose column vectors consist of three up- and down-type quarks respectively.

$$U' = \begin{pmatrix} u' \\ c' \\ t' \end{pmatrix}, \qquad D' = \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} \tag{1.5}$$

The weak charged current of the SM, in the basis of weak eigenstates, is

$$J_\mu = \frac{g}{\sqrt{2}} \overline{U}'_L \gamma_\mu D'_L + h.c., \tag{1.6}$$

where the subscript L represents the left chiral component of quark spinors [the quark spinors are operated by the left-handed projection operator, $(1-\gamma_5)/2$]. By definition, the charged current interactions are diagonal in the weak eigenstate basis. Consider U and D as mass eigenstate basis obtained from U' and D' through the following unitary transformations $U'_L = V_{UL} U_L$ and $D'_L = V_{DL} D_L$. The weak charged current in the mass basis becomes

$$\begin{aligned} J_\mu &= \frac{g}{\sqrt{2}} \bar{U}_L V^\dagger_{UL} \gamma_\mu V_{DL} D_L + h.c. \\ &= \frac{g}{\sqrt{2}} \bar{U}_L \gamma_\mu (V^\dagger_{UL} V_{DL}) D_L + h.c. \\ &= \frac{g}{\sqrt{2}} \bar{U}_L \gamma_\mu (V_{CKM}) D_L + h.c. \end{aligned} \tag{1.7}$$

From Eq. 1.7, it is clear that the coupling between mass eigenstates U_L and D_L is given by $(g/\sqrt{2})V_{CKM}$. Here, V_{CKM} is a unitary matrix, which is the previously introduced CKM matrix. Unlike the 2 × 2 Cabibbo matrix the CKM matrix elements are complex, and these complex coupling constants give rise to CP violation. The CKM matrix transforms mass eigenstates to weak eigenstates, as shown below.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_{\text{weak}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\text{mass}} \tag{1.8}$$

Generally, a 3 × 3 complex matrix has 9 real and 9 imaginary quantities, but the unitarity constraint ($V^\dagger_{CKM} V_{CKM} = 1$) and re-phasing leave three Euler angles

θ_{12}, θ_{23}, θ_{13} and one complex phase δ_{13}. These four parameters describe physics incorporated in the CKM matrix. A common and more convenient parametrization known as the Wolfenstein parametrization is described in terms of parameters λ, A, ρ and η:

$$V_{CKM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{1.9}$$

where $\lambda \equiv \sin\theta_{12}$ ($\theta_{12} \approx \theta_c$, Cabibbo angle), $A \equiv |V_{cb}|/\lambda^2$ and $V_{ub} \equiv A\lambda^3(\rho - i\eta)$

1.2 Physics beyond the Standard Model

Despite the consistency of the SM with all available data, it is well known that it cannot be a complete theory of all fundamental interactions. The model does not include the gravitational interaction and neither it has a suitable candidate for dark matter. Furthermore, the CP violation content of the SM is inadequate in explaining the matter-antimatter asymmetry observed in the universe. Flavor physics plays an important role in probing physics beyond the SM, which is complementary to the direct searches performed by energy-frontier experiments. Towards this end, CP violation in the charm sector and rare charm decays comprise two important probes.

1.3 CP violation in neutral D systems

If only the strong and electromagnetic interactions existed, D^0 and $\bar{D}^0$ mesons would be stable particles with a common mass. However, they decay via the weak interaction. None of the conservation laws prevent the transition of D^0 and $\bar{D}^0$ to a common state, leading to $D^0-\bar{D}^0$ mixing. The $|D^0\rangle$ and $\left|\bar{D}^0\right\rangle$ are the eigenstates of strong and electromagnetic interactions while in the weak interaction they mix and decay into other states. In this section we, develop a general formalism to deal with mixing and CP violation in the neutral D system. The formalism is generic such that the observable effect could be either from the SM or from new physics. The state of neutral D meson system in the $(D^0, \bar{D}^0)$ basis is given by:

$$|\psi(t)\rangle = a(t)\left|D^0\right\rangle + b(t)\left|\bar{D}^0\right\rangle, \tag{1.10}$$

where "t" is the proper time in the $D^0-\bar{D}^0$ frame. The time evolution of such system can be described by the Schrödinger equation:

$$i\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle \tag{1.11}$$

where H is the effective Hamiltonian and represented by a 2×2 matrix in this basis. Since both D^0 and $\bar{D}^0$ decay the evolution is not unitary, and hence the Hamiltonian H is not Hermitian. Any complex matrix can be expressed as a sum of Hermitian and anti-Hermitian parts. Therefore, H can be written as:

$$H = M - \frac{i}{2}\Gamma = \begin{pmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma_{1,1} & \Gamma_{1,2} \\ \Gamma_{2,1} & \Gamma_{2,2} \end{pmatrix}, \tag{1.12}$$

where both M and Γ are Hermitian. The eigenvalues of H are

$$\begin{aligned} \mu_H &= M_{11} - \frac{i}{2}\Gamma_{11} + \frac{1}{2}(\Delta m - \frac{i}{2}\Delta\Gamma), \text{and} \\ \mu_L &= M_{11} - \frac{i}{2}\Gamma_{11} - \frac{1}{2}(\Delta m - \frac{i}{2}\Delta\Gamma), \end{aligned} \tag{1.13}$$

Here, $\Delta m = m_H - m_L$ and $\Delta\Gamma = \Gamma_H - \Gamma_L$. The labels H and L stand for "heavy" and "light" respectively. The normalized eigenstates of the Hamiltonian are

$$\begin{aligned} |D_L\rangle &= p\left|D^0\right\rangle + q\left|\bar{D}^0\right\rangle, \\ |D_H\rangle &= p\left|D^0\right\rangle - q\left|\bar{D}^0\right\rangle, \end{aligned} \tag{1.14}$$

where p and q are complex. We study processes where D^0 or $\bar{D}^0$ decays to a final state f or to its CP conjugate $\bar{f}$. Any observable quantities ought to be independent of arbitrary phases of D^0, $\bar{D}^0$, f and $\bar{f}$. Other than the decay rates the phase invariant quantities that are relevant for CP violation are:

$$\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right|, \qquad \left|\frac{q}{p}\right|, \qquad \lambda_f \equiv \frac{q}{p}\frac{\bar{A}_f}{A_f}. \tag{1.15}$$

Where $A_f = \langle f \,|\, H \,|\, D^0\rangle$, $\bar{A}_f = \left\langle f \,\middle|\, H \,\middle|\, \bar{D}^0\right\rangle$ and $\bar{A}_{\bar{f}} = \left\langle \bar{f} \,\middle|\, H \,\middle|\, \bar{D}^0\right\rangle$.

1.3.1 Time evolution

Let us study the time evolution of an initial flavor eigenstate $|D^0\rangle$. Since the eigenstate $|D_H\rangle$ and $|D_L\rangle$ evolve independently without mixing, it is easier to write the evolution in terms of these states. In the beginning ($t = 0$),

$$\left|D^0(0)\right\rangle = \frac{1}{2p}(|D_L\rangle + |D_H\rangle). \tag{1.16}$$

At a later time t, the time evolution gives rise to

$$\begin{aligned} |D(t)\rangle &= \frac{1}{2p}(e^{-im_L t - \Gamma_L t/2}\,|D_L\rangle + e^{-im_H t - \Gamma_H t/2}\,|D_H\rangle) \\ &= g_+(t)\left|D^0\right\rangle - \frac{q}{p}g_-(t)\left|\bar{D}^0\right\rangle \end{aligned} \tag{1.17}$$

and similarly

$$\left|\bar{D}^0(t)\right\rangle = g_+(t)\left|\bar{D}^0\right\rangle - \frac{p}{q}g_-(t)\left|D^0\right\rangle, \tag{1.18}$$

where $g_\pm \equiv (1/2)(e^{-im_H t-\Gamma_H t/2} \pm e^{-im_L t-\Gamma_L t/2})$. Now the time-dependent decay rate of $D^0/\bar{D}^0$ decaying to f can be written as:

$$\begin{aligned}\frac{\frac{d\Gamma}{dt}[D^0(t)\to f]}{e^{-\Gamma t}N_f|A_f|^2} &= (1+|\lambda_f|^2)\cosh(\Delta\Gamma t/2) + (1-|\lambda_f|^2)\cos(\Delta m t)\\ &+ 2\mathrm{Re}(\lambda_f)\sinh(\Delta\Gamma t/2) + 2\mathrm{Im}(\lambda_f)\sin(\Delta m t),\end{aligned} \tag{1.19}$$

$$\begin{aligned}\frac{\frac{d\Gamma}{dt}[\bar{D}^0(t)\to f]}{e^{-\Gamma t}N_f|A_f|^2|p/q|^2} &= (1+|\lambda_f|^2)\cosh(\Delta\Gamma t/2) - (1-|\lambda_f|^2)\cos(\Delta m t)\\ &+ 2\mathrm{Re}(\lambda_f)\sinh(\Delta\Gamma t/2) - 2\mathrm{Im}(\lambda_f)\sin(\Delta m t).\end{aligned} \tag{1.20}$$

1.3.2 Types of CP violation

We can classify CP violation into three categories based on the phase invariant quantities given in Eq. (1.15).

1.3.2.1 CP violation in decay or direct CP violation

This asymmetry arises entirely due to the difference in the decay rates (amplitudes) of particle and antiparticle. Furthermore, it can occur for both neutral and charged mesons. The necessary condition for this kind of CP violation is

$$|\bar{A}_{\bar{f}}/A_f| \neq 1. \tag{1.21}$$

The observable CP violating quantity is

$$A_{\mathrm{d}} \equiv \frac{\Gamma(D^0 \to f) - \Gamma(\bar{D}^0 \to \bar{f})}{\Gamma(D^0 \to f) + \Gamma(\bar{D}^0 \to \bar{f})} = \frac{1 - |\bar{A}_{\bar{f}}/A_f|}{1 + |\bar{A}_{\bar{f}}/A_f|}. \tag{1.22}$$

The direct CP violation occurs if two amplitudes of comparable magnitude having different strong and weak phase contribute to a decay.

1.3.2.2 CP violation in mixing

It is possible to have CP violation even when $|\bar{A}_{\bar{f}}| = A_f$, if

$$|q/p| \neq 1 \tag{1.23}$$

Semileptonic decays of neutral pseudoscalar mesons provide a fertile ground for such a phenomenon. Let $P \to l^+X$ and $\bar{P} \to l^-X$ have same amplitude, while the amplitudes of $P \to l^-X$ and $\bar{P} \to l^+X$ vanish. The observation of a "wrong sign lepton" is, therefore, a signature of mixing. The difference in the probability of P and $\bar{P}$ decaying to the final state with wrong sign lepton is the measure of CP violation in mixing, given as:

$$A_{\mathrm{m}}(t) \equiv \frac{\frac{d\Gamma}{dt}[\bar{D}^0(t) \to l^+X] - \frac{d\Gamma}{dt}[D^0(t) \to l^-X]}{\frac{d\Gamma}{dt}[\bar{D}^0(t) \to l^+X] + \frac{d\Gamma}{dt}[D^0(t) \to l^-X]} = \frac{1 - |q/p|^4}{1 + |q/p|^4}. \tag{1.24}$$

1.3.2.3 CP violation in interference between mixing and decay

Even when there is no CP violation in decay alone ($|\bar{A}_f| = |A_f|$) nor in mixing ($|p/q| = 1$), there can be still CP violation if

$$\mathrm{Im}(\lambda_f) \neq 0 \tag{1.25}$$

This asymmetry arises for the cases where both D^0 and $\bar{D}^0$ decays to a common CP eigenstate, f_{CP}. The CP asymmetry can be defined as:

$$A_{\mathrm{i}}(t) \equiv \frac{\frac{d\Gamma}{dt}[\bar{D}^0(t) \to f_{CP}] - \frac{d\Gamma}{dt}[D^0(t) \to f_{CP}]}{\frac{d\Gamma}{dt}[\bar{D}^0(t) \to f_{CP}] + \frac{d\Gamma}{dt}[D^0(t) \to f_{CP}]}. \tag{1.26}$$

The resultant CP violation due to interference between mixing and decay is also known as indirect CP violation.

1.3.3 CP violation in $D^0 \to \pi^0\pi^0$

CP violation in the charm sector constitutes an excellent probe for new physics, as it is expected to be small in the SM owing to GIM suppression [4]. Therefore, theorists used to claim for long that “any significant deviation from a zero CP asymmetry in the charm decay would be a signature for new physics”. This topic got an immediate attention with the LHCb’s observation of a nonzero difference in CP violation asymmetry between the decays $D^0 \to \pi^+\pi^-$ and $D^0 \to K^+K^-$ [5], which was later supported by the CDF Collaboration [6]. Whether this is any sign of new physics or not is still under discussion. In the mean time, it has been suggested to precisely measure CP asymmetry in the isospin related channels such as $D^0 \to \pi^0\pi^0$ [7–9]. The $D^0 \to \pi^0\pi^0$ decay is a singly cabibbo suppressed decay, and various amplitudes contributing to this process are shown in Fig. 1.2. Indirect CP violation occurs due to interference between the tree amplitudes with and without mixing, while direct CP violation arises owing to the interference of tree and penguin amplitudes. It is clear from the possible amplitudes shown in Fig. 1.2 that both direct and indirect CP violation can occur for this mode. In addition to this, there is possibility for an interference of the singly cabibbo suppressed amplitude with new physics contribution , which can enhance CP violation effect.

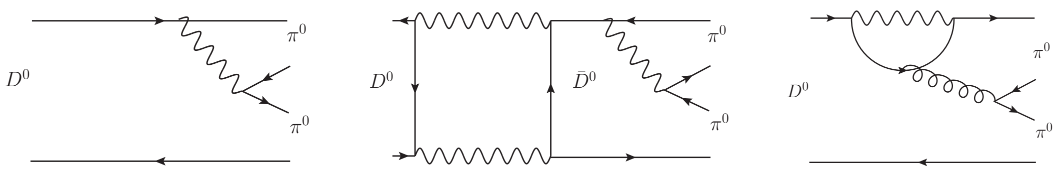


FIGURE 1.2: Feynman diagrams for various amplitudes contributing to $D^0 \to \pi^0\pi^0$ decay: (left) the $c \to d$ tree diagram, (middle) D^0–$\bar{D}^0$ mixing followed by the $\bar{c} \to \bar{d}$ tree diagram, and (right) the $c \to u$ loop (penguin) diagram.

1.4 Rare charm decay

Flavor Changing Neutral Current (FCNC) processes involve the change of flavor of quarks without altering their charge; an example is $c \to u\gamma$. These processes are forbidden at tree level in the SM but occur at higher orders starting at the one-loop level. The decays that proceed through electroweak FCNC diagrams are commonly referred to as rare decays. Absence of FCNC at tree level in the SM implies that the related processes provide a powerful tool for probing the structure of electroweak interactions. Some New Physics (NP) models like Minimal Supersymmetric Standard Model (MSSM) allow FCNC even at tree level and can thus enhance the decay branching fraction. This enhancement could be several orders of magnitude and acts as a clean window for probing NP.

In the SM, the FCNC decay $D^0 \to \gamma\gamma$ is contributed by both short-distance (SD) and long-distance (LD) effects. The SD is perturbative in nature and proceeds via the intermediate states involving virtual quarks. Analytical methods are employed to estimate its contribution. However, the major contribution to $D^0 \to \gamma\gamma$ is from the nonperturbative LD effect that relies on the vector meson dominance (VMD) and unitarity constraints. Unlike SD, here real hadrons $(\omega, \rho..)$ propagate in the intermediate states.

The rare decays of K and B mesons are well studied while they are not so well explored in the charm sector owing to very small expectation of FCNC in the SM. For charm decays, relatively lighter quarks, d and s, will propagate through the loop. Subsequently, the tiny values of $(m_q/M_W)^2$ make GIM suppression stronger in this sector. For the heavier b quark propagating in the loop, the small value of associated CKM matrix elements ($|V_{cb}^* V_{ub}|$) causes another suppression. In contrast, for K and B mesons the heavy t quark in the loop make the GIM suppression mild. The aforementioned two effects make the SD a minor contributor to rare charm decays. The LD effect lift the above suppression by few orders of magnitude. Since the LD contribution is non-perturbative in nature, it cannot be calculated by analytical methods.

1.5 SM prediction for $D^0 \to \gamma\gamma$

In the SM, the branching fraction for $D^0 \to \gamma\gamma$ decay is given by [10]:

$$\mathcal{B}(D^0 \to \gamma\gamma) = \frac{M_{D^0}^3 \tau_{D^0}}{64\pi} \left[|B(D^0 \to \gamma\gamma)|^2 + |C(D^0 \to \gamma\gamma)|^2\right] \tag{1.27}$$

Where, M_{D^0} and τ_{D^0} are mass and lifetime of D^0 mesons, B and C are parity-conserving and parity-violating invariant amplitudes, respectively.

1.5.1 SD contribution $c\bar{u} \to \gamma\gamma$

The quark-level transition $c \to u\gamma\gamma$ can arise either via a one-particle irreducible (1PI) diagram in which both photons arise from same vertex or via a one-particle reducible (1PR) diagram in which a photon is emitted either from the initial c quark or from the final u quark as shown in Fig. 1.3. The filled circle in Fig. 1.3 represents various loop-level diagrams for the $c \to u\gamma$ process. The process can proceed either via one-loop or two-loops as shown in Fig. 1.4. Since the important suppression factors are independent of gauge couplings, it is possible that higher

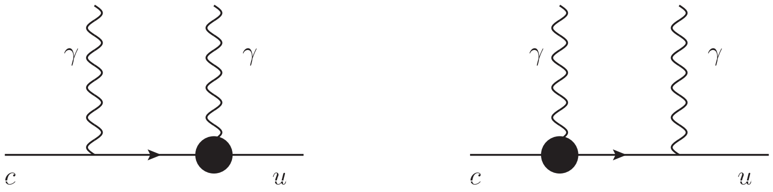


FIGURE 1.3: One-particle reducible contribution to $c \to u\gamma\gamma$.

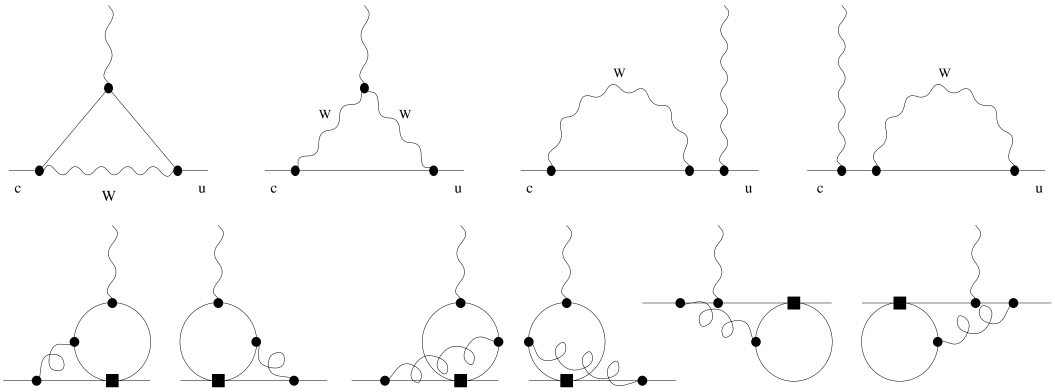


FIGURE 1.4: Loop level diagrams contributing to $c \to u\gamma$ process: (upper) one loop and (lower) two loop.

orders in the perturbation theory give dominant contributions to the radiative amplitude.

By inputting the two-loop $c \to u\gamma$ contributions (dominant one) to $c \to u\gamma\gamma$ amplitude (shown in Fig. 1.3) gives invariant amplitudes for $c\bar{u} \to \gamma\gamma$ as:

$$|B(c\bar{u} \to \gamma\gamma)| = |C(c\bar{u} \to \gamma\gamma)| = \frac{G_F \alpha}{3\sqrt{2}\pi} \frac{m_c}{M_D - m_c} f_D |A|, \tag{1.28}$$

where, $|A| \simeq 0.0047$,G_F is the Fermi coupling constant, α the fine structure constant, m_c charm quark mass, and f_D the form factor. Inserting these quantities in Eq. (1.27), we get the SD contribution as

$$\mathcal{B}(D^0 \to \gamma\gamma) \simeq 3 \times 10^{-11} \text{ [10]} \tag{1.29}$$

1.5.2 LD contribution $D^0 \to \gamma\gamma$

The LD contribution can be modeled using the vector meson dominance (VMD) and unitarity constraint with one or two particle intermediates state.

1.5.2.1 Vector meson dominance

The $D^0 \to \gamma\gamma$ amplitude can be seen as single vector meson dominance process (Fig. 1.5):

$$D^0 \to \gamma + \sum_k V_k^{0*} \to \gamma + \gamma \tag{1.30}$$

The invariant amplitudes based on the VMD model are

$$\begin{aligned} B(D^0 \to \gamma\gamma) &= \sum_i \frac{2e}{f_{V_i}} B_{V_i}\eta_i, \text{and} \\ C(D^0 \to \gamma\gamma) &= \sum_i \frac{2e}{f_{V_i}} C_{V_i}\eta_i, \end{aligned} \tag{1.31}$$

where f_V is the coupling for the $V^0 - \gamma$ conversion amplitude, the index "i" refers to a specific vector meson: ρ^0, ω^0, ϕ^0 etc., and $\eta_i \simeq 1/2$. The branching fraction estimated using VMD is:

$$\mathcal{B}^{(\text{VMD})}(D^0 \to \gamma\gamma) = (3.5^{+4.0}_{-2.6}) \times 10^{-8} \text{ [10]} \tag{1.32}$$

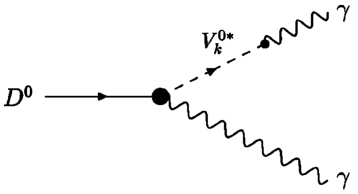


FIGURE 1.5: Single vector meson dominance contribution.

1.5.2.2 Single-particle unitarity contribution

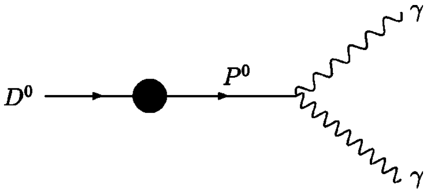


FIGURE 1.6: Weak mixing contribution to $D^0 \to \gamma\gamma$ decay.

In this process the D^0 mixes with a scalar meson and finally decays into a pair of photons as shown in Fig. 1.6. Further, the scalar meson could be a ground-state particle or a resonance. The respective branching fractions are given as below:

$$\mathcal{B}^{(\text{ground})}(D^0 \to \gamma\gamma) \simeq 3 \times 10^{-11} \text{ [10]} \tag{1.33}$$

$$\mathcal{B}^{(\text{resonance})}(D^0 \to \gamma\gamma) \sim 10^{-10} \text{ [10]} \tag{1.34}$$

1.5.2.3 Two-particle unitarity contribution

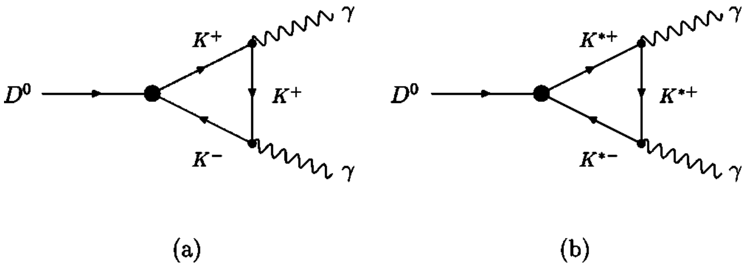


FIGURE 1.7: Two-particle unitarity contributions: (a) K^+K^- and (b) $K^{*+}K^{*-}$.

In this model, the D^0 decays to two photons via a two-particle intermediate state as shown in Fig. 1.7. The branching fraction due to this contribution is estimated to be:

$$\mathcal{B}^{K^+K^-}(D^0 \to \gamma\gamma) \sim 0.7 \times 10^{-8} \text{ [10]}. \tag{1.35}$$

Considering all above contributions, we anticipate a branching fraction for $D^0 \to \gamma\gamma$ in the neighborhood of 10^{-8} [10].

1.6 $D^0 \to \gamma\gamma$ at beyond Standard Model

The Minimal Supersymmetric Standard Model (MSSM) is an minimal extension to the Standard Model which considers only the minimum number of new particle states and new interactions consistent with phenomenology. Supersymmetry pairs every Standard Model fermion with a boson and vice versa. A typical contribution to the $c \to u\gamma$ transition in the minimal supersymmetric standard model (MSSM) is shown in Fig. 1.8. In this model, a squark (superpartner of the quark) is altered by a mass insertion, indicated by " × " in Fig. 1.8, and the exchange of squarks in loop leads to FCNC. Using the MSSM framework, $D^0 \to \gamma\gamma$ branching fraction is estimated to be:

$$\mathcal{B}(D^0 \to \gamma\gamma) \sim 6 \times 10^{-6} \text{ [11]}. \tag{1.36}$$

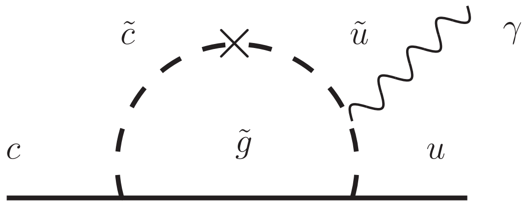


FIGURE 1.8: MSSM contribution to the $c \to u\gamma$ transition.

CHAPTER 2

Experimental Setup

The analyses presented in this thesis are based on the data collected by the Belle detector [12], located at an interaction point of the KEKB [13] asymmetric-energy e^+e^- collider. The KEKB is world's highest luminosity machine that was able to deliver data to Belle at an unprecedented instantaneous luminosity exceeding 2.11×10^{34} cm^{-2}s^{-1}. The Belle detector is designed and optimized to carry out a suite of measurements in various decay modes of B mesons to verify the Kobayashi-Maskawa mechanism [3] that explains CP violation in the standard model. It is a general-purpose magnetic spectrometer having a large solid angle coverage, which provides precise measurements of the B-meson decay vertex, excellent momentum resolution and particle identification capability. Major part of the data is accumulated at the $\Upsilon(4S)$ resonance in order to record as many B mesons as possible for the above CP violation study. However, the KEKB beam energies are tunable enabling Belle to collect data at $\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(3S)$ and $\Upsilon(5S)$ resonances. The latter samples are of special interest for hadron spectroscopy and to study the properties of B_s^0 mesons. Besides these, Belle has accumulated a copious amount of D mesons at various $\Upsilon(nS)$ resonances for studying charm physics.

2.1 KEKB

The KEK B-Factory (KEKB) is a two-ring, asymmetric-energy e^+e^- collider, which broke all previous records for integrated and instantaneous luminosity for a high energy accelerator. The KEKB achieved its highest peak luminosity, 2.1×10^{34} cm^{-2}s^{-1} [14] during June 2009 that was more than twice the design value (1×10^{34}cm^{-2}s^{-1}). As a result, Belle was able to accumulate over 1 ab^{-1} of data. An 8GeV electron ring (HER) and a 3.5GeV positron ring (LER) are installed side by side in a tunnel 11 m below the ground level, as shown in Fig.2.1. Each ring is about 3 km long and can store beam currents up to 2.0 A in the LER and 1.35 A in the HER with 508.9 MHz radio frequency (RF) acceleration systems. The two beams collide at the interaction point (IP) with a finite angle of 22 mrad in the horizontal plane. The nonzero crossing angle allows to fill all RF buckets with bunches without any risk of parasitic collision. The crossing angle also eliminates the need for separation dipole magnet at an expense of lower luminosity. In order to compensate the loss in luminosity, two superconducting crab cavities [12] were installed in each ring. In crab the crossing scheme, the bunches are kicked in the horizontal plane by transverse RF in the crab cavities so that they rotate and collide head-on at the IP. The scheme is illustrated in Fig. 2.2. The main parameters of KEKB are listed in Table 2.1.

2.2 Belle Detector

The Belle detector is configured around a 1.5 T superconducting solenoid and iron structure surrounding KEKB beams at the Tsukuba interaction region as shown in Fig. 2.3. The detector is composed of a number of subdetectors to perform vertexing, tracking and particle identification. The magnetic field enables the momentum measurement of charged particles. A silicon vertex detector (SVD) situated just

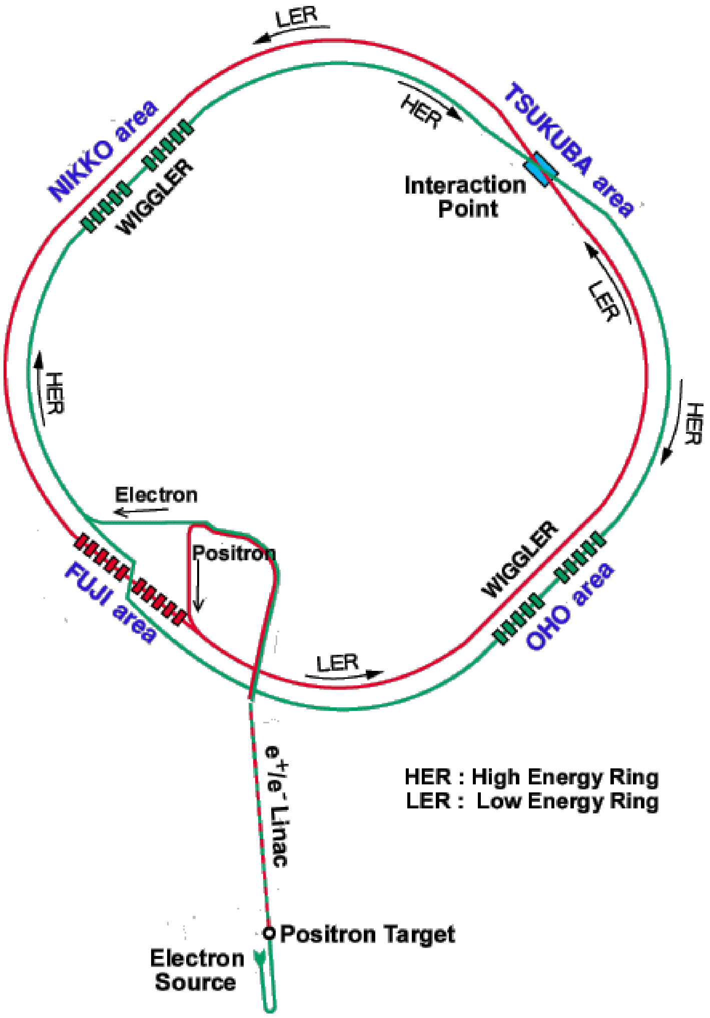
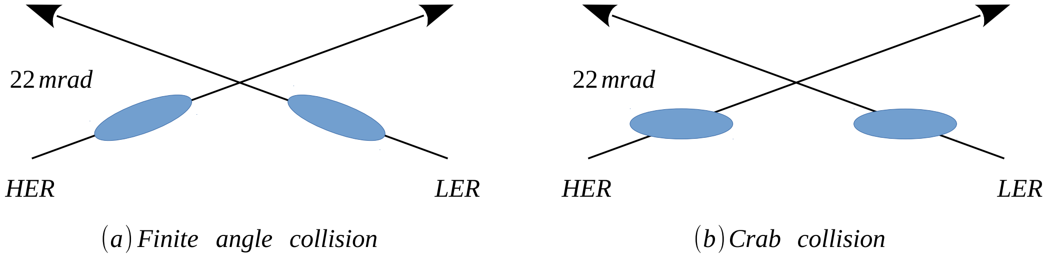


FIGURE 2.1: Schematic view of the KEKB collider [15].



(a) Finite angle collision

(b) Crab collision

FIGURE 2.2: Illustration of collision schemes at KEKB: (a) finite angle collisions and (b) crab crossing.

TABLE 2.1: Design value of main parameters of the KEKB accelerator [13].

Parameter	LER	HER	unit
Energy	3.5	8.0	GeV
Circumference	3.02		km
Luminosity	1×10^{34}		$\mathrm{cm}^{-2}\mathrm{s}^{-1}$
Crossing angle	± 11		mrad
Tune shifts (ξ_x/ξ_y)	0.039/0.052		
Beta function at IP (β_x^*/β_y^*)	0.33/0.01		m
Beam current	2.6	1.1	A
Natural bunch length (σ_z)	0.4		cm
Energy spread (σ_ε)	7.1×10^{-4}	6.7×10^{-4}	GeV
Bunch spacing (s_b)	0.59		m
Particle per bunch	3.3×10^{10}	1.4×10^{10}	
Emittance ($\varepsilon_x/\varepsilon_y$)	1.8×10^{-8}	3.6×10^{-10}	
RF voltage	6.5	12	MV
RF frequency	508.89		MHz
Harmonic number	5120		
Bending radius	16.3	104.5	m
Length of bending magnet	0.915	5.86	m

outside of the cylindrical beryllium beam pipe measures the decay vertices of B and D mesons. The charged particle tracking is performed by a central drift chamber (CDC) together with the SVD. Particle identification is provided by specific ionization (dE/dx) measurements in the CDC, Cherenkov photon yields from an aerogel threshold Cherenkov counter (ACC), and information from a time-of-flight counter (TOF) placed radially outside of the CDC. Electromagnetic showers are detected in an electromagnetic calorimeter (ECL), composed of an array of CsI(Tl) crystals. All these detector components are located inside the solenoid coil. Muons and K_L mesons are identified by arrays of resistive plate counters interspersed in the flux iron yoke of the solenoid.

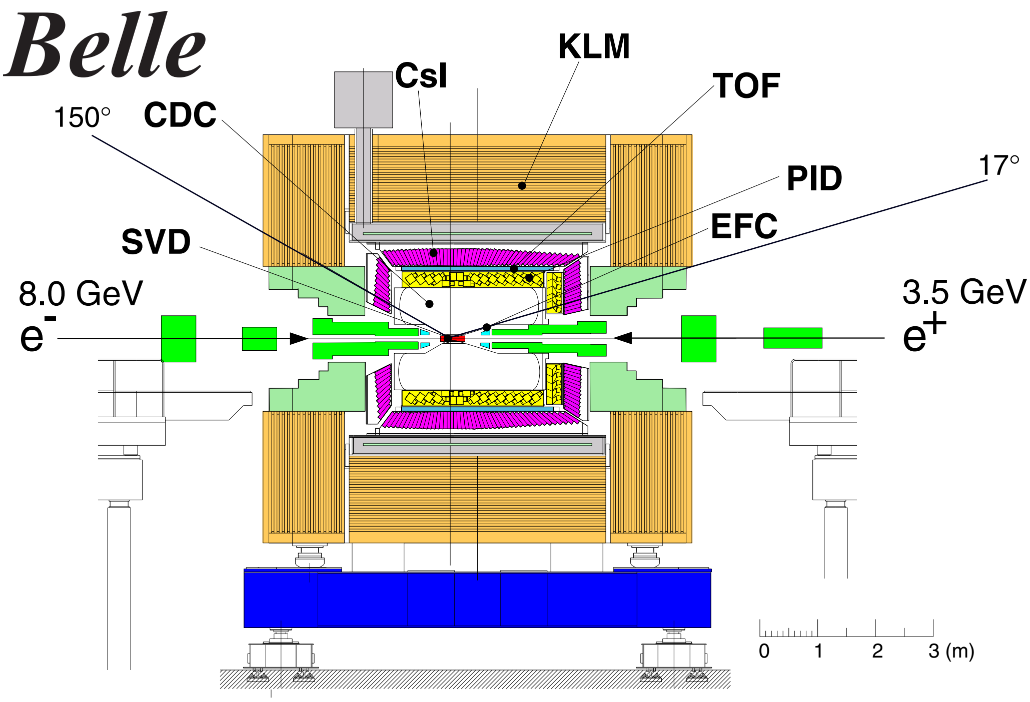


FIGURE 2.3: The layout of Belle detector [12].

2.2.1 Beam pipe and Interaction Point

At the IP the e^+ beam traverses anti-parallel to the $+z$ axis while the e^- beam crosses it with an inclination of 22 mrad. The beam pipe is made up of two coaxial beryllium pipes of thickness 0.5 mm leaving an annular gap of 2.5 mm between them [12]. It is 14.7 cm long with an inner radius of 20 mm. The beam pipe is made thin and placed close to the IP as possible in order to improve the vertex position resolution of charged tracks. Helium gas is circulating through the gap to cool the beam pipe to prevent potential damage to the inner layer of the SVD. The outer beryllium pipe is gold plated (20 μm) to reduce X-ray background. A cross-sectional view of the beam pipe is shown in Fig. 2.4. The superconducting final-focus quadrupole magnets (QCS) with a magnetic field of 4.8 T are located near the IP and common for both beams. Their axes are kept parallel to the incoming e^+ and e^- beams for reducing the synchrotron radiation. The normal conducting magnets, QC1's, are located outside QCS to provide vertical focus, only for the high-energy beam. A gold-plated radiation mask near the IP allows the synchrotron radiation generated by QCS and QC1 to pass through without hitting the beam-pipe. Several particle masks are also installed outside the beam pipe to reduce background.

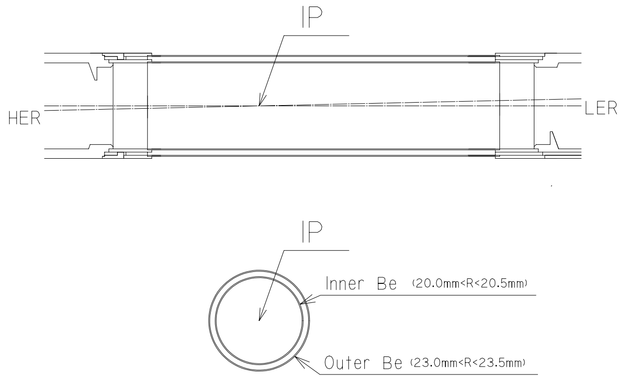


FIGURE 2.4: Cross-sectional view beam pipe at the interaction point [12].

2.2.2 Extreme Forward Calorimeter

The extreme forward Calorimeter (EFC) is placed very close to the IP, and composed of an array of radiation-hard Bismuth Germanate (BGO) crystals arranged into forward and backward cones around the beam pipe. The arrangement of BGO crystals in two cones are shown in Fig. 2.5. The angular coverage of the EFC ranges between 6.4° to 11.5° in the forward direction and 163.3° to 171.2° in the backward direction. It works as a beam mask to reduce background as well as a luminosity monitor for the Belle experiment (by measuring the rate of Bhabha events). In the EFC, each crystal has a trapezoidal shape kept inside 1mm thick stainless steel container and points towards the IP. Silicon photo-diodes are glued to the rear end of the crystal for detecting the scintillation light. The detector is segmented into 23 sections in ϕ and 5 sections in θ for both forward and backward cones, providing a good position resolution. The BGO has an excellent e/γ energy

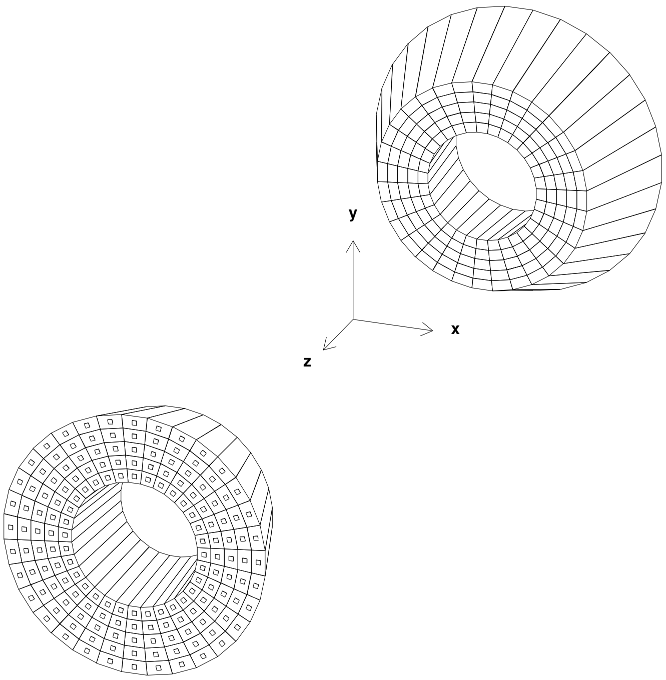


Figure 2.5: Configuration of BGO crystals in the Extreme Forward Calorimeter [12].

resolution of $(0.3-1)\%/\sqrt{E(\mathrm{GeV})}$ and is capable of detecting minimum ionizing particles with a large signal-to-noise ratio [16].

2.2.3 Silicon Vertex Detector

As described earlier, the main goal for Belle was to perform various time-dependent CP asymmetry measurements. Towards this goal, the time intervals are obtained by measuring the distance between the position of decay vertices. The SVD is designed to perform the vertex position measurements with a precision of 100 μm. Other than vertexing, its information are used in tracking by combining with the ones obtained from CDC. Since multiple columb scattering is the dominant source of vertex resolution degradation, the innermost layer of SVD is kept close to the IP. The initial version of the Silicon Vertex Detector (SVD 1) had 3 layers , which was later upgraded to 4 layer (SVD 2). A cross-sectional view of SVD1 is shown in Fig. 2.6. The support structures are low mass stiff and the readout electronics

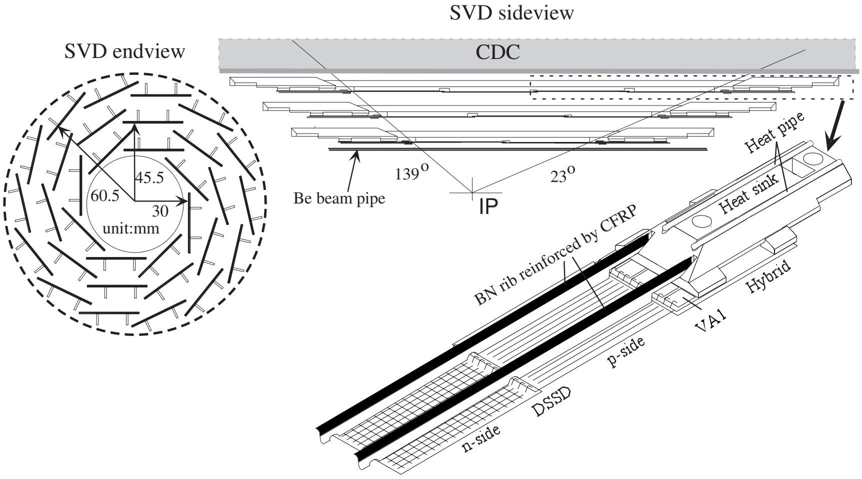


Figure 2.6: The configuration of the initial version of the silicon vertex detector (SVD1) [17].

are kept outside the tracking volume. The SVD covers about 86% of the total 4π solid angle. Each layer is constructed from ladders of double-sided silicon strip detectors (DSSDs). Each DSSD consists of 1280 sense strips and 640 readout pads on each side. The z-strip pitch is 42 μm while the ϕ-strip pitch is 25 μm. All sense strips are biased by a 25 MΩ bias resistor. The overall performance of the detector is verified by checking the consistency of the measured lifetime of D^0 mesons with its world average value [17].

2.2.4 Central Drift Chamber

The Central Drift Chamber (CDC) is a gas filled detector with a distinct arrangement of conducting wires. The charged particles are detected by the ionization of the gas followed by the collection of those ions in the wires. The CDC provides an efficient reconstruction of charged tracks as well as a precise determination of their momenta. It also performs the dE/dx measurement for particle identification and is an important component for triggering. The chamber is asymmetric in the z direction and has an angular coverage of $17° \leq \theta \leq 150°$. The inner

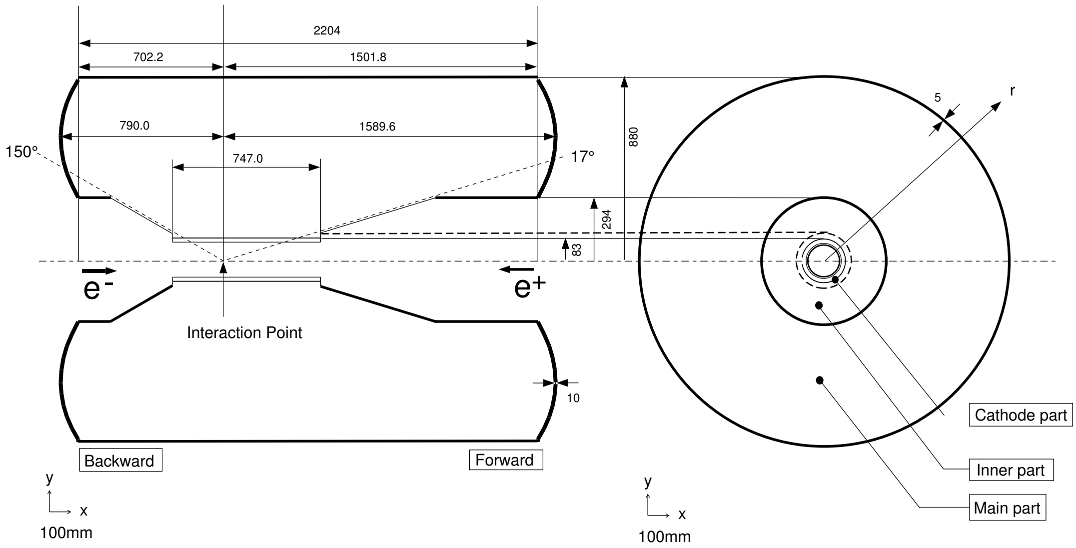


FIGURE 2.7: The structure of the central drift chamber [12].

radius is 80 mm without any wall for obtaining a good tracking efficiency for low-momentum tracks (this is particularly important in reconstructing the slow pions from D^{*+}). As shown in Fig. 2.7, the CDC has three parts: a cathode part, the inner part (small r region) with a conical shape to clear accelerator components, and the main part with a curved profile for minimizing the distortion caused by wire tension. The chamber has 50 cylindrical layers organized in 11 super-layers, each containing between three to six axial or small-angle-stereo layers, and three cathode strip layers. In total, the CDC has 8400 drift cells with individual cells being nearly square. Each cell is made up of aluminum field wires of $126\mu m$ and gold-plated tungsten sense wires of $30\mu m$. The wire configuration forming CDC cells is shown in Fig. 2.8. A low-Z helium-ethane gas mixture in the 1:1 ratio is used in oder to reduce multiple scattering. With this set up, we achieve a momentum resolution of $(\sigma_{p_T}/p_T) = (0.201 \pm 0.003)\% p_T \oplus (0.290 \pm 0.006)\%/\beta$ [18], where p_T is the transverse momentum in GeV/c and β is the velocity of the charged particle.

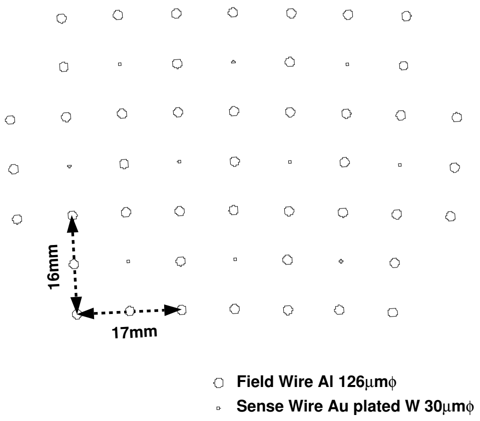


FIGURE 2.8: The wire configuration inside the central drift chamber [18].

2.2.5 Aerogel Cherenkov Counter

The aerogel cherenkov counter (ACC) is a powerful component of Belle used for a clear separation between the charged kaons and pions in the high-momentum region, 1.2 to 3.5 GeV/c. The Cherenkov radiation is produced by a dielectric material, silica aerogel (a colloidal form of $Si0_2$), due to traversing charged particle. The kaon-pion separation is achieved by measuring the difference in photon yield of the Cherenkov radiation produced by the charged particles with different masses. Here, the refractive index of the dielectric is chosen such that the pions give the radiation while the kaons do not. The silica aerogel tiles enclosed in a 0.2 mm thick aluminum box with approximate dimension of 12×12×12 cm^3 constitute a single ACC cell. Each cell is equipped with one or two fine-mesh photomultiplier tubes (FM-PMTs), which are specially designed to work in a 1.5 T magnetic field.

The ACC is installed in both the barrel (BACC) and forward endcap (EACC), as shown in Fig. 2.9. The BACC is comprised 960 cells that are segmented into 60 cells along the ϕ (z) direction. The refractive indices of the modules have a polar angle dependence, ranging from 1.010 to 1.028 in the backward direction (opposite the z axis). Note that aerogel with a smaller refractive index is required in the forward region, where momenta are largest. PMTs of three different diameters: 3, 2.5, and 2 inches, depending on refractive index, are attached with an ACC module in order to obtain uniform response for relativistic particles. The EACC has 228 cells arranged in five concentric rings with 60, 48, 48, 36, and 36 ϕ-segments. In the endcap region a single aerogel system with refractive index 1.030 is employed, which eliminates the need for an endcap TOF system. The BACC provides a 3σ K/π separation in the momentum range $1.0 < p < 3.6$ GeV/c, while the EACC provides the same separation in the range $0.7 < p < 2.4$ GeV/c.

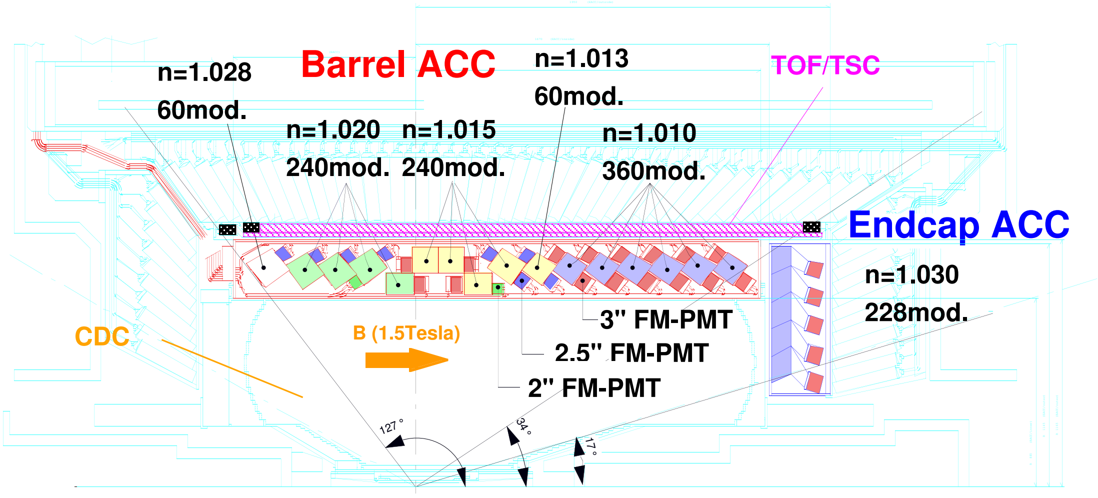


FIGURE 2.9: Arrangement of ACC modules in the Belle detector [12].

2.2.6 Time-of-flight counters

The time-of-flight detector (TOF) is a system of plastic scintillators used for the particle identification and triggering in the Belle detector. It comprises 64 barrel TOF/TSC (“trigger scintillation counter") modules. A TOF/TSC module consists of two trapezoidal shaped TOF counters and one TSC counter (in total 64,

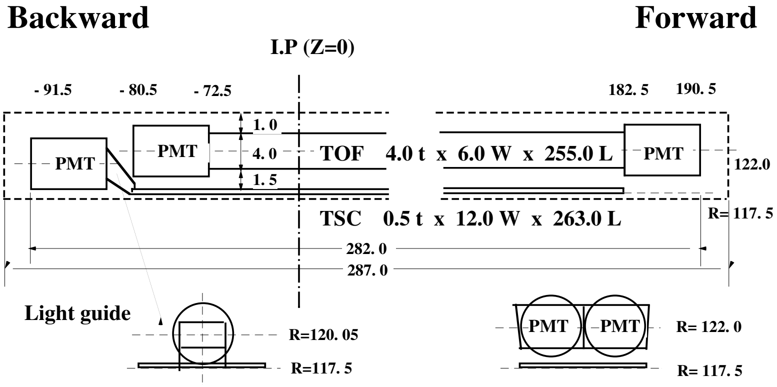


FIGURE 2.10: The dimensions of TOF/TSC module [19].

TSC and 128 TOF counters) with an intervening radial gap of 1.5 cm, which is introduced to suppress photon conversion background by taking a coincidence between the TSC and TOF counters. The dimension and arrangement of TOF/TSC modules is shown in Fig. 2.10. These modules cover a cylinder of radius 120 cm and length 255 cm with a polar angle coverage, $34\,^{\circ} < \theta < 121^{\circ}$. An FM-PMT is mounted directly on the TOF and TSC scintillation counters. Each TOF counter has a PMT at each end while the TSC counters have a PMT only at the backward end, connected via a light guide. The PMTs with 2-inch diameter, have 24 dynode stages and provide a gain of $3{\times}10^6$ at a high voltage below 2800 V. For a 1.2 m flight path, the TOF system with 100 ps time resolution provides a 2σ K/π separation up to 1.2 GeV/c [19]. The intrinsic resolution of TOF is found to be 80 ps, which is subject to many degradations; 10 ps due to 1.5 T magnetic field, 20 ps due to 4mm bunch length and jitter in the RF signal used as the reference time, and 20 ps from the readout electronics. By adding these uncertainties in quadrature, a time resolution of 96 ps is obtained, which meets the design goal of 100 ps.

2.2.7 Electromagnetic Calorimeter

The main purpose of the electromagnetic calorimeter (ECL) is to detect photons and electrons with high efficiency and measure their energies and positions with

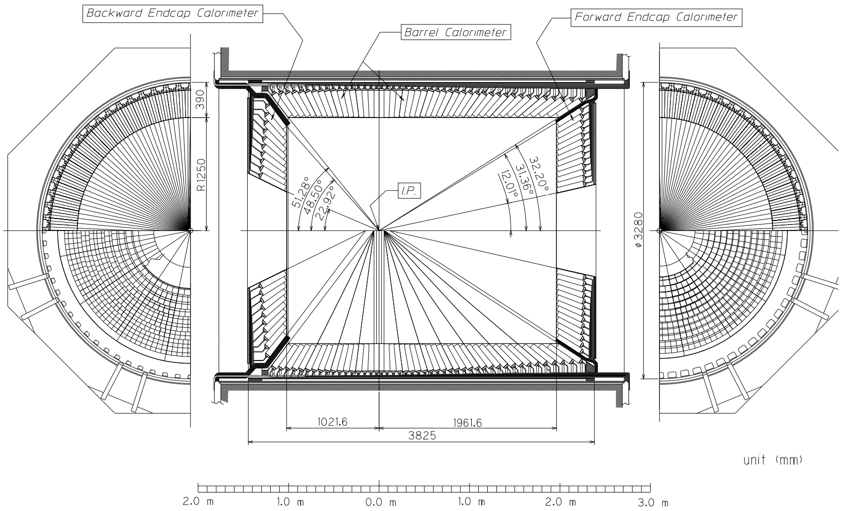
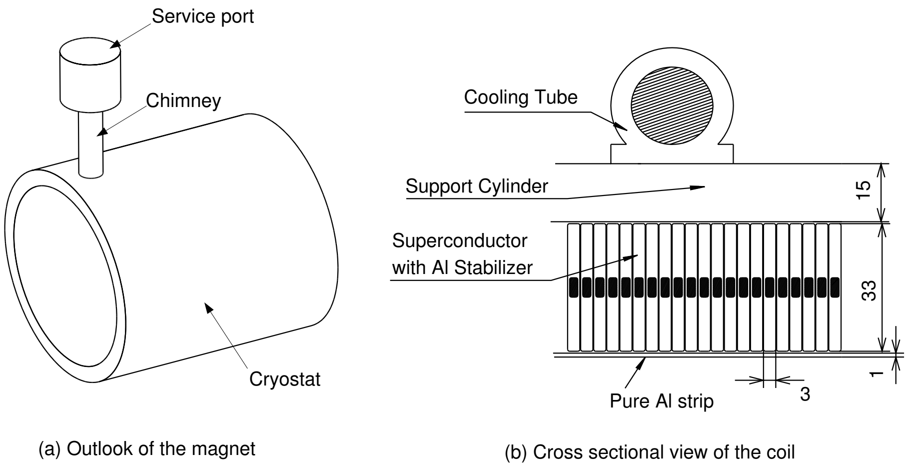


FIGURE 2.11: Overall structure of the electromagnetic calorimeter [12].

good resolution. In the ECL, a photon is identified as a cluster of energy deposits resulting from the electromagnetic shower while an electron as the cluster with an associated charged track. Good energy resolution results in a better hadron rejection. The detection of unstable particles like π^0 and η calls for a good separation between two daughter photons and a precise determination of the opening angle between them. These requirements are met with the use a fine-grained segmented array of CsI(Tl) crystals with silicon photodiode readout. They are organized into a barrel part with 6624 crystals and forward and backward endcaps with 1152 and 960 crystals, respectively [12]. A gap between the barrel and endcap region provides a pathway for the cables from inner subdetectors. A side- and end-view of ECL is shown in Fig. 2.11. All crystals are 30 cm (16.1 radiation length) long and assembled into a tower-like structure. Each of them is projected to the IP with a small tilt angle of 1.9° in both θ and ϕ directions to prevent any photon from escaping through the gaps between two adjacent crystals. The ECL covers 91% of the total solid angle. The energy scale of the system is calibrated by cosmic muons. With this setup, an energy resolution of 1.7% is obtained for $e^+e^- \to \gamma\gamma$ events. The π^0 and η are reconstructed from their two γ final states with mass resolutions of 4.8 and 12.1 MeV/c^2, respectively [20].

2.2.8 Detector Solenoid and Iron Structure



(a) Outlook of the magnet

(b) Cross sectional view of the coil

FIGURE 2.12: An outlook of the Belle solenoid and the cross-sectional view of the coil [12].

A superconducting solenoid provides a magnetic field of 1.5 T in a cylindrical volume of 3.4 m in diameter and 4.4 m in length. This magnetic field inside the tracking system enable the momentum measurement of charged particles. An overview of the solenoid magnet is shown in Fig. 2.12. Superconducting solenoids are installed on both sides of the IP to compensate deleterious effect of the field on the optical properties of KEKB beams. The magnet coil is surrounded by a multilayer structure consisting of iron plates that serves as the return path of magnetic flux as well as an absorber material for KLM. The flux return is organized into a barrel part and two end yokes. The solenoid provides 1.5T magnetic field at the nominal operating current of 4200 A. The magnetic field inside the tracking volume was mapped and has been used to improve the momentum resolution of charged tracks [21].

2.2.9 K_L^0 and Muon Detector

The KLM is the outermost part of Belle detector and designed for K_L^0 and μ detection, as its name suggests. It consists of a barrel and two endcap regions.

The barrel region covers a range from 45° to 125° in the polar angle. The forward and backward endcap extend this range to 20° to 155°. Basically, the KLM is an alternating layer of resistive plate counters (RPCs) and 4.7 cm-thick iron plates. An RPC module comprises of two parallel-plate electrodes (resistivity $> 10^{10}$ Ω-cm) separated by a gas filled gap. An ionizing particle traversing the gap results in a local discharge of the plates and induces a signal on external pickup strips, which can be used to record the location and time of ionization. A cross-sectional view of an RPC super-layer is shown in Fig. 2.13. The iron plates provide a total of 3.9 nuclear interaction lengths of material for a particle, making lightly interacting particles like muons, interact with the detector material. In addition to iron plates, the ECL provide another 0.8 nuclear interaction length and plays an important role in K_L^0 detection. The K_L^0 interacts with iron or ECL to produce a shower of ionizing particles. The location of this shower determines the direction of K_L^0. However, a precise measurement of K_L^0 energy is not possible due to large fluctuations in the size of the shower. Muons travel much further with smaller deflection on average than strongly interacting hadrons. The structure of both barrel and endcap parts is a repetition of 47 mm thick iron plates and 44 mm thick slots. RPC modules are placed into those slots. There are 15 layers in the barrel and 14 layers in the endcap. The barrel is divided azimuthally into octants while the endcap is divided into quadrants.

2.3 Trigger System

The trigger system is widely used in high energy physics to rapidly decide which events to keep as only a small fraction of the total can be recorded by the data acquisition system. The system selects “interesting" events, which occur at a relatively low rate using simple criteria (information from various subdetectors) that should be retained for the physics analysis.

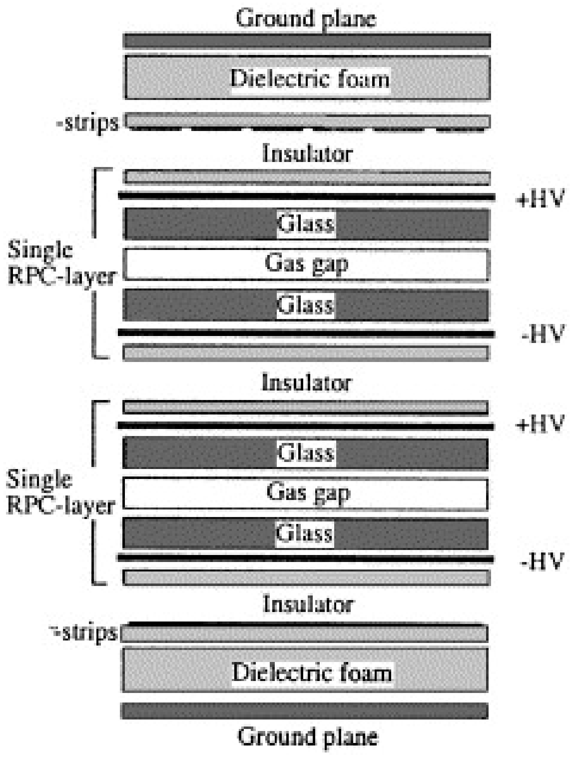


FIGURE 2.13: Schematic view of the cross-section of a KLM super-layer.

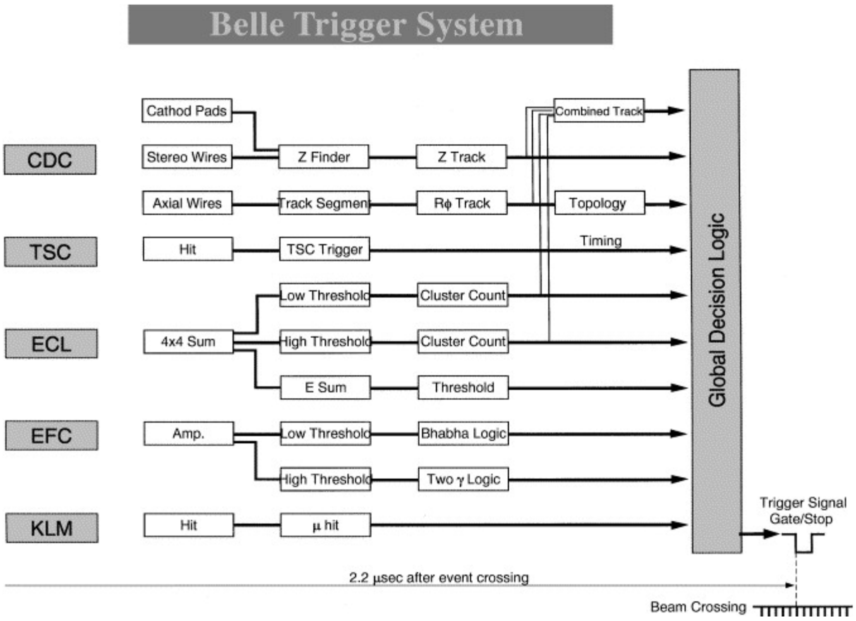


FIGURE 2.14: Trigger system of the Belle detector.

TABLE 2.2: Cross section and trigger rate for interesting events at a luminosity $10^{34}\text{cm}^{-2}\text{s}^{-1}$.

Physics process	Cross section (nb)	Rate(Hz)
$\Upsilon(4S) \to B\overline{B}$	1.2	12
$e^+e^- \to q\bar{q}$ $(q = u, d, s, c)$	2.8	28
$e^+e^- \to \mu^+\mu^-/\tau^+\tau^-$	1.6	16
Bhabha $(e^+e^- \to e^+e^-)$ $(\theta_{\text{lab}} \geq 17°)$	44	4.4[a]
$e^+e^- \to \gamma\gamma$ $(\theta_{\text{lab}} \geq 17°)$	2.4	0.24[a]
Two-photon events $(\theta_{\text{lab}} \geq 17°$, $p_T \geq 0.1$ GeV/c)	~ 15	~ 35[b]
Total	~ 67	~ 96

[a] prescaled by a factor 10^{-2}
[b] with restricted condition $p_T \geq 0.3$ GeV/c

Total cross sections for various interesting events and the corresponding trigger rates anticipated at luminosity, 10^{34} cm^{-2}-s^{-1}, are listed in Table 2.2. We need to accumulate Bhabha and $\gamma\gamma$ events to measure the luminosity and calibrate the detector response. The trigger rates for these two processes are prescaled by a factor 10^{-2} due their large cross section. The physics events have a trigger rate of 100 Hz while the simulation studies have shown that the beam related background has a rate $\sim$ 100 Hz. Hence for the luminosity $10^{34}\text{cm}^{-2}\text{s}^{-1}$, a trigger rate of $\sim$200 Hz is expected. The Belle trigger system is robust against unexpectedly high beam background rates.

In Belle the trigger system consists of a Level-1 hardware trigger and a Level-3 software trigger (implemented in the online computer farm). The Level-1 trigger comprises the subdetector trigger systems, and the central trigger system called the Global Decision Logic (GDL). A schematic diagram of the Belle Trigger system is shown in Fig. 2.14. The subdetector trigger systems are of two categories: track triggers and energy triggers. The CDC and TSC/TOF give trigger information from charged tracks. The innermost layers of CDC with cathode strips provide z-information for the trigger. The CDC provide $r-\phi$ and $r-z$ track information for trigger. The ECL gives a trigger signal based on the energy deposited in crystals.

In addition to this, signals from KLM and EFC give information on muons and two-photon events, respectively. The information from subdetectors are processed in parallel and later go to GDL, where all information are combined to characterize event type.

2.4 Data Acquisition System

The Belle Data Acquisition System (DAQ) is designed to record data up to 500 Hz at a typical event size of 30 kB. The DAQ reads the data from all subdetectors. All but for the SVD are read out by a TDC-based system which converts analog signals from subdetectors to digital by the charge-to-time (Q-to-T) conversion. In Q-to-T conversion the charge is converted to a pulse that has two edges where the leading edge shows the signal timing and the duration between two edges is the measure of charge. In order to handle large number of channels, the SVD has different readout system with all channels are digitized by flash ADC modules. The TDC modules did not have a pipe-line readout scheme, which results in large readout deadtime (around 30 μs). Several electronics upgrades were carried out in order to reduce the dead time during the latter parts of the experiment.

2.5 Particle Identification in Belle

The particle identification (PID) plays a key role in flavor tagging of neutral B and D mesons. The ideal characteristic of a PID system is high detection efficiency and low fake rate. The PID in Belle is performed by combining information from the ACC, TOF, CDC, ECL (for electron) and KLM (for muon).

2.5.1 Pion and Kaon Identification

The ACC, TOF and CDC are involved in distinguishing charged pion from kaons. These provide almost independent measurements and have good separation power in different momentum regions. The dE/dx measured in the CDC depicts a clear contrast between different particle species as shown in Fig. 2.15. The likelihood for each particle is defined assuming a Gaussian shape for the dE/dx distribution as:

$$P_i^{\mathrm{CDC}} = \frac{e^{-\chi_i^2/2}}{\sqrt{2\pi}\sigma_{dE/dx}}, \tag{2.1}$$

where $\sigma_{dE/dx}$ is resolution of dE/dx and χ_i^2 is given by

$$\chi_i^2 = \left(\frac{(dE/dx)_{\mathrm{meas}} - (dE/dx)_i}{\sigma_{dE/dx}} \right)^2 \tag{2.2}$$

where $(dE/dx)_{\mathrm{meas}}$ is the measured energy-loss and $(dE/dx)_i$ is the expected energy-loss for the ith particle species (e, μ, π, K, or p).

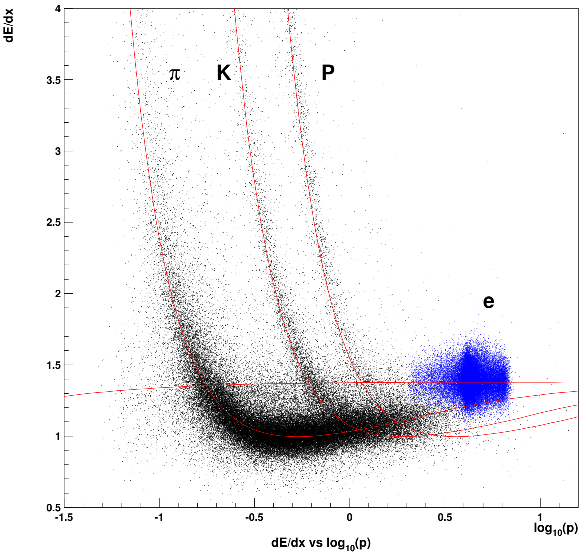


FIGURE 2.15: dE/dx vs momentum (in the log scale) for electron, pion, kaon and proton.

The flight time measurement in TOF provides a 3σ π/K separation below 1.2 GeV/c. The likelihood of the TOF system is defined assuming a Gaussian shape for the flight time distribution. The difference from the expected time assuming the ith particle is $\Delta_i^k = t_{meas.}^k - t_i^k$, where k is 0 or 1, which indicates PMT of two ends of a TOF counter. The χ^2 is defined as $\chi_i^2 = \Delta_i^T E^{-1} \Delta_i$, where Δ is a vector whose element is Δ_i^k and E is a 2×2 error matrix. Then the likelihood is given by

$$P_i^{\mathrm{TOF}} = \frac{e^{-\chi_i^2/2}}{\prod_{l=1}^{ndf} \sqrt{2\pi}\sigma_{\mathrm{TOF}}} \tag{2.3}$$

where σ_{TOF} is the TOF resolution.

The refractive index of ACC depends on the polar angle and is tuned for flavor tagging. As there is no TOF counter in the endcap, the ACC has to cover even a lower momentum region for the forward endcap. The likelihood of ACC for each particle species (P^{ACC}) is determined from the number of photoelectrons ($\mathrm{N_{pe}}$) distributions and probability distribution function (PDF) for each particle species expected from MC simulations.

Combining the individual, likelihood for each particle species, we get $P_i = P_i^{\mathrm{CDC}} \times P_i^{\mathrm{ToF}} \times P_i^{\mathrm{ACC}}$. In order to separate a signal particle i from a background particle j, the likelihood ratio is defined as $P(i:j) = P_i/(P_i + P_j)$. When particle i and j cannot be separated, $P(i:j)$ is set to0.5.

2.6 Belle Software and Computing System

As Belle is the highest luminosity detector, its computation system should be capable of promptly processing the data to provide them for user analyses. The Belle group developed a unique analysis framework called Belle Analysis Software (BASF) for different phases of event processing. It consists of many modules, written in C++, that are compiled as a shared object. The data handling is done with a traditional bank system, called PANTHER. A typical raw data event size is 35 kB (DST data) while a compact data set of an approximate size of 12 kB (mini-DST data) is produced for user analysis. Along with data we have Monte Carlo (MC) simulated events. The detector effect is included in simulations via the GEANT3 package. The background events in MC samples are calculated using beam data.

The Belle computing system comprises of four major elements. The first one is the KEKB computers, a principal system where data processing as well as analyses are performed. Then, we have a disk storage for keeping collision and MC data. For the network within Japan a super-SINET link is available since 2002. In addition, there are PC farms to enrich CPU for user analysis.

CHAPTER 3

Analysis Overview

This chapter provides an overview of the analysis strategy used in Belle experiment. We introduce various terms that are quite common within Belle and we shall use throughout the thesis. Further, we define some special variables used to discriminate interesting events.

3.1 Event Reconstruction

At KEKB, e^+e^- collisions lead to the production of many interesting bound states of quarks and antiquarks, and (free) leptons. The bound states include heavy resonances, mesons and baryons. Most of them are unstable and eventually decay to stable subatomic particles in a single step or several steps. Only 14 particles are seen by the detector while everything else are too short-lived to be detected. It include charged particles: $e^\pm$, $\mu^\pm$, $\pi^\pm$, $K^\pm$, and p ($\bar{p}$) and neutrals: γ, K_L, and n ($\bar{n}$). We reconstruct the unstable particles from their stable decay products, called daughters. The combination of all such short-lived intermediate states (and stable particles, where applicable) to reproduce an e^+e^- event is called event reconstruction, and it is achieved using some reconstruction algorithms. For example, consider a decay process $D^{*+} \to D^0(\to \pi^0\pi^0)\pi^+$. Each of the two π^0s mostly decays to a pair of photons. So the above decay chain is reconstructed from four photons and a charged pion. The reconstruction is performed based on a C++ program, called the reconstruction code. The code acts as a module in the BASF framework and takes the Belle data as input that in turn contain all types of events. The code reconstructs only the events under investigation and saves the output as an N-tuple in the “.root" format .

In this thesis we describe two analyses based on $D^0 \to \pi^0\pi^0$ and $D^0 \to \gamma\gamma$ decays. In both cases we require the D^0 (or $\overline{D}^0$) to come from a D^{*+} (or D^{*-}).

3.2 Signal and Background

The event reconstruction procedure described in Section 3.1 is generally not a 100% efficient algorithm. While the reconstruction is meant for a particular decay event, there could be a number of other processes (not exactly what we are looking for), that would pass our algorithm. For example, a decay $D^0 \to K^0_S\pi^0$ can be misidentified as a $D^0 \to \pi^0\pi^0$ event. There is a certain probability for a K^0_S to decay to a pair of π^0s ($30.69\pm0.05\%$ [26]). Misreconstruction of two π^0s, or loosing one of them in the reconstruction process for the $D^0 \to K^0_S(\pi^0\pi^0)\pi^0$ decay can mimic a $D^0 \to \pi^0\pi^0$ event. The correctly reconstructed events are called as *Signal* while the miss-reconstructed events that fake signal are known as *Background*.

3.3 Skimming

In Belle we have the collision data recorded by the detector, and few sets of MC simulated data where each set is statistically equivalent to the former. We run the reconstruction code on these two kinds of data to select the events under investigation. The data sample comprises a wide variety of events and is huge in size. On the other hand, during the analysis we need to run the reconstruction many times before converging on the final selection. The whole process takes a lot of time as well as computational resources. Therefore Belle uses an efficient procedure called skimming to avoid this unwanted wastage. In skimming, we generate a subset of the raw data by applying some basic selection criteria. These criteria are such that they reject many uninteresting events without any loss of signal. The skimming procedure is illustrated in Fig. 3.1, where the red dots represent signal while blue and green ones denote two types of backgrounds. The skimming is performed with a code that is similar to the reconstruction code, except that it gives an index file as the output. These index files contain the path to all events selected during skimming. The set of selection criteria used to produce skimmed data is called the skim-level selection. Once we are done with the skimming, we run the reconstruction code on the index files for further studies.

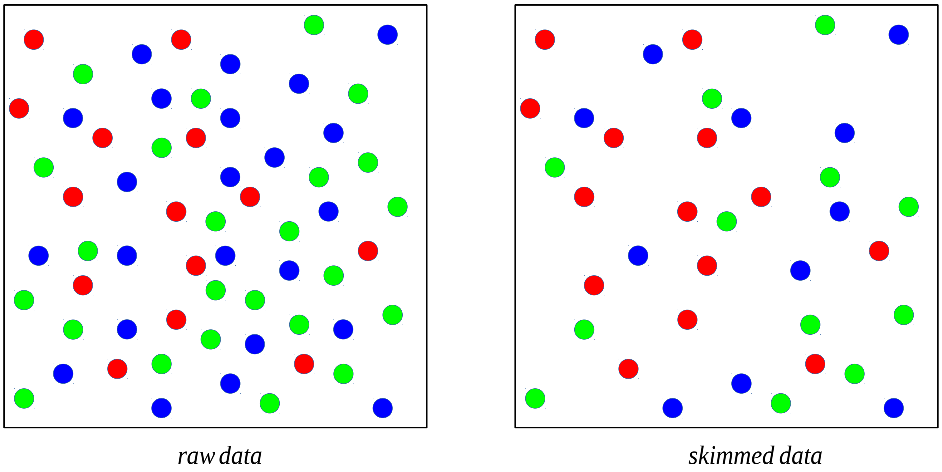


FIGURE 3.1: An illustration of the skimming procedure. The left plot represents the raw data sample with different types of events, while the right plot shows a skimmed sample in which the backgrounds (blue and green dots) are largely reduced largely without much loss to signal(red dots).

3.4 MC Simulated Data

The Belle detector records data by detecting and identifying stable subatomic particles produced in e^+e^- collisions. The quality of data depends on the ability of the detector to identify various processes without much misidentification. It is desirable to use the MC simulated data that have an extra advantage of containing the information without and with the detector effects. These simulated data are called generic MC events and produced by considering all possible outcomes of an e^+e^- collision, based on current physical theories. The generic MC samples are produced in two consecutive steps: generation and simulation. At the generation stage, events are randomly generated based on physics prediction. Then various detector effects are incorporated through a simulation process based on the GEANT3 package. The final generic samples are obtained by repeating the above procedures for individual processes followed by adding them according to their world-average branching fractions. We use these samples to determine various background processes as well as for the optimization of selection criteria. In addition to generic we use signal MC events that are the simulated data for a specific decay mode. The latter is generally used to determine the reconstruction efficiency for the given process.

3.5 MC Truth Matching

As we described in Section 3.4, the MC events are produced in a two-step process in which the generated events undergo a full simulation process to include detector effects. Even after reconstruction we can trace back the real physics process happened before including detector effect. The event information before adding the detector effect is called the generator level information. Using this information we can classify the selected events into either signal or various background processes.

This procedure of event classification is known as the MC truth matching, and can be done in the reconstruction code using some standard packages available in Belle.

3.6 Discriminating Variables

The detection and identification of subatomic particles in the Belle detector are achieved by the measurement of many observables like momentum, mass etc.. For every charge particle, we measure its momentum in the tracking system, and mass (or, velocity) using the PID system. For neutrals we use calorimetry to determine the energy of the particle. These information are combined appropriately during reconstruction, and different processes are characterized by the distinct values of these observables. In addition to the normal ones, we use many other special observables to distinguish signal from background.

3.6.1 $M(D^0)$

$M(D^0)$ is the reconstructed invariant mass of the D^0 candidate. Depending on the decay mode, this could be the invariant mass of twoπ^0's (for $D^0 \to \pi^0\pi^0$) or two γ's (for $D^0 \to \gamma\gamma$).

3.6.2 ΔM

ΔM is the mass difference between the D^{*+} and D^0 candidates. Irrespective of the final state, we use this variable for our studies.

3.6.3 $p^*(D^*)$

$p^*(D^*)$ is the momentum of the D^{*+} (or D^{*-}) candidate in the e^+e^- centre-of-momentum frame.

3.6.4 Impact Parameters

Selection criteria on spatial coordinates around the IP are applied to reject events from parasitic e^+e^- collisions. The dr and dz are the impact parameters perpendicular and parallel to the z axis of the Belle detector as shown in Fig 3.2. The first one is defined as the radial distance perpendicular to the z axis while the second is the distance along the z-axis about the IP ($z = 0$). The selection requirements on dr and dz retain the particles coming from a region near to the IP.

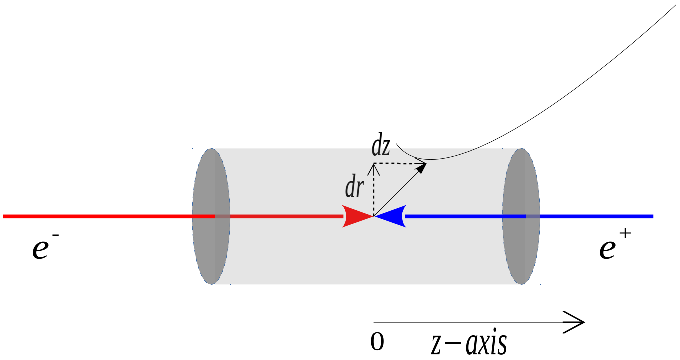


FIGURE 3.2: Schematic diagram of IP profile. The dr and dz are defined about the IP as shown here. The gray shaded region represent the cut: $dr = 1$ cm and $|dz| = 3$ cm.

3.6.5 E9/E25

The electrons and photons are detected in the CsI(Tl) crystals or, cells of the ECL. These detections are materialized via the production of an electromagnetic shower in the cells. Generally the shower is not confined to a single cell rather it is spread over a number of adjacent cells. The spread is almost uniform around a seed cell: the cell with the highest deposition of energy. E9/E25 is defined as the ratio of energy deposited in 3×3 crystals to that in 5×5 crystals about the seed, and it gives a measure of lateral spread of the shower. The shower can be initiated by hadrons (hadronic shower) and can be misidentified as an electromagnetic shower. However, the hadronic shower is relatively wider in the lateral dimension as compared to the electromagnetic shower. The E9/E25 variable can be used to distinguish between them. The electromagnetic showers are characterized by a higher value (typically in the range [0.85, 1]) while the hadronic ones have smaller values (ranging [0, 0.85]) of E9/E25.

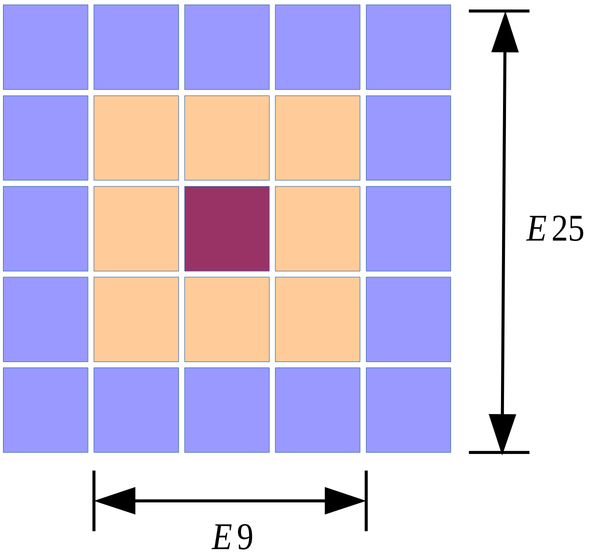


Figure 3.3: An illustration for the variable E9/E25. Figure shows the 5×5 crystal block (all cells within violet filled area), 3×3 crystal block (all cells inside the gold filled area) and the seed crystal (magenta square).

3.6.6 KID

The PID and various detector components contributing to it are described in Section. 2.5. Based on these we define a kaon identification likelihood (KID), given by the following ratio:

$$\text{KID} = \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi}, \tag{3.1}$$

where $\mathcal{L}_K$ ($\mathcal{L}_\pi$) is the likelihood of a charged track being due to a kaon (pion). The likelihood is calculated using dE/dx from the CDC, information from the TOF, and number of photoelectrons in the ACC. The KID is the probability for a given charged particle to be a kaon, as opposed to being a pion.

CHAPTER 4

CP violation in $\boldsymbol{D^0 \to \pi^0\pi^0}$

4.1 Introduction

In this chapter, we report a search for CP violation in the decay of neutral D mesons to the CP eigenstate $\pi^0\pi^0$. The measurement of time-dependent CP asymmetry is not feasible here since the final state comprises two neutral particles[1]. Assuming the total decay width to be same for particles and antiparticles, the time-integrated CP asymmetry is defined as:

$$A_{CP} = \frac{\Gamma(D^0 \to \pi^0\pi^0) - \Gamma(\overline{D}^0 \to \pi^0\pi^0)}{\Gamma(D^0 \to \pi^0\pi^0) + \Gamma(\overline{D}^0 \to \pi^0\pi^0)}, \tag{4.1}$$

where Γ represents the partial decay width. The A_{CP} term has got the following three contributions:

$$A_{CP} = A^d_{CP} + A^m_{CP} + A^i_{CP}, \tag{4.2}$$

where A^d_{CP} denotes the direct CP violation contribution which is decay-mode dependent, and A^m_{CP} and A^i_{CP} represent CP violation in mixing and due to interference between decays with and without mixing, respectively. These latter

---

[1] Time-dependent CP asymmetry measured in the bins of the lifetime of D mesons, which is achieved by measuring the decay vertices of the meson. For a final state comprising only neutral particles it is impossible to perform this study as one cannot reconstruct the decay vertices.

contributions are independent of the D^0 decay final states and related to the lifetime asymmetry $A_\Gamma = \frac{\tau(\bar{D}^0 \to \pi^0\pi^0) - \tau(D^0 \to \pi^0\pi^0)}{\tau(\bar{D}^0 \to \pi^0\pi^0) + \tau(D^0 \to \pi^0\pi^0)}$, as:

$$A_{CP}^m + A_{CP}^i = -A_\Gamma \tag{4.3}$$

A previous study of A_Γ by Belle [22] reported an asymmetry consistent with zero, $A_\Gamma = [-0.03 \pm 0.20(\text{stat}) \pm 0.08(\text{syst})]\%$.

Our study is the first attempt to measure $A_{CP}(D^0 \to \pi^0\pi^0)$ at the B-factories that uses the whole data sample (966 fb^{-1}) recorded by Belle. Earlier the CLEO Collaboration has reported a CP asymmetry of $(+0.1 \pm 4.8)\%$ using a 13.7 fb^{-1} data sample [23], whereas BaBar [24] has studied $D^0 \to \pi^0\pi^0$ with 471 fb^{-1} providing only its branching fraction.

In this analysis, we use a one dimensional fitting to the ΔM distribution for estimating A_{CP}. We repeat the A_{CP} measurement for $D^0 \to K_S^0\pi^0$, previously performed in Belle [25], to get a better understanding of the A_{CP} extraction procedure as well as to check the reconstruction and fitting codes. The previous results for this mode were $A_{CP} = (-0.28 \pm 0.19)\%$.

4.2 Method

The most important requirement to measure A_{CP} is to know the initial flavor of the D meson, whether it is a D^0 or a $\bar{D}^0$, known as flavor tagging. In our case, both D^0 and $\bar{D}^0$ decay to the same final state and hence we cannot distinguish between them by just looking at their final decay products. The flavor tagging is achieved by reconstructing the decay chain $D^{*+} \to D^0\pi_s^+$ (or $D^{*-} \to \bar{D}^0\pi_s^-$). The charge of the accompanying pion determines the flavor of the D meson. Here the pion is a low-momentum (or soft) one, which is indicated by a subscript "s". For extracting A_{CP} we measure asymmetry in the yields of D^{*+} and D^{*-}, called the raw asymmetry (A_{rec}), defined as:

$$A_{\text{rec}} = \frac{N(D^{*+}) - N(D^{*-})}{N(D^{*+}) + N(D^{*-})} \tag{4.4}$$

The above term can be written as a sum of three contributions:

$$A_{\text{rec}} = A_{FB} + A_{CP} + A_{\epsilon}^{\pi_s}, \tag{4.5}$$

where A_{FB} is the forward-backward asymmetry, A_{CP} is the intrinsic CP asymmetry and $A_{\epsilon}^{\pi_s}$ is the asymmetry in detection efficiencies of the oppositely charged soft pions. The A_{CP}, is independent of any kinematic variables.

The asymmetry $A_{\epsilon}^{\pi_s}$ can be determined using a self-tagged decay mode such as $D^0 \to K^-\pi^+$ ($\overline{D}^0 \to K^+\pi^-$). Here, unlike for $D \to \pi^0\pi^0$, the final state itself reveals the flavor of initial D meson while the flavor tagging based on soft pion is a complementary step. In the first step D^0 ($\overline{D}^0$) candidates are reconstructed from the $K\pi$ final states. The asymmetry in yields of D^0 and $\overline{D}^0$ is estimated as:

$$A^{\text{untag}} = A_{FB}^{D^0} + A_{CP}^{K\pi} + A_{\epsilon}^{K\pi} \tag{4.6}$$

In the second step, the D^0 ($\overline{D}^0$) is reconstructed from the $K\pi$ final state and subsequently combined with a $\pi_s^+(\pi_s^-)$ to form D^{*+} (D^{*-}). The asymmetry in yields of D^* candidates is estimated as:

$$A^{\text{tag}} = A_{FB}^{D^0} + A_{CP}^{K\pi} + A_{\epsilon}^{K\pi} + A_{\epsilon}^{\pi_s} \tag{4.7}$$

Thus $A_{\epsilon}^{\pi_s}$ can be obtained by subtracting A^{untag} from A^{tag}. Here the calculation is carried out in the two-dimensional bins of $[p(\pi_s),\ \cos\theta(\pi_s)]$, as the detection efficiency is a function of the pion momentum and also the detector acceptance for

Belle is asymmetric in θ. We can thus estimate the raw asymmetry for $D^0 \to \pi^0\pi^0$ in bins of $[p(\pi_s), \cos\theta(\pi_s)]$ and subtract $A_\epsilon^{\pi_s}$ to get the corrected asymmetry (the one corrected for detection efficiency asymmetry).

$$A_{\text{rec}}^{\text{cor}} = A_{\text{rec}} - A_\epsilon^{\pi_s} \tag{4.8}$$

The above asymmetry consists of two terms, the intrinsic CP asymmetry as well as the forward-backward asymmetry, A_{FB}, in the production of $D^{*\pm}$ mesons in $e^+e^- \to c\bar{c}$ arising from the $\gamma - Z^0$ interference and higher order QED effects. Since CP is conserved in the D^* production (it is a strong interaction process), the A_{FB} must be an odd function of $\cos\theta^*$, where θ^* is the polar angle of the D^* production in the centre-of-mass (CM) frame of the e^+e^- system, i.e.,

$$A_{FB}(\cos\theta^*) = -A_{FB}(-\cos\theta^*). \tag{4.9}$$

The above property of A_{FB} can be utilized to extract A_{CP} from $A_{\text{rec}}^{\text{cor}}$. The A_{CP} and A_{FB} are estimated by adding and subtracting bins at $\pm\cos\theta^*$ as

$$\begin{aligned} A_{FB} &= \frac{A_{\text{rec}}^{\text{cor}}(\cos\theta^*) - A_{\text{rec}}^{\text{cor}}(-\cos\theta^*)}{2}, \\ A_{CP} &= \frac{A_{\text{rec}}^{\text{cor}}(\cos\theta^*) + A_{\text{rec}}^{\text{cor}}(-\cos\theta^*)}{2}. \end{aligned} \tag{4.10}$$

4.3 Data samples

This analysis is based on a $966\,\text{fb}^{-1}$ data sample collected with the Belle detector at the $\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(3S)$, $\Upsilon(4S)$, and $\Upsilon(5S)$ resonances as well as the off-resonance data below the $\Upsilon(4S)$ and $\Upsilon(5S)$ peak. Although we are using the full data sample in the analysis, the generic MC events used for each stream (five different streams) correspond to only $\Upsilon(4S)$, $\Upsilon(4S)$ off-resonance, and $\Upsilon(5S)$ samples with integrated luminosities $710.5\,\text{fb}^{-1}$, $89.2\,\text{fb}^{-1}$, and $121.4\,\text{fb}^{-1}$, respectively. Studies using generic MC events such as optimization are performed separately for different CM energies while they are merged together for the estimation of $A_{\rm rec}$.

The signal decay sample is $D^{*+} \to D^0(\to \pi^0\pi^0)\pi_s^+$ while the $D^{*+} \to D^0(\to K^-\pi^+)\pi_s^+$ sample is used for estimating $A_\epsilon^{\pi_s}$ for the soft pion. As part of the study we do not estimate the $A_\epsilon^{\pi_s}$ map, instead we have used the ones calculated by B.R. Ko and M. Starič, given in Appendix A and B, respectively.

4.4 Peaking backgrounds in $D^0 \to \pi^0\pi^0$

The ΔM distribution for $D^0 \to \pi^0\pi^0$ after a loose cut on $M(D^0)$ exhibits a background peaking in the signal region. To find out the source of this peaking contribution, the generator level information in generic MC sample is studied. From the ΔM distribution shown in Fig. 4.1(left), we conclude that the peaking component is of the $D^0 \to K_S^0(\to \pi^0\pi^0)\pi^0$ origin. For this decay, the final state consists of three π^0's out of which two are combined to form a D^0 candidate. Since one π^0 is missing in the reconstruction, the peak is shifted to the lower side of the D^0 mass distribution as shown in Fig. 4.1(right). As the peaking background peaks outside the signal region in $M(D^0)$, a cut about the signal window can suppress much of the peaking background. Figure 4.1 shows that background of $K_S^0\pi^0$ origin has a peaking and a nonpeaking component. The peaking part corresponds to the D^0 candidates with both π^0's are from $D^0 \to K_S^0(\pi^0\pi^0)\pi^0$, whereas the nonpeaking part represents the D^0 candidates with only one π^0 is of $D^0 \to K_S^0(\pi^0\pi^0)\pi^0$ origin and the other is a random π^0. In order to remove the peaking contribution, we apply a tighter cut on the reconstructed D^0 mass, $1.758 < M(D^0) < 1.950$ GeV$/c^2$. The ΔM distribution after applying the cut on $M(D^0)$ is shown in Fig. 4.2 and one can see that the peaking component is largely reduced.

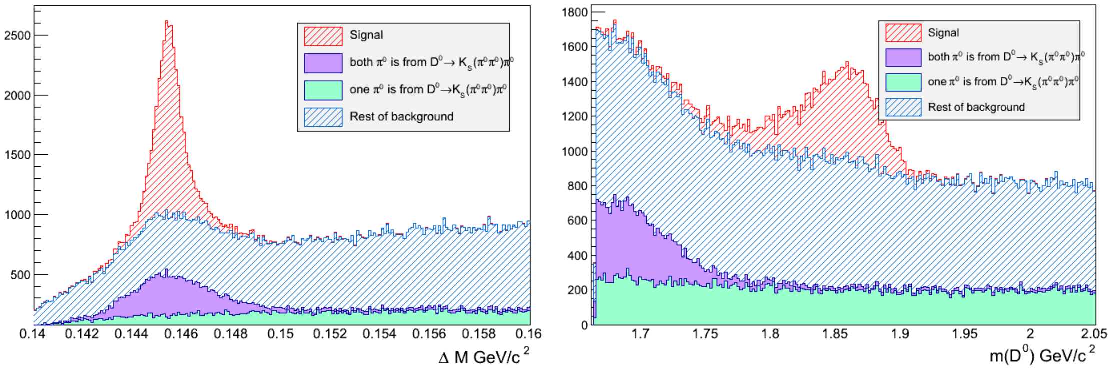


FIGURE 4.1: Distributions for $D^0 \to \pi^0\pi^0$: (left) ΔM distribution, where the $D^0 \to K^0_S\pi^0$ contribution is further classified into both π^0's coming from $D^0 \to K^0_S\pi^0$ and only one is of the $D^0 \to K^0_S\pi^0$ origin, and (right) the corresponding $M(D^0)$ distribution.

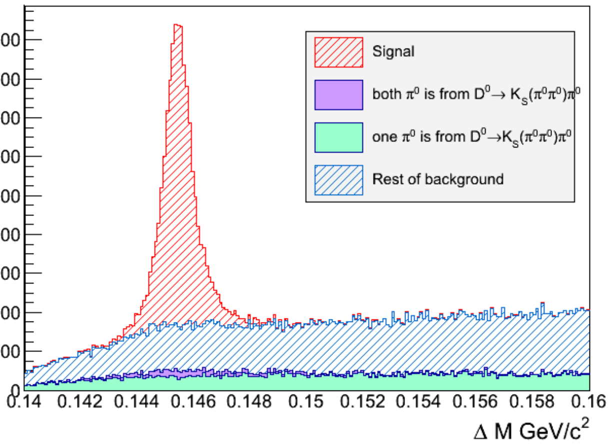


FIGURE 4.2: ΔM distribution for $D^0 \to \pi^0\pi^0$ sample after applying the cut $1.758 < M(D^0) < 1.930$ GeV$/c^2$.

4.5 Efficiency loss in π^0 reconstruction

The π^0 candidate in our selection is reconstructed from two photons, each of which is detected as an isolated cluster in the ECL. However, when one of them converts in the material of the detector before reaching the ECL, or when both of them are merged into a single cluster, the π^0 can't be reconstructed based on our standard reconstruction method. In this section, we describe these effects in detail and estimate the corresponding efficiency loss.

4.5.1 Converted photons

The conversion of a photon to an e^+e^- pair (pair production) is a phenomenon that affects the reconstruction efficiency of π^0. The neutral pion candidate for which one of the γ's converts in the material of the inner part of the detector (beam pipe, SVD, or CDC) will not be reconstructed by our selection. This effect is studied using the signal MC sample. Figure 4.3 shows two- and three-dimensional distribution of vertices of $\gamma \to e^+e^-$ from the generator level information. The efficiency loss due to this effect is proportional to the amount and density of material placed in front of the ECL. The efficiency loss for $D^0 \to \pi^0\pi^0$ reconstruction due to at least one photon being converted in the inner part of detector (up to a radius of 10 cm) is $\sim 25\%$ (efficiency loss per single photon is $\sim 6.5\%$).

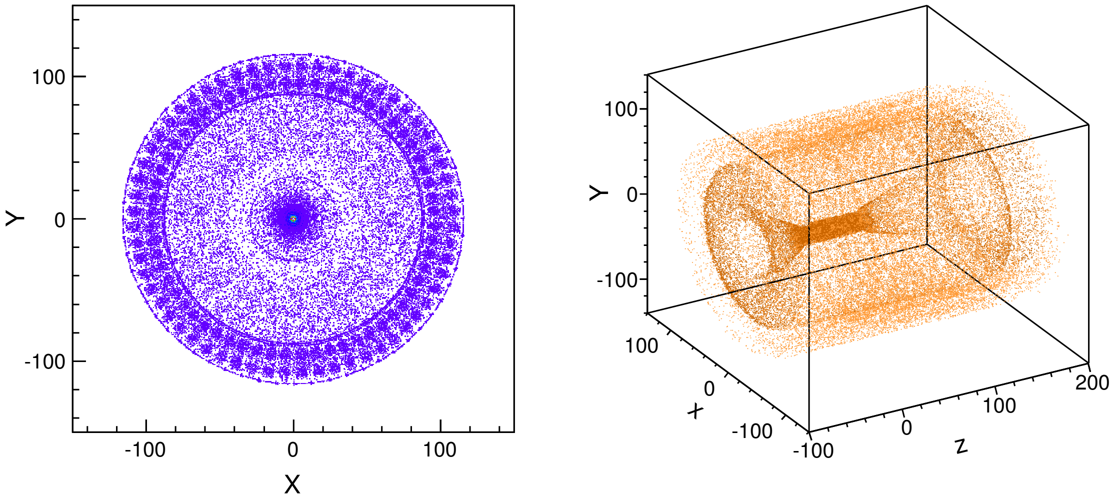


FIGURE 4.3: Vertex distribution at generator level for $\gamma \to e^+e^-$ in $D^* \to D^0\pi^+$, $D^0 \to \pi^0\pi^0$, $\pi^0 \to \gamma\gamma$.

4.5.2 Merged π^0's

The other major source of efficiency loss is due to merging of the two γ's, which are the decay products of the π^0. As the momentum of π^0 increases, the opening angle between two γ's decreases and they initiate showers in two adjacent cells of ECL eventually being detected as a single γ with high energy. This effect not only reduces the detection efficiency but also makes the cut on the shower shape variable $E9/E25 > 0.85$, the ratio of energy deposited by a shower in the 3×3 to 5×5 cell in the ECL, ineffective. Figure 4.4 shows a schematic of two overlapping ECL clusters for γ's coming from a high momentum π^0. The clustering algorithm assigns the overlapping cells, common to the two clusters, to one of them. So the cluster that gains an extra energy will have a smaller $E9/E25$. This effect can be seen in the correlation between the $E9/E25$ distribution of γ's in $\pi^0 \to \gamma\gamma$, for high momentum π^0 as shown in Fig 4.5(left). Figure 4.5(right) is the scatter plot between $E9/E25$ and the opening angle between two γ's in $\pi^0 \to \gamma\gamma$ decay ($\theta_{\gamma\gamma}$). If we apply a tight cut on $E9/E25$, say $E9/E25 > 0.85$, a considerable fraction of signal candidates having a smaller $\theta_{\gamma\gamma}$ (belong to high momentum π^0) will be lost. We estimate the efficiency loss by applying the cut $E9/E25 > 0.85$ to be 38%.

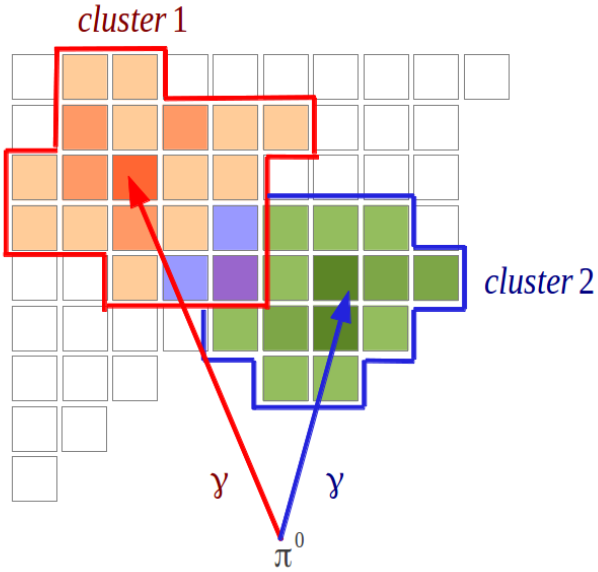


FIGURE 4.4: Schematic diagram of a high momentum $\pi^0 \to \gamma\gamma$ resulting in two adjacent ECL clusters.

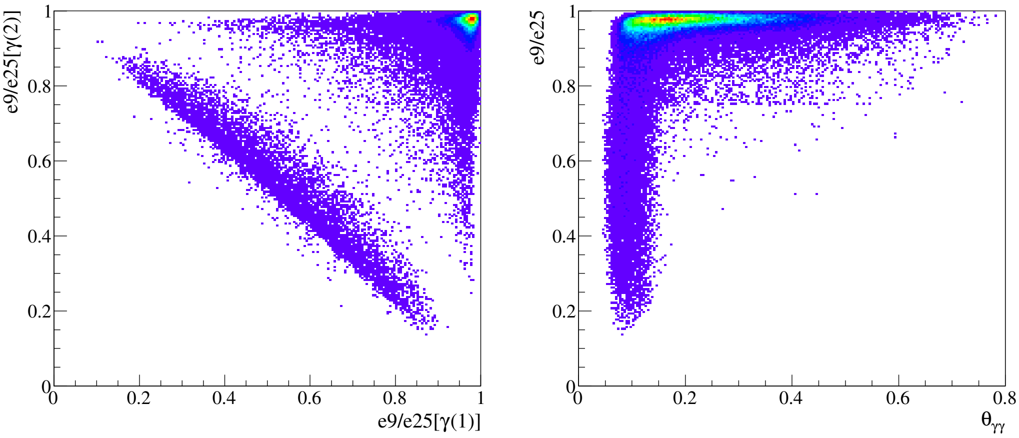


FIGURE 4.5: (left) $E9/E25$ of the two γ's from $\pi^0 \to \gamma\gamma$, (right) $E9/E25$ vs. $\theta_{\gamma\gamma}$.

4.6 Data-MC agreement

We check the overall agreement between data and MC events for the ΔM distribution. We define a signal region and sideband as follow (illustrated in Fig. 4.6):

$$\text{Signal region: } 0.144 < \Delta M < 0.147\,\text{GeV}/c^2$$
$$\text{Sideband: } 0.149 < \Delta M < 0.160\,\text{GeV}/c^2$$

A reasonable agreement for the signal component between data and MC events is observed in the signal region whereas the number of events observed in data clearly exceeds that expected from MC simulations in the sideband. Therefore, we need to correct for this mismatch while optimizing our selection cuts, described in the next section. The correction factor is simply defined as $(N_{\text{Data}}/N_{\text{MC}})$, estimated in the sideband (see Fig. 4.7).

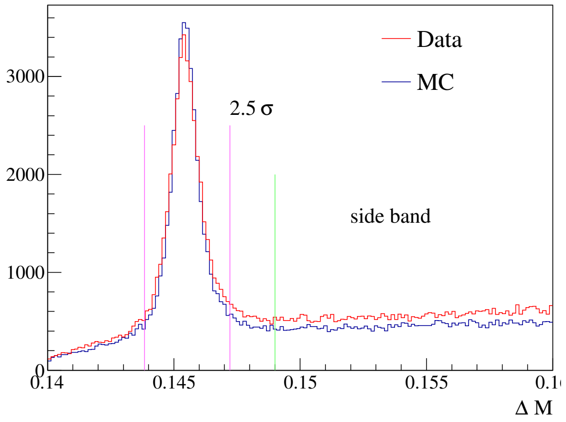


FIGURE 4.6: Signal region and sideband defined in the ΔM distribution.

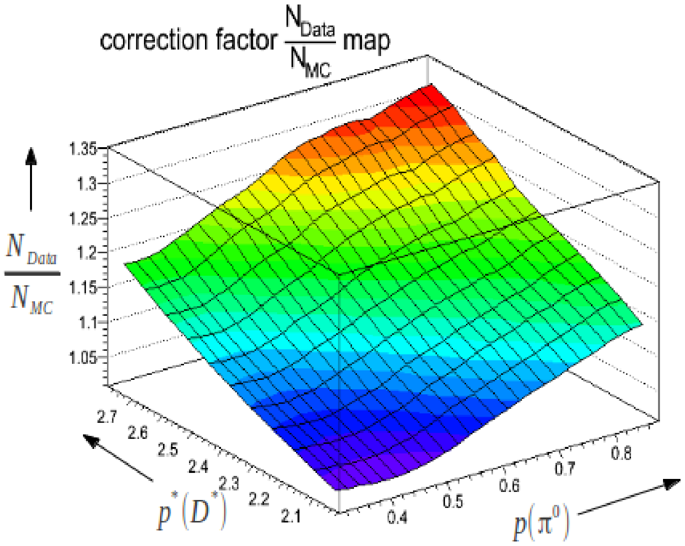


FIGURE 4.7: Correction factor for different $p(\pi^0)$ and $p^*(D^*)$ cut values.

4.7 Optimization of selection cuts

The selection optimization is performed by minimizing the expected error in the raw asymmetry:

$$\sigma_A = \frac{\sqrt{N_{\text{Sig}} + N_{\text{Bkg}}}}{N_{\text{Sig}}}, \tag{4.11}$$

while varying the cuts. Here, N_{Sig} and N_{Bkg} are the number of signal and background events, respectively, estimated from the generic MC sample in the signal region.Earlier to this, we have applied all skim level cuts retaining only the photon candidates with energy greater than 60 and 100 MeV for barrel and endcap regions of ECL, respectively. The MC truth matching information is used to distinguish signal from background.

Before doing optimization we first fine tune the width of the signal window. Figure 4.8 shows the result of signal-window optimization performed by varying $p(\pi^0)$ (momentum of π^0) while keeping $p^*(D^*)$ fixed. The optimization is carried out for different signal windows defined in terms of standard deviation of the corresponding ΔM distribution (σ_L and σ_R). The 2.5σ signal window used to estimate σ_A has an optimal point consistent with the blue curve [see Fig. 4.8(top right) and 4.8(bottom right)], which represents the error estimated by fitting the ΔM distribution in generic MC sample. Thus we chose the 2.5σ region as our optimal signal window.

A two-dimensional optimization is performed for $p^*(D^*)$ and $p(\pi^0)$ as there exists a strong correlation between them. Further, this is separately carried out for the $\Upsilon(4S)$ and $\Upsilon(5S)$ samples. We start the optimization from $p^*(D^*) > 2.5\,\text{GeV}/c$ for the $\Upsilon((4S))$ sample in order to suppress possible contamination of D^* candidates originating from the B-meson decays. The corresponding optimal cut position obtained are $p^*(D^*) > 2.5\,\text{GeV}/c$ and $p(\pi^0) > 0.64\,\text{GeV}/c$. The optimization result for $\Upsilon(4S)$ is shown in Fig. 4.9(left). A similar procedure is performed for the $\Upsilon(5S)$ sample after a cut $p^*(D^*) > 3.1\,\text{GeV}/c$ is applied. Subsequently, the optimal cuts $p^*(D^*) > 3.1\,\text{GeV}/c$ and $p(\pi^0) > 0.54\,\text{GeV}/c$ are obtained as shown in Fig. 4.9(right).

As mentioned in Section 4.6, there is a difference between data and MC expectation in the sideband. So a correction factor is estimated and applied during optimization to obtain cuts suitable for data. Figures 4.10 and 4.11 shows the correction

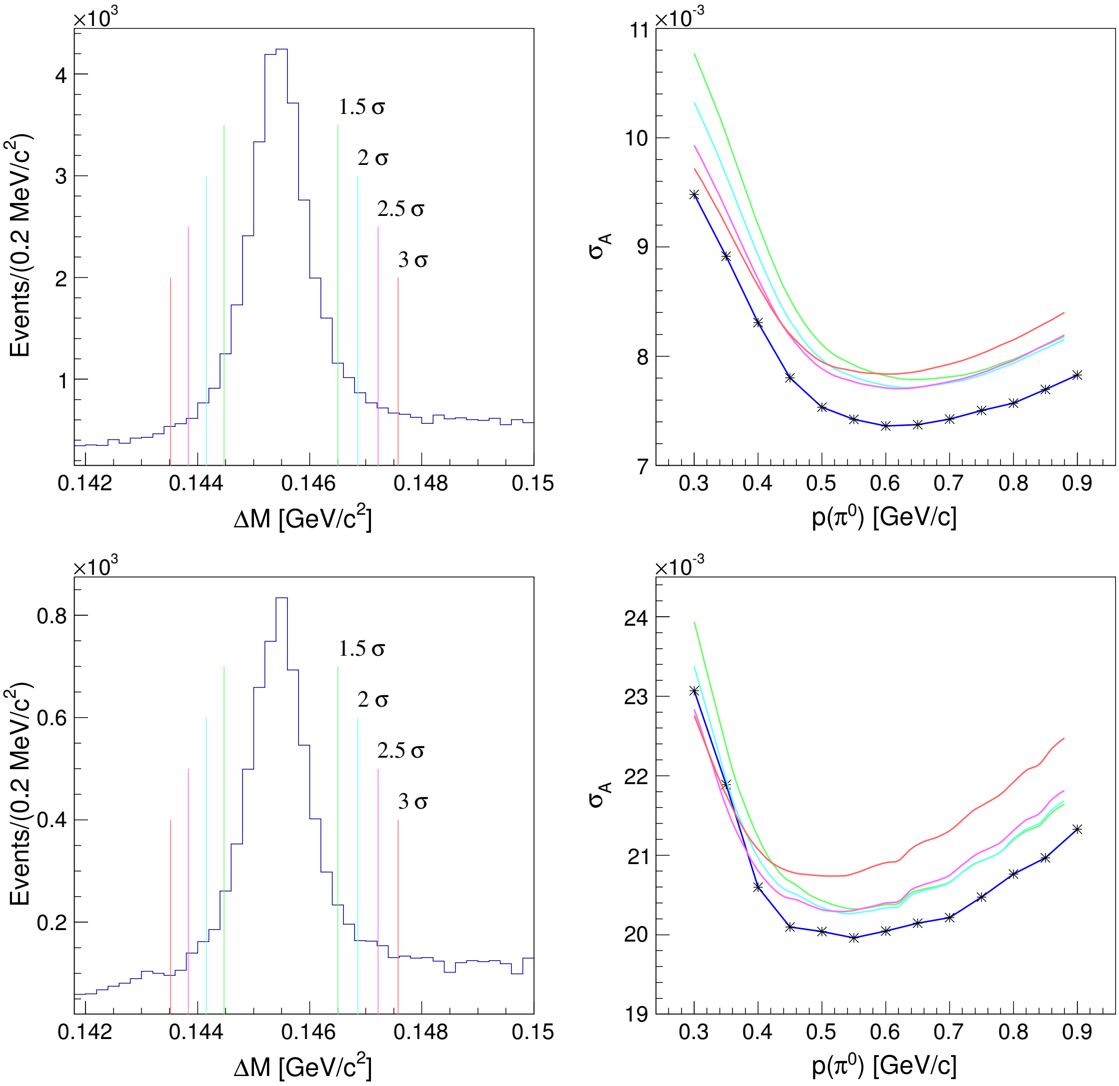


FIGURE 4.8: Signal window optimization: (top left) different signal windows in terms of standard deviation of the ΔM distribution for $\Upsilon((4S))$, (top right) optimization result for the $\Upsilon((4S))$ sample as a function of $p(\pi^0)$ keeping $p^*(D^*) > 2.5$ GeV/c fixed with different signal windows, (bottom left) different signal windows for the $\Upsilon((5S))$ sample, and (bottom right) optimization result for the $\Upsilon((5S))$ sample keeping $p^*(D^*) > 3.1$ GeV/c fixed. In both top right and bottom right the blue curve with star mark represents the optimal signal window.

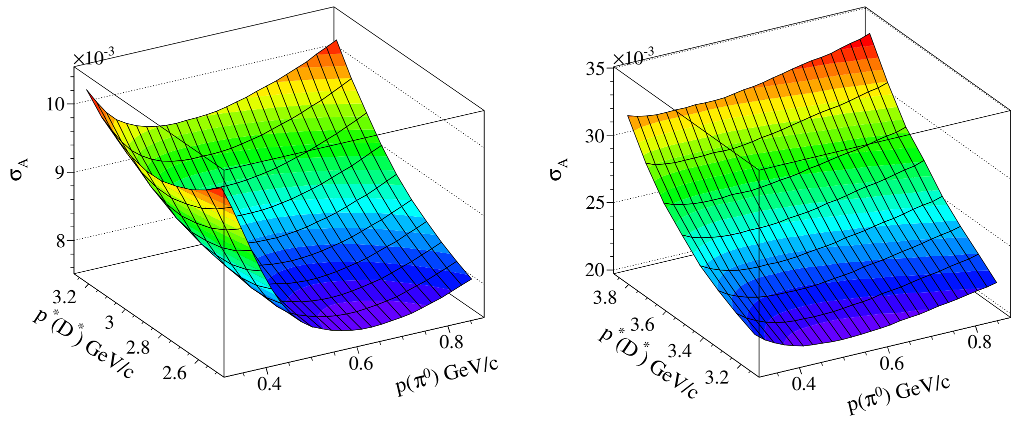


FIGURE 4.9: Estimated error σ_A for the (left) $\Upsilon(4S)$ and (right) $\Upsilon(5S)$ MC samples while varying $p(\pi^0)$ and $p^*(D^*)$.

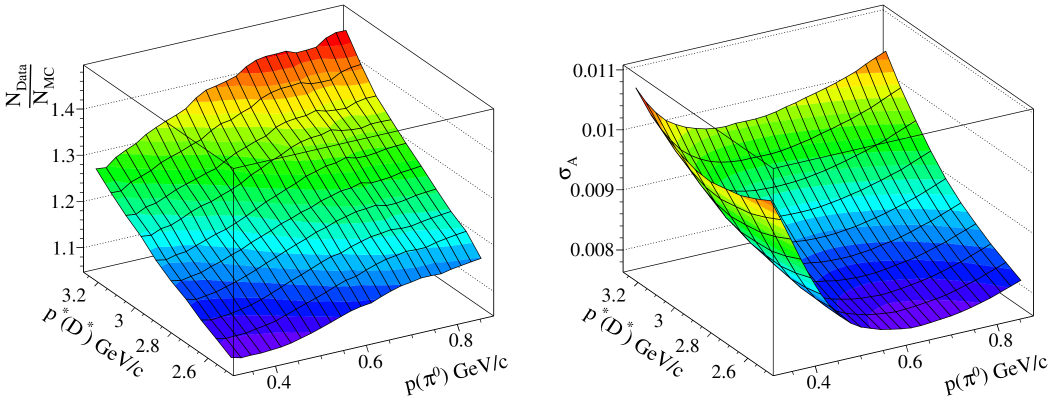


FIGURE 4.10: (left) Correction factor for $\Upsilon(4S)$ and (right) optimization by applying the correction factor.

factor and optimization result after applying this factor for the $\Upsilon(4S)$ and $\Upsilon(5S)$ samples, respectively. The application of correction factor slightly changes the error on raw asymmetry, σ_A, but does not impact the cut for $\Upsilon(4S)$ or $\Upsilon(5S)$.

4.8 Soft pion properties

As we use the ΔM distribution as the final variable to evaluate CP asymmetry, a narrow distribution of ΔM is preferred. Towards this end, the soft pion track

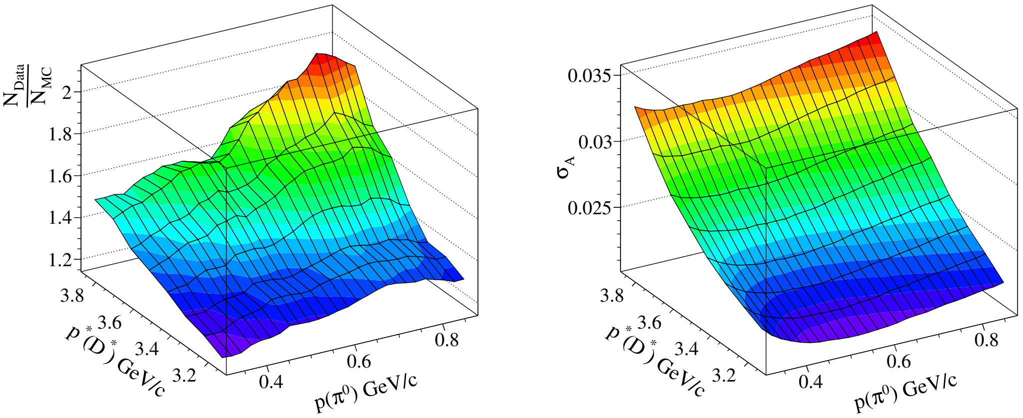


FIGURE 4.11: (left) Correction factor for $\Upsilon(5S)$ and (right) optimization by applying the correction factor.

is refitted to the known interaction point (IP) position obtained using the error matrix associated with the IP profile. The refitting improves the resolution of ΔM distribution as shown in Fig. 4.12.

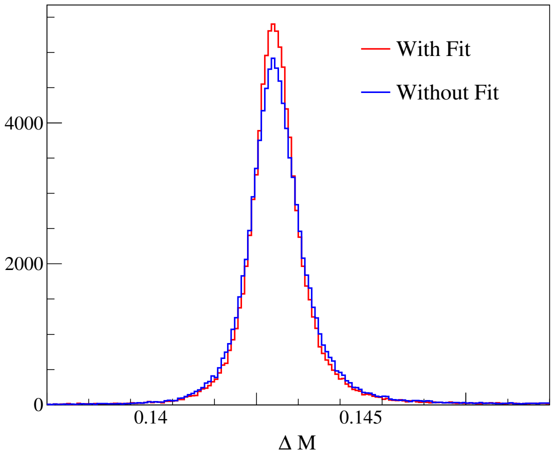


FIGURE 4.12: ΔM resolution before (blue) and after (red) fitting the slow pion to the IP.

The pion selection is dictated by the efficiency map used. In the case of Ref. [25], pion candidates with a KID values greater than 0.6 are rejected and there is no SVD hit requirement (the distribution is shown in Fig. 4.13). For Ref. [27], on the other hand, pion candidates with a KID value greater than 0.9 are rejected, and candidates with π_s having at least one SVD hit are retained. Also the map

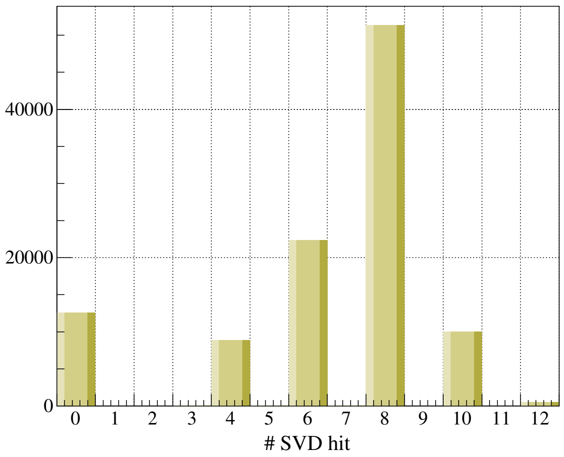


FIGURE 4.13: Distribution for the number of SVD hits associated to π_s.

requires a cut on the confidence level (CL >0.001) obtained by converting the χ^2 associated with vertex fitting of $\pi_s^{\pm}$.

4.9 Final set of cuts

The optimized set of cuts including those prescribed by Ref. [25] for the π_s selection are listed in Table 4.1. These cuts are applied at the reconstruction level and select a signal sample for further studies. We apply different cuts on $p^*(D^*)$ and $p(\pi^0)$ for different MC (or data) samples based on the CM energy of e^+e^- collision. The set of cuts for $\Upsilon(4S)$ is used for all data samples except for the case of $\Upsilon(5S)$.

4.10 Best candidate selection

After applying all selection cuts, the multiplicity of D^* candidates is found to be 6.3%, caused by fake π^0 or π_s. Figure 4.14 shows the candidate multiplicity distribution after applying optimized set of cuts at the reconstruction level. A signal MC study shows 44% of multiplicity is due to π_s while the remaining 56% is due to fake π^0. First, we need a suitable algorithm to pick the true $D^0 \to \pi^0\pi^0$

TABLE 4.1: Summary of optimized cuts including those prescribed by Ref. [25] for the π_s selection.

	Cut	
Variable	$\Upsilon(4S)$	$\Upsilon(5S)$
dr	1.0 cm	1.0 cm
$\|dz\|$	3.0 cm	3.0 cm
KID	< 0.6	< 0.6
$p^*(D^*)$	> 2.5 GeV/c	> 3.1 GeV/c
$M(D^0)$	(1.758,1.930) GeV/c^2	(1.758,1.930) GeV/c^2
$p(\pi^0)$	> 0.64 GeV/c	> 0.54 GeV/c
$M(\pi^0)$	(110,160) MeV/c^2	(110,160) MeV/c^2
E_γ^{barrel}	> 60 MeV	> 60 MeV
E_γ^{endcap}	> 100 MeV	> 100 MeV

among all D candidates. If there are more than one D^0 candidates in an event, we retain the one with the smallest $\sum \chi^2_{\pi^0}$ value, where $\chi^2_{\pi^0}$ is associated with the π^0 mass-constraint fit and sum is over two π^0 candidates. The $\sum \chi^2_{\pi^0}$ distribution for the correctly reconstructed signal as well as for all events are shown in Fig. 4.15. To improve our best candidate selection (BCS) efficiency, we have explored the possibility of using an independent variable, $M(D^0)$. Figure 4.16 shows that there is no correlation between $\sum \chi^2_{\pi^0}$ and $M(D^0)$ and hence the latter can be used as an additional criterion in the form of a minimum deviation of D^0 mass from the PDG value. In the end, out of all candidates in a given event the best D candidate is selected as the one having the lowest value of χ^2_{BCS}:

$$\chi^2_{\text{BCS}} = \sum \chi^2_{\pi^0} + \left(\frac{M(D^0) - m(D^0)_{\text{PDG}}}{\sigma_i}\right)^2 \tag{4.12}$$

where $i = L(R)$ if $M(D^0) - m(D^0)_{\text{PDG}}$ is $-$ve (+ve), where σ_L and σ_R are estimated from the signal MC sample. The second step is to pick up the best D^*

candidate from a set of candidates having the same χ^2_{BCS}. In such cases we select the candidate with the smallest impact parameter dr for π_s. The final BCS efficiency is 74%, estimated as:

$$\varepsilon_{\mathrm{BCS}} = \frac{\text{Number of correct candidates selected in BCS}}{\text{Number of events having candidate multiplicity} > 1} \tag{4.13}$$

The total number of D candidates in data, after applying the optimized cuts as well as BCS, is 34689 ± 325 and later presented in Fig. 4.21. The number is estimated by fitting the ΔM distribution with a signal shape modeled as the sum of a Gaussian and an asymmetric Gaussian function, combined with a background shape modeled as a threshold function (these functions and associated parameters are described in the following section).

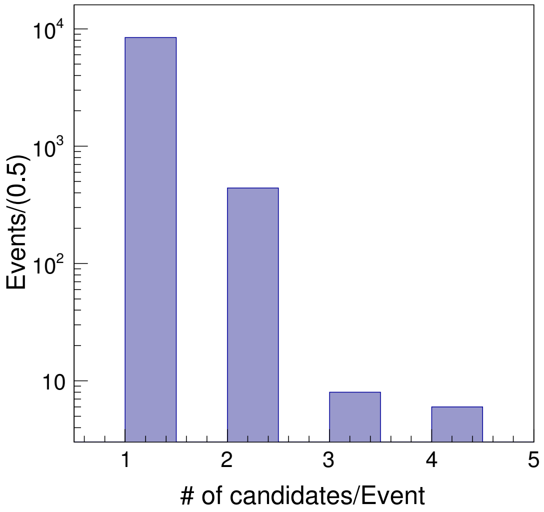


FIGURE 4.14: Candidate multiplicity distribution for $D^0 \to \pi^0\pi^0$ signal MC events.

4.11 A_{CP} estimation

A simultaneous fit of the ΔM distribution for D^{*+} and D^{*-} is used to estimate the asymmetry. The signal shape parameters are common both for the particle

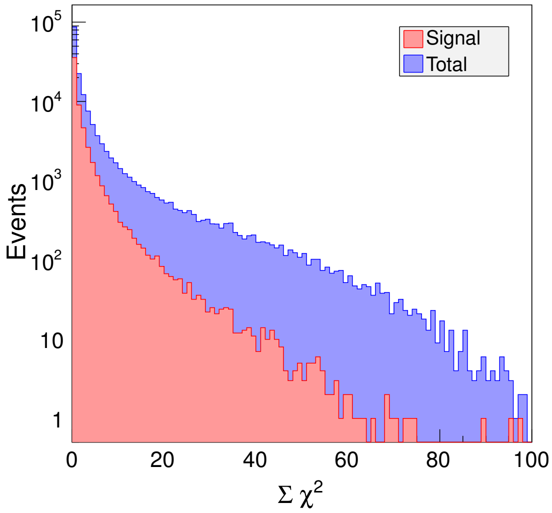


FIGURE 4.15: $\sum \chi^2$ distribution: the red filled histogram indicates the correctly reconstructed signal while the blue filled histogram denotes all events.

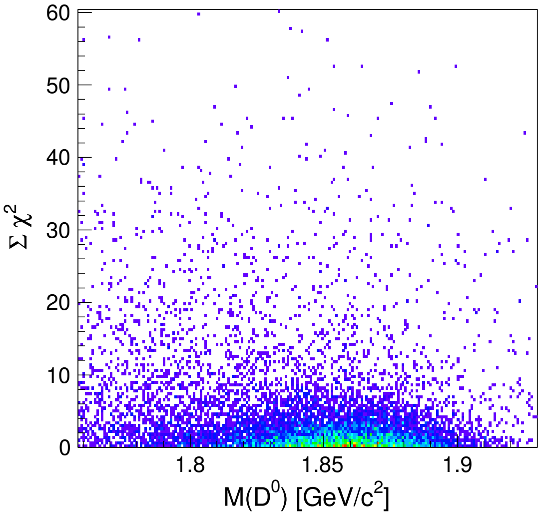


FIGURE 4.16: $M(D^0)$ vs. $\sum \chi^2_{\pi^0}$ distribution for signal MC events: there is no visible correlation between those two variables.

and antiparticle. The advantage of this method is that the full event statistics is utilized when the shape parameters are estimated in the fit. The signal component is modeled as a sum of a Gaussian and an asymmetric Gaussian function, with a common mean. A threshold function $f(x) = (x - m_0)^\alpha \exp[-\beta(x - m_0)]$, where m_0 is the nominal charged pion mass [26], and α and β are two shape parameters, is used to parametrize the background.

The raw asymmetry ($A_{\rm rec}$) obtained by fitting the first stream of generic MC events comprising $\Upsilon(4S)$, $\Upsilon(4S)$ off-resonance and $\Upsilon(5S)$ samples is shown in Fig. 4.17; the asymmetry has an associated error, $\sigma_A = 0.64\%$. The fit parameter $A_{\rm bkg}$ is the background asymmetry in the $D^0 \to \pi^0\pi^0$ decay. The $A_{\rm rec}$ and $A_{\rm bkg}$ values estimated for different MC streams are listed in Table 4.2.

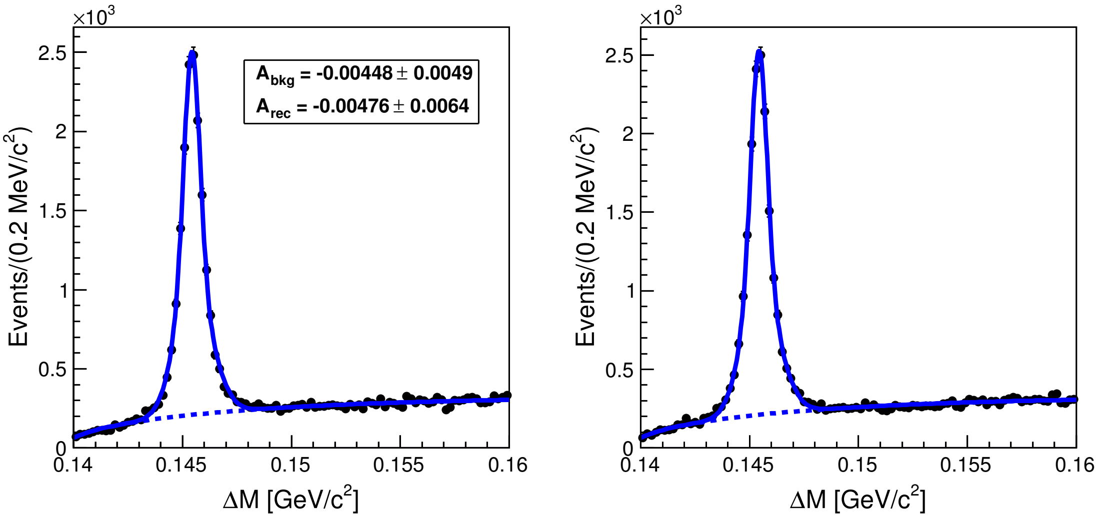


FIGURE 4.17: The raw asymmetry is estimated by simultaneously fitting the ΔM distributions for (left) D^0 and (right) $\overline{D}^0$ samples in the first MC stream.

4.12 $D^0 \to K_S^0\pi^0$ study

The result on CP asymmetry for $D^0 \to K_S^0\pi^0$ [25] is reproduced to check the robustness of the A_{CP} extraction procedure as well as to verify our reconstruction and fitting codes. The set of cuts is close to Ref. [25] by B. R. Ko (BK) that was

TABLE 4.2: A_{rec} and A_{bkg} estimated for the different MC streams available with the simultaneous fit method described in the text.

Stream	A_{rec} (%)	A_{bkg} (%)
0	-0.48 ± 0.64	-0.45 ± 0.49
1	$+0.28 \pm 0.65$	-2.63 ± 0.49
2	$+1.97 \pm 0.64$	-0.45 ± 0.49
3	$+0.10 \pm 0.64$	-0.54 ± 0.49
4	$+0.69 \pm 0.64$	-1.22 ± 0.50

based on using 791 fb^{-1} of data (KID< 0.6, no SVD hit requirement, $E9/E25$ cut applied, however we use $dr < 1.0$ cm and $|dz| < 3.0$ cm cuts already applied at the skim level). The signal yield found is 316845 ± 645 events for the data set of 791 fb^{-1} [see Fig. 4.18(top)], to be compared with the signal yield of 326303 ± 679 events of Ref. [25]. The signal yield for the full 966 fb^{-1} sample, without applying any $E9/E25$ cut, is 458527 ± 769 [see Fig. 4.18(bottom)].

The whole data sample is first divided into 10 bins in $\cos\theta^*$, where θ^* is the polar angle of D^{*+} production in the CM frame. The contents of these bins are further divided into 7×8 bins of $[p_T(\pi_s), \cos\theta(\pi_s)]$. We then calculate A_{rec} in each of these bins using a simultaneous fit. The raw asymmetry is subsequently corrected for the detection efficiency asymmetry by subtracting A^{π}_{ϵ} from the former [refer to Eq. (4.8)]. Now, the resulting $A^{\text{cor}}_{\text{rec}}$ is a function of $\cos\theta^*$, $\cos\theta(\pi_s)$ and $p_T(\pi_s)$. The corrected raw asymmetries are added up in the two-dimensional bins of $[p_T(\pi_s), \cos\theta(\pi_s)]$ to get 10 values of $A^{\text{cor}}_{\text{rec}}$ (the central values are added using $\sum_{i,j}(A_{ij}/\sigma^2_{ij})/\sum_{i,j}(1/\sigma^2_{ij})$ and uncertainties are added as $1/\sigma^2 = \sum_{ij}(1/\sigma^2_{ij})$ [28], where i and j denote the bin index in $p_T(\pi_s)$ and $\cos\theta(\pi_s)$, respectively) corresponding to 10 bins of $\cos\theta^*$. The A_{CP} and A_{FB} values are obtained by adding or subtracting $A^{\text{cor}}_{\text{rec}}$ in the bins having same $|\cos\theta^*|$ value as given in Eq. (4.10). The A_{CP} and A_{FB} obtained for the 791 fb^{-1} data are shown in Fig. 4.19. We get $A_{CP} = (-0.21 \pm 0.19)\%$, which is consistent with the earlier published result [25],

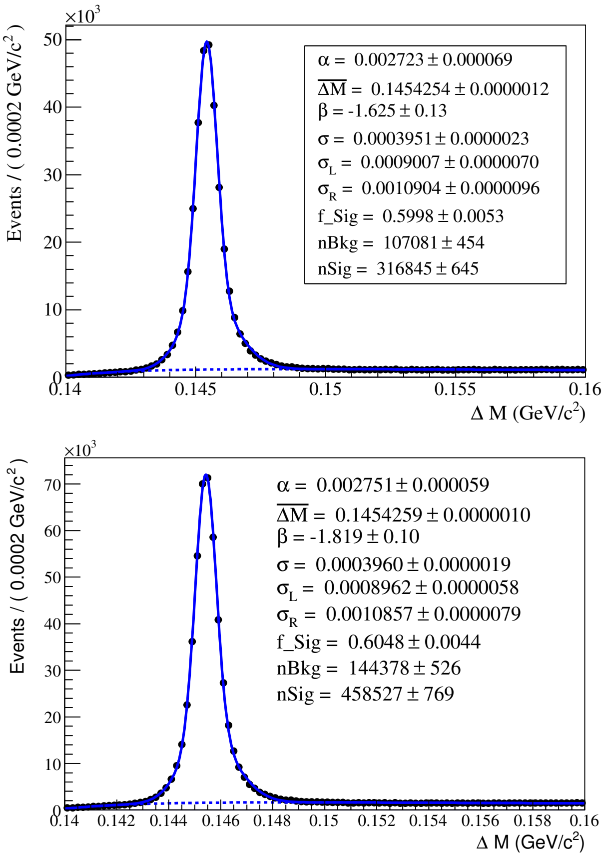


FIGURE 4.18: Signal yields for $D^0 \to K^0_S\pi^0$ sample: (top) 791 fb^{-1} data sample as used in Ref. [25], and (bottom) 966 fb^{-1} data sample.

$A_{CP} = (-0.28 \pm 0.19)\%$. With the whole data sample (966 fb^{-1}), the result is $A_{CP} = (-0.10 \pm 0.16)\%$ (Fig. 4.20).

As a cross check, we estimate A_{CP} in the full data sample using the selection and efficiency map of M. Starič (MS) for the slow pion. The binning of $[p(\pi_s), \cos\theta(\pi_s)]$ is 5×5. To verify that the efficiency map is properly applied, we compare the A_{CP} estimated without charged pion correction $[{A_{CP}}^{\rm uncorr} = (-0.21 \pm 0.16)\%]$ with the corrected CP asymmetry $[A_{CP} = (+0.00 \pm 0.17)\%]$; they are found to differ by $+0.21\%$. This can be compared to a similar exercise carried out by MK for $D \to K^+K^-$, where a difference of $+0.17\%$ is found, showing good agreement.

The difference in the A_{CP} central values between the above two selection (or map) analyses [$(-0.10 \pm 0.16)\%$] and [$(+0.02 \pm 0.15)\%$] is consistent with the systematic uncertainty assigned due to the efficiency map. The A_{CP} values estimated with different selection cuts, data sample and efficiency map are summarized in Table 4.3.

TABLE 4.3: Summary of A_{CP} values for different selection cuts, data sample (791 or 966 fb^{-1}), efficiency map used to correct for the detection asymmetry of the soft pion (from BK or MS) or no map used ("No").

Map	$\mathcal{L}(\mathrm{fb}^{-1})$	$(dr, \lvert dz\rvert)$ (cm)	$E9/E25$	# SVD	KID	CL	A_{CP} (%)
No	791	$(1.0 <, 3.0 <)$	> 0.85	-	< 0.6	-	-0.08 ± 0.19
BK	791	$(1.0 <, 3.0 <)$	> 0.85	-	< 0.6	-	-0.21 ± 0.19
No	966	$(1.0 <, 3.0 <)$	-	-	< 0.6	-	$+0.02 \pm 0.15$
BK	966	$(1.0 <, 3.0 <)$	-	-	< 0.6	-	-0.10 ± 0.16
No	966	$(1.0 <, 3.0 <)$	-	> 0	< 0.9	> 0.001	-0.21 ± 0.16
MS	966	$(1.0 <, 3.0 <)$	-	> 0	< 0.9	> 0.001	$+0.00 \pm 0.17$

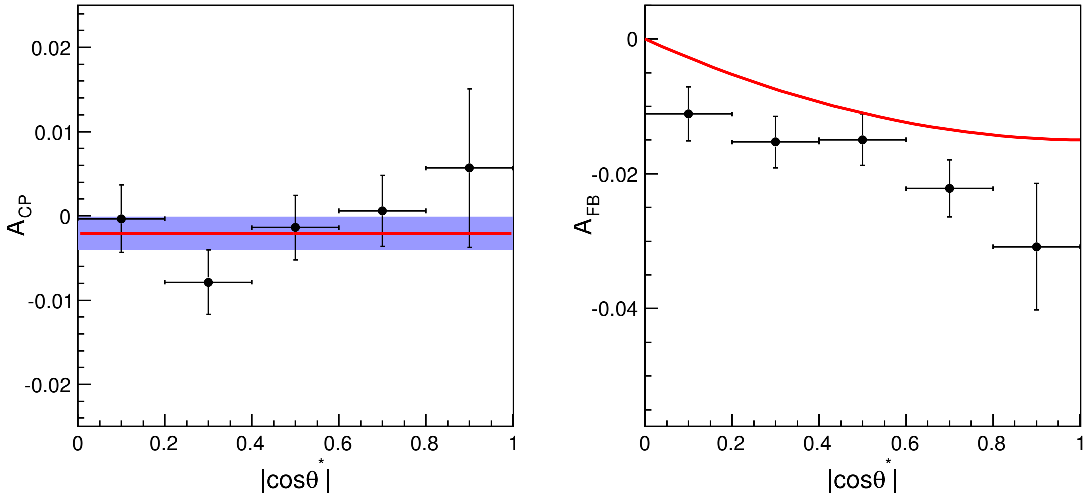


FIGURE 4.19: A_{CP} and A_{FB} estimated for the $D^0 \to K_S^0\pi^0$ mode in a condition similar to Ref. [25] using 791 fb^{-1} data sample.

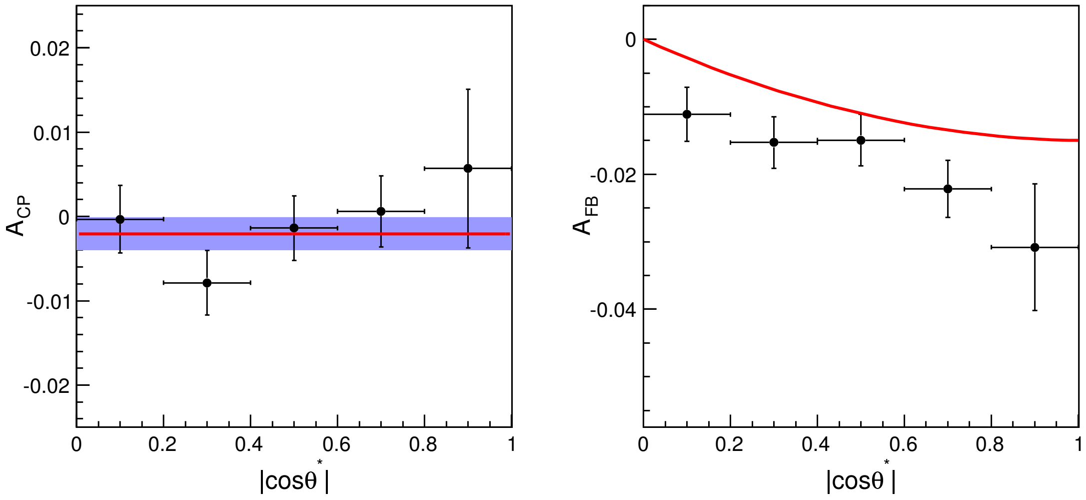


FIGURE 4.20: A_{CP} and A_{FB} are estimated for the $D^0 \to K_S^0\pi^0$ mode with 966 fb^{-1} data sample.

4.13 $D^0 \to \pi^0\pi^0$ study

Before estimating A_{CP} for $D^0 \to \pi^0\pi^0$, we need to decide which efficiency map to be used (and consequently which π_s selection to be applied). For this we estimate $A_{\rm rec}$ for few streams of generic MC events. The results are summarized in Table 4.4. The statistical errors obtained for BK's selection are systematically smaller. We understand the reason, that is MS's selection was optimized for time dependent CPV studies (SVD hits requirement, $CL > 0.001$) while BK's selection was fine tuned for time integrated CPV studies. After comparing the errors associated with $A_{\rm rec}$'s for both selections, we decide to adopt the map estimated by BK for Ref. [25]. The total signal yield for $D^0 \to \pi^0\pi^0$ is 34689±325 (Fig. 4.21)

The A_{CP} value for $D^0 \to \pi^0\pi^0$ is estimated in a similar way as for $D^0 \to K_S^0\pi^0$ (Section 4.12). Compared to the $K_S^0\pi^0$ case, however, the statistics available is ten times smaller, so the ΔM signal shape parameters cannot be let free in all bins during the estimation of asymmetry via the simultaneous fitting. This difficulty can be overcome by fixing the shape parameters of $D^0 \to \pi^0\pi^0$ in each bin to the corresponding value in $D^0 \to K_S^0\pi^0$ after correcting for the difference between the

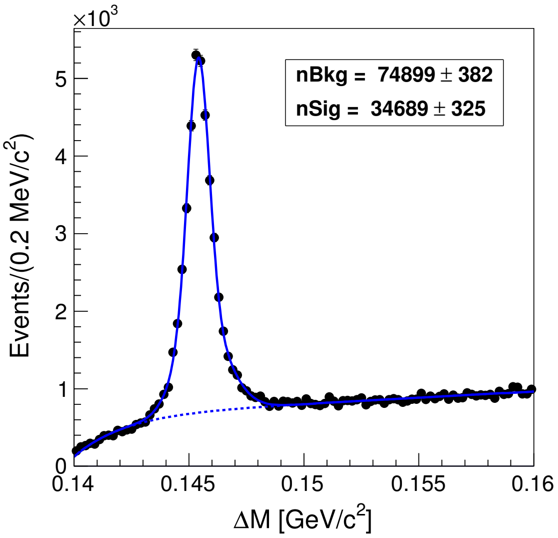


FIGURE 4.21: Signal yield for $D^0 \to \pi^0\pi^0$ with 966 fb^{-1} data sample.

TABLE 4.4: Summary of A_{rec} $(D^0 \to \pi^0\pi^0)$ values for different selection cuts used by MS or BK, for different MC streams.

Stream No:	$A_{\mathrm{rec}}^{\mathrm{MS}}$ (%)	$A_{\mathrm{rec}}^{\mathrm{BK}}$ (%)
0	-0.48 ± 0.64	-0.10 ± 0.62
1	-0.28 ± 0.65	$+0.45 \pm 0.62$
2	$+1.97 \pm 0.64$	$+2.21 \pm 0.62$
3	-0.10 ± 0.64	$+0.48 \pm 0.61$
4	$+0.69 \pm 0.64$	$+0.76 \pm 0.62$

two cases seen in signal MC events. This difference (shown for few typical bins in Fig. 4.22) can be determined as a correction factor as:

$$\frac{\sigma_i^{\mathrm{MC}}(\pi^0\pi^0)}{\sigma_i^{\mathrm{MC}}(K_S^0\pi^0)} \tag{4.14}$$

where σ_i is the resolution parameter for the i-th bin. In fact, there are three such factors corresponding to the three widths: σ, σ_L and σ_R. We estimate the correction factors separately in each bin and find them to be pretty much the same for all bins; they are 1.24, 1.11 and 1.14 respectively.

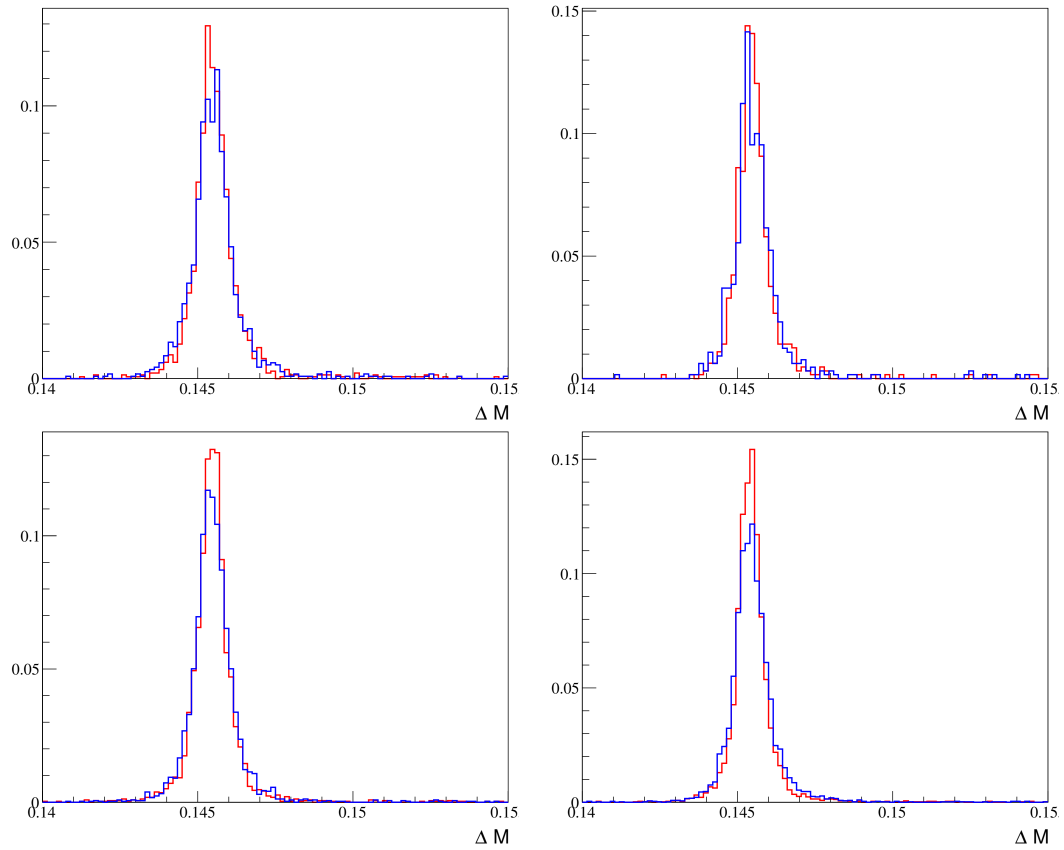


FIGURE 4.22: Comparison of normalized distribution of ΔM, in same bin, for $D^0 \to \pi^0\pi^0$ (blue) and $D^0 \to K_S^0\pi^0$ (red) in signal MC: (top left)$-0.8 \leq \cos\theta^* < 0.6$, $-0.2 \leq \cos\theta_{\pi_s} < 2$, $0.1 \leq p(\pi_s) < 0.2$ GeV/c (top right)$-0.8 \leq \cos\theta^* < 0.6$, $-0.2 \leq \cos\theta_{\pi_s} < 2$, $0.2 \leq p(\pi_s) < 0.3$ GeV/c (bottom left) $-0.6 \leq \cos\theta^* < 0.4$, $-0.2 \leq \cos\theta_{\pi_s} < 2$, $0.1 \leq p(\pi_s) < 0.2$ GeV/c (bottom right) $-0.6 \leq \cos\theta^* < 0.4$, $-0.2 \leq \cos\theta_{\pi_s} < 2$, $0.2 \leq p(\pi_s) < 0.3$ GeV/c

The signal ΔM shape depends mostly on the soft pion and is independent of $\cos\theta^*$. Therefore, the one-to-one correspondence between the $K_S^0\pi^0$ and $\pi^0\pi^0$ modes is taken in bins of $[p_T(\pi_s), \cos\theta(\pi_s)]$. The background shape parameters are kept free, as the nature of background is likely to be different in different decay modes. Figure 4.23 shows the A_{CP} estimated for $D^0 \to \pi^0\pi^0$. The value obtained $(-0.03 \pm 0.64)\%$ is consistent with null asymmetry. We also estimate $A_{CP}^{\rm uncorr} = (+0.08 \pm 0.64)\%$ and find the difference between $A_{CP}^{\rm uncorr}$ and A_{CP} consistent with the value in the $K_S^0\pi^0$ mode (Table 4.3). This confirms that the efficiency map has been properly applied.

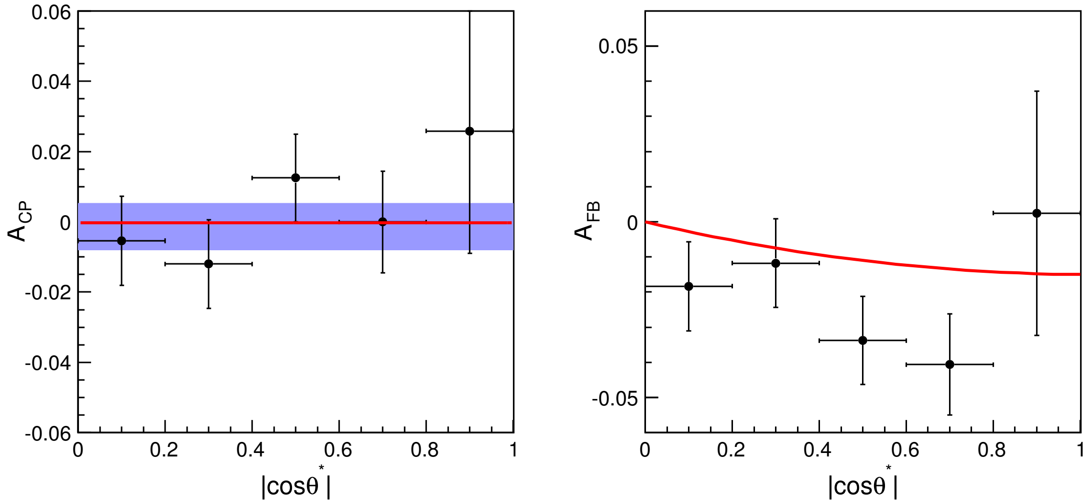


FIGURE 4.23: A_{CP} and A_{FB} values estimated for $D^0 \to \pi^0\pi^0$ in the final data sample.

4.14 Systematics

The main systematic uncertainty arises from the efficiency map, which has been estimated in previous analyses [25, 27] to be $\sim 0.1\%$. Various sources of systematics are described in the following sections and include the $\cos\theta^*$ binning (Section 4.14.1), the fixed shape parameters used in the fit (Section 4.14.2), and the efficiency map used to correct for $A_\epsilon^{\pi_s}$ (Section 4.14.3).

4.14.1 Binning

For estimating the nominal A_{CP}value, we have divided the whole data sample into 10 bins of $\cos\theta^*$, while the content of each $\cos\theta^*$ bin is further divided into 7×8 bins of $[p_T(\pi_s), \cos\theta(\pi_s)]$. The systematic error associated with these binnings can be determined by using $A_{CP}^{\rm uncorr}$ estimated with different number of $\cos\theta^*$ bins. For this purpose, we use 8, 10 (nominal), and 12 bins of $\cos\theta^*$. The variation of

A_{CP}^{uncorr} from its nominal value gives the error associated with binning. The results are listed in Table 4.5. The systematic uncertainty is 0.00% on the positive side and 0.07% on the negative side.

TABLE 4.5: Results obtained for different $\cos\theta^*$ binning.

N0. of $\cos\theta^*$ bins	A_{CP}^{uncorr} (%)	$A_{CP}^{\text{uncorr}} - A_{CP}^{\text{uncorr}}$(nominal) (%)
8	0.073 ± 0.635	-0.009
10	0.081 ± 0.636	
12	-0.015 ± 0.635	-0.066

4.14.2 Fitting

We fix five signal shape parameters in each bin to the corresponding values obtained for $D^0 \to K_S^0\pi^0$. These are the mean $(\overline{\Delta M})$, σ, σ_L, σ_R, and the relative fraction of the Gaussian to the asymmetric Gaussian component. The *CP* asymmetry obtained with these fixed parameters are referred to as A^{norm}, the nominal value. In order to calculate the systematic error due to fixed parameters, we vary each of them by the respective error while keeping other parameters fixed to their nominal values. By this variation we get two asymmetries $A_i^{\text{norm}-\text{err}}$ and $A_i^{\text{norm}+\text{err}}$, where i is the parameter index out of which we determine the variations of *CP* asymmetry as $A_i^{\text{norm}-\text{err}} - A^{\text{norm}}$ and $A_i^{\text{norm}+\text{err}} - A^{\text{norm}}$. These deviations can be of the same or opposite sign and are thus grouped in to two sets, one with positive deviation and the other with negative deviation. If both deviations are of the same sign, we take the largest value as the deviation in one side and put zero on the other. Results are summarized in Table 4.6. Finally each of these deviations are added in quadrature to get the total systematic error associated with fitting. This way we have obtained a systematics error of 0.03% on both positive and negative side.

Table 4.6: Systematic error associated with the fixed parameters.

Parameter	$A_{CP}^{\text{norm}-\text{err}} - A_{CP}^{\text{norm}}(\%)$	$A_{CP}^{\text{norm}+\text{err}} - A_{CP}^{\text{norm}}(\%)$	$+$	$-$
$\overline{\Delta m}$	-0.001	-0.002	—	0.002
σ	-0.011	0.003	0.003	0.011
σ_L	0.024	-0.012	0.024	0.012
σ_R	-0.015	0.017	0.017	0.015
Fraction	0.000	0.012	0.000	0.012
Total			0.030	0.025

4.14.3 Efficiency map

The systematic uncertainty associated with efficiency map is estimated in a similar way as described in Ref. [25]. The nominal value for A_{CP} is $-0.03\pm0.64\%$, which is obtained by applying a correction factor due to the detection efficiency asymmetry ($A_{\epsilon}^{\pi_s}$) in each of the $7 \times 8 = 56$ bins in $[p_T(\pi_s), cos\theta(\pi_s)]$. Now to calculate the systematics due to efficiency map, we vary $A_{\epsilon}^{\pi_s}$ by $\pm1\sigma$ (σ is the statistical error on $A_{\epsilon}^{\pi_s}$) in each of the above 56 bins and obtain the deviation in A_{CP} with respect to its nominal value. Adding all these deviations in quadrature gives us the final systematic error. We find the value to be ±0.07.

4.14.4 Total systematics

The final systematic uncertainty is obtained by adding all individual contributions described above in quadrature as summarized in Table 4.7. A total of $+0.07 - 0.10\%$ systematic error is obtained.

TABLE 4.7: Total systematic uncertainty for $D^0 \to \pi^0\pi^0$.

Systematics	+	−
Binning	0.000	0.066
Fitting	0.030	0.025
Eff. map	0.067	0.067
Total	0.073	0.097

4.15 Systematics in $D^0 \to K_S^0\pi^0$ study

We perform a similar systematics study as for the $D^0 \to \pi^0\pi^0$ mode. The systematics associated with binning is listed in Table 4.8. The corresponding error is 0.00% (0.02%) on the positive (negative) side. Unlike the $\pi^0\pi^0$ mode we vary all the parameters except for the "fraction" that is fixed to the value obtained from the signal MC sample. The systematics error due to the fixed fraction is estimated and mentioned in Table 4.9. By this way we obtain the systematic error associated with signal shape, 0.00% (0.01%) on the positive (negative) side. Systematics associated with efficiency map is estimated to be ±0.07 %.

TABLE 4.8: Systematic uncertainty associated with binning for $D^0 \to K_S\pi^0$.

No. of $\cos\theta^*$ bins	A_{CP}^{uncorr} (%)	$A_{CP}^{\text{uncorr}} - A_{CP}^{\text{uncorr}}$(nominal) (%)
8	0.012 ± 0.155	-0.006
10	0.018 ± 0.155	
12	0.002 ± 0.155	-0.017

TABLE 4.9: Systematic uncertainty due to the fixed parameters for $K_S^0\pi^0$ mode.

Parameter	$A_{CP}^{\text{norm}-\text{err}} - A_{CP}^{\text{norm}}(\%)$	$A_{CP}^{\text{norm}+\text{err}} - A_{CP}^{\text{norm}}(\%)$	+	−
Fraction	−0.001	−0.010	—	−0.010

TABLE 4.10: Total systematic error for $D^0 \to K_S\pi^0$.

Systematics	+	−
Binning	0.000	0.017
Fitting	0.000	0.010
Eff. map	0.066	0.066
Total	0.066	0.068

4.16 Conclusions

We have measured CP violation asymmetry in the decay $D^0 \to \pi^0\pi^0$ as $A_{CP} = (-0.03 \pm 0.64 \pm 0.10)\%$ using a data sample of $966\,\text{fb}^{-1}$. No significant asymmetry is observed and the measurement constitutes a significant improvement over the previous one by the CLEO Collaboration.

CHAPTER 5

Search for the rare decay $D^0 \to \gamma\gamma$

5.1 Introduction

In this chapter, we present a search for the $D^0 \to \gamma\gamma$ decay using Belle data. The results of previous searches by CLEO, *BABAR* and BESIII Collaborations are summarized in Table 5.1. None of these studies found any signal and *BABAR*'s limit continues to be the world's best upper limit. We employ herein nearly twice the data sample *BABAR* used and a sophisticated two-dimensional (2D) fitting procedure. These factors lead to a substantial improvement in sensitivity over the previous result [30].

TABLE 5.1: Summary of previous experimental results on $D^0 \to \gamma\gamma$.

Group	Data (fb^{-1})	Upper limit on $\mathcal{B}$ at 90% CL
CLEO	13.8	$< 2.9 \times 10^{-5}$ [29]
BABAR	470.5	$< 2.2 \times 10^{-6}$ [30]
BESIII	2.92	$< 3.8 \times 10^{-6}$ [31]

5.2 Data and MC samples

We use an 832fb^{-1} of data sample recorded at or near $\Upsilon(4S)$ and $\Upsilon(5S)$ resonances. In total, five streams of generic MC sample are used to study various background contributions, to optimize selection criteria, and to determine probability distribution functions (PDFs) for signal and backgrounds. Further we use two sets of signal MC samples comprising about 200k and 34k events generated at $\Upsilon(4S)$ and $\Upsilon(5S)$ resonances, for $D^0 \to \gamma\gamma$ and $D^0 \to K_S^0\pi^0$ decays. The relative size of these two samples is determined according to the integrated luminosity of the corresponding dataset (711 and 121 fb^{-1}, respectively). The two samples are merged to form a 234k combined signal MC sample.

Throughout this chapter, the charge-conjugate decay mode is implied unless stated otherwise. The signal sample is $D^{*+} \to D^0(\to \gamma\gamma)\pi_s^+$ while the $D^{*+} \to D^0(\to K_S^0\pi^0)\pi_s^+$ sample is used as the normalization mode. In addition, we use the control samples of $D^{*+} \to D^0(\to \phi\gamma)\pi_s^+$, $D^{*+} \to D^0(\to \pi^0\pi^0)\pi_s^+$ and $D^{*+} \to D^0(\to K_S^0\gamma)\pi_s^+$ to estimate data-MC correction factors and systematic uncertainties.

5.3 Reconstruction and analysis method

We reconstruct D^0 ($\overline{D}^0$) candidate from a pair of high energy photons having $E_\gamma > 200\,\text{MeV}$. The D candidate is then combined with a low momentum (“slow") pion π_s^+ (π_s^-) to form a D^{*+} (D^{*-}). Requiring the D^0 to come from a D^{*+} largely suppresses the background due to random combination of photons (“combinatorial"). We use the distributions of the reconstructed invariant mass of D^0 candidates $M(D^0)$ and ΔM to search for the decay. Here, the D^{*+} mesons mostly originate from $c\bar{c}$ continuum ($e^+e^- \to c\bar{c}$) via hadronization, which has an associated theory uncertainty $\sim 12.5\%$ [26]. This uncertainty restricts us from determining the number of D^*s with a desired accuracy. On the other hand, by measuring the $D^0 \to \gamma\gamma$ branching fraction with respect to a well measured mode such as $D^0 \to K^0_S\pi^0$ will alleviate this problem. The relative measurement also helps to cancel systematic uncertainties, common to both modes. The branching fraction for $D^0 \to \gamma\gamma$ is thus determined using the following relation:

$$\mathcal{B}(D^0 \to \gamma\gamma) = \frac{N(D^0 \to \gamma\gamma)}{N(D^0 \to K^0_S\pi^0)} \times \frac{\varepsilon_{K^0_S\pi^0}}{\varepsilon_{\gamma\gamma}} \times \mathcal{B}(D^0 \to K^0_S\pi^0)_{\mathrm{WA}}, \qquad (5.1)$$

where $N(D^0 \to \gamma\gamma)$ $[N(D^0 \to K^0_S\pi^0)]$ is signal yield, $\varepsilon_{\gamma\gamma}$ $[\varepsilon_{K^0_S\pi^0}]$ is the reconstruction efficiency for $D^0 \to \gamma\gamma$ $[D^0 \to K^0_S\pi^0]$, and $\mathcal{B}(D^0 \to K^0_S\pi^0)_{\mathrm{WA}}$ is world-average branching fraction [26] for $D^0 \to K^0_S\pi^0$.

5.4 Background classification

The selection cuts applied at the skim level are listed in Table 5.2. After applying these criteria, we identify various background contributions. The observed backgrounds can be broadly classified into three categories: physics, combinatorial and QED. The first one arises from specific physics processes, that are misidentified as signal. The background originating from the random combination of photons are called the combinatorial background. The QED backgrounds are those from various quantum electrodynamic processes such as $e^+e^- \to \gamma\gamma$ and $e^+e^- \to e^+e^-(\gamma)$.

The decay modes contributing as the physics background include $D^0 \to \pi^0\pi^0$, $D^0 \to \eta\pi^0$, $D^0 \to \eta\eta$, and $D^0 \to K^0_S(\pi^0\pi^0)\pi^0$. The π^0s (or ηs) in these final states mostly decay to a pair of photons. The above decays can be misidentified as $D^0 \to \gamma\gamma$ in two different ways. A pair of energetic photons of the π^0 (or η) origin can combine to mimic the signal. This can be suppressed by applying a dedicated π^0 (or η) veto [32] for each γ candidates. For this we calculate likelihoods by pairing each of the two signal photons with other photon candidates in an event. The second possibility is due to the merged clusters in the ECL. The merging effect

TABLE 5.2: Selection cuts applied at the skimming level.

Variable	Cut
dr	$< 1\,\mathrm{cm}$
$\lvert dz \rvert$	$< 3\,\mathrm{cm}$
$p^*(D^*)$	$> 2\,\mathrm{GeV}/c$
ΔM	$(135, 165)\,\mathrm{MeV}/c^2$
E_γ	$> 200\,\mathrm{MeV}$

in the ECL crystals are described in Section 4.5.2. Usually, such merged clusters are wider in the lateral dimension and thus can be suppressed by applying a cut on $E9/E25$ (> 0.85). Another process contributing to background is $D^0 \to K_L^0 \pi^0$, which arises due to the misidentification of the K_L^0 candidate as a γ.

The combinatorial background is mostly rejected at the reconstruction level by requiring D^0s to come from D^{*+} as described in Section 5.3. Further reduction is achieved due to the cut on $p^*(D^*)$. In order to suppress the QED background, we select only those events that have both charged track and neutral candidates (γ) more than four.

5.5 IP constraint of π_s

As described in Section 4.8, we fit the π_s track to the known interaction point (IP). Figure 5.2 shows the ΔM distribution before and after applying the fit; a pretty good improvement is observed in the resolution. Various selection cuts other than the skim level ones applied before optimization are listed in Table 5.3. The classification of backgrounds after applying these cuts for generic and signal MC events are shown in Fig. 5.1. These figures show the classification of various backgrounds on the basis of event type; they are further divided depending on the physics process. It is clear that the dominant contributor is the combinatorial background.

TABLE 5.3: Selection cuts applied before optimization to suppress backgrounds.

Variable	Cut
# of charged tracks	> 4
# of neutral candidates	> 4
$E9/E25$	> 0.85
Vertex fit for π_s	applied

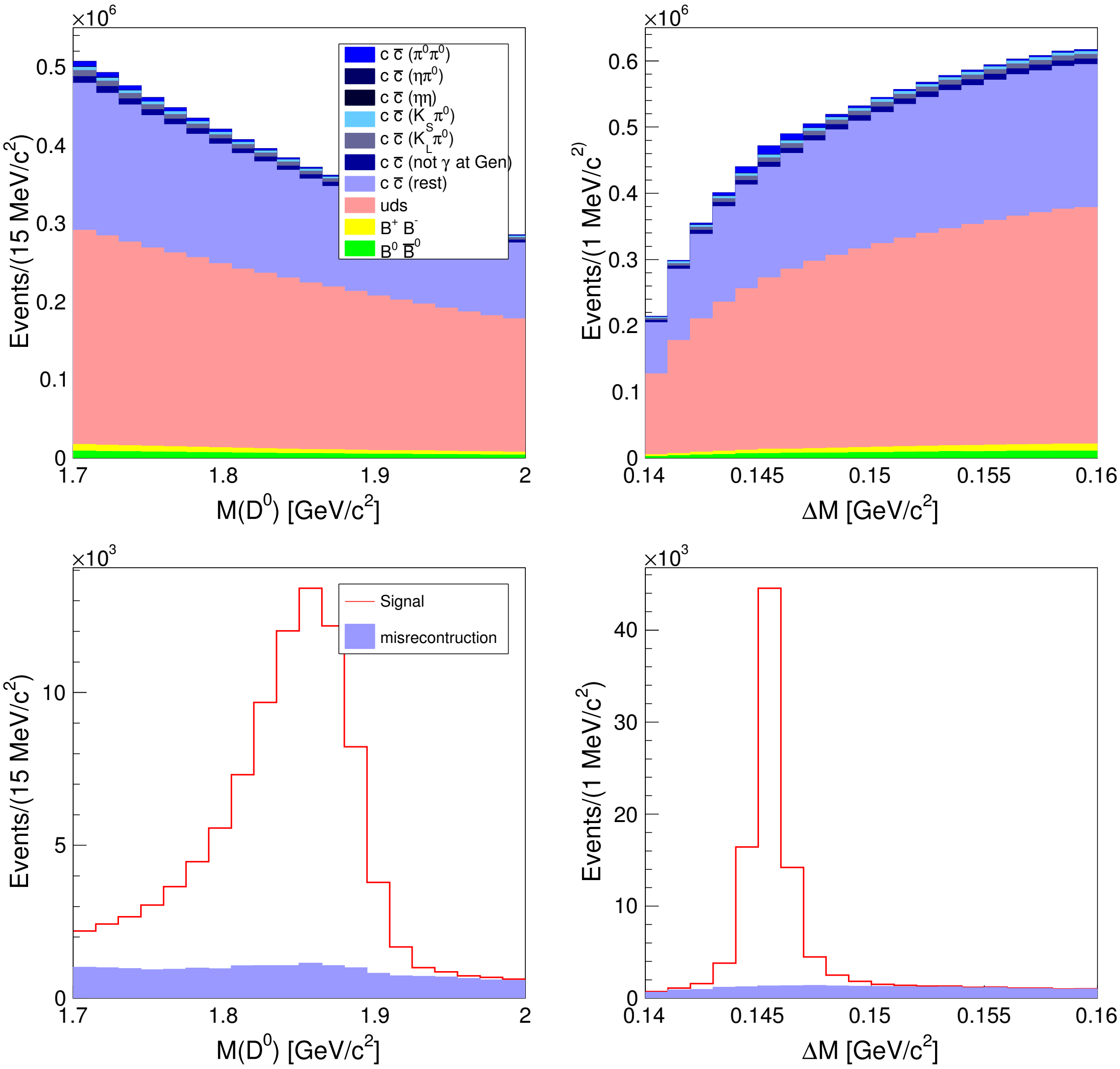


FIGURE 5.1: Classification of various backgrounds in the $m(D^0)$ and ΔM distributions after applying the cuts listed in Tables 5.2 and 5.3: (top left) and (top right) are for generic while (bottom left) and (bottom right) for signal MC events. In both the case, the $\Upsilon(4S)$ and $\Upsilon(5S)$ samples are merged.

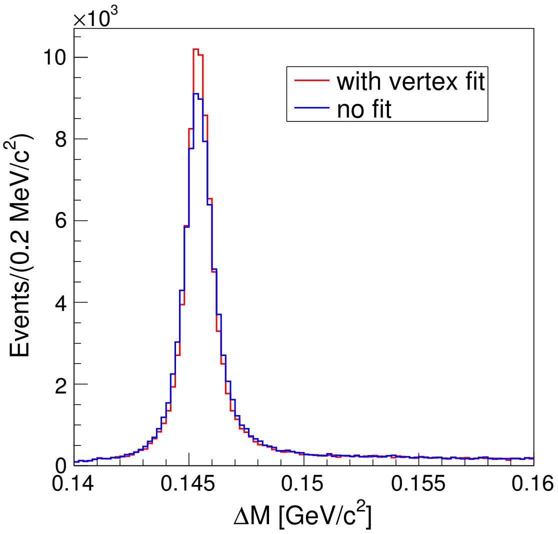


FIGURE 5.2: ΔM distribution before and after fitting π_s to the IP.

5.6 Data-MC agreement

We check the overall agreement between data and events MC for the $M(D^0)$ and ΔM distributions. We define a signal window: a $\pm 3\sigma$ region around the mean of these two distributions and a sideband: a region excluding the signal window. These regions are illustrated in Fig. 5.3) and given by:

$$\text{Signal region: } 1.7107 < M(D^0) < 1.9309\,\text{GeV}/c^2 \text{ and}$$
$$0.1434 < \Delta M < 0.1477\,\text{GeV}/c^2$$
$$\text{Sideband: } 1.95 < M(D^0) < 2.00 \text{ GeV}/\text{c}^2 \text{ or } 0.15 < \Delta M < 0.16\,\text{GeV}/c^2$$

The sideband $M(D^0)$ and ΔM distributions are shown in Fig. 5.4. Clearly, these distributions follow a similar trend in both data and simulations, although the number of observed events in data clearly exceeds that expected from the latter. Therefore we need to correct for this mismatch while optimizing our selection cuts. The correction factor is defined as $(N_{\text{Data}}/N_{\text{MC}})$, where N_{Data} (N_{MC}) is the number of data (MC) events in the sideband.

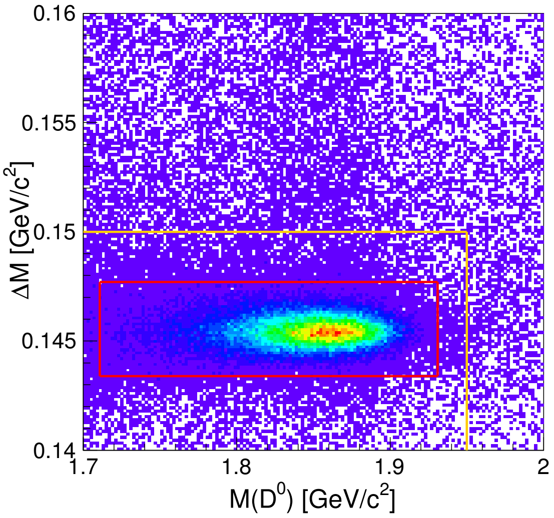


FIGURE 5.3: Illustration of the signal window and sideband. Region inside the red box is the signal window while the region above and right-side of the yellow line is the sideband.

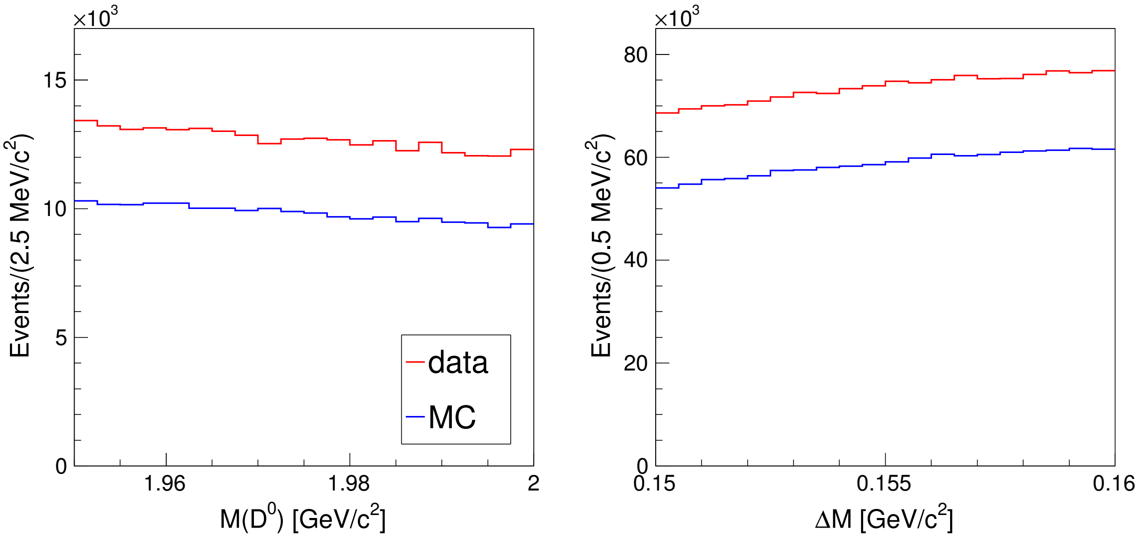


FIGURE 5.4: The comparison of data and MC in the sideband of $M(D^0)$ and ΔM.

5.7 Optimization of selection cuts

The selection cuts are finalized by an optimization procedure. As we are searching for a rare decay, we use a special procedure known as Punzi's method [33], based on both signal and generic MC events. The optimization is performed by maximizing a figure of merit (FOM) while varying the selection cut. The value for which the FOM is maximum is taken as the optimal cut position. The FOM for Punzi's method is [33]:

$$\mathrm{FOM} = \frac{\varepsilon(t)}{a/2 + \sqrt{B(t)}} \tag{5.2}$$

where ε and B are the signal efficiency and the number of expected background events for a given cut value t, and a is the desired confidence level in terms of standard deviations (σ) .We use signal MC events for estimating $\varepsilon(t)$ and the generic MC sample for $B(t)$. The correction for data-MC difference described in Section 5.6 is applied by multiplying $B(t)$ with the corresponding factor $(N_{\mathrm{Data}}/N_{\mathrm{MC}})$. The selection cuts that are pretty standard within Belle are applied without any optimization. The variables we optimize on are $p^*(D^*)$, $|E_{\gamma 1} - E_{\gamma 2}|/(E_{\gamma 1} + E_{\gamma 2})$, which is an asymmetry between the energy of two photons in $D^0 \to \gamma\gamma$, and Prob(π^0) [Prob(η)], the probability for the photon to come from a π^0 [η] decay. A comparison between signal and background events for all optimizing variables is provided in Fig. 5.5.

5.7.1 Optimization of $p^*(D^*)$ and $|E_{\gamma 1} - E_{\gamma 2}|/(E_{\gamma 1} + E_{\gamma 2})$

A two-dimensional optimization is performed for $p^*(D^*)$ and $|E_{\gamma 1} - E_{\gamma 2}|/(E_{\gamma 1} + E_{\gamma 2})$ in order to take care possible correlation between these two variables. For this, we simultaneously vary cuts on $p^*(D^*)$ and $|E_{\gamma 1} - E_{\gamma 2}|/(E_{\gamma 1} + E_{\gamma 2})$ distributions

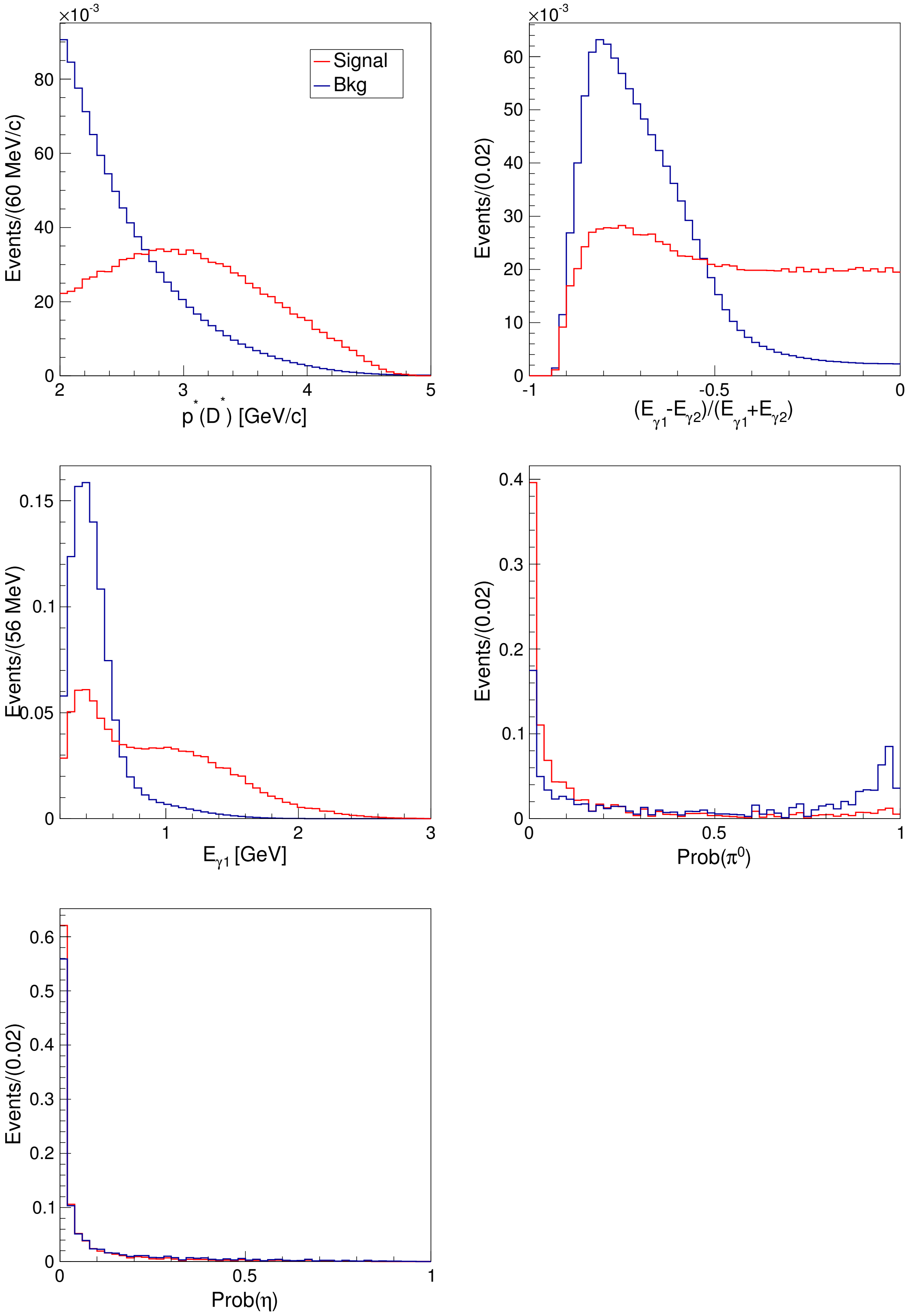


FIGURE 5.5: Signal and background distributions for the optimizing variables: (top left) $p^*(D^*)$, (top right) $|E_{\gamma 1} - E_{\gamma 2}|/(E_{\gamma 1} + E_{\gamma 2})$, (middle left) E_γ, (middle right) Prob(π^0), and (bottom) Prob(η).

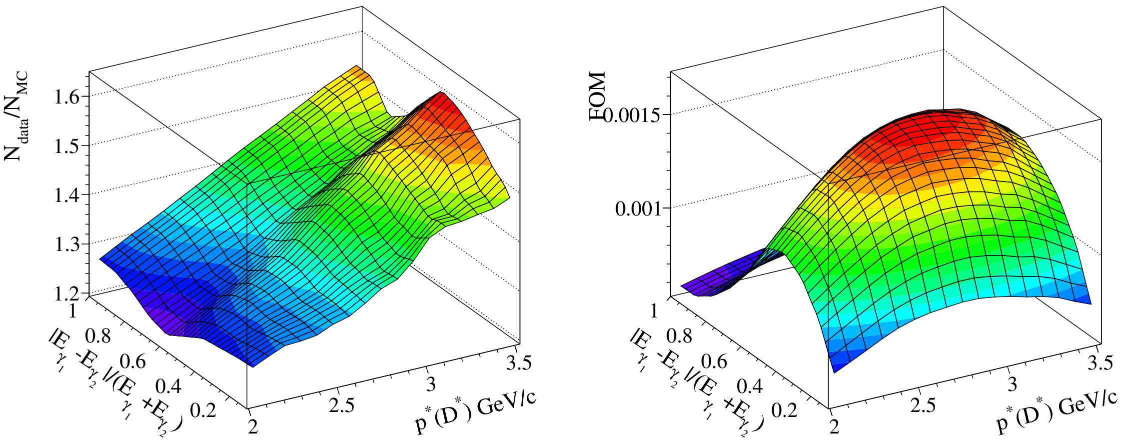


Figure 5.6: Correction factor and 2D optimization result obtained by applying the correction for $p^*(D^*)$ and $|E_{\gamma 1} - E_{\gamma 2}|/(E_{\gamma 1} + E_{\gamma 2})$ distributions.

to estimate the FOM in the signal window and the calibration factor $N_{\text{data}}/N_{\text{MC}}$ in the sideband. The set of cuts giving the highest value for FOM, after applying correction, is taken as the optimized set. The calibration factor and optimization result are shown in Fig. 5.6. The optimal cuts are $p^*(D^*) > 2.9\,\text{GeV}/c$ and $|E_{\gamma 1} - E_{\gamma 2}|/(E_{\gamma 1} + E_{\gamma 2}) < 0.5$.

5.7.2 Optimization of E_γ

The cut on the energy of second most-energetic photon candidates, E_γ, is optimized after applying the optimized set cuts on $p^*(D^*)$ and $|E_{\gamma 1} - E_{\gamma 2}|/(E_{\gamma 1} + E_{\gamma 2})$ distributions. The two photons in the $D^0 \to \gamma\gamma$ decay are sorted in ascending order of their energy, and so we perform optimization by varying cut on the lower-energy photon only. It is then obvious that the energy of other photon would be larger than the cut value. The calibration factor and optimization result are shown in Fig. 5.7 and the obtained cut is $E_\gamma > 0.9\,\text{GeV}$.

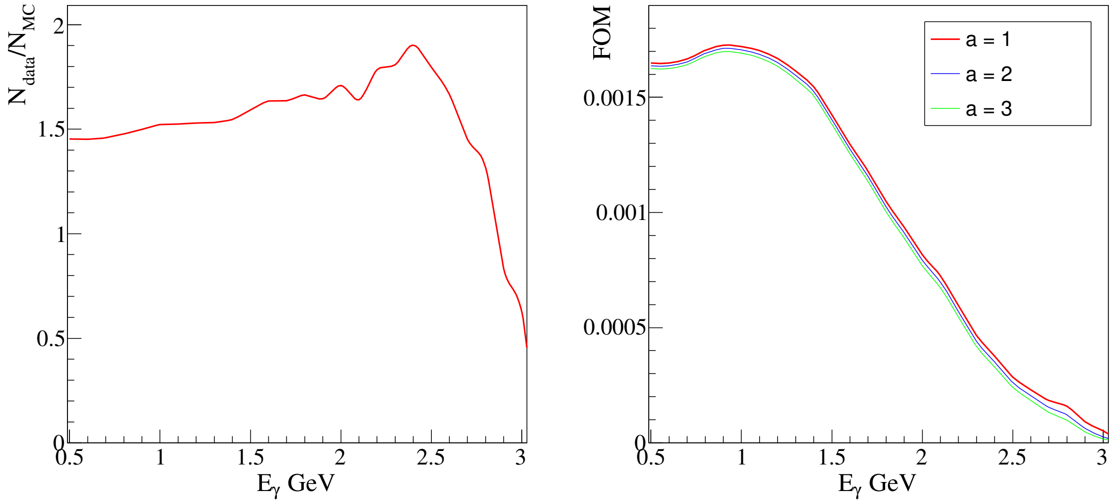


FIGURE 5.7: Correction factor and optimization result obtained by applying the correction for E_γ distribution.

5.7.3 Optimization of π^0 and η Vetoes

As described in Section 5.4, major physics background contributions are from $D^0 \to \pi^0\pi^0$, $D^0 \to \eta\pi^0$, $D^0 \to \eta\eta$ and $D^0 \to K_S^0(\pi^0\pi^0)\pi^0$. To suppress them, we use Koppenburg's π^0 and η veto [32] and optimize the cut on Prob(π^0) and Prob(η), the probability for a photon candidate to come from π^0 and η, respectively. Before this optimization, we apply the optimized cuts on $p^*(D^*)$, E_γ and $|E_{\gamma 1} - E_{\gamma 2}|/(E_{\gamma 1} + E_{\gamma 2})$. The optimization for Prob(π^0) is then performed by estimating the FOM for different cuts, simultaneously for both photons in $D^0 \to \gamma\gamma$. The correction factor and optimization results are shown in Fig. 5.8. The optimized cut obtained is Prob(π^0) < 0.15

The optimization result for the cut on Prob(η) using MC samples is shown in Fig. 5.9, where we don't find an obvious maximum as both signal and background events have similar distributions, as shown in the lower plot of Fig. 5.5. Thus we do not apply any cut on Prob(η).

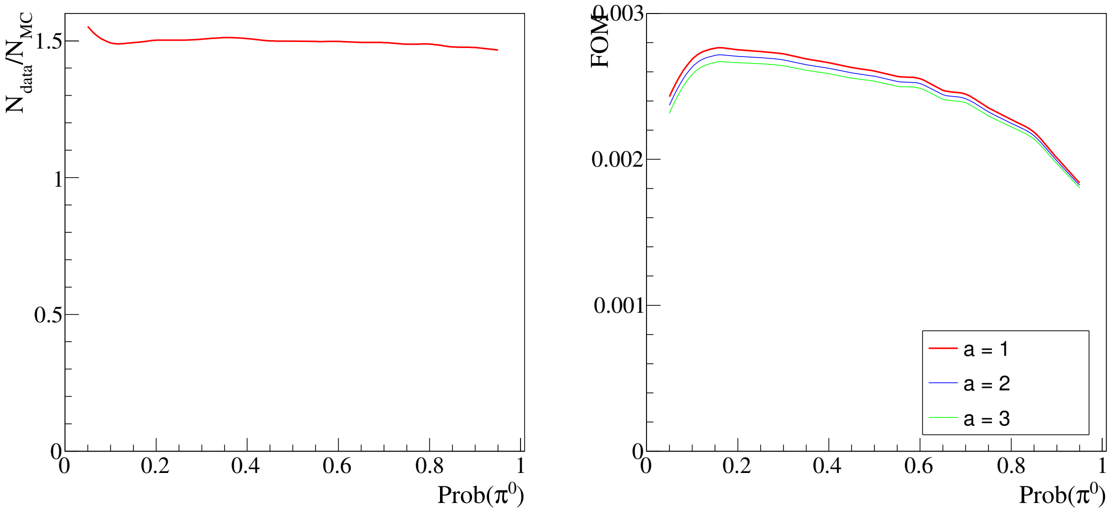


FIGURE 5.8: Correction factor and optimization result obtained by applying the correction for the Prob(π^0) distribution.

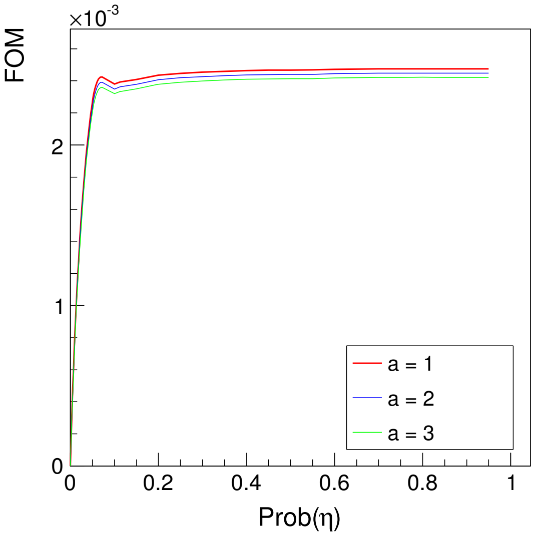


FIGURE 5.9: Optimization result obtained for Prob(η) using generic and signal MC events.

TABLE 5.4: Final set of selection cuts for $D^0 \to \gamma\gamma$ and the relative efficiency loss in signal and background with each of these cuts.

Variable	Cut	Signal eff. loss (%)	Bkg. eff. loss (%)
dr	< 1 cm	4.0	9.7
$\|dz\|$	< 3 cm	3.2	9.0
KID	< 0.6	0.2	2.6
# of charge track	> 4	5.7	2.5
# of neutral candidates	> 4	0.3	0.0
$E9/E25$	> 0.85	1.7	21.4
$p^*(D^*)$	> 2.9 GeV/c	47.1	84.2
$\frac{\lvert E_{\gamma 1} - E_{\gamma 2}\rvert}{(E_{\gamma 1} + E_{\gamma 2})}$	< 0.5	40.5	94.7
E_γ	> 0.9 GeV	6.2	28.7
Prob(π^0)	< 0.15	26.0	74.2

5.8 Best candidate selection

The candidate multiplicity distribution after applying the optimized set of cuts listed in Table. 5.4 are shown in Fig. 5.10. Only 2.7% of events have a multiplicity greater than one and it mostly arises due to D^0s produced due to a random combination of two photons. The best candidate is selected by retaining the one with the minimum value of transverse impact parameter, dr. The efficiency of best candidate selection (BCS), as defined in Eq. 4.13, is 70.5%. The $M(D^0)$ and ΔM distributions after applying an optimized set of cuts and BCS are shown in Fig. 5.11 (individual components are shown in Appendix D).

5.9 Fitting for the $D^0 \to \gamma\gamma$ decay

As described earlier, we use a 2D fitting to the $M(D^0)$ and ΔM distributions. The merged signal and five stream of generic MC samples are used to determine

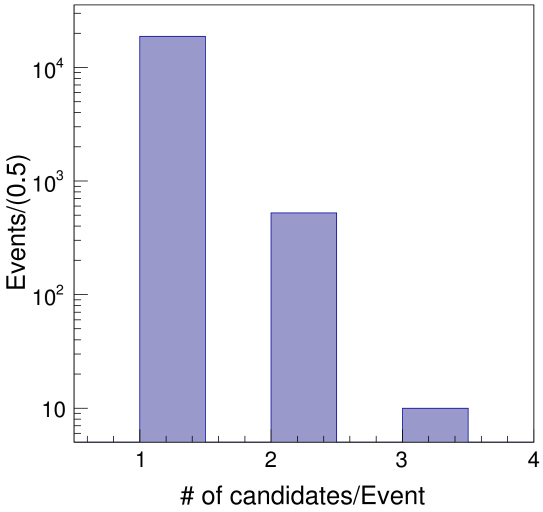


FIGURE 5.10: Candidate multiplicity distribution for $D^0 \to \gamma\gamma$ signal MC events.

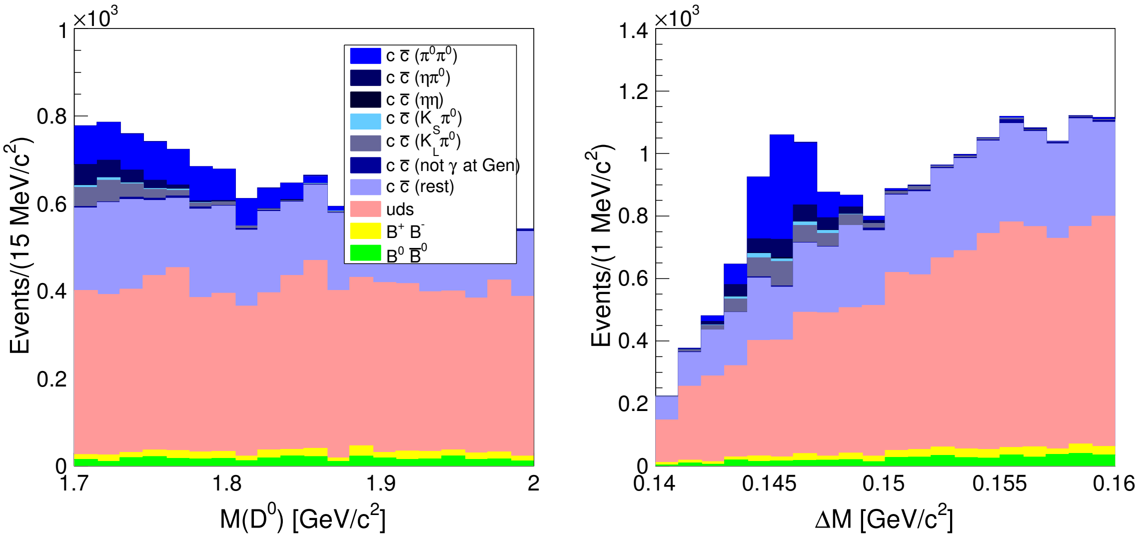


FIGURE 5.11: Classification of various backgrounds after applying the optimized set of cuts and best candidate selection.

the signal and peaking background PDF shapes, respectively. In general, a multi dimensional PDF is built of individual PDFs by considering the correlation between them. We check the correlation between the two fit variables as shown in Fig. 5.12. The upper left plot in Fig. 5.12 shows a 2D scatter histogram of ΔM vs. $M(D^0)$ for the $D^0 \to \gamma\gamma$ signal, from which it is not obvious that there is a strong correlation between them. We also check for this correlation in an alternate way by plotting the ΔM distribution in the bins of $M(D^0)$, as shown in the upper right plot of Fig. 5.12. The ΔM distribution shows a peak at the same position for different ranges of $M(D^0)$. The lower plots in Fig. 5.12 are the corresponding results for the peaking background. Unlike the signal, here we can see the width of the ΔM distribution has a considerable dependence on $M(D^0)$ while the mean is independent. Thus a correlated PDF is required to model the peaking background shape.

The signal PDF shape is described by a sum of a Gaussian and an asymmetric Gaussian function in both $M(D^0)$ and ΔM. An exponential and a threshold function are respectively used to model the small misreconstructed background contribution in these distributions. The 2D fit result for signal is shown in Fig. 5.13 from which we obtain a reconstruction efficiency of 7.3%.

A Gaussian function is used to model the peaking background in $M(D^0)$ while a function similar to signal is used in ΔM. In order to include the correlation between these variables, we parametrize the widths of ΔM as a function of $M(D^0)$, where the width of Gaussian is defined as $\sigma = \sigma 0(1 + k[M(D^0) - m(D^0)]^2)$, where m(D 0) is the world-average D^0 mass [26]. The two widths of the asymmetric Gaussian component are expressed as the multiples of σ. The result of the fit to the peaking background using the correlated PDF is presented in Fig. 5.14.

The remaining contribution is mainly from the combinatorial background of $c\bar{c}$ and uds origin. In order to model this component, we use a third-order Chebychev polynomial in $M(D^0)$ and a threshold function in the ΔM distribution. Figure 5.15

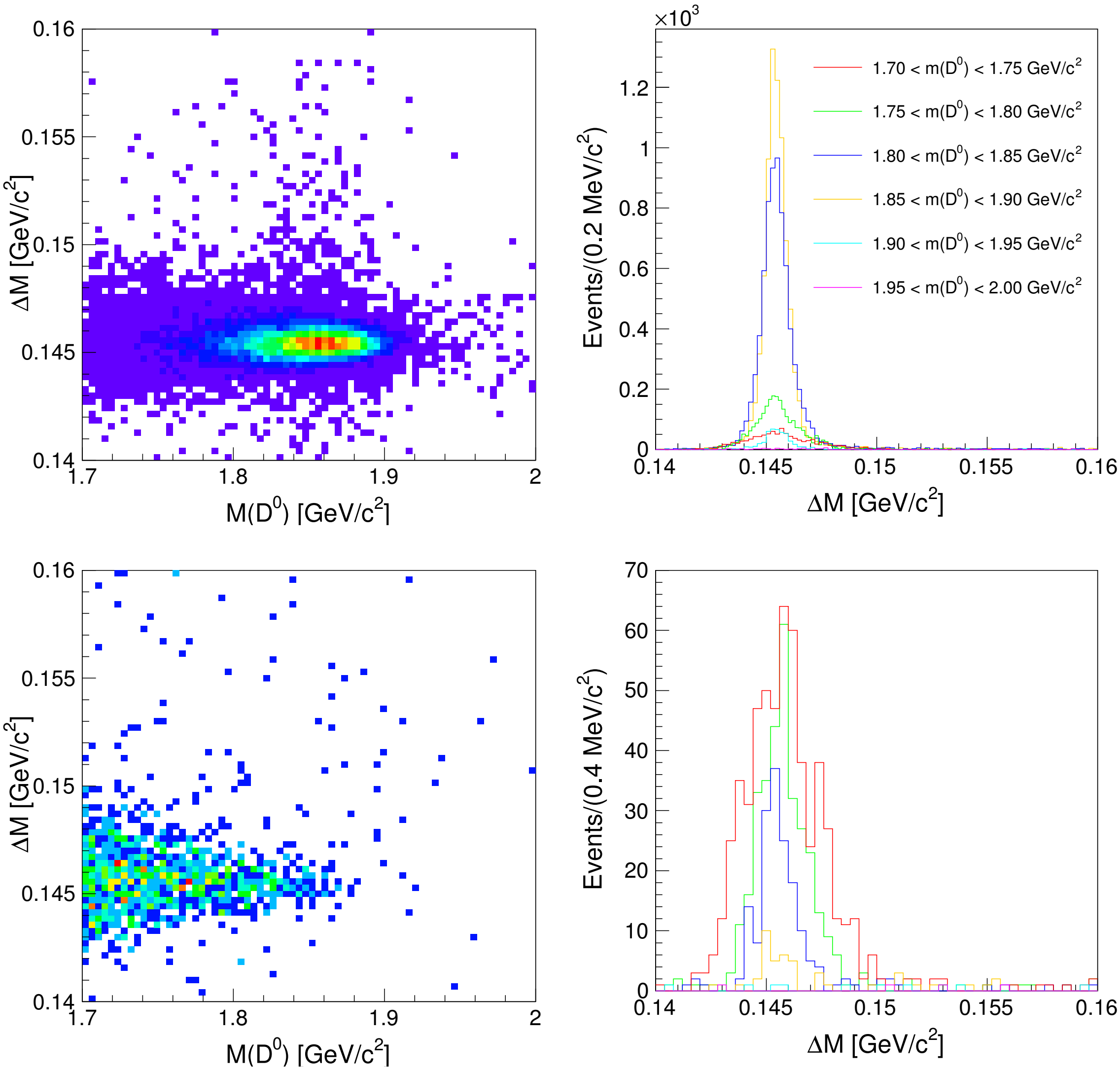


FIGURE 5.12: (upper left) 2D scatter plot between ΔM and $M(D^0)$ distributions, and (upper right) ΔM distribution in the bins of $M(D^0)$ in signal MC events. (lower left) and (lower right) show the corresponding plots for the peaking background.

shows the 2D fit result obtained for the combinatorial background using five stream of generic MC events.

A combined PDF for the 2D fit is constructed by adding the individual components for signal, peaking and combinatorial background. The shape parameters for signal and peaking background are fixed to their respective MC values, while the combinatorial parameters are varied during the fit. In addition, we vary the

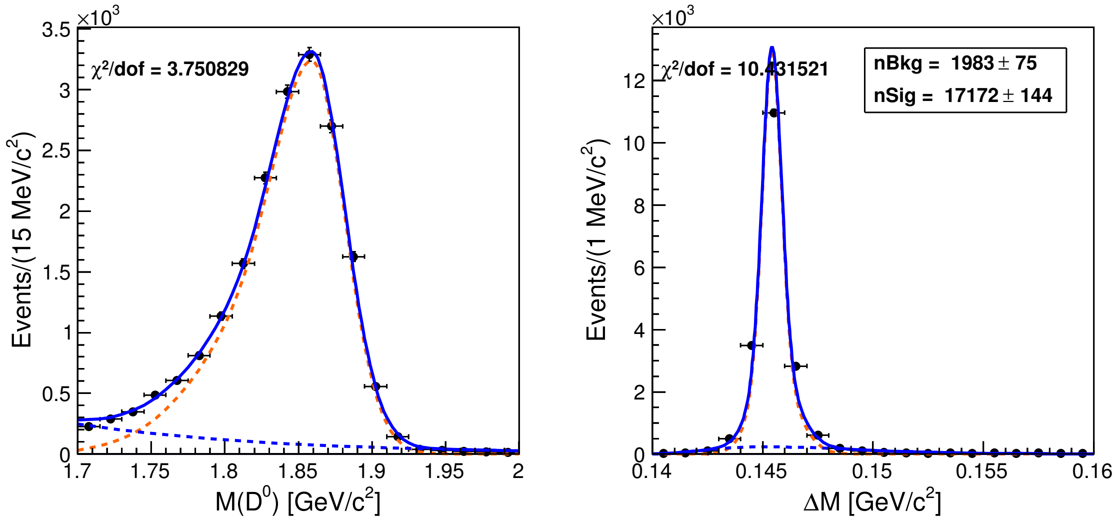


FIGURE 5.13: 2D fit to the signal MC sample of $D^0 \to \gamma\gamma$.

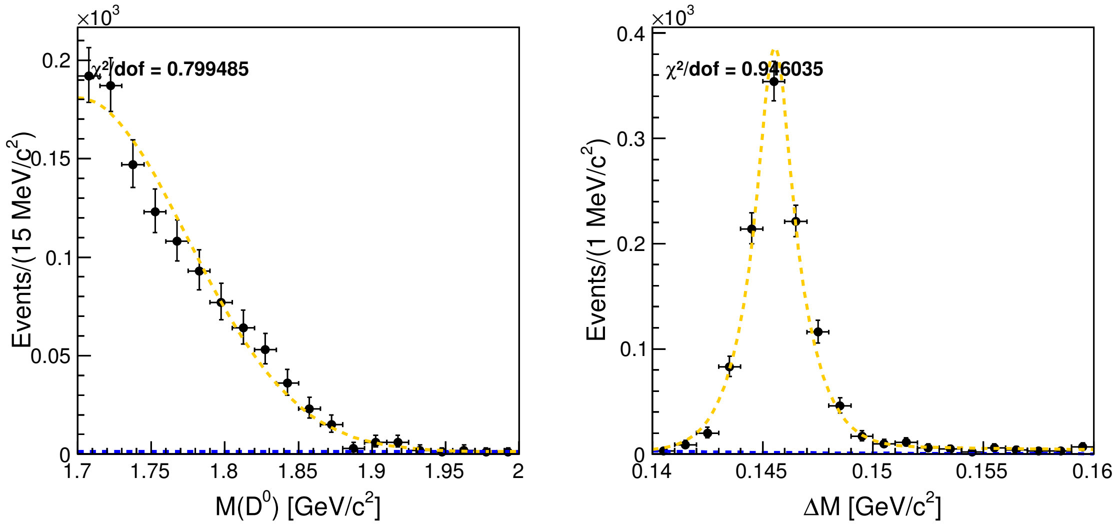


FIGURE 5.14: 2D fit for the peaking background using five streams of generic MC events.

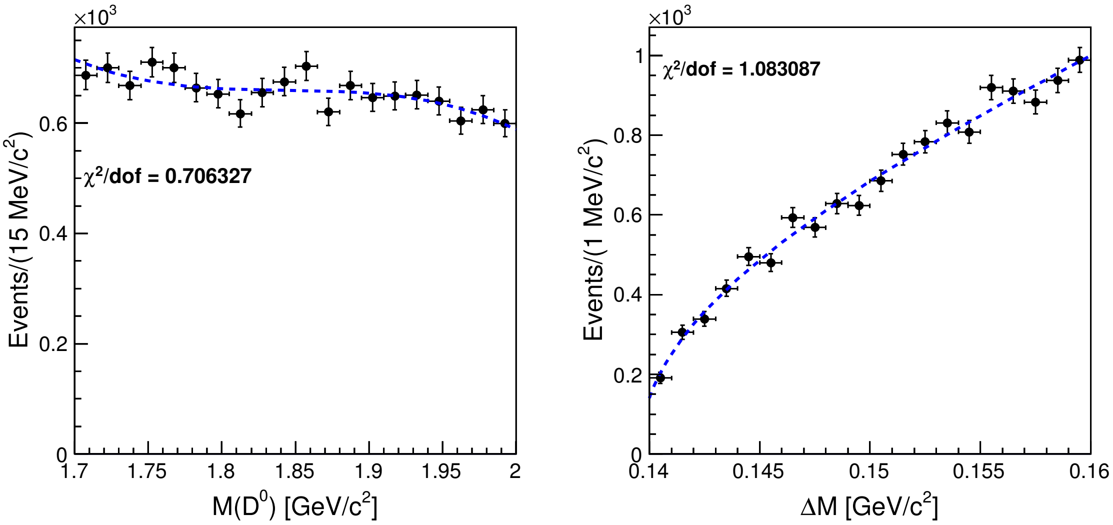


FIGURE 5.15: 2D fit for the combinatorial background using five streams of generic MC events.

TABLE 5.5: Summary of fit results to different streams of generic MC events.

Stream #	nSig	nBkg_peak	nBkg_comb
0	-4 ± 13	214 ± 28	2642 ± 56
1	-6 ± 12	249 ± 28	2632 ± 56
2	0 ± 12	216 ± 28	2684 ± 56
3	2 ± 12	203 ± 26	2528 ± 54
4	-13 ± 12	251 ± 29	2675 ± 56

normalizations of the three event components. Each of the MC streams, which are statistically equivalent to the data, are fitted using the combined PDF. As an example, the fit to 0^{th} stream generic MC is presented in Fig. 5.16, while the fit results to all streams are listed Table 5.5.

5.9.1 GSIM ensemble test

We performed a GSIM ensemble test to check the stability of our fitter. In this test, we pick the expected number of signal events ($\text{nSig}_{\text{expected}}$) from the signal

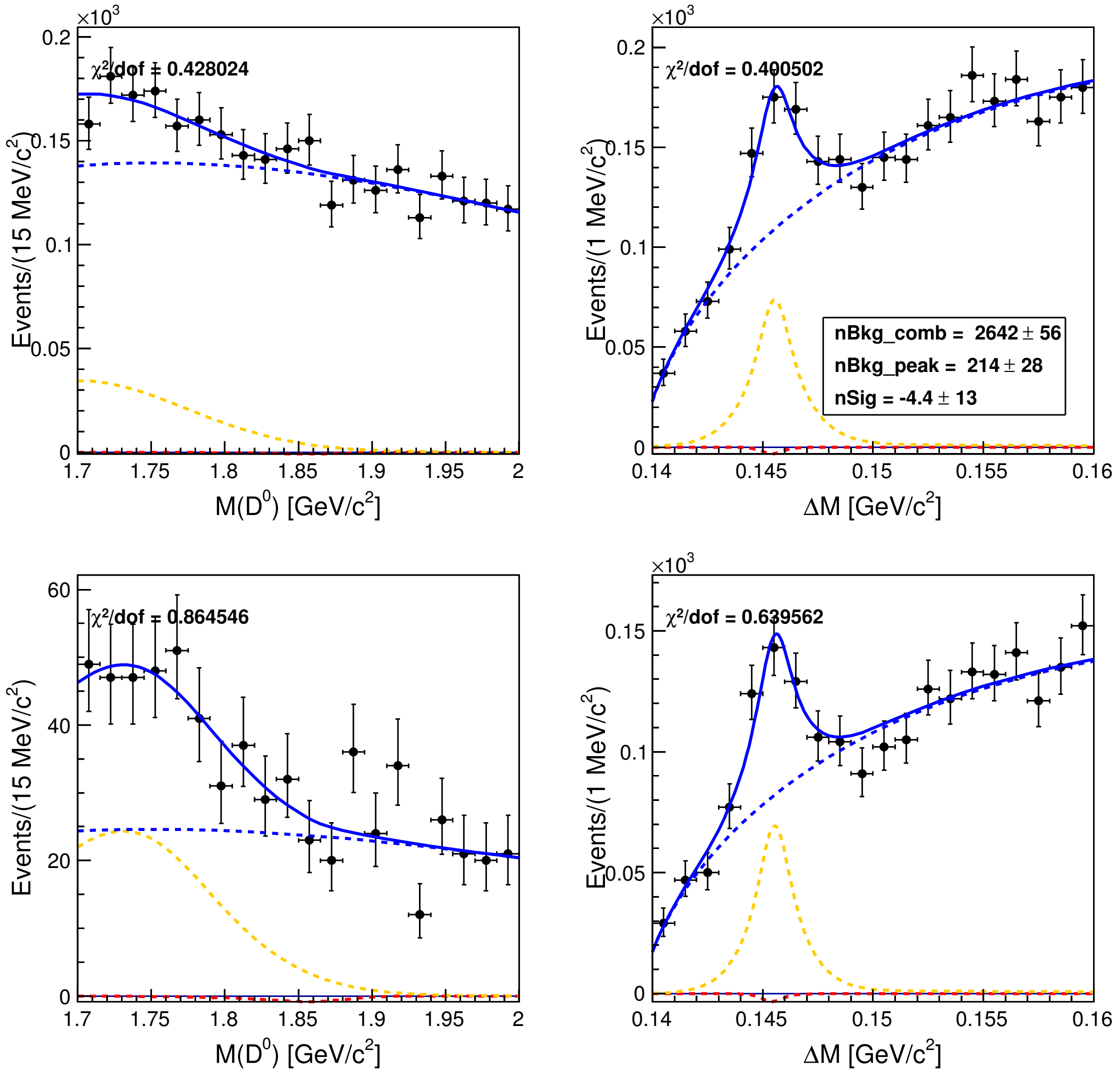


FIGURE 5.16: Results of the 2D fit based on a correlated PDF for the 0^{th} stream of generic MC events. Lower plots show the fit projection in one fit variable applying a $\pm 3\sigma$ signal region cut on the other.

MC sample and generate the background events using their PDF shapes, to prepare a combined pseudodata sample. The process is repeated for making several such independent samples[1]. Each sample is then fit with the PDF, while keeping the signal and background yields as well as the shape parameters of the combinatorial background free. The distribution of the fitted signal yields (nSig_{fit}) is fit

---

[1] Maximum number of such samples in a set is equal to the ratio of number of events in signal MC sample to $\text{nSig}_{\text{expected}}$.

with a Gaussian function to get the mean yield ($\text{nSig}_{\text{mean}}$) along with the associated statistical error. The closeness of $\text{nSig}_{\text{mean}}$ to $\text{nSig}_{\text{expected}}$ gives a measure of stability of the fitter while any considerable deviation will be considered as a bias in. We also check the linearity of our fitter by repeating the GSIM test for different signal yield hypotheses ($\text{nSig}_{\text{expected}} = 5, 10, 15...$). Then $\text{nSig}_{\text{mean}}$ is plotted versus $\text{nSig}_{\text{expected}}$ as shown in Fig. 5.17, which is fitted to a straight line yielding $\chi^2/\text{ndf} = 0.165$. It is clear from the plot that the fitter is linear over the considered range of signal yield.

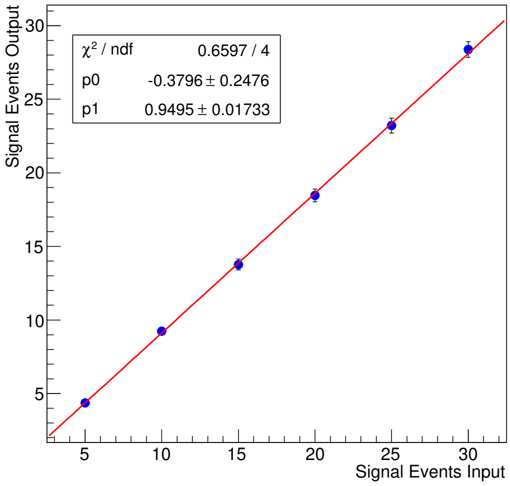


FIGURE 5.17: Linearity test performed for the $D^0 \to \gamma\gamma$ fitter.

5.10 Corrections for the data fit

As in the combined PDF we fix the shape parameters of the signal and peaking background to their MC values, we need to take care the possible data-MC difference. We use various control samples to obtain the correction factors and apply them to the fitter. The $D^0 \to \phi\gamma$ sample is used to correct for the signal shape in both $M(D^0)$ and ΔM distributions. On the other hand, a partially reconstructed sample of $D^0 \to \pi^0\pi^0$ is used to correct for the peaking background shape in

$M(D^0)$ while a reconstructed sample of the forbidden decay $D^0 \to K_S^0\gamma$ is used for correcting its ΔM distribution.

5.10.1 Correction factors from $D^0 \to \phi\gamma$

We chose the $D^0 \to \phi\gamma$ sample to correct for the signal PDF shapes in $D^0 \to \gamma\gamma$ after appropriate modifications. Figure 5.18 shows the projections of a three-dimensional (3D) fit to $M(D^0)$, ΔM and $\cos\theta_{\rm hel}$ distributions using this sample. For a $D^0 \to \phi(K^+K^-)X$ $(X = \pi^0, \gamma\ ...)$ decay, one can define the helicity angle, $\theta_{\rm hel}$, as the angle between the momentum of K^+ and the momentum of D^0 in the ϕ rest frame. The $D^0 \to \phi\gamma$ decay has a peaking contribution from $D^0 \to \phi\pi^0$. The signal follows a $\cos^2\theta_{\rm hel}$ distribution whereas the latter follows a $\sin^2\theta_{\rm hel}$ distribution. Therefore, we can use this shape difference to distinguish the two decays. In the 3D fit to model signal, we use a Crystal Ball function for $M(D^0)$, a sum of Gaussian and asymmetric Gaussian function for ΔM, and a $\cos^2\theta_{\rm hel}$ for the $\cos\theta_{\rm hel}$ distribution. As shown by the red dashed curve in Fig. 5.18, the tail parameters of the $M(D^0)$ and ΔM signal shapes are fixed from the $D^0 \to \phi\gamma$ MC sample.

For the peaking background (blue dashed curve in Fig. 5.18) from $D^0 \to \phi\pi^0$, we use a Crystal Ball function for $M(D^0)$, a sum of Gaussian and asymmetric Gaussian function for ΔM, and a $\sin^2\theta_{\rm hel}$ function for the $\cos\theta_{\rm hel}$ distribution. All shape parameters in $M(D^0)$ and ΔM are fixed to the corresponding values from the $D^0 \to \phi\pi^0$ MC sample. While fixing the peaking background shape, we apply corrections using correction factors obtained from the $D^0 \to K_S^0\gamma$ sample listed in Table 5.9. The combinatorial background (green dashed curve in Fig. 5.18) is modeled by the sum of an exponential and a constant function in $M(D^0)$, a threshold function in ΔM, and a sum of a constant and $\cos^2\theta_{\rm hel}$ function in the $\cos\theta_{\rm hel}$ distribution.

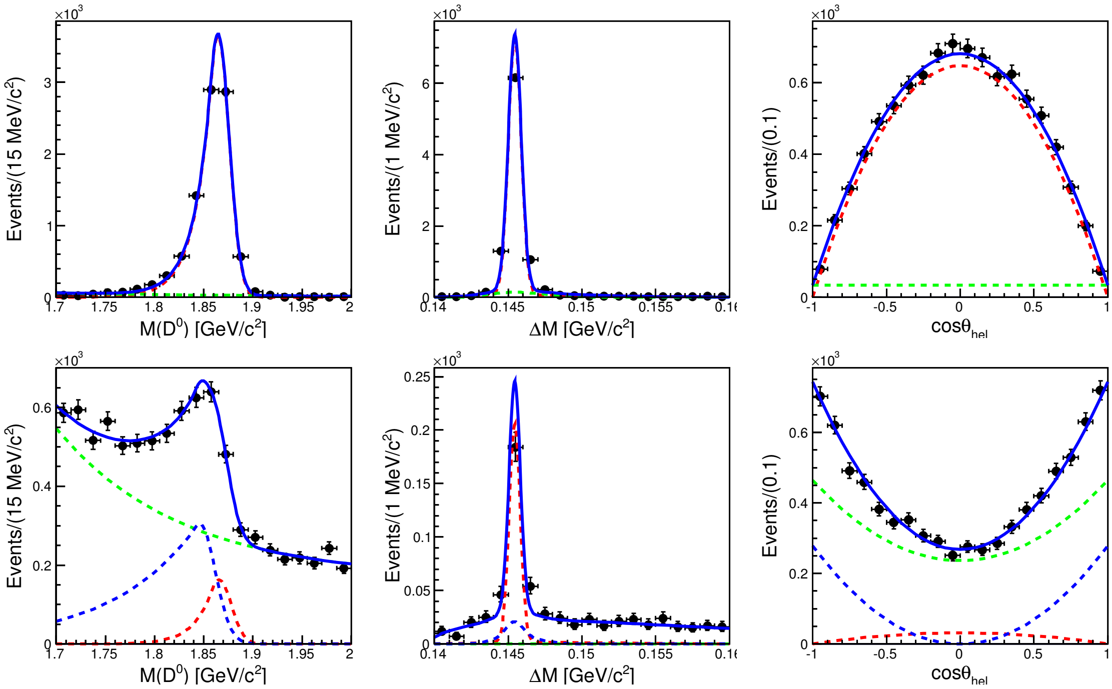


FIGURE 5.18: Results of the 3D fit with $M(D^0)$, ΔM and $\cos\theta_{\rm hel}$ for $D^0 \to \phi\gamma$: (upper) in MC events and (lower) in data.

TABLE 5.6: Correction factors obtained with the $D^0 \to \phi\gamma$ decay.

	$M(D^0)$	ΔM
Mean(MeV/c^2)	$+1.20 \pm 1.20$	-0.005 ± 0.029
Width	1.093 ± 0.089	0.861 ± 0.065

We fix the width of the signal model to the MC value as shown in Fig. 5.18(upper) and give it the freedom to vary through a multiplicative factor “width_mD0" (“width_DelM") for $M(D^0)$ (ΔM). Similarly, we fix the mean of the signal shape using the MC values and give it the freedom to vary through an addictive factor “mean_mD0" (“mean_DelM") for $M(D^0)$ (ΔM). 3D fit to the $D^0 \to \phi\gamma$ data sample is shown in Fig. 5.18(lower) and the obtained correction factors are listed in Table. 5.6.

5.10.2 Corrections factors from $D^0 \to \pi^0\pi^0$

The $D^0 \to \gamma\gamma$ decay has a peaking contribution from $D^0 \to \pi^0\pi^0$, as shown in Fig. 5.11. In the $D^0 \to \pi^0\pi^0$ decay, both the pions decay through $\pi^0 \to \gamma\gamma$. The misreconstruction of two most energetic photons from each π^0 can mimic a $D^0 \to \gamma\gamma$ event. A partially reconstructed $D^0 \to \pi^0\pi^0$ sample is thus used to calibrate the peaking background. We produce the $D^0 \to \gamma\gamma$ sample by picking up the most energetic photon (γ_{high}) from each of π^0 in $D^0 \to \pi^0\pi^0$ [neglect both low energy photons (γ_{low})]. An optimized set of cut is applied on high energy photons while the energy cut on the low energy photons is varied until a best match between the peaking background and the partially reconstructed $D^0 \to \pi^0\pi^0$ sample is achieved. A best match is found (see Fig. 5.19) at a cut $E_\gamma < 2$ GeV on γ_{low}. The details are given in Appendix **??**. The selection cuts for the $D^0 \to \pi^0\pi^0$ sample are summarized in Table 5.7.

By applying the cuts listed in Table 5.7, we produce a partially reconstructed sample of $D^0 \to \pi^0\pi^0$ using five streams of generic MC events. We fit the $M(D^0)$ distribution using an asymmetric Gaussian function as shown in Fig. 5.20(left). Then we fit the data sample using the PDF obtained from the MC sample by giving the freedom to its mean through an additive parameter and width through a multiplicative factor. The resultant correction factors are listed in Table 5.8.

5.10.3 Correction factors from $D^0 \to K_S^0\gamma$

We reconstruct the $D^0 \to K_S^0\gamma$ decay that is originally misreconstructed $D^0 \to K_S^0\pi^0$ and $D^0 \to K_S^0\eta$ decays. We apply the selection cuts listed in Table 5.4 to chose photons with similar kinematics as those in the $D^0 \to \gamma\gamma$ process. The $M(D^0)$ and ΔM distributions for $D^0 \to K_S^0\gamma$ are shown in Fig. 5.21. The sky-blue (pink) component denotes the $D^0 \to K_S^0\pi^0$ ($D^0 \to K_S^0\eta$) contribution. We can see that different contributions peak at different regions in $M(D^0)$ but at the same

TABLE 5.7: Optimized set of cuts for $D^0 \to \pi^0\pi^0$.

Variable	Cut value
dr	< 1 cm
$\|dz\|$	< 3 cm
$p^*(D^*)$	$> 2.9\,\mathrm{GeV}/c$
$p(\pi^0)$	$> 0.3\,\mathrm{GeV}/c$
$m(\pi^0)$	$(110, 160)\,\mathrm{MeV}/c^2$
For γ_{high}	
E_γ	$> 0.9\,\mathrm{GeV}$
$\|E_{\gamma 1} - E_{\gamma 2}\|/(E_{\gamma 1} + E_{\gamma 2})$	< 0.5
For γ_{low}	
E_γ	$< 0.21\,\mathrm{GeV}$

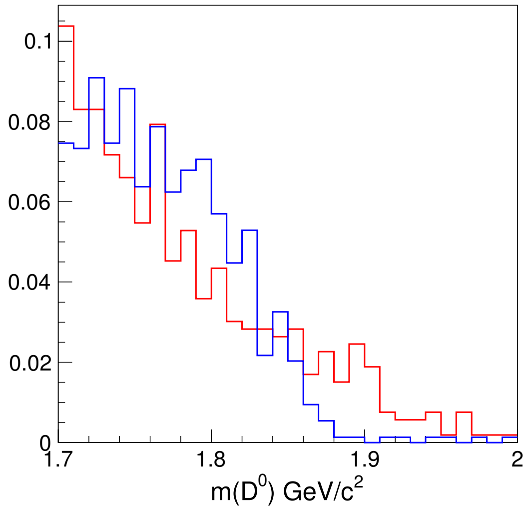


FIGURE 5.19: Comparison of the $D^0 \to \pi^0\pi^0$ component from the $D^0 \to \gamma\gamma$ (blue) and partially reconstructed $D^0 \to \pi^0\pi^0$ (red) sample.

TABLE 5.8: Correction factors obtained with the $D^0 \to \pi^0\pi^0$ decay.

	$M(D^0)$
Mean (GeV/c^2)	0.0119 ± 0.0025
Width	1.053 ± 0.020

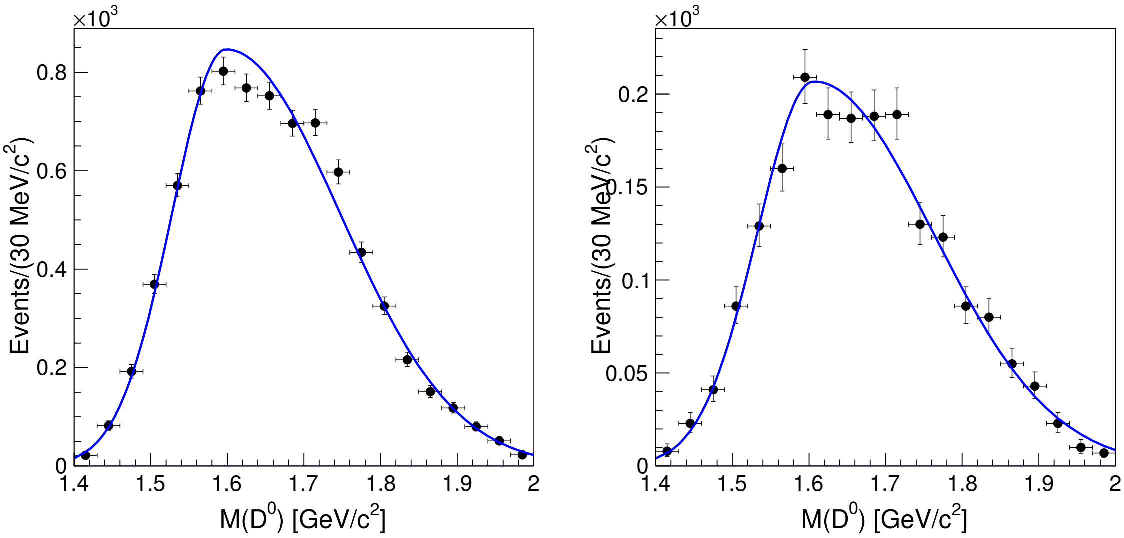


FIGURE 5.20: Fit to the partially reconstructed $D^0 \to \pi^0\pi^0$ sample: (left) MC and (right) data.

region in ΔM. In order to separate the $K^0_S\pi^0$ component, we study the $M(D^0)$ and ΔM distributions by applying the requirements $0.1425 < \Delta M < 0.1475\,\text{GeV}/c^2$ and $1.8 < M(D^0) < 1.9\,\text{GeV}/c^2$, respectively (see Fig. 5.22).

The $K^0_S\eta$ component in the ΔM distribution is completely removed by the cut as shown in Fig. 5.22(right). In $M(D^0)$, the $K^0_S\pi^0$ component is modeled as a sum of a Gaussian and an asymmetric Gaussian function with different mean, while we use an asymmetric Gaussian function for the $K^0_S\eta$ component. The ΔM distribution has only the $K^0_S\pi^0$ contribution under the peak as shown in Fig. 5.22(right); to model it we use a sum of a Gaussian and an asymmetric Gaussian function with common mean for signal and a threshold function for background.

Simultaneous fits are performed with the $D^0 \to K^0_S\gamma$ MC sample to determine the PDF shape parameters. The $D^0 \to K^0_S\gamma$ data sample is fitted with these PDFs by giving the freedom to the mean by an additive factor and to the width through a multiplicative factor. Figure 5.23 shows results of the fit to the data sample, and the correction factors are summarized in Table 5.9.

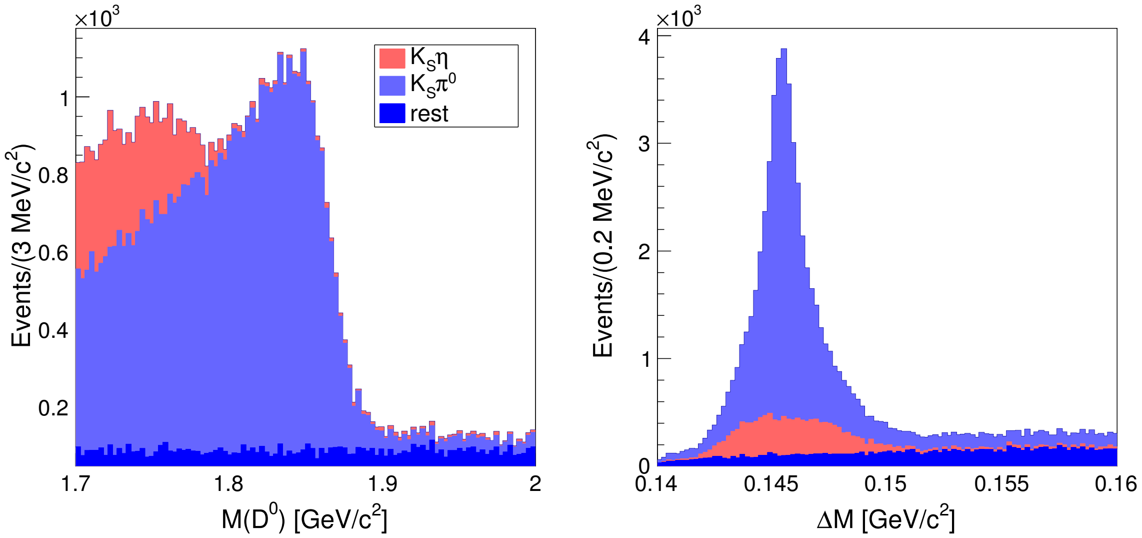


FIGURE 5.21: Classification of various contributions in the $D^0 \to K_S\gamma$ sample: (left) $M(D^0)$ and (right) ΔM.

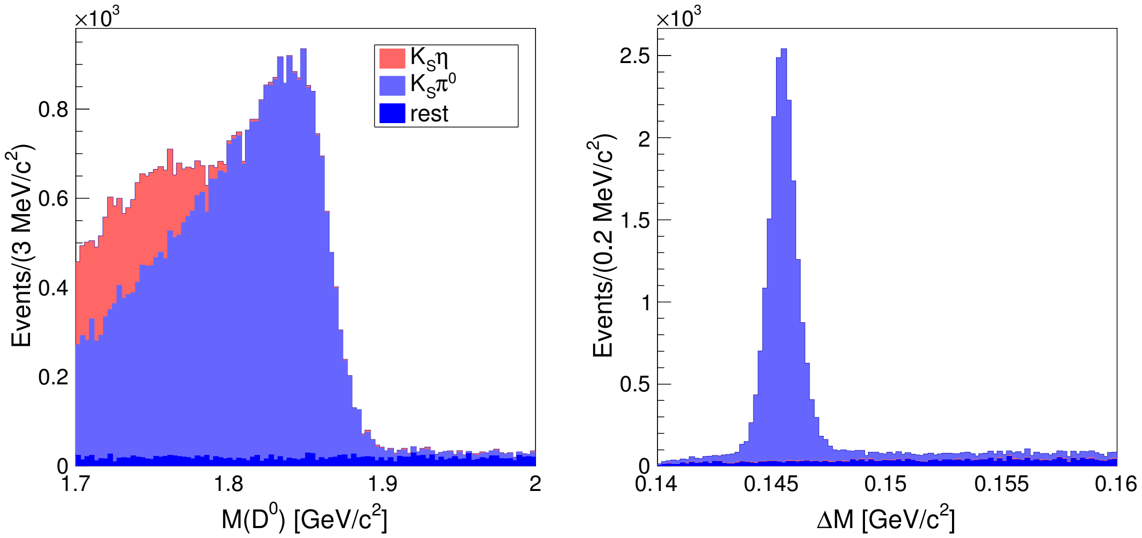


FIGURE 5.22: Classification of various contributions in the $D^0 \to K_S\gamma$ sample: (left) in $M(D^0)$ by applying the cut $0.1425 < \Delta M < 0.1475\,\mathrm{GeV}/c^2$ and (right) in ΔM with the cut $1.8 < M(D^0) < 1.9\,\mathrm{GeV}/c^2$.

TABLE 5.9: Correction factors obtained with the $D^0 \to K_S^0\gamma$ sample.

	$M(D^0)$	ΔM
Mean(MeV/c²)	1.700 ± 0.340	0.007 ± 0.005
Width	1.062 ± 0.014	1.078 ± 0.009

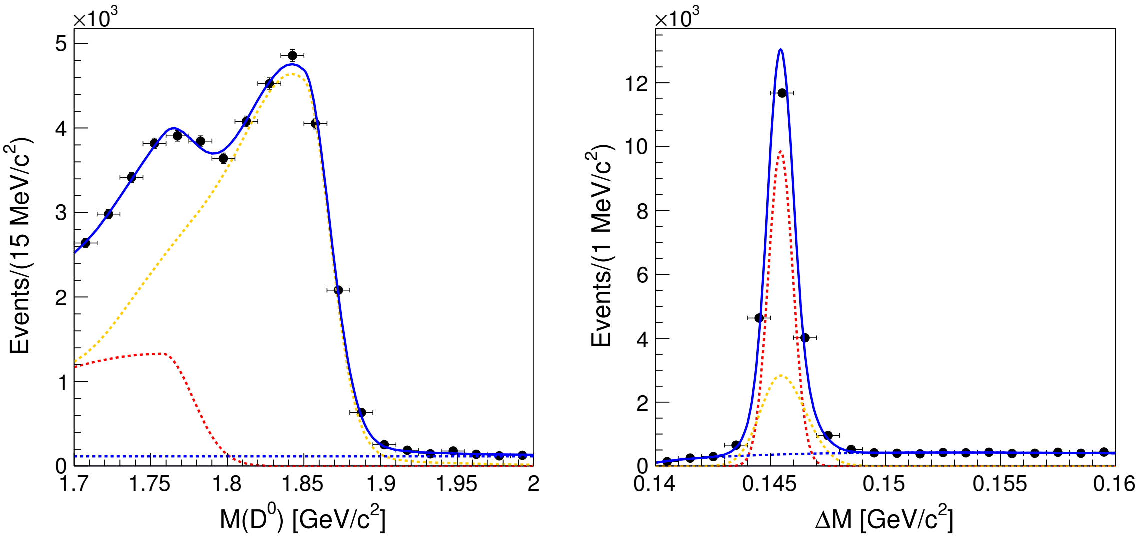


FIGURE 5.23: Fit to the $D^0 \to K^0_S\gamma$ data sample: (left) $M(D^0)$ and (right) ΔM.

5.11 Fit to data

For the data, we apply a cut on “ECL timing" for both signal photons, in addition to the cuts summarized in Table. 5.4. It is defined as the time interval between an e^+e^- collision and the moment the photons reach the ECL. The ECL timing of a typical $D^0 \to \gamma\gamma$ cluster falls in the range (9000, 11000)ns. So we retain only those events with the ECL timing of γ cluster falling in this range. This cut suppresses the off-time clusters[2] as well as reduces the level of combinatorial background. Note that this criterion is not applied to experiments 7 to 27, where the timing information is not available. Events that pass or are rejected by the ECL timing cut are shown in Fig. 5.24

For the signal PDF we apply correction factors listed in Table 5.6. As described earlier, these factors are estimated using $D^0 \to \phi\gamma$ that has one photon in the final state while we want to apply them to a decay with two photons. Thus the mean of $M(D^0)$ is shifted to twice, and the width is multiplied by the square, of

---

[2] These are the clusters arising from QED processes like $e^+e^- \to \gamma\gamma$ and $e^+e^- \to e^+e^-\gamma$.

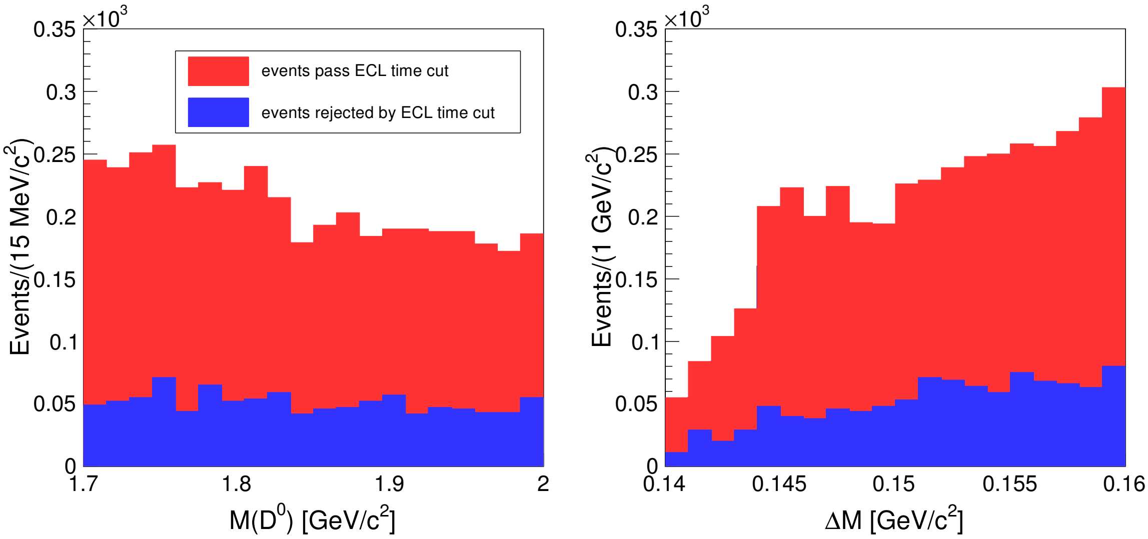


Figure 5.24: Events in data that pass and fail the ECL timing cut.

the corresponding correction factor. In case of ΔM the resolution is dominated by the soft pion and has very little dependence on the D^0 final state. So here we apply the corrections without any modification.

For the peaking background, we use correction factors obtained from $D^0 \to \pi^0\pi^0$ sample (listed in Table 5.8) for the $M(D^0)$ distribution, while its ΔM shape is corrected by using $D^0 \to K_S^0\gamma$ sample (given in Table 5.9). As described in Section 5.9, we use a special parametrization for this component in ΔM. In the parametrization the core width is represented by a parameter $\sigma 0$. So the ΔM PDF is corrected for by multiplying $\sigma 0$ with the correction factor.

Figure 5.25 shows results of the fit in data after applying proper corrections to different PDFs. We get a signal yield of 4 ± 15 events and 210 ± 32 and 2934 ± 59 events for the peaking and combinatorial background, respectively. Henceforth, the obtained signal yield (4) is referred to as $\mathrm{nSig}_{\mathrm{data}}$.

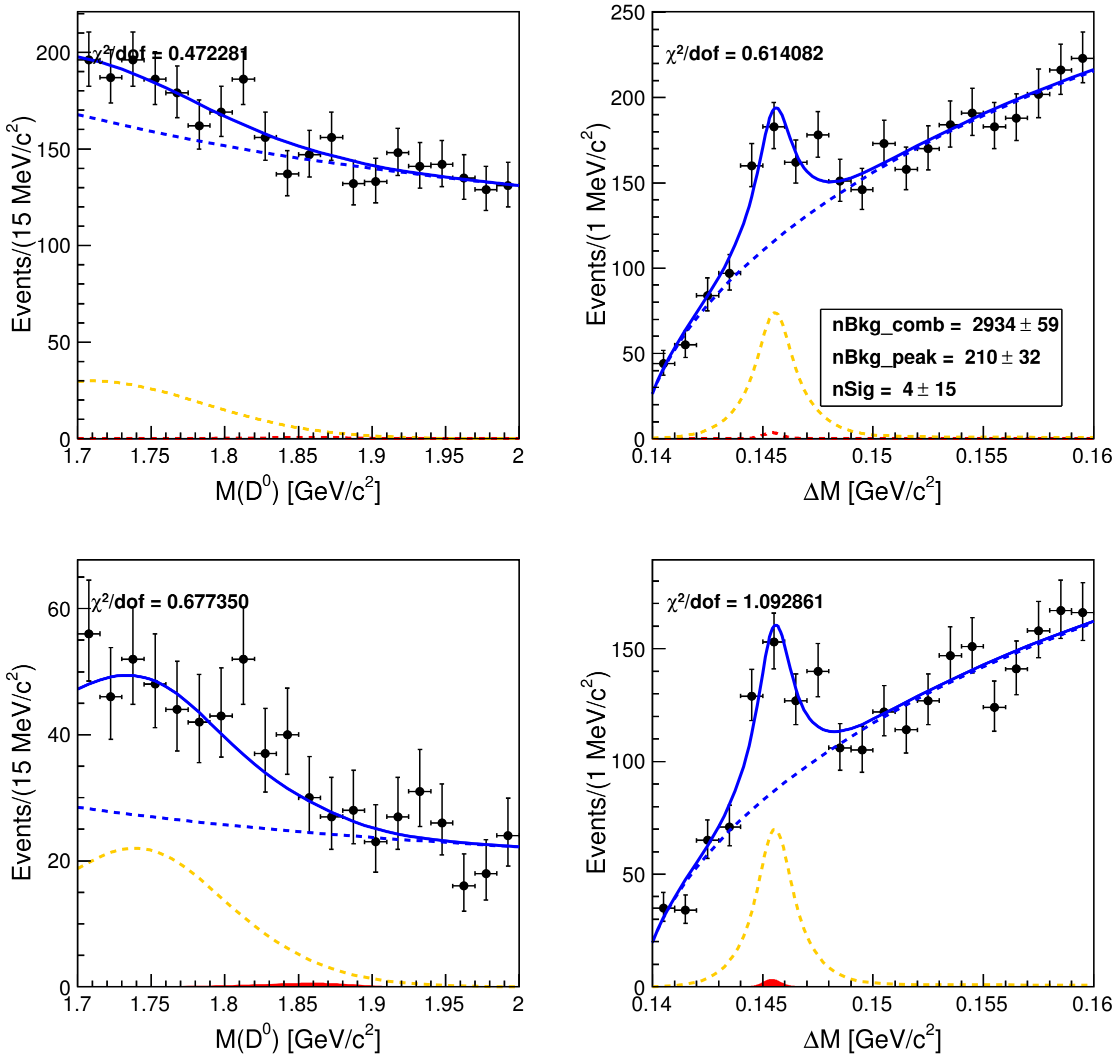


FIGURE 5.25: 2D fit results for $D^0 \to \gamma\gamma$ in data. Lower plots show the projections in one fitting variable while applying a $\pm 3\sigma$ signal region cut on the other one.

5.12 Normalization mode $D^0 \to K_S^0\pi^0$

For the $D^0 \to K_S^0\pi^0$ sample, we use the same set of cuts as in Ref. [25] that are listed in Table 5.10. A 2D fit is performed to obtain the signal efficiency and yield. The signal PDF shape is described by a sum of a Gaussian and an asymmetric Gaussian function in both $M(D^0)$ and ΔM. An exponential and a threshold function are respectively used to model the small misreconstructed

TABLE 5.10: Set of cuts applied on $D^0 \to K_S^0\pi^0$ sample.

Variable	Cut value
dr	< 1 cm
$\|dz\|$	< 3 cm
$p^*(D^*)$	$> 2.9\,\mathrm{GeV}/c$
$E_\gamma^{\mathrm{barrel}}$	$> 60\,\mathrm{MeV}$
$E_\gamma^{\mathrm{endcap}}$	$> 100\,\mathrm{MeV}$
$p(\pi^0)$	$> 0.5\,\mathrm{GeV}/c$
$m(\pi^0)$	$(110, 160)\,\mathrm{MeV}/c^2$
goodKs	applied

contribution in these distributions. We use a merged signal MC sample of 234k events (see Figure 5.26) to estimate the detection efficiency, which is found to be 7.2%. Figure 5.27 shows result of the 2D fit for the $D^0 \to K_S^0\pi^0$ data sample. We obtain a signal yield of 343050 ± 673 events.

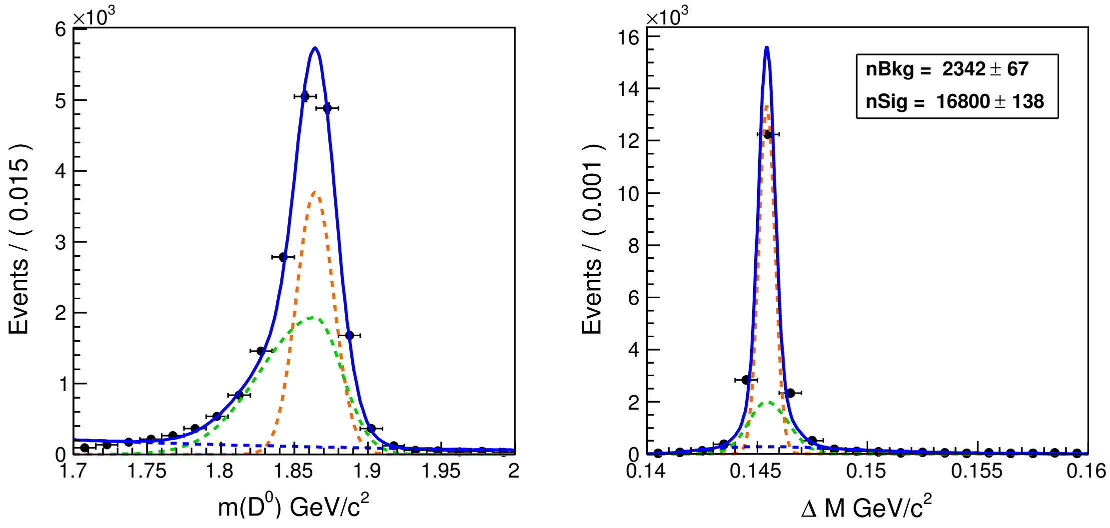


FIGURE 5.26: Results of the 2D fit to the $D^0 \to K_S^0\pi^0$ signal MC sample.

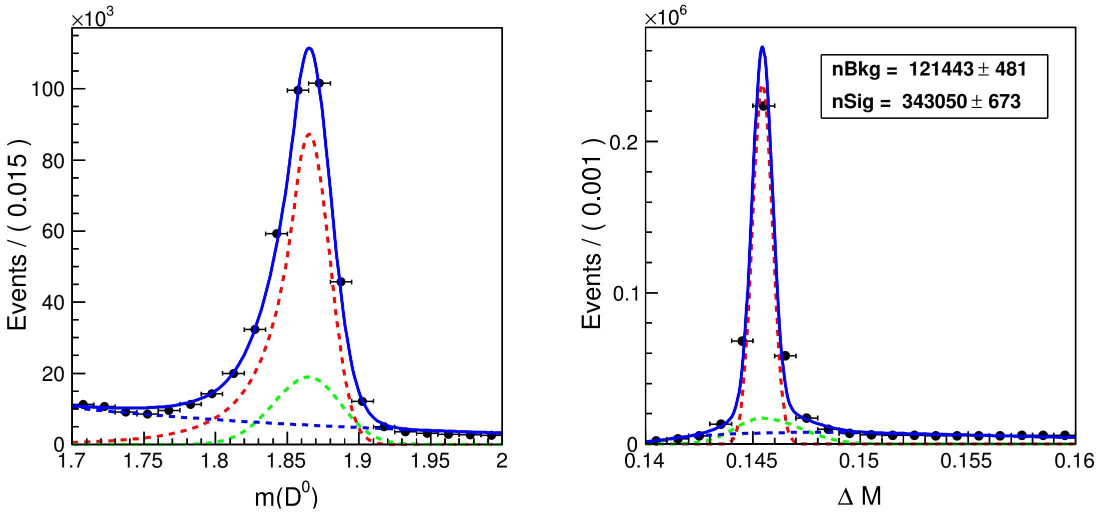


FIGURE 5.27: Results of the 2D fit to the $D^0 \to K^0_S\pi^0$ data sample.

5.13 Upper limit calculation

We use a frequentist method [34] to estimate an upper limit on the branching fraction for $D^0 \to \gamma\gamma$. In this method, we replace the signal yield $N(D^0 \to \gamma\gamma)$ in the expression for branching fraction [see Eq. (5.1)] with $N^{90\%}_{\rm UL}(D^0 \to \gamma\gamma)$, the yield at 90% confidence level. The UL expression for the branching fraction is given by:

$$\mathcal{B}^{90\%}_{\rm UL}(D^0 \to \gamma\gamma) < \frac{N^{90\%}_{\rm UL}(D^0 \to \gamma\gamma)}{N(D^0 \to K^0_S\pi^0)} \times \frac{\varepsilon_{K^0_S\pi^0}}{\varepsilon_{\gamma\gamma}} \times \mathcal{B}(D^0 \to K^0_S\pi^0)_{\rm WA} \tag{5.3}$$

We generate a set of 5000 toy MC samples based on the PDFs used in the data fit, by putting different nSig values (nSig$_{\rm i}$), viz., 1, 2, ... while fixing the background yields to the corresponding values obtained from data. Each toy MC sample is then fit with the same PDF (from which the samples are generated) by keeping the signal and background yields, and the shape parameters of the combinatorial background free to get the fitted signal yield, nSig$_{\rm fit}$. For each set we determine the percentage of toy MC samples having nSig$_{\rm fit}$ grater than $nSig_{data}$. The nSig$_{\rm i}$ value

TABLE 5.11: Summary of quantities for the estimation of upper limit on $\mathcal{B}(D^0 \to \gamma\gamma)$ in data.

Quantity	Value
$N_{\mathrm{UL}}^{90\%}(D^0 \to \gamma\gamma)$	24
$N(D^0 \to K_S^0\pi^0)$	343050
$\varepsilon_{K_S^0\pi^0}$	7.18%
$\varepsilon_{\gamma\gamma}$	7.34%
$\mathcal{B}(D^0 \to K_S^0\pi^0)_{\mathrm{WA}}$	1.19×10^{-2}

for which the percentage just crosses 90% is $N_{\mathrm{UL}}^{90\%}(D^0 \to \gamma\gamma)$. We find the latter to be 24, and substituting it in Eq. (5.3) we get $\mathcal{B}_{\mathrm{UL}}^{90\%}(D^0 \to \gamma\gamma) < 8.1 \times 10^{-7}$. Various quantities used for the upper limit calculation are listed in Table. 5.11.

5.14 Systematic uncertainties

The branching fraction measurement with respect to $D^0 \to K_S^0\pi^0$ cancels many systematic uncertainties such as the one associated with fragmentation of $c\bar{c}$ and with various cuts common to $D^0 \to \gamma\gamma$ and the reference mode. The estimation of the remaining contributions are described in the following sections.

5.14.1 PDF shape

We fix the signal and peaking background shape from the $D^0 \to \gamma\gamma$ generic MC sample, respectively. The possible difference in shape between data and simulations should be considered as a systematic uncertainty. To estimate the error due to PDF shape we take results of the fit in Fig. 5.25 as the nominal one. Then we vary the correction factors for the mean and width by $\pm 1\sigma$ separately for $M(D^0)$ and ΔM, and estimate the corresponding signal yields by fitting. The deviation of these yields with respect to the nominal value are calculated, and the +ve and

Table 5.12: Signal yields by varying correction factors by $\pm 1\sigma$.

	$M(D^0)$		ΔM	
	Yield	Deviation	Yield	Deviation
Signal PDF				
mean $-\sigma$	5 ± 15	1	4 ± 15	0
mean $+\sigma$	4 ± 15	0	4 ± 15	0
width $-\sigma$	3 ± 13	-1	4 ± 14	0
width $+\sigma$	5 ± 16	1	5 ± 15	1
Peaking background PDF				
mean $-\sigma$	6 ± 15	2	4 ± 15	0
mean $+\sigma$	3 ± 15	-1	4 ± 15	0
width $-\sigma$	7 ± 15	3	4 ± 15	0
width $+\sigma$	2 ± 15	-2	4 ± 15	0

$-$ve deviations are added in quadrature to get the systematic error. A summary of the calculation is given in Table. 5.12. The total systematic error due to PDF shape is $+4.0$ and -2.4 candidates.

5.14.2 Cut variation

This uncertainty is canceled out for the cuts that are common to $D^0 \to \gamma\gamma$ and the normalization mode $D^0 \to K_S^0\pi^0$ such as $p^*(D^*)$. For the remaining ones, different methods are used to estimate the associated systematic error. We use the $D^0 \to \phi\gamma$ control sample to calculate the uncertainty associated with the cuts on E_γ and Prob(π^0).

We estimate the ratio $N_{\rm Sig}/\varepsilon$, where $N_{\rm Sig}$ is the signal yield and ε is the detection efficiency, with and without the cut. The difference in the ratio with respect to the nominal value divided by the latter gives the systematic error. We find that the removal of cut on Prob(π^0) significantly increases the level of background and

TABLE 5.13: Summary of systematic errors associated with the cuts.

	Systematic error (%)	
	for 1γ	for 2γ
E_γ	± 2.0	± 4.0
Prob(π^0)	± 1.8	± 3.6
$\lvert E_{\gamma 1} - E_{\gamma 2}\rvert/(E_{\gamma 1} + E_{\gamma 2})$	—	± 4.2
		± 6.8

makes it difficult to extract signal yield. Thus we apply a loose cut Prob(π^0) $<$ 0.7, at which the fit converges. The estimated uncertainties are summarized in Table 5.13. The systematic error calculated using $D^0 \to \phi\gamma$ is doubled in order to get systematics for the signal.

The above method can't be used to estimate the uncertainty due to the $|E_{\gamma 1} - E_{\gamma 2}|/(E_{\gamma 1} + E_{\gamma 2})$ cut, since we don't have a control sample with two γs in the final state. For this we estimate the upper limit on the branching fraction without applying any cut on $|E_{\gamma 1} - E_{\gamma 2}|/(E_{\gamma 1} + E_{\gamma 2})$. The difference of the obtained value with respect to the nominal one divided by the latter is taken as the systematics. The systematic error is found to be -4.2%.

5.14.3 Systematic errors associated with candidate selections and $\mathcal{B}(D^0 \to K_S^0\pi^0)$

Further, there is systematic uncertainty due to the efficiency for photon detection, K_S^0 and π^0 reconstruction in case of the signal and normalization channels. The systematic error due to photon detection is about 2.2% for $E_\gamma = 1$GeV [35]. With two energetic photons in the signal final state, we allocate a 4.4% uncertainty. The uncertainty associated with K_S^0 reconstruction estimated with a sample of $D^{*+} \to D^0\pi_s^+$, $D^0 \to K_S^0(\pi^+\pi^-)\pi^+\pi^-$ is 0.7%. We obtain the systematic error due

TABLE 5.14: Summary of systematic errors associated with candidate selections and $\mathcal{B}(D^0 \to K_S^0\pi^0)$.

	Systematic error (%)
γ detection	± 4.4
K_S^0 reconstruction	± 0.7
π^0 identification	± 4.0
$\mathcal{B}(D^0 \to K_S^0\pi^0)$	± 3.3
Total	± 6.8

to π^0 reconstruction (4.0%) by comparing data-MC differences of the yield ratio between $\eta \to \pi^0\pi^0\pi^0$ and $\eta \to \pi^+\pi^-\pi^0$. The last one is the error on the branching fraction of the normalization channel $D^0 \to K_S^0\pi^0$ [26]. Table 5.14 lists various systematic sources along with their contributions.

5.14.4 Summary of systematic uncertainties

We combine systematic uncertainties due cut variation, candidate selections and $\mathcal{B}(D^0 \to K_S^0\pi^0)$ in quadrature and get the error $\pm 9.6\%$. We multiply this numbers with the nominal yield, 4, and later combine with systematics due to PDF shape. We get two numbers $(+\sqrt{(4 \times 0.096)^2 + (4)^2}$ and $-\sqrt{(4 \times 0.096)^2 + (2.4)^2})$, of which the largest value obtained, 4 events, is taken as the systematic uncertainty.

5.15 Including the systematic error in UL

To estimate upper limit, we generate a set of 5000 toy MC samples from the PDF with different values of nSig (0,1,5...) while keeping the background yields fixed. The percentage of toy samples in each set having $\text{nSig}_{\text{fit}} > 4$ (4 is the fitted signal yield) gives the confidence level. Now for including systematics to UL calculation, we generate a set of 5000 random numbers from a Gaussian function with mean zero and width equals to 4. Each of these numbers, say "a", is added to the 5000 $\mathrm{nSig_{fit}}$ values obtained from the earlier toy MC set. We estimate the percentage of toy samples in each set having $\mathrm{nSig_{fit}} + \mathrm{a} > 4$, and then follow the procedure described in Section 5.13 to obtain a new upper limit on the branching fraction, $\mathcal{B}^{90\%}_{UL+a}(D^0 \to \gamma\gamma)$. Subsequently we get $N^{90\%}_{\mathrm{UL}+a}(D^0 \to \gamma\gamma) = 25$ and $\mathcal{B}^{90\%}_{\mathrm{UL}+a}(D^0 \to \gamma\gamma) < 8.4 \times 10^{-7}$. We also estimate the expected upper limit on branching fraction (assuming $\mathrm{nSig_{data}} = 0$) as $\mathcal{B}^{90\%}_{UL+a}(D^0 \to \gamma\gamma) < 6.8 \times 10^{-7}$.

5.16 Conclusion

In summary, we have searched for the rare decay $D^0 \to \gamma\gamma$ using the full data sample recorded by the Belle experiment at or above the $\Upsilon(4S)$ resonance. In absence of a statistically significant signal, a 90% CL upper limit is set on its branching fraction at 8.5×10^{-7}. Our result constitutes the most restrictive limit on $D^0 \to \gamma\gamma$ to date and can be used to constrain NP parameter spaces. This FCNC decay will be probed further at the next-generation flavor factories including Belle II [36].

APPENDIX A

Efficiency map of M. Starič

The map for soft pion detection efficiency asymmetry, estimated separately for the SVD1 and SVD2 data samples is shown in bins of $[p(\pi_s), \cos\theta(\pi_s)]$ in Fig. A.1 (provided by M. Starič). In our case, SVD1 and SVD2 maps merged together by weighting them according to the respective integrated luminosities:

$$(A_{\epsilon}^{\pi_s})_{ij} = (A_{\epsilon}^{\pi_s}(SVD1))_{ij} \times \frac{\mathcal{L}_1}{\mathcal{L}_1 + \mathcal{L}_2} + (A_{\epsilon}^{\pi_s}(SVD2))_{ij} \times \frac{\mathcal{L}_2}{\mathcal{L}_1 + \mathcal{L}_2} \tag{A.1}$$

Here i, j are the indices vary from i, j=1,..,5 that indicate two-dimensional bins $[p(\pi_s), \cos\theta(\pi_s)]$. $A_{\epsilon}^{\pi_s}(SVD1)$ and $A_{\epsilon}^{\pi_s}(SVD2)$ are efficiency asymmetries for the SVD1 and SVD2 data samples, respectively. $\mathcal{L}_1$ and $\mathcal{L}_2$ are integrated luminosities for the SVD1 and SVD2 set of data.

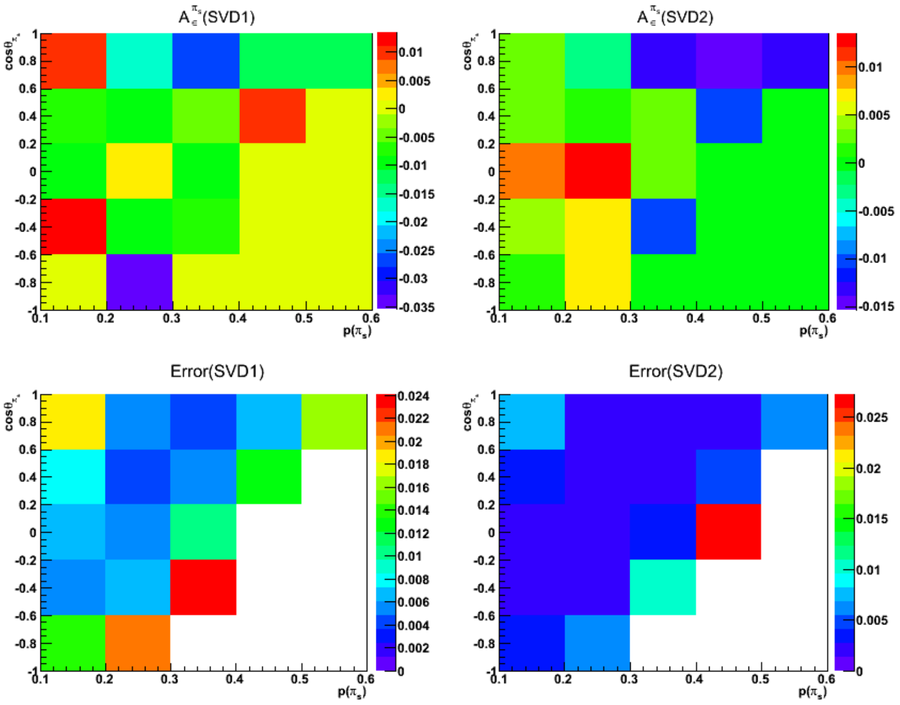


FIGURE A.1: $A_\epsilon^{\pi_s}$ map: upper plots are $A_\epsilon^{\pi_s}$ for SVD1 and SVD2 data samples separately, while the lower plots are associated errors.

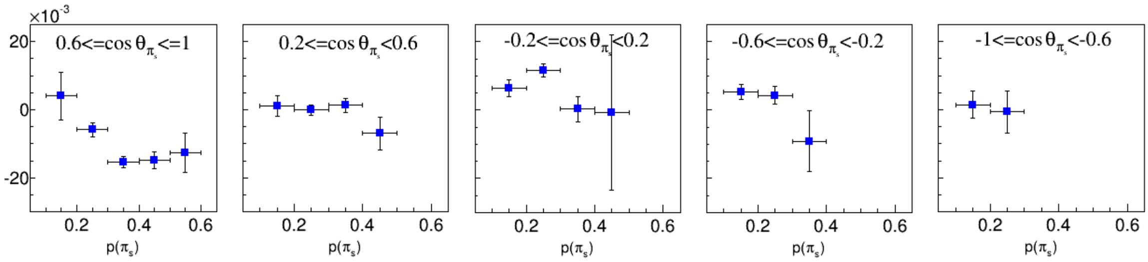


FIGURE A.2: $A_\epsilon^{\pi_s}$ map: after merging SVD1 and SVD2 data samples.

Table A.1: (upper) $A_\epsilon^{\pi_s}$ central values (in %) in bins of $[p(\pi_s), \cos\theta(\pi_s)]$, after SVD1 and SVD2 information are merged (lower) the error (in %) associated with $A_\epsilon^{\pi_s}$.

$(A_\epsilon^{\pi_s})_{ij}$ (%)		$p(\pi_s)$ (GeV/c)				
		[0.1,0.2)	[0.2,0.3)	[0.3,0.4)	[0.4,0.5)	[0.5,0.6]
$\cos\theta(\pi_s)$	[0.6,1]	0.402757	-0.588794	-1.53964	-1.48203	-1.25754
	[0.2,0.6)	0.120096	-0.010515	0.128475	-0.693619	0
	[-0.2,0.2)	0.642195	1.15899	0.0278915	-0.0757908	0
	[-0.6,-0.2)	0.541392	0.433606	-0.914378	0	0
	[-1,-0.6)	0.143174	-0.0564213	0	0	0

σ_{ij} (%)		p(π_s) (GeV/c)				
		[0.1,0.2)	[0.2,0.3)	[0.3,0.4)	[0.4,0.5)	[0.5,0.6]
$\cos\theta(\pi_s)$	[0.6,1]	0.705959	0.197722	0.169041	0.249957	0.581422
	[0.2,0.6)	0.303347	0.15755	0.205142	0.483471	0
	[-0.2,0.2)	0.243375	0.183584	0.377601	2.27402	0
	[-0.6,-0.2)	0.217086	0.249931	0.892603	0	0
	[-1,-0.6)	0.396584	0.619303	0	0	0

APPENDIX B

Efficiency map of B. R. Ko

TABLE B.1: $A_\epsilon^{\pi_s}$ central values (in %) in bins of $[p_T(\pi_s), \cos\theta(\pi_s)]$.

	$(A_\epsilon^{\pi_s})_{ij}$ (%)	$p_T(\pi_s)$ (GeV/c)							
		$[0.0, 0.15)$	$[0.15, 0.175)$	$[0.175, 0.195)$	$[0.195, 0.215)$	$[0.215, 0.235]$	$[0.235, 0.26]$	$[0.26, 0.295]$	$[0.295, \infty]$
$\cos\theta(\pi_s)$	$[-1, 0.25]$	0.309	1.065	1.194	0.353	-0.235	0.273	0.00.578	0.755
	$[-0.25, 0.1)$	0.221	0.00.910	0.756	1.715	1.107	1.031	-0.030	-0.043
	$[0.1, 0.35)$	-0.002	1.407	-0.322	1.227	0.809	0.693	0.170	0.037
	$[0.35, 0.55)$	0.566	-0.320	-0.783	0.061	0.100	-0.059	0.286	0.227

$(A_\epsilon^{\pi_s})_{ij}$ (%)		$p_T(\pi_s)$ (GeV/c)							
		$[0.0, 0.15)$	$[0.15, 0.175)$	$[0.175, 0.195)$	$[0.195, 0.215)$	$[0.215, 0.235]$	$[0.235, 0.26]$	$[0.26, 0.295]$	$[0.295, \infty]$
$\cos\theta(\pi_s)$	$[0.55, 0.7)$	2.052	-0.226	-2.169	-0.910	-0.487	-0.464	0.01.002	0.456
	$[0.7, 0.8)$	1.439	0.167	-2.084	-0.599	-0.394	-0.728	0.600	0.888
	$[0.8, 1)$	0.590	-0.391	-1.520	-1.531	-1.929	-0.521	-0.122	-0.249

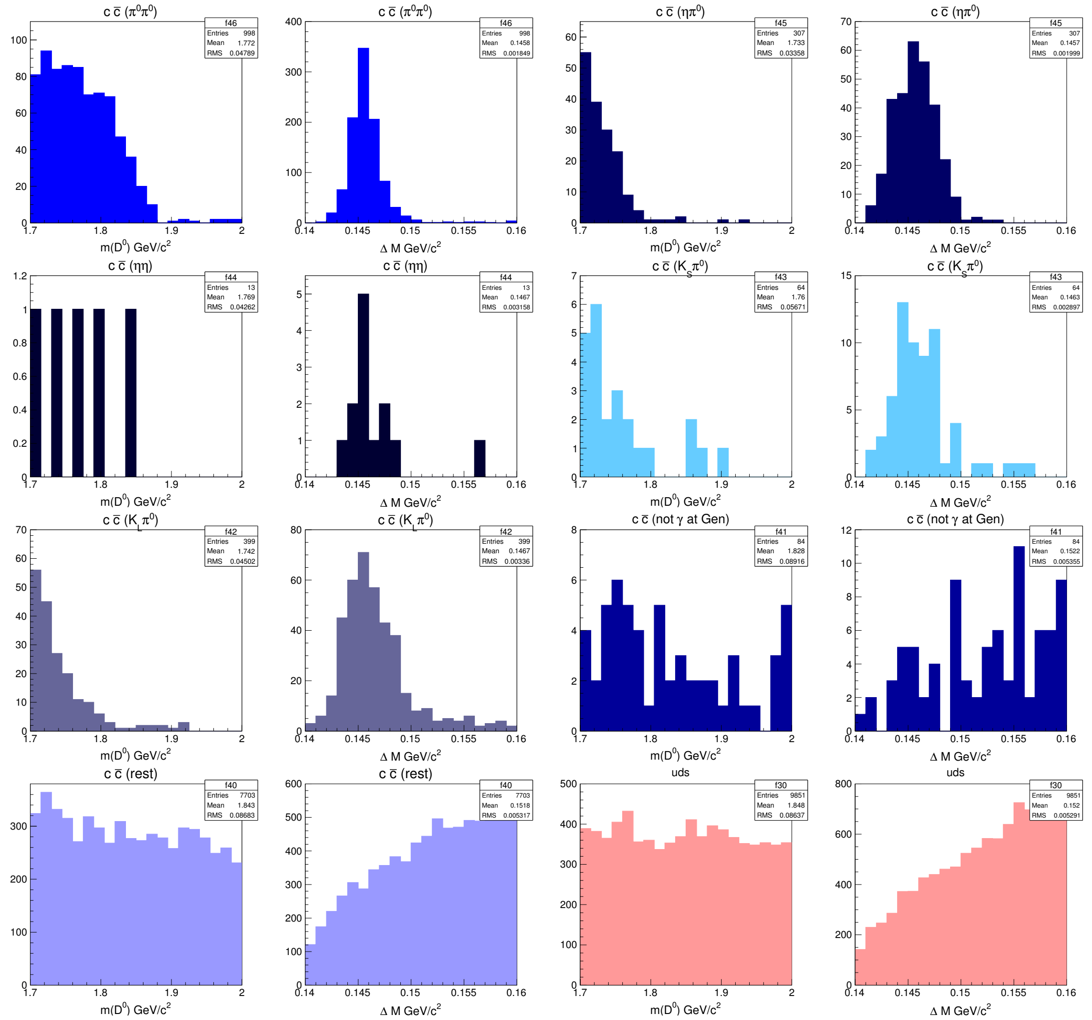
TABLE B.2: Error (in %) associated with $A_\epsilon^{\pi_s}$ in the bins of $[p_T(\pi_s), \cos\theta(\pi_s)]$.

σ_{ij}		$p(\pi_s)$ (GeV/c)							
		$[0, 0.15)$	$[0.15, 0.175)$	$[0.175, 0.195)$	$[0.195, 0.215)$	$[0.215, 0.235)$	$[0.235, 0.26)$	$[0.26, 0.295)$	$[0.295, \infty]$
$\cos\theta(\pi_s)$	$[-1, -0.25)$	0.281	0.378	0.458	0.520	0.630	0.716	0.899	1.691
	$[-0.25, 0.1)$	0.531	0.490	0.488	0.451	0.447	0.430	0.439	0.476

		$p(\pi_s)$ (GeV/c)							
	σ_{ij}	$[0, 0.15)$	$[0.15, 0.175)$	$[0.175, 0.195)$	$[0.195, 0.215)$	$[0.215, 0.235)$	$[0.235, 0.26)$	$[0.26, 0.295)$	$[0.295, \infty]$
$\cos\theta(\pi_s)$	$[0.1, 0.35)$	0.851	0.605	0.563	0.522	0.490	0.433	0.400	0.365
	$[0.1, 0.35)$	0.851	0.605	0.563	0.522	0.490	0.433	0.400	0.365
	$[0.35, 0.55)$	1.091	0.651	0.585	0.525	0.497	0.432	0.390	0.346
	$[0.55, 0.7)$	1.008	0.623	0.568	0.511	0.486	0.433	0.414	0.397
	$[0.7, 0.8)$	0.769	0.574	0.549	0.515	0.513	0.499	0.522	0.605
	$[0.8, 1]$	0.351	0.375	0.439	0.475	0.563	0.630	0.784	1.407

APPENDIX C

Background Distributions for $D^0 \to \gamma\gamma$



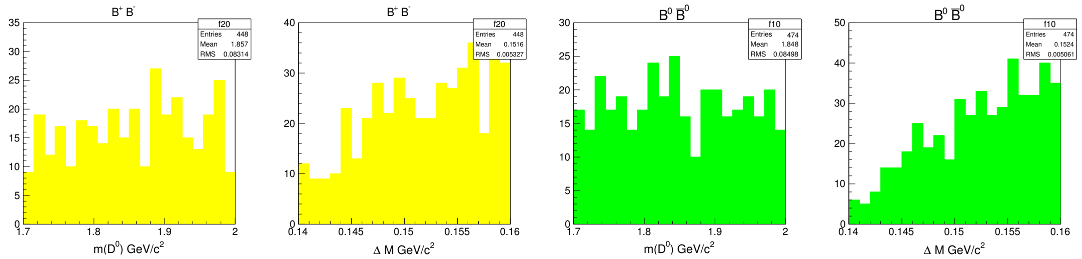


FIGURE C.1: $M(D^0)$ and ΔM distributions for various background contributions shown in Fig. 5.11.

APPENDIX D

Study of peaking background in $D^0 \to \gamma\gamma$ sample

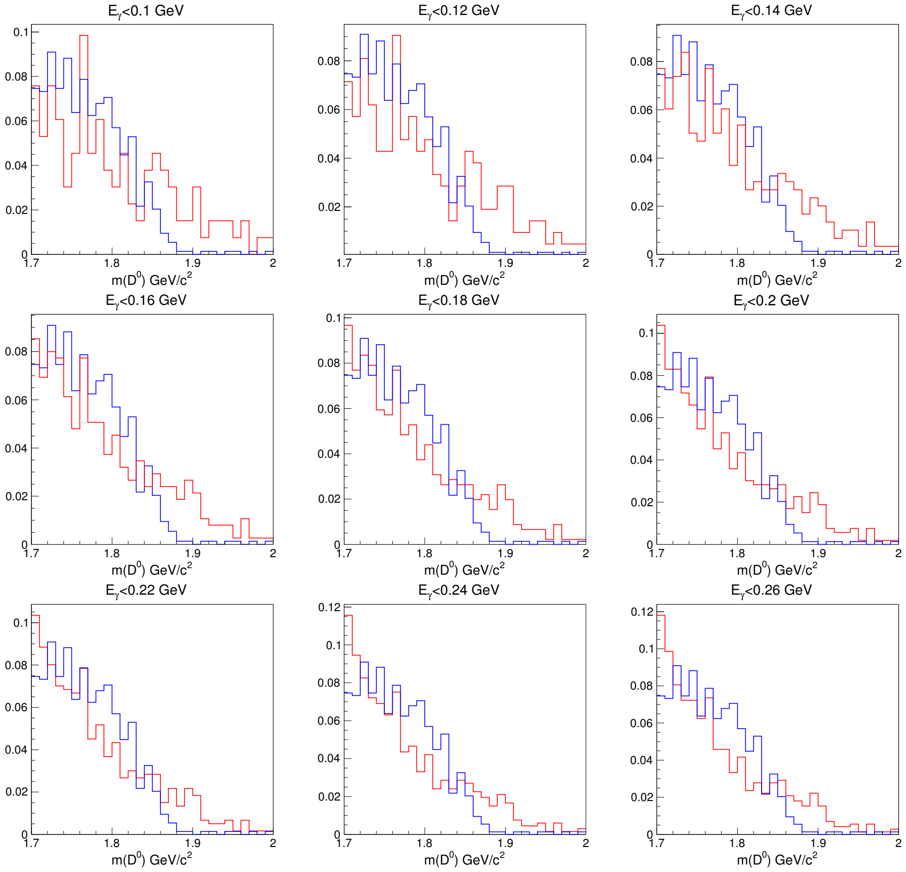


FIGURE D.1: Comparison of $D^0 \to \pi^0\pi^0$ component from $D^0 \to \gamma\gamma$ sample (blue) and partially reconstructed $D^0 \to \pi^0\pi^0$ sample (red) for different energy cuts on γ_{low}.

Bibliography

[1] N. Cabibb, Phys. Rev. Lett. **10**, 531 (1963).

[2] S. L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. D **2** (1970) 1285.

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

[4] S. L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev D. **2**, 1285 (1970).

[5] R. Aaij et al., (LHCb Collab.), Phys. Rev. Lett. **108**, 111602 (2012).

[6] M. Dorigo *et al.* (CDF Collab.), Phys. Rev. Lett. **109**, 111801 (2012).

[7] H-Y. Cheng and C-W. Chiang, Phys. Rev. D **85**, 034036 (2012), [arXiv:1201.0785].

[8] B. Bhattacharya, M. Gronau and J.L. Rosner, Phys. Rev. D **85**, 054014 (2012), [arXiv:1201.2351].

[9] G. Hiller, M. Jung and S. Schacht, Phys. Rev. D **87**, 014024 (2012), [arXiv:1211.3734].

[10] G. Burdman, E. Golowich, J. L. Hewett and S. Pakvasa, Phys. Rev. D **66** (2002) 014009 [hep-ph/0112235].

[11] S. Prelovsek and D. Wyler, Phys. Lett. B **500** (2001) 304

[12] A. Abashian et al. (Belle Collaboration), Nucl. Instrum. Methods Phys. Res, Sect. A **479**, 117 (2002).

[13] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res, Sect. A **499**, 1 (2003), and other papers included in this volume.

[14] H. Kichimi et al., JINST **5**, 03011 (2010).

[15] http://belle.kek.jp/belle/transparency/accelerator1.html.

[16] K. Ueno, S. K. Sahu, K. C. Peng, W. S. Hou and C. H. Wang, Nucl. Instrum. Meth. A **396** (1997) 103 [physics/9704013].

[17] G. Alimonti *et al.* [Belle Collaboration], Nucl. Instrum. Meth. A **453** (2000) 71.

[18] H. Hirano *et al.*, Nucl. Instrum. Meth. A **455** (2000) 294.

[19] H. Kichimi *et al.*, Nucl. Instrum. Meth. A **453** (2000) 315.

[20] K Miyabayashi, Nucl. Instrum. Meth. A **494** (2002) 298.

[21] N. Tan *et al.*, IEEE Trans. Nucl. Sci. **48** (2001) 900.

[22] M. Staric *et al.* [Belle Collaboration], Phys. Rev. Lett. **108** (2012) 071801 [arXiv:1110.0694 [hep-ex]].

[23] G. Bonvicini *et al.* (CLEO Collab.), Phys. Rev. D **63**, 071101 (2001).

[24] J.P. Lees (BaBar Collab.), Phys. Rev. D **85**, 091107 (2012), [arXiv:1110.6480].

[25] B. R. Ko *et al.* [Belle Collaboration], Phys. Rev. Lett. **106** (2011) 211801 [arXiv:1101.3365 [hep-ex]].

[26] K. A. Olive *et al.* (Particle Data Group), Chin. Phys. C **38** (2014) 090001.

[27] M. Staric *et al.* [Belle Collaboration], Phys. Lett. B **670** (2008) 190 [arXiv:0807.0148 [hep-ex]].

[28] Louis Lyons, *Statistics for Nuclear and Particle Physicists* (Cambridge University Press, 1989).

[29] T. E. Coan *et al.* [CLEO Collaboration], Phys. Rev. Lett. **90** (2003) 101801 [hep-ex/0212045].

[30] J. P. Lees *et al.* [BaBar Collaboration], Phys. Rev. D **85** (2012) 091107 [arXiv:1110.6480 [hep-ex]].

[31] M. Ablikim *et al.* [BESIII Collaboration], Phys. Rev. D **91** (2015) 11, 112015 [arXiv:1505.03087 [hep-ex]].

[32] P. Koppenburg *et al.* [Belle Collaboration], Phys. Rev. Lett. **93** (2004) 061803 [hep-ex/0403004].

[33] G. Punzi, eConf C **030908** (2003) MODT022 [arXiv:physics/0308063].

[34] Y. Liu *et al.* [Belle Collaboration], Phys. Rev. D **78** (2008) 011106 [arXiv:0805.3225 [hep-ex]].

[35] D. Dutta *et al.* [Belle Collaboration], Phys. Rev. D **91** (2015) 1, 011101 [arXiv:1411.7771 [hep-ex]].

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