Author/s:
Bloomfield, Tristan Joel

Title:
Measurement of Direct CP Asymmetry and Branching Fraction in B0→D0#0 and B+→D0#+ at the Belle Experiment

Date:
2019

Persistent Link:
https://hdl.handle.net/11343/234426

Measurement of Direct CP Asymmetry and Branching Fraction in $B^0 \to \overline{D^0}\pi^0$ and $B^+ \to \overline{D^0}\pi^+$ at the Belle Experiment

Tristan Bloomfield
https://orcid.org/0000-0001-9288-5069

Submitted in total fulfilment of the requirements of the degree of Doctor of Philosophy

September 2019

School of Physics
University of Melbourne

Abstract

This thesis describes the measurement of direct CP asymmetry and branching fraction for the hadronic B decays $B^0 \to \overline{D^0}\pi^0$ and $B^+ \to \overline{D^0}\pi^+$. The study uses the full dataset of 711 fb^{-1} collected at the $\Upsilon(4S)$ resonance by the Belle experiment at the KEKB accelerator in Tsukuba, Japan.

Event reconstruction, background suppression and modelling are first studied using Monte Carlo simulations, before yield and direct CP asymmetry are extracted in a three-dimensional unbinned extended maximum likelihood fit. $B^+ \to \overline{D^0}\pi^+$ is measured first as the control mode to validate the methodology, before same techniques are used on $B^0 \to \overline{D^0}\pi^0$. The measured branching fractions and direct CP asymmetries are:

$$\mathfrak{B}\left(B^0 \to \overline{D^0}\pi^0\right) = (2.69 \pm 0.06 \pm 0.09) \times 10^{-4},$$

$$A_{CP}\left(B^0 \to \overline{D^0}\pi^0\right) = (0.10 \pm 2.05 \pm 1.29) \times 10^{-2},$$

$$\mathfrak{B}\left(B^+ \to \overline{D^0}\pi^+\right) = (4.53 \pm 0.02 \pm 0.14) \times 10^{-3},$$

$$A_{CP}\left(B^+ \to \overline{D^0}\pi^+\right) = (0.19 \pm 0.36 \pm 0.60) \times 10^{-2},$$

for $B^0 \to \overline{D^0}\pi^0$ and $B^+ \to \overline{D^0}\pi^+$ respectively, where the first uncertainty is statistical and the second is systematic.

The represents the world's first measurement of direct CP asymmetry for $B^0 \to \overline{D^0}\pi^0$. This measurement of branching fraction of $B^0 \to \overline{D^0}\pi^0$ and $B^+ \to \overline{D^0}\pi^+$, and direct CP asymmetry of $B^+ \to \overline{D^0}\pi^+$ are the most precise to date, and consistent with the current world average values.

Declaration

This is to certify that:

- This thesis comprises only my original work towards the PhD except where indicated,
- Due acknowledgement has been made in the text to all other material used,
- The thesis is less than 100,000 words in length, exclusive of tables, bibliographies and appendices.

Tristan Bloomfield

Preface

Modern experimental particle physics is increasingly conducted as part of large international scientific collaborations. The cost, size and complexity of the apparatus required makes this sort of scientific work impossible in isolation. The outcomes on any individual piece of analysis would not be possible without the work of hundreds of collaborators in designing and operating the detector and accelerator, writing the software and algorithms for event reconstruction, studying the performance of these and other vital tasks.

The work presented in this thesis is detailed below

- Chapter 1 is an original overview of CP violation and the Standard Model, as well as the specific motivation and theoretical background for the decays studied.

- Chapter 2 provides an original summary of the Belle experiment and the KEKB accelerator

- Chapter 3-7 presents the author's original work on the measurement of branching fraction and direct CP asymmetry for $B^0 \to \overline{D^0}\pi^0$ and $B^+ \to \overline{D^0}\pi^+$ decays, conducted within the Belle collaboration under the supervision of A/Prof. Martin Sevior.

Unless stated otherwise, all information provided is the original work of the author.

Acknowledgements

First and foremost, I would like to thank my supervisor Martin Sevior for his guidance and support over a period that has probably been significantly longer than he bargained for when he took me on as a student. He is always happy to offer counsel and patient enough to put up knocking on his door every day for a month while write this thesis. Without both his expertise and optimism this work would not have progressed past me bashing my head against a wall of basf code back in my masters. I would also like to offer my gratitude to the other members of my supervisory panel, Andrew Melatos and Phillip Urquijo, whose advice, especially when things seemed to be heading off the rails, has helped me greatly.

Particle physics analysis work like this truly is a collaborative effort. Without the work of the hundreds of members of Belle and KEKB, over multiple decades, there would be no data and no tools with which to analyse it, so I offer my thanks to everyone who has been involved with the experiment. Thanks must go to all the Belle experts who have offered me advice on this analysis, especially HBD convener Gagan Mohanty, and my referees Minzu Wang, Alan Schwartz and Vitaly Vorobyev. Although not the topic of this thesis, I would also like to mention the Belle II computing team, specifically Hara-san, Ueda-san and Miyake-san. You made me feel incredibly welcome during my stay in KEK and always went out of your way to help. Working on Belle II computing was one of the most enjoyable parts of my PhD, and the skills I built have been useful both in my physics work and beyond.

I'd also like to thank my colleagues and friends in COEPP, firstly for all your physics advice, but most importantly for helping me keep a bit of SAN intact through hanging out and talking shit for hours. Shout outs to Eiasha, Millie, Brian, Laurence, Anton, David, Kim, Francis and anyone else I've shared an office with. Extra if you helped proofread this thesis. Special thanks to my fellow Belle HBD student and moe boats addict Chia-Ling. Without having someone to work through all the hurdles of physics analysis with I probably would have flipped a table and ragequit, and having a great friend in the office made it worth not becoming a hikikomori.

Last but certainly not least I would like to thank my family for their incredible support over the course of this PhD and my entire life. Without your encouragement and hard work to support me I would never have been able to start this degree let alone finish it.

This is hardly an exhaustive list of those I'd like to thank, so if I've left anyone out, my excuse is it's 4am and I still need to get up to do more proofreading tomorrow.

Contents

List of Figures

List of Tables

1 Introduction

When we look at the world around us, we see a universe that is made of matter. At first this does not seem to be in any way unusual, indeed if it were not the case we would not be here to observe the universe. But that a matter dominated universe seems natural could be seen as naïve anthropic bias. As a result of Lorentz invariance, in any relativistic quantum theory, particles come in particle-antiparticle pairs. Given the symmetries between particles and antiparticles (same mass and lifetime, but with opposite charges) and pair production ($\gamma\gamma \rightarrow e^+e^-, q\overline{q}, \dots$) of particle-antiparticle pairs, we expect to see equal amounts of matter and antimatter produced immediately after the big bang.

There are three possible interpretations of this: Matter-antimatter asymmetry is a local effect and there are large regions of antimatter in the universe; our previous expectation is wrong, and the universe began with more matter than antimatter; or the universe did start out symmetric and there are physical processes that produce this asymmetry over time. This first possibility has been largely discounted by cosmological observations. If there did exist regions on antimatter, the boundaries between them and matter regions would produce significant amounts annihilation events which we could observe as light [1].

To quantify this asymmetry we refer to baryon number. We assign a positive baryon number to quarks and the same magnitude but negative to the corresponding antiquark. The baryon asymmetry in the universe is defined as the difference in number of baryons, N_b, and antibaryons, $N_{\overline{b}}$, divided by their sum. As the amount of antibaryons observed in the universe today is negligible, and the main product of matter-antimatter annihilation are photons, this can be approximated as by the baryon to photon ratio,

$$\eta \equiv \left.\frac{N_b}{N_\gamma}\right|_{present} = \left.\frac{N_b - N_{\overline{b}}}{N_\gamma}\right|_{present} \sim 10^{-10}. \qquad (1.1)$$

These observations enable tests of models of the early universe [1]. Predictions from our current understanding of particle physics and big-bang cosmology models do give rise a non-zero η, but it is $\mathcal{O}(10^{-17})$, orders of magnitude smaller than is observed. There is still much we do not understand about how this asymmetry came to be.

In 1967 Andrei Sakharov proposed his famous three conditions for interactions to produce this baryon asymmetry [2]:

- Baryon number violation.

- C and CP-symmetry violation.
- Interactions taking place out of thermal equilibrium.

Interactions that violate baryon number are an obvious pre-requisite, but they are not sufficient. Both C-symmetry and CP-symmetry violation are required so that interactions resulting in an increase in baryon number are not counter-balanced by their conjugate interactions (that decrease baryon number by the same amount) occurring at the same rate. Lastly these interactions must occur out of equilibrium, or the reverse interaction will occur at the same rate cancelling out any asymmetry produced.

1.1 The Standard Model

The Standard Model (SM) of particle physics is a self-consistent theory that represents sciences' best effort to date in explaining the workings of the world at the most fundamental level. It casts all matter and interactions (excluding gravity which is not described by the SM) as particle fields in a quantum field theory (QFT) framework using the $SU(3) \times SU(2) \times U(1)$ gauge group. Particles are split into categories based on how they interact: 12 spin-$\frac{1}{2}$ fermions that make up all matter; 4 spin-1 Gauge bosons that mediate the fundamental forces; and 1 spin-0 scalar boson that gives mass to the particles. The fermions are further split into quarks and leptons, with quarks interacting via both the strong and electroweak forces, while leptons only interact via the electroweak force. These can be further be divided into generations (or families) based on the mass hierarchy of particles with the same charges. The properties of the fermions are summarised in Table 1.

The gauge bosons act as force carriers, being exchanged between particles to mediate interactions. Photons mediate the electromagnetic force, $W^{\pm}$ and Z^{0} weak interactions, and gluons (there are technically 8, carrying different colour/anti-colour charges) the strong force.

Lastly there is the scalar boson, the Higgs, that provides fundamental mass to particles through its couplings. The properties of the bosons are summarised in Table 2. Every particle p has a corresponding antiparticle $\overline{p}$ (note that some particles are their own antiparticle, such as the photon) which has the same properties but opposite quantum charges (such as electric charge).

Table 1: Properties of the quarks and leptons [3]

	Gen.	Name	Charge (e)	Mass (MeV)
Quarks	1st	Down (d)	$-\frac{1}{3}$	$4.67^{+0.48}_{-0.17}$
		Up (u)	$+\frac{2}{3}$	$2.16^{+0.49}_{-0.26}$
	2nd	Strange (s)	$-\frac{1}{3}$	93^{+11}_{-5}
		Charm (c)	$+\frac{2}{3}$	$1.27^{+0.02}_{-0.02} \times 10^3$
	3rd	Bottom (b)	$-\frac{1}{3}$	$4.18^{+0.03}_{-0.02} \times 10^3$
		Top (t)	$+\frac{2}{3}$	$172.9^{+0.4}_{-0.4} \times 10^3$
Leptons	1st	Electron (e)	-1	$0.511 \pm (3 \times 10^{-9})$
		Electron neutrino (ν_e)	0	$< 2.2 \times 10^{-6}$
	2nd	Muon (μ)	-1	$105.7 \pm (2 \times 10^{-6})$
		Muon neutrino (ν_μ)	0	< 0.19
	3rd	Tau (τ)	-1	$(1.78 \times 10^3) \pm 0.12$
		Tau neutrino (ν_τ)	0	< 18.2

Table 2: Properties of the bosons [3]

Name	Interaction	Charge (e)	Mass (MeV)
Photon ($\boldsymbol{\gamma}$)	Electromagnetic	0	0
Gluon ()	Strong	0	0
$\boldsymbol{W^{\pm}}$	Weak (charged)	± 1	$(80.4 \times 10^3) \pm 12$
$\boldsymbol{Z^0}$	Weak (neutral)	0	$(91.2 \times 10^3) \pm 2.1$
Higgs ($\boldsymbol{H^0}$)	Higgs field	0	$125.1 \pm 0.14 \times 10^3$

The Standard Model is probably the most rigorously tested theory in all of physics, and its predictions have been repeatedly substantiated by experiment. However, it is not a complete theory of fundamental physics. It is unable to account for gravitational force as described by general relativity, or the accelerating expansion rate of the universe. It does not include the phenomenon of neutrino oscillations that implies that neutrinos are not mass degenerate (the SM describes them as massless). And of particular relevance to this work, it cannot fully account for the baryon asymmetry observed in the universe.

1.2 CP-symmetry

The CP operation is a combination of the charge conjugation operator (C), which reverses all the charges of a particle, transforming it into its antiparticle, and the

parity operator (P), which reverses spatial direction such that $\vec{x} \overset{P}{\rightarrow} -\vec{x}$. For example, the CP operator acting on a left-handed electron will transform it into a right-handed positron,

$$e_L^- \overset{CP}{\rightarrow} e_R^+. \tag{1.2}$$

Handedness here refers to the helicity of a particle, that is the projection of a particles spin onto its momentum. A particle is said to be left-handed if it has negative helicity and right-handed for positive helicity.

Naively it seems that C violation should be enough to create a matter-antimatter asymmetry through its implication that

$$P \rightarrow f \neq \overline{P} \rightarrow \overline{f}. \tag{1.3}$$

But consider an interaction that violates C symmetry but is invariant under CP. Although it may not produce equal amounts of left-handed particles and left-handed antiparticles, the rate of production of left-handed particles and right-handed antiparticles will remain equal,

$$P \rightarrow f_L = \overline{P} \rightarrow \overline{f}_R. \tag{1.4}$$

Hence the requirement in the Sakharov conditions is that both C-symmetry and CP-symmetry must be violated.

For a long time both charge and parity were considered fundamental symmetries of physics. However in a 1956 review of experimental results by Lee and Yang [4] showed that while strong and electromagnetic interactions had been confirmed to be P-invariant, the same was not true in weak decays. Shortly after in 1957, parity violation was observed in the β-decay of cobalt-60 [5]. This violation of a seemingly fundamental symmetry was quite shocking; however it was realised that although P was violated in weak decays, the combination and C and P symmetries seemed to be conserved, and thus this must be the true symmetry between particles and antiparticles. This was the case until in 1964 the CP violating $K_L \rightarrow \pi^+\pi^-$ decay was observed by Cronin and Fitch [6]. With this discovery the symmetry between particles and anti-particles was thrown out, and a theoretical mechanism to explain this asymmetry was needed.

1.3 KM Mechanism

The mechanism by which CP violation is explained in the Standard Model was proposed by Kobayashi and Maskawa in 1973 [7]. Looking at the general form of a Standard Model Lagrangian,

$$\mathcal{L} = \sum_i (a_i \mathcal{O}_i(x) + h.c.)\,, \tag{1.5}$$

the interactions between particles are represented by the couplings, a_i, and the operators, $\mathcal{O}_i$, and the anti-particle interactions by their Hermitian conjugates ($h.c.$). From this we can see that for the equations that govern particle and anti-particle decays to differ, we must have at least one coupling with an irreducible complex phase. The gauge couplings between fermions and gauge bosons must be real, so these interactions must be CP invariant. In a world of massless particles without the Higgs boson, this would lead to the conclusion that CP should be a perfect symmetry (for the time being we are ignoring the strong sector $G^{\mu\nu,a}\tilde{G}^{a}_{\mu\nu}$ term).

When we introduce spontaneous symmetry breaking to give mass to particles through the Higgs field, we get the following Yukawa interactions between the quark fields and the Higgs:

$$\begin{aligned}\mathcal{L} = &-\sum_{i,j=1}^{n} (G_U)_{i,j}\left(\overline{U}_{i,L}, \overline{D}_{i,L}\right)\begin{pmatrix}\phi^0\\ -\phi^-\end{pmatrix} U_{j,R} \\ &-\sum_{i,j=1}^{n} (G_D)_{i,j}\left(\overline{U}_{i,L}, \overline{D}_{i,L}\right)\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} D_{j,R} + h.c.\end{aligned} \tag{1.6}$$

where n is the number of quark families, $G_{U[D]}$ is the up [down] Yukawa coupling, $\left(\overline{U}_L, \overline{D}_L\right)$ is the left handed quark doublet, U_R is the right-handed up singlet, D_R is the right-handed down singlet and ϕ are the components of the Higgs field. When the neutral Higgs acquires a vacuum expectation value, $\langle\phi^0\rangle = v$, the fermion mass matrices arise in the form:

$$\mathcal{M}_U = vG_U, \tag{1.7}$$

$$\mathcal{M}_D = vG_D, \tag{1.8}$$

Unlike gauge couplings, Yukawa couplings are in general complex and thus so are the mass matrices. As well as being an avenue for potential CP violation due to possible complex phases, the fact that the mass matrices are not diagonal means the flavour eigenstate and mass eigenstate of the quark fields are different.

To more clearly see the physical interpretation of this we look at how the electroweak Lagrangian transforms when we write the quark fields as mass eigenstates. The relevant gauge-quark components of Lagrangian are the charged current (CC),

$$\mathcal{L}_{CC} = \frac{g}{\sqrt{2}}\overline{U}_L \gamma_\mu D_L W^{-,\mu} + h.c. \tag{1.9}$$

which couples quarks to the W bosons, and the neutral current (NC),

$$\mathcal{L}_{NC} = \frac{g}{2\cos\theta_W}\left(\overline{U}_L\gamma_\mu U_L - \overline{D}_L\gamma_\mu D_L\right)Z^\mu, \tag{1.10}$$

coupling quarks and the Z boson. To see how quark field mass eigenstates relate to their flavour eigenstates, we use the singular value decomposition theorem that states any complex $n \times n$ matrix can be diagonalized by multiplying on either side by unitary matrices. To diagonalize the quark mass matrices,

$$\mathcal{M}_U^{diag} = T_{U,}\mathcal{M}_U S_U, \tag{1.11}$$

$$\mathcal{M}_D^{diag} = T_{D,}\mathcal{M}_D S_D. \tag{1.12}$$

The unitary matrices $T_{U[D]}$ transform the up [down] quark doublets from their flavour eigenstates to mass eigenstates as:

$$U_L^{mass} = T_U U_L, \tag{1.13}$$

$$D_L^{mass} = T_D D_L, \tag{1.14}$$

where $U[D]_L^{mass}$is the mass eigenstate of the left handed up [down] quark field. Applying this transforms the neutral current Lagrangian,

$$\begin{aligned}\mathcal{L}_{NC} &= \frac{g}{2\cos\theta_W}\left(\overline{U}_L\gamma_\mu U_L - \overline{D}_L\gamma_\mu D_L\right)Z^\mu \\ &= \frac{g}{2\cos\theta_W}\left(\overline{U}_L^{mass}T_U\gamma_\mu T_U^\dagger U_L^{mass} - \overline{D}_L^{mass}T_D\gamma_\mu T_D^\dagger D_L^{mass}\right)Z^\mu \\ &= \frac{g}{2\cos\theta_W}\left(\overline{U}_L^{mass}\gamma_\mu U_L^{mass} - \overline{D}_L^{mass}\gamma_\mu D_L^{mass}\right)Z^\mu,\end{aligned} \tag{1.15}$$

shows that it is unchanged and remains CP invariant. However, doing the same for the charged current,

$$\begin{aligned}\mathcal{L}_{CC} &= \frac{g}{\sqrt{2}}\overline{U}_L\gamma_\mu D_L W^{-,\mu} + h.c. \\ &= \frac{g}{\sqrt{2}}\overline{U}_L^{mass}T_U\gamma_\mu T_D^\dagger D_L^{mass}W^{-,\mu} + h.c. \\ &= \frac{g}{\sqrt{2}}\overline{U}_L^{mass}\gamma_\mu V_{CKM} D_L^{mass}W^{-,\mu} + h.c.\end{aligned} \tag{1.16}$$

shows the quark mixing matrix defined by

$$V_{CKM} = T_U T_D^\dagger. \tag{1.17}$$

This is the famous Cabibbo-Kobayashi-Maskawa (CKM) matrix [7] that leads to flavour changing in weak interactions and is a complex unitary matrix. However just because it is a complex matrix, does not mean the phases are necessarily irreducible.

From the unitarity condition $\sum_j V_{ij} V_{jk}^* = \delta_{ik}$, a general complex unitary $n \times n$ matrix has n^2 components, $\frac{1}{2}n(n-1)$ moduli and $\frac{1}{2}n(n+1)$ phases. We are free to arbitrarily rotate the quark phases as

$$U_i^{mass} = e^{i\phi_i} U_i^{mass}, \tag{1.18}$$

$$D_i^{mass} = e^{i\theta_i} D_i^{mass}, \tag{1.19}$$

which leads to V_{CKM} transforming as:

$$V_{CKM} \rightarrow \begin{bmatrix} e^{i\phi_i} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{i\phi_n} \end{bmatrix} V_{CKM} \begin{bmatrix} e^{i\theta_i} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{i\theta_n} \end{bmatrix}, \tag{1.20}$$

allowing for $2n-1$ phases for be rotated out. This leaves us with the total number of irreducible phases for an n family CKM matrix as $\frac{1}{2}(n-1)(n-2)$. When Kobayashi and Maskawa first published this in their 1973 paper [7] (for which they were later awarded the Nobel prize) only 3 quarks has been observed (although the existence of the charm quark had been predicted previously). However as can be seen, for only two families of quarks ($n = 2$) there is no complex phase to allow CP violation in weak decays, only a single mixing angle (the Cabbibo angle, θ_c). Thus, from this it was predicted there must be at least one more quark family to explain the observation of CP violation. This gives 3 mixing angles and a single irreducible phase. Four years later in 1977 the bottom quark was experimentally observed [8], but it wasn't until 1995 that the top quark was discovered at Fermilab [9, 10].

Combining this mechanism in the SM with models of big bang cosmology, the baryon-antibaryon asymmetry can be predicted as being $\mathcal{O}(10^{-17})$. This compares with observations of $\eta = \mathcal{O}(10^{-10})$. This large discrepancy tells us that while the SM theory can explain some CP violation, it cannot be the only source, and there must be fundamental physics we have yet to understand that has caused the universe to develop the way it has.

A brief note must be made on the $G^{\mu\nu,a}\tilde{G}^a_{\mu\nu}$ term in the strong interaction Lagrangian mentioned at the start of this section. This is a Gauge-invariant term of dimension-four, so should be included in our Standard Model Lagrangian as,

$$\mathcal{L}_{CP\ Strong} = \theta \frac{\alpha_s}{8\pi} G^{\mu\nu,a} \tilde{G}^a_{\mu\nu}. \tag{1.21}$$

This term violates both P and CP. However, constraints from measurements on the electric dipole moment of neutrons require $\theta \leq 10^{-10}$. This unnaturally small value for θ is known as the "Strong CP Problem" and is an unsolved problem in physics. For the rest of this work we focus only on the established theory of CP violation from the CKM matrix.

1.4 CKM Matrix

There are 3 families of quarks in the Standard Model (the minimum required to allow CP violation in weak interactions), leading to the CKM matrix being a 3×3 unitary matrix. This matrix relates the down-type part of the weak interaction quark doublet to the mass eigenstate of the down-type quarks,

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{bmatrix} d \\ s \\ b \end{bmatrix}, \tag{1.22}$$

with the magnitude of the element $|V_{ij}|$ being the relative probability of an up-type quark $i \in \{u, c, t\}$ coupling to a down-type quark $j \in \{d, s, b\}$ through a weak charged current interaction.

There is a multitude of ways in which the CKM matrix can be parameterised. The standard form adopted by the Particle Data Group (PDG) [3, 11] uses the three Euler angles in family space, θ_{12}, θ_{23}, θ_{13}, which describe the mixing between generations, and the complex phase δ_{13} that gives rise to CP violation.

$$V_{CKM} = \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{12}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{12}e^{i\delta_{13}} & c_{23}c_{13} \end{bmatrix}, \tag{1.23}$$

where $s_{ij} \equiv \sin\ \theta_{ij}$ and $c_{ij} \equiv \cos\theta_{ij}$. With this geometric representation the Euler angles define the size of the mixing between families, and from experiment we get a clear hierarchy of $1 \gg \sin\theta_{12} \gg \sin\theta_{23} \gg \sin\theta_{13}$. However, it is perhaps not the most elucidating form to see how this hierarchy reflects in the individual matrix elements. The Wolfenstein parametrisation [12] expresses the CKM matrix as a Taylor expansion of $\lambda \equiv \sin\theta_{12}$,

$$V_{CKM} = \begin{bmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{bmatrix} + \mathcal{O}(\lambda^4), \tag{1.24}$$

where A, ρ and η are all real numbers of order unity, defined as

$$\sin\theta_{23} = A\lambda^2 \tag{1.25}$$

$$\sin\theta_{13}e^{i\delta_{13}} = A\lambda^3(\rho + i\eta) \tag{1.26}$$

This form allows us to see the clear hierarchy of elements in orders of λ,

$$\underbrace{|V_{ud}| \approx |V_{cs}| \approx |V_{tb}|}_{\mathcal{O}(1)} \ll \underbrace{|V_{us}| \approx |V_{cd}|}_{\mathcal{O}(\lambda)} \ll \underbrace{|V_{cb}| \approx |V_{ts}|}_{\mathcal{O}(\lambda^2)} \ll \underbrace{|V_{ub}| \approx |V_{td}|}_{\mathcal{O}(\lambda^3)}. \tag{1.27}$$

It is noted that this form of the Wolfenstein parametrisation is only unitary up to $\mathcal{O}(\lambda^4)$. In order to satisfy unitarity conditions to all orders, the modified parameters $\overline{\rho}, \overline{\eta}$ are used,

$$\overline{\rho} + i\overline{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}, \tag{1.28}$$

which relate to the original parameters by

$$(\rho + i\eta) = (\overline{\rho} + i\overline{\eta})\frac{\sqrt{1 - A^2\lambda^4}}{\sqrt{1-\lambda^2}\left(1 - A^2\lambda^4(\overline{\rho} + i\overline{\eta})\right)}. \tag{1.29}$$

These differences only show at higher orders ($> \mathcal{O}(\lambda^4)$) so ρ, η and $\overline{\rho}, \overline{\eta}$ are often used interchangeably.

1.4.1 CKM Unitarity

The unitary condition $\sum_j V_{ij}\, V_{jk}^* = \delta_{ik}$ can be split into two conditions that inform us of different physical concepts.

$$\sum_i^3 \left|V_{ij}\right|^2 = 1, \qquad \forall j \in \{1,2,3\} \tag{1.30}$$

expresses the fundamental physical concept of weak universality, that the total size of the charged weak couplings of all the quarks is universal. For example, the sum of the couplings between an up-type quark u_i ($i \in \{u, s, t\}$) to all the down-type quarks is the same for all i.

The other part of the unitarity condition,

$$\sum_i^3 V_{ij}V_{ik}^* = \sum_i^3 V_{ji}V_{ki}^* = 0, \qquad \forall i, j \in \{1,2,3\} \tag{1.31}$$

can be expressed geometrically as six triangles in the complex plane, with side lengths related by the CKM matrix elements. The areas of these triangles are equal, and proportional to the size of CP violation in the CKM matrix. Of these triangles, two have sides all of the same order ($\mathcal{O}(\lambda^3)$),

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.32}$$

$$V_{td}V_{ud}^* + V_{ts}V_{us}^* + V_{tb}V_{ub}^* = 0. \tag{1.33}$$

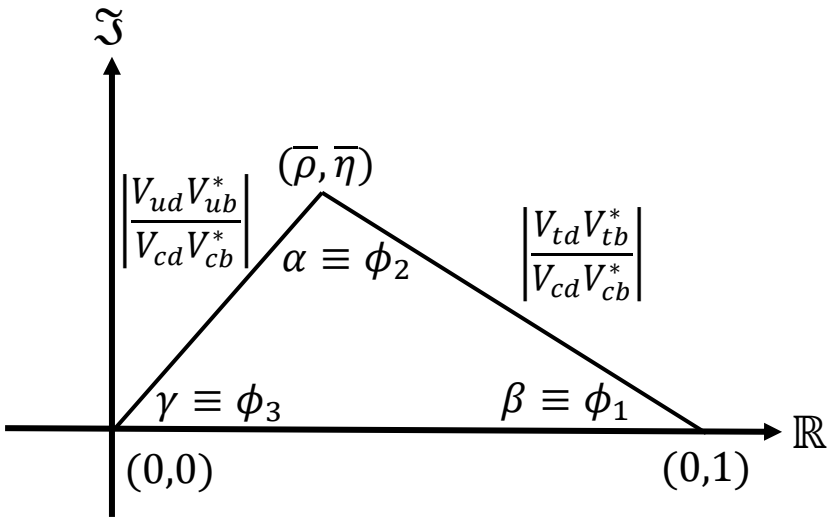


Figure 1: The normalised unitarity triangle in the complex plane.

Equation 1.32 is commonly used due to it relating the couplings present in B meson decays. In the Wolfenstein parametrisation, V_{cd} and V_{cb} lie on the real axis of the complex plane, so dividing by $V_{cd}V_{cb}^*$ normalises the equation so that the vertices of the triangle are located at (0,0), (0,1) and $(\overline{\rho},\overline{\eta})$, as shown in Figure 1. The angles of this triangle are therefore given by :

$$\beta \equiv \phi_1 \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \tag{1.34}$$

$$\alpha \equiv \phi_2 \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \tag{1.35}$$

$$\gamma \equiv \phi_3 \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right), \tag{1.36}$$

relating the relative phases of the matrix elements. By measuring these angles and sides lengths experimentally, the consistency of these constraints and the Standard Model theory from which they arise can be tested. The CKM-fitter group [13] perform global fits combining the results of numerous independent studies in various decay systems. The state of experimental constraints on these parameters (as of 2018) is summarised in Figure 2.

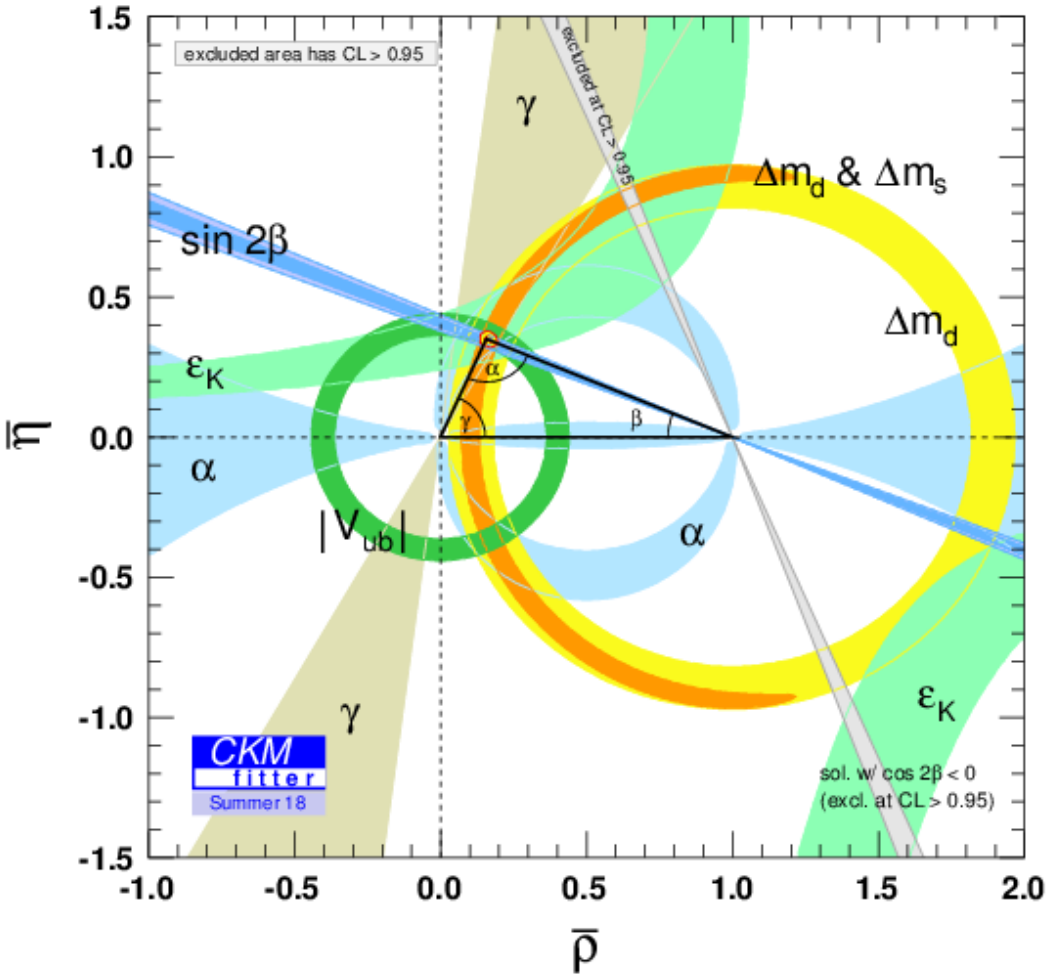


Figure 2: Current constraints on the parameters of the unitarity triangle from the global fit by the CKM-Fitter as of ICHEP 2018 [13].

1.5 Observing CP violation in Meson Systems

We quantify the amount of CP violation in a decay $P \to f$ by

$$A_{CP}(P \to f) \equiv \frac{\Gamma(P \to f) - \Gamma(\overline{P} \to \overline{f})}{\Gamma(P \to f) + \Gamma(\overline{P} \to \overline{f})}, \tag{1.37}$$

where $\Gamma(P \to f)$ is the partial width of the decay. Mesons are bound systems of a quark and antiquark pair. To see how CP violation can present itself as an experimentally observable quantity we adopt the formalism used in Bigi and Sanda [14] and look at a generic neutral meson P^0. From CPT-symmetry P^0 and $\overline{P^0}$ are mass degenerate, and can decay into each other weakly (as shown for the B^0 meson in Figure 3).

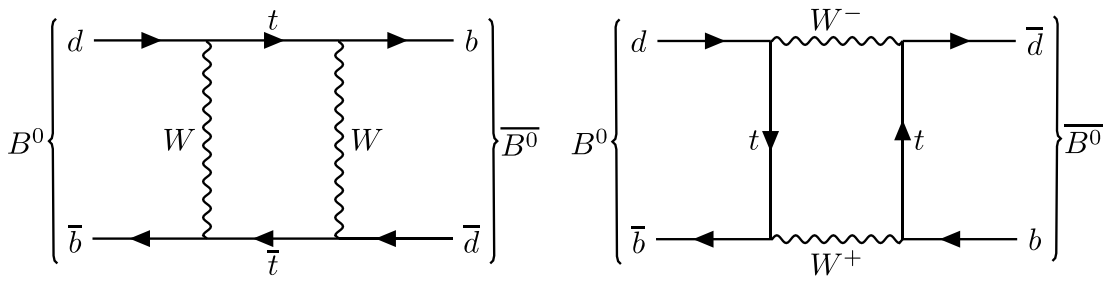


Figure 3: Feynman "Box" diagrams showing the process of $\boldsymbol{B^0 \to \overline{B^0}}$ oscillation.

We start with the time evolution Schrödinger equation for the P^0 and $\overline{P^0}$ system $|\Psi(t)\rangle = a(t)|P^0\rangle + b(t)|\overline{P^0}\rangle$ (restricted to P^0, $\overline{P^0}$ space):

$$i\frac{\partial}{\partial t}\Psi(t) = \mathcal{H}\Psi(t), \tag{1.38}$$

with the Hamiltonian

$$\mathcal{H} = M - \frac{i}{2}\Gamma, \tag{1.39}$$

where the mass matrix, M, and decay matrix, Γ, are both 2×2 Hermitian matrices. Solving the Schrödinger equation by diagonalizing the Hamiltonian we find the mass eigenstates:

$$|P_\pm\rangle = p|P^0\rangle \pm q|\overline{P^0}\rangle, \tag{1.40}$$

and corresponding eigenvalues:

$$\lambda_\pm = M_{11} - \frac{i}{2}\Gamma_{11} \pm \frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right), \tag{1.41}$$

where

$$\frac{q^2}{p^2} = \frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}. \tag{1.42}$$

These eigenvalues tell us the mass, $M_\pm$, and lifetime, $\Gamma_\pm$, of the eigenstates from the M and Γ contribution respectively,

$$M_\pm = \Re(\lambda_\pm), \tag{1.43}$$

$$\Gamma_\pm = 2\Im(\lambda_\pm). \tag{1.44}$$

Looking again at the Schrödinger equation for these eigenstates,

$$i\frac{\partial}{\partial t}|P_\pm\rangle = \lambda_\pm|P_\pm\rangle, \tag{1.45}$$

shows that their time evolution is given by

$$\left|P_{\pm}(t)\right\rangle = e^{i\lambda_{\pm}t}(p|P^0\rangle \pm q|\overline{P^0}\rangle\,). \tag{1.46}$$

Combining Equation 1.40 and Equation 1.46 we can find the time evolution of the flavour eigenstates:

$$|P^0(t)\rangle = f_+(t)|P^0\rangle + f_-(t)\frac{q}{p}|\overline{P^0}\rangle\,, \tag{1.47}$$

$$\left|\overline{P^0}(t)\right\rangle = f_+(t)|\overline{P^0}\rangle\ + f_-(t)\frac{p}{q}\ |P^0\rangle, \tag{1.48}$$

where

$$f_{\pm}(t) = \frac{1}{2}e^{-iM_\mu t}e^{-\frac{1}{2}\Gamma_\mu t}\left(e^{-\frac{i}{2}\Delta M t}e^{-\frac{1}{4}\Delta\Gamma t} \pm e^{\frac{i}{2}\Delta M t}e^{\frac{1}{4}\Delta\Gamma t}\right), \tag{1.49}$$

and $M_\mu \equiv \frac{M_+ + M_-}{2}$, $\Delta M \equiv M_+ - M_-$, $\Gamma_\mu \equiv \frac{\Gamma_+ + \Gamma_-}{2}$ and $\Delta\Gamma \equiv \Gamma_+ - \Gamma_-$.

The decay amplitudes for the processes P^0 and $\overline{P^0}$ into some final state f are given by

$$A(f) = \langle f|H_{\boldsymbol{W}}|P^0\rangle, \tag{1.50}$$

$$\overline{A}(f) = \langle f|H_W|\overline{P^0}\rangle, \tag{1.51}$$

respectively, where H_W is the weak Hamiltonian. The decay rates are thus

$$\begin{aligned}\Gamma(P^0(t)\to f) \propto |A(f)|^2 e^{-\Gamma_\mu t}\Bigg[&\left(1+\left|\lambda_f\right|^2\right)\cosh\frac{\Delta\Gamma}{2}t - \Re\left(\lambda_f\right)\sinh\frac{\Delta\Gamma}{2}t \\ &+\frac{1}{2}\left(1-\left|\lambda_f\right|^2\right)\cos\Delta M t - 2\Im\left(\lambda_f\right)\sin\Delta M t\Bigg],\end{aligned} \tag{1.52}$$

$$\begin{aligned}\Gamma\left(\overline{\mathrm{P}^0}(t)\to f\right) \propto \left|\overline{A}(f)\right|^2 e^{-\Gamma_\mu t}\Bigg[&\left(1+\frac{1}{\left|\lambda_f\right|^2}\right)\cosh\frac{\Delta\Gamma}{2}t - \Re\left(\frac{1}{\lambda_f}\right)\sinh\frac{\Delta\Gamma}{2}t \\ &+\frac{1}{2}\left(1-\frac{1}{\left|\lambda_f\right|^2}\right)\cos\Delta M t - 2\Im\left(\frac{1}{\lambda_f}\right)\sin\Delta M t\Bigg],\end{aligned} \tag{1.53}$$

where $\lambda_f = \frac{q}{p}\frac{\overline{A}(f)}{A(f)}$. As can be seen, the complete form of the time evolution of the decay are rather complex in their most general form. For clarity we classify CP violation into 3 cases [3, 14]: Direct CP violation, indirect CP violation and time dependent CP violation.

1.5.1 Direct CP Violation

Direct CP violation occurs when the amplitudes of a decay and its conjugate have different magnitudes,

$$|A(f)| \neq \left|\overline{A}\left(\overline{f}\right)\right|, \tag{1.54}$$

which results in CP violation without any requirement of mixing. This makes it the only CP violations observable in charged meson decays (without violating charge conservation). If mixing is not present, $\Delta M = \Delta\Gamma = 0$. The decay rates are simply proportional to the square of the decay amplitude and

$$A_{CP} = \frac{|A(P \to f)|^2 - \left|A\left(\overline{P} \to \overline{f}\right)\right|^2}{|A(P \to f)|^2 + \left|A\left(\overline{P} \to \overline{f}\right)\right|^2}, \tag{1.55}$$

with no time dependence. The conditions to achieve the inequality $|A(f)| \neq \left|\overline{A}\left(\overline{f}\right)\right|$ can be demonstrated by considering the most general form of the amplitude $A(f)$ as the sum of processes separable into a magnitude A_i, a CP violating phase ϕ_i, and a CP invariant phase δ_i,

$$A(f) = \sum_i A_i e^{i(\phi+\delta)}, \tag{1.56}$$

$$\overline{A}\left(\overline{f}\right) = \sum_i A_i e^{i(-\phi+\delta)}. \tag{1.57}$$

It is immediately obvious that in the case of a decay with only a single mechanism to induce the transition, no CP violation is possible. In the case of two processes however we find that

$$A_{CP} = \frac{2A_1A_2 \sin\Delta\delta \sin\Delta\phi}{A_1^2 + A_2^2 + 2A_1A_2 \cos\Delta\delta \sin\Delta\phi}, \tag{1.58}$$

where $\Delta\delta \equiv \delta_1 - \delta_2$ and $\Delta\phi \equiv \phi_1 - \phi_2$. For direct CP violation to occur there needs to exist a phase shift in both the CP violating and CP invariant phases between the two processes. As a note, the CP violating phase is generally called the weak phase, as it originates from weak interactions (the only source of CP violation in the SM Lagrangian as shown in chapter 1.3). The CP invariant phase is known as the strong phase as it is typically the result of final state interactions (FSIs) most heavily influenced by the strong force. Looking only at the magnitudes in Equation 1.58,

$$A_{CP} \propto \frac{2A_1A_2}{A_1 + A_2 + 2A_1A_2} = \frac{2\rho}{\rho^2 + 2\rho + 1}, \tag{1.59}$$

where $\rho = \frac{A_1}{A_2}$, shows A_{CP} is strictly increasing with $\rho \in [0,1]$. This means that in order to measure large CP violation, as well as both $\Delta\delta$ and $\Delta\phi$ approaching $\frac{\pi}{2}$, we

want the amplitudes of the two dominant processes with phase differences to be as close in magnitude as possible.

1.5.2 Indirect CP violation

To observe CP violation in neutral meson mixing, we look at flavour specific decays, where it is only possible for one flavour of neutral meson to decay into the final state

$$P^0 \to f \nleftarrow \overline{P^0}, \qquad P^0 \nrightarrow \overline{f} \leftarrow \overline{P^0}, \tag{1.60}$$

meaning

$$A(\overline{f}) = \overline{A}(f) = 0. \tag{1.61}$$

Substituting Equation 1.61 in Equation 1.52 and 1.53 gives

$$\Gamma(P^0(t) \to f) \propto e^{\Gamma_\mu t}|A(f)|^2 \left(\cosh\frac{\Delta\Gamma}{2} + \cos\Delta Mt\right), \tag{1.62}$$

$$\Gamma\left(P^0(t) \to \overline{f}\right) \propto e^{\Gamma_\mu t}\left|\frac{q}{p}\right|^2 \left|\overline{A}(\overline{f})\right|^2 \left(\cosh\frac{\Delta\Gamma}{2} - \cos\Delta Mt\right), \tag{1.63}$$

$$\Gamma\left(\overline{P^0}(t) \to f\right) \propto e^{\Gamma_\mu t}\left|\frac{p}{q}\right|^2 |A(f)|^2 \left(\cosh\frac{\Delta\Gamma}{2} - \cos\Delta Mt\right), \tag{1.64}$$

$$\Gamma\left(\overline{P^0}(t) \to \overline{f}\right) \propto e^{\Gamma_\mu t}\left|\overline{A}(\overline{f})\right|^2 \left(\cosh\frac{\Delta\Gamma}{2} + \cos\Delta Mt\right). \tag{1.65}$$

For an inclusive final state, such as semi-leptonic $\ell^- + X$ where X is all hadronic states, CPT invariance implies $A(f) = \overline{A}(\overline{f})$ so $\Gamma(P^0(t) \to f) = \Gamma\left(\overline{P^0}(t) \to \overline{f}\right)$ and the CP violation thus becomes

$$A_{CP} = \frac{\Gamma\left(P^0(t) \to \overline{f}\right) - \Gamma\left(\overline{P^0}(t) \to f\right)}{\Gamma\left(P^0(t) \to \overline{f}\right) + \Gamma\left(\overline{P^0}(t) \to f\right)} = \frac{1 - \left|\frac{p}{q}\right|^4}{1 + \left|\frac{p}{q}\right|^4}. \tag{1.66}$$

Here if $\left|\frac{p}{q}\right| \neq 1$ we have CP violation caused from the mixing between P^0 and $\overline{P^0}$ independent of the final state. It is perhaps interesting to note that the CP violation caused by time dependent oscillations is not itself time-dependent.

1.5.3 Time Dependent CP violation

Even if $|A(f)| = \left|\overline{A}(\overline{f})\right|$ and $\left|\frac{p}{q}\right| = 1$ there can still be CP violation in decays where the final state can be reached by either flavour of neutral particle, $P^0 \to f \leftarrow \overline{P^0}$. This CP violation is a result of interference between decays where mixing occurs,

$P^0 \to \overline{P^0} \to f$ and those where it doesn't, $P^0 \to f$. In general this will involve the full form of the time evolution from Equation 1.52 and 1.53, so we focus on the special case of B meson decays to a CP eigenstate. For the neutral B_d meson, current world average measurements [3] give

$$\frac{\Delta\Gamma}{\Gamma} = 0.001 \pm 0.010, \tag{1.67}$$

$$\left|\frac{p}{q}\right| = 1.001 \pm 0.008, \tag{1.68}$$

so we can approximate no indirect CP violation and no difference in lifetimes, $\left|\frac{p}{q}\right| = 1$, $\Delta\Gamma = 0$. If the final state is a CP eigenstate we also have $f = \overline{f}$. Our time evolution equations become

$$\begin{aligned}\Gamma(P^0(t) \to f) \propto |A(f)|^2 e^{-\Gamma_\mu t} \Big(&\big(1 + \rho_f\big) \\ &+ \frac{1}{2}\big(1 - \rho_f\big)\cos\Delta M t - 2\Im\big(\lambda_f\big)\sin\Delta M t\Big),\end{aligned} \tag{1.69}$$

$$\begin{aligned}\Gamma\left(\overline{P^0}(t) \to f\right) \propto \left|\overline{A}(f)\right|^2 e^{-\Gamma_\mu t} \Bigg(&\left(1 + \frac{1}{\rho_f}\right) \\ &+ \frac{1}{2}\left(1 - \frac{1}{\rho_f}\right)\cos\Delta M t - 2\Im\left(\frac{1}{\lambda_f}\right)\sin\Delta M t\Bigg),\end{aligned} \tag{1.70}$$

where $\rho_f = \frac{\left|\overline{A}(f)\right|^2}{|A(f)|^2}$. CP violation is therefore given by

$$A_{CP}(t) = \frac{\Gamma(P^0(t) \to f) - \Gamma\left(\overline{P^0}(t) \to f\right)}{\Gamma(P^0(t) \to f) + \Gamma\left(\overline{P^0}(t) \to f\right)} \tag{1.71}$$

$$= \frac{1 - \rho_f}{1 + \rho_f}\cos\Delta M t - \frac{2\Im\big(\lambda_f\big)}{1 + \rho_f}\sin\Delta M t\,, \tag{1.72}$$

which is dependent on time. This form splits the CP violation into two observables as the coefficients of cos and sin, where the former is a result of direct CP violation. Even in the case of no direct CP violation, $\rho_f = 1$, there is still the possibility of CP violation if the imaginary component of λ_f is non-vanishing,

$$A_{CP}(t) = \Im\big(\lambda_f\big)\sin\Delta M t\,. \tag{1.73}$$

This shows we can observe CP violation in systems without CP violating initial state mixing or a direct difference in decay amplitudes.

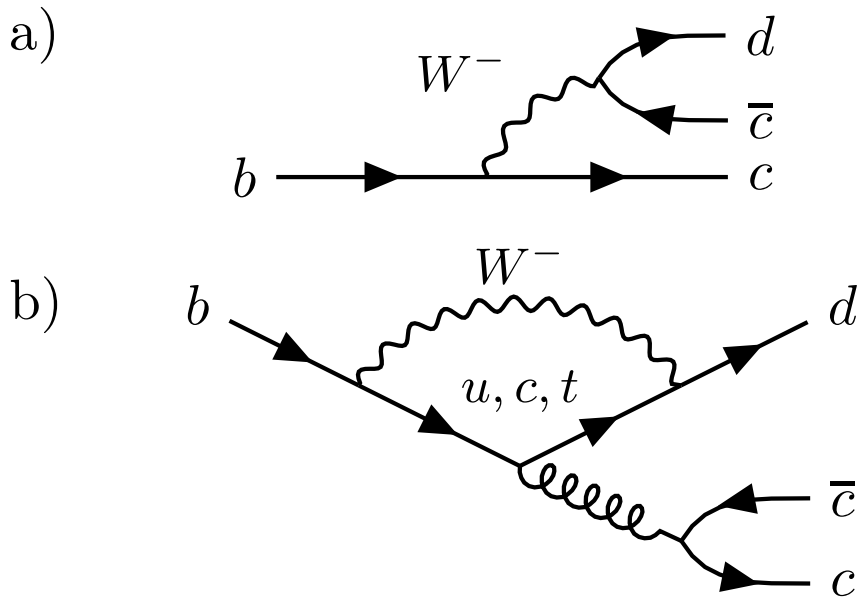


Figure 4: a) Tree diagram and b) Penguin diagram, for the $\boldsymbol{b \to c\bar{c}d}$ process.

1.6 Hadronic B Decays

Two-quark bound states which contain a $\overline{b}$ anti-quark and either a u, d, c or s quark are referred to as B mesons (B^+, B^0, B_c^+, B_s^o respectively), with $\overline{B}$ mesons ($B^-, \overline{B^0}, B_c^-, \overline{B_s^0}$) as their CP conjugates. These typically undergo two body hadronic decays through $\overline{b} \to \overline{q}q\overline{q}`$ $\left(q, \in \{u, d, s, c\};\ \overline{q} \in \{\overline{u}, \overline{d}, \overline{s}, \overline{c}\};\ \overline{q}` \in \{\overline{d}, \overline{s}\}\right)$ quark decays, with the non-b quark in the B meson acting as a spectator. Decays that proceed through a $b \to c$ quark transition are known as charmed B decays and are the most common, as this gives a b decay vertex with CKM element V_{cb}^* which is considerably larger (by a factor of λ) than V_{ub}^* element in charmless $b \to u$ decays.

As discussed in chapter 1.5.1, for direct CP violation to occur a decay needs contributions from multiple coherent processes that possess a difference in both the weak and strong phase. The strong phase difference of a process cannot be easily calculated as it is produced by final state interactions, but we can still estimate the potential for CP violation by comparing the Feynman amplitudes of the leading order process and that of the secondary process (the next largest that contains a different weak phase). There are two primary classes of diagrams that are involved in B meson decays: Tree and Penguin. Tree diagrams (example in Figure 4a) are those that contain no loops, the W boson emitted by the b quark simply decays to a quark anti-quark pair.

Penguin diagrams (example in Figure 4b) are those in which the emitted W boson is reabsorbed by the same quark line creating a flavour changing neutral current (FCNC). The quark side of the loop emits either a gluon (a gluonic penguin), Z boson or photon (an electroweak penguin). Relative to tree diagrams, penguin diagrams are suppressed by an additional mode dependant loop factor of $\mathcal{O}(0.2 - 0.3)$ [3]. It should be noted that penguin diagrams are only possible where the final state contains a quark anti-quark pair of the same flavour, as the gluon/photon/Z interaction cannot change flavour, and tree level FCNCs do not exist in the standard model. As the loop can contain any up-type quark (u, c, t), a penguin diagram like shown in Figure 4b is actually 3 distinct diagrams with different weak factors. Separating the CKM element factors from the rest of the diagram we have:

$$A_P(b \to q\bar{q}q') = V_{tb}V_{tq`}^{*}P_t + V_{cb}V_{cq`}^{*}P_c + V_{ub}V_{uq`}^{*}P_u\,, \qquad q` \in \{d, s\} \qquad (1.74)$$

where P_i is the non-weak component of the amplitude for a penguin with loop of flavour $i \in \{u, c, t\}$. Utilising the unitary relation (Equation 1.32) this can be reduced to a form containing two CKM factors such as:

$$A_P(b \to q\bar{q}q') = V_{tb}V_{tq`}^{*}(P_t - P_c) + V_{ub}V_{uq`}^{*}(P_u - P_c)\,, \qquad q` \in \{d, s\} \qquad (1.75)$$

where one of the CKM factors is the same as that of the tree diagram (if it exists) and another is different. For the purpose of studying CP violation the tree term and penguin term with the same weak factor can be combined in a "tree dominated component" as the leading term, leaving the "pure penguin component" with different weak phase as the secondary term.

1.7 $B^0 \to \overline{D^0}\pi^0$ Decays

The $B^0 \to \overline{D^0}\pi^0$ decay proceeds through the $b \to c\bar{u}d$ quark process. The tree level diagram, shown in Figure 5a, has the CKM component of its amplitude proportional to λ^2 arising from the $V_{cb}V_{ud}^{*}$ vertex elements. Because the final state does not include a same-flavoured quark anti-quark pair, no penguin diagram exists as a secondary component with differing weak phase.

Although for this study we look the D^0 decaying to a non-CP-eigen state (where the final state $f \neq \bar{f}$), the final states ($f = K^+\pi^-$ and $f = K^+\pi^-\pi^0$) can still be reached by both D^0 and $\overline{D^0}$ decays. Because of this, a secondary process exists in the form of the doubly Cabibbo suppressed (the CKM element of the amplitude is suppressed by a factor of λ^2 relative to the favoured process) diagram shown in Figure 5b. This contains CKM vertex elements $V_{cd}^{*}V_{ub}$ giving the necessary

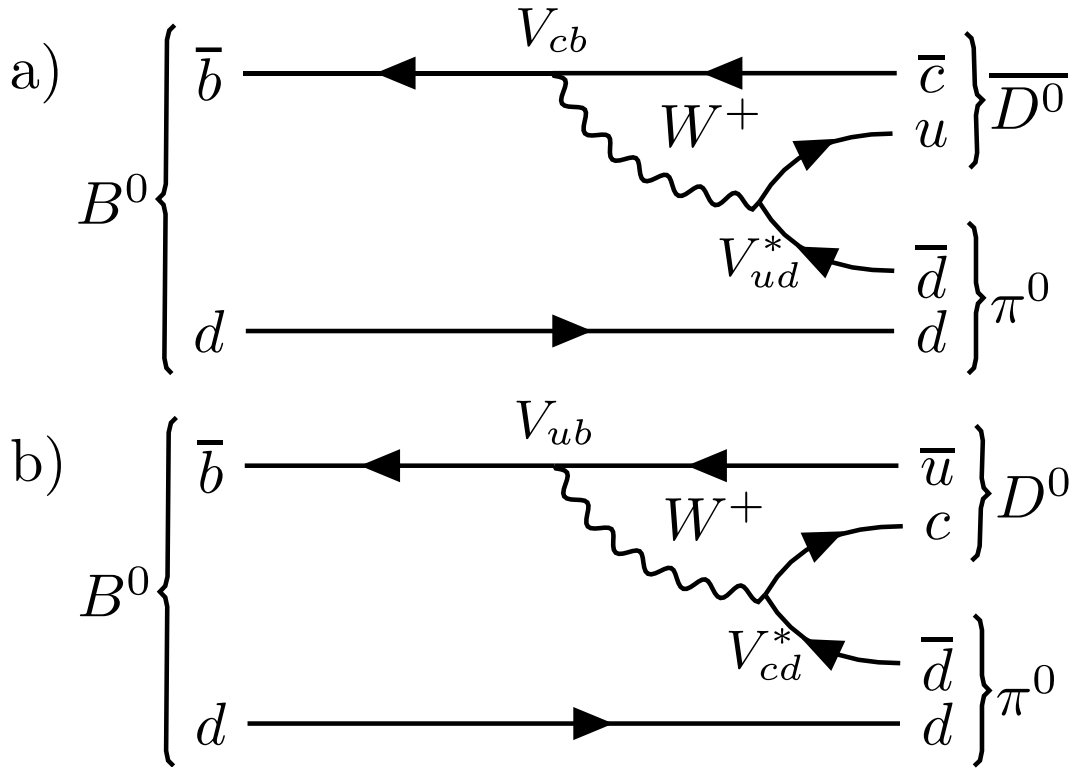


Figure 5: Feynman diagrams for a) the Cabibbo-favoured $\boldsymbol{B^0 \to \overline{D^0}\pi^0}$ decay; and b) the doubly Cabibbo supressed $\boldsymbol{B^0 \to D^0\pi^0}$ decay. Vertex CKM factors are labelled showing the $\boldsymbol{\frac{V_{cb}V_{ud}^*}{V_{cd}^*V_{ub}} \approx \lambda^2}$ suppression of the DCS process.

different weak phase, but also results in it being supressed by a factor of $\frac{V_{cb}V_{ud}^*}{V_{cd}^*V_{ub}} \approx \lambda^2 \approx 0.05$ compared to the Cabibbo-favoured term. Additionally, the decays $D^0 \to K^+\pi^-$ and $D^0 \to K^+\pi^-\pi^0$ are also supressed by a factor of λ^2 compared to the favoured decays $\overline{D^0} \to K^+\pi^-$ and $\overline{D^0} \to K^+\pi^-\,\pi^0$. The result of this is the amplitude of the secondary process is heavily suppressed compared to the primary process, meaning interference effects can be neglected and the theoretical prediction is no measurable CP violation.

Because the W-boson does not carry a colour charge, the quark anti-quark pair that is produced from it must be colour neutral. In the case of $B^0 \to \overline{D^0}\pi^0$, these do not hadronise into the same particle (u in the $\overline{D^0}$ and $\bar{d}$ in the π^0) and as a result, in order for the final state hadrons to remain colour neutral (as is required by QCD), the $\bar{b}$ and d quark (anti-quark) must also be of the same colour (anti-colour). For this reason, $B^0 \to \overline{D^0}\pi^0$ (and other similar decays) are referred to as colour suppressed. This restriction reduces the amplitude of colour suppressed decays by a factor of $\frac{1}{3}$ (and thus $\mathfrak{B}$ by $\frac{1}{9}$) compared to colour favoured decays.

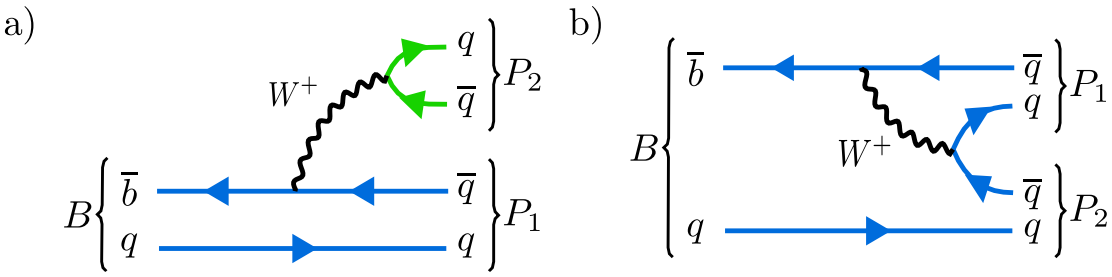


Figure 6: Examples of colour favoured and colour suppressed diagrams. In a), the quarks that form $\boldsymbol{P_2}$ can be a different colour/anticolour to those that form $\boldsymbol{P_1}$ and the initial $\boldsymbol{B}$. In b) both pairs of quarks must be the same colour/anticolour as the initial $\boldsymbol{B}$.

1.8 Motivation of Colour Suppressed $B^0 \to \overline{D^0}\pi^0$ Decays

Colour suppressed hadronic decays such as $B^0 \to \overline{D^0}\pi^0$ offer an interesting avenue of study for several reasons, both directly as experimental tests of theory and as a high precision control for other experimental investigations.

Previous results [15, 16] for branching fraction measurements of colour suppressed $B^0 \to \overline{D^0}h^0$ (where h^0 is a light neutral meson) decays show large deviations from the predictions of the "naïve" QCD factorisation model [17-19] where FSIs are neglected. Several theoretical models that aim to achieve a better description of these modes with large FSI contributions such as SCET (Soft Collinear Effective Theory) [20, 21] and pQCD (perturbative QCD) [22, 23] by accounting for rescattering effects between colour suppressed $B^0 \to \overline{D^0}h^0$ and colour allowed $B^0 \to \overline{D^{\pm}}h^{\mp}$ diagrams. Of these decays, $B^0 \to \overline{D^0}\pi^0$ has been shown [24] to have particularly dominant non-factorisable components, so tight experimental constraints on its properties are valuable in comparing different models of accounting for this FSIs.

In the standard model, $b \to c\overline{u}d$ transitions have no penguin terms owing to the lack of flavour changing neutral currents, and as such direct CP violation in these modes is expected to be minimal. Measurements of notable CP violation in one of these decays would be of significant interest and could hint at penguin contributions from BSM physics diagrams. Time-dependent analysis has measured $C\left(B^0 \to D_{CP}^{(*)}h^0\right) = (-2 \pm 8) \times 10^{-2}$ [25], consistent with this expectation, but time-independent studies of A_{CP} allow for more precise measurements of direct CP violation to test against Standard Model predictions.

A high precision measurement of branching fraction and A_{CP} for $B^0 \to \overline{D^0}\pi^0$ has further utility on top of comparison to theoretical predictions. $B^0 \to \overline{D^0}\pi^0$ is a

common control mode for use in much rarer charmless hadronic B decays such as $B \to \pi^0\pi^0$ and $B \to K_s\pi^0$. Precise measurement of properties for a control mode are important in giving a platform for the validation and refinement of analysis techniques prior to studying experimental data.

1.9 $B^+ \to \overline{D^0}\pi^+$ Decays

The $B^+ \to \overline{D^0}\pi^+$ decay is studied primarily as a control mode to validate the methodology used to analyse $B^0 \to \overline{D^0}\pi^0$. It is chosen because it is a very similar decay. Like $B^0 \to \overline{D^0}\pi^0$, $B^+ \to \overline{D^0}\pi^+$ proceeds through the $b \to c\bar{u}d$ quark process, with the difference being the change in spectator quark from a d to a u. This means at the quark level the diagram topology will be the same. It differs however, in that $B^+ \to \overline{D^0}\pi^+$ is not colour suppressed. It receives contribution from both colour favoured and colour suppressed diagrams, as there is a u quark in both the $\overline{D^0}$ and π^+ mesons, as shown in Figure 7. This gives a large increase to the decay amplitude of $B^+ \to \overline{D^0}\pi^+$, and its branching fraction is of order 10 times larger than $B^0 \to \overline{D^0}\pi^0$. This also contributes to it being excellent for a control study, as it will give a large signal yield to study modelling in detail.
As $B^+ \to \overline{D^0}\pi^+$ has the same quark level $b \to c\bar{u}d$ process and is studied in the same D^0 decays, it has no penguin diagram contributions, and interference effects from Cabibbo suppressed processes are also negligible. This leads to the same

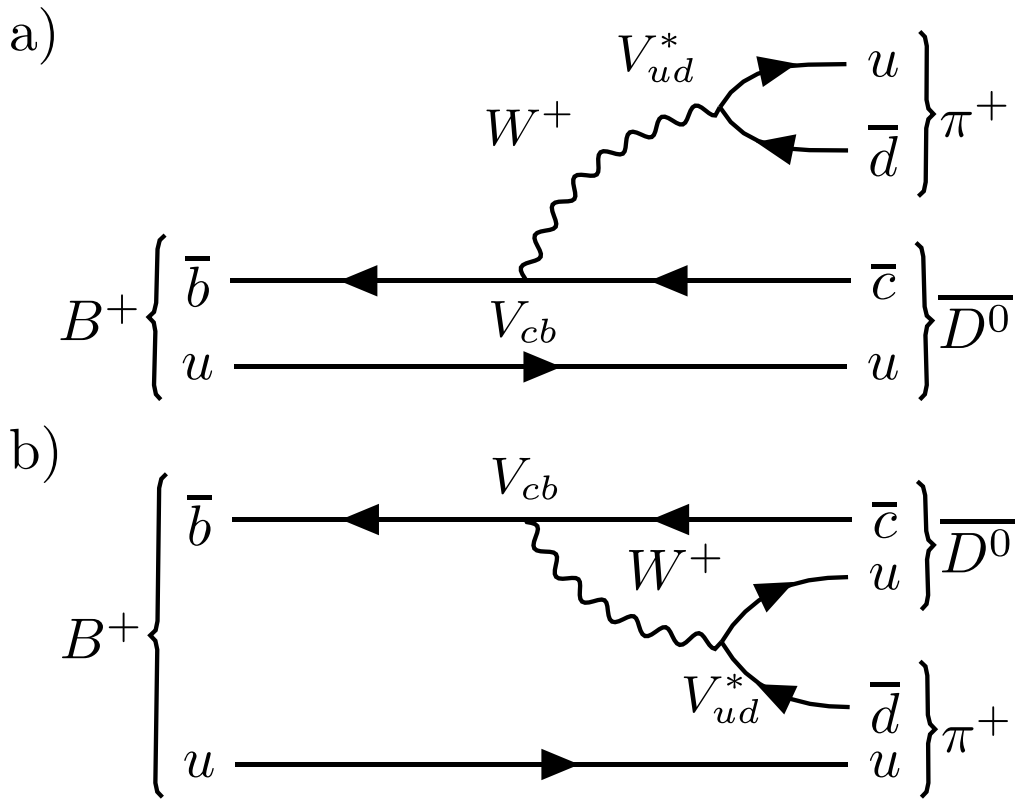


Figure 7: Tree level Feynman diagrams for the $B^+ \to \overline{D^0}\pi^+$ decay. a) is the colour favoured contribution, and b) shows the colour suppressed contribution.

theoretical prediction of no direct CP violation. The current PDG world average value [3] of

$$A_{CP,PDG} = (-0.7 \pm 0.7) \times 10^{-2} \quad (1.76)$$

is in agreement with this.

$B^+ \to \overline{D^0}\pi^+$ is another very common control mode in other studies of hadronic B decays. Having as accurate as possible measurements of its properties is important for increasingly precise studies in flavour physics to compare with when validating techniques. For this purpose, as well as serving as a control mode, a complete analysis is performed and the branching fraction and direct CP asymmetry for $B^+ \to \overline{D^0}\pi^+$ (with $\overline{D^0}$ decaying to the Cabibbo favoured final states $K^+\pi^-$ and $K^+\pi^-\pi^0$) are reported as a new measurement.

2 The Belle Detector and KEKB Accelerator

The High Energy Accelerator Research Organization, known as KEK, is a physics research institute located in Tsukuba, Japan. One of the main projects at KEK is the Belle experiment, with the aim of investigating CP violation, specifically in the B meson system. It is an international collaboration with members from around 20 countries and over 60 institutions. The two main facilities that make up the Belle experiment, the KEKB accelerator, and the Belle detector are briefly summarised in this chapter. The experiment collected data from 1999 until 2010, but the collaboration is still active as analysis of the data continues.

2.1 KEKB Accelerator

The KEKB accelerator [26, 27] is an asymmetric energy e^+e^- collider operated by KEK in Tsukuba, Japan. The Belle experiment is one of two so-called B factories (the other being the Babar experiment, based at SLAC [28]) due to the large number of B mesons produced for study. To meet this requirement, KEKB operated at a centre of mass (CoM) energy of $\sqrt{s} = 10.58\ GeV$, which corresponds to the rest mass of the $\Upsilon(4S)$ meson (a bound state of $b\overline{b}$). This produces B mesons through the interaction

$$e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\overline{B}. \tag{2.1}$$

The mass of the $\Upsilon(4S)$ is just slightly more than twice that of a B meson ($m_B = 5.28\ GeV$), and it decays to a $B\overline{B}$ pair more than 96% of the time. The production of neutral and charged B pairs is almost equal, with the ratio of branching fractions being [3]

$$\frac{\mathfrak{B}(\Upsilon(4S) \rightarrow B^+B^-)}{\mathfrak{B}\left(\Upsilon(4S) \rightarrow B^0\overline{B^0}\right)} = 1.058 \pm 0.024. \tag{2.2}$$

One of the main goals of the Belle experiment is to measure time dependent CP violation, which requires measurement of the difference in the decay time of the two B mesons. The timescales in this are very small, with the mean life-time of a B meson being $(1.519 \pm 0.004) \times 10^{-12}\ s$ [3]. In order for this to be measurable as the distance between decay vertices, the B mesons must be travelling very fast. This is the reason for the asymmetric energy of the collider.e^+e^- colliders is that B and $\overline{B}$ mesons are produced in equal amounts, allowing CP violation to be measured by counting the number of decay events and their CP conjugates, as

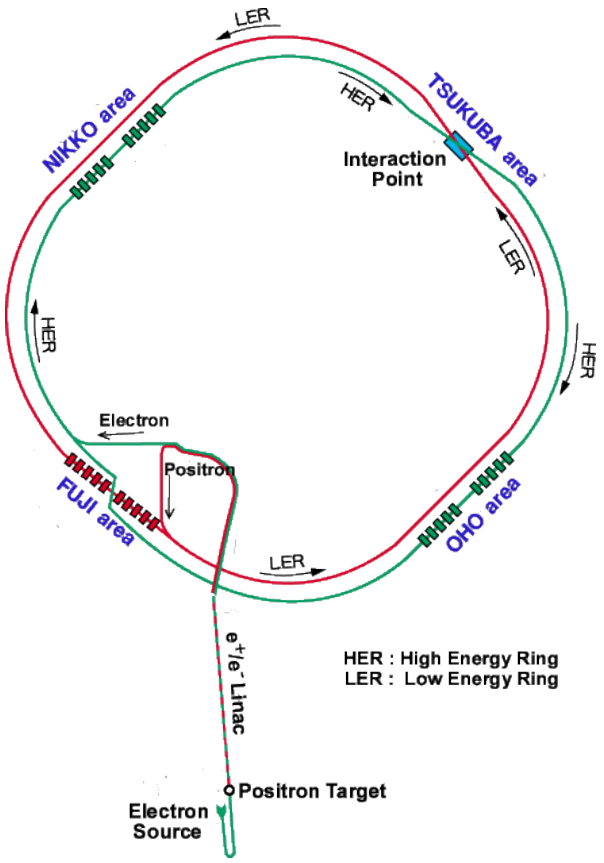


Figure 8: KEKB accelerator, showing the two storage rings, linear accelerator and Interaction point. [29]

$$\frac{\Gamma(P \to f) - \Gamma(\overline{P} \to \overline{f})}{\Gamma(P \to f) + \Gamma(\overline{P} \to \overline{f})} = \frac{N(P \to f) - N(\overline{P} \to \overline{f})}{N(P \to f) + N(\overline{P} \to \overline{f})}, \tag{2.3}$$

where $N(P \to f)$ is the number of events of the decay $P \to f$.

The KEKB accelerator has two storage rings: A High Energy Ring (HER) containing electrons travelling clockwise at 8.0 GeV, and a Low Energy Ring (LER) with positrons travelling anticlockwise at 3.5 GeV. These rings are around 3km in circumference with electrons and positrons being injected by a linear accelerator (linac) at collision energy. This asymmetric energy gives a Lorentz boost to the produced $\Upsilon(4S)$ of

$$\beta\gamma = \frac{E_- - E_+}{\sqrt{s}} \sim 0.425, \tag{2.4}$$

where E_- is the energy of the electron beam, E_+ is the energy of the positron beam and $\sqrt{s}$ is the CoM energy. This boost means that the average distance between the decay vertices of the B mesons in the lab frame is roughly 200 μm.

The beams consist of bunches of around 10^{11} particles, with the instantaneous luminosity defined as

$$L = \frac{N_B N_{e^+} N_{e^-} f}{A_{eff}}, \tag{2.5}$$

where N_B is the number of bunches in circulation, N_{e^+} and N_{e^-} are the number of positrons and electrons in each bunch respectively, f is the frequency of circulation, and A_{eff} is the overlapping cross-sectional area at the interaction point (IP). The design luminosity of KEKB was $1.0 \times 10^{34} cm^{-2} s^{-1}$, which was exceeded, setting the world record for instantaneous luminosity of $2.1 \times 10^{34} cm^{-2} s^{-1}$ [26]. The total time integrated luminosity recorded by Belle at the $\Upsilon(4S)$ resonance was $711\, fb^{-1}$, which corresponds to 772×10^6 $B\overline{B}$ pairs produced.

The two beams collide at the IP, around which the Belle detector is built. In order to reduce background from synchrotron radiation and parasitic collisions [26], separation-bend magnets are not used, and the beams cross at an angle of $22\, mrad$. This finite crossing angle results in a reduction in luminosity from the angle between the particle bunches, so in 2007 Belle installed a pair of crab cavities in each ring [30]. These crab cavities use a time varying RF field to rotate the e^+ and e^- bunches to ensure maximum overlap at the IP, as shown in Figure 9.

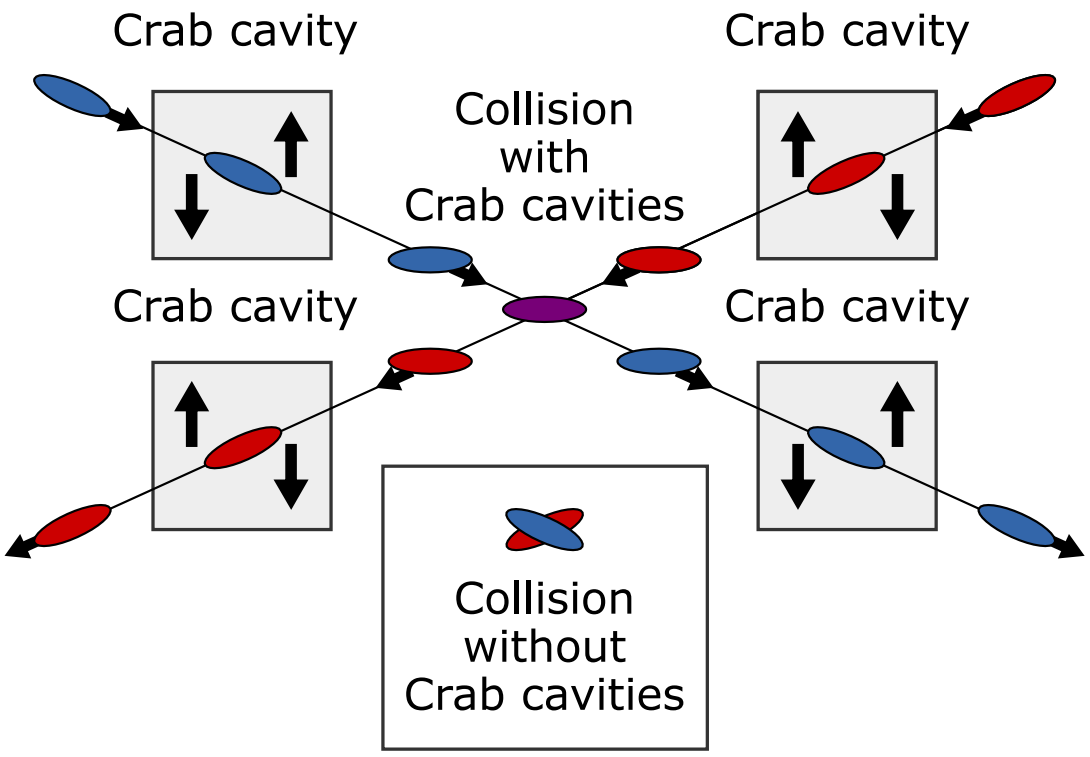


Figure 9: Diagram showing bunch rotation by crab cavities. [31]

2.2 BELLE Detector

When a particle collision happens, the full decay chain of particles cannot be directly observed, as the heavy hadrons will decay almost instantly. Only the lighter particles at the end of the chain have a long enough lifetime to be observed. A complex detector is required to measure a range of quantities from these final state particles in order to reconstruct the full decay.

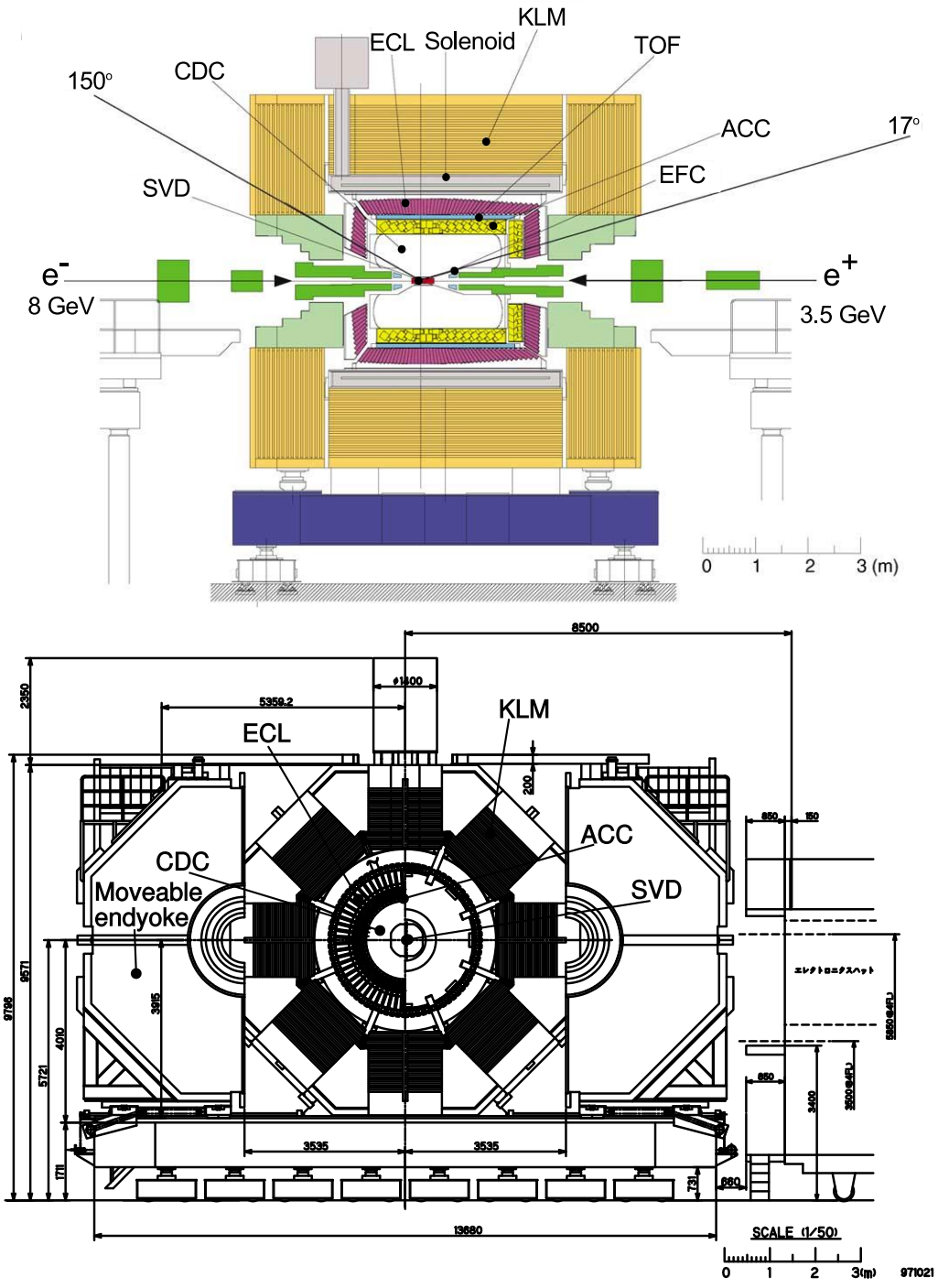


Figure 10: Cross-sectional schematic of the Belle detector from the side (top) and front (bottom) [32].

The Belle detector [32] is a multi-purpose spectrometer with a large solid angle coverage located around the IP of the KEKB accelerator. Moving outwards from the IP it contains: a silicon vertex detector (SVD), extreme forward calorimeter (EFC), central drift chamber (CDC), aerogel Cherenkov counters (ACC), time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL), located inside a solenoid that generates a 1.5 T magnetic field. Outside of this is the K_L and muon detector (KLM) that also provides magnetic flux return. Figure 10 is a schematic showing the location of these components in the Belle detector.

The co-ordinates used with the Belle detector place the origin at the IP, with the z-axis in the direction of the magnetic field (the same direction as the e^- beam line), x-axis horizontally away from the accelerator ring and y-axis vertically up. Typically, co-ordinates are given in polar form, with the polar angle θ being measured from the positive z-axis and the azimuthal angle ϕ from the positive x-axis.

2.2.1 IP and Beam Pipe

For the most accurate measurement of B vertices, the SVD should be placed as close to the IP as possible, both to reduce coulomb scattering effects and to improve the resolution of the SVD. However, the beam pipe can be subject to heating from energies of over a hundred watts, so it needs to shield the detector components from this heat. To meet these needs, the beam pipe at the IP is constructed from a double-walled beryllium cylinder with coolant pumped in between the walls. Beryllium was chosen for its low atomic number, which reduces the effect of coulomb scattering. Two versions of the beam pipe were used in conjunction with the 2 SVD setups. The first had an inner radius of 20.0 mm, outer radius of 23.5 mm and used helium gas as its coolant. The outer wall had a 20 μm thick gold foil covering to reduce synchrotron radiation. A replacement, smaller beam pipe was fitted in 2003 alongside the new SVD2. This has an inner radius of 15.0 mm, outer radius of 18.5 mm (with 10 μm of gold foil covering) and

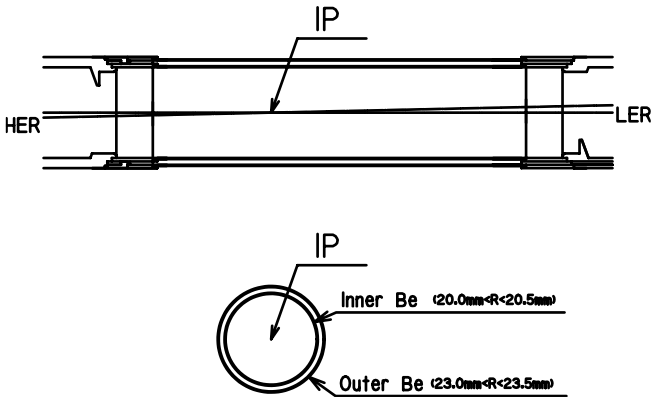


Figure 11: Cross-sectional schematic of the first Belle beam pipe [32].

used a liquid paraffin coolant [33]. A schematic of the first beampipe is shown in Figure 11.

2.2.2 Silicon Vertex Detector

The silicon vertex detector [32, 34, 35] is the innermost sub-detector of the Belle detector and is critical for measuring the location of B decay vertices so time dependent CP violation analyses can be conducted. It also plays a role in particle tracking and momenta determination. The SVD is able to reconstruct B decay vertices to a resolution of 100 μm, allowing for the measurement of decay length difference of B mesons, which is on average 200 μm [29].

The SVD consists of multiple layers of double-sided silicon strip detectors (DSSD). A DSSD is essentially a p-n junction with a reverse bias (75V) to create a full depletion zone between the p^+ doped silicon strip and the n^- doped silicon strip. When a charged particle passes through a DSSD and liberates electrons though ionisation, they will create an electric pulse as they hit the nearest p^+ strip, as will the positive holes in the nearest n^- strip. The p^+ strips are aligned with the z-axis, while the n^- strips are perpendicular, which allows charged particle hits to be located in the 2D plane of the DSSD. With multiple layers the trajectory can be reconstructed, and by looking at all charged particles the B decay vertex from which they originate can be calculated.

The original SVD installed in the Belle detector, SVD1, used three layers of DSSDs, as shown in Figure 12. The first layer contained 8 DSSD modules (called "ladders") located at a radius of 30 mm from the IP, the second layer contained 10 ladders at

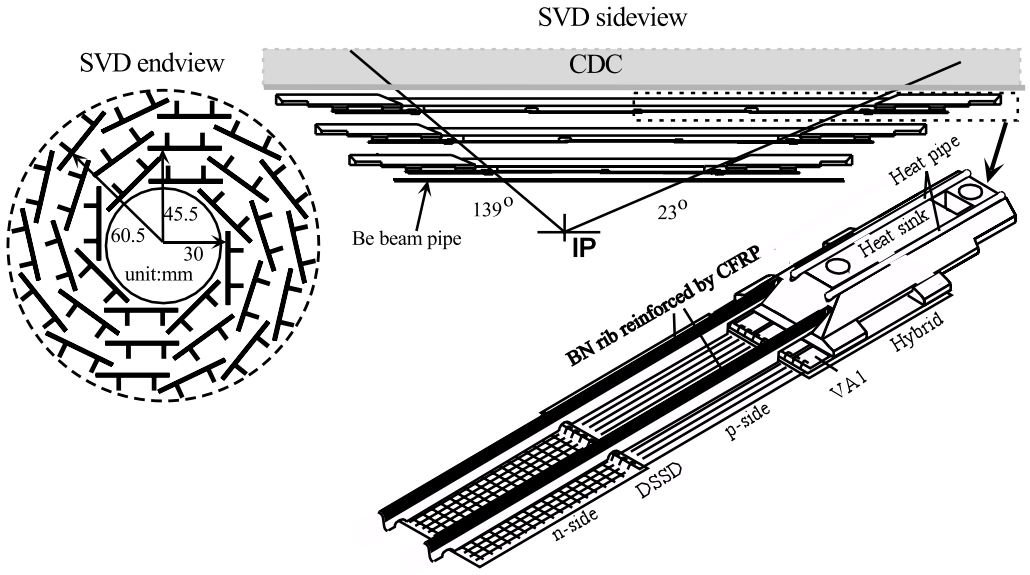


Figure 12: Cross-sectional view of SVD1 [32].

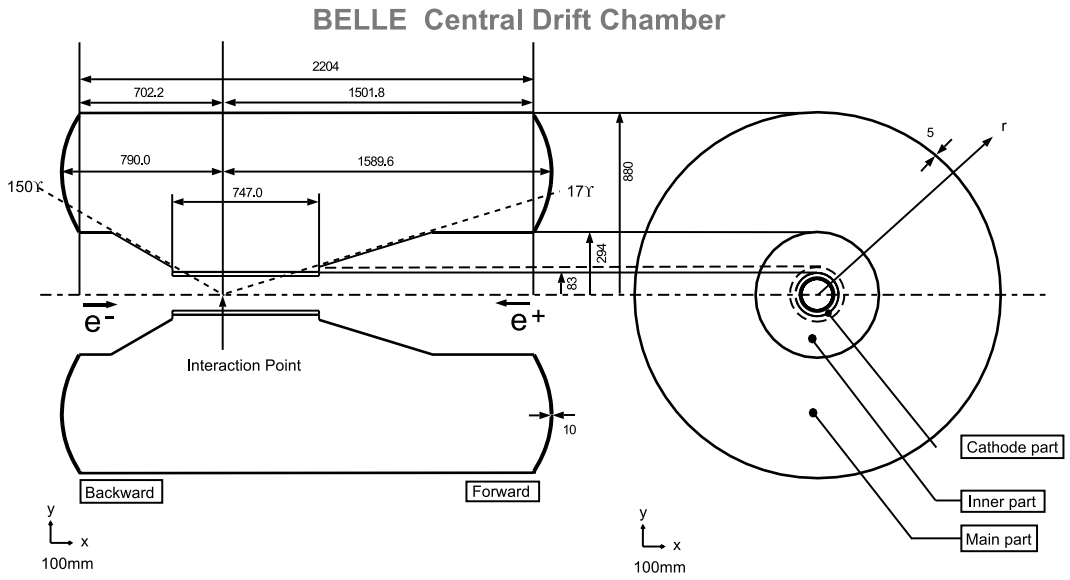


Figure 13: Layout and dimensions of the CDC [32].

a radius of 45.5 mm and the outer layer contained 14 ladders at a radius of 60.5 mm. This had an acceptance region of $23° < \theta < 139°$. In 2003, along with the new beam pipe, a new SVD was installed, SVD2 [35]. The narrower beam pipe allowed the inner layer to be closer to the IP, and SVD2 contained four layers. From innermost to outermost, the layers consisted of: 6 ladders at 20 mm radius, 12 ladders at 43.5 mm radius, 18 ladders at 70 mm radius and 18 ladders at 88 mm radius. SVD2 also had a larger angular acceptance of $17° < \theta < 150°$. This resulted in the SVD2 having an improvement in z-vertex resolution for low-momentum tracks of about 20% [35].

2.2.3 Central Drift Chamber

Because of the 1.5T magnetic field, charged particles will trace a helical trajectory as they move through the detector. The central drift chamber (CDC) [36] tracks the trajectory of charged particles and uses their curvature in the magnetic field to determine momentum. It also contributes to particle identification (PID). The momentum resolution requirement for particles with transverse momentum (p_T) greater than 100 MeV/c was

$$\frac{\sigma_{p_T}}{p_T} \sim 0.005\sqrt{1 + p_T^2}, \qquad (2.6)$$

where σ_{p_T} is the transverse momentum resolution. This requirement is met by the detector [32].

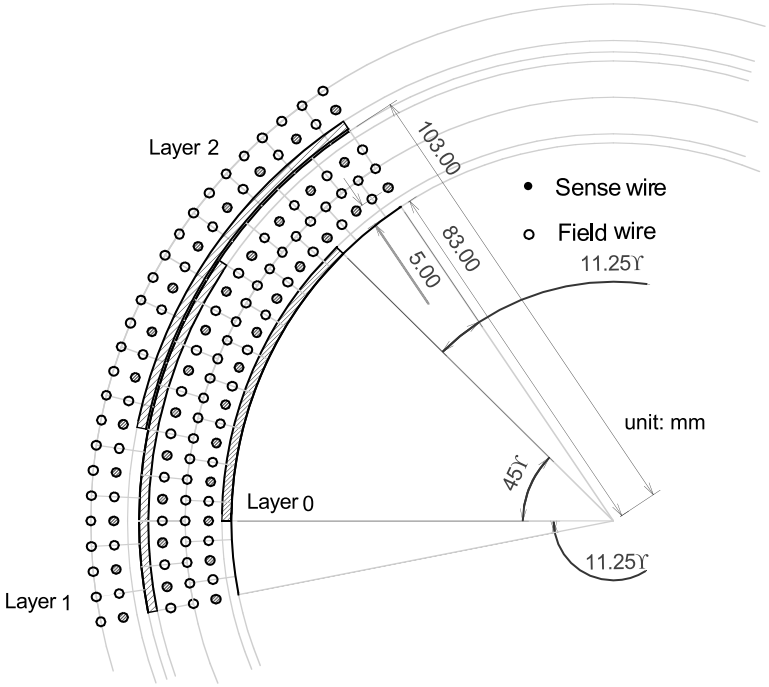


Figure 14: Drift cell structure in the CDC [32].

The CDC is filled with a mixture of 50% helium, 50% ethane gas, and consists of 8400 drift cells arranged in 50 layers, as shown in Figure 13 and Figure 14. Each cell contains an anode sense wire surrounded by cathode field wires. As a charged particle passes through the gas it will ionise it, freeing electrons which drift towards the wires. As these pick up speed they will further ionise the gas creating a cascade that is detected as an electric pulse. The size of this pulse is directly proportional to the number of electrons freed by the charged particle, and from this, to its energy loss as it traverses the gas. This energy loss can be related to the velocity of the particle by the Bethe formula [37],

$$-\frac{dE}{dx} = \frac{4\pi N_A z^2 e^5}{m_e v^2}\frac{Z}{A}\ln\left(\frac{2m_e v^2}{I(1-\beta^2)} - \beta^2\right), \qquad (2.7)$$

where $\frac{dE}{dx}$ is the loss in energy of the particle with distance, N_A is Avogadro's number, z is the charge of the particle, m_e and e are the mass and charge of an electron, v is the velocity of the particle, Z and A are the atomic number and atomic mass of the gas respectively, $\beta = \frac{v}{c}$, and I is the mean excitation potential of the gas. The energy loss depends on the velocity of the particle, which for a given momentum is related to its mass, and thus the type of particle it is. Figure 15 shows the relation between $\frac{dE}{dx}$ and momentum for various particles, indicating the good separation between particle types at low momentum ($< 0.8\ GeV$).

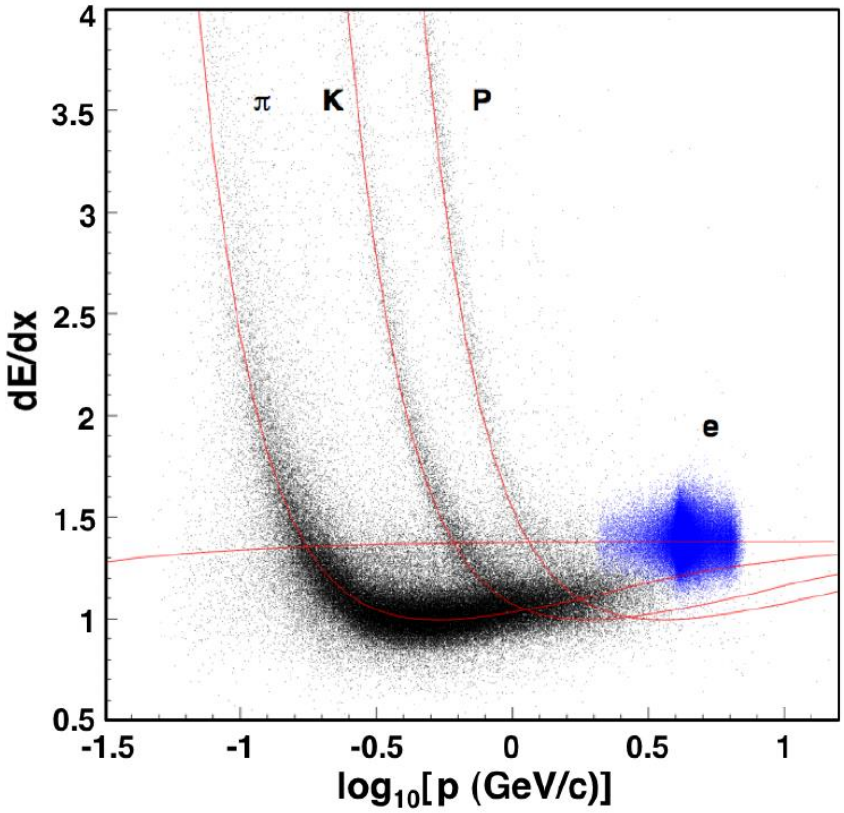


Figure 15: Relationship between $\boldsymbol{dE/dx}$ and momentum of particles. Points show measured values, with lines showing expected truncated mean [32].

The CDC contains both axial layers, where the wires are parallel to the beam line, and stereo layers, which are at a slight angle to allow the trajectory to the tracked in the z-direction as well. This slight angling of the stereo layers is important for this analysis as it plays a part in detector asymmetry, discussed further in chapter 4.5.2. The helium/ethane gas mixture is chosen for several reasons. As with the beryllium beam pipe, low atomic number is desirable to reduce coulomb scatterings. A low photoelectric cross-section also has the effect of reducing backgrounds caused by synchrotron radiation, and the ethane increases the rate of energy loss due to its larger electron density, resulting in increased accuracy of $\frac{dE}{dx}$ measurement.

2.2.4 Aerogel Cherenkov Counter

The aerogel Cherenkov counter (ACC) [38, 39] is an important part of PID in the Belle detector, with the main purpose of distinguishing between high momentum ($> 1.2 GeV/c$) pions and kaons that the CDC cannot identify based on $\frac{dE}{dx}$ information.

Cherenkov radiation is emitted by charged particles travelling through a dielectric medium with velocity greater than the speed of light in that medium [40],

$$\beta > \frac{1}{n}, \tag{2.8}$$

where n is the refractive index of the medium. The silica aerogel medium used in the ACC was chosen with a refractive index of $1.01 - 1.03$ depending on the polar angle of the detector element (see Figure 16), so that pions with a momentum of $1.2 - 3.5\ GeV$ will produce Cherenkov radiation, but kaons within the same momentum range will not.

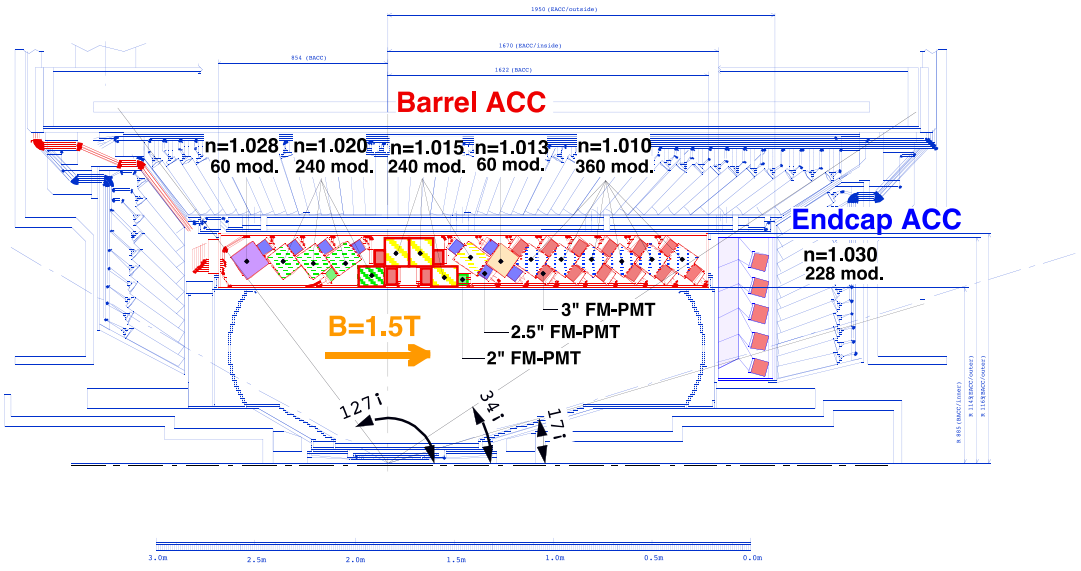


Figure 16: Side on layout of the ACC in the Belle detector. Barrel and endcap regions are shown, as is the refractive index used for different polar angles [32].

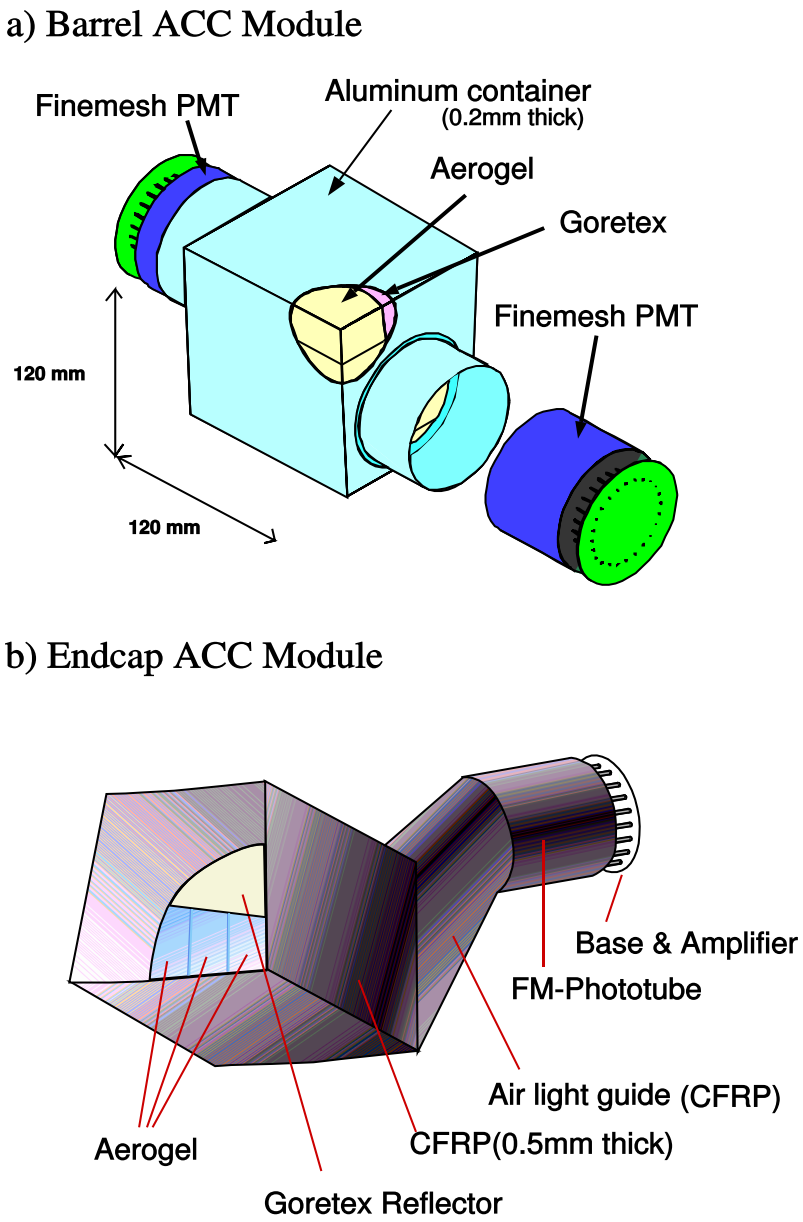


Figure 17: Design of the two types of ACC modules used in a) the barrel region and b) the endcap region [32].

The ACC is separated into two regions, the barrel and the endcap, as shown in Figure 16. The barrel region covers $33.3° < \theta < 127.9°$ and consists of 960 silica aerogel counter modules. The endcap region covers $13.6° < \theta < 33.4°$ and has 288 modules. Each module contains one or two fine mesh photo-multiplier tubes (FM-PMT) used to detect the Cherenkov light. These are chosen because they operate effectively in the strong 1.5T magnetic field of the detector. Construction of the ACC modules are shown in Figure 17.

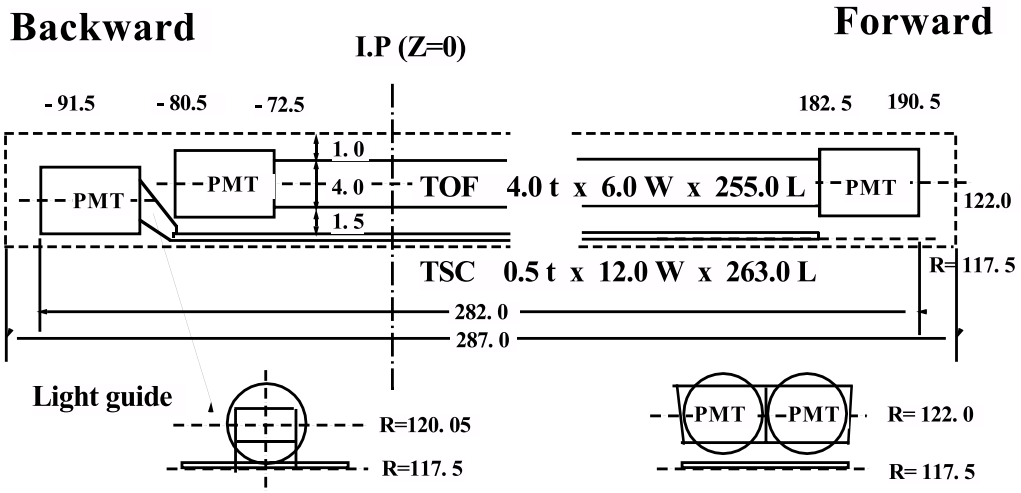


Figure 18: Layout of TOF module, including trigger scintillation counter (TSC) [32].

2.2.5 Time-of-Flight detector

The time-of-flight (TOF) detector [41] is another component of the Belle detector used for PID, specifically for low momentum particles ($p < 1.2\ GeV$) that constitute around 90% of particles produced by $B\overline{B}$ decays at Belle. The TOF has a rather simple purpose, to determine the time taken from the instant of collision for particles to reach it. From there the mass can simply be calculated from special relativity as

$$m = \frac{p}{c}\sqrt{\left(\frac{cT}{L}\right)^2 - 1}, \tag{2.9}$$

where m is the mass of the particle, p is its momentum, T is the time after collision, and L is the distance between its detection point in the TOF system and the IP.

The TOF detector consists of 128 modules, each containing three plastic scintillation counters, two for use by the TOF and one for the trigger system (see Section 2.2.10), with readout by FM-PMTs at either end (shown in Figure 18). The system achieves a time resolution of 100 ps, and the minimum momentum for particles to reach it is 0.28 GeV. Particles with momentum lower than this can only be identified with $\frac{dE}{dx}$ information from the CDC.

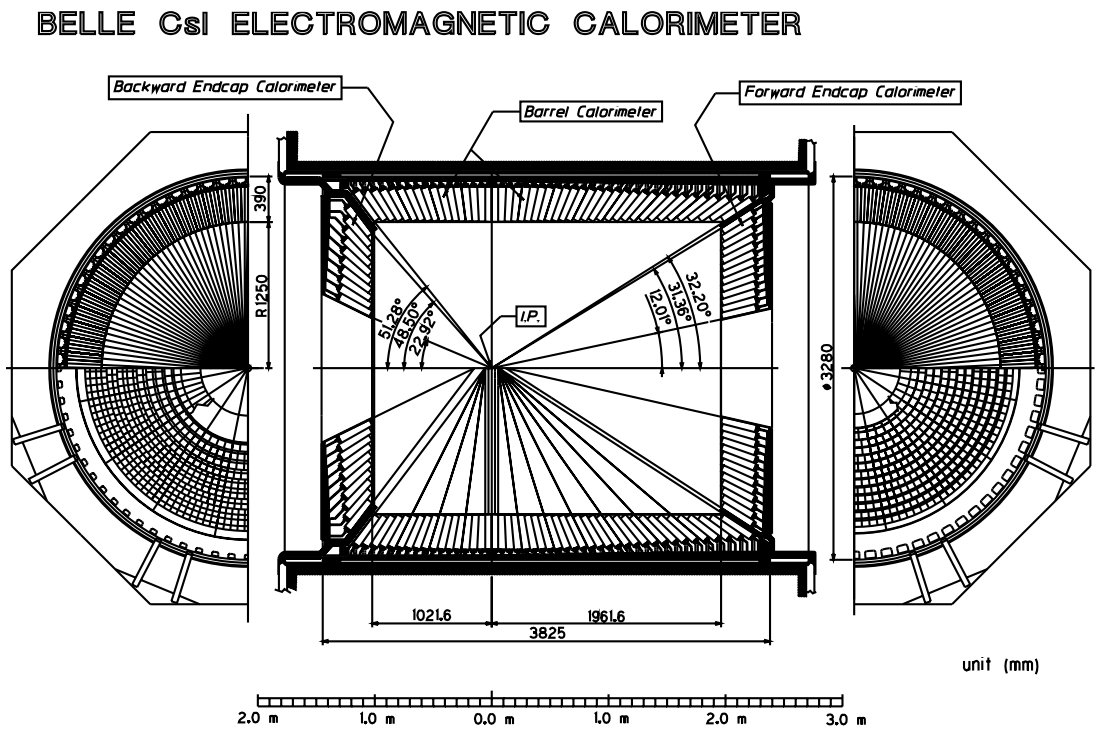


Figure 19: Schematic showing the layout of the barrel and endcap sections of the ECL [32].

2.2.6 Electromagnetic Calorimeter

The electromagnetic calorimeter (ELC) [42] absorbs the energy of photons and charged particles via electromagnetic interactions, so that their total energy is measured. The energy of the photons resulting from B decays can range from 20 MeV as products of neutral particles at the end of a long decay, to 8 GeV products of Bhabha events, so good energy resolution over a large energy range is required. Neutral particles like $\pi^0 s$, which immediately decay to photons, will not be detected by the CDC tracking system, so rely on the information from the ECL to measure their properties.

The ECL is constructed from Thallium-doped Caesium Iodide (CsI(Tl)) crystals. Photons interact with the material in the crystals, resulting in e^+e^- pair production, while electrons and positrons, (called just electrons for now on), emit photons via bremsstrahlung. The photons and electrons created by these interactions then undergo further pair production and bremsstrahlung, resulting in an electromagnetic cascade. This cascade generates scintillation light in proportion to its energy, which is detected by a pair of silicon photodiodes. Thallium-doping is used to shift the scintillation light into the visible spectrum for detection by the photodiodes.

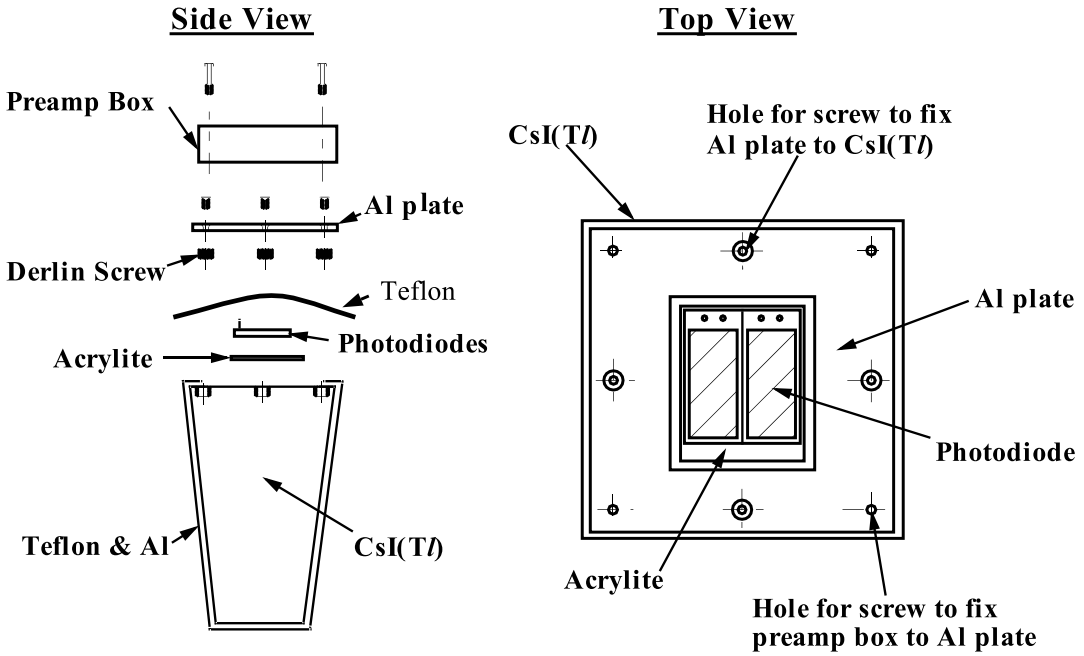


Figure 20: Construction of an ECL counter [32].

Photon deposits can be separated from those of electrons by matching the location of the deposit in the ECL to charged tracks in the CDC. Those without corresponding tracks are due to photons. Electrons can be distinguished from hadronic particles based on two quantities, $\frac{E}{p}$ and $\frac{E9}{E25}$ [43] (defined as follows). High momentum electrons deposit almost all their energy in the ECL, so energy measured by the ECL, E, and the momentum measured by the CDC, p, should be close to equal, giving $\frac{E}{p} = 1$. Hadronic particles will either interact elastically with minimal ionisation energy loss, or inelastically through strong interactions creating a hadronic shower, which will not be well contained, both leading to $\frac{E}{p} < 1$. Because of other energy loss effects, low momentum electrons have $\frac{E}{p} < 1$, so the shower shape is also used for particle discrimination.

$\frac{E9}{E25}$ is defined as the energy deposited in the nine closest ECL counters centred on the initial hit, divided by the energy deposited in the 25 closest ECL counters. Electron showers tend to be very narrow, so this will be close to 1. For particles that interact hadronically with the ECL, their showers will be much wider with leakage into neighbouring cells, giving $\frac{E9}{E25} < 1$.

Like the ACC system, the ECL is divided into two sections, the barrel and the endcap (the ECL has both a forward and backward endcap unlike the ACC). The barrel is 3.0 m long with an inner radius of 1.25 m, covering a polar angle region of $32.2° < \theta < 128.7°$, and contains 6624 crystals. The endcaps are located at

either end of the barrel extending the polar angle coverage to $12.4° < \theta < 155.1°$, with the forward (backward) endcap containing 1152 (960) crystals. This layout is shown in Figure 19. The ECL is finely segmented to give good position resolution on deposits, which is required to reconstruct high momentum $\pi^0 \to \gamma\gamma$ decays that will deposit energy close together.

The exact dimension of the CsI(Tl) crystals vary based on their location in polar angle, but are all 30 cm in length. This corresponds to roughly 16 radiation lengths for electrons [44], ensuring the majority of energy for electrons and photons will be contained within the calorimeter. Each crystal is placed in a Teflon and aluminium box to electrically insulate it and reduce light leakage, but this cannot be completely prevented. All the crystals are arranged pointing towards the IP, with the top containing the two photodiodes and preamp to detect the scintillation light. The construction of an individual ECL counter is shown in Figure 20.

2.2.7 Extreme Forward Calorimeter

The extreme forward calorimeter (EFC) [32] is an electromagnetic calorimeter located very close to beam pipe at the polar angles $6.4° < \theta < 11.5°$ and $163.3° < \theta < 171.2°$. In this region the EFC is subject to extreme levels of radiation, so is constructed from radiation-hard Bismuth Germanate ($Bi^4Ge^3O^{12}$) crystals, that trade the accuracy of the CsI(Tl) crystals used in the ECL for durability to radiation damage. The EFC is only used in certain physics analysis, such as $B \to \tau\nu$, where the extended calorimeter range is important. It also used for luminosity and beam monitoring, and to provide a mask against backgrounds for the CDC.

2.2.8 Superconducting Solenoid

A uniform 1.5T magnetic field is generated inside the Belle detector by the superconducting solenoid magnet [32] that surrounds all the components except the KLM. The magnetic field allows the CDC to measure momenta based on the curvature of particles, as well as the sign of their charge.

The solenoid is 4.4 m in length and 3.4 m in diameter, constructed from a single layer Niobium-Titanium-Copper superconducting coil around a high purity (99.99%) aluminium stabilizer cylinder. This is cooled to 4.5 K by the circulation of liquid helium through a stainless-steel cooling tube outside the coil.

2.2.9 K_L and Muon Detector

The outermost part of the Belle detector is the K_L and muon (KLM) detector [45]. As the name suggest, its primary role is the detection of K_L and $\mu^{\pm}$ particles, whose long life-time and high penetration means they typically do not decay in the inner detector nor deposit their energy fully in the ECL. Charged particles in the KLM are detected by glass resistive plate chambers (RPC), which consists of two high resistivity electrode plates in parallel, separated by a gap filled with a non-combustible gas. When a charged particle ionises the gas, it induces a discharge between the plates, which is picked up by external readout strips above and below (one with θ orientation, the other ϕ), giving a time and location of the particle passing through. Figure 21 shows the layout of a single detector layer in the barrel region, with two RPCs between the sets of readout strips for improved efficiency. Although the KLM can detect particles, it is unable to measure their energy.

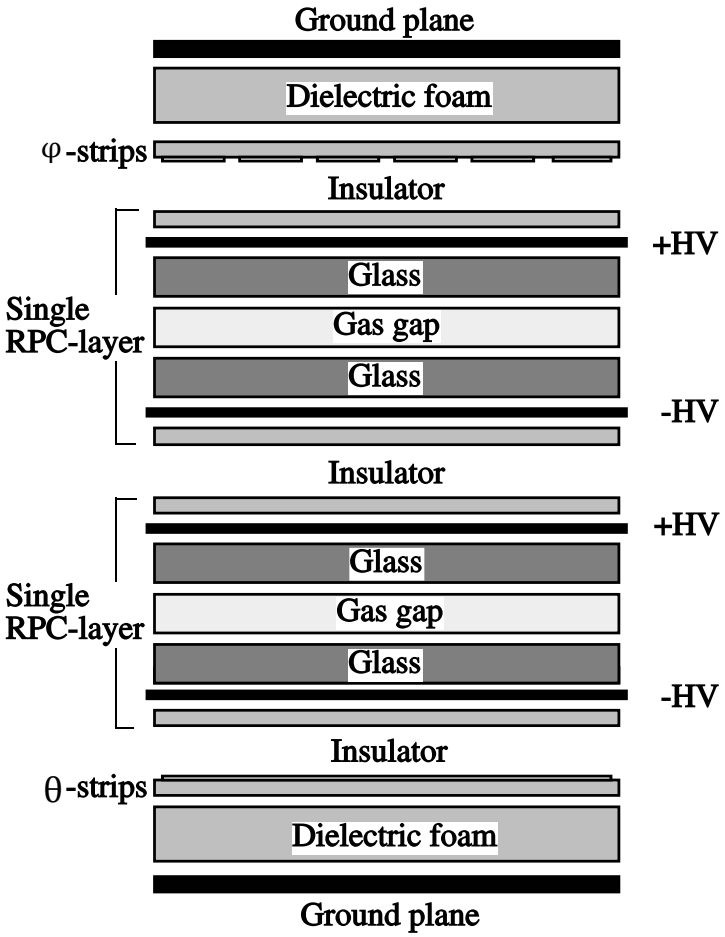


Figure 21: Cross-section of a KLM detector layer in the barrel region [32].

The KLM consists of 15 (14) detector layers alternating with 14 inactive layers in the barrel (endcap) region. These inactive layers are iron plates that double as the magnetic flux return for the solenoid. They provide 3.9 nuclear interaction lengths

(this is the mean distance travelled by a hadron before inelastic nuclear interaction) for K_L particles. Nuclear interactions generate showers of ionising particles detected by an RPC and absorbed by subsequent inactive layers. Discharge events that occur in a small number of layers that do not have a corresponding charge track in the CDC are identified as K_Ls and their direction of flight found by tracing back to the IP.

Muons do not interact via the strong force and will predominately lose energy via ionisation. This means they penetrate further through the KLM than hadronic particles that are both absorbed by the ECL and undergo nuclear scattering and absorption in the KLM. Consequently, muons are distinguished from light charged hadrons by long tracks of discharges in the KLM detector layers.

2.2.10 Trigger and Data Acquisition

If the recording of all data observed by the detector sub-systems was attempted, the bandwidth capacity of the data acquisition system (DAQ) would be overwhelmed. The majority of signals observed originate from background processes, such as beam interaction with residual gas in the beam pipe or synchrotron radiation. The job of the trigger system [46] is to discriminate between these background events and events of physical interest. These physically interesting events are primarily $e^+e^- \to \Upsilon(4S) \to B\overline{B}$, but also include $e^+e^- \to q\overline{q}$, $e^+e^- \to \ell^+\ell^-$ and two photon ($e^+e^- \to \gamma\gamma$) events. The Belle trigger system consists of two online triggers, the level-1 hardware trigger and the level-3 software trigger, and an offline level-4 software trigger.

The level-1 hardware trigger takes input from all subdetectors (except the SVD) for indications that events of interest took place [46]. The trigger signals from each of the subdetectors is then sent to the global decision logic (GDL) to assess and decide on whether to record the event or not. The primary basis for this decision is the number of charged tracks detected by the CDC and TOF, the total energy deposited in the ECL, and the identification of muons in the KLM. Detection of Bhabha or two-photon events in the EFC is also recorded, separate from physics data, for use in studies of the beam. The signal from the subdetectors must arrive within 1.85 μs of the event, with the final trigger decision after 2.2 μs. The efficiency of the level-1 trigger is close to 100%.

Events that pass the level-1 trigger are processed by the DAQ. The event builder combines information from the seven subsystems into a single record, which is passed to the online computer farm for conversion into an offline data format, as shown in Figure 22. This data is then processed by the level-3 software trigger to decide whether to record it to permanent storage. The level-3 trigger performs a

simple reconstruction in real-time, using a fast track fitting algorithm that allows the rejection of events which consist of tracks that do not originate from near the IP. This further reduces background events by a factor of two, while keeping the efficiency for hadronic events above 99%.

After events that pass the level-3 trigger are sent to storage, they are processed with basic reconstruction algorithms that provide the first steps for physics analysis, and also apply further selection criteria in the form of the level-4 offline trigger. This data is then stored as mini-DST (MDST) files for use in physics analysis.

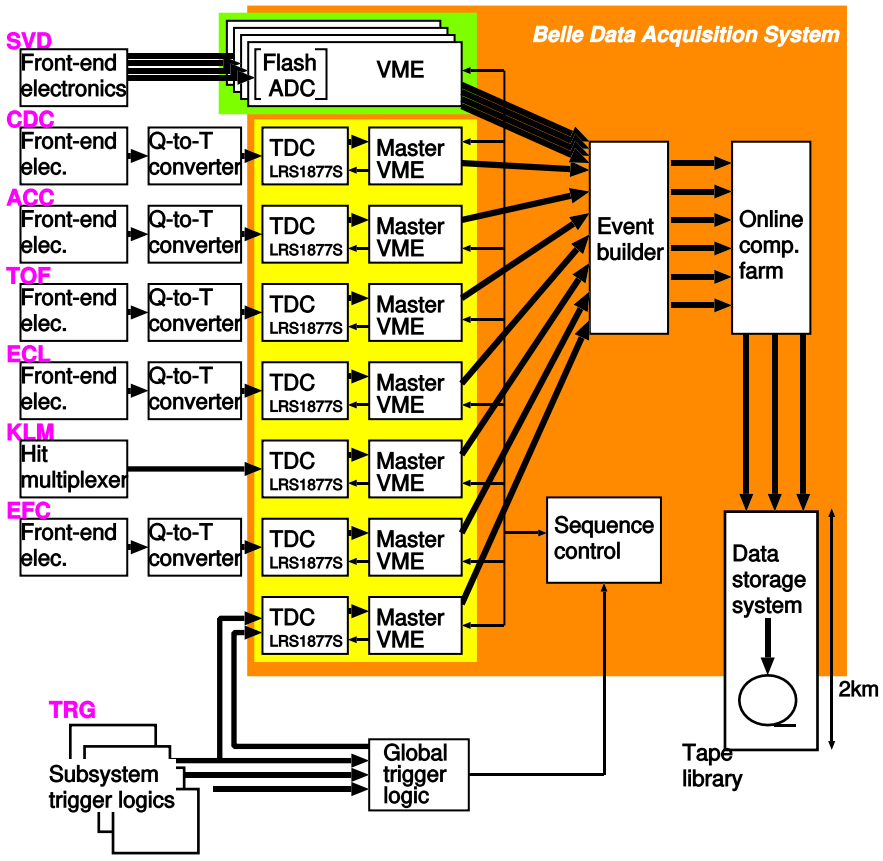


Figure 22: Overview of the dataflow in the DAQ system [32].

2.3 Belle Analysis Framework

The Belle analysis framework (BASF) [47] is the software framework though which the data gathered by the Belle detector is processed. It is used by both the DAQ and offline data processing as described in chapter 2.2.10; Monte Carlo generation; and event reconstruction for individual physics analysis. It is a modular framework written in C++ which allows users to combine their own modules for specific analysis tasks with common modules such as I/O and for interfacing with external software. Modules are written as class objects in C++,

but can also be written in C or Fortran and encapsulated as a class object by a wrapper in BASF. The architecture of BASF is described in Figure 23. The framework also includes a large C++ library for use in writing BASF modules. These include functions for common tasks in event reconstructions, such as vertex fitting, PID algorithms, flavour tagging, and calculation of Kakuno-Super-Fox-Wolfram (KSFW) moments [48] (see chapter 3.3.1).

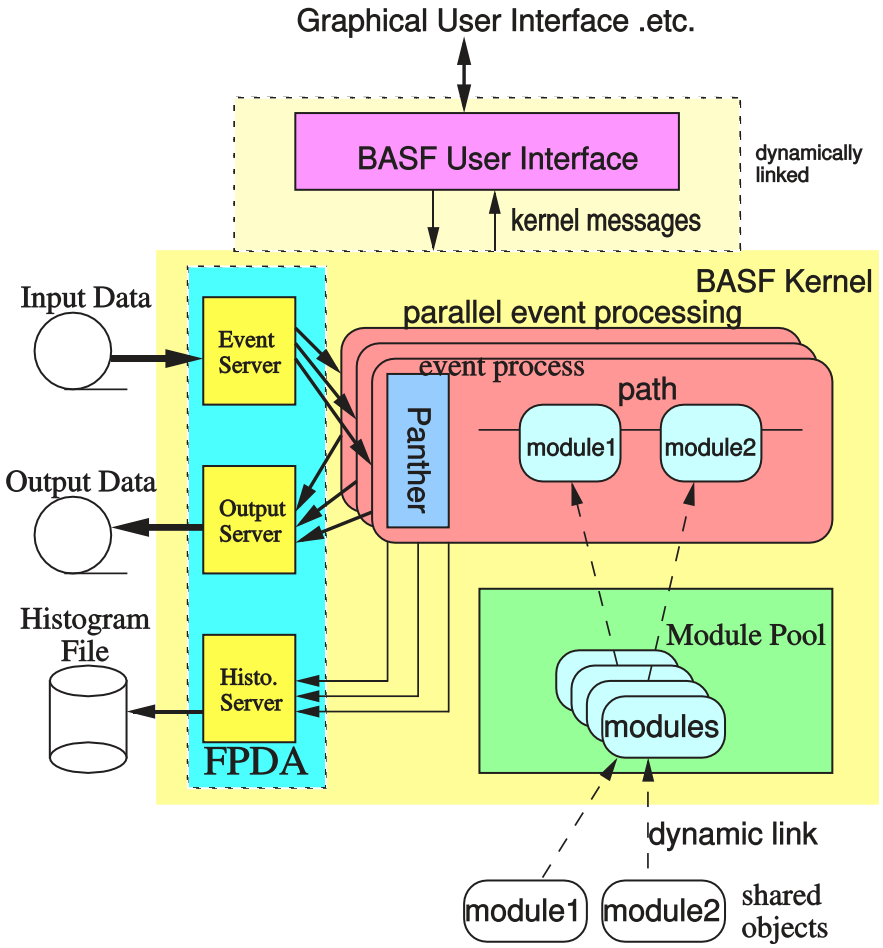


Figure 23: Schematic showing the BASF global architecture [47].

2.3.1 Particle Identification

For this analysis of $B \to \overline{D^0}\pi^0$, the separation of charged kaons and pions is fundamentally important for the reconstruction of the $\overline{D^0}$ from decay daughters. Using information from the subdetectors discussed in the preceding chapters, likelihood functions for each discriminating variable are constructed and combined into a likelihood ratio [29]. For distinguishing between hadrons, information is taken from the ACC, TOF and CDC and combined as

$$L(\alpha:\beta) = \frac{L_\alpha^{CDC} L_\alpha^{TOF} L_\alpha^{ACC}}{L_\alpha^{CDC} L_\alpha^{TOF} L_\alpha^{ACC} + L_\beta^{CDC} L_\beta^{TOF} L_\beta^{ACC}}, \tag{2.10}$$

for distinguishing α particles from β particles, where $L^{detector}$ is the likelihood function for a specific subdetector [49].

The likelihood function for the CDC component comes from the χ^2 of the measured value of $\frac{dE}{dx}$, parametrised as a Gaussian distribution,

$$L_{\alpha}^{CDC} = \frac{e^{-\frac{\chi^2}{2}}}{\sqrt{2\pi}\sigma_{\frac{dE}{dx}}(p)}\,, \tag{2.11}$$

$$\chi^2 = \left(\frac{\left(\frac{dE}{dx}\right)_{measured} - \left(\frac{dE}{dx}\right)_{\alpha(p)}}{\sigma_{\frac{dE}{dx}}(p)} \right)^2, \tag{2.12}$$

where $\left(\frac{dE}{dx}\right)_{measured}$ is the observed value from the CDC, $\left(\frac{dE}{dx}\right)_{\alpha(p)}$ is the expected value for a α particle with momentum p, and $\sigma_{\frac{dE}{dx}}(p)$ is the expected $\frac{dE}{dx}$ resolution of the CDC for a particle of momentum p.

Similarly, the likelihood function for the TOF component is a Gaussian distribution of a χ^2. The χ^2 is constructed as a vector of the difference in observed and expected time for each of the two PMTs in a counter,

$$\Delta = \begin{bmatrix} t_1^{obs} - t_1^{\alpha(p)} \\ t_2^{obs} - t_2^{\alpha(p)} \end{bmatrix}, \tag{2.13}$$

where t_i^{obs} is the time observed by the TOF and $t_i^{\alpha(p)}$ is the expected time for an α particle with momentum p, for the i^{th} PMT ($i \in \{1,2\}$). The error matrix E is constructed from the uncorrelated digital readout uncertainties of each PMT, and the correlated uncertainties due to the effects of tracking resolution. In the case of multiple counter hits the χ^2 values are summed, unless one value is very large, in which case it is ignored. This is combined to give a χ^2 and Gaussian likelihood distribution as

$$L_{\alpha}^{TOF} = \frac{e^{-\frac{\chi^2}{2}}}{\prod_{i=1}^{n}\sqrt{2\pi}\sigma_i}, \tag{2.14}$$

$$\chi^2 = \Delta^T E^{-1} \Delta, \tag{2.15}$$

where n is the number of PMT times included and σ_i is the resolution of each PMT time. It is noted that this formulation ignores the uncertainty correlations for the denominator, as they are quite small.

The ACC measures the number of photon-electrons (N_{pe}) in the photodiode resulting from Cherenkov radiation. While the expected means are quite simple, $N_{pe} = 0$ for kaons and some finite value of N_{pe} for pions, because of background effects the expected distributions are very complex. As such, the likelihood is defined by a binary function based on whether N_{pe} exceeds a certain threshold, N_{pe}^{th}, which is chosen to optimize separation based on a Monte Carlo study,

$$L_{ACC}^{\alpha} = \begin{cases} 1 - \epsilon(p), & N_{pe} < N_{pe}^{th} \\ \epsilon(p), & N_{pe} \geq N_{pe}^{th} \end{cases}, \tag{2.16}$$

where $\epsilon(p)$ is the expected efficiency for a particle of momentum p.

For discriminating between pions and kaons the likelihood ratio

$$L(K:\pi) = \frac{L_K^{CDC} L_K^{TOF} L_K^{ACC}}{L_\pi^{CDC} L_\pi^{TOF} L_\pi^{ACC} + L_K^{CDC} L_K^{TOF} L_K^{ACC}} \tag{2.17}$$

is used, sometimes called KID. A selection can then be made on the value of this to determine the species of a track, typically tracks with $L(K:\pi) > 0.6$ are identified as kaons, otherwise they are identified as pions. The performance of this KID algorithm was studied using the decay $D^{*+} \to D^0 \pi_s^+$ with $D^0 \to K^- \pi^+$ where the s subscript denotes a low momentum particle. This found that using the $L(K:\pi) > 0.6$ criteria stated above, the average $K^\pm$ ($\pi^\pm$) efficiency is around 88% (89%) with a $\pi^\pm$ ($K^\pm$) contamination of 8.5% (8.8%) [49, 50]. Figure 24 shows the distribution of $L(K:\pi)$ for $K^\pm$ and $\pi^\pm$ particles in this study, and Figure 25 shows the dependence of KID on particle momentum.

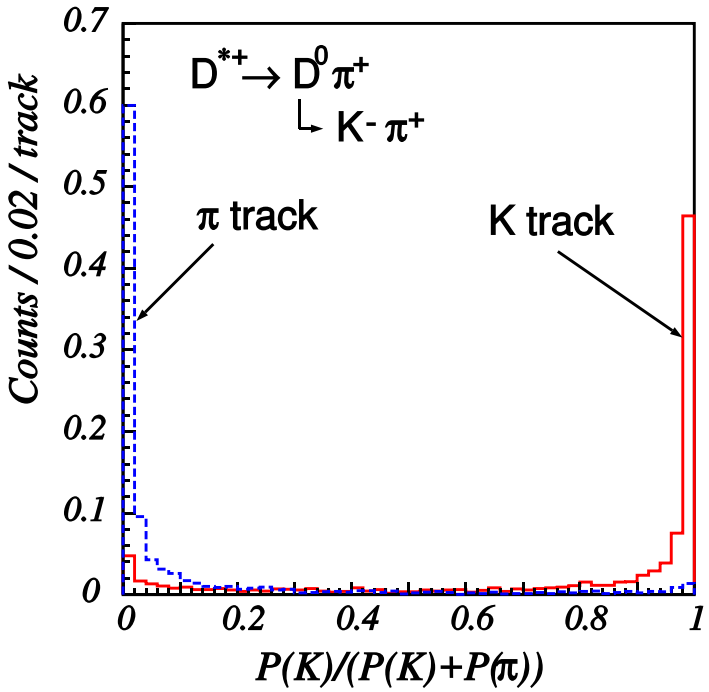
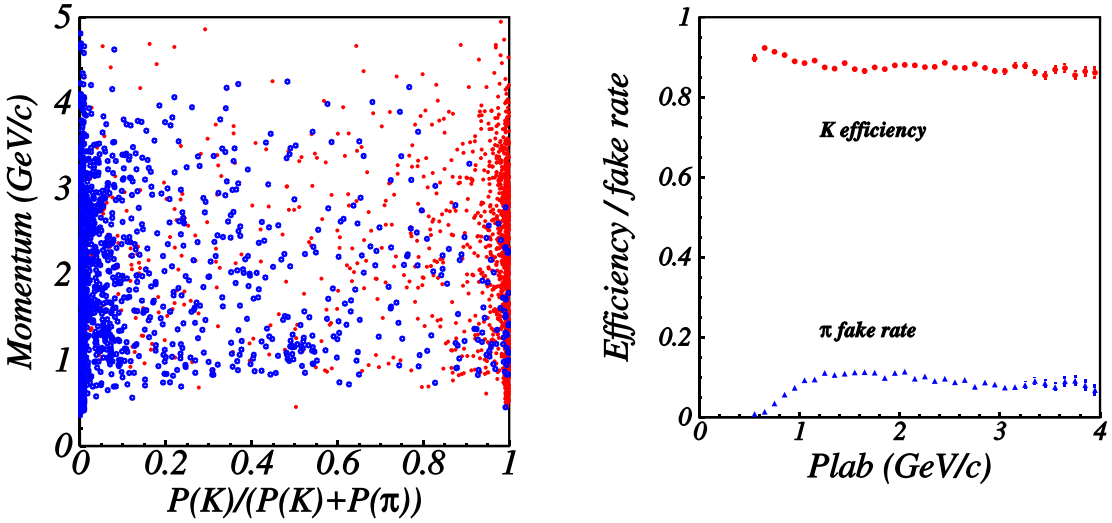


Figure 24: Distribution of $L(K:\pi)$ for charged tracks for D^{*+} tagged $D^0 \to K^-\pi^+$ data sample. Red, solid histogram is for $K^\pm$ particles and blue, dashed is for $\pi^\pm$ [32].



(a) (b)

Figure 25: Performance with respect to momentum for D^{*+} tagged $D^0 \to K^-\pi^+$ data sample. a) shows likelihood distribution for $K^\pm$ (red) and $\pi^\pm$ (blue) vs momentum; b) shows the $K^\pm$ efficiency (red) and $\pi^\pm$ contamination (blue) as a function of momentum for the criteria $L(K:\pi) > 0.6$ [32].

3 $B^0 \to \overline{D^0}\pi^0$ Analysis Strategy

This chapter describes the process by which the $B^0 \to \overline{D^0}\pi^0$ decay is reconstructed, backgrounds supressed, and data is fitted to extract the branching fraction and A_{CP}. This is a blind analysis where all methods are studied and verified using Monte Carlo (MC) simulated samples before the finalised methodology is applied to real experimental data. These simulated samples and verification steps are also described.

The particle decay $B^0 \to \overline{D^0}\pi^0$ is studied in two D^0 decay channels: $\overline{D^0} \to K^+\pi^-$ and $\overline{D^0} \to K^+\pi^-\pi^0$. These modes allow for the flavour of the D^0 to be approximately determined based off the charge of the kaon (see section 3.1 for corrections due to this approximation), do not contain a large number of final state particles and have large branching fractions ($\mathfrak{B}$). The current world average branching fractions for these decays are as follows [3]:

$$\mathfrak{B}\left(B^0 \to \overline{D^0}\pi^0\right) = (2.63 \pm 0.14) \times 10^{-4}\,, \tag{3.1}$$

$$\mathfrak{B}(D^0 \to K^-\pi^+) = (3.95 \pm 0.03) \times 10^{-2}\,, \tag{3.2}$$

$$\mathfrak{B}(D^0 \to K^-\pi^+\pi^0) = (1.44 \pm 0.05) \times 10^{-1}\,. \tag{3.3}$$

From these branching fractions and our expected reconstruction efficiencies, 5000 events are expected in the data sample collected by the Belle experiment.

The analysis procedure to extract A_{CP} and $\mathfrak{B}$ involves the following steps:

- Identify charged tracks as pions or kaons using the Belle Particle Identification systems.
- Identify ECL clusters as candidate γ's from π^0 decays.
- Reconstruct $\pi^0 \to \gamma\gamma$ decays.
- Reconstruct signal events in the two decay channels using BASF (Belle Analysis Framework).
- Perform selections on mass and momentum of particles to reduce the number of background events.
- Use the NeuroBayes [51] software package to define a continuum suppression variable (C_{NN}) and perform a selection to remove a significant amount of continuum background.
- Define probability distribution functions (PDFs) of the fitting variables for background and signal events.
- Perform a three-dimensional simultaneous fit to extract signal yield and A_{CP}.

3.1 Doubly Cabibbo Suppressed Decays

It is noted that determining the D^0 flavour using the kaon charge is an approximation, as the doubly Cabibbo suppressed (DCS) decays $\overline{B^0} \to \overline{D^0}\pi^0$, $D^0 \to K^+\pi^-$ and $D^0 \to K^+\pi^-\pi^0$ will result in the determined flavour being incorrect. These effects can be calculated using the ratio of wrong sign to right sign decay rates:

$$R = \frac{\mathfrak{B}_{\mathrm{WS}}}{\mathfrak{B}_{\mathrm{RS}}}. \tag{3.4}$$

$R\left(\overline{B^0} \to \overline{D^0}\pi^0\right)$ has not been measured so is approximated with the value for the colour flavoured decay,

$$R\left(\overline{B^0} \to D^+\pi^-\right) = (2.92 \pm 0.38 \pm 0.31) \times 10^{-4}\ [52]. \tag{3.5}$$

For the D^0 decays the PDG average values [3] are used:

$$R(D^0 \to K^+\pi^-) = (3.45 \pm 0.06) \times 10^{-3} \tag{3.6}$$

and

$$R(D^0 \to K^+\pi^-\pi^0) = (2.12 \pm 0.07) \times 10^{-3}. \tag{3.7}$$

The addition of DCS events will lead to an overestimate of the branching fraction of the Cabibbo favoured mode being measured. The correction factor to account for this is calculated as:

$$\sigma_{\mathfrak{B}} = \frac{1}{1 + R\left(\overline{B^0}\right) + R(D^0)} - 1, \tag{3.8}$$

where $R\left(\overline{B^0}\right)$ is from equation 3.5 and $R(D^0)$ is the average of equation 3.6 and 3.7, weighted by the expected number of events from each D^0 decay mode (see section 3.5). This gives a small correction to the observed value of $\mathfrak{B}$ of $(-0.313 \pm 0.007)\%$ which is taken into account in the final branching fraction.

For A_{CP}, imperfect flavour identification will lead to an underestimation compared to the true value. The correction factor for A_{CP} is calculated as:

$$\sigma_{A_{CP}} = 1 - \frac{1 - R\left(\overline{B^0}\right) - R(D^0)}{1 + R\left(\overline{B^0}\right) + R(D^0)}. \tag{3.9}$$

This gives a correction factor of $(+0.62 \pm 0.01)\%$. Because this is multiplicative with an already small value of A_{CP}, the absolute correction of $(+6.4 \pm 0.1) \times 10^{-6}$ is significantly smaller (by $\mathcal{O}(10^3)$) than the uncertainty on A_{CP}, so this correction is neglected.

3.2 Event Selection

3.2.1 Data Samples

This analysis uses the full Belle data sample of 711 fb^{-1}, corresponding to $(771.581 \pm 10.566) \times 10^6$ $B\overline{B}$ pairs gathered, at the KEKB e^+e^- collider at $\Upsilon(4S)$ resonance. This data is categorised into a series of experiments, defined as the periods between major shutdowns. All data in one experiment shares similar collider and detector conditions.

For the study of signal events, signal MC data sets were generating using EvtGen [53] and GSIM [54] containing 1 million $B\overline{B}$ events, with one B decaying as signal and the other decays as expected from previously measured branching fractions (generic decays). These events are generated simulating the conditions of each of the experiments, with the amount of events generated being proportional to the luminosity of each experiment. Four of these datasets are generated, two for each D^0 decay mode.

The EvtGen decay file is set to have an $\Upsilon(4S)$ decay into a $B^0\overline{B^0}$ pair, with one of these B mesons decaying generically, while the other signal B decays into $\overline{D^0}\pi^0/D^0\pi^0$ with even chance of either B^0 being the signal B. The $\overline{D^0}$ then decays into either $K^+\pi^-$ or $K^+\pi^-\pi^0$ depending on what mode is being studied.

Background studies used the official Belle MC, consisting of 6 streams of continuum events ($e^+e^- \to q\overline{q}$, $q = u, d, s, c$), 10 streams of generic $B\overline{B}$ events ($e^+e^- \to \Upsilon(4S) \to B^0\overline{B^0}/B^+B^-$ with $b \to c$ transitions) and 50 streams of charmless $B\overline{B}$ MC ($e^+e^- \to \Upsilon(4S) \to B^0\overline{B^0}/B^+B^-$ with one B undergoing charmless $b \to u$ transitions). A stream consists of the equivalent number of events to what was recorded during the full Belle run (711 fb^{-1}), with experimental conditions and luminosity proportions the same as the actual experiment.

3.2.2 Event Reconstruction

A list of final state particles ($\pi^\pm, K^\pm, \pi^0$) is made using information provided by both the hadron identification system (CDC, ACC, ToF) and tracking system (CDC and SVD). Charged particles are identified as either $\pi^\pm$ or $K^\pm$ based on the atc_pid (ACC, TOF, CDC Particle IDentification) algorithm. Charged particles with a $L(K:\pi)$ of less than 0.4 are identified as $\pi^\pm$'s and those over 0.6 as $K^\pm$'s, with

$$L(K:\pi) = \frac{L(K)}{L(K) + L(\pi)}, \tag{3.10}$$

where

$$L(X) = L_{ACC}(X) \times L_{TOF}(X) \times L_{CDC}(X) \tag{3.11}$$

is the likelihood of a given particle being of type X calculated by combining information from each identification subsystem, as discussed in detail in chapter 2.3.1.

π^0s are reconstructed through the $\pi^0 \to \gamma\gamma$ decay. First γ's are identified from clusters in the ECL that have no corresponding charged track and a deposited energy of $< 50\ (100)\ MeV$ in the barrel (endcap) region of the ECL. Pairs of γ's with invariant mass near that of the π^0 are added to the π^0 list and their momentum is calculated using a mass constrained fit. Additional gamma candidates are found from e^+e^- pairs that result from γ conversion in the inner detector [55].

D^0 candidates are reconstructed by any combination of $\pi^{\pm}$+$K^{\mp}$ or $\pi^{\pm}$+$K^{\mp}$+π^0 which have a combined invariant mass of between $1.80 GeV$ and $1.92 GeV$. B^0 candidates are then reconstructed by combining D^0 candidates with π^0s, taking care that the same π^0 is not taken as the daughter of both the B^0 and D^0.

The two primary kinematic variables used to constrain the reconstructed B^0 are beam constrained mass:

$$M_{BC} \equiv \sqrt{E_{Beam}^2 - p_B^2}, \tag{3.12}$$

and the energy difference:

$$\Delta E \equiv E_B - E_{Beam}, \tag{3.13}$$

where E_{Beam} is the centre of mass frame beam energy, E_B is the energy of the reconstructed B^0 and p_B is the momentum of the reconstructed B^0. A B^0 candidate is required to have $-0.2\ GeV < \Delta E < 0.2\ GeV$ and $5.255 GeV < M_{BC} < 5.288\ GeV$.

3.2.3 π^0 Energy Correction

Due to energy leakage in the ECL, the reconstructed π^0 energy will be lower than in reality. To compensate for this the reconstructed π^0 momentum is rescaled, taking $E_{\pi^0} = E_{beam} - E_{D^0}$, resulting in:

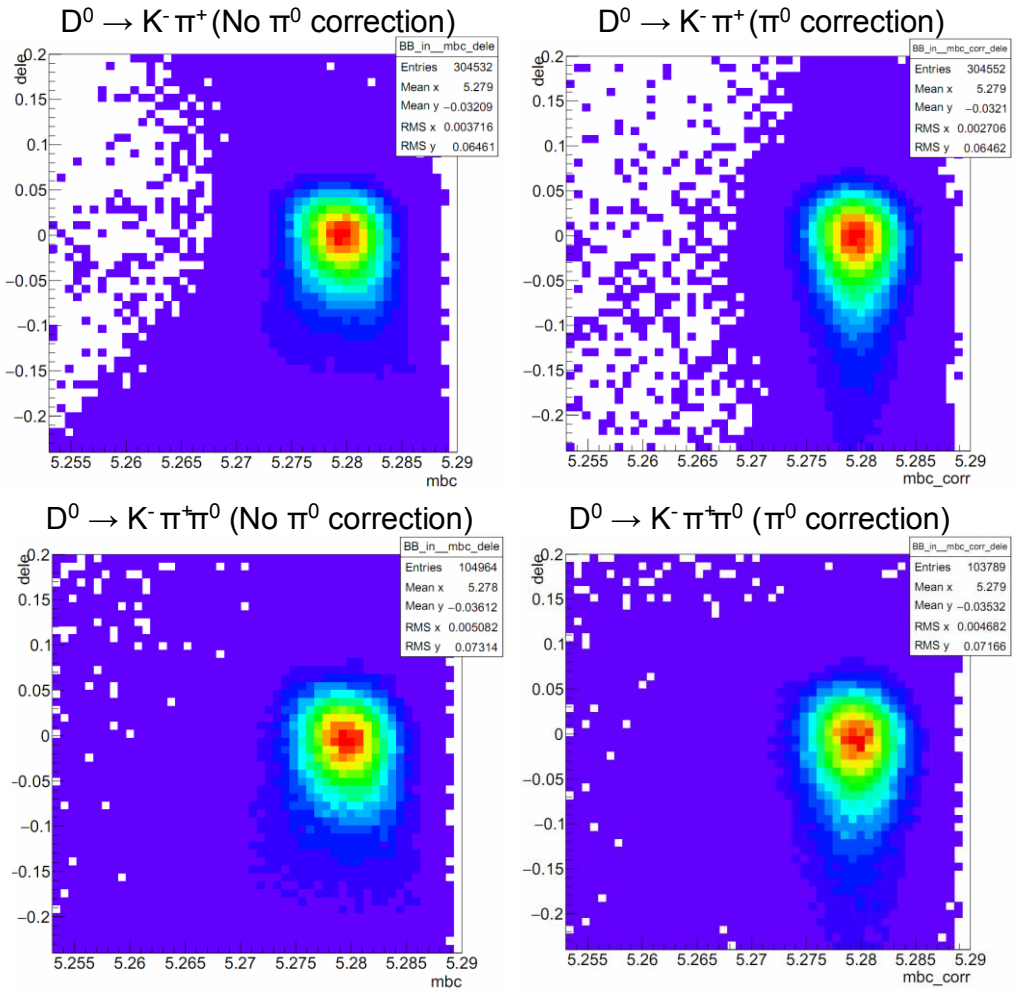


Figure 26: 2D histogram of $\boldsymbol{M_{BC}}$ vs $\boldsymbol{\Delta E}$ with (right) and without (left) $\boldsymbol{\pi^0}$ energy correction for the two $\boldsymbol{D^0}$ decay channels.

$$\kappa \equiv \frac{\sqrt{(E_{beam} - E_{D^0})^2 - M_\pi^2}}{|p_\pi|} \tag{3.14}$$

$$p_{\pi,corrected} \equiv \kappa \times p_\pi \tag{3.15}$$

This $p_{\pi,corrected}$ is then used to calculate a corrected M_{BC} (from now on simply referred to as M_{BC}). This momentum rescaling only affects M_{BC}, not ΔE, with the main aim to de-correlate the two variables. This procedure can only be applied to the π^0 that is the daughter of the B^0. The effect of this on the $M_{BC} \times \Delta E$ distribution is shown in Figure 26.

3.2.4 Best Candidate Selection

The expected number of events with 2 signal B decays is much less than 0.1, so any reconstructed events with more than one B^0 candidate are assumed to be due to mis-reconstruction rather than true double signal decays. In these cases, one of

the reconstructed B^0s is defined as the best candidate, which is taken as the signal B^0. This is done using

$$\Delta m(X) = m_{PDG}(X) - m(X), \tag{3.16}$$

where $m_{PDG}(X)$ is the PDG average mass for particle X and $m(X)$ is the reconstructed mass of particle X.

The best candidate is selected as the B^0 candidate which minimises $\Delta m(D^0)$. If there are multiple candidates with the same D^0, the candidate among these which minimises $\Delta m(\pi^0)$ is selected as the best candidate.

To check how well this best candidate selection is functioning, the reconstructed event is compared with the true decay chain from MC data. Because the π^0s in the final state are reconstructed from photon ECL showers, this is not straightforward, as there may be many interactions involving photons coming from a given π^0. A function to traceback the selected photons though the ECL cascade to their parent particle was written, which validates that both photons originate from the same source and that the source is a π^0 according to the genhep data. This function is not perfect however, and in a significant number of cases (around 25% of $\overline{D^0} \to K^+\pi^-\pi^0$ MC events) the required ECL data is unable to be retrieved, so full separation of self-cross feed (wrongly reconstructed signal events where some of the particles originate from the non-signal B) and true signal is not possible. The events which can be fully truth checked can still be studied to calculate the efficiency of the best candidate selection. Comparing the B^0 and D^0 multiplicities shows that multiple D^0 candidates are the predominant cause of multiple candidates. The increase in multiplicity in the $\overline{D^0} \to K^+\pi^-\pi^0$ mode is thus expected as a result of the combinatorics of the 3-body D^0 decay.

Table 3: Best candidate selection efficiency.

	$\overline{D^0} \to K^+\pi^-$	$\overline{D^0} \to K^+\pi^-\pi^0$
Average B^0 candidate multiplicity	1.304	2.185
Average D^0 candidate multiplicity	1.231	2.069
Percentage of events with multiple B^0 candidates	16%	47%
Percentage of events where correct B^0 candidate is chosen	96%	86%

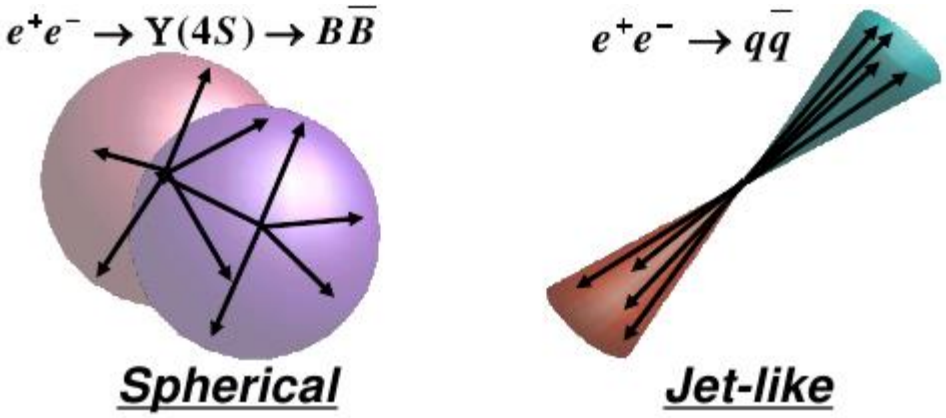


Figure 27: Different event shapes for continuum and $B\overline{B}$ events [56].

3.3 Continuum Suppression

Continuum events ($e^+e^- \to q\overline{q}$) are the dominant source of background in this analysis due to $e^+e^- \to q\overline{q}$ decays having 3 times the cross-section of $e^+e^- \to \Upsilon(4S)$, as well as the large number of light hadrons present in the final state of continuum events. In order to discriminate between continuum and $B\overline{B}$ events the differences in event topology are exploited. As the B meson mass is very close to half that of the $\Upsilon(4S)$, the B meson pair from a $\Upsilon(4S) \to B\overline{B}$ decay will have very low momentum, leading to a spherical distribution of final state particles. In contrast, continuum events have $q\overline{q}$ pairs produced with very high momentum, which results in a jet-like distribution. This is illustrated in Figure 27. In order to quantify the event shape into a discriminator, a number of event shape variables are measured and then combined in a multivariate analysis to give a single continuum suppression variable. The shape variables and multivariate analysis technique is discussed in this section.

3.3.1 KSFW

The most important variables used in differentiating continuum $q\overline{q}$ events and $B\overline{B}$ events are the KSFW (Kakuno-Super-Fox-Wolfram) moments [48].

The original Fox-Wolfram moment (R^2) is calculated as:

$$R_2 = \frac{\sum_{i,j}|p_i||p_j|P_2(\cos\theta_{i,j})}{\sum_{i,j}|p_i||p_j|}, \tag{3.17}$$

where i, j are final state photons and charged particles, p is momentum and

$$P_2(x) = \frac{1}{2}(3x^2 - 1) \tag{3.18}$$

is the 2nd order Legendre polynomial [48]. This quantity is defined in such a way that jet-like events (eg. $q\overline{q}$) will have an R_2 close to 1, and spherical events (eg. $B\overline{B}$) will have an R_2 close to 0. While this seems ideal, Monte Carlo testing shows that the peaks are quite far away from the ideal 0 and 1 for $B\overline{B}$ and $q\overline{q}$ events respectfully.

To improve the discriminating power, rather than using a single variable, multiple KSFW variables are combined into a Fisher discriminant [57] which can be optimised for distinguishing signal and background for a particular decay. The discrimination is maximised by calculating moments separately for signal-tag side particles pairs, as well as tag-tag side pairs (tag side refers to particles originating from the non-signal B), and also by considering charged, neutral and missing energy separately. The KSFW Fisher discriminant is defined as

$$KSFW = \sum_{j=C,N,m}\sum_{i=0}^{4}\alpha_{i,j}R_{i,j}^{SO} + \sum_{i=0}^{4}\beta_i R_i^{OO} + \gamma\sum|p_t|, \qquad (3.19)$$

where $\sum|p_t|$ is the scalar sum of all transverse momenta, $\alpha_{i,j}$, β_i and γ are Fisher coefficients. The KSFW moments are defined as:

$$R_{i,j}^{SO} = \frac{\sum_{k,l} Q_k Q_l |p_l| P_i(cos\theta_{kl})}{E_B - \Delta E}, \qquad \forall i \in \{1,3\} \qquad (3.20)$$

$$R_{i,j}^{SO} = \frac{\sum_{k,l} |p_l| P_i(\cos\theta_{kl})}{E_B - \Delta E}, \qquad \forall i \in \{0,2,4\} \qquad (3.21)$$

where $j = C, N, M$ refers to whether particles are charged, neutral or missing energy, k is all signal tracks of type j, l is all non-signal tracks of type j, and Q is the charge of the particle; and

$$R_i^{OO} = \frac{\sum_{k,l} Q_k Q_l |p_k||p_l| P_i(\cos\theta_{kl})}{E_B - \Delta E}, \qquad \forall i \in \{1,3\} \qquad (3.22)$$

$$R_i^{OO} = \frac{\sum_{k,l} |p_k||p_l| P_i(\cos\theta_{kl})}{E_B - \Delta E}, \qquad \forall i \in \{0,2,4\} \qquad (3.23)$$

where k, l are all non-signal B tracks [58].

There is a large correlation between these moments and the missing mass of the event (mm^2), so moments are calculated separately in 7 different mm^2 bins. This gives 17 variables in each of the 7 mm^2 bins, which are combined with the corresponding Fisher coefficients in a Fisher discriminant for each bin. The optimal values for these coefficients to best separate signal and $q\overline{q}$ background are then found by inputting one signal MC dataset and one stream of continuum MC into a training algorithm.

When an event's kinematic variables are input into the KSFW equation with the optimised Fisher coefficients, it will give a likelihood value between 0 and 1, with 0 indicating continuum-like shape, and 1 indicating signal-like shape. Because there may be slightly different kinematics between the two D^0 decay channels, 2 sets of Fisher coefficients are trained and used depending on which channel is employed in the signal reconstruction.

3.3.1 Additional Shape Variables

The other shape variables used are described below, with their distribution for $\overline{D^0} \to K^+\pi^-$ and $\overline{D^0} \to K^+\pi^-\pi^0$ modes shown in Figure 28 and Figure 29 respectively:

- $q \cdot r$:
  This is the flavour multiplied dilution factor output by the HAMLET (HAdrons Merged with LEptons for Tagging) module [59]. This module uses information from all particles not involved in the signal decay chain to determine the flavour of the B particle that does not undergo the signal decay (the tag-side B). It provides two variables as outputs, q, which is an integer indicating the flavour of the B according to the algorithm (+1 for B and -1 for $\overline{B}$), and r, the quality of the tagging by the algorithm ranging from 0 (entirely uncertain) to 1 (entirely certain). $q\overline{q}$ events are expected to have a $q \cdot r$ of close to 0 as there is no B to have flavour, while $B\overline{B}$ events are likely to peak near ± 1.
- $|\cos\theta_B|$, the B flight direction:
  This is the angle between the momentum of the signal B candidate and the e^+e^- beam line in the centre of mass frame. For $B\overline{B}$ events a $1 - \cos^2\theta_B$ distribution is expected, due to the $\Upsilon(4S) \to B\overline{B}$ decay being a vector particle decaying into two scalar particles. For $e^+e^- \to q\overline{q}$ events, a uniform distribution is expected.
- $|\cos\theta_{thr}|$, the thrust angle:
  This is the cosine of the angle between the thrust axis of the signal B candidate and the thrust axis of all other final state particles not involved in the signal decay. The thrust axis is the axis along which the longitudinal momenta of the final state particles defined as

$$T = \max_{\hat{x}} \frac{\sum_i |\overrightarrow{p_i} \cdot \hat{x}|}{\sum_i |\overrightarrow{p_i}|}, \tag{3.24}$$

  where i is summed over the final state particles, $\overrightarrow{p_i}$ is the momentum of the particle, T is the thrust and $\hat{x}$ is the thrust axis. Because of the spherical topology of $B\overline{B}$ events the signal distribution is expected to be uniform, while the jet-like nature of continuum events leads to a distribution peaking sharply at 1.

- Δz:
  The distance along the z-axis between the vertex of signal final state particles and vertex of charged tracks not reconstructed in the signal event. Due to the lifetime of the B meson, this distance is in general greater for signal events than continuum.

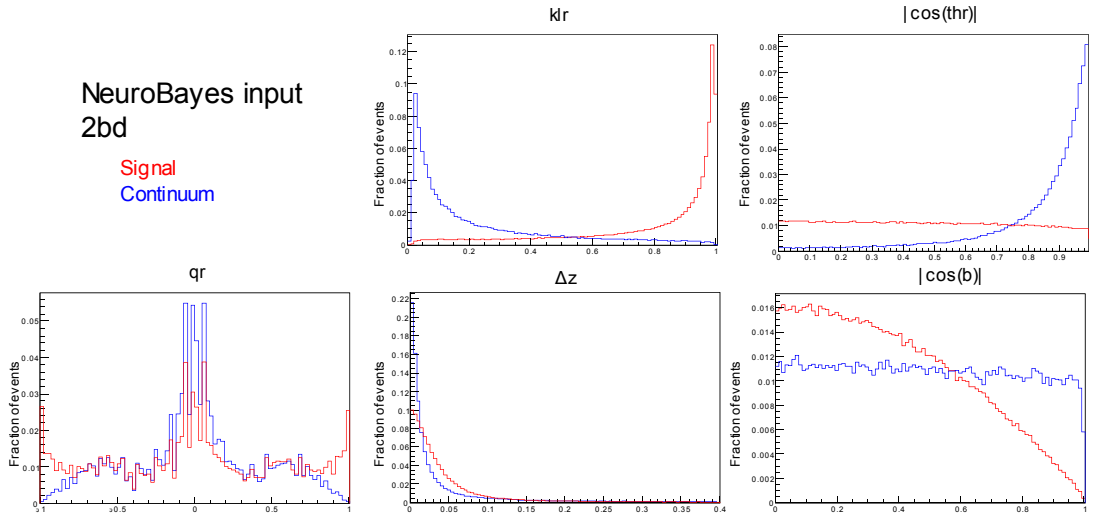


Figure 28: NeuroBayes input parameters for $\overline{D^0} \to K^+\pi^-$.

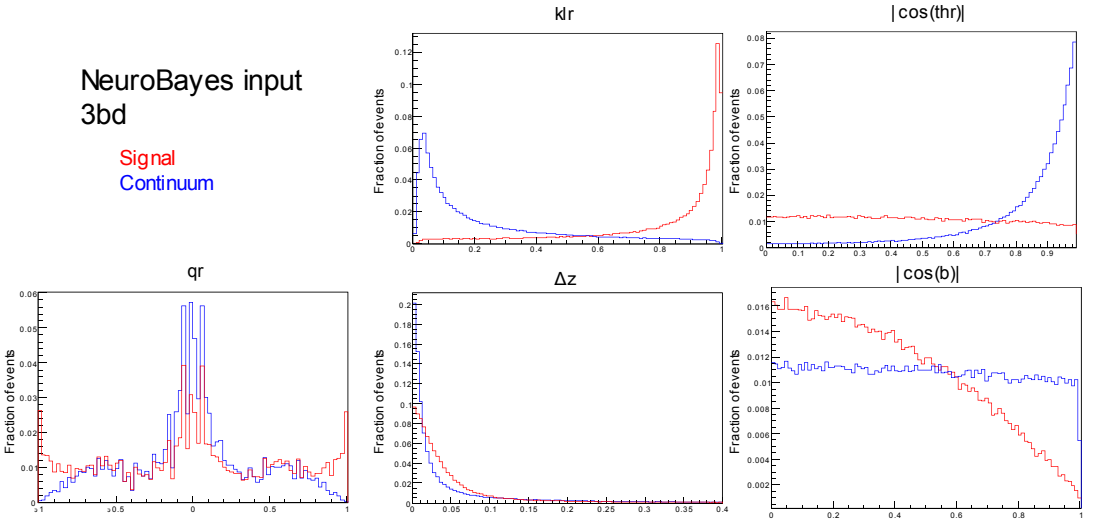


Figure 29: NeuroBayes input parameters for $\overline{D^0} \to K^+\pi^-\pi^0$.

3.3.2 NeuroBayes

NeuroBayes [51] is a software package that combines a neural network, variable pre-processing and other tools for multivariate analysis. The continuum suppression variable output by NeuroBayes (C_{NN}) has a range of (-1,1) with 1 being most signal like and -1 being most background like.

NeuroBayes training was conducted with a set of 100,000 signal MC and continuum events, separately for each decay channel. After the NeuroBayes expert was trained, it was verified by running over a different set of 1 million MC events and a full stream of continuum background, the output of which is shown in Figure 30. In order to maximise statistics, rather than placing an optimized selection requirement on C_{NN}, only loose selection is made, and a related variable is used in the fitting procedure later. The requirement of $C_{NN} > -0.05$ is used for both the $\overline{D^0} \to K^+\pi^-$ and $\overline{D^0} \to K^+\pi^-\pi^0$ channel, which results in 86% background reduction and 87% signal efficiency. To verify that the NeuroBayes network is not overtrained, the network is applied to an independent dataset of signal and continuum, and the efficiency is found to be the same (Figure 30). The C_{NN} distribution shown in Figure 31 is also the same for both datasets.

In order to allow the continuum suppression variable to be fitted with well-behaved Gaussian PDFs, C_{NN} is transformed into an appropriate shape using the formula:

$$C'_{NN} = \log\left(\frac{C_{NN} - C_{NN}^{cut}}{C_{NN}^{\max} - C_{NN}}\right), \tag{3.25}$$

where C_{NN}^{cut} is the selection requirement discussed previously (-0.05), and $C_{NN}^{\max}$ is the max value of C_{NN} obtained in the signal MC set used to verify the training. $C_{NN}^{\max}$ is used instead of 1 as the resulting distribution shape is more Gaussian. The distribution of C'_{NN} for the verification MC sets are shown in Figure 32.

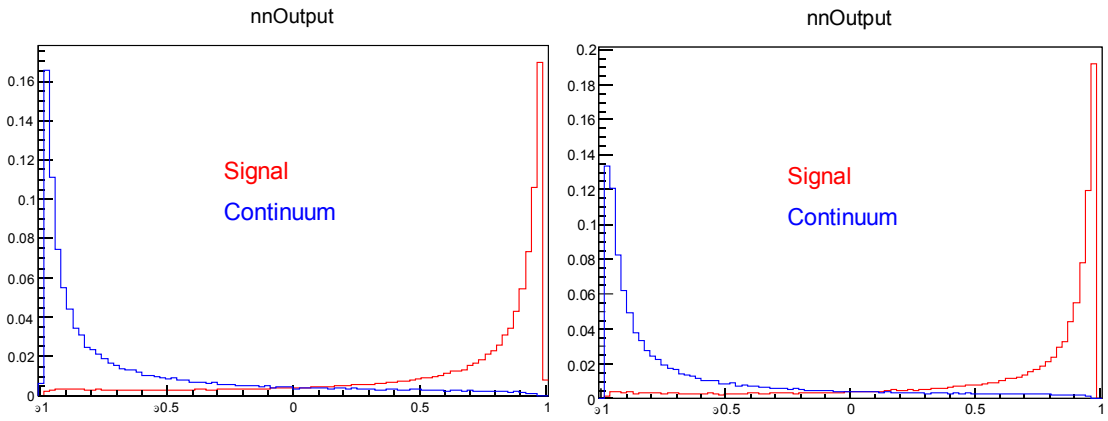


Figure 30: NeuroBayes output for verification datasets for $\overline{D^0} \to K^+\pi^-$ (Left) and $\overline{D^0} \to K^+\pi^-\pi^0$ (Right).

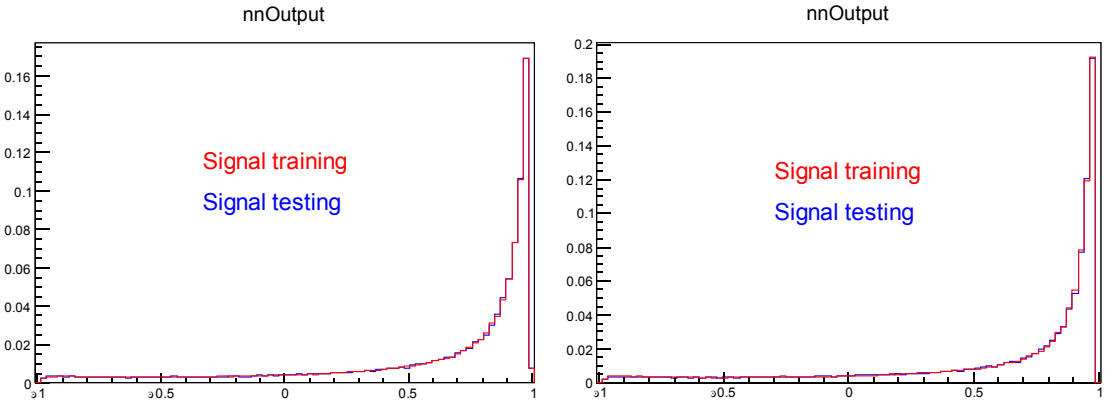


Figure 31: Neurobayes output comparing training and testing datasets for $\overline{D^0} \to K^+\pi^-$ (Left) and $\overline{D^0} \to K^+\pi^-\pi^0$ (Right).

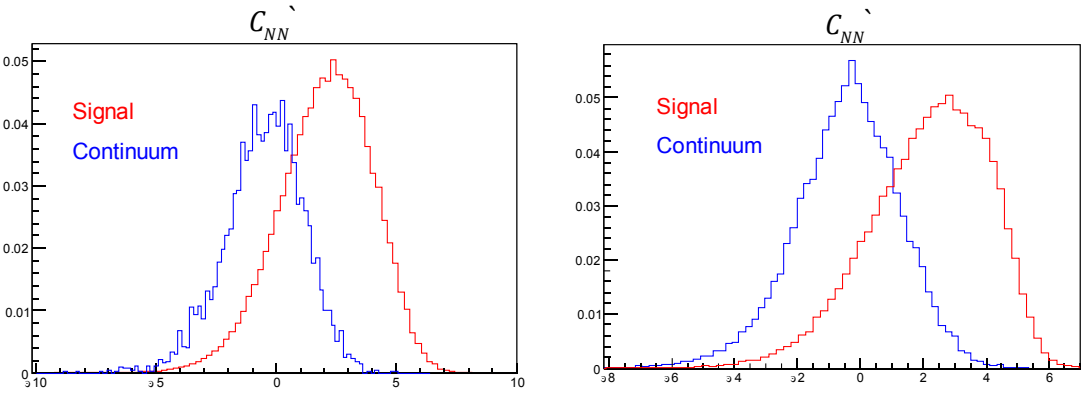


Figure 32: NeuroBayes transform, C'_{NN} for verification datasets for $\overline{D^0} \to K^+\pi^-$ (Left) and $\overline{D^0} \to K^+\pi^-\pi^0$ (Right).

3.4 $B\overline{B}$ Background

To reduce the amount of background from generic $B\overline{B}$ decays the following requirements are placed on the mass of the π^0s, M_{BC}, ΔE, final state particle momentum and the reconstructed D^0 mass:

Table 4: Selection criteria to reduce $B\overline{B}$ background.

Variable	Accepted Range
M_{BC}	5.253 – 5.288 GeV
ΔE	-0.2 – 0.2 GeV
$M(D^0)$	1.841 – 1.882 GeV
$M(\pi^0)(B^0\ daughter, K^+\pi^-)$	104.07 – 163.1 MeV
$M(\pi^0)(B^0\ daughter, K^+\pi^-\pi^0)$	105.77 – 164.37 MeV
$M(\pi^0)(D^0\ daughter, K^+\pi^-\pi^0)$	107.88 – 162.264 MeV
$p(K^\pm)$	0.3 – 3.5 GeV
$p(\pi^\pm)$	0.3 – 3.5 GeV
$p(\pi^0_{B^0})$	1.5 – 3.5 GeV
$p(\pi^0_{D^0})$	0.2 – 3.5 GeV

The D^0 mass selection criteria is based on $\pm 3\sigma$ of the $\overline{D^0} \to K^+\pi^-$ D^0 mass distribution in signal MC. The $\overline{D^0} \to K^+\pi^-$ selection criteria is used for both D^0 modes as the $\overline{D^0} \to K^+\pi^-\pi^0$ mass distribution is quite wide and a $\pm 3\sigma$ around this did not produce sufficient background reductions. For the π^0 mass selection criteria, they are also based on $\pm 3\sigma$ of the mass distribution of each π^0 separately.

The main components of the generic $B\overline{B}$ background are: combinatorial backgrounds, $B^0 \to \overline{D^0}\rho^0$ and $B^0 \to \overline{D^0}^*\left(\overline{D^0}[\gamma|\pi^0]\right)\pi^0$. The charmless $B\overline{B}$ background component is small and consists mainly of combinatorial background with some non-resonant B decays to the same final state.

3.5 Efficiencies

After all selections have been made, the reconstruction efficiency is calculated as $(27.53 \pm 0.04)\%$ for the $\overline{D^0} \to K^+\pi^-$ mode (1.07% of overall $B^0 \to \overline{D^0}\pi^0$ decays) and $(9.43 \pm 0.02)\%$ for the $\overline{D^0} \to K^+\pi^-\pi^0$ mode (1.34% of overall $B^0 \to \overline{D^0}\pi^0$ decays). The expected number of events for signal and each background type in the full Belle dataset of 711 fb^{-1} is shown in Table 5, with 4508 total signal events expected. This compares with the previous Belle analysis [15] using a $140fb^{-1}$ of data that found ~620 events.

Table 5: Expected number of events for $B^0 \to \overline{D^0}\pi^0$ reconstruction based on MC efficiencies.

	Expected Events $(\overline{D^0} \to K^+\pi^-)$	Expected Events $(\overline{D^0} \to K^+\pi^-\pi^0)$
Signal	2037	2471
Continuum	4322	19279
Generic $B\overline{B}$	6368	10675
Charmless $B\overline{B}$	235	474

3.6 Fitting

This section describes the process by which the yield (from which the branching fraction is calculated) and A_{CP} of the signal decay are extracted from the Belle dataset (or MC dataset used in validation). To do this, first probability distribution functions (PDFs) are fitted to MC data for signal and backgrounds (continuum, generic $B\overline{B}$ decays, charmless $B\overline{B}$ decays) separately for each D^0 decay channel for M_{BC}, ΔE, and C'_{NN}. These PDFs are then used in an unbinned maximum likelihood fit, with the signal yield and A_{CP} as parameters.

In order to extract the signal yield and A_{CP}, a simultaneous three-dimensional unbinned extended maximum likelihood fit is performed in M_{BC}, ΔE and C'_{NN} using the RooFit toolkit [60] in the ROOT [61] C++ framework. First, probability distribution functions (PDFs) are fitted to MC data for signal and backgrounds (continuum, generic $B\overline{B}$ decays, charmless $B\overline{B}$ decays) separately for each decay channel. Then, to extract the signal yield and A_{CP} from real data (or in the validation tests, a full stream of MC signal and background) these PDFs are fitted to data with floating normalisation constants for signal, generic $B\overline{B}$ and continuum backgrounds (charmless $B\overline{B}$ decay yield is fixed), and floating values for the continuum M_{BC} and ΔE PDF parameters. The region over which data is fitted is defined as:

$$5.255GeV < M_{BC} < 5.29GeV, \tag{3.26}$$

$$-0.2GeV < \Delta E < 0.2GeV, \tag{3.27}$$

$$-10 < C'_{NN} < 10. \tag{3.28}$$

3.6.1 PDF Shapes

Where possible, the 3-dimensional PDFs are modelled as the product of 1-dimensional PDFs:

$$\mathcal{P}(M_{BC}, \Delta E, C'_{NN}) = \mathcal{P}(M_{BC}) \times \mathcal{P}(\Delta E) \times \mathcal{P}(C'_{NN}), \tag{3.29}$$

However, there was found to be correlations between M_{BC} and ΔE in some event types that made this impossible, requiring

$$\mathcal{P}(M_{BC}, \Delta E, C'_{NN}) = \mathcal{P}(M_{BC}, \Delta E) \times \mathcal{P}(C'_{NN}), \quad (3.30)$$

to be used in certain cases. Attempts were made to fit 2D analytic PDFs to these distributions, but none could be found that modelled them satisfactorily. Instead PDFs were generated by modelling the MC data with a kernel density estimation (KEST) PDF. Details of this are discussed in chapter 3.6.2.

The following shapes were used for the M_{bc} and ΔE PDFs for the various event types:

- Signal: For the $\overline{D^0} \to K^+\pi^-$ mode, the M_{BC} distribution is modelled with a Crystal Ball function [62] while the ΔE distribution is modelled with the sum of a Crystal Ball function and a Gaussian. For the $\overline{D^0} \to K^+\pi^-\pi^0$ mode, there is still a strong correlation between M_{BC} and ΔE and no analytic 2D pdf could be found to fit the MC data. Instead a 2D Kernel Density Estimation PDF is used, generated from one of the signal MC samples.
- Generic $B\overline{B}$: Similarly to $\overline{D^0} \to K^+\pi^-\pi^0$ signal, there exist complex correlations between M_{BC} and ΔE in both modes, so a 2D Kernel Density Estimation PDF was generated from five streams of generic $B\overline{B}$ MC.
- Continuum: The M_{BC} distribution is fitted as an Argus function [63] and the ΔE distribution as a 1st order Chebyshev polynomial in both modes.
- Charmless $B\overline{B}$: As with generic $B\overline{B}$ and signal, charmless $B\overline{B}$ was modelled with a Kernel Density Estimation PDF in both modes, generated using 50 streams of charmless $B\overline{B}$ MC for M_{BC} and ΔE.

To model the C'_{NN} distribution, three summed Gaussians were used for all modes except continuum, which uses two summed Gaussians. The PDFs for all event types are shown in Figure 34-Figure 40 and summarised in Table 6.

Table 6: PDFs used to model $\boldsymbol{B^0 \to \overline{D^0}\pi^0}$.

	M_{BC}	ΔE	C'_{NN}
Signal ($\overline{D^0} \to K^+\pi^-$)	Crystal Ball fn.	Crystal Ball fn.	3 Gaussians
Signal ($\overline{D^0} \to K^+\pi^-\pi^0$)	2D kernel density estimation PDF		3 Gaussians
Generic $B\overline{B}$	2D kernel density estimation PDF		3 Gaussians
Continuum	ARGUS fn.	Chebyshev Polynomial	Continuum
Charmless $B\overline{B}$	2D kernel estimation histogram PDF		3 Gaussians

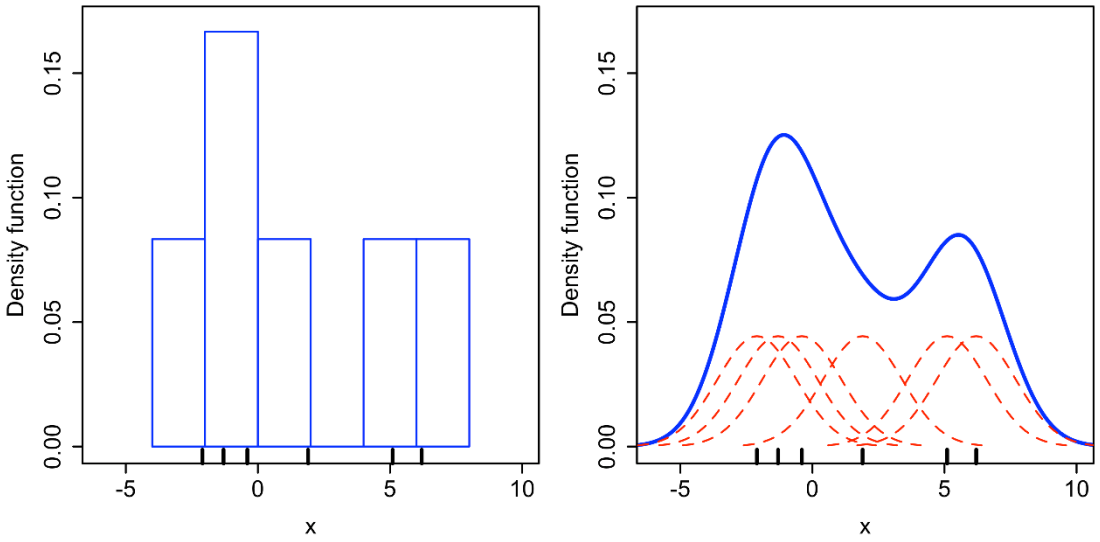


Figure 33: Example comparing a 1D histogram (left) to kernel density estimation (right) Figure shows a comparison between a histogram and a kernel density estimation for a sample of datapoints.. The black lines are datapoints, the red dashed lines are the kernel functions, and the blue line is the PDF that is the superposition of these.

3.6.2 Kernel Density Estimation PDFs

As mentioned, there are a number event types where well-fitting analytic PDFs could not be found due to complex correlations between M_{BC} and ΔE. To model these, kernel density estimations are used, generated from the MC samples. A kernel density estimation is similar to a histogram, however instead of assigning datapoints to discrete bins, each point is represented by a kernel function, and the superposition of these functions forms a continuous (depending on the functions) PDF. Gaussian distributions are used as the kernel function (as is common), the width of which is known as the bandwidth. This controls to the degree to which the PDF is smoothed or retains information. Roofit [60] is used as part of ROOT [61] to create the PDFs, and this uses an algorithm to dynamically vary the bandwidth of the kernels based on the local event density [64]. This means that for high density areas where the statistical uncertainty is low, a small bandwidth is chosen to preserve detailed shape information, while for low density areas where uncertainty is high, a large bandwidth is used to smooth out statistical fluctuations. Figure 33 shows a comparison between a histogram and a kernel density estimation for a sample of datapoints.

ROOT is unable to export 2D kernel density estimation PDFs for storage in a way that the fitting code can take it as input (and generation is very slow so cannot be done for every fit). To address this, the following method is used to store the kernel density estimation PDF data:

1. 2D kernel density estimation PDFs are generated from MC.
2. Kernel density estimation PDFs are then converted to TH2 histograms with a very large number of bins (320x320) to preserve as much data as possible and saved in a ROOT file.
3. The fitting code reads in this ROOT file, the TH2 is loaded and converted into a RooDataHist, then a RooHistPdf which is a format that can be used by the RooFit based fitter.

This technique allows the benefits of kernel density estimation PDFs with the speed and convenience of storage as a ROOT histogram.

3.6.3 Simultaneous Fitter

To extract the yield and A_{CP} of the signal from the dataset, an extended unbinned maximum likelihood fit is performed in the kinematic variables M_{BC} and ΔE and the continuum suppression variable C'_{NN} using Roofit. Data is divided into 4 bins based on the D^0 decay mode ($d \in \{D_{2b}, D_{3b}\}$) and flavour of the B^0 ($q = +1$ for B^0, $q = -1$ for $\overline{B}$), and the fit is performed simultaneously in all four bins. PDFs $\mathcal{P}_d^x(M_{BC}, \Delta E, C'_{NN})$ are defined for the event types, x: Signal (s), generic $B\overline{B}$ background (BB), continuum (c), and charmless $B\overline{B}$ (r), separately for each D^0 decay mode d: $\overline{D^0} \to K^+\pi^-$ (D_{2b}) and $\overline{D^0} \to K^+\pi^-\pi^0$ (D_{3b}) as discussed previously.

The likelihood function is defined as:

$$\mathcal{L} = \frac{e^{-\sum_x N^x}}{\prod_{q,d} N_{q,d}!} \times \prod_{q,d} \left[\prod_{j=1}^{N_{q,d}} \left(\sum_x f_d^x N^x \left[\frac{1 - q \times A_{CP}^x}{2} \right] \mathcal{P}_d^x \left(M_{\mathrm{BC}}^j, \Delta E^j, {C'_{NN}}^j \right) \right) \right], \quad (3.31)$$

where $N_{q,d}$ is the total number of events in bin (q, d); N^x is the fitted number of events of type $x \in \{s, BB, c, r\}$; A_{CP}^x is the fitted A_{CP} for event type x; $q \in \{+1, -1\}$ represents B^0 flavour; f_d^x is the fraction of events in bin $d \in \{D_{2b}, D_{3b}\}$ for type x, with $\sum_d f_d^x = 1$; and $M_{\mathrm{BC}}^j, \Delta E^j$ and ${C'_{NN}}^j$ are the values of $M_{BC}, \Delta E$ and C'_{NN} for event j.

Rather than the full likelihood function, it is perhaps easier to understand this by considering how the model simultaneously fits for the yield of each data type in the four separate bins. These four yields ($N_{q,d}$) are constrained by the total yield (N), A_{CP} and fraction of events that comes from each D^0 Decay mode f_d ($\sum_d f_d = 1$). These components can be related by the following simultaneous equations:

$$N_{+1,2b} = N \times f_{2b} \times \frac{(1 - A_{CP})}{2}, \quad (3.32)$$

$$N_{-1,2b} = N \times f_{2b} \times \frac{(1 + A_{CP})}{2}, \quad (3.33)$$

$$N_{+1,3b} = N \times f_{3b} \times \frac{(1 - A_{CP})}{2}, \quad (3.34)$$

$$N_{-1,3b} = N \times f_{3b} \times \frac{(1 - A_{CP})}{2}, \quad (3.35)$$

This relationship exists for the yields of all four event types.

The following parameters are floated as variables in the fit:

- $N^x, x \in \{s, BB, c\}$, the yield in each of the event types: Signal, generic $B\overline{B}$ background and continuum background.
- A_{CP}^s, the signal direct CP violation asymmetry.
- $f_d^x, x \in (BB, c, r)$ the fraction of events that are in the $\overline{D^0} \to K^+\pi^-$ and $\overline{D^0} \to K^+\pi^-\pi^0$ bins for backgrounds.
- The shape parameters of the continuum PDFs, $\mathcal{P}_{2b}^c(M_{BC})$, $\mathcal{P}_{3b}^c(M_{BC})$, $\mathcal{P}_{2b}^c(\Delta E)$, and $\mathcal{P}_{3b}^c(\Delta E)$.
- The calibration factors (see chapter 4.5.1) of the signal $\overline{D^0} \to K^+\pi^-$ PDFs, $\mathcal{P}_{2b}^s(M_{BC})$, $\mathcal{P}_{2b}^s(\Delta E)$.

The following parameters are held as fixed constraints:

- N^r, the yield in charmless $B\overline{B}$ background.
- f_d^s, the fraction of events that are in the $\overline{D^0} \to K^+\pi^-$ and $\overline{D^0} \to K^+\pi^-\pi^0$ bins for signal. This is set from the MC studies.
- All other PDF shapes (calibration factors) are set from MC studies (control mode studies) and are held fixed.
- $A_{CP}^x, x \in \{s, BB, c\}$, the A_{CP} for background events is fixed to 0, plus a correction for detector bias (see chapter 4.5.2).

These fixed parameters will contribute a systematic uncertainty that is studied in chapter 5.

The fitter uses MINUIT2 [66] with MIGRAD as a minimizer, and MINOS to calculate asymmetric uncertainties. MINUIT2 is called through the RooAbsPdf::fitTo function within RooFit to perform an extended likelihood fit. After a fit has completed, the status of MIGRAD and MINOS are checked to make sure both are good, as well as checking that the covariance matrix is full and accurate. If either of these checks fail, the floated parameters are reset to their initial values and randomised (by a Gaussian around their initial value with σ equal to their uncertainty) and then the fit is retried up to a maximum of 10 times. If this does not produce a good fit, fitting is deemed to have failed. All the final fits to data were confirmed as successful.

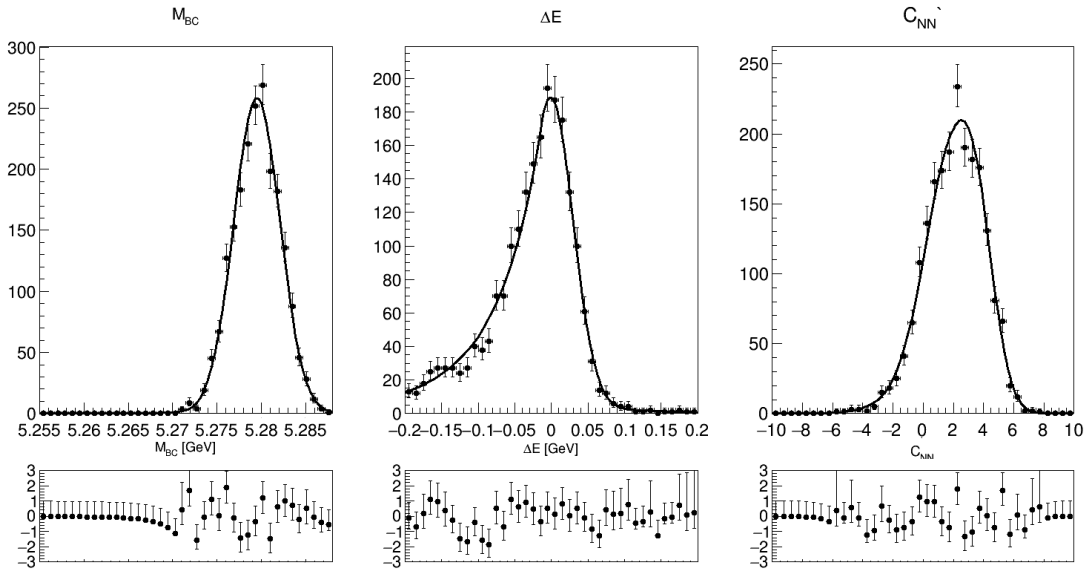


Figure 34: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C'_{NN}}$ PDFs for $\boldsymbol{\overline{D^0} \to K^+\pi^-}$ signal.

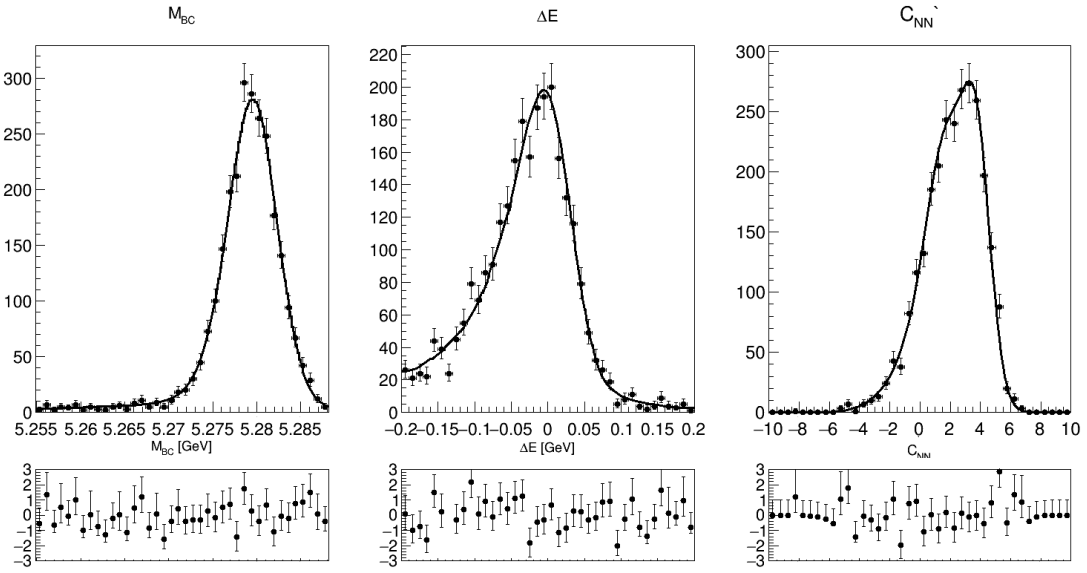


Figure 35: M_{BC}, ΔE and C'_{NN} PDFs for $\overline{D^0} \to K^+\pi^-\pi^0$ signal.

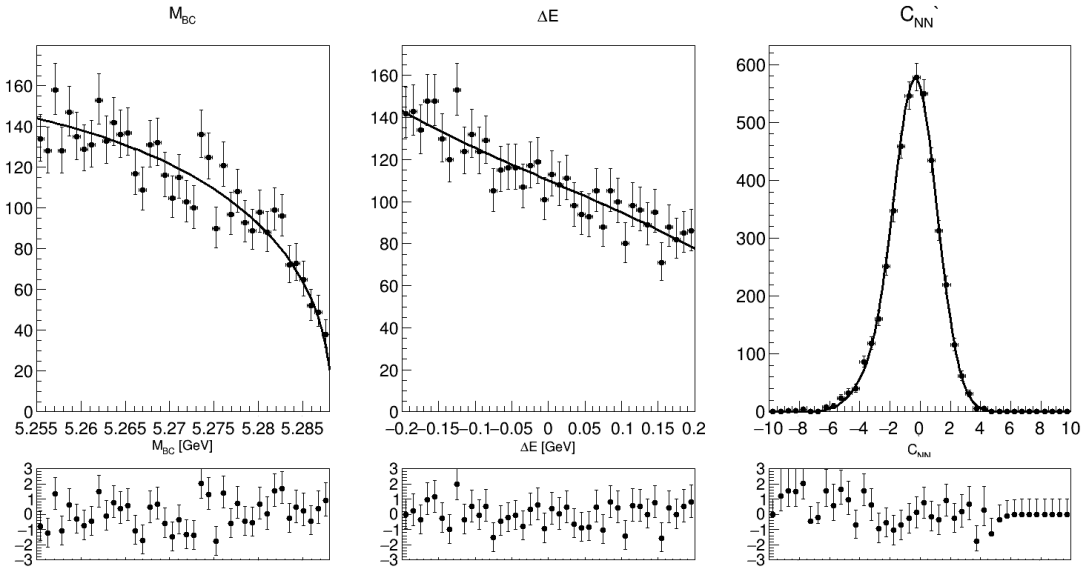


Figure 36: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C'_{NN}}$ PDFs for $\overline{\boldsymbol{D^0}} \rightarrow \boldsymbol{K^+\pi^-}$ continuum.

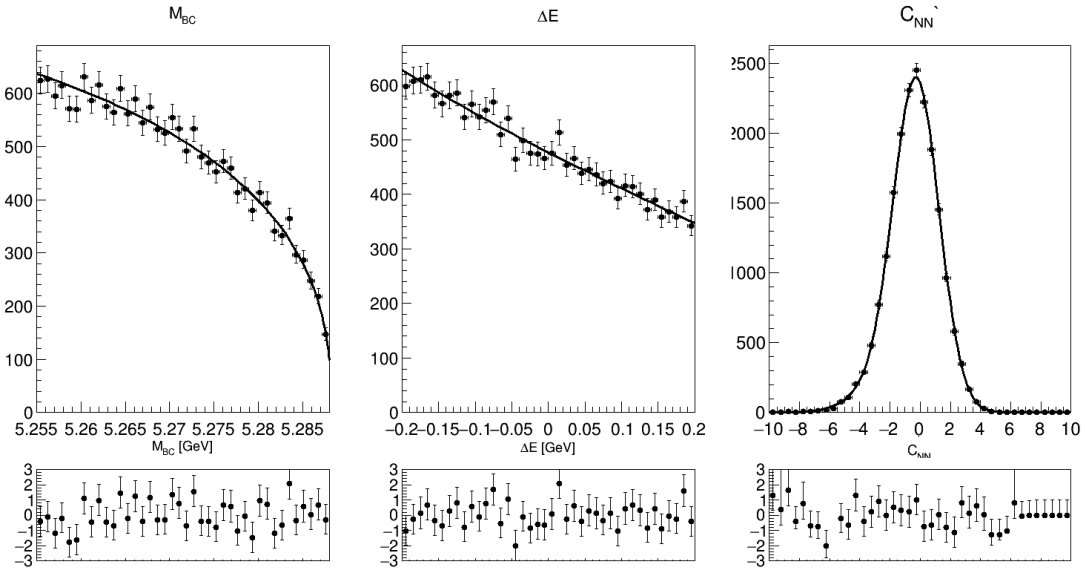


Figure 37: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C'_{NN}}$ PDFs for $\overline{\boldsymbol{D^0}} \rightarrow \boldsymbol{K^+\pi^-\pi^0}$ continuum.

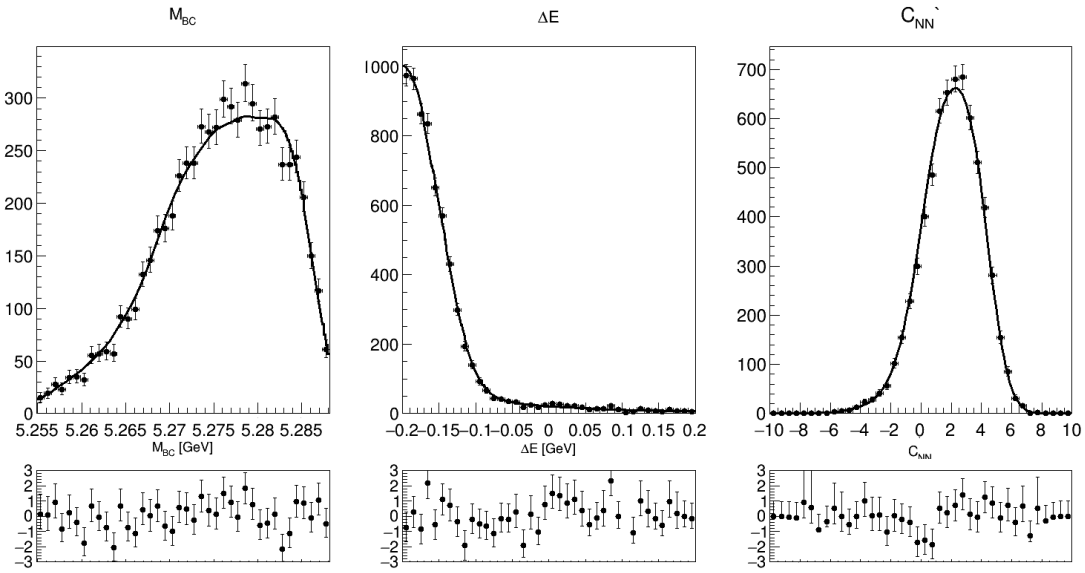


Figure 38: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C'_{NN}}$ PDFs for $\overline{\boldsymbol{D^0}} \rightarrow \boldsymbol{K^+\pi^-}$ generic $\boldsymbol{B\overline{B}}$ background.

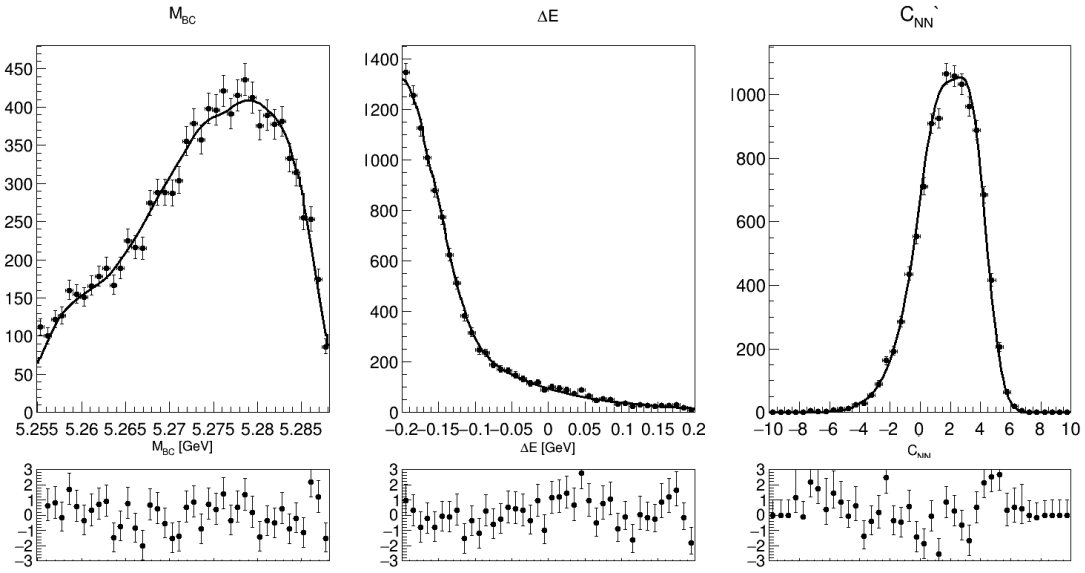


Figure 39: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C'_{NN}}$ PDFs for $\overline{\boldsymbol{D^0}} \rightarrow \boldsymbol{K^+\pi^-\pi^0}$ generic $\boldsymbol{B\overline{B}}$ background.

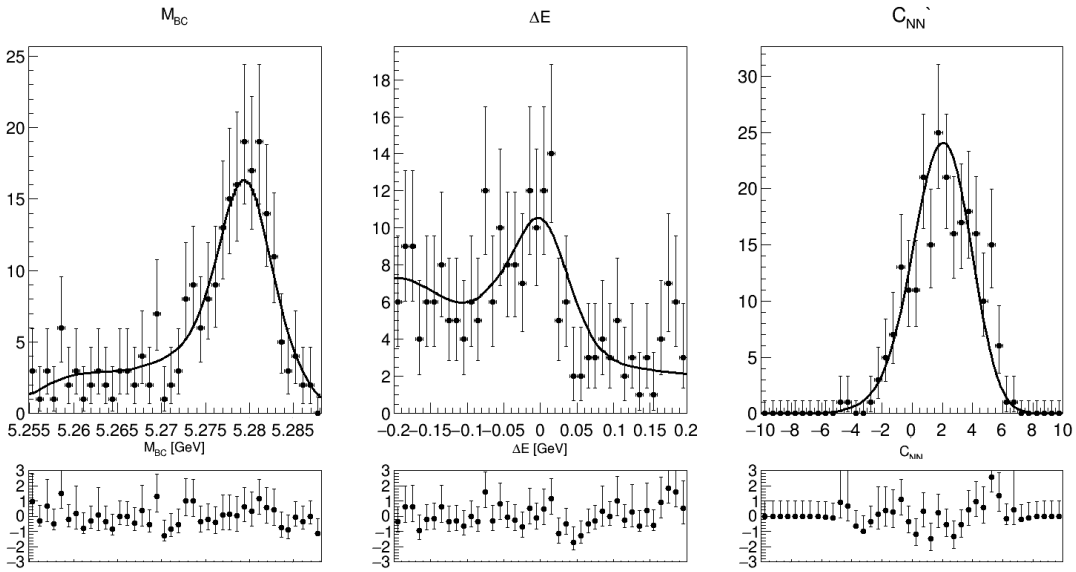


Figure 40: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C'_{NN}}$ PDFs for $\overline{\boldsymbol{D^0}} \rightarrow \boldsymbol{K^+\pi^-}$ charmless $\boldsymbol{B\overline{B}}$ background.

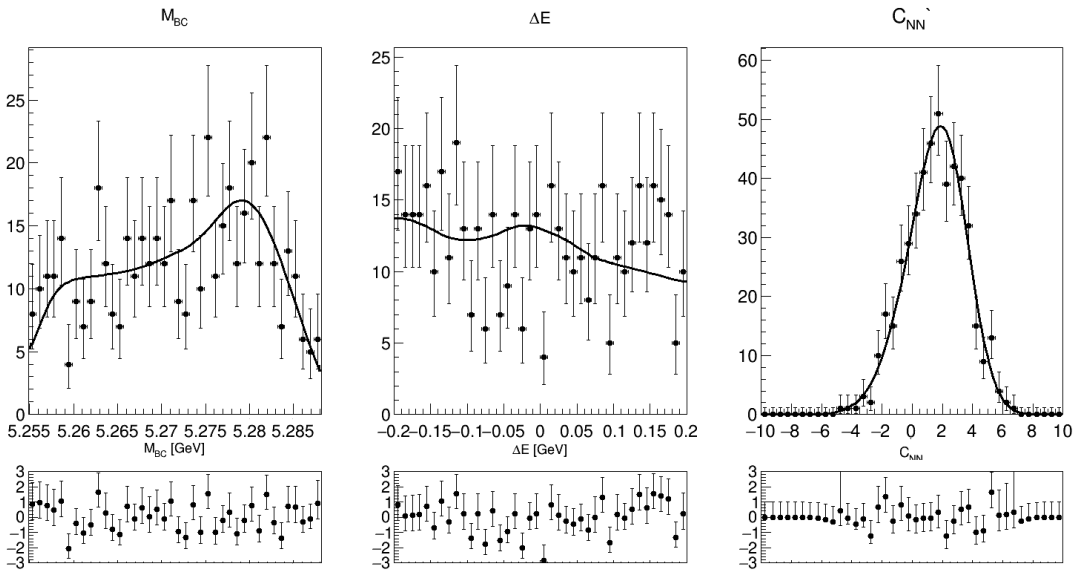


Figure 41: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C'_{NN}}$ PDFs for $\overline{\boldsymbol{D^0}} \rightarrow \boldsymbol{K^+\pi^-\pi^0}$ charmless $\boldsymbol{B\overline{B}}$ background.

3.7 Fitter Tests

Ensemble testing is conducted on the fitter to confirm that it is stable, has no significant bias and produces an accurate statistical uncertainty.

3.7.1 Toy Ensemble Test

A pull test is conducted primarily to verify the accuracy of the statistical uncertainty returned by the fitter, but it also provides information on any bias. This pull test is conducted by performing fits on 500 toy datasets and calculating the pull, defined as:

$$Pull = \frac{N_{fitted} - N_{input}}{\sigma}, \tag{3.36}$$

where N_{fitted} is the fitted signal yield, N_{input} is the expected signal yield in the dataset, and σ is the standard deviation of the fit. Ideally the mean of the pull distribution should be 0, representing the mean of the fitted yield being equal to the input yield, and the standard deviation of the pull (σ_{pull}) should be 1, indicating the statistical uncertainty reported by the fitter is the same as the standard deviation of the ensemble distribution. The datasets used are generated from the PDF shapes described in chapter 3.6.1, with the number of events of each type (signal, generic $B\overline{B}$ background, continuum background and charmless $B\overline{B}$ background) randomised by a Poisson distribution around the expected number from MC studies. The results of the pull test with N_{input} as the expected yield in data is shown in Figure 42. The pull test shows a mean pull of -0.016 ± 0.045, consistent with 0 and $\sigma_{pull} = 0.987 \pm 0.032$ consistent with unity, indicating no systematic bias in the fitter and an accurate uncertainty estimation respectively.

3.7.2 MC Ensemble Test

The same pull test described above is performed replacing the signal and charmless $B\overline{B}$ datasets generated from PDFs with data randomly selected from the MC data samples. The result of this test is shown in Figure 43. Doing this allows us to study our PDF modelling in the fitter. Unfortunately, generic $B\overline{B}$ and continuum background events cannot be selected from MC samples, as the MC datasets are not large enough to ensure the randomly selected samples are independent, so these events are still generated from PDFs. This test also shows a σ_{pull} consistent with unity, indicating good uncertainty estimation. There is a small pull found of 0.274 ± 0.045 indicating a very small bias most likely due to modelling inaccuracies in the signal PDFs.

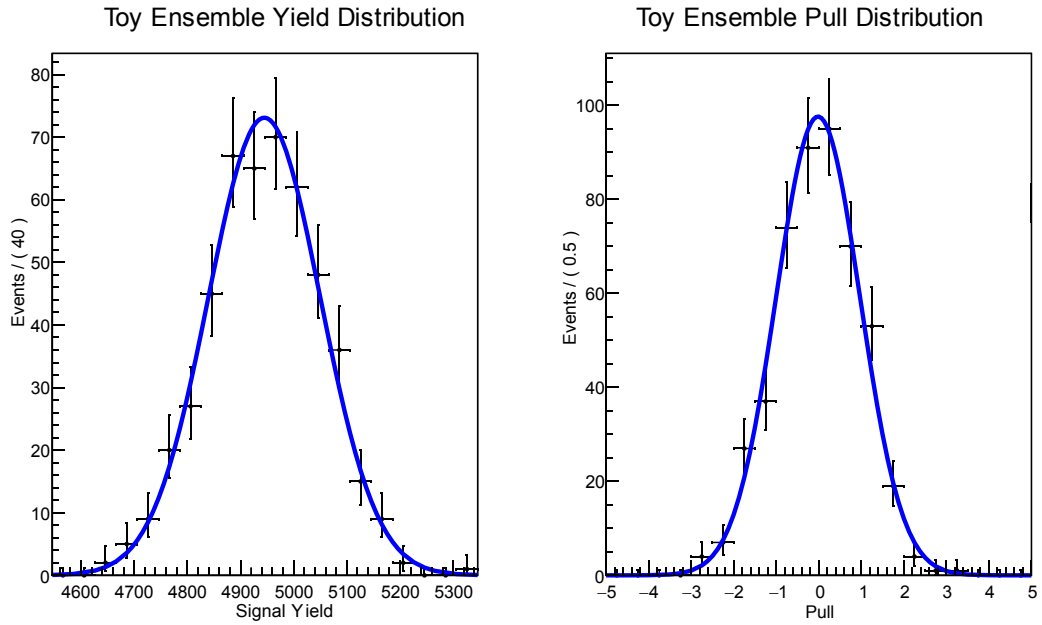


Figure 42: Yield (left) and corresponding pull (right) distribution for toy ensemble test.

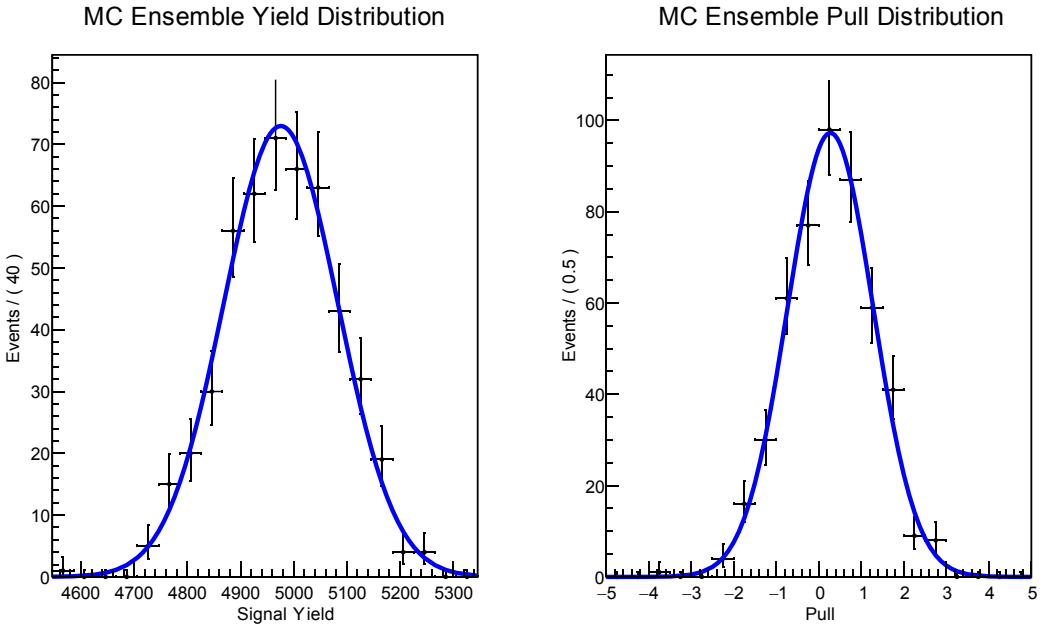


Figure 43: Yield (left) and corresponding pull (right) distribution for MC ensemble test.

A small number of tests were conducted with using each of the 6 streams of background for a pure MC test, and results were consistent within uncertainty (no true pull test could be conducted with this small sample size).

To quantify this bias, a linearity test is performed by conducting 5 MC pull tests at a range of different input yields, and then fitting a linear equation to the relation

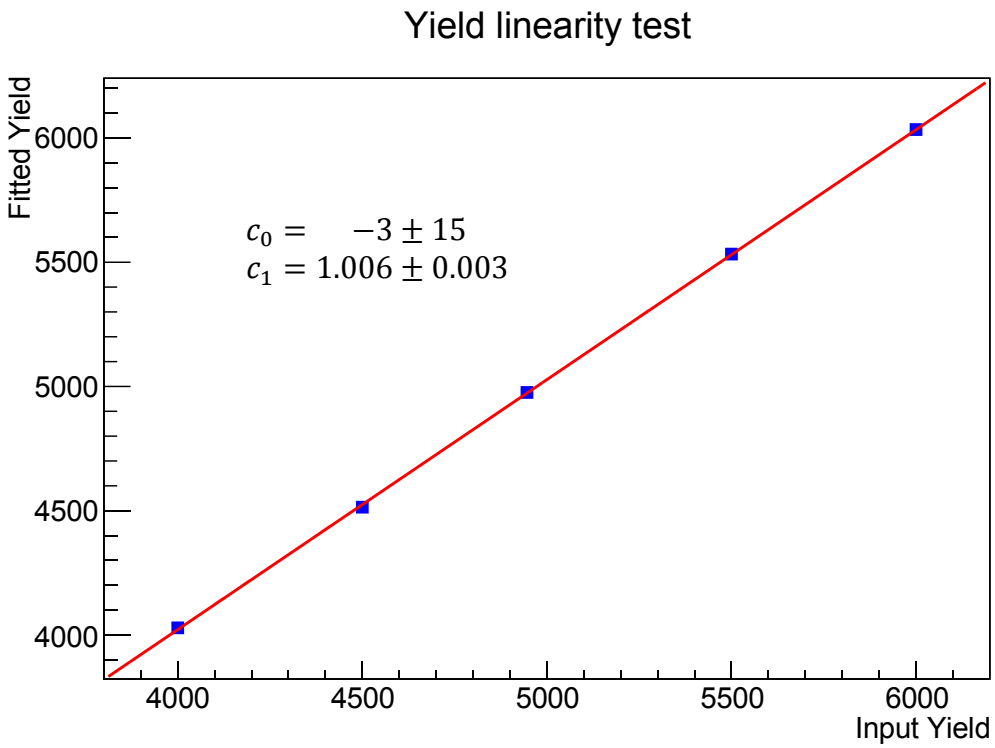


Figure 44: Linearity plot for Yield. Error bars are uncertainty in fitted mean of distributions but are too small to see.

between the fitted number of signal events and true number of signal events. A well-fitting linear relation is shown in Figure 44, and indicates a fitter bias of $+0.6 \pm 0.3\%$. As this bias is the same order as its uncertainty, it is not corrected for and instead it taken as a systematic uncertainty.

3.7.3 A_{CP} Ensemble Test

Toy (Figure 45) and MC (Figure 46) ensemble pull tests are also conducted on the fitted A_{CP} in the same way as signal yield. The toy ensemble test shows pull of 0.032 ± 0.047 consistent with 0, and $\sigma_{pull} = 1.017 \pm 0.035$ consistent with 1, indicating no bias in the fitter and an accurate statistical uncertainty. The MC ensemble test return a pull of 0.131 ± 0.049, a small bias again likely caused by imperfect signal PDFs, with $\sigma_{pull} = 1.056 \pm 0.037$ consistent with 1.

As with yield, in order to quantify any bias, a linearity test is conducted, using 9 different values for dataset A_{CP}. Figure 47 shows the fitted linear distribution which indicates a bias in fitted A_{CP} of $(+0.36 \pm 0.03) \times 10^{-2}$. This bias will be corrected for when A_{CP} is measured.

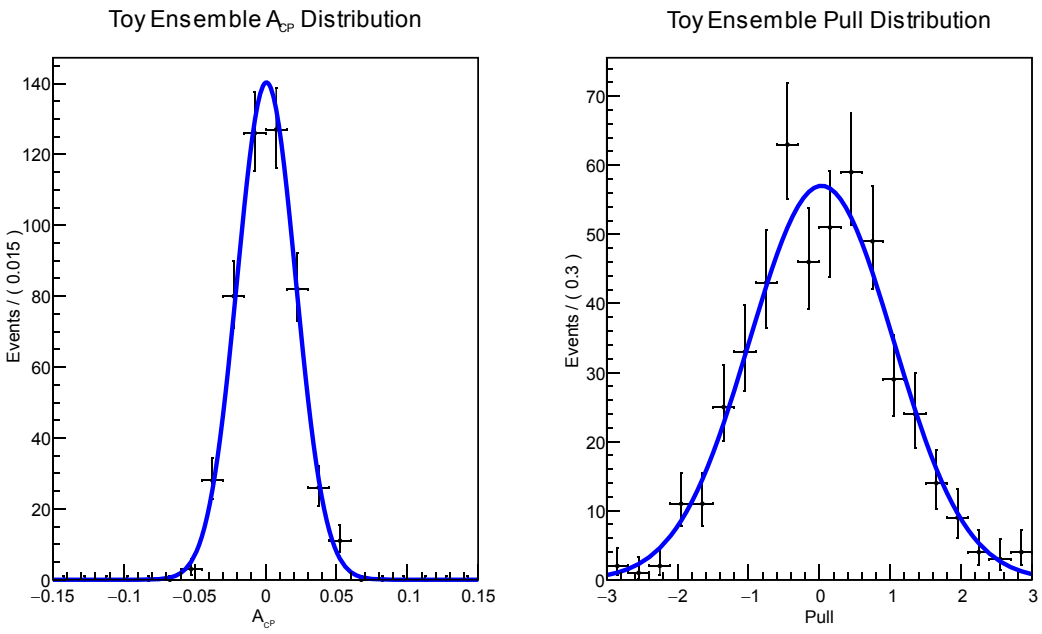


Figure 45: A_{CP} (left) and corresponding pull (right) distribution for toy ensemble test.

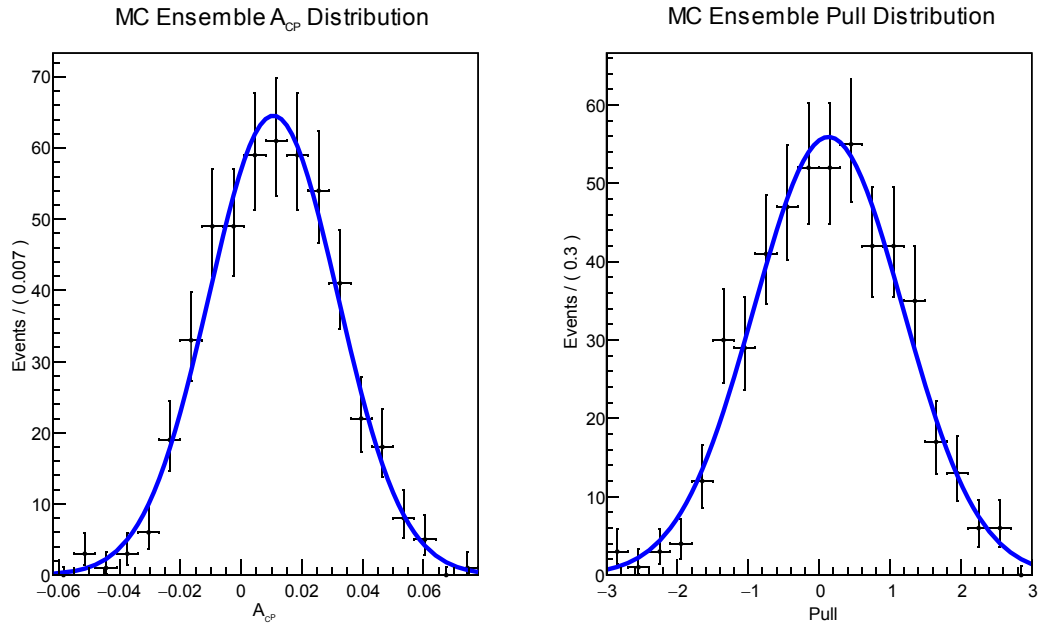


Figure 46: A_{CP} (left) and corresponding pull (right) distribution for MC ensemble test.

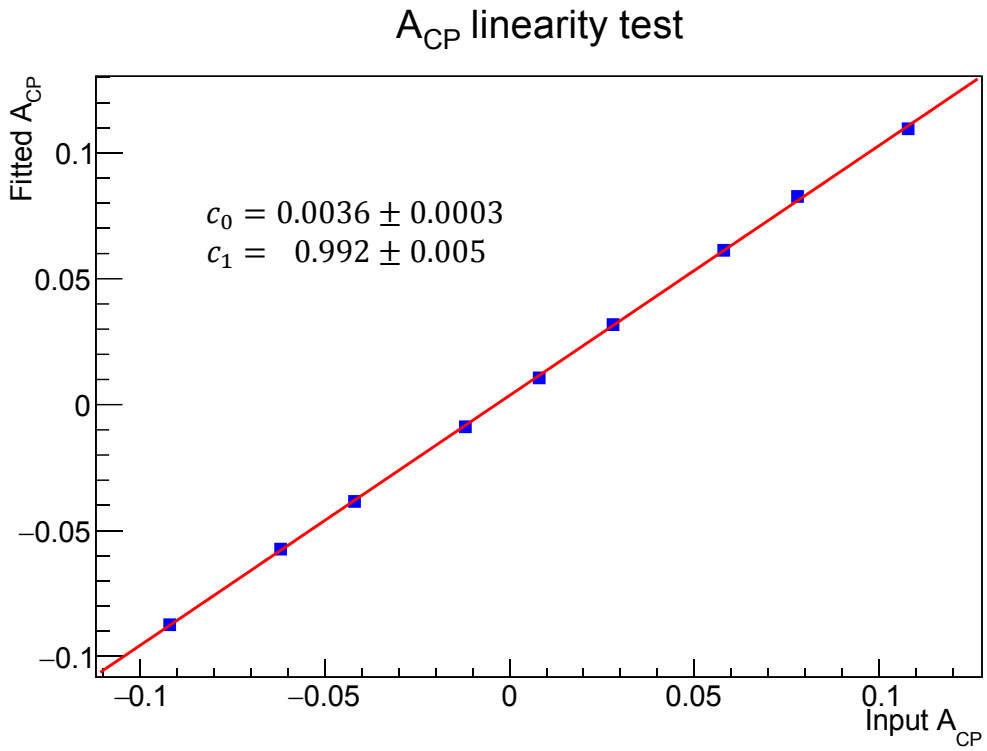


Figure 47: Linearity plot for $\boldsymbol{A_{CP}}$. Error bars are uncertainty in fitted mean of distributions and are too small to see.

4 $B^+ \to \overline{D^0}\pi^+$

As a way of validating the reconstruction efficiencies and fitting methodology, the same techniques are used to analyse a control mode with an already well measured branching fraction. As well as ensuring the fitted branching fraction and A_{CP} are consistent with existing measurements, this control mode is used to fit calibration factors to the shape parameters of the PDFs to account for any difference in variable distribution between MC and data. These calibration factors will then be applied in the $B^0 - \overline{D^0}\pi^0$ fit. Lastly, this control mode provides a data driven estimation of the systematic uncertainty associated with the A_{CP} correction for detector asymmetry (see chapter 5.2.3).

$B^+ \to \overline{D^0}\pi^+$ is chosen as the control mode due to its precisely measured branching fraction and having similar final state particles and kinematics (as the $\overline{D^0}$ is reconstructed in the same subdecay modes). The current world average value [3] of

$$\mathfrak{B}\left(B^+ \to \overline{D^0}\pi^+\right) = (4.68 \pm 0.13) \times 10^{-3} \quad (4.1)$$

gives an estimated yield of ~85000 events, which is roughly 19 times the expected amount of $B^0 \to \overline{D^0}\pi^0$ signal. This large yield ensures a high statistical accuracy for checking consistency of fitted branching fraction and A_{CP}.

Although $B^+ \to \overline{D^0}\pi^+$ is primarily studied here as a control mode, a full analysis is performed, including calculation of systematics uncertainties. Because of the extremely high yield and low statistical uncertainty, systematics analysis was needed to get a measure on how the result compares to previous measurement. As the full analysis is performed, the result of this study is also reported as a new measurement.

4.1 Doubly Cabibbo Suppressed Decays

As with $B^0 \to \overline{D^0}\pi^0$, there are DCS modes that will cause flavour tagging via the kaon charge to be incorrect. The wrong sign decay $B^+ \to D^0\pi^+$ is not only doubly Cabibbo suppressed, it is also colour suppressed unlike the right sign $B^+ \to \overline{D^0}\pi^+$ decay. This means that its effect will be negligible compared to that of the D^0 DCS modes. Using the PDG average values [3]

$$R(B^+ \to [K^+\pi^-]_D\pi^+) = (3.53 \pm 0.14) \times 10^{-3}, \quad (4.2)$$

and

$$R(B^+ \to [K^+\pi^-\pi^0]_D\pi^+) = (2.2 \pm 0.4) \times 10^{-3}, \quad (4.3)$$

it is found that the correction due for DCS effects for the branching fraction is $(-0.28 \pm 0.02)\%$ and for A_{CP} is $(+1.08 \pm 0.08) \times 10^{-5}$. As with $B^0 \to \overline{D^0}\pi^0$, a small correction is made for the effect on $\mathfrak{B}$, but the correction for A_{CP} is orders of magnitude smaller than the total uncertainty, so is neglected.

4.2 Reconstruction

Reconstruction of the $B^+ \to \overline{D^0}\pi^+$ decay is conducted in a similar manner to that of the $B^0 \to \overline{D^0}\pi^0$ decay, but replacing the B^0 daughter π^0 candidate with a $\pi^\pm$ from the charged particle list. After the below kinematic requirements and continuum suppression are accounted for, the efficiencies for each $\overline{D^0}$ decay channel are as follows:

$$\epsilon_{\overline{D^0}\to K^+\pi^-} = (33.08 \pm 0.04)\%, \quad (4.4)$$

$$\epsilon_{\overline{D^0}\to K^+\pi^-\pi^0} = (9.05 \pm 0.02)\%, \quad (4.5)$$

The number of expected events based of each type based off MC are shown in Table 7:

Table 7: Expected number of events of each type in $\boldsymbol{B^+ \to \overline{D^0}\pi^+}$ based on MC study.

Type	$\overline{D^0} \to K^+\pi^-$	$\overline{D^0} \to K^+\pi^-\pi^0$
$B^+ \to \overline{D^0}\pi^+$ Signal	44590	43633
Generic $B\overline{B}$ background	37391	40284
Continuum	5473	16689
Charmless $B\overline{B}$	337	532

4.2.1 Continuum Suppression

The same technique as used for the $B^0 \to \overline{D^0}\pi^0$ signal is used for the $B^+ \to \overline{D^0}\pi^+$ study (see chapter 3.3).

4.2.2 $B\overline{B}$ Background Reduction

The selection criteria used to reduce $B\overline{B}$ background are similar to those for $B^+ \to \overline{D^0}\pi^+$ (see chapter 3.4) and are shown in Table 8.

Table 8: Selection criteria for $B\overline{B}$ background reduction in $B^+ \to \overline{D^0}\pi^+$.

Variable	Accepted Range
M_{BC}	5.253 – 5.288 GeV
ΔE	−0.2 – 0.2 GeV
$M(D^0)$	1.85 – 1.88 GeV
$M\left(\pi^0_{B^0}\right)(K^+\pi^-)$	104.07 – 163.1 MeV
$M\left(\pi^0_{B^0}\right)(K^+\pi^-\pi^0)$	105.77 – 164.37 MeV
$M\left(\pi^0_{D^0}\right)(K^+\pi^-\pi^0)$	107.88 – 162.264 MeV
$p(K^\pm)$	0.3 – 3.5 GeV
$p(\pi^\pm)$	0.2 – 3.5 GeV
$p(\pi^\pm_{B^0})$	1.6 – 3.5 GeV

4.3 $B^+ \to \overline{D^0}\pi^+$ PDFs

The PDFs used to model $B^+ \to \overline{D^0}\pi^+$ are the same as those used for the $B^0 \to \overline{D^0}\pi^0$ signal mode. They are shown in Table 9.

Table 9: PDFs used to model $B^+ \to \overline{D^0}\pi^+$.

	M_{BC}	ΔE	C'_{NN}
Signal ($\overline{D^0} \to K^+\pi^-$)	Crystal Ball fn.	Crystal Ball fn. + Gaussian	3 Gaussians
Signal ($\overline{D^0} \to K^+\pi^-\pi^0$)	2D kernel estimation histogram PDF		3 Gaussians
Generic $B\overline{B}$ background	2D kernel estimation histogram PDF		3 Gaussians
Continuum	ARGUS fn.	Chebyshev Polynomial	2 Gaussians
Charmless $B\overline{B}$	2D kernel estimation histogram PDF		3 Gaussians

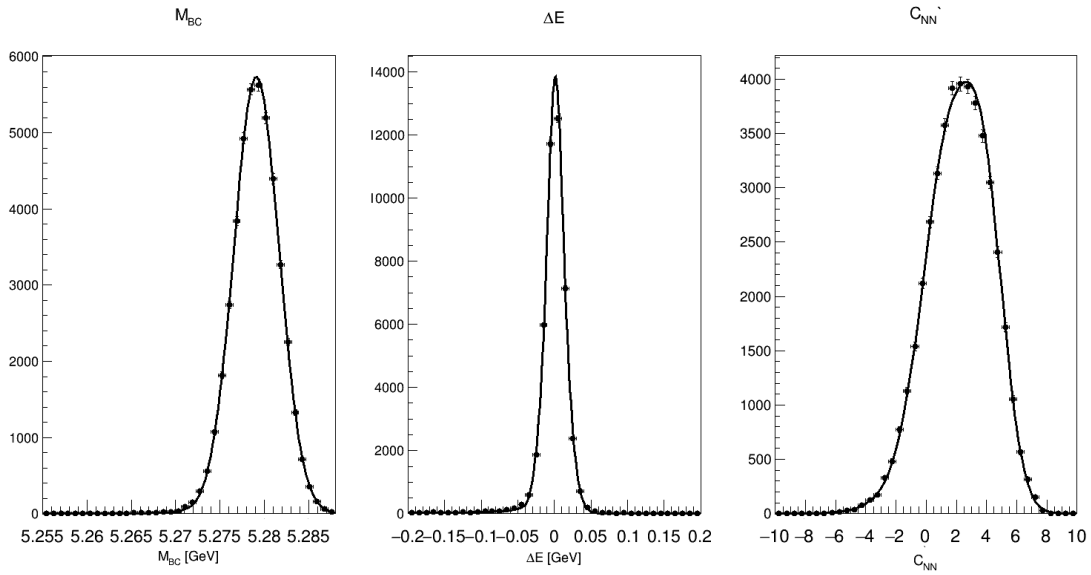


Figure 48: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C'_{NN}}$ PDFs for $\boldsymbol{\overline{D^0} \rightarrow K^+\pi^-}$ signal.

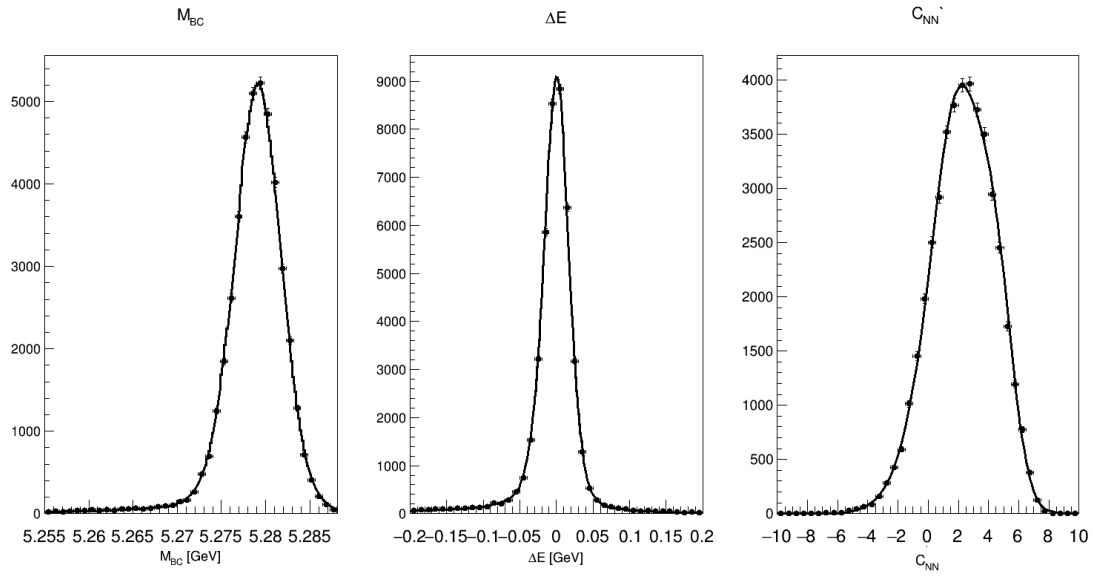


Figure 49: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C'_{NN}}$ PDFs for $\boldsymbol{\overline{D^0} \rightarrow K^+\pi^-\pi^0}$ signal.

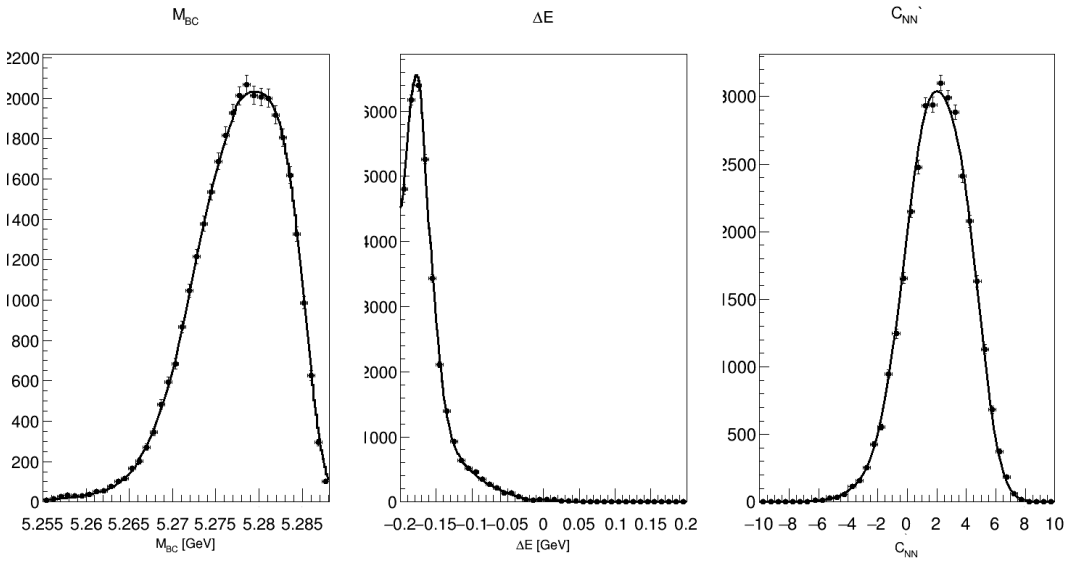


Figure 50: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C'_{NN}}$ PDFs for $\boldsymbol{\overline{D^0} \rightarrow K^+\pi^-}$ generic $\boldsymbol{B\overline{B}}$ background.

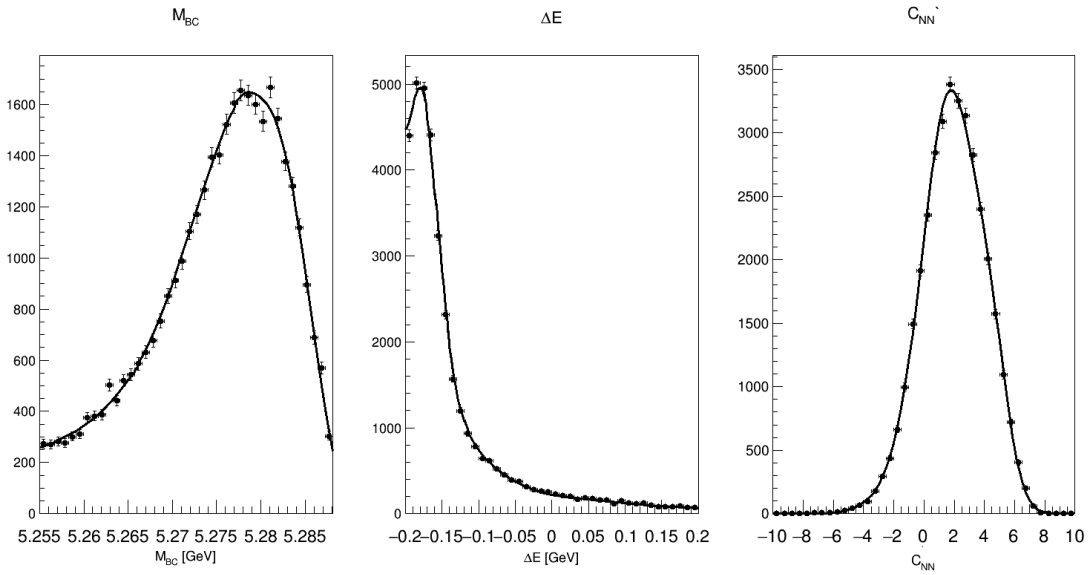


Figure 51: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C'_{NN}}$ PDFs for $\boldsymbol{\overline{D^0} \rightarrow K^+\pi^-\pi^0}$ generic $\boldsymbol{B\overline{B}}$ background.

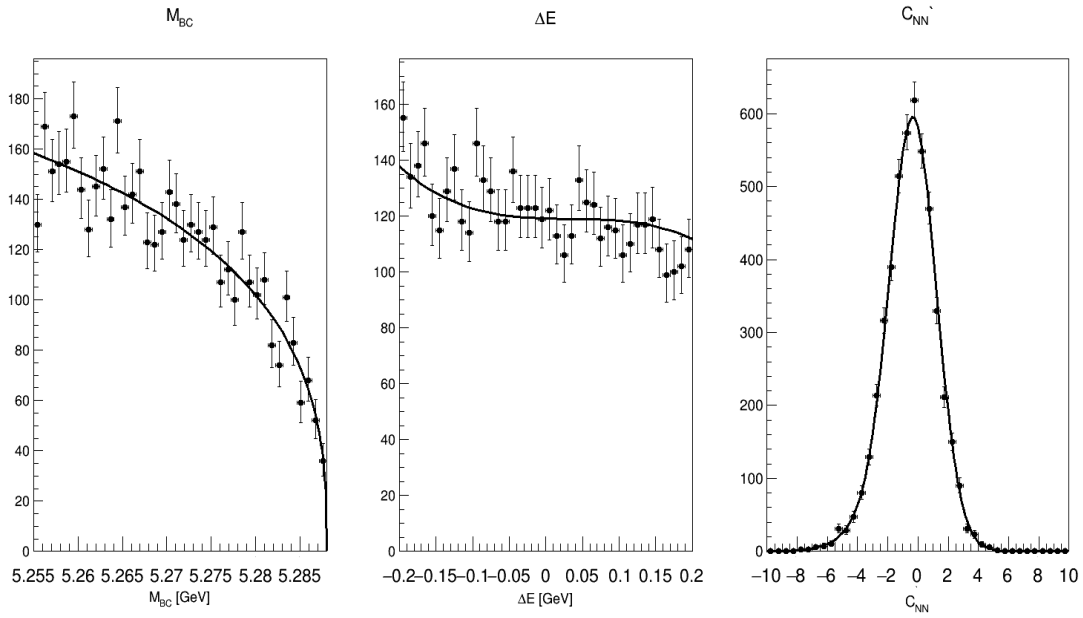


Figure 52: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C_{NN}`}$ PDFs for $\overline{\boldsymbol{D^0}} \rightarrow \boldsymbol{K^+\pi^-}$ continuum background.

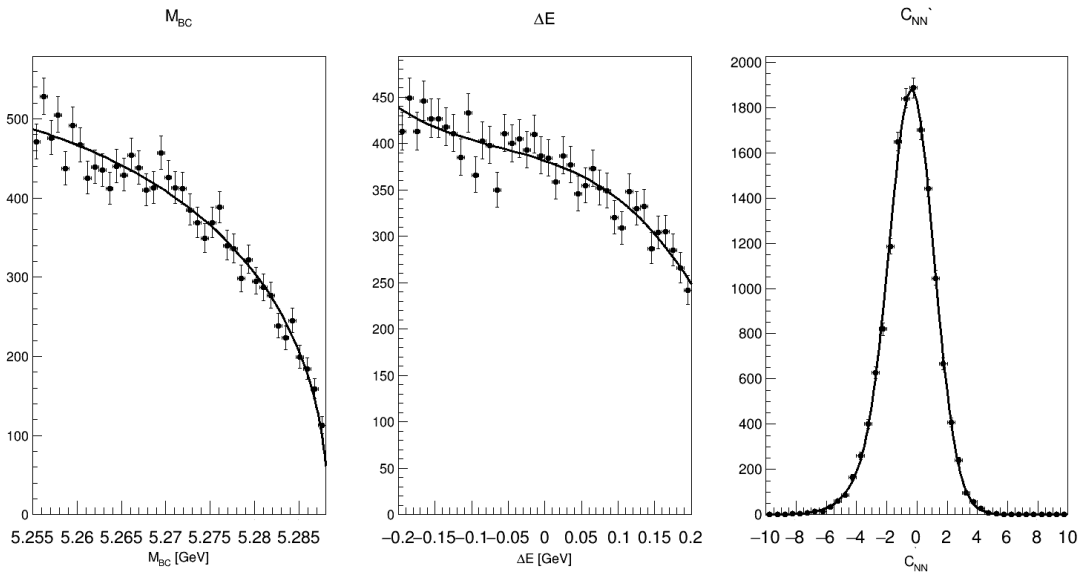


Figure 53: $\boldsymbol{M_{BC}}$, $\boldsymbol{\Delta E}$ and $\boldsymbol{C'_{NN}}$ PDFs for $\overline{\boldsymbol{D^0}} \rightarrow \boldsymbol{K^+\pi^-\pi^0}$ continuum background.

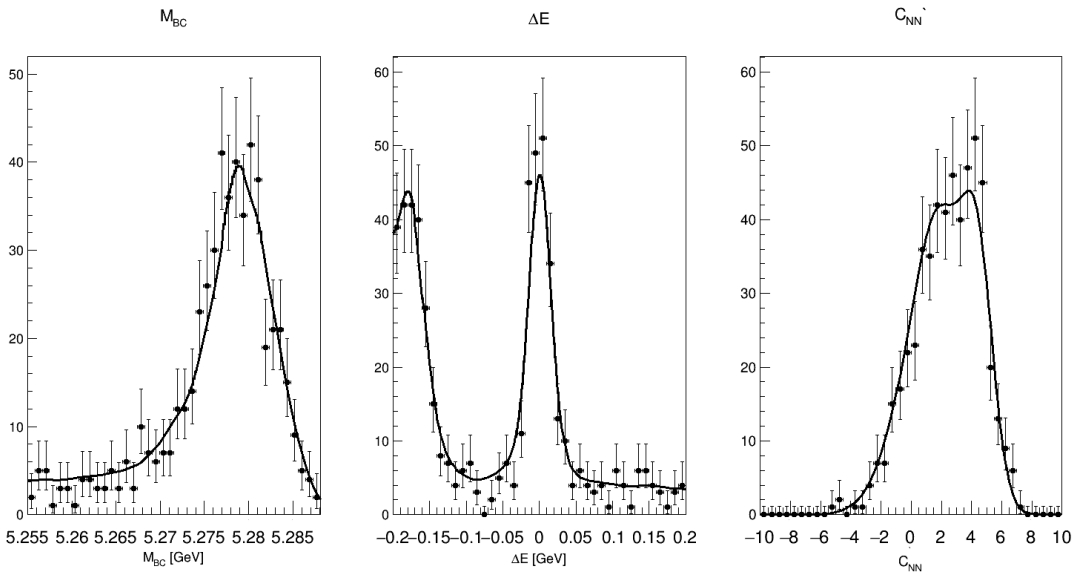


Figure 54: M_{BC}, ΔE and C'_{NN} PDFs for $\overline{D^0} \to K^+\pi^-$ charmless $B\overline{B}$ decay background.

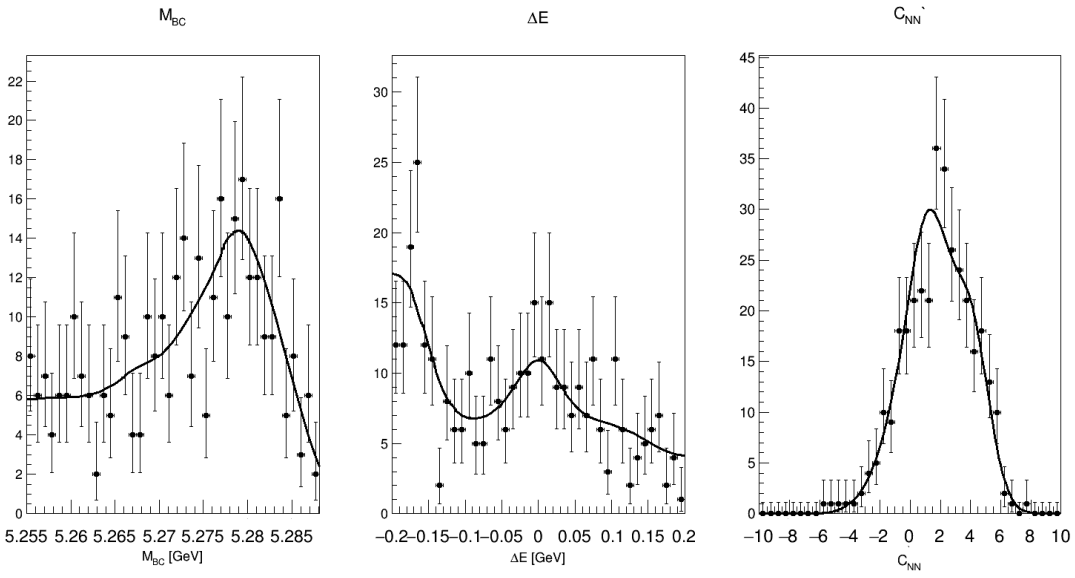


Figure 55: M_{BC}, ΔE and C'_{NN} PDFs for $\overline{D^0} \to K^+\pi^-\pi^0$ charmless $B\overline{B}$ decay background.

4.4 $B^+ \to \overline{D^0}\pi^+$ Fitter Bias Test

As with $B^0 \to \overline{D^0}\pi^0$, a linearity test is performed for both yield and A_{CP} using $B^+ \to \overline{D^0}\pi^+$ MC to check for fitting bias. 5 different yield inputs and 7 different A_{CP} inputs are fitted each 500 times using the same method described in chapter 3.6.

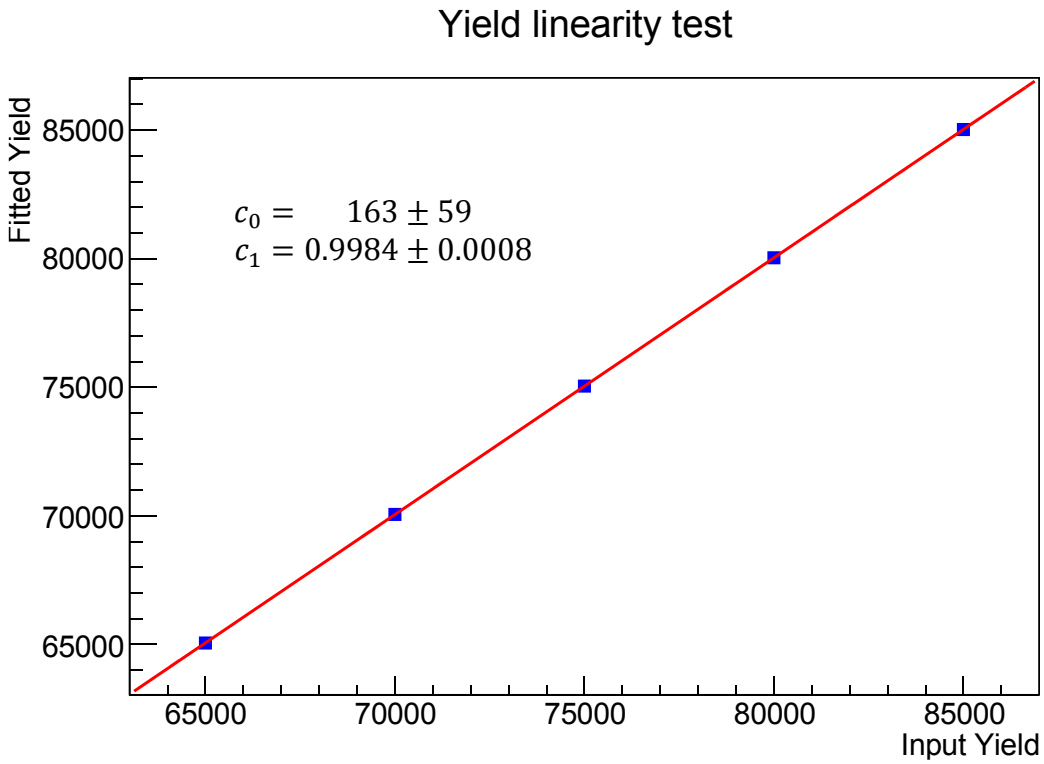


Figure 56: Linearity test for Yield using MC data.

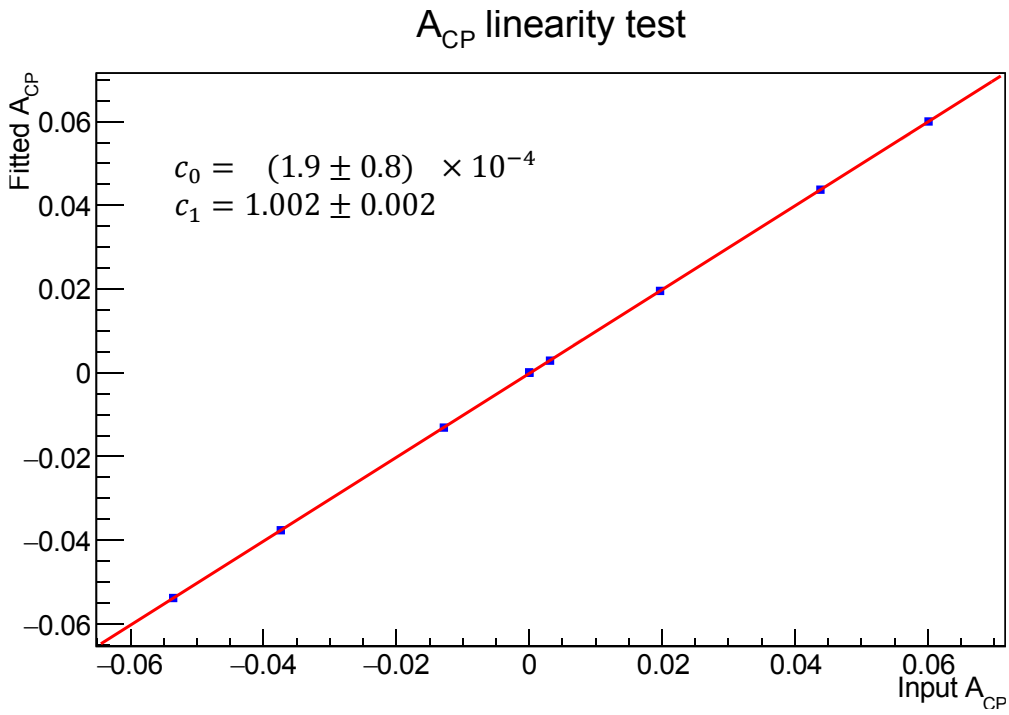


Figure 57: Linearity test for A_{CP} using MC data.

The results of these tests show that a dataset with a known yield of 75,000 events the fitted would be fitted as 75044 ± 118 which is consistent with no bias. Similarly, a dataset with known A_{CP} of 0.0200 would be fitted at 0.0198 ± 0.0002 which is also consistent with no bias.

4.5 $B^+ \to \overline{D^0}\pi^+$ Fit to Data

4.5.1 MC – Data Discrepancies

To account for any potential differences in the distribution of fitting variables between MC and data, additional parameters (calibration factors) are added to the PDFs to adjust the parameters by a small amount to take this difference into consideration.

Calibration factors are applied as mean shifts and width factors to the C'_{NN} Gaussians. The same calibration factor is applied to all components of the PDF (eg. The same factor for all Gaussians describing a C'_{NN} distribution).

It was found that floating a large number of parameters (yields, signal A_{CP}, calibration factors and continuum PDF parameters) at once in RooFit caused the MINOS uncertainty calculation to be unstable and regularly report errors. Due to this, the calibration factors for $D^0 \to K^+\pi^-$ and $D^0 \to K^+\pi^-\pi^0$ PDFs are fitted separately by performing a fit over the $B^+ \to \overline{D^0}\pi^+$ control mode data sample containing only events reconstructed in that D^0 mode. This reduces the number of parameters that MINUIT and MINOS have to deal with simultaneously, allowing for correct uncertainty calculations to be returned consistently. Fitted calibration factors are shown in Table 10.

These calibration factors are applied to the PDF used in fitting both $B^+ \to \overline{D^0}\pi^+$ and $B^0 \to \overline{D^0}\pi^0$. The C'_{NN} calibration factors are fixed as constant during the fit. Systematic uncertainties for the effect of these calibration factors are assigned as the difference in fitted results with and without them applied.

Table 10: Table of calibration factors from fitting the $B^+ \to \overline{D^0}\pi^+$ control mode.

Signal Calibration factors	
$\mu(C'_{NN})_{2bd}$	-0.040915 ± 0.0107
$\sigma(C'_{NN})_{2bd}$	1.0168 ± 0.00549
$\mu(C'_{NN})_{3bd}$	-0.054907 ± 0.0113
$\sigma(C'_{NN})_{3bd}$	1.0069 ± 0.00561
Generic $B\overline{B}$ Background Calibration factors	
$\mu(C'_{NN})_{2bd}$	-0.026961 ± 0.0117
$\sigma(C'_{NN})_{2bd}$	1.0184 ± 0.00602
$\mu(C'_{NN})_{3bd}$	-0.022443 ± 0.0133
$\sigma(C'_{NN})_{3bd}$	0.97554 ± 0.00682
$q\overline{q}$ Background Calibration factors	
$\mu(C'_{NN})_{2bd}$	0.17656 ± 0.0280
$\sigma(C'_{NN})_{2bd}$	1.0738 ± 0.0141
$\mu(C'_{NN})_{3bd}$	0.034005 ± 0.0196
$\sigma(C'_{NN})_{3bd}$	1.0348 ± 0.00968

The distribution of ΔE in data was found to deviate noticeably from the signal PDFs modelled on MC. Rather than using calibration factors, the mean and width of the PDFs for ΔE and M_{BC} in the $D^0 \to K^+\pi^-$ mode are floated in the fit. This means they will not contribute additional systematic uncertainties from having a fixed shape.

For the ΔE distribution in the $D^0 \to K^+\pi^-\pi^0$ mode, there are no parameters in the KEST histogram PDF to float. Instead, several KEST PDFs are generated with a variety of widths (by adding to ΔE a Gaussian generated random number, centred around 0 with σ between 0.005 GeV and 0.02 GeV). These were each tested in fits to $B^+ \to \overline{D^0}\pi^+$ data and $\sigma = 0.007\ GeV$ was found to give a PDF that was the closest match to the real data distribution. The effect of this modification is shown in Figure 58.

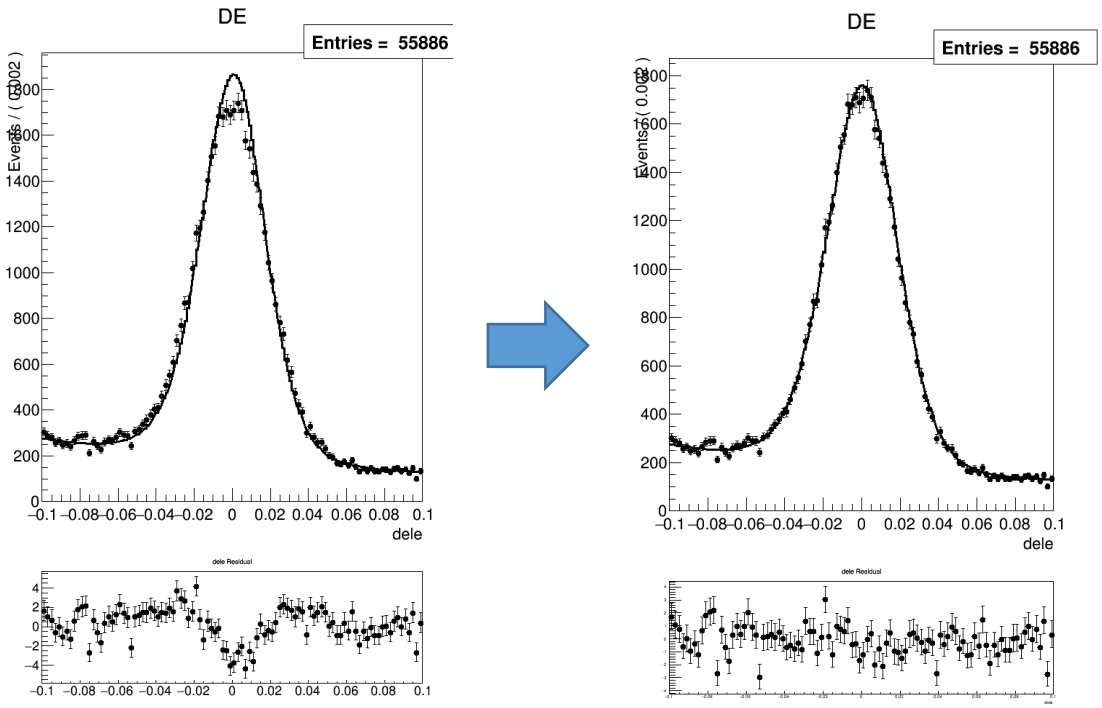


Figure 58: Effect of the KEST $\boldsymbol{\Delta E}$ modification on the fit to data for $\boldsymbol{B^+ \to \overline{D^0}\pi^+}$ ($\boldsymbol{D^0 \to K^+\pi^-\pi^0}$). Right shows the original PDF fitted to real data. Left shows the effect of modifying the $\boldsymbol{\Delta E}$ distribution with a Gaussian of width $\boldsymbol{0.007\ GeV}$.

4.5.2 A_{CP} Correction for Detector Asymmetry

Due to the nature of the detector (that it is constructed of matter) there will be a bias in the detection of matter and anti-matter particles. This is partly a result of the different nuclear cross-section between matter and anti-matter particles with the detector components. Our simulations attempt to take this into account, but they are not perfect. Another contribution to this bias is the way in which the CDC is constructed. The stereo wires are placed at a slight angle to the magnetic field, and as a result, particles with the same curvature but in opposite directions will have a different crossing angle, affecting their reconstruction efficiency slightly. To gauge the effect this has on the measured A_{CP} of $B^+ \to \overline{D^0}\pi^+$ the analysis was performed over a sideband region of data defined as $0.1 < \Delta E < 0.4$ & $5.255 < M_{BC} < 5.27$ which is composed almost entirely of continuum with an expected A_{CP} of 0.

$$A_{CP,sideband} = (1.78 \pm 0.38) \times 10^{-2}. \quad (4.6)$$

This value can be used to compensate for the detector asymmetry by fixing the background A_{CP} to this value and subtracting it from the fitted signal A_{CP} to get the true A_{CP}.

Looking at each D^0 decay channel separately in the sideband region gives:

$$A_{CP,sideband}\left(\overline{D^0} \to K^+\pi^-\right) = (0.67 \pm 0.65) \times 10^{-2}, \quad (4.7)$$

$$A_{CP,sideband}\left(\overline{D^0} \to K^+\pi^-\pi^0\right) = (2.92 \pm 0.36) \times 10^{-2}. \quad (4.8)$$

The difference in the two channels is because the detector asymmetry is a function of the momentum of the particles interacting with the detector. The three body decay will have a much higher distribution of low momentum particles than the two body decay. These values are not used in the correction (instead eq.4.6 is used) because of the simultaneous nature of the fit in both decay channels.

4.5.3 Branching Fraction Calculation

There is a significant difference in the size of systematic uncertainties that depend on individual D^0 decay modes. To ensure the accuracy of uncertainty calculation, the total branching fraction is calculated as the mean of the branching fraction in each decay mode, $\mathfrak{B}_d$ ($d \in \{\overline{\mathrm{D}^0} \to K^+\pi^-, \overline{\mathrm{D}^0} \to K^+\pi^-\pi^0\}$),

$$\begin{aligned}
\mathfrak{B} &= \mathrm{mean}\left(\mathfrak{B}_{\overline{\mathrm{D}^0}\to K^+\pi^-}, \mathfrak{B}_{\overline{\mathrm{D}^0}\to K^+\pi^-\pi^0}\right) \\
&= \mathrm{mean}\left(\frac{Y_{\overline{\mathrm{D}^0}\to K^+\pi^-}}{2 \times N_{B^+B^-} \times \epsilon_{\overline{\mathrm{D}^0}\to K^+\pi^-}}, \frac{Y_{\overline{\mathrm{D}^0}\to K^+\pi^-\pi^0}}{2 \times N_{B^+B^-} \times \epsilon_{\overline{\mathrm{D}^0}\to K^+\pi^-\pi^0}}\right) \\
&= \mathrm{mean}\left(\frac{f^S_{\overline{\mathrm{D}^0}\to K^+\pi^-} \times Y}{2 \times N_{B^+B^-} \times \epsilon_{\overline{\mathrm{D}^0}\to K^+\pi^-}}, \frac{f^S_{\overline{\mathrm{D}^0}\to K^+\pi^-\pi^0} \times Y}{2 \times N_{B^+B^-} \times \epsilon_{\overline{\mathrm{D}^0}\to K^+\pi^-\pi^0}}\right) \\
&= \frac{Y}{2 \times N_{B^+B^-}} \times \mathrm{mean}\left(\frac{f^S_{\overline{\mathrm{D}^0}\to K^+\pi^-}}{\epsilon_{\overline{\mathrm{D}^0}\to K^+\pi^-}}, \frac{f^S_{\overline{\mathrm{D}^0}\to K^+\pi^-\pi^0}}{\epsilon_{\overline{\mathrm{D}^0}\to K^+\pi^-\pi^0}}\right),
\end{aligned} \quad (4.9)$$

where Y is the total yield; Y_d is the yield for decay mode d. $N_{B^+B^-}$ is the number of B^+B^- pairs,

$$N_{B^+B^-} = \mathfrak{B}(\Upsilon(4S) \to B^+B^-) \times N_{B\overline{B}}, \quad (4.10)$$

with $N_{B\overline{B}} = (771.581 \pm 10.566) \times 10^6$ as the number of $B\overline{B}$ pairs in the full Belle dataset [33]. ϵ_d is the combined efficiency of decay mode d, defined as

$$\epsilon_d \equiv \mathfrak{B}(d) \times \epsilon_{PID}(d) \times \epsilon_{\pi^0}(d) \times \epsilon_{recon}(d) \times \epsilon_{track}(d), \quad (4.11)$$

where $\mathfrak{B}(d)$ is the branching fraction for decay mode d; $\epsilon_{PID}(d)$ is the PID efficiency correction for decay mode d; $\epsilon_{\pi^0}(d)$ is the π^0 efficiency correction for decay mode d; $\epsilon_{recon}(d)$ is the reconstructions efficiency for decay mode d (see chapter 4.2); and $\epsilon_{track}(d) = 1$, but has a systematic uncertainty associated with the number of charged tracks.

Instead of using a simple population weighted mean, to accurately address the effects of correlated and uncorrelated uncertainties of different magnitudes, a generalised weighted mean [67, 68] is used. The generalised weighted mean, $\bar{x}$, of a vector of values X is defined as

$$\bar{x} = \sigma_{\bar{x}}^2 (\mathcal{J}^T \Sigma^{-1} X), \tag{4.12}$$

where $\mathcal{J}$ is the design matrix $[1 \quad \cdots \quad 1]^T$, Σ is the covariance matrix for X, and $\sigma_{\bar{x}}^2$ is the variance of $\bar{x}$, defined by

$$\sigma_{\bar{x}}^2 = (\mathcal{J}^T \Sigma^{-1} \mathcal{J})^{-1}. \tag{4.13}$$

From Equation 4.9 and 4.12 the final branching fraction is calculated as

$$\mathfrak{B}\left(B^+ \to \overline{D^0}\pi^+\right) = \frac{Y}{2 \times N_{B^+B^-}} \times \bar{\epsilon}, \tag{4.14}$$

where $\bar{\epsilon}$ is the mean efficiency, defined from Equation 4.9 and 4.12 as,

$$\bar{\epsilon} = \sigma_{\bar{\epsilon}}^2 \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \Sigma^{-1} \left[\frac{f^S_{\overline{D^0} \to K^+\pi^-}}{\epsilon_{\overline{D^0} \to K^+\pi^-}}, \frac{f^S_{\overline{D^0} \to K^+\pi^-\pi^0}}{\epsilon_{\overline{D^0} \to K^+\pi^-\pi^0}} \right] \right). \tag{4.15}$$

The systematic uncertainties that make up the covariance matrix Σ will be discussed further in chapter 5.

4.5.4 Fit Results

Plots for the simultaneous fit of $M_{BC}, \Delta E$ and C'_{NN} are shown in Figure 59 - Figure 63. The final fitted values for signal yield and A_{CP} ($A_{CP,Raw}$ denotes the measured A_{CP} before it is corrected for detector bias) are as follows:

$$N_{sig} = 84537 \pm 306(stat.), \tag{4.16}$$

$$A_{CP,Raw} = \left(1.97 \pm 0.36(stat.)\right) \times 10^{-2}. \tag{4.17}$$

Using the methodology discussed above in chapter 4.5.3, the measured branching fraction for $B^+ \to \overline{D^0}\pi^+$ is:

$$\mathfrak{B}\left(B^+ \to \overline{D^0}\pi^+\right) = (4.53 \pm 0.02 \pm 0.14) \times 10^{-3}, \tag{4.18}$$

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

The main source of uncertainty is the systematics, the values used here are discussed further in chapter 5. The measured result is consistent with the PDG average branching fraction:

$$\mathfrak{B}\left(B^+ \to \overline{D^0}\pi^+\right)_{PDG} = (4.68 \pm 0.13) \times 10^{-3}. \tag{4.19}$$

It agrees with and supersedes the previous Belle result [69] of

$$\mathfrak{B}\left(B^+ \to \overline{D^0}\pi^+\right) = (4.34 \pm 0.10 \pm 0.23) \times 10^{-3}. \quad (4.20)$$

It is also consistent with and more precise than the BaBar result [70] of:

$$\left(B^+ \to \overline{D^0}\pi^+\right) = (4.90 \pm 0.07 \pm 0.22) \times 10^{-3}. \quad (4.21)$$

As a further cross check of the efficiencies, two more fits are performed using only one D^0 decay channel at a time.

$$\mathfrak{B}\left(B^+ \to \overline{D^0}(K^+\pi^-)\pi^+\right) = (4.59 \pm 0.02 \pm 0.15) \times 10^{-3}, \quad (4.22)$$

$$\mathfrak{B}\left(B^+ \to \overline{D^0}(K^+\pi^-\pi^0)\pi^+\right) = (4.46 \pm 0.02 \pm 0.23) \times 10^{-3}, \quad (4.23)$$

$$A_{CP,raw}\left(B^+ \to \overline{D^0}(K^+\pi^-)\pi^+\right) = \left(1.30 \pm 0.49(stat.)\right) \times 10^{-2}, \quad (4.24)$$

$$A_{CP,raw}\left(B^+ \to \overline{D^0}(K^+\pi^-\pi^0)\pi^+\right) = \left(2.78 \pm 0.53(stat.)\right) \times 10^{-2}. \quad (4.25)$$

Applying the A_{CP} correction for detector asymmetry from section 4.5.2 to the result for the full dataset gives:

$$A_{CP}\left(B^+ \to \overline{D^0}\pi^+\right) = (0.19 \pm 0.36 \pm 0.57) \times 10^{-2}. \quad (4.26)$$

The corrected A_{CP} for the individual channel fits are:

$$A_{CP}\left(B^+ \to \overline{D^0}(K^+\pi^-)\pi^+\right) = (0.66 \pm 0.49 \pm 0.89) \times 10^{-2}, \quad (4.27)$$

$$A_{CP}\left(B^+ \to \overline{D^0}(K^+\pi^-\pi^0)\pi^+\right) = (-0.15 \pm 0.53 \pm 0.68) \times 10^{-2}, \quad (4.28)$$

all of which are consistent with zero A_{CP} and with the PDG average of:

$$A_{CP,PDG} = (-0.7 \pm 0.7) \times 10^{-2}. \quad (4.29)$$

The result is also in agreement with both the previous Belle result [71] of

$$A_{CP}\left(B^+ \to \overline{D^0}\pi^+\right) = \left(-0.8 \pm 0.8(stat.)\right) \times 10^{-2}, \quad (4.30)$$

which it supersedes, and the LHCb study [72] of:

$$A_{CP}\left(B^+ \to \overline{D^0}\pi^+\right) = (-0.6 \pm 0.5 \pm 1.0) \times 10^{-2}. \quad (4.31)$$

An additional cross-check is made by performing a full fit on datasets split between SVD1 and SVD2. These fit results are also consistent with the main fit as well as the PDG average values.

$$\mathfrak{B}\left(B^+ \to \overline{D^0}\pi^+\right)_{SVD1} = \left(4.55 \pm 0.04(stat.)\right) \times 10^{-3}, \quad (4.32)$$

$$\mathfrak{B}\left(B^+ \to \overline{D^0}\pi^+\right)_{SVD2} = \left(4.52 \pm 0.02(stat.)\right) \times 10^{-3}, \quad (4.33)$$

$$A_{CP}\left(B^+ \to \overline{D^0}\pi^+\right)_{SVD1} = \left(0.5 \pm 1.2(stat.)\right) \times 10^{-2}, \quad (4.34)$$

$$A_{CP}\left(B^+ \to \overline{D^0}\pi^+\right)_{SVD2} = \left(0.1 \pm 0.8(stat.)\right) \times 10^{-2}. \quad (4.35)$$

As a double check on any potential discrepancy between the C'_{NN} PDFs and the C'_{NN} distribution in data, a comparison is made between the regular 3D fit and a 2D fit in M_{BC} and ΔE (without fitting to the C'_{NN} distribution). The difference in Yield was found to be $+54\ events$ (a deviation of 0.06%) and in A_{CP} -0.02×10^{-2}. These differences are within the bounds of the systematic uncertainties associated with the fitting bias and C'_{NN} calibration factors of $\pm 0.21\%$ and $\pm 0.02 \times 10^{-2}$ respectively, showing that the C'_{NN} PDFs are reasonable for use in the fit.

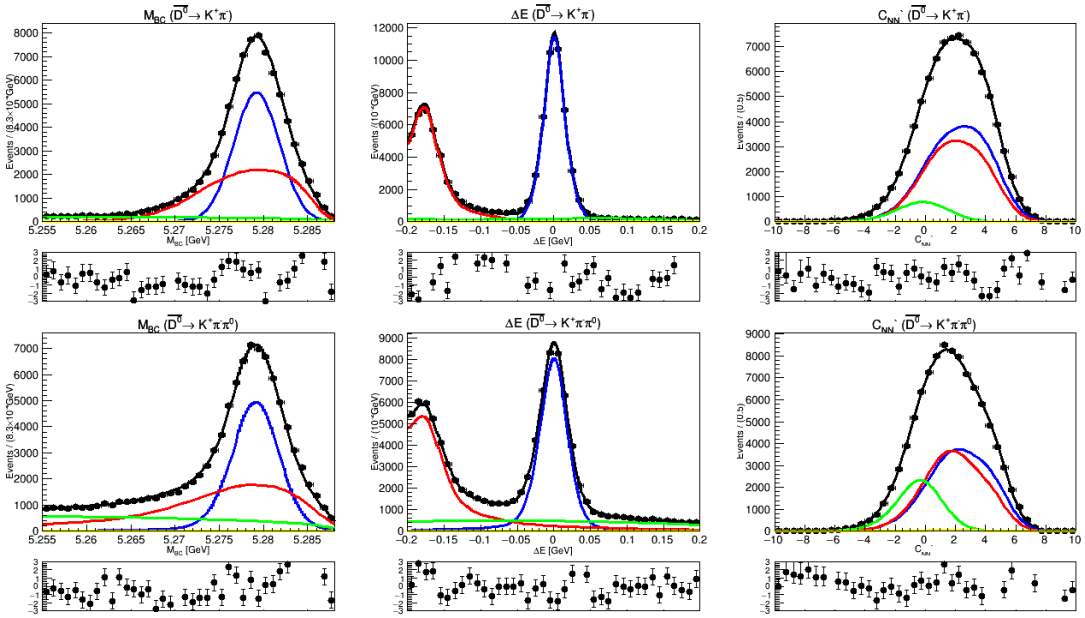


Figure 59: $B^+ \to \overline{D^0}\pi^+$ fit plots in the signal box region ($5.275 < M_{BC} < 5.285, -0.12 < \Delta E < 0.07, -1 < C'_{NN} < 6$) for M_{BC} (left), ΔE (middle) and C'_{NN} (right) split into $\overline{D^0} \to K^+\pi^-$ decays (top) and $\overline{D^0} \to K^+\pi^-\pi^0$ (bottom). The Blue line shows signal PDF, red shows generic $B\overline{B}$ background PDF, green shows continuum background PDF, yellow shows charmless $B\overline{B}$ background PDF, black is sum of all PDFs, points are data.

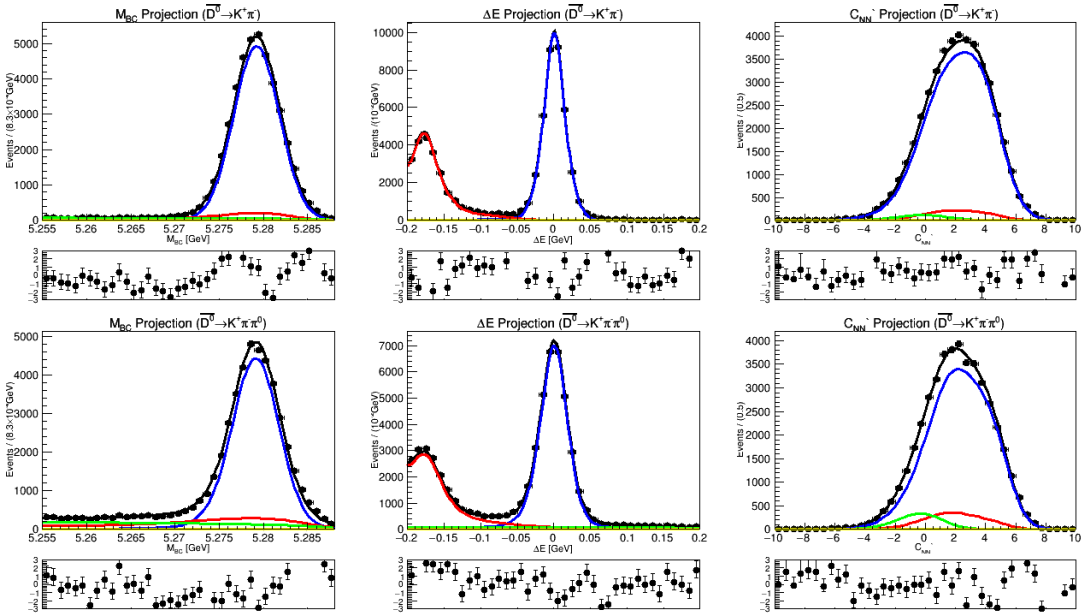


Figure 60: $B^+ \to \overline{D^0}\pi^+$ fit plots in the full fitting box for M_{BC} (left), ΔE (middle) and C'_{NN} (right) split into $\overline{D^0} \to K^+\pi^-$ decays (top) and $\overline{D^0} \to K^+\pi^-\pi^0$ (bottom). The Blue line shows signal PDF, red shows generic $B\overline{B}$ background PDF, green shows continuum background PDF, yellow shows charmless $B\overline{B}$ background PDF, black is sum of all PDFs, points are data.

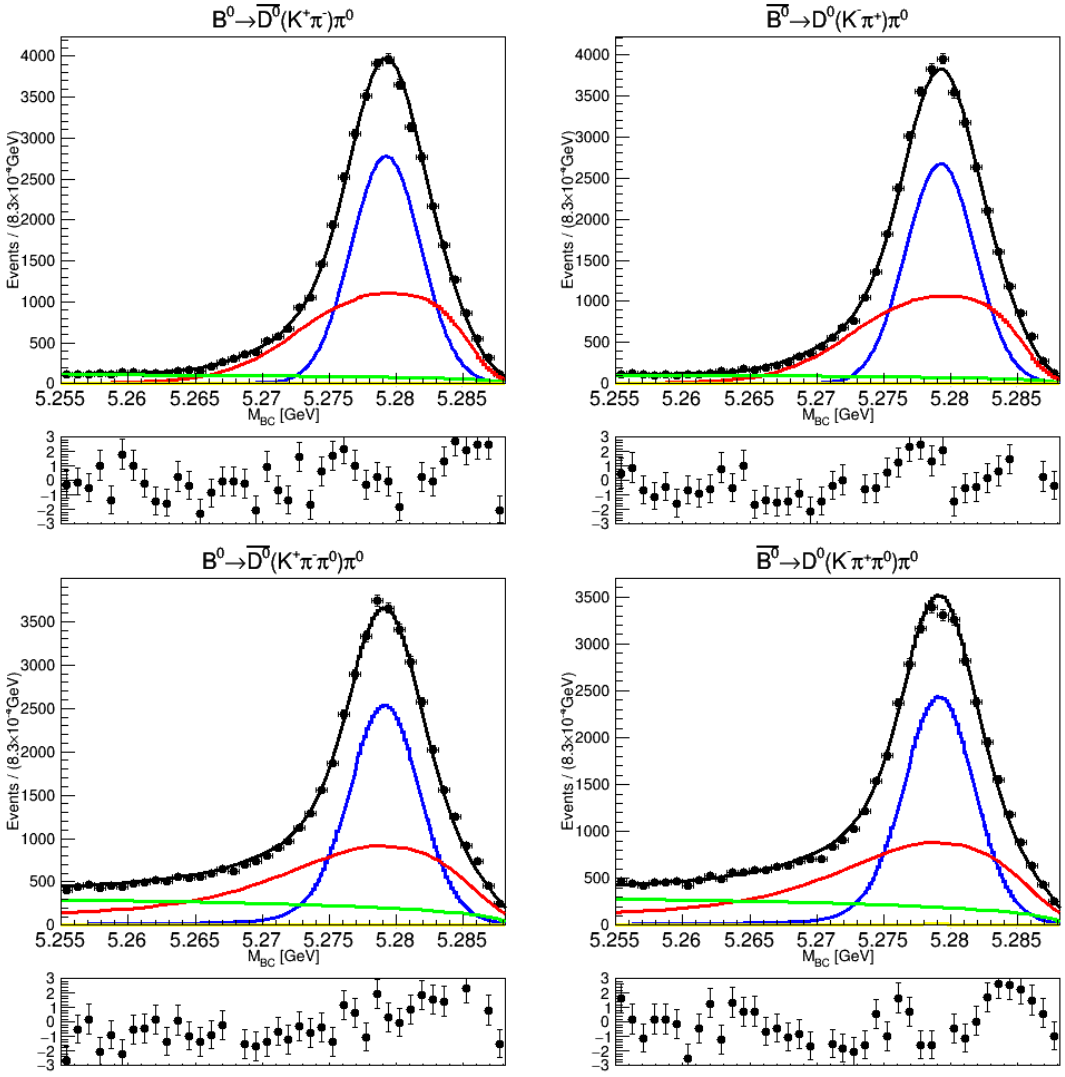


Figure 61: $B^+ \to \overline{D^0}\pi^+$ fit plots for M_{BC}, showing $B^+ \to \overline{D^0}(K^+\pi^-)\pi^+$ (top left), $B^- \to D^0(K^-\pi^+)\pi^-$ (top right), $B^+ \to \overline{D^0}(K^+\pi^-\pi^0)\pi^+$ (bottom left), $B^- \to D^0(K^-\pi^+\pi^0)\pi^-$ (bottom right). The Blue line shows signal PDF, red shows generic $B\overline{B}$ background PDF, green shows continuum background PDF, yellow shows charmless $B\overline{B}$ background PDF, black is sum of all PDFs, points are data.

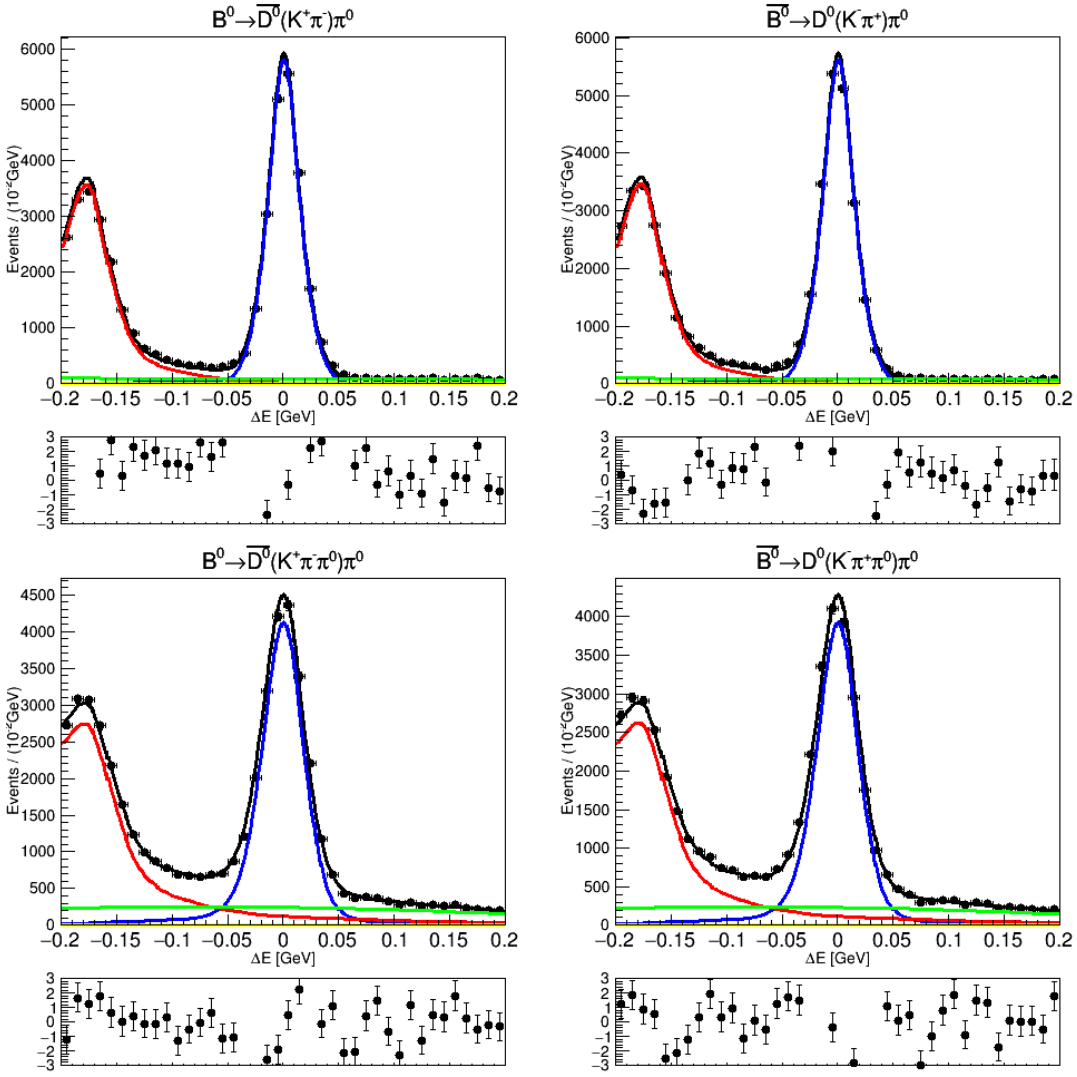


Figure 62: $B^+ \to \overline{D^0}\pi^+$ fit plots for ΔE, showing $B^+ \to \overline{D^0}(K^+\pi^-)\pi^+$ (top left), $B^- \to D^0(K^-\pi^+)\pi^-$ (top right), $B^+ \to \overline{D^0}(K^+\pi^-\pi^0)\pi^+$ (bottom left), $B^- \to D^0(K^-\pi^+\pi^0)\pi^-$ (bottom right). The Blue line shows signal PDF, red shows generic $B\overline{B}$ background PDF, green shows continuum background PDF, yellow shows charmless $B\overline{B}$ background PDF, black is sum of all PDFs, points are data.

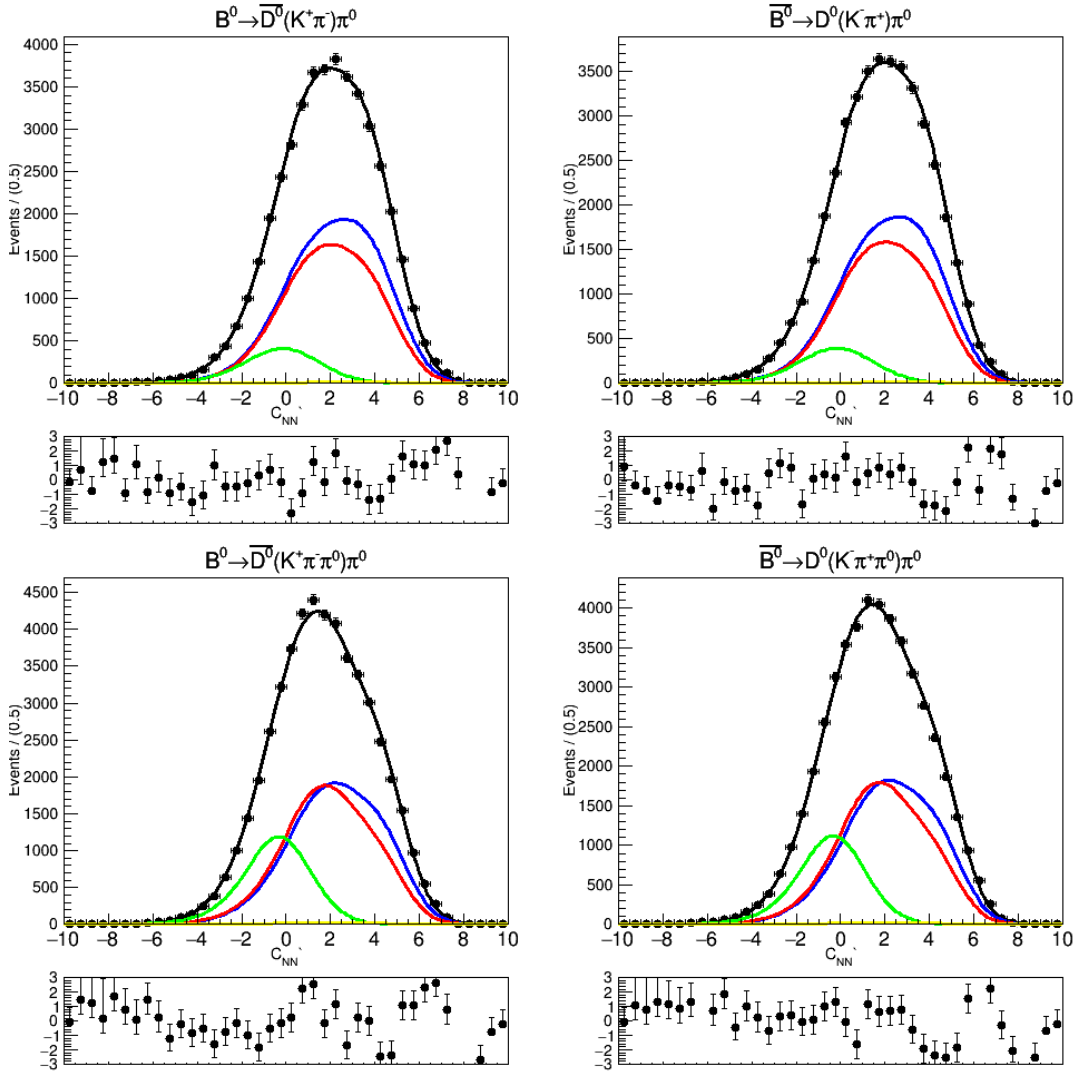


Figure 63: $B^+ \to \overline{D^0}\pi^+$ fit plots for C'_{NN}, showing $B^+ \to \overline{D^0}(K^+\pi^-)\pi^+$ (top left), $B^- \to D^0(K^-\pi^+)\pi^-$ (top right), $B^+ \to \overline{D^0}(K^+\pi^-\pi^0)\pi^+$ (bottom left), $B^- \to D^0(K^-\pi^+\pi^0)\pi^-$ (bottom right). The Blue line shows signal PDF, red shows generic $B\overline{B}$ background PDF, green shows continuum background PDF, yellow shows charmless $B\overline{B}$ background PDF, black is sum of all PDFs, points are data.

5 Systematics

This chapter is a discussion on the sources of systematic uncertainties and how they are estimated. The largest contributions to uncertainty are the D^0 sub-decay property measurements, the π^0 reconstruction efficiency, the PID correction efficiency and the A_{CP} detector bias. Uncertainties on branching fraction and A_{CP} are discussed individually.

5.1 Branching Fraction Systematic Uncertainties

Because a generalised weighted mean is used to calculate the final branching fraction in Equation 4.9 (chapter 4.5.3), the total systematic uncertainty cannot simply be calculated by adding all the individual terms in quadrature. The mean efficiency, $\bar{\epsilon}$, combines the terms defining the efficiency of the individual D^0 decay modes as a generalised weighted mean. The uncertainty for the mean efficiency is calculated from the covariance matrix of these terms, Σ, as

$$\sigma_{\bar{\epsilon}}^2 = (\mathcal{J}^T \Sigma^{-1} \mathcal{J})^{-1}, \tag{5.1}$$

as discussed in chapter 4.5.3. The systematic uncertainties related to these are: the charged track efficiency; the π^0 efficiency correction; the PID efficiency correction; the reconstruction efficiency; and the D^0 branching fraction of each mode.

From Equation (4.14), the mean efficiency is combined as a simple linear equation with the yield and number of B events, so the uncertainty on mean efficiency is added in quadrature with the remaining uncertainties in the standard way.

5.1.1 Charged Track Efficiency

Uncertainty from tracking efficiency is based on a previous study of high momentum ($P_t > 200 MeV/c$) tracks [73] (as used in [74]). Tracking efficiency is calculated as the ratio of partial and fully reconstructed D^* decays and compared between MC and data. The results of this study give a 0.35% uncertainty per charged track.

5.1.2 π^0 efficiency

Uncertainty from π^0 reconstruction efficiency is based on a previous study of π^0 efficiency from $\tau^- \to \pi^-\pi^0\nu_\tau$ [75] (as used in [76]). This found the ratio of data to MC π^0 efficiency to be $0.96 \pm 0.02\%$ per π^0.

5.1.3 Particle ID (PID)

A package provided by the PID group is used to calculate the correction for efficiency between MC and data [50] (as used in [74]). This is based off a previous study of the decay $D^{*+} \to D^0\pi_s^+$ with $D^0 \to K^-\pi^+$ where the s subscript denotes a low momentum particle. This slow pion means the decay can be reconstructed with a high signal to noise ratio without the use of PID selection criteria. The difference in PID efficiency between MC and data is measured in bins of momentum and polar angle. This data is compiled into a package to allow for the calculation of the average correction to PID efficiency for angle and momentum distributions of a specific decay. Efficiencies are calculated separately for $B^0 \to \overline{D^0}\pi^0$ and $B^+ \to \overline{D^0}\pi^+$ decays, and for each D^0 sub-decay mode. The correction for each particle is shown in Table 11 and Table 12.

Table 11: PID correction factors for $\boldsymbol{B^0 \to \overline{D^0}\pi^0}$.

Particle	$\overline{D^0} \to K^+\pi^-$	$\overline{D^0} \to K^+\pi^-\pi^0$
K^+	0.999 ± 0.009	1.005 ± 0.009
π^-	0.956 ± 0.009	0.970 ± 0.009
Total	0.954 ± 0.012	0.974 ± 0.013

Table 12: PID correction factors for $\boldsymbol{B^+ \to \overline{D^0}\pi^+}$.

Particle	$\overline{D^0} \to K^+\pi^-$	$\overline{D^0} \to K^+\pi^-\pi^0$
K^+	0.999 ± 0.009	1.006 ± 0.009
π^-	0.955 ± 0.009	0.969 ± 0.009
π^+ (D^0 daughter)	0.954 ± 0.009	0.954 ± 0.008
Total	0.910 ± 0.020	0.929 ± 0.020

5.1.4 D^0 Sub-Decay Branching Fractions

Uncertainty associated with the $\overline{D^0} \to K^+\pi^-$ and $\overline{D^0} \to K^+\pi^-\pi^0$ branching fractions, taken from the PDG average as $(3.95 \pm 0.03) \times 10^{-2}$ and $(14.4 \pm 0.5) \times 10^{-2}$ respectively. Based on the number of events reconstructed in each decay

mode, the combined uncertainty is 1.94% for $B^0 \to \overline{D^0}\pi^0$ and 1.76% for $B^+ \to \overline{D^0}\pi^+$.

5.1.5 Decay Reconstruction Efficiencies

Uncertainty associated with the reconstruction efficiency is measured using signal MC datasets. This is based on the Poisson statistics of the size of the dataset used. Efficiencies measured are:

$$\epsilon_{B^0 \to \overline{D^0}(K^+\pi^-)\pi^0} = 27.53 \pm 0.04\%, \tag{5.2}$$

$$\epsilon_{B^0 \to \overline{D^0}(K^+\pi^-\pi^0)\pi^0} = 9.43 \pm 0.02\%, \tag{5.3}$$

$$\epsilon_{B^+ \to \overline{D^0}(K^+\pi^-)\pi^+} = 33.08 \pm 0.04\%, \tag{5.4}$$

$$\epsilon_{B^+ \to \overline{D^0}(K^+\pi^-\pi^0)\pi^+} = 9.05 \pm 0.02\%, \tag{5.5}$$

5.1.6 Mean Efficiency

The five terms discussed in sections (5.1.1 to 5.1.5) are combined in a generalised weighted mean in the branching fraction calculation (Equation 4.9). The covariance matrix for these terms in each D^0 decay mode, Σ, is constructed considering correlations and is found to be

$$\Sigma_{B^0 \to \overline{D^0}\pi^0} = \begin{bmatrix} 1.48 & 2.40 \\ 2.40 & 6.77 \end{bmatrix}, \qquad \Sigma_{B^+ \to \overline{D^0}\pi^+} = \begin{bmatrix} 1.17 & 1.05 \\ 1.05 & 4.03 \end{bmatrix}, \tag{5.6}$$

where the matrix elements are related to the uncertainties in sections (5.1.1 to 5.1.5). Σ_{11} is the sum of the variance for all five terms in $\overline{D^0} \to K^+\pi^-$, and Σ_{22} is the same for $\overline{D^0} \to K^+\pi^-\pi^0$.

$$\begin{aligned} \Sigma_{12} = \Sigma_{21} = &\sqrt{\sigma^2_{PID}(K^+\pi^-) + \sigma^2_{\pi^0}(K^+\pi^-) + \sigma^2_{chg}(K^+\pi^-)} \\ &\times \sqrt{\sigma^2_{PID}(K^+\pi^-\pi^0) + \sigma^2_{\pi^0}(K^+\pi^-\pi^0) + \sigma^2_{chg}(K^+\pi^-\pi^0)}, \end{aligned} \tag{5.7}$$

is the product of the correlated standard deviations for both modes. The correlated terms are π^0 efficiency, PID correction efficiency and charged track efficiency. These 3 efficiencies are evaluated using the same measurements in both modes, so are treated as fully correlated.

Inverting the covariance matrices gives

$$\Sigma^{-1}_{B^0 \to \overline{D^0}\pi^0} = \begin{bmatrix} 1.60 & -0.57 \\ -0.57 & 0.35 \end{bmatrix}, \qquad \Sigma^{-1}_{B^+ \to \overline{D^0}\pi^+} = \begin{bmatrix} 1.12 & -0.29 \\ -0.29 & 0.32 \end{bmatrix}. \tag{5.8}$$

From this the uncertainty on mean efficiency, $\bar{\epsilon}$, is calculated as

$$\sigma_{\bar{\epsilon}}^2 = (\mathcal{J}^T \Sigma^{-1} \mathcal{J})^{-1}, \quad (5.9)$$

where $\mathcal{J}$ is the design matrix, $\mathcal{J} = [1 \quad 1]^T$. The uncertainty on $\bar{\epsilon}$ is calculated to be 2.44% for $B^0 \to \overline{D^0}\pi^0$ and 2.54% for $B^+ \to \overline{D^0}\pi^+$.

5.1.7 Number of $B\overline{B}$ pairs

This is the uncertainty associated with the measured number of $B\overline{B}$ pairs in the full dataset collected at Belle. This is 1.37% based on a dataset of $(771.581 \pm 10.566) \times 10^6$ $B\overline{B}$ pairs. This is based on the total integrated luminosity calculated utilising Bhabha decays in the barrel section of the detector [33].

5.1.8 $\mathfrak{B}(\Upsilon(4S))$

This is the uncertainty for the fraction of $B\overline{B}$ events at Belle that are neutral or charged. This is taken from the PDG average values [3] of

$$\mathfrak{B}\left(\Upsilon(4S) \to B^0\overline{B^0}\right) = 0.486 \pm 0.006, \quad (5.10)$$

$$\mathfrak{B}(\Upsilon(4S) \to B^+B^-) = 0.514 \pm 0.006. \quad (5.11)$$

5.1.9 DCS mode correction

This is the uncertainty associated with the correction to branching fraction resulting from the doubly Cabibbo suppressed decays as calculated in chapter 3.1. This is a 0.01% uncertainty for $B^0 \to \overline{D^0}\pi^0$ and a 0.02% uncertainty for $B^+ \to \overline{D^0}\pi^+$.

5.1.10 Fitting Bias

The bias found in the MC ensemble tests (see chapter 3.7.2) is quoted as an uncertainty. This is 0.6% and 0.2% for the $B^0 \to \overline{D^0}\pi^0$ and $B^+ \to \overline{D^0}\pi^+$ decays respectively.

5.1.11 $\overline{\mathbf{D^0}} \to K^+\pi^-$: $\overline{\mathbf{D^0}} \to K^+\pi^-\pi^0$ Yield Ratio

As the yield ratio between the D^0 decays is set based on the expected yields from MC and held as a fixed constraint during the fit, it contributes a systematic uncertainty. To take this into account 2 fits are performed, varying the fixed value

up and down by 1σ and taking the difference in signal yield as the uncertainty. This is found to be $^{+0.38\%}_{-0.31\%}$ for $B^0 \to \overline{D^0}\pi^0$ and $^{+0.19\%}_{-0.08\%}$ for $B^+ \to \overline{D^0}\pi^+$.

5.1.12 Fixed Background A_{CP}

Uncertainty from the fixed background A_{CP} constraints in the fit are estimated by varying them $\pm 1\sigma$ (using the uncertainty calculated in Section 5.2.3) and comparing the A_{CP} in the resultant fits. This is found to be 0.01% for $B^0 \to \overline{D^0}\pi^0$ and 0.05% for $B^+ \to \overline{D^0}\pi^+$.

5.1.13 C'_{NN} Calibration factors

To measure the uncertainty associated with the fixed C'_{NN} PDFs used in the fit, fitting is performed with and without the calibration factors applied to these PDFs. The difference between the yields of these fits is quoted as the uncertainty. This is found to be 0.34% for $B^0 \to \overline{D^0}\pi^0$ and 0.06% for $B^+ \to \overline{D^0}\pi^+$.

5.1.14 Modification of $\overline{\mathrm{D}^0} \to K^+\pi^-\pi^0$ $M_{BC} \times \Delta E$ PDF

The uncertainty from the ΔE correction to the $\overline{\mathrm{D}^0} \to K^+\pi^-\pi^0$ $M_{BC} \times \Delta E$ KEST PDF (described in Section 4.5.1) is obtained by comparing a fit using the corrected and uncorrected PDF. The difference in returned yield between the 2 fits is quoted as the uncertainty. This is found to be 0.63% for $B^0 \to \overline{D^0}\pi^0$ and 0.24% for $B^+ \to \overline{D^0}\pi^+$.

5.1.15 KEST fixed PDFs

Generic $B\overline{B}$, charmless $B\overline{B}$ and $\overline{D^0} \to K^+\pi^-\pi^0$ signal $M_{BC} \times \Delta E$ PDFs all use a fixed 2D KEST PDF. To estimate the uncertainty from this, an ensemble test is performed running 1000 fits over data, with each fit using a modified KEST PDF. The PDFs in the ensemble are generated by taking 40x40 binned histogram of MC dataset (used to generate the KEST originally) and applying a Gaussian fluctuation based on bin uncertainty to the corresponding area of the KEST histogram (as the KEST PDF is stored as a 320x320 histogram, each fluctuation will be applied to an 8x8 bin area). The uncertainty is taken as the standard deviation of this ensemble test. This is found to be 0.58% for $B^0 \to \overline{D^0}\pi^0$ and 0.05% for $B^+ \to \overline{D^0}\pi^+$.

5.1.16 Fixed Charmless $B\overline{B}$ Background Yields

Because the normalisation of the charmless $B\overline{B}$ background is fixed in the fit, uncertainty is estimated by varying this by $\pm 1\sigma$ and looking at the difference in fitted yield. This has negligible effect on the $B^+ \to \overline{D^0}\pi^+$ mode (<0.01%).
For $B^0 \to \overline{D^0}\pi^0$ this is found to be 0.26%.

5.2 A_{CP} Systematic Uncertainties

5.2.1 Fitting Bias

The uncertainty from the GSIM MC ensemble (chapter 3.7.2) tests is quoted. This is $\pm 0.03 \times 10^{-2}$ and $\pm 0.02 \times 10^{-2}$ for the $B^0 \to \overline{D^0}\pi^0$ and $B^+ \to \overline{D^0}\pi^+$ decays respectively.

5.2.2 D^0 Sub-Decay Branching Fractions

Uncertainty associated with the $\overline{D^0} \to K^+\pi^-$ and $\overline{D^0} \to K^+\pi^-\pi^0$ A_{CP}, taken from the PDG average is $(0.2 \pm 0.5) \times 10^{-2}$ and $(0.1 \pm 0.5) \times 10^{-2}$ respectively. Based on the number of events reconstructed in each decay mode, the combined uncertainty is 0.35×10^{-2} for $B^0 \to \overline{D^0}\pi^0$ and 0.35×10^{-2} for $B^+ \to \overline{D^0}\pi^+$.

5.2.3 A_{CP} Detector Bias

Uncertainty in the A_{CP} correction for detector bias is based off the statistics in the sideband used to measure it, as discussed in chapter 4.5.2. It is noted that this sideband measurement only contains continuum, and it is possible that $B\overline{B}$ events may differ based on different particle composition/momenta. It is also possible that kinematics may vary slightly between the sideband and signal peak. The comparison of the $B^+ \to \overline{D^0}\pi^+$ result to the theoretical expectation of no A_{CP} indicates this sideband correction is accurate, but to account for this data driven validation, the difference between the measured value and the theoretically expected value of 0 ($\pm 0.19 \times 10^{-2}$) is added in a quadratic sum. For $B^0 \to \overline{D^0}\pi^0$ this is $\pm 0.66 \times 10^{-2}$ and for $B^+ \to \overline{D^0}\pi^+$ this is $\pm 0.42 \times 10^{-2}$.

5.2.4 Fixed Background A_{CP}

Uncertainty from the fixed background A_{CP} constraints in the fit are estimated by varying them $\pm 1\sigma$ (using the uncertainty calculated in Section 5.2.3) and comparing the A_{CP} in the resultant fits. This is found to be 0.49×10^{-2} for $B^0 \to \overline{D^0}\pi^0$ and 0.06×10^{-2} for $B^+ \to \overline{D^0}\pi^+$. This uncertainty is correlated with the A_{CP} Detector Bias so it added linearly to it.

5.2.5 $\overline{D^0} \to K^+\pi^-$: $\overline{D^0} \to K^+\pi^-\pi^0$ Yield Ratio

Same as for yield, this is estimated by performing 2 fits varying the fixed yield ratio up and down by 1σ and taking the difference in A_{CP} as the uncertainty. This is found to be $^{+0.03}_{-0.02} \times 10^{-2}$for $B^0 \to \overline{D^0}\pi^0$ and $< 0.001 \times 10^{-2}$ for $B^+ \to \overline{D^0}\pi^+$.

5.2.6 C'_{NN} Calibration factors

As with yield, fitting is performed with and without the calibration factors applied to these PDFs. The difference between the A_{CP} of these fits is quoted as the uncertainty. This is found to be 0.06×10^{-2} for $B^0 \to \overline{D^0}\pi^0$ and 0.004×10^{-2} for $B^+ \to \overline{D^0}\pi^+$.

5.2.7 Modification of $\overline{D^0} \to K^+\pi^-\pi^0$ $M_{BC} \times \Delta E$ PDF

As with yield, estimated by fitting with and without the ΔE smearing correction. This is found to be 0.06×10^{-2} for $B^0 \to \overline{D^0}\pi^0$ and 0.006×10^{-2} for $B^+ \to \overline{D^0}\pi^+$.

5.2.8 KEST fixed PDFs

This is found using the same methodology as with yield (see chapter 5.1.15).
This is found to be 0.15×10^{-2} for $B^0 \to \overline{D^0}\pi^0$ and 0.15×10^{-2} for $B^+ \to \overline{D^0}\pi^+$.

5.2.9 Fixed Charmless $B\overline{B}$ Background Yields

Uncertainty is estimated by varying this by $\pm 1\sigma$ and taking half the difference in fitted yield. This was found to have a negligible ($< 0.001 \times 10^{-2}$) for both $B^0 \to \overline{D^0}\pi^0$ and $B^+ \to \overline{D^0}\pi^+$.

Table 13: Table of systematic uncertainties for branching fraction measurements.

	$B^0 \to \overline{D^0}\pi^0$	$B^+ \to \overline{D^0}\pi^+$
No. $B\overline{B}$	1.37%	1.37%
$\mathfrak{B}(\Upsilon(4S))$	1.23%	1.17%
DCS mode correction	0.01%	0.02%
Fit bias	0.60%	0.20%
Mean efficiency	2.44%	2.54%
$\overline{D^0} \to K^+\pi^-\pi^0$: $\overline{D^0} \to K^+\pi^-$ ratio	$+0.31\%$ -0.38%	$+0.19\%$ -0.08%
A_{CP} detector bias (backgrounds)	0.01%	0.05%
Calibration Factors (C'_{NN})	0.34%	0.06%
Modified KEST signal ($M_{BC}, \Delta E$)	0.63%	0.24%
KEST PDFs	0.58%	0.05%
Fixed Charmless $B\overline{B}$ Yields	$+0.26\%$ -0.27%	$< 0.01\%$
Total	$\mathbf{\pm 3.28\%}$	$\mathbf{\pm 3.13\%}$

Table 14: Table of systematic uncertainties for A_{CP} measurements. *denotes correlated variables.

	$B^0 \to \overline{D^0}\pi^0$ $(\times 10^{-2})$	$B^+ \to \overline{D^0}\pi^+$ $(\times 10^{-2})$
Fit bias	0.03	0.02
$\overline{D^0}$ decay A_{CP}	0.35	0.35
A_{CP} detector bias (signal)*	0.66	0.42
A_{CP} detector bias (backgrounds)*	$+0.49$ -0.49	$+0.03$ -0.03
$\overline{D^0} \to K^+\pi^-\pi^0$: $\overline{D^0} \to K^+\pi^-$ ratio	$+0.03$ -0.02	< 0.01
Calibration Factors (C'_{NN})	0.06	< 0.01
Modified KEST signal ($M_{BC}, \Delta E$)	0.06	< 0.01
KEST PDFs	0.15	< 0.01
Fixed Charmless $B\overline{B}$ Yields	< 0.01	< 0.01
Total	$\mathbf{\pm 1.22}$	$\mathbf{\pm 0.57}$

6 $B^0 \to \overline{D^0}\pi^0$ Results

This section details the results of the $B^0 \to \overline{D^0}\pi^0$ analysis on real data gathered at by the Belle experiment. Using the fitting methodology described previously in chapter 3.6, from the complete Belle dataset of $(771 \pm 10.57) \times 10^6$ $B\overline{B}$ pairs a signal yield of 4448 ± 97 events is obtained. Figure 65 shows the results of these fits to the full datasets. Figure 64 shows the fits with data projected into signal enhanced regions. Figure 66-Figure 68 show the results with the different charged kaon tags employed to distinguish $\bar{B}^0 \to D^0\pi^0$ from $B^0 \to \bar{D}^0\pi^0$.

6.1 Full Box Open Fit

The final fitted values for signal yield and A_{CP} are as follows:

$$N_{sig} = 4448 \pm 97(stat.), \tag{6.1}$$

$$A_{CP,Raw} = \left(1.48 \pm 2.05(stat.)\right) \times 10^{-2}. \tag{6.2}$$

Branching fraction is calculated using the method described in chapter 4.5.3,

$$\mathfrak{B}\left(B^+ \to \overline{D^0}\pi^+\right) = \frac{Y}{2 \times N_{B^0\overline{B^0}}} \times \bar{\epsilon}, \tag{6.3}$$

where Y is the total yield, $N_{B^0\overline{B^0}}$ is the number of $B^0\overline{B^0}$ pairs in the full Belle dataset, and $\bar{\epsilon}$ is the mean efficiency,

$$\bar{\epsilon} = \sigma_{\bar{\epsilon}}^2 \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \Sigma^{-1} \left[\frac{f^S_{\overline{D^0}\to K^+\pi^-}}{\epsilon_{\overline{D^0}\to K^+\pi^-}}, \frac{f^S_{\overline{D^0}\to K^+\pi^-\pi^0}}{\epsilon_{\overline{D^0}\to K^+\pi^-\pi^0}} \right] \right). \tag{6.4}$$

The mean efficiency and its uncertainty are calculated from the covariance matrix of the D^0 mode dependent efficiency terms as discussed in chapter 5.1.

From this, the measured branching fraction for $B^0 \to \overline{D^0}\pi^0$ is:

$$\mathfrak{B}\left(B^0 \to \overline{D^0}\pi^0\right) = (2.69 \pm 0.06 \pm 0.09) \times 10^{-4}, \tag{6.5}$$

where the first uncertainty is statistical and the second is systematic. This result is consistent with the current PDG average value [3] of:

$$\mathfrak{B}\left(B^0 \to \overline{D^0}\pi^0\right)_{\boldsymbol{PDG}} = (2.63 \pm 0.14) \times 10^{-4}. \tag{6.6}$$

It is in agreement with the previous Belle result [15] of

$$\mathfrak{B}\left(B^0 \to \overline{D^0}\pi^0\right) = \left(2.25 \pm 0.14(stat.) \pm 0.35(syst.)\right) \times 10^{-4} \tag{6.7}$$

and supersedes it. It also agrees with the measurement by the BaBar experiment [16] of

$$\mathfrak{B}\left(B^0 \to \overline{D^0}\pi^0\right) = \left(2.69 \pm 0.09(stat.) \pm 0.12(syst.)\right) \times 10^{-4}, \quad (6.8)$$

but achieves better statistical and systematic precision.

6.2 A_{CP} Correction for Detector Asymmetry

As discussed in chapter 4.5.2, a correction needs to be made to the measured value of A_{CP} to account for detector bias. The methodology to measure this correction for the $B^0 \to \overline{D^0}\pi^0$ decay is the same as for $B^+ \to \overline{D^0}\pi^+$. A sideband region of data defined as $0.1 < \Delta E < 0.4$ & $5.255 < M_{BC} < 5.27$ which is composed almost entire of continuum with an expected A_{CP} of 0.

$$A_{CP,sideband} = (1.02 \pm 0.64) \times 10^{-2}. \quad (6.9)$$

Using this value, the detector asymmetry can be compensated for by fixing the background A_{CP} and subtracted it from the fitted signal A_{CP} to get the true A_{CP}.

Looking at each D^0 decay channel separately in sideband gives:

$$A_{CP,sideband}\left(\overline{D^0} \to K^+\pi^-\right) = (-0.42 \pm 1.23) \times 10^{-2}, \quad (6.10)$$

$$A_{CP,sideband}\left(\overline{D^0} \to K^+\pi^-\pi^0\right) = (2.16 \pm 0.60) \times 10^{-2}. \quad (6.11)$$

Taking the A_{CP} correction for detector asymmetry and fitter bias previous discussed into account

$$A_{CP}\left(B^0 \to \overline{D^0}\pi^0\right) = (0.10 \pm 2.05 \pm 1.29) \times 10^{-2}. \quad (6.12)$$

This is consistent with the combined Belle/BaBar analysis of $B^0 \to D_{CP}^{(*)}h^0$ (a related measurement with different D^0 final states, for which similar values are expected) [25] which found

$$C\left(B^0 \to D_{CP}^{(*)}h^0\right) = (-2 \pm 8) \times 10^{-2}, \quad (6.13)$$

where $C(B \to f) = -A_{CP}(B \to f)$. This result is also consistent with the theoretically expected value of no A_{CP}.

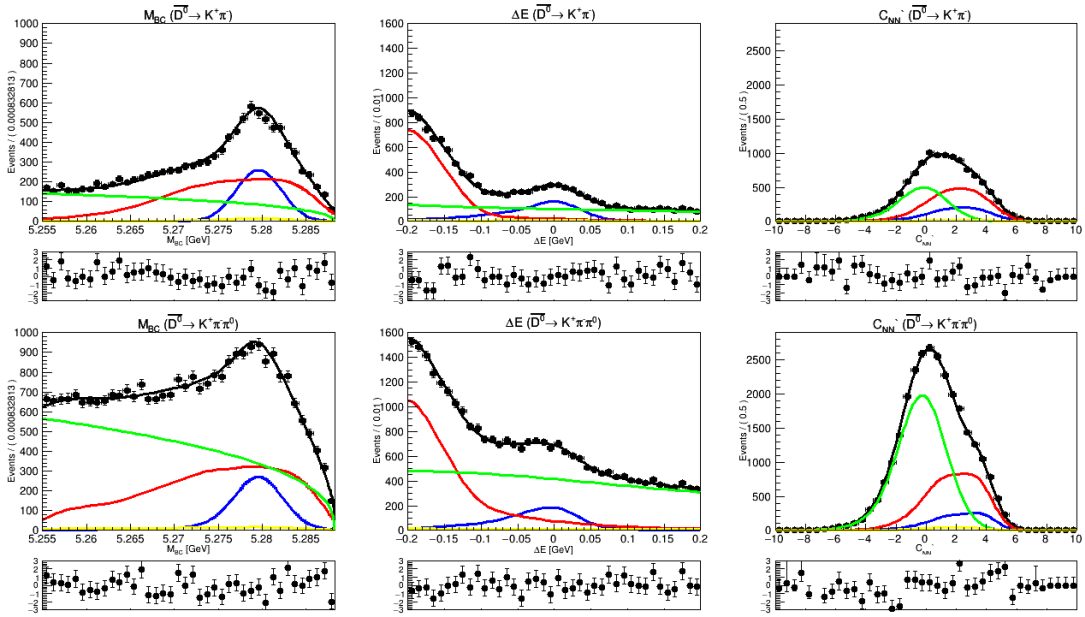


Figure 64: $B^0 \to \overline{D^0}\pi^0$ fit plots in the signal box region ($5.275 < M_{BC} < 5.285, -0.12 < \Delta E < 0.07, -1 < C'_{NN} < 6$) for M_{BC} (left), ΔE (middle) and C'_{NN} (right) split into $\overline{D^0} \to K^+\pi^-$ decays (top) and $\overline{D^0} \to K^+\pi^-\pi^0$ (bottom). Blue line shows signal PDF, red shows generic $B\overline{B}$ background PDF, green shows continuum background PDF, yellow shows charmless $B\overline{B}$ background PDF, black is sum of all PDFs, points are data.

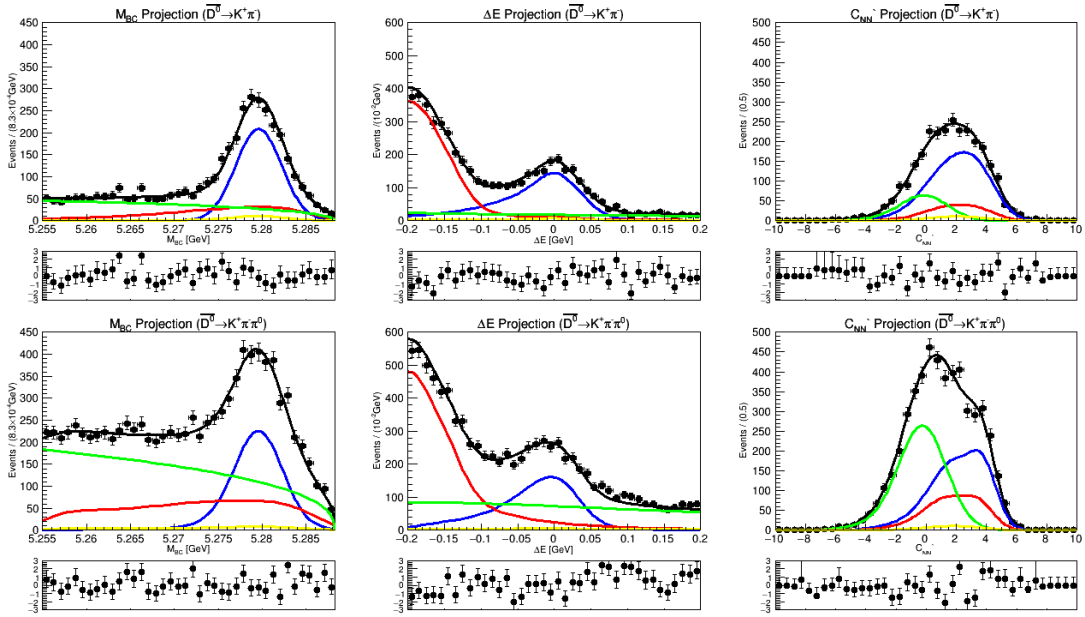


Figure 65: $B^0 \to \overline{D^0}\pi^0$ fit plots in the full fitting box for M_{BC} (left), ΔE (middle) and C'_{NN} (right) split into $\overline{D^0} \to K^+\pi^-$ decays (top) and $\overline{D^0} \to K^+\pi^-\pi^0$ (bottom). Blue line shows signal PDF, red shows generic $B\overline{B}$ background PDF, green shows continuum background PDF, yellow shows charmless $B\overline{B}$ background PDF, black is sum of all PDFs, points are data.

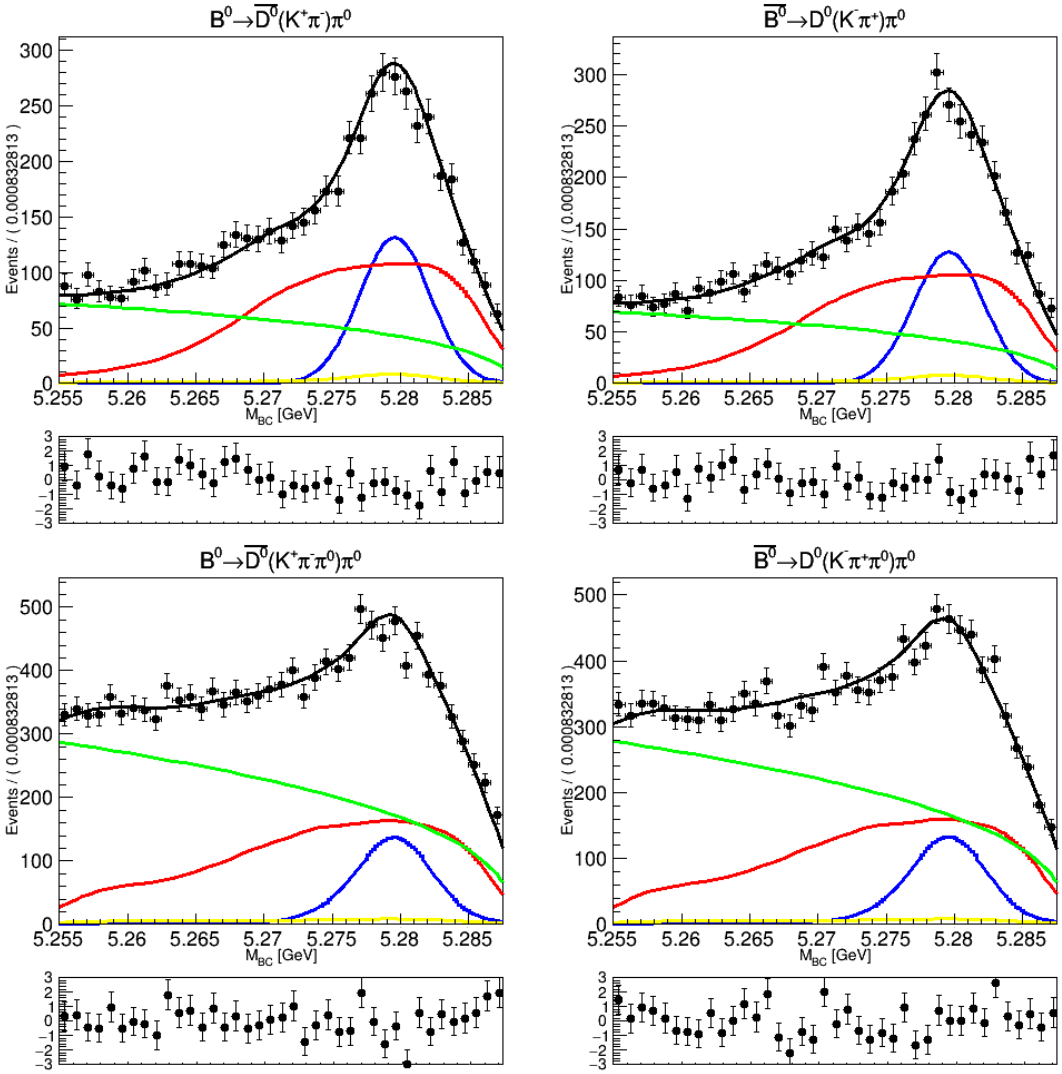


Figure 66: $B^0 \to \overline{D^0}\pi^0$ fit plots for M_{BC}, showing $B^0 \to \overline{D^0}(K^+\pi^-)\pi^0$ (top left), $B^0 \to D^0(K^-\pi^+)\pi^0$ (top right), $B^0 \to \overline{D^0}(K^+\pi^-\pi^0)\pi^0$ (bottom left), $B^0 \to D^0(K^-\pi^+\pi^0)\pi^0$ (bottom right). Blue line shows signal PDF, red shows generic $B\overline{B}$ PDF, green shows continuum background PDF, yellow shows charmless $B\overline{B}$ background PDF, black is sum of all PDFs, points are data.

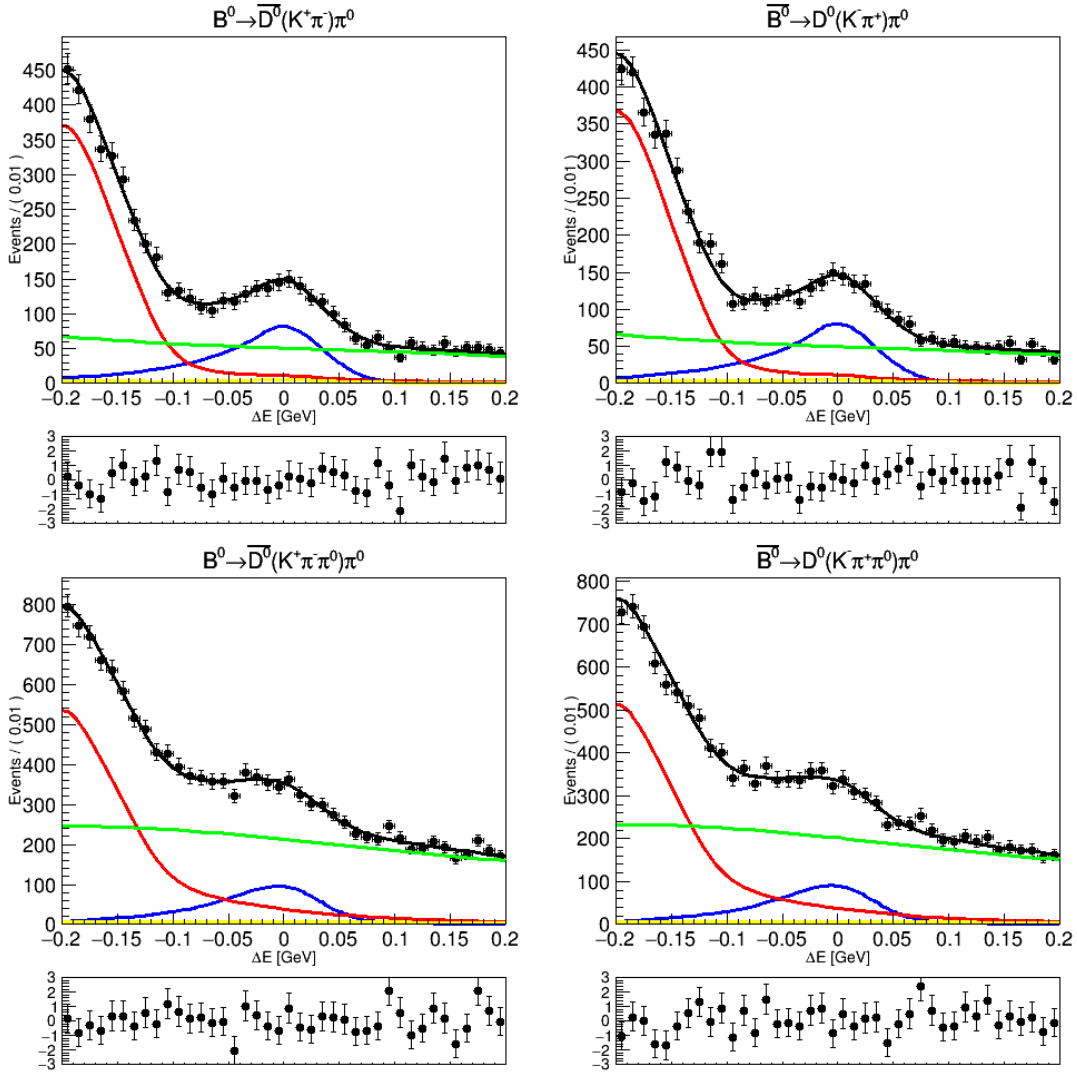


Figure 67: $B^0 \to \overline{D^0}\pi^0$ fit plots for ΔE, showing $B^0 \to \overline{D^0}(K^+\pi^-)\pi^0$ (top left), $B^0 \to D^0(K^-\pi^+)\pi^0$ (top right), $B^0 \to \overline{D^0}(K^+\pi^-\pi^0)\pi^0$ (bottom left), $B^0 \to D^0(K^-\pi^+\pi^0)\pi^0$ (bottom right). Blue line shows signal PDF, red shows generic $B\overline{B}$ background PDF, green shows continuum background PDF, yellow shows charmless $B\overline{B}$ background PDF, black is sum of all PDFs, points are data.

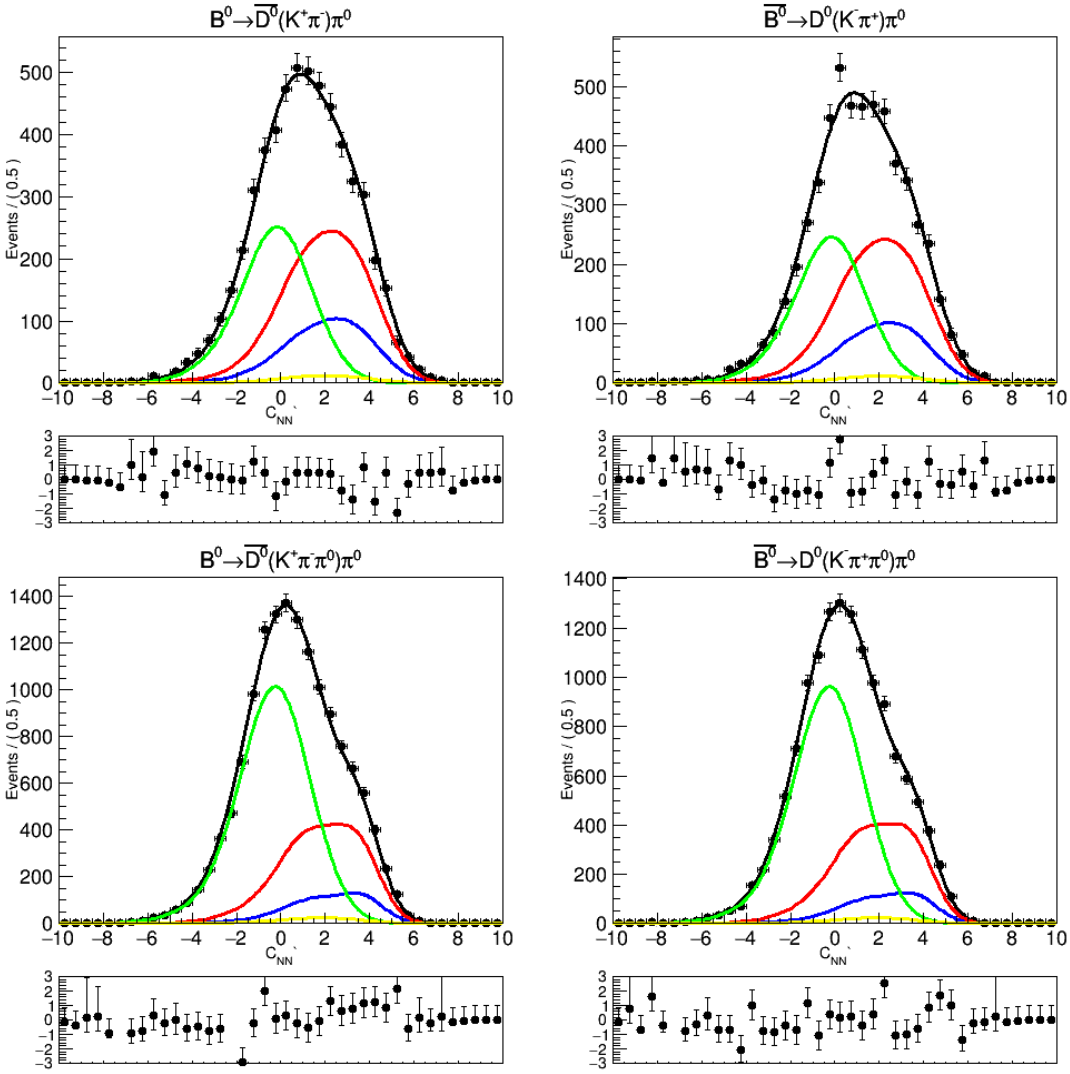


Figure 68: $B^0 \to \overline{D^0}\pi^0$ fit plots for C'_{NN}, showing $B^0 \to \overline{D^0}(K^+\pi^-)\pi^0$ (top left), $B^0 \to D^0(K^-\pi^+)\pi^0$ (top right), $B^0 \to \overline{D^0}(K^+\pi^-\pi^0)\pi^0$ (bottom left), $B^0 \to D^0(K^-\pi^+\pi^0)\pi^0$ (bottom right). Blue line shows signal PDF, red shows generic $B\overline{B}$ background PDF, green shows continuum background PDF, yellow shows charmless $B\overline{B}$ background PDF, black is sum of all PDFs, points are data.

7 Conclusion

This thesis has presented a study of the branching fraction and direct CP asymmetry of the hadronic B decay $B^0 \to \overline{D^0}\pi^0$ and $B^+ \to \overline{D^0}\pi^+$, using the full Belle dataset of 711 fb^{-1} at the $\Upsilon(4S)$ resonance. A three-dimensional fit in the kinematic variables M_{BC} and ΔE, and the continuum suppression variable C'_{NN} is utilised to extract the signal yield and A_{CP}.

The branching fraction and direct CP asymmetry for $B^0 \to \overline{D^0}\pi^0$ are measured as

$$\mathfrak{B}\left(B^0 \to \overline{D^0}\pi^0\right) = \left(2.69 \pm 0.06(stat.) \pm 0.09(syst.)\right) \times 10^{-4}, \quad (7.1)$$

$$A_{CP}\left(B^0 \to \overline{D^0}\pi^0\right) = \left(0.10 \pm 2.05(stat.) \pm 1.29(syst.)\right) \times 10^{-2}, \quad (7.2)$$

respectively. This is the world's first measurement of direct CP asymmetry in this mode and is consistent with the theoretical expectation of no direct CP violation. It also represents the most precise measurement of $\mathfrak{B}\left(B^0 \to \overline{D^0}\pi^0\right)$ to date. It is consistent with the previous Belle result [15] and supersedes it. It is also in agreement with the measurement by the BaBar experiment [16] but achieves better statistical and systematic precision.

The $B^+ \to \overline{D^0}\pi^+$ decay has been studied primarily as a means of verifying the analysis technique with a well measured decay, but is also reported as a new measurement. Its branching fraction and direct CP asymmetry are measured to be

$$\mathfrak{B}\left(B^+ \to \overline{D^0}\pi^+\right) = \left(4.53 \pm 0.02(stat.) \pm 0.14(syst.)\right) \times 10^{-3}, \quad (7.3)$$

$$A_{CP}\left(B^+ \to \overline{D^0}\pi^+\right) = \left(0.19 \pm 0.36(stat.) \pm 0.60(stat.)\right) \times 10^{-2}, \quad (7.4)$$

which is the most accurate measurement of these two values in the world. They agree with and supersede the previous Belle result for branching fraction [69] and direct CP asymmetry [71]. They are also consistent with and more precise than the BaBar measurement of branching fraction [70] and LHCb's study of direct CP asymmetry [72].

These results set a benchmark for future high precision B measurements and should prove valuable for the calibration of rare decay analyses at the 2nd generation B factories such as Belle II. These measurements also provide constraints on new physics theories introducing non-standard model flavour changing neutral currents that may induce A_{CP} through penguin interference.

Bibliography

1. Canetti, L., M. Drewes, and M. Shaposhnikov, *Matter and Antimatter in the Universe.* New J.Phys., 2012. **14**: p. 095012.
DOI: 10.1088/1367-2630/14/9/095012.
URL: https://arxiv.org/abs/1204.4186.

2. Sakharov, A.D., *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe.* Soviet Physics Uspekhi, 1991. **34**(5): p. 392-393.
DOI: 10.1070/pu1991v034n05abeh002497.
URL: http://dx.doi.org/10.1070/PU1991v034n05ABEH002497.

3. Tanabashi, M., et al., *Review of Particle Physics.* Phys.Rev., 2018. **D98**(3): p. 030001.
DOI: 10.1103/PhysRevD.98.030001.
URL: http://pdg.lbl.gov.

4. Lee, T.D. and C.-N. Yang, *Question of Parity Conservation in Weak Interactions.* Phys.Rev., 1956. **104**: p. 254-258.
DOI: 10.1103/PhysRev.104.254.

5. Wu, C.S., et al., *Experimental Test of Parity Conservation in Beta Decay.* Phys.Rev., 1957. **105**: p. 1413-1414.
DOI: 10.1103/PhysRev.105.1413.

6. Christenson, J.H., et al., *Evidence for the* 2π *Decay of the* K_2^0 *Meson.* Phys.Rev.Lett., 1964. **13**: p. 138-140.
DOI: 10.1103/PhysRevLett.13.138.

7. Kobayashi, M. and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction.* Prog.Theor.Phys., 1973. **49**: p. 652-657.
DOI: 10.1143/PTP.49.652.

8. Herb, S.W., et al., *Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions.* Phys.Rev.Lett., 1977. **39**: p. 252-255.
DOI: 10.1103/PhysRevLett.39.252.

9. Abe, F., et al., *Observation of top quark production in* $\bar{p}p$ *collisions.* Phys.Rev.Lett., 1995. **74**: p. 2626-2631.
DOI: 10.1103/PhysRevLett.74.2626.
URL: https://arxiv.org/abs/hep-ex/9503002.

10. Abachi, S., et al., *Search for high mass top quark production in* $p\bar{p}$ *collisions at* $\sqrt{s} = 1.8$TeV*.* Phys.Rev.Lett., 1995. **74**: p. 2422-2426.
DOI: 10.1103/PhysRevLett.74.2422.
URL: https://arxiv.org/abs/hep-ex/9411001.

11. Chau, L.-L. and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix.* Phys.Rev.Lett., 1984. **53**: p. 1802.
DOI: 10.1103/PhysRevLett.53.1802.

12. Wolfenstein, L., *Parametrization of the Kobayashi-Maskawa Matrix.* Phys.Rev.Lett., 1983. **51**: p. 1945. DOI: 10.1103/PhysRevLett.51.1945.

13. Charles, J., et al., *CP violation and the CKM matrix: Assessing the impact of the asymmetric* B *factories.* Eur.Phys.J., 2005. **C41**(1): p. 1-131. DOI: 10.1140/epjc/s2005-02169-1. URL: updated results and plots available at: http://ckmfitter.in2p3.fr.

14. Bigi, I.I. and A.I. Sanda, *CP Violation*. 2 ed. 2009, Cambridge: Cambridge University Press.

15. Blyth, S., et al., *Improved measurements of color-suppressed decays anti-*$B^0 \to D^0\pi^0, D^0\eta, D^0\omega, D^{*0}\pi^0, D^{*0}\eta$ and $D^{*0}\omega$. Phys.Rev., 2006. **D74**: p. 092002. DOI: 10.1103/PhysRevD.74.092002. URL: https://arxiv.org/abs/hep-ex/0607029.

16. Lees, J.P., et al., *Branching Fraction Measurements of the Color-Suppressed Decays* $\overline{B^0} \to D^{(*)0}\pi^0$, $D^{(*)0}\eta$, $D^{(*)0}\omega$, *and* $D^{(*)0}\eta'$ *and Measurement of the Polarization in the Decay* $\overline{B^0} \to D^{*0}\omega$. Phys.Rev., 2011. **D84**(3): p. 112007. DOI: 10.1103/PhysRevD.84.112007. URL: https://arxiv.org/abs/1107.5751.

17. Bauer, M., B. Stech, and M. Wirbel, *Exclusive non-leptonic decays of* D^-, D_s^- *and B-mesons.* Zeitschrift für Physik C Particles and Fields, 1987. **34**(1): p. 103-115. DOI: 10.1007/bf01561122. URL: https://doi.org/10.1007/BF01561122.

18. Deandrea, A., et al., *Two-body nonleptonic decays of* B *and* B_s *mesons.* Phys.Lett., 1993. **B318**: p. 549-558. DOI: 10.1016/0370-2693(93)91554-Z. URL: https://arxiv.org/abs/hep-ph/9308210.

19. Neubert, M. and A.A. Petrov, *Comments on color suppressed hadronic B decays.* Phys.Lett., 2001. **B519**: p. 50-56. DOI: 10.1016/S0370-2693(01)01093-0. URL: https://arxiv.org/abs/hep-ph/0108103.

20. Bauer, C.W., D. Pirjol, and I.W. Stewart, *Soft collinear factorization in effective field theory.* Phys.Rev., 2002. **D65**: p. 054022. DOI: 10.1103/PhysRevD.65.054022. URL: https://arxiv.org/abs/hep-ph/0109045.

21. Mantry, S., D. Pirjol, and I.W. Stewart, *Strong phases and factorization for color suppressed decays.* Phys.Rev., 2003. **D68**: p. 114009. DOI: 10.1103/PhysRevD.68.114009. URL: https://arxiv.org/abs/hep-ph/0306254.

22. Keum, Y.-Y., et al., *Nonfactorizable contributions to* $B \to D^{**(*)}M$ *decays.* Phys.Rev., 2004. **D69**: p. 094018. DOI: 10.1103/PhysRevD.69.094018. URL: https://arxiv.org/abs/hep-ph/0305335.

23. Lu, C.-D., *Study of color suppressed modes* $B^0 \to \overline{(D^{*0})}\eta'$. Phys.Rev., 2003. **D68**: p. 097502.
DOI: 10.1103/PhysRevD.68.097502.
URL: https://arxiv.org/abs/hep-ph/0307040.

24. Leganger, L.E. and J.O. Eeg, *Non-factorizable contribtion to* $\overline{B_d^0} \to \pi^0 D^0$. Phys.Rev., 2010. **D82**: p. 074007.
DOI: 10.1103/PhysRevD.82.074007.
URL: https://arxiv.org/abs/1003.3348.

25. Abdesselam, A., et al., *First Observation of CP Violation in* $\overline{B}^0 \to D_{CP}^{(*)} h^0$ *Decays by a Combined Time-Dependent Analysis of BABAR and Belle Data.* Phys.Rev.Lett., 2015. **115**(12): p. 121604.
DOI: 10.1103/PhysRevLett.115.121604.
URL: https://arxiv.org/abs/1505.04147v2.

26. Abe, T., et al., *Achievements of KEKB.* PTEP, 2013. **2013**: p. 03A001.
DOI: 10.1093/ptep/pts102.

27. Kurokawa, S. and E. Kikutani, *Overview of the KEKB accelerators.* Nucl.Instrum.Meth., 2003. **A499**: p. 1-7.
DOI: 10.1016/S0168-9002(02)01771-0.

28. Aubert, B., et al., *The BaBar detector.* Nucl.Instrum.Meth., 2002. **A479**: p. 1-116.
DOI: 10.1016/S0168-9002(01)02012-5.
URL: https://arxiv.org/abs/hep-ex/0105044.

29. Bevan, A.J., et al., *The Physics of the B Factories.* Eur.Phys.J., 2014. **C74**: p. 3026.
DOI: 10.1140/epjc/s10052-014-3026-9.

30. Abe, T., et al., *Compensation of the crossing angle with crab cavities at KEKB.* Conf.Proc., 2007. **C070625**: p. 27.
DOI: 10.1109/PAC.2007.4440328.

31. Eilfrei, *Use of crab caveties for the rotation of charged particle bunches in particle accelerators.* English Wikipedia 2015. Visited 2020-02-26.
URL: https://en.wikipedia.org/wiki/File:Crabcavity_n2.svg.

32. Abashian, A., et al., *The Belle Detector.* Nucl.Instrum.Meth., 2002. **A479**: p. 117-232.
DOI: 10.1016/S0168-9002(01)02013-7.

33. Brodzicka, J., et al., *Physics Achievements from the Belle Experiment.* PTEP, 2012. **2012**: p. 04D001.
DOI: 10.1093/ptep/pts072.

34. Ishino, H., *The Belle Silicon Vertex Detector.* Nucl.Instrum.Meth., 2005. **A549**: p. 16-19.
DOI: 10.1016/j.nima.2005.04.018.

35. Aihara, H., et al., *Belle SVD2 vertex detector.* Nucl.Instrum.Meth., 2006. **A568**: p. 269-273.
DOI: 10.1016/j.nima.2006.05.281.

36. Uno, S., *The BELLE central drift chamber.* Nucl.Instrum.Meth., 1996. **A379**: p. 421-423.
DOI: 10.1016/0168-9002(96)00555-4.

37. Bethe, H. and W. Heitler, *On the Stopping of Fast Particles and on the Creation of Positive Electrons.* Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, Volume 146, Issue 856, pp. 83-112, 1934. **146**: p. 83-112.
DOI: 10.1098/rspa.1934.0140.
URL: https://ui.adsabs.harvard.edu/abs/1934RSPSA.146...83B.

38. Iijima, T., et al., *Aerogel Cherenkov counter for the BELLE detector.* Nucl.Instrum.Meth., 2000. **A453**: p. 321-325.
DOI: 10.1016/S0168-9002(00)00652-5.

39. Sumiyoshi, T., et al., *Silica aerogel Cherenkov counter for the KEK B factory experiment.* Nucl.Instrum.Meth., 1999. **A433**: p. 385-391.
DOI: 10.1016/S0168-9002(99)00460-X.

40. Cherenkov, P.A., *Visible Emission of Clean Liquids by Action of γ Radiation.* Doklady Akademii Nauk SSSR, 1934. **2**: p. 451.

41. Kichimi, H., et al., *The BELLE TOF system.* Nucl.Instrum.Meth., 2000. **A453**: p. 315-320.
DOI: 10.1016/S0168-9002(00)00651-3.

42. Sagawa, H., *The BELLE CsI calorimeter.* Nucl.Instrum.Meth., 2000. **A453**: p. 259-261.
DOI: 10.1016/S0168-9002(00)00643-4.

43. Elliot, C., R.J. P, and E. Teresa, *Calorimetry In High Energy Physics - Proceedings Of The 7th International Conference*. 1998: World Scientific Publishing Company: p. 18-20.

44. Tsai, Y.-S., *Pair Production and Bremsstrahlung of Charged Leptons.* Rev.Mod.Phys., 1974. **46**: p. 815.
DOI: 10.1103/RevModPhys.46.815.

45. Abashian, A., et al., *The* K_L/μ *detector subsystem for the BELLE experiment at the KEK B factory.* Nucl.Instrum.Meth., 2000. **A449**: p. 112-124.
DOI: 10.1016/S0168-9002(99)01383-2.

46. Ushiroda, Y., et al., *Development of the central trigger system for the BELLE detector at the KEK B factory.* Nucl.Instrum.Meth., 1999. **A438**: p. 460-471.
DOI: 10.1016/S0168-9002(99)00823-2.

47. Itoh, R., *BASF - BELLE AnalysiS Framework*, in *9th International Conference on Computing in High-Energy Physics (CHEP 1997)*. 1997: Berlin, Germany.

48. Fox, G.C. and S. Wolfram, *Event Shapes in* e^+e^- *Annihilation.* Nucl.Phys., 1979. **B149**: p. 413.
DOI: 10.1016/0550-3213(79)90120-2.

49. Nakano, E., *Belle PID.* Nucl.Instrum.Meth., 2002. **A494**: p. 402-408. DOI: 10.1016/S0168-9002(02)01510-3.

50. Nishida, S., *Study of kaon and pion identification using inclusive* D^* *sample*, in *Belle Note **779***. 2005. (Belle internal document).

51. Feindt, M. and U. Kerzel, *The NeuroBayes neural network package.* Nucl.Instrum.Meth., 2006. **A559**: p. 190-194. DOI: 10.1016/j.nima.2005.11.166.

52. Das, A., et al., *Measurements of Branching Fractions for* $B^0 \to D_s^+ \pi^-$ *and* $\overline{B^0} \to D_s^+ K^-$. Phys.Rev., 2010. **D82**: p. 051103. DOI: 10.1103/PhysRevD.82.051103. URL: https://arxiv.org/abs/1007.4619.

53. Lange, D.J., *The EvtGen particle decay simulation package.* Nucl.Instrum.Meth., 2001. **A462**: p. 152-155. DOI: 10.1016/S0168-9002(01)00089-4.

54. Brun, R., et al., *GEANT 3.21.* CERN Report No. DD/EE/84-1, 1987.

55. Julius, T.M., *Measurement of the branching fraction of* $B^0 \to \pi^0\pi^0$ *decays using the final Belle dataset*. School of Physics, University of Melbourne, 2015. PhD. URL: http://hdl.handle.net/11343/58233.

56. Y.Unno. Image taken from Belle Analysis School presentation. (Private communication).

57. Fisher, R.A., *The Statistical Utilization of Multiple Measurements.* Annals of Eugenics, 1938. **8**(4): p. 376-386. DOI: 10.1111/j.1469-1809.1938.tb02189.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-1809.1938.tb02189.x.

58. Lee, S.H., et al., *Evidence for* $B^0 \to \pi^0\pi^0$. Phys.Rev.Lett., 2003. **91**: p. 261801. DOI: 10.1103/PhysRevLett.91.261801. URL: https://arxiv.org/abs/hep-ex/0308040.

59. Kakuno, H., et al., *Neutral B flavor tagging for the measurement of mixing induced CP violation at Belle.* Nucl.Instrum.Meth., 2004. **A533**: p. 516-531. DOI: 10.1016/j.nima.2004.06.159. URL: https://arxiv.org/abs/hep-ex/0403022.

60. Verkerke, W. and D. Kirkby *The RooFit toolkit for data modeling*. arXiv e-prints, 2003.

61. Antcheva, I., et al., *ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization.* Comput.Phys.Commun., 2009. **180**: p. 2499-2512. DOI: 10.1016/j.cpc.2009.08.005. URL: https://arxiv.org/abs/1508.07749.

62. Skwarnicki, T., *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*. 1986: DESY, F31-86-02.

63. Albrecht, H., et al., *Search for Hadronic* $b \to u$ *Decays.* Phys.Lett., 1990. **B241**: p. 278-282.
DOI: 10.1016/0370-2693(90)91293-K.

64. Cranmer, K.S., *Kernel estimation in high-energy physics.* Comput.Phys.Commun., 2001. **136**: p. 198-207.
DOI: 10.1016/S0010-4655(00)00243-5.
URL: https://arxiv.org/abs/hep-ex/0011057.

65. Drleft, *Comparison of a histogram and a kernel density estimate.* English Wikipedia 2010.
URL: https://en.wikipedia.org/wiki/File:Comparison_of_1D_histogram_and_KDE.png.

66. Winkler, M. and F. James, *MINUIT User's Guide.* 2004.

67. Rice, J.A., *Mathematical statistics and data analysis*. 2007, Belmont, CA: Thomson/Brooks/Cole.

68. Cox, M.G., et al., *The generalized weighted mean of correlated quantities.* Metrologia, 2006. **43**(4): p. S268-S275.
DOI: 10.1088/0026-1394/43/4/s14.
URL: http://dx.doi.org/10.1088/0026-1394/43/4/S14.

69. Kato, Y., et al., *Measurements of the absolute branching fractions of* $B^+ \to X_{c\bar{c}}K^+$ *and* $B^+ \to \overline{D^{(*)0}}\pi^+$ *at Belle.* Phys.Rev., 2018. **D97**(1): p. 012005.
DOI: 10.1103/PhysRevD.97.012005.

70. Aubert, B., et al., *Branching fraction measurement of* $B^0 \to D^{(*)} + \pi^-$, $B^- \to D^{(*)0}\pi^-$ *and isospin analysis of* $\overline{B} \to D^{(*)}$ *decays.* Phys.Rev., 2007. **D75**: p. 031101.
DOI: 10.1103/PhysRevD.75.031101.

71. Abe, K., et al., *Study of* $B^\pm \to D_{CP}K^\pm$ and $D^*_{CP}K^\pm$ *decays.* Phys.Rev., 2006. **D73**: p. 051106.
DOI: 10.1103/PhysRevD.73.051106.

72. Aaij, R., et al., *Observation of the suppressed ADS modes* $B^\pm \to \left[\pi^\pm K^\mp \pi^+ \pi^-\right]_D K^\pm$ *and* $B^\pm \to \left[\pi^\pm K^\mp \pi^+ \pi^-\right]_D \pi^\pm$. Phys.Lett., 2013. **B723**: p. 44-53.
DOI: 10.1016/j.physletb.2013.05.009.

73. Bhuyan, B., *High PT Tracking Efficiency Using Partially Reconstructed* D^*, in *Belle Note **1165***. 2010. (Belle internal document).

74. Hsu, C.-L., et al., *Measurement of branching fraction and direct CP asymmetry in charmless* $B^+ \to K^+K^-\pi^+$ *decays at Belle.* Phys.Rev., 2017. **D96**(3): p. 031101.
DOI: 10.1103/PhysRevD.96.031101.
URL: https://arxiv.org/abs/1705.02640.

75. Ryu, S., *Study of* π^0 *efficiency using* $\tau^- \to \pi^- \pi^0 \nu$ *events*, in *Belle Note **1224***. 2012. (Belle internal document).

76. Julius, T., et al., *Measurement of the branching fraction and CP asymmetry in* $B^0 \to \pi^0\pi^0$ *decays, and an improved constraint on* ϕ_2. Phys.Rev., 2017. **D96**(3): p. 032007.
DOI: 10.1103/PhysRevD.96.032007.
URL: https://arxiv.org/abs/1705.02083.